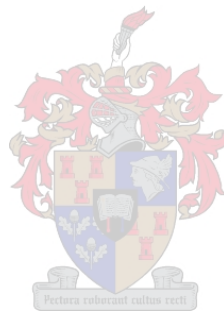


A data analysis demonstrator for managing customer experience in a partnering venture

by

Maryke Roos



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

Supervisor: Prof JF Bekker

April 2019

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.

Date: April 2019

Acknowledgements

I would like to give thanks to the one who made this entire process possible. He is the Light in my life, my Rock and my Saviour. As Psalm 118:29 says: ‘Give thanks to the Lord for he is good, his love endures forever.’

I would like to take this opportunity to thank the following people that the Lord has placed in my life and who have helped me. They have all contributed in their own unique way on this journey and I am grateful to them.

- Prof James Bekker, my supervisor, for all his support and guidance throughout this whole project.
- All my colleagues from the USMA research group for all their support. I would like to give special thanks to Marisa Walters and Zandaline Els for all their support, love and friendship.
- Ms. Anne Erikson, for editing my thesis and for also being willing to help me till the end.
- My parents, Tom and Janneke Roos for all their support, love and encouragement to further my studies.
- My boyfriend, Juan de Bruijne, for all his love and support and for understanding when my masters sometimes got higher priority.
- My sisters, Karike Roos and Rebakka Uys, for their friendship throughout the years and for encouraging me at each step in life.
- My brother, Arie Roos, for his love and support even when I do not always stay in touch.
- My two best friends, Alicia Potgieter and Anika van den Bout, for all their love, support and encouragement even when they did not always understand what I was talking about.

I would like to end off with the following bible verse: ‘Trust in the Lord with all your heart, and do not lean on your own understanding. In all your ways acknowledge Him, and He will make straight your paths.’ (Proverbs 3:5 – 6).

Abstract

Today's world can be better described as a digital world in which technology is becoming increasingly more dominant. In fact, the technologies of today are changing the biological, digital and physical worlds. This change has a huge impact on industry and how it operates and buzz words that have come forward during these times are Artificial Intelligence, Machine Learning and Big Data Analytics. In light of this the following question arises: 'How do we look after our customers by using these changes and technologies to our advantage?'

To investigate this question, a research study is conducted in which data analytics along with business partnering on a cross-functional platform are used to manage and improve a customer's experience. This is achieved by developing a capability demonstrator that will simulate customer activities on a full-scale platform in which data is captured and analysed.

The focus is placed on the domain of travel, in which a customer undertakes a journey which is solely planned, booked and managed by the demonstrator. While the customer travels, they engage with various collaborating enterprises and the demonstrator focuses on managing and improving the customer experience at various interaction points.

After the implementation of this demonstrator, evaluation has been done to ensure the demonstrator can effectively manage and improve customer experience by the use of data analytics in a partnering venture. Further analysis in the form of Machine Learning was applied on the trips simulated by the demonstrator. The analysis gave valuable insights into customer behaviour in the travel domain.

Opsomming

Die wêreld van vandag kan beter beskryf word as 'n digitale wêreld waarin tegnologie al hoe meer dominant word. Trouens, die tegnologieë van vandag verander die biologiese, digitale en fisiese wêreld. Hierdie verandering het 'n groot impak op die industrie en hoe die ondernemings funksioneer, en terme wat deesdae gebruik word is Kunsmatige Intelligensie, Masjienleer en Groot Data Analise. In die lig hiervan kan die volgende vraag gestel word: 'Hoe sien ons om na ons kliënte om sodoende hierdie verandering en tegnologie tot ons voordeel te gebruik?'

Om hierdie vraag te ondersoek, is 'n navorsingstudie gedoen waarin data analise saam met vennootskappe op 'n kruis-funksionele platform gebruik word om 'n kliënt se ervaring te bestuur en te verbeter. Dit word behaal deur 'n vermoë-demonstreerder te ontwikkel wat kliënte se aktiwiteite op 'n volskaalse platform sal simuleer waarin data geskep en ontleed word.

Die fokus is op die domein van reis, waarin 'n kliënt 'n reis neem wat uitsluitlik beplan, bespreek en bestuur word deur die demonstreerder. Terwyl die kliënt reis, het hulle interaksies met verskeie samewerkende ondernemings en die demonstreerder fokus op die bestuur en verbetering van 'n kliënt se ervaring op hierdie interaksie-punte.

Na die implementering van hierdie demonstreerder is evaluering gedoen om te verseker dat die demonstreerder 'n kliënt se ervaring effektief kan bestuur en verbeter deur die gebruik van data-analise in 'n vennootskap. Verdere analise in die vorm van masjienleer is toegepas op die reise wat deur die demonstreerder gesimuleer is. Die analise het insig gelewer oor die kliënt se gedrag in die domein van reis.

Contents

Nomenclature	xiii
1 Research proposal	1
1.1 Research background	1
1.2 Research statement	3
1.3 Research objectives	4
1.4 Research scope	4
1.5 Deliverable envisaged	5
1.6 Research methodology	5
1.7 Structure of thesis	7
1.8 Conclusion: Research proposal	7
2 Customer Experience and the management of it	9
2.1 Classification of customers	9
2.2 Customer Experience	12
2.2.1 Definition of Customer Experience	13
2.2.2 Understand the Customer Experience	16
2.2.3 How to measure Customer Experience	17
2.2.4 Importance of Customer Experience	20
2.3 Customer Experience Management	21
2.3.1 Customer Relationship Management	22
2.3.1.1 What is Customer Relationship Management	22
2.3.1.2 How to do Customer Relationship Management	23
2.3.2 Customer Experience Management	27
2.3.2.1 Definitions of Customer Experience Management	27
2.3.2.2 Why Customer Experience Management?	29
2.3.2.3 How to do Customer Experience Management	30
2.3.2.4 Challenges with implementing Customer Experience Management	34
2.3.3 Customer Relationship Management versus Customer Experience Management	35
2.4 The customer journey	36
2.5 360-Degree view of the customer	45
2.6 Conclusion: Customer experience and the management of it	48

3	Big Data and its analytics	49
3.1	Big Data	49
3.1.1	Overview of Big Data	49
3.1.2	Importance of Big Data	52
3.1.3	Challenges of Big Data	53
3.2	Big Data Analytics	58
3.2.1	Overview of Big Data Analytics	58
3.2.2	Big Data Analytics methodology	60
3.2.2.1	Framework for the methodology	60
3.2.2.2	Knowledge Management processes	62
3.2.2.3	Machine Learning	64
3.3	Conclusion: Big Data and its Analytics	72
4	Business partnering on a cross-functional platform	73
4.1	Business partnering	73
4.1.1	Overview of business partnering	73
4.1.2	Purpose of partnering	76
4.1.3	Enabler for business partnering	78
4.1.4	A business partnering example	83
4.2	Business partnering on a cross-functional platform	84
4.2.1	Overview of platforms	84
4.2.2	The cross-functional platform	86
4.3	The literature study integration	89
4.4	Conclusion: Business partnering on a cross-functional platform	91
5	The development of the Trip Planner Demonstrator	92
5.1	The Trip Planner Demonstrator	92
5.2	The system architecture for the Trip Planner Demonstrator	95
5.3	The Trip Planner Demonstrator database	99
5.3.1	Accommodation entities	103
5.3.2	Long Distance Transportation entities	105
5.3.3	Short Distance Transportation entities	106
5.3.4	Customer entities	110
5.3.5	Transaction entities	112
5.3.6	Other entities	113
5.4	The simulator for Trip Planner Demonstrator	114
5.4.1	Customer access system process	117

5.4.2	Trip planning process	118
5.4.2.1	Phase 1: Book LDT enterprise	121
5.4.2.2	Phase 2: Book accommodation enterprise	124
5.4.2.3	Phase 3: Book SDT enterprise	126
5.4.3	Customer travelling process	130
5.5	Conclusion: The development of the Trip Planner Demonstrator	135
6	Verification and evaluation of the Trip Planner Demonstrator	137
6.1	Introduction: Verification and evaluation	137
6.2	Verification of Trip Planner Demonstrator	138
6.3	Evaluation of Trip Planner Demonstrator	138
6.3.1	Evaluation: Customers travelling	140
6.3.2	Evaluation: Booking process	142
6.3.3	Evaluation: Customer Experience ratings	144
6.3.4	Evaluation: Changes due to bad ratings	145
6.4	Conclusion: Verification and evaluation of Trip Planner Demonstrator	151
7	Analysis of Trip Planner Demonstrator Data	152
7.1	Roadmap of analysis followed	152
7.2	Choosing Machine Learning tool and techniques	152
7.3	Preprocessing and transformation of data	153
7.3.1	Principal Component Analysis	154
7.3.2	Recency, Frequency and Monetary Analysis	156
7.4	Application of Machine Learning techniques	158
7.5	Insights and Knowledge	160
7.5.1	Analysis of customer behaviour	162
7.5.1.1	Analysis of accommodation customer behaviour	164
7.5.1.2	Analysis of LDT customer behaviour	165
7.5.1.3	Analysis of SDT customer behaviour	166
7.5.2	Analysis of Customer Experience	169
7.5.2.1	Analysis of overall Customer Experience	170
7.5.2.2	Analysis of individual Customer Experience	173
7.5.3	Analysis of transactions	182
7.6	Conclusion: Analysis of trips	186

CONTENTS

8 Conclusion and recommendations	188
8.1 Research summary	188
8.2 Self-assessment of work	190
8.3 Recommendations for future work	191
8.4 Conclusion	192
References	193
A Simulation of Trip Planner Demonstrator database	218
B Analysis of Trip Planner Demonstrator : Detail results	235

List of Figures

1.1	Research methodology	5
2.1	The wheat-to-flour and bread supply chain structure	11
2.2	The telco supply chain model	12
2.3	The world's oldest-known customer complaint	12
2.4	The delivery gap	15
2.5	Net Promoter Score	18
2.6	The CRM continuum	23
2.7	CRM model	24
2.8	The three dimensions of CRM	24
2.9	People dimensions: Layers of role players	26
2.10	Market readiness for CEM	29
2.11	Six factors for the customer journey	37
2.12	Outline of the customer journey model	39
2.13	Customer journey example	44
3.1	Frequency distribution of documents containing the term 'Big Data'	50
3.2	DIKW hierachy	59
3.3	Big Data Analytics	61
3.4	The evolution of knowledge management processes	64
4.1	Types of business partnering	75
4.2	Purpose of business partnering	77
4.3	Balancing act for business partnering	80
4.4	Properties of business partnering types	82
4.5	Platform effect	87
4.6	Business partnering on a cross-functional platform	88
4.7	BDA together with CEM	90
5.1	Components of the TPD	93
5.2	Overview of the actions of the TPD simulator	95
5.3	Object-Process diagram for the TPD architecture	97
5.4	Object-Process language for the TPD architecture	98
5.5	EERD for the TPD part A	101
5.6	EERD for the TPD part B	102

LIST OF FIGURES

5.7	Steps in a simulation study	115
5.8	The simulator concept model	116
5.9	Beta distributions for trip requirements	118
5.10	The trip planning process	119
5.11	Two cases of customer behaviour	120
5.12	The customer travelling process	130
6.1	Verification and evaluation	137
6.2	Histogram of frequency of trips taken per customer	140
6.3	Histogram of total trips taken	141
6.4	Three examples of customer journeys	150
7.1	Scatter plot representing the different dimensions of the data attributes	155
7.2	Weighted PCA applied to the customer accommodation dataset	156
7.3	RFM analysis for transactions	157
7.4	RFM analysis for overall CX	158
7.5	Silhouette plot for CX of accommodation	159
7.6	Silhouette plot for accommodation	160
7.7	Schematic for analysis of TPD data	161
7.8	k-Means clustering plot of customer behaviour	163
7.9	Pie charts for clustering of customer behaviour	164
7.10	Histogram for clusters of customer behaviour – Accommodation	165
7.11	Histogram for clusters of customer behaviour – LDT	166
7.12	Histogram for clusters of customer behaviour – SDT	167
7.13	Pie chart of clusters after RFM analysis	170
7.14	Clustered RFM analysis of overall CX	171
7.15	RFM values for overall CX clusters	172
7.16	k-Means clustering plot of CX	174
7.17	Pie charts for clustering of CX	175
7.18	Histogram for clusters of CX – Accommodation	176
7.19	Histogram for clusters of CX – LDT	177
7.20	Histogram for clusters of CX – SDT	178
7.21	Histogram for clusters of CX – Transactions	179
7.22	Pie chart of transaction clusters after RFM analysis	183
7.23	Clustered RFM Analysis of customer transactions	183
7.24	RFM plot for transactions of cluster 1	184
7.25	RFM plot for transactions of cluster 3	185

LIST OF FIGURES

B.1	Weighted PCA plot of customer LDT behaviour	236
B.2	Weighted PCA plot of customer SDT behaviour	236
B.3	Weighted PCA plot of CX – Accommodation	237
B.4	Weighted PCA plot of CX – LDT	237
B.5	Weighted PCA plot of CX – SDT	238
B.6	Weighted PCA plot of CX – Transactions	238
B.7	Silhouette for customer LDT behaviour	239
B.8	Silhouette customer SDT behaviour	239
B.9	Silhouette plot for CX – LDT	240
B.10	Silhouette for CX – SDT	240
B.11	Silhouette plot for transaction	241
B.12	Silhouette plot for CX – Transactions	241
B.13	Silhouette plot for overall CX	242

List of Tables

2.1	Historical perspective: Contributions to Customer Experience	14
2.2	CEM theme classification	32
2.3	Models and frameworks for the implementation of CEM	33
2.4	Comparison of CEM and CRM	36
2.5	Description of the journey phases	40
3.1	Value creation of BD - Quantity of articles	53
3.2	Value Creation of BD - The articles	54
3.3	Summary of BD challenges	55
3.4	Summary of BDA techniques	66
3.5	Classification techniques for BDA	68
3.6	Clustering techniques for BDA	69
3.7	Regression techniques for BDA	71
5.1	Object-Process diagram legend	96
5.2	Extended Entity-Relationship Diagram legend	100
5.3	Description of Accommodation entity	104
5.4	Description of booked Accommodation entity	104
5.5	Description of booked LDT entity	105
5.6	Description of LDT entity	106
5.7	Description of booked SDT entity	108
5.8	Description of SDT entity	109
5.9	Description of customer entity	110
5.10	Rate of CX entered	133
5.11	The λ values for the customer experience rating distribution	135
6.1	A summary of the total trips taken	142
6.2	Accuracy of booking process	143
6.3	Accuracy of CX rating	145
6.4	Changes due to CX Rating	145
6.5	Changes in booking of trips due to bad CX ratings	148
7.1	Dimensions of data attributes	155
7.2	A summary of the analysis done	161
7.3	Clusters for customer behaviour	168

LIST OF TABLES

7.4	RFM clusters detail for overall CX	173
7.5	Clusters for CX	180
7.6	RFM values for transactions	184
7.7	Customer's demographics of cluster 1	185
7.8	Customer's demographics of cluster 3	185
7.9	RFM clusters detail for transactions of customer	186
A.1	Distribution of accommodation type and rate	218
A.2	Customer table of first 10 entries	219
A.3	Summary of customer ages	219
A.4	Summary of customer Accommodation preferences	220
A.5	Summary of customer LDT preferences	221
A.6	Summary of customer loyalty programs	222
A.7	Summary of customer SDT preferences	222
A.8	Summary of product shop	225
A.9	Summary of SDT area link	226
A.10	Summary of stations and stores	227
A.11	Station information	228
A.12	Simulation of trip planner taxi rates	229
A.13	Areas considered for the TPD	230
A.14	Districts considered for the TPD	233
A.15	Provinces considered for the TPD	234
A.16	Area concatenated	234
B.1	Clusters of customer behaviour for accommodation	243
B.2	Clusters of customer behaviour for LDT	244
B.3	Clusters of customer behaviour for SDT	245
B.4	Clusters of CX for accommodation	246
B.5	Clusters of CX for LDT	247
B.6	Clusters of CX for SDT	248
B.7	Clusters of CX for transactions	249

Nomenclature

Acronyms

BD	Big Data
BDA	Big Data Analytics
BP	Business Partnering
CAC	Customer Acquisition Cost
CEM	Customer Experience Management
CES	Customer Effort Score
CIM	Customer Interaction Management
CKM	Customer Knowledge Management
CRISP	Cross-Industry Standard Process
CRM	Customer Relationship Management
CSAT	Customer Satisfaction
CX	Customer Experience
DIKW	Data-Information-Knowledge-Wisdom
EERD	Extended Entity-Relationship Diagram
ERD	Entity-Relationship Diagram
EU	European Union
GDPR	General Data Protection Regulation
IS model	Importance-satisfaction model
KDD	Knowledge Discovery in Databases
LDT	Long Distance Transportation
ML	Machine Learning
NPS	Net Promoter Score

NOMENCLATURE

OPD	Object-Process Diagram
OPL	Object-Process Language
OPM	Object-Process Methodology
PCA	Principal Component Analysis
POPI	Protection of Personal Information
RFM	Recency, Frequency, Monetary
SDT	Short Distance Transportation
SEMMA	Sample, Explore, Modify, Model and Assess
SERVQUAL	A multi-dimensional research instrument, designed to capture consumer expectations and perceptions of a service along the five dimensions that are believed to represent service quality.
SQM	Service Quality Management
TPD	Trip Planner Demonstrator
QFD model	Quality Function Deployment model

Chapter 1

Research proposal

The aim of the thesis is to conduct a study in which research and the application of engineering skills, tools and methods will be used to construct a capability *demonstrator*. The demonstrator will be the abstract concept of a digital system known as the *trip planner*. The study should give practical considerations to the industry partner based on how data analytics together with business partnering can be used to manage the experience of a customer.

The research proposal will give a background to the research study, the research statement, the objectives and the scope of the research study, the deliverable envisaged together with the contribution, the proposed research methodology and lastly the structure of the thesis document.

1.1 Research background

We live in a world known as the ‘digital world’, where technology is becoming more prominent and replacing the ‘old ways’ of doing things. There are many examples of how technology has taken over.

As a first example, paper is getting replaced by technology. More enterprises let their customers fill in forms online instead of using a paper form. There are also enterprises who give the customer the option whether they want to receive the receipt via SMS or email or if they want a printed receipt.

A second example is where flight tickets and entry tickets for sporting events, games and shows take the form of e-tickets and an increasing number of enterprises are adopting this into their businesses. If a passenger checks in for their flight, the boarding pass can then be viewed on their mobile device. When a student goes to watch a sports game, they can present the ticket on their mobile device instead of using a printed ticket.

A third example is when one drives on the freeway you can see feedback or updates on the billboards about traffic congestion, accidents, roadworks and more. These billboards give the opportunity for real-time updates.

A fourth example is that a customer can now order a wide range of products online, from bread and milk at Pick 'n Pay to a 110 inch Samsung TV from Takealot or even their dream king-size bed from Bed King.

The fact is that technology is evolving on a big scale and at a fast pace. It starts to control the lives of many people and people have access to many services at the touch of a button.

But the question remains: Are enterprises able to keep up with the demands of the so-called ‘digital world’ and are they still able to satisfy their customers’ needs? And how is the customer

1.1 Research background

looked after? Is the customer considered in these transformations? And if they are, to what extent does this occur?

Therefore, to keep up with the technological transformation of the world and customers' preferences, enterprises should shift from a business-centric to customer-centric view according to Rich (2015). To achieve this, enterprises should move away from the inside-out thinking approach to an outside-in thinking approach. In other words, they should see what the customers' needs are and how can they adjust to them while still bringing value to the market. When this aspect is understood, the enterprise should adjust their processes, systems and products or services accordingly. This will involve a total change in the current way of doing business, where enterprises only focus on their processes, systems and products/services and how to improve them they do not take the customers' preferences and needs into consideration.

Therefore, the purpose of this work is to conduct a study in which a *capability demonstrator* will be developed to show how customer experience can be managed by using data analytics and business partnering. The data analysis will be done based on using the customers' historical behavioural data and their preferences. The demonstrator will take the form of a digital information and decision-support system that is known as a *trip planner*. The problem to be solved is based on a case study.

The case study deals with how the proposed trip planner plans a trip for an individual. The individual will indicate where they want to go, what date they want to leave and when they want to return. The proposed trip planner will then book the trip and arrange for all activities required.

A fictional description of what typically happens is as follows.

Thandi lives in Cape Town and wants to attend her cousin's wedding just outside Durban this Saturday. She indicates to the trip planner that she wants to leave Cape Town on Thursday evening and needs to be back on Monday morning before work. The digital system will then book and pay for a return flight, transportation in and around Durban and accommodation outside Durban.

Thereafter the system will perform the following steps. On the Thursday prior to her departure, Thandi is informed of who will take her to the airport, what time she should be ready and what documentation she should have with her. When she arrives at the airport, she is already checked in and her boarding pass is available on her mobile device, therefore she only needs to drop off her bags. Thereafter she proceeds to the departure gate. On her way to the gate the system informs her about a special at her favourite coffee shop in the airport. She orders a coffee with a double chocolate muffin and pays for it by using a banking app.

Thandi finishes her coffee and chocolate muffin in the waiting area and then she boards the plane just in time for departure. After the aeroplane lands in Durban, the system informs Thandi that the Uber driver is waiting for her at the drop-and-go. Thandi rushes to the drop-and-go as she is tired. The Uber car then takes her to the guesthouse. On her way to the guesthouse, the owners are informed of her imminent arrival. They eventually receive her and take her to her room. She

1.2 Research statement

is quite impressed as the accommodation meets her requirements. After a long day Thandi goes to bed.

On the Friday morning she decides to go to do some sightseeing in Durban. Thandi uses *TripAdvisor* to plan her activities and notifies the system that she needs transportation to all her activities. The system book her transportation as the activities unfold and sends the required notifications to Thandi.

On the morning of the wedding, Thandi is awakened by the system that notifies her of who will pick her up in an hour's time. That gives her enough time to prepare and get ready for the wedding. The taxi arrives on time and she is dropped off at the venue. She enjoys her time with family and friends while attending the wedding. Later, Thandi notifies the system that she is ready to leave. Ten minutes later, the taxi arrives which gives her enough time to say her goodbyes.

On Sunday, Thandi decides to just spend time at the guesthouse and enjoy the relaxing atmosphere by the pool with her book in her hand. After a restful day, the system notifies her of the time her lift will pick her up the following day and that she should have a good night's of rest.

Early on the Monday morning, the taxi arrives in time to take her to the airport. Arriving at the airport, Thandi drop her bags off as she is already checked in and her boarding pass is again displayed on her mobile device. As Thandi approaches the boarding gate, the system informs her which breakfast places are open. She orders a quick breakfast with coffee and pays for it with her banking app. On the flight back home, Thandi takes a nap. When Thandi arrives at Cape Town airport, she is notified of who is waiting for her at the drop-and-go. She is taken directly to her work, due to a crash on the N2 which caused a delay. She did not therefore have time to stop at home.

The case study described a typical example of the functionality of the trip planner.

1.2 Research statement

The purpose of the study is to construct a *demonstrator* by using data analytics and business partnering to enable customer experience and the management of it. The demonstrator will be *implemented* via a *trip planner* which will record customer activities and experiences. The operations of the trip planner will be realised with a *simulator* which will simulate unique trips of many customers. The trip planner will use customers' historical behaviour and preferences to determine the best way to plan a complete trip for an individual to improve and manage that particular customer's experience.

1.3 Research objectives

The following objectives must be met:

1. *Conduct* thorough research to grasp the following concepts:
 - Customer Experience and Customer Experience Management.
 - Big Data and Big Data Analytics.
 - Business Partnering on a Cross-functional platform.
2. *Construct* an architecture and simulator that will form the basis of the Trip Planner Demonstrator.
3. *Simulate* data for the database of the Trip Planner Demonstrator.
4. *Simulate* unique trips for many customers.
5. *Analyse* customer journeys achieved by the Trip Planner Demonstrator.

1.4 Research scope

The *aim* of the study is that it should be shown how the *trip planner* can improve and control a customer's experience by using data analytics together with customer preferences and profiling. The assumptions have been made that the partnering venture is already set in place and that all legal issues have been taken care of. The system is built on top of a database, where the database integrates all entities provided by the business partners in the system. For the trip planner we assume that the following systems will have an impact on it:

- The *main form of transportation* to get the customer to their destination:
 - ◊ Aeroplane and
 - ◊ Bus.
- The type of *accommodation* preferred by the customer:
 - ◊ Backpackers Hostel,
 - ◊ Bed & Breakfast,
 - ◊ Guesthouse,
 - ◊ Hotel and
 - ◊ Self-catering units.

- The type of *transport* used at the customer's destination:
 - ◊ Car Rental,
 - ◊ Hailing App Taxi and
 - ◊ Normal Taxi.

1.5 Deliverable envisaged

The *outcome* of the study will show the industry partner, practical considerations when using business partnering on a cross-functional platform together with data analytics to improve and management customer experience.

The *contributions* of the study will be guidelines for Customer Experience Management on a cross-functional platform.

1.6 Research methodology

The study should show how the trip planner will be able to improve and control a customer's experience when they go on the trip. For the study to achieve this the following methodology will be used. The outline of the methodology can be seen in Figure 1.1.

The first step is to *formulate* the research problem. It is defined based on the research background, statement, objectives and scope. The detail of it can be seen in Sections 1.1 to 1.5.

The second step is to *conduct* a literature study which relates to objective 1. The purpose of the literature study is to understand how to design the demonstrator to fulfil the requirements and specifications of the trip planner. Therefore, the structure of the literature study will be as follows:

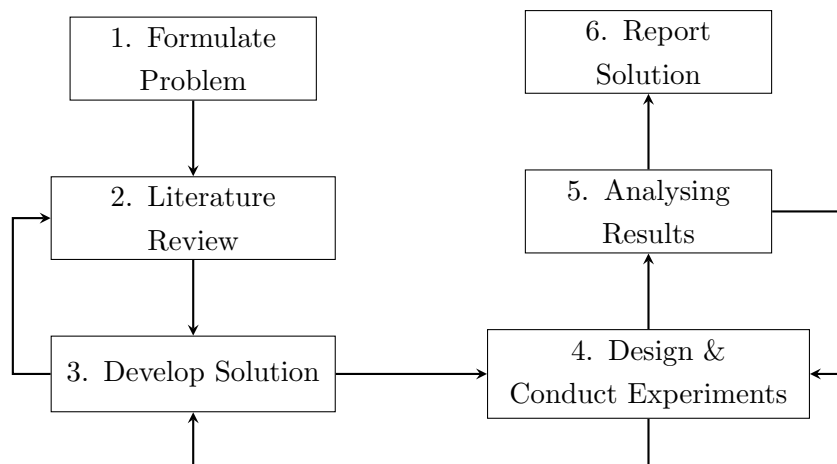


Figure 1.1: Research methodology

1. Customer Experience Management:

- What is customer experience and why is it important.
- How to manage customer experience.
- The relevance of the customer journey and how to construct it.
- How and what influences a customer.

2. Big Data Analytics:

- What is Big Data, the importance of it and issues around Big Data.
- Defining Big Data Analytics.
- The methodology of Big Data Analytics, with specific focus on Machine Learning.

3. Business partnering on a Cross-Functional Platform:

- What is the purpose of business partnering and what is needed to enable it.
- What are platforms.
- Determining how to link platforms and business partnering.

The third step is to *develop* a solution which relates to objective 2. The construction of the architecture and simulator for the demonstrator will be done based on the research conducted and requirements stated in the scope. During this step more aspects may be included to or excluded from the literature study as required.

The fourth step is to *design* and *conduct* experiments which relate to objectives 2 and 3. Data will be simulated according to the specifications and expectations of the industry partner. While data is simulated, changes can be made to the demonstrator in order for it to be able to handle the set of data. After input data has been simulated, the trip planner will simulate unique trips for many customers as set out by objective 4.

The fifth step is to *analyse* the results obtained from the implemented Trip Planner Demonstrator which relate to objective 5. In this step the customer journeys achieved from the trip planner will be analysed to determine whether they hold to be true to the purpose of the study and to determine the value of the Trip Planner Demonstrator. Changes might be made to the Trip Planner Demonstrator to obtain improved customer journeys or more data might be simulated for different results.

The final step is to *report* the obtained results and findings of the functionality of the Trip Planner Demonstrator.

1.7 Structure of thesis

The structure of the thesis document will be as follows:

Chapter 1: Research Proposal

The proposal for the study is given in this chapter and it includes the background of the research, the statement of research, what are the objectives, what is the scope of the study, what deliverables are envisaged and the research methodology that will be used.

Chapters 2 to 4: Literature Study

A literature study is given in this chapter based on the main research areas as mentioned in the research methodology. The literature study will cover a background of the topics and an in-depth focus on important aspects that are relevant to the study.

Chapter 5: Development of the Trip Planner Demonstrator

In this chapter the development and construction of the architecture and simulator for the Trip Planner Demonstrator will be explained based on how it was done. The methodology used for the development will also be incorporated as well as data requirements and objectives.

Chapter 6: Verification and Evaluation of the Trip Planner Demonstrator

In this chapter the verification and evaluation of the trip planner demonstrator will be provided. This will prove that the Trip Planner Demonstrator is built correctly and it replicates the right model.

Chapter 7: Analysis of Trip Planner Demonstrator Data

In this chapter, all the results obtained from the trip planner will be analysed and conclusions will be drawn from them. The purpose of the analysis is to determine the value of the Trip Planner Demonstrator.

Chapter 8: Conclusion of Study

A conclusion of the study is drawn up in this chapter. In this chapter, an explanation will be given on what was done in this study and whether it adheres to the scope and objectives as set out in Chapter 1.

1.8 Conclusion: Research proposal

The following aspects were discussed in this chapter. A background to the research was given based on why such a study is necessary and a description of the case study that will be used for the Trip Planner Demonstrator. After that the research statement was given based on what purpose of the study is. Then the objectives for the study were stated based on what will be met. The scope was then provided by explaining what the aim of the study is. After that, the deliverables envisaged were given as well as their contribution and a research methodology was provided as a guideline as

1.8 Conclusion: Research proposal

to how the study will be approached. The chapter then ended off with the structure of the thesis document which gave a clear view of how the study will be constructed and recorded.

In the next chapters a literature study will be done. The literature study will be done based on customer experience and the management of it, Big Data and the analytics of it and business partnering on a cross-functional platform.

Chapter 2

Customer Experience and the management of it

The research proposal was provided in Chapter 1 and the research methodology in Section 1.6. Due to the research methodology, a literature study needs to be conducted. In this chapter the first part of the literature study will be conducted on the Customer Experience together with Customer Experience Management. Research needs to be done on this to understand how the demonstrator should be constructed in order to know how a customer's experience can be managed by the trip planner.

Therefore, to understand the importance of Customer Experience and the Management of it, a literature study needs to be conducted based on the following aspects. The first aspect is to understand what a customer is. The second aspect is to investigate the Customer Experience, how it can be managed and why it is important. The third aspect is to understand how a Customer Experience can be management, why it should be managed and the challenges associated with it. The fourth aspect is to look at the customer journey and its significance to the Customer Experience. The last aspect is the 360-degree view of the customer. Although this topic of Customer Experience and the Management of it is more in the business environment, it is important to understand it to apply the right engineering tools and techniques.

2.1 Classification of customers

The first aspect of literature study is the Customer Experience and the Management of it. To get down to the depth of it, an important question to answer is, what is a customer?

The [Oxford English Dictionary \(1409\)](#) has amongst other, the following definition for a customer: “A purchaser of goods or services. In early use: specifically a person who regularly purchases from a particular business.”. [Investopedia \(2017\)](#) further supports this definition by defining a customer as “an individual or business that purchases the goods or services produced by a business”. From these two definitions, Definition 2.1 has been created, where an enterprise can be seen as “the culmination and co-existence of the business model and organisation to create and deliver the products and services” ([du Preez et al., 2015](#)).

Definition 2.1 (Customers). *A customer is any individual, group or business that buys and uses a product and/or services from an enterprise.*

2.1 Classification of customers

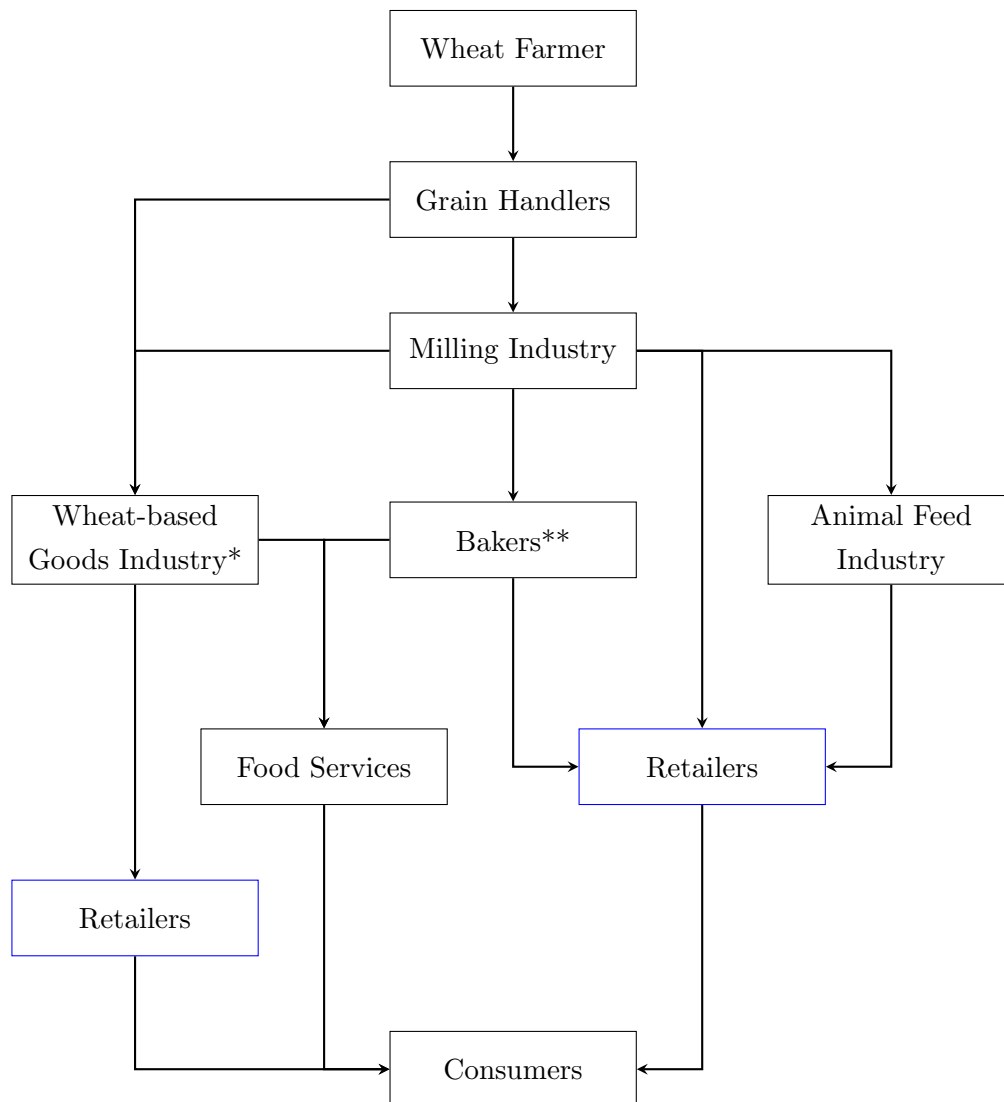
But to further understand what a customer is, it is important to know the different roles that a customer can take on. A good way to represent it is along a supply chain. The *wheat-to-flour and bread* supply chain will be used as an example. This supply chain is presented in Figure 2.1 .

Based on the *wheat-to flour and bread* supply chain, the grain handler is a customer for the wheat farmers as the grain handler buys wheat from the farmers to supply that to the milling industry and the wheat-based goods industry. These last two industries are therefore customers of the grain handlers. The milling industry will then supply meal, bran and flour to their customers who represent the wheat-based goods industry, bakers, the animal feed industry and the retailers. The retailers are therefore a customer of the milling industry, wheat-based goods industry, bakers and the animal feed industry. All the wheat-related products they buy from these suppliers, are then sold to their customers who are the consumers. The same goes for the food services, where they are the customers of the wheat-based goods industry and the bakers. After they buy their wheat products from them, they supply and sell it to their own customers who are the consumers. The supply chain ends off with the consumers because the consumer is the one who would use the end / final product.

Another way to represent the different roles a customer can take on is by using the *telecommunication (telco) industry* as example, where a telecommunication service is provided, such as data and telephony communications. The supply chain model for the *telco industry* can be seen in Figure 2.2. Based on this supply chain, one can see there are three different role players. The first role player is the vendor who is the producer of the technology and physical products. The second role player is the operator who is responsible for installing a capacity to provide the telecommunication service. The third role player is the city who uses the service and therefore creates a demand for the capacity that needs to be installed. In other words, the vendor supplies the necessary equipment to the operators. Therefore, the operators are customers of the vendor. The operators then use this equipment in order to install the right capacity to serve the demand of the city which varies over time. Therefore, the individuals in the city are the customers of the operators. The end user in this supply are the individuals in the city who consume this telco service.

Based on the two examples of the different roles a customer can fulfil in a product and service industry it is clear that a customer can be any party that buys a product and/or service and uses it, but at every instance in the supply chain the customer fulfil a different role. For the purpose of this study, the term ‘customer’ will refer to the end user of a product and/or service. The end user will be a ‘human’ user that act as an individual. The end user will always have a certain type of experience when he or she uses a product and/or service. Therefore, in the next section a Customer Experience will be investigated to understand what it is, how can it be measured and why is it important.

2.1 Classification of customers



* Includes biscuits, pasta, crackers, breakfast and cereals.

** Includes breads, speciality breads, pan loaves, rolls/buns, confectionery products, frozen dough and par-baked products

Figure 2.1: The wheat-to-flour and bread supply chain structure (Adapted from Barling *et al.* (2009); Moloisane (2004))

2.2 Customer Experience

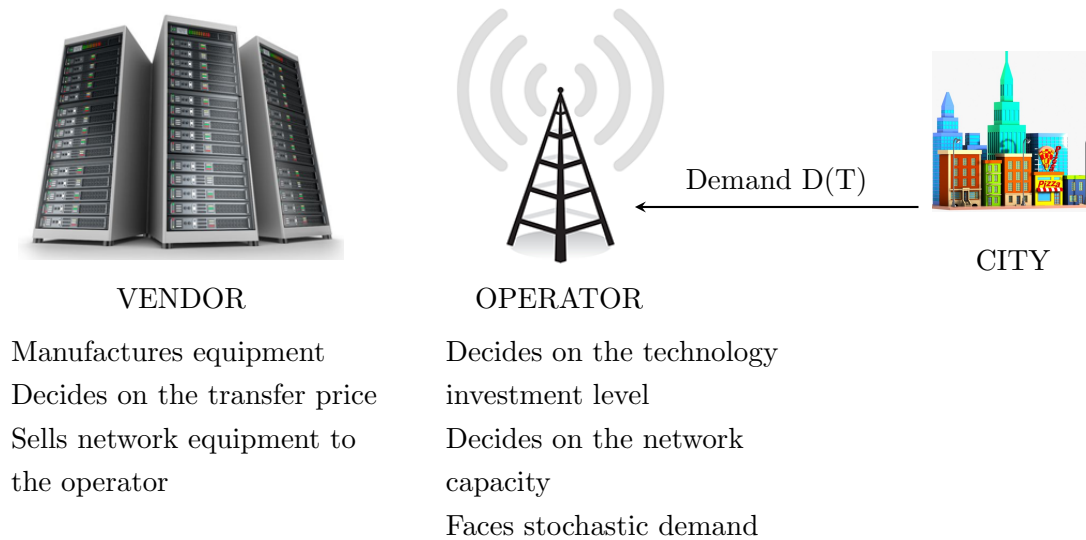


Figure 2.2: The telco supply chain model ([Çanakoğlu & Bilgiç, 2007](#))

2.2 Customer Experience

By understanding what a customer is and the different roles it can take on and by specifically considering the end user, the next question to answer is: what is meant by a ‘Customer Experience’?

The experience a customer has with a product and/or service has been present from the start of trading. The oldest recorded Customer Experience has been captured as a customer complaint. According to [Dodds \(2016\)](#), the oldest known customer complaint is in the British museum in London. The complaint is as follows: A customer, called Nanni, complained about the service he received for buying copper ingots from a seller, called Ea-nasir. Nanni was not happy with how Ea-nasir treated him and that the quality of the ingots was not as promised. He etched his complaint on a clay tablet. The recorded complaint can be seen in [Figure 2.3](#).



Figure 2.3: The world’s oldest-known customer complaint ([British Museum Blog, 2017](#))

2.2 Customer Experience

The world has evolved in such a way that customers' complaints can be recorded in various ways. Whether it is a letter sent to an enterprise or a social media post, customers can easily complain about the experience they had with a product and/or service or the enterprise itself. But, it is important to note that a Customer Experience is not necessarily only recorded by a customer complaint about a product and/or service or the enterprise itself. A customer complaint is only one way in which a customer can voice their opinion. It is not only limited to a specific enterprise that delivers an experience for a customer. It looks at the whole spectrum of enterprises who have an interaction with a customer and delivers an experience for that customer.

Therefore, to answer the question "What is meant by a Customer Experience?", the term will first be defined, then to how an enterprise should understand it, how it can be measured and the importance of it. These points will be discussed in the next subsections.

2.2.1 Definition of Customer Experience

Customer Experience (CX) might seem to be a new concept that has been making its existence over the past few years, but this is not true. This concept has been investigated for well over three decades, like the work done by [Holbrook & Hirschman \(1982\)](#), [Lebergott \(1993\)](#), [Pine & Gilmore \(1999\)](#), [Assury \(2002\)](#), and many more. [Verhoef et al. \(2009\)](#) have investigated the various themes from which CX has been studied from 1982 to 2008 and [Du Plessis & De Vries \(2016\)](#) have looked at the important work that was conducted in the CX domain from 1985 to 2015. The history of aspects that contribute to CX were defined by [Lemon & Verhoef \(2014\)](#) and it can be seen in [Table 2.1](#).

Therefore, there are various definitions and explanations in literature about CX, which shows that it can be defined from various angles. For the purpose of this study [Definition 2.2](#) will be used.

Definition 2.2 (Customer Experience). *Is the sum-total of experiences a customer has with a product and/or service. The experiences expands throughout the entire interaction a customer has with that particular product and/or service. That is from the first interaction until discontinued use ([Best et al., 2016](#); [Meyer & Schwager, 2007](#); [Richardson, 2010](#)).*

From this definition, CX can be analysed even further in order to understand the entire concept.

Table 2.1: Historical perspective: Contributions to Customer Experience Lemon & Verhoef (2014)

Time Frame	Topic Area	Contribution to CX	Representative Articles
1960s-1970s	Customer buying behaviour: process models	<ul style="list-style-type: none"> Encompasses path to purchase. Broad, experiential focus. Conceptual linkage models. Considered CX and customer decision-making as a process. 	Lawridge Robert & Steiner Gary (1961) Howard & Sheth (1969)
1970s	Customer satisfaction and loyalty	<ul style="list-style-type: none"> Identified key metrics to begin assessing overall CX. Empirical linkage models to identify key drivers. Assessed and evaluated customer perceptions and attitudes about an experience. 	Oliver (1980) Zeithaml (1988) Bolton & Drew (1991) Gupta & Zeithaml (2006)
1980s	Service Quality	<ul style="list-style-type: none"> Incorporated atmospherics and environment. Early journey mapping through blueprinting. Linked marketing and operations – focus on printing. Identified the specific context and elements of the CX. 	Parasuraman <i>et al.</i> (1988) Bitner (1990); Bitner (1992) Rust & Chung (2006) Bitner <i>et al.</i> (2008)
1990s	Relationship marketing	<ul style="list-style-type: none"> Expanded to B2B contexts. Identified key attitudinal drivers. Broadened the scope of customer responses considered in the CX. 	Dwyer <i>et al.</i> (1987) Morgan & Hunt (1994) Berry (1995)
2000s	Customer relationship management	<ul style="list-style-type: none"> Enable return-on-investment assessment. Identification of key touch points and drivers. Data-driven. Incorporate multi-channel aspects. Identified how specific elements of the CX influence each other and business outcomes. 	Reinartz & Kumar (2000); Verhoef (2003) Bolton <i>et al.</i> (2004); Reinartz <i>et al.</i> (2004) Rust <i>et al.</i> (2004); Payne & Fow (2005); Kumar & Reinartz (2006); Neslin <i>et al.</i> (2006); Kumar & Shah (2009)
2000s-2010s	Customer-centricity and customer focus	<ul style="list-style-type: none"> Customer perspective throughout organisation. Embedded the customer and customer data deeper into the organisation. Focused on redesigning CX from customer perspective. 	Sheth <i>et al.</i> (2000) Gulati & Oldroyd (2005) Shah <i>et al.</i> (2006)
2010s	Customer engagement	<ul style="list-style-type: none"> Recognised value of non-purchase interactions. Incorporated positive and negative attitudes, emotions, and behaviours. Conceptual platform to incorporate social media. More clearly recognised the customer's role in the experience. 	Libai <i>et al.</i> (2010); Van Doorn <i>et al.</i> (2010) Brodie <i>et al.</i> (2011) Kumar <i>et al.</i> (2010) Kumar <i>et al.</i> (2013) Hollebeek <i>et al.</i> (2014)

2.2 Customer Experience

It is important to note that an enterprise delivers an experience for their customers, whenever they interact with the customers in the target market. Whether the enterprise interacts with the customer via an advertising pamphlet, the selling of their product and/or service or the after-sale interaction. This contributes to the reason that it is very important for an enterprise to live up to their motto. Therefore, it is important to keep the following words of Abraham Lincoln in mind when dealing with customers: “*Actions speak louder than words*”. It is more valuable for an enterprise to deliver a good CX than the quality of their mottos (Dandridge, 2010). A phrase might attract a customer, but the CX will be reason whether a customer will stay or leave.

The problem is that what is expected by the customer is not necessarily delivered by the enterprise. This type of problem can be seen as a delivery gap, because there is a gap between what the enterprise believe they deliver and what the customer actually experienced. A survey done by Allen *et al.* (2005) delivered the following result: ‘When we recently surveyed 362 firms, we found that 80 percent believed they delivered a ‘superior experience’ to their customers. But when we then asked customers about their own perceptions, we heard a very different story. They said that only eight percent of firms were really delivering.’ This survey proves the fact that there is a delivery gap. Another way to look at this gap, as explained by Best *et al.* (2016), is that the delivery gap is the difference between the customer expectation and what the firm actually delivers. Figure 2.4 represents this delivery gap.

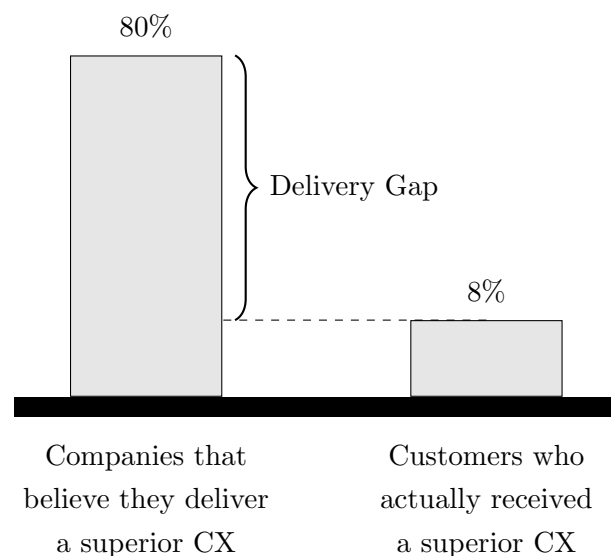


Figure 2.4: The delivery gap (Allen *et al.*, 2005)

For this delivery gap to be decreased, the enterprise should live up to what they promise they will deliver but also they should know how to deliver a superior CX. For an enterprise to know how to deliver a CX is not an easy task as all customers across the board have different expectations of what will be delivered and how it will be delivered.

According to [Payne & Frow \(2007\)](#) a customer will see the ‘perfect’ CX when he/she considers the experience of a product and/or service in the domain of the nature that they relate to and whether the experience can be achieved at an effective, competitive price. The word ‘perfect’ is placed in single quotes as the CX relates to a specific customer and how they experience the interaction with the product and/or service from their perspective.

2.2.2 Understand the Customer Experience

Now that CX has been defined, one must ask how does an enterprise come to the point where it will understand what the customer wants to experience?

First of all it is important to understand that the experience a customer has with the product and/or service is an emotional response, according to [Lehman \(2016\)](#). Therefore, an enterprise should be aware of their actions, because the smallest action can have a huge impact on how a customer perceives the experience. Therefore, an enterprise should think of how it wants to be treated if it were in the customer’s shoes.

Secondly, there is a link between what the customer actually experience and what they expect. According to [Dodds \(2016\)](#), an enterprise should consider eight aspects that are needed to integrate a customer expectation and experience. These eight aspects include:

1. *Value*: In terms of the price, the availability and functionality of the product and/or service. A customer has a certain expectation of the value he/she wants to receive from a product and/or service and the value of the product and/or service can determine how good (or bad) the CX will be.
2. *Help to create positive emotions*: It adds potential to extend sales beyond the initial purpose that the customer had and by doing that creating a positive emotion which will give more likelihood that the customer will buy the product and/or service again.
3. *Physical dynamics of the transaction*: It goes hand-in-hand with the saying that ‘the first impression is the last impression.’
4. *Promptness*: It is about how quick the enterprise is able to respond to their customers and how good their problem resolution is.
5. *Full transparency*: It refers to the information that should be readily available in a concise manner.
6. *Communication*: This is probably the most important aspect, as it is the only way in which an enterprise can truly connect with their customers and acts as a tool to understand their customers’ needs.

2.2 Customer Experience

7. *Post-sale customer service*: It is important to look after the customer, after the sale goes through as it forms part of the experience and the enterprise always needs to live up to the customer expectation.
8. *Encourage and direct customer loyalty*: To enhance customer loyalty, the enterprise should have at least the following elements:
 - Trustworthiness.
 - Dependable: ability to be counted or relied on.
 - Be responsive to customer needs.
 - Bring the right products and/or services to the market.

Lastly, when an enterprise examines the CX, it is important to do it over the *omnichannel* platform. [Best et al. \(2016\)](#) describe the omnichannel as a consistent and personalised experience of customers across all channels. In other words, the enterprise should be aware of which channels their customers use to interact with their products and/or services. The enterprise should deliver the same CX across all these channels. The omnichannel also helps the enterprise to get a better understanding on CX.

As a conclusion, CX is also dependent on the perception, emotions, unexpected behaviours and outside factors that will influence the customer. Customers are not robots, which makes CX a tricky concept. These aspects will be discussed in more depth in Section 2.5.

2.2.3 How to measure Customer Experience

How will an enterprise know when they deliver a good CX? Measurements are put in place to get a better understanding of the level of CX that the enterprise delivers. It is important to know that an appropriate tool should be used with the correct type of customer and type of outcome envisaged. The field known for measuring CX is *Customer Analytics*. This field will not be discussed in detail but the only focus will be on why does one measure CX and what measurements have been put in place for it.

The secret to measuring CX is to do it when a customer interacts with the product and/or service and enterprise. In other words, an enterprise should define definite points at which CX can be captured in the best way as it will give real-time data. Real-time data is important as it captures the CX as the world progresses. These points, known as touch points, are according to [Richardson \(2010\)](#), under the direct control of the enterprise, which in essence makes it easier for the enterprise to measure the CX.

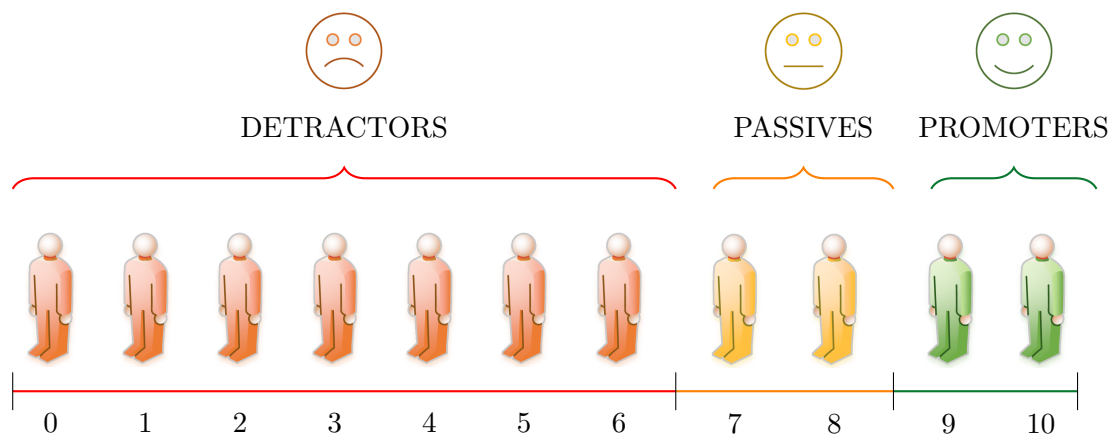
The next aspect to look at, is what metrics can be used to measure the CX? There are a lot of metrics to use on how to capture a customer satisfaction rating. The five most-used metrics are as follows ([Kierczak, 2017](#)):

2.2 Customer Experience

1. *Net Promoter Score* (NPS): The NPS is a tool that measures a customer's satisfaction and its link with business growth. In more simpler words, it determines how likely it is that a customer will recommend the enterprise. It basically classifies the customers on their rating of 'How likely will the customer recommend the enterprise to his/her family or friends?'. From here, three respondent groups are then classified, as seen in Figure 2.5.

- i. *Promoters*: The loyal customers, who will recommend the enterprise to anyone.
- ii. *Passive*: Customers who are satisfied with the enterprise but can switch to another enterprise due to competitive pressure.
- iii. *Detractors*: Customers who are unhappy with the enterprise and can cause damage to the enterprise.

The NPS is then determined by taking the Promoters and subtracting the Detractors from them. The score can range from 0 percent to 100 percent.



$$\text{NPS} = \% \text{ PROMOTERS} - \% \text{ DETRACTORS}$$

Figure 2.5: Net Promoter Score

2. *Customer Acquisition Cost* (CAC): The CAC measures the customer service quality together with the overall customer satisfaction. This metric is used to test an enterprise's acquisition channels and for customer segmentation. The CAC of a particular period can be calculated as follows: $CAC = (\text{Marketing Expenses}) / (\text{Number of customers gained})$.

3. *Customer Satisfaction* (CSAT): The CSAT is a tool that measures how satisfied a customer is with a specific experience they had with the enterprise, based on a score. It only captures specific instances of customer satisfaction and not overall customer satisfaction. The scores that the customer can give are based on a sliding scale for the level of satisfaction where they

2.2 Customer Experience

can be (i) very unsatisfied, (ii) unsatisfied, (iii) neutral, (iv) satisfied or (v) very satisfied. CSAT is then determined by calculating the overall rating by using the average from all the respondents.

4. *Customer Effort Score (CES)*: The CES is a tool that measures how much effort a customer had to put in to get the transaction completed. The less effort they put in, the higher the CES will be and the better customer satisfaction they had. They give a score based on the statement 'The organization made it easy for me to handle my issue', where the answers are on sliding scale of (i) strongly disagree, (ii) disagree, (iii) somewhat disagree, (iv) neutral, (v) somewhat agree, (vi) agree or (vii) strongly agree. CES is determined in the same way as CSAT.
5. *Customer churn rate*: This rate is the percentage of customers who discontinued their subscription to an enterprise either because they did not buy or use the enterprise's product and/or service or they cancelled a recurring product and/or service provided by the enterprise. The rate can be determined by $churn\ rate = (Total\ customers\ at\ end\ of\ period) / (Total\ customers\ over\ the\ same\ period)$.

CX can also be measured by looking at how customers interact on social media channels. Social media has a big influence on the world and a study by Perrin (2015) shows that 65 percent of adults are now using social network sites and that this has increased tenfold over the past decade. Therefore, social media can act as an active way of monitoring social media channels in order to gather information on customers' experiences with a product and/or service. Best *et al.* (2016) identified the following five metrics for monitoring CX:

1. *Consumer Activity Metrics*: This metric measures the activities the customer has actually done. An example is *social page views*, in other words how many customers viewed a specific page.
2. *Brand Reach Metrics*: This metric measures the audience connected to a brand. An example is *social connections*, in other words how many customers are connected on a specific page.
3. *Consumer Engagement Metrics*: This metric measures the actual impact of a communication on a consumer. An example is *engagement rate*, in other words, how many times did a customer engage with a specific page.
4. *Acquisition Metrics*: This metric measures the conversion of engagement into actually a favourable action towards the brand. An example is *visit duration*, in other words how long did a customer interact with a specific page.

5. *Conversion Metrics*: This metric measures the effectiveness of a social media activity to monetise CX. An example is *conversion rate*, it measures how effective a promotion is.

Other metrics have also been defined by [Anderson *et al.* \(2016\)](#), [Anderson \(2016\)](#), [Beard \(2014\)](#), [Lanoue \(2016\)](#), [Reznik \(2016\)](#) and [Zagorica \(2013\)](#).

Although the above-mentioned metrics are customer satisfaction metrics, it is important to note that there is a crucial difference between customer satisfaction and CX. One might believe that CX is just a fancy word for customer satisfaction but in reality it is not. There is a difference between the questions, “Are you satisfied with the product?” versus “How was the experience with the product?”. For example, let’s say a customer went on a roller coaster ride, the answer they might give for “Were you satisfied with the ride” might be way different compared to “How was your experience on the roller coaster?”.

Therefore, to successfully measure the CX, the customer satisfaction metrics can be used at every interaction point as well as asking the right questions at the right time. The recording of CX at every interaction point can then be recorded on the customer journey. The customer journey will be discussed in more depth in Section 2.4.

2.2.4 Importance of Customer Experience

A lot has been said on what CX is, how it can be measured and how an enterprise can understand what the customers’ expectations are. But why is it important to achieve the ‘perfect’ CX?

Firstly, the main functionality of achieving this CX, is that it will lead to an increase of profit for an enterprise and an increase in customer loyalty ([Payne & Frow, 2007](#)).

Secondly, another important aspect to keep in mind is, that the customer has more control today, due to the fact that technology is developing at a rapid rate, which gives customers the opportunity to know more about an enterprise’s products and/or services as well as their reputation. This happens outside the control of the enterprise, but the enterprise has to keep this in mind because customers also based their CX expectations on what they have learned by using the technology ([Anderson *et al.*, 2016](#)). Therefore, it is as if a customer should be part of the enterprise’s DNA, in other words the enterprise’s culture and strategy ([Lindgreen & Swaen, 2010](#)).

A third reason why CX is important, is because according to [Meyer & Schwager \(2007\)](#) it encompasses every aspect of what an enterprise is able to offer to customers. Therefore, in order to keep up with it, an enterprise should be able to measure it and adjust their strategy accordingly. An enterprise should also know how to proceed after they measured the CX and that is where Customer Experience Management plays a vital role. This will be discussed in more depth in Section 2.3.

Lastly, the biggest secret to CX is that an enterprise should ensure that the features of their products and/or services are set up in such a way that it will enhance the time a customer has

2.3 Customer Experience Management

with them and make it more enjoyable for the customer. It is important to keep the following three aspects in mind to build a unique CX, according to [Dandridge \(2010\)](#):

1. See everything you do through your customers' eyes.
2. Listen to your customers.
3. Empower the employee to make sure the customer is looked after.

CX can therefore not be overlooked in an enterprise. But how does one manage such an experience? In order to investigate this, Customer Experience Management will be discussed in the next section.

2.3 Customer Experience Management

In the previous section, the question as to what is meant by a CX has been answered. In this section, the question that needs to be answered is, how does one manage the CX?

For an enterprise to improve and look after their customers they should follow a customer-first marketing strategy. In other words, in everything they do they should put their customers first and move to a customer-centric view. According to [Rowe \(2017\)](#), a customer-first marketing strategy is “an approach to marketing that strives for the highest degree of customer satisfaction through deep understanding of customers' needs and wants and creates a value proposition with valuable products and services that exceeds their expectations.”

In literature there are various management approaches in place to follow this customer-first marketing strategy. These management structures include the following:

1. *Customer Relationship Management (CRM)*: It is an approach in which the relationship with current and potential future customers is managed, by analysing the customers' history with the enterprise to improve the business relationship with the customer.
2. *Customer Interaction Management (CIM)*: It is a software application in which the interactions between the enterprise and its customers are managed, by capturing the knowledge available to the customer. This approach is transaction-specific.
3. *Customer Knowledge Management (CKM)*: It is an approach in which the enterprise uses tools and practices to capture, store, organise, access and analyse customer data with the purpose of enhancing their sales, retention and engagement efforts.
4. *Customer Experience Management (CEM)*: Is the collection of processes an enterprise uses to track, oversee and organise every interaction between a customer and the enterprise throughout the customer lifecycle.

2.3 Customer Experience Management

5. *Service Quality Management (SQM)*: It is an approach in which the quality of the services delivered is managed by comparing it with the customer expectations.

Of these five management approaches, the two approaches most used by industry are CRM and CEM. CRM is the popular approach and has been present for over a decade, whereas CEM is a novel approach (Best *et al.*, 2016; Meyer & Schwager, 2007; Schmitt, 2003; Walden, 2017).

The most popular managing structure used by enterprises for managing the CX is CRM. It is also found to be the most popular structure in literature when specific references have been made to CX. Therefore, this section will discuss what CRM and CEM entail, where more focus will be placed on CEM.

The section will begin with a broad overview of CRM. Then CEM will be discussed, based on what it is, why it is used, how it can be done and the challenges associated with it. Afterwards a comparison will be drawn between CEM and CRM as well as why CEM is a better managing structure than CRM to manage a CX. This will also contribute to the fact that focus is placed on CEM and not CRM.

2.3.1 Customer Relationship Management

The most popular approach for the management of customer-first marketing strategy used in literature is that of CRM. This subsection gives a broad overview of this management approach for CX to give a background to CEM.

2.3.1.1 What is Customer Relationship Management

It is important to understand what CRM entails, before an overview can be given on how it should be done.

The definition of CRM as defined by Payne & Fow (2005) is that “CRM is a strategic approach that is concerned with creating improved shareholder value through the development of appropriate relationships with key customers and customer segments.” In other words, it focuses on the improvement of relationships with customers and therefore it is an internally focused approach. The enterprise will therefore develop the appropriate systems, processes and skills to manage these relationships.

Another definition of CRM as defined by Chen & Popovich (2003) is as follows, “We believe that CRM is not merely technology applications for marketing, sales and service, but rather, when fully and successfully implemented, a cross-functional, customer-driven, technology-integrated business process management strategy that maximises relationships and encompasses the entire organisation.” From this definition it can be seen that through improving the relationships with customers, the entire enterprise will also benefit from it. Acharyulu (2012) also mentions that this approach depends on

2.3 Customer Experience Management

business operations that are customer-centric and that operations should be managed in such a way to run these relationships efficiently and effectively.

These are only a few of the definitions mentioned in literature. There are more definitions available in literature that define CRM from different perspectives. Therefore, the definition by Payne & Fow (2005) of CRM along a continuum as can be seen in Figure 2.6 is a good way of establishing how broad CRM is and how widely it can be applied.

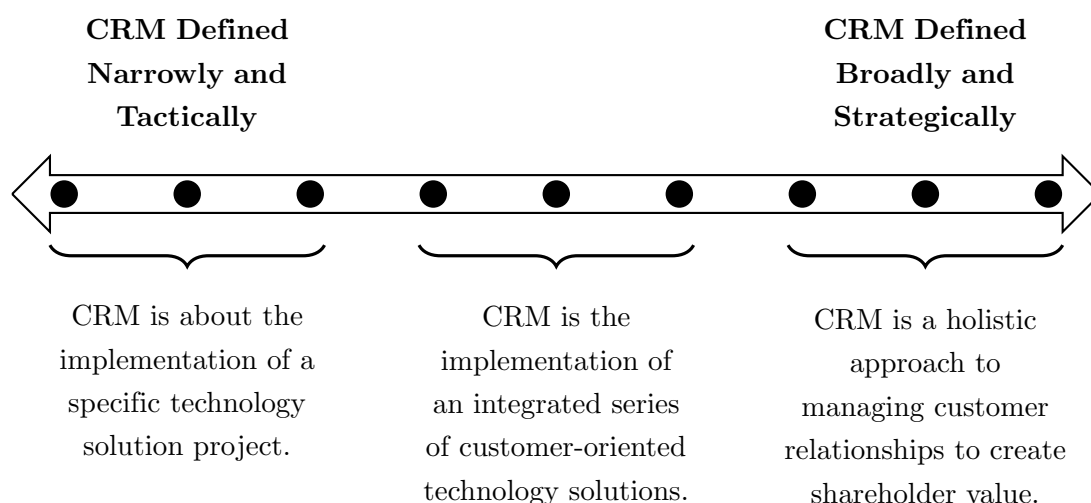


Figure 2.6: The Customer Relationship Management continuum (Payne & Fow, 2005)

2.3.1.2 How to do Customer Relationship Management

As can be seen in Figure 2.6, CRM can be applied in various ways in an enterprise. This creates the problem that CRM means different things for different role players in the business world.

Therefore, in order to overcome such a problem the CRM can be built on the following model as proposed by Winer (2001) which can be seen in Figure 2.7. The CRM Model describes seven components that need to be considered to develop a CRM solution that will work for the enterprise and its requirements. In other words, this model focuses on what an enterprise needs to know about their customers and how to use the information to develop a CRM-oriented enterprise. The seven components are as follows:

1. *Database of customer activity*: This database forms the foundation of the CRM solution.
2. *Analysis of database*: Analyse the database to determine customer segments or customer profiles.
3. *Decisions on which customers to target*: From the analysis results, determine which customers to target when performing marketing activities.

2.3 Customer Experience Management

4. *Tools to target customers:* Determine what tools and methods to use when marketing the products and/or services to the targeted customers.
5. *How to build relationships with target customers:* Determine what programs should be used for building and maintaining relationships with customers.
6. *Privacy Issues:* Important to protect customers' data and keep in mind the trade-off between improving relationships and the amount of information needed to do it.
7. *Measuring metrics for the success:* Develop customer-centric metrics to determine the success of the CRM solution and ensure that it will be able to give managers a better idea of how the CRM policies and programs are working.

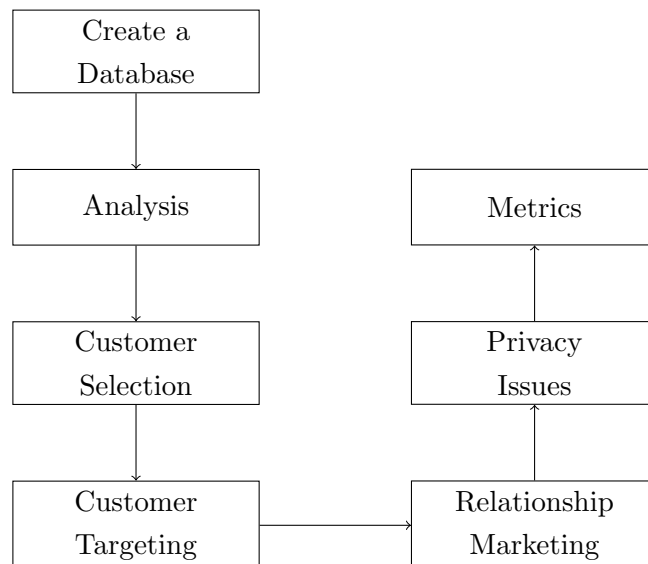


Figure 2.7: Customer Relationship Management model (Winer, 2001)

Another key aspect to keep in mind when creating a CRM solution for an enterprise is the three-dimensional aspect of it. The three dimensions as mentioned by Chen & Popovich (2003) are *people*, *processes* and *technology* as shown in Figure 2.8.

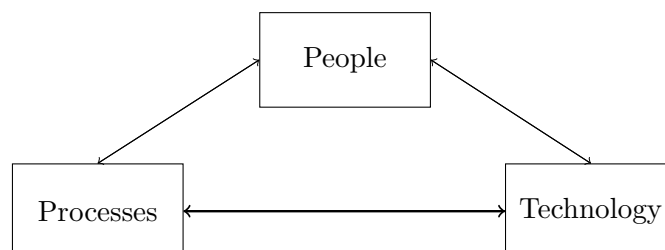


Figure 2.8: The three dimensions of Customer Relationship Management (Chen & Popovich, 2003)

2.3 Customer Experience Management

These three dimensions all have to act together in an environment where the following three characteristics of *customer-driven*, *technology-integrated* and *cross-functional* are present in an enterprise. Each of the three dimensions has their independent functionality but also need to act as an integrated unit. The three dimensions of CRM can be discussed as follows:

1. *People*: Consists of all human role players within the concept and control of CRM. The people dimension can be summarised in the *layer of role players* as can be seen in Figure 2.9.
 - (a) First layer: the *Customers* are the most important role-player. The enterprise should examine their existing and potential customers and determine in which customer segment they belong. This implies that CRM is an after-the-fact approach, since an enterprise can only examine a customer afterwards (Meyer & Schwager, 2007).
 - (b) Second layer: the *Employer and employees* are the building blocks of customer relationships. The employer should provide support and commitment throughout all aspects of CRM as they steer the enterprise in certain directions determined by their decisions. The employees on the other hand are the ones that ensure that the enterprise is steered in that direction. Therefore, it is important to strive for employee engagement as it has a significant impact on an employees' performance (Anitha, 2014). If an employee is more engaged in their work and in making decisions, they will automatically perform better and customer relationships can be improved. According to Payne & Fow (2005), employee engagement is a critical factor for the successful implementation of CRM.
 - (c) Third/outer layer: the *Organizational Culture* has an effect on the employers and employee layer. Rahimi (2017) points out that the organisational culture can either enable or disable the success of CRM in an enterprise. It is important to consider all the factors of the organisational culture that can have an impact on the employees.

