

Developing a joint space perceptual map with apple positions and consumer preference clusters

Market segmentation methodology tested on and applied to consumer preference ratings for ten apple varieties and positioning of new varieties



R. M. Hofstee – 940224351040

MCB-80436 - Master Thesis Marketing and Consumer behaviour

Management, Economics and Consumer studies

Wageningen, 13th of March 2020

Supervisor: dr. ir. F.J.H.M. Verhees

2nd Supervisor: dr. I.A.C.M. van der Lans



Colofon

Hofstee, R. M. (2020). *Developing a joint space perceptual map with apple positions and consumer preference clusters: Market segmentation methodology tested on and applied to consumer preference ratings for ten apple varieties and positioning of new varieties*. Wageningen University.

Keywords: apples, market segmentation, positioning, multidimensional scaling, multidimensional unfolding, cluster analysis, sensory attributes, consumer preference

© 2020 Wageningen, Stichting Wageningen Research, Wageningen Plant Research, P.O. Box 16, 6700 AA Wageningen, The Netherlands; T +31 (0)317 48 07 00; www.wur.eu/plant-research

Chamber of Commerce no. 09098104 at Arnhem VAT NL no. 8065.11.618.B01

PPS Duurzame Fruitintroducties Beter Begrepen (Sustainable Fruit Introductions Better Understood)
TKI T&U (TU 18157)

Work package 1. Overall Structure of the research infrastructure

MSc Thesis, Wageningen University & Research

Author:	R.M. Hofstee
Registration number:	940224351040
Study program:	Management, Economics and Consumer Studies
Chair group:	Marketing and Consumer Behaviour
Course code:	MCB-80436
Supervisor:	Dr. ir. F.J.H.M. Verhees
2 nd Supervisor:	Dr. I.A.C.M. van der Lans
Wageningen Research Open Crops:	P. A. Balk BSc & Ing. A.W. van der Heiden
Date of submission:	13 th of March 2020

The project Duurzame Fruitintroducties Beter Begrepen (TU-18157) receives financial support from the Top sector Tuinbouw & Uitgangsmaterialen. Within the Top sector, businesses, knowledge institutions and the government work together on innovations in the field of sustainable production of healthy and safe food and the development of a healthy, green living environment.

Preface

“Yes, this is a relevant technique in marketing research, but it will be explained in more depth in the more advanced courses” – Teachers at WageningenUR

Turns out I had to wait until I started doing my master thesis to actually gain in depth insight in some of the most relevant techniques used in marketing research!

I have come a long way to arrive at this point in my academic career and I do not regret a single step of it. I started doing my bachelors in Food Technology and somewhere in my second year I realized that my interests and capacities lie more with the consumer side of the food production chain. However, when I start something I will finish it, because that is the spirit I was brought up with. I finished my bachelors and did a premaster to be able to do the master Management, Economics and Consumer studies. I finally felt in place, even though all my beta friends made fun of me because I was now ‘living the LeBo live’.

I love to be able to apply my analytical and technical skills to these real world consumer issues and this thesis really gave me the opportunity to dive into some of the more advanced methods applied in marketing research. Statistics and methodology have always been my field of interest: turning numbers into meaning and gaining insights based on which decisions can be made.

In this research on apples I was also able to apply my experience of working in Field marketing. For my data gathering I went through the entire process from location planning, to logistics, to actually being present on a supermarket floor and asking consumers to try apples and fill in the questionnaire.

I want to thank my supervisor Frans, for his ongoing enthusiasm about the project and his realistic outlook on matters, and for putting the process into perspective. I sometimes doubted whether I was doing something new, because my project had so much overlap with my predecessor. Also, I want to thank my supervisor Ivo, because of the time he spent helping me with some statistical issues, coming up with ideas and the nice anecdotes that gave a personal touch to the statistical methods, which can be abstract from time to time. No one at the Leeuwenborch building has ever tried so many versions of PREFSCAL, I think. I also want to thank the people from Randwijk, especially Peter and Alma, for providing feedback on the project and for providing the fresh apples every day during the data gathering. They inspired me to switch from eating bananas to apples. I want to thank Job and Maikel from Brandwise BV, for providing materials that gave the instore data gathering a very professional touch, which was part of the success. Last but not least, I want to thank my study pal Nick, because he cut the apples in such a proper way and made the data gathering real fun, despite waking up very early every morning.

In my future career I would like to continue developing my skills in gathering and analysing sensory data with regards to consumer preference for food products and be able to consult companies in their decision making process regarding the marketing efforts for (new) products. I believe that the experience I gained in this thesis is a proper starting point for a promising career and I am looking forward to what the future holds.

Abstract

The fresh fruits and vegetables sector is a competitive sector in which new product development is key to gain competitive advantage. The selection of new apple varieties takes approximately twenty years and it is important to involve consumer preferences in the selection of new varieties. Therefore it was investigated how Dutch consumers perceive current winter apple varieties based on sensory attributes, also compared to findings on summer apple varieties and whether preference clusters could be distinguished. It was found that the Granny Smith, Golden Delicious and Fuji did not change their position in the map. Granny Smith and Golden Delicious were perceived as very different from the remaining eight apples. Golden Delicious was different in the sense that it lacked preferred attributes and Granny Smith was green, sour and tougher than the other apples. Fuji was rated to be the sweetest apple and the Jazz apple on average was the most preferred. Four preference clusters were found in this research and the significant attributes that were found were peel colour, glossiness, toughness, firmness of the peel, sourness and sweetness. These findings are in accordance with previous findings, except for that in this research peel colour and glossiness were found to be significant attributes. Based on the joint space perceptual map, a new apple variety could be best positioned at the cluster centroid of preference clusters one and/or three. A new variety would be either a greener apple that is average on glossiness, moderately high on toughness/firmness and has a balanced acidity, though more sweet than sour. Or it would be an apple that is averagely green, moderately high on glossiness, quite tough/firm and with a balanced acidity, though more sour than sweet.

Samenvatting

De tuinbouwsector (groente & fruit) is een competitieve sector waarin de ontwikkeling van nieuwe producten van groot belang is om een concurrerende positie in te nemen ten opzichte van concurrent-collega's. Het ontwikkelen van een nieuw ras duurt ongeveer twintig jaar en daarom is het van belang in een vroeg stadium consumenten voorkeuren te betrekken bij het selectieproces. Daarom is het onderzocht hoe Nederlandse consumenten huidige winter appels ervaren op basis van sensorische attributen, ook vergeleken met eerdere bevindingen m.b.t. zomer appels, en of consumenten konden worden gegroepeerd op basis van hun voorkeuren. Onderzoek wees uit dat de Granny Smith, Golden Delicious en Fuji hun positie behielden in de map. Granny Smith en Golden Delicious werden als erg verschillend ervaren t.o.v. de andere acht appels. Golden Delicious ontbrak het aan attributen die bepalend zijn voor consumenten voorkeur en Granny Smith werd vooral beoordeeld als een groene, zure, stevige appel. Fuji werd beschouwd als de zoetste appel en de Jazz appel was de appel die gemiddeld als beste werd beoordeeld. Vier segmenten met consumenten met een vergelijkbare voorkeur zijn geïdentificeerd, deze kunnen beschreven worden op basis van de volgende attributen: schilkleur, glans, stevigheid, stevigheid van het vruchtvlees, zuurheid en zoetheid. Deze bevindingen komen overeen met eerder gevonden resultaten, naast dat in dit onderzoek ook aangetoond is dat kleur en glans significante attributen zijn. Gebaseerd op de ruimte in de map, kunnen nieuwe rassen het beste gepositioneerd worden op de plek van de kern van de segmenten een en drie. Deze nieuwe appel zou dan een groenere appel zijn, die gemiddeld is qua glans, vrij stevig met een goede balans tussen zoet en zuur, doch liefst wat zoeter. Of het zou een groenere appel zijn, vrij glanzend en stevig met een goede balans tussen zoet en zuur, doch liefst wat zuurder.

Table of contents

Chapter 1: Introduction.....	1
Chapter 2: Pre-testing of segmentation methods.....	5
2.1. Perceptual mapping and multidimensional scaling (MDS)	5
2.2. Attribute vector fitting	5
2.3. Market segmentation.....	6
2.4. Segmentation method 1: two-stage approach using multiple linear regression and cluster analysis	8
2.4.1. Generating ideal points (Iacobucci, 2013).....	8
2.4.2. Cluster analysis.....	9
2.5. Segmentation method 2: two-stage approach using MDS PREFSCAL and cluster analysis	11
2.6. Comparing method 1 and 2.....	12
3.1. Study design	15
3.1.1. Sampling method	15
3.1.2. Stimuli.....	17
3.2. Data collection method / procedure.....	19
Chapter 4: Data analysis & results.....	22
4.1. Sample characteristics.....	22
4.2. Decompositional perceptual mapping	22
4.3. Attribute vector fitting	24
4.4. MDS PREFSCAL	27
4.5. Two-stage clustering procedure.....	28
4.6. Preference means and standard deviations.....	30
4.7. Cluster member characteristics.....	31
4.8. Ideal apples descriptions according to preference clusters.....	32
Chapter 5: Conclusion	36
Chapter 6: Discussion	38
6.1. Discussion of the results.....	38
6.1.1. Current findings and comparison with findings from Van den Broek (2019)	38
6.1.2. Comparison existing literature on consumer preference segments with segments found	38
6.2. Discussion on the data gathering process.....	38
6.2.1. Materials used	38
6.2.2. The questionnaire.....	39
6.2.3. Locations and external validity.....	39
6.2.4. Participant recruitment	39
6.2.5. Filling in the questionnaire	40

6.3. Discussion on the methodology, design and data analyses.....	41
Chapter 7: Recommendations.....	42
References.....	43
Appendices	46
Appendix 1 – Questionnaire.....	46
Appendix 2 – Mean dissimilarity ratings winter apples with standard deviation.....	52
Appendix 3 – Ideal point coordinates per respondent	53
Appendix 4 – Cluster membership per respondent.....	55
Appendix 5 - PROXSCAL – alternative data input.....	57

List of tables and figures

Table 1: classification segmentation bases, adapted from Wedel and Kamakura (2000)	6
Table 2: segmentation criteria, adapted from Kotler and Keller (2012)	7
Table 3: visiting appointments supermarkets	16
Table 4: apple varieties included in the research.....	17
Table 5: attributes that the apples were evaluated on.....	17
Table 6: incomplete design generated with crossdes	18
Table 7: apple pairs as rated by respondents per block.....	21
Table 8: sample characteristics, demographics and lifestyle	22
Table 9: lower triangle dissimilarity matrix.....	23
Table 10: : dimension coordinates apple varieties.....	24
Table 11: attribute means and standard deviation of the mean per apple variety	25
Table 12: outcome regression analyses attributes	26
Table 13: cluster centroids	28
Table 14: amount of respondents assigned to cluster	28
Table 15: predicted values attribute ratings ideal apples for the four segments.....	30
Table 16: mean preference scores per apple variety, sorted on preference.....	30
Table 17: mean preference ratings for each cluster per apple variety, sorted from most preferred to least preferred.....	31
Table 18: cluster 1 - Comparison mean attribute ratings and predicted values.....	32
Table 19: cluster 2 - Comparison mean attribute ratings and predicted values.....	33
Table 20: cluster 3 - Comparison mean attribute ratings and predicted values.....	33
Table 21: cluster 4 - Comparison mean attribute ratings and predicted values.....	34
Figure 1: summer apple varieties and ideal points - method 1 - three clusters	10
Figure 2: summer apple varieties and ideal points - method 1 - three clusters shown as centroids ...	10
Figure 3: summer apple varieties and ideal points - method 2 - four clusters	12
Figure 4: summer apple varieties and ideal points - method 2 - four clusters shown as centroids	13
Figure 5: spread of the selected supermarkets in the region	16
Figure 6: impression instore data gathering	19
Figure 7: scree plot normalized raw stress.....	23
Figure 8: decompositional perceptual map winter apple varieties	24
Figure 9: perceptual map including attribute vectors.....	26
Figure 10: joint space perceptual map including ideal points per respondent.....	27
Figure 11: joint space perceptual map winter apple varieties.....	28
Figure 12: joint space perceptual map winter apple varieties with ideal points per cluster.....	29
Figure 13: joint space perceptual map with four ideal apples per cluster and mirror images of vectors included	35

Chapter 1: Introduction

The fresh fruits and vegetables sector is a dynamic sector in which import and export of produce play a large role. In 2018 the values of export (including re-export) and the import of apples in the Netherlands was approximately in balance, respectively 261 and 264 million euros. Of the export value around one-third was produced in the Netherlands and two-third was imported and later exported (Dolman, Jukema & Ramaekers, 2019). This can be explained, because the Netherlands is an important and interesting trade hub due to its infrastructure, like the port in Rotterdam, which gives the Netherlands a logistical advantage (CBI, 2019). In 2019, the total apple production in the Netherlands was estimated to be 273,000 tons, with Elstar being the most widely produced apple (109 tons), followed by Jonagold (including Jonagored) with 67 tons. Also Kanzi, Rode Boskoop (Goudreinette), Junami and Golden Delicious are produced in the Netherlands (Statista, 2019). The availability of apples varies throughout the seasons and to serve Dutch consumers throughout the year, both imported apples and apples with a Dutch origin are offered on the market. This demand for constant supply of large quantities of fresh fruits and vegetables forces suppliers to collaborate in order to meet the requirements of the market (Kalaitzis & Van Dijk, 2005). Due to more liberalised international trade the competitive pressure has risen amongst firms operating in the fruits and vegetables sector and the *competitive rivalry* is high.

This can be illustrated by applying Porter's five forces framework (Porter, 2008) to the fruits and vegetables sector. Liberalised international trade is related to globalization, which increases the *threat of new entrants*, because entering the market has become easier due to a change in trade policies (Wrigley & Lowe, 2010). Low barriers of entry lead to a higher threat of new entrants (Porter, 2008). *Threat of substitute* products or services is high, because there are also other fruits that people consume in the same way as apples, for example citrus fruits and banana's; Harker, Gunson and Jaeger (2003) state that when the price of apples increases, consumers tend to shift to citrus fruits, for example. *Bargaining power of buyers* (e.g. retailers) is high, because consumers rely on supermarkets for their groceries, and therefore supermarkets demand only the best quality, higher quantities and lower prices to meet consumer requirements. To guarantee constant supply of good quality fresh produce, supermarkets invest in long term relationships with suppliers. In this power shift the importance of the wholesale sector has declined (Kalaitzis & Van Dijk, 2005). *Bargaining power of suppliers* can be high, depending on the type of supply. In the case of workforce as a supply to the sector, the bargaining power of suppliers is high. For example, the amount of migrant workers in the fresh fruits and vegetables sector has decreased while the sector is dependent on this type of workforce (Nieuwe Oogst, 2018). One of the causes of this, is the improved economic situation in the homeland of these workers (Volkskrant, 2019). Overall it can be concluded that the competitive rivalry of the fresh fruits and vegetables sector can be considered high (Porter, 2008).

To deal with this competitive rivalry, fresh fruits and vegetable producers can engage in new product development (NPD). As described by Van den Broek (2019) NPD is a key activity for businesses to deal with this competitive rivalry, because NPD makes it possible to reach competitive advantage and realize business growth. Dijksterhuis (2016) states that 50 to 75 percent of new products that are brought onto the market are removed before being profitable. Van Kleef, Van Trijp and Luning (2005) emphasize that it is important to involve consumer preference in the NPD process as early as possible, because it is a critical success factor. However, in the process of NPD 'listening to the voice of the customer' is often poorly executed. Nowadays, the fresh fruits and vegetables sector has acknowledged the need for involving consumer preference in the selection and positioning of new apple varieties (Onderzoeksvoorstel Topsector Tuinbouw & Uitgangsmaterialen (T&U), 2018).

The sector has found itself being challenged by the selection of apple varieties with future market potential due to the time it takes between the initial cross and the new variety reaching the market, which can be 15 to more than 20 years (Bowen, Blake, Tureček & Amyotte, 2019). In the case of products for which the time between the development and market introduction is long, like for apples, it is important to choose a proactive strategy. This strategy is focussed on uncovering future consumer wants, in contrast to a reactive strategy that is focussed on responding to dynamic market conditions and changing consumer preferences (Van Kleef, 2006). Uncovering future consumer wants is important, because once the product is introduced after twenty years, the market might have already changed by then. The development of decision making tooling for the selection and positioning of new apple varieties involving consumer preference is the first phase of a four-year research project called ‘Sustainable fruit introductions better understood’ which is commissioned and financed by the top sector T&U.

At the start of the research project, Van den Broek (2019) did research on the position of new apple varieties in relation to consumer preference and compared to current apple varieties. Van den Broek (2019) has shown by means of a literature study and selection criteria that certain sensory attributes were drivers of consumer preference for apples. It was found that within the appearance category, peel colour, size and glossiness were important attributes. Within the texture category, skin toughness, crispiness, firmness and juiciness were found to be relevant and in the flavour category, aroma, sourness and sweetness were selected as important drivers for consumer preference. Other attributes, like price, were left out of consideration, because the research focussed on the selection of new apple varieties to meet consumer wants with regards to preference for the apple based on the intrinsic sensory aspects of the apple. Respondents were asked to do sensory evaluations on five out of ten apple varieties. Proximity data, attribute ratings, preference data and demographic/lifestyle data were gathered. Based on this, multidimensional scaling and multiple linear regressions were applied to obtain a joint space perceptual map showing the positions of ten apple varieties harvested in the summer. One ideal vector was generated and it was found that the ideal apple for the set of respondents would be a smaller, glossy and juicy apple.

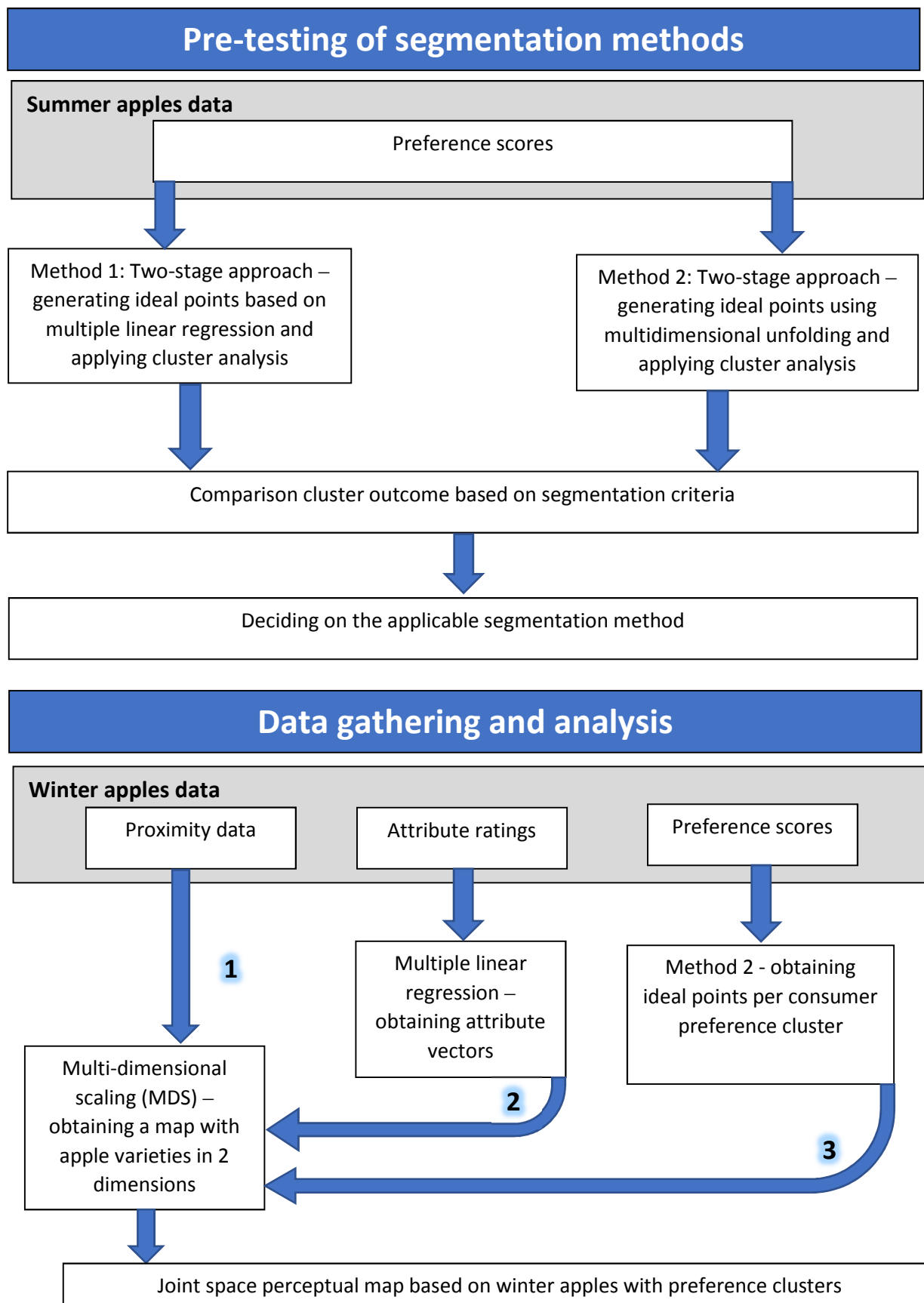
In this thesis the research by Van den Broek (2019) was built upon within the research project ‘Sustainable fruit introductions better understood’. Van den Broek (2019) included apple varieties available in Dutch supermarkets during the summer in the research, but in this research apples that are available in the winter were used, because the data gathering took place in January. The apples and related data will be referred to as summer apples (data) and winter apples (data). In the research by Van den Broek (2019) only one ideal vector for the entire set of respondents was generated, while it was also shown by means of a literature overview that consumer market segments with regards to apple preference exist (Dailliant-Spinnler, MacFie, Beyts & Hedderley, 1996 ; Bonany et al., 2014). As stated before, involving consumer preference is a key activity in NPD and one apple variety cannot serve the needs and wants of each customer. Therefore it is necessary to gain more insight in the market segments so that apples can be developed that meet the expectations of these consumer segments.

The overall objective of this research is to add to the methodology that is applied, by testing segmentation methodology to investigate how consumer preference clusters can be obtained that meet the segmentation criteria. The aim is to gain more insight into these market segments with regards to consumer preference for existing apple varieties. Therefore the main research question this papers aims to answer is *‘What is the position of winter apples based on consumer perceptions of these apples and what consumer preference clusters can be distinguished?’*. Three sub questions are posed to answer the main research question. First, what segmentation methods lead to preference

clusters of the summer apples data that meet the criteria for segmentation? The summer apples data will be used in the methodology testing to determine the segmentation method that will be used for the winter apples data. Second, what is the position of existing winter apple varieties based on consumer perceptions of these apples? And third, what preference clusters can be distinguished based on the winter apples data?

It is expected that consumer preferences vary over time and the results that are found are dependent on the time of the year in which the research takes place. Besides, the focus in this research is on the Dutch apple consumers and apples that are available in Dutch supermarkets, but the methods and results that are found can also be used and applied outside this scope, e.g. on consumers from outside the Netherlands and other fruits and vegetables.

In Chapter 2 the pre-testing of two segmentation methods will be discussed and applied to the summer apples data, in Chapter 3 the methodology and procedure for the winter apples data gathering will be presented. Chapter 4 shows the data analysis of the winter apples data and the results obtained from this analysis and the joint space perceptual map that was found. In Chapter 5 the conclusion is presented and the answers to the research questions will be given. In Chapter 6 the discussion will be given. Finally, in Chapter 7 the recommendations for further research will be discussed.



