An investigation into consumer buying behaviour of alcoholic beverages in specialist retailer outlets in South Africa

Deborah Vigario

Thesis presented in fulfilment of the requirements for the degree of
MCom (Quantitative Management)
in the Faculty of Ecomnomic and Management Sciences at Stellenbosch University

Supervisor: Professor Hannelie Nel

December 2019

Declaration

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

Date: December 1, 2019

ii

Abstract

The alcoholic beverage industry in South Africa is highly competitive, with thousands of brands competing for market share. Companies investing in advertising want to understand if they are getting a return on their investment and also ultimately, whether the advertising contributes towards changing consumers' buying behaviour and leads to increased revenue. The objective of this study is to examine the effectiveness and efficiency of branded advertising in specialist retailer outlets in the alcoholic beverage industry of South Africa, also making clear recommendations on how these methods can be used in the industry and be integrated into marketing strategy. The alcoholic beverage industry in South Africa is cluttered with brands and products. There is a lack of recorded information on advertising campaigns. The product life cycle (PLC) methodology is used to segment the market and determine a product's competitor set. The data envelopment analysis (DEA) method is used to determine the advertising efficiency. Regression analysis is used as a benchmark method to determine effectiveness. The results show that using the PLC and DEA methods in combination have meaningful results and meet the set objectives. It is also possible to overcome the practical industry problems of noise, clutter and availability of data, while providing market insights. The study will cover alcoholic beverages available in the formal retail market from January 2013 to December 2017.

Opsomming

Die alkoholiese drankbedryf in Suid-Afrika is uiters mededingend, met duisende handelsmerke wat meeding om markaandeel. Ondernemings wat in advertensies belê, wil verstaan of hulle 'n opbrengs op hul belegging kry, en uiteindelik ook of die advertering daartoe bydra om verbruikers se koopgedrag te verander en tot verhoogde omset lei. Die doel van hierdie studie is om die doeltreffendheid en doelmatigheid van handelsmerkadvertensies in spesialiswinkels in die alkoholiese drankbedryf in Suid-Afrika te ondersoek, en ook duidelike aanbevelings te maak oor hoe hierdie metodes in die bedryf gebruik kan word en in die bemarkingstrategie geïntegreer kan word. Die alkoholiese drankbedryf in Suid-Afrika is gelaai met handelsmerke en produkte. Daar is 'n gebrek aan aangetekende inligting oor advertensieveldtogte. Die produksielewensiklus (PLC) -metodologie word gebruik om die mark te segmenteer en die versameling van mededingers van 'n produk te bepaal, en die DEA-metode word gebruik om die advertensiedoeltreffendheid te bepaal. Regressie-analise word gebruik as 'n maatstaf om effektiwiteit te bepaal. Die resultate toon dat die gebruik van die PLC- en DEA-metodes in kombinasie noemenswaardige resultate lewer en aan die gestelde doelwitte voldoen. Dit spreek ook die praktiese industrieprobleme van geraas, rommel en dataverskaffing aan, terwyl dit insigte in die mark bied. Die studie dek alkoholiese drank wat beskikbaar is in die formele kleinhandelsmark vanaf Januarie 2013 tot Desember 2017.

vi

Acknowledgements

The author wishes to acknowledge the following people for their various contributions towards the completion of this work: The author of this paper wishes to acknowledge the following people for their various contributions and support towards the completion of this work:

- Professor Nel, for your unwavering guidance, support, care and patience;

- The Logistics department for the usage of facilities and support;

- Family for your continuous words of encouragement, interest and prayers; and

- All glory to our Heavenly Father, for granting me the opportunity and perseverance.

viii

viii

Contents

List of Abbreviations

Abbreviation	Definition
AB InBev	Anheuser-Busch InBev
ATL	Above the line advertising
AUP	Average unit price
BCG	Boston Consulting Group: Management consulting
BH	Brand-health or equity
BTL	Below the line advertising
CAGR	Compound annual growth rate
CFA	Confirmatory factor analysis
DEA	Data envelopment analysis
DGB	Douglas Green Bellingham
EFA	Exploratory factor analysis
FABs	Flavoured alcoholic beverages
FMCG	Fast moving consumer goods
GDP	Gross domestic product
JSE	Johannesburg stock exchange
KWV	Koöperatieve Wijnbouwers Vereniging (Co-operative Winemakers Union)
MAT	Moving annual trend
PCA	Principle component analysis
PLC	Product life cycle
RGBC	Really great brands company
RMSP	Root Mean Square Off-Diagonal Partials
RMSR	Root Mean Square Off-Diagonal Residuals
RTD	Ready to drink products
SA	South Africa
TV	Television

List of Figures

List of Tables

List of Symbols

Regression Analysis

Symbol	Definition
$\hat{\beta}_0$	The estimated intercept coefficient
$\hat{\beta}_i$	The estimated coefficient of the independent variable
$CE_{A,B}$	The cross price elasticity of product B on product A
ϵ_i	The error term
e_A	The own elasticity of product A
e_p	The price elasticity of demand
k	The number of independent variables that are needed to predict Y
n	The number of observations
$\%\Delta\mathbf{P_A}$	The percentage change in price for product A
$\hat{P}_A$	The average price of product A
$\hat{Q}_A$	The average quantity demanded of product A
$\%\Delta\mathbf{Q_A}$	The percentage change in quantity demanded of product A
R^2	The coefficient of determination
R_a^2	The adjusted R^2 takes into account the number of observations, n in the sample and the k number of independent variables
s_e	The standard error of estimates
$S_{\hat{\beta}_1}$	Standard error of $\hat{\beta}_1$
SSE	The sum od squares of the error terms
SSR	The sum of squares due to the regression
SST	The sum of squares of the total variation
VIF_Δ	The variance inflation factor, provides a ratio to measure multicollinearity
w	The current week ($w-1$ is the previous week)
x_i	The value of the independent variable of the i_{th} observation when using simple linear regression
x_{ij}	The value of the j_{th} independent variable for the i_{th} data point when using multiple linear regression
$\bar{x}$	The average value of all x_i's
y_i	The value of the dependent variable of the i_{th} observation when using simple linear regression
$\hat{y}_i$	The estimated value of y_i
$\bar{y}$	The average value of all y_i's

List of Symbols

Factor Analysis

Symbol	Definition
A	The factor score coefficient matrix $(p \times p)$
B'	The factor pattern in the matrix algebra common factor model $(q \times p)$
b_{kj}	The regression coefficient of the kth common factor for predicting the jth variable in the common factor model
$b_{1j}^2 + b_{2j}^2 + \ldots + b_{qj}^2$	Represent the communality of the jth variable in the common factor model
E	The unique factor in the matrix algebra common factor model
e_{ij}	The value of the ith observation on the jth unique factor in the common factor model
i	The ith observation in the set, $i = 1, \ldots, n$
j	The jth variable in the set, $j = 1, \ldots, p$
k	The k_{th} common factor, $k = 1, \ldots, q$
n	The number of observations
p	The number of variables
p_{ij}	The partial covariance matrix for the all the variables in the Kaiser-Meyer-Olkin measure
q	The number of common factors in the common factor model
S	The variance covariance matrix of the observed variables in the matrix algebra common factor model $(p \times p)$
s_{ij}	The sum of the squared correlation matrix for all the variables in the common factor model
s_{jk}	The covariance between the jth and k^{th} variables implied by the common factor model
$\mathbf{U}^2$	The diagonal covariance matrix of the unique factors in the matrix algebra common factor model
$[\mathbf{U}^2]_{jj}$	The uniqueness of the jth variable in the common factor model
X	The matrix of factor scores in the matrix algebra common factor model $(n \times q)$
x_{ik}	The value of the ith observation on the kth common factor in the common factor model
y_{ij}	The value of the ith observation on the jth variable in the common factor model
Y	The variable matrix $(n \times p)$
Z	The standardised Y variable matrix $(n \times p)$

List of Symbols

Cluster Analysis

Symbol	Definition
C	Cluster in the measurement of dissimilarity in cluster analysis
C_M	Formed from merging clusters C_K and C_L in the measurement of dissimilarity in cluster analysis
D_{KL}	The distance or dissimilarity measure between clusters C_K and C_L in the measurement of dissimilarity in cluster analysis
k	The number of cluster centres in non-hierarchical cluster analysis
N_K	The number of cases in C_K in the measurement of dissimilarity in cluster analysis
n	The number of cases in a cluster
p	The number of variables describing the cases characteristics in a cluster
p_i	The number of cases in the ith cluster in non-hierarchical cluster analysis
R^2	Is the coefficient of determination, this is a measure of determining the variation when clusters are formed
S^l	The set of n cases in l clusters of ordered pairs in cluster analysis
w	Is an estimation of the dimensionality of the variation between clusters
X_{ijk}	The value for variable k in case j belonging to cluster i in cluster analysis error sum of squares
$\tilde{x}_K$	The mean vector of cluster C_K in the measurement of dissimilarity in cluster analysis
$(\parallel x_i - c_j^{(t)} \parallel)$	The Euclidean distance between x_i and the cluster centre $c_j^{(t)}$ at iteration t

List of Symbols

Data Envelopment Analysis

Symbol	Definition
DMU_i	The decision making unit i, or entity i
I_{ij}	The value of the DMU_i on the input variable j
n_I	The number of input variables
n_O	The number of output variables
O_{ij}	The value of the DMU_i on the output variable j
v_j	The non-negative weight assigned to input variable j
w_j	The non-negative weight assigned to output variable j

CHAPTER 1

Introduction

Contents

A pertinent question for any company will be to determine whether the investment in advertising is valid and actually contributes towards changing consumers' buying behaviour and leads to increased revenue or market share. The fast moving consumer goods environment is very competitive and expensive to compete in. At the same time it becomes necessary for companies to reduce costs and become more efficient to obtain the required shareholder's return. The objective of this study is to examine the effectiveness and efficiency of branded advertising in specialist retailer outlets in the alcoholic beverage industry of South Africa. Branded advertising is the foundation of all brand building efforts and includes activities that seek to persuade consumers to change their buying behaviour or reinforce it [39]. The methodology followed in the study will be to investigate the relationships and insights that the product life cycle methodology and data envelopment analysis provide. These insights and relationships will be benchmarked against those that multiple regression analysis uncovers. The study will cover alcoholic beverages available in the formal retail market from January 2013 to December 2017.

1.1   Background

The alcoholic beverage industry in South Africa has had conservative growth in recent years, with a compound annual growth rate of 1.71% from 2010 to 2014 [43]. As companies in the industry strive to generate profits and shareholders' value under harsh economic conditions the question of how marketing activities affect revenue becomes more relevant. The global financial crisis of 2008 and the resulting recession that followed caused a reduction in gross domestic product (GDP) and this has elevated the competitiveness for brands, as consumers feel more pressure on their disposable income.

The alcoholic beverage landscape is segmented into two main sectors, on-consumption and off-consumption. On-consumption is the sale of alcohol that is to be consumed on the premises where it was purchased for example: bars, restaurants and taverns. Off-consumption is the sale of alcohol that is taken to another place to be consumed for example: purchased at a bottle store and consumed at home. Marketing campaigns are structured to reach consumers within these two broad categories. Marketing campaigns need to be measured to determine their effectiveness. The off-consumption environment is highly competitive between brand marketers. This competitive landscape makes determining and quantifying the effect that advertising has on a brand's revenue and the efficient mix of advertising a necessity for sustained market growth or dominance.

A brief history of the South African alcoholic beverage industry

The current landscape of the South African alcoholic beverage industry is heavily shaped by the past, but this is changing due to the entrance of global investments and opportunity. The South African alcoholic beverage industry can be broadly categorised to include the manufacturing, marketing and distribution of beer, wine and spirits. The industry's architecture has its roots in the past, over two decades ago, when the government had a high tolerance for monopolies in the market. The country was also isolated from trade, this led to a small number of companies with large market shares. This highly clustered environment means that a single company controls the majority of the market. In 2016 the research firm Nielson reported that 5 companies accounted for 82% of the value share of the total market [51], they are: AB InBev, Distell, Diageo, Heineken and Pernot-Ricard.

The spirits and beer segments are both categorised by a small number of companies producing, marketing and distributing the liquor. The local industry does showcase a significant number of market leading brands for example: Amarula Cream Liqueur and Castle Larger. There are also many spirits that are produced locally under international licence for example: Gordon's Gin and Smirnoff Vodka. The wine segment is different however, it is highly diversified with large established producers and a large number of independent wine estates and co-operatives.

The general beverage sector includes the South African alcoholic beverage industry, this can be broken down into three main segments which are: wine cellars, ready-to-drink (RTD) products, breweries or manufacturers and spirits distilleries. The wine segment can be further broken down into categories which are: sparkling wine, fortified wine, unfortified wine and perle wine. The RTD segment can be split into two main categories: beer and flavoured alcoholic beverages (FABs). The beer category can be further divided into sorghum beer, also know as 'traditional African beer' and malt beer. Craft beer is a new sub segment of the malt beer category. The FABs category includes a broad variety of products such as: spirit coolers, alcoholic energy drinks and ciders. The spirits segment is broken down into well defined categories which are: brown spirits (whisky, brandy, rum and cognac), white spirits (vodka, gin and cane) and liqueurs [33].

The South African alcoholic beverage industry faces a number of challenges that are consistent across the South African commercial landscape. The South African economy is a small open economy, that follows the growth of the global economy. Since 2011 a gap has started to open up between the local and global economic growth. This is due to constraints in the local environment and this gap is not forecasted to contract.

The leading factors to slowed growth in South Africa are: high inflation, low profitability and business confidence which remains low for the majority of non-durable goods retailers, high unemployment levels, constraints on infrastructure, availability of electricity and political uncertainty [26]. The South African alcoholic beverage industry also faces issues around competitiveness due to the global movement towards 'free trade', as well as legislation on black economic empowerment. The South African alcoholic beverage industry competes with other sectors for consumers' disposable income [26]. Over the last few decades the share of disposable income being spent on alcohol has declined. The main contributor to this decline has been the increase in spend on telecommunications [50]. The highly capital-intensive nature of the manufacturing, marketing and distribution of alcohol produces barriers to entry.

The market leaders

Figure 1.1 shows the time line of the establishment of the alcohol producers, distributors and manufacturers who are market leaders today.

Year of establishment of a company in the South African alcoholic beverage industry

1848	Edward Snell & Co
1895	South African Breweries
1918	KWV
1925	Stellenbosch Farmers Winery
1942	Douglas Green & Co
1945	Distillers Corporation
1947	Namaqua Wines
1985	Vinimark
1991	Douglas Green Bellingham
1994	Pernot-Ricard
1996	Really Great Brand Company
1997	Meridian Wine Merchants
1999	Halewood International South Africa
2000	Distell
2002	SABMiller
2004	Brandhouse
2015	Heineken, Namibian Breweries and Diageo
2016	AB InBev

FIGURE 1.1: *Year of establishment of companies in the alcoholic beverage industry in South Africa*

Over 150 years ago Edward Snell was on board a ship destined for Argentina, but the ship in distress docked in Cape Town. In 1848 Edward Snell moved to Kwa-Zulu Natal and established Edward Snell & Co, the company traded in imports and exports. In 1906 Vernon Hooper, Edward Snell's great nephew, bought out the interest from the Snell family. The company is still a family run business today. Edward Snell & Co is involved in the production and marketing of various spirits. Some of the brands in the portfolio include: Grants whisky, Glenfiddich whisky, Remy Martin cognac, Wellington brandy, Russian Bear vodka, Skyy vodka, Campari and Cape to Rio cane [22].

Just over 120 years ago in 1895, South African Breweries was founded in Johannesburg. Today they are the biggest manufacturer in the South African beer category. The former South African Breweries acquired the international Miller Brewing Co in 2002 and after this merger SABMiller became the second largest brewer in the world. In South Africa the key SABMiller brands include: Castle, Carling Black Label and Hansa Pilsener, internationally they own more than 150 brands. In 2016 AB InBev acquired SABMiller and they are now the biggest brewer in the world and poised to be one of the world's biggest fast moving consumer goods (FMCG) companies. The sorghum beer category can be broken down into two main competitors: United National Breweries and informally (home) brewed products. The clear or malt beer category is dominated by the former SABMiller, now AB InBev (90% market share), with some competition from Heineken on the premium end. In South Africa, AB InBev has ownership and control over its distribution network as well as raw material supply, the South African alcoholic beverage industry is dominated by the former South African Breweries [60].

The South African wine industry is far older than the first large co-operative, Koöperatieve Wijnbouwers Vereniging (KWV) or in English Co-operative Winemakers Union, which was established over 100 years ago. The South African wine industry started with the exploration of the Cape of Good Hope by the Dutch East India Company, with the first bottle of wine being produced in 1659. Jan Van Riebeek oversaw the plantation of vineyards, believing that eating grapes and drinking wine would save sailors from scurvy and other diseases during long voyages. Some 15 years later the Cape Governor, Simon van der Stel, who had a passion for wine, recruited French winemakers and purchased large farms of his own. Simon van der Stel also imported many different varietals of wine making grapes for his farms and also instituted a high standard of quality for wine produced in the Cape [71].

After Simon van der Stel's death, the wine industry in the Cape declined in quality and variety. Over the next 200 years the wine industry faced many challenges with farmers rather planting fruit trees and other more profitable crops. The farmers who continued to produce wine started planting high quantity yielding grape varietals. This lead to an over supply of wine and a low demand due to the quality.

In 1918 KWV was formed to defend the farmers through collective bargaining. KWV set a minimum price for wine and put in place a guarantee for farmers that they would purchase any excess wine that the farmers could not sell. This changed the South African wine industry as KWV's standing policy to purchase any excess wine reinforced a trend in very poor quality wine being produced in large quantities. KWV in turn was not using this low quality wine to sell but rather as an input into their Brandy production. KWV also gave the farmers an advantage in that machinery and technical knowledge could be pooled [71].

In 1925 Stellenbosch Farmers Winery was formed and the founder, Dr William Charles Winshaw, purchased land in Stellenbosch and started making natural or non-fortified wines. Over time Stellenbosch Farmers Winery through mergers and takeovers of other wholesalers and manufacturers such as: Monis of Paarl, VH Metterson, Nederburg and Sedgwick-Taylor resulted in Stellenbosch Farmers Winery becoming the producer and marketer of a large range of natural and fortified wines and spirits.

In 1945 Distillers Corporation was founded, quickly becoming the second biggest producing wholesaler after Stellenbosch Farmers Winery at that time. Similarly to Stellenbosch Farmers Winery the company also expanded through mergers and takeovers of companies like the Drostdy Co-operative Cellars and South African Distillers. In 1974 Distillers Corporation formed Bergkelder. Bergkelder was at the time

an original marketing concept which invited wine estates to make use of the corporation's bottling, sales and marketing expertise as well as maturation facilities.

In 1994 when the Apartheid era ended and the South African wine industry was open to exporting wine to the rest of the world, the quality of the production was quickly addressed. South Africa is now known as a county with boutique wineries and wine of excellent quality.

In 2000 Stellenbosch Farmers Winery and Distillers Corporation merged, both companies owned market leading brands and were listed on the Johannesburg stock exchange (JSE). Stellenbosch Farmers Winery produced and distributed wine, spirits and non-alcoholic drinks. Distillers Corporation also produced and distributed wine but the majority of the focus was on spirits, brandy in particular. The merged company is called Distell. In the spirits segment Distell is currently a market leader with brands that include: Klipdrift, Oude Meester, Viceroy, Van Ryn's brandy, Gordon's gin, Three Ships whisky, Scottish Leader whisky, Romanoff vodka and Amarula Cream liqueur. Distell also has the two market leading cider brands which are Hunters and Savanna. The company also owns a multitude of wine estates and brands including: Nederburg, JC Le Roux and Pongracz [21].

Douglas Green & Co was founded in Paarl selling reasonably priced quality wines, brandies, ports and sherries. In 1991 Douglas Green & Co merged with Bellingham cellars to form Douglas Green Bellingham (DGB). The DGB portfolio includes brands such as: Boschendal, Bellingham, Tall Horse, Strawberry Lips, Tang, Zappa Sambuca and The Redrock Brewing Company.

Pernot-Ricard is a French-based company and a world leader of the wine and spirits industry. Pernot-Ricard was established in South Africa in 1994, after the lift of the international trade sanctions. The company promotes both the group's international brands and manages local and regional brands, including Chivas Regal, Jameson whiskey, Absolut vodka, Heart rum, Olmeca tequila and G.H.Mumm champagne [27].

Really Great Brand Company (RGBC) was founded in 1996 and distributes premium spirits. RGBC is a small, independent and owner-managed company that has 25 premium brands in its portfolio including: Jack Daniels bourbon whisky, Dom Perignon champagne, Hennessey cognac and Moet & Chandon champagne [17].

Heineken was part of a joint venture in South Africa between: Namibian Breweries, Diageo (a large global beverage company) and Heineken. The joint venture was formed in July 2004 and called Brandhouse. The joint venture was dissolved in 2015, with the companies setting up independent manufacturing, marketing and distribution in South Africa. Brandhouse was selling and marketing brands including Heineken, Windhoek Lager, Guinness beer, J&B whisky, Dimple whisky, Johnnie Walker whisky, Smirnoff vodka, Smirnoff Spin and Smirnoff Double Black RTD [20]. After the joint venture was dissolved Diageo South African now sells and distributes all the brands previously held by the joint venture except for Heineken and Namibian Breweries brands which include: Heineken and Windhoek Lager [31].

The history of the South African alcoholic beverage industry and the success of previously SAB has consolidated the landscape of the beer segment in South Africa. The spirits segment has a vast offering of local and international brands to choose from. While the wine segment has an immense amount of local brands and some international brands. The South African alcoholic beverage landscape offers consumers a large variety of products to choose from, these products are competing with each other for market share

and brand loyalty.

The advertising landscape

Advertising is a means by which a brand can build relationships with consumers, for the main aim of advertising is to inform, persuade and remind consumers about the brand, in order to generate brand awareness. Marketing is not about finding the right consumer for the product, but finding the right product for the consumer. Lester Wunderman [39], a respected marketer said; "The chant of the Industrial Revolution was that of the manufacturer who said 'This is what I make, won't you please buy it.' The call of the Information Age is the consumer asking, 'This is what I want, won't you please make it.' ".

The advertising landscape is continually changing for example; the digital revolution started in the late 1980s and changed the marketing landscape forever. New capabilities were generated for both consumers and businesses. Consumers are able to use the internet and have increased buying power due to the ability to:

- compare prices and services;

- read information on product quality;

- read and generate reviews;

- not be limited by geography;

- hold a reverse auction where sellers compete for their business;

- have enormous variety that would not be possible in a physical store;

- shop at any time of the day or night;

- visit chat rooms and get the option of others; and

- shop from home or where ever they have internet access.

The alcoholic beverage industry is highly seasonal with most brands seeing an increase in volumes over the summer season. In Figure 1.2 the monthly off-consumption sales volumes for the South African alcoholic beverage industry are expressed as a percentage of the annual sales for that year, recorded by Nielsen [51] from 2015 to 2017. Figure 1.2 shows the seasonality of the alcoholic beverage industry, with peaks in sales over the summer season. Marketing campaigns all compete for consumers' attention during these periods.
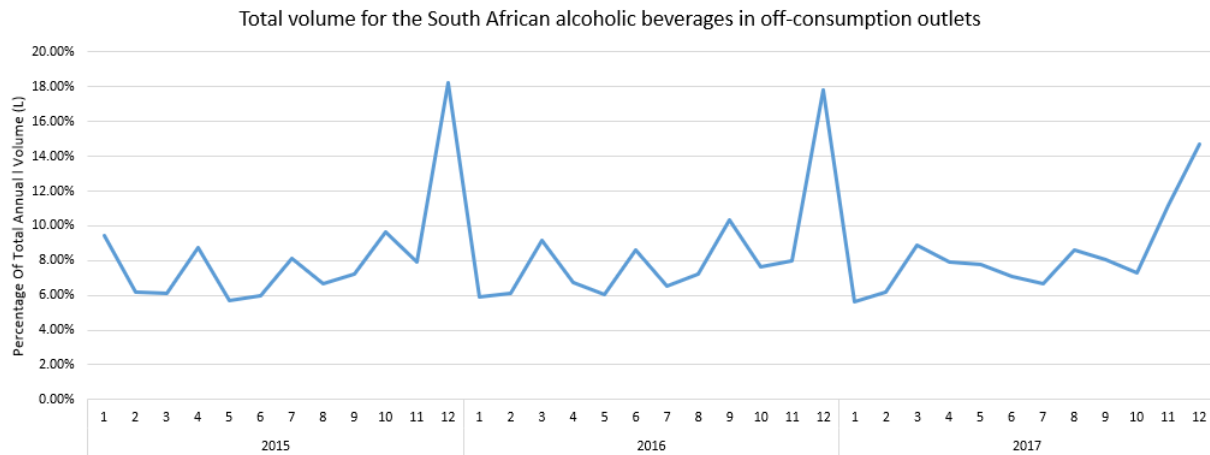
FIGURE 1.2: *Nielsen's volume report for the South African alcoholic beverages in off-consumption outlets [51]*

Advertising activity in retail outlets is constantly taking place and this makes it difficult to isolate the effect that an individual advertising event has on that brand's revenue, sales up-lift, or shift in consumer buying behaviour. There are many different ways that marketers can advertise their brand, some examples are:

- price discounting;

- value add (i.e. buy 2 and get 1 free, buy a 1L and get a free soft drink/mixer);

- instant gratification (i.e. spin and win an ice bucker, hat, t-shirt or glasses);

- main prize (i.e. win a trip to Thailand);

- gifting (i.e. gift box or tin);

- mass-media (i.e television (TV), radio, outdoor billboards);

- print (i.e. magazines, news paper);

- broadsheet (i.e. retailer print adds in news papers); and

- digital and social media (i.e. banner advertisements, Facebook).

From the launch of a brand to being culled, the brand progresses through a sequence of stages from introduction, growth, maturity and to decline. This process is known as the product life cycle (PLC). As a brand moves through these stages the marketing strategy and mix need to be updated.

Over a brand's life cycle the health of a brand or consumer sentiment can be measured through consumer surveys. The health of a brand can be affected by advertising and this change in consumer sentiment can be used as a measure of effectiveness of advertising. Brand-health measures are tracked over time to see if the consumer sentiment scores are improving or declining. The list below gives the 6 types of classifications of brand-health.

- Unprompted awareness: A respondent is asked to indicate all brands that you are aware of without any prompt by the interviewer.

- Aided awareness: A respondent is asked which brands you are aware of with the interviewer using a brand photo card to prompt answers.

- Ever drunk/consumed (Trial): A respondent is asked if you have ever tried the brand. Trial is needed to get into a consumer's regular usage repertoire.

- Regularly drunk/consumed: A respondent is asked which of the brands do you drink regularly. Regular usage is needed to building repeat purchases.

- Most often: A respondent is asked which one brand of alcohol would you say you drink most often. 'Most often' can be used as an indicator of brand loyalty.

- Power in the mind (PIM): This is a single share score that summarises the share a brand occupies in the minds of consumers. This score takes into account a consumer's desire or sentiment towards the brand and the measure links better than other methods with actual behaviour making it an indication of future behaviour.

These measures can be used as a proxy for consumer sentiment towards a brand. The advertising that a brand does will have an effect on these measures, advertising that resonates with the consumer should lead to the trend moving upward in the different measures. The brand-health scores are indicative of the health of a brand, indicating how consumers feel about a brand. These measures also show if the advertising is leading more consumers to consider the brand. The overall optimal spend on advertising is made up of a mixture of different mediums and activities, each aiming to affect consumer buying behaviour [39].

1.2   Problem description

Trying to reach consumers using all marketing methods is costly and will lead to overspending on advertising versus the revenue gain. The FMCG environment is competitive and expensive to compete in. At the same time it becomes necessary for companies to reduce costs and become more efficient to obtain the required shareholder's return. To achieve these goals companies need to find the most efficient ways to reach consumers. Factors that need to be taken into consideration when investigating and doing research within the alcoholic beverage industry are the clutter and noise in the market, data availability and accuracy.

- Clutter and noise in the market:
  - are due to the many activities taking place at once in the market. A brand can not be studied in isolation, as the effects of competitor advertising activities also have to be taken into account;
  - also include the influences of economic conditions, product life cycle, consumer sentiment, etc.; and
  - the alcoholic beverage industry is diversified with thousands of different brands and products.

- Data availability and accuracy:
  - the manufacturers of these different brands are reluctant to share information on how they spend their advertising budgets;

– the focus of the manufacturers has not been on collecting detailed information and the detail of many marketing campaigns is often not recorded at all; and

– in order to use methods like regression analysis, accurate historical time series data is needed on a detailed level that is often not available.

- The lack of accurate advertising spend information and the clutter within the market need to be taken into account when performing any investigation into this landscape.

The factors that needs to be considered when investigating and quantifying the effect that advertising has on sales or consumer buying behaviour in the alcoholic beverage industry are:

- determining a competitor set;

- defining insights; and

- integrating the results into strategy.

Factors to consider when determining a competitor set, defining relationships and insights are:

- what is the best method to group brands in order to use quantitative methods to evaluate their relationships?

- How are both the relationship between the products and the effort of advertising taken into account?

Factors to consider when integrating the results into industry or strategy are:

- how will the results be used?

- Can they be adopted into the strategy of a brand?

- How frequently can the results be updated, as the market landscape is ever changing.

These factors need to be taken into account when determining the competitors set and if a brand or product is spending advertising funds efficiently and effectively to generate sales uplift.

1.3   Objectives

The main objective of this research is to quantify and understand the relationship that various marketing activities have on the sales of a product or the change in market share of a product, thus representing a change in consumer buying behaviour. The objectives of this study is to gain insight into consumer buying behaviour in the South African alcoholic beverage industry and to analyse the relevant variables so that significant relationships can be investigated using quantitative methods to build a results driven efficient marketing strategy. The objective can be broken down into two parts:

1. Define a product's competitor set.

- Define a product's competitor set which is a set of products that interact with the product being investigated. Each brand or product within the alcoholic beverage industry is in a different life cycle, has a different consumer preference and used at different drinking occasions. The competitor set for a product should be evaluated against the relevant set of brands or products in the same life cycle stage. The product's life cycle or the growth potential of a product should be taken into account, as well as the market conditions.

- Investigate methods that address the problem of data availability, looking to use methods that do not require long periods of historical time series data such as:
  - factor analysis;
  - cluster analysis;
  - Boston Consulting Group growth share matrix; and
  - data envelopment analysis.

- Use regression analysis to establish a benchmark for the type of relationships and insights that can be observed from quantitative methods. Defining a competitor set by investigating significant relationships, taking market factors into account by calculating the price elasticities; and

- evaluate and combine the results of the methods mentioned above in order to obtain a final competitor set per product.

2. Determine if a brand or product's advertising spend is efficient and contributes towards changing consumers buying behaviour, thereby leading to an increase in revenue or market share.

- Determine the most efficient use of advertising that will result in the desired change in consumer behaviour, using data envelopment analysis.

- Quantify the relationship between the product being investigated and the competitor set, advertising variables and market factors.

- Evaluate and compare the results of both methods in order to obtain actionable insights and efficiencies; and

- make clear recommendations on how these methods can be used in the industry and integrated into strategy.

As part of the objective the most efficient use of advertising spend to change consumer buying behaviour will be investigated, as such efficiency needs to be defined and a clear distinction made between effectiveness and efficiency. Figure 1.3 illustrates the difference between effectiveness and efficiency, where effectiveness is defined as accomplishing a purpose and efficiency as accomplishing the purpose with the least amount of waste for the expected result.

In order to achieve these objectives and have results that can be effective in driving strategy, the time it takes to model the data and frequency with which the models can be updated and provide significant insights is important. The methodologies that will be investigated are methods that do not require historical time series data and can therefore be updated more often, to quantify the relationship that marketing activities have on brand revenue growth and responsively adapt strategies to shifts in the market. This implies the ability to rapidly quantifying the full mix of marketing activities and determining the optimal use to generate revenue growth for a brand, altering the consumers' buying behaviour.

	Inefficient	Efficient
Effective	Pursuing the correct goals, but inefficiently (costs are high)	Pursuing the correct goals and efficient (high ROI)
Ineffective	Pursuing the wrong goals and inefficient (low ROI)	Pursuing the wrong goals but cost efficient

FIGURE 1.3: *An illustration of efficiency vs. effectiveness [39]*

1.4 Methodology process

The methodology process followed in this study will be to investigate the relationships and insights that a product life cycle (PLC) methodology and data envelopment analysis (DEA) provide. These insights and relationships will be benchmarked against those that multiple regression analysis provides, which is currently the most commonly used technique. A causal research design will be followed to determine the relationship between consumer buying behaviour and internal and external market factors.

As a starting point for DEA a competitor set (a set of other products that have a relationship with the product being evaluated) needs to be determined. The PLC methodology will be used to group the brands or products and other products that they are competing with, into their respective life cycles, to determine competitor sets.

The PLC methodology is summarised as follows:

- due to the nature of the dataset the brands and products in the South African alcoholic beverage industry are by nature highly correlated. In order to overcome the high correlation within the dataset, it will be transformed using factor analysis to produce uncorrelated factor scores as an input for the cluster analysis;

- cluster analysis will be used to group the brands or products that have overlapping characteristics and similarities; and

- the Boston Consulting Group (BCG) growth share matrix will be used as a tool to determine each product's life cycle. This method seeks to group brands according to their growth potential benchmarked against market leaders.

The growth potential of a brand should correlate with its life cycle. In order to use the BCG growth share matrix, each product's biggest competitors need to be identified.

Regression analysis will also be used to determine a competitor set. The results will be compared with that of the PLC methodology. Once the competitor set has been defined, DEA can be used to draw insights about the efficiency of the marketing mix being used to drive sales or increase consumer buying behaviour. DEA can determine within the market which brands are spending their budget most efficiently and which particular type of advertising causes inefficiencies. The results from the DEA will be compared to the results when using multiple regression. The diagram in Figure 1.4 documents the process that will be followed in this research.

Problem:
The South African alcoholic beverage industry has thousands of competing brands. There are also many different advertising types and associated costs. Consumer buying behaviour is not only influenced by these factors but by economic factors as well. Companies also need to reduce costs and become more efficient to obtain the required shareholder's return.

Problem:
The alcoholic beverage industry has many different advertising types and campaigns running constantly. Clutter and noise in the market make it difficult for a brand to be studied in isolation, as the effects of competitor advertising activities also have to be taken into account. Marketing strategies are also influenced by economic conditions, a brand's life cycle and consumer sentiment.

Problem:
Data availability and accuracy is a problem due to the limited amount of available data on how marketers spend their advertising funds. The focus for most manufacturers in the alcoholic beverage industry has not been on recording event level detail about their advertising campaigns and therefore the detail of many marketing campaigns is often not recorded at all.

Main objective:
Gain insight into consumer buying behaviour in the South African alcoholic beverage industry. Analyse the dataset and variables so that significant relationships can be investigated using quantitative methods to build a results driven marketing strategy and efficiencies.

Objective 1:
Define a product's competitor set.

Objective 2:
Determine if a brand's advertising spend is efficient and contributes towards changing consumers' buying behaviour and leads to an increase in revenue or market share.

Problem:
The thousands of brands in the market cause clutter when performing analysis. Data is not easily available for all the products on the market.

Objective:
Define a product's competitor set which is a set of products that interact with the product being investigated. The product's life cycle or the growth potential of a product should be taken into account, as well as the market conditions.

Technique:
Product life cycle methodology:
- Factor analysis
- Cluster analysis
- BCG growth share matrix

Problem:
Validate the results of the PLC method using a benchmark method.

Objective:
Define a competitor set by investigating significant relationships, taking market factors into account.

Technique:
Price elasticity model using the log-log regression demand function.

Problem:
Review the results of the PLC methodology and the benchmark method.

Objective:
Evaluate and combine the results of both methods in order to obtain a final competitor set per product.

Technique:
Analysis of shared findings and commonalities between the PLC method and price elasticity model.

Problem:
Evaluate efficiency of marketing campaigns of a product within a competitor set.

Objective:
Determine the most efficient use of advertising that will result in the desired change in consumer behaviour.

Technique:
Data envelopment analysis

Problem:
Validate the results of the DEA method using a benchmark method.

Objective:
Quantify the relationship between the product being investigated and the competitor set, advertising variables and market factors.

Technique:
Determine the advertising mix that has a significant relationship with the product being investigated using multiple regression analysis.

Problem:
Review the results of the DEA methodology and the benchmark method.

Objective:
Evaluate and compare the results of both methods in order to obtain actionable insights and efficiencies

Technique:
Analysis of shared findings and differences in implementation of the results of DEA and regression analysis.

FIGURE 1.4: *Process flow for the research methodology investigating consumer buying behaviour*

1.5   Scope

The scope of this study will be the South African alcoholic beverage industry within a specialised retailer. The total market within this specialised retailer will be modelled to take into account the interplay of products across categories. For example, using volume and value reported by Nielsen the brandy category has vastly reduced in size over the last 10 years, while the total alcoholic beverage market has not reduced [51]. Thus brandy consumers have moved to other categories despite high advertising investment. In the last few years, however, due to price discounting the category has started to gain volume share, therefore switching consumers' buying behaviour from other categories back into brandy [51]. In order to gather real insights the full South African alcoholic beverage industry needs to be modelled within the specialised retailer.

Data availability and accuracy is a problem in the alcoholic beverage industry. An alcohol beverage producer, referred to as Company A, has given permission for their data to be used in this study. All the data that will be used in this study is secondary information.

In 2010 it became legal for supermarkets to sell spirits in their own liquor stores. Since then, retailers like Spar, Pick n Pay, Shoprite and Checkers have dramatically changed the retail environment. The specialist retailer outlets, liquor outlets linked to a food supermarket chain, now account for a large portion of the total markets' volume (13.8% share of off-consumption moving annual trend (MAT) as in Aug 2015 according to Nielsen [52]). This channel was previously dominated by independent outlets but specialist retailer stores are quickly growing in market share and expanding in number (13.44% growth on the previous year MAT, Aug 2015, according to Nielsen [52]). The specialist retailer's outlets are strategically located near the entrance of their food supermarkets, as they are segmented by location and income, thus providing customers of the same income level with an accessible wide range of alcoholic beverages. The full product range within the selected specialised retailer will be modelled, this is to assume the same amount of clutter and noise in the South African alcoholic beverage industry.

A spend tracker produced by Nielsen [52] provides industry spend per brand on mass-media, print and digital media. Company A does market research through an external research house [34] on brand equity measures and this data will also be available for analysis. From the internal records and surveys by Company A, the information on instant gratification and high value prizes will be provided. The information includes the Rand value spent on each advertising activity, the product and time when it was done.

The information that Company A provides will be used to investigate and quantify the effect that advertising has on consumer buying behaviour. This research will not investigate individual advertising events, rather the holistic effect that advertising has on consumer buying behaviour. Clear recommendations will be made on how these methods can be used in the industry and can be integrated into strategy.

1.6   Relevance of the study

This study will be of relevance to the alcoholic beverage industry, as it will direct alcohol producers on a process to quantify the relationship between their brand and products' advertising and their revenue. This information will be of importance in guiding manufacturers on how they should develop their brand's marketing strategies. The study aims to give advice on the maximum threshold of marketing spend per

advertising type, so as not to over invest in advertising but spend at the optimal level. In a very competitive environment the target of this study is to uncover relationships that will aid a manufacturer in gaining market share from competitor brands. Advertising accountability is becoming a requirement for success and to achieve this it is required that advertising efficiency be measured, maximised and benchmarked.

This study will investigate methods that require less detailed historical time series data than is needed for traditional regression analysis. Also, the methods being tested do not require long periods of historical data. This will be of relevance to the industry, as these models can be updated more frequently as changes in the market occur.

Many academic scholars like Bass [41], have questioned how much advertising spend is wasted and inefficient, not driving any consumer behaviour. This sentiment was mirrored in the quote by John Wanamaker, former U.S. Postmaster General saying, "I know half of my advertising is wasted, I just don't know which half" [41]. In the alcoholic beverage industry manufacturers like Company A are spending almost a billion rand on advertising and millions on consultants' advice and analysis. There is a great need in this competitive industry to understand what advertising is effective and efficient. Effective advertising resonates with the target audience, driving a change in consumer buying behaviour and sales volume. Efficient advertising also resonates with the target audience but also generates more revenue than spend on advertisings [39].

Company A had recently commissioned a marketing mix study with an industry leading consultancy. The executives that commissioned the study found that while the results were interesting, they did not lead to actionable output for the marketing teams. The consulting firm used multiple regression analysis to calculate price elasticities, determined significant relationships and return on investment of previous campaigns. The marketing teams failed to understand or translate the results into a better or more efficient strategy for future campaigns. The executives felt that this failure was because the results were very scientific and difficult for marketers to interpret, the study took over 9 months making the findings 'old' when presented.

The use of DEA in this study aims to address these issues, as DEA gives results that are easy to understand and provides clear actionable outputs. The method also does not rely on detailed historical time series data and can be set up and updated quickly and frequently. The marketing mix or advertising types are an essential part of any marketing strategy. Using these advertising types effectively is necessary to get a positive corresponding demand or buying behaviour from consumers at a profitable price [32]. The study will be significant to the alcoholic beverage industry as trade promotion management or marketing spend effectiveness becomes a sought after skill.

1.7 Layout of document

The objective of this research is to quantify and understand the relationship that various marketing activities have on the sales of a product or the change in market share of a product, this representing a change in consumer buying behaviour. Starting off, Chapter 2 examines methods that are currently being used in the alcoholic beverage industry to analyse consumer buying behaviour. Regression analysis is found to be the most common method and various types of regression analysis will be examined. The PLC methodology for determining a product's competitor set will be examined, including the grouping methods that will be used in the process. The DEA methodology will be used to determine the efficiency

of the advertising mix.

In Chapter 3 the different datasets will be reviewed. Each data source will be discussed separately, the dataset is either from survey collection or from transactional recordings of consumer purchases. The data and data sources will be described and relationships will be investigated in this chapter.

In Chapter 4 the two different methodologies with the objective to determine a product's competitor set will be modelled, reviewed and discussed. Firstly, the PLC methodology will be described and applied. This methodology includes using factor analysis, cluster analysis and the BCG growth share matrix. The results will be displayed and discussed. The benchmark or industry standard methodology of calculating the price elasticities using regression analysis, will also be used to determine a product's competitor set in Chapter 4. In Chapter 5 the results from the two methodologies will be compared and discussed.

Chapter 6 uses two methods to determine the efficiency of advertising efforts. Firstly, regression analysis will be used as a benchmark method for representing the industry standard and the results from the regression analysis will be displayed and discussed. Also in Chapter 6 the DEA methodology is applied and results will be displayed and discussed. Similarly to the previous chapters, in Chapter 7 the two methods will be compared and discussed. Finally in Chapter 8 the investigation will conclude with a summary of the findings. This chapter will be followed by the Appendix and Bibliography.

In summary the focus of this study will be to quantify the relationship between marketing activities and brand revenue. This investigation will determine what is the optimal spend on advertising that will generate the maximum sales volume uplift, eliminating overspend on advertising. The optimal spend on advertising is made up of a mixture of different mediums and activities, each aiming to affect consumer buying behaviour [39].

CHAPTER 2

Literature Review

Contents

The starting point for this research is to examine which methods are currently being used in the alcoholic beverage industry to analyse consumer buying behaviour. This investigation will start by examining the alcoholic beverage and advertising landscapes. The literature will be reviewed for the most commonly used method or benchmark method for these landscapes. The technique used to determine the competitor set methodology will also be examined.

2.1   The current industry landscape

Finding the optimal balance between the mix of the different advertising types and investment levels is not new to the alcoholic beverage industry, or any competitive commercial enterprise with a high investment in advertising. A simple search on the internet provides many firms providing services in 'advertising effectiveness measurement' such as: Nielsen [52], Boston Consulting Group [16], Ipsos [36],

IRi [19], Kantar Millward Brown [9], McKinsey & Company [42] and Skim [23], to name only a few.

Companies are realising that consumers are shifting and changing how they learn about and buy products. This rapidly changing landscape has produced both the need for companies to understand their consumers' behaviour and for other companies to provide this service. The relevance in any industry is to understand the effects that different variables have on consumer buying behaviour can be seen in this press statement [36]:

"The ability to collect and analyze digital data at extremely granular levels enables both marketers and their advertising partners to more successfully measure, predict and action the most effective and profitable means of optimizing each digital channel to achieve their business objectives. We are excited that Google has taken such a proactive approach in working with MMA and analytic companies within the marketplace in providing such a high level of objectivity and transparency."
Patrick Cummings, CEO of Marketing Management Analytics, 10 April 2017.

The methodology that companies like Nielsen [52], Ipsos [36] and IRi [19], use to provide 'advertising effectiveness measurement' is called market mix modelling. They use multivariate regression analysis on sales, advertising types and economic time series data to approximate and forecast the relationship of the different variables [52]. Advertising literature, academic texts and research show that multivariate regression analysis is the most common method used to evaluate the relationship between a product and the factors influencing purchasing behaviour.

In the book, "How brands grow: what marketers don't know"by Byron Sharp [70], he defines '10 laws' based on empirical generalisations. The author challenges conventional marketing 'thinking' with results from data rich investigations. He challenges the notion of marketing strategies aimed at customer loyalty and retention against those aimed at customer acquisition or trail. In the results on buying behaviour when studying the Coca-Cola brand he found that the average buyer purchases the Coca-Cola product 12 times a year, this average is very misleading with a skewed distribution due to some very heavy buyers. Looking at the data alternatively he found that the typical Coca-Cola buyer purchases the product just once or twice a year. This information can have a dramatic effect on a firm's marketing strategy of customer retention versus customer acquisition, also affecting the overall effectiveness of the marketing strategy and return on investment [70].

A recent study was performed by Terblanche-Smit, Du Preez and Van Der Spuy in 2015 in South Africa, measuring the impact of branded alcohol advertising and price [75]. They compared the impact price fluctuations had on a brand versus segment of consumption [75]. The investigation used multiple regression to investigate relationships between brand advertising spend, sales volume, market share, retail selling price and segment volume [75]. The results showed that the branded advertising has little to no effect on the consumer consumption behaviour. The retail selling price had a significant effect on market share [75]. The study did have limitations, in that data was not readily available and only the mass-media advertising expenditure was included. The study also focused on seeing a change in market share in the short term and did not look for a gradual increase in share which could be attributed to branded advertising. The research did conclude that it is imperative for brands in the market to make use of the marketing mix model to find an optimal combination of elements [75].

In 1991 Sethuraman performed an extensive meta-analysis of the ratio of price and advertising elasticity estimates [69]. Price elasticity is defined as the incremental change in sales quantity due to an incremental

change in the product's price [69]. Advertising elasticity is similarly defined as the incremental change in sales quantity due to an advertising campaign or incremental change in advertising spend. The research investigated if price discounts are more profitable than an increase in advertising exposure. A total of 16 studies were investigated, all these studies used regression analysis to calculate the price and advertising elasticity estimates [69]. The generalised findings when examining the literature are [69]:

- for both durable and non-durable goods, price elasticity is on average 20 times that of advertising elasticity;

- non-durable goods have approximately a 25 times higher price elasticity as compared to advertising elasticity, suggesting that many brand managers over invest in advertising on non-durable goods;

- durable goods have approximately a 5 times higher price elasticity than advertising elasticity, suggesting that brand managers can consider both advertising and price discounting as profitable strategies;

- for products in different PLC stages the ratio of price elasticity to advertising elasticity is higher for mature products, suggesting the benefit in adjusting marketing strategies according to the PLC;

- the elasticity ratio is different when changing the level of aggregation in the dataset; and

- the omission of a lagged sales variable produces systematically lower elasticity ratio values and therefore the lagged sales variables should be included in the model [69].

In the research of Sethuraman [69], brand management or marketing is broken down into two broad categories: firstly advertisements that can enhance an aspect of the brand or product and secondly price discounts that provide an instant incentive to purchase the brand. When analysing price discounts the model takes into account the opportunity cost from lost profit for selling to consumers who would have paid the higher price, as well as the cost of doing business with retailers who do not pass the full price discount to the consumers. The research finds that the higher a brand or product's price elasticity is the more the brand or product should be discounted. The profitability of this strategy is heavily influenced by a retailer passing on the full discount [69]. A noted limitation is that the study compares the short term effects of price elasticity and advertising elasticity. While price discounting is a short term strategy, advertising can be designed to have gains that are only realised over a long term [69].

In a study titled, "Mastering the mix: Do advertising, promotions, and sales force activities lead to differentiation?"published by Boulding and William in 1994 [6], regression analysis is used to compute price elasticity as a measure of differentiation from prior marketing activities. Their study investigated the type of content normally communicated by advertising, promotion and the sales force campaigns under three different pricing strategies. The first pricing strategy is to price the product above the industry, secondly to align the price of the product with industry prices and the third pricing strategy to undercut the industry prices and price the product far below the industry standard [6]. The investigation used data from both durable and non-durable consumer goods manufacturers, focusing on industries that frequently use advertising, promotions and sales force activities as part of their strategies. The finding is that unique communication and activities tailored to the pricing strategy does lead to an increase in brand differentiation [6].

Multiple regression analysis was used in the study recently completed by Mazhar Ali in 2016 [1], investigating the effectiveness of creative TV advertisements for high involvement products. High involvement products are products where an extensive thought process is involved in the purchase decision. The

consumer first gathers knowledge and considers a number of variables before finally making a purchase decision. The research objective was to measure the impact of creativity in advertisements on attitude toward advertisement, unaided recall and purchase intention [1]. Firstly, a binary logistic regression model was used and the results showed a small relationship between creativity in causing high advertisement recall and then a multiple regression analysis confirmed the positive relationship between creative advertisements and advertisement recall [1]. The final result showed that creativity in TV advertisements does increase the advertisement recall but this does not follow through to willingness to buy the product. The regression analysis showed that the creative and standard TV advertisements had the same effect on influencing purchasing behaviour [1]. The availability of data was noted as a limitation of the research, suggesting that more product categories and detailed information about the commercials and frequency could have resulted in findings that could be more generalised [1].

Another study also looking at the creative characteristics of advertisements, was published by Robinson, Wysocka and Hand in 2007 [59]. Their study investigates the impact of seven creative characteristics of banner advertisements on the effectiveness of advertising online [59]. The methodology followed used multiple regression analysis with binary dummy variables representing the seven creative characteristics. The dependent variable was the measure of click-through from the banner advertisements all the way to purchase. The researchers found their results to be in line with previous studies, concluding that the design elements of effective banner advertisements should be large in size, have informative copy about the product and the absence of a promotional incentive [59]. Two noted limitations of the study are firstly, [59] that the additional effects of brand building are not taken into account and secondly, that all banner advertisements in the sample came from one online gambling gaming portal.

In a recent study in 2017 by M Morgenstern, Z Li and JD Sargent titled, "The party effect: prediction of future alcohol use based on exposure to specific alcohol advertising content" [47], multiple regression analysis was again used to understand relationships. The study investigates whether young adolescents and young adults in the United States are more influenced by 'party', life-style and image-oriented alcohol advertising than by 'sports', 'masculinity', 'tranquillity' and 'product quality' [47]. The focus of the study was to test if life-style advertising was a predictor of future drinking behaviour, given that 42 percent of alcohol advertising in the United States is classified as life-style [47]. A representative sample of the market was tracked for 2 years using surveys. The results showed that with adolescents and young adults in the United States, higher exposure to alcohol advertisements using partying themes, does appear to have higher rates of alcohol use and binge drinking, independently of exposure to other alcohol advertising [47]. This study was one of the first to segment the research into both age and advertisement content categories. A limitation of the study is that while the results of the study are significant, they are not based on actual behaviour, they are based on claimed behaviour from a survey.

An observation that seems common across the literature is the limited amount of data available. This can in turn be denotative of the granular detailed level of data that is needed for regression analysis. In many industries data is not being collected or shared on a granular detailed level, making analysis incomplete or limited. The issues have been noted in the problem description in the previous chapter. Regression analysis will be discussed in detail in the next section.

2.2 Regression analysis to determine and quantify relationships

In its most simple form regression analysis is a statistical process for estimating the relationships among variables [81]. Regression analysis is a statistical modelling technique used to investigate relationships between a continuous dependent variable Y and one or more independent variables $X_1, X_2, X_3, \ldots, X_k$ [57]. The objective of regression analysis is to formulate a function that describes the relationship between the variables, such that it is possible to predict what value the dependent variable Y will take on for given values of X_i, the independent variables [57]. When using multiple regression, having more than one independent variable, regression analysis provides insight into how the average value of the dependent variable changes when one of the independent variables is varied, holding the other variables constant.

2.2.1 Regression analysis: simple linear regression

Using data points $(x_1, y_1), \ldots, (x_n, y_n)$, and assuming a linear relationship between X and Y, equation (2.1) models the linear relationship between the variables. In the equation ϵ_i is an error term representing the fact that not all the data points for (x_i, y_i) will fall perfectly on the regression line [81]. The error term ϵ_i is expected to average out over all the data in the set and the average is equal to zero. The relationship can be expressed as,

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i . \tag{2.1}$$

The values of the coefficients β_0 and β_1 are unknown and need to be estimated. Equation (2.1) can be re-written using the assumption that $\epsilon_i = 0$ [81],

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i , \tag{2.2}$$

where $\hat{y}_i$ is the estimated value of y_i using the estimates $\hat{\beta}_0$ and $\hat{\beta}_1$ for the values of β_0 and β_1 respectively. Estimating the values of β_0 and β_1 is a process of minimising the sum of squares of the errors or residuals, this is the difference between the actual data point (x_i, y_i) and the predicted values. The resulting values of $\hat{\beta}_0$ and $\hat{\beta}_1$ is given by equation (2.3), letting $\bar{y}$ be the average value of all y_i's and $\bar{x}$ be the average value of all x_i's [81]. The estimated values are,

$$\hat{\beta}_1 = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2} \ \text{ and } \ \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} . \tag{2.3}$$

Different components of variation are used to determine the adequacy of fit of the model. Three components of variation are the sum of squares of the total variation (SST), the sum of squares of the error terms (SSE) and the sum of squares due to the regression (SSR), where

$$SST = \sum(y_i - \bar{y})^2 , \tag{2.4}$$

$$SSE = \sum(y_i - \hat{y}_i)^2 , \text{ and} \tag{2.5}$$

$$SSR = \sum(\hat{y}_i - \bar{y})^2 . \tag{2.6}$$

A small value of SSE is an indication that the least squares regression line fits the dataset [57]. The relationship between the error variances can be seen in [57], where

$$SST = SSR + SSE .$$

Equation (2.7) gives the percentage of variation in Y explained by X [81] otherwise known as the coefficient of determination,

$$R^2 = \frac{SSR}{SST} , \tag{2.7}$$

where R^2 is a statistic with range ($0 \leq R^2 \leq 1$). R^2 is a goodness of fit measure and indicates the proportion of the total variation in the dependent variable Y, around its mean that is accounted for by the independent variable X in the estimated regression function [57].

The standard error of the estimates s_e, is a measure of accuracy of the prediction obtained from a regression model [57]. The standard error quantifies the amount of dispersion or variation in the actual data around the fitted regression line [57]. Letting n be the number of observations then,

$$s_e = \sqrt{\frac{SSE}{n-2}} \ . \tag{2.8}$$

The t-test in regression analysis tests if a significant relationship exists between the independent variable(s) and the dependent variable [81]. The hypothesis for β_1 is that,

$$H_0 : \beta_1 = 0 \text{ (No significant relationship exists between } X \text{ and } Y \text{)}$$
$$H_1 : \beta_1 \neq 0 \text{ (A significant relationship exists between } X \text{ and } Y \text{)} \ .$$

The t-statistic, which is the ratio of $\hat{\beta}_1$ to its standard error, should follow a t-distribution with $n-2$ degrees of freedom, and

$$t = \frac{\hat{\beta}_1}{S_{\hat{\beta}_1}} \ . \tag{2.9}$$

The corresponding p-value indicates the probability of obtaining an outcome that is more extreme than the observed test statistic value if $\beta_1 = 0$ [57]. A small p-value (typically ≤ 0.05) indicates strong evidence against the null hypothesis, therefore the null hypothesis is rejected. A large p-value (typically > 0.05) indicates weak evidence against the null hypothesis, therefore the null hypothesis is not rejected.

2.2.2 Regression analysis: multiple regression

Multiple regression analysis follows the same methodology as simple regression analysis but includes more than one independent variable. Let k be the number of independent variables that are needed to predict the dependent variable Y, given that there are n number of data points of the form $(y_i, x_{1i}, x_{2i}, \ldots, x_{ki})$, where x_{ji} is equal to the value of the jth independent variable for the ith data point [81], and letting y_i be the dependent variable for the ith data point, and $i = 1, \ldots, n$ and $j = 1, \ldots, k$. Equation (2.10) models the relationship between y and the k number of independent variables [81],

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \ldots + \beta_k x_{ki} + \epsilon_i \ . \tag{2.10}$$

Using the same methodology discussed previously for simple linear regression, the error terms, ϵ_i are assumed to average out over all the data in the set and the average is equal to zero. Also the estimated values of $\beta_0, \beta_1, \ldots, \beta_k$ are represented by $\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_k$ and again found by minimising the sum of squared error, this error is the difference between the actual data points and the predicted values. This gives the least squares regression equation,

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{1i} + \hat{\beta_2} x_{2i} + \ldots + \hat{\beta}_k x_{ki} \ . \tag{2.11}$$

This model will also need to be tested to determine if it adequately fits the dataset. The equations (2.4) to (2.6) are still valid and can be used. Equation (2.8) needs to be updated giving equation (2.12) where

the standard error of estimates s_e is amended to include k independent variables [81],

$$s_e = \sqrt{\frac{SSE}{n-k-1}} \ . \tag{2.12}$$

As with simple linear regression, the t-test in multiple regression is testing if a significant linear relationship exists between the independent variables and the dependent variable [81],

$$H_0 : \beta_i = 0 \text{ (No significant relationship exis its between } X_i \text{ and } Y)$$
$$H_1 : \beta_i \neq 0 \text{ (A significant relationship exis its between } X_i \text{ and } Y) \ .$$

The t-statistic which is the ratio of $\hat{\beta}_i$ to its standard error, should follow a t-distribution with $n - k - 1$ degrees of freedom [81]. As a general rule $\alpha = 0.05$ is commonly used as a level of significance in regression analysis [81].

As described in equation (2.7), the R^2 is a test statistic to obtain how well the model fits the dataset. In multiple regression, the R^2 will usually increase at least a little with every additional independent variable X_i added to the model. This is a problem because the value of R^2 can be systematically increased and inflated just by adding more independent variables to the model, regardless of fit [57]. Therefore, when using multiple regression analysis it is advised to use the adjusted R^2, R_a^2. The adjusted R^2 takes into account the number of observations n in the sample and the k number of independent variables in the model where,

$$R_a^2 = 1 - \left(\frac{SSE}{SST}\right)\left(\frac{n-1}{n-k-1}\right) \ . \tag{2.13}$$

R_a^2 decreases if insignificant variables are introduced into the model, where as R^2 can not decrease when additional independent variables are being added [57].

The F-statistic is a helpful statistic to determine if all the β_i are simultaneously equal to zero [57], or whether at least one of the β_i's is not equal to zero. The F-statistic with $k, n - k - 1$ degrees of freedom is defined by the equation [57],

$$F = \frac{SST/k}{SSE/(n-k-1)} \ . \tag{2.14}$$

It is important to always keep in mind that a simpler model is most often the best model. When building a multiple regression model, it is possible to include too many independent variables, this is called model 'overfit'. This can lead to insights being reflected in the results about the sample and they can not be generalised about the population [57]. As a general rule it is advised to include a maximum of one variable for every 10 observations [28]. When a model has 'overfit' present then the R^2 and R_a^2 will be inaccurate and converge towards one [28]. The R_a^2, while being the better measure for multiple regression, can not give accurate results when 'overfit' is present in the model.

Another problem with including too many independent variables is multicollinearity, when including variables that represent the same construct or factor [57]. Multicollinearity is when the independent variables in a regression model are highly correlated among themselves. When multicollinearity is present, the predicted value of β_i becomes unreliable and can often produce erroneous results [57]. Independent variables will need to be tested using correlation analysis for strong linear relationships, if a strong relationship exists then some of the independent variables need to be dropped from the model [81].

The variance inflating factor (VIF) measures the speed with which the variance and covariances increase. This indicator can be used to establish the presence of multicollinearity. The VIF score is calculated separately for each variable, relative to all other variables in the model. The VIF_Δ for variable j provides an index or ratio to measure multicollinearity [28],

$$VIF_\Delta = \frac{1}{(1 - r_j^2)} \tag{2.15}$$

where r_j^2 is the R_a^2 for the regression of X_j on the other covariates.

An assumption of regression analysis, simple or multiple, is that the error term should not be dependent on the independent variable(s) [81]. This assumption is called homoscedasticity, meaning that the sample population is consistent and does not have sub-populations. Heteroscedasticity is the absence of homoscedasticity. To test if heteroscedasticity is present, the error terms are plotted on the y-axis and the corresponding value of X_i on the x-axis, the graph should show no discernible pattern for the size of the errors to depend on the value of X_i [81]. If heteroscedasticity is present, a possible solution is using the log-log form, as this will often eliminate the heteroscedasticity [81]. In most statistical software packages the option to calculate White's heteroscedasticity-consistent variances and standard errors is an available function. This option should be used, the formulation is not covered in this research [28].

Another assumption of both simple and multiple regression is that the errors should be independent. This means that the value of one error term should tell you nothing about the subsequent error term value, or any other error term value [81]. Autocorrelation is observed when there is a pattern in the error terms and they are not independent. The Durbin-Watson test calculates the ratio of the sum of squared differences in successive residuals to the SSE [28], it can be used to test for autocorrelation. If ϵ_t is the residual associated with the observation at time t, then the Durbin-Watson d statistic is defined as,

$$d = \frac{\sum_{t=2}^{n}(\epsilon_t - \epsilon_{t-1})^2}{\sum_{t=1}^{n} \epsilon_t^2} \ . \tag{2.16}$$

In the Durbin-Watson test results, if there is no autocorrelation the output test statistic will be approximately equal to two [28]. If the Durbin-Watson output test statistic is close to zero then positive autocorrelation might be present and if the value is close to four, then negative autocorrelation might be present [28]. The most common type of autocorrelation is first-order autocorrelation, this is where the observed error is inclined to be influenced by the observed error that directly precedes it in the previous time period. When positive autocorrelation is present, it is an indication that a significant independent variable is missing from the model. Typically the parameter estimates are not affected by autocorrelation, however, the autocorrelation in the model is there because a variable is missing and including this variable in the model would change the parameter estimates.

Autocorrelation also causes the parameter estimates to vary further from the true value than the standard error indicates. Typically when autocorrelation is present the estimates of the standard error are smaller than the true values, this can give independent variables significant relationships when they are actually not significant [28]. Autocorrelation needs to be corrected, this can be done by introducing lagged variable(s) or using the generalised least squares method [28].

In many cases non-quantitative or qualitative factors influence the dependent variable, these independent factors can be represented by 'dummy variables'. Dummy variables are represented by a binary variable, for the presence of a characteristic of a variable a one and for the absence a zero is used. For example if the months of the year are used to represent seasonality, then let the number of months or categories, c,

be equal to 12. There always need to be $c - 1$ dummy variables included in the regression model [81]. The dummy variable that is left out the model is the reference dummy variable.

It is important to choose the correct or best fit regression equation. Plotting the relationship between the dependent and independent variables will assist in choosing the correct functional form to model. A data transformation might be needed for both the dependent and independent variables or either of the variables, depending on the relationship [81].

2.2.3   Regression analysis: price elasticity

An application of economic theory into statistics is the price elasticity of demand. In economics the relationship between supply and demand is defined. Economic theory of supply and demand finds that these forces form an equilibrium price for the supply and equilibrium quantity exchange of goods for the demand in a free market [45]. This relationship of supply and demand can be used to ascertain predictions in response to changes in the market place.

In the 1890's the concept of price elasticity of demand was formulated from utility theory [49]. Price elasticity is a measure of responsiveness in the quantity demanded to a change in price [45]. The price elasticity of demand of product A, e_A is the percentage change in the quantity demanded if the price of the product A changes by one percent, *ceteris paribus* for other market factors [45], or

$$e_A = \frac{\textit{Percentage change in quantity demanded of product A}}{\textit{Percentage change in price of product A}} \ . \tag{2.17}$$

It is important to note for interpretation of price elasticity that the elasticity is calculated using percentage change and not absolute change, this is because the two variables are not of the same unit of measure or scale [45]. Equation (2.17) is a ratio, this ratio is called the elasticity coefficient. Also important for the interpretation of results is the sign of the elasticity coefficient. A negative sign means that when the price increases the quantity demanded decreases.

Examination and approximation of price elasticity has many useful objectives. Once the price elasticity is understood on a product level, manufacturers and retailers can execute pricing strategies to take advantage of the relationships that price elasticity defines and quantifies. Pricing strategies can vary from choice of regular price, promotional price, product positioning, magnitude and duration of the promotions [82]. Price elasticity has been used in many studies to estimate the relative relationship of products and price [82], including studies in the alcoholic beverage industry [40].

Price elasticity can further be broken down into own elasticity and cross price elasticity. Own price elasticity of demand measures the relationship between a product's demand and its own price, all other market factors remaining constant. Cross price elasticity of demand measures the change in the demand for a quantity of a product relative to a change in price of another product, all other market factors remaining constant [45]. This is an important relationship, if there is no responsiveness or relationship between the two products the cross price elasticity will be zero [45].

Assuming there are two products in a market, product A and product B. These two products can either be perfect substitutes or complementary. Substitutes and complementary products are two subsets of a product's competitor set [39]. This is because both substitutes and complementary products are in a

consumer's consideration set when making a purchase. Substitutes are direct competitors, the consumer is choosing between purchasing either of the products [39]. Complementary products are products that are usually bought together, for example chips and dip. If the price of chips increases a consumer may forgo the dip and just purchase the chips, they are therefore complementary but still competing with each other. They are in the same free market and under the same market conditions, cost dynamics, continuous supply and in similar life cycle stage.

The formulation for own elasticity using the mid-point elasticity estimation method and the price elasticity defined in equation (2.17) is [82],

$$
\begin{aligned}
e_A ={}& \frac{\%\Delta\mathbf{Q_A}}{\%\Delta\mathbf{P_A}} \text{ , where} \\
\%\Delta\mathbf{Q_A} ={}& \frac{Q_{(A)_w} - Q_{(A)_{w-1}}}{Q_{(A)_{w-1}} - Q_{(A)_w}}/2 \quad \text{and} \\
\%\Delta\mathbf{P_A} ={}& \frac{P_{(A)_w} - P_{(A)_{w-1}}}{P_{(A)_{w-1}} - P_{(A)_w}}/2 \text{ ,}
\end{aligned}
\tag{2.18}
$$

where e_A is the own elasticity of product A, and $\%\Delta\mathbf{Q_A}$ the percentage change in quantity $\mathbf{Q}$ of product A and $\%\Delta\mathbf{P_A}$ the percentage change in price $\mathbf{P}$ of product A. Assuming that the market under study records information weekly, w refer the the current week and $w - 1$ the previous week, where $Q_{(A)_w}$ is the quantity demanded of product A in a week and $Q_{(A)_{w-1}}$ is the quantity demanded of product A in the preceding week. Also $P_{(A)_w}$ is the price of product A in a week and $P_{(A)_{w-1}}$ is the price of product A in the preceding week.

This definition of own price elasticity can also be rewritten as [82],

$$
e_A = (\Delta\mathbf{Q_A}/\Delta\mathbf{Q_P}) \times (\hat{P}_A/\hat{Q}_A) \text{ ,}
\tag{2.19}
$$

where $\hat{P}_A$ is the average price of product A and $\hat{Q}_A$ the average quantity demanded of product A. The cross price elasticity, $CE_{A,B}$, is the percentage change in the quantity demanded of product A, in response to a percentage change in price of product B [82], where $\%\Delta\mathbf{P_B}$ represents the percentage change in price of product. Therefore,

$$
\begin{aligned}
CE_{A,B} ={}& \frac{\%\Delta\mathbf{Q_A}}{\%\Delta\mathbf{P_B}} \text{ , where} \\
\%\Delta\mathbf{Q_A} ={}& \frac{Q_{(A)w} - Q_{(A)w-1}}{Q_{(A)w-1} - Q_{(A)w}}/2 \quad \text{and} \\
\%\Delta\mathbf{P_B} ={}& \frac{P_{(B)w} - P_{(B)w-1}}{P_{(B)w-1} - P_{(B)w}}/2 \text{ .}
\end{aligned}
\tag{2.20}
$$

The price elasticity can be calculated using regression analysis, specifically with the log-log demand function [81]. This transformation uses the natural logarithm of both the dependent and independent variables to calculate linear inputs into the model. Given the standard multiple regression line in equation (2.10), a log-log model demand function can be formulated for price elasticity of product A [82], where β_1 is the own elasticity of product A and β_2 is the cross price elasticity of change in product A's demand as a response to the change in product B's price [81].

$$
\ln\mathbf{Q_A} = \beta_0 + \beta_1 \ln\mathbf{P_A} + \beta_2 \ln\mathbf{P_B} + \epsilon \text{ .}
\tag{2.21}
$$

2.3   The technique used to determine a competitor set by clustering according to product life cycle

In the research to follow methods will be used to evaluate the effectiveness of different types of advertising. The starting point when using these methods is to determine the set of brands or products to be investigated. If all the products are used, then the dataset will be extremely large and calculations will be arduous and the results maybe inconclusive.

The definition of a competitor is a brand or product that meets the same consumer need [39]. In some industries it is a simple undertaking to determine a product's main competitor. An example would be the soft drink industry, Coca-Cola knows that PepsiCo is its biggest major competitor. In the alcoholic beverage industry it is not so simple, as there are a vast variety of similar brands and products.

This section will focus on a method to determine a product's competitor set. When a product is being evaluated, it needs to be against a defined set of other products that have a relationship with the product being evaluated, called the competitor set. If a competitor set is incomplete then techniques such as multivariate data analysis will yield inconclusive results.

2.3.1   Product life cycle

The alcoholic beverage industry landscape is changing all the time; new products enter, new packaging formats are used, products leave the market, consumer preference shifts and there are fluctuations in consumers' disposable income. Due to all these landscape changes, there is a flow of long-established or mature products gradually becoming less relevant to consumers and therefore a less popular purchase. At the same time, the demand for contemporary new product innovations increases rapidly after they are launched.

As the industry landscape and competitors change over time, the products' positioning and marketing strategy need to be updated. The PLC methodology relates to the life of a product as it moves through four distinct cycle stages with respect to the industry landscape, investment required and potential revenue. For a product to have a life cycle four facts need to be affirmed [39], they are:

- the life of a product is limited or finite;

- the product sales pass through distinct stages, each stage posing different challenges, possibilities and difficulties to the seller;

- during the PLC, profits will rise and fall depending on the stage; and

- during the PLC, depending on the stage, products need different advertising, investment, manufacturing, purchasing and human resource strategies.

The most common PLC curves are s-shaped. The s-shaped curve can be divided into four distinct cycle stages which are: introduction, growth, maturity and decline. The four stages, as defined by Jeffrey [37], are listed below.

- **Introduction**
  Sales growth is slow as the product has just been launched and has a low distribution. In this

stage there will be very little or no profit due to the high costs of introducing the product to the market. For marketers the first objective is to get acceptance of the product by consumers, this will calculate a demand for the product. Advertising investment will be moderate due to the high risk.

- **Growth**
  In the growth stage the product is fully distributed and accepted by the market. Sales growth is rapid and profits should dramatically improve. Advertising investment is substantial to embed the product in the market and ensure that it survives to maturity. The main objective for marketers in this stage is to gain brand or product loyalty. The growth stage is when the large loyal base is built for a product.

- **Maturity**
  The sales growth will start to slow down as the product reaches acceptance by most consumers. The product will still have a large sales base and yield substantial profits. Advertising investment will still be substantial in order to maintain the product at maturity for as long as possible. The sales force now needs to maintain the distribution footprint as the product's growth slows in the maturity phase, retailers will want to stock other faster moving products in their stores.

- **Decline**
  Sales start to decrease and the large base starts to dwindle. Advertising investment is minimised in order to maintain profitability.
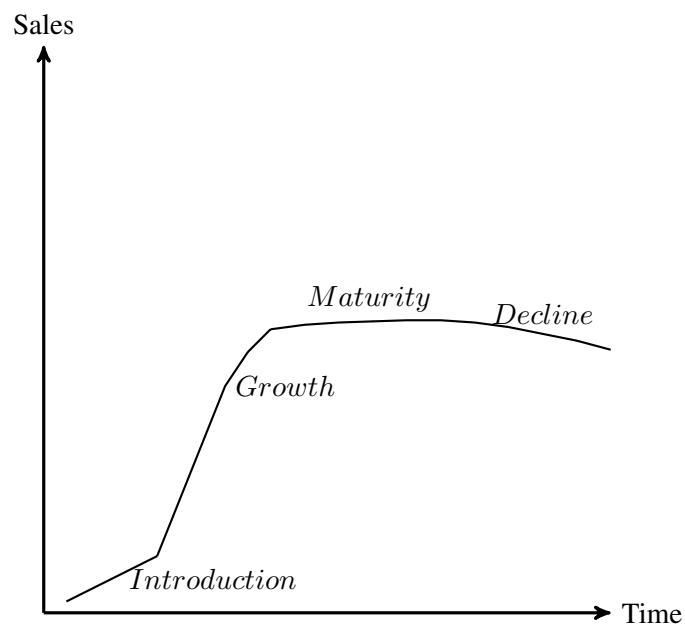
FIGURE 2.1: *Common product life cycle curve*

Figure 2.1 shows a common PLC. There are many different PLC 'curve shapes'. A product that enters the market and reaches maturity quickly, a fad, will have a very steep s-shaped curve. Other products like Coca-Cola have invested heavily and innovated in the maturity stage in order to go back into growth and extend the life of the product.

Brand owners understand the PLC and will typically invest according to the life cycle stage that the product is in. Brands in a specific life cycle stage should have similar investment levels and all will be competing with each other for market growth or share. The BCG growth share matrix is a common

technique used to group entities for example: firms, brands, or products into the four life cycle stages [39]. The BCG growth share matrix methodology requires each entities biggest competitor to be known.

The PLC methodology takes into account the consumer, customers in the market and competitors. The methodology of grouping according to product life cycle stage is best used to provide a framework to explain market changes or dynamics [18]. In new product development the PLC methodology is applied with the strategic intent to identify sales growth opportunities and maximise revenue [38]. The methodology has been used successfully in the FMCG industry as a prescriptive tool, providing insight into the different life cycle stages and their duration. When a company identifies which life cycle stage a product is in, the methodology can then be used as a guide for alternative strategies.

The PLC methodology can be used to predict future sales but it can also model marketing strategy. When the product life cycle stage is identified, then expectations about sales volumes and marketing strategy can be adjusted as per the description of the four life cycle stages [39]. Firms that monitor the landscape by using the PLC methodology are able to recognise changes and adapt to them. Using the PLC methodology can bring to light the effects that time and other changes have on a product and its potential growth [66].

Due to the simplicity of the product life cycle methodology it has been criticised [66]. The five main areas of criticism are:

- defining the appropriate level of aggregation for life cycle analysis;

- what are the underlying factors or forces that determine the parameters of the PLC;

- the utility of such analyses is frequently questioned by the subtle nature of the boundaries between the stages;

- the ability of a forecasting model to account for the driving forces during the various stages; and

- the role of the PLC in the formulation of strategy.

All of these criticisms can be overcome if proper research is performed to understand the underlying forces and processes within the industry. This environment can then be incorporated into the life cycle methodology [66]. The growth share matrix developed by the Boston Consulting Group is a methodology that can be used to segment or categorise entities into the four life cycle stages of the PLC [39], the Boston Consulting Group growth share matrix will be discussed in detail in the next section

2.3.2   Boston Consulting Group growth share matrix

The Boston Consulting Group (BCG) was established in 1963, in the United States of America, and is now one of the largest global consultancies. BCG is a management consulting firm and a world leader in business strategy. The BCG founder, Bruce Henderson, developed the growth share matrix in the 1970's as a portfolio management tool [16]. The original 'Boston Box' was developed as a tool to aid in strategic planning for multi-business-firms, the 'box' was used exclusively for taxonomy at first. As the method developed from the 'Boston Box' to the growth share matrix, it has been identified as a means to describe the typical product life-cycle of firms or brands [48].

In the article, "BCG Classics Revisited: The Growth Share Matrix" [58], the growth share matrix method is scrutinised to determine if after nearly 50 years, it is relevant to business today. The business landscape has changed over the past 50 years and one of the changes is that change happens more quickly. The growth share matrix is a 'snap-shot' in time and will need to be updated frequently in the current environment to keep up with the changes. The authors concluded that the method is still relevant and of use today, finding that; "Increasing changes certainly require companies to adjust how they apply the matrix. But it does not undercut the power of the original idea." [58].

There are numerous insights that can be gained from using the BCG growth share matrix. Firstly the graphical representation of the result into four taxonomies is visually easy to understand. The graphical display is a compelling and consolidated way to visualise the strengths and weaknesses of a firm's portfolio.

Secondly, the results give an indication into the capabilities for cash generation and general investment needs, this affecting the overall cash-flow of the firm. Thirdly, as each business unit is assigned to a quadrant, a unique strategy can be developed for that specific business unit [30].

The growth share matrix allocates each product according to its relative market share and growth rate into one of four quadrants they are: Question Marks, Stars, Cash Cows and Dogs. A successful product should travel from Question Marks to Stars and then to Cash Cows and finally becoming a Dog [39]. The four quadrants represent the life cycle of a product and are described below.

- **Question Marks**
  These are typically products that are in the introductory phase of the PLC. Products in this quadrant typically have high growth rates but low relative market share, typical of a new entrant. In this phase it is normal for a company to invest in advertising, to keep the product growing. The company can invest ahead of the product's profit to ensure that the product becomes a market leader.

- **Stars**
  These are typically products that are in the growth phase of the PLC. Brands that are Stars commonly have high growth rates and high relative market share. These products need to be fed in order to maintain their growth and they therefore usually have large advertising investment. It is not unusual for the product to be slightly cash-flow negative or neutral, this investment is made to push the product to reach the next stage.

- **Cash Cows**
  These are typically products that are in the maturity phase of the PLC. Cash Cows normally have slowing growth rates but large relative market share. These products ordinarily have substantial marketing investment to maintain them, however the advertising investment is not increasing year on year.

- **Dogs**
  These are typically products that are in the declining phase of the PLC. These products are declining in growth and relative market share. The advertising investment is normally very small and being tapered off. The company should not invest further in these products, as they do not have potential to go back to a previous stage under normal circumstances.

In Figure 2.2 the vertical axis of the matrix represents the annual market growth rate for the category

	Low Market Share	High Market Share
High Market Growth	Question Marks	Stars
Low Market Growth	Dogs	Cows

FIGURE 2.2: *BCG growth share matrix [39]*

within which the product is classified. For example, when calculating the annual market growth rate of of a leading brandy product, the brandy sub-category will be used as a proxy for the external profitability of the market. The brandy sub-category is a sub-set of brown spirits category and the spirits segment.

The annual market or category growth rate is calculated and not the product's growth rate [39], so that the category profitability is represented irrespective of the product's positioning within the market. The annual market growth rate can be calculated as,

$$Market\ Growth\ Rate = \frac{[Current\ Year\ Value] - [Previous\ Year\ Value]}{[Previous\ Year\ Value]} \times 100 \ . \qquad (2.22)$$

The horizontal axis of the matrix represents the relative market share for the product. Market share is an indicator of the product's strength within the market. Market share however does not indicate the relative size of the product within a particular category [30]. For example, if a leading brandy product has 10% market share within brandy, is this a substantial amount or not? The relative market share addresses this shortcoming. Relative market share is the product's market share relative to its leading competitor rather than shared to the category or segment. The relative market share can be calculated as,

$$Relative\ Market\ Share = \frac{[Current\ Year\ Sale\ Value]}{[Leading\ Competitor\ Current\ Year\ Sale\ Value]} \ . \qquad (2.23)$$

The relative market share measures the size or strength of a product relative to the closest competitor. A relative market share of 0.5 means that the product's sales value is half that of its closest competitor [30]. The 'normal' or not relative market share of a product is calculated by dividing the product's volume or value over that of the total market's volume or value for a respective time period.

The final calculation that needs to be done to complete the BCG growth share matrix, is determining the horizontal and vertical thresholds that define the four quadrants. This in essence means to define cut-off points that separate high growth from low growth products, similarly for relative share [30]. If the growth share matrix is for the alcoholic beverage industry, then the average growth rate for the industry can be used as the cut-off points for both the growth and share [30].

The BCG growth share matrix methodology requires each entity's largest competitor to be known. Cluster analysis as a methodology for grouping similar or competing entities will be discussed in detail in the next section.

2.3.3   Cluster analysis

Cluster analysis is used to classify entities into groups that are relatively homogeneous within the group and heterogeneous between the different groups, based on the values of their variables. Cluster analysis is a group of multivariate techniques that are used with the primary objective of grouping cases or sets. There any many different methods of cluster analysis but the main goal is always to position observations into groups with similar properties and creating groups that are as dissimilar as possible [44].

Cluster analysis uses the distance between observations to group the entities or cases. The method does not require anything to be known about the structure of the dataset, meaning that the number of clusters and cases belonging to those clusters are unknown.

The cluster analysis methodology has been used in many different fields of research, often the same method is 'discovered' in different fields and named differently but it is the same underlying technique. The methods have been used extensively in the fields of soil science, quantitative bio-geography, cladistics, inorganic molecular structures and also market segmentation [44].

There are three underlying fundamental concepts within cluster analysis, firstly the measurement of similarity, secondly how clusters are formed and finally how many clusters to form [7]. Similarity represents the distance or degree of relationship amongst the cases across all the characteristics defined by the variables. There are five commonly used distance measures that serve as criteria for grouping cases. The distance measures include the euclidean distance, squared euclidean distance, absolute euclidean distance, Chebyshev distance and Mahalanobis distance [44].

The second concept is how the clusters are formed, this is a process of simply identifying the two most similar cases and grouping them together. There are four main cluster forming methods they are: the hierarchical methods, partitioning algorithms, overlapping clustering procedures and ordination techniques [44]. The third concept is how many clusters should be formed, this is a process of using the hierarchical method, test statistics and the researcher's judgement [44].

The hierarchical method determines a treelike structure, called a dendrogram. There are two types of strategies for hierarchical clustering: agglomerative or divisive. The agglomerative technique starts with each case in its own cluster, at each successive step two clusters are merged until there is only one cluster [44]. They form non-overlapping cluster structures, meaning that once two cases are clustered together they will not be separated again. The total structure can be used as the solution or the researcher can select a level of interest to stop at.

The divisive method is the opposite of the agglomerative, it starts with all the cases in one cluster and at each step partitions them into more clusters, until they are each in their own cluster [44]. The hierarchical clustering methods are simple and a good starting point in developing an understanding of the structure and how many clusters there should be.

The partitioning or non-hierarchical methods require the number of clusters to be specified at the start. They do not follow a treelike structure approach, but allocate the cases to clusters based on their similarity. Partitioning methods produce distinct non-overlapping cluster structures. There are five distinct characteristics of partitioning methods.

Firstly, the starting partition or 'seed point' can either be randomly selected or specified. The second characteristic is the type of cluster assignment made to the cluster's centroid during analysis. Some methods will make a single evaluation of the cluster centroids, assigning each case to the nearest cluster centroid, while others will make multiple evaluations updating the centroids. The third characteristic is the statistical criterion used to assign the points to a cluster, this can range from a simple distance between two points to a complex matrix. The last two characteristics involve whether a fixed or variable number of clusters will be formed and how outliers will be treated. If a fixed number of cluster is formed, then outliers need to be forced into a cluster. With a variable number of clusters, outliers can be formed separately [44].

In cluster analysis it is useful to think of each data point or observation as the coordinates for a point in space. Distance and similarity are the two most important concepts in cluster analysis [7]. Some of the limitations of cluster analysis are also the reason there are so many different techniques. The different clustering methods have different treatment of the 'in-between cases', clusters with different densities, the dimensionality of the data, as well as the objective of the analysis. It can happen that one cluster has a high level of similarity between cases, while another cluster does not [7]. The method that is chosen should fit the dataset while keeping to the objective of clustering the data.

A second reason for the multitude of techniques is that the data itself may present various permutations that can be clustered. As an example, depending on the objective, a deck of cards can be clustered according to suit or number [7]. Another aspect of the various ways of clustering data is the starting point for the cluster centres. An important limitation of cluster analysis is that highly correlated datasets are a problem. Using other techniques like principal component analysis, canonical variates analysis, principal coordinates analysis and factor analysis can be used to affect a reduction in the apparent dimensionality of data [7]. When performing cluster analysis the definition of a cluster must be predefined, as the different clustering algorithms attempt to find different kinds of clusters [44]. The researcher needs to match the best algorithm with the dataset and objective.

Another limitation of cluster analysis is presented by the heuristic nature of the methods [44], this limitation is combined with the computational difficulties involved when datasets are large. There is an exponential increase in possible combinations as datasets increase in observations, this can make computation time very lengthy when looking for a global optimum. This means that each method will look for the optimal solution, using its own definition of a cluster and optimality but not all possible solutions will be tested [44].

There are several hundred cluster methods available [44]. SAS studio® software has many different methods such as [62]: average linkage, centroid method, complete linkage, density linkage, maximum likelihood for mixtures of spherical multivariate normal distributions with equal variances but possibly unequal mixing proportions, flexible-beta method, McQuitty's similarity analysis, median method, single linkage, two-stage density linkage, Ward's minimum-variance and disjoint cluster analysis (k-means model). This can make it problematic for the researcher to know which is the best method to use. Also, there is no assurance or guarantee that a clustering method will find an optional solution and the number of methods available adds to this complexity. The work done by Milligan and Cooper [39] is an aid to

mediating some of the limitations of cluster analysis, primarily in choosing a clustering technique. They used the monte carlo simulation analysis technique to review, group, categorise and compare the different clustering methods.

For hierarchical clustering they reviewed at least eleven different studies and for partitioning techniques five major studies were examined, the overlapping clustering procedures were also part of the research. The ordination techniques were also examined but are more correctly classified under factor analysis or principal component analysis. The Ward method and k-means method for hierarchical and partitioning techniques respectively were found to have an excellent recovery of cluster structures. Both methods yield the most consistent results over the different datasets and they were mostly the superior methods [44].

Scheibler and Schneider confirmed the results of Milligan and Cooper [65]. They concluded that when using euclidean distances both Ward's method and the k-means method using a non-random starting seed, ranked in the top performing methods. The study concluded that these two methods provided accurate solutions regardless of the type of similarity measure. They recommended these methods for application [65].

The cluster analysis method uses a measure of similarity or distance between observations to group the sets or entities [62], the scale of each variable describing an entity needs to be standardised so that each observation carries the same weight of importance.

For most standardisation methods the shape of the resulting clusters or groups needs to be known. For example, the standard linear transformation method works well when the resulting clusters are spherical, but only for certain instances of elliptical clusters [62]. The PROC ACECLUS procedure in SAS studio® software uses the approach provided by Art, Gnanadesikan, and Kettenring [62]. This approach does not require the number of clusters or the observation membership to a cluster to be known. The standard linear method uses the difference between the observations and the means, the PROC ACECLUS method uses a total sample sum of squares and crossproduct matrix for within clusters and between clusters to compute a pairwise difference between the observations.

Correlation analysis

In many quantitative methods it is important to quantify the strength of the relationship that variables have with each other as some methods such as cluster analysis are very sensitive to highly correlated entities. In the research done by Cohen in 1988 [46] on the statistical power of analysis for the behavioural sciences, standards were defined for correlation. A correlation coefficient that is less than 0.3 is an indication of a weak relationship, between 0.3 and 0.5 to be a moderate relationship and greater than 0.5 is an indication of a strong relationship [46].

The Pearson's correlation coefficient can be calculated between interval and ratio variables. There are two important numbers to observe when testing the relationship between variables: the correlation coefficient and the p-value. The correlation coefficients that have a value of 0.99 or higher indicate that two variables are almost perfectly correlated and at least one of the variables should be excluded from the cluster analysis model [46].

Using the t-test the hypothesis for testing whether a correlation is significant is

$$H_0 : \rho = 0 \text{ (There is no significant correlation between two variables)}$$

$$H_1 : \rho \neq 0 \text{ (There is a significant correlation between two variables)} .$$

The test statistic is,

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} , \tag{2.24}$$

where the p-value is determined by referring to a t-distribution with $n-2$ degrees of freedom [28], and r is the correlation coefficient which can range between -1 and 1. The p-value is an indicator of the significance of the results. The null hypothesis (H_0) states that no statistically significant correlation exists between two variables. A small p-value, typically less than 0.05, indicates strong evidence to reject the null hypothesis, or there is a statistically significant correlation between two variables. The null hypothesis (H_0) is not rejected if the p-value is greater than 0.05, indicating weak evidence against the null hypothesis [55]. The p-value is therefore used to determine whether the correlation coefficient is significant.

Measurement of similarity

Human beings perform cluster analysis all the time, it is a basic mental process of grouping similar items. These classifications are often done subjectively but with the development of algorithms, objective analysis can be performed [44]. The main definition of cluster analysis is that clusters should exhibit internal cohesion and external isolation from other clusters.

This means that similarity and dissimilarity need to be quantified. When plotting the data, observations with similar measurements will tend to be located spatially close together. Therefore, similarity or dissimilarity can be defined by distance-based measurements [7].

For a dissimilarity to be considered a distance, it needs to adhere to the following three conditions, where D_{ij} is the distance between observation i and j. Let x_i be the ith observation on a specific variable [7]. Then,

$$\begin{aligned} &D_{ij} \geq 0; \\ &D_{ij} = 0 \text{ if } x_i = x_j ; \\ &D_{ij} = D_{ji} ; \text{ and} \\ &D_{ia} + D_{ja} \geq D_{ij} . \end{aligned} \tag{2.25}$$

The first two conditions ensure that the observation has a measurable distance from other observations. The input data must be a coordinate or numeric variable, where the distance is symmetrical, it is the same distance from i to j and from j to i. The third conditions specifies that the distance is metric and avoids asymmetric dissimilarities, it is called the metric inequality [7].

Cluster analysis seeks to form groups based on similarities and differences within the dataset. The dataset can be formulated into a matrix with X represented by [79],

$$
X = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1p} \\ x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{np} \end{bmatrix},
$$

where n is the number of observations and p the number of variables describing the characteristics of the observations. The dataset is decomposed into a set of clusters based on the similarity and distance, such that S^l represents the set of n observations in l clusters [79],

$$
S^l = \{C_1, C_2, C_3, \dots C_l\}, \quad C_i \neq \oslash, \quad i = 1, \dots, l. \tag{2.26}
$$

Equation (2.26) mandates that all clusters or sets must contain at least one observations, no empty sets and that the clusters that are formed cover all the spaces that have observations. If that set of observations $o = \{A_1, A_2, A_3, \dots, A_n\}$, and any dissimilarity coefficient of observations D, then a set or cluster is classified a sub-set of p sets or clusters of observations o to which the method applies [79],

$$
\max_{i,j} D(A_i; A_j) < \min_{k,l} D(A_k; A_l),
$$

where $A_i, A_j, A_l \in 0$ and $A_k \notin p$. This constraint means that the maximum distance of any observations belonging to a set or cluster must always be less than the minimum distance of that observations from that cluster to another outside cluster [79].

Hierarchical clustering: Ward's minimum variance method

Ward's minimum variance method assumes that the dataset has a multivariate normal mixture when being clustered, as well as spherical covariance matrices and equal sampling probabilities. Ward's method runs through the dataset and at each iteration, the within-cluster sum of squares is minimised over all the partitions obtainable by merging two clusters from the previous iteration. The proportion of variance or semi-partial correlations make the sum of squares easier to interpret. The proportion of variance is equal to the sum of squares divided by the total sum of squares of two variables after controlling the effects of other variables [64].

Firstly defining the distance between clusters, let the distance between clusters C_K and C_L be represented as D_{KL} and letting cluster K contain n cases,

$$
D_{KL} = \frac{\| \tilde{x}_K - \tilde{x}_L \|^2}{\frac{1}{N_K} + \frac{1}{N_L}}, \tag{2.27}
$$

where D_{KL} is equal to any distance or dissimilarity measure between clusters C_K and C_L. Let the mean vector of cluster C_K be denoted as $\tilde{x}_K$, as is the mean vector of C_L denoted by $\tilde{x}_L$. Let N_K be equal to the number of cases in C_K and let N_L be equal to the number of cases in C_L. Also let n equal the number of cases and C_K be equal to the Kth cluster, of subset $\{1, 2, \dots, n\}$.

Let $d(x, y) = \frac{1}{2} \| x - y \|^2$ be equal to any distance or dissimilarity measure between observations and D_{KL} [62].

In equation (2.28) it is assumed that clusters C_K and C_L are merged to form C_M, D_{JM} gives the distance between the new cluster C_M and any other cluster C_J [62], where

$$D_{JM} = \frac{(N_J + N_K)D_{JK} + (N_J + N_L)D_{JL} - N_J D_{KL}}{N_J + N_M} \ . \tag{2.28}$$

A limitation of Ward's method is the tendency to produce clusters with the same number of cases, as well as joining clusters that have a small number of cases [62]. Another limitation is the method's sensitivity to outliers.

Non-Hierarchical clustering: k-Means method

In SAS studio® software the k-means method is evoked through the PROC FASTCLUS procedure. This simple unsupervised learning algorithm combines an effective method for determining initial clusters with a standard iterative algorithm, for minimising the sum of squared distances from the cluster means [10]. The PROC FASTCLUS procedure computes the Euclidean distances between quantitative variables and cluster centres.

The k-means algorithm begins with k initial cluster centroids $(c_1^{(0)}, c_2^{(0)}, c_3^{(0)}, \ldots, c_k^{(0)})$ or seeds, one centre for each cluster to form. The cases are assigned to cluster centres, such that each case belongs to only one cluster [5]. All the n cases or objects are compared with each seed for similarity by using the Euclidean distance between the case and the seed. Each case is assigned to the closest cluster seed, the objective function for the tth iteration is to minimise,

$$\sum_{i=1}^{n} \sum_{j=1}^{k} (\| x_i - c_j^{(t)} \|) \ , \tag{2.29}$$

where the number of cluster centres needs to be pre-specified prior to running the algorithm. The algorithmic steps for the k-means method begin with selecting the k cluster centres, then calculating the distance between each case and all the cluster centres and assigning each case to the nearest cluster centre [5]. Once the cluster centres are updated, the algorithm keeps repeating the calculation, updating the Euclidean distance for $j = 1, \ldots, k$ until each case is assigned to a cluster such that the sum of squared distances from the cluster means is minimised [5].

Using the k-means method the PROC FASTCLUS procedure in SAS studio® software generates non-random seeds or centre points for clusters, all observations are then divided into clusters based on their distance from the seeds. Each observation will belong to one and only one cluster. This procedure uses the Euclidean distance, so that the seed points are based on the least squares estimation. The procedure will run updating the seed points until it reaches complete convergence [54]. Each iteration reduces the least squares criterion until convergence is complete.

When generating random cluster seed points the k-means method is sensitive to outliers. These outliers will be visible in the results, as they will typically form clusters with only one case [10]. When determining the final value of k, these outliers need to be taken into account. Choosing a value of k that is too low will force the outliers into clusters with other cases and reduce the overall similarity within cluster groups. The k-means method is only recommended for datasets that have at least 100 or more observations [10].

One of the most challenging parts of cluster analysis using non-hierarchical methods, is knowing how many clusters should be formed [63]. When using the k-means method the number of clusters must be set at the start of the procedure, the procedure can be run multiple times with different values of k to determine the best value. This in itself can pose a problem of where to start with the testing. A hierarchical cluster method for example, Ward's minimum variance method, can be used to generate a starting point. The dendrogram produced from Ward's minimum variance method can give a good indication of the cluster structures, outliers and the distance between clusters.

Once a starting point has been established for k, significance testing statistics can be used as a guide for the final value of k. The cubic clustering criterion (CCC) can be used to approximate the number of clusters, based on minimising the within-cluster sum of squares [63]. The pseudo-F statistic can be a useful indicator of the number of clusters, as this statistic is an analysis of variance. The pseudo-F statistic gives an estimation of how well the clusters have formed by calculating the ratio of the variance between clusters and the variance within clusters [62]. Also, the pseudo-t^2 statistic gives an estimation of how far apart the clusters are after they are merged during each iteration of the algorithm [62].

The performance of the CCC test statistic has been evaluated by Milligan and Cooper [63] using the monte carlo method of evaluation. The results showed that the CCC ranked sixth, in terms of the best method of determining the number clusters. CCC tends to overestimate the optimal number of clusters [63]. When performing cluster analysis the CCC method should be used in conjunction with the pseudo-F and pseudo-t^2 statistics, to correctly identify the number of clusters.

In equations (2.30), (2.31) and (2.32) the formulation is given for the cubic clustering criterion (CCC), pseudo-F statistic and pseudo-t^2 statistic respectively, where

$$CCC = \ln\left(\frac{1 - E(R^2)}{1 - R^2}\right) \times \frac{\frac{\sqrt{nw}}{2}}{(0.001 + E(R^2))^{1.2}} , \qquad (2.30)$$

$$\text{pseudo-F} = \frac{\left(\frac{R^2}{G-1}\right)}{\left(\frac{1-R^2}{n-G}\right)} , \qquad (2.31)$$

$$\text{pseudo-}t^2 = \frac{B_{kl}}{\left(\frac{W_k + W_l}{n_k + n_l - 2}\right)} , \qquad (2.32)$$

$$\text{where } W_k \text{ is } \sum_{i \in C_k} \|x^i - \bar{x^k}\|^2 ,$$

$$\text{and } B_{kl} = W_m - W_k \text{ if } C_m = C_k \cup C_l ,$$

where R^2 is the coefficient of determination, this is a measure of determining the variation when clusters are formed. G is the number of clusters at any given level of the hierarchy and w is an estimation of the dimensionality of the variation between clusters. For the final cluster formed, C_k, the distance measure between observation x^i and the cluster centre x^k is represented by W_k, similarly for W_l. In equation (2.30), $E(R^2)$ is the expected value of R^2, n the total number of cases and w is an estimation of the dimensionality of the variation between clusters [63].

Using the goodness-of-fit criterion such as R^2, CCC, pseudo-F and pseudo-t^2 statistics will help overcome the challenge of determining k when performing cluster analysis. The output from the PROC FASTCLUS procedure will provide a graphical representation of the CCC, pseudo-F and pseudo-t^2 statistics.

For the CCC, a value per cluster greater than 3 is an indication of good cluster formations, while a value between 0 and 2 is an indication that there are potential clusters but not optimal. If the CCC value is a large negative number, then this indicates that there are significant outliers in the dataset [63]. When evaluating the pseudo-F statistics, the largest value should be used as this indicates well separated clusters [63].

Once the cluster analysis procedure is complete the results need to be evaluated for validity. Cluster validation is the procedure used to evaluate the goodness of the clustering algorithm's results [76]. There are three criteria for cluster validation [8], [13], [76]:

1. **internal cluster validation** which evaluates compactness or cluster cohesion, measuring how close are the objects within the same cluster;

2. **external cluster validation** validates the results against known labels or classes for the dataset by comparing the results of cluster analysis to the known external labels; and

3. **relative cluster validation** evaluates the cluster structure by varying the value of k, the number of clusters for the same algorithm, this method is often used to determine the number of clusters to form.

As highly correlated datasets are a problem when using cluster analysis [7], techniques like factor analysis can be used to reduce this affect, factor analysis will be discussed in detail in the next section.

2.3.4   Factor analysis

Factor analysis is a powerful technique used to refine and evaluate datasets. In the past the method has been used extensively in the field of psychology to find relationships between variables. In recent years the method has been used more broadly and has been applied to socio-economics, psychology, commerce, information systems, survey analysis and market segmentation [62] and many other fields with data that has an underlying construct [74]. Factor analysis gives a view of how the characteristics or attributes that describe the variables are grouped together, summarising large datasets according to the underlying relationships or constructs in the data [80]. It is an effective statistical technique that searches the data for joint variations in response to an unseen variable or variables [77].

Factor analysis uses the covariance of the observations for analysis and therefore uses a standard measure within the method. Therefore no standardisation needs to be performed on the dataset for factor analysis [72].

The factor analysis model expresses the variation and covariation within a set of observed variables, as a function of factors that capture the common variation [72]. Factor analysis uncovers latent variables that represent unobservable constructs or structures and these are referred to as factors. There are two main categories of factor analysis namely, exploratory factor analysis (EFA) and confirmatory factor analysis

(CFA). CFA is used as the name suggests to confirm a structure or construct, the researcher needs to already know how many factors are present. When the structure is unknown and the number of factors present is also unknown, then EFA will be used [74]. Using EFA imposes no preconceived or prior modelling knowledge on the result, therefore the underlying structure of factors can be identified [72].

The EFA method needs six basic assumptions about the dataset to be true or the method will not model the data correctly. The six assumptions are:

- firstly, that the measurement of the observations is at interval or ratio level, in that the difference between the observations is meaningful;

- secondly, that the dataset is from a random sample and properly represents the population [72];

- the third, fourth and fifth assumptions define the relationship between the variables, in that they need to have a linear relationship with a normal distribution and a bivariate normal distribution between each pair of observed variables; and

- the final assumption is that of multivariate normality [72].

Factor analysis has some limitations that need to be considered before using the method. The basic methodology of factor analysis relies on the correlations between the variables, this gives a description of the relationships, but the researcher still needs to overlay the causal insights. There are several methods for factor extraction such as maximum likelihood, unweighted least squares method, principal axis factoring and alpha factoring, to name a few [54].

The researcher will use a comparison of the output statistics to decide on the best extraction method to use, noting that information on the strengths and weaknesses of each method is not widely available [54]. In the work done by Fabrigar, Wegener, MacCallum and Strahan [24] in 1999 some guidelines are now available to determine the best extraction method. If the assumption of multivariate normality is being 'violated' then they recommend using principal axis factoring [24]. Other researchers have found that unweighted least squares is the best method to use when the assumption of multivariate normality is being 'violated' [54].

Another limitation is the general disagreement about the sample size needed. The recommendation by Suhr [72] is a minimum of 200, other subject matter authorities draw the line at a minimum of 500 [72], while others suggest a range from 100 to 1000 [74]. Evidence has shown that commonalities and the factor loading scores will determine the adequacy of the sample size [74]. The sample selection is assumed to be a random sample that is representative of the population and populations should not be pooled. There is also always a possibility that the observations are sample specific and not representative of the general population. Another limitation is that the dataset might not be normally distributed and not abide to the six basic assumptions discussed above [74].

Exploratory factor analysis algorithm

EFA is used to find the latent variables in the data when nothing is known about the constructs within the dataset, or how many dimensions are in the set of variables. CFA is used if the construct is known and

needs to be confirmed [15].

The extent to which observations vary together is call the covariance. A variance-covariance matrix shows the association between variables. The goal of EFA is to find groups of items that, when considered together, explain as much of the covariance as possible. Each of these groups of items is called a factor [15]. Each factor represents a certain amount of the overall variance observed in the variables. Together the factors should represent most of the variation in the variables.

Factor analysis starts off with creating the same number of factors as there are variables. The eigenvalue is a measure of how much of the variance of the observed variables each factor explains. The scree plot is a graphical representation of the eigenvalues, displaying the eigenvalue of each factor in descending order [54]. The factors that have an eigenvalue greater than 1, represent factors that explain more variance than a single variable does. The factor loading score of each variable represents the strength of the relationship of each variable to the factor. A factor's eigenvalue is the sum of the squared factor loading scores for variables on that factor [15], see equation (2.35). These factors can then be used instead of the variables to capture the relationship between the variables. The end result being that each aspect of a construct is represented by a single score or factor instead of an individual variable. The fewest number of factors that can explain the maximum amount of covariation should be chosen.

Factor extraction

The EFA process has three steps namely extraction, rotation and interpretation. Starting with extraction, a common factor is an unobservable variable that contributes to the variance of at minimum two of the observed variables [62], while a unique factor is also an unobservable variable but only contributes to the variance of one observed variable.

A common factor model [62] is represented by equation (2.33),

$$y_{ij} = x_{i1}b_{1j} + x_{i2}b_{2j} + \ldots + x_{iq}b_{qj} + e_{ij}, \tag{2.33}$$

where i, $(i = 1, \ldots, n)$, is the number of observations and j, $(j = 1, \ldots, p)$, the number of variables. Also y_{ij} is the value of the ith observation on the jth variable, and x_{ik} is the factor score value of the ith observation on the kth common factor $(k = 1, \ldots, q)$, and b_{kj} be the regression coefficient or factor loadings of the kth common factor for predicting the jth variable, also known as the factor pattern. The unique factor, e_{ij}, is the value of the ith observation on the jth unique factor and can also include an error term. The number of common factors is denoted by q.

For the common factor model to be valid two critical assumptions have been made. The first being that the unique factors are uncorrelated with each other and secondly, that they are also uncorrelated with a common factor. It is also assumed that common factors are uncorrelated with each other and have a measurable variance. Equation (2.34) is implied by the common factor model [62], to find the covariance s_{jk}, between the jth and kth variables, $j \neq k$, where

$$s_{jk} = b_{1j}b_{1k} + b_{2j}b_{2k} + \ldots + b_{qj}b_{qk} \,. \tag{2.34}$$

If the original variables are standardised in equation (2.34) the s_{jk} then yields correlations, so that the common factors explain correlations among observed variables. The variance of the jth variable is expressed as [62],

$$s_{jj} = 1 = b_{1j}^2 + b_{2j}^2 + \ldots + b_{qj}^2 + [\mathbf{U}^2]_{jj} \,, \tag{2.35}$$

where $b_{1j}^2 + b_{2j}^2 + \ldots + b_{qj}^2$ represents the communality of the jth variable and $[\mathbf{U}^2]_{jj}$ the uniqueness of the jth variable. The communality represents the proportion of the variance accounted for by the common factor. The residual correlation is that correlation that is not taken into account by the common factor. The residual correlation should be used as a measure to assess the model's fit, high values would mean an ill fitting model.

The common factor model in equation (2.33) can be simplified into a matrix algebra equation [62] written as,

$$Y = XB' + E , \tag{2.36}$$

Y is the $(n \times p)$ variable matrix, X is the $(n \times q)$ matrix of factor scores and B is the factor pattern and E the unique factor, B' is a $(q \times p)$ matrix and B is a $(p \times q)$ matrix.

To obtain the component score for each factor as linear combination of the observed variables the following formulas apply [73]:

- **Step 1:** Let y_{ij} be the observed values, $i = 1, \ldots, n$ and $j = 1, \ldots, p$. Standardise Y to obtain Z, an $(n \times p)$ matrix.

- **Step 2:** Determine S, a $(p \times p)$ matrix, which is the variance covariance matrix of Z.

- **Step 3:** Calculate V, the $(p \times q)$ eigenvector matrix depending on the chosen number of factors and

$$L = V'SV , \tag{2.37}$$

a $(q \times q)$ eigenvalue matrix.

- **Step 4:** The factor loadings B are calculated using,

$$B = V\sqrt{L} , \tag{2.38}$$

B is a $(p \times q)$ matrix.

- **Step 5:** The factor score coefficient matrix is,

$$A = S^{-1}B , \tag{2.39}$$

A is a $(p \times q)$ matrix; and

- **Step 6:** finally, the factor score matrix is calculated as,

$$X = ZA , \tag{2.40}$$

resulting in an $(n \times q)$ matrix.

It can be noted that $Z = XB'$ without the unique factor from equation (2.36).

The resulting factor scores X are the component scores, which will be used to represent each factor in further analysis. Typically, factor scores are computed and reported in a standardised form [62], where x is a factor score of the matrix $(n \times q)$, and B is the factor loading patter $(p \times q)$.

Factor rotation

Factor rotation is used to clarify relationships between variables and factors. The rotation process also provides options such as orthogonal or oblique rotation, each of these rotations also have different techniques that can be applied and should be evaluated to determine the best fit.

Factor rotation is a process where by the reference axis is manipulated, it is a non-singular linear transformation [72]. Factor rotation is used to clarify relationships between variables and factors, it does not change the relationships. The rotated pattern matrix consists of coefficients that are close to -1, 0 or 1, this makes interpretation easier. Factor loadings closer to 1 indicate a strong positive relationship, while close to -1 indicate a strong inverse relationship. The loadings close to 0 should also be examined to confirm that these variables are unrelated to the construct that is being formed. There are mainly two categories of rotation, they are: orthogonal and oblique [83]. Orthogonal rotation is a process where the axes are retained at a $90°$ angle to each other. In Oblique rotation different angles are used besides $90°$ to rotate the axes [72].

Orthogonal rotation assumes that the factors are uncorrelated with each other, this is not always the reality as often factors are correlated to a small degree [83]. There are many different types of orthogonal rotation, the two most commonly used rotations are Quartimax and Varimax. Quartimax rotation seeks to minimise the number of factors needed to explain each variable. Varimax rotation is used to minimise the number of variables with high loadings on multiple factors, making high loadings higher and small loadings smaller [83]. The factor loading score gives the magnitude of the loading, but not the statistical significance. The researcher can set out minimum cut-off thresholds for loading scores.

Oblique rotation is used when factors are assumed to be theoretically correlated [15]. Oblique rotation uses either a system of primary axes or a system of reference axes. The output is a pattern matrix that contains factor loadings as well as a correlation matrix, including correlation scores between factors [83]. There are many different types of oblique rotations, the two most commonly used methods are Oblimin and Promax. The Oblimin rotation seeks to simplify the output structure, while Promax is favoured due to computation speed with large datasets [83]. The output pattern matrix does not provide all the necessary information when extracting factors because the factors are correlated and it is not possible to measure the importance of a factor in explaining a variable. The factor structure and reference structures will assist in defining this relationship.

Sampling adequacy

The Kaiser-Meyer-Olkin measure of sampling adequacy is a measure of how well the data fits the factor analysis model [62]. The statistic is a measure of the proportion of variance among variables that might be common variance [2]. Equation (2.41) gives the equation for calculating the Kaiser-Meyer-Olkin score for variable j. The Kaiser-Meyer-Olkin score for variable j (KMO_j) is equal to the sum of the squared correlations r_{ij}, divided by the sum of the squared correlation matrix, plus the sum of the partial covariances p_{ij}.

$$KMO_j = \frac{\sum_{i \neq j} r_{ij}^2}{\sum_{i \neq j} r_{ij}^2 + \sum_{i \neq j} p_{ij}} \ . \tag{2.41}$$

The overall Kaiser-Meyer-Olkin measure of sampling adequacy can also be calculated using the model from equation (2.42). The overall Kaiser-Meyer-Olkin score (KMO) is equal to the sum of the squared

correlations r_{ij}, for all the variables divided by the sum of the squared correlation matrix, for all the variables plus the sum of the partial covariances p_{ij}, for the all the variables,

$$KMO = \frac{\sum_{i \neq j} \sum r_{ij}^2}{\sum_{i \neq j} \sum r_{ij}^2 + \sum_{i \neq j} \sum p_{ij}} \ , \tag{2.42}$$

taking the overall combination of the variables into account and where $i \neq j$.

As a general rule the overall Kaiser-Meyer-Olkin measure should be greater than 0.8 but scores greater than 0.6 are still acceptable [2].

The results from factor analysis can be validated for interpretability and theoretical stability. Using Hatcher's [53], [67] interpretability criteria, the four interpretability criteria are:

1. a factor should have at least 3 variables with significant loadings, a loading greater than 0.40 is considered significant [67];

2. variables that load onto to the same factor should share the same conceptual meaning;

3. variables that load onto different factors appear to measure or represent different constructs; and

4. the rotated factor pattern represents a simple construct, meaning that most variables have a relatively high loading (greater than 0.40 [67]) on only one factor and most factors have at least one variable with a relatively high loading to start with and the remaining loadings tapering off.

The results can also be tested for reliability using a consistency technique of splitting the sample and testing if the results sets are similar, provided the sample size is not small. This technique tests if the pattern in the results is maintained within the different sub-sets calculated by splitting the data-set and showing that the results are reliable [29].

2.4 Determining the efficiency of advertising using Data Envelopment Analysis

Data envelopment analysis (DEA) is a linear programming technique used to compare the efficiency of different entities. Efficiency is a measure of how well the inputs are used to generate outputs, compared to the other entities in the model. Farrell [25] published an article in 1957 on the measurement of productive efficiency. Twenty one years later this article was the base for the research done by Charnes, Cooper and Rhodes [12], (CCR_{DEA}), resulting in the DEA efficiency frontier being computed for the first time [41]. Data envelopment analysis has primarily been applied to the production and manufacturing environment, limited research has been done applying the method to advertising [41] until 2001.

Luo and Donthu published a study in 2001 using DEA to benchmark advertising efficiency in the United States across the top 100 firms by advertising spend. In their research they performed an extensive review of past studies on advertising efficiency, finding that most companies focused on the dynamics of the advertising variables or advertising mix and little on addressing efficiency [41]. The objective of the study was to measure advertising efficiency over time, using DEA as it allows for multiple inputs and outputs and calculates relative efficiency scores. The input variables used were mass-media advertising

types such as print, broadcast or TV and outdoor billboards and the output variables were sales and income. DEA was used in this study because of its ability to inform how the inputs and/ or outputs should be adjusted to transform inefficient advertisement into efficient advertisement [41].

As one of the first steps to insure that the inputs do have a relationship with the output variables, Pearson's first order correlation analysis was used to test for a relationship. The final result of the Luo and Donthu [41] study is that DEA is a useful tool in quantifying advertising efficiency and the method is practical for the industry allowing for multiple inputs and outputs. The overall results are consistent with previous studies showing that there are inefficiencies within advertising spend and concluded that a 20% reduction in spend would result in the same output [41].

Similar to Luo and Donthu [41], another study performed in 2012 also looked at advertising efficiencies over time using DEA. Hezekiah, Ramakhrishnan and Shaban [32], investigated the relative advertising efficiencies of 17 FMCG firms in India from 2006 to 2012 [32]. DEA was used in the study as it allows for competitors to be included in the model, a benchmark to be generated and again allowing for multiple inputs and outputs [32]. The study used data from the Centre for Monitoring Indian Economy, starting with the top 500 companies on the Bombay Stock Exchange. Firstly narrowing this down to 29 FMCG firms and then selecting a sample of 17 for the analysis based on the firms media expenditure. The input variable used in this study was an aggregate of the total advertising spend by a firm and the output variable was the sales revenue for the seven years under investigation. The results once more found inefficiency in the advertising spend, concluding that a 50% reduction in spend would result in the same output [32]. This study cited two limitations, the limited number of firms or entities used and the high level of aggregation on the advertising variables [32].

Most recently an extensive study into advertising spend efficiency among top United States advertisers has been published by Cheong, De Gregorio and Kim [14]. The alternative hypotheses are:

$H1_a$: from 1985 to 2012, the overall level of the efficiency of advertising spend among the
top 100 advertisers will decrease over time, (overspending); and

$H1_b$: from 1985 to 2012, the overall level of the efficiency of advertising spend among the
top 100 advertisers will increase over time, (smart managers).

DEA was utilised to measure the efficiency of the different firms. The input variables included in the study were magazine, newspaper, TV, radio, outdoor and internet, with the firms' sales revenue as the output variable [14]. The firms in the study are from a wide range of industries, product types and brands but are all in the top spenders on advertising. This study also tested the relationships of the inputs and outputs using Pearson's first order correlation before applying them in DEA. The results found that 61% of the firms were overspending and are thus inefficient [14]. The overall results supported hypothesis $H1_a$ as the results showed a 32% decrease in advertising efficiency over time [14].

In a study by Thore [78] in 1996 titled, "DEA and the management of the product cycle: The U.S. computer industry", a long term investigation is performed using the DEA methodology in conjunction with the PLC methodology formulated by Bass [4] to aid in creating a framework for interpreting the results. The research used information on the U.S computer industry during the decade from 1981 to

1990, utilising the investment in real capital and expenditure on research and development (R&D) as input variables. The output variables used were sales revenue, profits and market capitalisation of each firm. The research concluded that there is a relationship between efficiency and the product life cycle or stage [78]. The results of this investigation were consistent with other DEA studies examined by Thore [78], finding that most companies spend heavily to launch new entrants into the market and then struggle to maintain them, all the while continuing to launch a stream of new innovations, this was found to be inefficient [78]. A small group of companies like Apple did manage to manufacture products with long and substantial product life cycles and therefore remained on the efficiency frontier [78].

2.4.1   Data envelopment analysis

Data envelopment analysis will first be illustrated using a simple example before the next section on CCR DEA. The general concepts will be illustrated using the example of a doctor's practice with eight different doctors working at the practice, each doctor working 160 hours in a month. In Table 2.1 the number of examinations and surgeries performed by each doctor is shown. As each doctor only has 160 working hours in a week, there are trade-offs between the number of examinations and surgeries that any one doctor can do in a month [56].

