Investigating quality of data and the need for the restructuring of accident report form in South Africa

by

Oluwatimilehin Oluwaseun Okeowo

Thesis presented in fulfilment of the requirements for the degree of Master of Engineering (Research) (Engineering Management) in the Faculty of Engineering at Stellenbosch University

Pectora roborant cultus recti

UNIVERSITEIT
iYUNIVESITHI
STELLENBOSCH
UNIVERSITY

100
1918 · 2018

Supervisor: Dr Theuns Dirkse van Schalkwyk

March 2018

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own original work, that I am the sole author thereof [save to the extent explicitly otherwise stated], that reproduction and publication thereof by Stellenbosch University will not infringe any third-party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

O.O. Okeowo:

…………………………

March, 2018

…………………………

Date

Abstract

A reliable data quality system is a priority to all organisations or departments around the globe. Quality data is of great value to the agencies managing the road traffic system in South Africa in such areas as improvement of road safety complications, reconstruction of roads, advocating for adequate resources, strategising new public safety campaigns, and strengthening the enforcement of road policies to curb prevalent road accidents. The quality of the road accident data compiled at local level, according to complaints from the data users [data consumers], is alleged to be inaccurate, inadequate with regard to the relevant data demanded by the data users. Due to this effect, a need for a continuous and efficient assessment plan is considered at the local level [municipality level], where the preliminary data collection process is initiated, with the objective of uncovering and discarding the existence of errors or anomalies which frustrate the chance of raising the quality of the data processed at the provincial level.

In an attempt to have a detail understanding on these anomalies, however, objectives of this study are to investigate and evaluate the anomalies that exist during the reporting and recording of road accident data at the Stellenbosch Traffic Department under the jurisdiction of the Stellenbosch Municipality-Western Cape 024, and the necessity for the improvement of the Accident Report form to accommodate more information and better understanding of its application. These objectives are achieved through research questions designed to guide the methodology applied to execute the investigation. These questions cut across how simple and comprehensible the Accident Report form is to the users, and the practicality of the information provided in the form in understanding causes of road accidents and in managing resources in preventing road accidents in the Stellenbosch area. In this case, the methodology applied is based on the four-stage approach model, which includes the understanding and evaluation of the process engaged in the traffic department, in order to have a big picture about the vulnerable areas and the responsible factors affecting the quality of data collected through the use of Accident Report form. The methodology also incorporates the collection of the right data by assessing the information completed in the form through definite dimensions and metrics, and also a questionnaire was developed to substantiate both quantitative and qualitative data collection processes which involved the participation of the South African Police Service, as one of the authorised data collectors in the Stellenbosch.

With the aim of making this investigation a successful one, the analysis of the data collected was performed to determine the persistent accumulation of errors per data field in the Accident Report form, and the practical evaluation of data usability as applicable to the road accident occurrence within the Stellenbosch area, coupled with the analysis of the applicability of the Accident Report form through the understanding of the data collected with questionnaire.

These analyses are practically achieved through the introduction of the statistical base methods to measure the average scores and the frequency distribution of errors per field in each related factor defined, and also to display the interpretation of results achieved to illustrate the consequence of road accidents on the Stellenbosch environs.

Currently, the findings acquired demonstrate that huge proportion of errors were detected in the data completed in the Accident Report forms. The nature of errors determined indicates that human inaccuracies have paramount influence on the data completed in the form, which is attributable to the excessive level of negligence existing along the data processing line. This effect is ascribed to haphazard execution of the training and validation processes, and untimeliness in the processing of information, which probably complicate the level of quality of the data captured. A framework is developed as part of the objectives of this study to support an improved application of the quality dimensions along the road accident data processing line at the Stellenbosch Traffic Department. This framework strengthens the ability of both the reporting officers and the supervisory officers to adopt a zero-tolerance approach towards the existence of errors along the data processing line between the two authorised local departments.

The practicality of the data analysed as regard the consequence of road accidents on the Stellenbosch environs demonstrates that more road accidents occurred during the dawn periods [7:00 am to 7:59 am] and dusk periods [4:00 pm to 4:59 pm], most frequently in the winter periods, within a speed limit of sixty kilometres per hour, mostly on Fridays and Thursdays within the built-up areas on both the dual and single carriageways. The findings further indicate that road accidents occurrence involved more male drivers/cyclists, under the age of forty years, either more as South Africans or less as foreign nationals. However, the findings gathered from the survey analysis shows that some form users are facing challenges during the data collection activities, where crucial concerns were raised regarding the construction of a detailed accident sketch and description. This calls for a better training for the form users on how to construct a detailed sketch and description in accordance with the necessary features required. Moreover, some valuable and relevant improvements were determined in the areas like information arrangement, information interpretation, and addition of new information where necessary.

Opsomming

'n Betroubare data kwaliteit stelsel is 'n prioriteit vir enige organisasie of departement wêreldwyd. Gehalte data is belangrik en van groot waarde vir die bestuur van die padverkeer stelsel in Suid-Afrika, byvoorbeeld in die verbetering van padveiligheid, rekonstruksie van paaie, aanvra van voldoende hulpbronne, strategië vir nuwe openbare veiligheid veldtogte, en die afdwing en uitbouing van verkeersreëls om padongelukke te bekamp. Die klagtes van die data gebruikers met betrekking tot die gehalte van data wat gegenereer word deur die Padverkeersbestuurskorporasie is verwant aan die metode waarmee die verwerking van data hanteer word by elke betrokke vlak. Die behoefte aan 'n deurlopende assesseringsplan word bestudeer op die plaaslike munisipaliteit vlak, waar voorlopige data-insamelingsproses geïnisieer word, met die doel om foute of teenstrydighede te ontbloot, weg te doen daarmee, en 'n verhoging van die kwaliteit van die data wat verwerk word op provinsiale vlak mee te bring.

Een van die doelwitte van hierdie studie is om die teenstrydighede in verslagdoening en rekordhouding van verkeersinsidente op die plaaslike vlak in Suid-Afrika te bestudeer en te evalueer. Hierdie ondersoek noodsaak in-diepte insig in die prosedures wat geïmplementeer is binne die Stellenbosch Verkeersdepartement, as verantwoordelike departement insake Stellenbosch Munisipaliteit [SM-WC 024] prosedures, en die betrokkenheid van die Suid Afrikaanse Polisiediens in Stellenbosch, as een van die gemagtigde data versamelaars op plaaslike vlak. Dit noodsaak evaluasie van die bydraes van hierdie twee groot owerhede aangaande teenstrydighede in die data verwerkingsproses op plaaslike vlak, die analise van die primere data wat ingesamel is, en die bestudering van handgeskrewe ongeprosesseerde data op die ongeluksverslag vorms.

Om hierdie studie suksesvol af te handel, is belangrike data verwerking prosesse beskou; die noodsaaklikheid vir die verbetering van die ongeluksverslag vorm om meer inligting makliker vas te lê. Daar is ook 'n behoefte aan die ontwikkeling van 'n standaard raamwerk om die prestasie van die personeel en die eenhede wat betrokke is by die verwerking van data, te verbeter. Die ontwikkeling van die raamwerk is gebaseer op bevindings uit drie belangrike gebiede, naamlik die ontleding van die primere data, foute wat tydens datavaslegging gebeur, en die ontleding van kwalitatiewe data verkry deur vraelyste. Die vraelyste is voltooi deur twee tipes gebruikers op die plaaslike vlak, naamlik die verkeersbeamptes en polisiebeamptes.

Foute wat ontbloot is in die voltooide Ongeluk Verslag Vorms dui op aansienlike nalatigheid in die verwerking van data en prosesse tussen die twee plaaslike departemente. Die raamwerk wat ontwikkel is kan gebruik word om die toesighoudende beamptes by te staan om 'n zero toleransie benadering tot die ontstaan van foute in dataprosesse op plaaslike vlak te ondersteun. Die implementering van die raamwerk vereis toewyding en bereidwilligheid van

die twee departemente wat betrokke is en gereelde kommunikasie aangaande die voordele van die model aan hul onderskeie personeel in beheer van data-insameling stelsel, ten einde volhoubare verbetering in hulle data prosesse te realiseer.

Die verwerkte data van padongelukke op die Stellenbosch omgewing data toon dat meer padongelukke gedurende die dagbreek tydperke [07:00-07:59] en skemer tydperke [16:00-16:59], die meeste in die winter tydperke, binne 'n spoedgrens van sestig kilometer per uur, meestal op Vrydae en Donderdae in die beboude gebiede aan beide die dubbele en enkel rybane. Die bevindinge dui ook dat ongelukke meer gereeld voorkom met manlike bestuurders/fietsryers, onder die ouderdom van veertig jaar, en meer Suid-Afrikaners as buitelanders. Uitdagings in die data-insameling aktiwiteite sluit in die teken van 'n gedetailleerde skets ongeluk en beskrywing. Beter opleiding vir die vorm gebruikers oor hoe om 'n gedetailleerde skets en beskrywing saam te stel is, word aanbeveel. Daarbenewens is 'n paar waardevolle en relevante verbeterings bepaal in die gebiede soos inligting reëling, inligting interpretasie, en byvoeging van nuwe inligting waar nodig.

Acknowledgement

First of all, I would like to thank my family for their valuable and considerate supports throughout the completion of this research project. Broadly speaking, the success of this research project was made possible through the vital contributions from different individuals, journal articles, books and organisations such as Stellenbosch Traffic Department [STD] and South African Police Service [SAPS]. Finally, my candid appreciation is directed to my intelligent and amiable study leader, for his encouraging words, mutual understanding, collective reasoning and above it all, his technical advice towards the success of this project. Thank you for your valuable time, support, and guidance in making sure that this project is completed.

Dedications

This research project is dedicated to the glory of Almighty God for His grace bestowed on me from the inception of this study and to the conclusion of it. Also, dedicated to the memory of my late sister, *Oluwakemi 'Tunmise Okeowo,* who passed away in the middle of this study. May her gentle soul rest in perfect peace.

Table of contents

List of figures

List of tables

List of abbreviations

AASA	Automobile Association of South Africa
AB	Alcohol Breathalyser
ARF	Accident Report form
ARU	Accident Response Unit
ARN	Accident Register Number
BAC	Blood Alcohol Concentration
BCM	Buffalo City Municipality
CBRTA	Cross-Border Road Transport Agency
CPA	Criminal Procedural Act
CRO	Control Room Officer
DCO	Data Capturing Officer
DCS	Data Capturing Supervisor
DHA	Department of Home Affairs
DNF	Delivery Note Form
DoT	Department of Transport
eNaTIS	Electronic National Information System
GIS	Geographic Information System
IPAS	Integrated Provincial Accident System
MMT	Municipal/Metro, Traffic Department
MoT	Ministry of Transport
MRC/UNISA	Medical Research Council/University of South Africa
NATMAP	National Transport Master Plan
NDoH	National Department of Health
NDoT	National Department of Transport
NIMSS	National Injury Mortality Surveillance System
NRTA	National Road Traffic Act
PAIA	Promotion of Access to Information Act

POPI	Promotion of Personal Information
PTS	Public Transport Strategy
RAF	Road Accident Fund
RTI	Road Traffic Information
RTIA	Road Traffic Infringements Agency
RTA	Road Traffic Accident
RTMC	Road Traffic Management Corporation
RTMCA	Road Traffic Management Corporation Act
RTMS	Road Traffic Management System
S/CN	Serial/Capturing Number
SANRAL	South African National Road Safety Agency Limited
SAPS	South African Police Service
SM	Stellenbosch Municipality
SM-WC 024	Stellenbosch Municipality-Western Cape 024
SME	Speed Measurement Equipment
SOPs	Standard Operating Procedures
Stats SA	Statistics South Africa
STD	Stellenbosch Traffic Department
TCSP	Technical Committee for Standards and Procedures
TLVME	Traffic Light Violation Monitoring Equipment
TRAFMAN	Traffic Data Management Software
TSEs	Transportation Safety Engineers
UN/ECE	United Nations/Economic Commission for Europe
WHO	World Health Organisation

1. Introduction to the investigation

In this chapter, section 1.1 covers the introduction to the problems affecting the reporting and recording of road traffic accident [RTA] in South Africa, based on the *"complaints"* from the public about the quality level of road accident data collated at the local level, and transferred to the provincial level for further improvement. The effect of these problems on the data handled by the Road Traffic Management Corporation [RTMC], as the authorised agency managing the road traffic information [RTI], is concisely considered as part of the discussion. Moreover, the significance of the data analysis results to both the data producers and data users will be an additional part of the subject areas to be discussed in the section. However, the list of the likelihood problems preventing the capacity of the traffic management from producing a reliable road accident data is provided as part of the overview of the section 1.1.

Subsequently, the objectives of the study will be discussed in section 1.2, wherein the purpose of the study was observed along with the introduction to the necessary steps to engage in accessing the right data for this study. As part of the discussions in this section, the source of data used in the study was mentioned, to establish the importance of the study. The discussion is extended to key focus areas in the study, where the motive behind the technique to be adopted is briefly discussed. The purpose of considering the development of a framework to support the outcome of the analysis and findings is also discussed.

The last section in this chapter discussed the purpose of implementing the four-stage approach in the study. The considerable elements or factors involved in the process are introduced concisely in the section 1.2, along with the implementation process of the four-stage approach. As part of this section, the summary context of the following chapter is provided as concluding part of section 1.2, to coherently establish a clear understanding on the flow of the investigation to be carried out.

1.1 Problem statement

Presently, the level of inaccurate and insufficient information completed in the Accident Report form [ARF] affects the possibility of generating complete road accident data at the Stellenbosch Traffic Department [STD]. In addition, it is alleged that the road accident data compiled at the local level, processed at the provincial level and disseminated at the national level for the public use is marred with anomalies along the processing line (Vogel & Bester 2004; Sinclair 2011; Joseph 2013). However, the quality of road accident data is of utmost priority to both the data producer and data user in many ways. In this context, the data producer is identified as the RTMC, the custodian of the road accident data in South Africa (Adams 2001). On the other hand, the users of road accident data are identified as academics, insurers, police, lawyers, traffic engineers, safety engineers, safety educators, government etc. Additional discussions are provided on the significance of the road accident data to the data users in the subsection 2.2.1 in Chapter 2 below.

The performance of the data producer depends on the reliability of the data analysis. Hence, this enhances the capacity of the traffic management to know how to manage and distribute their resource capacities, and be capable of investigating complaints with the aim of improving their services to the public. On the contrary, the data users strongly depend on the reliability of the results obtained from the analysis carried out by the data producer. Thereafter, a post-analysis of the results is carried out on the statistical estimates and trends observed in the results authorised for public use.

It is important that the road accident data collated by the local authorities as analysed by the RTMC, offers better understanding on the areas that might remain problematic to the management. However, the procedures developed in coordinating the reliability of road accident data are considered as the responsibility of the data producer. The traffic management has been briefed on the importance of improving the quality level of the road accident data. This is as a result of the complaints from the data users on the anomalies affecting the quality of the data disseminated to the public annually. Primarily, the accumulation of anomalies along the procedures established for the processing of RTI, has been a concern to the data custodian. These problems prevented the data users from applying the results of the data analysis correctly, and thereby reduced their abilities to make vital decisions in their various areas of specialisation.

The complaints levelled against the quality level of road accident data by the data users cut across various units involved in the handling of the road accident data (O' Day 1993; Njord et al. 2005; Sinclair 2011; Joseph 2013). The actuality of the complaints alleged by the data users is connected to the process practically engaged in the department, and also to the application of the Accident Report form [ARF], which serves as the primary tool for data collection. According to the public complaints, the identified problems frustrating the ability of the traffic management include:

- Misinterpretation of information provided in the ARF.
- Incomplete data in the ARF.
- Deficiency of useful accident information.
- Lack of effective communication interface among the units involved in the evaluation of data.
- Negligence of responsibilities allocated to the units involved in the data processing.
- Incorrect application of codes before/during the data recording operations.

1.2 Objectives of the study and Research questions

The objectives of this study are to investigate and evaluate the anomalies that exist during the reporting and recording of road accident data at the Stellenbosch Traffic Department [STD] in Stellenbosch Municipality-Western Cape 024 [SM-WC 024], and the necessity for the improvement of the ARF, or the process to accommodate more information and better understanding of its application. The study will be carried out at the Stellenbosch Traffic Department [STD], as a case study on reporting and recording of road accident data at the local level in South Africa. This will assist in acquiring an in-depth understanding on the system implemented within the department. However, the study directly focuses on the procedures followed by the local traffic department to collate the unprocessed road accident data [raw data], which serves as the primary source of road accident data processed at the provincial level. In due course, this will offer a technical understanding on the level at which the ARF users or accident reporting officers observe the right application of the form while collecting data at the accident site.

As a result of the technical problems listed in the last paragraph of section 1.1, hence, a need for the development of a standard framework is required. The framework will be developed as an outcome of the investigation and evaluation procedures executed, based on the available traffic accident records accessed at the STD. The aim of the framework will be to motivate the efforts of all the units involved, with the objective of tackling anomalies existing along the data processing line at the local level. Thus, the framework will complement the existing procedures practically applied in the department.

The development of the framework will be established on the detailed overview of the role played by each unit involved. This facilitates the chance of identifying the areas or units that may contribute most to the existence of anomalies, thereby paving the way to find the problems preventing the department from achieving consistent road accident data processing system. This framework will cover areas that require the utmost improvement along the data processing system, as dictate by the results obtained from the analysis performed on the data assessed at the Accident Response Unit [ARU] in the STD. In this study, the framework developed will serve as a recommendation to the management on procedural structures to be used if proven relevant. Afterward, a decision will be made by the management whether to implement the system to support the existing structures.

In addition, if implemented, this framework will improve the quality of data disseminate to the public based on the primary improvements initiated at the local level. Ultimately, it will also solidify the relationship between the data producer and data users, by creating opportunity for valuable feedback that will be relevant towards the improvement of the quality level of the data collated. This research work will concentrate mainly on the procedures engaged in the STD for reporting and recording of the road accident data. With the purpose of achieving the above discussed objectives, however, some research questions were formulated to guide the methodology applied to execute the investigation. These questions are enumerated below as:

- Is the use of the ARF simple and comprehensible to the users?
- Is the information in the ARF appropriately completed to be of practical use in understanding causes of road accidents in the Stellenbosch area?
- Is the information in the ARF appropriately completed to be of practical use in managing resources in preventing road accidents in the Stellenbosch area?

Moreover, by evaluating these questions stepwise, valuable results will be attained and validated to generate satisfactory conclusions and recommendation for future research. A brief introduction to the methodology applied in this study is provided below. The methodology integrated the four-stage approach designed to facilitate the process of achieving the objectives of the study as completely discussed in Chapter 3. The four-stage approach are outlined below as:

- understanding and evaluating the *'process'* engaged in the STD by involving the persons and agencies involved;
- identifying the *'vulnerable areas'* and the *'responsible factors'* affecting the quality of data in the STD through a preliminary exploration of the data completed in the ARFs;
- collecting the *'right data'* from the available traffic records in the STD by assessing the information completed in the ARF, and the application of *'questionnaire'* to gain valuable understanding on the perceptions of the form users on the use of the ARF.
- analysis of the *'data collected'* to yield satisfactory results to guide the need for the development and implementation of a *'framework'* to justify the end product of the investigation.

Previous literatures studied will be provided in the subsequent chapter, which are centred on the standard definitions of the RTA and its consequence in South Africa. Other areas to be discussed are road accident data and its importance, procedures employed in the reporting and recording systems of road accident data, together with the role players' significance in the improvement of the quality of the road accident data gathered, and brief discussions on the probable areas that may be considered for sustainable improvement.

2. Overview of the investigation

This chapter comprises several sections covering the overview of the road accident data. Section 2.1 of this chapter discusses the background of RTA, by considering the effect of RTA across the world. This is buttressed by a discussion on the situation and the economic consequence of RTA in South Africa. However, as part of the supporting discussions in this chapter, the literatures carried out on the technique applied in the early era of road accident data processing, the limitations encountered, and the advantages of developing such technique during the era are provided in the Appendix A.

Furthermore, the following section discusses road accident data collection, which comprises the processes implemented for the transfer of road accident data across all the authorised units. The section exemplifies the procedures followed in collecting road accident data, right from the accident site to the next important stage of road accident data processing. Other areas to be discussed as part of this section are problems affecting road accident data collection, followed by the advantages and disadvantages of the road accident data collection [refer to Appendix A-A.6], and the description of the operations performed by the units handling the processing of road accident data.

The standard operating procedures [SOPs] designed for the processing of RTI is discussed as part of this chapter, alongside with the details about the stakeholders who manage the affairs of the road traffic system in South Africa [refer to Appendix A-A.9]. The chapter extends discussion on the procedural system engaged by the authorised personnel assigned to monitor the appropriate transaction of RTI [refer to Appendix A-A.10]. In addition, an overview description of the operations performed at the STD is considered. This provides insight into the responsibilities of the division managing the RTI in the local traffic department, and simultaneously strengthens an in-depth understanding on the operational system engaged by the local traffic department.

The last section of this chapter presents the data quality as the key part of the discussion. This section covers subsections surrounding the background of the data quality. The section offers knowledge on the interrelationship between the data producers and data users in the road traffic system. This part explains the benefits of considering feedbacks or critics from the data users in the road traffic system, with regard to the level of the quality of data disseminate to the public. The section further offers studies on the importance of data quality, problems with data quality and data quality components. Further discussions are accessible on the cases of underreporting in South Africa, literature studies performed on the improvement of data quality, and quality improvement strategies applied by previous researchers.

2.1 Road traffic accident [RTA]

In this present day, in most regions of the world, RTA is becoming everyday problem in the road transportation system compared to other sectors of transportation systems such as aircraft, railroad and sea. Many authors through research papers published (IRTAD 1998; Stats SA 2009; WHO 2009; WHO 2011; WHO 2013; RTMC 2014), developed peculiar definitions for the description of the RTA. However, according to a report published by IRTAD (1998), United Nations/Economic Commission for Europe [UN/ECE] defined RTA as “an accident which occurred or originated on a way or street open to the public traffic; which resulted in one or more persons being killed or injured and in which at least one moving vehicle was involved” (IRTAD 1998). Similarly, RTMC (2014) described RTA as “a single road traffic incident, regardless of the number of vehicles or persons involved in any particular accident. A road accident happens when two road users, regardless of mode of travel, try to occupy the same road space at the same time” (RTMC 2014).

Broadly speaking, the above quoted definition by the RTMC (2014) could be considered as the standard definition adopted in South Africa for RTA. In addition, the event of the RTA is apparently related to the collision or accident effect between two or more mobile objects such as an accident between multiple vehicles, a vehicle and a train, a vehicle and a cyclist, a vehicle and a pedestrian, and a vehicle and an animal, or immobile object and mobile object such as a vehicle and a fixed object, or a single vehicle overturned or a single vehicle left the road (IRTAD 1998; RTMC 2014).

Subsequently, in any case of reported RTA, the cause of the collision/accident together with the involvement of the driver are investigated by the police officer inspecting the scene of the accident. On the other hand, the technical aspect is carried out by the traffic engineers through the processing of the information collected from the accident scene, with the assistance of the accident reporting officer, who could be a traffic officer or police officer. Furthermore, a global report released by World Health Organisation (2013) linked the incessant occurrence of RTA to some key factors such as poor legislations, human errors, poor road maintenance and poor accident recording systems.

To complement the above statement, Sue Sinclair (2004) made a statement in a book regarding the significant of safety in transportation system that, “a key requirement of the road transport system is to provide safe and efficient travel for its users” (Sinclair 2004). The publication further provides a comprehensive understanding into the three major components of transport system, which are *the road user, the vehicle* and *the road environment.* In addition, the publication declared that failure of one or two of these three related components will certainly result to an upsurge in severe accident occurrence, which could be attributed to the interplay between human factors, vehicle factors and road environments (Sinclair 2004; Vogel & Bester 2005).

Predominantly, the interplay between these three factors can be administered or managed if the transportation departments or agencies keenly carried out their tasks accordingly, by enforcing the traffic law without any compromises, and to design, build and maintain the road infrastructures. If these tasks could be achieved, then the road users are anticipated to consider the demand of the road environment, by performing the task required of them to avoid accidents, through regular obedient towards the road safety rules and regulations.

Although, despite the influence of the aforementioned factors, the number of RTA recorded in the developed regions is much minimal to the number of RTA recorded in the developing regions. It is realised through studies carried out that, developed countries capitalised 60.0% of their safety programmes on human behavioural characteristics (Salmon et al. 2006). In fact, a huge percentage of road collisions or crashes/accidents taken place in these countries are understood to be instigated by human behavioural characteristics (World Health Organisation & The World Bank 2004; Jungu-Omara & Vanderschuren 2006; Salmon et al. 2006; Mikulik et al. 2007).

Apparently, the huge level of high-tech evolution, economic growth and impartial managerial structures placed the developed countries far ahead of developing countries in combating the problems of poor road safety systems. In other words, the developed countries possess the right tools and capacity to battle the incessant increase in the number of the RTA than many developing and underdeveloped countries (Derriks & Mak 2007; WHO 2009; WHO 2011; WHO 2013; WHO African Region 2013).

It is evident that road transportation development is often a priority to any country; on that point, a constant evaluation of the road traffic issues in any particular cities, provinces or countries is required. If this could be regularly executed, without a doubt, road accident occurrence rate could be practically lessened, and probably enhance the ability of the traffic engineers to have conversant ideas on the most problematic areas along the roads that could contribute to the occurrence of road accidents. These ideas could facilitate a better approach, with the intention of securing the safety of the road users.

In any case, transportation departments are mandated to deliver reliable and improved services to combat rampant occurrence of road fatalities, which are considered critical to a country's economic growth (Transportation Association of Canada 2006). These are the departments established to manage, and supervise the affairs of the transportation system in any countries around the globe. The only way these departments can achieve this is through a reliable standard process. The process can strengthen the safety plans developed by the departments, in an attempt to daunt any kind of deterrence that could be affecting the decision making on traffic matters.

Without question, the departments are important establishments that require dependent operations, with the intention of consolidating their service efficiency with valuable results to minimise criticisms from the public. Moreover, the obligations of the transportation departments can be depicted in five words such as monitoring, securing, constructing, maintaining and managing of all transit activities (Transportation Association of Canada 2006). If these five words are not applied within the context of a standard measure, therefore, it could result to an upsurge in road mortality rate, which could perhaps propel high demand for quality service by the public. In that case, the effective execution of these obligations is being handicapped due to lack of sustainable development in the areas such as transportation safety, reliable management structure, standard accident recording system, and others include road maintenance and construction procedures. Further understanding on the role players [stakeholders] in the transportation departments in South Africa will be provided in the Appendix A-A.9, alongside their assigned responsibilities towards ensuring a standard measure in the reduction of the daily occurrence of road accidents.

2.1.1 Global effect of RTA

The prevalent cases of road collisions resulting to serious injuries and death, has often being considered as a huge challenge to the public health (Stats SA 2009; WHO 2013). In fact, this has always been a major concern to the rest of the world. Without a doubt, many lives and properties have been destroyed at the detriment of the poor transportation management system. In a recent report issued by World Health Organisation (2013), the capacity of the transportation department was tasked on the need to stop or drastically minimise the rising number of road traffic injuries and deaths in most parts of the world. The global organisation further stressed that there is every possibility that daily increase in the number of vehicles on the roads may contribute to the rising number of road traffic injuries and deaths in most parts of the world (WHO African Region 2013). This effect could be as a result of the rapid development in global technology, wherein many different vehicle manufacturers have sprung up to compete with the existing ones, by producing sophisticated vehicles with reliable interior gadgets for the vehicle users.

Furthermore, prediction made by the World Health Organisation regarding road traffic injuries emphasised that; road traffic injuries would intensify from its present position as the 9th leading cause of death to 5th position compared to other types of cause of death by 2030 (Sinclair 2011; WHO 2013). This could be possible provided that no serious attentions are exercised towards the reduction of the road traffic injuries or fatalities in some regions around the world.

In essence, sloppy management activities in the transportation departments have caused a huge blow to the human capital, which has a direct impact on the economy of a country. In addition to the above statement, Tahera Anjuman et al (2007) acknowledged in a published article that many lives and properties have been destroyed, numerous number of road users are incapacitated due to the lack of improvement needed in reducing road traffic casualties (Anjuman et al. 2007).

Nevertheless, according to the statistical result released by the World Health Organisation (2013), only 15.0% countries as at 2012 have comprehensive legislation on the major risk factors to support road safety which is considered very low, whereas the remaining 85.0% countries without comprehensive legislation supporting road safety tend to have a high score of vulnerable road users as the major victim of RTA. In view of having vulnerable road users as the key affected factor in the road accidents, some developed countries like Netherlands, Norway, and United Kingdom have redesigned their roads and readjusted their regulations in order to accommodate the vulnerable road users (WHO 2009; WHO 2011; WHO 2013).

In summary, the above statement illustrates the degree at which developed countries tackle the issues of road accident occurrence rate than the developing countries. On the contrary, the involvement of the developing countries in tackling these problems is considered unimpressive according to the global evaluation report published by the World Health Organisation in 2013 on road safety issues (WHO 2013). The evaluation report revealed that developing countries have contributed low or little in minimising the effect of RTA on the road users and their localities (WHO 2009; Vanderschuren & Jobanputra 2011; WHO 2011; WHO African Region 2013). This implies that little or no improvement has been contributed so far by developing countries on the need to tackle consistent increase in the road fatalities and injuries compared to developed countries. This slack improvement in the developing countries contributes hugely to the global daily estimate of 16,000 deaths attributed to road traffic injuries since 2007, which is indeterminately expected to decrease in the subsequent years (Anjuman et al. 2007), as it accrued to the annual estimate of 1.3 million deaths over 4 years (WHO 2009; WHO 2011; WHO 2013; World Bank & World Health Organisation 2013).

Beyond question, the poor run of the developing countries in planning standard measures to reduce the incessant occurrence of the RTA, could be connected to the low performance rate in the area of reliable quality data assessment system. Due to this effect, road safety engineers, and other data users find it challenging to understand the practicality of the results obtained from the analysis performed on the road accident data processed. Nonetheless, road accident data are applied as the primary tool for developing reliable and appropriate countermeasures, in resolving the possible causes of accidents on the roads. More discussions will be provided on the road accident data in the subsequent sections.

2.1.2 State of RTA in South Africa

Most research projects carried out in the domain of RTA, were focused on the road traffic causalities, mainly on the need to lessen the occurrence of RTA. More so, few out of these articles, criticised the quality of data provided by the custodian of road accident data in South Africa (Lotter 2000; Vogel & Bester 2004; Vanderschuren & Jobanputra 2011; Sinclair 2011; Joseph 2013; Twala 2013), while more than 80.0% of the articles particularised on the suggested countermeasures that might be relevant in the reduction of accident fatalities, and the factors affecting the improvement of the road safety system in South Africa (Vanderschuren & Irvine 2002; Roux & Sinclair 2003; Sinclair 2004; Vanderschuren 2006; Jungu-Omara & Vanderschuren 2006; Vanderschuren & Jobanputra 2011; Sinclair & Murdoch 2012). Nevertheless, large number of road accidents has been recorded across all the nine provinces in South Africa for over a decade (Stats SA 2009; Stats SA 2010; Stats SA 2014). Evidently, it was publicised that land transport contributed largest proportion rate of 99.8% to the total transport accident deaths in South Africa (Stats SA 2009). However, stakeholders overseeing the affairs of the transportation safety in South Africa have managed over the years, to introduce many substantial ways of controlling the annual rise in RTA in all areas. Although, many of these approaches are acquired through the feasibility study carried out on the road safety systems employed in the developed countries, while others are locally developed to suit their environments. Some of the considerable ways used in controlling the rise of RTA are outlined below as:

- **Routine application of liquor/alcohol intake measuring device called Alcohol Breathalysers [AB];** which assists in determining the level of blood alcohol concentration [BAC] consumed by a driver/cyclist as required by the Technical Committee for Standards and Procedures for Traffic Control and Traffic Control Equipment [TCSP]. In addition, this step is taken by the law enforcement agency to discourage the drivers from indulging in the habit of risking their lives, and that of the other road users (Baguley 2001; van Niekerk 2012).
- **Persuading the use of seatbelts and helmets while driving;** to protect loss of lives and injuries in the event of any unplanned incidents (Jungu-Omara & Vanderschuren 2006).
- **Regular road checks routine;** to ascertain careful driving activities on all the South African roads (RTMC 2012).
- **Installation of Speed Measurement Equipment [SME] and Traffic Light Violation Monitoring Equipment [TLVME] at strategic places along the South African roads;** to compel the road users to obey speed limit directives as required by the

Technical Committee for Standards and Procedures for Traffic Control and Traffic Control Equipment [TCSP] (RTMC 2012).

- **Coordinating continual road safety education programmes;** to educate the road users on the dangers that maybe encountered on the road, as a result of human behavioural characteristics, such as *Safety Home* campaign programmes and other safety programmes (RTMC 2012).
- **Regular maintenance and reconstruction of roads;** to accommodate more road users according to their classes, and also to support better safety plans in the future (SANRAL 2008; SANRAL 2012).

In fact, according to the South African National Roads Agency Limited [SANRAL] annual report, estimate of about 1,000,000 [million] crashes/accidents occurred in a year in South Africa (SANRAL 2008; SANRAL 2012). The impact of these crashes/accidents resulted in more than 40 deaths a day, which accumulated to 14,000 deaths annually, wherein 20 people were permanently disabled and several hundred of serious injuries were suffered at the rate of 150,000 injuries per year (SANRAL 2012; RTMC 2012).

In addition, giving credit to the above statement, a conference paper submitted at 25th Southern Africa Transportation Conference in 2006 by Jungu-Omara & Vanderschuren (2006), affirmed that RTA claimed between 13,000 to 14,000 lives in South Africa. The study offered valuable awareness into the factors that led to the incessant increase in the RTA fatality in South Africa. Moreover, the authors disclosed that pedestrians are most vulnerable in the road accidents in South Africa. Although, the authors emphatically criticised issues such as human behavioural characteristics, vehicular conditions, and road maintenance issues as the paramount causes of accidents in South Africa.

In South Africa, according to the previous studies, highest number of fatalities and severity of road accidents are recorded mostly during festive periods, in every Christmas period in December particularly in the urban regions. From these studies, statistical information showed that 9,918 fatal road accidents were recorded in December 2002 alone, with 78.0% of the accidents attributed to human factors, while road factors and vehicle factors contributed 12.0% and 10.0% respectively (Jungu-Omara & Vanderschuren 2006; Stats SA 2009). Stats SA [Statistics South Africa] (2009) statistical report claimed that most road casualties resulting to death in South Africa happened within the vicinity of the accident victims' usual residences with a large estimate of 83.4% (Stats SA 2009).

Considering the above statement, Jungu-Omara & Vanderschuren (2006) suggested that South Africa as a developing country should emulate the international practices' ways of reducing fatalities and severity of RTA (Jungu-Omara & Vanderschuren 2006). The study further remarked the leads to the reduction of RTA in South Africa as: firstly, by identifying the prominent factors responsible for the harsh state of casualties in South Africa; and secondly, establish an educative programme to guide the public through the process of adjusting to the new and old traffic rules and regulations, in order to avoid unanticipated occurrence of accident; and thirdly, the traffic engineers should be ready to devise countermeasures towards drastic reduction in the RTA in South Africa (Jungu-Omara & Vanderschuren 2006).

2.1.3 Economic effect of RTA in South Africa

In the recent times, many global research reports published regarding the state of RTA in all regions of the world, emphasised that Africa region remains the only region with the highest number of RTA in the world (IRTAD 1998; WHO 2009; WHO 2013; WHO African Region 2013). Additionally, in a research report published by the World Health Organisation-Africa Region in 2013 on the situation of the road accident in Africa; an overview of the report demonstrated the rate at which accidents occur on the African roads, and the quality of improvement implemented in reducing the negative effects of the roads on the public [road users] (WHO African Region 2013). The report further proclaimed that the issue of poor growth in the road safety system in Africa can be attributed to the inability of many African countries to provide sufficient data, and resources to facilitate feasible improvement in the RTAs. Due to this effect, many countries in Africa lack the capacity to stabilise the output of the transportation managements in charge of road transportation safety, in making reliable decisions to combat the incessant occurrence of RTAs.

Furthermore, the statistical information made available in the report showed that middle-income countries suffer greatly in the rise of RTA with 20.1 deaths per 100,000 population (WHO African Region 2013), which is considered the highest in 2013. In fact, among the middle-income countries falls South Africa as the 2nd highest in the road fatality rate in Africa after Nigeria, with the highest fatality rates of 31.7 and 33.7 deaths per 100,000 population per year respectively as at 2013 (WHO African Region 2013). This is a serious issue that beckons for consistent improvement of the road safety policies guiding the use of the road, and the frequent monitoring of the affairs of the department in charge of the road transportation system in South Africa, in order to ensure safety of lives and properties. However, the public [road users] faulted the rise in the road accident fatality rate on several issues such as:

- **Incompetence of some traffic officials** –this factor incapacitates the effort of the local traffic department in charge of ensuring a consistent reportage of RTA (Joseph 2013).
- **Immoral behaviour of some traffic officials** –this factor disenables the effort of acquiring the right information regarding a particular accident. On account of this, many accident cases are underreported or unreported.

- **Inability of the RTMC to exert necessary strong leadership** –this issue renders the lower units inefficient towards their allocated duties. The effect of this slackness or negligence have cost the management the inability to get the best performance out of these units through the execution of the allocated projects in accordance with the standard procedures established (Adams 2001).
- **Lenience of the traffic management** –this factor undermines the effort of the management from enacting strong traffic laws towards ensuring consistency in the road safety system. This permits the road users to abuse the use of the road infrastructures outrageously at the expense of their own safety.
- **Decline involvement of the metros and municipalities** –this issue limits the ability of the RTMC from acquiring more reliable results regarding the road traffic issues. And it also affects the ability of the local authorities to support the initiatives set by the provincial authorities in ensuring stable operational procedures.
- **The poor handling of vehicle ownership** –this hampers the ability of the local department to ascertain the number of the vehicles that are roadworthy, and thereby supports the capacity of the local department in reducing the volumes of traffics faced by the road users every day (O' Day 1993).
- **The poor legislative system** –this allows the road users or public to violate the use of the road infrastructures to their own detriment.
- **The poor handling of road accident data** –this slows down the prospect of procuring adequate data required to support everyday activities like disbursement of resources, scheduling of duties, development of countermeasures to prevent accidents, and many more.

Even though, with the level of road mortality rate in South Africa, the country is still ranked among the middle-income countries in Africa with the advanced road safety systems. Regardless, one could envisage the reason behind the inconsistency in the accident reportage system in South Africa. In fact, the gravity of the issues highlighted above incapacitates the ability of the management to have uniform distribution of resources, funds and personnel to combat casualty in RTA in South Africa.

Actually, incessant occurrence of RTA on the South African roads has become the fear of the public, or the commuters plying the roads on a regular basis (Stats SA 2009; Stats SA 2010), despite a high number of good roads across the country. The annual increase in the number of casualties recorded in South Africa poses further threat to the health and economy of the country (Stats SA 2009). Moreover, in the course of lessening the effect of RTA on the public, South African government, through the Ministry of Transport [MoT] has suffered economic cost worth up to R133,000,000,000 [billion] annually as affirmed in the report published by SANRAL in 2012, but recently risen above that amount from the R210,000,000,000 [billion] to R300,000,000,000 [billion] (SANRAL 2012). Literally, this affects the South African economy, among other sectors that demand more distribution of resources. Therefore, it requires drastic measures to avoid further deepen economic cost effect in the transportation sector. Besides, a suggested way to carry out the reformation exercise, is to improvement the quality of the road accident data collected, validated and processed at any local, provincial and national traffic departments within South Africa.

2.2 Road accident data

Several countries persistently search for comprehensive knowledge on how to improve and stabilise their countries' transportation safety plans, with the purpose of having the substantial information to tackle the issue of RTA. Actually, across the globe, huge investments are continually made to ensure constant improvement of the road transportation system, as previously stated in the last paragraph in subsection 2.1.3 above. Obviously, quality road accident data plays a significant part in the decision-making process executed through the Ministry of Transport [MoT], because it provides a better opinion on the areas that require financial inputs.

The data collection process is seen as a reliable way of ensuring regular collation of relevant road accident data, required to enhance uniform distribution of resources and management improvement within the traffic department. In a different view, the process plays a crucial role in devising a preeminent approach in determining the possible causes of RTA, and concurrently offers suggestions on reliable ways of reducing frequent occurrence of RTA. Technically, road accident data is applied by the traffic experts to question why some situations are obstinate to change in the road traffic system. Admittedly, road accident data aids the ability of the traffic experts to gain practical understandings on the suitable countermeasures applicable in reducing the road traffic issues.

Moreover, the accuracy of the road accident data depends on the type of data collection techniques implemented; purposely to enhance the perfect reporting and recording of road traffic incidents across South Africa. Nonetheless, the data users find the accuracy of road accident data important in various ways (Ferrante et al. 1993; O' Day 1993; Rosman 2001; Njord et al. 2005). A report released by Suzette Thieman (1999) described the importance of road accident data as the only possible way of determining the need for improvement in the road safety management, and the type of precautionary measures required to be implemented, in order to avert the occurrence of accidents (Thieman 1999).

The report further emphasised the significance of road accident data in the establishment of a Geographic Information System [GIS], with the objective of relieving the stress undergone by the Transportation Safety Engineers [TSEs], regarding the technical analysis surrounding the location [street/road] where accidents occur most. More discussions are provided in the subsection 2.2.1 on the significance of road accident data to the data producers [RTMC] and the data users [public], whereas literature studies on road accident data collection and the techniques applied for data collection procedure is provided in the subsequent sections and other relevant details are provided in the Appendix A respectively.

Lack of sufficient information is classified among the problems frustrating the ability of RTMC from producing reliable road accident data, and also undermines the effort of the TSEs in developing sustainable methods to tackle accident occurrence. This particular problem is connected to three major areas where errors regularly transpired, such areas as inappropriate application of the ARF, accident reporting, and the accident recording of the RTI (Baguley 2001; Sinclair 2011). Additional information will be provided in subsection 2.2.4 on the areas that are empirically vulnerable while collating road accident data.

2.2.1 Significance of road accident data

A few practical applications reflecting the significance of the road accident data in the road transportation safety system was concisely mentioned in section 2.2. Road accident data is a significant constituent in transportation safety engineering. It is also applicable in many areas of transportation systems such as administrative management, road construction and maintenance, safety implementation programmes, safety education, research and development, and decision-making in distributing resource capacities.

However, the practicality of the road accident data in the aforementioned areas depends exclusively on the quality of the data sourced from the primary level. This aspect could be achieved through the establishment of consistent data collection procedures, reliable and productive management systems, and valuable feedback from the data users [data consumers] as a yardstick for data quality evaluation. Moreover, road accident data comprises the major part of the required elements for the annual statistics on road traffic casualties, as a prerequisite for financial budgeting through the national level to reinforce road safety initiative programmes. These programmes are established basically to solidify the public awareness campaign initiatives intended to guide the road users ahead of unforeseen dangers on the roads (O' Day 1993; Thieman 1999; Baguley 2001; Department for Transport 2011).

As earlier described in section 2.2 above, road accident data is valuable in the development of corrective measures devised by the TSEs, and also significant in the duty scheduling by the traffic departments, police departments, or law enforcement agencies in deploying patrols as a guide towards the operational tactics required in the reduction of road casualties (Department for Transport 2013). In fact, research institutes utilise the outcome of the data collated to develop better road safety measures to support the existing ones (O' Day 1993; Baguley 2001; Department for Transport 2011). In addition, Chris Baguley (2001) emphasised on the essentiality of improving the quality of the road accident data. He added that road accident data should be seen as the top priority for the transport department [RTMC], in order to promote the regular evaluation of the improvements expecting from the various activities executed (Baguley 2001).

Also, road accident data is considered as a vital tool in decision-making procedures at the local level, provincial level and national level. Categorically, at the local level, road accident data is applied to ascertain the exact areas where safety problems are becoming paramount along the roads within their jurisdictions. Observably, the possibility of identifying these exact areas depends on the reliability of the data collected at the level. This particular jurisdictional level establishes a solid relationship between the local authorised agencies such as police department and the traffic department, with the purpose of ensuring a reliable processing system for road accident data collection. Similarly, the local level establishes relationship with the provincial level managing the affair of the entire province (Baguley 2001), in order to facilitate successful distribution of the bulk road accident data collated through the accident reporting process.

Conversely, at the provincial level, the decisions concerning the traffic safety measures developed are made (Baguley 2001), and thus executing the appropriate testing and implementation of newly developed safety measures to ascertain their positive effects on the public. Unlike the local level, analysis of the data collated are carried out at the provincial level, through the use of an advanced road accident database system, with the determination of using the analysis results obtained to predict areas that required immediate improvement across the entire provincial area. The processed data made available are post-analysed by data users, who are professionals in the domain, to further evaluate the quality of the data through its practicality to determine the areas that require utmost improvement.

Unlike the provincial level, the national government oversees the national matters relating to the national road safety programmes (Baguley 2001; Sluis 2001). The decisions implemented at the national level are based on thorough evaluation of the national road safety reports prepared by the provincial level. Occasionally, the decision may transpire the need for new policies to be established, with the intention of supporting the existing ones, and as well the process for regular distribution of resources to curb unwanted behaviours of the road users on the road. On one hand, this strengthens the safety initiative programmes developed, and sponsored by the national government, with the purpose of educating road users on the need for the use of safety procedures and checks at their disposal.

2.2.2 Road accident data collection

Road accident data collection, in relation to RTA, could be described as a process that simply initiates the process of sourcing relevant information on a particular road accident through acceptable standard collection procedures established. The procedure is initiated from the application of ARF to the reportage of the accident, and then followed by the recording of the accident. The data collection procedures are relational data processing sequences, which connects the operations carried out in each unit involved. Figure 1 simply illustrates the four basic procedures followed in the processing of road accident data.

Basically, collection of road accident data should be thoroughly based on the definition of the data fields, because it is necessary to understand the purpose of the data, or what the data will be used for (Ehnes & Niu 2012). Generally, the relevant data elements [data fields] that are most paramount to the present state of the transportation safety in any country should be considered (Ehnes & Niu 2012). According to Ehnes & Niu (2012), "the more fields or entities are defined, the more related answers will be made available for evaluation; clearly, the administrative burden on reporters is increasing, making them less likely to deliver comprehensive and correct data" (Ehnes & Niu 2012). In addition, the data collection procedures were established according to the guidelines required to direct the units participating in the process through the activities involved. In South Africa, the procedures employed for the collection of the road accident data are administered by the RTMC as stated in section 1.1above. This particular department is recognised as one of the credible stakeholders authorised to manage the affairs of the road transportation systems in the country.

Nevertheless, Gorell (1997) stated that, "an efficient accident data collection and analysis system is a basic requirement for any country determined to tackle its road safety problem" (Gorell 1997). This statement demonstrates the initiative of gathering the relevant and accurate information, with the purpose of providing additional knowledge on the cause of RTA. The informative data fields provided in the data form [ARF] serve as guidelines to the reporting officer, through the process of obtaining relevant information for a particular type of accident. This depicts the significance of the ARF in the procedures designed for collection of road accident data. Concisely, the design and arrangement of the ARF have a vital role to play in the quality of the data collected. Additional information regarding the ARF is discussed in Appendix A.

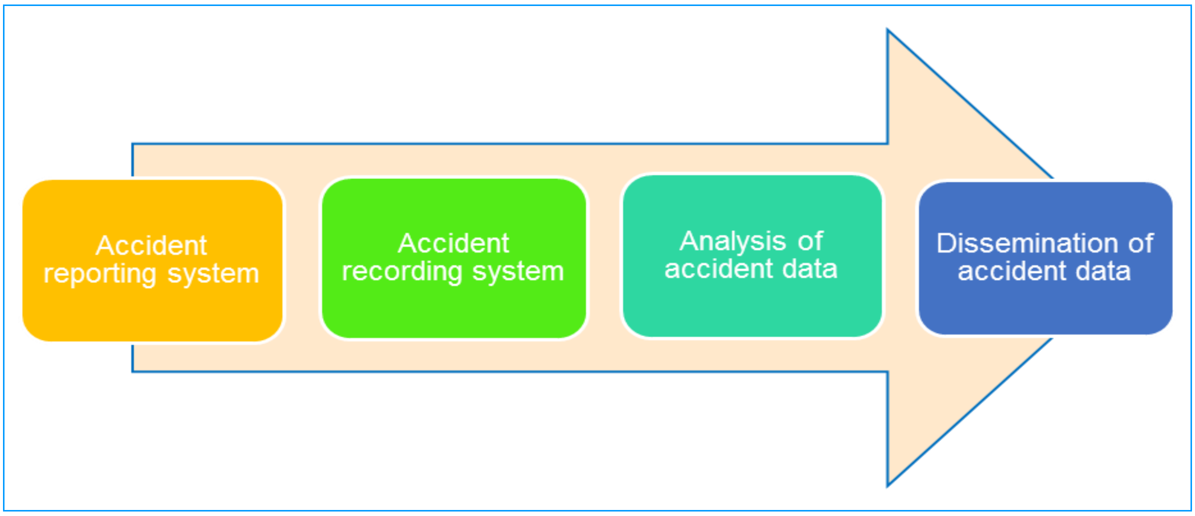

Figure 1: A diagram depicting the four basic procedures for road accident data processing

After completing the required information in the ARF, the completed form will be submitted for proper evaluation process, to determine the quality level of the data collected. The process provides an understanding on the level of quality of the data collected. The information provided in the ARF is reassessed in order to control the underreporting issues, which are categorised as missing data, insufficient information, misrepresentation and misinterpretation of information. Thereafter, the information on the form is captured into the road accident data capturing system, and validated for analysis purpose through other significant process sections such as *data pre-processing, data entry* and *data processing.* These units manage the processes of data coding and entering of data into the road accident database, and simultaneously performs the analysis of the data appropriately, to ascertain the level of improvement in various areas of the transportation system. The subsequent sections contain the discussion covering the four basic procedures in road accident data processing.

2.2.2.1 Accident reporting system

This procedural system involves the driver who was reportedly involved in a traffic accident, and the police officer or the traffic officer, who inspected the scene of the traffic accidents and reported them to the nearest Municipal/Metro, Traffic Department [MMT] office, or South African Police Service [SAPS] station. Basically, a driver involved in an accident along a road is compelled by the National Road Traffic Act No. 93 of 1996, to report his/her involvement to any of the nearest authorised local authorities mentioned above within 24hrs of the incident, and delegates will be dispatched instantly to monitor the circumstances on the accident scene. Thereafter, depending on the severity level of accident *[that is, only where case docket is opened/registered],* the traffic police is expected to measure the magnitude of the accident, and keep the record of the incident for reference purposes.

In the course of obtaining the necessary information as required, the police officer may determine the need for further investigations depending on the degree of severity of the accident. Primarily, a road accident involving death of a person or persons, is regularly handled by the police officer (Sluis 2001). Furthermore, the information collected necessitate a thorough review with the intention of detecting errors committed during the data collection process. The supervisory officer assigned to this task must be ready to draw attention of the reporting officer to errors detected. Actually, this process is performed with the intention of instilling accuracy and consistency as part of the strategic components that constitute a good quality data (Baguley 2001). If this could be achieved, minimal errors would be detected at the accident reporting section.

2.2.2.2 Accident recording system

In this particular unit, the information capturing unit at the local traffic department, or the MMT captures the information on the ARF, provided that the information is accurate and sufficient. At this level, the information is assessed, certified and captured before the daily quantity of the data collated could be transferred to the provincial department managed by the RTMC. However, operations like *data validation, data pre-processing, data entry* and *data processing* are concurrently carried out as the fundamental approaches of ascertaining the quality of information transformed into a complete data (Baguley 2001; Sluis 2001). The implementation of the data validation process stimulates the proper assessment of the information received from the MMT, and process it into data for analysis purpose.

Actually, the data capturing system designed for recording, updating and analysing of the road accident data offers more insight into the information received from the MMT. It provides a template arrangement of the information before it could be processed into data format. Moreover, some segments of the database, execute different tasks in the analysis of the road accident data. On one hand, this unit serves as the intermediary in the processing of road accident data collated from the accident location. Furthermore, the unit is recognised as the central link between the MMT and the national government. The unit executes tasks based on the feedback from both the local level and the national level.

2.2.2.3 Analysis of road accident data

This unit is placed under the supervision of the RTMC. The analysis of the road accident data intra-transfer from the *'Accident Recording System Unit'* to the *'Data Analysis Unit'* is performed within the RTMC bureau (O' Day 1993; Baguley 2001; Njord et al. 2005). This unit provides better understandings on the state of the RTA in any parts of the country (O' Day 1993). The analysis result may further reveal the most possible causes or contributory factors to different types of traffic accidents, and suggests appropriate steps that could be suitable to rectify or reduce the causes. Technically, the trend displays through the graphical illustrations of the analysis performed can predict the focus areas that required immediate improvement.

2.2.2.4 Dissemination of road accident data

This unit performs all the tasks involving circulation or dissemination of the annual analysis report regarding all types of accident such as fatality, serious and slight injuries, and other important areas that comprise of statistical data (Baguley 2001). In short, the unit evaluates the analysis of road accident data and disseminate the report of the analysis to the public for further evaluation. Actually, this unit educates the national government, authorised divisions and the public on the efficiency level of the safety measures implemented, and the need to strengthen the existing systems with the necessary resources to improve the road safety plans (Baguley 2001). Besides, the outcome of the analysis will reveal if the traffic regulations enacted are observed regularly by the road users. Thereafter, the annual analysis report will be made available to other bodies for references, such as researchers, lecturers, insurer, and amendment to legislation.

2.2.3 Elementary technique for accident reporting

The elementary technique developed for accident reporting system is similar in most countries. The technique begins with steps taken by the accident victim or the driver of the vehicle involved in the accident, who is compelled by the law to report the incident within a specific period of time [24hrs] to the nearest police station (Sluis 2001; RTMC 2007; Department for Transport 2013; RTMC 2014) as briefly mentioned in subsection 2.2.2.1. Afterwards, a police or traffic officer is authorised to document the accident information for significant motives, such as identification of responsible parties, formulation of countermeasures, legal actions, and government budgeting etc. However, in some countries, police is preferred to be in charge of all types of road accidents involving human injury and fatality except property *'damage only'* accidents (Sluis 2001; RTMC 2007; Department for Transport 2013).

Hence, the police officer or traffic officer is expected to complete the ARF with the utmost accuracy, and sincerity in agreement with the available facts at the location of the accident. In any case where few details of the incident are collected, further enquiry is to be carried out to substantiate the degree of insufficient information. Subsequently, a completed copy of the ARF is to be submitted or transferred to the nearest local traffic department overseeing the jurisdiction [SM-WC 024]. In a case where incomplete information is discovered, a filed copy of the incomplete ARF will be returned to the local police department for proper rectification.

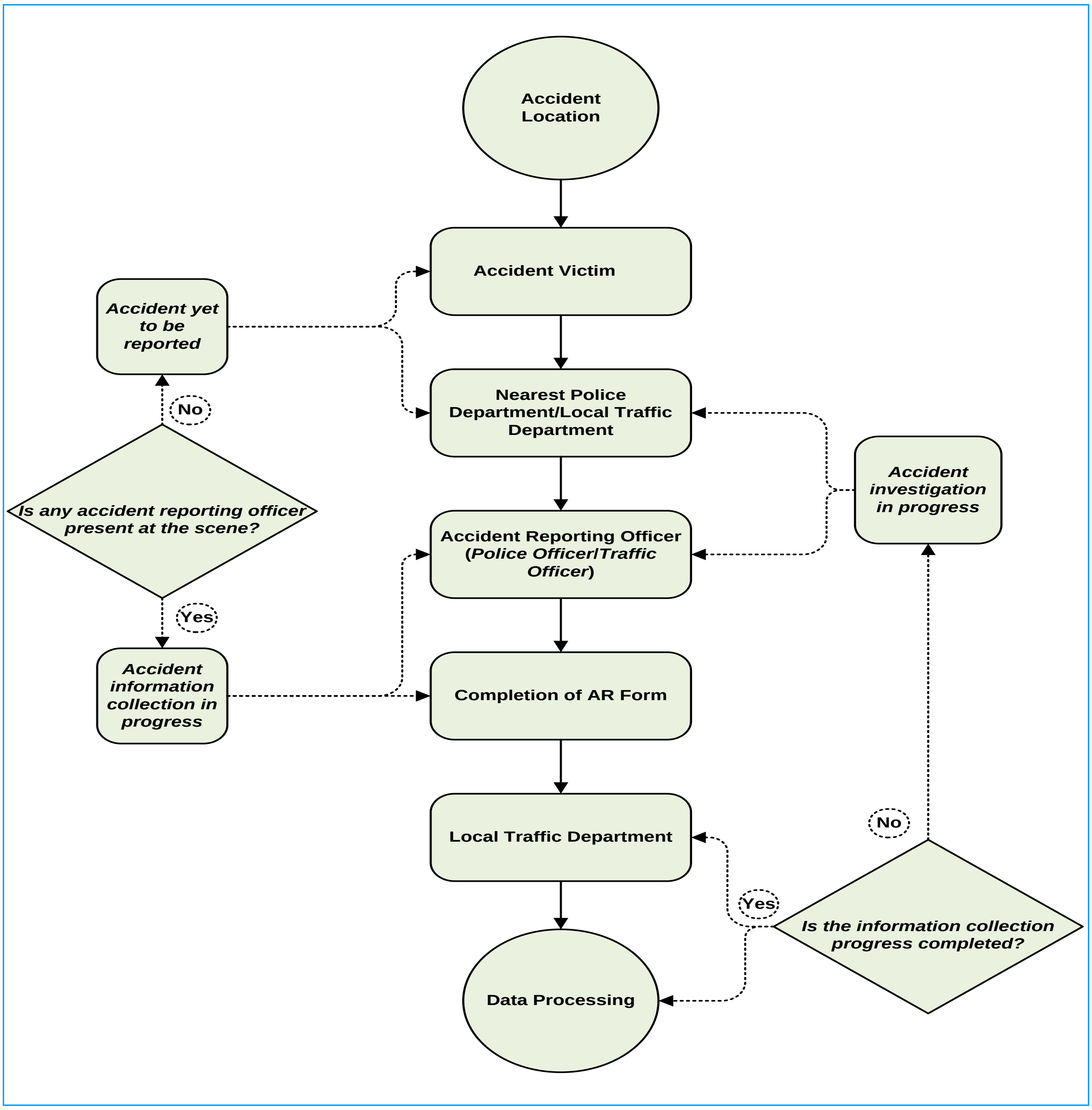

Figure 2: A model depicting the technique for accident reporting

At this point, the police officer who formerly reported the accident, is only authorised to implement the necessary changes expected, in agreement with the complaints of the DCO or prior to the suggested corrections requested by the supervisory officer at the local police department. After necessary the corrections have been made, rectified copies will be forwarded back to the DCO for final assessment before the road accident data could be captured into the local database system. Then, all data captured will be transferred to the provincial level as scheduled for further assessment procedures. At this level, the assessment is expected to determine the degree of reportage reflecting in the completed ARFs validated at the local traffic department. Thus, a complete definition of a completed ARF could be given as *'a data form that comprehends an extensive reportage of RTI, which is validated, approved* *and captured into the local dataset 'IPAS''*. Figure 2 shows a complete illustration of a common technique used for road accident reportage.

2.2.4 Collective problems in road accident data collection

In the section 1.1, a list of problems frustrating the efforts of the RTMC was presented. The alleged factors responsible for deficiencies in the road accident data handling were concisely illustrated in the Figure 3 below. These factors exist along the process line designed to support the efforts of the stakeholders involved in the procurement of a reliable road accident data quality, starting from the local or municipal level to other higher levels.

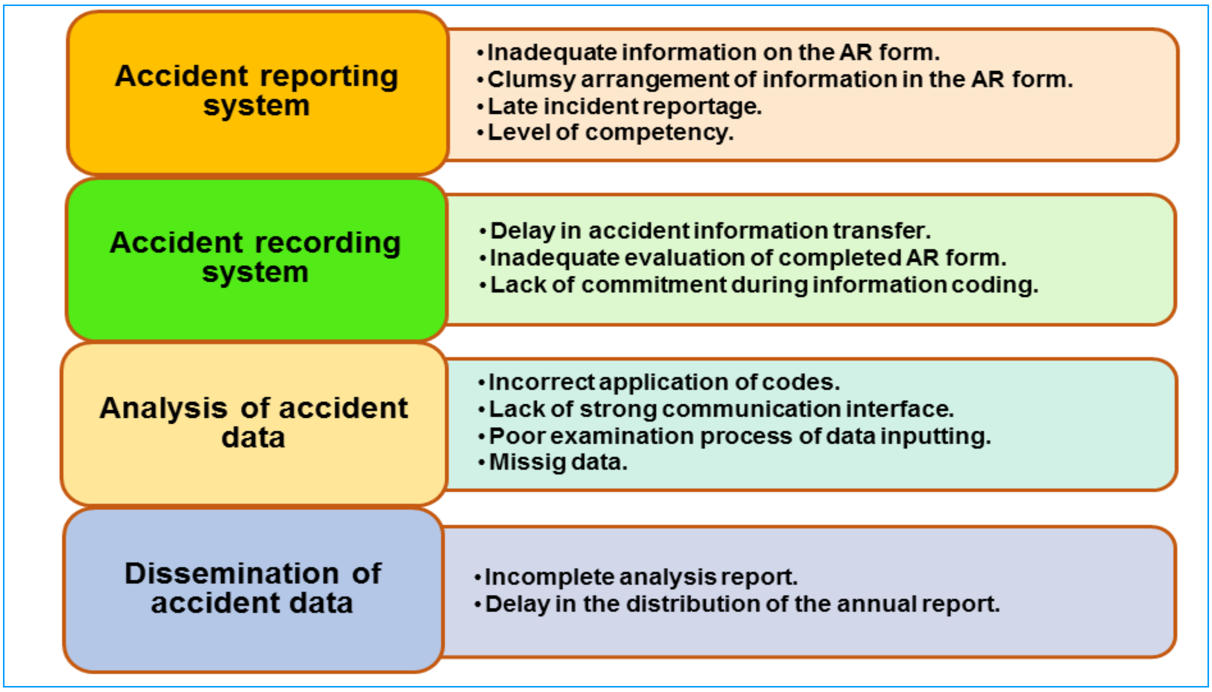

Figure 3: Illustrating the deficiencies experienced in each stage of road accident data handling

Although, it is yet to be practically proven that the listed problems below are responsible for the poor production of quality data at the STD. Ultimately, the process of determining the actual factors that are responsible for the poor quality of data, as claimed by the public [data users] can be achieved through the evaluation of the data acquired at the local traffic department, and also through survey techniques. The approach used in performing the two mentioned tasks, along with the expected results are discussed in the Chapter 3 and Appendix E.

Recall in subsection 2.2.2, a concise discussion was provided on the activities performed by the four basic procedures implemented in processing road accident data. However, the collective problems affecting the *accident reporting system* based on South African context are elaborated below.

- **Inadequate information on the ARF** –this is one of the factors observed under the accident reporting system, as one of the contributory factors in the issue of poor quality of road accident data. However, this particular factor deters the chance of gathering more detailed information on any reported road accidents (Vogel & Bester 2004).

- **Clumsy arrangement of the information in the ARF** –this factor is conceived as one of the problems facing the appropriate interpretation of the information provided in the ARF, and this makes the form boring to the users (Ehnes & Niu 2012).
- **Late incident reportage** –this action involves the initiatives of three persons in the reporting of a road accident within its period of occurrence. The problems identified here are the attitudes and commitments of the three parties in reporting the incident at the time of its occurrence. The three parties involved are briefly described below.
    - **The victim of the accident** –this person is compelled by the law to report his or her involvement in a road accident to the nearest police station within 24hrs.
    - **The police officer/traffic officer** –this officer is expected at the scene of the accident to gather or document information by completing the ARF accurately based on the dimension of casualty of the accident.
    - **The witness** –this person observes the occurrence of the accident at the time of its occurrence; and he is urged to report any occurrence of accident to the nearest police station, provided that the accident victim is not in the right state of mind to do so, and his willingness to offer additional information on the accident to the police officer or traffic officer in control of the circumstances.
- **Level of competency** –this factor demonstrates the degree at which the form users comprehend the application of the ARF. The influence of this particular factor in the accident reporting system, is marred by the inefficiency of the form users in applying the ARF appropriately.

The problems identified in the *accident recording system* are considered as the secondary problems, which affect the proper documentation of the data collected on the accident reported. Technically, the factors identified in this particular section contribute greatly to the unproductive analysis result at the provincial level. A clear description of the factors classified under the accident recording system is presented below with details.

- **Delay in accident information transfer** –this is attributed to the lack of consistent communication interface among the units involved in the processing of the road accident data (O' Day 1993). Inadequate communication contributed immensely to the inability of the authorised divisions to transfer compiled reports of the data collated in due time. This prevents thorough evaluation of the data collated.

- **Inadequate evaluation of completed ARF** –this is ascribed to the negligence attitudes exercised by the officers that execute the evaluation of the completed ARF (O' Day 1993).
- **Lack of commitment during information coding** –this is related to the backlogs encountered along the data processing stages. The first process stage of the data coding is performed at the local traffic departments by the DCOs, and the second stage is performed by Data pre-processing officer and Data entry officer in the provincial traffic departments. The actions perform by these two officers are based on the purpose of detecting and eliminating errors towards a quality road accident data (O' Day 1993). The two units raise the confidence of achieving a successful data analysis, by defining the possibility of accomplishing quality data through a rigorous assessment procedure. Under this section, brief description of the officers involved is provided below.
    - **Pre-processing unit** –this unit is required to carry out thorough examination of the reports transferred from the local traffic department, with the objective of determining any coding errors, and supply appropriate codes where necessary (O' Day 1993).
    - **Data entry unit** –this unit works further on the reports transferred from the pre-processing unit, and performs a cross-examination on the reports to determine if the coding is properly applied. Afterward, the officer in charge establishes a relationship between the new information and the existing information (O' Day 1993).

Earlier in this section, the probable problems affecting the processing of road accident data in the accident reporting and recording systems were concisely discussed, with details on the involvement of the officers that perform the operations within the units. Similarly, the problems affecting the analysis of the road accident data in the *analysis of accident data* cannot be left unmentioned. The type of problems affecting this specific unit are attributed to the sloppiness or operational ineffectiveness of the preceding tasks executed by the authorised officers in the Pre-processing and Data Entry Units. The performance of the officer in charge of the analysis of accident data can be measured through the degree of accuracy of the work output achieved in the first two units. Therefore, the outcome of the tasks executed by the Pre-processing and Data Entry Units determines the quality of the data to be analysed. However, this unit is challenged with technical problems elaborated below.

- **Incorrect application of codes** –this is actually attributed to the errors prior to the performances of the officers managing the coding of road accident data.

- **Lack of strong communication interface** –this kind of problem develops whenever there is interference in the operational relationship between the accident recording unit and the data analysis unit. The strength of the communication existence between these two units provides a better opportunity for continuity, in the process of stabilising the operational procedures established for data quality system.

- **Poor examination process of data inputting** –this problem frustrates the chance of achieving a smooth data analysis process. With this kind of problem, most times, there will be cases of mixing coding characters, which could hamper the ability of the data analyst from devising the easiest way of working through the data supplied from the preceding units.

- **Missing data** –this is one of the major problems preventing the chance of obtaining consistent road accident data. This problem is attributed to some other factors such as:

  - Poor application of the ARF,

  - Lack of commitment from the officers assigned to supervise the perfection of the data collated (O' Day 1993; Vogel & Bester 2004), and

  - Presence of impediment along the operational procedures implemented for data collation (Asia Injury Prevention Foundation et al. 2010b).

Moreover, in the *dissemination of accident data*, the service of the Analysis and Publication Unit is required on the procedures implemented for the dissemination of the analysis reports. This unit supervises the procedures carried out on the analysis of accident data. However, the performance of this particular unit depends on the Data Processing Unit. The communication effect between the units involved in the recording and analysis of the accident data influences the reliability of the analysis reports to be disseminated. Thus, the problems affecting the dissemination of analysis reports are highlighted below:

- **Incomplete analysis report** –this is as a result of the build-ups of anomalies along the process line, which could be linked with the failure to uncover errors that are undetected by the preceding units. Problems identified here are missing data, underreporting, data mix ups, poor of communication interface, and many more.

- **Delay in the release of the annual report** –this is ascribed to the sloppiness in the performance of the preceding units. This type of problem prevents the national level [National Department of Transport] from receiving all the provincial analysis reports as at due time, and thereby incapacitate the chance of publishing the annual reports with the right results at the scheduled period.

From another viewpoint, unreported accident cases like hit and run, could contribute to the problems frustrating complete compilation of accident data for annual publications. However, the inability of the police to obtain information regarding these cases can be enhanced through incident capturing device such as footage camera (Njord et al. 2005), where the party involved can be identified and interrogated if necessary. Other cases like incomplete accident investigations carried out by the police also lead to incomplete data collection, which may be due to adjournment of court hearings and lack of cooperation from the accident victim etc.

2.2.5 Road accident data handling units

Data handling can be described as the administrative procedures established to coordinate the processing of road accident data. The procedures are implemented through the establishment of the sequential data processing stages. Primarily, the stages are purposely implemented to transform road accident data from the unprocessed data to processed data before being disseminated to the public. In addition, the handling stages are illustrated in Figure 4 below[1], where the relationship between each stage determines the quality level of the data produced. These stages are designed to support the management of road accident data. The procedures followed in each stage of the data handling orchestrate the operations executed by each unit in determining the efficiency of each unit towards an active performance. Nonetheless, some of the units involved were partly mentioned in the preceding section.

2.2.5.1 Accident site

Accident site is the preliminary stage for data collection with the interactions between the accident victim or eyewitnesses and the police officer or traffic officer. As a result of this interaction, however, primary details like the location of the accident, time of the accident, date of the accident, and the driver involved in the accident are captured on the ARF.

2.2.5.2 Road accident data collection and investigation units

This unit comprises the police officer inspecting or reporting the incident, and the superior officer who queries any inaccuracies found in the data completed in the ARF. This unit carries out the preliminary data collection process and the investigation of the accident to support sufficient data supply in the road safety systems. In this unit, the two involving officers are discussed below in their order of functionalities in the handling of the reported RTA cases.

- **Police officer** –this officer stands as part of the road safety team in charge of the road accident data collection and investigation. A police officer is allocated by a superior

[1] The diagram illustrated in the Figure 4 is modified and adapted from James O' Day (1993).

officer to carryout thorough inspection on the accident reported. During this operation, the police officer is authorised to complete the ARF at the accident site (O' Day 1993; Sluis 2001). If there is any need for legal actions to be taken against one of the responsible parties, then the police officer will embark on a thorough investigation to substantiate the facts surrounding the accident case. After the completion of the ARF, then the police officer submits a copy of the report to a superior officer who will attest to the appropriateness of the data in the form.

- **Supervisory review unit** –the officer in charge of this particular unit performs his responsibility by supervising the completed ARF submitted by the police officer, with the aim of querying any inaccuracies in the completed form (O' Day 1993). Though, if any inaccuracies are discovered in the completed ARF, then the supervisory officer inquires the police officer who reports the incident to effect necessary corrections. On the contrary, if the data collected on the ARF is found accurate and complete, then a certified copy will be transferred to the Data Assessment Units for further assessment (O' Day 1993). The preliminary tasks are executed within the jurisdiction of the local level.

2.2.5.3 Data assessment units

The assessment of RTI is performed across the local level to the provincial level. In this unit, three separate divisions execute related tasks in the handling of the data collected. The interaction between these three units ensures quality data. Although, the three separate units are having one major purpose in common; which is the coding of the road accident data, either by analogue or digital means in the dataset. In addition, these units validate the quality level of the road accident data collated at the local level. More details are provided below on the tasks performed by each unit involved in the data assessment process.

Pre-processing unit –this is the first of the three units under the management of the Data Assessment Units. This unit performs a thorough examination process on the completed ARF received from the Supervisory Review Unit (O' Day 1993; Sluis 2001). The unit assists in the interpretation of gross errors, and also applies appropriate codes to where necessary; to lessen the workload to be executed at the next stages. Although, if errors are found, then the data pre-processing officer queries the ability of the preceding unit on the nature of errors discovered.

- **Data entry unit** –this is the second of the three units controlled by the Data assessment units. This unit functions as the intermediary between the Pre-processing Unit and the Data Processing Unit (O' Day 1993; Sluis 2001). In this particular unit, cross-examination is carried out on the codes allocated to the data prepared, or the coding structure applied by the Pre-processing Unit (O' Day 1993; Sluis 2001). This process is performed with utmost carefulness to ascertain the degree of consistency, and as well to determine the type of data to be processed in the subsequent stage before the data is entered into the dataset. Apparently, the communication effect between the Police Officer, Supervisory Review Unit, and the Pre-processing Unit determines the level of data sanitary in the assessment of the data processed in this unit.

- **Data processing unit** –the task execute in this unit is the continual assessment of the information [data] inputted or stored in the dataset at the Data Entry Unit, and as well determines the degree of reliability of the data assessed before reaching the Analysis and Publication Unit (O' Day 1993; Sluis 2001).

2.2.5.4 Decision making unit

This unit is considered as the superior of all units. The performance of this unit is measured based on the degree of consistency in the task completed by the Data Processing Unit. The unit includes the Analysis and Publications Unit, and the Public inclusive. The inclusion of the *Public* as part of the unit is to demonstrate the importance of the data users' feedbacks in the improvement of the data quality level towards a consistent output result. Most importantly, the opinions of the data users must always be valued by the data producers (Marshall & Rossman 1994; Wang & Strong 1996).

The data users continually perform post-analysis on the data, in order to ascertain its quality level in terms of the degree of completeness, consistency, and reliability of the data. Besides, the data users utilise the data for other purposes like researches, legal actions, innovations, budgeting and infrastructural developments etc.

On the contrary, the Analysis and Publications Unit, works further to compliment the task performed by the preceding units. The unit performs mostly three different tasks which are provided below as:

- Analysis of the data received from the Data Processing Unit.
- Publishes the analysis reports of the data evaluated; in order to disclose to the Public, the areas that require urgent improvements, and the ones that are currently undergoing improvements.
- Advocates for financial support from the government to support decision-making strategies (O' Day 1993; Njord et al. 2005); by presenting results obtained from the analysis performed before the Department of Transportation.

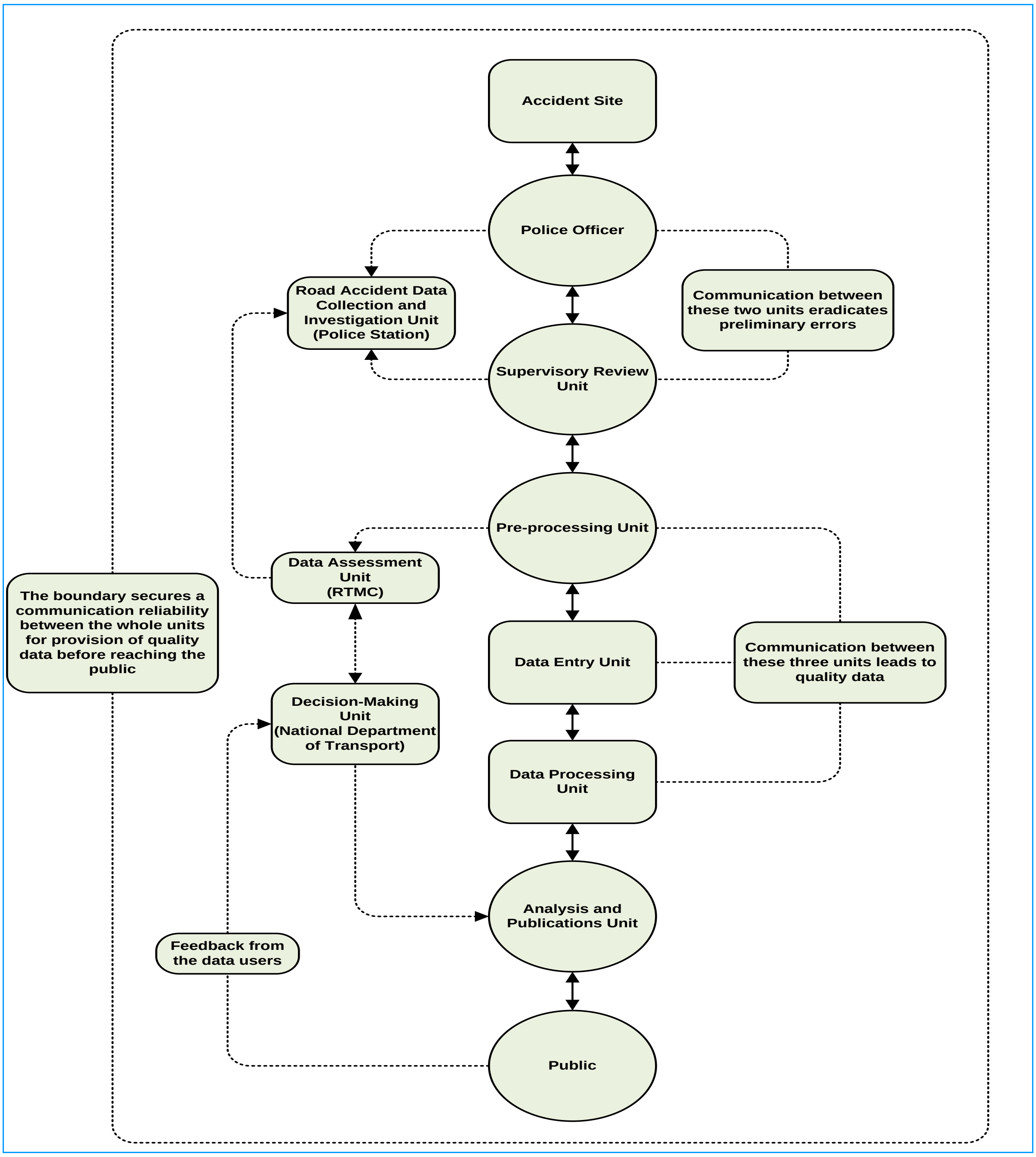

Figure 4: Process chart illustrating the communication interface among the data processing units

2.3 SOPs for road traffic information [RTI] management in South Africa

The SOPs for RTI management in South Africa are developed by RTMC, as approved in agreement with the Road Traffic Management Corporation Act [RTMCA], to guarantee a standard way of managing the RTI (RTMC 2012; RTMC 2014). The SOPs serve as the guidelines for all road traffic agencies or road traffic authorities involved in the acquisition of RTI, to facilitate a standard procedure for the production of quality road accident data. More so, the significance of achieving a consistent way of producing quality road accident data is fundamental to the RTMC, as the considerable agent to effective management of the road traffic system in South Africa. The key purpose of developing SOPs is discussed in the Appendix A-A.8.

2.4 Stellenbosch Traffic Department [STD]

The STD is a local division established to oversee the affairs of community matters relating to vehicle registration, learner's and driver's licence testing, payment of traffic fines and others such as law enforcement, card licence renewal, roadworthy testing, and public driving permits (Stellenbosch Traffic Department 2014). The department assists in restoring order to road traffic conditions within the municipality. The operations executed by this department are placed under the jurisdiction of the Stellenbosch Municipality [SM-WC 024]. The monthly reports of the activities executed are submitted to the municipal council for further evaluation, in order to ascertain the appropriate exploitation of resources within the municipality as authorised.

Furthermore, the operational relationship connecting the SM and the STD is displayed in Figure 5[2], wherein the units involved in the operations are shown in the organogram. Basically, the diagram illustrates the top officials in both the SM and the STD. The office of the municipal council is designated as the Director of Community and Protection Services, having the support of five managers under his/her supervision such as Manager Fire Services, Manager Fleet and Logistics, Manager Traffic Services, Manager Law Enforcement and Manager Community Services. The services of Financial Control Officer, Personal Assistant and Office Assistant [Messenger], are considered as the subordinate offices in the area of financial budgeting, advocating for financial improvement, resources distribution and circulation of memorandum within the department (Stellenbosch Traffic Department 2014; Stellenbosch Municipality 2012; Maree & Daniels 2007).

Generally, as depicted in Figure 5 below, the five managers operate in different areas, but they all have one relative operation, which is the protection of the community in the areas such as fire outbreak, road traffic circumstances, law enforcement, community services and vehicle fleeting. Furthermore, different offices are assigned under the direct supervision of the Traffic Services Manager, which is the key focus area in this investigation. The top office of the Traffic Services controls the office of the Head of Traffic Administration and Head of Traffic Law Enforcement. Actually, the services of the Support Assistants are always required by the top offices for constant circulation of memorandum, or conveying of information within the department when necessary.

---

[2] The diagram illustrated in Figure 5 is adapted from Stellenbosch Traffic Department (2014).

The operations carried out under the supervision of the Head of Traffic Administration are allocated to several units. Similar step is applicable to the office of the Head of Traffic Law Enforcement. Additionally, the office of the Head of Traffic Administration is responsible for vehicle testing, issuing of drivers' licence and vehicle registration. The office also has the service of court sections, which are responsible for the handling of the court cases pertaining to traffic offences, and a cashier section is established for the payment of traffic related offences. Although, the operations performed in the section for Head of Traffic Law Enforcement is classified under the supervision of the Superintendent of Training, Education & Accident Investigations and the Superintendent of Traffic Law Enforcement.

The Superintendent [Training, Education & Accident Investigations] controls other subordinate units such as the office of the Clerk and Control Room Officer [CRO]. Nonetheless, the CRO alert the standby officers on duty about any newly reported cases of traffic accidents, and convey feedback reports directly to the Superintendent [Training, Education & Accident Investigations]. On the other hand, the Superintendent [Traffic Law Enforcement] controls a large number of units designated as Assistant Superintendent, Traffic Wardens and others, such as Assistant Superintendent [Public Transport] and Assistant Superintendent [Speed & Electronic Enforcement]. These units, supervise the Traffic Officers and the Point Duty Officers under their jurisdictions except the Traffic Wardens, who works directly with the Superintendent [Traffic Law Enforcement]. The responsibilities discharged by all the units or divisions involved in the operations will be discussed in the subsection 2.4.1 below.

2.4.1 Responsibilities of the divisions in the Stellenbosch Traffic Department [STD]

In this section, the responsibilities discharged by each unit in the STD are discussed concisely with a clear understanding in the order of importance. Table 1[3] illustrates the services of the senior officers and their office assistance, whereas the Table 2 in the subsequent page shows the services discharged under the office of the Head of Traffic Administration, and similarly Table 3 displays the services discharged by each office placed under the jurisdiction of Head of Traffic Law Enforcement.

A full description of the responsibilities discharged by the divisions supervises by the Head of Traffic Administration is presented in the Table 2 below. Most of the responsibilities executed in these divisions are administrative tasks; implemented to facilitate the delivery of outstanding services to the users [public]. This division establishes a formidable relationship with the public,

[3] The contents in Table 1, Table 2 and Table 3 are adapted from Stellenbosch Traffic Department (2014).

Table 1: Responsibilities of the operational divisions of the STD

Stellenbosch Traffic Department	
Division	**Responsibilities**
Manager [Traffic Services]	▪ Coordinates the distribution of resources. ▪ Ensures the personnel effective management towards areas such as the traffic law enforcement services, driver's licence, vehicle registration, traffic fines etc. ▪ Compiles and implements the fire prevention plan and safety act. ▪ Manages the medium departmental construction activities.
Support Assistant	▪ Handles general administrative duties, and also files of information. ▪ Ensures uninterrupted circulation of the memorandum [information]. ▪ Arranges meetings, and writes meeting minutes of interdepartmental liaison.
Telephonist	▪ Answers and makes calls. ▪ Takes messages down. ▪ Directs calls to a particular office.
Office Assistant	▪ Cleans offices. ▪ Makes or arranges refreshments.
Gardener	▪ Assists the Office Assistant to arrange refreshments.
Head [Traffic Administration]	▪ Supervises the traffic administrative division. ▪ Communicates with the service providers towards achieving an effective and efficient service delivery. ▪ Plans budgeting and controls subordinates. ▪ Public liaison with the customers to ascertain public confidence in the service rendered.
Head [Traffic Law Enforcement]	▪ Supports the Chief Traffic Services in managing, coordinating policy implementation, control and supervision over the traffic law enforcement branch, including public transport, the accident bureau and the speed unit. ▪ Assists the Chief Traffic Services in planning and scheduling of operations, composition of an operating and capital budgeting for traffic law enforcement branch.

Table 2: Responsibilities assigned to the offices under the jurisdiction of Head of Traffic Administration

Head of Traffic Administration	
Division	**Responsibilities**
Management Representative for Vehicle Testing Centre	▪ Oversees the traffic administration of the vehicle testing station. ▪ Adherence to legal compliance.
Management Representative/Compliance Officer Driver's Licence	▪ Oversees the administrative work; by ensuring effective operation of the driver's licence departments. ▪ Performs administrative duties. ▪ Compiles administrative statistics in response to licence issuing. ▪ Ensures adherence to traffic law and policies.
Chief Clerk [Court Section]	▪ Oversees effective functioning of personnel at the court section. ▪ Manages administrative work regarding all traffic fines. ▪ Executes budget control for divisions. ▪ Performs public relation based on any vehicle related offences.
Chief Clerk [Motor Vehicle Registrations]	▪ Oversees administrative activities under his supervision and vehicle registration ▪ Compiles statistics based on PA 28 reports, and also ensures adherence to traffic law and policies.
Senior Examiner Vehicle Testing Station	▪ Tests vehicles for roadworthiness. ▪ Completes administrative work in response to the roadworthiness. ▪ Oversees tests done by examiners, and also ascertains the safety of the working environment.
Superintendent Driver's Licence Testing Centre	▪ Ensures the effective operation of the driver's licence testing centre. ▪ Oversees administrative activities and compilation of statistics pertaining to driver's licence testing.
Senior Clerk [Court Section]	▪ Monitors the collection of money and administration of court section. ▪ Oversees court case filing, and assists the Chief Clerk in training. ▪ Validates the appropriate execution of the tasks carry out by the administration.
Clerk [Court Section]	▪ Captures data and issuing of traffic fines. ▪ Files traffic violation cases. ▪ Executes public liaison and administrative support.
Cashier Bloemhof	▪ Receives and acknowledges payments.
Senior Clerk [Motor Vehicle Registration]	▪ Assists the Chief Clerk in motor vehicle registration on the day to day activities.
Clerk [Motor Vehicle Registration]	▪ Advocates for administrative support, public liaison, and assist compilation of statistics.
Examiner Vehicle Testing Station	▪ Assists the Senior Examiner Vehicle Testing Station in the testing of vehicles for road worthiness. ▪ Completes administrative task with regard to the road worthiness. ▪ Assists the Senior Examiner Vehicle Testing Station regards the safety of the working environment.
General Worker Vehicle Testing Station	▪ Semi-skilled labourers placed under training.
Driver's Licence Testing Officer	▪ Tests applicants according to set standards. ▪ Documents findings regarding the outcome of the training.
Clerk [Driver's Licence Testing Officer]	▪ Records data pertaining to booking of appointments. ▪ Receives payments, undertaking transactions, and filing documents. ▪ Attends to public queries, cash up and reconcile documents.

Table 3: Responsibilities assigned to the offices under the jurisdiction of the Head of Traffic Law Enforcement

Head of Traffic Law Enforcement	
Divisions	**Responsibilities**
Superintendent [Training, Education & Accident Investigations]	▪ Coordinates and manages division administrative activities. ▪ Oversees control room, and initiates public awareness concerning road safety. ▪ Oversees accident investigations; by liaising with the local police department to ascertain the level of completeness of the accident being investigated. ▪ Approves leave applications
Superintendent [Law Enforcement]	▪ Controls and supervises the personnel. ▪ Supervises the subordinates regularly to ensure proper execution of the tasks assigned to them. ▪ Coordinates emergency activities. ▪ Identifies danger areas along the roads. ▪ Carries out administrative duties. ▪ Controls the heavy traffic flow. ▪ Ensures public safety, roads inspection and enforcing the traffic laws.
CRO/DCO	▪ Carries out the duty of radio phoning activities. ▪ Communicates with the emergency services with regard to any reported accident. ▪ Operates the eNatis and TCS ▪ Executes administrative duties and establish public relation.
Clerk	▪ Administers the division's administrative tasks ▪ Ensures constant recordkeeping. ▪ Establishes solid public relation, and attends to the public queries.
Assistant Superintendent	▪ Controls and supervises the personnel under his jurisdiction. ▪ Identifies danger areas along the roads. ▪ Coordinates emergency activities. ▪ Executes administrative duties with regard to law enforcement activities. ▪ Regulates traffic density, enforces traffic law, and public safety inspection.
Traffic Wardens	▪ Stimulates the functioning of law enforcement and policing.
Assistant Superintendent [Public Transport]	▪ Supervises personnel, and also identifies hazardous areas on the roads. ▪ Coordinates emergency activities, administrative duties, and traffic control. ▪ Guarantees public safety and enforcing the traffic law.
Assistant Superintendent [Speed & Electrical Enforcement]	▪ Coordinates the division and the personnel. ▪ Oversees the Speed Unit back office. ▪ Coordinates Special Action and Alcohol monitoring. ▪ Oversees administrative tasks at the divisions, and public relation.
Traffic Officers	▪ Ensures the functioning of the law enforcement and policing. ▪ Improves the relationship between the law-enforcers and taxi-operators by establishing an effective land transport system, by the way, sensitise their participation through an integrated planning regulation and provision of a high standard service delivery. ▪ Renders the functioning of the administration within the Speed Unit. ▪ Performs public relations and adjudication based on the traffic offence committed by the road users such as speed enforcement.
Point Duty Officers	▪ Executes the traffic control duties during heavy flow of traffic.