Chapter 2: Pre-testing of segmentation methods

A methodological approach was followed and therefore in this chapter the theoretical background of the segmentation methods will be discussed. However, first perceptual mapping, multidimensional scaling (MDS) and attribute vector fitting will be shortly discussed, because these techniques have been applied in previous research (Van den Broek, 2019) and were applied in this research as well. Then, two segmentation methods will be applied on the summer apples data. Finally the two methods will be compared by using segmentation criteria. A reflection will be made which method is the most applicable to use on the winter apples data. In this way, an answer will be given to the first sub research question.

2.1. Perceptual mapping and multidimensional scaling (MDS)

Perceptual mapping is a technique that visually represents the position of products and/or brands in a dimensional space based on consumer perceptions of the products/brands (Iacobucci, 2013). The technique that is used to constitute a perceptual map is multidimensional scaling (MDS). MDS seeks to determine the position of products in a dimensional space based on (dis)similarity ratings that consumers give to pairs of products. The consumer is asked to rate pairs of products on their (dis)similarity, e.g. 'how different is apple 1 compared to apple 2?'. In practice the criteria based on which consumer rate the (dis)similarity of two products can differ across consumers (Jaeger, Wakeling & MacFie, 2000). Some consumers might rate the (dis)similarity based on the colour of apples and some might give a rating based on the size, sourness, or any other sensory attribute that an apple possesses. However, in MDS consumer perceptions are considered to be homogeneous and the criteria, e.g. sourness, colour, size based on which consumers rate the (dis)similarities are not asked and considered. The (dis)similarity ratings can also be called proximity data: proximity means closeness (Mair, Borg & Rusch, 2016) and the data represents how close or far away products are from one another in the mind of the consumer. In this research, a lower triangle dissimilarity matrix of mean respondent ratings of pairs of products is used as input for MDS.

In this research the program SPSS version 23 by IBM is used to run the MDS procedure. In SPSS several options for running MDS exist, namely PROXSCAL, PREFSCAL and ALSCAL. PROXSCAL offers the advantage of being able to minimize normalized raw stress, rather than strain. Stress and strain are both so-called loss functions (Buja, Swayne, Littman, Dean, Hofmann & Chen, 2008) and these loss functions in the process of running MDS are minimized. In ALSCAL, compared to PROXSCAL, the proximities are first transformed to squared distances and strain is then the loss function that is minimized (Busing, Commandeur, Heiser, Bandilla & Faulbaum, 1997). Normalized raw stress is generally preferred because it is a measure based on the distances, while strain is based on the squared distances (IBM, 2019). PREFSCAL will be discussed later in this research.

Running the PROXSCAL procedure will lead to dimension coordinates of apple varieties in a dimensional space showing the position of apple varieties based on consumer (dis)similarity ratings. Based on a plot of the normalized raw stress against the amount of dimensions it can be decided how many dimensions to choose for. A two-dimensional space is often preferred, because maps that consist of two dimensions are easier to interpret (Iacobucci, 2013). In the range from one to nine dimensions the goodness of fit increases / the badness of fit decreases between the model and the data, however also the risk of finding a degenerate solution increases (Busing, 2010). This will be elaborated on later.

2.2. Attribute vector fitting

Attribute vector fitting is applied additionally to be able to interpret the perceptual map and can be used to give meaning to the axes in the dimensional space. Mean attribute ratings, e.g. the mean

rating on sourness per apple variety, are used as dependent variables in a regression analysis. The dimensions obtained from the MDS are used as independent variables and running the regression will give estimates for the standardized regression weights, which are the coordinates of the attribute vector heads. In linear regression an estimate is made of an equation in the form of $Y = \beta_1 X_1 + \beta_2 X_2$, in which Y is the dependent variable and the X1 and X2 are the dimensions obtained from MDS. The standardized beta coefficients, β_1 and β_2 , can be plotted in the dimensional perceptual map that was obtained using MDS (Iacobucci, 2013) leading to attribute vectors that show the length and direction of the attributes vectors compared to the position of the apples. Compared to the unstandardized coefficients, the standardized coefficients are corrected for the variances of the attributes. Also so called normed beta weights can be used, the length of the vectors is then corrected so that the length of the attribute vectors for each vector is equal (Iacobucci, 2013).

2.3. Market segmentation

Perceptual mapping based on MDS and additional attribute vector fitting by using multiple linear regressions will lead to insights about the position of certain brands and products in the mind of the consumer, but this does not yet give any information about the brands and/or products that consumers prefer or about groups (segments) of consumers with the same preference. Consumer preferences are not the same as consumer perceptions. Consumer preferences are considered to be heterogeneous and they can be used as bases for market segmentation (Wedel & Kamakura, 2000). Preference formation can be viewed from the perspective of an information processing framework. When tasting a food sample, e.g. an apple, the sensory stimuli of the apple are detected and identified before being further processed. The interpretation of the identified sensory stimuli leads to a perceptual overview of product similarities and differences. Consumers compare this overview to a set of individual rules that govern their preferences and the affective evaluation of this comparison is then transformed into a preference score (Jaeger et al., 2000). This indicates that consumers need to be able to detect and identify certain stimuli before they can determine the similarities and differences and finally their preference for a product. In other words, consumer perception of stimuli is a process that is an antecedent to consumer preference formation.

Over time, several definitions of market segmentation have been proposed in literature. Kotler (1988) defined market segmentation as the set of activities undertaken to identify homogeneous consumer groups. It is relevant to identify homogeneous consumer groups, because one product cannot serve all customers (Wedel, 1990). The market can be segmented based on geographical location (e.g. a city, country), demographics (e.g. age, ethnicity), consumer behaviour (e.g. needs and benefits) and psychographics (e.g. values, attitudes and lifestyle) (Martin, 2011). Wedel and Kamakura (2000) proposed a classification of segmentation bases into general observable, general unobservable, product specific observable and product specific unobservable, see table 1. In this research the market was segmented based on consumer preferences for apple varieties, which can be considered a product specific unobservable segmentation base.

Table 1: classification segmentation bases, adapted from Wedel and Kamakura (2000)

	General	Product specific
Observable	Cultural, geographic and socio-economic variables	User status, usage frequency, store loyalty and patronage, situations
Unobservable	Psychographics, values, personality and lifestyles	Psychographic, perceptions of benefits and attributes, elasticities, preferences , intention

It was found that market segments for apples based on differences in preferences exist. Especially in the apple sector it is important to introduce a product that suits the needs of the target segment, because the market is hypercompetitive and constantly demanding new varieties of high quality (Bowen et al., 2019). Dailliant-Spinnler and colleagues (1996) found that consumer segmentation could be achieved based on whether a sweet, hard apple or a juicy, acidic apple was preferred. Bonany and colleagues (2014) have found by means of preference mapping that six clusters of consumers could be distinguished that were classified into two mega clusters: cluster A representing a group of consumers preferring sweet apples that were moderately juicy and crisp and cluster B representing a group preferring acidic, firm, juicy and crisp apples that were less sweet. In these researches preference mapping was used as a technique to obtain preference clusters, however no segmentation criteria are mentioned that assessed the effectiveness of the segmentation that was applied.

In literature several criteria are posed to test the effectiveness of the segmentation process. Dolnicar, Grün and Leisch (2018) provide an overview of publications in which several criteria for effective market segmentation are discussed. Amongst others, Kotler and Keller (2012) are mentioned. Kotler and Keller (2012) propose five segmentation criteria, namely whether the segments are measurable, substantial, accessible, differentiable and actionable, see table 2.

Table 2: segmentation criteria, adapted from Kotler and Keller (2012)

Criterion	Meaning
Measurable	the size, purchasing power, and characteristics of the segment can be measured.
Substantial	the segments are large and profitable enough to serve.
Accessible	the segments can be effectively reached and served.
Differentiable	the segments are distinguishable and respond differently to different marketing mix elements and programs.
Actionable	effective programs can be formulated for attracting and serving the segments.

Wedel and Kamakura (2000) propose six segmentation criteria, namely identifiability, substantiality, accessibility, stability, responsiveness and actionability. According to Wedel and Kamakura (2000), using preferences as segmentation basis leads to identifiable and substantial segments, meaning that the members of the segments are identifiable based on variables that are easily measured. In that sense identifiability is comparable to what Kotler and Keller (2012) define as measurable. Segments are substantial if they represent a large enough portion of the market to ensure profitability. These two segmentation criteria were used in this research to see whether segmentation of the respondents was effective. Accessibility, stability, responsiveness, differentiability and actionability are not used in this research, because it is not measured in any way whether the segments can be effectively reached, or how they respond to marketing efforts, for example. This is outside the scope of the research.

Methods for market segmentation can be classified in a priori and post hoc and in descriptive and predictive. A priori means that the number of segments was determined before the research and post hoc indicates that the number of segments depends on the data analysis (Wedel & Kamakura, 2000). Predictive segmentation methods have a set of dependent variables explained by independent variables and descriptive methods make no distinction between independent and

dependent variables. An example of an a priori descriptive method is using contingency tables. A priori predictive methods are cross-tabs and discriminant analysis and a post hoc predictive method is clusterwise regression, for example. Cluster analysis, which was applied in this research, can be considered a post hoc descriptive method (Wedel & Kamakura, 2000). This method was chosen, because no predefined idea exists of the segments that were to be found and therefore post hoc methods are more suitable.

Once segments are found based on preferences of respondents and the segmentation has been effective according to the criteria used, it is interesting to look further into the characteristics of the members that were assigned to the clusters and see whether there are similarities and differences with regards to the demographics and lifestyle data. When similarities within clusters and differences between clusters are found, this could be an indicator that segmentation based on demographic / lifestyle data could also provide useful information.

2.4. Segmentation method 1: two-stage approach using multiple linear regression and cluster analysis

2.4.1. Generating ideal points (Iacobucci, 2013)

In the research by Van den Broek (2019), the methodology for creating a perceptual map with attribute vectors as described in Iacobucci (2013) was followed. In Iacobucci (2013), also a method for obtaining ideal points for each respondent is described and this method was tested on the summer apples data.

The dimension coordinates (dim1 and dim2) for each apple variety obtained from MDS, were used as independent variables in multiple linear regressions, along with the sum of squares (SS) of dim1 and dim2, which is calculated via $SS = (dim1^2) + (dim2^2)$, for each apple variety. Iacobucci (2013) mentions to use the standardized dimension coordinates as input for the regression¹, however in this research the dimension coordinates as produced by MDS are used, because these coordinates give information on the scale on which apples positions differ from one another and standardizing them would diminish the possibility of being able to interpret the distances. In the multiple regressions per respondent, the preference scores for all apples per respondent are used as dependent variable.

The input for SPSS takes a form of apple varieties presented in rows with both the name of the variety and the number presented in two columns. Dim1, dim2 and SS are also presented in columns. This comes down to 10 rows and 5 columns. In the same file, the respondents preference ratings are presented in 54 columns, one for each respondent. Therefore the entire input file consists of 10 rows and 59 columns. Because not all respondents rated all ten apple varieties, there is missing data in the columns that show the preference ratings. For each single respondent a regression was run using the preference ratings as dependent variable and the dim1, dim2 and SS as independent variables. The unstandardized regression weights for dim1, dim2 and SS are used for further analysis. These are called b1, b2 and b3, respectively. The coordinates of the ideal points, new_b1 and new_b2, per respondent were calculated as follows: $new_b1 = b1 / (-2 * b3)$ and $new_b2 = b2 / (-2 * b3)$.

When plotting these coordinates new_b1 and new_b2 in a scatterplot it was found that outliers were present and the data was inspected visually to identify the outliers. Respondent 8, 15, 16, 19, 20, 35, 37 and 39 were chosen to be removed from the dataset, based on the visual inspection. The outliers were so far away from the other data points that the majority of points could not be distinguished

¹ In that case the calculation of SS as proposed would be correct (no subtraction of the mean is done). However, in this research the unstandardized values were used and the calculation of SS without mean subtraction was applied.

from one another. Removing the outliers lead to a visible group of data points. Also each apple variety was plotted into the same scatterplot. It was shown where the ideal points per respondent lie in the perceptual map compared to where the summer apple varieties lie. The coordinates of the ideal points were used as input for cluster analysis to be able to distinguish preference clusters in a two-step clustering approach, which will be discussed in the next paragraph.

2.4.2. Cluster analysis

hierarchical clustering (agglomerative and divisive)

The two-step approach to obtain preference clusters based on ideal points of respondents was also described by Wedel (1990) and the point of departure is using the ideal points per respondent as input for the cluster analysis. According to Aldenderfer, Mark and Roger (1984) a clustering method is “a multivariate statistical procedure that starts with a data set containing information about a sample of entities and attempts to reorganize these entities into relatively homogeneous groups.”

The ideal point coordinates were fed into an hierarchical cluster analysis, using the squared Euclidean distance and Ward’s method. Hierarchical clustering techniques can be divided in agglomerative and divisive techniques. In agglomerative techniques each respondent starts in his/her own cluster and in steps clusters are merged. In divisive techniques all respondents start in one big cluster and in steps the big clusters is divided into smaller clusters. Both techniques lead to hierarchies of clusters (Iacobucci, 2013), in this research agglomerative clustering was applied. Ward’s method of minimum variance is applied, which is a hierarchical clustering technique in which the variance within the cluster is minimized and the variance between clusters is maximized (Iacobucci, 2013). Ward’s method is suitable to use when the number of observations in each cluster is expected to be approximately equal and when there are no outliers (Ketchen & Shook, 1996). The Euclidean distance is the direct distance between two respondents which can be calculated by using the Pythagoras theorem. The squared Euclidean distance uses the same metric, but the square root is not taken. The (squared) Euclidean distance is the most widely used distance measure in cluster analysis and can be used in both hierarchical and k-means clustering (Kaufman & Rousseeuw, 2009).

An agglomeration schedule that shows the agglomeration coefficient per clustering step was generated, which was used as a criterion to decide upon the amount of clusters. Cluster members are assigned based on distances, and when the agglomeration coefficient becomes large, this indicates that cluster members are assigned to a cluster even though the distance is actually large. This why a one cluster solution has the largest agglomeration coefficient. Therefore, the amount of clusters before merger in the step where the agglomeration coefficient makes a large jump is the guideline for the amount of clusters that should be chosen. It was found that for the summer apples data the agglomeration coefficient made a jump when merging from 3 to 3 clusters and therefore a three cluster solution was chosen for. The ideal points are shown in figure 1, with 33 respondents assigned to cluster 1, 2 respondents assigned to cluster 2 and 11 respondents assigned to cluster 3. The cluster centroids from the hierarchical procedure were be saved and used as input for k-means clustering, which will be discussed in the next paragraph.

Hierarchical clustering as input for k-means clustering (non-hierarchical)

Charrad, Ghazzali, Boiteau and Niknafs (2015) state that often hierarchical clustering is applied to get an initial idea of the amount of clusters and then this number of clusters is fed into k-means clustering as the k amount of clusters. K-means clustering is a partitioning technique and it is an iterative process in which respondents are re-assigned to the k amount of clusters in steps and the cluster mean is constantly recalculated. The respondents are re-assigned to the cluster based on the distance between the respondent and the new cluster centroids. Once the iteration converges, this means no new optimal re-assigning of respondents can be done and the final clusters and cluster

memberships are obtained (Iacobucci, 2013). Based on the cluster memberships the respondents are grouped and then the mean of the ideal points of the members belonging to each cluster can be calculated to obtain the ideal points per cluster of respondents. These final cluster centres are also given by SPSS. The cluster centroids can then be plotted in the perceptual map to show the position of the ideal apples for the three consumer preference clusters, see figure 2.

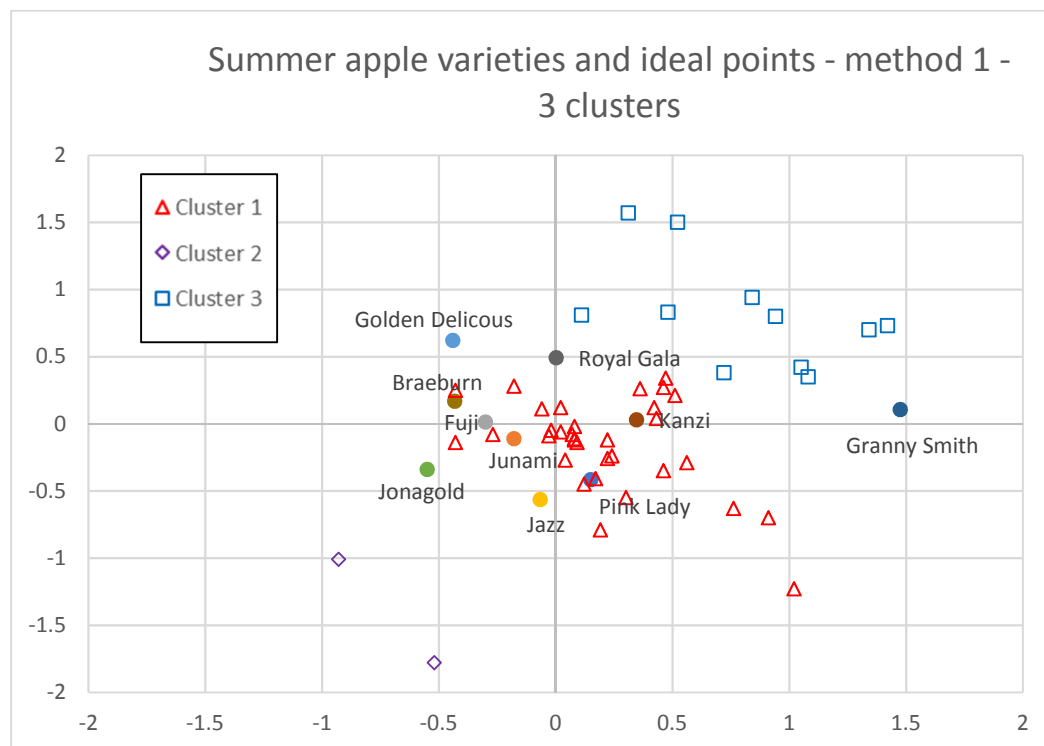


Figure 1: summer apple varieties and ideal points - method 1 - three clusters

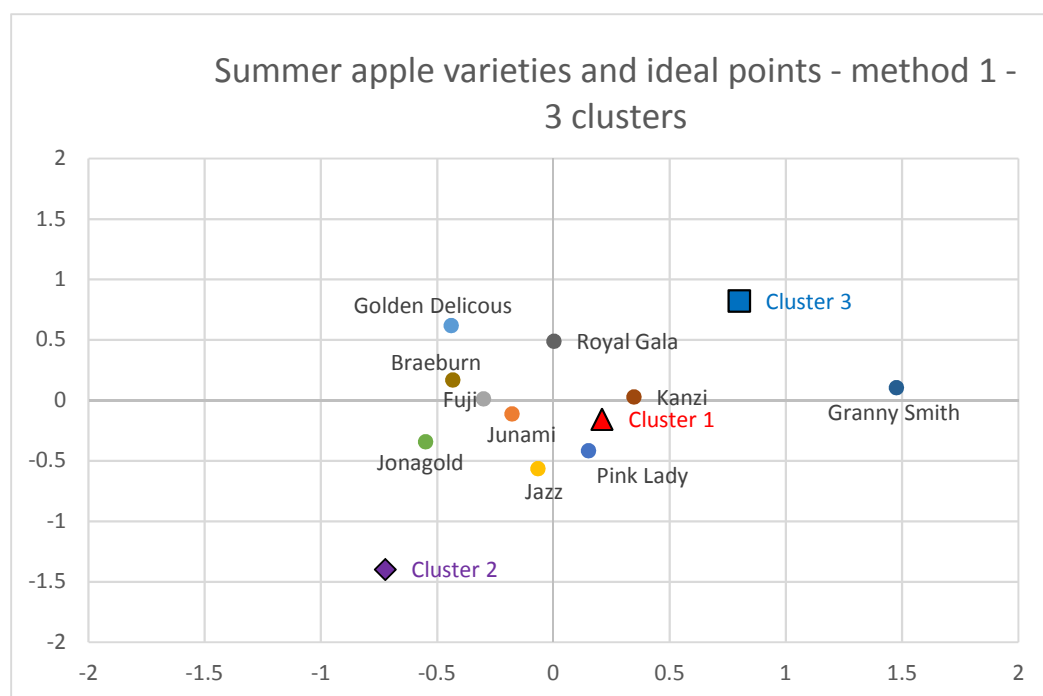


Figure 2: summer apple varieties and ideal points - method 1 - three clusters shown as centroids

2.5. Segmentation method 2: two-stage approach using MDS PREFSCAL and cluster analysis

Using multiple linear regression on the preference scores and dimensions obtained from MDS and the SS calculated from these dimensions, is one way to obtain ideal point coordinates, as shown in the previous method. However, SPSS also offers the possibility to obtain ideal points by using PREFSCAL. PREFSCAL offers a multidimensional unfolding analysis, which is a way of representing individual differences in preference judgements (Zhang & Takane, 2010).

The preference ratings are considered as a kind of dissimilarities between the apple varieties, that need to be transformed into distances. A transformation function is used for this, in the form of a linear equation, including an intercept and slope ($y = ax + b$, with a being the slope and b being the intercept). The dissimilarities are then transformed via that linear equation into distances and simultaneously the transformation function is matched to a criterion, which is the loss function. The loss function is to be minimized in the process of multidimensional unfolding, which is an iterative process. Mathematical support for this can be found in Busing (2010), however this is out of the scope of the research. Finally, an optimal solution for the transformation function is obtained via which the distances are obtained for the preference scores. Combining these distances leads to ideal points per respondent. These ideal points differ per respondent because the ratings and thus the distances differ and the apples that were rated by each respondent also differ.

PREFSCAL uses stress as a loss function, as well as PROXSCAL, however in PREFSCAL a penalty term on the intercept of the transformation function is added to prevent degenerate solutions (Busing, 2010). The penalized stress function contains parameters that influence the balance between the stress and penalty term and one of these parameters is the range. In SPSS the range can be set to a value by the user and setting the range to zero, means that the penalty is shut off. Changing the range has no effect when no intercept is estimated, because the penalty was included to influence the intercept and prevent degenerate solutions.

Solutions can be trivial, partially degenerate, absolutely degenerate, etc. but what all these solutions have in common is that the analysis gives solutions that are a perfect fit in terms of the loss function, but meaningless in terms of interpretation. Degeneracy of solutions is one of the big challenges in multidimensional unfolding (Busing, 2010). Degenerate solutions occur when transformations of the proximities are done and free estimation of the intercept and slope is allowed (Busing, 2010).

In this research, the data is entered with the respondents being represented on the rows and the apple varieties being represented in the columns. The preference ratings for the apples by the respondents are used for input in the PREFSCAL analysis and they are recoded from 1 – 7 to 6 – 0, so that the maximum value for liking is 0 and the program identifies this as the maximum value in the transformations. Because not all respondents rated all apples, there is missing data, however it is reported that PREFSCAL can deal with missing data that is caused by using a Balanced Incomplete Block Design (Busing, 2010). The proximities are then also changed from similarities to dissimilarities because the lowest value after recoding stands for highest liking. Transformation of the preference scores is achieved by use of a linear equation, minimizing the loss function, and no intercept is estimated. The range by default is set to 0.1 and is kept that way. The penalty term would prevent an undesirable intercept leading to identical transformations, which in turn would lead to degenerate solutions (Busing, 2010).