2.3 Customer Experience Management

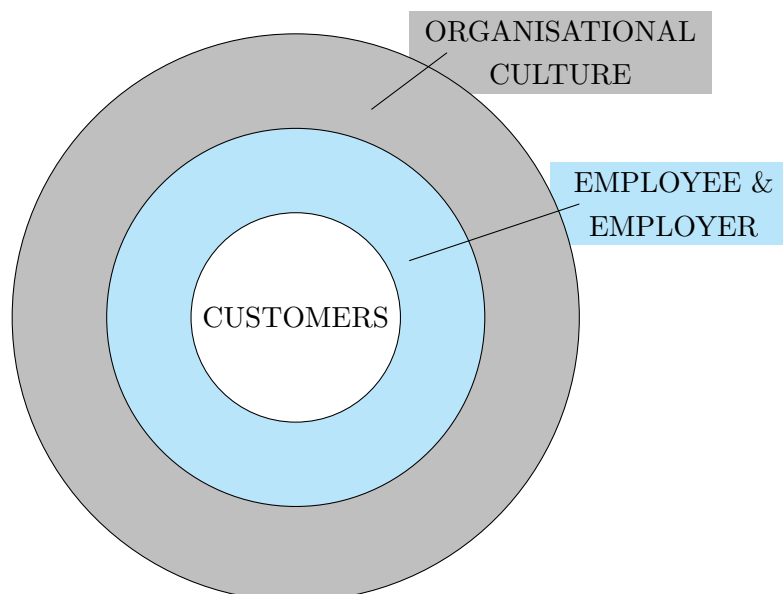


Figure 2.9: People dimensions: Layers of role players

2. *Processes*: This is all the processes that should be used in the implementation of a successful CRM approach. According to Kale (2004) it is very important to identify the correct strategic processes that need to be in place between an enterprise and its customers. The important thing with this dimension is to determine how to prioritise the components of the processes in order to bring about change which will improve the relationships with customers. These processes consist of the following five processes, according to Payne & Fow (2005):

- *Strategy Development process*: It controls how the enterprise strategy and customer strategy interrelate with one another.
- *Value Creation process*: It transforms the strategy development process into outputs that extract and deliver value to the enterprise with regards to, (i) value that the enterprise can provide, (ii) value that the enterprise can receive and (iii) the successful management of this value exchange.
- *Multichannel Integration process*: It creates value-adding activities with the customer by translating the outputs of the enterprise strategy and value creation process together.
- *Information Management process*: It generates customer insight and appropriate marketing responses by collecting, collating and use of customer data and information from all interaction points with the customer.
- *Performance Assessment process*: This consists of two components. The first is the macro view of the overall relationships which drive performance, as well monitoring it. The second is the micro view, which is a more detailed view of the metrics and key performance indicators.

2.3 Customer Experience Management

3. *Technology*: This concerns how it is used to enable and improve customer relationships. It is an important aspect as technological innovation is needed for CRM in order to be able to collect and analyse data on customers. The technology should be able to develop customer patterns, interpret customer behaviour, deliver products and/or services that enhances the customer's value, be able to respond effectively and efficiently to communication and lastly to develop predictive models about the customer. It is the enabler of understanding and predicting the customer. Another requirement as mentioned by [Winer \(2001\)](#) is that the technology should have online capabilities in order to keep up with the world and make it easier for the enterprise to understand their customers. Technology should also be able to access multiple channels and integrate the data gathered over these channels ([Payne & Fow, 2005](#)). The last aspect that technology should have is that technology should also protect the customers' data and it should know what data can be analysed.

From this section, it has been established that CRM is a customer relationship oriented management approach to CX. The CRM model should be adopted as a CRM solution to any enterprise and it should be adjusted to fulfil that specific enterprise's need while keeping in mind the three dimensions of CRM. In literature, many cases were investigated to determine how to apply CRM to different industry sectors. The application of CRM can be seen in [Acharyulu \(2012\)](#), [Venugopal & Priya \(2015\)](#), [Zineldin \(2005\)](#), [Krasnikov et al. \(2009\)](#), [Kahraman et al. \(2007\)](#), [Özgener & İraz \(2006\)](#) and others.

2.3.2 Customer Experience Management

CEM is an approach that has made its appearance over the past decade and has only been adapted by a few enterprises. As mentioned by [Goldenberg \(2017\)](#) only a handful of enterprises were able to implement it successfully. In order for CEM to be successful an enterprise should make it part of their DNA. As the term suggest, CEM is a strategic management approach in which an enterprise manages all the experiences a customer has with their products and/or services. By doing this, an enterprise steps into the shoes of the customer and makes decisions from their perspective. But in order to understand the power of CEM, CEM will firstly be defined in more depth. Secondly, it will be determined why CEM should be implemented. Thirdly, how to achieve CEM will be investigated by looking at different frameworks and models in literature on how to achieve CEM excellence. Lastly, the challenges enterprises currently face when implementing CEM will be determined.

2.3.2.1 Definitions of Customer Experience Management

As said before, CEM is a strategic management approach in which all the experiences of a customer during interactions with a product and/or service are managed. [Schmitt \(2003\)](#) defines the approach,

2.3 Customer Experience Management

that can take the form of a discipline, methodology and/or process to comprehensively manage a customer's cross-channel exposure, transactions and interaction with an enterprise, product, brand or service.

A formal definition that explains it better is that of [Goldenberg \(2017\)](#) which states that “CEM is the collection of business process and technology tools that companies leverage to manage all their customer interactions.” These interactions are a combination of interactions over time while the management of them should occur at the same time. These interactions are then recorded over a customer journey, which will be discussed in more depth in Section 2.4.

[Verhoef et al. \(2009\)](#) also points out that CEM is a strategy in which the CX should be engineered in such a way that value can be created both for the enterprise and the customer. In order to achieve this, an enterprise should shift their focus from the historical activities of the customer to the current experience of the customer. Therefore, an enterprise should shift from an inward focus, known as ‘*inside-out thinking*’, to an outward focus, known as ‘*outside-in thinking*’. The focus of the inside-out thinking is on departmental and operational goals and measures, whereas the focus of the outside-in thinking is on delivering high-value CX through continuous customer engagement ([Best et al., 2016](#)).

Therefore, as mentioned by [Morgan \(2015\)](#) “CEM urges a *sea change* where the organisation focus moves toward optimising each CX. In other words, one can say that CEM is a strategy that will transform a business-oriented enterprise to a customer-centric environment, where an enterprise creates profit by delighting the customer ([Best et al., 2016](#)). As indicated by [Parandker & Lokku \(2012\)](#) it is also used to minimise the gap between the enterprise and the customer in order to achieve a positive CX.

[Du Plessis & De Vries \(2016\)](#) see CEM as a “broad field that attempts deliberately to align the enterprise and various activities in an enterprise ultimately to deliver good CX that satisfy the enterprise's customers.” In other words when an enterprise implements CEM it is important to note that they have to look externally but also internally to understand how to adjust their processes, systems and skills to keep the customer as their first priority. Using CEM enables an enterprise to design, manage and optimise the end-to-end CX by providing the appropriate strategies, process models and information technology ([Sukwadi, 2015](#)). From all this information, Definition 2.3 has been formulated that will be used as a basis for the study.

Definition 2.3 (Customer Experience Management). *It is a strategical management approach that can be in the form of a discipline, methodology and/or process in which an enterprise manage a customer's interactions with the brand, products and/or services by using outside-in thinking and transforming their processes, systems and skills to be customer-centric instead of business-centric.*

2.3 Customer Experience Management

It is also important to note as was discussed in Section 2.2 that a customer always has an experience when they make use of a product and/or service or interact with an enterprise, whether the experience is good, bad or average. Therefore, to achieve the ‘superior CX’ it is important that the enterprise should be proactive. CEM is an approach in which the enterprise will be able to respond proactively to customers’ needs.

2.3.2.2 Why Customer Experience Management?

The reasons why CEM is a good strategical management approach to be implemented by any enterprise are as follows.

The first reason as defined by [Best et al. \(2016\)](#) is that there are five factors that show that the market is ready for CEM. Figure 2.10 represents these five factors. The first factor is *Market Maturity*, where the enterprises are maturing in such a way that the competitive nature changes. The second factor is *Economic Trends*, because these trends determine how customers spend their time and money. The third factor is *New Technologies*, where the technology consists of new capabilities that have the potential to improve CX. The fourth factor is *Customer Behaviour and Preferences*, which creates new risks and opportunities for enterprises as it changes. The fifth factor is *Shift in Competition*, that leads to quality improvement and more strategies that are customer oriented.

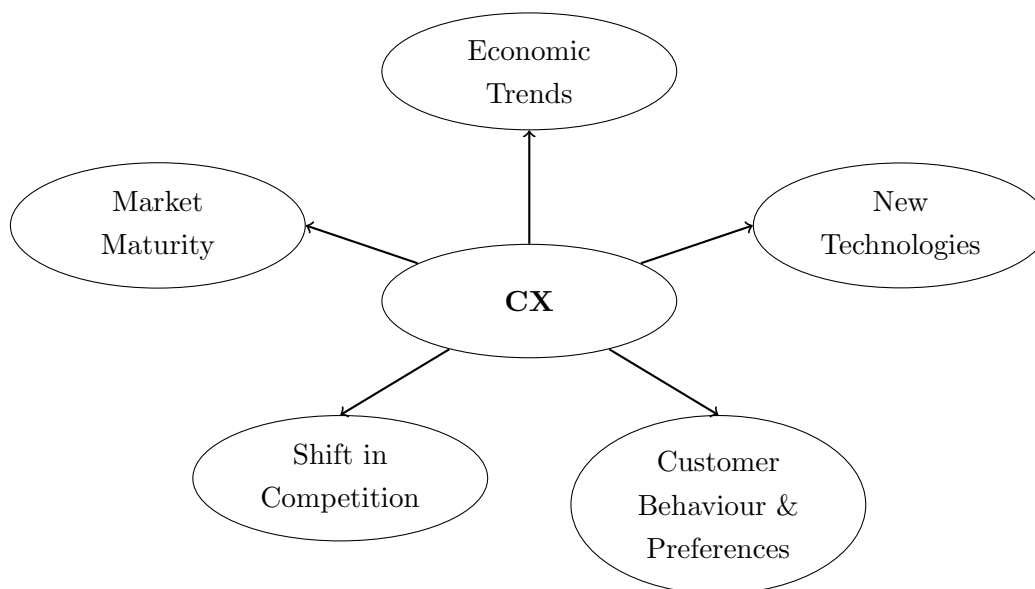


Figure 2.10: Market readiness for Customer Experience Management ([Best et al., 2016](#))

From these factors it can be seen that the market is ready for enterprises to implement CEM. But the key drivers for an enterprise to implement CEM can either be external or internally oriented, depending on the state of the enterprise.

The second reason why CEM is a good strategical management approach is because in today’s world view there exists a great need to understand the customer perspective and expectations and

2.3 Customer Experience Management

to develop solutions and strategies to determine how to address these needs. In other words, the enterprise is built in such a way that the customer is the first priority and will be served in the best way possible. CEM enables an enterprise to delight the customers and then to earn profit from that (Best *et al.*, 2016).

The third reason as mentioned by Parandker & Lokku (2012) is that CEM enables an enterprise to address the customers' individual needs and to understand the customers' behaviours. When an enterprise understands the customer's needs and behaviour it enables them to adjust the business model more efficiently in order to advance in the CX.

2.3.2.3 How to do Customer Experience Management

A very important element to the success of CEM is to ensure that all aspects and divisions of the enterprise support the CEM initiative and everyone is on board. This 'team work' begins at top management and spirals downwards. Therefore it is important to understand how to implement it successfully.

Rich (2015) came up with 12 actions that are needed to successfully implement CEM. It is important to note that these actions should not be done in isolation, but they are closely interlinked and require consistent and concerted effort. The 12 actions are as follows:

1. *Do not waste time, money and effort:* Use best practices, data management and domain frameworks to get the right CX.
2. *Always look from the outside in:* Everyone within an enterprise should approach everything from the customer's point of view.
3. *You have to have vision:* Top management should set the vision and the enterprise should continuously improve and be agile.
4. *Know what the customers want:* Customer journeys are vital for success, together with the fact that customer behaviours and preferences shift all the time at a micro level.
5. *You need a complete picture:* Develop and manage an enterprise-wide vision to deliver the right experience at all customer interaction points.
6. *Customers' data is an asset:* Good data management structures should be in place for collecting, collating, analysing and securing of customer data.
7. *Personalisation is popular:* Give recognition to each customer as each one is unique and look for a tailored experience.
8. *Simplify ruthlessly:* The need for processes that can be replicated and reused to simplify processes to give the customers what they want.

2.3 Customer Experience Management

9. *Partner management is critical*: Treat the customers as partners and ensure the four key elements of financial, contractual, operational and technical are present.
10. *Manage the programme*: Programme management and governance should be included in the business model.
11. *Invest in people*: The culture of the enterprise has a huge impact on the level of CX that will be delivered.
12. *Quick wins build confidence*: Prioritise what parts of enterprise should be addressed first to achieve the desired CX.

But in order to do it in the most efficient and effective way possible, the customer journey is a big enabler. The customer journey captures every experience a customer has with a product and/or service. This enables the enterprise to better understand the customer's needs and CEM can be implemented. The customer journey will be discussed in more depth in Section 2.4.

In order for an enterprise to make CX an important segment of the business, the following nine characteristics have to be kept in mind, as defined by [Anderson et al. \(2016\)](#):

1. *Branded Experience*: The CX delivered by the enterprise should match their brand. It is important to integrate these two aspects, to ensure that the brand lives up to the expected experience.
2. *Responsive Interface*: An enterprise should ensure that a customer can have the same experience across multiple devices and/or channels.
3. *Transparency and control*: An enterprise should stay ahead with the constantly changing environment.
4. *Team of experts*: A team of experts should be available to help an enterprise to deliver what they promise and to improve the CX.
5. *Business Objectives*: The enterprise should keep the CX in mind when creating business objectives and adapt accordingly.
6. *Recognition Technology*: The technology in an enterprise should be able to accommodate CX analytics data and how to give recognition for the business environment.
7. *Respondent Experience*: An enterprise should be open for feedback to continuously increase the CX.
8. *Agility and Flexibility*: An enterprise should be able to be agile and flexible in order to keep up with the marketplace.

2.3 Customer Experience Management

9. *Security and Availability*: With the emerging world of Big Data, it is a critical element to protect the customers' data 24/7.

The nine characteristics on how to achieve a good CX and the twelve actions required for implementation are the key to the successful implementation of CEM. But one might still wonder, what steps or stages an enterprise should go through to implement CEM that will suit the enterprise structure and needs. Various models and frameworks have been developed to cater for this. [Du Plessis & De Vries \(2016\)](#) have conducted a thematic and quantitative analysis on 23 CEM approaches (models and frameworks) that were created from 1993 to 2014 and have identified which themes were more prominent in these models and frameworks. A summary of the study can be seen in [Table 2.2](#).

Table 2.2: Customer Experience Management theme classification ([Du Plessis & De Vries, 2016](#))

Classification	Number of occurrences
CX design and implementation process:	
Customer understanding	20
Experience design	26
Experience implementation	22
Experience measurement	16
Organisational factors influencing CX:	
Leadership	10
Organisational design	4
Strategy	19
Culture	22
Systems / Technology / Processes	10
Grand Total	149

Based on the study of [Du Plessis & De Vries \(2016\)](#), it is clear that the 23 frameworks and models existing in literature are different. To discuss each approach in detail (adding to the fact that yet more approaches might have been developed) will be extensive. Therefore, [Table 2.3](#) represents a summary of six of the models and frameworks that are available for the implementation of CEM. These six approaches show that any type of enterprise can adapt the CEM approach. There are more models and frameworks available that contain more factors and adaptability to other industries as well.

Table 2.3: Models and frameworks for the implementation of Customer Experience Management

Model or Framework	Description	References
Company & Customer value framework	This conceptual framework shows how the CX and exchanged value are linked, as well as the interrelationships and mutual relationships with the main entities, where the main entities are the customers and the enterprise.	Gentile et al. (2007)
An Organising Framework	This framework represents how macro factors influence the retailers and the CX. It also looks at the retail CX, the factors that the enterprise can control and the market and financial metrics.	Grewal et al. (2009)
Conceptual Model of CX creation	This conceptual model is built based on the predecessor of and moderators of CX and it examines the need for CEM strategies to take these elements into account.	Verhoef et al. (2009)
Business Process Improvement Model with integrated CEM	This model helps an enterprise to improve the business processes to fit to the customers' needs. It links the CX to the business processes and translate the needs of the customer into technical design characteristics for the processes.	Both & van Rensburg (2010)
Conceptual CEM Framework	This framework integrates the SERVQUAL, IS and QFD models. The SERVQUAL methods are used to determine the customer needs based on five dimensions and it assesses the customer's perception of the experience. The IS model is used to identify potential areas for improvement. The QFD model is used to design or renovate with the purpose to focus on customer needs and service management requirements.	Sukwadi (2015)
Holistic CEM Framework	The objective is to provide guidance on how an enterprise should align themselves to be more customer-centric and how to implement customer-centric principles which should lead to an enhanced CX. The framework consists of two parts, the first being the CX design and implementation process and the second the building blocks for CX.	Du Plessis & De Vries (2016)

2.3 Customer Experience Management

2.3.2.4 Challenges with implementing Customer Experience Management

The challenges enterprises currently face in order to implement the CEM approach can be defined based on two categories; *business process* challenges and *system and technical* issues.

The top three *business process* challenges, have been identified by Rich (2015) as:

1. Getting a holistic view of the customer.
2. Understanding account profitability.
3. Meeting the conditions of service level agreements.

The biggest challenge for any enterprise is to gain a holistic view of the customer. This customer view is important as it enables an enterprise to get the right management and support structures in place to deliver the ‘superior’ CX. But due to legacy systems, data, issues around securing of data, interdepartmental collaboration, partnering problems and general process complexities it is hard to obtain the right view.

Another business process challenge, that does not form part of the top three but is important to mention, is the ability to understand the customer’s needs. Due to many factors that shape the market, it is becoming increasingly more difficult to predict what the customer wants. Parandker & Lokku (2012) also points out that the customer’s nature of perception is changing and they are vulnerable to several factors. Therefore, it is a challenge that can prevent enterprises to successfully implement CEM.

The top five *systems and technical* challenges, have been identified by Rich (2015) as:

1. *End-to-end control*: This is a challenge due to the diverse collection of systems and network elements available to collect data on customers.
2. *Systems complexity*: It is a challenge due to the breadth of customers, influencing systems and the long history available.
3. *Justifying investment*: It is a challenge due to fragmentation of priorities, the agility and returns on investments.
4. *Quality of analytical tools*: It is a challenge because more tools are available and it is difficult to recognise the correct tool to be used.
5. *Availability of critical skills*: It is a challenge as there are concerns around the complexity, agility and analytics capabilities together with an understanding of customers’ preferences and profitability.

2.3 Customer Experience Management

There are more challenges besides these mentioned here. But it is important that enterprises are able to identify the challenges they are facing when they want to implement CEM. Once the challenges have been identified, the enterprise should be able to come up with solutions on how to overcome these challenges.

2.3.3 Customer Relationship Management versus Customer Experience Management

In the previous subsections a broad overview was given of CRM and CEM. CEM was discussed based on what it is, why it is needed, how it should be done and the challenges faced by enterprises who want to implement CEM. Using this information the two management approaches can be compared.

CRM is a strategic management approach that enables an enterprise to critically assess their internal components after interactions with a customer as they determine what can be changed with regards to their processes, systems and skills. In other words, with CRM an enterprise focuses on developing the appropriate processes, systems and skills to manage the relationship with the customer. As mentioned by [Best *et al.* \(2016\)](#), with CRM an enterprise assesses how well the services, products, processes and people were performing, but the enterprise does not consider to what degree they are meeting their customers' expectations and how good the CX is.

CEM on the other hand is a strategic management approach that manages all the experiences a customer has with a product and/or service and the enterprise. As [Walden \(2017\)](#) points out, it is not managing an experience without a customer. Rather it determines how an enterprise should proactively respond to the drives and needs of a customer. As mentioned by [Best *et al.* \(2016\)](#) "it is an approach where a business is transformed from a business centric to customer centric approach." The enterprise looks at ways to deliver high value to their customers and change the processes, systems and skills accordingly. One might even say that the goal of CEM is to optimise the CX.

Therefore, there is a difference between CRM and CEM. A comparison between CRM and CEM can be seen in [Table 2.4](#). It is important to note that CEM is not necessarily a replacement for CRM, but CEM is built on top of good CRM processes and practices to take the enterprise a step closer to its customers. Another important distinction is that CEM is not only focusing on transactions, but it will rather go beyond that to build rich relationships with its customers ([Schmitt, 2003](#)).

For the purpose of this study, CEM will be used as the strategic management approach to manage the CX, because it is done in a proactive manner and it looks from the outside in. Therefore, with CEM an enterprise always put their customers first and adjust their processes, skills and systems in order to improve the customers' experience.

2.4 The customer journey

Table 2.4: Comparison of CEM and CRM (Meyer & Schwager (2007), modified)

	CRM	CEM
Strategy / Approach	“Inside-out” Thinking (business-centric)	“Outside-in” Thinking (customer-centric)
Focus	Inward	Outward
What	Captures and distributes what an enterprise knows about a customer.	Captures and distributes what the customer thinks of an enterprise.
When	After recorded customer interaction.	At customer interaction points.
Fundamental Principle	Measurement of what has happened in the past.	Drives a proactive approach to managing a customer.
How	<ul style="list-style-type: none"> • Point of Sale • Market Research • Tracking of Sales 	<ul style="list-style-type: none"> • Targeted studies • Observational Studies • Voice-of-customer research
Who	Customer-focused groups	Business of functional leaders
Relevance for Future Performance	Lagging: Drives cross-selling by bundling in products in demand with ones that are not.	Leading: Locates places to add offerings in the gaps between expectations and experiences.
What question to ask	“This is what we are doing, how well are we doing?”	“What is important to our customers, and how well are we doing?”

2.4 The customer journey

Now that a better understanding has been given as to what CX is and how it can be managed by using the CEM approach, it is important to understand what the customer wants and how a customer experiences the product and/or service which is produced or delivered by a particular enterprise over a distinct period of time. The best way to capture it, is by the use of the customer journey together with what the customer’s behaviour and preferences are on the journey. Both the CX and CEM sections have mentioned that a customer journey can be used to record the customers’ interaction points. In this section, the question that needs to be answered is, what does such a customer journey entail?

In Definition 2.2, it has been stated that a CX is the sum total of experiences. Therefore, it does not entail only a single encounter, but rather the total encounters a customer has with an enterprise and the product(s) and/or service(s) delivered by that particular enterprise. Therefore, to capture the overall CX, it is important to look at all the collective encounters over the entire life cycle (Du Plessis & De Vries, 2016). This is why the customer journey is such an important contributor to capture a CX.

2.4 The customer journey

Definition 2.4 will be used to define the customer journey. With this journey, an enterprise will be able to capture the customer's journey with the product and/or service as it unfolds. The customer journey, therefore, gives an enterprise the ability to fully understand the CX on a real-time basis as it captures the data about the customer's interactions with the product and/or service, as well as with the enterprise itself (Parandker & Lokku, 2012; Richardson, 2010; Ruzicka, 2017b).

Definition 2.4 (Customer Journey). *A journey a customer has with an enterprise over time during the purchase cycle, across multiple touch points and by the use of multiple channels (Lemon & Verhoef, 2014).*

When an enterprise tracks the customer journey, it is important to consider the six factors pointed out in Figure 2.11. These factors give an enterprise the ability to better understand its customer and be more effective in the business, product and operational planning as well as how the spending in the enterprise should be prioritised. According to Best *et al.* (2016) it is imperative that an enterprise should understand what factors influence the interactions a customer has and how the customer journey is influenced by it to create better offers and delight the customers more.

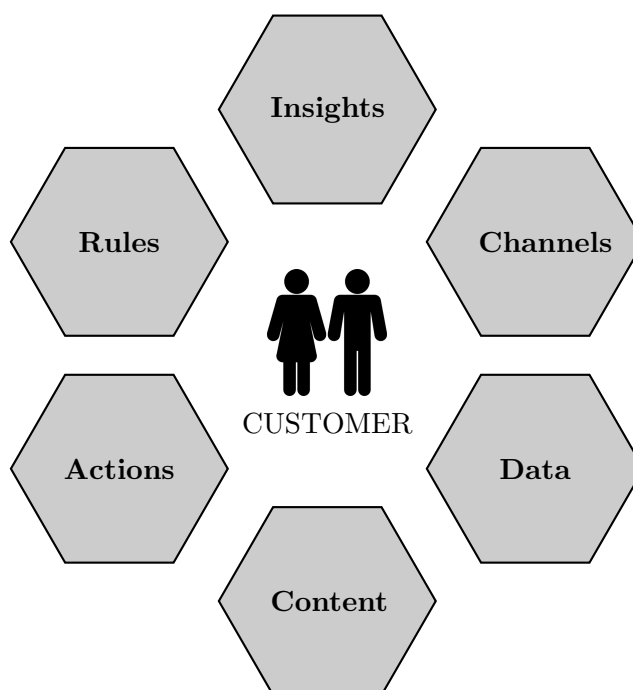


Figure 2.11: Six factors for the customer journey (Ruzicka, 2017a)

It is also important to understand and know that each customer has his/her own unique customer journey for each life cycle. In other words, customer journeys of different customers might have some similarities but they will never be exactly the same. This also counts for a customer who completes a

2.4 The customer journey

customer journey and starts a new journey at a later stage. There might be some similarities between these two journeys but it will never be exactly the same. The reason is that no two customers are the same and a customer can experience an event differently at specific points in time. But Section 2.5 will discuss this in more depth.

An outline of the model for the customer journey can be seen in Figure 2.12. A few variations exist in literature, like the process model as defined by Lemon & Verhoef (2014). For the purpose of this study, the model in Figure 2.12 will be used. This model is developed by the *TMForum* members, Batra & Kawecki (2014) and it helps an enterprise to understand how its customers interact with their product(s) and/or service(s) and the enterprise will be able to anticipate its customers' behaviour to improve their product(s) and/or service(s) delivery to improve the CX.

The *TMForum* customer journey model, also known as the *life-cycle experience model*, not only captures a customer journey across different phases. It also includes the channels which are used by the customer during the journey (*Channel* layer), the unique experiences a customer can have (*Experienced Events* layer) and the insights and knowledge that an enterprise can gain by applying advanced analytics algorithms while the customer is on their journey (*Insights & Knowledge* layer).

To track a customer journey it is important to know what the following entails, as identified by Definition 2.4:

1. the purchase cycle,
2. the touch points, and
3. the multiple channels.

The first aspect of the customer journey is the *purchase cycle*. The purchase cycle can be divided over three distinct phases namely (Batra & Kawecki, 2014; Lemon & Verhoef, 2014):

- (i) *Prepurchase (buying) phase*: This phase encompasses all the aspects and touch points when a customer is considering, making a decision and purchasing a product and/or service. It includes the interaction with the brands, category and environment before the purchase transaction is taken.
- (ii) *Purchase (using) phase*: This phase encompasses all the aspects and touch points when a customer actually uses the product and/or service. It includes the interaction to make payments, managing the account, seeking and receiving help.
- (iii) *Post-purchase phase (sharing) phase*: This phase encompasses all the aspects and touch points while the customer learns and uses the product and/or service, as well as sharing the experience with others until the discontinuation of use of the product and/or service. It includes the interaction with the brand after the purchasing and renewing of the product and/or service.

2.4 The customer journey

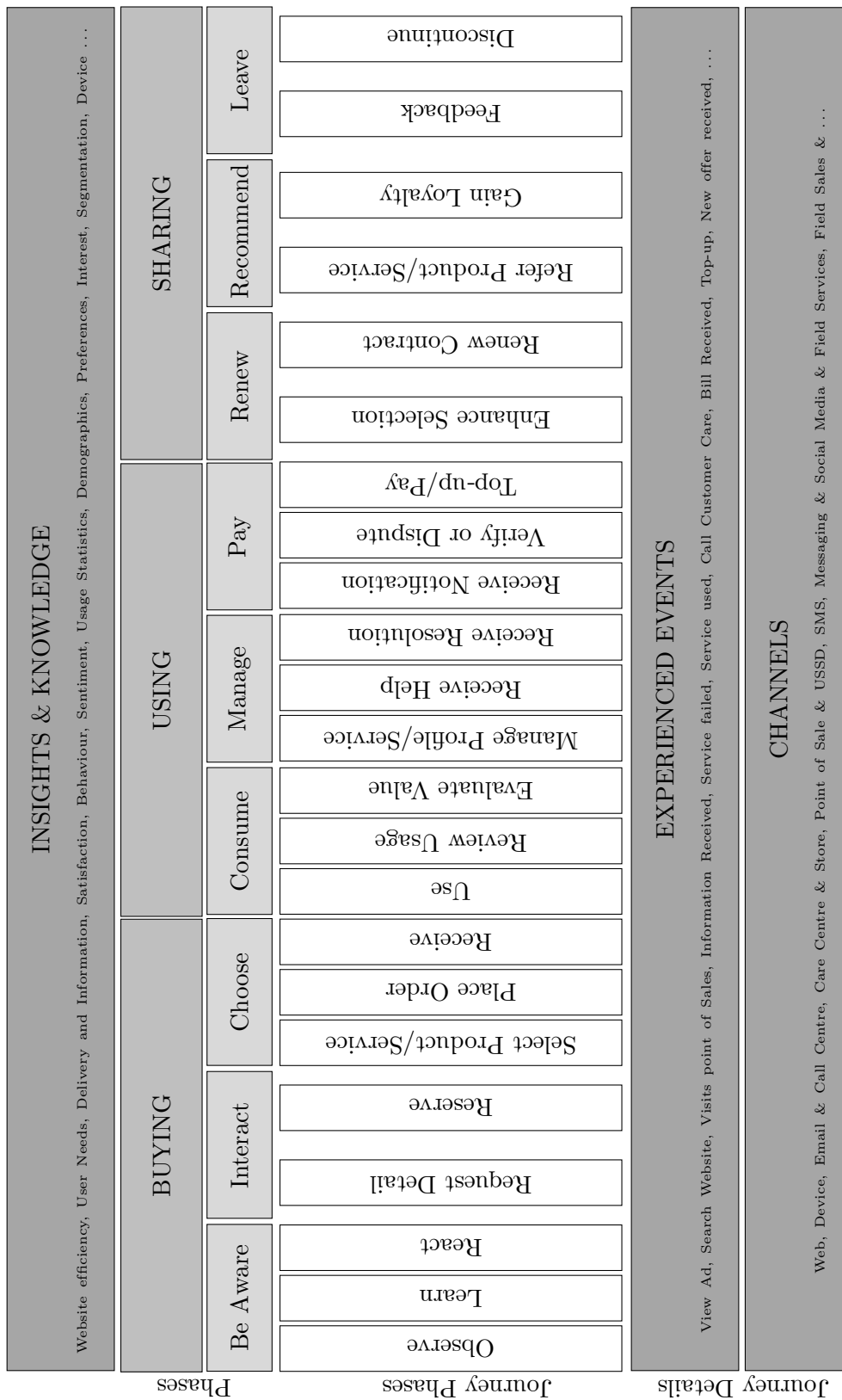


Figure 2.12: Outline of the customer journey model (Batra & Kaweck, 2014)

2.4 The customer journey

These three phases can then be further subdivided into nine phases, where the nine phases can then be divided into 23 *journey phases* as can be seen in Figure 2.12. A broad description of each of the nine overall phases and 23 journey phases can be seen in Table 2.5 . During these phases, the CX is recorded at various touch points. But the customer can also have an unique experience during any moment of these phases which is captured in the *Experienced Events* layers and will contribute to more insights to the CX.

Table 2.5: Description of the journey phases (Batra & Kaweck, 2014)

	PHASE	DESCRIPTION
	BE AWARE: The customer is becoming aware of the product and/or service.	
1.	Observe	Customer is implicitly or explicitly exposed to product and/or service.
2.	Learn	Customer is learning more about the offerings via directing an inquiry or viewing reviews & feedbacks.
3.	React	Customer has a natural reaction to the offerings.
	INTERACT: The customer begins to interact with the enterprise(s) via a channel of his/her choice.	
4.	Request Detail	Customer explicitly interact with the enterprise about the detail of the offer that is related to them receiving it.
5.	Reserve	Customer requests the enterprise to hold the inventory but no formal selection or order submission has been placed.
	CHOOSE: The customer selects a product and/or service and an enterprise.	
6.	Select Product / Service	Customer selects the offer they prefer.
7.	Place Order	Customer places an order with the enterprise.
8.	Receive	Customer receives the product/service at the time of selection or after selection.
	CONSUME: The customer makes decisions about the usage and personal experience of the product/service.	
9.	Use	Customer's experience revolves around the usability and consumption of the product/service.
10.	Review Usage	Customer reviews the product/service based on the usage.
11.	Evaluate Value	Customer evaluates the value and compares it with their expectation at the purchase point.
	MANAGE: The customer manage their account, make changes and seek help or assistance.	
12.	Manage Profile / Service	Customer interacts to manage his/her profile, product and/or service.
13.	Receive Help	Customer contacts enterprise to diagnose and solve problems that they may encounter with the product/service.
14.	Receive Resolution	Customer will request help for a resolution if their problems have not been solved.

Table 2.5 continues on next page

2.4 The customer journey

	PHASE	DESCRIPTION
	PAY: The customer receives and views charges and make necessary payments.	
15.	Receive Notification	Customer gets a notification from the enterprise about their balance or payment request.
16.	Verify or dispute	Customer verifies the notification and confirms/disputes the charges with the enterprise.
17.	Top-up / Pay	Customer tops up account or makes payments.
	RENEW: The customer decides about renewing the subscription.	
18.	Enhance Selection	Customer makes changes based on past and other's experiences.
19.	Renew Contract	Customer re-contracts the product/service from the enterprise once selection has been confirmed.
	RECOMMEND: The customer refers product/service & enterprise.	
20.	Refer Product / Service	Customer refers the product/service and enterprise (negatively or positively) based on their personal experience.
21.	Gain Loyalty	Customer gains loyalty with the enterprise which can be in monetary or non-monetary terms.
	LEAVE : The end of the customer and enterprise relationship.	
22.	Feedback	Customer provides feedback (implicitly or explicitly) when they leave the relationship with the enterprise.
23.	Discontinue	Confirms the cancellations of the relationship between the customer and enterprise.
End of Table 2.5		

The second aspect of the customer journey is the *touch points*. Touch points are when a customer interacts with an enterprise and it can be seen as the 'moments of truth' as it helps an enterprise to understand how, why and when a customer interacts with the enterprise itself (Meyer & Schwager, 2007; Opperman, 2017). Therefore, touch points are instances of direct contact with the product and/or service and with the representations of it. The representation can either be by the enterprise itself or some third party. Touch points therefore, have a huge impact on the measurement of CX, as they capture the CX throughout the journey at specific points in time and by doing that the customer journey is created on a real-time basis as the customer move through various touch points.

The touch points can be categorised into four categories(Lemon & Verhoef, 2014), namely :

- (i) *Brand-owned touch points*: When the customer's interactions are managed and designed by the enterprise and are under the enterprise's control. This includes any brand-controlled elements of marketing (*e.g.* packaging, price or attributes of the product) and brand-owned media (*e.g.* loyalty programs, websites or advertising).
- (ii) *Partner-owned touch points*: When the customer's interactions are jointly designed, managed or controlled by the enterprise itself together with one or more of its partners. The partners of the enterprise can be any of the following: communication channel providers, multichannel distribution partners, marketing agencies or loyalty program partners.

2.4 The customer journey

- (iii) *Customer-owned touch points*: When the customer is controlled by the customer's actions, where the enterprise, its partners or any other party has no control over it and cannot influence it. The experience is totally in control of the customer's own preferences and behaviour.
- (iv) *Social/external touch points*: When the experience of the customer is influenced by external touch points, which are controlled by other parties. This includes the environment the customer is in, other customers, influences of peers and independent information sources such as social media.

According to [Parandker & Lokku \(2012\)](#) it is important that an enterprise should determine whether the customer expectation was met at a specific touch point as it helps an enterprise understand what the customers' expectations are and what is important. Based on the categories mentioned above, it will increase the enterprise's ability to know what touch points they have control over and how they can influence it to improve the customer's overall experience.

The third aspect of the customer journey is the *multiple channels* through which a customer can interact with an enterprise. Historically, a customer had to complete their activities by using a limited set of channels. For example, a communication service provider only wanted a customer to call the contact centre or end up doing their transactions and activities in-store. But as identified by [Mitra & Kawecki \(2016b\)](#) the world has changed in such a way that the customer has the option on when and where to interact with a brand or enterprise and by using the channel of their choice. The environment in which enterprises operate today, encourages enterprises to take full advantage of all the multiple channels available to increase the number of opportunities to interact with the customers and to monitor these interaction points ([Gentilea et al., 2007](#); [Hamid & Akhir, 2014](#)). This gives the customer the opportunity to interact via their own preferred choice of channel, which ensures that the customer is placed in the centre and this contributes to the improvement of the overall CX ([Goldenberg, 2017](#)). But the enterprises should ensure that is integration between all these channels' platforms in order for the data to be consistent across all the channels and the data that flows across these channels ([Best et al., 2016](#)).

But what does such a customer journey look like? An example of a customer journey can be seen in [Figure 2.13](#). This is a simplified customer journey, where the 'experienced events' and 'insights and knowledge' layer have been left out. The black dots show the enterprise at which phases the customer has interacted with them (the touch points) and the line shows the order in which it was done. It is clear that the journey is not a sequential journey, but rather that the customer can bounce around different phases in a random order and the customer does not have to complete all the phases. It also indicates to the enterprise what channels the customer has used during this journey. The efficiency and effectiveness of these channels will also contribute to the CX. With this example it is crucial to notice that even though a dot represents a touch point it does not entail only

2.4 The customer journey

one interaction. The customer might have interacted several times with the enterprise and product and/or service at a specific point. This is where the ‘Experienced Events’ layer is useful as it captures a customer’s unique experience(s) at the different touch points during their journey. The ‘insights and knowledge’ layer gives the enterprise the ability to perform analytics to better understand the customer and improve the CX. The analytics are performed on the specific customer journey on a real-time basis. It is also important to note that each customer will have their own unique customer journey.

2.4 The customer journey

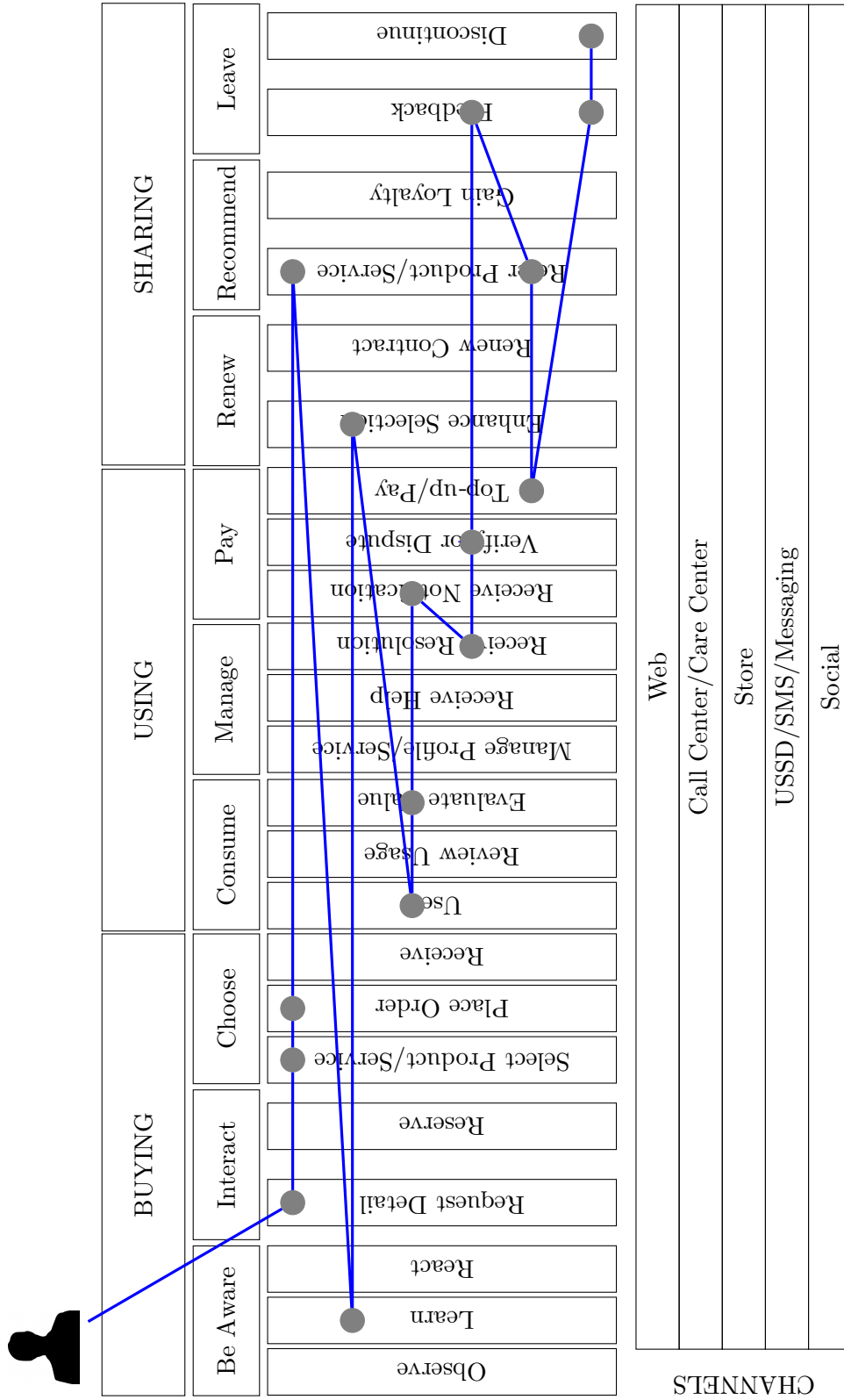


Figure 2.13: Customer journey example (Batra & Kaweck, 2014)

2.5 360-Degree view of the customer

To improve the CX, the customer journey alone will not suffice. As mentioned by Rich (2015) it is imperative that enterprises are able to *understand* these journeys while doing analysis in order to improve the customer journeys. Therefore, it is important to understand how and what influences the customer and how the customer's behaviour and preferences contribute to the CX. Therefore, the next question to answer is, what other factors influence a customer?

Customers are humans that have diverse opinions and expectations. This leads to the implication that a product and/or service might be perceived completely differently by two customers if they have both experienced the product and/or service objectively at the same level of quality. No CEM approach or the customer journey can prevent or cater for this. That is why the *360-degree view of the customer* is essential.

Definition 2.5 gives a comprehensive view of what this entails. It can be explained by placing the customer in the centre and viewing the customer from various angles. A key component of the *360-degree view of the customer* is the *customer sentiment*.

Definition 2.5 (360-Degree View of the Customer). *The 360-degree view of the customer includes anything that can be known about customers in order to improve the relationship to customers and users and give them an optimized and personalized experience” In other words it includes all that an enterprise can get to know the customer by (Mitra & Kawecki, 2016a).*

Customer sentiment is defined by Mitra & Kawecki (2016a) as the “differences between customers (specifically human users) in terms of their subjective perception of their experiences.” These perceptions can be in terms of the customer's diverse opinions and observations. It is dependent on the context, situation and arbitrary factors of the customer's surroundings. As stated by Richardson (2010), humans are not robots and they do not behave like robots. Humans are controlled by perception, emotions and their behaviours. Therefore, a customer's perception, emotion and (sometimes unexpected) behaviour has an inevitable effect on the CX. Verhoef *et al.* (2009) support this further by stating that a CX is holistic by nature and it involves the customer's cognitive, affective, emotional, social and physical responses to the enterprise and product and/or service.

The first aspect influencing a customer sentiment is *perception*. Perception as defined by Dodds (2016), is when the customer is guided by their emotional and psychological responses to stimuli presented when a customer is interacting with a product, service, brand or enterprise, as well as the environment in which the customer finds themselves at that particular instance. Gadman (1997) discusses on how an experience is triggered by the customer own perception. What a customer experiences is based on their *perception of what there is* rather than *what is actually there*. The

2.5 360-Degree view of the customer

words of [Russell \(2005\)](#) explains this in more depth when he says, “*All that I see, hear, taste, touch, smell and feel has been created from the data fed to me by my sensory organs. All I ever know of the world around are the images produced in the mind. I think I am seeing the tree ‘out there’, in the world around me. But all that I am actually experiencing is the image created in the mind.*” In other words, how customers perceive their experiences is based on how they have learnt to see the world and the triggers associated with it when they are experiencing something. These triggers are also programmed into human minds based on patterns, where the patterns determine how humans listen, read, observe and communicate. Perception is therefore highly subjective to the context and the situation. [Elliott & James \(1989\)](#) also point out that customers also have perceptions of what the enterprise will do, how it will be done and the characteristics of the contact person(s) from the enterprise.

The second aspect influencing customer sentiment is *emotions*. This includes the smells, sounds, taste and textures and environment as these elements have a direct impact on the emotions of the customer. It is important that the enterprise understand and grasp the feelings of its customers as well as the emotions associated with them ([Berry et al., 2012](#)). To influence the customer’s emotions, an enterprise should influence a customer’s affective system as well as the cognitive system. The enterprise will be able to do this by influencing a customer’s thinking or conscious mental processing. This will lead to customers engaging with the enterprise and generating moods and feelings in the customer which leads to the customer entering into a relationship with the enterprise ([Gentilea et al., 2007](#)). It is therefore crucial that the enterprise is careful with how they interact with customers. As [Elliott & James \(1989\)](#) points out, when a customer engages with an enterprise, they bring with them their own hopes, aims, dreams, tasks and intentions. The customers also experience feelings and moods during these interactions. The customer relates to the enterprise based on this. The enterprise should therefore have an understanding of the customers’ emotional and cognitive aspects.

The last aspect influencing customer sentiment is the *customer behaviour*. [Gentilea et al. \(2007\)](#), [Holbrook & Hirschman \(1982\)](#) and [Sukwadi \(2015\)](#) point out that aspects referring to the emotional and irrational side of customer behaviour need to be considered, because it is not only the rational ones that account for the whole experience. Therefore, customer behaviour plays a fundamental role in determining a customer’s preferences which influences their way of interacting with an enterprise to buy or use a particular product and/or service. [Puccinelli et al. \(2009\)](#) argue that the following topics provide a great insight into customer behaviour and the effect it has on CX:

1. *Goals, schema and information processing*: A customer wants to achieve a goal by purchasing a specific product and/or service. The goal depends on the customer’s needs and it influences the customer’s perceptions and behaviour.

2.5 360-Degree view of the customer

2. *Memory*: A customer's ability to encode, store and retrieve information will determine the customer's behaviour and the degree of influence it will have on the CX.
3. *Involvement*: The degree to which the customer is involved and engaged with the product and/or service, determines how the enterprise should engage with the customers.
4. *Attitudes*: A customer's attitude has an effect on the CX, depending on the context in which the customer finds himself and how long the customer's attitude endures over time.
5. *Affective processing*: Affect influences the customer's risk aversion and experimentation abilities. Therefore, it act as a motivator on whether a customer will engage and it influences how a customer will engage.
6. *Atmospheric*: A CX can be affected by the tangible and intangible design of the enterprise. The three factors that an enterprise should focus on are design, ambient and social.
7. *Attributions and choices*: Central to a customer behaviour are their attributions and choices. In other words, it is about the way a customer assigns a cause to an event and the effect the choices have.

There is more literature available on how customer behaviour influences the CX. For example, [Harris & Reynolds \(2003\)](#) investigated the effects that dysfunctional customer behaviour has in a service environment and [Söderlund \(1998\)](#) investigated what the relationship is between customer satisfaction and customer behaviour. Other publications include [Bamossy et al. \(2006\)](#); [Constantinides \(2004\)](#); [Ha & Perks \(2005\)](#); [Hernández et al. \(2010\)](#); [Jayawardhena \(2004\)](#); [Kristensen et al. \(1999\)](#) and [Solomon et al. \(2012\)](#).

Besides the customer sentiment, there are other components that also form part of the *360-degree view of the customer*. These components can be summarised as follows ([Gentilea et al., 2007](#); [Holbrook & Hirschman, 1982](#); [Verhoef et al., 2009](#)):

- *Environmental component*: The environment in which an enterprise interacts with the customers is important as it has an impact on the CX. The environment includes the social environment, stimuli of a customer, the communication and the content in which the interaction occurs. Customers can have a direct or indirect impact on each other and the stimuli of a customer affects the senses. Therefore it is important to look at the environment in which the interaction of a customer takes place as well as how it is done.
- *Lifestyle component*: The lifestyle of a customer is highly dependant on the value and belief systems of a customer. The value and belief system has an influence onto what degree the customers will involve and requires involvement from an enterprise. It also have an impact on how customers interact with the enterprise and the resources they will use.

2.6 Conclusion: Customer experience and the management of it

- *Relation component*: This component looks at customer relationships. It considers all relationships that can have an influence on the customer. This is the relationship with a customer and their social context, their relationship with other people, his/her ideal self and beyond.

From this it can be seen that it is crucial that the *360-degree view of the customer* needs to be considered when dealing with CX, because it is not only factors of the enterprise that have an effect on the CX, but also the customer sentiment, the environment, the lifestyle and the relationship components. If an enterprise manages to incorporate all the *360-degree view of the customer* components, they should be able to successfully implement the CEM approach in the business model.

2.6 Conclusion: Customer experience and the management of it

In this chapter the Customer Experience was discussed together with the Management of it. A thorough literature study was conducted in order to understand how a system can be created to incorporate the customer and their experience as well as how it should be managed. By doing this, a demonstrator can be constructed in order for it to be implemented via the trip planner. The customer's experience and the management of it was discussed based on the following aspects.

The first aspect was to determine what is a customer and to see their different roles in a supply chain. For the purpose of this study, the customer is the end-user/consumer.

The second aspect was to determine what a Customer Experience is. This was done based on defining what this is, understanding what a customer wants to experience, metrics that can be used for CX and the importance of CX for an enterprise.

The third aspect was to determine how to manage a CX. There are two popular strategic management approaches, namely Customer Relationship Management and Customer Experience Management. For the purpose of this study CEM will be used as it is a proactive approach in which the enterprise follows a customer-centric business model.

The fourth aspect was to determine what the customer journey is and the role it plays in the CX domain. The customer journey enables an enterprise to capture a customer's entire journey together with all the experiences at the different interaction points.

The last aspect was to determine what other factors influence a customer. By considering the 360-degree view, the customer is placed in the centre and viewed from various angles.

The next chapter will present a literature study on Big Data and the Analytics of it

Chapter 3

Big Data and its analytics

In the previous chapter the first part of the literature study, namely Customer Experience and the Management of it, was discussed. In this chapter the second part of the literature study will be discussed based on Big Data and its Analytics. A study needs to be conducted on this topic as the demonstrator will be using data analytics to determine how a customer's experience should be managed by the trip planner.

Therefore, to grasp what Big Data is and how the analytics can be used on it, the literature study will be conducted on the following. Big Data will be unpacked by looking at what Big Data actually entails, the importance of understanding what Big Data is and lastly the current issues related to Big Data. Once the Big Data term has been understood, the analytical part will be discussed by defining what Big Data Analytics is and what methodology should be used for it.

3.1 Big Data

As mentioned in Section 1.1 we live in a 'digital world' in which technology becomes more prominent and replaces the old ways of doing things. Due to this, more data gets stored on hardware and in the cloud. The result of this is that big sets of data are available but many do not know what to do with it. As pointed out by [Gandomi & Haider \(2015\)](#), due to the sudden rise in data much has been left unprepared, even to the extent that they fear it. In this section, the question as to what Big Data is will be unpacked.

3.1.1 Overview of Big Data

Big Data is a well-known term but it is not understood by many. In order to understand what Big Data (BD) is, it is good to know when and how this term emerged. From literature it is clear that the term emerged in the past century. There is documentation available on how *data* has emerged over the years. One such document is by Bernard Marr who has written an article about the history of data and how it has led the world to the *data age* ([Marr, 2017](#)). He looks at how humans stored and analysed data in the years *before Christ* and how it has built up from there. But there is more proof that BD is a term which emerged over the past century.

According to the [Oxford English Dictionary \(1980\)](#) the first reference to the word 'BD' was in the year 1980. [Jifa & Lingling \(2014\)](#) add on to this as they did a short history study on BD from the year 1994 to 2012. [Jifa & Lingling \(2014\)](#) point out that the first published work around BD was done by Fremont Rider. [Rider \(1994\)](#) discussed his concern around the rapid rate at which the size of university libraries is increasing. He estimated that Yale University will have roughly 0.2

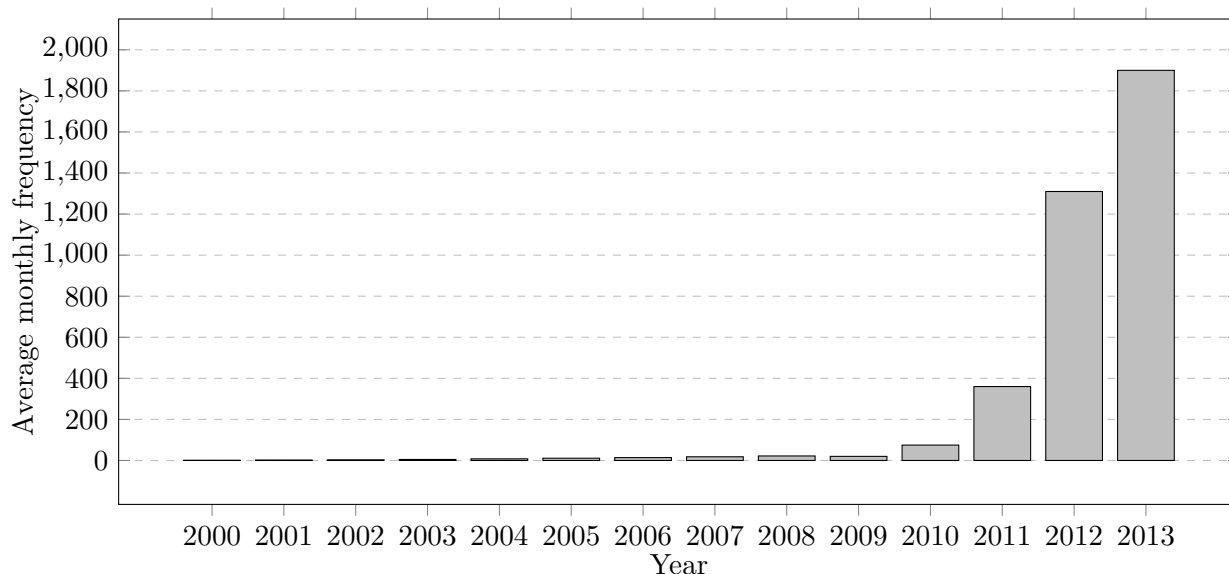


Figure 3.1: Frequency distribution of documents containing the term ‘Big Data’ in ProQuest Research Library (Gandomi & Haider, 2015)

billion volumes in the shelves by 2040 at a rate where the libraries double in size every 16 years. All these volumes will contribute to nearly three quarters of a million catalogue drawers (Coney, 1944). Therefore, it can be said that BD made its appearance in the academic world in the late twentieth century.

However, BD actually came to the attention of enterprises at the start of the digital transformation which led to digitisation in the twenty-first century (i-SCOOP, 2016). Due to this digital transformation enterprises had more data stored about their customers and the enterprise itself, either on their hardware or on the cloud which lead to the formation of BD. This phenomenon started in the past decade and a study done by Gandomi & Haider (2015) shows how the term became widespread in 2011. They also argue that the contribution to this is due to initiatives such as IBM and other leading technology enterprises who invested to help the market perform analytical techniques and methods on BD sets.

This can be further proved by the extensive research that was done in the BD domain. Figure 3.1 confirms this as it shows the frequency distribution of documents in the *ProQuest Research library* that contain the term ‘BD’ from the year 2000 to 2013.

By looking at Figure 3.1, it is clear that BD is a ‘BD’ domain on its own. Considering the term itself it can be seen that the first characteristics of BD is *size*, but what else categorises a BD set? Therefore, to answer this, another important question that needs to be answered is, when does one know that a dataset can be classified as a BD set? According to Anderson & Semmelroth (2015), BD refers to “sets of data that are far too massive to be handled with traditional hardware.” In other words, those sets of data that are too large to be analysed in traditional ways and advanced

hardware or software is needed. To understand BD in a broader sense, it is important to look at its definitions.

In literature, BD has been defined and redefined in various ways and as can be seen from the widespread of documentation available in Figure 3.1, a lot of studies have been done in and around the BD domain. The first authoritative definition of BD which became the industry standard, is defined by Laney (2001). He defined BD according to three Vs, where these three Vs are *Volume*, *Variety* and *Velocity*. The three Vs represent the following datasets' characteristics (Gandomi & Haider, 2015; Harvey, 2017; Katal *et al.*, 2013; Laney, 2001; McAfee & Brynjolfsson, 2012; Russom, 2011; Sagioglu & Sinanc, 2013):

1. *Volume*: It refers to the *magnitude* of data in terms of its breadth and length. The size can be one terabyte to one petabyte, but it can also be quantified by counting the number of records, transactions, tables or files. It is important to note that the size of the data gathered, is relative to how the term 'large' will be observed in the eyes of the enterprise.
2. *Variety*: It refers to the wide variety of *formats* in which data is captured. It can be further categorised into three main categories which are: *structured*, *semi-structured* or *unstructured* data. It also indicates the diversity of data in terms of its sources, data types and entities.
3. *Velocity*: It refers to the *rate* at which data is gathered and/or captured. In other words the speed at which it is streamed into the enterprise. It can also be quantified as the frequency of data generation or delivery.

Besides the three dimensions stipulated above, there are three more characteristics or V dimensions of BD. These dimensions include (i) *veracity*, which refers to the unreliability of data and can be seen as a subset of the variety dimension, (ii) *variability* which refers to the variation of the rates and can be seen as a subset of velocity and (iii) *value* which refers to the economic growth and insights of the data. Besides these six characteristics, more characteristics can be found in literature. Two examples are the work done by Diadmin (2015) and Vorhies (2014).

Based on all these characteristics, how can one define the term 'BD'? De Mauro *et al.* (2015) conducted a study in which they investigated and analysed 1 473 conference papers and articles purely to determine a consensual definition of BD. This definition will be used for the purpose of this study and is shown in Definition 3.1 .

Definition 3.1 (Big Data). *Big Data represents the information assets characterized by such a high volume, velocity and variety to require specific technology and analytical methods for its transformation into value (De Mauro et al., 2015).*

3.1.2 Importance of Big Data

Now that a better understanding has been provided on how BD emerged and what it is, the importance of it needs to be unpacked. This will be done to understand why BD matters to an enterprise.

The first contributor is that BD leads to *value creation*. Due to the fact that there is a big rise in the development of social media and the Internet of Things (IoT), enterprises have the ability to collect more information about their customers and customer data is increasing at an exponential factor. Therefore, data opens up various doors for enterprises to expand their limits on how they can predict and understand their customers. In other words, BD will have an ever-increasing impact on an enterprise and customers alike because they now have the potential to create value from all the data available (Chen *et al.*, 2014).

To further understand the impact of how BD leads to value creation, it is important to understand the dimensions in which BD creates it. Manyika *et al.* (2011) did a study based on five domains to get behind the idea of how BD creates value for an enterprise and to grasp the potential of BD. They have done it based on the following domains: (i) the United States health care domain, (ii) the European Union (EU) domain of public sector administration, (iii) the United States retail domain, (iv) the global personal location data domain, and (v) the global manufacturing domain. Based on this study they were able to identify the following five dimensions that leverage BD in order to create value for an enterprise. The five dimensions are as follows (Kaisler *et al.*, 2013; Manyika *et al.*, 2011):

1. **Creating transparency:** BD enables an enterprise to make the data more transparent for sharing amongst departments as well as accessible by relevant stakeholders. Therefore, BD is openly available for business and functional analysis.
2. **Enabling experimentation to discover needs, expose variability, and improve performance:** BD gives the ability and support for an enterprise to perform various experiments in order to provide more accurate and detailed performance data. This can be done in real time and non-real time. It gives the managers the opportunity to analyse performance variability and understand why variations are occurring.
3. **Segmenting populations to customise actions:** BD creates the opportunity for enterprises to tailor their products and/or services according to the customers' needs by using highly specified segmentations, which is done based on the customers' information.
4. **Replacing/supporting human decision-making with automated algorithms:** BD enables the enterprises to improve their decisions, minimise the risks associated with them and discover unknown insights which can be valuable.

5. **Innovating new business models, products and services:** by facilitating computer-assisted innovation and analysing the BD available to the enterprise, enterprises are able to come up with new business models, products and/or services.

Based on these five dimensions [Wamba *et al.* \(2015\)](#) did a review of articles to see the quantity of articles that have been published in which some of these dimensions are discussed. Table 3.1 gives an indication of how many articles have been published and Table 3.2 gives the reference to these articles. From this study, they have identified that the key dimension of value creation is that *BD enables the replacement and supporting of human decisions by using automated algorithms*.

Table 3.1: Quantity of articles published based on the types of value creation dimensions of Big Data ([Wamba *et al.*, 2015](#))

Dimension	Quantity
1. Creating transparency	17 articles
2. Enabling experimentation to discover needs, expose variability, and improve performance	28 articles
3. Segmenting populations to customize actions	20 articles
4. Replacing/supporting human decision-making with automated algorithms	35 articles
5. Innovating new business models, products, and services	25 articles

The second contributor to the importance of BD is the direct link that exists between BD and analysis of it. Due to this link BD has evolved from a business initiative to a business imperative ([Zikopoulos *et al.*, 2012](#)). In other words, at first BD created an opportunity for enterprises to give them the advantage to be a leading enterprise that creates more value, but for an enterprise to survive it is crucial to use algorithms and techniques on BD in order to get value from it.

Based on this it is clear that BD is an important aspect as it creates value for an enterprise across five dimensions, but to create this value it is crucial to analyse it in order to survive in the digital world today. The analysis of BD will be discussed in more depth in Section 3.2.

3.1.3 Challenges of Big Data

By understanding what BD is and its importance, it is also important to grasp the challenges enterprises face today with BD.

BD has been present for more than a decade but enterprises have really only given recognition to its value over the past decade as discussed in Section 3.1.1. As stated by [Katal *et al.* \(2013\)](#), BD goes beyond the storing of it on traditional hardware. It is about what an enterprise does with the BD available to them. Therefore, since BD is a relative new concept in enterprises and it is not the traditional hardware data as enterprises know it, it brings a few challenges.

Table 3.2: Articles published based on the types of value creation dimensions of Big Data (Wamba *et al.*, 2015)

Dimension	References
1. Creating transparency	Ann Keller <i>et al.</i> (2012); Beath <i>et al.</i> (2012); Boyd & Crawford (2012); Brown <i>et al.</i> (2011); Bughin <i>et al.</i> (2011); Chen <i>et al.</i> (2012); Cole <i>et al.</i> (2012); Davenport <i>et al.</i> (2012); Fisher <i>et al.</i> (2012); Griffin (2012); Huwe (2012); LaValle <i>et al.</i> (2011); McAfee & Brynjolfsson (2012); Schadt <i>et al.</i> (2010); Smith <i>et al.</i> (2012); Tankard (2012); Wagner (2012)
2. Enabling experimentation to discover needs, expose variability, and improve performance	Allen <i>et al.</i> (2012); Anderson & Blanke (2012); Ann Keller <i>et al.</i> (2012); Beath <i>et al.</i> (2012); Boja <i>et al.</i> (2012); Boyd & Crawford (2012); Brinkmann <i>et al.</i> (2009); Brown <i>et al.</i> (2011); Bughin <i>et al.</i> (2010); Chen <i>et al.</i> (2012); Cole <i>et al.</i> (2012); Davenport <i>et al.</i> (2012); Demirkan & Delen (2013); Fisher <i>et al.</i> (2012); Havens <i>et al.</i> (2012); Huwe (2012); Johnson (2012a); Kolker <i>et al.</i> (2012); LaValle <i>et al.</i> (2011); McAfee & Brynjolfsson (2012); Schadt <i>et al.</i> (2010); Siemens & Long (2011); Soares (2012); Sobek <i>et al.</i> (2011); Strawn (2012); Tankard (2012); Wagner (2012); White (2012)
3. Segmenting populations to customise actions	Ann Keller <i>et al.</i> (2012); Beath <i>et al.</i> (2012); Boyd & Crawford (2012); Bughin <i>et al.</i> (2010); Chen <i>et al.</i> (2012); Cole <i>et al.</i> (2012); Davenport <i>et al.</i> (2012); Demirkan & Delen (2013); Fisher <i>et al.</i> (2012); Griffin (2012); Highfield (2012); LaValle <i>et al.</i> (2011); McAfee & Brynjolfsson (2012); Schadt <i>et al.</i> (2010); Siemens & Long (2011); Smith <i>et al.</i> (2012); Soares (2012); Sobek <i>et al.</i> (2011); Tankard (2012); Wagner (2012)
4.Replacing/supporting human decision-making with automated algorithms	Allen <i>et al.</i> (2012); Anderson & Blanke (2012); Ann Keller <i>et al.</i> (2012); Beath <i>et al.</i> (2012); Boja <i>et al.</i> (2012); Boyd & Crawford (2012); Brown <i>et al.</i> (2011); Bughin <i>et al.</i> (2010, 2011); Chen <i>et al.</i> (2012); Cole <i>et al.</i> (2012); Dansion & Griffin (2012); Davenport <i>et al.</i> (2012); Demirkan & Delen (2013); Fisher <i>et al.</i> (2012); Gehrke (2012); Griffin (2012); Huwe (2012); Johnson (2012a,b); Kolker <i>et al.</i> (2012); Lane (2012); LaValle <i>et al.</i> (2011); McAfee & Brynjolfsson (2012); Meijer (2011); Ohata & Kumar (2012); Schadt <i>et al.</i> (2010); Siemens & Long (2011); Smith <i>et al.</i> (2012); Soares (2012); Sobek <i>et al.</i> (2011); Strawn (2012); Tankard (2012); Wagner (2012); White (2012)
5. Innovating new business models, products, and services	Ann Keller <i>et al.</i> (2012); Beath <i>et al.</i> (2012); Boyd & Crawford (2012); Brown <i>et al.</i> (2011); Bughin <i>et al.</i> (2010, 2011); Chen <i>et al.</i> (2012); Cole <i>et al.</i> (2012); Dansion & Griffin (2012); Davenport <i>et al.</i> (2012); Demirkan & Delen (2013); Fisher <i>et al.</i> (2012); Gehrke (2012); Griffin (2012); Huwe (2012); Johnson (2012a); Kolker <i>et al.</i> (2012); LaValle <i>et al.</i> (2011); McAfee & Brynjolfsson (2012); Ohata & Kumar (2012); Siemens & Long (2011); Soares (2012); Strawn (2012); Tankard (2012); Wagner (2012)

From literature the BD challenges can be categorised into five groups, as follows:

1. **Data Challenges:** It relates back to the challenges enterprises face with the characteristics of BD which are volume, variety, velocity, veracity, variability, and value.
2. **Process Challenges:** It relates back to the challenges enterprises encounter when the BD undergoes processing and the processes required for it.
3. **Management Challenges:** It relates back to challenges of how and where the management of the enterprises have control over. It includes the privacy, security, understanding and analysis of BD.
4. **Technical Challenges:** It relates back to challenges enterprises have with its technical capability to handle, process and store BD.
5. **Organisational Challenges:** It relates back to challenges on how the enterprises function and the control they have over BD.

These challenges can then be further unpacked as can be seen in Table 3.3. These challenges are only some BD challenges based on papers found in literature. There might be more challenges that can be added to this list. But, it is important for enterprises to acknowledge and identify these challenges in order to determine ways of overcoming them to extract the maximum value from their BD sets. As stated in Section 3.1.2, enterprises should have the ability to use the BD to their advantage in order to survive in this ‘digital world’.

Table 3.3: Summary of Big Data challenges

Challenge	Description	References
1. Data Challenges		
1.1 <i>Volume:</i> Dealing with data growth	The ever-growing volume of BD can be costly and difficult for enterprises.	Harvey (2017) ; Sivarajah <i>et al.</i> (2017)
1.2 <i>Variety:</i> Diversity of Data	The large volume of BD is captured in diverse formats and from diverse sources that can lead to inconsistency.	Sivarajah <i>et al.</i> (2017)
1.3 <i>Veracity:</i> Understanding of Data	Discrepancies can be found in the BD which makes the BD untrustworthy.	Sivarajah <i>et al.</i> (2017)
1.4 <i>Velocity:</i> High speed of Data	The rate at which BD is entering and adjusting to the current set of BD has an impact on the enterprise infrastructure.	Sivarajah <i>et al.</i> (2017)

Table 3.3 continues on next page

Challenge	Description	References
1.5. <i>Variability: Meaning of Data</i>	The ability to adjust to the rapid rate at which the meaning of BD is changing constantly.	Sivarajah <i>et al.</i> (2017)
1.6. <i>Value: Extracting Knowledge from Data</i>	To generate value from BD, the knowledge of BD should be extracted in the correct manner.	Sivarajah <i>et al.</i> (2017)
2. Process Challenges		
2.1. <i>Acquiring Data</i>	The process of acquiring BD at the current growth rate and from diverse sources can be daunting.	Katal <i>et al.</i> (2013); Labrinidis & Jagadish (2012); Sivarajah <i>et al.</i> (2017)
2.2. <i>Storing Data</i>	The ability to store the ever-growing volume of BD without discarding relevant information for value creation.	Kaisler <i>et al.</i> (2013); Katal <i>et al.</i> (2013); Labrinidis & Jagadish (2012); Sivarajah <i>et al.</i> (2017)
2.3. <i>Extracting and Cleaning of Data</i>	The process of cleaning and extracting BD in such a way that it will result in a big impact and add great value.	Harvey (2017); Labrinidis & Jagadish (2012); Sivarajah <i>et al.</i> (2017)
2.4. <i>Aggregating and Integrating Data</i>	The process of aggregation and integration of clean BD in order to enable BD analysis and modelling on it.	Harvey (2017); Labrinidis & Jagadish (2012); Sivarajah <i>et al.</i> (2017)
2.5. <i>Analysis of data</i>	Determining and applying the correct analysis methodology for the right type of BD set.	Chen <i>et al.</i> (2014); Kaisler <i>et al.</i> (2013); Katal <i>et al.</i> (2013); Labrinidis & Jagadish (2012); Sivarajah <i>et al.</i> (2017)
2.6. <i>Making data understandable</i>	Presenting the results for decision-makers in such a way that it can be visualised, read and understood.	Chen <i>et al.</i> (2014); Labrinidis & Jagadish (2012); Sivarajah <i>et al.</i> (2017)
3. Management Challenges		
3.1. <i>Leadership</i>	Leadership skills are crucial as managers need to identify opportunities with BD and respond to it in the correct way.	Kaisler <i>et al.</i> (2013); McAfee & Brynjolfsson (2012)
3.2. <i>Skills</i>	Management should identify the employees' skill set and talent(s) that they possess to get the right team working with the BD.	Harvey (2017); Katal <i>et al.</i> (2013); McAfee & Brynjolfsson (2012)

Table 3.3 continues on next page

Challenge	Description	References
3.3. <i>Energy Management</i>	The accessibility and expandability of BD should be ensured by using appropriate management mechanisms and system-level power consumption.	Chen et al. (2014)
3.4. <i>Decision-making</i>	The enterprise needs to give an opportunity to the right people to make relevant decisions based on the available BD.	McAfee & Brynjolfsson (2012)
3.5. <i>Privacy and Security</i>	The protection of BD by privacy policies and security is a major concern. The right BD policies should be in place to protect the BD at all time and cost.	Harvey (2017) ; Jagadish et al. (2014) ; Katal et al. (2013) ; Sivarajah et al. (2017) ; Wamba et al. (2015)
3.6. <i>Data Governance</i>	Governance should be in place to control the BD access and sharing of it.	Katal et al. (2013) ; Sivarajah et al. (2017) ; Wamba et al. (2015)
3.7. <i>Data Ownership</i>	Determining who owns and has a right to the BD.	Jagadish et al. (2014) ; Sivarajah et al. (2017)
4. Technical Challenges		
4.1. <i>Technology</i>	How technology can be used in the management, processing and analysis of BD.	Boyd & Crawford (2012) ; Cole et al. (2012) ; McAfee & Brynjolfsson (2012) ; Wamba et al. (2015)
4.2. <i>Time</i>	Support of real-time analysis and the enterprise's ability to keep up to date with BD.	Kaisler et al. (2013) ; Jagadish et al. (2014)
4.3. <i>Heterogeneous Data</i>	The ability to process and analyse the different forms and quality of BD (structured and unstructured BD).	Katal et al. (2013) ; Chen et al. (2014) ; Jagadish et al. (2014) ; Sivarajah et al. (2017)
5. Organisational Challenges		
5.1. <i>Cooperation</i>	To harvest the potential of BD, cooperation between experts in different fields in the enterprise is required.	Chen et al. (2014)
5.2. <i>Competitive differentiation through innovation</i>	BD brings innovative opportunities but an enterprise should know how to keep its competitiveness.	LaValle et al. (2011)

Table 3.3 continues on next page

Challenge	Description	References
5.3. <i>Cost and Operational Expenditure</i>	The enterprise should keep the rising demand for BD processing in mind, while at the same time being cost-effective when using BD and be encouraged to grow in the revenue margin.	LaValle et al. (2011) ; Sivarajah et al. (2017)
5.4. <i>Company culture</i>	The enterprise should adapt to data-driven decision culture.	McAfee & Brynjolfsson (2012)
5.5. <i>Industry Structure</i>	The enterprise should be able to understand the business value and realisation of BD.	Wamba et al. (2015)
End of Table 3.3		

Currently, most enterprises only store these BD sets but have limited resources or no clue on what to do with it. According to [Zikopoulos et al. \(2012\)](#), “no one has ever delivered a single penny of value out of merely storing data (excluding cloud hosts) and most enterprises do not have the ability to do more than just store these large sets of data.” Therefore, a methodology should be used to understand how to create value out of these datasets. Big Data Analytics is such a methodology which will be discussed in the next section.