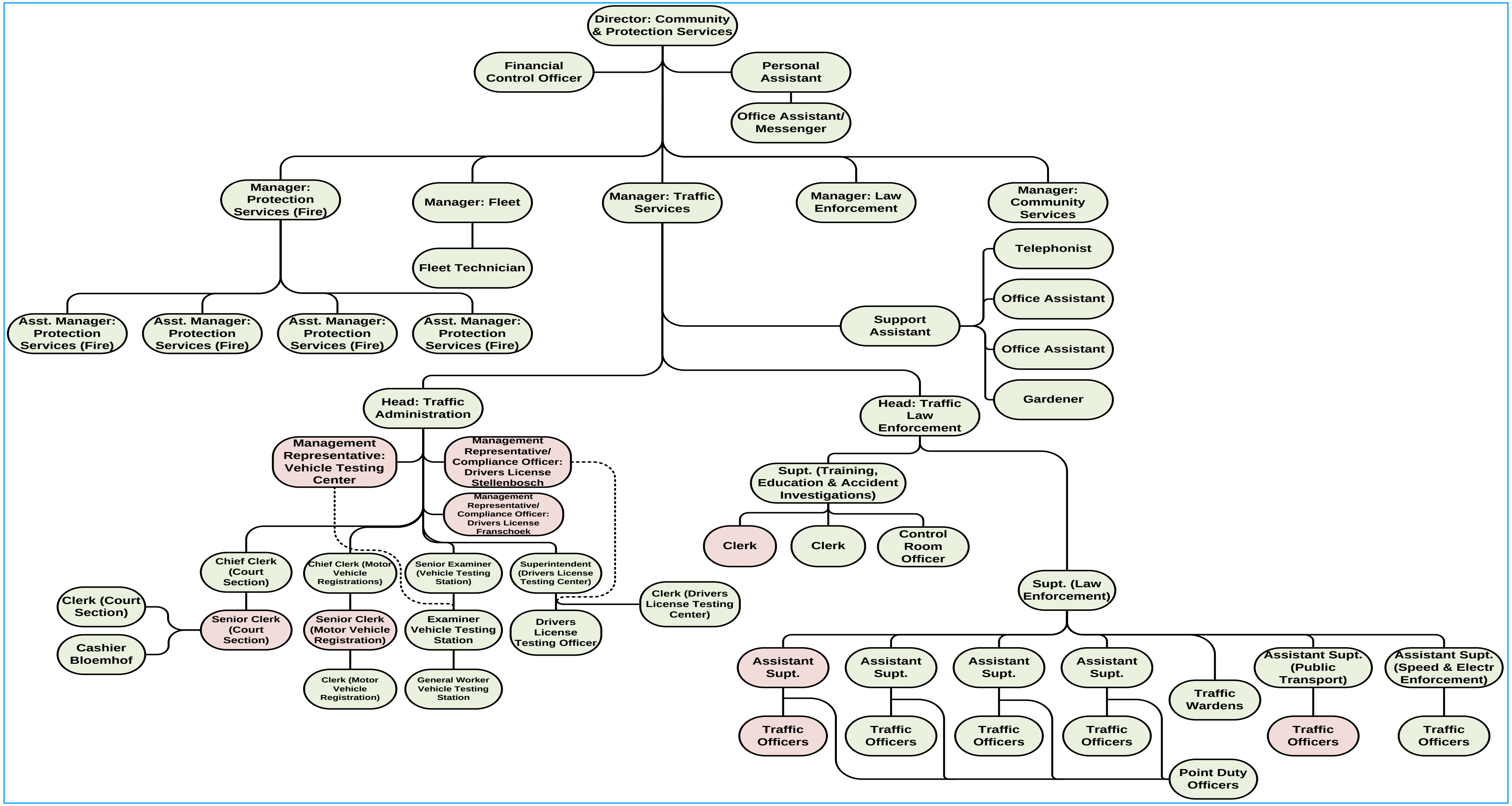

Figure 5: Organogram of the Stellenbosch Municipality [SM] connecting to the Stellenbosch Traffic Department [STD]

to encourage relevant feedbacks on the quality of services delivered. Correspondingly, the Table 3 illustrates the responsibilities discharged in the divisions under the management of the Head of Traffic Law Enforcement, where the research is currently being conducted.

Moreover, the key activities performed under the Head of Traffic Law Enforcement is basically for the purpose of minimising accidents on the roads. This purpose is achieved by ensuring constant inspection of the traffic flow, initiating safety programmes and enforcing the traffic laws, in agreement with the approved traffic regulations established by RTMC as enacted through RTMCA. For more details, refer to Table 3.

2.5 Data quality

Around the globe today, many organisations are yearning for a reliable means of improving the process system or SOP established for the acquisition of quality data, regardless of the type of operations executed within these organisations. Furthermore, pertaining to the RTA, an improved data quality system is considered as a leading remedy towards the possibility of minimising the occurrence of accidents on the roads in South Africa. Data quality can also be considered as the lead to data accuracy, wherein every of its components play important role in the assessment procedures designed for the data gathered (O' Day 1993; Vogel & Bester 2004).

The reliability of road accident data depends on the degree of consistency in acquiring relevant information at a sufficient level. Data quality discloses the importance of key factors to the data users as presented in the ARF. The need for the improvement of quality of road accident data paves the way for the standard development of relevant countermeasures required to establish reliable safety measures. According to Suzette Thieman (1999), as stated in a conference paper, “safety is a driving factor in access management, and the accident reports are the best indicator of the lack of safe roads” (Thieman 1999).

Furthermore, to affirm the statement declared above, James O' Day offered an insight into the meaning of quality as a compliment to a reliable data, and thereby describing quality as another word for accuracy, precision, timeliness, and completeness of data used to address the problems as important to the solutions (O' Day 1993). On the contrary, Richard Y. Wang & Diane M. Strong (1996; 2013) defined data quality as data that is fit for use by data users. Data quality is also defined in terms of its dimension as a set of quality attributes that represent a single concept or model of data quality (Wang & Strong 1996). James O' Day (1993) further placed emphasises on the significance of the road traffic quality data to several organisations such as the academic institutes, legal institutes, and traffic bodies in the areas that are paramount to their type of operations such as data collection, data analysis, data evaluation, and decision making processes, which is determined by the relational interface [communication] between the stages involved in the processing of data (O' Day 1993) as discussed in subsection 2.2.5.

Notwithstanding, many organisations based the outcome of their decisions on the quality level of the data processed. Due to this, it is imperative for these organisations, agencies, and departments to coordinate a process that could downsize the cost of poor or low quality data (Wang 2004). Moreover, Richard Wang (2004) advised data producers on the need to improve data quality, by treating data quality as a multidimensional concept beyond accuracy. In this concept, not only the accuracy should be considered as a major component required for the improvement of data quality. Other components such as completeness, consistency, timeliness should be considered significant in developing a standard data quality process.

Similarly, a statement declared by Sinclair (2011), offered a clear understanding on the reason not to consider accuracy as the only major element suitable for the improvement of data quality as regards the adjustment of database issues. Thus, the statement states that, "data quality is ultimately a consequence of data input accuracy, completeness, and consistency, and how relevant these are in addressing the objectives of the database" (Sinclair 2011).

Practically, by considering the above statement, a new data handling system can be developed, or perhaps by improving the existing system to support the appropriate handling of road accident data in the transportation system, which is considered more cost effective than developing a new system. According to Richard Wang (2004), the standard methods required to evaluate data quality have been accepted globally, along with the relevant tools needed to execute the evaluation process. The applications of these methods require an in-depth understanding of the operational procedures of the organisation involved. However, to buttress the above statement, Richard Wang (2004) declared that "understanding of the systems, processes, and management practices of an organisation has become as important as understanding its data" (Wang 2004).

In another article published, Richard Y. Wang et al (1995) stressed that "for organisations to be best served by their information systems, a high degree of data quality is required, and the need to ensure this quality has been addressed by researchers and practitioners for some time" (Richard Y. Wang et al. 1995). The statement illustrated that several methods have been developed by the experts in this domain, which many organisations have adopted to resolve the issues of poor data quality. These methods are applied by many enterprises but not at all corporate environments. Several enterprises have no idea of how to manage, reform or restructure the procedural systems used to acquire data within their organisation.

2.5.1 Interrelationship between data producers and data users

In spite of the rapid growth in technology since the past years, it is unfortunate that shortcomings mutilate the procedural system developed for the handling of the road accident data in South Africa. Many traffic accident statistics released annually to the public by the RTMC in South Africa were considered to be not error free by the data users. Conversely, from previous studies, few out of many datasets released to the public are found to be error free (Agran & Dunkle 1984; Ferrante et al. 1993; Rosman & Knuiman 1994; Rosman 2001; Baguley 2001; Vogel & Bester 2004; Sinclair 2011). However, this calls for overhauling or redesigning of the system handling the processing of the data for efficiency and effectiveness of a reliable result, depending on the capacity of the organisation involved.

In the third paragraph of section 0 above, data quality was described by Richard Y. Wang & Diane M. Strong (1996; 2013) as the type of data that is suitable and ready to be utilised for a designated purpose by the data users. This signifies that data users have a big role to play in ascertaining the quality level released to the public. Richard Y. Wang & Diane M. Strong (1996), suggested that a bond should be established between the data producers and data users (Wang & Strong 1996), so as to promote the steadiness of the quality of the data gathered through the means of feedbacks acquiring from the data users. This bond should be established on a close loop system, wherein the circulation of feedback will be held as a priority while disseminating information to the public as illustrated in the Figure 6[4].

The data users are in the best position to achieve a better end view on any of the data released to the public by the data producers. Richard Y. Wang and Diane M. Strong (1996) gave simple reason to support this by saying that "customer viewpoint of quality always judge whether or not a product is fit for use" (Wang & Strong 1996). Figure 6 shows that numerous data users *[data consumers]* are depending on the quality of RTI disseminated quarterly and annually by the data producer. The concern of the data users in accessing reliable and consistent end results of the RTI produced by the data producers, could facilitate the chance of improving the process applied in producing the road accident data through sufficient feedbacks. The feedback region serves as the interface where a communication medium is established between the data producer and data users as regards the South Africa context. Hence, the improvement of the data process can be built around the communication interface within the processing units and outside the processing units, based on the effective establishment of the feedback circulation on closed loop system as described earlier.

---

[4] The diagram displayed in Figure 6 is adapted from James O' Day (1993).

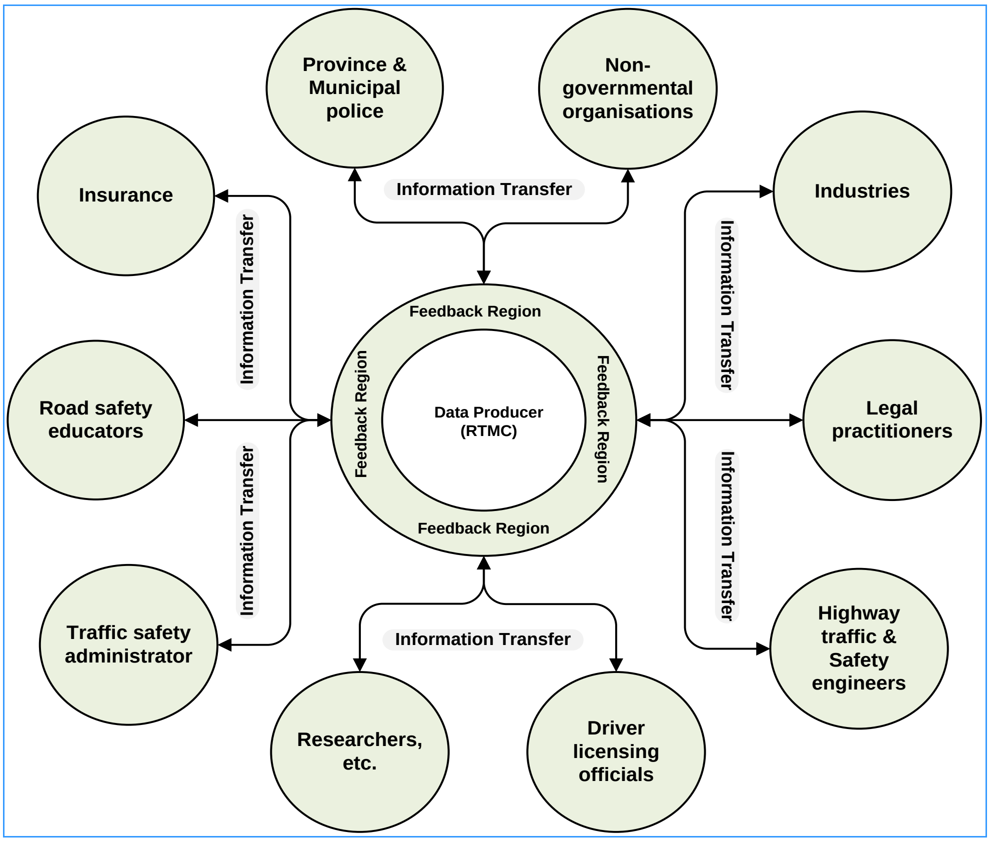

Figure 6: Illustrating the interface between the data producers and data users

2.5.2 Importance of data quality in road transportation system

In this section, the importance of data quality is based on the decision-making aspect regarding road transportation system only. This is much related to the significance of data collection in subsection 2.2.1. Data quality is the end result of the data collected after it has passed through rigorous analysis procedures. In other words, the end result of the data processed serves as the right tool for decision-making process, though, it depends on the degree of standard of the procedures applied. Quality data could be recognised as the foundation of decision-making in every aspect of any organisation. Data is considered as the reference point to many institutions' or countries' progresses when it comes to the understanding of the matters concerning the visibility of the institutions' or countries' improvement over the year. In essence, the quality level of a country's data determines her level of improvement in many areas such as population increment, economic stability, infrastructure, distribution of resources and gross development product; whereas in the aspect of road transportation system, it is used in many decision-making areas such as operation decisions, policy-making decisions, accident claim decisions, and transportation safety decisions etc. Some decision-making areas are described below according to their area of importance in the South African transportation system.

- **Decision making at the national level;** for policies and initiatives as applied by the National Transport Master Plan [Natmap], Public Transport Strategy [PTS] and Electronic National Traffic Information System [eNaTIS] (van Niekerk 2012).
- **Legislative decisions;** for enforcing the law that bound drink-driving, use of drugs as applied by National Road Traffic Act [NRTA] (Act No. 93 of 1996) and CPA, through the implementation of the Alcohol Breathalyser [AB] for the measuring of Blood Alcohol Concentration [BAC] (van Niekerk 2012).
- **Operational decisions;** for implementing a management system that preserves road infrastructures, improves road safety and increase productivity as applied by Road Traffic Management System [RTMS] (van Niekerk 2012).
- **Design and policy decisions;** for design, construction, management and maintenance of South Africa's national road network as applied by SANRAL (SANRAL 2008; SANRAL 2012).
- **Special study of incident;** for collection, processing and analyses of data before making the report available to the public as applied by RTMC (Vanderschuren & Jobanputra 2011).
- **Establishment of data capturing devices;** for the implementation of an electronic device to reduce car theft and cross-border crimes between Namibia and South Africa as applied by Cross-Border Road Transport Agency [CBRTA] (van Niekerk 2012).
- **Decision making on accident claims;** for confirming that a person is actually involved in accident during and after he is seeking for accident fund as applied by RAF (van Niekerk 2012).

2.5.3 Data quality problems

According to Vogel & Bester (2004), data quality is considered as the major solution to the reduction of accident occurrence (Vogel & Bester 2004). However, the problem affecting data quality in the RTA started from the information collation system [involving the local authorities] to the information processing system [involving the provincial authorities] (O' Day 1993; Baguley 2001; Njord et al. 2005). The data management sloppiness also contribute to the problems frustrating the acquisition of quality data such as inconsistency in information coverage, unclear interpretation of data, missing data elements, faulty entry procedures, response errors and locations representation (Vogel & Bester 2004; Sinclair 2011).

Besides, the quality of road accident data depends on the relational interface existing between the officers assigned or authorised to handle the reporting of incident, collection of information and recording of the information gathered. Another problem is the absent or poor link between the organisations involved in road accident data collection such as the hospital, insurance etc (WHO 2009; WHO 2011; Sinclair 2011). Consequently, the determination of the RTMC in achieving an error free road accident data, must be directed at the actions performed by the accident reporting officers, who gathered the unprocessed data [raw data] right from the sources. In addition, the insubstantial performance of the officers in charge of data processing contributed to the problems frustrating the objectives of the RTMC.

A statement by Wang et al (1995; 1996) affirmed that inaccuracy, inadequate data, missing information can have huge impacts both socially and economically on the status of a nation or an organisation (Richard Y Wang et al. 1995; Wang & Strong 1996). Futhermore, better undertsanding on how significant data quality is to any organisations was described by comparing its significance to difficulty in managing product quality. However, Wang et al (1995) consolidates the above statement by stating that, “while it is diffcult to manage product quality without understanding the attributes of the product which define its quality, it is also as difficult to manage data quality without understanding the characteristics that define data quality” (Richard Y Wang et al. 1995). This is applicable to road accident data quality system where the components that constitute the data quality system have to be understood clearly before quality of road accident data can be improved. This process defines the importance of understanding the operational responsibilities of the officers in charge of the system implemented for data acquisition in the STD.

Richard Y. Wang & Diane M. Strong (1996) further added that large numbers of database are marred with large numbers of errors, and some were perceived error minimal which is still considered a huge blow to the quality of data to be mined. To apply the practicality of the above statement to the RTA database in South Africa, however, Marion Sinclair (2011) suggested that the RTA database structure required reassessment to accommodate more data fields or information, as a means of reducing fidgety attitudes of the officers in charge of data entering to avoid more errors. Even so, organisations such as manufacturing companies, banks, retail companies are also encountering problems of poor data quality along the process of data entering (Wang & Strong 1996) .

An illustration was given by Richard Y. Wang & Diane M. Strong (Wang & Strong 1996) to support this statement, “a big New York bank found that data in its credit-risk management database were only 60.0% complete, necessitating double-checking by anyone using it, and more so, a major manufacturing company found that it could not access all sales data for a single customer because many different customer numbers were assigned to represent the same customer” (Wang & Strong 1996). In essence; the illustration given above, insinuates that in the process of achieving quality data, many factors relating to the cause of irregularities along the process should be double checked to determine certainty in the procedures applied. This could lessen the accumulation of errors committed along the production line designed for data acquisition.

Fundamentally, data quality enhancement should be based on the significance of the quality data to the data users (Wang & Strong 1996) as discussed in subsection 2.5.1 above. In addition, James O' Day (1993) advised users of road accident data on the limitations of the information which they are working with. Further explanation was laid on the expectation of uncertainty during data analysis, where the data is expected not to be a complete one or error free. Criticism was emphasised on the minimal involvement of many countries or states in performing any assessment as regards the completeness and accuracy of their road accident data (O’ Day 1993).

James O’ Day (1993) advised the government or the national transportation agencies on their position to establish a group or another department that should be allocated to oversee, crosscheck and supervise the quality level of road accident data (O’ Day 1993). This group or department will carry out further estimation of the road accident data, with the purpose of discarding any presence of inconsistency, starting from the reporting phase of the incident to the data entering phase.

2.5.4 Data quality components

Road accident data quality can be improved by considering the components that constitute quality data as relatively mentioned in section 0 and subsection 2.5.3 respectively. These components are used as directives towards a better way of attaining procedural steps for quality data assessment. They enhance the possibility of identifying the productive way of acquiring quality data in the process of assembling data, whereby paving the way for the possibility of identifying the particular areas that are problematic towards the acquisition of data.

According to statement made by James O’ Day (1993) in a science magazine, “from a quality point of view, accuracy and completeness are probably more important than timeliness, but excessive processing time should be avoided” (O’ Day 1993). On the contrary, from my personal viewpoint, accuracy and completeness as quality characteristics, should be built on timeliness. Simply because the expected outcome from the implementation of these two quality characteristics along the evaluation protocols depends on the time taken to process the data. This illustration could be used as a yardstick for determining the competence of the reporting officers. This could improve the rate at which the completed ARFs get to the DCO for further evaluations. Although, the productivity rate of an information system can be measured by the degree of quality of data produced and the satisfactory comments of the data users (Wang &

Strong 1996). In addition, an illustrative flowchart showing the relational connectivity between the quality components is shown in Figure 7[5].

- **Completeness of reportage** –this is the primary phase, wherein the evaluation process is ensued to ascertain the relevance of the data collected regarding all specified cases through the data collection method established (O' Day 1993; Vogel & Bester 2004; Sinclair 2011). Considering this aspect of quality component, the data supervisory officers, in charge of inspecting the completed ARF queries the experience of reporting officer regarding the application of the ARF, provided that errors are discovered.

- **Consistency of reportage** –this phase complements the evaluation performs in the first phase. The action performs here is basically to identify the factors that affect the degree of reportage (O' Day 1993; Sinclair 2011), which could be varied by jurisdiction, time, personal characteristics, weather, and other unspecified factors (O' Day 1993), to ascertain the flexibility and reliability of the database (Vogel & Bester 2004).

- **Missing data** –the activities perform in this phase strengthen the assurance of knowing the actuality of the problems or factors frustrating the consistency of accident reporting system, by identifying the probable causes of missing data attributing to the negligence of the officers executing the assessment of completeness of the reportage (O' Day 1993; Vogel & Bester 2004). Missing data can be described as unavailable data for analysis, which is classified important to the real-world scenarios (Batini et al. 2009). According to article published by Paolo Atzeni & Valeria De Antonellis in (1993), as cited by Batini et al (2009), "a data or value can be missing either because it exists, but it is not known, or because it does not exist, or because it is not known whether it exists" (Batini et al. 2009).

- **Consistency of interpretation** –this is principally one of the essential parts as regards application of the quality components. This certify the correlation of accident reportage across the authorised locals and provinces (Vogel & Bester 2004; O' Day 1993), to ascertain whether the accident coverage is reported in a similar way, across the localities or regions by different reporting officers regarding the basics information required as instructed in the ARF (O' Day 1993).

- **Right data** –this actually paves the way for better understanding of the analysis of the data, if only the right data are provided for actual purpose used for (Vogel & Bester

[5] The contents in the diagram illustrated in the Figure 7 are adapted from James O' Day (1993), Liesel Vogel & Christo Johannes Bester (2004), and Marion Sinclair (2011).

2004). An illustration is provided by James O' Day (1993) to support the understanding of this particular data quality component, which states as thus "for the road accident data to be most useful for Safety Authorities, vehicle characteristics is important; for the Highway Authorities, roadway details are important; for the Psychologists, personal details of the driver are important" (O' Day 1993).

- **Appropriate degree of detail** –this is useful in the area of broad analysis, wherein some variables or data elements need to be more detailed, so as to achieve an improved compilation and dissemination of information (Vogel & Bester 2004). Actually, this is dependent on the variables and the questions inquired. For instance; for *Weather Conditions and Visibility,* variables such as Clear, Overcast, Rain, Mist/Fog, Hail/Snow, Dust and others like Fire/Smoke, Severe Wind, and Unknown are relevant (O' Day 1993).

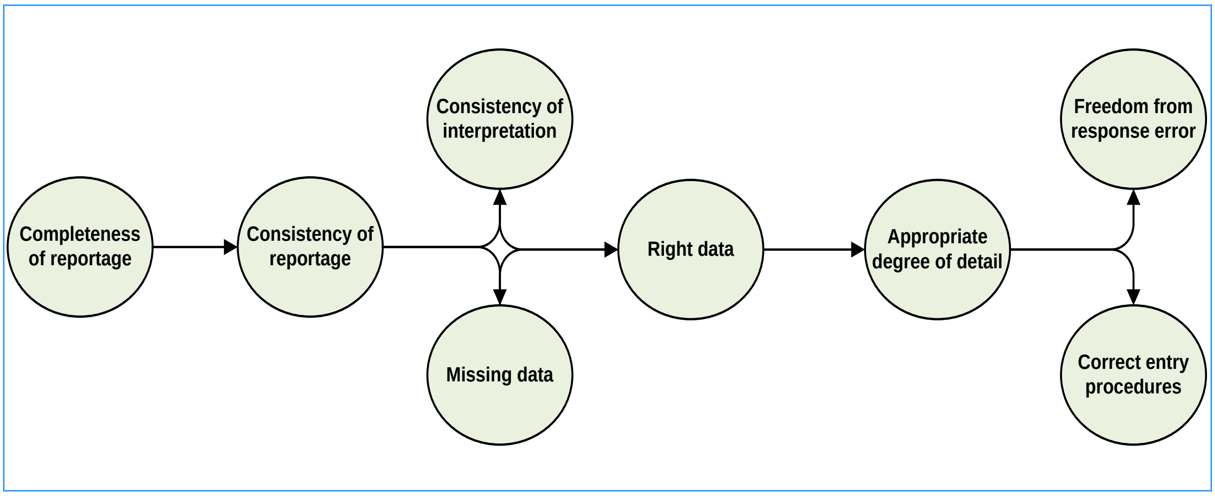

Figure 7: The relational connection between the quality components

- **Correct entry procedures** –this aspect scrutinises the procedure followed in the preceding phases. Here, further thorough assessment on the information provided in the ARF is executed, in order to ascertain the degree of quality of the information [data] inputting into the data set (O' Day 1993; Vogel & Bester 2004; Sinclair 2011). The information [data] received here is treated with utmost care to evaluate the tasks performed by the preceding phases. However, after evaluation process has been completed, then the officers in charge will embark on assigning appropriate codes to each relevant field where necessary.

- **Freedom from response error** –this aspect is basically to ascertain the degree of accuracy and completeness in reporting accidents, wherein RTI is observed to be incorrectly represented due to human errors which involved the physical measures of such accident scene, description of the location, including the entering process of the information into the road accident database (O' Day 1993; Vogel & Bester 2004). In addition, James O' Day (1993) offered a simple explanation of what '*Freedom from response error*' means, by stating that "when something was measured, was it measured correctly?" (O' Day 1993). An illustration that elucidate the inaccuracy level of the officers [human] reporting the incident.

However, it is important to control the quality of data by manual and automatic edit checks at the Correct Entry Procedures phase (O' Day 1993). Ultimately, analyses that identify shortcomings in the data gathered should be reported back to the data collectors, and the data entry officers so that the system will improve (O' Day 1993). The aforementioned quality components for data processing determine how accurate any data gathered can be, if processed in conformity to the practicality of the quality components. In addition, James O' Day (1993) suggested that the application of these quality components should be carried out manually to ascertain easy unearthing of shortcomings, and applied the corrections accordingly through automatic means if necessary.

2.5.5 Effect of underreporting in South Africa

The issue of underreporting in the road traffic accidents [RTAs] is rampant in most regions around the globe. However, many methods have been devised or developed to address this issue (Derriks & Mak 2007; Yamamoto et al. 2008). These methods were established for the improvement of the road accident reporting and road accident database system in both the developed, developing, and underdeveloped countries, by making the design simpler, comprehensive and robust for information capturing on collision incidents as mentioned in subsection 2.2.2. The collision report forms in the developed countries are substantially productive, sophisticated, understanding and simpler. The information structure of these report forms is preferable to the approved ones used in many developing and underdeveloped countries, since the information gathered or captured with the application of these forms are sufficient and reliable. This is due to the necessary measures taken or integrated into the procedural systems applied to monitor any occurrence of RTA.

Although, in many African countries, the availability of good quality of road accident data is a mirage, only South Africa can boast of regular data supply as cited in subsection 2.1.3 above. This ranks South Africa above most other African countries as having a relatively good safety system. This substantiates that the RTMC executes the objectives of improving the quality of data gathered from the primary source. Also, among African countries, South Africa, Zimbabwe, and Botswana are ahead in the application of better safety systems (Jacobs & Aeron-Thomas 2000), but the chance of generating sufficient and right data is often a difficult issue in these countries.

Among these three countries, South Africa is the only African country that applies the use of a GIS based road accident data system, while the other two countries based their safety approach on monitoring of highway accident locations (Jacobs & Aeron-Thomas 2000). This is implemented in order to have more data coverage in the road safety system. More so, South Africa is conceived as the only African country that has long established and well-respected research organisations with decades of experience in road safety research, while other countries appear unable to maintain a road safety research programme. These advanced research institutes placed South Africa ahead of other developing and underdeveloped countries in Africa.

Despite the achievement attained by South Africa in considering road safety as a priority, cases of underreporting of RTA by the police or other reporting officers are condemned by the public due to the unresolved annual rise in death tolls on the South African public roads. The South Africa collision recording system is criticised regularly for poor quality data production. An online journalist, Ray Joseph (2013) published an online report on the aim to uncover the worst South African drivers, by comparing two major cities [Cape Town and Johannesburg] in South Africa. The investigation was based on the findings obtained from the RTMC and insurance companies, based on the quality of the available data on accident claims and the average numbers of these claims. And also, the Automobile Association of South Africa [AASA] was considered for data sourcing.

In the process of the investigation, Ray Joseph affirmed that computer errors rendered some data unusable at the RTMC. To buttress the above statement, Ray Joseph (2013) further stated that collection and distribution of road accident data as administered by the RTMC is in disorder. This discredits the integrity of the available road accident data (Joseph 2013). This statement illustrates the extent of problems in the quality of data produced by RTMC. Though Ray Joseph is not an expert in the field, he carried out interview sessions with experts in this field to justify his findings. In addition, to support the criticism made by Ray Joseph regarding the status quo of the data processed by RTMC; Marion Sinclair (2011) stressed that “most estimates extracted from the Road Collision Database are basically on incomplete data” (Sinclair 2011).

Further affirmation was emphasised on the standard of the Road Collision Database practically used in the developing and underdeveloped countries in the same publication. Moreover, Sinclair (2011) confirmed the above statement by asserting that “it is obvious that most annual estimates gathered by the RTMC, or by traffic engineers, or safety experts, and road authorities in both developing and underdeveloped countries were considered little or no availability of data, compared to what produced or obtained in the developed countries, which is regarded as the most complete datasets” (Sinclair 2011).

Actually, the scope of the article published by Marion Sinclair (2011) only examines the overall structure and broad content of the Road Collision Database, with critical examination of what the database does accurately, and identify parts that accommodate errors. Some other investigators levelled complaints against RTMC as regards the road accident data produced annually across the entire nation. However, Media 24 investigations identified computer errors as a major problem frustrating the integrity of RTMC in providing reliable road accident data, which has rendered almost 20 years of national road accident data unusable (Joseph 2013), under the supervision of the RTMC as government's custodian of road accident data.

The World Health Organisation in 2002 reported that African Regions have the highest value of 20.8 per 100,000 population in Road Traffic Fatality compared to other regions of the world as cited by Marion Sinclair (2011). It was declared in the same article that 35 of 110 countries that participated in the exercise have an issue of incomplete data, most especially, middle-income countries and lower-income countries. The most complete datasets were provided by high-income countries which presented the lowest fatality rates, whereas little or no data were available from the poorest regions or countries in the world, which are perceived to have the highest fatality rates (Sinclair 2011). The database system designed for data custody in the developed countries are considered the best in terms of information robustness or sufficient data dissemination and accuracy, which energises their effort towards dramatic reduction rate in accident fatalities in their various regions.

Due to the above statement, one would ask a question that, *'why is the Road Collision Database of high-income countries far more productive than that of the middle-income and low-income countries?'* This question can be answered in terms of the outcome of the safety measures devised or developed to reduce road accidents occurrence on the roads in the high-income regions. This implies that the road safety systems cultivated in the high-income countries [developed countries] are more sophisticated than the road safety systems adopted in the middle-income [developing countries] and low-income countries [underdeveloped countries]. But the effect of the safety measures can only be assessed by the degree of efficiency and reliability of a complete database system or dataset. This is difficult to achieve in most middle-income and low-income countries, due to the substandard and/or obsolete procedural systems used in most of the regions attributable to the level of technological advancement. Data is found missing in middle-income and low-income countries due to cases of non-availability of data, level of incompetency, irrelevant information and human errors (Sinclair 2011).

However, further criticism was made by some individuals on the deteriorating level of data captured by the police, and the undermined process carried out by the RTMC. Sinclair (2011) mentioned in a conference article that the users of the road accident data have no confidence in the data disseminated quarterly and annually by the RTMC (Sinclair 2011). Actually, one of the criticisms was based on the recording protocol used in the classification of accident victim reported dead some hours/days after his involvement in an accident. With a clear clarification, in the developed countries like United States and France, anyone who dies within 30 days of a vehicle accident, through a regular follow up is considered as *fatality,* whereas the approach is different in South Africa context, only persons who die within six days after the main incident, are approved as accident fatality (RTMC 2007; Amoros et al. 2009; Joseph 2013). This policy has to be reviewed by the traffic management in order to improve the degree of adequacy of relevant data in the RTA. In South Africa, such casualty records of persons who died within 30 days or after six days can only be acquired from the hospital death records attributable to the motor-vehicle accidents.

Officially, management policies have a vast impact in the demerit of the National Road Collision Database, because these policies rendered some road traffic casualty cases impractical, just like the scenario discussed in the above paragraph. The due date for the confirmation of road traffic mortality should be extended beyond 6 days in South Africa, because there is always a possibility of the accident victim not surviving the impact of the accident after the specified number of days, and if approved, the insurance will benefit from the development and likewise the families of the accident victim. And most importantly, there will be more chances or opportunity of acquiring more data pertaining to the accident victim's details and other relevant information.

In addition, an expert in the field, Derek Luyt was interviewed by Ray Joseph (2013) with regards to the police improper documentation concerning regular follow up of deaths or road traffic mortality cases beyond the actual date of the incident. This is continually considered as a problematic issue due to strict policies implemented by the management of the road traffic issues (Joseph 2013). With further clarifications, concerning underreporting issues in South Africa, it is hard to base a fact on the level of the data obtained from the RTMC, just that the origin of the available data are questionable in many aspects, since the data show no clear representation (Joseph 2013). Moreover, Derek Luyt and Howard Demboysky mutual statements criticised that, "data available on RTA and fatalities in South Africa gives no coherent picture of what is happening on the South Africa's roads" (Joseph 2013). The two experts, as affirmed by Ray Joseph (2013), further categorised the data administered by the RTMC as immaterialised towards evaluation planning and decision-making processes established for the reduction of road fatalities in South Africa, which is faulted due to the degree of quality of the data gathered and processed.

From another perspective, two other researchers published a research paper on the road safety in Africa, majorly concentrated on Sub-Saharan African Region. The objective of the paper is to examine the existing data and information obtained on the road safety situation; thereby shortcomings in the data gathered could be identified and, priority needed to set ahead improvement towards a qualitative data system would be developed. Generally speaking, from their findings, it was discovered that sufficient awareness and interests were demonstrated towards crash incidents than non-crash incidents (Jacobs & Aeron-Thomas 2000). This is classified as one of the problems of underreporting facing both the middle-income and low-income countries.

Although the non-crash incidents are useful in the predetermination of traffic accidents, that is, the prone areas to accident occurrence and the possible causes. Hospitals and standby ambulances are known for keeping records of non-crash incidents, little or few numbers of non-crash incidents are recorded by the police, but affirmatively, hundreds of such cases are mostly recorded by the traffic department. Jacob and Aeron-Thomas (2000) further declared that only few countries among African countries with computerised or digitalised road accident data systems maximally utilise the data. This is categorically considered as a huge contribution to the worst state of road safety system in Africa compared to the rest of the world.

Actually, it is practically observable that many shortcomings are associated with the frustrations affecting the SOPs of the road accident data in South Africa. Recall from subsection 2.2.4 above, clear understanding was provided about the probable problems that may arise along each stage of data processing, and also the responsibilities of each unit involved were discussed in details. However, a summarised number of the problems discovered through the recent literature studies are listed below as thus:

- **Inability to obtain a complete and reliable road accident data in South Africa;** attributable to inadequate or irrelevant information concerning both nonfatal accidents and fatal accidents (Sinclair 2011).
- **Unreliable data collection and dissemination processes;** use of incorrect codes during data inputting, corrupt data hardware while migrating to a new database and lack of frequent back up by data processing team (Joseph 2013).
- **Lack of direct access to available data;** due to political influence and strict management policies (O' Day 1993; Njord et al. 2005).
- **Short term registering of accident fatalities in South Africa;** 6-day counting of accident fatality is another means of promoting the influence of underreporting in the quality of collision data (Amoros et al. 2009; Joseph 2013).

In short, the problems highlighted above are mainly connected to two major areas in the processing of road accident data. Thus, the two major areas are classified as; firstly, the inability of the police to deliver accurate reports from the primary source; and secondly, the inability of the data processing officers to disseminate error free data to the public. These two cases have cost the management the chance of producing a complete and reliable data for the convenience use of the public.

2.5.6 Concise synopsis of the data quality problems

The essentiality of improving the quality level of road accident data gathered at the local level, processed at the provincial level, and disseminated at the national level in South Africa is primacy to the RTMC, the data collection agencies, and the data users; who are the ultimate beneficiaries of the data. Despite this impulse, limited attention is devoted towards the handling of the road accident data gathered, which poses a huge threat to the sustainability of the transportation system in South Africa (Jungu-Omara & Vanderschuren 2006) as discussed in the Chapter 2 of this study.

To buttress the above statement, Stats SA (2009) emphasised in an annual report published in 2009 that, “without a reliable information, priorities for road traffic injury prevention, the impact and outcomes of such interventions cannot be rationally or satisfactorily determined” (Stats SA 2009). This statement reflects the significance of quality data in determining the degree of relevance of any supporting decisions established.

Apparently, road accident data is the base reference of ascertaining the weight or magnitude at which RTA occurred in any part of the world (Stats SA 2009). Although, with the attempt of improving the degree of consistency in the road accident data in South Africa, some other agents aside the RTMC were approved to manage the road accident data in South Africa. The approved agents are identified as the Statistics South Africa [Stats SA] and the National Injury Mortality Surveillance System [NIMSS] (Stats SA 2009). The NIMSS is an institution under the jurisdiction of the Medical Research Council/University of South Africa [MRC/UNISA], authorised to coordinate the processes developed for the acquisition of information regarding the non-natural deaths comprising the RTA deaths, and deaths caused by diseases, which are regularly obtained from the hospital records (Stats SA 2009). The ministry of transportation could establish a communication interface with the other authorised agents to consolidate the data produced by the RTMC for data-sufficiency purpose. In view of that, consistency in the production of road accident data could be determined to support decision-making processes in the ministry.

Substantial emphasises have been laid on the data shortcomings in the previous chapters, which affirmed the gravity of criticisms made by the data users. The public affirmed that the current situation of the road accident data administered by RTMC, only offers inaccurate, insufficient, and incomplete data (Sinclair 2011; Joseph 2013). Specifically, these shortcomings are ascribed to poor management problems, and thereby necessitate a practical exploration into the fundamental methods of acquiring the road accident data at the local traffic department. This study will instigate a reform process for proper acquisition of unprocessed road accident data [raw data]. In that case, an in-depth understanding into the operational procedures implemented at the ARU in the STD is valuable in this study. In general, the evaluation of the operational procedures will give valid suggestions into the problems thwarting the accurate processing of the unprocessed data. The organisational process of STD is discussed in the section 3.1 as the first step into the implementation of the first stage of the methodology applied in this study.

2.5.7 Literature study on data quality improvement

Data is regarded as significant information in RTA, and has huge impact on the possibility of reducing the critical rise in the RTA. Stats SA (2009) emphasised on the effect of data limitations towards obtaining the right result in terms of death records, such as incompleteness of background information on the deceased, inadequate specification of cause of death, and under-and-late registration of deaths (Stats SA 2009). Definitely, all these aforementioned limitations contributed to the poor data quality (Stats SA 2009).

However, many research works have been carried out on the improvement of road accident data quality several decades back. Some of these researches were based on the purpose of, firstly, making comparison of different data sources in order to establish a unified data source to stimulate consistency and completion of data (Agran & Dunkle 1984; Agran et al. 1990; Diana & Knuiman 1994; Rosman 2001); and secondly, others concentrate on how to improve the process involved in producing a reliable data to the public [data users], by identifying limitations affecting the dissemination of quality data to the public by the data producer [RTMC] (Shinar et al. 1983; Zegeer 1987; Agran et al. 1990; O' Day 1993; Ferrante et al. 1993; Marshall & Rossman 1994; Kevin Austin 1995; Johnson 1999; Jacobs & Aeron-Thomas 2000; Baguley 2001; Wang 2004; Engineering Advice and Services 2005; Njord et al. 2005; McDonald et al. 2009; Jeffrey et al. 2009; Asia Injury Prevention Foundation et al. 2010b; Sinclair 2011; Wang & Strong 2013).

Furthermore, some of these research works based their facts on the inability of the police officers or traffic officers to provide full details of an accident event. Few out of these articles further clarified that imprecisions such as omission or loss of relevant data fields has rendered most information or data provided in the ARF incomplete, while few others provide insights into procedures required to develop a practicable process suitable for acquiring quality data, by assuring a complete road accident data for public use. One of the benefits of the complete and accurate road accident data is to promote safety initiatives in the road transportation system. The accuracy of road accident data or information regarding the injuries and their causes is essential to road safety researchers, policy development and evaluation planning (Rosman 2001; Joseph 2013).

Lack of sufficient data has contributed immensely to the inability of devising a sustainable, systematic approach to tackle road accident occurrence. For instance; the increase in the RTA in the African Region, is considered as the inability of African nations to engage in the activities of devising a common way of sustaining and improving the existing process developed for road accident data gathering. And also, the sloppy handling of data in ensuring supply of accurate information is a prime factor to this effect (WHO African Region 2013; Joseph 2013; WHO 2013).

Diana L. Rosman (2001) emphatically estimated the rate of underreporting of accidents for different road user groups in Western Australia, where a road accident data comparison was performed between the hospital records and the police records. It was discovered that police way of keeping accident record is marred with much irregularities (Agran & Dunkle 1984; Zegeer 1987; Agran et al. 1990; O' Day 1993; Rosman & Knuiman 1994; K Austin 1995). As illustrated in the research article, the degree of reportage of injury severity by the police was indicated as less reliable type of accident record among others. A research work illustrated that both the police and hospital records are more adequate for the details acquisition as regards the drivers than the pedestrians (Rosman & Knuiman 1994).

In this context, the research is partly related to the investigation to be performed at the STD. The only observable difference is that the researchers established disparity between the police road traffic casualty reports and hospital accident reports, by ascertaining which among the two prominent data collectors is diligent in quality record keeping. Actually, the method used was based on the application of the probabilistic techniques to identify if there is any similarity establish between factors such as, injury type, severity and treatment, according to records obtained from both the police's and hospital's casualties records (Rosman 2001).

Generally, road accident data collected by the police are basically connected to three major areas which are accidents involving casualties, heavy traffic flow, vehicle hijacking, and use of influential substances such as alcohol, drugs etc. A statement made by Diana L. Rosman (2001) revealed that "police attend accident scenes if there are casualties or suspicion of alcohol involvement or if traffic flow is impeded" (Rosman 2001). However, during the research process, limitations to the linkage between the police and hospital reports were accentuated as lack of completeness, lack of accuracy, and lack of consistency. These three basic factors were discussed as the part of the components that constitute a quality data in subsection 2.5.4 above.

Furthermore, Diana L. Rosman (2001) made it clear that a major reason for art of matching traffic accident records is basically to deduce or minimize the existence of errors in both sources. This could be achieved by determining the gravity and thresholds of the errors committed, in order to clarify the differences and limitations in the comparison between the two data sources. Regarding this approach used by Diana L. Rosman in 2001, there are only a few details in the hospital records that could be linked to the police accident records. The details identified as *linkage factors* are road users [accident victims], accident locations, vehicles involved, genders, and number of people killed [casualties], and the circumstances of the road accidents which is also part of the records.

However, a report made available by Buffalo City Municipality (2005), regarding the reformation of the unproductive procedures implemented in processing the road accident data, offers the basic steps that are applicable in reforming the procedural systems designed for road traffic assessment. These steps are particularly considered in the eradication of the excessive accumulation of irregularities along the process. And also, the report illustrated the method considered in identifying factors that contribute to the rise or accumulation of irregularities. This method was implemented to evaluate the primary data sourcing phase with other subordinate data sourcing phases. The method used could be beneficial to the similar investigation to be carried out at the STD.

The method recommended a clinical way of cleansing the existing data through various initiatives, and setting new protocols to ascertain the appropriate capturing process for data collected (BCM 2005). However, two researchers, Ehnes & Niu (2012) specified in a published report, how reporting system and data collection system can be improved, alongside a standard procedural method for data evaluation and analysis. In this report, further emphasises are directed towards the basic principles of active reportage, data collection and evaluation procedures.

Emphatically, the method reveals the predominance level of some factors in the improvement of data processing system. Factors considered are the consultation of the stakeholders, design of reportage system, administrative and departmental structure, periodic maintenance and adjustment of the national database to accommodate more relevant data (Ehnes & Niu 2012). More so, this particular method, addresses areas such as regularisation of tools for data reporting, assessing the system used in data reporting and collection, and developing a reliable analysis platform for accident reported and data collected for appropriate evaluation process (Ehnes & Niu 2012).

In addition, a group of research institutes carried out similar method on how to strengthen the possibility of acquiring a reliable data processing system concerning RTA. The group illustrated the significance of reliable data and the assessment of the road safety data in a specific region (Asia Injury Prevention Foundation et al. 2010b). Moreover, the group dissected the key steps that are relevant towards the application of RTI to strengthening the existing road accident data systems, or the design and implementation of new ones (Asia Injury Prevention Foundation et al. 2010b). Thereafter, a valuable improvement approach towards the data processing system was discussed by the group. This particular approach was based on the quality level of data collected by the police through a set of evaluation processes, in order to check the reliability of the data and system performance. The approach consist of areas such as in-depth evaluation of existing systems, data collection tools, reporting requirements, training, and quality assurance (Asia Injury Prevention Foundation et al. 2010b).

A feasibility study was carried out on approachable methods that were implemented formerly by experts in this particular field. These methods were studied to aid the investigation carried out at the STD. The study will guide the evaluation of the degree of quality data collected by the local authorities as the primary data collectors. In this case, the completed ARFs [data forms] will be evaluated, with the purpose of determining the practicality of the data elements [variables] completed in the data forms to the real-world problems. In the course of executing this, the perception of the reporting officers on the application of the ARF will be ascertained through assessment procedures to satisfy the research questions in this study.

2.5.8 Conceptualised techniques for the data quality improvement

Few articles were published regarding the assessment of road accident data towards an improved quality data as earlier discussed. Generally, some articles discussed the effect of poor road accident data processing on the possibility of reducing the continual occurrence of accidents; while some only discussed the causes of poor quality of road accident data. More so, a few researchers pointed out areas or factors that contribute to poor quality of road accident data, but no suggestions were offered on how to tackle the problems.

Fundamentally, with the intention of achieving a quality data, it is imperative to assess the data sourcing procedures at the local level. Since the degree of data sanitary achieved at this level defines the quality output of the data processed at the provincial level. As a result of this, a practical evaluation of the procedures used in sourcing road accident data is essential. The evaluation process sensitises the quality level of data captured at the local level, through a practical assessment of the data with regard to the output results realised. Before this could be achieved, significantly, a feasible practice of appropriate assessment techniques designed for such case must be carried out. This demonstrates the basic purpose or importance of studying previous assessment approaches implemented by researchers, with reference to the improvement steps or procedures adopted in revitalising the quality level of data collected or processed at the local level.

Most articles or reports studied, performed the evaluation process for data quality improvement in different organisations. However, the practical understanding of the approach applied can be integrated into the procedures followed in the STD. This action galvanises the evaluation of the data collection procedures and the information/data representation on the data form. Many organisations are obliged to perform radical improvement of data quality as merit towards the appropriate decision-making process. Basically, the quality of data processed has a huge influence on the outcome of the decisions attained, regardless the level of improvement implemented towards the correctness of the data sourced.

According to Karr et al (2005), “data quality problems and actions are driven by decisions based on the data” (Karr et al. 2005). Two other researchers, Fisher & Kingma (2001) declared that the gravity of the decisions make is determined by the level of flawless of the information acquired through an unbiased process. If the expected outcome comes otherwise, then the process is considered corrupt or bias, which means it accommodates unobserved errors that required evaluation process (Fisher & Kingma 2001). This erroneous condition is mostly experienced when an organisation is incapable of revamping their information from an elementary data processing system to an advanced data processing system.

Moreover, a local municipality in the eastern part of South Africa published a report paper on the road accident data management procedure. The report paper encompassed the attributes of assessing the protocols implemented for road accident data collection within the East London area and King Williams Towns, under the jurisdiction of the Buffalo City Municipality [BCM]. From the report, it was depicted that the primary implementation of the assessment process was executed to determine or identify the problems preventing the department in acquiring quality data.

As a result of this, evaluation process implemented leaped from the duplication of road accident data, and detection of incorrect representation of accident locations to non-existence location coordinates (BCM 2005). The department established practicable steps to redress the problems mentioned. The evaluation process executed herein focused directly on the data collection tool and the RTI provided on the data form. The process was extended towards the establishment of GIS dataset, and protocols were restructured to ensure proper capturing of data (BCM 2005).

The first evaluation process begins with data cleansing process, which incorporates the correctional measures or practicable steps, ranging from data completeness checking, duplicated data eradication, location codes modification, and discarding of uncaptured accident reports to the integration of the historical data with the newly processed data (BCM 2005). After the implementation of these predominant steps; hence, procedures are developed and implemented to sustain the regular or periodic assessment of the application of the data form. According to the objectives stated in the report, the main purpose of implementing these protocols is to develop sustainable processes that ensure accurate data collection and capturing protocols at BCM. This strengthens the possibility of revitalising the process followed by the reporting officers towards data sourcing.

This evaluation process traversed from the monitoring of operations performed within and outside of the municipality, through a close monitoring of the officers' competence level towards a proper execution of allocated tasks. The possibility of achieving these outcomes was based on three attributes which are:

- Comprehensive understanding of the protocol followed.
- Efficiency of the protocol.
- Propose restructuring of the protocol.

These three attributes predict the assurance of achieving the objectives of the investigation. Another researcher, Batini et al (2009) provided layouts regarding the steps to follow in analysing and comparing data quality. The steps encompass the approach of developing a platform suitable for dimensional assessment of data quality. Unlike the approach used by the BCM (2005), where the investigation was carried out by the personnel of the department. In this context, the investigation team was exempted from studying the process adopted in the department, since personnel involved in the investigation have the knowledge of the process. Batini et al (2009) emphatically recommended that the most reliable way of resolving the poor quality process cultivated by an organisation is by being involved in the process. Obviously, this promotes fundamental knowledge about the procedural system adopted by the organisation. The information obtained by the researcher during the process will be sensitised towards the prospect of identifying the cause of the problem or perhaps the areas that are vulnerable. Moreover, Batini et al (2009) enumerated the steps involved in evaluating the problems affecting the quality of data in an organisation as:

- Organise the phases and steps that constitute a suitable research methodology platform.
- Develop strategies and techniques appropriate for the assessment and improvement of data quality level.
- Define the right dimensions and metrics that suit data quality assessment procedure.
- Identification of the cost types that maybe encountered in the course of performing the assessment procedure, such as cost associated with poor quality data, which is defined as the *process cost* caused by the data errors and the *opportunity cost* due to loss and/or missed revenues; another type is cost of assessment and improvement activities, which is also referred to as *direct cost* (Batini et al. 2009).
- Understand the nature of data considered in the research methodology; which defines the data type whether as primary data, which comprises unstructured and semi-structured data points, or secondary data which consist of structured data points.

- Understand the information transaction tools implemented for data processing, data modification and data management system in the research methodology (Batini et al. 2009).
- Perform close observation on the authorised personnel involved in the process developed for data updating.
- Evaluate the process followed in the data processing system.
- Evaluate the operations or services [tasks] performed in ensuring regular production of the data.

The attributes outlined above serve as paradigm for the methodology developed for the investigation to be performed in the STD. In the same article, Batini et al (2009) offered no illustrations on the procedure to be followed regarding the evaluation of organisations [authorities] involved, process followed, and operations or services performed as part of the research methodology, but rather explicated the significance and application of data quality dimensions in the data assessment process.

However, more details were offered concerning the first six attributes. The first attribute is *'phases and steps',* which constitutes the approach developed in acquiring background information about the organisation protocols and services, suitable dimensional measurements for assessing the quality level of data, and process improvement in selecting appropriate steps, strategies and techniques towards the possibility of achieving data quality goal (Batini et al. 2009). In essence, the first attribute galvanises the decision of channelling focus directly on the assessment procedures suitable for evaluating the problems rendering data quality inadequate. More so, this aspect measures the quality of data collected and the protocols followed through relevant dimensions (Batini et al. 2009).

The second attribute comprises the '*strategies and techniques*'. Under the second attribute, two types of strategies were recommended to guide the evaluation process, depending on the applicability of the strategies towards the process practiced in the organisation. The two types of strategies recommended were *'data-driven*' and *'process-driven'* strategies. Actually, strategies can be incorporated into the methods considered in the phases and steps. However, the data-driven strategy is described as *'an improvement process implemented for cultivating and improving quality of data processed through a direct modification of the data value* (Batini et al. 2009)*,'* while process-driven strategy is described as *'a considerable improvement strategy implemented for the redesigning of the processes that produce data'* (Batini et al. 2009).

The third attribute is recommended as *'dimension and metrics*'*,* which basically transpired the quality level of data, subject to sequential implementation of quality assessment processes [refer to subsection 2.5.4 above]. The quality dimensions are implemented to evaluate the level of purity of the data processed. This particular attribute is significant during information [data] outsourcing and information *[data]* transaction processes[6]. In addition, these dimensions are implemented according to their order of relevance to the research performed. This aspect is perceived as a key interest to many researchers in this field, as the process of uncovering the extent of poor quality data. The list of acceptable quality dimensions is provided in the table below according to the description provided by Batini et al (2009).

Table 4: Quality dimension elements

Quality dimensions	Descriptions
Accuracy	**Extent to which data are correct, reliable and approved (Wang & Strong 1996; Batini et al. 2009); data are adequately accurate provided that, such data collected are practicable to the real-world systems (Batini et al. 2009).**
Completeness	**Capacity of information system to represent every meaningful state of a real-world system (Lee et al. 2002; Batini et al. 2009; Wang & Strong 2013). Completeness is categorically recognised as the extent to which a data coverage covers all significant features as anticipated.**
Consistency	**"Violation of semantic rules define over a set of data items" (Batini et al. 2009). This is often experienced during data entering or data coding process. It is also applicable to the evaluation of the data collected with the use of data form, in form of unstructured and semi-structured data format.**
Time-related dimensions	
Currency	**Currency is defined by Redman (1996) in an article published, as the extent to which data is up-to-date or complete as cited by Batini et al (2009). This is defined as the time or period required in the entering process of data collated into the dataset.**
Volatility	**Volatility is simply defined by Jarke et al (2001) as "the time or period, for which information is valid in the real-world (Batini et al. 2009)"**
Timeliness	**Timeliness is described as the extent to which implementation processes are performed accordingly as regards data processing. Wang and Wand (1996) referred to Timeliness as "the extent to which the age of data is appropriate for the task at hand" (Batini et al. 2009).**

The fourth attribute recommended is *'costs'.* However, Batini et al (2009) described costs as "a relevant perspective considered in methodologies, due to the effects of low quality data on resource consuming activities" (Batini et al. 2009). From another viewpoint, costs can be defined in terms of the data collection as *'evaluation process implemented to quantify the energy, time and resources exhausted in the course of acquiring and improving quality of data or quality system for data processing'* (Batini et al. 2009). In conjunction with the last statement, Batini et al (2009) further emphasised that costs of low quality data is measured as the outcome of quality evaluation and improvement actions performed on data collated.

[6] Refer to section 3.2 for more understanding into the descriptions of both data sourcing and data transaction.

Due to this, management will be able to quantify the effect of poor quality data on the progress of the operations performed within the organisation, and as well, on the services rendered to the public [users]. For instance; cost of preventing regular occurrence of RTA on the South African roads is estimated to R300,000,000,000 [billion] from R210,000,000,000 [billion] annually (SANRAL 2008; SANRAL 2012), as cited in the subsection 2.1.3 above. English (1999) classified cost of poor quality into process and opportunity costs (Batini et al. 2009). Process costs are defined as *'the cost associated with the recapitulating and reassessing of whole data process due to errors discovered'* (Batini et al. 2009); while opportunity costs are referred to as *'the costs due to lost and missed revenues'* (Batini et al. 2009).

The fifth attribute is identified as the '*nature of data'.* This defines the type of data considering for data analysis. In this case, three types of data are considered, which are unstructured data, semi-structured data and structured data. The sixth attribute is considered as the *information transaction tools*. This is valuable in measuring the capacity of the tools involved in the data collection, data modification, and data management procedures.

The National Institute of Statistical Science [NISS] published a technical report illustrating the steps required in assessing the data quality warehoused in the database through statistical observations (Karr et al. 2005). The institute accentuated on the need to identify organisational issues as the first step to be itemised in the improvement checklist (Karr et al. 2005), which is similar to the recommendations offered by Batini et al (2009) regarding the appropriate methodologies towards the assessment of quality of the data produced. Moreover, many organisations possess the obligation to identify the organisational issues frustrating the process for producing quality data, but they are incapacitate from addressing such issues due to inadequate resources distribution (Karr et al. 2005).

Furthermore, Karr et al (2005) explained the fundamental steps to follow in addressing issues relating to data quality. The report structured these steps as part of the data quality dimensions and assessments. These steps can be considered as mechanisms for data quality improvement. Contrary to the methods discussed by Batini et al (2009), Karr et al (2005) focused directly on the data that are already refined or structured. In addition, the mechanisms are considerably implemented to determine the problems associating with data recording and analysis procedures as discussed in subsection 2.2.4 above. In this regard, the process cut off the assessment of the primary source of the data processed or inputted into the database. The mechanisms consist four essential steps identified as:

- **<u>Preliminary screening for data quality</u>** –This step was implemented to ascertain the quality effect of the primary assessment procedure adopted for the assessment of the data collected or gathered before being stored in the database. Karr et al (2005) emphatically specified that this particular step is mainly applicable to organisations that are incapably weak towards quality data production. Moreover, preliminary screening for data quality entails some correctional protocols initiated in reforming the quality of data collected. These correctional protocols are structured into four relevant processes, which are considered as *discarding of data*, *inadequate data resources*, *data quality improvement* and *data improvement not required* (Karr et al. 2005). These aforementioned protocols play important roles in ensuring the need for an improved implementation. The first of the correctional protocols, which is *discarding of data,* is implemented to ascertain the practicality of the data collated; followed by the *inadequate data resources*, which is implemented to check whether resources or means of improving data quality are ascertainable or obtainable; and thereafter, *data quality improvement* is implemented to justify the prudent implementation of the available resources in terms of personnel, money, time and initiatives; and the final protocol, which is *data improvement not required*, is implemented to quantify the level of improvement required, whether is low, high or no improvement required at all (Karr et al. 2005). This part of the mechanisms is made feasible through decision making process.

- **Exploratory analyses for data quality assessment** –this step is applied through appropriate analysis platform developed, wherein a vast overview or background of the quality of the data processed can be observed or acquired through statistical observations. The exploration of the data through descriptive statistics will showcase the practicality of the data acquired. Other important observations could be obtained by understanding the structure of the data [in terms of attributes, missing and incomplete data points, data grouping etc.], characteristics of variables or features, connections between variables and relational characteristics *[by checking for incorrect coding in the database during the creation of relational entities between variables]* (Karr et al. 2005).
- **Identification of anomalous data elements** –this aspect streamlines the irregularities that could be identified during data mining process without accessing the related database. The process can also be performed through critical observations by characterising the data collected. Beside the first approach, according to Karr et al (2005) anomalies can also be detected through the use of robust statistical approach. This approach directs attention on irregularities such as *identification of duplicate records*, *identification of omitted information* and *identification of regular misrepresentation of data* (Karr et al. 2005).
- **Selected methods for data quality improvement** –this step determines the right approach towards the data quality improvement. Karr et al (2005) affirmed the relevance or significance of this particular step in identifying any sort of inconsistencies that maybe existing in the database system. Obviously, this is applicable where

inconsistencies are existing among the data points and data definitions (Karr et al. 2005).

The application of these mechanisms is based on the practical understanding of the technical issues frustrating the quality of data stored in the database. From previous observations, it was depicted that human factors are extremely huge contributor of errors committed during data processing (O' Day 1993; Vogel & Bester 2004; Karr et al. 2005; Vogel & Bester 2005). According to a statement declared by Karr et al (2005), "data quality concerns are problems of large-scale machines and human generation of data, the assemblage of large data sets, data anomalies and organisational influences on data characteristics such as accuracy, timeliness and costs" (Karr et al. 2005).

Technically, errors are periodically, if not frequently committed at the preliminary level of data sourcing [data collection]. The desire to discard or rectify these errors guarantees minimal rate in errors committed during transfer of data from recording stage to analysis stage. More so, human factors should be prioritised as the major concern in the process of restructuring the degree of quality of the data assembled in any organisations. To buttress this statement further, Karr et al (2005) stressed that human factors are often the difficulty part in the quality data achievement because they are the controller of the data collection equipment. In order to complement the efforts of procuring quality data, Fisher and Kingma (2001) advised the data management to consider information *[data]* collected as a critical product, not simply as derivative of the process.

A group of research institutes published a book on how to design, improve and implement data systems (Asia Injury Prevention Foundation et al. 2010b). The book recommended necessary steps to be followed in designing, improving and implementing data systems, according to the type of operations and data processed in any particular organisation. The *design* as part of the approach is basically applicable to construction of new data systems, while *improvement* is based on the restructuring or reconstructing of the existing data system (Asia Injury Prevention Foundation et al. 2010b). Furthermore, in the particular book, the steps recommended were developed for the betterment of the road accident [crash] data system. The steps are developed to strengthen the capacity of the accident data processing system. The common similarity between the approaches presented in this book and other approaches discussed earlier is '*improvement*'. All the approaches particularised on the improvement of data quality, although the recommended approaches are slightly different from each other.

Nonetheless, the research group was organised by the World Health Organisation, with the aim of aiding the authorities or organisations in charge of road accident data management towards a standard road accident data record. In the book, the group recommended that a *working group* should be established, which consist of personnel who in one way or the other involved in the particular operation to be investigated (Asia Injury Prevention Foundation et al. 2010b). The research group recommends that the working group should include agencies, authorities and personnel that come together to design and develop the right objectives to fathom the best approach for the improvement of the data quality system adopted by an organisation. Furthermore, the integration of the working group into the improvement system, will facilitate a feasible and comprehensive description of activities to be executed to promote the development of correctional measurements towards a reliable result.

The involvement of the working group will catalyse the process of developing strategical methods necessary to revitalise the road accident data system. The working group provides checklist of relevant strategies to reinforce the output results of road safety data systems (Asia Injury Prevention Foundation et al. 2010b). The group suggested that relevant data elements, with accurate definitions of the accident features should be accommodated in the data collection tools such as ARF and database system.

Actually, the upgrading process suggested for the data collection tools can be executed while establishing a new data processing system or restructuring the existing system. In this regard, the extent of the decision making on whether to establish new data processing system or restructure the existing system depends on the outcome derived from the situational assessment carried out by the working group. According to Asia Injury Prevention Foundation et al (2010b), “when road accident data systems are improved, the reported number of injuries can rise, sometimes dramatically, because the system has become more effective at capturing events, or generating sufficient accident data” (Asia Injury Prevention Foundation et al. 2010b).

In addition, the group explicated the steps considered in reforming the existing system or establishing a new system. The checklist prepared by the working group suggested likelihood areas that require urgent improvement in the existing data processing system. This checklist contains the findings discovered and strategies developed towards a practical solution for improving data quality system. The implementation of the situational assessment is based on the goals and objectives required in the investigation actions such as:

- Studying the available data,
- Examining the quality and practicality of existing data,
- Defining improvement needed,
- Deciding on tools to be implemented for the collection of information, and
- Formulating the applicability of the data and appropriate dissemination (Asia Injury Prevention Foundation et al. 2010b).

The above itemised steps stimulate the methods used towards achieving the goals and objectives designed for the improvement of the existing system. Basically, these methods will promote the tendency of uncovering the problems affecting the data obtained from the existing system. The recommended steps can be integrated into the methodologies designed for the investigation to be performed in the STD.

A book published by Volunteer Estuary Monitoring (2006) elaborated on the importance of *data management, proper data interpretation* and *data presentation* in data processing system. In this book, some observable problems were identified which were attributed to human factors and the applicability of the data collection tools. This book suggests that appropriate management of data should be taken as one of the prioritised responsibilities of the personnel in charge of data processing system (Volunteer Estuary Monitoring 2006). The book further clarified that incomplete data and inaccurate data entry contribute to the problem of data interpretation (Volunteer Estuary Monitoring 2006). Theoretically, the effect of this on data processed rendered the data impractical, and destructs the tools implemented in fortifying the process.

The book analysed the relational relationship required to be established among the three significant personnel involved in data processing, which are *management personnel*, *preliminary data collector* and *data users* (Volunteer Estuary Monitoring 2006). The scholar emphasised on the need for the data management personnel to always be prepared towards any need for the restructuring of the structure or shape of the data collected by means of a sensible data management plan (Volunteer Estuary Monitoring 2006). In connection with the last statement, practical questions based on technical know-how can be developed and implemented, for instance: *how data will be processed*, *when it will be processed*, and *who will be responsible for each task* (Volunteer Estuary Monitoring 2006).

These questions can be applied to build up a feasible paradigm on how to tackle or identify the exact points or areas that may be vulnerable to poor data processing. In actual fact, understanding how a process works or operates provides a clear insight or in-depth understanding into likelihood problems that could be anticipated ahead of possible occurrence. The book described data interpretation as method of suggesting or providing a simple way of having clear understanding of the problem confronting quality of data (Volunteer Estuary Monitoring 2006). This aspect actually contributes tremendously in the process implemented in uncovering the issues affecting the data quality through the procedural steps enumerated below:

- Classifying variables that failed to meet quality criteria,
- Monitoring sources that fail to meet the criteria,
- Identifying the factors influencing the condition or possibility of achieving the quality criteria (Volunteer Estuary Monitoring 2006),
- Identifying the effect of incorrect data sources on the correct data sources, and

- Identifying the influence of the management sloppiness on the output or results obtained from the data or information that could be processed (Volunteer Estuary Monitoring 2006).

If the aforementioned procedural steps can be followed accordingly, then there is tendency of ascertaining the major problems, and the valuable chance of fathoming the right solution to resolve the problems. In the process of implementing the procedural steps, it is necessary to develop survey tools to support the investigation by seeking the opinion of the personnel in charge of data processing. Another group of scholars published a report on accident records systems. In the report, a similar layout of steps provided in a report published by Asia Injury Prevention Foundation et al (2010b) were recommended. These steps practically apply to the process of establishing new accident recording systems.

Besides, the group identified set of organisational problems that could hinder the success of acquiring quality data. In addition, this serves as the basis for the development of new accident recording system. The cost of establishing new system is much higher than the cost of improving the existing system. Therefore, it is advisable to work on the existing accident recording systems, because the benefit of developing a new recording system does not pay off immediately depending on the standard of the system (O' Day 1993; Njord et al. 2005; Asia Injury Prevention Foundation et al. 2010b).

The group recommended several actions anticipated to be executed towards a successful implementation of accident recording systems. The first action is the establishment of a state wide traffic records coordinating committee, which incorporates the corporate involvement of the stakeholders on right decisions making processes (Njord et al. 2005). This step is similar to the working group developed by Asia Injury Prevention Foundation et al (2010b). The coordinating committee setup comprises the combination of the data collectors, system managers, information technologists and safety analysts (Njord et al. 2005). The group further recommended the reason for integrating or valuing the feedback from the data users, which serves as part of the basis for measuring the standard of data processed (Njord et al. 2005). Ultimately, this includes the integration of the data collected by other authorised data collectors into the data collected by the local authorities. By considering this particular recommendation, thus, probability of obtaining sufficient data is possible. This will catalyse the rate of completeness of the available variables or features on the data form.

The remaining steps recommended for successful implementation of accident records systems are fundamental knowledge of traffic records systems and simplification of road accident data collection system (Njord et al. 2005). The group declared that, it is always important to have an in-depth understanding of the traffic records system to substantiate the investigation carry out. In this case, competent personnel in this field are recommended as part of the consultation process to redress inability of having valuable leads to the problems thwarting the improvement of data. Njord et al (2005) added that the features or data fields provided in the data form must be accurately defined, simple, rich and easy to comprehend, in order to be suitable for the purpose it is designed for.

The recommended strategies discussed in this section are evaluated, mastered and validated before being integrated into the proposed methodology developed for this research project. The strategies are perfectly observed and considered adequate for the investigation performed at the STD. The compatibility attributes of the strategies are checked before it is integrated into the methodology developed or designed for the ongoing investigation. In addition, these strategies are implemented to reinforce the methodology, in an attempt to practically fulfil the objectives of the study. The steps taken in implementing the right methods for the research project are elaborated in the following chapter.

3. Methodology of the study

In this section, the methodology applied as briefly mentioned in the section 1.2 covers four-stage approach. The four-stage approach is designed based on the integration of the literature studies performed on previous researches (BCM 2005; Karr et al. 2005; Volunteer Estuary Monitoring 2006; Batini et al. 2009; Asia Injury Prevention Foundation et al. 2010a; Asia Injury Prevention Foundation et al. 2010b; Ehnes & Niu 2012) relating to the investigation to be carried out at the Stellenbosch Traffic Department [STD], and also to satisfy the strategies required in achieving the objectives of the study (Batini et al. 2009). The implementation of the four-stage approach focused directly on the road accident data processing units in the department, by classifying the involvement of each unit in the operations. The tasks assigned to each of the four stages will be discussed in the subsequent paragraphs. However, a flowchart diagram illustrating the four-stage approach is presented in Figure 8, and strictly followed to ensure reliable complete research conclusions.

The first stage entails *'understanding'* of the operational procedures adopted at the STD. Concurrently, the process is *'evaluated'* to determine if data is consistently processed, and also to ascertain the standard of the operation. However, a comprehensive knowledge of the activities performed in the department will facilitate a clear understanding of the roles played by each unit towards achieving a possible outcome. This stage involves the study on the evaluation of the personnel and agencies involved in the processes as recommended by Asia Injury Prevention Foundation et al (2010a; 2010b) and Batini et al (2009) in the section 2.5.8.

Moreover, the second stage involves *'identification'* of both *vulnerable areas* and *responsible factors* that could be accountable for anomalies affecting the quality of road accident data collated at the STD. Nonetheless, preliminary exploration of data completed in the ARF will be utilised to determine the responsible factors contributing to the low quality level of road accident data at the STD as recommended by (Volunteer Estuary Monitoring 2006).

The third stage encompasses the study of available data by examining the quality and viability of the data represented in the ARF, and the necessary decisions required in selecting the right tool for data *'collection'* as recommended by (Batini et al. 2009; Asia Injury Prevention Foundation et al. 2010a; Asia Injury Prevention Foundation et al. 2010b). And also, a questionnaire is *'developed'* to aid the qualitative and quantitative data collection, with the purpose of gaining statistical information to validate the significance and findings of the study. The practicality of the questionnaire depends exclusively on the valuable opinions of the respondents [participants], as a way of establishing an analytical relationship between road accident data collected and the application of ARF (Karr et al. 2005). In this case, the *dimension* and *metrics* required in initiating the data collection process are appropriately defined, to aid simple transformation of metadata to structured data (Batini et al. 2009).

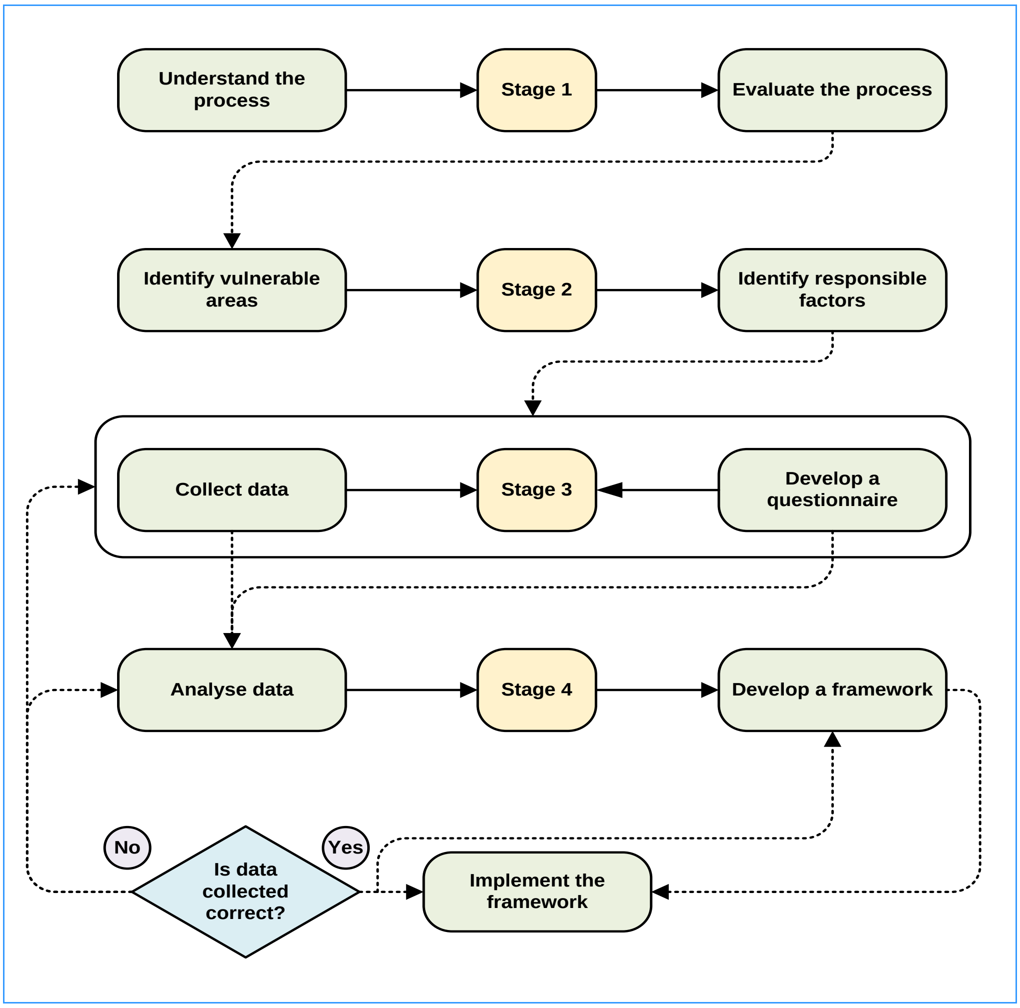

Figure 8: A flowchart illustrating the process approach of the four stages in research methodology

Last of the four-stage approach comprises three processes. The first process involves the analysis of the data collected to measure persistent accumulation of errors per data field in the Accident Report form (Karr et al. 2005; Volunteer Estuary Monitoring 2006), the practical analysis of data usability as applicable to the road accident occurrence within the Stellenbosch area, and the analysis of the applicability of the Accident Report form by the form users through the survey methodology (Karr et al. 2005). These analyses are achieved through the introduction of the statistical based methods like descriptive statistics to the measure average scores and the frequency distribution of errors per field in each related factor, and also to showcase the interpretation of results to illustrate the consequence of road accidents on the Stellenbosch environs. The set of scores realised from this process will be plotted on suitable charts as dictated by the structure of the scores, such as bar chart, pie chart, line chart, box-and-whisker chart and histogram. The second process entails the design of a framework in accordance with the trends realised from the analysis carried out. Then, the final process involves the implementation of the framework to supplement the existing approach practised in the traffic department.

3.1 Organisational process

In this section, the implementation of the four-stage approach is initiated through the first stage of the methodological approach, which involves the appropriate understanding and evaluating of the process engaged in the STD. Previously, in subsection 2.4.1 above, a full organogram of the Stellenbosch Municipality was presented with full details of obligations executed by the divisions under its jurisdiction. The relational interaction between all the divisions was also discussed. In this current section, a sectionalised part of the organogram comprising of the Traffic Services division, as shown in Figure 9 below[7] will be discussed to clarify the focus area of the study. The role played by each division constitutes the organisational process applied in stimulating the protocols for road accident data procurement.

The sequence showed in the Figure 9 depicts the flow of operations within the Traffic Services division. The Manager Traffic Services coordinates the distribution of resources, and constantly keeps the service efficiency functioning within the division. The Head of Traffic Law Enforcement executes allocated tasks according to instructions received through a dispatched memo from the office of the Manager of Traffic Services.

The activities performed by this office are categorically based on the functionalities of the other units under her jurisdiction. The office manages the road safety policy regulations enactment, and also coordinates the strategic way of guiding the road users from violating the rules and regulations of the NRTA in South Africa, and thereby raise the chance of reducing the occurrence of road accidents in the locality. Notwithstanding, this office gives directives to the Superintendent of Training, Education and Accident Investigation and Superintendent of Law Enforcement on the issues curtailing the progress of activities within the division in the area of traffic law enforcement, traffic control, operations planning and scheduling, and operational expenditure for execution of activities.

Additionally, the Superintendent handling the training, education and accident investigation oversees the offices of the CRO and the Clerk, both operating within the ARU. Moreover, the Superintendent manages the establishment of extensive programmes to educate the public on the road safety initiatives. Further responsibilities of the Superintendent are; to execute public awareness programmes on road safety, and supervision of the accident investigations[8] by regularly communicating with SAPS to ensure regular acquisition of ARF and assist in the process of sketch plans and photographs of any accident reported.

[7] The diagram shown in Figure 9 is adapted from Stellenbosch Traffic Department (2014).

[8] Read more about the responsibilities of Superintendent in the section 2.4.1.

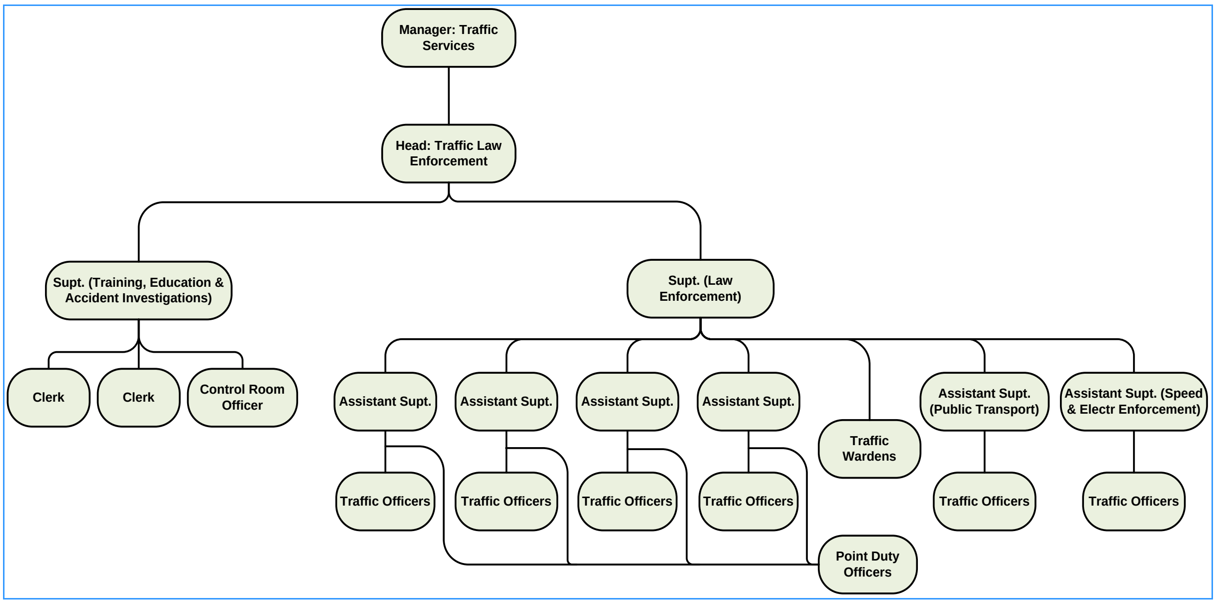

Figure 9: Sectioned organogram of Stellenbosch Traffic Department [STD] depicting the division in charge of road traffic problems

The office also manages any request that may be tendered by any external institutions for any investigation to be carried out in the division. The communication interface between the Superintendent of Training, Education and Accident Investigation and the CRO stimulates the degree of consistency in the appropriate processing of accident information, and also promotes rapid dissemination of information to the public regarding awareness of danger or occurrence of accident at any particular locations along the roads.

Moreover, the Superintendent of Traffic Law Enforcement coordinates the emergency activities and gives instructions directly to the Assistant Superintendents and Traffic Wardens, on the need to execute important tasks such as identifying danger areas along the roads, controlling heavy traffic flow, and lessening the chance of traffic casualties. The officers assist in regulating traffic in heavy congested areas or detective traffic signals, and execute duties as per schedule. These activities promote the public safety awareness through regular road inspection processes, and also enforcing traffic laws to reduce road accidents. All the aforementioned obligations are executed by the junior traffic officers as instructed.

The assurance of satisfactory processing of RTI across every bureau in the local traffic department, depends also on the effective communication between the offices of the Superintendents. The instructions given out on allocated tasks, are expected to be implemented accordingly by the officers appointed such tasks.

3.1.1 RTI handling unit

The handling of RTI in the STD is carried out at the ARU, where the initial processing of data is executed. This unit performs its operational activities, according to the instruction given by a senior Superintendent Officer. The first step applied in acquiring road traffic incident information is primarily initiated through the reportage of incident occurring within the Stellenbosch locality, which required the use of the ARF for data gathering in accordance with the instructions provided in the form.

The procedure of collecting the incident information [data] is assigned to the reporting officers, who could either be a traffic officer or a police officer depending on the severity of the incident and the exact period of the occurrence. Although, the discussion on the involvement of the police officer will be peripheral in this section, rather the involvement of the traffic officers will be detailed in order to fulfil the objectives of the study. However, extensive details on the procedural steps followed in acquiring RTI are provided in the following section.

3.1.1.1 RTI transaction

According to the sectional organogram displayed in Figure 9 above, it is illustrated diagrammatically that Traffic Officers partake in the prime process of the RTI. The Traffic Officers performed their responsibilities by sensitising the traffic rules and regulations accordingly towards the public through the appropriate enactment of the NRTA, and also attended to any reported incidents.

Moreover, other traffic officers like Traffic Wardens, perform their responsibilities by enforcing parking violations, and Point Duty Officers assist in regulating traffic within the SM-WC 024. Basically, the services of the Point Duty Officers are required mostly at the intersections or particular roads with recurrent heavy traffic within the municipality. Nonetheless, they can also assist in the completion of the ARF if any incident occurs at a nigh location.

The reportage of the incident is simply completed by the traffic officer, but more often by the police officer, as defined by the severity of the accident (RTMC 2007). Furthermore, the information [data] gathered by the traffic officer submitted to the Clerk in charge of capturing the ARF. Aside this, the Clerk performs other operations related to the registering of the number of completed ARFs[9] received, and also obtained the particulars of the officer who delivered the completed forms for reference purpose. Afterwards, the accumulated number of completed forms registered will be transferred to the CRO for further evaluation, who is also considered as the DCO. The operations execute within this office will be discussed in the next section.

[9] Refer to description of a completed ARF in the last paragraph in subsection 2.2.3.

3.1.1.2 RTI control room

The control room is the central unit for RTI transaction from the STD [local level] to the provincial level. Several operations are performed within the cubicle of the CRO, mainly the transfer of information. In this office, operations like data capturing of the completed ARFs are performed alongside other important tasks; like the receipt and registration of the completed ARFs, compiling, capturing and transferring of RTI and provision of the reports or complaints on the traffic accidents [also refer to Figure 10]. The abovementioned activities are simplified below in details, in terms of the duty assignments assigned to be performed within the office of the CRO, such duty assignments as:

- **Receiving and registering of the completed ARFs;**
  - Inspecting the level of accuracy of the information provided in the completed ARF; in case irregularities are discovered, then, the CRO issues query to that effect.
  - Ensuring the safekeeping of the documented files, and also performing the accurate placement of the files in the right order of arrangement, with the purpose of avoiding misplacement of completed ARF. This particular task is performed by the Clerk according to directives from the CRO.
  - Accurate referencing of the registered copy of the completed ARF. This task is assigned to the Clerk according to instruction given by the CRO.
- **Capturing and transferring of information [data] gathered at the local level to the provincial level [RTMC] for further examination;**
  - Creating a storage folder on the local data set to accommodate the magnitude of the scanned copies of the accident records, and thereby regularise the orderliness of the accident records.
  - Assigning codes to the completed ARF to be captured, and also performing the entering of the information provided on the ARF into the local dataset.
  - Compiling the necessary information that needed to be transferred to the RTMC or provincial traffic authority.
- **Attending to traffic accident issues by providing necessary information to support any on-going accident investigations and other related issues;**
  - Liaising with the police in order to confirm the conclusion of any accident investigations that might be delaying the delivery of the completed ARF.

- Assisting the accident victims or the insurance firms on funeral benefit claims or insurance benefit claims, by providing necessary information regarding a particular accident case.
- Informing or alerting the traffic officers or road safety patrol team on a newly reported incident that happens on the road, which required immediate attention.
- Notifying the fire/hazard and the emergency departments anytime accident takes place with the aid of preventing further escalation of the situation.
- Radio control with complaints, emergency services required and on-scene accident update.

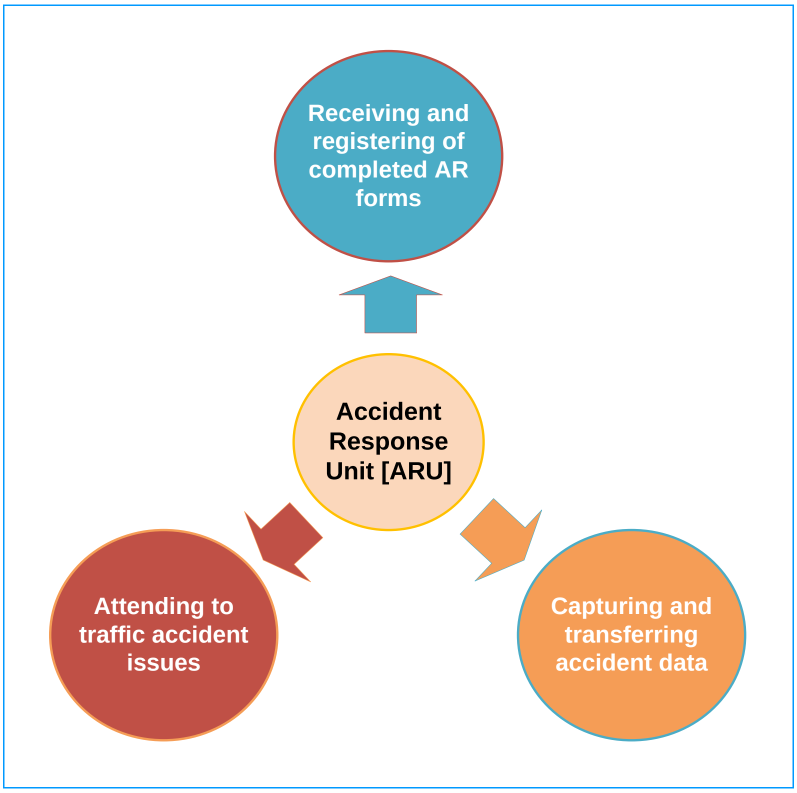

Figure 10: Operations performed in the Accident Response Unit [ARU]

On many occasions, the traffic officers personally gathered completed ARFs from the SAPS, in order to deter untimeliness in the compiling and transferring of information as scheduled.

3.1.1.3 RTI acquisition procedural model

In the previous section, the unit handling the RTI in the STD was discussed, in an attempt to provide a detailed understanding into the responsibilities discharged by the personnel involved in the processing of RTI. In the former discussion, the central operations performed by the CRO were mentioned, to clarify the importance of her position towards a quality RTI handling. The cubicle of the CRO is perceived as one of the key stages of information processing in the local traffic department. The officer is responsible for determining the quality level of the fundamental information that should be captured into the local data system. In addition, the valuable support of the Clerk in ensuring sustainable RTI processing system was not excluded.

Generally, it is essential to discuss the procedures used in the STD in acquiring and transferring RTI from the data outsourcing stage to the data processing stage within the jurisdiction of the local traffic department. However, according to the model displayed in the Figure 11 below, five police departments are classified under the jurisdiction of STD with regard to RTA matters. The five police departments considered are Cloetesville, Franschhoek, Groot-Drankenstein, Klapmuts and Stellenbosch. The diagram in Figure 11 illustrates the acquisition and transfer of RTI between the SM and SAPS involved in the management of the traffic issues within the Stellenbosch locality. The *dash-arrow* in the diagram indicates the direction of transferring the RTI, from one office to another through some checkpoints to determine the level of quality; while the *bold-arrow* indicates the districts, the departments and the offices involved in the process within the Stellenbosch locality.

The five police departments mentioned in the diagram constitute the locations prone to RTA within the jurisdiction of SM. Moreover, as diagrammatically illustrated in the Figure 11, the transfer of information commenced with the action carried out by the accident victim or the eyewitness in reporting or notifying a nearest police department in the districts where the accidents actually occurred. Occasionally, the STD/MMT is also notified about the incident, wherein the attention of other authorised agencies such as the emergency response [ambulance services] and protection services [fire] are required to salvage the situation on the location.

Right after the report of the incident, the nearest Police Department [SAPS] is expected to respond rapidly, by assigning Police Officers on rapid response duty to attend to the incident. In the process, the services of the Traffic Officers, Emergency Response [Accident] and Protection Services [Fire] are also required at the scene of the accident. Afterward, the dimension of the accident will be determined by the Police Officer, to ascertain the classification of the accident before commencing with the data collection process. The Police Officer determines if the accident is fatal or nonfatal, in order to clarify whether a case docket will be opened, and instantaneously commences with the collection of RTI [data] with the use of the

ARF, as the primal data collection tool. In some cases, where the traffic officer is present at the scene of the accident, then he/she decides to collect the information [data] at the incident. As earlier stated, it depends on the dimension of the incident. At this stage, the reporting officers are expected to be conversant with the procedures required in the completion of the ARF, in accordance with the training given as guideline towards the effective application of the form.

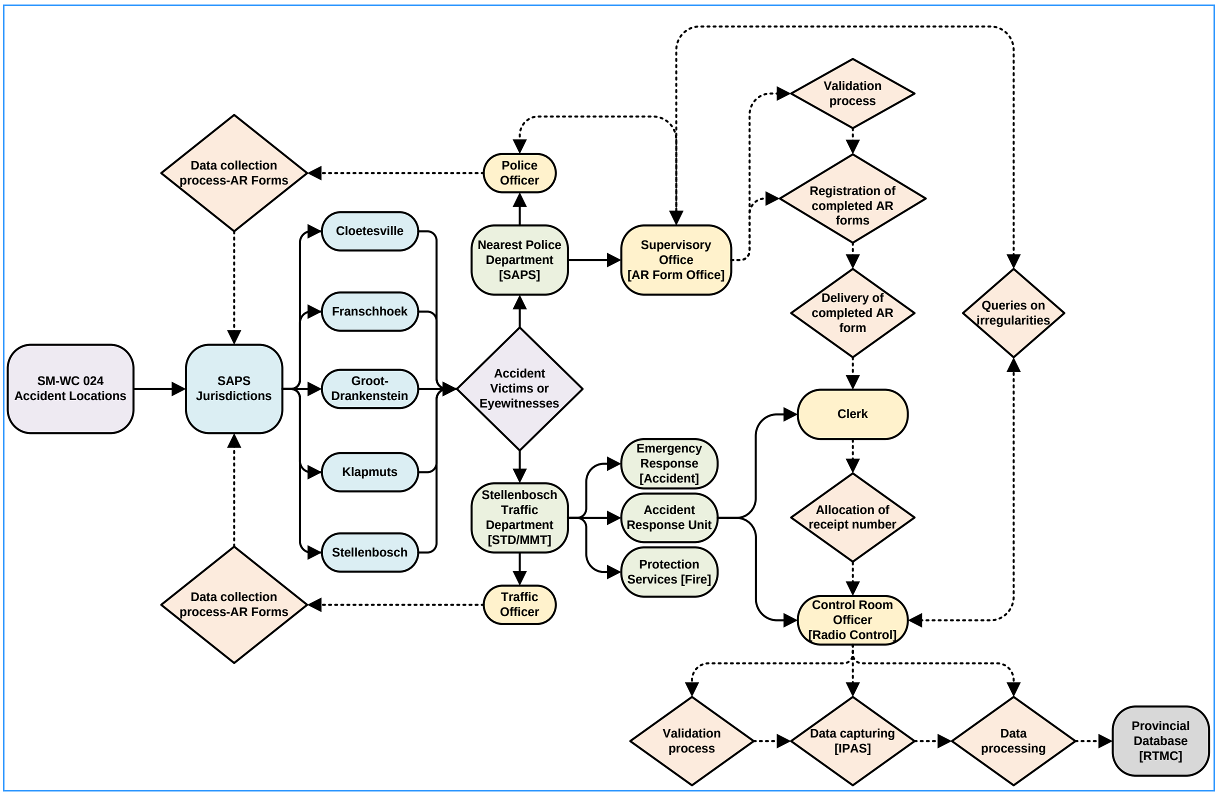

Figure 11: Procedural model for RTI acquisition in the Stellenbosch Traffic Department [STD]

After completing the ARF, the police officer submits a completed copy of the ARF to a Supervisory officer, who is in charge of the ARF office at the Police Department [SAPS]. Thereupon, the Supervisory officer performs necessary evaluation process to determine consistency, completeness, and accuracy of the reportage before registering the accident into the SAPS 176 Accident Register. The evaluation process performed by the Supervisory officer is classified as first validation process.

