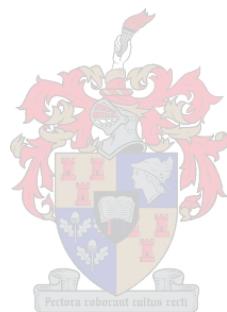


# **Understanding South African Chenin Blanc wine by using data mining techniques applied to published sensory data**

by

**Carlo Cesar Valente**



Thesis presented in partial fulfilment of the requirements for the degree of  
**Master of Science**

at  
**Stellenbosch University**  
Institute of Wine Biotechnology, Faculty of AgriSciences

*Supervisor:* Dr Helene Nieuwoudt  
*Co-supervisor:* Professor Florian Bauer

March 2016

## 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: March 2016

## Summary

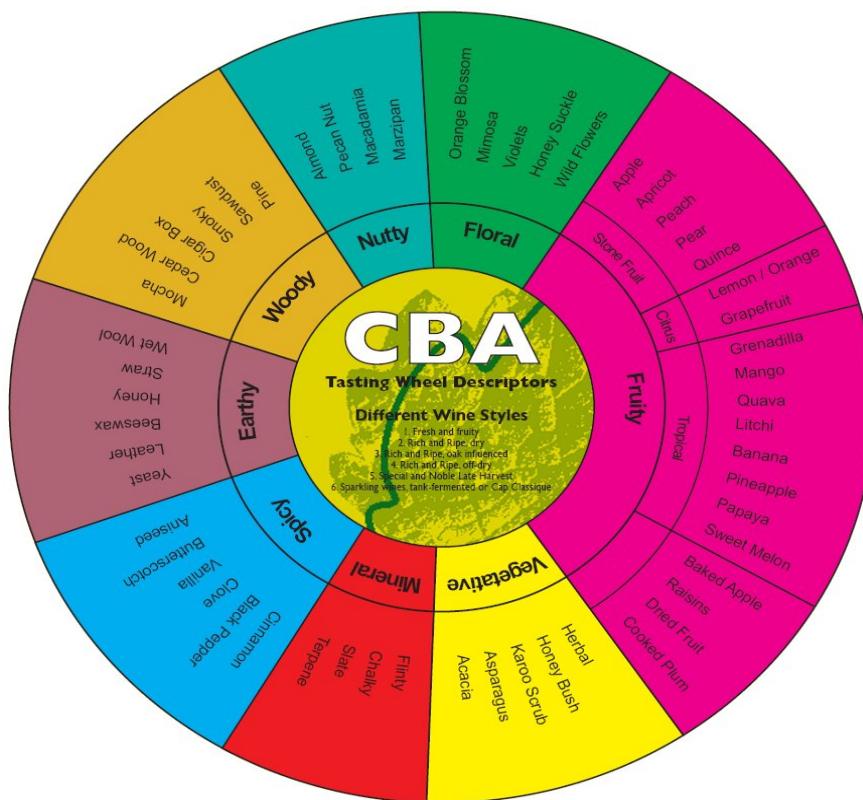
South African Chenin Blanc is the most planted grape cultivar in South Africa (SA) and is known for its versatility in wine sensory profiles. However, according to the South African wine industry, consumers are confused as to the different styles that make up Chenin Blanc wine. Currently, there are six different style classifications for South African Chenin Blanc wine that was proposed as a guideline by the Chenin Blanc Association (CBA). Previous research conducted at the University of Stellenbosch was aimed at evaluating these style classifications.

Previous results showed that, when using a small sample set of commercial Chenin Blanc, only two clear style categories could be identified – Fresh and Fruity and Rich and Ripe Wooded. However, there was a need to re-evaluate these style classifications using a larger sample set. The current study proposed a novel approach of using reputable and published text data on a Chenin Blanc wines as a dataset as opposed to acquiring a limited amount of Chenin Blanc wines for sensory evaluation.

The John Platter Wine Guide is South Africa's largest wine annual and provides a sensory description of each wine. Text data on Chenin Blanc was used over an investigation period of seven years (2007-2014) as it was felt that this could give a clear indication of the sensory properties within the cultivar. It would also provide us with an extremely large dataset to evaluate the current style classification of Chenin Blanc and this was one of the main aims of the study.

Computed text data analysis on 2746 Chenin Blanc wines showed that there were only two clear style categories identified – Fresh And Fruity and a Wooded style, which was conclusive with previous research conducted. Additionally, for the first time the entire sensory spectrum of South African Chenin Blanc was characterised and identifiable "cues", as mentioned in the wine industry, was established. For the purpose of validation, a second white wine cultivar, Sauvignon Blanc, was used to evaluate whether text data analysis on Chenin Blanc was conducted correctly.

Furthermore, the sensorial properties that were retrieved from the dataset for the characterization of the full sensory spectrum were later used for the re-evaluation and validation of the current Chenin Blanc Aroma Wheel. The Chenin Blanc Aroma Wheel was developed in 2005 and since then has never been re-evaluated. There has always been a need to re-evaluate the Aroma Wheel, but there has never been a large enough dataset to do so. This then served as a novel approach as it was the first time that such a task would be attempted. A frequency count of all the words used in the Platter Guide was used in conjunction with sensory attributes that currently made up the aroma wheel. An exclusion/inclusion attribute criterion of five citations was set. Attributes that were currently on the Aroma Wheel that were not cited at least five times were excluded and those that were currently not on the Aroma Wheel and were cited five times or more were included. Additional categories and sub-categories were included to better accommodate attributes that were previously presented in the wrong category. Figure 1 illustrates the current Chenin Blanc Aroma Wheel while figure 2 illustrates the validated Aroma Wheel as proposed in this study. Even though the two Aroma Wheels share similarities, it is evident that, for South African Chenin Blanc, the current aroma wheel is not an accurate representative of the white wine cultivar.



**Figure 1** Current Chenin Blanc Aroma Wheel as proposed by the Chenin Blanc Association (CBA)



**Figure 2** Validated and re-evaluated Chenin Blanc Aroma Wheel based on retrieved data from the John Platter Wine Guide.

## Opsomming

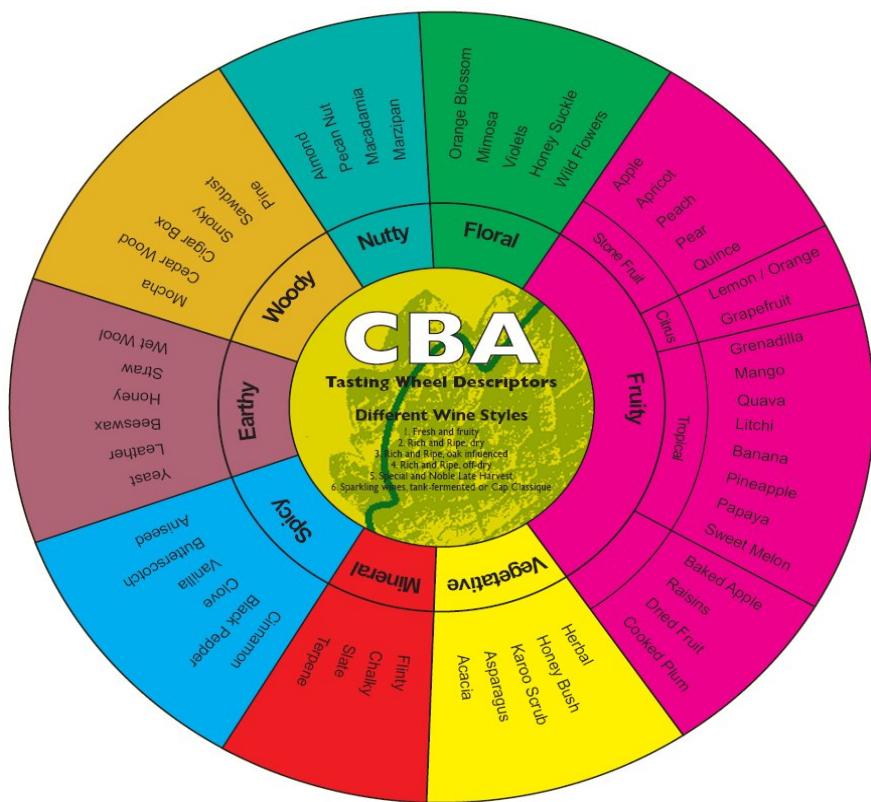
Suid-Afrikaanse Chenin Blanc is die mees volop verboude druifkultivar in Suid Afrika (SA) en is bekend vir veelsydigheid in terme van wyn sensoriese profiele. Nietemin is daar volgens die Suid-Afrikaanse wynindustrie 'n verwarring onder verbruikers rakende die verskillende Chenin Blanc wynstyle. Tans is daar ses verskillende stylklassifikasies vir Suid Afrikaanse Chenin Blanc wyne soos voorgestel in die Chenin Blanc Assosiasie (CBA) se riglyne. Vorige navorsing deur die Universiteit van Stellenbosch was gemik daarop om hierdie stylklasse te evalueer.

Vorige resultate het getoon dat daar binne 'n klein stel monsters van kommersiëële Chenin Blanc wyne slegs twee duidelike stylklasse geïdentifiseer kon word: Vars en Vrugtig sowel as Ryk en Ryp Gehout. Dit was noodsaaklik om hierdie beperking in stylklasse weer te evalueer binne die konteks van 'n groter stel monsters. Die huidige studie stel 'n nuwe aanslag voor waartydens betroubare en gepubliseerde teksdata oor Chenin Blanc wyn as 'n datastel gebruik word in plaas daarvan om 'n beperkte hoeveelheid Chenin Blanc wyne sensoriese te evalueer.

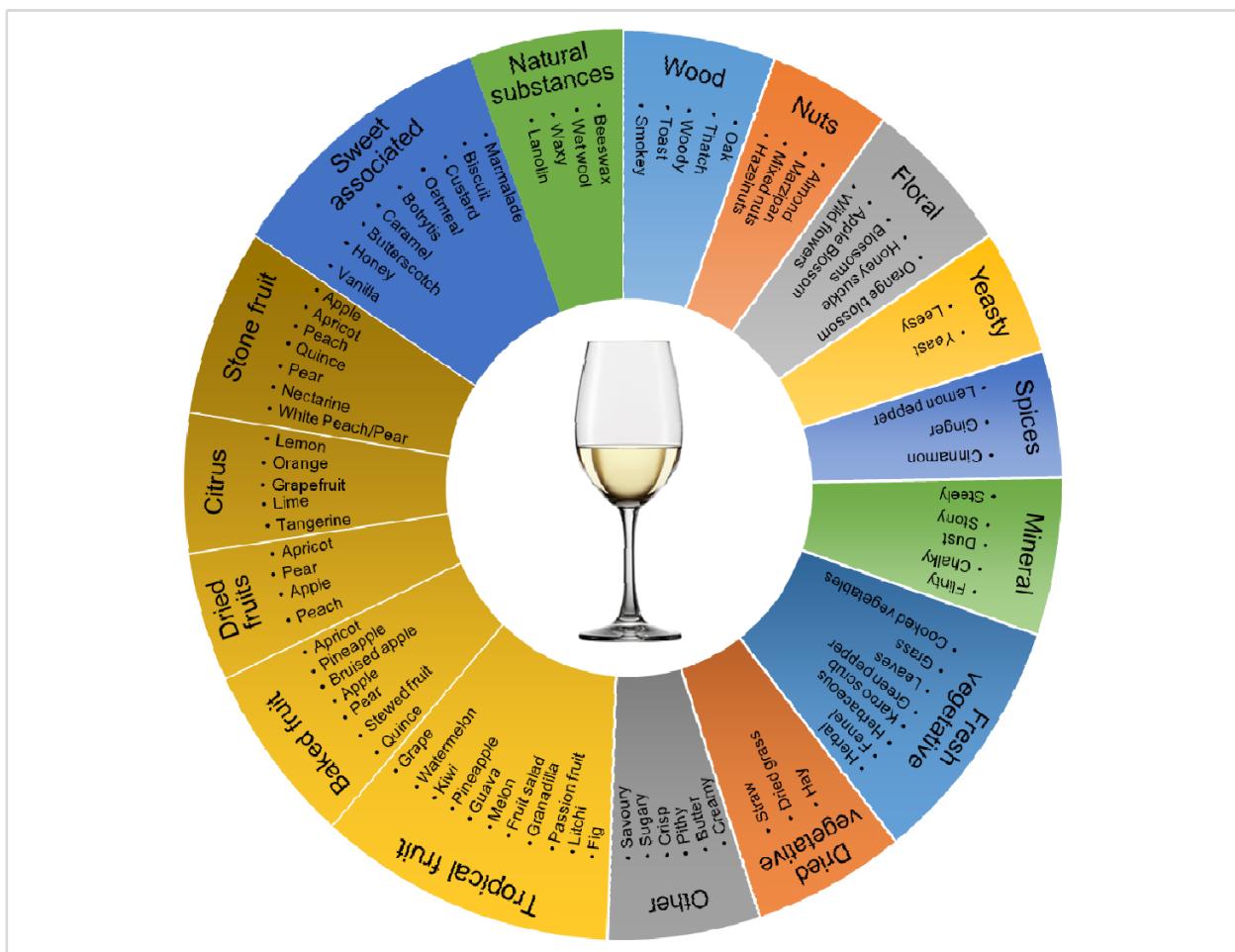
Die John Platter Wyngids is Suid Afrika se grootste jaarlikse wyngids wat 'n sensoriese beskrywing van elke wyn verskaf. Teksdata oor Chenin Blanc was gebruik oor 'n ondersoekperiode van sewe jaar (2007-2014) ten einde 'n duidelike aanduiding van die sensoriese eienskappe wat in Chenin Blanc voorkom te bekom.

Berekende teks data-analise op 2746 Chenin Blanc wyne het getoon dat daar wel slegs twee duidelike stylklasse geïdentifiseer kon word: Vars en Vrugtig sowel as Gehout. Hierdie bevinding is ooreenstemming met vorige bevindinge. Benewens is die hele sensoriese spektrum van Suid-Afrikaanse Chenin Blanc wyne gekarakteriseer en die identifiseerbare geur nuances soos gebruik deur die wynindustrie is bevestig. Ten einde die resultate te bekragtig is 'n tweede witwynkultivar, Sauvignon Blanc, deur dieselfde teks data-analise geanalyseer.

Verder is die sensoriese eienskappe wat verkry is vanuit die datastel vir die karakterisering van die volle sensoriese spektrum weer gebruik om die huidige Chenin Blanc Aroma Wiel te herevalueer en te bekragtig. Die Chenin Blanc Aroma Wiel was sedert dit in 2005 ontwikkel is nog nie weer hersien nie. Ten spyte van die behoefte om die Aroma Wiel te herevalueer kon die taak, weens 'n gebrek aan 'n datastel van voldoende grootte, tot en met nou nog nie verrig word nie. Hierdie studie is die eerste maal dat so 'n taak op hierdie wyse aangepak word. 'n Frekwensietelling van al die woorde wat in die John Platter Gids gebruik is is vergelyk met die sensoriese eienskappe wat teenwoordig is op die huidige Chenin Blanc Aroma Wiel. 'n Uitsluitings/Insluitings eienskapskriteria van ten minste vyf aanhalings was vasgestel. Gevolglik is eienskappe wat nie op die bestaande Aroma Wiel voorkom nie, maar ten minste vyf keer voorkom in die Platter Gids-data, ingesluit in die hersiene Aroma Wiel. Addisionele kategorieë en subkategorieë is ingesluit in die hersiene Aroma Wiel om eienskappe wat voorheen in verkeerde kategorieë geplaas is beter te akkomodeer. Figuur 1 wys die huidige Chenin Blanc Aroma Wiel terwyl Figuur 2 die hersiene Aroma Wiel soos voorgestel in hierdie studie illustreer. Hoewel daar ooreenkoms is tussen die twee Aroma Wiele is dit duidelik dat die huidige Aroma Wiel nie 'n akkurate weerspieëeling is van Suid Afrikaanse Chenin Blanc as 'n druifkultivar is nie.



**Figuur 3** Huidige Chenin Blanc Aroma Wheel soos voorgestel deur die Chenin Blanc Assosiasie (CBA)



**Figure 4** Hersiene en bekragtigde Chenin Blanc Aroma Wheel gebaseer op data vanuit die John Platter Wyngids

This thesis is dedicated to

My friends and family – thank you for all the support and love that you have shown me. I will be eternally grateful. Special recognition needs to be given to the lord for giving me the strength and drive to succeed.

“A dreamer is one who can only find his way by moonlight, and his punishment is that he sees the dawn before the rest of the world” – Oscar Wilde

This goes out to all my fans, without your support none of this could've happened.

## Biographical sketch

Carlo Cesar Valente was born in Cape Town, South Africa on 10 January 1991. He attended Ladybrand Primary School and matriculated at Ladybrand High School in 2008. Carlo obtained a BSc-degree in Human Life Sciences with Psychology in 2012 at the University of Stellenbosch. In 2013 he obtained his Honours degree in Medical Physiology – Reproductive (Andrology) research at the medical campus of the University of Stellenbosch. In 2014 Carlo enrolled of an MSc in Wine Biotechnology at the Institute for Wine Biotechnology, Stellenbosch University.

## Acknowledgements

I wish to express my sincere gratitude and appreciation to the following persons and institutions:

- **Dr Hélène Nieuwoudt**, Institute for Wine Biotechnology, Department of Viticulture and Oenology, Stellenbosch University, South Africa, for being my supervisor
- **Professor Florian Bauer**, Institute for Wine Biotechnology, Department of Viticulture and Oenology, Stellenbosch University, South Africa, for being my co-supervisor
- **Professor Dominique Valentin**, AgroSup Dijon, France, for being my co-supervisor and assisting with guidance and support
- **Ina Smith**, Chenin Blanc Association (CBA), thank you for your guidance and support.
- **Diner's Club – John Platter Wine Guide**, for the electronic database of the platters guide.
- **Professor Martin Kidd** for his assistance with statistical analysis.
- **Mrs Jeanne Brand, Mrs Valeria Panzeri, Ms Carla Weightman and Ms Renée Crous** for all their support and assistance with sensory and consumer work.
- **Pamela and Cesar Valente**, thank you for the support and encouragement.
- **Katia Valente**, thank you for the love and for always being there when I needed to talk to someone.
- **Simone Dames, Horatio Morgan and Mikhayla Williams**, thank you for the laughs and good times. More importantly thank you for always being there when I needed help whether it be with typing in of data or having someone to vent to.
- **My Distell family**, thank you for the support and guidance and allowing me time to complete this journey.
- **Jodie Hermans**, thank you for the love and care that you have shown me over these two years. I appreciate all the encouragement and assistance.
- **Ms Karin Vergeer**, Institute for Wine Biotechnology, Department of Viticulture and Oenology, Stellenbosch University, South Africa, for her assistance in the finalization of this thesis.

## Preface

This thesis is presented as a compilation of 5 chapters.

**Chapter 1 General Introduction and project aims**

**Chapter 2 Literature review**

The importance of wine sensory evaluation in modern wine industries

**Chapter 3 Research results**

Correspondence analyses based data mining approach to investigate and model the sensor characteristics of South African Chenin Blanc

**Chapter 4 Research results**

Validation and expansion of the South African Chenin Blanc aroma wheel to accommodate the sensory space and stylistic spectrum of South African Chenin Blanc wine

**Chapter 5 General discussion and conclusions**

## Table of Contents

<b>Chapter 1. General introduction and project aims</b>	<b>1</b>
1.1 Introduction	1
1.2 Project aims	2
1.3 Literature cited	3
<b>Chapter 2. Literature review: The importance of wine sensory evaluation in modern wine industries</b>	<b>5</b>
2.1 Wine beverage industry	6
2.1.1 Wine industry	6
2.2 South African wine producing provinces	7
2.3 Wine of Origin (WO) scheme	7
2.4 White wine cultivars in South Africa	8
2.5 Chenin Blanc	9
2.5.1 Background	9
2.5.2 SA Chenin Blanc Styles	9
2.6 Aroma wheel	10
2.7 Sensory perception	13
2.8 Application of sensory data for the wine industry	13
2.8.1 Identifying wine faults	13
2.8.2 Typicality	14
2.8.3 Regionality	15
2.8.4 Quality characterization	15
2.9 Data mining	16
2.9.1 Data mining evolution	16
2.9.2 Data mining application	16
2.9.3 Data mining task order and processing	16
2.10 General conclusiono and future impact	17
2.11 Literature cited	18
<b>Chapter 3. Research results: Correspondence analyses based data mining approach to investigate and model the sensory characteristics of South African Chenin Blanc</b>	<b>22</b>
3.1 Introduction	22
3.2 Materials and Methods	24
3.3 Results and Discussion	28
3.4 Conclusion	40
3.5 Literature cited	41

---

<b>Chapter 4. Research results: Validation and expansion of the South African Chenin Blanc Aroma Wheel to accommodate the sensory space and stylistic spectrum of South African Chenin Blanc wine</b>	<b>45</b>
4.1 Introduction	46
4.2 Materials and Methods	47
4.2.1 Current Chenin Blanc Aroma Wheel design	47
4.2.2 Data source and structure	48
4.2.3 Data processing	49
4.2.4 Data analysis	49
4.3 Results	49
4.4 Discussion	52
4.5 Conclusions	56
4.6 Literature cited	58
<b>Chapter 5. General discussion and conclusion</b>	<b>59</b>
5.1 General discussion and conclusions	60
5.2 Literature cited	63

---

**ADDENDUM A**

# Chapter 1

---

## General introduction and project aims

# 1. GENERAL INTRODUCTION AND PROJECT AIMS

## 1.1 INTRODUCTION

The South African wine industry aims to focus on producing wine of good quality (Foxcroft, 2009). For historical reasons, Chenin Blanc is the most planted grape cultivar in South Africa (SA). Chenin Blanc is frequently described as being of neutral character, and lacks distinguishing features such as the significant amounts of methoxypyrazines compounds that typically give Sauvignon Blanc a green-pepper characteristics (Parr *et al.*, 2013)

In SA, similar to countries such as the United States of America (USA) and New Zealand, Chenin Blanc wine (CB) had previously been used as an integral part of blended wines and often used for brandy distillation as referenced by Lawrence (2012) and Bougas (2014). In recent years, South African CB has continuously improved and diversified, and what was previously known as a workhorse cultivar (Bester, 2011) has now been recognised as a cultivar with significant potential to produce quality wines. However, unlike other white grape cultivars such as Sauvignon Blanc with its green and tropical poles (Parr *et al.*, 2010; Hoffmann *et al.*, 2014) and Chardonnay with its buttery aromas (Jaffré *et al.*, 2011), the sensorial perception and full sensory spectrum of CB is less well characterised. However, this situation presents an opportunity to explore new and improve existing wine styles. One of the main challenges involved in better characterising the sensory spectrum or sensory space of CB is the generation of a large enough dataset. The dataset would need to contain extensive or sufficient sensorial information on CB with the addition of extrinsic factors such as region and vintage to allow for a full representative and in-depth analysis of SA Chenin Blanc wines.

The Chenin Blanc Aroma Wheel which was put forward by the Chenin Blanc Association (CBA) in 2005 is a well-known and frequently used visual representation of the sensory attributes of this white wine cultivar. The aroma wheel attempts to reflect a holistic view of the full sensory spectrum of CB. However, the aroma wheel was not developed using scientific practices and methodologies and there has been no scientific validation of the aroma wheel.