Doctor	Number of examinations	Number of surgeries
1	48	68
2	12	80
3	35	76
4	31	71
5	20	70
6	20	105
7	36	53
8	15	65

TABLE 2.1: *Example of the outputs of each doctor after working* 160 *hours in a month*

DEA can be used to answer the question of which doctor is the most productive or has the best yield for their 160 hours of work per month. A scatter plot as shown in Figure 2.3 can assist in giving insight into which doctor is most efficient with their time. The scatter plot shows that doctor number six performed the most surgeries and doctor number one the most examinations. Doctor number six performed more examinations and surgeries than doctors two, five and eight. Doctor number one performed more examinations and surgeries than doctor number seven, while doctors two and four performed more surgeries than doctor one but less examinations.
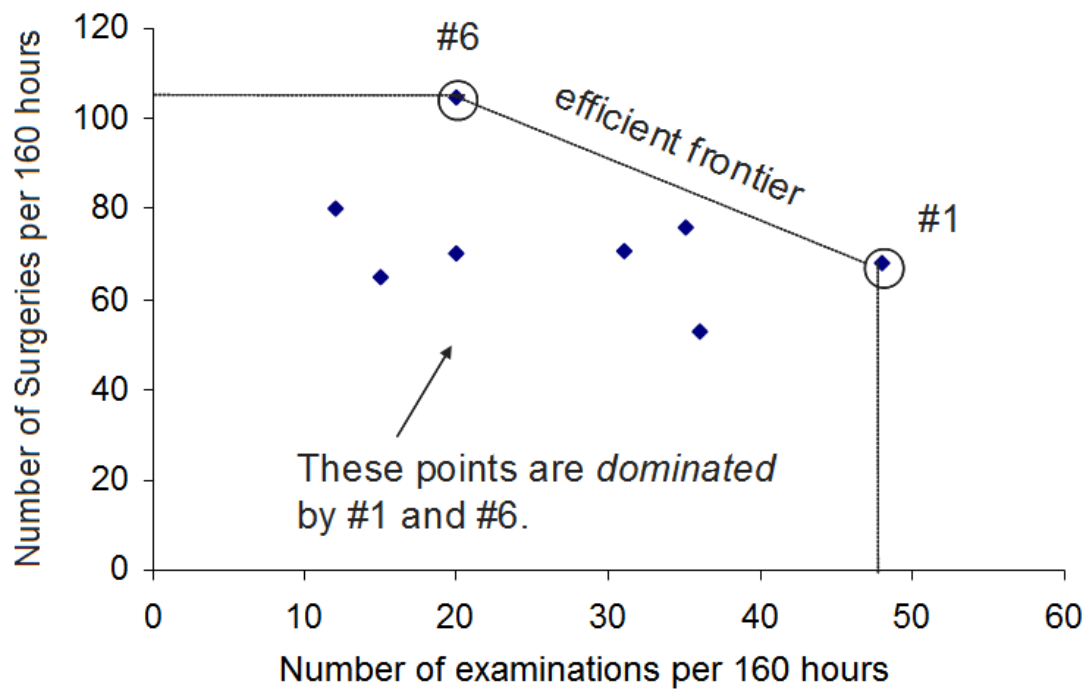
FIGURE 2.3: *Scatter plot of the number of examinations and surgeries performed in a* 160 *hour month for eight doctors [56]*

Figure 2.3 displays the efficient frontier, this is the line displaying the relationship between the outputs, where a doctor cannot increase an output without compensating a decrease in other outputs. Doctors one and six define the most efficient trade-off between the two outputs.

Using the efficient frontier that the outputs from doctors one and six have calculated, the relative distance to the frontier defines efficiency for the other doctors, as displayed in Figure 2.4. The 'nearest' efficient point on the efficient frontier defines a reference set of outputs and a linear combination of the reference set of inputs and outputs defines a hypothetical most efficient 'doctor'.
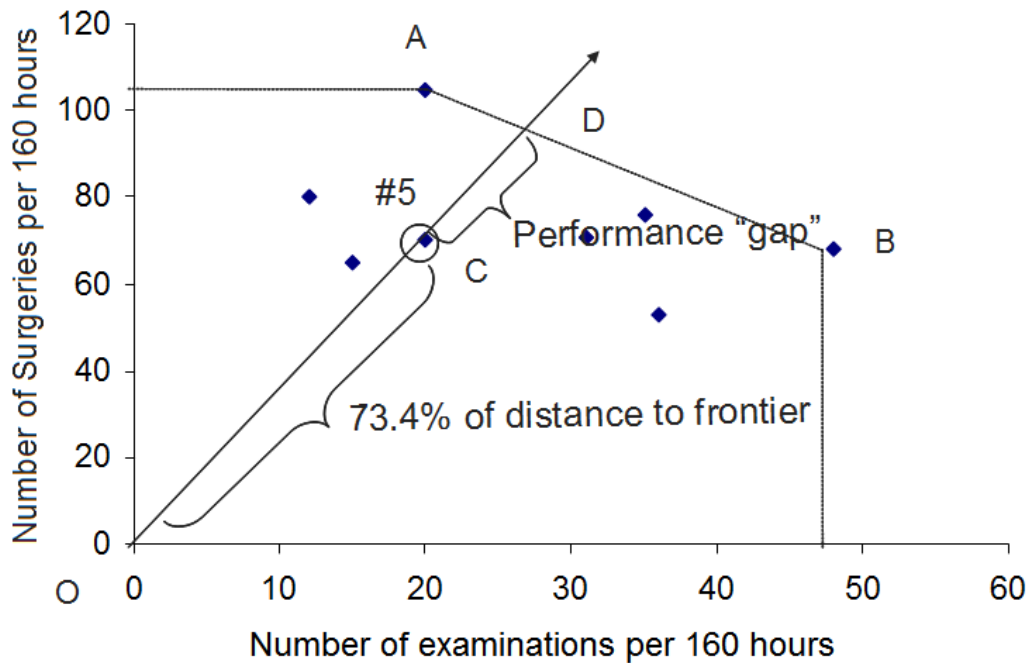
FIGURE 2.4: *Scatter plot displaying the efficiency frontier of for the eight doctors [56]*

The efficiency of doctor 5 is

$$\frac{the\ distance\ from\ O\ to\ C}{the\ distance\ from\ O\ to\ D}.$$

This simple example give an illustration into CCR data envelopment analysis that will be discussed in more detail in the section to follow.

2.4.2   CCR data envelopment analysis

DEA is a linear programming technique applied to compare the efficiency of different entities, that is how well are inputs utilised to calculate the outputs. DEA can be used to minimise the inputs needed to get the same output or to maximise the possible output. The linear programming model is run for each decision making unit (DMU) or entity separately. The model finds the minimum amount of inputs that will be used to produce the required output. The 'inefficient' DMUs are further analysed to identify the best input and output combinations that will make the DMU efficient. The best or most efficient DMUs can be identified and used as a benchmark, a weighted average of their inputs and outputs are calculated to determine target values for inputs and outputs of less efficient DMUs or entities [57].

Equation (2.43) shows the ratio used in DEA to determine efficiency. DEA uses relative efficiency, efficiency is relative to the information provided to the model. Efficiency is obtained by any DMU when no linear combination of other units in the model results in a composite unit that produces at least as much output using the same inputs [57].

$$\text{Efficiency} = \frac{Results\ Obtained}{Resources\ Used} \quad \text{or} \quad \frac{Weighted\ Sum\ of\ Outputs}{Weighted\ Sum\ of\ Inputs} \ . \tag{2.43}$$

Using the CCR DEA methodology as originally defined by Charnes, Cooper and Rhodes [12] with the objective of minimising the inputs and maximising the outputs, the efficiency of a DMU or entity i is defined by [57],

$$\text{Efficiency of } DMU_i = \frac{\sum_{j=1}^{n_O} O_{ij} w_j}{\sum_{j=1}^{n_I} I_{ij} v_j} \ , \tag{2.44}$$

where O_{ij} represents the value of the DMU_i, $i = 1, \ldots, d$, on output variable j, and I_{ij} represents the value of DMU_i on the input variable j, also letting w_j represents the non-negative weight assigned to the output variable j and v_j the non-negative weight assigned to input variable j [57]. The number of input variables is denoted by n_I and the number of output variables by n_O.

Formulating the linear programming model for any arbitrary unit i, starts with the objective function [57],

$$MAX\, z_i = \sum_{j=1}^{n_O} O_{ij} w_j \tag{2.45}$$

subject to

$$\sum_{j=1}^{n_O} O_{kj} w_j - \sum_{j=1}^{n_I} I_{kj} v_j \le 0 \ , k = 1, \ldots, \ d \ , \tag{2.46}$$

$$\sum_{j=1}^{n_I} I_{ij} v_j = 1, \text{for arbitrary unit } i, \text{and} \tag{2.47}$$

$$w_1, w_2, \ldots, w_{n_O}, v_1, v_2, \ldots, v_{n_I} \ge 0 \ . \tag{2.48}$$

This objective is computed for each DMU or entity separately, for $k = 1$ to the number of DMUs. The objective function is subject to three constraints: the efficiency constraint of each DMU (equation (2.46)), the sum of the weighted inputs for DMU, i must equal one (equation (2.47)) and the input and output weights can not be negative (equation (2.48)).

The efficiency constraints for each DMU in equation (2.46) ensures that each DMU can not be more than 100% efficient. These constraints are formulated by setting the weighted sum of the outputs minus the weighted sum of the inputs to be less than or equal to zero [57] for each DMU.

Unbounded solutions are situations where the objective function for a DMU is infinite, meaning that greater than 100% efficiency would be possible. To prevent this the sum of the weighted inputs for the DMU under investigation in equation (2.47) is set to equal one [57].

The foundation of the DEA method is to determine the non-negative weights that yield the maximum output while not letting the efficiency ratio exceed 100% for a DMU. This last constraint is very important because if the weights were allowed to be negative it would seriously jeopardise the usefulness of results in industry.

A 100% efficient ratio does not mean that these DMUs are operating in the best possible way or there is not a way to make them more efficient. The result means that no linear combination of the DMUs under investigation or in the model resulted in a composite unit that produces at least as much output using less

of the same amount of input [57]. The efficiency ratio is relative to the DMUs under investigation. The DEA method calculates an efficiency frontier line and all DMUs who lie below the line are inefficient, efficient DMUs will fall on the line. Unlike regression results that form a common average or regress to the mean, DEA will find a reference for each unit or entity when comparing it to the efficient frontier [41].

A limitation of DEA is the assumption that each DMU is capable of achieving the same level of performance or efficiency [57]. This makes the selection of DMUs that will be compared to each other very important, as this method is very susceptible to outliers. The input and output variables that are used need to be consistent for each DMU or entity. The model is also sensitive to the number of DMUs included in the set. As a general rule, the number of input variables multiplied by the number of output variables should be less than the number of DMUs [81].

Although DEA does have limitations it also has some advantages, in that it is a non-parametric procedure and makes no assumptions about the probability distribution of the variables being assessed [81]. The weighting or any other prior knowledge about the inputs is not necessary and the objective function does not assume any specific functional form relating inputs to outputs [81]. The scale or unit of measure does not need to be consistent across the inputs and outputs [57]. The method allows for multiple inputs and outputs to be used, as best describes the DMUs.

The most advantageous attribute of the DEA method is the creation of a composite unit, the inputs and outputs are a linear combination of the inputs and outputs of the efficient units [57]. It is possible to determine an efficient composite unit for a given inefficient unit as follows:

- **Step 1:** solve the DEA problem for the inefficient unit e;

- **Step 2:** calculate the shadow prices λ_k, $k = 1$ to the number of DMUs. λ_k will be equal to zero for inefficient units and zero or positive for efficient units;

- **Step 3:** calculate the input values for the composite unit associated with the inefficient unit e as follows,

$$\hat{I}_{ej} = \sum_{k=1}^{d} \lambda_k I_{kj} \ , \ j = 1, \ldots, n_I \ , \tag{2.49}$$

  for each input variable j, and I_{kj} is the input value of DMU k;

- **Step 4:** calculate the output values for the composite unit associated with inefficient unit e as follows,

$$\hat{O}_{ei} = \sum_{k=1}^{d} \lambda_k O_{ki} \ , \ i = 1, \ldots, n_O \ , \tag{2.50}$$

  for each output variable i and O_{ki} is the output value of DMU k.

In step 3 and 4, linear combinations of efficient units are calculated because $\lambda_k \geq 0$ for efficient units only. The input and output values of the composite unit can be used as a benchmark for the input and output values of the inefficient units.

There are many different methods for performing DEA [57], the method described above in equations (2.45) to (2.48) is also referred to as the CCR output orientated model, the CCR input orientated DEA method minimises the input. The BCC model considers variables return to scale and is obtained by adding a new constraint to the CCR input orientated model.

2.5   Chapter summary

This chapter provides an overview of techniques that are used to analyse consumer buying behaviour. Also, techniques to be used such as factor analysis, cluster analysis, BCG growth share matrix, regression analysis and DEA are described.

CHAPTER 3

Data

Contents

Company A has provided data on product sales and pricing from one of the largest supermarket retailers in South Africa. Company A also provides market research survey data on brand-health measurements from an external research house, information on mass-media advertising spend from the alcoholic beverage industry spend tracker produced by Nielsen [52] and internal data records on in-store promotions. Company A has given permission for all the data to be used in this study, the structures and attributes of the different datasets can be displayed but not the actual data records.

In this chapter the different data sources, their granularity and time frames will be reviewed. Each data source will be discussed separately, as well as aspects such as whether the dataset is either generated from survey collection of data or from transactional recordings of consumer purchases. The data and data sources will be described and relationships will be investigated in this chapter.

3.1   Data refinement

The dataset supplied by one of the largest supermarket retailers in South Africa includes information on 1126 liquor brands and 4738 products with sales in the period January 2013 to December 2017. The sales volume of the different brands ranges from 1 litre per month to over 2 million litres depending on the brand. The starting point for this research is to determine the set of brands or products that will be modelled. If the market in totality is used, then the dataset will be extremely large and calculations will be arduous.

The Pareto principle or more commonly known as the $80/20$ rule, states that 80% of the total sales volume will come from only 20% of the brands available [68]. Pareto was an Italian mathematician who made the conclusion when studying income levels that about 80% of the effect comes from 20% of the cause. The Pareto principle has been used extensively in marketing, with some researchers finding that as little as 4% of the brands available account for over 80% of the purchases [68].

In the alcoholic beverage industry the segments wine, spirits and RTDs are not homogeneous due to the alcohol content by volume. For this reason the Pareto principle is applied to the total market and per segment for both volume and value. A product is included if it is in the top 80% of the total market's share or of its segment's market share, this forming the cut-off point. The products selected that contribute to the top 80% volume and value, are 344 wine products, 135 spirits products and 115 RTDs. Table 3.1 gives an illustration of how the Pareto process is applied to the dataset.

After the Pareto Principle is applied the dataset is reduced to 594 products, which results in 91% of the volume and 86% of the value, coming from 13% of the products available. The 594 products will be used for regression analysis, factor analysis, cluster analysis and the BCG growth share matrix analysis. These are the methods that will be used to determine a product's competitor set, once the competitor set is determined, the competitor set will be used further in regression analysis and DEA.

Table indicating if a product, when ranked by share, is contributing to the top 80% by volume and value					
Product	Total market: volume share	Total market: value share	Segment: volume share	Segment: value share	Cut-off
Product 1	In top 80% share	In top 80% share	In top 80% share	In top 80% share	Keep
Product 2	Below top 80% share	Below top 80% share	Below top 80% share	In top 80% share	Keep
Product 3	In top 80% share	Below top 80% share	In top 80% share	Below top 80% share	Keep
Product 4	Below top 80% share	In top 80% share	Below top 80% share	In top 80% share	Keep
Product 5	Below top 80% share	Below top 80% share	In top 80% share	Below top 80% share	Keep
...	..	...	...	...	...
Product 4738	Below top 80% share	Below top 80% share	Below top 80% share	Below top 80% share	Remove

TABLE 3.1: *Example of the Pareto principle applied for the total market and per segment*

3.2 Volume, value and price

In 2010 it became legal for supermarkets to sell spirits in their own liquor stores. Since then retailers like Spar, Pick n Pay and Checkers have dramatically changed the alcoholic beverage landscape. These type of outlets are classified as specialist retailer outlets. Over the past 5 years these outlets have taken market share from small independent bottle stores and now account for over 13% [52] of the off-consumption market. Advertising spend is weighed towards these outlets due to the benefits of scale they offer brand marketers versus small independent bottle stores. Space to advertise in each outlet is limited and paid for by brand marketers.

Electronic till data has been collected from one of the largest supermarket retailers in South Africa, with over 400 liquor outlets. The 400 liquor outlets are located near to the entrance of their food supermarkets.

The liquor outlets have electronic till points and record all the sales information on a detailed granular level per transaction: product, quantity, price, date, time and store. The retailers then aggregate this data up to per week, store, segment, category, brand, product, quantity and price. This data will be used on a national level and is available from January 2013 up until December 2017. The dataset is a collection of consumers' actual purchases, therefore this information is not claimed behaviour as with a survey but actual purchase behaviour. Table 3.2, is an example of an extract of the dataset. The actual brand and product names as well as the numeric quantity for the products' volume, value and average unit price (AUP) are not shown here as Company A has a confidentiality agreement with the specialised supermarket retailer not to publish the figures.

In this research each product will be denoted by the segment and category that they belong to, also by an identification number and the product's packaging size. For example RTD-CIDER-P18-660ML, RTD-CIDER-P6-330ML, SPIRITS-LIQUEURS-P16-750ML and WINE-RED-P102-750ML.

Segment	Category	Product	Pack Size	Packaging Unit	Week Ending	Volume (L)	Value (R)	AUP (R)
BEER	BEER	P1	340 ML	6x340ML	08/01/2017	XXX	XXX	XXX
BEER	CIDER	P2	330ML	12x340ML	08/01/2017	XXX	XXX	XXX
BEER	FABS	P3	300ML	6x340ML	08/01/2017	XXX	XXX	XXX
SPIRITS	BRANDY	P4	750ML	24X300ML	08/01/2017	XXX	XXX	XXX
SPIRITS	CANE	P5	750ML	12x750ML	08/01/2017	XXX	XXX	XXX
SPIRITS	GIN	P6	750M	12x750ML	08/01/2017	XXX	XXX	XXX
SPIRITS	LIQUEURS	P7	750ML	12x750ML	08/01/2017	XXX	XXX	XXX
SPIRITS	RUM	P8	750ML	12x750ML	08/01/2017	XXX	XXX	XXX
SPIRITS	VODKA	P9	750 ML	12x750ML	08/01/2017	XXX	XXX	XXX
SPIRITS	WHISKY	P10	750 ML	12x750ML	08/01/2017	XXX	XXX	XXX
WINE	BAG IN BOX WINE	P11	5LTR	4X5LTR	08/01/2017	XXX	XXX	XXX
WINE	FORTIFIED WINE	P12	750ML	12x750ML	08/01/2017	XXX	XXX	XXX
WINE	PERLE WINE	P13	750ML	12x750ML	08/01/2017	XXX	XXX	XXX
WINE	RED WINE	P14	750ML	6x750ML	08/01/2017	XXX	XXX	XXX
WINE	ROSE WINE	P15	750ML	6x750ML	08/01/2017	XXX	XXX	XXX
WINE	SPARKLING WINE	P16	750ML	6x750ML	08/01/2017	XXX	XXX	XXX
WINE	WHITE WINE	P17	750ML	6x750ML	08/01/2017	XXX	XXX	XXX

TABLE 3.2: *Example of electronic till data that has been collected from one of the largest supermarket retailers in South Africa*

Figure 3.1 displays the total sales volume, value and AUP per week aggregated for the total liquor market for the three year period from January 2015 to December 2017. The graphs show that there is a lot of variation within the different months for volume and value. The volume and value variables in the data are seasonal. The AUP is more consistent across the months as price is not a highly seasonal variable. Seasonality is due to the increase in consumption and consumers over the festive holiday periods and month-end or pay day weekends.

Volume is measured in litres, to calculate the volume of a specific product the pack size is multiplied by

the number of units sold of that specific product,

$$Volume = \frac{Pack\ size(i.e 750ML)}{1000} \ \times \ number\ of\ units\ sold\ . \tag{3.1}$$

The AUP is the average unit price per week for a specific product, the product's price is averaged over all the outlets. The sales value per product, is calculated by multiplying the number of units of a specific product sold by the price of that product,

$$Sales\ Value = Number\ of\ Units\ Sold\ \times\ AUP\ . \tag{3.2}$$
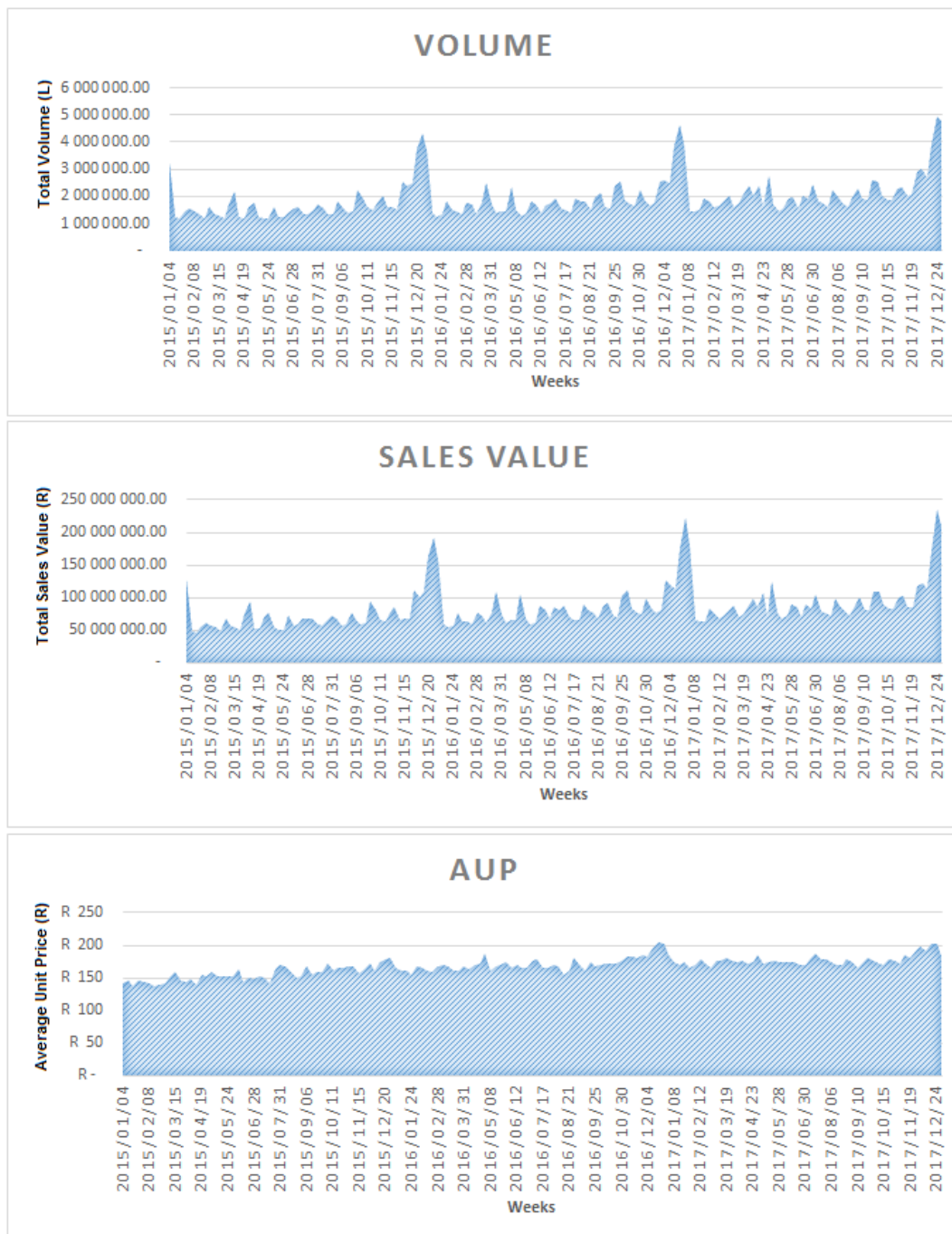
FIGURE 3.1: *Plots of the total sales volume, value and AUP per week aggregated for all alcoholic beverages for the three year period from January 2015 to December 2017*

The box and whisker plots are for the twelve month period from July 2016 to June 2017, for the variables volume, value and AUP and are displayed in Figure 3.2. The box and whisker plots for volume and value of individual products show that the alcoholic beverage landscape is skewed to the right, with numerous small volume and value products, these in comparison to the few products with relatively large volume and value. The plots show that there are a number of products that have a monthly total volume and value greater than the third quartile. The AUP does have a lot of deviation, as the pricing of products ranges from cheap basic products to luxury premium products and the distribution is also skewed to the right.
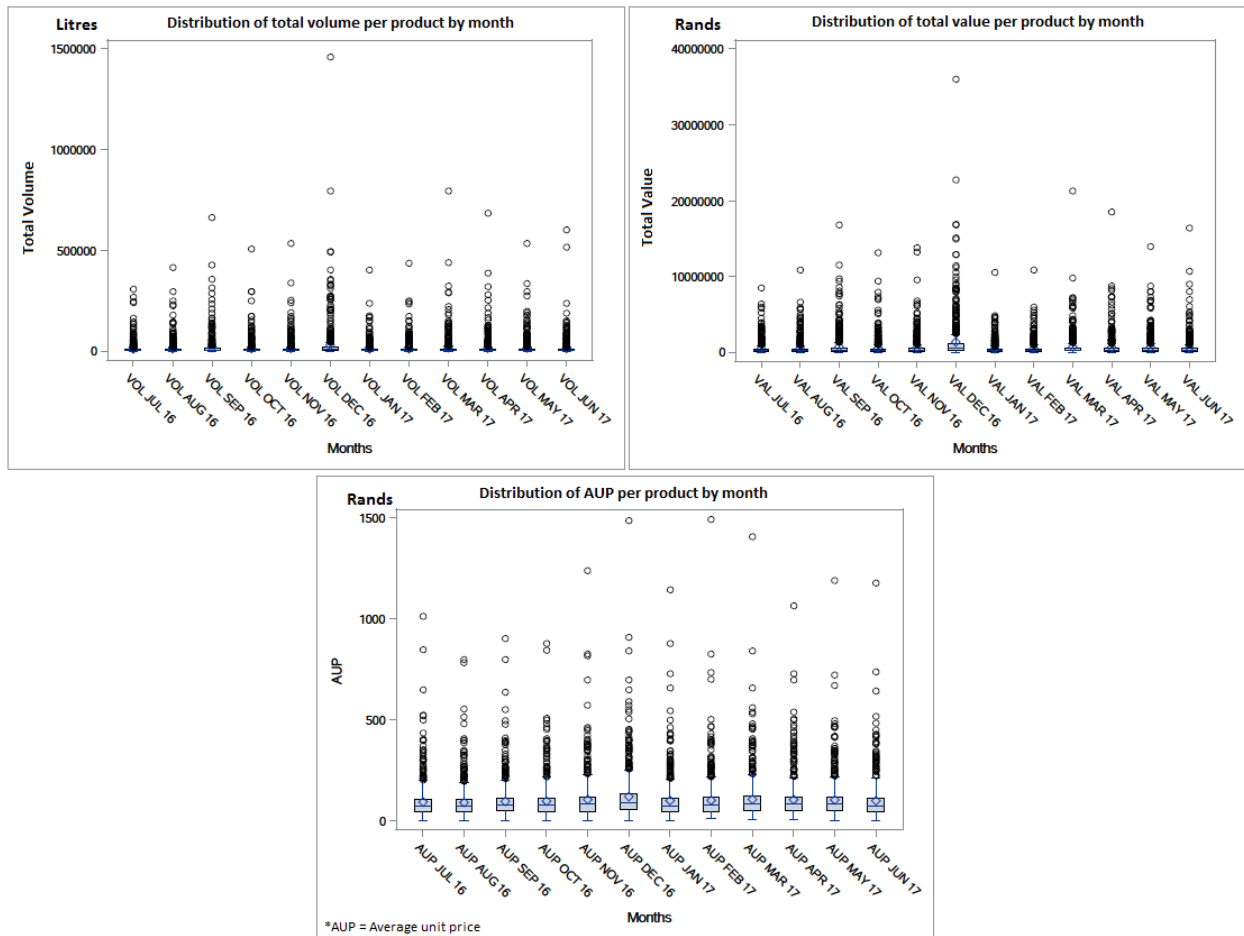
FIGURE 3.2: *Box and whisker plots of the numerical dataset for total volume, value and AUP for each product*

In Figure 3.3 the products are arranged from large to small with regards to total sales volume, total value and AUP values. Each product in the market is represented by a coordinate or 'dot' on the graph. Many of the methods that will be used, such as factor and cluster analysis, are sensitive to outliers in the data. In Figures 3.2 and 3.3 the variables total volume and total value each have one significant outlier. This is a market leading beer product, which has 9% of the total market volume share. Outliers should not always be excluded as they are often important to the investigation and can yield interesting results. As this beer product is a significant product on the market, it therefore can not be excluded from the analysis. Data transformations, such as the square root or natural logarithm transformations, can be used to standardise the dataset and reduce the effect of the outlier. The variable AUP also has outliers but to a lesser scale than total volume and total value, but will also benefit from a data transformation technique.
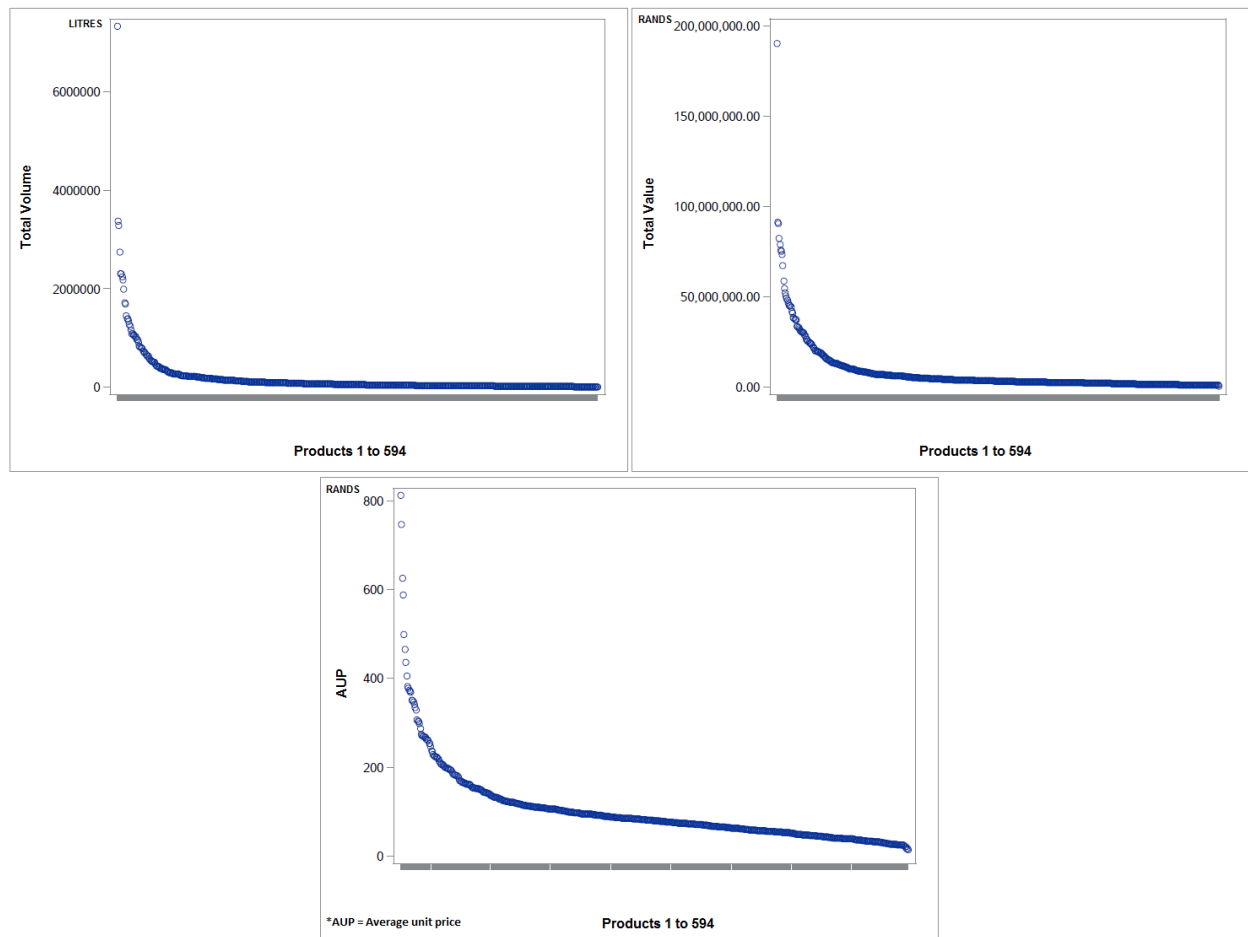
FIGURE 3.3: *Plots of volume, value and AUP, for each product in decreasing order of total volume, value and AUP for the period January 2017 to December 2017*

In Table 3.3 the Pearson's correlation coefficients and p-values are shown for the top 10 products by sales volume market share. Table 3.3 is just an extract but the results of the full dataset of 594 products follow the same trend, most products are highly correlated with each other.

In Table 3.3 the highlighted correlation coefficients are those that are greater than 0.5. Usually when products are competing with each other for consumer preference, an increase in sales for one product will mean a decrease in sales for the other competitor products, resulting in negative correlation coefficients. The results show that many of the 10 products by total sales volume market share are from the beer segment and have positive correlation coefficients. This shows that these products might not be competitors but rather complementary purchases, or the strong seasonal nature of the category is a more intense force than competition. The results show that most of the products observed have p-values ≤ 0.05 indicating that the correlations are significant, even if the correlation is not greater than 0.5.

Pearson's Correlation Coefficients And Corresponding P-Values

Product	WINE-BIB-P79-5LTR	WINE-BIB-P82-5LTR	RTD-BEER-P14-330ML	RTD-BEER-P35-440ML	RTD-BEER-P51-660ML	RTD-BEER-P21-340ML	RTD-BEER-P27-440ML	RTD-BEER-P58-750ML	RTD-BEER-P1-330ML	RTD-BEER-P20-340ML
WINE-BIB-P79-5LTR r	**1**	**0.84**	0.3	0.48	**0.52**	0.1	-0.14	0.26	0.43	0.49
WINE-BIB-P79-5LTR P-Value		0	0	0	0	0.23	0.07	0	0	0
WINE-BIB-P82-5LTR r		**1**	0.34	**0.58**	0.32	0.16	0.08	0.43	0.37	0.42
WINE-BIB-P82-5LTR P-Value			0	0	0	0.05	0.33	0	0	0
RTD-BEER-P14-330ML r			**1**	**0.71**	0.09	**0.54**	**0.51**	0.39	**0.66**	**0.73**
RTD-BEER-P14-330ML P-Value				0	0.26	0	0	0	0	0
RTD-BEER-P35-440ML r				**1**	0.3	0.45	0.35	**0.56**	**0.6**	**0.66**
RTD-BEER-P35-440ML P-Value					0	0	0	0	0	0
RTD-BEER-P51-660ML r					**1**	0.02	-0.59	0.32	0.21	0.25
RTD-BEER-P51-660ML P-Value						0.78	0	0	0.01	0
RTD-BEER-P21-340ML r						**1**	0.3	0.26	0.27	0.36
RTD-BEER-P21-340ML P-Value							0	0	0	0
RTD-BEER-P27-440ML r							**1**	0.23	0.28	0.32
RTD-BEER-P27-440ML P-Value								0	0	0
RTD-BEER-P58-750ML r								**1**	0.27	0.35
RTD-BEER-P58-750ML P-Value									0	0
RTD-BEER-P1-330ML r									**1**	**0.95**
RTD-BEER-P1-330ML P-Value										0
RTD-BEER-P20-340ML r										**1**
RTD-BEER-P20-340ML P-Value										

TABLE 3.3: *Pearson's correlation coefficients and p-values for the top 10 products by total sales volume market share for the period January 2017 to December 2017*

3.3   Brand-health

Company A uses an external research house which is one of the largest research agencies worldwide providing field market research, consumer insights and brand-health measurement [34]. Brand-health or brand equity has three basic components, they are: consumer perception, negative or positive effects and the resulting value [35]. This external research house has a large field research team that continuously, 48 weeks of the year, performs consumer surveys. The surveys are performed by face-to-face interviews,

using tablets as an aid. Only the 7 major cities in South Africa being: Johannesburg, Pretoria, Cape Town, Durban, East London, Port Elisabeth, Bloemfontein and Polokwane are included.

The surveys are administered to a representative sample of the South African population in the 7 major cities. The sample, while remaining representative of the area, could only include consumers who have consumed an alcoholic beverage in the last 7 days. Consumers who only drink beer and no other alcoholic beverage have been excluded as they skew the results, therefore the results of the beer brands will be under-read. The research team surveys 800 different consumers a week, adding up to a sample of 9600 consumers surveyed a year. The aim of the field survey research is to have a continuous study with the past 7 day 'non-(beer only)' or consumers who don't drink beer exclusively and are regular alcohol drinkers in urban metro South Africa, to monitor and track the alcohol landscape on key consumer health measures [34].

The survey questions are asked on a brand level for most brands, however if a brand has products that have a large market share or are different in taste or category then the questions are asked on a product level. For example a brand that has vodka products and RTD's will be separate in the survey questions.

Using consumer surveys of a representative sample of the alcohol consuming population, the consumer sentiment can be tracked and measured over time. The brand equity, image and other factors influencing a brand all contribute towards the consumer sentiment towards a brand. The brand imagery, market factors, market share and power in the mind are measures that have been tracked by an external research house [34].

Table 3.4 is an example of an extract of the twelve month rolling dataset for the period January 2017 to December 2017 provided by the external research house, the actual questions asked and responses are not shown. The questions asked are proprietary of the external research house and can not be published. Company A has not given permission for the questions or actual responses to be published in this research.

A sample of Brand-health survey questions expressed as a percentage of the sample population					
Product	Question 1	Question 2	Question 3	...	Question 78
WINE-BIB-P82	X.XX %	X.XX%	X.XX%	...	X.XX%
SPIRITS-LIQUEURS-P2	X.XX %	X.XX%	X.XX%	...	X.XX%
SPIRITS-WHISKY-P1	X.XX %	X.XX%	X.XX%	...	X.XX%
SPIRITS-WHISKY-P14	X.XX %	X.XX%	X.XX%	...	X.XX%
SPIRITS-RUM-P7	X.XX %	X.XX%	X.XX%	...	X.XX%
RTD-BEER-P2	X.XX %	X.XX%	X.XX%	...	X.XX%
RTD-BEER-P20	X.XX %	X.XX%	X.XX%	...	X.XX%
RTD-BEER-P22L	X.XX %	X.XX%	X.XX%	...	X.XX%
SPIRITS-GIN-P1	X.XX %	X.XX%	X.XX%	...	X.XX%
SPIRITS-WHISKY-P3	X.XX %	X.XX%	X.XX%	...	X.XX%
SPIRITS-BRANDY-P12	X.XX %	X.XX%	X.XX%	...	X.XX%
RTD-CIDER-P1	X.XX %	X.XX%	X.XX%	...	X.XX%
SPIRITS-WHISKY-P4	X.XX %	X.XX%	X.XX%	...	X.XX%
SPIRITS-WHISKY-P23	X.XX %	X.XX%	X.XX%	...	X.XX%
SPIRITS-WHISKY-P50	X.XX %	X.XX%	X.XX%	...	X.XX%
SPIRITS-WHISKY-P53	X.XX %	X.XX%	X.XX%	...	X.XX%
SPIRITS-WHISKY-P9	X.XX %	X.XX%	X.XX%	...	X.XX%
SPIRITS-BRANDY-P1	X.XX %	X.XX%	X.XX%	...	X.XX%
SPIRITS-BRANDY-P3	X.XX %	X.XX%	X.XX%	...	X.XX%
SPIRITS-BRANDY-P2	X.XX %	X.XX%	X.XX%	...	X.XX%

TABLE 3.4: *Example of an extract of Brand-health survey questions expressed as a percentage of the sample population for the period January 2017 to December 2017*

The brand-health survey questions can be categorised into four main groups: brand imagery, market factors, power in the mind and market share.

Brand imagery is developed over time through advertising campaigns with a consistent theme and is authenticated through the consumers' direct experiences. Brand imagery signifies what the brand currently stands for. It is a set of beliefs that consumers hold for a specific brand, it is the consumers' perception about the brand. In the research study there are 44 questions relating to brand imagery, (Q1 - Q44). The 'yes' responses are measured as a percentage of the sample population.

Market factors measure the positive or negative barriers for a consumer to purchase a brand. The factors that are measured are accessibility, product range, purchaser, promotions, information, influencer and price. In the research study there are 31 questions relating to market factors, (Q45 - Q76). The 'yes' responses are measured as a percentage of the sample population.

- Accessibility is a measure of how readily a consumer can find the brand in the places that a consumer frequents.

- Purchaser measures if the person consuming the brand is also the person shopping for the brand.

- The product range measures if the brand has the pack formats that consumers want and if the packaging is attractive for the consumers.

- Promotions measure if consumers are aware of the brand's promotional activity and if they like the promotions the brand is doing.

- Information is a measure of how much consumers feel that the brands 'talks to them'.

- Price measures what consumers think the price should be, is a brand too expensive or too cheap.

- Influencer measures the word of mouth around the brand and its social acceptability.

Market share, (Q77), refers to the 100% share across all brands, therefore if one brand gains share, another will lose share. The respondents give an indication of their stated consumption behaviour in the past and perceptions of different brands. **Power in the mind**, (Q78), is a measure that takes into account consumers' desire or sentiment towards the brand and the measure links better than other methods with actual behaviour, making it an indication of future behaviour.

As with the numeric interval variables volume, value and price, the ratio variables for the brand-health questions also have outliers. In Figure 3.4, each 'dot' or response represents the score for a product for a brand-health question. The responses are a percentage, each 'yes' response is counted and measured as a percentage of the sample population. In Figure 3.4 the box and whisker plots for the variables show that most products score between 0 to 0.2, with scattered outliers. These variables will also benefit from a data transformation.
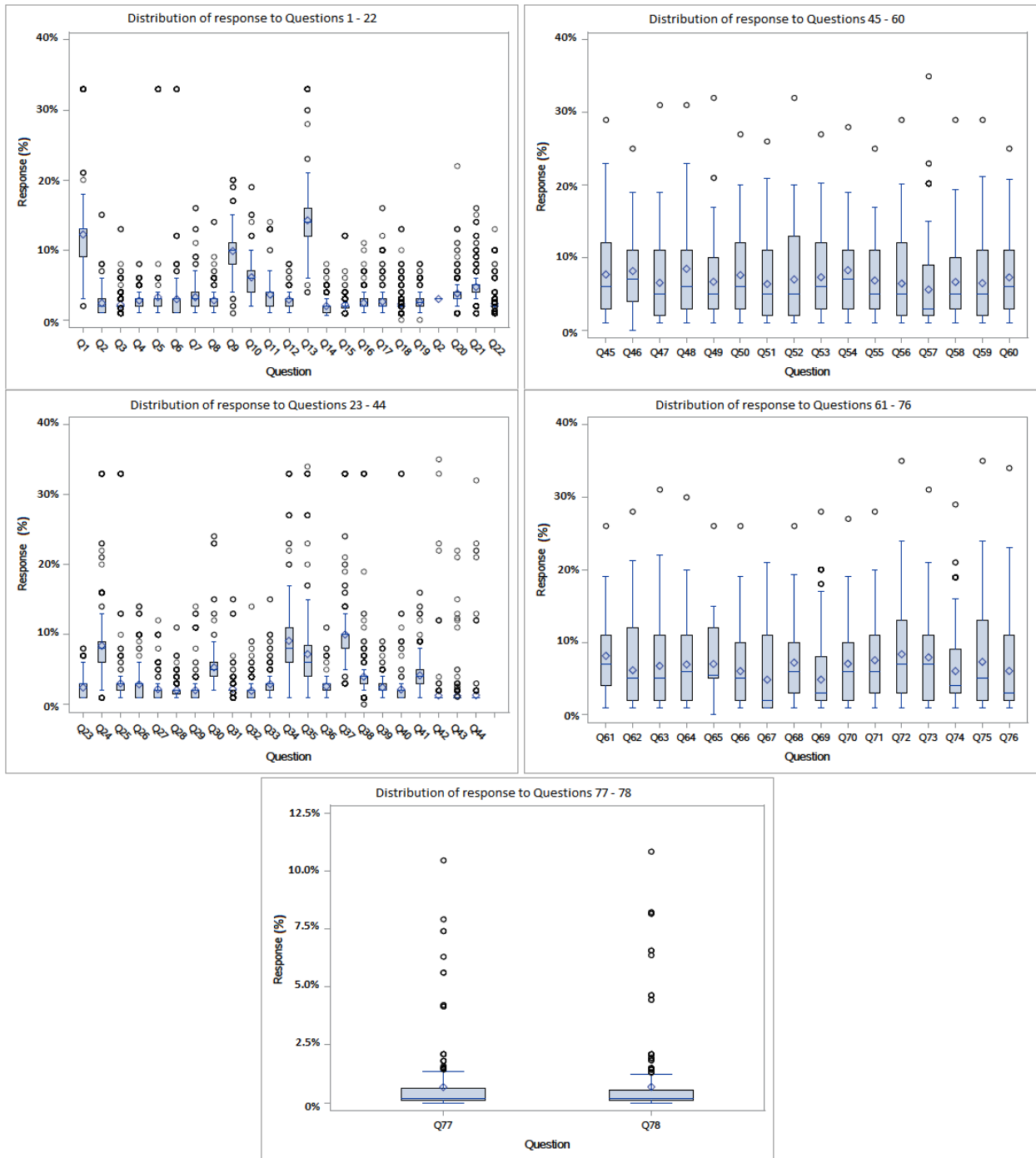
FIGURE 3.4: *Box and whisker plots displaying percentage responsiveness to each product's market factors (Q1-Q44), imagery (Q45-76), power in the mind (Q77) and market share (Q78) questions from an external research house [34]*

3.4   Brand attributes

Using the packaging information for each product, the brands can also be described by other attributes such as alcohol percentage, carbonation level, sugar content, pack size and taste profile. These are broad

categories but do give a factual description of the brand.

The categorical datasets provide additional attributes about the products. The categorical variables are of importance as they are descriptive attributes that consumers will use when deciding on a purchase, for example choosing between a sparkling or still product. The purchasing choices by consumers represent the repertoire set of a consumer [39]. The repertoire set of each individual consumer of a brand is aggregated to formulate the competitor set for that brand. A specific brand's competitor set are the other products or brands that most people will consider when making a purchase, if they purchase this specific brand or product.

The variables volume, value, AUP and brand-health are all interval variables. The categorical variables describing the brand attributes will be transformed into nominal variables. A (0 *or* 1) indicator variable will be used to identify a product having an attribute (1) or not possessing the attribute (0). Table 3.5 gives an overview of the brand attributes that will be used in this research.

Variable	Attribute	Type
Pack Size	Small Pack Wine	Indicator/Binary
	Medium Pack Wine	Indicator/Binary
	Large Pack Wine	Indicator/Binary
	Small Pack Spirits	Indicator/Binary
	Medium Pack Spirits	Indicator/Binary
	Large Pack Spirits	Indicator/Binary
	Small Pack RTDs	Indicator/Binary
	Medium Pack RTDs	Indicator/Binary
	Large Pack RTDs	Indicator/Binary
Alcohol percentage	Low alcohol	Indicator/Binary
	Mid-Alcohol	Indicator/Binary
	High Alcohol	Indicator/Binary
Carbonation level	Still	Indicator/Binary
	Carbonated	Indicator/Binary
	Sparkling	Indicator/Binary
Sugar content	Dry	Indicator/Binary
	Semi-Sweet	Indicator/Binary
	Sweet	Indicator/Binary
	Bitter	Indicator/Binary
Taste profile	Wine-Like	Indicator/Binary
	Beer-Like	Indicator/Binary
	Cider-Like	Indicator/Binary
	Fruity	Indicator/Binary
	Woody/Yeasty	Indicator/Binary
	Herbaceous/Green	Indicator/Binary
	Sweet-associated	Indicator/Binary
	Smoky/peaty	Indicator/Binary
	Smooth	Indicator/Binary
	Mouth coating/creamy	Indicator/Binary

TABLE 3.5: *Brand attribute variables names and type of variable*

3.5   Mass-media advertising

Industry spend per product on mass-media advertising types such as digital, magazine, newspaper, outdoor, radio and television (TV) adverts, is provided by Company A, as produced by Nielsen [52]. Mass-media or above the line (ATL) advertising is defined as advertising intended to reach a mass audience or the vast majority of the general public [39]. A simple way to differentiate between ATL advertising and below the line (BTL) is to ask the questions; does a person need to be in a specific geographic location to see or experience the advertisement? The ATL media type TV for example can be seen by many people in many different locations at the same time, the same is true for digital, magazine, newspaper and radio. Outdoor advertising is advertising on billboards, the side of a building or flag poles. While a person needs to be in a specific location to see the advertising, it can be experienced by many people at the same time. BTL advertising will be discussed in detail in the section below.

The information provided on the mass-media advertising types includes the Rand value spent on each advertising type by product and week. Kotler [39] wrote about the different types of ATL advertising variables, highlighting the advantages and limitations of each. Table 3.6 summarises his findings [39].

Type	Advantages	Limitations
Digital	Highly selective, interactive possibilities, two way communication and relatively low cost	High clutter and creative limitations on content
Magazine	High geographic and demographic selectivity, credibility and prestige and long life span	Long advertisement purchase lead time and no guarantee of position
Newspaper	Good local market coverage, broad acceptance and high believability	Short life span and low audience engagement
Outdoor	Flexibility, high repeat exposure and low cost	Limited audience selectivity and limitations on content
Radio	Mass use, high geographic and demographic selectivity and low cost	Audio presentation only, low audience attention and fleeting exposure
Television	Appealing to the senses and high reach	High clutter and less audience selectivity

TABLE 3.6: *Advantages and limitations of mass-media advertising types as defined by Philip Kotler [39]*

3.6   In-store advertising

BTL advertising happens at a specific geographic location, this is usually an off-consumption outlet such as a bottle store. An example would be if a company wanted to run a tasting drive, promoting their product and getting consumers to taste the product in-store. The company would not be able to go to all off-consumption outlets or bottle stores at the same time, it is unlikely that a company would have a budget large enough to go to all the outlets. The brand manager will need to pick a target group of outlets and then over a period of time have scheduled tastings in the different outlets. This means that a consumer needs to be in the specific outlet in which the tasting is being activated in at the time of the activation. Therefore this is not classified as a type of mass-media.

From their internal data Company A provides the information on BTL advertising in the specialised supermarket retailer. This information includes all advertising campaigns that ran in the stores of the specialised retailer. These are grouped into value adds, gifting, instant gratification, in-store engagement and main prizes. The information includes the Rand value spent on each advertising activity, the product

and time when it was done. Previously in Chapter 1.1 examples are given for each type of BTL advertising.

3.7 Chapter summary

In summary this chapter defines and describes the different datasets that Company A has provided. The interval variables volume, value, AUP and brand-health will all benefit from a data transformation due to highly correlated relationships and outliers. The categorical descriptive variables are displayed to show their level of differentiation. The mass-media spend types are described and compared according to Kotler's methodology [39]. The in-store advertising variables are also defined and described. The overall dataset of 4738 products is refined to 594 products using the Pareto principal.

CHAPTER 4

Determining a competitor set

Contents

This chapter describe the process and review of the product life cycle (PLC) methodology and regression analysis to determine a product's competitor set. The PLC methodology includes the use of factor analysis, cluster analysis and the BCG growth share matrix. The BCG growth share matrix will be used in this study to divide the products into four taxonomies or competitor sets. Products that are in the same quadrant should be in the same life cycle stage. Firms should make homogeneous investment decisions about these products and therefore the marketing efforts should be of the same scale.

Using the BCG growth share matrix requires the relative market share of each product to be known, thus in effect the biggest competitor needs to be known for each product. Cluster analysis will be used to group the products into small relative sets for the market share comparison. Cluster analysis requires the variables to be uncorrelated [7], therefore factor analysis will be used on the highly correlated variables to create uncorrelated scores representing the underlying relationship or pattern. Brands and products within the alcoholic beverage industry are highly correlated with each other by nature of the industry. Factor analysis can create uncorrelated scores so that distances between the observations are calculated more correctly in cluster analysis.

The sequential process that will be followed is to use factor analysis to group variables based on patterns of variation or correlation between the observations. Cluster analysis is used to group the products based on the proximity between the variables, using the output from factor analysis and the categorical variables. The BCG growth share matrix will then group the products into the four life cycle stages.

In Sub-chapter 4.4, multiple regression analysis will be used to test a sample of products, the sample selection will be discussed below. The regression analysis will be applied to identify if statistically significant relationships exist between a product and the rest of the products in the market. Price elasticities will be calculated to provide a framework for determining a product's competitor set using regression analysis. Figure 4.1 gives an overview of the process and procedures that will be followed in this chapter.
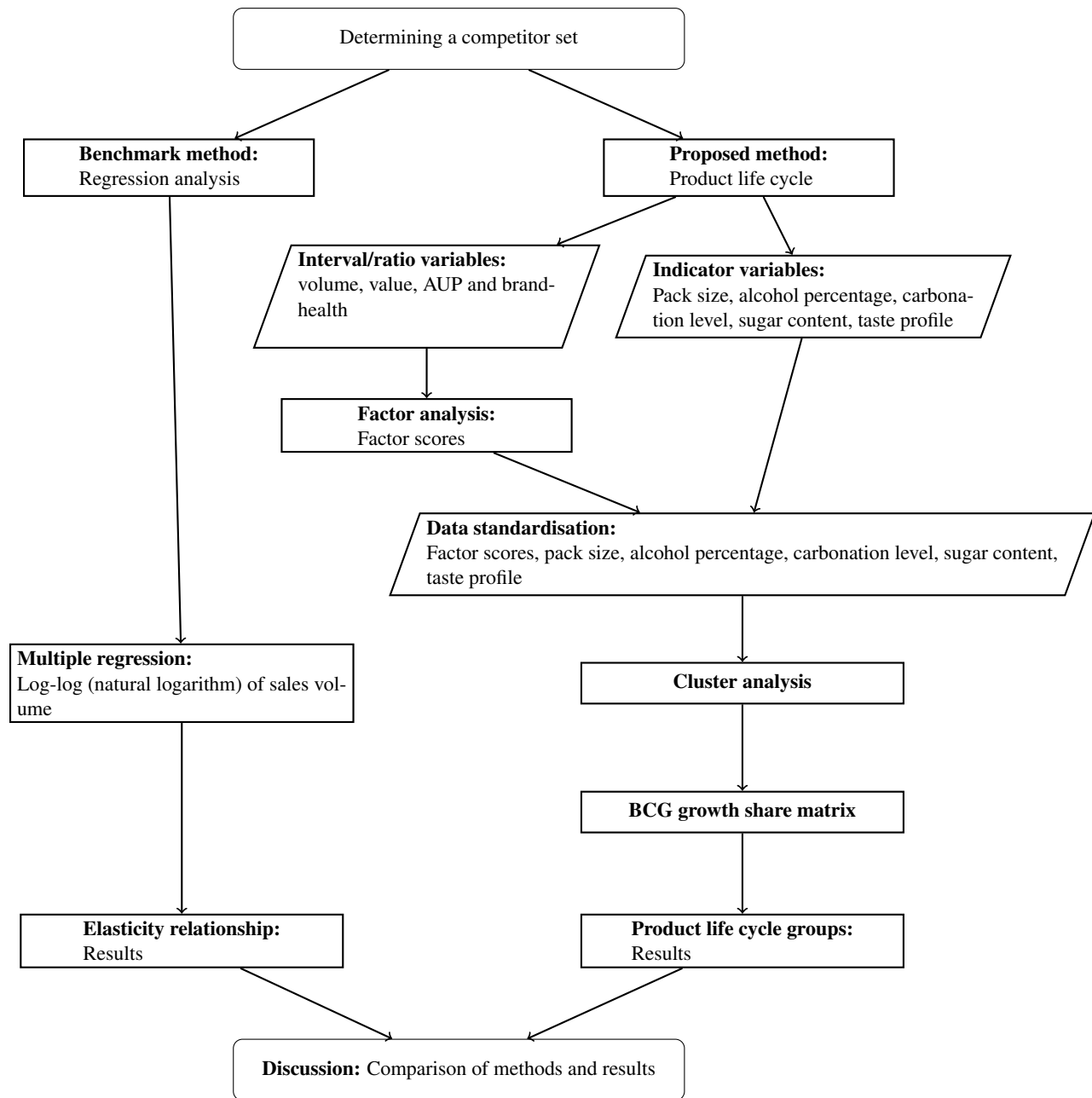
FIGURE 4.1: *Data and methodology procedure for determining a competitor set as described in Chapter 4*

4.1 Sample selection

The dataset has been refined to 594 products and these products will be used in the PLC methodology. This methodology will result in each product being assigned to a life cycle stage. The benchmark method of the log-log regression model will also be used to determine price elasticities to estimate a product's competitor set. The price elasticity methodology needs to be applied to each product in the market separately. The full market will be used as a base for the PLC methodology, but only a sample of products will be chosen to investigate their relationships using the log-log regression model.

When choosing a sample for the log-log regression model the objective is to have a diversified sample.

Four products have been chosen to be investigated to determine their competitor sets. The products are:

- RTD-CIDER-P18-660ML;

- RTD-CIDER-P6-330ML;

- SPIRITS-LIQUEURS-P16-750ML ; and

- WINE-RED-P102-750ML.

Two products have been chosen from the RTD segment due to the size and significance of the segment. RTD-CIDER-P18-660ML is a large pack size cider, while RTD-CIDER-P6-330ML is a medium pack size premium cider. SPIRITS-LIQUEURS-P16-750ML is a cream based spirit liqueur and WINE-RED-P102-750ML a well priced, good value for money red wine. These products have not been chosen for any significant reason other than that they had advertising exposure during the period of investigation. These four products will be used in the further investigation to compute the price elasticity estimates and when comparing the results of the two methods for determining a competitor set. This will make comparison of the different methods possible in the discussion.

Once the methodology for determining a competitor set is complete, the investigation continues by looking at the efficiency of the advertising efforts. The two methodologies will be applied using SPIRITS-LIQUEURS-P16-750ML as a showcase example of the process and results. SPIRITS-LIQUEURS-P16-750ML has been selected because it has a spread of advertising spend over most of the different types of advertising. Using SPIRITS-LIQUEURS-P16-750ML in both methodologies will allow for a direct comparison of the results in the discussion.

4.2 Method for determining a competitor set using the product life cycle methodology

The definition of a competitor is a group of brands or product that meets the same consumer need [39]. This sub-chapter will give the detailed procedures followed to group products into their four life cycle stages; taking into account the consumer, customers in the market and competitors. In order to group the products into competitor sets, factor analysis, cluster analysis and the BCG growth share matrix methodologies will be applied.

Factor analysis is the starting point for the interval and ratio variables volume, value, AUP and brand-health scores. The reason for this is because the descriptive variables volume, value and AUP need to be transformed into uncorrelated factor scores because they are highly correlated.

In Chapter 3.2 the Pearson's correlation coefficients for the different products and the variable volume over the period January 2013 up until December 2017 were tested and found to be highly correlated with each other. Cluster analysis could not be applied to this highly correlated dataset as cluster analysis is highly sensitive to correlated variables [7]. The brand-health scores were shown in Chapter 3.3, to have outliers and they will also benefit from the transformation into factor scores.

4.2.1   Exploratory factor analysis procedure

The variables in this study can be viewed as the consumer's response to the different products and brands available on the market. The sales volume and value are a reflection of how often and how many times a consumer purchased a product. The brand-health measures are a response to how consumers say they feel about the different products. Factor analysis looks for multiple observed variables that have a similar pattern of response, the similarity is due to the association of a latent or unseen variable. For example people may respond similarly to questions about a product's price, value for money offerings and bulk packaging, which are all associated with the latent variable affordability. Table 4.1 summaries the steps.

Procedural Step	Details
Step 1: Variable selection	Numerical variables and categorical variables
Step 2: Data distribution	Testing for a normal distribution
Step 3: Data relationships	Testing for a significant linear relationship and highly correlated variables
Step 4: Sampling adequacy	Kaiser-Meyer-Olkin measure of sampling adequacy
Step 5: Initial extraction	Unweighted least squares and principal axis factoring
Step 6: Retaining factors	Scree Test and Eigenvalues
Step 7: Rotation	Orthogonal and Oblique Rotation
Step 8: Results	Final models factor loading scores
Step 9: Validation	The results and forming meaningful constructs
Step 10: Reliability	Confirming by replicating the results
Step 11: Factor scores	Output to be used in cluster analysis

TABLE 4.1: *Factor analysis procedure decision sequence*

Step 1: Variable selection

Data cleaning is an important step in any analysis, as is the adequacy of the data for the method that will be used. In Chapter 3.1, the data cleaning process was addressed through Pareto analysis. This process left 594 products for evaluation and the variables volume, value, AUP and the brand-health scores.

In Figure 4.2 an example of the dataset that will be used for factor analysis is displayed. The categorical variables can not be used for factor analysis, as they are binary (0/1). One of the six basic assumptions of factor analysis discussed previously is that the observation measurements are interval or ratio, in that the difference between the observations is meaningful [72].

Product	Total 2017 Volume (L)	Total 2017 Value (L)	AUP (R)	Q1	Q2	Q3	Q4	Q5	…	Q78
SPIRITS-WHISKY-P5-1LTR	X.XX L	R X.XX	R X.XX	X %	X %	X %	X %	X %	…	X %
SPIRITS-WHISKY-P41-750ML	X.XX L	R X.XX	R X.XX	X %	X %	X %	X %	X %	…	X %
WINE-RED-P79-1.5LTR	X.XX L	R X.XX	R X.XX	X %	X %	X %	X %	X %	…	X %
WINE-BIB-P42-3LTR	X.XX L	R X.XX	R X.XX	X %	X %	X %	X %	X %	…	X %
WINE-BIB-P79-5LTR	X.XX L	R X.XX	R X.XX	X %	X %	X %	X %	X %	…	X %
WINE-RED-P71-750ML	X.XX L	R X.XX	R X.XX	X %	X %	X %	X %	X %	…	X %
WINE-BIB-P82-1.5LTR	X.XX L	R X.XX	R X.XX	X %	X %	X %	X %	X %	…	X %
WINE-BIB-P43-3LTR	X.XX L	R X.XX	R X.XX	X %	X %	X %	X %	X %	…	X %
WINE-BIB-P82-5LTR	X.XX L	R X.XX	R X.XX	X %	X %	X %	X %	X %	…	X %
WINE-ROSE-P12-750ML	X.XX L	R X.XX	R X.XX	X %	X %	X %	X %	X %	…	X %
WINE-BIB-P41-3LTR	X.XX L	R X.XX	R X.XX	X %	X %	X %	X %	X %	…	X %
WINE-BIB-P78-5LTR	X.XX L	R X.XX	R X.XX	X %	X %	X %	X %	X %	…	X %
WINE-PERLE-P6-300ML	X.XX L	R X.XX	R X.XX	X %	X %	X %	X %	X %	…	X %
WINE-PERLE-P5-300ML	X.XX L	R X.XX	R X.XX	X %	X %	X %	X %	X %	…	X %
WINE-SPARKLING-P22-750ML	X.XX L	R X.XX	R X.XX	X %	X %	X %	X %	X %	…	X %
SPIRITS-VODKA-P4-750ML	X.XX L	R X.XX	R X.XX	X %	X %	X %	X %	X %	…	X %
WINE-RED-P29-750ML	X.XX L	R X.XX	R X.XX	X %	X %	X %	X %	X %	…	X %
WINE-RED-P19-750ML	X.XX L	R X.XX	R X.XX	X %	X %	X %	X %	X %	…	X %
WINE-RED-P22-750ML	X.XX L	R X.XX	R X.XX	X %	X %	X %	X %	X %	…	X %
WINE-WHITE-P22-750ML	X.XX L	R X.XX	R X.XX	X %	X %	X %	X %	X %	…	X %
WINE-WHITE-P36-750ML	X.XX L	R X.XX	R X.XX	X %	X %	X %	X %	X %	…	X %
WINE-WHITE-P33-750ML	X.XX L	R X.XX	R X.XX	X %	X %	X %	X %	X %	…	X %
WINE-WHITE-P28-750ML	X.XX L	R X.XX	R X.XX	X %	X %	X %	X %	X %	…	X %
SPIRITS-LIQUEURS-P2-1LTR	X.XX L	R X.XX	R X.XX	X %	X %	X %	X %	X %	…	X %
SPIRITS-LIQUEURS-P16-750ML	X.XX L	R X.XX	R X.XX	X %	X %	X %	X %	X %	…	X %
RTD-BEER-P14-330ML	X.XX L	R X.XX	R X.XX	X %	X %	X %	X %	X %	…	X %
RTD-BEER-P35-440ML	X.XX L	R X.XX	R X.XX	X %	X %	X %	X %	X %	…	X %
RTD-BEER-P51-660ML	X.XX L	R X.XX	R X.XX	X %	X %	X %	X %	X %	…	X %
RTD-BEER-P13-330ML	X.XX L	R X.XX	R X.XX	X %	X %	X %	X %	X %	…	X %
RTD-BEER-P34-440ML	X.XX L	R X.XX	R X.XX	X %	X %	X %	X %	X %	…	X %
WINE-RED-P48-750ML	X.XX L	R X.XX	R X.XX	X %	X %	X %	X %	X %	…	X %

TABLE 4.2: *A sample of product's electronic till data that has been collected from one of the largest supermarket retailers in South Africa for the calendar year 2017 and the brand-health questions*

Step 2: Data distribution

One of the basic assumptions of EFA is defined by the relationship between the variables, in that they need to have a linear relationship with a normal distribution and a bivariate normal distribution between each pair of observed variables [72]. The 594 products that will be used for evaluation and the descriptive variables volume, value, AUP and the brand-health scores need to be tested to determine if they are normally distributed.

The PROC UNIVARIATE procedure in SAS studio® software is used to test if the variables are normally distributed. The sample size is less than 2000 and therefore the Shapiro-Wilks test should be used, if the sample size is larger then the Kolmogorov-Smirnov test should be used [62]. The hypothesis tested is,

$$H_0 : \text{There is no significant departure from normality}$$

$$H_1 : \text{There is a significant departure from normality.}$$

The null hypothesis (H_0) is not rejected if the p-value of the Shapiro-Wilks test is greater than 0.05. The frequency distribution graph as well as the test statistics for the normal distribution can be seen in Figure 4.2 and Table 4.3 for the variable volume. The variables volume, value, AUP and the brand-health questions were tested and all rejected the null hypothesis (H_0) as the p-values of the Shapiro-Wilks test statistics are less than 0.05. The conclusion is that there is a significant departure from normality.

In Figure 4.2, using the variable volume as an example, the histogram shows that the distribution is right skewed, the results were mostly the same for all the other variables. The lack of the normal distribution will need to be taken into account when choosing the factoring extraction method.
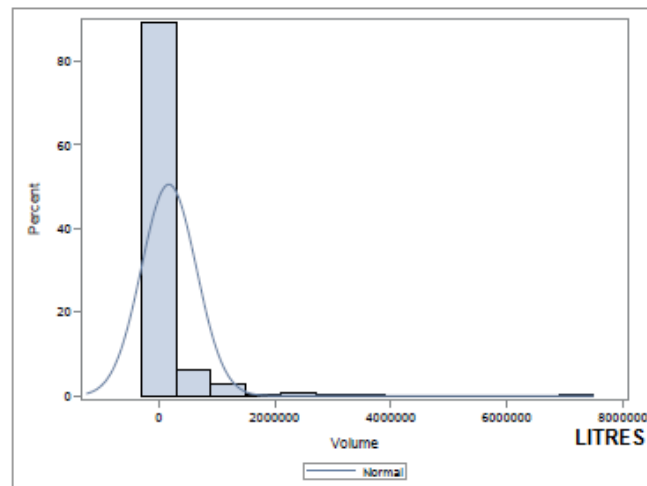
FIGURE 4.2: *Frequency distribution of the variable volume*

Tests for Normality		
Test	Statistic	p-Value
Shapiro-Wilks	$W = 0.063376$	< 0.0001
Kolmogorov-Smirnov	$D = 0.528257$	< 0.0100

TABLE 4.3: *The variable volume's normality test statistics*

Step 3: Data relationships

Correlation analysis will be used to test the strength of the relationship between the variables. Using Pearson's correlation coefficients the correlation can be between -1 and 1. The absolute value of the coefficient is used, any value close to 1 indicates that there is a significant relationship between the variables [64].

The correlation is calculated between all the interval and ratio variables using SAS studio® software and the PROC CORR procedure. According to Cohen [46], a value of 0.99 or higher is deemed to be highly correlated and at least one of the variables should be excluded from the factor analysis [46]. The correlation matrix in Table 4.4 below shows the resulting correlation coefficients for variables volume, value, AUP and market factor questions 1 to 5. In the appendix, Figures A.1 to A.28 display the resulting correlation coefficients and p-values for the full dataset.

The results, in the appendix Tables A.1 to A.28, show that none of the variables are perfectly correlated (1 or -1). The correlation coefficients which are greater than or equal to 0.99 or less than or equal to -0.99 identify variables that are highly correlated. The p-value of these variable also needs to be significant, typically less than 0.05. In the appendix, Figures A.1 to A.28 the results show that the following variables for imagery are highly correlated with each other: Q45, Q46, Q47, Q49, Q50, Q51, Q52, Q53, Q54, Q55, Q56, Q57, Q58, Q59, Q60, Q64, Q68, Q70, Q71, Q73, Q74, Q75 and Q76. The variables power in the mind, Q77 and market share, Q78, are also highly correlated with each other.

Highly correlated variables mean that at least two variables are accounting for the same underlying attribute or relationship. Using correlation analysis gives insight into the relationships between variables but does not assist in determining which variables should remain in the model. Factor analysis is used to refine the dataset in spite of the highly correlated variables while maintaining the underlying pattern in the data and producing uncorrelated factor scores.

	Volume	Value	AUP	Q1	Q2	Q3	Q4	Q5
		Pearson's Correlation Coefficients N = 594						
Volume r	1.00	0.81	-0.06	-0.11	-0.08	-0.01	-0.04	-0.07
Volume p-value		0.00	0.12	0.01	0.06	0.84	0.37	0.08
Value r		1.00	0.14	-0.14	-0.11	-0.04	-0.06	-0.10
Value p-value			0.00	0.00	0.01	0.28	0.12	0.01
AUP r			1.00	-0.01	-0.12	-0.07	-0.08	-0.13
AUP P-value				0.84	0.00	0.11	0.06	0.00
Q1 r				1.00	0.36	0.31	0.28	0.51
Q1 p-value					0.00	0.00	0.00	0.00
Q2 r					1.00	0.49	0.37	-0.07
Q2 p-value						0.00	0.00	0.11
Q3 r						1.00	0.35	-0.09
Q3 p-value							0.00	0.03
Q4 r							1.00	-0.06
Q4 p-value								0.11
Q5 r								1.00
Q5 p-value								

TABLE 4.4: *An extract of the Pearson's correlation coefficients and corresponding p-values for numerical variables*

In factor analysis the variables are required to have a linear relationship with each other [72]. The strength of the linear relationship can also be investigated by plotting the variables against each other using the PROC REG procedure in SAS studio® software. Figure 4.3 shows the results for volume against the variables value, AUP, Q1 and Q2 with regression test statistics. Observing the results for all the individual plots show that most of the variables do not have linear relationships with each other. As with

the results of the scatter plots for outliers, these variables will also benefit from a data transformation to induce a more linear relationship.
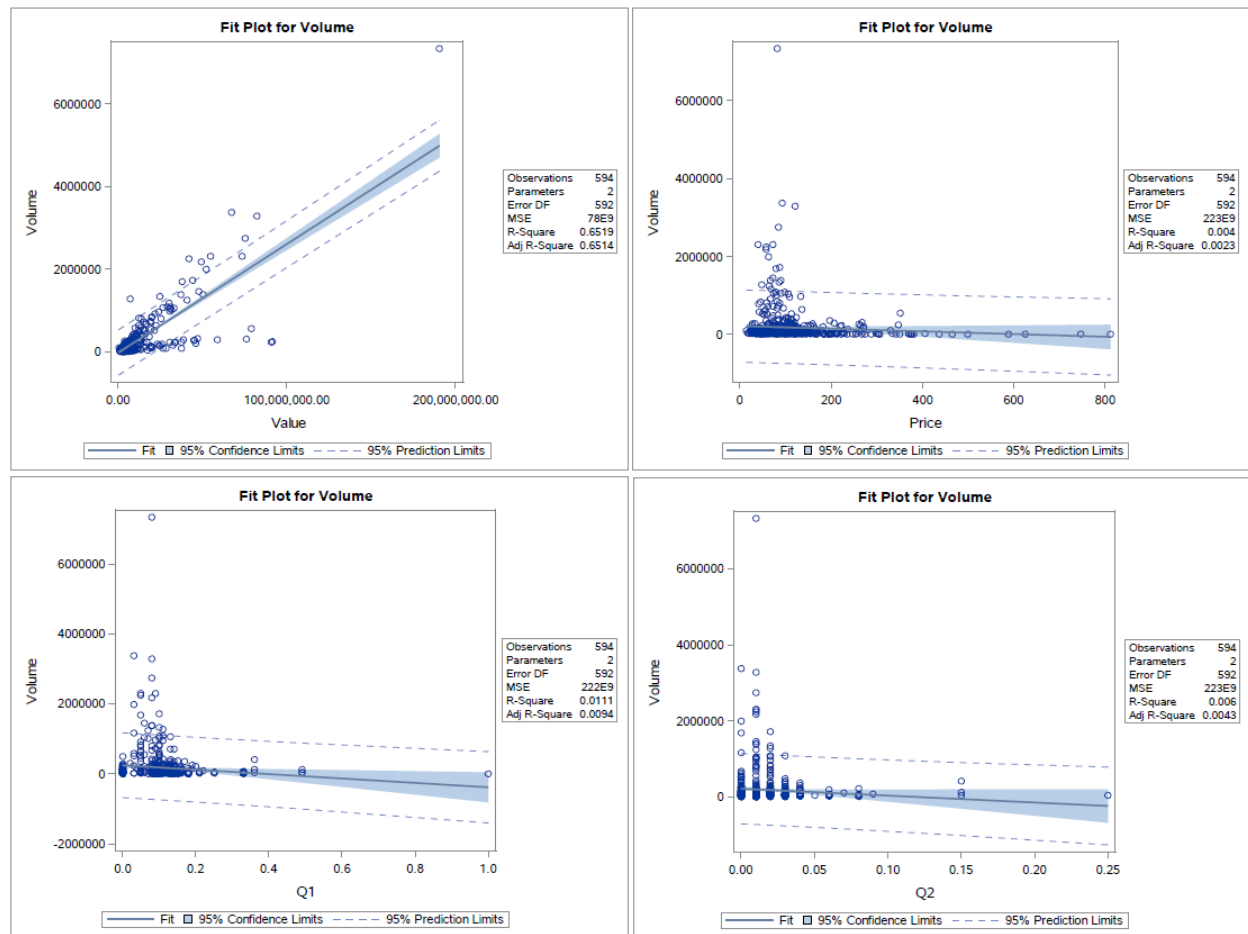
FIGURE 4.3: *The variable volumes plotted against the variables value, AUP, Q1 and Q2 with regression test statistics*

Step 4: Sampling adequacy

When determining which variables to exclude from the factor analysis, all criteria need to be taken into account such as: the variable's distribution, outliers, linearity and correlation. To determine holistically which variables should be used, the Kaiser-Meyer-Olkin measure of sampling adequacy will be used as a measure to test how well the data fits the factor analysis model [62].

The Kaiser-Meyer-Olkin measure of sampling adequacy, equation (2.41), will be used to test the validity of a variable for inclusion in further analysis. The statistic is a measure of the proportion of variance among variables that might be common variance [62], or variables that will successfully load onto factors.

As a general rule the overall Kaiser-Meyer-Olkin measure, equation (2.42), should be greater than 0.8 but scores greater than 0.6 are still acceptable [62]. Scores that are less than 0.5 indicate that there is a problem with the dataset and the method being used. Using the Kaiser-Meyer-Olkin measure of sampling

adequacy gives a summary of how small the partial correlations are relative to the ordinary correlations. A possible solution to an overall Kaiser-Meyer-Olkin score less than 0.5 would be to remove the variables that have individual Kaiser-Meyer-Olkin scores lower than 0.5 [62].

A minimum viable threshold for the Kaiser-Meyer-Olkin scores for individual variables can be set by the researcher. The threshold will determine if the variable will be included or excluded for further analysis. The threshold of a minimum score of 0.55 for an individual variable will be used. A score of less than 0.5 should be excluded but because the factor analysis output will be used in further analysis, a more conservative threshold or a higher threshold of 0.55 has been set as the cut-off. The dataset also has numerous correlated variables, indicating that it will benefit from a variable refinement.

In Table 4.5 the Kaiser-Meyer-Olkin measure scores are displayed for each variable, using equation (2.41). The overall Kaiser-Meyer-Olkin score is greater than 0.8, $\text{KMO} = 0.8451$, using equation (2.42). This means that the data does fit the model well. From Table 4.5, the variables with a Kaiser-Meyer-Olkin score less than 0.55 should be excluded as per the predetermined threshold. The variables scores will be rounded up to the second decimal. The variables with Kaiser-Meyer-Olkin scores less than 0.55 are Q10, Q15, Q20, Q21, Q27, Q32, Q35, Q40, Q42, Q43 and Q44 and they will be excluded.

The variables that are being excluded are all from the brand-health questions relating to market factors. Questions 1 to 9 all relate to different types of availability for example, is the product available in your local bottle store? Questions 10, 11 and 12 relate to the measure of influence the person has over the product being bought. Q10 will be removed from this construct. Questions 13 to 23 relate to the range or choices available within the brand for example, does the brand have enough different flavours available? Questions 15, 20 and 21 will be removed from this construct.

Questions 24 to 29 refer to the level of promotional activity the product has, or how many activations or marketing campaigns engage consumers. From this construct Q27 will be removed. Questions 30 to 33 relate to the level of information that is available about the brand. For example, not enough information is known about the brand to purchase it. From this construct Q32 will be removed. In Q34 to Q36 the questions relate to the adequacy of the price. From this construct Q35 will be removed. Questions 37 to 44 relate to the level of influence the brand has. For example the bar tender recommends this brand. From this construct Q40, Q42, Q43 and Q44 will be removed.

The cut-off for the Kaiser-Meyer-Olkin score removed Questions Q10, Q15, Q20, Q21, Q27, Q32, Q35, Q40, Q42, Q43 and Q44, but the correlation analysis found Questions Q45, Q46, Q47, Q49, Q50, Q51, Q52, Q53, Q54, Q55, Q56, Q57, Q58, Q59, Q60, Q64, Q68, Q70, Q71, Q73, Q74, Q75 and Q76 to be highly correlated with each other and the variables power in the mind, Q77 and market share, Q78, are also highly correlated with each other.

When comparing the results of correlation analysis and factor analysis, the results show that the factor analysis process found that some of the highly correlated variables are contributing towards unique variation in the model.

The factor analysis process can tolerate these highly correlated variables and will 'capture' the pattern in the data while producing an output of factor scores that are uncorrelated and can be used in cluster analysis.

Kaiser-Meyer-Olkin Measure Of Sampling Adequacy Overall Score =0.8451						
Variable	KMO Score	Variable	KMO Score	Variable	KMO Score
Volume 2017	0.8477	Q25	0.6257	Q52	0.9403
Value 2017	0.7733	Q26	0.6596	Q53	0.9321
AUP	0.5663	**Q27**	0.5212	Q54	0.9003
Q1	0.7592	Q28	0.6211	Q55	0.9079
Q2	0.6970	Q29	0.6013	Q56	0.8986
Q3	0.5731	Q30	0.5789	Q57	0.8977
Q4	0.6964	Q31	0.7244	Q58	0.9354
Q5	0.6528	**Q32**	0.3785	Q59	0.9208
Q6	0.6808	Q33	0.5829	Q60	0.9136
Q7	0.6701	Q34	0.5583	Q61	0.8937
Q8	0.7618	**Q35**	0.5290	Q62	0.9221
Q9	0.5872	Q36	0.6910	Q63	0.9392
Q10	0.5314	Q37	0.8500	Q64	0.8748
Q11	0.5985	Q38	0.6481	Q65	0.9035
Q12	0.7091	Q39	0.6926	Q66	0.9226
Q13	0.7943	**Q40**	0.4360	Q67	0.9012
Q14	0.6134	Q41	0.7364	Q68	0.9192
Q15	0.5131	**Q42**	0.0000	Q69	0.8895
Q16	0.5762	**Q43**	0.0000	Q70	0.9449
Q17	0.5536	**Q44**	0.0000	Q71	0.9043
Q18	0.6111	Q45	0.8352	Q72	0.9299
Q19	0.7021	Q46	0.9303	Q73	0.8934
Q20	0.4690	Q47	0.9037	Q74	0.9145
Q21	0.5303	Q48	0.8976	Q75	0.9158
Q22	0.6505	Q49	0.8992	Q76	0.9159
Q23	0.6823	Q50	0.9596	Q77	0.6604
Q24	0.6796	Q51	0.9479	Q78	0.6776

TABLE 4.5: *Kaiser-Meyer-Olkin measure of sampling adequacy scores for the categorical variables*

From testing the different factoring and rotation methods, it was observed that the Kaiser-Meyer-Olkin measure of sampling adequacy is consistent across the different extraction and rotation methods. For the results in Table 4.5 the principal axis factoring method with Varimax rotation was used, this process is described in the steps to follow. After the variables with individual KMO scores less than 0.55 have been removed, the overall score increases to 0.8871. This is a 0.042 improvement in the overall score.

Step 5: Initial extraction

The choice of data extraction method has some high-level guidelines but the researcher will need to test if the method used fits the dataset. Factor analysis is sensitive to outliers and the box and whisker and scatter plots shown in Chapter **??**, on the dataset show that some of the variables for example volume and value, do have significant outliers. Another assumption for factor analysis is that most variables are normally distributed, previously in Chapter 4.2.1, the Shapiro-Wilks tests show that the data is not normally distributed. The distribution of the variables is the best way to determine which extraction method will fit [54].

The maximum likelihood method is best suited for normally distributed variables. If the condition of multivariate normality is being violated then principal axis factoring or unweighted least squares are the best methods to use [54]. In general maximum likelihood is the most commonly used method.

The results from testing six different combinations of factoring methods are shown in Table 4.6, using PROC FACTOR procedure in SAS studio® software. Principal axis factoring and unweighted least squares methods are being tested as extraction methods. The initial estimation procedure for the commonalities and rotation methods will also be varied. Firstly, the squared multiple correlation (SMC) option sets the prior communality of each variable proportional to the R^2 of that variable with all other variables as an initial estimation. Secondly, if the correlation matrix is singular, then using the maximum squared multiple correlation (MAX) sets the prior communality of each variable to the maximum absolute correlation within any other variable [62]. The rotation types will also be varied between the Varimax (V) and Quartimax (Q) rotations.

In order to test the adequacy of the extraction method used, it is important to remember the objective of factor analysis to discover common factors. The factor analysis method attempts to take out as much variations as possible in the first factor. This can be used as a guide for method adequacy, the first factor should account for a high proportion of the variation [72]. Another general rule is that at least three variables need to have significant factor loading scores, scores greater than 0.3 [72]. Another measure for adequacy of fit is that the factor structure should explain most of the correlations, the root mean square off-diagonal residuals (RMSR) gives an indication of this. If the RMSR values are small then this indicates that the factor structure explains most of the variation. The root mean square off-diagonal partials (RMSP), when close to zero, are another indication that the factor structure explains most of the variation.

Method	Overall KMO Score	1st Factor Proportion	Loading > 0.3	RMSR	RMSP
Principle Axis					
SCM/V	0.8871	0.4529	1 Factor < 0.3	0.00185875	0.08096322
SCM/Q	0.8871	0.4529	2 Factor < 0.3	0.00185875	0.08096322
MAX/V	0.8871	0.5283	1 Factor < 0.3	0.01022066	0.15370663
MAX/Q	0.8871	0.5283	1 Factor < 0.3	0.01022066	0.15370663
Unweighted Least Squares					
SCM/V	0.8871	0.4333	1 Factor < 0.3	0.00031141	0.05180497
SCM/Q	0.8871	0.4529	1 Factor < 0.3	0.00031141	0.05180497
MAX/V	0.8871	0.5283	All Factors > 0.3	0.00230902	0.08170410
MAX/Q	0.8871	0.5283	All Factors > 0.3	0.00230902	0.08170410

TABLE 4.6: *Initial extraction results*

The results in Table 4.6 show that the MAX option produces a higher proportion of variation to be accounted for by the first factor, regardless of the extraction method. The overall result for the proportion of variation accounted for by the first factor did not distinguish either extraction method as a better fit. The unweighted least squared factor loading scores greater than 0.3 are an indication of the adequacy of the model when using the MAX option. The results in Table 4.6 show that the unweighted least squares method produces RMSR and RMSP values that are very close to zero. The unweighted least squares

method with the MAX option for the prior commonalities is therefore the more adequate fit as an extraction method.

Step 6: Retaining factors

The number of factors to retain for the final construct is an important decision in the process, too many factors can include unwanted error variance while conversely too few factors might not account for all the common variation. The objective is to determine which factors account for most of the variation in the data. There are three main criteria for determining how many factors to keep.

Firstly, the Kaiser's criterion which is to keep all factors with an eigenvalue greater than one [72]. This approach has been criticised for overestimating the number of factors to retain [83]. The Kaiser's criterion is better used in conjunction with the scree test.

The second criteria is the scree test, the scree plot (Figure 4.4) displays the eigenvalues associated with each factor in descending order. The number of factors to retain are those that are above or to the left of the point of inflection, which is the point where the curve starts to level off [54]. The third criterion is that factors should have at least 3 variables with significant loadings (greater than 0.3). Another measure that can also be used is the proportion of variation that is explained by each factor, a predetermine cut-off can be used.

Method	Kaiser Criterion	Scree Test	Factor > 3 variables loading
Principle Axis			
SCM/V	10 Factors	10 Factors	2 Factors > 3 loadings per factor
SCM/Q	10 Factors	10 Factors	2 Factors > 3 loadings per factor
MAX/V	8 Factors	8 Factors	4 Factors > 3 loadings per factor
MAX/Q	8 Factors	8 Factors	5 Factors > 3 loadings per factor
Unweighted Least Squares			
SCM/V	12 Factors	10 Factors	2 Factors > 3 loadings per factor
SCM/Q	10 Factors	10 Factors	2 Factors > 3 loadings per factor
MAX/V	8 Factors	8 Factors	3 Factors > 3 loadings per factor
MAX/Q	8 Factors	8 Factors	3 Factors > 3 loadings per factor

TABLE 4.7: *Results for the number of factors to retain when using Principal Axis and Unweighted Least Squares extraction methods and varying the prior commonalities and rotation methods*

In Table 4.7, the Kaiser criterion consistently has an equal or higher factor extraction than the scree test. The results for the scree tests are between 8 and 10 factors for all combinations of methods. Using both results, the minimum number of factors to retain is 8 and the maximum is 12. In Table 4.7, the variables with significant loadings (≥ 0.3) are the highest for the principal axis method with the MAX option. In the previous step the unweighted least squares method with the MAX option was determined to be the most adequate extraction method. The results for the unweighted least squares method with the MAX option and Quartimax rotation is displayed in Figure 4.4. The eigenvalues drop below 1 after the 8th factor and therefore 8 factors should be extracted.
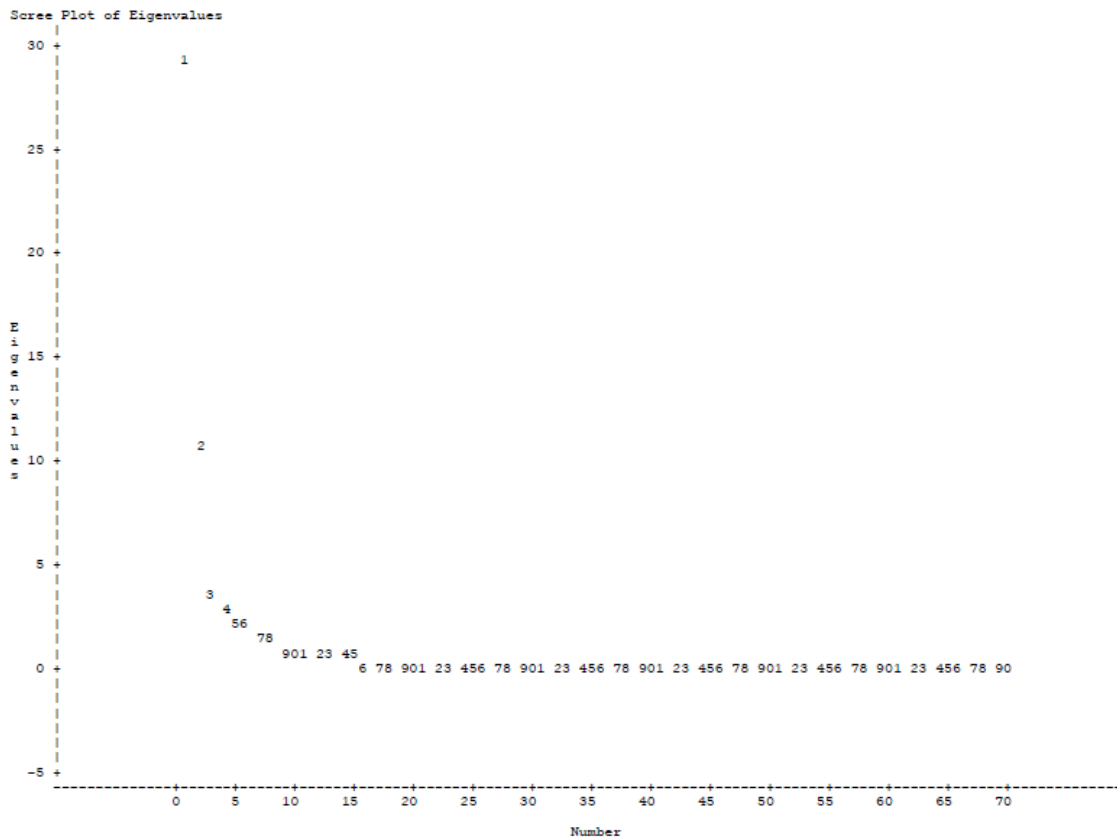
	Preliminary Eigenvalues: Total = 55.5133633 Average = 0.79304805			
Factor	**Eigenvalue**	**Difference**	**Proportion**	**Cumulative**
1	29.3279522	18.8731237	0.5283	0.5283
2	10.4548286	6.5552964	0.1883	0.7166
3	3.8995321	1.2579481	0.0702	0.7869
4	2.6415841	0.1987803	0.0476	0.8345
5	2.4428038	0.4140777	0.0440	0.8785
6	2.0287261	0.5629329	0.0365	0.9150
7	1.4657932	0.1697274	0.0264	0.9414
8	1.2960658	0.3552256	0.0233	0.9648
9	0.9408402	0.1275207	0.0169	0.9817
10	0.8133195	0.0733141	0.0147	0.9964
11	0.7400054	0.0370561	0.0133	1.0097
12	0.7029493	0.0771623	0.0127	1.0224
13	0.6257870	0.1429715	0.0113	1.0336
14	0.4828155	0.0657368	0.0087	1.0423
15	0.4170787	0.0792353	0.0075	1.0498
16	0.3378434	0.0205175	0.0061	1.0559
17	0.3173259	0.0783778	0.0057	1.0616
18	0.2389481	0.0704874	0.0043	1.0659

FIGURE 4.4: *Results from factor analysis using the unweighted least squares method with the MAX option and Quartimax rotation*

Step 7: Rotation

The initial or direct factor loadings can be ambiguous and difficult to interpret. Factor rotation is used to improve the interpretation and attain a simple structure where each variable loads onto as few factors as possible. The purpose of using factor analysis before cluster analysis is to create uncorrelated scores. Therefore the orthogonal rotation method will be used, Varimax and Quartimax methods will be tested to determine the best fit structure.

From the initial extraction results, the unweighted least squared method with the MAX option showed to be the best fit. The investigation for retaining factors resulted in 8 factors being optimal. From the results in Tables 4.6 and 4.7, for the unweighted least squares method with the MAX option, the Quartimax and Varimax methods of rotation produced the same result. The Quartimax rotation will be used as this method seeks to minimise the number of factors needed to explain each variable [83].

Step 8: Results

The final criterion is that the factor structure needs to form a conceptually meaningful construct [72]. After the factors have been extracted they need to be interpreted. This means assigning to each common factor or construct a name or description. This name needs to reflect the relationship between the variables. Factor interpretation is a subjective process. The researcher will need to rely on the understanding of the underlying subject matter in order to decide if the constructs formed are viable and explain the variables.

From the process described above and equation (2.33), the unweighted least squares method with the MAX option and Quartimax rotation to extract 8 factors will be used, the code for SAS studio® software is displayed in Figure A.1 in the appendix. The resulting factor pattern is displayed in Table 4.8 and 4.9. The variables are not in alphabetical or numerical order, they are in order of the weight of their factor loadings, equation (2.38). A variable will be considered 'loaded' onto a factor if the factor loading is greater than 0.3 [72].

Factor 1 starts with $Q68$ having a high loading, 0.97952 which is close to 1. The loadings decline down to $Q45$, for the first factor.

			Quartimax Rotated Factor Patter					
Variable	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8
Q68	**0.97952**	0.01732	0.00036	-0.01294	-0.01997	0.06418	0.04861	0.04951
Q64	**0.97805**	0.01857	-0.01337	-0.03386	-0.02888	0.09734	-0.00196	0.06465
Q49	**0.97546**	-0.01444	-0.05203	-0.00884	-0.02937	-0.00686	-0.02627	-0.04596
Q51	**0.97189**	-0.0162	0.02919	-0.03509	0.11343	0.03903	0.00994	-0.03163
Q66	**0.97047**	-0.00306	-0.02811	0.01634	0.01728	-0.00751	0.03028	0.00434
Q52	**0.97036**	-0.02368	-0.04057	0.04451	-0.01557	-0.08852	0.07032	-0.06474
Q73	**0.96935**	0.02898	-0.0138	-0.02862	-0.06372	0.08756	-0.00384	0.09971
Q56	**0.9691**	-0.02258	0.00985	-0.01868	0.12139	-0.00974	0.03326	-0.03278
Q70	**0.96893**	0.00888	-0.0477	0.00537	-0.05662	0.00499	-0.06535	0.02708
Q58	**0.96861**	-0.0092	-0.07087	0.05249	0.0476	-0.05495	0.00701	-0.09386
Q60	**0.96829**	-0.0056	0.00251	-0.04541	-0.01286	0.10786	-0.05533	0.00576
Q63	**0.96674**	0.00350	-0.01812	-0.0608	0.06582	0.13971	-0.05381	0.01909
Q71	**0.96669**	0.00207	-0.0118	-0.02202	-0.04202	0.05225	-0.02041	0.14301
Q72	**0.96456**	0.01639	0.02271	-0.03759	-0.08874	0.05048	0.03176	0.00829
Q53	**0.96394**	0.0017	0.1013	-0.02267	0.05032	0.07581	-0.05463	-0.00001
Q55	**0.96349**	-0.01755	0.09709	0.00905	-0.01863	-0.01639	0.00072	-0.06632
Q48	**0.96348**	0.03599	-0.03144	-0.01542	-0.11334	0.11677	-0.06863	0.06477
Q50	**0.96298**	0.00165	0.11451	-0.00294	0.01409	0.04118	-0.01387	-0.00187
Q54	**0.95675**	0.02612	0.02038	-0.02379	-0.13139	0.02239	0.01695	0.08848
Q47	**0.95672**	-0.02244	0.00536	-0.06745	0.08535	0.06468	-0.00439	-0.06215
Q59	**0.95183**	-0.0273	-0.05411	0.04733	0.12686	-0.09177	-0.00868	-0.08998
Q62	**0.95**	-0.0117	-0.01185	-0.04679	0.17785	0.10489	0.03244	-0.02717
Q67	**0.94403**	-0.11251	-0.02624	-0.02706	0.08046	-0.11164	-0.0244	-0.04305
Q74	**0.94114**	-0.03988	0.03812	0.03228	0.17405	-0.06153	0.00055	-0.07432
Q46	**0.92956**	0.04500	0.01348	0.01027	-0.19471	0.04922	0.03271	0.13026
Q61	**0.90414**	0.03738	0.06205	-0.04015	-0.25915	0.00879	0.06232	0.1002
Q76	**0.90398**	-0.11176	-0.08246	-0.00742	0.1564	-0.13328	-0.03591	-0.08211
Q75	**0.8987**	-0.04628	-0.06195	0.01367	0.16461	-0.16449	0.07228	-0.14616
Q69	**0.87668**	-0.0601	-0.06632	0.0023	0.27572	-0.05768	0.01415	-0.11143
Q57	**0.87464**	-0.08378	-0.02479	0.04442	0.2127	-0.15262	-0.01063	-0.12048
Q65	**0.86651**	-0.01986	-0.11219	0.06818	-0.13873	-0.11073	-0.03754	-0.00771
Q45	**0.81398**	0.01784	-0.06557	0.09423	-0.0637	-0.10459	0.02704	0.06108
Q25	-0.02285	**0.90133**	0.02118	0.17251	-0.00436	-0.02117	-0.19415	-0.01474
Q38	-0.04855	**0.89396**	0.03733	0.0685	-0.07384	-0.01466	-0.1265	-0.10817
Q5	-0.05789	**0.89161**	-0.00637	-0.1661	-0.00113	0.0800	0.05213	-0.00148
Q37	-0.0734	**0.80658**	0.28787	0.03272	-0.02899	0.08461	0.24008	0.17245
Q1	-0.09147	**0.74767**	0.25193	0.03904	-0.17976	0.37359	0.25069	0.06396
Q13	0.04006	**0.7456**	0.24701	0.07053	-0.16751	0.15446	0.38815	0.06838
Q24	-0.08011	**0.74088**	0.39166	0.1927	-0.10029	0.04003	-0.08616	0.28835
Q22	0.11064	**0.72893**	0.01789	0.15991	-0.00603	0.12141	0.37786	-0.05021
Q34	-0.09467	**0.57721**	0.24332	0.14045	-0.12931	-0.40707	-0.03895	-0.19723
Q19	-0.05537	0.17112	**0.76536**	-0.00976	-0.10641	-0.02196	0.17773	0.21585
Q41	0.0017	0.11791	**0.70048**	0.16416	-0.18946	-0.10946	-0.07037	0.25172
Q11	-0.0227	0.103	**0.6572**	0.21393	-0.04038	0.15007	-0.05623	-0.08667
Q9	-0.01026	0.07758	**0.64562**	0.1097	-0.13714	-0.1681	-0.01571	-0.41188
Q8	-0.14618	0.14453	**0.6208**	0.27076	-0.05675	0.37999	-0.00026	-0.21806
Q28	-0.19278	0.0752	**0.55361**	0.03623	-0.04903	0.12499	0.05199	0.20615
Q16	0.05263	0.02856	**0.54013**	0.02545	-0.04979	-0.06543	0.29657	-0.04255
Q36	0.04577	0.13263	**0.42738**	0.17009	-0.07848	0.29821	0.0643	-0.05286

TABLE 4.8: *Part 1: Quartimax rotated factor patter using the unweighted least squares with MAX option and 8 factors*

			Quartimax Rotated Factor Patter					
Variable	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8
Q39	-0.3426	0.0795	**0.39257**	0.17996	-0.08532	0.27392	0.02362	0.3026
Q30	0.0629	0.03042	0.21463	**0.8145**	-0.10139	0.00983	-0.1148	0.05575
Q31	-0.15725	0.08525	0.14828	**0.77551**	-0.02975	0.45067	-0.07972	0.06914
Q33	0.01227	0.11977	0.29769	**0.75183**	-0.0789	-0.02761	0.01991	-0.04327
Q18	-0.10051	0.09026	0.13893	**0.65527**	0.00137	0.07878	0.31086	-0.03231
Q26	-0.09298	0.13017	0.43784	**0.4764**	-0.09669	0.13982	0.06829	0.09001
Q78	0.20986	-0.12503	-0.21224	-0.13455	**0.83838**	0.01271	-0.08474	-0.02575
Q77	0.20469	-0.13086	-0.21741	-0.13902	**0.83183**	0.01085	-0.09014	-0.03388
Volume 2017	0.12988	-0.06569	-0.11359	-0.03876	**0.65126**	0.01917	-0.03224	0.01373
Value 2017	0.02407	-0.07957	-0.15077	0.0027	**0.5829**	-0.14117	0.01402	-0.06352
Q3	-0.08437	0.07224	0.06669	0.22862	0.03518	**0.59646**	0.4194	-0.14343
Q2	0.00087	0.19191	0.27949	0.28887	-0.16744	**0.59194**	0.20511	0.02475
Q23	-0.06275	0.15496	0.40449	0.1602	-0.16999	**0.53601**	-0.1018	-0.00347
Q7	0.08717	0.21288	0.13648	0.40068	-0.21641	**0.45239**	0.21414	0.03876
Q4	-0.15351	0.13702	-0.08092	0.1639	-0.29011	**0.43597**	-0.06308	0.34696
AUP	-0.05577	-0.01064	-0.03748	0.1158	-0.13867	**-0.39376**	0.06662	-0.14623
Q6	-0.06819	0.31114	0.26262	-0.11505	-0.01458	0.13285	**0.56064**	0.15018
Q17	0.00909	0.17782	0.14134	0.154	-0.15143	0.08615	**0.49995**	0.13567
Q12	0.12585	0.11438	0.27707	0.35679	-0.04243	-0.08433	**0.43536**	-0.1009
Q14	-0.23455	0.08009	0.20787	-0.03517	-0.16886	0.20042	0.1011	**0.66714**
Q29	-0.13681	0.08953	0.32223	0.40038	0.02301	-0.00164	0.28022	**0.50564**

TABLE 4.9: *Part 2: Quartimax rotated factor patter using the unweighted least squares with MAX option and 8 factors*

Step 9: Validation

Using Hatcher's [53] interpretability criteria the results from factor analysis can be validated. Firstly, examining the number of factors with loadings greater than 0.40 [67]. Table 4.10 displays the number of variables that load onto each factor and the number of variables with loadings greater than 0.40 . Only Factor 8 has less than three variables loaded, this factor will need to be carefully examined using the other validation criteria. There were only two variables with loadings that are below the cut-off value of 0.40, this is a strong indication that the results are valid and should be trusted.

	Validation of significant loadings							
Description	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8
Number of variables loaded in results	32	9	9	5	4	6	3	2
Number of variables loaded greater than 0.4	32	9	8	5	4	5	3	2

TABLE 4.10: *Validation of factor analysis results testing for significant loadings*

The second and third criteria for interpretability according to Hatcher [53] is the interpretation of the results, do the variables that have loaded onto a factor represent the same construct and are the constructs different from each other? Table 4.11 displays the factors and constructs formed based on the variables

that have loaded onto each factor.

Table 4.11 shows that the first factor represents positive feelings about the brand's imagery. Brand imagery is developed over time through advertising campaigns with a consistent theme and is authenticated through the consumers' direct experiences. The brand imagery signifies what the brand currently stands for. Factors 2, 3 and 4 are the inverse of Factor 1, representing the negative associations with the brands, representing the market factors and imagery as a measure of negative barriers for a consumer to purchase a brand. Each factor represents a different construct of negative association, Factor 2 represents consumers' feeling that the brands are not accessible to them due to distribution or price. Factor 3 represents the consumers' feelings towards the lack of packaging formats and taste profiles, while Factor 4 represents the consumers' feeling that the brand does not communicate with them.

Factors 5 and 6 represent the 'big brand' feeling. These are brands with large market share in both volume and value. These brands have strong levels of power in the mind of the consumers. This is a measure that takes into account consumers' desire or sentiment towards the brand, as well as the AUP. Factor 7 represents the ability of a consumer to find the brand in places that they frequent. Finally Factor 8 represents the construct of un-stylishness or not trendy, this factor represents the negative side for brands that are not in fashion. These brands are competing with other brands where being innovative is the most important variable for the consumer when choosing what to purchase.

Factor 8 was highlighted in the first of Hatcher's [53] interpretability criteria as being below the cut-off, however, when examining the factor against the second and third criteria it shows that the factor does form a meaningful construct. All the factors form meaningful constructs that are internally alike in the variables that have been grouped and externally dissimilar, this is another strong indication that the results are valid and can be trusted.

The fourth and final criteria for interpretability according to Hatcher [53] is that the rotated factor pattern represents simple constructs. In Tables 4.8 and 4.9 in the previous section the rotated factor pattern is displayed, this shows that variables Q3, Q7, Q23, Q26, Q29 and Q31 have greater than 0.4 loadings onto two factors. Variables Q39 and AUP do not have a high loading onto any factor. All of these variables were very close to the cut-off of 0.4 and based on the other validation criteria results they will not be excluded.

Overall the factors do follow the pattern of having at least one variable with a relatively high loading to start with and the remaining loadings tapering off. In total 88.6% of the variables have a loading of greater than 0.4 onto a unique factor. This result coupled with the previous validation steps show that the results are valid and can be trusted for further analysis.

Step 10: Reliability

The results can also be tested for reliability in terms of consistency, this method involves splitting the dataset to determine if the variables still load in the same pattern [29]. In order to create a comparison dataset the original dataset is randomly split into two groups. The random split is generated using the RAND() function in Microsoft Office 2013 Excel. Using the model set out in Figure A.1 in the appendix, factor analysis is repeated for the two sub-sets.

The results of the final factor analysis and the factor analysis on the two sub-sets can now be compared for consistency and reliability [29]. The comparison will be based on the variables and the factors that they have loaded onto. In Figure 4.5 the number of variables that have loaded onto each factor in the final result and in the two sub-sets is displayed, as well as the similarity of the loadings. The similarity measures if the same variable loads onto the same factor in the final results set as well as the validation sub-sets. The results show that similar loading patterns were observed in all three models, with a overall rate of 77.14% of similarity. This showing that the results are consistent and therefore reliable.

Factor	Construct	Structure
Factor 1	Enjoyment	Imagery and market factors: brand love or positive feelings towards the brands
		Q45, Q46, Q47, Q48, Q49, Q50, Q51, Q52, Q53, Q54, Q55, Q56, Q57, Q58, Q59, Q60, Q61, Q62, Q63, Q64, Q65, Q66, Q67, Q68, Q69, Q70, Q71, Q72, Q73, Q74, Q75 and Q76
Factor 2	Unreachable	Market factors: brand disfavour or negative feelings towards the brands due to availability and price accessibility
		Q1, Q5, Q13, Q22, Q24, Q25, Q34, Q37 and Q38
Factor 3	Dislike	Market factors: negative feelings towards the brand due to its packaging formates and taste
		Q8, Q9, Q11, Q16, Q19, Q28, Q36, Q39 and Q41
Factor 4	Isolation	Market factors: negative feelings towards the brand due to lack of information or promotions
		Q18, Q26, Q30, Q31 and Q33
Factor 5	Popular	Market share: big brands
		volume 2017, value 2017, Q77, and Q78
Factor 6	Accessibility	Market factors: accessible price
		AUP, Q2, Q3, Q4, Q7 and Q23
Factor 7	Availability	Availability
		Q6, Q12 and Q17
Factor 8	Un-stylish	Market factors: negative feelings towards brand positioning, packaging or relevance for to-day's trends
		Q14 and Q29

TABLE 4.11: *Resulting 8 factor structures from the unweighted least squares method with MAX option and Quartimax rotation*
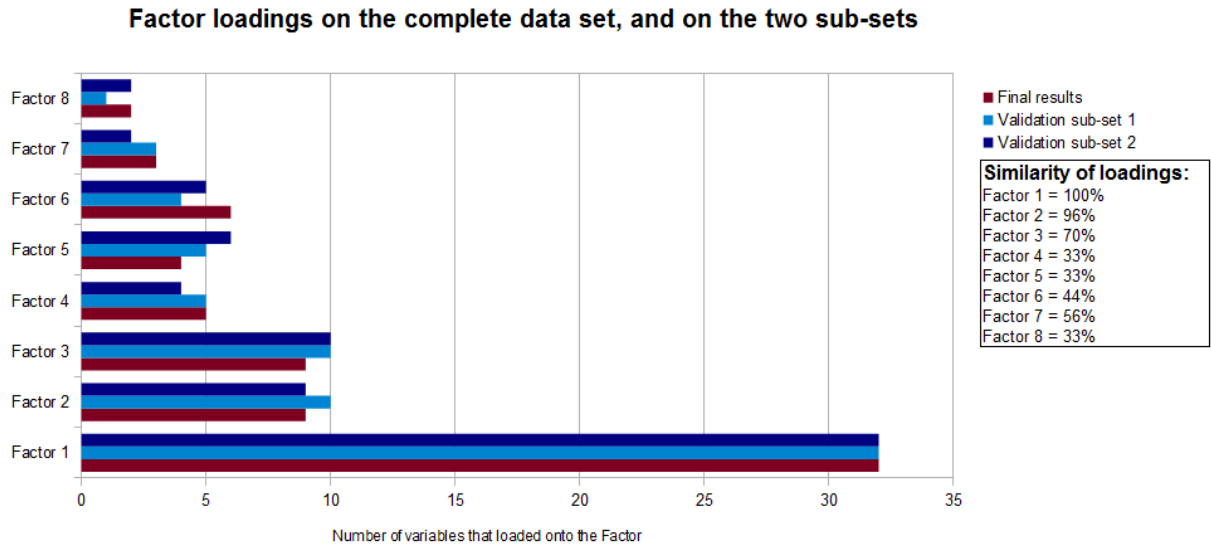
FIGURE 4.5: *Comparison of the factor loadings in the final results and the validation sub-sets*

Step 11: Factor scores

The OUT command in the PROC FACTOR procedure in SAS studio® software produces the factor scores, using equation (2.40). A factor score is a numerical value that informs a variable's relative spacing or standing on a latent factor, only the factors being retained are used. In Figure A.1 in the appendix, $n = 8$ is specified to inform the number of factors that should be retained. Each estimated factor score is computed as a linear combination of the standardised values of the variables that are retained factors [62], using equation (2.40). The factor scores from the factor analysis will be used to represent the variables volume, value, AUP and the brand-health variables when performing cluster analysis.

4.2.2   Cluster analysis procedure

This section describes the procedure used for cluster analysis, including Ward's minimum variance method and then the k-means method. The products will be grouped together into clusters based on their similar characteristics, also taking volume, value, AUP and the brand-health variables into account in the form of factor scores as a result of the factor analysis. The clusters will form sets that are representative of the consumer's repertoire sets.

Procedural Step	Details
Step 1: Sampling adequacy	Data cleaning and merging
Step 2: Data standardisation	Variable standardisation
Step 3: Starting point for determining k	Ward's minimum variance method
Step 4: Determining k	k-Means method
Step 5: Interpreting results	Interpret constructs for competitor sets
Step 6: Validation	Validate constructs for competitor sets

TABLE 4.12: *Decision sequence for the cluster analysis procedure*

Step 1: Sampling adequacy

The data cleaning process has already been addressed in the previous section on factor analysis. The variables volume, value, AUP and the brand-health score, were transformed into a factor scores for each product. After this process 594 products remain for evaluation with variables describing pack sizes, alcohol percentage, carbonation level, sugar content, taste profile and the eight factor scores. A sample of the data set available is shown in Tables 4.13 and 4.14, all these variables can be used for cluster analysis.

Product Pack Size	Small Pack Wine	Normal Pack Wine	Large Pack Wine	Small Pack Spirits	Normal Pack Spirits	Large Pack Spirits	Small Pack RTDs	Normal Pack RTDs	Large Pack RTDs
SPIRITS-WHISKY-P5-1LTR	0	0	0	0	0	1	0	0	0
SPIRITS-WHISKY-P41-750ML	0	0	0	0	1	0	0	0	0
WINE-BIB-P79-5LTR	0	0	1	0	0	0	0	0	0
WINE-BIB-P42-3LTR	0	0	1	0	0	0	0	0	0
WINE-BIB-P79-5LTR	0	0	1	0	0	0	0	0	0
WINE-RED-P71-750ML	0	1	0	0	0	0	0	0	0
WINE-BIB-P82-5LTR	0	0	1	0	0	0	0	0	0
WINE-BIB-P43-3LTR	0	0	1	0	0	0	0	0	0
WINE-BIB-P82-5LTR	0	0	1	0	0	0	0	0	0
WINE-ROSE-P12-750ML	0	1	0	0	0	0	0	0	0

Product Pack Size	Low Alcohol	Mid Alcohol	High Alcohol	Still	Carbonated	Sparkling	Dry	Semi-Sweet	Sweet	Bitter
SPIRITS-WHISKY-P5-1LTR	0	0	1	1	0	0	1	0	0	1
SPIRITS-WHISKY-P41-750ML	0	0	1	1	0	0	1	0	0	1
WINE-BIB-P79-5LTR	0	1	0	1	0	0	0	0	1	0
WINE-BIB-P42-3LTR	0	1	0	1	0	0	0	0	1	0
WINE-BIB-P79-5LTR	0	1	0	1	0	0	0	0	1	0
WINE-RED-P71-750ML	0	1	0	1	0	0	0	0	1	0
WINE-BIB-P82-5LTR	0	1	0	1	0	0	0	0	1	0
WINE-BIB-P43-3LTR	0	1	0	1	0	0	0	0	1	0
WINE-BIB-P82-5LTR	0	1	0	1	0	0	0	0	1	0
WINE-ROSE-P12-750ML	0	1	0	1	0	0	0	0	1	0

Product Pack Size	Wine Like	Beer Like	Cider Like	Fruity	Herbaceous	Woody/Yeasty	Sweet Associated	Smoky Peaty	Smooth	Mouth coating
SPIRITS-WHISKY-P5-1LTR	0	0	0	1	0	1	1	1	1	0
SPIRITS-WHISKY-P41-750ML	0	0	0	1	0	1	1	1	1	0
WINE-BIB-P79-5LTR	1	0	0	1	0	0	0	0	0	0
WINE-BIB-P42-3LTR	1	0	0	1	0	0	0	0	0	0
WINE-BIB-P79-5LTR	1	0	0	1	0	0	0	0	0	0
WINE-RED-P71-750ML	1	0	0	1	0	0	0	0	0	0
WINE-BIB-P82-5LTR	1	0	0	1	0	0	0	0	0	0
WINE-BIB-P43-3LTR	1	0	0	1	0	0	0	0	0	0
WINE-BIB-P82-5LTR	1	0	0	1	0	0	0	0	0	0
WINE-ROSE-P12-750ML	1	0	0	1	0	0	0	0	0	0

TABLE 4.13: *Part 1: A sample set of data showing the variables that will be used in the cluster analysis procedure*

Product Pack Size	Factor 1 Score	Factor 2 Score	Factor 3 Score	Factor 4 Score	Factor 5 Score	Factor 6 Score	Factor 7 Score	Factor 8 Score
SPIRITS-WHISKY-P5-1LTR	0.2261	-0.0061	-0.0112	0.834	-0.4086	-1.8113	0.087	-0.8073
SPIRITS-WHISKY-P41-750ML	0.2264	0.0012	-0.0014	0.8366	-0.3730	-1.8101	0.0885	-0.8225
WINE-BIB-P79-5LTR	-0.3609	-0.4555	-0.9712	-0.6986	0.5042	-0.0369	-1.2249	-0.1848
WINE-BIB-P42-3LTR	-0.3603	-0.4446	-0.9769	-0.6621	0.6362	-0.0841	-1.1794	-0.1712
WINE-BIB-P79-5LTR	-0.3600	-0.4449	-1.001	-0.6178	0.8094	-0.1432	-1.1343	-0.0850
WINE-RED-P71-750ML	-0.3608	-0.4531	-0.9686	-0.6966	0.5214	-0.0379	-1.2235	-0.1866
WINE-BIB-P82-5LTR	-0.3607	-0.4512	-0.9696	-0.6907	0.5378	-0.0461	-1.2150	-0.1879
WINE-BIB-P43-3LTR	-0.3598	-0.4359	-0.9744	-0.644	0.7172	-0.1049	-1.1582	-0.1713
WINE-BIB-P82-5LTR	-0.3581	-0.4068	-0.9844	-0.5493	1.1213	-0.2203	-1.0521	-0.1072
WINE-ROSE-P12-750ML	-0.3607	-0.4511	-0.9678	-0.6927	0.5388	-0.0423	-1.2189	-0.1875

TABLE 4.14: *Part 2: A sample set of data showing the variables that will be used in the cluster analysis procedure*

Step 2: Data standardisation

The PROC CLUSTER and PROC FASTCLUS procedures in SAS studio® software will be used in the next steps to execute cluster analysis using Ward's minimum distance method and the k-means method of clustering. The different variables that are to be used, are not all homogeneous across their units of measure. The variables describing the brand attributes are categorical, containing binary variables (0/1). The variables sales volumes, values and AUP as well as brand-health scores are interval, the latter set of variables has been transformed into eight factor scores, these uncorrelated values will be used for cluster analysis together with the brand attributes. These variables needs to be standardised and the data transformation will also help to reduce the effect of outliers.

The PROC ACECLUS method will be used in SAS studio® software to standardise the post-factor-analysis dataset and the categorical variables, alcohol percentage, carbonation level, sugar content, pack size and taste profile as described in Table 3.5. The procedure is used to standardise the variables to have a mean of zero and a variance of one [62], and the dataset is transformed such that the resulting within-cluster covariance matrix is spherical. In Figure A.2 in the appendix the code used in SAS studio® software to compute the data transformation is displayed.

Step 3: Starting point for determining the number of clusters: Ward's minimum variance method

Hierarchical clustering methods break down the data into sets or groups of products and then refines each partial decomposition to create the next set. Hierarchical clustering methods create non-empty subsets of the original dataset of products [79]. There are many different hierarchical clustering methods, each method has a different definition of similarity. For example in Ward's minimum distance method, equation (2.27), the distance between two clusters is defined as the analysis of variance sum of squares between two clusters summed over all the variables [62].

The objective in cluster analysis, as opposed to factor analysis, is to group 'similar' products together

when the underlying structure is unknown. This is carried out by using Ward's minimum distance method and the PROC CLUSTER procedure in SAS studio® software, as displayed in Figure A.3 in the appendix. Ward's minimum distance method is used to get an initial indication of the best possible value for k, the number of clusters. Using Ward's method the observations with the smallest distances between them are grouped together. Then the observations with the next smallest distances are added to each group, this process continues until all observations end up together in one large group. The cluster procedure is interpreted by observing the grouping history or pattern produced as the clusters were formed. In Figure A.3 in the appendix, the variables *Can1* to *Can37* are the standardised variables describing pack sizes, alcohol percentage, carbonation level, sugar content, taste profile and the factor scores.

In Figure 4.6 and Table 4.15, the results of the Ward's minimum distance method of clustering are displayed. The CCC values greater than 3 are an indication of good cluster formations [63], meaning that the items in a cluster are akin to each other and not associated to items in other clusters.

In Table 4.15, the results show that after the first cluster is formed, all subsequent clusters have a CCC value greater than 3 up until cluster 119. As the CCC value for the formation of cluster 119 is zero, this indicates that 119 is an upper limit for the number of clusters that should be formed.

When interpreting the CCC graphs, the peaks are a strong indication of a potentially good cluster formation or number of clusters that should be formed. In Figure 4.6 and using Table 4.15 as a reference, peaks or significant increases in CCC values are visible at clusters 3, 5, 8 and 25. The CCC does not have any large negative values, this indicates that there are no significant outliers [63]. The CCC test statistic needs to be interpreted in conjunction with the other test statistics.

When evaluating the pseudo-F statistics, the largest value should be used as this indicates well separated clusters [63]. In Figure 4.6, the pseudo-F statistic has only one peak and this is at 8 clusters.

The pseudo-t^2 statistics graph should be read from right to left, stopping at the first large peak [63]. The value after this peak should be used as a potential cluster. In Figure 4.6, there are small peaks at clusters 7, 48, 105, 111, 112 and 121. Reading from right to left, the first evident peak is at cluster 121. This showing that 122 clusters would be optimal according to the pseudo-t^2 statistics graph.