3.2 Big Data Analytics

In the previous section the question of what BD is, was answered. This was done based on its three main dimensions (Volume, Variety and Velocity), together with the importance of BD and the challenges of BD faced by enterprises. However, an enterprise will not be able to make a penny out of their BD if they do nothing with it. Therefore, the next question to answer is: How can value be extracted from BD? This section will investigate how it can be done by the use of analytics.

3.2.1 Overview of Big Data Analytics

Analytics is required in order to extract value from BD. Without analytics, the enterprises will not be able to do it and it will only waste storage space and money. However, to do analytics is not as simple as a click of a button. Therefore, it is important to understand what Big Data Analytics entails.

The phenomenon behind BDA and how it processes data to create value for an enterprise can be explained by the *Data-Information-Knowledge-Wisdom (DIKW) hierarchy*, also known as the ‘Knowledge Hierarchy’, the ‘Information Hierarchy’ or the ‘Knowledge Pyramid’ as can be seen in [Figure 3.2](#). The DIKW hierarchy represents how data, information, knowledge and (in some cases) wisdom are conceptualised with respect to one another as well as the processes involved for the transformation from one level to another ([Rowley, 2006](#)).

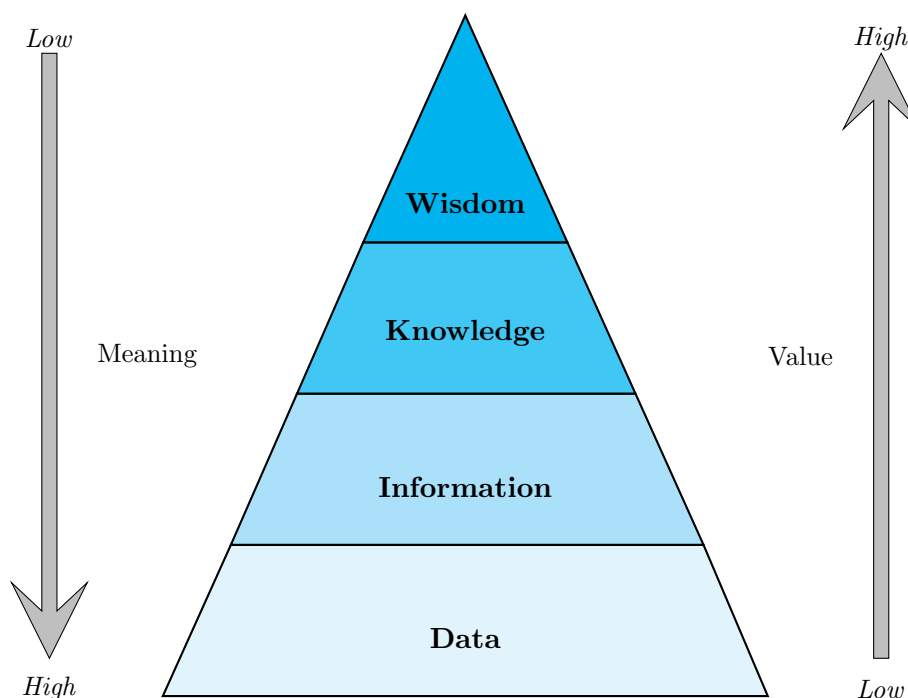


Figure 3.2: DIKW hierarchy (Chaffey & Wood, 2005)

The DIKW hierarchy first made its appearance in T.S. Eliot’s poem *The Rock*, that contains the line “*Where is the wisdom that we have lost in knowledge? Where is the knowledge that we have lost in information?*” (Eliot, 1934). From here on Ackoff (1989) defined the hierarchy by adding levels of understanding and intelligence. He also focused on the processes required for transformation between the levels. Chaffey & Wood (2005) added the meaning and value dimensions to the hierarchy as can be seen in Figure 3.2.

Therefore, the DIKW hierarchy represents that data can be transformed into information. From this information, knowledge can be gained and from the knowledge, wisdom can be created. The higher the level, the more value it will create but the less meaning it will have. By looking at this hierarchy, it is clear how one can create value from data. A way to achieve this is by performing analytics on the data, in other words Big Data Analytics (BDA).

According to Chen *et al.* (2015) BDA is an array of powerful analytical techniques and information technologies used by an enterprise to capitalise on the promise of creating value from BD. Based on this explanation, BDA can be formally defined as can be seen in Definition 3.2. The definition is a combination of definitions from Chen *et al.* (2015); Elgendy & Elragal (2014); Waller & Fawcett (2013). The advanced technologies that can be used to implement BDA include data management, open-source programming, statistical analysis, in-memory computing and visualisation tools for presenting the insights. The insights generated by BDA include hidden patterns, anomalies, unknown correlations and other actionable insights.

Definition 3.2 (Big Data Analytics). *Big Data Analytics is the process of examining Big Data by using advanced technologies in order to uncover useful information to help with making better decisions across the business processes.*

BDA can be further explained as follows. *An enterprise that has the ability to convert data into insights and intelligence effectively at the correct time and place throughout various levels in the enterprises has mastered the field of BDA* (Rich, 2015). Therefore, an enterprise should have the ability to generate insights which are non-trivial, previously unknown, implicit and potentially useful. Furthermore, they should understand that these insights have a direct impact on deciding or manipulating the current business strategy (Mahmood & Afzal, 2013).

3.2.2 Big Data Analytics methodology

Now that a definition has been provided on BDA, how does one perform this? There are numerous methodologies available for BDA, but for the purpose of this study a methodology based on how the USMA (2017) group perceives it will be used. Definition 3.3 is used for this methodology and it differs slightly from Definition 3.2, but the fundamental underlying the value creation from BD is present.

Definition 3.3 (Big Data Analytics Methodology). *Big Data Analytics is the entire methodology that is utilised to analyse Big Data sets in order to create value for an enterprise* (USMA, 2017).

To fully understand the BDA methodology proposed by the USMA (2017) research group, it will be explained by using the BDA Methodology Framework.

3.2.2.1 Framework for the methodology

The USMA (2017) research group defined a framework which explains and grasps every aspect of the proposed BDA Methodology. The BDA Methodology Framework can be seen in Figure 3.3.

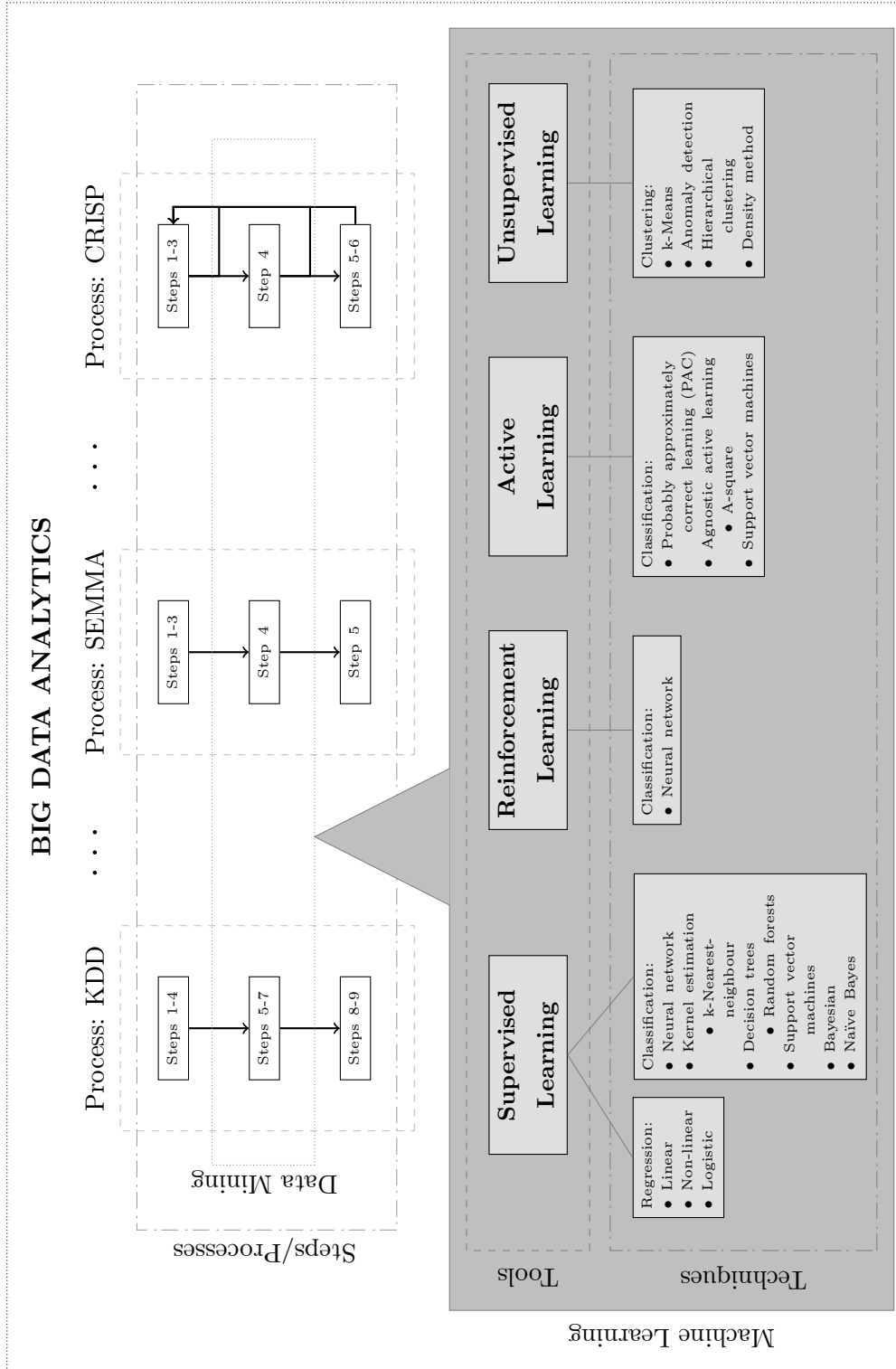


Figure 3.3: Big Data Analytics (USMA, 2017)

From this framework, the BDA methodology can be explained as follows. It is a methodology that is composed of various *knowledge management processes*. These processes are used to gain knowledge from BD in a structured way by undergoing several phases or steps.

The phases or steps, include the *data mining* phase or step which is responsible for the analytical part. The data mining phases or steps are the process of discovering patterns in data (Witten *et al.*, 2016). The tools and techniques used to perform these data mining phases or steps are done by *machine learning*. Machine Learning (ML) as defined by Faggella (2018) is “the science of getting computers to learn and act like humans do, and improve their learning over time in autonomous fashion, by feeding them data and information in the form of observations and real-world interactions.” In other words, ML is when a computer is programmed to learn from a set of input data fed to it (Ben-David & Shalev-Shwartz, 2014). By learning, the computer performs a process of converting experience into expertise or knowledge that can be used to generate value.

Therefore, BDA is a methodology that consists of knowledge management processes. These processes contain data mining phases or steps, which are responsible for performing ML, which consists of tools and techniques. By doing this, an enterprise can create value from their BD sets. The different aspects of the framework will now be discussed in more depth.

3.2.2.2 Knowledge Management processes

As explained earlier, the BDA methodology consists of various knowledge management processes. These processes are used to gain knowledge from BD in a structured way by undergoing various phases or steps. The three common processes used are (i) *Knowledge Discovery in Databases* (KDD), (ii) *Sample, Explore, Modify, Model and Assess* (SEMMA) and (iii) *Cross-Industry Standard Process for Data Mining* (CRISP-DM). A comparison of these three processes was done by Azevedo & Santos (2008).

The first process, KDD is defined by Fayyad *et al.* (1996b) as the “non-trivial process of identifying valid, novel, potentially useful and ultimately understandable patterns in data.” In other words, it is a process that extracts high-level knowledge from low-level data by using *data mining* methods, where the dataset requires preprocessing, sub-sampling and transformation (Fayyad, 1996). This process consists of the following five phases (Azevedo & Santos, 2008):

1. *Selection*: Creating or focusing on a dataset on which discovery will be performed.
2. *Preprocessing*: Cleaning and preprocessing of data to obtain consistent data.
3. *Transformation*: Using dimensionality reduction or transformation methods.
4. *Data Mining*: Search for patterns of interest in a particular representational form.
5. *Evaluation/Interpretation*: Evaluate or interpret the mined data.

The second process, SEMMA is developed by the *SAS Institute* and it forms part of the *SAS Enterprise Miner* suite whose core task is data mining ([Mariscal et al., 2010](#)). The SEMMA consists of a cycle with five stages ([Azevedo & Santos, 2008](#)):

1. *Sample*: The sampling of a dataset for significant information.
2. *Explore*: Explore the datasets for unanticipated trends and anomalies.
3. *Modify*: Create, select and transform variables to focus the model selection process.
4. *Model*: Search for a combination to give a reliable prediction of a desired outcome.
5. *Assess*: Evaluates the usefulness and reliability of the findings.

The third process, CRISP-DM is a hierarchical process model that consists of a set of tasks at four levels of abstraction, namely phase, generic task, specialised task and process instance ([Chapman et al., 2000](#)). CRISP-DM consist of a cycle with five stages ([Azevedo & Santos, 2008](#)):

1. *Business understanding*: Understanding the data mining project from a business perspective.
2. *Data understanding*: Familiarise with the data, identify the data quality problems and discovering of first insights.
3. *Data preparation*: Activities to construct the final dataset.
4. *Modelling*: Selection and applying various modelling techniques and calibrate parameters for optimal values.
5. *Evaluation*: Evaluation of the model(s) and execution steps are reviewed against the objectives.
6. *Deployment*: Organisation and presentation of the knowledge of the data.

The above-mentioned processes are the three common processes used in which data mining takes place. But there are more processes available in the academic domain and practice. These processes are mostly derived from KDD and CRISP-DM can also be used as determined by [Mariscal et al. \(2010\)](#). These processes can be seen in [Figure 3.4](#).

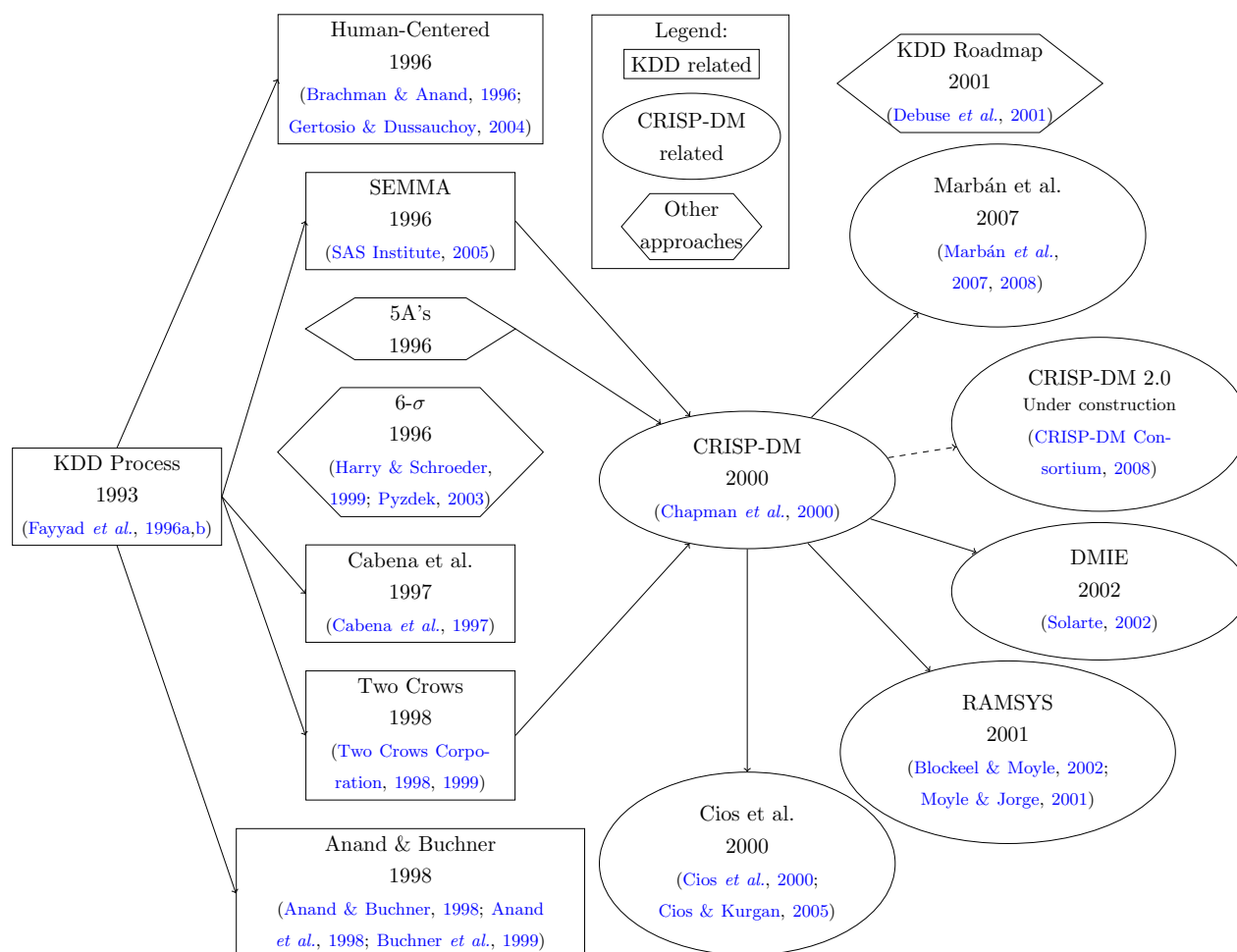


Figure 3.4: The evolution of knowledge management processes (Mariscal *et al.*, 2010)

3.2.2.3 Machine Learning

The knowledge management processes consist of several phases or steps, including the data mining phase or step. The data mining phase or step is the process of discovering patterns in data (Witten *et al.*, 2016). The tools and techniques used to perform these data mining phases or steps are done by ML. As mentioned before, ML is when a computer performs a set of processes on a set of data to generate value for a business, where the computer has been programmed in such a way that it can learn from a set of input data. As stated by Mitchell (1997) it uses concepts and results from a lot of fields which include statistics, artificial intelligence, philosophy, information theory, biology, cognitive science, computational complexity, and control theory.

Tools:

The four types of tools used for ML are (i) *supervised learning*, (ii) *reinforcement learning*, (iii) *active learning* or (iv) *unsupervised learning*. The two most common tools used are supervised and unsupervised learning (Ben-David & Shalev-Shwartz, 2014).

The different tools can be explained as follows (Ben-David & Shalev-Shwartz, 2014; Jiawei *et al.*, 2012; Schmidhuber, 2015):

- i *Supervised Learning* (SL): Test datasets are provided for training to develop the desired output. For example, when data about customers should be sorted based on their accommodation preference, test data with the category of whether the preference is hotel or guest house is given, to analyse the dataset.
- ii *Reinforcement Learning* (RL): The learner performs trial-and-error interactions to determine the desired output or behaviour within a dynamic environment. In other words, the learner's actions are evaluated by reward and penalty signals to determine what actions will yield the best reward.
- iii *Active Learning* (AL): The user partakes in an active role by interacting with the environment during the learning process. For example, the user should classify the accommodation preference chosen by the learner in order to understand what the customers' accommodation preference(s) will be.
- iv *Unsupervised Learning* (UL): No test datasets are provided but the goal is to come up with ways on how to cluster the data correctly. For example, here the analysis of customers' preference of accommodation type is sorted based on the type of accommodation but the learner should detect by itself whether the customer prefers a hotel or guest house.

Techniques:

The tools are composed of various data mining techniques as stipulated in Figure 3.3. For the each of the tools, either one or a combination of techniques can be used to perform ML. The technique used for the data mining phase is chosen based on the dataset that needs to be analysed. If necessary, more than one technique can be used on a dataset. The data mining techniques can be grouped into three data modelling types which are as follows:

- i *Classification*: It is predictive data analysis in which the data is divided into predefined classes with associated class labels.
- ii *Regression*: It is the process of exploring relationships in a dataset, between a dependent and one or more independent variable(s).
- iii *Clustering*: It is when data is divided into multiple clusters based on certain characteristics.

Table 3.4 represents the various BDA techniques, the group of tools under which it falls and sources that can be used to view how these techniques are used and what it contains. The definition and application for these various techniques can be seen in Tables 3.5, 3.6 and 3.7.

Table 3.4: Summary of Big Data Analytics techniques (USMA, 2017)

Techniques	Tools				Sources
	SL	RL	AL	UL	
1. CLASSIFICATION					
1.1 Decision Trees:					
1.1.1 Decision Trees	✓				Janikow (1998); Kim <i>et al.</i> (2006); Kotsiantis (2007); Larose (2014); Paramasivam <i>et al.</i> (2014); Rokach & Maimon (2014)
1.1.2 Classification and regression trees (CART)	✓				Breiman <i>et al.</i> (1984); Larose (2014); Lawrence & Wright (2001); Rokach & Maimon (2014); Steinberg & Colla (2009); Steynberg (2016); Wu <i>et al.</i> (2008)
1.1.3 C4.5 algorithm	✓				Hssina <i>et al.</i> (2014); Kotsiantis (2007); Larose (2014); Quinlan (2014); Wu <i>et al.</i> (2008)
1.1.4 Random Forest	✓				Breiman <i>et al.</i> (1984); Hastie <i>et al.</i> (2009); Jiawei <i>et al.</i> (2012)
1.2. Support Vector Machines (SVM)	✓		✓		Coussement & Van den Poel (2008); Huang <i>et al.</i> (2007); Jansen (2007); Kotsiantis (2007); Rechenhth (2014); Tomar & Agarwal (2013); Vapnik (1999); Wu <i>et al.</i> (2008)
1.3. Neural Networks	✓	✓			Bloom (2004); Chan (2005); Hastie <i>et al.</i> (2009); Izenman (2008); Jiawei <i>et al.</i> (2012); Kuo <i>et al.</i> (2006); Linoff & Berry (2011); Paliwal & Kumar (2009a); Petroulakis & Miaoudakis (2007)
1.4. Naive Bayes Vector	✓				Jiawei <i>et al.</i> (2012); Li (2015); Rechenhth (2014); Wu <i>et al.</i> (2008)
1.5. k-Nearest Neighbour	✓				Kotsiantis (2007); Larose (2014); Li (2015); Rechenhth (2014); Salkind (2007); Wu <i>et al.</i> (2008)
1.6. Rule-Based Classifiers	✓				Cakir & Aras (2012); Ishibuchi & Yamamoto (2005); Ishibuchi <i>et al.</i> (2005); Lawrence & Wright (2001)
2. CLUSTERING					
2.1. Clustering				✓	Aldenderfer & Blashfield (1984); Äyrämö & Kärkkäinen (2006); Chiu & Tavella (2008); Demšar & Zupan (2013); Izenman (2008); Jacob & Ramani (2012); Jiawei <i>et al.</i> (2012); Kuo <i>et al.</i> (2006); Madhulatha (2011); Pierson (2015); Rechenhth (2014)
2.2. Partitioning methods (Non-hierarchical) k-means / k-medoids				✓	Äyrämö & Kärkkäinen (2006); Berkhin (2006); Celebi (2014); Jiawei <i>et al.</i> (2012); Rajarajeswari & Ravindran (2015); Steinbach <i>et al.</i> (2000); Wu <i>et al.</i> (2008)

Table 3.4 continues on next page

3.2 Big Data Analytics

Techniques	Tools				Sources
	SL	RL	AL	UL	
2.3. Hierarchical methods Agglomerative (bottom-up) & Divisive (top-down) approach				✓	Berkhin (2006); Halkidi <i>et al.</i> (2001); Izenman (2008); Kamrani <i>et al.</i> (1993); Larose (2014); Madhulatha (2011); Mangiameli <i>et al.</i> (1996); Mathieu & Gibson (1993); Pierson (2015); Robles (1994); Steinbach <i>et al.</i> (2000)
2.4. Density-based methods DBSCAN / DENCLUE				✓	Berkhin (2006); Ester <i>et al.</i> (1996); Jiawei <i>et al.</i> (2012); Sander <i>et al.</i> (1998)
2.5. Grid-based methods STING / CLIQUE / BANG, etc.				✓	Berkhin (2006); Bounsaythip & Rinta-Runsala (2001); Ilango & Mohan (2010); Jiawei <i>et al.</i> (2012); Park & Lee (2004); Schikuta & Erhart (1997); Tang <i>et al.</i> (2009)
2.6. Self Organising Maps (SOM)				✓	Bounsaythip & Rinta-Runsala (2001); Himberg (2000); Lin & Wu (2007); Mangiameli <i>et al.</i> (1996); Tsiptsis & Chorianopoulos (2011); Vesanto (1999); Wu & Chow (2004)
3. REGRESSION					
3.1. Linear Regression Simple & Multiple Linear Regression	✓				Ben-David & Shalev-Shwartz (2014); Bishop (2006); Gera & Goel (2015); Leatherbarrow (1990); Paliwal & Kumar (2009b); Salkind (2007); Yang <i>et al.</i> (2017)
3.2. Non-Linear Regression	✓				Bates & Watts (1988); Chatterjee & Hadi (2006); Gallant (1975); Gera & Goel (2015); Leatherbarrow (1990); Riffenburgh (2011); Ruckstuhl (2010); Tellis (2006); Tellis & Ambler (2007)
3.3. Logistic Regression	✓				Barros & Hirakata (2003); Chatterjee & Hadi (2006); Hosmer Jr <i>et al.</i> (2013); Karp (1998); Montgomery <i>et al.</i> (2012); Peduzzi <i>et al.</i> (1996); Riffenburgh (2011); Salkind (2007)
End of Table 3.4					

Table 3.5: Classification techniques for Big Data Analytics

Classification Technique	Description	Application
Decision Trees: CART, C4.5 algorithm & Random Forest	Used for the interpretation of data with high computational accuracy. This technique recursively partitions data into discrete subcategories based on a specific variable's value and it is designed for problems in which preclassified variables can be used for the learning set.	Customer identification, Target customer analysis, Direct marketing, Loyalty programmes & One-to-one marketing
Support Vector Machines (SVM)	Used for the classification of data into two or more classes by the use of a hyperplane, where the hyperplane distinguishes between the classes to find the best classification function. This technique can be used for multiple class problems, where multiple hyperplanes will be used together with regression tasks.	One-to-One marketing, Text and hypertext categorisation, Pattern recognition & Bio-informatics
Neural Networks	Used for the extraction of linear combinations from input data as derived attributes and then to model the target as a non-linear function of these attributes.	Decision-making, Pattern recognition, Spam filtering, Face identification, Sequence recognition, Direct marketing & Segmentation
Naive Bayes Vector	Used for the probability rules to predict a class based on a set of independent properties. It is a combination of classification algorithms and all these algorithms share the assumption that all the properties of the data used to classify a certain class are independent of one another, for example the accommodation preference and the age of a customer.	Direct marketing, Pattern recognition, Spam filtering & Customer lifetime value
k-Nearest Neighbour	A class of k objects is defined in a learning set based on the objects that are closest to the test object. The label assigned to a class is based on the predominance of a specific class in this neighbourhood. The three key elements are (i) a set of labelled objects, (ii) the distance between objects and (iii) the number of k -nearest neighbours. It is used for classification, estimation and prediction problems.	Concept search, Outlier detection, Recommendation systems & Loyalty programs

Table 3.5 continues on next page

Classification Technique	Description	Application
Rule-Based Classifiers	A series of rules are developed to define a class based on using the right algorithms together with training data.	Concept search, Outlier detection, Recommendation systems & Loyalty programs
		End of Table 3.5

Table 3.6: Clustering techniques for Big Data Analytics

Clustering Technique	Description	Application
Clustering	Data is categorised into two or more groups (clusters) by using an appropriate algorithm with the goal of categorising data in such a way that there is similarity between the group members.	Segmentation, Product positioning, Selecting test markets
Partitioning methods	An iterative relocation algorithm is used to build clusters up by partitioning the datasets are at specific points to create. This is done until a 'optimal' partition situation has been attained and all the objects in the dataset will belong to a cluster.	object recognition, Recommendation systems & Grouping of items Algorithms create single set of clusters, most effective for small/medium datasets.
Hierarchical methods	The datasets are grouped into clusters by using a hierarchy method. This can be done based on two approaches, namely agglomerative (bottom-up) or divisive (top-down).	Analyse market entry strategies, Design group technology manufacturing cells, Decision support for large scale R&D projects & Increase the effectiveness of credit decisions.

Table 3.6 continues on next page

Clustering Technique	Description	Application
Density-based methods	<p>Groups of data are formed based on their density conditions. The clusters grow from data points that occur in the same area based the density of neighbouring objects (DB-SCAN) or according to a density function (DENCLUE).</p> <p>Determine the clusters based on the underlying attribute space. The algorithms contain both partitioning and hierarchical algorithms and it is proposed for spatial data mining. It divides the attribute space into a finite number of cells (grids) and operations are performed on this space to form clusters. The space is restricted to only a number of cells that form the grid structure.</p>	<p>The key idea is to group neighbouring objects according to its density to discover the clusters and the noise in a spatial database. It also identifies the outliers in a dataset.</p>
Grid-based methods	<p>Analyse datasets which were not manageable due to their dimensionality and/or size.</p>	
Self Organising Maps (SOM)	<p>The topological relationships between data objects are presented on a map as units, where the maps units are ordered in a regular one- or two-dimensional grid. A class of neural networks employs training algorithms that use self organisation to cluster the data on these maps.</p>	<p>Target customer analysis, Segmentation & Complaint management</p>

End of Table 3.6

Table 3.7: Regression techniques for Big Data Analytics

Regression Technique	Description	Application
Linear Regression	A statistical tool in which the relationship between two variables is modelled. It shows the relationship between some explanatory variables and a real valued outcome and how the one impacts the other.	Evaluate trends, Forecasting, Assess finance/insurance risks, Analyse marketing effectiveness & Customer lifetime value.
Non-linear Regression	A statistical tool which is very similar to linear regression, but the difference is that non-linear regression uses a curve instead of a line as best fit. The relationships between two or more independent variables are modelled by a non-linear function.	Assesses the effects of wear-in and wear-out on campaigns. This can be done by determining the effectiveness of advertisements on different age groups.
Logistic Regression	A statistical tool in which a dataset is analysed to determine the relationship between one or more independent variables and the dichotomous characteristic of interest (the dependent variable that is the response or outcome variable). It predicts the probability of presence of the characteristic of interest.	Loyalty programs, Credit scoring, Fraud detection, Segmentation & Direct marketing

3.3 Conclusion: Big Data and its Analytics

In this chapter Big Data and its Analytics was discussed. A thorough literature study had to be conducted in order to understand what can be done with all the information available about a customer and how the demonstrator can use analytics on it to manage and improve a customer experience. Therefore, Big Data and its Analytics was discussed based on the following aspects.

The first aspect was to establish what Big Data entails. It was done by determining how Big Data can be defined by various V dimensions and for the purpose of the study it volume, velocity and variety were selected. Once the term had been defined, the importance of Big Data was discussed and how it creates value for an enterprise was shown. The challenges that enterprises currently face were also discussed in order for an enterprise to effectively plan on what to do with their Big Data.

The second aspect was to establish how analytics can be used on Big Data. The general meaning behind Big Data Analytics was discussed broadly in order to understand what is meant by it. Keeping the meaning of Big Data Analytics in mind, a Big Data Analytics Methodology was defined by the USMA research group on how an enterprise can perform analytics on their Big Data. The methodology was discussed based on a framework.

The next chapter will present a literature study about *business partnering on a cross-functional platform*.

Chapter 4

Business partnering on a cross-functional platform

In the preceding chapters a literature study was conducted based on Customer Experience and the management of it, as well as Big Data and the Analytics of it. This chapter is about the last part of the literature study, which is business partnering and how to do it on a cross-functional platform. Research in this domain is required in order to understand how the demonstrator should be constructed for Customer Experience in a partnering venture. To establish this partnering venture, a business partnership is required, with the specific focus on how an architecture is required to do it on a cross-functional platform.

Therefore, this chapter will discuss business partnering and how to do it on a cross-functional platform (referred to as *business partnering on a cross-functional platform*). Firstly, business partnering will be discussed by providing an overview of it, the purpose of why it is needed and the enabler for it. Secondly, the chapter will discuss the *business partnering on a cross-functional platform* and how it acts as a technical platform that is information based. Lastly, an integration of the three main literature study topics will be provided.

4.1 Business partnering

In order to understand what is meant by *Business Partnering on a Cross-functional Platform*, the following question first needs to be answered: what is business partnering? The answer to this question will be discussed in the following subsections.

4.1.1 Overview of business partnering

Business partnering (BP): the term itself tells one that it is a partnership between businesses, in other words the relationship between two or more enterprises. Looking back in history, one might agree with the statement that BP has been in existence for many decades (Favier, 1998; Schulte Beerbühl, 2012). Since the start of trading, enterprises have exchanged customers and business ideas amongst one another to ensure that more customers use their products and/or services and to help one another to grow. This idea around BP can be explained by a fictional example.

The example is as follows: there are two tradesmen in a town called Wondrous in the year 1503. The one tradesman, known as Geoffrey, is a baker. The second tradesman, known as Hubert, is a shoemaker. Geoffrey and Hubert have a good relationship and came to the agreement that they would help each other with growth in the number of customers by giving customers that they refer

4.1 Business partnering

a small shopping discount. Since bread is a necessity and people require it on a daily basis, Geoffrey has more customers than Hubert. He always tells his customers that if they need new shoes they should go to Hubert as he is trustworthy. Hubert on the other side had a lot of travellers from various towns that came to him for shoe repairs and he always referred them to go buy something to eat from Geoffrey before they go back on their journeys as he made the best bread in town. Both of them even have each other's products in their stores and sell these products on behalf of each other. By having this mutual agreement, Geoffrey and Hubert are business partners.

This is just an example to show what BP can entail. But it is also a reality that BP can take on various forms and can be viewed from various angles. It can be seen as a business alliance, joint ventures, cross-marketing, franchising, co-manufacturing, co-operative activity and so on (Sheth & Parvatiyar, 1992). Therefore, to better understand what BP is, one can look at its definitions.

As shown in the previous two chapters, various definitions exist in literature to define a term. That is the same with BP. Two definitions that describe BP comprehensively are from Sheth & Parvatiyar (1992) and Ellram & Hendrick (1995). Although these two definitions are from two decades ago, they still capture the essence of BP. The definitions are as follows:

1. BP can be among competitors and non-competitors and it may exist for operational or strategic reasons (Sheth & Parvatiyar, 1992).
2. BP is an ongoing relationship between two (or more) firms that involves a commitment over an extended time period, and a mutual sharing of information and the risks and rewards of the relationship (Ellram & Hendrick, 1995).

From these two definitions, Definition 4.1 has been created to define what BP is.

Definition 4.1 (business partnering). *Business partnering is when a relationship exists between two or more enterprises, irrespective of whether they are competitors or non-competitors, in which these enterprises share information amongst one another over a time period whether it may be for operational or strategic reasons in order to share the rewards and risk of this relationship.*

From Definition 4.1 it can be seen that the two dimensions of a business partnership are *purpose* and *parties*. In other words, who are the parties involved in the partnerships and for what purpose is there a business partnership? The parties who can be involved are either competitors and/or non-competitors. The purpose for the partnership can be either for a strategic and/or operational purpose. Sheth & Parvatiyar (1992) designed a typology for business alliances based on these two dimensions. Before the typology is discussed, it is important to know what business alliance is. According to Hunt *et al.* (2002) business alliance exists when two or more businesses are in an

4.1 Business partnering

ongoing formal business relationship with the purpose of achieving a set of common goals. Therefore, business alliances are indirect BP.

BP can be categorised as one of four types as can be seen in Figure 4.1 and it can be explained as follows:

- *Cartels*: Formal agreements of partnering amongst competitors for operational purposes. An example of a cartel is two or more enterprises that partner with one another with the purpose of fixing prices, share a common infrastructure or control the supply of products.
- *Co-operative Arrangements*: Partnering amongst non-competitors for operational purposes. With this arrangement two or more enterprises share their costs and facilities in order to have operating efficiency. For example, these enterprises will modify their customers' or suppliers' systems or procedures and share relevant information amongst each other.
- *Competitive Alliances*: Ventures between competitors. When two or more enterprises partner with each other, but outside the relationship they still remain competitors. For example these enterprises will leverage combined capabilities and resources from each other.
- *Collaborative Enterprises*: Partnering of non-competitors for strategic purposes. The scope of this partnership is broad as it contains many functional areas of the enterprises. For example these enterprises can have joint marketing offers, technology or market development and joint products.

		Parties	
		Competitors	Non-competitors
Purpose	Strategic	Competitive Alliances	Collaborative Ventures
	Operational	Cartels	Co-operatives

Figure 4.1: Types of business partnering (Sheth & Parvatiyar, 1992)

Now that the BP types have been defined it is also important to realise that the business partners in these partnerships can fulfil one or more of the following types of roles (Burns *et al.*, 2014):

1. *Corporate Strategy*: Control over the 'strategic heart' of an enterprise and creates the link between the operational and strategic level.
2. *Change Management*: Pro-actively drives and leads enterprise change in a supportive manner.
3. *Customer Relationship Management*: Better understands the needs of a customer by the use of communication and information on a continuous basis.

4. *System Development*: Continuous influence on the design, management and development of the information systems.
5. *Risk Management*: Identify, monitor and measure the risks while at the same time averting these risks or minimising its damages if possible.

Keeping all the BP types and type of roles in mind it is important to understand that business partners are the key agents that will create value-adding business activities as well as the ability to drive the enterprise strategy to ensure continuous improvement and efficiency throughout the enterprise. From this it is clear that BP actually creates value. But why would an enterprise go into a partnership with other enterprises? The answer to this question lies within the next section.

4.1.2 Purpose of partnering

As we have seen in Definition 4.1, BP has two dimensions, namely parties and purpose. In this section the *purpose* of BP will be discussed to answer the question; why would an enterprise go into a partnership with other enterprises? Parise & Casher (2003) points out that relationships with other enterprises are essential for the *growth* and *survival* of the enterprise itself.

Sheth & Parvatiyar (1992) have identified eight purposes as can be seen in Figure 4.2. These purposes can be further divided into two categories based on its *strategic* and *operational* purposes, but to some extent these purposes can also overlap with each other and influence one another.

- The *strategic* purposes relate to the future of the business partnership. For example, the *growth opportunity* is a future purpose of partnering, since the enterprises will grow as the processes and relationship are evolving, but it is not happening at the current state of the partnership. These strategic purposes of BP will have an effect on the enterprise effectiveness which includes the competitiveness and future position of the enterprises.
- The *operation* purposes on the other hand represent the operational impact on the enterprise. These BP purposes affect the efficiency of the enterprise and will have an effect on the current status of the enterprise. For example, the *resource efficiency* of the enterprise will be immediately affected when the partnership is underway.

Gadman (1997) has identified a purpose of why BP can be seen as an essential part of an enterprise. The advantage is that when enterprises have a relationship amongst one another, which is mutually beneficial, it will contribute to the discovery of insights and creativity. This gives the enterprises a competitive advantage and it creates wealth by applying the gained knowledge from each other in new and innovative ways.

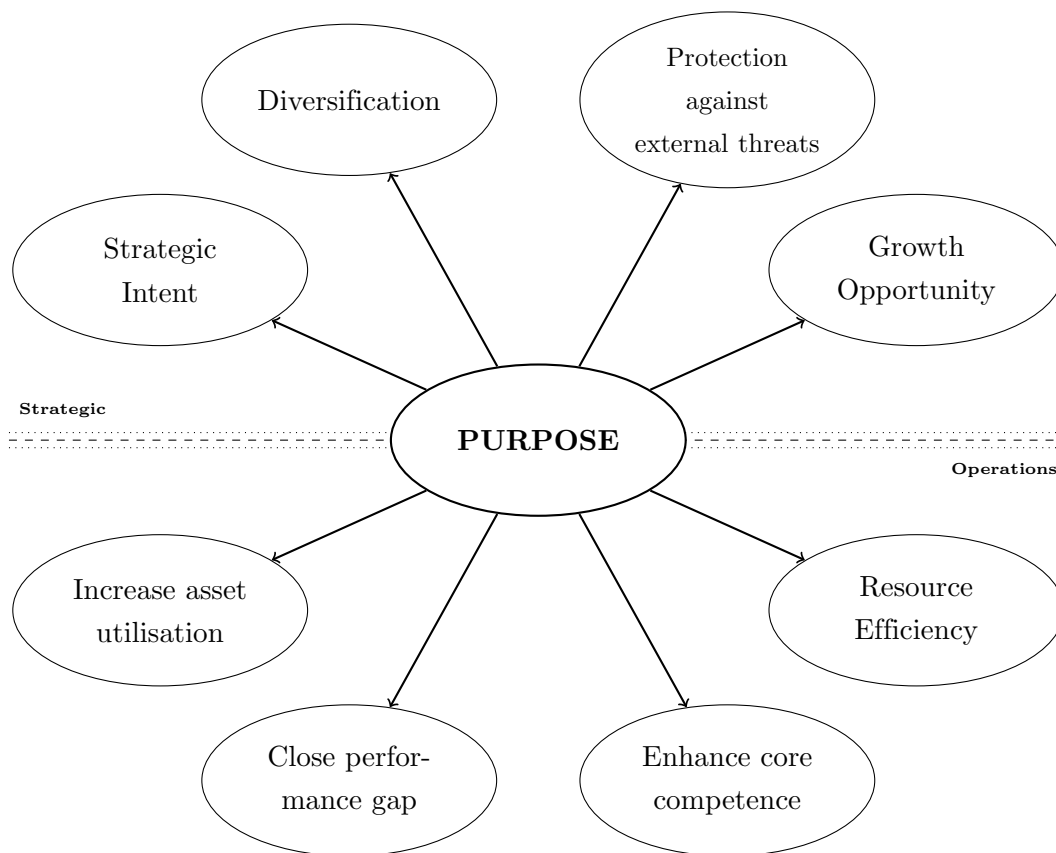


Figure 4.2: Purpose of business partnering (Sheth & Parvatiyar, 1992)

Another competitive purpose of BP is that it forms part of the strategic behaviour of an enterprise (Hasan *et al.*, 2002). Wucherer (2006) argues this further by stating the enterprises involved in the partnering are able to produce economically relevant innovations into the market and they improve their competitiveness and market position. By doing this in the correct way the enterprises are able to improve and secure their market positions.

Another purpose of BP is that it takes advantage of the intellectual capital of an enterprise, when taking into consideration the rapid change of the digital world and the knowledge one can gain of the information as we have seen in the DIKW hierarchy in Figure 3.2. Charles Savage said: “It is not just what we know, tacitly or explicitly, but how we knit together this knowledge. It is knowledge and knowledge taken together that provide the foundation for our next economy” (Gadman, 1997). Therefore, with BP, enterprises are able to exchange intellectual capital, whether it is data, information, knowledge and/or wisdom, amongst each other in order to contribute to the value creation of the enterprises itself as well as the economy.

4.1 Business partnering

Ronco & Ronco (2005) argue that BP is a method to give coordination to enterprises in a chaotic world. It is then further argued by showing what the purpose of BP is, based on the following five elements:

1. It aims to create *lasting* and *tangible processes* and *results*.
2. The starting point is to define the *issues* instead of looking at issues.
3. It requires the *involvement* from several enterprises at *multiple levels*.
4. It focuses on the *micro* and *macro* issues.
5. It is an *ongoing* relationship in which *lasting results* can be ensured.

BP has various purposes and it is important that an enterprise should know what the purpose is before they enter any partnership. It is important that the enterprise also keeps the five elements of BP in mind.

4.1.3 Enabler for business partnering

BP is when two or more enterprises have a relationship with each other in which the risks and rewards are shared. The purpose of it is that the enterprises will contribute to the discovery of new insights. But how does one enable such a relationship? This question will be answered in this section.

Although the following sounds obvious, it is a crucial contributor for the success of a business partnership. In order for BP to work, all parties involved should be willing to contribute to the relationship. Ellram & Hendrick (1995) looked at business partnerships between the buyer and seller. Based on this work, it was established that both the buyer and supplier should be actively involved to make a success of the partnership and both parties should be willing to change in order to improve the relationship. Together with this, Wucherer (2006) points out that BP is not only a cooperation between selected enterprises but it is an actual attitude change towards the way the enterprise does its daily activities. But it takes more than just the right attitude and willingness to contribute to the relationship. The enterprises have to “prove themselves as partners, make sure they qualify as partners and practice [sic] being one” (Wucherer, 2006).

Therefore, for such a relationship to be a success, the following elements should be present in the enterprises and the partnership (Burns *et al.*, 2014; Ellram & Hendrick, 1995; Gadman, 1997; Hasan *et al.*, 2002; Sheth & Parvatiyar, 1992; Venkatraman, 2015; Wucherer, 2006):

1. *At all levels of the participating enterprises, there should be a continuous and maturing culture of mutual cooperation.* It is important that everyone, not only the managers, is willing to work towards continuous improvement and communication is essential for this to work. The involved parties should be willing to help one another and share their risks.

4.1 Business partnering

2. *The enterprises should completely trust each other.* Trust is an essential element that should be present in a business partnership since it is the underlying determinant of managerial action which relates back to the control and acquisition of power for this relationship. The better the enterprises trust each other, the less managerial action for control will be necessary. Within this relationship it is also important that the enterprises help and support each other in uncertainties and the trust will help dealing with these uncertainties. Trusting one another can lead to deep insights into the need and visions of the enterprises and the partnership between them.
3. *All the enterprises should contribute to the relationship and play an active role in it.* The enterprises should have a high expectation for a long-term relationship and play their part in it. Together with trust, the enterprises should be loyal towards one another. Every enterprise in the partnership should have the willingness to make their own contributions to the partnership. The enterprises should use the operational knowledge in order to do that.
4. All the enterprises should be capable of
 - (a) *Knowing where the money is made in the enterprise.* It is important that the enterprises know at which transaction points value will be created and their focus should be shifted to these transactions. They should know and understand the key performance drivers in the enterprise that will potentially lead to value creation.
 - (b) *Investing in the relevant information and the integration of it.* For an enterprise to integrate the daily activities and the strategic goals, they should know in which information to invest and be able to integrate its performance view across time and departments. This is necessary since it enables an enterprise to see the links between activities and how these contribute to the financial and strategic performance of the enterprise itself.
 - (c) *Managing risks and opportunities by valuing trade-offs.* By doing this the enterprises involved should share their risks and rewards amongst each other. In order for this to work, loyalty is important, while at the same time each party should focus on continuous improvement. To further explain, the enterprises should be able to balance the trade-off between promoting initiatives that support the competitive positioning of the enterprise while at the same time they should enhance efficient behaviour and mindset that will drive the bottom-line profit targets. This balancing act can be seen in Figure 4.3.
 - (d) *Understand the partnering requirements and develop a roadmap to assess the feasibility of it.* The right choices should be taken regarding to how the partnering with other enterprises should be handled, to ensure the feasibility of being able to enter a partnership while at the same time being innovative in nature. To understand the partnering requirements in more depth it also requires the enterprise to look at the requirements, needs and

4.1 Business partnering

conditions of the other enterprise(s) who would also like to partake in the partnership. The rationale and plans of a partnership should be closely investigated to understand the benefits of entering a partnership.

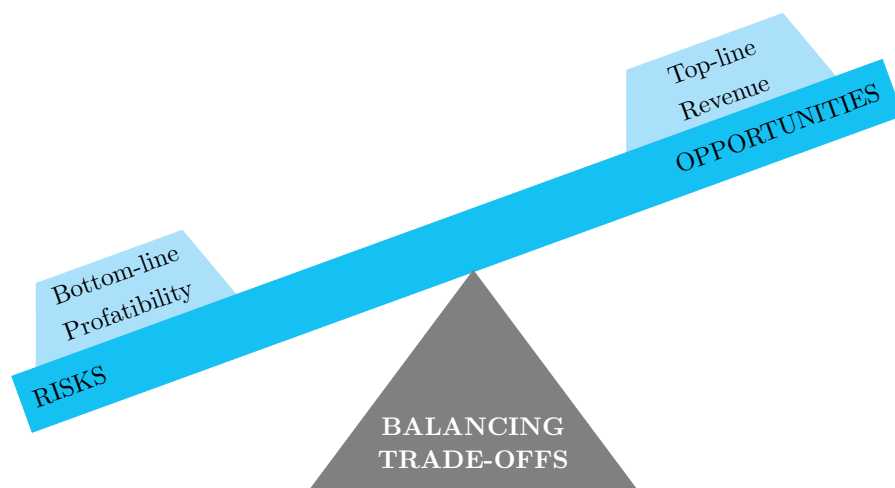


Figure 4.3: Balancing act for business partnering (Venkatraman, 2015)

5. *The enterprises should have the right hard and soft skills.* These skills are required to feature and support the previously mentioned elements. Skills must be applied in an effective way to emphasise the interest of the partnering venture by minimising conflict that can occur between relevant parties. The skills form part of the culture of the enterprises. Hard skills are, for example, to be IT proficient and have a broad business understanding. Soft skills are, for example, communication, interpersonal skills and strong conviction and persuasion. Effective communication is crucial for this relationship to be a success. Leadership skills are also essential to teach partners about decision-making with a focus on the finance and business values.
6. *The enterprises should be proactive by nature.* Another key to the success for partnering is that the enterprise should be pro-active in the digital world due to the ever increasing rate of change . The proactiveness of an enterprise goes hand-in-hand with the reason-and-logic way of thinking to which the enterprise should adapt to. Enterprises should be able to create a powerful synergy of insights in order to broaden their horizons and reveal new innovative possibilities.
7. *The enterprises should change their traditional way of thinking.* For an enterprise to make a success of the partnership they should move away from their traditional way of thinking into a more adaptive way of thinking. The enterprises should be more open to change and their thoughts around problem-solving should be in the right context.

4.1 Business partnering

Another enabler for BP is that the enterprise should change its traditional way of thinking. [Gadman \(1997\)](#) argues that an enterprise should change from a ‘clock logic’ way of thinking to a ‘reverse logic’ way of thinking. *Clock logic* refers to the way of thinking where the outputs and behaviours in the enterprise can be reduced to predictable equations and quantifiable relationships. By doing this, enterprises are in a situation where today’s problems come from yesterday’s solutions and today’s solutions will cause tomorrow’s problems, whereas the *reverse logic* way of thinking changes the perceptions, emotions and beliefs about the enterprise and others, through arguments that are based on reason and logic. By adapting to the ‘reverse logic’ way of thinking an enterprise acknowledges that the world created by men is both a reflection of humans’ inner patterns, perceptions and the relationship with the outer world. By doing this, an enterprise will be able to continually question the actions and process as well as recognising the suppliers, customers, competition and the enterprise itself as a seamless process. Therefore, by moving beyond the ‘clock logic’ way of thinking, the enterprises involved in a partnership should recognise and accept that they are a reflection of the marketplace and the market place is a reflection of the enterprise. Enterprises are then able to act as self-reflecting systems in which a strong connection exists between the enterprise itself is and how the world appears to the enterprise. For further reading on how enterprises should revert to ‘reverse logic’, good sources are the book by [Gadman \(1997\)](#), *Power Partnering: A strategy for Business Excellence in the 21st Century* and the book by [Ziegler et al. \(1994\)](#), *The Republic of Tea*.

It is also important to know what properties should be present in the business partnership in order for it to be a success based on the BP type it is. Four types of BP were identified in Section 4.1.1 as *competitive alliances*, *collaborative ventures*, *cartels* and *co-operatives*. [Sheth & Parvatiyar \(1992\)](#) have identified the properties that should be present to enable BP and four propositions on how to encourage a greater degree of autonomy, improved cross-functional cooperation and free flow of information amongst the enterprises, as can be seen in Figure 4.4. The description of the four propositions is as follows:

1. For the success of *strategic* BP, a great commitment amongst the enterprises should be present, the enterprises should be context-driven and the enterprises must be organised as a unique business venture in order to achieve its strategic effectiveness.
2. For the success of *operational* BP, a high degree of coordination is required and asset-sharing should happen on an ongoing basis. Another requirement is that the businesses processes requires re-engineering in order to achieve operational efficiency.
3. For the success of *competitive* BP, a high degree of management control is required, together with a limitation of specific functions, while at the same time ensuring mutual learning amongst the competitors.

4.1 Business partnering

4. For the success of *non-competitive* BP, a high degree of governance is required. Cooperation and learning should occur on a cross-functional basis, while communication on a free-flow basis should be encouraged.

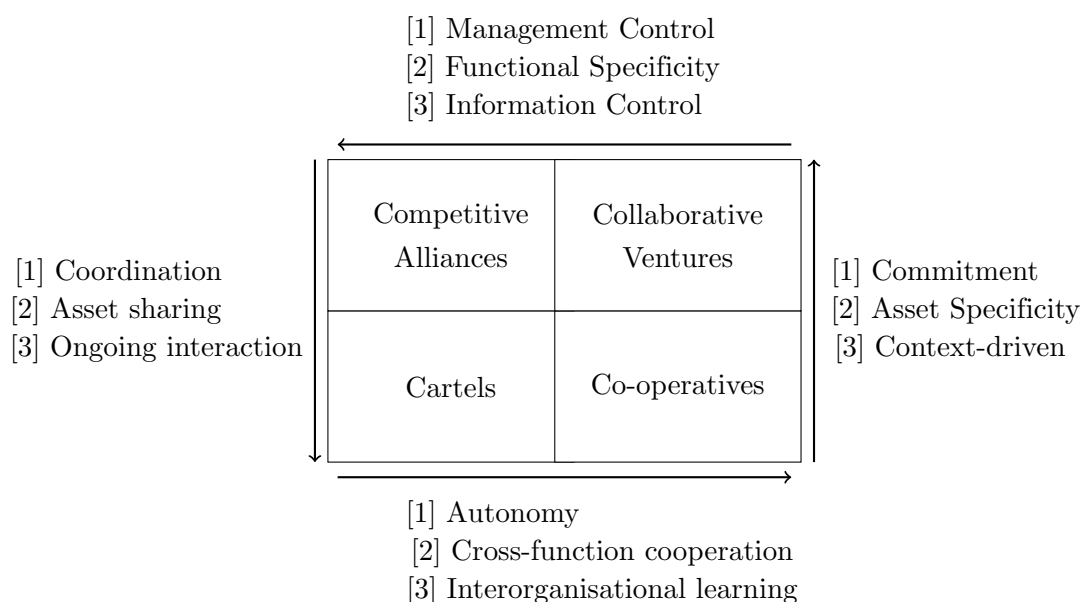


Figure 4.4: Properties of business partnering types (Sheth & Parvatiyar, 1992)

Based on these elements as described above, it is crucial that a *common framework of meaning* should be designed for the enterprises to work from in order to ensure that the partnership will be a success (Rich, 2015). This common framework of meaning is like a game. Each player who actively partakes in the game knows how the game works, how it progresses through various phases and they know the rules of the game.

As a concluding summary, to enable BP between enterprises it is important that the enterprises contribute to the relationship to the best of their ability. By doing this the enterprises should focus on the key elements necessary to enable it. These elements include the participation of all levels in the enterprise as well as its culture, to trust the enterprises who participates in this relationship, know the business processes inside and out, know what information is relevant and the enterprises should have the necessary hard and soft skills in place. Based on these elements it is important that the involved enterprises should derive a common framework in which they agree on how the partnership will work. The last important factor to enable BP is the type of partnership it will be. In order to implement BP successfully it is important to know for what purpose the enterprises enter the business partnership.

4.1.4 A business partnering example

Now that a thorough understanding has been given for BP, the purpose of it and the enabler for BP, it is important to show an example of how BP has been applied in practice. This will be done by looking at an example from *TMForum*.

TMForum is a non-profit association for enterprises in the telecommunications domain and they provide toolkits, amongst other things, for these enterprises to help them with problem-solving to maximise the success of the enterprise (TM Forum, 2017). One of these toolkits is the “*Big Data Analytics Guidebook*” which can be used by any enterprise as a best practice on how to implement BDA. Part of the tool kit contains *Use Cases* that gives an enterprise guidance based on a real world example of how BDA projects can be implemented and contribute towards value creation.

A good illustration of how BP contributes to value creation for an enterprise can be seen by looking at a TM Forum Big Data Use Case, *EP-PAM-1: Partner Value Optimisation* (Kerrigan *et al.*, 2017), which is a good example of how partnering can be done to create value optimisation.

To maximise the value of BP, BDA can be applied through optimised incentive programmes, commission rules and the settlement arrangements. In addition, the analytics can be used to increase profitability from partners usage by optimising new product and rate plan introduction as well as optimising consumer usage. BDA can also be applied to measure incentive plan effectiveness based on historical data and from that to forecast future sales revenues. Furthermore, analysis of past and future dealer sales performance can provide the ability to simulate incentive plans. To enable this the right data sources should be considered. The enterprises must also know what key activities and key resources are required.

In order for this partnership to work, the following data sources are required from all enterprises involved:

- Partner management data
- Billing and usage events
- Product catalogue
- Network data
- CRM data
- Purchase history

4.2 Business partnering on a cross-functional platform

The key activities required to enable the partnership are as follows:

- Measure the performance of existing incentive plans, commission rules and settlement arrangements.
- Forecast the future performance.
- Design optimised incentive plans, commission rules and settlement arrangements.

The key resources are based on the following data requirements that are related to:

- How the enterprises currently interact with their partners.
- The performance of those partners.
- The incentives they receive.
- The commissions they make.

If all these data requirements, key activities and key resources are implemented and considered by an enterprise, it is highly likely that the partnerships will be a success as the partner interaction process will be optimised. It will lead to a success due to the fact that issues are addressed by determining the way in which the enterprise interacts with their partners. This will then have a knock-on effect on the customers of the enterprises and it will strengthen the revenue that the enterprises receive by entering such a partnership.

4.2 Business partnering on a cross-functional platform

So far an overview of BP was given as well as the purpose of it and the enabler for business partnerships. In this section the following question needs to be answered: what is *business partnering on a cross-functional platform*? To understand it better, first an overview of platforms will be given to understand what is meant by a platform and how one can perceive it. Secondly, how BP can be done on a cross-functional platform will be discussed. The concept of *business partnering on a cross-functional platform* is synthesised by the author in which the partnership is combined with a platform.

4.2.1 Overview of platforms

In order to understand the *business partnering on a cross-functional platform*, it is important to understand what is a platform.

4.2 Business partnering on a cross-functional platform

Various forms of platforms exist in an enterprise, where some of the platform types can be identified as:

- Economical platform
- Social platform
- Political platform
- Software platform
- Strategy platform
- Product platform
- Business platform

For the purpose of this study as well as keeping in mind that it should be used for BP, only the *business platforms* will be considered since it is the most appropriate type of platform to use for BP. To further understand what a business platform is, it is necessary to define it and Definition 4.2 will be used.

Definition 4.2 (Business Platform). *A business platform consists of a set of related systems, processes, people and data that deliver (or expose) a set of business services via a standard set of application program interfaces (Turner, 2017b).*

Definition 4.2 in other words states that a business platform enables the logical partitioning or subdivision of an architecture into a set of related business services. These business services are provided by means of people, processes and systems. Therefore, this business platform enables an enterprise to handle change in an effective way without comprising the business processes. It also gives an enterprise the ability to enter other markets and it enhances its competitive ability.

Based on this it can be seen that a business platform consists of two key elements. Turner (2017a) defines the two elements of a business platform as follows:

1. *Platform Business Model*: A digital ecosystem or marketplace within an enterprise that connects the customers with the product and/or service in order to enable the ease of doing business.
2. *Platform-based IT Infrastructure*: An infrastructure in the enterprise that acts as a facilitator for the business model and supports the electronic marketplace.

4.2 Business partnering on a cross-functional platform

Most enterprises only have the one of these elements, but it is essential to possess both elements. The platform business model can evolve from the platform-based IT infrastructure and *vice versa*. When both of these elements are present, the enterprises will be able to use this business platform as a *digital ecosystem* in which new and more verticals can be enabled (Turner, 2017a). The term ‘vertical’ refers to the vertical market, which is the marketplace on which an enterprise offers products and/or services that will meet a specific need of a customer.

Based on this it can be concluded that a business platform gives an enterprise the ability to transform the business processes by the use of a set of organising principles and capabilities. These platforms exhibit an application layer that is used to design and execute new, high-level, scalable and effective products and/or services.

4.2.2 The cross-functional platform

As can be seen from the discussion thus far, BP has been present for many centuries and in Section 4.1.1 it was shown that BP has been present since the start of trading. To date a lot of key elements have been identified to show what enterprises should focus on to make a success of the BP. But moving beyond these elements, how can one use BP as a concept in today’s digital world where technology has become predominant and BD is everywhere?

Therefore, this concept of *business partnering on a cross-functional platform* is synthesised by the author, by taking the knowledge around BP and business platforms to determine how BP can occur on a cross-functional platform.

Figure 4.5, Turner (2017b) illustrates that if two phones are connected to each other, one connection is established. When five phones are connected to each other, 10 connections are established. But when 12 phones are connected to each other, 66 connections are established. In other words, as more phones are connected with one another, the more connections will be established. These connections grow at an exponential rate as more phones are connected. With this phenomenon, Turner (2017b) shows that the more parties that are signed on to a platform, the more people and companies will be reached and this will lead to more effectiveness as well as value creation. Therefore, in terms of the *business partnering on a cross-functional platform*, the more enterprises that will sign in to the cross-functional platform, the more connections there will be in this partnering venture.

Burns *et al.* (2014) mentions that business partnerships are created today for advisory and analytical purposes in which the integration of information from different parts of an enterprise can be done as well as between the partners. The reason is that enterprises are now experts in preparing, interpreting and using information/datasets. Together with this and the number of connections that can be established between enterprises, the BP can be performed on a cross-functional platform.

4.2 Business partnering on a cross-functional platform

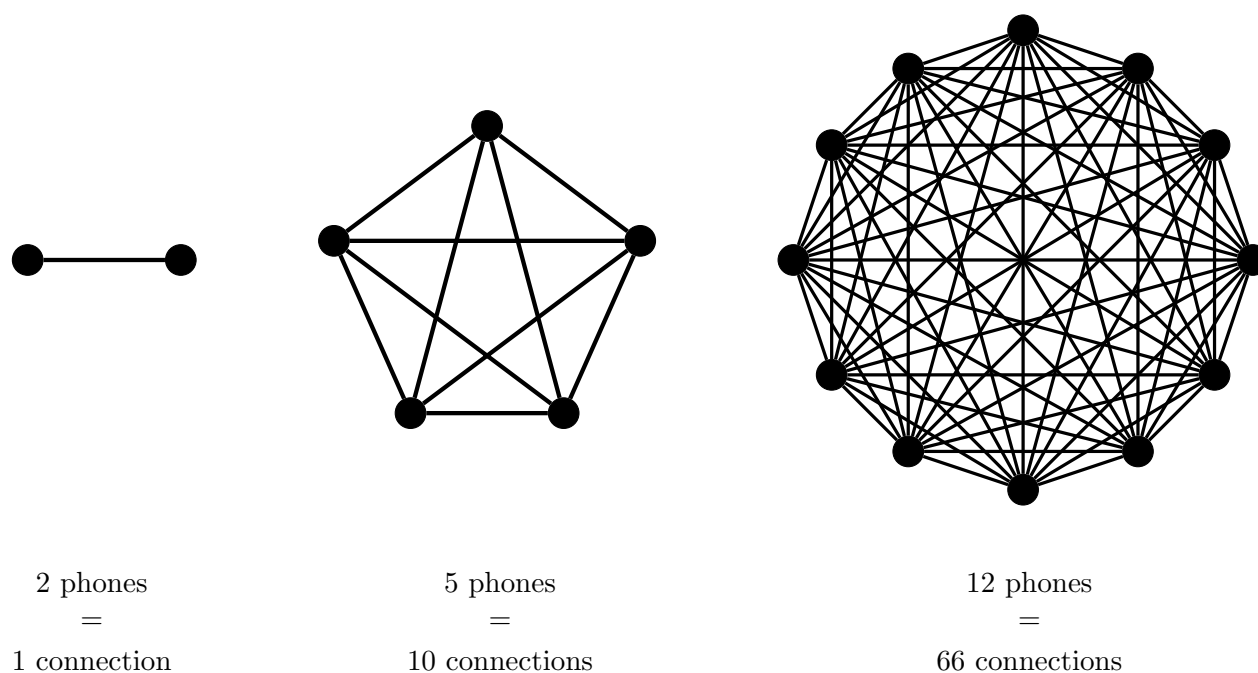


Figure 4.5: Platform effect (Turner, 2017b)

It will be a cross-functional platform in the sense that all these enterprises in the partnering venture will all participate in their own way to the partnership but at the same time they will all work and contribute together to make a success of the partnership. In other words, the enterprises will all sign up to the partnership for various functionalities, where these functionalities can be unique and/or the same. Within this platform the enterprises will exchange their information amongst one another. Keeping Figure 4.5 in mind, one can only imagine the amount of information that can be exchanged between the different enterprises as the number of enterprises signing up for such a platform increases.

This cross-functional platform will be a technological architecture that is information-based. In other words it will be a technological architecture which will enable the ability of the enterprises to share and exchange information amongst one another to add value to each independent enterprise as well as the owner of the platform. With the assumption that data governance and security has been taken care of and that the platform adhere to the Protection of Personal Information Act (POPI) or General Data Protection Regulation (GDPR) (EU, 2016; SA Government, 2013). To explain this further an example will be used with reference to Figure 4.6. The following six enterprises all sign up to this platform known as the *business partnering on a cross-functional platform*. The six enterprises are as follows:

- *XYZ Airlines*: It is the top airline company in country X that only flies domestically.
- *Airline Flies*: It is an airline company in country X that flies domestically and internationally.

4.2 Business partnering on a cross-functional platform

- *B&B Sleptight*: It is a family-operated company that has five bed-and-breakfast facilities situated in country X.
- *Busy Bee Hotel Group*: It is a hotel and leisure company situated in Country X with 87 hotel properties in five neighbouring countries.
- *Bank-it-up*: It is the oldest and most popular bank used by the citizens in Country X.
- *Ring ring mobile*: It is one of the top three telecommunication companies in Country X which provides voice, data, messaging and converged services to more than a third of the population in Country X.

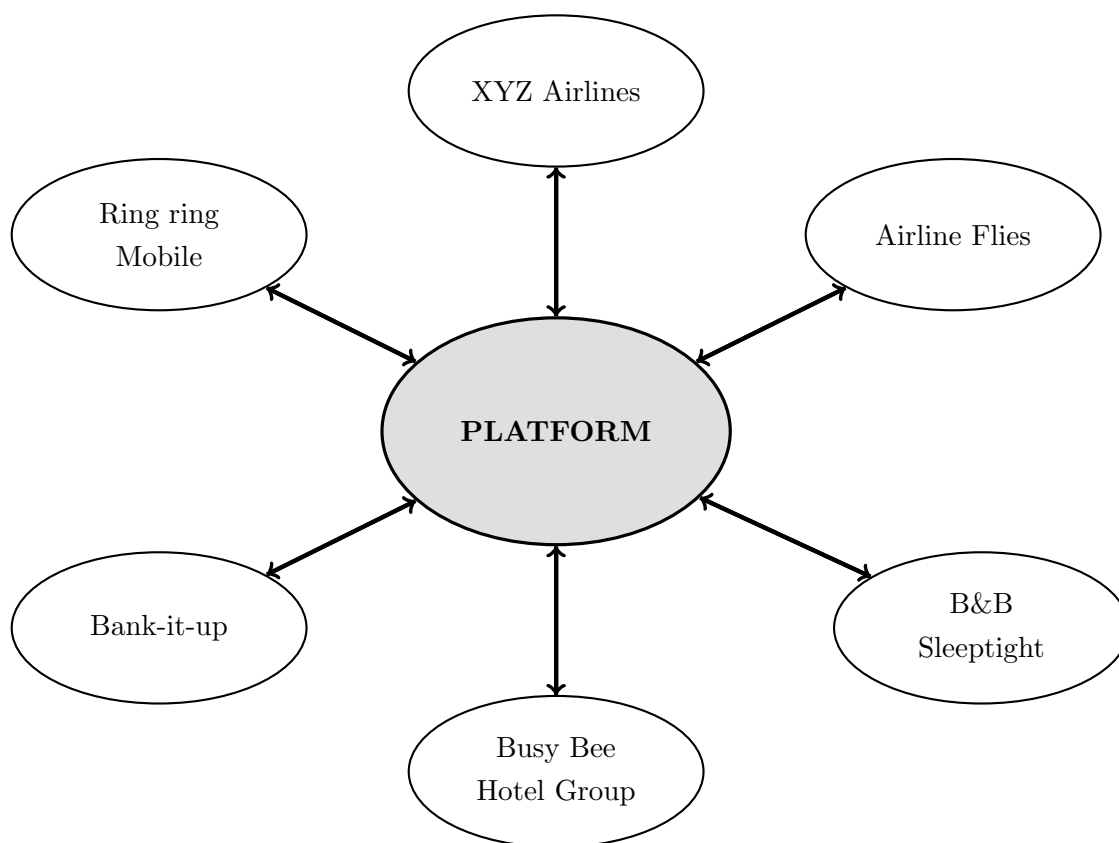


Figure 4.6: Business partnering on a cross-functional platform

Although *XYZ Airlines* and *Airline Flies* are competitors, they still opted in to the platform as they can share information with each other around their customers and make use of each other's business opportunities to create value. All five enterprises can also get information from *Bank-it-up* on how customers spend their money and the enterprises can learn what opportunities will be cost-effective for both the customer and the enterprise. *Ring ring mobile*, on the other hand, can track their customers' internet usage, as well as their location. They can then share with the other enterprises what customers spend most of their time on when they access the web and how the

4.3 The literature study integration

customers travel. *Busy bee Hotel Group* and *B&B Sleptight* can give information to the airline companies on how customers patterns are with checking in and out in order for the airlines to be able to better plan their flight times and destinations. So, the chain of events goes on, on how these enterprises create more business for each other by sharing non-sensitive information. By doing this on a platform, it gives all these enterprises a better opportunity to be able to generate more value for themselves and one another.

