

MASTER

Outdating of perishables at PLUS supermarkets; part of doing business?

Kock, R.A.L.

Award date:
2015

[Link to publication](#)

Disclaimer

This document contains a student thesis (bachelor's or master's), as authored by a student at Eindhoven University of Technology. Student theses are made available in the TU/e repository upon obtaining the required degree. The grade received is not published on the document as presented in the repository. The required complexity or quality of research of student theses may vary by program, and the required minimum study period may vary in duration.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain

Eindhoven, April 2015

**Outdating of perishables at PLUS
supermarkets; part of doing business?**

By
R.A.L. Kock

BSc Industrial Engineering (2013)
Identity number: 0656439

in partial fulfillment of the requirements for the degree of

Master of Science

in Operations Management & Logistics

Supervisors:

Ing. P. Van de Voorde

Dr. K.H. van Donselaar

Dr. ir. R.A.C.M. Broekmeulen

Hollander Logistiek, Hollander Barendrecht

Eindhoven University of Technology

Eindhoven University of Technology

TUE School of Industrial Engineering

Series Master Thesis Operations Management and Logistics

Subject headings: retail, supermarket, perishable products, assortment, waste, outdating

Abstract

This report describes which causes lead to outdating of perishables as a part of the total waste in PLUS supermarkets. The project is conducted in the convenience and in the meat products and salads ZB categories.

Data from PLUS Retail and Hollander Barendrecht like sales and outdating data is linked with observations and semi-structured interviews at several supermarkets throughout The Netherlands. Insight is given in the difference between supermarkets and products, the causes of outdating and ways to reduce the current amount of outdating.

Management summary

Fresh products are of increasing importance for Dutch supermarkets. Typical supermarkets are full of fresh products like vegetables, cheese and meat. But this illusion of abundance comes with an enormous cost. The total waste of products of Dutch supermarkets is estimated on 977 million Euros excluding VAT a year (Maarse, 2015). This number includes waste from all possible causes such as theft, outdating and broken products.

The total waste of products of PLUS supermarkets is estimated on 55,7 million Euros excluding VAT for 2014. This means that the PLUS supermarkets have an average waste of 2,70% of the total revenue. Most waste can be found in the flowers and plants category followed by the potatoes, fruit and vegetables category.

As the margins for supermarket retailers are very small and still decreasing due to an increasing importance of promotions, the percentage waste of the total revenue is an important factor for retailers. This project focused in specific on outdating of perishables as a part of the total waste in the supermarkets. It investigated which causes lead to outdating of these products and gives insights to HB, PLUS Retail and the supermarket managers. Data from PLUS Retail and HB like sales and outdating data is linked with observations and semi-structured interviews at several supermarkets throughout The Netherlands. The project is conducted in the convenience and in the meat products and salads ZB categories.

Opinion about outdating

Two types of supermarket manager were classified: grocer and entrepreneur. They have a total different opinion on outdating. Grocers typically want to get the lowest percentage outdating as possible, whereas entrepreneurs take outdating as a fact of doing business. Entrepreneurs argue that focusing too much on outdating will result in more out-of-stocks and an assortment that is too small. Some entrepreneurs even order more products than advised from the automated store ordering system to get a better store image. They simply accept that this will result in more outdating.

Waste of products is in the current situation the full responsibility of the supermarket manager. Some managers argue that this is not fair as the category management of PLUS Retail initially determines which products should be sold in the supermarkets and which products shouldn't. Often the category management makes shelf plans with a too broad assortment for the revenue that a supermarket has. This results in outdating of products. Because of this, the supermarket managers argue that the category management should be partly responsible for the waste in the PLUS supermarkets.

Opinion about selling products with a remaining shelf life of one day

Supermarkets can decide themselves whether they want to sell products with a remaining shelf life of one day or not. Following from the interviews, PLUS managers totally disagree about this topic. On the one hand managers, mostly grocers, argue that selling these products is necessary to run a profitable supermarket. On the other hand, the grocers and entrepreneurs that don't sell these products argue that selling these products is not in line with the core values of PLUS. 71,4% of the grocers sell these products in contrast to 46,7% of the entrepreneurs. In the supermarkets that sell these products, these products are mostly discounted on the whole or part of the day of expiration.

Opinion about discounting

Not in all PLUS supermarkets products with a short remaining shelf life are discounted. The managers who discount products see discounting products with a short remaining shelf life as a way to increase the chance that these products will be sold before they expire and in this way prevent outdating. Furthermore, managers argue that to prevent outdating using the right parameter settings in the automated store ordering system is way more important than discounting products. The way in which managers discount their products differs a lot between supermarkets.

Almost all interviewees argue that discounted products are not beneficial for the store image. In an estimated 14,5% of all PLUS supermarkets products are never or seldom discounted when they have a short remaining shelf life. Next to the decision to discount products or not, some managers decide to discount only products in certain categories.

Opinion about grabbing behavior

In all visited supermarkets grabbing behavior is observed. Grabbing behavior occurs when consumers buy products with a later best-before-date than the first product in the shelf with a shorter best-before-date given that there are two or more best-before-dates in the shelf.

The amount of observed grabbing behavior differs a lot between supermarkets, but we cannot compare supermarkets as no quantitative data is available about the grabbing behavior. Managers argue that part of the grabbing behavior can be prevented when customers have trust in the supermarket. However, some customers will always check the full shelf for products with the longest remaining shelf life. In 62,2% of the visited supermarkets rules exist for the amount of different best-before-dates that may be placed in the shelves.

Data analysis on supermarket level

Analysis on supermarket level gives in both categories clear insights in the factors that lead to outdating on supermarket level. In both analyses the average ratio between the case pack sizes and the average sales during the shelf lives of the products of a supermarket (supermarket QMuM ratio) is the most important factor for the outdating on supermarket level. The supermarket QMuM ratio has a strong positive relationship with the percentage outdating on supermarket level in both categories. This is in line with our hypothesis and the opinion of the interviewed supermarket managers. This positive relationship is stronger in the convenience category than in the meat products and salads category. The average remaining shelf life at arrival is more than two times higher for the meat products and salads category than for the convenience category. This explains why the supermarket QMuM ratio is more important for the convenience category.

The second most important factor in both categories is the average coefficient of variation of order size. This factor has in both categories a positive relationship with the percentage of outdating. This is in line with our hypothesis. This factor is changeable by the supermarkets by placing orders of the right order size. If the ordering pattern is more stable and the order size is more in line with the expected demand, the variation of the order size will decrease and the percentage of outdating will decrease according to our model.

The third most important factor in both categories is the average order size. This factor has in both categories a negative relationship with the percentage of outdating. This factor is totally under

control of a PLUS supermarket. No hypothesis was formulated about this variable as the average order size on its own doesn't say a lot about the outdating of a supermarket.

The percentage of order advises that is followed is significant in both categories. This factor is totally under control of a PLUS supermarket. However, in both categories the significance of this factor is just low enough to be included in our model. This means that we have to be careful about conclusions based on this factor. In the convenience category a strong positive relationship is found. The positive relationship found in the data is possibly coming from the fact that the automated store ordering system of PLUS doesn't include weather forecasts, promotions and seasonal effects in the order advises. Next to this, the system doesn't have a tradeoff between waste and service level. Delaying a planned order is likely to result in less outdating and a lower service level. Weather changes, promotions and seasonal effects will have an influence on the sales. Because of this, these effects should be predicted so that order advises can be changed only if needed. In the meat products and salads category we see a small negative relationship. The percentage of the total revenue coming from promotions in the convenience category is 1,6 times more than in the meat products and salads category. Next to this, the sales in the convenience category are more sensitive to weather changes. This two reasons combined explain why the order advises can be trusted more in the meat products and salads category than in the convenience category.

The category revenue share has in both categories a negative relationship with the percentage of outdating. This is in line with our hypothesis. However, in both categories the significance of this factor is just low enough to be included in our model. This means that we have to be careful about conclusions based on this factor. The negative relationship is much stronger in the convenience category than in the meat products and salads category.

Analysis on product level

Analysis on product level gives in both categories clear insights in the factors that lead to outdating on product level. In both analyses the average ratio between the case pack size and the average sales during the shelf lives of the product (product QMuM ratio) is the most important factor for the outdating on product level. The product QMuM ratio has a strong positive relationship with the percentage outdating on product level in both categories. This is in line with our hypothesis and the opinion of the interviewed supermarket managers.

Analysis on supermarket – product combination

Analysis on supermarket-product combinations gives in both categories clear insights in the factors that lead to outdating in a supermarket-product combination. In both analyses the ratio between the case pack size and the average sales during the shelf lives of the product (combination QMuM ratio) is the most important factor for the outdating of a combination. The combination QMuM ratio has a strong positive relationship with the percentage outdating in a combination in both categories. This is in line with our hypothesis and the opinion of the interviewed supermarket managers. This positive relationship is stronger in the convenience category than in the meat products and salads category. This is reasonable due to the differences in remaining shelf lives.

A second important factor in both categories is the natural logarithm of the average remaining shelf life. The natural logarithm of the average remaining shelf life has a negative relationship with the percentage outdating of a combination in both categories. This negative relationship is stronger in

the convenience category than in the meat products and salads category. This is reasonable due to the differences in remaining shelf lives.

Extra analysis on supermarket level

Both analyses on supermarket-product combinations show that supermarkets classified as grocers have on average a higher absolute and relative number of combinations with structural waste than supermarkets classified as entrepreneurs.

Both extra analysis on supermarket level with detailed information about whether they sell product with a remaining shelf life of one day, the type of manager, the discount percentage, the discount hours and the average distance didn't show any significant variables. So, we have to conclude that none of these factors has a significant influence on the percentage of outdating of perishables in the convenience or the meat products and salads category.

Outdating of perishables; part of doing business?

Based on the analyses in this report we can conclude that a large part of the outdating of perishables is coming from the width of the assortment of PLUS supermarkets. By choosing for a broad assortment of perishables, PLUS supermarkets choose to include perishables in the assortment with a large chance of outdating. So, part of the outdating of perishables is part of doing business in a PLUS supermarket.

However, this doesn't mean that the percentage of outdating should be accepted as a norm and taken for granted. PLUS Retail, HB and their entrepreneurs should still put effort in ways to reduce the current amount of outdating. Especially supermarkets with a high revenue should set a low norm for the percentage of outdating. They should be able to achieve a low percentage of outdating as they are able to achieve lower QMuM ratios for their products than supermarkets with a low revenue. In the following paragraph ways to reduce the amount of outdating will be explained. A lower amount of outdating will lead to a higher profitability of the supermarkets and a better image of PLUS as a sustainable supermarket chain.

Ways to reduce outdating

Some ways to reduce outdating of perishables can be formulated. As a first step, the registration of sales and outdating should become more accurate in order to be able to identify factors that cause outdating earlier. This also leads to more accurate settings in the automated store ordering system. The right parameter settings in the automated store ordering system are way more important than responding on left over products by discounting products.

The next step is to choose the right assortment for a supermarket. Following from the data analysis the ratios between the case pack sizes and the average sales during the shelf lives of the products (QMuM ratios) are the most important factors for outdating of perishables. This also follows from the fact that the 10% of the assortment consisting of the articles with the most outdating result in 36 till 50 percent of the total outdating per category. Shelf plans are now based on the meters of refrigerators instead of revenue of a supermarket. In some supermarkets this leads to products with a structural waste problem. Supermarkets should be critical about the width of the assortment that is possible in their supermarket. Products with a lot of outdating should only be kept in the assortment when the loss of customers coming for these products is larger than the savings in outdating.

The third step is only relevant for categories with strong weather, promotional or seasonal effects like the convenience category. The automated store ordering system of PLUS doesn't include weather forecasts, promotions and seasonal effects in the order advises. Because of this, these effects should be predicted so that order advises can be changed only if needed.

The fourth step is the right placing and handling of perishables. Unnecessary outdating due to grabbing behavior or wrong handling should be avoided. Smaller distances between shelves will prevent the undesirable grabbing behavior. Next to this, some shelves are stacked too full what will result in damaged products. Finally, some products were placed in a non-optimal environment which will speed up the decay process of the products.

The registration of sales and outdating, and the settings in the automated store ordering system are the responsibility of the supermarkets themselves. The choice of the right assortment is a common responsibility of the supermarket and PLUS Retail. Finally, HB is only for a small part responsible for the waste in the supermarkets. HB is only responsible for the waste if the delivered quality is not good or the remaining shelf life at arrival is shorter than agreed with PLUS Retail.

Preface

This report is the result of my master thesis in order to fulfill my master degree in the master Operations Management & Logistics at the Eindhoven University of Technology. Therefore I would like to express my gratitude towards some people that have helped me during my life as a student and during my master thesis project.

This projects represents the end of my life as a student. I have learned a lot during my Bachelor and Master. I am grateful that I have been able to engage in this study and for the opportunities it has given to me, both within and outside the curriculum. My student life was diverse and was characterized by meeting a lot of interesting persons, personal development and an international experience in Southern Africa.

I would like to thank several people who supported me throughout this project. First, I would like thank Peter van de Voorde and the other employees of Hollander Barendrecht for giving me the opportunity to perform my master thesis project at their company. I also would like to thank them for their flexibility, guidance and interesting discussions during my project. Second, I would like to thank the contact persons at PLUS Retail and all interviewed persons at PLUS supermarkets for their openness and for providing me a lot of information. Finally, I would like to thank my university supervisors, Karel van Donselaar and Rob Broekmeulen, for their time, flexibility, and feedback on my project.

For supporting me throughout my time as a student, I would like to thank my friends and parents. I would like to thank my friends for making my time as a student an unforgettable time. I would like to thank my parents for supporting me during my studies, not in the least by financial means. The freedom they gave me to make the best of my time as a student was the best support I could get.

Ruben Kock

Barendrecht, April 2015

Contents

Abstract	i
Management summary	ii
Preface	vii
1. Introduction	1
1.1 Report structure	1
1.2 Problem environment	1
1.3 Problem introduction	2
1.4 Literature review	3
1.4.1 Perishables	3
1.4.2 Causes of waste	3
1.4.3 Freshness of perishables	4
1.4.4 Expected consumer behavior	5
1.4.5 Observed consumer behavior	6
1.4.6 Pricing of perishables	6
1.4.7 Implications of consumer behavior on outdating	7
1.4.8 Consequences of changing the case pack size	8
1.4.9 Ways to force consumers to use FIFO instead of LIFO	9
1.5 Gaps in literature	9
2. Problem Definition	10
2.1 Motivation for project	10
2.2 Research question	10
2.3 Sub-questions	11
2.3.1 Analysis from interviews	11
2.3.2 Analysis from data	11
2.4 Research methodology	11
2.4.1 Analysis 1: analysis on supermarket level	12
2.4.2 Analysis 2: analysis on product level	13
2.4.3 Analysis 3: analysis on supermarket – product combination	13
2.4.4 Analysis 4: extra analysis on supermarket level	14
2.4.5 Left out variables	14
2.5 Hypotheses	16
2.5.1 Hypotheses for analysis 1: analysis on supermarket level	16
2.5.2 Hypotheses for analysis 2: analysis on product level	18
	viii

2.5.3	Hypotheses for analysis 3: analysis on supermarket - product combination	19
2.5.4	Hypotheses for analysis 4: extra analysis on supermarket level	19
2.6	Scope of research	20
2.6.1	Selection of categories	20
2.6.2	Selection of supermarkets for visits	22
2.6.3	Selection of supermarkets for data analysis	22
2.6.4	Selection of sales and outdating data for data analysis	22
2.7	Stakeholders of the project	23
3.	Analysis from interviews	23
3.1	Opinion about outdating and width of product assortment	23
3.2	Opinion about selling products with a remaining shelf life of one day	25
3.3	Opinion about discounting	25
3.3.1	Discount percentage	26
3.3.2	Discount hours	26
3.4	Opinion about grabbing behavior	27
4.	Analysis from data	27
4.1	Analysis of convenience category	27
4.1.1	Selection of supermarkets and data	28
4.1.2	Analysis 1: analysis on supermarket level	28
4.1.3	Analysis 2: analysis on product level	31
4.1.4	Analysis 3: analysis on supermarket – product combination	33
4.1.5	Analysis 4: extra analysis on supermarket level	35
4.2	Analysis of meat products and salads ZB category	35
4.2.1	Selection of supermarkets and data	36
4.2.2	Analysis 1: analysis on supermarket level	36
4.2.2	Analysis 2: analysis on product level	37
4.2.3	Analysis 3: analysis on supermarket – product combination	39
4.2.4	Analysis 4: extra analysis on supermarket level	41
5	Conclusions and recommendations	41
5.2	Analysis from interviews	41
5.2.2	Opinion about outdating	41
5.2.3	Opinion about selling products with a remaining shelf life of one day	42
5.2.4	Opinion about discounting	42
5.2.5	Opinion about grabbing behavior	43

5.3	Analysis from data	43
5.3.2	Analysis on supermarket level	43
5.3.3	Analysis on product level	45
5.3.4	Analysis on supermarket – product combination	47
5.3.5	Extra analysis on supermarket level	48
5.3.6	Outdating of perishables; part of doing business?	48
5.3.7	Ways to reduce outdating	49
6	Limitations	49
7	Further research	50
8	References	51
9	Appendices	54
9.1	Appendix A: percentage waste per category	54
9.2	Appendix B: description of variables	55
9.3	Appendix C: reasoning and calculation of variables	58
9.4.1	Orders	58
9.4.2	Sales	59
9.4.3	Remaining shelf life	61
9.4.4	Outdating	62
9.4.5	Percentages/ratios	62
9.4	Appendix D: regression analysis results	63
9.4.2	Convenience: analysis 1: analysis on supermarket level	63
9.4.3	Convenience: analysis 2: analysis on product level	66
9.4.4	Convenience: analysis 3: analysis on supermarket – product combination	68
9.4.5	Convenience: analysis 4: extra analysis on supermarket level	71
9.4.6	Meat products and salads: analysis 1: analysis on supermarket level	71
9.4.7	Meat products and salads: analysis 2: analysis on product level	74
9.4.8	Meat products and salads: analysis 3: analysis on supermarket – product combination	76
9.4.9	Meat products and salads: analysis 4: extra analysis on supermarket level	79

List of figures

Figure 1: Regulative cycle (Van Strien, 1997) 1
Figure 2: Percentage of waste per category 21
Figure 3: Distribution of outdating over assortment in convenience category 28
Figure 4: Distribution of outdating over assortment in meat products and salads category 35
Figure 5: Overview of percentage waste per category 54

List of tables

Table 1: Description of variables	55
Table 2: Model summary of analysis 1 in convenience category.....	63
Table 3: Regression coefficients of analysis 1 in convenience category	64
Table 4: Model summary of analysis 2 in convenience category.....	66
Table 5: Regression coefficients of analysis 2 in convenience category	66
Table 6: Model summary of final model in analysis 3 in convenience category.....	68
Table 7: Regression coefficients of final model in analysis 3 in convenience category	68
Table 8: Model summary of analysis 4 in convenience category.....	71
Table 9: Model summary of analysis 1 in meat products and salads category.....	71
Table 10: Regression coefficients of analysis 1 in meat products and salads category.....	72
Table 11: Model summary of analysis 2 in meat products and salads category.....	74
Table 12: Regression coefficients of analysis 2 in meat products and salads category.....	75
Table 13: Model summary of final model in analysis 3 in meat products and salads category.....	77
Table 14: Regression coefficients of final model in analysis 3 in meat products and salads category .	77
Table 15: Model summary of analysis 4 in meat products and salads category.....	79

1. Introduction

This chapter will provide an introduction of the research. First the report structure will be stated and then the problem environment will be discussed. Afterwards the problem will be introduced and an overview of the conducted literature review will be stated. Finally the gaps in literature will be described.

1.1 Report structure

The project is roughly structured in three parts: the problem definition phase, the analysis phase, and the drawing of conclusions and recommendations phase. These three parts are part of the regulative cycle (Van Strien, 1997), depicted in Figure 1. The remaining two stages are the intervention and the evaluation phase, but these were left out of the scope of the research due to time restrictions.



Figure 1: Regulative cycle (Van Strien, 1997)

To start with, a description of the problem situation at the PLUS supermarkets will be discussed in this chapter. In chapter 2, the research question, sub questions and the scope will be stated. Next, the analysis and diagnosis phase will be elaborated. This phase includes observations from supermarket visits and empirical data analysis. Finally, conclusions have been drawn based on the findings, and recommendations will be stated.

1.2 Problem environment

This research has been performed at Hollander Barendrecht B.V. (in sequence: HB) in cooperation with PLUS Retail B.V.

HB was founded in 1929 by Siem den Hollander and started as a ‘fruit and vegetables’ company in Rotterdam. In 1975 a rapid growth in the retail industry took HB to the current place of establishment, Barendrecht. Since 2008, after HB relocated to a new distribution center, it became a logistic service provider of both perishables (fruit and vegetables) and semi-perishable products (dairy, convenience goods) for retail. Nowadays, HB daily delivers refrigerated products to 258 PLUS supermarkets in the Netherlands. Furthermore, HB is a subsidiary of The Greenery, an international fresh produce company. HB strives to make the supply chain of perishables as efficient as possible to achieve the lowest supply chain costs and a maximum quality and freshness of the products for the final customer. For the context of the research it should be noted that perishables make up about

half of the sales in supermarkets (Ferguson and Ketzenberg, 2005), covering a large share in sales. For PLUS Retail B.V. the turnover figures show a similar ratio.

HB delivers products to the supermarkets of PLUS Retail B.V. The PLUS supermarkets are part of Sperwer Groep, which is a wholesale organization, started in 1928. PLUS Retail is one of the business units of the Sperwer Groep. The actual PLUS supermarket formula is established from the Plusmarkt introduced in 1988. In 1999 a change in market conditions caused a modification in formula management. This adjustment resulted in the current PLUS formula. PLUS Retail has the ambition to become the best supermarket in the Netherlands within two years from now (Smit, 2015). They distinguish themselves by the following core values:

- Attention: each PLUS supermarket is characterized by real dedication for customers and the best service;
- Quality: all products, employees and supermarkets have a high quality;
- Local: each PLUS supermarket offers a wide range of fresh and locally produced products at competitive prices and is strongly involved in the local community;
- Responsible: for each PLUS entrepreneur animal and environmentally friendly working is the first priority.

PLUS Retail is a company focused on the entrepreneurs. In contrast to most retailers in The Netherlands, the entrepreneurs of PLUS have a lot of freedom with regards to shelf plans, purchasing decisions and local promotions. The PLUS supermarkets had a total revenue of 2 billion Euros and a market share of 5,9 percent in 2014 (Meijssen, 2015). The total revenue increased with 3,1% in 2014 in a market that grew only 0,8%. The shop floor productivity was 160 Euros per square meter sales area. (Maarse and Te Pas, 2015). Almost all PLUS supermarkets are owned by entrepreneurs. The average net sales area is 995 square meters.

1.3 Problem introduction

Fresh products are of increasing importance for Dutch supermarkets. Typical supermarkets are full of fresh products like vegetables, cheese and meat. But this illusion of abundance comes with an enormous cost. The total waste of products of Dutch supermarkets is estimated on 977 million Euros excluding VAT a year (Maarse, 2015). This number includes waste from all possible causes such as theft, outdating and broken products.

The total waste of products of PLUS supermarkets is estimated on 55,7 million Euros excluding VAT for 2014. This means that the PLUS supermarkets have an average waste of 2,70% of the total revenue. Most waste can be found in the flowers and plants category followed by the potatoes, fruit and vegetables category. In Appendix A an overview can be found of the waste per product category.

As the margins for supermarket retailers are very small and still decreasing due to an increasing importance of promotions, the percentage waste of the total revenue is an important factor for retailers. This project focuses in specific on outdating of perishables as a part of the total waste in the supermarkets. It investigates which causes lead to outdating of these products and gives insights to HB, PLUS Retail and the supermarket managers.

1.4 Literature review

In this chapter the literature earlier discussed in Kock (2014) will be briefly summarized.

1.4.1 Perishables

This master thesis project is focused on perishable items as most of the waste at PLUS supermarkets is coming from these products. As Van Donselaar et al. (2006) describe “the main difference between perishables and non-perishables is the ‘Shelf Life’. The shelf life of a product is measured in days, counting from the day it is produced until the product becomes unacceptable for consumption or obsolete.” They further state that a perishable item should meet one of the following two criteria: “1. The high rate of deterioration at ambient storage conditions requires specific storage conditions at the store and/or at the consumer to slow the deterioration rate. 2. The obsolescence date of the product is such that reordering for the products with the same day is impractical.” (Van Donselaar et al., 2006). They classify perishables as products with a shelf life less than or equal to 30 days.

Van Donselaar et al. (2006) argue that next to this difference in shelf life there are also differences between perishables and non-perishables in terms of several key sales and logistic product characteristics. There is a significant difference in average sales per week, coefficient of variation of weekly sales, potential delivery frequency, average case pack size and minimum stock norm.