In Table 4.15, other test statistics, semi-partial R^2, R^2, approximate expected R^2 and ties are displayed. The semi-partial R^2 is a measure of uniformity of the merged clusters that have formed [62]. The semi-partial R^2 decreases as each new cluster is formed and at cluster 33 the semi-partial R^2 value drops to zero. This indicates that less than 33 clusters should be formed. The upper limit was 121 as defined by the pseudo-t^2 statistic but taking the semi-partial R^2 into account, the upper limit is now 33 clusters. This gives a possible range for k of 3 to 33.

The R^2 reflects the variation or difference between clusters [62], higher values are an indication of good cluster formation. In Table 4.15, the R^2 value increases to 0.99 by the formation of cluster 8. The approximate expected R^2 is calculated when requesting the CCC.

In Table 4.15, the ties column is blank indicating that there were no ties when clusters 1 to 50 were formed. At each level of the clustering process, pairs of clusters are identified with the minimum distance between them. Sometimes there can be two or more pairs of clusters with the same minimum

distance. This is called a 'tie' and it must be broken in some discretionary way [62]. If there are ties, then the results of the cluster analysis depend on the order of the dataset. Ties that occur in the beginning of cluster analysis usually have little effect on the later cluster formation. While ties that occur in the middle or latter parts of the cluster analysis should be investigated. The full results for the clusters 1 to 594 using Ward's minimum distance method show that after cluster 333 ties are present at all the subsequent clusters. This is an indication that a method like k-means will produce a better result, as the k-means algorithm repeats or iterates through the dataset recalculating for the best possible solution at each iteration and not keeping the result from the previous step as with hierarchical clustering methods.

The CCC, pseudo-F and pseudo-t^2 need to be used in conjunction to determine the best number of clusters. The CCC had the lowest number of clusters at 3, however all three of these test statistics indicate that the number of clusters should be in the range of 8 to 33. This range is a starting point for determining k using the k-means method.
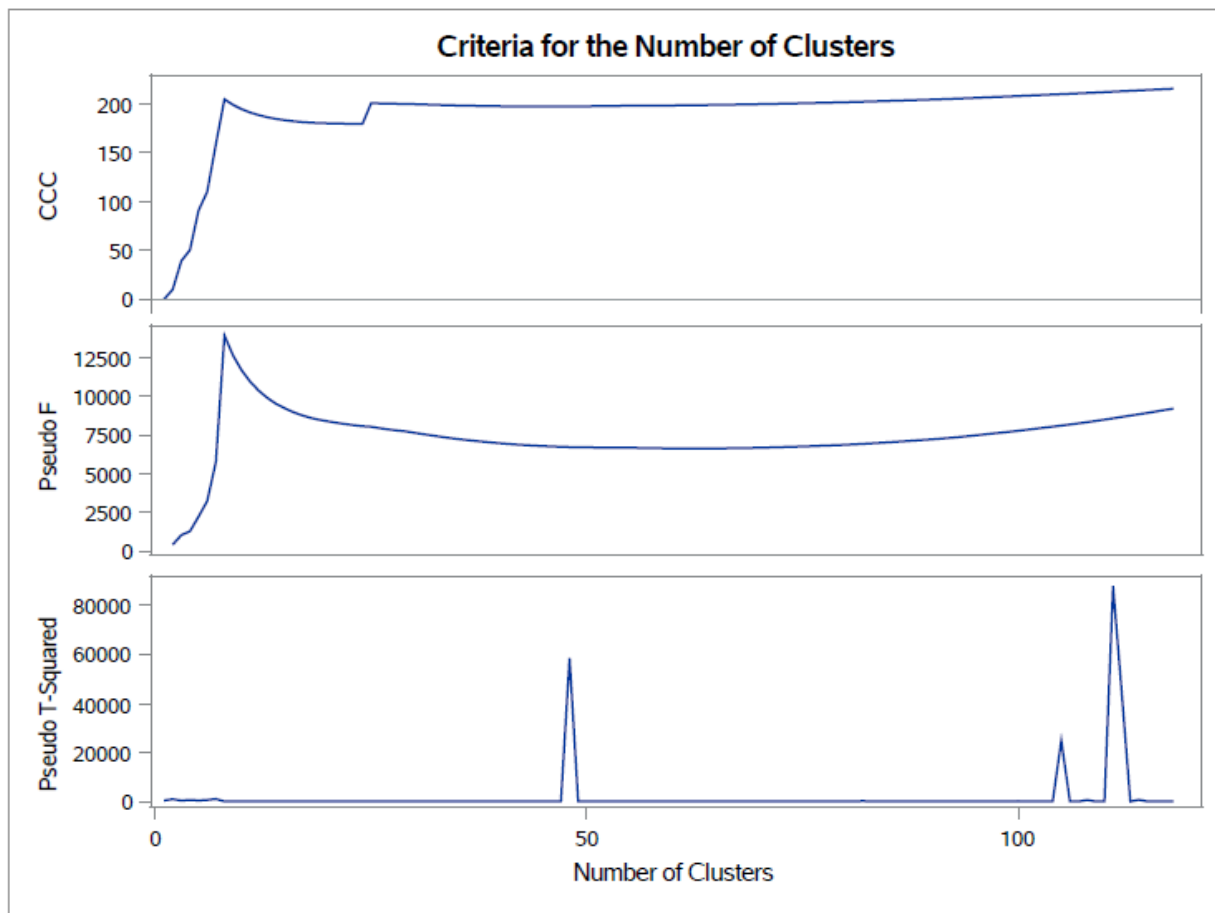
FIGURE 4.6: *Test statistics CCC, pseudo-F and pseudo-t^2 for hierarchical clustering using Ward's minimum distance method*

The PROC CLUSTER procedure in SAS studio® software using Ward's minimum distance method produces a dendrogram. A dendrogram is a graphical representation of the cluster structure formed. The horizontal axis of the dendrogram represents the distance or dissimilarity between clusters. The vertical axis represents the products. Each joining of two clusters is represented on the graph by a vertical line

joining the two horizontal lines. The horizontal position of the split, shown by the short vertical bar, gives the distance between the two clusters [54]. SAS studio® software is limited to a maximum of 200 products to be evaluated and displayed. In order to overcome this limit as the dataset has 594 products, *jmp* Statistical Discovery™ software from SAS [64] is used to produce Figure A.9 in the appendix. Figure A.9, is a basic dendrogram as labels are not able to be displayed due to the number of products but it does give a visual picture of the structure of the dataset.

No. Of Clusters	Frequency	Semi-partial R^2	R^2	Approximate Expected R^2	CCC	pseudo F	pseudo t^2	Ties
120	4	0	1	1	0	9404	2.8	
119	3	0	1	1	0	9307	4.4	
118	9	0	1	0.927	216	9214	32.9	
...	...	...	...	...	...	...	...	
50	5	0	0.998	0.871	198	6704	2.9	
49	5	0	0.998	0.87	198	6713	11.9	
48	12	0	0.998	0.869	197	6725	**59000**	
47	16	0	0.998	0.867	197	6742	8.6	
46	18	0	0.998	0.866	197	6764	15.1	
...	...	...	...	...	...	...	...	
40	15	0	0.998	0.856	198	6970	14.9	
39	19	0	0.998	0.854	198	7019	6.1	
38	29	0	0.998	0.852	198	7073	15.8	
37	75	0	0.998	0.85	198	7127	54	
36	18	0	0.998	0.848	198	7185	64.9	
35	28	0	0.998	0.846	198	7249	17.9	
34	49	0	0.998	0.844	199	7320	83.8	
33	21	0.0001	0.998	0.842	199	7398	6.4	
32	38	0.0001	0.998	0.839	199	7482	19.1	
31	24	0.0001	0.998	0.837	200	7565	9.6	
30	25	0.0001	0.997	0.834	200	7652	21	
29	37	0.0001	0.997	0.832	200	7742	13.9	
28	22	0.0001	0.997	0.829	200	7799	10	
27	98	0.0001	0.997	0.826	200	7867	36.5	
26	8	0.0001	0.997	0.823	201	7949	20.9	
25	45	0.0001	0.997	0.82	**201**	8040	21.2	
24	40	0.0001	0.997	0.817	180	8077	15.2	
23	60	0.0001	0.997	0.813	180	8140	45.1	
22	105	0.0001	0.997	0.81	180	8216	23.7	
21	134	0.0001	0.997	0.806	180	8295	23	
20	149	0.0001	0.996	0.801	180	8390	21.3	
19	198	0.0002	0.996	0.797	180	8489	25.7	
18	296	0.0002	0.996	0.792	181	8615	29.3	
17	304	0.0002	0.996	0.787	181	8778	26.6	
16	63	0.0002	0.996	0.781	182	8981	36.4	
15	91	0.0002	0.996	0.775	183	9232	27.5	
14	313	0.0002	0.995	0.768	185	9522	29.6	
13	335	0.0002	0.995	0.76	186	9896	26.3	
12	375	0.0002	0.995	0.752	189	10000	24.5	
11	466	0.0002	0.995	0.742	191	11000	24.3	
10	485	0.0002	0.994	0.731	195	12000	23.4	
9	507	0.0002	0.994	0.719	199	13000	22.7	
8	539	0.0002	0.994	0.704	**205**	**14000**	22.4	
7	541	0.0108	0.983	0.686	159	5754	**1038**	
6	543	0.0184	0.965	0.664	110	3231	609	
5	548	0.0269	0.938	0.635	**91.1**	2226	422	
4	552	0.0717	0.866	0.594	50.4	1274	638	
3	557	0.0896	0.777	0.528	**39.2**	1028	373	
2	578	0.3681	0.409	0.316	9.98	409	949	
1	594	0.4086	0	0	0	.	409	

TABLE 4.15: *An extract of the test statistics for clusters 1 to 50 using Ward's minimum distance method*

Step 4: Determining the number of clusters: k-Means method

The k-means algorithm is a disjointed cluster analysis procedure, based on the distance calculated from one variable to another for each product [54]. Unlike the hierarchical clustering methods, the k-means method does not form a tree-like structure.

The PROC FASTCLUS procedure in SAS studio® software generates non-random seeds or centre points for clusters, all observations are then divided into clusters based on their distance from the seeds, as per equation (2.29), and displayed in Figure A.4 in the appendix. The variables *Can1* to *Can37* are the standardised variables describing pack sizes, alcohol percentage, carbonation level, sugar content, taste profile and the factor scores.

In the previous section a starting point has been determined for k, with a possible range of 8 to 33 clusters. This range has been expanded to 3 to 33 due to the number of ties formed when using Ward's method. The larger range will ensure that more possible optimal solutions are tested. In Table 4.16, the results from the k-means method of clustering are displayed. The k-means method does not produce a CCC graph or the pseudo-t^2 in the output statistics, only the CCC values are displayed. In Figures 4.7 and 4.8 the results of the CCC and pseudo-t^2 output statistics have been graphed in Microsoft Office 2013 Excel.

As with hierarchical clustering, the CCC values greater than 3 are an indication of good cluster formations [63]. The results in Table 4.16, show that all the clusters in the range 3 to 33 have a CCC value greater than 3. The CCC does not have any large negative values, this indicates that there are no significant outliers [63]. In Table 4.16, the CCC values peak at $8, 9, 10$ and 11 clusters.

In Table 4.16, the pseudo-F statistic for $8, 9, 10$ and 11 clusters indicates well separated clusters as they have the largest values [63]. For the R^2 higher values are an indication of good cluster formation [62], the R^2 reflects the variation or difference between clusters. In Table 4.16, after 7 clusters the R^2 values are very similar and close to 0.99. The approximate expected R^2 is calculated when requesting the CCC.

According to the CCC, pseudo-F and R^2 test statistics $8, 9, 10$ and 11 clusters are all good cluster formations. The original range of 3 to 30 clusters has now been narrowed down to 8 to 11. The reduction in the range is helpful but a single value of k still needs to be determined.

No. Of Clusters	R^2	Approximate Expected R^2	CCC	pseudo-F
3	0.776653	0.52775	39.206	1027.55
4	0.866285	0.59427	50.413	1274.12
5	0.937951	0.63489	91.077	2225.87
6	0.93872	0.66382	82.929	1801.47
7	0.968631	0.68589	124.627	3020.93
8	**0.994031**	0.70372	**204.826**	**13942.22**
9	0.994232	0.71857	**198.882**	**12604.31**
10	0.994446	0.7312	**194.372**	**11619.35**
11	0.994530	0.74212	**189.651**	**10599.35**
12	0.994616	0.7517	**185.699**	9775.09
13	0.994783	0.76019	183.053	9232.4
14	0.994926	0.76781	180.673	8749.16
15	0.994976	0.77468	177.808	8190.66
16	0.995036	0.78093	175.356	7724.21
17	0.995170	0.78666	173.868	7430.57
18	0.995384	0.79193	173.401	7305.65
19	0.995457	0.79681	171.779	6999.2
20	0.995776	0.80134	172.874	7121.97
21	0.995733	0.80557	170.370	6684.89
22	0.995956	0.80953	170.846	6708.44
23	0.995967	0.81325	169.161	6409.99
24	0.996357	0.81676	171.915	6777.32
25	0.996270	0.82007	189.236	6332.13
26	0.996469	0.82322	190.176	6411.68
27	0.996441	0.82622	188.143	6106.25
28	0.996577	0.82908	188.441	6103.61
29	0.996660	0.83182	188.096	6020.99
30	0.996594	0.83444	185.706	5690.66
31	0.996704	0.83695	185.876	5675.38
32	0.996804	0.83936	185.982	5653.61
33	0.996588	0.84168	181.59	5120.56

TABLE 4.16: *Test statistics for 3 to 33 clusters using the k-means method of clustering*
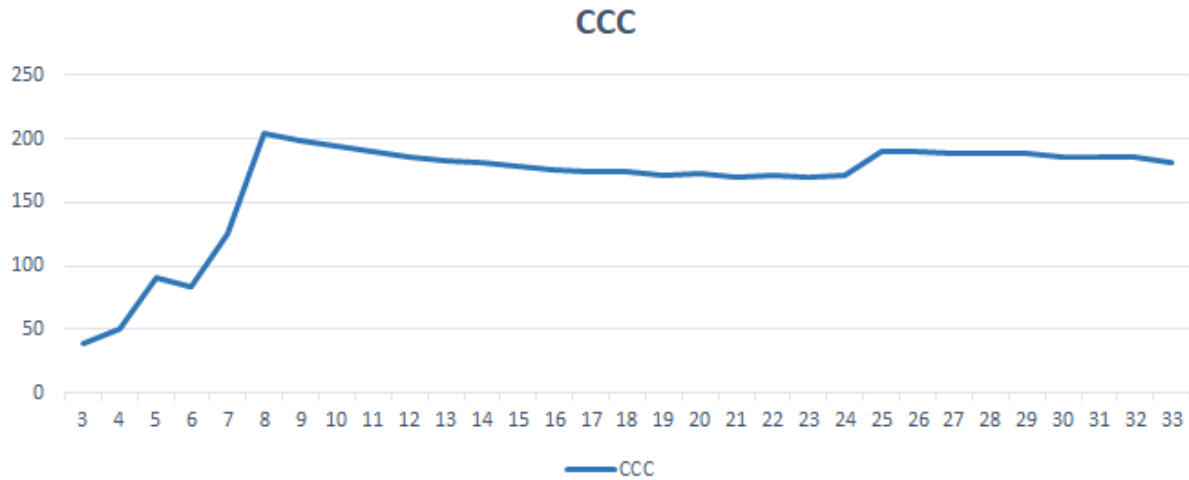
FIGURE 4.7: *Test statistics CCC for 3 to 33 clusters using the k-means method of clustering*
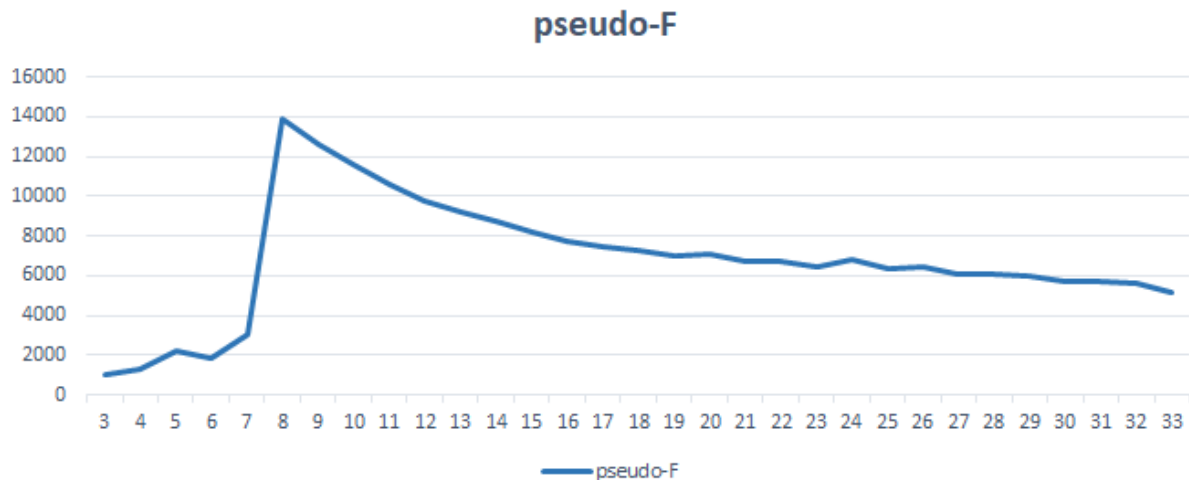
FIGURE 4.8: *Test statistics pseudo-F for 3 to 33 clusters using the k-means method of clustering*

The possible value of k has been reduced and now other attributes of the clusters need to be investigated to determine a single value of k. In Table 4.17, the frequency distribution of each cluster is displayed. The frequency distributions show the number of products that are members of each cluster.

From the results in Table 4.17, $k = 8$ resulted in 5 clusters with a low frequency ($frequency < 10$), 2 clusters with medium size frequencies ($10 < frequency < 100$), and 1 cluster with a large frequency ($frequency > 100$). The results are similar for $k = 9$, $k = 10$ and $k = 11$. The result for $k = 10$ is the most acceptable, as there are fewer clusters with only two members and almost an even distribution of low and medium frequency clusters. Also for $k = 10$, the formation does not include any clusters with only one product. This is important because these clusters will be used in the next section, where the formula requires each product in a cluster to have a market share relative to another product, hence

having only one product in a cluster would not be acceptable.

| | Cluster Formations | | | |
No. Of Clusters	Frequency k=8	Frequency k=9	Frequency k=10	Frequency k=11
1	4	485	2	4
2	539	4	21	24
3	16	54	16	15
4	2	2	4	7
5	5	21	30	16
6	21	16	459	5
7	2	5	5	5
8	5	5	5	2
9		2	2	2
10			50	21
11				493

TABLE 4.17: *The number of products per cluster for $k = 8, 9, 10, 11$ using the k-means method of clustering*

A single value of k has been determined, $k = 10$. In Figure A.4 in the appendix, the code for SAS studio® software for clustering using the k-means method and the code for $k = 10$ is shown. In Table 4.17 the final result for $k = 10$ clusters includes a cluster with a large frequency ($frequency > 100$). This cluster can be further refined by following the above procedure for the sub-set of products in Cluster 6 only.

Cluster 6 contains 459 products, a simple starting place for k when using the k-means method of clustering on the sub-set of products in Cluster 6 would be $k = 5$. If 493 is divided into five groups then the clusters would be of medium size, provided it is a roughly even split. After testing $k = 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16$ and 17, the output results show a spike in the CCC and pseudo-F statistic at $k = 16$. At $k = 16$ the clusters are more evenly distributed with the largest cluster having a frequency of 128, no clusters having only one member.

	Cluster Formations	
No. Of Clusters	Frequency k=10	Sub-set of Cluster 6 and k=16
1	2	6
2	21	15
3	16	40
4	4	31
5	30	8
6	459	26
7	5	5
8	5	82
9	2	128
10	50	4
11		32
12		5
13		55
14		9
15		2
16		11

TABLE 4.18: *The number of products per cluster for the original 10 clusters, as well as the number of products of the sub-set of Cluster 6*

The final frequency distribution when clustering with $k = 10$ clusters and clustering Cluster 6 again, is displayed in Table 4.18. The frequency distribution of the sub-set of Cluster 6 still contains a cluster with 128 products. Increasing the value of k from 16 onwards when using the k-means method only develops more clusters with single member products and does not dramatically decrease the size of the large cluster. Therefore further reduction of the results will not be performed.

Step 5: Interpreting results

The results of the cluster analysis are displayed in Table A.29 to Table A.35 in the appendix. The 10 clusters and the sub-set for Cluster 6 are shown in the tables. The results of the cluster analysis is that the set of 594 products have been grouped into 10 clusters or 26 smaller clusters when including the clusters and sub-sets.

In Table 4.19 each cluster is given a general description of the type of products that have been grouped together. The products are now grouped into competitor sets. These groups are important and can be used for further analysis. The external market factors must still be considered. The BCG growth share matrix will bring the clusters or competitor sets and the external market factors together.

Cluster	Sub-set	Description
1		Spirits coolers or premixed spirit RTDs in medium size packaging (330ML) .
		Ready-to-drink products with a spirits base in a single serving convenience orientated size packaging.
2		Beer, beerlike RTDs and ciders products in large packaging.
		Beer or similar taste profile products in a sharing or multiple serving size packaging.
3		Spirit and cream liquors and flavoured spirits.
		Flavoured spirits and liquor products that can be served as a single or with a mixer or in cocktails.
4		Small pack spirits (less than 750ML) for popular or high market share spirit brands.
		Popular spirit brands in a small packaging format to make the price accessible.
5		A variety of sizes of premixed spirit RTDs, fruit flavoured RTDs and affordable wine.
		Popular brands in single serving packaging sizes and also in large sharing packaging formats.
6		Popular products across the segments and packaging formats.
	6.1	Large pack format (1L) price accessible whiskeys.
	6.2	Affordable premium wine.
	6.3	Popular or high market share wines in medium or large packaging.
	6.4	Mostly brown spirits, brandy and whisky, in medium size packaging (750ML).
	6.5	Large pack format (1L) spirits which are popular or have high market share.
	6.6	Affordable whisky in medium size packaging.
	6.7	Generally large pack format (1L) white spirits.
	6.8	A large selection of products across the wine category and across the pack formats.
	6.9	A selection of products across the categories and also across the pack sizes.
	6.1	Speciality brown spirits and a gin in medium size packaging.
	6.11	Sparkling wines and champagnes.
	6.12	High energy medium pack cider RTDs and flavoured RTDs.
	6.13	Popular or high market share wines in various packaging sizes.
	6.14	Wines from the a large market share wine brand have clustered in medium size packaging.
	6.15	Affordable brown spirits in medium size packaging.
	6.16	Various RTDs in small and medium size packaging.
7		Premium products in the brown spirits and premium RTDs categories.
		High priced spirits in the Cognac category as well as high prices Cider and premixed spirits.
8		Small pack format wines.
		Sweet wine products in a single serving convenience orientated size packaging.
9		RTDs in a small pack format.
		Products for a grape-based RTD brand in a single serving convenience orientated size packaging with an intrinsic that is a semi-sweet sparkling wine.
10		Beer, beerlike RTDs, ciders and premixed flavoured drinks in small to medium pack format (less than 440ML).
		Ready-to-drink products with a beer base or taste profile in a single serving convenience orientated size packaging.

TABLE 4.19: *A description of the product grouped in the original 10 clusters, as well as the sub-set of Cluster 6 when using k-means method of clustering*

Step 6: Validation

Cluster analysis needs to be validated due to its heuristic nature, the procedure will always return a result even if there are no partitions present in the dataset being evaluated [44]. There are three criteria for cluster validation they are internal, external and relative cluster validation [8], [13], [76]. Relative cluster valuation will not be examined here as the process of varying the number of clusters was addressed in detail in the process of determining $k = 10$.
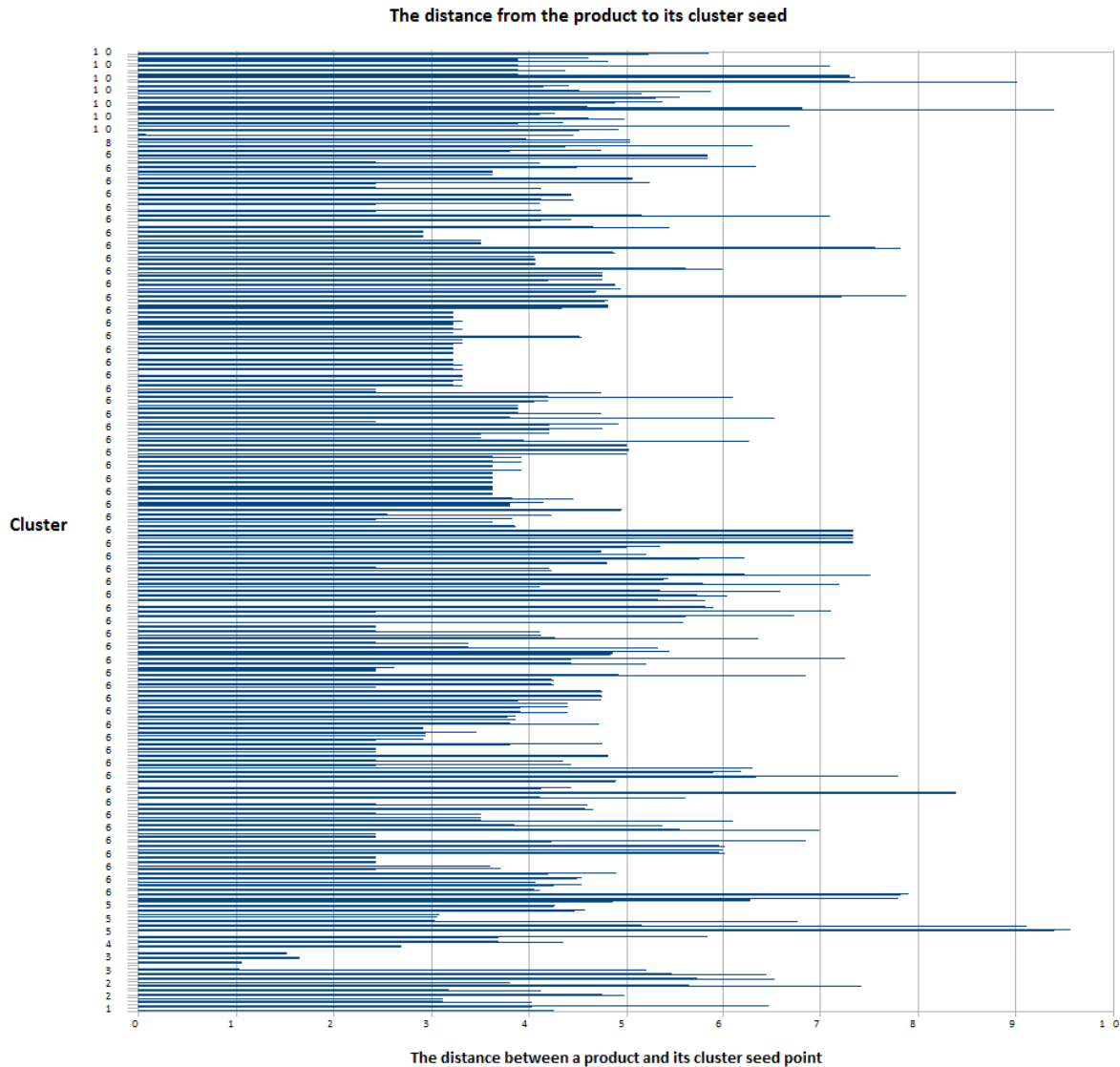
FIGURE 4.9: *Graphical display of the distance between each product and its cluster centre using the k-means method*

The first cluster validation criteria is to evaluate the internal cluster compactness or cluster cohesion, by reviewing the distance from a product to its cluster centre [76]. In Figure 4.9 the distance between each product and its cluster centre is plotted. The plot shows that there are no significant outliers and that all

the products are within the same range of distance from the cluster seeds. This is an indication that the results are valid and can be trusted. The researcher can decide on a cut-off when determining if there are outliers rather than a visual inspection of the distances.

The second criteria, external clustering validation, is to compare known partitions or segmentation within the industry and that of the results of cluster analysis [76]. In the alcoholic beverage industry as described previously, there are three main segments and each segment has sub-segments or categories.

Figure 4.10 displays the segments and categories within the alcoholic beverage industry, noting that the categories total to 10, this is also the number of clusters found in the results. This is evidence that there are partitions in the data and the cluster analysis produced a number of clusters that is consistent with the known number of partitions in the data.
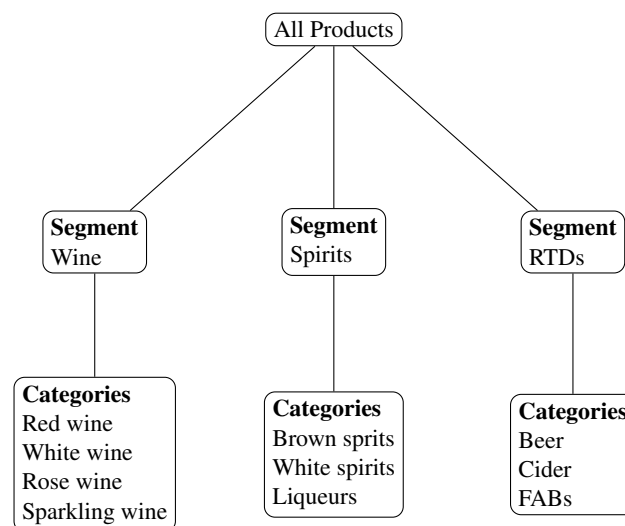
FIGURE 4.10: *Externally known segmentation within the dataset*

Another way of validating the results is to evaluate the clusters formed and overlay the known partitions in the data. Figure 4.11 displays each product as a data point or coordinate. Each product is plotted relative to its aggregated annual sales volume in litres for the calendar year 2017 and assigned on the x-axis to the cluster that it falls into. Each product is also colour coded to show which segment it belongs to. The plot shows that each cluster grouped products mostly from the same segment and within a similar volume band. This is further evidence to trust in the results from the cluster analysis.
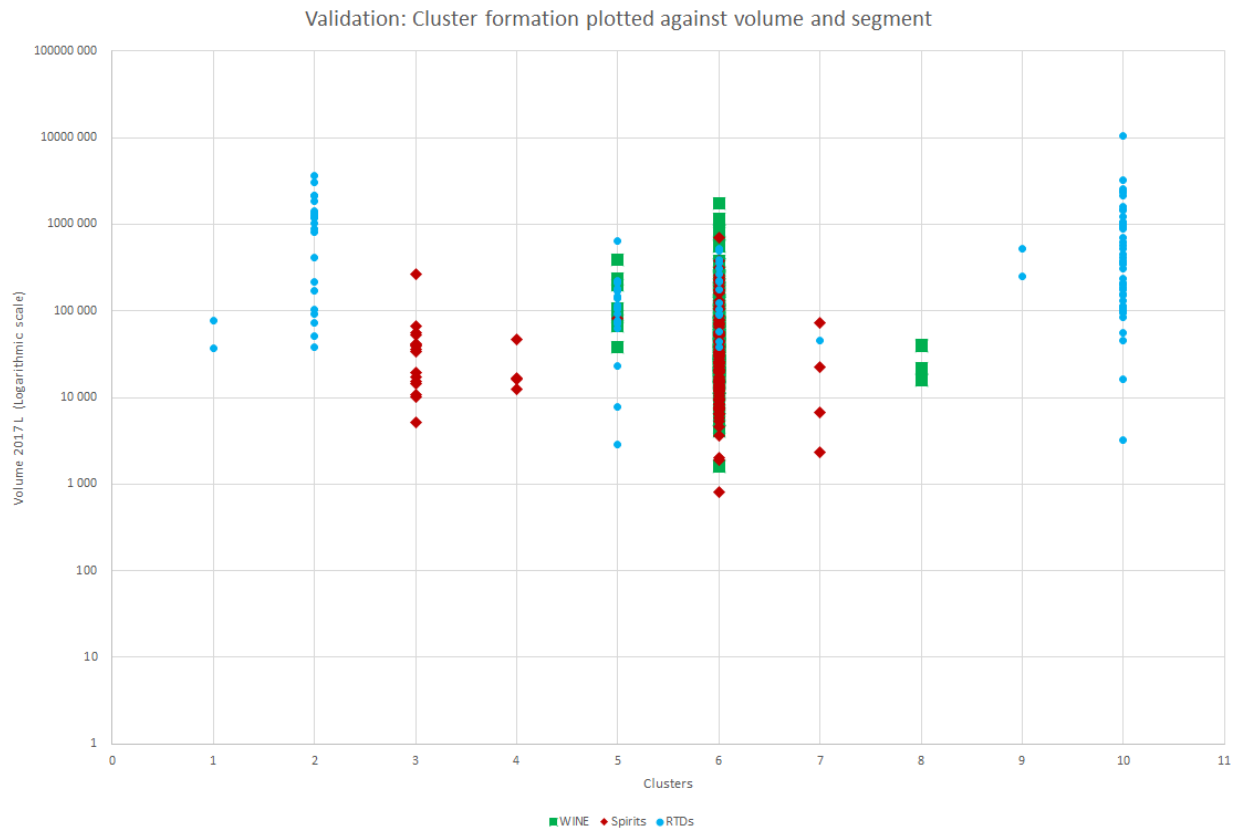
FIGURE 4.11: *Validation of the cluster pattern formation*

4.2.3   Boston Consulting Group growth share matrix procedure

At the essence of the BCG approach to the growth share matrix, is to represent a firm in terms of a portfolio of businesses, each business with its own unique contribution to profitability and growth of the overall firm. This approach views the firm as not one solidified entity but a group of independent distinct units that should formulate bespoke strategies [30]. Products can be seen as the 'businesses' within the portfolio of the alcoholic beverage industry.

The market growth rate for each product will be calculated as per equation (2.22). Using the total sales value for the 2017 calendar year from the electronic till data that has been collected from a supermarket retailer in South Africa, the results for each category are displayed in Table 4.20. Each product belongs to a category and the market growth rate of the category will form the y-axis in the BCG growth share matrix. If an alcohol category has a sub-category then the growth rate of the sub-category will be used for example, spirits is a segment and white spirits is a category of spirits, while cane, gin and vodka are sub-categories.

BCG Growth Share Matrix: Market Growth Rate	
Category	Growth Rate
BEER	16%
BAG IN BOX WINE	6%
BRANDY	4%
CANE	-4%
CIDER	10%
FABS	9%
FORTIFIED WINE	18%
GIN	48%
LIQUEURS	9%
PERLE WINE	-9%
RED WINE	9%
ROSE WINE	-1%
RUM	1%
SPARKLING WINE	8%
VODKA	19%
WHISKY	-1%
WHITE WINE	8%
TOTAL	9%

TABLE 4.20: *BCG growth share matrix: Market growth rates using total value for 2017 from the electronic till data that has been collected from one of the largest supermarket retailers in South Africa*

The relative market share for each product will be calculated as per equation (2.23), using the product with the highest sales value for the calendar year per cluster or sub-cluster as the leading competitor. The relative market share for each product will form the x-axis in the BCG growth share matrix.
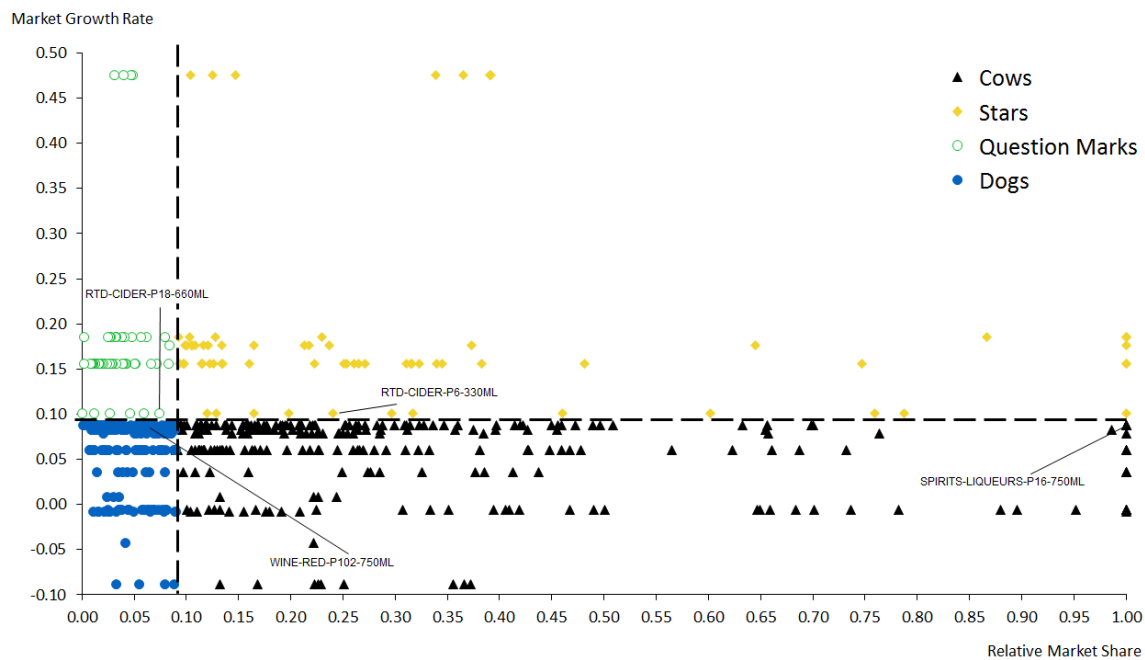
FIGURE 4.12: *BCG growth share matrix results*

The final results in Figure 4.12 and Figure 4.13 show how the products are distributed between the four taxonomies. The highest frequency of products is in the Cows quadrant or maturity phase of the PLC. This is to be expected, as the largest part of the market is made up of mature brands that have made it through the initial two phases. The second biggest group are the Dogs, these are brands that have come through maturity and the start to decline.

The PLC starts with new innovations, these are brands in the Question Mark phase and these represent the smallest frequency of brands. Those that become established and go into the growth phase become Stars and as these products increase in sales volume and value, they enter the Cows phase. The results support the PLC methodology and the frequency of products in each phase also follows the PLC methodology as indicated in Figure 4.13. In Figure 4.12 the 4 products that will be used as a sample for investigation have been highlighted, each product in the sample is in a different life cycle stage.

BCG Growth Share Matrix
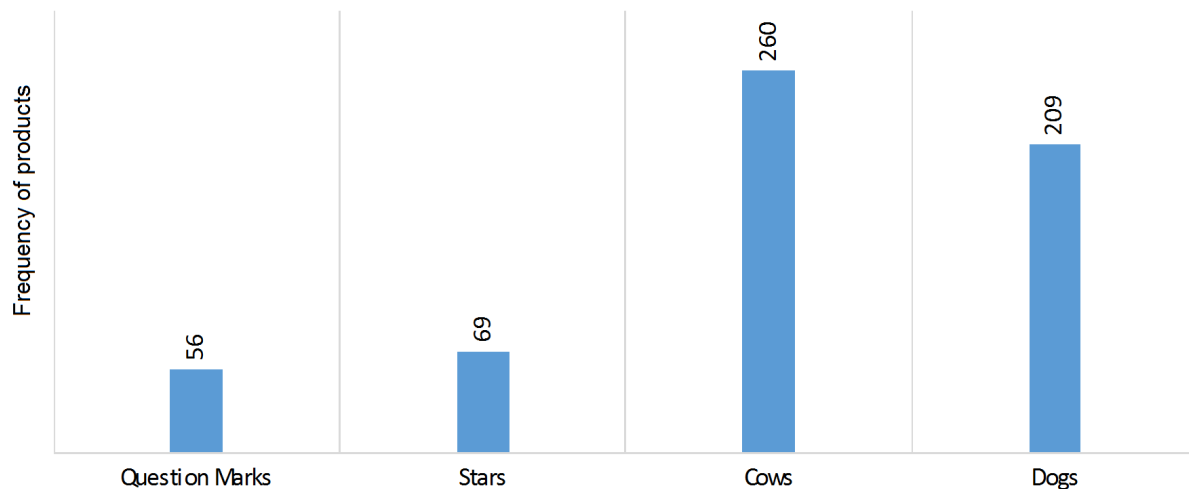
FIGURE 4.13: *BCG growth share matrix frequency distribution*

4.3   Results for determining a competitor set using the product life cycle methodology

The objective of using the PLC methodology is to determine a product's competitor set. A product is evaluated against a defined set of other products that have a relationship with the product being evaluated, called the competitor set. The market and other economic factors that are possible forces acting on a product, also need to be taken into account. Factor analysis, cluster analysis and the BCG growth share matrix have been used to define each product's life cycle stage and group them into competitor sets.

4.3.1 Factor analysis and the product life cycle

Factor analysis gives a view of how the characteristics or attributes that describe the variables are grouped together [80]. The resulting model, using the unweighted least squares method with the MAX option and Quartimax rotation to extract 8 factors, had a high overall KMO score (0.8871). The 8 constructs formed are: Enjoyment, Unreachable, Dislike, Isolation, Popular, Accessibility, Availability and Un-stylish.

In Figure 4.14 the relative size or width of each factor is displayed as a percentage of the 2017 total value share for the sum of all the products that are most strongly associated with that factor. Each product has a set of eight different factor scores, a score per factor, this result can be used to associate or allocate each product to a factor. The factor score that had the highest absolute value was taken to be the factor that the product associated most highly with. For example in Tables 4.13 and 4.14 the product SPIRITS-WHISKY-P5-1LTR would be allocated to Factor 6. In Figure 4.14 the chart also displays the percentage share that each alcohol category has within a factor (Red, White and Rose wine are rolled up into the group Wine to make reading the chart more legible).
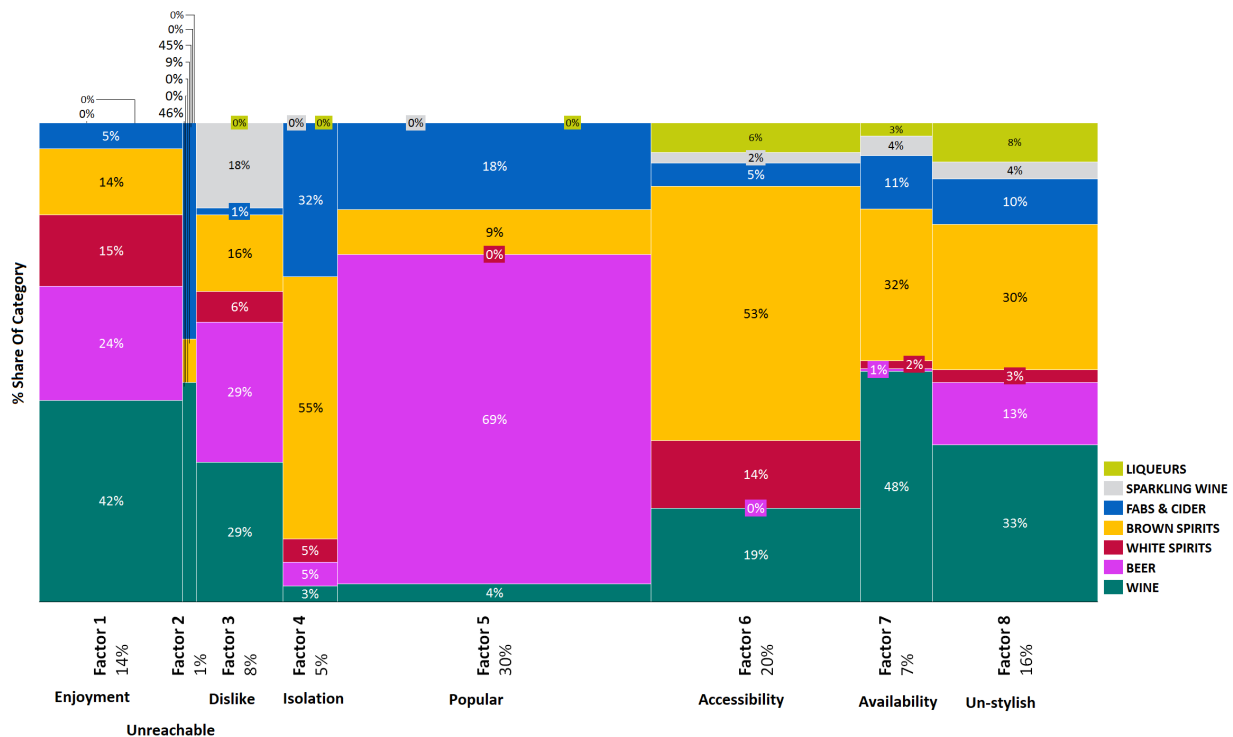
FIGURE 4.14: *A two-dimensional stacked chart of the products loaded onto each factor and the alcohol categories*

In Figure 4.14 wine has the highest percentage share of most factors except for Factors 4 and 5. This means that very few wine products had the highest association of factor scores with Factors 4 and 5, representing the constructs of Isolation and Popularity. From the factor scores it can be interpreted that for the wine category negative feelings towards wine products due to a lack of information about them or a low level of promotions are not characteristics that predominately defined wine products. Factor 7 shows that Availability is an important construct for wine.

Factor 5 is 30% of the total value share and the vast majority of the factor (69%) is made up of beer products that have factor scores associated. Factor 5 is a construct representing products that are popular or have a high market share. As beer is a very well established segment of the market, with many large volume share products, it is expected that beer has the highest percentage share on Factor 5 and has a small percentage share on Factors 4, 6 and 7 (Isolation, Accessibility and Availability).

White spirits are all about accessibility, enjoyment and stylishness as can been seen from the large amount of these products that have high association with Factors 1 and 6. This category has products at affordable price points that are value for money in premium looking packaging. Brown spirits products have a large percentage share for each factor, showing the large number of products in this category. Also that brown spirits is a broad category that ranges from very inexpensive low quality products to luxury expensive quality products.

FABs and ciders, like brown spirits have a high association of products on all factors, except for Factor 3 (dislike). This category has very few products compared to wines or spirits and shares with beer a high association of products on Factor 5, high market share products. The way the products have associated with the factors shows that most FABs and ciders are popular and no negative feelings towards the products due to availability, promotions, packaging, taste and price accessibility.

Unlike FABs and ciders, sparkling wine has a high association of products with Factor 3, showing that many sparkling wine products are difficult to find, lack in promotions, the taste can be limiting as well as the price. Liqueurs also do not make positive impressions with a high association of products and Factors 6, 7 and 8, showing that consumers find them difficult to find, price inaccessible and often un-stylish.

4.3.2   Cluster analysis and the product life cycle

The results showed, using the k-means method, that 10 clusters and 16 sub-sets (clusters within Cluster 6) were formed. Cluster analysis broke the market of products down into groups that are relatively homogeneous within the group and heterogeneous between the different groups. The groups or sets that cluster analysis produces will be further refined to be usable as the competitor sets, this is because the market factors have not been taken into account. The BCG growth share method brings the market forces into the analysis.

Keeping in mind that the alcoholic beverage industry is changing over time as new products enter, products leave the market and there are fluctuations in consumers' disposable income. Due to all these landscape changes, there is a flow of long-established or mature products gradually becoming less relevant to consumers and therefore a less popular purchase. At the same time, the demand for modern new product innovations increases rapidly after they are launched.

The BCG growth share matrix categorised the products into four taxonomies, Question Marks, Stars, Cows and Dogs. The results can now be viewed on many different levels. The results from the cluster analysis can be overlaid with the BCG growth share matrix, this will add another level to the BCG growth share matrix and give insight into the relationships within a quadrant or taxonomy.

The overall result of the BCG growth share matrix, representing the product life cycle stages of all the products, can be seen in Figure 4.12. The results will be examined in more detail now, looking into each

quadrant separately, moving in sequence with the life cycle.

The **introductory phase** of the PLC methodology or Question Marks quadrant of the BCG growth share matrix groups products that are typically new launches into the market, new packaging formats and product range extensions. These products are typically characterised by slow sales growth and low distribution. In this life cycle stage marketers are looking to get acceptance for the new product. Profits will be limited as explosive sales growth has not happened yet.

If a product is not a new entrant but is in the introductory phase, this product's category is in a growth phase but the product itself is not. This is where the category is experiencing high consumer interest or growth as consumers have shifted into this category, but the product, although established, has a small market share compared to its competitors. As with the new launches or innovation in this quadrant the marketers' objective is the same, to stimulate demand through awareness.

In Figure 4.15 and Figure A.10 in the appendix the results for the introductory life cycle stage are shown, firstly highlighting the clusters and then the segments (wine, spirits and RTDs). In each of the figures to follow the product's sales value for the calender year 2017, is represented by the relative size of the bubble. The results can be further viewed in more detail in Figures A.11 and Figures A.12 in the appendix, looking at the RTDs and spirits separately there were no wines in this quadrant.
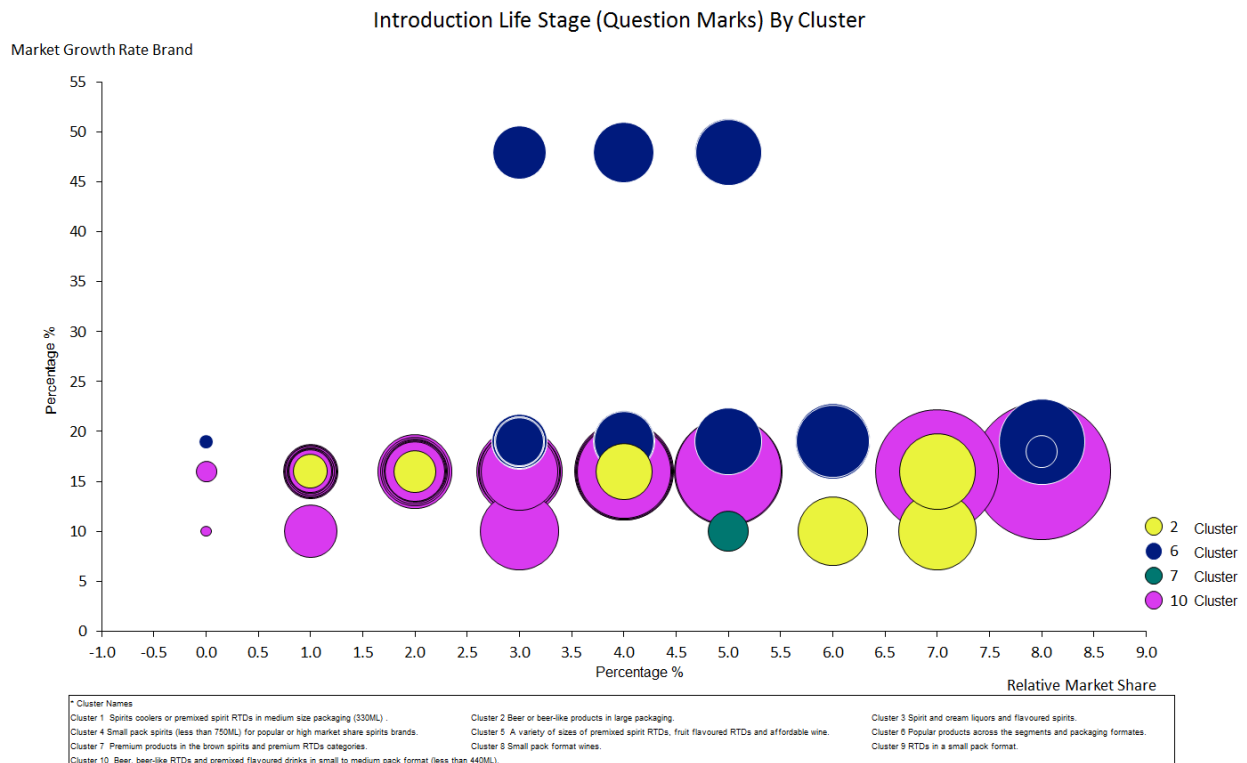
FIGURE 4.15: *Introduction life cycle stage by cluster*

The results for the introductory phase show most of the products are in the same range for market growth, except for the Gins (see Table 4.20). This is because the Gin category is experiencing an upswing. The

spirits that have been grouped are predominately range extensions for example, SPIRITS-VODKA-P15-750ML is a flavour extension of the vodka brands' range. RTDs follow the same trend, with predominately large pack format RTDs and flavour extensions added to the range.

In the introductory phase the products that have been grouped together are range extensions, pack extensions and low market share high category growth. This means that they are relatively small products in terms of market share but are in a category that is experiencing high growth. These products will all be competing with each other for the consumers' attention as a new entrant or a small brand in a growing category that consumers have not yet noticed. The common thread is that all of these products have potential for growth but they need consumers to try them, get distributed and become established.

As with the introductory phase, in the **growth phase** or Stars in the BCG growth share matrix, the categories are in high growth as well as the products. In the growth phase the products also have a high relative market share. In this life cycle stage product growth needs to be stimulated in order to maintain growth through to the maturity phase. Advertising investment is typically high, this investment is made to push the product to reach the next stage. Both sales and distribution should be increasing for products in the growth phase. The main objective for marketers in this life cycle stage is to gain brand or product loyalty. The growth life cycle stage is when a large loyal base is built for a product.

In Figure 4.16 and Figure A.13 in the appendix the results for the growth life cycle stage or Stars are displayed by cluster and segment respectively. The products which have been grouped at the top towards the left are characterised by low relative market share for this taxonomy and high category growth. These are products who are not small competitors but they are not the market leaders in a category that is experiencing high growth.

In Figure 4.16 and Figure A.13 in the appendix, the group of products in the bottom left are defined by low relative market share and category growth for this taxonomy. These are products that are mid-level products compared to competitors but they are not the market leaders, the category is growing but not as fast as other categories in the growth phase. The final group are the market leaders, this is the group of products on the bottom right. These products are the market leaders relative to competitors and their category is in growth, the category growth is less than other categories in this stage.
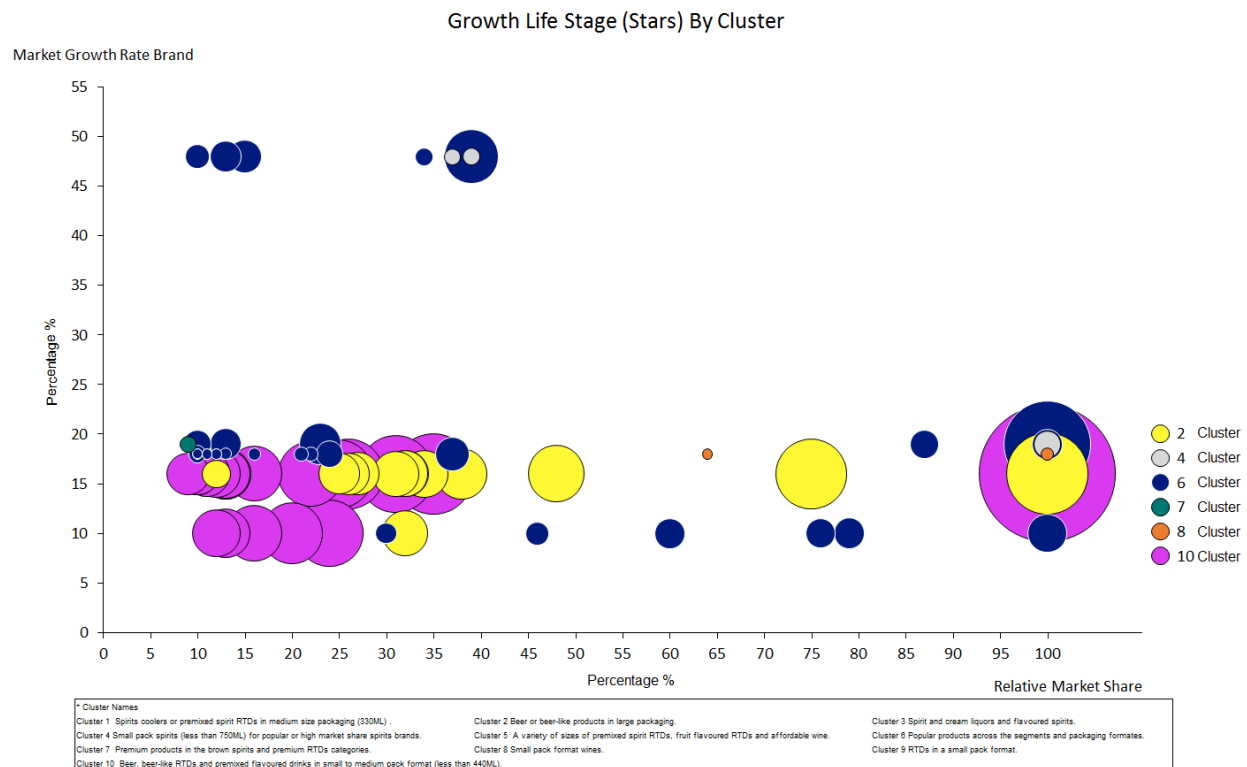
FIGURE 4.16: *Growth life cycle stage by cluster*

In Figure A.14 in the appendix, the RTD products with high sales have been grouped. The clusters reveal that they have also again been grouped according to the size of the product, Cluster 10 having the largest sales value products and then Cluster 2 and 6 respectively. Figure A.15 in the appendix, displays the results for spirits in the growth phase, there are only products from the white spirits category present. Figure A.16 in the appendix, shows that there are not many wines in the growth phase and the products that are there are all low priced and relatively high in alcohol.

Two market leaders have been grouped together in the growth phase: RTD-BEER-P24-340ML and SPIRITS-VODKA-P5-750ML. These two products together account for 7% value share and 10% volume share of the total market. This quadrant has products with high sales for their segment, showing that these are products that either have high frequency of purchase or market penetration. These products will be competing with each other to become popular or stay big brands. In the previous stage the products were competing to get noticed, now they are competing to become or stay popular.

Once a product has become established, the sales volumes are still high but the rate of growth slows down, this is the **maturity phase** or Cash Cows. The products have reached acceptance by the consumers. Loyal consumers are purchasing the products but they are not attracting new consumers all the time as they did in the growth phase.

The products will still have a large sales base and yield substantial profits. Advertising investment will still be substantial in order to maintain the products at maturity for as long as possible. Maintaining distribution is very important in this phase, if a product reduces its footprint it will quickly move into decline.

In Figure 4.17 and Figure A.17 in the appendix, the results are displayed for the products in the maturity life cycle stage. The maturity life cycle stage has a much higher frequency of products grouped than the previous two. The previous two stages also had very few wine products, while the maturity life cycle stage graph has many. In the growth phase the RTDs grouped had a predominately large market share, not just relative share, while in the maturity life cycle stage the smaller value share RTD products have been grouped. The spirits products are the opposite, having had predominately low value share products in the growth phase and large value share products in the maturity phase.
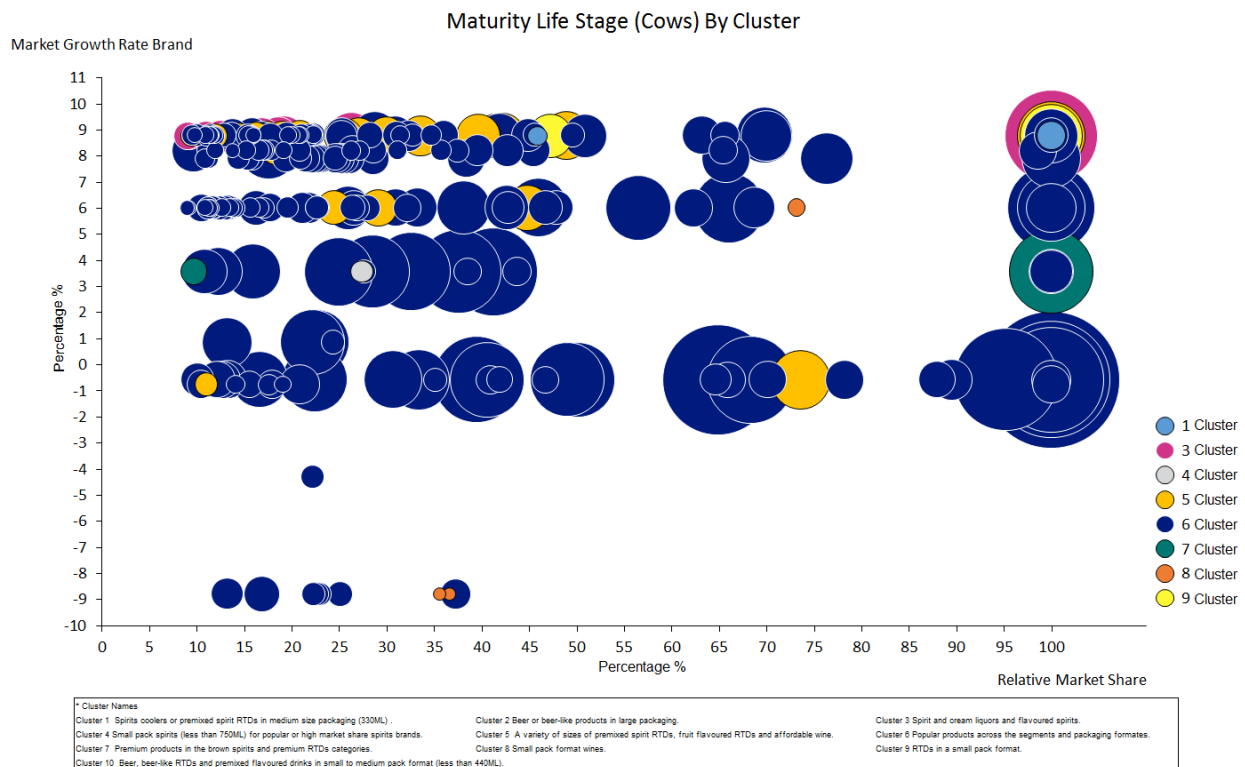
FIGURE 4.17: *Maturity life cycle stage by cluster*

In Figures A.18, A.19 and A.20 in the appendix, the RTDs are mostly grape and energy based drinks in small to medium sized packaging. The spirits products are predominately from the brown spirits category. Some products are declining in market growth, like the whisky products, meaning that while the product has a sizeable market share now it is unlikely to grow. In this quadrant there are well established products with a large market share, they are the market leaders for their clusters. As with whisky the perle wine category is in decline, while still having sizeable products.

All the products in the maturity phase are competing with each other to remain popular. In the introductory phase they competed for consumer attention to get sampled, then in the growth phase they competed to become part of the consumers' regular choice. In the maturity phase the products compete to stay relevant and hold or maintain the large share of the consumers' consumption behaviour. The products are competing with each other to maintain the space they take up on shelves in the retailers. Distribution remains very important in this life cycle stage.

The final stage of a product's life cycle is the **decline phase** or Dogs from the BCG growth share matrix.

These are products with small relative market shares, they are not the market leaders. They are also products in categories that have below market average growth or have gone into decline. This is when the large loyal consumer base built in the growth phase starts to diminish. At this life cycle stage marketers will lower investment as profits decline and are unlikely to increase again.

In Figure 4.18 and Figure A.21 in the appendix, the results for the products in the decline life cycle stage are displayed. As with the maturity phase, there are a large number of products grouped together in the decline phase. The decline phase does not mean that a product will die out in the short team, this phase can go on for a long time. In Figure A.22 in the appendix, the results show that there are few RTD products in this life cycle stage. Also as indicated in Figure A.23 in the appendix, there are less spirits in the decline phase than in the maturity phase but a large amount of wine products. In Figure A.24 in the appendix, some of the wine products like WINE-PERLE-P11-750ML, have a large value but the category is not growing and it is not the market leader.

All the products in the decline life cycle stage are competing for consumer preference shifts or a shift in market conditions. These products typically have little investment and marketing strategies are aimed at shifting consumers' preference to a once popular product, often a 'product or brand make over' or a new image is part of the strategy. These products are all competing to sustain their volumes and not lose share.
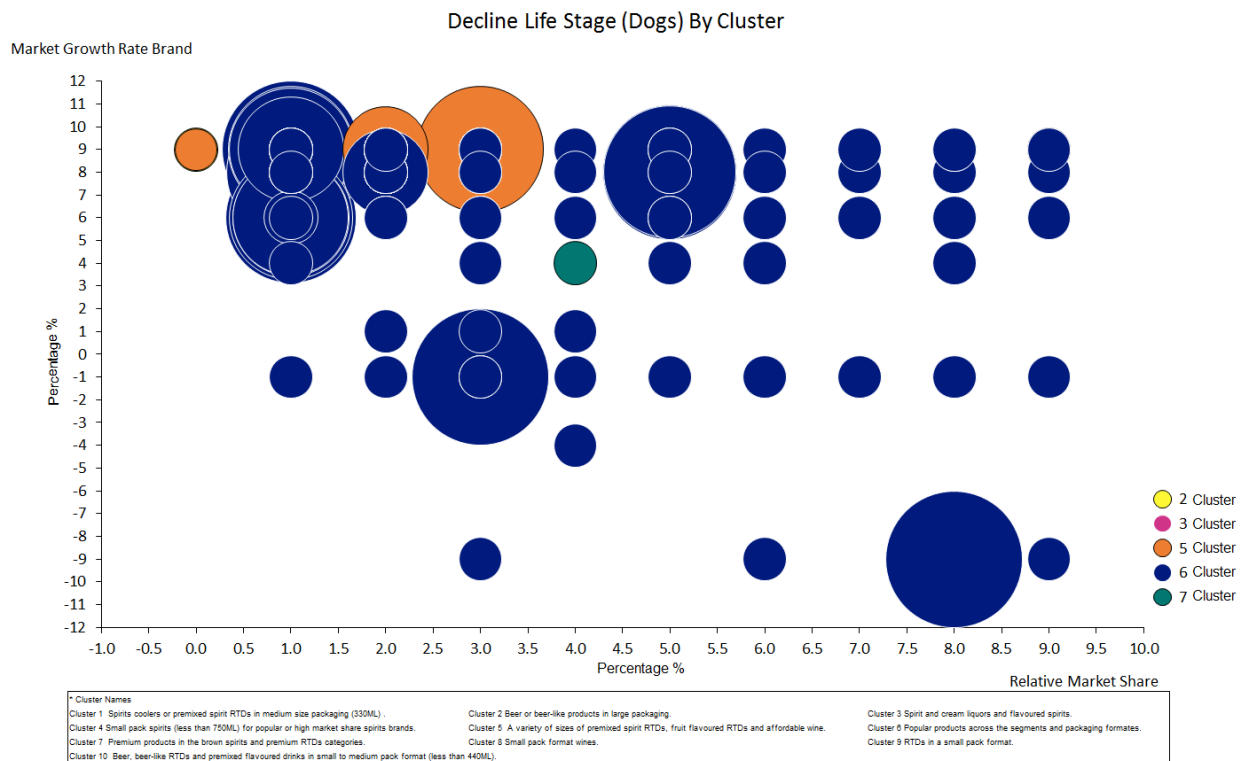
FIGURE 4.18: *Decline life cycle stage by cluster*

4.3.3 The product life cycle

Previously in Figure 4.13 the PLC results for all the products were shown, the results show that when graphing the full market an s-shaped curve is formed. When looking at the PLC by segment it becomes apparent that the different segments have unique shaped PLC curves. In Figure 4.19 the RTDs segment had a large number of products in both the introduction and growth life cycle stages but very few in the maturity and decline phase. This meaning that products tend to increase rapidly in sales, having a very steep life cycle curve and once they stop growing they pass though maturity and decline phases quickly. This is indicative of the innovation and new entrants, pack extensions and variants in this category. Marketers would invest a large amount in the launch and growth of these products.
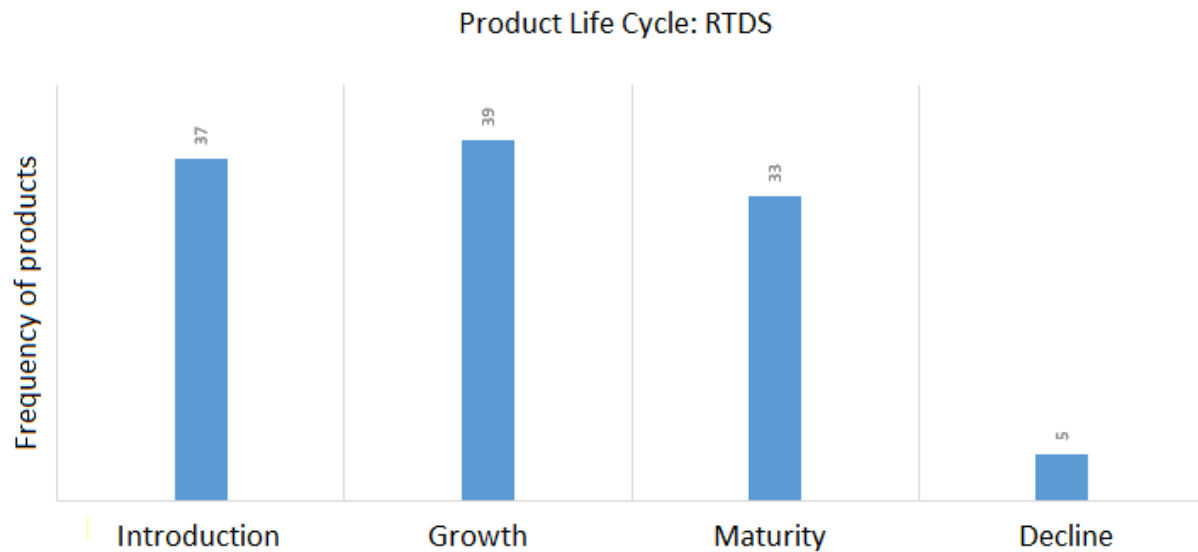
FIGURE 4.19: *PLC frequency distribution for RTD products*

In Figure 4.20 the spirits segment is characterised also by a very steep growth curve with few products in the growth phase and a number of products in the maturity and decline phases. This showing that innovation is high but not all products make it through to the growth phase and those that do have a long maturity and decline phase.

Product Life Cycle: Spirits
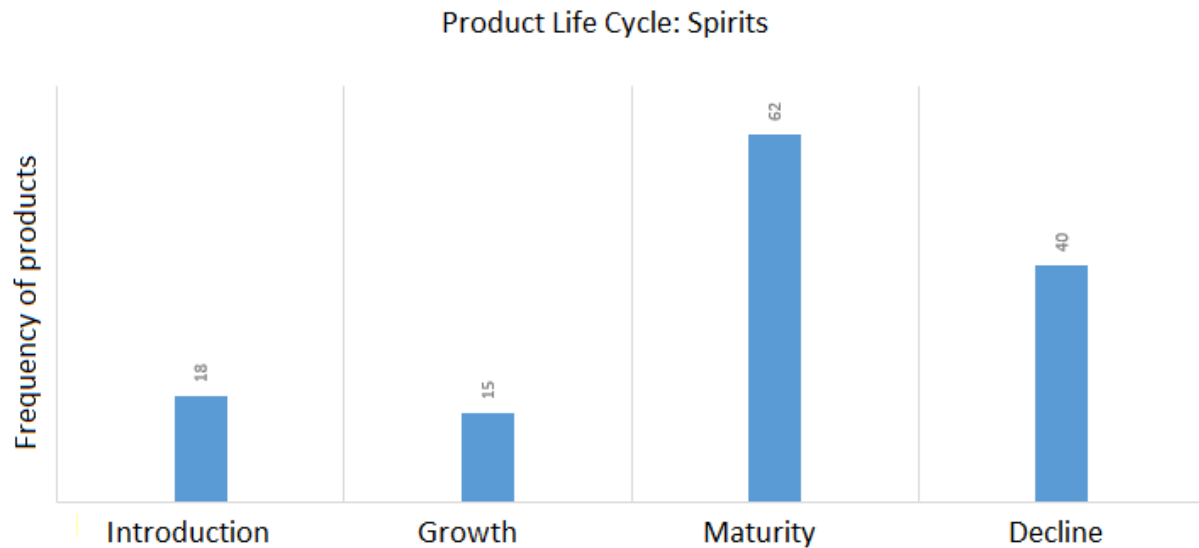
FIGURE 4.20: *PLC frequency distribution for spirits products*

The wine segment is different to the other two. Figure 4.21 shows that this segment has very few products in the introductory and growth life cycle stages and many in the maturity and decline life cycle stages. This shows that there has been little innovation in the wine category and products are either quickly accepted by the market and move through to the maturity stage or they do not make it.

Product Life Cycle: Wine
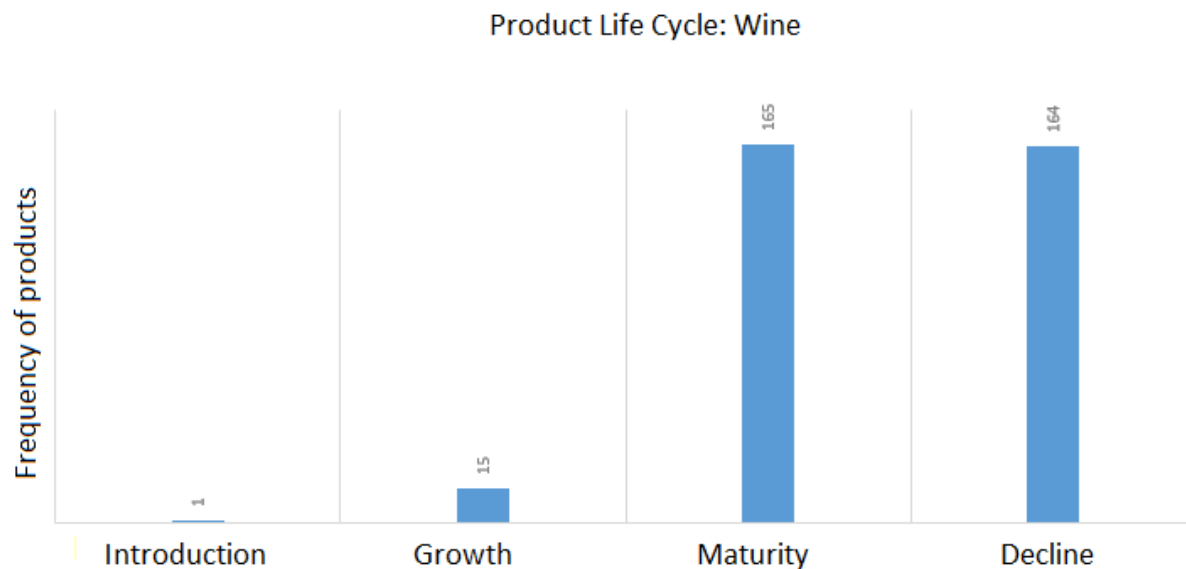
FIGURE 4.21: *PLC frequency distribution for wine products*

In Figure 4.22 the results are displayed showing the number of products in each taxonomy, cluster and the segment. This also displays the PLC of each segment described above. The results for each product indicating which factor, cluster, sub-cluster and taxonomy it belongs to are in Figure A.42 to Figure A.54

in the appendix.
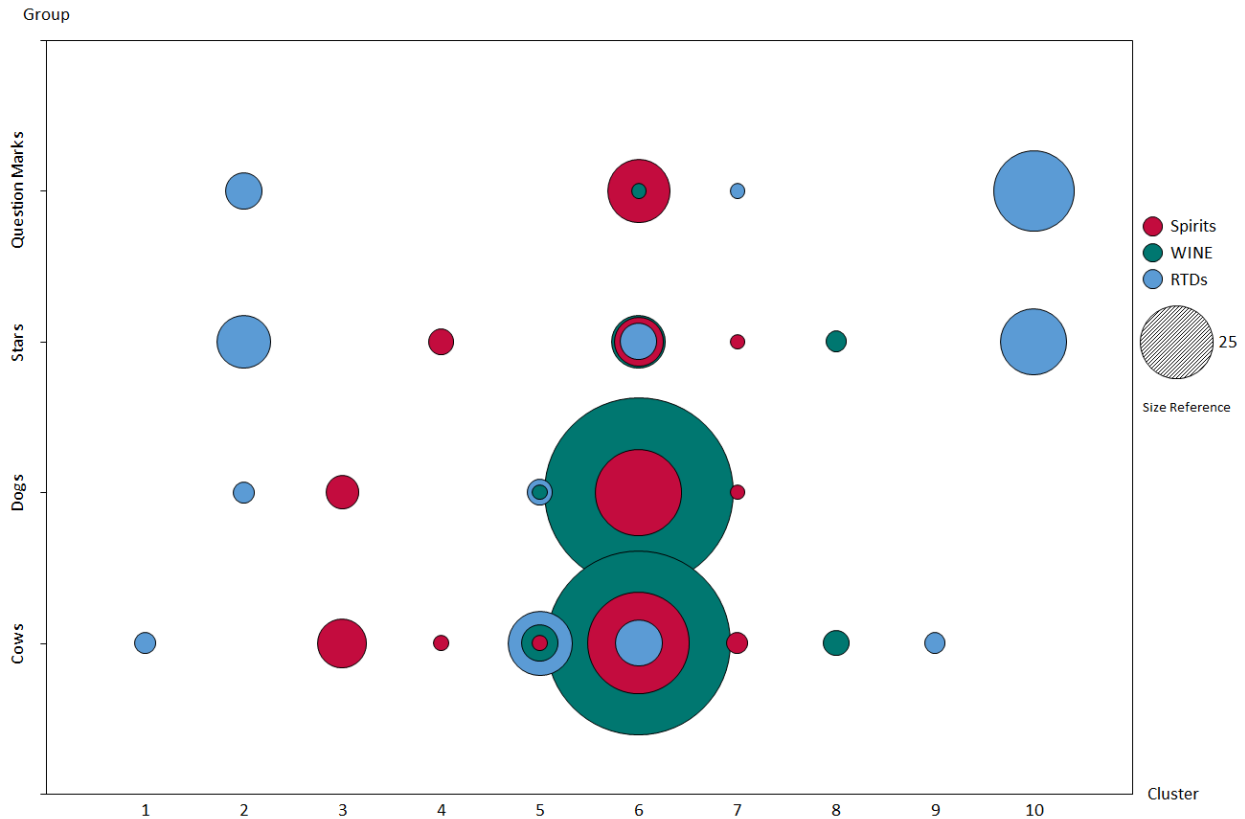
FIGURE 4.22: *Cluster analysis and BCG growth share grid overlaid with product frequency and segment*

4.4 Method for the price elasticity model to determine a competitor set

The price elasticity methodology will serve as an indication that a given product has a relationship with other products and consumers' buying behaviour [75], [40], [82]. This implies that a change in price of a product can be seen by the relationship with the change in volume of another product, a competitor. A competitor set is a group of products that meet the same consumer need [39]. As the price of one product changes, say increases, consumers purchase less of this product and more of the other products available. Consumers are substituting one product for another, as all these products meet the same need. This is the relationship that price elasticity is measuring.

The constant elasticity model is also known as the log-log model [28]. In this model described by equation (2.21) the slope or coefficient β_i measures the price elasticity of the dependent variable with respect to the independent variables [28]. The price elasticity of demand, e_p is the percentage change in the quantity demanded if the price of the product or good changes by one percent, *ceteris paribus* for other market factors [45]. If a dependent variable represented by the quantity demanded of a product, does not have a significant relationship with the independent variables, represented by the price changes of other products on the market, then this variable can be excluded from the possible competitor set.

Procedural Step	Details
Step 1: Sample cleaning	Data cleaning, merging and standardisation
Step 2: Data reduction	Industry Insider
Step 3: Model	Regression equation
Step 4: Limitations	Overfit and multicollinearity in independent variables
Step 5: Variable refinement	Removing overfit from the models
Step 6: Final model	Regression with refined set of independent variables

TABLE 4.21: *The Price elasticity model for determining a competitor set procedure decision sequence*

4.4.1   Step 1: Sample cleaning

Data cleaning should always be the first step in any analysis. In Chapter 3.1, the data cleaning process was addressed through Pareto analysis. This process left 594 products for evaluation and the variables for this regression analysis, volume and AUP. As discussed in Chapter 3.2, for each product the variable's volume and AUP will be used on a weekly level for the past 3 years, thus providing 157 data points, as displayed in Table 4.22 as an example of the dataset per product.

Previously in Chapter 4.2.1, the results show that many of the variables do not have a statistically significant linear relationship with each other. A data transformation is necessary, this will be addressed in the model used to define price elasticity. It is not necessary to merge any other variables or information with this dataset.

Year	Month	Week ending date	Volume (L)	AUP (R)
2015	1	2015/01/04	X.XX	X.XX
		2015/01/11	X.XX	X.XX
		2015/01/18	X.XX	X.XX
		2015/01/25	X.XX	X.XX
		2015/01/31	X.XX	X.XX
	2	2015/02/08	X.XX	X.XX
		2015/02/15	X.XX	X.XX
		2015/02/22	X.XX	X.XX
		2015/02/28	X.XX	X.XX
	3	2015/03/08	X.XX	X.XX
		2015/03/15	X.XX	X.XX
		2015/03/22	X.XX	X.XX
		2015/03/29	X.XX	X.XX
…	…	…	…	…
	10	2015/10/04	X.XX	X.XX
		2015/10/11	X.XX	X.XX
		2015/10/18	X.XX	X.XX
		2015/10/25	X.XX	X.XX
		2015/10/31	X.XX	X.XX
	11	2015/11/08	X.XX	X.XX
		2015/11/15	X.XX	X.XX
		2015/11/22	X.XX	X.XX
		2015/11/29	X.XX	X.XX
	12	2015/12/06	X.XX	X.XX
		2015/12/13	X.XX	X.XX
		2015/12/20	X.XX	X.XX
		2015/12/27	X.XX	X.XX
		2015/12/31	X.XX	X.XX

TABLE 4.22: *Example of one product's electronic till data that has been collected from one of the largest supermarket retailers in South Africa from January 2015 to December 2015*

4.4.2   Step 2: Data reduction

As defined in Chapter 4.1, the four products that will be used as a sample and investigated are: RTD-CIDER-P18-660ML, RTD-CIDER-P6-330ML, SPIRITS-LIQUEURS-P16-750ML and WINE-RED-P102-750ML. The dataset has been defined and there are 594 products available to be included in the model as independent variables, interviewing industry insiders is a first step.

A broad approach has been taken when refining the variables to be included after consulting with many industry insiders from Company A. When interviewing industry insiders the questions were kept on a high level, focusing on which categories they thought interacted or rivalled with the four products under investigation. The industry insiders all had very similar responses when defining a product's competitor set. For RTD-CIDER-P18-660ML the competing segments and categories are RTDs, large pack format wine and market leading spirits. For RTD-CIDER-P6-330ML the competing segments and categories are RTDs, liqueurs, premium whisky and sparkling wine. For SPIRITS-LIQUEURS-P16-750ML the competing segments and categories are spirits, large value share wine and RTDs. Finally for WINE-

RED-P102-750ML the competing segments and categories are wine, large value share spirits and RTDs.

Starting with 594 products available to be included as variables after the reduction, RTD-CIDER-P18-660ML has 249 variables, RTD-CIDER-P6-330ML has 206 variables, SPIRITS-LIQUEURS-P16-750ML has been reduced to 199 variables and WINE-RED-P102-750ML to 404 variables. The products that will be used in each model are listed in Tables A.55 to A.63 in the appendix.

4.4.3   Step 3: Model

The price elasticity model from equation (2.21) will be used to define the equations for the four sample products' price elasticity regression models. Starting with RTD-CIDER-P18-660ML, let the volume sales per week in litres be Q_{RCP18} and the average unit price be P_{RCP18}. Let β_{RCP18} be the own elasticity of RTD-CIDER-P18-660ML and β_{RBP14} the cross price elasticity of the change in RTD-CIDER-P18-660ML's demand as a response to the change in RTD-BEER-P14-330ML price [81], all other variables being equal. Let the set of independent variables listed in Table **??** starting with RTD-BEER-P14-330ML (P_{RBP14}) and ending with WINE-RED-P3-1.5LTR, (P_{WRP3}), be used in the model describing RTD-CIDER-P18-660ML's relationship to competitor products on the market. Then,

$$\ln Q_{RCP18} = \beta_0 + \beta_{RCP18} \ln P_{RCP18} + \beta_{RBP14} \ln P_{RBP14} + \ldots + \beta_{WRP3} \ln P_{WRP3} + \epsilon. \quad (4.1)$$

Similarly for RTD-CIDER-P6-330ML, let the volume sales per week in litres be Q_{RCP6} and the average unit price be P_{RCP6}. Let the set of independent variables listed in Table **??** starting with RTD-BEER-P14-330ML (P_{RBP14}) and ending with WINE-SPARKLING-P33-750ML be used in the model (P_{WSP33}). Then,

$$\ln Q_{RCP6} = \beta_0 + \beta_{RCP6} \ln P_{RCP6} + \beta_{RBP14} \ln P_{RBP14} + \ldots + \beta_{WSP33} \ln P_{WSP33} + \epsilon. \quad (4.2)$$

Moving to SPIRITS-LIQUEURS-P16-750ML, let the volume sales per week in litres be Q_{SLP16} and the average unit price be P_{SLP16}. Let β_{SLP16} be the own elasticity of SPIRITS-LIQUEURS-P16-750ML and β_{RFP21} the cross price elasticity of the change in SPIRITS-LIQUEURS-P16-750ML's demand as a response to the change in price of RTD-FABS-P21-275ML's [81], all other variables being equal. Let the set of independent variables listed in Table **??** starting with RTD-FABS-P21-275ML (P_{RFP21}) and ending with WINE-RED-P24-750ML be used in the model (P_{WRP24}). Then,

$$\ln Q_{SLP16} = \beta_0 + \beta_{SLP16} \ln P_{SLP16} + \beta_{RFP21} \ln P_{RFP21} + \ldots + \beta_{WRP24} \ln P_{WRP24} + \epsilon. \quad (4.3)$$

Following for WINE-RED-P102-750ML, let the volume sales per week in litres be Q_{WRP102} and the average unit price be P_{WRP102}. Let the set of independent variables listed in Tables A.61 and A.62 starting with RTD-FABS-P21-275ML (P_{RFP21}) and ending with WINE-FORTIFIED-P12-750ML be used in the model (P_{WFP12}). The price elasticity regression model uses the natural logarithm form of both the dependent and independent variables, therefore, transforming the datasets into linear forms and countering possible heteroscedasticity that might be present in the data [28]. Then,

$$\ln Q_{WRP102} = \beta_0 + \beta_{WRP102} \ln P_{WRP102} + \beta_{RFP21} \ln P_{RFP21} + \ldots + \beta_{WFP12} \ln P_{WFP12} + \epsilon. \quad (4.4)$$

4.4.4   Step 4: Limitations

As a general rule it is advised to include a maximum of one variable for every 10 observations to avoid overfit [28]. Each model will have 157 observations because the dataset is 3 years of weekly records.

This means that each model should not have more than 15 to 16 independent variables. This is a problem as the models described in equations (4.2) to (4.4) have between 199 and 404 independent variables. When a model is 'overfit' the R^2 and adjusted R^2 will be overestimated and tend towards 1. The results of the model will not be significant for the population [28].

Another problem is multicollinearity, as the products by nature will be highly correlated. For example in the RTD-CIDER-P18-660ML model, the competitor products RTD-CIDER-P13-440M and RTD-CIDER-P14-440ML (flavour extensions of a brand) are highly correlated (0.99) with each other. This is because both products are part of the same brand range and will typically get price discounts at the same time, most retailers are discounting a brand's range during a period. The AUP values of the products are variables in the RTD-CIDER-P18-660ML model and they will both be representing the same change or reaction in RTD-CIDER-P18-660ML volume to the product range's change in price. After testing the correlations between the variables it is confirmed that in each model several variables are correlated in the range of 0.9 to 0.99. This will introduce multicollinearity into the model and make the price elasticities β_i unreliable. Most researchers recommend using a correlation of 0.7 as a cut-off for multicollinearity [28].

The solution to reduce or remove multicollinearity in a model is to remove one of the highly correlated variables from the model as the variable that remains will represent the construct. This is not possible in the price elasticity models because of the methodology these models represent. If two variables are highly correlated with each other, they will make the model unreliable but removing one will be to remove an independent variable that might have had a significant relationship with the dependent variable. If for a product the independent variables are removed because of multicollinearity, in essence the researcher is choosing products to exclude from the competitor set, irrespective of the relationship with the dependent variable or product under investigation. Therefore it will not be possible to remove multicollinearity from the models in this way.

Heteroscedasticity should not be present in the models as both the dependent and independent variables have been transformed using the natural logarithm. The residuals will be plotted to test for heteroscedasticity. Autocorrelation in the models will need to be investigated with the Durbin-Watson test as well as with a graphical plot of the residuals over time.

The problems of overfit and multicollinearity will be addressed in two ways. Firstly, a reduction process will be performed on the independent variables for each of the four products under investigation, this will help to resolve the problem of overfit. Secondly, once the refined sets of independent variables have been defined for each product under investigation, each independent variable will be modelled separately with the dependent variable under investigation. The purpose of the separate modelling process is to overcome multicollinearity, avert heteroscedasticity and give each independent variable the opportunity to establish if it has a significant relationship with the dependent variable.

4.4.5   Step 5: Variable refinement

The variable refinement process can be done in many ways, two methods will be applied and the results examined and the most appropriate method will be selected. In the first method the number of products will be refined by surveying industry insiders and then using the step-wise regression analysis model in batches to further refine the dataset. The second method will be to use correlation analysis to refine the

variables, looking for the variables with the highest correlation to the product being investigated.

Variable refinement using stepwise regression analysis

The first variable refinement method will be the 'stepwise' method, implemented when running regression models in SAS studio® software. The procedure used will be to build the model up in groups of 50 variables at a time. Running the first 50 variables, only keeping the significant variables and adding the next 50 and continuing on like this through all the variables.

In SAS studio® software there are two variations of the standard regression analysis that were considered to reduce the effects of overfit. These two methods are forward-selection and the forward-selection stepwise method. The forward-selection method starts with no variables in the model and calculates the F-statistics for each variable that reflects its contribution to the model if that variable is included.

The p-values or significance of the F-statistics are compared to a value the researcher specifies in the model. Once the significant F-statistic is defined the variable that contributes most to the model are added. The model will continue running or iterating and adding variables with the largest F-statistic to the model, until no variable's F-statistics have a significance level greater than the defined cut-off level. Once a variable is added to the model it stays in the model. This method helps reduce the effect of overfit by evaluating the variables for significance before adding them to the model.

The other method is the forward selection stepwise method, this is a modification of the forward selection technique. Where the forward selection technique iterated through the variables remaining outside the model and not updating the variables included in the model, the stepwise method reviews the variables both inside the model and outside at each iteration, the stepwise method starts with an empty model and adds variables to the model based on calculating each variable's F-statistic reflecting that variable's contribution to the model. As with the forward selection using a defined significance level as a cut-off for inclusion in the model, the method iterates through the variables evaluating the F-statistic of both the variables already included in the model and those that are not. Variables already introduced in the model can be taken out or deleted if their significance level is below the defined cut-off. The model will stop when all the variables inside the model have a significance level greater than the cut-off and all the variables outside the model are not significant.

SAS studio® software provides other methods for example, the Maximum R^2 Improvement (MAXR) that similarly iterates through all the variables but adds the variable that has the largest increase in the model's R^2. This method and many others were not considered for this investigation because the model suffers from overfit and multicollinearity. The stepwise method is used as a refinement because it looks at the independent variable's level of significance, questioning; does this variable have a significant relationship with the dependent variable? Investigating and defining the group of independent variables that have a significant relationship with the dependent variable is the objectives of this analysis, this will define the competitor set.

The stepwise approach has been criticised because while it finds independent variables with significant relationships, it does not have a meaningful model selection criterion. The stepwise method does not evaluate all possible solutions, it only evaluates the best solution that will be optimal at that step [61]. Meaning that the model does not consider all the independent variable at once and then find an optimal

solution, it steps through the independent variables considering each one for the model as the model is at that time. The second independent variable will be considered for a model that possible only has the first independent variable included and so the third independent variable will be considered for a model that possible only has the first and second independent variables and so on until the model has stepped through all the independent variables. It is for this reason all the independent variables selected in each batch or addition of 50 independent variables will be used in the final model.

Another reason the stepwise method has been selected is because this gives the opportunity for independent variables to be taken out the model and added back, making the sequence in which a variable is evaluated for inclusion in the model irrelevant. In Figure A.5 in the appendix, the SAS studio® software code is displayed for running the first iteration for RTD-CIDER-P18-660ML. Once the model has run, the result will display the variables that have been kept in the model. These variables will be kept for the next iteration and the next 50 variables for evaluation will be added and the model run again. This process will be repeated until all the variables have had a chance for inclusion in the model. The model is iterating though all the variables and looking for those that add a significant relationship in this way the model is defining the product's competitor set.

As per the price elasticity regression models defined, both the dependent and independent variables need to first be transformed by their natural logarithm before being used in the model described in Figure A.5 in the appendix.

The results of the iterations of the 4 products of batches of 50 variables being added at a time are tabulated in Tables A.64 to A.68 in the appendix, Table 4.23 displays the results for SPIRITS-LIQUEURS-P16-750ML. After the two-step process has been computed to counter overfit, RTD-CIDER-P18-660ML will have 62 independent variables for inclusion in the final regression model when counting the unique products from the 6 batches. The final model for RTD-CIDER-P6-330ML's will have 45 unique variables from the 5 batches, SPIRITS-LIQUEURS-P16-750ML with 49 unique variables for inclusion from the 4 batches. Finally WINE-RED-P102-750ML with 63 unique variables from the 8 batches in the final regression model.

SPIRITS-LIQUEURS-P16-750ML Significant Independent Variables		
Product	Product	Product
Results Batch 1	**Results Batch 3**	**Results Batch 4**
SPIRITS-LIQUEURS-P16-750ML	SPIRITS-LIQUEURS-P16-750ML	SPIRITS-LIQUEURS-P16-750ML
WINE-RED-P22-750ML	SPIRITS-LIQUEURS-P2-1LTR	SPIRITS-LIQUEURS-P2-1LTR
SPIRITS-LIQUEURS-P2-1LTR	SPIRITS-RUM-P5-750ML	SPIRITS-RUM-P5-750ML
SPIRITS-RUM-P5-750ML	SPIRITS-BRANDY-P24-750ML	RTD-BEER-P44-500ML
WINE-RED-P23-750ML	RTD-BEER-P44-500ML	RTD-BEER-P24-340ML
SPIRITS-BRANDY-P24-750ML	RTD-BEER-P24-340ML	WINE-BIB-P48-3LTR
SPIRITS-GIN-P7-750ML	SPIRITS-BRANDY-P19-750ML	SPIRITS-WHISKY-P20-750ML
SPIRITS-RUM-P1-1LTR	SPIRITS-VODKA-P18-750ML	SPIRITS-LIQUEURS-P9-750ML
RTD-BEER-P44-500ML	WINE-BIB-P48-3LTR	WINE-SPARKLING-P31-750ML
RTD-BEER-P58-750ML	SPIRITS-WHISKY-P20-750ML	SPIRITS-BRANDY-P5-375ML
RTD-BEER-P24-340ML	SPIRITS-GIN-P2-200ML	SPIRITS-BRANDY-P22-750ML
SPIRITS-BRANDY-P19-750ML	SPIRITS-GIN-P3-375ML	WINE-RED-P96-750ML
SPIRITS-VODKA-P18-750ML	SPIRITS-LIQUEURS-P9-750ML	WINE-RED-P99-750ML
WINE-BIB-P48-3LTR	WINE-SPARKLING-P31-750ML	SPIRITS-RUM-P6-750ML
	SPIRITS-BRANDY-P5-375ML	SPIRITS-LIQUEURS-P11-750ML
Results Batch 2	SPIRITS-BRANDY-P22-750ML	SPIRITS-BRANDY-P23-750ML
SPIRITS-LIQUEURS-P16-750ML	WINE-RED-P96-750ML	WINE-BIB-P44-3LTR
SPIRITS-LIQUEURS-P2-1LTR	WINE-RED-P99-750ML	WINE-BIB-P40-3LTR
SPIRITS-RUM-P5-750ML	SPIRITS-RUM-P6-750ML	SPIRITS-VODKA-P9-750ML
SPIRITS-BRANDY-P24-750ML	SPIRITS-LIQUEURS-P11-750ML	RTD-CIDER-P10-330ML
RTD-BEER-P44-500ML	SPIRITS-BRANDY-P2-1LTR	WINE-FORTIFIED-P3-200ML
RTD-BEER-P58-750ML	SPIRITS-LIQUEURS-P13-750ML	SPIRITS-LIQUEURS-P8-750ML
RTD-BEER-P24-340ML	SPIRITS-BRANDY-P4-1LTR	RTD-FABS-P1-250ML
SPIRITS-BRANDY-P19-750ML	SPIRITS-BRANDY-P23-750ML	RTD-FABS-P23-300ML
SPIRITS-VODKA-P18-750ML	WINE-BIB-P44-3LTR	WINE-WHITE-P18-750ML
WINE-BIB-P48-3LTR	WINE-BIB-P40-3LTR	SPIRITS-WHISKY-P19-750ML
SPIRITS-WHISKY-P20-750ML	SPIRITS-VODKA-P9-750ML	SPIRITS-BRANDY-P13-750ML
SPIRITS-GIN-P2-200ML	RTD-CIDER-P10-330ML	WINE-RED-P24-750ML
SPIRITS-GIN-P3-375ML		SPIRITS-BRANDY-P3-1LTR
RTD-FABS-P6-250ML		SPIRITS-WHISKY-P48-750ML
RTD-FABS-P22-275ML		
SPIRITS-WHISKY-P23-750ML		
SPIRITS-LIQUEURS-P9-750ML		
WINE-SPARKLING-P31-750ML		
SPIRITS-WHISKY-P52-750ML		
SPIRITS-WHISKY-P49-750ML		
SPIRITS-WHISKY-P25-750ML		

TABLE 4.23: *SPIRITS-LIQUEURS-P16-750ML significant independent variables results as each batch of independent variables are added*

Variable refinement using correlation analysis

The price elasticity methodology is a well known method for determining and quantifying relationships between different products in a market. As shown above overfit is a serious problem when using regression analysis on the alcoholic beverage industry due to the vast variety of products available. Correlation

analysis is another method that can be used to define the relationship that the different products on the market have with each other.

The PROC CORR procedure in SAS studio® software can be used to compute the correlation coefficients for each product being investigated using its quantity demanded and the price of all the other products on the market. As with regression analysis both the quantity demanded and the price of all the other products on the market will be transformed using the natural logarithm, this is to standardise the dataset. In Figure A.6 in the appendix the SAS code is displayed for calculating the correlation coefficients for a product on the market relative to the products being investigated. The results can be ranked, the highest absolute values showing the strongest relationships. Only correlation coefficients with a p-value greater than 0.05 are significant and should be used.

In Figure 4.23 the distribution of the results for the correlation coefficients between RTD-CIDER-P18-660ML and all other products on the market are displayed. This distribution shows that there are 137 products that have an absolute value of correlation that is greater than 0.5, correlation coefficients that are greater than 0.5 is an indication of a strong relationship [46]. The results of the stepwise method for RTD-CIDER-P18-660ML and correlation analysis are shown in Tables A.69 and A.70 in the appendix, this showing that the two methods provide some overlap in the results but not the same sub-set.
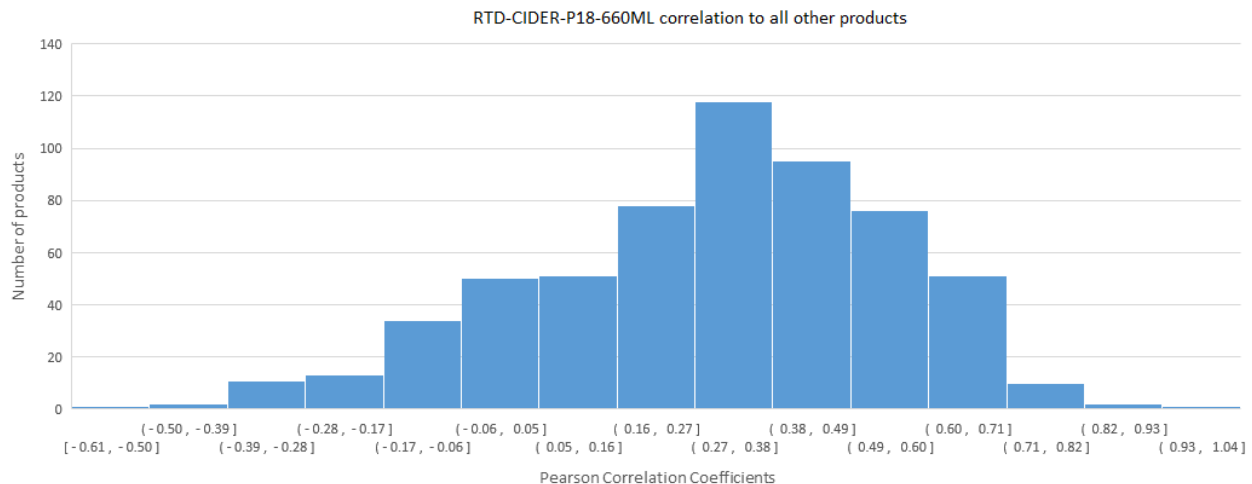
FIGURE 4.23: *Histogram showing the distribution of the correlation coefficients of all products relative to RTD-CIDER-P18-660ML*

In Figure 4.24 the distribution of the results for the correlation coefficients between RTD-CIDER-P6-330ML and all other products on the market are displayed. This distribution shows that there are 248 products that have an absolute value of correlation that is greater than 0.5, similarly in Figure 4.25 and Figure 4.26 the results are displayed for SPIRITS-LIQUEURS-P16-750ML and WINE-RED-P102-750ML showing that respectively they have a distribution of 225 and 100 products that have an absolute value of correlation that is greater than 0.5.
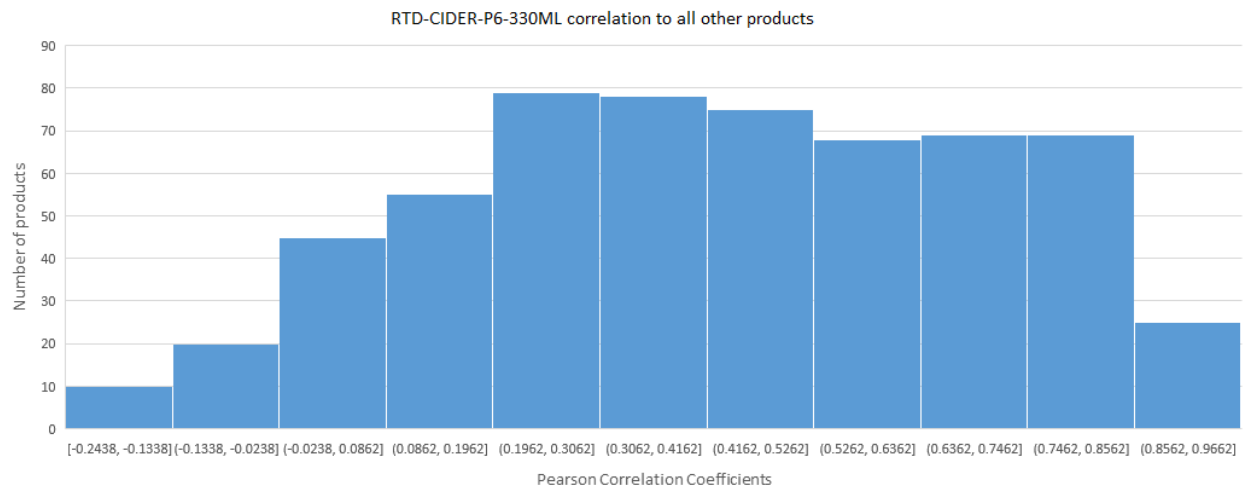
FIGURE 4.24: *Histogram showing the distribution of the correlation coefficients of all products relative to RTD-CIDER-P6-330ML*
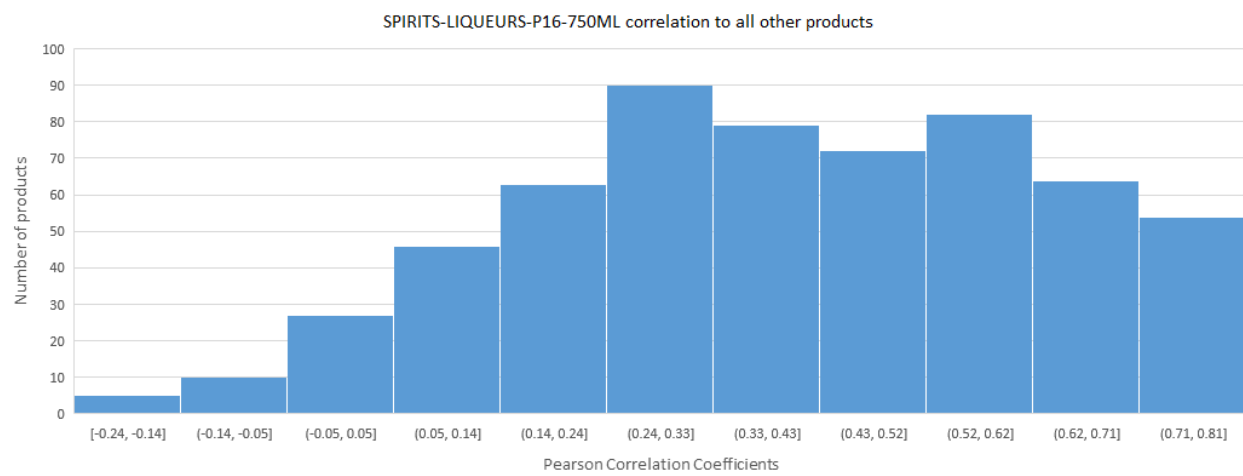
FIGURE 4.25: *Histogram showing the distribution of the correlation coefficients of all products relative to SPIRITS-LIQUEURS-P16-750ML*
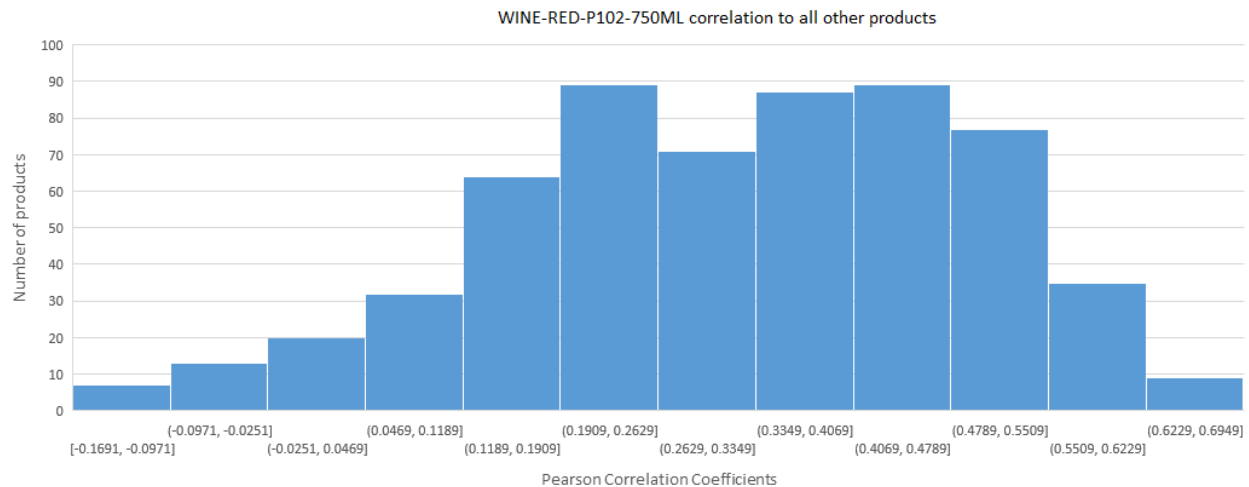
FIGURE 4.26: *Histogram showing the distribution of the correlation coefficients of all products relative to WINE-RED-P102-750ML*

The results of the stepwise method for RTD-CIDER-P18-660ML, RTD-CIDER-P6-330ML and WINE-RED-P102-750ML and their correlation coefficients are shown in TablesA.69 to A.73 in the appendix, the results for SPIRITS-LIQUEURS-P16-750ML are displayed below in Table 4.24. The results show that the two methods provide some overlap but do not produce the same sub-set of products.

Correlation analysis is a very simple solution to refining the variables but due to the highly correlated nature of the alcoholic beverage landscape, the results are not refined enough to use as they produce large sub-sets. In the stepwise method the results are tested for significance and the methodology is to determine the group of variables that have an effect on the dependent variables' quantity demanded [45]. In the stepwise method RTD-CIDER-P18-660ML will have 62 independent variables for inclusion in the final regression model while the correlation analysis provides 137 independent variables. Similarly for RTD-CIDER-P6-330ML, SPIRITS-LIQUEURS-P16-750ML and WINE-RED-P102-750ML with the stepwise method defines 45, 49 and 63 independent variables for inclusion in the final regression model respectively and the correlation analysis finds sub-sets of 137, 248 and 100 independent variables respectively.

The results from the stepwise method will be used for the final model as the price elasticity methodology using regression analysis is found to be recommended in literature [45]. Also the result set is smaller and therefore does not need to be refined.

In Figure 4.24 the results are shown for the significant independent from the stepwise method for product SPIRITS-LIQUEURS-P16-750ML and the rankings of the correlation coefficient by absolute strength. The products that had the highest correlations, ranks one to seven are not in this subset as these products did not have a significant relationship in the stepwise regressions analysis.

Product	Pearson Correlation Coefficients	p-value	Correlation Coefficients Ranking
SPIRITS-LIQUEURS-P11-750ML	0.78451	0.0001	8
SPIRITS-RUM-P6-750ML	0.76871	0.0001	13
SPIRITS-LIQUEURS-P9-750ML	0.75204	0.0001	23
SPIRITS-BRANDY-P24-750ML	0.74871	0.0001	26
WINE-SPARKLING-P31-750ML	0.73537	0.0001	39
SPIRITS-LIQUEURS-P2-1LTR	0.7325	0.0001	41
SPIRITS-LIQUEURS-P13-750ML	0.71971	0.0001	48
SPIRITS-RUM-P5-750ML	0.6966	0.0001	63
RTD-FABS-P6-250ML	0.69406	0.0001	65
SPIRITS-WHISKY-P19-750ML	0.69253	0.0001	67
SPIRITS-VODKA-P18-750ML	0.68689	0.0001	69
RTD-BEER-P24-340ML	0.66387	0.0001	87
SPIRITS-VODKA-P9-750ML	0.64793	0.0001	100
SPIRITS-BRANDY-P19-750ML	0.64222	0.0001	102
SPIRITS-BRANDY-P2-1LTR	0.63919	0.0001	104
SPIRITS-BRANDY-P3-1LTR	0.63515	0.0001	109
SPIRITS-BRANDY-P23-750ML	0.59963	0.0001	134
WINE-BIB-P40-3LTR	0.57993	0.0001	152
SPIRITS-GIN-P3-375ML	0.57706	0.0001	154
SPIRITS-WHISKY-P23-750ML	0.55653	0.0001	168
SPIRITS-GIN-P7-750ML	0.54124	0.0001	188
RTD-BEER-P58-750ML	0.52429	0.0001	199
SPIRITS-WHISKY-P52-750ML	0.5177	0.0001	203
SPIRITS-BRANDY-P22-750ML	0.51075	0.0001	209
SPIRITS-LIQUEURS-P8-750ML	0.50892	0.0001	210
SPIRITS-WHISKY-P25-750ML	0.4858	0.0001	233
WINE-BIB-P44-3LTR	0.45591	0.0001	253
RTD-FABS-P23-300ML	0.39859	0.0001	301
SPIRITS-BRANDY-P4-1LTR	0.39513	0.0001	304
SPIRITS-WHISKY-P49-750ML	0.39333	0.0001	306
WINE-WHITE-P18-750ML	0.3825	0.0001	316
RTD-FABS-P1-250ML	0.37172	0.0001	326
SPIRITS-WHISKY-P48-750ML	0.35975	0.0001	333
RTD-CIDER-P10-330ML	0.34036	0.0001	349
SPIRITS-GIN-P2-200ML	0.33386	0.0001	352
WINE-RED-P96-750ML	0.33156	0.0001	353
WINE-RED-P22-750ML	0.32838	0.0001	358
SPIRITS-BRANDY-P5-375ML	0.32043	0.0001	373
WINE-RED-P23-750ML	0.30865	0.0001	379
WINE-RED-P99-750ML	0.28476	0.0003	393
WINE-RED-P24-750ML	0.24525	0.0020	431
RTD-FABS-P22-275ML	0.24397	0.0021	434
SPIRITS-WHISKY-P20-750ML	0.23833	0.0026	442
SPIRITS-BRANDY-P13-750ML	0.21521	0.0068	459
WINE-BIB-P48-3LTR	0.20623	0.0096	468
RTD-BEER-P44-500ML	0.19368	0.0151	473
WINE-FORTIFIED-P3-200ML	0.1857	0.0199	482
SPIRITS-RUM-P1-1LTR	0.17745	0.0262	488

TABLE 4.24: *SPIRITS-LIQUEURS-P16-750ML significant independent variables results from the stepwise method ranked by correlation coefficient strength*

4.4.6   Step 6: Final model

An independent variable is kept for the final model if it was in the result set of significant variables for any batch, not just the final batch from the stepwise variable refinement method. The list of possible independent variables or competitor products have now been refined to those that have significant relationships in the stepwise model.

Remembering that the objective is to find the products that have a relationship with the product under investigation, it is for this reason that the independent variables will be modelled separately and to overcome multicollinearity. If the objective was to investigate how all the independent variables affect the dependent variables simultaneously then a different model would be needed which included all the independent variables simultaneously. The objective here however is much simpler, just to establish whether there is a relationship with the dependent variable irrespective of other independent variables. Also by running the model separately the highly correlated variables are not excluded from the final result set.

The models that will be computed will have the dependent variable's own price as an independent variable, the price of the dependent variable is also included as an independent variable in each model for the same reason; to provide stability to the model and increase the chances of the overall model being significant. It is assumed that each product would have a significant relationship with its own price. The models will each include one independent variable being tested from the results of the stepwise method as independent variables.

As noted earlier in the introduction, the alcoholic beverage industry is highly seasonal. It is for this reason that seasonal dummy variables have been added to each model, to give each model the best possibility of being significant as a model, so the individual variables' relationships can be studied and results trusted. The dummy variables will be binary and represent the different months of the year.

The price elasticity model from equation (2.21) will be used to define the equations for the individual price elasticity regression models. As an example, starting with RTD-CIDER-P18-660ML and the independent variable WINE-BIB-P79-1.5LTR, let the volume sales per week in litres be Q_{RCP18} and the average unit price be P_{RCP18}. Let β_{RCP18} be the own elasticity of RTD-CIDER-P18-660ML and β_{WBP79} the cross price elasticity of the change in RTD-CIDER-P18-660ML's demand as a response to the change in WINE-BIB-P79-1.5LTR [81], all other variables being equal. Let D_1 represent the month of January and D_2 the month of February, continuing to D_{11} representing November. The month of December is not included in the model as it is the reference variable for the dummy variables. Then,

$$\ln Q_{RCP18} = \beta_0 + \beta_{RCP18} \ln P_{RCP18} + \beta_{WBP79} \ln P_{WBP79} + \beta_{d1} D_1 + \beta_{d2} D_2 \ldots + \beta_{d11} D_{11} + \epsilon \,. \quad (4.5)$$

In Figure A.7 in the appendix the SAS studio® software code for multiple regression analysis of price elasticity is displayed for RTD-CIDER-P18-660ML and the first independent variable in the set.

Each model will have 13 independent variables so overfit should not be a problem any more. As noted earlier the dependent and independent variables have been transformed using the natural logarithm so heteroscedasticity should also not be a problem in the model. As only two non-binary dummy independent variables are used, multicollinearity should also not affect the model.

4.5 Results for the price elasticity model to determine a competitor set

The objective of this investigation is to group a selection of products that all meet the same consumer need as the product under investigation. This process has required many methods to refine the set of independent variables under consideration, the final models are simplified. The results for the individual regression models are displayed per independent variable in Tables 4.25 to 4.31.

4.5.1 Output statistics from regression analysis

The results are given for the dependent variable under investigation, evaluating the relationship between the dependent variable and the independent variables. The results for the own price elasticity of the dependent variable is not shown for each model, due to the simplistic nature of the models, analysing each independent variable separately. The results for the binary seasonal dummy variables are not displayed as they are not important to the research objective.

In Tables 4.25 to 4.31 the first observation is that all the models with a significant F-statistic and p-values well below the cut-off value of 0.05. Therefore, indicating that each model explains a significant portion of the variation in the dependent variables under investigation.

The adjusted R^2 should be used as an indication of model adequacy as these are multiple regression models. Models with low adjusted R^2's should be treated with caution. The low adjusted R^2 is a warning that although the model might give significant results, the independent variables only predict a small amount of variation or change in the dependent variable's behaviour. In Tables 4.25 to 4.31 the R^2 and adjusted R^2 for each model are displayed. The adjusted R^2's in the results for RTD-CIDER-P18-660ML range from 0.558 to 0.735, for RTD-CIDER-P6-330ML from 0.531 to 0.643, for SPIRITS-LIQUEURS-P16-750ML range from 0.671 to $0,782$ and finally for WINE-RED-P102-750ML the range is from 0.299 to 0.531.

For each product being investigated, the difference in each model's adjusted R^2 is due to the different independent variables that are being tested. The range of the adjusted R^2 for each of the 4 products should be investigated, for RTD-CIDER-P18-660ML, RTD-CIDER-P6-330ML and SPIRITS-LIQUEURS-P16-750ML the adjusted R^2 ranges are all with in a narrow field and also relative to each other. While the WINE-RED-P102-750ML price elasticity models have a range of adjusted R^2 values between 0.4013 to 0.531, there is one outlier. WINE-SPARKLING-P25-750ML has an adjusted R^2 value of 0.2997, this is an indication of an ill fitting model and this model should be excluded from the final results.