The CBA also provided guidelines on style classification based on the amount of residual sugar. There are six different style categories that include Fresh and Fruity, Rich and Ripe unwooded, Rich and Ripe wooded, Rich and Ripe slightly sweet, Sweet and Sparkling wines. Existing research on South African Chenin Blanc has, in particular, been focused on evaluating the different style classifications within CB while also investigating the correlation between chemical compounds and sensory attributes (Antwerpen, 2012; Bester, 2011; Hanekom, 2012). Results from these studies showed that, when evaluating dry and semi-dry wines, only two clear style categories, rich and ripe wooded and a style combining fresh and fruity and rich and ripe Unwooded, were reproducibly identifiable (Bester, 2011). Such data highlight the need for an in-depth investigation into the full style spectrum of CB. The above-mentioned research was however limited to relatively small sample sets, and in total some 150 wines were tested over a period of

---

three years. This comprises a small fraction of the, on average, 390 Chenin Blanc wines that are included annually in the John Platter Wine Guide. The current research project therefore attempted a different approach to obtain a holistic picture of the full Chenin Blanc wine style spectrum. Instead of selecting a small sample set consisting of commercial wines, South Africa's largest "sample/data set", the John Platter Wine Guide, was used to evaluate the characterisation of the vast majority of all single varietal Chenin Blanc wines in SA that were produced during a seven year period (2008-2014). The investigated period between 2008-2014 was selected as it this period was marked by an increase in international recognition of South African Chenin Blanc wines.

The John Platter Wine Guide, which is published annually, describes and evaluates the vast majority of all wines in South Africa. The guide provides detailed information on each wine including the region where the wine was produced and a sensory description of each wine. However, the data is not published in a structured manner for each wine resulting in a complex data set that mainly comprises of text data. Due to this complexity, a considerable amount of standardisation was required as the mining of wine sensory data has never been attempted on such a large scale.

Statistical methods such as Multidimensional scaling (MDS), Cluster analysis and Correspondence analysis (CA) were computed in order to get a visual representation of the full Chenin Blanc wine spectrum. However, the dataset first had to be standardised before any statistical analysis could be done (Figure 1). The results provided novel insights into the sensory descriptors that are used for South African Chenin Blanc due to the magnitude of the dataset and also provided insights into the correlations between different sensorial profiles. The results also provide some support for the concept of regionality and typicity within the wine producing regions in SA. The study aims to establish a framework for further research with regards to data processing and functionality, and for future expansion to other cultivars. The study is based on the novel holistic approach of analysing a dataset that was not compiled using conventional structured data capturing methodologies e.g. rapid sensory methods, and investigating whether or not any significant information could be derived from the data mining of the data.

This study forms part of a larger Chenin Blanc research program at the Institute of Wine Biotechnology (IWBT) and the department of Viticulture and Oenology (IWBT-DVO), University of Stellenbosch, in collaboration with the Chenin Blanc Association (CBA).

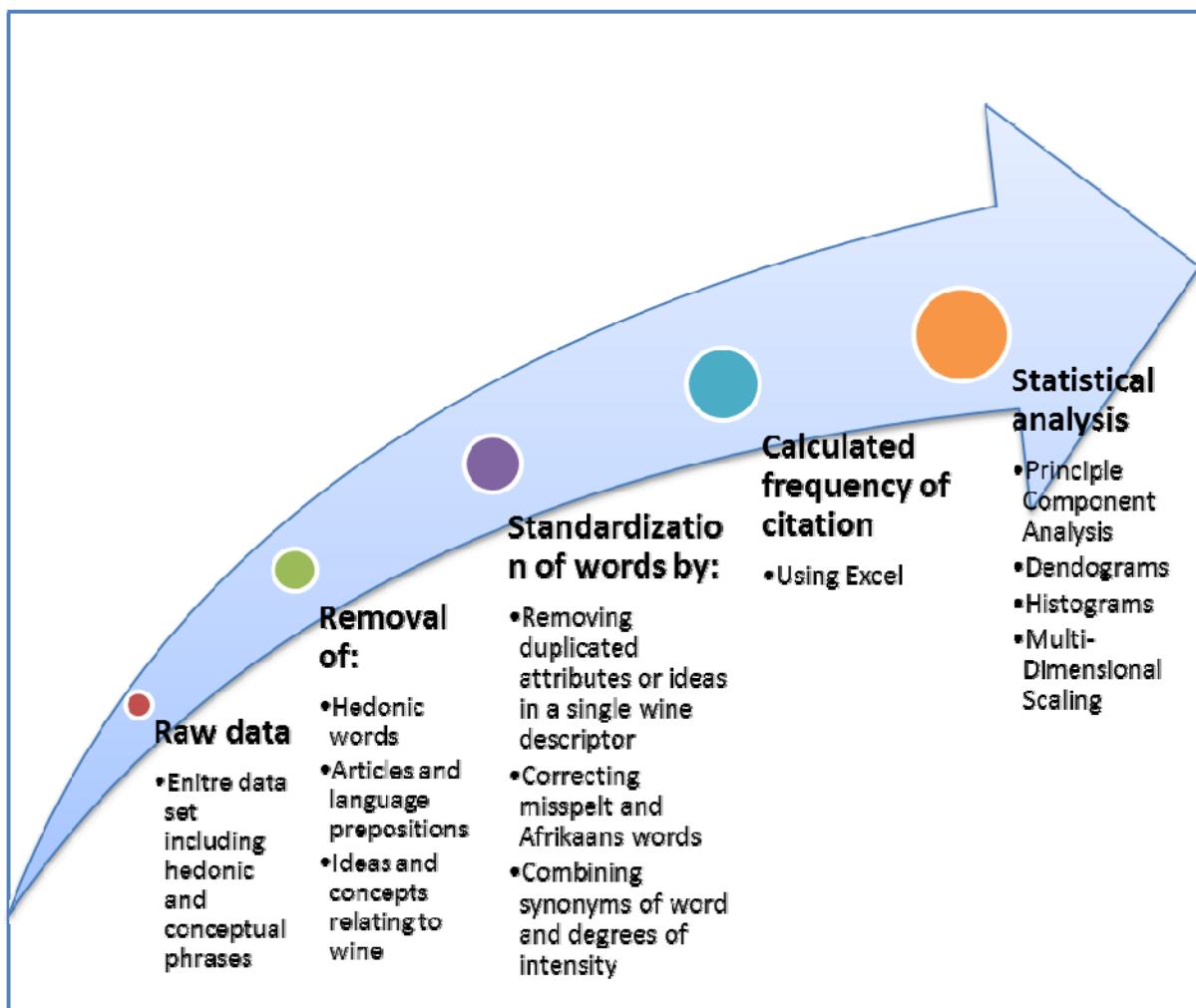
## **1.2 PROJECT AIMS**

---

The specific aims of the project were:

- 1) Model and characterize the sensory space of Chenin Blanc to identify possible relationships between sensory descriptors and regions using a correspondence analysis-based data mining approach

- 2) Obtain a list/profile of the most cited sensory attributes and explore the possibility of identifying industry specific “sensorial cues” that could be transferred in order to accurately characterise stylistic changes in the South African cultivar
- 3) Validate and elaborate on the current South African Chenin Blanc Aroma Wheel to accommodate the full style spectrum of South African Chenin Blanc wine



**Figure 1.1** Data mining protocol used before data analysis was done on the John Platter wine guide over the 7 year investigation period.

### 1.3 LITERATURE CITED

- Antwerpen L Van. (2012). Chemical and Sensory profiling of dry and semi-dry South African Chenin blanc wines.
- Bester I. (2011). Classifying South African Chenin blanc wine styles [Internet]. [cited 2014 Aug 13]. Available from: <https://ir1.sun.ac.za/handle/10019.1/18048>

- Foxcroft M. (2009). Growing the consumption of wine amongst emerging market consumers in South Africa. *Cape Wine Diss January* [Internet]. [cited 2014 Oct 21];(January). Available from: [http://www.capewineacademy.co.za/dissertations/capewinemaster\\_foxcroft\\_growing\\_wine\\_consumption.pdf](http://www.capewineacademy.co.za/dissertations/capewinemaster_foxcroft_growing_wine_consumption.pdf)
- Hanekom E. (2012). Chemical , sensory and consumer profiling of a selection of South African Chenin blanc wines produced from bush vines.
- Hoffmann M, Varela P, Ares G. (2014). Novel Techniques in Sensory Characterization Profiling and Consumer.
- Jaffré J, Valentin D, Meunier JM, Siliani A, Bertuccioli M, Le Fur Y. (2011). The Chardonnay wine olfactory concept revisited: A stable core of volatile compounds, and fuzzy boundaries. *Food Res Int* [Internet]. Elsevier Ltd **44**(1):456–64. Available from: <http://dx.doi.org/10.1016/j.foodres.2010.09.022>
- Parr W V., Schlich P, Theobald JC, Harsch MJ. (2013). Association of selected vinicultural factors with sensory and chemical characteristics of New Zealand Sauvignon blanc wines. *Food Res Int* [Internet]. Elsevier Ltd **53**(1):464–75. Available from: <http://dx.doi.org/10.1016/j.foodres.2013.05.028>
- Parr W V., Valentin D, Green J a., Dacremont C. (2010). Evaluation of French and New Zealand Sauvignon wines by experienced French wine assessors. *Food Qual Prefer.* **21**(1):56–64.
- Weightman CJ. (2014). Characterization of Chenin blanc wines produced by natural fermentation and skin contact : focus on application of rapid sensory profiling methods by.

# Chapter 2

---

## LITERATURE REVIEW

**The importance of wine sensory evaluation in modern wine industries**

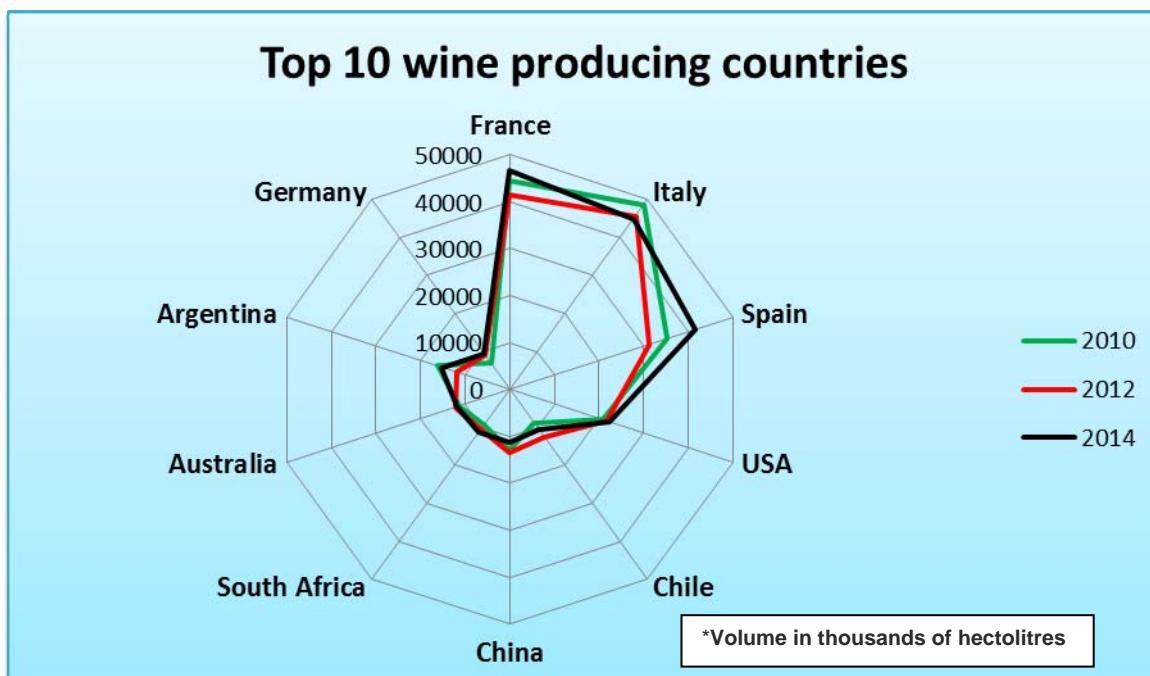
## 2. LITERATURE REVIEW: THE IMPORTANCE OF WINE SENSORY EVALUATION IN MODERN WINE INDUSTRIES

### 2.1 WINE BEVERAGE INDUSTRY

#### 2.1.1 Wine industry

Wine has been consumed by human societies for thousands of years and the earliest archaeological evidence of wine production dates back to as early as 6000BC (Q. D. Truong, 2011). Wine, was once viewed as luxury good, but according to Cortez et al (2009), since the mid 2000's, the social beverage has been consumed by a much wider range of consumers.

South Africa (SA) was previously not associated with being amongst the leading global wine producing nations, but since the late 2000's this notion has changed (Foxcroft, 2009; Bester, 2011) with the country, in 2014, being one of the largest wine producing countries in the Southern hemisphere. European nations such as Italy and France have a well established reputation for producing large volumes of wine with the French, in particular, being well known for the quality of wine they produce in addition to the high volumes. Figure 2.1 shows the top 10 wine producing countries over a five year period (in volume) and even though the production values are not as high as Italy, Spain and France, SA is establishing itself as an emerging entity in the global wine industry.



**Figure 2.1:** Graph illustrating the top 10 wine producing countries in the world from 2010-2014

The wine industry in SA is not only responsible for wine production, but also has a pivotal role in brandy production (Landau, 2011). Assisting the cause is just under 100 000 hectares of vines producing wine grapes that stretch over at least 800 km in length (SAWIS, 2014) and according to Bester (2011) this has helped the South African wine industry develop into one of the new world wine producing nations.

## **2.2 SOUTH AFRICAN WINE PRODUCING PROVINCES**

---

South Africa is divided into nine provinces, although not all the provinces are acclimatised to cultivate grapes that can be used for wine production - due to extreme heat conditions and a lack of rain fall. The Western Cape, which is situated on the South Western tip of Africa, is ideal due to the cooler mediterranean climate and optimal terrior portrayed in the region (Fisheries, 2013). However, there is an abundant emergence of new wineries in these areas due to the mediterranean climate, coupled with a copious amount of mountains, slopes and valleys. The climate together with the environmental conditions make the production of wine in the Western Cape increasingly unpredictable (Bester, 2011) and this has resulted in a particularly hostile and competitive wine industry. The competitive nature of the South African wine industry therefore encourages wine makers to invest in and employ new technological techniques that could give them a competitive advantage in an already overcrowded market environment.

## **2.3 WINE OF ORIGIN (WO) SCHEME**

---

Regarded as one of the most beautiful wine producing areas in the world, the Western Cape is further divided into regions which are made up of districts that in turn are made up of smaller wards. The Wine of Origin (WO) Scheme was developed by the Wine of South Africa (WOSA) to differentiate between areas in the Western Cape according to their geographical position. Figure 2.2 shows the map of the Western Cape that represents all of the wine producing areas and also indicates into which ward and ultimately which district they belong to. An updated and completed document of the Wine of Origin Scheme is attached as addendum A. In addition to the Western Cape's wine producing regions, the Northern Cape, which has a small wine producing region, can also be seen in figure 2.2.



Figure 2.2: South African Map showing the different wine producing areas including districts and wards

## 2.4 WHITE WINE CULTIVARS IN SOUTH AFRICA

Sauvignon Blanc and Chardonnay have been extensively studied and detailed information on the two white wine cultivars have been well documented ranging from the chemistry to the sensorial attributes associated with each (Spillman *et al.*, 2004; Ballester *et al.*, 2005, 2009; Campo *et al.*, 2008; Parr *et al.*, 2010). Chardonnay is well characterised to portray the toasted and wooded characteristics that can often be perceived as buttery or butterscotch (Q. D. Truong, 2011; Tao *et al.*, 2012) due to the malolactic fermentation it undergoes. While Sauvignon Blanc is intensely

documented for having two clear flavour spectrums being “green” and “tropical” (Parr *et al.*, 2013). However, due to Chenin Blanc’s neutral grape profile and ease at which it can be manipulated, no clear sensorial space or full sensory spectrum is evident and the cultivar is not as accurately characterised as some of the other white wine cultivars.

## 2.5 CHENIN BLANC

---

### 2.5.1 Background

Chenin Blanc acquired its name from Mont-Chenin in the French district of Touraine around approximately the 15th century and can be traced back to Anjou, around A.D. 845 (Stevenson & Sotheby’s (Firm), 2005). Chenin Blanc, also referred to as Steen, which was once regarded as the workhorse cultivar (Bester, 2011) is capable of producing award winning single cultivar wines that are able to compete with the Chenin Blanc from the Loire region in the South of France (Asimov, 2007; Bester, 2011).

Asimov (2007) was of the opinion that South Africa would be a welcome refuge in conjunction with its French counterpart, but was slightly disappointed that even though in 2012 South Africa had twice as many acres of land producing Chenin Blanc grapes than in the Loire, the wine produced lacked the signature floral, mineral and citrus flavours (Asimov, 2012).

According to Goodie (2011), South African Chenin Blanc has a typicity about it (Loubser, 2008). However, the neutral grape cultivar, which in 2008 made up 18.7% of South Africa’s grapes, lacks distinct primary grape flavours and can therefore be manipulated during wine making processes.

### 2.5.2 SA Chenin Blanc Styles

The neutral sensory profile that the Chenin Blanc grapes exhibit makes it particularly susceptible for manipulation in the attempt to create stylistic differences with a wide variety of aromas (Bester, 2011). In many instances this manipulation would be successful and beneficial to the white wine cultivar, but there have been documented cases where this has over-complicated the wine causing it to seem artificial – although this was just his opinion (Asimov, 2012)

As referenced by Bester (2011) there is an increasing confusion amongst consumers as to the different styles within the cultivar. The large variety of attributes that are present within a given bottle could be the main reason for the confusion as some consumers have very little knowledge of the complexity that certain attributes can cause. While the ineffective use of the back label only adds to the confusion.

A significant amount of research has been done with regard to the evaluation of style classification at the University of Stellenbosch and according to Loubser (2008), Bester (2011), Lawrence (2012) and Antwerpen (2012) the various styles of Chenin Blanc can be seen in table 2.1.

**Table 2.1** The different styles in South African Chenin Blanc based on residual sugar levels

Chenin Blanc style	Description
Sweet	> 30 g/L residual sugar
Rich and ripe – Slightly sweet	Between 9 and 30 g/L residual sugar
Rich and ripe – Wooded	Less than 9 g/L residual sugar
Rich and ripe – Unwooded	Less than 9 g/L residual sugar
Fresh and fruity	Less than 9 g/L residual sugar
Sparkling	Tank fermented or Cap Classique

However, studies have shown that for Dry Chenin Blanc (excluding sweet and sparkling) only two distinct styles could be clearly identified by a panel of experts; 1) Fresh and fruity, which was characterised by attributes such as guava, passion fruit, grapefruit, peach and apple, and 2) rich and ripe wooded which was characterised by attributes such as oak, nuts, spice, savoury and vanilla (Antwerpen, 2012).

## 2.6. AROMA WHEEL

---

The confusion exhibited by wine consumers with regards to stylistic classification within Chenin Blanc was a clear indication that not only was it of importance to provide clarification, but there needed to be a visual representation that would characterise the white wine cultivar.

Aroma wheels are used globally to give a visual representation of the sensory spectrum for a given product. The core purpose of an aroma wheel is to provide consumers with the insight to familiarise themselves with the product that can range from alcoholic beverages to food products (Reed, 2010; Association, 2011). The methodology behind the construction of an aroma wheel is based on first acquiring reference standards to align the panel that will be constructing the aroma wheel so that linguistic technicalities can be addressed before processing (Jolly & Hattingh, 2001; Jolley, 2014). Furthermore, there are various different techniques used to assemble a list of attributes that could potentially result in a product specific aroma wheel (Jolley, 2014).

In 2005 the South African Chenin Blanc Association (CBA) undertook the task of formulating an aroma wheel specifically for the white wine cultivar. The idea of constructing the Chenin Blanc Aroma Wheel arose after members of the CBA had a meeting to discuss the short-comings of the cultivar and how they were going to improve the reputation of a historic and market leading varietal.

According to Van Rooyen\* (2015), who was a key member of the team that was tasked with compiling the aroma wheel, Chenin Blanc was similar to other white wine varietals such as

Sauvignon Blanc and Chardonnay except it did not have distinct poles in which it could be differentiated into. This would be problematic as Chardonnay is well documented to be either wooded or unwooded and Sauvignon Blanc is known to have two clear poles being green and tropical. However, the methodological thinking behind constructing the list of attributes for Chenin Blanc was to assemble generic white wine sensory attributes that they felt would accurately represent the full sensory spectrum.

The CBA used three steps in order to compile a list of attributes that could potentially make up the Chenin Blanc Aroma Wheel.

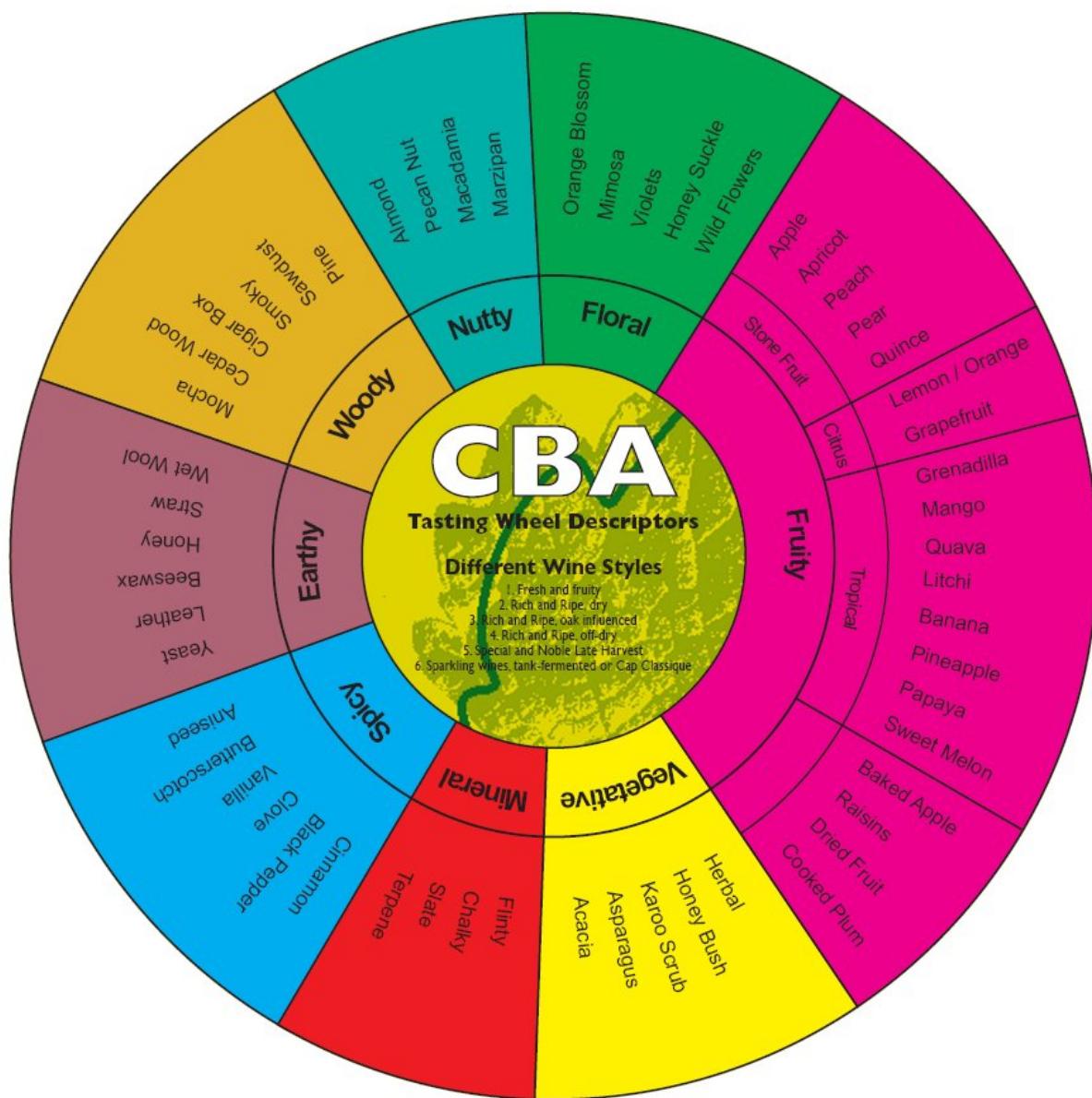
Step 1 was using pre-established aroma wheels of other white varietals and extracting attributes that they thought covered the full sensory spectrum or sensory space of Chenin Blanc.