In this research no intercept is estimated so the penalty term has no influence. Restrictions are put on the columns, meaning that the dimensions obtained from the PROXSCAL procedure representing the apple positions are entered in the analysis and PREFSCAL needs to estimate the ideal points

taking the apple positions into account. It is chosen to let SPSS display the table with Stress composition, showing the stress per respondent. A higher Stress indicates a worse fit and could be an indicator that a respondent should be removed as outlier. According to Kruskal (1964), there are values that indicate the goodness of fit of the multidimensional scaling process, a Stress of 0.2 or higher would indicate a poor fit. None of the respondents exceeded an individual value of 0.2 and thus all respondents were included. The ideal points per respondent as generated by PREFSCAL were used as input in a two stage clustering procedure which is performed in the same manner as in method 1. It was found that 4 clusters could be identified with respectively 8, 15, 13 and 18 members, see figure 3 and 4.

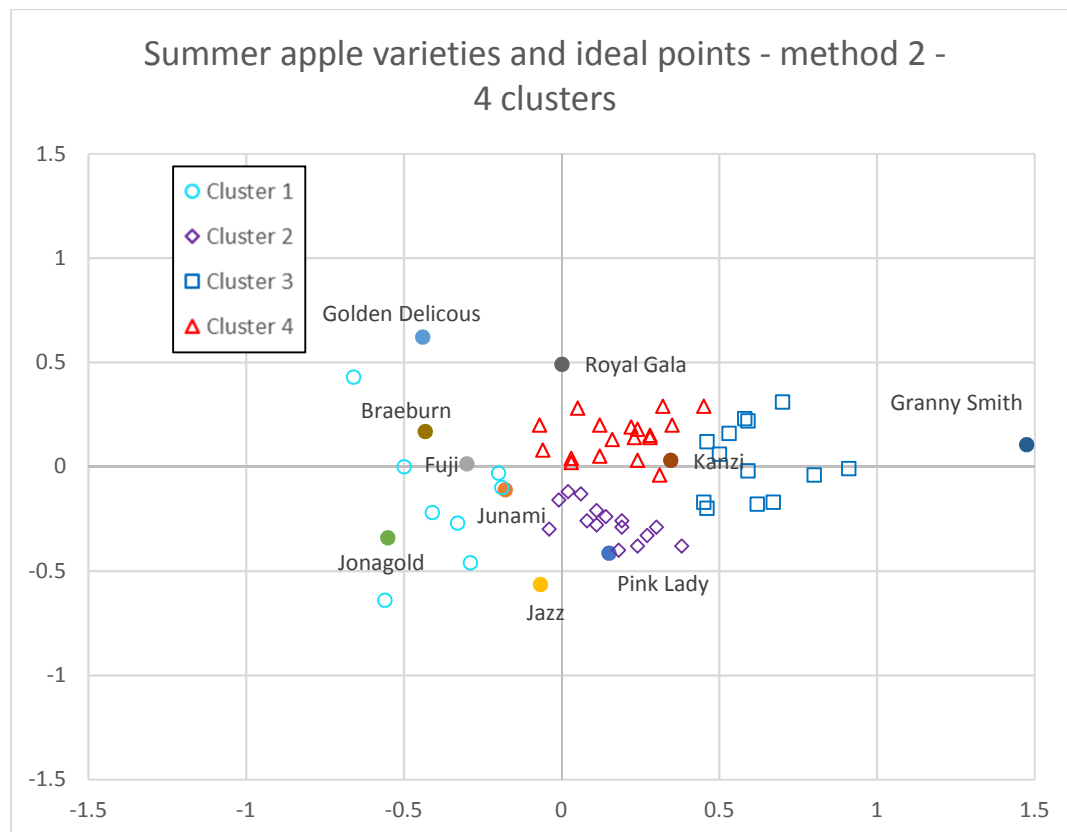


Figure 3: summer apple varieties and ideal points - method 2 - four clusters

2.6. Comparing method 1 and 2

When reflecting on these results obtained from method 1 and method 2, it can be said that following method 1 did not lead to substantial and measurable clusters. A cluster of two respondents emerges to the bottom of the map (not substantial), and when expecting these, they come across as outliers. Taking into account that already 8 respondents on a sample of 54 have been removed as outliers, it seems to be not wise to delete even more respondents from the data set, because this will lead to an unreliable representation of the results. Besides, a cluster of 11 respondents was found with a large spread on the top right of the map (not measurable), and one big cluster of 33 respondents around the majority of the apples was found, which seems to give no information about groups of respondents having a preference for certain apples, at most a non-preference for Granny Smith apples. Compared to method 1, method 2 leads to more substantial clusters, with respondents being more evenly spread across clusters and also visually clusters are easier to distinguish than when using method 1.

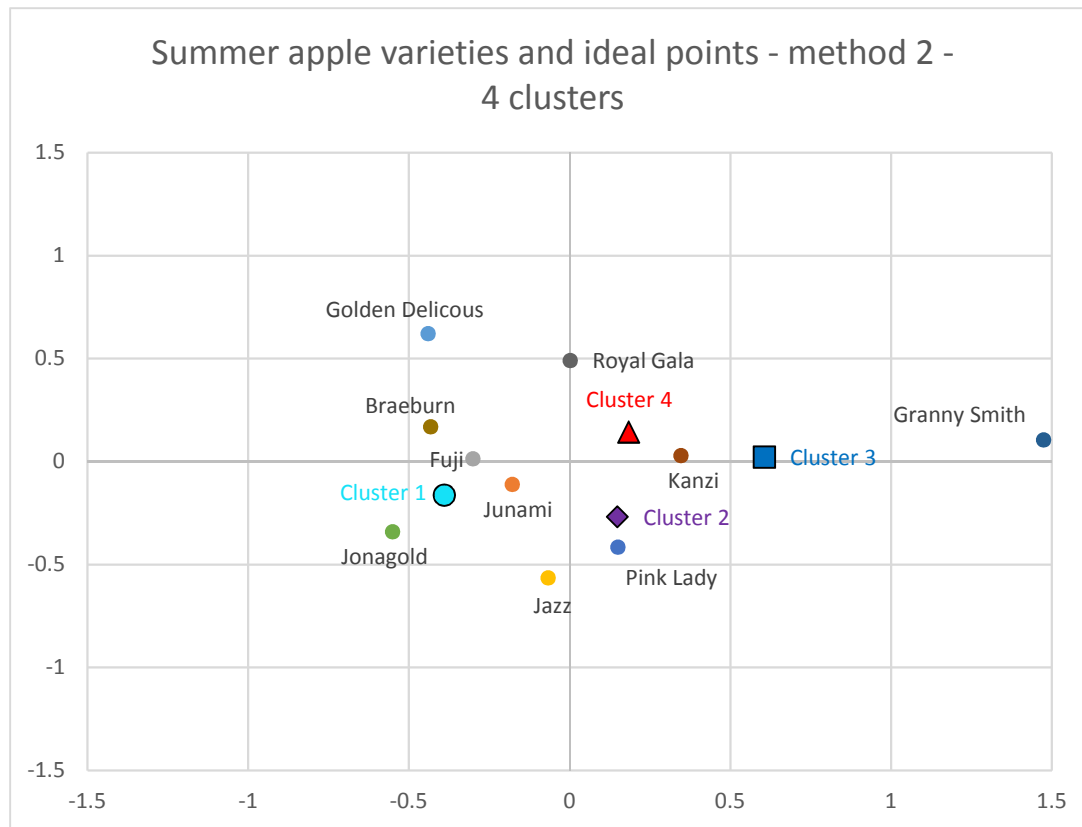


Figure 4: summer apple varieties and ideal points - method 2 - four clusters shown as centroids

In method 2 no respondent was identified as outlier and no respondent needed to be removed from the dataset. The reason for this is that in method 1, ideal points are calculated based on a regression of the dimensions that are obtained from MDS, and a manually calculated sum of squares. Calculating the sum of squares of the dimensions gives information about the spread of the dimensions. In a regression the preference scores are then fitted to the dimensions and the spread of the dimensions. After that, to obtain ideal points, the regression weights for a respondent, b_1 and b_2 , are corrected for the regression weight obtained for the SS, by dividing by $-2 * b_3$. One problem is that the sum of squares is calculated by using $(dim1^2) + (dim2^2)$, while normally the sum of squares is calculated by taking an observation, subtracting the mean of observations and then take the squared values and take the sum of the squared values. This is just a sum of the squared values of $dim1$ and $dim2$, but no mean is subtracted. Iacobucci (2013) states to do this, because standardized dimension coordinates are assumed to be used, with $\mu = 0$ and $\sigma = 1$. However, in this research the unstandardized values are used, in which the mean is not 0. What is also problematic with method 1, is that the regressions are done individually for each respondent, meaning that for each respondent a linear equation is fitted, based on the data available. Given the fact that respondents rated different combinations of apples and that for each respondent there are 5 missing values and 5 observations, the variation of outcomes of the individual multiple regressions will be large across respondents, leading to the outliers that were identified.

In method 2, the entire dataset of preference scores, including missing values, is transformed into distances by using a linear equation, by which values are assigned to the certain dissimilarities. For each respondent, the distances that are obtained are then used to calculate where the respondent would be in the given space, based on their distances to the given apple positions. This leads to the coordinates of the ideal points. Also, in method 2, there is an optimization step in which the outcome is matched to a fit criterion in an iterative process and this is lacking in method 1. In method 1,

multiple parameters are estimated (beta regression weights), while there are only 5 observations for each respondent, leading to little degrees of freedom left for further estimation. Method 2 in that sense is more statistically solid, since there are more observations included and more degrees of freedom left.

This chapter has answered to sub question SRQ1: ‘what segmentation methods lead to preference clusters of the summer apples data that meet the criteria for segmentation?’. Overall it can be stated that method 2 seems to lead to more substantial clusters, and is statistically more sound when it comes to missing data and thus this method will be used to analyse the winter apples data.

Chapter 3: Methodology

In this chapter the methodology and procedure for obtaining the winter apples data are discussed. Subsequently, the data analyses used to be able to get the results necessary to answer the remaining sub research questions and finally the main research question are discussed. This is a descriptive research in which primary data will be gathered by using surveys on sensory evaluations done by consumers, the data will be quantitative and analysed using SPSS version 23 by IBM.

3.1. Study design

This research focusses on the following research questions:

‘What is the position of winter apples based on consumer perceptions of these apples and what consumer preference clusters can be distinguished?’

SRQ 1: *What segmentation methods lead to preference clusters of the summer apples data that meet the criteria for segmentation?* This research question was answered in chapter 2. Methods were tested on the summer apples data and the clusters that were found were matched to segmentation criteria to find out what the most viable segmentation method was to be used on the winter apples data. It was found that using Segmentation method 2, in which MDS PREFSCAL and additional two-stage clustering is applied, leads to preference clusters that meet the posed criteria for segmentation.

SRQ 2: *What is the position of existing winter apple varieties based on consumer perceptions of these apples?* The methodology by Van den Broek (2019) was followed to obtain a perceptual map by using MDS PROXSCAL and including attribute vectors by applying attribute vector fitting via multiple linear regressions. Proximity data and attributes ratings were gathered to be able to answer this research question.

SRQ 3: *What preference clusters can be distinguished based on the winter apples data?* To answer this question, ideal points per respondent were generated by using MDS PREFSCAL and additional two-stage cluster analysis was applied to see whether preference clusters can be distinguished and if they meet the posed segmentation criteria. Preference data and demographic/lifestyle data was gathered to be able to answer this research question.

3.1.1. Sampling method

Selection of supermarkets

The conduction of sensory evaluations for the winter apples took place in Dutch supermarkets within a radius of maximally 35 km from Wageningen and 3 km minimally away from Wageningen in the period of the 6th of January 2020 to the 17th of January 2020. Supermarkets within a range of 3 km from Wageningen were excluded due to possibly biased inhabitants (students) and thus respondents.

Albert Heijn (AH) is market leader in the Netherlands with a market share of approximately 35%, followed by Jumbo, the biggest competitor of AH with 22% market share (Retailtrends, 2019). Because AH and Jumbo together have more than 50% market share, it was decided to include AH and Jumbo supermarkets. A calling list was composed including all AH and Jumbo supermarkets in the specified KM range. The supermarkets were contacted in the period of 25th of November 2019 to 20th of December 2019. A calling script was used, in which the researcher asked for the person that was responsible for instore activities in the supermarket. Often this was the retailer him/herself or the person responsible for ‘DKW’ (Droge Kruideniers Waren) or ‘AGF’ (Aardappelen, Groente en Fruit). A short introduction of the research was given in which the researcher explained that for this study, data needed to be gathered in Dutch supermarkets by letting customers taste apples and fill in a

questionnaire. It was asked whether it was possible to conduct the research in the contacted supermarket during one day in January 2020. Sometimes the responsible person immediately (dis)agreed, sometimes they asked for more information via an e-mail, sometimes the responsible person was not present at the time and the supermarket was contacted again at another moment in time. When a supermarket in a certain city/village was included in the research, the rest of the AH's and Jumbo's in that city/village were excluded, due to random sampling reasons and to prevent including the same respondents twice.

In total twelve supermarkets were selected. The spread of the supermarkets can be seen in figure 5. Ten supermarkets were chosen, that are shown with a marker with a black dot and two supermarkets were chosen as spare locations in the case the data could not be collected in the first ten supermarkets. These are shown with a marker with a white dot. The visited supermarkets are also shown in table 3.

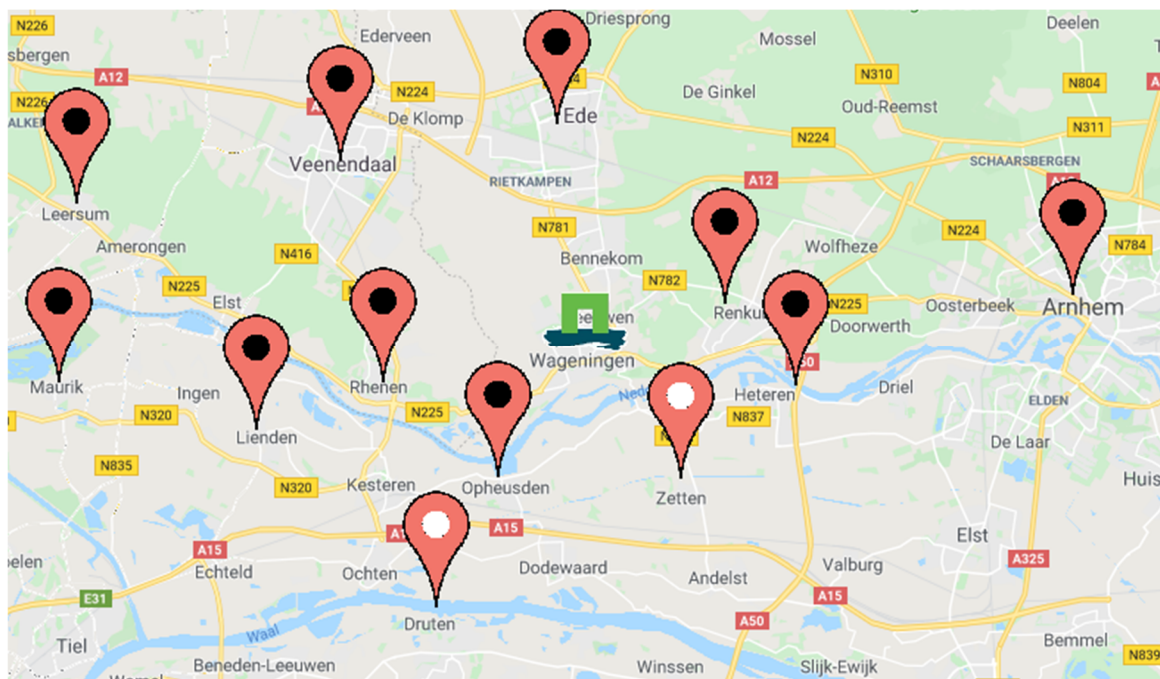


Figure 5: spread of the selected supermarkets in the region

Table 3: visiting appointments supermarkets

Date	Supermarket
6-1-2020	Jumbo Ronde Erf Veendam
7-1-2020	AH XL Keesomstraat Ede
8-1-2020	AH Dorpsstraat Renkum
9-1-2020	AH Swaenenstaete Opheusden
10-1-2020	Jumbo Plantsoenstraat Rhenen
13-1-2020	AH Fortunastraat Arnhem
14-1-2020	Jumbo Raadhuisstraat Maurik
15-1-2020	Jumbo Verbrughweg Lienden
16-1-2020	Jumbo Beemdhof Heteren
17-1-2020	AH Honingraat Leersum

Participants

Respondents were selected by asking the Dutch customers that were present in the selected supermarkets whether they wanted to participate in the research. The sample (N = 162) consisted of males (39.5%) and females (60.5%) that were born between 1936 and 2002.

3.1.2. Stimuli

Ten apple varieties were selected and supplied by the wholesaler of apples to Wageningen Research Open Crops. The apple varieties included in the research are shown in table 4. These apples were included, because in this time of year these apples are available in the Dutch supermarkets. On a daily basis the apples to be used were taken out of a cooler at Wageningen Research Open Crops a day in advance, and the apples were then picked up at Randwijk on a daily basis to maintain constant quality of the apples that were used. The attributes included in this research were comparable to those used by Van den Broek (2019) except that “skin toughness” was changed into “toughness” (of the whole apple). Besides, “firmness” (Van den Broek, 2019) was divided into “firmness of the peel” and “firmness of the fruit pulp”. This was done in accordance to the attributes rated by an expert panel, which is another research within the research project ‘Sustainable fruit introductions better understood’. The attributes included in this research are shown in table 5.

Table 4: apple varieties included in the research

number	Apple variety
1	Granny Smith
2	Jonagold
3	Elstar
4	Royal Gala
5	Pink Lady
6	Golden Delicious
7	Braeburn
8	Fuji
9	Jazz
10	Kanzi

Table 5: attributes that the apples were evaluated on

Classification	Category	Attribute
Search	Appearance	Peel colour
		Size
		Glossiness
Experience	Texture	Toughness
		Crispiness
		Firmness peel
		Firmness fruit pulp
		Juiciness
		Sourness
		Sweetness
	Flavour	Aroma

Incomplete design

In this research an incomplete design was used, meaning that not all ten varieties were rated by one respondent. This is due to the fact that too many apple combinations and too many attributes would

have to be rated, leading to sensory fatigue. A Balanced Incomplete Block Design (BIBD) was generated using RStudio 1.2.1335. More specifically, the package ‘crossdes’ was used to generate an incomplete data design with ten treatments in 54 blocks of 5 elements, see table 6.

Table 6: incomplete design generated with crossdes

Block	Apple 1	Apple 2	Apple 3	Apple 4	Apple 5
1	2	3	4	6	7
2	1	2	4	5	8
3	1	2	7	8	9
4	1	3	5	7	10
5	2	3	6	7	10
6	1	3	7	8	9
7	6	7	8	9	10
8	4	5	6	7	9
9	2	5	8	9	10
10	1	3	4	7	9
11	1	2	6	7	10
12	2	3	5	8	9
13	2	3	5	7	8
14	1	4	6	9	10
15	1	4	7	8	9
16	1	2	3	5	10
17	2	4	5	6	8
18	2	3	5	7	10
19	1	5	7	8	10
20	2	4	6	9	10
21	3	5	6	8	9
22	1	3	4	7	8
23	3	4	7	8	10
24	1	2	4	5	6
25	1	2	6	8	9
26	3	4	5	9	10
27	1	4	5	6	10
28	2	5	7	8	9
29	1	5	6	8	9
30	2	3	4	9	10
31	1	2	3	4	9
32	4	5	6	7	8
33	1	2	5	6	7
34	3	4	8	9	10
35	3	5	6	7	9
36	1	2	4	8	10
37	1	2	3	6	7
38	4	6	7	8	10
39	4	5	7	9	10
40	1	3	4	5	6
41	1	3	6	7	9
42	1	3	5	8	10
43	1	4	5	7	10
44	1	3	6	8	10
45	1	2	5	9	10
46	3	4	5	6	8
47	2	4	5	7	9
48	2	3	6	8	10
49	1	6	8	9	10
50	1	2	3	4	9
51	2	3	4	6	8
52	2	6	7	9	10
53	3	5	6	9	10
54	2	4	7	8	10

3.2. Data collection method / procedure

Every day the researcher and colleague were present around 08.30 in the supermarket, half an hour was needed to give the supermarket notice of presence and to set up the research in the supermarket. A sampling desk was set up, with a laptop placed on it and a barstool beside it, so a respondent had the possibility to sit when doing the evaluations, see figure 6. In practice, the sensory evaluations took place between 9.00 and 17.00. Effectively the researchers were gathering respondents and conducting the research during 7.5 hours a day, meaning that approximately 30 minutes were needed per respondent (including the recruiting) and one respondent was handled at the time.

Participants were approached when doing their groceries. The researchers, wearing a white blouse and a green apron and holding a serving tray with an apple, posed the question: ‘hello sir/madam, do you sometimes consume apples?’ When the answer is ‘yes’, the researcher asked: ‘would you like to taste and evaluate some apple varieties? You will help us finish our study and you can win a Bol.com gift card of 75 euros! It takes only ten minutes of your time’. It was chosen to wear a white blouse and green apron, because this outfit is used often in instore brand activations and consumers will associate the look with obtaining a free sample. The colour green was chosen for the apron, because it is expected that green is associated with fruits and vegetables and people that sell fruits.

It was explained what the respondent could expect from the questionnaire, if they needed help or had questions, the researchers answered the questions and helped the respondents. The whole apples were used to evaluate the appearance attributes (peel colour, size and glossiness) and one experience attribute (toughness), and the apples were simultaneously presented on round, white, paper plates labelled from one to five. The stickers showing the apple variety name were placed at the bottom of the apples to prevent the researchers from making mistakes. The respondents were instructed not to inspect the bottom of the apple. Two slices of apple of each variety were given to the respondent, when they finished evaluating the whole apples. The apple slices were used to evaluate the crispiness, firmness of the peel, firmness of the fruit pulp, juiciness, sourness, sweetness, and aroma.



Figure 6: impression instore data gathering



The respondents were instructed to start with the first apple and answer all questions for the first apple and then neutralize their taste and continue with the second apple and again answer all questions, until all five apples were evaluated. The reason for this is, that otherwise the respondent would have to taste (five apples times seven attributes) 35 times and this took too much of their time. Napkins were available, as well as unsalted crackers and water for neutralisation. All apple varieties were held in separate trays, out of sight of the respondents. The apples were cut behind the sampling desk out of sight of the respondents. In the time that respondents evaluated the whole apple, the slices were cut, so that the browning of the slices was minimized.

Questionnaire / survey

The questionnaire for this research was set up by using Qualtrics and the questionnaire by Van den Broek (2019) was used as a basis. The following changes to the questionnaire were made: a definition of the attribute 'aroma' was added, because in previous research it appeared this attribute was not clear to respondents. The order of the questions on the attributes aroma, sourness and sweetness has been changed to sweetness, sourness and aroma, because aroma captures the elements of flavour that are not covered by sweetness and sourness, so the order seemed more logical. The scale in the question on consumer preference was a 0 to 7 scale, which is corrected to a 1 to 7 liking scale, because this scale is more commonly used in social sciences. All scales were centred, except for the preference scale. This means that the indicator was centred in the middle at the value 50. This was done so that respondents were not biased towards one end of the scale. In the case of preference this was not done.