Product	F Value	Pr > F	R^2	AdjR^2	AdjR^2 Ranking	Durbin-Watson D	Parameter Estimate	Standard Error	t-Value	Pr > |t|	Parameter Estimate Ranking
WINE-RED-P79-1.5LTR	25.86	<.0001	0.7016	0.6745	10	**1.265**	4.2225	0.5899	7.16	<.0001	1
SPIRITS-WHISKY-P8-1LTR	28.16	<.0001	0.7191	0.6935	5	**1.266**	3.9994	0.5026	7.96	<.0001	2
SPIRITS-VODKA-P1-1LTR	22.99	<.0001	0.6764	0.647	21	**1.186**	3.8627	0.6429	6.01	<.0001	3
RTD-BEER-P50-660ML	34.34	<.0001	0.7574	0.7353	1	1.53	3.8409	0.3923	9.79	<.0001	4
RTD-CIDER-P18-660ML	17.61	<.0001	0.5947	0.5609	50	**0.86**	-3.7642	0.4167	-9.03	<.0001	Exclude
SPIRITS-WHISKY-P9-1LTR	30.4	<.0001	0.7343	0.7101	3	**1.271**	3.3331	0.3846	8.67	<.0001	5
RTD-CIDER-P16-500ML	21.6	<.0001	0.6626	0.632	24	**1.158**	3.3279	0.6202	5.37	<.0001	6
RTD-CIDER-P9-330ML	26.16	<.0001	0.704	0.6771	8	**1.206**	3.3031	0.4545	7.27	<.0001	7
WINE-BIB-P74-5LTR	24.95	<.0001	0.694	0.6662	15	**1.195**	3.2574	0.4781	6.81	<.0001	8
RTD-BEER-P55-750ML	26.05	<.0001	0.7031	0.6762	9	**1.156**	3.2503	0.4497	7.23	<.0001	9
WINE-BIB-P18-2LTR	29.46	<.0001	0.7281	0.7034	4	**1.356**	3.1256	0.3732	8.38	<.0001	10
RTD-FABS-P27-330ML	25.57	<.0001	0.6992	0.6719	11	**1.159**	3.0828	0.4373	7.05	<.0001	11
WINE-BIB-P69-5LTR	18.97	<.0001	0.6329	0.5996	30	**1.03**	2.9589	0.7665	3.86	0.0002	12
WINE-BIB-P21-2LTR	31.31	<.0001	0.74	0.7164	2	**1.38**	2.9284	0.3276	8.94	<.0001	13
RTD-BEER-P39-440ML	17.68	<.0001	0.6164	0.5816	39	**0.999**	-2.9284	1.0281	-2.85	0.005	Exclude
RTD-BEER-P57-750ML	23.41	<.0001	0.6803	0.6512	18	**1.017**	2.7886	0.4507	6.19	<.0001	14
WINE-FORTIFIED-P1-1LTR	24.16	<.0001	0.6872	0.6588	17	**1.14**	2.7479	0.4226	6.5	<.0001	15
RTD-CIDER-P15-440ML	20.38	<.0001	0.6495	0.6176	28	**1.004**	2.7437	0.5802	4.73	<.0001	16
RTD-CIDER-P17-660ML	16.41	<.0001	0.5986	0.5622	49	0.882	-2.7104	2.2813	-1.19	**0.2368**	Exclude
WINE-BIB-P19-2LTR	24.57	<.0001	0.6908	0.6627	16	**1.282**	2.5805	0.3871	6.67	<.0001	17
SPIRITS-BRANDY-P1-1LTR	18.4	<.0001	0.6258	0.5918	62	**0.861**	2.4181	0.7008	3.45	0.0007	Exclude
RTD-BEER-P29-440ML	19.01	<.0001	0.6334	0.6001	29	**0.992**	-2.3852	0.6137	-3.89	0.0002	Exclude
WINE-BIB-P66-5LTR	21.56	<.0001	0.6622	0.6315	25	**1.147**	2.2239	0.416	5.35	<.0001	18
RTD-BEER-P58-750ML	18.91	<.0001	0.6322	0.5988	31	**0.934**	1.75	0.4582	3.82	0.0002	Exclude
WINE-BIB-P30-3LTR	22.94	<.0001	0.6759	0.6464	22	**1.211**	1.5805	0.2641	5.99	<.0001	19
RTD-FABS-P11-275ML	17.83	<.0001	0.6184	0.5837	37	**1.006**	1.536	0.515	2.98	0.0034	20
RTD-BEER-P18-330ML	17.65	<.0001	0.616	0.5811	40	**0.927**	1.5176	0.5383	2.82	0.0055	Exclude
SPIRITS-WHISKY-P53-750ML	21.07	<.0001	0.657	0.6258	27	**1.194**	1.4243	0.2796	5.1	<.0001	21
RTD-BEER-P3-330ML	17.75	<.0001	0.6174	0.5826	38	**0.793**	-1.2812	0.4397	-2.91	0.0041	Exclude
RTD-BEER-P8-330ML	17.32	<.0001	0.6116	0.5763	42	**0.904**	1.1401	0.4563	2.5	0.0136	Exclude
RTD-FABS-P37-660ML	17.33	<.0001	0.6117	0.5764	41	**0.899**	-1.0634	0.4249	-2.5	0.0134	Exclude
RTD-BEER-P22-340ML	16.6	<.0001	0.6014	0.5652	46	0.858	-1.0612	0.6844	-1.55	**0.1232**	Exclude
RTD-BEER-P2-330ML	27.58	<.0001	0.7148	0.6889	6	**1.175**	0.7513	0.0968	7.76	<.0001	22
WINE-BIB-P79-5LTR	17.98	<.0001	0.6204	0.5859	36	**0.977**	0.6924	0.2226	3.11	0.0022	Exclude
RTD-BEER-P12-330ML	16.27	<.0001	0.5967	0.56	52	0.899	-0.5487	0.6536	-0.84	**0.4026**	Exclude
WINE-BIB-P35-3LTR	17.26	<.0001	0.6107	0.5753	43	**0.955**	0.5455	0.2248	2.43	0.0165	Exclude
WINE-PERLE-P4-2LTR	16.57	<.0001	0.601	0.5648	48	0.921	0.4571	0.3032	1.51	**0.1339**	Exclude
SPIRITS-LIQUEURS-P16-750ML	17.06	<.0001	0.608	0.5724	44	**0.829**	-0.4384	0.199	-2.2	0.0292	Exclude
RTD-BEER-P19-330ML	16.6	<.0001	0.6014	0.5652	47	0.884	-0.3146	0.2023	-1.56	**0.122**	Exclude
RTD-BEER-P14-330ML	16.18	<.0001	0.5953	0.5585	53	0.862	-0.2671	0.5595	-0.48	**0.6338**	Exclude
RTD-BEER-P1-330ML	16.18	<.0001	0.5953	0.5585	54	0.866	0.2278	0.4947	0.46	**0.6459**	Exclude
RTD-BEER-P31-440ML	18.7	<.0001	0.6296	0.596	32	**0.983**	0.2217	0.0604	3.67	0.0003	Exclude
WINE-BIB-P56-5LTR	16.3	<.0001	0.597	0.5604	51	0.871	-0.2184	0.2403	-0.91	**0.365**	Exclude

TABLE 4.25: *Part 1: The results of the multiple regression price elasticity models with RTD-CIDER-P18-660ML as the dependent variable*

Product	F Value	Pr > F	R^2	AdjR^2	AdjR^2 Ranking	Durbin-Watson D	Parameter Estimate	Standard Error	t-Value	Pr > |t	Parameter Estimate Ranking
SPIRITS-WHISKY-P11-750ML	16.15	<.0001	0.5949	0.5581	56	0.853	-0.1363	0.4795	-0.28	**0.7768**	Exclude
RTD-BEER-P16-330ML	16.15	<.0001	0.5949	0.5581	57	0.859	0.1286	0.4692	0.27	**0.7844**	Exclude
RTD-BEER-P56-750ML	25.33	<.0001	0.6973	0.6697	13	**1.074**	0.116	0.0167	6.96	<.0001	23
RTD-BEER-P41-440ML	16.14	<.0001	0.5948	0.5579	61	0.859	0.1131	0.7372	0.15	**0.8783**	Exclude
RTD-BEER-P43-500ML	22.84	<.0001	0.6749	0.6454	23	**0.972**	0.1109	0.0187	5.94	<.0001	Exclude
RTD-BEER-P34-440ML	16.15	<.0001	0.5949	0.5581	55	**0.871**	0.1073	0.4007	0.27	**0.7893**	Exclude
RTD-FABS-P38-660ML	21.31	<.0001	0.6595	0.6286	26	**1.081**	0.1057	0.0203	5.22	<.0001	24
RTD-CIDER-P14-440ML	25.37	<.0001	0.6976	0.6701	12	**1.079**	0.0933	0.0134	6.98	<.0001	25
RTD-CIDER-P3-330ML	25.31	<.0001	0.6971	0.6696	14	**1.071**	0.0916	0.0132	6.95	<.0001	26
RTD-FABS-P9-275ML	23.13	<.0001	0.6777	0.6484	20	**0.985**	0.0792	0.0131	6.07	<.0001	Exclude
RTD-FABS-P4-250ML	26.21	<.0001	0.7043	0.6775	7	**1.14**	0.0693	0.0096	7.28	<.0001	27
WINE-BIB-P55-5LTR	16.15	<.0001	0.5948	0.558	58	0.862	0.0629	0.306	0.21	**0.8376**	Exclude
RTD-BEER-P47-500ML	18.34	<.0001	0.6251	0.591	34	**0.867**	0.0622	0.0183	3.41	0.0009	Exclude
RTD-BEER-P28-440ML	18.21	<.0001	0.6234	0.5892	35	**0.885**	0.0608	0.0184	3.3	0.0012	Exclude
RTD-BEER-P44-500ML	18.54	<.0001	0.6276	0.5937	33	**0.884**	0.0599	0.0169	3.55	0.0005	Exclude
RTD-FABS-P3-250ML	23.37	<.0001	0.6799	0.6508	19	**0.97**	0.0494	0.008	6.17	<.0001	Exclude
WINE-WHITE-P1-5LTR	16.14	<.0001	0.5947	0.5579	59	0.855	-0.0487	0.3488	-0.14	**0.8892**	Exclude
WINE-ROSE-P2-1.5LTR	16.14	<.0001	0.5947	0.5579	60	0.857	-0.0329	0.3475	-0.09	**0.9247**	Exclude
RTD-BEER-P48-5LTR	16.87	<.0001	0.6053	0.5694	45	0.865	0.0272	0.0139	1.96	**0.0521**	Exclude

TABLE 4.26: *Part 2: The results of the multiple regression price elasticity models with RTD-CIDER-P18-660ML as the dependent variable*

In Table 4.25 to Table 4.31 the p-values of each independent variable in the models are shown. The level of significance of 0.05 will be used as a cut-off. All independent variables with corresponding p-values greater than 0.05 indicates that the null hypotheses states that the parameters are not significantly different from zero. The corresponding p-values greater than 0.05 of the independent variable have been highlighted in Table 4.25 to Table 4.31, these products should be excluded from the final results as they do not have a statistically significant relationship with the dependent variable.

Product	F Value	Pr > F	R^2	AdjR^2	AdjR^2 Ranking	Durbin-Watson D	Parameter Estimate	Standard Error	t-Value	Pr > |t|	Parameter Estimate Ranking
RTD-BEER-P24-340ML	20.73	<.0001	0.6533	0.6218	2	1.498	-2.9378	0.5027	-5.84	<.0001	1
RTD-BEER-P3-330ML	22.7	<.0001	0.6736	0.6439	1	1.554	-2.8889	0.4301	-6.72	<.0001	2
RTD-FABS-P7-275ML	18.24	<.0001	0.6238	0.5896	10	1.692	2.4271	0.5393	4.5	<.0001	3
RTD-FABS-P30-330ML	20.33	<.0001	0.6489	0.6169	5	1.707	2.2355	0.3959	5.65	<.0001	4
RTD-FABS-P20-275ML	18.74	<.0001	0.6301	0.5965	8	1.652	2.0302	0.4232	4.8	<.0001	5
RTD-FABS-P27-330ML	20.71	<.0001	0.6531	0.6215	3	1.706	1.9628	0.3365	5.83	<.0001	6
SPIRITS-LIQUEURS-P1-1LTR	18.04	<.0001	0.6212	0.5868	11	1.574	1.821	0.4163	4.37	<.0001	7
RTD-BEER-P58-750ML	19.23	<.0001	0.6361	0.6031	7	1.667	1.739	0.3425	5.08	<.0001	8
RTD-BEER-P20-340ML	16.87	<.0001	0.6053	0.5694	14	1.561	1.6936	0.4774	3.55	0.0005	9
SPIRITS-LIQUEURS-P11-750ML	15.87	<.0001	0.5906	0.5534	19	**1.347**	-1.2561	0.4748	-2.65	0.0091	10
WINE-SPARKLING-P7-750ML	16.18	<.0001	0.5953	0.5585	18	1.531	1.2197	0.4127	2.96	0.0037	11
RTD-BEER-P34-440ML	15.3	<.0001	0.5818	0.5438	25	**1.377**	-0.9347	0.4764	-1.96	0.0517	12
RTD-CIDER-P6-330ML	15.94	<.0001	0.5705	0.5348	35	**1.419**	-0.9081	0.4352	-2.09	0.0387	13
RTD-FABS-P1-250ML	16.94	<.0001	0.6063	0.5705	13	1.555	0.8822	0.2448	3.6	0.0004	14
RTD-BEER-P1-330ML	15	<.0001	0.5769	0.5384	31	1.457	0.8374	0.5727	1.46	**0.1459**	Exclude
RTD-BEER-P12-330ML	14.79	<.0001	0.5735	0.5347	36	1.414	0.7834	0.7852	1	**0.3201**	Exclude
SPIRITS-WHISKY-P32-750ML	15.21	<.0001	0.5802	0.5421	26	1.39	-0.7697	0.4236	-1.82	**0.0713**	Exclude
RTD-BEER-P39-440ML	14.68	<.0001	0.5716	0.5327	40	1.41	0.7317	1.2311	0.59	**0.5533**	Exclude
RTD-FABS-P11-275ML	14.98	<.0001	0.5767	0.5382	32	1.507	0.7128	0.4961	1.44	**0.153**	Exclude
WINE-SPARKLING-P10-750ML	15.81	<.0001	0.5897	0.5524	20	1.545	0.6858	0.2652	2.59	0.0107	15
SPIRITS-LIQUEURS-P12-750ML	15.09	<.0001	0.5784	0.5401	27	1.419	-0.6344	0.3878	-1.64	**0.1041**	Exclude
WINE-SPARKLING-P4-750ML	15.71	<.0001	0.5882	0.5508	21	1.575	0.6206	0.2507	2.48	0.0144	16
SPIRITS-WHISKY-P24-750ML	15.32	<.0001	0.582	0.544	24	1.516	0.5997	0.3027	1.98	0.0494	17
RTD-BEER-P21-340ML	14.76	<.0001	0.5731	0.5342	37	1.424	0.5974	0.6508	0.92	**0.3602**	Exclude
RTD-BEER-P27-440ML	14.84	<.0001	0.5743	0.5356	33	1.431	0.5719	0.5093	1.12	**0.2633**	Exclude
SPIRITS-LIQUEURS-P9-750ML	15.06	<.0001	0.5779	0.5396	29	1.379	-0.5566	0.3514	-1.58	**0.1154**	Exclude
WINE-SPARKLING-P31-750ML	15.03	<.0001	0.5774	0.539	30	1.355	-0.4589	0.3015	-1.52	**0.1302**	Exclude
RTD-BEER-P13-330ML	14.81	<.0001	0.5738	0.5351	34	1.401	-0.4554	0.4325	-1.05	**0.2942**	Exclude
RTD-CIDER-P2-330ML	14.63	<.0001	0.5709	0.5319	42	1.441	0.4216	1.2692	0.33	**0.7403**	Exclude
SPIRITS-LIQUEURS-P5-750ML	15.09	<.0001	0.5784	0.5401	28	1.348	-0.4115	0.2521	-1.63	**0.1048**	Exclude
RTD-BEER-P9-330ML	17.05	<.0001	0.6078	0.5721	12	1.507	-0.2945	0.08	-3.68	0.0003	18
SPIRITS-WHISKY-P45-750ML	14.73	<.0001	0.5725	0.5336	38	1.449	0.2637	0.3298	0.8	**0.4254**	Exclude
SPIRITS-WHISKY-P42-750ML	15.6	<.0001	0.5864	0.5488	23	**1.41**	0.204	0.0872	2.34	0.0206	19
RTD-FABS-P13-275ML	14.64	<.0001	0.5727	0.5335	39	1.415	-0.1238	0.3275	-0.38	**0.7061**	Exclude
RTD-BEER-P47-500ML	19.5	<.0001	0.6393	0.6065	6	1.64	0.0775	0.0149	5.22	<.0001	20
SPIRITS-WHISKY-P34-750ML	14.63	<.0001	0.5708	0.5318	43	1.419	0.0751	0.2475	0.3	**0.762**	Exclude
RTD-FABS-P6-250ML	16.65	<.0001	0.6022	0.566	16	**1.438**	0.0697	0.0207	3.37	0.001	21
RTD-FABS-P23-300ML	20.44	<.0001	0.6501	0.6183	4	1.644	0.0647	0.0114	5.7	<.0001	22
RTD-FABS-P33-440ML	15.64	<.0001	0.5871	0.5496	22	**1.428**	0.0606	0.0253	2.39	0.018	23
RTD-BEER-P48-5LTR	16.65	<.0001	0.6021	0.5659	17	1.51	0.0522	0.0155	3.37	0.001	24
RTD-FABS-P24-300ML	18.41	<.0001	0.626	0.592	9	1.535	0.0371	0.0081	4.6	<.0001	25
SPIRITS-WHISKY-P49-750ML	16.7	<.0001	0.6028	0.5667	15	1.463	0.0324	0.0095	3.41	0.0008	26
RTD-FABS-P38-660ML	14.67	<.0001	0.5714	0.5325	41	1.421	0.015	0.0274	0.55	**0.5855**	Exclude
RTD-BEER-P35-440ML	14.61	<.0001	0.5705	0.5315	45	1.42	0.0052	0.2881	0.02	**0.9859**	Exclude
SPIRITS-LIQUEURS-P7-750ML	14.62	<.0001	0.5706	0.5316	44	1.416	0.0045	0.0263	0.17	**0.8656**	Exclude

TABLE 4.27: *The results of the multiple regression price elasticity models with RTD-CIDER-P6-330ML as the dependent variable*

The parameter estimates in Table 4.25 to Table 4.31 are the $\hat{\beta}_i$ or cross price elasticities of the independent variables relative to the dependent variables. The parameter estimates are an indication of the strength of the relationship or lack of relationship and the parameter estimates sign an indication of the type of relationship.

The parameter estimates should not be used as actual price elasticities from these simple models, as the adjusted R^2 results showed that significant variables are missing from the models for example, promotions and other advertising. The coefficients are however comparable, as they are of the same scale and can be ranked to indicate the strength of their relationships relative to the dependent variable. Tables 4.25 to 4.31 are sorted according to the highest absolute parameter estimate coefficient, leaving out the products that have been excluded from the final set.

Product	F Value	Pr > F	R^2	AdjR^2	AdjR^2 Ranking	Durbin-Watson D	Parameter Estimate	Standard Error	t-Value	Pr > |t	Parameter Estimate Ranking
SPIRITS-BRANDY-P4-1LTR	38.16	<.0001	0.7763	0.7559	10	**1.417**	4.9706	0.7052	7.05	<.0001	1
SPIRITS-LIQUEURS-P2-1LTR	44.15	<.0001	0.8006	0.7824	1	1.528	4.4122	0.5159	8.55	<.0001	2
SPIRITS-BRANDY-P3-1LTR	35.68	<.0001	0.7644	0.743	14	**1.326**	4.4084	0.6972	6.32	<.0001	3
SPIRITS-RUM-P6-750ML	39.72	<.0001	0.7831	0.7634	7	1.513	4.1375	0.554	7.47	<.0001	4
SPIRITS-LIQUEURS-P16-750ML	27.8	<.0001	0.6985	0.6734	43	**1.07**	-3.7022	0.3095	-11.97	<.0001	5
RTD-FABS-P22-275ML	34.67	<.0001	0.7591	0.7372	17	1.48	3.6321	0.6053	6	<.0001	6
SPIRITS-BRANDY-P2-1LTR	40.29	<.0001	0.7855	0.766	6	1.468	3.3619	0.4414	7.62	<.0001	7
SPIRITS-BRANDY-P5-375ML	33.21	<.0001	0.7512	0.7286	19	**1.272**	3.1151	0.5661	5.5	<.0001	8
RTD-BEER-P24-340ML	28.72	<.0001	0.7231	0.6979	29	**1.042**	-2.5567	0.7179	-3.56	0.0005	9
SPIRITS-GIN-P2-200ML	41.72	<.0001	0.7914	0.7724	2	1.525	2.4944	0.3127	7.98	<.0001	10
RTD-BEER-P58-750ML	37.61	<.0001	0.7737	0.7531	12	1.455	2.4361	0.3535	6.89	<.0001	11
SPIRITS-WHISKY-P23-750ML	37.97	<.0001	0.7754	0.755	11	1.544	2.4338	0.3479	7	<.0001	12
SPIRITS-GIN-P3-375ML	38.38	<.0001	0.7772	0.757	9	**1.419**	2.2851	0.3214	7.11	<.0001	13
SPIRITS-RUM-P5-750ML	32.7	<.0001	0.7483	0.7254	20	1.492	2.133	0.4011	5.32	<.0001	14
SPIRITS-BRANDY-P24-750ML	32.18	<.0001	0.7452	0.7221	21	**1.325**	1.9209	0.375	5.12	<.0001	15
SPIRITS-RUM-P1-1LTR	41.12	<.0001	0.7889	0.7698	4	1.542	1.9095	0.244	7.83	<.0001	16
SPIRITS-WHISKY-P20-750ML	35.6	<.0001	0.764	0.7425	16	1.455	1.8514	0.294	6.3	<.0001	17
SPIRITS-WHISKY-P48-750ML	28.67	<.0001	0.7227	0.6975	30	**1.239**	1.7821	0.5043	3.53	0.0006	18
SPIRITS-VODKA-P9-750ML	35.63	<.0001	0.7641	0.7426	15	1.507	1.6667	0.2644	6.31	<.0001	19
SPIRITS-WHISKY-P19-750ML	29.16	<.0001	0.7261	0.7012	27	**1.26**	1.6473	0.4338	3.8	0.0002	20
SPIRITS-LIQUEURS-P11-750ML	28.47	<.0001	0.7213	0.696	31	**1.335**	1.5462	0.452	3.42	0.0008	21
SPIRITS-WHISKY-P25-750ML	28.15	<.0001	0.719	0.6935	33	**1.275**	1.4201	0.4396	3.23	0.0015	22
RTD-FABS-P1-250ML	31.3	<.0001	0.7399	0.7163	23	**1.267**	1.3587	0.2847	4.77	<.0001	23
WINE-BIB-P40-3LTR	29.69	<.0001	0.7297	0.7	28	**1.37**	1.2877	0.3171	4.06	<.0001	24
SPIRITS-VODKA-P18-750ML	26.92	<.0001	0.7099	0.6836	35	**1.197**	1.267	0.5338	2.37	0.0189	25
WINE-WHITE-P18-750ML	29.76	<.0001	0.7301	0.7056	24	**1.382**	1.254	0.3064	4.09	<.0001	26
WINE-FORTIFIED-P3-200ML	26.44	<.0001	0.7062	0.6795	37	1.084	1.1269	0.5822	1.94	**0.0549**	Exclude
SPIRITS-WHISKY-P52-750ML	28.18	<.0001	0.7192	0.6937	32	**1.227**	1.0328	0.3178	3.25	0.0014	27
WINE-BIB-P44-3LTR	29.7	<.0001	0.7297	0.7051	25	**1.381**	1.0264	0.2527	4.06	<.0001	28
SPIRITS-LIQUEURS-P13-750ML	26.47	<.0001	0.7064	0.6797	36	1.032	-1.0003	0.5087	-1.97	**0.0512**	Exclude
WINE-BIB-P48-3LTR	27.61	<.0001	0.7151	0.6892	34	**1.161**	0.9256	0.321	2.88	0.0045	29
SPIRITS-BRANDY-P23-750ML	26.2	<.0001	0.7043	0.6774	38	1.17	0.8779	0.5229	1.68	**0.0954**	Exclude
SPIRITS-BRANDY-P19-750ML	25.7	<.0001	0.7003	0.6731	44	1.142	0.6677	0.719	0.93	**0.3547**	Exclude
WINE-RED-P96-750ML	25.94	<.0001	0.7022	0.6751	40	1.155	0.5442	0.4083	1.33	**0.1847**	Exclude
WINE-RED-P99-750ML	25.91	<.0001	0.7019	0.6749	41	1.149	0.527	0.4096	1.29	**0.2003**	Exclude
WINE-RED-P24-750ML	26.13	<.0001	0.7038	0.6768	39	1.187	0.4575	0.2868	1.6	**0.1128**	Exclude
WINE-SPARKLING-P31-750ML	25.67	<.0001	0.7	0.6727	45	1.048	0.3688	0.4331	-0.85	**0.3959**	Exclude
WINE-RED-P22-750ML	25.79	<.0001	0.701	0.6738	42	1.125	0.316	0.2892	1.09	**0.2764**	Exclude
SPIRITS-GIN-P7-750ML	25.59	<.0001	0.6994	0.672	46	1.095	0.2533	0.3929	0.64	**0.5202**	Exclude
SPIRITS-BRANDY-P22-750ML	25.56	<.0001	0.6992	0.6718	48	1.079	0.2092	0.374	0.56	**0.5768**	Exclude
WINE-RED-P23-750ML	25.56	<.0001	0.6991	0.6718	47	1.075	0.1654	0.3001	0.55	**0.5824**	Exclude
SPIRITS-LIQUEURS-P8-750ML	29.31	<.0001	0.7271	0.7023	26	**1.09**	0.1246	0.0322	3.87	0.0002	30
RTD-CIDER-P10-330ML	41.64	<.0001	0.791	0.772	3	1.505	0.1055	0.0133	7.96	<.0001	31
RTD-BEER-P44-500ML	33.99	<.0001	0.7555	0.7333	18	**1.305**	0.097	0.0168	5.77	<.0001	32

TABLE 4.28: *Part 1: The results of the multiple regression price elasticity models with SPIRITS-LIQUEURS-P16-750ML as the dependent variable*

Product	F Value	Pr > F	R^2	AdjR^2	AdjR^2 Ranking	Durbin-Watson D	Parameter Estimate	Standard Error	t-Value	Pr > |t|	Parameter Estimate Ranking
RTD-FABS-P6-250ML	38.54	<.0001	0.7779	0.7578	8	**1.417**	0.0967	0.0136	7.15	<.0001	33
RTD-FABS-P23-300ML	37.34	<.0001	0.7725	0.7518	13	**1.321**	0.0897	0.0132	6.82	<.0001	34
SPIRITS-BRANDY-P13-750ML	31.74	<.0001	0.7426	0.7192	22	**1.243**	0.0574	0.0116	4.95	<.0001	35
SPIRITS-WHISKY-P49-750ML	40.59	<.0001	0.7868	0.7674	5	1.495	0.0552	0.0072	7.7	<.0001	36
SPIRITS-LIQUEURS-P9-750ML	25.48	<.0001	0.6985	0.6711	49	1.067	-0.0179	0.4501	-0.04	**0.9684**	Exclude

TABLE 4.29: *Part 2: The results of the multiple regression price elasticity models with SPIRITS-LIQUEURS-P16-750ML as the dependent variable*

In Tables 4.25 and 4.26 WINE-BIB-P79-1.5LTR has a positive coefficient, indicating that it is a substitute product for RTD-CIDER-P18-660ML. If the coefficient has a negative sign then the two products are complementary. A one percent increase in the price of WINE-BIB-P79-1.5LTR would result in a 4.3% increase in the volume of RTD-CIDER-P18-660ML, all other variables remaining constant. As the price of WINE-BIB-P79-1.5LTR increase, so consumers switch to RTD-CIDER-P18-660ML. Looking again at Tables 4.25 and 4.26 the own elasticity of RTD-CIDER-P18-660ML is -3.7642%, showing that a one percent increase in RTD-CIDER-P18-660ML's price will result in a 3.7642% decrease in volume. All the coefficients in Table 4.25 to Table 4.31 can be interpreted in the same way. Remembering to be cautious with these interpretations because of the simple nature of the models.

Product	F Value	Pr > F	R^2	AdjR^2	AdjR^2 Ranking	Durbin-Watson D	Parameter Estimate	Standard Error	t-Value	Pr > |t|	Parameter Estimate Ranking
RTD-FABS-P30-330ML	13.92	<.0001	0.5585	0.5184	4	1.86	1.9586	0.3321	5.9	<.0001	1
RTD-FABS-P27-330ML	14.16	<.0001	0.5628	0.5231	2	1.914	1.7019	0.2817	6.04	<.0001	2
SPIRITS-VODKA-P1-1LTR	11.1	<.0001	0.5023	0.4571	11	1.578	1.6439	0.4288	3.83	0.0002	3
WINE-FORTIFIED-P10-750ML	13.6	<.0001	0.5528	0.5121	5	1.789	1.6091	0.2824	5.7	<.0001	4
SPIRITS-WHISKY-P5-1LTR	14.04	<.0001	0.5607	0.5208	3	1.823	1.6078	0.2692	5.97	<.0001	5
RTD-BEER-P58-750ML	13.32	<.0001	0.5477	0.5065	6	1.871	1.4785	0.2677	5.52	<.0001	6
WINE-FORTIFIED-P11-750ML	14.58	<.0001	0.5699	0.5308	1	1.857	1.3285	0.2115	6.28	<.0001	7
WINE-RED-P102-750ML	9.86	<.0001	0.4512	0.4054	43	1.559	-1.286	0.1872	-6.87	<.0001	8
WINE-BIB-P66-5LTR	10.18	<.0001	0.4807	0.4335	17	1.741	0.9482	0.3328	2.85	0.005	9
WINE-PERLE-P7-1.5LTR	9.85	<.0001	0.4723	0.4244	20	1.599	-0.7812	0.3262	-2.4	0.0179	10
SPIRITS-GIN-P4-750ML	11.4	<.0001	0.509	0.4643	9	1.62	0.7561	0.1843	4.1	<.0001	11
WINE-SPARKLING-P25-750ML	6.13	<.0001	0.358	0.2997	63	1.374	-0.6926	0.1911	-3.62	**0.0004**	Exclude
WINE-FORTIFIED-P6-375ML	10.24	<.0001	0.4822	0.4351	15	1.698	-0.6524	0.223	-2.93	0.004	12
WINE-RED-P44-750ML	10.52	<.0001	0.4888	0.4423	13	1.664	-0.6054	0.1867	-3.24	0.0015	13
RTD-BEER-P24-340ML	9.45	<.0001	0.4621	0.4132	28	1.503	-0.5546	0.3249	-1.71	**0.09**	Exclude
WINE-RED-P79-750ML	10.36	<.0001	0.4849	0.4381	14	1.719	0.5443	0.1778	3.06	0.0026	14
SPIRITS-BRANDY-P23-750ML	9.6	<.0001	0.4659	0.4173	26	1.449	-0.5078	0.2557	-1.99	0.0489	15
WINE-BIB-P28-3LTR	10.77	<.0001	0.4948	0.4489	12	1.697	0.4975	0.1415	3.52	0.0006	16
WINE-RED-P39-750ML	9.86	<.0001	0.4728	0.4248	19	1.592	-0.4702	0.1944	-2.42	0.0168	17
WINE-FORTIFIED-P16-750ML	9.64	<.0001	0.467	0.4185	25	1.523	-0.4449	0.2162	-2.06	0.0414	18
WINE-RED-P68-750ML	9.9	<.0001	0.4737	0.4259	18	1.676	0.3907	0.1578	2.48	0.0144	19
WINE-BIB-P55-5LTR	9.41	<.0001	0.461	0.412	30	1.582	0.3336	0.207	1.61	**0.1092**	Exclude
WINE-RED-P22-750ML	9.84	<.0001	0.4722	0.4243	21	1.518	-0.329	0.1377	-2.39	0.0182	20
SPIRITS-WHISKY-P11-750ML	9.19	<.0001	0.4553	0.4057	42	1.615	0.3262	0.315	1.04	**0.3021**	Exclude
WINE-RED-P59-750ML	9.33	<.0001	0.4588	0.4096	34	1.641	0.3227	0.2272	1.42	**0.1577**	Exclude
WINE-BIB-P78-5LTR	9.75	<.0001	0.47	0.4218	23	1.6	-0.3199	0.1421	-2.25	0.0258	21
WINE-BIB-P44-3LTR	9.72	<.0001	0.4691	0.4208	24	1.68	0.3192	0.1453	2.2	0.0296	22
RTD-FABS-P1-250ML	9.32	<.0001	0.4587	0.4095	35	1.554	0.2972	0.2104	1.41	**0.16**	Exclude
WINE-RED-P33-750ML	9.31	<.0001	0.4585	0.4093	37	1.521	-0.295	0.2123	-1.39	**0.1668**	Exclude
WINE-WHITE-P58-750ML	9.78	<.0001	0.4706	0.4224	22	1.526	-0.267	0.1166	-2.29	0.0235	23
WINE-SPARKLING-P1-750ML	9.34	<.0001	0.4593	0.4101	32	1.505	-0.2659	0.1815	-1.46	**0.1453**	Exclude
WINE-RED-P19-750ML	9.34	<.0001	0.4591	0.41	33	1.662	0.2564	0.1766	1.45	**0.1488**	Exclude
WINE-RED-P61-750ML	9.32	<.0001	0.4587	0.4095	36	1.645	0.2336	0.1651	1.41	**0.1593**	Exclude
WINE-RED-P66-750ML	9.35	<.0001	0.4595	0.4104	31	1.628	0.2331	0.1569	1.49	**0.1394**	Exclude
WINE-RED-P7-750ML	9.18	<.0001	0.4548	0.4053	44	1.594	0.2301	0.2341	0.98	**0.3274**	Exclude
WINE-RED-P9-750ML	9.18	<.0001	0.4549	0.4053	45	1.532	-0.1761	0.1789	-0.98	**0.3268**	Exclude
WINE-BIB-P84-5LTR	9.17	<.0001	0.4546	0.405	46	1.534	-0.1722	0.1812	-0.95	**0.3437**	Exclude
WINE-SPARKLING-P2-750ML	9.43	<.0001	0.4617	0.4127	29	1.508	-0.1708	0.1022	-1.67	**0.0968**	Exclude
WINE-RED-P50-750ML	9.15	<.0001	0.454	0.4044	47	1.525	-0.1556	0.1809	-0.86	**0.3913**	Exclude
WINE-RED-P54-750ML	9.11	<.0001	0.4529	0.4032	49	1.598	0.1137	0.1691	0.67	**0.5023**	Exclude
WINE-RED-P49-750ML	9.09	<.0001	0.4525	0.4027	50	1.539	-0.1112	0.1908	-0.58	**0.5609**	Exclude
WINE-BIB-P45-3LTR	9.26	<.0001	0.4572	0.4078	38	1.54	-0.1014	0.0806	-1.26	**0.2105**	Exclude
WINE-BIB-P41-3LTR	9.12	<.0001	0.4534	0.4037	48	1.565	-0.0961	0.1271	-0.76	**0.4511**	Exclude
WINE-RED-P72-750ML	9.05	<.0001	0.4514	0.4015	57	1.546	-0.0919	0.3616	-0.25	**0.7999**	Exclude

TABLE 4.30: *Part 1: The results of the multiple regression price elasticity models with WINE-RED-P102-750ML as the dependent variable*

Product	F Value	Pr > F	R^2	AdjR^2	AdjR^2 Ranking	Durbin-Watson D	Parameter Estimate	Standard Error	t-Value	Pr > |t|	Parameter Estimate Ranking
WINE-RED-P11-750ML	9.07	<.0001	0.452	0.4022	53	1.55	-0.0871	0.1873	-0.47	**0.6426**	Exclude
WINE-RED-P37-750ML	9.23	<.0001	0.4562	0.4067	40	1.535	-0.0855	0.0745	-1.15	**0.2531**	Exclude
SPIRITS-WHISKY-P21-750ML	9.08	<.0001	0.4521	0.4023	52	1.575	0.0815	0.1646	0.5	**0.6213**	Exclude
WINE-BIB-P67-5LTR	9.08	<.0001	0.4522	0.4024	51	1.569	-0.0701	0.136	-0.52	**0.6071**	Exclude
WINE-PERLE-P8-1.5LTR	9.06	<.0001	0.4515	0.4017	54	1.569	0.0671	0.2187	0.31	**0.7594**	Exclude
WINE-BIB-P54-5LTR	10.21	<.0001	0.4813	0.4342	16	1.753	-0.059	0.0205	-2.88	0.0045	24
WINE-BIB-P52-3LTR	11.59	<.0001	0.5131	0.4688	8	1.705	0.0464	0.0109	4.26	<.0001	25
RTD-FABS-P23-300ML	12.22	<.0001	0.5263	0.4833	7	1.665	0.0454	0.0096	4.76	<.0001	26
WINE-BIB-P43-3LTR	9.06	<.0001	0.4515	0.4016	56	1.562	-0.04	0.1336	-0.3	**0.7654**	Exclude
WINE-RED-P87-750ML	9.05	<.0001	0.4513	0.4014	59	1.57	0.0373	0.1948	0.19	**0.8486**	Exclude
WINE-SPARKLING-P22-750ML	9.06	<.0001	0.4515	0.4017	55	1.561	0.0355	0.1125	0.31	**0.7533**	Exclude
WINE-RED-P23-750ML	9.05	<.0001	0.4514	0.4015	58	1.554	-0.0311	0.127	-0.24	**0.8068**	Exclude
WINE-RED-P32-750ML	9.05	<.0001	0.4513	0.4014	60	1.566	0.0283	0.1605	0.18	**0.8603**	Exclude
WINE-SPARKLING-P19-750ML	11.12	<.0001	0.5028	0.4576	10	1.643	0.0247	0.0064	3.85	0.0002	27
SPIRITS-BRANDY-P12-750ML	9.05	<.0001	0.4512	0.4014	61	1.564	0.0236	0.1702	0.14	**0.89**	Exclude
RTD-BEER-P45-500ML	9.46	<.0001	0.4623	0.4134	27	1.556	0.0174	0.0102	1.72	**0.0873**	Exclude
RTD-FABS-P36-440ML	9.25	<.0001	0.4569	0.4075	39	1.547	0.0172	0.0141	1.22	**0.2229**	Exclude
WINE-RED-P31-750ML	9.04	<.0001	0.4512	0.4013	62	1.552	-0.0149	0.1476	-0.1	**0.92**	Exclude
RTD-CIDER-P14-440ML	9.2	<.0001	0.4556	0.4061	41	1.56	0.0121	0.0112	1.07	**0.2845**	Exclude

TABLE 4.31: *Part 2: The results of the multiple regression price elasticity models with WINE-RED-P102-750ML as the dependent variable*

The Durbin-Watson D test statistics are shown in Tables 4.25 to Table 4.31. The results show that for the models of RTD-CIDER-P6-330ML, SPIRITS-LIQUEURS-P16-750ML and WINE-RED-P102-750ML, the Durbin-Watson D test indicates a very low possibility of autocorrelation.

Reading off a Durbin-Watson table, the lower bound for positive autocorrelation is 1.444 and the upper bound 1.814, at a 1% significance level. The confidence intervals for $n = 150$, rounding down from 157 observation and $k = 13$ where k is the number of independent variables, where the intercept does not equal zero.

The results for RTD-CIDER-P18-660ML in Tables 4.25 and 4.26 show that all but one of the models have Durbin-Watson D test statistics below 1.444. Only RTD-BEER-P50-660ML has a Durbin-Watson D test statistics in between the lower and upper bound. Of the remaining models nineteen have Durbin-Watson D test statistics below 1, these have been highlighted. These models should not be trusted because positive autocorrelation is present, they have been marked for exclusion. In Figure A.25 in the appendix, RTD-CIDER-P18-660ML and the model showing its interaction with RTD-BEER-P3-330ML, produces a Durbin-Watson D test statistics of 0.793. The 1st order autocorrelation is greater than 0.5, therefore this models will not be seen as significant and removed from the results. In Figure A.25 in the appendix, the upward trend in the plot of the predicted value against the residuals is a strong indication of autocorrelation.

The results for RTD-CIDER-P6-330ML in Table 4.27 show that six of the models have Durbin-Watson D test statistics above 1 but below 1.444. These models should be treated with caution but will not be excluded, remembering that 1st order autocorrelation occurs when the consecutive errors are correlated. 2nd order autocorrelation occurs when error terms two periods apart are correlated, and so on [28].

For each model the graphical display of the residuals versus the predicted values and the probability distributions of the residuals should be checked. The plot of the residuals versus the predicted values should have no pattern. The probability distribution of the residuals should follow a normal curve. Looking at RTD-CIDER-P6-330ML and the model showing its interaction with SPIRITS-LIQUEURS-P11-750ML, produces a Durbin-Watson D test statistics of 1.347. In Figure A.26 in the appendix, the graphical display of the output statistics for this price elasticity regression model are shown. The residuals plot does not show any obvious pattern and the probability distribution follows a normal curve, although slightly skewed to the left. The same follows for the other 5 variables with Durbin-Watson D test statistics below 1.444. There were no models with test statistics below one.

The results for SPIRITS-LIQUEURS-P16-750ML in Tables 4.28 and 4.29 displays that twenty three of the models have Durbin-Watson D test statistics above one but below 1.444. These models should be treated with caution but will not be excluded. There were no models with test statistics below one. The plots of the residuals and the probability distributions are used to assess if positive autocorrelation is present and affecting the models. For example in Figure 4.27 the output statistics as well as the Durbin-Watson D test is shown for SPIRITS-LIQUEURS-P16-750ML and the interaction with the independent variable SPIRITS-BRANDY-P4-1LTR. The results for the other twenty two models are similar, there is no obvious pattern in the residuals plot versus the predicted values and the probability distributions can be fitted to a normal curve, most slightly skewed to the right.

In Figure 4.27 the plot of residuals by predicted values in the upper-left corner of the diagnostics plots is used to investigate slight trends in the residuals. A fan-shaped trend might indicate the need for a variance-stabilizing transformation. A curved trend (such as a semicircle) might indicate the need for a quadratic term in the model. Since these residuals have no apparent trend, the analysis is considered to be acceptable.

A trend in the residuals would indicate non-constant variance in the data. The plots of residuals and standardised residuals versus predicted values exhibit no obvious patterns. The plot of externally standardised residuals (RStudent) by leverage values reveals that there are indications of a few outlying data points with high leverage that might be overly influencing the fit produced. The plot of Cook's D value by observation also indicates four highly influential observations, these are investigated and found to be important seasonal dummy variables and should be retained in the model. The normal quantile plot of the residuals and the residual histogram are consistent with the assumption of a linear relationship [61]. The points on the plot of the dependent variable versus the predicted values lie along or around a 45-degree line, indicating that the model predicts the behaviour of the dependent variable.

Finally the Fit-Mean vs. Residuals plot is interpreted by looking at the heights of the two plots. If the left side (Fit) is taller than the right side (residuals), then it means that the independent variables in the model explain a lot of the variation in the response variable. If the right side (residuals) are taller, then it means that there is still a lot of unexplained variation. The right side is taller, therefore the independent variables in the model explain a lot of the variation in the response variable.

The results for WINE-RED-P102-750ML in Tables 4.30 and 4.31 show that all the models have Durbin-Watson D test statistics above 1.444 but below 2. This indicating strongly that autocorrelation is not present in these models. The product's own price, a competitor and seasonality are enough to have a significant model void of autocorrelation. The adjusted R^2 does however indicate that not all the variance in the dependent variable is explain by price elasticities alone.

The final results are shown in Table 4.25 to Table 4.31, the models that have been excluded are shown as such in the final column ranking the relative strength of the coefficient parameter. These show that for RTD-CIDER-P18-660ML and WINE-RED-P102-750ML there are 27 products in their competitor sets. For RTD-CIDER-P6-330ML there are 26 products in its competitor set and SPIRITS-LIQUEURS-P16-750ML having 36 products.
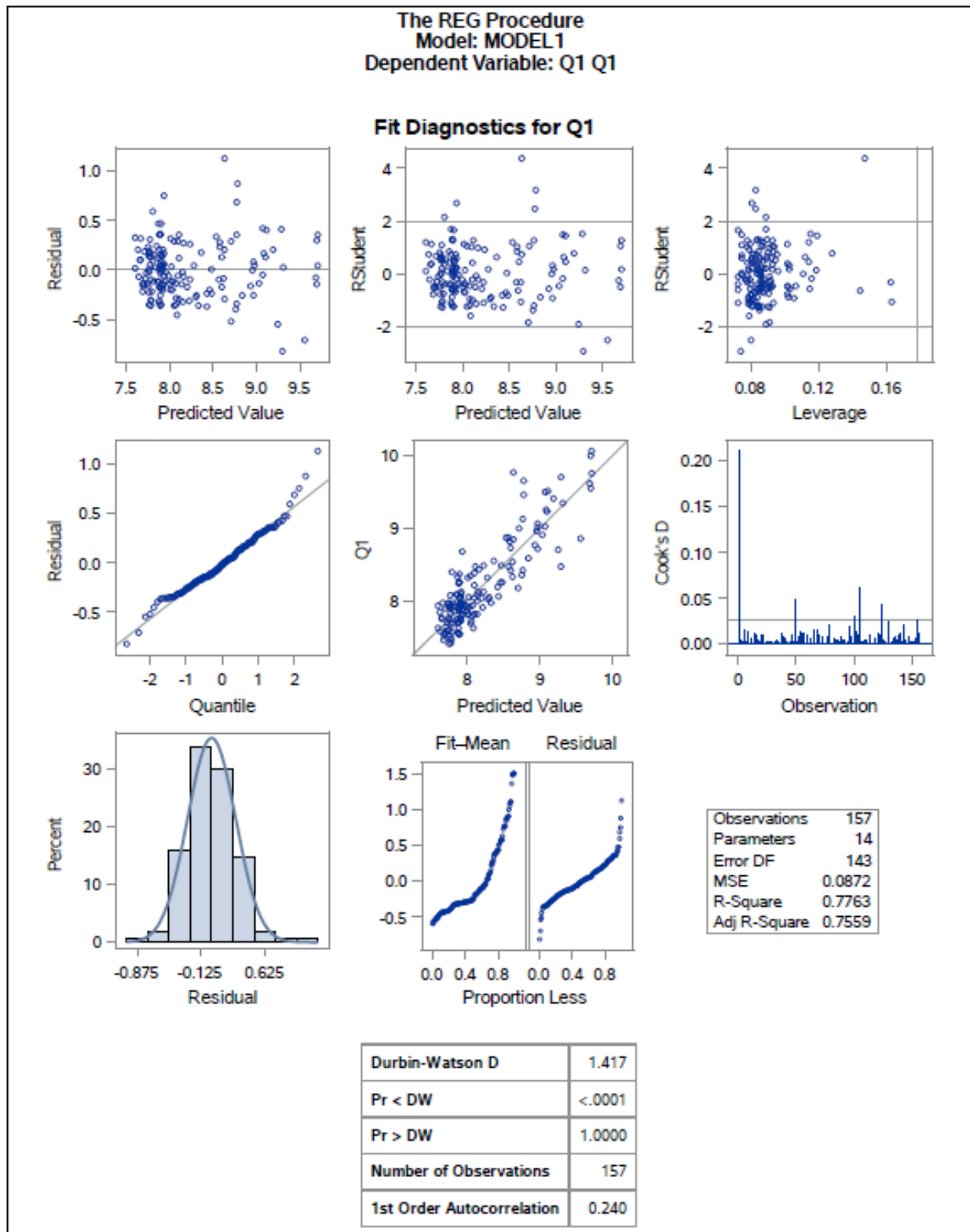
FIGURE 4.27: *SPIRITS-LIQUEURS-P16-750ML output statistics for price elasticity regression model with SPIRITS-BRANDY-P4-1LTR*

4.6   Chapter summary

Sub-chapter 4.2 demonstrates the PLC methodology using factor analysis, cluster analysis and the BCG growth share matrix. The PLC methodology segments all 594 products into one of four defined life cycle stages. Each life cycle stage has a well defined marketing objective and strategy. The PLC methodology is best used to provide a framework to explain market changes or dynamics [18].

The results reveal that the products are competing for the same type of consumer attention, either to attract consumers to make the initial purchase or to keep them making that same purchase taking the market and economic factors into account as well.

Using factor analysis, cluster analysis and the BCG growth share matrix gives the PLC methodology depth. Each taxonomy can be viewed in many different ways: as a whole, by factor, cluster, segment or category or any combination. The results give the researcher the ability to view the groupings under different lenses but ultimately achieving the objective of grouping the products into competitor sets.

After using these methods and investigating the results it would be recommended to also include other variables such as the own price elasticities, alcohol percentage by volume, product age, number of products in the brand's range, seasonality and additional questions in the surveys relating specifically to consumer drinking occasions and need states. A limitation of the factor and cluster analysis methods is that they are only as reliable and purposeful as the variables fed into them. If redundant, inconsequential or insignificant variables are used, they will cause these methods to find constructs and relationships that do not exist. It is therefore important to refine the variables used to define each product.

An advantage of using these methods is that many years of historical time series data is not needed, aggregated data for any consistent length across the variable can be used. Another advantage is that the information that is used for the variables does not need to be of the same scale, as a data transformation is part of the methodology. Finally the data does not need to be for any significant period of time for example, the information used could be only for the last 3 months to get a view of the market now, not taking into account any historical effects.

Sub-chapter 4.5.1 demonstrates the use of the log-log regression model (equation (2.21)) and the price elasticity methodology. The four sample products previously determined are modelled and the price elasticity estimates and their significance are calculated for each product thereby defining the products competitor set. Correlation analysis is tested as a possible variable refinement method but regression analysis is found to be the better method.

In Sub-chapters 4.5.1 and 4.4 the methodology followed required a few steps to reduce the number of products in the competitor sets, this is a limitation of the price elasticity using multiple regression methodology, as it is limited by the number of observations being in the ratio of 10 observation to 1 independent variable being included in the model. Another limitation of this method is that it requires the products to have had significant price variations during the period under investigation, some price movement is required for the change in demand to be visible. Also, large amounts of historical time series data is required. If the products in the dataset have been aggregated it is possible that the shifts and changes within the data would have been smoothed out and regression analysis would not be able to determine significant relationships.

An advantage of using the price elasticity methodology of determining products' relationships is that the results give scale and significance to each relationship. The results show if the relationship between the two products being investigated is significant or not (p-value of parameter estimates) and the scale of the relationship (parameter estimate score).

CHAPTER 5

Discussion of the price elasticity model and the product life cycle method

Contents

In the alcoholic beverage industry competition between brands and products is intense, as can be seen by the large amounts of money spent on advertising. Competitor products are business rivals who vie for the consumers' loyalty. These products usually have different features, benefits and prices but consumers see them as substitutes or in the same group under consideration when choosing a product to purchase.

Two methods have been used to determine the competitor sets of the four products. Multiple regression analysis using the price elasticity methodology has been used as a benchmark [75], [47], as this is the most common method used in the industry. The PLC methodology relates to the life cycle of a product as it moves through four distinct life cycle stages with respect to the industry landscape, investment required and potential revenue [39].

Firstly, regression analysis is a statistical process for estimating the relationships among variables [81]. The price elasticity methodology used own price elasticity of demand of a product to measure the relationship between a product's demand and its own price, all other market factors remaining constant. Cross price elasticity of demand measures the relationship between one product's volume and another product's price, measuring the change in the demand for a quantity of a product relative to a change in price of another product, all other market factors remaining constant [45]. This methodology does not directly take into account advertising expenditure or any strategy the company may have for the brand or product, these variables are not included as input into the model but their effect is observed in the resulting relationships. The growth of a segment or market factors outside of a product are also not directly

143

included in the price elasticity methodology.

The price elasticity methodology does however assume that there are many different factors acting on a brand or product and this can be viewed through the significance of relationships of the independent variable with the dependent variable and the magnitude of the parameter coefficients. For example, if a product has a high level of advertising that resonates with consumers, the brand-health will be good and advertising expenditure high. It would then be expected that this product would have low price elasticity relative to other products (this will be different for each industry). This brand has built a loyal consumer base, when other products reduce their price most consumers will stay loyal to the original brand and not substitute it. This would yield a low price elasticity as there is very little change in quantity of product A, due to changes in the price of product B. In this way the price elasticity does not directly take market factors, brand-health and advertising expenditure into account but the effects can be interpreted in the output.

The PLC methodology is a holistic approach, taking into account input variables relating to the market, product and consumers. The first step is factor analysis, which allows for a number of variables to be included and reduced down to a few constructs that represent those variables. Factor analysis groups products based on how they resonate with different variables representing the market factors, advertising expenditure and consumer sentiment. Cluster analysis then groups the products based on how they resonate with the constructs formed but relative to all other products. The BCG matrix incorporates the market potential and company strategy into the methodology.

5.1   A comparison of the price elasticity model and the product life cycle method

Table 5.1 is a comparison of the two methods, looking at the inputs, steps involved and output results. Firstly, the input data that is required for each methodology is investigated. Price elasticity models require detailed historical time series data, while the PLC methodology does not have these restrictions. The PLC methodology allows for any period of time's data to be included. The model could be run with data for the last three months for example, to get a view of the market now and thereby excluding historical effects. Another model could be run with the last twelve months' data and the two models can be compared for changes. The PLC methodology is versatile as it is not restricted by the type of data included as long as all the market factors are represented. As this methodology does not require detailed weekly time series data, data collection should be simpler as well as updating the model more frequently. The price elasticity model requires extensive historical data and the historical effects are taken into account in the output. Updating the model frequently does not yield new results for example, if the model is updated monthly adding a month's worth of data or 4 data points will not drastically change the output.

With the price elasticity model the number of independent variables needs to be reduced due to the problem of 'overfit' in multiple regression. This involves interviewing industry insiders to get a view of what they feel the possible competitor set should be. This is a subjective process and can leave unknown relationships undiscovered due to exclusion. The PLC method used factor analysis to refine the number of variables, this is not a subjective process using insiders' knowledge.

The computational burden to compute the PLC is less than when computing price elasticity. The PLC

methodology required the following models to be executed: one application for variable refinement factor analysis, eight applications to define the number of factors, one application for final factor results, one application to standardise the data, one application for Ward's method, thirty applications to define k for k-means and thirteen applications to test for the k for the sub clusters. This is a total of forty-five models that were applied in the steps to define the PLC for all the products.

The price elasticity model required the following models to be applied: twenty-three applications for stepwise variable refinement and two-hundred-and-nineteen applications for significance testing of variables. This is a total of two-hundred and two models that were applied to define the final competitor set of the four sample products being investigated.

The PLC methodology is able to segment the total market in a total of 45 computational sets or applications of models, the price elasticity methodology needs two-hundred-and-nineteen runs to segment four products, this will make segmenting the full market and frequently updating the results prohibitive.

The price elasticity methodology	The product life cycle methodology
Input Data Required	
- Historical time series and binary variables	- Categorical, ratio, interval and binary variables
- On a detailed weekly level	- Aggregated annual level
- At least 3 years of weekly data	- Any time period consisted across the variables
- Price fluctuations needed to determine relationships	- No special data requirements
- Limited number of variables can be included in the model (1 for every 10 observations)	- Large number of variables can be included
Model Steps	
- Subjective variable refinement (industry insiders)	- Factor analysis
- Variable refinement (stepwise and batches)	- Cluster analysis
- Individual models	- BCG growth share matrix
Model Computation	
- Heavy	- Medium
Output Size	
- RTD-CIDER-P18-660ML - 27 products	- RTD-CIDER-P18-660ML - 56 products
- RTD-CIDER-P6-330ML - 26 products	- RTD-CIDER-P6-330ML - 69 products
- SPIRITS-LIQUEURS-P16-750ML - 36 products	- SPIRITS-LIQUEURS-P16-750ML - 260 products
- WINE-RED-P102-750ML - 27 products	- WINE-RED-P102-750ML - 209 products
Results	
- Small groups	- Large groups
- Relationships tested for significance	- Relationships not tested for significance
- Results can be ranked in order of size of interaction (parameter coefficients)	- Results can be ranked in order of size of competitor (relative market share)

TABLE 5.1: *Comparison of price elasticity methodology and the product life cycle methodology focusing on the two models inputs, processes and results*

The price elasticity model requires significant price changes during the period under investigation for both the dependent and independent variables. In the market some products do not have frequent discounts and only inflation-based or periodic absolute price increases are made. These products will not produce significant results using the price elasticity methodology, the method will exclude these products as possible competitors as no significant shifts in volume due to price changes can be found. The PLC method does not require any special occurrences or events to take place during the time period used for investigation.

The final results for the PLC method is that the total market has been segmented and split into groups. These groups are large and the results from cluster analysis and factor analysis can be used to create smaller groups within the four main quadrants of the BCG growth share matrix. The results for price elasticity methodology are concise groups and they have been tested for significant relationships. The

price elasticity regression results can be ranked according to the 'largest competitor', the product that the dependent variable will lose the most volume to due to a change in its price. The results from the PLC method can also be ranked using the relative market share, but because the market share is relative and not absolute, there will be products with the same share.

The results are displayed in Figure 5.1 comparing the similarity or uniqueness in the findings. The results show that the price elasticity methodology and product life cycle methodology do not produce the same sub-set of products as a product's competitor set. RTD-CIDER-P18-660ML only has one overlapping product, RTD-BEER-P2-330ML, this is because the price elasticity methodology is looking for relationships with RTD-CIDER-P18-660ML and all products on the market and therefore finding a high interaction between high volume and big market share products, while the PCL methodology is taking into account that products in the cider category that do not have a high market share but are growing, this is why RTD-CIDER-P18-660ML is in the Question Mark life cycle stage.

Sharp [70] defines '10 laws', one being the duplication of purchase law. The law states "A brand's customer base overlaps with rival brands in line with its market share. If 30% of a brand's buyers also bought brand A in a period, then 30% of every rival brand's customers also bought brand A"[70]. This law is visible in the price elasticity methodologies' results, for RTD-CIDER-P18-660ML the results show that the product is interacting with all the large market share products. The PCL methodology however, also take market factors and product attributes into account and finds a set of products that RTD-CIDER-P18-660ML is competing with to become a big brand, not just the big brands that it is competing with.

FIGURE 5.1: *Analysis of the overlap or uniqueness in the results of the price elasticity and product life cycle methods competitor sets*

In Figure 5.2 the results are compared in terms of the categories that each competitor set products is allocated to, for RTD-CIDER-P18-660ML and the PLC methodology the results show that a majority of the competitor set are beer and vodka products.
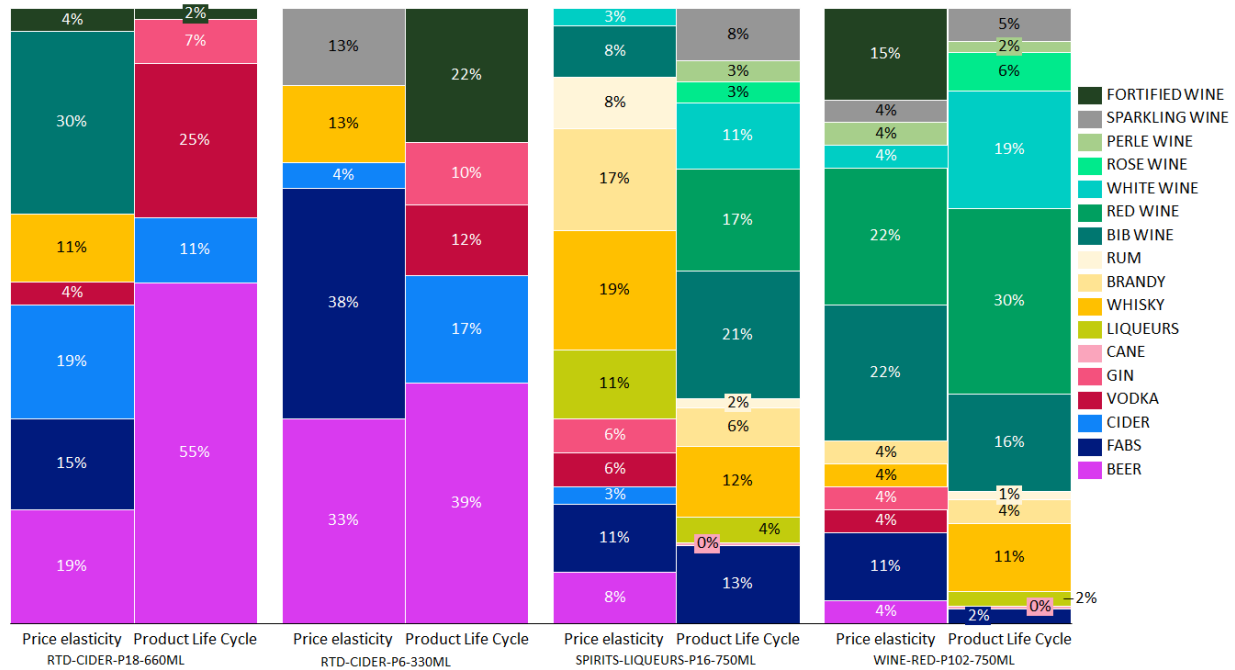
FIGURE 5.2: *Analysis of the overlap or uniqueness the price elasticity and product life cycle methods between categories*

Previously in Table 4.20 the market growth rates are shown for the different categories and this also shows that the beer and vodka categories are in growth. The PCL methodology is finding results that are consistent with consumer preference shift within the market and different categories, while the price elasticity methodology is finding results that are consistent with overall market share.

These two methodologies are both segmenting the products into competitor sets but they are taking fundamentally different approaches and factors into account, while a consumer making a purchasing decision is 'subjected' to all these factors. A consumer will be exposed to popular category trends by way of what other people are drinking, they will also be more aware of high market share products because these product tend to be advertised more and have high awareness levels. For this reason it is recommended to use both methodologies and combine the results to find the final competitor set that takes all the variables into account.

In Figures 5.1 and 5.2 the results for RTD-CIDER-P6-330ML, SPIRITS-LIQUEURS-P16-750ML and WINE-RED-P102-750ML all show a small overlap between the two methodologies.

SPIRITS-LIQUEURS-P16-750ML is in the Cow or maturity life cycle stage, this life cycle stage tends to have all the high market share products and therefore the price elasticity and PLC methodologies have a higher percentage overlap for this product.

RTD-CIDER-P18-660ML has competitor set products mostly in one category across both methodologies. RTD-CIDER-P6-330ML has a large percentage of its competitor set being beer in both methodologies. While the price elasticity methodologies finds competitor products in the FABs category, the PLC methodology finds competitor products in the cider category.

For all the products the results show that both methodologies are finding results that are different to each other but both represent factors consumers are having to take into account when making purchasing decisions and the most encompassing competitor set would be to combine the results of the two methodologies.

5.2   Combination of results: Final competitor set

The next step is to investigate advertising spend of a specific product relative to other products in the competitor set. The two methods to evaluate efficiency of advertising spend that will be used are DEA and multiple regression.

In Section 4.1 SPIRITS-LIQUEURS-P16-750ML was identified as the product that will be used as an example to investigate and compare results for the two methodologies. The competitor set of SPIRITS-LIQUEURS-P16-750ML needs to be included in the model to test the efficiency of SPIRITS-LIQUEURS-P16-750ML advertising expenditure relative to its competitors' efforts.

The requirement for the competitors set for the DEA model and the multiple regression model need to be taken into account now. The DEA model and the multiple regression model will be compared in the discussion. The multiple regression model needs to take into account all the forces acting on a product, so that these relationships can be tested for significance and strength. The competitor set of SPIRITS-LIQUEURS-P16-750ML needs to be included in the model to test the relationship between the products in the competitor set and SPIRITS-LIQUEURS-P16-750ML. DEA does not require the inputs or outputs to have a significant linear relationships or any other functional form.

After combining the results from the previous chapter for SPIRITS-LIQUEURS-P16-750ML, there are 22 products that overlap between the price elasticity methodology and PLC. All 22 will be used, as well as the top 10 products from both methodologies. For the price elasticity using multiple regression, the 10 products with the largest absolute coefficient parameter or price elasticities will be used. The PLC methodology, SPIRITS-LIQUEURS-P16-750ML is in the maturity life cycle stage and cluster 3, products with these two criteria will be used. This amounts to 32 products in the competitor set.

These products are:

- SPIRITS-LIQUEURS-P2-1LTR,

- SPIRITS-BRANDY-P24-750ML,

- SPIRITS-RUM-P1-1LTR,

- RTD-BEER-P24-340ML,

- WINE-BIB-P48-3LTR,

- SPIRITS-WHISKY-P20-750ML,

- SPIRITS-GIN-P2-200ML,

- RTD-FABS-P6-250ML,

- RTD-FABS-P22-275ML,

- SPIRITS-WHISKY-P23-750ML,

- SPIRITS-LIQUEURS-P9-750ML,

- SPIRITS-WHISKY-P25-750ML,

- SPIRITS-LIQUEURS-P12-750ML,

- SPIRITS-BRANDY-P5-375ML,

- SPIRITS-RUM-P6-750ML,

- SPIRITS-BRANDY-P2-1LTR,

- SPIRITS-BRANDY-P4-1LTR,

- WINE-BIB-P44-3LTR,

- WINE-BIB-P40-3LTR,

- SPIRITS-LIQUEURS-P8-750ML,

- SPIRITS-LIQUEURS-P7-750ML,

- SPIRITS-LIQUEURS-P6-750ML,

- RTD-FABS-P1-250ML,

- RTD-FABS-P23-300ML,

- SPIRITS-LIQUEURS-P3-750ML,

- SPIRITS-LIQUEURS-P4-750ML,

- WINE-WHITE-P18-750ML,

- SPIRITS-LIQUEURS-P5-750ML,

- SPIRITS-WHISKY-P19-750ML,

- SPIRITS-BRANDY-P13-750ML,

- SPIRITS-BRANDY-P3-1LTR and

- SPIRITS-LIQUEURS-P15-750ML.

The alcoholic beverage industry by nature generates highly correlated products. This is due to a vast number of similar product offerings, applying the same price discounts across a brands product range and having many products incorporated into the same promotion. The products are also influenced by the same external factors: weekends, public holidays, weather and consumer disposable income fluctuations to name a few. These factors make multicollinearity a strong possibility.

In Tables 5.2 and 5.3 the Pearson's correlation coefficients for SPIRITS-LIQUEURS-P16-750ML's competitor set are shown. The highly correlated products are highlighted and will need to be removed for multiple regression, the remainder are the products that will be used in the models.

The results show that to hedge against multicollinearity SPIRITS-LIQUEURS-P2-1LTR will need to be removed as it is highly correlated with:

- SPIRITS-RUM-P1-1LTR,
- SPIRITS-WHISKY-P20-750ML,
- SPIRITS-GIN-P2-200ML,

- RTD-FABS-P6-250ML,
- SPIRITS-BRANDY-P2-1LTR and
- SPIRITS-BRANDY-P3-1LTR.

From the price elasticity regression methodology results, SPIRITS-BRANDY-P4-1LTR is shown as SPIRITS-LIQUEURS-P16-750ML's biggest competitor, so this product should stay in the model. For this reason:

- SPIRITS-RUM-P1-1LTR,
- SPIRITS-WHISKY-P20-750ML,
- SPIRITS-GIN-P2-200ML,
- SPIRITS-BRANDY-P5-375ML,
- SPIRITS-BRANDY-P2-1LTR and
- SPIRITS-BRANDY-P3-1LTR,

will be removed from the model as they are highly correlated with SPIRITS-BRANDY-P4-1LTR.

Pearson's Correlation Coefficients N = 157

	SPIRITS-LIQUEURS-P2-1LTR	SPIRITS-LIQUEURS-P16-750ML	SPIRITS-BRANDY-P24-750ML	SPIRITS-RUM-P1-1LTR	RTD-BEER-P24-340ML	WINE-BIB-P48-3LTR	SPIRITS-WHISKY-P20-750ML	SPIRITS-GIN-P2-200ML	RTD-FABS-P6-250ML	RTD-FABS-P22-275ML	SPIRITS-WHISKY-P23-750ML	SPIRITS-LIQUEURS-P9-750ML	SPIRITS-WHISKY-P25-750ML	SPIRITS-LIQUEURS-P12-750ML	SPIRITS-BRANDY-P5-375ML	SPIRITS-RUM-P6-750ML	SPIRITS-BRANDY-P2-1LTR	SPIRITS-BRANDY-P4-1LTR
SPIRITS-LIQUEURS-P2-1LTR	1	0.3	0.6	**0.8**	0	0.5	**0.8**	**0.9**	**0.8**	**0.7**	**0.7**	0.3	0.4	0.4	0.6	**0.8**	**0.9**	**0.8**
SPIRITS-LIQUEURS-P16-750ML		1	0.3	0.1	0.1	0.2	0.1	0.1	0.1	0.1	0.2	0.2	0.1	0.2	0	0.3	0.3	0.2
SPIRITS-BRANDY-P24-750ML			1	0.6	-0.1	0.3	0.4	0.4	0.6	0.6	0.6	0.2	0.3	0.3	0.2	0.5	0.6	0.3
SPIRITS-RUM-P1-1LTR				1	0	0.3	**0.8**	**0.9**	0.7	**0.8**	**0.7**	0.3	0.3	0.4	**0.7**	**0.7**	**0.8**	**0.8**
RTD-BEER-P24-340ML					1	-0.1	0	-0.1	-0.3	-0.2	-0.1	0.3	0.1	0.3	0.2	0.1	0	0.1
WINE-BIB-P48-3LTR						1	0.5	0.4	0.5	0.4	0.3	0.2	0.1	-0.1	0.3	0.3	0.5	0.4
SPIRITS-WHISKY-P20-750ML							1	**0.8**	0.7	**0.7**	0.6	0.2	0.3	0.2	**0.7**	**0.7**	**0.8**	**0.8**
SPIRITS-GIN-P2-200ML								1	**0.8**	**0.8**	**0.7**	0.3	0.3	0.3	**0.7**	**0.8**	**0.8**	**0.8**
RTD-FABS-P6-250ML									1	**0.9**	0.6	0.2	0.3	0.1	0.5	0.6	**0.8**	**0.7**
RTD-FABS-P22-275ML										1	**0.7**	0.3	0.3	0.2	0.6	0.6	**0.7**	**0.7**
SPIRITS-WHISKY-P23-750ML											1	0.2	0.4	0.5	0.4	**0.7**	**0.7**	0.5
SPIRITS-LIQUEURS-P9-750ML												1	0.2	0.3	0.4	0.4	0.4	0.3
SPIRITS-WHISKY-P25-750ML													1	0.3	0.1	0.4	0.3	0.2
SPIRITS-LIQUEURS-P12-750ML														1	0.2	0.6	0.3	0.2
SPIRITS-BRANDY-P5-375ML															1	0.6	**0.7**	**0.8**
SPIRITS-RUM-P6-750ML																1	**0.8**	**0.7**
SPIRITS-BRANDY-P2-1LTR																	1	**0.9**
SPIRITS-BRANDY-P4-1LTR																		1

TABLE 5.2: *Part 1: Pearson's correlation coefficients for SPIRITS-LIQUEURS-P16-750ML's total competitor set*

Pearson's Correlation Coefficients N = 157

	WINE-BIB-P44-3LTR	WINE-BIB-P40-3LTR	SPIRITS-LIQUEURS-P8-750ML	SPIRITS-LIQUEURS-P7-750ML	SPIRITS-LIQUEURS-P6-750ML	RTD-FABS-P1-250ML	RTD-FABS-P23-300ML	SPIRITS-LIQUEURS-P3-750ML	SPIRITS-LIQUEURS-P4-750ML	WINE-WHITE-P18-750ML	SPIRITS-LIQUEURS-P5-750ML	SPIRITS-WHISKY-P19-750ML	SPIRITS-BRANDY-P13-750ML	SPIRITS-BRANDY-P3-1LTR	SPIRITS-LIQUEURS-P15-750ML
SPIRITS-LIQUEURS-P2-1LTR	0.6	0.6	0.5	0.5	0.3	**0.8**	0.5	0.4	0.4	0.5	0.3	0.6	0	**0.8**	0.4
SPIRITS-LIQUEURS-P16-750ML	0.1	0.3	0.1	0.1	0.1	0.2	0.1	0.2	0.3	0.2	0.2	0.2	-0.2	0.2	0.2
SPIRITS-BRANDY-P24-750ML	0.4	0.5	0.5	0.5	0.5	0.6	0.6	0.5	0.5	0.5	0.3	0.5	0	0.5	0.6
SPIRITS-RUM-P1-1LTR	0.6	0.5	0.4	0.4	0.3	**0.7**	0.5	0.4	0.5	0.5	0.3	0.5	0	**0.7**	0.4
RTD-BEER-P24-340ML	0	0	0.1	0.1	0	-0.2	-0.1	0.1	0.2	0.1	0.2	0	0	-0.1	0.1
WINE-BIB-P48-3LTR	0.3	0.4	0.2	0.2	0.1	0.6	0.6	0.2	0.2	0.3	0.2	0.3	-0.1	0.6	0.2
SPIRITS-WHISKY-P20-750ML	0.6	0.5	0.3	0.3	0.2	**0.7**	0.5	0.4	0.4	0.5	0.3	0.5	0.1	**0.8**	0.4
SPIRITS-GIN-P2-200ML	**0.7**	0.6	0.4	0.4	0.3	**0.8**	0.5	0.3	0.2	0.4	0.2	0.5	0.1	**0.8**	0.4
RTD-FABS-P6-250ML	**0.7**	**0.7**	0.4	0.4	0.4	**0.8**	**0.7**	0.3	0.3	0.5	0.3	0.5	0	**0.8**	0.4
RTD-FABS-P22-275ML	0.6	0.6	0.4	0.4	0.3	**0.7**	0.6	0.4	0.3	0.5	0.3	0.4	0	**0.8**	0.4
SPIRITS-WHISKY-P23-750ML	0.5	0.5	0.4	0.4	0.3	0.6	0.4	0.5	0.5	0.6	0.3	0.6	0	0.6	0.5
SPIRITS-LIQUEURS-P9-750ML	0.1	0.2	0.3	0.3	0.2	0.2	0.1	0.2	0.3	0.1	0.3	0.2	0.1	0.3	0.1
SPIRITS-WHISKY-P25-750ML	0.3	0.4	0.3	0.3	0.3	0.2	0.1	0.2	0.2	0.3	0.2	0.3	0.1	0.2	0.3
SPIRITS-LIQUEURS-P12-750ML	0.3	0.2	0.3	0.3	0.3	0.1	-0.1	0.4	0.4	0.4	0.2	0.2	0.1	0	0.3
SPIRITS-BRANDY-P5-375ML	0.5	0.5	0.2	0.2	0.1	0.5	0.3	0.2	0.2	0.3	0.2	0.2	0.2	**0.8**	0.2
SPIRITS-RUM-P6-750ML	0.6	0.5	0.5	0.5	0.4	0.6	0.3	0.4	0.5	0.5	0.3	0.5	0.2	**0.7**	0.5
SPIRITS-BRANDY-P2-1LTR	0.6	0.6	0.4	0.4	0.3	**0.8**	0.5	0.5	0.5	0.5	0.4	0.5	0	**0.9**	0.4
SPIRITS-BRANDY-P4-1LTR	0.6	0.5	0.3	0.3	0.1	**0.7**	0.4	0.3	0.3	0.3	0.3	0.3	0.1	**0.8**	0.2
WINE-BIB-P44-3LTR	1	0.6	0.3	0.3	0.2	0.6	0.4	0.3	0.3	0.5	0.1	0.4	0	0.6	0.3
WINE-BIB-P40-3LTR		1	0.3	0.3	0.3	0.6	0.6	0.2	0.2	0.5	0.2	0.5	0	0.6	0.4
SPIRITS-LIQUEURS-P8-750ML			1	1	**0.9**	0.4	0.3	0.4	0.3	0.3	0.2	0.3	0	0.3	0.4
SPIRITS-LIQUEURS-P7-750ML				1	**0.9**	0.4	0.3	0.4	0.3	0.3	0.2	0.3	0	0.3	0.3
SPIRITS-LIQUEURS-P6-750ML					1	0.4	0.3	0.3	0.2	0.2	0.1	0.2	0	0.2	0.3
RTD-FABS-P1-250ML						1	**0.7**	0.3	0.2	0.5	0.2	0.5	0	**0.8**	0.4
RTD-FABS-P23-300ML							1	0.2	0.2	0.5	0.2	0.5	-0.1	0.6	0.3
SPIRITS-LIQUEURS-P3-750ML								1	**0.9**	0.4	0.5	0.2	-0.1	0.4	0.4
SPIRITS-LIQUEURS-P4-750ML									1	0.5	0.5	0.2	-0.1	0.4	0.4
WINE-WHITE-P18-750ML										1	0.4	0.3	-0.1	0.4	0.5
SPIRITS-LIQUEURS-P5-750ML											1	0.2	0	0.3	0.3
SPIRITS-WHISKY-P19-750ML												1	-0.1	0.5	0.4
SPIRITS-BRANDY-P13-750ML													1	0.1	0.1
SPIRITS-BRANDY-P3-1LTR														1	0.3
SPIRITS-LIQUEURS-P15-750ML															1

TABLE 5.3: *Part 2: Pearson's correlation coefficients for SPIRITS-LIQUEURS-P16-750ML's total competitor set*

RTD-FABS-P6-250ML and RTD-FABS-P22-275ML are highly correlated, RTD-FABS-P22-275ML will be kept in the model as the product is more established in the market and RTD-FABS-P6-250ML

will be removed. SPIRITS-LIQUEURS-P8-750ML, SPIRITS-LIQUEURS-P7-750ML and SPIRITS-LIQUEURS-P6-750ML are all part of the same product range and highly correlated. SPIRITS-LIQUEURS-P8-750ML will be kept in the model as it has a higher correlation with SPIRITS-LIQUEURS-P16-750ML than the other two products, they will be removed.

SPIRITS-LIQUEURS-P3-750ML and SPIRITS-LIQUEURS-P4-750ML similarly are also part of the same product range and highly correlated. SPIRITS-LIQUEURS-P3-750ML is the more established product and will remain in the model, while SPIRITS-LIQUEURS-P4-750ML will be excluded. The relationship that the removed products have with SPIRITS-LIQUEURS-P16-750ML will be represented by the remaining competitor set included in the model.

This leaves the 21 independent variables for the competitor set, these products are:

- SPIRITS-BRANDY-P24-750ML,
- RTD-BEER-P24-340ML,
- WINE-BIB-P48-3LTR,
- RTD-FABS-P22-275ML,
- SPIRITS-WHISKY-P23-750ML,
- SPIRITS-LIQUEURS-P9-750ML,
- SPIRITS-WHISKY-P25-750ML,
- SPIRITS-LIQUEURS-P12-750ML,
- SPIRITS-RUM-P6-750ML,
- SPIRITS-BRANDY-P4-1LTR,
- WINE-BIB-P44-3LTR,
- WINE-BIB-P40-3LTR,
- SPIRITS-LIQUEURS-P8-750ML,
- RTD-FABS-P1-250ML,
- RTD-FABS-P23-300ML,
- SPIRITS-LIQUEURS-P3-750ML,
- WINE-WHITE-P18-750ML,
- SPIRITS-LIQUEURS-P5-750ML,
- SPIRITS-WHISKY-P19-750ML,
- SPIRITS-BRANDY-P13-750ML and
- SPIRITS-LIQUEURS-P15-750ML.

5.3 Chapter summary

The PCL methodology and the price elasticity methodology have been described, applied and the results were investigated and discussed. Finding that both methods produced significant results and both result sets having merit. Due to the worth and value that both methods produced the results are combined and a final competitor set is determined. In the following chapter the results will be used to determine the efficiency and effectiveness of advertising spend.

CHAPTER 6

Determining the effectiveness and efficiency of advertising efforts

Contents

This chapter describes the process and review of the use of multiple regression with dummy variables and data envelopment analysis to determine the effectiveness and efficiency of advertising efforts. As previously discussed DEA determines a measure of efficiency, while multiple regression determines significant relationships between a product's volume and effective marketing variables. The results from the previous chapter will be the starting point, as both methods require that a product's competitor set is known. SPIRITS-LIQUEURS-P16-750ML will be the product under investigation for both DEA and multiple regression with dummy variables. The same competitor set will be used for both methodologies.

Multiple regression with dummy variables has been established in the literary review as the industry standard for determining the effectiveness or significance of relationships in order to gain insights. DEA provides two insightful constructs, firstly, it determines if the product under investigation is efficient in its sales and absolute spend on the different advertising types. The second insight is provided by the composite units that are given as output for each inefficient product. The composite units give a best practice benchmark for each advertising variable [81] and sales value for each inefficient product. DEA creates a benchmark that is relative to the competitor set and therefore takes the market and other landscape factors into account. Each product in the competitor set is competing in the same landscape with the same influences and external factors.

The multiple regression model will use long periods of weekly historical time series data, while the DEA method does not require long periods of historical time series data. Both methodologies will be applied and compared for meaningful insights. Figure 6.1 gives an overview of the process and procedures that will be followed in this chapter.
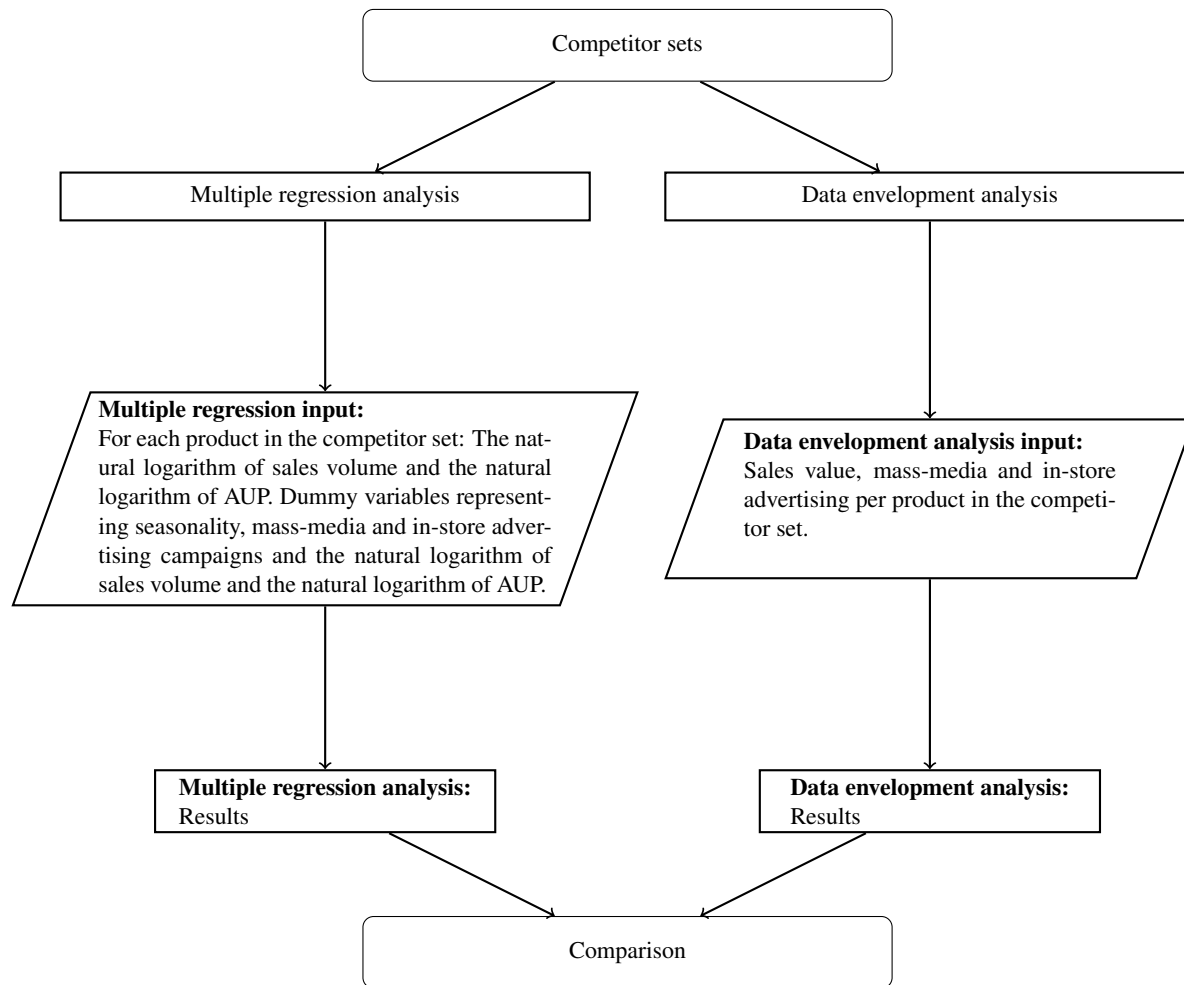
FIGURE 6.1: *Data and methodology procedure for determining efficiency and effectiveness of advertising efforts in Chapters 6*

6.1 Method for determining the significance of advertising efforts using regression analysis

Multiple regression analysis can be used to assist a brand team who wants to know how the demand for their company's product is related to the different types of advertising expenditure. Multiple regression analysis will be used to calculate the elasticity of demand with respect to advertising expenditure, that is the percent change in demand in response to a one percent change in the advertising budget [28]. This knowledge can be used when determining the 'a good' advertising budget [28].

In the elasticity demand model, the dependent variable is measured for responsiveness which will be represented by the product's volume. The independent variables are the product's own prices, the price movements of the products in the competitor set, the different types of advertising used and seasonality.

The interval variables, both dependent and independent, will be transformed using the natural logarithm, resulting in the elasticity demand model. Binary variables will be used in the model to describe the different advertising types and the seasonality.

The objective of this methodology is to use regression analysis to establish a benchmark for the type of relationships and insights that can be observed from quantitative methods. Regression analysis will be used to draw insights about the significance of the marketing mix being used to drive sales or increase consumer buying behaviour.

Procedural Step	Details
Step 1: Data	Describing and refining variables
Step 2: Model	Regression equation
Step 3: Model assessment	Standard error of estimate, coefficient of determination, F-test of the analysis of variance and VIF
Step 4: Final model	Regression with refined set of independent variables

TABLE 6.1: *Multiple regression procedure decision sequence*

6.1.1   Step 1: Data

As discussed previously SPIRITS-LIQUEURS-P16-750ML will be used as an example to investigate the method and result. The multiple regression model needs to take into account all the forces acting on a product, so that these relationships can be tested for significance and strength. In Tables 5.2 and 5.3 the Pearson's correlation coefficients are displayed for SPIRITS-LIQUEURS-P16-750ML's AUP and the AUP of the products in the competitor set. In the refinement of the competitor set the highly correlated products are removed to hedge against multicollinearity as done previously when determining the competitor set.

Advertising of products within the retail environment can be grouped into two main types: instant prizes and in-store engagement. Instant prizes also known as instant gratification, is when a consumer is rewarded for a purchase with a prize of some sort. Some examples of this type of advertising that ran during the period under investigation are: 'Buy a bottle of SPIRITS-LIQUEURS-P16-750ML and get a free packet of fudge' and 'Buy a bottle of SPIRITS-LIQUEURS-P16-750ML and get a free cocktail recipe booklet'. In-store engagement is not necessary telling the consumer 'buy and get', but rather talking about the brand with the consumer. Some examples of this type of advertising that ran during the period under investigation are the neck tags on the bottle highlighting the brand contributions to animal conservation efforts and in-store TV advertisements.

Another way to advertise in-store is to use gifting. This is to reward loyal consumers and create visibility and excitement so that new consumers will try the product. SPIRITS-LIQUEURS-P16-750ML had many different gifting offerings over the 3-year period. The gifts ranged from glasses, hats, bags to chocolates. The standard SPIRITS-LIQUEURS-P16-750ML bottle is packaged in a box including the gift, usually with a cellophane window for the gift to be visible.

Advertising out-of-store consists of digital, magazine, newspaper, outdoor, radio and TV advertisements. These are all mass-media channels and designed to reach a large audience. The advertising for SPIRITS-LIQUEURS-P16-750ML on these channels highlights the unique intrinsics of the products, its heritage and providence. They do not have a 'call to action' or 'go out and buy' or a pricing message. They are

rather building an emotional connection with consumers and the brand's image. SPIRITS-LIQUEURS-P16-750ML does not have spend on newspaper and radio advertising over the last three years, so these channels will not be included in the multiple regression model.
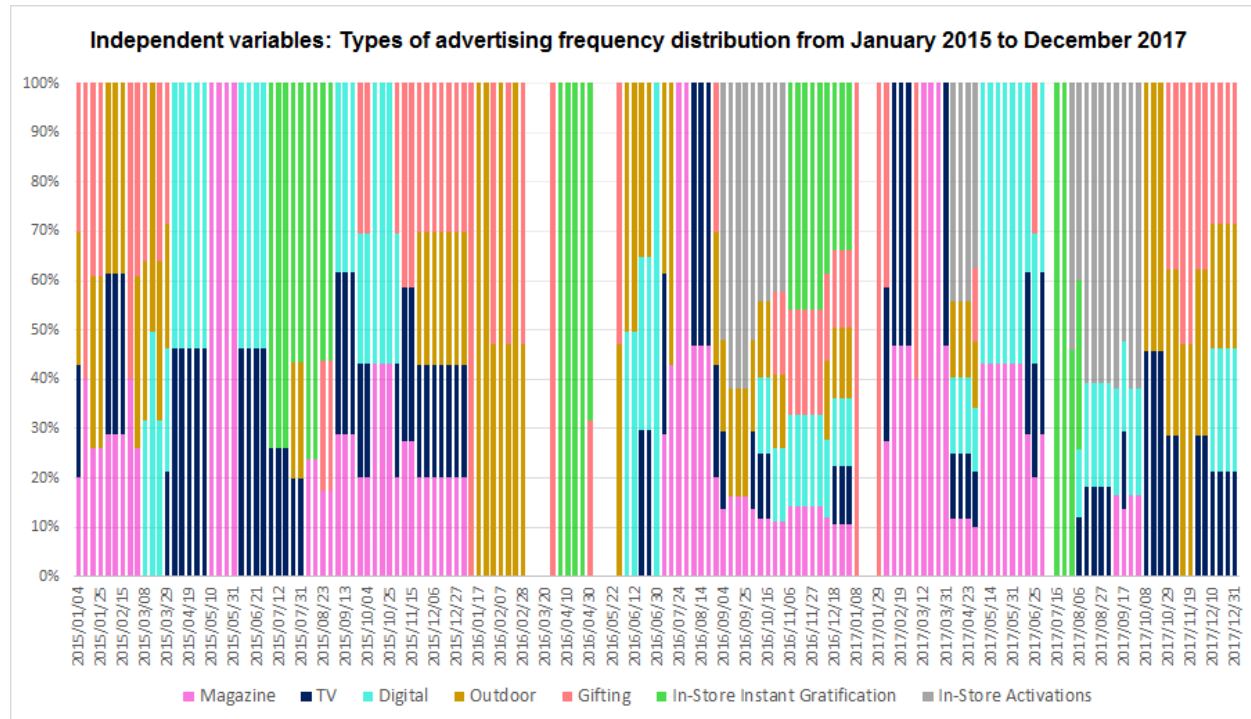
FIGURE 6.2: *Frequency distribution for each type of advertising variable for SPIRITS-LIQUEURS-P16-750ML over the 3-year period*

Figures 6.2 and 6.3 show the frequency distribution and contribution of each type of advertising over the 3-year period. The seasonality of the advertising is shown in Figure 6.2 with gifting occurring mostly over the summer months from the end of October to December. In-store instant gratification promotions happening on days of celebrations for example: Valentine's Day, Easter, Mother's Day and Father's Day. The mass-media advertising mediums have a higher percentage as these types of advertising have a much lower cost per reach than in-store advertising.
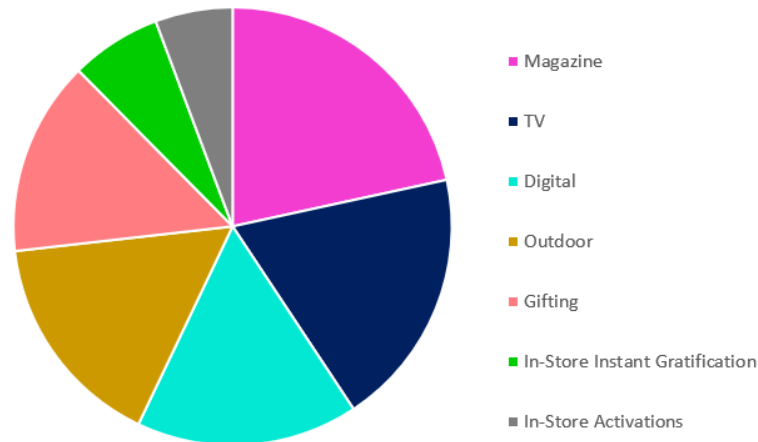
FIGURE 6.3: *Contribution to each type of advertising variable for SPIRITS-LIQUEURS-P16-750ML over the 3 year period*

Seasonality represents a few factors acting together, firstly, the changing seasons and with this comes a change in consumer consumption behaviour. The warmer summer months (November - April) bring more social occasions and traditional festive holidays. In Figure 6.4 the seasonality of the Liqueurs category is clearly displayed with the volumes spiking in the summer months. Seasonality is a strong factor in the alcoholic beverage landscape and needs to be represented in the model. Table 6.2 is an example of one year of data that will be used in the log-log model, using equation (2.21).

FIGURE 6.4: *Sales volume for the liqueurs category,January 2015 to December 2017*

Year	Month	Week ending date	SPIRITS-LIQUEURS-P16-750ML Volume	SPIRITS-LIQUEURS-P16-750ML AUP	SPIRITS-BRANDY-P24-750ML	RTD-BEER-P24-340ML	WINE-BIB-P48-3LTR	RTD-FABS-P22-275ML	SPIRITS-WHISKY-P23-750ML	SPIRITS-LIQUEURS-P9-750ML	SPIRITS-WHISKY-P25-750ML	SPIRITS-LIQUEURS-P12-750ML	SPIRITS-RUM-P6-750ML	SPIRITS-BRANDY-P4-1LTR	WINE-BIB-P44-3LTR	WINE-BIB-P40-3LTR	SPIRITS-LIQUEURS-P8-750ML	RTD-FABS-P1-250ML	RTD-FABS-P23-300ML	SPIRITS-LIQUEURS-P3-750ML	WINE-WHITE-P18-750ML	SPIRITS-LIQUEURS-P5-750ML	SPIRITS-WHISKY-P19-750ML	SPIRITS-BRANDY-P13-750ML	SPIRITS-LIQUEURS-P15-750ML	Gifting	In-Store instant gratification	In-store activations	TV	Outdoor	Magazine	Digital	Autumn	Winter	Spring
2015	1	2015/01/04	L	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	1	0	0	1	1	1	0	0	0	0
		2015/01/11	L	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	1	0	0	0	0	1	0	0	0	0
		2015/01/18	L	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	1	0	0	0	1	1	0	0	0	0
		2015/01/25	L	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	1	0	0	0	1	1	0	0	0	0
		2015/01/31	L	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	O	0	0	0	1	1	1	0	0	0	0
	2	2015/02/08	L	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	O	0	0	0	1	1	1	0	0	0	0
		2015/02/15	L	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	O	0	0	0	1	1	1	0	0	0	0
		2015/02/22	L	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	1	0	0	0	0	1	0	0	0	0
		2015/02/28	L	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	1	0	0	0	1	1	0	0	0	0
	3	2015/03/08	L	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	1	0	0	0	1	0	1	1	0	0
		2015/03/15	L	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	0	0	0	0	1	0	1	1	0	0
		2015/03/22	L	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	1	0	0	0	1	0	1	1	0	0
		2015/03/29	L	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	1	0	0	1	1	0	1	1	0	0
…	……																																		
	10	2015/10/04	L	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	1	0	0	1	0	1	1	0	0	1
		2015/10/11	L	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	O	0	0	0	0	0	1	1	0	0	1
		2015/10/18	L	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	O	0	0	0	0	0	1	1	0	0	1
		2015/10/25	L	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	O	0	0	0	0	0	1	1	0	0	1
		2015/10/31	L	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	1	0	0	1	0	1	1	0	0	1
	11	2015/11/08	L	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	1	0	0	1	0	1	0	0	0	1
		2015/11/15	L	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	1	0	0	1	0	1	0	0	0	1
		2015/11/22	L	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	1	0	0	1	1	1	0	0	0	1
		2015/11/29	L	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	1	0	0	1	1	1	0	0	0	1
	12	2015/12/06	L	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	1	0	0	1	1	1	0	0	0	0
		2015/12/13	L	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	1	0	0	1	1	1	0	0	0	0
		2015/12/20	L	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	1	0	0	1	1	1	0	0	0	0
		2015/12/27	L	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	1	0	0	1	1	1	0	0	0	0
		2015/12/31	L	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	1	0	0	1	1	1	0	0	0	0

TABLE 6.2: *Example of the electronic till data and the dummy variables representing the different types of advertising and seasonality to be used in the regression model to define relationships with SPIRITS-LIQUEURS-P16-750ML*

6.1.2   Step 2: Model

The price elasticity model from equation (2.21) will be used with the addition of binary dummy variables representing the different types of advertising and seasonality to define the relationship that SPIRITS-

LIQUEURS-P16-750ML has with these market factors [81], and the competitor set. The multiple regression model will represent the retail environment and all the different forces that could be acting on a consumer when making a purchasing decision. If a market force has a positive effect on the consumer and this leads to a purchase, this will be represented in the model as an increase in the quantity of SPIRITS-LIQUEURS-P16-750ML sold during that week or period. If a market force does not effect the sale quantity of SPIRITS-LIQUEURS-P16-750ML in the model this will be represented by independent variables not having a significant relationship with the dependent variable SPIRITS-LIQUEURS-P16-750ML.

After the dataset has been refined to remove multicollinearity the 21 independent variables for the competitor set are:

- SPIRITS-BRANDY-P24-750ML,
- RTD-BEER-P24-340ML,
- WINE-BIB-P48-3LTR,
- RTD-FABS-P22-275ML,
- SPIRITS-WHISKY-P23-750ML,
- SPIRITS-LIQUEURS-P9-750ML,
- SPIRITS-WHISKY-P25-750ML,
- SPIRITS-LIQUEURS-P12-750ML,
- SPIRITS-RUM-P6-750ML,
- SPIRITS-BRANDY-P4-1LTR,
- WINE-BIB-P44-3LTR,
- WINE-BIB-P40-3LTR,
- SPIRITS-LIQUEURS-P8-750ML,
- RTD-FABS-P1-250ML,
- RTD-FABS-P23-300ML,
- SPIRITS-LIQUEURS-P3-750ML,
- WINE-WHITE-P18-750ML,
- SPIRITS-LIQUEURS-P5-750ML,
- SPIRITS-WHISKY-P19-750ML,
- SPIRITS-BRANDY-P13-750ML and
- SPIRITS-LIQUEURS-P15-750ML.

The 7 independent variables to be included in the model to represent advertising are: gifting, in-store instant gratification, in-store activations, TV, outdoor, magazine and digital. These variables will be binary variables, a 1 indicates that the advertising medium was active during that week and a 0 indicates that it was not. Seasonality will be represented in the same way with binary variables for Autumn, Winter and Spring. Summer will not be included in the model as this is the reference dummy variable.

Let the dependent variable be the volume sales per week in litres of SPIRITS-LIQUEURS-P16-750ML, Q_{SLP16}, and the first independent variable, P_{SLP16}, be the average unit price of SPIRITS-LIQUEURS-P16-750ML, where β_{SLP16} is the own elasticity of SPIRITS-LIQUEURS-P16-750ML. The competitor set's AUP values per product will be denoted by the set $CS1, CS2, CS3 \ldots CS21$, as per the competitor set products listed in the paragraphs above. Let the set $A1, A2, A3 \ldots A7$ of independent variables represent the marketing factors, also as listed in the paragraph above and let the months of March, April and May be grouped as Autumn and June, July and August represent Winter. Spring is from September to November and Summer being December to February.

Equation (6.1) represents SPIRITS-LIQUEURS-P16-750ML and all the factors acting on the potential consumer when deciding whether to purchase this product or not. The variables β_{CS_i} represent the cross price elasticity of SPIRITS-LIQUEURS-P16-750ML in the competitor set. The variables β_{A_i} represent the relationship that sales volume of SPIRITS-LIQUEURS-P16-750ML has with the different advertising

types for $i = 1, \ldots, 7$. Finally the variables $\beta_{Seasonal_i}$ represent the difference between the percentage sales in a season and the percentage sales in Summer for $i = 1, \ldots, 3$.

$$
\begin{aligned}
\ln Q_{SLP16} =& \beta_0 + \beta_{SLP16} \ln P_{SLP16} + \beta_{CS1} \ln P_{CS1} + \beta_{CS2} \ln P_{CS2} + \ldots + \beta_{CS21} \ln P_{CS21} + \\
& \beta_{A1} Promo_{A1} + \beta_{A2} Promo_{A2} + \ldots + \beta_{A7} Promo_{A7} + \\
& \beta_{Autumn} Seasonal_{Autumn} + \beta_{Winter} Seasonal_{Winter} + \beta_{Spring} Seasonal_{Spring} + \epsilon
\end{aligned}
$$
$$(6.1)$$

In SAS studio® software, the procedures PROC REG can be used to run the multiple regression model. The dependent variable has been transformed using the natural logarithm as per the equation. The independent variables for AUP value of a product in the competitor set has also been transformed using the natural logarithm. In Figure A.8 in the appendix, the code is given to compute the multiple regression analysis.

6.1.3   Step 3: Model assessment

When building a multiple regression model it needs to be assessed to determine if the model fits the data, as this will determine if the results should be used.

In Figure 6.3 the standard error is small (9.97784), therefore being a first indication that this is a good model. The coefficient of determination or R^2 is a test statistic to obtain how well the model fits the dataset. R^2 is a goodness of fit measure and indicates the proportion of the total variation in the dependent variable Y, around its mean that is accounted for by the independent variables in the estimated regression function [57]. This means that 82.02% of the variation in SPIRITS-LIQUEURS-P16-750MLs quantity demanded is explained by the thirty two independent variables. When using multiple regression the adjusted R^2 is a better measure of fit as the number of independent variables is accounted for and the test statistic adjusted. The adjusted R^2 is high, with the model explaining 77.47% of the variation in SPIRITS-LIQUEURS-P16-750MLs quantity demanded.

The F-ratio is used to validate the model and the linear relationships that the dependent and independent variables have. A large value of the F-ratio indicates that most of the variation in the dependent variable is explained by the regression equation and that the model is valid. A small value of the F-ratio indicates that most of the variation in the dependent variable is unexplained. The F-ratio must first be tested for significance to assess if the test statistic can be trusted. In Figure 6.3 the p-value of the F-ratio is significant as it is less than 0.05. Since the F-value is 17.76 and the $F_{Critical} = 1.6835$, and the F-value is significant, it is reasonable to conclude that there is a great deal of evidence to infer that the model is valid.

The Durbin-Watson D test is used to test for first order autocorrelation, this being an indication that observations are time dependent. The Durbin-Watson D value = 1.437, an ideal score would be 2. A Durbin-Watson D value = 1.437 is above 1.414 at a 1% significance level as found in a standard Durbin-Watson table and it is therefore inconclusive if there is autocorrelation in the model.

When it is inconclusive if autocorrelation is present and the dataset contains time series data, then a lag variable can be introduced [81]. The dependent variable, SPIRITS-LIQUEURS-P16-750ML volume is lagged by one week and used as an independent variable. Using a lagged variable rids the model of unwanted biases and hedges against autocorrelation [81].

Source	DF	Sum of Squares	Mean Square	F Value	Pr > |F	Root MSE	Dependent Mean	R^2	AdjR^2	Coeff Var	Durbin-Watson D	1st Order Auto-correlation
Model	32	45.7377	1.4293	17.76	0.0001	0.28367	8.20559	0.8209	0.7747	3.45699	1.437	0.269
Error	124	9.97784	0.08047									
Corrected Total	156	55.71554										

TABLE 6.3: *Analysis of Variance output statistics for SPIRITS-LIQUEURS-P16-750ML multiple regression model*

Heteroscedasticity should not be present in the model as all the interval time series variables were transformed and standardised using the natural logarithm. In Figure 6.5 the residual plot of the predicted values and residuals shows no pattern, affirming the model is not affected by heteroscedasticity.

Multicollinearity, although reduced by removing the highly correlated variables, can still be present in the model. The variance inflation factors can be used to test for multicollinearity between the independent variables in the model. The variance inflation factor compares each variable to the model and scores it for multicollinearity. A VIF score that is less than 2 is a good indication of no multicollinearity. Scores that are between 2 and 10 should be investigated as there may be a problem and scores over 10 are a positive test for multicollinearity [28]. The solution to removing multicollinearity is to remove some of the highly correlated variables.

In Figure 6.4 the VIF scores are displayed for each independent variable. This model has more variables than the rule, 1 independent variable for every 10 observations, therefore, fewer than 15 to 16 independent variables should be ideal. A conservative cut-off of a score of 4 for the VIF value will be used as a threshold for removing independent variables. The independent variables: SPIRITS-BRANDY-P24-750ML, RTD-FABS-P22-275ML, SPIRITS-WHISKY-P23-750ML, SPIRITS-RUM-P6-750ML, RTD-FABS-P1-250ML and RTD-FABS-P23-300ML will be removed.

SPIRITS-BRANDY-P4-1LTR will not be removed due to the significant relationship this product has with SPIRITS-LIQUEURS-P16-750ML, as can be seen by the results of the PLC methodology. SPIRITS-BRANDY-P4-1LTR also has high correlation coefficients with the variables being removed, as can be seen in Tables 5.2 and 5.3 and will be left in the model to represent these competitor set variables.
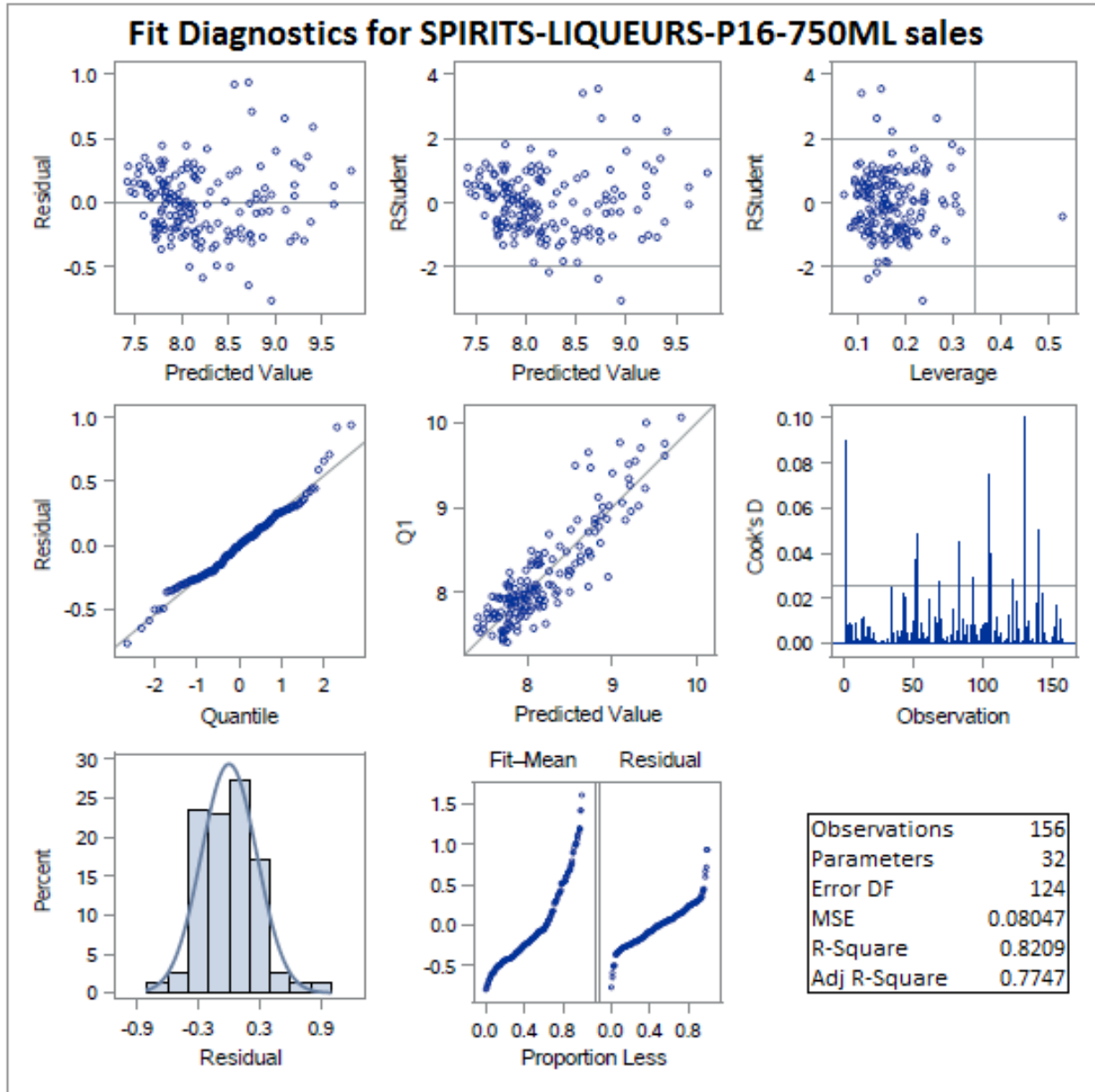
FIGURE 6.5: *Fit diagnostics for SPIRITS-LIQUEURS-P16-750ML multiple regression model*

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t	Variance Inflation Factor
Intercept	1	-17.52944	18.42991	-0.95	0.3434	0
SPIRITS-LIQUEURS-P16-750ML	1	-3.94629	0.32457	-12.16	0.0001	1.81707
SPIRITS-BRANDY-P24-750ML	1	-1.3744	0.61274	-2.24	0.0267	**4.2085**
RTD-BEER-P24-340ML	1	-2.32792	0.76353	-3.05	0.0028	1.74091
WINE-BIB-P48-3LTR	1	0.42378	0.37634	1.13	0.2623	2.28708
RTD-FABS-P22-275ML	1	-1.76315	1.26778	-1.39	0.1668	**5.49247**
SPIRITS-WHISKY-P23-750ML	1	0.34818	0.6426	0.54	0.5889	**4.37217**
SPIRITS-LIQUEURS-P9-750ML	1	-0.97539	0.4708	-2.07	0.0404	1.90504
SPIRITS-WHISKY-P25-750ML	1	0.36302	0.49098	0.74	0.4611	2.27781
SPIRITS-LIQUEURS-P12-750ML	1	-0.21504	0.64969	-0.33	0.7412	3.65814
SPIRITS-RUM-P6-750ML	1	3.42221	1.34636	2.54	0.0123	**7.30373**
SPIRITS-BRANDY-P4-1LTR	1	9.79663	1.26847	7.72	0.0001	5.03775
WINE-BIB-P44-3LTR	1	-0.81559	0.33355	-2.45	0.0159	2.46083
ROBERTSON S/DRY RED 3L	1	-0.56495	0.44837	-1.26	0.21	3.31159
SPIRITS-LIQUEURS-P8-750ML	1	0.5125	0.35872	1.43	0.1556	2.22406
RTD-FABS-P1-250ML	1	-1.11663	0.57857	-1.93	0.0559	**6.30174**
RTD-FABS-P23-300ML	1	2.49807	4.3802	0.57	0.5695	**4.49757**
SPIRITS-LIQUEURS-P3-750ML	1	0.61073	0.59714	1.02	0.3084	2.82743
WINE-WHITE-P18-750ML	1	0.50628	0.40099	1.26	0.2091	2.93998
SPIRITS-LIQUEURS-P5-750ML	1	-0.34736	0.38377	-0.91	0.3672	2.24376
SPIRITS-WHISKY-P19-750ML	1	-0.67395	0.56077	-1.2	0.2317	2.47174
SPIRITS-BRANDY-P13-750ML	1	-0.98593	1.34812	-0.73	0.466	1.65024
SPIRITS-LIQUEURS-P15-750ML	1	0.51729	0.78637	0.66	0.5119	2.43623
Gifting	1	0.11884	0.06295	1.89	0.0614	1.80125
In-Store Instant Gratification	1	0.22883	0.07986	2.87	0.0049	1.77176
In-Store Activation	1	0.21906	0.10066	2.18	0.0314	2.47199
TV	1	0.15682	0.05994	2.62	0.01	1.75168
Outdoor	1	0.07466	0.06385	1.17	0.2445	1.92953
Magazine	1	0.18882	0.06605	2.86	0.005	2.10333
Digital	1	0.14055	0.06382	2.2	0.0295	1.9366
Autumn	1	-0.1055	0.10478	-1.01	0.316	4.06717
Winter	1	-0.24413	0.09817	-2.49	0.0142	3.51042
Spring	1	-0.44611	0.10303	-4.33	0.0001	3.8666

TABLE 6.4: *Parameter Estimates output statistics for SPIRITS-LIQUEURS-P16-750ML multiple regression model*

6.1.4 Step 4: Final model

After the process of refining the model to remove multicollinearity, autocorrelation and overfit there are 26 independent variables. The competitor set of 15 independent variables that will be assessed for price elasticity are:

- RTD-BEER-P24-340ML,

- WINE-BIB-P48-3LTR,

- SPIRITS-LIQUEURS-P9-750ML,

- SPIRITS-WHISKY-P25-750ML,

- SPIRITS-LIQUEURS-P12-750ML,

- SPIRITS-BRANDY-P4-1LTR,

- WINE-BIB-P44-3LTR,

- WINE-BIB-P40-3LTR,

- SPIRITS-LIQUEURS-P8-750ML,

- SPIRITS-LIQUEURS-P3-750ML,

- WINE-WHITE-P18-750ML,

- SPIRITS-LIQUEURS-P5-750ML,

- SPIRITS-WHISKY-P19-750ML,

- SPIRITS-BRANDY-P13-750ML and

- SPIRITS-LIQUEURS-P15-750ML.

The lag variable of SPIRITS-LIQUEURS-P16-750ML sales volume, $QLAG_{SLP16}$ has been added to the model as an independent variable. The binary dummy independent variables representing the different advertising types and seasonality remain unchanged from the model in equation (6.1). The final model is represented by equation (6.2),

$$
\begin{aligned}
\ln Q_{SLP16} =& \beta_0 + \beta_{QLAG_{SLP16}} \ln QLAG_{SLP16} + \beta_{SLP16} \ln P_{SLP16} + \beta_{CS1} \ln P_{CS1} + \beta_{CS2} \ln P_{CS2} + \\
& \ldots + \beta_{CS15} \ln P_{CS15} + \beta_{A1} Promo_{A1} + \beta_{A2} Promo_{A2} + \ldots + \beta_{A7} Promo_{A7} + \\
& \beta_{Autumn} Seasonal_{Autumn} + \beta_{Winter} Seasonal_{Winter} + \beta_{Spring} Seasonal_{Spring} + \epsilon \, .
\end{aligned}
\tag{6.2}
$$

6.2 Results for determining the significance of advertising efforts using regression analysis

The objective of this investigation is to define and score the relationships that SPIRITS-LIQUEURS-P16-750ML has with its competitor set, different types of advertising and seasonality. As discussed in the previous chapter, the model first needs to be assessed for adequacy before the results can be trusted and interpreted.

6.2.1 Model fit

The same process needs to be followed as discussed previously in Section 6.1.3 when evaluating if the multiple regression model is an adequate fit. In Figure 6.5 the analysis of variance output statistics are displayed for the final model.

The R^2 is 81.41%, and the adjusted R^2 is 77.49%, this means that 77.49% of the variation in SPIRITS-LIQUEURS-P16-750MLs quantity demanded is explained by the twenty seven independent variables.

In Figure 6.5 the F-value is significant ($F_{Critical} = 1.55$), as the associated probability is less than 0.05. It is reasonable to conclude that there is a great deal of evidence to infer that the model is valid.

The Durbin-Watson D statistic is 1.589, the lower bound for positive autocorrelation is 1.414 at a 1% significance level. The test statistic is above the lower bound for the confidence interval and it is therefore inconclusive if there is autocorrelation in the model.

Source	DF	Sum of Squares	Mean Square	F Value	Pr > |F	Root MSE	Dependent Mean	R^2	AdjR^2	Coeff Var	Durbin-Watson D	1st Order Auto-correlation
Model	28	43.38147	1.60672	20.77	0.0001	0.27815	8.19563	0.8141	0.7749	3.39394	1.589	0.198
Error	128	9.90335	0.07737									
Corrected Total	156	53.28483										

TABLE 6.5: *Analysis of variance output statistics for the final SPIRITS-LIQUEURS-P16-750ML multiple regression model*

Heteroscedasticity should not be present in the model as all the interval time series variables were transformed and standardised using the natural logarithm. In Figure 6.6 the residuals are plotted showing no heteroscedasticity present in the model. Multicollinearity should also not be present in the model as it was dealt with in the model building method.
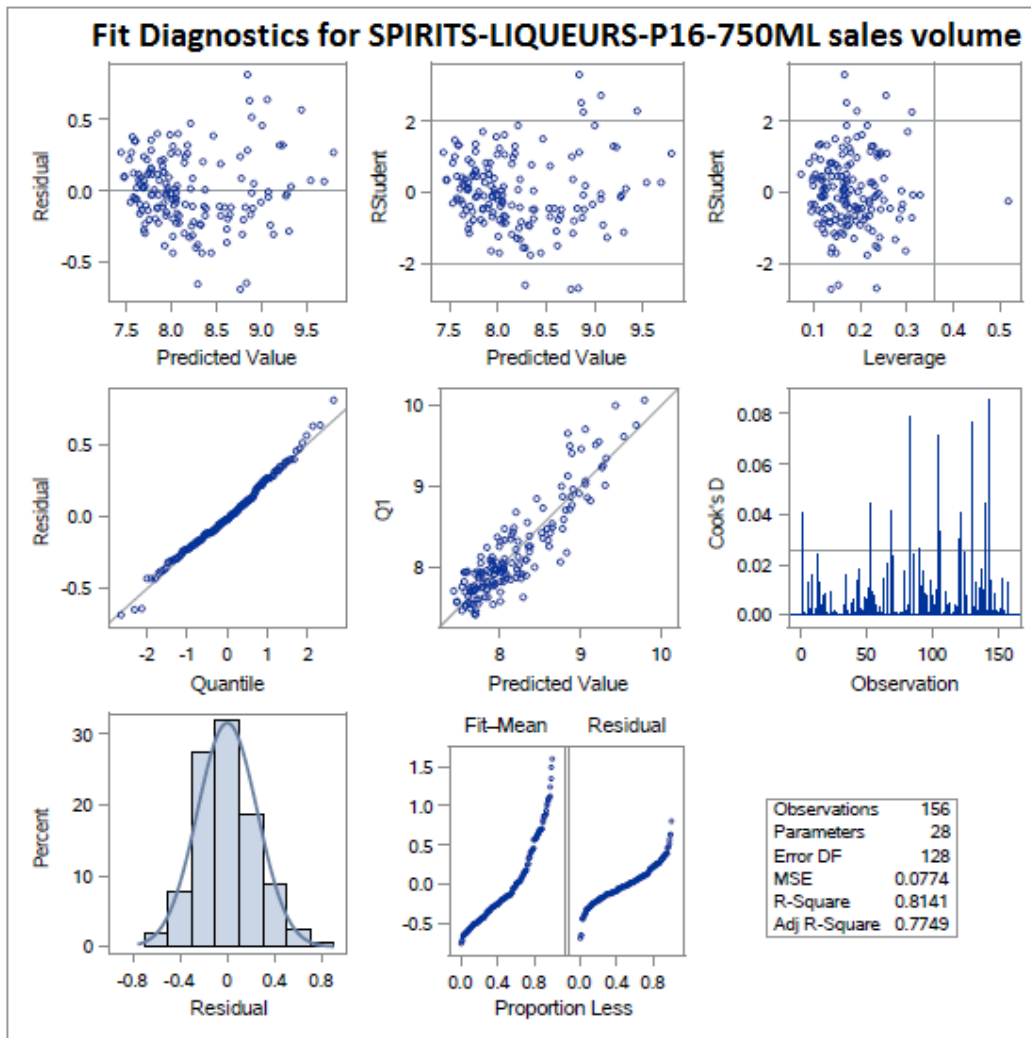
FIGURE 6.6: *Fit diagnostics for SPIRITS-LIQUEURS-P16-750ML final multiple regression mode*

6.2.2 Output statistics from regression analysis

In Figure 6.6 the VIF scores are shown, they are all below 4, this indicating the absence of multicollinearity. The analysis of variance test statistics show that there is reasonable evidence to conclude that the model is valid, now the individual coefficients can be interpreted.