It is important to note that every single byte of data exchanged on this platform is non-sensitive information to the platform with regards to the behaviour of their customers. In other words how the customers spend their money and time at that particular enterprise, how the customer react based on certain situations, what the customer complains about, what type of treatment the customer is looking for and so on, instead of giving identity numbers, cellphone numbers and physical addresses to one another. Based on this information provided by the different enterprises, the platform will then store and categorise this data for the other enterprises to use in order to create value. By doing this the enterprises partner with each other and the platform acts as the enabler for information to be exchanged and to create new revenue streams. It is important to note that non-competitors and competitors partner on this platform and the platform is open for any type of enterprise to opt in.

For *business partnering on a cross-functional platform* it is important to consider the privacy and security of data that will be exchanged on this platform. Privacy and security together with the governance in enterprises should be considered when designing this platform. The scope of this study does not include this factor. The assumption is made that it had already been considered and included when the *business partnering on a cross-functional platform* was designed.

Therefore, to answer the question “What is *business partnering on a cross-functional platform*?” It is an information-based platform that is enabled by the use of a technological architecture for various different enterprises to partner with each other. By entering this partnering venture the enterprises are willing to share and exchange data amongst one another by the use of this platform as an enabler to create new revenue streams.

4.3 The literature study integration

An extensive literature study was done in order to understand how to construct a demonstrator in order to implement it via the Trip Planner, where the purpose is to show how BP together with BDA can be done to improve and manage a CX. Therefore, the literature study consisted of three concepts/domains. Firstly, Customer Experience and its Management were researched in Chapter 2. Secondly, Big Data and its Analytics were researched in Chapter 3. Lastly, BP and the *business partnering on a cross-functional platform* were researched in this chapter. Now the last question to answer is, how do these three concepts integrate with each other to manage a CX in the best way possible?

4.3 The literature study integration

In Figure 4.7 the relationship between BDA and CEM can be seen. From Figure 4.7 it can be seen that BDA is used to analyse data about the customer. From this analysis, trends about the customer can be identified in terms of their expectations and behaviour. By using this knowledge, together with the principles of CEM, a CX can be improved to add value to the enterprise.

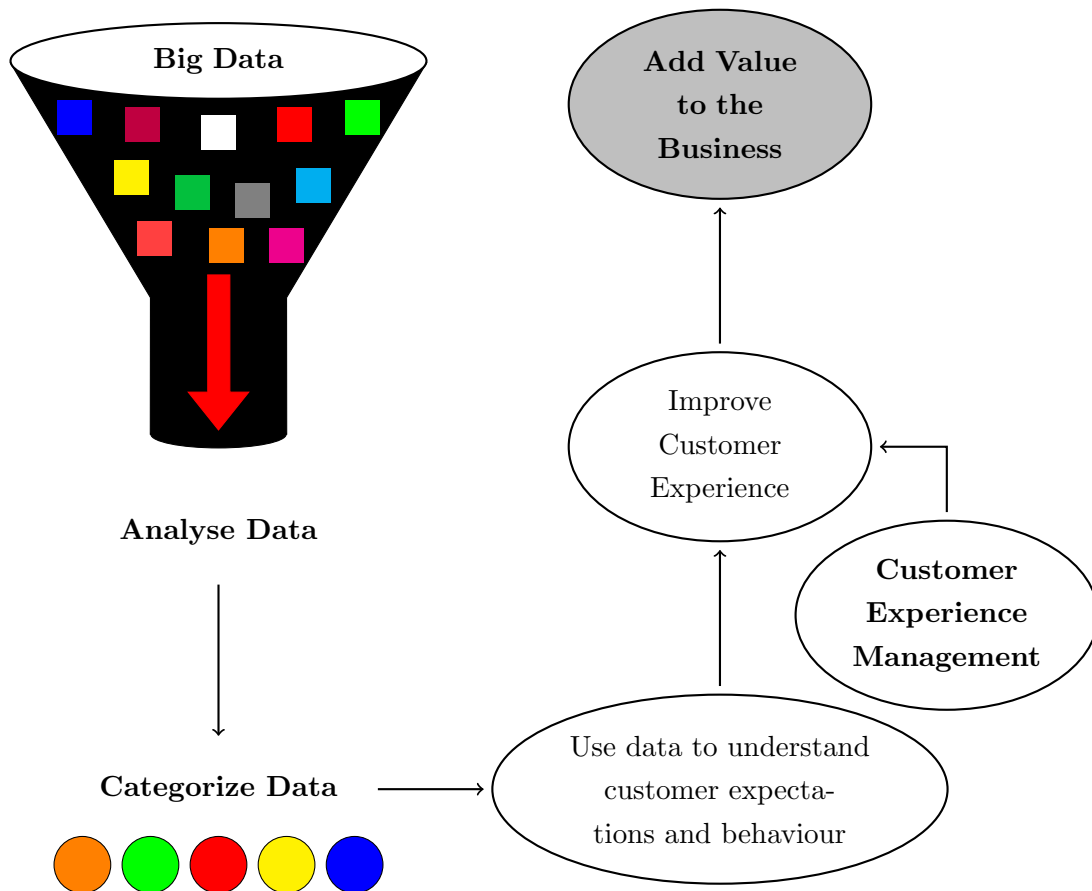


Figure 4.7: Big Data Analytics together with Customer Experience Management

But how does 'BP' fit in? BP acts as the enabler for enterprises to partner with each other to understand their customers by using customers' data on a cross-functional platform and by partnering, the enterprises are then able to improve the CX. Therefore, *business partnering on a cross-functional platform* gives the ability to perform BDA to improve a customer's experience.

As shown in the DIKW hierarchy diagram (Figure 3.2) and the purpose of BP, enterprises will exchange their intellectual capital amongst one another in order to add to value creation. The intellectual capital of the enterprises will be shared on the information platform as discussed in the previous section. This information platform will then act as the enabler for CEM by the use of BDA.

As a conclusion, the *Business Partnering on a Cross-functional Platform* acts as the enabler in order for information to be shared amongst enterprises. It is an information-based platform that is enabled by the users of a technological architecture. This information shared amongst the different

4.4 Conclusion: Business partnering on a cross-functional platform

partners on this information-based platform, will be stored in an information basis. BDA can then be performed in order to gain insights and knowledge about the customers. The enterprises can then apply CEM techniques on these insights and knowledge in order to improve and manage its customers' experience. An article by [Wucherer \(2006\)](#) shows how Siemens has adapted its business model to enable BP. By doing this it supports how BP can enable the ability to deliver maximum benefit to the customer and how an enterprise is able to support their customers through every stage.

4.4 Conclusion: Business partnering on a cross-functional platform

In this chapter, *business partnering on a cross-functional platform* was discussed. A thorough literature study had to be conducted in order to understand how business partnering will act as the enabler for a system in which customer experience needs to be managed by the use of Big Data Analytics. By doing this, more knowledge has been gained on how the demonstrator should be constructed. The *business partnering on a cross-functional platform* was discussed based on the following aspects.

The first aspect was to understand what business partnering entails. An overview of business partnering was provided, together with the purpose of why enterprises enter a business partnership and the enabler to enter such a partnership. A business partnering venture can only occur if two or more parties are involved and if there is a purpose to do it, where the overall purpose is that the enterprises will contribute to the delivery of new insights. The enabler of such a partnership consists of various elements, the most important of which is that enterprises should be willing to contribute to the partnership and know for what purposes they are entering it. After this an example was provided to show how business partnering can be applied in practice.

The second aspect was to understand what *business partnering on a cross-functional platform* is. This concept was synthesised by the author by using the idea of business partnering together with platforms. Therefore, first an overview was given of what a platform is with the focus on the business platform. After that the *business partnering on a cross-functional platform* was explained; that it is a technological architecture that is information-based.

The last aspect was to understand how all these different literature study domains discussed in Chapters 2, 3 and 4 integrate with one another. It was seen that *business partnering on a cross-functional platform* act as the enabler for BDA techniques to be applied in order to use CEM to improve and manage a customer experience.

The next chapter will discuss the demonstrator that is constructed for the purpose of this study.

Chapter 5

The development of the Trip Planner Demonstrator

In the preceding chapters, the research proposal was presented and a literature study was conducted based on Customer Experience and its Management, Big Data, and Analytics and Business Partnering on a Cross-functional Platform. This chapter will discuss the third aspect of the research methodology from Section 1.6, namely the development of the demonstrator. As mentioned in the research statement provided in Section 1.2, the demonstrator will show how Big Data and business partnering can be used as an enabler for customer experience and how to manage the customer experiences. This demonstrator will be implemented as a digital system known as the *trip planner*, where the trip planner will record various touch points, which consist of the customer activities and customer experiences. These touch points will be realised by the simulation of unique trips for many customers.

Therefore this chapter will discuss the various aspects required for the development of the *Trip Planner Demonstrator* (TPD). First, an overview of the TPD will be provided by looking at what it consists of. Secondly, a system architecture will be constructed to understand the needs and requirements of the TPD. Thirdly, a database for the TPD needs to be developed based on the system architecture to understand the information system required for it. Lastly, the simulator that acts as the enabler of the TPD's processes, needs to be developed based on the system architecture and database.

5.1 The Trip Planner Demonstrator

For the development and construction of the TPD, it is important to understand what it consists of. A graphical representation of the TPD can be seen in 5.1. The TPD consists of:

1. business partners,
2. a database,
3. a simulator and
4. a data analytics function.

The TPD is a *digital information and support system* whose functionality is supported by the four components. Each of the four components should be able to function on its own in isolation, but they should also be able to function as a unity in the overall system. In other words, the four components

5.1 The Trip Planner Demonstrator

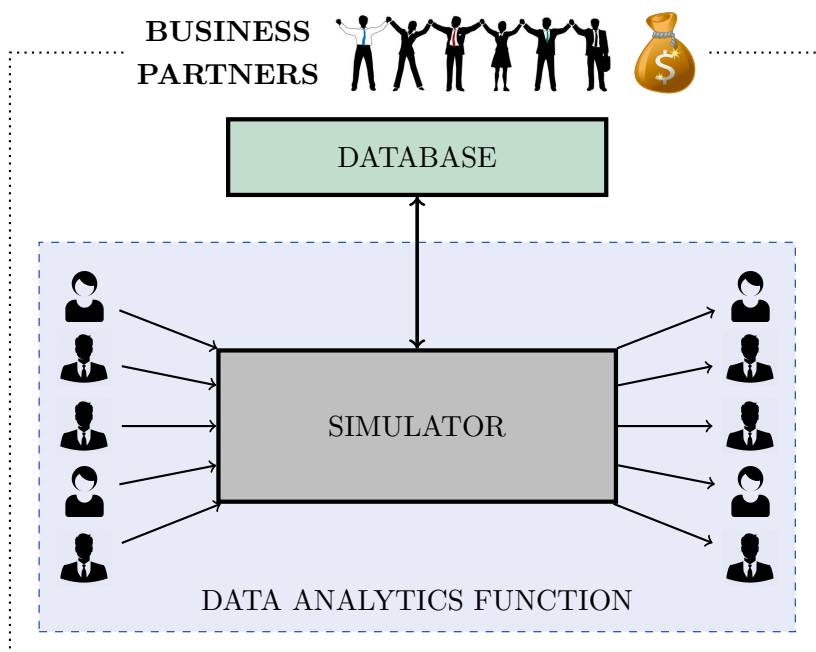


Figure 5.1: Components of the Trip Planner Demonstrator

should integrate with each other and function as a whole, while at the same time it should act independently. Therefore, the four components of the TPD perform as a digital information and support system by integrating with one another and being interdependent on each other. Each of these four components has their own unique functionalities but they play a vital role in the overall functionality of the system.

The first component of the TPD is the *business partners*. In order for the TPD to function as a digital information and support system, business partnerships are required. The enterprises who signed up for the partnerships are the enablers of the TPD as they share data on the cross-functional platform. Therefore, the cross-functional platform is required for the functionality of this system as data needs to be ingested, shared and stored amongst the business partners as discussed in Section 4.2.2. Therefore, as can be seen in Figure 5.1, the business partners act as a cornerstone for the overall functionality of the system. The business partners create the ability to share relevant customer information on the platform of the TPD. In other words, they will ensure that value is generated on the cross-functional platform of the TPD by sharing relevant data. These enterprises are also the ones who will reap the benefits of it as they will gain more business. Therefore, more business value will be generated for them.

It is important to note that non-sensitive customer data will be shared and that it will meet the requirements of the GDPR. The GDPR is a data protection and privacy regulation that has been set out by the EU that stipulates how personal data may be processed and the movement of this data (EU, 2016). In South Africa, the POPI act has been set in place and it adheres to the GDPR

5.1 The Trip Planner Demonstrator

(Giles, 2013; SA Government, 2013). Therefore, the customer data that is shared amongst business partners on the TPD platform will not include the customer's ID number, birth date, address, *etc.* The customer data that will be shared is with regards to where the customer stays while travelling, the aeroplane and/or bus enterprise they use, taxi's and/or car rental enterprises that the customer uses, *etc.* A more detailed description of the data will be provided in Section 5.3.

The second component of the TPD is the *database*. The purpose of the database is to store all relevant and non-sensitive customer data and the data which are required to enable the TPD processes. All these data entities should be stored and accessed in an efficient manner while ensuring accurate and consistent data at the same time. The *MS SQL Server* platform will be used as the database for the TPD. Section 5.3 will discuss the database and the data entities in more depth.

The third component of the TPD is the *simulator*. The purpose of the simulator is to simulate *unique* trips for *many* customers to show the functionality of the TPD. With reference to the case study provided in Section 1.1, the simulator will plan and simulate a trip for a customer based on the trip requirements (start date, end date, destination and budget) and their preferences and historical behaviour. The simulator should be able to simulate a unique trip for each customer that uses the system and for every time a customer uses it. In other words the simulator will plan a customer's trip and manage it while the customer takes a trip. The simulator will be constructed based on the architecture provided in Section 5.2. In Figure 5.2 the actions of the simulator are graphically presented. After a customer has entered their trip requirements, the simulator will plan and book a unique trip for the customer based on preferences and historical behaviour. Once the time has arrived for the trip to start, the customer will execute various interactions with the system until the trip has ended. It is important to note that since each customer undertakes a unique trip, the total touch points per trip will vary for each trip. The construction of the simulator will be discussed in Section 5.4.

The fourth and last component of the TPD is the *data analytics function*. The data analytics function has two purposes in the TPD. The first purpose is that data analytics should be performed when the trip is booked and planned for the customer. The second purpose is to perform analytics on the data available about a customer's trip, also known as the customer journey, to gain insights and knowledge about the customers. The knowledge and insights can be that a company can see what type of customer uses the services; what does this type of customer look like to be able to gain more leads and to determine if customers can be grouped according to their responses. It is important to note that for every trip a customer takes, the CX at each specific touch point is recorded and the various touch points create the customer journey as discussed in Section 2.4.

Analytics, in the form of ML techniques, should be performed on the recorded CX as well as historical behaviour and preferences for the planning of a trip. By performing ML it gives the TPD the ability to continuously improve a customer's experience when they take a trip and to determine

5.2 The system architecture for the Trip Planner Demonstrator

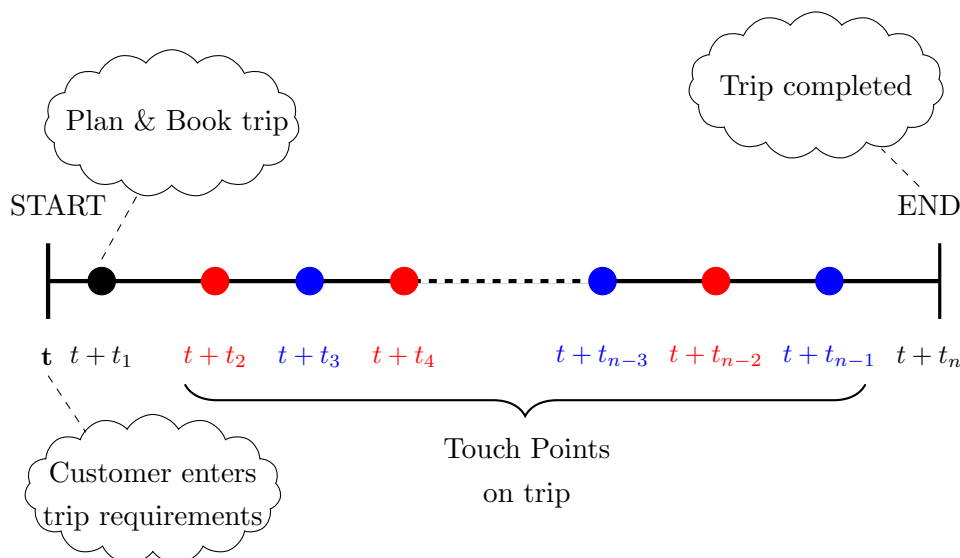


Figure 5.2: Overview of the actions of the Trip Planner Demonstrator simulator

its ability to successfully plan and execute a customer’s trip. To obtain these results, the appropriate ML tool and technique should be applied as discussed in Section 3.2.2.3. The data analytics which are used for the booking of the trips will be discussed in Section 5.4 and the analysis of the TPD will be discussed in Chapter 7.

Now that the components of the TPD have been discussed its construction can begin. The first aspect to consider is the system architecture, which will be discussed next.


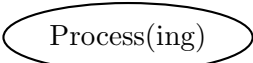






5.2 The system architecture for the Trip Planner Demonstrator

To develop the TPD it is important to first develop a *system architecture* to understand its *needs* and *requirements*. The term *system architecture* is defined by the [Open Group Architectural Framework \(2001\)](#) as “a set of elements depicted in an architectural model, and a specification of how these elements are connected to meet the overall requirements of an information system”.

The system architecture for the TPD is developed by using the *Object-Process Methodology* (OPM), which is an ISO Standard (ISO19450). OPM was developed by Dov Dori and can be defined as a methodology which “combines formal yet simple graphics with natural language sentences to express the function, structure, and behaviour of systems in an integrated, single model” ([Dori, 2002](#)). In other words it is a system architecture methodology in which a system’s behaviour and structure are represented in one model. OPM consists of two building blocks, namely *objects* and *processes*. A third OPM entity is *links* which is used to hold these building blocks together. A description of some of the OPM elements that will be used for the TPD can be seen in Table 5.1. A fourth OPM entity that is not used in the TPD is *state*, which describes objects and it cannot exist on its own.

5.2 The system architecture for the Trip Planner Demonstrator

Table 5.1: Object-Process diagram legend (Dori, 2002)

Symbol	Name	Description
	Object	A static object that has the potential of mental, physical, unconditional or stable existence.
	Process	The transformation that objects undergo by affecting, consuming, or generating them.
	Aggregation	Relates a whole to its parts.
	Generalisation	A general thing is related to its specializations.
	Instrument Link	A non-human that act as enabler in order for the process to occur.
	Agent Link	A human that act as enabler in order for the process to occur.
	Structural Link	It represents a meaningful association between two objects.
	Input/Output Link	It represent the connection between an object and a process to show how the object undergoes transformation or enables a process.

The OPM uses a set of *Object-Process Diagrams* (OPD) and *Object-Process Language* (OPL) to represent a system. The OPD is a graphical representation of the system architecture and the OPL is a collection of sentences which describes the corresponding OPD. The OPD for the TPD can be seen in Figure 5.3 and the corresponding OPL can be seen in Figure 5.4.

As can be seen from Figure 5.3 the TPD consists of two processes namely the *trip planning* process and the *customer travelling* process. The trip planning process is when the TPD uses the customer's trip requirements to plan a trip for the customer based on their historical data and preferences; whereas the customer travelling is when the TPD notifies the customer about the upcoming events of a specific trip and records the CX at the various touch points along the customer journey. The two processes will be discussed in more depth by using the OPD together with the OPL.

The first process, namely the *trip planning process*, is handled by a customer based on their trip requirements, historical data and preferences. This process requires data about the customer's historical behaviour with regards to accommodation, long distance form of transportation (referred to as Long Distance Transportation – LDT) and short distance form of transportation (referred to as Short Distance Transportation – SDT) in order to book a trip for a customer based on the customer's preferences, trip requirements and recordings of customer journeys. The customer's

5.2 The system architecture for the Trip Planner Demonstrator

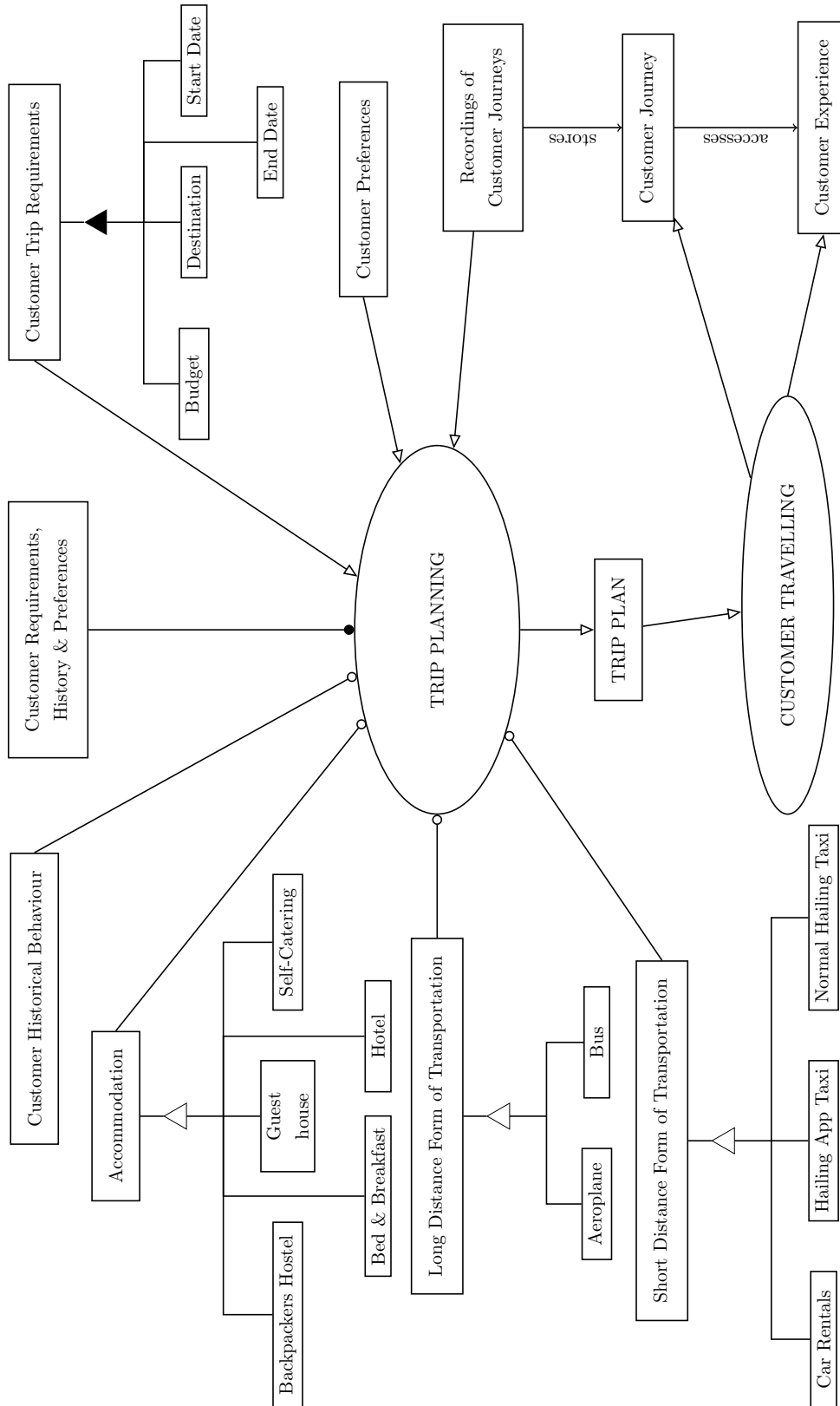


Figure 5.3: Object-Process diagram for the Trip Planner Demonstrator architecture

5.2 The system architecture for the Trip Planner Demonstrator

Accommodation is physical.

Long Distance Form of Transportation is physical.

Short Distance Form of Transportation is physical.

Customer Trip Requirements consists of Budget, Destination, End Date, and Start Date.

Customer Requirements, History & Preferences handles TRIP PLANNING.

Customer Journey accesses Customer Experience.

Recordings of Customer Journeys stores Customer Journey.

Aeroplane is a Long Distance Form of Transportation.

Bus is a Long Distance Form of Transportation.

Car Rentals is a Short Distance Form of Transportation.

Hailing App Taxi is a Short Distance Form of Transportation.

Normal Hailing Taxi is a Short Distance Form of Transportation.

Backpackers Hostel is an Accommodation.

Bed & Breakfast is an Accommodation.

Guest house is an Accommodation.

Hotel is an Accommodation.

Self-Catering is an Accommodation.

TRIP PLANNING requires Customer Historical Behaviour, Accommodation, Long Distance Form of Transportation, and Short Distance Form of Transportation.

TRIP PLANNING consumes Customer Preferences, Customer Trip Requirements, and Recordings of Customer Journeys.

TRIP PLANNING yields TRIP PLAN.

CUSTOMER TRAVELLING consumes TRIP PLAN.

CUSTOMER TRAVELLING yields Customer Journey and Customer Experience.

Figure 5.4: Object-Process language for the Trip Planner Demonstrator architecture

5.3 The Trip Planner Demonstrator database

historical behaviour also contains data about transactions made by the customers in order for the TPD to be able to make offers to the customer during the trip. Once the TPD has completed the trip planning process, the trip plan will be produced. The trip plan consists of an outline of the LDT, SDT and accommodation bookings.

It is important to note the following about the trip planning process. First, there is a fundamental difference between the LDT and SDT. The first type is the mode of transportation used to transport a customer from their home town to the destination, whereas the second type is the mode of transportation which the customer uses at the destination point and at their home town. Secondly, the TPD will only use the recordings of customer journeys if a customer has already completed at least one trip. If the customer has not been on a trip, the TPD cannot use the recordings of customer journeys. Thirdly, the OPL and OPD demonstrate what the trip requirements consist of and the types of accommodation, long distance forms of transportation and short distance forms of transportation which can be used.

The second process, namely the *customer travelling process*, uses the trip plan to execute the trip and it yields a customer journey and CX which are used to populate the touch points on the customer journey. Touch points include offers that can be made to customers along the trip based on their transactional history. Once a trip is completed, the customer journey for the specific trip is stored in the recordings of customer journeys.

It is important to note the following about the customer travelling process. Firstly, before each event of a trip, the customer will be notified by the system of the upcoming event. The notification will occur at the right time and place. Secondly, once an event (also known as touch point) is completed on a trip, the customer can record their experience. By recording the CX at these touch points, it gives feedback to the system on whether the CX has been managed successfully or not.

These two processes will be modelled in the simulator of the TPD, which will be discussed in Section 5.4. The information required and produced by the TPD will be stored in the database, which will be discussed next.

5.3 The Trip Planner Demonstrator database

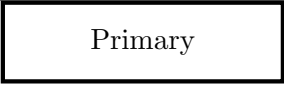
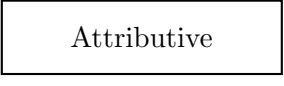
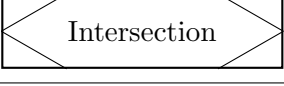
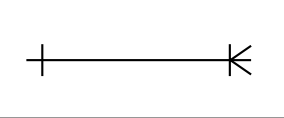
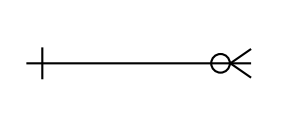
Keeping the system architecture in mind, the database required for the TPD can be constructed. In order to know what data entities to create for the database, an *Entity-Relationship Diagram* (ERD) has to be created based on the system architecture.

An ERD represents the *relationships* between *entities* stored in a database. The entities are any object or event one can collect data about and relationships are the association between the entities (Kendall & Kendall, 2014). One can further extend the ERD to an *Extended Entity-Relationship Diagram* (EERD) in which optionality is added to the relationships and intersection entities are added. Many conventions are available to draw the EERD and for this study the *crow's foot notation*

5.3 The Trip Planner Demonstrator database

will be used. A description of some of the elements used in the EERD for the TPD can be seen in Table 5.2.

Table 5.2: Extended Entity-Relationship Diagram legend (Kendall & Kendall, 2014)

Symbol	Name	Description
	Primary Entity	An entity that can exist on its own.
	Attributive Entity	An entity that requires information from a primary entity.
	Intersection Entity	An entity used to join two primary and/or attributive entities.
	One-to-many relationship (<i>mandatory</i>)	An instance on the one side must be related to at least one instance on the many side.
	One-to-many relationship (<i>optional</i>)	An instance on the one side may be related to no instance on the many side.

An EERD has been drawn up for the TPD. Since the EERD requires a great number of entities, the EERD was split into two parts. Therefore, the two EERD's should not be viewed in isolation, but rather as a unity to understand the overall EERD required for the database. The two EERD's can be seen in Figures 5.5 and 5.6. The entities that appear in both EERD's have been colour-coded for clarity.

Theoretically, the input data required for the TPD can be retrieved from the business partners who signed up for this system. Therefore, it is important that the TPD is built on top of a good cross-functional platform in order for the business partners to share non-sensitive data amongst one another. For the purpose of this study, simulated data will be used in order to overcome ethical clearance issues.

The input data required for the TPD will be populated by the use of simulation where the attributes of the data are based on statistics and information available on the internet. The Matlab software and Excel will be used to simulate and populate the input data. The data created when a customer embarks on the trip, will be simulated by the simulator which will be discussed in the Section 5.4. As mentioned in Section 5.1, MS SQL Server is used as the database server.

Each entity in the EERD will be discussed shortly by providing a brief description of the entity and what it entails. A description of how some of the input data was populated and simulated can be seen in Appendix A.

5.3 The Trip Planner Demonstrator database

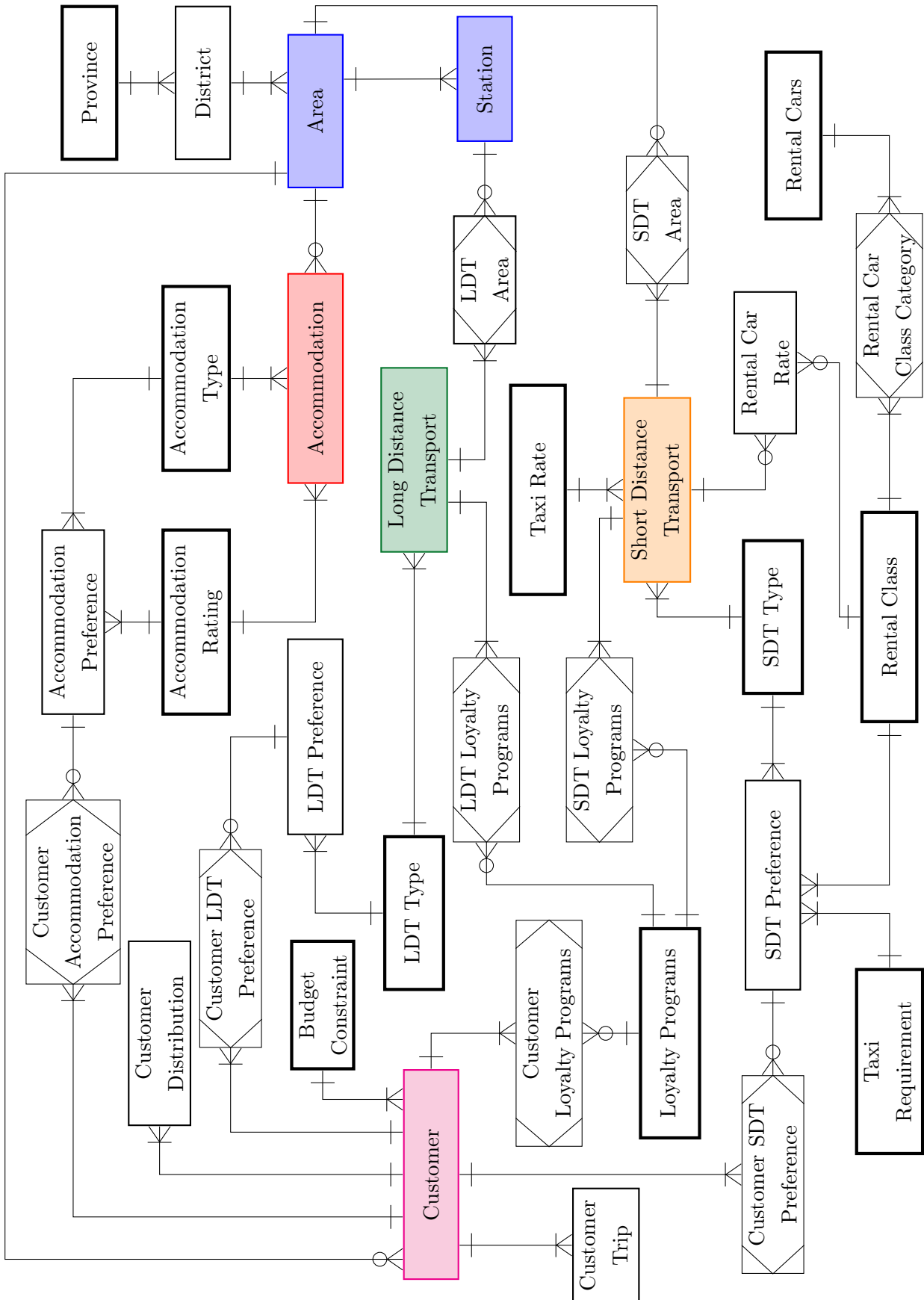


Figure 5.5: Extended entity relationship diagram for the Trip Planner Demonstrator part A

5.3 The Trip Planner Demonstrator database

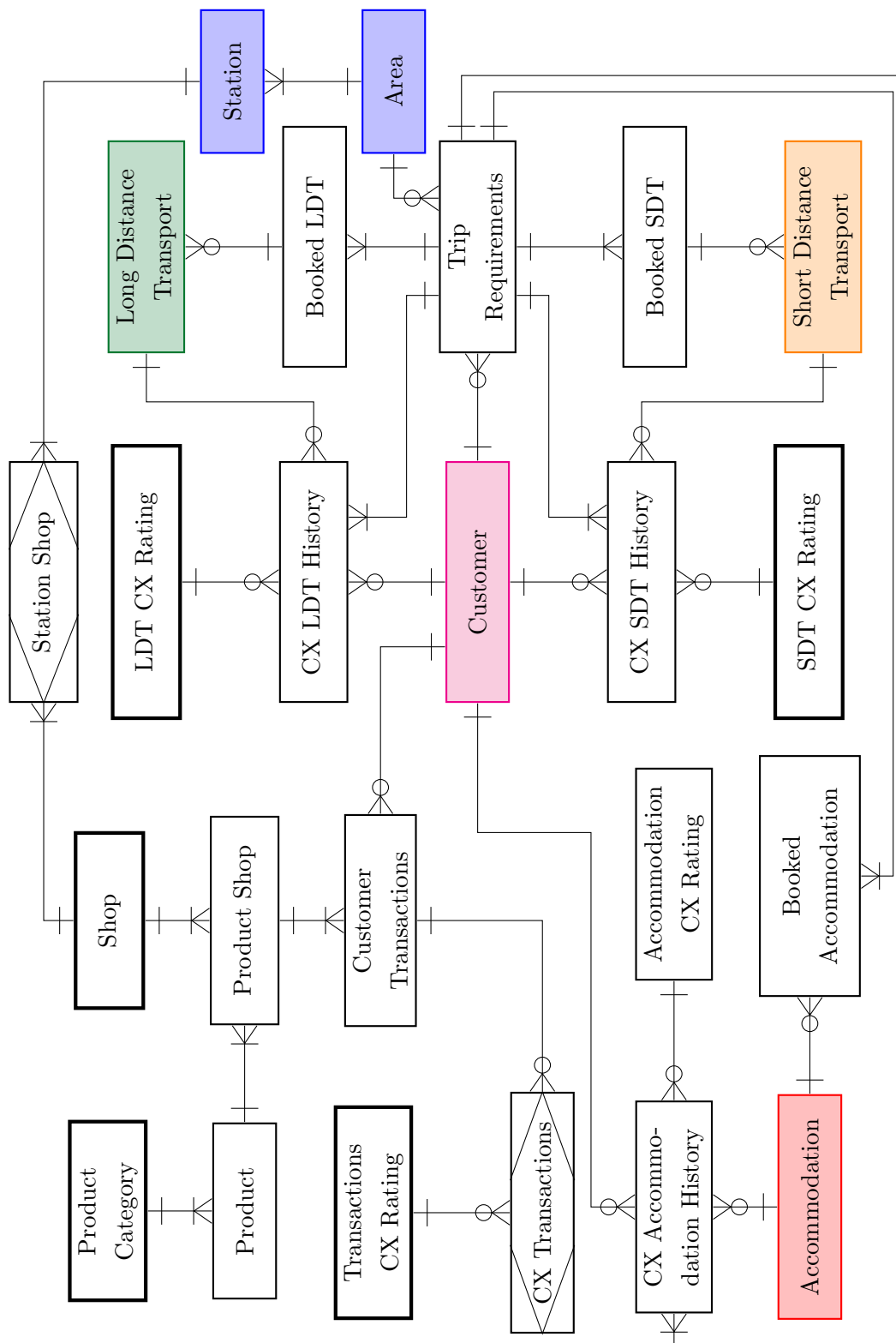


Figure 5.6: Extended entity relationship diagram for the Trip Planner Demonstrator part B

5.3 The Trip Planner Demonstrator database

5.3.1 Accommodation entities

The accommodation entities consist of all entities that relate to accommodation. These entities are required to plan and book a trip for a customer and it also relates to the customer travelling on a trip.

1. *Accommodation Rating*. This is a primary entity which captures the rating of an accommodation based on a scoring system such as the [Tourism Grading Council \(2018\)](#). The rating is an integer that can range from one star to five stars.
2. *Accommodation Type*. This is a primary entity which represents the type of accommodation that an accommodation enterprise can take on. An integer is assigned to the accommodation type and the different types of accommodation are backpackers, bed & breakfast, guest house, hotel or a self-catering hostel.
3. *Accommodation CX Rating*. This is a primary entity which represents the rating which a customer gives after staying at an accommodation place on the trip. The attributes contributing to this rating are the sleep quality, service and cleanliness of the facility. The rating is an integer scale where the values can be from one to five, where
 - (1) is the worst experience,
 - (2) bad or below average,
 - (3) average,
 - (4) good or above average and
 - (5) is an excellent experience.
 - (0) is another value stored in this entity, which showcase the situation when a customer did not enter a CX rating.
4. *Accommodation*. This entity is an attributive entity which stores all the information of the accommodation enterprises that sign up as a business partner to the TPD platform. Table 5.3 gives a description of the attributes used for this entity.
5. *Accommodation Preference*. This is an attributive entity and it captures the combination of accommodation type and accommodation rating which creates an unique accommodation preference. Since there are five accommodation types and five accommodation ratings, there are 25 accommodation preferences.
6. *Booked Accommodation*. This is an attributive entity which is used to store all the booking information regarding the accommodation enterprise booked for a customer's trip. Table 5.4 gives a description of the attributes used for this entity.

5.3 The Trip Planner Demonstrator database

Table 5.3: Description of Accommodation entity

	Attributes	Data Type	Description
1	ID	bigint	The primary key linked to each accommodation record.
2	Name	varchar(50)	The name of the accommodation record.
3	Rate per person per night	money	The costs to stay at the particular accommodation record.
4	Accommodation Rating	bigint	The foreign key that links back to the Accommodation Rating Table.
5	Accommodation Type	bigint	The foreign key that links back to the Accommodation Type Table.
6	Area	bigint	The foreign key that links back to the Area Table.

Table 5.4: Description of booked Accommodation entity

	Attributes	Data Type	Description
1	ID	bigint	The primary key linked to each Booked Accommodation instance.
2	Check-in or out	varchar(50)	Indication as to whether the customer should check-in or out.
3	Date and time	date	The check-in or out date and time.
4	Accommodation ID	bigint	The foreign key that links back to the Accommodation Table.
5	Trip ID	bigint	The foreign key that links back to the Trip Requirements Table.

5.3 The Trip Planner Demonstrator database

5.3.2 Long Distance Transportation entities

The LDT entities are all the entities that relate to the transportation required to get a customer from area x to area y , as well as the entities that will store the interactions of a customer with an LDT while travelling on a planned trip.

1. *LDT Type*. This is a primary entity which represents the types of LDT that are included in the TPD. An integer is assigned to uniquely identify the types of LDT which can either be an aeroplane or a bus.
2. *LDT CX Rating*. This is a primary entity which captures the rating a customer gives after using an LDT on their trip. The rating scale is the same as the Accommodation CX Rating and it also includes a value of 0. The attributes contributing to this rating are the airfares, in-flight service and on-time transport services.
3. *Booked LDT*. This is an attributive entity which is used to store all the booking information regarding the LDT enterprise booked for a customer's trip. Table 5.5 gives a description of the attributes used for this entity.

Table 5.5: Description of booked Long Distance Transportation entity

	Attributes	Data Type	Description
1	ID	bigint	The primary key linked to each Booked LDT instance.
2	Departure or Return	varchar(50)	Indication whether the LDT is departing from or returning to the customer's home area.
3	Date and time	date	The date and time of the flight.
4	LDT ID	bigint	The foreign key that links back to the LDT Table.
5	Trip ID	bigint	The foreign key that links back to the Trip Requirements Table.

4. *Long Distance Transport*. This is an attributive entity which represents all the LDT enterprises that sign up as a business partner to the TPD platform. Table 5.6 gives a description of the LDT's attributes.

5.3 The Trip Planner Demonstrator database

Table 5.6: Description of Long Distance Transportation entity

	Attributes	Data Type	Description
1	ID	bigint	The primary key linked to each LDT record.
2	Name	varchar(50)	The name of the LDT record.
3	LDT Type	bigint	The foreign key that links back to the LDT Type Table.

5. *LDT Preference*. This is an attributive entity which captures the preference that a customer may have for LDT. It is linked to the type of LDT, where it differentiates each type of LDT between a business or economy class. Therefore, an integer is assigned to each preference and the preferences consist of

- (i) aeroplane business class,
- (ii) aeroplane economy class,
- (iii) bus business class or
- (iv) bus economy class.

6. *LDT Area*. This is an attributive entity which captures the link between an LDT enterprise, the stations that it can travel to and from, and the costs. A station is a bus station or an airport. Each LDT enterprise is linked with two stations to create a route, in other words one can travel from the first assigned station to the second assigned station, and *vice versa*. Each LDT enterprise has at least one route. For that specific route, an economy and business class cost is associated with it. It is important to note that not every route has a business class option and not all LDT enterprises have a business class option.

7. *LDT Loyalty Programs*. This is an intersection entity which represents the link between an LDT enterprise and loyalty program(s). A loyalty program can be either a loyalty program provided by the specific enterprise or from an independent enterprise. For the latter part, a business partnership exists between an LDT and loyalty program enterprise. A LDT enterprise can be linked with more than one loyalty program and *vice versa*.

5.3.3 Short Distance Transportation entities

The SDT entities are all the entities that relate to transportation required when a customer is at their destination, when a customer needs to get to the station (bus station or airport) in their home area and all their transportation needs at the destination of the trip.

5.3 The Trip Planner Demonstrator database

1. *Rental Cars*. This is a primary entity which represents all types of cars that can be rented from a car rental enterprise. For example, the car can be a Hyundai i10, Nissan QashQai, or Mercedes C-class. The cars are assigned an integer value to uniquely identify the different rental cars.
2. *Rental Class*. This is a primary entity and it represents the category for a group of rental cars. The cars have been grouped by looking how current rental enterprises in South Africa classify their cars. For the purpose of this study the cars can be classified into any of the following eight classes:
 - (i) *Small Cars*: Hyundai i10 or similar
 - (ii) *Medium Cars*: Ford Fiesta ST or similar
 - (iii) *Large Cars*: Volkswagen Passat or similar
 - (iv) *Premium Cars*: Mercedes C-class or similar
 - (v) *Premium Plus Cars*: BMW i8 Roadster or similar
 - (vi) *People Carriers*: Volkswagen Caravelle or similar
 - (vii) *SUVs*: Toyota Land Cruiser or similar
 - (viii) *SUV Plus*: Peugeot 3008 or similar

Another class has been identified as ‘not rental cars’. This class is used for the SDT Preference which will be explained later on.

3. *Taxi Rate*. This is a primary entity which captures the taxi rate that is linked to the SDT entity. The rate consists of a base rate, rate per minute driving, rate per minute waiting and rate per km. 100 Rates were simulated, but one can add more rates. To explain what the taxi rates typically look like, here are three examples:
 - (a) Taxi Rate Number 1: R10 base rate, R0.70 rate per minute driving, R0 rate per minute waiting and R0 rate per km.
 - (b) Taxi Rate Number 2: R25 base rate, R0 rate per minute driving, R0.25 rate per minute waiting and R0.75 rate per km.
 - (c) Taxi Rate Number 3: R20 base rate, R0 rate per minute driving, R0 rate per minute waiting and R1.50 rate per km.

If the SDT type is car rentals, all the categories will be zero.

5.3 The Trip Planner Demonstrator database

Table 5.7: Description of booked Short Distance Transportation entities entity

	Attributes	Data Type	Description
1	ID	bigint	The primary key linked to each Booked SDT instance.
2	Home or Destination	varchar(50)	Indication whether the SDT is used at the destination or at home.
3	Date and time	date	The date and time that a taxi should transport a customer or when a rental car should be picked up or dropped off.
4	SDT ID	bigint	The foreign key that links back to the SDT Table.
5	Trip ID	bigint	The foreign key that links back to the Trip Requirements Table.

4. *SDT Type*. This is a primary entity which represents the SDT service type that an enterprise offers. The options are car rentals, hailing app taxis or normal taxis. The difference between the latter two is that the hailing app taxis are, for example *Uber* and the ‘normal’ taxis are, for example the *Yellow cab taxis*. The SDT types are uniquely identified by keys.
5. *Taxi Requirement*. This is a primary entity which captures when a customer will require a taxi service. The options for the taxi requirement are (i) at destination, (ii) at destination and home or (iii) none. An integer is assigned to each option to make it unique. A none option is also included for the SDT Preference which will be explained later on.
6. *SDT CX Rating*. This is a primary entity which captures the rating a customer gives after using an SDT on their trip. The rating scale is the same as the Accommodation CX and LDT CX Ratings and it also include a value of 0. The attributes contributing to this rating are the quality of the car, fee structure and all relevant processes followed by the SDT enterprise.
7. *Booked SDT*. This is an attributive entity which is used to store all the booking information with regards to the SDT enterprise booked for a customer’s trip. Table 5.7 gives a description of the attributes used for this entity.
8. *Rental Car Rate*. This is an attributive entity which captures the rental car rate that is linked to a specific rental car enterprise from the SDT table with the rental class. A rental car rate record consists of the rate per day, the age category, the SDT enterprise and the rental class. The rental car rates have been determined by looking at the rates of the current rental car organisations in South Africa and it depends on the SDT enterprise, rental class and the age

5.3 The Trip Planner Demonstrator database

of the customer. It is important to note that not all the SDT car rental enterprises will have all rental classes.

9. *Short Distance Transport*. This is an attributive entity which represents all the SDT enterprises that sign up as a business partner to the TPD platform. Table 5.8 gives a description of the SDT attributes.

Table 5.8: Description of Short Distance Transportation entity

	Attributes	Data Type	Description
1	ID	bigint	The primary key linked to each SDT record.
2	Name	varchar(50)	The name of the SDT record.
3	SDT Type	bigint	The foreign key that links back to the SDT Type Table.
4	Taxi Rate	bigint	The foreign key that links back to the Taxi Rate Table.

10. *SDT Preference*. This is an attributive entity which represents all the preferences a customer may have for SDT. The preferences are uniquely identified by an integer, where each preference is linked to the type of SDT, the taxi requirement and rental class. The entity consists of 12 records. A preference is for example,

- The one preference might be that a customer prefers car rentals. Since the preference is car rental, the taxi requirement will be none. But the rental class needs to be identified. For example, the customer might prefer a class one rental car.
- Another preference might be that a customer prefers hailing app taxi. Since the preference is a taxi, the rental class will be ‘not rental cars’. But the taxi requirement needs to be identified. For example, the customer might prefer to use a taxi only at the destination.

11. *SDT Loyalty Programs*. This is an intersection entity which represents the link between an SDT enterprise and loyalty program(s), where the loyalty program can be provided by the enterprise or due to a partnership between the SDT enterprise and loyalty program enterprise. An SDT enterprise can be linked to more than one loyalty program and *vice versa*.

12. *Rental Car Class Category*. This is an intersection entity which represents the link between the rental cars and the rental class. A car will be in at least one rental class and a rental class has more than one rental car.

5.3 The Trip Planner Demonstrator database

13. *SDT Area*. This is an intersection entity which represents the link between an SDT enterprise and the area in which it travels. An SDT enterprise can operate in many areas and an area can have many SDT enterprises.

5.3.4 Customer entities

The customer entities are all the entities that relate to the customers. In other words, all the entities that are used to store information about the customer, their preferences and historical behaviour and information gathered while the customer takes a trip.

1. *Customer*. This is an attributive entity which contains all the customers who sign up and use this system. A customer always has the option to opt out at any given point in time. Table 5.9 provides a description of the attributes of the customer that are used in the TPD.

Table 5.9: Description of customer entity

	Attributes	Data Type	Description
1	ID	bigint	The primary key linked to each customer.
2	Name	varchar(50)	The name of each customer.
3	Surname	varchar(50)	The surname of each customer.
4	Gender	bigint	The gender of each customer, where 1 is male and 2 is female.
5	Birth year	bigint	The birth year of each customer.
6	Area	bigint	The area is linked back to the area table and it represents the area the customer resides in.
7	Budget Constraint	bigint	The budget constraint is linked back to the budget constraint entity and it represents whether trip cost may be more than the specified trip budget or not.

2. *Customer Distribution*. This is an attributive entity which stores the parameters of the *Poisson* and *exponential* distributions used in the simulator. These two distributions are used to mimic the customer behaviour and will be discussed in more depth in Section 5.4. The entity consists of a beta value for the exponential distribution and a lambda value for accommodation, LDT, SDT and transactions. Each of these values is linked back to a customer as each customer will have their own individual set of parameters as no two customers are the same as discussed in Section 2.5.

5.3 The Trip Planner Demonstrator database

3. *Customer Trip*. This is an attributive entity which stores the latest trip information of a customer. It consists of the latest start date, latest end date and total trips completed using the TPD. This entity is linked back to a specific customer.
4. *CX Accommodation History*. This is an attributive entity which is used to store all the accommodation enterprises at which a customer has stayed. It also captures the recording of the CX rating at the instance when a customer rates their experience at an accommodation enterprise during the trip. The record consists of the customer who used the service, the specific accommodation being rated, the accommodation CX rating allocated, the specific trip when the accommodation is used and the date and time.
5. *CX LDT History*. This is an attributive entity which is used to store all the instances when a customer has used an LDT enterprise. It also captures the recording of a CX rating at the instance when a customer rates their experience with the LDT enterprise. The record consists of the customer who used the service, the specific LDT being rated, the LDT CX rating allocated, the specific trip when the LDT is used and the date and time.
6. *CX SDT History*. This is an attributive entity which is used to store all the instances when a customer has used a service provided by an SDT enterprise. It also captures the recording of a CX rating at the instance when a customer rates their experience with an SDT. The record consists of the customer who used the service, the specific SDT being rated, the SDT CX rating allocated, the specific trip when the SDT is used and the date and time. The allocation of the CX rating for SDT, LDT and accommodation will be discussed in Section 5.4, as well as the population of these tables.
7. *Customer Accommodation Preference*. This is an intersection entity which represents the link between a customer and an accommodation preference. A customer can have more than one accommodation preference and an accommodation preference can be linked to more than one customer.
8. *Customer LDT Preference*. This is an intersection entity which represents the link between a customer and an LDT preference. A customer can have more than one LDT preference and an LDT preference can be linked to more than one customer.
9. *Customer Loyalty Programs*. This is an intersection entity which represents the link between a customer and a loyalty program. A customer can be signed up to more than one loyalty program and a loyalty program can have more than one customer.

5.3 The Trip Planner Demonstrator database

10. *Customer SDT Preference*. This is an intersection entity which represents the link between a customer and an SDT preference. A customer can have more than one SDT preference and an SDT preference can be linked to more than one customer.
11. *CX Transactions*. This is an intersection entity which is used to store all the CX ratings which a customer has entered at all the instances when a customer has been made a transactional offer. In other words, it creates a link between a customer transaction and the transaction's CX rating entered by the customer.

5.3.5 Transaction entities

The transaction entities are all the entities that relate to the transactions made by a customer, the offers that the TPD can make when a customer takes a trip and entities that contribute to the population of transactions.

1. *Product Category*. This is a primary entity which captures the category of products. The product categories for the TPD are as follows,
 - (a) Coffees,
 - (b) Other Hot Drinks (for example; tea, hot chocolate and chai latte),
 - (c) Milkshakes,
 - (d) Sandwiches,
 - (e) Scones and
 - (f) Muffins.

The product categories can be expanded.

2. *Transactions CX Rating*. This is a primary entity which captures the CX rating a customer gives after a transaction has been made on their trip. The rating scale is the same as the Accommodation CX, LDT CX and SDT CX Ratings. The attributes contributing to this rating are the quality of the product, the ambience of the shop and service level.
3. *Customer Transactions*. This is an attributive entity which captures the transactional history of a customer. The transactions include the buying of food and drinks at restaurants, cafés and bars at any station (bus station or airport). This entity captures the customer who buys a product from a shop and the date. The simulation of the customer historical transactions data is discussed in Section 5.4 and Appendix A.
4. *Product*. This is an attributive entity which captures the products that a customer can buy. An integer is assigned to each product to uniquely identify the products.

5.3 The Trip Planner Demonstrator database

5. *Shop*. This is an attributive entity which captures all the shops that are located at the stations. An integer is assigned to each shop to uniquely identify the shops.
6. *Product Shop*. This is an attributive entity which represents the link between a shop and a product. A product and shop are linked and form a new unique primary key to be used in the customer transactions entity. At least one product is linked with one shop and *vice versa*.
7. *Station Shop*. This is an intersection entity which is used to create a link between the shop and station entity. Each shop is linked to at least one station and one station has at least one shop.

5.3.6 Other entities

Additional entities needed for the TPD to plan, book and manage a trip for a customer are discussed next.

1. *Budget Constraint*. This is a primary entity which represents the budget constraint for a customer. It can be either (1) hard constraint or (2) soft constraint. A hard constraint is when the total cost of the trip should be less or equal to the budget; whereas a soft constraint means that the trip cost may be more than the budget.
2. *Loyalty Programs*. This is a primary entity which represents the loyalty programs for which customers signed up and the business partnerships between the loyalty program company and SDT or LDT company. The only data entities that are considered to be linked to the loyalty programs are the car rental and hailing app taxis from the Short Distance Forms of Transportation and the Long Distance Forms of Transportation. The loyalty programs are used in the TPD since a customer buying behaviour is greatly influenced by it (Truth, 2017).
3. *Province*. This is a primary entity which contains the nine provinces of South Africa.
4. *Area*. This is an attributive entity which contains all the areas in South Africa.
5. *District*. This is an attributive entity which contains all the districts in South Africa.
6. *Station*. This is an attributive entity which represents the stations used by an LDT company. A station record consists of the station name and the area where it is located. Station is a collective name for the bus stations and airports.
7. *Trip Requirements*. This is an attributive entity which represents the trip requirements entered by a customer. A trip requirement consists of the customer, destination, budget and the start and end date of the trip.

5.4 The simulator for Trip Planner Demonstrator

Now that a short description has been given of each entity and relationships have been determined between the entities as can be seen in Figures 5.5 and 5.6, the simulator can be constructed. The simulator will be discussed next.

5.4 The simulator for Trip Planner Demonstrator

The simulator enables the processes of the TPD, which is the trip planning and customer travelling processes.

To construct the simulator, an appropriate methodology is required. The steps required for a simulation study have been defined by Banks *et al.* (1996) and can be seen in Figure 5.7.

From these steps, it can be concluded that steps 1 and 2 have already been completed in Chapter 1, step 3 has been completed in Sections 5.1 and 5.2 and step 4 has been performed in Section 5.3. The construction of the simulator will take place in step 5 in which a *model translation* is required. In other words, the simulator will model the processes of the TPD. The next chapter will discuss steps 7 to 10.

For the construction of the simulator it is important to make the necessary assumptions as well as develop a concept model. First, assumptions are required to simplify the model development and to ensure that the simulator will meet the goals of the study (Arena, 2012). These assumptions have been made throughout the model development, in other words during the construction of the system architecture, database and simulator.

Secondly, a concept model should be drawn up before the simulator can be translated into a computer model. The purpose of the concept model is to ensure that the simulator makes sense logically and it is easier to code the computer model and verify assumptions (Bekker, 2015). The concept model can either be in the form of pseudo code and/or a block diagram. The block diagram for the concept model of the simulator is shown in Figure 5.8.

As can be seen from the concept model it illustrates the logical flow of the TPD. After the concept model is drawn up, a software package should be chosen to convert the concept model into a computer model. After careful consideration, Matlab was chosen as it was seen as an easier software package to simulate unique trips for many customers and to evaluate the touch points for every trip which a system undergoes. Another reason for using Matlab is because it will be used for the data analytics function.

5.4 The simulator for Trip Planner Demonstrator

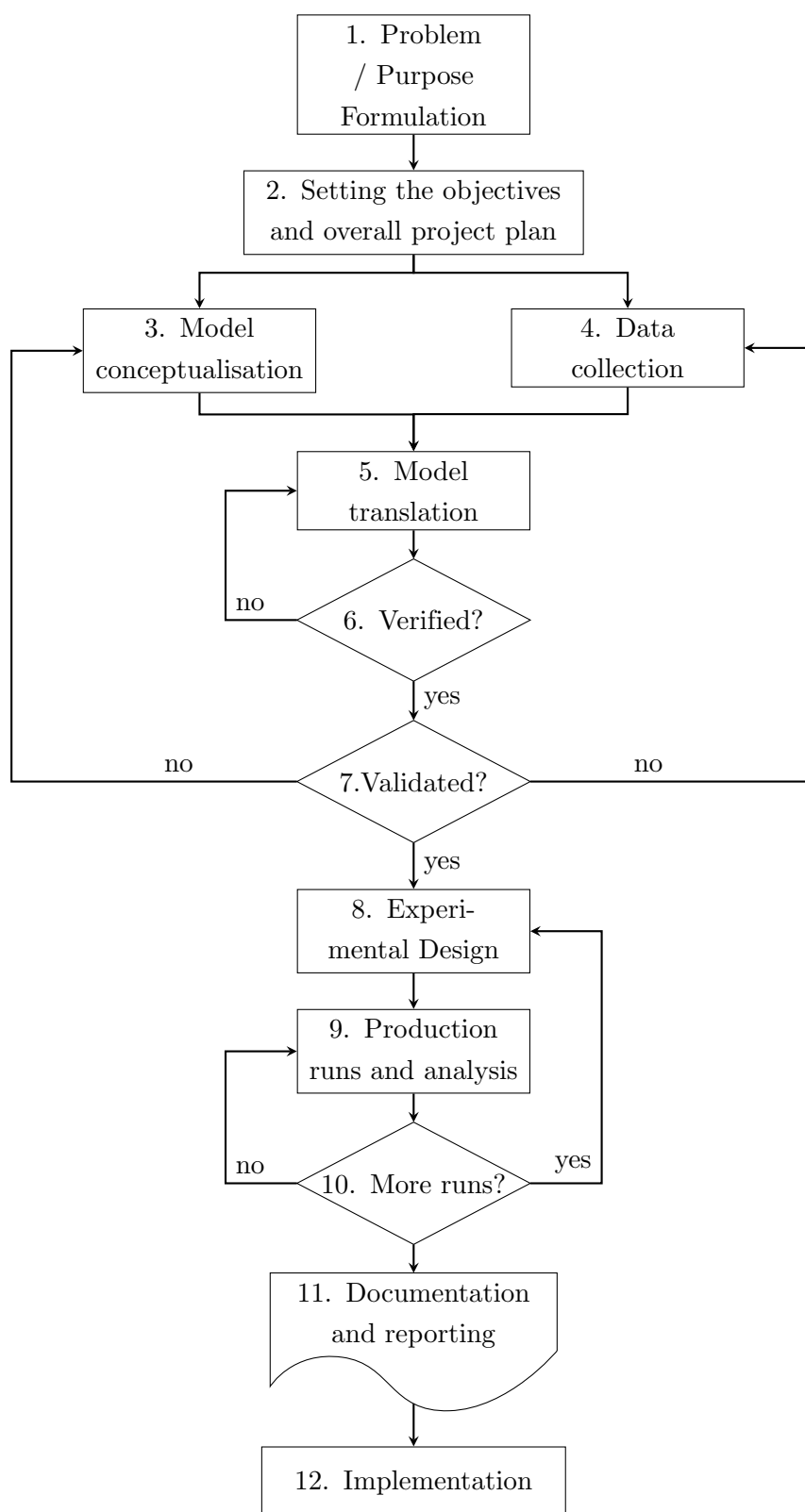


Figure 5.7: Steps in a simulation study (Banks *et al.*, 1996)

5.4 The simulator for Trip Planner Demonstrator

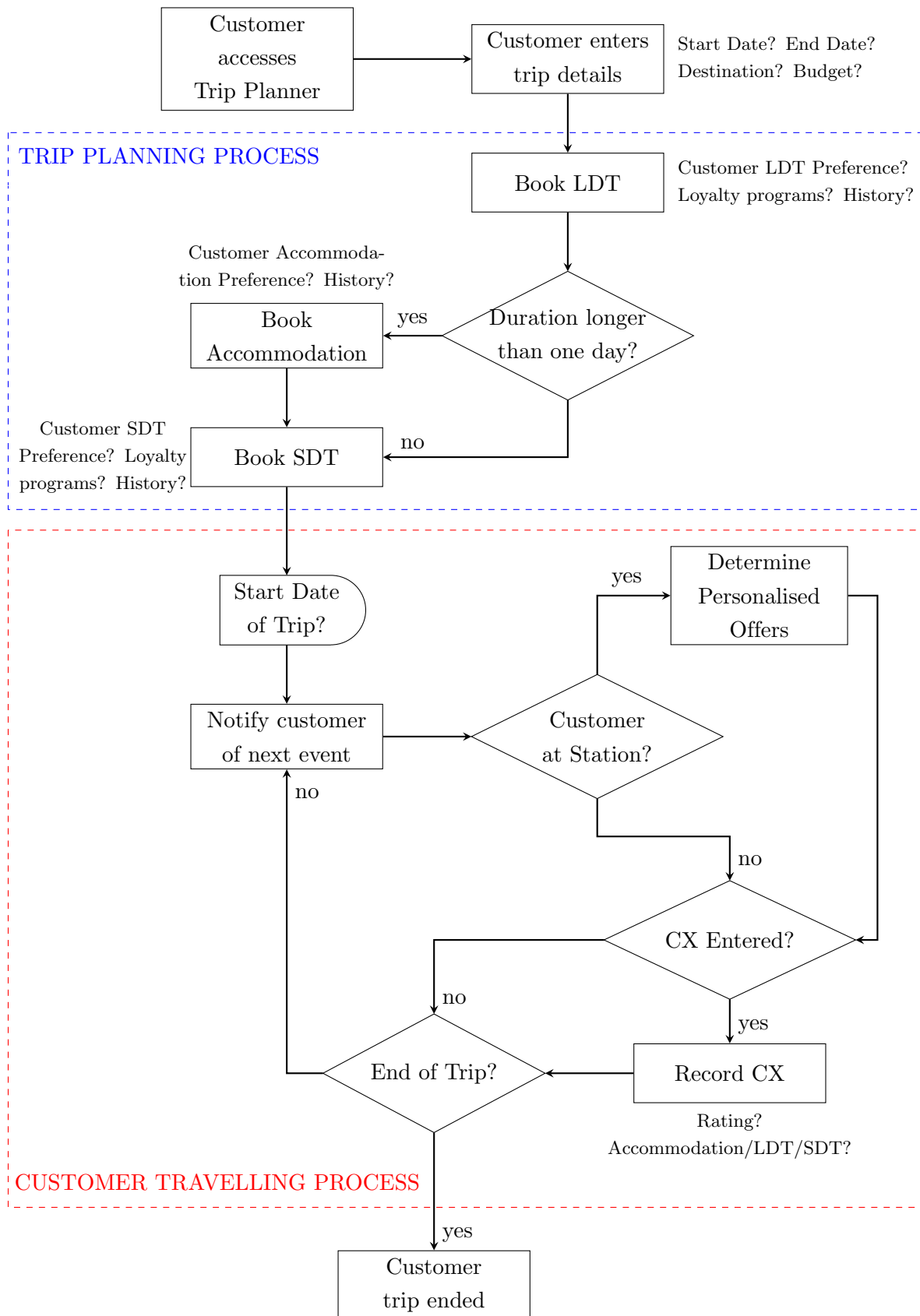


Figure 5.8: The simulator concept model

5.4 The simulator for Trip Planner Demonstrator

After the software package has been chosen, the construction of the TPD simulator can begin. As seen in the concept model in Figure 5.8 and the system architecture in Figure 5.3, the processes can be split into two, namely

1. *Trip planning* process, whose prime functionality is to plan and book a trip for a customer based on the trip requirements, history, preferences and loyalty program.
2. *Customer travelling* process, whose prime functionality is to record the touch points when a customer takes a trip.

For the trip planning and customer travelling processes to occur, the customer first has to access the system and enter their trip requirements. Since the TPD is a demonstrator, the simulator has to simulate these steps as well. The following subsections will provide detail on these processes. For the simulation of trips, information about trip statistics from Lake (2018) and Graft (2018) have been used.

5.4.1 Customer access system process

For the TPD to start with the trip planning process, the customer first has to access the system and enter the trip requirements. It is important to note the following,

- The TPD is for individuals and it cannot be used to plan a trip for a group of people.
- Only customers between the age of 18 and 70 can use this system as a person is of legal age from the age of 18 and only a small portion of seniors use technology (Anderson & Perrin).
- A customer has to sign up to use this system and they have the freedom to opt-out of the system at any given point in time.
- A customer can sign up to the system at any point in time. All relevant data relating to a customer will be pulled through once the customer signed up.

When a customer accesses the system and enter their trip requirements, the simulator will perform the following actions:

1. A customer is picked by using a *beta* distribution, where $\beta = \alpha = 1.5$ as can be seen in Figure 5.9.
2. A start date is chosen based on previous completed trips. If it is the customer's first trip, it is chosen based on a minimum start date. The end date is determined by the duration of the trip. The duration of the trip can be between one and 14 days and is determined by using a *beta* distribution, where $\beta = 1.5$ and $\alpha = 4$ as can be seen in Figure 5.9.

5.4 The simulator for Trip Planner Demonstrator

3. The destination will be determined by a trip area table. It is important to note that trips will only be booked from one district to another and where LDT routes are available. For example the system will book a trip from Cape Town to Beaufort West, but not from Cape Town to Stellenbosch. Therefore, the trip area table represents these options.
4. The budget of the trip is determined based on the duration of the trip together with the destination area of the trip and home area of the customer.

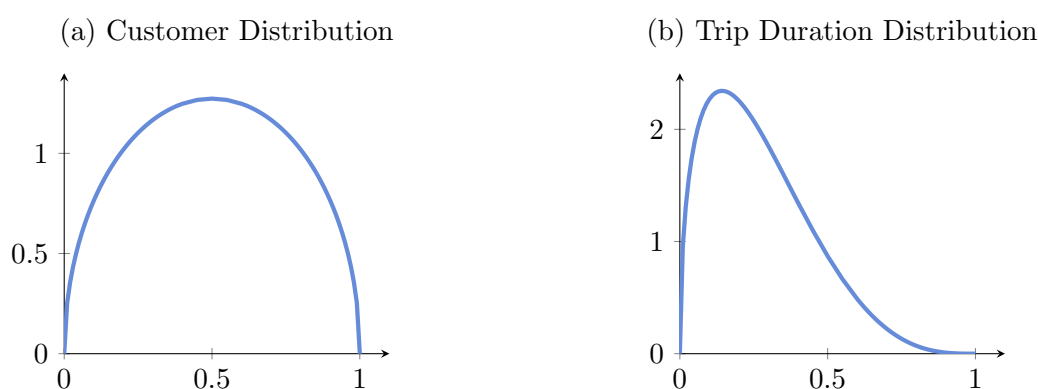


Figure 5.9: Beta distributions for trip requirements

Once the trip requirements have been determined, the simulator can start with the trip planning process.