1.4.2 Causes of waste

Waste in supermarkets is a structural problem. For example, in the United States in-store food losses were an estimated 43 billion pounds in 2008, equivalent to 10 percent of the total food supply at the retail level (Buzby et al., 2011). IG&H (Meijssen, 2011) estimate the amount of waste at supermarkets at one to three percent of the purchase value. They argue that the amount of waste mainly is on the profile and operational development of a store. For perishables in specific, Ferguson and Ketzenberg (2005) even argue that retailers can have losses up to 15 percent due to damage and spoilage. Next to that, Gunders (2012) argues that the retail model views waste as a part of doing business. Waste can have several causes. Van Burgh (2007) distinguishes five causes of waste:

1. Short shelf life in combination with low sales
2. Broken products
3. Ordering too much
4. Left over from promotion
5. Not attractive products (e.g. damaged or unpopular)

Most waste in his research in two Dutch supermarkets was coming from products with a short shelf life in combination with low sales. The necessary time to sell the case pack size was longer than the remaining shelf life of the products. Some of these products have a structural waste problem, but are strategic products for the supermarket, so they cannot be excluded from the assortment. Broken products can be a result of the consumers in the supermarkets or the employees. Waste in this category can partly be avoided when employees work more carefully. In the research of Van Burgh (2007), ordering too much was most of the time the result of an incorrect order amount coming from the automated store ordering (ASO) system due to a peak in demand in the previous period. This can be corrected by employees who recognize the unusual peak in demand.

Next to the above mentioned causes, Gunders (2012) identifies three other relevant causes of waste:

6. Overstocked product displays
7. Expectation of cosmetic perfection
8. Availability of fresh, ready food until closing

Most supermarket owners assume that customers buy more products from fully stocked shelves than from not fully stocked shelves. Because of this they sometimes order more products than necessary for the predicted demand. Another reason for waste is when supermarkets throw away products (mainly fruit and vegetables) that are not perfect. They expect that customers don't want to buy those items. The last reason stated by Gunders (2012) is that supermarkets want to offer prepared, ready-made food the whole day. This food can only be sold for some hours and at the end of the day the leftovers have to be thrown away. Another reason for waste in supermarkets is:

9. Theft

Theft can be committed by employees and customers. This cause of waste is hard to identify as products can be stolen during the transport from the distribution center to the supermarket or in the supermarket itself. Another difficulty is that monitoring possible theft is sometimes more costly than the value of the stolen goods itself. However, at some stores camera's or security guards are present to prevent theft.

1.4.3 Freshness of perishables

Entrup (2006) argues that product freshness is one of the most important buying criteria for consumers. Freshness has even replaced price as the primary concern of consumers regarding food. Consumers have a higher level of education, more knowledge about nutrition and more money available to spend for food products nowadays. For consumers, product freshness is a major part of product quality and will help to prevent health problems. According to Entrup (2006) consumers have two main sources of information concerning product freshness. Firstly, they can use their senses to evaluate the sensory quality of the product. Second, they can judge the freshness by checking the shelf life.

There is no generally accepted definition in literature for the term shelf life. In this report we use the definition issued by the Institute of Food Science and Technology according to which shelf life is defined as: "the time during which the food product will remain safe, be certain to retain the sensory, chemical, physical and microbiological characteristics, and comply with any label declaration of nutritional data". The shelf life is states in days and is determined by the manufacturer. Manufacturers have to spend considerable effort in determining the right shelf life as the products has to be of good quality at the end of the shelf life, but on the other hand consumers associate very long shelf lives with poor product quality (Kilcast & Subramaniam, 2000). Short shelf lives can for example improve the perceived freshness of ready-to-eat meals (Soethoudt et al., 2012).

Entrup (2006) argues that the actual shelf life depends on four major factors: formulation, processing, packaging and storage conditions. All of these factors are critical: however, their relative importance depends on the perishability of the product. Several mechanisms cause the deterioration of the food product. Extensive research on the decay of perishables can be found in e.g. Labuza and Taoukis (1990), Kilcast and Subramaniam (2000), and Hine (1987).

1.4.4 Expected consumer behavior

This chapter describes the expected consumer behavior with regards to withdrawal of perishables. Next chapter will be focused on the actual consumer behavior based on field research. Many research in the field of perishables inventory models, like Van de Ven (2014), assume that the withdrawal policy at the store is First In First Out (FIFO). This means that consumers at a store always take the oldest products from the shelves when there are products available with different best-before-dates.

The analysis of LIFO inventory systems is considerably less developed than that of FIFO systems (Cohen and Prastacos, 1981). According to Cohen and Prastacos (1981) the LIFO policy is observed in situations where the consumer has control of or influence on the issuing decision. For example, this is the case when consumers can choose between products with different best-before-dates. They predict higher outdateding under LIFO which acts to lower optimal inventories, however they argue that this difference is very small.

Nahmias (1982) supports this opinion and argues that in many real systems the user determines the issuing policy and when the utility of the new units is higher, LIFO issuing will be the result. For consumers in a supermarket the products with a later expiration date will have a higher utility, because the consumer has more days to use the perishable item. Nahmias (1982) argues that consumers will observe expiration dates on shelf items and will choose the newest item. Zhou and Yang (2003) also argue that consumers consider the loss in freshness and always prefer to buy last-in items and buy them even though they may be put at the back. So, they follow a LIFO policy. Cohen and Prastacos (1976) argue that the order-up-to-level is relatively insensitive to the choice of the issuing policy even though the optimal expected cost is significantly higher for LIFO. Nahmias (1982) argues that this proves that the relatively simple approximations derived in for example Nahmias (1976) which assume FIFO could also be used effectively in the more complex LIFO case.

Furthermore, Miranda and Kónya (2006) argue that consumers' approach to label and packaging information is different when purchasing perishables or non-perishables. Customers will check the best-before-date of a perishable product more often than of a non-perishable product. When purchasing perishables, consumers will respond to a short best-before-date of their preferred perishable product by switching to another brand, but are not prepared to buy a smaller pack of their preferred brand, even though they could replenish their stock in the next visit to the store. Miranda and Kónya (2006) argue there is no evidence that shoppers change their preferred items when they do not have enough available consumption time, even if prices are discounted, as this variable is not significant in influencing the estimated probability of examining the best-before-date.

Hoch and Deighton (1989) argue that the majority of consumer decision making occurs in the store. Dagnoli (1987) argues that only about 1/3 of the purchases are specifically planned in advance of visiting the store. Hoyer (1984) argues that consumers show a low level of involvement with most of the in-store decisions. They make their choices very quickly after minimal search and price comparison (Dickson and Sawyer, 1990).

Van Donselaar et al. (2012) argue that many consumers only visit a supermarket once a week. They will buy the items they intend to consume in the coming week. This behavior requires a remaining shelf life of at least a week. When such consumers easily can check the expiration dates on the

products and can find a product with a later expiration date, they will probably select this later product. This will result in LIFO withdrawal behavior.

1.4.5 Observed consumer behavior

This chapter focusses on the actual consumer behavior in supermarkets. Van Burgh (2007) conducted a field study about the grabbing behavior of consumers in two supermarkets in The Netherlands. We define grabbing behavior as: the number of bought products by consumers with a later best-before-date than the first product in the shelf with a shorter best-before-date as a percentage of the total sold products given that there are two or more best-before-dates in the shelf.

Van Burgh (2007) found significant differences of the grabbing behavior between stores. In one of the stores of his research he found 30 percent grabbing behavior, in another one 49,7 percent. He argues that the grabbing behavior is dependent on the space between shelves. Smaller space between the shelves makes it more difficult for the consumers to grab. The grabbing behavior also differs between product categories. For some product groups consumers are more critical about the best-before-date than for others. Furthermore, Van Burgh (2007) argues that the larger the number of items in the shelf, the less grabbing behavior occurs. At last, he argues that when the first product approaches its best-before-date grabbing behavior will increase.

Research by Broekmeulen and Van Donselaar (2014) shows that grabbing behavior is different between large and medium size supermarkets. They also show that grabbing behavior is dependent on the product category. They support Van Burgh (2007) by stating that grabbing behavior will increase when the first batch in the shelf is older. They further argue that sales rates will increase when supermarkets offer several batches, because of more choice for the customers. They argue that decreasing case pack sizes or minimal order quantities can reduce outdating at supermarkets.

Both Van Burgh (2007) and Broekmeulen and Van Donselaar (2014) state that the withdrawal behavior at supermarkets is a combination of FIFO and LIFO. Haijema (2011) supports this opinion and argues that the consumer behavior can be a combination of FIFO and LIFO, according to the customer's preference for a product of a specific age. Empirical research by Broekmeulen and Van Donselaar (2014) shows that the choice for FIFO or LIFO withdrawal is not randomly done by customers, but is dependent on their preferences. They argue that customers follow a satisficing choice model.

1.4.6 Pricing of perishables

Ferguson and Koenigsberg (2007) argue that retailers can mark down prices of items when the customers' perception of the quality of the leftover items is lower than that of their new items. They argue that how a firm should stock and price perishables depends on the characteristics of the perishable product and the consumer's perceived difference in quality between old and new item.

Ferguson and Koenigsberg (2007) classify three types based on the perceived quality level of the aged product by the consumer. Type 1 products have a constant perceived quality, but become unusable after a given date. The perceived quality level of type 2 products deteriorates continuously over time reaching a value of zero when a new version of the item becomes available. Finally, the perceived quality of type 3 products deteriorates over time, but does not reach a value of zero when a replenishment of new items arrive. However, the deterioration ensures that the customer values an older product lower than a newer one, thus retailers selling type 3 products must use price

differentiation between old and new products. This type 3 products can be divided in products who actual functionally deteriorate over time such as fresh products and products whose functionality does not degrade, but the customers' perceived utility of the product deteriorates over time.

Ferguson and Koenigsberg (2007) argue that carrying over a percentage of the old items is beneficial compared to never-carry or carry-all when (1) uncertainty over the market potential is high, (2) the cost to prepare the carried over unit for the market is low compared to the cost to purchase new units, and (3) the quality degradation of the unsold units is low. For most perishables in Dutch supermarkets the quality reduction of products per day is quite small. Only for items like bread the perceived quality reduction is so high that the perishables are not sold more than one day. For the other perishables in a supermarket, the cost of carrying old products is much lower than the purchasing value of a new item. Research at Dutch supermarkets showed that even when the perishables are sold with a discount the retailer still earns more with selling them than disposing them and selling only new products (Van de Ven, 2014). Furthermore, Nijs et al. (2001) prove that short- and long-run effectiveness of price promotions is greater for perishable goods than for other categories, so discounting aging perishables seems to work.

Tsiros and Heilman (2005) argue that managers should use different discount percentages for different categories as different categories involve different levels of product quality risks. For example the product quality risks for beef and chicken are quite high compared to other perishables and therefore managers should consider greater discounts to compensate for the greater risks associated with these products if they want to sell aging inventory. Discounting may not be as crucial for perishables with low product quality risks. Further, they argue that managers should weigh for each product category the trade-offs between the potential benefits of discounting to sell inventory and its potential negative effects on store image. Currently there is not sufficient literature to judge the effect of discounting on store image.

In contrast to Tsiros and Heilman (2005), Miranda and Kónya (2006) argue there is no evidence that shoppers change their preferred items when they do not have enough available consumption time, even if prices are discounted, as this variable is not significant in influencing the estimated probability of examining the best-before-date. This result came from a structured questionnaire among 473 randomly selected grocery shoppers in Australia. The empirical research of Tsiros and Heilman (2005) was conducted in the United States and maybe the perceptions on best-before-dates are different between countries.

Dutch supermarket managers have different pricing strategies for 'old' products. At some stores like AH products with a remaining shelf life of one day are marked down by 35%. At PLUS the mark down percentage differs between stores, for example some use 25%, others 30% and others 35%. Finally at other supermarkets 'old' products are disposed from the shelves and not marked down like at Jumbo.

1.4.7 Implications of consumer behavior on outdating

From the preceding chapters it is clear that the consumer behavior in supermarkets is not the same for each store. Different behavior leads to a different amount of outdating and shortages. Changing the ordering policy can reduce the amount of outdating and shortage. However, changing the order policy and thus the ASO is a time-consuming and costly process. Haijema (2011) states that "without

changing the order policies of such systems one may reduce outdating and shortages by issuing products in a sophisticated way.” (Haijema, 2011).

LIFO withdrawal by customers will always result in the same or a higher amount of outdating than FIFO withdrawal, because with LIFO withdrawal customers refuse to take older products and take newer products instead. Simulation results of Van Donselaar and Broekmeulen (2012) show that LIFO withdrawal will result in significant more outdating of perishables than FIFO withdrawal. However, empirical research of Broekmeulen and Van Donselaar (2014) at two Dutch supermarkets shows that in their setting LIFO withdrawal has a very limited effect on outdating. In their setting the sales rates increased when there were multiple batches and this results in less outdating and less lost sales. Note that LIFO behavior can only occur when there are multiple batches in the shelves.

1.4.8 Consequences of changing the case pack size

One solution to decrease the amount of outdating of perishables at Dutch supermarkets is optimizing the case pack size (Van de Ven, 2014). Changing the case pack size as Van de Ven (2014) suggests can have several consequences. One of the expected consequences is that the inventory of products will decrease when the case pack sizes are decreased. This will probably result in lower outdating. However, several investigations (i.e. Whitin 1957; Wolfe, 1968) show that a large inventory of a product will increase sales. A large inventory will increase the visibility of a product, gives confidence to the customer and gives the perception of being a popular product. So, we have to check if lower inventory levels have an influence on the demand for products. Van Donselaar and Broekmeulen (2012) and Weteling (2013) show with simulation that having a case pack size close to the mean demand during the product lifetime will result in high outdating. So, lowering the case pack size in these cases will result in lower outdating. Another consequence of changing the case pack size, as Van de Ven’s (2014) model suggests for several products, is that we will have a higher probability of products with different best-before-dates in the shelf. We have to check which consequences this has on the withdrawal behavior of customers.

One of the consequences of having different batches at the same time in the shelf is substitution between batches. Deniz et al. (2004) argue that suppliers can choose four different substitution policies for perishables with the implicit consent of the customer. These policies can also be viewed from a customer’s point of view. No-Substitution or Downward-Substitution are the most likely substitutions to happen. Some customers don’t want to make a substitution and for the customers who are willing to there is no problem to buy newer(fresher) products than they intended to buy. Full-Substitution is not very likely to happen as customers will probably not accept older products when they want products with a particular remaining shelf life. The exact amount of customers following each policy has to be determined in further research and is dependent on the category of the perishables (Van Woensel et al. 2007).

The previous paragraph described age-based substitution. Deniz et al. (2010) argue that next to age-sensitive customers, there are price-sensitive customers. These price-sensitive customers can be induced to purchase older products when they are marked down. Retailers mark down those products if their customers’ perception of the quality is lower than that of their new items. The consequence of the higher probability of products with different best-before-dates in the shelf on the amount of products that has to be marked downed is not known. We expect lower outdating, so

we also expect less products that has to be marked down. However, this has to be tested in real supermarkets.

As showed in this chapter, offering several batches at the same time will result in substitution effects and thus in demand diversion. It is important to note that retailers also have the choice to eliminate this diversion by not co-locating goods of different ages. In this way substitution is not possible for the customers. However, this choice requires that retailers take out all old products when they place a new batch in the shelf. This can result in a lot of outdateding.

1.4.9 Ways to force consumers to use FIFO instead of LIFO

Among others, Nahmias (2011) and Van Donselaar and Broekmeulen (2012) argue that FIFO withdrawal minimizes outdateding in the standard perishable inventory problem. According to Ferguson and Ketzenberg (2005) practitioners are well aware of the value in controlling inventory issuing with a FIFO issuing policy. As Haijema (2011) argues, the manager of a stock point often has some way to control the withdrawal process. Next to this, Deniz et al. (2010) argues that when different classes of customers have specific preferences for products of different ages, replenishment and issuance decisions should be made jointly.

When we notice LIFO or mixed withdrawal behavior at a supermarket we can think of ways to force consumers to use FIFO withdrawal. Stimulating FIFO behavior can be done by for example making the distances between the shelves smaller in order to make it more difficult for consumers to grab or replenishing the shelves just before an out of stock in order to prevent several batches (with different best-before-dates) in the shelf. In situations with only one batch in the shelf, consumers have no choice. Research by Van Burgh (2007) shows that this specific way of replenishment at four Dutch supermarkets already prevented 65 percent of the possible grabbing behavior. Furthermore, Ferguson and Ketzenberg (2005) argue that retailers make extensive investments to force customers to use FIFO withdrawal like installing rear-loading shelving systems and gravity wells in addition to extensive training and labor expenditures to ensure that perishables are continuously rotated. They argue that switching from LIFO or mixed withdrawal to FIFO issuing is generally more profitable than from sharing information.

1.5 Gaps in literature

The first gap in literature is the lack of investigations in which the expected causes of waste are linked with real data from supermarkets. Most literature focuses on perceived effects and uses simulations with assumptions about fill rates and shelf lives instead of real data. In this master thesis project real data will be investigated to see if the expected causes of waste really exist in Dutch supermarkets. Next to this more research is necessary about the influence of discounting product on the amount of outdateding of perishables. This project investigates whether discounting products has an influence on the amount of outdateding of perishables or not.

The second gap in literature is the lack of investigations in which the opinion of supermarket managers and entrepreneurs on outdateding of perishables is investigated. Only Gunders (2012) argues that the retail model in the US views waste as a part of doing business, but is it not sure whether this also holds for retailers in The Netherlands.

By combining real data with opinions from supermarkets employers and employees, and observations at several supermarkets, this project will add valuable insights to the current literature.

2. Problem Definition

This chapter will describe the problem situation of the outdateding of perishables at PLUS supermarkets in more detail. First the motivation for the project is given. Second, the research question and sub-questions will be described. Third, the research methodology and the hypotheses are explained. Finally the scope and the involved stakeholders of the problem will be described.

2.1 Motivation for project

The customers of today expect a wide range of perishable and non-perishable products in supermarkets. They want them to be of good quality and fairly priced. For the perishable items, good quality means fresh. Due to the short shelf life of perishable products and the stochastic demand of customers, the retailers of perishable products struggle with the outdateding of products. Because HB is the logistic service provider of PLUS Retail for most perishable items and most outdateding occurs with those perishable items, they want to help PLUS Retail to reduce the amount of outdateding in the PLUS supermarkets where they sent their products to.

In 2014 a student from the Eindhoven University of Technology was asked to investigate a way to reduce the amount of outdateding of perishables. In his master thesis, Van de Ven (2014) investigated the optimal case pack sizes for convenience products at PLUS supermarkets in order to reduce the amount of outdateding at the supermarkets. His objective was to reduce the cost of outdateding at the supermarkets of PLUS Retail by optimizing the case pack sizes for convenience products at HB, considering entailed additional supply chain operating costs. By using data collection from observations at HB and several PLUS supermarkets and actual sales data he developed a mathematical model to determine the total yearly costs in the supply chain. He combined his model with a model from Broekmeulen & Van Donselaar (2009) that identifies the optimal ordering quantity taking into account the age of the inventory.

Van de Ven (2014) found significant cost savings when the current inventory policy would be changed and the case pack sizes would be optimized. The operating costs at both HB and the supermarkets will increase, but the cost of outdateding will decrease much more. This will result in lower supply chain costs.

Both HB and PLUS Retail are enthusiastic about the results of Van de Ven's (2014) research. They want to conduct a follow-up study based on his results. Van de Ven's (2014) results are based on a mathematical model that is trying to approach the real world. Using a mathematical model requires making some assumptions about the real world. HB is interested in a closer approach to the real world to get more insight in the possible cost savings. There are several ways to get a closer approach to the real world. In the following chapters our approach for the desired follow-up will be explained.

2.2 Research question

Based on the results of Van de Ven (2014), HB wants to have a follow-up study conducted about the causes of the actual outdateding of perishables in PLUS supermarkets. The size and causes of the actual outdateding are not known yet. Observations at several PLUS supermarkets show that there are several differences between the supermarkets. Getting insight in these differences is important because these differences can have an impact on differences in the amount of outdateding of perishables in

each supermarket. HB wants to have investigated which in-store factors have an influence on the outdating of perishables. This leads to the following research question:

Which causes lead to outdating of perishables at PLUS supermarkets and how can the amount of outdating be reduced at those supermarkets?

In order to answer this question two phases for research have been determined: (i) analysis from interviews, and (ii) analysis from data. The sub-questions for these phases will now be stated.

2.3 Sub-questions

2.3.1 Analysis from interviews

1. What is the opinion of supermarket managers on outdating?
2. What is the opinion of supermarket managers on selling perishables with a short remaining shelf life?
3. What is the opinion of supermarket managers on discounting perishables with a short remaining shelf life?

2.3.2 Analysis from data

1. Which factors lead to outdating on supermarket level?
2. What is the reason of the differences in those factors between PLUS supermarkets?
3. Which of those factors can be changed by supermarket managers, PLUS Retail, HB or are not changeable?

4. Which factors lead to outdating on product level?
5. What is the reason of the differences in those factors between products?
6. Which of those factors can be changed by supermarket managers, PLUS Retail, HB or are not changeable?

7. Which factors lead to outdating in supermarket – product combinations?
8. What is the reason of the differences in those factors between combinations?
9. Which of those factors can be changed by supermarket managers, PLUS Retail, HB or are not changeable?

Question 1,2 and 3 of this paragraph will be answered in Analysis 1 and Analysis 4. Question 4,5 and 6 will be answered in Analysis 2. Question 7,8 and 9 will be answered in Analysis 3.

2.4 Research methodology

Several methods are available to get an answer on the research question and the sub-questions. The project is started by conducting semi-structured interviews with supermarket entrepreneurs, supermarket managers, department managers and store clerks. Entrepreneurs and managers will be interviewed, because they have insight in the total outdating of their store. Department managers will be interviewed, because they have insight in the total outdating of their department. Next to this, they are responsible for a good result on waste of products, next to revenue, margin and wages. Store clerks will be interviewed, because they fill the shelves and maybe also have insight in the grabbing behavior of customers.

The data will be analyzed by using MS Excel and Access, and IBM SPSS. Several regression analyses will be performed to find significant relations between variables. Data will be gathered at HB, PLUS Retail, PLUS supermarkets and from online data sources.

Several factors will be investigated to find possible relations between those factors and the amount of outdating of perishables. The factors with a possible impact on the outdating of perishables are coming from the literature as described in chapter 1.4 and from interviews with supermarket employees. Several factors that could have an influence of the outdating of perishables could not be taken into account due to unavailability of data or inaccuracy of data. These factors are described in chapter 2.4.5.

To be able to answer all sub-questions, regression analyses will be performed on three different aggregate levels. We will start with two high aggregate levels and continue with a low aggregate level. Level 1 will be an analysis on supermarket level where data about all products sold in a specific supermarket is combined. The same aggregate level will be used for an extra analysis per supermarket for the supermarkets that were visited during the project. Level 2 will be an analysis on product level where data about all supermarkets selling a specific product will be combined. Level 3 will be an analysis on supermarket product combination where data about a specific product supermarket combination is combined. All analyses will be performed two times; one time for the convenience category and one time for the meat products and salads category.

2.4.1 Analysis 1: analysis on supermarket level

The variables that will be used in the analysis on supermarket level, analysis 1, are:

1. Percentage outdating of total revenue ($perc_outdating_supermarket_{c,s}$)
2. Average order size ($avg_o_supermarket_{c,s}$)
3. Average coefficient of variation of order size ($avg_coefvar_o_supermarket_{c,s}$)
4. Average case pack size ($avg_Q_supermarket_{c,s}$)
5. Category revenue share ($crs_{c,s}$)
6. Average sales ($\mu_supermarket_{c,s}$)
7. Average coefficient of variation of sales ($avg_coefvar_sales_supermarket_{c,s}$)
8. Percentage of order advises that is followed ($oaf_{c,s}$)
9. Number of opening days per week (od_s)
10. Average disposable income in neighborhood of supermarket ($avgdi_s$)
11. Sales area (verkoopvloeroppervlakte in Dutch) (sa_s)
12. Average service level of HB ($avgserv_{c,s}$)
13. Average remaining shelf life ($avg_sl_supermarket_{c,s}$)
14. Natural logarithm of average remaining shelf life ($LN_avg_sl_supermarket_{c,s}$)
15. Supermarket QMuM ratio ($Q_mu_m_supermarket_{c,s}$)

The variables are described in more detail in Appendix B.

For the convenience and meat products and salads categories we have respectively 230 and 233 supermarkets for this analysis with 14 independent variables. Variable 14 is only a transformation of variable 13, so we don't have to count this variable as an extra variable for our analysis. So, we have respectively 17,69 and 17,92 observations per independent variable. The sample size required depends on the size of the effect and how much statistical power we want to detect these effects

(Field, 2009). Several rules of thumb exist about the minimal sample size. Most of them argue that more than 17 observations per independent variable is acceptable, but that we have to be careful about the outcomes of the regression analysis, because of the chance of overfitting. Overfitting means that when we have a lot of predictors and a limited number of samples, random sampling fluctuations will allow some linear combination of the predictors to match the predictand perfectly over the limited samples we have, but the correlations will disappear for a different set of samples. So, we can continue with the 14 independent variables in analysis 1, but have to be careful about the outcomes.