As mentioned before, skin toughness was rephrased to toughness and firmness was divided into firmness of the peel and firmness of the fruit pulp. A question on education level was added, research by Kelley, Hyde, Travis and Crassweller (2010) suggests that education level influences whether sweetness is considered an important attribute. In the questionnaire four types of data were gathered.

Questions one to eleven consisted of questions on the attributes of the apples, respondents were asked to give ratings on the before mentioned attributes on a 0 to 100 scale. Question twelve consisted of ten proximity judgements in which respondents are asked to rate apple pairs on how different they are. The pairs that the respondents had to judge are displayed in table 7. An incomplete design with ten apples was used in which one respondent rated only five apples. This means that the apple pairs that respondents from differing blocks needed to rate, differed for each block. However, the apple pairs are called 1 and 2, 1 and 3 ... 4 and 5 in table 7. For example, the first paired comparison for the respondents in block 1 is called 1 and 2, in practice a paired comparison is made between apple 2 (Jonagold) and apple 3 (Elstar). In the seventeenth block, the first paired comparison is also called 1 and 2, but in practice a comparison is made between apple 2 (Jonagold) and apple 4 (Royal Gala).

In question thirteen preference data was gathered by asking the respondent how well the apples that were tasted were liked on a 1 to 7 scale. Finally in question fourteen demographic / lifestyle data is asked: gender, year the respondent was born, how many times the respondent consumes apples monthly and the education level. The questionnaire is included in appendix 1.

Table 7: apple pairs as rated by respondents per block

	Apple 1 and 2	Apple 1 and 3	Apple 1 and 4	Apple 1 and 5	Apple 2 and 3	Apple 2 and 4	Apple 2 and 5	Apple 3 and 4	Apple 3 and 5	Apple 4 and 5
1	2 and 3	2 and 4	2 and 6	2 and 7	3 and 4	3 and 6	3 and 7	4 and 6	4 and 7	6 and 7
2	1 and 2	1 and 4	1 and 5	1 and 8	2 and 4	2 and 5	2 and 8	4 and 5	4 and 8	5 and 8
3	1 and 2	1 and 7	1 and 8	1 and 9	2 and 7	2 and 8	2 and 9	7 and 8	7 and 9	8 and 9
4	1 and 3	1 and 5	1 and 7	1 and 10	3 and 5	3 and 7	3 and 10	5 and 7	5 and 10	7 and 10
5	2 and 3	2 and 6	2 and 7	2 and 10	3 and 6	3 and 7	3 and 10	6 and 7	6 and 10	7 and 10
6	1 and 3	1 and 7	1 and 8	1 and 9	3 and 7	3 and 8	3 and 9	7 and 8	7 and 9	8 and 9
7	6 and 7	6 and 8	6 and 9	6 and 10	7 and 8	7 and 9	7 and 10	8 and 9	8 and 10	9 and 10
8	4 and 5	4 and 6	4 and 7	4 and 9	5 and 6	5 and 7	5 and 9	6 and 7	6 and 9	7 and 9
9	2 and 5	2 and 8	2 and 9	2 and 10	5 and 8	5 and 9	5 and 10	8 and 9	8 and 10	9 and 10
10	1 and 3	1 and 4	1 and 7	1 and 9	3 and 4	3 and 7	3 and 9	4 and 7	4 and 9	7 and 9
11	1 and 2	1 and 6	1 and 7	1 and 10	2 and 6	2 and 7	2 and 10	6 and 7	6 and 10	7 and 10
12	2 and 3	2 and 5	2 and 8	2 and 9	3 and 5	3 and 8	3 and 9	5 and 8	5 and 9	8 and 9
13	2 and 3	2 and 5	2 and 7	2 and 8	3 and 5	3 and 7	3 and 8	5 and 7	5 and 8	7 and 8
14	1 and 4	1 and 6	1 and 9	1 and 10	4 and 6	4 and 9	4 and 10	6 and 9	6 and 10	9 and 10
15	1 and 4	1 and 7	1 and 8	1 and 9	4 and 7	4 and 8	4 and 9	7 and 8	7 and 9	8 and 9
16	1 and 2	1 and 3	1 and 5	1 and 10	2 and 3	2 and 5	2 and 10	3 and 5	3 and 10	5 and 10
17	2 and 4	2 and 5	2 and 6	2 and 8	4 and 5	4 and 6	4 and 8	5 and 6	5 and 8	6 and 8
18	2 and 3	2 and 5	2 and 7	2 and 10	3 and 5	3 and 7	3 and 10	5 and 7	5 and 10	7 and 10
19	1 and 5	1 and 7	1 and 8	1 and 10	5 and 7	5 and 8	5 and 10	7 and 8	7 and 10	8 and 10
20	2 and 4	2 and 6	2 and 9	2 and 10	4 and 6	4 and 9	4 and 10	6 and 9	6 and 10	9 and 10
21	3 and 5	3 and 6	3 and 8	3 and 9	5 and 6	5 and 8	5 and 9	6 and 8	6 and 9	8 and 9
22	1 and 3	1 and 4	1 and 7	1 and 8	3 and 4	3 and 7	3 and 8	4 and 7	4 and 8	7 and 8
23	3 and 4	3 and 7	3 and 8	3 and 10	4 and 7	4 and 8	4 and 10	7 and 8	7 and 10	8 and 10
24	1 and 2	1 and 4	1 and 5	1 and 6	2 and 4	2 and 5	2 and 6	4 and 5	4 and 6	5 and 6
25	1 and 2	1 and 6	1 and 8	1 and 9	2 and 6	2 and 8	2 and 9	6 and 8	6 and 9	8 and 9
26	3 and 4	3 and 5	3 and 9	3 and 10	4 and 5	4 and 9	4 and 10	5 and 9	5 and 10	9 and 10
27	1 and 4	1 and 5	1 and 6	1 and 10	4 and 5	4 and 6	4 and 10	5 and 6	5 and 10	6 and 10
28	2 and 5	2 and 7	2 and 8	2 and 9	5 and 7	5 and 8	5 and 9	7 and 8	7 and 9	8 and 9
29	1 and 5	1 and 6	1 and 8	1 and 9	5 and 6	5 and 8	5 and 9	6 and 8	6 and 9	8 and 9
30	2 and 3	2 and 4	2 and 9	2 and 10	3 and 4	3 and 9	3 and 10	4 and 9	4 and 10	9 and 10
31	1 and 2	1 and 3	1 and 4	1 and 9	2 and 3	2 and 4	2 and 9	3 and 4	3 and 9	4 and 9
32	4 and 5	4 and 6	4 and 7	4 and 8	5 and 6	5 and 7	5 and 8	6 and 7	6 and 8	7 and 8
33	1 and 2	1 and 5	1 and 6	1 and 7	2 and 5	2 and 6	2 and 7	5 and 6	5 and 7	6 and 7
34	3 and 4	3 and 8	3 and 9	3 and 10	4 and 8	4 and 9	4 and 10	8 and 9	8 and 10	9 and 10
35	3 and 5	3 and 6	3 and 7	3 and 9	5 and 6	5 and 7	5 and 9	6 and 7	6 and 9	7 and 9
36	1 and 2	1 and 4	1 and 8	1 and 10	2 and 4	2 and 8	2 and 10	4 and 8	4 and 10	8 and 10
37	1 and 2	1 and 3	1 and 6	1 and 7	2 and 3	2 and 6	2 and 7	3 and 6	3 and 7	6 and 7
38	4 and 6	4 and 7	4 and 8	4 and 10	6 and 7	6 and 8	6 and 10	7 and 8	7 and 10	8 and 10
39	4 and 5	4 and 7	4 and 9	4 and 10	5 and 7	5 and 9	5 and 10	7 and 9	7 and 10	9 and 10
40	1 and 3	1 and 4	1 and 5	1 and 6	3 and 4	3 and 5	3 and 6	4 and 5	4 and 6	5 and 6
41	1 and 3	1 and 6	1 and 7	1 and 9	3 and 6	3 and 7	3 and 9	6 and 7	6 and 9	7 and 9
42	1 and 3	1 and 5	1 and 8	1 and 10	3 and 5	3 and 8	3 and 10	5 and 8	5 and 10	8 and 10
43	1 and 4	1 and 5	1 and 7	1 and 10	4 and 5	4 and 7	4 and 10	5 and 7	5 and 10	7 and 10
44	1 and 3	1 and 6	1 and 8	1 and 10	3 and 6	3 and 8	3 and 10	6 and 8	6 and 10	8 and 10
45	1 and 2	1 and 5	1 and 9	1 and 10	2 and 5	2 and 9	2 and 10	5 and 9	5 and 10	9 and 10
46	3 and 4	3 and 5	3 and 6	3 and 8	4 and 5	4 and 6	4 and 8	5 and 6	5 and 8	6 and 8
47	2 and 4	2 and 5	2 and 7	2 and 9	4 and 5	4 and 7	4 and 9	5 and 7	5 and 9	7 and 9
48	2 and 3	2 and 6	2 and 8	2 and 10	3 and 6	3 and 8	3 and 10	6 and 8	6 and 10	8 and 10
49	1 and 6	1 and 8	1 and 9	1 and 10	6 and 8	6 and 9	6 and 10	8 and 9	8 and 10	9 and 10
50	1 and 2	1 and 3	1 and 4	1 and 9	2 and 3	2 and 4	2 and 9	3 and 4	3 and 9	4 and 9
51	2 and 3	2 and 4	2 and 6	2 and 8	3 and 4	3 and 6	3 and 8	4 and 6	4 and 8	6 and 8
52	2 and 6	2 and 7	2 and 9	2 and 10	6 and 7	6 and 9	6 and 10	7 and 9	7 and 10	9 and 10
53	3 and 5	3 and 6	3 and 9	3 and 10	5 and 6	5 and 9	5 and 10	6 and 9	6 and 10	9 and 10
54	2 and 4	2 and 7	2 and 8	2 and 10	4 and 7	4 and 8	4 and 10	7 and 8	7 and 10	8 and 10

Chapter 4: Data analysis & results

4.1. Sample characteristics

In total 162 Dutch adults (i.e. 54 blocks x 3 respondents) that were recruited in 10 supermarkets in the Netherlands participated in the research. The sample characteristics are displayed in table 8. The gender balance of the sample was aimed for to be 50% male and 50% female in accordance to the population. In practice more female respondents were included in the research, because it appeared that there were usually more women doing groceries. Comparing the age distribution to the numbers from CBS shows that the age distribution of the sample is comparable to the Dutch population (CBS Statline, 2019). Furthermore, the consumption habits show that the majority of respondents consume 0 – 10 apples per month (40.7%). When looking at the education level of the sample, it is found that the majority of the adults has a finished education of MBO or higher (83.9%) of which 19.1% has finished WO.

Table 8: sample characteristics, demographics and lifestyle

Sample characteristics	Responses	N
Gender	Male	64 (39.5%)
	Female	98 (60.5%)
Year the respondent was born	1930 - 1939	2 (1.2%)
	1940 - 1949	10 (6.2%)
	1950 - 1959	26 (16.1%)
	1960 - 1969	38 (23.3%)
	1970 - 1979	35 (21.6%)
	1980 - 1989	23 (14.2%)
	1990 - 2002	28 (17.3%)
Consumption habits (times per month)	0 to 10	66 (40.7%)
	11 to 20	45 (27.8%)
	21 to 30	36 (22.2%)
	31 to 40	7 (4.3%)
	41 or more	8 (4.9%)
Education level	VMBO	11 (6.8%)
	HAVO	10 (6.2%)
	VWO	5 (3.1%)
	MBO	54 (33.3%)
	HBO	51 (31.5%)
	WO	31 (19.1%)

4.2. Decompositional perceptual mapping

It was aimed for to obtain a joint space perceptual map, in which the apple positions are displayed based on the dissimilarity ratings that were found. The dissimilarity ratings were given on a 0 – 100 scale and for the entire set of respondents the ratings were sorted for each apple pair, e.g. 1 and 2 ... 9 and 10 and the average was taken. The averages were combined into a lower triangle dissimilarity matrix as displayed in table 9. A higher number indicates that the apples are perceived as more dissimilar. This matrix was the input for MDS PROXSCAL. An overview of the mean dissimilarities and standard deviations are given in appendix 2, the standard deviations range from 17.08 to 27.14.

Running the MDS PROXSCAL procedure for one to nine dimensions gave values for the normalized raw stress, which were plotted against the number of dimensions, leading to the graph in figure 7. As stated before, in MDS a representation of ten apple positions can be displayed in a dimensional

space ranging from one to maximally nine dimensions. Depending on the amount of dimensions a stress value is obtained, because a stress function is minimized. A low stress indicates a good fit, meaning that the lower the stress value, the better the fit. When deciding on one dimension, the stress is high and towards two, three and four dimensions the stress decreases. The plot indicates that four or five dimensions would be suitable. However, four or five dimensions would not lead to an interpretable map, and the lower the normalized raw stress becomes, the larger the risk of obtaining a degenerate solution. Therefore it was chosen to decide on two dimensions.

Table 9: lower triangle dissimilarity matrix

	Granny Smith (1)	Jonagold (2)	Elstar (3)	Royal Gala (4)	Pink Lady (5)	Golden Delicious (6)	Braeburn (7)	Fuji (8)	Jazz (9)	Kanzi (10)
Granny Smith (1)	0.00									
Jonagold (2)	81.06	0.00								
Elstar (3)	72.11	54.11	0.00							
Royal Gala (4)	80.75	51.72	51.44	0.00						
Pink Lady (5)	75.42	55.28	54.03	56.14	0.00					
Golden Delicious (6)	74.92	63.67	74.36	66.69	65.03	0.00				
Braeburn (7)	71.86	55.50	58.50	65.72	56.17	73.97	0.00			
Fuji (8)	77.89	57.22	51.33	55.19	53.47	64.78	53.61	0.00		
Jazz (9)	71.33	60.42	51.86	53.22	59.06	67.44	64.58	53.42	0.00	
Kanzi (10)	71.89	62.56	58.83	60.89	63.39	68.78	57.44	64.11	58.33	0.00

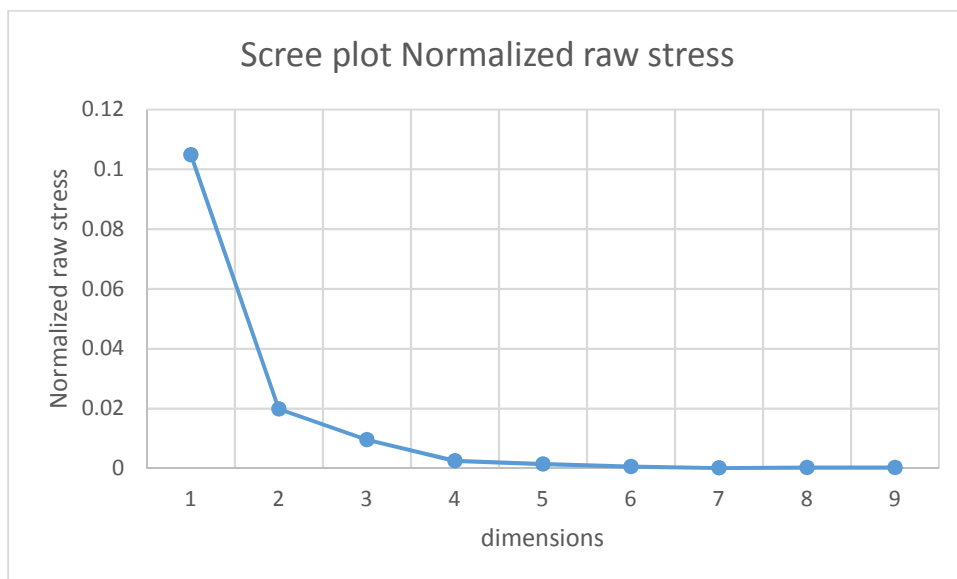


Figure 7: scree plot normalized raw stress

When deciding on two dimensions, the dimension coordinates per apple variety in table 10 are obtained. These coordinates represent the position of the apple varieties in the perceptual map.

Table 10: : dimension coordinates apple varieties

	Dim1	Dim2
Granny Smith	1.538	.169
Jonagold	-.238	.121
Elstar	-.108	-.310
Royal Gala	-.464	.078
Pink Lady	-.274	-.154
Golden Delicious	-.273	.961
Braeburn	.024	-.490
Fuji	-.397	-.151
Jazz	-.015	.046
Kanzi	.207	-.270

Plotting these coordinates lead to the decompositional perceptual map of the winter apples as displayed in figure 8. The axes were corrected to -2, 2, to show a realistic image of the distances between the apples and their position in the map. The Golden Delicious and Granny Smith lie further away from the rest of the apples. The closer the apples are to one another in the map, the more similar they are perceived by respondents.

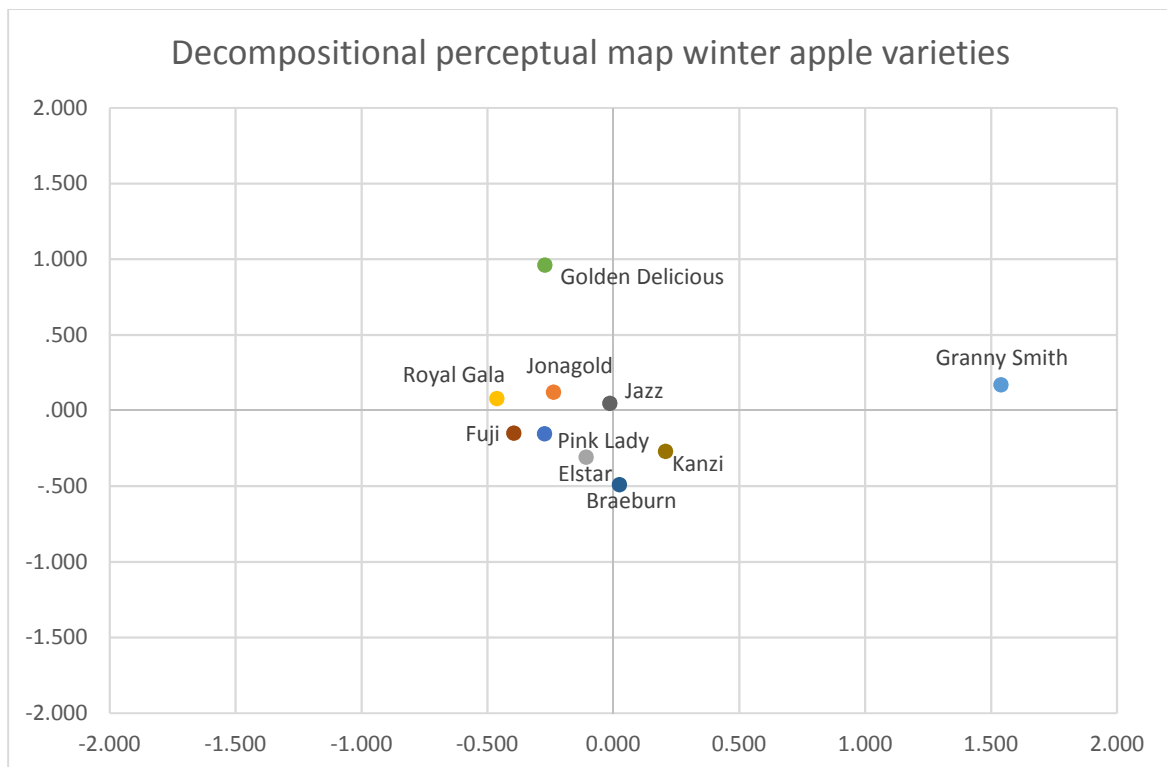


Figure 8: decompositional perceptual map winter apple varieties

4.3. Attribute vector fitting

To be able to interpret the perceptual map found in paragraph 4.2., attribute vector fitting was done. The attributes were rated on a 0 – 100 scale and for each attribute the mean and standard deviation of the means of the ratings for all respondents per apple variety were calculated and these means and standard deviations can be found in table 11. For example, looking at peel colour, Granny Smith

(11.14) and Golden Delicious (20.17) are considered to be the green apples, with Granny Smith being the most green and the Pink Lady (77.73) was considered to be the most red apple, on average. The standard deviation of the means were calculated to interpret whether there is agreement or disagreement across respondents about the attribute rating for a certain apple variety. A large deviation indicates disagreement and a small deviation indicates agreement. For example the standard deviation of the mean of glossiness for the Granny Smith (24.06), is the largest standard deviation found, it would indicate that there is more disagreement amongst respondents.

Table 11: attribute means and standard deviation of the mean per apple variety

	Peel colour	Size		Glossiness		Toughness		Crispiness		Firmness peel		
	Mean	St. dev.	Mean	St. dev.	Mean	St. dev.	Mean	St. dev.	Mean	St. dev.	Mean	St. dev.
Granny Smith	11.14	19.02	73.94	18.07	66.27	24.06	82.49	15.01	76.20	15.52	79.02	18.69
Jonagold	68.85	14.65	61.75	14.75	52.95	18.01	70.91	17.40	62.60	18.54	67.00	18.01
Elstar	61.37	16.62	60.37	15.53	56.10	17.95	68.17	19.58	71.63	15.40	71.54	13.44
Royal Gala	72.52	18.52	50.20	17.53	58.75	19.12	67.36	18.72	65.21	19.49	67.51	18.81
Pink Lady	77.73	14.41	61.67	14.62	61.15	18.89	68.89	17.84	67.88	18.48	71.05	16.39
Golden Delicious	20.17	18.55	65.37	14.06	35.64	20.20	65.84	17.17	53.25	21.25	58.47	19.51
Braeburn	67.57	17.63	68.95	15.92	65.89	15.75	73.93	15.26	60.81	23.49	67.21	18.64
Fuji	76.10	17.50	69.60	14.16	66.58	18.67	75.35	16.62	62.27	21.64	64.96	18.44
Jazz	76.94	15.39	61.25	18.06	73.26	15.78	75.10	16.58	73.41	15.71	75.35	14.42
Kanzi	66.84	15.22	58.12	15.01	60.83	18.31	71.90	15.86	66.22	19.33	68.37	18.66
	Firmness fruit pulp		Juiciness		Sourness		Sweetness		Aroma			
	Mean	St. dev.	Mean	St. dev.	Mean	St. dev.	Mean	St. dev.	Mean	St. dev.		
Granny Smith	77.37	15.09	66.84	18.09	73.93	20.64	34.75	20.61	52.20	23.34		
Jonagold	56.94	19.74	67.43	16.00	39.16	18.03	64.53	14.54	57.51	18.70		
Elstar	67.59	16.97	73.10	13.69	46.60	22.54	61.74	17.02	66.12	16.62		
Royal Gala	61.30	18.31	62.79	19.62	34.07	17.14	62.49	19.66	51.74	20.97		
Pink Lady	69.75	16.74	66.09	17.74	44.70	21.00	60.37	18.67	65.10	18.39		
Golden Delicious	48.64	17.79	61.42	17.82	32.58	18.26	56.72	22.50	48.30	19.61		
Braeburn	59.02	22.57	63.32	19.71	45.21	23.33	58.46	20.54	55.99	22.45		
Fuji	60.41	21.28	66.93	18.71	32.93	21.20	72.09	18.81	60.47	19.86		
Jazz	70.58	17.14	72.43	14.98	36.56	20.15	66.36	15.87	62.85	19.13		
Kanzi	61.52	20.80	69.19	16.28	58.36	19.99	52.59	21.73	62.40	18.57		

For each attribute a regression was done. The mean attribute ratings were used as dependent variables in the regression analysis and the dimensions obtained from the MDS were used as independent variables. The outcomes of the regression analyses are displayed in table 12, containing the regression weights representing the coordinates of the arrow heads of the attribute vectors. The

R^2 value is included, which is a measure for how well the attribute ratings are explained by the dimension coordinates and the higher the R^2 value, the better the data is explained by the regression formula (the model). The maximum value that R^2 can take is 1.0, indicating that the attribute ratings are explained perfectly in relation to the dimension scores via the regression formula. The R^2 value gives information of the strength of the relationship between the model and the dependent variable, but the F test statistics and the significance are also displayed to draw conclusions about the statistical significance. When the F-test is significant, this means that the defined model explains the data better than a model with no predictors (H_0). Peel colour, glossiness, toughness, firmness of the peel, sourness and sweetness are significant, since the p-value is smaller than $\alpha = 0.05$, meaning that the H_0 is rejected and H_a is kept. Plotting the beta weights, leads to the attribute vectors which are shown in figure 9.