5.4.2 Trip planning process

Based on the trip requirements, customer's historical behaviour and preferences, a trip can be planned and booked for the customer. It is important to note the following,

- When a booking is made for accommodation, LDT and SDT, it is assumed that it is available and not fully booked. No special check will be done for this as the system should automatically update the availability of these entities.
- The update of the database occurs automatically on a continuous basis.
- Customers are allowed to change their preferences at any given point in time.
- When a customer is using the system for the first time, historical behaviour can be gathered from business partners. For this study, the historical data was simulated by linking a customer respectively with accommodation, LDT and SDT enterprises based on their preferences and/or loyalty programs. The customer history was simulated by assuming that

5.4 The simulator for Trip Planner Demonstrator

- ◇ 5% of customers have no history,
- ◇ 10% of customers will use a different enterprise each time,
- ◇ 30% of customers will use the same enterprise each time if possible, and
- ◇ 55% of customers will use at least two or more enterprises.

This was done to simulate different types of customers.

Figure 5.10 shows the three phases that the simulator performs during the trip planning process which is taken from the concept model in Figure 5.8. All three phases include the actions which are required to book a full trip for a customer.

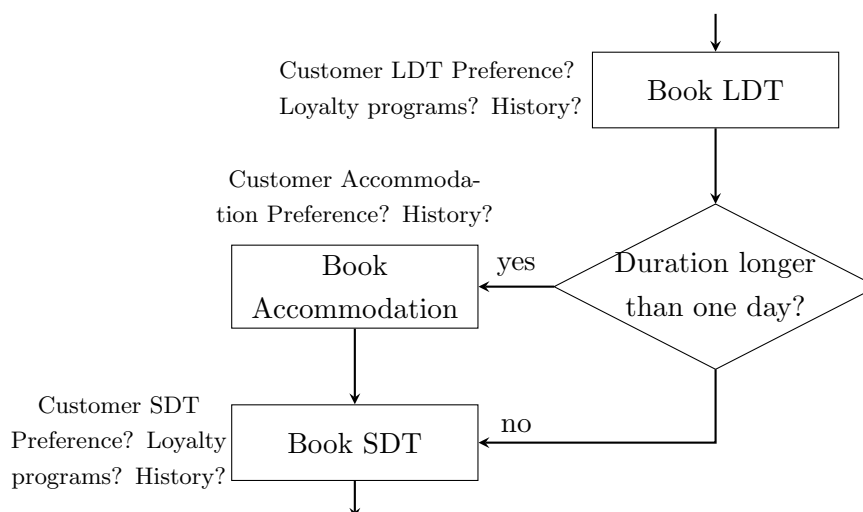


Figure 5.10: The trip planning process

For the booking of Accommodation, LDT and SDT the simulator uses the

1. customer history which includes previous CX entered when a trip was taken by that customer,
2. the customer's preference and
3. the loyalty programs for LDT and SDT.

The history table was populated before the implementation of the TPD as discussed earlier on and the CX was taken as an average CX (value of 3) because the assumption is made that only the history of the customer is available when they sign up to the system, but not their CX ratings.

Therefore, before the booking phases commence, the simulator first has to determine the average CX a customer has per Accommodation, LDT and SDT enterprise respectively. An appropriate method has to be chosen to take into consideration customers who might have a customer behaviour.

5.4 The simulator for Trip Planner Demonstrator

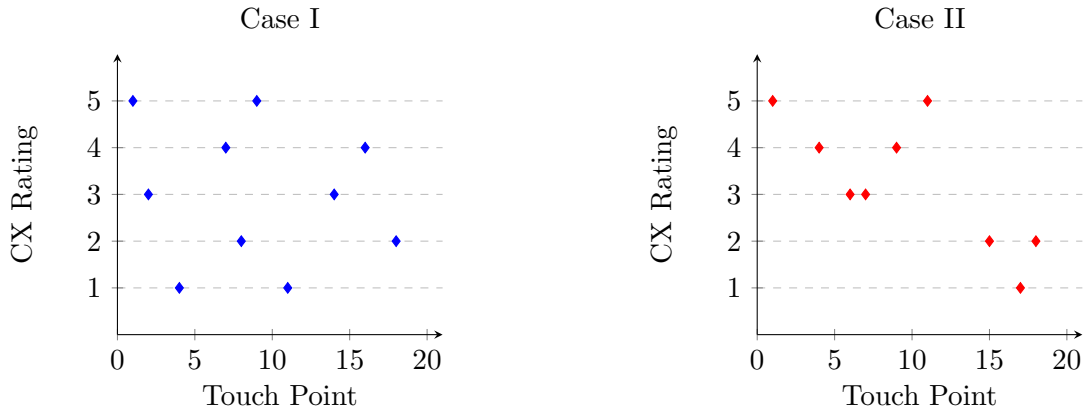


Figure 5.11: Two cases of customer behaviour

A method was proposed and is explained next with reference to Figure 5.11.

Case I and II in Figure 5.11 represent two different customer behavioural trends as to how and when a customer has rated their CX at various touch points with an enterprise. *Case I* represents a customer whose CX Rating fluctuates drastically from touch point to touch point. *Case II* represents a customer whose CX Rating has been average or above average during the first few touch points, but later their experiences were below average.

Therefore, in order to cater for these cases the *weighted average* method is used as shown in (5.1), where the touch points are divided into three groups.

1. Group one: The latest touch points which were rated (g_1).
2. Group two: The middle group of touch points which were rated (g_2).
3. Group three: The oldest touch points which were rated (g_3).

$$\begin{aligned}
 CX_Ave &= w_1 \frac{\sum_{i=1}^{g_1} CX_i}{g_1} + w_2 \frac{\sum_{i=g_1}^{g_2} CX_i}{g_2} + w_3 \frac{\sum_{i=g_2}^{g_3} CX_i}{g_3} \\
 \text{where, } g_1 &= \text{rounddown}\left(\frac{Total_CX}{3}\right) \\
 g_2 &= \text{round}\left(\frac{Total_CX}{3}\right) \\
 g_3 &= \text{roundup}\left(\frac{Total_CX}{3}\right).
 \end{aligned} \tag{5.1}$$

The CX_i represents the CX rating at touch point i . The weights assigned to each group in (5.1) are $w_1 = 0.6$, $w_2 = 0.25$ and $w_3 = 0.15$. These weights have been determined as the ‘best’ by testing various weights in Excel with different scenarios. The testing was done by determining how many errors have been made when an enterprise will be used again or not based on a $CX_Ave > 3$.

5.4 The simulator for Trip Planner Demonstrator

The recording of the CX at each touch point will be discussed in more depth in the customer travelling process, but it is important to note that when a customer does not enter a CX, a rating of 0 is recorded. Therefore, when the weighted average method is applied to determine the CX_{Ave} , it does not take into account the ratings with a value of 0.

However, when only one or two touch points have been rated, (5.1) cannot be used. To take care of these cases, the CX_{Ave} can be determined as follows:

1. If one touch point has been rated, the $CX_{Ave} = CX_1$.
2. If two touch points have been rated, the $CX_{Ave} = w_1 * CX_1 + (w_2 + w_3) * CX_2$.

After the average customer CX has been determined per Accommodation, LDT and SDT enterprise respectively, the booking phase can commence.

For the booking procedure to occur, an appropriate decision-making method should be used. For the purpose of the TPD, the simulator should book an enterprise, based on (i) the customer's historical data, (ii) the CX ratings entered by the customer, (iii) the preferences of the customer and (iv) the loyalty programs.

For the simulator to be able to do this, the decisions can be made by the use of rules.

Therefore, the trip planning process of the simulator should make use of a *Rule-based* application which incorporates *Rule-Based classification*. Rule-based classification is an appropriate ML technique since it uses **IF–THEN** rules to classify a data class (Han *et al.*, 2012). For the case of the trip planning process, it will classify which accommodation, LDT or SDT organisation to book for a customer.

Algorithms 1, 2 and 3 represent the pseudo code that is used to book the accommodation, LDT and SDT for a customer, respectively. In the algorithms the **SQL** indicates when the database in SQL needs to be accessed and the **blue words** indicate which attributes are specifically used to extract, sort or process the data of the customer. Each of the three booking phases will be discussed next.

5.4.2.1 Phase 1: Book LDT enterprise

The first phase of the trip planning process is to book the LDT enterprise that will be used by the customer on the trip. Algorithm 1 provides the steps which determine the LDT that needs to be booked for a customer. For the booking of an LDT enterprise, there are four steps required to determine the 'best-fit' LDT for a customer to give them a 'superior' CX during their trip.

The first step is to create a link between a customer and the LDT enterprises. It is done based on comparing the loyalty program(s) of a customer and their LDT preference(s) with the loyalty program(s) linked to an LDT and type of LDT. The customer's LDT preferences should be linked to the LDT type, before making this comparison.

5.4 The simulator for Trip Planner Demonstrator

The second step is to determine what LDTs are available to transport the customer to their desired destination. In other words, the available LDTs should be determined based on the route, where the route is determined by the customer's home area and the destination of the trip.

The third step is to determine all LDT options for the customer. To perform this step, the available LDTs are compared with customer LDT links, to determine if there is a match between an available LDT and customer LDT link. The following three scenarios can occur:

1. If the LDT corresponds, a value of 1 is assigned to show it is a 100% match.
2. If only the LDT type matches, a value of 2 is assigned to show that the type of LDT matches, but not the Loyalty Program.
3. If is no match is found, a value of 3 is assigned, to show that the LDT is not of the right type. For certain routes only buses or aeroplanes are available, so cases like these will happen when a customer prefers the one type of LDT, but only the other LDT type is available.

Once the *Option_Value* has been assigned to each available LDT, a comparison is made with the *Option_Value* and the average historical CX ratings entered for that specific LDT. If no CX rating is available, the *CX_Ave* is taken as 3. If the conditions are met as set out in Algorithm 1, the LDT is added to the corresponding options group and the cost of the LDT is determined based on the customer LDT preference and loyalty program. If a loyalty program is linked to the LDT, the cost will be discounted by 10%.

The fourth step is to choose and book an LDT. Therefore, after the final options have been determined, the LDT can be booked based on the LDT options and whether the customer has a hard or soft budget constraint, to get the 'best-fit' LDT option for the customer.

Algorithm 1 Ruled-Based Application for Booking of LDT

```

1: procedure CUSTOMER LDT LINK [Cust_LDT]
2:   Cust_Pref ← SQL Customer LDT Preference {Cust_ID, LDT_Pref_ID}
3:   Determine LDT_Type from LDT_Preference {LDT_Type_ID, LDT_Pref_ID}
4:   Cust_LP ← SQL Customer Loyalty Program links {Cust_ID, Loyalty_ID}
5:   LDT ← SQL LDT enterprises {LDT_ID, LDT_Type_ID, Loyalty_ID}
6:   Determine Customer LDT link based on LDT_Type_ID, Loyalty_ID
7: end procedure
8: procedure LDT ROUTES [Available_LDT]
9:   Determine routes {Area_Home_ID, Area_Destination_ID}
10:  Available_LDT ← SQL {dataStation_ID, Area_Home_ID, Area_Destination_ID}
11:  Sort the LDT_ID of Cust_LDT and Available_LDT
12: end procedure

```

5.4 The simulator for Trip Planner Demonstrator

```

13: procedure LDT OPTIONS
14: Cust LDT loop:
15:   Initialise  $i \leftarrow 1$ 
16:   while  $i \leftarrow$  length of Cust_LDT do
17: Available LDT loop:
18:   Initialise  $j \leftarrow 1$ 
19:   while  $j \leq$  length of Available_LDT do
20:     if LDT_ID of Cust_LDT and Available_LDT match then
21:       Assign Option_Value  $\leftarrow 1$ ,  $j \leftarrow 1$ ,  $i \leftarrow i + 1$  and go to Cust LDT loop
22:     else if LDT_Type_ID of Cust_LDT and Available_LDT match then
23:       Assign Option_Value  $\leftarrow 2$ ,  $j \leftarrow j + 1$  and go to Available LDT loop
24:     else
25:       Assign Option_Value  $\leftarrow 3$ ,  $j \leftarrow j + 1$  and go to Available LDT loop
26:     end
27:     if  $j >$  length of Available_LDT
28:        $j \leftarrow 1$ ,  $i \leftarrow i + 1$  and go to Cust LDT loop
29:     end
30:   end while Available LDT loop
31: end while Cust LDT loop
32: Group of LDT Options for customer
33:   Determine LDT options based on Option_Value and CX_Ave with a for loop
34:   IF Option_Value = 1 and (CX_Ave for LDT_ID)  $\geq 3$  THEN
35:     Determine cost based on LDT_ID, Loyalty_ID, LDT_Pref_ID and
     add to First Options
36:   IF Option_Value = 2 and (CX_Ave for LDT_ID)  $\geq 3$  THEN
37:     Determine cost based on LDT_ID, Loyalty_ID, LDT_Pref_ID and
     add to Second Options
38:   IF Option_Value = 3 and (CX_Ave for LDT_ID)  $\geq 3$  THEN
39:     Determine cost based on LDT_ID, Loyalty_ID, LDT_Pref_ID and
     add to Third Options
40: end procedure
41: procedure BOOK 'BEST FIT' LDT
42:   Book LDT
43:   IF First Options not empty and soft budget constraint THEN
44:     Book random LDT from First Options to the extend which the budget allows
45:   IF First Options not empty and hard budget constraint THEN
46:     Book LDT from First Options based on minimum cost
47:   IF Second Options not empty and soft budget constraint THEN
48:     Book LDT from Second Options based on maximum cost to the extend which the budget allows
49:   IF Second Options not empty and hard budget constraint THEN
50:     Book LDT from Second Options based on minimum cost
51:   IF Third Options not empty and hard or soft budget constraint THEN
52:     Book LDT from Third Options based on maximum cost to the extend which the budget allows
53: end procedure

```

5.4 The simulator for Trip Planner Demonstrator

5.4.2.2 Phase 2: Book accommodation enterprise

The second phase of the trip planning process is to book the accommodation enterprise that will be used on the trip. This phase will only happen if the duration of the trip is longer than one day, in other words when $end_date \geq start_date + day(1)$. Algorithm 2 provides the steps which determine the accommodation that needs to be booked for a customer. For the booking of an accommodation enterprise, there are four steps required to determine the ‘best-fit’ accommodation for a customer to give them a ‘superior’ CX during his trip.

The first step is to create a link between a customer and the accommodation enterprise. It is done based on comparing the accommodation preference of a customer with the accommodation type and rate.

The second step is to determine which accommodations are available. The available accommodations can be determined based on the destination of the trip.

The third step is then to determine the accommodation options available for the trip. A comparison is made between the available accommodation and the customer accommodation links, to determine how well the available accommodation type and rate match with the customer accommodation link. The following three scenarios can occur:

1. If the accommodation corresponds, a value of 1 is assigned to show it is a 100% match.
2. If only the accommodation type or rate match, a value of 2 is assigned to show that only one of the attributes corresponds with the accommodation preference.
3. If there is no match found, a value of 3 is assigned, to show the accommodation is not of the right type and rate.

Once the *Option-Value* has been assigned to each available accommodation, a comparison is made between the *Option-Value* and the average historical CX ratings entered for that specific accommodation. If no CX rating is available, the *CX_Ave* is taken as 3. If the conditions are met as set in Algorithm 2, the accommodation is added to the corresponding options group and the cost of the accommodation is then calculated.

The fourth and final step is to book an accommodation based on the *Option-Value* and whether the customer has a hard or soft budget constraint, to get the ‘best-fit’ accommodation option for the customer.

5.4 The simulator for Trip Planner Demonstrator

Algorithm 2 Rule-Based Application for Booking of Accommodation

```

1: procedure CUSTOMER ACCOMMODATION LINK [Cust_Acc]
2:   Cust_Pref  $\leftarrow$  SQL Customer Acc Preference {Cust_ID, Acc_Pref_ID}
3:   Determine Acc_Type & Acc_Rate from Acc_Preference {Acc_Type_ID, Acc_Rate_ID,
   Acc_Pref_ID}
4:   Accommodation  $\leftarrow$  SQL Accommodation enterprises {Acc_ID, Acc_Type_ID, Acc_Rate_ID}
5:   Determine Customer Acc link based on Acc_Type_ID, Acc_Rate_ID
6: end procedure
7: procedure ACCOMMODATION DESTINATION [Available_Acc]
8:   Available_Acc  $\leftarrow$  SQL for Destination {Area_ID}
9:   Sort the Acc_ID of Cust_Acc and Available_Acc
10: end procedure
11: procedure ACC OPTIONS
12: Cust Acc loop:
13:   Initialise i  $\leftarrow$  1
14:   while i  $\leq$  length of Cust_Acc do
15:   Available Acc loop:
16:     Initialise j  $\leftarrow$  1
17:     while j  $\leq$  length of Available_Acc do
18:       if Acc_ID of Cust_Acc. and Available_Acc match then
19:         Assign Option_Value  $\leftarrow$  1, j  $\leftarrow$  1, i  $\leftarrow$  i + 1 and go to Cust Acc loop
20:       else if Acc_Type_ID of Cust_Acc. and Available_Acc match then
21:         Assign Option_Value  $\leftarrow$  2, j  $\leftarrow$  1 and go to Available Acc loop
22:       else if Acc_Rate_ID of Cust_Acc. and Available_Acc match then
23:         Assign Option_Value  $\leftarrow$  2, j  $\leftarrow$  1 and go to Available Acc loop
24:       else
25:         Assign Option_Value  $\leftarrow$  3, j  $\leftarrow$  j + 1 and go to Available Acc loop
26:       end
27:       if j > length of Available_Acc then
28:         j  $\leftarrow$  1, i  $\leftarrow$  i + 1 and go to Cust Acc loop
29:       end
30:     end while Available Acc loop
31:   end while Cust Acc loop
32: Group of Accommodation Options for Customer
33:   Determine Acc options based on Option_Value and CX_Ave with a for loop
34:     IF Option_Value = 1 and (CX_Ave for Acc_ID)  $\geq$  3 THEN
35:       Determine total cost and add to First Options
36:     IF Option_Value = 2 and (CX_Ave for Acc_ID)  $\geq$  3 THEN
37:       Determine total cost and add to Second Options
38:     IF Option_Value = 3 and (CX_Ave for Acc_ID)  $\geq$  3 THEN
39:       Determine total cost and add to Third Option
40: end procedure

```

5.4 The simulator for Trip Planner Demonstrator

```

41: procedure BOOK ‘BEST FIT’ ACCOMMODATION OPTIONS
42:   Book Accommodation
43:   IF First Options not empty and soft budget constraint THEN
44:     Book random Accommodation from First Options to the extend which the budget allows
45:   IF First Options not empty and hard budget constraint THEN
46:     Book Accommodation from First Options based on minimum cost
47:   IF Second Options not empty and soft budget constraint THEN
48:     Book Accommodation from Second Options based on maximum cost to the extend which the
       budget allows
49:   IF Second Options not empty and hard budget constraint THEN
50:     Book Accommodation from Second Options based on minimum cost
51:   IF Third Options not empty and hard or soft budget constraint THEN
52:     Book Accommodation from Third Options based on maximum cost to the extend which the
       budget allows
53: end procedure

```

5.4.2.3 Phase 3: Book SDT enterprise

The third phase of the trip planning process is to book the SDT enterprise that will be used on the trip. Algorithm 3 provides the steps which determine the SDT that needs to be booked for a customer. For the booking of an SDT enterprise, there are four steps required to determine the ‘best-fit’ SDT for a customer to give them a ‘superior’ CX during their trip.

The first step is to create a link between a customer and the SDT enterprises. It is done based on comparing the loyalty program(s) of a customer and the SDT preference with the loyalty program(s) linked to an SDT and the type of SDT. The SDT type for the customer should be determined from the SDT preference, before making this comparison.

The second step is to determine whether the customer needs an SDT at home or not. Based on this, the available SDT can be determined by looking at the destination area and if applicable the home area.

The third step is to determine all the SDT options, for both the destination and home area if applicable. In other words all available SDT’s are then compared with Customer SDT links, to determine how well the available SDT matches with the Customer SDT link. The following three scenarios can happen:

1. If the SDT corresponds, a value of 1 is assigned to show it is a 100% match. If car rentals are preferred, the customer rental class preference is compared with the rental classes offered by the SDT, to determine if it is a match.
2. If only the SDT type matches, a value of 2 is assigned to show that the type of SDT matches, but not the loyalty program.

5.4 The simulator for Trip Planner Demonstrator

3. If no match found has been found, a value of 3 is assigned, to show the SDT is not of the right type or the rental class is not available. For certain areas not all three types of SDT are available and not all rental classes are offered by all car rental STDs enterprises. Therefore, there might be a chance that no match has been found.

Once the *Option_Value* has been assigned to each available SDT, a comparison is made with the *Option_Value* and the average historical CX ratings entered for that specific SDT. If no CX rating is available, the *CX_Ave* is taken as 3. If the conditions are met as set out in Algorithm 3, the cost of the SDT is determined based on the customer SDT preference and loyalty program. If a loyalty program is linked to the LDT, the cost will be discounted by 5%.

The fourth step is to choose and book an SDT for the destination and home area if applicable. It will be done based on the SDT Options and whether the customer has a hard or soft budget constraint, to get the ‘best fit’ SDT option for the customer.

Now that the trip has been planned and booked, the customer travelling process can start once the start date has been reached.

Algorithm 3 Rule-Based Application for Booking of SDT

```

1: procedure CUSTOMER SDT LINK [Cust_SDT]
2:   Cust_Pref ← SQL Customer SDT Preference {Cust_ID, SDT_Pref_ID}
3:   Determine SDT_Type from SDT-Preference {SDT_Type_ID, SDT_Pref_ID}
4:   if SDT_Preference not any form of taxi
5:     Determine rental class {SDT_Pref_ID, Rental_Class_ID}
6:   end
7:   Cust_LP ← SQL Customer Loyalty Program links {Cust_ID, Loyalty_ID}
8:   SDT ← SQL SDT organisations {SDT_ID, SDT_Type_ID, Loyalty_ID}
9:   Determine Customer SDT link based on SDT_Type_ID, Loyalty_ID, Rental_Class_ID
10: end procedure
11: procedure SDT AREA [Available_SDT & Home_SDT]
12:   if Taxi_Requirement is at home and destination
13:     Determine Home_SDT ← SQL {Area_Home_ID}
14:     Sort the SDT_ID of Home_SDT
15:   end
16:   Available_LDT ← SQL {Area_Destination_ID}
17:   Sort the SDT_ID of Cust_SDT and Available_SDT
18: end procedure
19: procedure SDT DESTINATION OPTIONS
20: Cust SDT loop:
21:   Initialise i ← 1
22:   while i ≤ length of text do
23: Available SDT loop:
24:   Initialise j ← 1

```

5.4 The simulator for Trip Planner Demonstrator

```

25:  while  $j \leq$  length of Available_SDT do
26:      if SDT.ID of Cust_SDT and Available_SDT then
27:          Assign Option_Value  $\leftarrow$  1,  $j \leftarrow$  1,  $i \leftarrow i + 1$  and go to Cust SDT loop
28:      else if SDT.Type.ID of Cust_SDT and Available_SDT then
29:          Assign Option_Value  $\leftarrow$  2,  $j \leftarrow j + 1$  and go to Available SDT loop
30:      else
31:          Assign Option_Value  $\leftarrow$  3,  $j \leftarrow j + 1$  and go to Available SDT loop
32:      end
33:      if  $j >$  length of Available_SDT
34:           $j \leftarrow$  1,  $i \leftarrow i + 1$  and go to Cust SDT loop
35:      end
36:  end while Available SDT loop
37:  end while Cust SDT loop
38:  Group of SDT Destination Options for Customer
39:  Determine SDT options based on Option_Value and CX_Ave with a for loop
40:      IF Option_Value = 1 and (CX_Ave for SDT.ID)  $\geq$  3 THEN
41:          Determine cost based on SDT.ID, Loyalty_ID, Taxi_Rate_ID, Rental_Rate_ID
          and add to First Options
42:      IF Option_Value = 2 and (CX_Ave for SDT.ID)  $\geq$  3 THEN
43:          Determine cost based on SDT.ID, Loyalty_ID, Taxi_Rate_ID, Rental_Rate_ID
          and add to Second Options
44:      IF Option_Value = 3 and (CX_Ave for SDT.ID)  $\geq$  3 THEN
45:          Determine cost based on SDT.ID, Loyalty_ID, Taxi_Rate_ID, Rental_Rate_ID
          and add to Third Options
46:  end procedure
47:  procedure SDT HOME OPTION IF APPLICABLE
48:  Cust SDT loop:
49:      Initialise  $i \leftarrow$  1
50:      while  $i \leq$  length of Cust_SDT do
51:  Home SDT loop:
52:      Initialise  $j \leftarrow$  1
53:      while  $j \leq$  length of Home_SDT do
54:          if SDT.ID of Cust_SDT and Home_SDT then
55:              Assign Option_Value  $\leftarrow$  1,  $j \leftarrow$  1,  $i \leftarrow i + 1$  and go to Cust SDT loop
56:          else if SDT.Type.ID of Cust_SDT and Available_SDT then
57:              Assign Option_Value  $\leftarrow$  2,  $j \leftarrow j + 1$  and go to Home SDT loop
58:          else
59:              Assign Option_Value  $\leftarrow$  3,  $j \leftarrow j + 1$  and go to Home SDT loop
60:          end
61:          if  $j >$  length of Home_SDT
62:               $j \leftarrow$  1,  $i \leftarrow i + 1$  and go to Cust SDT loop
63:          end
64:      end while Home SDT loop
65:  end while Cust SDT loop

```

5.4 The simulator for Trip Planner Demonstrator

```

66: Group of SDT Home Options for Customer
67:   Determine SDT options based on Option_Value and CX_Ave with a for loop
68:     IF Option_Value = 1 and (CX_Ave for SDT_ID)  $\geq$  3 THEN
69:       Determine cost based on SDT_ID, Loyalty_ID, Taxi_Rate_ID and add to First Home Options
70:     IF Option_Value = 2 and (CX_Ave for SDT_ID)  $\geq$  3 THEN
71:       Determine cost based on SDT_ID, Loyalty_ID, Taxi_Rate_ID and add to Second Home
       Options
72:     IF Option_Value = 3 and (CX_Ave for SDT_ID)  $\geq$  3 THEN
73:       Determine cost based on SDT_ID, Loyalty_ID, Taxi_Rate_ID and add to Third Home Options
74: end procedure
75: procedure BOOK 'BEST FIT' DESTINATION SDT
76:   Book SDT
77:     IF First Options not empty and soft budget constraint THEN
78:       Book random SDT from First Options to the extend which the budget allows
79:     IF First Options not empty and hard budget constraint THEN
80:       Book SDT from First Options based on minimum cost
81:     IF Second Options not empty and soft budget constraint THEN
82:       Book SDT from Second Options based on maximum cost to the extend which the budget allows
83:     IF Second Options not empty and hard budget constraint THEN
84:       Book SDT from Second Options based on minimum cost
85:     IF Third Options not empty and hard or soft budget constraint THEN
86:       Book SDT from Third Options based on maximum cost to the extend which the budget allows
87: end procedure
88: procedure BOOK 'BEST FIT' HOME SDT IF APPLICABLE
89:   Book SDT for Home
90:     IF First Home Options not empty and soft budget constraint THEN
91:       Book random SDT from First Home Options to the extend which the budget allows
92:     IF First Home Options not empty and hard budget constraint THEN
93:       Book SDT from First Home Options based on minimum cost
94:     IF Second Home Options not empty and soft budget constraint THEN
95:       Book SDT from Second Home Options based on maximum cost to the extend which the budget
       allows
96:     IF Second Home Options not empty and hard budget constraint THEN
97:       Book SDT from Second Home Options based on minimum cost
98:     IF Third Home Options not empty and hard or soft budget constraint THEN
99:       Book SDT from Third Home Options based on maximum cost to the extend which the budget
       allows
100: end procedure

```

5.4 The simulator for Trip Planner Demonstrator

5.4.3 Customer travelling process

After the trip planning phase, the customer travelling phase will start as soon as the start date has been reached. Figure 5.12 shows the phases that the simulator performs during the customer travelling process which is taken from the concept model in Figure 5.8. These phases include the actions which occur by the system and by the customer while a customer is travelling on the planned trip.

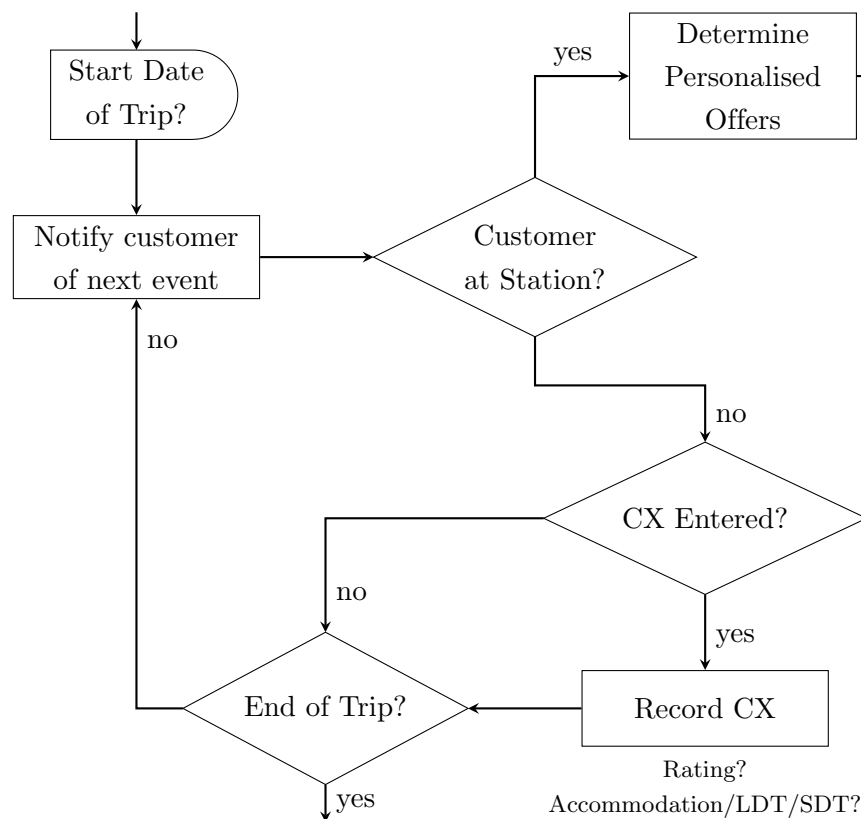


Figure 5.12: The customer travelling process

The first phase of the customer travelling process is the notifications which a customer will receive from the system. The purpose of the notifications is that a customer will get information regarding an upcoming event. The first notification which every customer will receive is when the start date has been reached. The notification reminders are about upcoming events, information regarding the upcoming events and offers. Notifications are there to inform the customer about upcoming events on their trips. The events might include any of the following:

- When the boarding pass for the LDT is available on their mobile device.
- What time the customer should be at the station.
- What time they will be picked up from home and dropped-off at the station and by whom.

5.4 The simulator for Trip Planner Demonstrator

- What time they will be picked up from the station and by whom.
- When it is time to board the LDT.
- Where to pick up the rental car and what documentation to have in place.
- Who will receive them at the accommodation.
- What time they should check out at the accommodation.

These are only a few examples of the notifications that can be sent to a customer. The notifications are unique for every trip and they are related to the upcoming events in the trip plan.

The second phase of the customer travelling process is a checkpoint. After a notification has been sent out, the simulator should check whether the customer is at a station. If the customer is at the station, the simulator should move to the personalised offer phase. However, if the customer is not at a station, the next phase will be the CX rating checkpoint.

The third phase of the customer travelling process is regarding personalised offers when the customer is at a station. The personalised offers are made based on the customer historical transactions. The *Customer Transactions* entity had to be populated before the implementation of the TPD to be able to make personalised offers as discussed in Appendix A. Therefore, the personalised offers are made by using the historical behaviour of the customer and by the use of a rule-based application. The pseudo code is presented in Algorithm 4.

Before the determination of personalised offers can occur, the CX Transactions history needs to be determined. The weighted average method from (5.1) will be used to determine the *CX_Ave* per *Product_Shop_ID* together with the total times (*CX_Count*) the customer has bought the specific product from a shop as presented in *Product_Shop_ID*. If no CX Rating was entered, an average value of 3 is used. After these two parameters have been determined, the personalised offers can be determined.

The number of offers are based on a *triangular distribution* where the upper limit is $b = 1$, the lower limit $a = 0$ and the mode $c = 0.5$. The total offers are then determined based on a random value *Tot_Trans_Offer* from the triangular distribution,

1. If $Tot_Trans_Offer \leq 0.25$, make one offer from categories 4 – 6.
2. If $Tot_Trans_Offer \leq 0.5$, make two offers: one from product categories 1 – 3, which are drink categories and the other one from product categories 4 – 6., which are food categories.
3. If $Tot_Trans_Offer > 0.5$, make one offer from categories 1 – 3.

5.4 The simulator for Trip Planner Demonstrator

Algorithm 4 Ruled-Based Application for Transactions

```

1: procedure CUSTOMER TRANSACTION HISTORY
2:   Cust_His ← SQL Customer Transactions {Cust_ID, Tran_ID}
3:   CX_His ← SQL Customer Transactions {Tran_ID, CX_ID}
4:   Determine from Cust_His and CX_His the CX_Ave and CX_Count for every
     Product_Shop_ID
5:   Link Category_ID to the Product_Shop_ID
6: end procedure
7: procedure SHOPS BASED ON STATIONS
8:   Shop ← SQL Shops at Station based on area {Area_Home_ID, Area_Dest_ID}
9:   Split Shop into Shop_HomeStation and Shop_DestinationStation
10: end procedure
11: procedure DETERMINE OFFER
12:   Compare Customer History with Shops depending on current station Home or Destination
     [Cust_Trans]
13:   Determine Tot_Trans-Offertext to know how many and what category offers to make
     [Trans-Offer]
14:   Compare Category_ID of Cust_Trans and Trans-Offer
15:   if a link has been found and (CX_Ave for Product_Shop_ID) ≥ 3 then
16:     Make Personalised Offer based on maximum CX_Count
17:   else if no link has been found or (CX_Ave for Product_Shop_ID) < 3 then
18:     Compare Product_ID of Cust_Trans and Trans-Offer
19:     if a link has been found and (CX_Ave for Product_ID) ≥ 3 then
20:       Make product offer from shop available at current station based
         on maximum CX_Count
21:     else if no link has been found or (CX_Ave for Product_ID) ≥ 3 then
22:       Make new product offer based on Category_ID offered from shops
         available at current station based on maximum CX_Count
23:     end
24:   end
25: end procedure
26: procedure PROCESS OFFER
27:   Determine if customer accepted offer based on Accept_Rate.
28:   if Offer Accepted then
29:     Record Transaction offer
30:   end
31: end procedure

```

5.4 The simulator for Trip Planner Demonstrator

The acceptance rate of the offer is determined by the type of offer made and a random value. The random value (*Accept_Rate*) is generated and the following conditions are considered to determine whether a customer has accepted the offer or not,

1. If an offer is made from category 4 – 6, the customer will accept the offer when $Accept_Rate < 0.4$.
2. If an offer is made from category 1 – 3, the customer will accept the offer when $Accept_Rate < 0.75$.

The fourth phase of the customer travelling process starts with another checkpoint. After the customer received an offer or when no offer has been made because the customer is not at a station, the simulator should check whether a customer has rated the touch point.

The touch point includes when a customer has accepted an offer, used an LDT service, used an SDT service or used an accommodation service. The rate at which a customer will enter their CX is determined by a random value *CX_E*. Table 5.10 presents the rates at which a customer will enter their CX, where the probability of customers entering a CX Rating is dependable on the touch point itself, whether it is for accommodation, LDT, SDT or transactions.

Table 5.10: Rate of CX entered

<i>Option_Value</i>	1	2	3
Accommodation	$CX_E < 0.4$	$CX_E < 0.5$	$CX_E < 0.6$
LDT	$CX_E < 0.25$	$CX_E < 0.3$	$CX_E < 0.4$
SDT	$CX_E < 0.33$	$CX_E < 0.4$	$CX_E < 0.5$
Transactions	$CX_E < 0.4$		

If one of these conditions is met, the customer will rate the touch point and the simulator will move to the recording of CX. However, it is important to determine what happens when a customer does not enter a CX rating, since there are various touch points along the trip. Two approaches can be followed:

1. Approach one: When a customer does not enter a rating, the system assumes that the customer had an average rating. Reason being that the customer was not impressed or unimpressed enough to rate the touch points.
2. Approach two: When a customer does not enter a rating, the system does not record the CX at that touch point and a value of 0 is assigned.

For the purpose of the TPD the second approach will be used since only the extreme points are of interest. Following the first approach, it will influence the extreme points and that will defeat the

5.4 The simulator for Trip Planner Demonstrator

purpose of the TPD system. Therefore, if a customer did not enter a CX rating, the next phase will be a checkpoint.

The fifth phase of the customer travelling process is the recording of the CX rating entered by a customer. The rating value will be determined by a distribution per Accommodation, LDT, SDT and Transactions: The distribution of the CX for accommodation, LDT, SDT and transaction will follow the *truncated Poisson distribution*,

$$P(x) = \frac{g(x)}{\sum_{x=0}^5 g(x)}$$

$$\text{where } g(x) = \frac{e^{-\lambda} \lambda^x}{x!}$$

and λ will differ for each customer per entity as can be seen in Table 5.11 and it will be in the following range: $2.5 < \lambda < 4.5$.

Since customer behaviour is unpredictable over time, it is important that the λ value per customer used for the above-mentioned distribution should fluctuate for every trip. In order to incorporate this, the change of λ has been determined per customer by using the following truncated exponential distribution,

$$P(\lambda) = \frac{\beta e^{-\lambda\delta}}{e^{-2.5\beta} - e^{-4.5\beta}} .$$

The value of β should also differ per customer as can be seen in Table 5.11 and it will be in the following range: $1.5 < \beta < 5$.

Based on these two distributions, the CX rating of a customer will be determined. Once a customer has entered their rating an extra step is added for the SDT. If the customer had a bad CX, in other words a rating of less than three, the simulator should change the SDT for the rest of the trip, only if it is a normal or hailing app taxi.

The sixth phase of the customer travelling process is another checkpoint. After the customer has entered a CX rating or not, the simulator will check whether the trip has been completed. If the trip has been completed, the customer trip will be signed off as completed and all details will be stored in SQL. However, if the trip has not come to an end the customer's travelling process will start again with the first phase which is the sending of notifications to the customer about the upcoming event.

5.5 Conclusion: The development of the Trip Planner Demonstrator

Table 5.11: The λ values for the customer experience rating distribution

	Customer				
	1	2	3	...	n
Accommodation	λ_{11}	λ_{12}	λ_{13}	...	λ_{1n}
LDT	λ_{21}	λ_{22}	λ_{23}	...	λ_{2n}
SDT	λ_{31}	λ_{32}	λ_{33}	...	λ_{3n}
Transactions	λ_{41}	λ_{42}	λ_{43}	...	λ_{4n}
Rate of Change	β_1	β_2	β_3	...	β_n

After the customer's travelling process has ended, the data around the completed trip will be stored in SQL and the simulator will simulate the next trip for the next customer. The simulator will take into account previous completed trips to ensure that changes are made in the booking phase if a customer had a bad experience at any given touch point on the trip. By repeating this process numerous times, the TPD will be able to effectively manage the CX when a customer takes a trip. It is also important to note the following:

- The update of the database occurs automatically on a continuous basis.
- All issues around governance and protection of data have been taken care of.

5.5 Conclusion: The development of the Trip Planner Demonstrator

In this chapter the development of the Trip Planner Demonstrator was discussed. The model development is the fun part of this study in which the Trip Planner Demonstrator was constructed based on the aim of this study. The aim is that a demonstrator for a digital information and support system, known as the Trip Planner, should demonstrate how a customer's experience can be managed and improved. This can be made possible by simulating unique trips of many customers. Therefore, the construction of the Trip Planner Demonstrator is required.

Firstly, it was important to understand what elements the Trip Planner Demonstrator consists of and its functionalities. It was done by drawing a graphical representation in which it was shown that the Trip Planner Demonstrator consists of four components. The first component was the business partners who are the enablers and value creators. The second component was the database that stores all the required data entities of the Trip Planner Demonstrator. The third component was the simulator whose responsibility it is to simulate the trips and put the Trip Planner Demonstrator in action. The last component was the data analytics function which is used in the simulator and which is responsible to gain knowledge and insights by the application of ML. The latter function will be discussed in the next chapter.

5.5 Conclusion: The development of the Trip Planner Demonstrator

Secondly, a system architecture was defined before the database and simulator of the Trip Planner Demonstrator can be developed. The system architecture was developed by using the Object-Process Methodology, in order to fully understand the needs and requirements of the Trip Planner Demonstrator. From the architecture it was shown that the Trip Planner Demonstrator consists of two processes, namely the trip planning and customer travelling process.

Thirdly, the database was constructed by using MS SQL Server as the platform. Before the database was built in MS SQL Server, an Extended Entity-Relationship Diagram had to be developed. By using this diagram, the data entities and the relationships between them were determined. After the Extended Entity-Relationship Diagram was completed, the data entities were created in MS SQL Server.

Lastly, the simulator of the Trip Planner Demonstrator was constructed. In order to develop the computer model, first a list of assumptions was made and a concept model was constructed. The list of assumptions was put together during the construction of the system architecture, database and simulator and the concept model was presented as a block diagram. The purpose of the simulator is to enable the trip planning and customer travelling process. Matlab was used as the software package for the computer model.

In the next chapter, the verification and validation of the simulator will be discussed, together with the knowledge and insights obtained from the data analytics function.

Chapter 6

Verification and evaluation of the Trip Planner Demonstrator

In the preceding chapters, the research proposal was given, a literature study was conducted and the development of the Trip Planner Demonstrator (TPD) was discussed. This chapter will look at the fourth aspect of the research methodology which is the verification and evaluation of the TPD. Verification and evaluation are required to ensure that the TPD is modelled correctly and that it replicates the concept of the system accurately and analysis is required to gain knowledge and insights from the TPD.

Therefore, this chapter will first discuss the verification. Secondly the evaluation conducted for the TPD will be discussed.

6.1 Introduction: Verification and evaluation

Verification and evaluation are two essential processes that occur during model development. With reference to Figure 5.8, verification is the first process of the two and has an influence on the model translation, whereas evaluation (also known as validation) has an influence on the model conceptualisation and data collection.

Even though these two processes occur at separate stages there is an overlap between them as illustrated in Figure 6.1. The overlap occurs as follows. While the model is being verified, a small part of the evaluation process occurs. Once the model has been verified, the model will be evaluated and changes are made as required. When changes have been applied, the model has to be verified again.

The two processes will be discussed in the next sections.

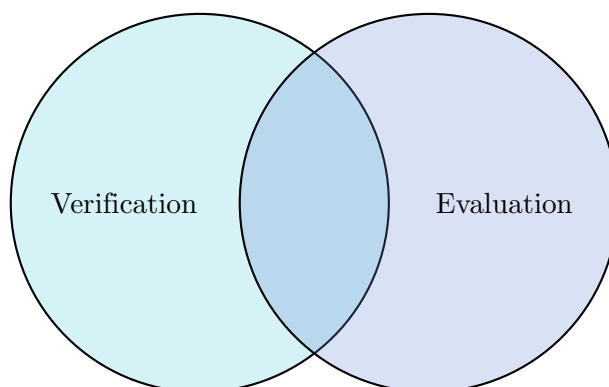


Figure 6.1: Verification and evaluation

6.2 Verification of Trip Planner Demonstrator

The *verification* process is a continuous process that occurs during all phases of the model development. The question addressed by the verification process is: ‘Was the model built correctly?’ and the main action performed is *debugging*. Other actions include correcting syntax issues, checking the logic of the model and correction of compiler errors (Bekker, 2015). After the model has been verified it does not mean that the model is error-free, but rather that no errors were detected by the test regime. Therefore, the focus of the verification process is to not prove model correctness but rather to find and correct errors.

The verification process was applied throughout the development of the TPD. The verification was also executed by tracing the actions performed by the simulator of the TPD, when only one customer has been simulated to go on a trip. By performing this step, debugging had to be done due to errors that occurred during execution while the logic of the simulator was tested at the same time. Various types of customers have been used to ensure that the TPD is built correctly. The scenarios that were used are:

- A customer who has no past CX and is completing a trip for the first time.
- A customer who has previous CX, but all of it was below a rating of 3.
- A customer who prefers an accommodation type and rate that is not available at the destination. For example, a customer who prefers a *5-star hotel*, but the destination of the trip only has *guest houses, bed & breakfast* and *2-star hotels*.
- A customer who prefers an LDT type, but it is not available between their home area and destination. For example, a customer who prefers an *aeroplane* but the customer is going to a destination for which only a *bus* is available for that specific route.
- A customer who prefers an SDT type that is not available at the destination. For example, the customer prefers a *car rental*, but the destination only has *normal taxis* and *hailing app taxis*.

After the verification process was completed, no errors were detected for the TPD using the current dataset. The next step is the evaluation of the TPD.

6.3 Evaluation of Trip Planner Demonstrator

The *evaluation* process is also known as the validation process. The *validation* process occurs at various stages of the model development to address the question: ‘Was the right model built?’. To answer this question, Law & Kelton (2000) recommend that the following three aspects need to be looked at:

6.3 Evaluation of Trip Planner Demonstrator

1. *Concept validity*: Is the system an adequate representation of the real world?
2. *Operational validity*: Is there a link between the model's generated data and the real world system's behavioural data?
3. *Credibility*: Does the model's end-user have confidence in the results generated by the model?

Based on these three questions it is clear that the validation process is required to check how well the model replicates the real-world system. Since this study does not have a 'real-world' system to compare it to, the author prefers to use the word *evaluation*. Reason being that by evaluating whether the TPD replicates and performs the actions of the concept model, the TPD has been 'validated'.

The evaluation process was done at the following stages of the TPD development:

1. After data was simulated for the database, to ensure that the simulated data complies with the statistics and information that was used for the simulation of data.
2. After the construction of the *trip planning* process of the simulator, to ensure that the simulator plans and books a trip for a customer correctly.
3. After the construction of the *customer travelling* process of the simulator, to ensure that the simulator is able to record the CX when a customer has entered a rating; an SDT taxi was changed after a bad CX; and the the travelling data has been stored correctly to be used again for the next trip.
4. After the model runs have been completed, to ensure that the overall model is able to plan and book a trip for a customer and execute the customer's trip.

The first evaluation was performed by comparing the simulated data in the database with the parameters used for the simulation of data.

The second evaluation was performed by looking at all the booking options available for the customers based on their trip requirements and whether the TPD was able to choose the best option for the customer based on their preferences and historical behaviour.

The third evaluation was performed by determining how well the TPD captures the customer travelling on a trip, the simulation of the transactional offers when a customer is at a station and the rating the customer gives the accommodation, LDT, SDT or transaction(s). The latter two evaluations were performed for extreme cases of 1 000 customers who exhibit the properties as presented by the verification scenarios.

The fourth evaluation of the TPD occurred after the final TPD was set in place and the model runs have commenced. The TPD ran for a period of *three simulated years*, where the assumption

6.3 Evaluation of Trip Planner Demonstrator

has been made that a customer goes on an average of *eight trips per year*. The following aspects were then evaluated,

- The distribution of customers that have been travelling and how many trips a customer has taken per year.
- The booking of the accommodation, LDT and SDT per trip compared to the destination of the trip, and the preferences of the customers applicable in the home area of the customer.
- The rate at which the customer has entered a CX rating for accommodation, LDT, SDT and transactions compared to the criteria values in Table 5.10.
- The changes that have been made in terms of accommodation, LDT and SDT during the next or the current trip due to bad CX per customer.

Each of these aspects will be discussed in the next section.

6.3.1 Evaluation: Customers travelling

The first aspect was to evaluate whether the TPD picks the customer in such a way that every customer does not take the same number of trips per year.

As shown in Figure 5.9, the simulator chooses a customer based on a *beta distribution*. A histogram is created to represent how many trips a customer takes over the three-year period as shown in Figure 6.2. The customers have been grouped into 50 classes and each class consists of 150 customers. It is clear from the histogram that every customer does not take the same number of trips and the shape of the histogram is close to that of a particular *beta distribution*.

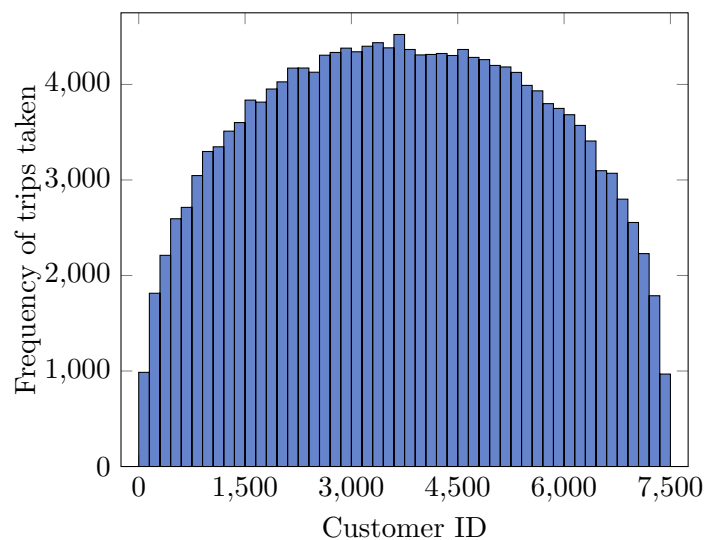


Figure 6.2: Histogram of frequency of trips taken per customer

6.3 Evaluation of Trip Planner Demonstrator

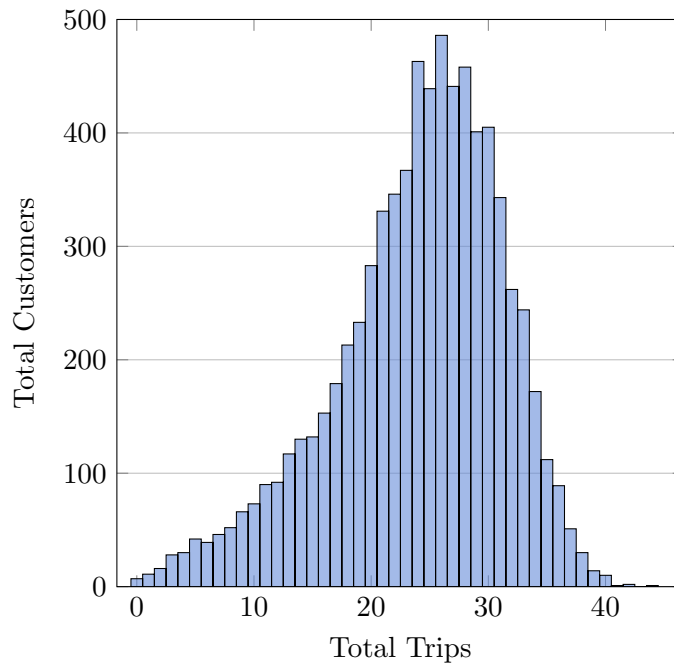


Figure 6.3: Histogram of total trips taken

Another histogram as shown in Figure 6.3 has been created to show the total customers who have taken a certain number of trips over the three years. This was done to further confirm that different behavioural trends have been simulated for the customers. Table 6.1 captures the detail of the total customers who have taken a certain number of trips per year over the three-year period. Based on the frequency of trips taken per customer it is clear that the TPD simulates a different number of trips per customer. By doing this the TPD caters for more- and less-frequent travelling customers. It has also been noted that seven out of the 7 500 customers do not book a trip through the TPD. This is not necessarily a mistake as a lot of users sign up for a system but in the end they do not use it.

6.3 Evaluation of Trip Planner Demonstrator

Table 6.1: A summary of the total trips taken

Total Trips	Average Trips per year	Total Customers	Total Trips	Average Trips per year	Total Customers
0	0	7	22	7.33	346
1	0.33	11	23	7.67	367
2	0.67	16	24	8	463
3	1	28	25	8.33	439
4	1.33	30	26	8.67	486
5	1.67	42	27	9	441
6	2	39	28	9.33	458
7	2.33	46	29	9.67	401
8	2.67	52	30	10	405
9	3	66	31	10.33	343
10	3.33	73	32	10.67	262
11	3.67	90	33	11	244
12	4	92	34	11.33	172
13	4.33	117	35	11.67	112
14	4.67	130	36	12	89
15	5	132	37	12.33	51
16	5.33	153	38	12.67	30
17	5.67	179	39	13	14
18	6	213	40	13.33	10
19	6.33	233	41	13.67	1
20	6.67	283	42	14	2
21	7	331	44	14.67	1

6.3.2 Evaluation: Booking process

The second aspect was to evaluate whether the TPD is able to book the accommodation, LDT and SDT correctly compared to the area(s) and preferences of a customer. The evaluation was done as follows:

1. *Accommodation:*

- The accommodation type and rating of the booked accommodation have been compared to the accommodation preference of the customer, where the preference consists of the accommodation type and rating.
- The area where the booked accommodation is located has been compared with the destination of the trip for the customer.

6.3 Evaluation of Trip Planner Demonstrator

2. Long Distance Transport:

- The type of the booked LDT has been compared to the LDT preference of the customer, where the preference consists of the LDT type and whether the customer prefers business or economy class.
- The route that the booked LDT caters for has been compared with the route between the destination area and home area of the customer.

3. Short Distance Transport:

- The SDT type and rental class of the booked SDT has been compared to the SDT preference of the customer, where the preference consists of the SDT type, the taxi requirement and rental class.
- The area where the booked SDT is located has been compared with the destination of the trip and if a taxi is required at home, the booked SDT has been compared with the home area of the customer.

After the evaluation was done, the results in Table 6.2 were obtained.

For the evaluation based on the area, a 100% accuracy rate has been obtained. This rate is as expected as the TPD should be able to book an accommodation, LDT and SDT accurately based on the area.

However, the TPD did not obtain a 100% rate based on the customer preference. Even though the ideal case should be that the TPD should do it 100% accurately, there are limitations. The reason is because not all types of entities are available for all areas and routes.

Table 6.2: Accuracy of booking process

Accuracy	Preference	Area
Accommodation	90%	100%
LDT	92%	100%
SDT	95%	100%

To further investigate this, the accuracy rate has been evaluated for accommodation, LDT and SDT. The following have been found:

- For *accommodation*: It makes sense that the accommodation entity has a less accurate rating. Taking into consideration that a customer has at least one out of 25 accommodation preferences, the destination should also have each accommodation preference. In other words for a 100% accuracy rate, the destination should have at least 25 accommodation entities that will satisfy the 25 preferences of a customer. For denser areas there is a greater probability that there will be at least one of each type, but for less dense areas it is not.

6.3 Evaluation of Trip Planner Demonstrator

- For *LDT*: It makes sense that it achieved the second worst accuracy rating since there are only two types of LDT available where the aeroplane does not cater for all routes and both types of LDT do not cater for economy and business class.
- For *SDT*: It makes sense that it achieved the best accuracy rating, since all areas have normal taxis. The error in the accuracy is contributed by the fact that not all car rental and hailing app taxis operate in all areas.

Another aspect that was evaluated for the booking process, was to compare the total trip cost with the budget. The following results have been obtained:

- A total of 27 409 of the 180 00 trips were above the budget.
- The total trip exceeded the budget between the range of 0.1% and 135.7%. For example if the budget was R1 000 the total trip cost was between R1 010 and R1 357.
- The portion of hard constraint customers was 4.37% whose trip cost was between the range of 0.1% and 2.79% more than the budget.

Ideally it should have been 0% for the hard constraint customers, but in the real world the actual cost ends up being a little higher than anticipated. For the soft constraint customers it is not a problem as they do not mind if the trip cost is higher than the budget. Therefore this error has been accepted and ignored.

Therefore, the author can state that the TPD is able to book a trip of a customer based on their preferences and historical behaviour with an average accuracy of 92.3%.

6.3.3 Evaluation: Customer Experience ratings

The third aspect was to evaluate whether the TPD can accurately record the CX entered by a customer on a trip.

The accuracy of the recording of CX has been established for all the touch points for which a customer can enter a rating. A comparison was made between the expected results and actual results. The results obtained can be seen in Table 6.3. It is important to note the following:

- The expected results have been determined by using the CX entering parameter as set out in Table 5.10 and multiplying it by the total option values for all the booked entities on the trips.
- The actual results have been obtained by using the database to see how many times a customer has entered a CX rating at a touch point.

6.3 Evaluation of Trip Planner Demonstrator

Table 6.3: Accuracy of Customer Experience rating

Touch Points	Total	Actual Rated	Expected Rated	Percent Error
Accommodation	180 000	80 478	77 562	3.76%
LDT	360 000	108 428	104 364	3.89%
SDT	466 498	162 108	161 874	0.12%
Transactions	341 063	135 743	136 425	-0.50%

The fluctuation is due to the fact that a random variable is used to determine whether the customer has entered a CX rating, which has an influence on the results.

Since the fluctuation obtained between the actual and expected results is less than 5% it has been accepted that the TPD accurately executes the customer travelling process.

6.3.4 Evaluation: Changes due to bad ratings

The fourth aspect was to evaluate whether a change was made in the booking of an accommodation, LDT or SDT entity due to a bad CX rating. For accommodation and LDT the change would only be implemented during the next trip, whereas with the SDT it can be either during the same trip if it is a taxi or for the next trip, irrespective of the SDT type. The evaluation results in Table 6.4 show the percentage changes in the booking of an entity versus the total CX ratings for that smaller than three. It is important to note that the total bad CX will be greater than the actual changes implemented due to the weighted average method that is used for the CX comparison during the trip planning process.

Table 6.4: Changes due to Customer Experience rating

		Percentage Changes	Percentage Bad CX
Accommodation		20.95%	37.32%
LDT		14.44%	41.14%
SDT	– on trip	8.2%	25.87%
	– after trip	12.4%	

It is also a good reflection of how a customer's behaviour fluctuates. If a customer always uses the same entity and they have one bad CX it does not necessarily mean that they do not want to use the service again. Cases like these are also a good opportunity for the relevant enterprise to address the reason behind the bad CX rating.

Table 6.5 presents 11 examples of customers who had a bad CX with a specific entity and for the next trip a new entity was assigned. The IDs used to uniquely identify the customers and entities in the database are used to present the cases. The first examples of each entity can be read as follows:

6.3 Evaluation of Trip Planner Demonstrator

- *LDT Change Example 1*: Customer 2596 went on a trip 21 April 2018 to Area 60. The system booked LDT nr 4 for them. However the customer was unhappy and gave them a rating of 2. For the next trip to Area 60 from 3 November 2018 to 15 November 2018, the system picked up that the customer had a bad experience and the system booked LDT nr 7 instead. After the trip was completed, the customer did not rate the touch point with this particular LDT. On the next trip of 29 July 2019 to 3 August 2019, the system picked up that the customer was unhappy when they used LDT nr 4. Therefore, the system booked LDT nr 7 again and the customer did not rate the touch point with this particular LDT after the trip completion. The last two trips indicated that the customer was happy with new LDT enterprise.
- *LDT Change Example 2 and 3*: These can be read in the same way as example 1.
- *Accommodation Change Example 1*: Customer 691 went on a trip 19 March 2018 to Area 189. The system booked Accommodation nr 435 for them. However the customer was unhappy and gave them a rating of 1. For the next trip to Area 189 from 15 January 2019 to 16 January 2019, the system picked up that the customer had a bad experience and the system booked Accommodation nr 450 instead. After the trip was completed, the customer did rate the touch point with this particular Accommodation as CX rating of 4. The last trip indicated that the customer was happy with new Accommodation enterprise.
- *Accommodation Change Example 2 and 3*: These can be read in the same way as example 1.
- *SDT After Trip Change Example 1*: Customer 1411 went on a trip to Area 140 from 24 April 2018 to 6 May 2018. The system booked SDT nr 18 for them. However the customer was unhappy and gave them a rating of 2. For the next trip to Area 62 from 26 May 2018 to 1 June 2018, the system picked up that the customer had a bad experience and the system booked SDT nr 16 instead. After the trip was completed, the customer was unhappy and gave them a rating of 1. On the next trip to Area 62 from 4 March 2019 to 9 March 2019, the system picked up that the customer was unhappy when they used SDT nr 18 and 16. Therefore, the system booked SDT nr 15 instead and the customer gave this particular SDT a CX rating of 4. On the next trip to Area 62 from 26 April 2020 to 4 May 2019, the system picked up that the customer was unhappy when they used SDT nr 18 and 16. Therefore, the system booked SDT nr 15 instead as the customer did enjoy it in the past and the customer gave this particular SDT a CX rating of 5. The last two trips indicated that the customer was happy with new SDT enterprise.
- *SDT After Trip Change Example 2*: These can be read in the same way as example 1.

6.3 Evaluation of Trip Planner Demonstrator

- *SDT During Trip Change Example 1*: Customer 2069 went on a trip to Area 34 from 20 September 2018 to 23 September 2018. The system booked SDT nr 15 for them. However the customer was unhappy after the first ride with them and gave them a rating of 2. For the next ride required, the system picked up that the customer was unhappy and the system changed the SDT to nr 14. After the second ride, the customer did not enter a rating and it was assumed that the customer is happy and no further changes are required.
- *SDT During Trip Change Example 2 and 3*: It can be read in the same way as example 1.

The last aspect of the evaluation is to present three cases of what a customer journey looks like. Figure 6.4 represents the three customer journeys. It is important to take the following conventions into consideration when looking at the customer journey:

- ■ replicates all the instances at which a customer interacts with the system.
- ■ replicates all the instances when a customer enters data into the system.
- ■ replicates all the instances when the system accesses data from the database with regards to the customer.
- ■ replicates all the actions performed by the system.
- ■ replicates all the output of the system with regards to the booked trip of a customer.
- *SD* replicates the start date for the required trip.
- *ED* replicates the end date for the required trip.
- *Booked Trip* represents the accommodation, LDT and SDT entity booked for the trip based on the ID associated with the relevant entity.
- t_{SD+i} represents the i time instances when an interaction occurs between the customer and TPD system.
- t_{ED} represents when the end date of the trip has been reached.

The trips have been selected at random while ensuring that at least one trip per year has been picked. Each customer has their own unique customer journey and it demonstrates when a customer interacts with the system and *vice versa*. As can be seen no trip is the same and it is unique for every customer. The TPD also simulates unique trips each time the same customer goes on a trip.

6.3 Evaluation of Trip Planner Demonstrator

Table 6.5: Changes in booking of trips due to bad CX ratings

	Customer_ID	Trip_ID	Start Date	End Date	Area_ID	Accommodation		LDT		SDT			
						ID	CX_Rating	ID	CX_Rating	ID	CX_Rating		
LDT CHANGE EXAMPLES:													
1	2 596	20668	21 04 2018	21 04 2018	60	-	-	4	2	-	-		
		45753	03 11 2018	15 11 2018				7	None				
		117241	29 07 2019	03 08 2019				7	None				
2	4 966	41709	25 10 2018	02 11 2018	198	-	-	5	2	-	-		
		125883	21 09 2019	26 09 2019				6	4				
		143591	03 11 2019	11 11 2019				6	4				
3	5 848	55420	13 12 2018	17 12 2018	189	-	-	6	2	-	-		
		110404	23 05 2019	26 05 2019				1	None				
		169066	21 08 2020	21 08 2020				1	5				
ACCOMMODATION CHANGE EXAMPLES:													
1	691	18842	19 03 2018	27 03 2018	189	-	-	1	-	-	-		
		152675	15 01 2020	16 01 2020				4					
2	5 061	35170	23 07 2018	30 07 2018	189	-	-	2	-	-	-		
		61480	18 01 2019	19 01 2019				5					
		161519	29 01 2020	02 02 2020				None					
3	6 683	69100	03 03 2019	06 03 2019	38	-	-	1	-	-	-		
		163871	21 06 2020	26 06 2020				5					
SDT CHANGE AFTER TRIP EXAMPLES:													
1	1 411	24249	24 04 2018	06 05 2018	140	-	-	-	-	-	18	2	
		27351	26 05 2018	01 06 2018							62	16	1
		64309	04 03 2019	09 03 2019							62	15	4
		163120	26 04 2020	04 05 2019							189	15	5

Table 6.5 continues on next page

6.3 Evaluation of Trip Planner Demonstrator

	Customer_ID	Trip_ID	Start Date	End Date	Area_ID	Accommodation		LDT		SDT		
						ID	CX_Rating	ID	CX_Rating	ID	CX_Rating	
2	3 311	25982	05 05 2018	07 05 2018	62					1	1	
		119859	07 08 2019	13 08 2019	62					5	3	
		172994	01 11 2020	05 11 2020	189					5	4	
SDT CHANGE DURING TRIP EXAMPLES:												
1	2 069	41 396	20 09 2018	23 09 2018	34						15	2
											14	None
2	3 848	105 842	09 04 2019	16 04 2019	86						96	2
											71	3
3	7 360	84 052	12 03 2019	19 03 2019	189						123	2
											65	3

End of Table 6.5

6.3 Evaluation of Trip Planner Demonstrator



Figure 6.4: Three examples of customer journeys

6.4 Conclusion: Verification and evaluation of Trip Planner Demonstrator

6.4 Conclusion: Verification and evaluation of Trip Planner Demonstrator

In this chapter the verification and evaluation of the TPD was discussed. The aim of the TPD is to show the abilities of how a customer experience can be managed and improved by the use of data analytics and *business partnering on a cross-functional platform*. Therefore, the model had to be verified and evaluated to ensure that the TPD is able to do it and analysis is required to show the value of such a system.

Firstly, the verification of the TPD was discussed shortly by stating what was done to verify the model. After verification the TPD has been declared as error-free within the given dataset.

Secondly, the evaluation of the TPD was discussed by explaining how it was done. The evaluation was done by evaluating whether the TPD can plan and book a trip and manage the trip while a customer is on the trip. Four evaluations were completed for the TPD. After the evaluation was completed, it can be stated with confidence that TPD is able to simulate unique trips for many customers by ensuring that the customer experiences are managed and improved.

Based on the verification and evaluation, the TPD has the ability to effectively plan and book an entire trip for a customer based on their trip requirements and by using their preferences and historical behaviour as an input to determine the optimal trip. The system also has the ability to manage and improve a customer experience while the customer is travelling. The next step of the process is to perform analysis on the output data of the TPD which will be discussed next.

Chapter 7

Analysis of Trip Planner Demonstrator Data

In the preceding chapters, the research proposal was given, a literature study was conducted and the construction and evaluation of the Trip Planner Demonstrator was completed. This chapter will look at the fifth aspect of the research methodology, which is the analysis of the data generated by the Trip Planner Demonstrator.

The analysis step is essential for this study as this is where the real value of the TPD lies. It has been proved by the verification and evaluation steps that the TPD can effectively manage and improve a CX in the travelling domain by incorporating all aspects of the customer using the database, where the database is built by integrating the data shared amongst the business partners on the cross-functional platform. But to gain insights from the trips completed by the customers and to generate more business value for the partners, data analytics is required.

This chapter will first provide a roadmap which was followed to perform the analysis. Secondly, it will discuss the techniques used to perform data analytics. Lastly, the insights and knowledge gained from performing the analysis will be discussed.

7.1 Roadmap of analysis followed

To extract value from the TPD, data analytics in the form of ML is an essential part of this study. A roadmap is required to set out the steps required to perform ML on the dataset gathered by the trips completed on the TPD.

The roadmap followed during the analysis process is based on the knowledge management processes discussed in Section 3.2.2.3 and it consists of the following steps:

1. Determine which ML tool and technique(s) to use.
2. Preprocessing and transformation of the data for analytical purposes.
3. Performing ML techniques on the dataset(s).
4. Determining knowledge and insights from the analysis.

7.2 Choosing Machine Learning tool and techniques

The first step is to determine the appropriate ML tool. There are four tools to choose from, namely (i) supervised learning, (ii) reinforcement learning, (iii) active learning or (iv) unsupervised learning.

7.3 Preprocessing and transformation of data

To determine which tool to use, one should determine what the purpose is for applying ML.

As mentioned in Section 5.1, the TPD has a data analytics function to apply analytics after a customer has completed a trip in order to learn more about the customer's behaviour at the touch points. This will shed light on the various types of customers that use this system. By determining the unique group of customers, one will be able to know what types of services to propose to them and predict a customer's behaviour more accurately. Therefore, the ML model should be able to learn from the customer data without a test dataset and *unsupervised learning* is the appropriate tool to use for it.

After the tool selection, the appropriate technique should be chosen. The techniques that can be used are grouped as (i) classification, (ii) regression or (iii) clustering. A combination of techniques or only one technique can be used. The chosen technique depends on the dataset that needs to be analysed. Since the analytical purpose of the TPDs data function is to learn the type of customer that uses the system and how they rate the touch points, *clustering* is a good technique to use. Clustering has the ability to use an algorithm to categorise a dataset into two or more groups by finding similarities between the groups.

Now that the ML tool and technique have been chosen, the next step is to preprocess and transform the data.

7.3 Preprocessing and transformation of data

The data gathered after a customer has completed a trip has to be processed and transformed to get useful insights and results from applying clustering techniques on the dataset.

The first step of the preprocessing and transformation of the data was to determine what data will be used for the analysis. The following datasets were chosen to be included in the analysis:

- Customer historical behaviour with respect to the type of accommodation, type of LDT and type of SDT.
- Customer historical behaviour with respect to the CX ratings entered for the accommodation, LDT, SDT and transactional offers.
- Customer historical behaviour with respect to the transactional offers.

The second step was to extract the right views from the database in MS SQL Server into Matlab to get the data ready for preprocessing. When the views were extracted, careful consideration had to take place to ensure that the right attributes are included for analysis. After the views were imported into Matlab, further preprocessing was necessary. Actions performed to preprocess the data include:

7.3 Preprocessing and transformation of data

- Determine the age category per customer from the customer's birth year attribute. The age of the customer was determined by using the birth year of the customer and the current year. The following age categories were used:
 1. Category 1: 18 to 19 years
 2. Category 2: 20 to 24 years
 3. Category 3: 25 to 29 years
 4. Category 4: 30 to 34 years
 5. Category 5: 35 to 39 years
 6. Category 6: 40 to 44 years
 7. Category 7: 45 to 49 years
 8. Category 8: 50 to 54 years
 9. Category 9: 55 to 59 years
 10. Category 10: 60 to 64 years
 11. Category 11: 65 to 69 years
 12. Category 12: 70+ years
- Determine the province per customer by linking the province to the area in which the customer resides.
- Link the type of accommodation, type of LDT and type of SDT respectively with each accommodation, LDT and SDT in the database. This should be done for all accommodation, LDT, SDT used per customer on the trips and for all CX ratings entered by the customers.

After the processing was completed, the data was transformed. Two approaches were followed for the transformation of the data. The first approach is Principal Component Analysis (PCA) and the second approach is the recency, frequency and monetary (RFM) approach.

7.3.1 Principal Component Analysis

The first approach that was used to transform the customer data was PCA and the reason for this approach will be explained.

As can be seen in Table 7.1, the dimensions of the attributes of the data differ significantly. A 3D scatter plot is created for the first 1 000 customers based on the customer gender, customer province and accommodation type as shown in Figure 7.1. It is clear that the different dimensions will have an impact on the clusters.