2.4.2 Analysis 2: analysis on product level

The variables that will be used in the analysis on product level, analysis 2, are:

1. Percentage outdated of total revenue (*perc_outdating_product_p*)
2. Average order size (*avg_o_product_p*)
3. Average coefficient of variation of order size (*avg_coefvar_o_product_p*)
4. Case pack size (*Q_product_p*)
5. Average sales per day (*μ_product_p*)
6. Average coefficient of variation of sales per day (*avg_coefvar_sales_product_p*)
7. Average remaining shelf life (*avg_sl_product_p*)
8. Natural logarithm of average remaining shelf life (*LN_avg_sl_product_p*)
9. Product QMuM ratio (*Q_μ_m_product_p*)

The variables are described in more detail in Appendix B.

For the convenience and meat products and salads categories we have respectively 224 and 276 products for this analysis with 8 independent variables. Variable 8 is only a transformation of variable 7, so we don't have to count this variable as an extra variable for our analysis. So, we have respectively 32 and 39,43 observations per independent variable. As described before, the sample size required depends on the size of the effect and how much statistical power we want to detect these effects (Field, 2009). Field (2009) describes that more than 224 observations for 7 predictors is enough when we expect large or medium effects. Maybe we are not able to find small effects with this number of observations.

2.4.3 Analysis 3: analysis on supermarket – product combination

The variables that will be used in the analysis on supermarket – product combination, analysis 3, are:

1. Percentage outdated of total revenue (*perc_outdating_combination_{p,s}*)
2. Average order size (*avg_o_combination_{p,s}*)
3. Coefficient of variation of order size (*coefvar_o_combination_{p,s}*)
4. Case pack size per product (*Q_combination_{p,s}*)
5. Average sales per day (*μ_combination_{p,s}*)
6. Coefficient of variation of sales per day (*coefvar_sales_combination_{p,s}*)
7. Average remaining shelf life (*avg_sl_combination_{p,s}*)
8. Natural logarithm of average remaining shelf life (*LN_avg_sl_combination_{p,s}*)
9. Combination QMuM ratio (*Q_μ_m_combination_{p,s}*)

The variables are described in more detail in Appendix B.

For the convenience and meat products and salads categories we have respectively 38794 and 47794 combinations for this analysis with 8 independent variables. Variable 8 is only a transformation of variable 7, so we don't have to count this variable as an extra variable for our analysis. So, we have respectively 5542 and 6827,71 observations per independent variable. As described before, the sample size required depends on the size of the effect and how much statistical power we want to detect these effects (Field, 2009). Field (2009) describes that more than 38000 observations for 7 predictors is enough when we expect small, medium or large effects. So, the number of observations is large enough for analysis 3.

2.4.4 Analysis 4: extra analysis on supermarket level

Next to the above mentioned three analyses, a fourth analysis will be conducted only for the supermarkets that will be visited. For this analysis the extra data will be used that is gathered at the supermarket visits. We start with the significant variables from analysis 1 and add the six new variables to the model.

The variables that will be used in the extra analysis on supermarket level, analysis 4, are:

1. Percentage outdated of total revenue ($perc_outdating_supermarket_{c,s}$)
2. Value of significant variables and constant from analysis 1
3. Average distance in centimeter between top of a product and next shelf ($avgdistance_{c,s}$)
4. Presence of springs in the shelves ($springs_{c,s}$)
5. Discount percentage ($dp_{c,s}$)
6. Number of hours of being able to sell discounted products ($dh_{c,s}$)
7. Selling products with one day remaining shelf life ($productstoday_s$)
8. Supermarket manager type ($manager_s$)

The variables are described in more detail in Appendix B.

For the convenience and meat products and salads categories we have 34 supermarkets for this analysis with 6 new independent variables. So, we have 5,66 observations per variable. As described before, the sample size required depends on the size of the effect and how much statistical power we want to detect these effects (Field, 2009). Several rules of thumb exist about the minimal sample size. All of them argue that 34 observations for 6 predictors is too low. This is because of the chance of overfitting. A common approach that controls overfitting by keeping the number of independent variables to a minimum is using the stepwise method. This is the method we will use for analysis 4.

2.4.5 Left out variables

Some other factors can also have an influence on the amount of outdated of perishables in supermarkets. These are: type of customer, number of supermarkets in neighborhood of supermarket, shelf meters for perishables, service levels, average inventory level, place of the product in the shelf, quality of replenishment process, quality of ordering process, storing conditions in the supermarket and transport, quality of products at arrival and the amount of theft per supermarket.

Customers at PLUS are classified in customer groups. The customers in different customer groups will have different preferences for perishables and this will result in different buying patterns. For example, customers with a big family will buy bigger packages than single households. The

percentage of customers per customer group per supermarket is confidential and could not be used for this project. So, we need to leave these groups out of our analyses.

Traditionally consumers buy products in several supermarkets in their neighborhood. Having more than one supermarket in a neighborhood increases the chance that a consumer will buy products in more than one supermarket. By buying perishables during promotion at another store, the promotions of competing supermarkets will have an impact on the sales in the supermarket. This will probably result in higher outdating of the perishables that are promoted in nearby supermarkets. The number of supermarkets in the catchment area of each PLUS supermarket is confidential and could not be used for this project. So, we need to leave this factor out of our analyses.

PLUS supermarkets can buy their own refrigerators and because of this they decide how many meters of the shelf space is available for perishables. When the number of meters of shelf space is too much for the revenue in perishables, outdating will occur due to a low turnover per meter. The shelf space for perishables of each PLUS supermarket is confidential and could not be used for this project. So, we need to leave these factor out of our analyses.

We do know the sales area of each supermarket. However, the sales area cannot be seen as an alternative for the meters of refrigerators as there is no strict link between the sales area and the meters of refrigerators. Managers decide themselves how many meters of refrigerators they want in their supermarket.

Several service levels can be used by retailers to measure their performance towards customers. For example, the fill rate, measures the proportion of total demand within a reference period that is delivered without delay. When setting a higher desired fill rate for products, the probability of out of stocks will decrease, but the probability of outdating of products will increase. The settings for the service levels will have an influence on outdating, but these settings are not available for the PLUS supermarkets. So, we need to leave these levels out of our analyses.

Having a high inventory level compared to the average sales will result in a higher amount of outdating. A relatively high inventory level is likely to be found in stores with low average demand per day for perishables. They will place more products in the shelves than they should do for these products because they want a good store image. As the parameters in the automated store ordering system are not available, the average inventory level in each supermarket cannot be calculated. Interviews show that supermarkets change a lot in the initial settings of the automated store ordering systems, so making assumptions about the settings is not possible. So, we need to leave these levels out of our analyses.

The place of the product in the shelf will have an influence on outdating, because products on the lowest or highest shelf are less visible for consumers and this will result in lower demand. This lower demand will probably result in more outdating. However, changing the location of a product with a lot of outdating is not that easy, as the shelf management is done with a focus on revenue and margin. Replacing a product to a better location will probably result in decreased demand for another product and this is not desirable. Next to this, the location of products differs a lot in different supermarkets as most supermarket managers don't follow the shelf plans made by PLUS Retail. So, we have to leave this factor out of our analyses.

The quality of the replenishment process also has an influence on the amount of outdating. The store clerks have to fill the shelves according to FIFO, but sometimes they fill according to LIFO and this will result in higher outdating. The quality of this process is hard to judge as in the supermarkets a lot of mostly part-time store clerks are working. Next to this, one can assume that store clerks that don't do their job according to their job requirements will be fired or their work contract will not be extended. So, a bad quality of the replenishment process will be solved in quite a short time. Because of this, this factor will be left out of the regression analyses.

The quality of the ordering process also has an influence on the amount of outdating. Ordering too much will result in outdating. The quality of the ordering process is dependent on the store clerk with this task. The quality of this process is hard to judge and will therefore be left out the regression analyses. Note that we do know the percentage of order advises that is followed by a supermarket in a category.

The storing condition of the perishables in the store and at transport can also have influence on the amount of outdating. Especially for vegetables and fruit the temperature and humidity of the storing environment has a big influence on the decay process of the product. As it is not possible to judge the storing conditions at each fruit and vegetables department, each backroom and at each transport, this factor will be left out the regression analyses.

The quality of products when they arrive at the supermarkets also has an influence on the amount of outdating at supermarkets. This especially holds for vegetables and fruit as these products are very sensitive to decay during transport or storage. When the perishables have a bad quality when they arrive at the supermarket, some products already have to be thrown away before they can even be sold or they have to be sold in a short time. This will result in higher outdating. As all PLUS supermarkets get their perishables from the same distribution center we assume that the quality of products at arrival doesn't differ between supermarkets. Therefore this factor will be left out of the regression analyses.

The waste of products is not always placed in the right group of waste and some supermarkets don't register the amount of stolen products. Because of this, perishables that are stolen will incorrectly be registered as outdated products or stock differences. Supermarkets with a high amount of theft will have a higher amount of waste, but this effect cannot be analyzed as there is no accurate data about this factor. Therefore this factor will be left out of the regression analyses.

2.5 Hypotheses

Some hypotheses can be formulated about the effect of the factors on the amount of outdating of perishables.

2.5.1 Hypotheses for analysis 1: analysis on supermarket level

Hypothesis 1: *The higher the average coefficient of variation of order size of a supermarket, the higher the amount of outdating.*

When supermarkets have a high average coefficient of variance of order size, they experience strongly fluctuating sales or order products based on emotion or intuition without statistical basis. Both factors will result in a higher chance of overstocking and therefore outdating.

Hypothesis 2: *Having a small revenue share of a category compared to the total revenue of all perishables will result in a higher amount of outdating.*

Having a small revenue share of a category compared to the total revenue of all perishables implies that one has a category that is not so popular compared to the total product group of perishables. The supermarket will however have at least a base assortment of perishables for each category and will have higher outdating of the perishables in this specific category, because the average demand in this category is quite low.

Hypothesis 3: *The higher the average sales per product per supermarket per day, the lower the amount of outdating.*

Higher average sales implies higher sales for the products in the supermarket and thus more popular products. Being a popular product implies a higher probability that all items will be sold before they expire.

Hypothesis 4: *The higher the percentage of order advises that is followed, the smaller the amount of outdating.*

Following the order advises of the automated store ordering system will prevent that orders are based on emotion or intuition without statistical basis. In this way a higher percentage of order advises that is followed will result in a smaller amount of outdating.

Hypothesis 5: *A supermarket that is open 7 days a week will have a smaller amount of outdating than a supermarket that is open 6 days a week.*

Being open 7 days a week will result in more potential to sell perishables than being open 6 days a week. This increased potential will result in a smaller amount of outdating.

Hypothesis 6: *The higher the average disposable income of consumers in the neighborhood of a supermarket, the lower the amount of outdating of perishables.*

Having a higher disposable income results in a larger group of affordable perishables for consumers. Consumers that have a low disposable income will not have the opportunity to buy relatively expensive perishables. Supermarkets in poor or rich neighborhoods will have more or less the same assortment, but the supermarkets in poor neighborhoods will not sell a lot of expensive perishables. This will result in a higher amount of outdating in supermarkets located in poor neighborhoods.

Hypothesis 7: *Supermarkets with a small sales area will have more outdating than supermarkets with a large sales area.*

Supermarkets with a small sales area will have more outdating than supermarkets with a large sales area, because they will attract less consumers and have a relatively low revenue while having still quite a broad assortment of perishables.

Hypothesis 8: *The higher the service level of HB for a supermarket, the higher the amount of outdating.*

The service level of HB for a supermarket shows the percentage of case packs that are delivered compared to the ordered case packs. The supermarkets that will get their goods in the afternoon have on average a lower service level, because their products will be picked at the warehouse at the end of the picking process and some product get out of stock during the picking process. Supermarkets with a lower service level will have more empty shelves and less inventory in the

supermarkets. This lower inventory will result in less product to be sold and thus in less products that can outdate.

Hypothesis 9: *The higher the average remaining shelf life per supermarket, the lower the amount of outdating.*

The remaining shelf lives of products when they are placed in shelves also has an influence on the amount of outdating at supermarkets. All PLUS supermarkets get their perishables from the same distribution center and most perishables are not stored in the backroom of the supermarkets. However, the delivery frequency of perishables to the distribution center of HB differs a lot depending on the product. For example, some products are delivered two times a day and other products are delivered only once a week. Next to this, not all supermarkets order products with same frequency. Stores with a high revenue will have a higher order frequency. These reasons will result in different remaining shelf lives of products that arrive at different supermarkets. Stores with a higher average remaining shelf life will have less outdating of the perishables.

Hypothesis 10: *The higher the supermarket QMuM ratio, the higher the amount of outdating.*

The supermarket QMuM ratio calculates per supermarket the average ratio between the case pack sizes and the average sales during the shelf lives of the products of a supermarket. When this ratio is low, the case pack sizes of the products of a supermarkets are relatively small compared to the sales during the shelf lives of the products. This means that there is a high chance that the case packs will be sold within the lifetime of the products. Having a high ratio means that the case pack sizes is relatively large compared to the sales during the shelf lives of the products. This will result in a higher chance of left over products for a case pack and thus in a higher amount of outdating. Supermarkets with a low ratio will have a lower chance of outdating for their assortment.

For the average order size and the average case pack size no hypotheses are drawn because of the fact that the average order size and the average case pack size on their own don't have a lot of value for predicting the outdating in a supermarket. For example, a large case pack size only results in more outdating when the average sales are low and/or the shelf life of the products is short. For the average coefficient of variation of sales no hypothesis can be drawn. When a certain coefficient of variation of sales is expected due to the week pattern it doesn't have to result in outdating as orders are based on these expectations.

2.5.2 Hypotheses for analysis 2: analysis on product level

Most of the hypotheses for the analysis on supermarket level will hold as well for the analysis on product level. Hypothesis 1, 3, 9 and 10 are slightly changed to be useful for the analysis on product level.

Hypothesis 11: *Having a high average coefficient of variation of order size per product will result in more outdating.*

When products have a high average coefficient of variance of order size, products experience strongly fluctuating sales or products are ordered by supermarkets based on emotion or intuition without statistical basis. Both factors will result in a higher chance of overstocking and therefore outdating.

Hypothesis 12: *The higher the average sales per product per supermarket per day, the lower the amount of outdating.*

Higher average sales implies higher sales for the product in supermarkets and thus a more popular product. Being a popular product implies a higher probability that all items will be sold before they expire.

Hypothesis 13: *The higher the average remaining shelf life per product, the lower the amount of outdating.*

Products with a long shelf life will have more opportunity to be sold during their lifetime than products with a short shelf life. Because of this, less outdating is expected for products with a long average remaining shelf life.

Hypothesis 14: *The higher the product QMuM ratio, the higher the amount of outdating.*

The product QMuM ratio calculates per product the ratio between the case pack size and the average sales during the shelf lives of a product in all supermarkets. When this ratio is high, the case pack size is relatively small compared to the sales during the shelf lives of the product in all supermarkets. This means that there is a high chance that the case packs will be sold within the lifetimes of the products. Having a low ratio means that the case pack size is relatively large compared to the sales during the shelf lives of the product in all supermarkets. This will result in a higher chance of left over products for a case pack and thus in a higher amount of outdating. Products with a low ratio will have a lower chance of outdating compared to other products.

As explained in the previous chapter, no hypotheses are drawn for the average order size, the case pack size and the coefficient of variance of sales.

2.5.3 Hypotheses for analysis 3: analysis on supermarket - product combination

Most of the hypotheses for the analyses on supermarket and product level will hold as well for the analysis on supermarket - product combination. Hypothesis 11, 12 and 13 hold in the same way for the analysis on supermarket – product combination. Hypothesis 14 is slightly changed to be useful for the analysis on supermarket - product combination.

Hypothesis 15: *The higher the combination QMuM ratio, the higher the amount of outdating.*

The combination QMuM ratio calculates per supermarket – product combination the ratio between the case pack size and the average sales during the shelf lives of a product in a supermarket. When this ratio is high, the case pack size is relatively small compared to the sales during the shelf lives of the product in a supermarket. This means that there is a high chance that the case packs will be sold within the lifetimes of the products. Having a low ratio means that the case pack size is relatively large compared to the sales during the shelf lives of the product in a supermarket. This will result in a higher chance of left over products for a case pack and thus in a higher amount of outdating. Combinations with a low ratio will have a lower chance of outdating compared to other combinations.

As explained in the chapter 2.5.1, no hypotheses are drawn for the average order size, the case pack size and the average coefficient of variance of sales.

2.5.4 Hypotheses for analysis 4: extra analysis on supermarket level

Hypothesis 16: *The smaller the distance in centimeter between the top of a product and the next shelf, the smaller the amount of outdating.*

As known from literature, grabbing behavior exists in the product group of perishables. Grabbing behavior will result in higher outdating, as consumers will pick products with a longer remaining shelf life and don't follow the FIFO withdrawal policy that should be followed to minimize outdating. Decreasing the distance in centimeter decreases the opportunity for consumer to grab. This will result in a smaller amount of outdating.

Hypothesis 17: *The presence of a spring in the shelves will result in a lower amount of outdating.*

In some supermarkets springs are present in the shelves to avoid grabbing behavior of customers and to ensure well-presented products. Having springs will result in lower grabbing behavior and thus in a lower amount of outdating of perishables.

Hypothesis 18: *The higher the discount percentage, the lower the amount of outdating.*

A higher discount percentage will stimulate more consumers to buy the discounted product than a lower percentage. In this way, a higher percentage will result in a lower amount of outdating.

Hypothesis 19: *The higher the number of hours of being able to sell discounted products, the lower the amount of outdating.*

A higher number of hours of being able to sell discounted products increases the opportunity of selling discounted products. This holds for regular customers as for customers who come specific to a supermarket to buy discounted products. The higher opportunity will result in more discounted products that will be sold and in this way in less outdating.

Hypothesis 20: *Supermarkets selling products with one day remaining shelf life will have less outdating than supermarkets that don't sell these products.*

The selling period in supermarkets that don't sell products with one day remaining shelf life is one day shorter than the selling period in the supermarkets that do sell these products. This smaller selling period will result in more outdating as products have to be removed from the shelves earlier.

Hypothesis 21: *Supermarkets classified as grocer will have less outdating than supermarkets classified as entrepreneur.*

As will be explained in chapter 3 grocers and entrepreneurs have a total different opinion about outdating. As grocers typically want to get the lowest percentage outdating as possible, whereas entrepreneurs take outdating as a fact of doing business, grocers are expected to have less outdating.

2.6 Scope of research

To ensure that the project is possible in the available six months an appropriate scope has to be determined. As described earlier we will only focus on perishable items.

2.6.1 Selection of categories

The perishable products at HB can be classified into fourteen different categories:

1. Potatoes, fruit and vegetables;
2. Precut vegetables;
3. Convenience products;
4. Sandwiches;
5. Cheese;
6. Meat;

7. Meat products and salads ZB;
8. Meat products and salads AV/AVA;
9. Fish;
10. Bulk dairy products;
11. Dairy products;
12. Yellow fats;
13. Pastry;
14. Plants and flowers;

Some products are sold according to their weight and are cut in the supermarkets. These products can be found in category 5 and 8. As not all the supermarkets get the same products and the data of these products is difficult to analyse, these product will be left out of the master thesis project. Fish (category 9) will be left out the analysis as these products are very sensitive to seasonal demand. Pastry (category 13) will be left out of the project as some supermarkets order them at a local supplier instead of at the cross-dock supplier. Plants and flowers (category 14) will be left out the analysis as these are non-food products.

The products in category 2, 4, 9, 13, 14 and some of the products in category 5 and 6 are coming from cross-dock suppliers. At HB less information is available about those products. From this products only the minimum remaining shelf life at arrival at the supermarket is known. Because of this, these categories will be out the analyses.

After these decisions, categories 1, 3, 7, 10, 11 and 12 are still available for this project. As can be seen in Figure 2 below, category 1 (potatoes, fruit and vegetables) has from this remaining categories the highest percentage of waste. However, in the data of PLUS Retail the data of category 1 and 2 of HB is combined. Next to this, the registration of sales and outdating in this category is not accurate as a significant percentage of the sales is done by weight instead of predetermined price per consumer unit. Finally, this category is very extensive. Combining all these findings, we have to decide to leave this category out of our project.

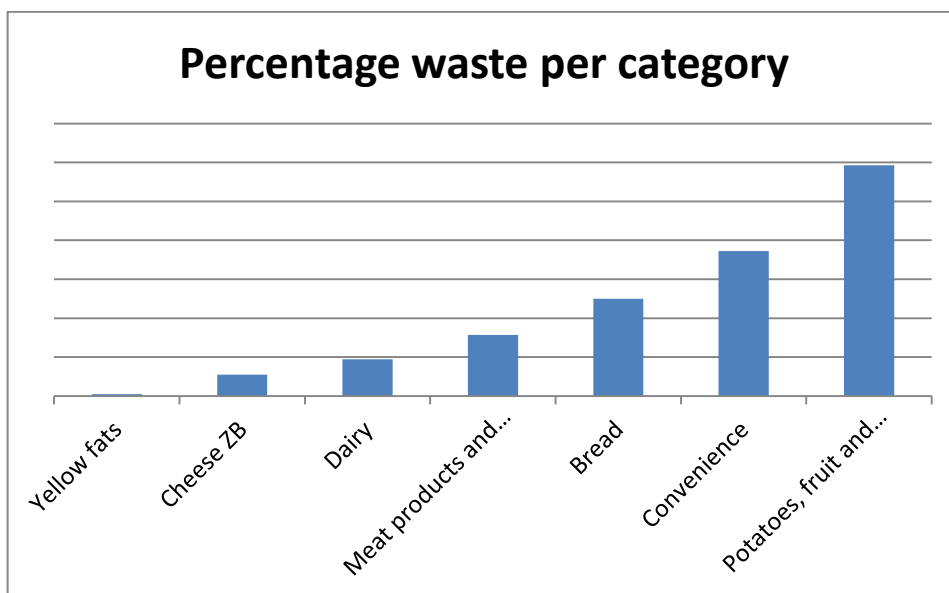


Figure 2: Percentage of waste per category

The category with the second highest percentage of waste of the remaining categories is convenience. This category will be investigated in this project.

If we go further to the left we see that bread also has a high percentage of waste. These products are not coming from HB so we leave them out of our analyses. The next category to the left is meat products and salads ZB. This category will be investigated as well in this project. Due to time restriction no further categories will be analyzed in this project.

2.6.2 Selection of supermarkets for visits

Several supermarkets were visited for interviews and measurements. Explorative research at the beginning of the project showed that the percentage of waste was not equal in the provinces of The Netherlands. Next to this, PLUS doesn't have the same amount of supermarkets in each provinces. Because of this reason, supermarkets were visited throughout the whole country.

In the end 37 supermarkets from 33 different entrepreneurs were visited. 1 supermarket was owned by PLUS Retail. The average sales area of the visited supermarkets was 988 square meters. This is only 7 square meters less than the average PLUS supermarket. Both very small and very large supermarkets were visited. The average revenue share of the convenience category of the total revenue share of the product categories of HB for the sample was 0,08% higher than the average revenue share of this category in the total population. The average revenue share of the meat products and salads category of the total revenue share of the product categories of HB for the sample was 0,39% lower than the average revenue share of this category in the total population. The differences between the sample of supermarkets that was visited and the total population are quite small. So, the sample is a good representation of the total population.

2.6.3 Selection of supermarkets for data analysis

The accuracy of the available data is very important for the accuracy of the results of our analyses. Because of this, supermarkets that don't register the outdating of products accurately will be left out the analysis. Examples of supermarkets from where we cannot use the data are supermarkets that don't register all the outdating or scan articles on a wrong article code.

Our first selection of supermarkets is based on the number of opening days during the period of the available data. Supermarkets that were only open during a part of the year were left out our analyses. Furthermore, we delete supermarkets with a lot of inaccurate data. Supermarkets with a high percentage of sales in non-scan items cannot be used, because we don't know which articles are exactly sold if an article is scanned as non-scan. Supermarkets with more than 10% of the sales in consumer units in non-scan items will be left out our analysis. Next to a correct way of scanning products at the cash desks, a correct way of scanning products that are discounted or thrown away is important for this project. Because of this, we delete supermarkets where the following ratio is higher than the average ratio for all supermarkets + one standard deviation.

$$Ratio = \frac{\text{Stock corrections in consumer units}}{\text{Waste due to outdating in consumer units}} \quad (1)$$

2.6.4 Selection of sales and outdating data for data analysis

In contrast with Van Donselaar et al. (2006), products in the convenience and meat products and salads categories with a shelf life of more than 30 days are seen as perishables as well and thus included in this project. In these products outdating occurs as well.