Table 12: outcome regression analyses attributes

	β_1	β_2	R^2	F	Sign.
Peel colour	-0.660	-0.637	0.842	18.667	0.002
Size	0.562	0.055	0.319	1.640	0.261
Glossiness	0.302	-0.697	0.577	4.779	0.049
Toughness	0.785	-0.240	0.674	7.231	0.020
Crispiness	0.585	-0.399	0.501	3.517	0.088
Firmness peel	0.679	-0.391	0.613	5.548	0.036
Firmness fruit pulp	0.623	-0.396	0.545	4.187	0.064
Juiciness	0.151	-0.428	0.206	0.906	0.447
Sourness	0.911	-0.241	0.889	28.054	0.000
Sweetness	-0.897	-0.168	0.833	17.431	0.002
Aroma	-0.197	-0.676	0.496	3.441	0.091

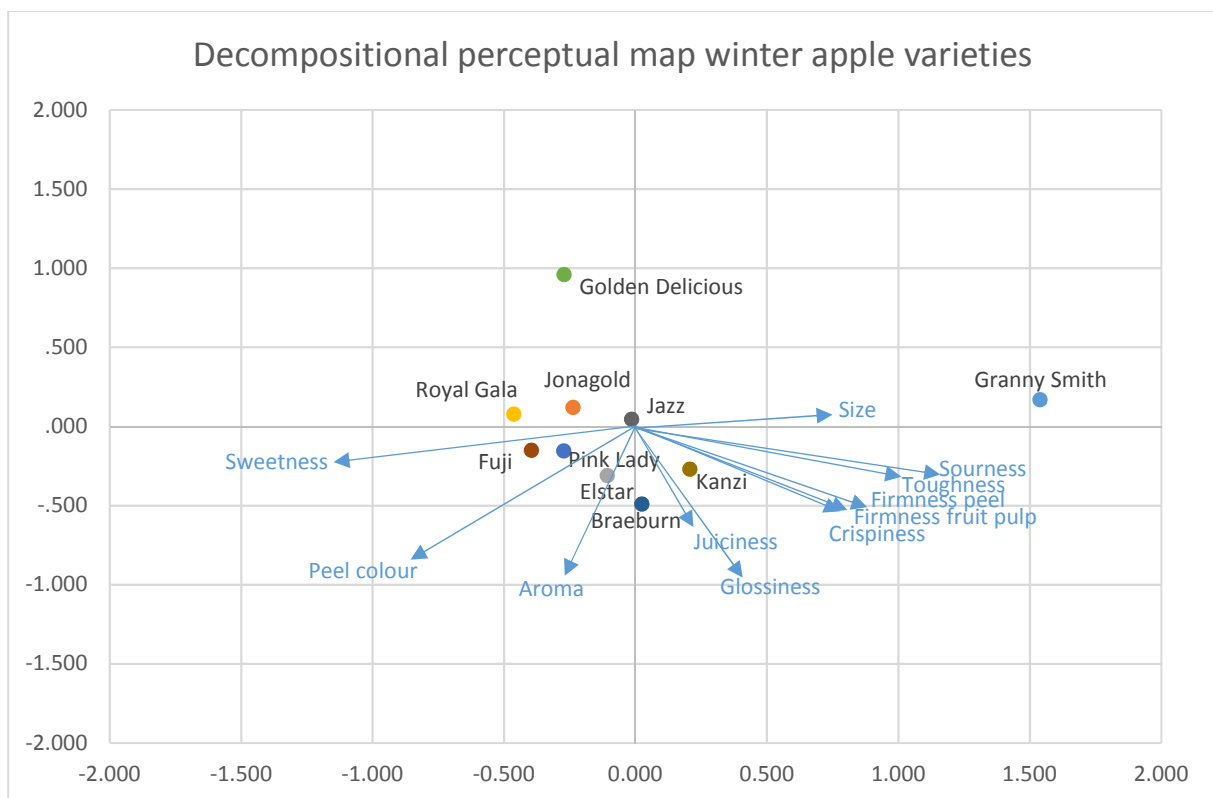


Figure 9: perceptual map including attribute vectors

The sourness vector ($R^2 = 0.889$) and sweetness vector ($R^2 = 0.833$) are the significant attribute vectors that lie closest to the x-axis, indicating that these can be used to name the x-axis. Sourness can be used to name the part of the x-axis to the right of the origin and sweetness to name the part of the x-axis to the left of the origin. The most significant vector that is found in the direction of the y-axis is peel colour, which ranges from green to red. This can be used to name the y axis, with green towards the top of the y-axis and red towards the bottom of the y-axis. This is due to the mirror image of the vector that can be made when mirroring the vector over the origin.

4.4. MDS PREFSCAL

To obtain ideal points per respondent, the preference data was handled as described in Chapter 2, method 2. When inspecting the stress per respondent, it was found that respondent 11, 21, 38, 47, 73, 84, 85, 93, 107 and 151 had individual mean stress values of 0.2 and higher up to a maximum of 0.338. According to Kruskal (1964), stress values of 0.2 and higher indicate a poor fit. The normalized raw stress for the PREFSCAL solution is 0.176. It was tried to remove above mentioned respondents from the dataset and assess the overall fit (normalized raw stress). The normalized stress became 0.154, which seems to be a marginal improvement. Taking into account that the normalized raw stress initially was lower than 0.2 and this marginal improvement that was 0.176 to 0.154, it was decided to keep all respondents in the data set, despite some having individual stress levels exceeding 0.2.

The ideal points per respondent can be found in Appendix 3, ideal_x represent the x coordinate and ideal_y the y coordinate of the ideal point. Plotting these ideal points in the perceptual map lead to the joint space perceptual map in figure 10 showing the ideal points as grey stars. The ideal points per respondent are quite evenly spread across the map, but never go beyond the Golden Delicious or the Granny Smith. Besides, at some points in the map the density seems to increase, meaning that there are multiple respondents with a comparable preference at that point in the map.

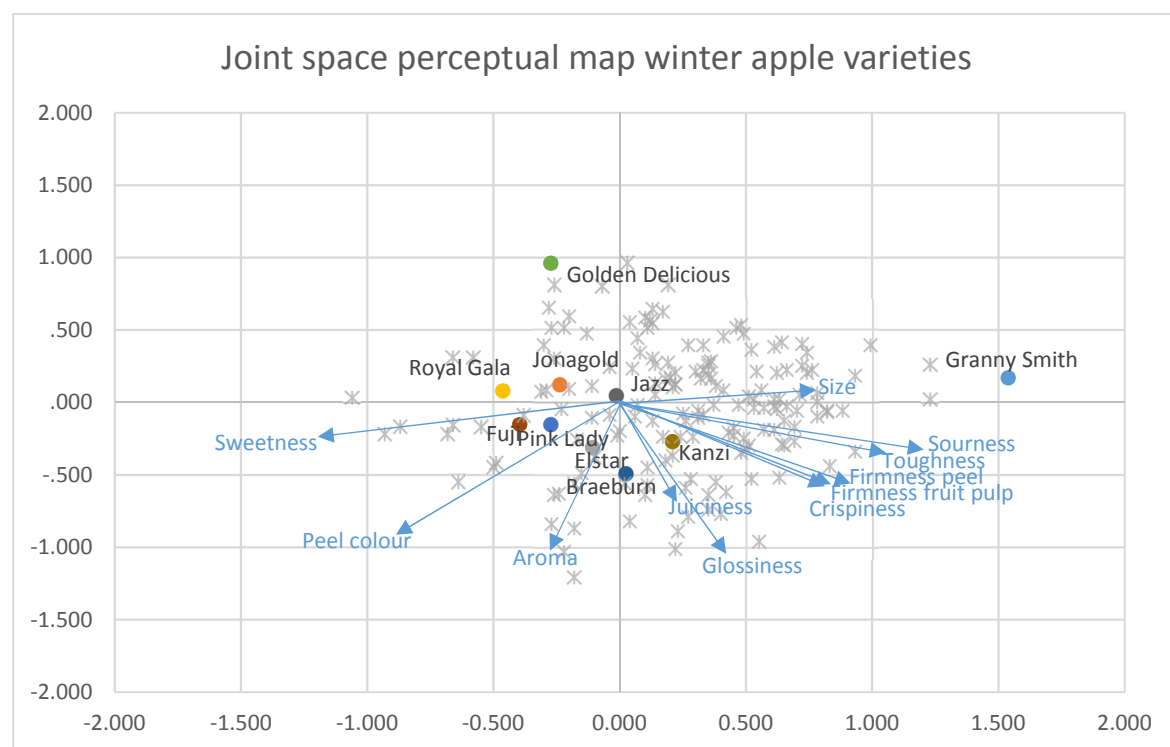


Figure 10: joint space perceptual map including ideal points per respondent

4.5. Two-stage clustering procedure

The ideal points obtained by using MDS PREFSCAL were used as input for a two-stage clustering procedure as described in Chapter 2, method 2. The respondents were assigned to the clusters as displayed in Appendix 4 and the cluster centroids are shown in table 13. In table 14 the amount of respondents assigned to cluster 1, 2, 3 and 4 are displayed.

Table 13: cluster centroids

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
X coordinate	.12	.11	.62	-.52
Y coordinate	.35	-.62	-.02	-.11

Table 14: amount of respondents assigned to cluster

	Respondents assigned to cluster (N)
Cluster 1	52
Cluster 2	32
Cluster 3	60
Cluster 4	18

Plotting the respondents in accordance to their cluster membership leads to figure 11. In figure 12 the centroids of the preference clusters are shown. The densities of the ideal points seem to be well presented in the cluster centroids of cluster 1 and 3, but to a lesser extent in the centroids of cluster 2 and 4, of which the spread seems larger as well.

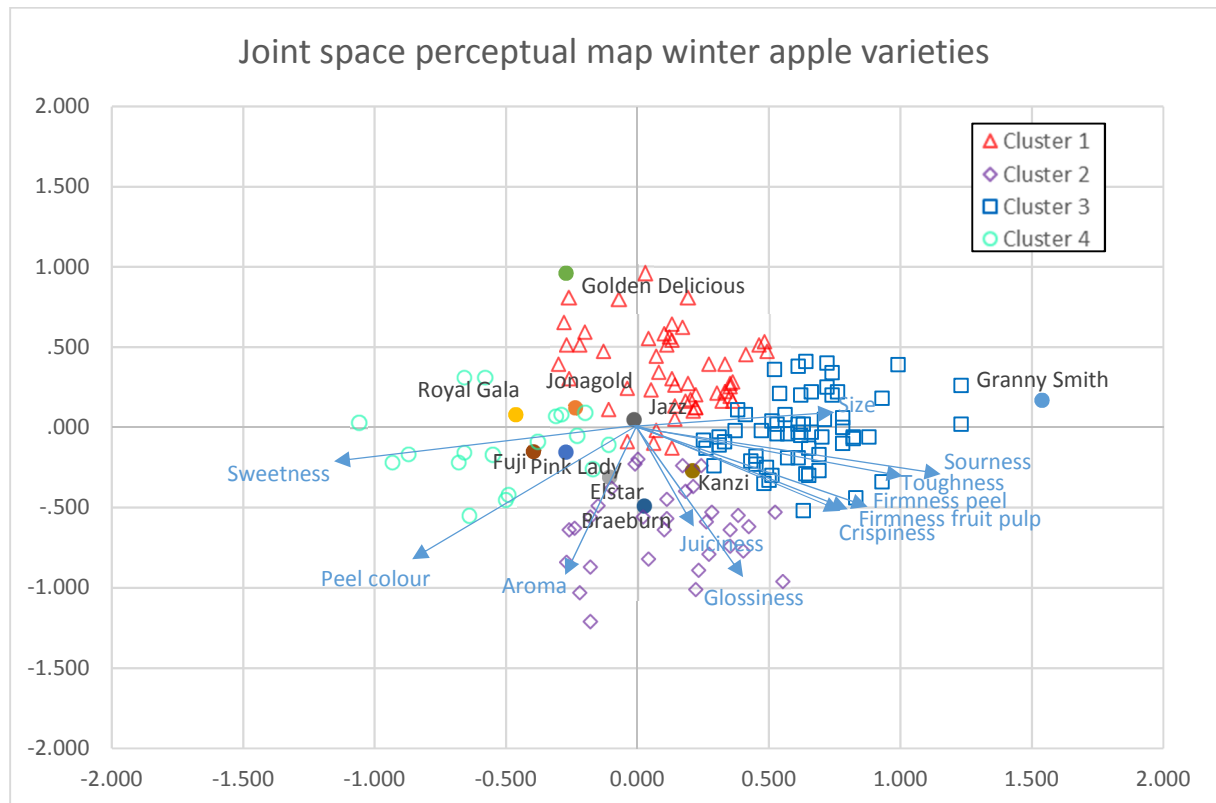


Figure 11: joint space perceptual map winter apple varieties

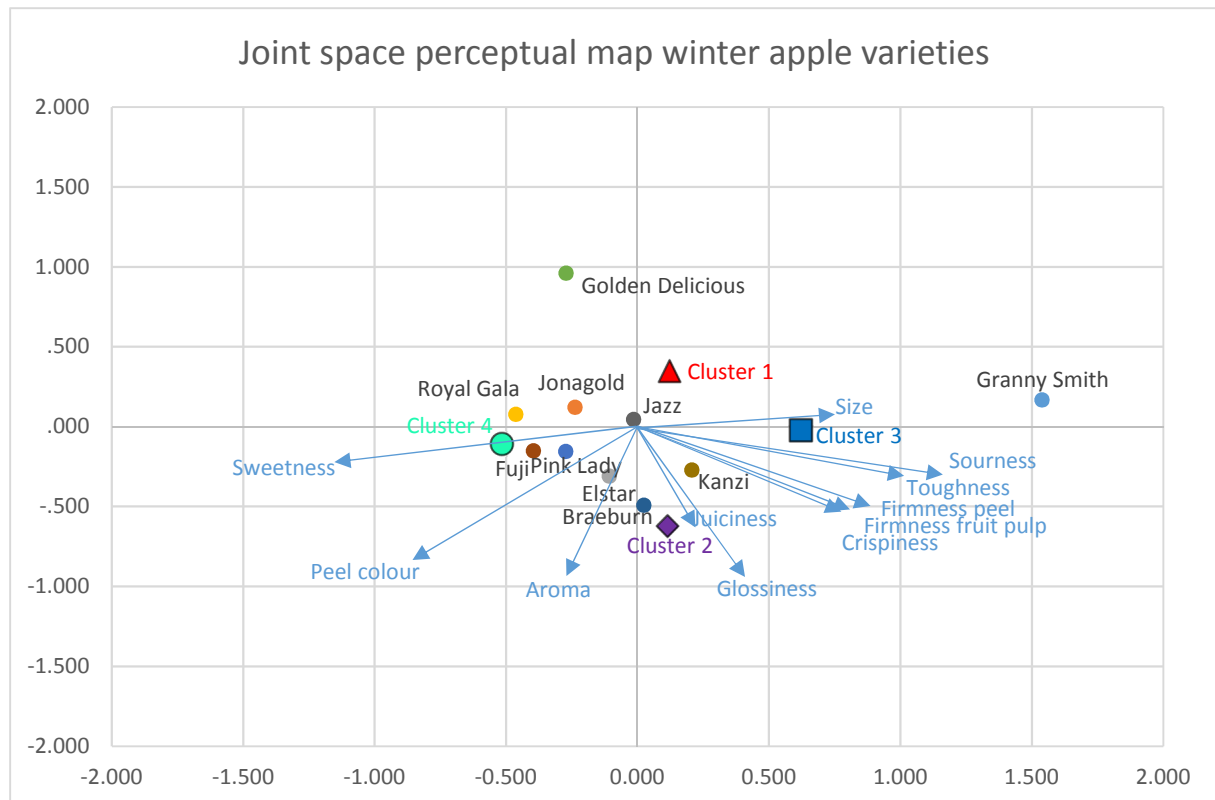


Figure 12: joint space perceptual map winter apple varieties with ideal points per cluster

When applying the segmentation criteria to the clusters found in this research, it was found that the respondents are divided quite evenly across the clusters and each cluster contains enough respondents to be substantial, though cluster 4 contains 18 respondents and in that sense is less substantial. In cluster 1 a more dense group of ideal points is visible and some other ideal points moving to the top of the map.

The cluster centroids represent the ideal apple positions per segment of which the mean attribute ratings are not known, because they cannot be measured. The positions of the ten apple varieties are known, because the values were obtained by using PROXSCAL. The mean attribute ratings of the ten apples are known, because the attributes were scored by respondents and a mean was calculated. With this data available, it is possible to run a regression analysis to obtain predicted values for the mean attribute ratings of the ideal apples.

The apple variety names were in a column, including Ideal apple 1, 2, 3 and 4. Next, the x coordinates and y coordinates of the apple positions, for the ten apple varieties and the four ideal apples are presented in two columns. Then, eleven columns containing the mean attribute ratings for the ten apple varieties are added. There is no data for the four ideal apples in these eleven columns. Then, a regression analysis is run for each attribute with the mean attribute ratings as dependent variable and the apple coordinates used as independent variables. It is asked to save the predicted values. These predicted values for the mean attribute ratings of the four ideal apples are presented in table 15.

Table 15: predicted values attribute ratings ideal apples for the four segments

	Peel colour	Size	Glossiness	Toughness	Crispiness	Firmness peel	Firmness fruit pulp	Juiciness	Sourness	Sweetness	Aroma
Ideal apple 1	43.15	64.25	54.13	71.79	64.44	67.94	61.57	65.63	44.21	55.59	54.39
Ideal apple 2	80.74	63.29	71.60	74.62	70.94	73.27	69.38	69.66	51.71	59.87	64.50
Ideal apple 3	43.80	67.16	63.52	76.24	70.36	73.31	68.98	67.68	57.43	49.39	57.22
Ideal apple 4	78.30	59.64	58.93	68.85	63.15	66.22	59.65	66.89	34.61	67.56	60.48

Values ranging from 0 to 20 are considered low, values from 21 to 40 are considered moderately low, values from 41 to 60 are considered average, values from 61 to 80 are considered moderately high and values from 81 to 100 are considered high ratings. These definitions are given for description purposes of the ideal apples.

4.6. Preference means and standard deviations

It is also possible to look at the mean preference scores for each apple variety, which were displayed in table 16. The preference scores were given on a 1 to 7 scale and a higher number indicates a higher preference.

Table 16: mean preference scores per apple variety, sorted on preference

Apple variety	Mean, std. dev.
Jazz	5.07, 1.72
Elstar	4.86, 1.74
Pink Lady	4.78, 1.72
Kanzi	4.72, 1.90
Fuji	4.64, 1.83
Braeburn	4.36, 1.78
Royal Gala	4.36, 1.71
Jonagold	3.86, 1.67
Granny Smith	3.53, 2.12
Golden Delicious	2.95, 1.57

Subsets of the respondents were made based on their cluster membership. The mean preference scores for each apple variety per cluster was calculated and it was found that between the cluster differences exist in the preference for certain apple varieties. The means and standard deviation of the means can be found in table 17. The apples that were scored highest on preference by the members in cluster 1 were the Jazz, Pink Lady and Jonagold, and the most disliked apples were the Golden Delicious, Braeburn and Granny Smith. The most preferred apples in cluster 2 were the Braeburn, Elstar and Kanzi, and the most disliked apples were Jonagold, Golden Delicious and Granny Smith. In cluster 3 the most preferred apples were Granny Smith, Kanzi and Jazz, and the most disliked apples were Royal Gala, Jonagold and Golden Delicious. The most preferred apples in cluster 4 were Fuji, Royal Gala and Elstar, and the least preferred apples are Golden Delicious, Kanzi and

Granny Smith. In the joint space perceptual map, the ideal point of cluster 1 lies close to Jonagold and Jazz, the ideal point of cluster 2 lies close to Braeburn and Kanzi, the ideal point of cluster 3 lies close to Granny Smith and Kanzi, the ideal point of cluster 4 lies closest to Fuji and Royal Gala. This indicates that the data is well presented in the joint space perceptual map.

Table 17: mean preference ratings for each cluster per apple variety, sorted from most preferred to least preferred

Cluster 1		Cluster 2		Cluster 3		Cluster 4	
	Mean, std. dev.		Mean, std. dev.		Mean, std. dev.		Mean, std. dev.
Jazz	6.00, 1.36	Braeburn	6.24, 0.66	Granny Smith	5.62, 1.02	Fuji	6.42, 1.00
Pink Lady	5.08, 1.19	Elstar	5.41, 2.00	Kanzi	5.45, 1.77	Royal Gala	6.22, 1.09
Jonagold	4.79, 1.58	Kanzi	5.09, 1.81	Jazz	4.77, 1.41	Elstar	5.73, 1.01
Royal Gala	4.67, 1.41	Fuji	4.87, 1.92	Elstar	4.50, 1.98	Pink Lady	5.50, 1.43
Elstar	4.52, 1.40	Pink Lady	4.80, 1.79	Pink Lady	4.19, 2.06	Jonagold	5.33, 1.21
Fuji	4.26, 1.61	Jazz	4.25, 1.81	Fuji	4.11, 1.83	Braeburn	4.50, 1.93
Kanzi	4.04, 1.60	Royal Gala	3.65, 2.01	Braeburn	4.00, 1.55	Jazz	4.30, 2.11
Golden Delicious	4.00, 1.31	Jonagold	3.20, 1.54	Royal Gala	3.92, 1.38	Golden Delicious	3.82, 1.89
Braeburn	3.50, 1.63	Golden Delicious	1.57, 0.76	Jonagold	3.50, 1.54	Kanzi	3.67, 2.35
Granny Smith	2.38, 1.35	Granny Smith	1.50, 0.85	Golden Delicious	2.52, 1.25	Granny Smith	1.25, 0.50

A one-way ANOVA was run to compare whether there were significant differences in preference for the ten apple varieties between the four clusters. Preference was rated on a 1 to 7 scale, with a higher number indicating a higher liking and is thus an ordinal variable. Normally distributed residuals can be used to check for the normality assumption in ANOVA, however ANOVA is robust against the normality assumption, therefore this check is left out of consideration. Another assumption of ANOVA is that equal variances are assumed. To check whether equal variances can be assumed, Levene's test was done. The H_0 of Levene's test is that equal variances can be assumed. It was found that the p-values for Jonagold (0.254, df = 3, 0.858), Golden Delicious (1.672, df = 3, 0.180), Fuji (2.175, df = 3, 0.098) and Kanzi (1.470, df = 3, 0.229) exceeded $\alpha = 0.05$, meaning that the variances of the preference scores of Jonagold, Golden Delicious, Fuji and Kanzi can be considered homogeneous between groups (equal variances can be assumed). This means that only these apples can be used in the ANOVA analysis. It was found in the ANOVA analysis that none of the p-values exceeded $\alpha = 0.05$ meaning that the H_0 is kept. This indicated that for these apples, there is a significant difference in preference between the four clusters.