7.3 Preprocessing and transformation of data

Table 7.1: Dimensions of data attributes

Data Attributes	Dimensions
Gender of customer	2
Age category of customer	12
Province of customer	9
Accommodation type	5
LDT type	2
SDT type	3
CX rating	5

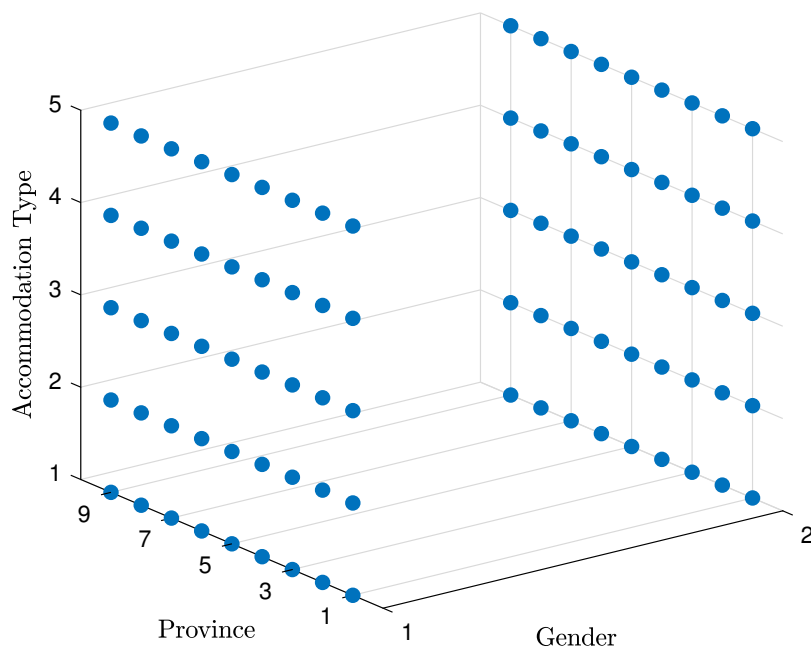


Figure 7.1: Scatter plot representing the different dimensions of the data attributes

7.3 Preprocessing and transformation of data

To overcome the problem of the difference in dimensions, PCA can be used. PCA is a dimensionality reduction technique in which a new coordinate system is established by searching for orthogonal directions (Sorzano *et al.*, 2014). PCA can therefore be used to transform the data before clustering begins, as it has the ability to remove the cluster structure of the dataset. One can take the step even further and apply the weighted PCA on the dataset. By performing the weighted PCA method, the inverse variable variance is used as weights in the analysis and leads to more accurate results (Yue & Tomoyasu, 2004).

For the purpose of this study, the weighted PCA has been used as it gives more accurate clusters than the normal PCA. In Figure 7.2 the weighted PCA has been applied on the dataset used in Figure 7.1. By comparing these two figures it is clear that the weighted PCA gives a better dataset to be used for clustering. The weighted PCA plots used during the analysis of all datasets can be viewed in Appendix B.

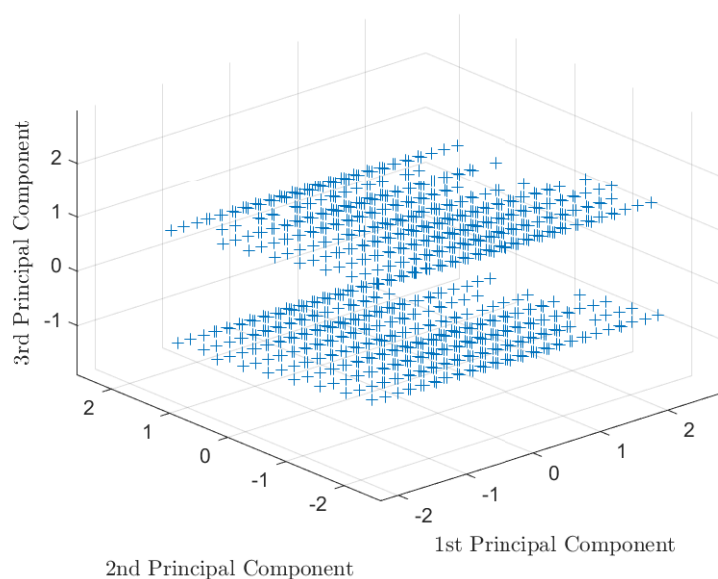


Figure 7.2: Weighted Principal Component Analysis applied to the customer accommodation dataset

7.3.2 Recency, Frequency and Monetary Analysis

The second approach that was used to transform the customer data was RFM. The RFM stands for:

- (R) – Recency: The recency of a customer’s purchase behaviour.
- (F) – Frequency: How many times have a customer purchased something.
- (M) – Monetary: The average monetary expenditure a customer has spend on a purchase.

7.3 Preprocessing and transformation of data

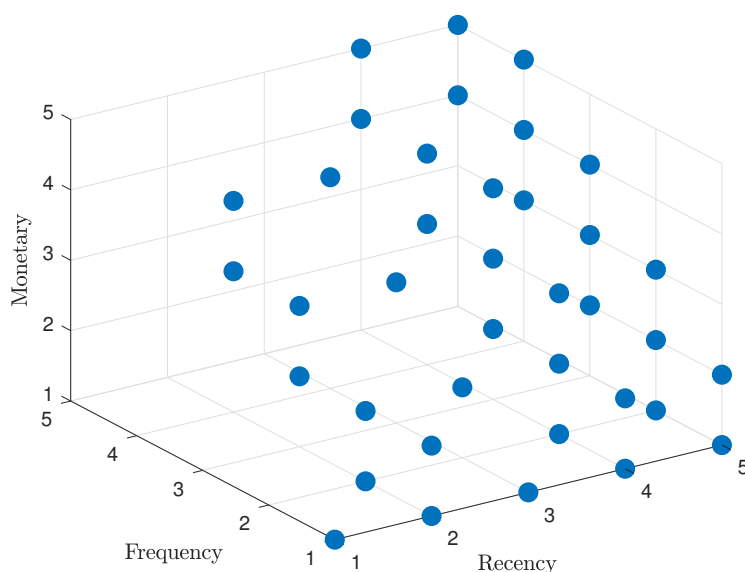


Figure 7.3: Recency, Frequency and Monetary analysis for transactions

The RFM approach gives an enterprise the ability to understand the customer behaviour and group their customers according to these values (Chen *et al.*, 2005; Tsiptsis & Chorianopoulos, 2011). Based on these values one will then be able to apply clustering techniques on the transformed datasets.

The RFM approach was used for analysing the transactions of the customer in the context of the TPD. The *recency* value is the latest purchase made by the customer after the customer has accepted the proposed offer when they are at a station. The *frequency* value is the total transactional purchases made by a customer over all the trips. The *monetary* value is the total money spent on the transactional offer. A purchasing price has been assigned per product per shop to determine the monetary value. The transformed transactional data for all customers can be seen in Figure 7.3

The RFM approach was also used for analysing the average CX ratings entered by the customer for Accommodation, LDT, SDT and transactions over all trips. The *recency* value is taken as the date when the last touch point has been rated by a customer. The *frequency* value is taken as the total touch points that have been rated up to that point. The *monetary* value is taken as the average CX entered over all touch points. Even though CX is not a monetary expenditure for a customer, it is a monetary expenditure for an enterprise as their revenue will increase when they have more satisfied customers. The transformed CX data can be seen in Figure 7.4.

Once the data was transformed by the use of PCA and RFM, the third step of the analysis roadmap can be performed. This step is to perform the chosen ML technique(s) on the dataset to gain knowledge and insight, which will be discussed next.

7.4 Application of Machine Learning techniques

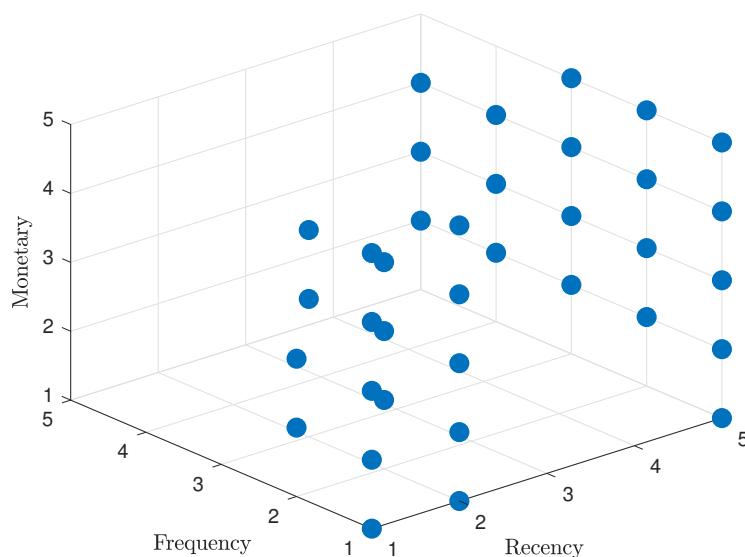


Figure 7.4: Recency, Frequency and Monetary analysis for overall Customer Experience

7.4 Application of Machine Learning techniques

The chosen ML technique has been determined in Section 7.2 as clustering. It will be used as the aim of the data analytics function to have the ability to group the customers based on their behavioural trends.

To determine which cluster technique will be used, a similar scenario is faced as with choosing the ML tool. There is more than one clustering technique available and only one, or more than one technique can be used together. The clustering technique was chosen based on tutorials which the author did in the Matlab environment and by using a trial-and-error approach to see what technique will be the best for the purpose of the study.

The clustering technique used in the end was the *k-Means* clustering technique as it is one of the fastest distance-based clustering techniques which can handle any dimension or size dataset (Tsipitsis & Chorianopoulos, 2011).

To determine the ‘optimal’ number of clusters to be used, the evaluation function in Matlab was used. The evaluation function in Matlab uses the silhouette value. The silhouette value rank the clusters in the range $[-1, 1]$ to determine what total number of clusters will give good clusters and which will not. The silhouette values measure how close a point in one cluster is based on the neighbouring clusters. The greater the value, the better the cluster will be. A silhouette value of greater than 0.5 will produce good clusters (Lletuí *et al.*, 2004).

Therefore, the total clusters used in k-Means clustering was determined by using the silhouette values. Figure 7.5 will be used as an example on how the total clusters can be chosen. As can be

7.4 Application of Machine Learning techniques

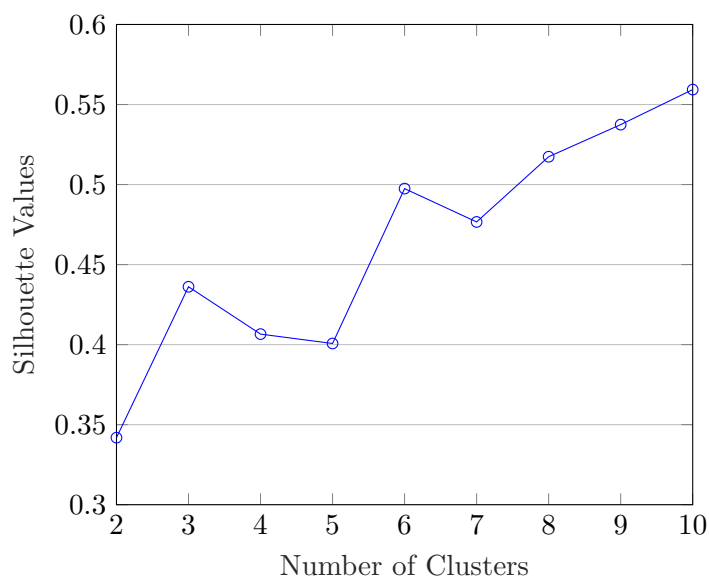


Figure 7.5: Silhouette plot for Customer Experience of accommodation

seen, for a total number of clusters of seven or less, the silhouette value is less than 0.5. Therefore, to get an optimal number of clusters, the minimum number of clusters should be eight. However, cluster nine and ten yield even better silhouette values which indicate that these clusters can also be chosen. However, there exists a trade-off between a good silhouette value and the computational power needed to analyse more clusters. Therefore, for the purpose of the study the minimum number of clusters will be chosen to demonstrate the ability that lays within the TPD.

The silhouette value plot can also indicate why the weighted PCA is better to use compared to the PCA. The silhouette value plot for the accommodation will be used for demonstrative purposes. As can be seen in Figure 7.6, for smaller clusters, the normal PCA yields a better silhouette value. However, as the number of clusters increases the silhouette value for the weighted PCA increases and the silhouette value for the normal PCA decreases. Therefore, the weighted PCA is indeed a better tool to use for bigger clusters. The silhouette value plots used during the analysis of all datasets can be viewed in Appendix B.

After the analysis has been completed, insights and knowledge can be gained, which will be discussed next.

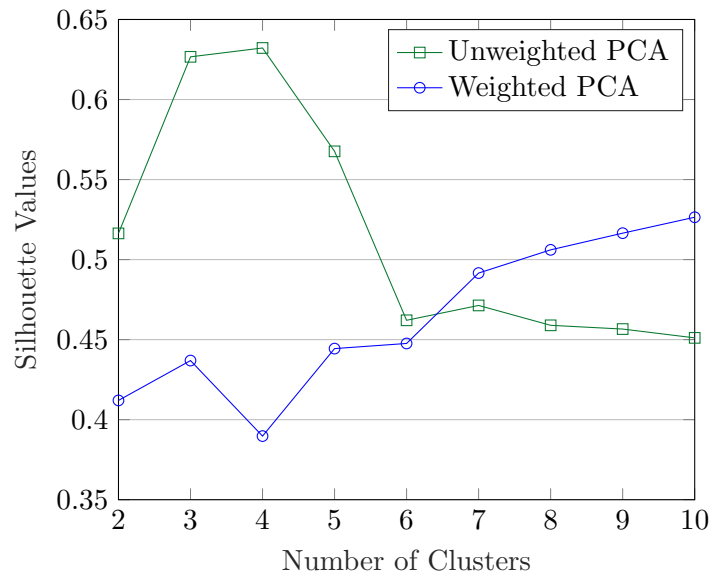


Figure 7.6: Silhouette plot for accommodation

7.5 Insights and Knowledge

The last aspect of the study is to determine what knowledge and insights can be gathered from the analysis applied on the TPD data.

The road to get to the insights and knowledge phase has been discussed in detail above and can be summarised in the schematic in Figure 7.7. Table 7.2 summarises the analysis that was performed on the data obtained from the TPD of the customers travelling on a trip.

The insights and knowledge gained from applying the ML techniques will be discussed based on the analysis done for

- Customer Behaviour,
- Customer Experience and
- Transactions.

During the analysis only the 7493 customers who completed at least one trip have been considered. The three analyses done will be discussed in the next sections.

7.5 Insights and Knowledge

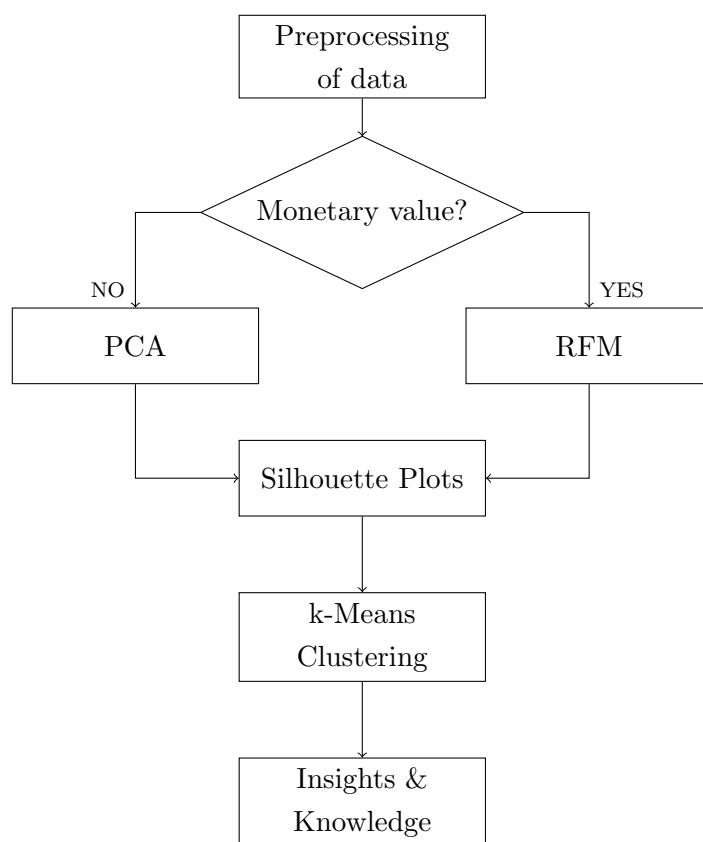


Figure 7.7: Schematic for analysis of Trip Planner Demonstrator data

Table 7.2: A summary of the analysis done

	PCA	RFM	# Clusters	k-Means
Customer Behaviour				
Accommodation	X	N/A	8	X
LDT	X	N/A	7	X
SDT	X	N/A	8	X
Customer Experience				
Accommodation	X	N/A	8	X
LDT	X	N/A	7	X
SDT	X	N/A	8	X
Transaction	X	N/A	8	X
Overall	N/A	X	4	X
Transactions	N/A	X	6	X

7.5.1 Analysis of customer behaviour

For the first analysis, clustering was applied on the customer behaviour trends for accommodation, LDT and SDT. Analysis was done on these entities to determine whether a unique group of customers uses a certain type of accommodation, LDT or SDT.

The first step was to perform the weighted PCA to reduce the dimensionality of the dataset. The results obtained from it can be seen in Figures 7.2, B.1 and B.2.

The second step was to find the ‘optimal’ number of clusters for each entity with the use of the silhouette plots, as shown in Figures 7.6, B.7 and B.8. For the accommodation and SDT the total number of clusters was chosen as eight and seven for LDT.

The third step was to perform k-Means clustering on these entities to determine the clusters. The graphical representation of the clusters obtained for each of the three entities can be seen in Figure 7.8 and the pie charts which demonstrate how the customers are distributed over the clusters can be seen in Figure 7.9.

After the clusters were obtained, they were analysed to see what knowledge and insights they contain. Tables B.1, B.2 and B.3 summarise the distribution of all customers over the clusters to see if there is a distinct group of customers that prefer a certain type of accommodation, LDT or SDT.

Each entity will be discussed by looking at two clusters and in the end a summary table will be provided for all the clusters per entity.

7.5 Insights and Knowledge

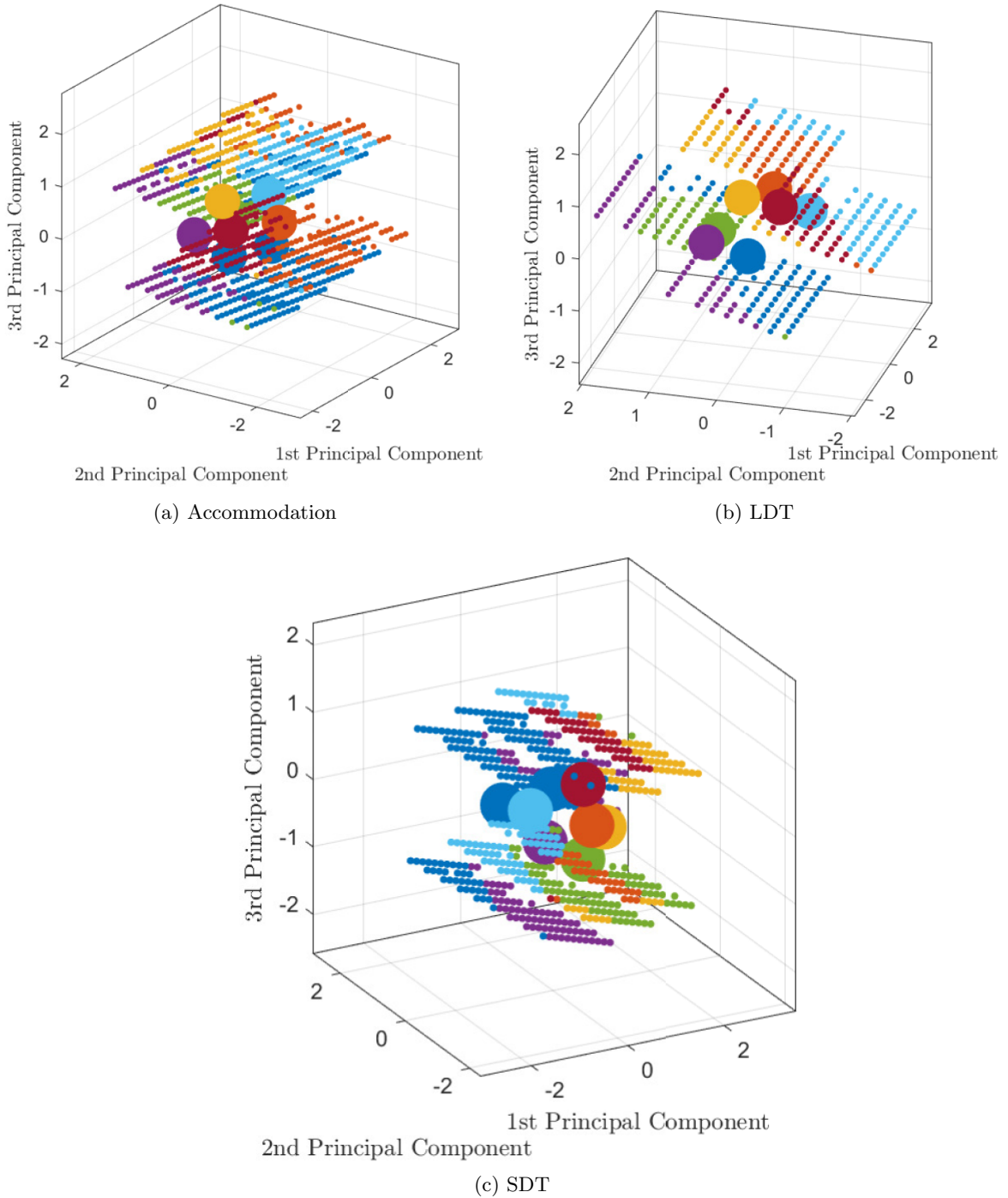


Figure 7.8: k-Means clustering plot of customer behaviour

7.5 Insights and Knowledge

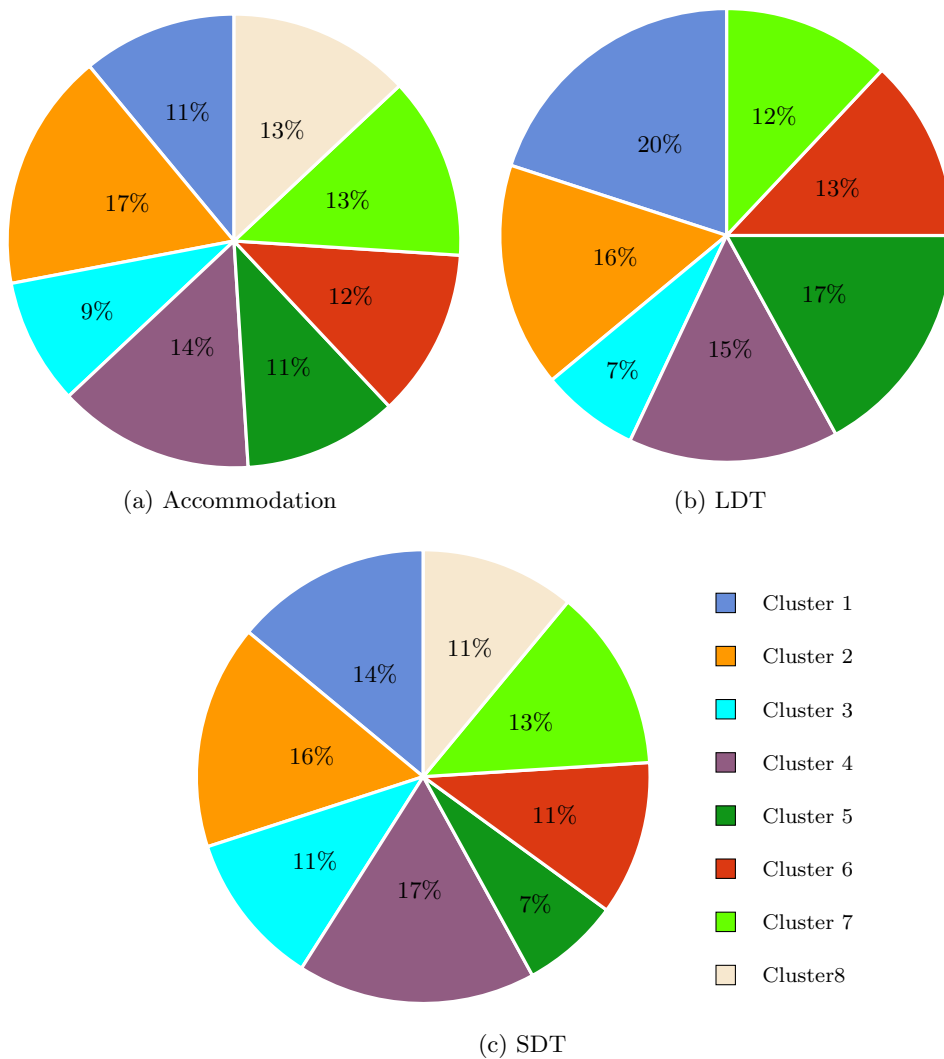


Figure 7.9: Pie charts for clustering of customer behaviour

7.5.1.1 Analysis of accommodation customer behaviour

The first entity, Accommodation, has been clustered using eight clusters. Figure 7.10 shows the histogram of the total customers per cluster. The biggest cluster is cluster 2 which has 1 270 customers and the smallest cluster is cluster 3 which has 648 customers. All the other clusters are distributed in-between. If one looks at the demographics of the clusters, it can be seen that the customers can be divided amongst their attributes per cluster.

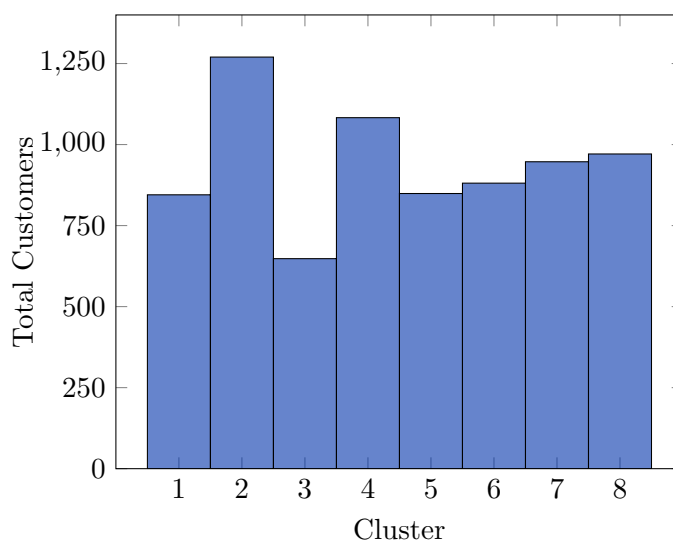


Figure 7.10: Histogram for clusters of customer behaviour – Accommodation

Two of the eight clusters will be highlighted and they have the following features:

1. *Cluster 2* consists of customers who prefer either bed & breakfasts, guesthouses or backpackers.

These customers:

- Are predominantly males,
- Reside all over South Africa except in the North West and the Western Cape and
- Are all older than 19.

2. *Cluster 6* consists of customers who prefer either Hotel or self-catering units. These customers:

- Are only females and
- Reside all over South Africa except in the Northern Cape, North West and the Western Cape and
- Are all younger than 45.

7.5.1.2 Analysis of LDT customer behaviour

The second entity, LDT, has been clustered using seven clusters. Figure 7.11 shows the histogram of the total customers per cluster. The biggest cluster is cluster 1 which has 1 447 customers and the smallest cluster is cluster 3 which has 507 customers. All the other clusters are distributed in-between.

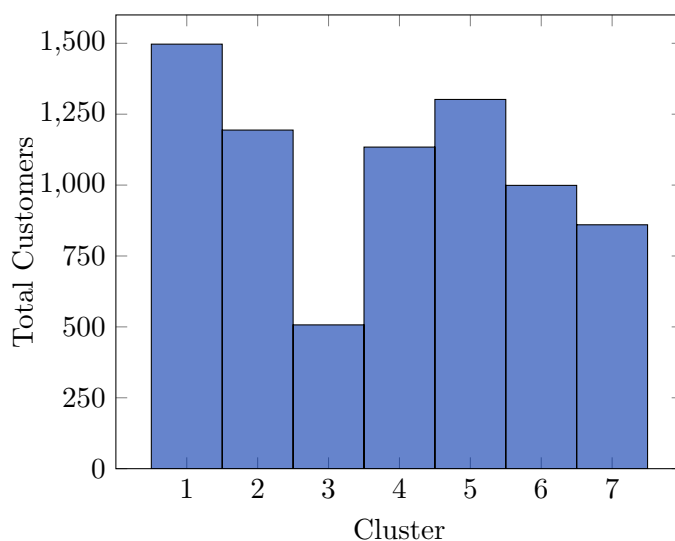


Figure 7.11: Histogram for clusters of customer behaviour – LDT

Two of the seven clusters will be highlighted and they have the following features:

1. *Cluster 1* consists of customers who prefer an aeroplane. These customers:
 - Are predominantly females,
 - Reside in the Eastern Cape, Free State, Gauteng and KwaZulu-Natal and
 - Are 30 years and older.

2. *Cluster 3* consists of customers who prefer a bus. These customers:
 - Are predominantly females,
 - Reside all over South Africa except in the Eastern Cape, Free State, Gauteng and KwaZulu-Natal and
 - Are all between 40 and 54 years old.

7.5.1.3 Analysis of SDT customer behaviour

The third entity, SDT, has been clustered using eight clusters. Figure 7.12 shows the histogram of the total customers per cluster. The biggest cluster is cluster 4 which has 1 220 customers and the smallest cluster is cluster 5 which has 527 customers. All the other clusters are distributed in-between. If one look at the demographics of the clusters, it can be seen that the customers can be divided amongst their attributes per cluster.

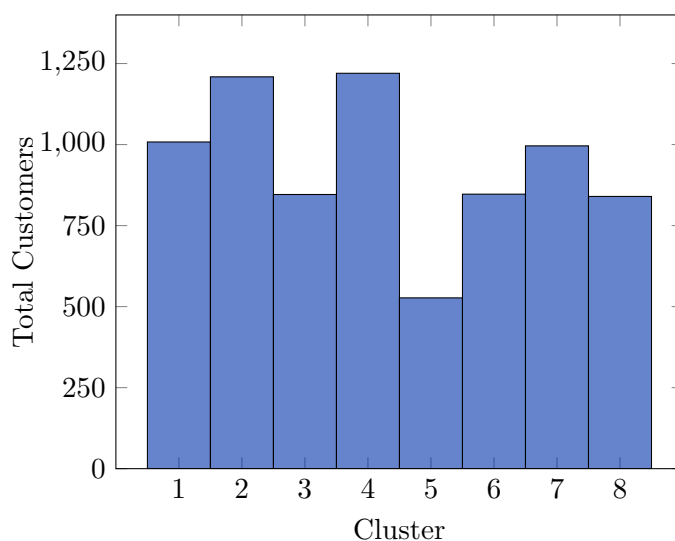


Figure 7.12: Histogram for clusters of customer behaviour – SDT

Two of the eight clusters will be highlighted and they have the following features:

1. *Cluster 1* consists of customers who prefer normal taxis and hailing app taxis. These customers:
 - Are only females,
 - Reside all over South Africa except in Northern Cape, North West and the Western Cape and
 - Are younger than 50 years.
2. *Cluster 4* consists of customers who prefer a car rental or hailing app taxi. These customers:
 - Are only males,
 - Reside all over South Africa except in the North West and the Western Cape and
 - Are older than 19 years.

Based on this analysis, it is clear that by using the data generated by the TPD through the trips that have been completed, insights and knowledge can be gained on what type of entity a customer uses more than other. Table 7.3 summarises the clusters for all the entities. The next analysis is to look at the CXs of touch points.

7.5 Insights and Knowledge

Table 7.3: Clusters for customer behaviour

Cluster	Gender	Age	Province	Type
Accommodation				
1	Only Males	Older than 29 years	Eastern Cape, Free State, Gauteng and KwaZulu-Natal	Hotels and Self-Catering Units
2	Predominantly Males	Older than 19 years	All provinces excluding North West and Western Cape	Backpackers, Bed & Breakfast and Guest Houses
3	Predominantly Females	Older than 39 years	Eastern Cape, Free State, Gauteng and KwaZulu-Natal	Hotels and Self-Catering Units
4	Only Females	Younger than 55 years	All provinces excluding Northern Cape, North West and Western Cape	Bed & Breakfast and Guest Houses
5	Predominantly Males	Younger than 65 years	All provinces excluding Eastern Cape and Gauteng	Hotels and Self-Catering Units
6	Only Females	Younger than 45 years	All provinces excluding Northern Cape, North West and Western Cape	Hotels and Self-Catering Units
7	Both genders	Younger than 60 years	All provinces excluding Eastern Cape and Gauteng	Bed & Breakfast and Guest Houses
8	Both genders	Younger than 55 years	All provinces excluding Eastern Cape and Gauteng	Guesthouses, Hotels and Self-Catering Units
LDT				
1	Predominantly Males	Older than 29 years	Eastern Cape, Free State, Gauteng and KwaZulu-Natal	Aeroplanes
2	Only Females	Between age of 20 and 59	All provinces excluding Northern Cape, North West and Western Cape	Buses
3	Predominantly Females	Between age of 45 and 59	Limpopo, Mpumalanga, Northern Cape, North West and Western Cape	Buses
4	Both Genders	Younger than 60 years	All provinces excluding Eastern Cape, Free State and North West	Aeroplanes
5	Only Females	Younger than 60 years	All provinces excluding Northern Cape, North West and Western Cape	Aeroplanes

Table 7.3 continues on next page

7.5 Insights and Knowledge

Cluster	Gender	Age	Province	Type
6	Predominantly Males	Between age of 25 and 49 and age of 55 and 75	All provinces excluding North West and Western Cape	Buses
7	Predominantly Males	Younger than 60 years	All provinces excluding Eastern Cape and Free State	Buses
SDT				
1	Only Females	Younger than 50 years	All provinces excluding Northern Cape, North West and Western Cape	Normal Taxis and Hailing App Taxis
2	Both genders	All ages	All provinces excluding Eastern Cape, Free State and Gauteng	Car Rental and Hailing App Taxis
3	Only Males	All ages	All provinces excluding Eastern Cape, Free State and Gauteng	Normal Taxis and Hailing App Taxis
4	Only Males	Older than 19 years	All provinces excluding North West and Western Cape	Car Rental and Hailing App Taxis
5	Predominantly Females	Older than 19 years	Eastern Cape, Free State, Gauteng and KwaZulu-Natal	Car Rental and Hailing App Taxis
6	Predominantly Males	Older than 39 years	All provinces excluding North West	Normal Taxis and Hailing App Taxis
7	Predominantly Females	Younger than 55 years	All provinces excluding North West and Western Cape	Car Rental and Hailing App Taxis
8	Both genders	All ages	All provinces excluding North West and Western Cape	Only Normal Taxis

End of Table 7.3

7.5.2 Analysis of Customer Experience

For the second analysis, clustering was applied on the CX ratings at touch points. Two types of analyses were done as to determine what group of customers rate a service and how the customers rate an entity. The first analysis used RFM and looked at the overall CX, whereas the second analysis used PCA and looked at each entity.

7.5.2.1 Analysis of overall Customer Experience

For the first analysis, RFM was used to transform the data and the following steps were performed.

The first was to perform RFM and the results as shown in Figure 7.3 was determined. The second step was to find the ‘optimal’ number of clusters by the use of the silhouette plot as can be seen in Figure B.13. After this step, the analysis was done for four and eight clusters. In the end, four clusters were chosen as it gave the ‘best’ number of clusters.

The graphical representation of the clusters obtained for each of the three entities can be seen in Figure 7.14 and the pie charts which demonstrate how the customers are distributed over the clusters can be seen in Figure 7.13. It is clear that cluster four is the biggest.

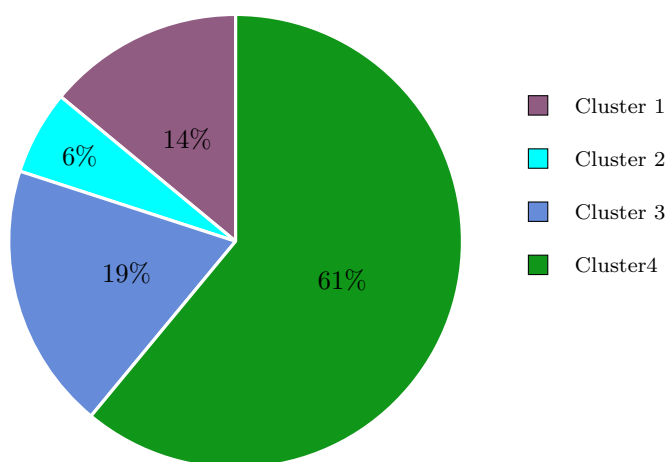


Figure 7.13: Pie chart of clusters after Recency, Frequency and Monetary analysis

The RFM values for all four clusters can be seen in Figure 7.15. The following can be concluded per customer:

1. *Cluster 1* represents customers who have travelled recently, but do not rate the touch points frequently. However, when they do rate the touch points, it will be a bad rating of either one or two.
2. *Cluster 2* is the smallest cluster and it represents customers who either travelled recently or not, and do not rate frequently. But when they do rate it can be any value with the majority of it being three.
3. *Cluster 3* represents customers who travelled recently and when they travel they give frequent ratings, where the ratings will usually be three.
4. *Cluster 4* is the biggest cluster and it represents customers who have travelled recently, but they do not rate the touch points frequently and if they do rate the touch points it will usually be a rating of three, otherwise a rating of two is given.

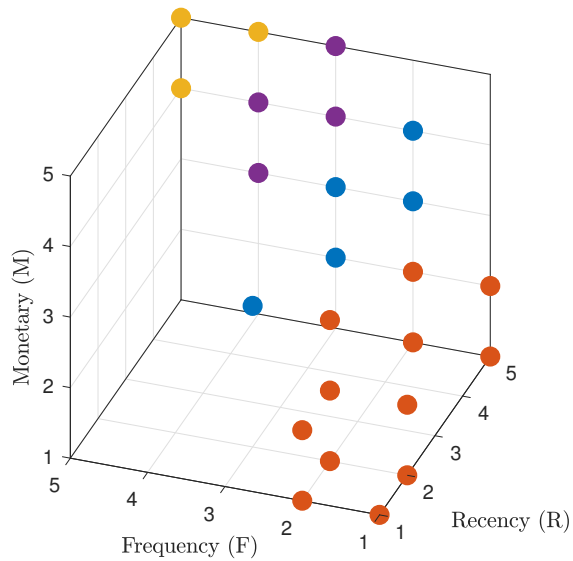


Figure 7.14: Clustered Recency, Frequency and Monetary analysis of overall Customer Experience

After the RFM clustering was done, the clustering was applied to the attributes of the customers to see whether the clusters can be defined according to the customer attributes. The distribution of customers per cluster can be seen in Table 7.4.

When one tries to determine the customer's attributes per cluster, it can be seen that there is no distinct difference between the clusters. For example, when looking at the gender for all clusters the ration between the males versus females are all similar.

Therefore, in order to determine the type of customers that rate the touch points PCA can be used. The second analysis is required for the CX ratings of accommodation, LDT, SDT and transactions.

7.5 Insights and Knowledge

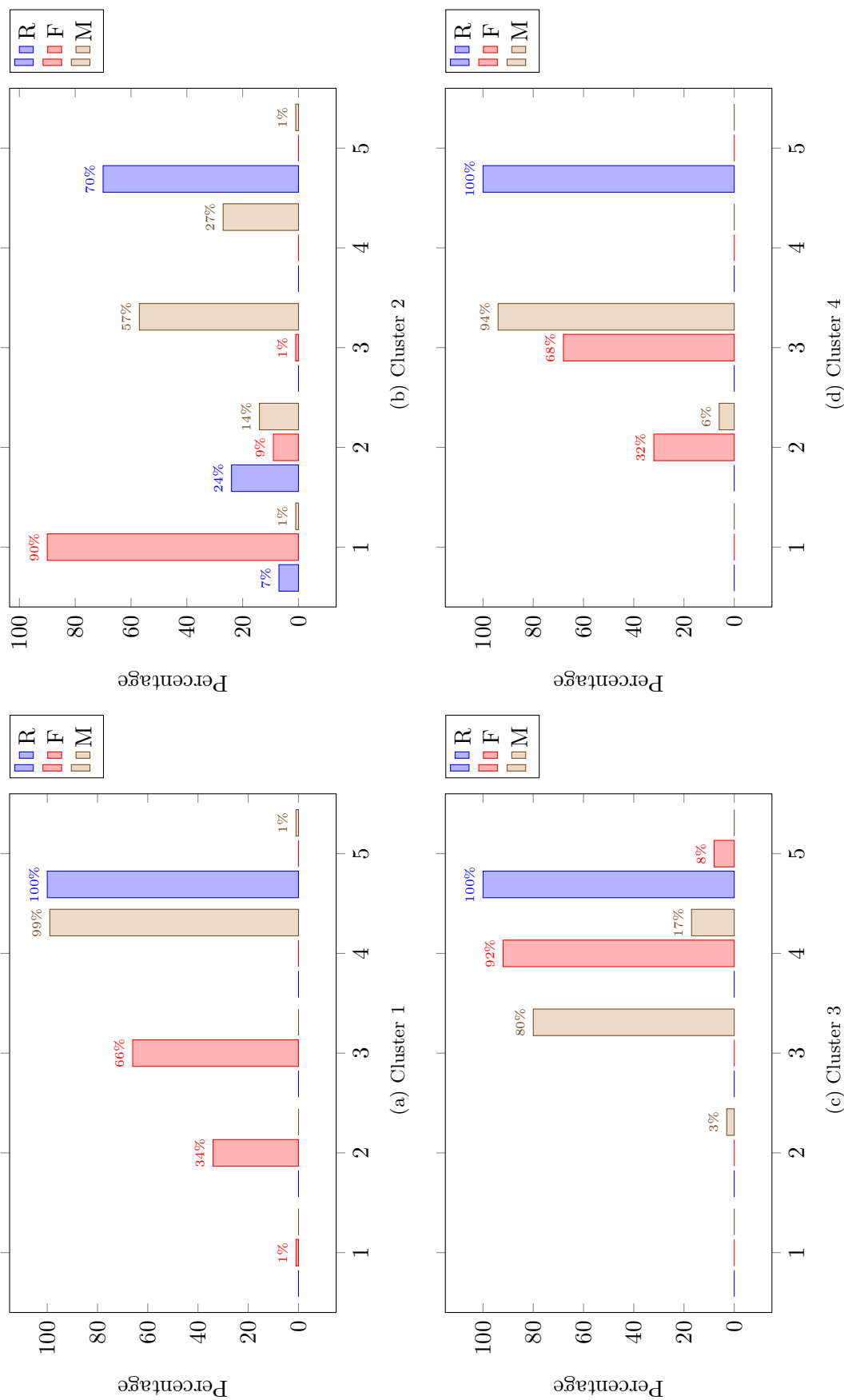


Figure 7.15: Recency, Frequency and Monetary values for overall Customer Experience clusters

7.5 Insights and Knowledge

Table 7.4: Recency, Frequency and Monetary clusters detail for overall Customer Experience

Cluster		1	2	3	4
Gender	Male	6.89%	3.22%	8.61%	30.29%
	Female	7.39%	3.1%	9.4%	31.11%
Age Category	18 – 19 years	0.91%	0.45%	1.31%	3.82%
	20 – 24 years	2.07%	1.15%	2.82%	9.97%
	25 – 29 years	2.08%	0.72%	2.58%	8.77%
	30 – 34 years	1.84%	0.93%	2.55%	8.5%
	35 – 39 years	1.47%	0.63%	1.77%	5.75%
	40 – 44 years	1.41%	0.65%	1.48%	6.27%
	45 – 49 years	1.12%	0.47%	1.24%	4.35%
	50 – 54 years	1.15%	0.52%	1.55%	4.64%
	55 – 59 years	0.77%	0.31%	0.87%	3.1%
	60 – 64 years	0.61%	0.19%	0.92%	2.76%
	65 – 69 years	0.4%	0.09%	0.33%	1.55%
70+ years	0.44%	0.2%	0.59%	1.92%	
Province	Eastern Cape	1.2%	0.69%	0.15%	5.1%
	Free State	0.56%	0.44%	0.08%	2.32%
	Gauteng	4.32%	0.21%	8.29%	19.7%
	KwaZulu-Natal	3.23%	0.85%	5.07%	14.09%
	Limpopo	0.25%	0.41%	0%	0.83%
	Mpumalanga	0.4%	0.59%	0.05%	2.18%
	Northern Cape	0.87%	1.19%	0.04%	2.94%
	North West	0.19%	0.33%	0%	0.53%
	Western Cape	3.26%	1.59%	4.32%	13.72%

7.5.2.2 Analysis of individual Customer Experience

For the second analysis PCA was used to transform the data and the following steps were performed.

The first step was to perform the weighted PCA on it to reduce the dimensionality and the results obtained from it can be seen in Figures B.3, B.4, B.5 and B.6.

The second step was to find the ‘optimal’ numbers of clusters for each entity by the use of the silhouette plots, as can shown in Figures 7.5, B.9, B.10 and B.12. For all the entities eight clusters were chosen.

The third step was to perform k-Means clustering on these entities to determine what the clusters will consist of. The clusters obtained for the three entities can be seen in Figure 7.16 and the pie charts which demonstrate the distribution of customers over the clusters can be seen in Figure 7.17.

7.5 Insights and Knowledge

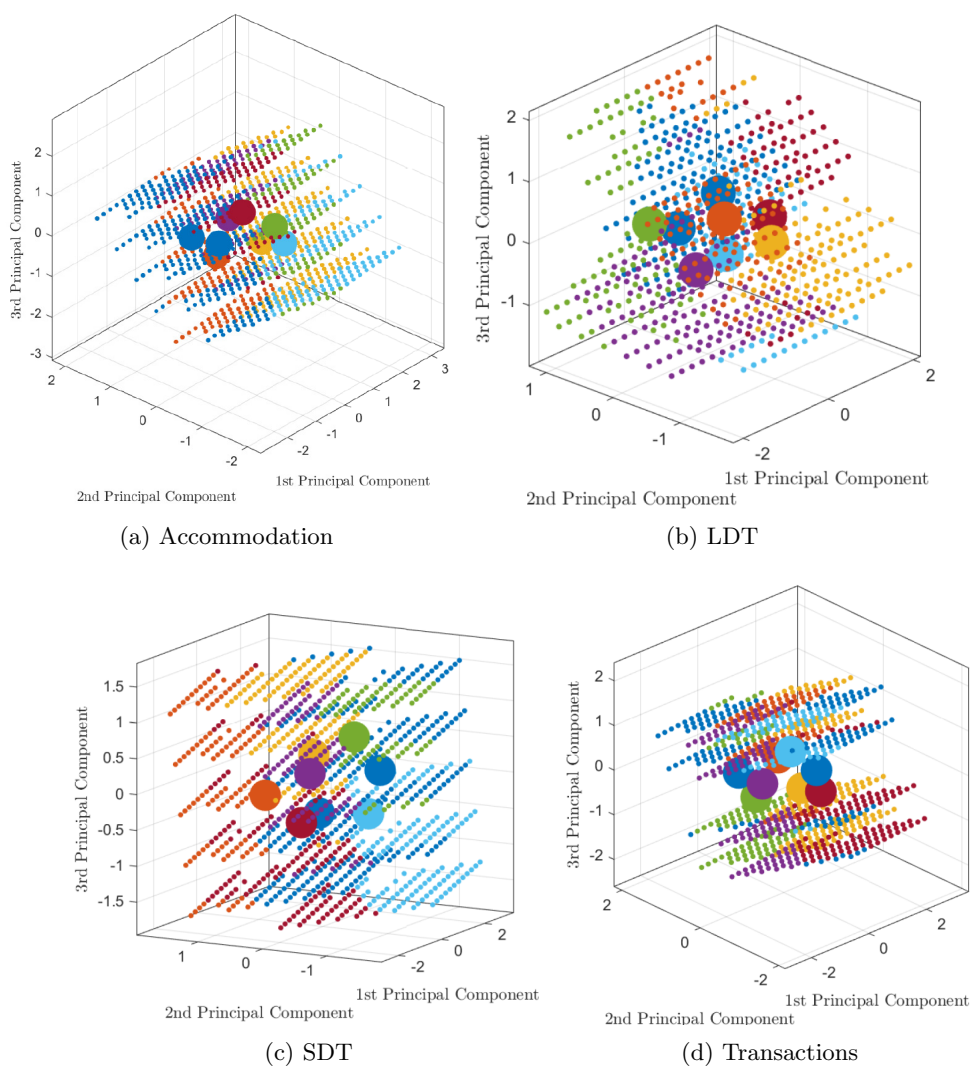


Figure 7.16: k-Means clustering plot of Customer Experience

After the clusters were obtained, the clusters have been analysed to see what knowledge and insights they contain. Tables B.4, B.5, B.6 and B.7 summarise the distribution of all customers over the clusters to see if there is a distinct group of customers that prefer a certain type of accommodation, LDT or SDT.

The CX of each entity will be discussed by looking at two clusters and in the end a summary table will be provided for all the clusters per entity.

7.5 Insights and Knowledge

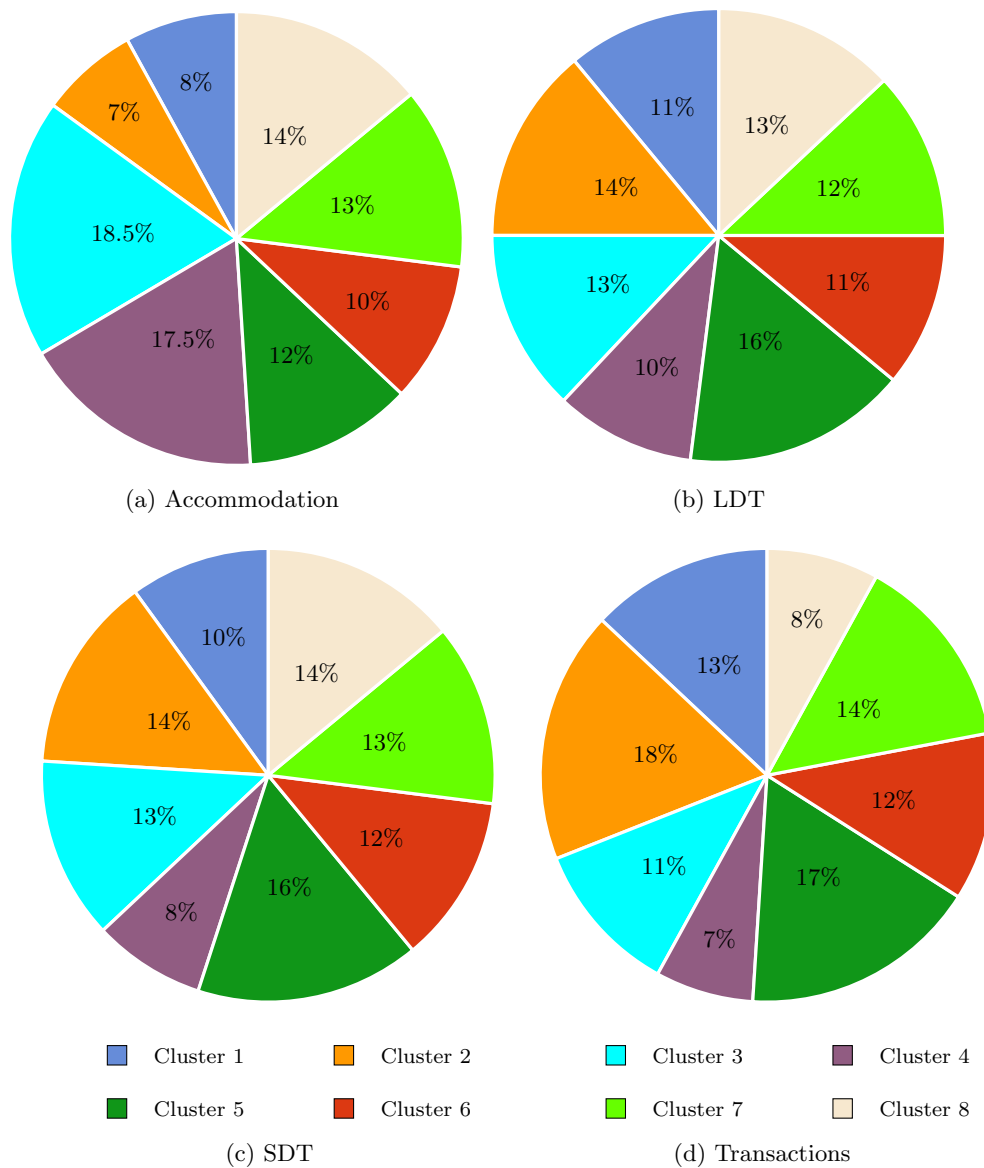


Figure 7.17: Pie charts for clustering of Customer Experience

Accommodation:

The first entity, CX for accommodation, has been clustered using eight clusters. Figure 7.18 shows the histogram of the total customers per cluster. The biggest cluster is cluster 3 which has 1 383 customers and second is cluster 4 which has 1 321 customers. The smallest cluster is cluster 2 which has 555 customers and second smallest is cluster 1 with 561 customers. All the other clusters are distributed in-between.

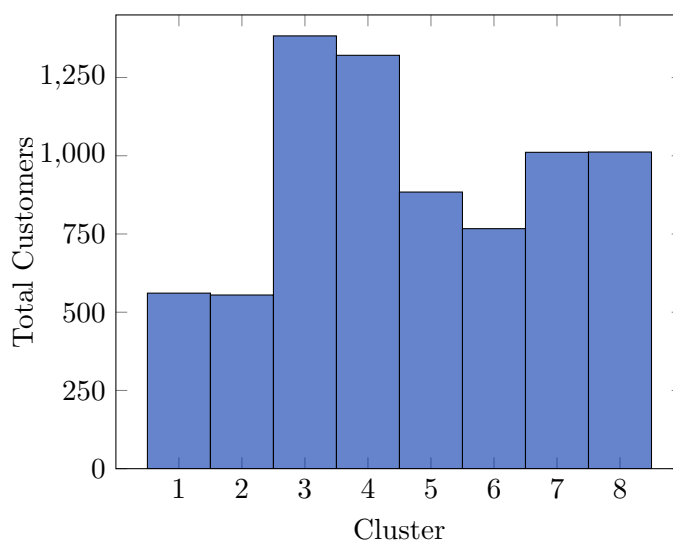


Figure 7.18: Histogram for clusters of Customer Experience – Accommodation

Two of the eight clusters will be highlighted and they have the following features:

1. *Cluster 4* consists of customers who give any rating except the lowest rating. These customers:
 - Are only females,
 - Reside all over South Africa except in Mpumalanga, Northern Cape, North West and the Western Cape and
 - Are all younger than 60 years.
2. *Cluster 8* consists of customers who give all ratings. These customers:
 - Are only males,
 - Reside in Mpumalanga, Northern Cape, North West and the Western Cape and
 - Are of all ages.

LDT:

The second entity, CX for LDT, has been clustered using eight clusters. Figure 7.19 shows the histogram of the total customers per cluster. The biggest cluster is cluster 5 which has 1213 customers and the smallest cluster is cluster 4 which has 754 customers. All the other clusters are distributed in-between. If one look at the demographics of the clusters, it can be seen that the customers can be divided amongst their attributes per cluster.

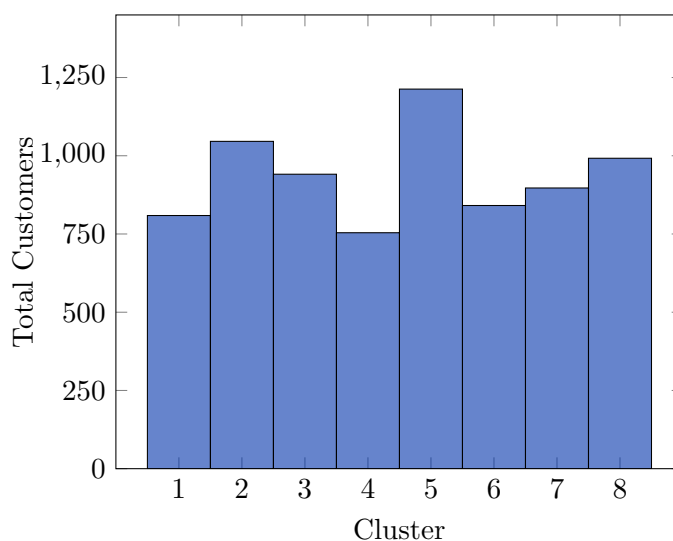


Figure 7.19: Histogram for clusters of Customer Experience – LDT

Two of the eight clusters will be highlighted and they have the following features:

1. *Cluster 3* consists of customers who give a rating of average and higher. These customers:

- Are only males,
- Reside all over South Africa except in the North West and the Western Cape and
- Are all younger than 55 years.

2. *Cluster 7* consists of customers who give all ratings. These customers:

- Are predominantly females,
- Reside all over South Africa except in the Eastern Cape, Free State and Gauteng and
- Are of all ages.

SDT:

The third entity, CX for SDT, has been clustered by using eight clusters. Figure 7.20 shows the histogram of the total customers per cluster. The biggest cluster is cluster 4 which has 1 220 customers and the smallest cluster is cluster 5 which has 527 customers. All the other clusters are distributed in-between. If one look at the demographics of the clusters, it can be seen that the customers can be divided amongst their attributes per cluster.

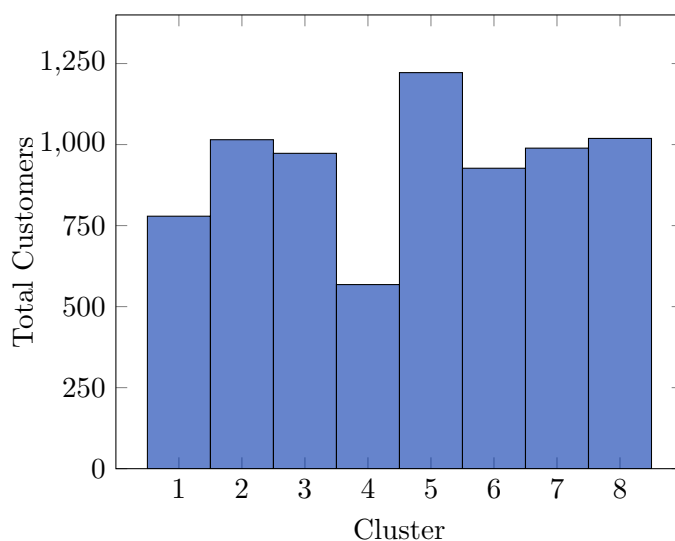


Figure 7.20: Histogram for clusters of Customer Experience – SDT

Two of the eight clusters will be highlighted and they have the following features:

1. *Cluster 3* consists of customers who give a rating of average and higher. These customers:
 - Are only males,
 - Reside all over South Africa except in the North West and the Western Cape and
 - Are younger than 55 years.
2. *Cluster 5* consists of customers give a rating of average and lower. These customers:
 - Are only females,
 - Reside all over South Africa except in the North West and the Western Cape and
 - Are of all ages.

Transactions:

The fourth entity, CX for transactions, has been clustered using eight clusters. Figure 7.21 shows the histogram of the total customers per cluster. The biggest cluster is cluster 2 which has 1 363 customers. The smallest cluster is cluster 4 which has 563 customers and second is cluster 8 which has 573 customers. All the other clusters are distributed in-between. If one look at the demographics of the clusters, it can be seen that the customers can be divided amongst their attributes per cluster.

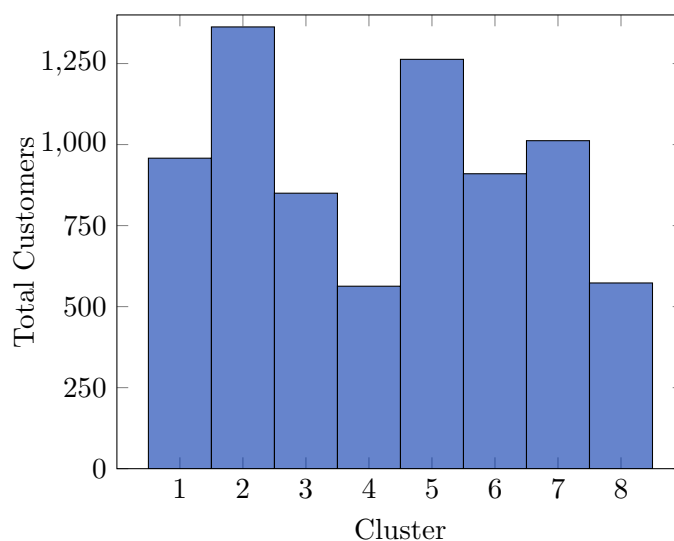


Figure 7.21: Histogram for clusters of Customer Experience – Transactions

Two of the eight clusters will be highlighted and they have the following features:

1. *Cluster 1* consists of customers who give rating of above average or excellent. These customers:

- Are only females,
- Reside all over South Africa except in Mpumalanga, Northern Cape, North West and the Western Cape and
- Are younger than 50 years.

2. *Cluster 8* consists of customers who give below average or lowest. These customers:

- Are only females,
- Reside in Mpumalanga, Northern Cape, North West and the Western Cape and
- Are all older than 35 years.

Based on both the analyses done, it is clear that by using the trip data generated by the TPD, insights and knowledge can be gained on what type of entity a customer rates and what their rating behaviour is. Table 7.5 summarises the clusters for all the entities. The next analysis is to look at the transactions.

7.5 Insights and Knowledge

Table 7.5: Clusters for Customer Experience

Cluster	Gender	Age	Province	CX Rating
Accommodation				
1	Only Females	Younger than 60 years	Mpumalanga, Northern Cape, North West and Western Cape	All CX ratings excluding lowest
2	Only Females	All ages	Mpumalanga, Northern Cape, North West and Western Cape	All CX ratings
3	Only Females	All ages	Eastern Cape, Free State, Gauteng, KwaZulu-Natal and Limpopo	All CX ratings
4	Only Females	Younger than 60 years	Eastern Cape, Free State, Gauteng, KwaZulu-Natal and Limpopo	All CX ratings excluding lowest
5	Only Males	All ages	Eastern Cape, Free State, Gauteng and KwaZulu-Natal	All CX ratings
6	Only Males	Older than 24 years	All provinces excluding North West and Western Cape	All CX ratings
7	Only Males	Younger than 50 years	All provinces excluding North West and Western Cape	All CX ratings excluding lowest
8	Only Males	All ages	Mpumalanga, Northern Cape, North West and Western Cape	All CX ratings
LDT				
1	Only Females	All ages	Eastern Cape, Free State, Gauteng and KwaZulu-Natal	All CX ratings excluding lowest
2	Predominantly Males	All ages	Mpumalanga, Northern Cape, North West and Western Cape	All CX Ratings
3	Only Males	Younger than 55 years	All provinces excluding North West and Western Cape	Average, Good (above average) and Excellent CX ratings
4	Predominantly Males	Older than 19 years	Eastern Cape, Free State, Gauteng and KwaZulu-Natal	All CX ratings excluding lowest
5	Predominantly Females	All ages	All provinces excluding North West and Western Cape	Lowest, Bad (below average) and Average CX Ratings
6	Predominantly Females	All ages	All provinces excluding North West	Lowest, Bad (below average) and Average CX Ratings

Table 7.5 continues on next page

7.5 Insights and Knowledge

Cluster	Gender	Age	Province	CX Rating
7	Predominantly Females	All ages	KwaZulu-Natal, Limpopo, Mpumalanga, Northern Cape, North West and Western Cape	All CX Ratings
8	Only Females	Younger than 55 years	All provinces excluding North West and Western Cape	Average, Good (above average) and Excellent CX ratings
SDT				
1	Only Males	All ages	Eastern Cape, Free State, Gauteng and KwaZulu-Natal	All CX Ratings excluding lowest
2	Only Females	All ages	Mpumalanga, Northern Cape, North West and Western Cape	All CX Ratings
3	Only Males	Younger than 55 years	All provinces excluding North West and Western Cape	Average, Good (above average) and Excellent CX ratings
4	Only Females	All ages	Eastern Cape, Free State, Gauteng and KwaZulu-Natal	All CX Ratings excluding lowest
5	Only Females	All ages	All provinces excluding North West and Western Cape	Lowest, Bad (below average) and Average CX Ratings
6	Only Males	All ages	All provinces excluding North West	Lowest, Bad (below average) and Average CX Ratings
7	Predominantly Males	All ages	All provinces excluding Eastern Cape, Free State and Gauteng	All CX Ratings
8	Only Females	Younger than 60 years	All provinces excluding North West and Western Cape	Average, Good (above average) and Excellent CX ratings
Transactions				
1	Only Females	Younger than 55 years	Eastern Cape, Free State, Gauteng, KwaZulu-Natal and Limpopo	Good (above average) and Excellent CX ratings
2	Only Males	Younger than 60 years	All provinces excluding North West and Western Cape	Average, Good (above average) and Excellent CX ratings
3	Only Females	Younger than 65 years	Eastern Cape, Free State, Gauteng, KwaZulu-Natal and Limpopo	Bad (below average), Average and Good (above average) CX ratings

Table 7.5 continues on next page

7.5 Insights and Knowledge

Cluster	Gender	Age	Province	CX Rating
4	Only Females	Younger than 60 year	Mpumalanga, Northern Cape, North West and Western Cape	Average, Good (above average) and Excellent CX ratings
5	Only Males	Between age of 25 and 65	All provinces excluding North West and Western Cape	All CX Ratings
6	Only Females	35 years and older	Eastern Cape, Free State, Gauteng, KwaZulu-Natal and Limpopo	All CX Ratings
7	Only Males	Younger than 60 years	Mpumalanga, Northern Cape, North West and Western Cape	All CX Ratings
8	Only Females	35 years and older	Mpumalanga, Northern Cape, North West and Western Cape	Low and Bad (below average) CX Ratings

End of Table 7.5

7.5.3 Analysis of transactions

For the third analysis, clustering was applied on the customer transactions data. Analysis was done for these entities to determine valuable insights and knowledge from the customer transactions.

The first step was to perform RFM analysis to find the buying behaviour of customers and the results are shown in Figure 7.3. The second step was to find the ‘optimal’ number of clusters by the use of the silhouette plot as can be seen in Figure B.11. After this step the analysis was done for eight clusters.

The graphical representation of the clusters obtained for each of the three entities can be seen in Figure 7.23 and the pie charts which demonstrate how the customers are distributed over the clusters can be seen in Figure 7.22. It is clear that cluster four is the biggest.

The RFM values for all eight clusters can be seen in Table 7.6. After the RFM clustering was done, the clustering was applied to the attributes of the customers to see whether the clusters can be defined according to the customer attributes.

Two clusters will be discussed to explain the clustering of transactions. The first is cluster 1. The RFM values for this cluster can be seen in Figure 7.24. Based on the RFM values it can be said that customers in cluster 1 buy recently when they travel, but they do not buy frequently. When they do buy products at the shops they spend an average amount of money on them. The customer demographics of this cluster shown in Table 7.7 can be used.

The second cluster is cluster 3. The RFM values for this cluster can be seen in Figure 7.25. Based on the RFM values it can be said that customers in cluster 3 buy recently when they travel, but they buy less frequently than the customers of cluster 1. When they do buy products at the

7.5 Insights and Knowledge

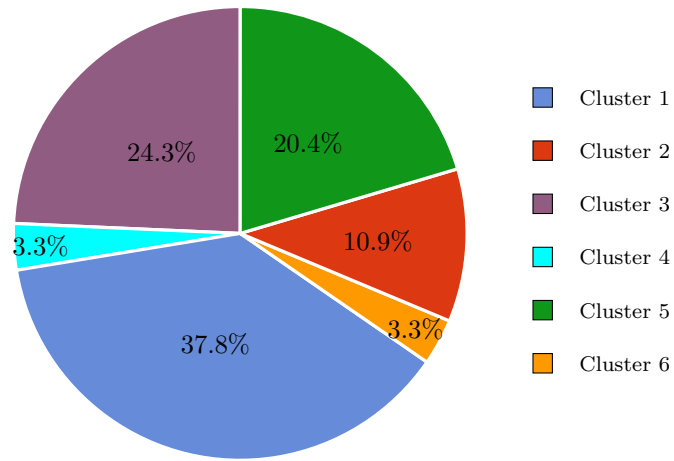


Figure 7.22: Pie chart of transaction clusters after Recency, Frequency and Monetary analysis

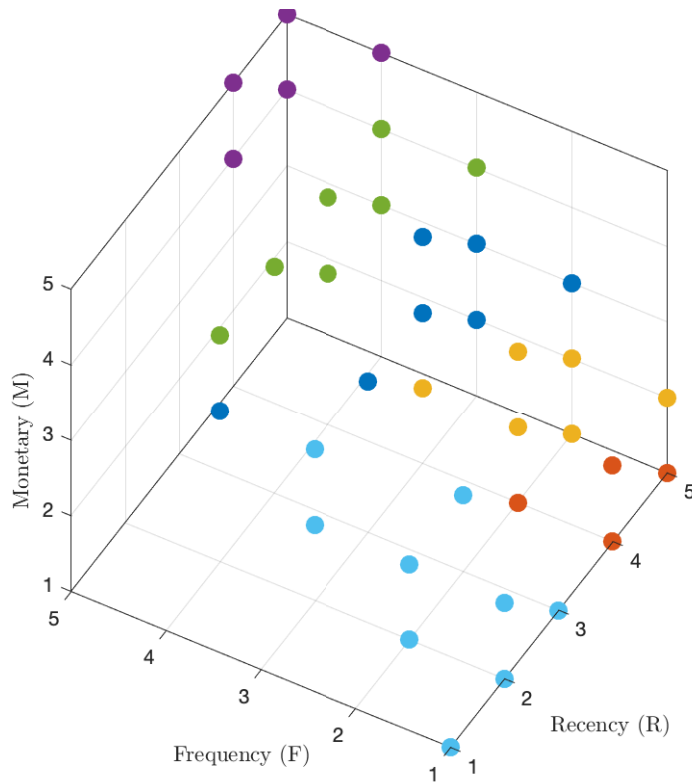


Figure 7.23: Clustered Recency, Frequency and Monetary Analysis of customer transactions

7.5 Insights and Knowledge

Table 7.6: Recency, Frequency and Monetary values for transactions

Cluster		1	2	3	4	5	6
Recency	1	0%	0%	0%	0%	0%	15%
	2	0%	0%	0%	0%	0%	49%
	3	0%	0%	0%	0%	0%	1%
	4	7%	3%	15%	2%	2%	35%
	5	93%	97%	85%	98%	98%	0%
Frequency	1	0%	0%	10%	0%	0%	78%
	2	2%	0%	68%	0%	0%	17%
	3	98%	0%	22%	0%	5%	5%
	4	0%	100%	0%	0%	95%	0%
	5	0%	0%	0%	100%	0%	0%
Monetary	1	0%	0%	12%	0%	0%	80%
	2	0%	0%	88%	0%	0%	16%
	3	100%	100%	0%	0%	0%	4%
	4	0%	0%	0%	56%	99%	0%
	5	0%	0%	0%	44%	1%	0%

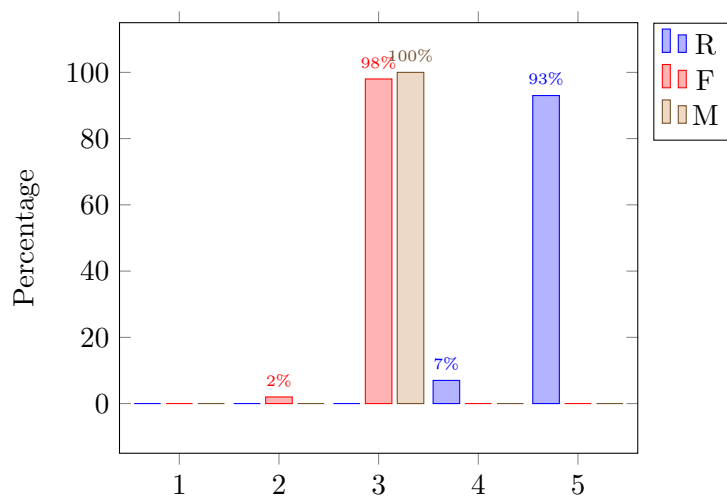


Figure 7.24: Recency, Frequency and Monetary plot for transactions of cluster 1

7.5 Insights and Knowledge

Table 7.7: Customer’s demographics of cluster 1

Gender:		Age:						
Female	Male	15 – 24	25 – 34	35 – 44	45 – 54	55 – 64	65+	
52%	48%	22%	28%	19%	15%	10%	6%	
Province								
1	2	3	4	5	6	7	8	9
7.7%	3.1%	36.9%	24.6%	0.5%	2.5%	2.7%	0.3%	21.8%

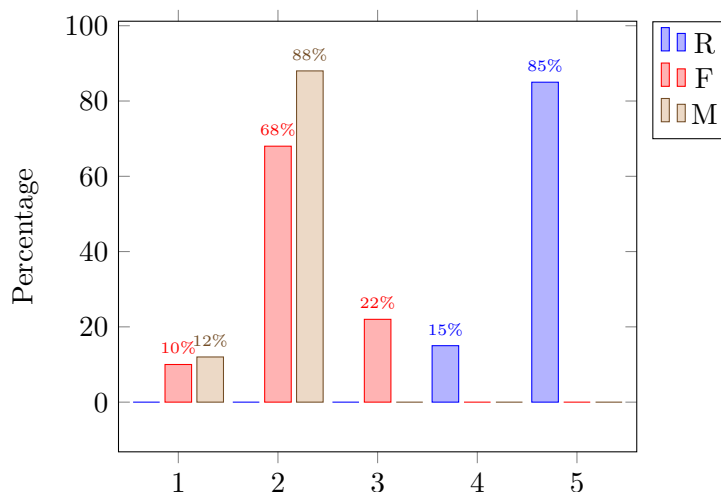


Figure 7.25: Recency, Frequency and Monetary plot for transactions of cluster 3

shops they also do not spend a lot of money on it. The customer demographics of this cluster are shown in Table 7.8.