Perishables that are on promotion in all supermarkets will be left out of scanning data, because their demand pattern differs a lot from regular items and they normally have a different remaining shelf life when arriving at the supermarkets. Next to the countrywide promotions, there are also local promotions. These promotions will also be left out of the regression, because their demand pattern also differs a lot from regular items. Promotional products that are thrown away cannot be left out of the outdating data, as it is not registered if the waste is coming from regular or promotional items.

Local items will be left out of the sales and outdating data. Discounted products that are sold are left out the sales data as these products cannot be linked to the regular products.

The outdating data contains different reason codes for waste. We are not interested in waste coming from causes like tasting session or production waste. We only take into account waste with the following reason codes: 'Overcode', 'Bederf', 'Afprijzen voorraad correctie' and 'Derving voorraad correctie'. Note that part of the stock corrections that we take into account is coming from theft by customers or employees. Most supermarkets don't register all the products that are thrown away, so if we only take the registered waste, the amount of outdating is largely underestimated.

2.7 Stakeholders of the project

There are a number of stakeholders involved in this project. First of all, PLUS Retail, because they provide all scanning and sales data necessary for this project. The results of the project can have consequences for PLUS Retail as they can provide advises to their entrepreneurs. Second, HB, because they provide all data about the deliveries of the perishables to the supermarkets. Reducing the amount of outdating in the PLUS supermarkets will reduce the amount of products they have to send to the supermarkets and in this way to a lower revenue for HB. Finally, the entrepreneurs of PLUS supermarkets as they can decide which products to order and how much. Results from this project can change their order decisions.

3. Analysis from interviews

As described in chapter two 37 supermarkets from 33 different entrepreneurs were visited. 1 supermarket was owned by PLUS Retail. In this chapter the opinion of the interviewees is described.

3.1 Opinion about outdating and width of product assortment

From the interviews with supermarket entrepreneurs, supermarket managers, department managers and store clerks, two types of managers that run a supermarket appear. In this report the two types are defined as:

1. Grocer: manager with a (very) strong focus on outdating; tries to get the percentage outdating as low as possible by changing assortment or settings in the automated store ordering system. Products with structural waste will be deleted from the assortment after a short period of time. Shelves don't have to be full of products during the whole day; just before delivery from HB shelves may be quite empty. Manager expects that customers will not easily go to another supermarket if a specific product is not in the assortment or out of stock.
2. Entrepreneur: manager without a strong focus on outdating; if outdating is below a set norm, the outdating is acceptable. More focus on revenue than on outdating. Products with structural waste will be deleted from the assortment after a relatively long period of time

and only if they add insufficient value to the assortment of the supermarket and a good alternative is available in the resulting assortment. Store image is very important and customers should have the same choice independent of the time of visiting the supermarket during the day. Manager expects that customers will easily go to another supermarket if a specific product is not in the assortment or out of stock.

As can be seen, the defined grocer and entrepreneur have a total different opinion on outdating. Grocers typically want to get the lowest percentage outdating as possible, whereas entrepreneurs take outdating as a fact of doing business. Entrepreneurs argue that focusing too much on outdating will result in more out-of-stocks and an assortment that is too small. Some entrepreneurs even order more products than advised from the automated store ordering system to get a better store image. They simply accept that this will result in more outdating.

7 out of the 37 managers (18,9%) are classified as grocers. The remaining 30 of the 37 managers (81,1%) are classified as entrepreneurs. We cannot extrapolate this percentage to all PLUS supermarkets as this classification can only be made after an interview with the manager of the supermarket. Grocers seem to be active in smaller supermarkets than entrepreneurs.

If the specific products have outdating several times during a period of a month managers speak of structural waste. They can do several things in case of a structural waste problem:

1. Change the place of the product in the shelf
2. Adjust the minimum stock level in the automated store ordering system
3. Order the products manually instead of by the automated store ordering system
4. Accept the structural waste
5. Remove the product from the assortment

Managers classified as grocers tend to remove products with structural waste earlier from the assortment than managers classified as entrepreneurs. However, both types want to give new products a chance to get known by the customers. Some managers decide not to follow shelf plans from PLUS because they think some products will cause structural waste. They don't want to make customers satisfied with new products and having to remove the products after a short period because the products cause structural waste.

Waste of products is in the current situation the full responsibility of the supermarket manager. Some managers argue that this is not fair as the category management of PLUS Retail initially determines which products should be sold in the supermarkets and which products shouldn't. Because of this, the category management should be partly responsible for the waste in the PLUS supermarkets. Often the category management makes shelf plans with a too broad assortment for the revenue that a supermarket has. Shelf plans are now based on the meters of refrigerators instead of revenue of a supermarket. This results in outdating of products. Note that the amount of meters of refrigerators could not be used in our model as explained in chapter 2.4.5. Next to the category management, managers argue that HB is only for a small part responsible for the waste in the supermarkets. HB is only responsible for the waste if the delivered quality is not good or the remaining shelf life at arrival is shorter than agreed with PLUS Retail.

3.2 Opinion about selling products with a remaining shelf life of one day

In the past PLUS Retail had the slogan: “In PLUS supermarkets no products with a remaining shelf life of one day can be found”. Of course day fresh products like newspapers and bread were excluded from this slogan. Today, PLUS supermarkets can decide themselves whether they want to sell products with a remaining shelf life of one day or not. Following from the interviews, PLUS managers totally disagree about this topic.

On the one hand managers, mostly grocers, argue that selling these products is necessary to run a profitable supermarket. These managers argue that when they decide to not sell these products, their outdating will increase in such a way that the viability of their supermarkets is not guaranteed. 71,4% of the grocers sell these products in contrast to 46,7% of the entrepreneurs. In the supermarkets that sell these products, these products are mostly discounted on the whole or part of the day of expiration. On the other hand, the grocers and entrepreneurs that don't sell these products argue that selling these products is not in line with the core values of PLUS as formulated in chapter 1.2. They argue that customers should be able to trust a PLUS supermarket and be able to pick any product from the shelves without having to check the expiration date first. They argue that customers pay higher prices in a PLUS supermarket than in a discount supermarket and because of this should expect more quality of the products and should trust the supermarket more. In particular if these products are sold on the day of expiration for the regular price, the managers that don't sell these products argue that this is not fair to the customers.

3.3 Opinion about discounting

Not in all PLUS supermarkets products with a short remaining shelf life are discounted. The managers who discount products see discounting products with a short remaining shelf life as a way to increase the chance that these products will be sold before they expire and in this way prevent outdating. Furthermore, managers argue that to prevent outdating using the right parameter settings in the automated store ordering system is way more important than discounting products. The way in which managers discount their products differs a lot between supermarkets.

Almost all interviewees argue that discounted products are not beneficial for the store image. That is why managers will set up a local promotion if a lot of products are left over from countrywide promotions. The opinion that discounting products is not beneficial for the store image is also the reason why some managers decide to don't discount at all. In 5 of the 37 visited supermarkets (13,5%), products are never or seldom discounted.

Analyzing the waste data of the Convenience category comes to an estimated 15,8% of the supermarkets that don't discount at all or discount not a significant amount in this category. This is calculated by taking the supermarkets where the following ratio is lower than the average ratio for all supermarkets minus one time the standard deviation.

$$Ratio = \frac{\text{number of discounted products}}{\text{total waste except discounted products}} \quad (2)$$

Note that the number of discounted products include all products that are discounted independent from whether they are sold or thrown away eventually. The number of total waste includes waste from all reason codes except waste from discounted products.

Using the same criteria for the Meat products ZB category comes to an estimated 14,3% of the supermarkets that don't discount at all or discount not a significant amount in this category. Based on these numbers we can estimate that in around 14,5% of all PLUS supermarkets products are never or seldom discounted when they have a short remaining shelf life.

Next to the decision to discount products or not, some managers decide to discount only products in certain categories. Only two supermarkets don't discount products in the meat category and a high number of seven supermarkets don't discount products in the potatoes, fruit and vegetables category. Supermarkets that normally don't discount products, choose to discount products in the meat category as the percentage of outdated is quite high in this category. In the potatoes, fruit and vegetables category, managers mainly decide to not discount products because of the store image.

3.3.1 Discount percentage

For the supermarkets that do discount their products with a short remaining shelf life, the discount percentage of the perishables differs a lot between and sometimes also within supermarkets. The minimum discount percentage observed during the supermarket visits is 25%, the maximum discount percentage observed is 50%. Supermarkets that choose 25% argue that this percentage is high enough to convince consumers to buy the items. Supermarkets that choose 50% typically try to sell discounted products as soon as possible to keep a good store image.

Most supermarkets use one discount percentage for all categories of perishables. However, some managers argue that for some categories a lower or higher percentage is better. For example, some managers argue that for dairy products a lower discount percentage is better, as these products have a lower margin. On the other hand some products need, according to some managers, a higher percentage to convince consumers to buy them, like pre-cut vegetables or meat.

Most supermarkets use one discount percentage for all products in a category. However, some supermarkets use a higher discount percentage of 50% for products that are special or quite expensive like fish or special meat. They argue that when discounting those products with the normal discount percentage of 35% in their supermarkets, these products will not be sold, because they are still too expensive for consumers who normally buy average priced articles.

Next to this, employees can change the resulting price from choosing a certain discount percentage. Sometimes this calculated price is changed to a more convincing price for customers. For example, the calculated price of €2,03 will be changed to €1,98 to convince customers.

3.3.2 Discount hours

In the supermarkets that do discount perishables, the moment of discounting differs. The number of discount hours is the total opening hours of a supermarket in which discounted products are sold. For example, if employees discount products at 3 pm and remove the unsold discounted products at 8 pm, there are 5 discount hours for the products in this category. When employees discount products with a remaining shelf life of more than one day, a product has more than 12 discount hours. For example, if employees discount products on day 1 at 5 pm with a remaining shelf life of 2 days and these products are removed on day 2 at 8 pm, there are $3+12=15$ discount hours for the products in this category.

Supermarkets choose the moment of discounting products mostly based on available employees and intuition. This means that discounting products is not done at the same time each day of the week.

Department managers and store clerks have the freedom to decide at what time of the day they discount product based on guidelines from the supermarket manager. Products are not discounted at the same time each day of the week as supermarkets don't want to sell products for a discounted price which would be sold for a normal price as well. On days with a lot of products with a short remaining shelf life the products will be discounted earlier than on days with only a small amount of products with a short remaining shelf life. In the table below the guidelines from the supermarket managers can be found. As can be seen, perishables are discounted in many cases earlier than on the day of expiration.

3.4 Opinion about grabbing behavior

The interviewees were asked about the grabbing behavior they experience in their supermarket and if this has an influence on the amount of outdating of perishables. As defined in chapter 1.4.5 grabbing behavior occurs when consumers buy products with a later best-before-date than the first product in the shelf with a shorter best-before-date given that there are two or more best-before-dates in the shelf.

In all supermarkets grabbing behavior is observed. The amount of observed grabbing behavior differs a lot between supermarkets, but we cannot compare supermarkets as no quantitative data is available about the grabbing behavior. Managers argue that part of the grabbing behavior can be prevented when customers have trust in the supermarket. However, some customers will always check the full shelf for products with the longest remaining shelf life. Especially managers of supermarkets in lower-income neighborhoods notice a lot of grabbing behavior. They argue that their customers have plenty of time to check all the products in the shelves. In 23 of the 37 visited supermarkets (62,2%) rules exist for the amount of different best-before-dates that may be placed in the shelves. In these 23 supermarkets, the managers want at maximum one or two different dates in the shelves. They will keep products in the stockroom if necessary.

Some supermarkets have springs in the shelves of perishables. Only in the categories cheese, convenience, meat products and salads ZB, potatoes, fruit and vegetables springs can be found. In the table below information about springs in the shelves can be found. Managers argue that the choice for springs is mainly based on store image and financial motives rather than on preventing grabbing behavior. In some categories only a few or no springs are placed because pushing the products is harmful for the products. This holds for example for the convenience products such as pancakes and hotdogs and the vegetable products like lettuce and mushrooms.

4. Analysis from data

In this chapter the gathered data at HB, PLUS Retail, PLUS supermarkets and from online data sources will be analyzed.

4.1 Analysis of convenience category

Our first analysis will be conducted within the convenience category. We use sales, delivery and outdating data from the 22nd of October 2013 till the 19th of October 2014.

A first look at the data shows that the convenience category is relatively small viewed from a revenue perspective, but important from a waste perspective for PLUS supermarkets.

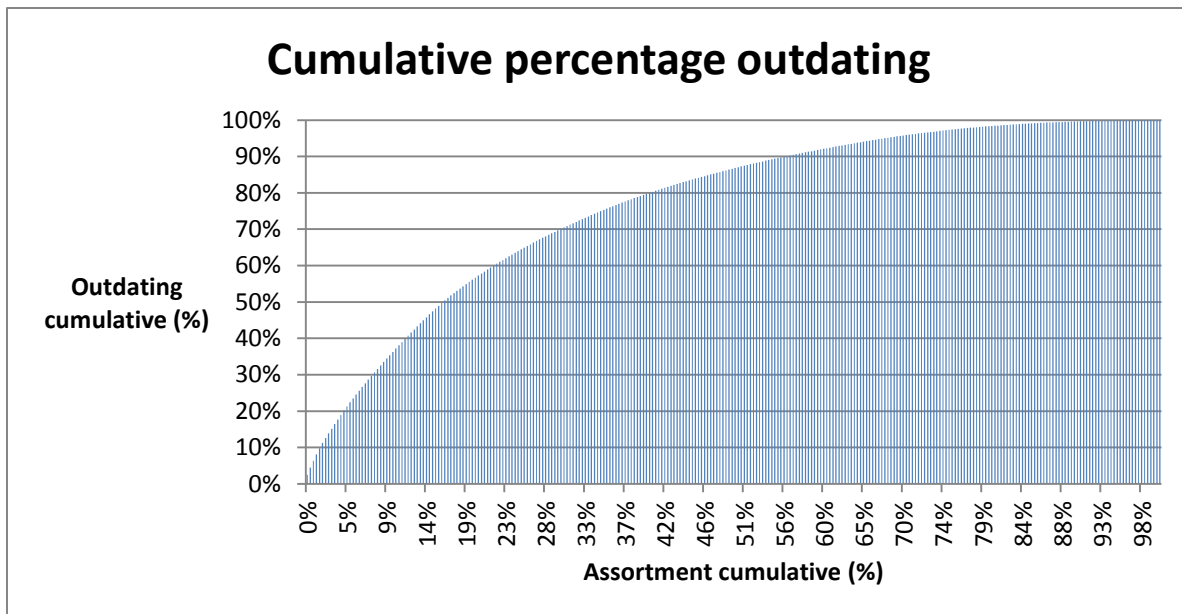


Figure 3: Distribution of outdated over assortment in convenience category

This graph shows the distribution of the outdated in this category over the assortment. As can be seen in the graph, some articles contribute much more to the total outdated in this category than others. The 10% of the assortment consisting of the articles with the most outdated results in 36% of the total outdated in this category. In contrast, the second half of the assortment with the articles that have the lowest outdated per article results only in 13% of the outdated in this category.

4.1.1 Selection of supermarkets and data

As described in chapter 2.5.3 some supermarkets and data will be removed from the dataset. Supermarkets will be removed in case of too less opening days during the period of the available data, in case of a too high percentage of non-scan items or in case of too much unclassifiable waste. 27 out of the 259 supermarkets were deleted from the dataset.

As described in chapter 2.6.4 perishables that are on countrywide or local promotion will be left out of the scanning data. Promotional products that are thrown away cannot be left out of the outdated data, as it is not registered if the waste is coming from regular or promotional items. Local items will be left out of the sales and outdated data. Discounted products that are sold are left out the sales data as these products cannot be linked to the regular products. For the rest of this chapter only waste with the reason codes 'Overcode', 'Bederf', 'Afprijzen voorraad correctie' or 'Derving voorraad correctie' will be taken into account.

4.1.2 Analysis 1: analysis on supermarket level

We start running a multiple regression analysis with the percentage of outdated as dependent variable and the other 14 variables as described in chapter 2.4.1 as independent variables. In all regression analyses in this report we are searching for the model with the highest adjusted R^2 as this represents the model with the highest percentage of variation in the dependent variable that is explained by the model. We use the adjusted R^2 instead of the R^2 because the adjusted R^2 takes into account the number of variables included in the model. We can use hierarchical regression when predictors are selected based on past work (Field, 2009). As all mentioned literature predicts outdated on product level and not on supermarket level, we cannot use hierarchical regression.

Next, we can use forced entry when we have good theoretical reasons for including the predictors. This is not the case for outdating on supermarket level, so we cannot use this method. As we are not sure which variables have an effect on the of outdating we use the backward method. This is done because forward selection is more likely than backward elimination to exclude predictors involved in suppressor effects (Field, 2009). As explained in chapter 2.4.1 the number of variables is relatively small for the number of observations, so we have to be careful about the outcomes.

Using the backward method we find that the supermarket QMuM ratio, percentage of order advises that is followed, average remaining shelf life, category revenue share, average order size, average coefficient of variation of order size and the average sales are significant. However, the average sales and the supermarket QMuM ratio are highly correlated. This is expected as the average sales is part of the supermarket QMuM ratio.

Separate regression analyses show that the average sales should be removed for a higher adjusted R square. Running the regression analysis in the same way with the remaining 13 independent variables shows that seven variables are significant: the supermarket QMuM ratio ($b=44,989$), percentage of order advises that is followed ($b=5,993$), average remaining shelf life ($b=0,818$), category revenue share ($b=-35,132$), average order size ($b=-1,393$), average coefficient of variation of order size ($b=8,452$) and sales area ($b=0,001$). This model gives a high adjusted R square of 0,564. All assumptions applicable for multiple regression analyses are met. More information about testing the assumptions of the multiple regression analysis and about the results can be found in Appendix D. Due to confidentiality not all relevant tables are included in the Appendix.

According to our analysis the higher the supermarket QMuM ratio, the higher the percentage of outdating per supermarket. An increase in the percentage of order advises that is followed leads to a higher percentage of outdating. An increase in the average remaining shelf life leads to a higher percentage outdating. An increase in the category revenue share leads to a lower percentage outdating. An increase in the average order size leads to a lower percentage of outdating. An increase in the average coefficient of variation of order size leads to a higher percentage of outdating. Finally, an increase in the sales area will lead to a higher percentage of outdating.

We continue analysis 1 by judging why the differences in the significant variables exist between the supermarkets and if the values of the variables can be changed. The size of the standardized regression coefficients tells us which factors are the most important factors for predicting the percentage of outdating on supermarket level. The most important factor for the percentage of outdating is the supermarket QMuM ratio. There is a strong positive relationship between the supermarket QMuM ratio and the percentage outdating of a supermarket ($b=44,989$).

The strong positive relationship ($b=44,989$) is in line with our hypothesis. The supermarket QMuM ratio can partly be influenced by a supermarket and partly by HB. By removing products from the assortment with a high product QMuM ratio, the supermarket QMuM ratio of a supermarket decreases and the percentage outdating will decrease as well. Another solution is to increase the average sales by getting more customers in the supermarket. This is however difficult to realize. A third solution is to decrease the case pack size, so that a supermarkets needs less days to sell a case pack. This has however implications for HB and their suppliers as well. This will probably also result in more different best before dates in the shelves. The fourth solution is to increase the average remaining shelf life. This can only be done by HB. They should get the products with the longest

remaining shelf life of their suppliers and should keep just enough stock to serve all demand from the supermarkets. The average remaining shelf life is very difficult to change at acceptable costs.

The second most important factor is the average coefficient of variation of order size. The strong positive relationship ($b=8,452$) is in line with our hypothesis. This factor is changeable by the supermarkets by placing orders of the right order size. If the ordering pattern is more stable and the order size is more in line with the expected demand, the variation of the order size will decrease and the percentage of outdating will decrease according to our model.

The third most important factor is the average order size. This factor is totally under control of a PLUS supermarket. No hypothesis is formulated about this variable as the average order size on its own doesn't say a lot about the outdating of a supermarket. The negative relationship ($b=-1,393$) is likely to be coming from the differences in revenue of supermarkets. The products in the assortments have different case pack sizes and not all supermarkets order one case pack size of a product per order. Large supermarkets will order more than one case pack size of a product and small supermarkets will only order one case pack size of a product.

The fourth most important factor is the average remaining shelf life. This factor is only changeable by HB. They should get the products with the longest remaining shelf life of their suppliers and should keep just enough stock to serve all demand from the supermarkets. This factor is very difficult to change at acceptable costs. The positive relationship ($b=0,818$) is in contrast to our hypothesis. We expect that supermarkets with a longer average remaining shelf life have a lower percentage of outdating. The positive relationship in the data is possibly coming from the different assortments that PLUS supermarkets have. Each product has its own average remaining shelf life and a unique combination of products will lead to a unique average remaining shelf life per supermarket.

The fifth most important factor is the sales area. The positive relationship ($b=0,001$) is in contrast to our hypothesis. An increase of the sales area of 100 square meters will result in an increase of 0,1 percent of the percentage of outdating. This positive relationship is possibly coming from the fact that large supermarkets have a higher turnover per square meter, but also a broader assortment with some unpopular products which result in outdating.

The sixth most important factor is the percentage of order advises that is followed. This factor is totally under control of a PLUS supermarket. The strong positive relationship ($b=5,993$) is in contrast to our hypothesis. The significance of this factor is 0,038 which means that the significance of this factor is just low enough to be included in our model. This means that we have to be careful about conclusions based on this factor. We expect that ordering based on statistical basis will result in less outdating than ordering based on emotion and intuition. The positive relationship found in the data is possibly coming from the fact that the automated store ordering system of PLUS doesn't include weather forecasts, promotions and seasonal effects in the order advises. Next to this, the system doesn't have a tradeoff between waste and service level. Delaying a planned order is likely to result in less outdating and a lower service level. Weather changes, promotions and seasonal effects will have an influence on the sales. Because of this, these effects should be predicted so that order advises can be changed only if needed.

The seventh most important factor is the category revenue share. The strong negative relationship ($b=-35,132$) is in line with in our hypothesis. The significance of this factor is 0,099 which means that

the significance of this factor is just low enough to be included in our model. This means that we have to be careful about conclusions based on this factor. This factor is partly changeable by a PLUS supermarket. By having the right assortment for the customers, they will buy more convenience products and the category revenue share will increase. However, research by Jongen (2013) showed that the category revenue share has a strong relationship with the customer groups of a supermarket. These customer groups are not changeable as they are mainly dependent on the location of the supermarket.

For the average coefficient of variation of sales, the number of opening days, the average disposable income and the service level of HB no significant relationships are found. This doesn't mean that there is absolutely no relationship, it only means that with our analysis the relationship cannot be found. This is probably because the effect of each of these variables on the percentage of outdating is too small to be noticeable with our small number of observations.

4.1.3 Analysis 2: analysis on product level

We continue with running a multiple regression analysis on product level with the percentage of outdating as dependent variable and the other 8 variables as described in chapter 2.4.2 as independent variables. As described in chapter 4.1.2 we can use hierarchical regression when predictors are selected based on past work. Some of the independent variables we use for analyzing the outdating on product level are coming from literature. However, some of them like the sales area are not mentioned in literature. This is why we decide to use the backward method. As explained in chapter 2.4.2 the number of observations is enough when we expect large or medium effects. Maybe we are not able to find small effects with this number of observations.

Using the backward method we find that all independent variables are significant. However, the natural logarithm of the average remaining shelf life and the average remaining shelf life are obviously highly correlated. Next to this, the case pack size has a very high VIF of 13,6. So, the case pack size has to be removed from the model.

Separate regression analyses show that the average remaining shelf life should be removed for a higher adjusted R square. Running the regression analysis again with the remaining six variables shows that five variables are significant: the natural logarithm of the average remaining shelf life ($b=-24,409$), the product QMuM ratio ($b=24,135$), the average coefficient of variation of sales per day ($b=27,960$), the average sales ($b=10,645$) and the average coefficient of variation of order size ($b=-23,751$). This models gives a relatively low adjusted R square of 0,296. All assumptions applicable for multiple regression analyses are met. More information about testing the assumptions of the multiple regression analysis and about the results can be found in Appendix D. Due to confidentiality not all relevant tables are included in the Appendix.

According to our analysis the longer the average remaining shelf life, the lower the percentage of outdating per product. The higher the product QMuM ratio, the higher the percentage of outdating per product. An increase in the average coefficient of variation of sales per day leads to a higher percentage of outdating. An increase in the average sales leads to a higher percentage of outdating. Finally, an increase of the average coefficient of variation of order size leads to a lower percentage of outdating.

Next to this regression analyses we check how many products have a structural waste problem. As defined in chapter 1.4.2 for these product the average time needed to sell the case pack size was longer than the average remaining shelf life of the products. This holds for four products in this category.

We continue analysis 2 by judging why the differences in the significant variables exist between the products and if the values of the variables can be changed. The size of the standardized regression coefficients tells us which factors are the most important factors for predicting the percentage of outdating on product level. The most important factor for outdating on product level is the natural logarithm of the average remaining shelf life.