Step 2 was inviting a group of wine makers and experts that were very familiar with the cultivar to list a few attributes that they thought represented Chenin Blanc.

The final step was taking a pre-established brand aroma wheel and carefully selecting attributes that they thought might represent the sensory space of Chenin Blanc (Jolly & Hattingh, 2001; Van Rooyen, 2015).

An accumulated attribute list of between 90-100 sensory attributes was compiled by the CBA for further processing. These attributes had to then be re-analysed in order to select only the ones that was thought to accurately represent the cultivar. Using literature on the formation of aroma wheels from all over the world the CBA developed an inclusion and exclusion criteria that was implemented to select which attributes would constitute the aroma wheel. The final compiled list of 90-100 attributes were then given to another, larger group of experts that had to individually stipulate whether they thought the attributes should be included or excluded. Once this was completed they then collectively decided on the 56 attributes that would make up the Chenin Blanc Aroma Wheel and allocated each attribute into a family or group that they felt accurately characterised the collective group of attributes e.g. Fruity, Woody, Nutty (figure 2.3).

## Different attributes used to describe the Chenin Blanc and the formation of the Chenin Blanc Aroma Wheel



**Figure 2.3:** Chenin Blanc Aroma Wheel showing the 6 different styles as well as the different sensorial attributes put forward by the Chenin Blanc Association (CBA)

Even though there has been no experimental or scientific design attached to the development of the Chenin Blanc Aroma Wheel, it is widely used both locally and internationally by wine experts and consumers. There is, however, a need for the Chenin Blanc Aroma wheel to be scientifically evaluated and validated and this formed the foundation of chapter 4 of this thesis.

The Chenin Blanc Aroma Wheel as well as other aroma wheels is important when looking at consumer sensory perception as it serves as a means of covering all the sensorial characteristics of a cultivar

\* Personal communication with Jan Van Rooyen (2015)

## 2.7 SENSORY PERCEPTION

---

Sensory perception, according to the gene ontology consortium, occurs in organisms capable of performing neurophysiological processing of the stimuli in their environment, and covers the processes of hearing, vision, taste and smell. Various studies have been conducted on how factors can influence both sensory perception and liking. One such study was conducted by Fernández-vázquez et al (2015) where it was reported that that colour variation in orange juice altered the liking and perceived sweetness, sourness and overall intensity (Fernández-vázquez et al., 2013).

Sensory perception can be influenced by intrinsic and extrinsic factors and this has also shown to influence quality perception (Sáenz-Navajas et al., 2014; Rahman & Reynolds, 2015). Intrinsic factors or cues are those related to the factors that cannot be changed or altered once the final product has been packaged or bottled and this varies in products. For wine, key intrinsic factors include mouthfeel, flavour and aroma, but the divergence in wine quality drivers are not the same for all wine cultivars and are not only driven by intrinsic factor (Sáenz-Navajas et al., 2013, 2014). However, consumers mostly use intrinsic cues when purchasing wine as opposed to extrinsic cues (Rahman & Reynolds, 2015).

Extrinsic factors or cues refer to properties which are not physically part of the wine itself, but make up the external properties of the product ranging from packaging or labelling to price (Sáenz-Navajas et al., 2013). Extrinsic factors also play an important role in quality perception and can be divided into two classes. 1) cues related to the product itself like wine of origin and oak maturation and 2) cues that can be altered without changing the product like price and branding (Chrea et al., 2011). Factors such as wine of origin, bottle weight (Piqueras-Fiszman & Spence, 2012), back label information, label aesthetic and the presence/absence of awards of/in wine all play a key role in sensory perception (Sáenz-Navajas et al., 2014; Tang et al., 2015).

Sensory perception is often used in sensory science to compile datasets by using rapid sensory methods in order to retrieve text or quantitative information (Hoffmann et al., 2014; Jolley, 2014; Weightman, 2014). The information that is retrieved can then be used for industrial and scientific purposes (Garbez et al., 2015).

## 2.8 APPLICATION OF SENSORY DATA FOR THE WINE INDUSTRY

---

### 2.8.1 Identifying wine faults

Sensory perception can reveal certain characteristics in wine that can be seen as positive or negative. The positive aromas in wine can range from fresh tropical to wood complexity – depending on the wine cultivar, while a wine fault is a less desirable or negative odour. Wine fault is an unpleasant defect in wine that occurs during poor wine making practices or storage

conditions (Jackson, 2008). An example of a wine fault is *Brettanomyces bruxellensis* ("Brett Characteristics"). Brett causes a chemical reaction in wine when it reacts with phenolics which produce a very distinct odour. The leather horsey odour is often considered as a fault in wine and is generally perceived as a negative characteristic by both consumers and experts (Lattey *et al.*, 2010; Tempère *et al.*, 2014).

Another common white wine fault is 2,4,6-trichloroanisole (TCA) which is one of many compounds that is recognised by means of sensory perception. The mouldy taint that can often be perceived in wine is more commonly known as cork taint and occurs when chloroanisoles and bromoanisoles such as TCA contaminate the cork or winery materials. Cork taint is the main reason for wine consumer rejection (Cravero *et al.*, 2015) and due to the low sensitivity threshold for TCA, consumers often pick up the mouldy or wet cardboard odour that can also resemble bell peppers (Stevenson & Sotheby's (Firm), 2005; Jackson, 2008). Table 2.2 shows additional faults that can be perceived in wine together with the chemical compound or molecule responsible for the sensory description

**Table 2.2** Chemical compounds and their sensory description perceived as a wine fault

Compound/molecule	Sensory description
Trans-2-Hexanal	Vegetal
Acetaldehyde (ethanal)	Rotten apple
Acetic acid	Vinegar
Ethyl acetate	Glue
Decanoic acid	Soap
SO <sub>2</sub>	Sulphur
H <sub>2</sub> S	Rotten egg
Ethanethiol (mercaptan)	Onion
Dimethyl sulphide	Cauliflower
Ethyl-4-phenol	Horse
2-Ethyl-fenchol	Mouldy-earthly
Trichloroanisole (TCA)	Cork

## 2.8.2 Typicality

Typicality refers to the concept that a grape cultivar or region can be characterised by a set group of variables/descriptors or when a particular wine accurately represents its category or concept (Parr *et al.*, 2010, 2013). Sensory perception plays a key role in the validation of wine cultivars as several studies have shown that some cultivars are characterised by one or more groups of attributes or styles (Ballester *et al.*, 2005).

Wine or cultivar typicality reflects both its origin and varietal purity as it forms a template for which other wines are measured against in terms of quality perception (Parr *et al.*, 2010; Jaffré *et al.*, 2011). This becomes increasingly meaningful as brand typicality becomes significant in

consumer purchase behaviour and liking, especially when consumers are not knowledgeable about extrinsic factors such as wine of origin (Tang *et al.*, 2015)

### 2.8.3 Regionality

Regionality, which is a niche term in wine sensory, refers to the reputation a wine region has for producing wines with a particular style (Easingwood *et al.*, 2011). Regionality has emerged as one of the main drivers for brand marketing of wines in countries such as Australia and New Zealand due to its beneficial impact in other countries (Easingwood *et al.*, 2011). This follows the success of regions having iconic status in terms of wine quality due to, amongst others, the region in which they reside. Bordeaux, the Napa Valley as well as the Barossa Valley are all examples of regions that are able to stand as a brand on their own due to the typicality they possess in terms of regionality (Johnson & Bruwer, 2007).

Consumer acceptance of regional wine is based on the sensorial characteristics of a given wine. Both consumer purchase behaviour and consumer liking has been shown to be positively correlated to wine of origin (Sáenz-Navajas *et al.*, 2014). The authenticity associated with regionality plays an important role in consumer liking, as studies have shown that consumers are far less interested in terror difference as opposed to the region the terror is in (Mouton, 2006).

### 2.8.4 Quality characterization

Sensory perception of wine has been shown to be positively correlated to wine quality. Different sensorial characteristics will be seen as positive or negative for different wine cultivars, although mouldy taint and wet cardboard have been associated with being a fault and in turn has greatly lowered the liking and quality in most wines (Jackson, 2008). Tannat is a red grape variety that is grown in the South American country of Uruguay. Sensorial attributes were generated by a trained panel of experts, and for this cultivar specifically, found that yeasty and earthy aromas were positively correlated with lower quality perception while dried fruit and berries were positively correlated with a higher perceived quality rating (Varela & Gámbaro, 2006).

Intrinsic characteristics in wine differ from cultivar to cultivar and the green sensorial spectrum in Sauvignon Blanc may not be sought after in other cultivars where “green” attributes are not what characterise the style in terms of typicality.

Sensory perception of wine is often used in the research field in order to explore the outcomes of using different viticultural practices. Altering viticultural practices can lead to a change in chemical compounds that can influence the sensorial characteristics of wine. However, for constructive application of any results, a certain amount of data capturing and data mining would need to be completed before analysis can commence.

## 2.9. DATA MINING

---

### 2.9.1 Data mining evolution

During the past decade advances in information technology has meant that more people are now capable of accessing information on the internet. Studies have shown that in America alone, 22% of all time spent on the internet is spent on social networks such as Facebook, Twitter and other blogs (Wang *et al.*, 2012). This is an indication that there is even more data being uploaded to the internet ranging in magnitude of data to various fields of research. This has led to the development of techniques used to access or mine all the data.

Data mining is a process of trying to discover potential patterns from a large amount of data. It is a multi-disciplinary topic that is based on logics in database systems (Paramasivam *et al.*, 2014). More simply put, data mining is the technique whereby a large data set is filtered to extract useful information while overlooking data that might not be relevant for a particular study (Li, 2015).

### 2.9.2 Data mining application

In the field of biology, data mining has made it possible to do a host of things ranging from sequencing entire genomes both cheaply and quickly to measuring potential interactions between genes in a single or parallel experiment. Due to the advances in data mining and the effectiveness of the tool, new research areas in biology has arisen – computational biology and systems biology (Ahnert, 2013). Other fields where data mining techniques have been used include the banking and financial industry where customers are targeted for marketing segments based on systematic data analyses done on financial records. The retail industries also use data mining techniques to identify purchasing habits and predict consumer trends while telecommunication companies also use data mining techniques in order to identify fraudulent patterns and unusual trends (Li, 2015).

### 2.9.3 Data mining task order and processing

Data mining can be described as a 4 class task; classification, clustering, regression and association (Li, 2015). Paramasivam et al (2014), found that in addition to the 4 class task found by Li in 2015 there are two more tasks namely; text mining and link analysis.

**Classification** - This requires arranging the data into predefined groups

<b>Clustering</b>	-	Which is similar to classification except the groups are not predefined, but are grouped based on similarity
<b>Regression</b>	-	This involves trying to find a function which models the data with the least amount of error
<b>Association rule</b>	-	Requires searching for relationships between variables
<b>Text mining</b>	-	This involves searching unstructured data sets and is used in determining important concepts (Claster <i>et al.</i> , 2010; Coutinho <i>et al.</i> , 2013).
<b>Link analysis</b>	-	This focuses on tracking different populations via segmentation using connectivity between objects.

Furthermore it can be described as a 6 step process as illustrated in figure 2.4 (Paramasivam *et al.*, 2014). In recent times it is commonly agreed that data mining is an essential step in the process of knowledge discovery in databases, or KDD (Li, 2015).



**Figure 2.4:** Stepwise process of data mining

There are several different types of data mining techniques and each one has derivatives of itself. The data mining technique that will be used is dependent on factors such as the size of the data, the type of data and the type of results that the researcher will want to obtain. Examples of data mining techniques are machine learning, visual data mining, neural networks, pattern recognitions and signal processing.

## 2.10 GENERAL CONCLUSION AND FUTURE IMPACT

---

The study is based on contributing to a better understanding of the sensory profile of South African Chenin Blanc wine. The research conducted will be beneficial to the wine academic sector as it will attempt to validate previous studies done on style classification. The study will also assist in setting the foundation for further hypothesis to be based on especially if exploring Chenin Blanc. However, a similar study could be done on other wine cultivars. The results also have a potential impact on the wine industry as for the first time the sensory spectrum of Chenin Blanc will be characterised and made available.

Furthermore, we aim to apply new data analysis tools and processing of a large, unstructured dataset to investigate whether we can exploit the retrieved data. By doing so, we foresee that this

will potentially provide novel insights for the South African wine industry with regards to sensory perception and the description of Chenin Blanc.

## 2.11 LITERATURE CITED

---

- Ahnert SE. (2013). Network analysis and data mining in food science: the emergence of computational gastronomy. *Flavour* [Internet]. Flavour **2**(1):1. Available from: Flavour
- Antwerpen L Van. (2012). Chemical and Sensory profiling of dry and semi-dry South African Chenin blanc wines.
- Asimov E. (2007). Wines of The Times - South Africa's Trove of an Elusive Grape - Review - NYTimes.com [Internet]. [cited 2014 Oct 21]. Available from: [http://www.nytimes.com/2007/03/07/dining/reviews/07wine.html?ref=southafrica&\\_r=0](http://www.nytimes.com/2007/03/07/dining/reviews/07wine.html?ref=southafrica&_r=0)
- Asimov E. (2012). South African Chenin Blancs Aren't What They Used to Be - Wines of The Times - NYTimes.com. [cited 2014 Oct 23]; Available from: [http://www.nytimes.com/2012/07/18/dining/reviews/south-african-chenin-blancs-arent-what-they-used-to-be-wines-of-the-times.html?pagewanted=all&\\_r=0](http://www.nytimes.com/2012/07/18/dining/reviews/south-african-chenin-blancs-arent-what-they-used-to-be-wines-of-the-times.html?pagewanted=all&_r=0)
- Association SM. (2011). A comprehensive guide to Japanese Sake.
- Ballester J, Abdi H, Langlois J. (2009). The odor of colors: Can wine experts and novices distinguish the odors of white, red, and rosé wines? *Chemosens* ... [Internet]. [cited 2014 Aug 13];203–13. Available from: <http://link.springer.com/article/10.1007/s12078-009-9058-0>
- Ballester J, Dacremont C, Le Fur Y, Etiévant P. (2005). The role of olfaction in the elaboration and use of the Chardonnay wine concept. *Food Qual Prefer.* **16**(4):351–9.
- Bester I. (2011). Classifying South African Chenin blanc wine styles [Internet]. [cited 2014 Aug 13]. Available from: <https://ir1.sun.ac.za/handle/10019.1/18048>
- Campo E, Do B V., Ferreira V, Valentin D. (2008). Aroma properties of young Spanish monovarietal white wines: A study using sorting task, list of terms and frequency of citation. *Aust J Grape Wine Res.* **14**(2):104–15.
- Chrea C, Melo L, Evans G, Forde C, Delahunty C, Cox DN. (2011). An investigation using three approaches to understand the influence of extrinsic product cues on consumer behavior: An example of Australian wines. *J Sens Stud.* **26**(1):13–24.
- Claster W, Shanmuganathan S, Sallis P. (2010). Wine Tasting and a Novel Approach to Cluster Analysis. *2010 Fourth Asia Int Conf Math Model Comput Simul* [Internet]. Ieee :152–7. Available from: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5489636>
- Coutinho F, Lang a, Mitschang B. (2013). Making Social Media Analysis more efficient through Taxonomy Supported Concept Suggestion. *Btw* [Internet]. (Cci):457–76. Available from: <http://subs.emis.de/LNI/Proceedings/Proceedings214/457.pdf>
- Cravero MC, Bonello F, Pazo Alvarez MDC, Tsolakis C, Borsa D. (2015). The sensory evaluation of 2,4,6-trichloroanisole in wines. *J Inst Brew* [Internet]. **121**(3):411–7. Available from: <http://doi.wiley.com/10.1002/jib.230>
- Easingwood C, Lockshin L, Spawton A. (2011). The Drivers of Wine Regionality. *J Wine Res.* **22**(1):19–33.
- Fernández-vázquez R, Hewson L, Fisk I, Vila DH, Jose F, Mira H, et al. (2013). Colour influences sensory perception and liking of orange juice. *Flavour.* **3**(8):1–8.

Fisheries AF&. (2013). A PROFILE OF THE SOUTH AFRICAN WINE MARKET VALUE CHAIN.

Foxcroft M. (2009). Growing the consumption of wine amongst emerging market consumers in South Africa. *Cape Wine Diss January* [Internet]. [cited 2014 Oct 21];(January). Available from: [http://www.capewineacademy.co.za/dissertations/capewinemaster\\_foxcroft\\_growing\\_wine\\_consumption.pdf](http://www.capewineacademy.co.za/dissertations/capewinemaster_foxcroft_growing_wine_consumption.pdf)

Garbez M, Galopin G, Sigogne M, Favre P, Demotes-Mainard S, Symoneaux R. (2015). Assessing the visual aspect of rotating virtual rose bushes by a labeled sorting task. *Food Qual Prefer* [Internet]. Elsevier Ltd **40**:287–95. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S095032931400127X>

Goodie J. (2011). South Africa's white wines [Internet]. [cited 2014 Oct 21]. Available from: <http://www.wineanorak.com/sawhites.htm>

Hoffmann M, Varela P, Ares G. (2014). Novel Techniques in Sensory Characterization Profiling and Consumer.

Jackson RS. (2008). 11 - Sensory Perception and Wine Assessment [Internet]. Wine Sci. (Third Ed. Available from: <http://www.sciencedirect.com/science/article/pii/B9780123736468500147>

Jaffré J, Valentin D, Meunier JM, Siliani A, Bertuccioli M, Le Fur Y. (2011). The Chardonnay wine olfactory concept revisited: A stable core of volatile compounds, and fuzzy boundaries. *Food Res Int* [Internet]. Elsevier Ltd **44**(1):456–64. Available from: <http://dx.doi.org/10.1016/j.foodres.2010.09.022>

Johnson R, Bruwer J. (2007). Regional brand image and perceived wine quality: the consumer perspective. *Int J Wine Bus Res.* **19**(4):276–97.

Jolley B. (2014). Development Of quality control tools and a taste prediction model for Rooibos.

Jolly NP, Hattingh S. (2001). A Brandy Aroma Wheel for South African Brandy \*. **22**(1):27–8.

Landau L. (2011). A profile of the South African wine market value chain [Internet]. Available from: <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle>No+Title#0>

Lattey KAA, BRAMLEY BRR, FRANCIS ILL. (2010). Consumer acceptability, sensory properties and expert quality judgements of Australian Cabernet Sauvignon and Shiraz wines. *Aust J Grape Wine Res* [Internet]. [cited 2015 Sep 16];**16**(1):189–202. Available from: <http://doi.wiley.com/10.1111/j.1755-0238.2009.00069.x>

Li Y. (2015). DATA MINING: CONCEPTS, BACKGROUND AND METHODS OF INTEGRATING UNCERTAINTY IN DATA MINING [Internet]. Available from: [http://www.ccsc.org/southcentral/E-Journal/2010/Papers/Yihao final paper CCSC for submission.pdf](http://www.ccsc.org/southcentral/E-Journal/2010/Papers/Yihao%20final%20paper%20CCSC%20for%20submission.pdf)

Loubser F. (2008). Chenin blanc table wine in South Africa. *Cape Wine Master Diss March* [Internet]. [cited 2014 Oct 21];(March). Available from: [http://capewineacademy.co.za/dissertations/CheninBlanc\\_Table\\_Wines.pdf](http://capewineacademy.co.za/dissertations/CheninBlanc_Table_Wines.pdf)

Mouton G. (2006). TERROIR - THE FOOTPRINT OF GREAT WINES. (July).

Paramasivam V, Yee TS, Dhillon SK, Sidhu AS. (2014). A methodological review of data mining techniques in predictive medicine: An application in hemodynamic prediction for abdominal aortic aneurysm disease. *Biocybern Biomed Eng* [Internet]. Korea Institute of Oriental Medicine **34**(3):139–45. Available from: <http://dx.doi.org/10.1016/j.bbe.2014.03.003>

Parr W V., Schlich P, Theobald JC, Harsch MJ. (2013). Association of selected vinicultural factors with sensory and chemical characteristics of New Zealand Sauvignon blanc wines. *Food Res Int* [Internet]. Elsevier Ltd **53**(1):464–75. Available from: <http://dx.doi.org/10.1016/j.foodres.2013.05.028>

Parr W V., Valentin D, Green J a., Dacremont C. (2010). Evaluation of French and New Zealand Sauvignon wines by experienced French wine assessors. *Food Qual Prefer.* **21**(1):56–64.

- Piqueras-Fiszman B, Spence C. (2012). The weight of the bottle as a possible extrinsic cue with which to estimate the price (and quality) of the wine? Observed correlations. *Food Qual Prefer* [Internet]. Elsevier Ltd **25**(1):41–5. Available from: <http://dx.doi.org/10.1016/j.foodqual.2012.01.001>
- Q.D. Truong. (2011). The global wine industry [Internet]. Available from: <http://www.scribd.com/doc/124908342/Wine-Industry-Morgan-Stanley>
- Rahman I, Reynolds D. (2015). Wine: Intrinsic attributes and consumers' drinking frequency, experience, and involvement. *Int J Hosp Manag* [Internet]. Elsevier Ltd **44**:1–11. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S0278431914001443>
- Reed S. (2010). Sensory analysis of chocolate liquor - Sophisticated sciences such as sensory and flavor chemistry are needed to understand flavour development.
- Van Rooyen J. (2015). Formation of the Chenin Blanc Aroma Wheel.
- Sáenz-Navajas M-P, Ballester J, Pêcher C, Peyron D, Valentin D. (2013). Sensory drivers of intrinsic quality of red wines. *Food Res Int* [Internet]. Elsevier Ltd [cited 2014 Aug 13];**54**(2):1506–18. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S0963996913005358>
- Sáenz-Navajas M-P, Ballester J, Peyron D, Valentin D. (2014). Extrinsic attributes responsible for red wine quality perception: A cross-cultural study between France and Spain. *Food Qual Prefer* [Internet]. Elsevier Ltd [cited 2014 Aug 13];**35**:70–85. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S0950329314000238>
- SAWIS. (2014). South African Wine Industry Statistics (No.38) [Internet]. p. 1–44. Available from: [http://www.sawis.co.za/info/download/Engels\\_2014.pdf](http://www.sawis.co.za/info/download/Engels_2014.pdf)
- Spillman PJ, Sefton M a, Gawel R. (2004). The effect of oak wood source, location of seasoning and coopering on the composition of volatile compounds in oak-matured wines. *Aust J Grape Wine Res* [Internet]. **10**(3):216–26. Available from: <http://dx.doi.org/10.1111/j.1755-0238.2004.tb00025.x>
- http://dx.doi.org/10.1111/j.1755-0238.2004.tb00026.x
- Stevenson T, Sotheby's (Firm). (2005). The Sotheby's Wine Encyclopedia.
- Tang VCM, Tchetchik A, Cohen E. (2015). Perception of wine labels by Hong Kong Chinese consumers. *Wine Econ Policy* [Internet]. Elsevier **4**(1):12–21. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S2212977415000113>
- Tao Y, Sun D-W, Górecki A, Błaszcza W, Lamparski G, Amarowicz R, et al. (2012). Effects of high hydrostatic pressure processing on the physicochemical and sensorial properties of a red wine. *Innov Food Sci Emerg Technol*. **19**:545–54.
- Tempère S, Cuzange E, Schaaper MH, de Lescar R, de Revel G, Sicard G. (2014). “Brett character” in wine: Is there a consensus among professional assessors? A perceptual and conceptual approach. *Food Qual Prefer*. **34**:29–36.
- Varela P, Gámbaro A. (2006). Sensory descriptive analysis of Uruguayan Tannat wine: Correlation to quality assessment. *J Sens Stud*. **21**(2):203–17.
- Wang S, Huang Q, Jiang S, Tian Q, Qin L. (2012). Nearest-neighbor method using multiple neighborhood similarities for social media data mining. *Neurocomputing* [Internet]. Elsevier **95**:105–16. Available from: <http://dx.doi.org/10.1016/j.neucom.2011.06.039>
- Weightman CJ. (2014). Characterization of Chenin blanc wines produced by natural fermentation and skin contact : focus on application of rapid sensory profiling methods by.