By comparing the customer demographics of both clusters 1 and 3, it is clear that there is no distinct difference between the two and also that all attributes are distributed, evenly over the two clusters. The distribution of customers per cluster can be seen in Table 7.9.

When one tries to determine the customers’ attributes per cluster, it can be seen that there is no distinct difference between the clusters. For example, when looking at the gender for all clusters the ratio between the males versus females are all similar. Therefore, there is no distinct difference

Table 7.8: Customer’s demographics of cluster 3

Gender:		Age:						
Female	Male	15 – 24	25 – 34	35 – 44	45 – 54	55 – 64	65+	
52%	48%	23%	27%	21%	16%	9%	4%	
Province								
1	2	3	4	5	6	7	8	9
14.1%	7%	12.1%	15.6%	5%	7.8%	13%	3%	22.4%

7.6 Conclusion: Analysis of trips

Table 7.9: Recency, Frequency and Monetary clusters detail for transactions of customer

Cluster		1	2	3	4	5	6
Gender	Male	48%	52%	48%	48%	50%	52%
	Female	52%	48%	52%	52%	50%	48%
Age Category	18 – 19 years	6%	7%	6%	7%	7%	6%
	20 – 24 years	16%	15%	17%	15%	15%	20%
	25 – 29 years	13%	15%	14%	18%	15%	12%
	30 – 34 years	15%	12%	13%	16%	13%	16%
	35 – 39 years	9%	10%	10%	10%	9%	9%
	40 – 44 years	10%	11%	10%	6%	9%	12%
	45 – 49 years	7%	7%	8%	6%	6%	6%
	50 – 54 years	8%	6%	8%	8%	8%	8%
	55 – 59 years	6%	5%	5%	4%	5%	4%
	60 – 64 years	4%	4%	4%	7%	5%	3%
	65 – 69 years	3%	3%	2%	1%	2%	1%
	70+ years	3%	3%	3%	3%	3%	3%
Province	Eastern Cape	8%	2%	14%	0%	1%	13%
	Free State	3%	1%	7%	0%	1%	8%
	Gauteng	37%	45%	12%	46%	45%	4%
	KwaZulu-Natal	25%	26%	16%	25%	29%	16%
	Limpopo	0%	0%	5%	0%	0%	4%
	Mpumalanga	2%	1%	8%	0%	0%	6%
	Northern Cape	3%	1%	13%	0%	0%	20%
	North West	0%	0%	3%	0%	0%	6%
	Western Cape	22%	24%	22%	28%	24%	24%

between the attributes of the customers.

Therefore, the RFM only indicated the buying behaviour of customers when they travel. Since the focus of the study is on management and improvement of CX this will not be analysed further.

7.6 Conclusion: Analysis of trips

In this chapter data analytics was applied to the trip data generated by the TPD. This step was required to show the knowledge and insights contained in this system. It is important to note that this insight and knowledge are not used in the TPD, however, the business partners that use this system can use it to improve their marketing campaigns (Mitik *et al.*, 2017). The analysis was therefore done to show the capabilities of the TPD and the value it has for business partners.

First, the roadmap of analysis that was used for this study was discussed. The roadmap captures

7.6 Conclusion: Analysis of trips

the steps that were followed to generate knowledge and insights from the TPR.

Secondly, the first step of the roadmap was discussed to show how a ML tool and technique was chosen. Clustering, which is unsupervised learning, was selected.

Thirdly, the second step of the roadmap was discussed to show how the datasets were preprocessed and how RFM and PCA were used to transform the data.

Fourthly, the third step of the roadmap was discussed in which k-Means was chosen as a clustering technique and it was determined that the silhouette value will be used to determine the total number of clusters to be used for analysis.

Lastly, the fourth step of the roadmap was discussed which contains the analyses of the data sets. All the clusters were highlighted and these clusters can be used by business partners to improve marketing strategies and to understand the customer behaviour in the travel domain.

The next chapter will conclude this study.

Chapter 8

Conclusion and recommendations

In the preceding chapters, the research study has unfolded from a proposal to a capability demonstrator. Therefore, this chapter will serve as a conclusion to this thesis in which the aim was to conduct a study in which research and the application of engineering skills, tools and methods were used to construct a capability demonstrator.

This chapter will first provide a summary of the work completed. Secondly, a self-assessment of the work will be presented. Lastly, recommendations for future work will be provided.

8.1 Research summary

In Chapter 1 the research proposal was provided. In the research proposal the case study was introduced of how a fictitious customer named Thandi was travelling to Durban and back without having to plan, book or manage the trip herself. The case study was used to conceptualise the purpose of the study which was to develop a demonstrator that uses data analytics in a partnering venture to manage and increase a customer's experiences in the travelling domain. The research objectives were stated along with the methodology that was used to achieve these objectives within the research scope.

Before the author looked at the development of the demonstrator, a thorough literature study had to be conducted across three domains. Firstly, in Chapter 2 the concept of Customer Experience was investigated together with the management approaches available for it and aspects that can influence a customer's single and overall experiences. It was important here that an engineer should be able to understand how human aspects have to be incorporated into systems.

Secondly, in Chapter 3 the concept of Big Data was investigated to understand what it is and what analytical tools are available for it. A broad literature study had to be conducted to ensure that the author was able to fully grasp the field of Big Data Analytics.

Lastly, in Chapter 4 the concept of *Business Partnering on a Cross-functional Platform* was investigated to understand how the platform would enable the functionality of the demonstrator. Even though the study excludes the practical part of the platform, it is important to investigate this concept since the business partners form the cornerstone for the functionality of the system.

After the literature review was conducted, Objective 1 was achieved. In Chapter 5 the focus moved to the development of the demonstrator. A demonstrator was developed to show how to manage and improve a customer experience in the travel domain by the use of data analytics in a partnering venture. The demonstrator was implemented as a digital support and information system known as the trip planner. Therefore, the resulting system is called the *Trip Planner Demonstrator*

8.1 Research summary

in which unique trips for many customers were simulated by the use of customer preferences and historical behaviour. The Trip Planner Demonstrator consisted of four components namely, (i) business partners, (ii) a database, (iii) a simulator and (iv) a data analytics function. An architecture for this system was designed by the use of the Object-Process Methodology and this architecture was used as reference for the development of the Trip Planner Demonstrator.

For the business partners, it was assumed that all aspects were already set in place, that all legal issues had been taken care of and that data regulations such as the *GDPR* and the *POPI* Act had been adhered to.

For the database, MS SQL Server was used as the database server and all input data required for the Trip Planner Demonstrator was simulated to avoid ethical clearance issues. The data entities in the database have been created as set out by the Extended Entity-Relationship Diagram and the input data entities were simulated in the Matlab environment.

For the simulator, Matlab was used to implement the actions performed by the Trip Planner Demonstrator. The simulator is a fundamental component of the Trip Planner Demonstrator since it enables all processes required to plan, book and manage a customer's trip.

The data analytics function of the Trip Planner Demonstrator had two purposes. The first purpose was that data analytics were required to enable the planning process for a customer's trip based on their preferences and historical behaviour. The second purpose was its use in the analysis of the Trip Planner Demonstrator data.

After the Trip Planner Demonstrator was developed, Objectives 2 to 4 were achieved. In Chapter 6 the focus moved to the verification and evaluation of the Trip Planner Demonstrator.

The verification was a continuous process which was performed during all phases of the development of the Trip Planner Demonstrator. The main actions performed were debugging and a step-by-step verification for various types of customer who could use the Trip Planner Demonstrator. After the verification process, no errors have been detected for the Trip Planner Demonstrator in the given dataset.

The evaluation of the Trip Planner Demonstrator occurred after (i) the trip planning process, (ii) customer travelling process and (iii) the runs completed by the Trip Planner Demonstrator. The evaluation was performed to ensure the Trip Planner Demonstrator had been built correctly. Comparisons were made by comparing the expected results versus the actual results. After the evaluation process, the author can say with confidence that the Trip Planner Demonstrator was able to effectively manage and improve a customer experience when they go on a trip.

After the Trip Planner Demonstrator has been confirmed as 'the right model has been built right', the analysis of trips completed by the Trip Planner Demonstrator commenced. In Chapter 7 the focus moved to the analysis of the unique trips generated for many customers.

The roadmap followed for the analysis was to first preprocess the data. After preprocessing, the data was transformed by the use of the Principal Component Analysis or the Recency-Frequency-Monetary analysis, followed by k-Means clustering. The results from the k-Means clustering were used to cluster customers based on their behaviour in terms of the type of accommodation, long distance transportation or short distance transportation they used and how they rated their touch points regarding accommodation, long distance transportation, short distance transportation and transactions. Based on these clusters, valuable insights on customer behaviour in the travelling domain were achieved. After the analysis was completed, Objective 5 was achieved.

8.2 Self-assessment of work

This section presents a self-assessment of the work done during this study. The author thoroughly enjoyed finding new ways in which customer experience can be improved and managed by using engineering skills and techniques. This field of study contributes to the industrial engineering domain since a system was understood, implemented and integrated across different functionalities. In fact an industrial engineer is a key candidate to design and implement such a system.

It is important that the reader should understand that the digital information and support system represented by the Trip Planner Demonstrator goes beyond the typical apps or websites that are currently available for managing a customer trip in the travel domain.

Examples of existing systems are *TripIt*, *TripCase* and *WorldMate*. These systems integrate the itinerary of the flight, car rental, accommodation, restaurant bookings *etc.* into a single master itinerary. Based on the master itinerary the system then informs a customer on a real-time basis of all the upcoming events as the trip progresses. In other words, these systems manage a trip for a customer based on the itineraries booked and provided by the customer.

Examples of another type of system are *Expedia*, *Travelstart* and *Cheapflights*. These systems can book a trip for a customer, where a trip can consist of a flight, hotel, car or any combination of these three options. The customer makes their own decision on which options they would prefer and the system only integrates all options available for the date, time and place specified by the customer.

However, the Trip Planner Demonstrator developed by the author combines these systems and even goes beyond their capabilities. The Trip Planner Demonstrator books a trip for a customer based on their preferences and historical behaviour without the customer having to make any decision. The trip includes the travel arrangement to get to the destination, travel arrangements required at the destination and at home and the type of accommodation required. The system then manages the experiences a customer has on the trip, by informing them about upcoming events, suggesting drinks and food when the customer is at an airport or bus station and applying changes as necessary when a customer is unhappy.

8.3 Recommendations for future work

The author also learnt a lot with regards to the following:

1. The crucial importance of always striving to deliver the ‘superior customer experience’. If an enterprise does not have any customers it will close its doors because customers are the key driver for any type of enterprise. Therefore, the enterprise culture, employees, managers, *etc.* should all strive to satisfy the customers’ needs. Since a customer’s behaviour is influenced by so many different aspects, it is also one of the most sensitive ‘mechanisms’ to understand in any enterprise. However, that should not prevent an enterprise from catering for their customers.
2. The power that lies in the field of data analytics and how important it is to first preprocess and transform data before applying any type of analysis on it. The author has also learnt that data analysis is a trial-and-error exercise before the ‘best’ data analytics tool is found.

The shortcomings of this study is that the author had to simulate data which does not necessarily reflect a realistic view of real-life data. Even though statistics and information available online have been used, a lot of assumptions were made by the author from her own reference and how she perceives things to be. Another shortcoming is that since only the gender, area of residence and birth year of the customer were used, the clusters obtained by the analysis are very broad. If one uses more customer attributes, better clusters can be obtained.

8.3 Recommendations for future work

This section presents a few suggestion for potential future work to further the research:

- Currently the bookings are only done based on the customer’s preferences and history. To improve the booking, customer profiles can be included by using the techniques demonstrated in Chapter 1. A study done by another student in the same research group looked at customer super-profiling (Walters, 2018) which illustrates that customer super-profiles are a better tool to use for profiling a customer.
- The value which can be generated by the Trip Planner Demonstrator system for all relevant parties. A financial impact analysis can be completed to determine how this system will generate new revenue streams for both the owner of such a system and the enterprises who sign up as business partners to the system.
- The technical side of what the business partnering on a cross-functional platform should consist of. For the purpose of the study it was assumed that all requirements of this platform had been taken care of as suggested in Chapter 4. However, this study can be expanded to include this aspect.

8.4 Conclusion

This chapter concludes the research study in which it was shown that the Trip Planner Demonstrator can effectively manage and improve a customer's experiences through the use of data analysis in a partnering venture. This chapter provided a concise summary of the overall research study, a summary of the self-assessment of the work completed and recommendations for possible future work.

References

- ACHARYULU, G. (2012). Leveraging customer relationship management (CRM) in corporate hospital supply chain. *The IUP Journal of Supply Chain Management*, **4(1)**, 72 – 87. [22](#), [27](#)
- ACKOFF, R. (1989). From data to wisdom. *Journal of Applied Systems Analysis*, **16**, 3 – 9. [59](#)
- AIRBNB (2018). <https://www.airbnb.com/>, Accessed: 2018-04-12. [218](#)
- ALDENDERFER, M. & BLASHFIELD, R. (1984). *Cluster analysis*. Quantitative applications in the social sciences, Sage Publications. [66](#)
- ALLEN, B., BRESNAHAN, J., CHILDERS, L., FOSTER, I., KANDASWAMY, G., KETTIMUTHU, R., KORDAS, J., LINK, M., MARTIN, S., PICKETT, K. *et al.* (2012). Software as a service for data scientists. *Communications of the ACM*, **55(2)**, 81 – 88. [54](#)
- ALLEN, J., REICHHELD, F.F., HAMILTON, B. & MARKEY, R. (2005). Closing the delivery gap. <http://www2.bain.com/bainweb/pdfs/cms/hotTopics/closingdeliverygap.pdf>, Accessed: 2017-05-08. [15](#)
- ANAND, S. & BUCHNER, A. (1998). *Decision Support Using Data Mining*. Financial Times Management. [64](#)
- ANAND, S.S., PATRICK, A.R., HUGHES, J.G. & BELL, D. (1998). A data mining methodology for cross sales. *Knowledge-based System Journal*, **10**, 449 – 461. [64](#)
- ANDERSON, A. & SEMMELROTH, D. (2015). *Statistics for Big Data For Dummies*. New Jersey: John Wiley & Sons. [50](#)
- ANDERSON, D. (2016). The 4 Most Important Customer Experience Metrics. <https://www.iperceptions.com/blog/most-important-customer-experience-metrics>, Accessed: 2017-04-28. [20](#)
- ANDERSON, D., HOUPERT, J., FRASER, J., COCHRANE, L., TOMOIKA, N., AUSSANT, P. & BRAÜN, W. (2016). The Definitive Guide to Customer Experience: Your step-by-step guide to constructing a winning CX program. <https://www.iperceptions.com/en/the-definitive-guide-to-cx-iperceptions>, Accessed: 2017-04-11. [20](#), [31](#)
- ANDERSON, M. & PERRIN, A. (????). Technology use among seniors. <http://www.pewinternet.org/2017/05/17/technology-use-among-seniors/>, Accessed: 2018-03-28. [117](#)

REFERENCES

- ANDERSON, S. & BLANKE, T. (2012). Taking the long view: from e-science humanities to humanities digital ecosystems. *Historical Social Research/Historische Sozialforschung*, 147 – 164. [54](#)
- ANITHA, J. (2014). Determinants of employee engagement and their impact on employee performance. *International Journal of Productivity and Performance Management*, **63(3)**, 308 – 323. [25](#)
- ANN KELLER, S., KOONIN, S.E. & SHIPP, S. (2012). Big data and city living – what can it do for us? *Significance*, **9(4)**, 4 – 7. [54](#)
- ARENA, P. (2012). The role of assumptions. <http://duckofminerva.com/2012/12/the-role-of-assumptions.html>, Accessed: 2018-05-16. [114](#)
- ASSURY, L. (2002). *Pursuing Happiness: American Consumers in the Twentieth Century*. CMP Media. [13](#)
- ÄYRÄMÖ, S. & KÄRKKÄINEN, T. (2006). Introduction to partitioning-based clustering methods with a robust example. *Reports of the Department of Mathematical Information Technology. Series C, Software engineering and computational intelligence 1/2006*. [66](#)
- AZEVEDO, A. & SANTOS, M.F. (2008). KDD, SEMMA and CRISP-DM: a parallel overview. *IADIS European Conference Data Mining*, 182 – 185. [62](#), [63](#)
- BAMOSSY, G., SOLOMON, M., ASKEGAARD, S. & HOGG, M.K. (2006). *Consumer Behaviour A European Perspective*. Prentice Hall Europe, 3rd edition. [47](#)
- BANKS, J., CARSON, J.S. & NELSON, B.L. (1996). *Discrete-Event System Simulation*. Pearson, 2nd edition. [114](#), [115](#)
- BARLING, D., SHARPE, R. & LANG, T. (2009). Traceability and ethical concerns in the UK wheat-bread chain: from food safety to provenance to transparency. *International Journal of Agricultural Sustainability*, **7(4)**, 261 – 278. [11](#)
- BARROS, A.J. & HIRAKATA, V.N. (2003). Alternatives for logistic regression in cross-sectional studies: an empirical comparison of models that directly estimate the prevalence ratio. *BMC medical research methodology*, **3(1)**, 21. [67](#)
- BATES, D.M. & WATTS, D.G. (1988). *Nonlinear regression analysis and its applications*. Wiley Online Library. [67](#)
- BATRA, R. & KAWECKI, A. (2014). Customer Experience Management – lifecycle model. *TM Forum Framework Best Practice*. [38](#), [39](#), [40](#), [44](#)

REFERENCES

-
- BEARD, R. (2014). Customer Experience does not have to be a guessing game: Measure these 3 metrics. <http://blog.clientheartbeat.com/customer-experience-metrics/>, Accessed: 2017-04-28. 20
- BEATH, C., BECERRA-FERNANDEZ, I., ROSS, J. & SHORT, J. (2012). Finding value in the information explosion. *MIT Sloan Management Review*, **53(4)**, 18. 54
- BEKKER, J. (2015). Short notes on aspects of Discrete-event Simulation. Unpublished notes. 114, 138
- BEN-DAVID, S. & SHALEV-SHWARTZ, S. (2014). *Understanding Machine Learning: From Theory to Algorithms*. Cambridge University Press. 62, 64, 65, 67
- BERKHIN, P. (2006). A survey of clustering data mining techniques. In *Grouping multidimensional data*, 25 – 71, Springer. 66, 67
- BERRY, L.L. (1995). Relationship marketing of services – growing interest, emerging perspectives. *Journal of the Academy of marketing science*, **23(4)**, 236 – 245. 14
- BERRY, L.L., CARBONE, L.P. & HAECKEL, S.H. (2012). Managing the total customer experience. *MIT Sloan Management Review*, **43(3)**, 84 – 89. 46
- BEST, I., SHAMIR, S., AMALFITANO, M., COTTON, S., KAWECKI, A., CHOWDHURY, S., MITRA, S. & SENDEL, R. (2016). Customer Experience Management – Introduction and Fundamentals. *TM Forum Framework Best Practice*. 13, 15, 17, 19, 22, 28, 29, 30, 35, 37, 42
- BISHOP, C.M. (2006). *Pattern recognition and machine learning*. Springer. 67
- BITNER, M.J. (1990). Evaluating service encounters: the effects of physical surroundings and employee responses. *The Journal of Marketing*, 69 – 82. 14
- BITNER, M.J. (1992). Servicescapes: The impact of physical surroundings on customers and employees. *The Journal of Marketing*, 57 – 71. 14
- BITNER, M.J., OSTROM, A.L. & MORGAN, F.N. (2008). Service blueprinting: a practical technique for service innovation. *California management review*, **50(3)**, 66 – 94. 14
- BLOCKEEL, H. & MOYLE, S. (2002). Collaborative data mining needs centralised model evaluation. 64
- BLOOM, J.Z. (2004). Tourist market segmentation with linear and non-linear techniques. *Tourism Management*, **25(6)**, 723 – 733. 66

REFERENCES

- BOJA, C., POCOVNICU, A. & BATAGAN, L. (2012). Distributed parallel architecture for “big data”. *Informatica Economica*, **16(2)**, 116. [54](#)
- BOLTON, R.N. & DREW, J.H. (1991). A multistage model of customers’ assessments of service quality and value. *Journal of consumer research*, **17(4)**, 375 – 384. [14](#)
- BOLTON, R.N., LEMON, K.N. & VERHOEF, P.C. (2004). The theoretical underpinnings of customer asset management: A framework and propositions for future research. *Journal of the Academy of Marketing Science*, **32(3)**, 271 – 292. [14](#)
- BOTH, G. & VAN RENSBURG, A. (2010). Proposed business process improvement model with integrated experience management. *South African Journal of Industrial Engineering*, **21(1)**, 45 – 57. [33](#)
- BOUNSAITHIP, C. & RINTA-RUNSALA, E. (2001). Overview of data mining for customer behavior modeling. *VTT Information Technology Research Report, Version*, **1**, 1 – 53. [67](#)
- BOYD, D. & CRAWFORD, K. (2012). Critical questions for big data: provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society Journal*, **15(5)**, 662 – 679. [54](#), [57](#)
- BRACHMAN, R.J. & ANAND, T. (1996). *Advances in Knowledge Discovery and Data Mining*, chap. The process of knowledge discovery in databases, 37 – 57. American Association for Artificial Intelligence. [64](#)
- BREIMAN, L., FRIEDMAN, J., STONE, C.J. & OLSHEN, R.A. (1984). *Classification and regression trees*. CRC press. [66](#)
- BRINKMANN, B.H., BOWER, M.R., STENGEL, K.A., WORRELL, G.A. & STEAD, M. (2009). Large-scale electrophysiology: acquisition, compression, encryption, and storage of big data. *Journal of neuroscience methods*, **180(1)**, 185 – 192. [54](#)
- BRITISH MUSEUM BLOG (2017). 29 things you (probably) didn’t know about the British Museum. <http://blog.britishmuseum.org/29-things-you-probably-didnt-know-about-the-british-museum/>, Accessed: 2017-03-29. [12](#)
- BRODIE, R.J., HOLLEBEEK, L.D., JURIC, B. & ILIC, A. (2011). Customer engagement: conceptual domain, fundamental propositions, and implications for research. *Journal of Service Research*, **14(3)**, 252 – 271. [14](#)

REFERENCES

- BROWN, B., CHUI, M. & MANYIKA, J. (2011). Are you ready for the era of ‘big data’. *McKinsey Quarterly*, **4(1)**, 24 – 35. [54](#)
- BUCHNER, A.G., MULVENNA, M.D., ANAND, S.S. & HUGHES, J.G. (1999). An internet-enabled knowledge discovery process. In *Proceedings of the 9th international database conference, Hong Kong*, vol. 1999, 13–27. [64](#)
- BUGHIN, J., CHUI, M. & MANYIKA, J. (2010). Clouds, big data, and smart assets: Ten tech-enabled business trends to watch. *McKinsey quarterly*, **56(1)**, 75 – 86. [54](#)
- BUGHIN, J., LIVINGSTON, J. & MARWAHA, S. (2011). Seizing the potential of ‘big data’. *McKinsey Quarterly*, **4**, 103 – 109. [54](#)
- BURNS, J., WARREN, L. & OLIVEIRA, J. (2014). Business partnering : Is it all that good ? *Controlling & Management Review*, **58(2)**, 36 – 41. [75](#), [78](#), [86](#)
- CABENA, P., HADJINIAN, P., STADLER, R., VERHEES, J. & ZANASI, A. (1997). *Discovering data mining: from concept to implementation*. Prentice Hall PTR New Jersey. [64](#)
- CAKIR, O. & ARAS, M.E. (2012). A recommendation engine by using association rules. *Procedia-Social and Behavioral Sciences*, **62**, 452 – 456. [66](#)
- ÇANAĞOĞLU, E. & BILGIÇ, T. (2007). Analysis of a two-stage telecommunication supply chain with technology dependent demand. *European journal of operational research*, **177**. [12](#)
- CELEBI, M.E. (2014). *Partitional clustering algorithms*. Springer. [66](#)
- CHAFFEY, D. & WOOD, S. (2005). *Business Information Management: Improving Performance Using Information Systems*. Financial Times Prentice Hall. [59](#)
- CHAN, C.C.H. (2005). Online auction customer segmentation using a neural network model. *International Journal of Applied Science and Engineering*, **3(2)**, 101 – 109. [66](#)
- CHAPMAN, P., CLINTON, J., KERBER, R., Khabaza, T., REINARTZ, T., SHEARER, C. & WIRTH, R. (2000). CRISP-DM 1.0: Step-by-step data mining guide. techreport, DaimlerChrysler, SPSS, OHRA and NCR. [63](#), [64](#)
- CHATTERJEE, S. & HADI, A.S. (2006). *Regression analysis by example*. John Wiley & Sons, 4th edition. [67](#)
- CHEN, D.Q., PRESTON, D.S. & SWINK, M. (2015). How the use of big data analytics affects value creation in supply chain management. *Journal of Management Information Systems*, **32(4)**, 4 – 39. [59](#)

REFERENCES

-
- CHEN, H., CHIANG, R.H. & STOREY, V.C. (2012). Business intelligence and analytics: from big data to big impact. *MIS quarterly*, 1165 – 1188. [54](#)
- CHEN, I. & POPOVICH, K. (2003). Understanding customer relationship management (CRM). *Business Process Management Journal*, **9**, 672 – 688. [22](#), [24](#)
- CHEN, M., MAO, S. & LIU, Y. (2014). Big data: A survey. *Mobile Networks and Applications*, **19(2)**, 171 – 209. [52](#), [56](#), [57](#)
- CHEN, M.C., CHIU, A.L. & CHANG, H.H. (2005). Mining changes in customer behavior in retail marketing. *Expert Systems with Applications*, **28(4)**, 773 – 781. [157](#)
- CHIU, S. & TAVELLA, D. (2008). Chapter 7 - Introduction to Data Mining. In *Data Mining and Market Intelligence for Optimal Marketing Returns*, 137 – 192, Butterworth-Heinemann, Boston. [66](#)
- CIOS, K., TERESINSKA, A., KONIECZNA, S., POTOCKA, J. & SHARMA, S. (2000). Diagnosing myocardial perfusion from pect bull's-eye maps — a knowledge discovery approach. *IEEE Engineering in Medicine and Biology Magazine*, **19**, 17 – 25. [64](#)
- CIOS, K.J. & KURGAN, L.A. (2005). *Advanced Techniques in Knowledge Discovery and Data Mining*, chap. Trends in data mining and knowledge discovery, 37 – 57. Springer. [64](#)
- COLE, J.B., NEWMAN, S., FOERTTER, F., AGUILAR, I. & COFFEY, M. (2012). Breeding and Genetics Symposium: Really big data: Processing and analysis of very large data sets. *Journal of Animal Science*, **90(3)**, 723 – 733. [54](#), [57](#)
- COMPUTICKET TRAVEL (2018). <https://www.computickettravel.com/>, Accessed: 2018-04-12. [224](#)
- CONEY, D. (1944). *The Southwestern Social Science Quarterly*, **25(3)**, 222 – 224. [50](#)
- CONSTANTINIDES, E. (2004). Influencing the online consumer's behavior: the web experience. *Internet Research*, **14(2)**, 111 – 126. [47](#)
- COUSSEMENT, K. & VAN DEN POEL, D. (2008). Churn prediction in subscription services: An application of support vector machines while comparing two parameter-selection techniques. *Expert systems with applications*, **34(1)**, 313 – 327. [66](#)
- CRISP-DM CONSORTIUM (2008). The CRISP-DM blog. <http://crispdm.wordpress.com>, Accessed: 2017-08-02. [64](#)

REFERENCES

-
- DANDRIDGE, M. (2010). Customer Experience Management. <http://ewweb.com/sales/customer-experience-management>, Accessed: 2017-04-04. 15, 21
- DANSION, F. & GRIFFIN, J. (2012). Analytics and the cloud-the future is here. *Financial Executive*, **28(9)**, 97 – 99. 54
- DAVENPORT, T.H., BARTH, P. & BEAN, R. (2012). How big data is different. *MIT Sloan Management Review*, **54(1)**, 43. 54
- DE MAURO, A., GRECO, M. & GRIMALDI, M. (2015). What is Big Data? A consensual definition and a review of key research topics. *AIP Conference Proceedings*, **1644**, 97 – 104. 51
- DEBUSE, J., DE LA IGLESIA, B., HOWARD, C. & RAYWARD-SMITH, V. (2001). Building the kdd roadmap. In *Industrial Knowledge Management*, 179–196, Springer. 64
- DEMIRKAN, H. & DELEN, D. (2013). Leveraging the capabilities of service-oriented decision support systems: Putting analytics and big data in cloud. *Decision Support Systems*, **55(1)**, 412 – 421. 54
- DEMŠAR, J. & ZUPAN, B. (2013). Orange: Data mining fruitful and fun-a historical perspective. *Informatika*, **37(1)**. 66
- DIADMIN (2015). Characteristics of big data – part one. <http://www.dataintensity.com/characteristics-of-big-data-part-one/>, Accessed: 2017-05-11. 51
- DODDS, D. (2016). What is customer experience? *The Huffington Post*. 12, 16, 45
- DORI, D. (2002). *Object-Process Methodology: A holistic systems approach*. Springer. 95, 96
- DU PLESSIS, L. & DE VRIES, M. (2016). Towards a hollistic customer experience management framework for enterprises. *South African Journal of Industrial Engineering*, **27(3)**, 23 – 36. 13, 28, 32, 33, 36
- DU PREEZ, N., ESSMAN, H., LOUW, L., SCHUTTE, C., MARAIS, S. & BAM, W. (2015). *Enterprise Engineering textbook*. Stellenbosch University — Industrial Engineering — Enterprise Engineering Group. 9
- DWYER, F.R., SCHURR, P.H. & OH, S. (1987). Developing buyer-seller relationships. *The Journal of marketing*, 11 – 27. 14
- ELGENDY, N. & ELRAGAL, A. (2014). Big data analytics : A literature review paper. In *Industrial Conference on Data Mining*, Springer. 59
- ELIOT, T.S. (1934). *The rock*. Harcourt, Brace. 59

REFERENCES

- ELLIOTT, R. & JAMES, E. (1989). Varieties of client experience in psychotherapy: An analysis of the literature. *Clinical Psychology Review*, **9**, 443 – 467. [46](#)
- ELLRAM, L.M. & HENDRICK, T.E. (1995). Partnering characteristics: A dyadic perspective. *Journal of Business Logistics*, **16(1)**, 41 – 64. [74](#), [78](#)
- ESTER, M., KRIEGEL, H.P., SANDER, J., XU, X. *et al.* (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. In *Kdd*, vol. 96(34), 226 – 231. [67](#)
- EU (2016). Regulation (EU) 2016/679 of the european parliament and of the council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC. *Official Journal of the European Union*, **L 119**, 1 – 88. [87](#), [93](#)
- FAGGELLA, D. (2018). What is machine learning? - an informed definition. <https://www.techemergence.com/what-is-machine-learning/>, Accessed: 2018-10-01. [62](#)
- FAVIER, J. (1998). *Gold & spices: the rise of commerce in the Middle Ages*. Holmes & Meier London. [73](#)
- FAYYAD, U., PIATETSKY-SHAPIRO, G. & SMYTH, P. (1996a). The kdd process for extracting useful knowledge from volumes of data. *Communication of the ACM*, **39**, 27 – 34. [64](#)
- FAYYAD, U.M. (1996). Data mining and knowledge discovery: making sense out of data. *IEEE Expert*, **11(5)**, 20 – 25. [62](#)
- FAYYAD, U.M., PIATETSKY-SHAPIRO, G., SMYTH, P. & UTHURUSAMY, R. (1996b). *Advances in knowledge discovery and data mining*, chap. From Data Mining to Knowledge Discovery: An Overview, 1 – 36. MIT Press: Cambridge. [62](#), [64](#)
- FISHER, D., DELINE, R., CZERWINSKI, M. & DRUCKER, S. (2012). Interactions with big data analytics. *interactions*, **19(3)**, 50 – 59. [54](#)
- GADMAN, S. (1997). *Power partnering : a strategy for business excellence in the 21st century*. Butterworth-Heinemann, Boston, Mass. [45](#), [76](#), [77](#), [78](#), [81](#)
- GALLANT (1975). Nonlinear regression. *The American Statistician*, **29(2)**, 73 – 81. [67](#)
- GANDOMI, A. & HAIDER, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, **35(2)**, 137 – 144. [49](#), [50](#), [51](#)
- GEHRKE, J. (2012). Quo vadis, data privacy? *Annals of the New York Academy of Sciences*, **1260(1)**, 45 – 54. [54](#)

REFERENCES

-
- GENTILEA, C., SPILLER, N. & NOCI, G. (2007). How to sustain the Customer Experience: An overview of experience components that co-create value with the customer. *European Management Journal*, **25(5)**, 395 – 410. [33](#), [42](#), [46](#), [47](#)
- GERA, M. & GOEL, S. (2015). Data mining-techniques, methods and algorithms: A review on tools and their validity. *International Journal of Computer Applications*, **113(18)**. [67](#)
- GERTOSIO, C. & DUSSAUCHOY, A. (2004). Knowledge discovery from industrial databases. *Journal of Intelligent Manufacturing*, 29 – 37. [64](#)
- GILES, J. (2013). What does the GDPR mean for the POPI act? <https://www.michalsons.com/blog/gdpr-mean-popi-act/19959>, Accessed: 2018-09-19. [94](#)
- GOLDENBERG, B. (2017). CXM: Give your customers the experiences they want. *CRM Magazine*, **21(4)**, 6. [27](#), [28](#), [42](#)
- GRAFT, A. (2018). Travel and Tourism Statistics: The ultimate collection. <https://blog.accessdevelopment.com/tourism-and-travel-statistics-the-ultimate-collection#other>, Accessed: 2018-06-11. [117](#)
- GREWAL, D., LEVYB, M. & KUMAR, V. (2009). Customer Experience Management in Retailing: An organizing framework. *Journal of Retailing*, **85(1)**, 1 – 14. [33](#)
- GRIFFIN, R. (2012). Using big data to combat enterprise fraud: to combat fraud, more organizations are thinking big-employing new approaches around big data to convert the volumes of information available into useful insight and real action. *Financial executive*, **28(10)**, 44–48. [54](#)
- GULATI, R. & OLDROYD, J.B. (2005). The quest for customer focus. *Harvard Business Review*, **83(4)**, 92 – 101. [14](#)
- GUPTA, S. & ZEITHAML, V. (2006). Customer metrics and their impact on financial performance. *Marketing science*, **25(6)**, 718 – 739. [14](#)
- HA, H.Y. & PERKS, H. (2005). Effects of consumer perceptions of brand experience on the web: brand familiarity, satisfaction and brand trust. *Journal of Consumer Behaviour*, **4(6)**, 438 – 452. [47](#)
- HALKIDI, M., BATISTAKIS, Y. & VAZIRGIANNIS, M. (2001). On clustering validation techniques. *Journal of intelligent information systems*, **17(2)**, 107 – 145. [67](#)
- HAMID, N.R.A. & AKHIR, R.M. (2014). Beyond technology-based customer relationship management: It is total customer experience management. *Research in Business and Economic Journal*, 1 – 16. [42](#)

REFERENCES

-
- HAN, J., KAMBER, M. & PEI, J. (2012). *Data Mining Concepts and Techniques*. Morgan Kaufmann, Waltham, 3 edition. 121
- HARRIS, L.C. & REYNOLDS, K.L. (2003). The Consequences of Dysfunctional Customer Behavior. *Journal of Service Research*, **6(2)**, 144 – 161. 47
- HARRY, M. & SCHROEDER, R. (1999). *Six Sigma, the Breakthrough Management Strategy Revolutionizing the World's Top Corporations*. Currency. 64
- HARVEY, C. (2017). Big data. <http://www.datamation.com/big-data/what-is-big-data.html>, Accessed: 2017-06-02. 51, 55, 56, 57
- HASAN, R., BUCKLEY, P.J. & GLAISTER, K.W. (2002). International joint ventures: partnering skills and cross cultural issues. *Long Range Planning*, **35**, 113 – 134. 77, 78
- HASTIE, T., TIBSHIRANI, R. & FRIEDMAN, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction*. Springer, 2nd edition. 66
- HAVENS, T.C., BEZDEK, J.C., LECKIE, C., HALL, L.O. & PALANISWAMI, M. (2012). Fuzzy c-means algorithms for very large data. *IEEE Transactions on Fuzzy Systems*, **20(6)**, 1130 – 1146. 54
- HERNÁNDEZ, B., JIMÉNEZ, J. & MARTIN, M.J. (2010). Customer behavior in electronic commerce: The moderating effect of e-purchasing experience. *Journal of Business Research*, **63(9)**, 964 – 971, advances in Internet Consumer Behavior & Marketing Strategy. 47
- HIGHFIELD, T. (2012). Talking of many things: using topical networks to study discussions in social media. *Journal of technology in human services*, **30(3-4)**, 204 – 218. 54
- HIMBERG, J. (2000). A SOM based cluster visualization and its application for false coloring. In *Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks (IJCNN 200)*, vol. 3, 587 – 592, IEEE. 67
- HOLBROOK, M.B. & HIRSCHMAN, E.C. (1982). The experiential aspects of consumption: Consumer fantasies, feelings, and fun. *The Journal of Consumer Research*, **9(2)**, 132 – 140. 13, 46, 47
- HOLLEBEEK, L.D., GLYNN, M.S. & BRODIE, R.J. (2014). Consumer brand engagement in social media: Conceptualization, scale development and validation. *Journal of interactive marketing*, **28(2)**, 149 – 165. 14
- HOSMER JR, D.W., LEMESHOW, S. & STURDIVANT, R.X. (2013). *Applied logistic regression*. John Wiley & Sons, volume 398. 67

REFERENCES

- HOTELS.COM (2018). <https://za.hotels.com/>, Accessed: 2018-04-12. 218
- HOWARD, J.A. & SHETH, J.N. (1969). *The theory of buyer behavior*. Wiley New York, Volume 14. 14
- HSSINA, B., MERBOUHA, A., EZZIKOURI, H. & ERRITALI, M. (2014). A comparative study of decision tree id3 and c4. 5. *International Journal of Advanced Computer Science and Applications*, **4(2)**, 13 – 19. 66
- HUANG, J.J., TZENG, G.H. & ONG, C.S. (2007). Marketing segmentation using support vector clustering. *Expert systems with applications*, **32(2)**, 313 – 317. 66
- HUNT, S.D., LAMBE, C.J. & WITTMANN, C.M. (2002). A Theory and Model of Business Alliance Success. *Journal of Relationship Marketing*, **1(1)**, 17 – 35. 74
- HUWE, T.K. (2012). New technology, new workflows, new ways to collaborate. *Computers in libraries*, **32(3)**, 20 – 22. 54
- I-SCOOP (2016). Digitization, digitalization and digital transformation: the differences. <https://www.i-scoop.eu/digitization-digitalization-digital-transformation-disruption/>, Accessed: 2017-09-20. 50
- ILANGO, M. & MOHAN, D.V. (2010). A survey of grid based clustering algorithms. *International Journal of Engineering Science and Technology*, **2(8)**, 3441 – 3446. 67
- INVESTOPEDIA (2017). Customer. <http://www.investopedia.com/terms/c/customer.asp>, Accessed: 2017-05-11. 9
- ISHIBUCHI, H. & YAMAMOTO, T. (2005). Rule weight specification in fuzzy rule-based classification systems. *IEEE transactions on fuzzy systems*, **13(4)**, 428 – 435. 66
- ISHIBUCHI, H., YAMAMOTO, T. & NAKASHIMA, T. (2005). Hybridization of fuzzy GBML approaches for pattern classification problems. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, **35(2)**, 359 – 365. 66
- IZENMAN, A.J. (2008). *Modern Multivariate Statistical Techniques: regression, classification, and Manifold Learning*. Springer, volume 1 edn. 66, 67
- JACOB, S.G. & RAMANI, R.G. (2012). Data mining in clinical data sets: a review. *Foundation of Computer Science*, **4(6)**. 66
- JAGADISH, H., GEHRKE, J., LABRINIDIS, A., PAPAKONSTANTINOY, Y., PATEL, J.M., RAMAKRISHNAN, R. & SHAHABI, C. (2014). Big data and its technical challenges. *Communications of the ACM*, **57(7)**, 86 – 94. 57

REFERENCES

-
- JANIKOW, C.Z. (1998). Fuzzy decision trees: issues and methods. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, **28(1)**, 1 – 14. [66](#)
- JANSEN, S. (2007). Customer segmentation and customer profiling for a mobile telecommunications company based on usage behavior. *A Vodafone Case Study*. [66](#)
- JAYAWARDHENA, C. (2004). Personal values' influence on e-shopping attitude and behaviour. *Internet Research*, **14(2)**, 127 – 138. [47](#)
- JIawei, H., KAMBER, M. & PEI, J. (2012). *Data Mining: Concepts and Techniques*. Morgan Kaufman, 3rd edition. [65](#), [66](#), [67](#)
- JIFA, G. & LINGLING, Z. (2014). Data, DIKW, Big Data and Data Science. *Procedia Computer Science*, **31**, 814 – 821. [49](#)
- JOHNSON, B.D. (2012a). The secret life of data. *The Futurist*, **46(4)**, 20. [54](#)
- JOHNSON, J.E. (2012b). Big data + big analytics = big opportunity: big data is dominating the strategy discussion for many financial executives. as these market dynamics continue to evolve, expectations will continue to shift about what should be disclosed, when and to whom. *Financial Executive*, **28(6)**, 50 – 54. [54](#)
- KAHRAMAN, C., OLY NDUBISI, N., KOK WAH, C. & NDUBISI, G.C. (2007). Supplier-customer relationship management and customer loyalty: The banking industry perspective. *Journal of Enterprise Information Management*, **20(2)**, 222 – 236. [27](#)
- KAISLER, S., ARMOUR, F., ESPINOSA, J.A. & MONEY, W. (2013). Big Data: Issues and challenges moving forward. In *2013 46th Hawaii International Conference on System Sciences*, 995 – 1004. [52](#), [56](#), [57](#)
- KALE, S.H. (2004). CRM failure and the seven deadly sins. *Marketing Management*, **13**, 42 – 46. [26](#)
- KAMRANI, A.K., PARSAEI, H.R. & CHAUDHRY, M.A. (1993). A survey of design methods for manufacturing cells. *Computers & industrial engineering*, **25(1-4)**, 487 – 490. [67](#)
- KARP, A.H. (1998). Using logistic regression to predict customer retention. In *Proceedings of the Eleventh Northeast SAS Users Group Conference..* [67](#)
- KATAL, A., WAZID, M. & GOUDAR, R.H. (2013). Big data: Issues, challenges, tools and good practices. In *2013 6th International Conference on Contemporary Computing, IC3 2013*, 404 – 409. [51](#), [53](#), [56](#), [57](#)

REFERENCES

-
- KENDALL, K.E. & KENDALL, J.E. (2014). *Systems Analysis and Design*. Pearson Education, 9th edition. 99, 100
- KERRIGAN, M., DONG, W., KAWECKI, A., NACHMAN, S., STOCK, C., SENDEL, R., AKSELA, M. & MITRA, S. (2017). Big Data Analytics Guidebook: Use Cases. *TM Forum Framework Best Practice*. 83
- KIERCZAK, L. (2017). 5 Crucial Customer Satisfaction Metrics. <https://survicate.com/customer-satisfaction/metrics/>, Accessed: 2017-04-28. 17
- KIM, S.Y., JUNG, T.S., SUH, E.H. & HWANG, H.S. (2006). Customer segmentation and strategy development based on customer lifetime value: A case study. *Expert systems with applications*, **31(1)**, 101 – 107. 66
- KOLKER, E., STEWART, E. & ÖZDEMİR, V. (2012). DELSA Global for “big data” and the Bioeconomy: Catalyzing collective innovation. *Industrial Biotechnology*, **8(4)**, 176 – 178. 54
- KOTSIANTIS, S.B. (2007). Supervised machine learning: A review of classification techniques. In *Proceedings of the 2007 Conference on Emerging Artificial Intelligence Applications in Computer Engineering: Real World AI Systems with Applications in eHealth, HCI, Information Retrieval and Pervasive Technologies*, 3 – 24, IOS Press, Amsterdam, The Netherlands, The Netherlands. 66
- KRASNIKOV, A., JAYACHANDRAN, S. & KUMAR, V. (2009). The impact of customer relationship management implementation on cost and profit efficiencies: evidence from the US commercial banking industry. *Journal of marketing*, **73(6)**, 61 – 76. 27
- KRISTENSEN, K., MARTENSEN, A. & GRONHOLDT, L. (1999). Measuring the impact of buying behaviour on customer satisfaction. *Total Quality Management*, **10**, 602 – 614. 47
- KUMAR, V. & REINARTZ, W. (2006). *Customer Relationship Management: A Databased Approach*. New York: John Wiley & Sons. 14
- KUMAR, V. & SHAH, D. (2009). Expanding the role of marketing: from customer equity to market capitalization. *Journal of Marketing*, **73(6)**, 119 – 136. 14
- KUMAR, V., PETERSEN, J.A. & LEONE, R.P. (2010). Driving profitability by encouraging customer referrals: who, when, and how. *The Journal of Marketing*, **74(5)**, 1 – 17. 14
- KUMAR, V., BHASKARAN, V., MIRCHANDANI, R. & SHAH, M. (2013). Practice prize winner—creating a measurable social media marketing strategy: increasing the value and ROI of intangibles and tangibles for hokey pokey. *Marketing Science*, **32(2)**, 194 – 212. 14

REFERENCES

-
- KUO, R., AN, Y., WANG, H. & CHUNG, W. (2006). Integration of self-organizing feature maps neural network and genetic k-means algorithm for market segmentation. *Expert systems with applications*, **30(2)**, 313 – 324. [66](#)
- LABRINIDIS, A. & JAGADISH, H.V. (2012). Challenges and Opportunities with Big Data. In *Proceedings of the VLDB Endowment*, vol. 5(12), 2032 – 2033. [56](#)
- LAKE, R. (2018). Business Travel Statistics: 23 speedy facts to know. <https://www.creditdonkey.com/business-travel-statistics.html>, Accessed: 2018-05-03. [117](#)
- LANE, J. (2012). O privacy, where art thou?: Protecting privacy and confidentiality in an era of big data access. *Chance*, **25(4)**, 39 – 41. [54](#)
- LANEY, D. (2001). 3D Data Management: Controlling data volume, velocity and variety. *META Group Research*, **6**, 4. [51](#)
- LANOUE, S. (2016). Customer experience metrics: a brief guide on how to measure CX. <https://www.usertesting.com/blog/2016/05/19/customer-experience-metrics/>, Accessed: 2017-04-28. [20](#)
- LAROSE, D.T. (2014). *Discovering knowledge in data: an introduction to data mining*. John Wiley & Sons, 2nd edition. [66](#), [67](#)
- LAUVIDGE ROBERT, J. & STEINER GARY, A. (1961). A model for predicting measurements of advertising effectiveness. *Journal of Marketing*, **25**. [14](#)
- LAVALLE, S., LESSER, E., SHOCKLEY, R., HOPKINS, M.S. & KRUSCHWITZ, N. (2011). Big data, analytics and the path from insights to value. *MIT Sloan Management Review*, **52(2)**, 21 – 32. [54](#), [57](#), [58](#)
- LAW, A.M. & KELTON, W.D. (2000). *Simulation modeling and analysis*. McGraw-Hill, 3 edition. [138](#)
- LAWRENCE, R.L. & WRIGHT, A. (2001). Rule-based classification systems using classification and regression tree (CART) analysis. *Photogrammetric engineering and remote sensing*, **67(10)**, 1137 – 1142. [66](#)
- LEATHERBARROW, R.J. (1990). Using linear and non-linear regression to fit biochemical data. *Trends in biochemical sciences*, **15(12)**, 455 – 458. [67](#)
- LEBERGOTT, S. (1993). *Pursuing Happiness: American Consumers in the Twentieth Century*. Princeton, NJ: Princeton University Press. [13](#)

REFERENCES

- LEHMAN, F. (2016). We all are our company's customer experience. http://www.huffingtonpost.com/fritz-lehman/we-all-are-our-companys-c_b_11389540.html, Accessed: 2017-04-04. 16
- LEKKEŚLAAP (2018). <https://www.lekkeslaap.co.za/>, Accessed: 2018-04-12. 218
- LEMON, K.N. & VERHOEF, P.C. (2014). Understanding Customer Experience Throughout the Customer Journey. *Journal of Marketing*, **80(6)**, 69 – 96. 13, 14, 37, 38, 41
- LI, R. (2015). Top 10 data mining algorithms, explained. <http://www.kdnuggets.com/2015/05/top-10-data-mining-algorithms-explained.html>, Accessed: 2017-07-27. 66
- LIBAI, B., BOLTON, R., BÜGEL, M.S., DE RUYTER, K., GÖTZ, O., RISSELADA, H. & STEPHEN, A.T. (2010). Customer-to-customer interactions: broadening the scope of word of mouth research. *Journal of Service Research*, **13(3)**, 267 – 282. 14
- LIN, G.F. & WU, M.C. (2007). A SOM-based approach to estimating design hyetographs of ungauged sites. *Journal of Hydrology*, **339(3-4)**, 216 – 226. 67
- LINDGREEN, A. & SWAEN, V. (2010). Corporate social responsibility. *International Journal of Management Reviews*, **12(1)**, 1 – 7. 20
- LINOFF, G.S. & BERRY, M.J. (2011). *Data mining techniques: for marketing, sales, and customer relationship management*. John Wiley & Sons. 66
- LLETUÍ, R., ORTIZ, M., SARABIA, L. & SÁNCHEZ, M. (2004). Selecting variables for k-means cluster analysis by using a genetic algorithm that optimises the silhouettes. *Analytica Chimica Acta*, **515(1)**, 87 – 100, Paper presented at the 5th Colloquium Chemioetricum Mediterraneum. 158
- MADHULATHA, T.S. (2011). Comparison between k-means and k-medoids clustering algorithms. *Advances in Computing and Information Technology*, 472 – 481. 66, 67
- MAHMOOD, T. & AFZAL, U. (2013). Security analytics : Big data analytics for cybersecurity. In *Second National Conference on Information Assurance (NCIA)*, 129 – 134. 60
- MANGIAMELI, P., CHEN, S.K. & WEST, D. (1996). A comparison of SOM neural network and hierarchical clustering methods. *European Journal of Operational Research*, **93(2)**, 402 – 417. 67
- MANYIKA, J., CHUI, M., BROWN, B., BUGHIN, J., DOBBS, R., ROXBURGH, C. & BYERS, A.H. (2011). Big data: The next frontier for innovation, competition, and productivity. *McKinsey Global Institute*. 52

REFERENCES

-
- MARBÁN, O., MARISCAL, G., MENASALVAS, E. & SEGOVIA, F.J. (2007). An engineering approach to data mining projects. *Lecture Notes in Computer Science 4881*, 578 – 588. 64
- MARBÁN, O., SEGOVIA, J., MENASALVAS, E. & FERNANDEZ-BAIZAN, C. (2008). Towards data mining engineering: a software engineering approach. *Information Systems Journal*. 64
- MARISCAL, G., ÓSCAR MARBÁN & FERNÁNDEZ, C. (2010). A survey of data mining and knowledge discovery process models and methodologies. *The Knowledge Engineering Review*, **25(2)**, 137 – 166. 63, 64
- MARR, B. (2017). A brief history of big data everyone should read. <https://www.weforum.org/agenda/2015/02/a-brief-history-of-big-data-everyone-should-read/>, Accessed: 2017-09-19. 49
- MATHIEU, R.G. & GIBSON, J.E. (1993). A methodology for large-scale R&D planning based on cluster analysis. *IEEE Transactions on Engineering Management*, **40(3)**, 283 – 292. 67
- MCAFEE, A. & BRYNJOLFSSON, E. (2012). Big Data : The management revolution. *Harvard Business Review*, **2012**, 1 – 9. 51, 54, 56, 57, 58
- MEIJER, E. (2011). The world according to LINQ. *Communications of the ACM*, **54(10)**, 45 – 51. 54
- MEYER, C. & SCHWAGER, A. (2007). Understanding Customer Experience. *Harvard business review*, **8(2)**, 117 – 126. 13, 20, 22, 25, 36, 41
- MITCHELL, T. (1997). *Machine Learning*. McGraw-Hill. 64
- MITIK, M., KORKMAZ, O., KARAGOZ, P. & TOROSLU, F., ISMAIL HAKKI YUCEL (2017). Data mining approach for direct marketing of banking products with profit/cost analysis. *The Review of Socionetwork Strategies*, **11(1)**, 17 – 31. 186
- MITRA, S. & KAWECKI, A. (2016a). Guidebook on 360-Degree view of a customer. *TM Forum*. 45
- MITRA, S. & KAWECKI, A. (2016b). Omni channel guidebook. *TM Forum Framework Best Practice*. 42
- MOLOISANE, P. (2004). Strategy document for the South African wheat to bread value chain. *Walterberg Municipality*, Compiled for the Wheat Steering Committee. 11
- MONTGOMERY, D.C., PECK, E.A. & VINING, G.G. (2012). *Introduction to linear regression analysis*. John Wiley & Sons, volume 821. 67

REFERENCES

- MORGAN, R. (2015). Customer experience management: Here's how to create 'advocates' for your dealership through positive sales encounters. *Business Insights*, **Nov**, 63 – 66. [28](#)
- MORGAN, R.M. & HUNT, S.D. (1994). The commitment-trust theory of relationship marketing. *The Journal of marketing*, 20 – 38. [14](#)
- MOYLE, S. & JORGE, A. (2001). RAMSYS-A methodology for supporting rapid remote collaborative data mining projects. In *ECML/PKDD01 Workshop: Integrating Aspects of Data Mining, Decision Support and Meta-learning (IDDM-2001)*. [64](#)
- MUNICIPALITIES OF SOUTH AFRICA (2018). <https://municipalities.co.za/>, Accessed: 2018-03-05. [230](#)
- NESLIN, S.A., GREWAL, D., LEGHORN, R., SHANKAR, V., TEERLING, M.L., THOMAS, J.S. & VERHOEF, P.C. (2006). Challenges and opportunities in multichannel customer management. *Journal of Service Research*, **9(2)**, 95 – 112. [14](#)
- OHATA, M. & KUMAR, A. (2012). Big data: a boon to business intelligence. *Financial Executive*, **28(7)**, 63 – 65. [54](#)
- OLIVER, R.L. (1980). A cognitive model of the antecedents and consequences of satisfaction decisions. *Journal of marketing research*, 460 – 469. [14](#)
- OPEN GROUP ARCHITECTURAL FRAMEWORK (2001). <http://www.opengroup.org/public/arch/>, Accessed: 2018-05-29. [95](#)
- OPPERMAN, J. (2017). Why is customer journey mapping so important? https://www.sas.com/en_gb/insights/articles/marketing/why-is-customer-journey-mapping-important.html/, Accessed: 2017-11-06. [41](#)
- OXFORD ENGLISH DICTIONARY (1409). Definition: 'customer'. <http://www.oed.com.ez.sun.ac.za/>, Accessed: 2017-05-11. [9](#)
- OXFORD ENGLISH DICTIONARY (1980). Definition: 'big data'. <http://www.oed.com.ez.sun.ac.za/>, Accessed: 2017-05-23. [49](#)
- ÖZGENER, Ş. & İRAZ, R. (2006). Customer relationship management in small–medium enterprises: The case of Turkish tourism industry. *Tourism Management*, **27(6)**, 1356 – 1363. [27](#)
- PALI WAL, M. & KUMAR, U.A. (2009a). Neural networks and statistical techniques: A review of applications. *Expert systems with applications*, **36(1)**, 2 – 17. [66](#)

REFERENCES

- PALIWAL, M. & KUMAR, U.A. (2009b). A study of academic performance of business school graduates using neural network and statistical techniques. *Expert Systems with Applications*, **36(4)**, 7865 – 7872. [67](#)
- PARAMASIVAM, V., YEE, T.S., DHILLON, S.K. & SIDHU, A.S. (2014). A methodological review of data mining techniques in predictive medicine: An application in hemodynamic prediction for abdominal aortic aneurysm disease. *Biocybernetics and Biomedical Engineering*, **34(3)**, 139 – 145. [66](#)
- PARANDKER, S.R. & LOKKU, D. (2012). Customer Experience Management. In *Third International Conference on Services in Emerging Markets Customer*, 44 – 49. [28](#), [30](#), [34](#), [37](#), [42](#)
- PARASURAMAN, A., ZEITHAML, V.A. & BERRY, L.L. (1988). SERVQUAL: A multiple-item scale for measuring consumer perception of service quality. *Journal of retailing*, **64(1)**, 12. [14](#)
- PARISE, S. & CASHER, A. (2003). Alliance portfolios: Designing and managing your network of business-partner relationships. *Academy of Management Executive*, **17(4)**, 25 – 39. [76](#)
- PARK, N.H. & LEE, W.S. (2004). Statistical grid-based clustering over data streams. *Acm Sigmod Record*, **33(1)**, 32 – 37. [67](#)
- PAYNE, A. & FOW, P. (2005). A strategic framework for customer relationship management. *Journal of marketing*, **69(4)**, 167 – 176. [14](#), [22](#), [23](#), [25](#), [26](#), [27](#)
- PAYNE, A. & FROW, P. (2007). Towards the ‘perfect’ customer experience. *Journal of Brand Management*, **15(2)**, 89 – 101. [16](#), [20](#)
- PEDUZZI, P., CONCATO, J., KEMPER, E., HOLFORD, T.R. & FEINSTEIN, A.R. (1996). A simulation study of the number of events per variable in logistic regression analysis. *Journal of clinical epidemiology*, **49(12)**, 1373 – 1379. [67](#)
- PERRIN, A. (2015). Social media usage: 2005 – 2015. *Pew Research Center*, **October**. [19](#)
- PETROULAKIS, N. & MIAOUDAKIS, A. (2007). An application of neural networks in market segmentation. In *PanHellenic Conference in New Technologies and Marketing*, 185 – 190. [66](#)
- PIERSON, L. (2015). *Data science for dummies*. John Wiley & Sons. [66](#), [67](#)
- PINE, J.B. & GILMORE, J.H. (1999). The experience economy: Work is theatre and every business a stage. *Harvard Business School Press, Boston, Massachusetts*. [13](#)
- PUCCINELLI, N.M., GOODSTEINB, R.C., GREWAL, D., PRICE, R., RAGHUBIRE, P. & STEWART, D. (2009). Customer Experience Management in Retailing: Understanding the buying process. *Journal of Retailing*, **85(1)**, 15 – 30. [46](#)

REFERENCES

-
- PYZDEK, T. (2003). *The Six Sigma Handbook*. McGraw-Hill, 2nd edition. 64
- QUINLAN, J.R. (2014). *C4. 5: programs for machine learning*. Elsevier. 66
- RAHIMI, R. (2017). Organizational culture and customer relationship management: A simple linear regression analysis. *Journal of Hospitality Marketing & Management*, **26(4)**, 1 – 7. 25
- RAJARAJESWARI, A. & RAVINDRAN, R.M. (2015). A comparative study of k-means, k-medoid and enhanced k-medoid algorithms. *International Journal of Advance Foundation and Research in Computer (IJAFRC)*, **2(8)**, 7 – 10. 66
- RECHENTHIN, M.D. (2014). *Machine-learning classification techniques for the analysis and prediction of high-frequency stock direction*. The University of Iowa. 66
- REINARTZ, W., KRAFFT, M. & HOYER, W.D. (2004). The customer relationship management process: Its measurement and impact on performance. *Journal of marketing research*, **41(3)**, 293 – 305. 14
- REINARTZ, W.J. & KUMAR, V. (2000). On the profitability of long-life customers in a noncontractual setting: An empirical investigation and implications for marketing. *The Journal of marketing*, **64(4)**, 17 – 35. 14
- REZNIK, Y. (2016). These metrics will eliminate the guesswork from measuring Customer Experience. <https://www.nanorep.com/six-essential-metrics-measuring-customer-experience>, Accessed: 2017-04-28. 20
- RICH, R. (2015). Customer Experience and Analytics in a Digital World. *TM Forum Framework Best Practice*. 2, 30, 34, 45, 60, 82
- RICHARDSON, A. (2010). Understanding customer experience. *Harvard business review*. 13, 17, 37, 45
- RIDER, F. (1994). *The Scholar and the Future of the Research Library: A Problem and Its Solution*. Hadham Press. 49
- RIFFENBURGH, R. (2011). *Statistics in Medicine*. Elsevier Science. 67
- ROBLES, F. (1994). International market entry strategies and performance of United States catalog firms. *Journal of Interactive Marketing*, **8(1)**, 59 – 70. 67
- ROKACH, L. & MAIMON, O. (2014). *Data mining with decision trees: theory and applications*. World scientific. 66

REFERENCES

-
- RONCO, W.C. & RONCO, J.S. (2005). *The Partnering Solution: A Powerful Strategy for Managers, Professionals, and Employees at All Levels*. The Career Press, New Jersey. 78
- ROWE, S.D. (2017). Customer first! Businesses do better with customer-first marketing. *CRM Magazine*, **21(2)**, 14. 21
- ROWLEY, J. (2006). The wisdom hierarchy: representations of the DIKW hierarchy. *Journal of Information Science*, **33(2)**, 163 – 180. 58
- RUCKSTUHL, A. (2010). Introduction to nonlinear regression. 67
- RUSSELL, P. (2005). Consciousness and Reality. <http://www.peterrussell.com/Reality/RHTML/R2.php>, Accessed: 2017-11-08. 46
- RUSSOM, P. (2011). Big data analytics. *TDWI Best Practices Report*. 51
- RUST, R.T. & CHUNG, T.S. (2006). Marketing models of service and relationships. *Marketing Science*, **25(6)**, 560 – 580. 14
- RUST, R.T., LEMON, K.N. & ZEITHAML, V.A. (2004). Return on marketing: Using customer equity to focus marketing strategy. *The Journal of marketing*, **68(1)**, 109 – 127. 14
- RUZICKA, N. (2017a). Customer centricity: Analytics in action. <https://inform.tmforum.org/customer-centricity/2017/09/customer-centricity-analytics-action/>, Accessed: 2017-10-10. 37
- RUZICKA, N. (2017b). Make data do the work. <https://inform.tmforum.org/customer-centricity/2017/08/make-data-work/>, Accessed: 2017-10-10. 37
- SA GOVERNMENT (2013). Protection of Personal Information Act 4 of 2013. <https://www.gov.za/documents/protection-personal-information-act>, Accessed: 2018-09-19. 87, 94
- SAGIROGLU, S. & SINANC, D. (2013). Big data: A review. *2013 International Conference on Collaboration Technologies and Systems*, 42 – 47. 51
- SALKIND, N.J. (2007). *Encyclopedia of measurement and statistics*. Sage, volume 1. 66, 67
- SANDER, J., ESTER, M., KRIEGEL, H.P. & XU, X. (1998). Density-based clustering in spatial databases: The algorithm GDBSCAN and its applications. *Data mining and knowledge discovery*, **2(2)**, 169 – 194. 67
- SAS INSTITUTE (2005). SEMMA data mining methodology. <http://www.sas.com/technologies/analytics/datamining/miner/semma.html>., Accessed: 2017-08-02. 64

REFERENCES

- SCHADT, E.E., LINDERMAN, M.D., SORENSON, J., LEE, L. & NOLAN, G.P. (2010). Computational solutions to large-scale data management and analysis. *Nature Reviews Genetics*, **11**(9), 647. [54](#)
- SCHIKUTA, E. & ERHART, M. (1997). The bang-clustering system: Grid-based data analysis. In *International Symposium on Intelligent Data Analysis*, 513 – 524, Springer. [67](#)
- SCHMIDHUBER, J. (2015). Deep learning in neural networks: An overview. *Neural networks*, **61**, 85 – 117. [65](#)
- SCHMITT, B.H. (2003). *Customer experience management : a revolutionary approach to connecting with your customers*. New York : Wiley. [22](#), [27](#), [35](#)
- SCHULTE BEERBÜHL, M. (2012). Networks of the hanseatic league. <http://ieg-ego.eu/en/threads/european-networks/economic-networks/margrit-schulte-beerbuehl-networks-of-the-hanseatic-league>, Accessed: 2017-09-21. [73](#)
- SHAH, D., RUST, R.T., PARASURAMAN, A., STAELIN, R. & DAY, G.S. (2006). The path to customer centricity. *Journal of service research*, **9**(2), 113 – 124. [14](#)
- SHETH, J.N. & PARVATIYAR, A. (1992). Towards a theory of business alliance formation. *Scandinavian International Business Review*, **3**, 71 – 87. [74](#), [75](#), [76](#), [77](#), [78](#), [81](#), [82](#)
- SHETH, J.N., SISODIA, R.S. & SHARMA, A. (2000). The antecedents and consequences of customer-centric marketing. *Journal of the Academy of Marketing Science*, **28**(1), 55 – 66. [14](#)
- SIEMENS, G. & LONG, P. (2011). Penetrating the fog: Analytics in learning and education. *EDUCAUSE Review*, **46**(5), 30. [54](#)
- SIVARAJAH, U., KAMAL, M.M., IRANI, Z. & WEERAKKODY, V. (2017). Critical analysis of Big Data challenges and analytical methods. *Journal of Business Research*, **70**, 263 – 286. [55](#), [56](#), [57](#), [58](#)
- SMITH, M., SZONGOTT, C., HENNE, B. & VON VOIGT, G. (2012). Big data privacy issues in public social media. In *2012 6th IEEE International Conference on Digital Ecosystems Technologies (DEST)*, 1 – 6, IEEE. [54](#)
- SOARES, L. (2012). The rise of big data. *EDUCAUSE Review*, **47**(3), 60. [54](#)
- SOBEK, M., CLEVELAND, L., FLOOD, S., KELLY HALL, P., KING, M.L., RUGGLES, S. & SCHROEDER, M. (2011). Big data: large-scale historical infrastructure from the minnesota population center. *Historical methods*, **44**(2), 61 – 68. [54](#)

REFERENCES

- SÖDERLUND, M. (1998). Customer satisfaction and its consequences on customer behaviour revisited: The impact of different levels of satisfaction on word-of-mouth, feedback to the supplier and loyalty. *International Journal of Service Industry Management*, **9(2)**, 169 – 188. [47](#)
- SOLARTE, J. (2002). A proposed data mining methodology and its application to industrial engineering. *Master's thesis, University of Tennessee, Knoxville*. [64](#)
- SOLOMON, M., RUSSELL-BENNETT, R. & PREVITE, J. (2012). *Consumer Behaviour*. Pearson Higher Education AU, 3rd edition. [47](#)
- SORZANO, C.O.S., VARGAS, J. & MONTANO, A.P. (2014). A survey of dimensionality reduction techniques. *arXiv preprint arXiv:1403.2877*. [156](#)
- STATS SA (2018). <http://www.statssa.gov.za/>, Accessed: 2018-03-05. [218](#)
- STEINBACH, M., KARYPIS, G., KUMAR, V. *et al.* (2000). A comparison of document clustering techniques. In *KDD workshop on text mining*, vol. 400(1), 525 – 526, Boston. [66](#), [67](#)
- STEINBERG, D. & COLLA, P. (2009). *The top ten algorithms in data mining*, vol. 9, chap. 10: CART: classification and regression trees, 179 – 200. CRC Press. [66](#)
- STEYNBERG, R. (2016). *A framework for identifying the most likely successful underprivileged tertiary bursary applicants*. Ph.D. thesis, p.H.D. thesis, Stellenbosch University. [66](#)
- STRAWN, G.O. (2012). Scientific research: How many paradigms?. *EDUCAUSE Review*, **47(3)**, 26. [54](#)
- SUKWADI, R. (2015). Utilizing customer experience management framework to create a delightful service experience. *International Journal of Industrial Engineering and Management*, **6(1)**, 29 – 42. [28](#), [33](#), [46](#)
- TANG, W., SANVILLE, E. & HENKELMAN, G. (2009). A grid-based bader analysis algorithm without lattice bias. *Journal of Physics: Condensed Matter*, **21(8)**, 084204. [67](#)
- TANKARD, C. (2012). Big data security. *Network security*, **2012(7)**, 5 – 8. [54](#)
- TAXI CALCULATOR (2018). <https://www.taxi-calculator.com/>, Accessed: 2018-03-12. [228](#)
- TELLIS, G.J. (2006). Modeling marketing mix. *Handbook of marketing research*, 506 – 522. [67](#)
- TELLIS, G.J. & AMBLER, T. (2007). *The SAGE handbook of advertising*. Sage. [67](#)
- TM FORUM (2017). About tm forum. <https://www.tmforum.org/about-tm-forum/>, Accessed 2017-11-01. [83](#)

REFERENCES

- TOMAR, D. & AGARWAL, S. (2013). A survey on data mining approaches for healthcare. *International Journal of Bio-Science and Bio-Technology*, **5**(5), 241 – 266. 66
- TOURISM GRADING COUNCIL (2018). <https://www.tourismgrading.co.za/>, Accessed: 2018-03-05. 103
- TRAVELSTART (2018). <https://www.travelstart.co.za/>, Accessed: 2018-03-19. 224
- TRUTH (2017). South african loyalty landscape 2017. <https://truth.co.za/wp-content/uploads/Truth-Whitepaper-October-2017.pdf>, Accessed: 2018-05-09. 113, 221
- TRUTH & BRANDMAPP (2018). South african loyalty landscape 2018/19. <http://truth.co.za/wp-content/uploads/Truth-BrandMapp-Loyalty-Whitepaper-201819.pdf>, Accessed: 2018-08-20. 221
- TSIPTSIS, K.K. & CHORIANOPOULOS, A. (2011). *Data mining techniques in CRM: inside customer segmentation*. John Wiley & Sons. 67, 157, 158
- TURNER, A. (2017a). How do platform principles apply to telcos ? <https://inform.tmforum.org/internet-of-everything/2017/08/platform-principles-apply-telcos/>, Accessed: 2017-11-01. 85, 86
- TURNER, A. (2017b). Platforms: Design, deploy, diversify, repeat. <https://inform.tmforum.org/features-and-analysis/2017/05/platforms-design-deploy-diversify-repeat/>, Accessed: 2017-11-01. 85, 86, 87
- TWO CROWS CORPORATION (1998). *Introduction to Data Mining and Knowledge Discovery*. Two Crows Corporation, 2nd edition. 64
- TWO CROWS CORPORATION (1999). *Introduction to Data Mining and Knowledge Discovery*. Two Crows Corporation, 3rd edition. 64
- UBER ESTIMATE (2018). <http://uberestimate.com/prices/Johannesburg/>, Accessed: 2018-03-12. 228
- USMA (2017). USMA Working Group, Unit for Systems Modelling and Analysis. 60, 61, 66
- VAN DOORN, J., LEMON, K.N., MITTAL, V., NASS, S., PICK, D., PIRNER, P. & VERHOEF, P.C. (2010). Customer engagement behaviour: Theoretical foundations and research directions. *Journal of Service Research*, **13**(3), 253 – 266. 14
- VAPNIK, V. (1999). *The Nature of Statistical Learning Theory*. Springer Science & Business Media. 66

REFERENCES

- VENKATRAMAN, S. (2015). Business partnering - when does it work ? *Strategic Finance*, 47 – 53. 78, 80
- VENUGOPAL, P. & PRIYA, A. (2015). The impact of customer service on Customer Relationship Management. *Global Management Review*, **10(1)**, 139 – 152. 27
- VERHOEF, P.C. (2003). Understanding the effect of customer relationship management efforts on customer retention and customer share development. *The Journal of marketing*, **67(4)**, 30 – 45. 14
- VERHOEF, P.C., LEMON, K.N., PARASURAMANC, A., ROGGEVEEND, A., TSIROS, M. & SCHLESINGER, L.A. (2009). Customer Experience Creation: Determinants, Dynamics and Management Strategies. *Journal of Retailing*, **85(1)**, 31 – 41. 13, 28, 33, 45, 47
- VESANTO, J. (1999). SOM-based data visualization methods. *Intelligent data analysis*, **3**, 111 – 126. 67
- VORHIES, W. (2014). How Many "V's" in Big Data? the Characteristics that Define Big Data. <https://www.datasciencecentral.com/profiles/blogs/how-many-v-s-in-big-data-the-characteristics-that-define-big-data>, Accessed: 2017-05-11. 51
- WAGNER, E. (2012). Realities learning professionals need to know about analytics. *T and D*, **66(8)**, 54 – 58. 54
- WALDEN, S. (2017). *Customer Experience Management Rebooted: Are you an experience brand or an efficiency brand?*. London: Macmillan Publishers. 22, 35
- WALLER, M.A. & FAWCETT, S.E. (2013). Data Science, Predictive Analytics, and Big Data : A revolution that will transform supply chain design and management. *Journal of Business Logistics*, **34(2)**, 77 – 84. 59
- WALTERS, M. (2018). Development and demonstration of a customer super-profiling tool utilising data analytics for alternative targeting in marketing campaigns. 191
- WAMBA, S.F., AKTER, S., EDWARDS, A., CHOPIN, G. & GNANZOU, D. (2015). How 'big data' can make big impact: Findings from a systematic review and a longitudinal case study. *International Journal of Production Economics*, **165**, 234 – 246. 53, 54, 57, 58
- WHITE, M. (2012). Digital workplaces: Vision and reality. *Business information review*, **29(4)**, 205 – 214. 54

REFERENCES

- WINER, R.S. (2001). A framework for Customer Relationship Management. *California Management Review*, **43(4)**, 89 – 105. [23](#), [24](#), [27](#)
- WITTEN, I.H., FRANK, E., HALL, M.A. & PAL, C.J. (2016). *Data Mining: Practical Machine Learning Tools and Techniques*. Elsevier, 4th edition. [62](#), [64](#)
- WU, S. & CHOW, T.W. (2004). Clustering of the self-organizing map using a clustering validity index based on inter-cluster and intra-cluster density. *Pattern Recognition*, **37(2)**, 175 – 188. [67](#)
- WU, X., KUMAR, V., ROSS QUINLAN, J., GHOSH, J., YANG, Q., MOTODA, H., MCLACHLAN, G.J., NG, A., LIU, B., YU, P.S., ZHOU, Z.H., STEINBACH, M., HAND, D.J. & STEINBERG, D. (2008). Top 10 algorithms in data mining. *Knowledge and Information Systems*, **14(1)**, 1 – 37. [66](#)
- WUCHERER, K. (2006). Business partnering - a driving force for innovation. *Industrial Marketing Management*, **35(1)**, 91 – 102. [77](#), [78](#), [91](#)
- YANG, L., LIU, S., TSOKA, S. & PAPAGEORGIOU, L.G. (2017). A regression tree approach using mathematical programming. *Expert Systems with Applications*, **78**, 347 – 357. [67](#)
- YUE, H.H. & TOMOYASU, M. (2004). Weighted principal component analysis and its applications to improve fdc performance. In *Decision and Control, 2004. CDC. 43rd IEEE Conference on*, vol. 4, 4262 – 4267, IEEE. [156](#)
- ZAGORICA, A. (2013). 5 Customer Experience metrics every successful company tracks. resreport, Buffer Social Blog, Accessed: 2017-04-28. [20](#)
- ZEITHAML, V.A. (1988). Consumer perceptions of price, quality, and value: a means-end model and synthesis of evidence. *The Journal of marketing*, 2 – 22. [14](#)
- ZIEGLER, M., ROSENZWEIG, B. & ZIEGLER, P. (1994). *The Republic of Tea: The Story of the Creation of a Business, as Told Through the Personal Letters of Its Founders*. Bantam Doubleday Dell Publishing Group. [81](#)
- ZIKOPOULOS, P., DE ROOS, D., PARASURAMAN, K., DEUTSCH, T., GILES, J. & CORRIGAN, D. (2012). *Harness the Power of Big Data*. McGraw-Hill. [53](#), [58](#)
- ZINELDIN, M. (2005). Quality and customer relationship management (CRM) as competitive strategy in the swedish banking industry. *The TQM magazine*, **17(4)**, 329 – 344. [27](#)

Appendix A

Simulation of Trip Planner Demonstrator database

This appendix provides detail about how some of the entities in the database of the Trip Planner Demonstrator (TPD) have been simulated. The focus is only on the input data required for the TPD and it does not include how the output data of the simulator was simulated. The details around the output data were discussed in Section 5.4.