For the varieties where equal variances could not be assumed, the Welch test and Brown-Forsythe test were applied. The outcome of the Welch test resulted in a p-value for Pink Lady (0.179) larger than $\alpha = 0.05$, meaning that the H_0 is kept, indicating there is no significant difference in liking between the four clusters for Pink Lady. The outcome of the Brown-Forsythe test resulted in p-values for Elstar (0.069) and Pink Lady (0.118), meaning that next to the Pink Lady, the preference for Elstar is not significantly different between groups when using Brown-Forsythe.

4.7. Cluster member characteristics

The respondents were assigned to clusters based on their ideal point coordinates and in this paragraph the cluster members for each cluster were inspected based on the demographic / lifestyle data that was gathered. Crosstabs and a Chi-square test were used to assess whether cluster membership and a certain demographic / lifestyle variable were related. The cluster membership

was put in columns and for each demographic / lifestyle variable, e.g. gender, age, consumption habit and education level, a crosstab was generated with the demographic / lifestyle variable presented in rows. The H_0 of the Chi-square test is that cluster membership is associated with gender, age, consumption habit or education level. The p-value for the Chi-square test for gender was found to be 0.485 which is larger than $\alpha = 0.05$, meaning that the H_0 could be kept. This indicates there is no association between gender and cluster membership. It was found that for age, consumption habits and education level more than 20% of the cells had an expected count less than 5, indicating that the results of the Chi-square test for these variables is not valid. Therefore Fisher's exact test was used. It was found that for the variables age and education level the procedure could not be run. The time limit was set to 5 and later to 30 minutes, it could be that the tables were too large (7 rows x 4 columns for age and 6 rows x 4 columns for education level). At the time being, the researcher found no solution for this. The p-value for consumption habit (0.417) was smaller than $\alpha = 0.05$, which means that H_0 is kept and no relation is found between cluster membership and consumption habit.

4.8. Ideal apples descriptions according to preference clusters

Based on the results presented in the previous paragraphs the four clusters and the ideal apple for each cluster can be described as follows.

Cluster 1: greener, average glossy, moderately high tough/firm apple with balanced acidity (more sweet than sour)

Cluster 1 contains 52 respondents, it is the second largest cluster, and it can be found in the upper part slightly to the right in the map. The two most preferred apples of this cluster are the Jazz (6.00, 1.36) and the Pink Lady (5.08, 1.19).

Table 18: cluster 1 - Comparison mean attribute ratings and predicted values

Attribute	Comparison based on mean attribute ratings and predicted value
Peel colour	Greener than the Jazz and Pink Lady and overall score average on peel colour. <i>Ideal apple 1 (43.25) < Jazz (76.24) < Pink Lady (77.73)</i>
Glossiness	More matte than the Pink Lady and overall score average on glossiness. <i>Ideal apple 1 (54.13) < Pink Lady (61.15) < Jazz (73.26)</i>
Toughness	Between the Pink Lady and the Jazz and overall score moderately high on toughness. <i>Pink Lady (68.89) < Ideal apple 1 (71.79) < Jazz (75.10)</i>
Firmness of the peel	Softer peel than the Pink Lady and score moderately high on firmness of the peel. <i>Ideal apple 1 (67.94) < Pink Lady (71.05) < Jazz (75.35)</i>
Sourness	Comparable to the Pink Lady and more sour than the Jazz and score overall average on sourness. <i>Jazz (36.56) < Ideal apple 1 (44.21) < Pink Lady (44.70)</i>
Sweetness	Less sweet than the Pink Lady and the Jazz and overall score average on sweetness. <i>Ideal apple 1 (55.59) < Pink Lady (60.37) < Jazz (66.36)</i>

In table 18 a comparison is made on the significant attributes between the mean attribute ratings and predicted values for the attribute ratings of the ideal apple. The ideal apple for cluster 1 would be greener, average on glossiness, moderately high on toughness/firmness and have a balanced acidity, tough more sweet than sour.

Cluster 2: red apple, with moderately high glossiness, tougher/firmer, with a balanced acidity (slightly sweeter than sour)

Cluster 2 contains 32 respondents and is positioned to the bottom of the map, slightly to the right. The two most preferred apples in this cluster were the Braeburn (6.24, 0.66) and Elstar (5.41, 2.00).

Table 19: cluster 2 - Comparison mean attribute ratings and predicted values

Attribute	Comparison based on mean attribute ratings and predicted value
Peel colour	More red than the Braeburn and Elstar and overall score moderately high on peel colour. <i>Elstar (61.37) < Braeburn (67.57) < Ideal apple 2 (80.74)</i>
Glossiness	Glossier than the Elstar (average glossy) and Braeburn (moderately high glossiness) and overall score moderately high on glossiness. <i>Elstar (56.10) < Braeburn (65.89) < Ideal apple 2 (71.6)</i>
Toughness	Somewhat tougher than the Braeburn and score overall moderately high on toughness. <i>Elstar (68.17) < Braeburn (73.93) < Ideal apple 2 (74.62)</i>
Firmness of the peel	Firmness of the peel somewhat higher than the Elstar and score moderately high on firmness of the peel. <i>Braeburn (67.21) < Elstar (71.54) < Ideal apple 2 (73.27)</i>
Sourness	More sour than the Elstar and score average on sourness. <i>Braeburn (45.21) < Elstar (46.60) < Ideal apple 2 (51.71)</i>
Sweetness	Somewhat sweeter than the Elstar and score moderately high and sweetness. <i>Braeburn (58.46) < Ideal apple 2 (59.87) < Elstar (61.74)</i>

In table 19 a comparison is made on the significant attributes between the mean attribute ratings of the Braeburn and Elstar and predicted values for the attribute ratings of the ideal apple. The ideal apple of cluster 2 can be described as a red apple, with moderately high glossiness, which is tougher/firmer, and has a balanced acidity, tough slightly sweeter than sour.

Cluster 3: Average green, moderately high glossy, quite tough/firm apple with balanced acidity (more sour than sweet)

Cluster 3 contains 60 respondents and is the largest cluster that was found. The cluster is found in the middle of the right side of the map. The two most preferred apples in this cluster were the Granny Smith (5.62, 1.02) and Kanzi (5.45, 1.77).

Table 20: cluster 3 - Comparison mean attribute ratings and predicted values

Attribute	Comparison based on mean attribute ratings and predicted value
Peel colour	Less green than the Granny Smith and less red than the Kanzi, overall average score on peel colour <i>Granny Smith (11.14) < Ideal apple 3 (43.80) < Kanzi (66.84)</i>
Glossiness	Glossier than the Kanzi and less glossy than the Granny Smith, overall moderately high score on glossiness <i>Kanzi (60.83) < Ideal apple 3 (63.52) < Granny Smith (66.27)</i>
Toughness	Tougher than the Kanzi, overall score moderately high on toughness <i>Kanzi (71.90) < Ideal apple 3 (76.24) < Granny Smith (82.49)</i>
Firmness of the peel	Firmer peel than the Kanzi, not as firm as the Granny Smith, overalls score moderately high on firmness of the peel <i>Kanzi (68.37) < Ideal apple 3 (73.31) < Granny Smith (79.02)</i>
Sourness	Comparable sourness to the Kanzi, overall score average on sourness <i>Ideal apple 3 (57.43) < Kanzi (58.36) < Granny Smith (73.93)</i>
Sweetness	Sweeter than the Granny Smith, but less sweet than the Kanzi, overall score average on sweetness <i>Granny Smith (34.75) Ideal apple 3 (49.39) < Kanzi (52.59)</i>

In table 20 a comparison is made on the significant attributes between the mean attribute ratings of the Granny Smith and Kanzi and predicted values for the attribute ratings of the ideal apple of cluster 3. The ideal apple of cluster 3 can be described as an apple that is average green, moderately high on glossiness, quite tough/firm and with balanced acidity, though more sour than sweet. On most

attributes the ideal apple scores in between the Granny Smith and Kanzi apple, and that is also the position it takes in the joint space perceptual map.

Cluster 4: red, average glossy, moderately high tough/firm apple that is more sweet than sour

Cluster 4 contains 18 respondents and is the smallest cluster that was found. The cluster is positioned in the middle of the left side of the map. The two most preferred apples in this cluster were the Fuji (6.42, 1.00) and the Royal Gala (6.22, 1.09).

Table 21: cluster 4 - Comparison mean attribute ratings and predicted values

Attribute	Comparison based on mean attribute ratings and predicted value
Peel colour	More red than the Fuji and score moderately high on peel colour <i>Royal Gala (72.52) < Fuji (76.10) < Ideal apple 4 (78.3)</i>
Glossiness	Comparable to glossiness of Royal Gala, overall score average on glossiness <i>Royal Gala (58.75) < Ideal apple 4 (58.93) < Fuji (66.58)</i>
Toughness	Somewhat tougher than Royal Gala, less tough than Fuji, overall score moderately high on toughness <i>Royal Gala (67.36) < Ideal apple 4 (68.85) < Fuji (75.35)</i>
Firmness of the peel	Firmer peel than Fuji, slightly less than Royal Gala, overall score moderately high on firmness of the peel <i>Fuji (64.96) < Ideal apple 4 (66.22) < Royal Gala (67.51)</i>
Sourness	Comparable to the sourness of Royal Gala, overall score moderately low on sourness <i>Fuji (32.93) < Royal Gala (34.07) < Ideal apple 4 (34.61)</i>
Sweetness	Sweetness in between the Royal Gala and Fuji, overall score moderately high on sweetness <i>Royal Gala (62.49) < Ideal apple 4 (67.56) < Fuji (72.09)</i>

In table 21 a comparison is made on the significant attributes between the mean attribute ratings of the Fuji and Royal Gala and predicted values for the attribute ratings of the ideal apple of cluster 4. The ideal apple of cluster 4 can be described as a red apple, that scores average on glossiness, is quite tough/firm and more sweet than sour.

Another way to look at the data presented above, is by looking at the positions of the ideal apples in the perceptual map. Each attribute vector can be mirrored in the origin, to obtain a mirror image of the vector. For example, in figure 13 it can be seen that the arrow for peel colour that points towards the left bottom quadrant, represents a red colour, because the highest scores represent a red colour. When plotting a mirror image of this vector, it points towards to top right quadrant of the map, which represents a green colour, because the lower scores were used to indicate green.

This image can be used to identify attributes that a certain ideal apple would possess the most, in a visual way. The attribute vector is mirrored in the origin and then the apples need to be imagined as if they are moved over a perpendicular line towards the vector (either the actual vector or the mirror image of the vector) until the apples actually lie on the vector. For example, when imagining the ideal apples on the vector sweetness (both the actual vector and the mirror vector), it can be seen that in the direction moving from sweet towards not sweet, the ideal apple 4 is on the vector first, followed by 2, 1 and 3, with ideal apple 3 then being the most 'not sweet'. This is also in accordance to the predicted values that were found for sweetness for the four ideal apples.

In other words, figure 13 can also be used to describe the ideal apples, regardless of the existing brands, in contrast to what was done in table 18 to table 21.

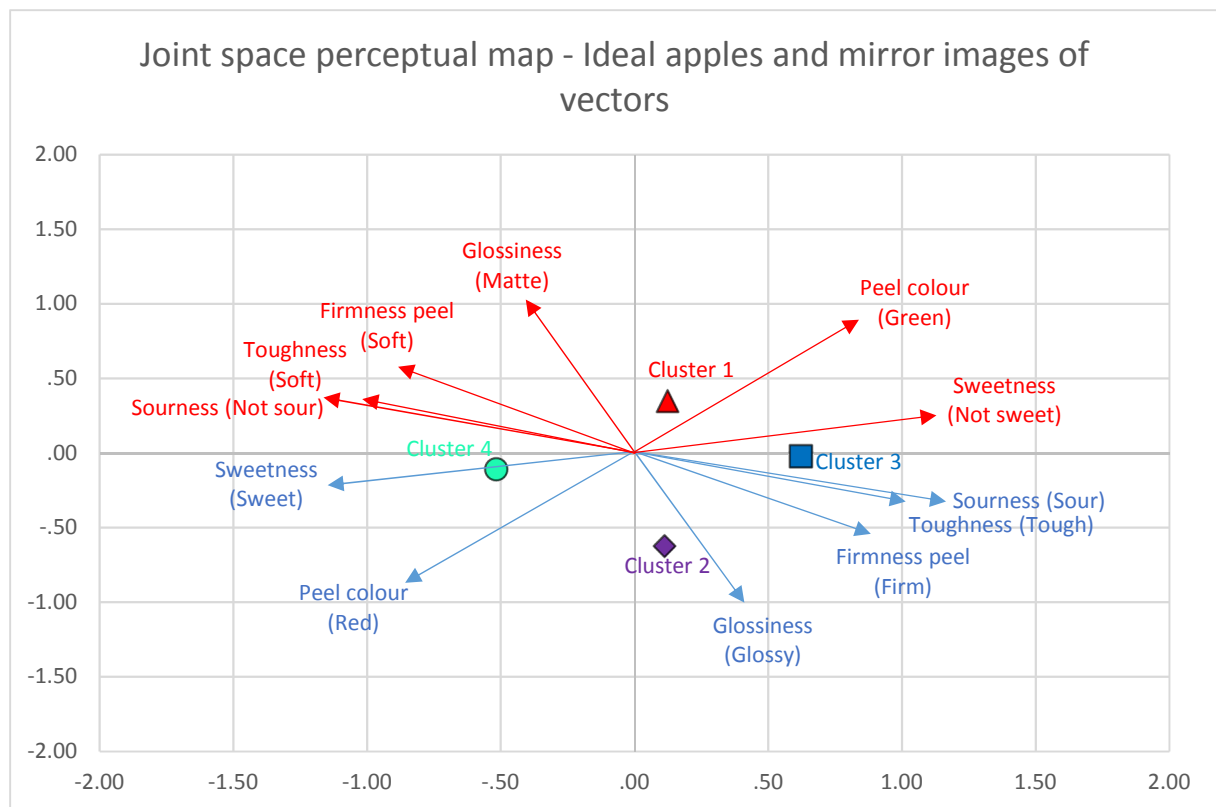


Figure 13: joint space perceptual map with four ideal apples per cluster and mirror images of vectors included

Chapter 5: Conclusion

This research has aimed to answer the following main research question:

‘What is the position of winter apples based on consumer perceptions of these apples and what consumer preference clusters can be distinguished?’

SRQ 1: *What segmentation methods lead to preference clusters of the summer apples data that meet the criteria for segmentation?*

In this research two segmentation methods have been tested and applied to summer apples data. In the first place ideal points were generated by means of multiple linear regressions for each respondent with the preference scores on apple varieties and apple position coordinates obtained from multidimensional scaling as predictors. It was found that this method lead to outliers and this is most probably caused by the fact that individual linear relations are estimated, not taking into account the entire set of respondents. Besides, the fact that missing data was present lead to a small number of degrees of freedom left for estimation and a large error around the obtained ideal points is expected. The alternative method, applying multidimensional unfolding, led to clusters that were more substantial compared to the first method and in literature it was found the method could handle missing data. No outliers were identified and the obtained results were well interpretable, met de posed segmentation criteria and were not expected to be degenerate. Applying MDS PREFSCAL and additional two-stage clustering thus leads to preference clusters.

SRQ 2: *What is the position of existing winter apple varieties based on consumer perceptions of these apples?*

The position of the existing winter apple varieties were found by using the dissimilarity ratings and MDS PROXSCAL. It was found that Granny Smith and Golden Delicious were very dissimilar to the majority of the apples. Due to attribute vector fitting it could be seen that the Golden Delicious is considered as an apple that lacks attributes that consumers like in apples and there is consensus amongst respondents about the disliking of this apple variety. The Granny Smith is considered dissimilar because it is very green, tough, firm and acidic. There is disagreement amongst consumers when it comes to liking of this apple, some like the apple very much and others not at all. The other eight apple varieties are positioned closer to one another and this indicates that they are considered more similar.

SRQ 3: *What preference clusters can be distinguished based on the winter apples data?*

Four preference clusters were found in this research by applying multidimensional unfolding via PREFSCAL and additional two-stage clustering. The segmentation base was preference and significant differences between the clusters with regards to gender and consumption habit were not found. The clusters can be described as follows. The ideal apple for cluster 1 would be greener, average on glossiness, moderately high on toughness/firmness and have a balanced acidity, tough more sweet than sour. The ideal apple of cluster 2 can be described as a red apple, with moderately high glossiness, which is tougher/firmer, and has a balanced acidity, tough slightly sweeter than sour. The ideal apple of cluster 3 can be described as an apple that is average green, moderately high on glossiness, quite tough/firm and with balanced acidity, though more sour than sweet. The ideal apple of cluster 4 can be described as a red apple, that scores average on glossiness, is quite tough/firm and more sweet than sour.

The results indicated that the respondents in cluster 4 are served already, because the Fuji is positioned very close to the centroid of that cluster. This is also the case for cluster 2, which is

positioned very close to the Braeburn apple. The members of cluster 1 would most likely choose for the Jazz apple, though here is some room to develop a new variety. This would then be a greener, apple, that is averagely glossy, moderately high on toughness/firmness and has a balanced acidity, tough is more sweet than sour.

The most room for development can be found in the direction of cluster three. No apple is positioned really close to the centroid of cluster three, indicating the chance that competition is smaller when a new apple would be positioned there. This apple would be an apple that is average green, scores moderately high on glossiness, quite tough/firm and with a balanced acidity, though more sour than sweet. This apple would be in between the Kanzi and the Granny Smith apple. Cluster 1 and 3 are also the largest clusters found in this research, so developing apples that would suit the needs of these segments, would also likely to be profitable.

Summarizing the answers on the three sub questions answers the main research question '*What is the position of winter apples based on consumer perceptions of these apples and what consumer preference clusters can be distinguished?*' This research has sought to give information about the perception of current apple varieties, the segments that exist and to obtain a joint space perceptual map showing where new apple varieties could be positioned and what attributes would be preferred. This was done so that consumer preferences for apples become more incorporated in the selection of new apple varieties. The Granny Smith and Golden Delicious take in a very different positions than the other apples. The remaining apples are considered to be more similar. In this research four preference clusters are found that are characterized by the preference based on peel colour, glossiness, toughness, firmness of the peel, sourness and sweetness. Based on the findings in this research and the perceptual map obtained, the best place to position a new apple variety would be at the centroids of cluster one and three. When looking at the perceptual map as a whole, there is more room for differentiation, since now the majority of apples is found close to one another, though at the borders of the segments there is more room for positioning new varieties. The ideal apple positions are determined based on the preference for existing apples, but it could be that developing new varieties with more distinct characteristics, will lead to increased choice for consumers.

Chapter 6: Discussion

6.1. Discussion of the results

6.1.1. Current findings and comparison with findings from Van den Broek (2019)

In this research a joint space perceptual map containing ten winter apple positions, eleven attribute vectors and four preference clusters with centroids was found. Comparing these results to what was found in previous research (Van den Broek, 2019), shows that the position of some apples in the currently found perceptual map take in similar positions as in the previously found perceptual map. This is the case for the apple varieties Golden Delicious, Granny Smith and Fuji. Also the direction of the majority of the attribute vectors found in previous research compared to current research is similar, this is the case for all vectors except for aroma and size.

These differences in apple positions for previous and current research can be explained by the following: in previous research another set of apples was presented, leading to different subsets presented to respondents. Consumers tend to judge apples relatively to one another and make comparisons between them, leading to differing positions. Another explanation is that the quality of apples varies throughout the year, because it is a seasonal product, leading to differences in attributes. In that sense, it is expected that when doing this research repeatedly in time, the apple positions will differ across the map over time.

What draws attention however is that the position of some apples barely changed. Regardless of the time of year, the quality and the set presented to respondents, there seems to be consensus about the position of these apples. This is also confirmed by the fact that the direction of the vectors have barely changed. The fact that the vector for aroma differs can be explained by consumers not finding aroma a clear attribute. Despite the added definition of aroma, in the conduction of the sensory evaluations it became clear that ‘aroma’ was still not clear. Therefore it is proposed to remove aroma as attribute in future research. The direction of the attribute vector for size can differ, because the size of the apples is dependent on the harvest and can vary a lot across seasons. Other attributes vary as well, but it could be that size varies so much that there is a noticeable difference for consumers.

6.1.2. Comparison existing literature on consumer preference segments with segments found

Four segments were found that can be divided based on preference for colour, acidity, toughness/firmness and glossiness. This is in accordance to the findings of Dailliant-Spinnler and colleagues (1996), who found that consumer segmentation could be achieved based on acidity, hardness and juiciness and the findings of Bonany and colleagues (2014) who have found that segmentation could be achieved based on sweetness, juiciness, crispiness and firmness. In this research crispiness and juiciness were not found to be significant attributes, and colour and glossiness were found to be significant attributes.

Other drivers for consumer behaviour with regards to buying and consumption of apples were mentioned by respondents as well and considering these could be important. These two important drivers are price and goal of use. These will be discussed in the recommendations, because it is not an official outcome of this research.

6.2. Discussion on the data gathering process

6.2.1. Materials used

The apples that were used were bought at once and come from the same batch. They were stored during the data gathering in coolers at Wageningen Research Open Crops to maintain the quality. However, during the second week some apples seem to have degraded in quality and when cutting

some apples they seemed to turn brown quicker than in the previous week. Apples are a fresh product and thus degradation of quality is inevitable.

6.2.2. The questionnaire

In setting up the questionnaire a mistake was made in the preference rating scale, an extra 4 was included by accident leading to an 8 point rating scale from 1 to 7 instead of a 7 point rating scale. The third day of the data gathering this was corrected in Qualtrics. The answers that were given up until that moment were overwritten by Qualtrics and corrected to the adjusted scale. An independent sample t-test was done between the first 33 responses that were gathered in the days the erroneous scale was used and the rest of the responses where the right scale was used. There was no statistically significant differences between the responses given on the erroneous scale and the right scale and thus it was decided to keep the preference data of the 33 responses despite the mistake that was made in the scale.

In the questionnaire the scales were centred and when respondents wanted to answer 50 out of 100, they would have to click on the 50, otherwise the questionnaire did not continue to the next page. What sometimes happened is that respondents dragged the arrow slightly more to the right, compared to staying exactly in the middle, to make sure they answered the question. Also with regards to how the questionnaire is set up: the response on the question on how many apples are consumed must be limited to numerical input only, because now sometimes responded answered '2 kg a month' and the questionnaire would allow this.

6.2.3. Locations and external validity

The data was gathered in ten different supermarkets within 35 km's from Wageningen, predominantly in villages and also within the 'Betuwe', which is a region known for its connection with fruits and vegetables. This could impact the external validity of the research with regards to generalizing the findings to the entire Dutch population. A solution for this could be to conduct the research in multiple provinces of the Netherlands and also include larger cities. Besides, more respondents could be recruited.