# Chapter 3

---

## RESEARCH RESULTS

**Correspondence analyses based data mining approach to investigate and model the sensory characteristics of South African Chenin Blanc**

### 3. RESEARCH RESULTS: CORRESPONDENCE ANALYSES BASED DATA MINING APPROACH TO INVESTIGATE AND MODEL THE SENSOR CHARACTERISTICS OF SOUTH AFRICAN CHENIN BLANC

#### 3.1 INTRODUCTION

South African Chenin Blanc has become an emerging force in recent times with export figures showing that the cultivar is now becoming increasingly sought after – 46.6 million litres being exported in 2011 compared to the 53.7 million litres in 2013 (SAWIS, 2014). This is particularly encouraging for the wine industry since the cultivar was previously known not to produce top quality wines, but can now stand on its own as a single cultivar and even be produced organically – wines that do not contain sulphur dioxide/sulphites (Clark 2007, Garaguso & Nardini 2015). The concept of drinking organic wines is more appealing to some consumers as research has shown that consumers are more likely to purchase products if there is minimal damage to the environment (Barber & Taylor, 2013). The emergence of Chenin Blanc can be attributed to the dedication of wine makers to improve the wine quality, but recognition also needs to be given to organisations that promote the cultivar itself.

Within South Africa the Chenin Blanc Association (CBA) as well as The John Platter Wines Guides are examples of the type of organisations that promote Chenin Blanc both locally and on the international front. According to the CBA there are six different styles of Chenin Blanc in South Africa namely; Fresh & Fruity, Rich and Ripe Wooded, Rich and Ripe Unwooded, Rich and Ripe slightly sweet, Sparkling and Sweet (Bester, 2011) Although research has been mainly focused on identifying the different styles within the cultivar, nothing has been done on trying to model the sensory attributes of South African Chenin Blanc. This might be due to the fact that a large enough data set is needed to accurately model any cultivar and there are not many of these data sets available.

Food and wine companies need information on how consumers perceive their products and this can be acquired by using rapid sensory profiling techniques such as Descriptive Analyses (DA), free text descriptions and frequency of citation (Anzanello *et al.*, 2011; Kostov *et al.*, 2014) – refer to Varela and Ares for further reading (Varela & Ares, 2012). These types of methods form part of a structured matrix in order to acquire sensorial information and doing statistical analysis on these types of datasets can be validated by the science behind the methodology. However, when using an unstructured dataset e.g. not giving any instructions or guidelines to generate the data, many questions arise whether the dataset is feasible and whether significant information can be extracted. To process the validity of such an unstructured data set certain data mining techniques would need to be applied.

Data mining is the practice of examining large pre-existing databases in order to generate new information and involves different techniques. Examples of data mining techniques include decision trees, association rule mining, and neural networks, and each of these techniques analyses data in different ways (Yang & Chen, 2015). Due to the advances in information technology it is now easier to mine large data sets (Cortez *et al.*, 2009; Ronowicz *et al.*, 2015). This has become increasingly important as there is a need to assess wines with the aim of exercising quality control - as this is part of the certification process and can improve wine making (Cortez *et al.*, 2009; Appalasamy, 2012).

The aim of data mining is to extract high levels of knowledge from raw data sets that have been compiled together using rapid sensory techniques and also serves as a means to filter through variables and remove data that might have little to no influence on the data set (Ortiz-Servin *et al.*, 2015; Ronowicz *et al.*, 2015). The advantage of using data mining techniques is that the variable ,which differs in fields of study, is not fixed and can range from anything from sensory attributes in sensory science to lung cells in cancer diagnosis (Yang & Chen, 2015). It also offers a broader perspective of the profiling within social and cultural networks in food science (Ahnert, 2013) and more specific to this study a broader perspective on the sensory profile within wine.

The John Platter Wine Guide, which is SA's largest and most accredited wine dataset was used to extract information on 7 years worth of single varietal Chenin Blanc (2008-2014). An unstructured approach of data generation was used as the assessors were not given a set of descriptors to look for and simply stated what they perceived to be in the wine by using both smell and taste. Data processing plays a vital part before any data mining techniques can be used (Appalasamy, 2012) and the same applied for this study. Due to the nature of the dataset and the complexity of the unstructured matrix in which the data was collected, a large amount of standardisation had to be done in order to get the dataset into a structured form for analysis.

As previously stated, research has been done with the aim of identifying the different styles within Chenin Blanc, and a very small data set was used to draw conclusion (Antwerpen, 2012; Hanekom, 2012). Never before has a study been conducted to analyse wines using a data set that included more than 2500 wines over a period of 7 years (2008-2014). The main aim of the paper was to explore the novelty of the topic and attempt to model Chenin Blanc by capturing the different attributes in the hope that this can give us a clearer indication of the sensory space within the cultivar. For characterisation purposes only positively perceived sensory attributes were considered. This would also give us a better understanding of how the cultivar is perceived.

Additional objectives were set to: 1) see whether combinations of attributes are most likely associated with subset groups such as: regions or wine of origin due to the fact that Chenin Blanc can grow in different climates and areas and is easily adaptable to its environment; 2) Determine, using the dataset, what the most cited sensory attributes for South African Chenin Blanc are; 3) Determine if there were any sensory "cues" that represented the wooded and unwooded styles in Chenin Blanc and 4) Identify whether the data set can give us a clear answer regarding style

classification. This is a novel approach as it is the first time a complete data set will be used to evaluate the sensory space within a specific cultivar in South Africa. Similar to a study done by Ballester et al (2005) on French Chardonnay, a key overall objective was set to accurately characterise, not only the sensory space, but also other aspects of the Chenin Blanc cultivar and potentially verify the existence of a Chenin Blanc wine concept (Ballester et al., 2005). The scientific novelty also provides the advantage that the assessors were unaware that the data set would be analysed. Thus doing data mining on a pre-established data set for sensory attributes would be beneficial as it rules out any biased opinion that could be preconceived had the assessors known prior to rating the wine. The reason behind doing the study is to give scientific knowledge back to the industrial and research sector with the hope of finally giving the Chenin Blanc cultivar a recognisable identity.

## 3.2 MATERIALS AND METHODS

---

### John Platter Wine Guide

The John Platter Wine Guide is South Africa's leading wine guide, and has used a number of different assessors over the last couple of years. All the assessors that currently are, or previously were part of the expert panel, have significant experience in the wine industry and many are judges at international wine competitions. A single assessor tastes all the wines of a particular winery for a maximum of three years and the number of assessors varies from 15-20 per year. Even though the assessors are very experienced and have a significant reputation in the wine industry they are not trained or calibrated in order to perceive the same aroma and in turn call it the same thing. Their opinion will therefore be based on individual perspectives, and the descriptors used will reflect these perspectives and thus having 15-20 assessors a year is important as trends being extracted from the unstructured dataset should be significant due to the nature of how the data was generated.

### Wines

Data analysis was only done on South African Chenin Blanc and all the other cultivars were omitted – these include blends that contain Chenin Blanc. Over the period of the 7 years (2008-2014), 2746 wines were tasted. Table 3.1 illustrates the amount of wines being analysed for each year

**Table 3.1** The amount of Chenin Blanc wines that were analysed each year and the total number over the investigation period

Year	Number of wines
2008	336
2009	357
2010	371
2011	380
2012	408
2013	441
2014	453
<b>Total</b>	<b>2746</b>

### Sensorial methods and evaluation in Platter

Cellars from all over the Western Cape send in either tank samples or final bottled products to Platter to be rated. Within South Africa the Western Cape wine producing areas are divided into regions. The regions are then made up by a few smaller districts and that is then made up by even smaller wards (Full list attached in the addendum). The wines first need to be checked to make sure that it was not rated in a previous edition of Platter. However, there are two exceptions to this rule. The first one being that the wine may be tasted and rated for the current edition of platter if a tank sample was sent in the previous year and it is now a final bottled product. The second exception would be that special permission would need to granted to producers who feel their wines might have been unfairly rated or the wine has went on to win numerous awards due to the quality it now possess because of extended aging.

Once the wines are submitted the expert panel sample the wine and give it a star rating out of five and also attach a small description of the wine. Due to the fact that assessors taste the wines in the comfort of their own homes to ensure a relaxed environment and to avoid sensory fatigue a few challenges have arisen. No list of attributes is given and the assessors simply state what they perceive to be in the wine. This has however led to wide variety of words being used to describe the wine – where not all the words were sensorial attributes but included holistic and global views e.g. “picnic wine”, “perfect with fish”. Moreover, on average, assessors used 14 words to describe each wine where only 1.32 was sensorial attributes.

### Standardisation of variables

The description list or “blurb” as mentioned in the platter guide had to be first cleaned up manually to remove words that were not going to be used for the study e.g. ideal for picnics, summer wine. Function words or stop words were also removed from the data such as articles, and prepositions (Bécue-Bertaut 2014). The blurb then had to be cleaned up a few more times to single out attributes and work out discrepancies such as spelling mistakes and the use of a different language (Afrikaans in one case). Duplicates of words within the same blurb were also altered as long as it did not cause loss of credibility e.g. if the same attribute was mentioned twice in the description of the wine.

Once the data had been cleaned up the attributes were standardised in order to combine synonyms (table 3.2) and omit unnecessary data. This led to 266 different attributes being used to describe the Chenin Blanc cultivar. These attributes varied from *acacia and onion skin* to *white stone fruit and Turkish delight*.

In addition to the 266 different attributes being used there was also cases where the same attribute is spoken of, but different descriptions are used or where the same description is used but the assessors may be thinking of a different attribute e.g. Passion fruit being called Granadilla (Kostov et al., 2014). Other complications and challenges came up as assessors would mention things as “more fruit compared to previous vintages”. For the purpose of this study comparative statements were avoided as far as possible and merely stated that the previous as well as the current vintage were “fruity”. Once this was completed all the words were inserted into an excel sheet.

**Table 3.2** Extract of words that had to be “Cleaned up” or standardized from what they were originally described by in the data set.

Standardized word	Variation of word
Apple	Green apple, ripe apple, golden delicious apple, granny smith apple, fleshy apple, crisp apple, apple fruit salad, bruised apple, yellow apple, red apple,
Bubble-gum	Bubblegum, Bubblegum hint, Bubblegummy
Peach	Styling peach, Peach concentration, Peach nuances, Peach flavours, Peach kernel, peachy, flush, sweet-ripe, Peachiness, Peach toned, Gentle peach, sun ripe peach,
Grass	Uncomplicated grass, whafty grass, dusty grass, dry grass
Citrus	Perfumed citrus, Citrus notes, Citrus flavours, Citrus peel, Citrus intensity, Citrus concentration, Citrus zest, fine Citrus, lively Citrus

### Clarification of words

Some of the words had either a negative or a positive connotation attached to them. For the purpose of this study only positive connotations were used for analysis and the rest were discarded. Due to the novelty of the study there is no pre-established protocol as how to approach negative connotations and we decided to omit it as it did not serve the aim of modelling the cultivar. The following words had to be standardised with this mind

- |             |   |
|-------------|---|
| <b>Oak</b>  | – All descriptions that mentioned the presence of oak on the pallet or on the nose were included in the study.  |
| <b>Acid</b> | – Using the descriptor “acid” had to be carefully analysed and this was only included when the assessors commented on the good acidity and presence thereof. Cases where the assessors mentioned that the wines were lacking acidy were simply overlooked for this study. |

<b>Fresh</b>	– Fresh was used in all cases where it was used as a descriptor on its own. If an attribute was described as being a “fresh apple” only the attribute “apple” would be used and the “fresh” would be overlooked
<b>Rich</b>	– The word rich was used as both an olfactory description and as a mouth feel. The word was only included as an olfactory attribute if it represented the entire blurb e.g. “rich fruit”
<b>Balance</b>	– Balance was included as a mouth feel as well as an overall perception of the wine on a sensory level.
<b>Lees characteristics –</b>	Lees was used if the blurb indicated that the wine had lees treatment. It was also used if the assessors had indicated that lees treatment had enhanced the wine or if they could pick up on the treatment.

Another problem that arose was that the data had to be simplified in a way that could be easily computed and therefore some data was lost. This was problematic as by reducing complete sentences to single words or concepts there was the possibility of losing some original meaning and thus impoverish the corpus. Table 3.3 illustrates the “blurb” or description of different wines where there is a common idea, but different words are being used.

**Table 3.3** Illustrates how a common idea can be lost in the description due to the lexicon being used

1. *“Easy going tropical quaffer perfect for summer lunches”.*
2. *“shy, but opens up to delicious ripe peach, apricot notes, oak aromas still prominent but will integrate. Complexity from four different vineyards, three soil types & differing ripeness levels, all artfully blended”*

“Blurb” 1 and 2 in the table 3.3 are extracts from the actual data that were obtained. When doing the statistical analyses, in the first instance, the only word that we were interested in is “tropical”. In the second instance there are a lot more words that would be retrieved. Looking specifically at the words “peach” and “apricot” it could be said that these fruit are indeed tropical fruit, but because the word “tropical” was not used, this would not add to its respective tally. Even though the two blurbs have a similar idea on the fruitiness of the wine they are not tallied the same.

## Statistical methods

### Frequency of citation

- Frequency of citation is a technique used in various fields including sensory science. Citation frequency is the amount of times a specific product is mentioned from a data set (Campo *et al.*, 2010; Jaeger *et al.*, 2015; Picard *et al.*, 2015). We used frequency of citation to see how many times attributes/descriptors were used in the data set over the 7 years (2008-2014)

### Multi dimensional scaling (MDS)

- MDS is based on finding out the similarity and differences between products. The processed data yields a model reflecting product similarities as proximities of points in space (Lawless *et al.*, 1995). This is useful as the spatial configuration may give insights into important dimensions of product differences and clusters or categories that may exist (Lawless *et al.*, 1995; Johnson *et al.*, 2013; Garbez *et al.*, 2015). After computing a frequency of citation task on the obtained data, a correspondence analysis was computed followed by an MDS test

### Cluster analyses

- Cluster analyses is a technique used to partition a data set into a undetermined number of subgroups so that the members in one group would be closely related to each other (Claster *et al.*, 2010; Chrea *et al.*, 2011) or more simple segregates objects into groups based upon similarity (Brotzman *et al.*, 2015). In this paper we did a cluster analyses on the different attributes we retrieved from the data set to determine whether some of the attributes were more closely related than others.

### Correspondence analysis (CA)

- Correspondence analysis (CA) is a multivariate-graphical method which looks at the correspondence or symmetric association between row and column or categorical variables (McEwan & Schlich, 1991; Beh *et al.*, 2011; Vidal *et al.*, 2015). Using the data obtained after doing a frequency of citation task, a CA test was computed.

### Classification and regression trees (CART)

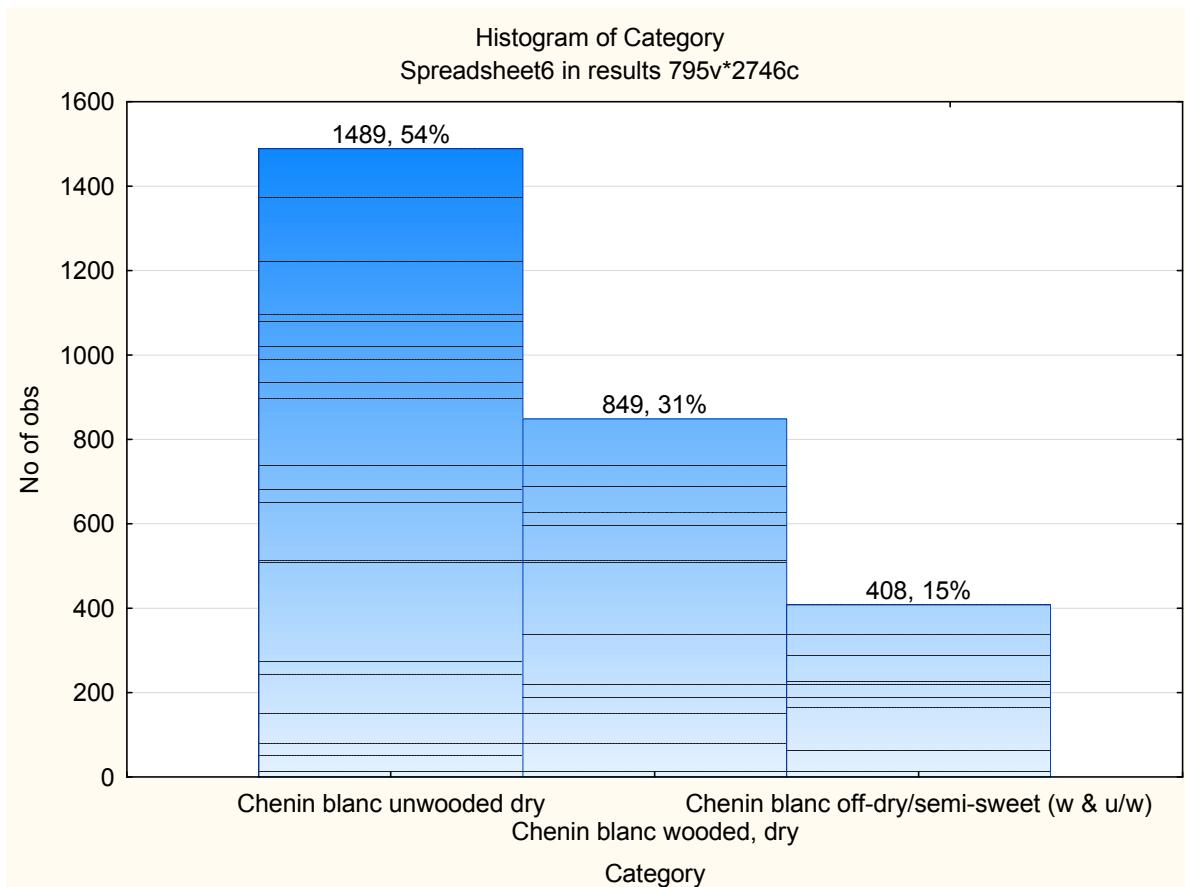
- CART is based on an algorithm that is used for building and evaluating regression trees and predicts continuous (regression) and categorical predictor variables (classification). The analysis has a number of advantages over other classification methods, including multivariate logistic regression (Mahjoobi & Etemad-Shahidi, 2008; Ronowicz *et al.*, 2015). We used the CART method to derive rules from the data to describe the relationship between attributes and styles (Mining, 2009).

## 3.3 RESULTS AND DISCUSSION

---

### Style classification as per dataset

A histogram was computed and according to the dataset that was obtained, there are 3 different styles of Chenin Blanc (figure 3.1). These styles are; wooded, unwooded and off-dry/semi-sweet. The results show that 54% of all the wines that were submitted to Platter (data source) were unwooded dry while 31% of the wines were classified as being wooded. The remaining 15% of wines were classified as being in the off-dry/ semi-sweet category.



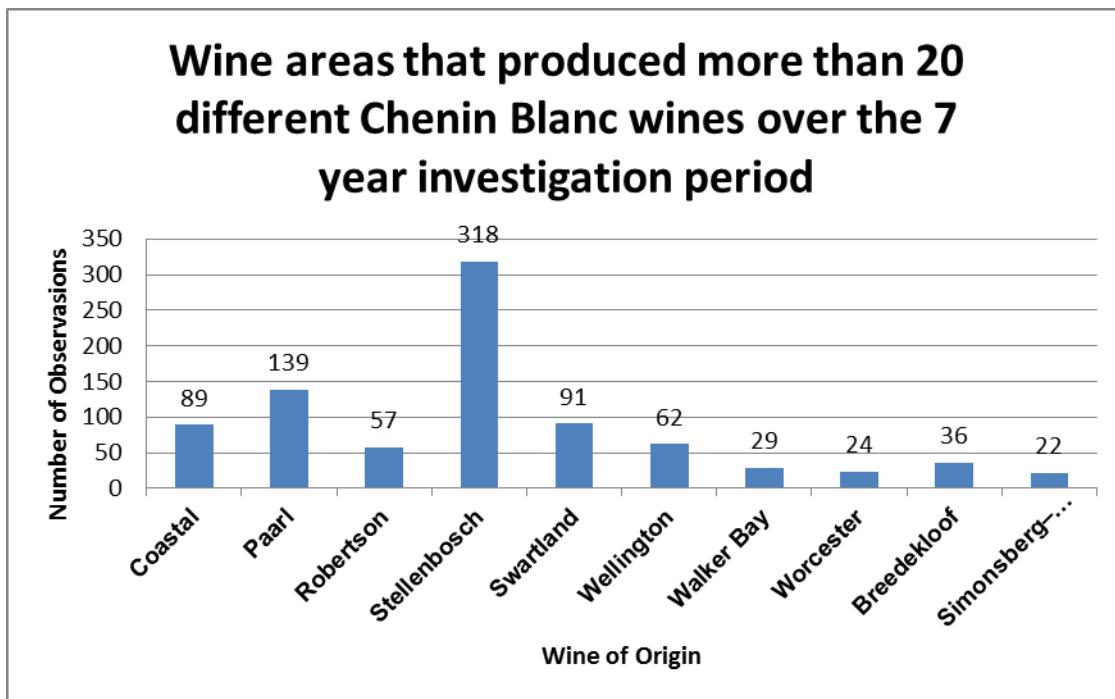
**Figure 3.1** Shows the frequency of the three different styles for Chenin Blanc that was described in the data set.