1. The *accommodation* entity was populated by ensuring that there is at least three accommodation enterprises per area and the more dense populated areas will have more accommodation records. The rate per person per night has been simulated based on (i) the type of accommodation it is, (ii) the rating of the accommodation and (iii) the area where it is located. Prices from [Airbnb \(2018\)](#), [LekkeSlaap \(2018\)](#) and [Hotels.com \(2018\)](#) have been used as a guideline. The same cost scheme is applied to all accommodation enterprises. In other words all accommodation enterprises have a per night rate for one person, since the system only books a trip for one individual. A total of 603 accommodation enterprises were created for the model and Table A.1 provides a summary of how many of each type and rate was created.
2. The *Customer* entity was populated based on information provided on [Stats SA \(2018\)](#). Table A.2 represents the first 10 customers simulated. In total 7 500 customers have been simulated for the TPD system, with the assumption that they are between 18 and 70 years old. The budget constraint has been allocated in such a way that 15% of customers have a soft constraint and 85% have a hard constraint. The age distribution of the customer can be seen in Table A.3 and 51% of the customers are female and 49% are male.

Table A.1: Distribution of accommodation type and rate

		1	2	3	4	5	TOTAL
		1 Star	2 Stars	3 Stars	4 Stars	5 Stars	
1	Backpackers	3	5	13	10	3	34
2	Bed & Breakfast	8	17	34	25	12	96
3	Guest House	9	21	43	28	19	120
4	Hotel	16	48	60	42	19	185
5	Self-catering Hostel	19	32	52	37	28	168
TOTAL		55	123	202	142	81	603

Table A.2: Customer table of first 10 entries

Customer ID	First Name	Surname	Gender	Birth Year	Area
1	Barack	Mthimunye	Male	1974	10
2	Jettie	Maremba	Male	1999	208
3	Keishla	Phooko	Female	1995	199
4	Huso	Mdakane	Female	1990	140
5	Wanjiku	Mpelane	Female	1996	114
6	Carson	Seedat	Male	1989	130
7	Dennis	Adendorff	Male	1961	190
8	Osogo	Madia	Male	1955	21
9	Daisha	Mashao	Female	1993	74
10	Kwanza	Monnenyana	Male	1994	150
⋮	⋮	⋮	⋮	⋮	⋮

Table A.3: Summary of customer ages

Birth Year	Male	Female
1994 – 2000	23%	22%
1984 – 1993	29%	27%
1974 – 1983	20%	19%
1964 – 1973	14%	16%
1954 – 1963	9%	10%
1948 – 1953	5%	6%

3. The *Customer accommodation Preference* entity was simulated in such a way that

- 20% of customers have one preference,
- 35% of customers have two preferences,
- 30% of customers have three preferences and
- 15% of customers have four preferences.

Table A.4 gives the overall percentage of preferences linked to customers and there is a total of 18 000 links.

Table A.4: Summary of customer Accommodation preferences

Preference ID	Acc Type ID	Acc Rate ID	Distribution
1	1	1	0.13%
2	1	2	0.63%
3	1	3	1.13%
4	1	4	0.38%
5	1	5	0.25%
6	2	1	1.25%
7	2	2	6.25%
8	2	3	11.25%
9	2	4	3.75%
10	2	5	2.5%
11	3	1	1.37%
12	3	2	6.88%
13	3	3	12.37%
14	3	4	4.13%
15	3	5	2.75%
16	4	1	0.75%
17	4	2	3.75%
18	4	3	6.73%
19	4	4	2.25%
20	4	5	1.5%
21	5	1	1.5%
22	5	2	7.5%
23	5	3	13.5%
24	5	4	4.5%
25	5	5	3%

4. The *Customer LDT Preference* entity was simulated in such a way that 30% of the customers have one LDT preference and 70% of the customers have two preferences. When a customer has two preferences the option can either be

- Aeroplane economy class and bus economy class or
- Aeroplane economy class and bus business class.

A customer who prefers aeroplane business class, might also prefer a bus business or economy class. Since the preferences were simulated that only 10% of the ‘aeroplane’ customers prefer business class, these links have been excluded. The distribution of the customer preferences for LDT can be seen in Table A.5.

Table A.5: Summary of customer LDT preferences

LDT Type	Distribution	LDT Preference	Distribution
Aeroplane	56%	Business Class	10%
		Economy Class	90%
Bus	44%	Business Class	25%
		Economy Class	75%

5. The *Customer Loyalty Program* entity was simulated based on a study done by Truth (2017). The following considerations were taken into account:

- The age category of the customer.
- The gender of the customer.
- The total loyalty programs that can be linked to a customer.

Table A.6 provides a summary of the total customers linked to the loyalty programs. For this study, 29 loyalty programs were included in the database. In October 2018 a new white paper on loyalty amongst South Africans has been published by Truth & BrandMapp (2018).

6. The *Customer SDT Preference* entity was simulated according to the distribution as set out in Table A.7 and it was simulated such that 30% of the customers have one SDT preference and 70% of the customers have two SDT preferences. For customers with two preferences, it was simulated that

- 40% of customers prefer car rental and hailing app taxi,
- 35% of customers prefer car rental and normal taxis and
- 25% of customers prefer hailing app and normal taxis.

Table A.6: Summary of customer loyalty programs

Total Loyalty Programs	Total Customers	Distribution
0	1045	16%
2	323	5%
3	645	10%
4	1614	25%
5	1775	27%
6	1130	18%
7	645	10%
8	323	5%

Table A.7: Summary of customer SDT preferences

SDT Preference	SDT Type	Taxi Requirement	Rental Class	Distribution
1	Car Rental	None	Small Cars	15%
2	Car Rental	None	Medium Cars	15%
3	Car Rental	None	Large Cars	8%
4	Car Rental	None	Premium	0.6%
5	Car Rental	None	Premium Plus	0.4%
6	Car Rental	None	People Carriers	0.6%
7	Car Rental	None	SUV's	0.3%
8	Car Rental	None	SUV Plus	0.1%
9	Hailing App Taxi	At Destination	Not Rental Cars	12%
10	Hailing App Taxi	At Home & Destination	Not Rental Cars	22.5%
11	Normal Taxi	At Destination	Not Rental Cars	8%
12	Normal Taxi	At Home & Destination	Not Rental Cars	17.5%

7. The *Customer Transactions* entity was simulated by grouping the customer as big spenders and low spenders, as well as a wide variety of spenders and low variety of spenders. A big spender is a customer who buys a product at least once per week, where the low spender buys a product at most once per month. A wide variety spender buys different products and a low variety spender buys the same products each time. Therefore, the customer transactions were simulated as follows for a three-year period:

- (a) *Wide variety, big spender*: A total of 15% of customers are wide variety, big spenders. These customers were linked to five product categories and buy at least three products per category. The transactional data for this customer was simulated in such a way that they buy either one or two products per week and each time it will be
- a product from category 1 to 3 (drink category) and a product from category 4 to 6 (food category),
 - a product from category 1 to 3 (drink category) or
 - a product from category 4 to 6 (food category).
- (b) *Wide variety, low spender*: A total of 40% of customers are wide variety, low spenders. These customers were linked to four product categories and buy at least three products per category. The transactional data for this customer was simulated in such a way that they buy two products once per month and each time it will be a product from category 1 to 3 (drink category) and a product from category 4 to 6 (food category).
- (c) *Low variety, big spender*: A total of 35% of customers are low variety, big spenders. These customers were linked to two product categories and buy at most three products per category. The transactional data for this customer was simulated in such a way that they buy either one or two products per week and each time it will be
- a product from category 1 to 3 (drink category) and a product from category 4 to 6,
 - a product from category 1 to 3 (drink category) or
 - a product from category 4 to 6 (food category).
- (d) *Low variety, low spender*: A total of 10% of customers are low variety, low spenders. These customers were linked to two product categories and buy at most three products per category. The transactional data for this customer was simulated in such a way that they buy one product per month and each time it will be,
- a product from category 1 to 3 (drink category) or
 - a product from category 4 to 6 (food category).

-
8. The *LDT Area* entity was populated by using routes of bus and aeroplane enterprises in South Africa. Bus routes that have been used include the bus routes of *Greyhound*, *Intercape*, *Citi-liner*, *Transa*, etc. Flight routes that have been used include the flight routes of *Kulula*, *Mango*, *FlySafair*, *British Airways* etc. The total routes as well as the option of whether a business class is available or not has been simulated based on the current bus and aeroplane information available on the enterprises websites. The costs used per entity is an average cost and [Travelstart \(2018\)](#) and [Computicket Travel \(2018\)](#) have been used as a guideline. The *Long Distance Transportation* entity consists of eight aeroplane and six bus enterprises.
9. The *LDT Loyalty Programs* entity was populated by creating a link between at least one LDT enterprise and one loyalty program (keep in mind, there is a “no loyalty program” option). The links have been created as follows:
- One bus and one aeroplane enterprise have no loyalty programs.
 - Two bus and two aeroplane enterprises have one loyalty program.
 - One bus and two aeroplane enterprises have two loyalty programs.
 - One bus and one aeroplane enterprise have three loyalty programs.
 - One bus and one aeroplane enterprise have four loyalty programs.
 - One aeroplane enterprise has five loyalty programs.
10. The *Product Shop* entity was simulated by looking at products currently offered at airports and by creating a link between these products and a shop. It was simulated by looking at the product category and by ensuring that every shop will not offer every single product. A total of 56 products were created and 22 shops for this model. After the simulation, 774 links were created. [Table A.8](#) provides a snapshot of how the entity was populated.

Table A.8: Summary of product shop

Product Shop ID	Shop ID	Product ID	Category ID
1	2	1	1
2	3	1	1
3	4	1	1
⋮	⋮	⋮	⋮
158	17	11	2
159	18	11	2
160	1	12	2
⋮	⋮	⋮	⋮
344	16	25	3
345	19	25	3
346	2	26	4
347	3	26	4
⋮	⋮	⋮	⋮
580	22	40	5
581	3	41	5
582	4	41	5
⋮	⋮	⋮	⋮
772	15	56	6
773	16	56	6
774	19	56	6

11. The *SDT Area* entity was simulated by ensuring that each area has at least one taxi enterprise (normal or hailing app taxis) and that the more dense areas will have more SDT enterprises operating in that specific area. Table A.9 provides a summary of the total areas in which the type of SDT operates. For clarity, when looking at the first row it can be seen that five car rental enterprises operate in 23 areas.
12. The *SDT Loyalty Programs* entity was simulated in such a way that 85% of the SDTs have no loyalty programs and the rest are linked to at least one loyalty program. Only car rentals and hailing app taxis can be linked to a loyalty program. The reason is because most taxis do not make use of loyalty programs.

Table A.9: Summary of SDT area link

SDT Type	Total Area's	Total SDT
Car Rental	23	5
	88	8
Hailing App Taxi	8	3
	19	2
	23	2
	79	4
	87	4
	88	1
Normal Taxi	1	10
	2	1
	4	4
	7	3
	8	2
	10	1
	24	17
	34	1
	45	24
	121	36

13. The *Short Distance Transport* entity was populated in such a way that 12.5% are hailing app taxis, 10% are car rentals and 77.5% are normal taxis and it was randomly linked with a taxi rate, where the car rental enterprises will have a zero rate for all rate categories. A total of 128 SDT enterprises were created for the model. The following assumptions were made for car rentals:

- the deposit or authorisation fee that is held against the customer account is not taken into consideration since it is only credited against the customer account in the case of accidents, theft or damages.
- all the enterprises group their cars with the same car class category approach, but not all the car rental enterprises necessarily cater for all the car classes.

14. The *Station* entity was populated based on the information gathered when the LDT Area entity was created. Table A.10 contains a snapshot of how the Station entity was populated. The station entity was populated with 157 stations across South Africa for the model and it is linked back to the area.

15. The *Station Shop* entity linked the 22 shops to 157 stations and it was simulated based on the following:

- All bus stations have three shops.
- The three biggest airports in South Africa have all shops, which are OR Tambo, King Shaka and Cape Town international airport.
- The other airports have either six, 11 or 17 shops, where the allocated shops depended on the area. The more dense the area, the more shops there will be at the station.

Table A.10: Summary of stations and stores

Station ID	Station Name	Area ID
1	Arathusa Safari Airport	67
2	Bram Fischer International Airport	21
3	Cape Town International Airport	103
4	King Shaka International Airport	40
5	East London Airport	1
6	George Airport	111
7	Hoedspruit Airport	61
8	Kimberley Airport	79
9	Lanseria International Airport	39
10	Londolozi Airstrip	67
⋮	⋮	⋮
148	Vryburg Bus Station	99
149	Worcester Bus Station	104
150	Total Garage (Busbuckridge)	67
151	Excel Garage Grand National Hotel Warden	28
152	Hoedspruit Greyhound Office	61
153	Phalaborwa Greyhound Office	60
154	Graaff-Reinet Bus Station	17
155	Uitenhage Bus Station	2
156	Knysna Bus Station	113
157	Riviersonderend Bus Station	117

By creating the links as above, 651 links were created. Table A.11 provides details about how many stations were linked with a shop.

Table A.11: Station information

Shop ID	Total Stations	Shop ID	Total Stations
1	30	12	36
2	21	13	32
3	37	14	34
4	32	15	32
5	33	16	33
6	33	17	29
7	21	18	17
8	30	19	34
9	30	20	31
10	19	21	30
11	32	22	25

16. The *Taxi Rate* entity was simulated based on rates from [Uber Estimate \(2018\)](#) and [Taxi Calculator \(2018\)](#). The rates have been seen as universal and not dependent on the area due to the fact that the TPD only estimates costs during the booking of the trip. It is important to note that some of the taxi rate categories can be zero and others non-zero. The simulation of the rates has been based on the information given in Table A.12. The values of the rates have been simulated based on the minimum and maximum value as well as taking into consideration the other rates that apply to the option. The zero and one values in the option rows indicate whether the rate will be used or not, respectively. The distribution determines the total rates that have to be simulated. For this model 100 rates have been created.

Table A.12: Simulation of trip planner taxi rates

	Min Fare	Base Fare	Driving Min Rate	Waiting Min Rate	KM Rate	Distribution
Min Value	R10	R5	R0.50	R0.50	R5	-
Max Value	R100	R100	R2.50	R2.50	R40	
At least four taxi rates						
Option 1	1	1	1	0	1	22%
Option 2	1	1	0	1	1	8%
Option 3	1	0	1	1	1	6%
Option 4	0	1	1	1	1	19%
At least three taxi rates						
Option 1	1	1	0	0	1	2%
Option 2	1	0	1	0	1	9%
Option 3	1	0	0	1	1	3%
Option 4	0	1	1	0	1	19%
Option 5	0	1	0	1	1	2%
At least two taxi rates						
Option 1	1	0	0	0	1	6%
Option 2	0	1	0	0	1	3%
Option 3	1	0	1	0	0	1%

The last entity which need to be looked at is the *Area* entity. The area entity was populated based on the local municipalities of South Africa and can be seen in Table A.13. The area entity was then linked backed to the *District* entity. The district entity was populated based on the metropolitan and district municipalities ([Municipalities of South Africa, 2018](#)) and can be seen in Table A.14. The district was then linked backed to the *Province* entity. The province entity was populated based on nine provinces of South Africa as seen in Table A.15. For the model only the areas which have at least one LDT enterprise linked to it, have been used. Table A.16 gives a concatenated view of the areas considered. South Africa can be divided up into 213 areas and 52 districts but only 121 areas and 51 districts are used.

Table A.13: Areas considered for the Trip Planner Demonstrator

Area ID	Area Name
1	Buffalo City Metropolitan
2	Nelson Mandela Bay Metropolitan
3	Umzimvubu Local
4	Amahlathi Local
5	Mbhashe Local
6	Mnquma Local
7	Ngqushwa Local
8	Raymond Mhlaba Local
9	Engcobo Local
10	Enoch Mgijima Local
11	Intsika Yethu Local
12	Inxuba Yethemba Local
13	Walter Sisulu Local
14	King Sabata Dalindyebo Local
15	Mhlontlo Local
16	Blue Crane Route Local
17	Dr Beyers Naudé Local
18	Kouga Local
19	Koukamma Local
20	Makana Local
21	Mangaung Metropolitan
22	Moqhaka Local
23	Masilonyana Local
24	Matjhabeng Local
25	Nala Local
26	Tokologo Local
27	Dihlabeng Local
28	Maluti-A-Phofung Local
29	Mantsopa Local
30	Setsoto Local
31	Kopanong Local
Table A.13 continues on next page	

Area ID	Area Name
32	Mohokare Local
33	City of Ekurhuleni Metropolitan
34	City of Johannesburg Metropolitan
35	City of Tshwane Metropolitan
36	Emfuleni Local
37	Lesedi Local
38	Merafong City Local
39	Mogale City Local
40	eThekwini Metropolitan
41	Newcastle Local
42	Greater Kokstad Local
43	Ubuhlebezwe Local
44	uMzimkhulu Local
45	KwaDukuza Local
46	City of uMhlathuze Local
47	Mthonjaneni Local
48	Ray Nkonyeni Local
49	Umdoni Local
50	Umuziwabantu Local
51	Mpofana Local
52	Msunduzi Local
53	uMngeni Local
54	Big 5 Hlabisa Local
55	Endumeni Local
56	Alfred Duma Local
57	AbaQulusi Local
58	eDumbe Local
59	Polokwane Local
60	Ba-Phalaborwa Local
61	Maruleng Local
62	Makhado Local
63	Musina Local
64	Bela-Bela Local
65	Modimolle-Mookgophong Local
66	Mogalakwena Local
67	Bushbuckridge Local
68	City of Mbombela Local
69	Nkomazi Local
70	Dr Pixley Ka Isaka Seme Local
71	Govan Mbeki Local
72	Lekwa Local
73	Mkhondo Local
74	Msukaligwa Local
75	Emakhazeni Local
Table A.13 continues on next page	

Area ID	Area Name
76	Emalahleni Local
77	Steve Tshwete Local
78	Magareng Local
79	Sol Plaatje Local
80	Ga-Segonyana Local
81	Gamagara Local
82	Kamiesberg Local
83	Khai-Ma Local
84	Nama Khoi Local
85	Emthanjeni Local
86	Thembelihle Local
87	Ubuntu Local
88	Umsobomvu Local
89	Dawid Kruiper Local
90	Kai !Garib Local
91	Kgatelopele Local
92	Tsantsabane Local
93	Moses Kotane Local
94	Rustenburg Local
95	City of Matlosana Local
96	JB Marks Local
97	Maquassi Hills Local
98	Lekwa-Teemane Local
99	Naledi Local
100	Mahikeng Local
101	Ramotshere Moiloa Local
102	Tswaing Local
103	City of Cape Town Metropolitan
104	Breede Valley Local
105	Drakenstein Local
106	Stellenbosch Local
107	Beaufort West Local
108	Laingsburg Local
109	Prince Albert Local
110	Bitou Local
111	George Local
112	Hessequa Local
113	Knysna Local
114	Mossel Bay Local
115	Oudtshoorn Local
116	Swellendam Local
117	Theewaterskloof Local
118	Bergrivier Local
119	Cederberg Local
Table A.13 continues on next page	

Area ID	Area Name
120	Matzikama Local
121	Swartland Local
End of Table A.13	

Table A.14: Districts considered for the Trip Planner Demonstrator

ID	District Name	ID	District Name
1	Buffalo City Metropolitan	27	uMzinyathi District
2	Nelson Mandela Bay Metropolitan	28	uThukela District
3	Alfred Nzo District	29	Zululand District
4	Amathole District	30	Capricorn District
5	Chris Hani District	31	Mopani District
6	Joe Gqabi District	32	Vhembe District
7	OR Tambo District	33	Waterberg District
8	Sarah Baartman District	34	Ehlanzeni District
9	Mangaung Metropolitan	35	Gert Sibande District
10	Fezile Dabi District	36	Nkangala District
11	Lejweleputswa District	37	Frances Baard District
12	Thabo Mofutsanyana District	38	John Taolo Gaetsewe District
13	Xhariep District	39	Namakwa District
14	City of Ekurhuleni Metropolitan	40	Pixley Ka Seme District
15	City of Johannesburg Metropolitan	41	ZF Mgcawu District
16	City of Tshwane Metropolitan	42	Bojanala Platinum District
17	Sedibeng District	43	Dr Kenneth Kaunda District
18	West Rand District	44	Dr Ruth Segomotsi Mompati District
19	eThekweni Metropolitan	45	Ngaka Modiri Molema District
20	Amajuba District	46	City of Cape Town Metropolitan
21	Harry Gwala District	47	Cape Winelands District
22	iLembe District	48	Central Karoo District
23	King Cetshwayo District	49	Eden District
24	Ugu District	50	Overberg District
25	uMgungundlovu District	51	West Coast District
26	uMkhanyakude District		

Table A.15: Provinces considered for the Trip Planner Demonstrator

ID	Province
1	Eastern Cape
2	Free State
3	Gauteng
4	KwaZulu-Natal
5	Limpopo
6	Mpumalanga
7	Northern Cape
8	North West
9	Western Cape

Table A.16: Area concatenated

Area	District	Province	Area	District	Province	Area	District	Province	Area	District	Province
1	1	1	32	13	2	62	32	5	92	41	7
2	2	1	33	14	3	63	32	5	93	42	8
3	3	1	34	15	3	64	33	5	94	42	8
4	4	1	35	16	3	65	33	5	95	43	8
5	4	1	36	17	3	66	33	5	96	43	8
6	4	1	37	17	3	67	34	6	97	43	8
7	4	1	38	18	3	68	34	6	98	44	8
8	4	1	39	18	3	69	34	6	99	44	8
9	5	1	40	19	4	70	35	6	100	45	8
10	5	1	41	20	4	71	35	6	101	45	8
11	5	1	42	21	4	72	35	6	102	45	8
12	5	1	43	21	4	73	35	6	103	46	9
13	6	1	44	21	4	74	35	6	104	47	9
14	7	1	45	22	4	75	36	6	105	47	9
15	7	1	46	23	4	76	36	6	106	47	9
16	8	1	47	23	4	77	36	6	107	48	9
17	8	1	48	24	4	78	37	7	108	48	9
18	8	1	49	24	4	79	37	7	109	48	9
19	8	1	50	24	4	80	38	7	110	49	9
20	8	1	51	25	4	81	38	7	111	49	9
21	9	2	52	25	4	82	39	7	112	49	9
22	10	2	53	25	4	83	39	7	113	49	9
23	11	2	54	26	4	84	39	7	114	49	9
24	11	2	55	27	4	85	40	7	115	49	9
25	11	2	56	28	4	86	40	7	116	50	9
26	11	2	57	29	4	87	40	7	117	50	9
27	12	2	58	29	4	88	40	7	118	51	9
28	12	2	59	30	5	89	41	7	119	51	9
29	12	2	60	31	5	90	41	7	120	51	9
30	12	2	61	31	5	91	41	7	121	51	9
31	13	2									

Appendix B

Analysis of Trip Planner Demonstrator : Detail results

This appendix provides tables and graphs that were created during the analysis of the TPD. They were not included in the main document because of the detail they contain. The content is, however, important for support of the TPD performance analysis.

The tables and figures are as follows:

- The weighted PCA plot has been created for the analysis of all the customer behaviour and experience analyses and the weighted PCA plots are as follows for each entity:
 1. Customer LDT Behaviour – Figure [B.1](#)
 2. Customer SDT Behaviour – Figure [B.2](#)
 3. Customer Experience for Accommodation – Figure [B.3](#)
 4. Customer Experience for LDT – Figure [B.4](#)
 5. Customer Experience for SDT – Figure [B.5](#)
 6. Customer Experience for Transactions – Figure [B.6](#)
- The silhouette plots created to determine the required number of clusters during the analysis of all the customer behaviour and experience, and the weighted PCA plots are as follows for each entity:
 1. PCA for Customer LDT Behaviour – Figure [B.7](#)
 2. PCA for Customer SDT Behaviour – Figure [B.8](#)
 3. PCA for Customer Experience for LDT – Figure [B.9](#)
 4. PCA for Customer Experience for SDT – Figure [B.10](#)
 5. RFM Analysis of Transactions – Figure [B.11](#)
 6. PCA of Customer Experience of Transactions – Figure [B.12](#)
 7. RFM for Customer Experience for all Touch points – Figure [B.13](#)

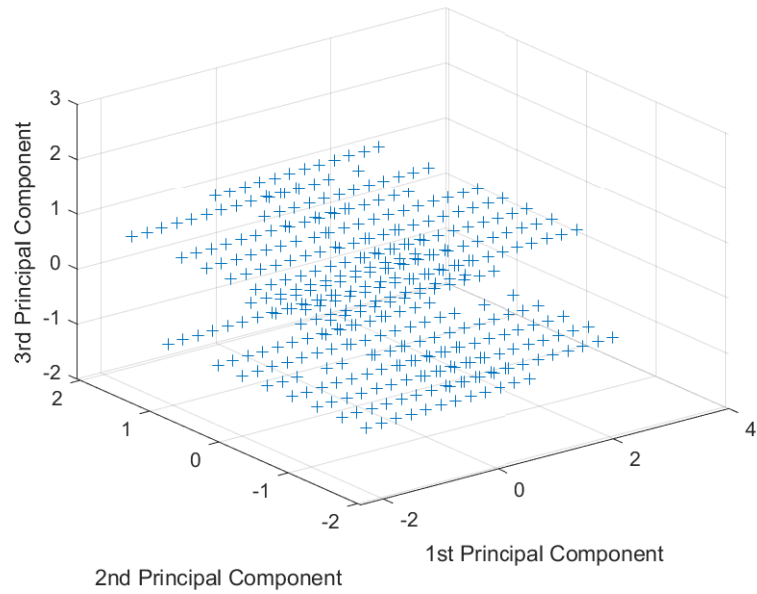


Figure B.1: Weighted PCA plot of customer LDT behaviour

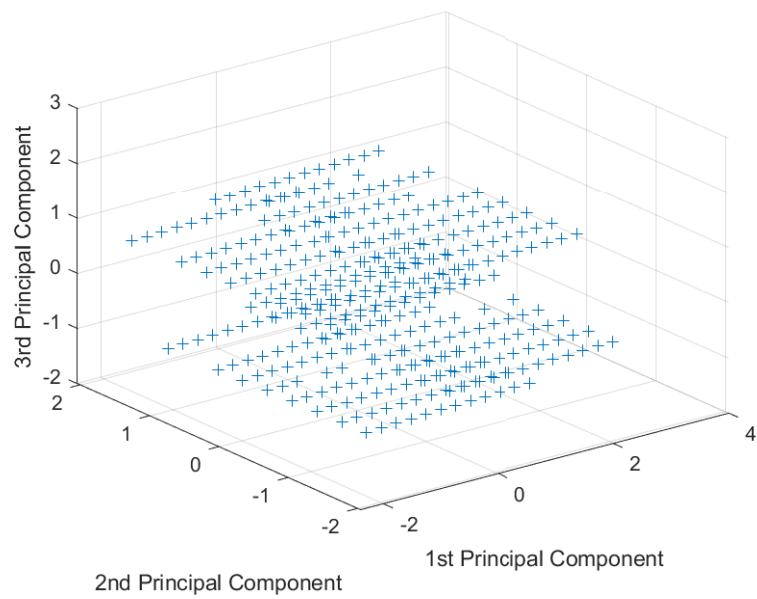


Figure B.2: Weighted PCA plot of customer SDT behaviour

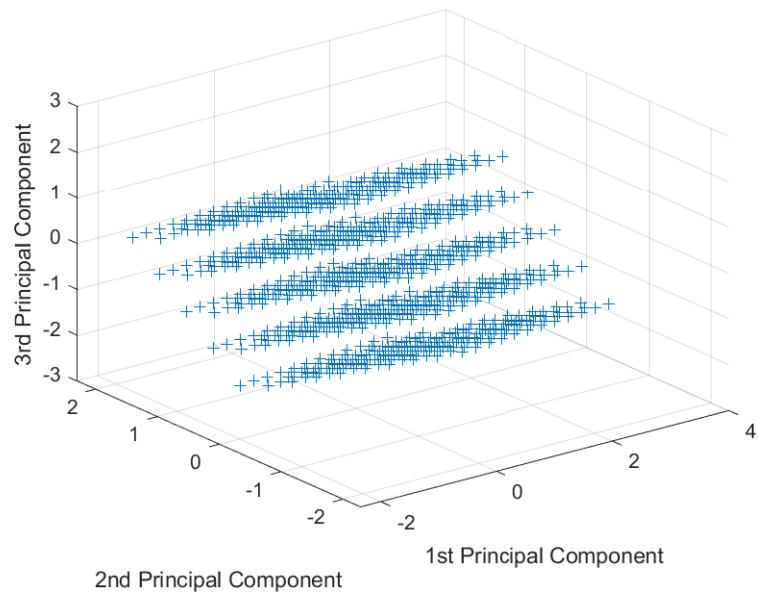


Figure B.3: Weighted PCA plot of CX – Accommodation

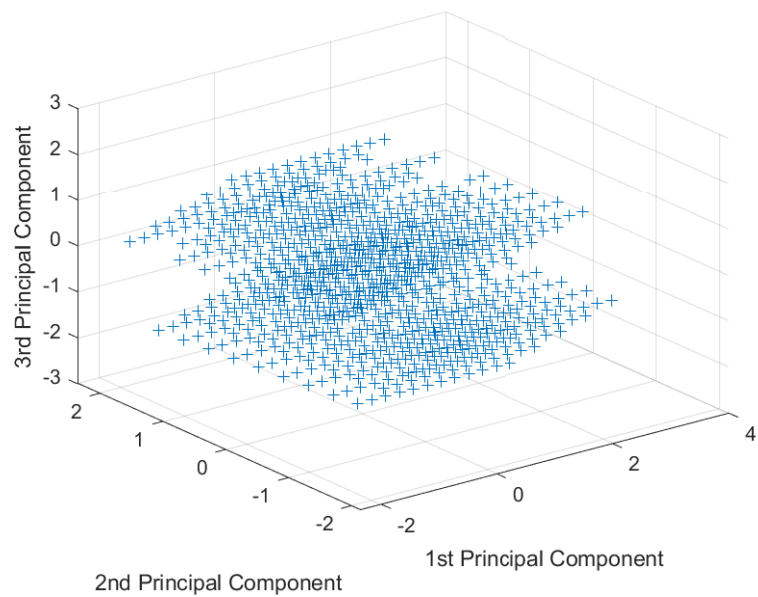


Figure B.4: Weighted PCA plot of CX – LDT

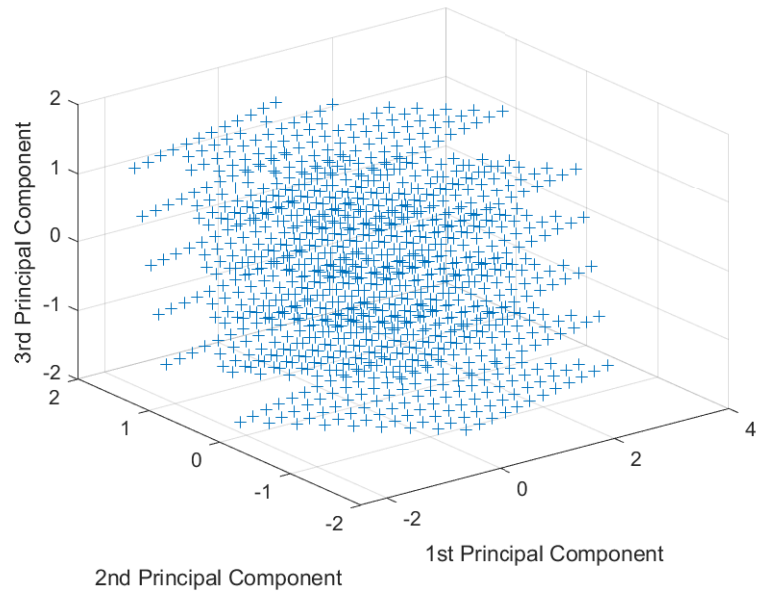


Figure B.5: Weighted PCA plot of CX – SDT

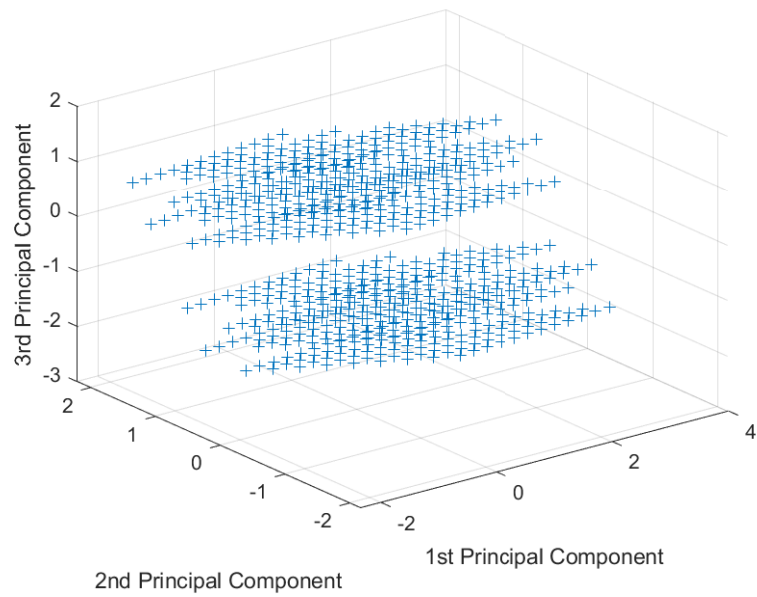


Figure B.6: Weighted PCA plot of CX – Transactions

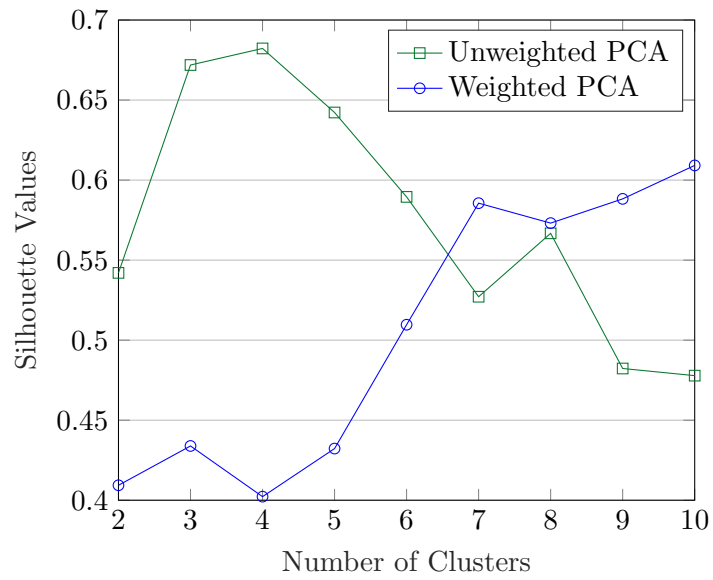


Figure B.7: Silhouette for customer LDT behaviour

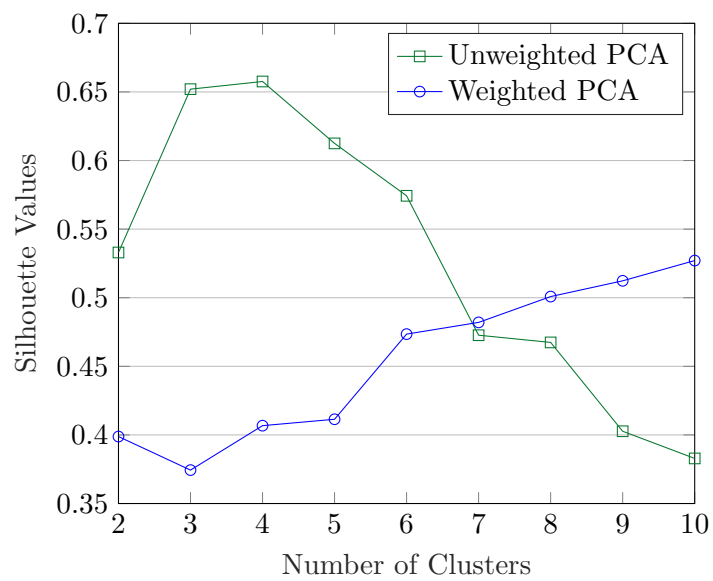


Figure B.8: Silhouette customer SDT behaviour

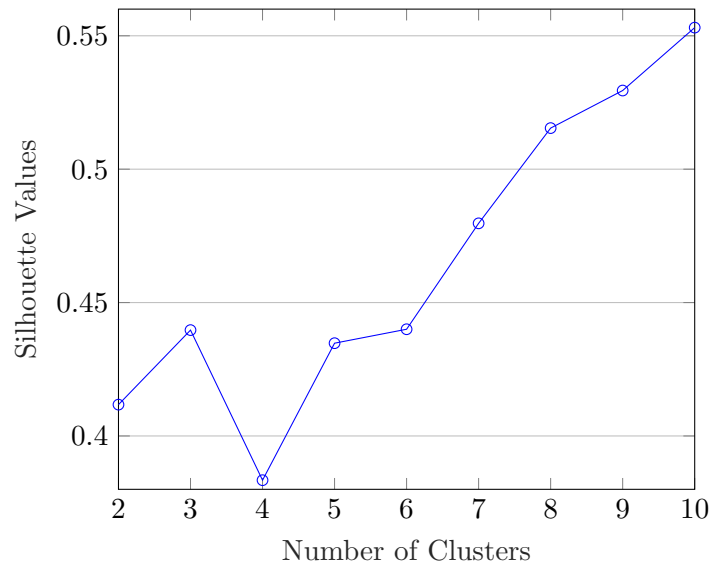


Figure B.9: Silhouette plot for CX – LDT

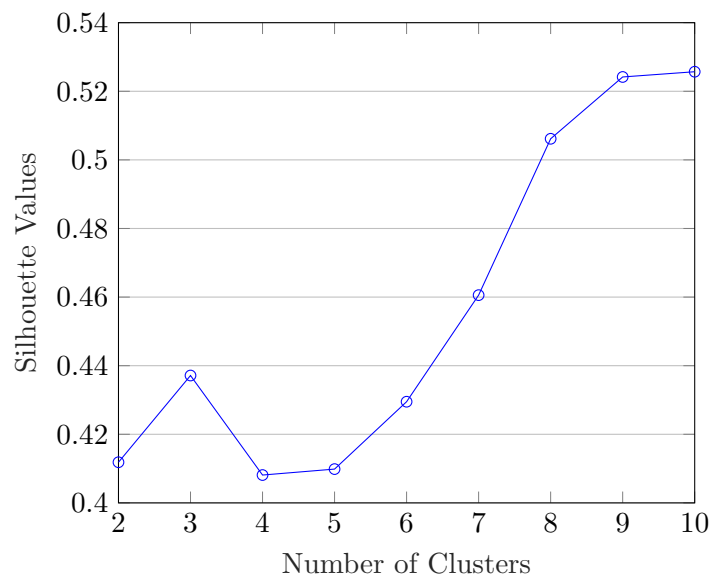


Figure B.10: Silhouette for CX – SDT

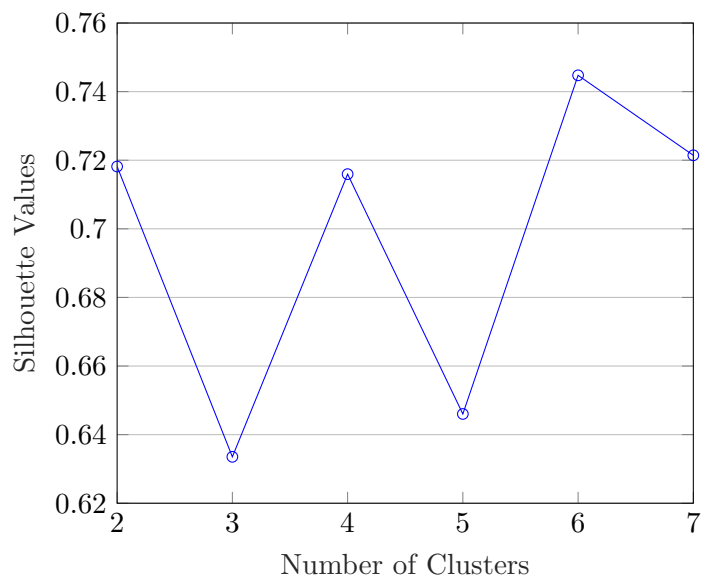


Figure B.11: Silhouette plot for transaction

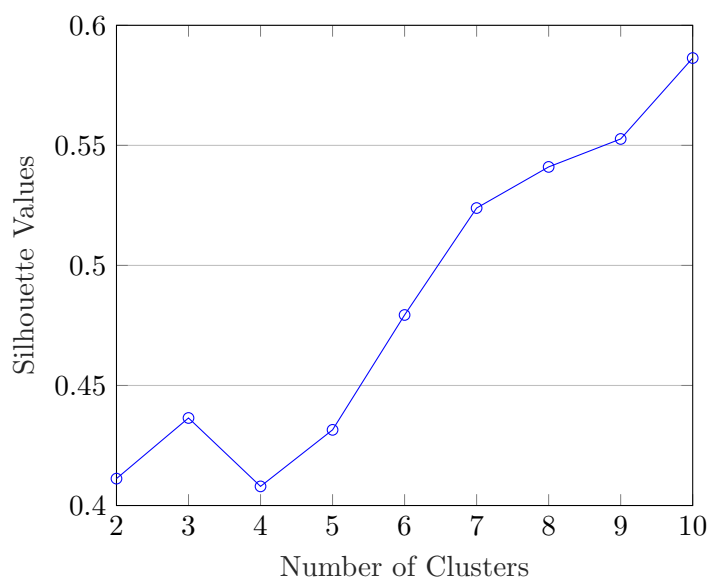


Figure B.12: Silhouette plot for CX – Transactions

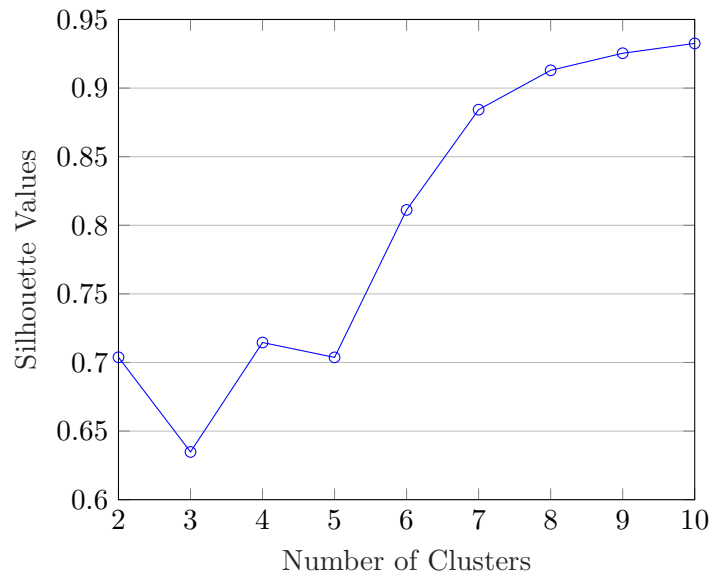


Figure B.13: Silhouette plot for overall CX

- The k-Means clustering applied on the customer behavioural data formed the clusters summarised in the following tables.
 1. The clusters for the customer behaviour for accommodation – Table [B.1](#)
 2. The clusters for the customer behaviour for LDT – Table [B.2](#)
 3. The clusters for the customer behaviour for SDT – Table [B.3](#)
- The k-Means clustering applied on the customer experience data formed the clusters summarised in the following tables.
 1. The clusters for the customer experience for accommodation – Table [B.4](#)
 2. The clusters for the customer behaviour for LDT – Table [B.5](#)
 3. The clusters for the customer behaviour for SDT – Table [B.6](#)
 4. The clusters for the customer behaviour for Transactions – Table [B.7](#)

Table B.1: Clusters of customer behaviour for accommodation

Cluster		1	2	3	4	5	6	7	8
Gender	1	10.7%	13.9%	1%	0%	10.2%	0%	8.2%	5%
	2	0%	2.9%	8%	14.5%	1.2%	11.8%	4.6%	8%
Age Category	1	0%	0%	0%	1.19%	0.66%	1.26%	1.35%	1.69%
	2	0%	1.83%	0%	2.73%	2.15%	2.8%	2.74%	3.3%
	3	0%	1.95%	0%	2.34%	2.1%	2.51%	1.92%	2.66%
	4	1.14%	2.08%	0%	2.37%	1.74%	2.77%	1.38%	2.32%
	5	1.63%	1.57%	0%	1.72%	1.07%	1.68%	1.13%	0.87%
	6	2.56%	1.68%	1.05%	1.57%	0.62%	0.65%	1.18%	0.87%
	7	1.78%	1.11%	1.41%	1.35%	0.53%	0%	0.94%	0.48%
	8	1.99%	1.26%	1.68%	1.34%	0.65%	0%	0.83%	0.61%
	9	1.16%	1.81%	1.1%	0%	0.74%	0%	0.61%	0%
	10	1.55%	1.69%	1.1%	0%	0.5%	0%	0%	0%
	11	0.81%	0.71%	0.66%	0%	0%	0%	0%	0%
	12	1.08%	0.84%	0.85%	0%	0%	0%	0%	0%
Province	1	0.85%	1.30%	1.8%	0.98%	0%	0.74%	0%	0%
	2	0.51%	0.53%	0.67%	0.59%	0%	0.33%	0%	0%
	3	5.64%	9.79%	4.06%	7.53%	3.67%	5.63%	0.31%	0.24%
	4	3.47%	4.83%	2.41%	4.8%	2.95%	4.29%	1.53%	0.5%
	5	0%	0.11%	0%	0.22%	0.14%	0.22%	0.11%	0.30%
	6	0%	0.12%	0%	0.48%	0.45%	0.58%	0.44%	0.19%
	7	0%	0.13%	0%	0%	0.58%	0%	1.30%	1.14%
	8	0%	0%	0%	0%	0.11%	0%	0.25%	0.22%
	9	0%	0%	0%	0%	3.55%	0%	8.84%	10.57%
Type	1	0%	0.89%	0%	0%	0%	0%	0%	0%
	2	0%	8.84%	0%	7%	0%	0%	7.48%	0%
	3	0%	7.90%	0%	7.1%	0%	0%	5.39%	1.17%
	4	4.38%	0%	2.87%	0%	1.76%	4.07%	0%	4.47%
	5	5.84%	0%	5.91%	0%	10%	7.72%	0%	7.24%

Table B.2: Clusters of customer behaviour for LDT

Cluster		1	2	3	4	5	6	7
Gender	1	16.6%	0.0%	1.5%	8.5%	0.0%	11.6%	10.7%
	2	2.7%	14.9%	6.6%	6.6%	17.3%	2%	1%
Age Category	1	0%	0%	0%	2.27%	1.36%	0%	0.91%
	2	0%	3.03%	0%	3.63%	2.99%	0.42%	2.58%
	3	0%	2.63%	0%	2.30%	2.7%	1.67%	1.86%
	4	3.32%	2.80%	0%	1.98%	2.77%	1.34%	2%
	5	2.64%	2.23%	0%	1.27%	1.81%	1.56%	0.76%
	6	2.73%	2.11%	1.28%	1.20%	1.72%	1.83%	0.7%
	7	2.09%	1.94%	1.09%	0.91%	1.47%	0%	0.48%
	8	2.33%	2.16%	0.99%	0.99%	1.49%	1.13%	0.66%
	9	1.73%	1.48%	0.0%	0.69%	1.10%	0.71%	0.79%
	10	3.04%	0.00%	0.0%	0%	0.0%	1.66%	0%
	11	1.96%	0.00%	0.00%	0%	0%	0.92%	0%
	12	2.56%	0.00%	0.00%	0%	0%	1.26%	0%
Province	1	1.07%	1.83%	0%	0%	0.74%	2.30%	0%
	2	0.66%	1.16%	0%	0%	0.38%	0.96%	0%
	3	11.51%	6.52%	0%	0.71%	9.96%	6.48%	1.68%
	4	5.71%	5.34%	0%	1.44%	5.74%	3.66%	2.92%
	5	0%	0.32%	0.01%	0.06%	0.14%	0.10%	0.22%
	6	0%	0.18%	0.50%	0.44%	0.45%	0.01%	0.55%
	7	0%	0%	1.23%	0.55%	0%	0.12%	1.28%
	8	0%	0%	0.28%	0%	0%	0%	0.31%
	9	0%	0%	5.79%	11.91%	0%	0%	4.78%
Type	1	19.24%	0%	0%	15.1%	17.4%	0%	0%
	2	0%	15.11%	7.81%	0%	0%	13.62%	11.72%

Table B.3: Clusters of customer behaviour for SDT

Cluster		1	2	3	4	5	6	7	8
Gender	1	0%	13.89%	4.74%	16.24%	0.75%	11.12%	1.17%	5.65%
	2	13.45%	8.84%	0%	0%	6.29%	0.17%	12.12%	5.57%
Age Category	1	1.38%	1.07%	1.31%	0%	0%	0%	2.25%	0.39%
	2	3.19%	2.79%	2.36%	2.93%	0.12%	0%	2.48%	1.97%
	3	2.83%	2.08%	2.22%	3.18%	0.12%	0%	2.1%	1.73%
	4	3.07%	2.06%	1.91%	2.96%	0.18%	0%	2.15%	1.42%
	5	1.88%	1.34%	1.32%	1.31%	0.08%	1.93%	1.5%	0.14%
	6	0.7%	1.37%	1.21%	1.59%	0.38%	2.04%	1.33%	1.08%
	7	0.4%	0.9%	0.93%	0.95%	0.55%	1.47%	1.06%	0.96%
	8	0%	1.02%	1.04%	1.19%	1.57%	1.60%	0.42%	1.16%
	9	0%	0.66%	0.47%	0.66%	1.31%	1.32%	0%	0.68%
	10	0%	0.29%	0.31%	0.74%	1.16%	1.24%	0%	0.84%
	11	0%	0.19%	0.18%	0.32%	0.77%	0.59%	0%	0.4%
	12	0%	0.17%	0.21%	0.47%	0.78%	1.11%	0%	0.46%
Province	1	0.77%	0%	0%	1.16%	1.38%	0.51%	0.72%	1.11%
	2	0.39%	0%	0%	0.43%	0.6%	0.23%	0.47%	0.47%
	3	6.44%	0%	0%	9.12%	3.26%	5.25%	6.98%	5.6%
	4	5.31%	0.61%	0.35%	5.19%	1.81%	4.03%	4.04%	3.76%
	5	0.16%	0.14%	0.05%	0.11%	0%	0.08%	0.23%	0.07%
	6	0.38%	0.57%	0.51%	0.13%	0%	0.18%	0.48%	0.11%
	7	0%	1.2%	1.17%	0.16%	0%	0.22%	0.37%	0.09%
	8	0%	0.33%	0.24%	0%	0%	0%	0%	0%
	9	0%	11.1%	11.15%	0%	0%	0.78%	0%	0%
Type	1	0%	9.54%	0%	13.39%	3.49%	0%	13.08%	0%
	2	6.28%	4.38%	4%	2.91%	3.55%	4.5%	0.21%	0.0%
	3	7.17%	0%	9.45%	0%	0%	6.79%	0%	11.19%

Table B.4: Clusters of CX for accommodation

Cluster		1	2	3	4	5	6	7	8
Gender	1	0%	0%	0%	0%	11.81%	10.26%	13.46%	13.48%
	2	7.49%	7.41%	18.45%	17.64%	0%	0%	0%	0%
Age Category	1	0.86%	0.17%	0.37%	1.9%	0.41%	0%	2%	0.86%
	2	2.1%	0.41%	0.91%	4.41%	2.43%	0%	3.17%	2.28%
	3	1.3%	1.01%	2%	2.7%	1.5%	0.67%	2.95%	2.14%
	4	1.08%	0.86%	2.17%	2.79%	1.58%	0.78%	2.39%	2.16%
	5	0.73%	0.69%	1.66%	1.71%	0.86%	0.91%	1.72%	1.34%
	6	0.8%	0.78%	1.55%	1.61%	1.36%	1.56%	0.78%	1.23%
	7	0.26%	0.89%	1.34%	1.43%	0.77%	1.11%	0.45%	0.93%
	8	0.22%	0.88%	2.37%	0.66%	1.3%	1.36%	0%	0.96%
	9	0.15%	0.55%	1.58%	0.42%	0.78%	0.82%	0%	0.76%
	10	0%	0.5%	1.96%	0%	0.45%	1.33%	0%	0.41%
	11	0%	0.33%	1.15%	0%	0.19%	0.56%	0%	0.22%
	12	0%	0.34%	1.40%	0%	0.17%	1.15%	0%	0.21%
Province	1	0%	0%	1.52%	1.34%	1.44%	0.69%	0.87%	0.00%
	2	0%	0%	0.82%	0.69%	0.46%	0.29%	0.51%	0%
	3	0%	0%	9.31%	9.31%	6.25%	4.91%	7.06%	0%
	4	0%	0%	6.55%	6%	3.62%	3.8%	4.44%	0%
	5	0%	0%	0.26%	0.28%	0%	0.12%	0.19%	0%
	6	0.68%	0.46%	0%	0%	0%	0.21%	0.39%	0.71%
	7	0.91%	0.79%	0%	0%	0%	0.17%	0.11%	1.3%
	8	0.15%	0.13%	0%	0%	0%	0%	0%	0.31%
	9	5.74%	6.02%	0%	0%	0%	0%	0%	11.22%
CX Rating	1	0%	2.42%	5.69%	0%	1.35%	4.47%	0.0%	2%
	2	0.95%	3.16%	7.91%	2.1%	4.99%	4.21%	0.59%	3.82%
	3	0.42%	0.39%	1.09%	0.97%	0.8%	0.49%	0.69%	0.76%
	4	3.2%	1.18%	2.78%	7.72%	3.07%	1.00%	6.23%	4.09%
	5	2.92%	0.25%	0.99%	6.85%	1.59%	0.09%	5.96%	2.82%

Table B.5: Clusters of CX for LDT

Cluster		1	2	3	4	5	6	7	8
Gender	1	0%	13.85%	11.26%	16.2%	5.58%	1.42%	0.6%	0.00%
	2	12%	1.4%	0%	1.05%	10.29%	6.33%	7.76%	12.23%
Age Category	1	0.35%	0.98%	1.32%	0%	0.88%	0.3%	1.14%	1.44%
	2	1.48%	2.29%	3.43%	0.15%	2.16%	0.84%	2.28%	3.31%
	3	1.29%	2.15%	2.77%	0.18%	1.92%	1.14%	2%	2.79%
	4	0.55%	1.82%	2.25%	0.5%	2%	1.84%	2%	2.69%
	5	0.42%	1.11%	1.73%	0.35%	1.58%	1.45%	1.34%	1.61%
	6	1.37%	1.33%	0.8%	0.87%	1.51%	1.62%	1.22%	0.93%
	7	0.91%	1.07%	0.46%	1.04%	1.28%	1.02%	0.82%	0.53%
	8	1.30%	1.09%	0.23%	1.29%	1.54%	1.26%	1%	0.34%
	9	0.85%	0.69%	0%	1.02%	1%	0.76%	0.74%	0%
	10	0.85%	0.5%	0%	0.97%	1.04%	0.96%	0.33%	0%
	11	0.36%	0.28%	0%	0.58%	0.65%	0.49%	0.12%	0%
	12	0.64%	0.21%	0%	0.7%	0.74%	0.72%	0.18%	0%
Province	1	1.24%	0%	0.84%	1.66%	0.68%	0.76%	0%	0.6%
	2	0.56%	0%	0.37%	0.44%	0.69%	0.35%	0%	0.34%
	3	5.82%	0%	6.51%	3.66%	8.18%	5.97%	0%	6.66%
	4	2.78%	0%	4.42%	1.92%	6.05%	4.58%	0.24%	5.02%
	5	0%	0%	0.16%	0%	0.25%	0.1%	0.1%	0.21%
	6	0%	0.37%	0.47%	0%	0.27%	0.21%	0.59%	0.5%
	7	0%	1.32%	0.22%	0%	0.18%	0.18%	1.11%	0.22%
	8	0%	0.28%	0%	0%	0%	0%	0.29%	0%
	9	0%	11.54%	0%	0%	0%	0.24%	10.85%	0%
CX Rating	1	0.00%	1.96%	0%	0%	5.88%	5.63%	1.84%	0%
	2	3.01%	3.79%	0%	0.58%	9.33%	6.35%	3.81%	0%
	3	1.84%	1.25%	0.82%	0.94%	1.06%	0.39%	1.21%	1.11%
	4	4.04%	3.96%	5.97%	4.3%	0%	0%	3.77%	6.31%
	5	1.5%	2.58%	6.20%	1.86%	0%	0%	2.56%	6.13%

Table B.6: Clusters of CX for SDT

Cluster		1	2	3	4	5	6	7	8
Gender	1	10.42%	0%	12.98%	0%	0%	12.38%	13.15%	0%
	2	0%	13.56%	0%	7.58%	16.27%	0%	0.05%	13.6%
Age Category	1	0.35%	0.96%	1.34%	0.05%	0.90%	0.31%	1.09%	1.47%
	2	1.47%	2.34%	3.42%	0.14%	2.12%	0.82%	2.31%	3.29%
	3	1.29%	2.15%	2.78%	0.16%	1.9%	1.15%	2.03%	2.80%
	4	0.56%	1.82%	2.27%	0.49%	2.01%	1.84%	2.01%	2.69%
	5	0.42%	1.12%	1.71%	0.35%	1.57%	1.44%	1.34%	1.62%
	6	1.38%	1.34%	0.78%	0.86%	1.53%	1.61%	1.24%	0.92%
	7	0.91%	1.05%	0.46%	1.04%	1.29%	1.02%	0.82%	0.55%
	8	1.31%	1.1%	0.22%	1.26%	1.56%	1.24%	0.99%	0.24%
	9	0.85%	0.71%	0%	0.99%	1.0%	0.75%	0.75%	0%
	10	0.86%	0.49%	0%	0.98%	1.02%	0.96%	0.33%	0%
	11	0.36%	0.27%	0%	0.57%	0.65%	0.49%	0.12%	0%
	12	0.65%	0.2%	0%	0.7%	0.72%	0.73%	0.17%	0%
Province	1	1.27%	0%	0.87%	1.61%	0.69%	0.79%	0%	0.6%
	2	0.46%	0%	0.37%	0.45%	0.59%	0.35%	0%	0.36%
	3	5.81%	0%	6.52%	3.66%	8.17%	5.96%	0%	6.65%
	4	2.8%	0%	4.41%	1.85%	6.03%	4.55%	0.23%	5.04%
	5	0%	0%	0.16%	0%	0.24%	0.1%	0.1%	0.22%
	6	0%	0.39%	0.44%	0%	0.27%	0.21%	0.6%	0.5%
	7	0%	1.31%	0.21%	0%	0.37%	0.2%	1.12%	0.23%
	8	0%	0.28%	0%	0%	0%	0.0%	0.3%	0%
	9	0%	11.56%	0%	0%	0%	0.23%	10.86%	0%
CX Rating	1	0%	1.96%	0%	0%	5.87%	5.64%	1.79%	0%
	2	3%	3.8%	0%	0.56%	9.31%	6.34%	3.79%	0%
	3	1.85%	1.27%	0.81%	0.95%	1.15%	0.38%	1.22%	1.12%
	4	4.05%	3.95%	5.97%	4.27%	0%	0%	3.83%	6.34%
	5	1.52%	2.57%	6.20%	1.81%	0%	0%	2.57%	6.14%

Table B.7: Clusters of CX for transactions

Cluster		1	2	3	4	5	6	7	8
Gender	M	0%	18.19%	0%	0%	17.08%	0%	13.66%	0%
	F	12.78%	0%	11.14%	7.53%	0%	12.22%	0%	7.39%
Age Category	1	1.28%	2%	1.06%	0.88%	0%	0%	0.83%	0%
	2	2.93%	4.69%	2.4%	2.07%	0%	0%	2.32%	0%
	3	2.54%	2.86%	1.42%	1.23%	2.88%	0%	2.17%	0%
	4	2.63%	2.52%	1.42%	1.08%	3.03%	0%	2.17%	0%
	5	1.81%	1.86%	1.18%	0.75%	2.55%	2.18%	1.41%	0.68%
	6	0.7%	2.02%	1%	0.81%	1.69%	2.22%	1.23%	0.66%
	7	0.62%	1.29%	0.86%	0.28%	1%	1.29%	0.89%	0.81%
	8	0.27%	0.59%	1.37%	0.26%	2%	1.42%	1%	0.93%
	9	0%	0.36%	0.85%	0.15%	2%	1.21%	0.76%	1.46%
	10	0%	0%	0.43%	0%	1.93%	1.51%	0%	1.58%
	11	0%	0%	0%	0%	0%	1%	0%	0.74%
	12	0%	0%	0%	0%	0%	1.39%	0%	0.52%
Province	1	0.84%	1.38%	1.07%	0%	1.48%	0.87%	0%	0%
	2	0.51%	0.56%	0.55%	0%	0.59%	0.49%	0%	0%
	3	6.66%	9.47%	6.05%	0%	8.75%	5.99%	0%	0%
	4	4.58%	6.14%	3.37%	0%	5.89%	4.7%	0%	0%
	5	0.19%	0.17%	0.1%	0%	0.12%	0.17%	0%	0%
	6	0%	0.42%	0%	0.66%	0.11%	0%	0.65%	0.42%
	7	0%	0.03%	0%	0.89%	0.15%	0%	1.37%	0.76%
	8	0%	0%	0%	0.12%	0%	0%	0.4%	0.15%
	9	0%	0%	0%	5.83%	0%	0%	11.26%	6.04%
CX Rating	1	0%	0%	0%	0%	5.75%	3.52%	2.25%	4.78%
	2	0%	0%	2.95%	0%	1.99%	5.38%	4.15%	3.64%
	3	0%	2.27%	2.13%	1.04%	4.98%	0.88%	1.73%	0%
	4	6.47%	8.59%	6.03%	3.37%	2.54%	1.73%	2.93%	0%
	5	6.32%	7.32%	0%	3.12%	0.83%	0.42%	2.86%	0%