In every supermarket the lay out was different, and the store lighting and shadow casted on the apples was somewhat different. This could have influenced the ratings of the search attributes, because these are judged visually. A supermarket is not comparable to a sensory booth in which light, temperature, etc. can be regulated.

6.2.4. Participant recruitment

In the process of participant recruitment it was effective to put the focus on 'can you help us with our research', consumers seemed to be more willing to participate. A downfall of this is that participants are recruited that do not have time, but are afraid to say no, which could impact the answers given by these respondents. Besides, participants are approached while shopping for groceries, which is for them a habitual process in which they are guided by cues. Participating in a research like this forces participants into a more rational mode, which could have impacted the judgements given. Participants that take part in the research in the morning, can differ from participants that take part during lunchtime, participants that are hungry might give other answers than participants that are satiated. These are factors that can impact sensory judgements, but these are out of control of the researchers. Because the reward offered in this research was quite high, it is possible participants only participated to have a chance to win the reward.

6.2.5. Filling in the questionnaire

At the beginning of the questionnaire, respondents would look at the apples and try to guess which apples they needed to judge. This indicates that variety name / brand name is important and the look of certain apples evokes the name, especially in the case of Granny Smith. It is also expected that the look of the apple gave consumers a predefined idea about the apple, for example using the green colour as a cue for acidity, leading to green apples being judged as more acidic than they are in reality. This has to do with confirmation bias. People are looking for confirmation of what they already know. Some respondents were focussed on liking and preference and expressed this in the beginning of the questionnaire, for example at the first question on peel colour: 'is this about how much I like the colour?'. Where necessary the researchers corrected this by explaining it was about to what extent a certain attribute was perceived.

Older participants had difficulties working with the laptop and mouse. This sometimes lead to elderly taking a long time participating and answering all the questions. Participants (male/female) in the presence of young children sometimes participated, but the young children would distract them by talking and/or crying. This could have led to respondents with a parent role to fill in the questionnaire more hurried so they could give attention to their child.

During the questionnaire, some respondents showed sensory fatigue. When respondents arrived at the pairwise comparison they would express 'even more questions?', 'how long will this take?', 'how many questions after this?'. Also respondents said that at a certain point they experienced difficulties in tasting the difference, because they tasted too much. Crackers and water were offered for neutralization, however respondents could not be forced to use this, because this would give a compelling character to their voluntary participation. This leads to the impression that 5 apples could be too much for one respondent to judge in the setting that was chosen for. People in supermarkets have limited time and willingness to spend time, thus in future research it is advised to reduce the amount of apples, shorten the questionnaire or ask other questions to obtain other types of data that lead to apple positions in a dimensional space. Reducing the amount of apples for each respondent would lead to more missing data. Reducing the amount of apples that are assessed in the research is also an option, for example by including only the most sold apples, because a new apple variety would preferably be positioned in the map in a competitive place relative to other well sold apples. Another option is to exclude the apples of which the position in the map stays constant over time, as mentioned before.

During the part on attribute ratings, respondents were instructed to answer all attributes for one apple and then continue to the next apple. It could be an idea for the next research to ask all attributes per apple, instead of asking to rate all apples per attribute. Otherwise the respondents need to keep scrolling back up in the questionnaire. Letting respondents taste five apples per attribute leads to many tasting moments and this takes too much of their time and leads to clouded responses. Another point of attention with regards to the attribute ratings is that biting in a whole apple is another experience than apple slices. Some respondents explained they would peel the apple prior to consumption. The attribute aroma was, despite the definition given, still unclear to respondents. Often it was confused with liking of the apple, some respondents tried to smell the apple scent. With regards to the question on colour, respondents were not sure how to judge an apple that was bi-coloured (yellow-red), it is proposed to add 'yellow' to the middle of the scale at the question on peel colour. At the end of the questionnaire the researchers let respondents guess what apples they tasted, the respondents enjoyed guessing the apples, because it involved them more personally. Also they liked to talk about their favourite apple. Many respondents mentioned the Elstar as the apple they always buy and some respondent mentioned that they cared about the

price of apples the most. Some respondents mentioned purpose of use (apple pie, apple for kids to take to school), origin and whether the apples were locally produced or not and some varieties that were not included were mentioned, namely Cox and Welland.

Despite the time respondents needed to spend and the sensory fatigue that sometimes was visible, it is believed that the data that was gathered in this research is of high quality and also reliable. One respondent was handled at the time, meaning that the researchers could intervene and help where necessary. Also, because the respondents might have felt under observation, they took the job seriously.

6.3. Discussion on the methodology, design and data analyses

In the research design an incomplete design is used, this means that not all apples were judged by all respondents, leading to missing data. This also leads to subsets of apples being presented to respondents. An implication of this is that the respondents judge the apple relatively to the other four in the subset of apples, not taking into account the other apples. It could be the case that a respondent has a favourite apple, which is not included in the subset, leading to a distorted image of the preference. For example, a consumer might have to judge five red apples, while their favourite apple is the Granny Smith. However, it is also expected that the larger the sample, the less the outcome will be affected by the incomplete design. Another implication of the design is that the Granny Smith was apple number 1, and in the design often the first apple respondents needed to rate. It is possible that the apples tasted after Granny Smith are judged as sweeter and less acidic than they are in reality, despite neutralizing with water and / or cracker. To prevent the effect of presenting an acidic, hard apple on the judgement of the remaining apples, it is proposed to mirror the design half way, meaning that where apple 1 would be judged as first apple in the first half of the design, in the second half it would be the last apple.

A remark that needs to be made with regards to the data analysis, it that there are better methods available for analysing these types of data, but applying these was not within the capabilities of the researcher based on the level of education experienced. For example, clusterwise regression by working with the R software package Flexmix, fuzzy clustering and working with probabilities of cluster membership were found to be better methods, according to literature. These will be discussed in the recommendations. In this research a lower triangle dissimilarity matrix with averaged proximities per apple pair was used as input for the PROXSCAL procedure. However, another way of using PROXSCAL was tried as well, this is discussed in Appendix 5.

Chapter 7: Recommendations

In the discussion several limitations and weaknesses of this research were identified. Therefore the following changes and alternative methods are proposed to improve the research.

It is proposed to leave out unclear attributes (e.g. aroma), or attributes that seem to mean the same, because they are close to one another (firmness peel, firmness fruit pulp, toughness), this also reduces the amount of sensory evaluations for the respondents.

Distinguishing the difference between the more similar apple varieties might be a difficult task for respondents in general. It is proposed to let a set of respondents do a free sorting task in which only the red apples, so to speak, are included to investigate whether respondents are well able to distinguish between the varieties and based on which attributes. According to Chollet, Valentin and Abdi (2014) data from these type of experiments can also be analysed using MDS. When it is found that respondents are able to distinguish the differences, it is proposed to continue with the current method of dissimilarity judgements in pairs, however with a reduced amount of apples to be judged by one respondent. If respondents are not that well able to distinguish differences, other methods to obtain dissimilarity judgements should be investigated.

With regards to clustering respondents based on their preference, also other methodology and software programs exist. It is proposed to recruit a researcher that has more knowledge on sensometrics and is familiar with the software program R. The method that could be used then is clusterwise regression. Clusterwise regression uses regression and cluster analysis simultaneously. According to Wedel (1990) clusterwise regression can overcome the short-comings of the two-stage approach as applied in this research. It was shown that with regards to clustering the preferences of consumers correctly based on an artificial data set, using clusterwise regression lead to 47.9% of preferences clustered correctly compared to 43.4% in the two-stage approach (Wedel & Kistemaker, 1989), which is an improvement. In R a package is available, namely Flexmix (Leisch, 2004), which allows for clusterwise regression.

What is also interesting is that when inspecting the cluster members based on demographic and lifestyle data there is no significant difference between the clusters with regards to gender and consumption habit, even though in literature it is found that clusters can be found in the preference of apples based on these segmentation bases. It could be investigated to group the respondents based on their demographics/lifestyle characteristics and see whether significant differences in preferences are found. Also a method needs to be found to figure out significant differences between groups with regards to age and education level, this was not found in this research.

Price, origin, willingness to pay, and the way apples are packaged are outside the scope of this research project, but it is recommended to investigate these as well and this can be achieved by doing an analysis of trade-offs: a Conjoint analysis (Jaeger, Hedderley & MacFie, 2001). As some respondents said: 'I really like the Kanzi apple, but for that money I can buy twice as much Elstar'. Next to price, also purpose of use was mentioned by respondents as drivers in the decision making process for a certain apple variety. Also, respondents mentioned brands that were not included in this research like Cox and Welland. These could be interesting to include in next researches.

References

- Aldenderfer, Mark S., and Roger K. Blashfield (1984), *Cluster Analysis*, Newbury Park, CA: Sage, p.7.
- Bonany, J., Brugger, C., Buehler, A., Carbó, J., Codarin, S., Donati, F., ... & Höller, I. (2014). Preference mapping of apple varieties in Europe. *Food Quality and Preference*, 32, 317-329.
- Bowen, A. J., Blake, A., Tureček, J., & Amyotte, B. (2019). External preference mapping: A guide for a consumer-driven approach to apple breeding. *Journal of sensory studies*, 34(1), e12472.
- Buja, A., Swayne, D. F., Littman, M. L., Dean, N., Hofmann, H., & Chen, L. (2008). Data visualization with multidimensional scaling. *Journal of Computational and Graphical Statistics*, 17(2), 444-472.
- Busing, F. M. T. A., Commandeur, J. J., Heiser, W. J., Bandilla, W., & Faulbaum, F. (1997). PROXSCAL: A multidimensional scaling program for individual differences scaling with constraints. *Softstat*, 97, 67-74.
- Busing, F. M. T. A. (2010). *Advances in multidimensional unfolding*. Doctoral thesis, Leiden University.
- CBI. (2019). *Exporting fresh fruit and vegetables to the Netherlands*. Den Haag: Ministry of Foreign Affairs. Retrieved from <https://www.cbi.eu/node/2835/pdf/>, accessed on 8-9-2019
- CBS Statline. (2019). *Bevolking; geslacht, leeftijd en burgerlijke staat, 1 januari*. Retrieved from <https://opendata.cbs.nl/statline/#/CBS/nl/dataset/7461BEV/table?fromstatweb>, accessed on 5-3-2020
- Charrad, M., Ghazzali, N., Boiteau, V., & Niknafs, A. (2015). "NbClust: An R Package for Determining the Best Number of Clusters in a Data Set." *Journal of Statistical Software* 61: 1–36.
- Chollet, S., Valentin, D., & Abdi, H. (2014). Free sorting task. *Novel Techniques in Sensory Characterization and Consumer Profiling*. Valera P. & Ares, G. (Eds.). Boca Raton: Taylor and Francis, 207-227.
- Daillant-Spinnler, B., MacFie, H. J. H., Beyts, P. K., & Hedderley, D. (1996). Relationships between perceived sensory properties and major preference directions of 12 varieties of apples from the southern hemisphere. *Food quality and preference*, 7(2), 113-126.
- Dijksterhuis, G. (2016). New product failure: Five potential sources discussed. *Trends in food science & technology*, 50, 243-248.
- Dolman, M.A., G.D. Jukema, P. Ramaekers (EDS.), 2019. *De Nederlandse landbouwexport in 2018 in breder perspectief*. Wageningen, Wageningen Economic Research, Rapport 2019-001
- Dolnicar, S., Grün, B., & Leisch, F. (2018). Step 2: Specifying the Ideal Target Segment. In *Market Segmentation Analysis* (pp. 31-37). Springer, Singapore
- Harker, F. R., Gunson, F. A., & Jaeger, S. R. (2003). The case for fruit quality: an interpretive review of consumer attitudes, and preferences for apples. *Postharvest biology and technology*, 28(3), 333-347.
- Iacobucci, D. (2013). *Marketing models: multivariate statistics and marketing analytics*. Mason, OH: South-Western.
- IBM. (2019). *Multidimensional scaling: Choosing PROXSCAL*. Retrieved from https://www.ibm.com/support/knowledgecenter/en/SSLVMB_23.0.0/spss/categories/choosing_proxscal.html, accessed on 4-11-2019

- Jaeger, S. R., Wakeling, I. N., & MacFie, H. J. (2000). Behavioural extensions to preference mapping: the role of synthesis. *Food Quality and Preference*, 11(4), 349-359.
- Jaeger, S. R., Hedderley, D., & MacFie, H. J. (2001). Methodological issues in conjoint analysis: a case study. *European Journal of Marketing*.
- Kalaitzis, P. & Van Dijk, G. (2005), "*Supply chain analysis of the fruit and vegetables market in The Netherlands*", MEDFROL Project no: SSPE-CT-2004-502459 (STREP). Retrieved from <http://medfrol.maich.gr/documentation/view/reports/wp2-deliverables/Supply%20chain%20analysis%20of%20the%20fruit%20and%20vegetable%20market%20in%20The%20EU> , accessed on 8-9-2019
- Kaufman, L., & Rousseeuw, P. J. (2009). *Finding groups in data: an introduction to cluster analysis* (2009 ed., Vol. 344). John Wiley & Sons. pp. 4-23
- Kelley, K., Hyde, J., Travis, J., & Crassweller, R. (2010). Assessing consumer preferences of scab-resistant apples: a sensory evaluation. *HortTechnology*, 20(5), 885-891.
- Ketchen, D. J., & Shook, C. L. (1996). The application of cluster analysis in strategic management research: an analysis and critique. *Strategic management journal*, 17(6), 441-458.
- Kotler, P. (1988), *Marketing Management*, Prentice-Hall, Englewood Cliffs, New Jersey.
- Kotler, P. T., & Keller, K. L. (2012) *Marketing Management*, 14th Edition. Editorial Pearson. pp. 252-255
- Kruskal, J. B. (1964). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, 29(1), 1-27.
- Leisch, F. (2004). Flexmix: A general framework for finite mixture models and latent glass regression in R.
- Mair, P., Borg, I., & Rusch, T. (2016). Goodness-of-fit assessment in multidimensional scaling and unfolding. *Multivariate behavioral research*, 51(6), 772-789.
- Martin, G. (2011). The importance of marketing segmentation. *American Journal of Business Education (AJBE)*, 4(6), 15-18.
- Nieuwe Oogst. (2018). Minder arbeidsmigranten werkzaam in tuinbouw. Retrieved from <https://www.nieuweoogst.nl/nieuws/2018/12/11/minder-arbeidsmigranten-werkzaam-in-tuinbouw>, accessed 5-3-2020
- Onderzoeksvoorstel T&U. (2018). *Duurzame Fruitintroducties Beter Begrepen*. Wageningen: WR OT, Topsector T&U. Received by e-mail from dr. ir. F.J.H.M. Verhees on 3-7-2019
- Porter, M. E. (2008). The five competitive forces that shape strategy. *Harvard business review*, 86(1), 25-40
- Retailtrends. (2019). Retrieved from <https://retailtrends.nl/news/58403/aantal-winkels-jumbo-volgend-jaar-boven-de-700>, accessed on 22-11-2019
- Statista. (2019). Total apple production in the Netherlands from 2015 to 2019, by type (in 1,000 tons). Retrieved from <https://www.statista.com/statistics/659973/apple-production-in-the-netherlands-by-type/> , accessed on 13-12-2019

Van den Broek, S. J. (2019). Sustainable Food Introductions Better Understood; Positioning of new apple varieties in relation to consumer preferences and compared to current apple varieties. Wageningen Research.

Van Kleef, E., Van Trijp, H. C., & Luning, P. (2005). Consumer research in the early stages of new product development: a critical review of methods and techniques. *Food quality and preference*, 16(3), 181-201.

Van Kleef, E. (2006). *Consumer research in the early stages of new product development: issues and applications in the food domain*.

Volkskrant. (2019). Agrariërs worstelen om Poolse werknemers vast te houden: 'Beter loon is niet langer genoeg'. Retrieved from <https://www.volkskrant.nl/nieuws-achtergrond/agrariers-worstelen-om-poolse-werknemers-vast-te-houden-beter-loon-is-niet-langer-geenog~bb0beba2/>, accessed on 5-3-2020

Wedel, M. (1990). *Clusterwise regression and market segmentation: developments and applications* (Doctoral dissertation, Wedel).

Wedel, M. and Kamakura, W. A. (2000) *Market Segmentation: Conceptual and Methodological Foundations* (2nd Ed.) Kluwer Academic Publishers, Boston, Dordrecht, London.

Wedel, M., & Kistemaker, C. (1989). Consumer benefit segmentation using clusterwise linear regression. *International Journal of Research in Marketing*, 6(1), 45-59.

Wrigley, N., & Lowe, M. (2010). The globalization of trade in retail services. *Organisation for Economic Co-operation and Development*.

Zhang, Z., & Takane, Y. (2010). Statistics: multidimensional scaling. *International Encyclopedia of Education*, 304-311.

Appendices

Appendix 1 – Questionnaire



Introduction

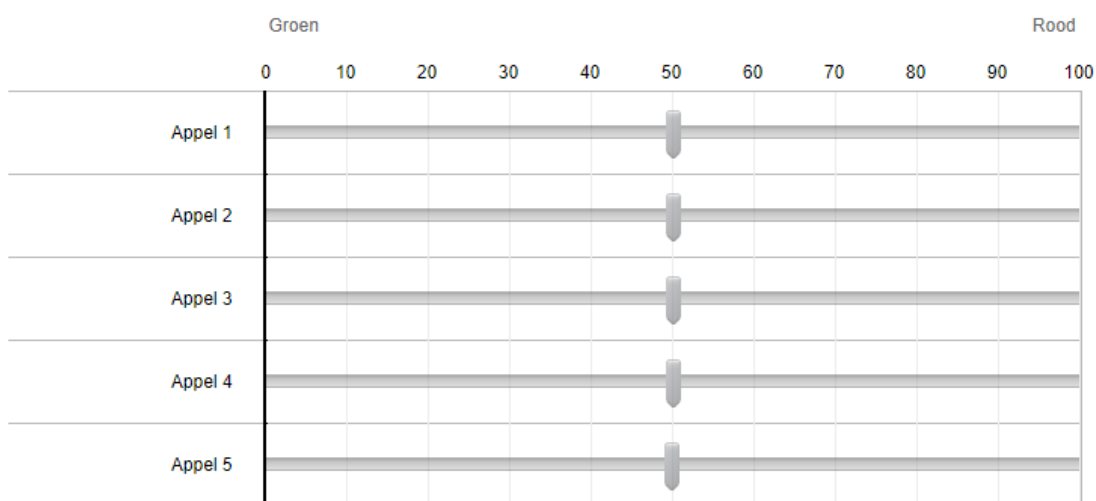
Beste deelnemer,

bedankt voor uw deelname aan dit onderzoek over appels.

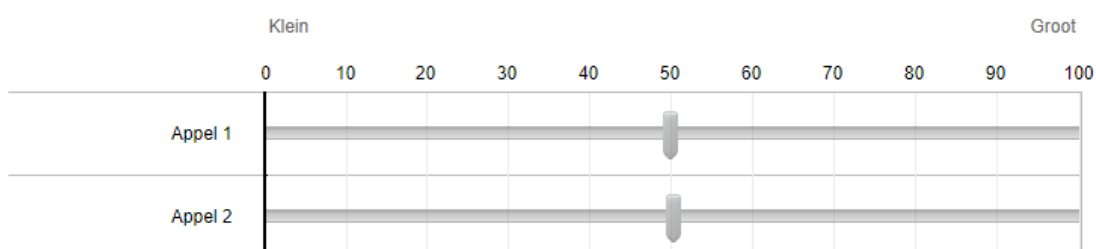
Dit onderzoek maakt deel uit van mijn Master Thesis die ik schrijf om mijn opleiding af te ronden aan Wageningen University & Research (WUR). Het proeven van de appels en het invullen van de vragenlijst duurt ongeveer 10 minuten. Beantwoord de vragen alstublieft eerlijk, er zijn geen goede of foute antwoorden. De resultaten worden uitsluitend voor dit onderzoek gebruikt en de antwoorden blijven anoniem. Als u vragen hebt over dit onderzoek, kunt u contact opnemen met Sanne via sanne.hofstee@wur.nl

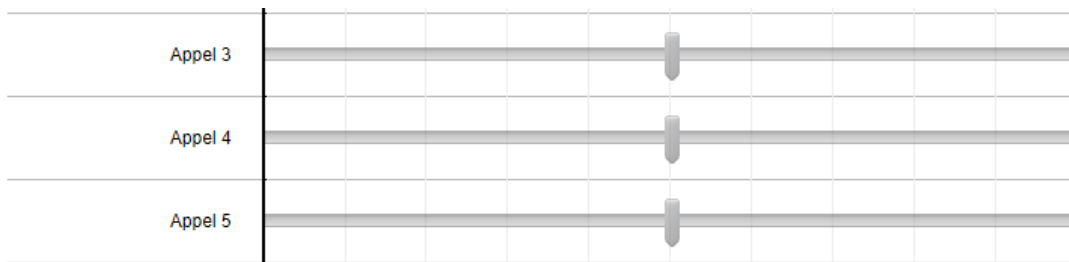
Voordat u de appels gaat proeven, wil ik u vragen om de appels eerst goed te **bekijken**. U mag de appels aanraken, maar **eet ze niet**.