### Wine analysis and interpretation

\*A complete list of regions, districts and wards is included as addendum A

### Wine producing areas in the Western Cape

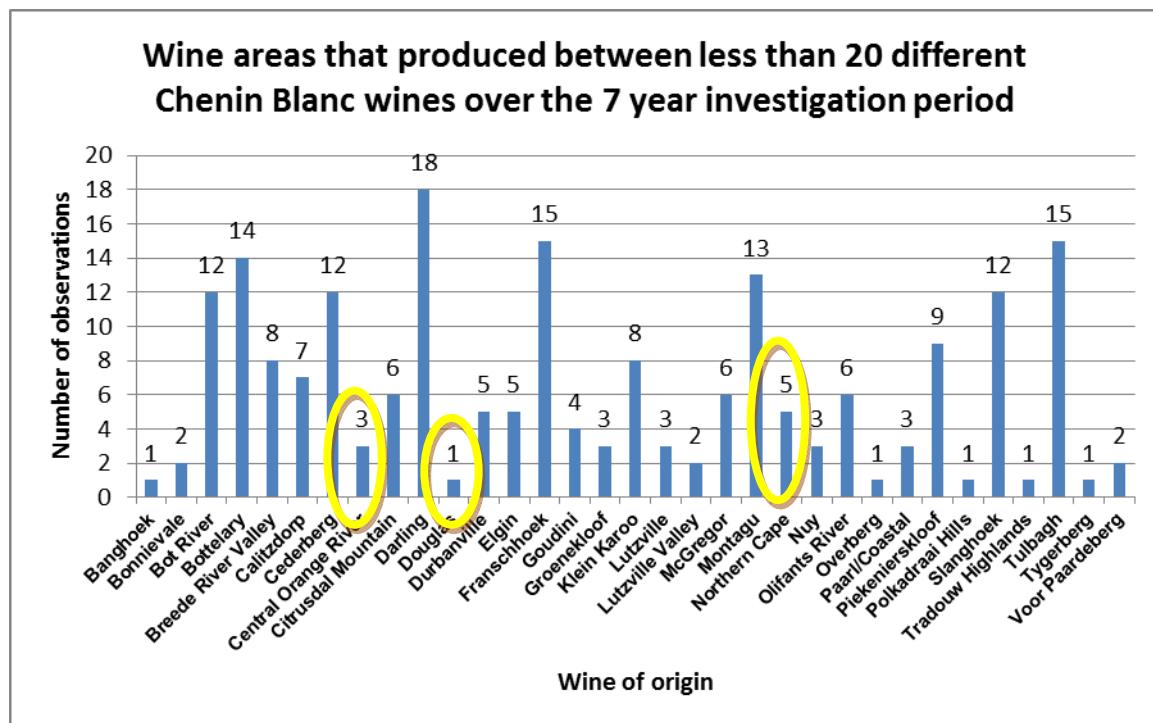
- Including regions, districts and wards



**Figure 3.2** Bar graph showing the different areas in South Africa that produced more than 20 different Chenin Blanc Wines over the investigation period from 2008-2014

Figure 3.2 shows the 10 wine producing areas in South Africa that produced more than 20 different Chenin Blanc wines over the investigation period. An inclusion criterion of at least 20 wines was set as this represents the 10 highest producing areas. Stellenbosch is responsible for the highest amount of Chenin Blanc being produced while Paarl is second. However, Stellenbosch produces more than double the amount that the Paarl region produces. Out of the 10 wine areas the Simonsberg-Stellenbosch district produces the least amount of different Chenin Blanc wines.

The reason for doing this analysis was to see whether there were specific areas in the Western Cape that the Chenin Blanc grapes thrive under and if certain areas were not capable of producing that many different wines because of external conditions. Never before has a study been conducted in South Africa that looks specifically at the different wine areas in the Western Cape or South Africa. As previously said the wine industry has grown over the last decade and this is supported by research showing that more people are consuming wine worldwide. This is the reason why we chose to use an investigation period of 7 years between the years 2008 and 2014. Chenin Blanc can grow in different environments and under different conditions and figure 3.2 illustrates this well. Within the wide expanse of Chenin Blanc growing areas is a vast range of macroclimate and vineyard soil types influenced by the unique geography of the area. This includes several inland mountain chains and valleys. Within the Stellenbosch region alone, there are more than 50 unique soil types with a Mediterranean climate (Stevenson & Sotheby's (Firm), 2005) and this could be a reason for the large amounts of different Chenin Blanc wines being produced in the district.



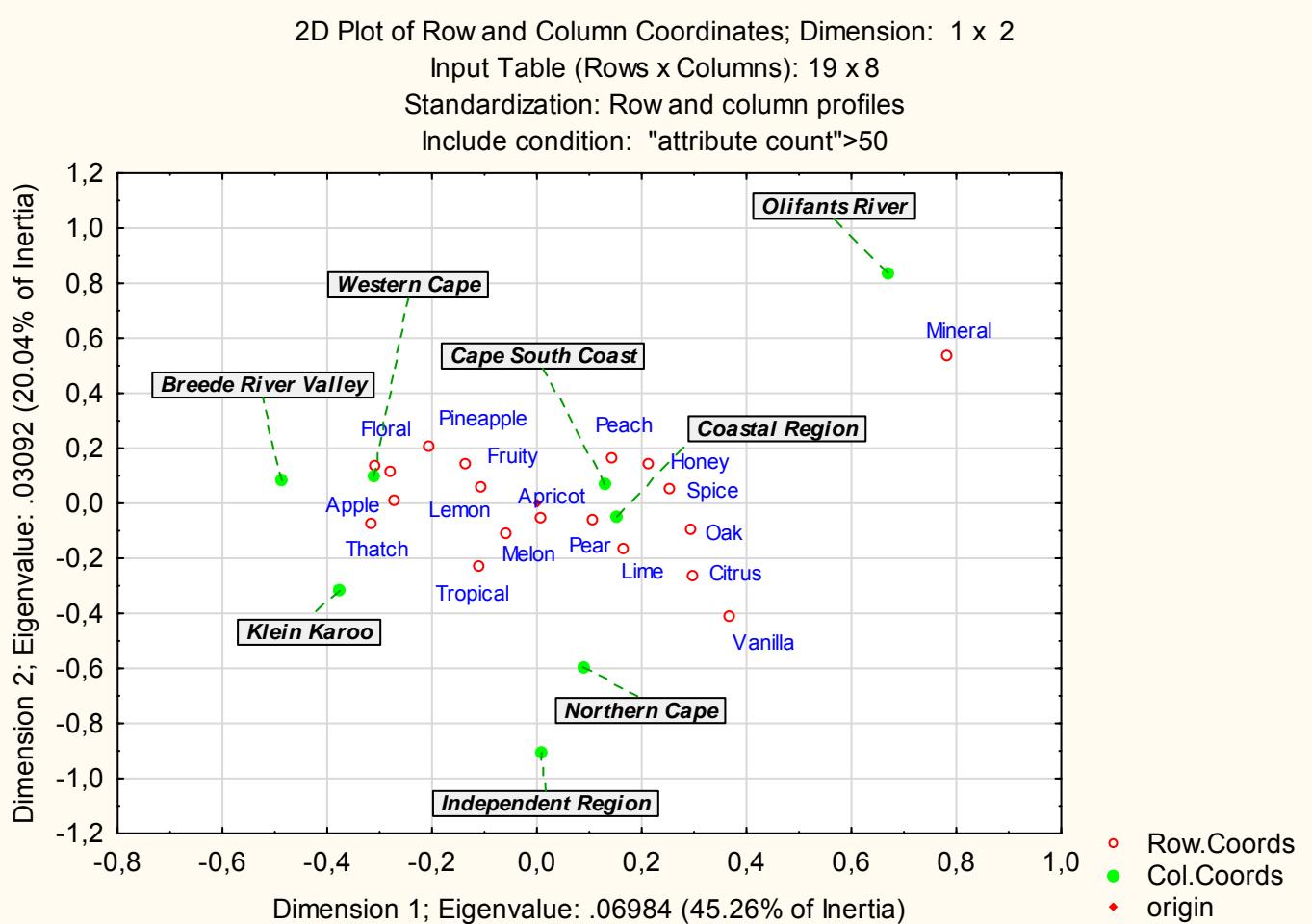
**Figure 3.3** Bar graph showing the different areas in South Africa that produced less than 20 different Chenin Blanc Wines over the investigation period from 2008-2014

When looking at areas that produced less than 20 different Chenin Blanc wines over the 7 years (2008-2014) we saw that Darling produced the most with 18. Franschhoek and Tulbagh were joint second with 15 and as can be seen in figure 3.3 the rest are fairly close together. Notably, more than half of the wine producing areas in the graph on average produced less than 1 Chenin Blanc a year and some only produced one varietal over the entire investigation period. Figure 3 serves to represent the smaller areas or wards in the Western Cape and it can therefore be expected that they do not produce as many different wines as some of the bigger wine producing areas, alternatively it could be that in certain regions, Chenin Blanc is not the preferred growing cultivar and that producers may focus their attention to red wine cultivars. However, the focal point of the graph is to illustrate that although Bot River, Cederberg and Slanghoek only produced 12 different varietals of Chenin Blanc over the investigation period it was still more than the entire Northern Cape had produced – which was 9 (Northern Cape as a province (5), Douglasdale (1) and Central Orange River (3)).

### Wine typicity based on geographical areas

The 6 wine regions in South Africa (Addendum A) were looked at to see whether the concept of typicity or regionality could be validated. The concept of wine typicity or typicality has been employed to express when a particular wine accurately represents its category or concept (Parr et al., 2010). For this analysis, the regions were not used as a single entity, but were a compilation of the wards and districts that represent the region e.g. wines produced in the Goudini and Slanghoek

wards were tallied as Breedekloof and this in turn was tallied as Breede river valley (Addendum A). The region labelled as “No Region” by the South African Wine Industry wine of origin scheme was changed to “independent region” so that it would not be associated with having no wine of origin. However, in some cases there was no wine of origin listed in the data set and these were labelled as “No wine of origin”. There were a few instances where, Western Cape as a province was used as the wine of origin. The amounts of wines represented by the 6 wine regions in addition to the Western Cape, Northern Cape and Lack of wine of origin specified were as follows: Breede River valley 169, Cape South Coast 50, Coastal Region 843, Klein Karoo 31, Olifants River 29, Independent Region 12, Western Cape 317, Northern Cape 9 and regions were the wine of origin was not specified 1286. Figure 3.4 shows that when looking at attributes that have been mentioned more than 50 times in the dataset there were some regions that more closely related with attributes. Both the Northern Cape and the independent regions do not have any significant correspondence with any of the attributes. Olifants River is the main driver of minerality in South Africa and this is of particular interest as the concept of minerality is appealing to most wine makers and researches alike. Breede River Valley and the Klein Karoo have a significant correspondence with attributes such as floral, apple and thatch while the Cape South Coast and the Coastal region have more significant correspondence with tropical fruits such as peach, pear and apricot. More notably, The Coastal and Cape South Coast Regions are the main drivers of the more woodsy attributes that are associated with Chenin Blanc.



**Figure 3.4** Correspondence analyses plot illustrating the 6 wine regions in South Africa and its association with the attributes that represent the Chenin Blanc cultivar as mentioned by the John Platter Wine Guide from 2008-2014.

### Wine sensory analysis based perceived data

**Table 3.4** Shows the 39 sensory attributes for Chenin Blanc that were mentioned 50 times or more over the investigation period of 2008-2014

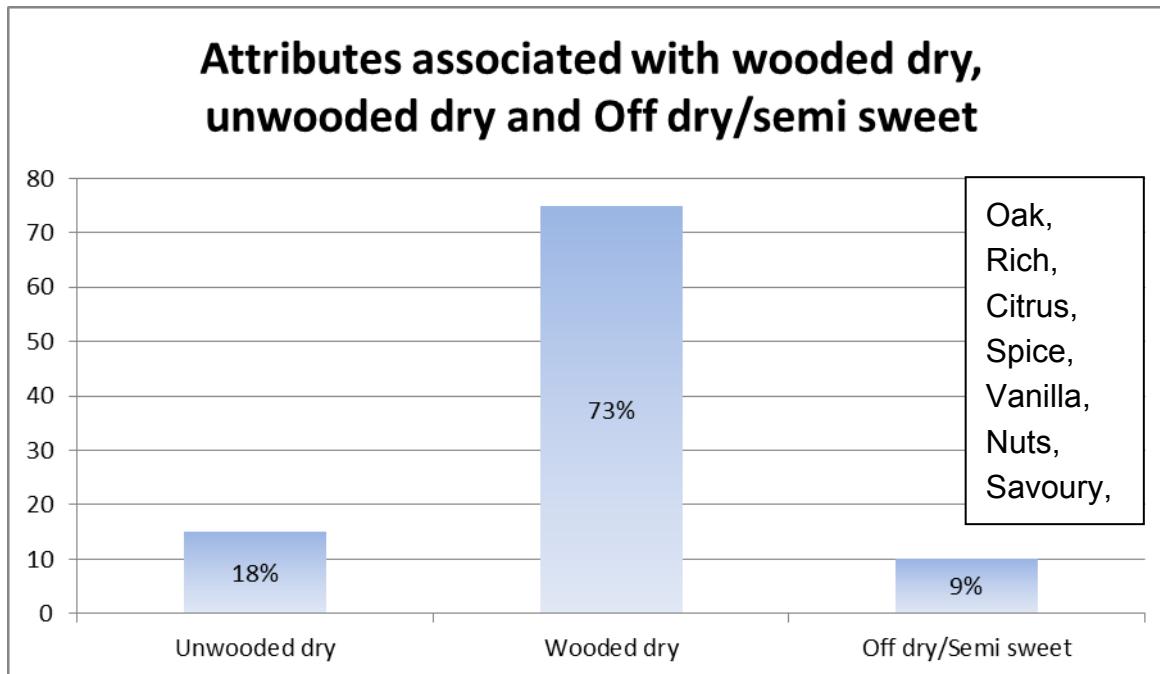
Oak	553	Thatch	111
Acid	525	Mineral	107
Apple	440	Sweet	106
Fresh	395	Citrus	102
Tropical	307	Complex	100
Dry	300	Creamy	97
Fruit	297	Spice	97
Ripe fruit	283	Off dry	93
Rich	253	Apricot	92
Melon	242	Lemon	86
Crisp	232	Quince	85
Balance	214	Juicy	84
Peach	183	Round	80
Honey	172	Vanilla	77
Light	168	Full	70
Pear	149	Savoury	68
Lees characteristics	148	Nuts	63
Pineapple	136	Guava	62
Floral	131	Almond	52
Lime	113		

Table 3.4 shows the attributes that were cited more than 50 times over the 7 years investigation period. Oak was the attribute that was cited most frequently followed by acid. These are the only two attributes that were cited over 500 times. Melon is the most cited tropical fruit and Lemon is the most cited citric fruit. With regards to the use of oak and acid, it was only tallied if the assessors used it in a positive connotation. In some cases it was mentioned that a certain wine was lacking in acidity and this was not tallied up. The reason for this was that it did not meet the aim of modelling the cultivar. Based on the results from table 4 it could be said that a typical Chenin Blanc would have hints of oak and have good acidity. The frequency of tropical and apple also means that there could be the possibility of two poles being a fresh and fruity and wooded Chenin Blanc. This observation is consistent with previous research that was done – even though the previous research was done on a much smaller scale using only a handful of wines (Botha, 2010; Antwerpen, 2011; Hanekom, 2011).

### **Application to modelling Chenin Blanc in terms of attributes and style**

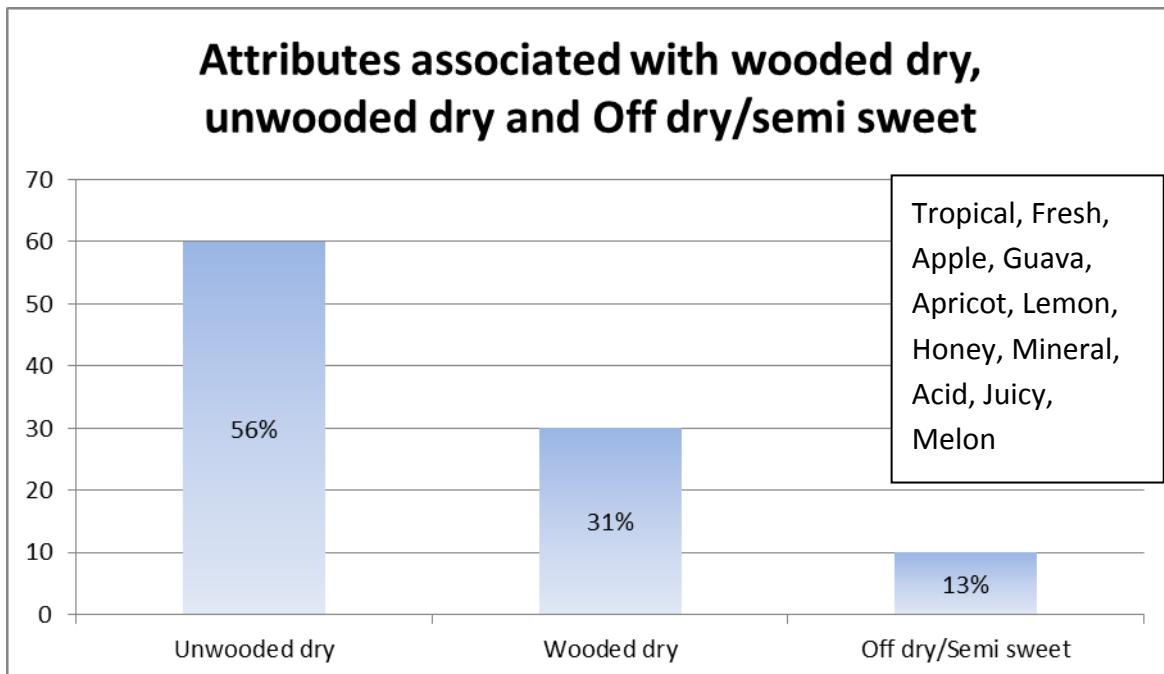
#### **Sensorial “Cues” for Chenin Blanc and style classification**

A cue, in psychological term, refers to a stimulus that is often not consciously perceived, which results in a specific behavioural response (Hanekom *et al.*, 2013). Cues can be divided into two groups, extrinsic (e.g. region) and intrinsic (attributes) cues (Hauteville *et al.*, 2005; Sáenz-Navajas *et al.*, 2013). This then leads us to the question, what cues drive a cultivar? And more specifically for Chenin Blanc, are these cues based on sensory attributes, regionality (Parr *et al.*, 2010) or styles (Hanekom, 2012)? No previous research has been done in order to identify sensorial cues for Chenin Blanc.



**Figure 3.5** Bar graph showing how attributes such as Oak, Rich, Citrus, Spice, Vanilla and nuts are related to the three different styles that were proposed in the data set.

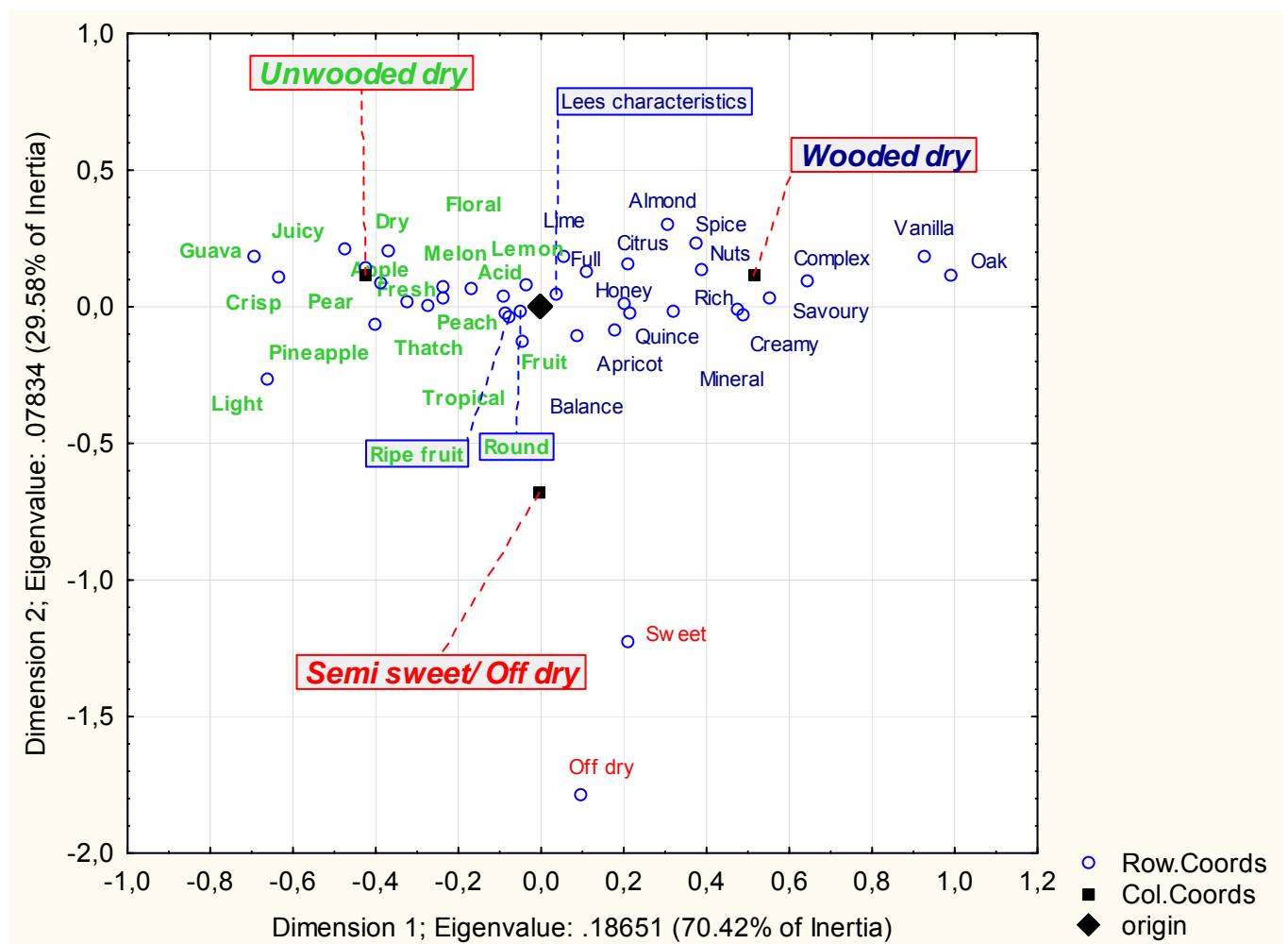
Figure 3.5 shows a bar graph representing the three different styles within the Chenin Blanc cultivar (as proposed by the data set) after computing a CART test. When attributes such as oak, rich, citrus, spice, vanilla, nuts, creamy, savoury and almond were used to describe the wine it was associated with being an unwooded Chenin Blanc 18% of the time and an off dry semi/sweet Chenin Blanc 9% of the time. The graph shows that when these attributes were perceived in the wine, 73% of the time the wine would be a wooded Chenin Blanc. This gives us the first real indication as what the sensorial cues for the Chenin Blanc wooded style is.



**Figure 3.6** Bar graph showing how attributes such as tropical, fresh, apple, guava, apricot, lemon, honey and mineral are related to the three different styles that were proposed in the data set.