The strong negative relationship ($b=-24,409$) is in line with our hypothesis. The average remaining shelf life of a product is only changeable by HB. They should get the products with the longest remaining shelf life of their suppliers and should keep just enough stock to serve all demand from the supermarkets. This factor is very difficult to change at acceptable costs.

The second most important factor is the average coefficient of variation of sales per day. No hypothesis is formulated about this factor as the average coefficient of variation of sales on its own doesn't say a lot about the outdating of a product. When a certain coefficient of variation of sales is expected due to the week pattern it doesn't have to result in outdating as orders are based on these expectations. This factor is changeable by the supermarket and by PLUS Retail. By stimulating sales on days with low sales the existing week pattern will be attenuated and the average coefficient of variation of sales will be lowered. We can't say whether the percentage of outdating will change when the average coefficient of variation of sales is changed.

The third most important factor is the product QMuM ratio. The strong positive relationship ($b=24,135$) is in line with our hypothesis. The product QMuM ratio is partly changeable by the supermarket, partly by PLUS Retail and partly by HB. By removing the product from the assortment of a supermarket with a high combination QMuM ratio for this specific product, the product QMuM ratio for this specific product decreases and the percentage outdating will decrease as well. Another solution is to increase the average sales by stimulating the product. This is however difficult to realize and will probably influence sales of other products. A third solution is to decrease the case pack size, so that a supermarkets needs less days to sell a case pack. This has however implications for HB and their suppliers as well. This will probably also result in more different best before dates in the shelves. The fourth solution is to increase the average remaining shelf life. This can only be done by HB. They should get the products with the longest remaining shelf life of their suppliers and should keep just enough stock to serve all demand from the supermarkets. The average remaining shelf life is very difficult to change at acceptable costs.

The fourth most important factor is the average coefficient of variation of the order size of a product. This factor is totally under control of a PLUS supermarket. The strong negative relationship ($b=-23,751$) is in contrast to our hypothesis. This factor is changeable by the supermarkets by placing orders of the right order size. According to our model, the percentage of outdating will decrease when the average coefficient of variation of order size increases. This is not correct.

The fifth most important factor is the average sales of a product. The strong positive relationship ($b=10,645$) is in contrast to our hypothesis. The significance of this factor is 0,040 which means that

the significance of this factor is just low enough to be included in our model. This means that we have to be careful about conclusions based on this factor. A positive relationship is not possible as higher average sales can never lead to a higher percentage of outdating. The only possible reason for the found relationship is that products are sold in specific supermarkets and not in all supermarkets. Comparing products with a different number of selling points maybe leads to a wrong comparison and wrong results.

4.1.4 Analysis 3: analysis on supermarket – product combination

We continue with running a multiple regression analysis on each supermarket product combination with the percentage of outdating as dependent variable and the other 8 variables as described in chapter 2.4.3 as independent variables. For the same reasons as for the analysis on product level we decide to use the backward method. As explained in chapter 2.4.3 the number of observations is large enough to find large, medium and even small effects.

Using the backward method we find that all independent variables are significant. However, the natural logarithm of the average remaining shelf life and the average remaining shelf life are obviously highly correlated. Separate regression analyses show that the average remaining shelf life should be removed for a higher adjusted R square. We see that the adjusted R square of this model with seven independent variables is only 0,184. This is low, so a deeper analysis is needed.

The total dataset of supermarket – product combinations is split per article and some articles are removed from the dataset. After this the total dataset of remaining supermarket – product combinations is split per supermarket. Some supermarkets are removed from the dataset. Details about this process can be found in Appendix D.

A new multiple regression analyses is performed with the remaining 196 out of the 224 articles and 220 out of the 230 supermarkets. Using the backward method we find that all independent variables are significant. However, the natural logarithm of the average remaining shelf life and the average remaining shelf life are obviously highly correlated. Separate regression analyses show that the average remaining shelf life should be removed for a higher adjusted R square. We see that the adjusted R square of this model with seven independent variables is a relatively low 0,211. Not all assumptions applicable for multiple regression analyses are fully met. The assumption of normally distributed errors is not fully met. More information about testing the assumptions of the multiple regression analysis and about the results can be found in Appendix D. Due to confidentiality not all relevant tables are included in the Appendix.

According to our analysis the higher the combination QMuM ratio, the higher the percentage of outdating per product ($b=58,474$). An increase in the natural logarithm of the average shelf life of a product leads to a lower percentage of outdating ($b=-15,155$). An increase in the coefficient of variation of order size leads to lower percentage of outdating ($b=-16,724$). An increase in the average sales leads to a higher percentage of outdating ($b=4,534$). An increase of the coefficient of variation of sales leads to a lower percentage of outdating ($b=-2,806$). An increase in the average order size leads to a higher percentage of outdating ($b=0,250$). Finally, an increase of the case pack size of a product leads to a higher percentage of outdating ($b=-2,930$).

Next to this regression analyses we check how many supermarket - products combinations have a structural waste problem. For these combinations the average time needed to sell the case pack size

in a supermarket was longer than the average remaining shelf life of the products. This holds for 2182 combinations in this category. Grocers have on average a higher absolute and relative amount of combinations with structural waste than entrepreneurs.

We continue analysis 3 by judging why the differences in the significant variables exist between the combinations and if the values of the variables can be changed. The size of the standardized regression coefficients tells us which factors are the most important factors for predicting the percentage of outdating of a combination. The most important factor for the percentage of outdating is the combination QMuM ratio.

The strong positive relationship ($b=58,474$) is in line with our hypothesis. As explained in chapter 4.1.3 the combination QMuM ratio for a product is partly changeable by the supermarket, partly by PLUS Retail and partly by HB. Ways to change this variable can be found in chapter 4.1.3 as well.

The second most important factor is the case pack size of a product. This factor is changeable by HB and their suppliers. Decreasing the case pack size will increase the chance that a case pack will be sold within the remaining shelf life of a case pack. However, smaller case packs will lead to extra handling cost at the suppliers, HB and the PLUS supermarkets. Next to this, it is likely that smaller case packs will lead to more batches in the shelves and thus in more grabbing behavior as noticed by the supermarket managers. As described in chapter 2.5.2 no hypothesis is formulated about this factor as the case pack size on its own don't say a lot about the outdating of a product. For example, a large case pack size only results in more outdating when the average sales are low and/or the remaining shelf life of the products is short.

The third most important factor is the natural logarithm of the average remaining shelf life. The strong negative relationship ($b=-15,155$) is in line with our hypothesis. The average remaining shelf life is only changeable by HB. The third most important factor is the coefficient of variation of order size. The strong negative relationship ($b=-16,724$) is in contrast to our hypothesis. This factor is totally under control of a PLUS supermarket. Ways to change these factors can be found in chapter 4.1.3.

The fourth most important factor is the coefficient of variation of order size. This factor is totally under control of a PLUS supermarket. The strong negative relationship ($b=-16,724$) is in contrast to our hypothesis. This factor is changeable by the supermarkets by placing orders of the right order size. According to our model, the percentage of outdating will decrease when the average coefficient of variation of order size increases. This is not correct.

The fifth most important factor is the average sales. The positive relationship ($b=4,534$) is in contrast to our hypothesis. A positive relationship is not possible as higher average sales can never lead to a higher percentage of outdating. The only possible reason for the found relationship is that products are sold in specific supermarkets and not in all supermarkets. Comparing products with a different number of selling points maybe leads to a wrong comparison and thus a wrong conclusion.

The sixth most important factor is the coefficient of variation of sales. As described in chapter 4.1.3, no hypothesis is formulated about this factor as the average coefficient of variation of sales on its own doesn't say a lot about the outdating of a product. This factor is changeable by the supermarket and by PLUS Retail.

The seventh most important factor is the average order size. This factor is totally under control of a PLUS supermarket. As described in chapter 2.5.2 no hypothesis is formulated about this factor as the average order size on its own don't say a lot about the outdating of a product. For example, a large order size only results in more outdating when the average sales are low and/or the remaining shelf life of the products is short.

4.1.5 Analysis 4: extra analysis on supermarket level

We continue our analysis of the convenience category by running a multiple regression analysis with as dependent variable the percentage of outdating per supermarket. The seven variables that were relevant for the convenience category in analysis 1 and the six new variables as described in chapter 2.4.4 will be used as independent variables.

We are allowed to use the enter method, as we have good theoretical reasons for including those predictors (Field, 2009). However we go from a dataset with 230 supermarket to a dataset of 34 supermarkets. The *b*-values are likely to change when going from 230 to 34 supermarkets, so we have to use the *b*-values following from the model of analysis 1. We also use the constant of analysis 1. This leads to a model with an adjusted R² of 0,519. Next we use the stepwise method to test which of the six new variables could be included in the model. As described in chapter 2.4.4, using the backward method is not possible for analysis 4 as we only have 34 observations. Using the stepwise method shows that none of the six variables will be included. The results can be found in Appendix D.

So, we have to conclude that none of the six new variables has a significant influence on the percentage of outdating of perishables in the convenience category. This doesn't mean that there is absolutely no relationship, it only means that with our analysis the relationship cannot be found. This is probably because the effect of each of the six new variables on the percentage of outdating is too small to be noticeable with our small number of observations.

4.2 Analysis of meat products and salads ZB category

We continue the analysis chapter with the meat products and salads ZB category. We use sales, delivery and outdating data from the 21nd of October 2013 till the 19th of October 2014.

A first look at the data shows that the meat products and salads category is relatively small viewed from a revenue perspective, but important from a waste perspective for PLUS supermarkets.

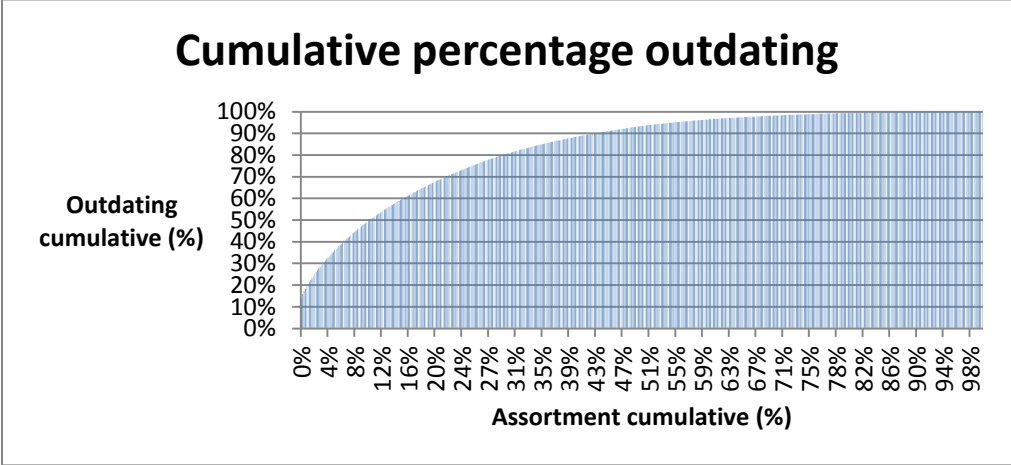


Figure 4: Distribution of outdating over assortment in meat products and salads category

This graph shows the distribution of the outdated in this category over the assortment. As can be seen in the graph, some articles contribute much more to the total outdated in this category than others. The 10% of the assortment consisting of the articles with the most outdated results in 50% of the total outdated in this category. In contrast, the second half of the assortment with the articles that have the lowest outdated per article results only in 6% of the outdated in this category.

4.2.1 Selection of supermarkets and data

Selection of supermarkets and data will be done in the same way as for the convenience category. In the meat products and salads category 26 out of the 259 supermarkets were deleted from the dataset.

4.2.2 Analysis 1: analysis on supermarket level

A first look at the data of this category shows that one supermarket has an average percentage outdated of that is more than nine times higher than the other supermarkets. We classify this supermarket as an outlier and remove it from the dataset. Next, we start running a multiple regression analysis with the percentage of outdated as dependent variable and the other 14 variables as described in chapter 2.4.1 as independent variables. As described in chapter 4.1.2 we cannot use hierarchical or forced entry regression. As we are not sure which variables have an effect on the percentage of outdated we use the backward method. This is done because forward selection is more likely than backward elimination to exclude predictors involved in suppressor effects (Field, 2009). As explained in chapter 2.4.1 the number of variables is relatively small for the number of observations, so we have to be careful about the outcomes.

Using the backward method we find that the supermarket QMuM ratio, percentage of order advises that is followed, category revenue share, average order size, average coefficient of variation of order size, average coefficient of variance of sales, average remaining shelf life and the natural logarithm of the average remaining shelf life are significant. However, the average remaining shelf life and the natural logarithm of the average remaining shelf life are highly correlated. This is expected as the natural logarithm of the average remaining shelf life is only a transformation of the average remaining shelf life.

Separate regression analyses show that both the average remaining shelf life and its natural logarithm will not be included in the regression model when they are without their highly correlated variable. Running the regression analyses again with the remaining 12 variables shows that six variables are significant: the supermarket QMuM ratio ($b=26,782$), percentage of order advises that is followed ($b=-0,034$), category revenue share ($b=-0,142$), average order size ($b=-0,318$), average coefficient of variation of order size ($b=5,055$) and the average coefficient of variation of sales ($b=1,314$). This model gives a high adjusted R square of 0,453. All assumptions applicable for multiple regression analyses are met. More information about testing the assumptions of the multiple regression analysis and about the results can be found in Appendix D. Due to confidentiality not all relevant tables are included in the Appendix.

According to our analysis the higher the supermarket QMuM ratio, the higher the percentage of outdated per supermarket. An increase in the percentage of order advises that is followed leads to a lower percentage of outdated. An increase in the category revenue share leads to a lower percentage outdated. An increase in the average order size leads to a lower percentage of outdated. An increase in the average coefficient of variation of order size leads to a higher percentage of

outdating. Finally, an increase in the average coefficient of variation of sales will lead to a higher percentage of outdating.

We continue analysis 1 by judging why the differences in the significant variables exist between the supermarkets and if the values of the variables can be changed. The size of the standardized regression coefficients tells us which factors are the most important factors for predicting the percentage of outdating on supermarket level. The most important factor for the percentage of outdating is the supermarket QMuM ratio.

The strong positive relationship ($b=26,782$) is in line with our hypothesis. As described in chapter 4.1.2 the supermarket QMuM ratio can partly be influenced by a supermarket and partly by HB. Ways to change the supermarket QMuM ratio can be found in chapter 4.1.2 as well.

The second and third most important factors are the average coefficient of variation of order size ($b=5,055$) and the average order size ($b=-0,318$). The significance of the average order size is 0,036 which means that the significance of this factor is just low enough to be included in our model. This means that we have to be careful about conclusions based on this factor. The direction of the two relationships is the same as in analysis 1 in the convenience category; the only difference is the size of the effects. These factors are described in more detail in chapter 4.1.2.

The fourth most important factor is the average coefficient of variation of sales. As described in chapter 4.1.3, no hypothesis is formulated about this factor as the average coefficient of variation of sales on its own doesn't say a lot about the outdating of a product. Ways to change this factor can be found in chapter 4.1.3 as well.

The fifth most important factor is the category revenue share ($b=-0,142$). The significance of this factor is 0,032 which means that the significance of this factor is just low enough to be included in our model. This means that we have to be careful about conclusions based on this factor. The direction of this relationship is the same as in analysis 1 in the convenience category; the only difference is the size of the effect. This factor is described in more detail in chapter 4.1.2.

The sixth most important factor is the percentage of order advises that is followed. The small negative relationship ($b=-0,034$) is in line with our hypothesis. The significance of this factor is 0,048 which means that the significance of this factor is just low enough to be included in our model. This means that we have to be careful about conclusions based on this factor. We expect that ordering based on statistical basis will result in less outdating than ordering based on emotion and intuition. This is found in the data from the PLUS supermarkets in the meat products and salads category.

For the number of opening days, the average disposable income, the average sales and the service level of HB no significant relationships are found. This doesn't mean that there is absolutely no relationship, it only means that with our analysis the relationship cannot be found. This is probably because the effect of each of these variables on the percentage of outdating is too small to be noticeable with our small number of observations.

4.2.2 Analysis 2: analysis on product level

We continue with running a multiple regression analysis on product level with the percentage of outdating as dependent variable and the other 8 variables as described in chapter 2.4.2 as independent variables. As described in chapter 4.1.2 we can use hierarchical regression when

predictors are selected based on past work. Some of the independent variables we use for analyzing the outdating on product level are coming from literature. However, some of them like the order size are not mentioned in literature. This is why we decide to use the backward method. As explained in chapter 2.4.2 the number of observations is enough when we expect large or medium effects. Maybe we are not able to find small effects with this number of observations.

Using the backward method we find that the product QMuM ratio, average coefficient of variation of sales, average sales, case pack size and the average order size are significant. The average order size and the case pack size have a VIF of 6,390 and 6,030 respectively. This is quite high, but not a problem for our analysis as this strong correlation is logical as most supermarkets order one case pack size when ordering a product. We decide to leave both variables in our model.

So, our model has five independent variables: the product QMuM ratio ($b=43,742$), the average coefficient of variation of sales ($b=4,455$), the average sales ($b=2,484$), the case pack size ($b=-1,006$) and the average order size ($b=0,239$). This model gives a high adjusted R square of 0,693. All assumptions applicable for multiple regression analyses are met. More information about testing the assumptions of the multiple regression analysis and about the results can be found in Appendix D. Due to confidentiality not all relevant tables are included in the Appendix.

According to our analysis the higher the product QMuM ratio, the higher the percentage of outdating per product. An increase in the average coefficient of variation of sales per day leads to a higher percentage of outdating. An increase in the average sales leads to a lower percentage of outdating. An increase in the case pack size leads to a lower percentage of outdating. Finally, an increase of the average order size leads to a higher percentage of outdating.

Next to this regression analyses we check how many products have a structural waste problem. As defined in chapter 1.4.2 for these products the average time needed to sell the case pack size was longer than the average remaining shelf life of the products. This holds for six products in this category.

We continue analysis 2 by judging why the differences in the significant variables exist between the products and if the values of the variables can be changed. The size of the standardized regression coefficients tells us which factors are the most important factors for predicting the percentage of outdating on product level. The most important factor for outdating on product level is the product QMuM ratio. The strong positive relationship ($b=43,742$) is in line with our hypothesis. As described in chapter 4.1.3. the product QMuM ratio is partly changeable by the supermarket, partly by PLUS Retail and partly by HB. Ways to change this factor can be found in chapter 4.1.3.

The second most important factor is the case pack size of a product. No hypothesis is formulated about this variable as the case pack size on its own doesn't say a lot about the outdating of a product. The negative relationship ($b=-1,006$) is likely to come from the differences in assortment of supermarkets. Supermarkets have different assortments and maybe products with relatively small or large case pack sizes are not sold in a representative amount of the PLUS supermarkets. This factor is changeable by HB and their suppliers. Decreasing the case pack size will increase the chance that a case pack will be sold within the remaining shelf life of a case pack. However, smaller case packs will lead to extra handling cost at the suppliers, HB and the PLUS supermarkets. Next to this, it is likely

that smaller case packs will lead to more batches in the shelves and thus in more grabbing behavior as noticed by the supermarket managers.

The third most important factor is average sales of a product. The positive relationship ($b=2,484$) is in contrast to our hypothesis. A positive relationship is not possible as higher average sales can never lead to a higher percentage of outdated. The only possible reason for the found relationship is that products are sold in specific supermarkets and not in all supermarkets. Comparing products with a different number of selling points can maybe lead to a wrong comparison and thus a wrong conclusion.

The fourth most important factor is the average coefficient of variation of sales per day. As described in chapter 4.1.3, no hypothesis is formulated about this factor as the average coefficient of variation of sales on its own doesn't say a lot about the outdateding of a product. Ways to change this factor can be found in chapter 4.1.3 as well.

The fifth most important factor is the average order size of a product. This factor is totally under control of a PLUS supermarket. No hypothesis is formulated about this variable as the average order size on its own doesn't say a lot about the outdateding on product level. The positive relationship ($b=0,239$) is likely to come from the differences in assortment of supermarkets. Note that the significance of this factor is also quite low (0,071). The products in the assortments have different case pack sizes and not all supermarkets order one case pack size of a product per order.

For the average coefficient of variation of order size no significant relationship was found. This doesn't mean that there is absolutely no relationship, it only means that with our analysis the relationship cannot be found. This is probably because the effect of each of these variables on the percentage of outdateding is too small to be noticeable with our small number of observations.

4.2.3 Analysis 3: analysis on supermarket – product combination

We continue with running a multiple regression analysis on each supermarket product combination with the percentage of outdateding as dependent variable and the other 8 variables as described in chapter 2.4.3 as independent variables. For the same reasons as for the analysis on product level we decide to use the backward method. As explained in chapter 2.4.3 the number of observations is large enough to find large, medium and even small effects.

Using the backward method we find that six independent variables are significant. However, the natural logarithm of the average remaining shelf life and the average remaining shelf life are obviously highly correlated. Separate regression analyses show that the average remaining shelf life should be removed for a higher adjusted R square. We see that the adjusted R square of this model with five independent variables is only 0,208. This is low, so a deeper analysis is needed.

The total dataset of supermarket – product combinations is split per article and some articles are removed from the dataset. After this the total dataset of remaining supermarket – product combinations is split per supermarket and some supermarkets are removed from the dataset. Details about this process can be found in Appendix D.

A new multiple regression analyses is performed with the remaining 253 out of the 281 articles and 223 out of the 233 supermarkets. Using the backward method we find that seven independent variables are significant. However, the natural logarithm of the average remaining shelf life and the

average remaining shelf life are obviously highly correlated. Separate regression analyses show that the average remaining shelf life should be removed for a higher adjusted R square. We see that the adjusted R square of the final model with five independent variables is a relatively low 0,240. Not all assumptions applicable for multiple regression analyses are fully met. The assumption of normally distributed errors is not fully met. More information about testing the assumptions of the multiple regression analysis and about the results can be found in Appendix D. Due to confidentiality not all relevant tables are included in the Appendix.

According to our analysis the higher the combination QMuM ratio, the higher the percentage of outdating per product ($b=22,212$). An increase in the natural logarithm of the average remaining shelf life of a product leads to a lower percentage of outdating ($b=-7,880$). An increase of the coefficient of variation of sales leads to a higher percentage of outdating ($b=6,538$). An increase in the coefficient of variation of order size leads to lower percentage of outdating ($b=-5,843$). An increase in the average order size leads to a higher percentage of outdating ($b=0,184$).

Next to this regression analyses we check how many supermarket - products combinations have a structural waste problem. For these combinations the average time needed to sell the case pack size in a supermarket was longer than the average remaining shelf life of the products. This holds for 3456 combinations in this category. Grocers have on average a higher absolute and relative amount of combinations with structural waste than entrepreneurs.

We continue analysis 3 by judging why the differences in the significant variables exist between the combinations and if the values of the variables can be changed. The size of the standardized regression coefficients tells us which factors are the most important factors for predicting the percentage of outdating of a combination. The most important factor for the percentage of outdating is the combination QMuM ratio. The strong positive relationship ($b=22,212$) is in line with our hypothesis. As explained in chapter 4.1.3 the combination QMuM ratio for a product is partly changeable by the supermarket, partly by PLUS Retail and partly by HB. Ways to change this variable can be found in chapter 4.1.3 as well.

The second most important factor is the coefficient of variation of sales. As described in chapter 4.1.3, no hypothesis is formulated about this factor as the average coefficient of variation of sales on its own doesn't say a lot about the outdating of a product. This factor is changeable by the supermarket and by PLUS Retail. Ways to change these two factors can be found in chapter 4.1.3.

The third most important factor is the natural logarithm of the average remaining shelf life. The strong negative relationship ($b=-7,880$) is in line with our hypothesis. The average remaining shelf life is only changeable by HB. As described in chapter 4.1.3, this factor is very difficult to change at acceptable costs.

The fourth most important factor is the coefficient of variation of order size. The negative relationship ($b=-5,843$) is in contrast to our hypothesis. This factor is changeable by the supermarkets by placing orders of the right order size. According to our model, the percentage of outdating will decrease when the average coefficient of variation of order size increases. This is not correct. The fifth most important factor is the average order size. This factor is totally under control of a PLUS supermarket. No hypothesis is formulated about this variable as the average order size on its own

doesn't say a lot about the outdating on product level. The positive relationship ($b=0,184$) is likely to come from products with a large case pack size and a high percentage of outdating.

For the average sales no significant relationship was found. This doesn't mean that there is absolutely no relationship, it only means that with our analysis the relationship cannot be found. This is probably because the effect of the variable on the percentage of outdating is too small to be noticeable with our number of observations.

4.2.4 Analysis 4: extra analysis on supermarket level

We continue our analysis of the meat products and salads category by running a multiple regression analysis with as dependent variable the percentage of outdating per supermarket. The six variables that were relevant for the meat products and salads category in analysis 1 and the six new variables as described in chapter 2.4.4 will be used as independent variables.