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t	Variance Inflation Factor
Intercept	1	-6.35775	7.79032	-0.82	0.416	0
Lag Q1	1	0.14431	0.04811	3	**0.0033**	1.59503
SPIRITS-LIQUEURS-P16-750ML	1	-3.63582	0.32379	-11.23	**0.0001**	1.77634
SPIRITS-BRANDY-P4-1LTR	1	8.83099	1.00323	8.8	**0.0001**	3.13325
RTD-BEER-P24-340ML	1	-1.11527	0.74133	-1.5	**0.0135**	1.54624
WINE-BIB-P48-3LTR	1	0.23423	0.33885	0.69	0.4907	1.91174
SPIRITS-LIQUEURS-P9-750ML	1	-0.80749	0.44224	-1.83	0.0702	1.72112
SPIRITS-WHISKY-P25-750ML	1	0.53733	0.46362	1.16	0.2486	2.09749
SPIRITS-LIQUEURS-P12-750ML	1	0.8158	0.45906	1.78	0.0779	1.86092
WINE-BIB-P44-3LTR	1	-0.83281	0.3157	-2.64	**0.0094**	2.24798
WINE-BIB-P40-3LTR	1	-0.40053	0.43461	-0.92	0.3585	3.11866
SPIRITS-LIQUEURS-P8-750ML	1	0.2283	0.31666	0.72	0.4722	1.78437
SPIRITS-LIQUEURS-P3-750ML	1	0.67122	0.54893	1.22	0.2237	2.47114
WINE-WHITE-P18-750ML	1	0.17657	0.38225	0.46	0.6449	2.7741
SPIRITS-LIQUEURS-P5-750ML	1	0.82277	0.34616	2.38	**0.0189**	1.89806
SPIRITS-WHISKY-P19-750ML	1	-0.7202	0.47034	-1.53	0.1282	1.79554
SPIRITS-BRANDY-P13-750ML	1	-1.26261	1.25107	-1.01	0.3148	1.47805
SPIRITS-LIQUEURS-P15-750ML	1	-0.29075	0.6922	-0.42	0.6752	1.95653
Gifting	1	0.07518	0.06066	1.24	**0.0431**	1.72019
BTL Instant gratification	1	0.31141	0.07333	4.25	**0.0001**	1.55191
BTL In-store engagement	1	0.05204	0.08999	0.58	0.564	2.05229
TV	1	0.14003	0.0582	2.41	**0.0176**	1.70617
Outdoor	1	0.08993	0.0616	1.46	0.1467	1.85104
Magazine	1	0.11503	0.06088	1.89	**0.0411**	1.84886
Digital	1	0.20914	0.05657	3.7	**0.0003**	1.57519
Autumn	1	-0.1049	0.09268	-1.13	0.2598	3.30189
Winter	1	-0.1932	0.0914	-2.11	**0.0365**	3.15815
Spring	1	-0.32897	0.09254	-3.55	**0.0005**	3.23774

TABLE 6.6: *Parameter estimates output statistics for the final SPIRITS-LIQUEURS-P16-750ML multiple regression model*

In Table 6.6 the output statistics are displayed for the independent variables. The intercept is $\beta_0 = -6.35775$, this is the average demand for SPIRITS-LIQUEURS-P16-750ML when all independent variables are zero. In multiple regression it is often misleading to try to interpret the β_0, especially when the value is zero or less. The lag variable does not need to be interpreted as it has been used to hedge against autocorrelation and not to investigate a relationship.

In Table 6.6 the price elasticity relationships of SPIRITS-LIQUEURS-P16-750ML to it competitor set are displayed. The relationship between SPIRITS-LIQUEURS-P16-750ML quantity demanded and its

own price elasticity is described by $\beta_{SLP16} = -3.64$. If SPIRITS-LIQUEURS-P16-750ML increases its price by 1% then the quantity demanded for SPIRITS-LIQUEURS-P16-750ML will decrease by 3.64%, if all other factors remain constant. There is sufficient evidence at the 5% significance level to infer that the SPIRITS-LIQUEURS-P16-750ML price does have a negative linear relationship with the quantity demanded as the p-value is less than 0.05 and therefore significant.

In Table 6.6 products, WINE-BIB-P48-3LTR, SPIRITS-LIQUEURS-P9-750ML, SPIRITS-WHISKY-P25-750ML, SPIRITS-LIQUEURS-P12-750ML, WINE-BIB-P40-3LTR, SPIRITS-LIQUEURS-P8-750ML, SPIRITS-LIQUEURS-P3-750ML, WINE-WHITE-P18-750ML, SPIRITS-WHISKY-P19-750ML, SPIRITS-BRANDY-P13-750ML and SPIRITS-LIQUEURS-P15-750ML all have p-values that are greater than 0.05 and therefore do not have a significant linear relationship with SPIRITS-LIQUEURS-P16-750MLs quantity demanded.

The competitor set products, RTD-BEER-P24-340ML, SPIRITS-BRANDY-P4-1LTR, WINE-BIB-P44-3LTR and SPIRITS-LIQUEURS-P5-750ML all have p-values less than 0.05 and therefore there is sufficient evidence at the 5% significance level to conclude that each of these products are linearly related to SPIRITS-LIQUEURS-P16-750MLs quantity demanded.

6.2.3 Price elasticities

Using the independent variables parameter estimates from Table 6.6, if RTD-BEER-P24-340ML increase its price by 1% then the quantity demanded for SPIRITS-LIQUEURS-P16-750ML will decrease by 1.12%, if all other factors remain constant. If SPIRITS-BRANDY-P4-1LTR increases its price by 1% then the quantity demanded for SPIRITS-LIQUEURS-P16-750ML will increase by 8.83%, if all other factors remain constant. If WINE-BIB-P44-3LTR increase its price by 1% then the quantity demanded for SPIRITS-LIQUEURS-P16-750ML will decrease by 0.83%, if all other factors remain constant. If SPIRITS-LIQUEURS-P5-750ML increase its price by 1% then the quantity demanded for SPIRITS-LIQUEURS-P16-750ML will increase by 0.82%, if all other factors remain constant.

The price elasticities not only quantify the change in demand for SPIRITS-LIQUEURS-P16-750ML but they also give insight into the relationship SPIRITS-LIQUEURS-P16-750ML has with its competitor set. The products WINE-BIB-P48-3LTR, SPIRITS-LIQUEURS-P9-750ML, SPIRITS-WHISKY-P25-750ML, SPIRITS-LIQUEURS-P12-750ML, WINE-BIB-P40-3LTR, SPIRITS-LIQUEURS-P8-750ML, SPIRITS-LIQUEURS-P3-750ML, WINE-WHITE-P18-750ML, SPIRITS-WHISKY-P19-750ML, SPIRITS-BRANDY-P13-750ML and SPIRITS-LIQUEURS-P15-750ML are all the competitor set of SPIRITS-LIQUEURS-P16-750ML's but do not have a significant effect on the quantity demanded of SPIRITS-LIQUEURS-P16-750ML. Put another way, these products are not causing consumers to change their purchasing behaviour in favour or against SPIRITS-LIQUEURS-P16-750ML.

The products that do have a significant relationship with SPIRITS-LIQUEURS-P16-750ML can now be classed as either substitutes or complementary. The substitutes products have a positive parameter estimate coefficient, SPIRITS-BRANDY-P4-1LTR and SPIRITS-LIQUEURS-P5-750ML and the complementary products have a negative parameter estimate coefficient, RTD-BEER-P24-340ML and WINE-BIB-P44-3LTR.
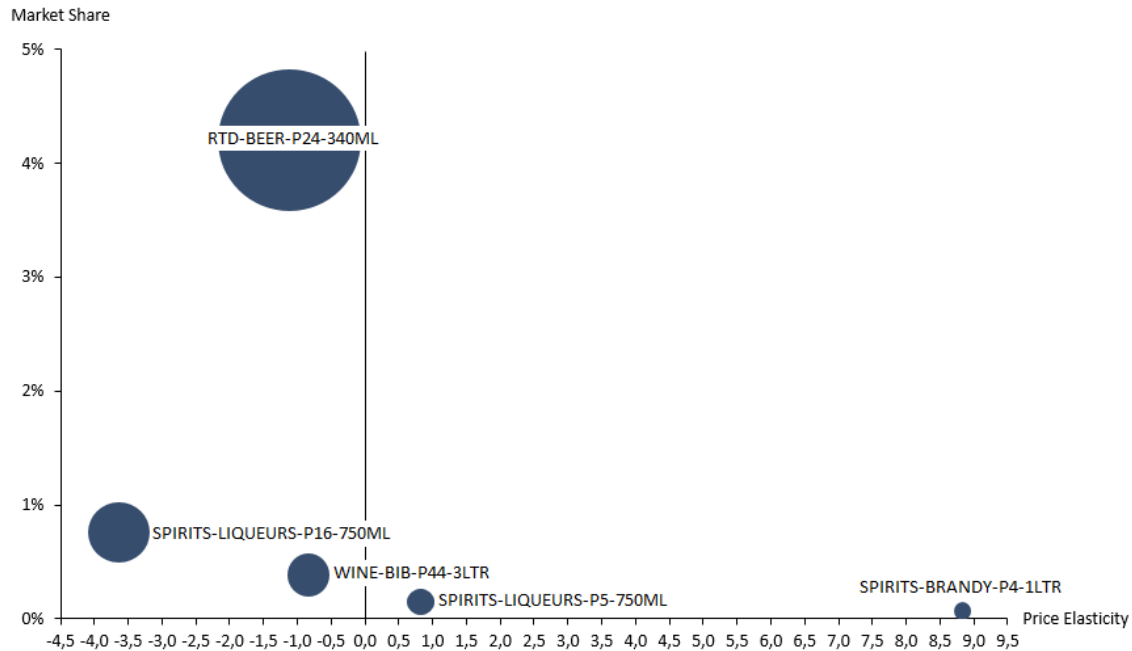
FIGURE 6.7: *SPIRITS-LIQUEURS-P16-750ML and its significant competitor sets price elasticity and market share*

In Figure 6.7 the price elasticities of the significant independent variables are plotted against their market share. The market share is a value share for 2017 for the total alcoholic beverage market in the specialised retailer. The size of the bubble of each product represents the relative sales value for the product for 2017 in the specialised retailer. RTD-BEER-P24-340ML has relatively large market share, large sale value and low elasticity, this shows that the product is 'healthy'. It has a high market share, inferring that either a lot of consumers are purchasing the product or it is being purchased very frequently, possibly both. The low price elasticity is relative to the demand for SPIRITS-LIQUEURS-P16-750ML, this shows that consumers will rather reduce the amount of SPIRITS-LIQUEURS-P16-750ML that they are purchasing than RTD-BEER-P24-340ML. As these products are complements, consumers are purchasing RTD-BEER-P24-340ML and SPIRITS-LIQUEURS-P16-750ML together, if the price of RTD-BEER-P24-340ML increases, the consumer would rather purchase the RTD-BEER-P24-340ML at the higher price and forgo the amount of SPIRITS-LIQUEURS-P16-750ML that they can afford to purchase.

From Figure 6.7 it can be seen that SPIRITS-LIQUEURS-P16-750ML has a higher market share than SPIRITS-BRANDY-P4-1LTR, WINE-BIB-P44-3LTR and SPIRITS-LIQUEURS-P5-750ML. SPIRITS-LIQUEURS-P16-750ML has a complementary relationship with WINE-BIB-P44-3LTR due to the lower market share and low elasticity of WINE-BIB-P44-3LTR it can be inferred that this is a less common complementary product with SPIRITS-LIQUEURS-P16-750ML purchases than RTD-BEER-P24-340ML.

SPIRITS-BRANDY-P4-1LTR and SPIRITS-LIQUEURS-P5-750ML are both substitutes for SPIRITS-LIQUEURS-P16-750ML. SPIRITS-BRANDY-P4-1LTR has a low market share, low sales value and a high price elasticity, these are not indications of a 'healthy' product. This shows that consumers who are purchasing both SPIRITS-BRANDY-P4-1LTR and substituting it for SPIRITS-LIQUEURS-P16-750ML are price sensitive. Price is driving the reason for the purchasing behaviour of consumers. SPIRITS-

LIQUEURS-P5-750ML also has a low market share and sale value but conversely it has a low price elasticity. Consumers are substituting SPIRITS-LIQUEURS-P5-750ML and SPIRITS-LIQUEURS-P16-750ML but they are not as sensitive to price as the SPIRITS-BRANDY-P4-1LTR consumers.

In Figure 6.8 the relationship between the amount spent on advertising using different types of advertising and the demand for SPIRITS-LIQUEURS-P16-750ML for the final model as per equation (6.2) are displayed. The graph shows that in-store instant gratification had the largest effect on increasing sales volume for the amount spent on this type of advertising, all other variables remaining constant. Gifting has the smallest effect on sales for the amount of spend relative to the other advertising types.

This result is in line with other known methodologies for advertising, advertising types that have a strong 'call to action' or 'purchase now' message have the greatest effect on sales. When measuring the increase in sales in this way, the long term gradual effect that advertising has on a consumer's purchases is not captured [39]. T.V advertising tends not to have an immediate impact on sales volumes but as consumers are reached by the advertising repeatedly creating impressions, increasing brand awareness can lead to purchases in the future [39].
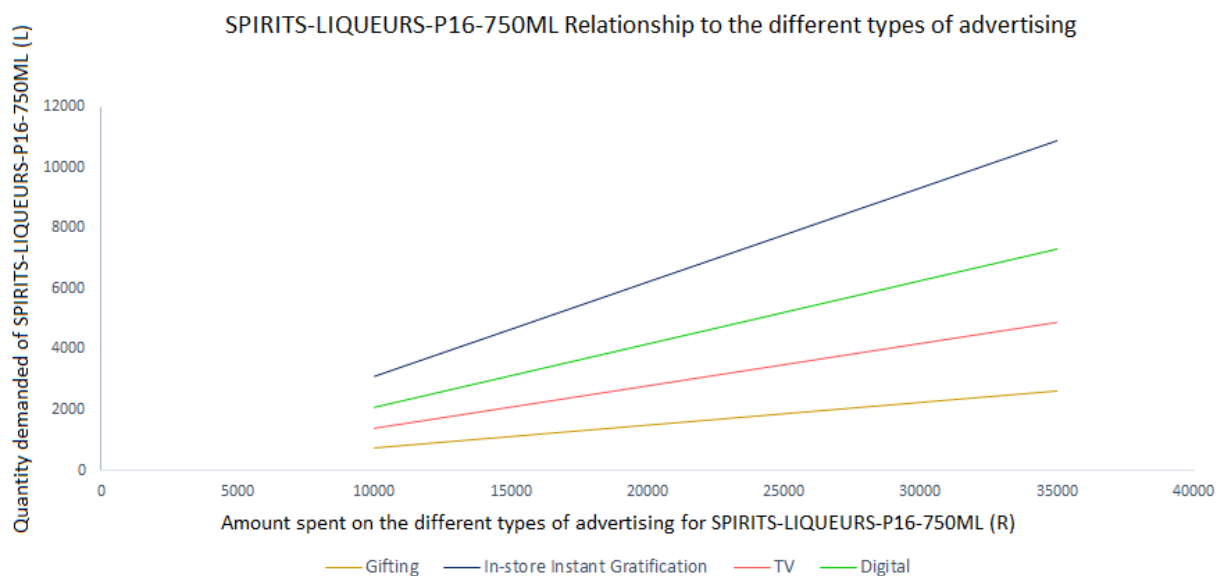
FIGURE 6.8: *The relationship between the amount spent on advertising SPIRITS-LIQUEURS-P16-750ML and the quantity demanded*

In Table 6.6 the relationship of the quantity demanded for SPIRITS-LIQUEURS-P16-750ML and seasonality is displayed showing that Autumn, Spring and Winter have negative parameter estimates, meaning that Summer has the greatest effect on SPIRITS-LIQUEURS-P16-750ML sales volume.

The multiple regression model can be used to forecast the sales volumes of SPIRITS-LIQUEURS-P16-750ML, the model can also be used to test pricing scenarios and the possible effects on sales volumes. The model can also be used to test thresholds for changes in SPIRITS-LIQUEURS-P16-750ML advertising variables. The regression model provides a view of the importance of the different seasons to the product, this information can be used for forecasting or in developing advertising or pricing strategies.

6.3   Method for determining the efficiency of advertising efforts using data envelopment analysis

DEA can be used to categorise products as efficient or inefficient with regards to advertising spend. Efficiency is defined by how well the inputs were used to generate the output relative to all the products included in the model. The inputs for each product will be the advertising types: cinema, digital, gifting, magazine, newspaper, outdoor, TV and radio. These inputs to the model are an aggregation of the spend per advertising type for the calendar year 2016 and 2017. The output for each product is the product's own sales value. The output sales value will be an aggregation of the sales value for each product at the retailer's till point for the calendar year 2016 and 2017. Two separate DEA models will be computed, modelling 2016 and 2017 to gain insights into changes in efficiency over time.

Procedural Step	Details
Step 1: Competitor set	Defining the competitor set
Step 2: Data	Defining inputs and outputs
Step 3: Model	Linear programming

TABLE 6.7: *Data envelopment analysis procedure decision sequence*

6.3.1   Step 1: Competitor set

The refined competitor set used previously in Section 6.1.4 will be used for DEA. This group of refined products is used so that the results of the multiple regression model and the DEA are comparable. The 15 competitor products that SPIRITS-LIQUEURS-P16-750ML will be measured against for efficiency are:

- RTD-BEER-P24-340ML,
- WINE-BIB-P48-3LTR,
- SPIRITS-LIQUEURS-P9-750ML,
- SPIRITS-WHISKY-P25-750ML,
- SPIRITS-LIQUEURS-P12-750ML,
- SPIRITS-BRANDY-P4-1LTR,
- WINE-BIB-P44-3LTR,
- WINE-BIB-P40-3LTR,

- SPIRITS-LIQUEURS-P8-750ML,
- SPIRITS-LIQUEURS-P3-750ML,
- WINE-WHITE-P18-750ML,
- SPIRITS-LIQUEURS-P5-750ML,
- SPIRITS-WHISKY-P19-750ML,
- SPIRITS-BRANDY-P13-750ML and
- SPIRITS-LIQUEURS-P15-750ML.

6.3.2   Step 2: Data

The amount spent on each advertising variable or type will be used as an input to the model. Advertising products within the retail environment or BTL can be grouped into two main types: instant gratification and in-store engagement. Instant prizes or instant gratification for the consumer is when a consumer is

rewarded for a purchase with a prize. Some examples of this type of advertising that ran during the period under investigation are 'Buy a bottle of SPIRITS-LIQUEURS-P16-750ML and get a free tot glass' and 'Buy a bottle of SPIRITS-LIQUEURS-P16-750ML and get R5 free airtime'.

In-store engagement is advertising that is engaging and 'talking' about the brand with the consumer. Some examples of this type of advertising that ran during the period under investigation are neck tags on the bottle highlighting the brands conservation efforts and in-store TV advertisements. This information is provided by Company A from their own internal records. In the DEA model the aggregated spend in Rands for the calendar year 2016 and 2017 on each advertising type will be used as an input.

Gifting is also another in-store type of advertising. Gifting can have many objectives for a brand as it can attract new consumers because of the gift. It can also be used as a way for the brand team to reward loyal consumers and create visibility of the product. The amount of money each brand owner spent on gifting is not known, this would be the cost of the value added to the 'normal stock'. The value added is the cost of the gift item and the gift packaging material. As a proxy for this unknown cost, the sale value at the retailer till point of the product's gift pack will be used. If a brand owner created a large quantity of gifting, then there will be a large amount of gifts that sold at the retailer, if only a small amount was spent on gifting and only a few gifts were created then only a small amount could have been sold. In the DEA model the aggregated spend in Rands for the calendar year 2016 and 2017 on gifting will be used as an input. Company A provides this information from the till point data that they purchase from specialist retailers.

Advertising out of store consists of digital, magazine, newspaper, outdoor, radio and TV adverts. These are all typical ATL mass-media channels and designed to reach a large audience. The advertising content for SPIRITS-LIQUEURS-P16-750ML on these mediums primarily focuses on the brand's unique intrinsics, its heritage and providence. Due to this type of content these ATL adverts are not created with the objective of driving volume but rather to build an emotional connection with consumers. In the DEA model the aggregated spend in Rands for the calendar year 2016 and 2017 will be used. This information is provided by Company A who purchases the information from an industry spend tracker produced by Nielsen [52].

In Figure 6.8 the input and output variables are shown for each product per year. The actual amounts are not displayed as this is sensitive information and Company A has only given permission for the advertising spend values to be used for research but not published. All inputs and outputs are Rand values.

SPIRITS-LIQUEURS-P16-750ML competitor set inputs and output for 2016 or 2017

Product	BTL Instant gratification	BTL In-store engagement	Gifting Value	Digital	Magazine	Newspaper	Outdoor	Radio	TV	Sales Value
SPIRITS-LIQUEURS-P16-750ML	X	X	X	X	X	-	X	-	X	X
RTD-BEER-P24-340ML	X	X	-	X	-	X	X	X	X	X
WINE-BIB-P48-3LTR	-	-	X	X	X	X	-	-	X	X
SPIRITS-LIQUEURS-P9-750ML	X	-	X	X	-	-	-	-	X	X
SPIRITS-WHISKY-P25-750ML	X	X	X	X	X	X	X	X	X	X
SPIRITS-LIQUEURS-P12-750ML	-	-	X	X	X	-	-	X	-	X
SPIRITS-BRANDY-P4-1LTR	X	-	X	X	-	X	-	X	X	X
WINE-BIB-P38 AND P44-3LTR	X	-	-	X	X	X	-	-	-	X
SPIRITS-LIQUEURS-P8-750ML	X	-	X	X	-	X	X	X	X	X
SPIRITS-LIQUEURS-P3-750ML	-	-	X	X	-	-	X	X	-	X
WINE-WHITE-P18-750ML	-	-	X	X	X	X	-	-	-	X
SPIRITS-LIQUEURS-P5-750MLL	X	X	X	X	-	-	-	-	X	X
SPIRITS-WHISKY-P19-750ML	X	-	X	X	X	X	X	X	X	X
SPIRITS-BRANDY-P13-750ML	-	-	-	X	-	-	-	X	-	X
SPIRITS-LIQUEURS-P15-750ML	-	X	X	X	-	-	-	-	-	X

TABLE 6.8: *SPIRITS-LIQUEURS-P16-750ML competitor set inputs and output variables for 2016 or 2017 for data envelopment analysis*

The output that will be used in the model is the sale value, as collected by the retailer's till point for each product and aggregated for the calendar year 2016 and 2017 separately. The sales value is a proxy for consumer buying behaviour, it is assumed that if the advertising or gifting offering is effective then consumers will purchase the product.

Examining the inputs first, it is possible for a brand owner to increase or decrease their spend on an advertising type. It might be limited on how much a brand owner can increase there spend, as the company might have budget constraints but it is theoretically possible. The same for the output of sales value, it is possible for the sales value of a product to increase or decrease. Again the sales value that a product can increase and grow in volume too might be limited due to supply constraints or willing consumers but it is theoretically possible. This is an important concept for DEA when developing the inputs and outputs for the model [57].

6.3.3 Step 3: Model

When building a DEA model, it does not need to be tested for fit or have specific functional form requirements. The DEA model is capable of handling multiple inputs and outputs. It can uncover relationships that remain hidden in other methods [81].

Formulating the linear programming model starts with the objective function [57],

$$MAX_{Product} : \sum_{j=1}^{1} \text{Sales Value}_{Product} w_j \ , \tag{6.3}$$

where the objective is to maximise the sales value of the product in the competitor set being evaluated or the efficiency, where efficiency is the weighted outputs to weighted inputs and restricted to a range of 0 to 1. Sales Value$_{Product}$ represent the output or sales value of the product in the competitor set being evaluated for the calendar year. w_j, be the weight assigned to the output for the product being evaluated. The model will be computed separately for each product in the competitor set, giving each product a turn to be evaluated for efficiency. The inputs and output weights are allowed to vary between models. This allows each product to find weights that will maximise their efficiency score. The models for the calendar year 2016 and 2017 will be computed separately so that the results can be compared.

The objective function is subject to three constraints also described previously in equations (2.46), (2.47) and (2.48). The input variables v_j are BTL instant gratification, in-store engagement, gifting NSV, digital, magazine, newspaper, outdoor, radio and TV advertising. The output variable is the sales value of each product in the competitor set for the calendar year 2016 or 2017.

Microsoft Office 2013 Excel with the Solver add-in enabled will be used to run the model. The DEA model will be executed separately for each product to calculate their relative efficiencies. Figure 6.9 shows an example of how the model will be set up in Excel using Solver. The weighted inputs are the sum of each input multiplied by its weight. As there is only one output for each product this is the multiplication of the output and the weight for the output. The column labelled 'Difference' is the difference between the weighted output and weighted inputs for each product. Using the INDEX function in Microsoft Office 2013 Excel calls each product and set the weighted sum of the inputs for the product equal to 1, this cell is labelled constraints. The INDEX function is also used to calculate the weighted sum of the output for the product being evaluated so that this can be maximised using the Solver add-in.
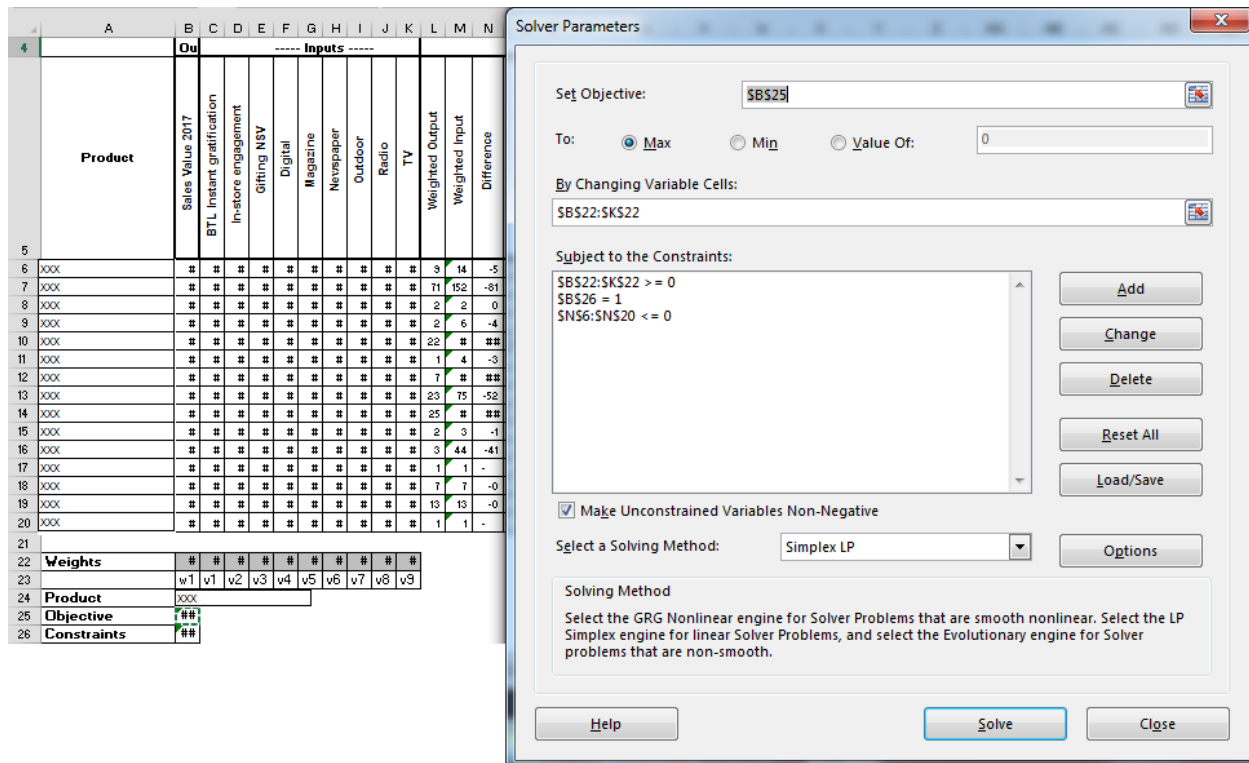
FIGURE 6.9: *Linear Programming using Microsoft Excel Office 2013*

The underlying principal of DEA is to find the efficiently frontier, this means that no other linear combination of inputs could result in at least as much or more output as defined in equation (6.3). This linear combination of the efficient DMUs or products create a composite unit, this provides valuable information to the inefficient DMUs or products on how they can be optimised to become efficient [57]. In this way DEA creates a best practice group that can be used for benchmarking as defined in equation (2.50). When using Microsoft Office 2013 Excel in the Solver results there is an option for sensitivity analysis, this produces shadow variables or composite units that can inform inefficient products on how to be more efficient. The efficient units will become reference units for the inefficient units. The linear combination of the efficient units forms a hypothetical composite units with shadow prices as the weights [57].

6.4 Results for determining the efficiency of advertising efforts using data envelopment analysis

The DEA model does not need to be assessed for adequacy of fit as no functional form is required for the objective function. The results are displayed in Table 6.9 for the two DEA models, one for SPIRITS-LIQUEURS-P16-750ML and its competitor set for 2016 and the other for 2017.

6.4.1 DEA results with 2016 data

The results for 2016 show that RTD-BEER-P24-340ML, WINE-BIB-P48-3LTR, SPIRITS-LIQUEURS-P9-750ML, WINE-BIB-P44-3LTR, WINE-BIB-P40-3LTR, SPIRITS-LIQUEURS-P8-750ML, WINE-WHITE-P18-750ML, SPIRITS-LIQUEURS-P5-750ML, SPIRITS-WHISKY-P19-750ML, SPIRITS-BRANDY-

P13-750ML and SPIRITS-LIQUEURS-P15-750ML are 100% efficient, while the remaining products are operating less efficiently including SPIRITS-LIQUEURS-P16-750ML.

The DEA efficiency measurement does not mean that these products are operating in the most optimal efficient way, it simply means that no linear combination of adversing inputs could produce as much output with the given amount of input relative to the other products in the model [57]. This highlights how important it is to have defined the correct competitor set. In contrast for the inefficient products, SPIRITS-LIQUEURS-P16-750ML, SPIRITS-WHISKY-P25-750ML, SPIRITS-LIQUEURS-P12-750ML, SPIRITS-BRANDY-P4-1LTR and SPIRITS-LIQUEURS-P3-750ML the DEA model will show using the composite units that there is a more optimal mix of advertising expenditure to achieve the same level of output or revenue.

6.4.2 DEA results with 2017 data

In Table 6.9 the results for 2017 show that SPIRITS-LIQUEURS-P8-750ML went from efficient in 2016 to inefficient in 2017. Also, the products SPIRITS-LIQUEURS-P12-750ML, SPIRITS-BRANDY-P4-1LTR and SPIRITS-LIQUEURS-P3-750ML changed from inefficient in 2016 to efficient in 2017. In Table 6.9 it can be seen that SPIRITS-LIQUEURS-P16-750ML and SPIRITS-WHISKY-P25-750ML are both inefficient in 2016 and 2017 in their advertising spend and mix. The spend of both of these products was not reduced from 2016 to 2017. Conversely, the spend on advertising of SPIRITS-LIQUEURS-P12-750ML, SPIRITS-BRANDY-P4-1LTR and SPIRITS-LIQUEURS-P3-750ML which were all inefficient in 2016 was reduced in 2017 and changed to having relative efficient advertising spend.

6.4.3 DEA results comparison of 2016 and 2017 data

In Table 6.9 the results for SPIRITS-LIQUEURS-P8-750ML show that the advertising spend was increased by 49% from 2016 to 2017. The comparison between the results for 2016 and 2017 show that a reduction in advertising spend results in efficiency while an increase in advertising spend results in an inefficiency. The study published by Luo and Donthu [41] and that of Hezekiah, Ramakhrishnan and Shaban [32] support this finding that most advertisers are over investing and that reducing their advertising spend would not result in a reduction in revenue, but rather increased efficiency.

Product	2016 Efficiency	2016 Result	2016 Spend Ranking	2017 Efficiency	2017 Result	2017 Spend Ranking	% Change in spend 2016 - 2017
SPIRITS-LIQUEURS-P16-750ML	0.88	Inefficient	3	0.64	Inefficient	4	0%
RTD-BEER-P24-340ML	1.00	Efficient	2	1.00	Efficient	2	14%
WINE-BIB-P48-3LTR	1.00	Efficient	14	1.00	Efficient	15	-84%
SPIRITS-LIQUEURS-P9-750ML	1.00	Efficient	7	1.00	Efficient	7	-30%
SPIRITS-WHISKY-P25-750ML	0.52	Inefficient	1	0.46	Inefficient	1	10%
SPIRITS-LIQUEURS-P12-750ML	0.05	Inefficient	8	1.00	**Efficient**	10	-56%
SPIRITS-BRANDY-P4-1LTR	1.00	Inefficient	5	1.00	**Efficient**	5	-19%
WINE-BIB-P38 AND P44-3LTR	1.00	Efficient	13	1.00	Efficient	12	69%
SPIRITS-LIQUEURS-P8-750ML	1.00	Efficient	4	0.26	**Inefficient**	3	49%
SPIRITS-LIQUEURS-P3-750ML	0.50	Inefficient	9	1.00	**Efficient**	9	-36%
WINE-WHITE-P18-750ML	1.00	Efficient	15	1.00	Efficient	14	93%
SPIRITS-LIQUEURS-P5-750MLL	1.00	Efficient	10	1.00	Efficient	8	9%
SPIRITS-WHISKY-P19-750ML	1.00	Efficient	6	1.00	Efficient	6	-16%
SPIRITS-BRANDY-P13-750ML	1.00	Efficient	12	1.00	Efficient	13	-20%
SPIRITS-LIQUEURS-P15-750ML	1.00	Efficient	11	1.00	Efficient	11	0%

SPIRITS-LIQUEURS-P16-750ML and its competitor sets results for 2016 and 2017 DEA

TABLE 6.9: *SPIRITS-LIQUEURS-P16-750ML and its competitor sets results for 2016 and 2017 data envelopment analysis*

In Table 6.9 only the relative changes in spend can be shown and discussed as per the consent of Company A. For SPIRITS-LIQUEURS-P16-750ML the advertising spend was increased between 2016 and 2017 on, BTL instant gratification, in-store engagement, gifting and digital while reducing spend on magazine and TV advertising. SPIRITS-LIQUEURS-P8-750ML's overall advertising spend from 2016 to 2017 grew, spending more on BTL instant gratification, digital, outdoor, radio and TV while decreasing spend on gifting and newspaper advertising.

The products that reduced their over-all advertising spend between 2016 and 2017 became efficient, starting with SPIRITS-LIQUEURS-P12-750ML which had an over-all reduction in spend on all its advertisings types, gifting, digital and magazine. SPIRITS-BRANDY-P4-1LTR shrunk down on its spending year-on-year by 19% by lowering spend on digital, newspaper and TV advertising while increasing gifting spend. SPIRITS-LIQUEURS-P3-750ML minimised advertising spend on BTL instant gratification, in-store engagement and digital while, increasing advertising spend on gifting, outdoor and radio with an overall reduction in year on year spend of 36%.

One of the benefits of using DEA is the resulting information provided by the composite units or shadow prices. In Table 6.10 the hypothetical Shadow prices for SPIRITS-LIQUEURS-P16-750ML are shown for 2016 and 2017 respectively.

Product	2016 Shadow prices	2017 Shadow prices
SPIRITS-LIQUEURS-P16-750ML	0.00	0.00
RTD-BEER-P24-340ML	0.00	0.00
WINE-BIB-P48-3LTR	0.00	0.00
SPIRITS-LIQUEURS-P9-750ML	0.00	0.00
SPIRITS-WHISKY-P25-750ML	0.00	0.00
SPIRITS-LIQUEURS-P12-750ML	0.00	0.00
SPIRITS-BRANDY-P4-1LTR	0.00	0.04
WINE-BIB-P38 AND P44-3LTR	0.00	0.00
SPIRITS-LIQUEURS-P8-750ML	0.00	0.00
SPIRITS-LIQUEURS-P3-750ML	0.00	0.00
WINE-WHITE-P18-750ML	0.01	0.00
SPIRITS-LIQUEURS-P5-750ML	0.42	0.27
SPIRITS-WHISKY-P19-750ML	0.00	0.01
SPIRITS-BRANDY-P13-750ML	0.00	0.00
SPIRITS-LIQUEURS-P15-750ML	0.57	0.68

TABLE 6.10: *Shadow prices for SPIRITS-LIQUEURS-P16-750ML relative to its competitor sets for 2016 and 2017*

As the actual rand value of the different input and output variables can not be disclosed, the percentage change from the initial Rand values to the optimal hypothetical composite unit Rand values will be discussed. Therefore the composite unit can not be shown.

In 2016 the hypothetical composite unit for SPIRITS-LIQUEURS-P16-750ML has an overall spend reduction of 26%. The composite unit for SPIRITS-LIQUEURS-P16-750ML in 2016 has a 55% reduction in BTL instant gratification spend, a 77% reduction in in-store engagement advertisings spend, a 12% reduction in gifting spend, a 89% reduction in digital advertising spend, the removal of all spend on magazine advertising, a 40% reduction in outdoor spend and finally a 20% reduction in TV advertising spend. Newspaper and radio advertising did not have any spend for SPIRITS-LIQUEURS-P16-750ML and therefore no recommendation is provided for these advertising types. The sales value will remain unchanged for the composite unit.

In 2017 the findings for SPIRITS-LIQUEURS-P16-750ML are very similar to the previous year. The results for SPIRITS-LIQUEURS-P16-750ML in 2017 show the weighted average of the input in a ratio of, 4% of SPIRITS-BRANDY-P4-1LTR, plus 27% of SPIRITS-LIQUEURS-P5-750MLL, plus 1% of Three Ships Select 750M, plus 68% of SPIRITS-LIQUEURS-P15-750ML produces a hypothetical composite unit for SPIRITS-LIQUEURS-P16-750ML with equal output but less inputs require.

In 2017 the hypothetical composite unit for SPIRITS-LIQUEURS-P16-750ML has an overall spend reduction of 42%, while he sales value will remain unchanged. The composite unit for SPIRITS-LIQUEURS-P16-750ML in 2016 is very similar to the results for 2017. The composite unit in 2017 has a 36% reduction in BTL instant gratification spend, a 88% reduction in in-store engagement advertisings spend, a 36% reduction in gifting spend, a 86% reduction in digital advertising spend, again the removal of all spend on magazine advertising, a 28% reduction in outdoor spend and finally a 36% reduction in TV advertising spend. Similarly newspaper and radio advertising did not have any spend

for SPIRITS-LIQUEURS-P16-750ML in 2107 and therefore no recommendation is provided for these advertising types.

Type of advertising	2016 % Reduction in spend	2017 % Reduction in spend
BTL instant gratification	55%	36%
In-store engagement	77%	88%
Gifting NSV	12%	36%
Digital	89%	86%
Magazine	No spend	No spend
Newspaper		
Outdoor	40%	28%
Radio		
TV	20%	36%
Total	26%	42%

TABLE 6.11: *Reduction in advertising spend required when using the composite units for SPIRITS-LIQUEURS-P16-750ML in 2016 and 2017 to calculate the optimal spend*

The DEA method using the composite unit has created a clear benchmark for the product SPIRITS-LIQUEURS-P16-750ML. The results giving clear guidelines on how much should be spent in total and the ratio of the advertising mix.

The results for SPIRITS-LIQUEURS-P16-750ML between 2016 and 2017 are very similar, in that the product remains inefficient and has a similar pattern of advertising spend. The findings published by Cheong, De Gregorio and Kim [14] also found that the overall level of advertising spending efficiency among the top 100 advertisers will decrease over time. Firms that are 'overspending' tend to continue overspending [14]. The DEA methodology can be employed to stop the cycle of overspending and assist in finding the optimal spend level.

6.5   Chapter summary

Sub-chapter 6.1 describes the dataset that will be used to compute the multiple regression model with dummy variables. The independent variables are refined to remove multicollinearity, autocorrelation and overfit from the final model. The final model has 26 independent variables representing the competitor sets AUP and the seven different advertising types and seasonality.

Sub-chapter 6.2 provides insights into the significance of advertising efforts using multiple regression analysis. The adequacy of the model is reviewed, as well as the output statistics. The model is found to be significant and a good fit to the data. The results show that SPIRITS-LIQUEURS-P16-750ML has a complementary relationship with the products RTD-BEER-P24-340ML and WINE-BIB-P44-3LTR and a substitute relationship with the products SPIRITS-BRANDY-P4-1LTR and SPIRITS-LIQUEURS-P5-750ML. The results show that in-store instant gratification had the largest effect on increasing sales volume for the amount spent. The advertising variable in-store engagement and outdoor did not have significant relationships with SPIRITS-LIQUEURS-P16-750ML quantity demanded and are therefore not

effective types of advertising spend for SPIRITS-LIQUEURS-P16-750ML and the quantity demanded in specialised retailer outlets.

The seasonal variables showed that summer is a very important time for SPIRITS-LIQUEURS-P16-750ML to reach consumers and drive up its sales volume.

Sub-chapter 6.3 described the dataset that will be used for DEA and the set-up of the DEA model and the process used to complete the DEA model.

In Sub-chapter 6.4 the results for the DEA methodology are examined, finding SPIRITS-LIQUEURS-P16-750ML to be over investing in advertising spend in both 2016 and 2017. The composite unit produced from the models give a clear guideline on how advertising spend should be reduced in the marketing mix in order for the advertising spend to be more efficient. In the next chapter the result of both the price elasticity model with dummy variables and data envelopment analysis will be compared and discussed.

CHAPTER 7

Discussion of the price elasticity model with dummy variables and DEA

Contents

The overarching objective of this research is to investigate methods that require less detailed historical time series data and can therefore be updated more often to quantifying the relationship that marketing activities have on brand revenue growth. Another objective is to produce results that create actionable insights that can be used in marketing strategy. This means to be able to quantify the full mix of marketing activities and determine the optimal use to generate revenue growth for a brand altering the consumers' buying behaviour.

7.1   A comparison of the price elasticity model with dummy variables and the DEA method

DEA is a method that does not require detailed historical time series data and it is simple to update the model and compute it. The DEA model can have variables of any scale as long as they are consistent across the products, these variables do not need to be transformed and standardised. In Table 7.1 a comparison of the price elasticity model with dummy variables and DEA methodologies focusing on the two

models inputs, processes and results is displayed.

In both methods the internal and external market factors need to be represented by variables and the competitor set. The regression analysis output statistics will give an indication if the model is not built adequately to represent the products' environment that is being modelled, DEA does not have statistical output measures and therefore no indication of the significance of the results. If an important variable is left out of the DEA model the results will only reflect the best possible outcome from the input given and not indicate that something is missing.

The price elasticity model with dummy variables	Data envelopment analysis
Input data required	
- Long periods of historical time series data and binary variables	- Interval (Ratio could also be used)
- On a detailed level	- Aggregated
- Data transformation required to bring the variables to the same scale	- Variables can be of any unit
- At least 3 years of weekly data	- Any time period consisted across the variables
- Price fluctuations and changes in advertising spend and type are needed to determine relationships	- No special data requirements
- Limited number of variables can be included in the model (1 for every 10 observations)	- With caution, sensitivity to outliers, large amount of variables can be included
Model set-up	
- The independent variables need to describe the internal and external market factors	- The external market factors are taken into account by the model due to the results being relative to the other products
- One model for the time period under investigation	- Numerous models are possible for creating comparisons and comparing change over selected time periods
Results	
- Elasticity coefficients describing the strength of the variables' relationships	- Relative efficiency scores for each product
- P-values give meaning to significance of relationships	- Composite units to inform the most efficient marketing mix
- Statistical output indicators and measures	- Optimal values for each inefficient product's marketing mix in the variables scale
- Results can be difficult for a marketer to interpret	- Results are simple to understand and interpret
Actionable insights	
- Information about relationships	- Benchmark spend per variable determined
- Forecasting model and scenario testing	- Scenario testing

TABLE 7.1: *Comparison of the price elasticity model of with dummy variables and data envelopment analysis methodologies focusing on the two models' inputs, processes and results*

The multiple regression model requires a long period of historical time series data, this means that updating the model frequently will not yield largely significant changes in findings. The DEA model can be updated frequently as any period of time can be used as long as it is consistent across the variables.

The DEA model can be used to model different marketing strategies and campaigns. The results will be examined to evaluate if they provide insight into quantifying the full marketing mix and determining an optimal level of spend to drive revenue growth.

The results from the price elasticity model with dummy variables gives insight into the effectiveness of the different marketing types. If a type of advertising has a significant effect on the sales of the product then it is effective in changing the quantity demanded of the product. This gives the marketer a directional starting point for developing their marketing strategy, they know which types of marketing to use or not to use.

The absolute size of the elasticity coefficient also gives scale or strength to the relationship between the advertising variables and the quantity demanded of the product. The marketer can use this in their strategy as a guide for the ratio in which to split there advertising budget between the marketing types.

The DEA method gives clear actionable insights into the optimal marketing mix. The use of the composite units produces a benchmark efficient mix of spend per advertising type that will yield the maximum quantity demanded for the product. The marketer now has the most efficient marketing mix, this information can also be used as a ratio to fit the marketers budget constraints.

The DEA method uses all the marketing types that the competitor set used in the make-up of the model, this gives the final optimal marketing mix the 'opportunity' to find a variable that the product under investigation did not use or spend money on as par of the final result. The price elasticity model with dummy variables will only give insight into the relationships that the product under investigation has with advertising variables that it has used or spent money on in the past.

7.2   The optimal marketing mix

Marketers desire to know an optimal marketing mix in order to be both effective and efficient. Marketers strive to create effective advertising that resonates with the target audience, driving a change in consumer buying behaviour and increase sales volume while simultaneously being efficient in advertising spend, such that the advertising generates more revenue than spend on advertisings [39].

In the price elasticity model with dummy variables effectiveness is measured by the independent variable having a significant relationship with the dependent variables, the strength of the relationship is indicated by the absolute value of the coefficients. This relationship measures the responsiveness of the quantity demanded for a product due to changes in the price and advertising variables of the competitor set.

The DEA method is a measure of the relative efficiency, the inputs, the amount spent on each advertising variable, in relation to the output variables, the sales value. The DEA method produces an optimal marketing mix that is easily understandable and implementable by marketers.

7.3   Chapter summary

The price elasticity model with dummy variables and the DEA method have been described, applied and the results investigated and discussed. The findings are that while the price elasticity model with dummy variables does give insight into the relationships that change consumer buying behaviour or the quantity demanded for the dependant variables, DEA provides an optimal marketing mix.

CHAPTER 8

Conclusion

Contents

This study has investigated the effectiveness and efficiency of branded advertising in specialist retailer outlets in the alcoholic beverage industry of South Africa. After examination the landscape in South Africa was found to be highly competitive with a multitude of diverse brands and products all competing for market share. A somewhat unchanging market place with the the country's politics causing large monopolies with little international competition. This stable environment has rapidly changed over recent years with major international companies like AB InBev taking over SABMiller, the largest alcohol producer in South Africa, generating fierce competition.

The advertising landscape has also undergone changes with the advent of the 'digital revolution'. These changes have brought new ways for marketers to connect with consumers but it has also given consumers the ability to research and compare products far more easily [39]. These changes have increased the rivalry in the market place and the need for marketers to have efficient marketing strategies for their brands or products to be successful in this shifting competitive environment.

Companies in the South African alcoholic beverage industry are spending millions of Rands on advertising per year [52] but this landscape has been found to be difficult to analyse and generate insights for marketing strategies. A cause of this difficulty was identified as the clutter or noise in the market, there are a vast array of brands and products available, all changing prices and spending on different advertising types at different times, making branding or marketing decisions difficult. Due to the competitive nature of the industry, information on marketing strategies and advertising spending was often found to not be available.

This study has investigated the use of the PLC methodology and the DEA method in order to identify methods that can be updated frequently that do not need long periods of time series historical data and can generate actionable insights that will drive an efficient marketing strategy. Regression analysis has been established as the benchmark methodology. The literary study shows that regression analysis is a commonly used method. Regression analysis does have limitations, one being that the method requires

long periods of historical time series data, but a common finding across the research is that availability of long periods of historical time series data is rare in this industry.

This investigation has shown the PLC methodology and DEA can be used that take the cluttered landscape and data availability into consideration while determining the best possible mix of advertising types to use in order to increase a brand's revenue.

8.1   Determining a competitor set

In order to change consumer buying behaviour, the 'need' that your brand or product fills must be determined as well as the group of other brands or products that are for-filling the same 'need'. This information will be valuable in building a marketing strategy. The PLC methodology was investigated as a method that can segment the entire market into competitor sets and does not require long periods of historical time series data. The PLC methodology involves several steps including factor analysis, cluster analysis and the BCG growth share matrix.

The PLC methodology was applied to categorize the products into four life cycle stages. Therefore providing marketers information that can be used to decide on the scale of investment needed for a brand or product. The result of this methodology also gives insight into the consumers buying behaviour and sentiment towards a brand or product. As the full market is segmented information about competitor brands and products are also reviled.

Factor analysis was used successfully to generate eight factor constructs and uncorrelated factor scores. The input variables to factor analysis were highly correlated and numerous, this method achieved the goal of creating uncorrelated factor scores that represented the variables' volume, value, AUP and the brand-health scores. The method did not require long periods of time series data, any period of data could be used as long as it was consistent across the variables. The factor constructs formed gave insight into how the attributes or variables are grouped together and what underling construct do consumers associate products with.

The constructs formed were: Enjoyment, Unreachable, Dislike, Isolation, Popular, Accessibility, Availability and Un-stylish. These constructs can give marketers valuable insights into what they need to focus on in their brand strategies. Factor analysis is simple to set up and apply and can be updated frequently with the latest data. This method allocates all products on the market a factor score and by virtue a highest association to a factor, giving marketers knowledge not only about their brand or product but also about competitor products and the make-up of the market landscape as a whole, as presented in Chapter 4.3 and Figure 4.14. This knowledge gives a marketer insight into which factor or construct the products with the highest market share are associated with, showing where consumer interests, preferences or dislikes are, therefore shaping consumer buying behaviour.

Cluster analysis was used to group the products based on similarities in the variables describing pack size, alcohol percentage, carbonation level, sugar content, taste profile and the eight factor scores. This method formed 10 clusters and 16 sub-clusters, after testing the results they were found to be valid. The cluster analysis procedure provided important groups that were used to determine each product's relative competitor for the BCG growth share matrix methodology. Similarly to factor analysis, the cluster

analysis procedure did not require long periods of time series data, any consistent period of data could be used and be frequently updated.

Cluster analysis produced groups that can also add further insight into the alcoholic beverage landscape and information on brand-health of the different products. The full market is segmented in cluster analysis and the groups all have defining features. In Chapter 4, Table 4.19 the defining attributes for each group are tabulated. Cluster 1 for example has 'convenience' as a defining characteristic for this group of products. All the products are in packaging that does not require them to be decanted into a glass, they are a single serving. Marketers who are developing strategies for products in Cluster 1 will now have insight into the importance of 'convenience' for consumers who are purchasing their product and competitor products as an input to their marketing strategy.

Factor and cluster analysis are useful methods on their own as they overcome the problems of noise, clutter in the market, data availability and accuracy on a detailed granular level. Both methods can be used to add insights to marketing strategy, the insights are not necessary 'actionable' insights but should be valuable in building a marketing strategy based on a better understanding of the market and consumer buying behaviour and sentiment.

The BCG growth share matrix methodology was applied to bring the external market factors into the PLC methodology. As with factor and cluster analysis this method segments the full market and the procedure did not require long periods of time series data, any consistent period of data could be used and be frequently updated. The results were found to be consistent with the distribution expected for the industry.

The BCG growth share matrix produced results that highlighted categories like white spirits that are experiencing high growth, the results also highlighted products like SPIRITS-CANE-P2-750ML that is in the 'Dogs' quadrant. The result show that while all the other white spirits are growing, the Cane category is declining. Mainstay Cane did not make the cut-off to be included in this study during the data refinement process, it was however reported in, 'Die Burger' newspaper in Stellenbosch on 21 December 2018 [3], that the brand would not be produced any more due to poor sales in the Cane category. Marketers using the BCG growth share matrix can have insight into these shifts in consumer buying behaviours and develop strategies to either re-stimulate consumers' interest or migrate consumers to another product.

The PLC methodology and its sub-methods provide a wealth of information that can be used as input to marketers about the alcoholic beverage landscape and all the different products available. The PLC methodology provides levels of detail on the competitor sets, with the four quadrants that can also be segmented into smaller groups by cluster or factor. When each life cycle stage was examined it was found that the different segments have unique shaped PLC curves, similar to the findings of Sethuraman [69].

The RTDs segment showed a large amount of products in the introductory and growth phases creating PLC curves that are very steep. As to be expected from a steep PLC cure, this category also has a large number of new innovations, new entrants, pack extensions and variants. The landscape looks to currently be shaped by marketing strategies that are focused on launching new brands and keeping consumers interested in 'old' brands with new packaging or range innovation.

The spirits segment also has a steep PLC curve but with a spread of products in the growth and decline

life cycle stages. The level of innovation in this segment is high with lots of products in the introduction stage but the number of products taper off along the curve, showing that not all products are able to move on from the introduction stage. This gives insight into the level of competition in the spirits segment. The products in the maturity life cycle stage tend to have a large market share relative to the segment and is an indication for marketers that getting a product through the introduction life cycle stage can yield reward.

The wine segment is the opposite of the RTDs segment with most products being in the maturity and decline life cycle stages. Wine products have long gradual PLC curves, showing a segment with little innovation where products are either long established or accepted quickly by the market and move rapidly to the maturity life cycle stage.

The price elasticity model was applied to represent a product's relationship with other products and consumers' buying behaviour as a benchmark to the PLC methodology. The method did require long periods of historical time series data and can not be updated frequently to yield dramatically different results due to the long periods of data that forms part of the model.

The methodology was not simple to set up and had many data refinement steps in order to compute the final model. The results did not segment the full market but only the product under investigation, this is a limitation of the method as the results are not in context of the full market. It will be very arduous to calculate the own price elasticities for each product.

The results were validated and found to be significant, isolating the products that have a relationship with the product under investigation and ranking the strengths of those relationships. The price elasticity of a product and those of its competitor set can be valuable information to have when building a marketing strategy. Products that have a relatively high own price elasticity are indicative of products that are sensitive to changes in price, a small change in price can result in a large shift in demand. This also gives insight into the product's brand-health, showing that it might be weak as consumers will easily switch in and out of this product due to price changes. Conversely for a product with low price elasticity, consumers are loyal and changes in price will not greatly shift demand. For example, for a product with a high price elasticity, marketers can apply this knowledge to their brand's strategy, possibly focusing on gaining consumer loyalty or exploiting the high elasticity and dropping the price to gain volume.

The results from the process of determining a competitor set were used in the subsequent sections to define the relationships that variables have on the sales uplift of a brand representing a change in consumer buying behaviour. In the next section both the DEA method and multiple regression analysis with binary dummy variables are sensitive to giving insignificant results if the incorrect competitor set are modelled. DEA is particularly susceptible to this as the method creates a ratio of relativity to the products included in the model. Due to the sensitivity of the methods to follow and as both the PLC methodology and price elasticity methodology use fundamentally different approaches to determining a competitor set, they take different market factors into account and model different types of relationships, the results showed it was best to be conservative and combine the two methodologies to obtain a final the competitor set.

8.2 Defining the relationships that variables have on the sales uplift

After the products' competitor sets was defined, an investigation into consumer buying behaviour was performed using the relationship that the products in the competitor set, the different advertising types and market factors have on sales value. Multiple regression analysis with binary dummy variables was found in the literature to be a commonly used method and therefore was used as a benchmark, this methodology was compared to DEA. The problem of having limited data available can be addresses by using the DEA techniques and DEA provides actionable insights. The results for the DEA method were consistent with the expectations from the literature, finding that this model was simple to construct and produced an optimal marketing mix.

Multiple regression analysis with binary dummy variables was used to quantify if a product's quantity demanded had a significant relationship with price changes in the competitor set, the different types of advertising used and seasonality.

The multiple regression analysis with binary dummy variables required long periods of time series historical data on not only the quantity demanded and pricing of the product and its competitor set but also the advertising types used. Long periods of time series data on when each brand or product flighted a T.V advertisement or had a promotion in store is not easily available. This is a limitation of this methodology.

After applying the regression model the results showed that SPIRITS-LIQUEURS-P16-750ML the product being used for this investigation had a significant relationship with SPIRITS-BRANDY-P4-1LTR and SPIRITS-LIQUEURS-P5-750ML as substitutes, RTD-BEER-P24-340ML and WINE-BIB-P44-3LTR as complementary products and the advertising types: gifting, in-store instant gratification, T.V., Magazine and digital as effective advertising spend.

The regression model gives marketers insight into the significance or effectiveness of a relationship but also the strength and contextual direction of the relationship between the dependent and independent variables. The model does not give marketers contextual information on the full marketing landscape and marketers will need to incorporate the effective independent variables, representing effective marketing campaigns into their strategies. If a type of advertising was not used during the period under investigation no insights are provided on this advertising type.

The DEA method was applied to find an efficiency frontier, the optimal combination of input variables to maximise the output or sales revenue. This method did not require long periods of historical time series data, any period of data could be used as long as it was consistent across the input and output variables, also any unit of measure could be used and be frequently updated. The results were consistent with the findings of other researchers such as, Luo and Donthu [41] and Cheong, De Gregorio and Kim [14], showing that the inefficient products in 2016 become efficient, by reducing the overall advertising spend in 2017.

The results of the DEA method and regression analysis were directionally similar, both finding that outdoor is an ineffective and efficient type of advertising spend for SPIRITS-LIQUEURS-P16-750ML. The DEA method provides the optimal marketing mix as an output, giving actionable insights for marketers on how to efficiently spend their advertising budgets.

DEA has been found to be the method that met all the objectives set out for this research. The DEA method can be used to analyse different marketing campaigns for efficiency because of the ability to vary the period of the dataset, different marketing strategies can also be modelled. As the DEA model is simple to build and compute it can be applied frequently to varying time periods to monitor changes in trends or shifts in consumers' buying behaviour. Being aware of the shifts in consumer buying behaviour and which adversing types are causing the shift will add valuable insight to marketing strategies.

A limitation of the DEA method is that it will always produce a result, even if the incorrect competitor set is modelled or advertising types are missing. The DEA method does not provide output statistics to validate how well the model represents that landscape and fits the data set.

The price elasticity methodology was found to also produce a informative result, while not as simple and insightful as DEA. One limitation of the DEA model built is that products price fluctuations or price discounting strategies by the marketer are not directly represented in the model. As future research the own price elasticity of each product in the competitor set could be used as an additional output variable in the DEA model and as an input the amount spent on price discounting. This inclusion will give marketers the ability to also gain insights about how their advertising expenditure is affected by the product and its competitor sets pricing strategy.

8.3   Future work

For future research a feature selection model should be investigated for reducing the number of variables as inputs to the models. In a study by Girish in 2014 [11], an overview is provided into some of the common methods present in literature. Feature selection is the process of selecting a subset of variables from the input data set which can efficiently describe the entire input data set, thereby reducing the effects from noise or irrelevant variables while still maintaining the significant variables [11]. Feature selection methods provide a way to reducing computation time, improving forecasting or predictive performance and a better understanding of the data [11]. Filter, Wrapper and the Embedded methods should be tested as a variable reduction process before regression analysis and could also be tested against factor and cluster analysis in the PLC methodology.

APPENDIX A

Appendix

A.1 Code for SAS studio software procedure

A.1.1 Exploratory factor analysis procedure

The code for SAS studio® software is displayed in Figure A.1, from Equation 2.33, using the unweighted least squares method with the MAX option and Quartimax rotation to extract 8 factors.

```
 proc factor data=data.base method=uls mineigen=0 priors=max reorder rotate=quartimax corr msa
scree residuals heywood n=8 out=data.factorresults;
var Volume Value AUP
    Q1 Q2 Q3 Q4 Q5 Q6 Q7 Q8 Q9
    Q11 Q12 Q13 Q14 Q16 Q17 Q18 Q19
    Q22 Q23 Q24 Q25 Q26 Q28 Q29 Q30
    Q31 Q33 Q34 Q36 Q37 Q38 Q39
    Q41 Q45 Q46 Q47 Q48 Q49 Q50
    Q51 Q52 Q53 Q54 Q55 Q56 Q57 Q58 Q59 Q60
    Q61 Q62 Q63 Q64 Q65 Q66 Q67 Q68 Q69 Q70
    Q71 Q72 Q73 Q74 Q75 Q76 Q77 Q78;
run;
```

FIGURE A.1: *Code in SAS studio® software for factor analysis using the PROC FACTOR procedure*

A.1.2   Cluster analysis procedure

The code for SAS studio® software to compute the data transformation is displayed in Figure A.2. The PROC ACECLUS method will be used in SAS studio® software to standardise the post-factor-analysis dataset and the categorical variables, alcohol percentage, carbonation level, sugar content, pack size and taste profile.

```
proc aceclus data=data.factorresults out=data.data.basestd proportion=.03;
var SmallPackWine MediumPackWine LargePackWine
    SmallPackSpirits MediumPackSpirits LargePackSpirits
    SmallPackRTDs MediumPackRTDs LargePackRTDs
    Still Carbonated Sparkling Lowalcohol MidAlcohol HighAlcohol
    Dry SemiSweet Sweet Bitter WineLike BeerLike CiderLike
    Fruity WoodyYeasty HerbaceousGreen SweetAssociated SmokyPeaty Smooth
    MouthCoatingCreamy
    Factor1 Factor2 Factor3 Factor4 Factor5 Factor6 Factor7 Factor8;
run;
```

FIGURE A.2: *Code in SAS studio® software for standardising the dataset using PROC ACECLUS procedure*

The code for SAS studio® software to compute Ward's minimum distance method is displayed in Figure A.3. The variables *Can1* to *Can37* are the standardised variables describing pack sizes, alcohol percentage, carbonation level, sugar content, taste profile and the factor scores.

```
proc cluster data=data.basestd outtree=data.Tree method=ward ccc pseudo;
var Can1 Can2 Can3 Can4 Can5 Can6 Can7 Can8 Can9 Can10
    Can11 Can12 Can13 Can14 Can15 Can16 Can17 Can18 Can19 Can20
    Can21 Can22 Can23 Can24 Can25 Can26 Can27 Can28 Can29 Can30
    Can31 Can32 Can33 Can34 Can35 Can36 Can37;
id Product Pack;
run;
```

FIGURE A.3: *Code in SAS studio® software for clustering using Ward's minimum distance method*

The code for SAS studio® software to compute the k-means method of clustering using the PROC FASTCLUS is displayed in Figure A.4. The variables *Can1* to *Can37* are the standardised variables describing pack sizes, alcohol percentage, carbonation level, sugar content, taste profile and the factor scores.

```
proc fastclus data=data.basestd out=data.ClusResult maxclusters=10 maxiter=100;
var Can1 Can2 Can3 Can4 Can5 Can6 Can7 Can8 Can9 Can10
    Can11 Can12 Can13 Can14 Can15 Can16 Can17 Can18 Can19 Can20
    Can21 Can22 Can23 Can24 Can25 Can26 Can27 Can28 Can29 Can30
    Can31 Can32 Can33 Can34 Can35 Can36 Can37;
run;
```

FIGURE A.4: *Code in SAS studio® software for clustering using the k-means method*

A.1.3   Stepwise regression analysis

The code for SAS studio® software to compute the first iteration for RTD-CIDER-P18-660ML is displayed in Figure A.5.

```
proc reg data=datafile.RTD-CIDER-P18-660ML;
model Q1=P1 . . . P594 ;
       / selection=stepwise details=summary;
run;
```

FIGURE A.5: *Code in SAS studio® software for stepwise regression analysis*

A.1.4   Correlation analysis

The code for SAS studio® software to compute the correlation coefficients for a product on the market relative to the products being investigated are displayed in Figure A.6.

```
proc corr data=datafile.RTD-CIDER-P18-660ML reorder;
var Q1
with P1 P2 P3 . . . P594
run;
```

FIGURE A.6: *Code in SAS studio® software for correlation analysis*

A.1.5   Price elasticity multiple regression analysis

The code for SAS studio® software to compute multiple regression analysis of price elasticity is displayed for RTD-CIDER-P18-660ML and the first independent variable in the set in Figure A.7.

```
proc reg data=datafile.RTD-CIDER-P18-660M;
model Q18=P18 P79 D1 D2 D3 D4 D5 D6 D7 D8 D9 D10 D11
         / DW;
run;
```

FIGURE A.7: *Code in SAS studio® software for multiple regression analysis of price elasticity*

A.1.6   Multiple regression analysis of SPIRITS-LIQUEURS-P16-750ML

The code for SAS studio® software to compute multiple regression analysis with binary dummy variables is displayed in Figure A.8.

```
proc reg data=datafile.SPIRITS-LIQUEURS-P16-750ML;
model Q16= P1 . . . P594
                Gifting BTLINTG BTLACT TV Outdoor Magazine Digital
                Autumn Winter Spring
run;
```

FIGURE A.8: *Code in SAS studio® software for multiple regression analysis of SPIRITS-LIQUEURS-P16-750ML*

A.2   Data: Data relationships

The correlation matrices showing the resulting correlation coefficients and p-values.

| | Pearson's Correlation Coefficients N = 594 Prob $> \lvert r \rvert$ under H0: Rho=0 | | | | | | | | | | |
	Volume	Value	AUP	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8
Volume r	1.00	0.81	-0.06	-0.11	-0.08	-0.01	-0.04	-0.07	-0.06	-0.07	-0.05
Volume p-value		0.00	0.12	0.01	0.06	0.84	0.37	0.08	0.14	0.09	0.26
Value r	0.81	1.00	0.14	-0.14	-0.11	-0.04	-0.06	-0.10	-0.08	-0.08	-0.08
Value p-value	0.00		0.00	0.00	0.01	0.28	0.12	0.01	0.04	0.05	0.04
AUP r	-0.06	0.14	1.00	-0.01	-0.12	-0.07	-0.08	-0.13	0.05	0.01	-0.02
AUP p-value	0.12	0.00		0.84	0.00	0.11	0.06	0.00	0.23	0.80	0.60
Q1 r	-0.11	-0.14	-0.01	1.00	0.36	0.31	0.28	0.51	0.73	0.38	0.32
Q1 p-value	0.01	0.00	0.84		0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q2 r	-0.08	-0.11	-0.12	0.36	1.00	0.49	0.37	-0.07	0.03	0.56	0.44
Q2 p-value	0.06	0.01	0.00	0.00		0.00	0.00	0.11	0.48	0.00	0.00
Q3 r	-0.01	-0.04	-0.07	0.31	0.49	1.00	0.35	-0.09	0.03	0.53	0.36
Q3 p-value	0.84	0.28	0.11	0.00	0.00		0.00	0.03	0.54	0.00	0.00
Q4 r	-0.04	-0.06	-0.08	0.28	0.37	0.35	1.00	-0.06	-0.09	0.41	0.31
Q4 p-value	0.37	0.12	0.06	0.00	0.00	0.00		0.11	0.03	0.00	0.00
Q5 r	-0.07	-0.10	-0.13	0.51	-0.07	-0.09	-0.06	1.00	0.23	-0.10	-0.01
Q5 p-value	0.08	0.01	0.00	0.00	0.11	0.03	0.11		0.00	0.02	0.77
Q6 r	-0.06	-0.08	0.05	0.73	0.03	0.03	-0.09	0.23	1.00	0.05	0.04
Q6 p-value	0.14	0.04	0.23	0.00	0.48	0.54	0.03	0.00		0.19	0.37
Q7 r	-0.07	-0.08	0.01	0.38	0.56	0.53	0.41	-0.10	0.05	1.00	0.43
Q7 p-value	0.09	0.05	0.80	0.00	0.00	0.00	0.00	0.02	0.19		0.00
Q8 r	-0.05	-0.08	-0.02	0.32	0.44	0.36	0.31	-0.01	0.04	0.43	1.00
Q8 p-value	0.26	0.04	0.60	0.00	0.00	0.00	0.00	0.77	0.37	0.00	
Q9 r	-0.04	-0.01	0.05	0.29	0.07	0.05	0.06	-0.10	0.34	0.07	0.20
Q9 p-value	0.38	0.79	0.23	0.00	0.10	0.27	0.15	0.02	0.00	0.11	0.00
Q10 r	-0.06	-0.02	0.08	0.38	0.12	0.09	0.16	-0.07	0.39	0.11	0.31
Q10 p-value	0.12	0.61	0.04	0.00	0.00	0.02	0.00	0.07	0.00	0.01	0.00
Q11 r	-0.03	-0.08	-0.10	0.12	0.19	0.07	0.15	-0.01	-0.04	0.19	0.47
Q11 p-value	0.54	0.07	0.02	0.00	0.00	0.09	0.00	0.75	0.37	0.00	0.00
Q12 r	-0.01	0.00	0.03	0.15	-0.01	-0.00	-0.01	-0.07	0.23	0.01	0.02
Q12 p-value	0.77	0.99	0.47	0.00	0.84	0.91	0.87	0.07	0.00	0.85	0.64
Q13 r	-0.09	-0.10	-0.04	0.70	0.28	0.31	0.17	0.54	0.40	0.33	0.33
Q13 p-value	0.02	0.02	0.39	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q14 r	-0.13	-0.15	-0.11	0.14	0.23	0.25	0.36	-0.04	-0.04	0.24	0.01
Q14 p-value	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.36	0.34	0.00	0.83
Q15 r	-0.05	-0.02	0.07	0.12	0.08	0.28	0.12	-0.07	0.02	0.08	0.27
Q15 p-value	0.25	0.56	0.10	0.00	0.06	0.00	0.00	0.07	0.61	0.06	0.00
Q16 r	-0.01	-0.04	-0.03	0.16	0.19	0.38	0.21	-0.03	-0.01	0.20	0.36
Q16 p-value	0.72	0.31	0.43	0.00	0.00	0.00	0.00	0.49	0.81	0.00	0.00
Q17 r	-0.07	-0.08	-0.07	0.42	-0.01	-0.01	-0.09	0.26	0.64	0.01	-0.06

TABLE A.1: *Part 1: Pearson's correlation coefficients and p-values for numerical variables*

	Pearson's Correlation Coefficients N = 594 Prob > | r | under H0: Rho=0										
	Volume	Value	AUP	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8
Q17 p-value	0.08	0.05	0.10	0.00	0.84	0.77	0.03	0.00	0.00	0.72	0.14
Q18 r	-0.06	-0.03	0.02	0.16	0.35	0.47	0.25	-0.12	0.01	0.37	0.27
Q18 p-value	0.12	0.42	0.59	0.00	0.00	0.00	0.00	0.00	0.89	0.00	0.00
Q19 r	-0.04	-0.06	-0.04	0.39	0.06	-0.04	0.02	-0.03	0.51	0.02	0.14
Q19 p-value	0.36	0.16	0.32	0.00	0.13	0.32	0.57	0.46	0.00	0.66	0.00
Q20 r	-0.02	-0.02	-0.03	0.07	0.16	0.14	0.26	-0.06	-0.07	0.39	0.11
Q20 p-value	0.58	0.69	0.40	0.09	0.00	0.00	0.00	0.12	0.07	0.00	0.01
Q21 r	-0.06	-0.02	0.14	0.19	0.27	0.30	0.35	-0.14	-0.05	0.47	0.29
Q21 p-value	0.15	0.63	0.00	0.00	0.00	0.00	0.00	0.00	0.23	0.00	0.00
Q22 r	-0.04	-0.04	-0.03	0.48	-0.06	-0.02	-0.12	0.91	0.23	-0.03	-0.06
Q22 p-value	0.31	0.38	0.51	0.00	0.16	0.65	0.00	0.00	0.00	0.49	0.14
Q23 r	-0.07	-0.13	-0.09	0.26	0.36	0.27	0.31	-0.00	0.02	0.47	0.58
Q23 p-value	0.07	0.00	0.02	0.00	0.00	0.00	0.00	0.95	0.62	0.00	0.00
Q24 r	-0.09	-0.10	-0.05	0.24	0.13	0.09	0.10	0.32	-0.06	0.11	0.21
Q24 p-value	0.02	0.01	0.27	0.00	0.00	0.03	0.01	0.00	0.15	0.01	0.00
Q25 r	-0.06	-0.07	-0.04	0.39	0.02	-0.00	0.01	0.69	-0.07	0.02	0.07
Q25 p-value	0.15	0.10	0.36	0.00	0.58	0.91	0.89	0.00	0.09	0.56	0.08
Q26 r	-0.06	-0.05	-0.00	-0.05	0.08	-0.01	0.01	-0.07	-0.03	0.05	0.09
Q26 p-value	0.13	0.18	0.96	0.22	0.04	0.80	0.75	0.08	0.46	0.20	0.03
Q27 r	-0.02	-0.05	0.01	-0.05	0.05	-0.09	0.05	-0.07	-0.04	-0.03	0.27
Q27 p-value	0.71	0.19	0.78	0.23	0.19	0.02	0.27	0.08	0.34	0.40	0.00
Q28 r	-0.04	-0.08	-0.06	0.14	0.17	0.14	0.26	-0.03	-0.01	0.13	0.40
Q28 p-value	0.28	0.05	0.15	0.00	0.00	0.00	0.00	0.48	0.90	0.00	0.00
Q29 r	-0.06	-0.06	-0.07	0.03	0.10	0.18	0.16	-0.08	-0.03	0.11	0.02
Q29 p-value	0.12	0.14	0.11	0.47	0.02	0.00	0.00	0.04	0.52	0.01	0.70
Q30 r	-0.04	-0.03	0.10	0.27	0.08	0.04	0.17	-0.09	0.29	0.15	0.13
Q30 p-value	0.37	0.47	0.02	0.00	0.04	0.28	0.00	0.03	0.00	0.00	0.00
Q31 r	-0.06	-0.08	-0.04	0.16	0.36	0.34	0.41	-0.09	-0.06	0.40	0.41
Q31 p-value	0.15	0.04	0.39	0.00	0.00	0.00	0.00	0.03	0.18	0.00	0.00
Q32 r	-0.02	-0.02	0.13	0.49	0.04	-0.02	0.05	-0.03	0.60	0.05	0.12
Q32 p-value	0.70	0.67	0.00	0.00	0.31	0.67	0.25	0.46	0.00	0.18	0.00
Q33 r	-0.03	-0.01	0.04	-0.02	0.02	-0.00	0.11	-0.07	-0.05	0.09	0.08
Q33 p-value	0.53	0.76	0.33	0.55	0.63	0.99	0.01	0.11	0.22	0.03	0.06
Q34 r	-0.11	0.00	0.35	0.35	0.13	0.01	0.07	0.39	-0.04	0.18	0.29
Q34 p-value	0.01	0.97	0.00	0.00	0.00	0.88	0.08	0.00	0.30	0.00	0.00
Q35 r	-0.13	-0.01	0.37	0.33	0.05	-0.04	0.00	0.41	-0.03	0.09	0.20
Q35 p-value	0.00	0.83	0.00	0.00	0.21	0.33	0.95	0.00	0.50	0.02	0.00
Q36 r	-0.00	-0.03	-0.08	0.18	0.38	0.22	0.29	-0.06	-0.05	0.42	0.53
Q36 p-value	0.98	0.47	0.04	0.00	0.00	0.00	0.00	0.15	0.22	0.00	0.00
Q37 r	-0.12	-0.11	-0.05	0.73	0.19	0.08	0.21	0.65	0.45	0.22	0.26
Q37 p-value	0.00	0.01	0.24	0.00	0.00	0.07	0.00	0.00	0.00	0.00	0.00
Q38 r	-0.09	-0.07	-0.02	0.44	0.09	0.08	0.14	0.68	-0.06	0.08	0.13
Q38 p-value	0.03	0.11	0.65	0.00	0.03	0.05	0.00	0.00	0.15	0.05	0.00
Q39 r	-0.04	-0.06	-0.05	0.44	0.09	0.03	0.14	-0.02	0.49	0.07	0.19
Q39 p-value	0.28	0.17	0.20	0.00	0.03	0.49	0.00	0.57	0.00	0.10	0.00
Q40 r	-0.06	-0.07	-0.03	0.33	0.13	0.14	-0.03	0.49	0.43	0.15	-0.01
Q40 p-value	0.15	0.09	0.43	0.00	0.00	0.00	0.40	0.00	0.00	0.00	0.85

TABLE A.2: *Part 2: Pearson's correlation coefficients and p-values for numerical variables*

	Volume	Value	AUP	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8
Pearson's Correlation Coefficients N = 594 Prob > | r | under H0: Rho=0											
Q41 r	-0.10	-0.08	0.05	0.17	0.30	-0.04	0.33	-0.05	-0.02	0.32	0.47
Q41 p-value	0.02	0.05	0.19	0.00	0.00	0.39	0.00	0.22	0.63	0.00	0.00
Q42 r	0.02	0.02	-0.00	0.00	-0.02	0.05	-0.01	-0.03	-0.01	-0.04	-0.00
Q42 p-value	0.67	0.69	0.94	0.99	0.59	0.27	0.90	0.46	0.90	0.36	0.94
Q43 r	0.08	0.02	-0.02	-0.12	0.05	0.14	0.21	-0.07	-0.19	0.12	0.04
Q43 p-value	0.06	0.65	0.69	0.00	0.27	0.00	0.00	0.07	0.00	0.00	0.38
Q44 r	0.03	-0.00	0.03	0.09	0.41	0.26	0.07	-0.17	-0.06	0.39	0.44
Q44 p-value	0.51	0.99	0.49	0.03	0.00	0.00	0.08	0.00	0.15	0.00	0.00
Q45 r	0.12	0.03	0.04	0.01	0.11	0.16	0.13	-0.05	-0.04	0.19	0.14
Q45 p-value	0.00	0.42	0.32	0.78	0.01	0.00	0.00	0.19	0.29	0.00	0.00
Q46 r	0.05	-0.02	0.07	0.04	0.13	0.14	0.18	-0.05	-0.04	0.23	0.18
Q46 p-value	0.22	0.67	0.07	0.37	0.00	0.00	0.00	0.23	0.34	0.00	0.00
Q47 r	0.13	0.05	0.04	0.01	0.08	0.16	0.11	-0.05	-0.04	0.17	0.14
Q47 p-value	0.00	0.26	0.37	0.82	0.04	0.00	0.01	0.19	0.30	0.00	0.00
Q48 r	0.10	0.00	-0.01	0.03	0.15	0.17	0.17	-0.04	-0.04	0.22	0.16
Q48 p-value	0.02	0.97	0.76	0.41	0.00	0.00	0.00	0.32	0.38	0.00	0.00
Q49 r	0.11	0.03	0.05	0.02	0.09	0.16	0.13	-0.05	-0.04	0.19	0.14
Q49 p-value	0.01	0.46	0.26	0.67	0.02	0.00	0.00	0.19	0.31	0.00	0.00
Q50 r	0.13	0.04	0.04	0.00	0.10	0.15	0.11	-0.05	-0.04	0.18	0.13
Q50 p-value	0.00	0.34	0.33	0.92	0.02	0.00	0.01	0.18	0.35	0.00	0.00
Q51 r	0.15	0.06	0.04	-0.00	0.08	0.14	0.09	-0.06	-0.04	0.16	0.12
Q51 p-value	0.00	0.14	0.31	0.92	0.06	0.00	0.04	0.16	0.33	0.00	0.00
Q52 r	0.10	0.03	0.08	-0.00	0.07	0.14	0.09	-0.06	-0.04	0.17	0.13
Q52 p-value	0.01	0.45	0.05	0.96	0.07	0.00	0.02	0.13	0.29	0.00	0.00
Q53 r	0.13	0.05	0.07	-0.00	0.09	0.14	0.11	-0.06	-0.04	0.18	0.12
Q53 p-value	0.00	0.23	0.11	0.99	0.04	0.00	0.01	0.13	0.29	0.00	0.00
Q54 r	0.08	-0.01	0.04	0.03	0.14	0.15	0.16	-0.05	-0.04	0.23	0.19
Q54 p-value	0.06	0.87	0.37	0.42	0.00	0.00	0.00	0.24	0.37	0.00	0.00
Q55 r	0.09	0.03	0.08	0.01	0.08	0.16	0.12	-0.06	-0.05	0.19	0.13
Q55 p-value	0.02	0.52	0.05	0.83	0.04	0.00	0.00	0.13	0.26	0.00	0.00
Q56 r	0.14	0.06	0.07	-0.01	0.07	0.13	0.08	-0.06	-0.04	0.16	0.12
Q56 p-value	0.00	0.12	0.09	0.81	0.08	0.00	0.06	0.13	0.30	0.00	0.00
Q57 r	0.14	0.07	0.10	-0.02	0.03	0.14	0.05	-0.07	-0.05	0.12	0.08
Q57 p-value	0.00	0.08	0.02	0.60	0.49	0.00	0.25	0.10	0.26	0.00	0.04
Q58 r	0.13	0.06	0.08	-0.00	0.07	0.14	0.10	-0.06	-0.04	0.17	0.12
Q58 p-value	0.00	0.15	0.04	0.97	0.11	0.00	0.02	0.14	0.27	0.00	0.00
Q59 r	0.15	0.07	0.07	-0.01	0.06	0.15	0.07	-0.06	-0.04	0.15	0.11
Q59 p-value	0.00	0.07	0.10	0.82	0.17	0.00	0.07	0.14	0.30	0.00	0.01
Q60 r	0.13	0.03	-0.00	0.02	0.12	0.17	0.14	-0.05	-0.04	0.19	0.13
Q60 p-value	0.00	0.40	0.99	0.71	0.00	0.00	0.00	0.22	0.31	0.00	0.00
Q61 r	0.01	-0.04	0.10	0.04	0.14	0.14	0.18	-0.05	-0.03	0.23	0.19
Q61 p-value	0.81	0.34	0.02	0.34	0.00	0.00	0.00	0.22	0.43	0.00	0.00
Q62 r	0.18	0.08	-0.00	-0.00	0.08	0.16	0.09	-0.05	-0.04	0.15	0.10
Q62 p-value	0.00	0.06	0.95	0.92	0.06	0.00	0.03	0.21	0.31	0.00	0.01
Q63 r	0.16	0.05	-0.03	0.01	0.10	0.17	0.11	-0.04	-0.04	0.18	0.12
Q63 p-value	0.00	0.22	0.47	0.83	0.01	0.00	0.01	0.28	0.37	0.00	0.00
Q64 r	0.11	0.03	0.04	0.02	0.10	0.16	0.14	-0.05	-0.04	0.19	0.14
Q64 p-value	0.01	0.50	0.31	0.66	0.01	0.00	0.00	0.21	0.32	0.00	0.00

TABLE A.3: *Part 3: Pearson's correlation coefficients and p-values for numerical variables*

	Volume	Value	AUP	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8
Pearson's Correlation Coefficients N = 594 Prob > | r | under H0: Rho=0											
Q65 r	0.03	0.00	0.20	0.00	0.06	0.10	0.14	-0.08	-0.05	0.19	0.13
Q65 p-value	0.48	0.91	0.00	0.93	0.15	0.02	0.00	0.06	0.19	0.00	0.00
Q66 r	0.10	0.04	0.12	-0.00	0.07	0.13	0.10	-0.07	-0.04	0.18	0.13
Q66 p-value	0.02	0.34	0.00	1.00	0.11	0.00	0.01	0.10	0.27	0.00	0.00
Q67 r	0.13	0.05	0.09	-0.01	0.06	0.14	0.07	-0.06	-0.04	0.15	0.12
Q67 p-value	0.00	0.19	0.04	0.77	0.17	0.00	0.11	0.13	0.33	0.00	0.00
Q68 r	0.12	0.03	0.02	0.01	0.11	0.15	0.12	-0.05	-0.04	0.19	0.14
Q68 p-value	0.00	0.45	0.71	0.78	0.01	0.00	0.00	0.22	0.34	0.00	0.00
Q69 r	0.17	0.10	0.10	-0.02	0.02	0.13	0.06	-0.07	-0.04	0.12	0.10
Q69 p-value	0.00	0.02	0.02	0.63	0.58	0.00	0.18	0.11	0.29	0.01	0.02
Q70 r	0.11	0.03	0.05	0.02	0.10	0.16	0.13	-0.05	-0.04	0.19	0.14
Q70 p-value	0.01	0.52	0.27	0.66	0.02	0.00	0.00	0.21	0.35	0.00	0.00
Q71 r	0.11	0.02	0.02	0.02	0.12	0.16	0.16	-0.05	-0.04	0.21	0.13
Q71 p-value	0.01	0.64	0.71	0.61	0.00	0.00	0.00	0.22	0.29	0.00	0.00
Q72 r	0.08	0.00	0.04	0.02	0.13	0.15	0.14	-0.05	-0.04	0.22	0.17
Q72 p-value	0.04	0.91	0.30	0.58	0.00	0.00	0.00	0.23	0.34	0.00	0.00
Q73 r	0.10	0.01	0.02	0.02	0.12	0.15	0.15	-0.05	-0.04	0.21	0.15
Q73 p-value	0.02	0.79	0.56	0.61	0.00	0.00	0.00	0.23	0.35	0.00	0.00
Q74 r	0.15	0.07	0.07	-0.01	0.05	0.14	0.08	-0.06	-0.04	0.15	0.11
Q74 p-value	0.00	0.08	0.09	0.79	0.20	0.00	0.07	0.13	0.28	0.00	0.01
Q75 r	0.13	0.07	0.13	-0.01	0.05	0.12	0.08	-0.07	-0.05	0.15	0.13
Q75 p-value	0.00	0.07	0.00	0.74	0.20	0.00	0.06	0.07	0.26	0.00	0.00
Q76 r	0.14	0.07	0.09	-0.01	0.05	0.13	0.09	-0.07	-0.05	0.14	0.09
Q76 p-value	0.00	0.10	0.03	0.72	0.24	0.00	0.03	0.10	0.23	0.00	0.02
Q77 r	0.57	0.50	-0.04	-0.16	-0.12	-0.03	-0.08	-0.11	-0.09	-0.12	-0.07
Q77 p-value	0.00	0.00	0.30	0.00	0.00	0.51	0.06	0.01	0.03	0.00	0.08
Q78 r	0.59	0.50	-0.06	-0.15	-0.11	-0.02	-0.08	-0.09	-0.08	-0.10	-0.07
Q78 p-value	0.00	0.00	0.18	0.00	0.01	0.65	0.04	0.03	0.04	0.01	0.07

TABLE A.4: *Part 4: Pearson's correlation coefficients and p-values for numerical variables*

	Q9	Q10	Q11	Q12	Q13	Q14	Q15	Q16	Q17	Q18	Q19	Q20
	Pearson's Correlation Coefficients N = 594 Prob > |r| under H0: Rho=0											
Volume r	-0.04	-0.06	-0.03	-0.01	-0.09	-0.13	-0.05	-0.01	-0.07	-0.06	-0.04	-0.02
Volume p-value	0.38	0.12	0.54	0.77	0.02	0.00	0.25	0.72	0.08	0.12	0.36	0.58
Value r	-0.01	-0.02	-0.08	0.00	-0.10	-0.15	-0.02	-0.04	-0.08	-0.03	-0.06	-0.02
Value p-value	0.79	0.61	0.07	0.99	0.02	0.00	0.56	0.31	0.05	0.42	0.16	0.69
AUP r	0.05	0.08	-0.10	0.03	-0.04	-0.11	0.07	-0.03	-0.07	0.02	-0.04	-0.03
AUP p-value	0.23	0.04	0.02	0.47	0.39	0.01	0.10	0.43	0.10	0.59	0.32	0.40
Q1 r	0.29	0.38	0.12	0.15	0.70	0.14	0.12	0.16	0.42	0.16	0.39	0.07
Q1 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.09
Q2 r	0.07	0.12	0.19	-0.01	0.28	0.23	0.08	0.19	-0.01	0.35	0.06	0.16
Q2 p-value	0.10	0.00	0.00	0.84	0.00	0.00	0.06	0.00	0.84	0.00	0.13	0.00
Q3 r	0.05	0.09	0.07	-0.00	0.31	0.25	0.28	0.38	-0.01	0.47	-0.04	0.14
Q3 p-value	0.27	0.02	0.09	0.91	0.00	0.00	0.00	0.00	0.77	0.00	0.32	0.00
Q4 r	0.06	0.16	0.15	-0.01	0.17	0.36	0.12	0.21	-0.09	0.25	0.02	0.26
Q4 p-value	0.15	0.00	0.00	0.87	0.00	0.00	0.00	0.00	0.03	0.00	0.57	0.00
Q5 r	-0.10	-0.07	-0.01	-0.07	0.54	-0.04	-0.07	-0.03	0.26	-0.12	-0.03	-0.06
Q5 p-value	0.02	0.07	0.75	0.07	0.00	0.36	0.07	0.49	0.00	0.00	0.46	0.12
Q6 r	0.34	0.39	-0.04	0.23	0.40	-0.04	0.02	-0.01	0.64	0.01	0.51	-0.07
Q6 p-value	0.00	0.00	0.37	0.00	0.00	0.34	0.61	0.81	0.00	0.89	0.00	0.07
Q7 r	0.07	0.11	0.19	0.01	0.33	0.24	0.08	0.20	0.01	0.37	0.02	0.39
Q7 p-value	0.11	0.01	0.00	0.85	0.00	0.00	0.06	0.00	0.72	0.00	0.66	0.00
Q8 r	0.20	0.31	0.47	0.02	0.33	0.01	0.27	0.36	-0.06	0.27	0.14	0.11
Q8 p-value	0.00	0.00	0.00	0.64	0.00	0.83	0.00	0.00	0.14	0.00	0.00	0.01
Q9 r	1.00	0.74	0.27	0.87	0.39	-0.03	0.13	0.18	0.44	0.13	0.72	0.08
Q9 p-value		0.00	0.00	0.00	0.00	0.49	0.00	0.00	0.00	0.00	0.00	0.05
Q10 r	0.74	1.00	0.26	0.36	0.54	0.06	0.32	0.30	0.49	0.23	0.81	0.13
Q10 p-value	0.00		0.00	0.00	0.00	0.18	0.00	0.00	0.00	0.00	0.00	0.00
Q11 r	0.27	0.26	1.00	-0.02	0.20	0.03	0.14	0.43	-0.13	0.24	0.13	0.32
Q11 p-value	0.00	0.00		0.66	0.00	0.49	0.00	0.00	0.00	0.00	0.00	0.00
Q12 r	0.87	0.36	-0.02	1.00	0.18	-0.02	-0.03	-0.00	0.33	-0.03	0.48	-0.02
Q12 p-value	0.00	0.00	0.66		0.00	0.55	0.44	0.90	0.00	0.52	0.00	0.65
Q13 r	0.39	0.54	0.20	0.18	1.00	0.21	0.23	0.33	0.57	0.30	0.49	0.27
Q13 p-value	0.00	0.00	0.00	0.00		0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q14 r	-0.03	0.06	0.03	-0.02	0.21	1.00	0.24	0.26	0.02	0.11	0.06	0.42
Q14 p-value	0.49	0.18	0.49	0.55	0.00		0.00	0.00	0.65	0.01	0.17	0.00
Q15 r	0.13	0.32	0.14	-0.03	0.23	0.24	1.00	0.34	0.04	0.44	0.05	0.04
Q15 p-value	0.00	0.00	0.00	0.44	0.00	0.00		0.00	0.38	0.00	0.21	0.37
Q16 r	0.18	0.30	0.43	-0.00	0.33	0.26	0.34	1.00	-0.04	0.22	0.08	0.22
Q16 p-value	0.00	0.00	0.00	0.90	0.00	0.00	0.00		0.28	0.00	0.04	0.00
Q17 r	0.44	0.49	-0.13	0.33	0.57	0.02	0.04	-0.04	1.00	-0.02	0.63	-0.04
Q17 p-value	0.00	0.00	0.00	0.00	0.00	0.65	0.38	0.28		0.58	0.00	0.37
Q18 r	0.13	0.23	0.24	-0.03	0.30	0.11	0.44	0.22	-0.02	1.00	-0.00	0.36
Q18 p-value	0.00	0.00	0.00	0.52	0.00	0.01	0.00	0.00	0.58		0.92	0.00
Q19 r	0.72	0.81	0.13	0.48	0.49	0.06	0.05	0.08	0.63	-0.00	1.00	0.04
Q19 p-value	0.00	0.00	0.00	0.00	0.00	0.17	0.21	0.04	0.00	0.92		0.36
Q20 r	0.08	0.13	0.32	-0.02	0.27	0.42	0.04	0.22	-0.04	0.36	0.04	1.00
Q20 p-value	0.05	0.00	0.00	0.65	0.00	0.00	0.37	0.00	0.37	0.00	0.36	

TABLE A.5: *Part 5: Pearson's correlation coefficients and p-values for numerical variables*

	Q9	Q10	Q11	Q12	Q13	Q14	Q15	Q16	Q17	Q18	Q19	Q20
	Pearson's Correlation Coefficients N = 594 Prob > | *r* | under H0: Rho=0											
Q21 r	0.08	0.18	0.20	-0.02	0.36	0.26	0.22	0.35	-0.05	0.23	-0.02	0.34
Q21 p-value	0.06	0.00	0.00	0.71	0.00	0.00	0.00	0.00	0.26	0.00	0.66	0.00
Q22 r	-0.12	-0.13	-0.12	-0.05	0.54	-0.08	-0.03	-0.02	0.31	-0.06	-0.09	-0.13
Q22 p-value	0.00	0.00	0.00	0.18	0.00	0.05	0.43	0.60	0.00	0.17	0.04	0.00
Q23 r	0.12	0.24	0.36	0.00	0.27	0.20	0.24	0.14	-0.04	0.18	0.09	0.21
Q23 p-value	0.00	0.00	0.00	0.98	0.00	0.00	0.00	0.00	0.37	0.00	0.03	0.00
Q24 r	0.43	-0.00	0.12	0.59	0.31	0.18	0.15	0.10	-0.04	0.14	-0.00	0.07
Q24 p-value	0.00	0.97	0.00	0.00	0.00	0.00	0.00	0.01	0.37	0.00	0.96	0.09
Q25 r	-0.05	-0.10	0.09	-0.04	0.47	-0.01	-0.07	0.02	-0.08	0.07	-0.04	0.06
Q25 p-value	0.20	0.02	0.02	0.34	0.00	0.72	0.11	0.55	0.06	0.09	0.33	0.15
Q26 r	0.54	-0.04	0.04	0.77	-0.03	-0.00	-0.03	-0.03	0.01	0.08	-0.01	-0.01
Q26 p-value	0.00	0.30	0.28	0.00	0.53	0.95	0.43	0.40	0.76	0.04	0.79	0.87
Q27 r	0.06	0.21	0.17	-0.05	0.08	-0.08	0.25	0.02	0.10	0.09	0.09	-0.16
Q27 p-value	0.14	0.00	0.00	0.23	0.05	0.05	0.00	0.57	0.02	0.03	0.03	0.00
Q28 r	0.14	0.29	0.37	-0.03	0.15	0.31	0.49	0.36	-0.05	0.25	0.10	0.11
Q28 p-value	0.00	0.00	0.00	0.40	0.00	0.00	0.00	0.00	0.18	0.00	0.02	0.01
Q29 r	0.04	0.06	0.19	-0.03	0.22	0.48	0.49	0.30	0.04	0.56	0.01	0.50
Q29 p-value	0.34	0.17	0.00	0.42	0.00	0.00	0.00	0.00	0.38	0.00	0.76	0.00
Q30 r	0.46	0.00	0.12	0.62	-0.01	0.01	0.03	0.02	-0.07	0.18	-0.02	0.12
Q30 p-value	0.00	0.94	0.00	0.00	0.76	0.76	0.42	0.61	0.09	0.00	0.71	0.00
Q31 r	0.04	0.01	0.25	0.00	0.19	0.10	0.08	-0.09	-0.04	0.52	-0.02	0.22
Q31 p-value	0.30	0.86	0.00	0.91	0.00	0.01	0.05	0.03	0.31	0.00	0.63	0.00
Q32 r	0.03	0.08	0.09	-0.02	0.00	-0.00	0.08	0.06	-0.04	0.05	0.03	0.06
Q32 p-value	0.48	0.06	0.03	0.67	0.96	0.98	0.04	0.17	0.32	0.27	0.52	0.17
Q33 r	0.56	-0.02	0.04	0.81	-0.05	-0.03	0.02	-0.01	-0.05	0.08	-0.02	0.03
Q33 p-value	0.00	0.66	0.31	0.00	0.20	0.44	0.64	0.77	0.27	0.04	0.68	0.45
Q34 r	0.09	0.22	0.17	-0.05	0.43	-0.01	0.16	0.14	-0.14	0.14	0.02	0.03
Q34 p-value	0.04	0.00	0.00	0.24	0.00	0.86	0.00	0.00	0.00	0.00	0.58	0.41
Q35 r	0.07	0.18	0.14	-0.04	0.40	-0.04	0.12	0.12	-0.14	0.10	0.02	-0.02
Q35 p-value	0.07	0.00	0.00	0.28	0.00	0.33	0.00	0.00	0.00	0.02	0.67	0.67
Q36 r	0.11	0.22	0.32	-0.00	0.24	0.16	0.23	0.21	-0.04	0.21	0.07	0.25
Q36 p-value	0.01	0.00	0.00	0.92	0.00	0.00	0.00	0.00	0.35	0.00	0.08	0.00
Q37 r	0.40	0.54	0.16	0.22	0.84	0.12	0.05	0.13	0.56	0.07	0.54	0.12
Q37 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.21	0.00	0.00	0.07	0.00	0.00
Q38 r	-0.02	0.06	0.01	-0.05	0.50	-0.01	0.01	0.03	-0.07	0.06	-0.01	0.01
Q38 p-value	0.64	0.11	0.77	0.25	0.00	0.80	0.81	0.48	0.09	0.12	0.85	0.72
Q39 r	0.70	0.77	0.10	0.48	0.49	0.10	0.04	0.03	0.63	0.02	0.91	0.06
Q39 p-value	0.00	0.00	0.01	0.00	0.00	0.02	0.31	0.40	0.00	0.55	0.00	0.17
Q40 r	-0.03	0.03	0.00	-0.05	0.37	0.07	-0.04	0.08	0.50	0.06	-0.00	0.02
Q40 p-value	0.51	0.49	0.98	0.19	0.00	0.10	0.30	0.06	0.00	0.12	0.96	0.55
Q41 r	0.18	0.35	0.46	-0.01	0.31	0.26	0.27	0.37	-0.07	0.18	0.19	0.35
Q41 p-value	0.00	0.00	0.00	0.89	0.00	0.00	0.00	0.00	0.09	0.00	0.00	0.00
Q42 r	-0.00	0.05	-0.07	-0.02	0.00	0.03	-0.05	0.01	-0.02	-0.03	-0.01	0.00
Q42 p-value	0.99	0.22	0.11	0.56	0.99	0.49	0.19	0.83	0.66	0.54	0.72	0.96
Q43 r	-0.19	-0.09	0.11	-0.21	0.05	0.22	0.14	0.18	-0.07	0.12	-0.16	0.25
Q43 p-value	0.00	0.03	0.01	0.00	0.27	0.00	0.00	0.00	0.10	0.00	0.00	0.00
Q44 r	0.09	0.23	0.33	-0.07	0.20	0.06	0.23	0.20	0.00	0.24	0.05	0.15
Q44 p-value	0.03	0.00	0.00	0.07	0.00	0.17	0.00	0.00	0.93	0.00	0.20	0.00

TABLE A.6: *Part 6: Pearson's correlation coefficients and p-values for numerical variables*

Pearson's Correlation Coefficients N = 594 Prob $> \mid r \mid$ under H0: Rho=0											
	Q9	Q10	Q11	Q12	Q13	Q14	Q15	Q16	Q17	Q18	Q19	Q20
Q45 r	0.08	0.12	0.17	0.04	0.11	0.09	0.18	0.27	-0.04	0.11	0.06	0.16
Q45 p-value	0.04	0.00	0.00	0.37	0.01	0.03	0.00	0.00	0.29	0.01	0.17	0.00
Q46 r	0.10	0.16	0.19	0.04	0.15	0.15	0.21	0.30	-0.04	0.13	0.08	0.20
Q46 p-value	0.01	0.00	0.00	0.33	0.00	0.00	0.00	0.00	0.36	0.00	0.06	0.00
Q47 r	0.08	0.12	0.16	0.03	0.10	0.04	0.18	0.24	-0.05	0.11	0.05	0.13
Q47 p-value	0.04	0.00	0.00	0.44	0.02	0.29	0.00	0.00	0.22	0.01	0.22	0.00
Q48 r	0.08	0.12	0.20	0.03	0.13	0.13	0.19	0.30	-0.04	0.12	0.07	0.18
Q48 p-value	0.05	0.00	0.00	0.42	0.00	0.00	0.00	0.00	0.35	0.00	0.09	0.00
Q49 r	0.09	0.13	0.17	0.04	0.11	0.08	0.19	0.27	-0.05	0.13	0.05	0.15
Q49 p-value	0.03	0.00	0.00	0.38	0.01	0.06	0.00	0.00	0.27	0.00	0.20	0.00
Q50 r	0.09	0.12	0.17	0.04	0.11	0.07	0.18	0.28	-0.04	0.11	0.06	0.15
Q50 p-value	0.04	0.00	0.00	0.34	0.01	0.09	0.00	0.00	0.30	0.01	0.15	0.00
Q51 r	0.09	0.11	0.16	0.04	0.08	0.03	0.17	0.24	-0.05	0.11	0.05	0.11
Q51 p-value	0.04	0.01	0.00	0.34	0.04	0.42	0.00	0.00	0.24	0.01	0.24	0.01
Q52 r	0.09	0.13	0.16	0.04	0.10	0.05	0.18	0.26	-0.05	0.12	0.05	0.13
Q52 p-value	0.02	0.00	0.00	0.31	0.02	0.21	0.00	0.00	0.23	0.00	0.23	0.00
Q53 r	0.09	0.12	0.15	0.04	0.09	0.07	0.18	0.26	-0.05	0.12	0.05	0.14
Q53 p-value	0.03	0.00	0.00	0.31	0.02	0.09	0.00	0.00	0.27	0.00	0.24	0.00
Q54 r	0.10	0.14	0.20	0.04	0.14	0.12	0.20	0.31	-0.04	0.12	0.08	0.18
Q54 p-value	0.02	0.00	0.00	0.35	0.00	0.00	0.00	0.00	0.34	0.00	0.07	0.00
Q55 r	0.09	0.13	0.16	0.04	0.11	0.08	0.18	0.26	-0.05	0.13	0.05	0.15
Q55 p-value	0.02	0.00	0.00	0.30	0.01	0.05	0.00	0.00	0.24	0.00	0.21	0.00
Q56 r	0.09	0.12	0.16	0.04	0.08	0.04	0.17	0.24	-0.05	0.11	0.05	0.12
Q56 p-value	0.03	0.00	0.00	0.31	0.04	0.31	0.00	0.00	0.23	0.01	0.23	0.00
Q57 r	0.08	0.10	0.12	0.04	0.05	0.01	0.16	0.19	-0.05	0.11	0.02	0.09
Q57 p-value	0.04	0.01	0.00	0.33	0.19	0.87	0.00	0.00	0.19	0.01	0.54	0.03
Q58 r	0.09	0.13	0.15	0.04	0.09	0.05	0.19	0.24	-0.05	0.12	0.04	0.12
Q58 p-value	0.03	0.00	0.00	0.32	0.03	0.22	0.00	0.00	0.24	0.00	0.30	0.00
Q59 r	0.09	0.12	0.15	0.04	0.08	0.03	0.18	0.24	-0.05	0.11	0.04	0.11
Q59 p-value	0.03	0.00	0.00	0.34	0.05	0.52	0.00	0.00	0.22	0.01	0.32	0.01
Q60 r	0.08	0.11	0.18	0.04	0.11	0.10	0.18	0.28	-0.04	0.12	0.06	0.16
Q60 p-value	0.05	0.01	0.00	0.36	0.01	0.02	0.00	0.00	0.31	0.00	0.15	0.00
Q61 r	0.11	0.17	0.20	0.04	0.15	0.16	0.22	0.32	-0.04	0.15	0.08	0.20
Q61 p-value	0.01	0.00	0.00	0.33	0.00	0.00	0.00	0.00	0.34	0.00	0.05	0.00
Q62 r	0.07	0.08	0.15	0.04	0.07	0.04	0.15	0.23	-0.05	0.09	0.04	0.11
Q62 p-value	0.10	0.05	0.00	0.39	0.07	0.31	0.00	0.00	0.25	0.03	0.33	0.01
Q63 r	0.07	0.08	0.17	0.03	0.09	0.08	0.15	0.26	-0.04	0.09	0.05	0.14
Q63 p-value	0.10	0.04	0.00	0.45	0.02	0.06	0.00	0.00	0.30	0.04	0.20	0.00
Q64 r	0.09	0.12	0.17	0.04	0.11	0.10	0.19	0.27	-0.05	0.12	0.06	0.16
Q64 p-value	0.04	0.00	0.00	0.38	0.01	0.01	0.00	0.00	0.26	0.00	0.18	0.00
Q65 r	0.10	0.17	0.13	0.04	0.11	0.12	0.22	0.23	-0.05	0.17	0.04	0.18
Q65 p-value	0.01	0.00	0.00	0.28	0.01	0.00	0.00	0.00	0.23	0.00	0.34	0.00
Q66 r	0.10	0.15	0.15	0.04	0.10	0.07	0.20	0.24	-0.05	0.14	0.05	0.14
Q66 p-value	0.02	0.00	0.00	0.28	0.02	0.09	0.00	0.00	0.22	0.00	0.23	0.00
Q67 r	0.09	0.12	0.15	0.04	0.08	0.04	0.19	0.24	-0.05	0.12	0.05	0.11
Q67 p-value	0.02	0.00	0.00	0.28	0.05	0.31	0.00	0.00	0.25	0.00	0.24	0.01
Q68 r	0.08	0.11	0.18	0.04	0.11	0.08	0.17	0.28	-0.04	0.10	0.06	0.15
Q68 p-value	0.05	0.01	0.00	0.36	0.01	0.04	0.00	0.00	0.32	0.01	0.16	0.00

TABLE A.7: *Part 7: Pearson's correlation coefficients and p-values for numerical variables*

	Q9	Q10	Q11	Q12	Q13	Q14	Q15	Q16	Q17	Q18	Q19	Q20
	Pearson's Correlation Coefficients N = 594 Prob $>$ $\mid r \mid$ under H0: Rho=0											
Q69 r	0.09	0.11	0.12	0.04	0.05	-0.01	0.17	0.19	-0.05	0.11	0.02	0.08
Q69 p-value	0.03	0.01	0.00	0.33	0.18	0.88	0.00	0.00	0.19	0.01	0.56	0.05
Q70 r	0.09	0.13	0.17	0.04	0.11	0.09	0.19	0.27	-0.04	0.12	0.05	0.16
Q70 p-value	0.04	0.00	0.00	0.39	0.01	0.03	0.00	0.00	0.31	0.00	0.18	0.00
Q71 r	0.08	0.12	0.17	0.04	0.12	0.13	0.17	0.27	-0.04	0.12	0.06	0.19
Q71 p-value	0.06	0.00	0.00	0.37	0.00	0.00	0.00	0.00	0.32	0.00	0.17	0.00
Q72 r	0.09	0.14	0.19	0.04	0.13	0.10	0.20	0.28	-0.04	0.13	0.07	0.17
Q72 p-value	0.03	0.00	0.00	0.38	0.00	0.02	0.00	0.00	0.31	0.00	0.08	0.00
Q73 r	0.08	0.12	0.19	0.04	0.12	0.11	0.18	0.29	-0.04	0.11	0.07	0.17
Q73 p-value	0.04	0.00	0.00	0.38	0.00	0.01	0.00	0.00	0.33	0.01	0.11	0.00
Q74 r	0.09	0.11	0.14	0.04	0.08	0.03	0.17	0.23	-0.05	0.11	0.04	0.11
Q74 p-value	0.03	0.01	0.00	0.31	0.06	0.40	0.00	0.00	0.23	0.01	0.32	0.01
Q75 r	0.10	0.14	0.14	0.04	0.08	0.02	0.20	0.22	-0.05	0.13	0.04	0.12
Q75 p-value	0.02	0.00	0.00	0.29	0.04	0.66	0.00	0.00	0.23	0.00	0.32	0.00
Q76 r	0.08	0.11	0.13	0.04	0.07	0.05	0.18	0.21	-0.05	0.12	0.04	0.13
Q76 p-value	0.04	0.01	0.00	0.34	0.07	0.20	0.00	0.00	0.19	0.00	0.38	0.00
Q77 r	-0.06	-0.09	-0.05	-0.03	-0.14	-0.15	-0.04	-0.11	-0.11	-0.06	-0.04	-0.06
Q77 p-value	0.15	0.03	0.25	0.50	0.00	0.00	0.36	0.01	0.01	0.14	0.32	0.14
Q78 r	-0.06	-0.09	-0.05	-0.03	-0.13	-0.15	-0.03	-0.10	-0.11	-0.05	-0.04	-0.05
Q78 p-value	0.15	0.03	0.23	0.47	0.00	0.00	0.44	0.02	0.01	0.23	0.29	0.25

TABLE A.8: *Part 8: Pearson's correlation coefficients and p-values for numerical variables*

	Q21	Q22	Q23	Q24	Q25	Q26	Q27	Q28	Q29	Q30	Q31	Q32
Volume r	-0.06	-0.04	-0.07	-0.09	-0.06	-0.06	-0.02	-0.04	-0.06	-0.04	-0.06	-0.02
Volume p-value	0.15	0.31	0.07	0.02	0.15	0.13	0.71	0.28	0.12	0.37	0.15	0.70
Value r	-0.02	-0.04	-0.13	-0.10	-0.07	-0.05	-0.05	-0.08	-0.06	-0.03	-0.08	-0.02
Value p-value	0.63	0.38	0.00	0.01	0.10	0.18	0.19	0.05	0.14	0.47	0.04	0.67
AUP r	0.14	-0.03	-0.09	-0.05	-0.04	-0.00	0.01	-0.06	-0.07	0.10	-0.04	0.13
AUP p-value	0.00	0.51	0.02	0.27	0.36	0.96	0.78	0.15	0.11	0.02	0.39	0.00
Q1 r	0.19	0.48	0.26	0.24	0.39	-0.05	-0.05	0.14	0.03	0.27	0.16	0.49
Q1 p-value	0.00	0.00	0.00	0.00	0.00	0.22	0.23	0.00	0.47	0.00	0.00	0.00
Q2 r	0.27	-0.06	0.36	0.13	0.02	0.08	0.05	0.17	0.10	0.08	0.36	0.04
Q2 p-value	0.00	0.16	0.00	0.00	0.58	0.04	0.19	0.00	0.02	0.04	0.00	0.31
Q3 r	0.30	-0.02	0.27	0.09	-0.00	-0.01	-0.09	0.14	0.18	0.04	0.34	-0.02
Q3 p-value	0.00	0.65	0.00	0.03	0.91	0.80	0.02	0.00	0.00	0.28	0.00	0.67
Q4 r	0.35	-0.12	0.31	0.10	0.01	0.01	0.05	0.26	0.16	0.17	0.41	0.05
Q4 p-value	0.00	0.00	0.00	0.01	0.89	0.75	0.27	0.00	0.00	0.00	0.00	0.25
Q5 r	-0.14	0.91	-0.00	0.32	0.69	-0.07	-0.07	-0.03	-0.08	-0.09	-0.09	-0.03
Q5 p-value	0.00	0.00	0.95	0.00	0.00	0.08	0.08	0.48	0.04	0.03	0.03	0.46
Q6 r	-0.05	0.23	0.02	-0.06	-0.07	-0.03	-0.04	-0.01	-0.03	0.29	-0.06	0.60
Q6 p-value	0.23	0.00	0.62	0.15	0.09	0.46	0.34	0.90	0.52	0.00	0.18	0.00
Q7 r	0.47	-0.03	0.47	0.11	0.02	0.05	-0.03	0.13	0.11	0.15	0.40	0.05
Q7 p-value	0.00	0.49	0.00	0.01	0.56	0.20	0.40	0.00	0.01	0.00	0.00	0.18
Q8 r	0.29	-0.06	0.58	0.21	0.07	0.09	0.27	0.40	0.02	0.13	0.41	0.12
Q8 p-value	0.00	0.14	0.00	0.00	0.08	0.03	0.00	0.00	0.70	0.00	0.00	0.00
Q9 r	0.08	-0.12	0.12	0.43	-0.05	0.54	0.06	0.14	0.04	0.46	0.04	0.03
Q9 p-value	0.06	0.00	0.00	0.00	0.20	0.00	0.14	0.00	0.34	0.00	0.30	0.48
Q10 r	0.18	-0.13	0.24	-0.00	-0.10	-0.04	0.21	0.29	0.06	0.00	0.01	0.08
Q10 p-value	0.00	0.00	0.00	0.97	0.02	0.30	0.00	0.00	0.17	0.94	0.86	0.06
Q11 r	0.20	-0.12	0.36	0.12	0.09	0.04	0.17	0.37	0.19	0.12	0.25	0.09
Q11 p-value	0.00	0.00	0.00	0.00	0.02	0.28	0.00	0.00	0.00	0.00	0.00	0.03
Q12 r	-0.02	-0.05	0.00	0.59	-0.04	0.77	-0.05	-0.03	-0.03	0.62	0.00	-0.02
Q12 p-value	0.71	0.18	0.98	0.00	0.34	0.00	0.23	0.40	0.42	0.00	0.91	0.67
Q13 r	0.36	0.54	0.27	0.31	0.47	-0.03	0.08	0.15	0.22	-0.01	0.19	0.00
Q13 p-value	0.00	0.00	0.00	0.00	0.00	0.53	0.05	0.00	0.00	0.76	0.00	0.96
Q14 r	0.26	-0.08	0.20	0.18	-0.01	-0.00	-0.08	0.31	0.48	0.01	0.10	-0.00
Q14 p-value	0.00	0.05	0.00	0.00	0.72	0.95	0.05	0.00	0.00	0.76	0.01	0.98
Q15 r	0.22	-0.03	0.24	0.15	-0.07	-0.03	0.25	0.49	0.49	0.03	0.08	0.08
Q15 p-value	0.00	0.43	0.00	0.00	0.11	0.43	0.00	0.00	0.00	0.42	0.05	0.04
Q16 r	0.35	-0.02	0.14	0.10	0.02	-0.03	0.02	0.36	0.30	0.02	-0.09	0.06
Q16 p-value	0.00	0.60	0.00	0.01	0.55	0.40	0.57	0.00	0.00	0.61	0.03	0.17
Q17 r	-0.05	0.31	-0.04	-0.04	-0.08	0.01	0.10	-0.05	0.04	-0.07	-0.04	-0.04
Q17 p-value	0.26	0.00	0.37	0.37	0.06	0.76	0.02	0.18	0.38	0.09	0.31	0.32
Q18 r	0.23	-0.06	0.18	0.14	0.07	0.08	0.09	0.25	0.56	0.18	0.52	0.05
Q18 p-value	0.00	0.17	0.00	0.00	0.09	0.04	0.03	0.00	0.00	0.00	0.00	0.27
Q19 r	-0.02	-0.09	0.09	-0.00	-0.04	-0.01	0.09	0.10	0.01	-0.02	-0.02	0.03
Q19 p-value	0.66	0.04	0.03	0.96	0.33	0.79	0.03	0.02	0.76	0.71	0.63	0.52
Q20 r	0.34	-0.13	0.21	0.07	0.06	-0.01	-0.16	0.11	0.50	0.12	0.22	0.06
Q20 p-value	0.00	0.00	0.00	0.09	0.15	0.87	0.00	0.01	0.00	0.00	0.00	0.17
Q21 r	1.00	-0.07	0.28	0.04	-0.03	-0.04	-0.09	0.13	0.25	0.05	0.19	0.03

Pearson's Correlation Coefficients N = 594 Prob > | r | under H0: Rho=0

TABLE A.9: *Part 9: Pearson's correlation coefficients and p-values for numerical variables*