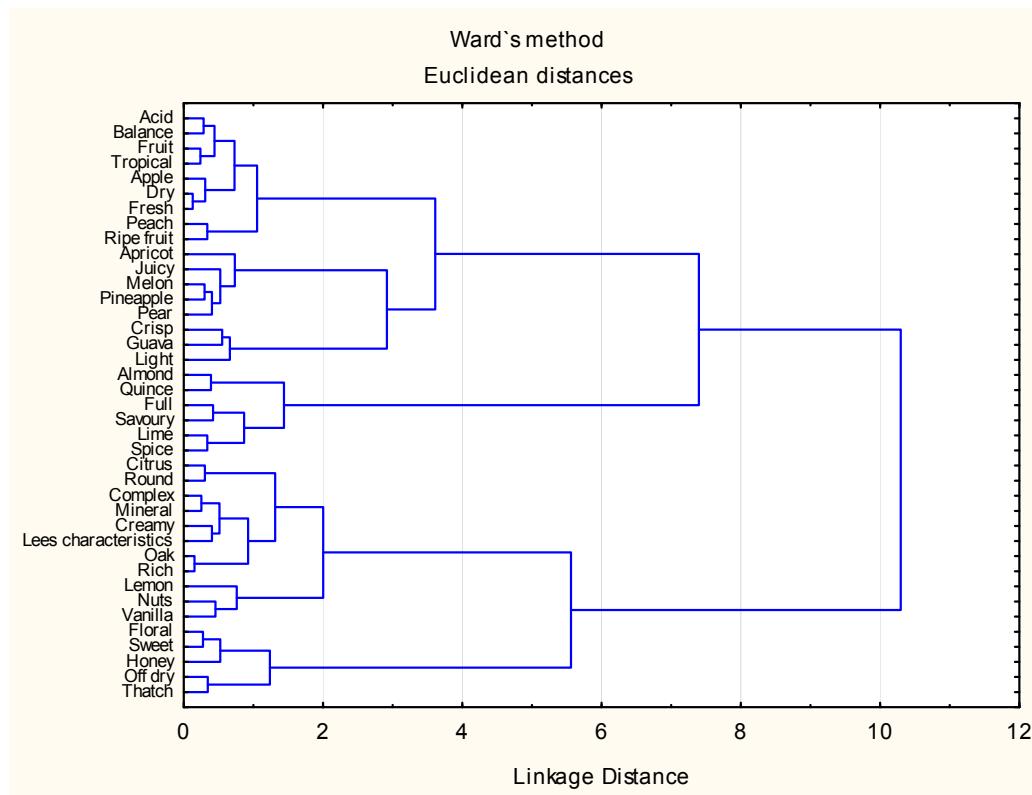
Figure 3.6 shows a bar graph representing the three different styles within the Chenin Blanc cultivar (as proposed by the data set) after computing a CART test. This grouping seemed to be more complex than that of the wooded style. More attributes were used to try and characterise the unwooded style – which is expected due to the large amounts of attributes describing the cultivar as a whole. When observing words such as: tropical, acid, balance, fresh, juicy, crisp, melon, peach, ripe fruit, apple, guava, floral, apricot, dry, pear, thatch, round, fruit, lime, light, honey, lees characteristics, full, lemon, pineapple, quince and mineral we found that 56% of the time if the attribute was used it was associated with being an unwooded Chenin Blanc. 31% of the time when the sensory attributes were used the wine was associated with being a wooded Chenin Blanc and only 13% of the times were the attributes were used was the wine associated with being an off dry/semi-sweet Chenin Blanc. This gives us the first real indication as what the sensorial cues for the Chenin Blanc unwooded style is. However, this is dubious as according to the CBA there should be another style (Rich and Ripe unwooded) present within the unwooded category. The off dry/semi-sweet style (graph not shown) was only characterised by two attributes. The two sensory attributes, off dry and semi-sweet were associated with the off dry/semi sweet style 61% of the time as opposed to the 17% unwooded and the 22% of the wooded dry.

## Correspondence and Cluster Analysis



**Figure 3.7:** Two dimensional correspondence analysis on the sensory attributes for Chenin Blanc against Style classification as proposed by the data set.

Figure 3.7 represents a two dimensional view of the correspondence analyses that was done on the platter data. The figure illustrates the correspondence between all the attributes and also how closely they are related to the three different styles. The attributes strongly illustrate the sensory space with an explained variance of 70.42%. For the sake of having a clear observation of the graph, the labels that overlapped each other were slightly moved. The original location of the attribute is still shown by the dotted line connecting to the origin point (in this case the small circle).

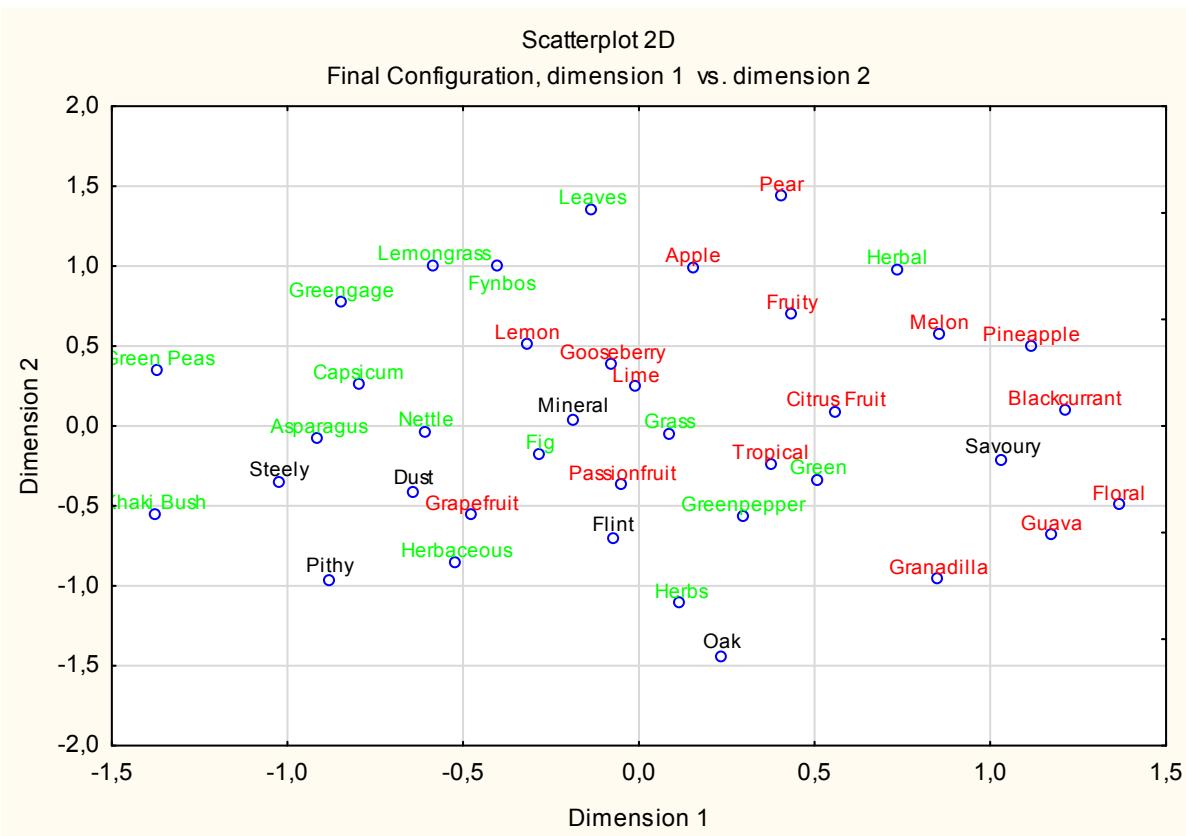


**Figure 3.8** Cluster analysis on 2 dimensions of the sensory attributes within Chenin Blanc using multidimensional scaling

Figure 3.8 illustrates a 2-dimensional view of the clusters that reside within the attributes after computing a cluster analysis. Depending on what linkage distance you looking at can determine the amount of groups or clusters within the study. When looking at the linkage distance 4 we can clearly see that four clusters are evident and likewise when looking at linkage distant 8 we can only see two clusters present. The Cluster analysis validates previous research done that concluded that there were two clear style groups namely Chenin Blanc Wooded and Chenin Blanc Unwooded while the Rich and Ripe Unwooded is not particularly evident.

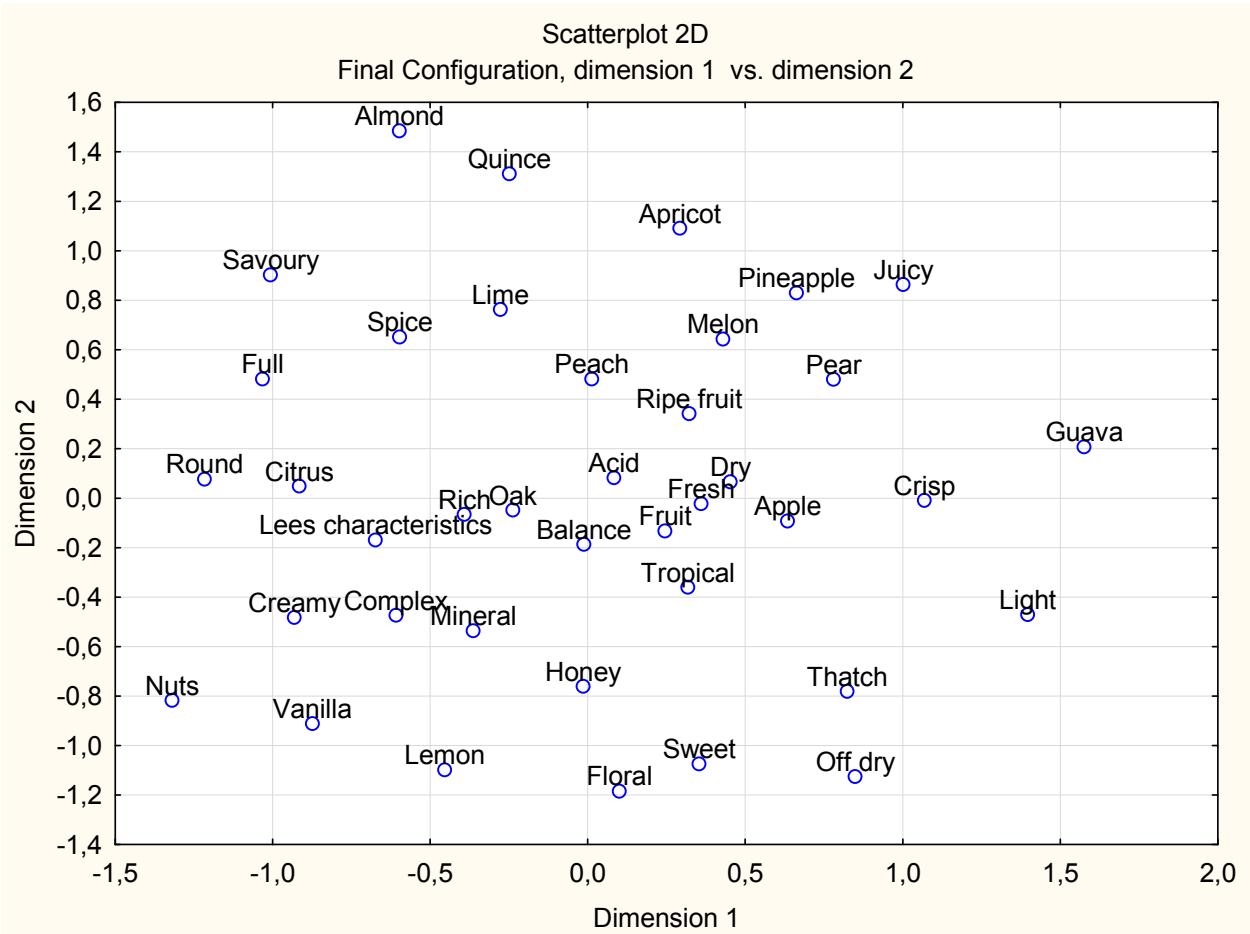
### Sauvignon Blanc validation

The novelty of the study meant that there was no way to accurately compare our results with previously published results. We then had to find another way to validate the results and did another correspondence analysis on a cultivar that is very well documented and modelled, Sauvignon Blanc. The same standardization protocol and methods was done on 4352 Sauvignon Blanc wines over the same 7 year investigation period. The data was taken from the same source as the Chenin Blanc data – John Platter Wine Guide. As previously stated, Sauvignon Blanc is well documented as having two clear poles being green and tropical.



**Figure 3.9** A multidimensional scaling plot of the sensory attributes in Sauvignon Blanc using 4352 wines over a 7 year investigation period (2008-2014)

Figure 3.9 clearly shows the two poles within the Sauvignon Blanc cultivar. On the left hand side of the plot the green spectrum is evident with attributes such as green peas, asparagus and khaki bush (coloured in Green). The middle of the plot shows the combination of spectrums with attributes such as mineral and gooseberry showing the diversity. The right hand side of the plot shows the tropical end of the spectrum with pineapple, guava, melon and blackcurrant being typical of a tropical Sauvignon Blanc (coloured in red). The plot does accurately model the sensory space of Sauvignon Blanc and this serves as means of validation to the Chenin Blanc data and results as it is done by the same assessors and using the same database.



**Figure 3.10** A multidimensional scaling plot of the sensory attributes in Chenin Blanc using 2746 wines over a 7 year investigation period (2008-2014)

As can be seen in figure 3.10 the modelling of Chenin Blanc is extremely difficult due to the fact that it does not have distinct pole classification. But based on the correspondence analysis that is previously shown the only thing we can deduce is that some attributes are more closely related to style modelling and classification than others. The left hand side of the plot is dominated by more wooded Chenin Blanc characteristics such as Vanilla, Nuts and savoury while the right hand side of the plot is more tropical.

### 3.4 Conclusion

The study was based on doing statistical analyses on more than 2746 Chenin Blanc wines in South Africa. Sensory descriptors on Chenin Blanc wines were obtained and a variety of different methods were used to see whether it was possible to model the complete cultivar on a larger scale than has ever been done using a reliable data base over an extended period of time. These results can be relevant to all wine makers and also to the field of sensory research as a whole as it gives in-depth knowledge on what South African Chenin Blanc is perceived as and whether or not this is the direction that wine makers want to move forward in.

The fact that there are so many sensory attributes associated with the cultivar makes it harder for the average consumer to process what is in the wine and also to understand the profile of the wine by looking at the label. There is an increased need for a simpler way to explain the profile of Chenin Blanc and hence the need for it to be. The data and results were validated by using a well-known and pre-modelled cultivar - Sauvignon Blanc. This shows that the data from the John Platter Wine Guide as well as the assessors are reliable and accurate when it comes to profiling wine.

Given the magnitude of the data it would be recommended to do more in depth research on a both cultivars and explore other variables such as price and star rating (Out of 5).

### 3.5 LITERATURE CITED

---

- Ahnert SE. (2013). Network analysis and data mining in food science: the emergence of computational gastronomy. *Flavour* [Internet]. Flavour **2**(1):1. Available from: Flavour
- Antwerpen L Van. (2011). Chemical and Sensory profiling of dry Chenin blanc wines styles Chenin blanc. (November).
- Antwerpen L Van. (2012). Chemical and Sensory profiling of dry and semi-dry South African Chenin blanc wines.
- Anzanello MJ, Fogliatto FS, Rossini K. (2011). Data mining-based method for identifying discriminant attributes in sensory profiling. *Food Qual Prefer* [Internet]. Elsevier Ltd [cited 2014 Sep 30];**22**(1):139–48. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S0950329310001552>
- Appalasamy P. (2012). Classification-based data mining approach for quality control in wine production. *J Appl Sci*.
- Ballester J, Dacremont C, Le Fur Y, Etiévant P. (2005). The role of olfaction in the elaboration and use of the Chardonnay wine concept. *Food Qual Prefer*. **16**(4):351–9.
- Barber N a., Taylor DC. (2013). Experimental approach to assessing actual wine purchase behavior. *Int J Wine Bus Res* [Internet]. **25**(3):203–26. Available from: <http://www.emeraldinsight.com/10.1108/IJWBR-2012-0013>
- Bécue-bertaut M. (2014). Tracking verbal-based methods beyond conventional descriptive analysis in food science bibliography . A statistical approach. *Food Qual Prefer* [Internet]. Elsevier Ltd **32**:2–15. Available from: <http://dx.doi.org/10.1016/j.foodqual.2013.08.010>
- Beh EJ, Lombardo R, Simonetti B. (2011). A European perception of food using two methods of correspondence analysis. *Food Qual Prefer* [Internet]. Elsevier Ltd **22**(2):226–31. Available from: <http://dx.doi.org/10.1016/j.foodqual.2010.10.001>
- Bester I. (2011). Classifying South African Chenin blanc wine styles [Internet]. [cited 2014 Aug 13]. Available from: <https://ir1.sun.ac.za/handle/10019.1/18048>
- Botha H. (2010). and Chemical Profiling of South African Natural Sweet Chenin blanc wine.
- Brotzman RL, Cook NB, Nordlund K, Bennett TB, Gomez Rivas a., Döpfer D. (2015). Cluster analysis of Dairy Herd Improvement data to discover trends in performance characteristics in large Upper Midwest dairy herds. *J Dairy Sci* [Internet]. :3059–70. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S0022030215001095>
- Campo E, Ballester J, Langlois J, Dacremont C, Valentin D. (2010). Comparison of conventional descriptive analysis and a citation frequency-based descriptive method for odor profiling: An application to

- Burgundy Pinot noir wines. *Food Qual Prefer* [Internet]. Elsevier Ltd **21**(1):44–55. Available from: <http://dx.doi.org/10.1016/j.foodqual.2009.08.001>
- Chrea C, Melo L, Evans G, Forde C, Delahunty C, Cox DN. (2011). An investigation using three approaches to understand the influence of extrinsic product cues on consumer behavior: An example of Australian wines. *J Sens Stud.* **26**(1):13–24.
- Claster W, Shanmuganathan S, Sallis P. (2010). Wine Tasting and a Novel Approach to Cluster Analysis. *2010 Fourth Asia Int Conf Math Model Comput Simul* [Internet]. Ieee :152–7. Available from: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5489636>
- Cortez P, Cerdeira A, Almeida F, Matos T, Reis J. (2009). Modeling wine preferences by data mining from physicochemical properties. *Decis Support Syst* [Internet]. Elsevier B.V. [cited 2014 Sep 30];**47**(4):547–53. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S0167923609001377>
- Garaguso I, Nardini M. (2015). Polyphenols content, phenolics profile and antioxidant activity of organic red wines produced without sulfur dioxide/sulfites addition in comparison to conventional red wines. *Food Chem* [Internet]. Elsevier Ltd **179**:336–42. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S0308814615001806>
- Garbez M, Galopin G, Sigogne M, Favre P, Demotes-Mainard S, Symoneaux R. (2015). Assessing the visual aspect of rotating virtual rose bushes by a labeled sorting task. *Food Qual Prefer* [Internet]. Elsevier Ltd **40**:287–95. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S095032931400127X>
- Hanekom E. (2011). Chenin blanc Wine Style Hélène Nieuwoudt Chenin blanc wine style diversity. (November):10–1.
- Hanekom E. (2012). Chemical , sensory and consumer profiling of a selection of South African Chenin blanc wines produced from bush vines.
- Hanekom E, Tredoux A, Muller N. (2013). Chenin blanc label cues : do they influence consumer liking and choice of wine ? (May):2–4.
- Hauteville FD, Perrouty JP, Perrouty JP. (2005). Intrinsic and extrinsic quality cues : The moderating effect of expertise on consumer judgements on the quality of wines.
- Jaeger SR, Beresford MK, Paisley AG, Antúnez L, Vidal L, Cadena RS, et al. (2015). Check-all-that-apply (CATA) questions for sensory product characterization by consumers: Investigations into the number of terms used in CATA questions. *Food Qual Prefer* [Internet]. Elsevier Ltd **42**:154–64. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S0950329315000348>
- Johnson TE, Hasted A, Ristic R, Bastian SEP. (2013). Multidimensional scaling (MDS), cluster and descriptive analyses provide preliminary insights into Australian Shiraz wine regional characteristics. *Food Qual Prefer* [Internet]. Elsevier Ltd **29**(2):174–85. Available from: <http://dx.doi.org/10.1016/j.foodqual.2013.03.010>
- Kostov B, Bécue-Bertaut M, Husson F. (2014). An original methodology for the analysis and interpretation of word-count based methods: Multiple factor analysis for contingency tables complemented by consensual words. *Food Qual Prefer* [Internet]. Elsevier Ltd **32**:35–40. Available from: <http://dx.doi.org/10.1016/j.foodqual.2013.06.009>
- Lawless HT, Sheng N, Knoops SSCP. (1995). Multidimensional scaling of sorting data applied to cheese perception. *Food Qual Prefer.* **6**(2):91–8.
- Mahjoobi J, Etemad-Shahidi a. (2008). An alternative approach for the prediction of significant wave heights based on classification and regression trees. *Appl Ocean Res* [Internet]. Elsevier Ltd **30**(3):172–7. Available from: <http://dx.doi.org/10.1016/j.apor.2008.11.001>
- McEwan J a., Schlich P. (1991). Correspondence analysis in sensory evaluation. *Food Qual Prefer.* **3**(1):23–

36.

- Mining D. (2009). Classification and Regression Trees. (November).
- Ortiz-Servin JJ, Cadenas JM, Pelta D a., Castillo A, Montes-Tadeo J-L. (2015). Nuclear fuel lattice performance analysis by data mining techniques. *Ann Nucl Energy* [Internet]. **80**:236–47. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S030645491500081X>
- Parr W V., Valentin D, Green J a., Dacremont C. (2010). Evaluation of French and New Zealand Sauvignon wines by experienced French wine assessors. *Food Qual Prefer*. **21**(1):56–64.
- Picard M, Tempere S, de Revel G, Marchand S. (2015). A sensory study of the ageing bouquet of red Bordeaux wines: A three-step approach for exploring a complex olfactory concept. *Food Qual Prefer* [Internet]. Elsevier Ltd **42**:110–22. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S0950329315000233>
- Ronowicz J, Thommes M, Kleinebudde P, Krysi Ski J. (2015). A data mining approach to optimize pellets manufacturing process based on a decision tree algorithm. **73**:44–8.
- Sáenz-Navajas MP, Campo E, Sutan A, Ballester J, Valentin D. (2013). Perception of wine quality according to extrinsic cues: The case of Burgundy wine consumers. *Food Qual Prefer*. **27**(1):44–53.
- SAWIS. (2014). SAWIS [Internet]. [cited 2015 Apr 28]. Available from: <http://www.sawis.co.za/>
- Stevenson T, Sotheby's (Firm). (2005). The Sotheby's Wine Encyclopedia.
- Varela P, Ares G. (2012). Sensory profiling, the blurred line between sensory and consumer science. A review of novel methods for product characterization. *Food Res Int* [Internet]. Elsevier Ltd **48**(2):893–908. Available from: <http://dx.doi.org/10.1016/j.foodres.2012.06.037>
- Vidal L, Tárrega A, Antúnez L, Ares G, Jaeger SR. (2015). Comparison of Correspondence Analysis based on Hellinger and chi-square distances to obtain sensory spaces from check-all-that-apply (CATA) questions. *Food Qual Prefer* [Internet]. Elsevier Ltd **43**:106–12. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S0950329315000634>
- Yang H, Chen Y-PP. (2015). Data mining in lung cancer pathologic staging diagnosis: Correlation between clinical and pathology information. *Expert Syst Appl* [Internet]. Elsevier Ltd **42**(15-16):6168–76. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S0957417415002134>

# Chapter 4

---

## RESEARCH RESULTS

**Validation and expansion of the South African Chenin Blanc Aroma Wheel to accommodate the sensory space and stylistic spectrum of South African Chenin Blanc wine**

## **4. RESEARCH RESULTS : VALIDATION AND EXPANSION OF THE SOUTH AFRICAN CHENIN BLANC AROMA WHEEL TO ACCOMMODATE THE SENSORY SPACE AND STYLISTIC SPECTRUM OF SOUTH AFRICAN CHENIN BLANC WINE**

### **4.1 INTRODUCTION**

Chenin Blanc, the versatile white wine grape cultivar (Hanekom, 2012), has developed such that it is consumed more in recent times in comparison to previous years (Fisheries, 2013; WOSA, 2014). With 18% of the total hectares of grapevines planted being Chenin Blanc, it is important for the wine industry that the cultivar is well represented and understood (Botha, 2010; Bester, 2011; Antwerpen *et al.*, 2013; Fisheries, 2013; WOSA, 2014).