Q1. Beoordeel de appels op hun **schil kleur**

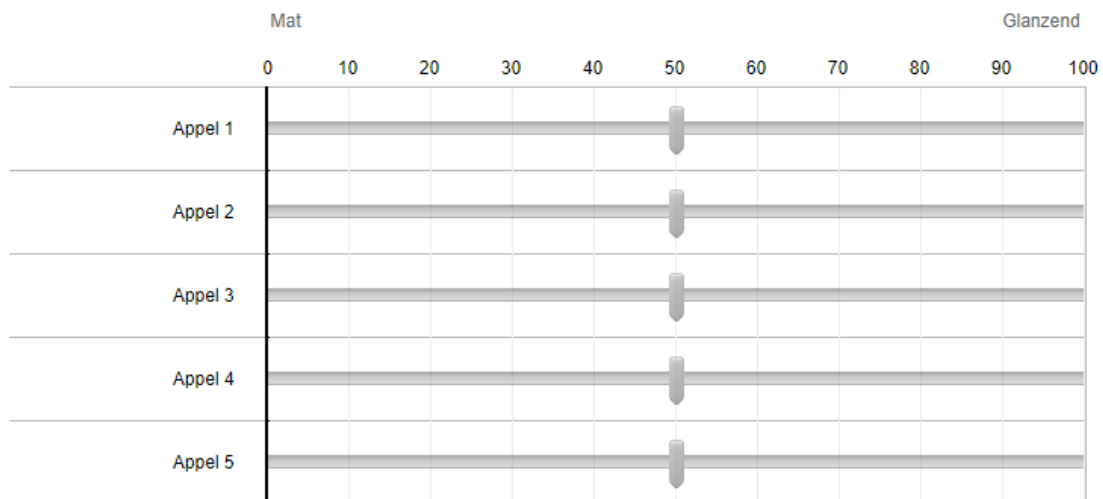


Q2. Beoordeel de vijf appels op hun **grootte**

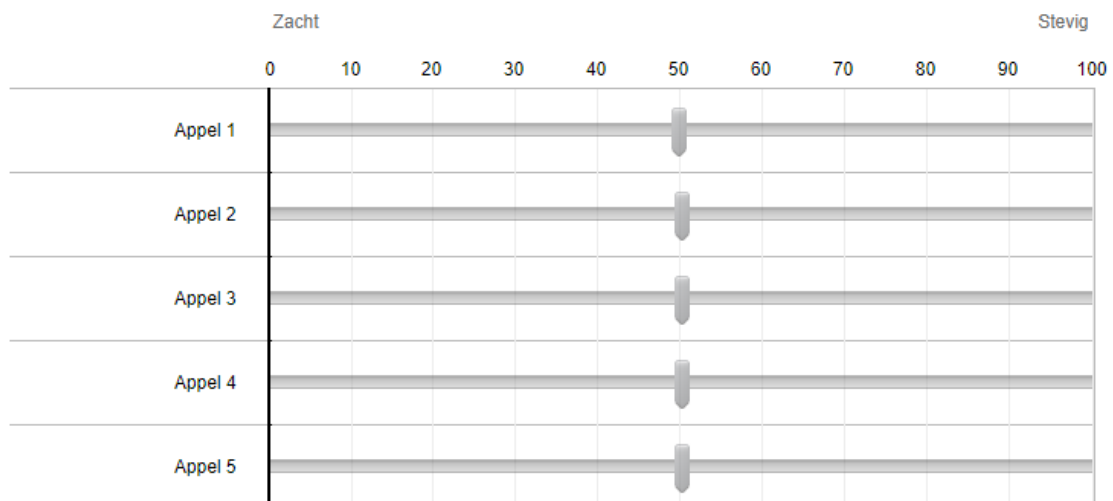




Q3. Beoordeel de vijf appels op hun **glans**



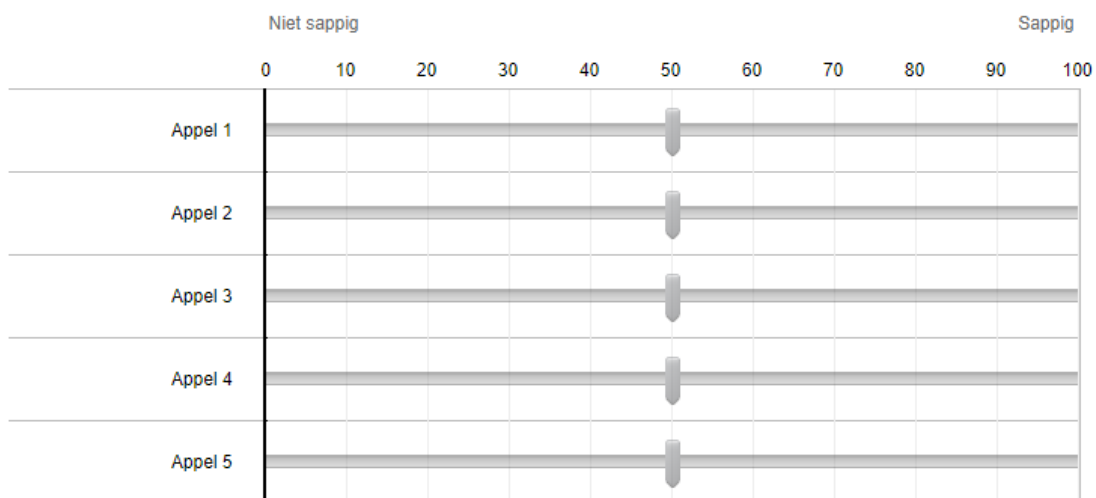
Q4. Beoordeel de vijf appels op hun **stevigheid**



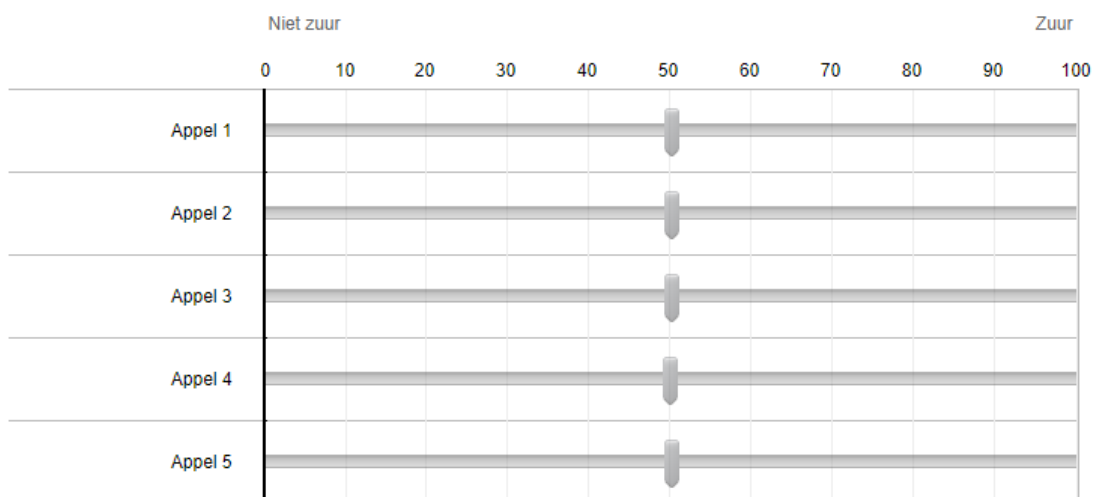
Vraag de onderzoeker om appelpartjes. U krijgt van elke appel twee partjes. U mag deze appelpartjes nu eten. Neem de tijd, u kunt doorgaan met eten terwijl u de resterende vragen beantwoordt. U kunt crackers en water gebruiken om uw smaak tussentijds te neutraliseren.

Q5. Beoordeel de appels op hun **knapperigheid**

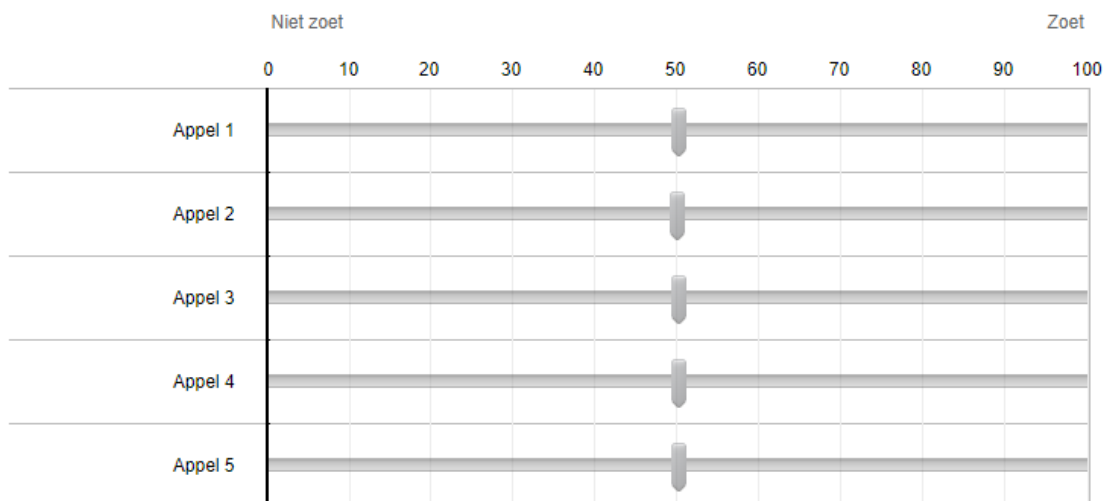
Q8. Beoordeel de appels op hun **sappigheid**



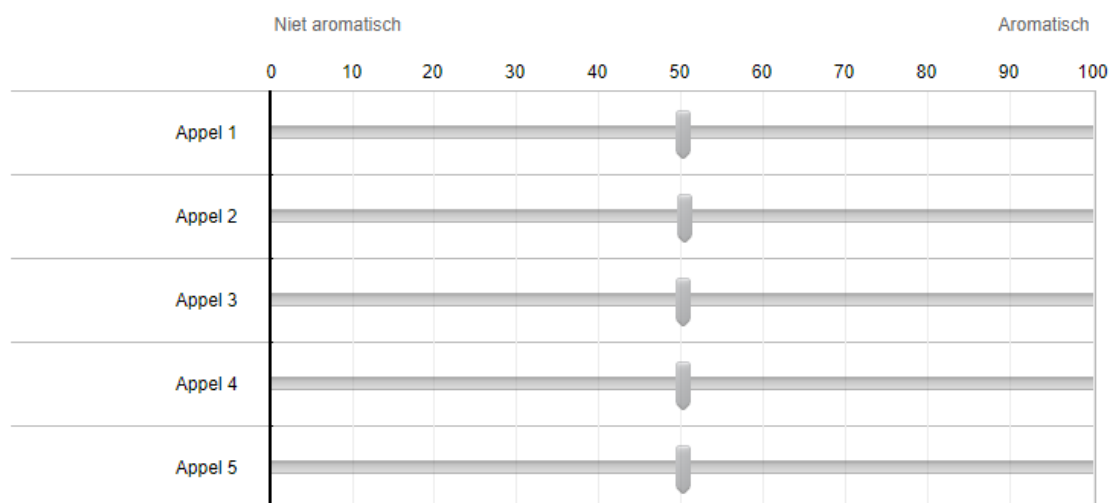
Q9. Beoordeel de appels op hun **zuurheid**



Q10. Beoordeel de appels op hun **zoetheid**

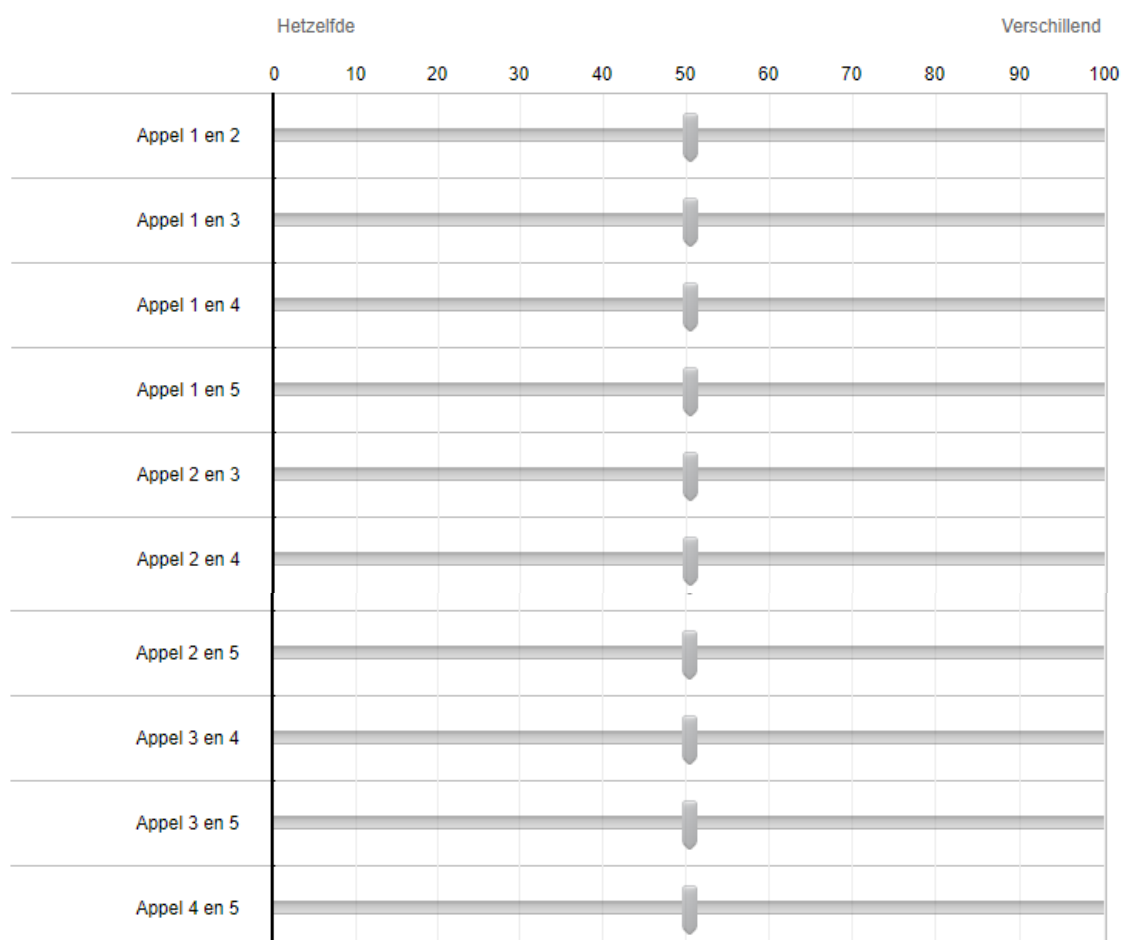


Q11. Beoordeel de appels op hun **aroma**. Aroma omvat alles wat met de smaak van de appel te maken heeft, behalve zoetheid en zuurheid.



Q12. Hoe verschillend vindt u de appels?

Gebruik uw eigen criteria waarop u deze verschillen baseert. Als u denkt dat twee appels erg hetzelfde zijn, sleept u de pijl naar links; als u denkt dat twee appels heel verschillend zijn, sleept u de pijl naar rechts.



Q13. Welke appels hebben uw voorkeur?

(Geef een hogere score voor een hogere voorkeur)

	Minste voorkeur	1	2	3	4	5	6	Meeste voorkeur	7
Appel 1									
Appel 2									
Appel 3									
Appel 4									
Appel 5									

Q14. Wat is uw geslacht?

- ☐ Man
☐ Vrouw

Q15. Wat is uw geboortjaar?

Q16. Hoeveel keer per maand eet u appels?

Q17. Wat is uw hoogst genoten opleiding?

- ☐ VMBO
☐ HAVO
☐ VWO
☐ MBO
☐ HBO
☐ WO

Dit is het einde van deze vragenlijst. Heel erg bedankt voor uw deelname!

Als u een korte samenvatting van de resultaten van deze studie wilt ontvangen, of als u vragen of opmerkingen heeft, neemt u dan gerust contact met mij op (sanne.hofstee@wur.nl).

Wilt u kans maken op de Bol.com gift card ter waarde van 75 euro? Vul dan hieronder uw naam en e-mail adres in. De winnaar wordt binnen 4 weken na afronden van het onderzoek bekend gemaakt.

Naam:

E-mail adres:

Appendix 2 – Mean dissimilarity ratings winter apples with standard deviation

Pair	Mean	Std.dev.
1 and 2	81.06	23.44
1 and 3	72.11	22.34
1 and 4	80.75	20.15
1 and 5	75.42	20.78
1 and 6	74.92	21.25
1 and 7	71.86	24.55
1 and 8	77.89	17.08
1 and 9	71.33	19.98
1 and 10	71.89	25.98
2 and 3	54.11	22.48
2 and 4	51.72	22.02
2 and 5	55.28	22.68
2 and 6	63.67	27.14
2 and 7	55.50	24.68
2 and 8	57.22	25.24
2 and 9	60.42	19.53
2 and 10	62.56	23.23
3 and 4	51.44	23.83
3 and 5	54.03	21.52
3 and 6	74.36	22.88
3 and 7	58.50	22.19
3 and 8	51.33	23.68
3 and 9	51.86	26.80
3 and 10	58.83	20.29
4 and 5	56.14	19.30
4 and 6	66.69	25.33
4 and 7	65.72	20.41
4 and 8	55.19	24.16
4 and 9	53.22	25.96
4 and 10	60.89	24.18
5 and 6	65.03	23.89
5 and 7	56.17	26.44
5 and 8	53.47	22.21
5 and 9	59.06	24.33
5 and 10	63.39	19.64
6 and 7	73.97	22.94
6 and 8	64.78	21.24
6 and 9	67.44	21.30
6 and 10	68.78	20.18
7 and 8	53.61	19.92
7 and 9	64.58	21.23
7 and 10	57.44	23.78
8 and 9	53.42	20.91
8 and 10	64.11	22.10
9 and 10	58.33	22.39

Appendix 3 – Ideal point coordinates per respondent

Respondent	Ideal_x	Ideal_y
1	0.13	-0.13
2	0.11	-0.57
3	0.27	0.39
4	0.51	-0.3
5	0.37	-0.02
6	-0.18	-0.56
7	0.18	0.16
8	0.61	-0.19
9	0.71	0.05
10	-0.3	0.39
11	0.51	0.04
12	0.07	0.44
13	0.64	-0.29
14	0.35	-0.64
15	-0.27	-0.84
16	-0.13	0.47
17	0.41	0.45
18	0.83	-0.44
19	0.35	0.19
20	0.35	0.25
21	-0.58	0.31
22	-0.2	0.09
23	0.38	-0.55
24	0.29	-0.24
25	0.04	-0.82
26	0.41	0.08
27	0.48	0.53
28	0.22	0.12
29	-0.26	0.3
30	0.19	0.81
31	0.69	-0.17
32	0.66	-0.03
33	0.47	-0.02
34	-0.68	-0.22
35	0.57	-0.19
36	-0.18	-0.87
37	-1.06	0.03
38	-0.18	-1.21
39	0.55	-0.96
40	-0.29	0.08
41	0.78	-0.1
42	-0.11	0.11
43	0.78	0.06
44	0.35	0.27
45	0.54	0.21
46	-0.22	-1.03
47	1.23	0.02

Respondent	Ideal_x	Ideal_y
82	0.49	-0.25
83	-0.2	0.59
84	0.42	-0.62
85	0.48	-0.35
86	0.45	-0.23
87	0.28	-0.53
88	0.53	-0.04
89	0.23	-0.89
90	0.82	-0.07
91	0.13	0.3
92	0.99	0.39
93	0.66	0.22
94	0.17	-0.24
95	0.11	0.51
96	-0.15	-0.49
97	-0.1	-0.38
98	-0.55	-0.17
99	0.25	-0.08
100	-0.87	-0.17
101	-0.64	-0.55
102	0.52	-0.53
103	0.93	0.18
104	0.88	-0.06
105	-0.38	-0.09
106	-0.17	-0.26
107	0.03	0.96
108	0.32	0.16
109	0.61	0.38
110	0.63	0.02
111	0.76	0.22
112	0.27	-0.79
113	0.45	-0.18
114	-0.23	-0.05
115	-0.66	0.31
116	-0.28	0.65
117	-0.22	0.51
118	0.33	-0.09
119	-0.01	-0.23
120	0.18	-0.4
121	0.04	0.55
122	0.08	0.34
123	0.82	-0.06
124	-0.04	0.24
125	0.22	0.12
126	0.21	0.1
127	0.19	0.27
128	0.4	-0.77

48	0.72	0.4
49	0.14	0.05
50	0.36	0.28
51	0.14	0.26
52	0.43	-0.21
53	1.23	0.26
54	-0.66	-0.16
55	0.93	-0.34
56	0.14	0.13
57	0.7	-0.06
58	0.57	-0.04
59	0.31	-0.06
60	0.26	-0.13
61	0.61	-0.03
62	-0.49	-0.42
63	-0.93	-0.22
64	0.07	-0.02
65	0.22	-1.01
66	0.02	-0.56
67	0.1	0.58
68	0.26	-0.59
69	0.12	0.56
70	0.11	-0.45
71	0	-0.2
72	0.74	0.2
73	0.36	0.16
74	0.06	-0.1
75	0.46	0.51
76	-0.26	0.81
77	0.13	0.64
78	0.64	0.41
79	0.38	0.11
80	0.17	0.62
81	0.33	0.22

129	0.65	-0.3
130	-0.11	-0.11
131	0.62	-0.05
132	-0.04	-0.09
133	0.72	0.25
134	-0.27	0.51
135	0.1	-0.64
136	0.33	0.39
137	-0.5	-0.45
138	0.53	0.02
139	0.24	-0.24
140	-0.24	-0.63
141	0.13	0.54
142	0.21	-0.37
143	0.05	0.23
144	0.5	-0.33
145	0.61	0.02
146	0.52	0.36
147	0.78	0
148	0.35	-0.74
149	-0.26	-0.64
150	0.3	0.21
151	0.74	0.34
152	-0.31	0.07
153	0.69	-0.27
154	0.56	0.08
155	0.2	0.17
156	0.65	-0.13
157	-0.07	0.8
158	0.49	0.47
159	0.22	0.2
160	0.31	-0.11
161	0.63	-0.52
162	0.62	0.2

Appendix 4 – Cluster membership per respondent

Respondent	Cluster
1	1
2	2
3	3
4	3
5	1
6	2
7	1
8	3
9	3
10	1
11	3
12	1
13	3
14	2
15	2
16	1
17	3
18	3
19	1
20	1
21	1
22	1
23	2
24	1
25	2
26	1
27	3
28	1
29	1
30	1
31	3
32	3
33	3
34	4
35	3
36	2
37	4
38	2
39	2
40	1
41	3
42	1
43	3
44	1
45	3
46	2
47	3

Respondent	Cluster
82	3
83	1
84	2
85	3
86	3
87	2
88	3
89	2
90	3
91	1
92	3
93	3
94	1
95	1
96	2
97	1
98	4
99	1
100	4
101	4
102	2
103	3
104	3
105	1
106	1
107	1
108	1
109	3
110	3
111	3
112	2
113	3
114	1
115	1
116	1
117	1
118	1
119	1
120	1
121	1
122	1
123	3
124	1
125	1
126	1
127	1
128	2

48	3
49	1
50	1
51	1
52	3
53	3
54	4
55	3
56	1
57	3
58	3
59	1
60	1
61	3
62	4
63	4
64	1
65	2
66	2
67	1
68	2
69	1
70	1
71	1
72	3
73	1
74	1
75	3
76	1
77	1
78	3
79	1
80	1
81	1

129	3
130	1
131	3
132	1
133	3
134	1
135	2
136	3
137	4
138	3
139	1
140	2
141	1
142	1
143	1
144	3
145	3
146	3
147	3
148	2
149	2
150	1
151	3
152	1
153	3
154	3
155	1
156	3
157	1
158	3
159	1
160	1
161	2
162	3

Appendix 5 - PROXSCAL – alternative data input

In this research it was tried to put the dissimilarity data into PROXSCAL in another way than a lower triangle dissimilarity matrix with manually calculated averaged dissimilarities for the rated pairs. The data was structured as follows. The respondents were presented in one column and the rated pairs and rating for the pairs were represented in rows. For example, respondent 1 rated pair 2 and 3 and gave a score of 68 for the dissimilarity, then the data was structured as:

Respondent	Apple x	Apple y	score
1	2	3	68

This was done for each rating by all respondents. This gave the opportunity to keep the effect of individual scores and the variance in the analysis, instead of taking the average of the scores on each pair. When running this procedure in SPSS PROXSCAL, it was important to specify that the proximities are in a single column, and not in a matrix. When running the analysis, an error occurred, and weights needed to be specified for PROXSCAL to be able to deal with the missing data. Therefore artificial data was created: for each respondent 5 additional pairs were added (1 and 2, 3 and 4, 5 and 6, 7 and 8, 9 and 10) with score 0. An additional column with weights was added, 1 for the actual data and 0 for the artificial data. When running the analysis with these weights specified, there was an output. This output, however was not interpretable. Apple varieties that were not considered comparable in any way were found close to one another in the map. Theoretically this method should give better results, because the variance for each respondent is taken into account and the stress per respondent could be inspected to identify respondents having a high stress value. In practice however, this method does not lead to an interpretable solution. It could be interesting to look further into this method in future research. For the sake of being able to compare the results found in this research with previous research it was chosen to work with a lower triangle dissimilarity matrix. Besides, the results obtained from that analysis provided a more interpretable solution.