We are allowed to use the enter method, as we have good theoretical reasons for including those predictors (Field, 2009). However we go from a dataset with 233 supermarket to a dataset of 34 supermarkets. The b -values are likely to change when going from 233 to 34 supermarkets, so we have to use the b -values following from the model of analysis 1. We also use the constant of analysis 1. This leads to a model with an adjusted R^2 of $-0,031$. This negative value is coming due to one supermarket with a very high percentage outdating. This supermarket is removed from the dataset. The new model has an adjusted R^2 of $0,599$. Next we use the stepwise method to test which of the six new variables could be included in the model. As described in chapter 2.4.4, using the backward method is not possible for analysis 4 as we only have 33 observations per variable. Using the stepwise method shows that none of the six variables will be included. The results can be found in Appendix D. Due to confidentiality not all relevant tables are included in the Appendix.

The only variable that maybe has an effect on the amount of outdating is the average distance between the top of a product and the next shelf, this variable has a significance of $0,060$. Adding this variable to our model shows that a higher average distance leads to less outdating. This is not possible, because grabbing behavior will stay the same or will be conducted more when the average distance is increased.

So, we have to conclude that none of the six new variables has a significant influence on the percentage of outdating of perishables in the meat products and salads category. This doesn't mean that there is absolutely no relationship, it only means that with our analysis the relationship cannot be found. This is probably because the effect of each of the six new variables on the percentage of outdating is too small to be noticeable with our small number of observations.

5 Conclusions and recommendations

In this chapter the research question and sub-questions will be answered by combining the analysis from interviews and the analysis from data.

5.2 Analysis from interviews

5.2.2 Opinion about outdating

Two types of supermarket manager were classified: grocer and entrepreneur. They have a total different opinion on outdating. Grocers typically want to get the lowest percentage outdating as

possible, whereas entrepreneurs take outdating as a fact of doing business. Entrepreneurs argue that focusing too much on outdating will result in more out-of-stocks and an assortment that is too small. Some entrepreneurs even order more products than advised from the automated store ordering system to get a better store image. They simply accept that this will result in more outdating.

7 out of the 37 managers (18,9%) are classified as grocers. The remaining 30 of the 37 managers (81,1%) are classified as entrepreneurs. We cannot extrapolate this percentage to all PLUS supermarkets as this classification can only be made after an interview with the manager of the supermarket. Grocers seem to be active in smaller supermarkets than entrepreneurs.

Waste of products is in the current situation the full responsibility of the supermarket manager. Some managers argue that this is not fair as the category management of PLUS Retail initially determines which products should be sold in the supermarkets and which products shouldn't. Often the category management makes shelf plans with a too broad assortment for the revenue that a supermarket has. Shelf plans are now based on the meters of refrigerators instead of revenue of a supermarket. This results in outdating of products. Because of this, the supermarket managers argue that the category management should be partly responsible for the waste in the PLUS supermarkets.

5.2.3 Opinion about selling products with a remaining shelf life of one day

Supermarkets can decide themselves whether they want to sell products with a remaining shelf life of one day or not. Following from the interviews, PLUS managers totally disagree about this topic. On the one hand managers, mostly grocers, argue that selling these products is necessary to run a profitable supermarket. On the other hand, the grocers and entrepreneurs that don't sell these products argue that selling these products is not in line with the core values of PLUS. 71,4% of the grocers sell these products in contrast to 46,7% of the entrepreneurs. In the supermarkets that sell these products, these products are mostly discounted on the whole or part of the day of expiration.

5.2.4 Opinion about discounting

Not in all PLUS supermarkets products with a short remaining shelf life are discounted. The managers who discount products see discounting products with a short remaining shelf life as a way to increase the chance that these products will be sold before they expire and in this way prevent outdating. Furthermore, managers argue that to prevent outdating using the right parameter settings in the automated store ordering system is way more important than discounting products. The way in which managers discount their products differs a lot between supermarkets.

Almost all interviewees argue that discounted products are not beneficial for the store image. That is why managers will set up a local promotion if a lot of products are left over from countrywide promotions. The opinion that discounting products is not beneficial for the store image is also the reason why some managers decide to don't discount at all. In an estimated 14,5% of all PLUS supermarkets products are never or seldom discounted when they have a short remaining shelf life. Next to the decision to discount products or not, some managers decide to discount only products in certain categories.

For the supermarkets that do discount their product with a short remaining shelf life, the discount percentage of the perishables differs a lot between and sometimes also within supermarkets. The minimum discount percentage observed during the supermarket visits is 25%, the maximum discount

percentage observed is 50%. The average discount percentage per category in a supermarket is around 35%. This is also the most chosen percentage in all category – supermarket combinations.

Also the moment of discounting differs between supermarkets. The number of discount hours is the total opening hours of a supermarket in which discounted products are sold. Supermarkets choose the moment of discounting products mostly based on available employees and intuition. Products are not discounted at the same time each day of the week as supermarkets don't want to sell products for a discounted price which would be sold for a normal price as well.

5.2.5 Opinion about grabbing behavior

In all supermarkets grabbing behavior is observed. Grabbing behavior occurs when consumers buy products with a later best-before-date than the first product in the shelf with a shorter best-before-date given that there are two or more best-before-dates in the shelf.

The amount of observed grabbing behavior differs a lot between supermarkets, but we cannot compare supermarkets as no quantitative data is available about the grabbing behavior. Managers argue that part of the grabbing behavior can be prevented when customers have trust in the supermarket. However, some customers will always check the full shelf for products with the longest remaining shelf life. Especially managers of supermarkets in lower-income neighborhoods notice a lot of grabbing behavior. They argue that their customers have plenty of time to check all the products in the shelves. In 62,2% of the visited supermarkets rules exist for the amount of different best-before-dates that may be placed in the shelves. In these supermarkets, the managers want at maximum one or two different dates in the shelves. They will keep products in the stockroom if necessary.

Some supermarkets have springs in the shelves of perishables. Only in the categories cheese, convenience, meat products and salads ZB, potatoes, fruit and vegetables springs can be found. Managers argue that the choice for springs is mainly based on store image and financial motives rather than on preventing grabbing behavior. In some categories no springs are placed because pushing the products is harmful for the products.

5.3 Analysis from data

5.3.2 Analysis on supermarket level

Analysis on supermarket level gives in both categories clear insights in the factors that lead to outdating on supermarket level. In both analyses the average ratio between the case pack sizes and the average sales during the shelf lives of the products of a supermarket (supermarket QMuM ratio) is the most important factor for the outdating on supermarket level. The supermarket QMuM ratio has a strong positive relationship with the percentage outdating on supermarket level in both categories. This is in line with our hypothesis and the opinion of the interviewed supermarket managers. This positive relationship is stronger in the convenience category than in the meat products and salads category. The average shelf life at arrival is more than two times higher for the meat products and salads category than for the convenience category. This explains why the supermarket QMuM ratio is more important for the convenience category.

The supermarket QMuM ratio can partly be influenced by a supermarket and is partly unchangeable. By removing products from the assortment with a high product QMuM ratio, the supermarket

QMUM ratio for a supermarket decreases and the percentage outdating will decrease as well. Another solution is to increase the average sales by getting more consumers in the supermarket. This is however difficult to realize. A third possible solution is to decrease the case pack size, so that a supermarket needs less days to sell a case pack. This has however implications for HB and their suppliers as well. This will probably also result in more different best before dates in the shelves. The fourth solution is to increase the average remaining shelf life. This can only be done by HB. They should get the products with the longest remaining shelf life from their suppliers and should keep just enough stock to serve all demand from the supermarkets. The average remaining shelf life is very difficult to change at acceptable costs.

The second most important factor in both categories is the average coefficient of variation of order size. This factor has in both categories a positive relationship with the percentage of outdating. This is in line with our hypothesis. This factor is changeable by the supermarkets by placing orders of the right order size. If the ordering pattern is more stable and the order size is more in line with the expected demand, the variation of the order size will decrease and the percentage of outdating will decrease according to our model.

The third most important factor in both categories is the average order size. This factor has in both categories a negative relationship with the percentage of outdating. This factor is totally under control of a PLUS supermarket. No hypothesis is formulated about this variable as the average order size on its own doesn't say a lot about the outdating of a supermarket. The negative relationships are likely to be coming from the differences in revenue of supermarkets. The products in the assortments have different case pack sizes and not all supermarkets order one case pack size of a product per order. Large supermarkets will order more than one case pack size of a product and small supermarkets will only order one case pack size of a product.

The percentage of order advises that is followed is significant in both categories. This factor is totally under control of a PLUS supermarket. However, in both categories the significance of this factor is just low enough to be included in our model. This means that we have to be careful about conclusions based on this factor. In the convenience category we see a strong positive relationship which is in contrast to our hypothesis. We expect that ordering based on statistical basis will result in less outdating than ordering based on emotion and intuition. The positive relationship found in the data is possibly coming from the fact that the automated store ordering system of PLUS doesn't include weather forecasts, promotions and seasonal effects in the order advises. Next to this, the system doesn't have a tradeoff between waste and service level. Delaying a planned order is likely to result in less outdating and a lower service level. Weather changes, promotions and seasonal effects will have an influence on the sales. Because of this, these effects should be predicted so that order advises can be changed only if needed. In the meat products and salads category we see a small negative relationship. The percentage of the total revenue coming from promotions in the convenience category is 1,6 times more than in the meat products and salads category. Next to this, the sales in the convenience category are more sensitive to weather changes. This two reasons combined explain why the order advises can be trusted more in the meat products and salads category than in the convenience category.

The category revenue share has in both categories a negative relationship with the percentage of outdating. This is in line with our hypothesis. However, in both categories the significance of this

factor is just low enough to be included in our model. This means that we have to be careful about conclusions based on this factor. The negative relationship is much stronger in the convenience category ($b=-35,132$) than in the meat products and salads category ($b=-0,142$). The average remaining shelf life at arrival is more than two times higher for the meat products and salads category than for the convenience category. This explains why the category revenue share is more important for the convenience category. This factor is partly changeable by a PLUS supermarket. By having the right assortment for the customers, they will buy more products in a category and the category revenue share will increase. However, part of the category revenue share is not changeable as the category revenue share has a strong relationship with the customer groups of a supermarket. These customer groups are not changeable as they are mainly dependent on the location of the supermarket.

Furthermore, some variables are only significant for one of the two categories. The average remaining shelf life has a positive relationship with the percentage of outdating in the convenience category. The positive relationship is in contrast to our hypothesis. We expect that supermarkets with a longer average remaining shelf life have a lower percentage of outdating. The positive relationship in the data is possibly coming from the different assortments that PLUS supermarkets have. Each product has its own average remaining shelf life and a unique combination of products will lead to a unique average remaining shelf life per supermarket.

The sales area has a small positive relationship with the percentage of outdating in the convenience category. An increase of the sales area of 100 square meters will result in an increase of 0,1 percent of the percentage of outdating. This positive relationship is in contrast to our hypothesis. This positive relationship is possibly coming from the fact that large supermarkets have a higher turnover per square meter, but also a broader assortment with some unpopular products which result in outdating.

The average coefficient of variation of sales has a negative relationship with the percentage of outdating in the meat products and salads category. No hypothesis was formulated about this factor as the average coefficient of variation of sales on its own doesn't say a lot about the outdating of a product. When a certain coefficient of variation of sales is expected due to the week pattern it doesn't have to result in outdating as orders are based on these expectations. Because of this no further attention will be paid to this variable.

5.3.3 Analysis on product level

Analysis on product level gives in both categories clear insights in the factors that lead to outdating on product level. Three factors are significant in both analyses. In both analyses the average ratio between the case pack size and the average sales during the shelf lives of the product (product QMuM ratio) is an important factor for the outdating on product level. The product QMuM ratio has a strong positive relationship with the percentage outdating on product level in both categories. This is in line with our hypothesis and the opinion of the interviewed supermarket managers.

The product QMuM ratio is partly changeable by the supermarket, partly by PLUS Retail and partly by HB. By removing the product from the assortment of a supermarket with a high combination QMuM ratio for this product, the product QMuM ratio for a product decreases and the percentage outdating will decrease as well. Another solution is to increase the average sales by stimulating the product. This is however difficult to realize and will probably influence sales of other products. A third solution

is to decrease the case pack size, so that a supermarket needs less days to sell a case pack. This has however implications for HB and their suppliers as well. This will probably also result in more different best before dates in the shelves. The fourth solution is to increase the average remaining shelf life. This can only be done by HB. They should get the products with the longest remaining shelf life from their suppliers and should keep just enough stock to serve all demand from the supermarkets. The average remaining shelf life is very difficult to change at acceptable costs.

The average coefficient of variation of sales has a positive relationship with the percentage of outdating in both categories. No hypothesis was formulated about this factor as the average coefficient of variation of sales on its own doesn't say a lot about the outdating of a product. When a certain coefficient of variation of sales is expected due to the week pattern it doesn't have to result in outdating as orders are based on these expectations. Because of this no further attention will be paid to this variable.

The average sales has a positive relationship with the percentage of outdating in both categories. This is in contrast to our hypothesis. A positive relationship is not possible as higher average sales can never lead to a higher percentage of outdating. The only possible reason for the found relationship is that products are sold in specific supermarkets and not in all supermarkets. Comparing products with a different number of selling points can maybe lead to a wrong comparison and thus a wrong conclusion.

Furthermore, some variables are only significant for one of the two categories. The natural logarithm of the average remaining shelf life has a negative relationship with the percentage of outdating in the convenience category. The strong negative relationship is in line with our hypothesis. The average remaining shelf life of a product is only changeable by HB. They should get the products with the longest remaining shelf life from their suppliers and should keep just enough stock to serve all demand from the supermarkets. This factor is very difficult to change at acceptable costs. As the average shelf life at arrival is more than two times higher for the meat products and salads category than for the convenience category, it is reasonable that this variable is only significant for the convenience category.

The average coefficient of order size has in a negative relationship with the percentage of outdating in the convenience category. This factor is totally under control of a PLUS supermarket. The strong negative relationship is in contrast to our hypothesis. This factor is changeable by the supermarkets by placing orders of the right order size. If the order size is more in line with the expected demand, the variation of the order size will decrease and the percentage of outdating will increase according to our model. This is not correct.

The case pack size has a negative relationship with the percentage of outdating in the meat products and salads category. No hypothesis was formulated about this variable as the average order size on its own doesn't say a lot about the outdating of a product. The negative relationship is likely to come from the differences in assortment of supermarkets. This factor is changeable by HB and their suppliers. Decreasing the case pack size will increase the chance that a case pack will be sold within the remaining shelf life of a case pack. However, smaller case packs will lead to extra handling cost at the suppliers, HB and the PLUS supermarkets. Next to this, it is likely that smaller case packs will lead to more batches in the shelves and thus in more grabbing behavior as noticed by the supermarket managers.

The average order size has a small positive relationship with the percentage of outdating in the meat products and salads category. This factor is totally under control of a PLUS supermarket. No hypothesis was formulated about this variable as the average order size on its own doesn't say a lot about the outdating on product level. The positive relationship is likely to come from the differences in assortment of supermarkets. The products in the assortments have different case pack sizes and not all supermarkets order one case pack size of a product per order.

5.3.4 Analysis on supermarket – product combination

Analysis on supermarket-product combinations gives in both categories clear insights in the factors that lead to outdating in a supermarket-product combination. Five factors are significant in both analyses. In both analyses the ratio between the case pack size and the average sales during the shelf lives of the product (combination QMuM ratio) is the most important factor for the outdating of a combination. The combination QMuM ratio has a strong positive relationship with the percentage outdating in a combination in both categories. This is in line with our hypothesis and the opinion of the interviewed supermarket managers. This positive relationship is stronger in the convenience category than in the meat products and salads category. The average remaining shelf life at arrival is more than two times higher for the meat products and salads category than for the convenience category. This explains why the combination QMuM ratio is more important for the convenience category. Ways to change the combination QMuM ratio can be found in chapter 5.3.3.

The second factor that is significant in both categories is the natural logarithm of the average remaining shelf life. The natural logarithm of the average remaining shelf life has a negative relationship with the percentage outdating in a combination in both categories. This negative relationship is stronger in the convenience category than in the meat products and salads category. This is reasonable due to the differences in remaining shelf lives. More information about changing this factor for a supermarket-product combination can be found in chapter 5.3.3.

The third factor that is significant in both categories is the coefficient of variation of sales. No hypothesis was formulated about this factor as the average coefficient of variation of sales on its own doesn't say a lot about the outdating of a product. When a certain coefficient of variation of sales is expected due to the week pattern it doesn't have to result in outdating as orders are based on these expectations. Because of this no further attention will be paid to this variable.

The fourth factor that is significant in both categories is the coefficient of variation of order size. The coefficient of variation of order size has a negative relationship with the percentage outdating in a combination. This negative relationship is stronger in the convenience category than in the meat products and salads category. The negative relationships are in contrast to our hypothesis. This factor is changeable by the supermarkets by placing orders of the right order size. If the order size is more in line with the expected demand, the variation of the order size will decrease and the percentage of outdating will increase according to our model. This is not correct.

The fifth factor that is significant in both categories is the average order size. This factor is totally under control of a PLUS supermarket. No hypothesis was formulated about this factor as the average order size on its own doesn't say a lot about the outdating of a product.

Furthermore, two factors are only significant for the convenience category. The average sales has a positive relationship with the percentage of outdating of a combination. This is in contrast to our

hypothesis. A positive relationship is not possible as higher average sales can never lead to a higher percentage of outdating. The only possible reason for the found relationship is that products are sold in specific supermarkets and not in all supermarkets. Comparing products with a different number of selling points can maybe lead to a wrong comparison and thus to a wrong conclusion.

The case pack size has a positive relationship with the percentage of outdating of a combination in the convenience category. This factor is changeable by HB and their suppliers. Decreasing the case pack size will increase the chance that a case pack will be sold within the remaining shelf life of a case pack. However, smaller case packs will lead to extra handling cost at the suppliers, HB and the PLUS supermarkets. Next to this, it is likely that smaller case packs will lead to more batches in the shelves and thus in more grabbing behavior as noticed by the supermarket managers. No hypothesis was formulated about this factor as the case pack size on its own doesn't say a lot about the outdating of a product. For example, a large case pack size only results in more outdating when the average sales are low and/or the shelf life of the products is short.

5.3.5 Extra analysis on supermarket level

Both analyses on supermarket-product combinations show that supermarkets classified as grocers have on average a higher absolute and relative number of combinations with structural waste than supermarkets classified as entrepreneurs.

Both extra analysis on supermarket level with detailed information about whether they sell product with a remaining shelf life of one day, the type of manager, the discount percentage, the discount hours and the average distance didn't show any significant variables. So, we have to conclude that none of these factors has a significant influence on the percentage of outdating of perishables in the convenience or the meat products and salads category. This doesn't mean that there is absolutely no relationship, it only means that with our analysis the relationship cannot be found. This is probably because the effect of each of these factors on the percentage of outdating is too small to be noticeable with our small number of observations.

5.3.6 Outdating of perishables; part of doing business?

Based on the analyses in this report we can conclude that a large part of the outdating of perishables is coming from the width of the assortment of PLUS supermarkets. By choosing for a broad assortment of perishables, PLUS supermarkets choose to include perishables in the assortment with a large chance of outdating. So, part of the outdating of perishables is part of doing business in a PLUS supermarket.

However, this doesn't mean that the percentage of outdating should be accepted as a norm and taken for granted. PLUS Retail, HB and their entrepreneurs should still put effort in ways to reduce the current amount of outdating. Especially supermarkets with a high revenue should set a low norm for the percentage of outdating. They should be able to achieve a low percentage of outdating as they are able to achieve lower QMuM ratios for their products than supermarkets with a low revenue. In the following paragraph ways to reduce the amount of outdating will be explained. A lower amount of outdating will lead to a higher profitability of the supermarkets and a better image of PLUS as a sustainable supermarket chain.

5.3.7 Ways to reduce outdating

Based on the aforementioned conclusions some ways to reduce outdating of perishables can be formulated. As a first step, the registration of sales and outdating should become more accurate in order to be able to identify factors that cause outdating earlier. This also leads to more accurate settings in the automated store ordering system. The right parameter settings in the automated store ordering system are way more important than responding on left over products by discounting products.

The next step is to choose the right assortment for a supermarket. Following from the data analysis the ratios between the case pack sizes and the average sales during the shelf lives of the products (QMUM ratios) are the most important factors for outdating of perishables. This also follows from the fact that the 10% of the assortment consisting of the articles with the most outdating result in 36 till 50 percent of the total outdating per category. Shelf plans are now based on the meters of refrigerators instead of revenue of a supermarket. In some supermarkets this leads to products with a structural waste problem. Supermarkets should be critical about the width of the assortment that is possible in their supermarket. Products with a lot of outdating should only be kept in the assortment when the loss of customers coming for these products is larger than the savings in outdating.

The third step is only relevant for categories with strong weather, promotional or seasonal effects like the convenience category. The automated store ordering system of PLUS doesn't include weather forecasts, promotions and seasonal effects in the order advises. Because of this, these effects should be predicted so that order advises can be changed only if needed.

The fourth step is the right placing and handling of perishables. Unnecessary outdating due to grabbing behavior or wrong handling should be avoided. Based on the regression analyses no relationship could be found between the average distance between products and the next shelf and the percentage outdating. However, employees argue that a greater distance between the shelves will make it easier for customers to check the remaining shelf lives of products and perform grabbing behavior. Smaller distances will prevent this undesirable behavior. Next to this, some shelves are stacked too full what will result in damaged products. Finally, some products were placed in a non-optimal environment which will speed up the decay process of the products.

Note that an important condition for the success of the formulated approach is that the delivered products should be of good quality and the remaining shelf life is as long as possible.

The registration of sales and outdating, and the settings in the automated store ordering system are the responsibility of the supermarkets themselves. The choice of the right assortment is a common responsibility of the supermarket and PLUS Retail. Finally, HB is only for a small part responsible for the waste in the supermarkets. HB is only responsible for the waste if the delivered quality is not good or the remaining shelf life at arrival is shorter than agreed with PLUS Retail.

6 Limitations

One of the limitations of this project is that the average sales are determined by using point of sales data from PLUS Retail. In other words, the assumption is made that the sales exactly represent the demand. However, the average sales don't take into account the sales of discounted products. Next to this, for the days with zero sales for a supermarket-product combination it was not known if there

was really no demand or that there was demand but no products were available. For example, products could be out of stock. In situations where the number of days without sales was higher than the average remaining shelf life at arrival, these days were noted as days where no product was available and excluded from the average sales calculations. In other cases we assume that there were products available, but no demand. The effect of the demand in the week after promotion is not taken into account. Lower sales after a promotion period will result in an incorrect estimate of the average sales and the coefficient of variation of sales. Finally, the project did not take into account how many Sundays a supermarket was open during the period of our data. For supermarkets that were not the full year or not at all open on Sundays the average sales were slightly underestimated.

The second limitation of this project is the data accuracy of the waste data. Not all waste is registered in the correct waste group. Most supermarkets don't register all the products that are thrown away, so if we only take the registered waste, the amount of outdating is largely underestimated. Because of this, we also include the stock differences in our outdating data. We assume that all waste registered as stock differences can be seen as outdating. However, part of the stock differences is not correct registered waste such as broken products or theft by employees or customers. Because of this, in the end the outdating in supermarkets was slightly overestimated.

7 Further research

The first recommendation for future research is to carry out a similar project for more categories. Especially categories with a high percentage of waste such as the potatoes, fruit and vegetables category and the meat category are interesting categories to investigate.

The second recommendation is to visit more supermarket throughout the whole country to check whether the found results are representative for all PLUS supermarkets. A new analysis is then possible in which can be tested whether factors as discount percentage and type of manager have a significant effect on the percentage of outdating.

The third recommendation is to include some variables that were left out our analysis in future analyses. The left out variables can be found in chapter 2.4.5. For example, the quality of the ordering and replenishment process probably have an effect on the percentage of outdating. These possible effects are interesting to investigate.

The fourth recommendation is to conduct more empirical research about the consumer choice with regard to perishables. Current literature assumes that more inventory will increase the visibility of the product, gives confidence to the customer and will in this way stimulate sales. Field studies have to be done to check whether this holds for perishables or not.