Organisations such as the Chenin Blanc Association (CBA) and Pinotage Association (PO) were established to, amongst other things, enhance and promote the cultivar with which they are affiliated with. An essential part of promoting the cultivar was to develop a document that could give consumers an indication of what the cultivar comprised of based on sensory profiles.

These sensory profiles are combined to form an Aroma wheel. Aroma wheels are used globally to provide a visual representation of a given product. The formulation of an aroma wheel is based acquiring specific reference standards that are perceived to be representative of the product and then having a panel evaluate them in terms of similarity – to the product (Jolly & Hattingh, 2001). As previously stated at section 3 of the literature review, there are conventional ways of developing or formulating an aroma wheel.

However, the Chenin Blanc aroma wheel, which was developed in 2005 by a few members of the CBA was not developed using conventional practices. The manner in which they retrieved data and developed the wheel was unconventional, but was primarily based on previous literature. Using literature on the sensory aromas within white wine, a few descriptors were selected and compiled to form the wheel (Rooyen, 2015).

The aroma wheel was not compiled using a scientific method and even though this is the case, it is still used both in South Africa and internationally as a good reference point to the sensory profile of Chenin Blanc. The Chenin Blanc aroma wheel that was put forward by the CBA is made up of eight groups with four sub-groups and contains sensory descriptors that attempts to cover the entire sensory space within the Chenin Blanc cultivar.

Sensory space is a term that refers to the full dimension of attributes/descriptors within a cultivar, but previous studies have shown that this is not easy to achieve or determine as various factors influence the broad term which is “sensory space” (Campo *et al.*, 2008).

Since the establishment of the Chenin Blanc aroma wheel in 2005 it has not been re-evaluated or validated and the possibility of it being outdated is likely as the cultivar has evolved in the past couple of years with regards to style classification and the emergence of a more mineral profile in some cases (Bester, 2011; Antwerpen, 2012; Hanekom, 2012) . The John platter Wine Guide is South Africa's bestselling wine annual and is well respected both locally and internationally (Platter, 2015). The guide is a representation of all the South African wines produced in a given year and is therefore the single biggest database with regards to wine information that can be found in the South.

A study was conducted by Jolly *et al* (2001) where they attempted and successfully compiled a brandy aroma. For his study, Jolly used a descriptor list developed specifically for brandy consisting of 61 terms (compiled using previous literature) and added a further 50 terms after consulting with brandy experts. Since brandy is distilled from grape wine, 94 wine descriptors were also included. Subsequent to that, 21 members from the brandy industry were selected to decide on the most common attributes/descriptors from the compiled list (Jolly & Hattingh, 2001).

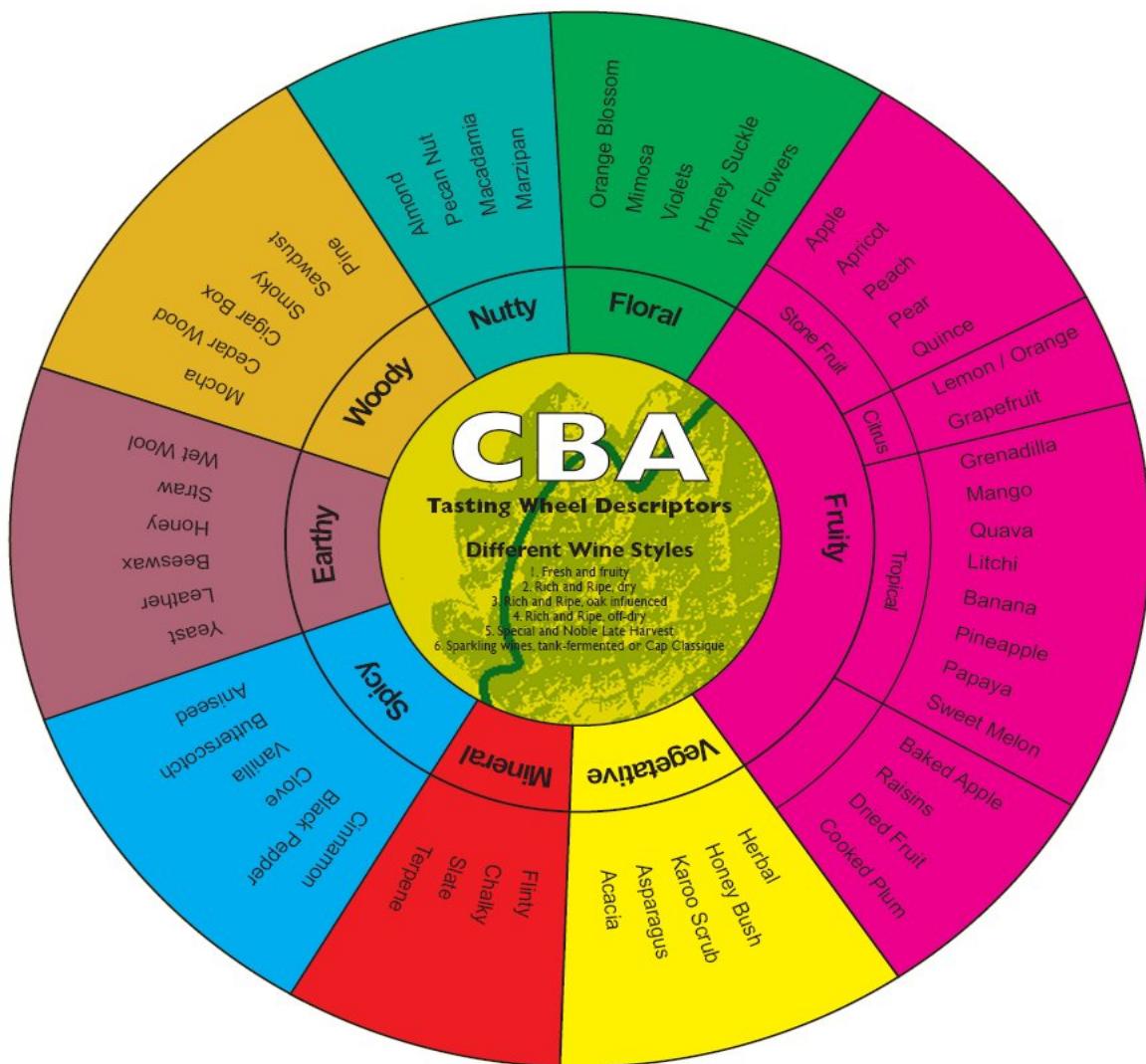
The approach we propose is a novel one and is based on revising and validating the existing Chenin Blanc Aroma Wheel and potentially developing a new one that represents the styles in the South African cultivar using a neutral, accredited and reliable source, the John Platter Wine Guide. Contrary to the work done by Jolly *et al* (2001), the most common attributes/descriptors are not determined by a panel of experts selecting from a pre-established list, but rather by a statistical program computing 7 years' worth of data (2008-2014).

## 4.2 MATERIALS AND METHODS

---

### 4.2.1 Current Chenin Blanc Aroma Wheel design

The existing Chenin Blanc aroma wheel put forward by the CBA contains eight groups namely; Floral, Nutty, Woody, Earthy, Spicy, Mineral, Vegetative and Fruity. The fruity category is further divided into 4 subgroups being; Stone fruit, Citrus, Tropical and an unnamed group containing cooked and dried fruits (figure 4.1)



**Figure 4.1:** The aroma wheel compiled in 2005 by the Chenin Blanc Association that represent the sensory space of Chenin Blanc wines

#### 4.2.2 Data source and structure

Data was retrieved from the John Platter Wine Guide on Chenin Blanc from 2008-2014 ( $n = 2746$ ). The guide provides both a sensory blurb – produced by a trained panel of experts, as well as a star rating out of 5. The sensory blurb contains attributes/descriptors that are perceived in the wine and also includes information varying from the mouth feel to the idle dishes to pair with some of the wines.

The process in which the John Platter Wine Guide goes from acquiring the wines from wine farms to publication is a long and carefully planned out one that spans over a 3-4 month period. The trained panel of experts that analyse the wines give an unbiased and more importantly an accurate and precise representation of each wine. This makes the database the only reliable and accredited source

to accurately re-evaluate any aroma wheel for South African wines. It is also the best means available to span and determine the entire sensory space within a grape cultivar

#### **4.2.3 Data processing**

In order to perform statistical analysis on the acquired data matrix the cleaning up and standardization procedure described in chapter 3 - section 3.2 (Methods and materials – Standardisation of variables) was followed.

#### **4.2.4 Data analysis**

A frequency of citation sheet was compiled with the statistical program STATISTICA. The frequency of all the words listed in the existing Chenin Blanc aroma wheel were analysed and verified to see whether it should remain on the wheel or not..

### **4.3 RESULTS**

---

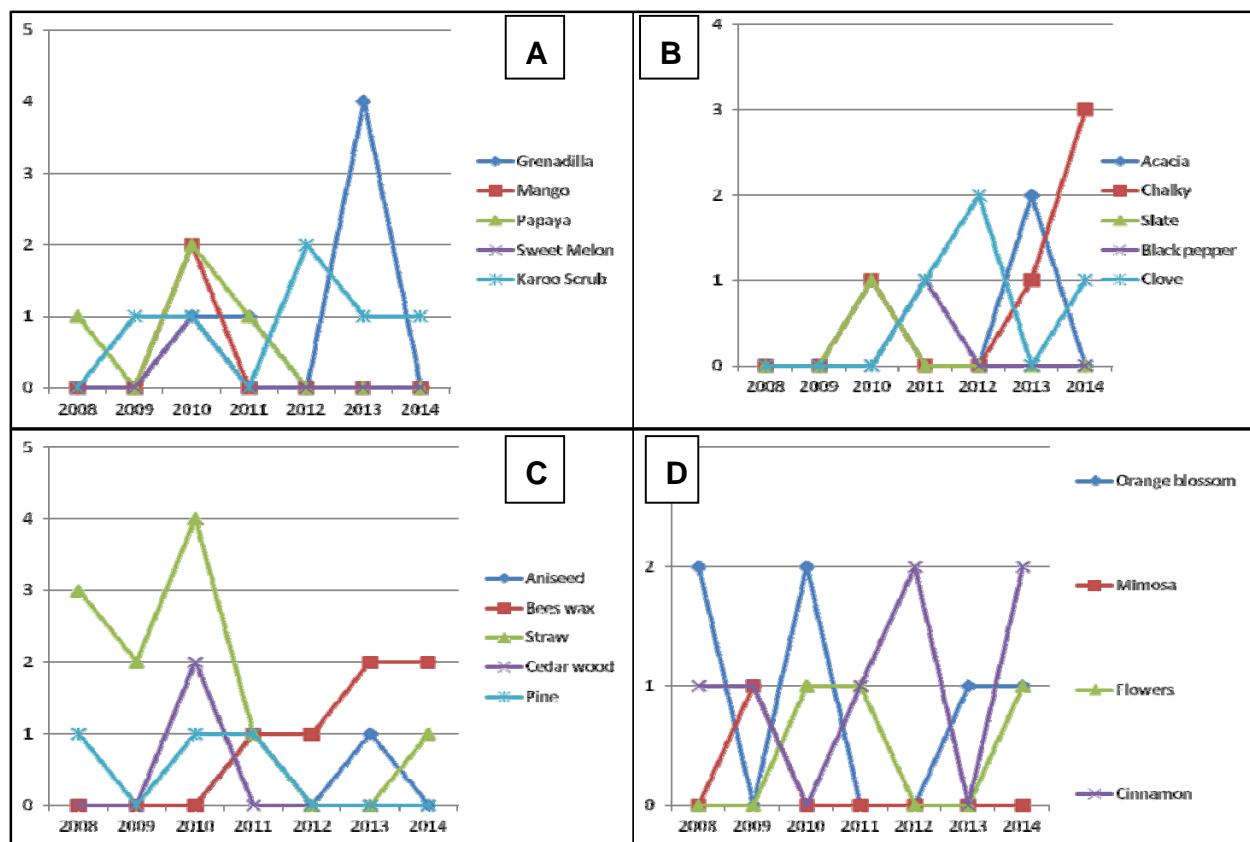
The frequency of citation table extracted from STATISTICA showed that with a dataset of 2746 different Chenin Blanc wines over the 7 year investigation period, more than 266 unique attributes were used by the expert panel when profiling South African Chenin Blanc wines between 2008 and 2014. The sensory space within the cultivar covers a variety of spectrums ranging from tropical to green orientated as well as more wooded profiles with floral descriptors. Cardamom, baked bread, camphor and butterscotch are just some of the descriptors that form part of what was known as the chameleon cultivar (Lawrence, 2012).

As diverse as the Chenin Blanc Aroma Wheel is, after the frequency of citation table was compiled on the existing attributes on the aroma wheel it was observed that there were descriptors in the existing Chenin Blanc Aroma Wheel that were never used. In addition when compiling the frequency of citation for all the words used in the database it was observed that there were descriptors that were frequently used, but were not included on the current aroma wheel.

The descriptors that are currently on the existing Chenin Blanc aroma wheel were all categorised in groups according to the frequency of citation over the 7 year investigation period (2008-2014). All the descriptors that were not cited at all are placed in the first column labelled “0”. Likewise all the descriptors that were cited between 1-4 is labelled in the corresponding group and so forth – as can be seen in table 4.1.

**Table 4.1:** Frequency of citation of all the words currently on the Chenin Blanc Aroma Wheel using the data set spanning over the 7 year investigation period (2008-2014)

	Frequency of citation						
	0	1-4	5 – 10	11 – 20	21 – 30	31 - 100	More than 101
Attributes/ descriptors	Asparagus Banana Cigar box Cooked plum Honey bush Leather Macadamia Mocha Pecan nut Raisins Saw dust Terpene Violet	Acacia Aniseed Black pepper Cedar wood Clove Mango Mimosa Papaya Pine Slate Sweet melon	Bees wax Chalky Cinnamon Flowers Grenadilla/ Granadilla Karoo scrub Orange blossom Straw	Butterscotch Flinty Herbal Litchi Marzipan Orange Smoky Wet wool	Grapefruit Honey suckle Yeasty	Almond Apricot Baked apple Dried fruit Guava Lemon Quince Vanilla	Apple Honey Peach Pear Pineapple



**Figure 4.2:** Illustrates the attributes that were cited 0 – 10 times over the 7 year investigation period (2008-2014).

Figure 4.2 A) shows the descriptors that have been cited 10 times or less, granadilla is the only attribute that is cited more than twice in a single year and this has only happened once over the 7 year investigation period. Besides for the granadilla there is no significant trend in the citation of attributes. Figure 4.2 B illustrates a similar trend, however since 2012, chalky has had an upward trend, being cited more each year. Figure 4.2 C illustrates that Straw, which was cited 4 times in 2010, was not cited at all during 2012 and 2013 and was only cited once in 2014 while beeswax which was not cited at all from 2008-2010 has been increasing in citation since 2011. Cinnamon and Orange Blossom in figure 4.2 D fluctuate annually as they are both cited twice in a given year and then not at all in the following. Apart from beeswax and chalky, which show an increasing trend in citation, all the other attributes are being cited randomly and in an inconsistent pattern.

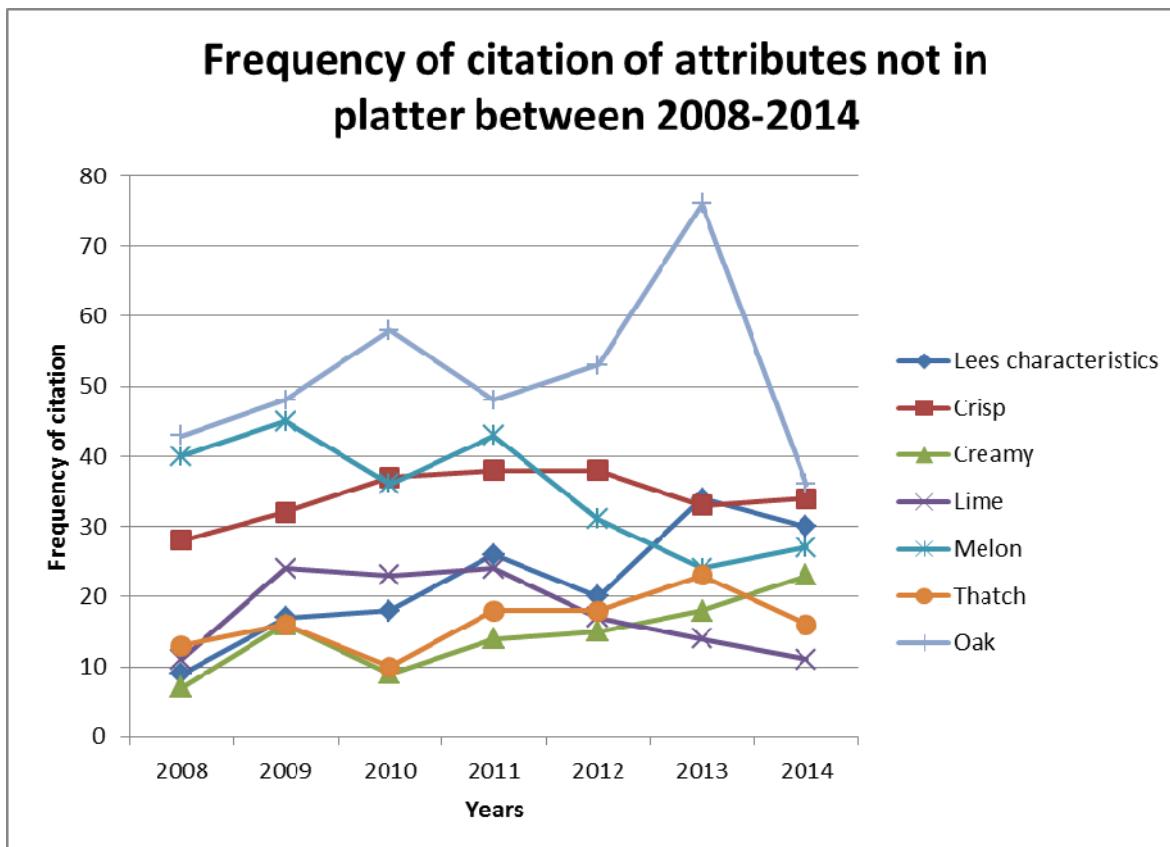
Fifty-four descriptors have been cited numerous times by the panel of experts and have not been included in the aroma wheel. Descriptors such as Oak and melon have been cited over 200 times, but there is no allocated slot for them in the current Chenin Blanc Aroma Wheel. All the descriptors that have been cited in the John Platter Wine Guide which are not on the existing aroma wheel have been tabulated (table 4.2) in the same format that was followed for table 4.1.

**Table 4.2:** Frequency of citation of all the words that are currently not on the Chenin Blanc Aroma Wheel using the data set spanning over the 7 year investigation period (2008-2014)

	Frequency of citation							
	5-10	11 – 20	21 – 30	31 – 40	41 – 50	50 – 100	More than 100	
Attributes/ descriptors	Apple Blossom Blossoms Boiled sweets Fennel Fig Fynbos Ginger Hazelnut Herbaceous Honeycomb Lanolin Stewed quince Stoney Tangerine	Biscuit Butter Caramel Custard Dried apricot Dried grass Dried peach Dust Kiwi Marmalade Nectarine Oatmeal Passion fruit	Bruised apple Hay	Botrytis Grape Grass Leaves Pithy Steely Toast Waxy White peach Woody	Fruit salad Sugary	-	Creamy Nuts Savoury Spice	Crisp Floral Lees characteristics Lime Melon Mineral Oak Ripe fruit Thatch

Figure 4.3 shows the frequency of citation of some of the attributes that were not mentioned in the John Platter Wine Guide between the years 2008-2014. The graph illustrates how certain descriptors have been mentioned more each year – such as creamy and lees characteristics. It also illustrates

how the descriptors that were once cited frequently have started to decrease over the past few years e.g. melon and lime. From 2008 through to 2013 the frequency of citation of Oak was increasing, but as can be seen in the graph the presents of oak in 2014 has greatly decreased. These descriptors were selected to illustrate that although they are not on the current aroma wheel, they are consistently present in South African Chenin Blanc wine.



**Figure 4.3:** Frequency of citation of the attributes in the John Platter Wine Guide that is not currently on the Chenin Blanc Aroma wheel

## 4.4 DISCUSSION

The sensory profile within the Chenin Blanc cultivar is very diverse and this can be seen in table 4.1. Words such as pineapple and litchi clearly indicate the tropical spectrum of the cultivar. There is also a floral spectrum with attributes/descriptors such as honey suckle being cited .This is to be expected as floral notes and tropical fruit are both signature elements in Chenin Blanc (Asimov, 2012). The emergence of flinty, chalky and slatey gives us a better insight into the possibility of a new style being developed and future work should be conducted to focus solely on the mineral aspect.

However, even though the sensory space is so diverse and represented by so many descriptors there is the possibility that, for South African Chenin Blanc, the existing aroma wheel does not give an accurate enough representation of the full sensory spectrum of South African Chenin Blanc wine. Figure 4.1 is the existing aroma wheel that is currently being used. The wheel attempts to cover the entire sensory spectrum of Chenin Blanc, but there are descriptors that should potentially be removed due to the fact that they were very rarely cited in 7 years' worth of data (2008-2014). Table 4.2 shows the frequency of the descriptors that have been cited in the John Platter Wine Guide that are currently on the aroma wheel.

The descriptors that have never been cited during the investigation period include Banana, cigar box and Mocha. These do not give a good representation of the sensory space within Chenin Blanc and should possibly be removed from the Chenin Blanc aroma wheel.

The category 1-4 starts to give an indication of the sensory space, but perhaps not enough to warrant a place on the wheel whereas category 5-10 gives a better indication. However, in the case of granadilla the frequency of citation is not accurate enough as granadilla and passion fruit are perceived as the same fruit and the dataset allowed two variations of grammar – Grenadilla and Granadilla.

The category 11-20 gives a clearer indication of the sensory space of South African Chenin Blanc. Even though the descriptors have not been cited that frequently, they do start to represent descriptors that could play a key role in style classification. The attributes/descriptors that have been cited more than 20 times warrant a spot on the wheel as they do represent the signature elements in Chenin Blanc e.g. floral.

Table 4.3 shows the frequency of citation of the descriptors that have been perceived by the expert panel, but are not included on the current aroma wheel. A large amount attributes/descriptors were not included in the table as they were not cited between 5 - 10 times – a cut off of 5 citations were set due to the fact that it was felt that they did not significantly represent the sensory space of South African Chenin Blanc yet. The category 11-20 has attributes/descriptors that would more likely be a good representation of the sensory space within Chenin Blanc as opposed to those cited less than 10 times. The dried fruits ( apricot and peach) gives an indication that perhaps there should a specific group allocated for dried fruits and possibly have the sub groups being the different fruit e.g. peach, pear, apple and apricot. It was also observed in the category the presence of passion fruit, which as previously stated, can also be perceived as granadilla/grenadilla. For this reason it might be beneficial to combine the two attributes/descriptors in the same way as “lemon/orange” was in the existing aroma wheel.

The bruised apple and hay that fall into the next category are new descriptors that have emerged. Previously the bruised apple would have to fall into the “apple” sub-group in the fruits. The hay

however would not be able to be associated with anything else. The same can be said for the dried grass in the category 11-20.