In any case where the Supervisory officer finds the report of the accident sufficient, then certified copies of the completed ARFs will be transferred to the local traffic department through the use of a Delivery Note Form [DNF]. Otherwise, a query will be issued with reference to the poor quality of accident coverage, which compels the reporting officer to implement the necessary corrections as indicated by the Supervisory officer. The DNF is established to support process uniformity while transferring RTI along the information processing line, from one local authority to another local authority (Sluis 2001). In some cases, the local traffic department requested for completed ARFs from the local police department, in an attempt to initiate a rapid process of the RTI as scheduled on daily and monthly basis.

The DNF is delivered to the ARU, and received by the Clerk, who allocates the receipt number of the completed forms, and also register the details of the Police Officer delivering the completed ARFs. These steps are followed to ensure consistency in the record keeping process for reference purposes. The Clerk checks the validity of the forms completed by determining the particulars of the Police Officers, who recorded the accidents on the ARF.

Subsequently, the Clerk submits the evaluated copies for proper examination, with the aim of ascertaining the level of correctness of the information provided in the completed ARFs. In any case, if errors are discovered at the cubicle of CRO/DCO, then a query will be issued to the SAPS by particularising the exact areas or data fields that require immediate corrections. Once the returned copies are declared valid, then certified copies will be captured. Thereafter, the CRO/DCO determines the next processing stage of RTI depending on the quality of the information. However, the sequential operations performed by the CRO/DCO within the ARU as illustrated in the Figure 11, are validating of RTI, capturing of RTI into the IPAS, and transferring of RTI to the next stage of data processing.

3.2 Operational procedures evaluation

In this section, the second stage of the four-stage approach is carried out in accordance with the research methodology flowchart illustrated in Figure 8. The implementation of this particular stage facilitates the determination of the vulnerable areas and responsible factors. The need for improving the quality of data processed at the provincial level and disseminated at the national level, requires an in-depth investigation into the process applied or followed in acquiring the unprocessed data at the local level, which is the basis for the primary acquisition of the RTI before being processed. Although, while observing the operational procedures for the acquisition and transferring of RTI as discussed in the subsection above; however, some problems were observed according to the statement made by the Clerk, which are attributed to the degree of inconsistency along the operational procedures.

According to Figure 11, two stages of validation process were implemented in the reportage of RTA at the local level. The first stage of validation process implemented, was executed at any of the five police departments, depending on the location of the accident within the area under the jurisdiction of SM-WC 024; while the second stage validation process was carried out at the local traffic department by the Superintendent, who is authorised by the provincial authorities to execute the assessment and capturing of the RTI. Basically, the purpose of the two stages of the validation process, in the police departments and traffic department, is to scrutinise the completed ARFs for any existence of inconsistencies, incompleteness and inaccuracies in the reportage of the RTAs, and also to ensure reliability in the task performed at the initial stage, with the aim of strengthening the quality of RTI captured.

The problems observed while processing RTI through the two local authorised departments are categorised into two types. These problems are mostly encountered in the areas showed in orange colour in Figure 11 according to findings obtained. The classification of these problems, as described below, is based on the structure of operations performed by the units involved in the process.

- **Problems encountered during RTI outsourcing** –these problems are encountered during the collection of data at the accident location. These particular problems are due to any circumstances that influence the capacity of the reporting officers to be able to collect relevant and sufficient data.
- **Problems encountered during RTI transaction** –these problems are encountered during information [data] transfer, within the two local authorities managing the preliminary data collection process of RTA within the SM.

Furthermore, the aforementioned problems are encountered in the three main information sourcing stages, which are highlighted below with detailed clarifications:

- **The accident location** –this stage is characterised by three parties who are; the accident victim, the eyewitness [if any], and the reporting officer [Police Officer or Traffic Officer]. This particular stage determines the level of purity of the information prearranged for processing. Although, there are likely problems observed at this stage that could deter the level of purity of the RTI outsourced. These problems are outlined below as:
    - ¤ **The late response to incident reportage** –this could be encountered due to delay response in the accident reporting. In this case, the chance of acquiring information depends on the prompt response from the authorised local departments. The major difficulty experienced from this problem is the inability of the reporting officer to acquire sufficient RTI, which is attributed to cases such as; the reluctance of the accident victim from cooperating with the police *[by not disclosing his/her relevant information],* the condition of the accident victim *[at the period of the accident],* and/or refusal of the eyewitness *[from informing the nearest local authorised department about the incident, and also from supporting the accident reportage with valuable and tangible information].*
    - ¤ **The poor understanding of the accident circumstances** –this is attributed to the level of understanding of the accident case by the reporting officer. In some cases, this particularly leads to the incorrect determination of the accident

dimension, which causes difficulty in the classification of accident types. This may extend to assemblage of incorrect information or omission of the essential parts of the ARF.

- ¤ **The inappropriate application of the ARF** –this problem arises through the incorrect applicability of the ARF in the process of acquiring RTI. This is one of the main issues thwarting the quality of RTI collated by the local authorities. The major problem encountered here is the inability of the reporting officer to collect the right information; for example, incorrect indication and constant omission of answer options. This is as a result of the difficulty that maybe experienced in the understanding of the practicality of the features provided in the ARF as related to the accident environment.

- **The local police department [SAPS]** –this stage comprises the problems that could be frustrating the quality of RTI at the SAPS. In addition, the problems identified here may contribute to the poor handling of RTI at the local police department.

  - ¤ **The incompetence of the reporting officer** –this problem renders the reporting officer ineffective, wherein he or she lacks the technical know-how of applying the appropriate procedure towards the acquisition of information. However, this particular problem could be ascribed to insufficient training as regards the RTA issues, and it may be due to poor level of intelligence of the reporting officers towards the application of the training skills acquired. This contributes to the difficulty of obtaining the right information for the right classification of RTA.

  - ¤ **The inadequate supervision proficiency** –this issue is absolutely one of the strongest major causes of undiscovered errors in the reportage of RTA. The ultimate problem experiencing here, is the inability of controlling or detecting anomalies at the primary stage of information [data] assemblage at the local police department [SAPS]. This case is attributed to the lack of commitment from the part of the Supervisory officer, who is assigned to supervise or inspect the degree of completeness of RTI provided in the completed ARF. For more clarifications, this problem may lead to delay in RTI transfer, due to inadequate supervision skills demonstrated by the Supervisory officer.

  - ¤ **Lack of effective communication** –this particular problem arises due to poor relational interface between the Supervisory officer and the reporting officer [police officer]. As a result of this, the possibility of effecting or making corrections regarding the RTI provided in the ARF is truncated. The consequence or adverse effect of this particular problem could result to incompleteness of RTI, and inconsistency in the processing of information.

- **Late delivery of completed ARFs** –this is a huge issue in the processing of RTI across the two authorised local departments in SM. The delivery of completed ARFs on time is observed as an issue that requires immediate improvement as stated by DCO in charge of the control room at STD. According to her, the problem daunts the swiftness of determining the quality level of RTI provided in the completed forms due to prompt demands from the provincial department [RTMC] on a daily, quarterly and monthly basis. Basically, this problem could continue to recur, unless a new officer is allocated the responsibilities of managing the prompt delivery of completed ARF to the STD. This problem caused the failure to meet up with the demands of the provincial department.

- **The local traffic department [MMT]** –this stage accommodates the problems confronting the personnel in the road accident data capturing unit, regarding the acquisition of quality RTI to promote the standard of safety measures that could be relevant in minimising the occurrence of road accidents in the Stellenbosch area. The probable problems identified at this processing stage of RTI are highlighted below:

  - **Incorrect allocation of reference number** –this kind of problem is attributed to the errors that could arise from the cubicle of the Clerk, who assists the CRO in the RTA matters, by acknowledging the receipt of the completed ARFs from the local police department. The major problem here is basically the mix-up or misplacement of accident record files [completed ARFs] subject to incorrect allocation numbers. The cause of this problem can be ascribed to factors such as inexperience and lack of commitment.

  - **Delay in response to queries** –this problem arises due to the negligence attitudes exercised in response to queries issued as regards the poor coverage of RTA. This frustrates the ability of the CRO to procure sufficient information on time.

  - **Inadequate validation process** –this is a cogent issue that requires urgent improvement. This problem is similar to the issue of *'inadequate supervision proficiency'* which was previously discussed among the problems frustrating quality RTI. This particular problem is encountered within the cubicle of the CRO, which could be due to the gravity of unresolved errors invalidating the practicality of the RTI. This problem contributes to the quality of RTI.

- ¤ **Inappropriate information [data] capturing procedure** –this is an uncommon issue in the ARU. Although, data capturing is part of the operations executed by the CRO. The likelihood cause of this problem may arise from fatigue; which could as a result of workload volumes, and this could trigger an unnecessary muddle up of tasks, and simultaneously leads to skipping over necessary steps require in data capturing process.
- ¤ **Delay in information [data] transferring procedures** –this encompasses the issues enumerate below:
  - Unresolved problems regarding the response to queries issued (O' Day 1993).
  - Improper codification approach towards the assemblage of the information gathered, which may lead to loss of information (O' Day 1993).
  - Lack of commitment from the personnel towards their operational tasks, which may lead to inconsistency along the procedural operations (O' Day 1993).

3.3 Data collection procedures

In this section, the third stage of the four-stage approach is implemented to enhance the procedures required in collecting the data intended for the qualitative and quantitative analyses of the usable data and the applicability of the Accident Report form. This section covers the procedures designed in acquiring the available traffic accident records at the STD. Steps applied in collecting, assessing and analysing data accessed at the local traffic department will be consecutively discussed in the subsequent sections. Meanwhile, the limitations encountered while assessing the available data cannot be shelved without being observed in this study. The limitations rendered some information void and inconsiderable for practical analysis of the data collected. More clarifications are provided in the section 9.2 regarding the determinable factors that rendered such information impracticable for the research purpose.

In the process of accessing the availability of data, the officers in charge of the accident records provided necessary guidelines to aid a successful data collection process. The guidelines are anticipated to prevent any unnecessary confusion while accessing the accident records. However, the guidelines extend to necessary areas that could support a productive research project at the department, such as:

- Moral conduct on the part of the investigator at the commencement of the research project to the last day of the data acquisition at the department.

- Confidentiality of the accident information; the investigator is expected to exercise necessary discretion towards the information acquired.
- The arrangement of the record files should in no cause be altered, in order to avoid any misplacement of the completed ARFs during the on-going assessment actions.
- The information provided in the completed ARFs should in no cause be altered unless officially instructed to do so.

However, access was only granted to the unprocessed data [raw data] available on a complete accident record year, while other recent records were classified unavailable due to official reasons relating to pending case dockets opened at the SAPS, or uncompleted investigation regarding some particular road casualty cases.

3.3.1 Research data

In this research, data acquired are categorised as primary data (Hox & Boeije 2005), sourced directly from the completed ARF records kept in the accident record cabinets in the ARU. The procurement of the research data was based on the configuration of the information provided on the completed ARF, which is the combination of both semi-structured data and unstructured data. Hence, the semi-structured data is described as *'a set of data partially structured or organised into data points, which still required further assessment to validate the usability of the data assembled for practical analysis';* while unstructured data is characterised as *'a set of data acquired through text format, where a determinable approach was implemented to structure the text into data, in order to have a clear knowledge of the dataset'* (Batini et al. 2009; Schoenbach 2014).

Broadly speaking, in the context of this research, the group of data collected are completely raw data [unprocessed data], which require a sequential evaluation process of determining the practicality level of the data towards the motive of this study. The process comprises methods implemented to assess, acquire, assemble, and analyse the data. However, a further discussion on the implementation of the evaluation process is provided in the subsequent subsection.

In addition, it is essential to have a substantial view of the research objectives in the type of data procured through the evaluation process. As a result of this, two types of data were formed from the set of data assembled. The two types of data formed are captured and non-captured [uncaptured] data. Captured data, in this perspective, is described as *'a form of unprocessed data completed correctly in the ARF, which has gone through the validation processes before it is captured into the local dataset';* while non-captured data is described as *'omission or* *inappropriate completion of data in the completed ARF, subjected to appropriate evaluation process to determine its degree of practicability'* (Schoenbach 2014).

Quintessentially, the significance of procuring these two types of data is basically to implement the objectives of this research, *which necessitate a complete evaluation of the anomalies thwarting the reporting and recording of road accident data in South Africa, and the necessity for the improvement of the ARF.* The practicality of the data will assist the traffic management to have an in-depth understanding into the problem challenging the level of services rendered to the public. And thereby, boost their degree of readiness to develop substantial plans to limit the daily occurrence rate of RTA in the Stellenbosch locality. The evaluation of the data will strengthen the capacity of having sufficient data to tackle RTA.

On the other hand, the problem of data mishandling is a huge concern to the local traffic management. This necessitates the request for the evaluation of the non-captured data in order to have a clear view of the key problems frustrating the probability of attaining a data quality level. The evaluation of the non-captured data is established on the definition of the forms of error found in the completed ARFs [refer to Chapter 4].

3.3.2 Evaluation process

Previously, the problems frustrating the appropriate acquisition of quality data at the local level were elaborated. This actualises the degree of viability of the procedures designed in assessing the reliability of the RTI transfer to the provincial level. Besides, a definite understanding of the procedural operations carried out at the STD was discussed, along with the units handling the RTI in the department. The transaction of RTI across the two authorised local departments was concisely described, to elucidate the transfer of information from one cubicle to another.

As part of the discussions that followed, central unit handling the assessment and capturing of completed ARFs was observed. Primarily, the discussion was based on the information acquisition procedural model, where the concise description of the information transaction from the five police departments to the local traffic department was observed. However, from the section that followed, a broad insight into the problems encountered during the information transaction amongst the two authorised local departments was explained.

The evaluation process implemented in this section is centralised on three significant phases which are data acquisition, data assessment and data analysis (Batini et al. 2009). According to Figure 12 below, the three phases comprise some basic technical steps as subordinates, which are part of the mechanisms required in achieving the objectives of the investigation. More so, these steps offer broader understanding into the practicality of the data fetched from the accident records kept in the ARU. In addition, the dimensions applied in establishing the methods of measuring the road accident data are attributed to factors that define the type of data to be assembled.

Furthermore, the evaluation process covers the procedures followed in acquiring road accident data through the implementation of assessment protocols. This approach will enhance the possibility of determining the degree of consistency, timeliness and accuracy in accident reportage within the Stellenbosch locality over the duration of a year. The purpose of applying this process was necessitated through the definition of attributes or features provided in the ARFs, which are classified as metadata yet to be processed. Metadata can simply be described as *'a process whereby attributes of a particular object or objects are structured into complete data'* (Volunteer Estuary Monitoring 2006; Schoenbach 2014). To be more specific, in terms of road accidents, the process is practically used to structure information regarding the details of accidents, through the application of accident technical questions (Baguley 2001; Sinclair 2011) formulated for the design of the ARF such as described below:

- **Where** did the accident happen *[location]*?
- **When** did the accident happen *[time, date and day]*?
- **Who** was/were involved in the accident *[driver, passenger, pedestrian and object]*?
- **What** causes the accident *[possible causes and contributory factors]*?
- **How** did the accident happen *[sketches and description]*?
- **Why** did accident happen *[technical reasoning]*?

The accident technical questions are relationally and simply structured in the ARF to compensate the classification of accident details. The questions are structured into metadata, in a numeric format, to support simple collection of data from the ARF. In brief, the evaluation process combined the applicability of the *data acquisition, data assessment* and *data analysis* in processing the raw data gathered as stated earlier in the third paragraph in this section. The three phases are simply implemented to structure the raw data through the use of a statistical analysis template as the data exploration tool. The tool is utilised to achieve the implementation of the three-phase evaluation in processing the raw data. The analysis template developed by using this tool was considered to suit the type of evaluation process to be implemented.

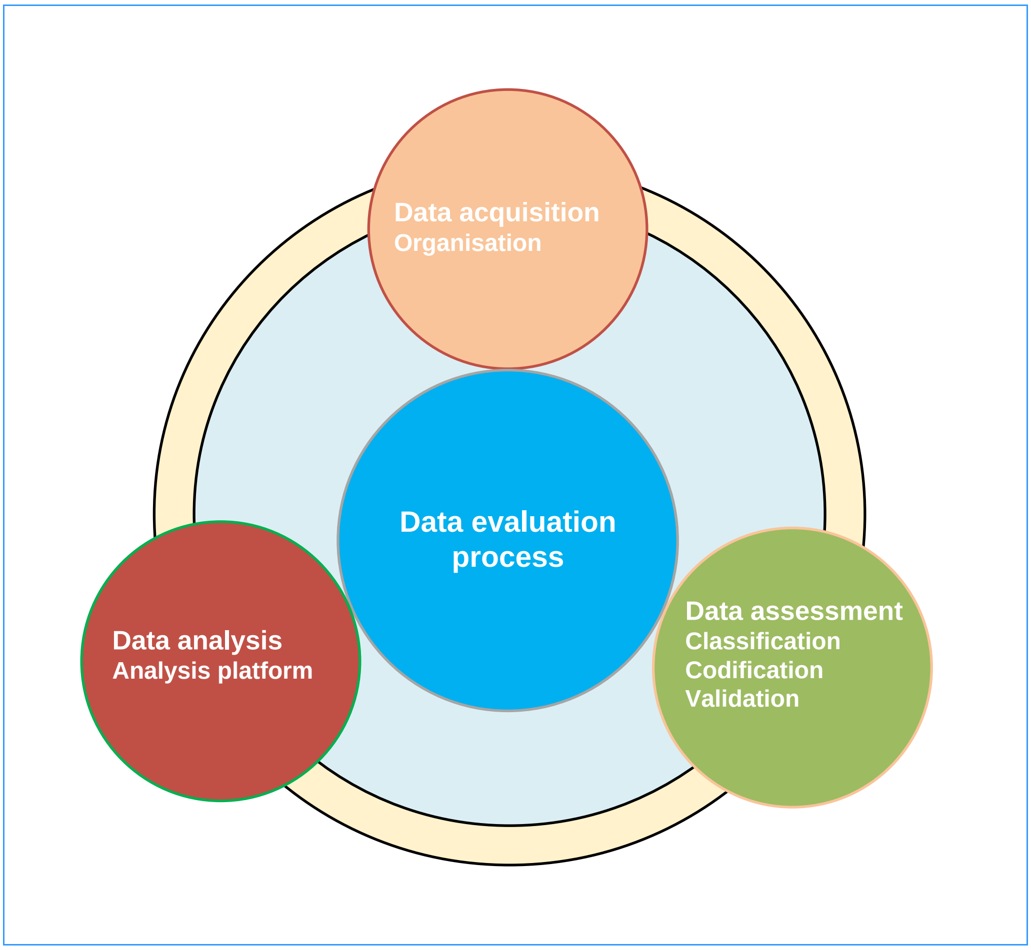

Figure 12: Three-phase evaluation approach applied in assessing the completed ARFs

A simplified practical model is developed to support a full illustration on the three-phrase evaluation approach presented earlier in the Figure 12 above. In the diagram displayed in the Figure 13 below, the first stage is the *data evaluation process,* which involves three phases namely *data acquisition, data assessment* and *data analysis.* Under each of the phases, different tasks are executed in order to stimulate the evaluation process implemented towards an achievable and reliable data exploration result.

The first task is executed as the data acquisition, which is initiated through an appropriate organisation of the data collected, with the purpose of having a clear understanding into the type of data to be organised. The second task that follows is performed in the data assessment, which is initiated through the implementation of the *data classification, data codification* and *data validation* sequentially, with regard to procurement of data and determination of data fitness for analysis (Lee et al. 2002; Volunteer Estuary Monitoring 2006; Batini et al. 2009).

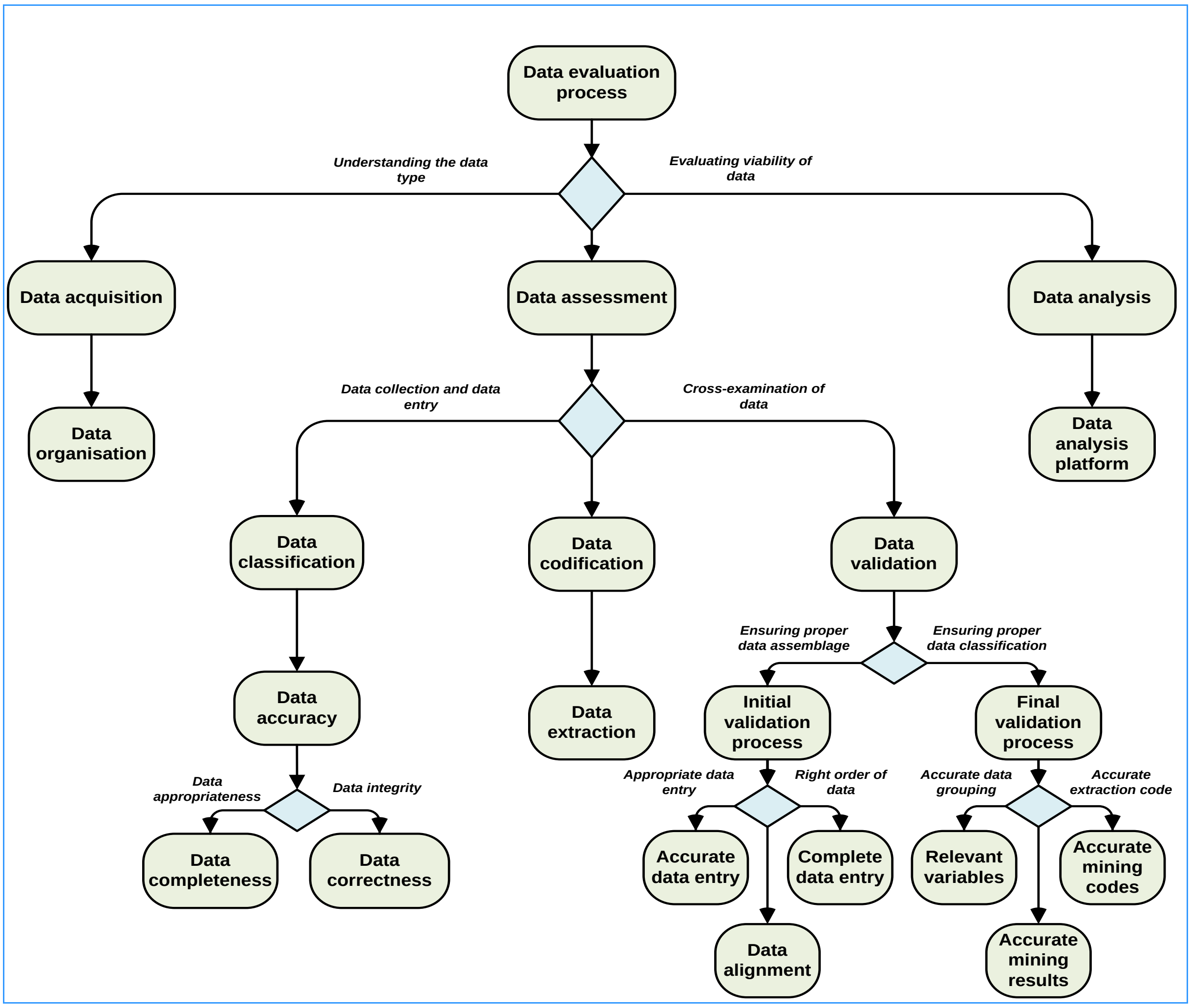

Figure 13: Practical model for the application of the assessment procedure for the research data

Moreover, from the same diagram, it is illustrated that some secondary tasks are executed directly in the data classification, data codification and data validation. First of all, from the data classification viewpoint, three classification processes were implemented that initialise the process of examining data precision level, and the reliability of the data assembled for the next stage. This process prepares the data for codes application towards the extraction of the usable data for analysis. Data codification is a process that encompasses the use of numerical codes to acquire or extract data points from the assembly of data set (Aitken et al. 2004).

After this process, a data validation process follows which constitutes two secondary stages. The two secondary stages implemented are categorised as initial validation process and final validation process. Each one of the validation processes consists three-process branches to quantify the outcome of the previous processes towards an attainable result. Tasks executed under the data validation are primarily to ensure appropriate data assemblage, and appropriate data classification to avert mix up during data assemblage (Batini et al. 2009).

To simplify the above statement, under the initial validation process, two major tasks were executed in an attempt to discourage any inaccuracy during data entry, by checking the completion and alignment of data. In a similar way, under the final validation process, two tasks were also executed to complement the task executed in the initial validation stage. The motive of performing these tasks is to ensure accurate data grouping and accurate extraction code to daunt incorrect application of the right codes. The concluding phase of the evaluation process is *Data analysis*, wherein a data analysis platform is developed to support a simple analysis process towards a reliable outcome. More simplifying details on the phases involved in the evaluation process will be discussed in the following section.

3.3.2.1 Data acquisition

In the literature study presented in Chapter 2, some practical definitions of data collection were provided to support the procedure of acquiring road accident data. Thus, in terms of road accident data, data collection could be described as 'a process that requires technical approaches towards the assemblage of the data fit for decision–making in the distribution of resources to combat the regular RTA occurrence, or improve managerial operations towards effective and efficient results' (O' Day 1993; Baguley 2001; Njord et al. 2005).

Conversely, from the structure viewpoint, attributes/data fields in the ARF [data form] are related variables formalised for the improvement of the road safety in South Africa. Primarily, the initiation of data acquisition procedure paves the way for the accurate organisation of the raw data collected. Data acquired are spread over 12 months of accident records documented in 2012. The data was entered manually into excel. The procedure followed in organising the data is discussed to details in the next section.

3.3.2.2 Data organisation

The data organisation pattern considered in this section is simply based on the arrangement of the data fields provided in the ARF. The data fields are already structured into metadata in the data form, by applying numerical codes for the answers to aid easy data collection process, as stated earlier in the last paragraph in subsection 3.3.2 above. Although, not all the data fields or answers are coded numerically, such data fields as *accident location*[10]*, accident sketch,* and *accident description* are in text format which required to be structured into a useful data format. In this section, the first step taken in the data collection process is by considering the design or development of a data collection template that suit the type of data to be collected.

[10] The analysis on the *'accident location'* is not made available due to the absence of the location coordinates on all the completed ARFs assessed.

The data collection template prepares the way for other evaluation processes such as classification, codification and validation of data. Otherwise stated, the data collection template serves as subordinate for the three-phase evaluation in ensuring quality data processing. Notwithstanding, these assessment processes will be discussed in the subsection 3.3.2.3 below.

First of all, the organisation of data collected is attributed to the type or nature of the data mined. This type of data is categorised as primary data. According to the nature of the data collected, a template developed aids such data assessment activities as easy extraction of data, accurate grouping of data, perfect data analysis, and also to reflect the integrity and quality of data mined with reference to the arrangement of the features in the completed ARF.

The design of the template incorporates relevant features or data fields in the ARF. The features are tabulated to avoid unnecessary disorder of data, by considering the accident register number [ARN] and the serial/capturing number [S/CN] as the reference points. This step supports the monitoring of consistency along the assemblage process. Thereafter, a technical preview process was established for quality purpose to unravel the errors that might exist in the data acquired.

Absolutely, the benefit of processing raw data in this investigation strengthens the chance of understanding the most interesting parts in the ARF to its users. However, this could reveal the areas that are contributing to poor quality accident reportage, and thereby discloses the incapacity of the management to observe such important circumstances. It is essential to comprehend the data fields or features in the ARF, in order to stimulate the process of developing a suitable template for data classification. More details will be presented in the subsequent sections regarding the implementation of the data assessment in ensuring accurate, sufficient and practical data required to accomplish the objectives of the investigation. In this particular section, data will be managed through the implementation of subordinate processes by considering the application of dimensional attributes for the data assessment protocol.

3.3.2.3 Data assessment

In this aspect of the evaluation process, three basic technical steps are implemented in the assessment process. The first step is referred to as *data classification,* where all available data are collected and entered into a designed data collection template. Thus, this step is executed through the procedure based on definite dimensions and metrics for data assessment (Batini et al. 2009). The procedure is basically developed to ascertain the availability of the raw data gathered from the actual number of reported accidents in 2012. The second step is referred to as *data codification*, where data is prepared for easy and quick data analysis. The codification process is applied to initialise the extraction procedure for both captured data and non-captured data. This step is exclusively carried out after the implementation of the data organisation and data classification processes.

The definite description of captured data and non-captured data will be discussed as part of the dimensions and metrics used in determining the quality level of data collected. Moreover, the final step implemented in the assessment process is *data validation*, where all the data classified and coded are validated to determine the integrity and quality of the data accumulated for the purpose of the investigation. The data validation process will examine the clarity of having the assemblage of the right data. Besides, the process will ascertain the correct application of numerical codes applied in the data extraction process and also determines if there is no disarray in the process of classifying the data acquired at the initial step of this section.

3.3.2.4 Data classification

This is the first step applied in categorising the defined variables for the data assessed. In this section, the data organised in the template developed are categorised into grouped variables according to the number of relevant data fields provided in the ARF. Previously, the benefit of developing the data organisation template was mentioned, which is considered as a guide to the classification of data into suitable groups with reference to the attributes provided in the data form. Data classification could be described as *'a process of arraying data in assemblage order of classes or groups of characters that constitute the data'* (Batini et al. 2009; Volunteer Estuary Monitoring 2006).

Basically, the classification process entails the collection of data and entering of data into the template developed. The basic method exploited in this section establishes a definite assessment procedure that could reflect the practicality of the data collected. Along the process, a data classification assessment exercise was defined based on the marked or indicated answers in the completed ARFs. The classification of the answers is placed directly under each field to ensure correctness and completeness to measure the degree of accuracy. With the aim of sustaining this process, the ARN and S/CN were entered in the first two columns of the template to validate the accurate entering of the relevant data. These served as reference points to guide the recovery of incorrect data entering, and as well, to prevent any possible disarray along the process.

Although, answers marked or indicated for each attribute or character in the completed ARFs were categorically assembled in data point format to stimulate correlations among the classes or groups of observations obtained. Since the data mined from the reported accident records are exclusively raw data; then, the need for proper classification to aid smooth extraction procedure is essential. Hence, this will deter any needless duplication of data, missing data or data omission before the data codification process follows.

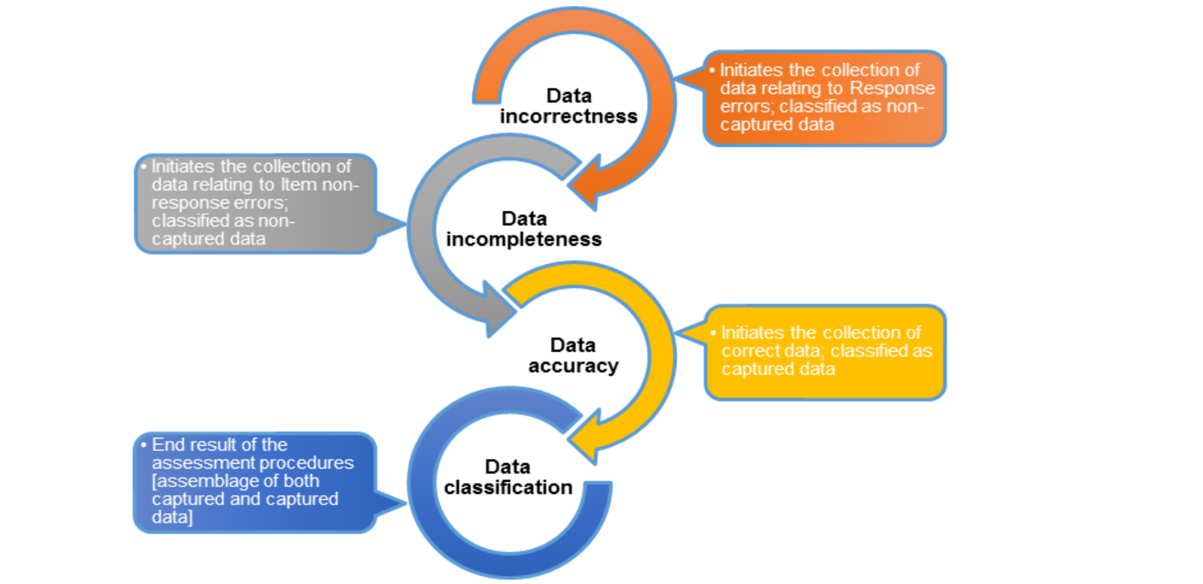

Figure 14: Dimensional metrics used for data assessment

From the template structure perspective, sections developed for attributes or data fields were demarcated to have a clear correlation between them. Fundamentally, the template aids the transformation of unprocessed data to processed data, by being able to distinctly categorise the captured data from the non-captured data. It further offers a clear understanding into the structure of the data collected.

The dimensional metrics applied in structuring the data points into practical data are demonstrated in the subsequent sections along with the methods implemented. Figure 14 depicts the process of Data Classification metrics.

3.3.2.4.1 Accuracy

This particular dimension technique is used to ascertain the degree of correctness or precision of the information represented in the completed ARFs. The correct entry of information in the completed ARF is regarded as *captured data,* while the incorrect entry of information in the completed ARF is considered as *non-captured data.* These categories of data are mainly observed and assessed through the appropriate application of the two other dimension techniques. All the information completed properly in the completed ARFs are entered into the template developed for data structuring process. Refer to subsection 3.3.1 above for more clarifications on the descriptions of both the captured data and non-captured data.

Moreover, with regard to the unstructured data in the completed ARF, it is necessary to define a simple and better means of structuring the data into data points. On account of this, metadata format was developed to transform unstructured data into data points for easy data collection and recording procedures. The three main unstructured data in the completed ARF are the accident location, accident sketch and accident description.

In addition, the written addresses of accident locations were not provided with any written coordinates on the completed ARFs, which indicates deficiency in the application of any advanced GPS device in acquiring the right coordinates for the road networks. If the coordinates are acquired, then the data will be analysed and plotted on the map-graph to show the exact locations where road accidents occurred mostly on the geographical location of Stellenbosch area. This approach will depict the most vulnerable streets or roads where accidents occurred mostly in the Stellenbosch area. Although, the possibility of carrying out analysis based on the different locations of accidents is disenabled due to the inaccessibility of data on the GPS coordinates.

Conversely, a different approach was applied in structuring the information regarding the accident sketch into metadata. Appropriate dimensions were established to quantify the sketch with most required information (Batini et al. 2009). The actual dimensions used to evaluate the accuracy of the sketch are considered as ‘informative sketch’ and ‘non-informative sketches’. The definitions developed as regards the two ways of quantifying the accident sketch is based on the instructions provided in the instruction leaflet of the data form. Thus, informative sketch is described as *‘an accident sketch that is completely drawn in accordance with the instructions provided in the ARF’;* while non-informative sketch is described as *‘an accident sketch that is presented without any consideration of the instructions provided in the ARF’* (RTMC 2007).

A similar approach was used for accident description. Categorically, the dimensions established in quantifying the accident description were based on the quality of written statement presented in the completed ARF. The metrics considered here are ‘informative description’, ‘non-informative description’ and ‘unclear description’. A definite description of each metric is provided to guide the understanding of the approach used to structure the text format into data points. Hence, informative description is described as *‘a written accident statement that is presented in an interpretative manner, which can simply be correlated to the informative sketch provided on the ARF’;* whereas non-informative description is defined as *‘a written accident statement presented with lack of descriptive features not in accordance with the informative sketch’;* although, unclear description is practically described as, *‘a written accident statement that has mix up languages, spelling errors and perhaps incorrect use of words’* (RTMC 2007).

3.3.2.4.2 Completeness

This assessment stage is focused on the determination of the degree of coverage of an accident as regard the information omitted in the completed ARFs. This measurement technique was used to determine the degree of item non-response error of the fields/elements in the completed ARFs. During the evaluation process, it was discovered that some factors actually contributed to inadequate RTI. Factors discovered as the major contributors to inadequate RTI are highlighted below as:

- Management negligence;
- Unreported accident cases.
- Late accident response.
- Reluctance in the reportage of accident.
- Inadequate competence in correlating accident related findings with the information/answers provided in the ARF.

Ultimately, the procurement of the non-captured data in this research is actually to understand the commitment of the reporting officer towards the possibility of procuring or gathering sufficient and complete data, and to identify areas that have the poorest data coverage among the data fields provided in the report form.

3.3.2.4.3 Consistency

This is the last of the three basic dimensions implemented for acquiring raw data from the available accident records. *Consistency,* as part of the assessment procedure, was applied to ascertain the degree of coverage of accident as regards correlation in the interpretation of information provided in the completed ARFs. To shed more light on the previous statement, the consistency will practically determine the feasibility or usefulness of all the information [data points] that are denoted in the ARF for the research purpose.

Fundamentally, the metric applied is basically to measure consistency in the interpretation of the information indicated or provided in the ARF either as *'incorrect'* or *'undefined'*. Nevertheless, the *'incorrect'* and *'undefined'* information left uncorrected are observed as errors, either as *item non-response errors or response errors.* For instance, *cases* such as: *'incorrect classification of answers, incorrect correlation of findings with the answers provided on the ARF, marking of double answers, muddle up answers, incorrect indication of answers,* and *text misspell are all considered as errors due to response errors* (O' Day 1993), while cases such as: *omission or missing information, that is, failure to indicate or write down information about the driver or cyclist, vehicle or motorcycle, location,* and many more are all considered as errors due to *item non-response errors* (O' Day 1993). More details regarding the analysis of errors discovered will be provided in Chapter 4.

Errors classified above can be attributed to the inability of the ARF users to carry out tasks according to the instructions provided on the instruction leaflet or manual. This suggests that the response to uniformity in the completion of the data form is very low; however, the purity of information provided in the data form determines the quality or reliability of data processed.

Therefore, it is essential to always measure the degree of usability of the ARF by the reporting officers.

3.3.2.5 Data codification

Data codification could be described as ‘the process implemented in transforming information into data points, which are considered appropriate for statistical analysis purpose’ (Schoenbach 2014). Furthermore, this aspect is considered very significant in the analysis of data assessed. It prepares data for technical analysis in the aspect of data grouping [assemblage] and graphical illustrations. The procedure was introduced after both the unstructured and semi-structured data has been captured into the dataset developed in the excel [data analysis tool]; although, all the data were gathered in a monthly sequence. Quintessentially, the data codification procedure is applied to extract data points from the dataset into data groups within the same data analysis tool. The dataset contains the data points of both the unstructured and semi-structured data measured, by using a specified dimensional metric as previously discussed in subsection 3.3.2.2 above.

Moreover, data points were extracted into various variable groups classified under each data field presented in the ARF, through the application of statistically computed formulas. The form of variables processed in this research is considered as Categorical Variables. This is a type of variable where observations are counted and grouped into various data groups of different variables for analysis purpose (Decoster 2006). The extraction of captured and non-captured data was successfully carried out through the application of the data codification process in the data analysis tool. The purpose of carrying out the extraction procedure is to organise the data into groups, interpret the data by understanding the practicality of the data, and analyse the data to acquire practical solutions to the real-world problems.

3.3.2.5.1 Extraction of captured data

The process introduced in mining the captured data into groups of categorical variables was achieved based on the appropriate assemblage of the variables in each field. The numerical formulas applied were formulated in accordance with the understanding of the type of data points prepared for analysis. In addition, the captured data consists the grouping of all available that are classified as fit and usable for analysis purpose. In this case, number of observable variables mined from the dataset were based on the counting of the number of occurrences of each categorical variable, while some were obtained by formulating the class interval, frequency and percentiles depending on the type of data.

Practically, numerical formulas applied are based on the nature and interpretation of the data categorised in each variable. Excel functions as *CountA* and *Countblank, Countif, Sumif* and *Frequency* are integrated together to fathom the best way to classify the data into categorical variables. *CountA* and *Countblank* were combined to have a complete number of reported accidents per month, with a full count of the dates and days that these accidents were reported. On the other hand, the introduction of both *Countif* and *Sumif,* is applied to obtain the number of observations of some important variables. For instance, *Countif* is applied in a case like; *total accidents occurred in weekdays = number of occurrence of each weekday that road accident occurred.* Similar data extraction approach is applied to variables or data fields as *Built-up area, Speed limit on road, Drivers'/cyclists'' countries, Gender, Severity of injury, Road type, Vehicle type* and many other important variables.

Considering the application of *Sumif,* two different variables were integrated to generate one possible outcome. The approach compares two ways of generating one possible result, but this depends on the reliability of data points acquired. For instance, *Sumif* is applied in cases like; *'total number of vehicles in accidents in weekdays'* and '*total number of vehicles involved in accidents in a month.*' The relational connectivity considered for the extraction of right data points for the two cases is the *'number of vehicles involved in accidents',* where day of week was considered as a reference base in order to actualise the integrity of the data assembled. This further assist in disclosing the extent of variations in the data points representing the 'number of vehicles involved'. The outcome of this approach is benchmarked with a direct sum of the total data points pertaining to single-vehicle and multiple-vehicle accidents, which yielded similar results.

The last numerical formula used in extracting data points is *Frequency.* This formula was considered to classify data fields that require the formulation of class intervals. The class intervals determine a complete distribution of data belonging to any particular variable or field. For instance, considering the *Time of accident,* all the registered periods of road accident occurrence are sorted with the intention of having a better view of the data acquired. In this case, accidents are reported 24 hours in a day, which indicates that accident occurrence periods are different. As a result of this, a class interval difference of one hour was established between the periods that accidents usually occurred, explicitly within the time interval of 12:59:00 am to 11:59:00 pm. This approach was used to produce a better way of generating data points for the periods at which accidents occurred in Stellenbosch area. Similar numerical approach was established to determine the distribution of the ages of accident victims. The class interval ranges from 20 to 100, with reference to a class interval of difference of 20, which dictates the appropriate way of grouping the ages of the accident victims.

3.3.2.5.2 Extraction of non-captured data

The mining of non-captured data was based on the definite criteria defined in terms of errors found in the ARF, such errors as item non-response errors, that is, omission of relevant data; and response errors, that is, incorrect completion of relevant data. The criteria constitute assessment of the non-captured data based on the void or inappropriate completion of relevant data in the ARF. Although, the degree of sufficiency of the road accident data depends exclusively on errors reduction. The importance of mining the non-captured data paves way for simplification of the problems associating with errors found in the ARF. The mining process constitutes the application of numerical formulas as done in previous section. The numerical formulas applied in mining non-captured data are *Countif, Sumif* and *Countifs.* The application of the *Countif* is similar to the approach used in the previous section. The numerical formula introduced is applicable to a single variable, which requires no reference base. For instance, mining the errors in the Day of week; *total non-captured data in weekdays=count of Sundays, "void", Mondays, "void" …. Saturdays, "void".* In addition, a similar approach was applied to compute other variables like *Accident date, Time of accident, etc.*

Other numerical recipes applied are *Sumif* and *Countifs.* These two formulas required the use of reference base to incorporate all essential features into the process of mining all the available errors in terms of item non-response errors and response errors, which literally means that void cells are counted or summed up for all non-captured data. The reference base was fragmented into five relevant mining metrics, which are *'one vehicle', two vehicles', 'three vehicles', 'four vehicles and more'* and *'no data'* as the relational connectivity relating variables together for a complete and accurate mining of non-captured data in each variable. In this approach, a 'one-vehicle accident' corresponding to void data [empty cells] in a variable is extracted; whereas similar approach is applied to other mining metrics, except 'no data' where only void data or empty cells in both the reference base and the require variable or field are extracted. In essence, 'no data' could be referred to as the empty cells in the reference base corresponding to empty cells found in the variables or fields subject to measurement.

3.3.2.6 Data validation

In the data validation segment, two validation stages were implemented for the evaluation of the raw data assemblage and data extracted into various groups as discussed in subsection 3.3.2 above. The two validation stages are categorised as the initial validation and final validation stages. However, the initial validation process consists the three assessment processes illustrated in the Figure 15. The three assessment processes implemented are discussed below:

- **<u>Accurate data entry</u>** –this validation process was implemented during the entering of data into the dataset developed. The process was carried out to crosscheck the accurate entry of the data collected from the completed ARFs [data forms]. This aspect was particularised to determine the degree of correctness or precision of the data entered. Although, most incorrect inputting of data encountered during this period is associated with the use of old ARF [refer to section 9.2], which excluded some relevant information that the new ARF had included. The step taken in rectifying this error was made carried out by using the ARN and S/CN to trace the completed ARF with errors.

- **Complete data entry** –this aspect of the validation process was implemented with the intention of determining the degree of completeness of the data collected. The process was carried out by crosschecking the level of correlations of the indicated answers in the completed ARFs to the data collected or inputted into the dataset. Actually, the benefit of implementing this process is to ascertain the degree of sufficiency and accuracy of the data assembled.

- **Data alignment** –this is the last of the three validation processes categorised under the initial validation process. This particular one was implemented to crosscheck the right order of the arrangement or grouping of the variable data points. This aspect is carried out to validate the right grouping of the required data elements [variables].

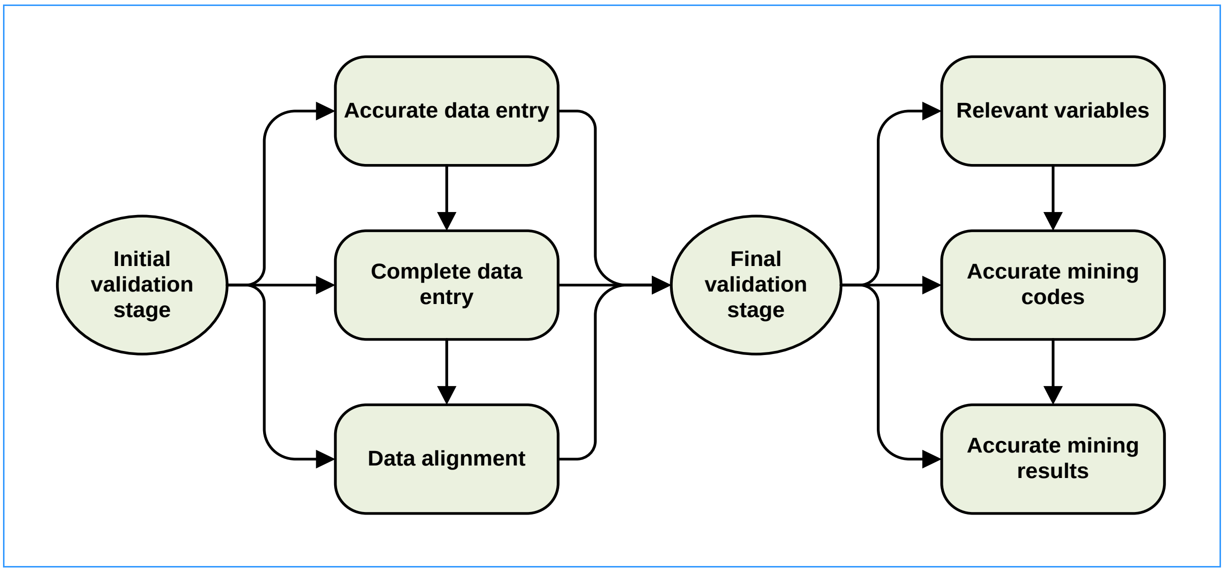

Figure 15: Data validation process stages

However, the *initial validation process* was enhanced by the introduction of the final validation process. In actual fact, the process carried out in the *initial validation process* ensures that all the data assembled in dataset are error free, reliable and practicable. Moreover, this primary stage prepared all the data for final validation process. At this level, the final validation process was implemented to complement the performance of the initial validation process, by ensuring that right variables are grouped, right codes are applied, and expected results are accomplished as illustrated in the Figure 15. Furthermore, these steps are highlighted below with concise details.

- **<u>Relevant variables</u>** –this particular validation process was implemented to ascertain if the right variables are arranged or grouped for the right purpose. The process evaluates the possibility of detecting assemblage or grouping of incorrect variables, which could prevent the practicality of the investigation, and thereby affects the determination of achieving the objectives of the study. In this case, the importance of detecting the errors is to pave the way for the right application of the numerical codes towards achieving the right results.
- **<u>Accurate mining codes</u>** –This determines if right extraction codes are implemented for the right outcome, in relation to the definition of the variables under which the data points are to be grouped. This specific validation process will ascertain if the right data points are obtained for the right variables.
- **<u>Accurate mining results</u>** –this is the last stage of the final validation process. The process is implemented to check if accurate data are acquired through the appropriate application of statistical codes in the extraction process. Actually, this process was initialised to determine the situation where statistical codes were perchance applied incorrectly for the extraction of unrelated data points for the right variables.

In this section, the implementation of the validation process was enhanced through appropriate arrangement of the data fields or attributes in the ARF. This establishes uniformity in the data grouping, through the right application of codes.

3.3.2.7 Data analysis

This section is considered as the concluding part of the evaluation process, which is implemented in exploring the quality of the data extracted. In another statement, the process is specifically considered as the practical way of evaluating the viability of all the data grouped in each variable. This approach is demonstrated through a descriptive analysis of data, and also through the graphical illustration of data analysed. In essence, the graphical illustration stimulates the chance of determining the spread of the scores *[data value]*. Data analysis can be referred to as *'a method of managing and transforming the feasibility of the data acquired, assessed and organised into a reliable result'* (Volunteer Estuary Monitoring 2006; Batini et al. 2009; Ehnes & Niu 2012). The major part of this section has been elucidated in the subsection 3.3.2.5 above.

On account of gratifying the application of data analysis, a platform was developed for data analysis to substantiate the tasks performed by data acquisition and data assessment in the evaluation process. However, this platform is referred to as *'data analysis platform'*, which serves as the medium of implementing the descriptive analysis of data, with the decision of exploring the relevance of the data grouped into variables. The platform is also designed to accommodate the implementation of the data coding process. Fundamentally, in this research, the data analysis is implemented to carry out tasks such as:

- Differentiating the practical data from the impractical data,
- Procuring the analysis results that display the significance or practicality of the data fields or variables, and
- Comprehending the correlations and disparities between the related data fields.

The above-mentioned tasks were used as yardstick to develop the approach to uncover the mystery behind the huge loss of data or quality of data gathered by the reporting officer in the process of completing the ARF.

3.3.2.7.1 Data analysis platform

This platform is developed in an excel to explicate the importance and reliability of the data assembled. The medium is established to have a clear view regarding the realism and distributions of the data points. Nonetheless, in this section, descriptive analyses were performed to ascertain the distribution of scores across each variable such as measures of central tendency, measures of dispersion, and quartile ranges [percentiles] etc. Furthermore, the platform was developed to support graphical representation of the distributions of the variables. Besides, the platform also enhances a feasible correlation background of variables, in order to determine the dependence between different variables.

Ultimately, the platform provides an overview of the results of the data analysed in a more concise manner for better understanding. The extraction processes discussed in the subsection 3.3.2.5 above, were instantaneously executed alongside the analysis of the data. As regards this process, all variables were analysed according to the data type. In some cases, class interval needs to be established in order to have a clear understanding into the meaning of the data assembled, while some analysis cases were based on the direct application of measures of dispersion and measures of central of tendency to show the gravity of the data distributions. In the case of class interval, the frequency of the data was obtained, also the histogram and the quartile ranges of the data were analysed to improve the clarity of the distributions.

However, the analysis result obtained demonstrates the distribution of variables as illustrated in analysis platform. These particular results exhibited the purposeful of the data distributed across each variable, and as well clarified the difference between the captured data and the non-captured data [uncaptured data] as regards the findings discovered on the quality of data collected by the reporting officers. The practical outcomes of the analysis are discussed in the following chapters in both tabular and graphical layouts.

3.3.2.7.2 Mathematical approach for the Average estimates of the non-captured data

The grouping of the non-captured data per variable were further evaluated by determining the average estimates of the non-captured data as grouped into three related factors. Statistically, the average estimate of the non-captured obtained in each field, is simply referred to as the outcomes of the *'ratio of the sum of non-captured data in each field per month, to the amount of non-captured data in each field* (Schmuller 2009)*'* [refer to the formula below].

$$\overline{\mathrm{X}} = \frac{\sum \mathrm{X}}{\mathrm{N}}$$ ....Equation 1 (Montgomery & Runger 2007; Schmuller 2009; Lane et al. 2013)

$\overline{\mathrm{X}}$ = Mean/average of the non-captured data per field.

$\sum \mathrm{X}$ = Sum of non-captured data per month.

N = Number of non-captured data summed up.

3.3.2.7.3 Numerical approach for the Histogram of the non-captured data

As part of the evaluation process performed on the non-captured data extracted, a histogram is considered as 'the graphical approach suitable for illustration of the frequency distributions of ranges of scores (Montgomery & Runger 2007; Lane et al. 2013; Lane 2015). Besides, the frequency distributions of non-captured data estimates acquired in all the data fields in each related factor are displayed in two separated tables in all the analysis sections discussed in this study.

However, the purpose of separating the eight variables/data fields into two separate tables, offers a feasible understanding into the frequency distributions of the non-captured data across all the data fields in each related factor. In addition, in most cases, the range difference between the maximum and minimum scores is much wider due to high variation in the rate at which data is missing or mishandled in each field; hence, it will be difficult to have a sensible class interval/bin for the computation of the frequency distributions (Montgomery & Runger 2007).

In the course of obtaining the appropriate class width for the formation of the class interval/bin, two different methods were considered for similar results. The first of the two is the *Sturgis' rule,* which is described as a feasible way of setting the class intervals *as close as possible* to Equation 2 below.

$$Q = 1 + Log_2[N]$$ ....Equation 2

that is, rounded to the nearest integer (Lane et al. 2013; Lane 2015); where $Log_2[N]$ is referred to as the *base 2 Log* of number of observations (Lane 2015). The above written formula can also be rewritten in another statistical format as illustrated in the Equation 3 below.

$Q = 1 + 3.3Log_{10}[N]$……………….Equation 3

where $Log_{10}[N]$ is the *Log base 10* of the number of observations and *Q* is referred to as the *number of class intervals/bins* in the histogram (Cimbala 2014; Lane 2015).

The second of the two is the *Rice rule,* which is statistically calculated to establish the number of interval *twice the cube root* of the number of observations (Cimbala 2014). Hence, the formula statistically generates the equation below.

$\mathrm{Q} = 2[N]^{\frac{1}{3}}$ ....Equation 4

'Q' was previously defined in *Sturgis' rule* and *N* is referred to as the *number of observations* (Cimbala 2014; Lane 2015). These two methods ensure the feasibility of achieving a better choice of class width to determine the number of class intervals/bins for an absolute frequency distribution (Lane 2015).

4. Analyses and findings in Errors and Accident related factors

This chapter discusses the errors detected through visual evaluation of the completed ARFs assessed from the STD. In an attempt to have a better understanding of the errors committed by the reporting officers, then the researcher develops an official classification of errors found in the completed ARF as presented in section 4.1. However, most errors committed are connected to the inaccuracy, incompleteness, inconsistency, and untimeliness existing along the RTI processing line.

Due to the anomalies discovered in the process, three forms of error are categorised in order to assist with the visual assessment procedure. The three forms of error are unit non-response error, item non-response error and response error. According to the classification of the three forms of error, however, all errors were successfully categorised accordingly.

With the aim of completing this chapter, the outcomes of the analytical breakdown performed on the usable data in each field are presented in the subsequent sections. The discussion commences with the interpretation of the registered number of accidents occurred in the Stellenbosch area in 2012, followed by the interpretation of Accident related factors comprising data fields, such as Accident date and Day of week, Time of accident, Accident type, Severity of injury and Summary of persons involved, coupled with analysis and findings from the two contributory factors present in the Appendix B such as, Light condition and Weather conditions and visibility. These data fields[factors] were discussed in order to have a practical view or an idea of the data analysed as related to real-world systems, by showing the practical level of the usable data points acquired through the means of data graphing and data grouping.

In addition, the interpretation process was based on the understanding of the data, process involved, and integration of the results into the real-world systems. In this case, therefore, the importance of both the captured data and non-captured data were considered as the only link towards a better understanding of the data quality problem encountered at the local level. The interpretation of the results will reveal the severity of the problems affecting the quality level of the data gathered from the reportage of RTA in the Stellenbosch area. This expression will establish the effect of the missing data on the practicality of usable data in the analysis of real-world problems in the RTA.

4.1 Practical exploration of errors

This part is important in the need to improve the quality of the road accident data collected. The chance of finding errors was permitted through the evaluation processes implemented in Chapter 3. Actually, this procedure provides some significant clues to the problems frustrating the completeness and correctness of the data collected from the Accident Response Unit

[ARU], where the accident records are warehoused in the Stellenbosch Traffic Department [STD].

Nonetheless, the visual exploration carried out here, was subject to a thorough screening of the information presented in the completed ARF, which serves as the data form for the coverage of accidents. The screening process extends to all sections in the data form, in order to pave the way for proper assessment of the significant factors considered in this research project. During this process, several errors were uncovered which were associated with the application of the ARF. In some cases, as regards these shortcomings, different types of errors detected connect to how comprehensible and logical the ARF appears to the users, or perhaps the degree of competency of the form users in using the form.

Collectively, errors determined through visual screening are classified into three forms of error, which are enumerated below as:

- Unit non-response error,
- Item non-response error, and
- Response error.

The first form of error, known as the ‘Unit non-response error’, is described thus as, “the failure to obtain an accident report for a particular case that satisfies the definition for inclusion, that is, one that meets the defined threshold level” (O’ Day 1993; CIHI 2009; Azimaee et al. 2014). The second form of error, ‘Item non-response error’ is defined as, “the failure to record any value [data] for a particular variable or field in the ARF, e.g. failure to write down the driver’s age” (O’ Day 1993; CIHI 2009; Azimaee et al. 2014). The last of the three forms of error, ‘Response error’ is referred to as “the recording or indication of incorrect value [data] for any variables, e.g. listing a person’s age as 20 instead of 30, or indicating a speed limit as 6km/hr instead of 60 km/hr” (O’ Day 1993; CIHI 2009; Azimaee et al. 2014). Among the three forms of error, only two forms are consistently observed during the assessment of the completed ARF, which are grouped under the ‘Item non-response error’ and ‘Response error’.

As earlier stated in the previous paragraph, the possibility of detecting these errors is accredited to the standard evaluation procedures followed, which facilitate a complete implementation of the data assessment processes such as *data accuracy, data correctness* and *data completeness.* The classification of the assessment processes constitutes the procedure applied in defining the errors. Fundamentally, data correctness was implemented to define the errors due to *‘Response error’;* while data completeness was implemented to define the errors due to *‘Item non-response error’* (O’ Day 1993; CIHI 2009; Azimaee et al. 2014). The integration of both the data correctness and data completeness assessment processes determines the degree of accuracy of the data assembled.

Table 5: Categories of errors discovered through visual evaluation

Exploration of errors detection	
Response errors	
• **Errors detected in *Light condition* and *Weather conditions and visibility,* are connected to where answers indicated/marked are not related to each other, such errors were discovered between the answers indicated as *'daylight'* and *'overcast'.***	• Errors detected in Speed limit on road, are connected to indication of the unspecified speed limit on the road such cases as 6km/hr.
• **Errors detected in *Traffic control type,* where the option like *'not a junction, crossing or barrier line'* is interpreted as unavailable, but in the segment for *Accident sketch* and *Accident description,* it was indicated that the accident occurred at no junction or barrier line.**	• Errors detected in *Vehicle manoeuvre/ what driver was doing* and *Traffic control type,* are connected to incorrect specification of answers of *'reverse'* and *'parking'.* These answers were written/indicated under *Traffic control type* instead of *Vehicle manoeuvre/what driver was doing?*
• **Errors detected in *Accident type* and *Number of vehicles involved*, are connected to where answers indicated are not related to each other, such errors are discovered between the answers indicated as *'sideswipe'* and indicated *number of vehicles involved*. The indicated option shows sideswipe, which mostly involved two vehicles at the junction section, but the number of vehicles indicated shows that it is a single-vehicle accident at the intersection.**	• Errors detected in *Number of vehicles involved and Accident type*, are connected to incorrect indication of answers as regards *accident type,* where option indicated is '*accident with pedestrian',* but under *Number of vehicles involved*, it is indicated that the accident involved two vehicles.
• **Errors detected in *Accident type, Accident sketch* and *Accident description,* are connected to mix up information between the indicated/marked answers such as 'accident with fixed/other object', 'accident drawing and accident statement', where it is illustrated in the accident sketch and description that the accident involved a pedestrian, but indicated in the Accident type as accident involving fixed object.**	• Errors detected in *Speed limit on road* and *Number of vehicles involved*, are connected to incorrect indication of *speed limit* in the option box created for the *number of vehicles involved*, and otherwise indicates number of vehicles involved in the option box created for the *speed limit.*
• **Errors detected in *Accident location,* are connected to misspelling of the street/road names.**	• Errors detected in *Accident date and Day of week,* are connected to incorrect completion of *incorrect date* for the *right day*, and the *incorrect day* for the *right date*.
Item non-response errors	
• **Omission cases were found in all sections in the completed ARF. Many answer options were left unmarked/emptied, which may be due to some circumstances preventing the ARF users from collecting or understanding some information represented in the form. Data omission can also be considered as missing data.**	

Errors caused by the inability of the form users to identify the right answers for the right information is connected to '*incorrect correlation of the accident findings with the answers specified in the data form*'. According to a statement made by the DCO, some of these errors are committed as a result of inexperience and less concentration from the users of the data form. Aside this, errors are also committed as a result of the reporting officers violating the instructions provided to guide them, as well as other conditions encountered along the process of data collection.

Moreover, errors explored here ranged from the '*indication of double answers*', wherein the form users violate the instruction provided to mark one option only; '*omission of valuable information*', wherein necessary information required in RTA are ignored; and *'incorrect indication of relevant answers'*, wherein the information provided on the data form are misinterpreted to the findings discovered at the accident scene. As a result of this, the errors are classified under non-captured data [uncaptured data], provided that the errors are not in any way rectified before the data is being processed [structured].

These errors are repeated along the process of data collection. Other categories of errors discovered are tabulated in the Table 5[11]. This table shows the collective arrangement of the errors uncovered through practical exploration processes. The DCO, who manages the capturing of veritable road accident data added that, some identified limitations as discussed in the section 9.2, hampered the chance of fetching sufficient and accurate information [data] from the completed ARFs. In addition, it is understood that the completed ARFs with these limitations are mostly the photocopies delivered to the STD from the SAPS. Besides, some other considerable factors like information composition, information misinterpretation, and insufficient information in the ARFs are perceived as part of the problems contributing to the accumulation of errors during the data collection process. The suggested improvements to these three problems will be discussed as part of the conclusion in this investigation.

4.2 Introduction to the analysis and findings in Accident related factors

This section concentrates on the interpretation of the analysis results carried out on the *'Accident related variables'* and the *'Contributory factor related variables'*, which caused or contributed to accidents. Furthermore, the interpretation of the analysis results demonstrates the practicality of the data collected, which could provide key insight into the factors contributing most to occurrence of road accidents in the Stellenbosch community. The principle behind this is to determine whether the data collected at the local level give a true reflection of the real-world system as regards the RTA. In addition, the results/findings discussed in this section are based on the true reflection of the data categorised under the data fields characterised in the data form. The data fields are characterised according to the criteria used in defining the classification of both the captured data and non-captured data. In this regard, the captured data[12] are regarded as *'the documented data considered fit for analysis purpose'* (Wang & Strong 1996; Lee et al. 2002; Wang 2004; Wang & Strong 2013).

The analysis of the fit and unfit data in the Accident related factors are discussed graphically in order to have a more practical idea about the problems affecting the quality of road accident data. The number of registered road accidents reported at the STD throughout the year 2012 will be discussed first in this section. On the other hand, concerning the other related factors,

---

[11] The concept used in classifying the forms of error are adapted from James O' Day (1993), CIHI (2009), and Azimaee et al (2014).

[12] Refer to section 3.3.1 for the proper definitions of both captured data and non-captured data.

the data points gathered are analysed in order to demonstrate the effect or influence of each field on the three key factors which are *the road user, the vehicle* and *the road environment.*

4.2.1 Analysis of the Registered road accidents

Registered accidents are approved as the number of reported RTA over a period of time, which are documented by the authorised local departments responsible for the road traffic matters. Additionally, *'Registered accidents*' are evaluated for accuracy, validated for completeness and acknowledged for documentation at the local level [STD] or provincial level before the collated RTI could be transferred to RTMC for further evaluation. Despite the assessment processes executed, irregularities still frustrate the integrity of the data processed. However, with the objective of clarifying the integrity of the data collected and analysed, a benchmark initiative was carried out in comparing the data obtained in this study with the reported road accident data of the City of Cape Town Metropolitan Municipality (Traffic Accident Statistics 2005).

Table 6 illustrates the total number of road accidents that occurred in both SM and the City of Cape Town Metropolitan Municipality. In this table, estimated sum of 2,451 road accidents were registered in SM in 2012, compared to that of City of Cape Town Metropolitan Municipality with a huge total estimate of 85,398 road accidents in 2005 (Traffic Accident Statistics 2005). The large number of road accidents in the City of Cape Town Metropolitan Municipality every month compared to SM, is logically due to Cape Town area being much larger and having many more activities day to day. Both previous factors increase the rate at which road accidents occur in the city. Cape Town has more activities as large vehicular movements, more daily business transactions, a higher population rate, and more infrastructures. However, the total amount of road accidents realised in SM considering the population size of over 250,000 inhabitants (Stellenbosch Municipality 2012), within the aforesaid year reflect a huge number of RTAs resulting from different accident types.

In the table, according to the scores arrayed in frequency and percentage rates, monthly accident occurrence rates in both municipalities indicate that May was considered as the worst month with highest occurrence score of 7,525 [8.8%] road accidents in the City of Cape Town Metropolitan Municipality in 2005 (Traffic Accident Statistics 2005), unlike in SM with highest occurrence score of 253 [10.3%] road accidents in March as the worst month for road accidents occurrence in 2012. Similar observations show that the fairest month with lowest occurrence of road accidents in the City of Cape Town Metropolitan Municipality was January with 6,340 [7.4%] (Traffic Accident Statistics 2005), while December was observed as the fairest month with lowest occurrence of 143 [5.8%] road accidents in SM.

With huge number of accidents registered in March in Stellenbosch locality, more errors are likely to be committed in the data fields covered. In addition, the outcomes obtained from the analysis of non-captured data in each field demonstrate that some data fields are omitted, while some are completely misrepresented and misinterpreted during the data collection actions. Besides, the average estimate of road accidents per month in SM in 2012 is 204, which is approximately 34 times smaller to the average estimate of 7,117 road accidents per month in the City of Cape Town Metropolitan Municipality (Traffic Accident Statistics 2005).

Table 6: Comparison of Registered road accidents per month between the Stellenbosch Municipality and the City of Cape Town Metropolitan Municipality

Number of Registered road accidents per month				
Month	**Stellenbosch Municipality in 2012**		**City of Cape Town Municipality in 2005**[13]	
	Frequency [in hundreds]	**Percentage estimates [%]**	**Frequency [in thousands]**	**Percentage estimates [%]**
Jan	161	6.6%	6340	7.4%
Feb	236	9.6%	6900	8.1%
Mar	253	10.3%	7075	8.3%
Apr	186	7.6%	7494	8.8%
May	230	9.4%	7525	8.8%
Jun	182	7.4%	7467	8.7%
Jul	167	6.8%	6960	8.2%
Aug	233	9.5%	7509	8.8%
Sep	201	8.2%	6803	8.0%
Oct	238	9.7%	7301	8.5%
Nov	221	9.0%	7120	8.3%
Dec	143	5.8%	6904	8.1%

The tabular array of scores under the SM in the Table 6 is demonstrated in the chart presented in the Figure 16. The scores confirmed that the most occurrence periods of RTA in the Stellenbosch area fall between periods of busy activities such as academic and non-holiday periods[14], which perhaps amplified the population size of the road users and number of vehicles travelling daily on the Stellenbosch roads; whereas in the City of Cape Town Metropolitan Municipality, wet and winter are considered as the consequence of bad weather and visibility conditions causing poor visibility and wet roads across the city as affirmed by the Traffic Accident Statistics (2002; 2005). The chart demonstrates the trends in the road accidents occurrence per month in SM.

According to the variations observed in the chart, the number of road accidents varies from one month to another. The number of accidents reported rise and fall sequentially within the

[13] Data extracted from a report published by the Traffic Accident Statistics (2005) on Registered road traffic accidents in the City of Cape Town Metropolitan Municipality in 2005.

[14] Refer to the subsequent section for further analysis on the period that accident occurred most in the Stellenbosch community.

period of two to four months in 2012. From the period of January to April, the trend observed shows an upsurge in February and shifts slightly towards March, which later falls in April. Similarly, from the period of May to August, the pattern observed appears to be convex wherein the rate at which accidents are registered within this period falls dramatically towards July and a distinct increase is observed in August. In addition, a corresponding trend was observed along the period of September to December. The pattern illustrates another increment from September to October and a reduction from November to December.

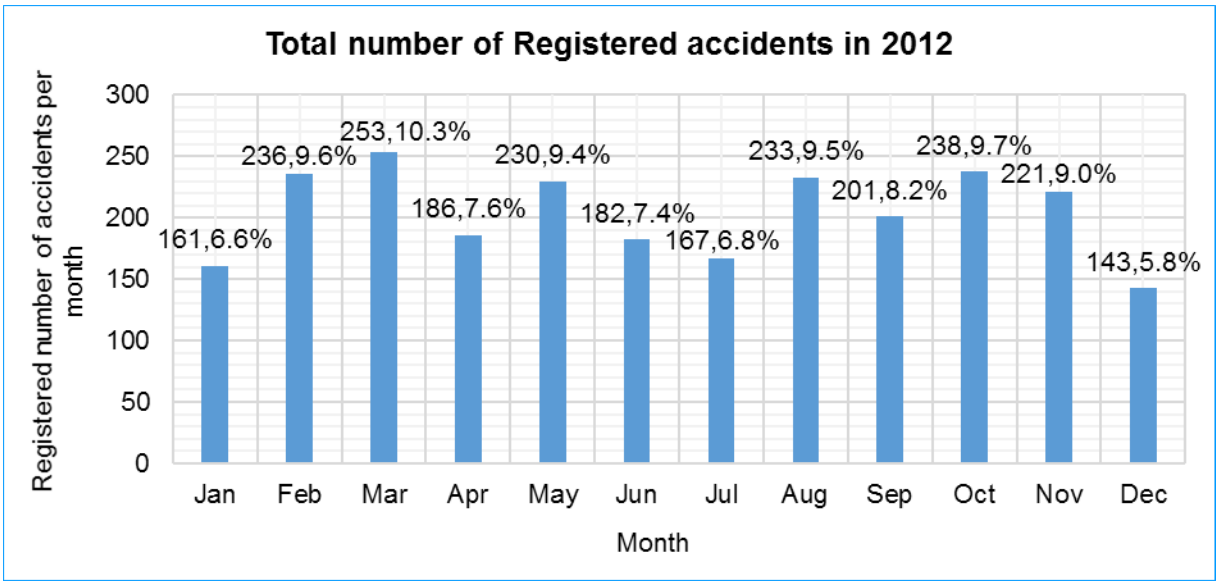

Figure 16: Total estimates of the Registered road accidents in 2012

Theoretically, observations carried out reveal that some determinable factors influenced the monthly occurrence of road accidents in SM. The factors are enumerated below as:

- Periodic increase in the population size of road users residing within or outskirt of Stellenbosch area, such as drivers, cyclists, and pedestrians (O' Day 1993).
- Periodic increase in the number of vehicles; attributed to the increase in the number of residents of Stellenbosch (O' Day 1993).
- Intensity and tenacity of heavy traffic flow across the Stellenbosch's main roads; which could be accredited primarily to the impatience of the road users (Roux & Sinclair 2003; Jungu-Omara & Vanderschuren 2006; Sinclair & Murdoch 2012)
- Lenient law enforcement protocols exerted by the law enforcement agencies including the police officers(O' Day 1993).
- Inability of any accident victim in reporting his involvement in an accident to the nearest local authorities on the exact day, date and time of occurrence; which could be attributed to factors such as lack of insurance welfares, hit and run cases, unregistered vehicles, and many more (O' Day 1993).

4.2.2 Analysis of the Accident date and Day of week

This section covers the analysis performed on both the *Accident date* and *Day of week* as part of the data fields or variables provided in the data form. The presentation of the results obtained from these two data fields is virtually related, in accordance with the relevance of the data interpretation in the real-world problem. From my personal perception, *Accident date* can be referred to as '*a specific date or day of the month an accident occurs, wherein a prompt reporting of the incident to the nearest authorised local department is mandatory'*, while the *Day of week* can be referred to as *'a definite or specific day in a week that an accident occurs as mandated to be reported'*. Table 7 displays the estimated scores pertaining to the analysis of the actual dates or days of the months when road accidents occurred.

The estimated scores demonstrate the number of the times road accidents occurred on each date or day of the month allocated to a particular weekday in a month. Furthermore, the few *'zeros*' *[specify in yellow colour]* among the scores in the table indicate that no accidents were reported on the specified dates, while the empty cells indicate that some months are not up to 31 days. However, the variations observed in the scores arrayed in the Table 7 imply that some scores are ≥10 accidents per date/day of the month, while other scores are <10 accidents per date/day of the month. Since the dates or days of the months characterise the weekdays in 12 months, therefore, a simple illustration was established to identify the weekdays with the most occurrences of accidents, by considering the scores that fall within the range of scores ≥10 accidents per date/day of the month as a suitable sample. The scores in this range are validated, in order to practically determine the specific factors or conditions that contribute to the cause of accidents on these weekdays.

From the Table 7, it is observed that the two largest amount of accidents per date/day of the month were deduced on 11th of May with a total amount of 21 [0.87%] accidents per date/day of the month, and on 25th of May with a total 22 [0.91%] accidents per date/day of the month. According to the dates or days of the months' arrangement in the 2012 calendars, the two aforesaid dates or days of the months were allocated to two Fridays in May, which is probably due to the beginning of the winter period. Preliminarily, this implies that more accidents occurred on Friday than any other day in the weekday. In addition, to support the initial findings provided above, an analysis of the dates or days of the months allocated to weekdays as per scores that fall within the range of scores ≥10 accidents per date/day of the month was carried out. The selected dates or days of the months correspond to the scores that fall within the specified range of scores ≥10 accidents per date/day of the month, and were validated in the 2012 calendars to determine the days that contributed most to the amount of the times accidents occurred in the same year. Based on the outcome obtained from the validation,

Table 7: Accident date estimates in road accident occurrence

Total estimates of Accident date in road accident occurrence in 2012													
Date/Day	**Jan**	**Feb**	**Mar**	**Apr**	**May**	**Jun**	**Jul**	**Aug**	**Sep**	**Oct**	**Nov**	**Dec**	**Total scores**
1st	2	10	9	8	2	8	3	10	14	7	6	8	**87**
2nd	0	12	9	5	1	6	0	7	7	7	7	6	**67**
3rd	0	8	9	3	19	4	4	10	3	5	10	4	**79**
4th	2	7	7	6	10	7	6	6	7	12	2	7	**79**
5th	4	5	7	1	3	10	2	2	6	6	6	10	**62**
6th	3	12	6	1	2	10	5	12	5	10	5	10	**81**
7^{h}	5	6	16	5	5	5	10	4	13	6	12	10	**97**
8th	5	6	5	2	11	10	5	11	9	6	10	8	**88**
9th	3	14	11	6	9	7	3	3	2	8	8	4	**78**
10th	2	10	16	6	6	2	4	5	4	10	18	6	**89**
11th	3	1	4	6	21	4	3	6	6	8	3	7	**72**
12th	3	10	8	10	8	6	2	4	5	12	8	8	**84**
13th	7	9	3	7	4	7	4	5	5	6	5	4	**66**
14th	5	9	8	8	4	2	5	9	5	5	5	8	**73**
15th	4	7	10	2	5	5	10	8	9	8	8	4	**80**
16th	1	7	12	7	2	4	2	8	5	7	13	1	**69**
17th	6	12	6	9	3	5	4	9	7	8	2	5	**76**
18th	10	12	5	5	6	5	6	12	9	10	1	3	**84**
19th	2	3	11	10	8	6	9	7	12	17	10	0	**95**
20th	3	8	11	12	1	7	12	4	9	6	9	6	**88**
21st	1	7	5	2	9	5	6	6	8	5	8	5	**67**
22nd	2	8	5	2	9	8	9	8	5	5	7	2	**70**
23rd	15	5	4	10	4	8	2	4	5	10	5	2	**74**
24th	9	3	4	11	6	2	2	11	1	6	6	3	**64**
25th	8	8	6	3	22	4	8	12	3	6	3	1	**84**
26th	12	5	8	14	10	3	2	5	6	6	6	2	**79**
27th	5	4	5	6	4	4	7	6	11	11	12	1	**76**
28th	9	11	8	7	6	6	9	8	7	3	14	3	**91**
29th	7	10	15	4	7	7	7	7	4	7	6	3	**84**
30th	9		11	5	8	9	3	8	4	5	6	1	**69**
31st	13		7		9		9	9		7		5	**59**
Total scores	**160**	**229**	**251**	**183**	**224**	**176**	**163**	**226**	**196**	**235**	**221**	**147**	

Friday tops the list of the days with the higher scores falling within the range of scores ≥10 accidents per date/day of the month [refer to Table 8 below].

Table 8 displays the number of times each weekday occurs within the range of scores ≥10 accidents per date/day of the month, accordant with the observations performed by using 2012 calendar as benchmark. From Table 8, it is interpreted that Friday has the highest frequency score of 17 times for a certain number of RTAs in 2012. Correspondingly Thursday, Wednesday, and Saturday are among the weekdays that occurred more frequently with frequency scores of 12 times, 14 times and 11 times respectively for a certain number of RTAs. During this process, the public holiday periods were also observed to ascertain if the dates/days with scores ≥10 accidents fall among the dates/days for the holiday periods [indicated in green colour] analysed in the Table 7. Apparently, according to the validation outcome performed on the holiday periods observed in SM, only 12th of April *'Family Day'* recorded the number of road accidents [≥10 accidents], unlike in the City of Cape Town Metropolitan Municipality where 22nd of March *'Public Holiday'* recorded the highest number of road accidents [258 accidents][15] (Traffic Accident Statistics 2005). This illustrates that large number of holiday periods observed in Stellenbosch contributed minimally to the number of road accidents per dates/days.

From the technical point of view, the data points obtained from the analysis result of the Accident date can also be utilised to illustrate the exact dates/days that accidents are frequently reported to the STD. In addition, the results illustrated in the Table 8, can be utilised by the division in charge of road traffic matters to understand the rate at which accidents occur per weekday, and the prospect of being able to actualise the exact number of officers required to carry out a prompt control of any situations that may lead to occurrence of RTAs in the weekdays-with most occurrences of RTA. The general conclusion on this illustration will be supported with the results acquired through the analysis of the *Day of week* in the subsequent paragraphs.

With the purpose of computing the expected number of road accidents per date; thus, an average estimate of the scores presented in the Table 7 was statistically obtained. It is indicated that an average estimate of 6 road accidents occurred per date/day with an estimate of 0.27%, which is three times smaller than the estimate of 0.91% actualised from an estimated score of 22 road accidents recorded on a particular date/day in May 2012.

[15] See the benchmark table for Road Accidents during Holiday Periods in the Appendix B.3.

Table 8: The analysis of the Accident date with scores ≥10 accidents per date

Weekday with scores ≥10 accidents per date	Number of frequencies for weekday with scores ≥10 accidents per date
Sunday	1
Monday	6
Tuesday	8
Wednesday	12
Thursday	14
Friday	17
Saturday	11

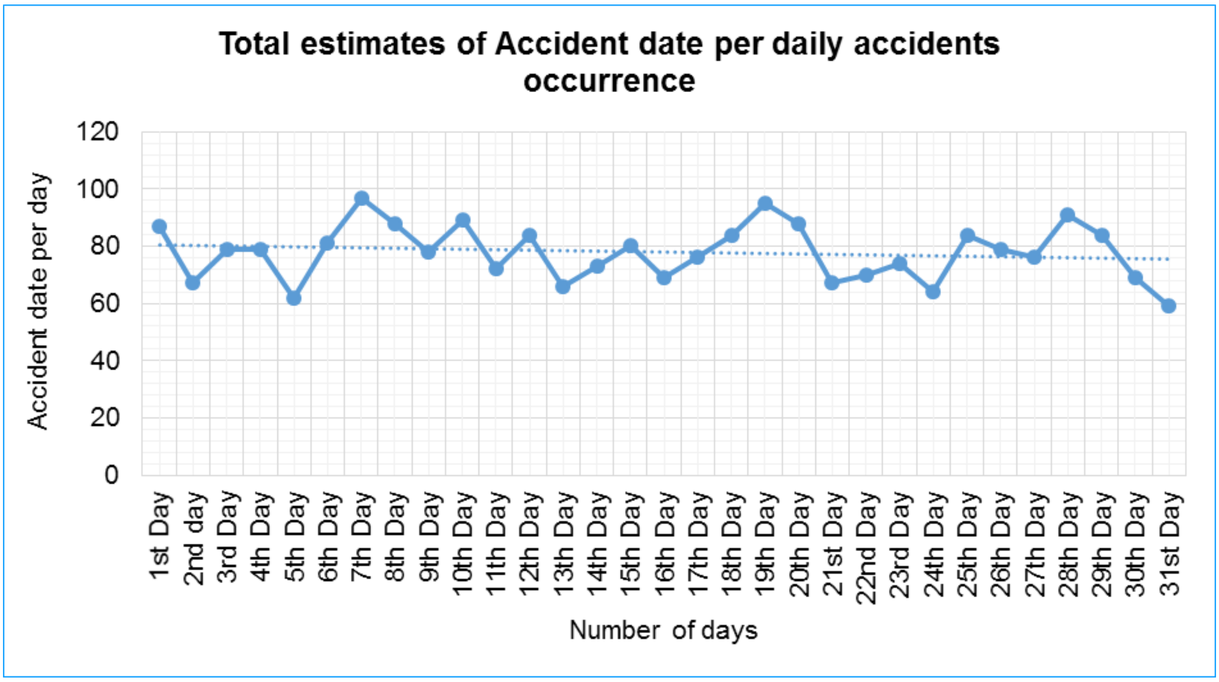

Figure 17: Total estimates of the Accident date per day

Similarly, a monthly aggregate of the data points assembled in the Table 7 was analysed to determine the months with most observations of road accidents occurrence. The interpretation of the observations observed yields a highest estimate of 251 [10.4%] and a lowest estimate of 147 [6.1%] accidents throughout the 31 days in March and December respectively. The small amount realised in December indicates that long public holiday periods were observed [see Table 7], which demonstrates a drastic reduction in the number of road users and traffic congestions experienced during a no-holiday period within the Stellenbosch locality.

Conversely, an average estimate of 201 [8.3%] accidents occurred within the 31 days of every month according to the result calculated in Accident date. Thus, this percentage estimate is equal to the average estimate obtained for the number of registered accidents illustrated in subsection 4.2.1 above, but the total estimates for each month are not equal but close in ranges. This illustrates that some data were rendered impractical or void due to the presence of anomalies during the assessment process performed on the raw data, such as missing data, incorrect dates[16], and many others.

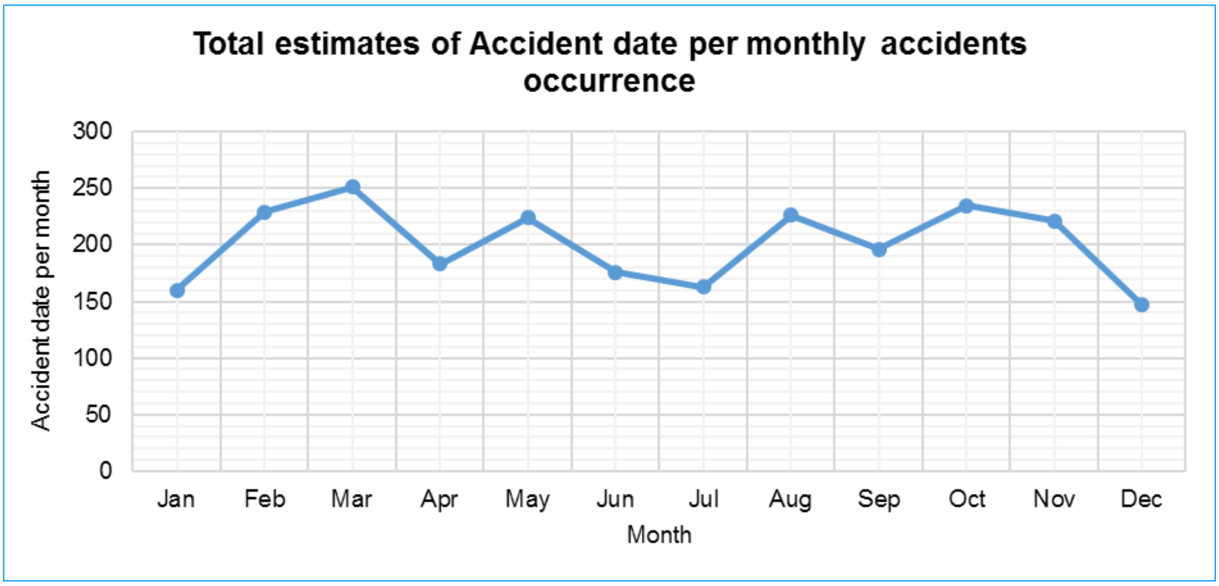

Figure 18: Total estimates of the Accident date per month

Two charts are plotted from the results of the analysis performed on the data points displayed in the Table 7. The first chart displayed in Figure 17 represents the daily illustration of the *Accident date* calculated, while the second chart in Figure 18 represents the monthly illustration of the same field. The trends shown in the two charts are not similar in structure, but both charts offer practical understanding towards the purpose of describing the usefulness of this data field in the analysis of the RTA, by specifying the rate at which accidents occurred on each date in a month to validate the actual weekday and date with most occurrence of RTA. Besides, the trends display in the charts produced a sensible distribution of the data calculated for both daily and monthly illustrations of Accident dates. Actually, these illustrations could assist the traffic management to understand the practical contribution of the Accident dates towards an improved understanding of the accident occurrence rate within the Stellenbosch locality.

In the first chart [see Figure 17], a daily illustration of the road accidents was demonstrated. The trend line drawn across the chart signifies that the scores are very close in range, which implies that fewer errors are committed occasionally by the ARF users. The trend line also shows the exact range where the average estimate falls. In this case, the average estimate indicates how consistently the accident dates are being represented in the ARF by the reporting officers.

---

[16] Refer to section 3.3.2.7.2 for further details.

By observing the chart, the points above the average estimate mostly constitute the data points gathered from the weekdays, such as Wednesdays, Thursdays, Fridays, and Saturdays according to the practical observations performed on the 2012 calendar year, while the points below the trend line are from the data points gathered mostly on Sundays, Mondays and Tuesdays. Among the weekdays with the scores below the trend line, thus only Tuesdays recorded high figures closer to the average. This indicates that in some years to come, Tuesdays could be among the weekdays with high records of accidents, if drastic measures are not appropriately implemented. This part will be elucidated further in the *'Day of week'* to support this illustration.

However, if the data points indicated in the line chart displayed in the Figure 18 is converted to bar chart; thus, a similar trend corresponding to the structure demonstrated in the chart displayed in the Figure 16[17] could be obtained. Absolutely, if the suggested illustration is possibly carried out, the chart graphically appears like the chart demonstrated in the Figure 16. Contrarily, the scores presented in the Table 7 are slightly smaller compared with the scores presented in the Table 6 in the subsection 4.2.1 above, due to the influence of missing data in the Accident date. The structure of the trend shows that March has the highest score, with a score slightly above 250, and December appears the lowest with a score slightly below 150.

On the other hand, considering the *Day of week*, the table set below appears simpler than the Table 7 displayed earlier. Basically, because the variables considered here are weekdays only, which are concisely demonstrated towards the determination of the weekday that recorded most occurrence of RTA in the Stellenbosch area, as similarly demonstrated in the *Accident date.* In addition, the data points gathered from this field can also be utilised to illustrate how accurately the weekdays are represented in the ARF in accordance with the respective dates/days that accidents occurred.

The purpose of analysing the data points assembled in the Table 9, is to benchmark the analysis presented in the *Accident date.* In the table, the weekdays are tabulated in rows, with the intention of having a compact structure or distribution of the data points. On the other hand, while the months are tabulated in columns, with the intention of having a clear overview of the data points over the year. As a result of the structure, the distribution of the data points presented a range of data points between 7 to 58 road accidents per weekdays.

The distribution demonstrates a great number of frequency observations of data points within the range of ≥30 mostly from Wednesday to Saturday. Primarily, the distribution of the data points illustrates the total count of weekdays represented in the completed AR form. This

[17] Refer to subsection 4.2.1 for more clear observations.

illustration can also offer an understanding into the rate at which accidents occur depending on the degree of completeness and correctness of representing the weekday in the data form.