8 References

- Broekmeulen, R.A.C.M., van Donselaar, K.H. "A heuristic to manage perishable inventory with batch ordering, positive lead-times, and time-varying demand". *Computers & Operations Research* 2009; 36(11); p. 3013-3018.
- Broekmeulen, R.A.C.M., van Donselaar, K.H. "A satisficing choice model of withdrawal behavior of perishable items". ISIR 2014
- Buzby, J. C., Hyman, J., Stewart, H., Wells, H. F. "The Value of Retail-and Consumer-Level Fruit and Vegetable Losses in the United States". *Journal of Consumer Affairs*, 2011, 45(3), 492-515. Chicago
- Cohen, M. A. "Analysis of Single Critical Number Order Policies in Perishable Inventory Theory". *Operations Research*. 1976, 24, 726-741.
- Cohen, M.A., Prastacos, G.P. "Critical number ordering policy for LIFO perishable inventory systems". *Computers & Operations Research*. 1981;8(3):185-95.
- Dagnoli, J. "Impulse Governs Shoppers". *Advertising Age*, 1987, October 5, 93
- Deniz, B., Scheller-Wolf, A., Karaesmen, I. "Managing inventories of perishable goods: The effect of substitution". GSIA Working Paper# 2004-E55, Carnegie-Mellon University, Pittsburgh, PA.
- Deniz, B., Karaesmen, I., Scheller-Wolf, A. "Managing perishables with substitution: inventory issuance and replenishment heuristics". *Manufacturing & Service Operations Management*, 2010, 12(2), 319-329.
- Dickson, Peter R., Sawyer, A.G. "The Price Knowledge and Search of Supermarket Shoppers". *Journal of Marketing*, 1990, 54 (July), 42-53.
- Entrup, M. L. *Advanced planning in fresh food industries: integrating shelf life into production planning*. Springer. 2006.
- Ferguson, M.E., Ketzenberg, M.E. *Information sharing to improve retail product freshness of perishables*. Georgia Institute of Technology. 2005.
- Ferguson, M. E., Koenigsberg, O. "How should a firm manage deteriorating inventory?". *Production and Operations Management*, 2007, 16(3), 306-321.
- Field, A. (2009). *Discovering statistics using SPSS*. Sage publications.
- Gunders, D. "Wasted: How America is losing up to 40 percent of its food from farm to fork to landfill." Natural Resources Defense Council Issue Paper. Chicago. 2012
- Haijema, R. "Optimal issuing of perishables with a short fixed shelf life". In: *Computational Logistics* (pp. 160-169). Springer Berlin Heidelberg. 2011.
- Hine, D. J. "Shelf-life prediction". In: *Modern Processing, Packaging and Distribution Systems for Food* (pp. 62-85). Springer US. 1987.

- Hoch, Stephen J., Deighton, J.A. "Managing What Consumers Learn From Experience". *Journal of Marketing*, 1989, 53 (2), 1-20.
- Hoyer, W.D. "An Examination of Consumer Decision Making for a Common Repeat Purchase Product". *Journal of Consumer Research*, 1984, 11 (3), 822-831.
- Jongen, L.J. Dervingreductie convenience productgroep. Bachelor Eindproject. Eindhoven University of Technology. 2013.
- Kilcast, D., Subramaniam, P. (Eds.). *The stability and shelf-life of food*. Elsevier. 2000.
- Kock, R.A.L. Literature study. Master thesis preparation 1. Eindhoven University of Technology. 2014.
- Labuza, T. P., Taoukis, P. S. "The relationship between processing and shelf life". In: *Foods for the 90's*. pp73-105. 1990.
- Nahmias, S. "Myopic Approximations for the Perishable Inventory Problem". *Management Science*. 1975, 22, 1002-1008.
- Nahmias, S. "Perishable Inventory Theory: A Review". *Operations Research*, 1982, Vol. 30, No. 4, pp. 680-708
- Nahmias, S. *Perishable Inventory Systems*. New York: Springer. 2011.
- Nijs, V. R., Dekimpe, M. G., Steenkamps, J. B. E., Hanssens, D. M. "The category-demand effects of price promotions". *Marketing science*, 2001, 20(1), 1-22.
- Maarse, H. (2015, January 24). Derving in supermarkten bijna €1 miljard. Distrifood, pp 1.
- Maarse, H., Te Pas, H. (2015, March 28). Plus wint in survival of the fittest. Distrifood, pp 7.
- Meijssen, J. (2011, September 7). Belang van derving neemt toe. Retrieved from: <http://www.distrifood.nl/Formules/Algemeen/2011/9/Belang-van-derving-neemt-toe-DIS140819W/>
- Meijssen, J. (2015, January 30). Marktaandeelen: Lidl loopt uit op Aldi. Distrifood. Retrieved from: <http://www.distrifood.nl/Formules/Algemeen/2015/1/Marktaandeelen-Lidl-loopt-uit-op-Aldi-1695412W/>
- Miranda, M. J., Kónya, L. "Promoting brands of perishable items on the promise of generous availability of consumption time". *Journal of Targeting, Measurement and Analysis for Marketing*, 2006, 14(3), 238-248.
- Soethoudt, J., van der Sluis, A., Waarts, Y. & Tromp, S. *BO-Voedselverspilling*. 2012.
- Smit, R. (2015, March 25). Supermarket Plus zegt meer op de kleintjes te gaan letten. *Het Financieele Dagblad*, pp 14.
- Tsiros, M., Heilman, C. M. "The effect of expiration dates and perceived risk on purchasing behavior in grocery store perishable categories". *Journal of Marketing*, 2005, 69(2), 114-129.

- Van Burgh, V. R. Verbetering bestelkwaliteit automatisch bestelsysteem voor beperkt houdbare producten bij een Nederlandse supermarktketen: onderzoek naar de huidige prestatie van het automatische bestelsysteem en het graaigedrag van supermarktklanten ten behoeve van mogelijke verbeteringen. Technische Universiteit Eindhoven. 2007.
- Van Donselaar, K., van Woensel, T., Broekmeulen, R. A. C. M., Fransoo, J. "Inventory control of perishables in supermarkets". *International Journal of Production Economics*, 2006, 104(2), 462-472.
- Van Donselaar, K. H., Broekmeulen, R. A.C.M. "Approximations for the relative outdating of perishable products by combining stochastic modeling, simulation and regression modeling". *International Journal of Production Economics*, 2012, 140(2), 660-669.
- Van de Ven, F.J.L. Reducing the amount of outdating for the stores of PLUS Retail by optimizing the case pack sizes for convenience products at Hollander Barendrecht BV, considering entailed additional supply chain operating costs. Eindhoven University of Technology. 2014.
- Van Strien, P. "Towards a methodology of psychological practice." *Theory and Psychology*, 1997, Vol.7 (5), pp. 683 - 700.
- Van Woensel, T., Van Donselaar, K., Broekmeulen, R.A.C.M, Fransoo, J. "Consumer responses to shelf out-of-stocks of perishable products". *International Journal of Physical Distribution & Logistics Management*, 2007, 37(9), 704-718
- Weteling, R. Improving Freshness: The effects of week patterns and opening terms on outdating at Albert Heijn. Technische Universiteit Eindhoven. 2013
- Whitin, T.M. The theory of inventory management. Princeton University Press, Princeton, NJ. 1957.
- Wolfe, H.B. "A model for control of style merchandise". *Industrial Management Review*, 1968, 9, pp 69.
- Zhou, Y., & Yang, S. L. "An optimal replenishment policy for items with inventory-level-dependent demand and fixed lifetime under the LIFO policy". *Journal of the Operational Research Society*, 2003, 54(6), 585-593.

9 Appendices

9.1 Appendix A: percentage waste per category

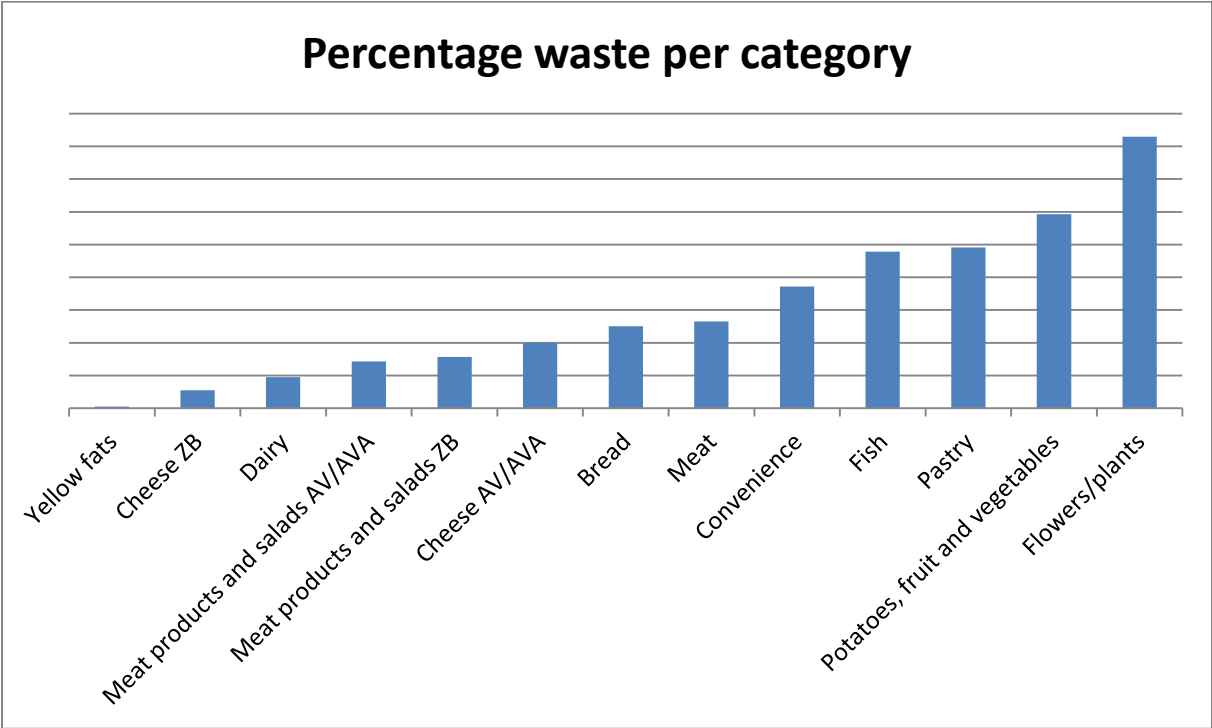


Figure 5: Overview of percentage waste per category

9.2 Appendix B: description of variables

Table 1: Description of variables

c	Category number (1=convenience, 2= meat products and salads ZB)
d	Day number
i	Shelf number
p	Product number
s	Supermarket number
w	Week number
$crs_{c,s}$	Revenue share of category c in supermarket s
$dp_{c,s}$	Discount percentage in category c in supermarket s
$dh_{c,s}$	Discount hours in category c in supermarket s
$manager_s$	Type of manager (grocer=0, entrepreneur=1) in supermarket s
np_c	Number of different products in category c
$np_{c,s}$	Number of different products in category c in supermarket s
ns_p	Number of supermarkets for product p
$o_{d,p,s}$	Order in consumer units on day d for product p for supermarket s
$oa_{c,s,w}$	Total number of automatic order advises in week w for category c in supermarket s
$oac_{c,s,w}$	Total number of changed automatic order advises in week w for category c in supermarket s
$oaf_{c,s}$	Percentage of automatic order advises that is followed in category c in supermarket s
od_s	Number of opening days per week for supermarket s (6 days=0, 7 days=1)
$ol_{d,p,s}$	Order line on day d for product p for supermarket s
$productstoday_s$	Selling products with one day remaining shelf life (no=0, yes =1) in supermarket s
$promdays1_{p,s}$	Number of days with countrywide promotion for product p in supermarket s
$promdays2_{p,s}$	Number of days with local promotion for product p in supermarket s
$Q_product_p$	Case pack size for product p
$Q_combination_{p,s}$	Case pack size for product p in supermarket s
sa_s	Sales area (verkoopvloeroppervlakte in Dutch) of supermarket s
$sales_units_{d,p,s}$	Sales in consumer units on day d for product p in supermarket s without countrywide or local promotions
$sales_euros_{d,p,s}$	Sales in Euros on day d for product p in supermarket s without countrywide or local promotions

$sellingdays1_{p,s}$	Regular selling days for product p in supermarket s
$sellingdays2_{p,s}$	Days with positive sales for product p in supermarket s
$sellingdays3_{p,s}$	Days with zero sales for product p in supermarket s
$sellingperiod_{p,s}$	Selling period in days for product p in supermarket s
$springs_{c,s}$	Presence of springs in category c in supermarket s (no=0, yes=1)
$tna_{p,s}$	Number days in which product p is temporary not available in supermarket s
$outdating_euros_{d,p,s}$	Waste in Euros on day d for product p in supermarket s from outdating registered with reason codes: 'Overcode', 'Bederf', 'Afprijzen voorraad correctie' and 'Derving voorraad correctie'.
$total_outdating_euros_supermarket_{c,s}$	Total waste in Euros in category c in supermarket s
$total_outdating_euros_product_p$	Total waste in Euros for product p
$total_outdating_euros_combination_{p,s}$	Total waste in Euros for product p in supermarket s
$total_sales_euros_supermarket_{c,s}$	Total sales in Euros in category c in supermarket s
$total_sales_euros_product_p$	Total sales in Euros for product p
$total_sales_euros_combination_{p,s}$	Total sales in Euros for product p in supermarket s
$total_sales_units_supermarket_{c,s}$	Total sales in consumer units in category c in supermarket s
$total_sales_units_product_p$	Total sales in consumer units for product p
$total_sales_units_combination_{p,s}$	Total sales in consumer units for product p in supermarket s
$perc_outdating_supermarket_{c,s}$	Percentage outdating in category c in supermarket s
$perc_outdating_product_p$	Percentage outdating for product p
$perc_outdating_combination_{p,s}$	Percentage outdating for product p in supermarket s
$avg_o_supermarket_{c,s}$	Average order size in category c in supermarket s
$avg_o_product_p$	Average order size for product p
$avg_o_combination_{p,s}$	Average order size for product p in supermarket s
$stdev_o_combination_{p,s}$	Standard deviation of order size for product p in supermarket s
$total_ol_combination_{p,s}$	Total number of order lines for product p in supermarket s
$avg_coefvar_o_supermarket_{c,s}$	Average coefficient of variation of order size in category c in supermarket s
$avg_coefvar_o_product_p$	Average coefficient of variation of order size for product p

$coefvar_o_combination_{p,s}$	Coefficient of variation of order size for product p in supermarket s
$avg_Q_supermarket_{c,s}$	Average case pack size in category c in supermarket s
$\mu_supermarket_{c,s}$	Average sales in consumer units per day for product p in supermarket s
$\mu_product_p$	Average sales in consumer units per day for product p
$\mu_combination_{p,s}$	Average sales in consumer units for product p in supermarket s
$stdev_sales_combination_{p,s}$	Standard deviation of sales in consumer units for product p in supermarket s
$avg_coefvar_sales_supermarket_{c,s}$	Average coefficient of variation of sales in consumer units in category c in supermarket s
$avg_coefvar_sales_product_p$	Average coefficient of variation of sales in consumer units for product p
$coefvar_sales_combination_{p,s}$	Coefficient of variation of sales in consumer units for product p in supermarket s
avg_di_s	Average disposable income in neighborhood of supermarket s
$avg_serv_{c,s}$	Average service level of HB in category c in supermarket s
$avg_sl_supermarket_{c,s}$	Average remaining shelf life at arrival in category c in supermarket s
$LN_avg_sl_supermarket_{c,s}$	Natural logarithm of average remaining shelf life at arrival in category c in supermarket s
$avg_sl_product_p$	Average remaining shelf life at arrival for product p
$LN_avg_sl_product_p$	Natural logarithm of average remaining shelf life at arrival for product p
$avg_sl_combination_{p,s}$	Average remaining shelf life at arrival for product p in supermarket s
$LN_avg_sl_combination_{p,s}$	Natural logarithm of average remaining shelf life at arrival for product p in supermarket s
$Q_mu_m_supermarket_{c,s}$	QMUM ratio in category c in supermarket s
$Q_mu_m_product_p$	QMUM ratio for product p
$Q_mu_m_combination_{p,s}$	QMUM ratio for product p in supermarket s
$avg_distance_{c,s}$	Average distance in category c in supermarket s

9.3 Appendix C: reasoning and calculation of variables

In this chapter all necessary calculations will be showed for calculating the values of the variables mentioned in chapter 9.3.

9.4.1 Orders

$$avg_o_combination_{p,s} = \frac{\sum_{d=0}^{\infty} o_{d,p,s}}{\sum_{d=0}^{\infty} ol_{d,p,s}} \quad (3)$$

$$avg_o_supermarket_{c,s} = \frac{\sum_{p=0}^{\infty} avg_o_combination_{p,s}}{np_{c,s}} \quad \text{if } p \in c \quad (4)$$

Equation 4 will be calculated two times per supermarket. One time for all products in the convenience category ($c=1$) and one time for all products in the meat products and salads category ($c=2$). To make a fair comparison between supermarkets, we divide the sum of the average order sizes for all products p in the assortment of supermarket s by the number of products that supermarket s has in its assortment for category c . Calculating the average order size of supermarket s by dividing the sum of ordered units of supermarket s by the sum of the order lines of supermarket s should not result in a fair comparison because the differences in assortment between supermarkets will not be taken into account in this way.

$$avg_o_product_p = \frac{\sum_{s=0}^{\infty} avg_o_combination_{p,s}}{ns_p} \quad (5)$$

To make a fair comparison between products, we divide the sum of the average order sizes for all supermarkets s ordering product p by the number of supermarkets ordering product p . Calculating the average order size per product by dividing the sum of ordered units of product p in all supermarkets s by the sum of the order lines of product p in all supermarket s should not result in a fair comparison because the differences in ordering decisions of the supermarkets will not be taken into account in this way.

$$total_ol_combination_{p,s} = \sum_{d=0}^{\infty} ol_{d,p,s} \quad (6)$$

$$stdev_o_combination_{p,s} = \sqrt{\frac{\sum_{d=0}^{\infty} (o_{d,p,s} - avg_o_combination_{p,s})^2}{total_ol_combination_{p,s}}} \quad (7)$$

$$coefvar_o_combination_{p,s} = \frac{stdev_o_combination_{p,s}}{avg_o_combination_{p,s}} \quad (8)$$

$$avg_coefvar_o_supermarket_{c,s} = \frac{\sum_{p=0}^{\infty} coefvar_o_combination_{p,s}}{np_{c,s}} \quad \text{if } p \in c \quad (9)$$

Equation 9 will be calculated two times per supermarket. One time for all products in the convenience category ($c=1$) and one time for all products in the meat products and salads category ($c=2$). To make a fair comparison between supermarkets, we divide the sum of the coefficient of variation of the order size for all products p in the assortment of supermarket s by the number of products that supermarket s has in its assortment for category c . Calculating the average coefficient of variation of the order size of supermarket s by taking the square root of the squared sum of differences between the order size of product p on day d for supermarket s and the average order size of supermarket s divided by the sum of the order lines of supermarket s should not result in a

fair comparison because the differences in assortment between supermarkets will not be taken into account in this way.

$$avg_coefvar_o_product_p = \frac{\sum_{s=0}^{\infty} coefvar_o_combination_{p,s}}{ns_p} \quad \text{if } p \in c \quad (10)$$

To make a fair comparison between products, we divide the sum of the coefficient of variation of the order size of product p in all supermarkets s by the number of supermarkets ordering product p . Calculating the average coefficient of variation of the order size of product p by taking the square root of the squared sum of differences between the order size of product p on day d for supermarket s and the average order size of product p divided by the sum of the order lines for product p should not result in a fair comparison because the differences in assortment between supermarkets will not be taken into account in this way.

$$avg_Q_supermarket_{c,s} = \frac{\sum_{p=0}^{\infty} Q_product_p}{np_{c,s}} \quad (11)$$

Equation 11 will be calculated two times per supermarket. One time for all products in the convenience category ($c=1$) and one time for all products in the meat products and salads category ($c=2$). To make a fair comparison between supermarkets, we divide the sum of case pack sizes for all products p in the assortment of supermarket s by the number of products that supermarket s has in its assortment for category c . Any other weighted or unweighted average should not result in a fair comparison because the differences in assortment between supermarkets will not be taken into account in a fair way.

$$Q_combination_{p,s} = Q_product_p \quad (12)$$

As for all supermarkets the case pack size of product p is the same, the above equation holds.

$$oaf_{c,s} = \frac{\sum_{w=0}^{52} (oac_{c,s,w} - oac_{c,s,w})}{\sum_{w=0}^{52} (oac_{c,s,w})} \quad (13)$$

Equation 13 will be calculated two times per supermarket. One time for all products in the convenience category ($c=1$) and one time for all products in the meat products and salads category ($c=2$).

9.4.2 Sales

$$total_sales_euros_supermarket_{c,s} = \sum_{d=0}^{\infty} \sum_{p=0}^{\infty} sales_euros_{d,p,s} \quad \text{if } p \in c \quad (14)$$

Equation 14 will be calculated two times per supermarket. One time for all products in the convenience category ($c=1$) and one time for all products in the meat products and salads category ($c=2$).

$$total_sales_euros_product_p = \sum_{d=0}^{\infty} \sum_{s=0}^{\infty} sales_euros_{d,p,s} \quad (15)$$

$$total_sales_euros_combination_{p,s} = \sum_{d=0}^{\infty} sales_euros_{d,p,s} \quad (16)$$

$$total_sales_units_supermarket_{c,s} = \sum_{d=0}^{\infty} \sum_{p=0}^{\infty} sales_units_{d,p,s} \quad \text{if } p \in c \quad (17)$$

Equation 17 will be calculated two times per supermarket. One time for all products in the convenience category ($c=1$) and one time for all products in the meat products and salads category ($c=2$).

$$total_sales_units_product_p = \sum_{d=0}^{\infty} \sum_{s=0}^{\infty} sales_units_{d,p,s} \quad (18)$$

$$total_sales_units_combination_{p,s} = \sum_{d=0}^{\infty} sales_units_{d,p,s} \quad \text{if } p \in c \quad (19)$$

$$sellingdays1_{p,s} = sellingperiod_{p,s} - toa_{p,s} - promdays1_{p,s} - promdays2_{p,s} \quad (20)$$

$$sellingdays2_{p,s} = \sum_{d=0}^{\infty} d \text{ if } sales_units_{d,p,s} > 0 \quad (21)$$

$$sellingdays3_{p,s} = sellingdays1_{p,s} - sellingdays2_{p,s} \quad (22)$$

$$\mu_combination_{p,s} = \frac{total_sales_units_combination_{p,s}}{sellingdays1_{p,s}} \quad (23)$$

$$\mu_supermarket_{c,s} = \frac{\sum_{p=0}^{\infty} \mu_combination_{p,s}}{np_{c,s}} \quad \text{if } p \in c \quad (24)$$

Equation 24 will be calculated two times per supermarket. One time for all products in the convenience category ($c=1$) and one time for all products in the meat products and salads category ($c=2$). To make a fair comparison between supermarkets, we divide the sum of average sales for all products p in the assortment of supermarket s by the number of products that supermarket s has in its assortment for category c . Calculating the average sales of supermarket s by dividing the total sales of supermarket s by the sum of the selling days1 for all products p of supermarket s should not result in a fair comparison because the differences in assortment between supermarkets will not be taken into account in a correct way.

$$\mu_product_p = \frac{\sum_{s=0}^{\infty} \mu_combination_{p,s}}{ns_p} \quad (25)$$

To make a fair comparison between products, we divide the sum of the average sales for all supermarkets s selling product p by the number of supermarkets selling product p . Calculating the average sales of product p by dividing the total sales of product p in all supermarkets s by the sum of the selling days1 of product p for all supermarkets s should not result in a fair comparison because the differences in sales of the supermarkets will not be taken into account in this way.

$$stdev_sales_combination_{p,s} = \sqrt{\frac{((0 - \mu_combination_{p,s})^2 \cdot sellingdays3_{p,s} + \sum_{d=0}^{\infty} (sales_units_{d,p,s} - \mu_combination_{p,s}))^2}{sellingdays1_{p,s}}} \quad \text{if } sales_units_{d,p,s} > 0 \quad (26)$$

The standard deviation of the sales of product p in supermarket s consists of two parts. We calculate the difference between the average sales in consumer units for product p in supermarket s and zero for the days without sales for product p in supermarket s . And we calculate the differences between the sales in consumer units on day d for product p in supermarket s without nationwide or local promotions and the average sales in consumer units for product p in supermarket s . This is done because in the sales data the days with zero sales are not correctly registered.

$$coefvar_sales_combination_{p,s} = \frac{stdev_combination_{p,s}}{\mu_combination_{p,s}} \quad (27)$$

$$avg_coefvar_sales_supermarket_{c,s} = \frac{\sum_{p=0}^{\infty} coefvar_sales_combination_{p,s}}{np_{c,s}} \quad \text{if } p \in c \quad (28)$$

Equation 28 will be calculated two times per supermarket. One time for all products in the convenience category ($c=1$) and one time for all products in the meat products and salads category ($c=2$). To make a fair comparison between supermarkets, we divide the sum of the coefficient of variation of the sales for all products p in the assortment of supermarket s by the number of products that supermarket s has in its assortment for category c . Any other weighted or unweighted average should not result in a fair comparison because the differences in assortment between supermarkets will not be taken into account in a fair way.

$$avg_coefvar_sales_product_p = \frac{\sum_{s=0}^{\infty} coefvar_sales_combination_{p,s}}{ns_p} \quad (29)$$

To make a fair comparison between products, we divide the sum of the coefficient of variation of the sales for all supermarkets s selling product p by the number of supermarkets selling product p . Any other weighted or unweighted average should not result in a fair comparison because the differences in sales of the supermarkets will not be taken into account in this way.