Looking at the other attributes/descriptors that have been cited more than 30 times, it is evident that they form a vital part of the sensory space for South African Chenin Blanc. The fact that 9 attributes/descriptors have been cited more than 100 times calls for a re-evaluation of the South African Chenin Blanc Aroma Wheel. The below mentioned points form the basis of what could constitute a new Chenin Blanc aroma wheel

- Removing all attributes that have been cited less than 5 times in the 7 year investigation period (2008-2014) from the existing Chenin Blanc wheel. This includes all the attributes from asparagus – sweet melon in the first two categories of table 4.1.
- The 11 groups making up the new aroma wheel would be:
  - Wood
  - Nuts
  - Floral
  - Spices
  - Mineral
  - Fresh Vegetative
  - Dried Vegetative
  - Sweet Associated
  - Other
  - Yeasty
  - Fruits
    - Stone fruit
    - Citrus
    - Tropical
    - Dried fruit
    - Baked/Cooked fruit

**Table 4.3:** The table shows the existing and new groups of the proposed Chenin Blanc aroma wheel including the attributes/descriptors that have been added and removed.

Groups and sub-groups	Attributes/descriptors added	Attributes/descriptors removed
Wood	Oak Thatch Toast Woody	Mocha Cigar box Saw dust Cedar wood Pine
Nutty	Mixed nuts Hazelnut	Pecan Macadamia
Floral	Apple Blossom Blossoms	Violets Mimosa
Earthy	Waxy Lanolin	Leather
Spicy	Lemon pepper Ginger	Black pepper Clove Aniseed
Mineral	Steely Dust Stoney	Terpene Slate
Fresh vegetative	Green pepper Cooked vegetables Grass Herbaceous Fennel Leaves	Honey bush Asparagus Acacia
Dried vegetative	Straw Fynbos Hay Dried grass	None
Other	Creamy Crisp Sugary Savoury Butter	Cooked plum Raisin

**Table 4.3 (cont.)**

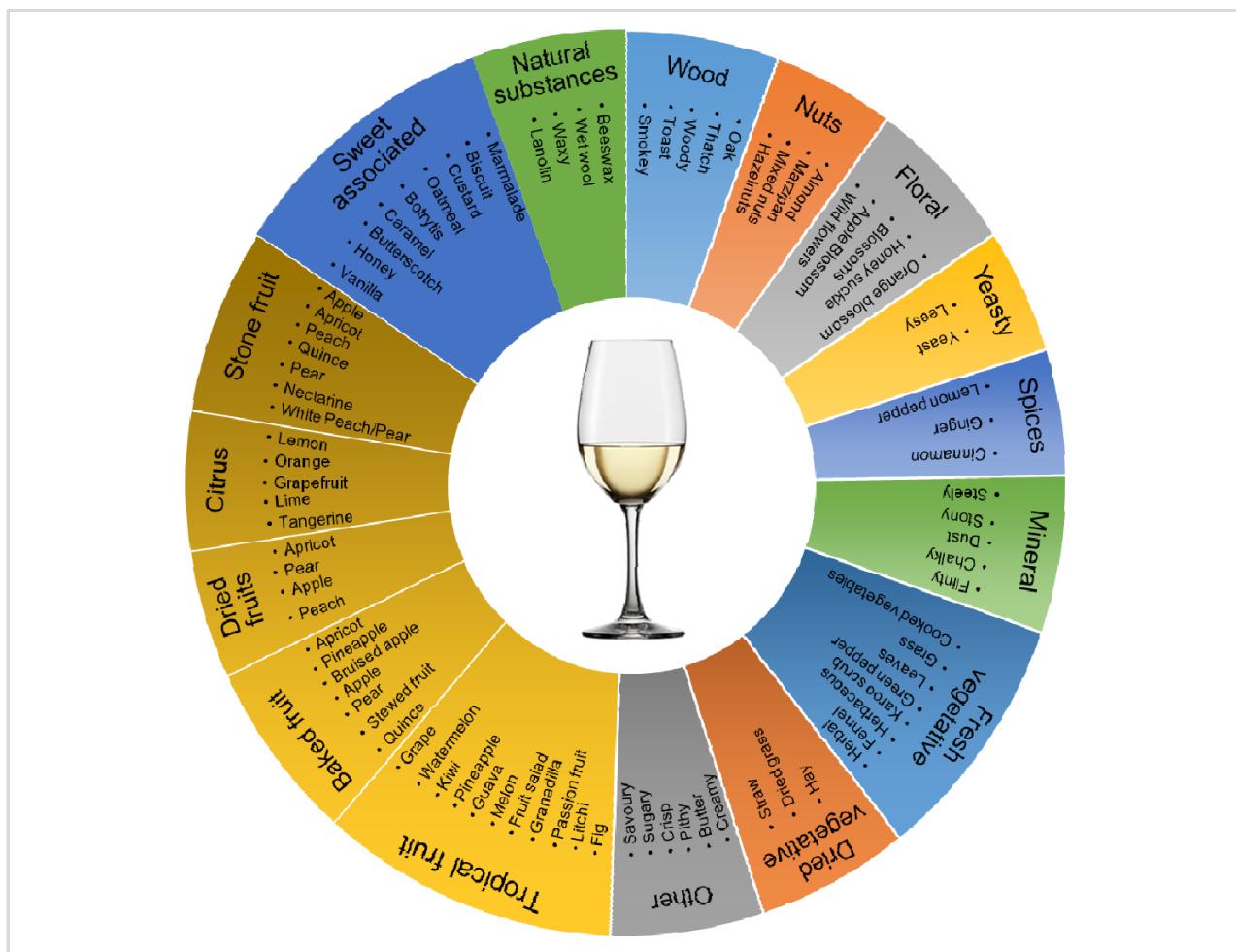
Fruits	Attributes/descriptors added	Attributes/descriptors removed
1. Stone fruit	Nectarine White peach White pear Pithy	None
2. Citrus fruit	Lemon/orange becomes single entities Lime Tangerine	None
3. Tropical fruit	Fruit salad Grape Melon Watermelon Kiwi Fig Passion fruit	Banana Papaya Sweet melon Mango
4. Dried fruit	Apricot Pear Apple Peach	None
5. Baked/cooked fruit	Stewed fruit Apple Pear Apricot Pineapple Bruised apple Stewed quince	None
6. Yeast	Yeast Leesy	
7. Sweet Associated	Botrytis Boiled sweets Honeycomb Biscuit Custard Marmalade Oatmeal	

## 4.5 CONCLUSIONS

---

One of the main factors influencing beneficial consumer behaviour in the wine market is the confusion that consumers are currently experiencing with regard to understanding the sensory attributes that make up the wine and with this comes complexity. The high numbers of attributes/descriptors that cover the sensory space in South African Chenin Blanc makes it hard for the average consumer to

mentally profile the wines based on information on the bottle. Adding to this, we have also found that the current Chenin Blanc aroma wheel does not accurately represent South African Chenin Blanc. This could also be due to the evolution of styles and wine making techniques. Future work would need to be done in a few years to re-evaluate the Chenin Blanc aroma wheel as it is evident that as time passes so does the evolution of the wines. Figure 4.4 shows how the Chenin Blanc Aroma wheel would look if using the inclusion/exclusion criterion of at least 5 citations over the 7 year investigation period.



**Figure 4.4** New validated Chenin Blanc aroma wheel using the John Platter Wine Guide as a data set over the 7 year investigation period (2008-2014)

## 4.6 LITERATURE CITED

---

- Antwerpen L Van. (2012). Chemical and Sensory profiling of dry and semi-dry South African Chenin blanc wines.
- Antwerpen L Van, Hanekom E, Bester I, Muller N, Tredoux A, Africa S, et al. (2013). Feedback : Comprehensive three-year ( 2010 – 2012 ) research project on SA Chenin blanc wine : Part 1. (March):2012–3.
- Asimov E. (2012). South African Chenin Blancs Aren't What They Used to Be - Wines of The Times - NYTimes.com. [cited 2014 Oct 23]; Available from: [http://www.nytimes.com/2012/07/18/dining/reviews/south-african-chenin-blancs-arent-what-they-used-to-be-wines-of-the-times.html?pagewanted=all&\\_r=0](http://www.nytimes.com/2012/07/18/dining/reviews/south-african-chenin-blancs-arent-what-they-used-to-be-wines-of-the-times.html?pagewanted=all&_r=0)
- Bester I. (2011). Classifying South African Chenin blanc wine styles [Internet]. [cited 2014 Aug 13]. Available from: <https://ir1.sun.ac.za/handle/10019.1/18048>
- Botha H. (2010). and Chemical Profiling of South African Natural Sweet Chenin blanc wine.
- Campo E, Do B V., Ferreira V, Valentín D. (2008). Aroma properties of young Spanish monovarietal white wines: A study using sorting task, list of terms and frequency of citation. *Aust J Grape Wine Res.* **14**(2):104–15.
- Fisheries AF&. (2013). A PROFILE OF THE SOUTH AFRICAN WINE MARKET VALUE CHAIN.
- Hanekom E. (2012). Chemical , sensory and consumer profiling of a selection of South African Chenin blanc wines produced from bush vines.
- Jolly NP, Hattingh S. (2001). A Brandy Aroma Wheel for South African Brandy \*. **22**(1):27–8.
- Lawrence N. (2012). Volatile metabolic profiling of SA Chenin blanc fresh and fruity and rich and ripe wine styles : Development of analytical methods for flavour compounds ( aroma and flavour ) and application of chemometrics for resolution of complex analytical measurement.
- Platter. (2015). Platter's - About Platter's South African Wine Guide [Internet]. [cited 2015 Apr 14]. Available from: <http://www.wineonaplatter.com/about>
- WOSA. (2014). Wines of South Africa [Internet]. [cited 2014 Oct 27]. Available from: [http://www.wosa.co.za/sa/stats\\_worldwide.php](http://www.wosa.co.za/sa/stats_worldwide.php)

# Chapter 5

---

## GENERAL DISCUSSION AND CONCLUSION

## 5. GENERAL DISCUSSION AND CONCLUSIONS

### 5.1 GENERAL DISCUSSION AND CONCLUSIONS

---

Chenin Blanc is the most planted grape cultivar in South Africa (SA) and the quality of single cultivar wines produced from this grape has recently gained international recognition. However, the complete sensory domain of the cultivar is not yet well understood. The factor that probably contributed most to this situation is a lack of a large enough number of wines that could be considered representative of this wine style category. A considerable amount of research has been done on the chemical properties and resulting wine styles of both Sauvignon Blanc and Chardonnay which has assisted both grape cultivars to establish identifiable cultivar specific characteristics. The unique green pepper extreme accompanied by a complete and opposite tropical style has been a critical aspect in consumers identifying and understanding the different styles of Sauvignon Blanc. The same applies to Chardonnay wine styles that exhibit either a wooded, buttery profile or a more subtle soft fruit one. However, the same stylistic understanding does not apply to Chenin Blanc. This is ironic due to the fact that it is the most planted grape cultivar but according to Antwerpen (2012), was the least researched wine variety in SA at the time. This neutral grape cultivar does not contain methoxypyrazines and therefore lacks certain flavour compounds that could possibly give it these unique identifiable characteristics.

Previous research conducted at the University of Stellenbosch was focused on evaluating the style classifications that have been set forward, as guidelines, by the Chenin Blanc Association (CBA). Results showed that for dry and semi-sweet Chenin Blanc wines there were only two clearly differentiating style categories, namely wood and un-wooded. Additional work was also done on identifying the full sensory spectrum or sensory space of Chenin Blanc by using a small sample set of commercial wines. All the research conducted thus far has used relatively small data sets comprising of commercial wines and ranging in size between 15 and 20 wines with a focus on sensory and chemical characterisation. Acquiring a larger sample set would result in logistical complications and sensory evaluation by a trained panel also has its limitations such as, amongst others, sensory fatigue e.g. nose saturation and exhaustion (Campo *et al.*, 2008; Johnson *et al.*, 2013). There was however a need for an in-depth study into the sensory attributes that characterised Chenin Blanc and also to validate the style classification guidelines using a large data set.

The novel approach of evaluating style classification in addition to identifying the full sensory spectrum of Chenin Blanc using a large dataset was dependant on the acquisition of a large dataset. The John Platter Wine Guide, which is SA's largest wine document, is comprised of more than a thousand wines produced each year, across regions, price points and wine styles. This document or wine annual contains detailed information of each wine and also provides a sensory description that is generated by a panel of experts. The methodology behind how the sensory

description is generated is not conventional and therefore does not follow a structured format. The panel of experts are not directed on what to say nor are they influenced by a set of descriptors. However, even with the complexity of the unstructured information, with the correct and appropriate data processing tools the electronic document received was standardised into a dataset that comprised of 2746 Chenin Blanc wines over the selected 7 year investigation period of 2008-2014.

This study provides a stepwise methodology for processing a large dataset consisting of text. Once the electronic document was received a large amount of standardisation had to be done which included the removal of text data that did not represent a sensory attribute. Subsequently, variations of a single concept had to be standardised to one term e.g. appley and apple-like was standardised to apple. Further data mining processes including the removal of negative concepts also had to be concluded before a complete structured dataset was obtained and could be statistically analysed

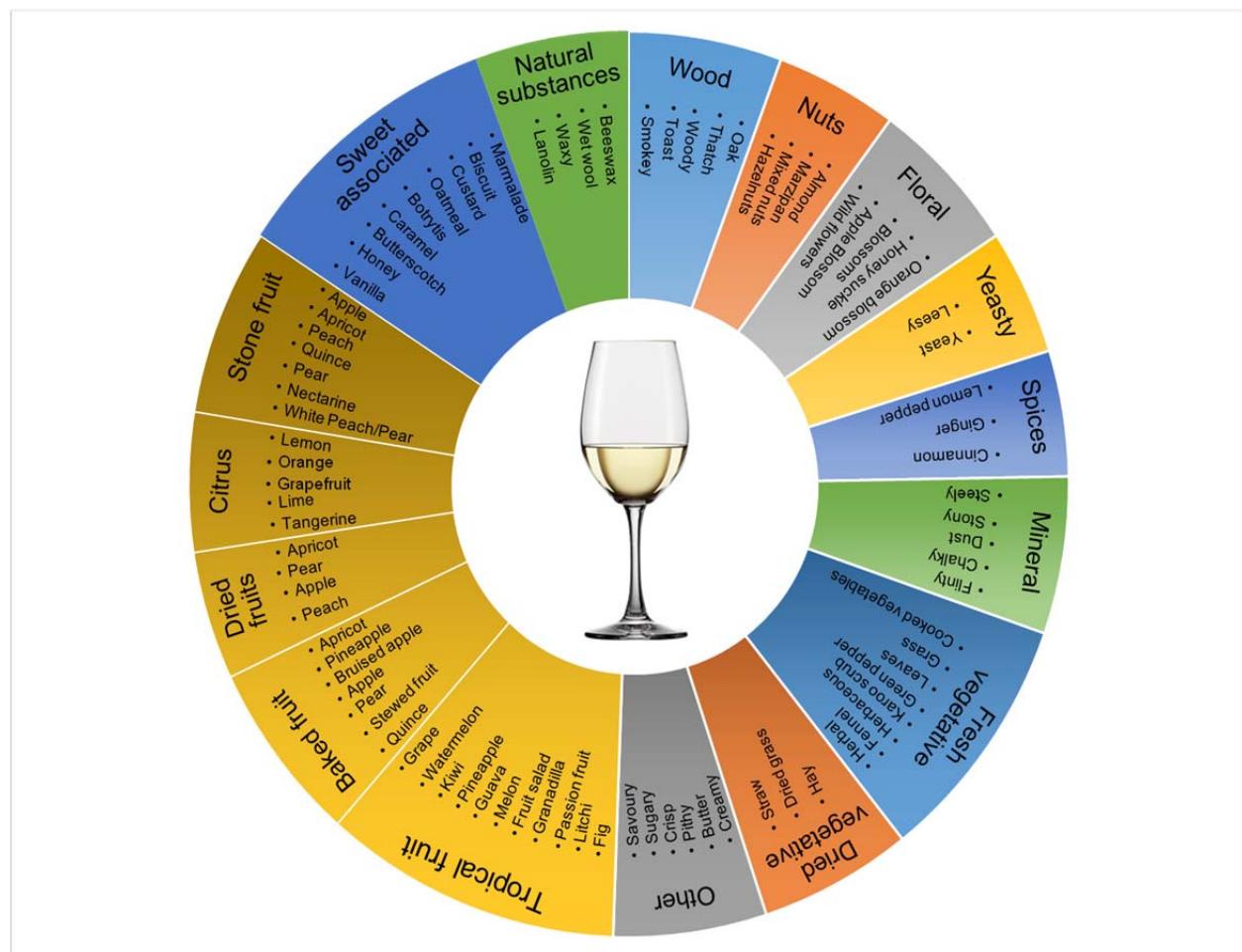
With a key problem of acquiring a large dataset on Chenin Blanc being solved this unique and novel approach was focused on characterising and identifying the full sensory spectrum of South African Chenin Blanc using a data mining approach applied to a large dataset of commercial wines. The results showed that a wooded and an unwooded style category could be identified which is in agreement with previous research findings (Bester, 2011; Antwerpen, 2012; Hanekom, 2012). However, from the dendrogram processed for style classification it was evident that within the unwooded category there is a branched family or sub-group that contains heavier, more sweet associated attributes e.g. quince, while also having savoury notes

The main aim of identifying the full sensory spectrum of Chenin Blanc was also achieved, but it must be noted the only attributes that were cited more than 50 times over the investigation period were used, as it was considered to more accurately characterise the sensory space of Chenin Blanc.

The identification of the full sensory spectrum within Chenin Blanc yielded interesting findings. There were certain attributes, such as oak, that were cited considerably more times than others over the selected 7 year investigation period and under closer inspection it was found that there were some attributes that were only mentioned once or twice. These findings were the foundation to the objective of updating the current Chenin Blanc Aroma Wheel.

The Chenin Blanc Aroma Wheel was formulated in 2005 by a few members of the Chenin Blanc association. Upon further investigation it was found that the aroma wheel, which is widely used in SA, was not scientifically developed for South African Chenin Blanc, but was developed as a generic white wine aroma wheel that was thought to be closely related to Chenin Blanc. Since the development of the wheel in 2005 it has never been validated or re-evaluated. This study suggests an updated version of the Chenin Blanc aroma wheel based on a large dataset representative of South African Chenin Blanc wines. Using the dataset that was compiled for characterisation of the sensory spectrum of Chenin Blanc, a contingency table was calculated on

all the sensory attributes that were mentioned during the selected 7 year investigation period. All the attributes that are currently on the Chenin Blanc Aroma Wheel where analysed to see whether a place on the wheel was validated. It was found that out of the 55 attributes that make up the wheel, 13 of them had not been cited at least once during the investigation period while 11 other attributes were only cited between 1-4 times. Furthermore, there were 13 attributes that were cited more than 50 times over the investigation period that were not included in the current aroma wheel. Attributes such as 'oak', 'melon' and 'crisp' were cited well over 200 times. These findings led to the formation of a new Chenin Blanc Aroma Wheel that is specific for the South African cultivar. All the attributes currently on the aroma wheel that were mentioned less than five times were removed, while attributes that were not initially on the aroma wheel were included if they were cited more than five times. Additional segments on the aroma wheel were made to include Dried fruit, Natural substances, Yeasty, Dried vegetative, Baked fruit and a segment termed Other. The proposed aroma wheel can be seen as figure 5.1



**Figure 5.1** New validated and reconstructed Chenin Blanc Aroma Wheel based on data retrieved from the John Platter Wine Guide during the selected 7 year investigation period (2008-2014)

To conclude, the work presented in this thesis was aimed at attempting to characterise SA's most planted grape cultivar. The study has successfully identified the full sensory spectrum of South African Chenin Blanc in addition to also identifying possible stylistic cues (attributes) or terms that is important in the wine industry. These cues also give South African Chenin Blanc an identifiable character although as previously said they are style dependant. The data processing methodology can also be used as the foundation for future work and similar studies can be conducted on other grape cultivars as well. Future studies could also focus on doing a more in-depth focus on the concept of regionality as this study only briefly touched on it and also including lattice relationships to explain concepts and variables such as wine of origin, estate wines or the difference between young and old vines.

## 5.2 LITERATURE CITED

---

- Antwerpen L Van. (2012). Chemical and Sensory profiling of dry and semi-dry South African Chenin blanc wines.
- Bester I. (2011). Classifying South African Chenin blanc wine styles [Internet]. [cited 2014 Aug 13]. Available from: <https://ir1.sun.ac.za/handle/10019.1/18048>
- Campo E, Do B V., Ferreira V, Valentin D. (2008). Aroma properties of young Spanish monovarietal white wines: A study using sorting task, list of terms and frequency of citation. *Aust J Grape Wine Res.* **14**(2):104–15.
- Hanekom E. (2012). Chemical , sensory and consumer profiling of a selection of South African Chenin blanc wines produced from bush vines.
- Johnson TE, Hasted A, Ristic R, Bastian SEP. (2013). Multidimensional scaling (MDS), cluster and descriptive analyses provide preliminary insights into Australian Shiraz wine regional characteristics. *Food Qual Prefer* [Internet]. Elsevier Ltd **29**(2):174–85. Available from: <http://dx.doi.org/10.1016/j.foodqual.2013.03.010>

## ADDENDUM A

Production areas defined in terms of the wine of origin scheme.

Geographical unit: Western Cape

Region	District	Ward
Breede River Valley	Breedekloof	Goudini, Slanghoek
	Robertson	Agerkliphoogte, Bonnievale, Boesmanrivier, Eilandia, Hoopsrivier, Klaasvoogds, Le Chasseur, McGregor, Vinkrivier
	Worcester	Aan-De-Doorns (repeated 8 June 2014), Hex River Valley, Nuy, Scherpenheuvel
Cape South Coast	Cape Agulhas	Elim
	Elgin	No ward
	Overberg	Elandskloof, Klein River, Greyton, Theewater
	Plettenberg Bay	No ward
	Swellendam	Buffeljags, Malgas, Stormsvlei
	Walker Bay	Bot River, Hemel-en-Aarde Ridge, Hemel-and-Aarde Valley, Sunday's Glen, Upper Hemel-and-Aarde Valley, Stanford Foothills
	No district	Herbetsdale, Napier, Stilbaai East
Coastal Region	Cape Peninsula	Constantia, Hout Bay
	Cape Point (repeated 3 January 2014)	No ward
	Darling	Groenekloof
	Franschhoek/Franschhoek Valley	No Ward
	Paarl	Simonsberg-Paarl, Voor Paardeberg
	Stellenbosch	Banghoek, Bottelary, Devon Valley, Jonkershoek Valley, Papegaaiberg, Polkadraai Hills Simonsberg-Stellenbosch
	Swartland	Malmesbury, Riebeekberg, St Helena Bay
	Tulbach	No ward
	Tygerberg	Philidelphia, Durbanville
	Wellington	No ward

<b>Region</b>	<b>District</b>	<b>Ward</b>
Klein Karoo	Calitzdorp	No ward
	Langeberg-Garcia	No ward
	No district	Montagu, Outeniqua, Tradouw, Tradouw Highlands, Upper Langkloof
Olifants River	Citrusdal Mountain	Piekenierskloof
	Citrusdal Valley	No ward
	Lutzville Valley	Koekenaap
	No district	Bamboes Bay, Spuitdrift, Vredendal
No Specified Region	Ceres Plateau	Ceres
	No district	Cederberg, Lamberts Bay, Prins Albert Valley