Table 9: Day of week estimates in road accident occurrence

Days	Total estimates of Day of week in road accident occurrence in 2012												Total scores	Percentage estimates
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec		
Sun	18	22	21	15	10	14	29	17	19	18	7	12	**202**	**8.9%**
Mon	26	33	28	31	24	22	11	31	20	33	26	22	**307**	**13.5%**
Tue	30	29	26	24	32	22	18	24	20	39	26	15	**305**	**13.4%**
Wed	22	38	39	20	26	25	22	39	25	33	36	17	**342**	**15.1%**
Thu	18	38	44	31	41	16	16	24	28	36	34	19	**345**	**15.2%**
Fri	18	30	44	25	58	32	30	41	35	35	36	23	**407**	**17.9%**
Sat	18	29	43	26	28	34	30	34	36	31	31	22	**362**	**15.9%**
Total scores	**150**	**219**	**245**	**172**	**219**	**165**	**156**	**210**	**183**	**225**	**196**	**130**		

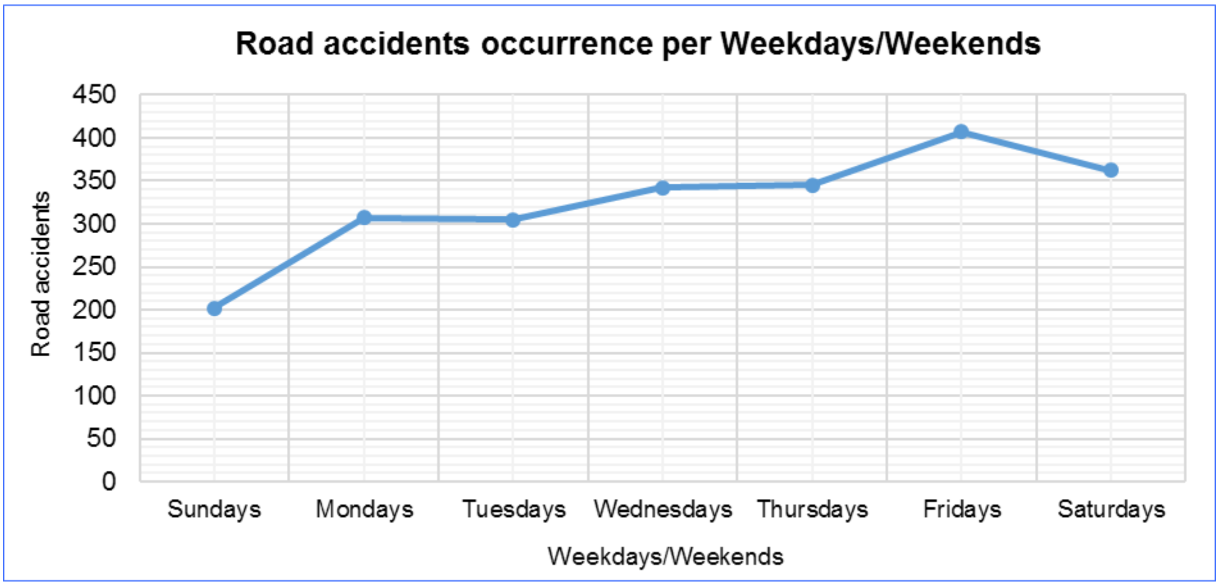

Figure 19: Total estimates of road accidents per Weekdays/Weekends

In this context, among the four weekdays with high frequency observations within the range of ≥30, only Fridays have the highest total score of 407 [17.9%] of road accidents followed by Saturdays and Thursdays with total scores of 362 [15.9%] and 345 [15.2%], compared with the findings obtained from the analysis carried out in the City of Cape Town Metropolitan Municipality, where Fridays were considered the worst weekdays with a total score of 14,965 [17.52%] of road accidents followed by Saturdays and Mondays with total scores of 12,726

[14.90%] and 12,494 [14.63%][18] (Traffic Accident Statistics 2005). This illustration implies that in Stellenbosch Municipality, RTA occurred mostly towards the weekends since many activities are engaged by the road users in these particular periods, such as drink-driving, substance abuse, traffic law violation, over speeding, and many others. Nonetheless, in the case of the City of Cape Town Metropolitan Municipality, similar activities are engaged but on Mondays high traffic congestion rate is a probable issue around the City of Cape Town Metropolitan Municipality as specified in the report by the Traffic Accident Statistics (2005).

From the same table, in the Stellenbosch area, Sundays, among the weekdays have the lowest total score of 202 [8.9%], with a percentage estimate marginally smaller than the half of the percentage estimate obtained in Fridays. Similarly, in the City of Cape Town Metropolitan Municipality, Sundays contributed the least road accidents with a total score of 8,854 [10.39%] (Traffic Accident Statistics 2005). This illustration demonstrates dramatic reduction in RTA occurrence rate on Sundays in both areas compared to other weekdays, which perhaps could be ascribed to declining involvement in the influential factors such as road traffic congestion, substance abuse, traffic law violation, and many other factors. In addition, considering the SM, this particular analysis justified the relevance of accident occurrence per dates/days and weekdays in the analysis of RTAs that occurred within the area.

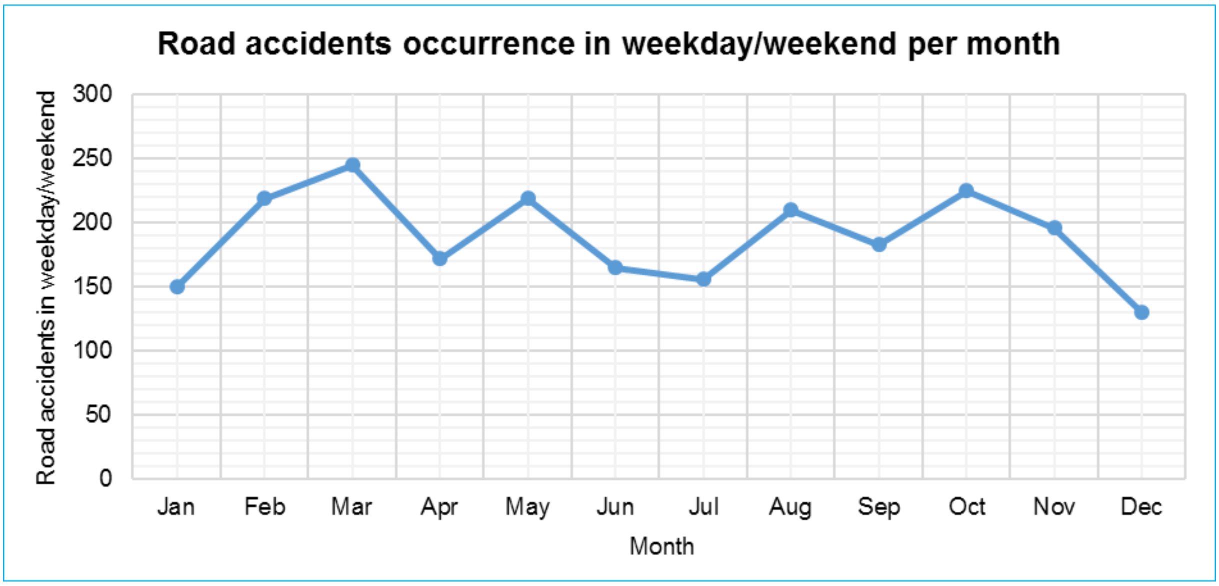

Figure 20: Total estimates of Day of week per month

In the Figure 19, an increasing trend from Mondays to Fridays is observed, which explains how busy the roads could be during these periods. Predominantly, in these periods, many road users do not exercise patience when it comes to heavy traffic matters. The shape of chart in

[18] See the benchmark table for Road Accidents by Weekdays between the two areas [regions] in the Appendix B.3.

the Figure 20 is similar to the shape of the chart presented earlier in Figure 18, but the magnitude of the scores is apparently not equal at all. It is observed that the Accident date is more represented in the data form than the Day of week.

It can also be concluded from the shape of the two charts that data is collected simultaneously, despite errors being made. This creates a gap between the degree of data sufficiency and the data capturing in actuating a reliable decision-making process. Whenever there is insufficient data supply, the possibility of attaining improvement is limited, and sometimes considered impracticable. This problem could be attributed to the attitude of the accident victim and the reporting officers in the acquisition and provision of the right information. For instance, situation wherein the accident victim [driver/cyclist] could not remember the right date and weekday that the accident actually occurs, and also the insensibility of the reporting officer by incorrectly representing or forgetting to represent the information in the ARF.

4.2.3 Analysis of the Time of accident

This section describes the result of the analysis performed basically on the exact times or periods when RTAs occur most in the Stellenbosch Municipality. This analysis can be utilised to understand or interpret the peak periods when traffics are most intensifying. And also, it can help the traffic management to identify the appropriate way of distributing or managing the resources proposed for road safety programmes to stimulate safety of lives and properties. In essence, this can be implemented to advocate for dividend distribution of resources in terms of safety tools [equipment], deployment of sufficient amount of safety officers, and allocation of traffic patrol team to ensure safety of the residents and their properties across the Stellenbosch Municipality [SM].

The table set below displays the data points obtained from the analysis of the actual time an accident occurred in the Stellenbosch area in the same year. The times spread across 12 months, with the time-interval of 1-hour difference in 24 hours, in order to accommodate all the data points acquired in the analysis. From the Table 10, the actual time-intervals that accidents occurred most fall within the time-intervals of 7:00:00 am to 6:59:00 pm [approximately 12 hours], while the actual time-intervals that accidents occurred less fall within the time-interval of 7:00:00 pm to 6:59:00 am [approximately 12 hours] according to the results obtained. The frequency distribution displayed in the Table 10 demonstrate that time-intervals with total scores above 100 accidents fall within the time-interval range of 7:00:00 am to 6:59:00 pm, which cover 12 hours out of the 24 hours a day. Similar illustration confirmed that time-intervals with total scores below 100 accidents fall within the time-interval range of 7:00:00 pm to 6:59:00 am, which covers the remaining 12 hours.

Table 10: Time of accident estimates in road accident occurrence

Time-Interval	Total estimate of Time of accident in road accident occurrence in 2012												Total scores
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	
00:59:00 AM	2	5	6	5	3	6	4	5	4	2	3	1	**46**
1:59:00 AM	1	4	5	1	3	2	3	1	1	3	2	1	**27**
2:59:00 AM	1	1	4	8	5	4	12	1	2	1	0	2	**41**
3:59:00 AM	3	5	5	2	2	2	4	5	1	2	1	1	**33**
4:59:00 AM	2	0	1	1	3	3	0	5	1	4	3	2	**25**
5:59:00 AM	1	1	1	0	2	2	1	1	3	2	2	0	**16**
6:59:00 AM	4	1	2	2	9	3	9	5	1	4	7	2	**49**
7:59:00 AM	11	15	16	13	15	19	10	19	14	10	13	6	**161**
8:59:00 AM	7	14	13	11	16	10	10	20	20	14	10	9	**154**
9:59:00 AM	7	10	9	5	6	8	7	7	17	11	15	10	**112**
10:59:00 AM	10	11	11	8	11	10	6	9	16	13	14	9	**128**
11:59:00 AM	6	16	22	17	10	11	9	10	15	14	11	12	**153**
12:59:00 PM	12	12	27	7	14	13	6	12	10	19	15	9	**156**
1:59:00 PM	12	14	13	11	13	16	10	21	19	13	14	13	**169**
2:59:00 PM	12	8	12	17	19	13	14	15	15	20	11	10	**166**
3:59:00 PM	7	20	18	10	17	8	9	10	9	23	12	7	**150**
4:59:00 PM	15	27	21	16	16	9	13	27	10	24	22	9	**209**
5:59:00 PM	19	15	24	17	22	7	11	18	12	16	28	6	**195**
6:59:00 PM	6	18	11	8	11	7	7	14	10	14	11	7	**124**
7:59:00 PM	2	10	7	5	13	11	4	8	4	8	5	4	**81**
8:59:00 PM	4	9	8	7	7	5	6	2	3	7	4	7	**69**
9:59:00 PM	6	8	3	6	0	4	3	9	5	2	3	4	**53**
10:59:00 PM	2	3	6	5	9	3	7	4	4	6	8	5	**62**
11:59:00 PM	6	5	6	3	3	2	0	3	3	3	3	5	**42**

From a more simplify way, according to the data points assembled in the table, the lowest accident record times fall within the time-interval of one hour [5:59:00 am] with a total record of 16 accidents, which explains that road accident occurred 16 times within the specified period. And also, the highest accident record times fall within the time-interval of one hour [4:59:00 pm] with a total record of 209 accidents, which likewise explains that accident occurred 209 times within the specified period; thus, this actual time-interval is considered as the peak period over the 12 months in 2012. Additional finding discovered from the Table 10 is the actual periods with none record of road accident occurrence. In the table, the fewer *'zero'* data points [indicated in green colour] suggest that no accident occurs in some specified periods.

Moreover, the actual time-interval with no accident occurrence are categorised mostly among the time-intervals with less accidents occurrence. By observing the data points, it is deduced that the *zero accident occurrence periods* are found within the time-intervals of 2:59:00am, 4:59:00am, 5:59:00am, 9:59:00pm and 11:59:00pm. These time-intervals are observed within the late-night periods, under February, April, May, July, November and December respectively. In the chart displayed in Figure 21, it is deduced that fewer accidents were recorded or reported in the midnight, but in the daylight, more accidents were recorded or reported.

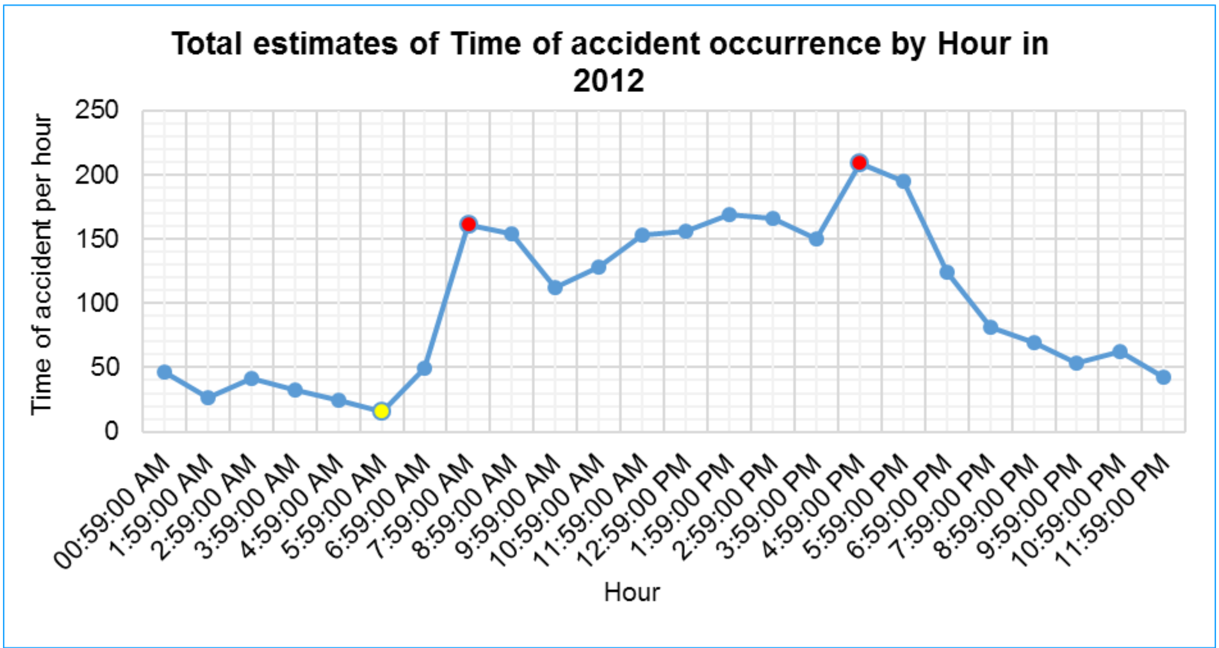

Figure 21: Total estimates of the Time of accident by hour in 2012

The distribution of the data points shows some variations around the time-intervals as displayed in the chart above. The variations demonstrate several periods of accident occurrence. Specifically, the result displayed in the chart illustrates three major points indicating the peak periods and the most extreme period for the lowest road occurrence within the Stellenbosch area. However, the two red spots are the two major peaks in the chart. The red spot to the left side of the chart indicates the peak period attained in the *dawn time* [7:00:00

am to 7:59:00 am]. This is the exact period when high number of residents and workers travel to their various destinations in SM (Sinclair & Murdoch 2012), mostly during the workdays. However, similar result is observed in the analysis illustrated in the City of Cape Town Metropolitan Municipality, where most road accidents also occurred in the workdays between 7:00 am to 8:00 am (Traffic Accident Statistics 2005). These illustrations signify the morning peak periods in both the SM and the City of Cape Town Metropolitan Municipality [refer to the benchmark table in the Appendix B.3].

In addition, the red dot to the right side of the chart indicates that another peak period was attained in the *dusk time* [4:00:00 pm to 4:59:00 pm]. This is the exact period when residents return to their various homes after much work done in SM (Sinclair & Murdoch 2012), more often during the workdays. On the contrary, in the case of the City of Cape Town Metropolitan Municipality, it is illustrated that high number of road accidents were recorded between 5:00 pm to 6:00 pm during the workdays, at the exact period when most workers and residents were returning to their various home (Traffic Accident Statistics 2005). These illustrations signify the evening peak periods in both municipalities as shown in the Table 49 in Appendix B.3. Besides, the yellow spot indicates the period with fewer accident occurrence. The spot falls among the periods with lowest number of accidents occurrence attained.

4.2.4 Analysis of the Accident type

This section discusses the type of road accidents occurred in the Stellenbosch locality in accordance with the true reflection of the data assembled in the Table 11 below. The field consist of 11 variables defined according to their designated purposes. The analysis result obtained here could be applied to measure some factors relating to the relevance of the data fields in the *Accident related factors*, such factors as:

- Key factors leading to accident occurrence-relevant to the insurance companies, traffic engineers, and traffic management agencies,
- Conformation of the accident locations-relevant to the traffic engineers and road construction management agencies, and
- Determination of robustness of the vehicles manufactured-relevant to the vehicle manufacturers.

To virtually achieve the above-mentioned practical applications, thus, the level of viability of the data gathered necessitates the implementation of a standard dimension to avert incorrect and missing data during data collection actions. This particular dimension was implemented to produce the data points in the table below.

Table 11: Accident type estimates in road accident occurrence

Total estimates of Accident type in road accident occurrence in 2012											
Month	**Head/ Rear End**	**Head On**	**Sideswipe [opposite directions]**	**Sideswipe [same direction]**	**Approach at Angle [both travelling straight]**	**Single Vehicle [left the road]**	**Single Vehicle *[overturned]***	**Accident with Pedestrian**	**Accident with Animal**	**Accident with Train**	**Accident with Fixed/Other Object**
Jan	49	3	18	17	3	4	3	7	3	0	15
Feb	84	2	29	33	15	4	1	10	4	0	16
Mar	59	3	33	29	20	5	5	6	5	0	11
Apr	54	3	25	21	11	5	3	7	2	0	8
May	73	0	28	27	19	4	2	9	5	0	16
Jun	56	0	18	14	6	5	3	9	7	0	17
Jul	50	4	21	13	9	7	1	7	3	0	15
Aug	77	3	21	27	13	3	3	2	4	1	11
Sep	45	3	28	25	14	4	5	6	0	0	11
Oct	70	3	36	21	20	2	3	8	3	0	15
Nov	60	2	27	24	15	6	3	7	2	0	13
Dec	33	1	15	20	3	4	4	8	2	0	7
Total scores	**710**	**27**	**299**	**271**	**148**	**53**	**36**	**86**	**40**	**1**	**155**

The table contains the data points assembled for the 11 categorical variables in the *Accident type*. The relevance of the 11 categorical variables in realising the significance of Accident type in the analysis of road accidents is demonstrated, through the appropriate interpretation of the data points assembled in the table above. The analysis result obtained is presented in a graphical chart in the Figure 22, which offers a better understanding into the distribution structure of the data points. The distribution of data points shows that few types of accident occurred some distances away from intersection [junction] areas.

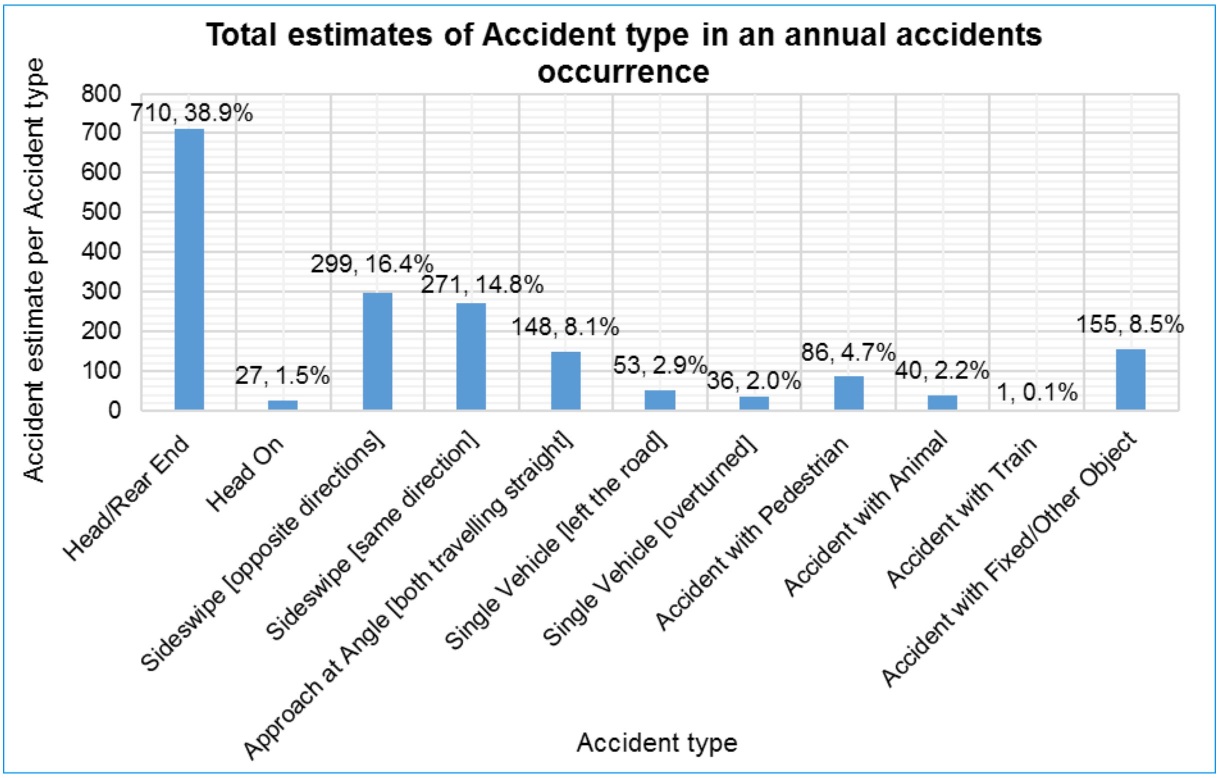

Figure 22: Total estimates of the Accident type in 2012

According to the array of data points in the table above, among the 11 variables, *'Head/Rear End'* recorded the largest number of road accidents at 38.9% [710 accidents] in SM, whereas similar result attained demonstrates that *'Head/Rear End'* occurred most regularly in the City of Cape Town Metropolitan Municipality which accounted for 47.5% [21,440][19] of all the 11 variables in the Accident type (Traffic Accident Statistics 2005). It is understood that this accident type occurred frequently at the road intersections, mostly when traffic rules are violated, such as red-light running, incorrect vehicle overtaking and violating pedestrian crossing rules (Sinclair & Murdoch 2012).

[19] Refer to the benchmark table in the Appendix B.3 for additional information on Accident type.

Further analysis result on SM demonstrates that *'Sideswipe [opposite directions]'* contributed second largest number of road accidents at 16.4% [299 accidents], followed by *'Sideswipe [same directions]'* which accounts for 14.8% [271 accidents]. However, in the City of Cape Town Metropolitan Municipality, findings illustrated that Sideswipe [same direction] occurred regularly as the second largest, followed by 'Accident with Fixed/Other Object' which accounts for 13.1% [5,926 accidents] (Traffic Accident Statistics 2005). In the case of SM, the lowest of all the 11 variables is *'Accident with train',* which has a score of 1 [0.1%] accident case.

This accident type is considered unusual in SM. The circumstances leading to this accident type could be as a result of sloppy operational responsibilities carried out by the two departments during the period of the incident. This particular incident only occurred in August, and it may take another decade for such incident to reoccur. Among these categorised variables, however, '*Accident with pedestrian'* placed 6th position with a total score of 86 [4.7%] accidents, but rather placed in the 4th position in the City of Cape Town with a total score of 5,426 [13.1%] accidents (Traffic Accident Statistics 2005). This shows that pedestrians are at a higher risk in the City of Cape Town Metropolitan Municipality than in SM.

The chart displayed above demonstrates an improved view of the data points to stimulate a simple observation of other relating results. In the chart, a few data points on accidents relating to '*Single vehicle [overturned]'* was observed, with a total score of 36 [2.0%] accidents. This type of accident case infrequently occurs in the Stellenbosch area, but rather considered as part of the relevant analysis results required in the Accident related factors. With a clear clarification, the result further insinuates that most accident cases reported to the STD happened mostly during traffic congestions, when the drivers approach or leave the controlled or uncontrolled junction (Sinclair & Murdoch 2012). Although, a better interpretation of the result is supported with the results derived from the *Road type* and *Junction type.*

From a different perspective, the impact angles connecting the vehicles involved together infer that most accidents occurred at the junction [intersection] areas (Mansfield et al. 2008; Sinclair & Murdoch 2012). The description stating how the accident occurred, and the sketch showing the occurring effect of the accident in the completed ARF support the illustration given above. In addition, this illustration can also be supported by the result acquired in the *Vehicle manoeuvre*[20]*,* which describes the extra activities engaged by the drivers at period of the incident.

[20] Refer to Appendix D for the analysis result of the Vehicle manoeuvre.

4.3 Introduction to the analysis of the non-captured data in Accident related factors

This section discusses the result produced by using two different methods of descriptive statistical analysis. The first method is based on the analysis of the average estimates of the non-captured data[21]. This method is applied by computing the average estimates of mismanaged data across all the related data fields categorised in the *Accident related factors.* The motive behind the application of this method, is to have a feasible interpretation of the average estimates of all data elements left uncaptured in a specific field during road accidents reportage.

In contrast, the second of the two methods, is applied to have a practicable illustration of the rate at which a particular field is being mismanaged during data collection actions. The estimates of count per non-captured data range in *Accident related factors* are graphically illustrated in the subsection 4.3.2 below. This process is achieved through the application of histogram as a statistical tool applied, to illustrate the distribution shape of the non-captured data. In this context, a simplified description of the non-captured data was provided earlier in Chapter 3, in which discussion on the suitable interpretation of the type of errors committed in the completion of ARF was presented.

Furthermore, the estimated scores presented in this section constitute the amount of data errors assembled as the non-captured data. The illustration of the result will obtain a better understanding into the factors identified as the key cause of data loss in the *Accident related factors.*

4.3.1 Average estimates of the non-captured data in Accident related factors

The scores arrayed in the Table 12 demonstrate variations in the average estimates of non-captured data amongst the eight data fields in Accident related factors. The variations illustrate the degree at which data were mismanaged in the eight related variables/data fields during the road accident data collection process. The result shows that there are some abnormalities contributing to the mishandling of road accident data during the data collection activities. For more details on the simple mathematical approach to the average estimates of non-captured data refer to subsection 3.3.2.7.2.

---

[21] For more information on the non-captured data refer to subsection 3.3.2.5.2.

However, under Accident related factors, a total average of 324 data is mismanaged due to poor data quality processing. From the set of scores grouped in the table below, with an in-depth understanding, it is observed that fewer number of variables/data fields in the Accident related factors lost a large number of road accident data, unlike other related factors discussed in the chapters 5 and 6. The results obtained from the analysis reveal that a huge average estimate of 156 [48.3%] data is mismanaged in *Summary of persons involved* as presented in the Figure 23 below, which is virtually up to the half value of the total average of data mismanaged in the Accident related factors. This huge amount is in excess of 47.3% of the least average estimate of 1.0% data that are considered missing in the *Weather conditions and visibility* according to the results displayed in the table below.

Table 12: Estimated average numbers of non-captured accident data in Accident related factors

Average estimates of non-captured accident data in Accident related factors		
Accident related factors	**Average estimates**	**Percentage estimates**
Accident date	4	1.2%
Day of week	15	4.6%
Time of accident	4	1.1%
Accident type	53	16.3%
Severity of injury	84	26.0%
Summary of persons involved	156	48.3%
Light condition	5	1.5%
Weather conditions and visibility	3	1.0%
Total average estimate	**324**	

In addition, the results further reveal that *Severity of injury* and *Accident type* also contributed large amounts of non-captured data among other remaining five variables/data fields with average estimates of 84 [26.0%] and 53 [16.3%] respectively. Huge amount of data mishandled are lost to factors such as:

- Lack of consistency along the reporting system,
- Inability to acquire relevant information concerning the type of accident *[which might be due to the severity level of the accident],* and
- Lack of commitment from the accident reporting officers.

This illustration reveals that little or no awareness is directed towards the significance of these three variables/ data fields by the form users, such as the reporting officers [police and traffic officers], the supervisory officer, and the DCO during the data collection protocols. Definitely, this may affect some important activities like decision-making in the transportation safety systems, effective management structure, and quality distribution of safety resources and tools, where quality road accident data is relevant in the improvement of processes. However, this consequence maybe averted by necessitating regular supervision to deter incorrect representation and exclusion of relevant variables/data fields. If considered, therefore, the traffic management could advocate for necessary resources required to reduce the rates at which data are mismanaged during data collection process.

According to the results displayed in the chart below, it is observed that *'the magnitude of the registered/reported number of accidents does not determine the reduction rate in the amount of the data loss during the data collection activities.'* Clearly, under the *Summary of persons involved,* the average estimate calculated demonstrates higher scores of data loss per monthly registered/reported number of road accidents, which fall within the range scores of 100 and above unlike other related data fields. This illustrates that less or no commitment was exercised by the reporting officers towards the understanding of this particular field, which could be ascribed to poor interpretation of information. This further insinuates that little or no supervision is implemented periodically to determine or check the competency of the reporting officers in the coverage of road accident data.

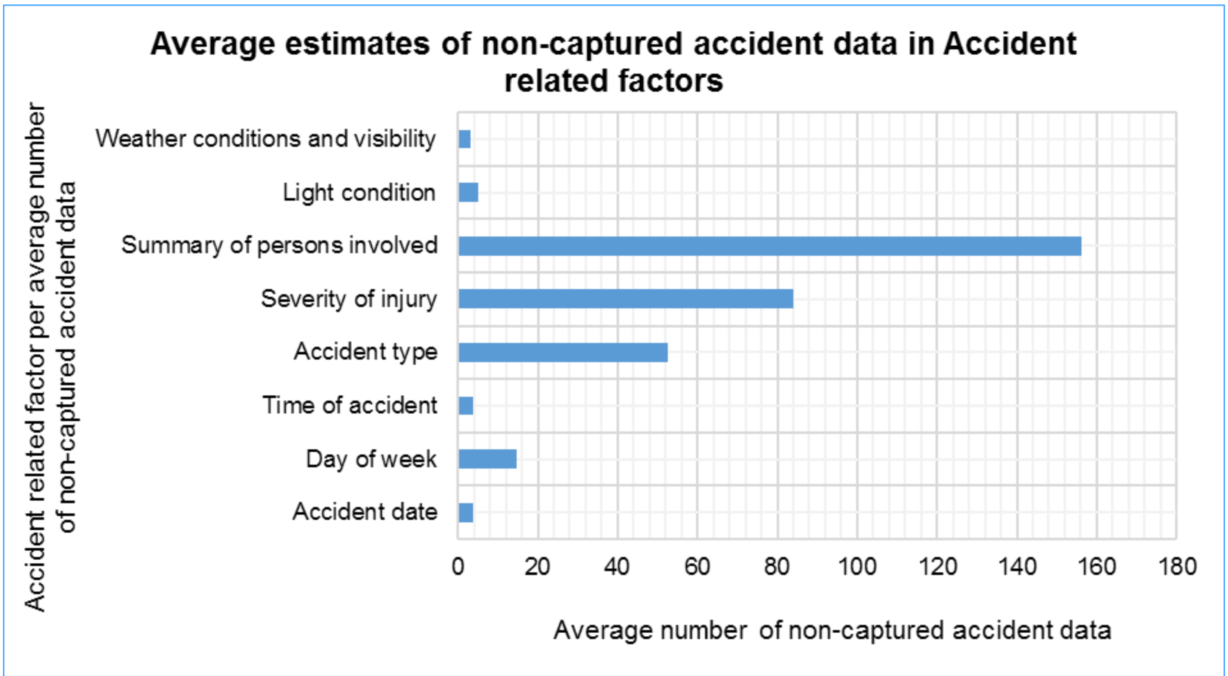

Figure 23: Total average estimates of non-captured accident data in Accident related factors

Quite the reverse, a minimal loss of data is demonstrated in few of the eight related variables/data fields grouped under the Accident related factors. From the chart above, it is understood that *Weather conditions and visibility* produces the least amount of missing data, with an average estimate of 3 [1.0%] non-captured data. And also, an equal average estimate of 4 non-captured data are calculated for both *Time of accident* and *Accident date,* with a close percentage estimates of 1.1% and 1.2% respectively. The magnitude of the average estimates of non-captured data calculated indicates high reflection of a reliable interpretation of the three data fields mentioned above during road accident data collection.

More so, according to the analysis, it is understood that more awareness, commitments, and composures were directed towards a complete collection of data pertaining to the four variables/data fields with the lowest average estimates of non-captured data. It appears that many reporting officers find the information about the four data fields less difficult to acquire[22]. Ultimately, this infers that a large number of the reporting officers understand the interpretation of these four data fields than any other data fields grouped in the Accident related factors.

Technically, *'the lesser, the number of non-captured data that are acquired; the higher, the number of captured data; and the better, the practicality of the usable data'.* In other words, '*the higher, the number of omitted data or errors discovered per field, the greater its consequence on the practicality of the data captured per field';* although, it depends on the degree of sufficiency of the quantity of data completed in the ARF. This illustration literally explains that, *'the complete number of data captured correctly [usable data], determines the extent of the practicality of a particular field towards the analysis of a real-world problem in the RTA*'.

4.3.2 Histogram analysis of the non-captured data in Accident related factors

The set of estimated scores arranged in the Table 13 are considered for the frequency analysis of the non-captured data extracted from the Accident related factors. From the Table 13, the estimated scores demonstrate observable variations in each field. The amount of data lost in some related data fields like *Accident date, Time of accident, Light condition* and *Weather conditions and visibility* per month, is considered minimal compared to the size of data lost in *Day of week, Accident type,* and also *Severity of injury* and *Summary of persons involved* as displayed in Table 13.

A complete collection of data pertaining to Accident date is observed in November and December, while similar improvement is also observed in Weather conditions and visibility in February. This demonstrates the desire of the traffic management in agitating for zero tolerance towards any kinds of error committed during the road accident data collection.

The methods applied in computing the frequency distribution shown in both Figure 24 and Figure 25 are mathematically illustrated in the subsection 3.3.2.7.3. During the process, some

---

[22] Refer to Chapter 7 for more relevant information on the perception of the reporting officers concerning their understanding on the data fields.

parameters were considered in the formation of the class intervals that produced the histograms, such parameters as maximum, minimum, count, range, and class width. In the two tables, the empty cells indicate complete collection of road accident data, that is, no detection of error in any of the data fields. The frequency distributions generated are diagrammatically illustrated in the Figure 24 and Figure 25. The two charts present the frequency distribution of the non-captured data across all the data fields in the Accident related factors. In this case, *the distribution of the estimated scores are based on the count per non-captured data range.*

Table 13: Estimates of non-captured data for data fields with least missing data in Accident related factors

Total estimates of non-captured data for least missing data in Accident related factors					
Months	**Accident date**	**Day of week**	**Time of accident**	**Light condition**	**Weather conditions and visibility**
Jan	1	11	3	4	2
Feb	7	17	4	4	
Mar	2	7	2	9	3
Apr	3	14	16	1	6
May	5	11	1	6	2
Jun	5	14	4	5	4
Jul	4	11	2	5	4
Aug	5	23	2	2	6
Sep	4	18	2	7	1
Oct	3	13	3	8	5
Nov		25	4	6	2
Dec		13	1	3	1

The chart below demonstrates a *J-shaped* distribution with a positively skewed type of histogram, wherein the peak score is positioned at the lowest range as indicated on the horizontal axis. The ranges with fewer frequencies appeared farther to the right side of the histogram. In addition, the distribution further indicates that high count of non-captured data was attained within the range of 1-4 non-captured data. This particular range produces the highest frequency of 30 counts. This frequency demonstrates the amount of times this particular range of non-captured data was discovered in each field.

It is further observed that an estimated frequency of 14 counts is generated within a range of 5-8 non-captured data, which is practically equal to half of the estimated frequency acquired within the range of 1-4. This illustration explains that five to eight data [non-captured] are habitually mismanaged at the rate of 14 times per accident reporting. However, only four data fields fall within these two ranges, such data fields as *Accident date, Time of accident, Light condition* and *Weather conditions and visibility.* With a clear explanation, the result acquired within the range of 1-4 non-captured data, reveals that it is easier to commit errors at the rate of one to four data on a daily accident reportage activity. Similar illustration is applicable to the histogram of large scores of non-captured presented in the Figure 25.

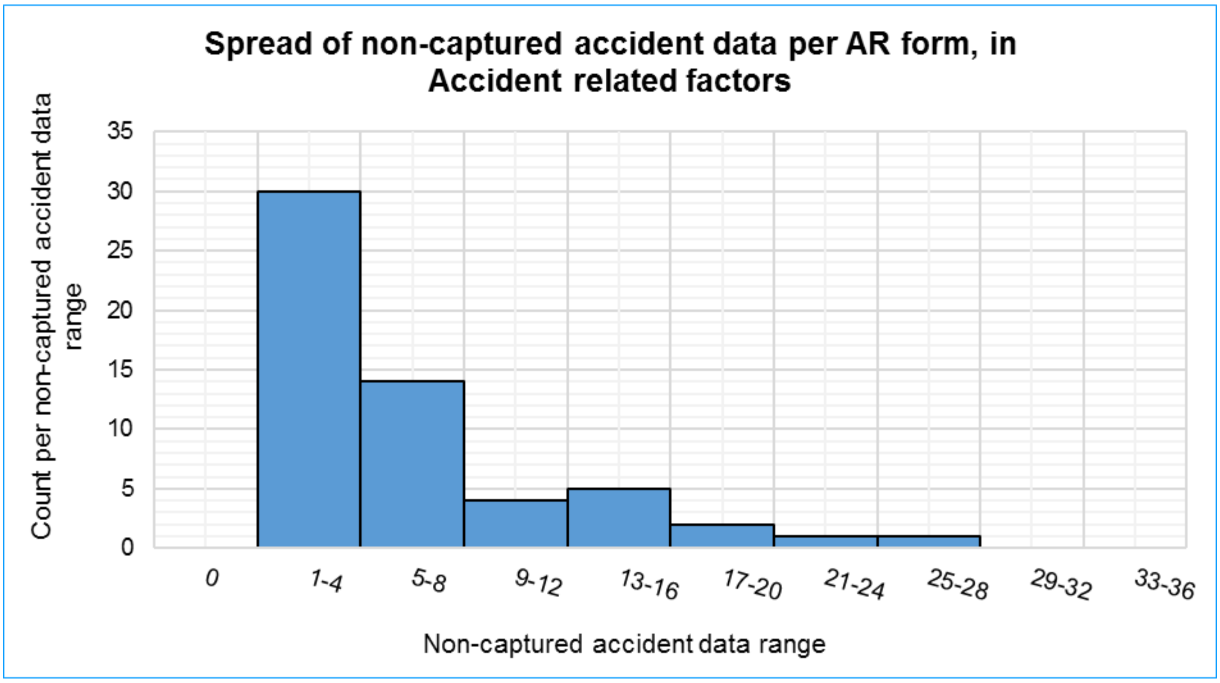

Figure 24: Distribution of non-captured data for data fields with least missing data in Accident related factors

Other results observed show that a lower estimated frequency of 4 counts was attained within the range of 9-12 non-captured data. Another lower estimated frequency of 5 counts was acquired within the range of 13-16 non-captured data, which is considerably higher than the result observed within the range of 17-20 non-captured data. An equal frequency estimate of one count falls within the ranges of 21-24 and 25-28 of non-captured data respectively. In this context, with regard to small scores of non-captured data, errors leading to these ranges of non-captured data are restricted to one particular field that suffers a huge missing data at the range of 21 to 28 non-captured data per daily reportage of road accidents.

From a close observation, as displayed in the Table 13, the estimates of the non-captured data accumulated in the *Day of week*, reflect high estimated scores of non-captured data. This illustrates that few reporting officers find it difficult to acquire the accurate day that a road accident occurs. Generally, throughout the year, errors committed in the *Day of week* is considered frequent. Since the scores acquired spread across only the high ranges of non-captured data farther to the right side of the frequency distribution chart. The problem contributing to insufficient data in the Day of week is issue of late reportage of accident, where the accident was never reported on the exact day of occurrence, but only in some days later.

The table below presents the assemblage of the scores acquired for the three other related data fields, with large scores of non-captured data in the Accident related factors. The distribution of these scores exhibits variations across the 12 months. Besides, the size of the least score observed in this group is extremely greater than the corresponding high score observed in the Table 13, therefore, this necessitate the formulation of a suitable class interval. Additionally, the size of the class width calculated determines the formation of the class interval, alongside the estimated scores considered for the number of classes through the comparison of the results obtained from both the *Sturgis' rule* and *Rice rule* application.

Table 14: Estimates of non-captured data for data fields with most missing data in Accident related factors

Total estimates of non-captured data for most missing data in Accident related factors			
Months	**Accident type**	**Severity of injury**	**Summary of persons involved**
Jan	39	70	133
Feb	39	87	60
Mar	79	93	203
Apr	50	74	165
May	48	97	190
Jun	47	69	155
Jul	38	68	124
Aug	68	89	188
Sep	58	89	165
Oct	57	98	206
Nov	63	105	170
Dec	46	69	117

From the Table 14, considering the size of the estimated scores grouped in each data field, only the *Summary of persons involved* produces extremely larger scores of non-captured data. The distribution shape obtained produces a positively skew distribution of estimated scores of non-captured data farther to the right side of the distribution chart with a distinct peak as demonstrated in the Figure 25 below.

However, highest peak of 10 counts is observed within the range of 53-78 non-captured data. This frequency estimate constitutes the data mismanaged in the *Accident type* and *Severity of injury.* Similarly, an equal frequency estimate of 7 counts is observed within the range of 27-52 and 79-104 non-captured data respectively, which is majorly the accumulation of non-captured data contributed by both the *Accident type* and *Severity of injury.* This explicates that data pertaining to these two related data fields are regularly mishandled at high degree of non-captured data, compared to the *Summary of persons involved* with high estimated scores of non-captured data distributed within the ranges of fewer frequency scores.

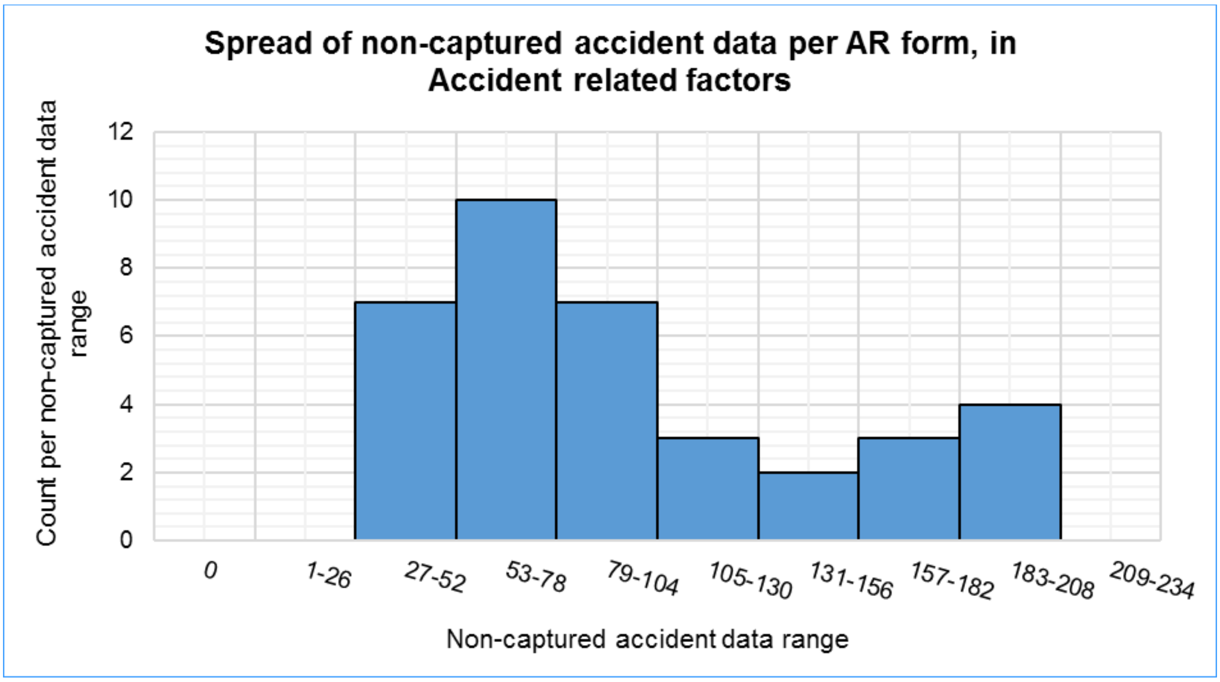

Figure 25: Distribution of non-captured data for data fields with most missing data in Accident related factors

Conversely, a frequency estimate of 4 counts is observed as the highest estimated score acquired within the ranges with large amount of non-captured data. This reflects the amount of times that a high amount of non-captured data is discovered in a particular field in the Accident related factors.

5. Analyses and findings in Road related factors

This chapter discusses the contributions of the *Road related factors* to the accidents that happened in 2012 according to the result of analysis performed on the unprocessed data assembled. This discussion emphasises the fitness level of the data acquired in the Road related factors, by having an in-depth knowledge into the practicality of the data analysed. In the process of achieving this, thus it is necessary to understand the relationship between these data fields, with the purpose of establishing a practical interpretation of the result obtained.

The Road related factors are classified into five data fields, which are the *Speed limit on road* and *Built-up area, Road type* and *Junction type, Road surface related conditions, Road impediment related conditions,* and two others such as *Road control related conditions and Road layout related conditions.* The Road surface related conditions constitute some relevant data fields like *Road surface type and Quality of road surface,* which are discussed correspondingly as provided in the data form; while under the Road impediment related conditions such relevant data fields as *Road marking visibility, Obstructions and Overtaking control* are also discussed in the Appendix C.

Similarly, two important data fields are grouped under the Road control related conditions, which are *Traffic control type* and *Condition of road signs*. The Road layout related conditions consist of relevant data fields such as *Direction of road and Position of vehicle before accident.* Further discussions in relation to these four data fields are provided in the Appendix C-C.3.1 & C.3.2, to support the analysis and findings regarding the usability of the data assembled in Road related factors.

All of the above-mentioned related factors have several distinct variables, which are combinations of both *'possible causes'* and *'contributory factors'*. Additionally, the interpretation of the result obtained demonstrates the periodic contribution of the Road related factors to the occurrence of RTA. As part of the analysis carried out in this section, the histogram distribution and average amount of non-captured data acquired in each field are discussed along with the overall estimate of the usable data. The outcome of the interpretation will support a preliminary understanding into the application of the ARF. This particularly offers hints on why some variables are poorly represented in the data form, in connection with the circumstances affecting the reporting officers from acquiring relevant information.

5.1 Introduction to the analysis of the Speed limit on road and Built-up area

This section produces the outcome of the analysis performed on the *Speed limit on road* and the *Built-up area* in the two separate tables presented in subsections 5.1.1 and 5.1.2 below.

From the Table 15, the specifications for the *Speed limit on road* are provided in the first column of the table, which are considered as the categorical variables. The speed is measured over the '*actual distance travelled per unit time [kilometres per hour]'*. Actually, the implementation of the speed limit specifications primarily depends on the related data fields like the *Built-up area* and the *Road type*. On the other hand, the *Built-up area* can be considered as the area with the largest number of settlers, infrastructures, and individual properties. Moreover, the rate at which accidents occur along a built-up area necessitates the key implementation of the speed limit, with the purpose of reducing predetermined occurrence of the RTA in order to controlling the magnitude of the speed engaged by the drivers/motorcyclists travelling on the roads.

5.1.1 Analysis of the Speed limit on road

The table below comprises the data points obtained from the analysis performed on the speed limit specifications represented in the completed ARF. The speed limit specifications ranged from 10km/hr to 120km/hr, as otherwise specified through road signs which depends on the type of road and the location of the road. In essence, according to the description of the *Built-up area* in the section 5.1 above, however, the need for applying a specified speed limit along any particular roads requires a reliable road accident data to justify the implementation by considering factors, such as:

- Rapid infrastructural development *[determines by the level of the revenue generated],*
- Design of the road *[determines by the developing level of the area, whether in a rural or an urban area],* and
- Congestion level of the road *[determines by the number of the road users and vehicles].*

From the table set below, it is observed that a total estimate score of 1,390 [77.0%] accidents occurred within the speed limit range of 60km/hr in 2012. This estimate is positioned as the highest among other road speed limits specified, followed by the speed limit range of 100km/hr and 80km/hr with total estimate scores of 177 [10.0%] and 158 [9.0%] accidents respectively. The score calculated for 60km/hr suggests that most accidents occurred within the built-up areas, that is, areas with a huge concentration of the residents such as local streets and avenues.

On the other hand, the total scores calculated for the speed limits beyond 60km/hr demonstrate a few occurrences of accidents on the highway, urban freeway and rural roads located within/around Stellenbosch locality. The set of speed limits from 10km/hr to 60km/hr and below are strictly applied to the local streets and avenues in the built-up area to control the speed engaged by the drivers, in order to deter the rate at which road accidents occur within this area.

The categories of accidents occurring in the parking lots and the fuel/gas station [garage] are categorised under the speed limits below 30km/hr, which covers a total estimate of 19 [1.1%] casualties as displayed in the Figure 27.

Table 15: Estimates of road accident at specified Speed limit

Total estimates of road accident occurrence at specified Speed limit in 2012													
Speed Limit *[km/hr]*	**Jan**	**Feb**	**Mar**	**Apr**	**May**	**Jun**	**Jul**	**Aug**	**Sep**	**Oct**	**Nov**	**Dec**	**Total**
120km/h	0	0	0	1	0	1	0	0	1	0	0	0	**3**
100km/h	10	11	24	16	9	10	8	12	20	17	23	17	**177**
80km/h	17	17	11	16	14	8	7	8	13	12	19	16	**158**
60km/h	77	126	161	123	125	100	76	122	106	154	140	80	**1390**
40km/h	4	2	3	4	7	1	1	3	5	5	3	1	**39**
30km/h	2	1	1	2	2	3	2	0	2	5	1	0	**21**
20km/h	1	0	1	0	2	1	1	2	1	4	0	4	**17**
10km/h	1	1	0	0	0	0	0	0	0	0	0	0	**2**
Total scores	**112**	**158**	**201**	**162**	**159**	**124**	**95**	**147**	**148**	**197**	**186**	**118**	

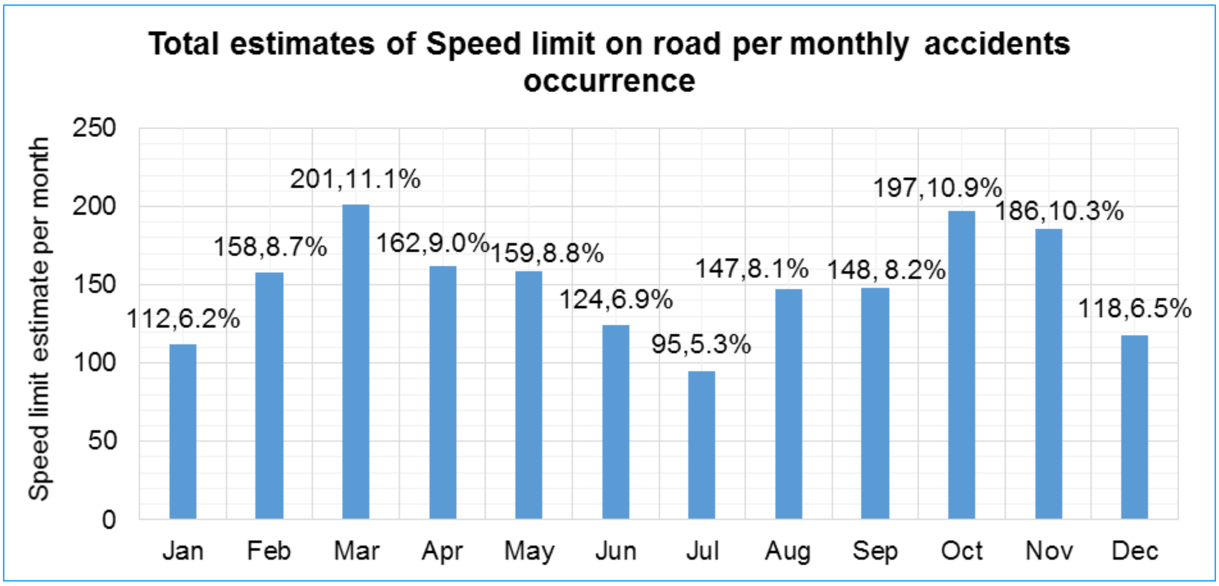

Figure 26: Total estimates of the Speed limit on road per month

Most of these cases are attributed to situations where a vehicle owner met his/her vehicle dented, or similar situation where a vehicle reversed into another vehicle. Additional problem observed is attributed to the attitudes of the road users not observing the speed cautionary signs along the road side.

By observing the monthly estimates of the total scores displayed in the Table 15 above, it is deduced that more data pertaining to the specified speed limits were represented in March based on the degree of accuracy and completeness exercised during the process of collecting the data. This suggests provision for more data depending on the efficiency or performance level of the reporting officers. *The lesser the RTA cases reported daily/monthly, the higher the probability of committing lesser errors.* This condition depends on the commitment of the reporting officers towards the completion of the ARF.

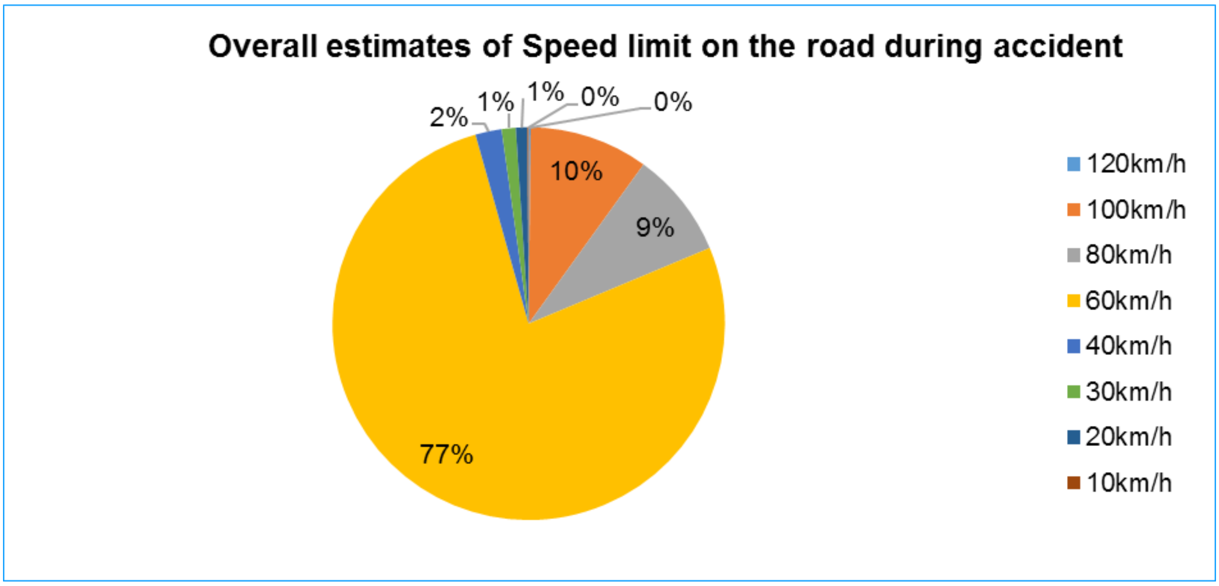

Figure 27: Total estimates of the specified Speed limit at accident location in 2012

The importance of having a complete and regular indication of speed limit in the ARF could aid the ability of the traffic engineers and the road construction agencies, in having the general view of the average speed attained by the road users after any road traffic accidents occurrence (Bester & Geldenhuys 2007).

5.1.2 Analysis of the Built-up area

The table below illustrates the rate at which accidents occurred inside/outside the built-up areas within the Stellenbosch Municipality [SM]. The data collected are characterised into two variables, which are *Built-up area accidents* and *Non-built up area accidents*. The scores assembled in the Table 16 demonstrate that more road accidents occurred within the Built-up areas than the Non-built up areas, where large numbers of Stellenbosch residents inhabited.

In the chart displayed in Figure 28, total estimates of 1,244 [83.0%] accidents occurred within the Built-up area, at a speed limit of 60km/hr and below. Although this amount demonstrates an unequal rate to the total amount realised in the speed limit within the range of 60km/hr and below. This is due to the huge rate of data loss to uncompleted data elements in both Built-up area and Non-built-up area accidents. The high rise in the number of accidents occurring at a speed limit of 60km/hr can be attributed to factors, such as violation of traffic rules and nonchalant attitudes of the road users.

Table 16: Built-up and non-built up areas estimates in road accident occurrence

Total estimates of road accidents occurrence in Built-up and Non-built up areas in 2012		
Months	**Built-up area**	**Non-built up area**
Jan	68	18
Feb	113	13
Mar	135	24
Apr	114	25
May	102	21
Jun	89	16
Jul	49	15
Aug	112	17
Sep	103	23
Oct	148	29
Nov	123	29
Dec	88	30

In contrast, a small amount of 260 [17.0%] accidents occurred in the non-built up area in Stellenbosch locality. This estimate demonstrates number of road accidents that occurred on the roads located in the sparsely settlement, mostly at a speed limit above 60km/hr. The chart presented in the Figure 29 demonstrates variations determined in the monthly estimates of the data points assembled in the table above for both built-up and non-built up areas. A gradual increase was observed in the rate at which accidents occurred in the built-up areas. This indicates an increment rate of 3.6% from January to February, which later dropped to 1.8% from February to March. This illustration suggests busy traffic conditions during these periods, which is ascribed to the peak periods identified in the section 4.2.3 above, with a slight growth in the number of road users.

However, a deep drop was observed from April to July, which could be attributed to the public holidays observed within these periods as previously discussed in the subsection 4.2.2 above. From August, a huge increase in the amount of the road accidents was observed in the built-up area, which is linked to the busy academic sessions in the schools and other business activities; however, the increment slightly dropped in September probably due to the adjustment towards road policies by the road users. An extreme increase in the amount of accidents occurred in the built-up area was observed in October, which gradually dropped through November to December. The decline trend observed could be ascribed to less road users, which is due to public holidays observed in December.

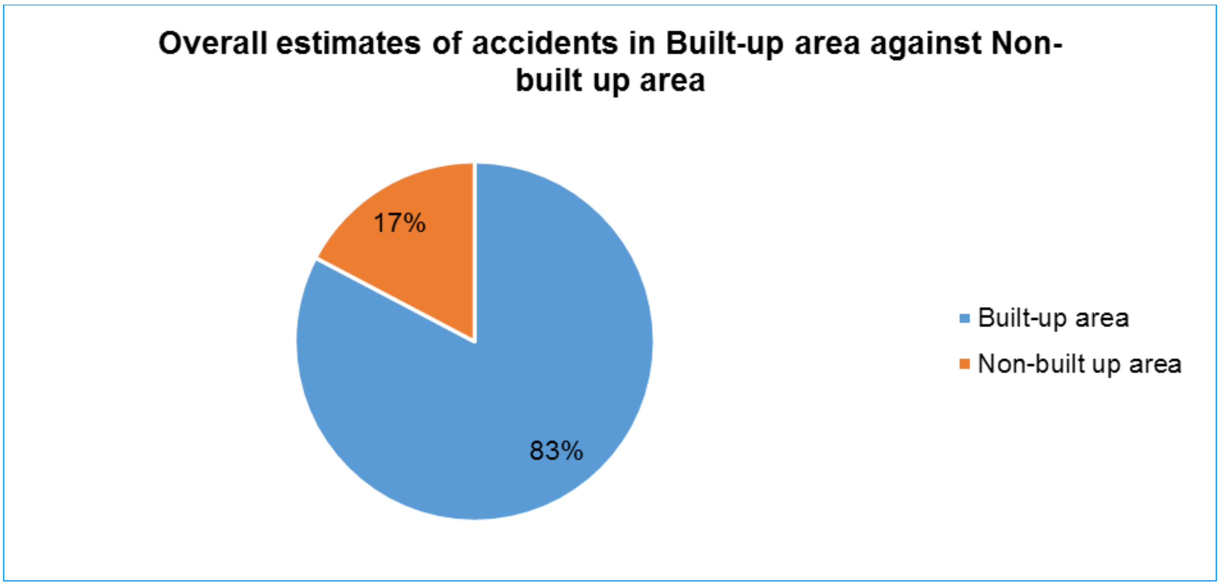

Figure 28: Total percentage estimates of accidents in Built-up area against Non-built up area in 2012

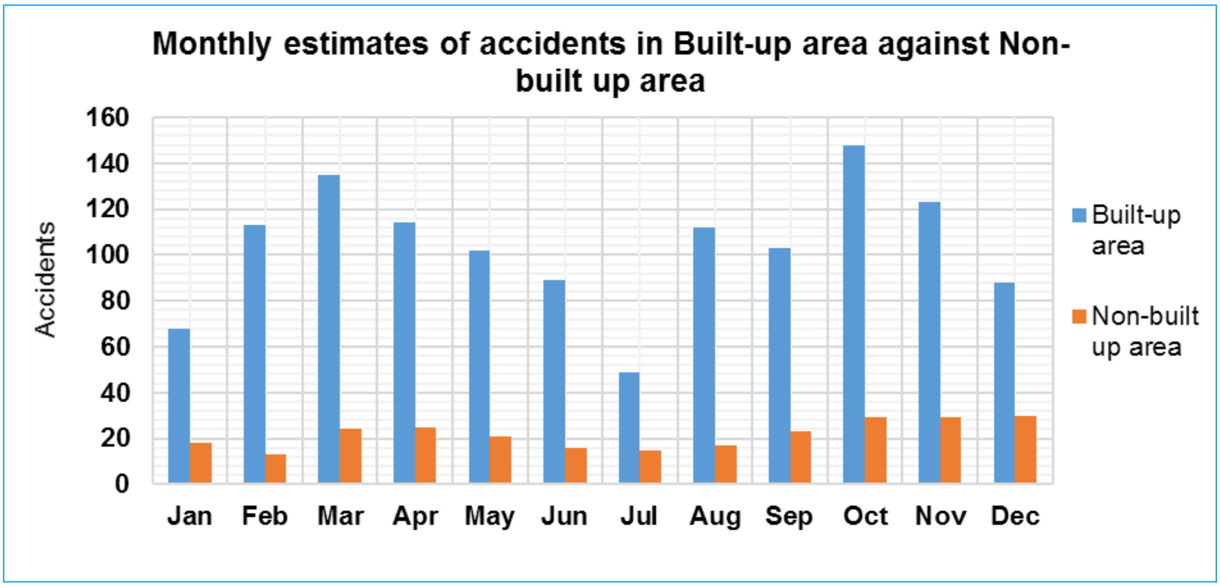

Figure 29: Total estimates of accidents in Built-up area against Non-built up area per month in 2012

On the contrary, in the case of non-built up area, the variations observed are different to the one demonstrated formerly in the built-up area. Unlike the built-up area, a slight drop was observed in February with a total estimate difference of 5 accidents, followed by a gradual increase from March, extending to April as graphically revealed in the chart above. This increment dropped with reduction rate of 1.5% between May and June, and 0.4% between June and July. However, a constant increase in trend was observed from August, extending to December with equal amount of 29 accidents in both October and November.

The chart above demonstrates trends with similar trajectory to the chart plotted for the *Speed limit on road* in Figure 26. However, more data were accurately and completely represented in the *Speed limit on road*, in excess of the data represented in the *Built-up area.* This demonstrates that inconsistencies marred the completion of the data collected with regard to the two related data fields. From a different perspective, one can simply conclude that users find the acquisition of information on the Speed limit more significant and easier to comprehend than the Built-up area during the completion of the ARF.

5.2 Introduction to the analysis of the Road type and Junction type

This section discusses the categories of roadway types travelled by the road users and the junction types intersected while travelling on the road on a daily basis, as parts of the considerable contributors to the occurrence of accidents in the Stellenbosch area. However, it is understood that more than one junction or intersection are encountered along the roadways in Stellenbosch. Therefore, estimation of the effect of the road and junction layouts on the road users should be carried out regularly. The results obtained in the analysis of the two data fields establishes a relationship relating the two data fields together with regard to occurrence of road traffic casualties.

5.2.1 Analysis of the Road type

Road type consist of eight variables, which are defined according to the categories of the roadways in the Stellenbosch locality where accidents actually occurred. The first of the eight variables is *Freeway,* followed by other variables like *On/off ramp, Dual carriageway, Single carriageway, One way,* and three other remaining variables such as *Other, On-road parking/rank,* and *Off-road parking/rank.* The data points assembled in the table set below demonstrate close scores between the '*dual carriageway'* and *'single carriageway'.* According to the table displayed in this section, only *dual carriageway* and *single carriageway* generated 82.8% of the total percentage estimates calculated in Road type, while the remaining proportion constitutes the six other variables as demonstrated in the Figure 30.

As graphically illustrated in the chart below, estimated proportions of 42.0% [675] and 41.0% [662] road accidents occurred on both the *dual*-and *single*-carriageways respectively, within a high-speed limit of 60km/hr, depending on the location of the accidents. Among the remaining variables, it is observed that an estimated proportion of 5.0% [87] road accidents occurred on *Off-road parking/rank*, followed by *Freeway* with the road accident estimate of 5.0% [83]. Observably, the result realised from the analysis of the *Freeway* demonstrates that fewer accidents occurred along the traffic with specified speed limits greater than 60km/hr.

Table 17: Road type estimates in road accident occurrence

Total estimates of road accident occurrence on specified Road type in 2012								
Months	**Freeway**	**On/Off Ramp**	**Dual Carriageway**	**Single Carriageway**	**One Way**	**Other**	**On-road Parking/ Rank**	**Off-road Parking/ Rank**
Jan	5	0	23	50	2	4	6	6
Feb	10	0	46	60	3	3	3	5
Mar	14	1	81	62	4	4	4	9
Apr	10	1	53	49	2	3	4	7
May	7	1	44	57	3	0	1	5
Jun	8	0	55	39	4	1	1	4
Jul	9	2	26	37	3	1	2	6
Aug	4	0	76	52	5	2	5	8
Sep	6	1	49	64	1	1	2	4
Oct	2	3	95	65	5	2	8	12
Nov	4	0	83	72	1	1	4	11
Dec	4	0	44	55	2	0	3	10
Total scores	**83**	**9**	**675**	**662**	**35**	**22**	**43**	**87**

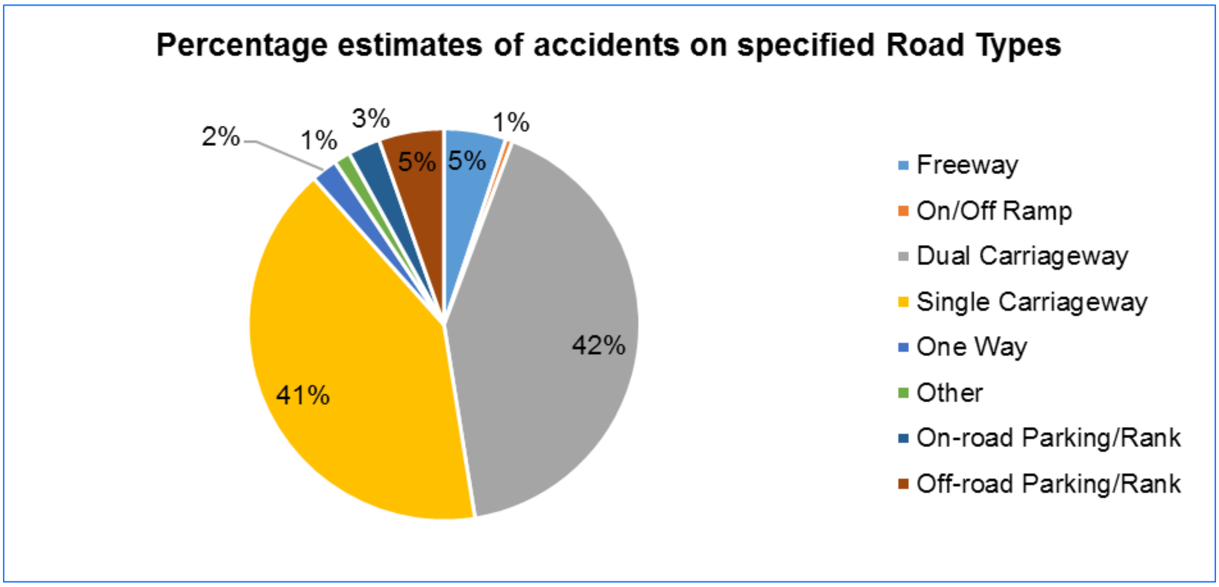

Figure 30: Total estimates of Road type in 2012

It is understood that speed limit is proportional to change or reduction along the *dual* and *single carriageways* while approaching the built-up areas to ensure safety for all the road users, from a higher speed limit to a lower speed limit (Langford & O'Hare 2005). Despite the implementation of the speed limit cautionary signs along the road sides, however, more accidents occurred everyday along the two road types mentioned above in the Stellenbosch locality. The suggested factors responsible for such large number of casualties along these two road types are identified as:

- Violation of specified speed limits along the road,
- Poor observation of the speed limit cautionary signs along the roadsides,
- Influence of alcohol and drugs used by the driver while driving,
- Lack of adequate speed detecting devices along the busy roads (Njord et al. 2005),
- Lack of sufficient quality data to support adequate distribution of the road safety resources (O' Day 1993; Vogel & Bester 2004), and
- Lack of commitments from the authorised officers in charge of cautioning the behaviours of the road users along these roads.

5.2.2 Analysis of the Junction type

Junction type consist of 12 related variables, which are considered as part of the probable contributors of road casualties within the Stellenbosch area. It is observed that not all junctions are controlled in the Stellenbosch locality. Some junctions are automatically or manually[23] controlled by the robots or road signs, depending on how busy the intersections could be during intense traffic activities around the locality. Most road types with controlled junctions are commonly located in the main roads, such as the *dual* and *single carriageways,* while other roads like *one-way* have their junctions controlled mostly by a *stop sign*. This suggests that the risk posed by the uncontrolled junctions are higher than any other type of controlled junctions within the locality. This effect could increase the amount of road casualties occurring at the uncontrolled intersections (Sinclair & Murdoch 2012).

During personal communications, a high-ranking traffic officer, added that most junction types without a control robot are controlled by the implementation of the stop signs. In addition, the officer further emphasised that car drivers and motorcyclists are advised to cautiously observe the stop signs to avoid road traffic casualties. The stop sign regulates traffic based on a real-time process depending exclusively on the behaviours of the drivers or motorcyclists while arriving at the intersection point. It authorises the first arriving vehicle at the intersection, to depart ahead of the other vehicles in the same queue, in the order of a sequential structure to avoid any accident. In the context of the traffic regulatory sign, the safety of the pedestrians is also considered by implementing the pedestrian crossing signs. More in-depth details are provided in the Appendix C-C.2.1, where the analysis carried out demonstrates the contribution

[23] Some junctions are manually controlled based on the instincts of the road users on their understanding towards the road signs.

Table 18: Junction type estimates in road accident occurrence

Estimates of road accident occurrence at Junction type in 2012												
Months	**Cross Roads**	**T-junction**	**Staggered Junction**	**Y-junction**	**Circle**	**Level Crossing**	**Not a junction or crossing**	**On ramp/slipway**	**Off ramp/slipway**	**Pedestrian Crossing**	**Property Driveway/ Access**	**Other**
Jan	12	10	3	2	1	0	20	0	0	0	2	9
Feb	23	9	4	2	4	0	32	0	0	3	2	5
Mar	29	8	4	0	5	0	72	0	0	0	2	9
Apr	31	4	3	0	1	1	58	2	0	0	0	5
May	26	12	2	0	3	0	49	0	0	1	3	1
Jun	21	6	6	1	6	0	45	0	0	0	0	0
Jul	17	4	4	0	5	0	20	0	1	0	2	4
Aug	21	3	4	2	5	4	74	0	0	0	1	2
Sep	27	8	2	0	7	0	46	0	0	0	3	2
Oct	29	27	7	0	7	0	85	0	1	0	1	6
Nov	31	18	8	0	6	0	88	0	0	1	2	3
Dec	22	14	4	0	2	0	41	0	0	1	0	5
Total scores	**289**	**123**	**51**	**7**	**52**	**5**	**630**	**2**	**2**	**6**	**18**	**51**

of *Traffic control type* and other related data fields in the road traffic casualties.

From the Table 18, an estimated proportion of 51.0% [630] road accidents occurred at the locations that are neither junctions nor crossing. This estimate is more than the estimated proportion actualised in *Crossroads,* with a percentage difference of 27.8% as similarly displayed in Figure 31. The result presented with regard to the *Not a junction or crossing*,

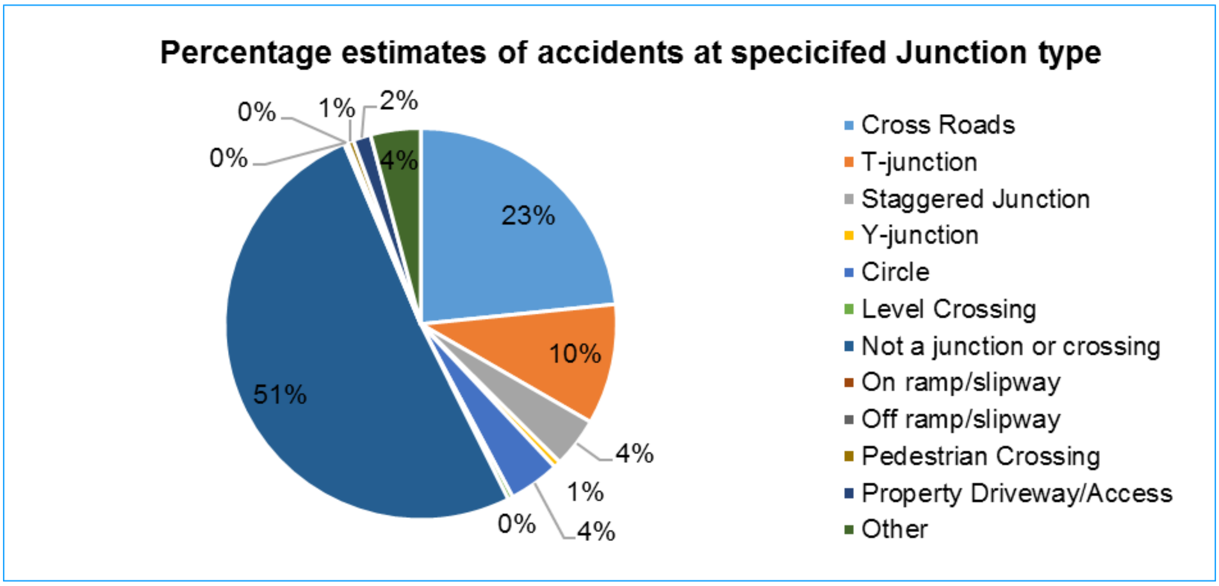

Figure 31: Total estimates of Junction type in 2012

insinuates that most accidents occurred at a point where the vehicle is not approaching any intersections or crossways along the roads in the Stellenbosch locality. In addition, a total proportion of 10.0% [123] road accidents occurred at the *T-junction,* which is far above the sum estimate of road accidents that occurred at both *Staggered junction*, *Y-junction* and *Circle*. This result is attributed to three main factors, which are driver's actions, characteristics of red-light runners, and conditions leading to red-light running (Sinclair & Murdoch 2012).

5.3 Introduction to the analysis of the Road surface related conditions

This section discusses the contribution of road surface in the cause of RTA in the Stellenbosch area. Two data fields are discussed in this section, which are relatively connected by one entity named *'surface of the road'*, which demonstrated the relational attributes between them. The two data fields considered in this section are *Road surface type and Quality of road surface.* These two related data fields are considered as part of the possible contributors of the occurrence of RTA within the Stellenbosch area. In this case, the *Road surface type* consist of five variables which are *Concrete, Tarmac, Gravel, Dirt*, and *Other,* while the *Quality of road surface* consist of six variables which are *Good, Bumpy, Pothole, Cracks, Corrugated* and *Other*. The analysis of the two related data fields is presented in the subsequent subsections below.

5.3.1 Analysis of the Road surface type

Table 19 contains data estimates of the five categorical variables in the *Road surface type.* The classification of the *'road surface type'* on the road, where accidents occurred is evaluated through the analysis procedures carried out in this section. The purpose of this analysis could provide sufficient hints on the probable causes of the RTA by understanding this particular data field, in accordance with the correlations of the other related data fields in the *Road surface related conditions.*

Table 19: Estimated scores of Road surface type at the time of road accident occurrence

Total number of accidents on Road surface type in 2012					
Months	**Concrete**	**Tarmac**	**Gravel**	**Dirt**	**Other**
Jan	14	137	2	1	1
Feb	4	221	0	0	0
Mar	8	233	4	1	1
Apr	7	173	1	0	0
May	6	218	0	0	2
Jun	5	166	2	0	4
Jul	5	155	0	0	2
Aug	3	222	4	0	1
Sep	6	189	1	2	0
Oct	7	216	2	1	0
Nov	9	206	1	0	3
Dec	3	130	2	0	1
Total scores	**77**	**2266**	**19**	**5**	**15**

From the Table 19, a set of high scores is aligned directly under the variable *'Tarmac',* which indicates that more accidents occurred on the tarmac roads since most local roads and interprovincial roads are constructed with tarmac. The data points grouped in the table set above, represent the number of accidents on the Road surface type from January to December according to the structure of the information indicated in the completed ARF.

In the chart below, an estimated proportion of 95.0% [2,266] road accidents occurred on the tarmac roads in SM in 2012, which is 90.0% above other variables. In addition, similar findings were demonstrated in the City of Cape Town Metropolitan Municipality with an estimated proportion of 88.4% [75,527] road accidents (Traffic Accident Statistics 2005). Among the remaining variables, only *Concrete* demonstrated a slight larger proportion in both SM and the City of Cape Town Metropolitan Municipality [see Appendix C.4 for more information] compared to other variables such as *Gravel, Dirt* and *Other*. Road accidents that occurred along these road surfaces, can be attributed to incidents that happened at parking lots, garages and rural farm roads.

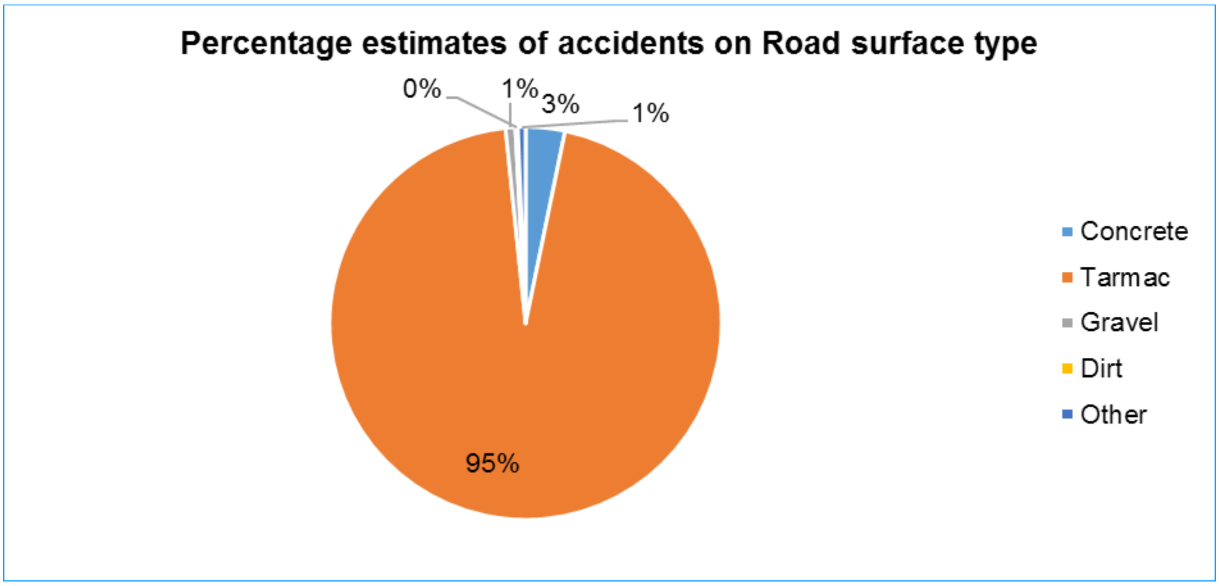

Figure 32: Estimated percentages of the road surface type at the time of road accidents occurrence in 2012

5.3.2 Analysis of the Quality of road surface

Table 20 contains all accidents scores calculated for the six variables grouped in the *Quality of road surface*. The result obtained in this section indicates that most accidents occurred on good roads, while a small quota of the result is acquired from the remaining five variables. However, an approximated proportion of 97.1% [2,355] road accidents occurred along *good* quality surface roads within the Stellenbosch locality in 2012. This huge estimate could only be incited mostly by three major factors, suchlike vehicular conditions, human behavioural characteristics and environmental challenges (Vogel & Bester 2005).

Table 20: Estimated scores of Quality of road surface at the time of road accidents occurrence

Quality of road surface in road accidents occurrence in 2012						
Months	**Good**	**Bumpy**	**Pothole**	**Cracks**	**Corrugated**	**Other**
Jan	155	2	2	0	0	0
Feb	234	0	0	1	0	0
Mar	243	2	3	1	0	2
Apr	173	2	2	0	0	0
May	219	0	0	0	0	2
Jun	170	2	0	0	1	4
Jul	153	3	3	0	0	3
Aug	217	5	0	2	0	3
Sep	217	5	0	2	0	3
Oct	224	4	0	0	0	3
Nov	212	2	1	0	0	3
Dec	138	1	0	1	0	0
Total scores	**2355**	**28**	**11**	**7**	**1**	**23**

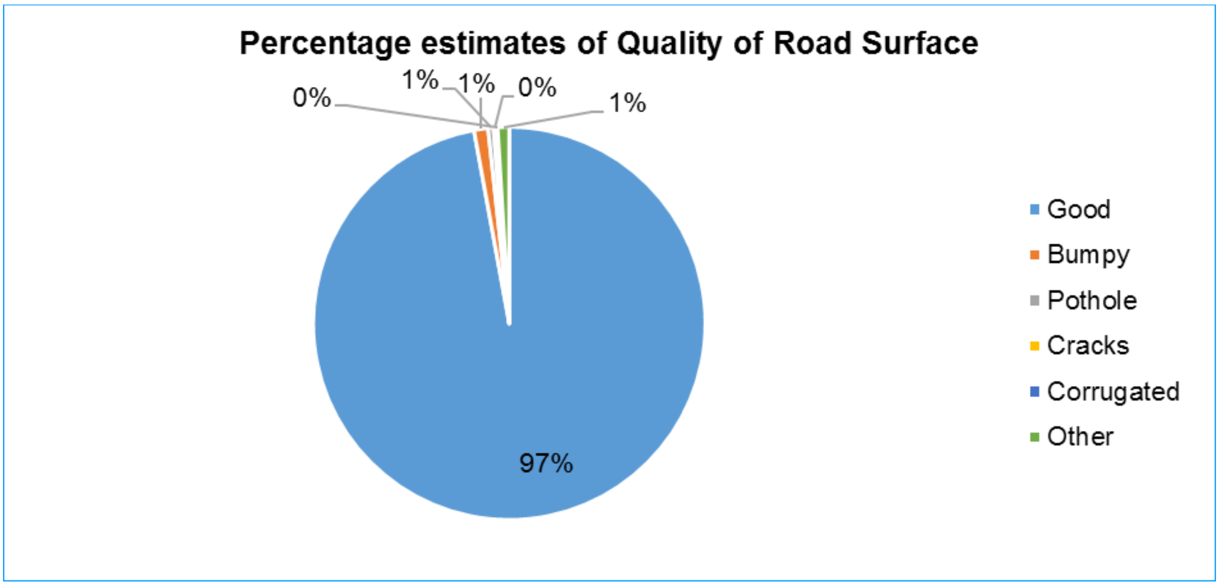

Figure 33: Estimated percentages of the road surface quality at the time of road accidents occurrence in 2012

Additionally, an approximated estimate of 1.2% [28] road accidents occurred along the *bumpy* quality of road surface in the same year, followed by small estimates of road accidents that occurred on the poor road surface conditions, such as *pothole, cracks,* and *corrugated* road conditions with approximated sum estimates of 1.0% [see graphical illustration above]. Besides, an estimated proportion of 0.9% [23] road accidents are connected to indeterminate road surface conditions specified as *other*. The implementation of the result acquired from the analysis of the *Quality of road surface,* probably may require a revisit to the location of the accident to simplify the clues leading to the incident.

The classification of these variables reveals that accidents connected to the *'Bumpy'* road surface is located within the built-up area with a speed limit below 60km/hr, while other variables like *Pothole, Cracks,* and *Corrugated* road surface can be located in any areas within a speed limit below or above 60km/hr. The accidents attributed to this set of variables could be ascribed to two main factors enumerated below:

- Inability of the road users [drivers/motorcyclists] to observe any danger ahead while driving on a high-speed on the road, and
- Inability of the road traffic management to educate the road users [drivers/motorcyclists] as per the condition of the road they travel upon.

5.4 Introduction to the analysis of the non-captured accident data in Road related factors

In this section, the results of the average estimates and histograms of the non-captured data analysed in the *Road related factors* are discussed. The methods implemented in computing the two descriptive statistical analyses were explained to details in the subsections 4.3.1 and 4.3.2 above. In addition, average estimates computed illustrate the mean score of all the non-captured data assembled for each field in the *Road related factors* over the period of 12 months [see Table 21]. The mean scores represent the expected score of unmanaged data detected in each field [see Figure 34].

In the case of the histogram, the necessary approach applied to develop the required parameters was earlier discussed in the subsection 4.3.2 above. The results obtained from the two descriptive statistical analysis are accurately represented in the charts presented in the Figure 34, Figure 35 and Figure 36. The motive behind the separation of the data fields into two tables in the analysis of the histogram was simplified in the subsection 4.3.2 above, along with the procedures required towards a practical distribution of the accidents scores procured across the thirteen variables/data fields [see Table 22 and Table 23].

5.4.1 Average estimates of the non-captured accident data in Road related factors

This section discusses the average estimates of the non-captured data computed for the thirteen related data fields categorised under the Road related factors. The total average estimate of the data mismanaged in the Road related factors is slightly higher than the total average estimate data mismanaged in the Accident related factors [refer to subsection 4.3.1]. Considering the variation observed in the mean scores assembled in the Table 21, one can simply discover that minimal amount of data is mismanaged in some certain variables/data fields, like *Road surface type, Quality of road surface, Road surface,* and *Road marking visibility.*

The mean scores computed for the aforementioned variables/data fields demonstrate that a minimum quota of data was inappropriately represented in the ARF for the four data fields during the data collection activities. *Road surface*, among the aforementioned variables, lost an average estimate of 4 [0.7%] data points every month to anomalies existing along the data collection procedures. This amount is considered as the lowest mean score among the four variables, followed by the *Quality of road surface* with a mean score of 5 [1.0%] non-captured data. Also, two other road related data fields demonstrated a fair minimum loss of data, such data fields as the *Road surface type* and *Road marking visibility* yielding mean scores of 6 [1.2%] and 7 [1.5%] non-captured data respectively.

Table 21: Estimated average numbers of non-captured accident data in Road related factors

Average estimates of non-captured accident data in Road related factors		
Road related factors	**Average estimates**	**Percentage estimates**
Speed limit	53	11.0%
Built-up area	78	16.1%
Road type	70	14.5%
Junction type	105	21.5%
Road surface type	6	1.2%
Quality of road surface	5	1.0%
Road surface	4	0.7%
Road marking visibility	7	1.5%
Obstructions	23	4.6%
Overtaking control	33	6.8%
Traffic control type	77	15.8%
Road signs clearly visible	12	2.4%
Condition of road signs	14	2.8%
Total average estimate	**486**	

Some other variables/data fields generated higher average estimates of non-captured data [see Figure 34]. These estimates reflect a great loss of road accident data during the data evaluation process. This implies that some of these categorical variables [data fields] in the Road related factors are left incomplete in the ARF during the data collection process. In addition, in Figure 34, the variations observed demonstrate a huge inconsistency in the representation of the data fields. Among all the data fields with missing data, *Junction type* is the most frequently omitted variable/data field, with an average number of 105 [21.5%] non-captured data. Other variables/data fields with high amounts of non-captured data are *Built-up area, Traffic control type, Road type,* and *Speed limit,* with average estimates of 78 [16.1%], 77 [15.8%], 70 [14.5%] and 53 [11.0%] non-captured data respectively.

Some other variables/data fields, like *Overtaking control, Obstructions, Road signs clearly visible,* and *Condition of road signs* produced less than 34 non-captured data each. As displayed in Figure 34, it is clear that less data entries are lost in the road surface related data fields compared to other related data fields in the Road related factors. This indicates that most reporting officers have better insight into the data fields in the road surface related data fields.

On the contrary, in the case of the variables/data fields with large average estimates of non-captured data, such data fields as the *Road type, Junction type, Built-up area* and *Speed limit on road*; one could conclude that an insubstantial approach was exercised towards the entering of data. This led to high amounts of data omission and misrepresentation issues attributable to such errors as *'Item non-response error*' and *'response error'* detected during the assessment of the completed ARF.

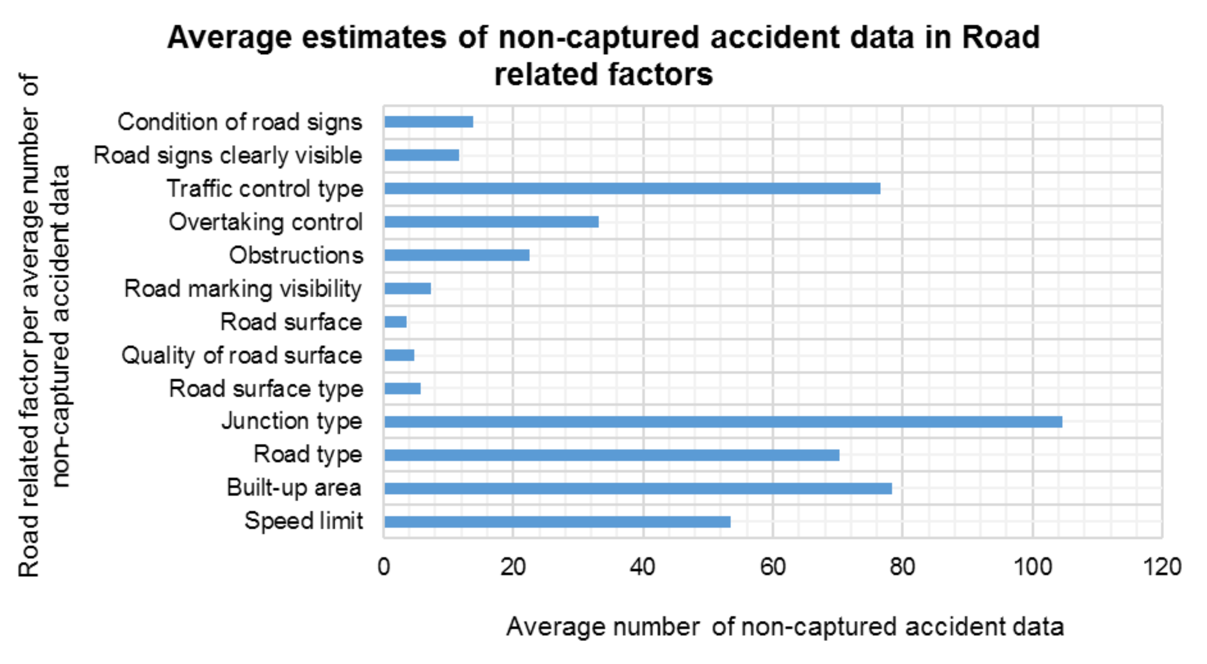

Figure 34: Total average estimates of non-captured accident data in Road related factors

Besides, a large percentage of errors discovered are ascribed to the late reportage of RTAs; where some accidents are reported to the local traffic department at a later time after the actual period of the incident. This condition, most times, reduces the ability to acquire more information about the incident, unless a reporting officer reaches the accident scene right in time, before the vehicles involved are towed away, or the accident victims are moved away from the scene. This case is defined by the situation encounter at the scene of the accident. Some road users find it difficult to distinguish the responsibilities of the traffic department from that of the police department when it comes to the issue of accident reportage, because many times police's attention have been called to incidents involving property *'damage only'.*

Actually, many instances of accidents occurring in parking lots are not properly represented in the data form. In essence, these accident cases are not perfectly indicated in the case of *Road type* and *Junction type.* The answer options specified in the two related data fields are expected to correlate with the options specified under the *Vehicle manoeuvre*. This simply infers that it is difficult for some reporting officers to comprehend the proper interpretation of some accident circumstances. This contributes to the inadequate completion of ARF regarding some relevant data fields. Incomplete data hampers the decision-making process, and hinders the distribution of resources to combat accidents caused by these particular related data fields.