	Q21	Q22	Q23	Q24	Q25	Q26	Q27	Q28	Q29	Q30	Q31	Q32
	Pearson's Correlation Coefficients N = 594 Prob $> \mid r \mid$ under H0: Rho=0											
Q21 p-value		0.10	0.00	0.35	0.41	0.38	0.03	0.00	0.00	0.20	0.00	0.50
Q22 r	-0.07	1.00	-0.08	0.32	0.69	-0.05	-0.15	-0.09	-0.02	-0.08	-0.07	-0.06
Q22 p-value	0.10		0.04	0.00	0.00	0.24	0.00	0.03	0.65	0.06	0.07	0.17
Q23 r	0.28	-0.08	1.00	0.17	0.05	0.05	0.28	0.41	0.02	0.13	0.54	0.08
Q23 p-value	0.00	0.04		0.00	0.22	0.23	0.00	0.00	0.67	0.00	0.00	0.05
Q24 r	0.04	0.32	0.17	1.00	0.52	0.79	0.12	0.25	0.22	0.59	0.16	-0.01
Q24 p-value	0.35	0.00	0.00		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.85
Q25 r	-0.03	0.69	0.05	0.52	1.00	0.02	-0.06	-0.02	0.10	0.02	0.17	-0.04
Q25 p-value	0.41	0.00	0.22	0.00		0.66	0.18	0.57	0.01	0.64	0.00	0.37
Q26 r	-0.04	-0.05	0.05	0.79	0.02	1.00	0.14	0.00	0.06	0.75	0.14	-0.01
Q26 p-value	0.38	0.24	0.23	0.00	0.66		0.00	0.94	0.13	0.00	0.00	0.74
Q27 r	-0.09	-0.15	0.28	0.12	-0.06	0.14	1.00	0.27	-0.12	0.00	0.11	0.07
Q27 p-value	0.03	0.00	0.00	0.00	0.18	0.00		0.00	0.00	0.96	0.01	0.11
Q28 r	0.13	-0.09	0.41	0.25	-0.02	0.00	0.27	1.00	0.31	0.02	0.12	0.07
Q28 p-value	0.00	0.03	0.00	0.00	0.57	0.94	0.00		0.00	0.69	0.00	0.07
Q29 r	0.25	-0.02	0.02	0.22	0.10	0.06	-0.12	0.31	1.00	0.14	0.27	0.05
Q29 p-value	0.00	0.65	0.67	0.00	0.01	0.13	0.00	0.00		0.00	0.00	0.27
Q30 r	0.05	-0.08	0.13	0.59	0.02	0.75	0.00	0.02	0.14	1.00	0.29	0.56
Q30 p-value	0.20	0.06	0.00	0.00	0.64	0.00	0.96	0.69	0.00		0.00	0.00
Q31 r	0.19	-0.07	0.54	0.16	0.17	0.14	0.11	0.12	0.27	0.29	1.00	0.03
Q31 p-value	0.00	0.07	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00		0.51
Q32 r	0.03	-0.06	0.08	-0.01	-0.04	-0.01	0.07	0.07	0.05	0.56	0.03	1.00
Q32 p-value	0.50	0.17	0.05	0.85	0.37	0.74	0.11	0.07	0.27	0.00	0.51	
Q33 r	0.02	-0.05	0.06	0.73	0.01	0.92	-0.01	-0.00	0.05	0.81	0.16	0.01
Q33 p-value	0.65	0.22	0.17	0.00	0.84	0.00	0.87	0.91	0.23	0.00	0.00	0.84
Q34 r	0.23	0.39	0.07	0.34	0.57	-0.03	0.12	0.22	-0.07	-0.01	0.00	0.06
Q34 p-value	0.00	0.00	0.11	0.00	0.00	0.49	0.00	0.00	0.10	0.78	1.00	0.13
Q35 r	0.14	0.42	-0.02	0.32	0.59	-0.03	0.10	0.12	-0.09	-0.03	-0.07	0.05
Q35 p-value	0.00	0.00	0.65	0.00	0.00	0.45	0.02	0.00	0.02	0.54	0.10	0.25
Q36 r	0.43	-0.09	0.51	0.13	0.00	0.02	0.19	0.50	0.12	0.07	0.32	0.07
Q36 p-value	0.00	0.02	0.00	0.00	0.94	0.64	0.00	0.00	0.00	0.09	0.00	0.07
Q37 r	0.14	0.60	0.25	0.29	0.50	-0.05	0.06	0.16	-0.01	0.01	0.14	0.04
Q37 p-value	0.00	0.00	0.00	0.00	0.00	0.27	0.18	0.00	0.77	0.88	0.00	0.33
Q38 r	-0.01	0.64	0.03	0.45	0.85	-0.03	0.01	-0.00	-0.07	-0.01	0.08	-0.01
Q38 p-value	0.74	0.00	0.50	0.00	0.00	0.46	0.72	0.94	0.10	0.87	0.04	0.85
Q39 r	0.04	-0.08	0.19	0.03	-0.03	-0.01	0.06	0.12	0.01	0.02	0.12	0.03
Q39 p-value	0.37	0.05	0.00	0.54	0.54	0.89	0.14	0.00	0.84	0.70	0.00	0.49
Q40 r	-0.02	0.51	0.03	-0.08	-0.08	-0.04	-0.05	-0.01	-0.03	-0.05	-0.06	-0.02
Q40 p-value	0.57	0.00	0.42	0.05	0.04	0.37	0.24	0.85	0.43	0.20	0.14	0.59
Q41 r	0.48	-0.13	0.38	0.16	0.01	0.03	0.29	0.45	0.22	0.14	0.17	0.19
Q41 p-value	0.00	0.00	0.00	0.00	0.79	0.41	0.00	0.00	0.00	0.00	0.00	0.00
Q42 r	-0.02	-0.00	0.01	-0.03	-0.02	-0.02	0.01	0.02	-0.01	-0.00	-0.02	0.03
Q42 p-value	0.70	0.98	0.89	0.51	0.65	0.57	0.80	0.63	0.77	0.99	0.62	0.47
Q43 r	0.22	-0.09	0.11	-0.10	-0.10	-0.13	0.13	0.14	0.19	-0.13	0.11	-0.10
Q43 p-value	0.00	0.03	0.01	0.01	0.02	0.00	0.00	0.00	0.00	0.00	0.01	0.02
Q44 r	0.24	-0.16	0.43	0.02	-0.07	0.05	0.44	0.24	0.03	0.05	0.34	0.04
Q44 p-value	0.00	0.00	0.00	0.58	0.09	0.24	0.00	0.00	0.43	0.23	0.00	0.32

TABLE A.10: *Part 10: Pearson's correlation coefficients and p-values for numerical variables*

| | Pearson's Correlation Coefficients N = 594 Prob > \| r \| under H0: Rho=0 | | | | | | | | | | | |
	Q21	Q22	Q23	Q24	Q25	Q26	Q27	Q28	Q29	Q30	Q31	Q32
Q45 r	0.16	-0.03	0.11	-0.01	-0.00	-0.02	0.12	0.14	0.04	0.06	0.07	0.05
Q45 p-value	0.00	0.42	0.01	0.76	0.91	0.69	0.00	0.00	0.28	0.14	0.11	0.24
Q46 r	0.21	-0.03	0.16	0.02	0.01	-0.01	0.13	0.18	0.08	0.07	0.08	0.05
Q46 p-value	0.00	0.44	0.00	0.69	0.74	0.90	0.00	0.00	0.06	0.08	0.05	0.26
Q47 r	0.13	-0.03	0.09	-0.02	-0.01	-0.02	0.12	0.13	0.03	0.06	0.06	0.06
Q47 p-value	0.00	0.43	0.04	0.60	0.72	0.67	0.00	0.00	0.49	0.14	0.14	0.17
Q48 r	0.18	-0.04	0.15	0.00	-0.00	-0.02	0.12	0.17	0.07	0.06	0.08	0.05
Q48 p-value	0.00	0.33	0.00	0.98	0.99	0.70	0.00	0.00	0.09	0.13	0.05	0.24
Q49 r	0.16	-0.03	0.09	-0.01	-0.01	-0.01	0.11	0.15	0.05	0.06	0.06	0.06
Q49 p-value	0.00	0.49	0.02	0.78	0.87	0.72	0.01	0.00	0.22	0.14	0.16	0.18
Q50 r	0.14	-0.03	0.08	-0.01	-0.00	-0.01	0.11	0.14	0.04	0.06	0.05	0.05
Q50 p-value	0.00	0.54	0.05	0.76	0.96	0.76	0.01	0.00	0.31	0.16	0.26	0.24
Q51 r	0.12	-0.03	0.06	-0.02	-0.01	-0.02	0.12	0.14	0.02	0.06	0.05	0.05
Q51 p-value	0.00	0.52	0.13	0.57	0.78	0.67	0.00	0.00	0.71	0.18	0.19	0.19
Q52 r	0.14	-0.02	0.07	-0.02	-0.01	-0.02	0.12	0.13	0.03	0.06	0.06	0.05
Q52 p-value	0.00	0.55	0.07	0.64	0.88	0.70	0.00	0.00	0.52	0.13	0.15	0.18
Q53 r	0.14	-0.03	0.07	-0.02	-0.00	-0.02	0.11	0.13	0.03	0.06	0.05	0.05
Q53 p-value	0.00	0.54	0.11	0.68	0.96	0.67	0.01	0.00	0.40	0.15	0.19	0.23
Q54 r	0.19	-0.03	0.15	0.01	0.01	-0.01	0.13	0.17	0.07	0.07	0.08	0.05
Q54 p-value	0.00	0.42	0.00	0.80	0.81	0.90	0.00	0.00	0.09	0.08	0.05	0.19
Q55 r	0.17	-0.03	0.07	-0.01	-0.00	-0.01	0.11	0.13	0.04	0.07	0.06	0.05
Q55 p-value	0.00	0.53	0.08	0.73	0.93	0.72	0.01	0.00	0.28	0.10	0.12	0.20
Q56 r	0.13	-0.02	0.05	-0.02	-0.01	-0.02	0.12	0.13	0.03	0.06	0.05	0.05
Q56 p-value	0.00	0.54	0.19	0.57	0.80	0.65	0.00	0.00	0.52	0.17	0.25	0.20
Q57 r	0.10	-0.02	0.00	-0.04	-0.02	-0.02	0.10	0.09	-0.00	0.05	0.04	0.05
Q57 p-value	0.01	0.66	0.94	0.34	0.68	0.55	0.01	0.02	0.97	0.21	0.31	0.21
Q58 r	0.14	-0.02	0.07	-0.02	-0.01	-0.02	0.12	0.14	0.03	0.06	0.06	0.05
Q58 p-value	0.00	0.59	0.09	0.63	0.86	0.65	0.00	0.00	0.50	0.15	0.17	0.22
Q59 r	0.13	-0.02	0.04	-0.03	-0.01	-0.02	0.11	0.12	0.02	0.05	0.04	0.05
Q59 p-value	0.00	0.58	0.29	0.49	0.74	0.60	0.01	0.00	0.66	0.19	0.32	0.19
Q60 r	0.16	-0.03	0.09	-0.01	-0.01	-0.02	0.11	0.15	0.06	0.06	0.07	0.05
Q60 p-value	0.00	0.40	0.03	0.78	0.86	0.65	0.01	0.00	0.17	0.15	0.10	0.21
Q61 r	0.20	-0.03	0.16	0.02	0.02	0.00	0.14	0.17	0.08	0.08	0.08	0.05
Q61 p-value	0.00	0.44	0.00	0.60	0.64	0.98	0.00	0.00	0.04	0.05	0.04	0.20
Q62 r	0.12	-0.03	0.05	-0.03	-0.01	-0.02	0.10	0.12	0.02	0.05	0.05	0.05
Q62 p-value	0.00	0.51	0.21	0.51	0.73	0.58	0.02	0.00	0.65	0.23	0.19	0.22
Q63 r	0.12	-0.03	0.09	-0.02	-0.01	-0.02	0.10	0.14	0.05	0.05	0.05	0.05
Q63 p-value	0.00	0.43	0.03	0.66	0.78	0.62	0.02	0.00	0.26	0.21	0.19	0.23
Q64 r	0.15	-0.03	0.10	-0.01	-0.00	-0.02	0.12	0.15	0.05	0.06	0.07	0.05
Q64 p-value	0.00	0.42	0.01	0.82	0.94	0.71	0.00	0.00	0.27	0.13	0.10	0.23
Q65 r	0.20	-0.03	0.12	-0.01	0.01	-0.02	0.14	0.14	0.04	0.08	0.09	0.04
Q65 p-value	0.00	0.53	0.00	0.88	0.88	0.67	0.00	0.00	0.30	0.07	0.03	0.35
Q66 r	0.15	-0.03	0.08	-0.01	-0.00	-0.01	0.13	0.14	0.03	0.06	0.06	0.05
Q66 p-value	0.00	0.54	0.06	0.73	0.98	0.72	0.00	0.00	0.46	0.12	0.14	0.21
Q67 r	0.12	-0.02	0.05	-0.02	-0.01	-0.02	0.12	0.12	0.03	0.06	0.05	0.05
Q67 p-value	0.00	0.62	0.22	0.60	0.86	0.67	0.00	0.00	0.49	0.18	0.27	0.19
Q68 r	0.15	-0.03	0.10	-0.01	-0.00	-0.01	0.11	0.14	0.05	0.06	0.07	0.05
Q68 p-value	0.00	0.47	0.01	0.78	0.98	0.72	0.01	0.00	0.27	0.14	0.11	0.22

TABLE A.11: *Part 11: Pearson's correlation coefficients and p-values for numerical variables*

| | Pearson's Correlation Coefficients N = 594 Prob $>$ $\mid r \mid$ under H0: Rho=0 | | | | | | | | | | | |
	Q21	Q22	Q23	Q24	Q25	Q26	Q27	Q28	Q29	Q30	Q31	Q32
Q69 r	0.10	-0.02	0.01	-0.03	-0.01	-0.02	0.12	0.11	-0.01	0.05	0.04	0.05
Q69 p-value	0.01	0.70	0.73	0.43	0.73	0.61	0.00	0.01	0.81	0.21	0.31	0.19
Q70 r	0.16	-0.03	0.10	-0.01	-0.00	-0.02	0.12	0.15	0.05	0.06	0.06	0.05
Q70 p-value	0.00	0.46	0.01	0.84	0.95	0.69	0.00	0.00	0.25	0.14	0.12	0.22
Q71 r	0.18	-0.04	0.12	-0.00	-0.00	-0.02	0.11	0.16	0.06	0.06	0.09	0.05
Q71 p-value	0.00	0.36	0.00	0.91	0.98	0.70	0.01	0.00	0.17	0.11	0.04	0.23
Q72 r	0.17	-0.03	0.14	-0.00	0.00	-0.01	0.13	0.15	0.06	0.07	0.08	0.05
Q72 p-value	0.00	0.40	0.00	0.95	0.94	0.79	0.00	0.00	0.13	0.09	0.05	0.22
Q73 r	0.17	-0.03	0.13	-0.00	0.00	-0.01	0.13	0.16	0.06	0.06	0.07	0.05
Q73 p-value	0.00	0.41	0.00	0.97	0.99	0.80	0.00	0.00	0.17	0.13	0.11	0.24
Q74 r	0.14	-0.02	0.04	-0.03	-0.01	-0.02	0.11	0.11	0.02	0.05	0.04	0.05
Q74 p-value	0.00	0.63	0.37	0.49	0.79	0.63	0.01	0.01	0.62	0.19	0.32	0.20
Q75 r	0.15	-0.02	0.05	-0.02	-0.01	-0.01	0.14	0.11	0.01	0.06	0.05	0.05
Q75 p-value	0.00	0.69	0.23	0.59	0.87	0.78	0.00	0.01	0.77	0.13	0.22	0.22
Q76 r	0.13	-0.03	0.04	-0.03	-0.01	-0.02	0.11	0.11	0.02	0.06	0.05	0.05
Q76 p-value	0.00	0.53	0.35	0.47	0.73	0.58	0.01	0.01	0.56	0.18	0.25	0.25
Q77 r	-0.12	-0.07	-0.08	-0.13	-0.11	-0.08	-0.02	-0.09	-0.08	-0.07	-0.10	-0.04
Q77 p-value	0.00	0.11	0.04	0.00	0.01	0.06	0.71	0.03	0.05	0.11	0.01	0.33
Q78 r	-0.12	-0.05	-0.08	-0.12	-0.09	-0.08	-0.02	-0.09	-0.07	-0.07	-0.11	-0.04
Q78 p-value	0.00	0.18	0.05	0.00	0.03	0.05	0.64	0.04	0.07	0.10	0.01	0.35

TABLE A.12: *Part 12: Pearson's correlation coefficients and p-values for numerical variables*

	Q33	Q34	Q35	Q36	Q37	Q38	Q39	Q40	Q41	Q42	Q43	Q44
	Pearson's Correlation Coefficients N = 594 Prob > | r | under H0: Rho=0										
Volume r	-0.03	-0.11	-0.13	-0.00	-0.12	-0.09	-0.04	-0.06	-0.10	0.02	0.08	0.03
Volume p-value	0.53	0.01	0.00	0.98	0.00	0.03	0.28	0.15	0.02	0.67	0.06	0.51
Value r	-0.01	0.00	-0.01	-0.03	-0.11	-0.07	-0.06	-0.07	-0.08	0.02	0.02	-0.00
Value p-value	0.76	0.97	0.83	0.47	0.01	0.11	0.17	0.09	0.05	0.69	0.65	0.99
AUP r	0.04	0.35	0.37	-0.08	-0.05	-0.02	-0.05	-0.03	0.05	-0.00	-0.02	0.03
AUP p-value	0.33	0.00	0.00	0.04	0.24	0.65	0.20	0.43	0.19	0.94	0.69	0.49
Q1 r	-0.02	0.35	0.33	0.18	0.73	0.44	0.44	0.33	0.17	0.00	-0.12	0.09
Q1 p-value	0.55	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.99	0.00	0.03
Q2 r	0.02	0.13	0.05	0.38	0.19	0.09	0.09	0.13	0.30	-0.02	0.05	0.41
Q2 p-value	0.63	0.00	0.21	0.00	0.00	0.03	0.03	0.00	0.00	0.59	0.27	0.00
Q3 r	-0.00	0.01	-0.04	0.22	0.08	0.08	0.03	0.14	-0.04	0.05	0.14	0.26
Q3 p-value	0.99	0.88	0.33	0.00	0.07	0.05	0.49	0.00	0.39	0.27	0.00	0.00
Q4 r	0.11	0.07	0.00	0.29	0.21	0.14	0.14	-0.03	0.33	-0.01	0.21	0.07
Q4 p-value	0.01	0.08	0.95	0.00	0.00	0.00	0.00	0.40	0.00	0.90	0.00	0.08
Q5 r	-0.07	0.39	0.41	-0.06	0.65	0.68	-0.02	0.49	-0.05	-0.03	-0.07	-0.17
Q5 p-value	0.11	0.00	0.00	0.15	0.00	0.00	0.57	0.00	0.22	0.46	0.07	0.00
Q6 r	-0.05	-0.04	-0.03	-0.05	0.45	-0.06	0.49	0.43	-0.02	-0.01	-0.19	-0.06
Q6 p-value	0.22	0.30	0.50	0.22	0.00	0.15	0.00	0.00	0.63	0.90	0.00	0.15
Q7 r	0.09	0.18	0.09	0.42	0.22	0.08	0.07	0.15	0.32	-0.04	0.12	0.39
Q7 p-value	0.03	0.00	0.02	0.00	0.00	0.05	0.10	0.00	0.00	0.36	0.00	0.00
Q8 r	0.08	0.29	0.20	0.53	0.26	0.13	0.19	-0.01	0.47	-0.00	0.04	0.44
Q8 p-value	0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.85	0.00	0.94	0.38	0.00
Q9 r	0.56	0.09	0.07	0.11	0.40	-0.02	0.70	-0.03	0.18	-0.00	-0.19	0.09
Q9 p-value	0.00	0.04	0.07	0.01	0.00	0.64	0.00	0.51	0.00	0.99	0.00	0.03
Q10 r	-0.02	0.22	0.18	0.22	0.54	0.06	0.77	0.03	0.35	0.05	-0.09	0.23
Q10 p-value	0.66	0.00	0.00	0.00	0.00	0.11	0.00	0.49	0.00	0.22	0.03	0.00
Q11 r	0.04	0.17	0.14	0.32	0.16	0.01	0.10	0.00	0.46	-0.07	0.11	0.33
Q11 p-value	0.31	0.00	0.00	0.00	0.00	0.77	0.01	0.98	0.00	0.11	0.01	0.00
Q12 r	0.81	-0.05	-0.04	-0.00	0.22	-0.05	0.48	-0.05	-0.01	-0.02	-0.21	-0.07
Q12 p-value	0.00	0.24	0.28	0.92	0.00	0.25	0.00	0.19	0.89	0.56	0.00	0.07
Q13 r	-0.05	0.43	0.40	0.24	0.84	0.50	0.49	0.37	0.31	0.00	0.05	0.20
Q13 p-value	0.20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.99	0.27	0.00
Q14 r	-0.03	-0.01	-0.04	0.16	0.12	-0.01	0.10	0.07	0.26	0.03	0.22	0.06
Q14 p-value	0.44	0.86	0.33	0.00	0.00	0.80	0.02	0.10	0.00	0.49	0.00	0.17
Q15 r	0.02	0.16	0.12	0.23	0.05	0.01	0.04	-0.04	0.27	-0.05	0.14	0.23
Q15 p-value	0.64	0.00	0.00	0.00	0.21	0.81	0.31	0.30	0.00	0.19	0.00	0.00
Q16 r	-0.01	0.14	0.12	0.21	0.13	0.03	0.03	0.08	0.37	0.01	0.18	0.20
Q16 p-value	0.77	0.00	0.00	0.00	0.00	0.48	0.40	0.06	0.00	0.83	0.00	0.00
Q17 r	-0.05	-0.14	-0.14	-0.04	0.56	-0.07	0.63	0.50	-0.07	-0.02	-0.07	0.00
Q17 p-value	0.27	0.00	0.00	0.35	0.00	0.09	0.00	0.00	0.09	0.66	0.10	0.93
Q18 r	0.08	0.14	0.10	0.21	0.07	0.06	0.02	0.06	0.18	-0.03	0.12	0.24
Q18 p-value	0.04	0.00	0.02	0.00	0.07	0.12	0.55	0.12	0.00	0.54	0.00	0.00
Q19 r	-0.02	0.02	0.02	0.07	0.54	-0.01	0.91	-0.00	0.19	-0.01	-0.16	0.05
Q19 p-value	0.68	0.58	0.67	0.08	0.00	0.85	0.00	0.96	0.00	0.72	0.00	0.20
Q20 r	0.03	0.03	-0.02	0.25	0.12	0.01	0.06	0.02	0.35	0.00	0.25	0.15
Q20 p-value	0.45	0.41	0.67	0.00	0.00	0.72	0.17	0.55	0.00	0.96	0.00	0.00
Q21 r	0.02	0.23	0.14	0.43	0.14	-0.01	0.04	-0.02	0.48	-0.02	0.22	0.24
Q21 p-value	0.65	0.00	0.00	0.00	0.00	0.74	0.37	0.57	0.00	0.70	0.00	0.00

TABLE A.13: *Part 13: Pearson's correlation coefficients and p-values for numerical variables*

	Q33	Q34	Q35	Q36	Q37	Q38	Q39	Q40	Q41	Q42	Q43	Q44
	Pearson's Correlation Coefficients N = 594 Prob > $\mid r \mid$ under H0: Rho=0											
Q22 r	-0.05	0.39	0.42	-0.09	0.60	0.64	-0.08	0.51	-0.13	-0.00	-0.09	-0.16
Q22 p-value	0.22	0.00	0.00	0.02	0.00	0.00	0.05	0.00	0.00	0.98	0.03	0.00
Q23 r	0.06	0.07	-0.02	0.51	0.25	0.03	0.19	0.03	0.38	0.01	0.11	0.43
Q23 p-value	0.17	0.11	0.65	0.00	0.00	0.50	0.00	0.42	0.00	0.89	0.01	0.00
Q24 r	0.73	0.34	0.32	0.13	0.29	0.45	0.03	-0.08	0.16	-0.03	-0.10	0.02
Q24 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.54	0.05	0.00	0.51	0.01	0.58
Q25 r	0.01	0.57	0.59	0.00	0.50	0.85	-0.03	-0.08	0.01	-0.02	-0.10	-0.07
Q25 p-value	0.84	0.00	0.00	0.94	0.00	0.00	0.54	0.04	0.79	0.65	0.02	0.09
Q26 r	0.92	-0.03	-0.03	0.02	-0.05	-0.03	-0.01	-0.04	0.03	-0.02	-0.13	0.05
Q26 p-value	0.00	0.49	0.45	0.64	0.27	0.46	0.89	0.37	0.41	0.57	0.00	0.24
Q27 r	-0.01	0.12	0.10	0.19	0.06	0.01	0.06	-0.05	0.29	0.01	0.13	0.44
Q27 p-value	0.87	0.00	0.02	0.00	0.18	0.72	0.14	0.24	0.00	0.80	0.00	0.00
Q28 r	-0.00	0.22	0.12	0.50	0.16	-0.00	0.12	-0.01	0.45	0.02	0.14	0.24
Q28 p-value	0.91	0.00	0.00	0.00	0.00	0.94	0.00	0.85	0.00	0.63	0.00	0.00
Q29 r	0.05	-0.07	-0.09	0.12	-0.01	-0.07	0.01	-0.03	0.22	-0.01	0.19	0.03
Q29 p-value	0.23	0.10	0.02	0.00	0.77	0.10	0.84	0.43	0.00	0.77	0.00	0.43
Q30 r	0.81	-0.01	-0.03	0.07	0.01	-0.01	0.02	-0.05	0.14	-0.00	-0.13	0.05
Q30 p-value	0.00	0.78	0.54	0.09	0.88	0.87	0.70	0.20	0.00	0.99	0.00	0.23
Q31 r	0.16	0.00	-0.07	0.32	0.14	0.08	0.12	-0.06	0.17	-0.02	0.11	0.34
Q31 p-value	0.00	1.00	0.10	0.00	0.00	0.04	0.00	0.14	0.00	0.62	0.01	0.00
Q32 r	0.01	0.06	0.05	0.07	0.04	-0.01	0.03	-0.02	0.19	0.03	-0.10	0.04
Q32 p-value	0.84	0.13	0.25	0.07	0.33	0.85	0.49	0.59	0.00	0.47	0.02	0.32
Q33 r	1.00	-0.02	-0.02	0.02	-0.02	-0.00	-0.00	-0.06	0.06	-0.02	-0.12	0.01
Q33 p-value		0.70	0.64	0.55	0.61	0.95	0.94	0.15	0.14	0.71	0.00	0.84
Q34 r	-0.02	1.00	0.97	0.22	0.45	0.63	-0.01	-0.05	0.36	0.03	-0.07	0.21
Q34 p-value	0.70		0.00	0.00	0.00	0.00	0.88	0.19	0.00	0.42	0.11	0.00
Q35 r	-0.02	0.97	1.00	0.01	0.42	0.64	-0.02	-0.05	0.27	0.03	-0.11	0.12
Q35 p-value	0.64	0.00		0.85	0.00	0.00	0.58	0.26	0.00	0.47	0.01	0.00
Q36 r	0.02	0.22	0.01	1.00	0.20	0.01	0.10	-0.01	0.48	0.08	0.18	0.48
Q36 p-value	0.55	0.00	0.85		0.00	0.86	0.02	0.78	0.00	0.06	0.00	0.00
Q37 r	-0.02	0.45	0.42	0.20	1.00	0.60	0.58	0.39	0.37	-0.03	-0.12	0.07
Q37 p-value	0.61	0.00	0.00	0.00		0.00	0.00	0.00	0.00	0.44	0.00	0.09
Q38 r	-0.00	0.63	0.64	0.01	0.60	1.00	-0.00	-0.04	0.10	-0.00	-0.11	-0.01
Q38 p-value	0.95	0.00	0.00	0.86	0.00		0.96	0.29	0.02	0.95	0.01	0.78
Q39 r	-0.00	-0.01	-0.02	0.10	0.58	-0.00	1.00	-0.05	0.17	-0.03	-0.17	-0.02
Q39 p-value	0.94	0.88	0.58	0.02	0.00	0.96		0.27	0.00	0.44	0.00	0.67
Q40 r	-0.06	-0.05	-0.05	-0.01	0.39	-0.04	-0.05	1.00	-0.04	-0.03	0.05	0.05
Q40 p-value	0.15	0.19	0.26	0.78	0.00	0.29	0.27		0.36	0.52	0.18	0.25
Q41 r	0.06	0.36	0.27	0.48	0.37	0.10	0.17	-0.04	1.00	-0.01	0.11	0.30
Q41 p-value	0.14	0.00	0.00	0.00	0.00	0.02	0.00	0.36		0.82	0.01	0.00
Q42 r	-0.02	0.03	0.03	0.08	-0.03	-0.00	-0.03	-0.03	-0.01	1.00	-0.00	0.05
Q42 p-value	0.71	0.42	0.47	0.06	0.44	0.95	0.44	0.52	0.82		0.98	0.26
Q43 r	-0.12	-0.07	-0.11	0.18	-0.12	-0.11	-0.17	0.05	0.11	-0.00	1.00	0.21
Q43 p-value	0.00	0.11	0.01	0.00	0.00	0.01	0.00	0.18	0.01	0.98		0.00
Q44 r	0.01	0.21	0.12	0.48	0.07	-0.01	-0.02	0.05	0.30	0.05	0.21	1.00
Q44 p-value	0.84	0.00	0.00	0.00	0.09	0.78	0.67	0.25	0.00	0.26	0.00	
Q45 r	0.04	0.13	0.09	0.29	0.02	0.00	-0.03	0.01	0.19	-0.03	0.31	0.31
Q45 p-value	0.31	0.00	0.04	0.00	0.68	0.91	0.46	0.82	0.00	0.51	0.00	0.00

TABLE A.14: *Part 14: Pearson's correlation coefficients and p-values for numerical variables*

	Pearson's Correlation Coefficients N = 594 Prob > | r | under H0: Rho=0											
	Q33	Q34	Q35	Q36	Q37	Q38	Q39	Q40	Q41	Q42	Q43	Q44
Q46 r	0.05	0.18	0.13	0.32	0.05	0.03	-0.02	0.02	0.24	-0.03	0.32	0.33
Q46 p-value	0.22	0.00	0.00	0.00	0.25	0.48	0.59	0.56	0.00	0.46	0.00	0.00
Q47 r	0.04	0.14	0.09	0.28	0.01	-0.00	-0.04	0.00	0.17	-0.01	0.28	0.31
Q47 p-value	0.35	0.00	0.03	0.00	0.87	1.00	0.38	0.97	0.00	0.77	0.00	0.00
Q48 r	0.04	0.11	0.06	0.32	0.03	0.02	-0.03	0.01	0.21	-0.03	0.32	0.33
Q48 p-value	0.32	0.01	0.13	0.00	0.45	0.69	0.54	0.77	0.00	0.46	0.00	0.00
Q49 r	0.04	0.14	0.10	0.28	0.02	0.01	-0.03	0.01	0.18	-0.03	0.29	0.30
Q49 p-value	0.34	0.00	0.02	0.00	0.71	0.88	0.44	0.85	0.00	0.45	0.00	0.00
Q50 r	0.04	0.14	0.09	0.28	0.01	0.00	-0.04	0.01	0.18	-0.03	0.29	0.31
Q50 p-value	0.33	0.00	0.02	0.00	0.82	0.97	0.34	0.83	0.00	0.46	0.00	0.00
Q51 r	0.04	0.14	0.09	0.26	0.01	-0.01	-0.03	-0.00	0.16	-0.02	0.27	0.29
Q51 p-value	0.33	0.00	0.03	0.00	0.90	0.87	0.43	0.97	0.00	0.58	0.00	0.00
Q52 r	0.04	0.17	0.12	0.26	0.01	0.00	-0.04	0.00	0.18	-0.03	0.28	0.30
Q52 p-value	0.28	0.00	0.00	0.00	0.77	0.94	0.37	0.91	0.00	0.54	0.00	0.00
Q53 r	0.04	0.15	0.11	0.27	0.01	0.00	-0.04	0.01	0.17	-0.03	0.29	0.30
Q53 p-value	0.29	0.00	0.01	0.00	0.83	0.97	0.39	0.87	0.00	0.54	0.00	0.00
Q54 r	0.05	0.16	0.11	0.33	0.04	0.02	-0.02	0.02	0.24	-0.04	0.31	0.33
Q54 p-value	0.24	0.00	0.01	0.00	0.30	0.54	0.58	0.66	0.00	0.39	0.00	0.00
Q55 r	0.05	0.17	0.12	0.27	0.02	0.01	-0.03	0.01	0.18	-0.04	0.29	0.29
Q55 p-value	0.26	0.00	0.00	0.00	0.61	0.77	0.44	0.75	0.00	0.38	0.00	0.00
Q56 r	0.04	0.15	0.11	0.26	0.00	-0.01	-0.03	0.00	0.17	-0.02	0.28	0.29
Q56 p-value	0.31	0.00	0.01	0.00	0.92	0.89	0.40	0.92	0.00	0.57	0.00	0.00
Q57 r	0.04	0.16	0.12	0.21	-0.01	-0.01	-0.04	0.00	0.12	-0.01	0.25	0.25
Q57 p-value	0.39	0.00	0.00	0.00	0.83	0.74	0.39	0.99	0.00	0.74	0.00	0.00
Q58 r	0.04	0.16	0.12	0.25	0.01	-0.00	-0.03	0.00	0.16	-0.03	0.28	0.28
Q58 p-value	0.30	0.00	0.00	0.00	0.83	0.97	0.44	0.94	0.00	0.47	0.00	0.00
Q59 r	0.04	0.15	0.10	0.24	-0.00	-0.01	-0.04	0.00	0.15	-0.03	0.27	0.28
Q59 p-value	0.34	0.00	0.01	0.00	0.96	0.84	0.39	0.97	0.00	0.49	0.00	0.00
Q60 r	0.04	0.12	0.06	0.29	0.02	0.00	-0.03	0.01	0.18	-0.03	0.30	0.30
Q60 p-value	0.34	0.00	0.13	0.00	0.65	0.93	0.50	0.83	0.00	0.44	0.00	0.00
Q61 r	0.05	0.21	0.17	0.30	0.06	0.04	-0.02	0.03	0.26	-0.04	0.31	0.33
Q61 p-value	0.19	0.00	0.00	0.00	0.15	0.30	0.63	0.45	0.00	0.38	0.00	0.00
Q62 r	0.03	0.10	0.04	0.26	-0.01	-0.02	-0.03	-0.01	0.14	-0.02	0.28	0.27
Q62 p-value	0.42	0.02	0.28	0.00	0.90	0.65	0.48	0.86	0.00	0.71	0.00	0.00
Q63 r	0.03	0.08	0.03	0.28	0.00	-0.01	-0.03	-0.00	0.16	-0.02	0.29	0.29
Q63 p-value	0.43	0.05	0.51	0.00	0.95	0.84	0.42	0.98	0.00	0.66	0.00	0.00
Q64 r	0.04	0.15	0.10	0.28	0.02	0.01	-0.03	0.01	0.18	-0.03	0.30	0.30
Q64 p-value	0.31	0.00	0.02	0.00	0.60	0.81	0.52	0.87	0.00	0.44	0.00	0.00
Q65 r	0.06	0.23	0.21	0.23	0.04	0.03	-0.02	0.02	0.20	-0.04	0.29	0.27
Q65 p-value	0.16	0.00	0.00	0.00	0.33	0.46	0.61	0.65	0.00	0.34	0.00	0.00
Q66 r	0.05	0.20	0.16	0.25	0.02	0.01	-0.03	0.01	0.19	-0.03	0.28	0.30
Q66 p-value	0.24	0.00	0.00	0.00	0.62	0.79	0.49	0.83	0.00	0.47	0.00	0.00
Q67 r	0.04	0.16	0.12	0.25	0.01	-0.00	-0.03	0.00	0.16	-0.02	0.27	0.30
Q67 p-value	0.30	0.00	0.00	0.00	0.90	0.93	0.40	0.92	0.00	0.63	0.00	0.00
Q68 r	0.04	0.12	0.07	0.29	0.02	0.00	-0.03	0.00	0.19	-0.03	0.30	0.31
Q68 p-value	0.30	0.00	0.08	0.00	0.69	0.92	0.42	0.92	0.00	0.45	0.00	0.00
Q69 r	0.04	0.16	0.12	0.21	-0.01	-0.02	-0.03	-0.00	0.12	-0.02	0.24	0.24
Q69 p-value	0.37	0.00	0.00	0.00	0.83	0.67	0.45	0.91	0.00	0.67	0.00	0.00

TABLE A.15: *Part 15: Pearson's correlation coefficients and p-values for numerical variables*

	Q33	Q34	Q35	Q36	Q37	Q38	Q39	Q40	Q41	Q42	Q43	Q44
	Pearson's Correlation Coefficients N = 594 Prob $> \mid r \mid$ under H0: Rho=0											
Q70 r	0.04	0.14	0.10	0.28	0.02	0.01	-0.03	0.01	0.18	-0.02	0.30	0.31
Q70 p-value	0.31	0.00	0.02	0.00	0.65	0.88	0.49	0.83	0.00	0.63	0.00	0.00
Q71 r	0.04	0.13	0.07	0.30	0.03	0.01	-0.02	0.01	0.19	-0.04	0.31	0.30
Q71 p-value	0.28	0.00	0.07	0.00	0.46	0.77	0.66	0.81	0.00	0.35	0.00	0.00
Q72 r	0.05	0.15	0.10	0.32	0.03	0.02	-0.03	0.02	0.21	-0.04	0.31	0.34
Q72 p-value	0.24	0.00	0.01	0.00	0.47	0.59	0.41	0.69	0.00	0.32	0.00	0.00
Q73 r	0.04	0.13	0.09	0.30	0.03	0.01	-0.03	0.01	0.20	-0.03	0.31	0.32
Q73 p-value	0.30	0.00	0.04	0.00	0.54	0.77	0.46	0.74	0.00	0.48	0.00	0.00
Q74 r	0.04	0.15	0.11	0.24	0.00	-0.01	-0.04	0.00	0.15	-0.03	0.27	0.27
Q74 p-value	0.34	0.00	0.01	0.00	0.97	0.85	0.39	0.96	0.00	0.52	0.00	0.00
Q75 r	0.05	0.19	0.15	0.24	0.01	0.00	-0.04	0.01	0.17	-0.04	0.27	0.29
Q75 p-value	0.24	0.00	0.00	0.00	0.86	0.96	0.37	0.89	0.00	0.39	0.00	0.00
Q76 r	0.04	0.16	0.12	0.23	0.00	-0.00	-0.04	-0.00	0.15	-0.03	0.27	0.28
Q76 p-value	0.33	0.00	0.00	0.00	0.96	0.92	0.39	1.00	0.00	0.49	0.00	0.00
Q77 r	-0.04	-0.11	-0.12	-0.02	-0.17	-0.12	-0.08	-0.07	-0.15	0.02	0.11	0.04
Q77 p-value	0.35	0.01	0.00	0.70	0.00	0.00	0.06	0.11	0.00	0.62	0.01	0.34
Q78 r	-0.04	-0.11	-0.12	-0.01	-0.16	-0.11	-0.08	-0.06	-0.15	0.02	0.11	0.04
Q78 p-value	0.33	0.01	0.00	0.83	0.00	0.01	0.05	0.13	0.00	0.62	0.01	0.34

TABLE A.16: *Part 16: Pearson's correlation coefficients and p-values for numerical variables*

| | Pearson's Correlation Coefficients N = 594 Prob $> \mid r \mid$ under H0: Rho=0 | | | | | | | | | | | |
	Q45	Q46	Q47	Q48	Q49	Q50	Q51	Q52	Q53	Q54	Q55	Q56
Volume r	0.12	0.05	0.13	0.10	0.11	0.13	0.15	0.10	0.13	0.08	0.09	0.14
Volume p-value	0.00	0.22	0.00	0.02	0.01	0.00	0.00	0.01	0.00	0.06	0.02	0.00
Value r	0.03	-0.02	0.05	0.00	0.03	0.04	0.06	0.03	0.05	-0.01	0.03	0.06
Value p-value	0.42	0.67	0.26	0.97	0.46	0.34	0.14	0.45	0.23	0.87	0.52	0.12
AUP r	0.04	0.07	0.04	-0.01	0.05	0.04	0.04	0.08	0.07	0.04	0.08	0.07
AUP p-value	0.32	0.07	0.37	0.76	0.26	0.33	0.31	0.05	0.11	0.37	0.05	0.09
Q1 r	0.01	0.04	0.01	0.03	0.02	0.00	-0.00	-0.00	-0.00	0.03	0.01	-0.01
Q1 p-value	0.78	0.37	0.82	0.41	0.67	0.92	0.92	0.96	0.99	0.42	0.83	0.81
Q2 r	0.11	0.13	0.08	0.15	0.09	0.10	0.08	0.07	0.09	0.14	0.08	0.07
Q2 p-value	0.01	0.00	0.04	0.00	0.02	0.02	0.06	0.07	0.04	0.00	0.04	0.08
Q3 r	0.16	0.14	0.16	0.17	0.16	0.15	0.14	0.14	0.14	0.15	0.16	0.13
Q3 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q4 r	0.13	0.18	0.11	0.17	0.13	0.11	0.09	0.09	0.11	0.16	0.12	0.08
Q4 p-value	0.00	0.00	0.01	0.00	0.00	0.01	0.04	0.02	0.01	0.00	0.00	0.06
Q5 r	-0.05	-0.05	-0.05	-0.04	-0.05	-0.05	-0.06	-0.06	-0.06	-0.05	-0.06	-0.06
Q5 p-value	0.19	0.23	0.19	0.32	0.19	0.18	0.16	0.13	0.13	0.24	0.13	0.13
Q6 r	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.05	-0.04
Q6 p-value	0.29	0.34	0.30	0.38	0.31	0.35	0.33	0.29	0.29	0.37	0.26	0.30
Q7 r	0.19	0.23	0.17	0.22	0.19	0.18	0.16	0.17	0.18	0.23	0.19	0.16
Q7 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q8 r	0.14	0.18	0.14	0.16	0.14	0.13	0.12	0.13	0.12	0.19	0.13	0.12
Q8 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q9 r	0.08	0.10	0.08	0.08	0.09	0.09	0.09	0.09	0.09	0.10	0.09	0.09
Q9 p-value	0.04	0.01	0.04	0.05	0.03	0.04	0.04	0.02	0.03	0.02	0.02	0.03
Q10 r	0.12	0.16	0.12	0.12	0.13	0.12	0.11	0.13	0.12	0.14	0.13	0.12
Q10 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00
Q11 r	0.17	0.19	0.16	0.20	0.17	0.17	0.16	0.16	0.15	0.20	0.16	0.16
Q11 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q12 r	0.04	0.04	0.03	0.03	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04
Q12 p-value	0.37	0.33	0.44	0.42	0.38	0.34	0.34	0.31	0.31	0.35	0.30	0.31
Q13 r	0.11	0.15	0.10	0.13	0.11	0.11	0.08	0.10	0.09	0.14	0.11	0.08
Q13 p-value	0.01	0.00	0.02	0.00	0.01	0.01	0.04	0.02	0.02	0.00	0.01	0.04
Q14 r	0.09	0.15	0.04	0.13	0.08	0.07	0.03	0.05	0.07	0.12	0.08	0.04
Q14 p-value	0.03	0.00	0.29	0.00	0.06	0.09	0.42	0.21	0.09	0.00	0.05	0.31
Q15 r	0.18	0.21	0.18	0.19	0.19	0.18	0.17	0.18	0.18	0.20	0.18	0.17
Q15 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q16 r	0.27	0.30	0.24	0.30	0.27	0.28	0.24	0.26	0.26	0.31	0.26	0.24
Q16 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q17 r	-0.04	-0.04	-0.05	-0.04	-0.05	-0.04	-0.05	-0.05	-0.05	-0.04	-0.05	-0.05
Q17 p-value	0.29	0.36	0.22	0.35	0.27	0.30	0.24	0.23	0.27	0.34	0.24	0.23
Q18 r	0.11	0.13	0.11	0.12	0.13	0.11	0.11	0.12	0.12	0.12	0.13	0.11
Q18 p-value	0.01	0.00	0.01	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.01
Q19 r	0.06	0.08	0.05	0.07	0.05	0.06	0.05	0.05	0.05	0.08	0.05	0.05
Q19 p-value	0.17	0.06	0.22	0.09	0.20	0.15	0.24	0.23	0.24	0.07	0.21	0.23
Q20 r	0.16	0.20	0.13	0.18	0.15	0.15	0.11	0.13	0.14	0.18	0.15	0.12
Q20 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00
Q21 r	0.16	0.21	0.13	0.18	0.16	0.14	0.12	0.14	0.14	0.19	0.17	0.13
Q21 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

TABLE A.17: *Part 17: Pearson's correlation coefficients and p-values for numerical variables*

	Q45	Q46	Q47	Q48	Q49	Q50	Q51	Q52	Q53	Q54	Q55	Q56
	Pearson's Correlation Coefficients N = 594 Prob $>$ $|r|$ under H0: Rho=0											
Q22 r	-0.03	-0.03	-0.03	-0.04	-0.03	-0.03	-0.03	-0.02	-0.03	-0.03	-0.03	-0.02
Q22 p-value	0.42	0.44	0.43	0.33	0.49	0.54	0.52	0.55	0.54	0.42	0.53	0.54
Q23 r	0.11	0.16	0.09	0.15	0.09	0.08	0.06	0.07	0.07	0.15	0.07	0.05
Q23 p-value	0.01	0.00	0.04	0.00	0.02	0.05	0.13	0.07	0.11	0.00	0.08	0.19
Q24 r	-0.01	0.02	-0.02	0.00	-0.01	-0.01	-0.02	-0.02	-0.02	0.01	-0.01	-0.02
Q24 p-value	0.76	0.69	0.60	0.98	0.78	0.76	0.57	0.64	0.68	0.80	0.73	0.57
Q25 r	-0.00	0.01	-0.01	-0.00	-0.01	-0.00	-0.01	-0.01	-0.00	0.01	-0.00	-0.01
Q25 p-value	0.91	0.74	0.72	0.99	0.87	0.96	0.78	0.88	0.96	0.81	0.93	0.80
Q26 r	-0.02	-0.01	-0.02	-0.02	-0.01	-0.01	-0.02	-0.02	-0.02	-0.01	-0.01	-0.02
Q26 p-value	0.69	0.90	0.67	0.70	0.72	0.76	0.67	0.70	0.67	0.90	0.72	0.65
Q27 r	0.12	0.13	0.12	0.12	0.11	0.11	0.12	0.12	0.11	0.13	0.11	0.12
Q27 p-value	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.01	0.00	0.01	0.00
Q28 r	0.14	0.18	0.13	0.17	0.15	0.14	0.14	0.13	0.13	0.17	0.13	0.13
Q28 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q29 r	0.04	0.08	0.03	0.07	0.05	0.04	0.02	0.03	0.03	0.07	0.04	0.03
Q29 p-value	0.28	0.06	0.49	0.09	0.22	0.31	0.71	0.52	0.40	0.09	0.28	0.52
Q30 r	0.06	0.07	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.07	0.07	0.06
Q30 p-value	0.14	0.08	0.14	0.13	0.14	0.16	0.18	0.13	0.15	0.08	0.10	0.17
Q31 r	0.07	0.08	0.06	0.08	0.06	0.05	0.05	0.06	0.05	0.08	0.06	0.05
Q31 p-value	0.11	0.05	0.14	0.05	0.16	0.26	0.19	0.15	0.19	0.05	0.12	0.25
Q32 r	0.05	0.05	0.06	0.05	0.06	0.05	0.05	0.05	0.05	0.05	0.05	0.05
Q32 p-value	0.24	0.26	0.17	0.24	0.18	0.24	0.19	0.18	0.23	0.19	0.20	0.20
Q33 r	0.04	0.05	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.05	0.05	0.04
Q33 p-value	0.31	0.22	0.35	0.32	0.34	0.33	0.33	0.28	0.29	0.24	0.26	0.31
Q34 r	0.13	0.18	0.14	0.11	0.14	0.14	0.14	0.17	0.15	0.16	0.17	0.15
Q34 p-value	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q35 r	0.09	0.13	0.09	0.06	0.10	0.09	0.09	0.12	0.11	0.11	0.12	0.11
Q35 p-value	0.04	0.00	0.03	0.13	0.02	0.02	0.03	0.00	0.01	0.01	0.00	0.01
Q36 r	0.29	0.32	0.28	0.32	0.28	0.28	0.26	0.26	0.27	0.33	0.27	0.26
Q36 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q37 r	0.02	0.05	0.01	0.03	0.02	0.01	0.01	0.01	0.01	0.04	0.02	0.00
Q37 p-value	0.68	0.25	0.87	0.45	0.71	0.82	0.90	0.77	0.83	0.30	0.61	0.92
Q38 r	0.00	0.03	-0.00	0.02	0.01	0.00	-0.01	0.00	0.00	0.02	0.01	-0.01
Q38 p-value	0.91	0.48	1.00	0.69	0.88	0.97	0.87	0.94	0.97	0.54	0.77	0.89
Q39 r	-0.03	-0.02	-0.04	-0.03	-0.03	-0.04	-0.03	-0.04	-0.04	-0.02	-0.03	-0.03
Q39 p-value	0.46	0.59	0.38	0.54	0.44	0.34	0.43	0.37	0.39	0.58	0.44	0.40
Q40 r	0.01	0.02	0.00	0.01	0.01	0.01	-0.00	0.00	0.01	0.02	0.01	0.00
Q40 p-value	0.82	0.56	0.97	0.77	0.85	0.83	0.97	0.91	0.87	0.66	0.75	0.92
Q41 r	0.19	0.24	0.17	0.21	0.18	0.18	0.16	0.18	0.17	0.24	0.18	0.17
Q41 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q42 r	-0.03	-0.03	-0.01	-0.03	-0.03	-0.03	-0.02	-0.03	-0.03	-0.04	-0.04	-0.02
Q42 p-value	0.51	0.46	0.77	0.46	0.45	0.46	0.58	0.54	0.54	0.39	0.38	0.57
Q43 r	0.31	0.32	0.28	0.32	0.29	0.29	0.27	0.28	0.29	0.31	0.29	0.28
Q43 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q44 r	0.31	0.33	0.31	0.33	0.30	0.31	0.29	0.30	0.30	0.33	0.29	0.29
Q44 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q45 r	1.00	0.97	0.98	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
Q45 p-value		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

TABLE A.18: *Part 18: Pearson's correlation coefficients and p-values for numerical variables*

	Q45	Q46	Q47	Q48	Q49	Q50	Q51	Q52	Q53	Q54	Q55	Q56
	Pearson's Correlation Coefficients N = 594 Prob > | r | under H0: Rho=0											
Q46 r	0.97	1.00	0.94	0.97	0.96	0.97	0.93	0.96	0.96	0.99	0.97	0.94
Q46 p-value	0.00		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q47 r	0.98	0.94	1.00	0.96	0.99	0.98	0.98	0.98	0.97	0.97	0.98	0.98
Q47 p-value	0.00	0.00		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q48 r	0.99	0.97	0.96	1.00	0.98	0.98	0.96	0.97	0.97	0.99	0.97	0.96
Q48 p-value	0.00	0.00	0.00		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q49 r	0.99	0.96	0.99	0.98	1.00	0.98	0.98	0.99	0.98	0.98	0.99	0.98
Q49 p-value	0.00	0.00	0.00	0.00		0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q50 r	0.99	0.97	0.98	0.98	0.98	1.00	0.99	0.99	0.99	0.98	0.98	0.99
Q50 p-value	0.00	0.00	0.00	0.00	0.00		0.00	0.00	0.00	0.00	0.00	0.00
Q51 r	0.99	0.93	0.98	0.96	0.98	0.99	1.00	0.99	0.99	0.96	0.98	0.99
Q51 p-value	0.00	0.00	0.00	0.00	0.00	0.00		0.00	0.00	0.00	0.00	0.00
Q52 r	0.99	0.96	0.98	0.97	0.99	0.99	0.99	1.00	0.99	0.98	0.99	0.99
Q52 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00		0.00	0.00	0.00	0.00
Q53 r	0.99	0.96	0.97	0.97	0.98	0.99	0.99	0.99	1.00	0.98	0.99	0.99
Q53 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00		0.00	0.00	0.00
Q54 r	0.99	0.99	0.97	0.99	0.98	0.98	0.96	0.98	0.98	1.00	0.98	0.97
Q54 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00		0.00	0.00
Q55 r	0.99	0.97	0.98	0.97	0.99	0.98	0.98	0.99	0.99	0.98	1.00	0.99
Q55 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00		0.00
Q56 r	0.99	0.94	0.98	0.96	0.98	0.99	0.99	0.99	0.99	0.97	0.99	1.00
Q56 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
Q57 r	0.96	0.90	0.96	0.91	0.96	0.95	0.98	0.98	0.97	0.92	0.97	0.98
Q57 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q58 r	0.99	0.95	0.98	0.96	0.99	0.99	0.99	0.99	0.99	0.97	0.99	0.99
Q58 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q59 r	0.98	0.93	0.98	0.95	0.98	0.98	1.00	0.99	0.99	0.96	0.99	1.00
Q59 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q60 r	0.99	0.95	0.97	0.98	0.98	0.99	0.99	0.98	0.99	0.98	0.98	0.98
Q60 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q61 r	0.96	0.99	0.93	0.96	0.95	0.95	0.92	0.95	0.95	0.98	0.96	0.93
Q61 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q62 r	0.98	0.91	0.97	0.95	0.97	0.98	0.99	0.97	0.98	0.95	0.97	0.98
Q62 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q63 r	0.98	0.93	0.98	0.98	0.98	0.98	0.98	0.97	0.97	0.97	0.97	0.98
Q63 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q64 r	1.00	0.97	0.99	0.99	1.00	0.99	0.98	0.99	0.99	0.99	0.99	0.99
Q64 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q65 r	0.91	0.94	0.89	0.88	0.91	0.90	0.88	0.92	0.92	0.92	0.94	0.91
Q65 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q66 r	0.98	0.96	0.97	0.95	0.98	0.97	0.97	0.99	0.99	0.97	0.99	0.98
Q66 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q67 r	0.98	0.94	0.97	0.95	0.98	0.98	0.99	0.99	0.99	0.96	0.99	0.99
Q67 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q68 r	1.00	0.96	0.98	0.99	0.99	1.00	0.99	0.99	0.99	0.99	0.98	0.99
Q68 p-value	0.00	0.00	Stellenbosch 0.00	0.00	0.00	https0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q69 r	0.95	0.89	0.95	0.90	0.96	0.95	0.98	0.97	0.97	0.91	0.96	0.98
Q69 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

TABLE A.19: *Part 19: Pearson's correlation coefficients and p-values for numerical variables*

	Q45	Q46	Q47	Q48	Q49	Q50	Q51	Q52	Q53	Q54	Q55	Q56
Pearson's Correlation Coefficients $N = 594$ Prob $> \mid r \mid$ under H0: Rho=0												
Q70 r	1.00	0.97	0.98	0.98	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
Q70 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q71 r	0.99	0.97	0.97	0.98	0.99	0.99	0.98	0.98	0.99	0.99	0.98	0.98
Q71 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q72 r	0.99	0.98	0.97	0.99	0.98	0.98	0.95	0.97	0.97	0.99	0.98	0.96
Q72 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q73 r	1.00	0.98	0.98	0.99	0.99	0.99	0.97	0.98	0.98	1.00	0.98	0.98
Q73 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q74 r	0.98	0.93	0.98	0.95	0.98	0.98	0.99	0.99	0.99	0.96	0.99	1.00
Q74 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q75 r	0.97	0.94	0.97	0.93	0.97	0.97	0.97	0.98	0.98	0.95	0.98	0.98
Q75 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q76 r	0.98	0.94	0.98	0.94	0.98	0.97	0.98	0.99	0.98	0.95	0.99	0.99
Q76 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q77 r	0.08	0.01	0.10	0.06	0.07	0.07	0.08	0.06	0.07	0.04	0.05	0.10
Q77 p-value	0.05	0.72	0.01	0.14	0.08	0.07	0.04	0.14	0.10	0.37	0.20	0.01
Q78 r	0.09	0.02	0.11	0.07	0.08	0.08	0.09	0.06	0.08	0.04	0.06	0.11
Q78 p-value	0.03	0.66	0.01	0.11	0.06	0.05	0.03	0.12	0.07	0.31	0.16	0.01

TABLE A.20: *Part 20: Pearson's correlation coefficients and p-values for numerical variables*

	Pearson's Correlation Coefficients N = 594 Prob $> \mid r \mid$ under H0: Rho=0											
	Q57	Q58	Q59	Q60	Q61	Q62	Q63	Q64	Q65	Q66	Q67	Q68
Volume r	0.14	0.13	0.15	0.13	0.01	0.18	0.16	0.11	0.03	0.10	0.13	0.12
Volume p-value	0.00	0.00	0.00	0.00	0.81	0.00	0.00	0.01	0.48	0.02	0.00	0.00
Value r	0.07	0.06	0.07	0.03	-0.04	0.08	0.05	0.03	0.00	0.04	0.05	0.03
Value p-value	0.08	0.15	0.07	0.40	0.34	0.06	0.22	0.50	0.91	0.34	0.19	0.45
AUP r	0.10	0.08	0.07	-0.00	0.10	-0.00	-0.03	0.04	0.20	0.12	0.09	0.02
AUP p-value	0.02	0.04	0.10	0.99	0.02	0.95	0.47	0.31	0.00	0.00	0.04	0.71
Q1 r	-0.02	-0.00	-0.01	0.02	0.04	-0.00	0.01	0.02	0.00	-0.00	-0.01	0.01
Q1 p-value	0.60	0.97	0.82	0.71	0.34	0.92	0.83	0.66	0.93	1.00	0.77	0.78
Q2 r	0.03	0.07	0.06	0.12	0.14	0.08	0.10	0.10	0.06	0.07	0.06	0.11
Q2 p-value	0.49	0.11	0.17	0.00	0.00	0.06	0.01	0.01	0.15	0.11	0.17	0.01
Q3 r	0.14	0.14	0.15	0.17	0.14	0.16	0.17	0.16	0.10	0.13	0.14	0.15
Q3 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00
Q4 r	0.05	0.10	0.07	0.14	0.18	0.09	0.11	0.14	0.14	0.10	0.07	0.12
Q4 p-value	0.25	0.02	0.07	0.00	0.00	0.03	0.01	0.00	0.00	0.01	0.11	0.00
Q5 r	-0.07	-0.06	-0.06	-0.05	-0.05	-0.05	-0.04	-0.05	-0.08	-0.07	-0.06	-0.05
Q5 p-value	0.10	0.14	0.14	0.22	0.22	0.21	0.28	0.21	0.06	0.10	0.13	0.22
Q6 r	-0.05	-0.04	-0.04	-0.04	-0.03	-0.04	-0.04	-0.04	-0.05	-0.04	-0.04	-0.04
Q6 p-value	0.26	0.27	0.30	0.31	0.43	0.31	0.37	0.32	0.19	0.27	0.33	0.34
Q7 r	0.12	0.17	0.15	0.19	0.23	0.15	0.18	0.19	0.19	0.18	0.15	0.19
Q7 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q8 r	0.08	0.12	0.11	0.13	0.19	0.10	0.12	0.14	0.13	0.13	0.12	0.14
Q8 p-value	0.04	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00
Q9 r	0.08	0.09	0.09	0.08	0.11	0.07	0.07	0.09	0.10	0.10	0.09	0.08
Q9 p-value	0.04	0.03	0.03	0.05	0.01	0.10	0.10	0.04	0.01	0.02	0.02	0.05
Q10 r	0.10	0.13	0.12	0.11	0.17	0.08	0.08	0.12	0.17	0.15	0.12	0.11
Q10 p-value	0.01	0.00	0.00	0.01	0.00	0.05	0.04	0.00	0.00	0.00	0.00	0.01
Q11 r	0.12	0.15	0.15	0.18	0.20	0.15	0.17	0.17	0.13	0.15	0.15	0.18
Q11 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q12 r	0.04	0.04	0.04	0.04	0.04	0.04	0.03	0.04	0.04	0.04	0.04	0.04
Q12 p-value	0.33	0.32	0.34	0.36	0.33	0.39	0.45	0.38	0.28	0.28	0.28	0.36
Q13 r	0.05	0.09	0.08	0.11	0.15	0.07	0.09	0.11	0.11	0.10	0.08	0.11
Q13 p-value	0.19	0.03	0.05	0.01	0.00	0.07	0.02	0.01	0.01	0.02	0.05	0.01
Q14 r	0.01	0.05	0.03	0.10	0.16	0.04	0.08	0.10	0.12	0.07	0.04	0.08
Q14 p-value	0.87	0.22	0.52	0.02	0.00	0.31	0.06	0.01	0.00	0.09	0.31	0.04
Q15 r	0.16	0.19	0.18	0.18	0.22	0.15	0.15	0.19	0.22	0.20	0.19	0.17
Q15 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q16 r	0.19	0.24	0.24	0.28	0.32	0.23	0.26	0.27	0.23	0.24	0.24	0.28
Q16 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q17 r	-0.05	-0.05	-0.05	-0.04	-0.04	-0.05	-0.04	-0.05	-0.05	-0.05	-0.05	-0.04
Q17 p-value	0.19	0.24	0.22	0.31	0.34	0.25	0.30	0.26	0.23	0.22	0.25	0.32
Q18 r	0.11	0.12	0.11	0.12	0.15	0.09	0.09	0.12	0.17	0.14	0.12	0.10
Q18 p-value	0.01	0.00	0.01	0.00	0.00	0.03	0.04	0.00	0.00	0.00	0.00	0.01
Q19 r	0.02	0.04	0.04	0.06	0.08	0.04	0.05	0.06	0.04	0.05	0.05	0.06
Q19 p-value	0.54	0.30	0.32	0.15	0.05	0.33	0.20	0.18	0.34	0.23	0.24	0.16
Q20 r	0.09	0.12	0.11	0.16	0.20	0.11	0.14	0.16	0.18	0.14	0.11	0.15
Q20 p-value	0.03	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.01	0.00
Q21 r	0.10	0.14	0.13	0.16	0.20	0.12	0.12	0.15	0.20	0.15	0.12	0.15
Q21 p-value	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

TABLE A.21: *Part 21: Pearson's correlation coefficients and p-values for numerical variables*

	Q57	Q58	Q59	Q60	Q61	Q62	Q63	Q64	Q65	Q66	Q67	Q68
	Pearson's Correlation Coefficients N = 594 Prob $> \mid r \mid$ under H0: Rho=0											
Q22 r	-0.02	-0.02	-0.02	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.02	-0.03
Q22 p-value	0.66	0.59	0.58	0.40	0.44	0.51	0.43	0.42	0.53	0.54	0.62	0.47
Q23 r	0.00	0.07	0.04	0.09	0.16	0.05	0.09	0.10	0.12	0.08	0.05	0.10
Q23 p-value	0.94	0.09	0.29	0.03	0.00	0.21	0.03	0.01	0.00	0.06	0.22	0.01
Q24 r	-0.04	-0.02	-0.03	-0.01	0.02	-0.03	-0.02	-0.01	-0.01	-0.01	-0.02	-0.01
Q24 p-value	0.34	0.63	0.49	0.78	0.60	0.51	0.66	0.82	0.88	0.73	0.60	0.78
Q25 r	-0.02	-0.01	-0.01	-0.01	0.02	-0.01	-0.01	-0.00	0.01	-0.00	-0.01	-0.00
Q25 p-value	0.68	0.86	0.74	0.86	0.64	0.73	0.78	0.94	0.88	0.98	0.86	0.98
Q26 r	-0.02	-0.02	-0.02	-0.02	0.00	-0.02	-0.02	-0.02	-0.02	-0.01	-0.02	-0.01
Q26 p-value	0.55	0.65	0.60	0.65	0.98	0.58	0.62	0.71	0.67	0.72	0.67	0.72
Q27 r	0.10	0.12	0.11	0.11	0.14	0.10	0.10	0.12	0.14	0.13	0.12	0.11
Q27 p-value	0.01	0.00	0.01	0.01	0.00	0.02	0.02	0.00	0.00	0.00	0.00	0.01
Q28 r	0.09	0.14	0.12	0.15	0.17	0.12	0.14	0.15	0.14	0.14	0.12	0.14
Q28 p-value	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q29 r	-0.00	0.03	0.02	0.06	0.08	0.02	0.05	0.05	0.04	0.03	0.03	0.05
Q29 p-value	0.97	0.50	0.66	0.17	0.04	0.65	0.26	0.27	0.30	0.46	0.49	0.27
Q30 r	0.05	0.06	0.05	0.06	0.08	0.05	0.05	0.06	0.08	0.06	0.06	0.06
Q30 p-value	0.21	0.15	0.19	0.15	0.05	0.23	0.21	0.13	0.07	0.12	0.18	0.14
Q31 r	0.04	0.06	0.04	0.07	0.08	0.05	0.05	0.07	0.09	0.06	0.05	0.07
Q31 p-value	0.31	0.17	0.32	0.10	0.04	0.19	0.19	0.10	0.03	0.14	0.27	0.11
Q32 r	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.04	0.05	0.05	0.05
Q32 p-value	0.21	0.22	0.19	0.21	0.20	0.22	0.23	0.23	0.35	0.21	0.19	0.22
Q33 r	0.04	0.04	0.04	0.04	0.05	0.03	0.03	0.04	0.06	0.05	0.04	0.04
Q33 p-value	0.39	0.30	0.34	0.34	0.19	0.42	0.43	0.31	0.16	0.24	0.30	0.30
Q34 r	0.16	0.16	0.15	0.12	0.21	0.10	0.08	0.15	0.23	0.20	0.16	0.12
Q34 p-value	0.00	0.00	0.00	0.00	0.00	0.02	0.05	0.00	0.00	0.00	0.00	0.00
Q35 r	0.12	0.12	0.10	0.06	0.17	0.04	0.03	0.10	0.21	0.16	0.12	0.07
Q35 p-value	0.00	0.00	0.01	0.13	0.00	0.28	0.51	0.02	0.00	0.00	0.00	0.08
Q36 r	0.21	0.25	0.24	0.29	0.30	0.26	0.28	0.28	0.23	0.25	0.25	0.29
Q36 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q37 r	-0.01	0.01	-0.00	0.02	0.06	-0.01	0.00	0.02	0.04	0.02	0.01	0.02
Q37 p-value	0.83	0.83	0.96	0.65	0.15	0.90	0.95	0.60	0.33	0.62	0.90	0.69
Q38 r	-0.01	-0.00	-0.01	0.00	0.04	-0.02	-0.01	0.01	0.03	0.01	-0.00	0.00
Q38 p-value	0.74	0.97	0.84	0.93	0.30	0.65	0.84	0.81	0.46	0.79	0.93	0.92
Q39 r	-0.04	-0.03	-0.04	-0.03	-0.02	-0.03	-0.03	-0.03	-0.02	-0.03	-0.03	-0.03
Q39 p-value	0.39	0.44	0.39	0.50	0.63	0.48	0.42	0.52	0.61	0.49	0.40	0.42
Q40 r	0.00	0.00	0.00	0.01	0.03	-0.01	-0.00	0.01	0.02	0.01	0.00	0.00
Q40 p-value	0.99	0.94	0.97	0.83	0.45	0.86	0.98	0.87	0.65	0.83	0.92	0.92
Q41 r	0.12	0.16	0.15	0.18	0.26	0.14	0.16	0.18	0.20	0.19	0.16	0.19
Q41 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q42 r	-0.01	-0.03	-0.03	-0.03	-0.04	-0.02	-0.02	-0.03	-0.04	-0.03	-0.02	-0.03
Q42 p-value	0.74	0.47	0.49	0.44	0.38	0.71	0.66	0.44	0.34	0.47	0.63	0.45
Q43 r	0.25	0.28	0.27	0.30	0.31	0.28	0.29	0.30	0.29	0.28	0.27	0.30
Q43 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q44 r	0.25	0.28	0.28	0.30	0.33	0.27	0.29	0.30	0.27	0.30	0.30	0.31
Q44 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q45 r	0.96	0.99	0.98	0.99	0.96	0.98	0.98	1.00	0.91	0.98	0.98	1.00
Q45 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

TABLE A.22: *Part 22: Pearson's correlation coefficients and p-values for numerical variables*

	Q57	Q58	Q59	Q60	Q61	Q62	Q63	Q64	Q65	Q66	Q67	Q68
Pearson's Correlation Coefficients N = 594 Prob > | r | under H0: Rho=0												
Q46 r	0.90	0.95	0.93	0.95	0.99	0.91	0.93	0.97	0.94	0.96	0.94	0.96
Q46 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q47 r	0.96	0.98	0.98	0.97	0.93	0.97	0.98	0.99	0.89	0.97	0.97	0.98
Q47 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q48 r	0.91	0.96	0.95	0.98	0.96	0.95	0.98	0.99	0.88	0.95	0.95	0.99
Q48 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q49 r	0.96	0.99	0.98	0.98	0.95	0.97	0.98	1.00	0.91	0.98	0.98	0.99
Q49 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q50 r	0.95	0.99	0.98	0.99	0.95	0.98	0.98	0.99	0.90	0.97	0.98	1.00
Q50 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q51 r	0.98	0.99	1.00	0.99	0.92	0.99	0.98	0.98	0.88	0.97	0.99	0.99
Q51 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q52 r	0.98	0.99	0.99	0.98	0.95	0.97	0.97	0.99	0.92	0.99	0.99	0.99
Q52 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q53 r	0.97	0.99	0.99	0.99	0.95	0.98	0.97	0.99	0.92	0.99	0.99	0.99
Q53 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q54 r	0.92	0.97	0.96	0.98	0.98	0.95	0.97	0.99	0.92	0.97	0.96	0.99
Q54 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q55 r	0.97	0.99	0.99	0.98	0.96	0.97	0.97	0.99	0.94	0.99	0.99	0.98
Q55 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q56 r	0.98	0.99	1.00	0.98	0.93	0.98	0.98	0.99	0.91	0.98	0.99	0.99
Q56 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q57 r	1.00	0.98	0.99	0.95	0.88	0.97	0.94	0.96	0.90	0.97	0.98	0.95
Q57 p-value		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q58 r	0.98	1.00	0.99	0.98	0.94	0.98	0.97	0.99	0.92	0.99	0.99	0.98
Q58 p-value	0.00		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q59 r	0.99	0.99	1.00	0.98	0.91	0.99	0.97	0.98	0.89	0.98	0.99	0.98
Q59 p-value	0.00	0.00		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q60 r	0.95	0.98	0.98	1.00	0.93	0.98	0.99	0.99	0.87	0.96	0.98	0.99
Q60 p-value	0.00	0.00	0.00		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q61 r	0.88	0.94	0.91	0.93	1.00	0.89	0.91	0.96	0.95	0.96	0.93	0.95
Q61 p-value	0.00	0.00	0.00	0.00		0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q62 r	0.97	0.98	0.99	0.98	0.89	1.00	0.99	0.97	0.84	0.95	0.98	0.98
Q62 p-value	0.00	0.00	0.00	0.00	0.00		0.00	0.00	0.00	0.00	0.00	0.00
Q63 r	0.94	0.97	0.97	0.99	0.91	0.99	1.00	0.98	0.84	0.94	0.97	0.99
Q63 p-value	0.00	0.00	0.00	0.00	0.00	0.00		0.00	0.00	0.00	0.00	0.00
Q64 r	0.96	0.99	0.98	0.99	0.96	0.97	0.98	1.00	0.92	0.98	0.98	0.99
Q64 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00		0.00	0.00	0.00	0.00
Q65 r	0.90	0.92	0.89	0.87	0.95	0.84	0.84	0.92	1.00	0.96	0.91	0.89
Q65 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00		0.00	0.00	0.00
Q66 r	0.97	0.99	0.98	0.96	0.96	0.95	0.94	0.98	0.96	1.00	0.98	0.97
Q66 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00		0.00	0.00
Q67 r	0.98	0.99	0.99	0.98	0.93	0.98	0.97	0.98	0.91	0.98	1.00	0.98
Q67 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00		0.00
Q68 r	0.95	0.98	0.98	0.99	0.95	0.98	0.99	0.99	0.89	0.97	0.98	1.00
Q68 p-value	0.00	0.00	Stellenbosch	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
Q69 r	0.99	0.98	0.99	0.94	0.87	0.97	0.94	0.95	0.88	0.96	0.97	0.94
Q69 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

TABLE A.23: *Part 23: Pearson's correlation coefficients and p-values for numerical variables*

	Q57	Q58	Q59	Q60	Q61	Q62	Q63	Q64	Q65	Q66	Q67	Q68
Pearson's Correlation Coefficients N = 594 Prob $>$ $\mid r \mid$ under H0: Rho=0												
Q70 r	0.96	0.99	0.99	0.99	0.95	0.98	0.98	1.00	0.91	0.98	0.98	0.99
Q70 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q71 r	0.95	0.98	0.98	0.99	0.95	0.98	0.98	0.99	0.91	0.97	0.97	0.99
Q71 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q72 r	0.92	0.96	0.95	0.97	0.98	0.94	0.96	0.99	0.92	0.96	0.96	0.98
Q72 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q73 r	0.93	0.98	0.97	0.99	0.97	0.96	0.98	0.99	0.91	0.97	0.97	0.99
Q73 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q74 r	0.99	1.00	1.00	0.97	0.92	0.98	0.97	0.98	0.91	0.98	0.99	0.98
Q74 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q75 r	0.98	0.99	0.98	0.95	0.93	0.95	0.94	0.97	0.94	0.99	0.98	0.96
Q75 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q76 r	0.99	0.99	0.99	0.96	0.92	0.97	0.96	0.98	0.93	0.98	0.99	0.97
Q76 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q77 r	0.09	0.08	0.10	0.06	-0.02	0.12	0.11	0.07	0.02	0.06	0.08	0.07
Q77 p-value	0.02	0.04	0.02	0.12	0.56	0.00	0.01	0.11	0.60	0.13	0.06	0.08
Q78 r	0.10	0.09	0.10	0.07	-0.02	0.13	0.12	0.07	0.02	0.07	0.08	0.08
Q78 p-value	0.02	0.03	0.01	0.07	0.58	0.00	0.00	0.08	0.61	0.11	0.04	0.06

TABLE A.24: *Part 24: Pearson's correlation coefficients and p-values for numerical variables*

Pearson's Correlation Coefficients N = 594 Prob $> \mid r \mid$ under H0: Rho=0										
	Q69	Q70	Q71	Q72	Q73	Q74	Q75	Q76	Q77	Q78
Volume r	0.17	0.11	0.11	0.08	0.10	0.15	0.13	0.14	0.57	0.59
Volume p-value	0.00	0.01	0.01	0.04	0.02	0.00	0.00	0.00	0.00	0.00
Value r	0.10	0.03	0.02	0.00	0.01	0.07	0.07	0.07	0.50	0.50
Value p-value	0.02	0.52	0.64	0.91	0.79	0.08	0.07	0.10	0.00	0.00
AUP r	0.10	0.05	0.02	0.04	0.02	0.07	0.13	0.09	-0.04	-0.06
AUP p-value	0.02	0.27	0.71	0.30	0.56	0.09	0.00	0.03	0.30	0.18
Q1 r	-0.02	0.02	0.02	0.02	0.02	-0.01	-0.01	-0.01	-0.16	-0.15
Q1 p-value	0.63	0.66	0.61	0.58	0.61	0.79	0.74	0.72	0.00	0.00
Q2 r	0.02	0.10	0.12	0.13	0.12	0.05	0.05	0.05	-0.12	-0.11
Q2 p-value	0.58	0.02	0.00	0.00	0.00	0.20	0.20	0.24	0.00	0.01
Q3 r	0.13	0.16	0.16	0.15	0.15	0.14	0.12	0.13	-0.03	-0.02
Q3 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.51	0.65
Q4 r	0.06	0.13	0.16	0.14	0.15	0.08	0.08	0.09	-0.08	-0.08
Q4 p-value	0.18	0.00	0.00	0.00	0.00	0.07	0.06	0.03	0.06	0.04
Q5 r	-0.07	-0.05	-0.05	-0.05	-0.05	-0.06	-0.07	-0.07	-0.11	-0.09
Q5 p-value	0.11	0.21	0.22	0.23	0.23	0.13	0.07	0.10	0.01	0.03
Q6 r	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.05	-0.05	-0.09	-0.08
Q6 p-value	0.29	0.35	0.29	0.34	0.35	0.28	0.26	0.23	0.03	0.04
Q7 r	0.12	0.19	0.21	0.22	0.21	0.15	0.15	0.14	-0.12	-0.10
Q7 p-value	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01
Q8 r	0.10	0.14	0.13	0.17	0.15	0.11	0.13	0.09	-0.07	-0.07
Q8 p-value	0.02	0.00	0.00	0.00	0.00	0.01	0.00	0.02	0.08	0.07
Q9 r	0.09	0.09	0.08	0.09	0.08	0.09	0.10	0.08	-0.06	-0.06
Q9 p-value	0.03	0.04	0.06	0.03	0.04	0.03	0.02	0.04	0.15	0.15
Q10 r	0.11	0.13	0.12	0.14	0.12	0.11	0.14	0.11	-0.09	-0.09
Q10 p-value	0.01	0.00	0.00	0.00	0.00	0.01	0.00	0.01	0.03	0.03
Q11 r	0.12	0.17	0.17	0.19	0.19	0.14	0.14	0.13	-0.05	-0.05
Q11 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.25	0.23
Q12 r	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	-0.03	-0.03
Q12 p-value	0.33	0.39	0.37	0.38	0.38	0.31	0.29	0.34	0.50	0.47
Q13 r	0.05	0.11	0.12	0.13	0.12	0.08	0.08	0.07	-0.14	-0.13
Q13 p-value	0.18	0.01	0.00	0.00	0.00	0.06	0.04	0.07	0.00	0.00
Q14 r	-0.01	0.09	0.13	0.10	0.11	0.03	0.02	0.05	-0.15	-0.15
Q14 p-value	0.88	0.03	0.00	0.02	0.01	0.40	0.66	0.20	0.00	0.00
Q15 r	0.17	0.19	0.17	0.20	0.18	0.17	0.20	0.18	-0.04	-0.03
Q15 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.36	0.44
Q16 r	0.19	0.27	0.27	0.28	0.29	0.23	0.22	0.21	-0.11	-0.10
Q16 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.02
Q17 r	-0.05	-0.04	-0.04	-0.04	-0.04	-0.05	-0.05	-0.05	-0.11	-0.11
Q17 p-value	0.19	0.31	0.32	0.31	0.33	0.23	0.23	0.19	0.01	0.01
Q18 r	0.11	0.12	0.12	0.13	0.11	0.11	0.13	0.12	-0.06	-0.05
Q18 p-value	0.01	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.14	0.23
Q19 r	0.02	0.05	0.06	0.07	0.07	0.04	0.04	0.04	-0.04	-0.04
Q19 p-value	0.56	0.18	0.17	0.08	0.11	0.32	0.32	0.38	0.32	0.29
Q20 r	0.08	0.16	0.19	0.17	0.17	0.11	0.12	0.13	-0.06	-0.05
Q20 p-value	0.05	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.14	0.25
Q21 r	0.10	0.16	0.18	0.17	0.17	0.14	0.15	0.13	-0.12	-0.12
Q21 p-value	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

TABLE A.25: *Part 25: Pearson's correlation coefficients and p-values for numerical variables*

	Q69	Q70	Q71	Q72	Q73	Q74	Q75	Q76	Q77	Q78
Pearson's Correlation Coefficients N = 594 Prob > | r | under H0: Rho=0
Q22 r	-0.02	-0.03	-0.04	-0.03	-0.03	-0.02	-0.02	-0.03	-0.07	-0.05
Q22 p-value	0.70	0.46	0.36	0.40	0.41	0.63	0.69	0.53	0.11	0.18
Q23 r	0.01	0.10	0.12	0.14	0.13	0.04	0.05	0.04	-0.08	-0.08
Q23 p-value	0.73	0.01	0.00	0.00	0.00	0.37	0.23	0.35	0.04	0.05
Q24 r	-0.03	-0.01	-0.00	-0.00	-0.00	-0.03	-0.02	-0.03	-0.13	-0.12
Q24 p-value	0.43	0.84	0.91	0.95	0.97	0.49	0.59	0.47	0.00	0.00
Q25 r	-0.01	-0.00	-0.00	0.00	0.00	-0.01	-0.01	-0.01	-0.11	-0.09
Q25 p-value	0.73	0.95	0.98	0.94	0.99	0.79	0.87	0.73	0.01	0.03
Q26 r	-0.02	-0.02	-0.02	-0.01	-0.01	-0.02	-0.01	-0.02	-0.08	-0.08
Q26 p-value	0.61	0.69	0.70	0.79	0.80	0.63	0.78	0.58	0.06	0.05
Q27 r	0.12	0.12	0.11	0.13	0.13	0.11	0.14	0.11	-0.02	-0.02
Q27 p-value	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.01	0.71	0.64
Q28 r	0.11	0.15	0.16	0.15	0.16	0.11	0.11	0.11	-0.09	-0.09
Q28 p-value	0.01	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.03	0.04
Q29 r	-0.01	0.05	0.06	0.06	0.06	0.02	0.01	0.02	-0.08	-0.07
Q29 p-value	0.81	0.25	0.17	0.13	0.17	0.62	0.77	0.56	0.05	0.07
Q30 r	0.05	0.06	0.06	0.07	0.06	0.05	0.06	0.06	-0.07	-0.07
Q30 p-value	0.21	0.14	0.11	0.09	0.13	0.19	0.13	0.18	0.11	0.10
Q31 r	0.04	0.06	0.09	0.08	0.07	0.04	0.05	0.05	-0.10	-0.11
Q31 p-value	0.31	0.12	0.04	0.05	0.11	0.32	0.22	0.25	0.01	0.01
Q32 r	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	-0.04	-0.04
Q32 p-value	0.19	0.22	0.23	0.22	0.24	0.20	0.22	0.25	0.33	0.35
Q33 r	0.04	0.04	0.04	0.05	0.04	0.04	0.05	0.04	-0.04	-0.04
Q33 p-value	0.37	0.31	0.28	0.24	0.30	0.34	0.24	0.33	0.35	0.33
Q34 r	0.16	0.14	0.13	0.15	0.13	0.15	0.19	0.16	-0.11	-0.11
Q34 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01
Q35 r	0.12	0.10	0.07	0.10	0.09	0.11	0.15	0.12	-0.12	-0.12
Q35 p-value	0.00	0.02	0.07	0.01	0.04	0.01	0.00	0.00	0.00	0.00
Q36 r	0.21	0.28	0.30	0.32	0.30	0.24	0.24	0.23	-0.02	-0.01
Q36 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.70	0.83
Q37 r	-0.01	0.02	0.03	0.03	0.03	0.00	0.01	0.00	-0.17	-0.16
Q37 p-value	0.83	0.65	0.46	0.47	0.54	0.97	0.86	0.96	0.00	0.00
Q38 r	-0.02	0.01	0.01	0.02	0.01	-0.01	0.00	-0.00	-0.12	-0.11
Q38 p-value	0.67	0.88	0.77	0.59	0.77	0.85	0.96	0.92	0.00	0.01
Q39 r	-0.03	-0.03	-0.02	-0.03	-0.03	-0.04	-0.04	-0.04	-0.08	-0.08
Q39 p-value	0.45	0.49	0.66	0.41	0.46	0.39	0.37	0.39	0.06	0.05
Q40 r	-0.00	0.01	0.01	0.02	0.01	0.00	0.01	-0.00	-0.07	-0.06
Q40 p-value	0.91	0.83	0.81	0.69	0.74	0.96	0.89	1.00	0.11	0.13
Q41 r	0.12	0.18	0.19	0.21	0.20	0.15	0.17	0.15	-0.15	-0.15
Q41 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q42 r	-0.02	-0.02	-0.04	-0.04	-0.03	-0.03	-0.04	-0.03	0.02	0.02
Q42 p-value	0.67	0.63	0.35	0.32	0.48	0.52	0.39	0.49	0.62	0.62
Q43 r	0.24	0.30	0.31	0.31	0.31	0.27	0.27	0.27	0.11	0.11
Q43 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01
Q44 r	0.24	0.31	0.30	0.34	0.32	0.27	0.29	0.28	0.04	0.04
Q44 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.34	0.34
Q45 r	0.95	1.00	0.99	0.99	1.00	0.98	0.97	0.98	0.08	0.09
Q45 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.03

TABLE A.26: *Part 26: Pearson's correlation coefficients and p-values for numerical variables*

	Q69	Q70	Q71	Q72	Q73	Q74	Q75	Q76	Q77	Q78
Q46 r	0.89	0.97	0.97	0.98	0.98	0.93	0.94	0.94	0.01	0.02
Q46 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.72	0.66
Q47 r	0.95	0.98	0.97	0.97	0.98	0.98	0.97	0.98	0.10	0.11
Q47 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01
Q48 r	0.90	0.98	0.98	0.99	0.99	0.95	0.93	0.94	0.06	0.07
Q48 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.14	0.11
Q49 r	0.96	0.99	0.99	0.98	0.99	0.98	0.97	0.98	0.07	0.08
Q49 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.08	0.06
Q50 r	0.95	0.99	0.99	0.98	0.99	0.98	0.97	0.97	0.07	0.08
Q50 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.07	0.05
Q51 r	0.98	0.99	0.98	0.95	0.97	0.99	0.97	0.98	0.08	0.09
Q51 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.03
Q52 r	0.97	0.99	0.98	0.97	0.98	0.99	0.98	0.99	0.06	0.06
Q52 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.14	0.12
Q53 r	0.97	0.99	0.99	0.97	0.98	0.99	0.98	0.98	0.07	0.08
Q53 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.10	0.07
Q54 r	0.91	0.99	0.99	0.99	1.00	0.96	0.95	0.95	0.04	0.04
Q54 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.37	0.31
Q55 r	0.96	0.99	0.98	0.98	0.98	0.99	0.98	0.99	0.05	0.06
Q55 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.20	0.16
Q56 r	0.98	0.99	0.98	0.96	0.98	1.00	0.98	0.99	0.10	0.11
Q56 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01
Q57 r	0.99	0.96	0.95	0.92	0.93	0.99	0.98	0.99	0.09	0.10
Q57 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.02
Q58 r	0.98	0.99	0.98	0.96	0.98	1.00	0.99	0.99	0.08	0.09
Q58 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.03
Q59 r	0.99	0.99	0.98	0.95	0.97	1.00	0.98	0.99	0.10	0.10
Q59 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.01
Q60 r	0.94	0.99	0.99	0.97	0.99	0.97	0.95	0.96	0.06	0.07
Q60 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.12	0.07
Q61 r	0.87	0.95	0.95	0.98	0.97	0.92	0.93	0.92	-0.02	-0.02
Q61 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.56	0.58
Q62 r	0.97	0.98	0.98	0.94	0.96	0.98	0.95	0.97	0.12	0.13
Q62 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q63 r	0.94	0.98	0.98	0.96	0.98	0.97	0.94	0.96	0.11	0.12
Q63 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00
Q64 r	0.95	1.00	0.99	0.99	0.99	0.98	0.97	0.98	0.07	0.07
Q64 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.11	0.08
Q65 r	0.88	0.91	0.91	0.92	0.91	0.91	0.94	0.93	0.02	0.02
Q65 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.60	0.61
Q66 r	0.96	0.98	0.97	0.96	0.97	0.98	0.99	0.98	0.06	0.07
Q66 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.13	0.11
Q67 r	0.97	0.98	0.97	0.96	0.97	0.99	0.98	0.99	0.08	0.08
Q67 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.06	0.04
Q68 r	0.94	0.99	0.99	0.98	0.99	0.98	0.96	0.97	0.07	0.08
Q68 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.08	0.06
Q69 r	1.00	0.96	0.94	0.90	0.92	0.99	0.98	0.98	0.10	0.11
Q69 p-value		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.01

Pearson's Correlation Coefficients N = 594 Prob $> |r|$ under H0: Rho=0

TABLE A.27: *Part 27: Pearson's correlation coefficients and p-values for numerical variables*

	Q69	Q70	Q71	Q72	Q73	Q74	Q75	Q76	Q77	Q78
Q70 r	0.96	1.00	0.99	0.98	0.99	0.98	0.97	0.98	0.05	0.06
Q70 p-value	0.00		0.00	0.00	0.00	0.00	0.00	0.00	0.22	0.16
Q71 r	0.94	0.99	1.00	0.98	0.99	0.98	0.96	0.97	0.06	0.07
Q71 p-value	0.00	0.00		0.00	0.00	0.00	0.00	0.00	0.14	0.09
Q72 r	0.90	0.98	0.98	1.00	0.99	0.95	0.95	0.96	0.07	0.08
Q72 p-value	0.00	0.00	0.00		0.00	0.00	0.00	0.00	0.09	0.07
Q73 r	0.92	0.99	0.99	0.99	1.00	0.97	0.96	0.96	0.06	0.07
Q73 p-value	0.00	0.00	0.00	0.00		0.00	0.00	0.00	0.15	0.11
Q74 r	0.99	0.98	0.98	0.95	0.97	1.00	0.99	0.99	0.11	0.11
Q74 p-value	0.00	0.00	0.00	0.00	0.00		0.00	0.00	0.01	0.01
Q75 r	0.98	0.97	0.96	0.95	0.96	0.99	1.00	0.99	0.10	0.11
Q75 p-value	0.00	0.00	0.00	0.00	0.00	0.00		0.00	0.01	0.01
Q76 r	0.98	0.98	0.97	0.96	0.96	0.99	0.99	1.00	0.11	0.11
Q76 p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00		0.01	0.01
Q77 r	0.10	0.05	0.06	0.07	0.06	0.11	0.10	0.11	1.00	0.99
Q77 p-value	0.02	0.22	0.14	0.09	0.15	0.01	0.01	0.01		0.00
Q78 r	0.11	0.06	0.07	0.08	0.07	0.11	0.11	0.11	0.99	1.00
Q78 p-value	0.01	0.16	0.09	0.07	0.11	0.01	0.01	0.01	0.00	

Pearson's Correlation Coefficients N = 594 Prob $> |r|$ under H0: Rho=0

TABLE A.28: *Part 28: Pearson's correlation coefficients and p-values for numerical variables*

A.3 Method for determining a competitor set using the product life cycle methodology: Cluster analysis procedure

A.3.1 Ward's minimum variance method

The PROC CLUSTER procedure in SAS studio® software using Ward's minimum distance method produces a dendrogram. A dendrogram is a graphical representation of the cluster structure formed. SAS studio® software is limited to a maximum of 200 products to be evaluated and displayed. In order to overcome this limit as the dataset has 594 products, *jmp* Statistical Discovery™ software from SAS is used to produce Figure A.9 .
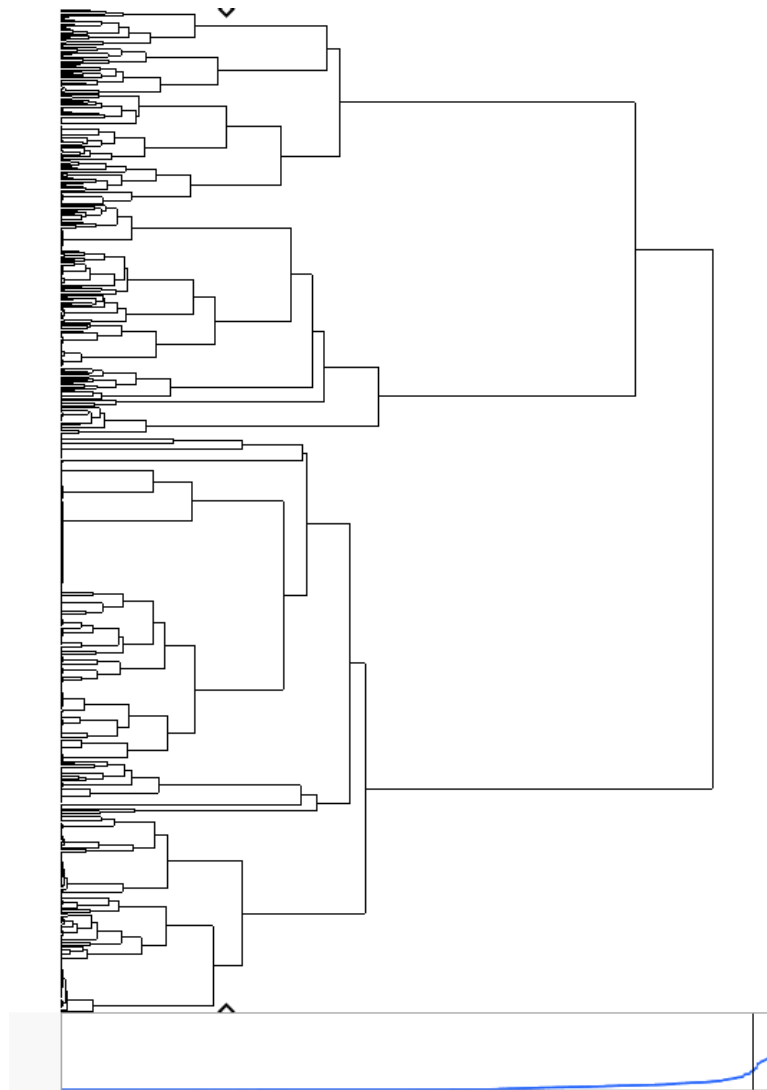
FIGURE A.9: *Dendrogram displaying the cluster structure of the products using the descriptive variables pack sizes, alcohol percentage, carbonation level, sugar content, taste profile and factor scores*

A.3.2   Interpreting results

Cluster	Sub-cluster	Product	Cluster	Sub-cluster	Product
1		RTD-FABS-P27-330ML	6	3	WINE-BIB-P82-5LTR
1		RTD-FABS-P30-330ML	6	3	WINE-ROSE-P12-750ML
2		RTD-BEER-P51-660ML	6	3	WINE-BIB-P41-3LTR
2		RTD-FABS-P38-660ML	6	3	WINE-BIB-P78-5LTR
2		RTD-BEER-P44-500ML	6	3	WINE-RED-P72-750ML
2		RTD-BEER-P58-750ML	6	3	WINE-RED-P85-750ML
2		RTD-BEER-P43-500ML	6	3	WINE-ROSE-P13-750ML
2		RTD-BEER-P55-750ML	6	3	WINE-WHITE-P47-750ML
2		RTD-BEER-P45-500ML	6	3	WINE-BIB-P81-5LTR
2		RTD-BEER-P50-660ML	6	3	WINE-BIB-P80-5LTR
2		RTD-BEER-P59-750ML	6	3	WINE-RED-P95-750ML
2		RTD-BEER-P46-500ML	6	3	WINE-PERLE-P1-1LTR
2		RTD-BEER-P57-750ML	6	3	WINE-BIB-P63-5LTR
2		RTD-BEER-P48-5LTR	6	3	WINE-BIB-P69-5LTR
2		RTD-BEER-P49-650ML	6	3	WINE-BIB-P74-5LTR
2		RTD-CIDER-P18-660ML	6	3	WINE-PERLE-P3-1LTR
2		RTD-CIDER-P17-660ML	6	3	WINE-PERLE-P4-2LTR
2		RTD-BEER-P47-500ML	6	3	WINE-PERLE-P13-750ML
2		RTD-BEER-P56-750ML	6	3	WINE-ROSE-P1-1.5LTR
2		RTD-FABS-P37-660ML	6	3	WINE-BIB-P44-3LTR
2		RTD-CIDER-P16-500ML	6	3	WINE-ROSE-P5-750ML
2		RTD-BEER-P52-660ML	6	3	WINE-FORTIFIED-P2-1LTR
2		RTD-BEER-P53-660ML	6	3	WINE-FORTIFIED-P5-2LTR
3		SPIRITS-LIQUEURS-P2-1LTR	6	3	WINE-FORTIFIED-P16-750ML
3		SPIRITS-LIQUEURS-P16-750ML	6	3	WINE-RED-P3-1.5LTR
3		SPIRITS-LIQUEURS-P1-1LTR	6	3	WINE-BIB-P49-3LTR
3		SPIRITS-LIQUEURS-P9-750ML	6	3	WINE-RED-P93-750ML
3		SPIRITS-LIQUEURS-P12-750ML	6	3	WINE-BIB-P50-3LTR
3		SPIRITS-LIQUEURS-P14-750ML	6	3	WINE-ROSE-P19-750ML
3		SPIRITS-LIQUEURS-P11-750ML	6	3	WINE-WHITE-P60-750ML
3		SPIRITS-LIQUEURS-P10-750ML	6	3	WINE-ROSE-P18-750ML
3		SPIRITS-LIQUEURS-P13-750ML	6	3	WINE-RED-P4-1.5LTR
3		SPIRITS-LIQUEURS-P8-750ML	6	3	WINE-WHITE-P2-1.5LTR
3		SPIRITS-LIQUEURS-P7-750ML	6	3	WINE-WHITE-P2-5LTR
3		SPIRITS-LIQUEURS-P6-750ML	6	4	SPIRITS-RUM-P5-750ML
3		SPIRITS-LIQUEURS-P3-750ML	6	4	SPIRITS-WHISKY-P27-750ML
3		SPIRITS-LIQUEURS-P4-750ML	6	4	SPIRITS-WHISKY-P16-750ML
3		SPIRITS-LIQUEURS-P5-750ML	6	4	SPIRITS-WHISKY-P18-750ML
3		SPIRITS-LIQUEURS-P15-750ML	6	4	SPIRITS-WHISKY-P29-750ML
4		SPIRITS-GIN-P2-200ML	6	4	SPIRITS-BRANDY-P19-750ML
4		SPIRITS-GIN-P3-375ML	6	4	SPIRITS-BRANDY-P18-750ML
4		SPIRITS-BRANDY-P5-375ML	6	4	SPIRITS-BRANDY-P21-750ML
4		SPIRITS-VODKA-P3-200ML	6	4	SPIRITS-BRANDY-P20-750ML
5		RTD-FABS-P9-275ML	6	4	SPIRITS-WHISKY-P30-750ML
5		RTD-FABS-P8-275ML	6	4	SPIRITS-WHISKY-P28-750ML

TABLE A.29: *Part1: A description of each product and cluster membership for the original 10 clusters, as well as the sub-set of Cluster 6 when using k-means method of clustering*

Cluster	Sub Cluster	Product	Cluster	Sub Cluster	Product
5		RTD-FABS-P13-275ML	6	4	SPIRITS-WHISKY-P31-750ML
5		RTD-FABS-P31-440ML	6	4	SPIRITS-BRANDY-P17-750ML
5		RTD-FABS-P12-275ML	6	4	SPIRITS-WHISKY-P22-750ML
5		RTD-FABS-P10-275ML	6	4	SPIRITS-WHISKY-P21-750ML
5		WINE-BIB-P84-5LTR	6	4	SPIRITS-WHISKY-P50-750ML
5		WINE-WHITE-P68-750ML	6	4	SPIRITS-BRANDY-P7-750ML
5		WINE-ROSE-P4-1.5LTR	6	4	SPIRITS-BRANDY-P22-750ML
5		SPIRITS-WHISKY-P11-750ML	6	4	SPIRITS-BRANDY-P6-750ML
5		RTD-FABS-P14-275ML	6	4	SPIRITS-BRANDY-P16-750ML
5		RTD-FABS-P28-330ML	6	4	SPIRITS-RUM-P6-750ML
5		RTD-FABS-P32-440ML	6	4	SPIRITS-BRANDY-P9-750ML
5		RTD-FABS-P17-275ML	6	4	SPIRITS-RUM-P3-750ML
5		RTD-FABS-P16-275ML	6	4	SPIRITS-RUM-P4-750ML
5		RTD-FABS-P19-275ML	6	4	SPIRITS-BRANDY-P15-750ML
5		RTD-FABS-P18-275ML	6	4	SPIRITS-BRANDY-P23-750ML
5		RTD-FABS-P2-250ML	6	4	SPIRITS-WHISKY-P32-750ML
5		RTD-FABS-P4-250ML	6	4	SPIRITS-WHISKY-P38-750ML
5		RTD-FABS-P3-250ML	6	4	SPIRITS-BRANDY-P11-750ML
5		RTD-FABS-P1-250ML	6	4	SPIRITS-BRANDY-P13-750ML
5		RTD-FABS-P24-300ML	6	4	SPIRITS-BRANDY-P10-750ML
5		RTD-FABS-P29-330ML	6	5	SPIRITS-WHISKY-P1-1LTR
5		RTD-FABS-P34-440ML	6	5	SPIRITS-WHISKY-P4-1LTR
5		RTD-FABS-P23-300ML	6	5	SPIRITS-WHISKY-P8-1LTR
5		RTD-FABS-P33-440ML	6	5	SPIRITS-WHISKY-P10-1LTR
5		WINE-BIB-P51-3LTR	6	5	SPIRITS-BRANDY-P1-1LTR
5		WINE-RED-P105-750ML	6	5	SPIRITS-BRANDY-P2-1LTR
5		WINE-BIB-P52-3LTR	6	5	SPIRITS-BRANDY-P4-1LTR
5		WINE-WHITE-P66-750ML	6	5	SPIRITS-BRANDY-P3-1LTR
6	1	SPIRITS-WHISKY-P5-1LTR	6	6	SPIRITS-WHISKY-P41-750ML
6	1	SPIRITS-WHISKY-P2-1LTR	6	6	SPIRITS-WHISKY-P15-750ML
6	1	SPIRITS-WHISKY-P3-1LTR	6	6	SPIRITS-WHISKY-P14-750ML
6	1	SPIRITS-WHISKY-P9-1LTR	6	6	SPIRITS-WHISKY-P37-750ML
6	1	SPIRITS-WHISKY-P7-1LTR	6	6	SPIRITS-WHISKY-P43-750ML
6	1	SPIRITS-WHISKY-P6-1LTR	6	6	SPIRITS-WHISKY-P20-750ML
6	2	WINE-ROSE-P17-750ML	6	6	SPIRITS-WHISKY-P36-750ML
6	2	WINE-WHITE-P58-750ML	6	6	SPIRITS-WHISKY-P24-750ML
6	2	WINE-WHITE-P59-750ML	6	6	SPIRITS-WHISKY-P44-750ML
6	2	WINE-RED-P90-750ML	6	6	SPIRITS-WHISKY-P45-750ML
6	2	WINE-RED-P88-750ML	6	6	SPIRITS-WHISKY-P42-750ML
6	2	WINE-ROSE-P16-750ML	6	6	SPIRITS-WHISKY-P52-750ML
6	2	WINE-RED-P92-750ML	6	6	SPIRITS-WHISKY-P49-750ML
6	2	WINE-WHITE-P57-750ML	6	6	SPIRITS-WHISKY-P51-750ML
6	2	WINE-WHITE-P52-750ML	6	6	SPIRITS-WHISKY-P25-750ML
6	2	WINE-RED-P91-750ML	6	6	SPIRITS-WHISKY-P34-750ML
6	2	WINE-RED-P5-750ML	6	6	SPIRITS-WHISKY-P46-750ML
6	2	WINE-WHITE-P4-750ML	6	6	SPIRITS-WHISKY-P33-750ML
6	2	WINE-RED-P7-750ML	6	6	SPIRITS-WHISKY-P40-750ML

TABLE A.30: *Part2: A description of each product and cluster membership for the original 10 clusters, as well as the sub-set of Cluster 6 when using k-means method of clustering*

Cluster	Sub Cluster	Product	Cluster	Sub Cluster	Product
6	2	WINE-RED-P6-750ML	6	6	SPIRITS-WHISKY-P12-750ML
6	2	WINE-WHITE-P3-750ML	6	6	SPIRITS-WHISKY-P19-750ML
6	3	WINE-RED-P79-1.5LTR	6	6	SPIRITS-WHISKY-P35-750ML
6	3	WINE-BIB-P42-3LTR	6	6	SPIRITS-WHISKY-P47-750ML
6	3	WINE-BIB-P79-5LTR	6	6	SPIRITS-WHISKY-P17-750ML
6	3	WINE-RED-P71-750ML	6	6	SPIRITS-WHISKY-P39-750ML
6	3	WINE-BIB-P82-1.5LTR	6	6	SPIRITS-WHISKY-P48-750ML
6	3	WINE-BIB-P43-3LTR	6	7	SPIRITS-CANE-P1-1LTR
6	7	SPIRITS-RUM-P1-1LTR	6	9	WINE-BIB-P72-5LTR
6	7	SPIRITS-GIN-P1-1LTR	6	9	WINE-BIB-P65-5LTR
6	7	SPIRITS-VODKA-P1-1LTR	6	9	WINE-BIB-P64-5LTR
6	7	SPIRITS-VODKA-P2-1LTR	6	9	WINE-RED-P66-750ML
6	8	WINE-RED-P19-750ML	6	9	WINE-RED-P64-750ML
6	8	WINE-RED-P22-750ML	6	9	WINE-RED-P57-750ML
6	8	WINE-PERLE-P8-1.5LTR	6	9	WINE-WHITE-P37-750ML
6	8	WINE-PERLE-P2-1LTR	6	9	SPIRITS-GIN-P7-750ML
6	8	WINE-PERLE-P10-750ML	6	9	WINE-WHITE-P24-750ML
6	8	WINE-PERLE-P7-1.5LTR	6	9	SPIRITS-CANE-P2-750ML
6	8	WINE-PERLE-P11-750ML	6	9	SPIRITS-RUM-P2-750ML
6	8	WINE-PERLE-P9-1.5LTR	6	9	SPIRITS-RUM-P7-750ML
6	8	WINE-PERLE-P12-750ML	6	9	WINE-RED-P40-750ML
6	8	WINE-FORTIFIED-P1-1LTR	6	9	SPIRITS-VODKA-P17-750ML
6	8	WINE-FORTIFIED-P8-750ML	6	9	WINE-RED-P62-750ML
6	8	WINE-FORTIFIED-P10-750ML	6	9	SPIRITS-VODKA-P18-750ML
6	8	WINE-BIB-P83-5LTR	6	9	SPIRITS-VODKA-P20-750ML
6	8	WINE-BIB-P84-1.5LTR	6	9	WINE-RED-P65-750ML
6	8	WINE-ROSE-P4-1.5LTR	6	9	WINE-WHITE-P23-750ML
6	8	WINE-WHITE-P42-750ML	6	9	WINE-RED-P50-750ML
6	8	WINE-BIB-P47-3LTR	6	9	WINE-RED-P31-750ML
6	8	WINE-RED-P84-750ML	6	9	WINE-WHITE-P12-750ML
6	8	WINE-BIB-P46-3LTR	6	9	WINE-RED-P26-750ML
6	8	WINE-WHITE-P55-750ML	6	9	WINE-RED-P79-750ML
6	8	WINE-BIB-P45-3LTR	6	9	WINE-WHITE-P1-5LTR
6	8	WINE-RED-P87-750ML	6	9	WINE-WHITE-P56-750ML
6	8	WINE-BIB-P48-3LTR	6	9	WINE-RED-P89-750ML
6	8	WINE-WHITE-P54-750ML	6	9	WINE-ROSE-P2-1.5LTR
6	8	WINE-WHITE-P43-750ML	6	9	WINE-ROSE-P15-750ML
6	8	WINE-ROSE-P14-750ML	6	9	WINE-WHITE-P50-750ML
6	8	WINE-RED-P80-750ML	6	9	WINE-RED-P86-750ML
6	8	WINE-RED-P83-750ML	6	9	WINE-WHITE-P49-750ML
6	8	WINE-WHITE-P51-750ML	6	9	WINE-WHITE-P67-750ML
6	8	WINE-WHITE-P26-750ML	6	9	WINE-RED-P101-750ML
6	8	WINE-RED-P47-750ML	6	9	WINE-ROSE-P20-750ML
6	8	WINE-WHITE-P27-750ML	6	9	WINE-RED-P56-750ML
6	8	WINE-SPARKLING-P16-750ML	6	9	WINE-RED-P27-750ML
6	8	WINE-SPARKLING-P19-750ML	6	9	WINE-WHITE-P15-750ML
6	8	WINE-RED-P44-750ML	6	9	WINE-RED-P30-750ML
6	8	WINE-RED-P43-750ML	6	9	WINE-RED-P36-750ML

TABLE A.31: *Part3: A description of each product and cluster membership for the original 10 clusters, as well as the sub-set of Cluster 6 when using k-means method of clustering*

Cluster	Sub Cluster	Product	Cluster	Sub Cluster	Product
6	8	WINE-FORTIFIED-P11-750ML	6	9	WINE-WHITE-P19-750ML
6	8	WINE-FORTIFIED-P9-750ML	6	9	WINE-RED-P34-750ML
6	8	WINE-FORTIFIED-P4-2LTR	6	9	WINE-RED-P55-750ML
6	8	WINE-FORTIFIED-P14-750ML	6	9	WINE-WHITE-P40-750ML
6	8	WINE-BIB-P3-1LTR	6	9	WINE-WHITE-P35-750ML
6	8	WINE-BIB-P26-3LTR	6	9	SPIRITS-GIN-P9-750ML
6	8	WINE-BIB-P61-5LTR	6	9	SPIRITS-GIN-P4-750ML
6	8	WINE-RED-P8-750ML	6	9	WINE-ROSE-P9-750ML
6	8	WINE-BIB-P5-1LTR	6	9	WINE-WHITE-P25-750ML
6	8	WINE-BIB-P23-3LTR	6	9	WINE-WHITE-P10-750ML
6	8	WINE-BIB-P60-5LTR	6	9	SPIRITS-GIN-P8-750ML
6	8	WINE-BIB-P27-3LTR	6	9	WINE-WHITE-P32-750ML
6	8	WINE-BIB-P62-5LTR	6	9	WINE-ROSE-P11-750ML
6	8	WINE-BIB-P2-1LTR	6	9	WINE-RED-P58-750ML
6	8	WINE-BIB-P22-3LTR	6	9	WINE-WHITE-P34-750ML
6	8	WINE-BIB-P55-5LTR	6	9	SPIRITS-GIN-P6-750ML
6	8	WINE-BIB-P1-1LTR	6	9	WINE-WHITE-P13-750ML
6	8	WINE-BIB-P24-3LTR	6	9	WINE-RED-P35-750ML
6	8	WINE-BIB-P59-5LTR	6	9	WINE-ROSE-P8-750ML
6	8	WINE-BIB-P57-5LTR	6	9	WINE-RED-P20-750ML
6	8	WINE-RED-P11-750ML	6	9	WINE-WHITE-P38-750ML
6	8	WINE-BIB-P56-5LTR	6	9	WINE-RED-P54-750ML
6	8	WINE-BIB-P4-1LTR	6	9	WINE-RED-P76-750ML
6	8	WINE-BIB-P28-3LTR	6	9	SPIRITS-VODKA-P21-750ML
6	8	WINE-BIB-P73-5LTR	6	9	WINE-RED-P12-750ML
6	8	WINE-RED-P10-750ML	6	9	WINE-BIB-P67-5LTR
6	8	WINE-BIB-P25-3LTR	6	9	WINE-BIB-P71-5LTR
6	8	WINE-BIB-P58-5LTR	6	9	WINE-BIB-P70-5LTR
6	8	WINE-RED-P9-750ML	6	9	WINE-RED-P49-750ML
6	8	WINE-BIB-P54-5LTR	6	9	WINE-RED-P73-750ML
6	8	WINE-FORTIFIED-P7-750ML	6	9	WINE-WHITE-P53-750ML
6	8	WINE-RED-P45-750ML	6	9	WINE-WHITE-P46-750ML
6	8	WINE-FORTIFIED-P15-750ML	6	9	WINE-RED-P74-750ML
6	8	WINE-RED-P37-750ML	6	9	WINE-WHITE-P48-750ML
6	8	WINE-WHITE-P21-750ML	6	9	WINE-RED-P82-750ML
6	8	WINE-WHITE-P30-750ML	6	9	WINE-WHITE-P41-750ML
6	8	WINE-WHITE-P29-750ML	6	9	WINE-RED-P52-750ML
6	8	WINE-RED-P32-750ML	6	9	WINE-RED-P69-750ML
6	8	WINE-RED-P46-750ML	6	9	SPIRITS-VODKA-P22-750ML
6	8	WINE-WHITE-P18-750ML	6	9	WINE-RED-P16-750ML
6	8	WINE-RED-P41-750ML	6	9	SPIRITS-VODKA-P9-750ML
6	8	WINE-FORTIFIED-P13-750ML	6	9	SPIRITS-VODKA-P11-750ML
6	8	WINE-RED-P24-750ML	6	9	SPIRITS-VODKA-P7-750ML
6	8	WINE-WHITE-P11-750ML	6	9	SPIRITS-VODKA-P10-750ML
6	8	WINE-ROSE-P7-750ML	6	9	SPIRITS-VODKA-P8-750ML
6	8	WINE-FORTIFIED-P12-750ML	6	9	WINE-RED-P68-750ML
6	9	SPIRITS-VODKA-P4-750ML	6	9	SPIRITS-VODKA-P16-750ML
6	9	WINE-RED-P29-750ML	6	9	SPIRITS-VODKA-P12-750ML

TABLE A.32: *Part4: A description of each product and cluster membership for the original 10 clusters, as well as the sub-set of Cluster 6 when using k-means method of clustering*

Cluster	Sub Cluster	Product	Cluster	Sub Cluster	Product
6	9	WINE-WHITE-P22-750ML	6	9	SPIRITS-VODKA-P14-750ML
6	9	WINE-WHITE-P36-750ML	6	9	SPIRITS-VODKA-P15-750ML
6	9	WINE-WHITE-P33-750ML	6	9	SPIRITS-VODKA-P13-750ML
6	9	WINE-WHITE-P28-750ML	6	9	SPIRITS-VODKA-P5-750ML
6	9	WINE-RED-P48-750ML	6	9	SPIRITS-GIN-P10-750ML
6	9	WINE-RED-P51-750ML	6	9	WINE-BIB-P77-5LTR
6	9	WINE-RED-P59-750ML	6	9	WINE-BIB-P76-5LTR
6	9	WINE-RED-P39-750ML	6	9	WINE-BIB-P75-5LTR
6	9	WINE-ROSE-P10-750ML	6	9	WINE-RED-P67-750ML
6	9	WINE-WHITE-P39-750ML	6	9	WINE-RED-P81-750ML
6	9	SPIRITS-VODKA-P19-750ML	6	9	WINE-WHITE-P45-750ML
6	9	WINE-BIB-P68-5LTR	6	9	WINE-RED-P75-750ML
6	9	WINE-RED-P78-750ML	6	13	WINE-BIB-P18-2LTR
6	9	WINE-WHITE-P44-750ML	6	13	WINE-RED-P17-750ML
6	9	WINE-RED-P77-750ML	6	13	WINE-BIB-P20-2LTR
6	9	SPIRITS-GIN-P5-750ML	6	13	WINE-BIB-P33-3LTR
6	9	WINE-BIB-P66-5LTR	6	13	WINE-WHITE-P5-750ML
6	9	WINE-RED-P70-750ML	6	13	WINE-BIB-P8-1LTR
6	9	WINE-WHITE-P31-750ML	6	13	WINE-BIB-P15-1LTR
6	9	WINE-RED-P63-750ML	6	13	WINE-BIB-P40-3LTR
6	9	WINE-WHITE-P17-750ML	6	13	WINE-BIB-P12-1LTR
6	9	WINE-RED-P33-750ML	6	13	WINE-BIB-P13-1LTR
6	9	WINE-RED-P61-750ML	6	14	WINE-RED-P97-750ML
6	9	WINE-WHITE-P16-750ML	6	14	WINE-RED-P100-750ML
6	9	WINE-RED-P60-750ML	6	14	WINE-WHITE-P62-750ML
6	9	WINE-RED-P53-750ML	6	14	WINE-WHITE-P63-750ML
6	10	SPIRITS-WHISKY-P23-750ML	6	14	WINE-RED-P96-750ML
6	10	SPIRITS-WHISKY-P26-750ML	6	14	WINE-RED-P99-750ML
6	10	SPIRITS-WHISKY-P53-750ML	6	14	WINE-RED-P94-750ML
6	10	SPIRITS-GIN-P11-750ML	6	14	WINE-WHITE-P61-750ML
6	11	WINE-SPARKLING-P22-750ML	6	14	WINE-RED-P98-750ML
6	11	WINE-SPARKLING-P9-750ML	6	15	SPIRITS-BRANDY-P8-750ML
6	11	WINE-SPARKLING-P23-750ML	6	15	SPIRITS-WHISKY-P13-750ML
6	11	WINE-SPARKLING-P24-750ML	6	16	RTD-FABS-P11-275ML
6	11	WINE-SPARKLING-P21-750ML	6	16	RTD-FABS-P7-275ML
6	11	WINE-SPARKLING-P8-750ML	6	16	RTD-FABS-P6-250ML
6	11	WINE-SPARKLING-P5-750ML	6	16	RTD-FABS-P22-275ML
6	11	WINE-SPARKLING-P6-750ML	6	16	RTD-FABS-P5-250ML
6	11	WINE-SPARKLING-P31-750ML	6	16	RTD-FABS-P26-330ML
6	11	WINE-SPARKLING-P27-750ML	6	16	RTD-FABS-P25-330ML
6	11	WINE-SPARKLING-P7-750ML	6	16	RTD-CIDER-P2-330ML
6	11	WINE-SPARKLING-P10-750ML	6	16	RTD-CIDER-P10-330ML
6	11	WINE-SPARKLING-P1-750ML	6	16	RTD-FABS-P15-275ML
6	11	WINE-SPARKLING-P25-750ML	6	16	RTD-CIDER-P4-330ML
6	11	WINE-SPARKLING-P26-750ML	7		SPIRITS-BRANDY-P24-750ML
6	11	WINE-SPARKLING-P12-750ML	7		SPIRITS-BRANDY-P12-750ML
6	11	WINE-SPARKLING-P13-750ML	7		SPIRITS-BRANDY-P14-750ML
6	11	WINE-SPARKLING-P14-750ML	7		SPIRITS-VODKA-P6-750ML

TABLE A.33: *Part5: A description of each product and cluster membership for the original 10 clusters, as well as the sub-set of Cluster 6 when using k-means method of clustering*