5.4.2 Histogram analysis of the non-captured accident data in Road related factors

In this section, typically, two separate tables containing the estimated scores of the non-captured data discovered in the *Road related factors* are presented, with the purpose of computing a suitable choice of the class interval/bin for a reliable result. Ultimately, an appropriate distribution shape, offering practical insight into the structure of the estimated scores assembled was achieved through the selection of a suitable class interval. The scores organised in Table 22 and Table 23 illustrate monthly accumulation of the non-captured data observed in the related data fields.

Table 22: Estimates of non-captured data for data fields with least missing data in Road related factors

Total estimates of non-captured data for least missing data in Road related factors							
Months	**Road surface type**	**Quality of road surface**	**Road surface**	**Road marking visibility**	**Obstructions**	**Road signs clearly visible**	**Condition of road signs**
Jan	6	2	2	5	5	10	9
Feb	11	1		5	41	9	8
Mar	6	2		7	25	18	18
Apr	5	9	6	10	16	5	9
May	4	9		5	19	10	14
Jun	5	5	3	11	20	16	17
Jul	5	5	2	3	11	11	10
Aug	3	6	4	7	26	11	11
Sep	3	5	5	12	26	13	16
Oct	12	7	4	9	36	10	23
Nov	3	3	3	8	27	13	16
Dec	7	3	3	5	19	14	14

From the Table 22, within the 12-month period, only *road surface* among all the thirteen related data fields is perfectly and absolutely completed in February, March and May during the data collection proceedings. The outcome of the accuracy is denoted as empty cells *[no errors are detected]* in the Table 22. The magnitude of the non-captured data grouped in the *Road surface related data fields,* such data fields as *road surface type, quality of road surface* and *road surface* reflects minimal omissions/errors.

The table above presents set of estimated scores generated for the frequency analysis of non-captured data per field, within the frequency range of 1 to 41 non-captured data. The estimates offer a simple formation of the class interval [bin] towards a practical understanding of the frequency distributions of the non-captured data.

However, the distribution shape obtained in this section is similar to the distribution shape presented in Figure 24, but the slight difference in the shape shows a *gap* in the distribution actualised. Observations from Figure 35, reveals that the distribution shape further exhibited a peak score of 30 counts within the lowest range of non-captured data. This produces a J-shaped distribution chart, and positively skewed to the right side of the chart, wherein each bar represents the amount of times a particular range of non-captured data is counted.

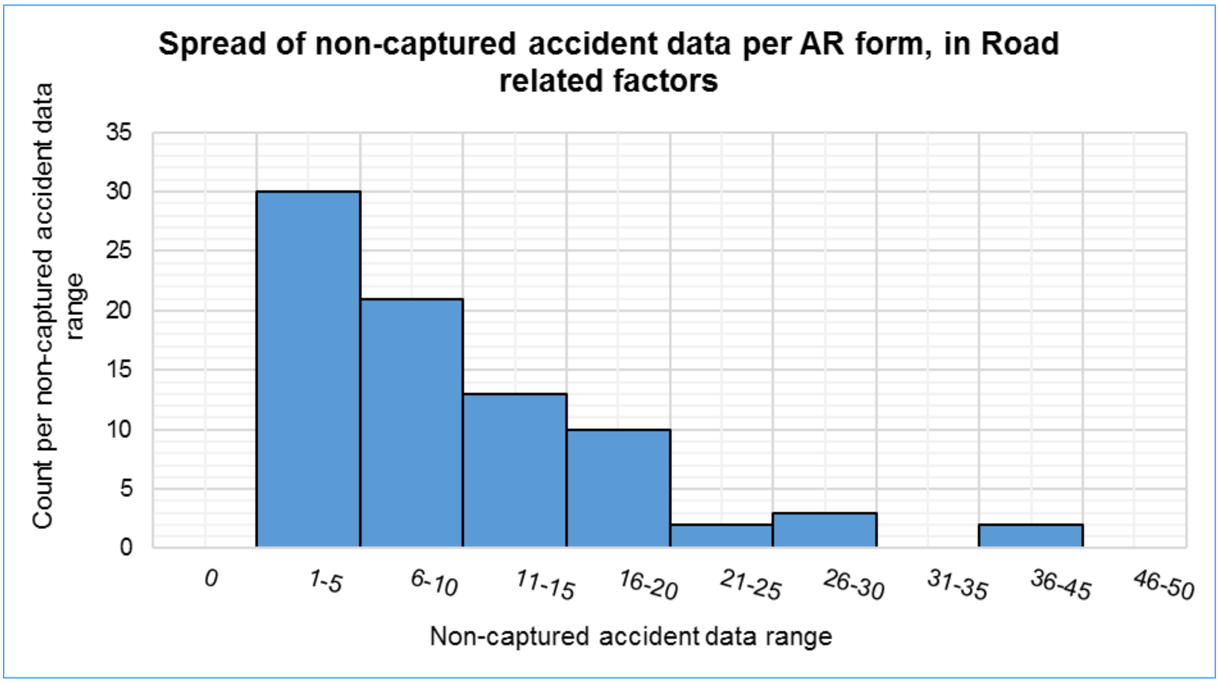

Figure 35: Distribution of non-captured data for data fields with least missing data in Road related factors

The result displayed in the chart demonstrates that high counts per non-captured data clustered within the ranges of 1 to 20 non-captured data per month. However, fewer counts per non-captured data are observed within the ranges of 21 to 45 non-captured data, which appeared farther to the right tail of the histogram. Out of the seven data fields presented in the Table 22, only three data fields contributed massively to the ranges with large scores of non-captured data per month, such data fields as *Obstruction, Road signs clearly visible,* and *Condition of road sign.*

Conversely, other six related data fields characterised with large scores of non-captured data in the *Road related factors* are displayed in the Table 23. The lowest and highest scores observed in this table are 20 and 156 non-captured data. The distribution shape presents a clustered distribution of high scores without a gap unlike the previous chart.

The frequency distributions with high counts fall mostly within the high ranges of non-captured data. From a graphical illustration, highest peak of 19 counts demonstrate a large estimate of data mismanaged within the range of 61-80 non-captured data. Observably, a frequency estimate of 14 counts is observed within the range of 41-60 non-captured data, followed by frequency estimate of 12 counts within the ranges of 21-40 and 101-120 non-captured data respectively. A frequency estimate of 10 counts is observed within the range of 81-100 non-captured data. The high frequency estimates presented in Figure 36, reveals that high amount of data is mismanaged in the Road related factors.

Table 23: Estimates of non-captured data for data fields with most missing data in Road related factors

Total estimates of non-captured data for most missing data in Road related factors						
Months	**Speed limit**	**Built-up area**	**Road type**	**Junction type**	**Overtaking control**	**Traffic control type**
Jan	49	75	68	103	24	61
Feb	78	110	104	156	39	67
Mar	51	92	78	124	34	113
Apr	22	47	59	79	20	82
May	71	107	114	130	31	84
Jun	58	77	70	97	30	81
Jul	72	103	81	110	20	51
Aug	86	104	81	117	34	82
Sep	53	75	73	106	52	69
Oct	41	61	46	75	50	94
Nov	35	64	45	104	44	69
Dec	25	25	25	54	21	67

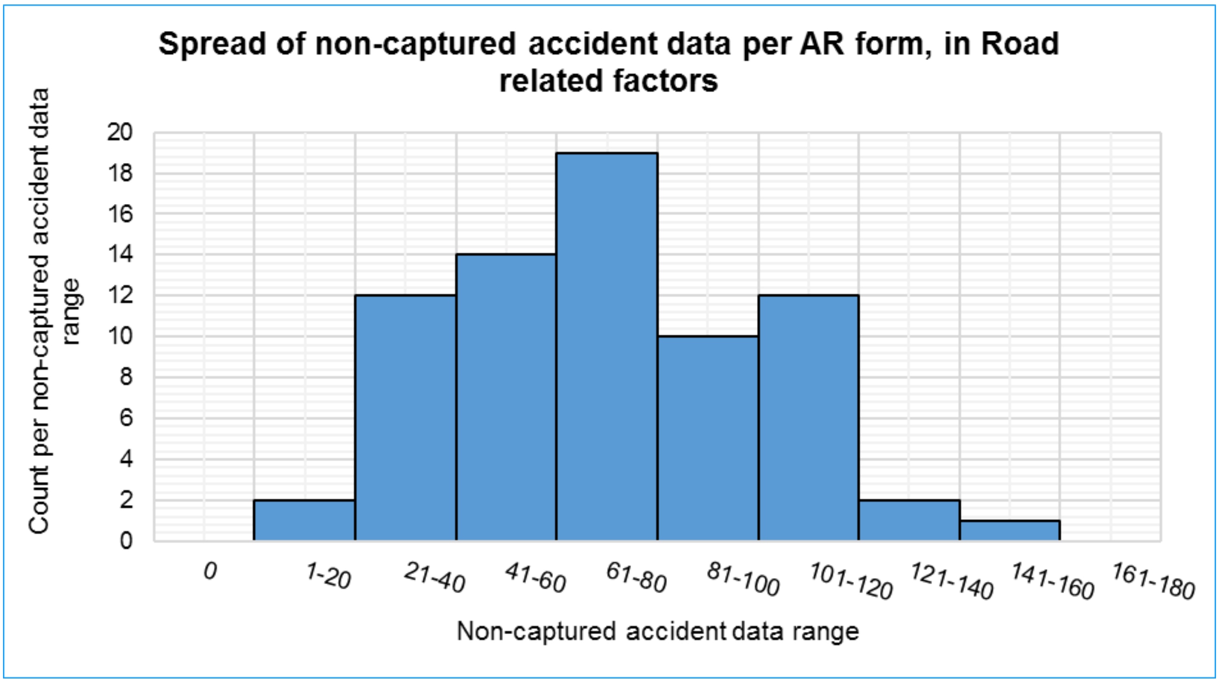

Figure 36: Distribution of non-captured data for data fields with most missing data in Road related factors

On the contrary, some ranges produced fewer counts with a frequency estimates below 4 counts. From the illustration given in the chart above, fewer frequency estimate of non-captured data are observed within the range of 1-20, 121-140, and 141-160 respectively. The results obtained demonstrate the incompetence of the reporting officer towards completing of the six data fields, with high counts of non-captured data in Road related factors because many data elements are misinterpreted, and respectively omitted in the process of data collection.

Among the six data fields presented in the table above, *built-up area, road type, junction type* and *traffic control type* contributed extremely high scores of non-captured data [refer to subsection 5.4.1 above]. The consequence of omitting or excluding a high estimates of data elements, rendered much of the data gathered ineffectual. This limits the possibility of correlating the findings with the actual circumstances of the road accidents.

6. Analyses and findings in Human related factors

This chapter discusses the analysis of the personal attributes of the persons [drivers/cyclists] involved in the RTA within the Stellenbosch locality, along with the necessary details required to determine the degree of safety conditions. In this section, it is necessary to analyse the usable data in the *Human related factors*. In addition, the discussion encompasses the relational data fields categorised under the Human related factors, such data fields as *Ages of drivers/cyclists, Nationalities of drivers/cyclists, Gender and Race of drivers/cyclists, Safety measures related factors*, and *Vehicle manoeuvre [what driver was doing].* Additional analysis and findings as regards some of these related factors are provided in the Appendix D.

The analysis of the human related factors identifies the major areas that require close observation or attention of the road traffic management. In essence, some factors like age, gender, race and nationality, can be used to measure the most vulnerable group in the RTAs in the Stellenbosch area.

However, as part of this chapter, the non-captured data will be analysed to provide more insight into the degree to which data are mismanaged in the *Human related factors*. This is achieved through the calculation of the average estimate of non-captured data and the formulation of a suitable class interval to support a practical frequency distribution.

6.1 Analysis of Ages of drivers/cyclists involved

This section discusses the ages of the persons involved in the RTAs in the Stellenbosch area in 2012. The data points presented in the Table 24 represent the ages of both the drivers and cyclists involved in the RTAs, which is characterised with the ages of the foreign national and the South African drivers/cyclists. The data was analysed by applying descriptive statistics wherein the mean and distribution of the ages of the road users involved is simply determined.

In the process, a class interval difference of 20 was considered in the computation of the distribution of the age groups, as calculated between the intervals of 20 to 100[24]. This approach was considered in order to include all known age groups involved in the traffic accidents as indicated in the form. A clear distribution pattern of the data points is shown in Figure 37 below. However, from the distribution of the scores presented in the Table 24, one can simply observe that age groups between 21-40 and 41-60[25] are more frequently represented in the data form.

[24] Refer to subsection 3.3.2.5.1 for more details.

[25] Both age groups are combination of youths and middle-ages.

The data specifies that the aforementioned age groups characterised the group of drivers/cyclists that largely involved in the RTAs in SM. In addition, few records of drivers/cyclists grouped within the age group of 81-100 are obtained since this age group limitedly used the roads, unlike the two age groups mentioned previously.

In this section, two charts are plotted to illustrate the ages of the drivers/cyclists involved in the RTAs within the Stellenbosch area in 2012. The chart shown in Figure 37 is known as *a box and whisker chart,* which demonstrates a wide spread of the age distribution of the drivers/cyclists involved in road accidents from the period of January to December the same year. More so, the chart presents the *maximum age per month, minimum age per month, and quartiles' age [first quartile age, median age, and third quartile age] per month.* Alternatively, the second chart demonstrates overall estimates of the age distributions of drivers/cyclists based on the percentage scores presented in the last column of the Table 24.

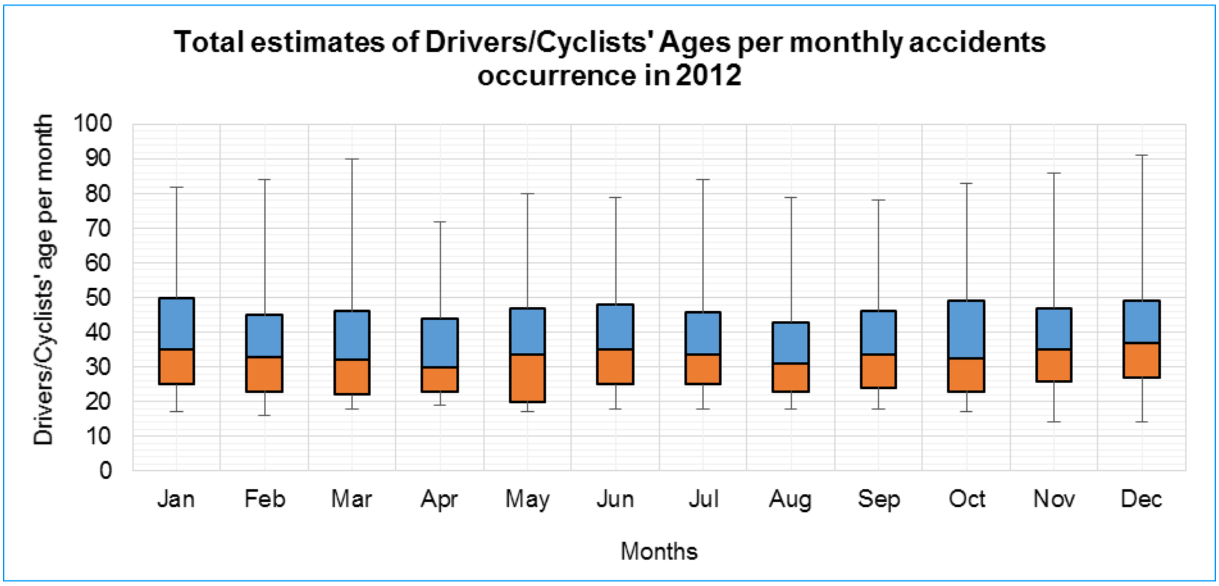

Figure 37: Total estimates of the Drivers/Cyclists' ages per month

From the third tabular section of the table, the scores calculated for maximum age, minimum age, and quartiles' age are arrayed in monthly order to illustrate the summarised distribution of the ages of the drivers/cyclists involved in the RTAs. The maximum ages obtained in each month fall within the ranges of 61-80 and 81-100 age groups, which illustrates that some old age drivers/cyclists are involved in RTAs. However, considering the minimum ages, values obtained in each month fall within the range of 14-20 age group. According to the result, it is understood that a young cyclist was involved in road accidents in Stellenbosch in the same year, whose age is below 15 years.

Table 24: Drivers/cyclists' ages estimates in road accident occurrence

Total estimates of Drivers/Cyclists' ages involved in the RTA occurrence in 2012														
Age groups	**Jan**	**Feb**	**Mar**	**Apr**	**May**	**Jun**	**Jul**	**Aug**	**Sep**	**Oct**	**Nov**	**Dec**	**Total scores**	**Percentage estimates**
	Frequency													
14-20	17	37	55	27	29	18	12	40	21	35	22	10	**323**	**9.4%**
21-40	107	175	182	159	171	139	125	184	164	179	165	109	**1859**	**54.4%**
41-60	67	94	104	74	107	79	72	75	76	93	84	59	**984**	**28.8%**
61-80	15	22	26	11	15	14	12	22	24	27	27	23	**238**	**7.0%**
81-100	3	3	4	0	0	0	1	0	0	2	2	1	**16**	**0.5%**
Mean ages	38	36	36	35	36	37	37	35	37	37	38	40		
Standard Error	1.04	0.82	0.80	0.82	0.78	0.85	0.93	0.80	0.88	0.83	0.87	1.08		
Maximum, Minimum and Quartiles														
Max ages	82	84	90	72	80	79	84	79	78	83	86	91		
Min ages	17	16	18	19	17	18	18	18	18	17	14	14		
First Quartile [Q1] ages	25	23	22	23	20	25	25	23	24	23	26	27		
Median [Q2] ages	35	33	32	30	33.5	35	33.5	31	33.5	32.5	35	37		
Third Quartile [Q3] ages	50	45	46	44	47	48	45.75	43	46	49	47	49		
Box and Whisker														
Q1-Min	8	7	4	4	3	7	7	5	6	6	12	13		
Q1	25	23	22	23	20	25	25	23	24	23	26	27		
Q2-Q1	10	10	10	7	13.5	10	8.5	8	9.5	9.5	9	10		
Q3-Q2	15	12	14	14	13.5	13	12.25	12	12.5	16.5	12	12		
Max-Q3	32	39	44	28	33	31	38.25	36	32	34	39	42		

Observably, in the second tabular section of the table, the mean scores [mean ages] and the standard errors of the age distributions are presented from January to December. The scores displayed in the mean ages demonstrate a slight increment above the scores presented for the median ages. Actually, from the age distribution, average ages calculated fall within 35 to 40 years of age across the 12 months. These average ages fall absolutely within the age group of 21-40, that is, group of youths and partially middle-ages (Traffic Accident Statistics 2002; Sinclair & Murdoch 2012; IRTAD 2013). This insinuates that youths and middle age drivers/cyclists are predominantly involved in the RTAs in the Stellenbosch area than any other age groups.

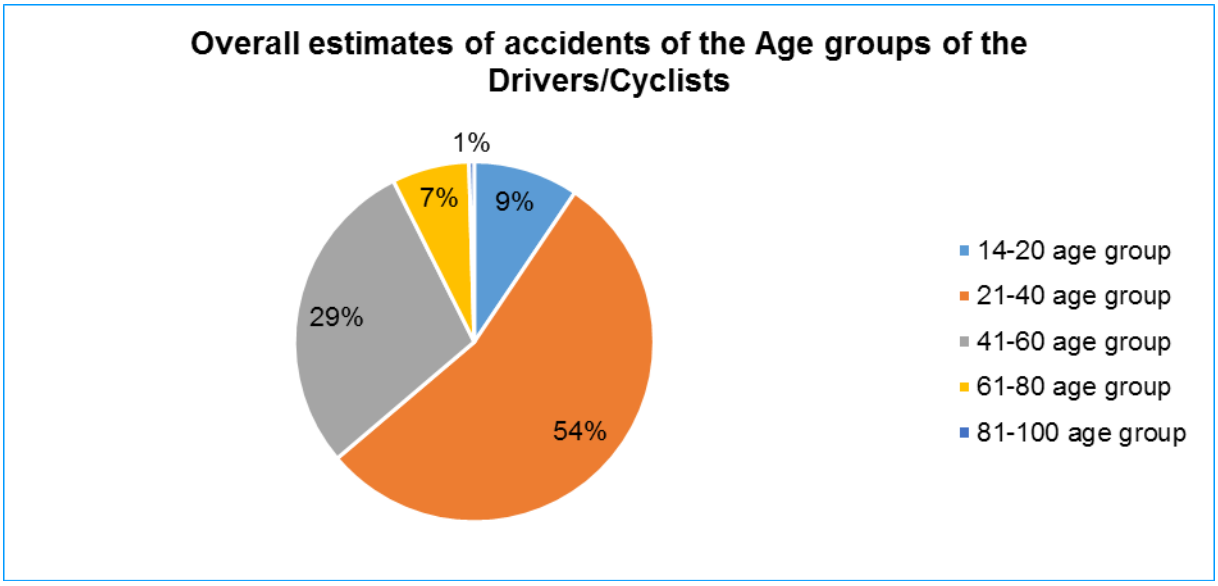

Figure 38: Total percentage estimates of accidents of the Age-groups of the Drivers/Cyclists in 2012

The same statement can be supported by the results demonstrated in the overall estimates of the age groups presented in the Figure 38. In the chart, it is illustrated that 54.0% of the drivers/cyclists involved in the road accidents within SM falls within 21-40 age group. This validates the actual age groups comprising a high number of students, manual workers and residents who are drivers/cyclists travelling around the Stellenbosch area. This certain age group contributed to the high traffic volumes, rise in the number of the road users, and high number of road accidents which may be as a result of incompetence, substance abuse, traffic policies violation etc. A total estimate of 29.0% of the drivers/cyclists fall within the age group of 41-60, which are exclusively the middle-age drivers/cyclists. With the intention of validating the dominance of this age group on the road usage in another municipality, however, findings obtained by Traffic Accident Statistics (2005) illustrated that 26-35 and 36-45 age groups are most vulnerable age groups in the City of Cape Town Metropolitan Municipality (Traffic Accident Statistics 2002; Traffic Accident Statistics 2005). Although an even interval was not considered across the age groups in the report, but practical findings are acquired [refer to

Appendix D.2 for drivers' Age classification]. The remaining 17.0% quota of the overall estimates is occupied by three other age groups, which are age groups with limited involvement in the RTAs in the Stellenbosch area, such age groups as 14-20, 61-80, and 81-100.

6.2 Analysis of Countries of drivers/cyclists involved

This section presents the analysis of the drivers'/cyclists' nationality as part of the influential factors contributing to the cause of accidents in the Stellenbosch area due to different driving systems. Table 25 comprises the analysis of both of the local drivers and foreign drivers' contribution to the RTAs occurrence on the Stellenbosch roads. The analysis of the result obtained is based on the concept of *'who lives and works within the Stellenbosch locality',* with the consideration of travellers from countries like Germany, Namibia, Netherlands, Zimbabwe *'who only visit the locality during holidays, or travel to workplace/business associates.'* It is understood through a complaint made by a superior traffic officer that most foreign nationals driving on the Stellenbosch roads have little knowledge of the driving system in South Africa. This issue is due to the different driving systems adopted by various countries. The traffic officer further suggested that most of these foreigners are considered to be students, tourists, and small-scale business owners.

Table 25: Estimates of Countries of drivers/cyclists involved in road accident

Total estimates of foreign and local drivers/cyclists involved in the RTA in 2012		
Foreign and local drivers/cyclists	**Scores**	**Percentage estimates**
Foreign drivers	143	5%
South African drivers [Stellenbosch drivers]	2939	95%
Total score	**3082**	
Total estimates of Countries of drivers/cyclists involved in the RTA in 2012		
Foreign drivers/cyclists	**Scores**	**Percentage estimates**
Germany	20	14%
Namibia	30	21%
Netherlands	11	8%
Zimbabwe	27	19%
Others	55	38%
Total score	**143**	

With the purpose of curtailing the involvement of the foreign drivers/cyclists in the RTA, the Stellenbosch Traffic Department [STD] urged them to undergo the driver's training to strength their level of understanding towards the driving system applied in South Africa. The results displayed in the table above are divided into two tabular sections. The first tabular section contains the comparison of both the indigenous and the foreign drivers/cyclists, while the second tabular section contains the percentage estimates of the foreign drivers/cyclists only, by ascertaining the exact nationalities that predominantly involved in the RTAs in the Stellenbosch locality. The results obtained are displayed in the Figure 39 and Figure 40.

In the first tabular section, as presented in the chart below, a small number of foreign drivers is involved in the RTA in the Stellenbosch area in 2012. This illustrates that only 5.0% of accidents have foreign drivers/cyclists involved in the Stellenbosch locality throughout this period. However, a large amount of South African drivers, probably Stellenbosch indigenous drivers, was documented as the highest victims of the road accidents in the area. Statistically, the quota reflects the result computed for the South African drivers, which is 19 times greater than the result computed for the foreign drivers.

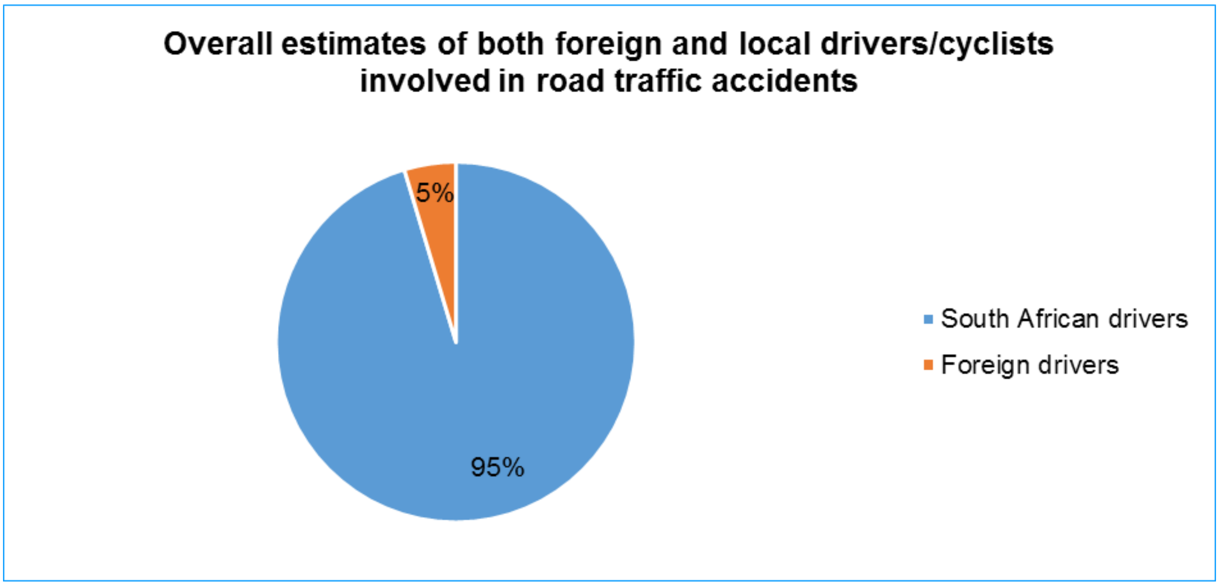

Figure 39: Total estimates of foreign and local drivers/cyclists involved in accidents in 2012

From a simplified analysis of the foreign drivers/cyclists involved, according to the results illustrated in Figure 40, Namibian drivers recorded the highest percentage estimate of 21.0% in the accidents, followed by Zimbabwean and German drivers with percentage estimates of 19.0% and 14.0% respectively. Clearly, the high involvement of the three nationalities in the occurrence of RTAs could be attributed to the immigration influx effect, and the need to travel.

Additionally, Netherlands among all the rest of the countries, contributed 8.0% in the accidents. This quota is not the lowest of all, but higher than the percentage estimate computed for each nationality grouped under the variable identified as *Others,* such countries as Australia, Nigeria, France, Belgium, and many others as United States, United Kingdom, Spain, Zambia

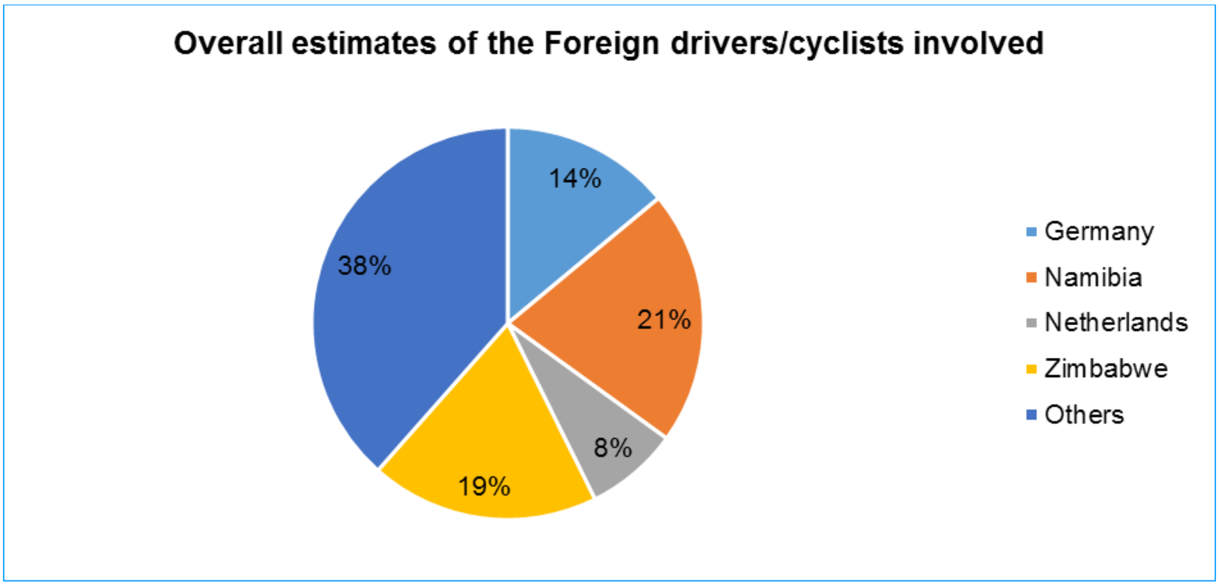

Figure 40: Total estimates of the Foreign drivers/cyclists involved in accidents in 2012

etc. This particular variable contains the sum of percentage estimates of 38.0%, which is the combination of the percentage estimates of all other nationalities with minimal involvement in the road accidents in Stellenbosch locality. Among the countries listed above, only few of them adopted driving system similar to that of South Africa.

6.3 Introduction to the analysis of the drivers'/cyclists' Gender and Race

This section discusses the Gender and Race of the drivers/cyclists involved in the RTAs in the Stellenbosch locality. The analysis results generated from the two related data fields are presented in the subsections below. The *Gender* consist of three related variables which are the *Male, Female,* and *Unknown*. These three related variables are analysed statistically according to their involvement in the RTAs. In addition, considering the *Race* of the drivers/cyclists, six related variables are defined in accordance with their contributions in the RTAs. The six variables considered here are *Asian, Black, Coloured, White, Other* and *Unknown*. However, among the six related variables mentioned, *Other* and *Unknown* are defined according to the condition of the accident, and the discretion requested by the victim not to disclose the *race* he/she belongs to.

6.3.1 Analysis of the Gender of drivers/cyclists involved

In this section, the scores arrayed in Table 26 represent the outcome of the analysis performed on the three related variables grouped in the Gender analysis. Observably, the result obtained demonstrates that 66.0% [2,244] *Male* drivers/cyclists residing or working in the Stellenbosch locality were involved in the accidents. This amount is double the total score calculated for

Female drivers/cyclists, which confirms that *Male* drivers/cyclists are more vulnerable to road accidents (Sinclair & Murdoch 2012). In fact, this could be ascribed to the behavioural characteristics of the *Male* drivers/cyclists by being impatient on the road, or not being careful while driving on Stellenbosch roads.

In addition to this statement, *Female* drivers/cyclists living or working in the Stellenbosch area are considered less vulnerable to the RTAs than their male counterpart [see Figure 41]. In the chart, a total estimate of 33.0% [1,137] *Female* drivers/cyclists were reported to be involved in the RTAs in 2012. This illustrates that *Female* drivers/cyclists are considered more cautious while travelling on the road than their *Male* counterparts. This result can also be argued to be because the Male drivers/cyclists outnumbered the Female drivers/cyclists in the Stellenbosch locality. A less significant amount of undefined gender identity of the drivers/cyclists involved in road accidents is calculated for Unknown [see Figure 41].

Table 26: Gender estimates in road accident occurrence

Total estimates of Gender in the RTA occurrence			
Month	**Gender categories**		
	Male	**Female**	**Unknown**
Jan	143	73	2
Feb	199	126	2
Mar	246	114	1
Apr	151	102	0
May	201	120	2
Jun	179	76	2
Jul	157	64	2
Aug	223	106	2
Sep	183	100	0
Oct	233	101	2
Nov	193	102	1
Dec	136	53	0
Total scores	**2244**	**1137**	**16**

Figure 42 presents the monthly distribution of the scores calculated for the *Gender* categories. The chart demonstrates the variations observed in the monthly estimates of the three variables grouped in the *Gender*. The highest number of *Male* drivers/cyclists involved in RTAs in the Stellenbosch locality are recorded mostly in March, with a total estimate of 7.2% [246] followed by October and August, with total estimates of 6.9% [233] and 6.6% [223] respectively.

Considering the Female drivers/cyclists, fewer accidents are recorded from January to December. According to the result, the highest number of *Female* drivers/cyclists involved in road accidents was observed in February with a total percentage estimate of 3.7% [126].

Actually, in both March and May, similar estimates of 3.4% [114] and 3.5% [120] were documented respectively.

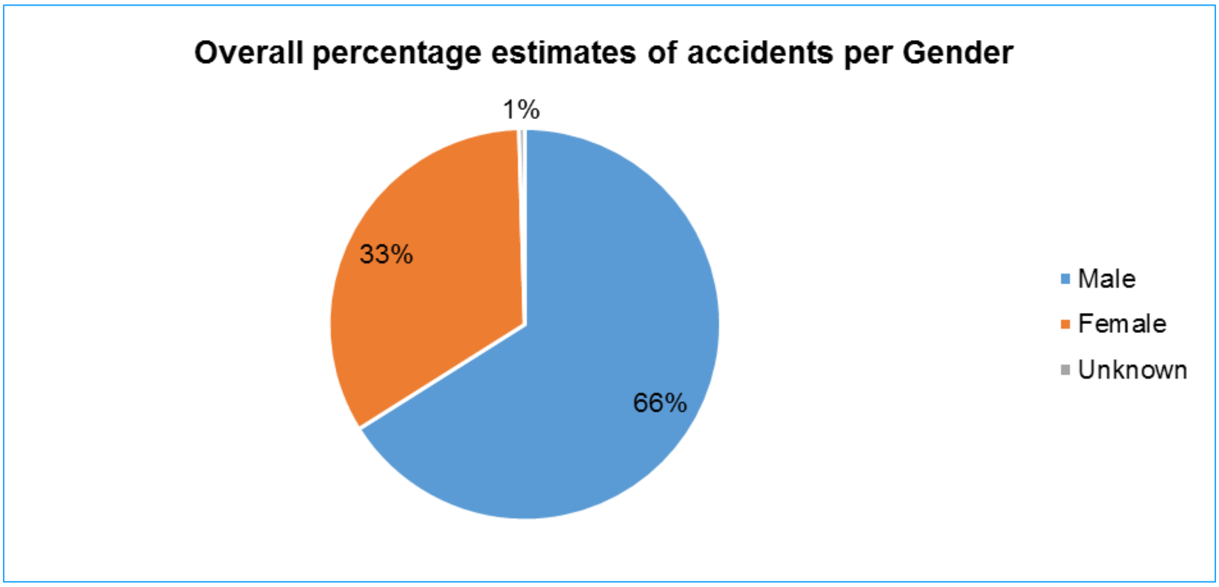

Figure 41: Total percentage estimates of accidents per Gender in 2012

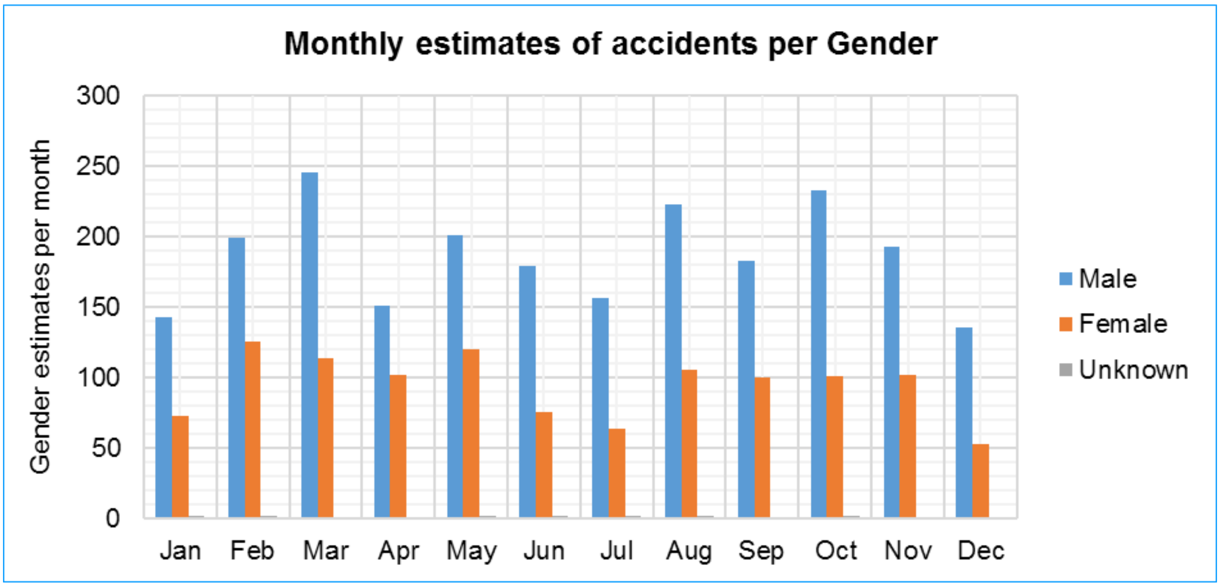

Figure 42: Total estimates of monthly accidents per Gender in 2012

The trend observed in the chart demonstrates an increment in February for both *Male* and *Female* drivers/cyclists. Although, the involvement of the *Male* drivers/cyclists in the road casualties in the Stellenbosch locality dropped in April while the involvement of the *Female* drivers/cyclists dropped in March. A similar trend pattern was observed from April to September for both *Male* and *Female* drivers/cyclists. From September to November, a practically constant number of *Female* drivers/cyclists involved in the road accidents were observed. On the other hand, within the same periods, the variations observed in the involvement of the *Male* drivers/cyclists in the road accidents demonstrated a gradual decline in trend. Overall, the involvement of the *Male* drivers/cyclists are much higher than that of their female counterpart in 2012.

6.3.2 Analysis of the Race of drivers/cyclists involved

This section discusses the outcome of the analysis carried out basically on the group of people involved mostly in the road accidents within the Stellenbosch area in 2012. This particular field consist of six related variables, and it represents the definitive identification of the race of people involved in the road accidents in the Stellenbosch area. These variables define the categories of the drivers and cyclists, who are inhabitants, visiting and/or working in the Stellenbosch locality. The six related variables considered here are *Asian, Black, Coloured, White, Other* and *Unknown.* Regarding the description of Other and Unknown consult the section 6.3 above.

Table 27: Race estimates in road accident occurrence

Total estimates of Race in the RTA occurrence in 2012						
Months	**Asian**	**Black**	**Coloured**	**White**	**Other**	**Unknown**
Jan	2	25	74	107	2	2
Feb	4	32	75	212	1	2
Mar	3	34	92	207	2	2
Apr	1	32	53	167	0	1
May	1	31	82	208	0	2
Jun	1	26	94	127	1	2
Jul	0	30	68	117	1	1
Aug	0	38	86	203	0	1
Sep	2	41	103	183	0	0
Oct	2	39	74	168	0	1
Nov	2	39	74	168	0	1
Dec	0	34	49	102	0	0
Total scores	**18**	**401**	**924**	**1969**	**7**	**15**

The analysis performed in this section reveals that *White* drivers/cyclists are mostly involved in RTAs, more than any other group inhabiting, visiting and/or working in the Stellenbosch locality. Evidently, this group of people are considered the second highest inhabitant to the Coloured people in the Stellenbosch Municipality in terms of population estimates published in 2001, 2006, and 2007 (Maree & Daniels 2007; Vanderschuren & Jobanputra 2011; Stellenbosch Municipality 2012). They are exclusively the major race group with the highest involvement in the road accidents within the locality. This could be benchmarked with the available estimation of the vehicle ownership records warehoused in the unit responsible for motor vehicle registrations.

The accidents scores generated for the six variables are assembled in Table 27. Observably, a total estimate of 59.0% [1,969] is calculated for accidents involving White drivers/cyclists, followed by both the Coloured and the Black drivers/cyclists with total estimates of 28.0% [924] and 12.0% [401] each. More so, a total sum of 1.0% [40] is calculated for the remaining three variables [see Figure 43 below]. This suggests that the *White* and the *Coloured* drivers/cyclists travel the roads around the Stellenbosch locality more frequently than any other race group (Vanderschuren & Jobanputra 2011).

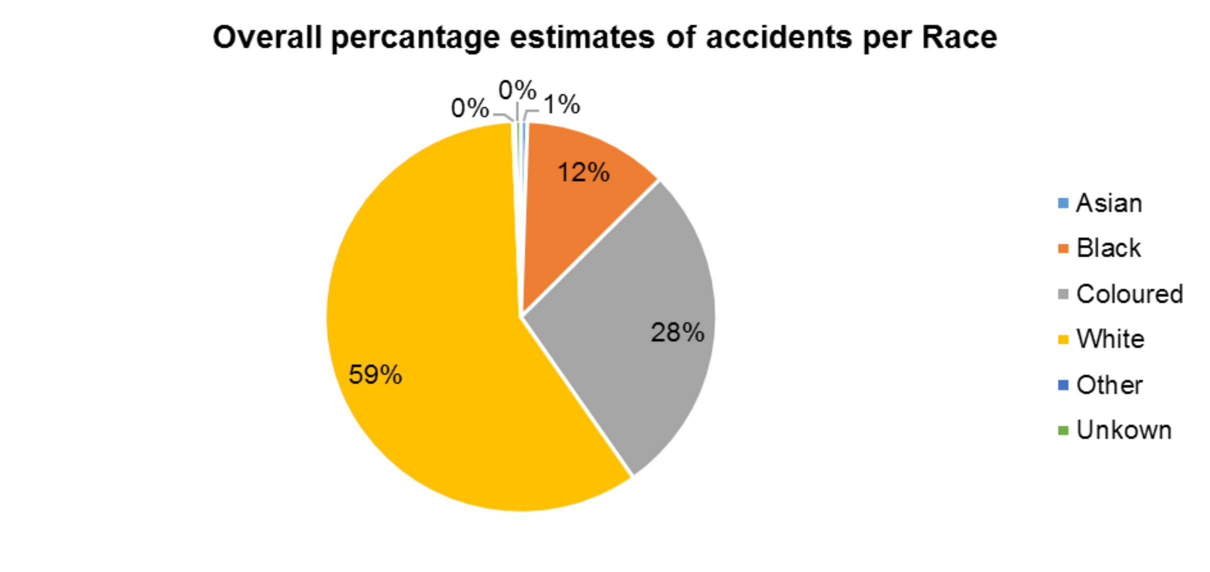

Figure 43: Total percentage estimates of accidents per Race in 2012

A distribution chart is plotted to illustrate the spread of the scores in the table above across 12 months. The trend pattern demonstrated in Figure 44 offers understanding into the variations observed in the monthly estimates of data assembled in the six related variables. Actually, the discussion is centred on the three major variables with significant estimates, such variables as White, Coloured, and Black drivers/cyclists. Observably, an increase in the Coloured and Black drivers'/cyclists' involvement in the road accidents was observed from January to March. A slight increment in the White drivers/cyclists is observed from January to February.

Among the three variables, only the White and Coloured drivers/cyclists exhibited similar trends in the chart displayed above. Apparently, the highest involvement of the *White* drivers/cyclists in the occurrence of road accidents in 2012 are recorded in February with a score of 212 drivers/cyclists. This amount covers a percentage estimate of 6.4% out of the total estimates calculated across 12 months for all the six related variables. From March to May, a convex curve indicating variations in the involvement of both the White and Coloured drivers/cyclists in the road crashes in the Stellenbosch locality is demonstrated. However, it is a different case for the Black drivers/cyclists, instead a gradual decrease is observed.

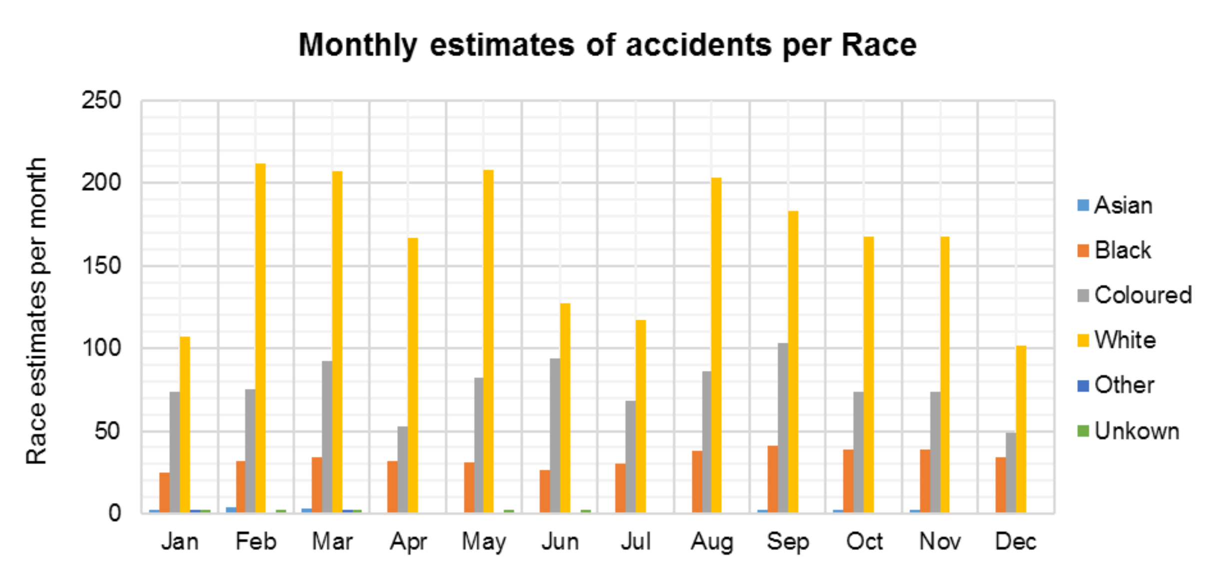

Figure 44: Total estimates of monthly accidents per Race in 2012

A drastic decline in the involvement of the White drivers/cyclists in the road crashes is observed from May to July. In the case of *Coloured* and *Black* drivers/cyclists, convex and concave trends are demonstrated from May to July. On the other hand, towards the end of the year, equal amounts of data points, at different estimates, were observed in the six related data fields from October and November. The highest involvement of the Coloured and Black drivers/cyclists in the road crashes was recorded in September, covering percentage estimates of 3.1% [103] and 1.2% [41] respectively across 12 months.

In addition to this, the overall sum of captured data calculated for *Gender,* is not equal to the overall sum of the captured data computed for *Race.* Therefore, a small estimate of 1.0% is lost in the categories of Race. This consequence is as a result of inconsistencies experienced along the process of completing the accident details provided in the ARF.

6.4 Introduction to the analysis of the non-captured data estimates in Human related factors

This section discusses the average estimates and the frequency distribution [histogram] of the non-captured data discovered in relation to the assessment of the related data fields grouped under the *Human related factors*. The methods used to analyse the average estimates and the histogram of the non-captured data in the Human related factors were previously explained in the subsections 4.3.1 and 4.3.2 above[also see subsection 3.3.2.5.2 for full details]. In the subsequent sections, the interpretation of the findings deduced from the average estimates and histogram of the non-captured data will be discussed in details. The average estimates cut across 10 related data fields, with the purpose of computing an approximated amount of non-captured data per each field. Similarly, the histogram developed provides a better understanding into the distribution of the non-captured data estimates across the 10 related data fields. However, two tables containing the estimated scores of the non-captured data are presented in subsection 6.4.2 below.

6.4.1 Average estimates of non-captured data in Human related factors

This section presents a discussion on the average estimates of the unmanaged data determined during the assessment of the Human related factors. The discussion basically covers a complete analysis of the unmanaged data explored in the 10 related data fields grouped under the Human related factors [see Table 28]. Correspondingly, a graphical illustration of the analysis result is demonstrated in the chart presented in Figure 45.

The total average estimate deduced from the Human related factors is considerably greater than the summation of the total average estimates calculated in both the *Accident related factors* and *Road related factors.* Thus, a total average estimate of 1,124 non-captured data generated demonstrates a high loss of large number of data points. These errors, perhaps, are attributable to the omission of information and the incorrect interpretation of the information acquired at the scene of the accident. These issues could aggravate the difficulty in accumulating sufficient data, and the chance of acquiring accurate interpretation details of the accident as indicated in the data form.

Table 28: Estimated average numbers of non-captured accident data in Human related factors

Average estimates of non-captured accident data in Human related factors		
Human related factors	**Average estimates**	**Percentage estimates**
Ages of drivers/cyclists	86	7.6%
Nationality of drivers/cyclists	115	10.2%
Gender of drivers/cyclists	91	8.1%
Race of drivers/cyclists	94	8.3%
Driving/learner licence type	189	16.8%
Seatbelt fitted/helmet present	113	10.1%
Seatbelt/helmet definitely used	124	11.0%
Liquor/drug use *[suspected]*	147	13.1%
Liquor/drug use *[evidentiary tested]*	153	13.6%
Vehicle manoeuvre	12	1.1%
Total average estimate	**1124**	

Sometimes, the problems might be as a result of the severity of the crash incident, in which the conditions of the accident victims are severely injured, or a case where the accident victims have left the scene. These conditions, thereby, incapacitate the possibility of acquiring necessary information. According to the process flowchart presented in Figure 59 in Appendix A-A.10.1, it is illustrated that anytime a fatal accident is reported to the nearest SAPS/MMT, a criminal case docket is expected to be filed for investigation purposes, depending on the agreement terms between the drivers involved.

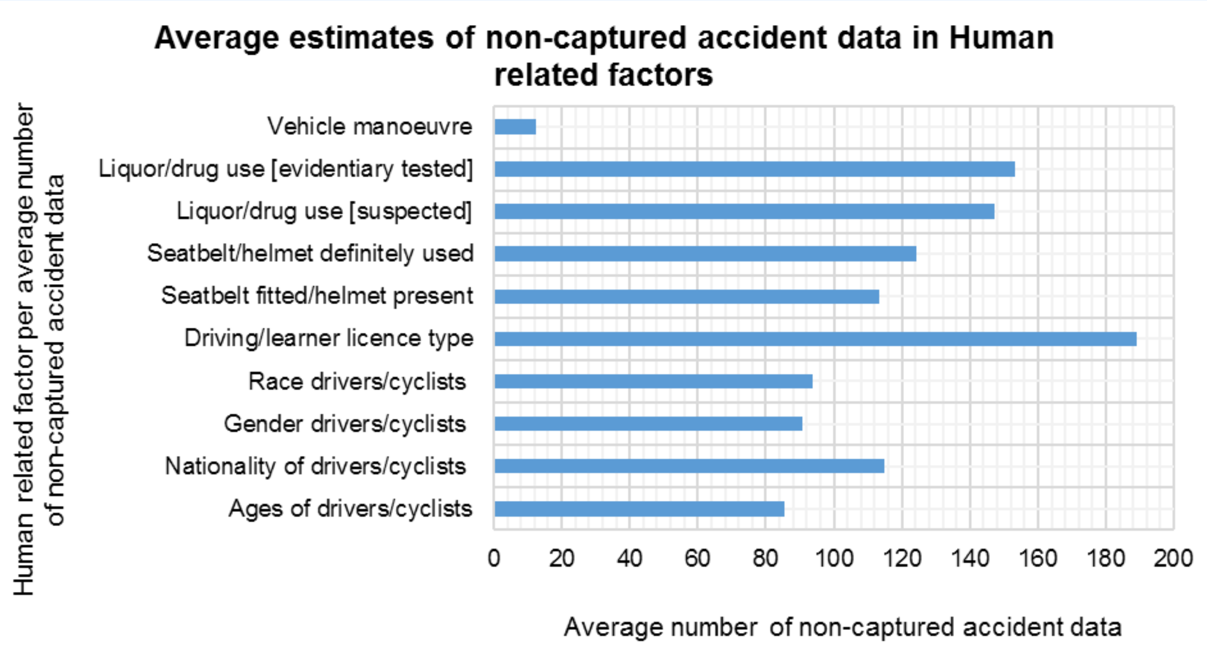

Figure 45: Total average estimates of non-captured accident data in Human related factors

From the table above, each field in the Human related factors corresponds with an average score. It is discovered that the Vehicle manoeuvre produces the lowest average estimate of 12 [1.1%] non-captured data per month amongst the remaining related data fields [see Figure 45]. This actual estimate is the second lowest to the *Road signs clearly visible* in the overall comparison of all the average estimates of non-captured data in the three related factors considered in this study. In this section, other related data fields with slightly lower amount of non-captured data points are *Ages of drivers/cyclists, Gender of drivers/cyclists,* and *Race of drivers/cyclists* with a sum average estimate of 271 [24.0%] non-captured data.

Considering the data fields with the extremely high average estimates of non-captured data, however, *Driving/learner licence type* demonstrates the highest average estimate of 189 [16.8%] non-captured data per month amongst all the data fields in the Human related factors. Other data fields with high loss of data are *Liquor/drug use [evidentiary], Liquor/drug use [suspected],* and *Seatbelt/helmet definitely used,* with average estimates of 153 [13.6%], 147 [13.1%], and 124 [11.0%] per month respectively.

According to the findings acquired concerning the high degree of omission and incorrect completion of information in Human related factors; thus, it is clear that the acquisition of information regarding the *Driving/learner licence type* is ascribed primarily to two factors, such as:

- Driving on the road without a driving/learner licence, and

- Inability to interpret the international driving licences.

The types of the driving licence used by the foreign drivers are dissimilar to that approved in South Africa, in terms of information arrangement and codes allocation to enhance information classification of the driver and the vehicle. This effect, thereof, could lead to difficulty in accomplishing an accurate interpretation of such field, due to insufficient training to support the interpretation of the international driving licence types. In the case of the *Liquor/drug use [evidentiary tested]* and *Liquor/drug use [suspected],* however, it is problematic to acquire information regarding the measuring of the liquor/drug use due difficulties as:

- Late reportage and disappearing of the accident victims from the accident scene, and
- Poor educational level, resulting to poor practical knowledge on how to appropriately operate the Alcohol Breathalyser.

6.4.2 Histogram analysis of non-captured data in Human related factors

This section discusses the frequency distribution estimate for the 10 related data fields in the Human related factors. Similar to the approach applied in the subsections 4.3.2 and 5.4.2 above, two separate tables were created to support the understanding of the data distribution. Each table contains the estimated scores of the non-captured data obtained from the analysis of all the data fields in the Human related factors.

However, the set of scores assembled in Table 29, demonstrate a huge difference between the highest and lowest scores across the four data fields. In the table, an extreme score of 126 non-captured data is considered as the highest score, while a lowest score of 7 non-captured data is found. The huge difference between the two scores indicate that some data elements are completed easier and more often in the ARF than many other data fields.

The approach used to compute the class intervals of the two histograms obtained in this section is based on the mathematical approach explained in the subsection 3.3.2.5.2. The distribution of the estimated scores assembled in the table below demonstrates a wide range of non-captured data in three data fields, namely the *Ages of drivers/cyclists, Gender of drivers/cyclists,* and *Race of drivers/cyclists.* These are the four factors with lowest average scores in Figure 45. From the same table, a complete and accurate completion of *Vehicle manoeuvre* is observed in April indicating a zero-data loss. A graphical illustration of the estimated scores is presented in Figure 46, in order to reveal the rate at which data is lost.

In the chart, a distribution gap is observed within the range of 35-51 non-captured data, with no frequency score within the specified range. Ultimately, the clustered part of the distribution produced the highest peak score of 13 counts within the range of 94-110 non-captured data. The peak score illustrates that large estimates of data are regularly uncaptured in the Human related factors between the range of 94 to 110, which is considered extremely high compared to the results achieved in both the Accident related factors and Road related factors in the analysis of the non-captured data with large scores. The peak score further demonstrates the amount of times a particular range of non-captured data is observed in the Human related factors. In addition, the clustered part of the distribution falls within the ranges with higher class interval of non-captured data, indicating a predominant high number of missing data points in the three data fields.

Table 29: Estimates of non-captured data for data fields with least missing data in Human related factors

Total estimates of non-captured data for least missing data in Human related factors				
Months	**Ages of drivers/cyclists**	**Gender of drivers/cyclists**	**Race of drivers/cyclists**	**Vehicle manoeuvre**
Jan	80	71	69	7
Feb	105	111	112	13
Mar	96	104	126	22
Apr	64	82	64	
May	64	97	96	12
Jun	78	74	80	11
Jul	74	73	79	7
Aug	109	101	105	13
Sep	91	92	87	17
Oct	100	100	106	14
Nov	105	110	123	12
Dec	61	74	78	8

However, to the left side of the gap, a frequency estimate of one count was observed within the range of 18-34 non-captured. This range illustrates the lowest frequency observation of the amount of errors discovered within the rate of 18 to 34 non-captured data in each field monthly. Further observation shows a high frequency score of 10 counts within the range of 1-17 non-captured data per field, illustrating a few counts of errors committed in each field per month. For the range 1 to 17 the non-captured data consisted entirely in the field *'Vehicle manoeuvre'.*

On the other hand, Table 30 contains the six data fields with the highest scores from Figure 45. The distribution of the estimated scores presents wide scores between 79 [lowest score] and 244 [highest score]. This set of scores partially dictates the size of the class interval. Aside from the statement given above, the size of the scores also determines the degree at which data are being mishandled in these data fields.

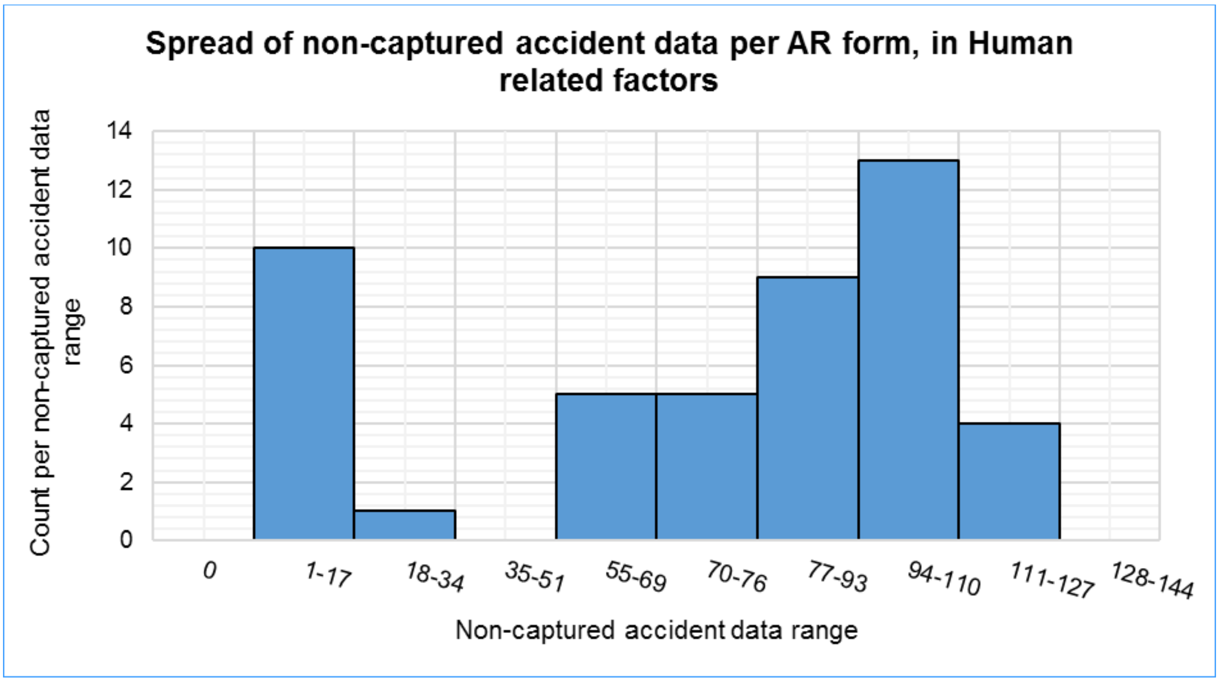

Figure 46: Distribution of non-captured data for data fields with least missing data in Human related factors

The six data fields are Nationality of drivers/cyclists, Driving/learner licence type, Seatbelt fitted/helmet present, Seatbelt/helmet definitely used, Liquor/drug use [suspected] and Liquor/drug use [evidentiary tested]. The chart presented below, demonstrates a distribution shape with the highest peak score of 18 counts within the range of 113-133 non-captured data.

Table 30: Estimates of non-captured data for data fields with most missing data in Human related factors

Total estimates of non-captured data for most missing data in Human related factors						
Months	**Nationality of drivers/cyclists**	**Driving/ learner licence type**	**Seatbelt fitted/helmet *present***	**Seatbelt/helmet *definitely used***	**Liquor/drug use *[suspected]***	**Liquor/drug use [evidentiary tested]**
Jan	88	152	95	112	114	128
Feb	133	217	125	144	167	169
Mar	143	233	161	184	212	220
Apr	85	177	108	123	137	143
May	118	221	115	139	159	177
Jun	105	169	104	113	139	145
Jul	102	166	87	99	117	120
Aug	129	244	113	121	159	162
Sep	112	131	125	136	150	149
Oct	150	206	118	130	155	157
Nov	135	219	121	97	156	162
Dec	79	134	89	93	101	106

The shape of the distribution produced is positively skewed to the right side of the chart. This distribution demonstrates a fewer frequency scores farther to the right side of the chart, which comprises the high class-interval ranges of non-captured data. In addition, the distribution represents the occurrence rate of a score classified within a specific range of non-captured data. From 175 and higher, frequency scores less than 5 counts are observed.

Further practical observations were performed to determine other useful results in the chart. As a result of this, from a clear indication, on the left side of the chart, a low frequency score of 5 counts is observed within the range of 71-91 non-captured data. This illustrates that fewer amount of errors is discovered at the rate of 71 to 91 non-captured data occasionally in each section during the data collection activities.

Ultimately, due to the range difference between the lowest and highest scores, it is clear that some users struggle to complete these data fields. Evidently, among other related factors discussed earlier in this study, it can be assumed that related data fields grouped in the Human related factors are largely misrepresented during any data collection activities. According to the investigation performed, complains are levelled against the refusal of the road users such as drivers, cyclists, and pedestrians in disclosing their personal details, such relevant information as age, residential/home address, nationality, and ID type/ID number. This effect contributed to the huge extent of non-captured data in the Human related factors.

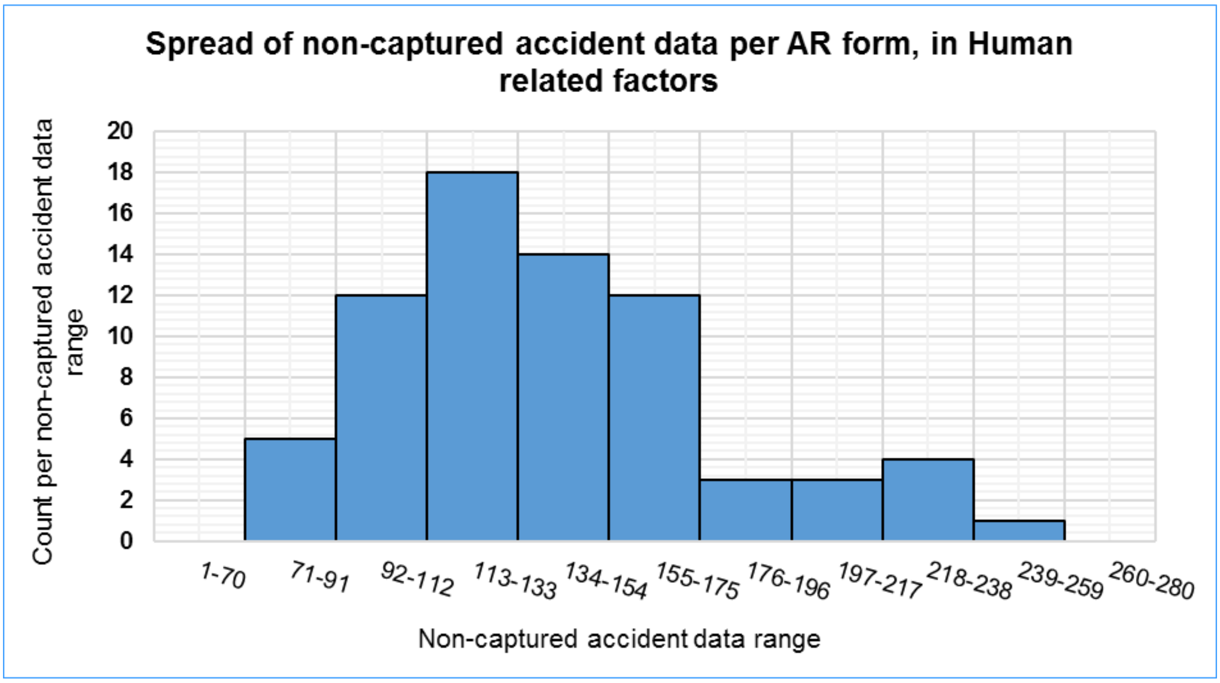

Figure 47: Distribution of non-captured data for data fields with most missing data in Human related factors

The ages of some road users, who are South Africans, were managed to be captured through the information arranged on their personal driving licence, but major concerns were raised against the inability of obtaining the ages of the foreign drivers/cyclists. More so, the issue of data mishandling could also be ascribed to the unexpected number of road accidents reported each day, which could influence the reliability of the reporting process due to factors like unreliable deployment, poor training, ineffective response, and incompetency. Besides, the irregularities affecting the collection of the *right data*, or the inability to avoid errors habitually contributes to the huge amount of uncaptured data in each field.

In that case, the reporting officers should be conversant with the necessary steps required to accumulate the *right data* pertaining to all the relevant data fields. Additionally, the reporting officers are advised to have a firm approach towards the accident victims, to facilitate valuable cooperation from them in order to acquire sufficient data for road accident analysis.

7. Survey analyses and findings

This chapter discusses the results of the responses demonstrated in the questionnaire by the selected participants, who willingly partook in the survey exercise carried out at both the STD and the local police department [SAPS] in the Stellenbosch. Before the commencement of the survey exercise, a sample size of 20 participants was considered, for both males and females who have participated in the reportage of RTA formerly or recently around the Stellenbosch locality. The survey exercise requested the participation of 10 reporting officers per local authorised department; that is, 10 participants from the local traffic department [STD], and other 10 participants from the local police department [SAPS]. The survey exercise is based on the availability and readiness of the 20 participants, coupled with a general assessment of the participants towards the survey questions inquired. A substantial descriptive analysis of the steps involved in the development of the survey is provided in the Appendix E. The description of the questionnaire cuts across areas such as;

- Purpose of the questionnaire.
- Development of the questionnaire.
- Validation of the questionnaire.
- Selection of participants.
- Completion of questionnaire.

However, each area contributed to the success of achieving an end result in the processing of the questionnaire. The development of the survey questions is basically to substantiate the significance and feasibility of the research project undertook at the STD, and also to improve the primary way of collecting the road accident data, with the need for the restructuring of the ARF.

In addition, the development of the questionnaire covers three key areas such as practicality of the ARF, reformation of the ARF, and application difficulty of the ARF[26]. The survey questions are categorised in accordance with the exact factors contributing to the poor data collection process. Besides, the responses/answers to the questions will clarify the purpose of the questionnaire towards the need of supporting the practical understanding of the identified factors that led to the poor data collection as discussed as part of the preceding analysis chapters.

---

[26] Complete flow diagram of the process is illustrated in the Figure 78 in Appendix E-E.1.

According to the results gathered from the analysis of the structured data, it is discovered that the data/information provided in the ARF is marred with anomalies categorised under the two forms of error discussed in the beginning of Chapter 4. A four-stage section is established in the questionnaire, with the purpose of differentiating the relevance of each section in the survey exercise. The *Section I* of the questionnaire comprises general information as regards the participants, such as;

- Description of the participants-his/her profession, age, sex, and educational level.
- Extent of involvement of a participant in the application of ARF.
- Extent of involvement of a participant in the inspection of an accident.
- Extent to which a participant understands the procedure of the ARF.

However, *Section II* of the questionnaire presents the questions developed for the technical acquisition of the information regarding the clumsiness of the data fields provided in the ARF to the participants during the collection of road accident data. On the contrary, *Section III* of the questionnaire is ultimately reserved for two significant data fields in the ARF, which are *Accident sketch* and *Accident description*. Survey questions developed for these two data fields are directed towards the inability of obtaining an accurate sketch of a road accident, and a comprehensive description explaining how an accident occurs.

With the intention of establishing more facts on the response options marked or indicated in the previous sections by the participants, however, an *Opinion section* was developed to complement other survey questions that are not provided with answers. This particular section is preferably to be answered by the participants with their natural instincts, according to their peculiar perceptions about the questions asked. This establishes a crystallised analytical relationship between the road accident data collated and the application of the ARF through a distinctive observation of the participants towards some individual opinions. The results obtained from the analysis of the questionnaire are discussed in the subsequent sections below; thus, the result will be used to validate findings obtained in the preceding chapters where necessary. This will establish a correlation between the questionnaire results and the practical results of the road accident data illustrated in both Chapter 4, Chapter 5 and Chapter 6.

7.1 Analyses and findings on the training and error related questions

This section discusses the analysis performed on the response options obtained from the preliminary information provided in the *Section I* of the questionnaire. Twenty participants [sample size] are expected to participate in this survey exercise as mentioned earlier, but only 19 participants out of the actual sample size demonstrated their interest in the *Section I* of the questionnaire, while the remaining one declined his/her participation due to some undisclosed circumstances. From the results gathered, it was observed that 10 participants indicated their designation as Traffic Officers. Also, 8 out of the 9 participants from the SPD/SAPS indicated their designation as Police Officers, while one out of the 9 participants indicated none of the options.

However, out of the 19 participants that partook in the survey exercise, only 7 male participants were involved. This number is quite smaller than that of their female counterparts, with a number of 12 female participants. According to the overall estimates, it was discovered that 11 of the participants claimed that their ages fall between the ranges of 41-45 years old, while 2 participants out of the remaining 8 participants claimed that their ages fall between the ranges of 36-40 years old. From the remaining 6 participants, 5 participants claimed that their ages fall between 31-35 years old, while the last participant's age falls in the age-group between 26-30 years. According to the estimates given above, understandably, the average age of the participants is 40 years. In the tables, it may be observed that not all respondents answered every question. However, the data was considered useful as is.

Table 31: Participants' observations about the training and error related questions in the application of the ARF

Estimates of responses to the training and error related questions			
Survey questions	**Response options**		
	Yes	**No**	**Not sure**
1. **Do you intend to further your education?**	9	5	5
2. **While completing the ARF, do you suspect that you might commit any errors?**	9	2	3
3. **Do you undertake any relevant training course on how to use the ARF in any accident reporting cases?**	7	8	0
4. **Were you provided any manual guide or instruction manual during the training course?**	5	3	2

As part of the research procedures, the educational level of these officers was also enquired during the survey exercise, in order to have a clear view of their level of competence towards the ability of structuring a reliable data collection process. It was discovered that 7 participants are Graduates, while 11 participants claimed that they possess Grade 12, but only one participant out of the 19 participants prefers not to disclose his educational level. This set of participants asserted that they have been working with the local traffic department for the period of 6 years and beyond. This substantiates their level of experience, and the ability to be competent enough to provide quality data.

Based on the result of the analysis performed, it is understood that lower rank officers were mostly deployed to attend to any crash incidents. Perhaps, this may necessitate thorough supervision to measure their level of confidence towards an improved supply of complete data at the local level. If measured, no doubt, it could boost the chance of reducing the amount of errors committed, in form of *item non-response error* and *response error*.

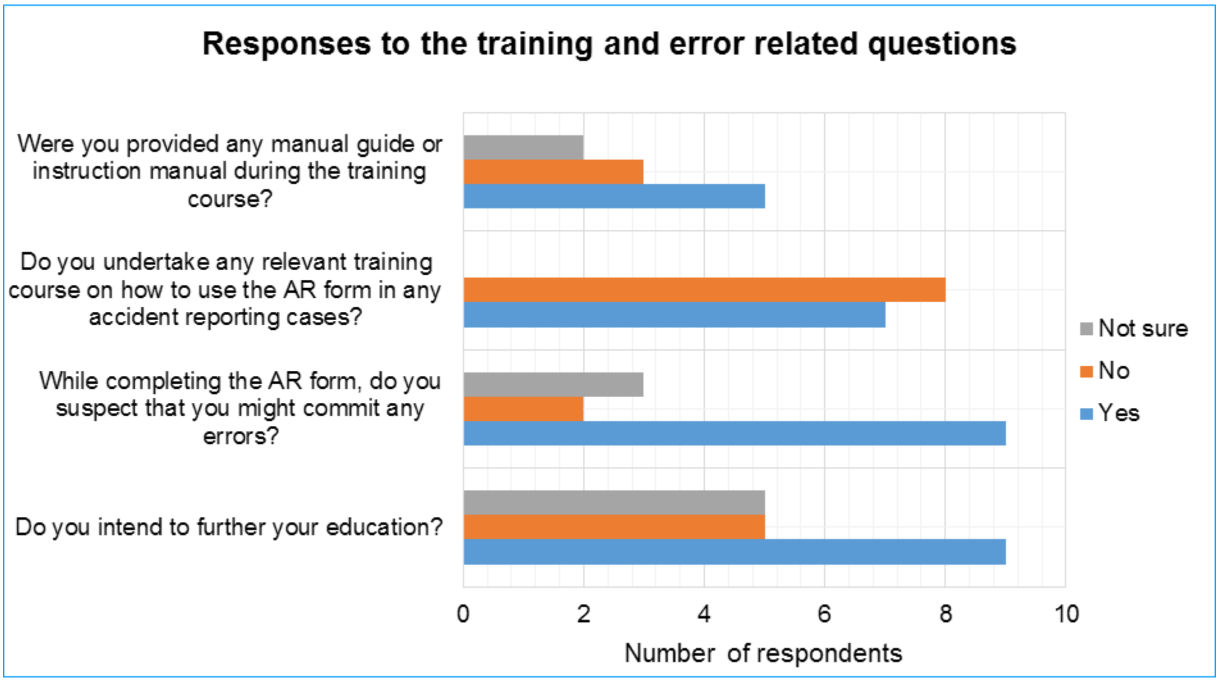

Figure 48: Total estimates of Participants' responses to the training and error related questions

From the results demonstrated in the chart above [see Figure 48], a total estimate of 9 [47.4%] participants have the desire to further their education, while a total estimate of 5 [26.3%] participants declined the desire of furthering their education beyond their current educational level. Similarly, another estimate of 5 [26.3%] participants indicated that they were not certain of furthering their education beyond their present level.

The above illustration is further simplified as thus; 5 participants out of the 9 participants who affirmed to the desire of furthering their education are considered to be among the participants with Grade 12, whereas 3 participants out of the remaining 4 participants are considered to be Graduates. This suggests that low-ranking officers have the desire to broaden their knowledge with the intention of improving their performance rates. If this desire could be transformed into reality, then the possibility of achieving a reliable outcome regarding the appropriateness of road traffic matters could be envisaged. Conversely, 6 participants with Grade 12 and 4 participants with Graduate degrees, asserted that they never had the desire, neither were they sure of furthering their education beyond their current educational level.

From the same chart, other results demonstrated are *participants' affirmations to the tendency of committing errors during the completion of the ARF, participants' affirmation to whether any training was implemented,* and *participants' affirmations to provision of manual guide or instruction manual as a directive during data collection process.* The results demonstrated in the chart indicate that 9 [64.3%] participants affirmed that there is a definite tendency of committing an error in the process of completing the ARF, while a total sum of 3 [21.4%] participants declared that they are not absolutely certain about the tendency of committing an error in the process. In addition, 2 [14.3%] participants declined that the tendency of committing an error in the process of completing the ARF is impractical.

In a similar context, according to the responses obtained in one of the survey questions asked in *Section I* of the questionnaire as regards *how often do you suspect that you commit errors when completing the ARF?* However, 11 participants claimed that they could commit 1-3 errors in the process of completing the ARF. On the contrary, 3 participants declared that they never committed any errors during the similar process. Only 1 participant claimed that she has the tendency of committing more than 15 errors during the same process. The case of committing more than 15 errors during data collection proceedings is not encouraging, therefore it should not be a regular issue since it poses a huge effect on the adequate supply of road accident data. Observably, this analysis ascertained the participants' level of conversance with the process, because it actually transpired their degree of experience in the data collection process.

It is deduced from the same chart that 7 [47.0%] participants declared that training was actually undertaken in the process of learning how to use ARF when inspecting a reported accident within the Stellenbosch locality. This shows that a certain number of participants based their personal experience of applying the ARF on the process of studying the manual guide/instruction manual provided. In contrast, a higher estimate of 8 [58.8%] participants affirmed that no training was being undertaken to guide them through the process of applying the ARF properly during any reported accidents.

Considering the provision of manual guide/instruction manual, the result displayed in the chart above demonstrates that half amount of the participants affirmed to have been provided with the manual guide/instruction manual, while the remaining half of the participants affirmed that neither was any manual guide/instruction manual provided, nor were they certain whether any manual guide/instruction manual was provided.

A pilot result obtained from one of the survey questions demonstrates *how long the training is scheduled.* This result showed that 8 participants claimed that a few days training section was provided for the application of ARF. Only one participant claimed that 5-6 months training section was provided. The illustration benchmarks the third survey question presented in the third row of Table 31 above, which supports the tendency of committing an error in the process of completing an ARF. A few days training is not enough to guide the reporting officers on how to complete data form, because high number of Grade 12 officers are among the officers deploy to monitor any road traffic incidents. This literally stipulates that most reporting officers deployed to accident scene required lot more experiences along the process, with a constant supervision by a superior officer.

7.2 Analyses and findings on the understanding and arrangement of the ARF

Questions analysed in this section include the challenges facing the understanding of the ARF in accordance with the indicated response options in the questionnaire. These questions were designed to determine the knowledge of the participants [reporting officers] regarding the information provided in the ARF. Besides, the questions inquired are considered as the key link to determine whether the data form requires any improvement, to accommodate more relevant information needed to resolve the issue of insufficient supply of road accident data. However, the questions considered are presented in the Table 32 below, with response options comprising five simplified options, such as *Strongly disagree, Slightly disagree, Neither agree, nor disagree, Slightly agree, and Strongly agree*. These options provide an insight into the perceptions of the participants as regards their responses to the questions inquired.

In the table below, the response results displayed indicate that less than 19 participants demonstrated their interest in the three survey questions, a total number of 8 [53.3%] participants strongly and slightly disagree with the knowledge of finding ARF challenging to understand. In addition, a total number of 3 [20.0%] participants neither agree, nor disagree with the possibility of having challenges with the application of the ARF. The first illustration shows that some reporting officers find the ARF more simple and comprehensive, while the second illustration indicates that some reporting officers rather remain unconvinced as regards the question inquired. On the contrary, a total number of 4 [26.7%] participants slightly and strongly agree that, they find the application of the ARF more challenging to understand. This could be attributed to the responses discussed in the section 7.1 above, as regards the provision for ARF application training and their level of experience.

For the second survey question, the results obtained reflect a close score across the response options considerably, compared to the results obtained from both the first and third survey questions. From the result, a total number of 5 [33.3%] participants strongly and slightly disagree that the arrangement of the information in the ARF instigates error during the road accident data collection process, while a total amount of 6 [40.0%] participants slightly and strongly agree that the information arrangement contributes to the errors committed. This shows that some certain number of participants [reporting officers] considered the arrangement of the information in the ARF not confounding towards data collection process, while some certain number of participants [reporting officers] considered it confounding.

Table 32: Participants' reactions to the understanding and arrangement of the information in the ARF

Estimates of responses to the understanding and arrangement of information in the ARF					
Survey questions	Response options				
	Strongly disagree	Slightly disagree	Neither agree, Nor disagree	Slightly agree	Strongly agree
1. Do you find ARF challenging to understand?	3	5	3	3	1
2. Do you think that the arrangement of the information in the ARF could cause errors?	3	2	4	4	2
3. Do you think the arrangement of the information requires restructuring?	1	5	1	5	2

Additionally, an estimate of 4 [26.7%] participants neither agrees, nor disagrees that the arrangement of the information in the ARF contributes to the errors committed during the data collection process. This result suggests that some participants have no indication on whether the arrangement of the information in the ARF has a huge influence on the errors committed. In essence, this further affirms that some reporting officers have little knowledge on the applications of the ARF, and neither were they in any period been at the RTA inspection scene in a prolonged period of time. As a result of this, some relevant questions were developed to strengthening the results acquired in the second survey question. The participation level of the participants in the inspection or reporting of RTA was measured, in order to determine their level of understanding towards the application of the ARF. These questions are crystallized to fathom a way of identifying how often the participants partake in the process.

According to one of the survey questions that states *how regularly do you participate in inspecting or reporting a RTA?* Thus, the responses acquired from this particular question demonstrate that a total percentage estimate of 36.8% [7] participants claimed that they have never been assigned to inspect any road accidents. They rather operate as administrative officers in charge of reports assessment and information processing in the two authorised local departments.

In contrast, a percentage estimate of 63.2% [12] participants claimed that they have partaken in the inspection process of RTA in the period of 1-5 days ago, 1-4 weeks ago and more than 5 months ago. To buttress this illustration, an estimate of 71.4% [10] participants affirmed that they have partaken in the completion of the ARF during a particular RTA inspection process. Similarly, an equal amount of 14.3% [2] participants respectively affirmed that they occasionally complete the ARF, while other equal number of participants declared that they have never participated in the completion of the ARF.

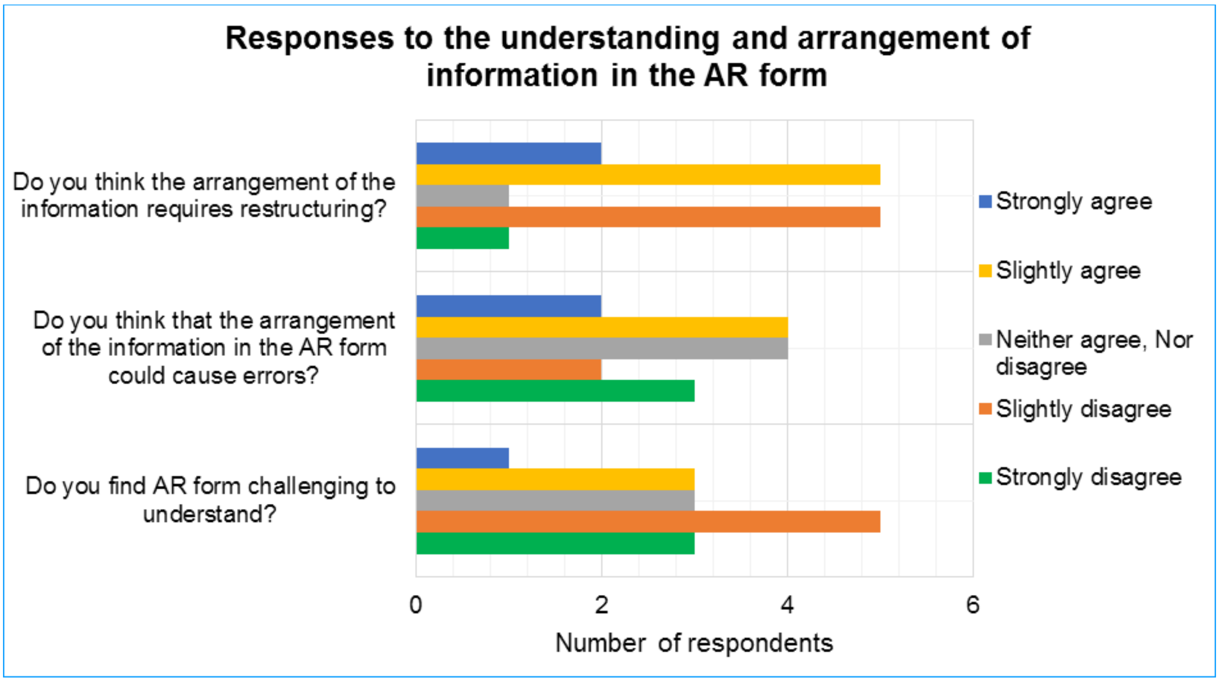

Figure 49: Total estimates of Participants' reactions to the understanding/arrangement of information in the ARF

The third survey question in the last row of the Table 32 above, was developed to support the responses obtained from the second survey question. The purpose of the third survey question, is to clarify a key need for the restructuring of the information arrangement in the ARF. However, the responses obtained from this particular question demonstrate that equal number of participants 5 [35.7%] slightly disagree and slightly agree on the need to restructure the data form. A total estimate of 2 [14.3%] participants strongly agree that the data form required reformation, while only one participant [7.1%] strongly disagree.

Observably, in Figure 49 above, it is simply understood that some participants are not convinced about the need to restructure the arrangement of the information in the data form. From the table displayed above, only 2 participants, each in both second and third questions, strongly affirmed to the effect of the information arrangement on the application of the ARF, and the need for the restructuring of the information. This indicates that a fewer number of participants sturdily find the ARF faulty in the area of information arrangement.

7.3 Analyses and findings on the time-taken to complete ARF for different accident types

This section encompasses three related questions developed for the purpose of measuring the participants' responses on the time-taken to complete an ARF. Five definitive variables are considered as the dimensional metrics, with respect to the period required of a reporting officer to complete an ARF during an accident inspection. According to the consultation carried out at the STD, a superior traffic officer acknowledged that, *"it is required of a reporting officer to* *complete a data form within a set range of time, mostly not more than 10 minutes per form, but it depends on the circumstances of the accident".*

As a result of the above statement, the required ranges of time-interval considered are presented along with the survey questions in the Table 33 below. The questions were practically developed to determine how long it could take a reporting officer in minutes, to complete a data form during the inspection of different conditions of the accident, such as Single-vehicle accident, Multiple-vehicle accident, and Single-vehicle and a Pedestrian accident.

Table 33: Participants' responses to the completion of the ARF during different accident types

Total estimates of responses to time-taken to complete ARF for different accident types					
Survey questions	**Response options**				
	1-5mins	**6-15mins**	**16-30mins**	**31-40mins**	**46mins & more**
1. **How long does it take you to complete an ARF for *Single-vehicle* accident?**	4	7	3	1	0
2. **How long does it take you to complete an ARF for *Multiple-vehicle* accident?**	1	4	4	3	3
3. **How long does it take you to complete an ARF for *Single-vehicle* and a *Pedestrian* accident?**	1	5	6	1	2

During the survey analysis, it was observed that only 15 participants demonstrated their interests in the three questions. The result of the analysis performed on the first question explains the responses of the participants regarding the time-taken to complete an ARF for Single-vehicle accident. From the result obtained, 7 [46.7%] participants claimed that they use 6 to 15 minutes to complete an ARF for a Single-vehicle accident inspection. Subsequently, 4 [26.7%] participants claimed that they use lesser time to complete an ARF for a Single-vehicle accident inspection at the rate of 1 to 5 minutes. A sum estimates of 26.7% participants affirmed that they normally use above 15 minutes to complete an ARF during the reportage of an accident involving single vehicle only. Utilising this amount of time to complete an ARF for a Single-vehicle accident indicates the level of competence of the reporting officers towards the prospect of acquiring error free reports.

Conversely, from the second question, the result obtained demonstrates variations in the responses indicated by the participants about the time-taken to complete an ARF for Multiple-vehicle accident. From the result, it is observed that a large number of participants use more time to complete an ARF for Multiple-vehicle accident as illustrated in the chart presented below [see Figure 50]. Illustratively, an equal amount of 4 [26.7%] participants asserted that they use between 6 to 15 minutes and 16 to 30 minutes respectively to complete an ARF for

Multiple-vehicle accident. Similarly, a total amount of 6 [40.0%] participants affirmed that they use above 30 minutes to complete an ARF for same type of accident.

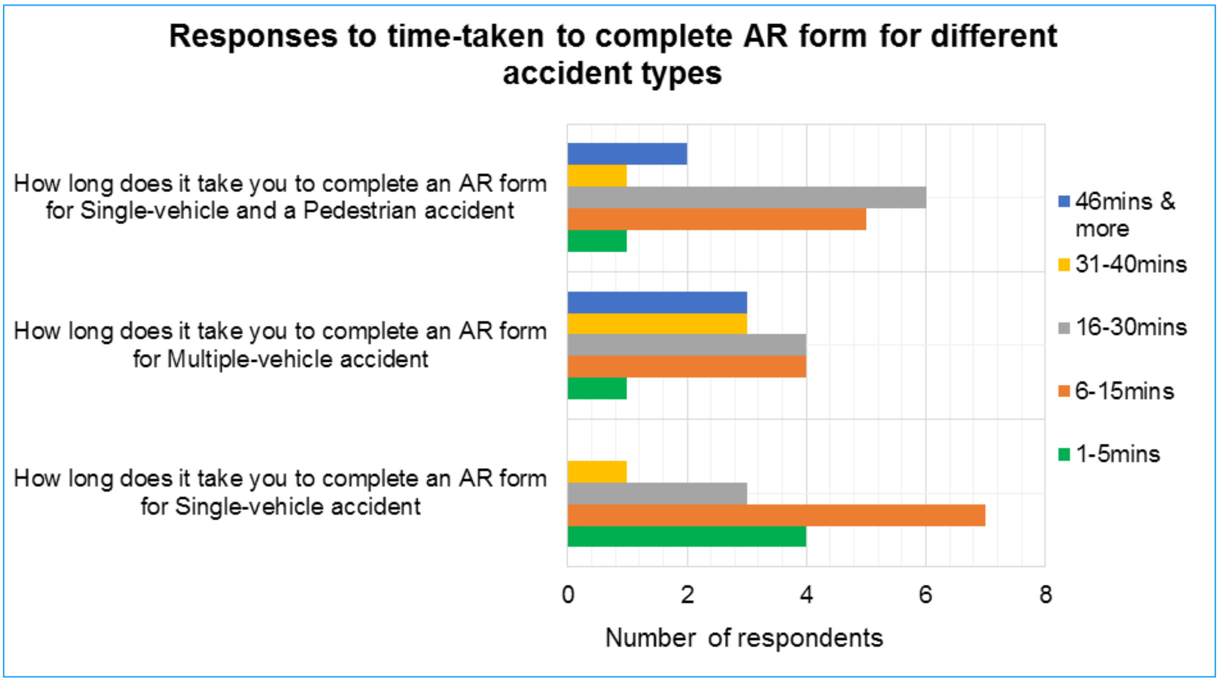

Figure 50: Total estimates of the Participants' responses to time-taken in completing ARF for different accident types

The last of the three questions is based on the time-taken to complete an ARF for an accident involving a single vehicle and a pedestrian. The responses obtained show that 6 [40.0%] participants asserted that they use 16 to 30 minutes to complete an ARF for Single-vehicle and a Pedestrian accident. However, 33.3% [5] of the participants affirmed that they use 6 to 15 minutes to complete an ARF for same type of accident. Only 6.7% [one participant] affirmed that they use 1 to 5 minutes to complete an ARF during the inspection of Single-vehicle and a Pedestrian accident, while a total percentage estimate of 20.0% [3] participants claimed that they use more than 30 minutes to complete an ARF for a Single-vehicle and a Pedestrian accident.

Generally, by comparison, the overall estimates calculated from the results established that most participants found it appropriate to use between 6 to 15 minutes, as the appropriate time-interval required to complete an ARF for the three forms of accident. In addition, a total estimate of 28.9% [13] participants affirmed that they use 16 to 30 minutes to complete the ARF for the three forms of accident. This outcome is second to the estimated value of 35.6% [16] calculated for participants that use 6-15 minutes. Other results show that 13.3% [6] participants claimed that they use 1 to 5 minutes' time-interval to complete the ARF for the three forms of accident. However, an equal estimate of 11.1% [5] participants claimed that they use above 30 minutes to complete the ARF for the three forms of accident.

The overall analyses show the preferred time-interval for the completion of the ARF in the three forms of accident. The best time-interval observed is between 6 to 15 minutes. Although, the estimated value acquired in the time-interval of 16 to 30 minutes demonstrate a close result to the estimated value acquired in the time-interval of 6 to 15 minutes. This illustration presents a valuable reason to consider the two ranges of time-interval to complete an ARF for the three forms of accident.

A time-interval of 6 to 15 minutes is considered suitable for the completion of ARF during the inspection of a Single-vehicle accident. On the other hand, a time-interval of 16 to 30 minutes should be considered for both Multiple-vehicle and Single-vehicle and a Pedestrian accident. These instances could strengthen the ability to reduce the rate at which errors are being committed during the accident inspection period. More so, it could stimulate the opportunity of acquiring more information from the accident victims depending on the level of severity of the accident. The selected time-intervals also justify the level of experience and conversance of the participants towards the application of the ARF.

7.4 Analyses and findings on the completion of ARF for different weather conditions and times of day

This section presents the results of the analysis performed on the three related survey questions presented in the Table 34 below, and the variations observed in the results are illustrated graphically in the chart presented in the Figure 51 below. These questions substantiate the savvy on the degree of difficulty experience by the reporting officers while completing an ARF in the day-time, night-time and rainy period. Four response options were considered as the dimensional metrics for structuring the responses/answers towards achieving accurate results. The four response options considered are as follows: *Very challenging, Challenging, Easy,* and *Very easy.* These options define the simplicity behind the idea of corresponding the responses/answers to the three related survey questions.

From the first question, as presented in the table below, it is simply observed that a high percentage estimate of 40.0% [6] participants affirmed that completing an ARF in the day-time is easy, while a lower percentage estimate of 20.0% [3] participants claimed that it is very easy to complete an ARF in day-time. These results clarify that a total amount of 9 participants has no difficulties in completing an ARF in the day-time. This illustration clarifies their level of experience in the application of the data form.

Table 34: Participants' responses to the completion of the ARF for different weather conditions and times of day

Total estimates of responses to the completion of the ARF for different weather conditions and times of day				
Survey questions	**Response options**			
	Very challenging	**Challenging**	**Easy**	**Very easy**
1. How difficult is it to complete ARF during day-time?	1	5	6	3
2. How difficult is it to complete ARF during night-time?	5	6	3	1
3. How difficult is it to complete ARF during rainy period?	5	7	2	1

By comparing other results, a percentage estimate of 33.3% [5] of the participants asserted that completing an ARF in the day-time is challenging, while only one participant affirmed that he/she finds the completion of an ARF very challenging in the day-time. These results illustrate that a total amount of 6 participants has the impression that completing an ARF in the day-time is quite challenging. The comparison of the two qualitative results shows that completion of an ARF in the day-time is never considered clumsy to a large percentage of participants.

The second question basically demonstrates the results pertaining to the perception of the participants on the completion of the ARF in the night-time. From the graphical illustration displayed below, 6 [40.0%] participants affirmed that they find the completion of the ARF challenging in the night-time, with additional estimates of 5 [33.3%] participants affirming that the completion of an ARF in the night-time is very challenging. In contrast, a few number of participants claimed that completing the ARF in the night-time is easy. This yields an estimate value of 3 [20.0%] participants, with an additional estimate of 6.7% [one participant] asserting that the completion of an ARF seems very easy in the night-time.

The analysis confirms that a large number of participants experienced difficulties during the process of completing the form in the night-time, while a nominal number of participants declared that they experience no difficulties in completing the form in the same period. Primarily, in the night periods, the police officers are primarily informed about any road accidents before the traffic officers. Nevertheless, the traffic officers operate in the night period as rapid response team. In this particular period of time, fewer amount of information is acquired, due to the extent of severity of the road accident and some other factors affecting the situation of such incident. Actually, if such accident is considered as property *'damage only',* however, it is required of the accident victim to register the accident at the nearest local traffic department for further investigation, or except a criminal case docket is required to be opened/registered for a criminal investigation.

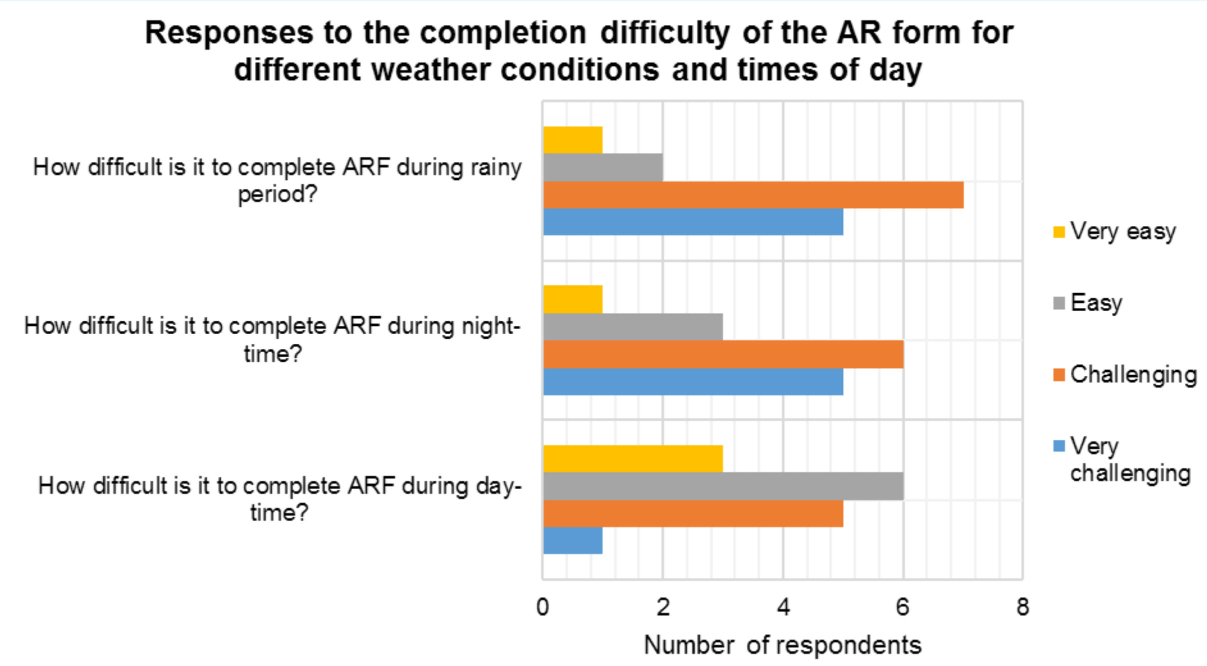

Figure 51: Total estimates of the Participants' responses to the completion of the ARF for different weather conditions and times of day

The third question focuses directly on the perceptions of the participants towards the completion of an ARF in the rainy period. The reactions of the participants to this particular question, seems slightly similar to the analysis result obtained in the night period. In conformity to the results acquired, it is observed that a percentage estimate of 46.7% [7] participants agreed that it is challenging while completing an ARF in the rainy period, whereas an additional percentage estimate of 33.3% [5] participants claimed that it is very challenging to complete an ARF in the rainy period. Contrariwise, a sum estimates of 3 [20.0%] participants claimed that they find it easy and much easier to complete an ARF during the rainy period. This result reflects the perceptions of the participants on the impact of performing road accidents inspection during the rainy period.

From a viewpoint, correspondingly, an overview of the responses/answers displayed in the Table 34 produced an overall estimate of all the responses to the three questions. According to the result obtained, it is observed that 40.0% [18] of the participants claimed that completing an ARF in the day-time, night-time, and rainy period is *challenging,* while a percentage estimate of 24.4% [11] participants claimed that completing an ARF during the three periods mentioned is *very challenging.*

On the other hand, a percentage estimate of 24.4% [11] participants affirmed that completing an ARF during day-time, night-time and rainy period is *easy,* while an estimated percentage of

11.1% [5] participants declared that completing an ARF in the three periods is *very easy*. By general comparison, it is obvious that a large number of the participants find it difficult to have an accurate completion of the ARF during the day-time, night-time and rainy period. The difficulties encountered could possibly be attributable to such factors as enumerated below:

- Poor correlation of facts and findings to complement the options provided in the ARF (Kayi 2007).
- Lack of cooperation from the accident victims e.g. hit and run scenario.
- Lack of necessary tools to execute operations accurately during these periods of time (Thieman 1999; Kayi 2007).
- Unlit streets/roads which may affects the ability to acquire relevant traffic road accident data in the night periods.
- Failure to perform a thorough observation at the location of the accident, which leads to insufficient data/information coverage (Kayi 2007).

7.5 Analyses and findings on the difficulty of acquiring data in the accident factors

This section discusses the analysis performed on the perceptions of the participants on the difficulty of acquiring information regarding all the categorical data fields in the ARF. This section represents the Section II of the questionnaire, wherein 15 relevant data fields in the ARF are considered. In the Section II, the data fields in the ARF are provided as the answer options in the subsidiary questions[27]. The first survey question is centred on the idea of understanding the perception of the participants on any particular factor with data acquisition difficulty, while the subsidiary question provides response options to support the positive response indicated in the first question inquired. However, the response rate shows that only 14 participants demonstrated interest in the Section II, while the remaining number of participants ignored this section.

The distribution of the set of scores displayed in the table below is graphically illustrated in Figure 52. The results obtained, though, yielded no indication of data acquisition difficulty in the *Light condition*. In essence, the results generated from the affirmative responses indicated that 6 participants, at most, struggle with the data acquisition in accident factors as displayed in the Table 35 except the *Light condition.*

[27] Read more details about the Section II in the Appendix E-E.8.

Table 35: Participants responses to the completion of ARF for all accident factors in RTA

Responses to the difficulty of acquiring data regarding accident factors in RTA			
Accident factors	**Response options**		
	Difficult	**Not difficult**	**Not sure**
Accident primary details	1	12	1
Accident location	3	11	
Driver's particulars	6	6	2
Road details	3	11	
Light condition		13	1
Weather conditions and visibility	1	11	1
Road signs	1	11	2
Vehicle details	1	12	1
Accident type	2	10	1
Hit and Run	1	11	2
Summary of persons involved	2	11	1
Pedestrians/Passengers	6	6	2
Pedestrians and Cyclists only	4	5	2
Special observations	2	10	2
Dangerous goods only	3	11	
Total scores	**36**	**151**	**18**
Percentage estimates	**17.6%**	**73.7%**	**8.8%**

According to the finding, the factors with the most response rates are the *Driver's particulars* and *Pedestrians/Passengers,* both appear more difficult to acquire during data collection proceedings. Other factors with response rates within the range of 3-4 responses are *Accident location, Road details, Pedestrians and Cyclists only,* and *Dangerous goods only*, while other factors with fewer response rates within the range 1-2 responses are Accident primary details [such as accident date, time of accident, day of week etc.], *Weather conditions and visibility, Road signs, Vehicle details, Accident type, Hit and Run, Summary of persons involved,* and *Special observations.*

This particular discussion clarifies that fewer number of participants have difficulty in acquiring relevant information, as regards some factors/data fields provided in the ARF. Correspondingly, 13 participants at most indicate that they have no difficulty in obtaining data regarding the accident factors. Conversely, 2 participants at most demonstrate their objectiveness towards the question inquired about all the factors except three other factors, like *Accident location, Road details,* and *Dangerous goods only.*

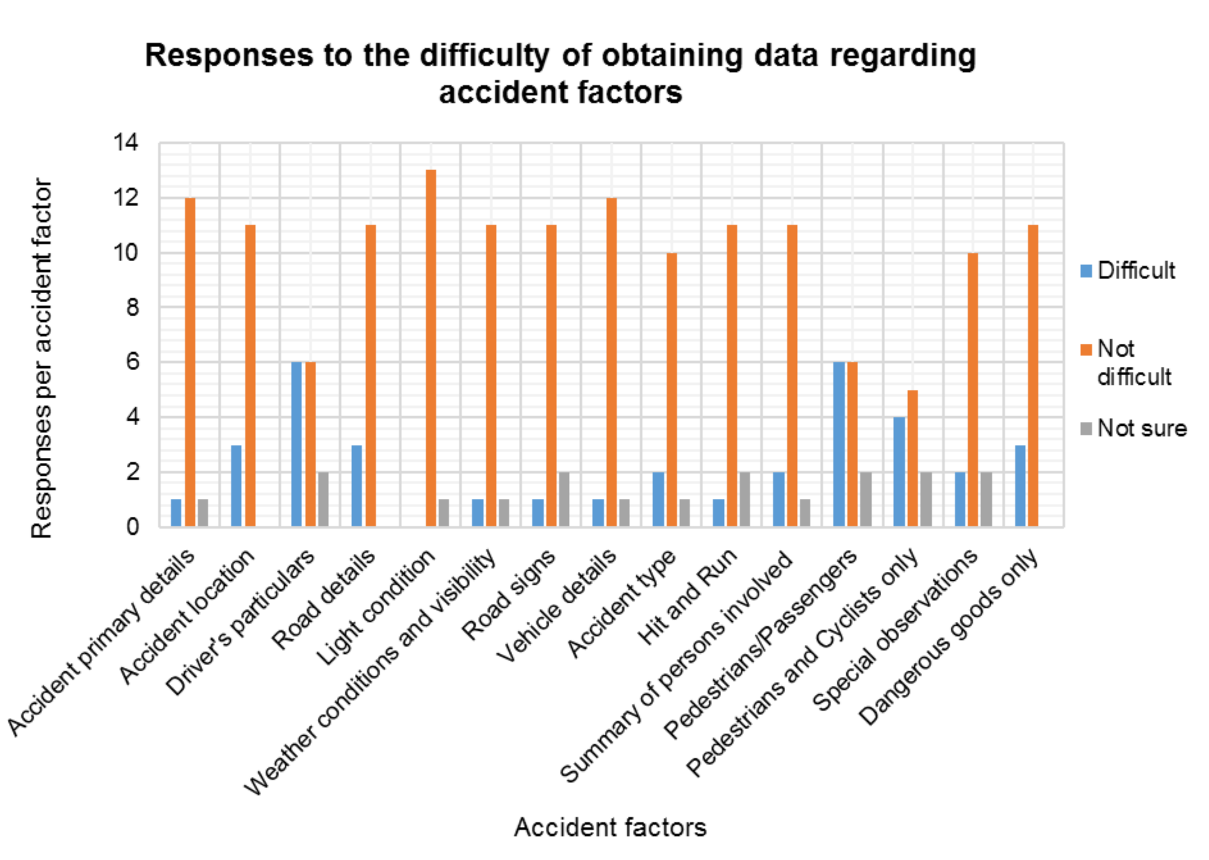

Figure 52: Total estimates of the Participants' responses to the completion of the ARF for accident factors

Overall estimate computed shows that a total percentage estimate of 73.7% responses affirmed to *'Not difficult',* to demonstrate their perceptions towards the difficulty of procuring road accident data for some factors/data fields. This amount represents the most indicated/marked response options in the Section II of the questionnaire, compared to other two response options *'Difficult'* and *'Not sure'*, with total percentage estimates of 17.6% and 8.8% respectively. The results indicate that a high number of participants have no problems with the acquisition of data/information as regards all the accident factors/data fields.

Ultimately, the participation estimates of all the participants in this particular section of the questionnaire is evaluated graphically in Figure 53. The chart basically presents the average response estimates between the traffic officers and the police officers as regards the questions inquired in the Section II. The result obtained shows that police officers have fewer participation rate than the traffic officers. In other words, thus, an equal average estimate of 2 participants [traffic officers and police officers] indicate *'Difficult'* per question inquired in the Section II. This average estimate reflects the expected number of participants responding to a particular question inquired, by indicating/marking their choices of response options.

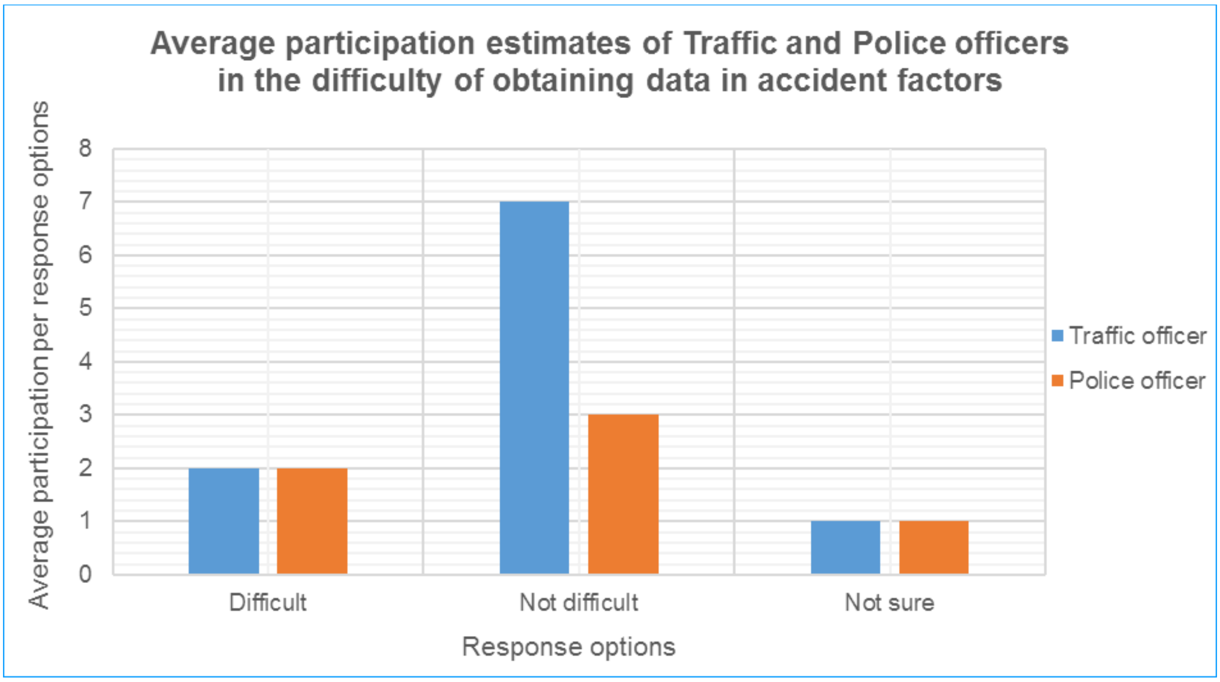

Figure 53: Participation estimates of the Police and Traffic Officers in the accident factors

Conversely, considering both the negative and indeterminate responses, the average participation estimates observed establishes that more traffic officers indicated *'Not difficult',* while an equal amount of traffic and police officers indicated indeterminate response *'Not sure'.* This illustration further demonstrates the average estimates of the traffic officers indicating negative response *'Not difficult'* per question inquired in Section II, in excess of 4 participants to the average estimates of the police officers, meanwhile an equal average estimate of one participant each demonstrates indeterminate decision per question inquired. Overall result shows that average participation estimates of 10 traffic officers indicate all options per question in the Section II, whereas only average participation estimates of 6 police officers indicate all options per question.

7.6 Analyses and findings on the circumstances affecting the acquisition of an informative Accident sketch

This section discusses the results obtained from the analysis of the observable conditions affecting the possibility of providing or acquiring an informative *Accident sketch*[28] during a road accident inspection. In addition, the section is centred on the needed to practically understanding the perceptions of the participants as regards their inability to provide a clear and explanatory *Accident sketch*. With regard to this effect, four related survey questions were

[28] Refer to subsection 3.3.2.4.1 for a descriptive definition of an informative Accident sketch.

developed with suitable response options. The response options guide the participants in determining the right answer to a particular survey question.

Table 36: Participants' responses to the completion of the ARF for Accident sketch

Responses to circumstances affecting the acquisition of Accident sketch			
Survey questions	**Response options**		
	Yes	**No**	**Not sure**
1. Do you undergo any special training on the procedures implemented in drawing the Accident sketch details of the accident?	6	7	1
2. Does the training undergone help you to understand the procedures implemented in drawing the sketch details of the accident?	7	0	0
3. Do any circumstances affect your ability to produce a good technical sketch details of the accident?	4	8	1
4. Do you find it difficult to obtain required features in the Accident sketch as instructed in the ARF?	5	9	0

However, analysis performed on the participants' choice of response options yielded the results displayed graphically in the Figure 54, which constituted the opinions of the traffic officers and police officers. The first question is based on the notion that the participants undergone a special training on the procedures required to produce a complete sketch of the road accident. Fundamentally, it is instructed in the instruction leaflet developed for the ARF that, only the authorised officers are permitted to produce a sketch of the accident. In other word, it is unlawful or unofficial if any accident victim carries out this particular task.

Observably, the outcome of the analysis performed on the first question shows that 7 participants affirmed to the negative response *'No',* by confirming that they undergone no special training on the procedures implemented in constructing a detailed Accident sketch. On the contrary, an estimate of 6 participants affirmed to the positive response *'Yes'*, by asserting that they undergone training on how to construct a detailed sketch of road accidents, while only one participant was objectively unconvinced about the question. The difference between the estimated values of both the positive and negative responses demonstrates a close estimate in the number of participants. This illustration suggests a good reason for the reassessment of the procedures implemented for *Accident sketch,* in order to improve the chance of acquiring a detailed sketch of road accidents to support the exact replica of the sketch in the GPS system-if any available.

The analysis performed on the second question supports the results gathered from the first question based on the positive response *'Yes'* only. The result shows that only 7 participants confirmed that the training undergone strengthens their ability to clearly understand the procedures implemented in constructing a detailed sketch of a road accident. However, from the analysis of the first question, it was noted that only 6 participants demonstrated an affirmative response. Obviously only those who were trained for the sketch, would be able to answer this question.

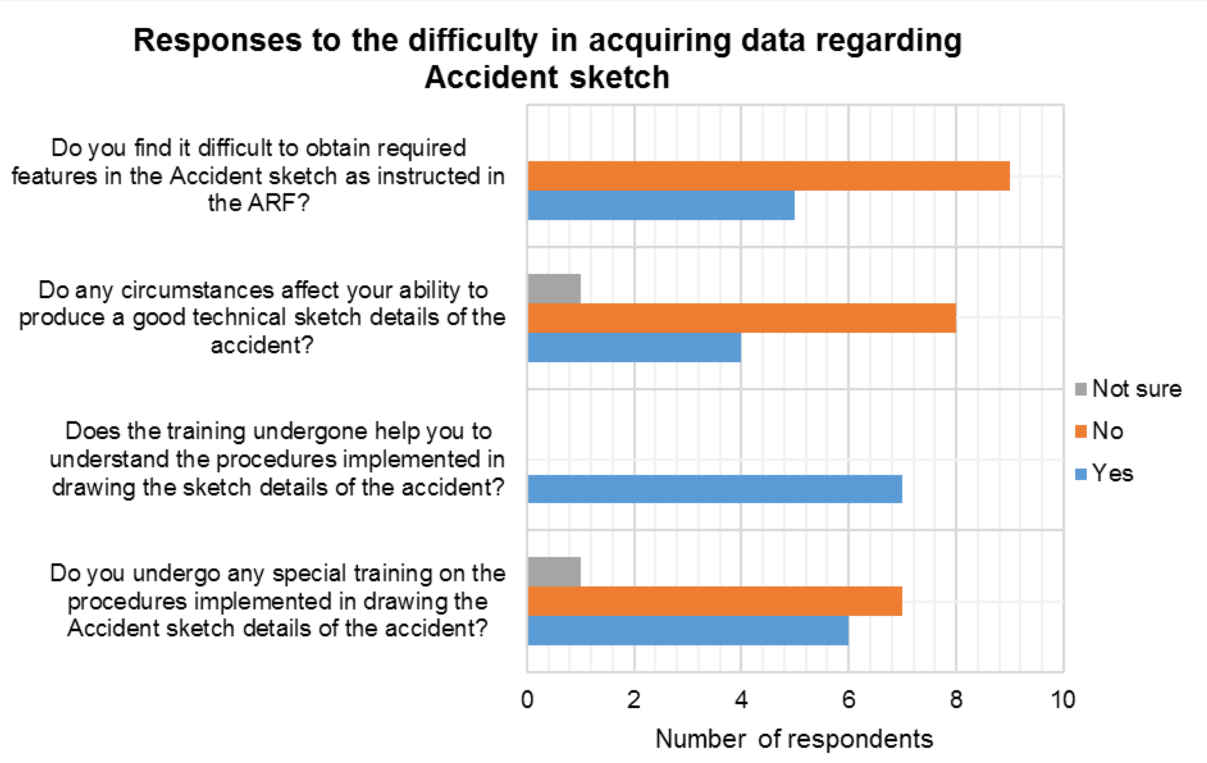

Figure 54: Total estimates of the Participants' responses to the completion of the ARF for Accident sketch

Relatively, the result obtained in the second question yielded 7 participants, which is excess of one participant to the number of participants that affirmed to positive response in the first question. This further demonstrates the lack of understanding from one of the participants towards the question inquired. Perhaps, this could be one of the problems encounter in the interpretation of information in the ARF. In this case, the additional participant, is a product of incorrect or poor interpretation of the questions.

The third question is based on the concept of identifying whether there are conditions contributing to the inability of producing a technical sketch details of a road accident. Thirteen participants showed personal interest towards this particular segment. The result obtained demonstrates that only 4 participants affirmed to the positive response of being constrained by certain circumstances affecting their ability to produce a good technical drawing of a road accident as presented in Figure 54.

In contrast, 8 participants declared that no conditions whatsoever have affected their ability to produce a good technical drawing of a road accident, while only one participant is indeterminate about the question. This insinuates that fewer number of participants are not confident at sketching, while a large number of participants are confident in their abilities to produce a good technical drawing.

With the objective of clarifying the circumstances affecting the ability of the four participants from constructing a detailed technical sketch of a road accident; thus, blank spaces are provided for valuable opinions from the four participants to support their responses. In the process, a series of useful opinions was provided by the participants to clarify their view on the circumstances contributing to their inability of constructing an informative technical sketch to support the description of a road accident.

From these opinions, some participants disclosed that *'they lack the required skills to construct a good sketch',* while some others divulged that *'inadequate or improper training handicapped them from producing a good sketch'.* Some participants further claimed that *'a situation wherein the vehicle involved is not found at the scene of the accident, or a situation where the injured persons already moved or left from the scene of the accident, then the desire of acquiring the necessary information to construct a good sketch is handicapped'*.

The last of the four questions is centred on the difficulty to obtain required features that constitute an informative technical sketch. The features are defined according to the requirements stated in the sketch section of the ARF. The required features constituting a detailed sketch of a road accident are enumerated as thus;

- Direction North must be indicated with an arrow symbol and letter 'N',
- Indicate the direction of travel of the vehicles involved,
- Positions of vehicles involved,
- Reference number of each vehicle,
- Symbolise the pedestrian involved,
- Alleged point of impact, tyre marks, fixed point[s], and
- Other object[s] involved.

According to the result displayed under this particular question in the Table 36, it is understood that 5 participants affirmed to the positive response *'Yes',* that they encountered difficulty to obtain the required features. Contrary to this, 9 participants affirmed to the negative response *'No',* that they had no trouble in indicating the required features in the road accident sketch.

However, a remark made by one of the participants, shows that some of the reporting officers lack the ability of constructing a proper sketch with the required features. This inability, if possible, might prompt some of the reporting officers to task the accident victims to make a sketch of the road accident, provided he/she is in good condition to do so. This attitude has adverse effect on the possibility of achieving unbiased sketch. Further observations show that *'lack of enough space for a clear and detailed sketch in the ARF, late reporting of accident [leads to missing information], inability of reconstructing the scene [due to the absence of the accident victims and vehicles involved]';* all these aforementioned factors have impact on the ability of constructing an informative sketch of a particular accident.

7.7 Analyses and findings on the circumstances affecting the acquisition of an informative Accident description

In this section, the results displayed in Table 37 constitute the analysis performed on the response options indicated by the participants regarding the circumstances affecting the acquisition of an informative *Accident description*[29]. The five survey questions are comparatively formulated, with the intention of having a general understanding of the likelihood areas that contribute to the difficulty confronting the ability of obtaining a detailed description of a road accident. The survey questions presented in this section is similar in pattern to the survey questions presented in the section 7.6 above, but based on different motives. A graphical illustration of the results is presented in Figure 55.

The first question focuses on the purpose of finding out whether the participants undergone any special training on how to write a good descriptive report of a road accident. From the table, it is observed that 5 participants out of 14 interested participants affirmed that they undergone necessary training, that guides them through the protocol of generating or writing a detailed description of a road accident. However, the remaining number of participants affirmed to the negative response *'No',* demonstrating that they are not trained on how to produce a detailed description of a road accident. By comparison, pertaining to the availability of any special training, it is observed that great number of participants demonstrated their concerns in producing a detailed description of a road accident than a detailed sketch of a road accident.

The second question is formulated to finalise the influence of the training undergone on the writing of a detailed description of a road accident. In essence, this question is viewed as a subordinate to the affirmative response in the first question, by clarifying the benefit of the training in the process of writing an informative *Accident description*. From the result obtained, only one participant out of 5 participants declined the benefit of training in understanding the procedures implemented in writing the description of road accident. In contrast, the remaining 4 participants affirmed that the training contributed tremendously to their ability to compose a

[29] Refer to subsection 3.3.2.4.1 for a descriptive definition of an informative Accident description.

detailed *Accident description,* to enhance the preliminary interpretation of the sketch demonstrates in the *Accident sketch* segment.

Table 37: Participants' responses to the completion of the ARF for Accident description

Responses to circumstances affecting the acquisition of Accident description			
Survey questions	**Response options**		
	Yes	**No**	**Not sure**
1. Do you undergo any special training on the procedures implemented in writing the description of the accident?	5	9	0
2. Does the training undergone helps you to understand the procedures implemented in writing the description statement of the accident?	4	1	0
3. Do you prefer the accident victim, provided s/he is in safe condition, to write the accident description?	3	11	0
4. Do you find it difficult to obtain an accurate description regarding any accident as instructed in the ARF?	6	8	0
5. Do you think the language barrier contributes to the problems of obtaining an accurate description of the accident?	8	4	0

The third question is basically directed at the concept of considering whether some participants [reporting officers] prefer the accident victim of a reported accident to write the description of a road accident. From the findings gathered, as displayed in Table 37, only 3 participants out of 14 interested participants, confirmed that they preferably permit an accident victim to write the description of a road accident, provided he/she is in a good condition. Although the remaining 11 participants affirmed to the negative response that they do not prefer the accident victim to write the description of a road.

The instruction leaflet comprising the general information on the procedures required in reporting a road accident and the completion of the ARF states that, *"driver of a vehicle involved is authorised to complete the form where no criminal case docket has been registered, provided he/she is in a good condition to do so"* (RTMC 2007). This can further be elucidated that a driver of a particular vehicle involved in a road accident that caused the death of one or more road users, or where a criminal case docket is filed should never for any reason whatsoever be allowed to write the description statement of any road accidents.

In that case, the opinion of the participants with affirmative responses, were sought by providing blank spaces for necessary comments to buttress their reasons for preferring accident victim to complete the form. One of the comments made by the participants reflects that driver of a vehicle that involved in the road accident is permitted to complete the form, despite that a criminal case docket is opened/filed. One of the participants stated that *"this particular scenario happens mostly during late reportage, wherein the reporting officer has no* *clue about the accident, therefore, only the accident victim of the accident could provide any useful information regarding how the incident happens".*

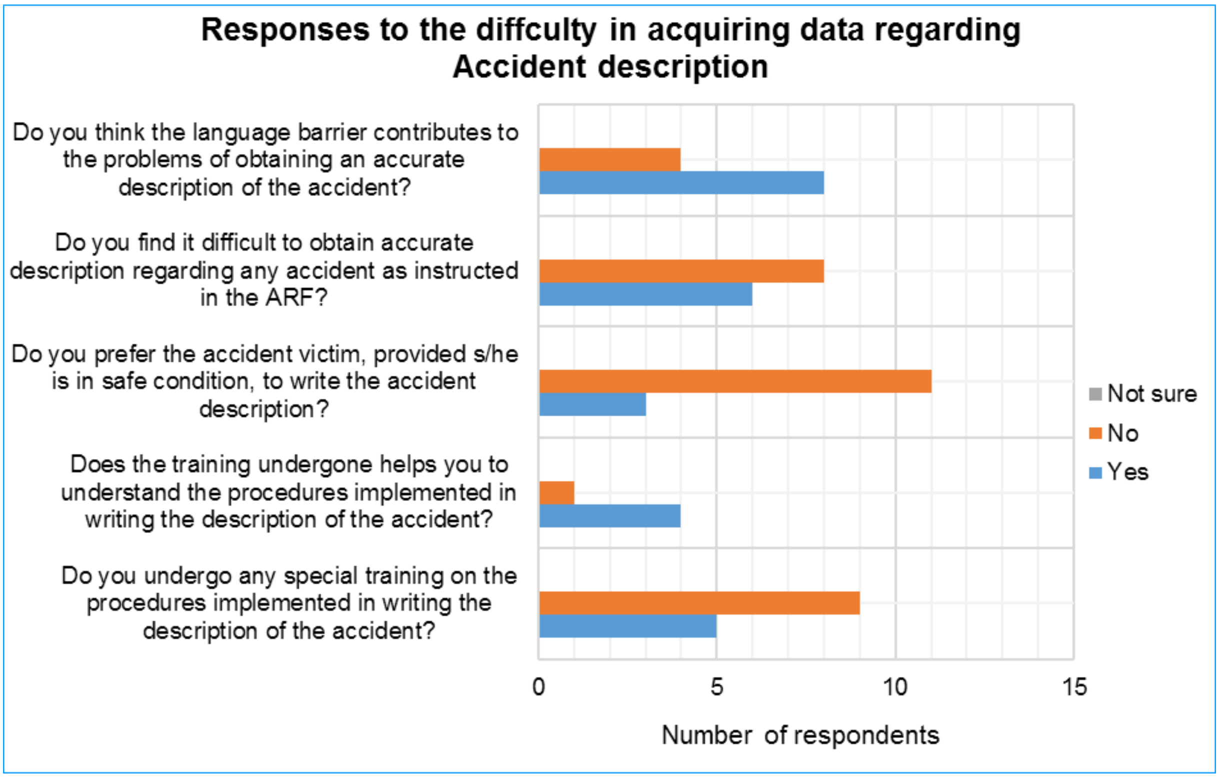

Figure 55: Total estimates of the Participants responses to the completion of the ARF for Accident description

Another participant added that, *"reporting officers are most times accused of writing incorrect information, or being biased as regards the writing of the description of an accident, or being blamed for insurance disagreement in issuing an accident benefit claim due the inappropriate detail of the description"*. Comprehensibly, the first comment does not warrant any reason for authorising a driver to complete the ARF, unless the reporting officer in charge is neither committed, nor competent to perform his/her duty diligently.

The fourth question is based on the idea of clarifying the perceptions of the participants towards the difficulties confronting their ability of writing a detailed description of a road accident. Technically, from the analysis result presented in the fourth row of Table 37, it is observed that 8 participants that they never encountered any difficulty in obtaining an accurate description of a road accident, while 6 participants claimed that they encountered difficulty in obtaining an accurate description of a road accident. The result shows slight variations between the positive and negative responses, which stipulates that a fewer number of the participants agreed to the incapability of composing a detailed accident description.

Moreover, the above result is substantiated with the observations derived from the comments made by some participants, as subordinate description of the primary problems opposing the process of producing an accurate description of a road accident. According to one of the comments made, a participant stated that *"some drivers are alcoholic, confused, and under shock",* while another participant added that, *"some are aggressive and non-cooperative during a road accident reportage".* The link between the two comments stated above shows that it is always difficulty to get RTA information from a driver who is not in right state of mind. During this period, drunk drivers are perceived to be extremely disordered and very delicate from being ready to provide relevant information concerning any reported accidents.

Further comments from the participants point out other areas contributing to the difficulties thwarting the ability to produce an accurate description of a road accident. One participant emphasised that feasible problems causing underreporting in the accident reporting is mostly *'a case of hit and run accident, and also a serious accident wherein everybody is injured'.* Additionally, *'a situation where the driver involved in a road accident died on the spot, and there were no eyewitnesses to provide relevant information on how the accident happened',* was accentuated by another participant as part of the difficulties preventing the acquisition of an accurate description of a road accident.

Similarly, one participant claimed that, *'a scenario where the accident victims are in the hospital for a long period of time, deter the chance of obtaining required accident description information'.* In this context, the reporting officer responsible for this case is expected to follow-up regularly on the current condition of the accident victim. This process could assist in determining the right time to acquire some relevant information about the accident. One participant criticises the attitudes of some reporting officers towards their responsibilities by alleging that, *'members are not interested and motivated to complete an ARF properly'.* Actually, this is a management issue that requires a feasible management approach, wherein the performance rate of the reporting officers could regularly be measured and periodically compensated with incentives, in order to motivate their abilities and readiness towards achieving a reliable data acquisition result.

The last of the five questions is developed to support the discussion presented formerly in the third question. This assists in evaluating whether language barrier is considered as part of the difficulties preventing the possibility of achieving an accurate description of a road accident. According to the analysis result obtained, 8 participants out of the 12 interested participants declared that they have issues with the language barrier, while the remaining amount affirmed to the negative response that they have no issues with the language barrier. The interpretation above shows that language barrier is a big issue in the composition of a detailed road accident.

8. A conceptual framework developed for the RTI

In this section, a concise procedural process followed at the local level will be reconsidered in the evaluation of the challenges discovered in the process. The procedures implemented in acquiring and transferring RTI at the local level is performed daily and weekly[30], based on the procedural operations carried out at this level, such procedures as:

- Recordkeeping of completed ARFs.
- Compilation and arrangement of completed ARFs.
- Obtaining scan copies of the completed ARFs.
- Gathering of completed forms from various data sources.
- Performing validation process on the completed ARFs.
- Confirming the approval of the completed ARFs.
- Capturing of the road accident data represented in the completed ARFs.

Notwithstanding, with all the above-mentioned procedures, one will inquire the reason for high irregularities, resulting to high loss of road accident data discovered in the completed ARFs. The large amount of non-captured data discovered, as analysed in the Chapter 4, Chapter 5 and Chapter 6 of this study, could be attributed to the weaknesses existing along the sequence of the road accident data collection, evaluation and analysis.

The current procedural process adopted at the local level in processing RTI, simply supports the implementation of validation process every Friday of the week, with the aim of unfolding any existence of anomalies that could possibly render the data completed in the ARF inadequate. At the local level, the validation process implemented is based on the application of two quality dimensions, applied along the process of compiling the ARFs at the local police departments [SAPS], and transferring the forms to the STD with the use of DFN, and receive by the DCO for further processing (RTMC 2014). The process is only implemented to evaluate the degree of correctness and integrity of the data collected. The region demarcated in the flowchart with *orange dash lines*, are not considered in this research, since the investigation carried out is limited to the operations executed at the local level as stated in the section 1.2 above.

[30] see Table 38 in the Appendix A-A.10.3.

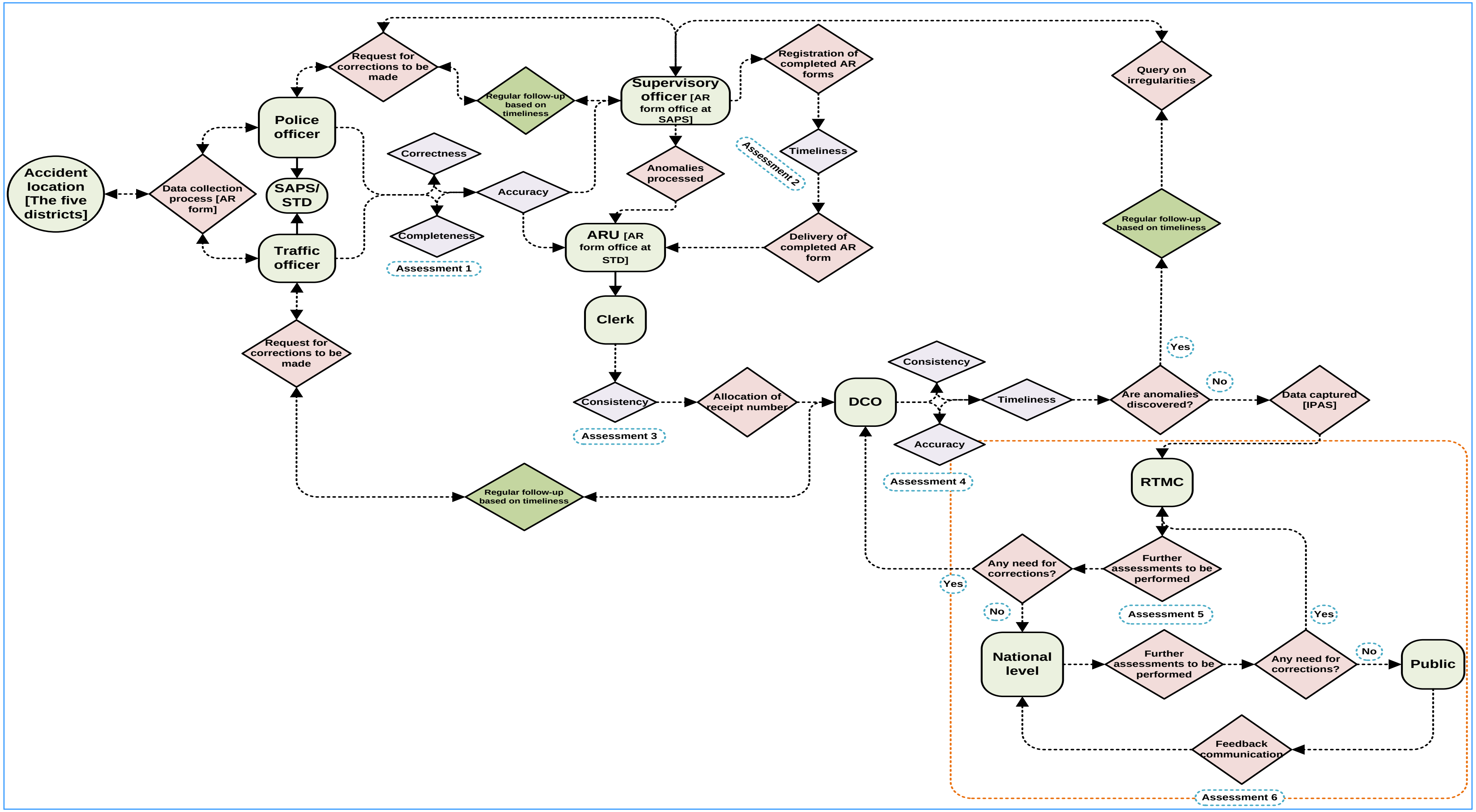

Figure 56: A simple flowchart for an improved application of the quality dimensions along the road accident data production process at the local level

The analyses demonstrated in the previous chapters revealed great number of anomalies in the data completed in the ARF. Anomalies yielding to item non-response and non-response errors, where errors that constitute *incorrect completion of data, incorrect interpretation of data,* and *omission of data* are uncovered during this process. Among the anomalies mentioned above, only the omission of data, which is basically referred to as the *item non-response error,* is predominately encountered throughout this assessment process. However, close to 80.0% of the problems thwarting the quality of the road accident data in South Africa are subsidised at the local level, since the inception of the accident reportage and data collection proceedings start at this specific level.

Ultimately, to tackle these anomalies a framework was developed to strengthen the existing protocol followed by the local authorities, in determining and identifying the areas or units that contribute huge anomalies. Nevertheless, the objectives of developing the framework were earlier stated in the section 1.2. On the other hand, the problems identified in each area of data processing were promptly and prudently evaluated, in order to ascertain the extent of the integrity of the data collected.

However, a flowchart illustration of the framework is presented in Figure 56, with the indications of the units and the officers involved. The flowchart also includes the introduction of additional quality dimensions, and the relational interfaces connecting the units involved. The integration of the quality dimensions into the framework is considered with the aim of extending the corrective measures from the initial point of data sourcing to the final point of the data transfer [at the local level].

8.1 Appropriate implementation of completeness, correctness and accuracy in the data collection and inspection processes

Primarily, road accident data collection is carried out based on the information obtain from an accident victim or an eyewitness of a road accident. In this context, the key problems encountered at this level revolved around response attitudes of the parties involved, and the level of competence of the reporting officer inspecting the accident. According to the findings obtained, it is understood that these problems can be controlled by assigning a higher rank or an experienced reporting officer to guide the junior officers, or perhaps the newly trained reporting officers through the process. Actually, this could alleviate the unexpected amount of incomplete and incorrect accident reports, and could also reduce the large amount of data exclusion *[omission].* In that case, the degree of readiness of all the allotted officers should regularly be checked to determine their level of commitment towards achieving adequate reports of any RTA.

The prime action to be performed, in conform to the illustration of the framework, will be implemented at the ARF office in the local police departments [SAPS] in the SM-WC 024, wherein the police officer who inspects a fatal accident submit a complete copy of the report of the accident to a Supervisory officer for verification. Herein, the framework developed suggests an improved set of procedures to the Supervisory officer, with the aim of improving the quality of the accident reports processed at the STD. The conceptualisation of the improved procedures is based on the integration of the quality dimensions like *completeness, correctness,* and *accuracy* into the formal procedures used for data collection as shown in the Figure 56 above. The functionality of the framework depends on the appropriate application of the three quality dimensions, and perhaps it could also be supported with a high degree of communication between the reporting officer and the Supervisory officer.

The implementation of the three quality dimensions is based on two related questions, which are expected to be followed by the Supervisory officer, before any conclusion could be reached on the degree of accuracy of a particular accident report as submitted by the reporting officer. In the RTA context, the two questions are formulated as:

- How adequate is the RTI gathered or collected with the use of ARF?
- How relevant is the RTI completed for each data element?

The basis of the two questions above is connected to the bond between the three quality dimensions. Primarily, pertaining to the two questions, the application of the three quality dimensions requires the involvement of both the police officer [accident reporting officer] and the supervisory officer. In this case, the degree of *completeness* of the RTI for each data element must be validated accordingly, to ascertain the level of relevance of the data completed in the ARF. This paves the way for the unproblematic documentation of the RTI, cleansed from any errors involving missing data and incorrect completion of data.

After a proper evaluation action is exercised with regard to the degree of completeness of the report; thus, the level of adequacy of the RTI should be validated properly through the implementation of *correctness* as a suitable quality dimension required to ascertain a standard measure of the RTA. Correctness measures the degree of reportage of the RTA, with respect to the appropriateness of the data in the ARF. In essence, it assists in measuring various errors committed in terms of item non-response and response errors (CIHI 2009).

The appropriate implementation of these two quality dimensions yields *accuracy* (Batini et al. 2009), which measures the outcome of the validation discussed in the preceding paragraph [see Figure 57]. Accuracy measures the degree of completeness and correctness of the definition of the data elements in the completed ARF. *The degree of accuracy of the RTI depends on the performance of the police officer/reporting officer, who inspects the accident* *and completes the ARF; and it also depends on the diligence of the Supervisory officer, who executes the validation process* (O' Day 1993; Sluis 2001).

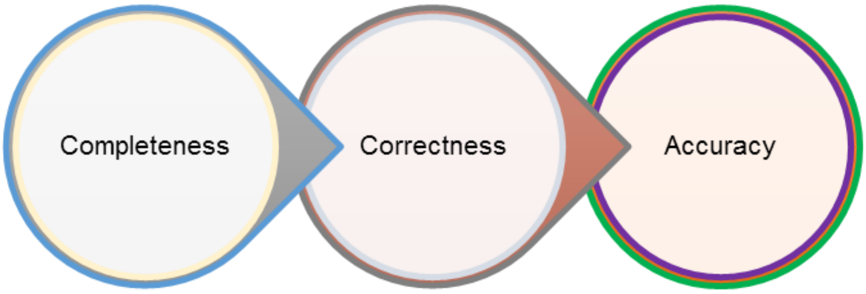

Figure 57: The sequential process of the three quality dimensions implemented in the first assessment

The strength of the validation procedures exercised by the Supervisory officer, is categorically directed on the sufficiency and precise provision of the necessary data required in the ARF. However, the appropriate execution of these procedures is considered inadequate at the local level; thus, this led to poor quality of the reports received by the DCO. As a result of this, the persistence of errors in the reports submitted at the local traffic department, heightens the amount of work to be done by the DCO. The three key problems identified in this section of road accident data are enumerated below:

- Unreadiness to produce error free reports,
- Low level of diligence demonstrated by the Supervisory officer towards sustaining a quality data assessment, and
- Inability to rectify errors detected according to the report from the Supervisory officer/DCO.

The appropriate implementation of the three quality dimensions discussed in this section will support process continuity in addressing the abovementioned key problems discovered in the first section of reporting and recording of road accident data by executing the tasks enumerated below:

- Regular communication of the relevance of quality dimensions to reporting officers before executing any reportage activities.
- Regular communication of the essentiality of the relevant data fields in the form to the reporting officers.

- Duty deployment should be based on the level of experience, competency, and the desire to be part of the inspection team.
- Regular follow-ups must be adopted to strengthening the quick processing of report form.

To drive the improvement process, the management or authorised superior officials in charge of the data validation process should demonstrate the willingness, and the necessary urge to implement the suggested tasks above in bolstering the quality reportage of RTA at the local level. In other words, the desire to validate the degree of correctness, completeness, and accuracy of the RTI reports appropriately, without skipping over any errors, strengthens the processing of data at the preliminary level. Moreover, since the implementation of correctness and completeness are parallel, an improved outcome will be achieved, which could enhance the purpose of actions highlighted below:

→ Uncovering of incorrect completion of data fields in the ARF, due to measurable factors frustrating the process.

→ Measuring the extent of appropriateness and commitment of the reporting officers.

→ Comparison of the degree of reportage under different climatic conditions such as rain, winter, spring and summer; and also, under different times of the day such as night-time, day-time and rainy time; and lastly, during different accident types single-vehicle accident, multiple-vehicle accident, and single-vehicle and a Pedestrian accident etc.

→ Distinct familiarity with the reliability and integrity of the reports submitted on a regular basis.

8.2 Appropriate implementation of timeliness in the data transfer

In this segment, *timeliness* of the appropriate transfer of data collected within the local level is considered. Timeliness is an important quality dimension that determines the actual time taken to process a particular information within an organisation. It is stated earlier in section 2.5.4 above, that the outcome of both the accuracy and completeness should be based on the time taken to process the information distributed within the local level, since *timeliness* defines the extent of appropriateness in the implementation of the three quality dimensions. The key purpose of applying *timeliness* along the information processing line is to check any delay that may be frustrating the transfer of information within the two authorised departments. In essence, it assists in accelerating the rate of data processing alongside some other important activities that could stabilise the system, such activities as valuing criticism to project improvement, appropriate distribution of data collection materials, and allocation of competent officers to facilitate reliable reportage.

In addition, this particular quality dimension measures the level of purity and usability of the data gathered in connection with the period of processing, from the first stage of data processing to the next stage of data processing (Lee et al. 2002; Wang 2004; Wang & Strong 2013; Azimaee et al. 2014). Timeliness determines the promptness to process data collected and query issued within the set timeframe (CIHI 2009; Azimaee et al. 2014). This process is frustrated due to the delay encountered in excessive processing time on the Supervisory officer's desk, which is due to the late response to corrections requested of the police officer who is responsible for the incident reportage, or pending road accident investigation reports. As a result of this delay, the chance of accelerating the required time to process the accident reports received by the DCO is truncated before the compiled reports reach the provincial level in due time.

In this context, *timeliness is observed as a vital tool in assessing the degree of responsiveness to the quality assessment of the information processed.* This quality dimension is established between the offices of the Supervisory officer and the DCO in the framework, with the purpose of facilitating a prudent validation process of the data collected through a solid relational communication.

The promptness in the transfer of information/data between the two offices, determines the level of appropriateness of the validation process applied. However, activities such as proper completion of the validation process, registration of the completed ARF, delivery of the completed ARF, and reviewing of the delivery and registration details, must be regularly executed in accordance with the implementation of the *timeliness* by deterring unnecessary workload in the next data assessment. Most importantly, by habitually observing a specified timeframe for daily delivery of completed ARF, actually it is necessary to work within the set period of time in order to avoid any delay in the weekly collation of validated reports at the local level.

From a viewpoint, the pertinence of *timeliness* between the offices of the Supervisory officer and the DCO should be built on strong communication interface, in order to facilitate a timely collation and dissemination of data at the provincial level. The application of this particular quality dimension at the initial data collection stage is likewise conceived important, because it all depends exclusively on the situation of the accident inspected, and the competency level of the reporting officers [refer to survey analysis in 7.3 above]. The essentiality of *timeliness* should be exercised regularly by the Supervisory officer to substantiate the extent of appropriate implementation of the first assessment, which comprises the combination of the completeness, correctness and accuracy.

Ultimately, the advantage of *timeliness* at the local level stimulates the possibility of measuring the proficiency of the units or departments involved in the processing of road accident data. If this could be achieved, then consistent reporting and dissemination of road traffic reports in due time at both the provincial and national levels would be possible. The two key problems identified in this section demonstrate the inappropriate implementation of *timeliness* along the transfer of information/data between the Supervisory officer and DCO are enumerated below;

- Excessive processing time of reports, which resulted in more backlogs of untreated workloads or reports.
- Late response to corrections or queries, which resulted in more backlogs of untreated errors in the completed ARFs.

In order to resolve these problems, valuable related tasks enumerated below should be executed to sustain continuity in the execution of *timeliness* between these two officials.

- Consolidating the communication interface between the two officials.
- Essentiality of prompt processing of information should always be a priority.

8.3 Appropriate implementation of consistency, accuracy and timeliness in the data capturing

This section involves three assessment stages, which are assessment three, four and five as depicted in the Figure 56 above. From the diagram, assessment three and four constitute the combination of the three related quality dimensions, such as consistency, accuracy and timeliness in the processes between the DCO and RTMC. Moreover, in the STD, the data capturing office comprises two officials, the Clerk and the DCO. The relevance of this office in ensuring quality data at the local level cannot be ignored. In this office, the first part of the assessment is carried out by the Clerk[31], who is supposed to crosscheck the appropriate delivery details of the completed ARFs, and also to allocate registration numbers to the completed ARFs.

In the process of executing his/her tasks, the Clerk is authorised to monitor the *consistency* along the course of collating the registered ARFs, with the intention of avoiding incomplete delivery of completed ARFs as listed in the DNF submitted by the reporting officers. This particular error can lead to misplacement and duplication of the completed ARFs, which could cause capturing of similar data for two different accidents.

The ultimate findings compiled by the Clerk, through the implementation of the third assessment will be forwarded to the DCO for proper verifications before a drastic measure can be taken. These findings are conceived as errors detected along the registration of the

[31] Refer to section 2.4.1 for more details on the responsibilities of the Clerk.

completed ARFs. The errors are simply detected only when some details deviate from procedures established for *consistency* validation, such errors as incorrect or missing details about:

→ the police station in which area the accident occurred,

→ the accident location address,

→ the responsible police officer for the accident reportage, and

→ the police station where the accident was reported and/or ARF was completed.

The DCO receives the compiled documents, and carry out the fourth assessment based on the primary complaints filed by the Clerk regarding the errors in the documents. The DCO is the intermediary between the local level and the provincial level, in terms of quality assurance of the RTI at the SM-WC 024. He/she is authorised to take decisions as regards any anomalies observe during the transfer of RTI within the local level. Although, some decisions that are beyond the jurisdiction of the DCO is handled by a superior traffic officer, managing the activities relating to training, education, and accident investigations.

The DCO is expected to verify the *consistency of reportage* and *interpretation* of the data fields as represented in the ARF. At present the responsibilities of the DCO is carried out by the Clerk. The key problems identified are:

¤ Burden of other responsibilities executed by the DCO/Clerk;

¤ Competency of the officials in charge of information processing; and

¤ Obsolete systems for information processing, which could lead to unnecessary delay in the recording, compiling and transferring of information caused by factors suchlike corrupt system, process incapacity, and stagnant progress.

The end products of these problems could only be tackled through the appropriate implementation of the recommended quality dimensions along the data processing line. In this case, *consistency* merited the initial validation process, followed by *accuracy* and *timeliness*. There are some quality components categorised under these quality dimensions, which define the usefulness of implementing a reliable data quality control system, in order to demonstrate the understanding on *how* and *where* it is necessary to implement a certain quality dimension.

The quality components like *consistency of reportage* and *interpretation* are perfectly suitable for the validation of the reports transfer from the SAPS. Firstly, considering the implementation of *consistency of reportage;* the DCO is tasked to identify the factors frustrating the degree of reportage. These factors prevent accurate reportage of accident in terms of *'incompleteness of data fields/elements'*, which leads to discovery of *missing data* (O' Day 1993; CIHI 2009; Azimaee et al. 2014). Secondly, the implementation of the *consistency of interpretation* validates the similarity of accident reports deliver from all the authorised units under the jurisdiction of SM-WC 024 (O' Day 1993; CIHI 2009; Azimaee et al. 2014). However, this is possible since a common ARF is applied in South Africa. The procedure verifies whether the reportage depicts any correlations in the reports received from the other authorised local units within the jurisdiction, in order to avoid duplication of reports. This procedure paves the way for an effective validation based on the ability of the DCO to adequately *'spot check'* the similarity in the reports.

Literally, at the DCO's office, validation process advances with the implementation of *accuracy* to ascertain the appropriateness of the data collected, followed by the implementation of *timeliness* to determine the promptness in transferring the data collated from one assessment stage to another (CIHI 2009; Azimaee et al. 2014). In this context, implementation of accuracy in this stage, measures the performance of the verification process initially executed at the local police department by the Supervisory officer, in accordance with the complaints filed by the Clerk.

The implementation of *accuracy* extensively promotes the relevance of the data quality components like the *completeness of reportage, right data, appropriate degree of detail, correct entry procedures,* and *freedom from response error.* The proper assessment of these data quality components determines the degree of accuracy that can be accomplished during the data validation process.

In the process of validating the degree of accuracy of the data collected, the DCO is tasked to validate the *completeness of reportage* first, in order to determine the relevance of information represented in the ARFs. The outcome realised will be applied to measure the ability of the reporting officers to comprehend the appropriate application of the ARF. At this point, the examination of the *right data* follows immediately, to determine if the data collected are appropriately detailed for the actual purpose before the completed ARFs are approved to be scanned and properly documented. Thereafter, the DCO will forward a query to the responsible police department in any of the five police departments in the same jurisdiction, based on the incorrect completion of information/data in the completed ARF. A prompt response to the query is required to be carried out in order to avoid any further delay in the transfer of information to the provincial level.

Right after the implementation of the accuracy, then *timeliness* is implemented in this assessment to guide a prompt processing of the inaccuracies detected, since the change to be made will be re-evaluated before a compiled report could be forwarded to the provincial office. It is recommended that the implementation of the quality dimensions should be carried out regularly with caution, in order to avoid skipping over the inaccuracies and inconsistencies marring the acquisition of quality data. The compiled reports are forwarded to the provincial level/RTMC, where further assessments are carried out on the reports received from the local level.

The actual tasks executed to address the above listed problems in supporting a continuous validation process to yield reliable quality data are enumerated below:

- Assigning a relief officer.
- Consecutive periodic training on newly acquired data processing software.
- Regular update of data processing systems and regular backup of all information.

9. Summary of findings

The findings acquired in this study answered the three research questions formulated in the section 1.2, regarding the simplicity and comprehensibility of the ARF to the users, and the appropriate completion of the ARF by the users to attain a practical use in understanding causes of road accidents and managing resources in preventing road accidents in the Stellenbosch area. In that case, the simplicity and comprehensibility of the ARF were evaluated along with the appropriate completion of the ARF, through the findings gathered from the observations of the form users and the analysis of the errors committed during the process of data collection.

The simplicity and comprehensibility of the ARF determines the chance of producing adequate and decipherable data to improve poor quality data, since the degree of incomprehensibility of the form to its users contributes categorically to the errors committed along the sequential procedures developed for the collection of data. The problem of incomprehensibility was measured through the survey completed, which answered the first research question. Although, based on the critical evaluations carried out on the data or information represented in the ARFs, findings confirmed the actual data fields that seemed incomprehensible to the form users. The evaluation of the survey offers some valuable understanding through the opinions of some selected participants, with regard to the degree of simplicity and comprehensibility of the form to its users.

According to the survey result, a large number of participants claimed that they absolutely comprehend the information provided in the ARF, while a few participants claimed that they are having challenges with the understanding of the information in the form. Due to this, a certain number of participants acknowledged the need to re-examine the arrangement of the information in the ARF. The split of opinions corresponded to the participant's department, that is Police or Traffic. This suggested that the training differed between the departments. To buttress this statement, a number of participants claimed that they undergone the necessary training ahead of any incident reportage, while others declared that no particular training was undergone before the exercise. It could be understood that training provided is effective, but the programme is not administered consistently.

The adequacy level of the training given to the form users was further evaluated in the area of accident sketch and accident description. From the findings acquired, it is understood that poor training hampers the ability of some participants, in constructing a good accident sketch and a good written description of a particular road accident. Although, a large number of participants asserted that they benefit from the training undertaken, but the impact of the training is not reflected in their ability to produce informative descriptions and sketches of a particular road accident.

Also, a small number of participants agreed that they lack the necessary skills required to construct a detailed sketch of a road accident. This result is slightly different from the result obtained in the analysis of the accident description. From the analysis, a higher number of participants declared their uneasiness in producing a detailed description of road accidents than a detailed sketch of road accidents. One of the key problems affecting the appropriate composition of a detailed description of road accident is a *language barrier*.

In addition, general observations were carried out to ascertain the key data fields in the ARF that appeared difficult to comprehend during data collection activities. Findings reveal that fewer number of participants demonstrated their uneasiness towards the difficulty of collecting adequate data for some data fields in the ARF. Amongst the data fields in the ARF, only the *Light condition* was considered not challenging.

The importance of understanding the first research question extends to the evaluation of the participants' level of competence in completing the accident information appropriately. In the process, the time taken by the participants to complete an ARF for three forms of accident, such as Single-vehicle accident, Multiple-vehicle accident, and Single-vehicle and a Pedestrian accident was determined.

According to the findings obtained, a large number of participants preferred to use 6 to 15 minutes, as their best time to complete the ARF for the three forms of accident; though some other participants declared that they prefer to use 16 to 30 minutes to complete an ARF for the same reason. In this case, the accuracy and timeliness of the process depend completely on the circumstances of the accident, because a fatal accident will demand more time whether it involves only one vehicle or multiple vehicles. Nevertheless, the condition of the accident victim predicts the swiftness of completing a form rightly. More so, the circumstances surrounding the incident could be worsened if such incident occurs in an adverse weather condition or night time. Perhaps, this could extend the time required to collect data concerning any type of accident. Therefore, the chance of committing errors during adverse weather or at night is probably higher.

Thus, as part of the circumstances that encourage the making of errors during data capture in the ARF, the influence of weather conditions is evaluated. By considering this, three major weather conditions were considered as suitable metrics for measuring the observations of the participants on the completion of ARF, such weather conditions as day-time, night-time and rainy period. The omissions of data obtained reveals that a great number of participants find the completion of ARF challenging in any of the three weather conditions. Comparison of the individual results indicates that more concerns were demonstrated towards the difficulty of completing ARF in the rainy period than in the other two weather conditions. The probable factors contributing to the difficulty experienced by the participants during these periods were enumerated in the last paragraph of section 7.4.

The above results were substantiated with the practical exploration of errors discovered in the information completed in the ARFs. The errors discovered vary from one form to another, which could be the inability of the form users to obtain the right data for a particular accident event, or the inexperience, less concentration and violation of instructions by the form users while completing the form. The actions mentioned were evaluated to determine the existence of errors during the data evaluation process. The process implemented was based on the right definitions of the three forms of error with regard to RTA; for instance, errors described as *response error* are categorised as the indication of double answer options [where it is instructed to indicate only one answer option], and incorrect indication of relevant answers. On the other hand, errors attributed to *item non-response error* are identified as the omission of valuable information, although the case of *unit non-response error* is unsubstantiated in this study.

The classification of errors enhances the ability to measure the occurrence rate of non-captured data by obtaining the average rate of the data lost per field, and the distribution analysis of the data lost. According to the results obtained, for the Accident related factors there were total average estimate of 324 missing data points monthly. Also, for Road related factors, there were total average estimate of 486 missing data points monthly. Lastly, for Human related factors, there were total average estimate of 1,124 missing data points per month. In Accident related factors, above 90.0% of this average estimate is contributed largely by three data fields, namely Summary of persons involved, Severity of injury and Accident type. On the other hand, in the Road related factors, such data fields as Junction type, Built-up area, Traffic control type, Road type and Speed limit on the road comprised most of the missing data. A similar observation was attained in the Human related factors, where large disparities were observed in the data completed in the ARF. Many data fields were left empty, which led to a huge amount of item non-response errors assembled as part of non-captured data.

The distribution analysis of the non-captured data showed that data errors were mostly counted between the ranges of 1-5 errors, 61-80 errors, 94-110 errors and 113-133 errors in the three related factors, but the accumulation of errors vary monthly. In view of the disparities discovered in the data completed in the ARF, however, the perceptions of the form users were sought in order to have an overview of how often a form user commits errors when completing the ARF. According to the survey result, 1-3 errors are frequently committed by some participants, whereas only one participant confirmed the possibility of committing more than 15 errors in one data collection exercise. Other remaining participants demonstrated their level of intolerance towards the possibility of committing an error, whereas the findings obtained reveal the tendency of committing a large amount of errors in the process of collecting data.

With the proportion of the errors discovered in this study, one could simply conclude that great amount of data were inappropriately completed and excessively omitted. The omission of data is ascribed to non-response from the accident victims, or perhaps late reporting of road accidents to the nearest authorised local department, which inhibits the ability of the reporting officer to have a full understanding on how the accident actually occurred. At this point, the competency of the reporting officer is tested towards the ability to determine the actual cause of the accident.

In this context, little or no attention are currently exercised towards the need for correcting or reducing the issue of data mishandling. Additionally, large amounts of data omissions are restricted mainly to some particular data fields, which are perceived to be problematic to the form users due to some explainable circumstances. Also, slack data management along the preliminary level of data sourcing at the local level contribute greatly to the amount of data lost.

Descriptive statistics were used and interpreted accordingly to satisfy the other two research questions in this study. In Accident related factors, analysis of the practical data shows that 2,451 road accidents were registered in 2012, which resulted in few number of fatal injury and several nonfatal injuries mostly in the daytime [daylight]. This yielded an average value of 204 registered road accidents per month. Moreover, the descriptive statistics of both the Accident date and the Day of week were carried out to determine whether the data collected establishes any relationship between the dates and weekdays that accidents occurred. This revealed the weekday with the most occurrence of RTAs. This enhances the ability to determine the specific factors or conditions that could be contributing to the cause of accidents on the identified weekdays. The results affirmed that a large number of road accident cases is reported on *Fridays* in the Stellenbosch area in 2012. Also, in the analysis of the Accident date, Thursday recorded the second highest amount of road accidents among the weekdays.

Whilst exploiting the data field, Day of week, Saturday recorded the weekday with second highest amount of road accidents among the rest of the weekdays. Based on this result, however, the two results contradict each other due to large existence of item non-response errors in the data fields. The two results further established that among the remaining weekdays, Sunday recorded lowest amount of road accidents. Conclusively, the quality of the data assembled in the Day of week field, suffers from a large amount of missing data, which is due to users' omitting in the data form.

In the case of the Time of accident, findings obtained establish that the actual time-intervals with the high occurrence of RTAs in the Stellenbosch area fall within the time-interval ranges of 7:00:00 am to 6:59:00 pm and low occurrence of RTAs between 7:00:00 pm to 6:59:00 am. In a more simplified way, the lowest occurrence time of road accidents are the early hours before the daybreak with a lowest estimate of 16 road accidents, while the most occurrence time of road accidents were recorded during the early hours before night-time with a highest estimate of 209 road accidents. As part of the findings, zero accident occurrence periods were observed mostly in the early hours ahead of the daybreak and in the late hours of the night-time. This elucidates that more road accidents occurred in the Stellenbosch area between the dawn and dusk periods in 2012, than at night.

Further analysis performed explains the relevance of the two peak periods in supporting the knowledge of the distribution of rush-hour traffic within a set period of time. It is concluded that traffic congestion contributed to a large portion of road accidents recorded between the dawn and dusk periods, most especially in the clear weather condition. The two peak periods signify the effect of large movement of the residents and workers within and outside the Stellenbosch area.

The finding obtained from the analysis of the Accident type shows that a high amount of road accidents occurred at the road intersections [junctions] within the Stellenbosch area. The result demonstrated that most accidents occurred while drivers arrived or departed from the intersections. Accidents occurring at the intersection, happened mostly as a result of some traffic offences like red-light running, wrong-side vehicle overtaking, and violation of pedestrian crossing rules. The analysis of this particular field led to the identification of causal factors leading to each type of accident.

Other results obtained indicates that large number of accidents registered or recorded at the Stellenbosch Traffic Department were contributed by the Road related factors, as one of the causal factors of the road accidents. From the findings discovered, it is realised that large amount of road accidents predominantly occurred within a speed limit of 60km/hr in the built-up areas that is densely populated. More so, other findings demonstrate that fewer road accidents occurred on the highways, urban freeways, and rural roads located in the non-built up areas. Contrasting data points acquired in the Speed limit on the road and the Built-up area indicated substantial data incompleteness due to item non-response error and response error discovered. In actual fact, the number of the data points acquired in the two data fields is expected to be equal if these errors are avoided. This effect indicates a degree of inconsistency in the process of achieving accurate and complete procurement of accurate road accident data at the municipality [local] level.

Results achieved also show that great number of road accidents registered in 2012, occurred mainly on the *dual* and *single carriageways* within the speed limit of 60km/hr depending on the locations of the accidents. Considering the Junction type, it is logical that risks posed by uncontrolled junctions are larger, compared to the controlled junctions around the Stellenbosch area. This explains the higher rate at which road accidents occurred at the uncontrolled intersections. Human errors, in making decisions at these junctions, are a main contributor.

To strengthen the above illustration, the estimated proportion of road accidents attributed to the *'not a junction or crossing'* indicates that most of the accidents occurred at a certain distance away from the intersections [junctions], probably along a straight road. Thus, at that point of incident the road does not connect with any intersection or crossway. In the case of intersection accidents, the potential factors responsible are enumerated in the last paragraph in subsection 5.2.2. It is further uncovered that robot controlled intersection also recorded a large number of road accidents at 32.0% along with stop sign controlled intersection with 13.0% in SM.

More so, the findings derived from the analysis of the road surface, indicate that more than 90.0% of the road accidents occurred on the tarmac surface, having an estimated proportion above 95.0% of the road surface quality to be in good condition. This insight leads to the realisation that despite the good condition of the roads, road accidents along the Stellenbosch roads are inevitable due to such factors as vehicular conditions, human behavioural characteristics, and environmental challenges. Moreover, further findings discovered in this study show that 85.0% of the road accidents occurred on the roads with good road marking visibility, with 91.0% of the road signs considered good. In addition, fewer number of road accidents [3.0%] are due to road obstruction issues and 90.0% of the road accidents in Stellenbosch occurred on straight road direction, with 85.0% of the road accidents indicating that vehicle[s] is/are on correct road lane position before the incident[s].

In the evaluation of the Human related factors, as one of the causal factors of road accidents in SM, the findings show that the age group within the range of 21-40 demonstrated the highest percentage [54.0%] of drivers/cyclists involved in road accidents in the area. Further analysis shows that the number of drivers/cyclists aged between 41-60 involved in road accidents is significantly lower at 29.0%. The lowest frequency distribution of drivers/cyclists involved in the monthly occurrence of road accidents falls within the range of ages 81-100, with a percentage estimate of 0.5%.

The similarity observed between the mean and median ages, as displayed in the Table 24 in section 6.1 and the box-and-whisker plot in Figure 37, validates the predominance of the drivers/cyclists within 21-40 age-group as the most vulnerable road users in the SM. Basically, the particular age-group with the highest frequency estimate comprises students, manual workers and tourists; depending on their periods of visiting. These sets of people contribute tremendously to the traffic issues, daily increase in the volume of road users, and the rampant occurrence of road accidents in the area, because the municipality is generally known for farming, tourism, academic and business transactions.

The above illustration is further strengthened with the analysis of the drivers'/cyclists' countries, to ascertain the degree of involvement of South Africans and foreign nationals in road accidents registered in the Stellenbosch area. Findings show that less foreign nationals were involved in road accidents, since a large proportion [95.0%] of South Africans were involved in road accidents in the municipality. The remaining estimate constitutes a quota reflecting the contribution of foreign nationals from countries like Namibia, Zimbabwe, Germany and Netherlands. This suggests that limited knowledge of the South African driving systems and road use regulations contribute to the involvement of foreign nationals in the road accidents.

In addition, the results obtained from the analysis of gender and race groups clarify the involvement of the individual groups of people in road accidents. Observably, in the gender analysis, male drivers/cyclists are more vulnerable to road accidents than female drivers/cyclists in the area. However, some data points were lost to undefined gender identity and errors. The analysis of the race group shows that white drivers/cyclists are more involved in road accidents than any other race group in the municipality. This relates to more white drivers/cyclists on the roads than the other related groups. Further observation shows that the highest peak of white drivers'/cyclists' involvement in road accidents was observed in February, while for the other two prominent groups in the Stellenbosch locality such as Coloured and Black (African) drivers/cyclists, both demonstrated their highest peak in September.

Other significant findings gathered demonstrate that a fair number of drivers/cyclists at 56.0% definitely complied with the use of seatbelt/helmet during the road accidents, while a minimal percentage of drivers/cyclists was confirmed without the use of seatbelt/helmet with 36.0% of the data not accessible. From more findings, it is realised that a minimal percentage [1.0%] of drivers/cyclists was evidentiarily confirmed to be under the influence of liquor/drug use. More so, 42.2% of road accidents occurred when vehicles involved were travelling straight along the roads in the area, and other activities like turning right, stationary [waiting in traffic], reversing and parked are part of the major contributors to road accident occurrence in SM.

9.1 Conclusions

The quality of data completed on the Accident Report forms [ARFs] collated at the Stellenbosch Traffic Department [STD], necessitate a significant investigation into the factors or problems preventing the necessary improvements that should be implemented. With the aim of achieving this, however, the investigation was initially carried out at the Accident Response Unit [ARU], where the annual accident reports were warehoused under the management of the Stellenbosch Traffic Department. The investigation was executed based on the need to clarify the "simplicity and comprehensibility" of the Accident Report forms to the users, and the appropriateness of the data completed in the form to evaluate the practical use of the data in “understanding causes of road accidents” and “managing resources in preventing road accidents” in the Stellenbosch area.

Stellenbosch Traffic Department and the South African Police Service [SAPS] in the Stellenbosch Municipality-Western Cape 024 played vital roles in achieving the objectives of this investigation, since a mutual relationship exists between the two departments to ensure an appropriate validation of the data collected, starting from the point of reporting to the point of recording by preventing the presence of anomalies along the data processing line at the local level. Despite the practicality of the validation, degree of completeness and correctness of the data completed in the Accident Report forms suffered huge anomalies, which were uncovered by determining the factors contributing to the accumulation of errors thwarting the quality level of the road accident data completed in the form.

In the process, five major factors were determined as the key contributors to the persistent anomalies uncovered along the reporting and recording of the road accident information in the Stellenbosch Municipality [SM]. The five factors determined are *human behavioural characteristics, data collection tool, competence of the personnel, reliability of the data collection procedures and consistency of the validation processes.* In this case, the ‘human behavioural characteristics’ defines the readiness of accident victims or eyewitnesses to assist in the procurement of relevant information, and the commitment of the reporting officers in indicating accurate information. In addition, the ‘data collection tool’ determines the level of simplicity of the form users based on the comprehensiveness and completeness of the information sufficiency. Also, the ‘competence of the personnel’ demonstrates the ability to comprehend the information and the zeal to acquire the right information appropriately. The ‘reliability of the data collection procedures’ describes the efficient level of the procedural system implemented at the local traffic department. And the last factor, ‘consistency of the validation processes’ determines the degree of steadiness in substantiating quality of data and promptness in the transfer of data for further assessment.

The above-mentioned factors were measured through a visual exploration of the data in the form, analysis of the practical and impractical data and the development of a survey to seek the opinion of the form users. According to the findings, series of related activities executed at the local level were affected by poor communication between the locally authorised departments. The existence of poor communication propagates more errors and untimeliness in the delivery of reports. Due to this, the probability of securing quality data is lowered between these locally authorised departments. This issue frustrated various preceding quality improvement efforts that were attempted. In the process of data capture to electronic means, the existence of many errors was detected in the data completed in the Accident Report forms. At the local level, errors determined indicates that human inaccuracies have paramount influence on the data completed in the Accident Report form. No single completed Accident Report form was found free of errors during this investigation, since many relevant data fields in the form are left either unrepresented or misrepresented due to the competence level of the form users.

This effect is attributed to haphazard execution of training and validation processes, which probably led to the errors committed, and perhaps lack of rapid response to the corrections requested on the resubmission of report forms within the local authorised departments. In regard to the excessive influence of the human inaccuracies in the aspect of the data acquisition, a survey exercise was initiated to substantiate the involvement of the form users in the accumulation of errors. The findings gathered from the survey exercise enhanced the determination of the right strategy employed in developing a suitable framework. According to the survey responses, some data fields are considered challenging during data collection activities to 17.6% form users, such data fields as accident location, driver's particulars, road details, pedestrians'/passengers' details etc. This is basically attributed to factors like lack of sufficient training, incompetence of the form users, hit and run cases, late reporting of incident, lack of cooperation from the accident victims *[only in good state of mind]* or eyewitnesses, and inability to persuade accident victims for more relevant information.

Besides, crucial concerns were raised regarding the construction of a detailed accident sketch and description, because many accident sketches in the Accident Report forms were rendered very poorly. Though, it is advantageous to analyse the accident sketch and description when attempting to understand the actual cause of a particular accident. Such analysis could result in valuable clues towards the development of appropriate countermeasures by the traffic engineers. In essence, the composition of a well written accident statement guides the understanding of the police, the traffic and the insurance concerning the primary cause of the accident. However, it is more essential for the traffic management to adopt a better system to train the form users [reporting officers], particularly on how to construct a detailed sketch and description in accordance with the necessary features required as enumerated in the section 7.6 above.

Ultimately, it is determined that due to lack of comprehensive training and differences in the training between the locally authorised departments, not all form users comprehend the application of the form. This has a huge effect on the sufficiency of the right data in the Accident Report form. Recently, new blank Accident Report forms were being distributed across the municipalities in the Western Cape Province. These newly improved forms are designed to reduce the effect of lost copies of the completed forms, in order to boost prompt information processing at the local level and the proper handling of the forms. The Accident Report forms come with carbon copy duplicates attached to each of the forms. Although, this improvement does not completely resolve the issue of errors committed during the completion of the Accident Report form. Also, access to the online road accident database should be made possible to assist the improvement of information disseminated to the public.

A general review carried out on the simplicity and comprehensibility of the Accident Report form suggested some valuable improvements in areas like information arrangement and interpretation, and addition of new information which perhaps required the exclusion and inclusion of information in an attempt to promote effective application of the form. According to the findings deduced, the current arrangement in the Accident Report form is not relationally structured, to aid relative understanding or interpretation of the information or data fields in the form to the users. The arrangement was rather based on the utilisation of the available space in the form to accommodate as much as possible information to a feasible extent, which causes omission of some field entries during data collection activities, such data fields as *Road type* and *Junction type.* Hence, this makes it easy for the form users to skip over them due to their current position in the form, since the attention of the form users are more directed to the particulars of the accident victim than many other crucial information. This issue has led to a huge number of item non-response errors [data omission] with regard to the two data fields. In addition, the review shows that necessary modifications are significant, in which the application of the accident technical questions could assists in guiding the re-arrangement of some data fields [see subsection 3.3.2 above]. The accident technical questions are interrelated in the process of understanding every aspect relating to the cause of any particular accident.

Apparently, the information in the ARF is simple to comprehend but some variables like *'dawn/dusk',* which literally have two different interpretations, were assumed as same interpretation in the form. Similarly, accident data fields like *'vehicle manoeuvre/what driver was doing'* distinctly can boost sufficiency of road accident data, if separated and modified to measure additional data fields. To be more specific, the data field title *'what driver was doing'* can be simplified into *'driver's actions',* to represent well defined cases of distractions that predominantly lead to road accidents in the Stellenbosch locality, such cases as *passenger influence, radio/CD, looking at scenery, driver fatigue, cell phone, eating* and *drinking,* and many more others. The significance of these data elements can simply be used to enhance the result currently derived from the *'vehicle manoeuvre'.* Though it might be difficult to measure vehicle manoeuvres, its great feasibility depends on the grade of devices or facilities at the disposal of the traffic management. On the other hand, the inclusion of valuable information or data fields like *'number of casualties'* and *'insurance details'* in the Accident Report form were recommended based on the opinions of the form users. Addition of the *'number of casualties'* in the top section of the first page in the Accident Report form could deter data insufficiency in the *'severity of injury'* and *'summary of persons involved',* which contributed extremely to the extent of uncaptured data in the two data fields. In similar way, the inclusion of *'insurance details'* would reinforce the chance of discouraging underreporting of data fields.

Currently the validation process practised in the locally authorised departments simply focuses directly on the assurance of accomplishing a set of operational procedures. Although, none of the operations executed along the validation process was skipped over, but lack of commitment and tenuous practices discredited the efficiency of the system. From all indications, the existence of anomalies discovered along the process substantiate the motive for the development a framework to illustrate key problems with inappropriate implementation of quality dimensions and tasks executed to address the problems in minimising the mishandling of the road accident data in the locally authorised departments. The framework will be used to drive the appropriate implementation of the recommended quality dimensions to strengthen the capacity required and improve the data processing line within the locally authorised departments. Therefore, the framework assists in exposing errors in the completed Accident Report forms promptly, and dictates the right actions required to stabilise the proper processing of the road traffic information from one unit to another. Furthermore, the framework provides checklists comprising suggested solutions to current problems which could prevent some imminent problems. The checklists consist of the sequence of responsibilities to be executed, to sustain the continuous application of the quality dimensions along the data processing line. The performance of this framework, will depend on the commitment and the readiness of the units involved in the handling of the road traffic information [RTI] at the local level.

Despite the errors in the Accident Report forms, the data collected and analysed was still of use in "understanding causes of road accidents" and also in "managing resources in preventing road accidents" in the Stellenbosch Municipality [SM]. According to statistics obtained, the number of road traffic accidents reported in the SM-WC 024 in 2012, led to the death of 11 road users, one of the lowest fatality records among municipalities under the jurisdiction of the Western Cape Province. More than half of the accidents that occurred within the Stellenbosch locality are linked to periods of busy activities, such periods as academic resumption and non-holiday periods, which necessitate the consistent monitoring of the behaviours of the road users to curtail road accidents within the locality. The statistical results demonstrated that fewer than 2.0% of road accidents are related to the abuse use of liquor or drug. And also, fewer than 10.0% road accidents are due to absence of use of safety belt [seatbelt] and safety helmet [see Appendix D-D.1.2 & D.1.4].

In the Stellenbosch Municipality, as shown through statistics, road accidents occurred mostly on Fridays, but the position of the second highest weekday was unsure between Thursdays and Saturdays due to lack of data correlation between the Accident date and Day of week. This is so because many option boxes were left unmarked in the Accident Report forms which led to incomplete data. It is further noticed that among the holiday periods in the dates/days, only “Family Day” recorded 10 road accidents. This demonstrated some consistency in the data collected, but the completeness and correctness of the data collected were mediocre. The benefit of analysing accident date/day fields, is that it offers the traffic management the opportunity to quantify the rate at which road accidents are reported and documented daily. Also, this could guide an accurate distribution of resources, in terms of human capacity and financial aids, and appropriately motivate sufficient resources to control the occurrence of road accidents within the Stellenbosch district.

To support the above findings, it is understood that traffic periods with large number of reported accident cases fall within the *dawn period* between 7:00:00-7:59:00 am, and in *dusk period* between 4:00:00-4:59:00 pm. These findings indicate that residents travel mostly within these periods around the Stellenbosch area. This literally means that accidents reported in the hours of the daylight, occurred mostly after sunrise and before sunset in the daytime. This recommends that traffic management should be more observant of the periods and the weekdays in which traffic accidents occur most, specially periods with intensified traffic congestion, because most of these road accidents occurred on the dual and single carriageways in the Stellenbosch area, particularly at the intersection points within a speed limit of 60 km/hr in the built-up areas. Analysing the *Time of accident* enhances the capacity of the traffic management in identifying the specific periods, that probably contribute most to high occurrence of RTA, and the exact period when traffics are most intense.

Furthermore, assessment of the contributory factors through descriptive statistics demonstrate that most road accidents happened on tarmac road surfaces with good conditions, while fewer were linked to poor road surface conditions like concrete, gravel and dirt. More so, other statistical results showed that a great number of road accidents were recorded in the daylight, predominantly in the summer periods, than at any other seasonal periods regardless of any particular condition of the road surface. In the case of other variables, such as *dusk/dawn* and *night-with or without streetlight,* a constant increase in the number of road accidents were observed mostly within the winter periods, and slightly across other seasonal periods including spring and autumn. Other results indicated that a large proportion of these accidents occurred during clear weather conditions with great visibility. The possible instigating factors leading to road accidents in a clear weather conditions with great visibility are identified as speeding, influence of alcohol and drugs, driver-impatience, vehicular condition and many more. The remaining smaller proportion constitutes records of road accidents that occurred during raining and cloudy weather conditions with less or poor visibility.

Evidently, among the races involved in road accidents, according to the statistical facts gathered, white drivers/cyclists are most vulnerable in the prevalent occurrence of road accidents in the Stellenbosch Municipality-Western Cape 024, perhaps as a result of their dominance in the use of the roads more than any other race groups. Findings illustrate that the proportion of white drivers/cyclists involved in the road accidents in 2012, is in excess of 18.0% greater than the sum proportion of the remaining drivers/cyclists in the other race groups. The analysis of the gender identities and nationalities demonstrated that more male drivers/cyclists, either as South Africans or foreign nationals, are involved in road accidents than their female counterparts. Statistical evidence shows that mostly South Africans, irrespective of their race groups, were involved in road accidents in the Stellenbosch while fewer foreign nationals were involved. The involvement of foreign nationals in accidents may be attributed to road use regulations and different driving systems. In addition to this, the investigation shows that drivers/cyclists under the age of 40 years are more susceptible to road accidents than any other age groups in the municipality.

9.2 Limitations of research

In the process of evaluating the available data, some limitations were encountered which rendered some valuable information unusable and void. The limitations hampered the possibility of obtaining relevant and complete data from the completed ARFs. Furthermore, a number of information could not be recovered because they were excised or cut off from the completed ARFs. These problems could be attributed to managerial deficiencies, which may be due to an ineffective inspection from the officers supervising the affairs of the accident records.

Thus, the list of limitations encountered along the evaluation process is provided below with detailed explanations, according to the extent of their effects on the data assessment.

- **Use of old ARF;** the use of old ARF excluded or excised new features added to strengthen the sufficiency of relevant information, which are observed as significant elements to the possibility of devising formidable countermeasures to curtail accident occurrence. Under close observation, a number of information was missing in the ARFs, which contribute to the inability of the accident reporting officers to achieve complete incident coverage. This impact, directly rendered some information impractical for research motives. The dimensions of the variables considered paramount for data grouping are affected, thereby reduce the magnitude of data that could be gathered. This limitation was occasionally encountered along the data collection process.
- **Duplicate copy;** during the evaluation process, duplicate copies of some completed ARFs were found. These particularly caused the duplication of data collected during

data assemblage, and were sorted out by crosschecking and comparing the registration number of the reported accidents.

- **Punctured answers;** this occurs when completed ARFs are being photocopied, which are mostly accident coverage performed by the police officer as explained by a senior traffic officer, who works in the ARU. According to observation carried out, some answers were found perforated while trying to create a record file for accident reported daily. Apparently, this issue rendered many marked answers useless, and simultaneously disallowed complete collation of indicated answers for data analysis.
- **Ineligible writing and spelling errors;** these two limitations influenced the possibility of structuring some written attributes into metadata through allocation of numeric codes to sufficient analysis. Some few cases were discovered in the completed ARFs, typically in the areas such as accident location, accident sketch, and accident description. In addition, some spelling errors were committed through the written statements, which are not reflecting the accurate meaning of the statement portrayed. These issues are associated with two key factors considered as level of literacy and lack of sufficient training on how to construct an interpretative statement. The abovementioned factors affect the technicality of completed ARF submitted. This peculiar problem consumes more time of the DCO to discover the right words or actual meaning of the statement made in the completed ARFs, with the aim of determining the accurate data for the right purpose.
- **Misplacement;** this specific issue imposed unnecessary delay along the process of collecting data. The issue of misplacement of completed ARFs are considered as errors due to personal influences such as fatigue at work, misconception of orderliness, and lackadaisical attitudes of the personnel towards their tasks.
- **Availability of the survey participants;** this issue affects the timely completion of the questionnaire distributed to the participants who are experienced officers in the reporting of the road accidents. In addition, it is very difficult to encourage these officers to participate in this exercise because they are always on duty.