Cluster	Sub Cluster	Product	Cluster	Sub Cluster	Product
6	11	WINE-SPARKLING-P29-750ML	7		RTD-CIDER-P5-330ML
6	11	WINE-SPARKLING-P32-750ML	8		WINE-PERLE-P6-300ML
6	11	WINE-SPARKLING-P28-750ML	8		WINE-PERLE-P5-300ML
6	11	WINE-SPARKLING-P30-750ML	8		WINE-BIB-P53-500ML
6	11	WINE-SPARKLING-P20-750ML	8		WINE-FORTIFIED-P3-200ML
6	11	WINE-SPARKLING-P18-750ML	8		WINE-FORTIFIED-P6-375ML
6	11	WINE-WHITE-P20-750ML	9		RTD-FABS-P21-275ML
6	11	WINE-SPARKLING-P2-750ML	9		RTD-FABS-P20-275ML
6	11	WINE-SPARKLING-P3-750ML	10		RTD-BEER-P14-330ML
6	11	WINE-SPARKLING-P4-750ML	10		RTD-BEER-P35-440ML
6	11	WINE-SPARKLING-P17-750ML	10		RTD-BEER-P13-330ML
6	11	WINE-SPARKLING-P15-750ML	10		RTD-BEER-P34-440ML
6	11	WINE-SPARKLING-P11-750ML	10		RTD-BEER-P7-330ML
6	11	WINE-SPARKLING-P33-750ML	10		RTD-BEER-P2-330ML
6	12	RTD-FABS-P36-440ML	10		RTD-BEER-P21-340ML
6	12	RTD-FABS-P35-440ML	10		RTD-BEER-P27-440ML
6	12	RTD-CIDER-P14-440ML	10		RTD-BEER-P32-440ML
6	12	RTD-CIDER-P3-330ML	10		RTD-BEER-P1-330ML
6	12	RTD-CIDER-P13-440ML	10		RTD-BEER-P20-340ML
6	13	WINE-RED-P28-750ML	10		RTD-BEER-P26-440ML
6	13	WINE-RED-P23-750ML	10		RTD-BEER-P3-330ML
6	13	WINE-RED-P38-750ML	10		RTD-BEER-P24-340ML
6	13	WINE-RED-P42-750ML	10		RTD-BEER-P31-440ML
6	13	WINE-WHITE-P14-750ML	10		RTD-BEER-P22-340ML
6	13	WINE-RED-P104-750ML	10		RTD-BEER-P38-440ML
6	13	WINE-RED-P106-750ML	10		RTD-BEER-P23-340ML
6	13	WINE-RED-P103-750ML	10		RTD-BEER-P39-440ML
6	13	WINE-WHITE-P64-750ML	10		RTD-BEER-P25-355ML
6	13	WINE-ROSE-P21-750ML	10		RTD-BEER-P4-330ML
6	13	WINE-WHITE-P65-750ML	10		RTD-BEER-P28-440ML
6	13	WINE-RED-P102-750ML	10		RTD-BEER-P4-330ML
6	13	WINE-WHITE-P8-750ML	10		RTD-BEER-P10-330ML
6	13	WINE-BIB-P6-1LTR	10		RTD-BEER-P29-440ML
6	13	WINE-BIB-P16-2LTR	10		RTD-BEER-P16-330ML
6	13	WINE-BIB-P29-3LTR	10		RTD-BEER-P37-440ML
6	13	WINE-RED-P14-750ML	10		RTD-BEER-P17-330ML
6	13	WINE-RED-P1-1.5LTR	10		RTD-BEER-P40-440ML
6	13	WINE-RED-P21-750ML	10		RTD-CIDER-P9-330ML
6	13	WINE-ROSE-P6-750ML	10		RTD-CIDER-P15-440ML
6	13	WINE-RED-P2-1.5LTR	10		RTD-CIDER-P7-330ML
6	13	WINE-RED-P25-750ML	10		RTD-CIDER-P1-330ML
6	13	WINE-WHITE-P9-750ML	10		RTD-CIDER-P11-440ML
6	13	WINE-BIB-P19-2LTR	10		RTD-CIDER-P8-330ML
6	13	WINE-BIB-P34-3LTR	10		RTD-BEER-P6-330ML
6	13	WINE-WHITE-P6-750ML	10		RTD-BEER-P36-440ML
6	13	WINE-BIB-P21-2LTR	10		RTD-BEER-P15-330ML
6	13	WINE-WHITE-P7-750ML	10		RTD-BEER-P8-330ML

TABLE A.34: *Part6: A description of each product and cluster membership for the original 10 clusters, as well as the sub-set of Cluster 6 when using k-means method of clustering*

Cluster	Sub Cluster	Product	Cluster	Sub Cluster	Product
6	13	WINE-BIB-P14-1LTR	10		RTD-CIDER-P6-330ML
6	13	WINE-BIB-P39-3LTR	10		RTD-CIDER-P12-440ML
6	13	WINE-BIB-P9-1LTR	10		RTD-BEER-P5-330ML
6	13	WINE-BIB-P37-3LTR	10		RTD-BEER-P9-330ML
6	13	WINE-BIB-P32-3LTR	10		RTD-BEER-P30-440ML
6	13	WINE-BIB-P11-1LTR	10		RTD-BEER-P12-330ML
6	13	WINE-BIB-P30-3LTR	10		RTD-BEER-P33-440ML
6	13	WINE-BIB-P35-3LTR	10		RTD-BEER-P42-440ML
6	13	WINE-BIB-P7-1LTR	10		RTD-BEER-P19-330ML
6	13	WINE-BIB-P17-2LTR	10		RTD-BEER-P41-440ML
6	13	WINE-BIB-P31-3LTR	10		RTD-BEER-P18-330ML
6	13	WINE-RED-P13-750ML			
6	13	WINE-BIB-P38-3LTR			
6	13	WINE-RED-P18-750ML			
6	13	WINE-BIB-P10-1LTR			
6	13	WINE-BIB-P36-3LTR			
6	13	WINE-RED-P15-750ML			

TABLE A.35: *Part7: A description of each product and cluster membership for the original 10 clusters, as well as the sub-set of Cluster 6 when using k-means method of clustering*

A.3.3   Validation

Product	Cluster	Distance	Product	Cluster	Distance
RTD-FABS-P27-330ML	1	4.27	WINE-WHITE-P22-750ML	6	2.446
RTD-FABS-P30-330ML	1	4.27	WINE-WHITE-P36-750ML	6	2.4459
RTD-BEER-P51-660ML	2	4.0393	WINE-WHITE-P33-750ML	6	2.4466
RTD-FABS-P38-660ML	2	6.4725	WINE-WHITE-P28-750ML	6	2.4466
RTD-BEER-P44-500ML	2	3.3812	WINE-RED-P48-750ML	6	2.4464
RTD-BEER-P58-750ML	2	4.046	WINE-RED-P51-750ML	6	2.4458
RTD-BEER-P43-500ML	2	3.1297	WINE-PERLE-P8-5LTR	6	6.0135
RTD-BEER-P55-750ML	2	3.1206	WINE-PERLE-P2-1LTR	6	6.013
RTD-BEER-P45-500ML	2	5.4553	WINE-PERLE-P10-750ML	6	5.9533
RTD-BEER-P50-660ML	2	4.9968	WINE-PERLE-P7-5LTR	6	6.0098
RTD-BEER-P59-750ML	2	4.7735	WINE-PERLE-P11-750ML	6	5.9495
RTD-BEER-P46-500ML	2	4.1972	WINE-PERLE-P9-5LTR	6	6.0119
RTD-BEER-P57-750ML	2	4.1387	WINE-PERLE-P12-750ML	6	5.9528
RTD-BEER-P48-5LTR	2	3.1788	RTD-FABS-P11-275ML	6	5.8941
RTD-BEER-P49-650ML	2	3.0931	SPIRITS-RUM-P5-750ML	6	4.2504
RTD-CIDER-P18-660ML	2	7.4246	SPIRITS-WHISKY-P27-750ML	6	6.8428
RTD-CIDER-P17-660ML	2	5.6641	SPIRITS-WHISKY-P15-750ML	6	5.1945
RTD-BEER-P47-500ML	2	3.8194	WINE-RED-P59-750ML	6	2.4466
RTD-BEER-P56-750ML	2	3.7515	WINE-RED-P39-750ML	6	2.4456
RTD-FABS-P37-660ML	2	6.5309	WINE-ROSE-P10-750ML	6	2.4467
RTD-CIDER-P16-500ML	2	5.7424	WINE-WHITE-P39-750ML	6	2.4472
RTD-BEER-P52-660ML	2	4.7357	SPIRITS-WHISKY-P1-1LTR	6	6.9835
RTD-BEER-P53-660ML	2	6.4532	SPIRITS-WHISKY-P16-750ML	6	5.5607
SPIRITS-LIQUEURS-P2-1LTR	3	5.4839	SPIRITS-WHISKY-P18-750ML	6	5.5756
SPIRITS-LIQUEURS-P16-750ML	3	2.4143	SPIRITS-VODKA-P19-750ML	6	5.3683
SPIRITS-LIQUEURS-P1-1LTR	3	5.2135	WINE-RED-P28-750ML	6	3.8521
SPIRITS-LIQUEURS-P9-750ML	3	1.0437	WINE-RED-P23-750ML	6	3.8571
SPIRITS-LIQUEURS-P12-750ML	3	1.0621	SPIRITS-WHISKY-P14-750ML	6	6.0942
SPIRITS-LIQUEURS-P14-750ML	3	1.0461	WINE-BIB-P68-5LTR	6	3.5184
SPIRITS-LIQUEURS-P11-750ML	3	1.0633	WINE-BIB-P72-5LTR	6	3.5182
SPIRITS-LIQUEURS-P10-750ML	3	1.0729	WINE-BIB-P65-5LTR	6	3.5196
SPIRITS-LIQUEURS-P13-750ML	3	1.0489	WINE-BIB-P64-5LTR	6	3.5199
SPIRITS-LIQUEURS-P8-750ML	3	1.6519	WINE-RED-P66-750ML	6	2.4467
SPIRITS-LIQUEURS-P7-750ML	3	1.6509	WINE-RED-P64-750ML	6	2.447
SPIRITS-LIQUEURS-P6-750ML	3	1.6511	WINE-RED-P72-750ML	6	4.6622
SPIRITS-LIQUEURS-P3-750ML	3	1.5389	WINE-RED-P85-750ML	6	4.5893
SPIRITS-LIQUEURS-P4-750ML	3	1.5204	WINE-ROSE-P13-750ML	6	4.6119
SPIRITS-LIQUEURS-P5-750ML	3	3.8065	WINE-WHITE-P47-750ML	6	4.6089
SPIRITS-LIQUEURS-P15-750ML	3	1.849	WINE-RED-P57-750ML	6	2.4466
SPIRITS-GIN-P2-200ML	4	2.7131	WINE-WHITE-P37-750ML	6	2.4458
SPIRITS-GIN-P3-375ML	4	2.7093	SPIRITS-GIN-P7-750ML	6	5.4539
SPIRITS-BRANDY-P5-375ML	4	5.1596	WINE-SPARKLING-P9-750ML	6	5.6145
SPIRITS-VODKA-P3-200ML	4	4.3663	SPIRITS-WHISKY-P37-750ML	6	4.1208
RTD-FABS-P9-275ML	5	3.6932	WINE-ROSE-P17-750ML	6	8.4007
RTD-FABS-P8-275ML	5	3.691	WINE-WHITE-P58-750ML	6	8.3807
RTD-FABS-P13-275ML	5	3.6871	WINE-WHITE-P59-750ML	6	8.3798
RTD-FABS-P31-440ML	5	5.8374	WINE-WHITE-P24-750ML	6	2.4452

TABLE A.36: *Part 1: Display of the distance between each product and its cluster centre using the k-means method*

Product	Cluster	Distance	Product	Cluster	Distance
RTD-FABS-P12-275ML	5	3.6851	SPIRITS-WHISKY-P43-750ML	6	4.1362
RTD-FABS-P10-275ML	5	3.6837	WINE-FORTIFIED-P1-1LTR	6	4.4371
WINE-BIB-P84-5LTR	5	9.4058	WINE-FORTIFIED-P8-750ML	6	3.8172
WINE-WHITE-P68-750ML	5	9.5641	SPIRITS-CANE-P1-1LTR	6	7.5308
WINE-ROSE-P4-5LTR	5	9.4275	SPIRITS-CANE-P2-750ML	6	4.8833
SPIRITS-WHISKY-P11-750ML	5	9.1102	WINE-SPARKLING-P23-750ML	6	4.9156
RTD-FABS-P14-275ML	5	5.1805	WINE-SPARKLING-P24-750ML	6	4.9157
RTD-FABS-P28-330ML	5	5.1828	SPIRITS-RUM-P2-750ML	6	6.3449
RTD-FABS-P32-440ML	5	6.7636	SPIRITS-RUM-P1-1LTR	6	7.7977
RTD-FABS-P17-275ML	5	3.0537	SPIRITS-RUM-P7-750ML	6	6.1567
RTD-FABS-P16-275ML	5	3.7297	RTD-FABS-P7-275ML	6	5.8903
RTD-FABS-P19-275ML	5	3.0689	WINE-BIB-P81-5LTR	6	6.1914
RTD-FABS-P18-275ML	5	3.0901	WINE-BIB-P80-5LTR	6	6.1882
RTD-FABS-P2-250ML	5	4.276	WINE-RED-P95-750ML	6	6.2953
RTD-FABS-P4-250ML	5	4.4895	WINE-RED-P40-750ML	6	2.4458
RTD-FABS-P3-250ML	5	4.59	SPIRITS-WHISKY-P29-750ML	6	4.4494
RTD-FABS-P1-250ML	5	3.4589	SPIRITS-VODKA-P17-750ML	6	5.3566
RTD-FABS-P24-300ML	5	4.2653	SPIRITS-BRANDY-P19-750ML	6	4.3622
RTD-FABS-P29-330ML	5	4.2716	WINE-RED-P62-750ML	6	2.4466
RTD-FABS-P34-440ML	5	5.9649	SPIRITS-BRANDY-P18-750ML	6	4.2263
RTD-FABS-P23-300ML	5	4.8667	SPIRITS-VODKA-P18-750ML	6	4.8345
RTD-FABS-P33-440ML	5	6.2777	SPIRITS-VODKA-P20-750ML	6	4.828
WINE-BIB-P51-3LTR	5	7.8101	WINE-SPARKLING-P21-750ML	6	4.8976
WINE-RED-P105-750ML	5	7.8945	WINE-RED-P65-750ML	6	2.4456
WINE-BIB-P52-3LTR	5	7.8136	WINE-WHITE-P23-750ML	6	2.4459
WINE-WHITE-P66-750ML	5	7.9098	WINE-RED-P50-750ML	6	2.4436
SPIRITS-WHISKY-P5-1LTR	6	6.3464	WINE-RED-P31-750ML	6	2.4517
SPIRITS-WHISKY-P41-750ML	6	4.1229	WINE-FORTIFIED-P10-750ML	6	3.8204
WINE-BIB-P79-5LTR	6	4.0556	WINE-PERLE-P1-1LTR	6	4.7761
WINE-BIB-P42-3LTR	6	4.136	WINE-WHITE-P12-750ML	6	4.2271
WINE-BIB-P79-5LTR	6	4.257	WINE-RED-P26-750ML	6	2.4474
WINE-RED-P71-750ML	6	4.5344	WINE-RED-P79-750ML	6	2.926
WINE-BIB-P82-5LTR	6	4.0737	WINE-WHITE-P1-5LTR	6	3.4789
WINE-BIB-P43-3LTR	6	4.1869	WINE-WHITE-P56-750ML	6	2.9476
WINE-BIB-P82-5LTR	6	4.4987	WINE-RED-P89-750ML	6	2.9393
WINE-ROSE-P12-750ML	6	4.5432	WINE-ROSE-P2-5LTR	6	3.4785
WINE-BIB-P41-3LTR	6	4.0973	WINE-ROSE-P15-750ML	6	2.9461
WINE-BIB-P78-5LTR	6	4.1929	WINE-WHITE-P50-750ML	6	2.9323
WINE-SPARKLING-P22-750ML	6	4.9062	WINE-RED-P86-750ML	6	2.9354
SPIRITS-VODKA-P4-750ML	6	5.8746	WINE-WHITE-P49-750ML	6	2.9315
WINE-RED-P29-750ML	6	2.4488	WINE-WHITE-P67-750ML	6	4.732
WINE-RED-P19-750ML	6	3.7112	WINE-BIB-P83-5LTR	6	3.8055
WINE-RED-P22-750ML	6	3.6224	WINE-BIB-P84-5LTR	6	3.7962
WINE-RED-P101-750ML	6	3.871	SPIRITS-WHISKY-P25-750ML	6	6.2221
WINE-ROSE-P4-5LTR	6	3.7956	WINE-RED-P35-750ML	6	4.2432
WINE-ROSE-P20-750ML	6	3.8702	WINE-ROSE-P8-750ML	6	4.2416
WINE-WHITE-P42-750ML	6	4.6964	WINE-RED-P20-750ML	6	4.2188
WINE-BIB-P47-3LTR	6	4.4117	WINE-WHITE-P38-750ML	6	2.4467
WINE-RED-P84-750ML	6	3.9069	WINE-WHITE-P26-750ML	6	4.7962
WINE-BIB-P46-3LTR	6	4.4154	WINE-RED-P47-750ML	6	4.802

TABLE A.37: *Part 2: Display of the distance between each product and its cluster centre using the k-means method*

Product	Cluster	Distance	Product	Cluster	Distance
WINE-WHITE-P55-750ML	6	3.9071	WINE-WHITE-P27-750ML	6	4.7992
WINE-BIB-P45-3LTR	6	4.4103	WINE-SPARKLING-P16-750ML	6	5.7655
WINE-RED-P87-750ML	6	3.9068	WINE-SPARKLING-P19-750ML	6	5.7652
WINE-BIB-P48-3LTR	6	4.4044	SPIRITS-BRANDY-P1-1LTR	6	6.2176
WINE-WHITE-P54-750ML	6	3.895	SPIRITS-BRANDY-P7-750ML	6	3.9235
WINE-RED-P90-750ML	6	4.7431	SPIRITS-BRANDY-P22-750ML	6	5.221
WINE-RED-P88-750ML	6	4.7452	WINE-SPARKLING-P12-750ML	6	4.7411
WINE-ROSE-P16-750ML	6	4.7568	WINE-SPARKLING-P13-750ML	6	4.7418
WINE-RED-P92-750ML	6	4.7427	WINE-SPARKLING-P14-750ML	6	4.7414
WINE-WHITE-P57-750ML	6	4.7592	SPIRITS-BRANDY-P6-750ML	6	5.0102
WINE-WHITE-P52-750ML	6	4.7736	SPIRITS-BRANDY-P16-750ML	6	5.3515
WINE-RED-P91-750ML	6	4.7413	SPIRITS-BRANDY-P8-750ML	6	7.0216
SPIRITS-BRANDY-P21-750ML	6	4.3651	WINE-RED-P97-750ML	6	7.34
WINE-RED-P56-750ML	6	2.4465	WINE-RED-P100-750ML	6	7.3339
WINE-RED-P27-750ML	6	4.253	WINE-WHITE-P62-750ML	6	7.3348
WINE-WHITE-P15-750ML	6	4.2493	WINE-WHITE-P63-750ML	6	7.3341
WINE-RED-P30-750ML	6	4.2525	WINE-RED-P96-750ML	6	7.341
WINE-RED-P36-750ML	6	4.2532	WINE-RED-P99-750ML	6	7.3393
WINE-WHITE-P19-750ML	6	4.2504	WINE-RED-P94-750ML	6	7.3551
WINE-RED-P34-750ML	6	4.2535	WINE-WHITE-P61-750ML	6	7.3365
SPIRITS-WHISKY-P2-1LTR	6	6.8588	WINE-RED-P98-750ML	6	7.3388
SPIRITS-WHISKY-P20-750ML	6	4.9214	WINE-RED-P38-750ML	6	3.8685
SPIRITS-BRANDY-P20-750ML	6	4.3134	WINE-RED-P42-750ML	6	3.868
WINE-RED-P55-750ML	6	2.4457	WINE-WHITE-P14-750ML	6	3.8523
WINE-WHITE-P40-750ML	6	2.4465	WINE-RED-P44-750ML	6	3.6269
WINE-WHITE-P35-750ML	6	2.6177	WINE-RED-P43-750ML	6	3.6261
SPIRITS-GIN-P9-750ML	6	4.8693	WINE-RED-P54-750ML	6	2.4463
SPIRITS-WHISKY-P36-750ML	6	5.219	WINE-RED-P76-750ML	6	3.8294
SPIRITS-WHISKY-P30-750ML	6	4.4348	SPIRITS-VODKA-P21-750ML	6	4.8251
SPIRITS-WHISKY-P28-750ML	6	4.5586	SPIRITS-RUM-P6-750ML	6	4.2489
SPIRITS-WHISKY-P31-750ML	6	4.4326	WINE-RED-P12-750ML	6	2.5602
SPIRITS-GIN-P1-1LTR	6	7.2566	WINE-SPARKLING-P29-750ML	6	5.9174
SPIRITS-GIN-P4-750ML	6	5.7027	WINE-SPARKLING-P32-750ML	6	4.9573
WINE-WHITE-P43-750ML	6	4.8573	WINE-SPARKLING-P28-750ML	6	4.966
WINE-ROSE-P14-750ML	6	4.8616	WINE-SPARKLING-P30-750ML	6	5.9522
WINE-SPARKLING-P8-750ML	6	5.4479	WINE-FORTIFIED-P11-750ML	6	3.8236
WINE-SPARKLING-P5-750ML	6	5.3317	WINE-FORTIFIED-P9-750ML	6	3.8151
WINE-SPARKLING-P6-750ML	6	5.3318	SPIRITS-WHISKY-P34-750ML	6	4.1658
WINE-RED-P80-750ML	6	3.3903	SPIRITS-WHISKY-P46-750ML	6	4.1284
WINE-RED-P83-750ML	6	3.3888	WINE-FORTIFIED-P4-2LTR	6	4.4624
WINE-WHITE-P51-750ML	6	3.3863	WINE-FORTIFIED-P14-750ML	6	3.8271
WINE-ROSE-P9-750ML	6	2.4462	WINE-BIB-P3-1LTR	6	3.6299
WINE-WHITE-P25-750ML	6	2.4457	WINE-BIB-P26-3LTR	6	3.6309
SPIRITS-WHISKY-P3-1LTR	6	6.3622	WINE-BIB-P61-5LTR	6	3.627
SPIRITS-WHISKY-P24-750ML	6	4.2888	WINE-RED-P8-750ML	6	3.9321
WINE-WHITE-P10-750ML	6	4.1357	WINE-BIB-P5-1LTR	6	3.6302
SPIRITS-GIN-P8-750ML	6	4.7902	WINE-BIB-P23-3LTR	6	3.6325
SPIRITS-WHISKY-P44-750ML	6	4.1209	WINE-BIB-P60-5LTR	6	3.6294
WINE-WHITE-P32-750ML	6	2.447	WINE-BIB-P27-3LTR	6	3.6304
WINE-ROSE-P11-750ML	6	2.4444	WINE-BIB-P62-5LTR	6	3.6285

TABLE A.38: *Part 3: Display of the distance between each product and its cluster centre using the k-means method*

Product	Cluster	Distance	Product	Cluster	Distance
WINE-RED-P58-750ML	6	2.4462	WINE-BIB-P2-1LTR	6	3.6313
WINE-WHITE-P34-750ML	6	2.4462	WINE-BIB-P22-3LTR	6	3.6324
SPIRITS-BRANDY-P17-750ML	6	4.2068	WINE-BIB-P55-5LTR	6	3.6299
RTD-FABS-P6-250ML	6	5.5917	WINE-BIB-P1-1LTR	6	3.6301
RTD-FABS-P22-275ML	6	5.6031	WINE-BIB-P24-3LTR	6	3.6311
RTD-FABS-P36-440ML	6	6.6989	WINE-BIB-P59-5LTR	6	3.6279
RTD-FABS-P5-250ML	6	5.6212	WINE-BIB-P57-5LTR	6	3.6274
RTD-FABS-P35-440ML	6	6.7287	WINE-RED-P11-750ML	6	3.9322
SPIRITS-GIN-P6-750ML	6	4.7946	WINE-BIB-P56-5LTR	6	3.6293
WINE-WHITE-P13-750ML	6	2.4462	WINE-BIB-P4-1LTR	6	3.6299
SPIRITS-WHISKY-P4-1LTR	6	7.1121	WINE-BIB-P28-3LTR	6	3.6312
SPIRITS-WHISKY-P22-750ML	6	5.8031	WINE-BIB-P73-5LTR	6	3.6349
RTD-FABS-P26-330ML	6	5.89	WINE-RED-P10-750ML	6	3.9316
SPIRITS-WHISKY-P23-750ML	6	5.8277	WINE-BIB-P25-3LTR	6	3.631
SPIRITS-WHISKY-P26-750ML	6	5.771	WINE-BIB-P58-5LTR	6	3.6265
SPIRITS-WHISKY-P8-1LTR	6	7.0137	WINE-RED-P9-750ML	6	3.932
SPIRITS-WHISKY-P21-750ML	6	5.8241	WINE-BIB-P54-5LTR	6	3.6291
SPIRITS-WHISKY-P50-750ML	6	5.3372	WINE-RED-P104-750ML	6	5.0157
WINE-SPARKLING-P31-750ML	6	5.7435	WINE-RED-P106-750ML	6	5.1372
WINE-SPARKLING-P27-750ML	6	6.0382	WINE-RED-P103-750ML	6	5.0207
WINE-SPARKLING-P7-750ML	6	5.7341	WINE-WHITE-P64-750ML	6	5.0238
WINE-SPARKLING-P10-750ML	6	4.9713	WINE-ROSE-P21-750ML	6	5.0309
WINE-SPARKLING-P1-750ML	6	6.592	WINE-WHITE-P65-750ML	6	5.0176
WINE-SPARKLING-P25-750ML	6	5.3572	WINE-RED-P102-750ML	6	5.0129
WINE-SPARKLING-P26-750ML	6	5.3549	SPIRITS-GIN-P11-750ML	6	7.2818
SPIRITS-WHISKY-P45-750ML	6	4.1261	SPIRITS-BRANDY-P2-1LTR	6	6.2581
SPIRITS-WHISKY-P10-1LTR	6	7.1954	SPIRITS-BRANDY-P9-750ML	6	3.9499
SPIRITS-WHISKY-P53-750ML	6	5.7993	WINE-BIB-P67-5LTR	6	3.5201
SPIRITS-WHISKY-P42-750ML	6	5.472	WINE-BIB-P71-5LTR	6	3.5217
SPIRITS-WHISKY-P52-750ML	6	5.3938	WINE-BIB-P70-5LTR	6	3.5218
SPIRITS-WHISKY-P49-750ML	6	5.4384	WINE-BIB-P63-5LTR	6	4.2193
SPIRITS-WHISKY-P51-750ML	6	4.8967	WINE-BIB-P69-5LTR	6	4.2255
SPIRITS-WHISKY-P9-1LTR	6	7.514	WINE-BIB-P74-5LTR	6	4.2274
WINE-PERLE-P3-1LTR	6	4.763	SPIRITS-VODKA-P14-750ML	6	4.7623
WINE-PERLE-P4-2LTR	6	4.7572	SPIRITS-VODKA-P15-750ML	6	4.7631
WINE-PERLE-P13-750ML	6	4.2306	SPIRITS-VODKA-P13-750ML	6	4.7619
WINE-SPARKLING-P20-750ML	6	4.9203	SPIRITS-VODKA-P2-1LTR	6	7.5918
WINE-RED-P49-750ML	6	2.4449	SPIRITS-VODKA-P5-750ML	6	5.994
WINE-SPARKLING-P18-750ML	6	5.138	RTD-FABS-P15-275ML	6	5.6178
WINE-WHITE-P20-750ML	6	6.5225	WINE-RED-P37-750ML	6	4.0754
WINE-FORTIFIED-P7-750ML	6	3.8056	WINE-WHITE-P21-750ML	6	4.073
WINE-SPARKLING-P2-750ML	6	4.7422	WINE-WHITE-P30-750ML	6	4.0772
WINE-SPARKLING-P3-750ML	6	4.7395	WINE-WHITE-P29-750ML	6	4.0677
WINE-RED-P73-750ML	6	3.8972	WINE-RED-P32-750ML	6	4.073
WINE-WHITE-P53-750ML	6	3.8996	WINE-RED-P46-750ML	6	4.0778
WINE-WHITE-P46-750ML	6	3.8999	WINE-WHITE-P18-750ML	6	4.067
WINE-RED-P74-750ML	6	3.8967	WINE-RED-P41-750ML	6	4.079
WINE-WHITE-P48-750ML	6	3.8958	WINE-SPARKLING-P11-750ML	6	4.8804
WINE-RED-P82-750ML	6	3.8985	SPIRITS-GIN-P10-750ML	6	4.8724
WINE-RED-P45-750ML	6	4.054	RTD-CIDER-P4-330ML	6	7.0803

TABLE A.39: *Part 4: Display of the distance between each product and its cluster centre using the k-means method*

Product	Cluster	Distance	Product	Cluster	Distance
SPIRITS-RUM-P3-750ML	6	4.2062	RTD-CIDER-P14-440ML	6	7.8117
SPIRITS-RUM-P4-750ML	6	4.2473	RTD-CIDER-P3-330ML	6	7.5671
RTD-FABS-P25-330ML	6	6.1097	RTD-CIDER-P13-440ML	6	8.1977
SPIRITS-BRANDY-P15-750ML	6	4.1998	WINE-BIB-P77-5LTR	6	3.518
SPIRITS-BRANDY-P4-1LTR	6	6.684	WINE-BIB-P76-5LTR	6	3.5186
SPIRITS-BRANDY-P23-750ML	6	4.7438	WINE-BIB-P75-5LTR	6	3.5199
WINE-WHITE-P41-750ML	6	2.4474	WINE-RED-P67-750ML	6	2.4471
WINE-RED-P52-750ML	6	2.4482	WINE-RED-P81-750ML	6	2.9279
WINE-RED-P69-750ML	6	2.447	WINE-WHITE-P45-750ML	6	2.9233
WINE-WHITE-P8-750ML	6	3.3309	WINE-RED-P75-750ML	6	2.935
WINE-BIB-P6-1LTR	6	3.2283	WINE-RED-P78-750ML	6	2.9269
WINE-BIB-P16-2LTR	6	3.2282	WINE-WHITE-P44-750ML	6	2.9247
WINE-BIB-P29-3LTR	6	3.2267	WINE-RED-P77-750ML	6	2.9252
WINE-RED-P14-750ML	6	3.3359	SPIRITS-GIN-P5-750ML	6	5.4507
WINE-RED-P1-5LTR	6	3.2262	WINE-BIB-P66-5LTR	6	4.6739
WINE-RED-P21-750ML	6	3.3249	WINE-RED-P70-750ML	6	4.2332
WINE-ROSE-P6-750ML	6	3.3252	WINE-WHITE-P31-750ML	6	2.4462
WINE-RED-P2-5LTR	6	3.2243	SPIRITS-WHISKY-P40-750ML	6	4.1291
WINE-RED-P25-750ML	6	3.3224	SPIRITS-WHISKY-P32-750ML	6	4.4331
WINE-WHITE-P9-750ML	6	3.3282	SPIRITS-WHISKY-P38-750ML	6	4.4289
WINE-BIB-P19-2LTR	6	3.2282	SPIRITS-WHISKY-P12-750ML	6	7.0857
WINE-BIB-P34-3LTR	6	3.2248	SPIRITS-WHISKY-P19-750ML	6	5.1727
WINE-WHITE-P6-750ML	6	3.3317	WINE-RED-P63-750ML	6	2.4469
WINE-BIB-P21-2LTR	6	3.2282	WINE-WHITE-P17-750ML	6	2.4457
WINE-WHITE-P7-750ML	6	3.3311	SPIRITS-WHISKY-P35-750ML	6	4.1354
WINE-BIB-P14-1LTR	6	3.2219	SPIRITS-WHISKY-P7-1LTR	6	6.3466
WINE-BIB-P39-3LTR	6	3.2221	SPIRITS-WHISKY-P47-750ML	6	4.1207
WINE-BIB-P9-1LTR	6	3.2262	WINE-RED-P33-750ML	6	2.4448
WINE-BIB-P37-3LTR	6	3.2244	WINE-RED-P3-5LTR	6	4.1287
WINE-BIB-P32-3LTR	6	3.2244	WINE-BIB-P49-3LTR	6	4.1325
WINE-BIB-P11-1LTR	6	3.2261	WINE-RED-P93-750ML	6	4.4543
WINE-BIB-P30-3LTR	6	3.2278	WINE-BIB-P50-3LTR	6	4.1353
WINE-BIB-P35-3LTR	6	3.2272	WINE-ROSE-P19-750ML	6	4.465
WINE-BIB-P7-1LTR	6	3.2273	WINE-WHITE-P60-750ML	6	4.4494
WINE-BIB-P17-2LTR	6	3.2272	WINE-ROSE-P18-750ML	6	4.4434
WINE-BIB-P31-3LTR	6	3.2246	WINE-RED-P4-5LTR	6	4.1492
WINE-RED-P13-750ML	6	3.3345	WINE-ROSE-P3-5LTR	6	4.1685
WINE-BIB-P38-3LTR	6	3.2156	WINE-WHITE-P2-5LTR	6	4.1297
WINE-RED-P18-750ML	6	3.3284	WINE-RED-P61-750ML	6	2.4465
WINE-ROSE-P6-750ML	6	4.5337	WINE-WHITE-P16-750ML	6	2.4464
WINE-BIB-P44-3LTR	6	4.5163	WINE-RED-P60-750ML	6	2.4468
WINE-ROSE-P5-750ML	6	4.7406	SPIRITS-WHISKY-P17-750ML	6	5.2532
WINE-BIB-P10-1LTR	6	3.2261	WINE-SPARKLING-P33-750ML	6	5.1952
WINE-BIB-P36-3LTR	6	3.2274	SPIRITS-BRANDY-P11-750ML	6	5.0787
WINE-RED-P15-750ML	6	3.3361	SPIRITS-BRANDY-P13-750ML	6	5.0761
WINE-BIB-P18-2LTR	6	3.2283	WINE-FORTIFIED-P13-750ML	6	3.8241
WINE-RED-P17-750ML	6	3.3279	WINE-RED-P24-750ML	6	3.6281
WINE-BIB-P20-2LTR	6	3.2269	WINE-WHITE-P11-750ML	6	3.6423
WINE-BIB-P33-3LTR	6	3.2212	WINE-ROSE-P7-750ML	6	3.6403
WINE-WHITE-P5-750ML	6	3.3285	SPIRITS-BRANDY-P3-1LTR	6	6.5304

TABLE A.40: *Part 5: Display of the distance between each product and its cluster centre using the k-means method*

Product	Cluster	Distance	Product	Cluster	Distance
WINE-BIB-P8-1LTR	6	3.2262	SPIRITS-BRANDY-P10-750ML	6	4.5128
WINE-BIB-P15-1LTR	6	3.2331	SPIRITS-WHISKY-P6-1LTR	6	6.3467
WINE-BIB-P40-3LTR	6	3.219	SPIRITS-WHISKY-P39-750ML	6	4.1229
WINE-SPARKLING-P4-750ML	6	5.9453	SPIRITS-WHISKY-P48-750ML	6	4.1204
WINE-BIB-P12-1LTR	6	3.2224	WINE-RED-P53-750ML	6	2.4456
WINE-BIB-P13-1LTR	6	3.2216	WINE-RED-P5-750ML	6	5.8369
SPIRITS-VODKA-P22-750ML	6	4.8807	WINE-WHITE-P4-750ML	6	5.8469
WINE-RED-P16-750ML	6	4.3457	WINE-RED-P7-750ML	6	5.8371
SPIRITS-VODKA-P9-750ML	6	4.8245	WINE-RED-P6-750ML	6	5.8388
SPIRITS-VODKA-P11-750ML	6	4.8208	WINE-WHITE-P3-750ML	6	5.8487
SPIRITS-VODKA-P1-1LTR	6	6.9197	WINE-FORTIFIED-P12-750ML	6	3.8246
SPIRITS-VODKA-P7-750ML	6	4.7924	SPIRITS-BRANDY-P24-750ML	7	4.7443
SPIRITS-VODKA-P10-750ML	6	4.8191	SPIRITS-BRANDY-P12-750ML	7	4.5276
SPIRITS-VODKA-P8-750ML	6	4.8145	SPIRITS-BRANDY-P14-750ML	7	4.3765
RTD-CIDER-P2-330ML	6	7.2236	SPIRITS-VODKA-P6-750ML	7	6.3108
RTD-CIDER-P10-330ML	6	7.8836	RTD-CIDER-P5-330ML	7	8.1754
SPIRITS-WHISKY-P13-750ML	6	8.6734	WINE-PERLE-P6-300ML	8	5.0575
WINE-FORTIFIED-P2-1LTR	6	4.6874	WINE-PERLE-P5-300ML	8	5.058
WINE-FORTIFIED-P5-2LTR	6	4.7032	WINE-BIB-P53-500ML	8	3.987
WINE-FORTIFIED-P16-750ML	6	4.9466	WINE-FORTIFIED-P3-200ML	8	4.4586
WINE-FORTIFIED-P15-750ML	6	3.8329	WINE-FORTIFIED-P6-375ML	8	4.4555
WINE-SPARKLING-P17-750ML	6	4.8873	RTD-FABS-P21-275ML	9	0.0753
WINE-SPARKLING-P15-750ML	6	4.8894	RTD-FABS-P20-275ML	9	0.0753
WINE-RED-P68-750ML	6	2.4471	RTD-BEER-P14-330ML	10	4.5326
SPIRITS-WHISKY-P33-750ML	6	4.2065	RTD-BEER-P35-440ML	10	4.9272
SPIRITS-VODKA-P16-750ML	6	4.7623	RTD-BEER-P13-330ML	10	6.8278
SPIRITS-VODKA-P12-750ML	6	4.7505	RTD-BEER-P34-440ML	10	6.6937
RTD-BEER-P8-330ML	10	3.8898	RTD-BEER-P17-330ML	10	4.164
RTD-BEER-P2-330ML	10	4.3592	RTD-BEER-P40-440ML	10	4.4197
RTD-BEER-P21-340ML	10	4.8163	RTD-CIDER-P9-330ML	10	8.635
RTD-BEER-P27-440ML	10	4.9972	RTD-CIDER-P15-440ML	10	9.0067
RTD-BEER-P32-440ML	10	4.6283	RTD-CIDER-P7-330ML	10	7.2971
RTD-BEER-P1-330ML	10	3.9535	RTD-CIDER-P1-330ML	10	6.6043
RTD-BEER-P20-340ML	10	4.1199	RTD-CIDER-P11-440ML	10	7.3536
RTD-BEER-P26-440ML	10	4.2903	RTD-CIDER-P8-330ML	10	7.2876
RTD-BEER-P3-330ML	10	6.4633	RTD-BEER-P7-330ML	10	3.9016
RTD-BEER-P24-340ML	10	9.3936	RTD-BEER-P36-440ML	10	4.621
RTD-BEER-P31-440ML	10	6.8175	RTD-BEER-P15-330ML	10	4.384
RTD-BEER-P22-340ML	10	4.6094	RTD-BEER-P9-330ML	10	3.8996
RTD-BEER-P38-440ML	10	5.1595	RTD-CIDER-P6-330ML	10	6.7562
RTD-BEER-P23-340ML	10	4.8877	RTD-CIDER-P12-440ML	10	7.0975
RTD-BEER-P39-440ML	10	5.3689	RTD-BEER-P6-330ML	10	3.8874
RTD-BEER-P25-355ML	10	5.3015	RTD-BEER-P10-330ML	10	3.8816
RTD-BEER-P5-330ML	10	5.306	RTD-BEER-P30-440ML	10	4.8239
RTD-BEER-P28-440ML	10	5.55	RTD-BEER-P12-330ML	10	3.887
RTD-BEER-P4-330ML	10	4.9737	RTD-BEER-P33-440ML	10	4.6177
RTD-BEER-P11-330ML	10	5.166	RTD-BEER-P42-440ML	10	5.373
RTD-BEER-P29-440ML	10	5.8702	RTD-BEER-P19-330ML	10	5.2249
RTD-BEER-P16-330ML	10	4.5293	RTD-BEER-P41-440ML	10	5.8618
RTD-BEER-P37-440ML	10	5.177	RTD-BEER-P18-330ML	10	5.4983

TABLE A.41: *Part 6: Display of the distance between each product and its cluster centre using the k-means method*

A.4   Results for determining a competitor set using the product life cycle methodology

A.4.1   Cluster analysis and the product life cycle

In Figure A.10 the results for the introductory life cycle stage are shown, firstly highlighting the clusters and then the segments (wine, spirits and RTDs). In each of the figures to follow the product's sales value for the calender year 2017, is represented by the relative size of the bubble. The results can be further viewed in more detail in Figures A.11 and Figures A.12.
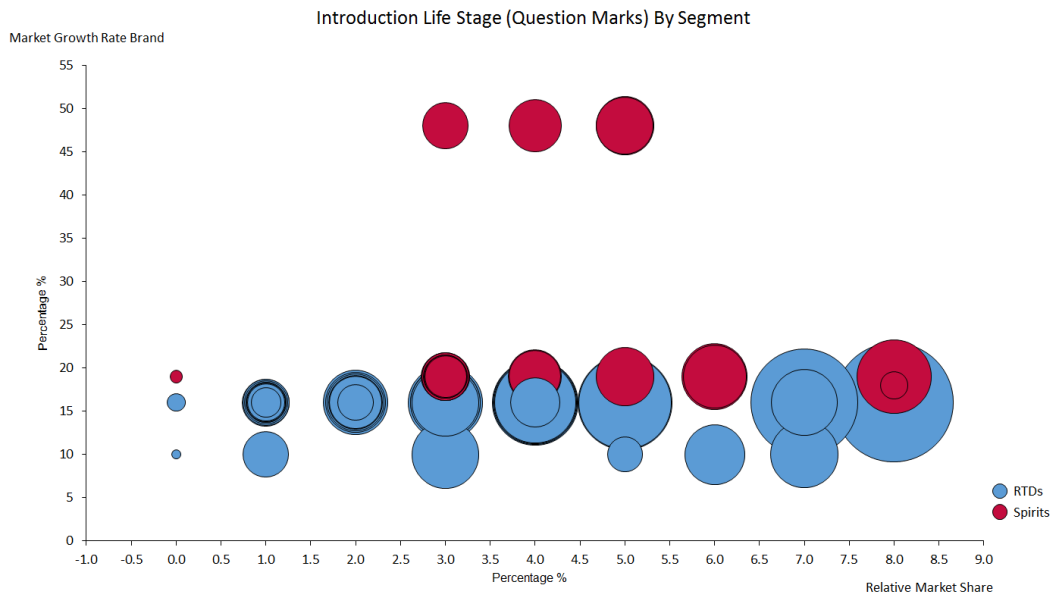
FIGURE A.10: *Introduction life cycle stage by segment*
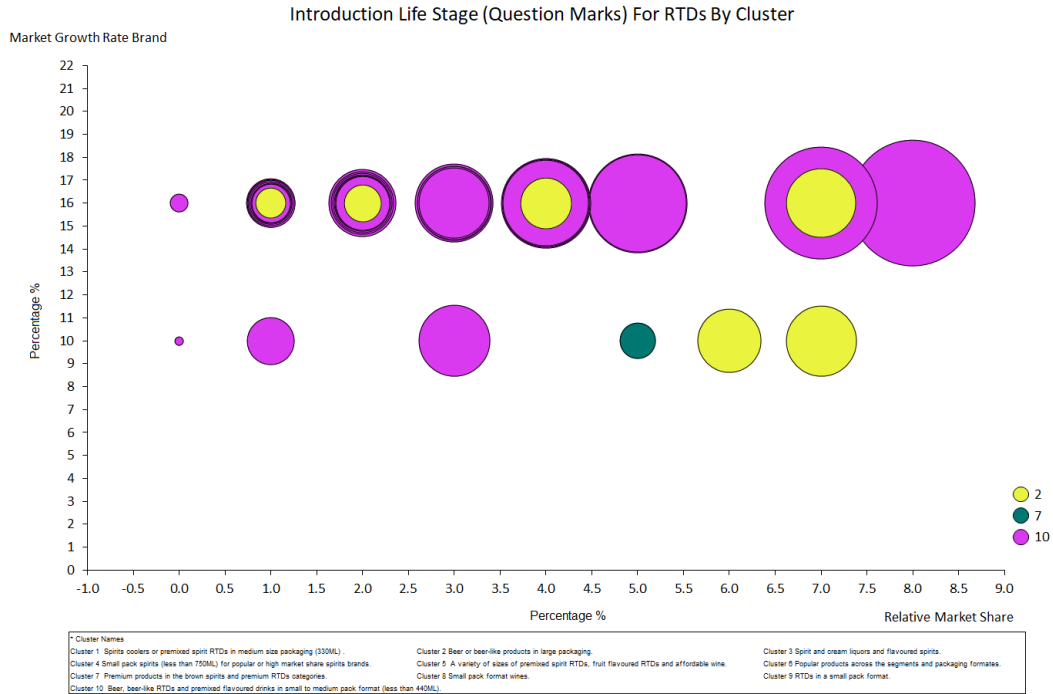
FIGURE A.11: *Introduction life cycle stage for RTDs by cluster*
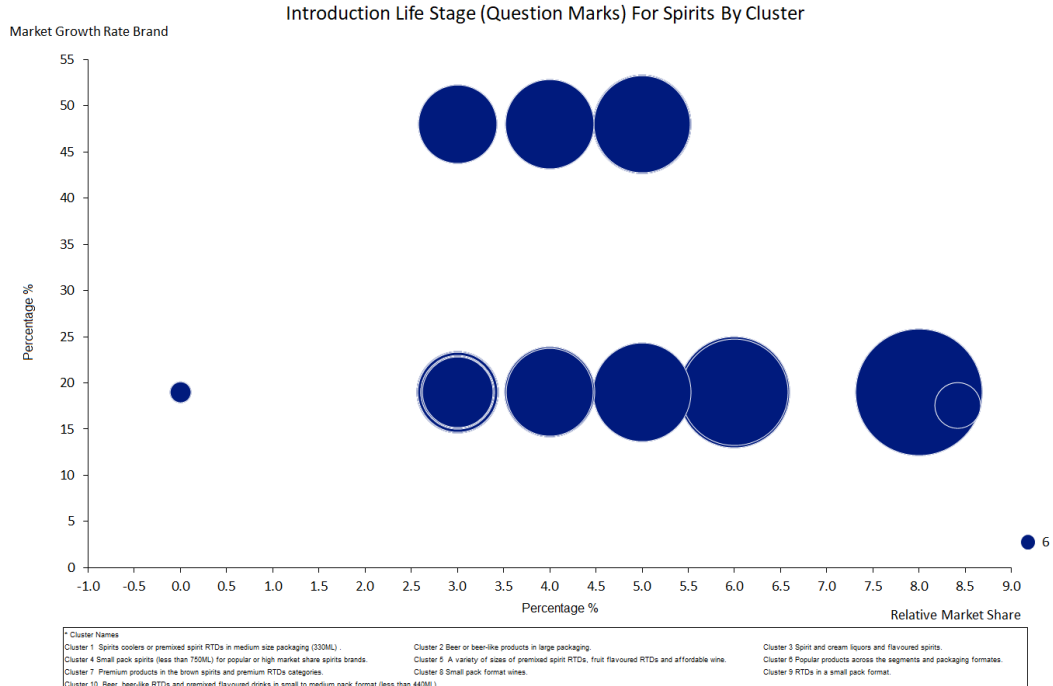
FIGURE A.12: *Introduction life cycle stage for spirits by cluster*

In Figure A.13 the results for the growth life cycle stage or Stars are displayed by cluster and segment respectively. In Figures A.14, A.15 and A.16 the results are displayed per segment.
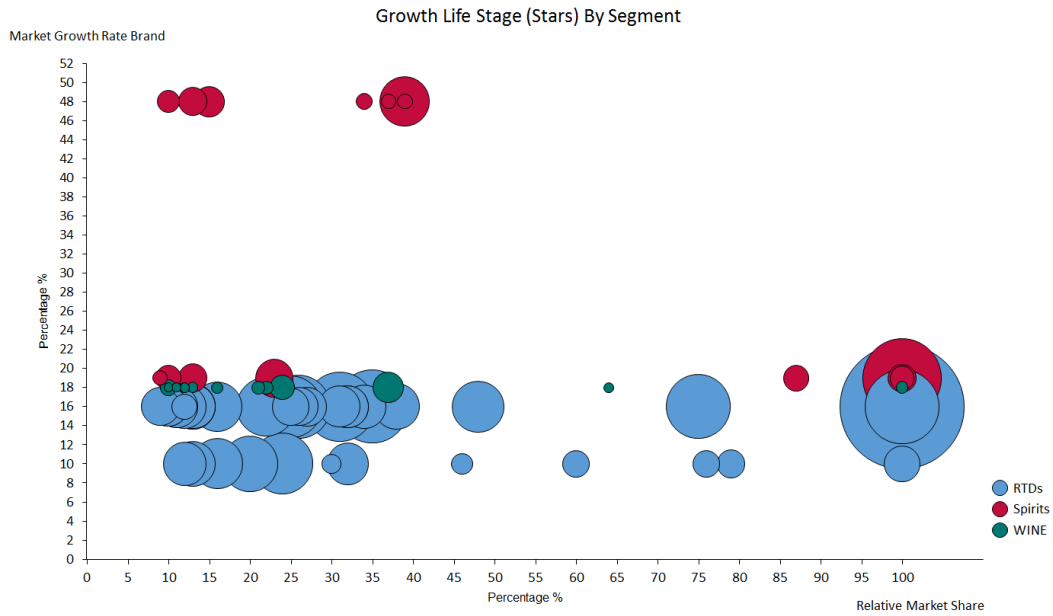
FIGURE A.13: *Growth life cycle stage by segment*
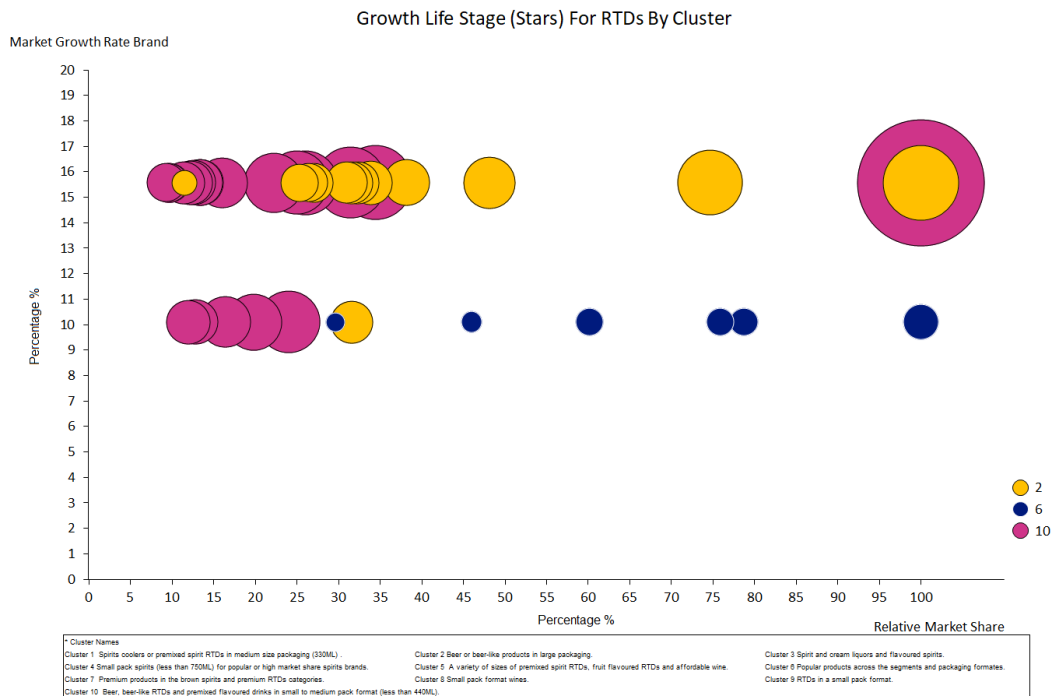
FIGURE A.14: *Growth life cycle stage for RTDs by cluster*
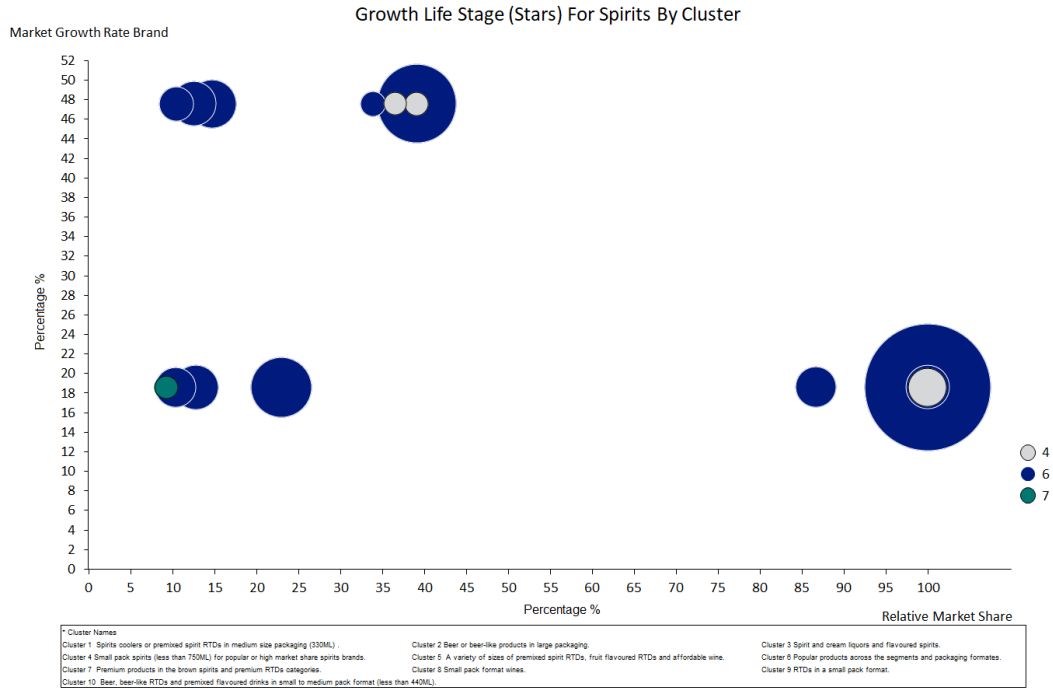
FIGURE A.15: *Growth life cycle stage for spirits by cluster*
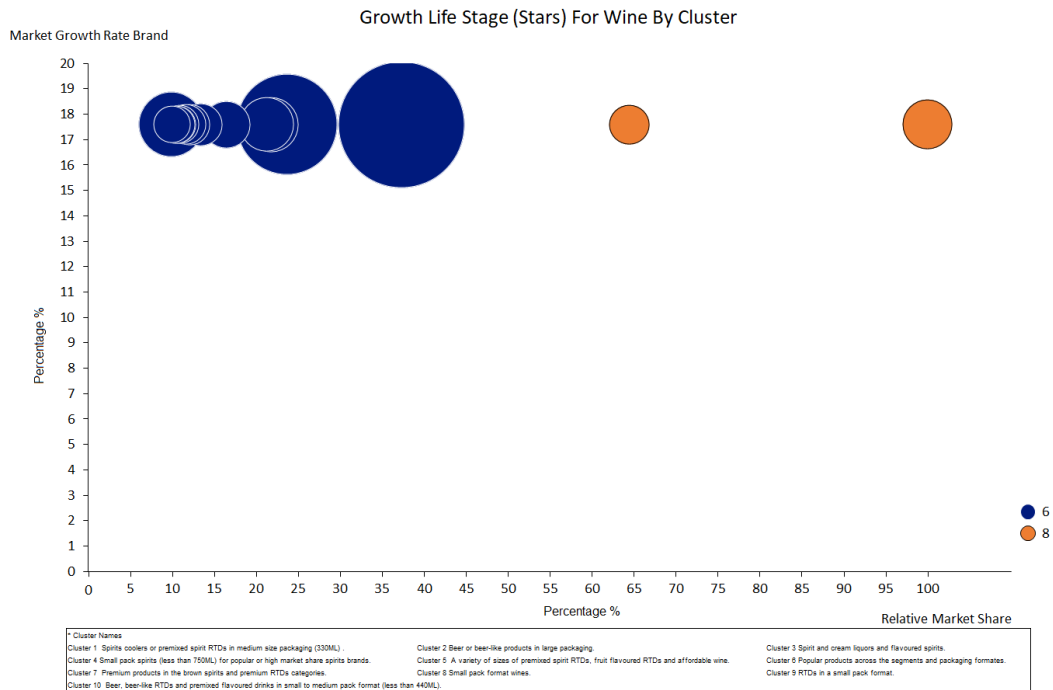
FIGURE A.16: *Growth life cycle stage for wine by cluster*

In Figure A.17 the results are displayed for the products in the maturity life cycle stage. The maturity life cycle stage has a much higher frequency of products grouped than the previous two. In Figures A.18, A.19 and A.20 the results are displayed per segment.
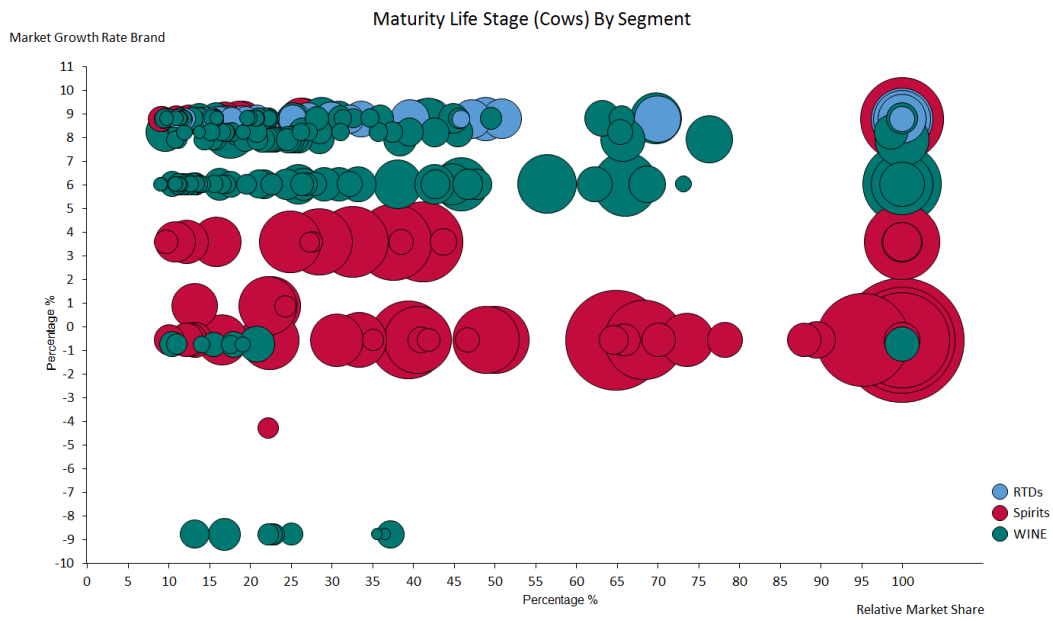
FIGURE A.17: *Maturity life cycle stage by segment*
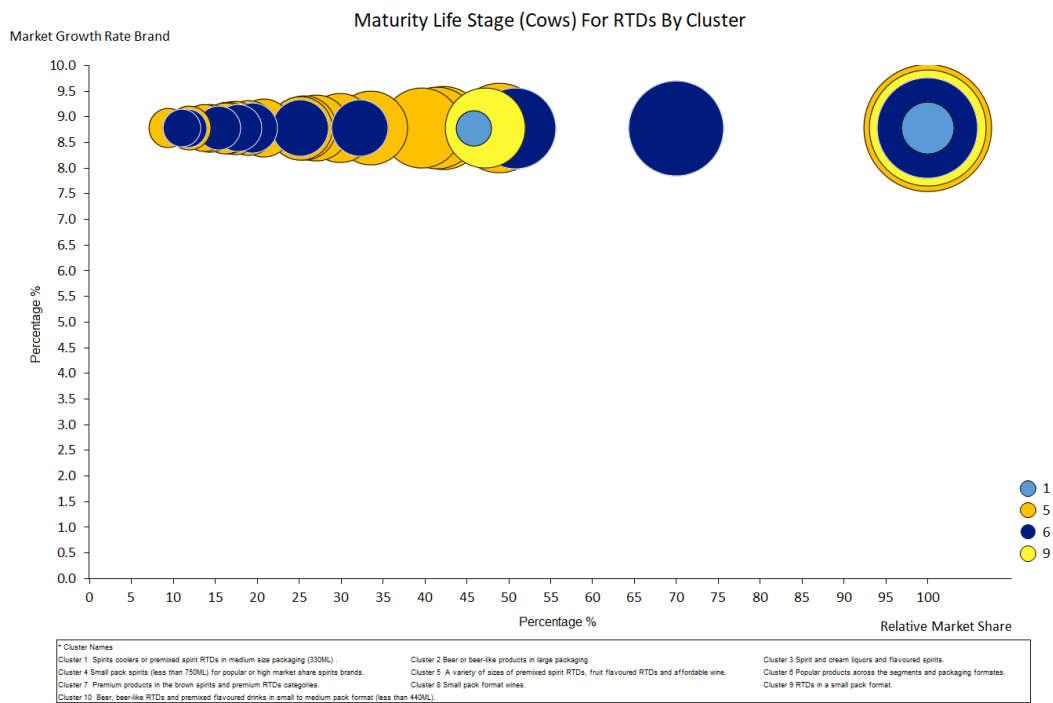
FIGURE A.18: *Maturity life cycle stage for RTDs by cluster*
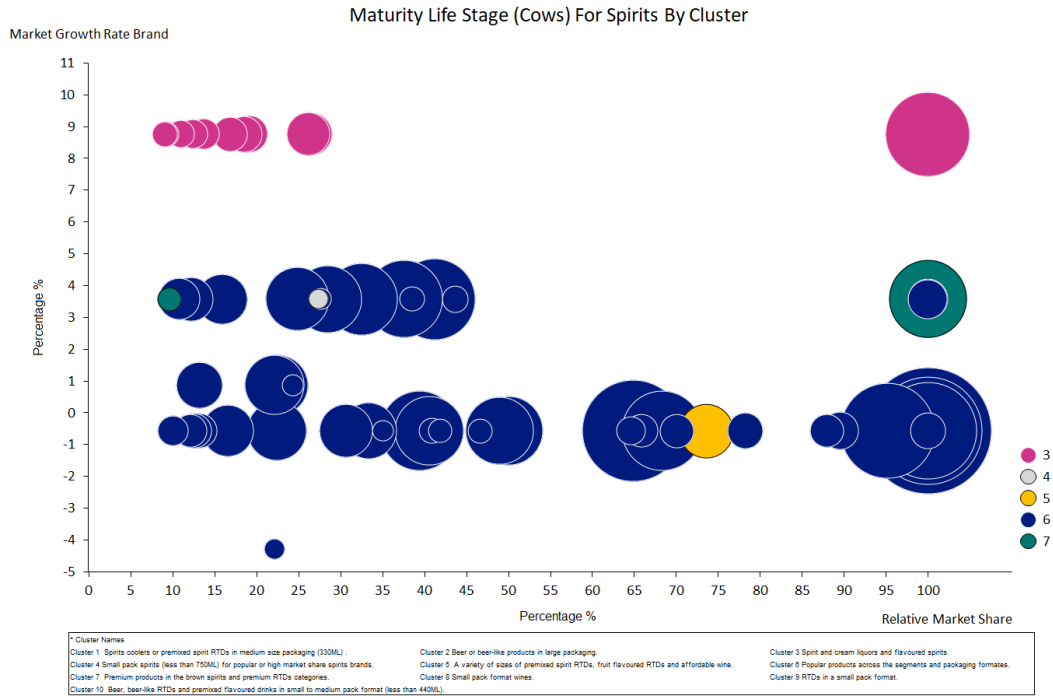
FIGURE A.19: *Maturity life cycle stage for spirits by cluster*
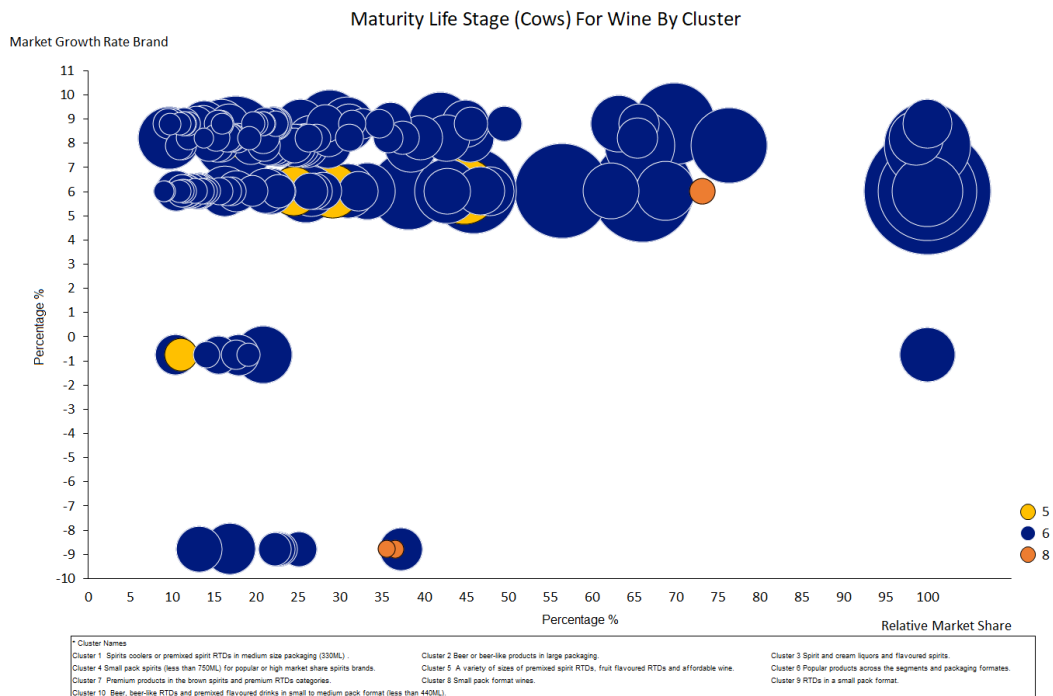
FIGURE A.20: *Maturity life cycle stage for wine by cluster*

In Figure A.21 the results for the products in the decline life cycle stage are displayed. As with the maturity phase, there are a large number of products grouped together in the decline phase. In Figures A.22, A.23 and A.24 the results are displayed per segment.
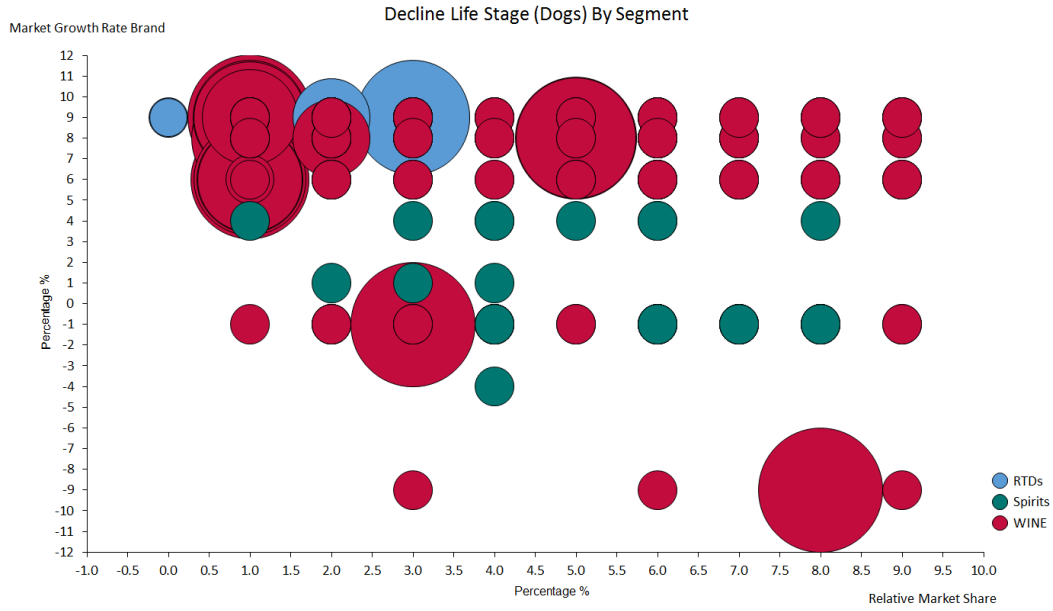
FIGURE A.21: *Decline life cycle stage by segment*
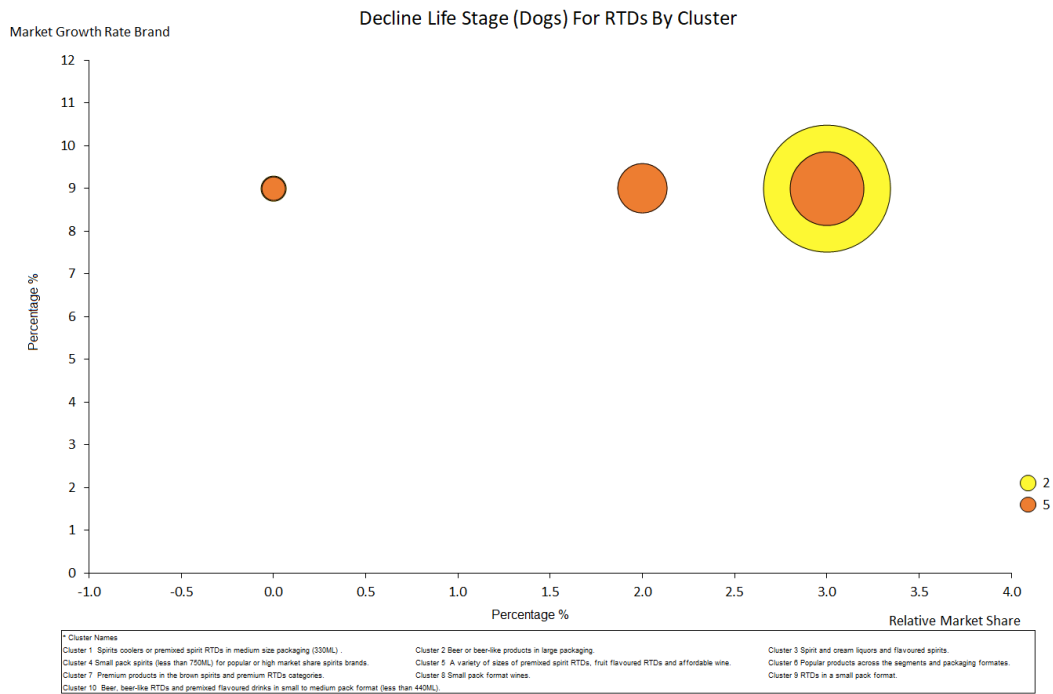
FIGURE A.22: *Decline life cycle stage for RTDs by cluster*
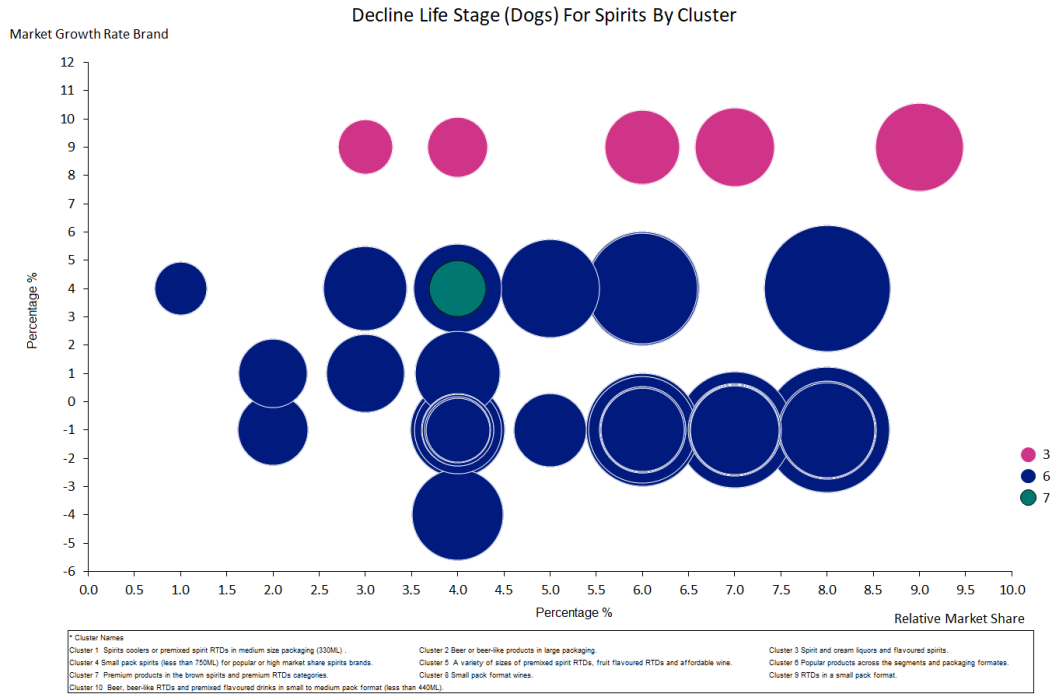
FIGURE A.23: *Decline life cycle stage for spirits by cluster*
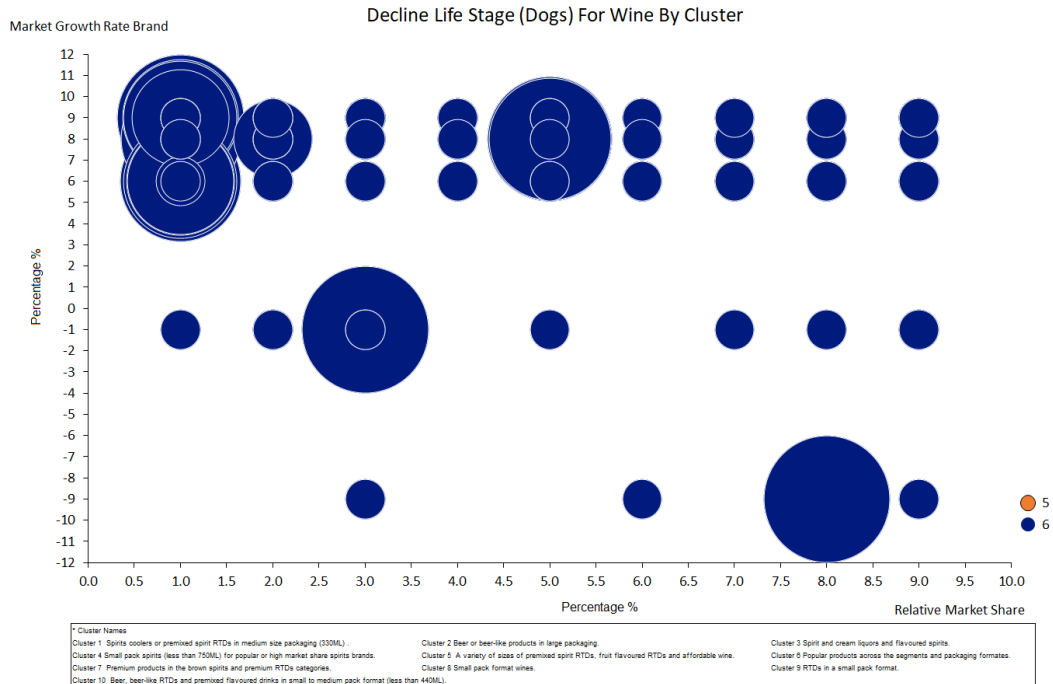
FIGURE A.24: *Decline life cycle stage for wine by cluster*

A.4.2   The product life cycle

The results from the PLC methodology and from the industry has created many categories that each product can be allocated to. The alcoholic beverage industry breaks products down into segments and then into categories. The results from the factor analysis produced factor scores, each product can be allocated to a factor by using the highest absolute factor score as a indicator. Cluster analysis divides the products into clusters and finally the BCG growth share matrix separates the products into four quadrants. Each product can also be described by the product life cycle stage that it has been allocated to. The tables below show each product and which category it belongs to.

			Factor, Cluster, Sub-Cluster and Taxonomy Results			
Product	Segment	Category	Factor	Cluster	Sub-Cluster	Taxonomy
RTD-FABS-P30-330ML	RTDs	FABS & CIDER	Factor 8	1		Maturity (Cows)
RTD-FABS-P27-330ML	RTDs	FABS & CIDER	Factor 8	1		Maturity (Cows)
RTD-BEER-P45-500ML	RTDs	BEER	Factor 5	2		Growth (Stars)
RTD-BEER-P58-750ML	RTDs	BEER	Factor 5	2		Growth (Stars)
RTD-BEER-P55-750ML	RTDs	BEER	Factor 5	2		Growth (Stars)
RTD-BEER-P44-500ML	RTDs	BEER	Factor 5	2		Growth (Stars)
RTD-BEER-P43-500ML	RTDs	BEER	Factor 5	2		Growth (Stars)
RTD-BEER-P46-500ML	RTDs	BEER	Factor 1	2		Growth (Stars)
RTD-CIDER-P16-500ML	RTDs	FABS & CIDER	Factor 5	2		Growth (Stars)
RTD-BEER-P49-650ML	RTDs	BEER	Factor 5	2		Growth (Stars)
RTD-BEER-P56-750ML	RTDs	BEER	Factor 1	2		Growth (Stars)
RTD-BEER-P47-500ML	RTDs	BEER	Factor 1	2		Growth (Stars)
RTD-BEER-P57-750ML	RTDs	BEER	Factor 1	2		Growth (Stars)
RTD-BEER-P50-660ML	RTDs	BEER	Factor 5	2		Growth (Stars)
RTD-BEER-P59-750ML	RTDs	BEER	Factor 8	2		Growth (Stars)
RTD-CIDER-P18-660ML	RTDs	FABS & CIDER	Factor 5	2		Introduction (Question Marks)
RTD-BEER-P51-660ML	RTDs	BEER	Factor 5	2		Introduction (Question Marks)
RTD-CIDER-P17-660ML	RTDs	FABS & CIDER	Factor 5	2		Introduction (Question Marks)
RTD-BEER-P48-5LTR	RTDs	BEER	Factor 5	2		Introduction (Question Marks)
RTD-FABS-P37-660ML	RTDs	FABS & CIDER	Factor 1	2		Decline (Dogs)
RTD-BEER-P52-660ML	RTDs	BEER	Factor 5	2		Introduction (Question Marks)
RTD-BEER-P53-660ML	RTDs	BEER	Factor 7	2		Introduction (Question Marks)
RTD-FABS-P38-660ML	RTDs	FABS & CIDER	Factor 6	2		Decline (Dogs)
SPIRITS-LIQUEURS-P16-750ML	SPIRITS	LIQUEURS	Factor 8	3		Maturity (Cows)
SPIRITS-LIQUEURS-P9-750ML	SPIRITS	LIQUEURS	Factor 6	3		Maturity (Cows)
SPIRITS-LIQUEURS-P2-1LTR	SPIRITS	LIQUEURS	Factor 8	3		Maturity (Cows)
SPIRITS-LIQUEURS-P5-750ML	SPIRITS	LIQUEURS	Factor 7	3		Maturity (Cows)
SPIRITS-LIQUEURS-P6-750ML	SPIRITS	LIQUEURS	Factor 6	3		Maturity (Cows)
SPIRITS-LIQUEURS-P3-750ML	SPIRITS	LIQUEURS	Factor 6	3		Maturity (Cows)
SPIRITS-LIQUEURS-P7-750ML	SPIRITS	LIQUEURS	Factor 6	3		Maturity (Cows)
SPIRITS-LIQUEURS-P15-750ML	SPIRITS	LIQUEURS	Factor 6	3		Maturity (Cows)
SPIRITS-LIQUEURS-P8-750ML	SPIRITS	LIQUEURS	Factor 6	3		Maturity (Cows)
SPIRITS-LIQUEURS-P12-750ML	SPIRITS	LIQUEURS	Factor 6	3		Maturity (Cows)
SPIRITS-LIQUEURS-P4-750ML	SPIRITS	LIQUEURS	Factor 6	3		Maturity (Cows)
SPIRITS-LIQUEURS-P11-750ML	SPIRITS	LIQUEURS	Factor 6	3		Decline (Dogs)
SPIRITS-LIQUEURS-P13-750ML	SPIRITS	LIQUEURS	Factor 6	3		Decline (Dogs)
SPIRITS-LIQUEURS-P10-750ML	SPIRITS	LIQUEURS	Factor 6	3		Decline (Dogs)
SPIRITS-LIQUEURS-P14-750ML	SPIRITS	LIQUEURS	Factor 6	3		Decline (Dogs)
SPIRITS-LIQUEURS-P1-1LTR	SPIRITS	LIQUEURS	Factor 6	3		Decline (Dogs)
SPIRITS-VODKA-P3-200ML	SPIRITS	WHITE SPIRITS	Factor 8	4		Growth (Stars)
SPIRITS-GIN-P2-200ML	SPIRITS	WHITE SPIRITS	Factor 1	4		Growth (Stars)
SPIRITS-GIN-P3-375ML	SPIRITS	WHITE SPIRITS	Factor 1	4		Growth (Stars)
SPIRITS-BRANDY-P5-375ML	SPIRITS	BRANDY	Factor 6	4		Maturity (Cows)
RTD-FABS-P33-440ML	RTDs	FABS & CIDER	Factor 2	5		Maturity (Cows)
SPIRITS-WHISKY-P11-750ML	SPIRITS	WHISKY	Factor 4	5		Maturity (Cows)
RTD-FABS-P1-250ML	RTDs	FABS & CIDER	Factor 7	5		Maturity (Cows)
WINE-BIB-P84-5LTR	WINE	WINE	Factor 6	5		Maturity (Cows)
RTD-FABS-P23-300ML	RTDs	FABS & CIDER	Factor 2	5		Maturity (Cows)

TABLE A.42: *Part 1: Results for determining a competitor set by clustering according to product life cycle stage displaying factors, clusters and taxonomy*

		Factor, Cluster, Sub-Cluster and Taxonomy Results				
Product	Segment	Category	Factor	Cluster	Sub-Cluster	Taxonomy
RTD-FABS-P17-275ML	RTDs	FABS & CIDER	Factor 8	5		Maturity (Cows)
RTD-FABS-P24-300ML	RTDs	FABS & CIDER	Factor 4	5		Maturity (Cows)
RTD-FABS-P14-275ML	RTDs	FABS & CIDER	Factor 1	5		Maturity (Cows)
RTD-FABS-P32-440ML	RTDs	FABS & CIDER	Factor 1	5		Maturity (Cows)
WINE-BIB-P52-3LTR	WINE	WINE	Factor 4	5		Maturity (Cows)
RTD-FABS-P19-275ML	RTDs	FABS & CIDER	Factor 8	5		Maturity (Cows)
RTD-FABS-P29-330ML	RTDs	FABS & CIDER	Factor 4	5		Maturity (Cows)
RTD-FABS-P2-250ML	RTDs	FABS & CIDER	Factor 4	5		Maturity (Cows)
WINE-BIB-P51-3LTR	WINE	WINE	Factor 4	5		Maturity (Cows)
RTD-FABS-P16-275ML	RTDs	FABS & CIDER	Factor 8	5		Maturity (Cows)
WINE-WHITE-P68-750ML	WINE	WINE	Factor 6	5		Maturity (Cows)
RTD-FABS-P9-275ML	RTDs	FABS & CIDER	Factor 6	5		Maturity (Cows)
WINE-WHITE-P66-750ML	WINE	WINE	Factor 4	5		Maturity (Cows)
RTD-FABS-P3-250ML	RTDs	FABS & CIDER	Factor 4	5		Maturity (Cows)
RTD-FABS-P10-275ML	RTDs	FABS & CIDER	Factor 6	5		Maturity (Cows)
RTD-FABS-P12-275ML	RTDs	FABS & CIDER	Factor 6	5		Maturity (Cows)
RTD-FABS-P28-330ML	RTDs	FABS & CIDER	Factor 1	5		Maturity (Cows)
RTD-FABS-P13-275ML	RTDs	FABS & CIDER	Factor 6	5		Maturity (Cows)
RTD-FABS-P18-275ML	RTDs	FABS & CIDER	Factor 8	5		Maturity (Cows)
WINE-ROSE-P4-1.5LTR	WINE	WINE	Factor 6	5		Maturity (Cows)
RTD-FABS-P31-440ML	RTDs	FABS & CIDER	Factor 6	5		Maturity (Cows)
WINE-RED-P105-750ML	WINE	WINE	Factor 4	5		Decline (Dogs)
RTD-FABS-P8-275ML	RTDs	FABS & CIDER	Factor 6	5		Decline (Dogs)
RTD-FABS-P4-250ML	RTDs	FABS & CIDER	Factor 8	5		Decline (Dogs)
RTD-FABS-P34-440ML	RTDs	FABS & CIDER	Factor 4	5		Decline (Dogs)
SPIRITS-BRANDY-P12-750ML	SPIRITS	BRANDY	Factor 3	7		Maturity (Cows)
SPIRITS-BRANDY-P24-750ML	SPIRITS	BRANDY	Factor 8	7		Maturity (Cows)
SPIRITS-VODKA-P6-750ML	SPIRITS	WHITE SPIRITS	Factor 1	7		Growth (Stars)
RTD-CIDER-P5-330ML	RTDs	FABS & CIDER	Factor 4	7		Introduction (Question Marks)
SPIRITS-BRANDY-P14-750ML	SPIRITS	BRANDY	Factor 3	7		Decline (Dogs)
WINE-FORTIFIED-P6-375ML	WINE	FORTIFIED	Factor 7	8		Growth (Stars)
WINE-BIB-P53-500ML	WINE	WINE	Factor 8	8		Maturity (Cows)
WINE-FORTIFIED-P3-200ML	WINE	FORTIFIED	Factor 7	8		Growth (Stars)
WINE-PERLE-P5-300ML	WINE	PERLE	Factor 1	8		Maturity (Cows)
WINE-PERLE-P6-300ML	WINE	PERLE	Factor 1	8		Maturity (Cows)
RTD-FABS-P21-275ML	RTDs	FABS & CIDER	Factor 8	9		Maturity (Cows)
RTD-FABS-P20-275ML	RTDs	FABS & CIDER	Factor 8	9		Maturity (Cows)
RTD-BEER-P24-340ML	RTDs	BEER	Factor 5	10		Growth (Stars)
RTD-BEER-P42-440ML	RTDs	BEER	Factor 5	10		Growth (Stars)
RTD-BEER-P17-330ML	RTDs	BEER	Factor 5	10		Growth (Stars)
RTD-BEER-P20-340ML	RTDs	BEER	Factor 5	10		Growth (Stars)
RTD-BEER-P3-330ML	RTDs	BEER	Factor 5	10		Growth (Stars)
RTD-CIDER-P6-330ML	RTDs	FABS & CIDER	Factor 5	10		Growth (Stars)
RTD-BEER-P21-340ML	RTDs	BEER	Factor 5	10		Growth (Stars)
RTD-CIDER-P9-330ML	RTDs	FABS & CIDER	Factor 5	10		Growth (Stars)
RTD-CIDER-P1-330ML	RTDs	FABS & CIDER	Factor 5	10		Growth (Stars)
RTD-BEER-P16-330ML	RTDs	BEER	Factor 1	10		Growth (Stars)
RTD-BEER-P40-440ML	RTDs	BEER	Factor 5	10		Growth (Stars)

TABLE A.43: *Part 2: Results for determining a competitor set by clustering according to product life cycle stage displaying factors, clusters and taxonomy*

		Factor, Cluster, Sub-Cluster and Taxonomy Results				
Product	Segment	Category	Factor	Cluster	Sub-Cluster	Taxonomy
RTD-BEER-P19-330ML	RTDs	BEER	Factor 5	10		Growth (Stars)
RTD-CIDER-P15-440ML	RTDs	FABS & CIDER	Factor 5	10		Growth (Stars)
RTD-BEER-P14-330ML	RTDs	BEER	Factor 8	10		Growth (Stars)
RTD-BEER-P25-355ML	RTDs	BEER	Factor 4	10		Growth (Stars)
RTD-CIDER-P11-440ML	RTDs	FABS & CIDER	Factor 5	10		Growth (Stars)
RTD-BEER-P10-330ML	RTDs	BEER	Factor 3	10		Growth (Stars)
RTD-BEER-P35-440ML	RTDs	BEER	Factor 8	10		Growth (Stars)
RTD-BEER-P1-330ML	RTDs	BEER	Factor 5	10		Growth (Stars)
RTD-BEER-P15-330ML	RTDs	BEER	Factor 3	10		Growth (Stars)
RTD-BEER-P31-440ML	RTDs	BEER	Factor 5	10		Introduction (Question Marks)
RTD-BEER-P2-330ML	RTDs	BEER	Factor 5	10		Introduction (Question Marks)
RTD-BEER-P4-330ML	RTDs	BEER	Factor 3	10		Introduction (Question Marks)
RTD-BEER-P27-440ML	RTDs	BEER	Factor 5	10		Introduction (Question Marks)
RTD-BEER-P39-440ML	RTDs	BEER	Factor 8	10		Introduction (Question Marks)
RTD-BEER-P29-440ML	RTDs	BEER	Factor 3	10		Introduction (Question Marks)
RTD-BEER-P41-440ML	RTDs	BEER	Factor 3	10		Introduction (Question Marks)
RTD-BEER-P18-330ML	RTDs	BEER	Factor 3	10		Introduction (Question Marks)
RTD-BEER-P5-330ML	RTDs	BEER	Factor 1	10		Introduction (Question Marks)
RTD-BEER-P9-330ML	RTDs	BEER	Factor 1	10		Introduction (Question Marks)
RTD-BEER-P23-340ML	RTDs	BEER	Factor 8	10		Introduction (Question Marks)
RTD-CIDER-P8-330ML	RTDs	FABS & CIDER	Factor 1	10		Introduction (Question Marks)
RTD-BEER-P33-440ML	RTDs	BEER	Factor 1	10		Introduction (Question Marks)
RTD-BEER-P13-330ML	RTDs	BEER	Factor 8	10		Introduction (Question Marks)
RTD-BEER-P28-440ML	RTDs	BEER	Factor 3	10		Introduction (Question Marks)
RTD-BEER-P7-330ML	RTDs	BEER	Factor 1	10		Introduction (Question Marks)
RTD-BEER-P12-330ML	RTDs	BEER	Factor 1	10		Introduction (Question Marks)
RTD-BEER-P22-340ML	RTDs	BEER	Factor 3	10		Introduction (Question Marks)
RTD-BEER-P34-440ML	RTDs	BEER	Factor 3	10		Introduction (Question Marks)
RTD-BEER-P4-330ML	RTDs	BEER	Factor 3	10		Introduction (Question Marks)
RTD-BEER-P32-440ML	RTDs	BEER	Factor 1	10		Introduction (Question Marks)
RTD-CIDER-P7-330ML	RTDs	FABS & CIDER	Factor 1	10		Introduction (Question Marks)
RTD-BEER-P26-440ML	RTDs	BEER	Factor 5	10		Introduction (Question Marks)
RTD-BEER-P37-440ML	RTDs	BEER	Factor 1	10		Introduction (Question Marks)
RTD-BEER-P30-440ML	RTDs	BEER	Factor 1	10		Introduction (Question Marks)
RTD-BEER-P36-440ML	RTDs	BEER	Factor 1	10		Introduction (Question Marks)
RTD-BEER-P8-330ML	RTDs	BEER	Factor 1	10		Introduction (Question Marks)
RTD-BEER-P6-330ML	RTDs	BEER	Factor 1	10		Introduction (Question Marks)
RTD-BEER-P38-440ML	RTDs	BEER	Factor 3	10		Introduction (Question Marks)
RTD-CIDER-P12-440ML	RTDs	FABS & CIDER	Factor 5	10		Introduction (Question Marks)
SPIRITS-WHISKY-P2-1LTR	SPIRITS	WHISKY	Factor 6	6	1	Maturity (Cows)
SPIRITS-WHISKY-P9-1LTR	SPIRITS	WHISKY	Factor 6	6	1	Maturity (Cows)
SPIRITS-WHISKY-P3-1LTR	SPIRITS	WHISKY	Factor 8	6	1	Maturity (Cows)
SPIRITS-WHISKY-P6-1LTR	SPIRITS	WHISKY	Factor 6	6	1	Maturity (Cows)
SPIRITS-WHISKY-P7-1LTR	SPIRITS	WHISKY	Factor 6	6	1	Maturity (Cows)
SPIRITS-WHISKY-P5-1LTR	SPIRITS	WHISKY	Factor 6	6	1	Maturity (Cows)
WINE-ROSE-P17-750ML	WINE	WINE	Factor 7	6	2	Maturity (Cows)
WINE-WHITE-P52-750ML	WINE	WINE	Factor 7	6	2	Maturity (Cows)
WINE-WHITE-P57-750ML	WINE	WINE	Factor 7	6	2	Maturity (Cows)

TABLE A.44: *Part 3: Results for determining a competitor set by clustering according to product life cycle stage displaying factors, clusters and taxonomy*

Factor, Cluster, Sub-Cluster and Taxonomy Results						
Product	Segment	Category	Factor	Cluster	Sub-Cluster	Taxonomy
WINE-RED-P88-750ML	WINE	WINE	Factor 7	6	2	Maturity (Cows)
WINE-WHITE-P59-750ML	WINE	WINE	Factor 7	6	2	Maturity (Cows)
WINE-WHITE-P58-750ML	WINE	WINE	Factor 7	6	2	Maturity (Cows)
WINE-RED-P90-750ML	WINE	WINE	Factor 7	6	2	Maturity (Cows)
WINE-RED-P92-750ML	WINE	WINE	Factor 7	6	2	Maturity (Cows)
WINE-RED-P7-750ML	WINE	WINE	Factor 7	6	2	Maturity (Cows)
WINE-RED-P5-750ML	WINE	WINE	Factor 7	6	2	Maturity (Cows)
WINE-RED-P6-750ML	WINE	WINE	Factor 7	6	2	Maturity (Cows)
WINE-WHITE-P3-750ML	WINE	WINE	Factor 7	6	2	Maturity (Cows)
WINE-ROSE-P16-750ML	WINE	WINE	Factor 7	6	2	Maturity (Cows)
WINE-RED-P91-750ML	WINE	WINE	Factor 7	6	2	Maturity (Cows)
WINE-WHITE-P4-750ML	WINE	WINE	Factor 7	6	2	Maturity (Cows)
WINE-BIB-P82-5LTR	WINE	WINE	Factor 5	6	3	Maturity (Cows)
WINE-BIB-P79-5LTR	WINE	WINE	Factor 7	6	3	Maturity (Cows)
WINE-BIB-P44-3LTR	WINE	WINE	Factor 8	6	3	Maturity (Cows)
WINE-BIB-P78-5LTR	WINE	WINE	Factor 7	6	3	Maturity (Cows)
WINE-BIB-P43-3LTR	WINE	WINE	Factor 7	6	3	Maturity (Cows)
WINE-FORTIFIED-P16-750ML	WINE	FORTIFIED	Factor 7	6	3	Growth (Stars)
WINE-RED-P85-750ML	WINE	WINE	Factor 6	6	3	Maturity (Cows)
WINE-BIB-P42-3LTR	WINE	WINE	Factor 7	6	3	Maturity (Cows)
WINE-FORTIFIED-P2-1LTR	WINE	FORTIFIED	Factor 7	6	3	Growth (Stars)
WINE-WHITE-P2-1.5LTR	WINE	WINE	Factor 8	6	3	Maturity (Cows)
WINE-PERLE-P4-2LTR	WINE	PERLE	Factor 1	6	3	Maturity (Cows)
WINE-BIB-P41-3LTR	WINE	WINE	Factor 7	6	3	Maturity (Cows)
WINE-RED-P4-1.5LTR	WINE	WINE	Factor 8	6	3	Maturity (Cows)
WINE-RED-P72-750ML	WINE	WINE	Factor 6	6	3	Maturity (Cows)
WINE-PERLE-P3-1LTR	WINE	PERLE	Factor 1	6	3	Maturity (Cows)
WINE-BIB-P63-5LTR	WINE	WINE	Factor 1	6	3	Maturity (Cows)
WINE-ROSE-P19-750ML	WINE	WINE	Factor 8	6	3	Maturity (Cows)
WINE-FORTIFIED-P5-2LTR	WINE	FORTIFIED	Factor 7	6	3	Growth (Stars)
WINE-ROSE-P12-750ML	WINE	WINE	Factor 7	6	3	Decline (Dogs)
WINE-BIB-P82-1.5LTR	WINE	WINE	Factor 7	6	3	Decline (Dogs)
WINE-ROSE-P5-750ML	WINE	WINE	Factor 8	6	3	Decline (Dogs)
WINE-BIB-P69-5LTR	WINE	WINE	Factor 1	6	3	Decline (Dogs)
WINE-WHITE-P2-5LTR	WINE	WINE	Factor 8	6	3	Decline (Dogs)
WINE-BIB-P50-3LTR	WINE	WINE	Factor 8	6	3	Decline (Dogs)
WINE-RED-P93-750ML	WINE	WINE	Factor 8	6	3	Decline (Dogs)
WINE-ROSE-P1-1.5LTR	WINE	WINE	Factor 8	6	3	Decline (Dogs)
WINE-BIB-P74-5LTR	WINE	WINE	Factor 1	6	3	Decline (Dogs)
WINE-RED-P3-1.5LTR	WINE	WINE	Factor 8	6	3	Decline (Dogs)
WINE-RED-P71-750ML	WINE	WINE	Factor 7	6	3	Decline (Dogs)
WINE-BIB-P80-5LTR	WINE	WINE	Factor 6	6	3	Decline (Dogs)
WINE-BIB-P49-3LTR	WINE	WINE	Factor 8	6	3	Decline (Dogs)
WINE-PERLE-P13-750ML	WINE	PERLE	Factor 1	6	3	Decline (Dogs)
WINE-WHITE-P60-750ML	WINE	WINE	Factor 8	6	3	Decline (Dogs)
WINE-WHITE-P47-750ML	WINE	WINE	Factor 6	6	3	Decline (Dogs)
WINE-BIB-P81-5LTR	WINE	WINE	Factor 6	6	3	Decline (Dogs)

TABLE A.45: *Part 4: Results for determining a competitor set by clustering according to product life cycle stage displaying factors, clusters and taxonomy*

		Factor, Cluster, Sub-Cluster and Taxonomy Results				
Product	Segment	Category	Factor	Cluster	Sub-Cluster	Taxonomy
WINE-PERLE-P1-1LTR	WINE	PERLE	Factor 1	6	3	Decline (Dogs)
WINE-RED-P95-750ML	WINE	WINE	Factor 6	6	3	Decline (Dogs)
WINE-ROSE-P18-750ML	WINE	WINE	Factor 8	6	3	Decline (Dogs)
WINE-RED-P79-1.5LTR	WINE	WINE	Factor 7	6	3	Decline (Dogs)
WINE-ROSE-P13-750ML	WINE	WINE	Factor 6	6	3	Decline (Dogs)
SPIRITS-WHISKY-P21-750ML	SPIRITS	WHISKY	Factor 6	6	4	Maturity (Cows)
SPIRITS-WHISKY-P22-750ML	SPIRITS	WHISKY	Factor 8	6	4	Maturity (Cows)
SPIRITS-BRANDY-P23-750ML	SPIRITS	BRANDY	Factor 5	6	4	Maturity (Cows)
SPIRITS-WHISKY-P16-750ML	SPIRITS	WHISKY	Factor 8	6	4	Maturity (Cows)
SPIRITS-BRANDY-P7-750ML	SPIRITS	BRANDY	Factor 6	6	4	Maturity (Cows)
SPIRITS-BRANDY-P9-750ML	SPIRITS	BRANDY	Factor 6	6	4	Maturity (Cows)
SPIRITS-BRANDY-P16-750ML	SPIRITS	BRANDY	Factor 4	6	4	Maturity (Cows)
SPIRITS-BRANDY-P10-750ML	SPIRITS	BRANDY	Factor 7	6	4	Maturity (Cows)
SPIRITS-WHISKY-P29-750ML	SPIRITS	WHISKY	Factor 6	6	4	Maturity (Cows)
SPIRITS-RUM-P3-750ML	SPIRITS	RUM	Factor 4	6	4	Maturity (Cows)
SPIRITS-WHISKY-P30-750ML	SPIRITS	WHISKY	Factor 6	6	4	Maturity (Cows)
SPIRITS-BRANDY-P22-750ML	SPIRITS	BRANDY	Factor 7	6	4	Maturity (Cows)
SPIRITS-BRANDY-P11-750ML	SPIRITS	BRANDY	Factor 8	6	4	Maturity (Cows)
SPIRITS-BRANDY-P13-750ML	SPIRITS	BRANDY	Factor 8	6	4	Maturity (Cows)
SPIRITS-BRANDY-P19-750ML	SPIRITS	BRANDY	Factor 4	6	4	Decline (Dogs)
SPIRITS-WHISKY-P32-750ML	SPIRITS	WHISKY	Factor 6	6	4	Decline (Dogs)
SPIRITS-WHISKY-P27-750ML	SPIRITS	WHISKY	Factor 4	6	4	Decline (Dogs)
SPIRITS-BRANDY-P15-750ML	SPIRITS	BRANDY	Factor 4	6	4	Decline (Dogs)
SPIRITS-WHISKY-P50-750ML	SPIRITS	WHISKY	Factor 6	6	4	Decline (Dogs)
SPIRITS-BRANDY-P18-750ML	SPIRITS	BRANDY	Factor 4	6	4	Decline (Dogs)
SPIRITS-WHISKY-P28-750ML	SPIRITS	WHISKY	Factor 6	6	4	Decline (Dogs)
SPIRITS-BRANDY-P6-750ML	SPIRITS	BRANDY	Factor 7	6	4	Decline (Dogs)
SPIRITS-WHISKY-P38-750ML	SPIRITS	WHISKY	Factor 6	6	4	Decline (Dogs)
SPIRITS-BRANDY-P17-750ML	SPIRITS	BRANDY	Factor 4	6	4	Decline (Dogs)
SPIRITS-WHISKY-P31-750ML	SPIRITS	WHISKY	Factor 6	6	4	Decline (Dogs)
SPIRITS-RUM-P5-750ML	SPIRITS	RUM	Factor 4	6	4	Decline (Dogs)
SPIRITS-BRANDY-P20-750ML	SPIRITS	BRANDY	Factor 4	6	4	Decline (Dogs)
SPIRITS-RUM-P6-750ML	SPIRITS	RUM	Factor 4	6	4	Decline (Dogs)
SPIRITS-WHISKY-P18-750ML	SPIRITS	WHISKY	Factor 8	6	4	Decline (Dogs)
SPIRITS-RUM-P4-750ML	SPIRITS	RUM	Factor 4	6	4	Decline (Dogs)
SPIRITS-BRANDY-P21-750ML	SPIRITS	BRANDY	Factor 4	6	4	Decline (Dogs)
SPIRITS-BRANDY-P3-1LTR	SPIRITS	BRANDY	Factor 7	6	5	Maturity (Cows)
SPIRITS-WHISKY-P4-1LTR	SPIRITS	WHISKY	Factor 8	6	5	Maturity (Cows)
SPIRITS-WHISKY-P8-1LTR	SPIRITS	WHISKY	Factor 6	6	5	Maturity (Cows)
SPIRITS-WHISKY-P1-1LTR	SPIRITS	WHISKY	Factor 8	6	5	Maturity (Cows)
SPIRITS-BRANDY-P2-1LTR	SPIRITS	BRANDY	Factor 6	6	5	Maturity (Cows)
SPIRITS-WHISKY-P10-1LTR	SPIRITS	WHISKY	Factor 2	6	5	Maturity (Cows)
SPIRITS-BRANDY-P4-1LTR	SPIRITS	BRANDY	Factor 7	6	5	Maturity (Cows)
SPIRITS-BRANDY-P1-1LTR	SPIRITS	BRANDY	Factor 6	6	5	Maturity (Cows)
SPIRITS-WHISKY-P25-750ML	SPIRITS	WHISKY	Factor 6	6	6	Maturity (Cows)
SPIRITS-WHISKY-P17-750ML	SPIRITS	WHISKY	Factor 7	6	6	Maturity (Cows)
SPIRITS-WHISKY-P14-750ML	SPIRITS	WHISKY	Factor 1	6	6	Maturity (Cows)

TABLE A.46: *Part 5: Results for determining a competitor set by clustering according to product life cycle stage displaying factors, clusters and taxonomy*

Factor, Cluster, Sub-Cluster and Taxonomy Results						
Product	Segment	Category	Factor	Cluster	Sub-Cluster	Taxonomy
SPIRITS-WHISKY-P15-750ML	SPIRITS	WHISKY	Factor 1	6	6	Maturity (Cows)
SPIRITS-WHISKY-P24-750ML	SPIRITS	WHISKY	Factor 8	6	6	Maturity (Cows)
SPIRITS-WHISKY-P19-750ML	SPIRITS	WHISKY	Factor 8	6	6	Maturity (Cows)
SPIRITS-WHISKY-P20-750ML	SPIRITS	WHISKY	Factor 6	6	6	Maturity (Cows)
SPIRITS-WHISKY-P35-750ML	SPIRITS	WHISKY	Factor 6	6	6	Maturity (Cows)
SPIRITS-WHISKY-P12-750ML	SPIRITS	WHISKY	Factor 8	6	6	Maturity (Cows)
SPIRITS-WHISKY-P40-750ML	SPIRITS	WHISKY	Factor 6	6	6	Maturity (Cows)
SPIRITS-WHISKY-P42-750ML	SPIRITS	WHISKY	Factor 8	6	6	Maturity (Cows)
SPIRITS-WHISKY-P41-750ML	SPIRITS	WHISKY	Factor 6	6	6	Decline (Dogs)
SPIRITS-WHISKY-P44-750ML	SPIRITS	WHISKY	Factor 6	6	6	Decline (Dogs)
SPIRITS-WHISKY-P33-750ML	SPIRITS	WHISKY	Factor 6	6	6	Decline (Dogs)
SPIRITS-WHISKY-P49-750ML	SPIRITS	WHISKY	Factor 6	6	6	Decline (Dogs)
SPIRITS-WHISKY-P48-750ML	SPIRITS	WHISKY	Factor 6	6	6	Decline (Dogs)
SPIRITS-WHISKY-P45-750ML	SPIRITS	WHISKY	Factor 6	6	6	Decline (Dogs)
SPIRITS-WHISKY-P34-750ML	SPIRITS	WHISKY	Factor 6	6	6	Decline (Dogs)
SPIRITS-WHISKY-P39-750ML	SPIRITS	WHISKY	Factor 6	6	6	Decline (Dogs)
SPIRITS-WHISKY-P36-750ML	SPIRITS	WHISKY	Factor 6	6	6	Decline (Dogs)
SPIRITS-WHISKY-P46-750ML	SPIRITS	WHISKY	Factor 6	6	6	Decline (Dogs)
SPIRITS-WHISKY-P52-750ML	SPIRITS	WHISKY	Factor 6	6	6	Decline (Dogs)
SPIRITS-WHISKY-P43-750ML	SPIRITS	WHISKY	Factor 6	6	6	Decline (Dogs)
SPIRITS-WHISKY-P51-750ML	SPIRITS	WHISKY	Factor 6	6	6	Decline (Dogs)
SPIRITS-WHISKY-P47-750ML	SPIRITS	WHISKY	Factor 6	6	6	Decline (Dogs)
SPIRITS-WHISKY-P37-750ML	SPIRITS	WHISKY	Factor 6	6	6	Decline (Dogs)
SPIRITS-VODKA-P2-1LTR	SPIRITS	WHITE SPIRITS	Factor 8	6	7	Growth (Stars)
SPIRITS-VODKA-P1-1LTR	SPIRITS	WHITE SPIRITS	Factor 1	6	7	Growth (Stars)
SPIRITS-GIN-P1-1LTR	SPIRITS	WHITE SPIRITS	Factor 1	6	7	Growth (Stars)
SPIRITS-RUM-P1-1LTR	SPIRITS	RUM	Factor 4	6	7	Maturity (Cows)
SPIRITS-CANE-P1-1LTR	SPIRITS	WHITE SPIRITS	Factor 3	6	7	Maturity (Cows)
WINE-BIB-P73-5LTR	WINE	WINE	Factor 3	6	8	Maturity (Cows)
WINE-BIB-P83-5LTR	WINE	WINE	Factor 6	6	8	Maturity (Cows)
WINE-RED-P22-750ML	WINE	WINE	Factor 1	6	8	Maturity (Cows)
WINE-BIB-P62-5LTR	WINE	WINE	Factor 3	6	8	Maturity (Cows)
WINE-BIB-P61-5LTR	WINE	WINE	Factor 3	6	8	Maturity (Cows)
WINE-BIB-P48-3LTR	WINE	WINE	Factor 6	6	8	Maturity (Cows)
WINE-WHITE-P54-750ML	WINE	WINE	Factor 6	6	8	Maturity (Cows)
WINE-RED-P24-750ML	WINE	WINE	Factor 1	6	8	Maturity (Cows)
WINE-BIB-P60-5LTR	WINE	WINE	Factor 3	6	8	Maturity (Cows)
WINE-WHITE-P43-750ML	WINE	WINE	Factor 6	6	8	Maturity (Cows)
WINE-WHITE-P18-750ML	WINE	WINE	Factor 1	6	8	Maturity (Cows)
WINE-PERLE-P8-1.5LTR	WINE	PERLE	Factor 3	6	8	Maturity (Cows)
WINE-RED-P32-750ML	WINE	WINE	Factor 1	6	8	Maturity (Cows)
WINE-BIB-P59-5LTR	WINE	WINE	Factor 3	6	8	Maturity (Cows)
WINE-RED-P44-750ML	WINE	WINE	Factor 1	6	8	Maturity (Cows)
WINE-WHITE-P29-750ML	WINE	WINE	Factor 1	6	8	Maturity (Cows)
WINE-RED-P43-750ML	WINE	WINE	Factor 1	6	8	Maturity (Cows)
WINE-BIB-P58-5LTR	WINE	WINE	Factor 3	6	8	Maturity (Cows)
WINE-BIB-P45-3LTR	WINE	WINE	Factor 6	6	8	Maturity (Cows)
WINE-BIB-P28-3LTR	WINE	WINE	Factor 3	6	8	Maturity (Cows)

TABLE A.47: *Part 6: Results for determining a competitor set by clustering according to product life cycle stage displaying factors, clusters and taxonomy*

Factor, Cluster, Sub-Cluster and Taxonomy Results						
Product	Segment	Category	Factor	Cluster	Sub-Cluster	Taxonomy
WINE-PERLE-P2-1LTR	WINE	PERLE	Factor 3	6	8	Maturity (Cows)
WINE-SPARKLING-P19-750ML	WINE	SPARKLING	Factor 3	6	8	Maturity (Cows)
WINE-PERLE-P9-1.5LTR	WINE	PERLE	Factor 3	6	8	Maturity (Cows)
WINE-BIB-P55-5LTR	WINE	WINE	Factor 3	6	8	Maturity (Cows)
WINE-PERLE-P12-750ML	WINE	PERLE	Factor 3	6	8	Maturity (Cows)
WINE-BIB-P5-1LTR	WINE	WINE	Factor 3	6	8	Maturity (Cows)
WINE-RED-P80-750ML	WINE	WINE	Factor 6	6	8	Maturity (Cows)
WINE-PERLE-P10-750ML	WINE	PERLE	Factor 3	6	8	Maturity (Cows)
WINE-RED-P11-750ML	WINE	WINE	Factor 3	6	8	Maturity (Cows)
WINE-RED-P37-750ML	WINE	WINE	Factor 1	6	8	Maturity (Cows)
WINE-FORTIFIED-P8-750ML	WINE	FORTIFIED	Factor 6	6	8	Growth (Stars)
WINE-FORTIFIED-P15-750ML	WINE	FORTIFIED	Factor 7	6	8	Growth (Stars)
WINE-RED-P46-750ML	WINE	WINE	Factor 1	6	8	Maturity (Cows)
WINE-SPARKLING-P16-750ML	WINE	SPARKLING	Factor 3	6	8	Maturity (Cows)
WINE-WHITE-P42-750ML	WINE	WINE	Factor 8	6	8	Maturity (Cows)
WINE-BIB-P47-3LTR	WINE	WINE	Factor 6	6	8	Maturity (Cows)
WINE-WHITE-P21-750ML	WINE	WINE	Factor 1	6	8	Maturity (Cows)
WINE-ROSE-P14-750ML	WINE	WINE	Factor 6	6	8	Maturity (Cows)
WINE-BIB-P57-5LTR	WINE	WINE	Factor 3	6	8	Maturity (Cows)
WINE-WHITE-P51-750ML	WINE	WINE	Factor 6	6	8	Maturity (Cows)
WINE-WHITE-P27-750ML	WINE	WINE	Factor 1	6	8	Maturity (Cows)
WINE-BIB-P4-1LTR	WINE	WINE	Factor 3	6	8	Maturity (Cows)
WINE-WHITE-P26-750ML	WINE	WINE	Factor 1	6	8	Maturity (Cows)
WINE-FORTIFIED-P1-1LTR	WINE	FORTIFIED	Factor 1	6	8	Growth (Stars)
WINE-RED-P10-750ML	WINE	WINE	Factor 3	6	8	Maturity (Cows)
WINE-RED-P41-750ML	WINE	WINE	Factor 1	6	8	Maturity (Cows)
WINE-BIB-P56-5LTR	WINE	WINE	Factor 3	6	8	Maturity (Cows)
WINE-RED-P8-750ML	WINE	WINE	Factor 3	6	8	Maturity (Cows)
WINE-WHITE-P11-750ML	WINE	WINE	Factor 1	6	8	Maturity (Cows)
WINE-BIB-P27-3LTR	WINE	WINE	Factor 3	6	8	Maturity (Cows)
WINE-ROSE-P7-750ML	WINE	WINE	Factor 1	6	8	Maturity (Cows)
WINE-BIB-P23-3LTR	WINE	WINE	Factor 3	6	8	Maturity (Cows)
WINE-BIB-P46-3LTR	WINE	WINE	Factor 6	6	8	Maturity (Cows)
WINE-FORTIFIED-P11-750ML	WINE	FORTIFIED	Factor 5	6	8	Growth (Stars)
WINE-BIB-P3-1LTR	WINE	WINE	Factor 3	6	8	Maturity (Cows)
WINE-FORTIFIED-P10-750ML	WINE	FORTIFIED	Factor 5	6	8	Growth (Stars)
WINE-WHITE-P30-750ML	WINE	WINE	Factor 1	6	8	Maturity (Cows)
WINE-RED-P83-750ML	WINE	WINE	Factor 6	6	8	Maturity (Cows)
WINE-FORTIFIED-P4-2LTR	WINE	FORTIFIED	Factor 5	6	8	Growth (Stars)
WINE-RED-P19-750ML	WINE	WINE	Factor 1	6	8	Maturity (Cows)
WINE-BIB-P2-1LTR	WINE	WINE	Factor 3	6	8	Maturity (Cows)
WINE-RED-P9-750ML	WINE	WINE	Factor 3	6	8	Maturity (Cows)
WINE-BIB-P26-3LTR	WINE	WINE	Factor 3	6	8	Maturity (Cows)
WINE-RED-P47-750ML	WINE	WINE	Factor 1	6	8	Maturity (Cows)
WINE-FORTIFIED-P12-750ML	WINE	FORTIFIED	Factor 5	6	8	Growth (Stars)
WINE-BIB-P84-1.5LTR	WINE	WINE	Factor 6	6	8	Maturity (Cows)
WINE-FORTIFIED-P9-750ML	WINE	FORTIFIED	Factor 6	6	8	Growth (Stars)
WINE-FORTIFIED-P14-750ML	WINE	FORTIFIED	Factor 5	6	8	Growth (Stars)

TABLE A.48: *Part 7: Results for determining a competitor set by clustering according to product life cycle stage displaying factors, clusters and taxonomy*

Factor, Cluster, Sub-Cluster and Taxonomy Results						
Product	Segment	Category	Factor	Cluster	Sub-Cluster	Taxonomy
WINE-FORTIFIED-P13-750ML	WINE	FORTIFIED	Factor 5	6	8	Growth (Stars)
WINE-RED-P45-750ML	WINE	WINE	Factor 1	6	8	Maturity (Cows)
WINE-BIB-P1-1LTR	WINE	WINE	Factor 3	6	8	Maturity (Cows)
WINE-BIB-P22-3LTR	WINE	WINE	Factor 3	6	8	Maturity (Cows)
WINE-BIB-P24-3LTR	WINE	WINE	Factor 3	6	8	Decline (Dogs)
WINE-ROSE-P4-1.5LTR	WINE	WINE	Factor 6	6	8	Decline (Dogs)
WINE-WHITE-P55-750ML	WINE	WINE	Factor 6	6	8	Decline (Dogs)
WINE-BIB-P25-3LTR	WINE	WINE	Factor 3	6	8	Decline (Dogs)
WINE-PERLE-P7-1.5LTR	WINE	PERLE	Factor 3	6	8	Decline (Dogs)
WINE-RED-P87-750ML	WINE	WINE	Factor 6	6	8	Decline (Dogs)
WINE-FORTIFIED-P7-750ML	WINE	FORTIFIED	Factor 6	6	8	Introduction (Question Marks)
WINE-RED-P84-750ML	WINE	WINE	Factor 6	6	8	Decline (Dogs)
WINE-PERLE-P11-750ML	WINE	PERLE	Factor 3	6	8	Decline (Dogs)
WINE-BIB-P54-5LTR	WINE	WINE	Factor 3	6	8	Decline (Dogs)
SPIRITS-VODKA-P5-750ML	SPIRITS	WHITE SPIRITS	Factor 6	6	9	Growth (Stars)
SPIRITS-GIN-P4-750ML	SPIRITS	WHITE SPIRITS	Factor 1	6	9	Growth (Stars)
SPIRITS-VODKA-P7-750ML	SPIRITS	WHITE SPIRITS	Factor 1	6	9	Growth (Stars)
SPIRITS-RUM-P7-750ML	SPIRITS	RUM	Factor 4	6	9	Maturity (Cows)
WINE-WHITE-P67-750ML	WINE	WINE	Factor 6	6	9	Maturity (Cows)
SPIRITS-GIN-P5-750ML	SPIRITS	WHITE SPIRITS	Factor 6	6	9	Growth (Stars)
SPIRITS-RUM-P2-750ML	SPIRITS	RUM	Factor 1	6	9	Maturity (Cows)
SPIRITS-VODKA-P4-750ML	SPIRITS	WHITE SPIRITS	Factor 4	6	9	Growth (Stars)
SPIRITS-GIN-P7-750ML	SPIRITS	WHITE SPIRITS	Factor 6	6	9	Growth (Stars)
SPIRITS-VODKA-P12-750ML	SPIRITS	WHITE SPIRITS	Factor 3	6	9	Growth (Stars)
WINE-WHITE-P48-750ML	WINE	WINE	Factor 6	6	9	Maturity (Cows)
WINE-RED-P20-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-BIB-P77-5LTR	WINE	WINE	Factor 1	6	9	Decline (Dogs)
SPIRITS-VODKA-P8-750ML	SPIRITS	WHITE SPIRITS	Factor 1	6	9	Introduction (Question Marks)
SPIRITS-VODKA-P18-750ML	SPIRITS	WHITE SPIRITS	Factor 6	6	9	Introduction (Question Marks)
WINE-WHITE-P12-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
SPIRITS-VODKA-P9-750ML	SPIRITS	WHITE SPIRITS	Factor 1	6	9	Introduction (Question Marks)
WINE-WHITE-P56-750ML	WINE	WINE	Factor 6	6	9	Decline (Dogs)
WINE-RED-P75-750ML	WINE	WINE	Factor 6	6	9	Decline (Dogs)
WINE-ROSE-P15-750ML	WINE	WINE	Factor 6	6	9	Decline (Dogs)
SPIRITS-GIN-P9-750ML	SPIRITS	WHITE SPIRITS	Factor 6	6	9	Introduction (Question Marks)
SPIRITS-VODKA-P10-750ML	SPIRITS	WHITE SPIRITS	Factor 1	6	9	Introduction (Question Marks)
WINE-RED-P16-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
SPIRITS-GIN-P6-750ML	SPIRITS	WHITE SPIRITS	Factor 6	6	9	Introduction (Question Marks)
SPIRITS-CANE-P2-750ML	SPIRITS	WHITE SPIRITS	Factor 6	6	9	Decline (Dogs)
SPIRITS-VODKA-P17-750ML	SPIRITS	WHITE SPIRITS	Factor 6	6	9	Introduction (Question Marks)
SPIRITS-GIN-P10-750ML	SPIRITS	WHITE SPIRITS	Factor 6	6	9	Introduction (Question Marks)
WINE-RED-P89-750ML	WINE	WINE	Factor 6	6	9	Decline (Dogs)
WINE-WHITE-P36-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-RED-P50-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
SPIRITS-VODKA-P11-750ML	SPIRITS	WHITE SPIRITS	Factor 1	6	9	Introduction (Question Marks)
WINE-RED-P49-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-RED-P74-750ML	WINE	WINE	Factor 6	6	9	Decline (Dogs)