The consequence of these limitations affects the findings in the aspect of data quality level and the adequate participation of the reporting officers to support adequate understanding of the form users and the problems encountered while completing the form.

9.3 Future research work

In due course, a future research topic would be to extend this investigation to the provincial level, where the bulk processing of the road traffic information is being executed daily, and the use of electronic data processing is paramount. At this level, the performance of the units involved can be evaluated to determine the cause of automated errors or anomalies existing along the data processing line, and suggest a valuable and feasible improvement.

10. References

Adams, M., 2001. *The Road Traffic Management Corporation: A Case Study*. Rand Afrikaans University.

Agran, P.F., Castillo, D.N. & Winn, D.G., 1990. Limitations of data compiled from police reports on pediatric pedestrian and bicycle motor vehicle events. *Accident Analysis & Prevention*, 22(4), pp.361–370.

Agran, P.F. & Dunkle, D.E., 1984. A Comparison of Reported and Noncrash Events. *Accident Analysis & Prevention*, 17(1), pp.7–13.

Aitken, A. et al., 2004. *Handbook on Improving Quality by Analysis of Process Variables*, European Commission EUROSTAT.

Alsop, J. & Langley, J., 2001. Under-reporting of motor vehicle traffic crash victims in New Zealand. *Accident Analysis & Prevention*, 33(3), pp.353–359. Available at: http://www.sciencedirect.com/science/article/pii/S000145750000049X [Accessed February 13, 2014].

Amoros, E. et al., 2009. *Reporting on Serious Road Traffic Casualties: Combining and Using Different Data Sources to Improve Understanding of Non-Fatal Road Traffic Crashes*,

Anjuman, T. et al., 2007. ICME2007-AM-30 Road Traffic Accident: A Leading Cause of the Global Burden of Public Health Injuries and Fatalities. In *International Conference on Mechanical Engineering*. Dhaka, Bangladesh: Accident Research Center (ARC), BUET, pp. 29–31.

Asia Injury Prevention Foundation et al., 2010a. How to Conduct A Situational Assessment 2. In *Data systems: A Road Safety Manual for Decision-Makers and Practitioners*. pp. 21–45.

Asia Injury Prevention Foundation et al., 2010b. How to Design, Improve and Implement Data Systems 3. In *Data systems: A Road Safety Manual for Decision-Makers and Practitioners*. pp. 49–76.

Atzeni, P. & Antonellis, V. De, 1993. Relational Database Theory. *Benjamin/Cummings*.

Austin, K., 1995. The identification of mistakes in road accident records: Part 2, casualty variables. *Accident Analysis & Prevention*, 27(2), pp.277–282. Available at: http://www.sciencedirect.com/science/article/pii/000145759400066U.

Austin, K., 1995. The identification of mistakes in road accident records: Part 2, Casualty variables. *Accident; analysis and prevention*, 27(2), pp.277–82. Available at: http://www.ncbi.nlm.nih.gov/pubmed/7786394.

Azimaee, M. et al., 2014. *MCHP Data Quality Framework*, Winnipeg, Manitoba, Canada. Available at: http://www.umanitoba.ca/faculties/medicine/units/community_health_sciences/departmental_units/mchp/protoco l/media/Data_Quality_Framework.pdf.

Baguley, C., 2001. The Importance of A Road Accident Data System and Its Utilisation. In *International Symposium on Traffic Safety Strengthening and Accident Prevention*.

Nanjing, China: Transport Research Laboratory, United Kingdom, pp. 1–20.

Batini, C. et al., 2009. Methodologies for Data Quality Assessment and Improvement. *ACM Computing Survey*, 41(3), pp.1–52. Available at: http://doi.acm.org/10.1145/1541880.1541883.

BCM, 2005. *Accident Data Management Protocol*, East London, South Africa.

Bester, C.J. & Geldenhuys, F., 2007. Speed Trends on Major Roads in South Africa. In *Proceedings of the 26th Southern African Transport Conference (SATC2007)*. Pretoria, South Africa: Document Transformation Technologies cc, pp. 600–609.

CIHI, 2009. *Accuracy, timeliness, comparability, usability, relevance: The CIHI Data Quality Framework*, Ottawa, Ontario Canada: Canadian Institute for Health Information. Available at: www.cihi.ca.

Cimbala, J.M., 2014. Histograms. *Histograms*, pp.1–4. Available at: https://www.mne.psu.edu [Accessed December 5, 2015].

Decoster, J., 2006. *Testing Group Differences using T-tests, ANOVA, and Nonparametric Measures*, 348 Gordon Palmer Hall, Box 870348, Tuscaloosa, AL 35487-0348: Department of Psychology, University of Alabama. Available at: http://www.stat-help.com/notes.html.

Department for Transport, 2013. *STATS19 Road Accident Dataset*, London, United Kingdom.

Department for Transport, 2011. *STATS20 - Instructions for the Completion of Road Accident Reports from Non-CRASH Sources*, England, United Kingdom.

Derriks, H.M. & Mak, P.M., 2007. Underreporting of Road Traffic Casualties. In *Transportation Research Board of the National Academies*. Rotterdam 3000 BA, Netherland, pp. 1–39. Available at: https://www.oecd.org/unitedstates/.

Diana, R. & Knuiman, M.W., 1994. A Comparison of Hospital and Police Road Injury Data. *Accident Analysis and Prevention*, 26(2), pp.215–222.

Ehnes, H. & Niu, S., 2012. *Improvement of National Reporting, Data Collection and Analysis of Occupational Accidents and Diseases*, CH-1211 Geneva 22, Switzerland.

Engineering Advice and Services, 2005. *Buffalo City Road Safety Plan: Accident Data Management Protocol*,

English, L.P., 1999. *Improving Data Warehouse and Business Information Quality: Methods for Reducing Costs and Increasing Profits*, John Wiley & Sons, Inc.

Ferrante, A.M., Rosman, D.L. & Knuiman, M.W., 1993. The Construction of A Road Injury Database. *Accident Analysis & Prevention*, 25(6), pp.659–665.

Fisher, C.W. & Kingma, B.R., 2001. Criticality of Data Quality as Exemplified in Two Disasters. *Journal of Management Information Systems*, 39, pp.109–116.

Gorell, R.S.J., 1997. Accident Data Collection and Analysis: The Use of MAAP in the Sub-Saharan Region of Africa. In *3rd African Road Safety Congress*.

Hox, J. & Boeije, H., 2005. Data Collection (Primary vs Secondary). *Encyclopedia of Social Measurement*, Volume 1, pp.593–599.

IRTAD, 1998. *IRTAD Special Report: Definitions and Data Availability: - Compilation and Evaluation of A-Level Roads and Hospitalised Victims in OECD Countries - Accident and Injury Definitions*, Bundesanstalt fur Strassenwesen (BASt), Referat U2, Bruderstrasse 53, D-51427 Bergisch Gladbach, Germany. Available at: http://www.oecd.org/dataoecd/24/1/2103123.pdf.

IRTAD, 2013. *Road Safety Annual Report 2013: An International Expert Network and Database on Road Safety Data*, International Transport Forum, 2 Rue Andre Pascal, 75775 Paris Cedex 16, France. Available at: www.internationaltransportforum.org.

Jacobs, G. (Dr) & Aeron-Thomas, A., 2000. *Africa Road Safety Review-Final Report*, United State of America.

Jarke, M. et al., 2001. *Fundamentals of Data Warehouses* 2nd ed., New York, NY, USA: Spring-Verlag New York, Inc. Secaucus, NJ.

Jeffrey, S. et al., 2009. An Evaluation of Police Reporting of Road Casualties. *Injury prevention: Journal of the International Society for Child and Adolescent Injury Prevention*, 15(1), pp.13–8. Available at: http://injuryprevention.bmj.com/cgi/content/long/15/1/13 [Accessed February 13, 2014].

Johnson, C., 1999. Improving Accident Reports. In Glasgow, Scotland: SAFECOMP'99.

Joseph, R., 2013. Who are South Africa's Worst Drivers? The Truth doesn't Lie in the Data. , (4th July), pp.1–8.

Jungu-Omara, I.. & Vanderschuren, M.J.W.., 2006. Ways of Reducing Accidents on South African Roads. In *Proceedings of the 25th Southern African Transport Conference (SATC 2006)*. Pretoria, South Africa: Document Transformation Technologies cc, pp. 454–464.

Karr, A.F. et al., 2005. *Data Quality: A Statistical Perspective*, 19 T.W. Alexander Drive, PO Box 14006, Research Triangle Park, NC 27709-4006. Available at: www.niss.org.

Kayi, C., 2007. Current Practices in Recording Road Traffic Accidents in Kenya: Implications of Data Quality on Analysis of Road Accidents Using GIS. In *11th World Conference on Transportation Research*. University of California, Berkeley: Kenyatta University, Department of Geography, Nairobi, Kenya, pp. 1–21.

Lane, D.M., 2015. Histograms. *Graphing Distributions*, pp.1–3. Available at: http://onlinestatbook.com/2/graphing_distributions/histograms.html [Accessed December 5, 2015].

Lane, D.M. et al., 2013. Introduction to Statistics. *Online Statistics Education: A Multimedia Course of Study*, pp.1–694. Available at: http://onlinestatbook.com/ [Accessed September 25, 2015].

Langford, J. & O'Hare, M., 2005. Road Safety Benefits of 50km/h Urban Speed Limits. , pp.1–9.

Lee, Y.W. et al., 2002. AIMQ: A Methodology for Information Quality Assessment. *Information and Management*, 40, pp.133–146.

Lotter, H.J.., 2000. Road Safety Performance Measurement in South Africa. In *South African Transport Conference: Action in Transport for the New Millennium*. South Africa:

Document Transformation Technologies cc, p. 13.

Mansfield, H. et al., 2008. *Analysis of the On the Spot (OTS) Road Accident Database*, Great Minster House, 76 Marsham Street, London, United Kingdom. Available at: www.dft.gov.uk.

Maree, L. & Daniels, D., 2007. *Stellenbosch Municipality: Integrated Development Plan: 2nd Generation-Original*, Stellenbosch, Cape Town. Available at: www.stellenbosch.gov.za.

Marshall, C. & Rossman, G., 1994. Data Collection Methods. In *Designing Qualitative Research*. pp. 78–107.

McDonald, G., Davie, G. & Langley, J., 2009. Validity of police-reported information on injury severity for those hospitalized from motor vehicle traffic crashes. *Traffic injury prevention*, 10(2), pp.184–190.

Mikulik, J. et al., 2007. *Road Accident Investigation Guidelines for Road Engineers*, Czech Republic.

Montgomery, D.C. & Runger, G.C., 2007. *Applied Statistics and Probability for Engineers* Permission., United States of America: John Wiley & Sons, Inc. Available at: http://www.wiley.com/go/permissions.

van Niekerk, L., 2012. Pocket Guide to South Africa. *Government Communication and Information System*.

Njord, J.R. et al., 2005. *Crash Records Systems: A Synthesis of Highway Practice*, Washington DC, United States of America: National Cooperative Highway Research Program. Available at: www.TRB.org.

O' Day, J., 1993. Accident Data Quality. *Synthesis of Highway Practice 192*, pp.1–41.

Redman, T.C. & Godfrey, B.A., 1996. *Data Quality for the Information Age*, Available at: www.artechhouse.com/international.

Rosman, D.L., 2001. The Western Australian Road Injury Database (1987 – 1996): Ten Years of Linked Police, Hospital and Death Records of Road Crashes and Injuries. *Accident Analysis & Prevention*, 33(December 1988), pp.81–88. Available at: www.elsevier.com/locate/aap.

Rosman, D.L. & Knuiman, M.W., 1994. A Comparion of Hospital and Police Road Injury Data. *Accident Analysis & Prevention*, 26(2), pp.215–222.

Roux, A. & Sinclair, M., 2003. Pedestrian Crossing Behaviour along A Primary Access Corridor in Stellenbosch. , (2003), pp.243–253.

RTMC, 2007. *South Africa AR Form Draft*, Pretoria, South Africa.

RTMC, 2012. *South African Road Safety Audit Manual*,

RTMC, 2014. *Standard Operating Procedure for Road Traffic Information Management*, Pretoria, South Africa.

Salmon, P., Regan, M. & Johnston, I., 2006. *Human Error and Road Transport: Phase Two - A Framework for An Error Tolerant Road Transport System*, SANRAL, 2012. *2012 Annual Report*, 48 Tambotie Avenue, Val de Grace, Pretoria.

SANRAL, 2008. *Annual Report 2008*, Pretoria, South Africa. Available at: www.nra.co.za.

Schmuller, J., 2009. *Statistical Analysis with Excel for Dummies* Second Edi., 111 River Street, Hoboken, NJ 07030-5774: Wiley Publishing, Inc. Available at: http://www.wiley.com/go/permissions.

Schoenbach, V.J., 2014. Data Analysis and Interpretation; Concepts and Techniques for Managing, Editing, Analyzing and Interpreting Data from Epidemiologic Studies. In *Epidemiologic Studies*. pp. 451–499. Available at: www.epidemiolog.net.

Shinar, D., Treat, J.R. & McDonald, S.T., 1983. The validity of police reported accident data. *Accident Analysis & Prevention*, 15(3), pp.175–191.

Sinclair, M., 2011. Road Injury Data in South Africa - An Assessment of the National Road Collision Database. In *30th Southern African Transport Conference (SATC 2011)*. Pretoria, South Africa: Document Transformation Technologies cc, pp. 262–274.

Sinclair, M. & Murdoch, M., 2012. Driver Compliance with Traffic Signals in Stellenbosch. In *31st Southern African Transport Conference (SATC 2011)*. Pretoria, South Africa: Document Transformation Technologies cc, pp. 512–525.

Sinclair, S., 2004. *Accident Reduction Guide*, Level 7, 260 Elizabeth Street, Surry Hills, NSW 2018: Road Safety, Licensing & Vehicle Management. Available at: www.rta.nsw.gov.au.

Sluis, J. van der, 2001. *Traffic Safety Information in South Africa: How to Improve the National Accident Register*, 2260 BB Leidschendam, The Netherlands.

Stats SA, 2009. *Road Traffic Accident Deaths in South Africa, 2001 – 2006: Evidence from Death Notification*, Pretoria, South Africa. Available at: www.statssa.gov.za.

Stats SA, 2014. *Statistical Release Mortality and Causes of Death in South Africa, 2011: Findings from Death Notification*, 170 Thabo Sehume Street, Pretoria 0002, Private Bag X44, Pretoria, South Africa. Available at: www.statssa.gov.za.

Stats SA, 2010. *Statistical Release Mortality and Causes of Death in South Africa, 2007: Findings from Death Notification*, Pretoria, South Africa. Available at: www.statssa.gov.za.

Stellenbosch Municipality, 2012. *Stellenbosch Municipality Annual Report*, Stellenbosch, Cape Town. Available at: www.stellenbosch.gov.za.

Stellenbosch Traffic Department, 2014. Directorate: Community and Protection Services. *Department Organogram*. Available at: www.stellenbosch.gov.za [Accessed May 20, 2003].

Thieman, S., 1999. Accident Data Use and Geographic Information System (GIS). In *Sixth National Conference on Transportation Planning for Small and Medium-Sized Communities*. 2101 O' Neil Avenue, Cheyenne, WY 82001, pp. 1–7.

Traffic Accident Statistics, 2002. *City of Cape Town: Road Traffic Accident Statistics: 2001*, City of Cape Town, Vlaeberg 8018, South Africa.

Traffic Accident Statistics, 2005. *Traffic Accident Report City of Cape Town 2005*, City of Cape Town, Vlaeberg 8018, South Africa.

Transportation Association of Canada, 2006. *Performance Measures for Road Networks: A Survey of Canadian Use*, Canada.

Twala, B., 2013. Extracting Grey Relational Systems from Incomplete Road Traffic Accidents Data: The Case of Gauteng Province in South Africa. *Expert Systems*, 0(0), p.12. Available at: http://doi.wiley.com/10.1111/exsy.12035 [Accessed January 30, 2014].

Vanderschuren, M., 2006. *Intelligent Transport Systems for South Africa*.

Vanderschuren, M. & Jobanputra, R., 2011. Hazardous Road Safety Location Analysis: A Case Study of the Western Cape. In *Proceedings of the 30th Southern African Transport Conference (SATC 2011)*. Pretoria, South Africa: Document Transformation Technologies cc, pp. 233–242.

Vanderschuren, M.J.W. & Irvine, C.., 2002. Is South Africa Ready for An International Safety Comparison? In *21st Annual South African Transport Conference: Towards Building Capacity and Accelerating Delivery*. South Africa: Document Transformation Technologies cc, p. 10.

Vogel, L. & Bester, C.J., 2005. A Relationship between Accident Types and Causes. In *Proceedings of the 24th Southern African Transport Conference (SATC 2005)*. Pretoria, South Africa: Document Transformation Technologies cc, pp. 233–241.

Vogel, L. & Bester, C.J., 2004. Accident Data Quality in the Western Cape., pp.1–8.

Volunteer Estuary Monitoring, 2006. Voluntary Estuary Monitoring Manual: Data Management, Interpretation, and Presentation. In *A Methods Manual*. United States: Environmental Protection Agency, pp. 1–23. Available at: http://www.epa.gov/owow/estuaries/monitor/.

Wang, R.Y., 2004. Data Quality: Theory in Practice. In *EPA 23rd Annual National Conference on Managing Environmantal Quality Systems*. pp. 1–4.

Wang, R.Y., Reddy, M.P. & Kon, H.B., 1995. Toward Quality Data: An Attribute-Based Approach. , 13, pp.349–372.

Wang, R.Y., Storey, V.C. & Firth, C.P., 1995. A Framework for Analysis of Data Quality Research. *IEEE Transaction on Knowledge and Data Engineering*, 7(4), pp.624–640.

Wang, R.Y. & Strong, D.M., 1996. Beyond Accuracy: What Data Quality Means to Data Consumers. *Journal of Management Information Systems*, 12(4), pp.5–34.

Wang, R.Y. & Strong, D.M., 2013. Beyond Accuracy: What Data Quality Means to Data Consumers. , 12(4), pp.5–33. Available at: http://www.jstor.org/stable/40398176.

Wang, R.Y. & Wand, Y., 1996. Anchoring Data Quality Dimensions in Ontological Foundations. *Communications of the ACM*, 39(11), pp.86–95.

WHO, 2011. *2nd Global Status Report*, 20 Avenue Appia Geneva 27 Switzerland. Available at: http://www.who.int/violence_injury_prevention/road_safety_status/2009/en/index.html.

WHO, 2010. *Data systems: A Road Safety Manual for Decision-Makers and Practitioners*, 10 Avenue Appia, 1211 Geneva 27, Switzerland.

WHO, 2013. *Global Status Report on Road Safety*, 20 Avenue Appia, Geneva 27, Switzerland.

WHO, 2009. *Global Status Report on Road Safety: Time for Action*, 20 Avenue Appia Geneva Switzerland. Available at: www.who.int/violence_injury_prevention.

WHO African Region, 2013. *Road Safety in the WHO African Region: The Facts 2013*, Geneva, Switzerland. Available at: http://www.who.int/entity/violence_injury_prevention/publications/road_traffic/UN_GA_resolution-54-255-en.pdf.

World Bank & World Health Organisation, 2013. *Supporting a decade of action*, World Health Organsation, 20 Avenue Appia, 1211 Geneva 27, Switzerland. Available at: http://www.who.int/about/licensing/copyright_form/en/index.html.

World Health Organisation & The World Bank, 2004. *The Global Impact: World Report on Road Traffic Injury Prevention*,

Yamamoto, T., Hashiji, J. & Shankar, V.N., 2008. Underreporting in traffic accident data, bias in parameters and the structure of injury severity models. *Accident; analysis and prevention*, 40(4), pp.1320–9. Available at: http://www.sciencedirect.com/science/article/pii/S0001457508000237 [Accessed February 5, 2014].

Zegeer, C. V, 1987. Uses and Limitations of Police Accident Data. *Transportation*.

Appendix A: Additional literatures on road accident data collection

A.1 Early era's technique for road accident data collection

The early commencement of road accident data collection was dated back to 1912 in United States (O' Day 1993). The data collection begun at the local level, where accumulations of data were considered to be pin-maps in the office of the traffic engineers or police chiefs, with the use of coloured pin to identify different accident events and the exact accident locations (O' Day 1993).

Moreover, the exact spots on the map, where these coloured pins were positioned, were used to initiate the decision-making processes on where engineering improvements are most requisite. Besides, it is also used to promote law enforcement campaigns towards a better conduct of the road users while travelling on the roads (O' Day 1993). During this era, data management systems were not that effective, compared to this present time where many approaches of resolving issues concerning accident events have been developed and employed to tackle big data analysis.

A.2 Benefits of the early era's technique for road accident data collection

The technique used in the early era has a reward of easy interpretation. The data were considered accurate, complete and detailed, since the officials are fully committed to their tasks, with their individual involvement in the development and implementation of the old technique (O' Day 1993).

Although, after some years, an advanced technique for managing road accident data was developed, which involved the early gathering of unprocessed and processed data at the three authorised levels, with the aim of enumerating failures in the traffic system. The technique required the use of card sorters, which was the kind of computer system used to handle data in this era (O' Day 1993). Specifically, during this era, data analysis was established primarily on the number of persons injured or killed in a particular group of the population, at a specific accident location (O' Day 1993).

The data collection technique implemented, was majorly designed, developed and improved to obtain a quality data system which could offer better result on accuracy, consistency and interpretation of the result obtained. The magnitude of the data assembled and processed in the early era is much smaller, compared to the magnitude of data assembled and processed at this present time due to the everyday increase in the population of the road users. This permits considerate implementation of the quality components, as the foundation for reliable processing of data quality.

Understandably, reliability of data processing system could be achieved through an optimised design process. The process enhances the actions engaged in the collection of the relevant data during the completion of the ARF, and also enhances the other vital processes involved before entering the data into the database system.

A.3 Limitations in the early era's technique for road accident data collection

The efficiency of the old technique was attributed to the commitment demonstrated by the traffic officials towards their tasks. Despite the ability of the old technique to yield accurate results, however, some limitations were encountered that necessitate the reason for advanced techniques. Clearly, the old technique is limited to conditions like:

- **Increase in population of a country or a particular geographical location;** which demands for more infrastructures (O' Day 1993).
- **Increase in the number of vehicles;** both registered and unregistered vehicles (WHO 2010; O' Day 1993).
- **Rapid transformation of the rural to urban;** which could be attributed to growth in internally generated revenues and rapid development of infrastructures leading to the high increase in the number of immigrants.
- **Lack of advanced data collection facilities to cover large areas;** due to the ineffective economic growth of a country (WHO 2010).

The aforementioned limitations postulate new techniques to data collection in the recent years where technology has transformed many obsolete ideas to better ones. This led to the development of advanced computer to resolve the issue of big data processing for a reliable result.

A.4 Accident [crash] report form

Accident Report form [crash report form] is designed for the collection of relevant data required in obtaining appropriate information about the cause of an accident involving single vehicle, multiple vehicles, motorcycles, bicycles, vehicle and fixed object, and information about the fatality involving road users such as the driver, passengers, pedestrians and many other relevant information.

According to a report published by the Department for Transport in United Kingdom, in charge of Stats19 Road Accident Dataset in the region, an Accident Report form could be defined as "a form designed to collect set of collision [accident] data or a wide variety of information about the accident such as time, date, location, road conditions together with the vehicles and casualties involved and contributory factors to the accident as interpreted by the police" (Department for Transport 2011; Department for Transport 2013). This particular data collection process is performed by a police officer or a traffic officer right at the scene of the accident, whenever any type of accident injury is reported either by the victim of the accident or by the eyewitness present at the scene of the accident (Department for Transport 2013).

The development of the Accident Report form paves the way for accurate reporting and recording of accident information. It serves as the manual process of keeping accident records before the information could be processed into a complete data, through the data compilation processes authorised by the RTMC, as a *lead authorised agency* managing the road accident data in South Africa.

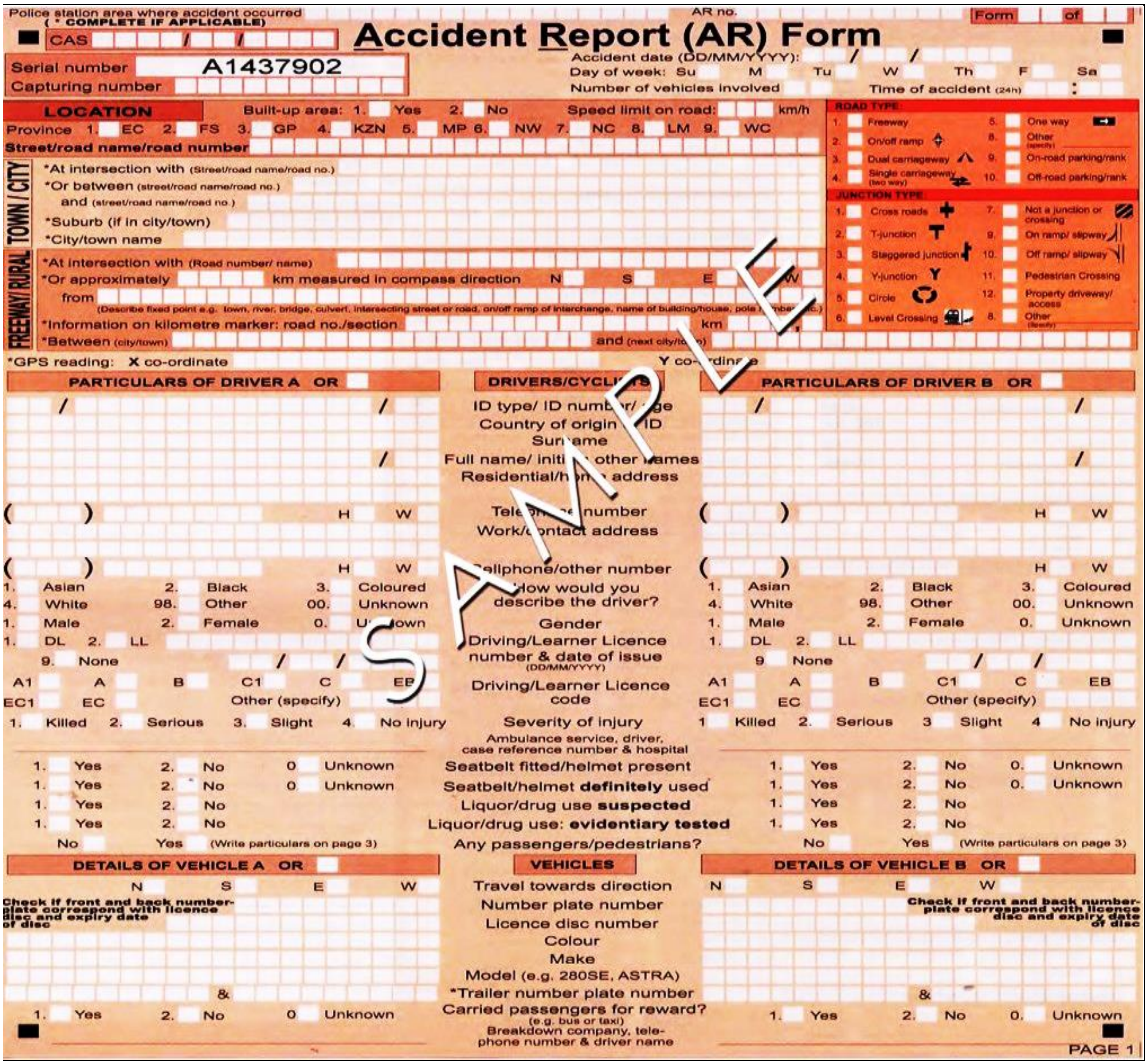

Figure 58: A sample picture of the ARF used in South Africa

In essence, information gathered by using the ARF required reliable recordkeeping process for reference purpose, as it may be requested by the police, insurers, lawyers and many other data users for other important motives. The form serves as a prerequisite way of keeping accident record intact. A copy of the ARF used in South Africa is depicted in the Figure 58[32].

A.5 Description and purpose of ARF

The design of the ARF used in South Africa, is based on technical questions required to formulate complete answers towards the problems thwarting safety of the road users in the country. More so, these questions provide a better technique of developing a reliable data system for electronic record keeping and analysis. In South Africa context, the procedures implemented in the design of the ARF and data collection systems, are classified as the overall responsibilities of the RTMC as authorised by the Department of Transport [DoT]. Moreover, ARF is designed in line with the international standard, though the data form still requires basic improvements to accommodate more information, in order to strengthen the need for sufficient data and reliable decision-making process.

The type of Accident Report forms used in the developed countries are considered robust and opulent with more essential features, that could accommodate additional information required, since further research discoveries are unfolded in the area of RTA on daily basis (O' Day 1993; Department for Transport 2011; Ehnes & Niu 2012; Department for Transport 2013). The Accident Report forms are designed simple, comprehensive, and informative (O' Day 1993; Johnson 1999; WHO 2010) compared to the categories of the Accident Report form used in most developing countries. Consequently, these characteristics constituted more accurate, adequate and reliable road accident data gathered in the developed countries compared to the road accident data gathered in the developing countries (Ferrante et al. 1993; Rosman & Knuiman 1994; Rosman 2001; WHO African Region 2013; WHO 2013).

The data system implemented for data recordkeeping in many developed countries, is considered standard in terms of its degree of technological improvement, reliable information and data dissemination technique, which thereby complement the efforts of the developed countries towards a consistent reduction plan for accident fatalities (Rosman & Knuiman 1994). Fundamentally, one could say that dataset is a replica of ARF, but preferably in an advance way (Baguley 2001). Thus, the dataset is considered as the electronic storage system designed to perform operations such as data assemblage, data storage, and data analysis of information [structured and semi structured data], which are manually gathered through the use of ARF.

[32] The picture shown in the Figure 58 is adapted from RTMC (2007).

In short, the description of information or instruction required in the Accident Report form should be simple, brief and clear to perform functions enumerated below:

- Perfect definitions of each field or factor where necessary.
- Convenient use of the data form by the form users [accident reporting officers].
- Facilitate the initial process of the road accident data collected.

For more clarifications, the purpose of developing ARF could be attributed to four significant reasons highlighted below:

- **Identification of the parties involved;** to provide details about the accident victims, the number of persons involved or casualties' details, vehicle particulars and other relevant information. Actually, this information is very relevant to the insurance companies to facilitate quick processing of accident funds.
- **Identification of the causal factors;** to offer technical understanding into the identification of the possible causes, and the contributory factors that may be responsible for such accident.
- **Planning for intervention to resolve safety issues;** to facilitate a resourceful approach to tackle the problems frustrating the road safety systems.
- **Advocating for financial support;** to ease the cost of implementing sustainable improvements through new ideas that may be developed based on the findings obtained from the analysis performed.

A.6 Advantages of road accident data collection

The advantages of collecting road accident data provide further understanding into fundamental motives of developing SOP, to guide the authorities involved in managing the road traffic issues through their assigned responsibilities. However, the benefits elucidate the significance of road accident data to the road traffic authorities, government and public in general. The advantages in collecting the road accident data are outlined below as:

- **Assists the government in decision making process regarding the national budgeting structure** (WHO 2010); with the aim of having the knowledge of distributing wealth and resources across the nation.
- **Supports the detailed insight about a nation's population**; in order to have a view of what figure to be expecting in the subsequent years.
- **Encourages the development of a country** (WHO 2010); in terms of infrastructures, basic amenities and intensive care programmes.

- **Serves as a tip to the road safety engineers** (O' Day 1993); in determining suitable measures to curb the recurrence of accident on the road.
- **Serves as a determining factor in developmental matters**; to link the urban areas with the rural areas, in terms of economic growth.
- **Assists in determining where development is paramount in the area of safety on the road** (O' Day 1993); in order to deter incessant occurrence of accidents.
- **Supports the government and agencies in charge of developing road safety rules** (Baguley 2001; Njord et al. 2005); through the feasibility of analysis report received on the areas that contribute mostly to accident occurrence.
- **Aids decision making processes** (O' Day 1993; Njord et al. 2005); assisting the police and the insurance companies to understand the base of the legal action filed against accident victims.

A.7 Disadvantages of road accident data collection

The disadvantages in the road accident data collection illustrate the probable problems that could arise from poor data collection process. The drawbacks are associated with the level of the SOPs implemented, to ensure a smooth and reliable data procurement process. The disadvantages in collecting the road accident data are outlined below as:

- **Feckless handling of the road accident data** (Batini et al. 2009); along the road accident data processing system.

- **Poor analytical methods** (Mansfield et al. 2008); in determining the best measures suitable for accident reduction.

- **Inept application of the capturing devices, or measuring devices in strategic areas** (Njord et al. 2005); identified as the key areas contributing to accident recurrence.

- **Inadequate training of the personnel in charge of data handling**; cost the management poor reporting and recording of accident matters.

- **Nonchalant attitudes of the accident victims and witnesses** (Alsop & Langley 2001); in cooperating with the traffic police during investigation process, thereby leads to poor quality and insufficient data.

- **Inability of the form users to understand the application of the ARF** (Asia Injury Prevention Foundation et al. 2010b); which leads to loss of relevant data.

- **Failure of the management in adopting an advanced data collation system**; which affects the ability of the management to project a method of acquiring relevant data.

- **Lack of comprehensive ARF** (Ehnes & Niu 2012); with poor sectional arrangement and inadequate information, which generates poor data collation.

A.8 Purpose of standard operating procedures [SOPs]

The SOPs developed aid some operational motives that are very paramount to the RTMC. These procedures assist in coordinating a consistent way of producing quality data system. The primary purposes of developing the SOPs are provided below:

- Serving as set of instructions required in the execution of several tasks involved in the production of RTI (RTMC 2014).
- Strengthening the procedures contributing to the production of RTI, such as accident notification or identification, collation and analysis, capturing and quality assurance of road accident data (RTMC 2014).
- Offering directives on how to manage the services rendered by each unit involved in the reporting, recording, distribution and security management of RTI (RTMC 2014).
- Monitoring the steps involved in the transacting of information, in order to stabilise the dissemination of the right information within the sequence of the units involved.

A.9 Road traffic accident [RTA] stakeholders in South Africa

In view of the RTI handling in South Africa, several stakeholders are involved in different units or divisions with regard to their allotted responsibilities. Some stakeholders are authorised as road traffic agents by the transportation ministry of South Africa. Primarily, some serve as assistants in motivating the process of acquiring a reliable RTI; while others aid the enforcement of the traffic laws in order to daunt the road users from violating the traffic rules and regulations.

In addition, the impacts of these departments in generating quality data, and also in reducing the inclined number of RTA recorded daily in South Africa, require continual assessment to ascertain the state of road safety in the country. Besides, the support of the research institutes is valuable in this context, in order to cultivate new safety systems to stimulate the existing ones. Even though, the transportation departments and agencies are working hard, both at the local and provincial levels to minimise the occurrence of RTA. Despite the determination to minimise the prevalent occurrence of RTA in South Africa, however, the daily reportage of RTA cases is still considered worrisome due to factors such as:

- sloppiness in the relational operations of the authorities involved (Adams 2001),
- mishandling of information processing (RTMC 2012),
- unyielding attitude of some road users towards traffic rules (Jungu-Omara & Vanderschuren 2006), and

- regulations and soft traffic laws (Vanderschuren & Irvine 2002; RTMC 2012).

Nonetheless, in sustaining the ability of the traffic management in acquiring sufficient knowledge on technical know-how, -why, -where, and -when an incident recurs, thus, a reliable and sustainable data set structure need to be established, to boost an in-depth understanding into any accident matters. For that reason, all the units or departments involved in the processing of RTI deserve to be carried along; in order to achieve a better information processing system towards a proficient and sustainable quality data. Naturally, this will strengthen the appropriate processing of the unprocessed road accident data collected from a reported incident. Furthermore, the responsibilities of these stakeholders are presented in the subsequent paragraphs.

- **National Department of Transport [NDoT]** –this department is in charge of all general regulations of transportation system, including any other related matters in South Africa such as public transport, rail transportation, civil aviation, shipping system, freight and motor vehicles. This department is in control of disbursing funds to stimulate the activities performed by the other lower cadres, in reducing the problems of transportation system. In addition, the department makes decision over the road traffic affairs, and also responsible for disseminating the publication of the reports compiled by the provincial traffic management (RTMC 2014).
- **Road Traffic Management Corporation [RTMC]** –concise descriptions of this department have earlier been discussed in the prior sections. This particular department is a principal agency in the general matters concerning road safety, by coordinating and motivating the activities executed in the area of strategic planning, safety regulations, advocating for financial support, and stimulating continual improvement of the road safety system. Moreover, the department also assists in facilitating the enforcement of road traffic laws towards the reduction of RTA, and thereby develops safety programs to protect the lives and properties of the road users. In addition, the department stabilises the affairs of the RTI by solidifying the communication interface between all the units, in order to avoid poor quality processing of RTI (RTMC 2014; Adams 2001). Above all, the department is recognised for its involvement in overseeing the data assessment which involves operations such as data pre-processing, data entering and data processing, and transfer of completed copy of analysis reports to the NDoT.
- **South African Police Service [SAPS]** –this department assists in securing and stabilising the accident scene. In some cases, the department assists in coordinating the traffic flow, completes ARF at the scene of the accident, and also embarks on an accident investigation for alleged criminal offences.

- **Road Traffic Infringements Agency [RTIA]** –this agency develops reports on the road traffic violations, or related offences and submit complete reports to the superior agency for further evaluation.
- **Road Accident Fund [RAF]** –this agency validates the necessary information regarding accident claims; by liaising with the police department and the insurance company in order to acquire other necessary information regarding the accident victim.
- **South African National Road Safety Agency Limited [SANRAL]** –this department supplies relevant information regarding the road traffic volumes and incidents.
- **Toll Concessionaires** –this division assists in supplying information regarding RTI.
- **National Traffic Police, Provincial, Metropolitan and Local Traffic Authorities** –these departments collaborate with the other departments in locating any incident areas at the local level, and simultaneously assists in regulating traffic volumes, controlling and securing an accident scene. Notwithstanding, the department also helps in consolidating the participation of other local traffic agencies, by mobilising emergency services and SAPS in protecting lives and properties. And the department is also involved in the recording of RTI, and embarks on investigation of accidents by strengthening the effort of reducing the cause of RTA.
- **National Department of Health [Hospitals, Pathologists, Paramedics etc.] [NDoH]** –this department helps in validating any issues regarding road traffic injuries and casualties due to RTA.
- **Department of Home Affairs [DHA]** –this department processes the death certificate issued by the Department of Health, regarding casualties due to RTA.
- **Insurance Companies/Associations** –this department participates in information sharing, and communicates with the relevant department concerning accident benefit claims.
- **Towing Companies/Associations** –this department also assists in sharing information on road traffic matters, and the correct application of the information to avoid unnecessary problems. Besides, the association assists in the towing or conveying of vehicles involved in accidents, depending on the level of damages done on the vehicles.
- **Media Houses/Organisation** –this organisation assists in information dissemination and communication across local levels through provincial levels to national level.

A.10 Synopsis of RTI production in South Africa

A concise description of the processes involved in the production of the RTI in South Africa is discussed in the subsequent sections. These processes are considered as the standard procedures established for the acquisition of quality data (RTMC 2014), as approved by the RTMC and other authorised agencies.

A.10.1 Attending to the accident scene and data collection

The flowchart, displayed in the Figure 59[33], demonstrates the sequential procedures followed in acquiring the RTI, starting from the point of data collection to the point reports compilation. Furthermore, the process is initiated through an authorised official, who controls the perturbed condition at the scene of the accident, and simultaneously collects the data based on the findings acquired.

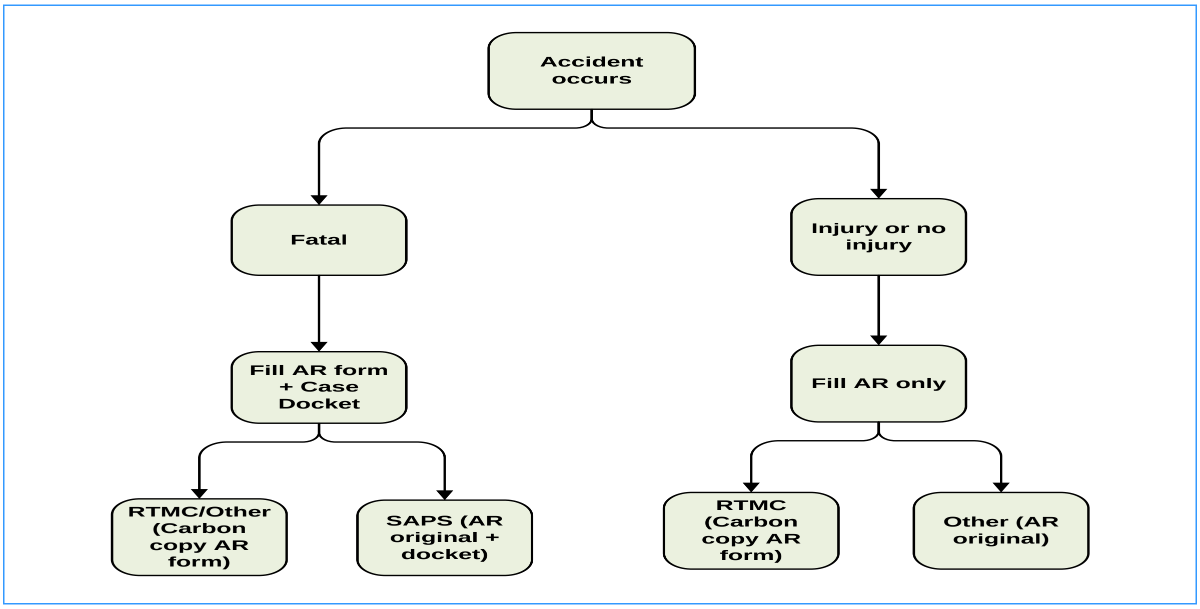

Figure 59: Process flowchart demonstrating the steps taken in acquiring RTI in South Africa

According to the flow of the process, the acquisition of data is distinctly defined by the categories of severity of injury, such as fatal and nonfatal injuries. The nonfatal injuries are categorised as accident severity involving an injury and/or no injury cases, whereby no persons involved in the accident were registered dead within six days; whereas, in the case of fatal injury, the accident severity involved the death of a person at instant of accident occurrence or within six days from the period of the accident occurrence, either as a driver or as a passenger[s] of a particular vehicle that involved in an accident, even as a pedestrian who involved in an accident, practicably a case docket maybe registered.

A.10.2 Road accident data validation and capturing in South Africa

[33] The diagrams illustrated in both Figure 59 and Figure 60 are adapted from RTMC (2014).

The flowchart demonstrated in the Figure 60, establishes a complete procedure followed in validating and capturing the RTI in accordance with the categories of accident severity. The process flow indicates the transfer of information from the local level to the provincial level and/or RTMC, where data gathered were captured in the Traffic Data Management Software [TRAFMAN] and other data capturing system like Integrated Provincial Accident System [IPAS].

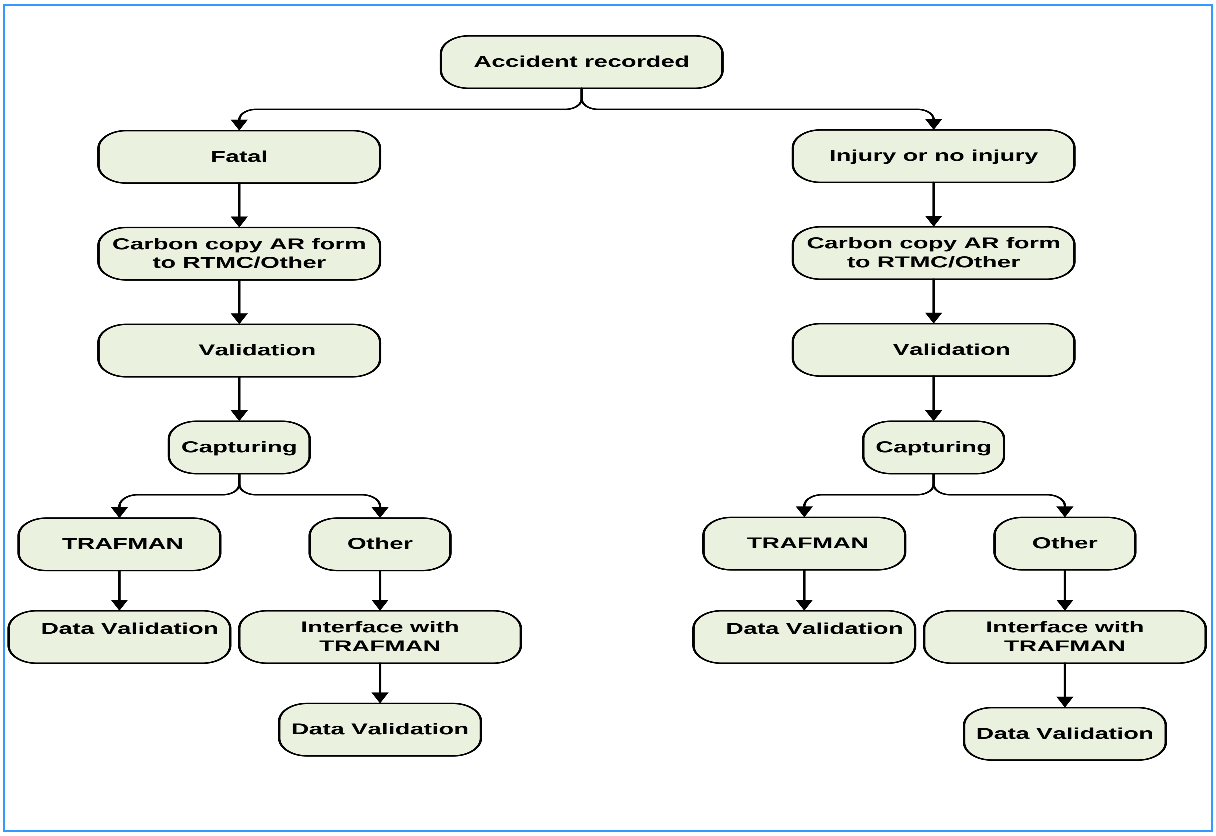

Figure 60: Process flowchart demonstrating the steps taken in processing RTI in South Africa

However, this particular system is similar in all nine provinces in South Africa, as the process implemented or designed for the capturing, validating and transferring of all the data involving RTAs. As indicated in the diagram above, in any case where an authority prefers a different capturing system, definitely an interface must be established in line with the TRAFMAN system.

A simple illustration of the process diagram shows that information gathered from the accident scene is completed in the ARF and a carbon copy is made and transferred to the RTMC or any other traffic authorities. Validation and capturing processes are concurrently executed at each stage involved, to ascertain the degree of quality of the data gathered for a particular purpose.

A.10.3 Road accident data handling procedure

This section discusses the services rendered by the officers handling the road accident data in South Africa in accordance with the findings acquired in STD. In subsection 2.2.5, a sameness procedure was discussed regarding the units that handle the processing of road accident data, but in a general approach. The Table 38[34] presents a detailed processing of RTI at the local, provincial and national levels.

Table 38: Operations perform by the local level, provincial level and national level regarding processing of RTI

Local Level [Police Departments and Traffic Authorities]	
Daily [activities performed every day]	**Weekly [activities performed every Friday as specified]**
• Keep records of submitted ARFs intact. • Arrange the collated ARFs for immediate transfer. • Proper transfer of completed ARFs to the cluster.	• Collection of the completed ARFs by the data collectors at the various data sources. • Perform first stage validation by crosschecking the completeness and accuracy of the complete ARFs. • Confirm the approval of the completed ARFs to be collected.
Provincial Level [Consist of Capturing Centres]	
Weekly [activities performed every week]	**Monthly [activities performed every month]**
• Approve the receipt of the delivery form of the completed ARFs from the collector. • Perform a second stage validation simultaneously, by checking for accuracy, correctness, timeliness etc.	• Prepare the reports on the content of information captured. • Transferring captured data to the relevant units for further evaluation.
National Level [Road Traffic Reports Collation Centre]	
Monthly [activities performed every month]	
• Assembling of data imported from all the Provinces. • Perform third stage validation instantaneously, by checking for accuracy, reliability, relevancy, accessibility, integrity, comparability, coherence etc. • Gather reports prepared and deliver them to the Quality Assurance Committee. • Initiate the dissemination of reports to the public *[data users].* • Encourage continuous procedural system for information procurement.	

The processing of information is carried out on daily, weekly and monthly basis at the local level, which comprises the police departments and traffic authorities through the provincial level, which comprises the data pre-processing, data entry and data processing centres and to the national level, which compiled the reports and prepare a complete annual report covering all the departments for public assessment (RTMC 2014).

A.10.4 Road accident data capturing officer's responsibilities

The main task performed here is basically to ensure an accurate collation process of the completed ARFs. The Data Capturing Officer [DCO] captured the data represented in the

[34] The contents in the Table 38, Table 39, Table 40, Table 41 and Table 42 are adapted from RTMC (2014).

completed ARFs, sometimes by manually entering the information or making a scan copy. Although, the first step taken by the DCO to execute this task, is by signing off the completed ARFs, and simultaneously collated and captured the data into the local dataset.

In addition, the officer further executes some other tasks, by suggesting any necessary corrections required with regard to the data gathered from several sources. The DCO prepares a weekly and monthly concise report, and submit a complete copy to the supervisor for approval, and later file up the ARFs accordingly (RTMC 2014). A detailed summary of the responsibilities of the DCO is provided in the table below.

Table 39: Responsibilities of the Data Capturing Office in the processing of the RTI

Responsibilities of the Data Capturing Officer
Daily and Weekly Tasks
• Approval of the ARFs received. • Capturing data presented on ARFs manually or by scanning. • Ensuring the rectification of errors discovered with several sources. • Ensuring the safekeeping of the collective folder of the ARFs captured.
Weekly and Monthly Tasks
• Produces a concise report of data captured every week and month. • Handing over a copy of the report to the supervisor for approval. • Registering compiled copies of completed ARFs captured.

A.10.5 Road accident data capturing supervisor's responsibilities

Data Capturing Supervisor [DCS] ensures a standard process for data quality assurance, and instigates the effective utilisation of local use of information. The supervisor is primarily the provider of unused ARF for appropriate collection of road accident data, and as well, instil a strict follow up of the SOP by subjecting the junior officers to training for appropriate processing of the road accident data accordingly (RTMC 2014).

In addition, the supervisor performs other tasks such as distribution of stationery, stocks and equipment to other personnel who request for them through proper application procedures. The supervisor also performs compilation and assembling of the accident reports received, by ensuring an appropriate definition to data factors and indicators. Moreover, the supervisor gives a directive on a suitable filing arrangement for the approved accident reports [completed ARFs].

Other tasks carried out by the supervisor is fundamentally based on the need to set up a standard training on basic evaluation processes, suchlike data features and indicators documentation, data quality assessment and data utilisation process for junior officers assigned to data collection and collation, as considered more substantial in the possibility of achieving an improved quality data.

Table 40: Responsibilities of the Data Capturing Supervisor in the processing of the RTI

Responsibilities of the Data Capturing Supervisor
Daily, Weekly and Monthly Tasks
• Ensuring a complete process for acquisition of RTI. • Provision of adequate resources to support steadiness of the SOP. • Distribution of writing materials, blank ARFs etc. • Handling the important telephone calls and fax messages. • Applying appropriate definitions for the data features and indicators. • Follow up any discrepancies with various sources. • Distributing the circular by email, hardware and software, junior staff to their appropriate positions, and ascertaining the uninterrupted internet connections.
Weekly and Monthly Tasks
• Organising relevant and proper training section for junior staff regarding the data elements, data quality assessment and data collection techniques. • Instructing the new staff on the procedure established for the handling of the RTI. • Handling the periodic supervision of the activities performed under his or her jurisdiction. • Performing appropriate distribution of stocks and equipment according to applications submitted for procurement request.

Accordingly, the superior ascertains the degree of standard orientation of the new staff on daily, weekly and monthly activities scheduled for execution, towards a reliable process for minimal errors (RTMC 2014). A concise description of the responsibilities discharged by the DCS is provided in the Table 40 in accordance with the findings obtained from the STD.

A.10.6 Quality assurance

In this section, compiling of accident reports are subject to sequential procedures required to ensure the steady improvement of the RTI, as exactly enumerated in the Table 41. These procedures are incorporated into a road traffic quality assurance framework, designed and supervised periodically by the appropriate quality assurance committee (RTMC 2014).

Nevertheless, the procedures constitute the quality assessment components like relevance, accuracy, completeness, timeliness, accessibility, interpretability, and many others like comparability and coherence, methodological soundness, and integrity (RTMC 2014). The above-mentioned quality assessment components are applied to ascertain constant improvement of the quality of the road accident data gathered through different sources. For a concise explanation, see Figure 7.

Table 41: Actions performed by the Data Quality Assurance Unit in ensuring smooth data processing

Actions performed by Data Quality Assurance Unit
• Check out the consistency level of the completed ARFs captured. • Comparison of the hardcopy of completed ARFs with the softcopy in the electronic captured data • Executing tasks according to guidelines established by the RTMC towards a prudent way of keeping road traffic records intact. • Set up a reliable committee that deliberates on the practicality and reliability of the data or information collated before being transferred into the next section for further evaluation process. • Performing the primary validation of the information gathered every month by ensuring absence of duplicate copies, misplacements etc. • Clarifying the identification of the irregularities discovered by informing the previous unit while spot checking the completed ARFs. • Follow up with the corrections anticipated to be rectified by the previous unit before such forms could be transferred to the next stage for further evaluation. • Ensuring the proper application of the quality assessment components along the process line developed for the road accident data. • Distributing authorised road traffic reports across the other units within the province, before such reports are sanctioned for public dissemination.

A.10.7 Reporting and dissemination

The activities regarding reporting and dissemination of RTI are managed by the RTMC, as an authorised custodian of the road accident data in South Africa. The process involved a periodical schedule according to the report types and timeframes for the publications. The periodic schedule is provided in the Table 42, which depicts the periodic preparation of the road traffic reports before it could be disseminated to the public.

Table 42: Periodic publishing of road traffic completed reports

Report Type	Time fame
Calendar Report [January to December of each year]	**February [Subsequent year]**
Quarterly report	August, November, February, and May
Festive season report	February
Easter report	June

A.10.8 Access to RTI

The procedure for permission to access RTI is carried out through a writing application, which is subject to the provisions of the Promotion of Access to Information Act [PAIA] 1 of 2000, and any relevant information legislative prescripts (RTMC 2014). In addition, the permission to information, which requires personal information, will be administered with regard to the prescripts of Promotion of Personal Information [POPI] Act 4 of 2013.

A.10.9 Records management

The reports produced annually are properly archived as instructed through the policies and procedures established to guide the application of such information. The process is guided by the National Archives and Records Service of South Africa Act, No 43 of 1996 [as amended] (RTMC 2014). This permitted other data users to have the opportunity to access past records and make a genuine reference to the earlier records or reports disseminated by the RTMC.

A.10.10 Information security

This particular section provides details on the measures taken by the RTMC, to ensure the security of the RTI before and after the reports have been published for public assessment. In essence, according to RTMC, the authenticity of the information is based on integrity and confidentiality, by taking necessary measures to prevent abusive use of RTI in the publicity, loss of information and damage of information; due to an unauthorised destruction of information which might be initiated by unlawful access to the information.

Appendix B: Additional analyses and findings on Accident related factors

B.1 Analysis of the Severity of injury and Summary of persons involved

This section discusses the two most related data fields in the Accident related factors. The discussion is centred on the correlation between the data processed in the *‘Severity of injury’* and *‘Summary of persons involved’*. These two categorised data fields have a different purpose in the data collection procedure; though, both are related in interpretation, but one encompasses details of all the road users involved in road accidents.

Actually, the *‘Severity of injury’* consist of four different categorical variables, characterised in collecting the details concerning the drivers and cyclists only, which excludes the pedestrians and passengers (Department for Transport 2011; Mansfield et al. 2008; RTMC 2007). On the contrary, the *‘Summary of persons involved’* is characterised with similar variables as that of Severity of injury to consolidate the injury data pertaining to all the road users, including the estimated injury data represented in the *'Severity of injury'* for the drivers and cyclists involved in road accidents (Department for Transport 2011; Mansfield et al. 2008; RTMC 2007).

The four variables are grouped into fatal injury and nonfatal injury. The *'Killed'* injury is categorised directly under fatal injury, while the *'Serious*' injury, *'Slight'* injury and *'No injury'* can be categorised directly under nonfatal injury (RTMC 2014). Moreover, under *'Serious injury*', there is a tendency that person involved may not survive the accident, but it wholly depends on the severity of the accident which may require hospitalisation of the victim. In the case of the '*Slight*' injury, the severity of the accident is not as severe as the one experienced under the '*Killed*' and *'Serious'* injury. Also, in the *'No injury',* the severity of the accident is considered very minimal compared with the other three variables. In this case, the *'Severity of injury'* provides preliminary estimates of the data relating to road accident injuries that involved only the driver and cyclist, excluded the pedestrian and passenger as displayed in Table 43 below. In addition, the Table 44 displays the data acquired in the road accident injuries for all the road users are categorised under the *'Summary of persons involved'*, such road users like the driver, cyclist, passenger and pedestrian. The interpretation of the injury data will exemplify the number of road users involved in different type of road accident injuries.

Table 43: Severity of injury estimates in road accident occurrence

Total estimates of Severity of injury in RTA occurrence in 2012				
Months	**Severity categories**			
	Killed	**Serious**	**Slight**	**No Injury**
Jan	0	5	22	193
Feb	0	8	27	317
Mar	0	6	44	324
Apr	1	3	28	231
May	0	4	22	296
Jun	1	2	21	237
Jul	0	1	18	211
Aug	0	4	22	317
Sep	0	7	26	254
Oct	0	0	4	52
Nov	0	5	22	273
Dec	1	5	17	170
Total scores	**3**	**50**	**273**	**2875**

The table above presents the set of injury data analysed in the Severity of injury concerning the drivers and cyclists involved in road accident injuries. From the table, the array of scores displayed in the last column, represents a large number of unharmed victims over a year, which contains only the drivers and cyclists. According to the scores in the table, only 3 road accident injury cases led to the death of 3 people [drivers/cyclists], with the lowest percentage estimate of 0.1% in 2012.

On the contrary, an estimated percentage of 2.0% road accident injuries resulted to *'Serious'* injury in the same year. This result reflects the impact of the road accident injuries on the accident victims, that is, the resultant of the accident on the drivers and cyclists involved. Observably, this analysis specifies that highest score of *'Serious'* injury was recorded in February, followed by September and March respectively as shown in the table above. Further observations reveal that, a rise in the amount of *'Serious'* injury over the same year dropped in April, July and October respectively, as shown in the same table.

Additionally, zero fatal injury was achieved within nine months in the variable '*Killed'* except in April, June and December, as displayed in the first column of the table. Although in the case of *'Serious'* injury, zero nonfatal injury was achieved only in October. The interpretation of the result suggests that a smaller number of persons was killed on the Stellenbosch roads in 2012. Besides, the result further reflects the efficiency of the traffic response unit and the outstanding performance of the emergency unit towards the determination to save. This supports minimal accident injuries resulting from daily occurrence of road accidents in Stellenbosch.

In the chart set below, an approximated estimate of 90.0% nonfatal injury, is recorded for *'No injury'* in the Severity of injury. The estimated result obtained in the *'No injury',* demonstrates that most registered accidents in Stellenbosch locality in 2012 are not frequently high speed related accidents, but rather as accidents that occurred within the built-up area with speed limit of ≤60kmlhr.

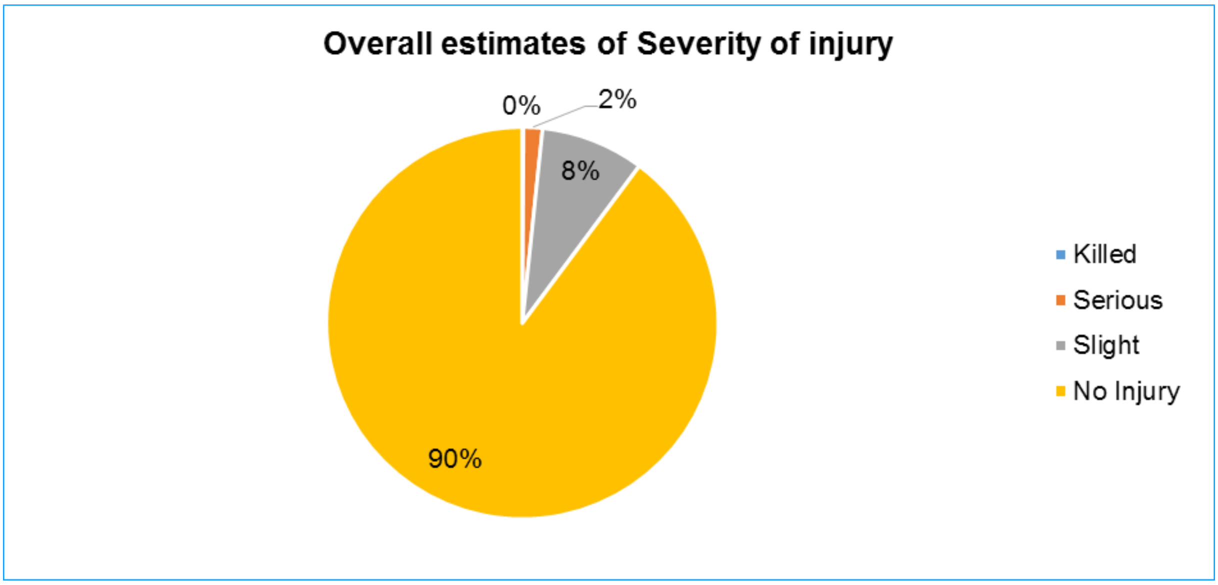

Figure 61: Total estimates of the Severity of injury in 2012

The table below displays the injury data acquired in the *'Summary of persons involved',* which consolidates the result obtained in the *'Severity of injury'.* The 'Summary of persons involved', as previously described in this section, could be referred to as the total estimate of the injury data with regard to the involvement of all the road users [pedestrians, passengers, drivers and cyclists] in road accidents. In that case, the injury data displayed in the table, is expected to be equal or more than the injury data acquired in the *'Severity of injury', depending on the* inclusion of the injury details of other road users like passengers and pedestrians; however, this is relatively not the case.

From the table, it is understood that the amount of fatal injury recorded in the Severity of injury, increased from 3 fatal injuries to 11 fatal injuries with an incremental rate of 57.2% in the *'Number of persons dead'.* Considering the remaining variables, however, nonfatal injury data presented in the *'Number of persons seriously injured'*, demonstrates an increment rate of 5.0% from 50 injury data to 60 injury data. The slight increment observed shows that some injury data are missing.

Quite the reverse, the injury data arrayed in the *'Number of persons slightly injured'* and *'Number of persons not injured'* as displayed in the last two columns of the table, drastically declined instead of gradually increasing. Therefore, a decline rate of 58.6% was observed in the total amount of injury data arrayed in the *'Number of persons slightly injured'* with an estimate of 2,056 missing data, decreasing from 2,875 injury data to 819 injury data. Similar observation was surmised in the total amount of injury data grouped in the *'Number of persons not injured'*. In this case, a decline rate of 23.0% was observed in the injury data assembled in the *'Number of persons slightly injured'* with an estimate of 102 missing data, decreasing from 273 injury data to 171 injury data.

Table 44: Summary of persons involved estimates in road accident occurrence

Total estimates of Summary of persons involved in RTA occurrence in 2012				
Months	**Number of Persons Dead *[Killed]***	**Number of Persons seriously Injured**	**Number of Persons slightly Injured**	**Number of Persons not Injured**
Jan	2	3	8	55
Feb	1	11	38	92
Mar	1	11	23	81
Apr	0	2	3	41
May	0	1	13	97
Jun	0	5	6	44
Jul	0	3	9	66
Aug	0	2	13	93
Sep	2	6	12	56
Oct	1	5	12	66
Nov	2	5	30	88
Dec	2	6	4	40
Total scores	**11**	**60**	**171**	**819**

According to the definition of the *'Summary of persons involved'*, absolutely, a concise estimate of the total scores obtained should be an accurate reflection of all the accident injury data pertaining to all the road users. Hence, the result demonstrated in the Figure 62 is not a true reflection of the result demonstrated in the Figure 61, with the addition of other relevant road users such as mentioned earlier. This signifies lack of effective inspection procedures from the authorised supervisors in charge of validating the completed ARFs.

From a different perspective, as displayed in the two charts, it is construed that despite the effect of the missing data, however, all variables still preserved their positions accordingly. Observably, in the chart displayed in the Figure 62, a total estimate of 1.0% persons were killed in 2012. This estimate is improved through the inclusion of the fatal injury details of the passengers and pedestrians. Similar improvement was observed in the number of road users who were seriously injured. An improved estimate of 6.0% were actualised as the total estimate of all the road users that were seriously injured in the road accidents in Stellenbosch in 2012.

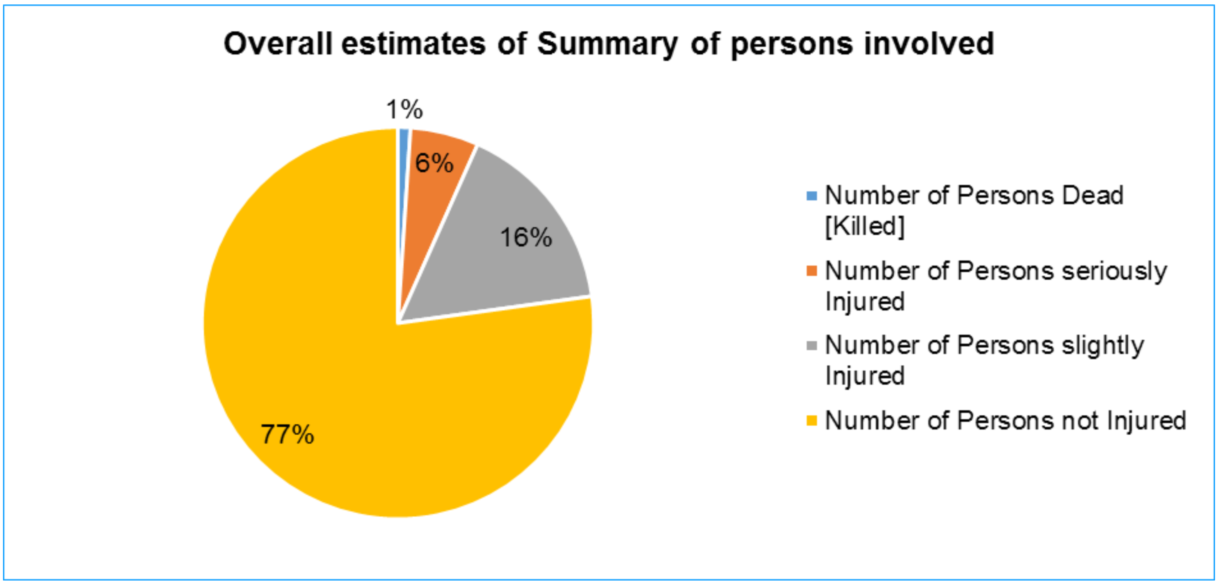

Figure 62: Total estimates of the Summary of persons involved in 2012

In the case of persons slightly injured and not injured, no improvement was observed in the estimate obtained due to huge amount of missing data. Substantially, from the same chart, an approximated percentage rate of 16.0% were realised as the total estimate of all the road users that were slightly injured, while a total of 77.0% persons were not in any way injured in any accident injuries registered in 2012. In spite of the anomalies discovered, the results obtained in each variable illustrate that more road users are uninjured in road accidents, and a few percentage estimates of road users is considered slightly injured. Similar illustration indicates that fewer number of road users are killed and seriously injured in road accidents in Stellenbosch the same year.

The problem of data insufficiency, in this context, could be attributed to misinterpretation of data fields or variables, which is perceived as a result of the insensible attitudes of the form users *[accident reporting officials],* towards the inability to include the injury data gathered from the accident injury details of the drivers and cyclists, with the injury data of the passengers and pedestrians as instructed in the ARF. This effect has rendered some information unfit for decision making actions in the road safety systems.

B.2 Contributory factors related analysis in the Accident related factors

This section presents the analysis results obtained from the two major contributory factors categorised directly under the Accident related factors. The two related data fields have their distinct impacts on the cause of accidents in the Stellenbosch area. A practical interpretation of the result will be explicated in this section, to justify the major causes of accidents within the vicinity. The first discussion centres on the *'Light condition'*, while the subsequent discussion focuses on the '*Weather conditions and visibility*'. This discussion will showcase or demonstrate the influence or effect of the *'light'* and *'weather'* as one of the potential factors contributing to the causes of environmental accidents.

B.2.1 Analysis of the Light condition

The table set below contains the five variables grouped under the Light condition. Actually, these five variables are designated as part of the contributing factors that may instigate the occurrence of road accidents along the Stellenbosch roads. The analysis of these factors could in many ways assist the traffic engineers, in understanding the reasons behind the occurrence of road accident in a logical way. According to the data present in the table below, it is understood that more road accidents occurred in the daylight period. As generally understood, daylight can be described as the '*exact period after the sunrise and before the sunset, as long as there is still a clear light before sundown, and sometimes refer to as daytime*'.

From the simple definition given above, one can easily rationalise that these periods fall between the accident occurrence periods analysed in the subsection 4.2.3 above. As illustrated in the chart shown in Figure 21 in the same subsection, the two peak periods indicated with red colour falls in the beginning of the daytime and towards the end of the daytime. This validates the high scores realised in the variable *'Daylight'* over a year.

In addition, in the chart below, the result demonstrates that an estimate of 76.0% road accidents occurred in the *Daylight,* which is more than the estimated scores realised in the other variables. The estimated scores actualised in the *Night-lit by street light* and *Night-unlit*, demonstrated fair results in terms of accident occurrence rate, with each variable having percentage estimate of 16.0% and 6.0% respectively. These illustrations explain that large number of road accidents recorded in the *Daylight* is as result of intense traffic congestions around the Stellenbosch area (Sinclair & Murdoch 2012), while fewer number of road accidents recorded in the Night-lit by street light, Night-unlit and Dawn/Dusk is as a result of light traffic volumes in the area (Traffic Accident Statistics 2002; Traffic Accident Statistics 2005).

Table 45: Light condition estimates in road accident occurrence

Total estimates of Light condition in RTA occurrence in 2012					
Months	**Daylight**	**Night [Lit by Street Lights]**	**Night [Unlit]**	**Dawn/Dusk**	**Other**
Jan	126	22	6	3	0
Feb	187	31	9	5	0
Mar	187	34	19	3	1
Apr	138	31	15	1	0
May	151	49	13	10	1
Jun	117	41	10	8	0
Jul	100	41	15	6	0
Aug	164	42	17	7	1
Sep	154	23	12	5	0
Oct	189	29	8	4	0
Nov	185	22	7	1	0
Dec	113	15	9	1	2
Total scores	**1811**	**380**	**140**	**54**	**5**

However, other relevant variables like *'Dawn/Dusk'* and *'Other'* contributed lowest to the occurrence rate of accident in Stellenbosch locality, at rate of 2.0% and 0.2% respectively. Meanwhile the result obtained during the dusk/dawn is not impressive because confusion might ensue while interpreting the meaning of this particular variable. The grouping of two different words for one result is not promoting the chance of acquiring more data. This is observed with the intention of enhancing the technical efficiency of the traffic engineers with the desire of discovering more facts about the cause of the accident.

In actual fact, dusk means nightfall, which is the time of the day that simultaneously follows sunset; while dawn means daybreak, which is the first light of the day after a night. This recommends that accidents that occur in the nightfall is different from accidents that occur in the daybreak. Therefore, the desire to separate this single variable will strengthen a feasible analysis process towards a reliable result.

In this section, overview of the monthly estimates of the Light condition is displayed graphically in the Figure 64. The monthly distribution of the data points actualised in five variables reflect accumulation of high values from May to August as shown in the Table 45. Understandably, these four months fall within the winter and spring periods. This insinuates that more accidents occurred mostly in the two seasons than any other seasons. To substantiate this further, the results actualised in the Light condition, is integrated into the four major seasons in South Africa, such as autumn, winter, spring and summer.

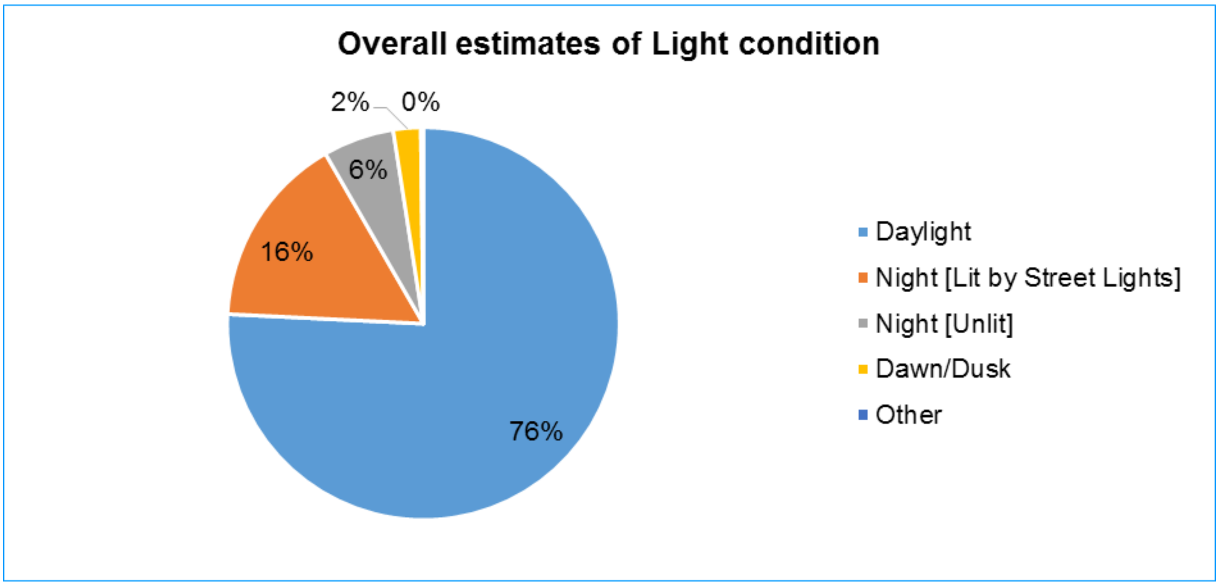

Figure 63: Total estimates of the Light condition in 2012

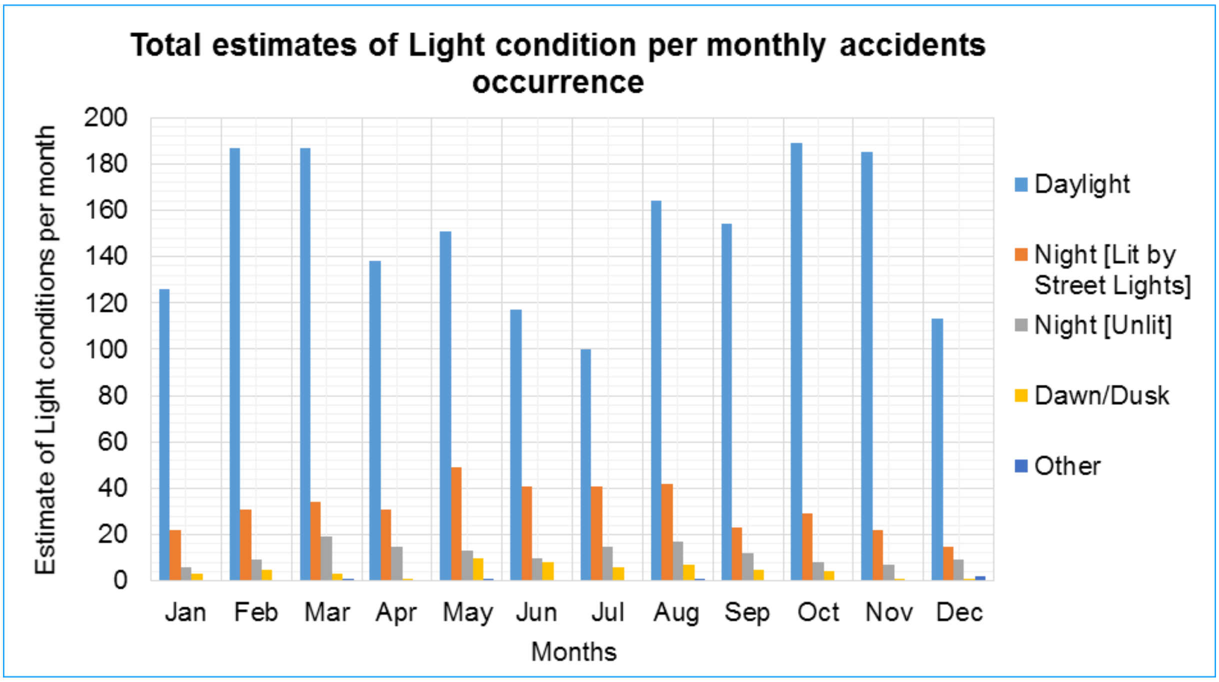

Figure 64: Total estimates of the Light condition per month

In the daylight, high occurrence of road accidents was observed in October, February, March and November respectively. Clearly, these months fall within the summer periods. Other results obtained illustrate that high estimate of road accidents that occurred in the night periods along the roads with lit or unlit street lights fall mostly within the winter, spring and autumn periods, that is in the month of March, April, June, July and August respectively.

Observably, as regards the *'Dawn/dusk',* the result achieved reflects similar trends as that of the result obtained in the *Night-lit by street lights*. This indicates that high estimates of road accidents occurrence in the dawn or dusk fall within the winter and spring periods only. However, the last of the five variables does not has a direct technical title like other variables. Although, it is an option designated for undefined variable in the Light condition. Therefore, in this context, the result is considered insignificant.

From these clarifications, considering the occurrence of road accident in the night time, winter period contributed greatly to the high number of accidents recorded in the night than any other seasons. Literally, in the winter period, night falls so quickly and daybreak comes late in the morning. Thus, not all road users have greater visibility or vision of seeing what lies ahead of them because of the paramount weather problems during this particular period.

B.2.2 Analysis of the Weather conditions and visibility

In this section, the influence of *Weather conditions and visibility* as a considerable or likelihood contributor to accident occurrence is analysed. Moreover, this particular field consist of nine related variables arranged in the table below. The definition of these variables is related to the environmental conditions, which might be imputed to environmental pollutions and seasonal changes. The purpose of analysing this particular field, is to gain more in-depth understandings into the uncertainties resulting in road accidents in the Stellenbosch locality.

Table 46: Weather conditions and visibility estimates in road accident occurrence

Total estimates of Weather conditions & visibility in RTA occurrence in 2012									
Months	**Clear**	**Overcast**	**Rain**	**Mist/Fog**	**Hail/Snow**	**Dust**	**Fire/Smoke**	**Severe Wind**	**Unknown**
Jan	153	2	3	0	0	0	0	0	1
Feb	229	3	2	1	0	0	0	0	1
Mar	215	18	16	2	0	0	2	0	2
Apr	151	9	20	1	0	0	0	0	0
May	155	24	46	4	0	1	0	0	0
Jun	100	39	36	1	0	1	0	0	4
Jul	85	26	48	2	0	1	0	0	1
Aug	159	15	55	0	0	0	0	0	2
Sep	144	20	33	3	0	1	0	0	1
Oct	186	20	24	2	0	0	0	1	2
Nov	201	7	5	0	0	2	0	2	0
Dec	135	3	2	0	0	0	0	0	2
Total scores	**1913**	**186**	**290**	**16**	**0**	**6**	**2**	**3**	**16**

From the spread of the scores showed in the chat below, an estimate of 79.0% road accidents occurred in the *clear* weather condition. This is a huge amount compared to other variables in the same table. Recall, in subsection 4.2.3 above, it was revealed that most inhabitants of Stellenbosch area travelled regularly between the time of 7:00:00 am to 6:59:00 pm more than any other periods. Thus, between these times, clear condition of weather and good visibility are always experienced, except in the winter periods. As a result of this, there is tendency of achieving such huge proportion of accidents in the clear weather condition.

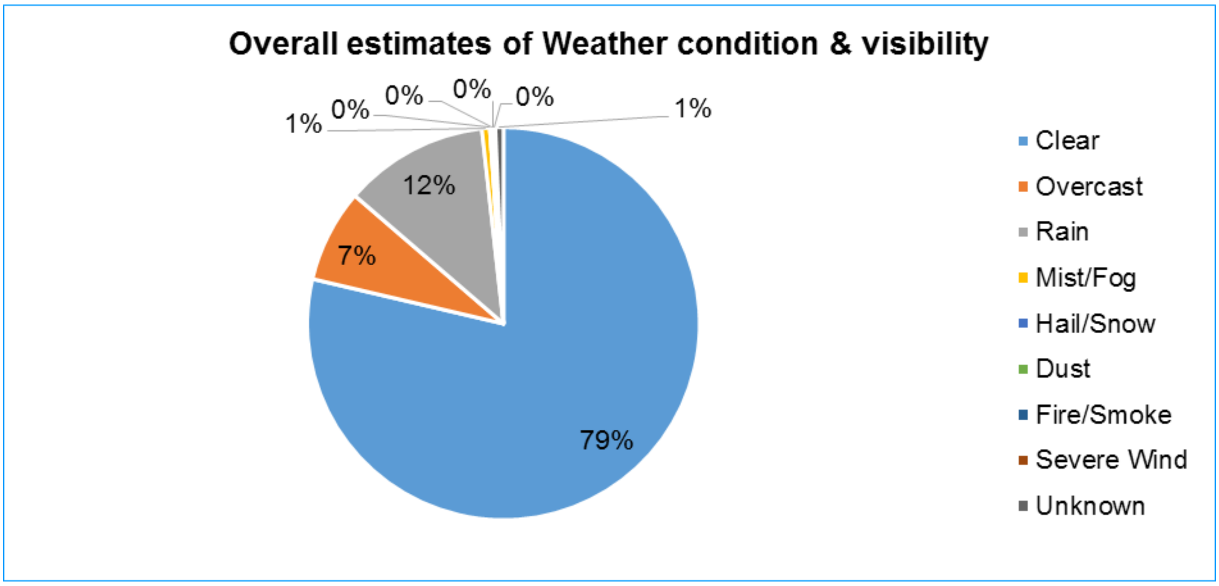

Figure 65: Total estimates of the Weather conditions and visibility in 2012

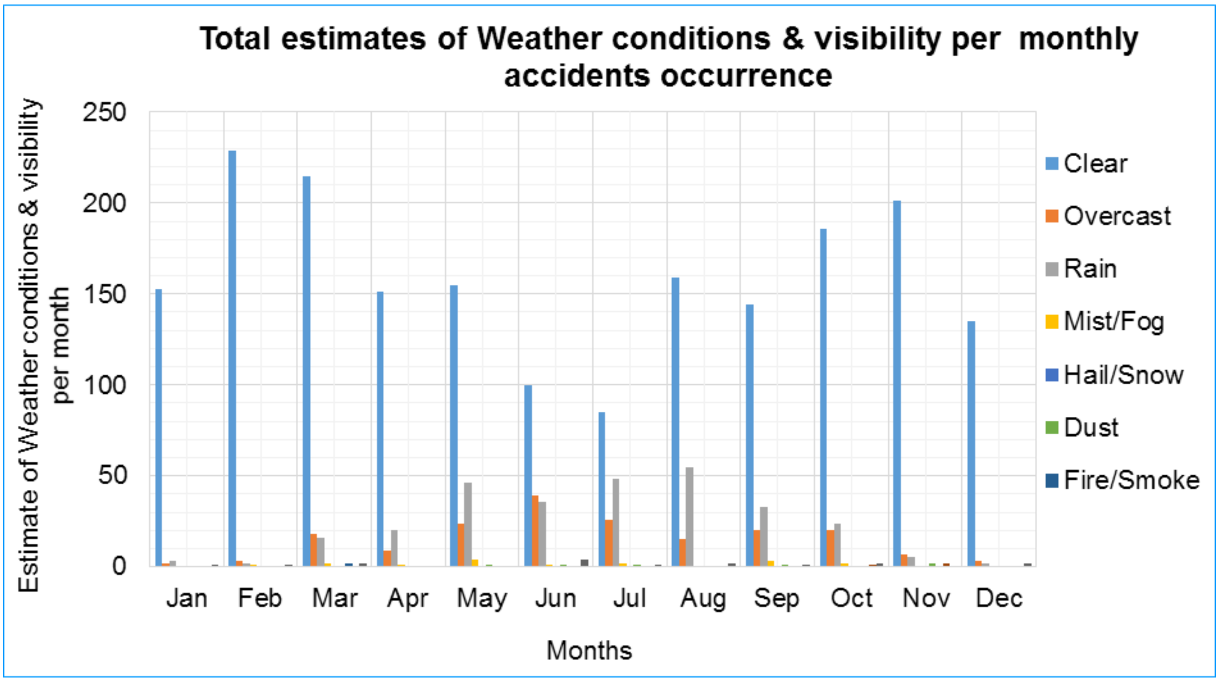

Figure 66: Total estimates of the Weather Condition & Visibility per month

In addition, a total estimate of 12.0% road accidents occurred in the rainy condition; while in the overcast weather condition, an estimated proportion of 7.0% road accidents occurred, which might affect the visibility of some drivers/cyclists. However, other variables contributed an approximated estimate of 2.0%, which is marginally smaller to the estimated scores obtained in the first three variables.

The monthly estimated scores assembled in the Table 44, is plotted in the chart above. The chart presented similar structure to the chart showed in the Figure 64. According to the result obtained, only three weather conditions contributed greatly to the occurrence of road accidents from March to October in 2012. On the contrary, few records of the accidents were attributed to other variables like Fire/smoke, Severe wind, Dust, Mist/fog and Unknown. And also, none was recorded for Hail/snow through the year.

B.3 Benchmark table for Holiday Periods, Weekdays, Peak periods and Accident type

Table 47: Comparison of validated holiday dates/days in road traffic accidentsc between the Stellenbosch Municipality and the City of Cape Town Metropolitan Municipality

Road accidents occurrence during Holiday Periods						
Public Holiday	**Stellenbosch in 2012**			**City of Cape Town in 2005[35]**		
	Date	**Day of week**	**Accidents**	**Date**	**Day of week**	**Accidents**
New Year's Day	1-Jan	Sunday	2	1-Jan	Saturday	163
Human Rights Day	21-Mar	Wednesday	5	21-Mar	Monday	146
Public Holiday	22-Mar	Thursday	5	22-Mar	Tuesday	258
Good Friday	6-Apr	Friday	1	9-Apr	Saturday	227
Family Day	12-Apr	Thursday	10	12-Apr	Tuesday	227
Freedom Day	27-Apr	Friday	6	27-Apr	Wednesday	195
Worker's Day	1-May	Tuesday	2	1-May	Sunday	180
Youth Day	16-Jun	Saturday	4	16-Jun	Thursday	162
National Women's Day	9-Aug	Thursday	3	9-Aug	Tuesday	156
Heritage Day	24-Sep	Monday	1	24-Sep	Saturday	195
Day of Reconciliation	16-Dec	Sunday	1	16-Dec	Friday	215
Christmas Day	25-Dec	Tuesday	1	25-Dec	Sunday	117
Day of Goodwill	26-Dec	Wednesday	2	26-Dec	Monday	130
Public Holiday	27-Dec	Thursday	1	27-Dec	Tuesday	126

[35] Data extracted from a report published by the Traffic Accident Statistics (2005) on the validation of Holiday Periods of road traffic accidents in the City of Cape Town Metropolitan Municipality in 2005.

Table 48: Comparison of road traffic accidents per Weekdays between the Stellenbosch Muicipality and the City of CapeTown Metropolitan Municipality

Road accidents occurrence per Weekdays				
Weekday	Stellenbosch in 2012		City of Cape Town in 2005[36]	
	Total scores	Percentage estimates	Total scores	Percentage estimates
Sundays	202	8.9%	8854	10.39%
Mondays	307	13.5%	12494	14.63%
Tuesdays	305	13.4%	12298	14.40%
Wednesdays	342	15.1%	12064	14.12%
Thursdays	345	15.2%	11997	14.04%
Fridays	407	17.9%	14965	17.52%
Saturdays	362	15.9%	12726	14.90%

[36] Data extracted from a report published by the Traffic Accident Statistics (2005) on road traffic accidents per Weekday in the City of Cape Town Metropolitan Municipality in 2005.

Table 49: Comaprison of road traffic accidents Peak periods between the Stellenbosch Municipality and the City of Cape Town Metropolitan Municipality

Peak periods in Time of road accident occurrence														
Stellenbosch Municipality in 2012														
Hour	**Month**												**Total**	**Percent**
	Jan	**Feb**	**Mar**	**Apr**	**May**	**Jun**	**Jul**	**Aug**	**Sep**	**Oct**	**Nov**	**Dec**		
7:59:00 AM	11	15	16	13	15	19	10	19	14	10	13	6	**161**	**6.7%**
4:59:00 PM	15	27	21	16	16	9	13	27	10	24	22	9	**209**	**8.6%**

City of Cape Town Municipality in 2005[37]									
Hour	**Weekdays**							**Total**	**Percent**
	Sundays	**Mondays**	**Tuesdays**	**Wednesdays**	**Thursdays**	**Fridays**	**Saturdays**		
07:00-08:00	157	1164	1159	1022	946	1010	261	**5719**	**7.0%**
17:00-18:00	570	1015	1080	1068	1073	1110	706	**6620**	**8.0%**

[37] Data extracted from a report published by the Traffic Accident Statistics (2005) on road traffic accidents per Peak periods in Time of Accident in the City of Cape Town Metropolitan Municipality in 2005.

Table 50: Comparison of road traffic accidents by Accident type between the Stellenbosch Muncipality and the City of Cape Town Metropolitan Municipality

Road accidents occurrence by Accident type				
Road surface type	Stellenbosch in 2012		City of Cape Town in 2005[38]	
	Total scores	Percentage estimates	Total scores	Percentage estimates
Head/Rear End	710	38.9%	21440	47.5%
Head On	27	1.5%	911	2.0%
Sideswipe [opposite directions]	299	16.4%	6396	14.2%
Sideswipe [same direction]	271	14.8%	40	0.1%
Approach at Angle [both travelling straight]	148	8.1%	3295	7.3%
Single Vehicle [left the road]	53	2.9%	N/A	N/A
Single Vehicle [overturned]	36	2.0%	1009	2.2%
Accident with Pedestrian	86	4.7%	5426	12.0%
Accident with Animal	40	2.2%	648	1.4%
Accident with Train	1	0.1%	63	0.1%
Accident with Fixed/Other Object	155	8.5%	5926	13.1%

Appendix C: Additional analyses and findings on Road related factors

C.1 Introduction to the analysis of the Road impediment related conditions

The three data fields considered in this section are *'Road marking visibility', 'Obstructions'* and *'Overtaking control'.* However, only two data fields out of the aforementioned three data fields were discussed in this section. Furthermore, the section focuses only on the analysis of the Road impediment related conditions. The discussion offers insight into the effect of some factors that may be hindering the safety of the road users in terms of the road marking visibility details, obstructions details, and overtaking control details on a particular accident location on the Stellenbosch roads.

In this section, *Obstructions* consist of five variables while the *Overtaking control* comprises of four categorical variables, which are offered as the *answer options* in the data form. The effect of the three data fields in the accident occurrence is primarily based on the accurate correlation of the evidences found at the scene of the accident with the variables representing the right answer options to the evidences. Therefore, it is necessary for a reporting official inspecting an accident, to exhibit a competent capability of relating the information provided in the ARF with the findings found at the location of a particular accident.

[38] Data extracted from a report published by the Traffic Accident Statistics (2005) on road traffic accidents by Accident type in the City of Cape Town Metropolitan Municipality in 2005.

C.1.1 Analysis of the Road marking visibility

The table set below presents the outcome of the analysis performed on the *Road marking visibility*. The four variables identified in this particular field are *Unknown, Good, Not good,* and *N/A [not available]*. These four variables defined the conditions of the road marking as one of the probable contributors to the cause of the road accidents in the Stellenbosch area.

Table 51: Road marking visibility estimates in road accident occurrence

Total estimates of Road marking visibility in RTA occurrence in 2012				
Months	**Unknown**	**Good**	**Not good**	**N/A**
Jan	3	134	3	16
Feb	4	205	3	19
Mar	5	220	4	16
Apr	5	145	5	18
May	5	197	3	20
Jun	9	135	5	22
Jul	6	129	3	26
Aug	8	185	7	26
Sep	4	163	3	19
Oct	4	195	3	27
Nov	1	181	1	30
Dec	1	122	2	13
Total scores	**55**	**2011**	**42**	**252**

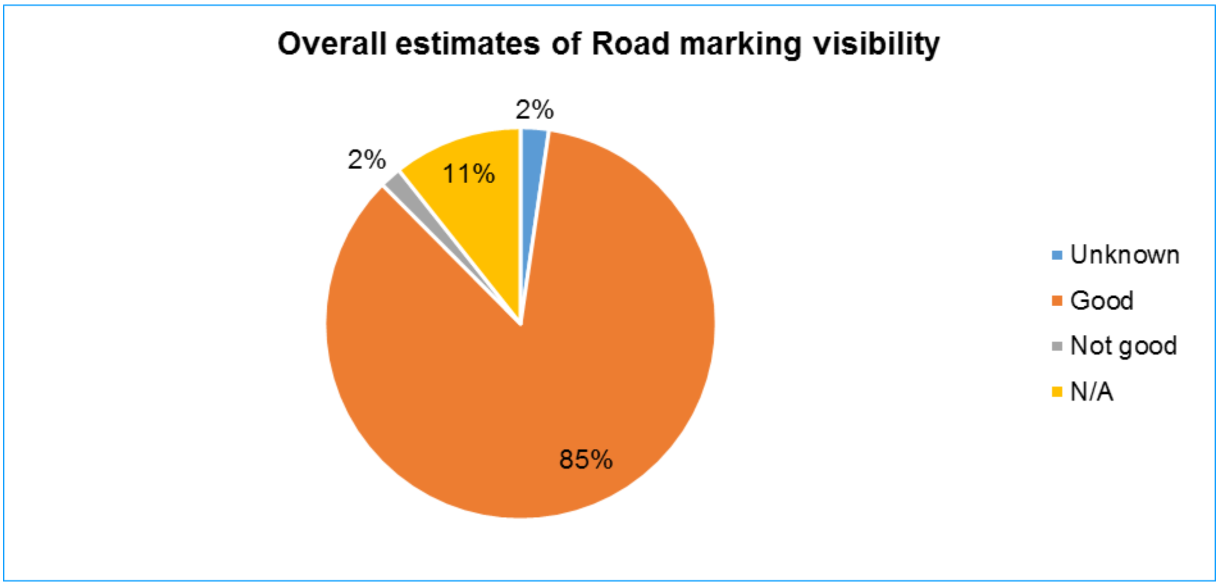

Figure 67: Total estimates of the Road marking visibility in 2012

However, many roads within the Stellenbosch jurisdiction are marked to initialise safety precautions, and periodically maintained to guide the composure of the drivers or road users towards the proper usage of the road. From the table set above, few estimated scores indicate that the visibility of road markings at some accident locations within the Stellenbosch locality was poor, unavailable and unknown in some cases, while a huge proportion of accident cases indicates that the visibility of road markings at some accident locations in the locality was in good condition.

As graphically presented in the chart above, it is absolutely understood that large estimates of road accidents occurred along the Stellenbosch roads with good road marking visibility. Statistically, an estimated percentage of 85.0% road accidents that occurred in the Stellenbosch area, laid no claim to defective road markings. In contrast, an approximated estimate of 2.0% road accidents showed that the visibility of the road markings is poor. A sum estimates of 13.0% road accidents demonstrated that the road marking visibility is some cases is found inaccessible and unknown.

C.1. 2 Analysis of the Obstructions

This section discusses the effect of the *Obstructions* as part of the causal factors of road accidents. As previously stated in the introductory section, five variables are considered in this field, in which only three variables out of the main five are defined accordingly to the probable circumstances that might cause or result in road accident along the Stellenbosch roads. The three variables defined are *Accident site, Roadworks* and *Roadblock,* and the remaining two variables with undefined circumstances are classified as *Other* and *None*.

Table 52: Obstructions estimates in road accident occurrence

Total estimates of Obstructions in RTA occurrence in 2012					
Months	**Accident Site**	**Roadworks**	**Roadblock**	**Other**	**None**
Jan	2	3	0	4	134
Feb	4	2	1	9	195
Mar	2	10	1	11	201
Apr	0	0	0	5	165
May	1	0	0	5	203
Jun	3	5	0	1	152
Jul	2	0	0	5	148
Aug	1	5	1	7	193
Sep	2	3	0	9	161
Oct	1	0	0	5	196
Nov	0	3	0	7	184
Dec	0	1	0	2	121
Total scores	**18**	**32**	**3**	**70**	**2053**

In the chart below, sum approximated estimate of 3.0% showed that on-going road maintenance, road barriers and accident sites contributed minimally to the circumstances leading to many road accidents in the Stellenbosch area. On the other hand, a great number of figures indicated that no obstructions, in any way, led to the circumstances of an estimated proportion of 94.0% road accidents; although, a reasonable estimate of 3.0% was classified indeterminate with regard to whether any obstruction led to the road accidents reported. In this case, other relevant obstructions circumstances are not made available in the ARF, therefore, a space for specification of such findings is provided in the form.

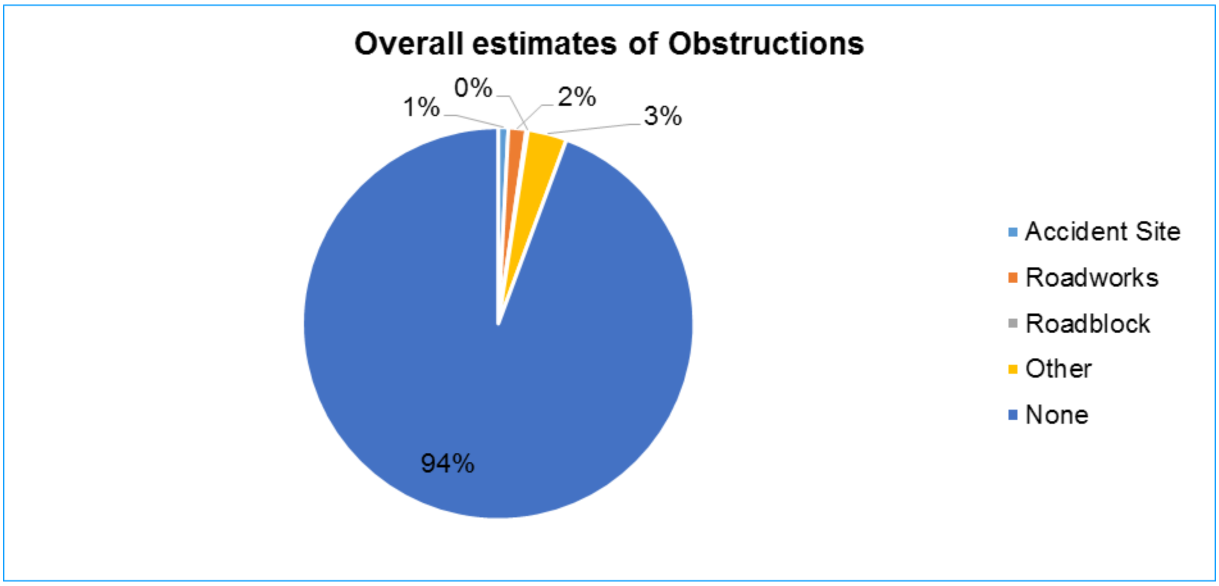

Figure 68: Total estimates of the Obstructions in 2012

In many cases, according to findings acquired during the data assessment actions, several spaces designed for specification purposes were unused, despite the need for specifying the possible circumstances leading to the incident, in order to acquire more relevant data under obstructions.

C.2 Introduction to the analysis of the Road control related conditions

This section centralises on the discussion of the results acquired in the *Road control related conditions.* In this discussion, two data fields are considered, in demonstrating the involvement of the road control conditions in the occurrence of road accidents. However, the two related data fields discussed the *Traffic control type* and *Condition of road signs*. Traffic control type consist of thirteen variables, which define the type of traffic control used, and the condition of both the controlled and uncontrolled intersections within the Stellenbosch area.

The analysis of the thirteen variables offer results on the circumstances at the intersection, along the road where the accidents occurred. On the part of the Condition of road signs, four variables are discussed, which define the current state of the road signs. The four variables illustrated the conditional functionalities of the road signs.

C.2.1 Analysis of the Traffic control type

The table set below contains the estimates acquired from the analysis of the *Traffic control type*. From the same table, variables like *Robot, Stop sign, Uncontrolled junction, Not a junction-crossing or barrier line* and *Barrier line* produced high estimated scores. Observably, from Figure 69, an estimated proportion of 34.0% road accidents occurred at some locations away from the intersections or barrier line. With further clarifications, this illustration implies that a high number of accidents occurred mostly, on a straight road over a certain specific kilometre or mile away from any junction, crossing and barrier line.

In addition, from the chart below, another high estimate of 32.0% road accidents occurred at the intersections controlled by *robot* within the Stellenbosch area. This estimate is approximately 2.0% lesser than the estimated proportion obtained in *'Not a junction, crossing or barrier line'*. Similarly, *'Uncontrolled junction*' without a stop sign produced an estimated proportion of 8.0%, which is approximately 5.0% lesser than the estimated proportion calculated for intersections controlled by a *stop sign*. However, as part of the results obtained, a small estimate of 2.0% road accidents was connected to traffic control type like pedestrian crossing in Stellenbosch area. This estimated proportion illustrates that fewer road accident cases were occurred as a result of poor observations of the pedestrian crossing by the road users, predominantly within the built-up areas in the Stellenbosch locality. This result may influence the road accident estimates obtained from the *'Accident with pedestrian*' in the section 4.2.4 above.

The estimated proportions of the junctions controlled by *robots*, *stop sign* and uncontrolled junctions reflect some part of the results demonstrated in the '*Accident type'* discussed in subsection 4.2.4 above. Clearly, type of accidents like head/rear end, head on, sideswipe-*opposite directions,* sideswipe-*same directions,* and accident with pedestrian are preponderantly encountered at both junctions controlled by robots and stop sign, uncontrolled junctions and pedestrian crossing.

However, in the subsection 5.2.2 above, variables like *'not a junction or crossing'* and *'pedestrian crossing'* are among the twelve variables discussed in the analysis of *Junction type*. In the section for Junction type, more data were collected for *'not junction or crossing'* than in the section for Traffic control type. This particular variable is repeated twice under these two different data fields, which are relatively connected with the exclusion of the '*barrier line'*. Meanwhile, regarding the *'pedestrian crossing'*, more accidents data were represented in the section for Traffic control type than in the Junction type.

Table 53: Traffic control type estimates in road accident occurrence

Total estimates of Traffic control type in RTA occurrence in 2012													
Months	**Robot**	**Stop Sign**	**Yield Sign**	**Officer**	**Officer plus Robot**	**Uncontrolled Junction**	**Not a Junction, Crossing or Barrier Line**	**All Robots out of order**	**Some Robots out of order**	**Flashing Robots *[Red/Yellow]***	**Boom**	**Pedestrian Crossing**	**Barrier Line**
Jan	28	13	3	0	0	10	30	0	9	1	0	1	5
Feb	62	19	4	0	0	13	44	1	12	0	0	4	10
Mar	32	24	4	1	0	10	53	1	0	0	0	4	11
Apr	37	14	4	0	0	9	35	0	0	0	1	2	2
May	64	15	4	1	0	12	40	0	0	0	1	4	5
Jun	28	10	2	0	0	11	38	0	0	0	0	3	9
Jul	33	11	4	0	0	11	46	0	0	0	0	3	8
Aug	45	13	9	1	1	8	62	0	0	0	1	2	9
Sep	38	22	4	1	1	8	49	0	0	0	0	2	7
Oct	46	23	4	0	1	13	43	1	0	0	0	2	11
Nov	54	18	11	1	0	12	44	1	0	0	0	4	7
Dec	27	10	2	0	0	5	27	0	0	0	0	0	5
Total scores	**494**	**192**	**55**	**5**	**3**	**122**	**511**	**4**	**21**	**1**	**3**	**31**	**89**

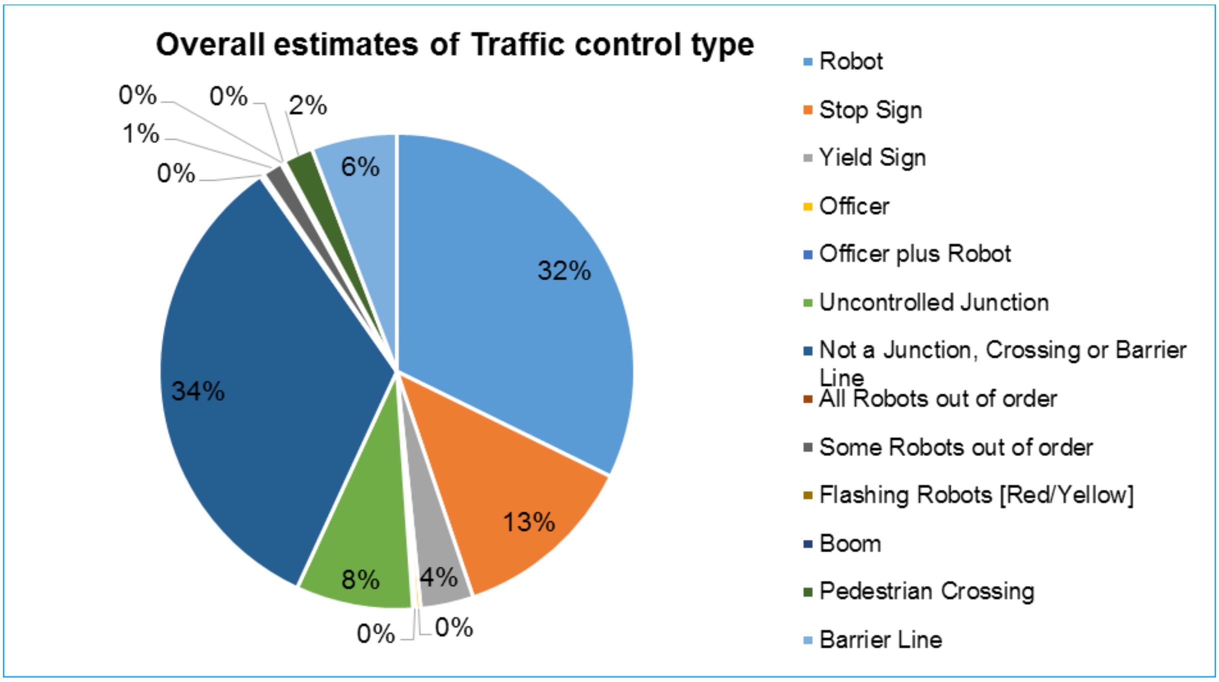

Figure 69: Total estimates of the Traffic control type in 2012

In this context, the two illustrations show that the quality of the data collected by the form users are unreliable. In other word, *'the consequence of inconsistency in the manner at which data is omitted or misrepresented, from one field to another, contributes at large to the impracticality of the data collected'.* In addition, it should be noted that *pedestrian crossing* should be excluded from the Junction type, instead it is better off as a variable in the Traffic control type because it serves better as a traffic control.

C.2.2 Analysis of the Condition of road signs

This section presents feedback on the condition of the road signs at the point of the incident occurrence. In this field, four variables were discussed, which describe the likelihood conditions of the road cautionary signs. From the table below, it is observed that approximated estimate of 91.0 % [2,082] road accidents claimed that cautionary road signs along the Stellenbosch roads were in good condition at the time of the incident occurrence. This states that the road signs are absolutely effective enough to offer a cautionary guide to the road users before the event.

Alternatively, a small approximation of 1.0% dictates that the road signs were not in good at the time of the incident occurrence, while an estimated proportion of 8.0% confirmed that no information was made available as regards the condition of the cautionary road signs at the incident occurrence, which might be as a result of late incident reportage. In addition, a smaller approximated estimate of 0.1% indicated that the cautionary signs on the roads are damaged or preferably considered missing at the point of incident occurrence.

Table 54: Condition of road signs estimates in road accident occurrence

Total estimates of Condition of road sign in RTA occurrence in 2012				
Months	**Good**	**Not Good**	**Damaged or Missing**	**N/A**
Jan	139	2	0	11
Feb	217	1	0	10
Mar	218	2	0	14
Apr	159	4	0	11
May	205	2	0	9
Jun	144	2	0	19
Jul	138	4	0	14
Aug	192	6	2	22
Sep	167	1	1	16
Oct	200	2	0	13
Nov	187	0	0	16
Dec	116	0	0	13
Total scores	**2082**	**26**	**3**	**168**

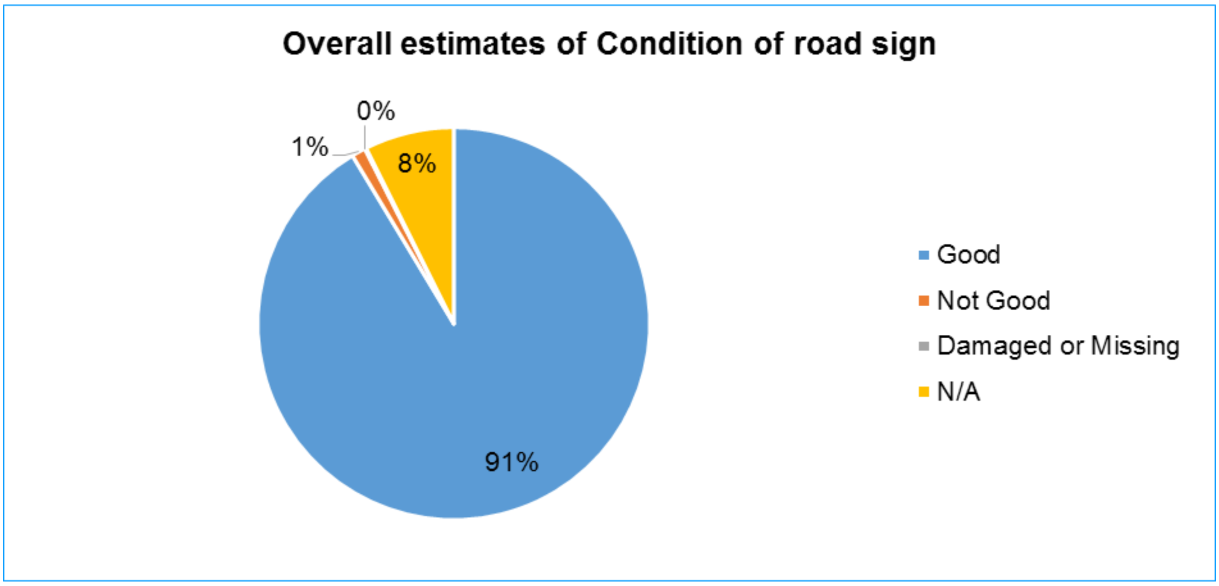

Figure 70: Total estimates of the Conditions of road sign in 2012

More importantly, a need for more informative and advanced cautionary road signs are significant, as part of the strategies required to stimulate the reduction of RTA on the Stellenbosch roads.

C.3 Introduction to the analysis of the Road layout related conditions

This section discusses the influence of the road layout as one of the probable contributors to the cause of accidents in the Stellenbosch area. This particular factor establishes a relation between the *road direction* and the *position of the vehicles before the accident* occurred. This offers insight into the areas that link, the *speed engaged* by the vehicles while travelling through the *direction of the road* with the *position of the vehicles.*

Three variables are designated under the road direction, which are *Straight*, *Curving* and *Sharp curve [90^0 bend],* but alternatively, the vehicles' position before incident occurrence consist of six variables that covered sufficient details regarding this particular field. The six variables considered are *Correct road lane, Wrong road lane,* and *Wrong side of road,* and other variables are *Road shoulder, On-road parking bay* and *Off-road parking bay*.

C.3.1 Analysis of the Direction of road

In this section, results acquired from the analysis of the three variables categorised in the direction of road are presented in both graphical format and tabular array. In Stellenbosch, according to the estimated scores in the table below, most road accidents reported in 2012, occurred in a straight direction of the road with an estimated proportion of 90.0% [1,900], which is two times larger than the sum estimates of the two other variables. Observably, a total estimate of 8.0% road accidents occurred along the curve direction of the road, and also, a nominal estimate of 2.0% road accidents occurred at the exact sharp curve with 90^0 bend.

Table 55: Direction of road estimates in road accident occurrence

Total estimates of Direction of road in RTA occurrence in 2012			
Months	**Straight**	**Curving**	**Sharp Curve *[90-degree bend]***
Jan	121	19	1
Feb	189	12	6
Mar	192	15	1
Apr	145	11	3
May	180	11	2
Jun	140	11	0
Jul	130	13	7
Aug	176	25	3
Sep	150	14	5
Oct	186	15	5
Nov	178	19	1
Dec	113	7	2
Total scores	**1900**	**172**	**36**

From the above illustrations, it is deduced that curve and sharp curve road directions contributed minimally to the road accidents occurrence in the Stellenbosch area. Therefore, some accident cases attributed to sharp curve, may be as a result of two combinational factors like *excessive speed* and *skidding angle,* which could probably lead to the overturning of a vehicle along the 90^0 bend.

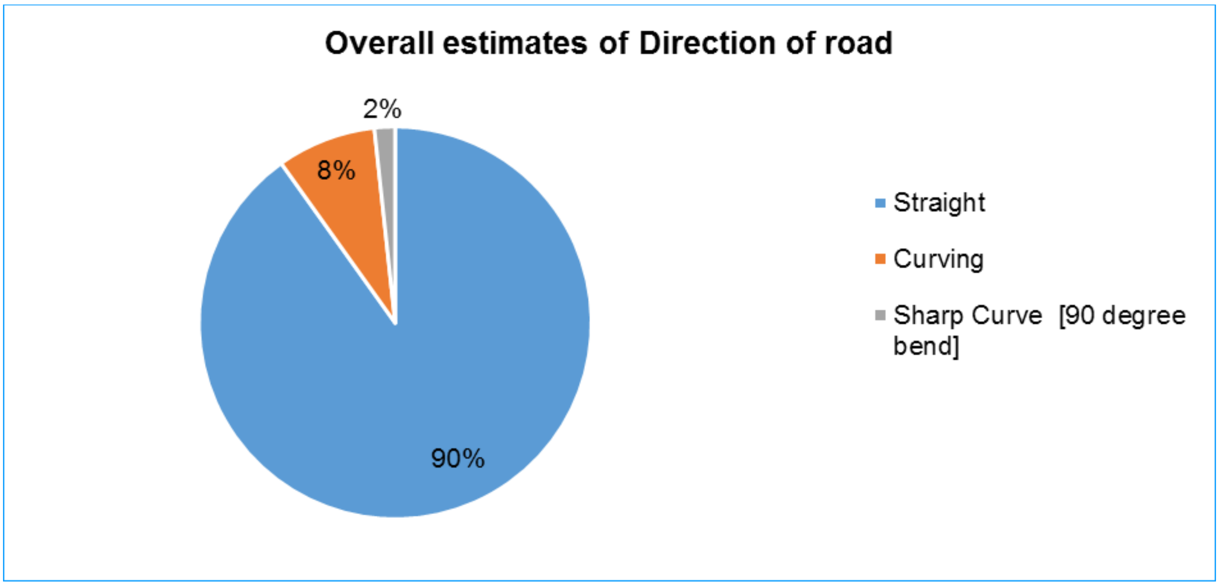

Figure 71: Total estimates of the Direction of road in 2012

These results relationally associate with the results obtained in the Accident type analysis displayed in the Table 11 in subsection 4.2.4 above. Road accidents occurring on the straight and curving directions of the road, could be associated with other variables like Head/rear end, Head on, Sideswipes-*both opposite and same directions*, Approach at angle-*both travelling straight,* and Single vehicle-*left the road.*

C.3.2 Analysis of the Position of vehicle before accident

This section presents a discussion on the results derived from the analysis performed on the six categorical variables in the *position of vehicles before accident*. The discussion offers view on the position of the vehicle before the accident occurred, with the aim of strengthening the interpretation of the circumstances surrounding the incident. In the ARF, this field is created for both single vehicle and multiple vehicle accidents, in which the positional directions of the vehicles before the accident can be understood. These positions are attributed to two main factors, specifically, when vehicles are in *motion* or *stationary* depending on what the driver is doing at the time of the incident.

In addition, if an accident involves multiple vehicles, it is hypothetically understood that the vehicles may be in the same direction or opposite directions before the incident occurred. Thus, two similar or different answer options will be indicated in the ARF for the multiple-vehicles accident, but only one answer option is indicated for a single-vehicle accident in the form. However, the major anomaly discovered here is *data omission,* which may be as a result of the disappearance of the driver involved at the scene of the accident.

In the table below, a huge estimate of 85.0% [1,989] claims that most vehicles involved in road accidents were on the correct road lane before the incident occurred, probably when a vehicle is stationary while waiting in traffic and suddenly bash from behind. Other results show that some vehicles were on the *wrong road lane* and *on-road parking bay,* right before the incident occurred, with an equal approximation of 4.0% each. According to the discussion had with a traffic official, road accidents resulting from wrong road lane, may be connected mostly, to the behavioural characteristics of the vehicles' drivers.

Table 56: Position of vehicles before accident estimates in road accident occurrence

Total estimates of Position of vehicles before accident in the RTA occurrence in 2012						
Months	**Correct Road Lane**	**Wrong Road Lane**	**Wrong Side of Road**	**Road Shoulder**	**On-road Parking Bay**	**Off-road Parking Bay**
Jan	120	7	5	2	12	8
Feb	195	6	5	6	10	7
Mar	189	10	3	6	8	10
Apr	155	8	3	4	6	5
May	186	7	3	7	4	7
Jun	152	5	3	2	1	7
Jul	136	14	3	3	5	7
Aug	195	8	2	3	12	5
Sep	162	11	2	5	8	8
Oct	190	3	2	8	13	9
Nov	190	10	15	1	1	0
Dec	119	4	2	3	6	8
Total scores	**1989**	**93**	**48**	**50**	**86**	**81**

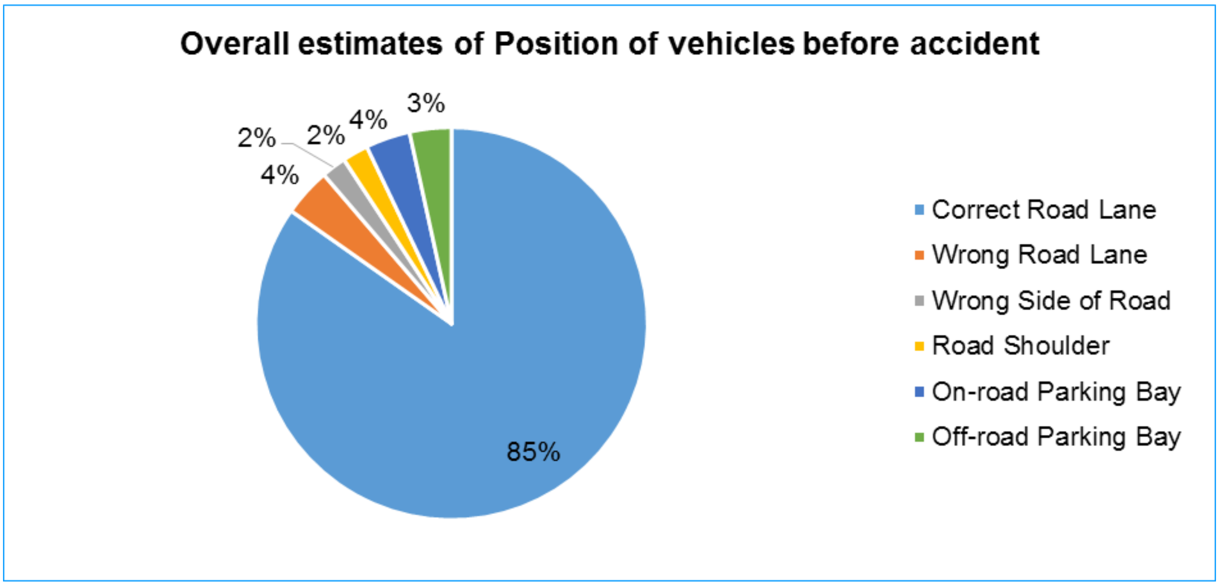

Figure 72: Total estimates of the Position of vehicles before accident in 2012

However, from the other remaining results, a sum estimates of 7.0% was acquired from the analysis of such variables as *wrong side of road, road shoulder* and *off-road parking bay.*

Among these three variables, *off-road parking bay* produced results indicating circumstances that happened in places like home car park, mall parking area, workplace parking lot etc.

C.4 Benchmark table for Road Surface type in Road related factors

Table 57: Comparison of road traffic accidents by Road surface type between the Stellenbosch Municipality and the City of Cape Town Metropolitan Municipality

Road accidents occurrence by Road surface type				
Road surface type	**Stellenbosch in 2012**		**City of Cape Town in 2005**[39]	
	Total scores	**Percentage estimates**	**Total scores**	**Percentage estimates**
Concrete	77	3.2%	4454	5.2%
Tarmac	2266	95.1%	75527	88.4%
Gravel & Dirt	24	1.0%	734	0.9%
Other/Unknown	15	0.6%	4683	5.5%

Appendix D: Additional analyses and findings on Human related factors

D.1 Introduction to the analysis of the Safety measures related conditions

This section discusses the importance of safety regulations in deterring the probable causes of road accidents on the roads within and outside the built-up areas in the Stellenbosch locality. The section encompasses four related variables, such as *Seatbelt fitted/helmet present, Seatbelt/helmet definitely used, Liquor/drug use suspected* and *Liquor/drug use evidentiary tested*.

These four variables are included in the ARF, as a means of measuring the risks that associate with driving or riding without any use of seatbelt or helmet within the Stellenbosch locality. Also, the variables are applicable in determining whether the use of liquor/drug contribute to the cause of road accidents in the Stellenbosch locality, with the aim of ensuring the safety of the driver or to avert any predetermine road traffic injury or occurrence of accident.

These four variables assist in quantifying the degree of safety of the drivers and the cyclists that are involved in the road accidents. The inspection of these variables in a reported accident is implemented by the accident reporting officials, either a police officer or a traffic officer based on the severity of the accident.

[39] Data extracted from a report published by the Traffic Accident Statistics (2005) on road traffic accidents by Road surface type in the City of Cape Town Metropolitan Municipality in 2005.

D.1.1 Analysis of the Seatbelt fitted/helmet present

This section discusses the results produced from the analysis of the answer options categorised in the *Seatbelt fitted/helmet present.* Three related variables are designated as the answer options suitable to guide the simple data collection of this particular field. The distribution of the estimated scores presented in the Table 58 demonstrates estimated high scores in the column for the variable, *'Yes'.* This particular variable specifies whether the vehicles/motorcycles involved in the road accidents had seatbelt fitted or helmet present.

As demonstrated in the chart below an estimated score of 61.0% vehicles/motorcycles involved in the road accidents in the Stellenbosch area, adequately had seatbelt fitted and helmet present. Moreover, this illustration shows the tendency of having a minimal fatality cases. This validates the result obtained in the severity of injury presented in the Appendix B-B.1.1, which declared that only 3 [0.1%] persons, that is, the drivers/cyclists were reportedly killed in road accidents in Stellenbosch.

Table 58: Seatbelt fitted/helmet present estimates in road accident occurrence

Total estimates of Seatbelt fitted/helmet *present* in the RTA occurrence in 2012			
Months	**Yes**	**No**	**Unknown**
Jan	111	21	63
Feb	225	10	76
Mar	172	19	113
Apr	149	12	66
May	182	22	100
Jun	137	16	70
Jul	131	17	61
Aug	177	13	129
Sep	148	12	88
Oct	186	19	112
Nov	185	20	77
Dec	100	11	62
Total scores	**1903**	**192**	**1017**

Considering the variable *'No'*, an indication on lack of seatbelt fitted or helmet present in the vehicles/motorcycles involved in the road accidents. According to the magnitude of the estimated scores obtained in this particular variable, it is observed that a small proportion of 6.0% vehicles/motorcycles involved in the road accidents, had no seatbelt fitted or helmet present. As result of this, perhaps, the lives of the persons involved are endangered or unsafe, due to the lack of seatbelt fitted or helmet present.

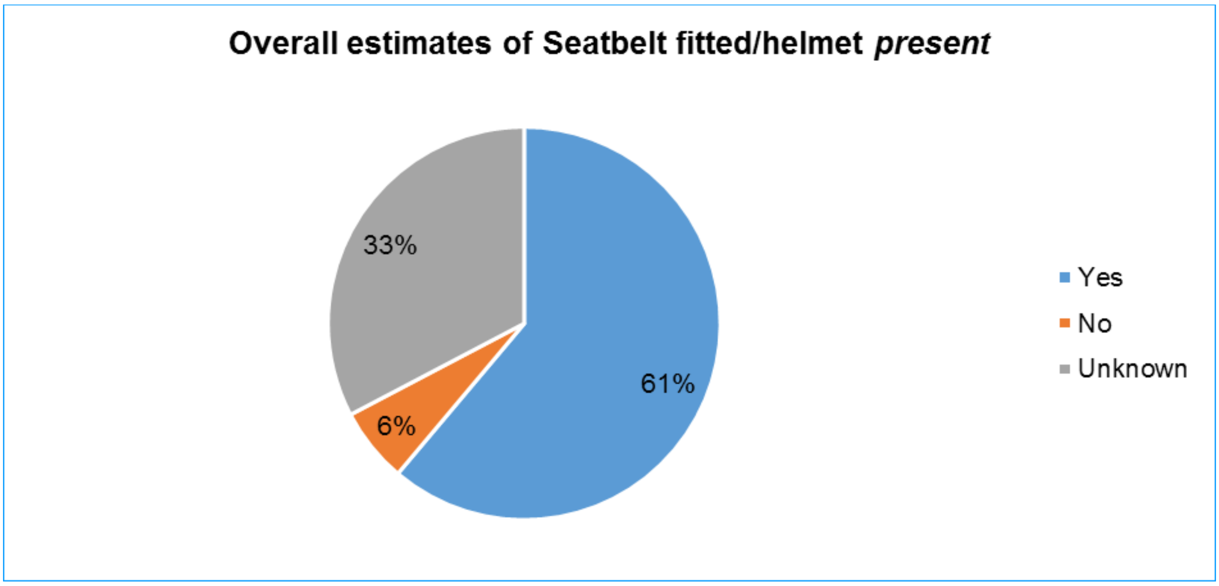

Figure 73: Total estimates of the Seatbelt fitted/helmet present in 2012

The last of the three variables is *'Unknown',* which reflects estimated high scores of 33.0% due to undisclosed or indeterminate information on whether the vehicles/motorcycles involved in any road accidents, have seatbelt fitted or helmet present. This particular case could be attributed to factors, such as late reportage of many accident events, the discretion of information and hit and run cases. However, this may affect the possibility of acquiring many key information during the completion of the ARF.

D.1.2 Analysis of the Seatbelt fitted/helmet definitely used

This section discusses the results acquired from the analysis of the *Seatbelt/helmet definitely used.* The results support the objective of whether the seatbelt fitted/helmet present is definitely applied right before the incident occurred. However, similar variables or answer options are also considered as the right metrics in measuring whether the drivers or cyclists make use of the safety seatbelt or safety helmet, in accordance with the safety instructions enacted by the law enforcement agencies. The use of safety belts and safety helmets by the drivers and the cyclists contribute a reasonable proportion to the reduction of severe injury during RTAs.

In addition, the relevance of this field is to determine or measure the right data for the purpose of defining the significance of knowing whether a seatbelt fitted or helmet present is considerably utilised. The consideration of the variable *'Yes'* is to explore how many drivers and cyclists used their safety belt and safety helmet during any accidents in the Stellenbosch locality. Equally, variable *'No'* is considered as another metric or dimension applied to explore how many drivers/cyclists do not use safety belt and safety helmet during any accidents.

Table 59: Seatbelt/helmet definitely used estimates in road accident occurrence

Total estimates of Seatbelt/helmet definitely used in the RTA occurrence in 2012			
Months	**Yes**	**No**	**Unknown**
Jan	90	24	64
Feb	196	21	77
Mar	132	23	126
Apr	123	14	74
May	158	22	100
Jun	116	24	74
Jul	119	19	59
Aug	157	15	139
Sep	123	16	98
Oct	167	23	115
Nov	165	23	84
Dec	93	11	65
Total scores	**1639**	**235**	**1075**

The third variable, *'Unknown',* is considered as the right metric/dimension required for measuring the indeterminate information regarding whether the drivers/cyclists used safety belt/helmet during the accident. This particular variable is used to justify a specific condition in the RTA, where some necessary data cannot be covered due to some circumstances, such as late incident reporting, injury level of the drivers/cyclists as described in the last paragraph of the previous section [see Appendix D-D.1.1].

The table above presents the estimated scores generated from the analysis of the three variables considered in determining whether accident victims used their seatbelt/helmet regularly. In the chart below, an estimated proportion of 56.0% [1,639] indicated that large number of drivers/cyclists involved in the road accidents within the Stellenbosch locality used seatbelt and helmet. This actual statistic could be achieved through an effective and reliable inspection procedure by the law enforcement agents on a daily basis.

On the contrary, a total estimate of 8.0% declared that few number of drivers/cyclists involved in road accidents within the same locality ignored the use of seatbelt/helmet. This actual estimate is six times smaller than the amount of the drivers/cyclists that complied with the use of seatbelt/helmet while travelling on the roads. Considering the third variable, however, a

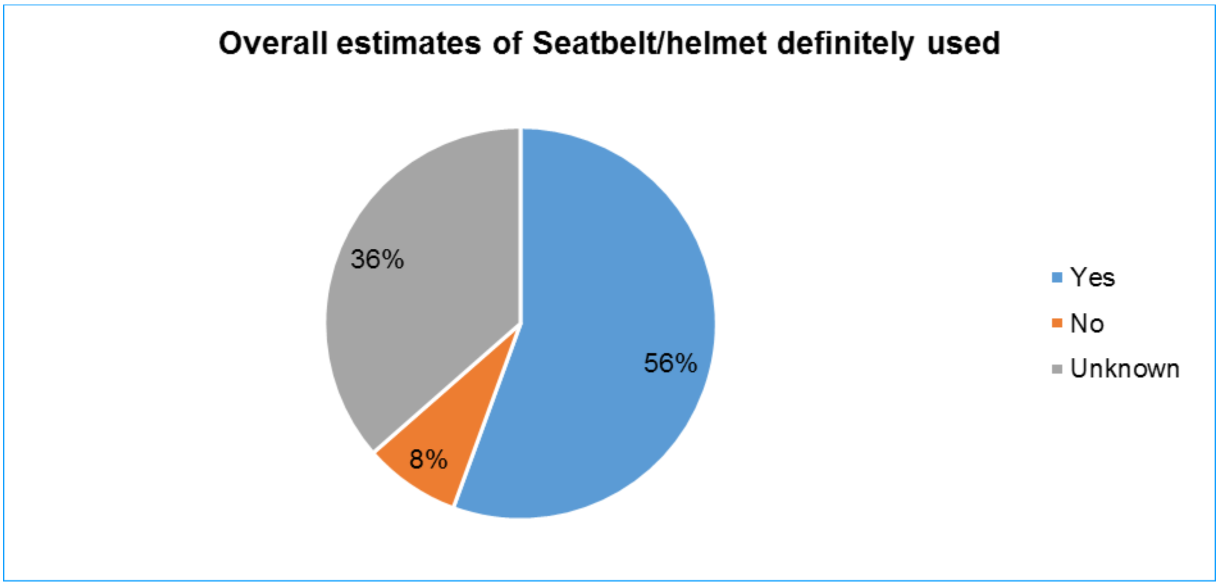

Figure 74: Total estimates of the Seatbelt/helmet definitely used in 2012

proportional estimate of 36.0% data was lost to indeterminate circumstances, where no information declared that neither the drivers used seatbelt, nor the cyclists used helmet during a road accident in the locality.

D.1.3 Analysis of the Liquor/drug use [suspected]

This section discusses the approach implemented in determining the contributing factors to the occurrence of road accidents. This particular field is considered as the metric suitable to examine the suspicion use of liquor/drug by the drivers/cyclists involved in the road accidents in the Stellenbosch locality. The examination process necessitates the application of liquor/drug use detecting device, to ascertain the alcohol/drug consumption rate as a result of the accident. Basically, this procedure could be justified through the means of carryout liquor/drug evidentiary testing, to clarify whether drivers/cyclists involved was under the influence of alcohol or drugs.

In the table below, two variables are considered in grouping the right data pertaining to whether a driver/cyclist involved in the road accidents is suspicion of liquor or drug use. The result displayed in the variable *'Yes'*, confirmed the tendency of suspecting the influence of alcohol use or drug use as the potential cause of the road accidents in Stellenbosch. Under this specific variable, as illustrated in the chart presented in Figure 75, a nominal proportion of 1.0% indicates that some drivers/cyclists were suspected of liquor/drug use at the point of the incident.

Nonetheless, this particular result does not present the exact estimates of the drivers/cyclists found guilty of using influential substances. Relatively, a determined outcome of the drivers/cyclists found guilty of the liquor/drug use are subsequently discussed in the following section. Quite the reverse, the variable *'No'* demonstrates the estimates of the unsuspected drivers/cyclists considered free of liquor/drug use, in accordance with the outcome of the test carried out. As presented in the chart below, an estimated proportion of 99.0% shows that large number of drivers/cyclists involved in the road accidents were not suspected of liquor/drug use.

Table 60: Liquor/drug use suspected estimates in road accident occurrence

Total estimates of Liquor/drug use *[suspected]* in the RTA occurrence in 2012		
Months	**Yes**	**No**
Jan	1	175
Feb	5	266
Mar	7	247
Apr	3	195
May	3	186
Jun	2	186
Jul	1	178
Aug	1	271
Sep	2	221
Oct	3	276
Nov	3	244
Dec	1	160
Total scores	**32**	**2605**

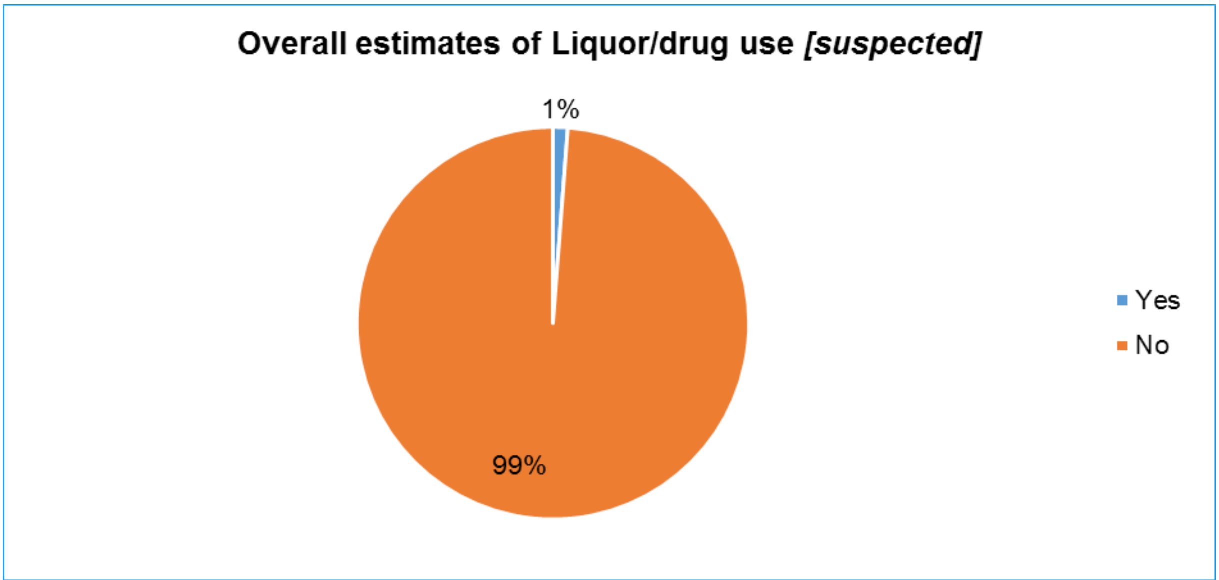

Figure 75: Total estimates of the Liquor/drug use *[suspected]* in 2012

Due to the large number of unsuspected influence of alcohol or drug use on the drivers/cyclists, it is concluded that some other factors like high speed, drivers'/cyclists' impatience, vehicular condition and traffic rules encroachment.

D.1.4 Analysis of the Liquor/drug use [evidentiary tested]

This section basically discusses the results generated from the statistical analysis of the data points gathered in the *Liquor/drug use [evidentiary tested].* This particular field consist of two variables, which relatively confirmed the evidentiary result of suspected use of liquor or drug. This process requires the service of the *Alcohol Breathalysers [AB]*, in measuring the level of the liquor concentration in the blood of the drivers/cyclists involved in an accident, as mentioned in subsection 2.1.2 above. This device measures the concentration level of the blood alcohol in grams per millilitre [g/mL].

Understandably, the nominal proportion illustrates that fewer road accidents occurred as a result of the influence of liquor and drug use in the Stellenbosch locality. A regular exercise carried out by the traffic law enforcement agency, ensures a drastic reduction in the drink-driving related circumstances. This action is executed through substantial public campaigns, and also by active and regular corrective measures to discourage the road users from indulging in such dreadful act.

Table 61: Liquor/drug use [evidentiary tested] estimates in road accident occurrence

Total estimates of Liquor/Drug Use *[evidentiary tested]* in the RTA occurrence in 2012		
Months	**Yes**	**No**
Jan	0	162
Feb	3	266
Mar	5	221
Apr	1	191
May	3	181
Jun	1	181
Jul	0	176
Aug	1	268
Sep	5	219
Oct	3	274
Nov	2	239
Dec	1	155
Total scores	**25**	**2533**

From the table above, a group of scores generated from the analysis of the two variables considered in confirming the evidentiary test of liquor/drug use is presented. However, in the chart below, a small estimate of 1.0% declared that minimal number of drivers/cyclists were tested positive to liquor/drug use, as disclosed by the results collated from the application of the Alcohol Breathalysers.

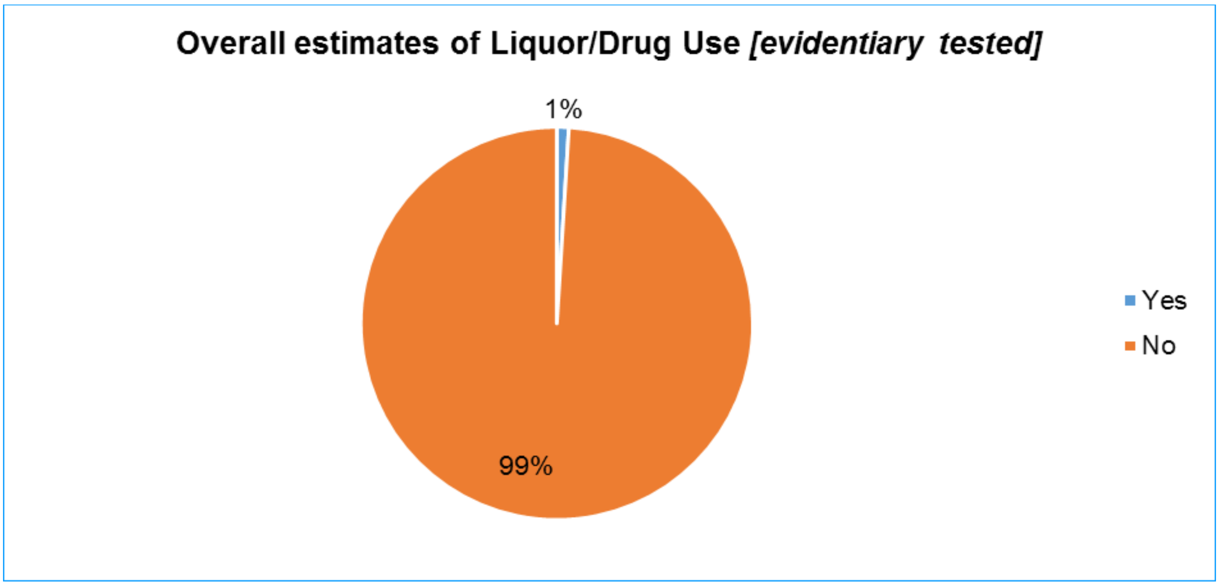

Figure 76: Total estimates of the Liquor/drug use *[evidentiary tested]* in 2012

In addition, a large estimated proportion of 99.0% road accidents were not connected to use of liquor/drug by the drivers/cyclists in Stellenbosch. Quintessentially, this estimate demonstrates that huge number of drivers/cyclists were not evidentially found guilty of liquor/drug use during an incident inspection.

D.1.5 Analysis of the Vehicle manoeuvre [what driver was doing]

This section presents a practical understanding into the data analysed based on the actions engaged by the drivers before the accidents. This particular field consist of twenty related variables, which are attributed to the actions involved by the drivers when the accident occurred. The table below contains set of scores produced from the analysis of the *Vehicle manoeuvre*. The understanding of the results generated reflects a practical level of the data analysed in this particular field, with regard to the actions engaged by the drivers earlier before the accident occurred.

According to the estimated scores presented in the table below, observably, an estimated score of 1,423 [42.2%], demonstrates that huge number of road accidents occurred mostly when the drivers/cyclists are travelling straight. Although, a nominal proportion of 0.1% indicates that lowest number of road accidents were attributed to merging of two roads into one-way, in which the impatience of the road users is being tested.

Some proportions of the results presented in the chart below are attributed to traffic rules and regulations violation by the drivers/cyclists, such as traffic lights encroachment, road signs encroachment, over speeding etc (Sinclair & Murdoch 2012). In addition, among other related

Table 62: Vehicle manoeuvre/*what driver was doing* estimates in road accident occurrence

Total estimates of Vehicle Manoeuvre/what driver was doing in RTA occurrence in 2012																				
Months	**Turning Right**	**Turning Left**	**U-turn**	**Enter Traffic Flow**	**Merging**	**Diverging**	**Overtaking [pass to right]**	**Overtaking [pass to left]**	**Travelling Straight**	**Reversing**	**Sudden Start**	**Sudden Stop**	**Busy Parking**	**Changing Lane**	**Swerving**	**Slowing Down**	**Avoiding Object**	**Stationary [waiting in traffic]**	**Parked**	**Other**
Jan	23	12	3	2	0	1	3	2	84	22	0	2	2	2	2	7	0	25	23	6
Feb	38	21	0	5	0	1	3	0	145	27	5	10	2	4	5	6	4	37	26	0
Mar	42	13	4	2	1	6	4	9	148	23	2	6	5	3	3	4	3	31	33	2
Apr	33	10	1	4	1	2	0	2	117	26	3	13	2	6	3	3	0	22	21	2
May	38	15	3	3	1	2	4	1	131	28	4	5	0	4	2	2	1	43	19	1
Jun	29	7	1	3	0	1	4	0	105	24	0	10	0	2	4	1	1	28	12	1
Jul	19	15	3	2	0	2	1	1	91	24	2	3	1	0	4	1	1	21	18	6
Aug	38	11	3	4	0	1	3	0	136	27	1	7	2	3	4	8	3	46	21	0
Sep	36	17	3	5	0	2	7	4	112	26	0	9	1	3	5	3	4	24	21	6
Oct	45	16	1	3	0	3	5	0	138	32	2	3	2	1	3	2	1	36	34	3
Nov	41	13	3	3	0	0	7	2	126	22	1	8	3	2	4	5	3	48	24	4
Dec	21	6	2	2	0	2	1	0	90	15	4	4	0	1	2	1	2	16	20	0
Total scores	**403**	**156**	**27**	**38**	**3**	**23**	**42**	**21**	**1423**	**296**	**24**	**80**	**20**	**31**	**41**	**43**	**23**	**377**	**272**	**31**

variables like *Turning right, Turning left, Reversing, Stationary-awaiting in traffic* and *Parked,* contributed a huge estimated proportion respectively to the degree at which accidents occurred in the Stellenbosch locality.

Statistically, an estimated proportion of 11.9% road accidents resulted through the action engaged while one of the vehicles involved was turning to the right side of the road. However, other actions that led to road accidents, such as *Stationary-awaiting in traffic, Reversing, Parked* and *Turning left* contributed estimated proportions of 11.2%, 8.8%, 8.1% and 4.6% respectively. Other remaining variables contributed a sum proportion of 13.2%, which is approximately three times smaller than the estimated score presented in the categorical variable, *Travelling straight*.

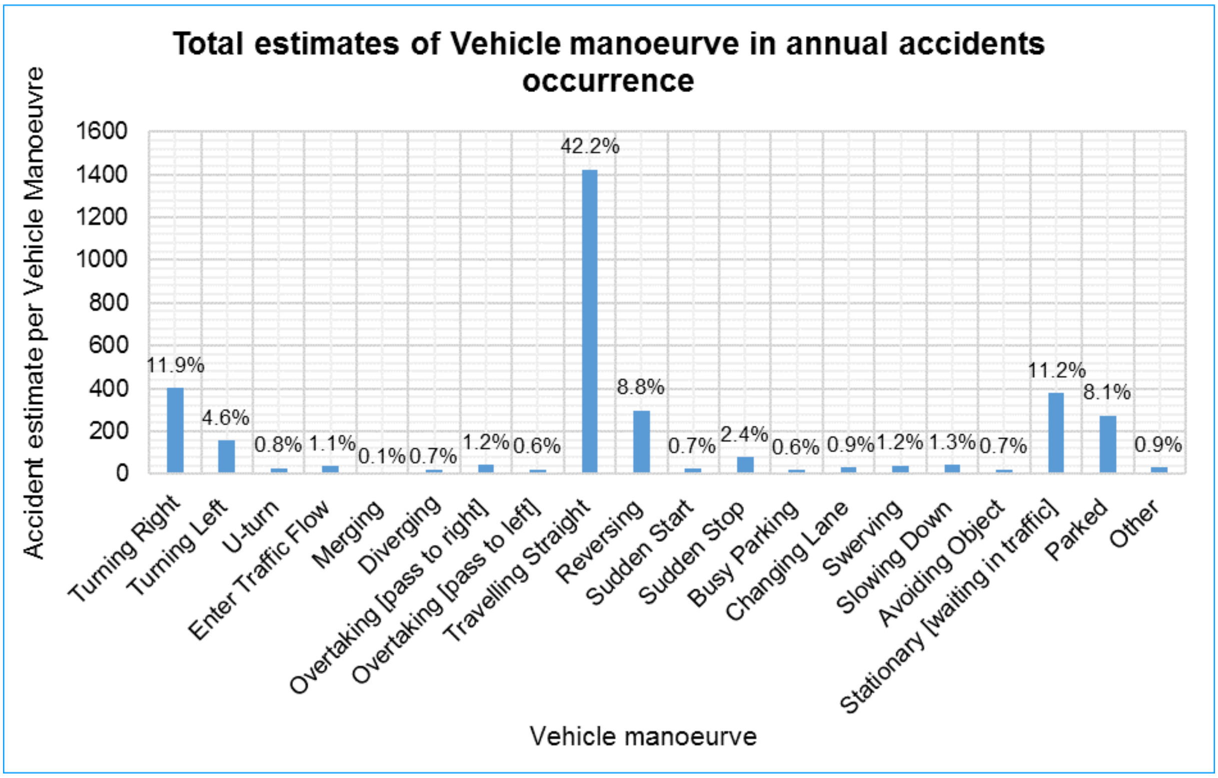

Figure 77: Total estimates of the Vehicle manoeuvre/*what driver was doing* in 2012

Understandably, the practicality of the results discussed in this section supports the understanding of the illustrations presented in the section 4.2.4 above and Appendix C-C.2.1. The results acquired in the variables like '*Turning right'* and *'Turning left'* can be applied to support the results obtained in the other related variables such as Head/rear end, Sideswipe-*opposite directions* and Sideswipe-*same direction.* On the other hand, accident cases like *'Reversing'* can lead to accident types like Head/rear end, while *'Merging'* and *'Travelling straight'* can also lead to Sideswipe-*same direction* and Approach at angle-*both travelling straight* respectively.

D.2 Benchmark table for drivers' Age classifications

Table 63: Comparison of drivers' involvement in the road traffic accidents by Age classifications between Stellenbosch Municipality and the City of Cape Town Metropolitan Municipality

Road accidents occurrence by driver's Age classifications					
Stellenbosch in 2012			City of Cape Town in 2005[40]		
Age groups	Total scores	Percentage estimates	Age groups	Total scores	Percentage estimates
14-20	323	9.4%	0-15	47	0.0%
21-40	1859	54.4%	16-17	218	0.0%
41-60	984	28.8%	18-21	7061	5.0%
61-80	238	7.0%	22-25	10878	9.0%
81-100	16	0.5%	26-35	29451	25.0%
			36-45	25671	22.0%
			46-55	16212	13.0%
			56-65	8328	6.0%
			66-75	3436	2.0%
			>75	1138	1.0%

Appendix E: Procedures adopted in developing the investigation survey

E.1 Questionnaire development

In this section, the purpose of developing a questionnaire, as part of the methodological approach adopted in this study is discussed. Fundamentally, the motive of developing or designing the questionnaire reflects the realism of the investigation carried out at the STD, with regard to the practical evaluation of the information acquired, through the application of the ARF. Previously in the Appendix A, a synopsis discussion on the ARF was presented to clarify the significance of the data collection tool in the processing of RTI by the two local authorities.

Similarly, as part of this discussion, some other important contexts surrounding the integration of the questionnaire into the investigation carried out is discussed, such agendas as:

- *consideration of relevant participants [respondents],*
- *validation of the questionnaire,*
- *completion of the questionnaire, and*
- *analysis of the questionnaire data.*

[40] Data extracted from a report published by the Traffic Accident Statistics (2005) on road traffic accidents by driver's Age classifications.

These contexts are represented in the flowchart illustrated in the Figure 78 below, which presents a simple understanding of the processes considered in the development of the questionnaire.

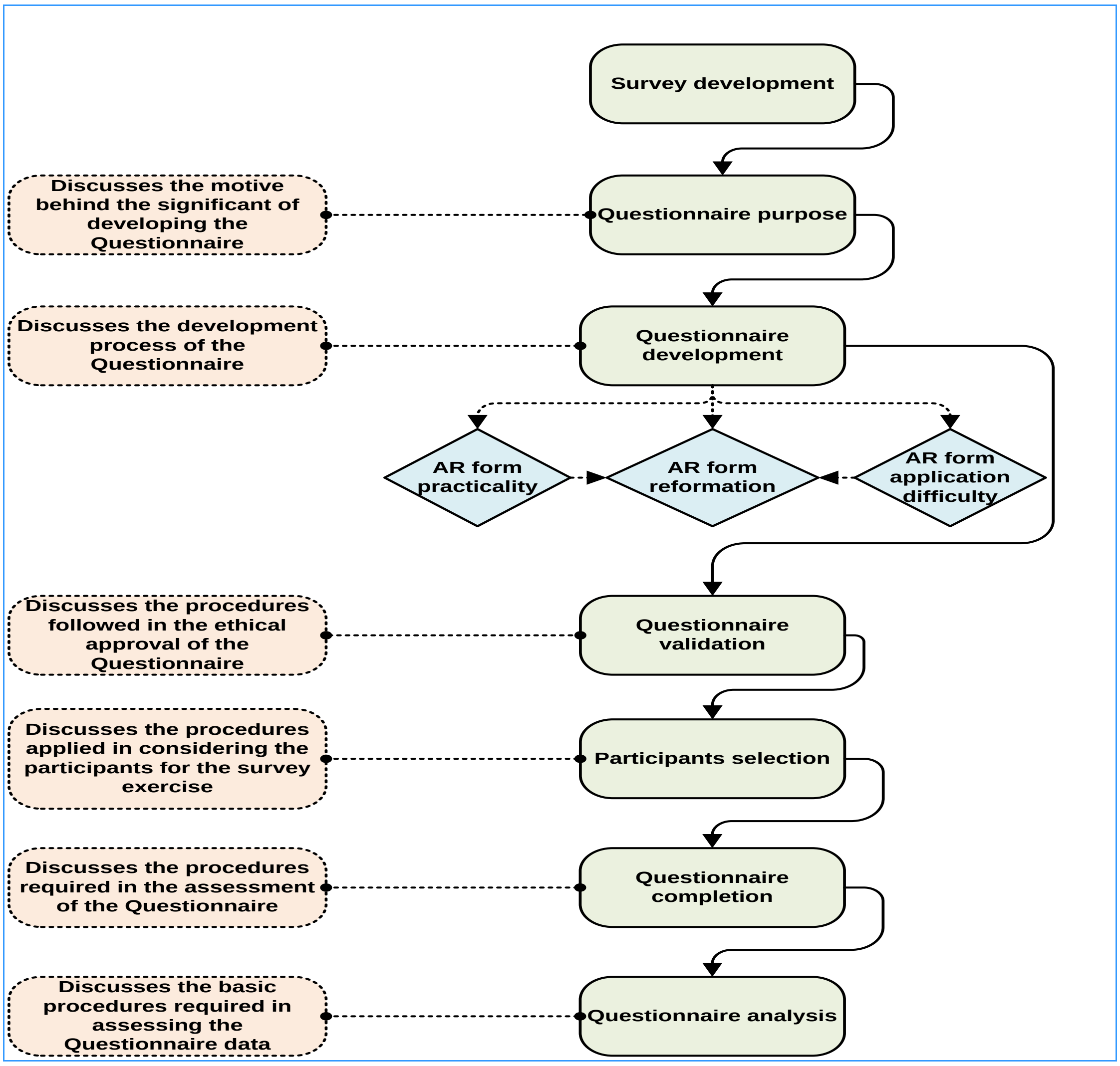

Figure 78: The process flowchart for the questionnaire development

E.2 Purpose of the questionnaire

However, as part of this investigation, the application of the ARF necessitates the need to seek the opinions of its users. Due to this, a questionnaire covering relevant survey questions pertaining to poor collection of road accident data was designed and implemented. Ultimately, there is a crucial need for investigating the degree of applicability of the data form. This substantiates the process of developing a standard questionnaire to promote the significance and practicality of the investigation.

In essence, the purpose of developing the questionnaire is to validate the result of the analysis of the data gathered from the STD. Nonetheless, the feasibility of the questionnaire will crystallise the analytical relationship between the road accident data collected and the application of the ARF. The questions structured include areas that are considered as valuable leads to the problem of inadequate or low-quality data affecting the road accident data.

E.3 Questionnaire development

The development of the questionnaire is introduced, with the objective of determining the areas concerning the application of the ARF in collecting road accident data. The survey questions provided in the questionnaire are classified into three classes, namely:

- **Practicality of the ARF** -this particular part illustrates survey questions, with relevant answers reflecting the capacities or potentials of the ARF users, in conjunction with their level of understanding towards the application of the form. However, this part will reflect the perceptions of the form users, with regard to the correct understanding of the information provided in the form. Additionally, this actualises the procedural approach of the form users towards the right application of the form.
- **Difficulty encountered while applying the ARF** -in this part, survey questions were developed in accordance with the information provided in the ARF, with the aim of determining the parts in the form that the users find most difficulty during data collection. If some answers can be provided for this set of questions, then, the problems associating with the increase in the errors committed can be uncovered.
- **Need for ARF reform** -this aspect covers the structuring of the survey questions that seek for the personal opinions of the form users regarding the standard level of ARF. This part also stimulates the valuable contributions of the form users, with regard to the restructuring of ARF through the suggested opinions provided.

The survey questions provided in the questionnaire were categorised into four separate sections. Nonetheless, the four sections entail the arrangement of the aforementioned classes of survey questions. Each survey question is unified with option boxes, allocated with numerical code to enhance the ability of structuring the answers marked into metadata for appropriate evaluation of the data collected. Also, additional spaces were provided for personal opinions of the form users. The questionnaire is designed to provide answers to some technical questions structured in connection with the opinions of the ARF users *[participants].*

E.4 Questionnaire validation

The procedures followed in validating the standard of the questionnaire involved a committee of validators, including the Researcher *[investigator],* Researcher's Supervisor, University Ethical Standard, Stellenbosch Traffic Department [STD] and Provincial Police Department [SAPS]. As illustrated in the Figure 79 below, the validation procedure is inaugurated by the researcher, who develops the questionnaire according to its purpose of utilisation. The researcher submitted a copy of the questionnaire to his or her supervisor for a preliminary evaluation process, in order to determine if the survey questions presented or provided in the questionnaire are relevant to the investigation [research] performed. Instantaneously, after the preliminary process has been completed, an evaluated copy of the questionnaire is submitted to the University Ethical Standard committee for ethical verification and approval, with the aim of ascertaining the standard of the questionnaire.

This particular process determines if the questionnaire contains any irrelevant questions or statements that may portray nepotism, criticism and indiscretion. The researcher's supervisor was included as part of the panel reviewing the standard level of the questionnaire. The panel reviewed the degree of risk of the questionnaire, whether it is a low risk, medium risk and high risk. At this level, a positive outcome of the validation necessitates the approval of the questionnaire by the University Ethical Standard committee. Once the panel approves the questionnaire, printed copies will be sent to the STD for approval, and extra copies would be sent to the Local Police Department [SAPS] for similar purposes.

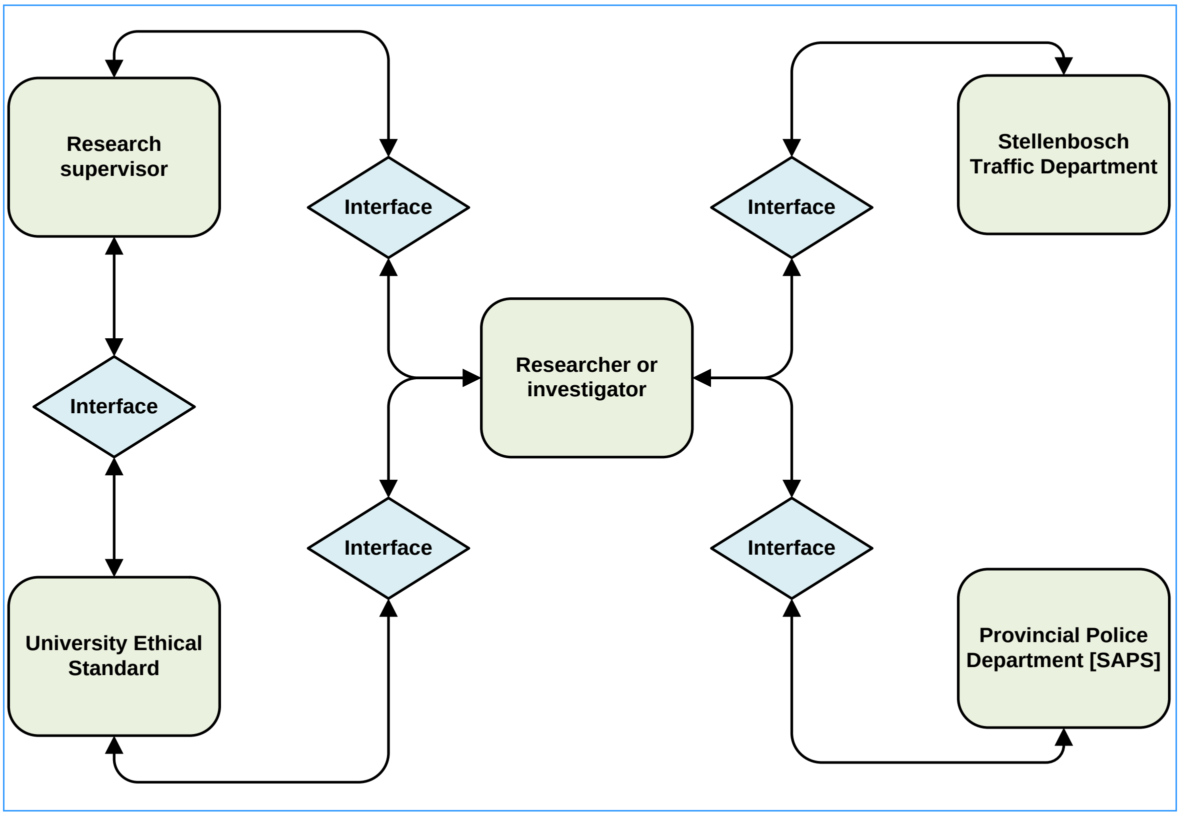

Figure 79: Procedures implemented in validating the standard of the questionnaire developed

The two departments validated the questionnaire according to the policy enacted, which required the submission of the application or request letter to guarantee the acceptance of the questionnaire as part of the data collection tools. Ultimately, after the approval of the exercise by the two departments, thereafter, a certain sample size of the questionnaire was distributed to the two departments for completion. After a comprehensive completion of the questionnaire, then, the researcher collates the completed forms for data grouping and analysis.

Although, the limitation experienced in this case, is the delay in getting the necessary approvals from the two local authorities in time. For this reason, the procedure taken in seeking approval for this survey exercise at the Provincial Police Department [SAPS] was time consuming. In the process, the local police department [SAPS] in the Stellenbosch area, referred the approval of the questionnaire to the Provincial Police Department in Cape Town, because the department lack the jurisdiction to approve such application.

Actually, this process deterred the time required to complete the research within the specified period. This limitation is due to the strict policies exercised towards the possibility of securing approval for the survey exercise. However, the *interface* indicated in the diagram above, refers to the communication relationship established between the institutions involved in the process over the approval of the survey exercise.

E.5 Selected participants

The criteria adopted in selecting the right participant for survey exercise are:

- Accident reporting officer who attended an accident scene previously or recently.
- Accident reporting officer who participated in the completion of ARF previously or recently.

Considerably, the selected participants for the exercise are the form users, who were in one way or another participated in the completion of ARF during any accident events, within the Stellenbosch Municipality [SM-WC024]. The participants are briefed about the procedures required to complete the questionnaire, through the instructions provided on the cover page of the form. Though, photocopies of the ARF were distributed along with the forms, in order to guide the participant through the process of understanding the survey questions in the form.

E.6 Questionnaire completion

The completed copies of the questionnaires will be collated by the researcher, in an attempt to avoid any misplacement or loss of any completed copy of the questionnaire form. Moreover, the completed questionnaires were evaluated by applying similar process implemented in the subsection 3.3.2 above. Answer options marked correctly in the questionnaires will be appropriately assessed, assembled, and analysed. The incorrect answers/responses will be classified as voids to avoid muddling up of analysis.

E.7 Questionnaire analysis

The analysis of the questionnaire is based on similar assessment procedures implemented in the subsection 3.3.2 above. The answers/responses options will be collected, structured and assembled by using a structured data extraction template. However, questions presented in the questionnaire are generalised, in order to have a complete understanding of the challenges frustrating the ability of some reporting officers to acquire sufficient and right data for RTA analysis.

E.8 Questionnaire completed by the STD and SAPS

Questionnaire

Measuring the application of the Accident Report form

Questionnaire set up by:	Oluwatimilehin .O. Okeowo Industrial Engineering Master's Student Stellenbosch University
Email Address:	16798597@sun.ac.za
Approved by:	University Ethical Committee
Date:	18/09/2015
Supervisor:	Theuns Dirkse van Schalkwyk [MEng] Senior Lecturer/Consultant Department of Industrial Engineering Stellenbosch University
Email:	theuns@sun.ac.za
Website:	www.ie.sun.ac.za

Purpose of the Questionnaire

This survey is mainly developed to substantiate the significance and feasibility of the research project currently being undertaken at the Stellenbosch Traffic Department regarding the improvement of accident data collated by the accident reporting officers, and the need for the restructuring of the AR form, which serves as one of the data collection tools. The questions provided in the questionnaire encompass areas that are considered as leads to the anticipating solution to the problem of inadequate or low-quality data affecting the traffic accident data disseminated to the data users in the area of decision-making.

However, according to recent findings acquired through the traffic accident data warehoused at the Accident Response Unit in the Stellenbosch Traffic Department, it is evidently revealed that certain information provided in the AR form is not accurately reported. Some information is not well represented, wherein the relevant information meant to enhance adequate data quality is omitted. Aside from this, additional errors are committed while completing the AR form which are due to other reasons as might be shown in the questionnaire answers.

Consequently, in order to validate the result of the analysis of the data warehoused at the Accident Response Unit in the Stellenbosch Traffic Department, valuable opinions of the traffic accident reporting officers, and that of the traffic police officers are required by completing the questionnaire to establish a crystallised analytical relationship between traffic accident data collated and the application of the AR form.

Information Privacy

The information provided in the questionnaire is developed according to the ethical code of the University of Stellenbosch guiding procedures required in performing any research task. The code of ethics for research at the university prevent researchers from asking the respondent questions that may reflect nepotism, misconduct, and in anyway expose the affiliated institution's privacy. In essence, the privacy of the respondent and the affiliated institution will be treated with utmost sincerity and respect.

Instructions

1. The respondent is instructed to indicate his/her answers by marking the boxes with two cross lines as shown here ☒.
2. The respondent is advised NOT to mark more than one option, where it is instructed to mark only one.
3. An opinion section is provided to the respondent at the end of Section III for the respondent's personal opinions to be made.
4. For easy understanding and verifying purposes, a copy of AR form is attached to the questionnaire to guide the respondent throughout the process of completing the questionnaire.
5. The respondent is urged to mark accurate options relevant to the information provided in the questionnaire.

Section I

- Are you a Traffic Officer/Police Officer/Peace Officer? *(mark ONE only)*
  1. ☐Traffic Officer 2. ☐Police Officer
  3. ☐Peace Officer 0. ☐None of the options
- What sex/gender are you? *(mark ONE only)*
  1. ☐Male 2. ☐Female 0. ☐Prefer not to disclose
- Among the age-group provided below, where does your age falls? *(mark ONE only)*
  1. ☐18-25 2. ☐26-30
  3. ☐31-35 4. ☐36-40
  5. ☐41-45 6. ☐46-50
  7. ☐51-55 8. ☐56 and older
  0. ☐Prefer not to disclose
- What grade level did you complete in school? *(mark ONE only)*
  1. ☐Grade 8 2. ☐Grade 9
  3. ☐Grade 10 4. ☐Grade 11
  5. ☐Grade 12 6. ☐Graduate
  0. ☐Prefer not to disclose
- Do you intend to further your education? *(mark ONE only)*
  1. ☐Yes 2. ☐No 0. ☐Not sure
- How long have you been working for the Traffic Department/Police Department? *(mark ONE only)*
  1. ☐1-5 months 2. ☐6 months-1 year
  3. ☐2-3 years 4. ☐4-5 years
  5. ☐6-7 years 6. ☐8-9 years
  7. ☐10 years and more 0. ☐Prefer not to disclose
- How regularly do you participate in inspecting or reporting a road traffic accident? *(mark ONE only)*
  1. ☐On no occasion 2. ☐Yes, occasionally (monthly)
  3. ☐Yes, often (weekly) 4. ☐Yes, very often (daily)
- If yes, when was the last time you participated in inspecting or reporting a road traffic incident? *(mark ONE only)*
  1. ☐1- 5 days ago 2. ☐1- 4 weeks ago
  3. ☐1- 4 months ago 4. ☐More than 5 months ago
- During the inspection or reporting process, do you fill-in the AR form? *(mark ONE only)*
  1. ☐Yes 2. ☐No 3. ☐Occasionally
- Do you in any circumstances find the AR form challenging to understand or interpret during any incident reporting? *(mark ONE only)*
  1. ☐Strongly disagree 2. ☐Slightly disagree 3. ☐Neither agree, Nor disagree
  4. ☐Slightly agree 5. ☐Strongly agree

- While filling in the AR form, do you suspect that you might commit any errors? *(mark ONE only)*

  1. ☐Yes 2. ☐No 0. ☐Not sure

- How often do you suspect that you commit these errors when fill-in the AR form? *(mark ONE only)*

  1. ☐Never 2. ☐1 – 3 errors3. ☐4 – 8 errors
  4. ☐9 – 15 errors 5. ☐More than 15 errors

- Do you think the arrangement of the information in the AR form causes or contributes to the errors made while fill-in in the AR form? *(mark ONE only)*

  1. ☐Strongly disagree 2. ☐Slightly disagree 3. ☐Neither agree, Nor disagree
  4. ☐Slightly agree 5. ☐Strongly agree

- Do you think the arrangement of the information in the AR form requires restructuring? *(mark ONE only)*

  1. ☐Strongly disagree 2. ☐Slightly disagree 3. ☐Neither agree, Nor disagree
  4. ☐Slightly agree 5. ☐Strongly agree

- Do you undertake any relevant training course on how to use the AR form in any accident reporting cases? *(mark ONE only)*

  1. ☐Yes 2. ☐No 0. ☐Not sure

- If yes, for how long is the training scheduled? *(mark ONE only)*

  1. ☐a few days 2. ☐1 - 3 weeks 3. ☐1 - 2 months
  4. ☐3 - 4 months 5. ☐5 - 6 months 6. ☐7 months and more

- During this training, were you provided a manual guide or instruction manual that guides you on the procedural use of AR form during the training? *(mark ONE only)*

  1. ☐Yes 2. ☐No 0. ☐Not sure

- How long does it take you to fill-in AR forms for accident involving a **Single** vehicle? *(mark ONE only)*

  1. ☐1– 5mins 2. ☐6– 15mins 3. ☐16– 30mins 4. ☐31– 45mins 5. ☐46mins & more

- How long does it take you to fill-in AR forms for accident involving **Multiple** vehicles? *(mark ONE only)*

  1. ☐1– 5mins 2. ☐6– 15mins 3. ☐16– 30mins 4. ☐31– 45mins 5. ☐46mins & more

- How long does it take you to fill-in AR forms for accident involving a **Single** vehicle and a **Pedestrian**? *(mark ONE only)*

  1. ☐1– 5mins 2. ☐6– 15mins 3. ☐16– 30mins 4. ☐31– 45mins 5. ☐46mins & more

- How challenging is it to find the details of the accident that is required by the AR form, during the day-time? *(mark ONE only)*

  1. ☐Very challenging 2. ☐Challenging
  3. ☐Easy 4. ☐Very easy

- How challenging is it to find the details of the accident that is required by the AR form, during the night-time? *(mark ONE only)*

  1. ☐Very challenging 2. ☐Challenging
  3. ☐Easy 4. ☐Very easy

- How challenging is it to find the details of the accident that is required by the AR form, during a rainy period? *(mark ONE only)*

  1. ☐Very challenging 2. ☐Challenging
  3. ☐Easy 4. ☐Very easy

- Do you in any circumstances prefer that the accident victim (driver or pedestrian), provided that s/he is in a safe condition, fill-in the AR form on your behalf while at the accident scene? *(mark ONE only)*

  1. ☐Yes 2. ☐No 0. ☐ Prefer not to disclose

- If yes, which of the factors provided below influence your decision. *(may mark MORE than one)*

  1. ☐Difficulties in finding the details of the accident as required by the AR form

2. ☐Difficulties in acquiring an accurate Accident Sketch and Accident Description
3. ☐Difficulties in acquiring Passengers/Pedestrians' information
4. ☐Difficulties in acquiring Driver's particulars
5. ☐Inadequate training regarding the application of the AR form

- In most cases, some important details are omitted while fill-in the AR form. Do you think that the factors outlined below could be responsible for the omissions, or inaccurate fill-in of the AR form? *(may mark MORE than one)*

1. ☐Inability to acquire the relevant details at the time of the incident
2. ☐Difficulties in the understanding of the information requested in the AR form
3. ☐Lack of interest, commitment and fatigue while fill-in the AR form
4. ☐Failure to identify the right information on the AR form concerning the accident
5. ☐Delay in the reporting process of the accident
6. ☐Lack of cooperation from the accident victims
7. ☐Lack of sufficient information on the AR form

Section II (Refer to your Accident Report form)

- Do you find it difficult to obtain information regarding any of these primary **ACCIDENT** details such as: Accident Date, Day of Week, Number of Vehicles involved, and Time of Accident? *(mark ONE only)*

1. ☐Yes 2. ☐No 0. ☐Not sure

- If yes, mark the exact area in the options provided below. *(may mark MORE than one)*

1. ☐Accident Date 2. ☐Day of Week
3. ☐Number of Vehicles involved 4. ☐Time of Accident

- Do you find it difficult to obtain information regarding the Accident **LOCATION**? *(mark ONE only)*

1. ☐Yes 2. ☐No 0. ☐Not sure

- If yes, mark the exact area in the options provided below. *(may mark MORE than one)*

1. ☐Built-Up Area 2. ☐Speed Limit on Road
3. ☐Province 4. ☐Street/Road Name/Road Number & Nodes
5. ☐Town/City 6. ☐Freeway/Rural
7. ☐GPS Reading [*X & Y Coordinates*]

- Do you find it difficult to obtain information regarding the **DRIVER'S** particulars? *(mark ONE only)*

1. ☐Yes 2. ☐No 0. ☐Not sure

- If yes, mark the exact area by marking the options provided below. *(may mark MORE than one)*

1. ☐ID Type/ID Number/Age 2. ☐Country of Origin of ID
3. ☐Race 4. ☐Gender
5. ☐Driving/Learner Licence 6. ☐Driving/Learner Licence Code
7. ☐Severity of Injury 8. ☐Seatbelt Fitted/Helmet Present
9. ☐Seatbelt/Helmet definitely used 10. ☐Liquor/Drug use suspected
11. ☐Liquor/Drug Use: evidentiary tested 12. ☐Passengers/Pedestrians Presence

- Do you find it difficult to obtain information regarding the **ROAD** details? *(mark ONE only)*

1. ☐Yes 2. ☐No 0. ☐Not sure

- If yes, mark the exact area in the options provided below. *(may mark MORE than one)*

1. ☐Road Type 2. ☐Junction Type
3. ☐Road Surface Type 4. ☐Quality of Road Surface
5. ☐Road Surface 6. ☐Road Marking Visibility
7. ☐Obstructions 8. ☐Overtaking Control

- Do you find it difficult to obtain information regarding the **LIGHT** Condition? *(mark ONE only)*

1. ☐Yes 2. ☐No 0. ☐Not sure

- Do you find it difficult to obtain information regarding the **WEATHER CONDITIONS** & **VISIBILITY**? *(mark ONE only)*

1. ☐Yes 2. ☐No 0. ☐Not sure

- Do you find it difficult to obtain information regarding the **ROAD SIGNS**? *(mark ONE only)*

1. ☐Yes 2. ☐No 0. ☐Not sure

- If yes, mark the exact area by marking the options provided below. *(may mark MORE than one)*

1. ☐Traffic Control 2. ☐Road Signs Clearly Visible
3. ☐Condition of Road Signs 4. ☐Direction of Road
5. ☐Flat or Sloped

- Do you find it difficult to obtain information regarding the **VEHICLE** details? *(mark ONE only)*

1. ☐Yes 2. ☐No 0. ☐Not sure

- If yes, mark the exact area in the options provided below. *(may mark MORE than one)*

1. ☐Vehicle Type 2. ☐Vehicle Position before Accident
3. ☐Vehicle Manoeuvre/Driver Action 4. ☐Vehicle Damage

- Do you find it difficult to obtain information regarding the **ACCIDENT TYPE**? *(mark ONE only)*

1. ☐Yes 2. ☐No 0. ☐Not sure

- Do you find it difficult to obtain information regarding the **HIT** & **RUN**? *(mark ONE only)*

1. ☐Yes 2. ☐No 0. ☐Not sure

- Do you find it difficult to obtain information regarding the **SUMMARY** of **PERSONS INVOLVED**? *(mark ONE only)*

1. ☐Yes 2. ☐No 0. ☐Not sure

- Do you find it difficult to obtain information regarding the **PEDESTRIANS/PASSENGERS'** particulars? *(mark ONE only)*

1. ☐Yes 2. ☐No 0. ☐Not sure

- If yes, mark the exact area in the options provided below. *(may mark MORE than one)*

1. ☐ID Type/ID Number/Age 2. ☐Country of Origin of ID
3. ☐Race 4. ☐Gender
5. ☐Driving/Learner Licence 6. ☐Driving/Learner Licence Code
7. ☐Severity of Injury 8. ☐Seatbelt Fitted/Helmet Present
9. ☐Seatbelt/Helmet definitely used 10. ☐Liquor/Drug use suspected
11. ☐Liquor/Drug Use: evidentiary tested

- Do you find it difficult to obtain information regarding the **PEDESTRIANS** and **CYCLISTS ONLY**? *(mark ONE only)*

1. ☐Yes 2. ☐No 0. ☐Not sure

- If yes, mark the exact area by marking the options provided below. *(may mark MORE than one)*

1. ☐Position 2. ☐Location
3. ☐Manoeuvre 4. ☐Pedestrian Action
5. ☐Colour of Clothing

- Do you find it difficult to obtain information regarding the **SPECIAL OBSERVATIONS**? *(mark ONE only)*

1. ☐Yes 2. ☐No 0. ☐Not sure

- If yes, mark the exact area in the options provided below. *(may mark MORE than one)*

1. ☐Tyre Burst 2. ☐Lights
3. ☐Reflector Quality 4. ☐Chevron

- Do you find it difficult to obtain information regarding the **DANGEROUS GOODS ONLY**? *(mark ONE only)*

  1. ☐Yes 2. ☐No 0. ☐Not sure

- If yes, mark the exact area by marking the options provided below. *(May mark MORE than one)*

  1. ☐Dangerous Good 2. ☐Spillage
  3. ☐Vapour/Gas Emission 4. ☐Dangerous Goods Place

Section III (Accident Sketch and Accident Description)

This section comprises questions regarding the accident sketch and accident description. The questions in each subsection are provided with the relevant options to support the understanding of the questions, and where necessary for the respondent to provide his/her opinion regarding the questions.

Accident Sketch

- Do you undergo any special training regarding the procedures involved in drawing the accident sketch details of the accident? *(mark ONE only)*

  1. ☐Yes 2. ☐No 0. ☐Not sure

- If yes, did the training given help you to understand the procedures involved in drawing the sketch details of the accident? *(mark ONE only)*

  1. ☐Yes 2. ☐No 0. ☐Not sure

- Have you ever experienced any circumstance(s) affecting your ability to produce a good technical sketch details of the accident? *(mark ONE only)*

  1. ☐Yes 2. ☐No 0. ☐Not sure

- If yes, give the reason(s) to support your response below. *(write them in the spaces provided below)*

  1. ________________
  2. ________________
  3. ________________

- Do you think it is difficult to obtain information regarding the features of an accident sketch as instructed in the AR form? *(mark ONE only)*

  1. ☐Yes 2. ☐No 0. ☐Not sure

- If yes, give the reason(s) to support your response below. *(write them in the spaces provided below)*

  1. ________________
  2. ________________
  3. ________________

Accident Description

- Do you undergo any special training regarding the procedures involved in writing the description? *(mark ONE only)*

  1. ☐Yes 2. ☐No 0. ☐Not sure

- If yes, did the training given help in understanding the procedures involved in writing the description of the accident?

  1. ☐Yes 2. ☐No 0. ☐Not sure

- Do you prefer the accident victim (driver/pedestrian), provided s/he is in a safe condition, to write the description on how the accident happened? *(mark ONE only)*

  1. ☐Yes 2. ☐No 0. ☐Not sure

- If yes, give reason for your decision for preferring the accident victim to write the description. *(write them in the spaces provided below)*

  1. ________________
  2. ________________
  3. ________________

- Have you in any circumstances find it difficult to obtain an accurate description of the accident? *(mark ONE only)*

1. ☐Yes 2. ☐No 0. ☐Not sure

- If yes, give the reason(s) to support your response below. *(write them in the spaces provided below)*

  1. ______
  2. ______
  3. ______

- Do you think the language barrier contributes to the problems of obtaining an accurate description of the accident? *(mark ONE only)*

  1. ☐Yes 2. ☐No 0. ☐Not sure

Opinions Section

- What other relevant information do you think should be included in the AR form? *(write them in the spaces provided below)*

  1. ______
  2. ______
  3. ______
  4. ______
  5. ______
  6. ______

- What information do you think is not necessary in the AR form? *(write them in the spaces provided below)*

  1. ______
  2. ______
  3. ______
  4. ______
  5. ______
  6. ______

- What other circumstances do you think could be contributing to the problems of acquiring complete information regarding the use of AR form? *(write them in the spaces provided below)*

  1. ______
  2. ______
  3. ______
  4. ______
  5. ______
  6. ______

- What other improvements would you recommend? *(write them in the spaces provided below)*

  1. ______
  2. ______
  3. ______
  4. ______
  5. ______
  6. ______

Any suggestions *(write your suggestions in the spaces provided below)*

F.9 Final report submitted by Stellenbosch Traffic Department

STELLENBOSCH

STELLENBOSCH • PNIEL • FRANSCHHOEK

MUNISIPALITEIT • UMASIPALA • MUNICIPALITY

INNOVATION CAPITAL

Ref: J. Waldis (021-808 8811)
Traffic Services

15 June 2016

Stellenbosch University
Faculty of Engineering
Private Bag XI
Matieland
Stellenbosch
7602
Fax: +27 21 808 3822

Dear Sir/Madam

RE: STUDENT - OLUWATIMILEHIN OLUWASEUN OKEOWO

Above mentioned student currently completing his Master's Degree in the Department of Industrial Engineering, opted to analyze accident data captured by Stellenbosch Traffic Services.

Topic: Investigating quality of data and the need for the restructuring of accident report form in South Africa.

I have proofread the thesis with minor changes to the structures and reporting lines within the department as well as in South Africa. I am pleased with the manner in which the analysis was done and incorporated throughout the document.

The proposals have been discussed with various authorities responsible for road safety, where after a request was received to further evaluate the findings.

Kind regards

J.K.WALDIS
TRAFFIC MANAGER