TABLE A.49: *Part 8: Results for determining a competitor set by clustering according to product life cycle stage displaying factors, clusters and taxonomy*

		Factor, Cluster, Sub-Cluster and Taxonomy Results				
Product	Segment	Category	Factor	Cluster	Sub-Cluster	Taxonomy
SPIRITS-VODKA-P19-750ML	SPIRITS	WHITE SPIRITS	Factor 6	6	9	Introduction (Question Marks)
WINE-ROSE-P11-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-BIB-P76-5LTR	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-WHITE-P10-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
SPIRITS-VODKA-P13-750ML	SPIRITS	WHITE SPIRITS	Factor 3	6	9	Introduction (Question Marks)
SPIRITS-VODKA-P14-750ML	SPIRITS	WHITE SPIRITS	Factor 3	6	9	Introduction (Question Marks)
WINE-RED-P35-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-RED-P69-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
SPIRITS-GIN-P8-750ML	SPIRITS	WHITE SPIRITS	Factor 6	6	9	Introduction (Question Marks)
WINE-WHITE-P22-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-RED-P86-750ML	WINE	WINE	Factor 6	6	9	Decline (Dogs)
WINE-RED-P70-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-RED-P12-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-RED-P81-750ML	WINE	WINE	Factor 6	6	9	Decline (Dogs)
WINE-WHITE-P15-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-RED-P73-750ML	WINE	WINE	Factor 6	6	9	Decline (Dogs)
SPIRITS-VODKA-P22-750ML	SPIRITS	WHITE SPIRITS	Factor 6	6	9	Introduction (Question Marks)
WINE-RED-P33-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
SPIRITS-VODKA-P15-750ML	SPIRITS	WHITE SPIRITS	Factor 3	6	9	Introduction (Question Marks)
SPIRITS-VODKA-P21-750ML	SPIRITS	WHITE SPIRITS	Factor 6	6	9	Introduction (Question Marks)
WINE-ROSE-P20-750ML	WINE	WINE	Factor 6	6	9	Decline (Dogs)
WINE-RED-P78-750ML	WINE	WINE	Factor 6	6	9	Decline (Dogs)
WINE-RED-P65-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-RED-P26-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-WHITE-P17-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
SPIRITS-VODKA-P20-750ML	SPIRITS	WHITE SPIRITS	Factor 6	6	9	Introduction (Question Marks)
WINE-RED-P76-750ML	WINE	WINE	Factor 6	6	9	Decline (Dogs)
WINE-RED-P52-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-BIB-P72-5LTR	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-WHITE-P24-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-RED-P55-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-RED-P56-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-RED-P30-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-WHITE-P33-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-WHITE-P1-5LTR	WINE	WINE	Factor 6	6	9	Decline (Dogs)
WINE-WHITE-P19-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-BIB-P68-5LTR	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-ROSE-P2-1.5LTR	WINE	WINE	Factor 6	6	9	Decline (Dogs)
WINE-RED-P29-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-WHITE-P53-750ML	WINE	WINE	Factor 6	6	9	Decline (Dogs)
WINE-BIB-P75-5LTR	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-RED-P27-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-RED-P51-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-WHITE-P50-750ML	WINE	WINE	Factor 6	6	9	Decline (Dogs)

TABLE A.50: *Part 9: Results for determining a competitor set by clustering according to product life cycle stage displaying factors, clusters and taxonomy*

		Factor, Cluster, Sub-Cluster and Taxonomy Results				
Product	Segment	Category	Factor	Cluster	Sub-Cluster	Taxonomy
WINE-WHITE-P41-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-RED-P77-750ML	WINE	WINE	Factor 6	6	9	Decline (Dogs)
WINE-BIB-P66-5LTR	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-WHITE-P40-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-WHITE-P31-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-RED-P53-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-WHITE-P49-750ML	WINE	WINE	Factor 6	6	9	Decline (Dogs)
WINE-WHITE-P13-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-RED-P39-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-WHITE-P44-750ML	WINE	WINE	Factor 6	6	9	Decline (Dogs)
WINE-RED-P36-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-WHITE-P23-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-WHITE-P25-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-ROSE-P8-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-RED-P40-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-ROSE-P10-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-WHITE-P46-750ML	WINE	WINE	Factor 6	6	9	Decline (Dogs)
WINE-WHITE-P34-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-RED-P58-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-RED-P31-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-RED-P79-750ML	WINE	WINE	Factor 6	6	9	Decline (Dogs)
WINE-RED-P82-750ML	WINE	WINE	Factor 6	6	9	Decline (Dogs)
WINE-RED-P34-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-WHITE-P39-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-RED-P57-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-RED-P59-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-WHITE-P45-750ML	WINE	WINE	Factor 6	6	9	Decline (Dogs)
WINE-RED-P101-750ML	WINE	WINE	Factor 6	6	9	Decline (Dogs)
WINE-WHITE-P16-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-RED-P62-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-RED-P63-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-WHITE-P37-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-RED-P66-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-BIB-P65-5LTR	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-WHITE-P38-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-RED-P61-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-WHITE-P28-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-ROSE-P9-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-RED-P54-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-RED-P48-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-WHITE-P35-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-RED-P60-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-BIB-P64-5LTR	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-WHITE-P32-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-RED-P64-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-RED-P67-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-BIB-P71-5LTR	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-BIB-P67-5LTR	WINE	WINE	Factor 1	6	9	Decline (Dogs)

TABLE A.51: *Part 10: Results for determining a competitor set by clustering according to product life cycle stage displaying factors, clusters and taxonomy*

			Factor, Cluster, Sub-Cluster and Taxonomy Results			
Product	Segment	Category	Factor	Cluster	Sub-Cluster	Taxonomy
WINE-BIB-P70-5LTR	WINE	WINE	Factor 1	6	9	Decline (Dogs)
WINE-RED-P68-750ML	WINE	WINE	Factor 1	6	9	Decline (Dogs)
SPIRITS-VODKA-P16-750ML	SPIRITS	WHITE SPIRITS	Factor 3	6	9	Introduction (Question Marks)
SPIRITS-WHISKY-P53-750ML	SPIRITS	WHISKY	Factor 5	6	10	Maturity (Cows)
SPIRITS-WHISKY-P23-750ML	SPIRITS	WHISKY	Factor 6	6	10	Maturity (Cows)
SPIRITS-GIN-P11-750ML	SPIRITS	WHITE SPIRITS	Factor 7	6	10	Growth (Stars)
SPIRITS-WHISKY-P26-750ML	SPIRITS	WHISKY	Factor 6	6	10	Decline (Dogs)
WINE-SPARKLING-P1-750ML	WINE	SPARKLING	Factor 8	6	11	Maturity (Cows)
WINE-SPARKLING-P27-750ML	WINE	SPARKLING	Factor 7	6	11	Maturity (Cows)
WINE-SPARKLING-P31-750ML	WINE	SPARKLING	Factor 6	6	11	Maturity (Cows)
WINE-SPARKLING-P2-750ML	WINE	SPARKLING	Factor 3	6	11	Maturity (Cows)
WINE-SPARKLING-P7-750ML	WINE	SPARKLING	Factor 6	6	11	Maturity (Cows)
WINE-WHITE-P20-750ML	WINE	WINE	Factor 1	6	11	Maturity (Cows)
WINE-SPARKLING-P25-750ML	WINE	SPARKLING	Factor 3	6	11	Maturity (Cows)
WINE-SPARKLING-P29-750ML	WINE	SPARKLING	Factor 8	6	11	Maturity (Cows)
WINE-SPARKLING-P5-750ML	WINE	SPARKLING	Factor 3	6	11	Maturity (Cows)
WINE-SPARKLING-P6-750ML	WINE	SPARKLING	Factor 3	6	11	Maturity (Cows)
WINE-SPARKLING-P30-750ML	WINE	SPARKLING	Factor 8	6	11	Maturity (Cows)
WINE-SPARKLING-P12-750ML	WINE	SPARKLING	Factor 3	6	11	Maturity (Cows)
WINE-SPARKLING-P20-750ML	WINE	SPARKLING	Factor 3	6	11	Maturity (Cows)
WINE-SPARKLING-P14-750ML	WINE	SPARKLING	Factor 3	6	11	Maturity (Cows)
WINE-SPARKLING-P13-750ML	WINE	SPARKLING	Factor 3	6	11	Maturity (Cows)
WINE-SPARKLING-P9-750ML	WINE	SPARKLING	Factor 6	6	11	Maturity (Cows)
WINE-SPARKLING-P11-750ML	WINE	SPARKLING	Factor 3	6	11	Maturity (Cows)
WINE-SPARKLING-P33-750ML	WINE	SPARKLING	Factor 3	6	11	Maturity (Cows)
WINE-SPARKLING-P3-750ML	WINE	SPARKLING	Factor 3	6	11	Maturity (Cows)
WINE-SPARKLING-P15-750ML	WINE	SPARKLING	Factor 3	6	11	Maturity (Cows)
WINE-SPARKLING-P8-750ML	WINE	SPARKLING	Factor 6	6	11	Maturity (Cows)
WINE-SPARKLING-P26-750ML	WINE	SPARKLING	Factor 3	6	11	Decline (Dogs)
WINE-SPARKLING-P17-750ML	WINE	SPARKLING	Factor 3	6	11	Decline (Dogs)
WINE-SPARKLING-P28-750ML	WINE	SPARKLING	Factor 3	6	11	Decline (Dogs)
WINE-SPARKLING-P32-750ML	WINE	SPARKLING	Factor 3	6	11	Decline (Dogs)
WINE-SPARKLING-P21-750ML	WINE	SPARKLING	Factor 3	6	11	Decline (Dogs)
WINE-SPARKLING-P18-750ML	WINE	SPARKLING	Factor 3	6	11	Decline (Dogs)
WINE-SPARKLING-P10-750ML	WINE	SPARKLING	Factor 3	6	11	Decline (Dogs)
WINE-SPARKLING-P24-750ML	WINE	SPARKLING	Factor 3	6	11	Decline (Dogs)
WINE-SPARKLING-P4-750ML	WINE	SPARKLING	Factor 8	6	11	Decline (Dogs)
WINE-SPARKLING-P23-750ML	WINE	SPARKLING	Factor 3	6	11	Decline (Dogs)
WINE-SPARKLING-P22-750ML	WINE	SPARKLING	Factor 3	6	11	Decline (Dogs)
RTD-FABS-P36-440ML	RTDs	FABS & CIDER	Factor 4	6	12	Maturity (Cows)
RTD-CIDER-P13-440ML	RTDs	FABS & CIDER	Factor 3	6	12	Growth (Stars)
RTD-CIDER-P3-330ML	RTDs	FABS & CIDER	Factor 3	6	12	Growth (Stars)
RTD-CIDER-P14-440ML	RTDs	FABS & CIDER	Factor 6	6	12	Growth (Stars)
RTD-FABS-P35-440ML	RTDs	FABS & CIDER	Factor 4	6	12	Maturity (Cows)
WINE-BIB-P15-1LTR	WINE	WINE	Factor 8	6	13	Maturity (Cows)
WINE-RED-P106-750ML	WINE	WINE	Factor 7	6	13	Maturity (Cows)
WINE-BIB-P40-3LTR	WINE	WINE	Factor 8	6	13	Maturity (Cows)
WINE-RED-P23-750ML	WINE	WINE	Factor 1	6	13	Maturity (Cows)

TABLE A.52: *Part 11: Results for determining a competitor set by clustering according to product life cycle stage displaying factors, clusters and taxonomy*

Factor, Cluster, Sub-Cluster and Taxonomy Results						
Product	Segment	Category	Factor	Cluster	Sub-Cluster	Taxonomy
WINE-BIB-P14-1LTR	WINE	WINE	Factor 8	6	13	Maturity (Cows)
WINE-BIB-P38-3LTR	WINE	WINE	Factor 8	6	13	Maturity (Cows)
WINE-RED-P25-750ML	WINE	WINE	Factor 8	6	13	Maturity (Cows)
WINE-BIB-P13-1LTR	WINE	WINE	Factor 8	6	13	Maturity (Cows)
WINE-RED-P21-750ML	WINE	WINE	Factor 8	6	13	Maturity (Cows)
WINE-BIB-P39-3LTR	WINE	WINE	Factor 8	6	13	Maturity (Cows)
WINE-BIB-P12-1LTR	WINE	WINE	Factor 8	6	13	Maturity (Cows)
WINE-ROSE-P21-750ML	WINE	WINE	Factor 8	6	13	Maturity (Cows)
WINE-BIB-P33-3LTR	WINE	WINE	Factor 8	6	13	Maturity (Cows)
WINE-RED-P2-1.5LTR	WINE	WINE	Factor 8	6	13	Maturity (Cows)
WINE-WHITE-P14-750ML	WINE	WINE	Factor 1	6	13	Maturity (Cows)
WINE-ROSE-P6-750ML	WINE	WINE	Factor 8	6	13	Maturity (Cows)
WINE-RED-P103-750ML	WINE	WINE	Factor 8	6	13	Maturity (Cows)
WINE-BIB-P37-3LTR	WINE	WINE	Factor 8	6	13	Maturity (Cows)
WINE-BIB-P11-1LTR	WINE	WINE	Factor 8	6	13	Maturity (Cows)
WINE-BIB-P34-3LTR	WINE	WINE	Factor 8	6	13	Maturity (Cows)
WINE-RED-P1-1.5LTR	WINE	WINE	Factor 8	6	13	Maturity (Cows)
WINE-BIB-P8-1LTR	WINE	WINE	Factor 8	6	13	Maturity (Cows)
WINE-BIB-P9-1LTR	WINE	WINE	Factor 8	6	13	Maturity (Cows)
WINE-BIB-P10-1LTR	WINE	WINE	Factor 8	6	13	Maturity (Cows)
WINE-BIB-P31-3LTR	WINE	WINE	Factor 8	6	13	Maturity (Cows)
WINE-RED-P17-750ML	WINE	WINE	Factor 8	6	13	Maturity (Cows)
WINE-WHITE-P64-750ML	WINE	WINE	Factor 8	6	13	Maturity (Cows)
WINE-BIB-P32-3LTR	WINE	WINE	Factor 8	6	13	Maturity (Cows)
WINE-WHITE-P5-750ML	WINE	WINE	Factor 8	6	13	Maturity (Cows)
WINE-BIB-P36-3LTR	WINE	WINE	Factor 8	6	13	Maturity (Cows)
WINE-RED-P13-750ML	WINE	WINE	Factor 8	6	13	Maturity (Cows)
WINE-RED-P42-750ML	WINE	WINE	Factor 1	6	13	Maturity (Cows)
WINE-RED-P18-750ML	WINE	WINE	Factor 8	6	13	Decline (Dogs)
WINE-RED-P28-750ML	WINE	WINE	Factor 1	6	13	Decline (Dogs)
WINE-BIB-P17-2LTR	WINE	WINE	Factor 8	6	13	Decline (Dogs)
WINE-RED-P104-750ML	WINE	WINE	Factor 8	6	13	Decline (Dogs)
WINE-BIB-P29-3LTR	WINE	WINE	Factor 8	6	13	Decline (Dogs)
WINE-BIB-P7-1LTR	WINE	WINE	Factor 8	6	13	Decline (Dogs)
WINE-RED-P14-750ML	WINE	WINE	Factor 8	6	13	Decline (Dogs)
WINE-WHITE-P6-750ML	WINE	WINE	Factor 8	6	13	Decline (Dogs)
WINE-BIB-P35-3LTR	WINE	WINE	Factor 8	6	13	Decline (Dogs)
WINE-RED-P102-750ML	WINE	WINE	Factor 8	6	13	Decline (Dogs)
WINE-WHITE-P9-750ML	WINE	WINE	Factor 8	6	13	Decline (Dogs)
WINE-RED-P15-750ML	WINE	WINE	Factor 8	6	13	Decline (Dogs)
WINE-BIB-P20-2LTR	WINE	WINE	Factor 8	6	13	Decline (Dogs)
WINE-BIB-P30-3LTR	WINE	WINE	Factor 8	6	13	Decline (Dogs)
WINE-WHITE-P65-750ML	WINE	WINE	Factor 8	6	13	Decline (Dogs)
WINE-RED-P38-750ML	WINE	WINE	Factor 1	6	13	Decline (Dogs)
WINE-BIB-P19-2LTR	WINE	WINE	Factor 8	6	13	Decline (Dogs)
WINE-WHITE-P7-750ML	WINE	WINE	Factor 8	6	13	Decline (Dogs)
WINE-BIB-P21-2LTR	WINE	WINE	Factor 8	6	13	Decline (Dogs)

TABLE A.53: *Part 12: Results for determining a competitor set by clustering according to product life cycle stage displaying factors, clusters and taxonomy*

	Factor, Cluster, Sub-Cluster and Taxonomy Results					
Product	Segment	Category	Factor	Cluster	Sub-Cluster	Taxonomy
WINE-BIB-P6-1LTR	WINE	WINE	Factor 8	6	13	Decline (Dogs)
WINE-BIB-P16-2LTR	WINE	WINE	Factor 8	6	13	Decline (Dogs)
WINE-WHITE-P8-750ML	WINE	WINE	Factor 8	6	13	Decline (Dogs)
WINE-BIB-P18-2LTR	WINE	WINE	Factor 8	6	13	Decline (Dogs)
WINE-RED-P94-750ML	WINE	WINE	Factor 2	6	14	Maturity (Cows)
WINE-RED-P96-750ML	WINE	WINE	Factor 2	6	14	Maturity (Cows)
WINE-WHITE-P61-750ML	WINE	WINE	Factor 2	6	14	Maturity (Cows)
WINE-RED-P97-750ML	WINE	WINE	Factor 2	6	14	Maturity (Cows)
WINE-RED-P99-750ML	WINE	WINE	Factor 2	6	14	Maturity (Cows)
WINE-WHITE-P62-750ML	WINE	WINE	Factor 2	6	14	Maturity (Cows)
WINE-RED-P98-750ML	WINE	WINE	Factor 2	6	14	Maturity (Cows)
WINE-RED-P100-750ML	WINE	WINE	Factor 2	6	14	Maturity (Cows)
WINE-WHITE-P63-750ML	WINE	WINE	Factor 2	6	14	Maturity (Cows)
SPIRITS-BRANDY-P8-750ML	SPIRITS	BRANDY	Factor 3	6	15	Maturity (Cows)
SPIRITS-WHISKY-P13-750ML	SPIRITS	WHISKY	Factor 3	6	15	Maturity (Cows)
RTD-CIDER-P2-330ML	RTDs	FABS & CIDER	Factor 6	6	16	Growth (Stars)
RTD-FABS-P6-250ML	RTDs	FABS & CIDER	Factor 4	6	16	Maturity (Cows)
RTD-CIDER-P10-330ML	RTDs	FABS & CIDER	Factor 7	6	16	Growth (Stars)
RTD-FABS-P22-275ML	RTDs	FABS & CIDER	Factor 4	6	16	Maturity (Cows)
RTD-CIDER-P4-330ML	RTDs	FABS & CIDER	Factor 6	6	16	Growth (Stars)
RTD-FABS-P7-275ML	RTDs	FABS & CIDER	Factor 8	6	16	Maturity (Cows)
RTD-FABS-P11-275ML	RTDs	FABS & CIDER	Factor 8	6	16	Maturity (Cows)
RTD-FABS-P26-330ML	RTDs	FABS & CIDER	Factor 8	6	16	Maturity (Cows)
RTD-FABS-P25-330ML	RTDs	FABS & CIDER	Factor 1	6	16	Maturity (Cows)
RTD-FABS-P5-250ML	RTDs	FABS & CIDER	Factor 4	6	16	Maturity (Cows)
RTD-FABS-P15-275ML	RTDs	FABS & CIDER	Factor 1	6	16	Maturity (Cows)

TABLE A.54: *Part 13: Results for determining a competitor set by clustering according to product life cycle stage displaying factors, clusters and taxonomy*

A.5   Method for the price elasticity model to determine a competitor set

Product	Product	Product
RTD-BEER-P14-330ML	RTD-CIDER-P5-330ML	WINE-ROSE-P4-1.5LTR
RTD-BEER-P35-440ML	RTD-CIDER-P6-330ML	WINE-BIB-P47-3LTR
RTD-BEER-P51-660ML	RTD-CIDER-P12-440ML	WINE-BIB-P46-3LTR
RTD-BEER-P13-330ML	RTD-CIDER-P16-500ML	WINE-BIB-P45-3LTR
RTD-BEER-P34-440ML	RTD-CIDER-P2-330ML	WINE-BIB-P48-3LTR
RTD-FABS-P11-275ML	RTD-CIDER-P10-330ML	WINE-BIB-P49-3LTR
RTD-BEER-P7-330ML	RTD-FABS-P2-250ML	WINE-BIB-P50-3LTR
RTD-FABS-P21-275ML	RTD-FABS-P4-250ML	WINE-RED-P4-1.5LTR
RTD-FABS-P20-275ML	RTD-FABS-P3-250ML	WINE-WHITE-P2-1.5LTR
RTD-FABS-P9-275ML	RTD-FABS-P1-250ML	WINE-WHITE-P2-5LTR
RTD-FABS-P8-275ML	RTD-FABS-P24-300ML	WINE-FORTIFIED-P4-2LTR
RTD-FABS-P13-275ML	RTD-FABS-P29-330ML	WINE-BIB-P3-1LTR
RTD-FABS-P31-440ML	RTD-FABS-P34-440ML	WINE-BIB-P26-3LTR
RTD-FABS-P38-660ML	RTD-FABS-P23-300ML	WINE-BIB-P61-5LTR
RTD-FABS-P12-275ML	RTD-FABS-P33-440ML	WINE-BIB-P5-1LTR
RTD-FABS-P10-275ML	RTD-BEER-P5-330ML	WINE-BIB-P23-3LTR
RTD-FABS-P27-330ML	RTD-FABS-P15-275ML	WINE-BIB-P60-5LTR
RTD-FABS-P7-275ML	RTD-BEER-P9-330ML	WINE-BIB-P27-3LTR
RTD-BEER-P2-330ML	RTD-BEER-P30-440ML	WINE-BIB-P62-5LTR
RTD-BEER-P21-340ML	RTD-CIDER-P4-330ML	WINE-BIB-P2-1LTR
RTD-BEER-P27-440ML	RTD-CIDER-P14-440ML	WINE-BIB-P22-3LTR
RTD-BEER-P44-500ML	RTD-CIDER-P3-330ML	WINE-BIB-P55-5LTR
RTD-BEER-P58-750ML	RTD-CIDER-P13-440ML	WINE-BIB-P1-1LTR
RTD-BEER-P32-440ML	RTD-BEER-P12-330ML	WINE-BIB-P24-3LTR
RTD-BEER-P1-330ML	RTD-BEER-P33-440ML	WINE-BIB-P59-5LTR
RTD-BEER-P20-340ML	RTD-BEER-P42-440ML	WINE-BIB-P57-5LTR
RTD-BEER-P26-440ML	RTD-BEER-P52-660ML	WINE-BIB-P56-5LTR
RTD-BEER-P43-500ML	RTD-BEER-P19-330ML	WINE-BIB-P4-1LTR
RTD-BEER-P55-750ML	RTD-BEER-P41-440ML	WINE-BIB-P28-3LTR
RTD-BEER-P3-330ML	RTD-BEER-P53-660ML	WINE-BIB-P73-5LTR
RTD-BEER-P24-340ML	RTD-BEER-P18-330ML	WINE-BIB-P25-3LTR
RTD-BEER-P31-440ML	SPIRITS-LIQUEURS-P16-750ML	WINE-BIB-P58-5LTR
RTD-BEER-P45-500ML	SPIRITS-WHISKY-P1-1LTR	WINE-BIB-P54-5LTR
RTD-BEER-P50-660ML	SPIRITS-WHISKY-P14-750ML	WINE-BIB-P67-5LTR
RTD-BEER-P22-340ML	SPIRITS-WHISKY-P2-1LTR	WINE-BIB-P71-5LTR
RTD-BEER-P38-440ML	SPIRITS-WHISKY-P11-750ML	WINE-BIB-P70-5LTR
RTD-BEER-P23-340ML	SPIRITS-WHISKY-P3-1LTR	WINE-BIB-P63-5LTR
RTD-BEER-P39-440ML	SPIRITS-BRANDY-P12-750ML	WINE-BIB-P69-5LTR
RTD-BEER-P59-750ML	SPIRITS-WHISKY-P4-1LTR	WINE-BIB-P74-5LTR
RTD-BEER-P25-355ML	SPIRITS-WHISKY-P22-750ML	WINE-PERLE-P3-1LTR
RTD-BEER-P4-330ML	SPIRITS-WHISKY-P8-1LTR	WINE-PERLE-P4-2LTR
RTD-BEER-P28-440ML	SPIRITS-WHISKY-P21-750ML	WINE-BIB-P6-1LTR
RTD-BEER-P4-330ML	SPIRITS-WHISKY-P53-750ML	WINE-BIB-P16-2LTR
RTD-BEER-P10-330ML	SPIRITS-WHISKY-P9-1LTR	WINE-BIB-P29-3LTR
RTD-BEER-P29-440ML	SPIRITS-WHISKY-P25-750ML	WINE-RED-P1-1.5LTR
RTD-BEER-P16-330ML	SPIRITS-BRANDY-P1-1LTR	WINE-RED-P2-1.5LTR

TABLE A.55: *Part 1: Possible competitor set for RTD-CIDER-P18-660ML after refinement by industry insiders*

Product	Product	Product
RTD-BEER-P37-440ML	SPIRITS-BRANDY-P7-750ML	WINE-BIB-P19-2LTR
RTD-BEER-P46-500ML	SPIRITS-BRANDY-P8-750ML	WINE-BIB-P34-3LTR
RTD-BEER-P57-750ML	SPIRITS-BRANDY-P4-1LTR	WINE-BIB-P21-2LTR
RTD-BEER-P17-330ML	SPIRITS-BRANDY-P23-750ML	WINE-BIB-P14-1LTR
RTD-BEER-P40-440ML	SPIRITS-VODKA-P1-1LTR	WINE-BIB-P39-3LTR
RTD-BEER-P48-5LTR	SPIRITS-WHISKY-P13-750ML	WINE-BIB-P9-1LTR
RTD-BEER-P49-650ML	SPIRITS-VODKA-P2-1LTR	WINE-BIB-P37-3LTR
RTD-CIDER-P9-330ML	SPIRITS-VODKA-P3-200ML	WINE-BIB-P32-3LTR
RTD-CIDER-P15-440ML	SPIRITS-VODKA-P5-750ML	WINE-BIB-P11-1LTR
RTD-CIDER-P18-660ML	SPIRITS-WHISKY-P17-750ML	WINE-BIB-P30-3LTR
RTD-CIDER-P7-330ML	SPIRITS-BRANDY-P3-1LTR	WINE-BIB-P35-3LTR
RTD-FABS-P6-250ML	WINE-RED-P79-1.5LTR	WINE-BIB-P7-1LTR
RTD-FABS-P22-275ML	WINE-BIB-P42-3LTR	WINE-BIB-P17-2LTR
RTD-FABS-P36-440ML	WINE-BIB-P79-5LTR	WINE-BIB-P31-3LTR
RTD-FABS-P5-250ML	WINE-BIB-P82-1.5LTR	WINE-BIB-P38-3LTR
RTD-FABS-P35-440ML	WINE-BIB-P43-3LTR	WINE-ROSE-P1-1.5LTR
RTD-CIDER-P1-330ML	WINE-BIB-P82-5LTR	WINE-BIB-P44-3LTR
RTD-CIDER-P11-440ML	WINE-BIB-P41-3LTR	WINE-BIB-P10-1LTR
RTD-CIDER-P17-660ML	WINE-BIB-P78-5LTR	WINE-BIB-P36-3LTR
RTD-CIDER-P8-330ML	WINE-PERLE-P8-1.5LTR	WINE-BIB-P18-2LTR
RTD-FABS-P30-330ML	WINE-PERLE-P2-1LTR	WINE-BIB-P20-2LTR
RTD-FABS-P26-330ML	WINE-PERLE-P7-1.5LTR	WINE-BIB-P33-3LTR
RTD-FABS-P14-275ML	WINE-PERLE-P9-1.5LTR	WINE-BIB-P8-1LTR
RTD-FABS-P28-330ML	WINE-BIB-P68-5LTR	WINE-BIB-P15-1LTR
RTD-FABS-P32-440ML	WINE-BIB-P72-5LTR	WINE-BIB-P40-3LTR
RTD-BEER-P6-330ML	WINE-BIB-P65-5LTR	WINE-BIB-P12-1LTR
RTD-BEER-P36-440ML	WINE-BIB-P64-5LTR	WINE-BIB-P13-1LTR
RTD-BEER-P47-500ML	WINE-FORTIFIED-P1-1LTR	WINE-FORTIFIED-P2-1LTR
RTD-BEER-P56-750ML	WINE-BIB-P81-5LTR	WINE-FORTIFIED-P5-2LTR
RTD-BEER-P15-330ML	WINE-BIB-P80-5LTR	WINE-FORTIFIED-P6-375ML
RTD-BEER-P8-330ML	WINE-PERLE-P1-1LTR	WINE-BIB-P77-5LTR
RTD-FABS-P17-275ML	WINE-WHITE-P1-5LTR	WINE-BIB-P76-5LTR
RTD-FABS-P16-275ML	WINE-ROSE-P2-1.5LTR	WINE-BIB-P75-5LTR
RTD-FABS-P19-275ML	WINE-BIB-P83-5LTR	WINE-BIB-P66-5LTR
RTD-FABS-P18-275ML	WINE-BIB-P84-1.5LTR	WINE-BIB-P51-3LTR
RTD-FABS-P25-330ML	WINE-BIB-P84-5LTR	WINE-BIB-P52-3LTR
RTD-FABS-P37-660ML	WINE-ROSE-P4-1.5LTR	WINE-RED-P3-1.5LTR

TABLE A.56: *Part 2: Possible competitor set for RTD-CIDER-P18-660ML after refinement by industry insiders*

Product	Product	Product
RTD-BEER-P14-330ML	RTD-BEER-P40-440ML	RTD-FABS-P15-275ML
RTD-BEER-P35-440ML	RTD-BEER-P48-5LTR	RTD-BEER-P9-330ML
RTD-BEER-P51-660ML	RTD-BEER-P49-650ML	RTD-BEER-P30-440ML
RTD-BEER-P13-330ML	RTD-CIDER-P9-330ML	RTD-CIDER-P4-330ML
RTD-BEER-P34-440ML	RTD-CIDER-P15-440ML	RTD-CIDER-P14-440ML
RTD-FABS-P11-275ML	RTD-CIDER-P18-660ML	RTD-CIDER-P3-330ML
RTD-BEER-P7-330ML	RTD-CIDER-P7-330ML	RTD-CIDER-P13-440ML
RTD-FABS-P21-275ML	RTD-FABS-P6-250ML	RTD-BEER-P12-330ML
RTD-FABS-P20-275ML	RTD-FABS-P22-275ML	RTD-BEER-P33-440ML
RTD-FABS-P9-275ML	RTD-FABS-P36-440ML	RTD-BEER-P42-440ML
RTD-FABS-P8-275ML	RTD-FABS-P5-250ML	RTD-BEER-P52-660ML
RTD-FABS-P13-275ML	RTD-FABS-P35-440ML	RTD-BEER-P19-330ML
RTD-FABS-P31-440ML	RTD-CIDER-P1-330ML	RTD-BEER-P41-440ML
RTD-FABS-P38-660ML	RTD-CIDER-P11-440ML	RTD-BEER-P53-660ML
RTD-FABS-P12-275ML	RTD-CIDER-P17-660ML	RTD-BEER-P18-330ML
RTD-FABS-P10-275ML	RTD-CIDER-P8-330ML	SPIRITS-WHISKY-P41-750ML
RTD-FABS-P27-330ML	RTD-FABS-P30-330ML	SPIRITS-LIQUEURS-P2-1LTR
RTD-FABS-P7-275ML	RTD-FABS-P26-330ML	SPIRITS-LIQUEURS-P16-750ML
RTD-BEER-P2-330ML	RTD-FABS-P14-275ML	SPIRITS-WHISKY-P27-750ML
RTD-BEER-P21-340ML	RTD-FABS-P28-330ML	SPIRITS-WHISKY-P15-750ML
RTD-BEER-P27-440ML	RTD-FABS-P32-440ML	SPIRITS-WHISKY-P16-750ML
RTD-BEER-P44-500ML	RTD-BEER-P6-330ML	SPIRITS-WHISKY-P18-750ML
RTD-BEER-P58-750ML	RTD-BEER-P36-440ML	SPIRITS-WHISKY-P14-750ML
RTD-BEER-P32-440ML	RTD-BEER-P47-500ML	SPIRITS-WHISKY-P37-750ML
RTD-BEER-P1-330ML	RTD-BEER-P56-750ML	SPIRITS-WHISKY-P43-750ML
RTD-BEER-P20-340ML	RTD-BEER-P15-330ML	SPIRITS-WHISKY-P29-750ML
RTD-BEER-P26-440ML	RTD-BEER-P8-330ML	SPIRITS-WHISKY-P20-750ML
RTD-BEER-P43-500ML	RTD-FABS-P17-275ML	SPIRITS-WHISKY-P36-750ML
RTD-BEER-P55-750ML	RTD-FABS-P16-275ML	SPIRITS-WHISKY-P30-750ML
RTD-BEER-P3-330ML	RTD-FABS-P19-275ML	SPIRITS-WHISKY-P28-750ML
RTD-BEER-P24-340ML	RTD-FABS-P18-275ML	SPIRITS-WHISKY-P31-750ML
RTD-BEER-P31-440ML	RTD-FABS-P25-330ML	SPIRITS-WHISKY-P11-750ML
RTD-BEER-P45-500ML	RTD-FABS-P37-660ML	SPIRITS-WHISKY-P24-750ML
RTD-BEER-P50-660ML	RTD-CIDER-P5-330ML	SPIRITS-WHISKY-P44-750ML
RTD-BEER-P22-340ML	RTD-CIDER-P6-330ML	SPIRITS-WHISKY-P22-750ML
RTD-BEER-P38-440ML	RTD-CIDER-P12-440ML	SPIRITS-WHISKY-P26-750ML
RTD-BEER-P23-340ML	RTD-CIDER-P16-500ML	SPIRITS-WHISKY-P23-750ML
RTD-BEER-P39-440ML	RTD-CIDER-P2-330ML	SPIRITS-LIQUEURS-P1-1LTR
RTD-BEER-P59-750ML	RTD-CIDER-P10-330ML	SPIRITS-LIQUEURS-P9-750ML
RTD-BEER-P25-355ML	RTD-FABS-P2-250ML	SPIRITS-WHISKY-P21-750ML
RTD-BEER-P4-330ML	RTD-FABS-P4-250ML	SPIRITS-WHISKY-P50-750ML
RTD-BEER-P28-440ML	RTD-FABS-P3-250ML	SPIRITS-WHISKY-P45-750ML
RTD-BEER-P4-330ML	RTD-FABS-P1-250ML	SPIRITS-WHISKY-P53-750ML
RTD-BEER-P10-330ML	RTD-FABS-P24-300ML	SPIRITS-WHISKY-P42-750ML
RTD-BEER-P29-440ML	RTD-FABS-P29-330ML	SPIRITS-WHISKY-P52-750ML
RTD-BEER-P16-330ML	RTD-FABS-P34-440ML	SPIRITS-WHISKY-P49-750ML
RTD-BEER-P37-440ML	RTD-FABS-P23-300ML	SPIRITS-WHISKY-P51-750ML
RTD-BEER-P46-500ML	RTD-FABS-P33-440ML	SPIRITS-WHISKY-P25-750ML
RTD-BEER-P57-750ML	RTD-BEER-P5-330ML	SPIRITS-LIQUEURS-P12-750ML

TABLE A.57: *Part 1: Possible competitor set for RTD-CIDER-P6-330ML after refinement by industry insiders*

Product	Product
SPIRITS-WHISKY-P34-750ML	WINE-SPARKLING-P21-750ML
SPIRITS-WHISKY-P46-750ML	WINE-SPARKLING-P8-750ML
SPIRITS-LIQUEURS-P14-750ML	WINE-SPARKLING-P5-750ML
SPIRITS-LIQUEURS-P11-750ML	WINE-SPARKLING-P6-750ML
SPIRITS-LIQUEURS-P10-750ML	WINE-SPARKLING-P31-750ML
SPIRITS-LIQUEURS-P13-750ML	WINE-SPARKLING-P27-750ML
SPIRITS-WHISKY-P13-750ML	WINE-SPARKLING-P7-750ML
SPIRITS-WHISKY-P33-750ML	WINE-SPARKLING-P10-750ML
SPIRITS-LIQUEURS-P8-750ML	WINE-SPARKLING-P1-750ML
SPIRITS-LIQUEURS-P7-750ML	WINE-SPARKLING-P25-750ML
SPIRITS-LIQUEURS-P6-750ML	WINE-SPARKLING-P26-750ML
SPIRITS-LIQUEURS-P3-750ML	WINE-SPARKLING-P16-750ML
SPIRITS-LIQUEURS-P4-750ML	WINE-SPARKLING-P19-750ML
SPIRITS-LIQUEURS-P5-750ML	WINE-SPARKLING-P12-750ML
SPIRITS-WHISKY-P40-750ML	WINE-SPARKLING-P13-750ML
SPIRITS-WHISKY-P32-750ML	WINE-SPARKLING-P14-750ML
SPIRITS-WHISKY-P38-750ML	WINE-SPARKLING-P32-750ML
SPIRITS-WHISKY-P12-750ML	WINE-SPARKLING-P29-750ML
SPIRITS-WHISKY-P19-750ML	WINE-SPARKLING-P28-750ML
SPIRITS-WHISKY-P35-750ML	WINE-SPARKLING-P30-750ML
SPIRITS-WHISKY-P47-750ML	WINE-SPARKLING-P20-750ML
SPIRITS-WHISKY-P17-750ML	WINE-SPARKLING-P18-750ML
SPIRITS-WHISKY-P39-750ML	WINE-SPARKLING-P2-750ML
SPIRITS-LIQUEURS-P15-750ML	WINE-SPARKLING-P3-750ML
SPIRITS-WHISKY-P48-750ML	WINE-SPARKLING-P4-750ML
WINE-SPARKLING-P22-750ML	WINE-SPARKLING-P17-750ML
WINE-SPARKLING-P9-750ML	WINE-SPARKLING-P15-750ML
WINE-SPARKLING-P23-750ML	WINE-SPARKLING-P11-750ML
WINE-SPARKLING-P24-750ML	WINE-SPARKLING-P33-750ML

TABLE A.58: *Part 2: Possible competitor set for RTD-CIDER-P6-330ML after refinement by industry insiders*

Product	Product	Product
RTD-FABS-P21-275ML	SPIRITS-BRANDY-P21-750ML	SPIRITS-RUM-P6-750ML
RTD-FABS-P20-275ML	SPIRITS-WHISKY-P2-1LTR	SPIRITS-WHISKY-P34-750ML
RTD-BEER-P44-500ML	SPIRITS-WHISKY-P20-750ML	SPIRITS-WHISKY-P46-750ML
RTD-BEER-P58-750ML	SPIRITS-BRANDY-P20-750ML	SPIRITS-LIQUEURS-P14-750ML
RTD-BEER-P55-750ML	SPIRITS-GIN-P9-750ML	SPIRITS-GIN-P11-750ML
RTD-BEER-P24-340ML	SPIRITS-WHISKY-P36-750ML	SPIRITS-LIQUEURS-P11-750ML
RTD-BEER-P45-500ML	SPIRITS-WHISKY-P30-750ML	SPIRITS-LIQUEURS-P10-750ML
RTD-FABS-P6-250ML	SPIRITS-WHISKY-P28-750ML	SPIRITS-BRANDY-P2-1LTR
RTD-FABS-P22-275ML	SPIRITS-WHISKY-P31-750ML	SPIRITS-BRANDY-P9-750ML
RTD-FABS-P36-440ML	SPIRITS-GIN-P1-1LTR	SPIRITS-LIQUEURS-P13-750ML
RTD-FABS-P30-330ML	SPIRITS-GIN-P2-200ML	SPIRITS-RUM-P3-750ML
RTD-FABS-P17-275ML	SPIRITS-GIN-P3-375ML	SPIRITS-RUM-P4-750ML
RTD-CIDER-P2-330ML	SPIRITS-GIN-P4-750ML	SPIRITS-VODKA-P6-750ML
RTD-CIDER-P10-330ML	SPIRITS-WHISKY-P11-750ML	SPIRITS-BRANDY-P15-750ML
RTD-FABS-P1-250ML	SPIRITS-WHISKY-P3-1LTR	SPIRITS-BRANDY-P4-1LTR
RTD-FABS-P24-300ML	SPIRITS-WHISKY-P24-750ML	SPIRITS-BRANDY-P23-750ML
RTD-FABS-P23-300ML	SPIRITS-GIN-P8-750ML	SPIRITS-VODKA-P22-750ML
RTD-FABS-P33-440ML	SPIRITS-BRANDY-P12-750ML	SPIRITS-VODKA-P9-750ML
RTD-CIDER-P3-330ML	SPIRITS-BRANDY-P14-750ML	SPIRITS-VODKA-P11-750ML
RTD-CIDER-P13-440ML	SPIRITS-WHISKY-P44-750ML	SPIRITS-VODKA-P1-1LTR
RTD-BEER-P42-440ML	SPIRITS-BRANDY-P17-750ML	SPIRITS-VODKA-P7-750ML
SPIRITS-WHISKY-P5-1LTR	SPIRITS-GIN-P6-750ML	SPIRITS-VODKA-P10-750ML
SPIRITS-WHISKY-P41-750ML	SPIRITS-WHISKY-P4-1LTR	SPIRITS-VODKA-P8-750ML
SPIRITS-VODKA-P4-750ML	SPIRITS-WHISKY-P22-750ML	SPIRITS-WHISKY-P13-750ML
SPIRITS-LIQUEURS-P2-1LTR	SPIRITS-WHISKY-P26-750ML	SPIRITS-WHISKY-P33-750ML
SPIRITS-LIQUEURS-P16-750ML	SPIRITS-WHISKY-P23-750ML	SPIRITS-VODKA-P16-750ML
SPIRITS-RUM-P5-750ML	SPIRITS-LIQUEURS-P1-1LTR	SPIRITS-VODKA-P12-750ML
SPIRITS-WHISKY-P27-750ML	SPIRITS-LIQUEURS-P9-750ML	SPIRITS-VODKA-P14-750ML
SPIRITS-WHISKY-P15-750ML	SPIRITS-WHISKY-P8-1LTR	SPIRITS-VODKA-P15-750ML
SPIRITS-WHISKY-P1-1LTR	SPIRITS-WHISKY-P21-750ML	SPIRITS-VODKA-P13-750ML
SPIRITS-WHISKY-P16-750ML	SPIRITS-WHISKY-P50-750ML	SPIRITS-VODKA-P2-1LTR
SPIRITS-WHISKY-P18-750ML	SPIRITS-WHISKY-P45-750ML	SPIRITS-VODKA-P3-200ML
SPIRITS-VODKA-P19-750ML	SPIRITS-WHISKY-P10-1LTR	SPIRITS-VODKA-P5-750ML
SPIRITS-BRANDY-P24-750ML	SPIRITS-WHISKY-P53-750ML	SPIRITS-LIQUEURS-P8-750ML
SPIRITS-WHISKY-P14-750ML	SPIRITS-WHISKY-P42-750ML	SPIRITS-LIQUEURS-P7-750ML
SPIRITS-GIN-P7-750ML	SPIRITS-WHISKY-P52-750ML	SPIRITS-LIQUEURS-P6-750ML
SPIRITS-WHISKY-P37-750ML	SPIRITS-WHISKY-P49-750ML	SPIRITS-LIQUEURS-P3-750ML
SPIRITS-WHISKY-P43-750ML	SPIRITS-WHISKY-P51-750ML	SPIRITS-LIQUEURS-P4-750ML
SPIRITS-CANE-P1-1LTR	SPIRITS-WHISKY-P9-1LTR	SPIRITS-LIQUEURS-P5-750ML
SPIRITS-CANE-P2-750ML	SPIRITS-WHISKY-P25-750ML	SPIRITS-GIN-P10-750ML
SPIRITS-RUM-P2-750ML	SPIRITS-LIQUEURS-P12-750ML	SPIRITS-GIN-P5-750ML
SPIRITS-RUM-P1-1LTR	SPIRITS-BRANDY-P1-1LTR	SPIRITS-WHISKY-P40-750ML
SPIRITS-RUM-P7-750ML	SPIRITS-BRANDY-P5-375ML	SPIRITS-WHISKY-P32-750ML
SPIRITS-WHISKY-P29-750ML	SPIRITS-BRANDY-P7-750ML	SPIRITS-WHISKY-P38-750ML
SPIRITS-VODKA-P17-750ML	SPIRITS-BRANDY-P22-750ML	SPIRITS-WHISKY-P12-750ML
SPIRITS-BRANDY-P19-750ML	SPIRITS-BRANDY-P6-750ML	SPIRITS-WHISKY-P19-750ML
SPIRITS-BRANDY-P18-750ML	SPIRITS-BRANDY-P16-750ML	SPIRITS-WHISKY-P35-750ML
SPIRITS-VODKA-P18-750ML	SPIRITS-BRANDY-P8-750ML	SPIRITS-WHISKY-P7-1LTR
SPIRITS-VODKA-P20-750ML	SPIRITS-VODKA-P21-750ML	SPIRITS-WHISKY-P47-750ML

TABLE A.59: *Part 1: Possible competitor set for SPIRITS-LIQUEURS-P16-750ML after refinement by industry insiders*

Product	Product
SPIRITS-WHISKY-P17-750ML	WINE-SPARKLING-P31-750ML
SPIRITS-BRANDY-P11-750ML	WINE-SPARKLING-P27-750ML
SPIRITS-BRANDY-P13-750ML	WINE-SPARKLING-P1-750ML
SPIRITS-BRANDY-P3-1LTR	WINE-RED-P97-750ML
SPIRITS-BRANDY-P10-750ML	WINE-WHITE-P62-750ML
SPIRITS-WHISKY-P6-1LTR	WINE-RED-P96-750ML
SPIRITS-WHISKY-P39-750ML	WINE-RED-P99-750ML
SPIRITS-LIQUEURS-P15-750ML	WINE-RED-P94-750ML
SPIRITS-WHISKY-P48-750ML	WINE-WHITE-P61-750ML
WINE-BIB-P79-5LTR	WINE-RED-P98-750ML
WINE-BIB-P43-3LTR	WINE-BIB-P61-5LTR
WINE-BIB-P82-5LTR	WINE-BIB-P60-5LTR
WINE-BIB-P78-5LTR	WINE-BIB-P62-5LTR
WINE-PERLE-P6-300ML	WINE-BIB-P73-5LTR
WINE-PERLE-P5-300ML	WINE-RED-P106-750ML
WINE-RED-P22-750ML	WINE-SPARKLING-P2-750ML
WINE-PERLE-P8-1.5LTR	WINE-BIB-P44-3LTR
WINE-RED-P23-750ML	WINE-BIB-P15-1LTR
WINE-ROSE-P17-750ML	WINE-BIB-P40-3LTR
WINE-BIB-P83-5LTR	WINE-BIB-P53-500ML
WINE-BIB-P84-5LTR	WINE-FORTIFIED-P3-200ML
WINE-BIB-P48-3LTR	WINE-FORTIFIED-P6-375ML
WINE-WHITE-P54-750ML	WINE-FORTIFIED-P16-750ML
WINE-WHITE-P57-750ML	WINE-RED-P32-750ML
WINE-WHITE-P52-750ML	WINE-WHITE-P18-750ML
WINE-WHITE-P43-750ML	WINE-RED-P24-750ML

TABLE A.60: *Part 2: Possible competitor set for SPIRITS-LIQUEURS-P16-750ML after refinement by industry insiders*

Product	Product	Product
RTD-FABS-P21-275ML	SPIRITS-BRANDY-P23-750ML	WINE-BIB-P64-5LTR
RTD-FABS-P20-275ML	SPIRITS-VODKA-P1-1LTR	WINE-RED-P66-750ML
RTD-FABS-P27-330ML	SPIRITS-WHISKY-P13-750ML	WINE-RED-P64-750ML
RTD-BEER-P44-500ML	SPIRITS-VODKA-P2-1LTR	WINE-RED-P80-750ML
RTD-BEER-P58-750ML	SPIRITS-VODKA-P3-200ML	WINE-RED-P72-750ML
RTD-BEER-P55-750ML	SPIRITS-VODKA-P5-750ML	WINE-RED-P85-750ML
RTD-BEER-P24-340ML	SPIRITS-WHISKY-P7-1LTR	WINE-ROSE-P13-750ML
RTD-BEER-P45-500ML	SPIRITS-WHISKY-P17-750ML	WINE-WHITE-P10-750ML
RTD-FABS-P6-250ML	SPIRITS-BRANDY-P3-1LTR	WINE-RED-P57-750ML
RTD-FABS-P22-275ML	SPIRITS-WHISKY-P6-1LTR	WINE-WHITE-P37-750ML
RTD-FABS-P36-440ML	WINE-RED-P79-1.5LTR	WINE-SPARKLING-P9-750ML
RTD-FABS-P30-330ML	WINE-BIB-P42-3LTR	WINE-ROSE-P17-750ML
RTD-FABS-P17-275ML	WINE-BIB-P79-5LTR	WINE-WHITE-P58-750ML
RTD-CIDER-P2-330ML	WINE-RED-P71-750ML	WINE-WHITE-P59-750ML
RTD-CIDER-P10-330ML	WINE-BIB-P82-1.5LTR	WINE-WHITE-P24-750ML
RTD-FABS-P1-250ML	WINE-BIB-P43-3LTR	WINE-FORTIFIED-P1-1LTR
RTD-FABS-P24-300ML	WINE-BIB-P82-5LTR	WINE-FORTIFIED-P8-750ML
RTD-FABS-P23-300ML	WINE-ROSE-P12-750ML	WINE-SPARKLING-P23-750ML
RTD-FABS-P33-440ML	WINE-BIB-P41-3LTR	WINE-SPARKLING-P24-750ML
RTD-CIDER-P14-440ML	WINE-BIB-P78-5LTR	WINE-BIB-P81-5LTR
RTD-CIDER-P3-330ML	WINE-PERLE-P6-300ML	WINE-BIB-P80-5LTR
RTD-CIDER-P13-440ML	WINE-PERLE-P5-300ML	WINE-RED-P95-750ML
SPIRITS-WHISKY-P5-1LTR	WINE-SPARKLING-P22-750ML	WINE-RED-P40-750ML
SPIRITS-LIQUEURS-P16-750ML	WINE-RED-P29-750ML	WINE-RED-P62-750ML
SPIRITS-WHISKY-P15-750ML	WINE-RED-P19-750ML	WINE-SPARKLING-P21-750ML
SPIRITS-WHISKY-P1-1LTR	WINE-RED-P22-750ML	WINE-RED-P65-750ML
SPIRITS-WHISKY-P16-750ML	WINE-WHITE-P22-750ML	WINE-WHITE-P23-750ML
SPIRITS-WHISKY-P14-750ML	WINE-WHITE-P36-750ML	WINE-RED-P50-750ML
SPIRITS-WHISKY-P2-1LTR	WINE-WHITE-P33-750ML	WINE-RED-P31-750ML
SPIRITS-GIN-P2-200ML	WINE-WHITE-P28-750ML	WINE-FORTIFIED-P10-750ML
SPIRITS-GIN-P3-375ML	WINE-RED-P48-750ML	WINE-PERLE-P1-1LTR
SPIRITS-GIN-P4-750ML	WINE-RED-P51-750ML	WINE-WHITE-P12-750ML
SPIRITS-WHISKY-P11-750ML	WINE-PERLE-P8-1.5LTR	WINE-RED-P26-750ML
SPIRITS-WHISKY-P3-1LTR	WINE-PERLE-P2-1LTR	WINE-RED-P79-750ML
SPIRITS-WHISKY-P24-750ML	WINE-PERLE-P10-750ML	WINE-WHITE-P1-5LTR
SPIRITS-BRANDY-P12-750ML	WINE-PERLE-P7-1.5LTR	WINE-WHITE-P62-750ML
SPIRITS-WHISKY-P4-1LTR	WINE-PERLE-P11-750ML	WINE-WHITE-P56-750ML
SPIRITS-WHISKY-P22-750ML	WINE-PERLE-P9-1.5LTR	WINE-RED-P89-750ML
SPIRITS-WHISKY-P23-750ML	WINE-PERLE-P12-750ML	WINE-ROSE-P2-1.5LTR
SPIRITS-WHISKY-P8-1LTR	WINE-RED-P59-750ML	WINE-RED-P99-750ML
SPIRITS-WHISKY-P21-750ML	WINE-RED-P39-750ML	WINE-ROSE-P15-750ML
SPIRITS-WHISKY-P10-1LTR	WINE-ROSE-P10-750ML	WINE-WHITE-P50-750ML
SPIRITS-WHISKY-P53-750ML	WINE-WHITE-P39-750ML	WINE-RED-P86-750ML
SPIRITS-WHISKY-P9-1LTR	WINE-RED-P28-750ML	WINE-WHITE-P49-750ML
SPIRITS-WHISKY-P25-750ML	WINE-RED-P23-750ML	WINE-WHITE-P67-750ML
SPIRITS-BRANDY-P7-750ML	WINE-BIB-P68-5LTR	WINE-BIB-P83-5LTR
SPIRITS-BRANDY-P8-750ML	WINE-BIB-P72-5LTR	WINE-BIB-P84-1.5LTR
SPIRITS-BRANDY-P2-1LTR	WINE-SPARKLING-P5-750ML	WINE-BIB-P84-5LTR
SPIRITS-BRANDY-P4-1LTR	WINE-BIB-P65-5LTR	WINE-WHITE-P68-750ML

TABLE A.61: *Part 1: Possible competitor set for WINE-RED-P102-750ML after refinement by industry insiders*

Product	Product	Product
WINE-RED-P101-750ML	WINE-ROSE-P9-750ML	WINE-BIB-P5-1LTR
WINE-ROSE-P4-1.5LTR	WINE-WHITE-P25-750ML	WINE-BIB-P23-3LTR
WINE-ROSE-P4-1.5LTR	WINE-WHITE-P32-750ML	WINE-BIB-P60-5LTR
WINE-SPARKLING-P32-750ML	WINE-ROSE-P11-750ML	WINE-BIB-P27-3LTR
WINE-ROSE-P20-750ML	WINE-RED-P58-750ML	WINE-BIB-P62-5LTR
WINE-WHITE-P42-750ML	WINE-WHITE-P34-750ML	WINE-BIB-P2-1LTR
WINE-BIB-P47-3LTR	WINE-WHITE-P13-750ML	WINE-BIB-P22-3LTR
WINE-SPARKLING-P30-750ML	WINE-SPARKLING-P31-750ML	WINE-BIB-P55-5LTR
WINE-RED-P84-750ML	WINE-SPARKLING-P27-750ML	WINE-BIB-P1-1LTR
WINE-BIB-P46-3LTR	WINE-SPARKLING-P7-750ML	WINE-BIB-P24-3LTR
WINE-WHITE-P55-750ML	WINE-SPARKLING-P10-750ML	WINE-BIB-P59-5LTR
WINE-BIB-P45-3LTR	WINE-SPARKLING-P1-750ML	WINE-BIB-P57-5LTR
WINE-RED-P87-750ML	WINE-SPARKLING-P25-750ML	WINE-RED-P11-750ML
WINE-BIB-P3-1LTR	WINE-SPARKLING-P26-750ML	WINE-BIB-P56-5LTR
WINE-BIB-P48-3LTR	WINE-RED-P35-750ML	WINE-BIB-P4-1LTR
WINE-WHITE-P54-750ML	WINE-ROSE-P8-750ML	WINE-BIB-P28-3LTR
WINE-RED-P90-750ML	WINE-RED-P20-750ML	WINE-BIB-P73-5LTR
WINE-RED-P8-750ML	WINE-WHITE-P38-750ML	WINE-RED-P10-750ML
WINE-RED-P88-750ML	WINE-RED-P47-750ML	WINE-BIB-P25-3LTR
WINE-ROSE-P16-750ML	WINE-WHITE-P26-750ML	WINE-BIB-P58-5LTR
WINE-RED-P92-750ML	WINE-SPARKLING-P16-750ML	WINE-RED-P9-750ML
WINE-WHITE-P57-750ML	WINE-SPARKLING-P19-750ML	WINE-BIB-P54-5LTR
WINE-WHITE-P52-750ML	WINE-WHITE-P27-750ML	WINE-RED-P104-750ML
WINE-RED-P91-750ML	WINE-SPARKLING-P12-750ML	WINE-RED-P106-750ML
WINE-RED-P56-750ML	WINE-SPARKLING-P13-750ML	WINE-RED-P103-750ML
WINE-RED-P27-750ML	WINE-SPARKLING-P14-750ML	WINE-WHITE-P64-750ML
WINE-WHITE-P15-750ML	WINE-RED-P97-750ML	WINE-ROSE-P21-750ML
WINE-RED-P30-750ML	WINE-RED-P100-750ML	WINE-WHITE-P65-750ML
WINE-RED-P36-750ML	WINE-WHITE-P63-750ML	WINE-RED-P102-750ML
WINE-WHITE-P19-750ML	WINE-RED-P96-750ML	WINE-BIB-P67-5LTR
WINE-RED-P34-750ML	WINE-RED-P94-750ML	WINE-BIB-P71-5LTR
WINE-RED-P55-750ML	WINE-WHITE-P61-750ML	WINE-BIB-P70-5LTR
WINE-WHITE-P40-750ML	WINE-RED-P98-750ML	WINE-BIB-P63-5LTR
WINE-WHITE-P35-750ML	WINE-RED-P38-750ML	WINE-BIB-P69-5LTR
WINE-BIB-P49-3LTR	WINE-RED-P42-750ML	WINE-BIB-P74-5LTR
WINE-RED-P93-750ML	WINE-WHITE-P14-750ML	WINE-PERLE-P3-1LTR
WINE-BIB-P50-3LTR	WINE-RED-P44-750ML	WINE-PERLE-P4-2LTR
WINE-ROSE-P19-750ML	WINE-RED-P43-750ML	WINE-PERLE-P13-750ML
WINE-WHITE-P60-750ML	WINE-RED-P54-750ML	WINE-SPARKLING-P20-750ML
WINE-ROSE-P18-750ML	WINE-RED-P76-750ML	WINE-RED-P49-750ML
WINE-RED-P4-1.5LTR	WINE-RED-P12-750ML	WINE-SPARKLING-P18-750ML
WINE-WHITE-P2-1.5LTR	WINE-SPARKLING-P29-750ML	WINE-WHITE-P20-750ML
WINE-WHITE-P2-5LTR	WINE-SPARKLING-P28-750ML	WINE-FORTIFIED-P7-750ML
WINE-WHITE-P43-750ML	WINE-FORTIFIED-P11-750ML	WINE-SPARKLING-P2-750ML
WINE-ROSE-P14-750ML	WINE-FORTIFIED-P9-750ML	WINE-SPARKLING-P3-750ML
WINE-SPARKLING-P8-750ML	WINE-FORTIFIED-P4-2LTR	WINE-RED-P73-750ML
WINE-SPARKLING-P6-750ML	WINE-FORTIFIED-P14-750ML	WINE-WHITE-P53-750ML
WINE-RED-P83-750ML	WINE-BIB-P26-3LTR	WINE-WHITE-P46-750ML
WINE-WHITE-P51-750ML	WINE-BIB-P61-5LTR	WINE-RED-P74-750ML

TABLE A.62: *Part 2: Possible competitor set for WINE-RED-P102-750ML after refinement by industry insiders*

Product	Product	Product
WINE-WHITE-P48-750ML	WINE-BIB-P40-3LTR	WINE-FORTIFIED-P13-750ML
WINE-RED-P82-750ML	WINE-BIB-P53-500ML	WINE-RED-P24-750ML
WINE-RED-P45-750ML	WINE-SPARKLING-P4-750ML	WINE-WHITE-P11-750ML
WINE-WHITE-P41-750ML	WINE-BIB-P12-1LTR	WINE-ROSE-P7-750ML
WINE-RED-P52-750ML	WINE-BIB-P13-1LTR	WINE-WHITE-P47-750ML
WINE-RED-P69-750ML	WINE-RED-P16-750ML	WINE-RED-P53-750ML
WINE-WHITE-P8-750ML	WINE-FORTIFIED-P2-1LTR	WINE-RED-P5-750ML
WINE-BIB-P6-1LTR	WINE-FORTIFIED-P3-200ML	WINE-WHITE-P4-750ML
WINE-BIB-P16-2LTR	WINE-FORTIFIED-P5-2LTR	WINE-RED-P7-750ML
WINE-BIB-P29-3LTR	WINE-FORTIFIED-P6-375ML	WINE-RED-P6-750ML
WINE-RED-P14-750ML	WINE-FORTIFIED-P16-750ML	WINE-WHITE-P3-750ML
WINE-RED-P1-1.5LTR	WINE-FORTIFIED-P15-750ML	WINE-FORTIFIED-P12-750ML
WINE-RED-P21-750ML	WINE-SPARKLING-P17-750ML	
WINE-ROSE-P6-750ML	WINE-SPARKLING-P15-750ML	
WINE-RED-P2-1.5LTR	WINE-RED-P68-750ML	
WINE-RED-P25-750ML	WINE-RED-P37-750ML	
WINE-WHITE-P9-750ML	WINE-WHITE-P30-750ML	
WINE-BIB-P19-2LTR	WINE-WHITE-P21-750ML	
WINE-BIB-P34-3LTR	WINE-WHITE-P29-750ML	
WINE-WHITE-P6-750ML	WINE-RED-P32-750ML	
WINE-BIB-P21-2LTR	WINE-RED-P46-750ML	
WINE-WHITE-P7-750ML	WINE-WHITE-P18-750ML	
WINE-BIB-P14-1LTR	WINE-RED-P41-750ML	
WINE-BIB-P39-3LTR	WINE-SPARKLING-P11-750ML	
WINE-BIB-P9-1LTR	WINE-BIB-P77-5LTR	
WINE-BIB-P37-3LTR	WINE-BIB-P76-5LTR	
WINE-BIB-P32-3LTR	WINE-BIB-P75-5LTR	
WINE-BIB-P11-1LTR	WINE-RED-P67-750ML	
WINE-BIB-P30-3LTR	WINE-RED-P81-750ML	
WINE-BIB-P35-3LTR	WINE-WHITE-P45-750ML	
WINE-BIB-P7-1LTR	WINE-RED-P75-750ML	
WINE-BIB-P17-2LTR	WINE-RED-P78-750ML	
WINE-BIB-P31-3LTR	WINE-WHITE-P44-750ML	
WINE-RED-P13-750ML	WINE-RED-P77-750ML	
WINE-BIB-P38-3LTR	WINE-RED-P60-750ML	
WINE-RED-P18-750ML	WINE-BIB-P66-5LTR	
WINE-ROSE-P1-1.5LTR	WINE-RED-P70-750ML	
WINE-BIB-P44-3LTR	WINE-WHITE-P31-750ML	
WINE-ROSE-P5-750ML	WINE-RED-P63-750ML	
WINE-BIB-P10-1LTR	WINE-WHITE-P17-750ML	
WINE-BIB-P36-3LTR	WINE-BIB-P51-3LTR	
WINE-RED-P15-750ML	WINE-RED-P105-750ML	
WINE-BIB-P18-2LTR	WINE-BIB-P52-3LTR	
WINE-RED-P17-750ML	WINE-WHITE-P66-750ML	
WINE-BIB-P20-2LTR	WINE-RED-P33-750ML	
WINE-BIB-P33-3LTR	WINE-RED-P3-1.5LTR	
WINE-WHITE-P5-750ML	WINE-RED-P61-750ML	
WINE-BIB-P8-1LTR	WINE-WHITE-P16-750ML	
WINE-BIB-P15-1LTR	WINE-SPARKLING-P33-750ML	

TABLE A.63: *Part 3: Possible competitor set for WINE-RED-P102-750ML after refinement by industry insiders*

A.5.1 Variable refinement using stepwise regression analysis

Product	Product
Results Batch 1	**Results Batch 3**
RTD-CIDER-P18-660ML	SPIRITS-LIQUEURS-P16-750ML
WINE-RED-P79-1.5LTR	RTD-FABS-P11-275ML
WINE-BIB-P79-5LTR	RTD-FABS-P9-275ML
SPIRITS-LIQUEURS-P16-750ML	RTD-FABS-P38-660ML
RTD-BEER-P14-330ML	RTD-BEER-P44-500ML
RTD-BEER-P34-440ML	RTD-BEER-P1-330ML
RTD-FABS-P11-275ML	RTD-BEER-P31-440ML
RTD-FABS-P9-275ML	RTD-BEER-P50-660ML
RTD-FABS-P38-660ML	RTD-BEER-P39-440ML
WINE-FORTIFIED-P1-1LTR	WINE-WHITE-P1-5LTR
RTD-FABS-P27-330ML	RTD-BEER-P28-440ML
RTD-BEER-P2-330ML	RTD-BEER-P29-440ML
RTD-BEER-P44-500ML	RTD-BEER-P57-750ML
RTD-BEER-P58-750ML	RTD-BEER-P48-5LTR
RTD-BEER-P1-330ML	RTD-CIDER-P17-660ML
RTD-BEER-P43-500ML	SPIRITS-WHISKY-P8-1LTR
RTD-BEER-P55-750ML	SPIRITS-WHISKY-P53-750ML
	SPIRITS-WHISKY-P9-1LTR
Results Batch 2	SPIRITS-BRANDY-P1-1LTR
RTD-CIDER-P18-660ML	RTD-BEER-P47-500ML
SPIRITS-LIQUEURS-P16-750ML	RTD-BEER-P56-750ML
RTD-BEER-P14-330ML	WINE-BIB-P55-5LTR
RTD-FABS-P11-275ML	WINE-BIB-P56-5LTR
RTD-FABS-P9-275ML	RTD-CIDER-P18-660ML
RTD-FABS-P38-660ML	SPIRITS-LIQUEURS-P16-750ML
RTD-FABS-P27-330ML	RTD-FABS-P11-275ML
RTD-BEER-P44-500ML	RTD-FABS-P9-275ML
RTD-BEER-P1-330ML	RTD-FABS-P38-660ML
RTD-BEER-P3-330ML	RTD-BEER-P44-500ML
RTD-BEER-P31-440ML	RTD-BEER-P31-440ML
RTD-BEER-P50-660ML	RTD-BEER-P50-660ML
RTD-BEER-P22-340ML	RTD-BEER-P39-440ML
RTD-BEER-P39-440ML	WINE-WHITE-P1-5LTR
WINE-WHITE-P1-5LTR	RTD-BEER-P28-440ML
WINE-ROSE-P2-1.5LTR	RTD-BEER-P29-440ML
RTD-BEER-P28-440ML	RTD-BEER-P57-750ML
RTD-BEER-P29-440ML	RTD-BEER-P48-5LTR
SPIRITS-WHISKY-P11-750ML	RTD-CIDER-P17-660ML
RTD-BEER-P16-330ML	SPIRITS-WHISKY-P8-1LTR
RTD-BEER-P57-750ML	SPIRITS-WHISKY-P53-750ML
RTD-BEER-P48-5LTR	SPIRITS-WHISKY-P9-1LTR
RTD-CIDER-P9-330ML	SPIRITS-BRANDY-P1-1LTR
RTD-CIDER-P15-440ML	RTD-BEER-P47-500ML
	RTD-BEER-P56-750ML
Results Batch 3	WINE-BIB-P55-5LTR
RTD-CIDER-P18-660ML	WINE-BIB-P56-5LTR

TABLE A.64: *Part 1: RTD-CIDER-P18-660ML significant independent variables results as each batch of independent variables are added*

Product	Product
Results Batch 4	**Results Batch 6**
WINE-BIB-P69-5LTR	SPIRITS-LIQUEURS-P16-750ML
WINE-BIB-P74-5LTR	RTD-FABS-P11-275ML
WINE-PERLE-P4-2LTR	RTD-FABS-P38-660ML
RTD-BEER-P8-330ML	RTD-BEER-P44-500ML
RTD-FABS-P37-660ML	RTD-BEER-P31-440ML
WINE-BIB-P19-2LTR	RTD-BEER-P39-440ML
WINE-BIB-P21-2LTR	WINE-WHITE-P1-5LTR
WINE-BIB-P30-3LTR	RTD-BEER-P28-440ML
WINE-BIB-P35-3LTR	RTD-BEER-P29-440ML
WINE-BIB-P18-2LTR	RTD-BEER-P57-750ML
SPIRITS-VODKA-P1-1LTR	RTD-BEER-P48-5LTR
	SPIRITS-WHISKY-P8-1LTR
Results Batch 5	SPIRITS-WHISKY-P53-750ML
RTD-CIDER-P18-660ML	SPIRITS-BRANDY-P1-1LTR
SPIRITS-LIQUEURS-P16-750ML	RTD-BEER-P47-500ML
RTD-FABS-P11-275ML	WINE-BIB-P55-5LTR
RTD-FABS-P38-660ML	WINE-BIB-P56-5LTR
RTD-BEER-P44-500ML	WINE-PERLE-P4-2LTR
RTD-BEER-P31-440ML	RTD-FABS-P37-660ML
RTD-BEER-P39-440ML	WINE-BIB-P19-2LTR
WINE-WHITE-P1-5LTR	WINE-BIB-P30-3LTR
RTD-BEER-P28-440ML	WINE-BIB-P18-2LTR
RTD-BEER-P29-440ML	RTD-FABS-P4-250ML
RTD-BEER-P57-750ML	RTD-FABS-P3-250ML
RTD-BEER-P48-5LTR	RTD-CIDER-P14-440ML
RTD-CIDER-P17-660ML	RTD-CIDER-P3-330ML
SPIRITS-WHISKY-P8-1LTR	RTD-BEER-P41-440ML
SPIRITS-WHISKY-P53-750ML	RTD-BEER-P18-330ML
SPIRITS-BRANDY-P1-1LTR	
RTD-BEER-P47-500ML	
WINE-BIB-P55-5LTR	
WINE-BIB-P56-5LTR	
WINE-PERLE-P4-2LTR	
RTD-FABS-P37-660ML	
WINE-BIB-P19-2LTR	
WINE-BIB-P30-3LTR	
WINE-BIB-P18-2LTR	
SPIRITS-VODKA-P1-1LTR	
RTD-CIDER-P16-500ML	
RTD-FABS-P4-250ML	
RTD-FABS-P3-250ML	
RTD-CIDER-P14-440ML	
RTD-CIDER-P3-330ML	
RTD-BEER-P12-330ML	
WINE-BIB-P66-5LTR	
RTD-BEER-P19-330ML	
RTD-BEER-P41-440ML	
RTD-BEER-P18-330ML	

TABLE A.65: *Part 2: RTD-CIDER-P18-660ML significant independent variables results as each batch of independent variables are added*

Product	Product	Product
Results Batch 1	**Results Batch 3**	**Results Batch 5**
RTD-CIDER-P6-330ML	RTD-BEER-P1-330ML	RTD-CIDER-P6-330ML
RTD-BEER-P35-440ML	RTD-BEER-P24-340ML	RTD-BEER-P35-440ML
RTD-BEER-P13-330ML	RTD-BEER-P39-440ML	RTD-CIDER-P6-330ML
RTD-BEER-P34-440ML	RTD-BEER-P48-5LTR	RTD-BEER-P35-440ML
RTD-FABS-P11-275ML	RTD-FABS-P30-330ML	RTD-FABS-P11-275ML
RTD-FABS-P20-275ML	SPIRITS-LIQUEURS-P9-750ML	RTD-FABS-P38-660ML
RTD-FABS-P13-275ML	WINE-SPARKLING-P31-750ML	RTD-BEER-P27-440ML
RTD-FABS-P38-660ML	WINE-SPARKLING-P7-750ML	RTD-BEER-P58-750ML
RTD-FABS-P27-330ML	WINE-SPARKLING-P10-750ML	RTD-BEER-P24-340ML
RTD-FABS-P7-275ML	SPIRITS-WHISKY-P45-750ML	RTD-BEER-P48-5LTR
RTD-BEER-P21-340ML	SPIRITS-WHISKY-P42-750ML	WINE-SPARKLING-P31-750ML
RTD-BEER-P27-440ML	SPIRITS-WHISKY-P49-750ML	WINE-SPARKLING-P7-750ML
RTD-BEER-P58-750ML	SPIRITS-LIQUEURS-P12-750ML	SPIRITS-WHISKY-P45-750ML
RTD-BEER-P1-330ML	RTD-BEER-P47-500ML	SPIRITS-WHISKY-P42-750ML
RTD-BEER-P20-340ML	SPIRITS-WHISKY-P34-750ML	SPIRITS-WHISKY-P49-750ML
RTD-BEER-P3-330ML	SPIRITS-LIQUEURS-P11-750ML	RTD-BEER-P47-500ML
RTD-BEER-P24-340ML		SPIRITS-WHISKY-P34-750ML
	Results Batch 4	SPIRITS-LIQUEURS-P11-750ML
Results Batch 2	RTD-CIDER-P6-330ML	WINE-SPARKLING-P4-750ML
RTD-CIDER-P6-330ML	RTD-BEER-P35-440ML	RTD-CIDER-P2-330ML
RTD-BEER-P35-440ML	RTD-FABS-P11-275ML	SPIRITS-LIQUEURS-P7-750ML
RTD-BEER-P13-330ML	RTD-FABS-P38-660ML	RTD-FABS-P1-250ML
RTD-BEER-P34-440ML	RTD-BEER-P27-440ML	RTD-FABS-P24-300ML
RTD-FABS-P11-275ML	RTD-BEER-P58-750ML	RTD-FABS-P23-300ML
RTD-FABS-P20-275ML	RTD-BEER-P24-340ML	RTD-FABS-P33-440ML
RTD-FABS-P13-275ML	RTD-BEER-P48-5LTR	RTD-BEER-P9-330ML
RTD-FABS-P38-660ML	WINE-SPARKLING-P31-750ML	SPIRITS-LIQUEURS-P5-750ML
RTD-BEER-P27-440ML	WINE-SPARKLING-P7-750ML	RTD-BEER-P12-330ML
RTD-BEER-P58-750ML	SPIRITS-WHISKY-P45-750ML	SPIRITS-WHISKY-P32-750ML
RTD-BEER-P1-330ML	SPIRITS-WHISKY-P42-750ML	
RTD-BEER-P24-340ML	SPIRITS-WHISKY-P49-750ML	
RTD-BEER-P39-440ML	RTD-BEER-P47-500ML	
SPIRITS-WHISKY-P24-750ML	SPIRITS-WHISKY-P34-750ML	
RTD-BEER-P48-5LTR	SPIRITS-LIQUEURS-P11-750ML	
RTD-FABS-P6-250ML	WINE-SPARKLING-P4-750ML	
RTD-FABS-P30-330ML	RTD-CIDER-P2-330ML	
SPIRITS-LIQUEURS-P1-1LTR	SPIRITS-LIQUEURS-P7-750ML	
SPIRITS-LIQUEURS-P9-750ML	RTD-FABS-P1-250ML	
	RTD-FABS-P24-300ML	
Results Batch 3	RTD-FABS-P23-300ML	
RTD-CIDER-P6-330ML	RTD-FABS-P33-440ML	
RTD-BEER-P35-440ML	RTD-BEER-P9-330ML	
RTD-FABS-P11-275ML	SPIRITS-LIQUEURS-P5-750ML	
RTD-FABS-P13-275ML	RTD-BEER-P12-330ML	
RTD-FABS-P38-660ML	SPIRITS-WHISKY-P32-750ML	
RTD-BEER-P27-440ML		
RTD-BEER-P58-750ML		

TABLE A.66: *RTD-CIDER-P6-330ML significant independent variables results as each batch of independent variables are added*

Product	Product
Results Batch 1	**Results Batch 3**
WINE-RED-P102-750ML	SPIRITS-GIN-P4-750ML
SPIRITS-WHISKY-P5-1LTR	SPIRITS-WHISKY-P11-750ML
WINE-BIB-P43-3LTR	SPIRITS-BRANDY-P12-750ML
WINE-BIB-P41-3LTR	
WINE-BIB-P78-5LTR	**Results Batch 4**
WINE-SPARKLING-P22-750ML	WINE-RED-P102-750ML
WINE-RED-P19-750ML	WINE-RED-P66-750ML
WINE-RED-P22-750ML	RTD-BEER-P58-750ML
WINE-PERLE-P8-1.5LTR	RTD-BEER-P24-340ML
WINE-PERLE-P7-1.5LTR	RTD-BEER-P45-500ML
WINE-RED-P59-750ML	WINE-RED-P50-750ML
WINE-RED-P39-750ML	WINE-FORTIFIED-P10-750ML
WINE-RED-P23-750ML	WINE-BIB-P45-3LTR
WINE-RED-P66-750ML	WINE-RED-P87-750ML
	SPIRITS-GIN-P4-750ML
Results Batch 2	SPIRITS-WHISKY-P11-750ML
WINE-RED-P102-750ML	SPIRITS-BRANDY-P12-750ML
SPIRITS-WHISKY-P5-1LTR	RTD-FABS-P36-440ML
WINE-PERLE-P8-1.5LTR	RTD-FABS-P30-330ML
WINE-RED-P66-750ML	SPIRITS-WHISKY-P21-750ML
WINE-RED-P72-750ML	WINE-SPARKLING-P1-750ML
WINE-WHITE-P58-750ML	WINE-SPARKLING-P25-750ML
RTD-FABS-P27-330ML	WINE-SPARKLING-P19-750ML
RTD-BEER-P58-750ML	WINE-RED-P44-750ML
RTD-BEER-P24-340ML	
RTD-BEER-P45-500ML	**Results Batch 5**
WINE-RED-P50-750ML	WINE-RED-P102-750ML
WINE-RED-P31-750ML	WINE-RED-P66-750ML
WINE-FORTIFIED-P10-750ML	RTD-BEER-P58-750ML
WINE-RED-P79-750ML	RTD-BEER-P24-340ML
WINE-BIB-P84-5LTR	RTD-BEER-P45-500ML
	WINE-RED-P50-750ML
Results Batch 3	WINE-FORTIFIED-P10-750ML
WINE-RED-P102-750ML	WINE-BIB-P45-3LTR
SPIRITS-WHISKY-P5-1LTR	WINE-RED-P87-750ML
WINE-RED-P66-750ML	RTD-FABS-P36-440ML
WINE-RED-P72-750ML	RTD-FABS-P30-330ML
RTD-FABS-P27-330ML	SPIRITS-WHISKY-P21-750ML
RTD-BEER-P58-750ML	WINE-SPARKLING-P1-750ML
RTD-BEER-P24-340ML	WINE-SPARKLING-P25-750ML
RTD-BEER-P45-500ML	WINE-SPARKLING-P19-750ML
WINE-RED-P50-750ML	WINE-RED-P44-750ML
WINE-FORTIFIED-P10-750ML	WINE-RED-P54-750ML
WINE-RED-P79-750ML	WINE-FORTIFIED-P11-750ML
WINE-BIB-P45-3LTR	WINE-BIB-P55-5LTR
WINE-RED-P87-750ML	WINE-RED-P11-750ML

TABLE A.67: *Part 1: WINE-RED-P102-750ML significant independent variables results as each batch of independent variables are added*

Product	Product
Results Batch 5	**Results Batch 7**
WINE-BIB-P28-3LTR	SPIRITS-VODKA-P1-1LTR
WINE-RED-P9-750ML	WINE-FORTIFIED-P6-375ML
WINE-BIB-P54-5LTR	WINE-FORTIFIED-P16-750ML
WINE-BIB-P67-5LTR	WINE-RED-P68-750ML
	RTD-FABS-P1-250ML
Results Batch 6: Significant Independent Variables	RTD-FABS-P23-300ML
WINE-RED-P102-750ML	WINE-RED-P37-750ML
WINE-RED-P66-750ML	WINE-RED-P32-750ML
RTD-BEER-P58-750ML	
RTD-BEER-P24-340ML	**Results Batch 8**
RTD-BEER-P45-500ML	WINE-RED-P102-750ML
WINE-FORTIFIED-P10-750ML	RTD-BEER-P58-750ML
WINE-RED-P87-750ML	RTD-BEER-P24-340ML
RTD-FABS-P36-440ML	RTD-BEER-P45-500ML
RTD-FABS-P30-330ML	WINE-FORTIFIED-P10-750ML
WINE-SPARKLING-P1-750ML	WINE-RED-P87-750ML
WINE-SPARKLING-P25-750ML	RTD-FABS-P36-440ML
WINE-SPARKLING-P19-750ML	RTD-FABS-P30-330ML
WINE-RED-P44-750ML	WINE-SPARKLING-P1-750ML
WINE-RED-P54-750ML	WINE-RED-P54-750ML
WINE-FORTIFIED-P11-750ML	WINE-SPARKLING-P2-750ML
WINE-BIB-P55-5LTR	SPIRITS-BRANDY-P23-750ML
WINE-BIB-P28-3LTR	WINE-BIB-P44-3LTR
WINE-BIB-P54-5LTR	SPIRITS-VODKA-P1-1LTR
WINE-RED-P49-750ML	WINE-FORTIFIED-P6-375ML
WINE-SPARKLING-P2-750ML	WINE-FORTIFIED-P16-750ML
SPIRITS-BRANDY-P23-750ML	WINE-RED-P68-750ML
	RTD-FABS-P23-300ML
Results Batch 7	WINE-RED-P37-750ML
WINE-RED-P102-750ML	WINE-RED-P32-750ML
WINE-RED-P66-750ML	RTD-CIDER-P14-440ML
RTD-BEER-P58-750ML	WINE-BIB-P66-5LTR
RTD-BEER-P24-340ML	WINE-BIB-P52-3LTR
RTD-BEER-P45-500ML	WINE-RED-P33-750ML
WINE-FORTIFIED-P10-750ML	WINE-RED-P61-750ML
WINE-RED-P87-750ML	WINE-RED-P7-750ML
RTD-FABS-P36-440ML	
RTD-FABS-P30-330ML	
WINE-SPARKLING-P1-750ML	
WINE-RED-P54-750ML	
WINE-FORTIFIED-P11-750ML	
WINE-BIB-P28-3LTR	
WINE-BIB-P54-5LTR	
WINE-SPARKLING-P2-750ML	
SPIRITS-BRANDY-P23-750ML	
WINE-BIB-P44-3LTR	

TABLE A.68: *Part 2: WINE-RED-P102-750ML significant independent variables results as each batch of independent variables are added*

A.5.2   Variable refinement using correlation analysis

Product	Pearson Correlation Coefficients	p-value	Correlation Coefficients Ranking
RTD-CIDER-P17-660ML	0.9508	0.0001	1
RTD-BEER-P55-750ML	0.8693	0.0001	2
RTD-BEER-P50-660ML	0.7985	0.0001	4
RTD-BEER-P58-750ML	0.7299	0.0001	10
WINE-WHITE-P1-5LTR	0.7186	0.0001	12
WINE-BIB-P69-5LTR	0.6808	0.0001	18
RTD-BEER-P39-440ML	0.6787	0.0001	22
RTD-BEER-P41-440ML	0.677	0.0001	24
RTD-FABS-P37-660ML	0.6615	0.0001	30
RTD-BEER-P28-440ML	0.6505	0.0001	36
SPIRITS-VODKA-P1-1LTR	0.6499	0.0001	37
RTD-CIDER-P16-500ML	0.6497	0.0001	38
RTD-CIDER-P15-440ML	0.6409	0.0001	44
RTD-CIDER-P14-440ML	0.6384	0.0001	46
RTD-FABS-P9-275ML	0.6317	0.0001	48
WINE-PERLE-P4-2LTR	0.6284	0.0001	51
RTD-CIDER-P3-330ML	0.6282	0.0001	52
RTD-BEER-P57-750ML	0.6272	0.0001	55
RTD-FABS-P3-250ML	0.6055	0.0001	62
SPIRITS-WHISKY-P11-750ML	0.6052	0.0001	64
WINE-BIB-P74-5LTR	0.6026	0.0001	66
WINE-ROSE-P2-1.5LTR	0.5806	0.0001	84
RTD-BEER-P56-750ML	0.572	0.0001	91
RTD-BEER-P3-330ML	0.5697	0.0001	92
RTD-BEER-P12-330ML	0.5519	0.0001	102
RTD-BEER-P43-500ML	0.5506	0.0001	104
RTD-BEER-P29-440ML	0.5331	0.0001	120
WINE-FORTIFIED-P1-1LTR	0.5043	0.0001	134
SPIRITS-LIQUEURS-P16-750ML	0.5002	0.0001	137
WINE-RED-P79-1.5LTR	0.4705	0.0001	160
RTD-BEER-P19-330ML	0.4671	0.0001	165
RTD-BEER-P44-500ML	0.4566	0.0001	169
RTD-BEER-P47-500ML	0.4344	0.0001	188
WINE-BIB-P55-5LTR	0.4286	0.0001	196
SPIRITS-WHISKY-P53-750ML	0.4172	0.0001	207
RTD-BEER-P48-5LTR	0.4085	0.0001	221
SPIRITS-WHISKY-P8-1LTR	0.4004	0.0001	226
RTD-FABS-P11-275ML	0.3994	0.0001	227
RTD-FABS-P38-660ML	0.3777	0.0001	241
WINE-BIB-P79-5LTR	0.3382	0.0001	294
RTD-FABS-P4-250ML	0.3321	0.0001	298
RTD-CIDER-P9-330ML	0.316	0.0001	311
WINE-BIB-P56-5LTR	0.2892	0.0002	344

TABLE A.69: *Part 1: RTD-CIDER-P18-660ML significant independent variables results from the stepwise method ranked by correlation coefficient strength*

Product	Pearson Correlation Coefficients	p-value	Correlation Coefficients Ranking
RTD-BEER-P14-330ML	0.2828	0.0003	356
RTD-BEER-P2-330ML	0.2482	0.0017	392
WINE-BIB-P19-2LTR	0.2073	0.0092	425
RTD-BEER-P1-330ML	0.1915	0.0163	436
WINE-BIB-P21-2LTR	-0.1875	0 0.018	439
RTD-BEER-P18-330ML	0.1718	0.0314	454
SPIRITS-WHISKY-P9-1LTR	0.1479	0.0646	468
RTD-BEER-P16-330ML	0.1404	0.0794	474
RTD-BEER-P22-340ML	-0.0936	0.2430	514
RTD-BEER-P34-440ML	0.0931	0.2461	515
WINE-BIB-P66-5LTR	0.0903	0.2609	517
WINE-BIB-P35-3LTR	0.0724	0.3675	529
RTD-BEER-P31-440ML	0.0695	0.3872	532
WINE-BIB-P30-3LTR	0.0549	0.4946	543
RTD-FABS-P27-330ML	0.0406	0.6135	556
WINE-BIB-P18-2LTR	0.0325	0.6866	564
SPIRITS-BRANDY-P1-1LTR	0.021	0.7943	578
RTD-BEER-P8-330ML	0.0103	0.8982	587

TABLE A.70: *Part 2: RTD-CIDER-P18-660ML significant independent variables results from the stepwise method ranked by correlation coefficient strength*

Product	Pearson Correlation Coefficients	p-value	Correlation Coefficients Ranking
WINE-SPARKLING-P7-750ML	0.9072	0.0001	3
RTD-CIDER-P2-330ML	0.90392	0.0001	5
SPIRITS-LIQUEURS-P11-750ML	0.90113	0.0001	6
SPIRITS-LIQUEURS-P5-750ML	0.88092	0.0001	16
RTD-FABS-P13-275ML	0.86605	0.0001	19
RTD-FABS-P7-275ML	0.8614	0.0001	22
RTD-FABS-P11-275ML	0.84348	0.0001	29
RTD-BEER-P35-440ML	0.83472	0.0001	33
SPIRITS-LIQUEURS-P9-750ML	0.82456	0.0001	39
WINE-SPARKLING-P31-750ML	0.80876	0.0001	53
RTD-BEER-P24-340ML	0.80303	0.0001	60
RTD-FABS-P24-300ML	0.80043	0.0001	61
SPIRITS-LIQUEURS-P12-750ML	0.78281	0.0001	69
RTD-FABS-P20-275ML	0.78002	0.0001	70
RTD-BEER-P20-340ML	0.77676	0.0001	73
RTD-BEER-P39-440ML	0.7722	0.0001	77
RTD-BEER-P12-330ML	0.76062	0.0001	83
RTD-BEER-P3-330ML	0.74408	0.0001	95
RTD-BEER-P34-440ML	0.71893	0.0001	107
RTD-FABS-P6-250ML	0.71312	0.0001	110
RTD-BEER-P1-330ML	0.71056	0.0001	111
RTD-BEER-P48-5LTR	0.70032	0.0001	117
WINE-SPARKLING-P10-750ML	0.68467	0.0001	131
WINE-SPARKLING-P4-750ML	0.68181	0.0001	134
SPIRITS-LIQUEURS-P7-750ML	0.66557	0.0001	146
RTD-FABS-P33-440ML	0.63473	0.0001	164
SPIRITS-WHISKY-P42-750ML	0.568	0.0001	200
RTD-FABS-P23-300ML	0.53131	0.0001	226
SPIRITS-WHISKY-P32-750ML	0.52011	0.0001	233
SPIRITS-WHISKY-P34-750ML	0.50975	0.0001	241
RTD-BEER-P21-340ML	0.49341	0.0001	253
RTD-BEER-P58-750ML	0.42761	0.0001	297
RTD-FABS-P1-250ML	0.41414	0.0001	308
RTD-BEER-P13-330ML	0.41381	0.0001	309
RTD-BEER-P9-330ML	0.38603	0.0001	326
SPIRITS-LIQUEURS-P1-1LTR	0.32997	0.0001	360
SPIRITS-WHISKY-P45-750ML	0.32809	0.0001	362
SPIRITS-WHISKY-P49-750ML	0.3233	0.0001	366
RTD-FABS-P30-330ML	0.3224	0.0001	367
RTD-BEER-P47-500ML	0.31206	0.0001	378
SPIRITS-WHISKY-P24-750ML	0.28517	0.0003	400
RTD-BEER-P27-440ML	0.28185	0.0003	403
RTD-FABS-P27-330ML	0.24702	0.0018	432
RTD-FABS-P38-660ML	0.21905	0.0058	453

TABLE A.71: *RTD-CIDER-P6-330ML significant independent variables results from the stepwise method ranked by correlation coefficient strength*

Product	Pearson Correlation Coefficients	p-value	Correlation Coefficients Ranking
WINE-FORTIFIED-P11-750ML	0.63488	0.0001	7
WINE-FORTIFIED-P10-750ML	0.60504	0.0001	14
WINE-PERLE-P8-1.5LTR	0.59198	0.0001	21
WINE-SPARKLING-P25-750ML	0.59181	0.0001	22
WINE-BIB-P43-3LTR	0.56482	0.0001	36
WINE-SPARKLING-P1-750ML	0.55359	0.0001	43
SPIRITS-WHISKY-P21-750ML	0.53175	0.0001	63
SPIRITS-BRANDY-P23-750ML	0.53056	0.0001	64
SPIRITS-BRANDY-P12-750ML	0.52022	0.0001	81
WINE-BIB-P41-3LTR	0.51363	0.0001	86
WINE-RED-P33-750ML	0.49451	0.0001	107
WINE-PERLE-P7-1.5LTR	0.4787	0.0001	122
WINE-RED-P39-750ML	0.47089	0.0001	133
WINE-RED-P72-750ML	0.45888	0.0001	147
RTD-BEER-P24-340ML	0.45817	0.0001	148
SPIRITS-VODKA-P1-1LTR	0.45727	0.0001	149
WINE-BIB-P66-5LTR	0.45192	0.0001	157
WINE-RED-P50-750ML	0.43364	0.0001	177
WINE-SPARKLING-P2-750ML	0.42685	0.0001	181
RTD-FABS-P1-250ML	0.42679	0.0001	182
WINE-BIB-P55-5LTR	0.41264	0.0001	206
WINE-SPARKLING-P22-750ML	0.40244	0.0001	213
WINE-RED-P22-750ML	0.38942	0.0001	227
WINE-BIB-P78-5LTR	0.36537	0.0001	256
RTD-FABS-P36-440ML	0.36484	0.0001	257
RTD-FABS-P30-330ML	0.363	0.0001	259
SPIRITS-WHISKY-P5-1LTR	0.36096	0.0001	262
WINE-BIB-P44-3LTR	0.34642	0.0001	285
SPIRITS-WHISKY-P11-750ML	0.32794	0.0001	306
WINE-RED-P19-750ML	0.3261	0.0001	308
RTD-BEER-P58-750ML	0.31503	0.0001	322
SPIRITS-GIN-P4-750ML	0.31116	0.0001	326
WINE-BIB-P84-5LTR	0.30108	0.0001	335
WINE-RED-P54-750ML	0.28254	0.0003	352
WINE-RED-P87-750ML	0.28202	0.0003	354
WINE-RED-P37-750ML	0.26505	0.0008	365
WINE-RED-P44-750ML	0.26163	0.0009	372
WINE-RED-P31-750ML	0.25771	0.0011	378
WINE-SPARKLING-P19-750ML	0.24738	0.0018	389
WINE-FORTIFIED-P6-375ML	0.2463	0.0019	391
WINE-RED-P9-750ML	0.23151	0.0035	409
WINE-FORTIFIED-P16-750ML	0.22933	0.0039	412
RTD-CIDER-P14-440ML	0.22844	0.0040	414
WINE-BIB-P67-5LTR	0.22695	0.0043	418
WINE-RED-P23-750ML	0.21556	0.0067	431
WINE-RED-P66-750ML	0.20348	0.0106	443
RTD-FABS-P27-330ML	0.19094	0.0166	457
WINE-RED-P11-750ML	0.18593	0.0197	461

TABLE A.72: *Part 1: WINE-RED-P102-750ML significant independent variables results from the stepwise method ranked by correlation coefficient strength*

Product	Pearson Correlation Coefficients	p-value	Correlation Coefficients Ranking
WINE-RED-P49-750ML	0.18171	0.0228	465
WINE-BIB-P54-5LTR	0.16461	0.0394	481
WINE-RED-P7-750ML	0.15771	0.0485	489
WINE-WHITE-P58-750ML	0.15617	0.0508	493
WINE-BIB-P28-3LTR	0.1548	0.0529	495
RTD-FABS-P23-300ML	0.14987	0.0610	499
WINE-RED-P59-750ML	-0.1485	5 0.063	500
WINE-RED-P61-750ML	0.14167	0.0768	508
WINE-RED-P32-750ML	0.13673	0.0877	512
WINE-RED-P68-750ML	0.1257	0.1167	520
WINE-BIB-P45-3LTR	0.12508	0.1186	522
WINE-RED-P79-750ML	-0.0815	9 0.309	547
WINE-BIB-P52-3LTR	-0.0771	9 0.336	550
RTD-BEER-P45-500ML	0.00667	0.9339	589

TABLE A.73: *Part 2: WINE-RED-P102-750ML significant independent variables results from the stepwise method ranked by correlation coefficient strength*

A.6   Results for the price elasticity model to determine a competitor set

A.6.1   Output statistics from regression analysis
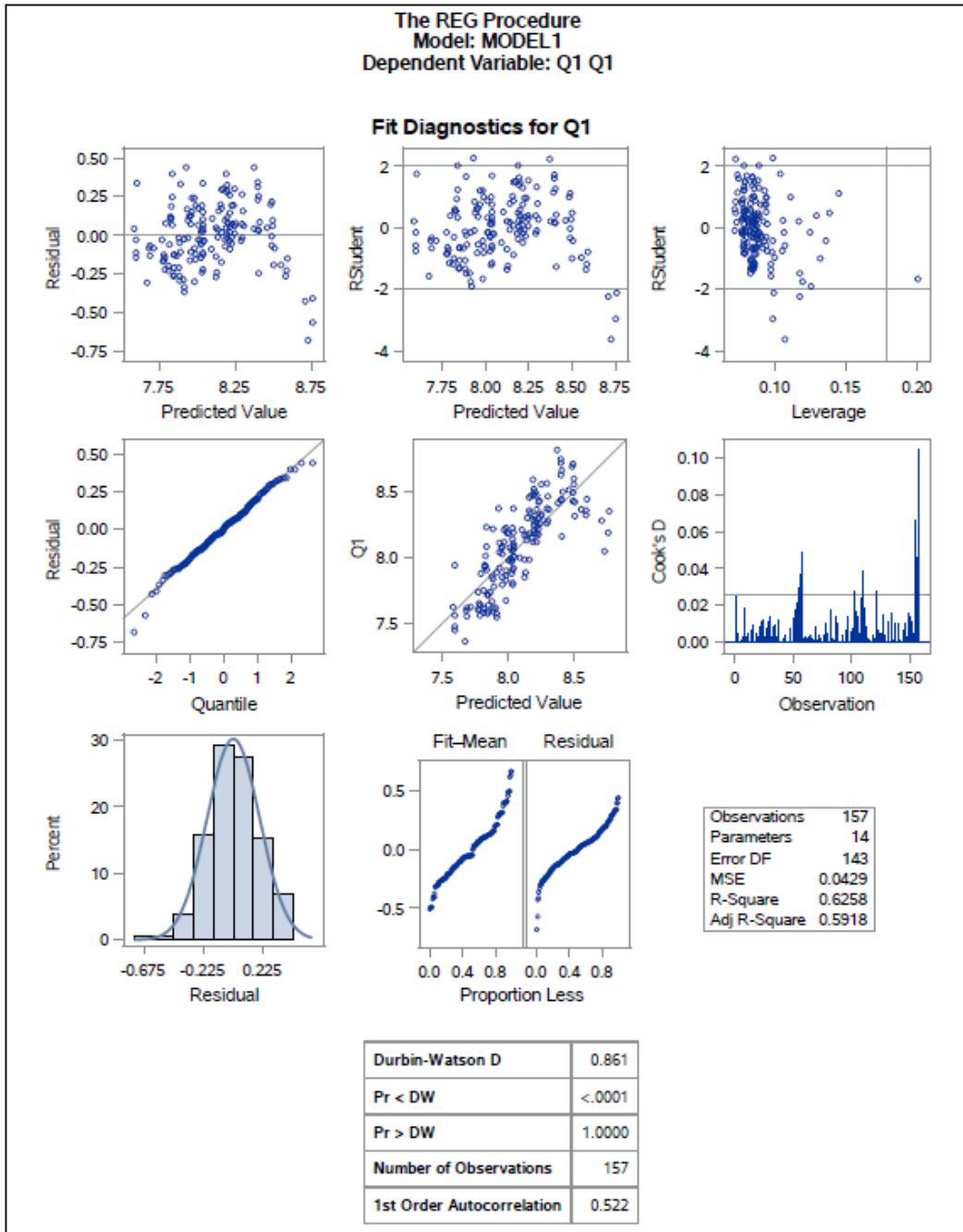
FIGURE A.25: *RTD-CIDER-P18-660ML output statistics for price elasticity regression model with RTD-BEER-P3-330ML*
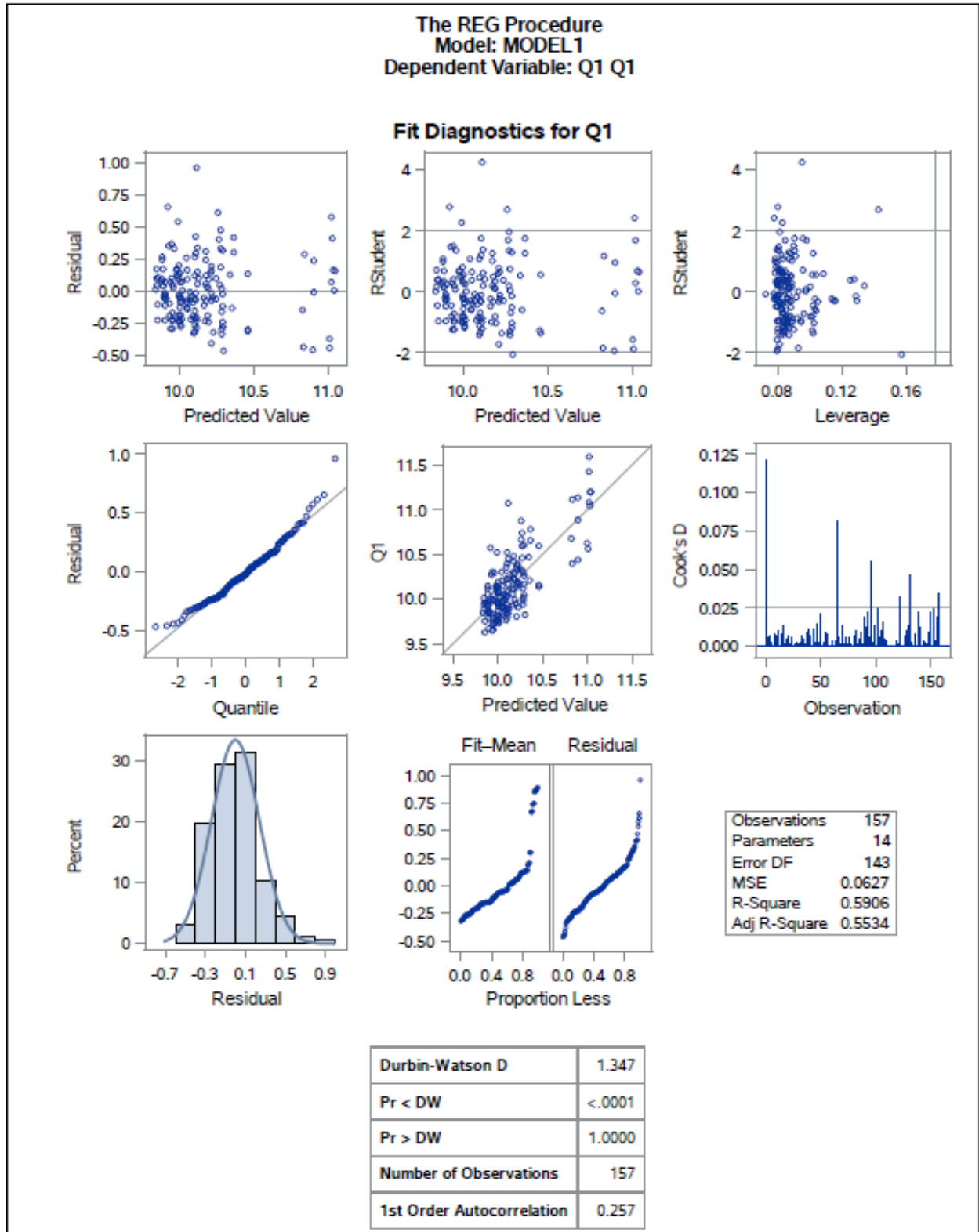
FIGURE A.26: *RTD-CIDER-P6-330ML output statistics for price elasticity regression model with SPIRITS-LIQUEURS-P11-750ML*

Bibliography

[1] ALI M, 2016, *Evaluating advertising effectiveness of creative television advertisements for high involvement products*, IDEAS Working Paper Series from RePEc, Available from `http://search.proquest.com/docview/1809236083/`.

[2] BA C, 1977, *A study of a measure of sampling adequacy for factor-analytic correlation matrices*, Multivariate Behavioral Research, **12(1)**.

[3] BAND-JONKER N, 2018, *Distell laat Mainstay cane se bootjie wegseil . . . vir altyd*, Die Burger Newspaper, **21 December**.

[4] BASS FM, 1969, *A new product growth for model consumer durables*, Management Science, **15(5)**, pp. 215–228.

[5] BEYNON M and DAUNT K, 2012, *Cluster analysis: An example analysis on personality and dysfunctional customer behaviour*, pp. 359–378 in *Quantitative Modelling in Marketing and Management*, pp. 359–378. World Scientific Publishing Co.

[6] BOULDING W, 1994, *Mastering the mix: Do advertising, promotion, and sales force activities lead to differentiation?*, Journal of Marketing Research (JMR), **31(2)**, pp. 159–173.

[7] BRATCHELL N, 1989, *Cluster analysis*, Chemometrics and Intelligent Laboratory Systems, **6**, pp. 105–125.

[8] BROCK G, PIHUR V, DATTA S and DATTA S, 2008, *An R package for cluster validation*, Journal of Statistical Software, Articles, **25(4)**, pp. 1–22, Available from `https://www.jstatsoft.org/v025/i04`.

[9] BROWN KM, *Company website*, Available from `http://www.millwardbrown.com/mb-global/what-we-do`.

[10] BUSH RJ and SINCLAIR SA, 1991, *A multivariate model and analysis of competitive strategy in the U.S. hardwood lumber industry*, Forest Science, **37(2)**, pp. 481–499.

[11] CHANDRASHEKAR G, 2014, *A survey on feature selection methods*, Computers and Electrical Engineer, **40(1)**, pp. 16–29.

[12] CHARNES A, 1978, *Measuring the efficiency of decision making units*, European Journal of Operational Research, **2(6)**, pp. 429–445.

[13] CHARRAD M, GHAZZALI N, BOITEAU V and NIKNAFS A, 2014, *An R package for determining the relevant number of clusters in a data set*, Journal of Statistical Software, Articles, **61(6)**, pp. 1–36, Available from `https://www.jstatsoft.org/v061/i06`.

[14] CHEONG Y, DE GREGORIO F and KIM K, 2014, *Advertising spending efficiency among top U.S. advertisers from 1985 to 2012: Overspending or smart managing?*, Journal of Advertising, **43(4)**, pp. 344–358.

[15] CHUMNEY F, *Exploratory factor analysis: A brief introduction*, Available from `https://www.youtube.com/watch?v=rQZiy7GE-LM`.

[16] COMPANY BC, *Company website*, Available from `http://www.bcg.com`.

[17] COMPANY RGB, *Company website*, Available from `www.rgbc.co.za/`.

[18] DAY GS, 1981, *The product life cycle: Analysis and applications issues*, Journal of Marketing, **45(4)**, pp. 60 – 67, Available from `http://search.ebscohost.com.ez.sun.ac.za/login.aspx?direct=true&db=buh&AN=4998159&site=ehost-live&scope=site`.

[19] DELIVERED IG, *Company website*, Available from `https://www.iriworldwide.com/en-SA/solutions/analytics-and-in-market-execution/Marketing-mix`.

[20] DIAGEO, *Company website*, Available from `http://www.diageo.com`.

[21] DISTELL, *Company website*, Available from `https://www.distell.co.za/`.

[22] ESNELL, *Company website*, Available from `http://www.esnell.co.za/`.

[23] EXPERTS SDB, *Company website*, Available from `https://skimgroup.com/`.

[24] FABRIGAR LR, WEGENER DT, MACCALLUM RC and STRAHAN EJ, 1999, *Evaluating the use of exploratory factor analysis in psychological research*, Psychological Methods, **4**, pp. 272–299.

[25] FARRELL MJ, 1957, *The measurement of productive efficiency [with discussion]*, Journal of the Royal Statistical Society, Series A, **120**, pp. 253–291.

[26] FNB BER, *FNB BER consumer confidence index 2015 q3*, Available from `https://www.ber.ac.za/BER%20Documents/BER-press-release/?doctypeid=1068&year=2015/`.

[27] GREEN D and CO, *Company website*, Available from `http://www.dgb.co.za/`.

[28] GUJARATI DN, 2003, *Basic econometrics*, 4th Edition, McGraw Hill, Boston.

[29] HARRELL FE JR, 2006, *Regression modeling strategies*, Springer-Verlag, Berlin, Heidelberg.

[30] HAX AC and MAJLUF NS, 2003, *Strategic management an integrated perspective*, Prentice Hall.

[31] HEINEKEN, *Company website*, Available from `http://www.heineken.com/za/`.

[32] HEZEKIAH L, RAMAKHRISHNAN M and SHABAN M, 2016, *Advertising media efficiency of FMCG firms in india: An empirical investigation using data envelopment analysis*, IUP Journal of Management Research, **15(1)**, pp. 37–47, Available from `http://search.proquest.com/docview/1772424335/`.

[33] HOLTZKAMPF E, 2015, *Liquor consumption patterns in South Africa July 2015*, SAWIS South African Wine Industry Information and Systems Yearly Report, **Yearly Report**.

[34] HOUSE R, 2017, *Alcoholic beverage brand-health monitor monthly report July 2017*, Research House South Africa Monthly Report, **Monthly Report**.

[35] INVESTOPEDIA, *Brand equity*, Available from `http://www.investopedia.com/terms/b/brandequity.asp`.

[36] IPSOS, *Company website*, Available from `https://www.ipsos.com/en-us/knowledge/new-services/marketing-mix-model`.

[37] JEFFREY RL, 1975, *Development of a product life cycle portfolio concept for strategic planning purposes*, Thesis (M.B.L.)–University of South Africa.

[38] JENSEN W, 1982, *Grocery products exhibit different life cycle curve*, Marketing News: American Marketing Association, p. 6.

[39] KOTLER P, 2003, *Marketing management*, 11th Edition, Prentice Hall, Upper Saddle River, N.J.

[40] LIU F and TIAN G, 2011, *Is the demand for alcoholic beverages in developing countries sensitive to price? Evidence from China*, International Journal of Environmental Research and Public Health, **8(6)**, pp. 2124–2131, Available from `https://doaj.org/article/1b3a469ef41d4d13b1efa968d87fb72b`.

[41] LUO X and DONTHU N, 2001, *Benchmarking advertising efficiency*, Journal of Advertising Research, **41(6)**, pp. 7–18, Available from `http://search.proquest.com/docview/205093279/`.

[42] MCKINSEY, *Company website*, Available from `https://www.mckinsey.com/za/our-work`.

[43] METTYEAR D, 2015, *Domestic volume report 2015*, IWSR International Wines and Spirits Record, **Yearly Report**.

[44] MILLIGAN GW and COOPER MC, 1987, *Methodology review: Clustering methods*, Applied Psychological Measurement, **11(4)**, pp. 329–354, Available from `http://dx.doi.org/10.1177/014662168701100401`.

[45] MOHR P, 2008, *Economics for South African students*, 4th Edition, Van Schaik, Pretoria.

[46] MOOI E and SARSTEDT M, 2011, *A concise guide to market research: The process, data, and methods using IBM SPSS Statistics*, Available from `http://www.springer.com/gp/book/9783642539640`.

[47] MORGENSTERN M, LI Z and SARGENT JD, 2017, *The party effect: prediction of future alcohol use based on exposure to specific alcohol advertising content. (report)*, Addiction, **112(1)**.

[48] MORRISON A and WENSLEY R, 1991, *Boxing up or boxed in? a short history of the Boston consulting group share growth matrix*, Journal of Marketing Management, **7(2)**, pp. 105–129, Available from `http://dx.doi.org/10.1080/0267257X.1991.9964145`.

[49] MULUGETA D, GREENFIELD JM, BOLEN T and CONLEY LK, *Paper 425-2013 Price-and cross-price elasticity estimation using SAS*.

[50] NAUMANN E, 2015, *The South African liquor industry June 2015*, Consumer and Corporate Regulation Division CCRD DTI Yearly Report, **Yearly Report**.

[51] NIELSEN, 2015, *Off consumption alcoholic beverage monthly sales volume report June 2015*, Nielsen Monthly Report, **Monthly Report**.

[52] NIELSEN, *Company website*, Available from `http://www.nielsen.com/ssa/en/solutions/advertising-effectiveness.html`.

[53] O'ROURKE N and HATCHER L, 2013, *A step-by-step approach to using SAS system for factor analysis and structural equation modeling*, SAS Institute Inc.

[54] OSBORNE JW and BANJANOVIC ES, 2016, *Exploratory factor analysis with SAS ®*, SAS Institute Inc.

[55] PAGANO RR, 1998, *Understanding statistics in the behavioural sciences, 5th Edition*, Brooks/Cole Publishing Company.

[56] PINEDO M, *Department of information, operations and management sciences*, Available from `https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=2&cad=rja&uact=8&ved=2ahUKEwicuqWfnMbjAhXXQxUIHbVTCbgQFjABegQIABAC&url=http%3A%2F%2Fwww.stern.nyu.edu%2Fom%2Ffaculty%2Fpinedo%2Fofs%2Fdownload%2FDEA.ppt&usg=AOvVaw1Ii7M8uyLrVpabIdMy_MLB`.

[57] RAGSDALE CT, 2007, *Managerial decision modeling*, 6th Edition, South-Western, Mason, Ohio.

[58] REEVES M, MOOSE S and VENEMA T, *BCG classics revisited: The growth share matrix*, Available from `https://www.bcg.com/publications/2014/growth-share-matrix-bcg-classics-revisited.aspx`.

[59] ROBINSON H, WYSOCKA A and HAND C, 2007, *Internet advertising effectiveness: The effect of design on click-through rates for banner ads*, International Journal of Advertising, **26(4)**, pp. 527–541.

[60] SABMILLER, *Company website*, Available from `http://www.sab.co.za`.

[61] SAS, *Base SAS(R) 9.2 procedures guide: The REG procedure, second edition*, Available from `https://support.sas.com/documentation/cdl/en/statug/63033/HTML/default/viewer.htm#chap0_toc.htm`.

[62] SAS, *SAS Institute Inc SAS/STAT(R) 9.2 user's guide, second edition*, Available from `https://support.sas.com/documentation/cdl/en/statug/63033/PDF/default/statug.pdf`.

[63] SAS, *SAS ®Enterprise Miner™14.3: Cubic clustering criterion*, Available from `http://documentation.sas.com/?docsetId=emref&docsetTarget=n1dm4owbc3ka5jn11yjkod7ov1va.htm&docsetVersion=14.3&locale=en`.

[64] SAS, *Base SAS(R) 9.4 procedures guide: The CORR procedure, fourth edition*, Available from `http://support.sas.com/documentation/cdl/en/procstat/68142/HTML/default/viewer.htm#procstat_corr_overview.htm`.

[65] SCHEIBLER D and SCHNEIDER W, 1985, *Monte carlo tests of the accuracy of cluster analysis algorithms: A comparison of hierarchical and non hierarchical methods*, Multivariate Behavioral Research, **20(3)**, pp. 283–304, Available from `http://dx.doi.org/10.1207/s15327906mbr2003_4`, pMID: 26781967.

[66] SCHELLINCK DA, 1983, *Effect of time on a marketing strategy*, Industrial Marketing Management, **12(2)**, pp. 83–88, Available from `http://doi.org/10.1016/0019-8501(83)90016-0`.

[67] Schönrock-Adema J, Heijne-Penninga M, van Hell EA and Cohen-Schotanus J, 2009, *Necessary steps in factor analysis: Enhancing validation studies of educational instruments. the PHEEM applied to clerks as an example*, Medical Teacher, **31(6)**, pp. e226–e232, Available from `https://doi.org/10.1080/01421590802516756`, pMID: 19089728.

[68] Schultz DE, 2009, *PARETO PARED*, Marketing News, **43(18)**, p. 24, Available from `http://search.ebscohost.com.ez.sun.ac.za/login.aspx?direct=true&db=buh&AN=49107123&site=ehost-live&scope=site`.

[69] Sethuraman R, 1991, *An analysis of the tradeoff between advertising and price discounting*, Journal of Marketing Research (JMR), **28(2)**, pp. 160–175.

[70] Sharp B and Institute TEB, 2010, *How brands grow: what marketers don't know by Byron Sharp and the researchers of the Ehrenberg-Bass Institute*, Oxford University Press South Melbourne, Vic.

[71] South Africa WWO, *Industry website*, Available from `https://www.wosa.co.za/The-Industry/History/Three-Centuries-of-Cape-Wine/Cape-Wines-Before-the-20th-Century/Timeline/`.

[72] Suhr D (Ed), 2014, *Exploratory or confirmatory factor analysis.* SAS Institute Inc, Statistics and Data Analysis, Available from `https://support.sas.com/resources/papers/proceedings/proceedings/sugi31/200-31.pdf`.

[73] Tabachnick BG and Fidell LS, 2013, *Using multivariate statistics: Pearson New International Edition*, volume 15, Pearson Education UK, Harlow.

[74] Taherdoost H, Sahibuddin S and Jalalioon N (Eds), 2014, *Exploratory factor analysis; concepts and theory.* Universiti Teknologi Malaysia, Advances in Applied and Pure Mathematics, Available from `http://www.wseas.us/e-library/conferences/2014/Gdansk/MATH/MATH-49.pdf`.

[75] Terblanche-Smit M, Du Preez R and Van Der Spuy T, 2015, *Measuring the impact of branded alcohol advertising and price on brand versus segment consumption*, The Clute Institute.

[76] Theodoridis S and Koutroumbas K, 2008, *Pattern recognition*, 4th Edition, Academic Press, Inc., Orlando, FL, USA.

[77] Thompson B, 2004, *Exploratory and confirmatory factor analysis: Understanding concepts and applications*, 1st Edition, American Psychological Association, Washington, DC.

[78] Thore S, 1996, *DEA and the management of the product cycle: The U.S. computer industry*, Computers and Operations Research, **23(4)**, pp. 341–357.

[79] Trebuňa P and Halčinová J, 2013, *Mathematical tools of cluster analysis*, Applied Mathematics, **04(05)**, p. 814–816.

[80] Tucker LR and MacCallum RC, 1997, *Exploratory Factor Analysis*, Unpublished manuscript, Ohio State University, Columbus, Available from `http://www.ffzg.unizg.hr/psiho/phm/nastava/Book_Exploratory%20Factor%20Analysis.PDF`.

[81] Winston WL, 2004, *Operations research: Applications and algorithms*, 4th Edition, Thomson, Australia.

[82] Wyn Owen G, 2012, *Applying point elasticity of demand principles to optimal pricing in management accounting*, The International Journal of Applied Economics and Finance, **6**, pp. 89–99.

[83] YONG AG and PEARCE S, 2013, *A beginner's guide to factor analysis: Focusing on exploratory factor analysis*, Tutorials in Quantitative Methods for Psychology, **9(2)**, pp. 70–94.