9.4.3 Remaining shelf life

$$avg_sl_combination_{p,s} = \frac{\sum_{d=0}^{\infty} o_{d,p,s} \cdot sl_{d,p,s}}{\sum_{d=0}^{\infty} o_{d,p,s}} \quad (30)$$

$$avg_sl_supermarket_{c,s} = \frac{\sum_{p=0}^{\infty} avg_sl_combination_{p,s}}{np_{c,s}} \quad \text{if } p \in c \quad (31)$$

Equation 31 will be calculated two times per supermarket. One time for all products in the convenience category ($c=1$) and one time for all products in the meat products and salads category ($c=2$). To make a fair comparison between supermarkets, we divide the sum of the average remaining shelf lives for all products p in the assortment of supermarket s by the number of products that supermarket s has in its assortment for category c . Any other weighted or unweighted average should not result in a fair comparison because the differences in assortment between supermarkets will not be taken into account in a fair way.

$$avg_sl_product_p = \frac{\sum_{s=0}^{\infty} avg_sl_combination_{p,s}}{ns_p} \quad (32)$$

To make a fair comparison between products, we divide the sum of the average remaining shelf lives for all supermarkets s selling product p by the number of supermarkets selling product p . Any other weighted or unweighted average should not result in a fair comparison because the differences in remaining shelf lives of the supermarkets for product p will not be taken into account in this way.

$$LN_avg_sl_combination_{p,s} = LN(avg_sl_combination_{p,s}) \quad (33)$$

$$LN_avg_sl_supermarket_{c,s} = LN(avg_sl_supermarket_{c,s}) \quad (34)$$

$$LN_avg_sl_product_p = LN(avg_sl_product_p) \quad (35)$$

9.4.4 Outdating

$$total_outdating_euros_supermarket_{c,s} = \sum_{d=0}^{\infty} \sum_{p=0}^{\infty} outdating_euros_{d,p,s} \quad \text{if } p \in c \quad (36)$$

Equation 36 will be calculated two times per supermarket. One time for all products in the convenience category ($c=1$) and one time for all products in the meat products and salads category ($c=2$).

$$total_outdating_euros_product_p = \sum_{d=0}^{\infty} \sum_{s=0}^{\infty} outdating_euros_{d,p,s} \quad (37)$$

$$total_outdating_euros_combination_{p,s} = \sum_{d=0}^{\infty} outdating_euros_{d,p,s} \quad (38)$$

9.4.5 Percentages/ratios

$$perc_outdating_supermarket_{c,s} = \frac{total_outdating_euros_supermarket_{c,s}}{total_sales_euros_supermarket_{c,s}} \quad (39)$$

$$perc_outdating_product_p = \frac{total_outdating_euros_product_p}{total_sales_euros_product_p} \quad (40)$$

$$perc_outdating_combination_{p,s} = \frac{total_outdating_euros_combination_{p,s}}{total_sales_euros_combination_{p,s}} \quad (41)$$

$$Q_mu_m_supermarket_{c,s} = \frac{avg_Q_supermarket_{c,s}}{\mu_supermarket_{c,s} \cdot avg_sl_supermarket_{c,s}} \quad (42)$$

$$Q_mu_m_product_p = \frac{Q_product_p}{\mu_product_p \cdot avg_sl_product_p} \quad (43)$$

$$Q_mu_m_combination_{p,s} = \frac{Q_combination_{p,s}}{\mu_combination_{p,s} \cdot avg_sl_combination_{p,s}} \quad (44)$$

The QMuM ratios calculate the ratios between the case pack size and the average sales during the shelf lives of products in a supermarket. This is done because of the fact that the case pack size, the average sales and the remaining shelf lives on their own don't have a lot of value for predicting the outdating in a supermarket. For example, a large case pack size only results in more outdating when the average sales are low and/or the remaining shelf life of the product is short.

9.4 Appendix D: regression analysis results

9.4.2 Convenience: analysis 1: analysis on supermarket level

Table 2: Model summary of analysis 1 in convenience category

Model Summary				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,768 ^a	,590	,566	1,959
2	,768 ^b	,590	,568	1,955
3	,768 ^c	,590	,569	1,951
4	,768 ^d	,589	,570	1,949
5	,766 ^e	,586	,569	1,951
6	,763 ^f	,582	,567	1,956
7	,760 ^g	,578	,564	1,963

a. Predictors: (Constant), Q_Mu_M_supermarket, Order_advises_followed, Servicelevel_HB, Avg_disposable_income, Avg_Q_supermarket, Openingdays, Avg_shelflife_supermarket, Category_revenue_share, Avg_o_supermarket, Sales_area, Avg_coefvar_sales_supermarket, Avg_coefvar_o_supermarket, Ln_avg_shelflife_supermarket

b. Predictors: (Constant), Q_Mu_M_supermarket, Order_advises_followed, Servicelevel_HB, Avg_disposable_income, Openingdays, Avg_shelflife_supermarket, Category_revenue_share, Avg_o_supermarket, Sales_area, Avg_coefvar_sales_supermarket, Avg_coefvar_o_supermarket, Ln_avg_shelflife_supermarket

c. Predictors: (Constant), Q_Mu_M_supermarket, Order_advises_followed, Servicelevel_HB, Avg_disposable_income, Openingdays, Avg_shelflife_supermarket, Category_revenue_share, Avg_o_supermarket, Sales_area, Avg_coefvar_o_supermarket, Ln_avg_shelflife_supermarket

d. Predictors: (Constant), Q_Mu_M_supermarket, Order_advises_followed, Servicelevel_HB, Openingdays, Avg_shelflife_supermarket, Category_revenue_share, Avg_o_supermarket, Sales_area, Avg_coefvar_o_supermarket, Ln_avg_shelflife_supermarket

e. Predictors: (Constant), Q_Mu_M_supermarket, Order_advises_followed, Servicelevel_HB, Openingdays, Avg_shelflife_supermarket, Category_revenue_share, Avg_o_supermarket, Sales_area, Avg_coefvar_o_supermarket

f. Predictors: (Constant), Q_Mu_M_supermarket, Order_advises_followed, Servicelevel_HB, Avg_shelflife_supermarket, Category_revenue_share, Avg_o_supermarket, Sales_area, Avg_coefvar_o_supermarket

g. Predictors: (Constant), Q_Mu_M_supermarket, Order_advises_followed, Avg_shelflife_supermarket, Category_revenue_share, Avg_o_supermarket, Sales_area, Avg_coefvar_o_supermarket

Table 3: Regression coefficients of analysis 1 in convenience category

Coefficients ^a							
Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
	B	Std. Error	Beta			Tolerance	VIF
7 (Constant)				-3,735	,000		
Avg_o_supermarket	-1,393	,443	-,243	-3,148	,002	,320	3,122
Avg_coefvar_o_supermarket	8,452	2,836	,270	2,981	,003	,232	4,309
Category_revenue_share	-35,132	21,183	-,080	-1,658	,099	,817	1,225
Order_advises_followed	5,993	2,864	,093	2,092	,038	,967	1,034
Sales_area	,001	,001	,125	2,245	,026	,616	1,624
Avg_shelflife_supermarket	,818	,253	,155	3,238	,001	,825	1,211
Q_Mu_M_supermarket	44,989	3,174	,768	14,176	,000	,649	1,542

a. Dependent Variable: Percentage_outdating_supermarket

9.4.2.1 Checking assumptions of multiple linear regression analysis

Field (2009) describes nine assumptions that must be true to draw conclusions about a population based on a multiple regression analysis. These assumptions will be tested for analysis 1.

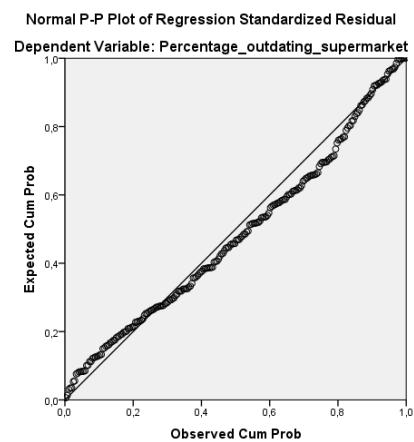
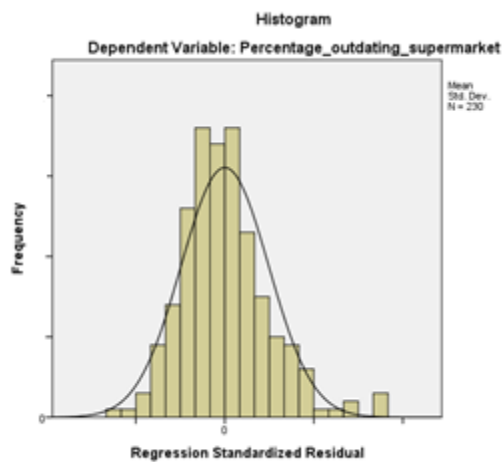
1. Variable types : All variables predictor variables should be quantitative or categorical and the outcome variable must be quantitative, continuous and unbounded
This assumption holds as we have one categorical and 13 quantitative predictors. The dependent variable is quantitative, continuous and unbounded.
2. Non-zero variance: the predictors should have some variation in value
This assumption holds for all 14 predictors.
3. No perfect multicollinearity: there should be no perfect linear relationship between two or more of the predictors
This assumptions holds for our final model as can be seen in table 17 above. All VIF values are below 10 and the average VIF value is 2,01 which is acceptable.
4. Predictors are uncorrelated with ‘external variables’
This assumption means that there should be no external variables that correlate with any of the variables included in the regression model. We are not able to test this assumption as about a lot of possible relevant variables for our model no data is available.
5. Homoscedasticity: at each level of the predictor variables, the variance of the residual terms should be constant

As can be seen in the graphs that SPSS provides, the dots are evenly dispersed around zero. This means that there is homoscedasticity in the variance of the residual terms of all predictors. So, this assumption is met.

6. Independent errors: for any two observations the residual terms should be uncorrelated. Our final model gives an Durbin-Watson statistic of 1,792. This is acceptable according to Field (2009).

7. Normally distributed errors

The histogram as displayed below looks like a normal distribution. In the second figure the dots are not far from the line and this indicates a small deviation from normality. From both figures can be concluded that the residuals are normally distributed and our assumption is met.



8. Independence

Each of our values come from a separate entity, so all of the values of the outcome variables are independent. So, this assumption is met.

9. Linearity

The partial plot for the supermarket QMuM ratio shows a clear linear relationship. The partial plots for the other predictors show less linear relationships. However, the partial plots mostly follow a straight line and are not showing any other relationships like an exponential or curve linear relationship. The deviations from the straight line are coming from supermarkets whose percentage of outdated differs a lot from similar supermarkets. These points may not be removed as they are coming from real supermarkets. As none of the partial plots show another relationship than a linear relationship, this assumption is met.

9.4.3 Convenience: analysis 2: analysis on product level

Table 4: Model summary of analysis 2 in convenience category

Model Summary				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,558 ^a	,312	,293	21,526
2	,558 ^b	,311	,296	21,481

a. Predictors: (Constant), Ln_avg_shelflife_product, Mu_product, Q_Mu_M_product, Avg_coefvar_o_product, Avg_o_product, Avg_coefvar_sales_product

b. Predictors: (Constant), Ln_avg_shelflife_product, Mu_product, Q_Mu_M_product, Avg_coefvar_o_product, Avg_coefvar_sales_product

Table 5: Regression coefficients of analysis 2 in convenience category

Coefficients ^a								
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
2	(Constant)				2,671	,008		
	Avg_coefvar_o_product	-23,751	9,062	-,183	-2,621	,009	,644	1,552
	Mu_product	10,645	5,144	,183	2,069	,040	,403	2,479
	Avg_coefvar_sales_product	27,960	7,730	,374	3,617	,000	,296	3,378
	Q_Mu_M_product	24,135	8,164	,212	2,956	,003	,616	1,622
	Ln_avg_shelflife_product	-24,409	4,886	-,385	-4,996	,000	,533	1,877

a. Dependent Variable: Percentage_outdating

9.4.3.1 Checking assumptions of multiple linear regression analysis

The same assumptions as for analysis 1 will be tested for analysis 2.

1. Variable types : All variables predictor variables should be quantitative or categorical and the outcome variable must be quantitative, continuous and unbounded
This assumption holds as we have 8 quantitative predictors. The dependent variable is quantitative, continuous and unbounded.
2. Non-zero variance: the predictors should have some variation in value
This assumption holds for all 8 predictors.
3. No perfect multicollinearity: there should be no perfect linear relationship between two or more of the predictors

This assumptions holds for our final model as can be seen in table 19 above. All VIF values are below 10 and the average VIF value is 2,18 which is acceptable.

4. Predictors are uncorrelated with 'external variables'

This assumption means that there should be no external variables that correlate with any of the variables included in the regression model. We are not able to test this assumption as about a lot of possible relevant variables for our model no data is available.

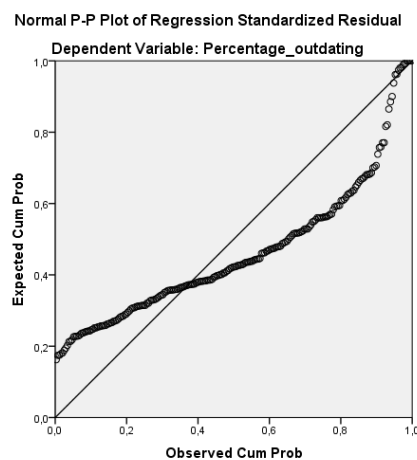
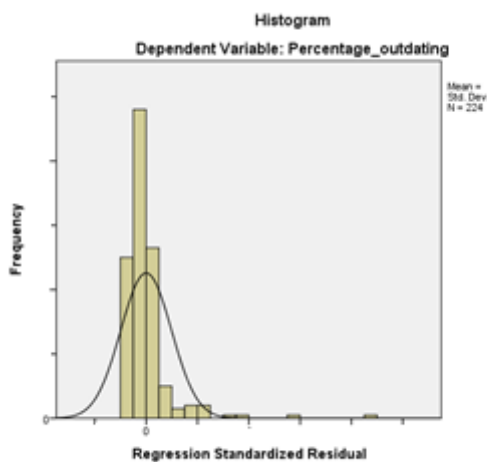
5. Homoscedasticity: at each level of the predictor variables, the variance of the residual terms should be constant

As can be seen in the graphs that SPSS provides, the dots are evenly dispersed around zero. In each figure some dots that with a big distance from zero can be found, these are outliers of products which deviate a lot from similar products. These points may not be removed as they are coming from real products. The distribution of the points is acceptable and this means that there is homoscedasticity in the variance of the residual terms of all predictors. So, this assumption is met.

6. Independent errors: for any two observations the residual terms should be uncorrelated
Our final model gives an Durbin-Watson statistic of 1,712. This is acceptable according to Field (2009).

7. Normally distributed errors

The histogram as displayed below shows that the residuals are positively skewed. The normal probability plot also shows up deviation from normality. The deviation is acceptable, so we can conclude that our assumption is met.



8. Independence

Each of our values come from a separate entity, so all of the values of the outcome variables are independent. So, this assumption is met.

9. Linearity

All partial plots don't show clear linear relationships. However, the partial plots mostly follow a straight line and are not showing any other relationships like an exponential or curve linear relationship. The deviations from the straight line are coming from products whose percentage of outdated differs a lot from similar products. These points may not be removed as they are coming from real products. As none of the partial plots show another relationship than a linear relationship, this assumption is met.

9.4.4 Convenience: analysis 3: analysis on supermarket – product combination

The total dataset of supermarket – product combinations is split per article. So, we run per article a multiple regression analysis on each supermarket product combination using the backward method. For the 224 articles 281 different models were found. The adjusted R square of the models varies from 0,000 till 0,803. Articles with an adjusted R square of below 0,10 for their best regression model were deleted from the dataset. This results in 196 remaining articles. We run a new multiple regression and find a model with eight independent variables. Some of them are highly correlated. The adjusted R square of this model is 0,201. This is only a small improvement of the base model, so we continue our deep analysis.

We split the total dataset of remaining supermarket – product combinations per supermarket. So, we run per supermarket a multiple regression analysis on each supermarket product combination. For the 230 supermarkets 468 different models were found. The adjusted R square of the models varies from 0,026 till 0,654. Supermarkets with an adjusted R square of below 0,10 for their best regression model were deleted from the dataset. This results in 220 remaining supermarkets.

Table 6: Model summary of final model in analysis 3 in convenience category

Model Summary				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,460 ^a	,212	,211	41,5357475101 77970

a. Predictors: (Constant), Q_Mu_M_combination, Ln_avg_shelflife_combination, Coefvar_o_combination, Mu_combination, Avg_o_combination, Coefvar_sales_combination, Q_combination

Table 7: Regression coefficients of final model in analysis 3 in convenience category

Coefficients ^a							
Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
	B	Std. Error	Beta			Tolerance	VIF
1 (Constant)				28,453	,000		
Avg_o_combination	,250	,119	,019	2,094	,036	,304	3,288
Coefvar_o_combination	-16,724	,901	-,106	-	,000	,734	1,363
Q_combination	-2,930	,189	-,157	-	,000	,236	4,246
Mu_combination	4,534	,549	,058	8,257	,000	,484	2,065
Coefvar_sales_combination	-2,806	,677	-,033	-4,144	,000	,389	2,571
Ln_avg_shelflife_combination	-15,155	,779	-,126	-	,000	,572	1,748

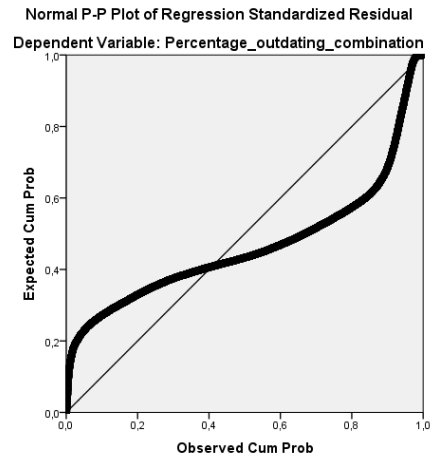
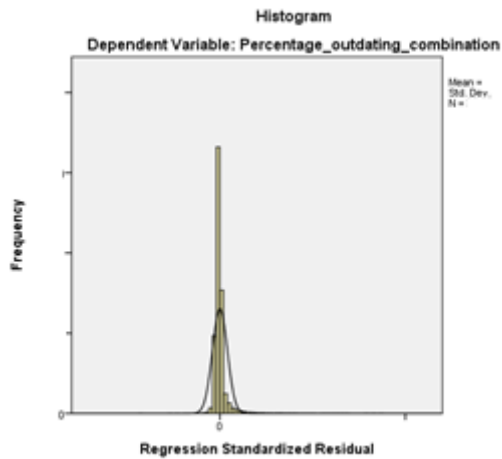
Q_Mu_M_combination	58,474	1,083	,442	53,992	,000	,358	2,794
--------------------	--------	-------	------	--------	------	------	-------

a. Dependent Variable: Percentage_outdating_combination

9.4.4.1 Checking assumptions of multiple linear regression analysis

The same assumptions as for analysis 1 and 2 will be tested for analysis 3.

1. Variable types : All variables predictor variables should be quantitative or categorical and the outcome variable must be quantitative, continuous and unbounded
This assumption holds as we have 8 quantitative predictors. The dependent variable is quantitative, continuous and unbounded.
2. Non-zero variance: the predictors should have some variation in value
This assumption holds for all 8 predictors.
3. No perfect multicollinearity: there should be no perfect linear relationship between two or more of the predictors
This assumptions holds for our final model as can be seen in table 21 above. All VIF values are below 10 and the average VIF value is 2,58 which is acceptable.
4. Predictors are uncorrelated with 'external variables'
This assumption means that there should be no external variables that correlate with any of the variables included in the regression model. We are not able to test this assumption as about a lot of possible relevant variables for our model no data is available.
5. Homoscedasticity: at each level of the predictor variables, the variance of the residual terms should be constant
As can be seen in the graphs that SPSS provides, the dots are evenly dispersed around zero. In each figure some dots that with a big distance from zero can be found, these are outliers of combinations which deviate a lot from similar combinations. These points may not be removed as they are coming from real combinations. The distribution of the points is acceptable and this means that there is homoscedasticity in the variance of the residual terms of all predictors. So, this assumption is met.
6. Independent errors: for any two observations the residual terms should be uncorrelated
Our final model gives an Durbin-Watson statistic of 1,621. This is acceptable according to Field (2009).
7. Normally distributed errors
The histogram as displayed below shows that the residuals are normally distributed. However, the normal probability plot shows up quite a lot deviation from normality. This means that our assumption is not fully met.



8. Independence

Each of our values come from a separate entity, so all of the values of the outcome variables are independent. So, this assumption is met.

9. Linearity

Most partial plots don't show clear linear relationships. However, the partial plots mostly follow a straight line and are not showing any other relationships like an exponential or curve linear relationship. The deviations from the straight line are coming from combinations whose percentage of outdated differs a lot from similar combinations. These points may not be removed as they are coming from real combinations. For the partial plot of the case pack size it is clear that no clear linear relationship can be found as the case pack size is always an integer number and for some integer values like five no combinations exist. As none of the partial plots show another relationship than a linear relationship, this assumption is met.

9.4.5 Convenience: analysis 4: extra analysis on supermarket level

Table 8: Model summary of analysis 4 in convenience category

Model Summary				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,731 ^a	,534	,519	1,92488305642 4347

a. Predictors: (Constant), analyse1

9.4.6 Meat products and salads: analysis 1: analysis on supermarket level

Table 9: Model summary of analysis 1 in meat products and salads category

Model Summary				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,690 ^a	,477	,448	1,020
2	,690 ^b	,476	,450	1,018
3	,690 ^c	,476	,453	1,016
4	,690 ^d	,476	,454	1,014
5	,688 ^e	,473	,454	1,014
6	,686 ^f	,471	,454	1,014
7	,684 ^g	,468	,453	1,015

a. Predictors: (Constant), Q_Mu_M_supermarket, Avg_Q_supermarket, Servicelevel_HB, Openingdays, Avg_disposable_income, Order_advises_followed, Category_revenue_share, Sales_area, Avg_o_supermarket, Avg_coefvar_sales_supermarket, Avg_coefvar_o_supermarket, Mu_supermarket

b. Predictors: (Constant), Q_Mu_M_supermarket, Avg_Q_supermarket, Servicelevel_HB, Openingdays, Order_advises_followed, Category_revenue_share, Sales_area, Avg_o_supermarket, Avg_coefvar_sales_supermarket, Avg_coefvar_o_supermarket, Mu_supermarket

c. Predictors: (Constant), Q_Mu_M_supermarket, Avg_Q_supermarket, Servicelevel_HB, Openingdays, Order_advises_followed, Category_revenue_share, Avg_o_supermarket, Avg_coefvar_sales_supermarket, Avg_coefvar_o_supermarket, Mu_supermarket

- d. Predictors: (Constant), Q_Mu_M_supermarket, Avg_Q_supermarket, Openingdays, Order_advises_followed, Category_revenue_share, Avg_o_supermarket, Avg_coefvar_sales_supermarket, Avg_coefvar_o_supermarket, Mu_supermarket
- e. Predictors: (Constant), Q_Mu_M_supermarket, Openingdays, Order_advises_followed, Category_revenue_share, Avg_o_supermarket, Avg_coefvar_sales_supermarket, Avg_coefvar_o_supermarket, Mu_supermarket
- f. Predictors: (Constant), Q_Mu_M_supermarket, Openingdays, Order_advises_followed, Category_revenue_share, Avg_o_supermarket, Avg_coefvar_sales_supermarket, Avg_coefvar_o_supermarket
- g. Predictors: (Constant), Q_Mu_M_supermarket, Order_advises_followed, Category_revenue_share, Avg_o_supermarket, Avg_coefvar_sales_supermarket, Avg_coefvar_o_supermarket

Table 10: Regression coefficients of analysis 1 in meat products and salads category

Coefficients ^a							
Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
	B	Std. Error	Beta			Tolerance	VIF
7 (Constant)				1,380	,169		
Avg_o_supermarket	-,318	,151	-,184	-2,105	,036	,309	3,234
Avg_coefvar_o_supermarket	5,055	1,462	,311	3,458	,001	,292	3,420
Category_revenue_share	-,142	,066	-,108	-2,152	,032	,931	1,074
Avg_coefvar_sales_supermarket	1,314	,574	,154	2,287	,023	,522	1,915
Order_advises_followed	-,034	,017	-,099	-1,987	,048	,945	1,058
Q_Mu_M_supermarket	26,782	3,055	,624	8,766	,000	,467	2,143

a. Dependent Variable: Percentage_outdating_supermarket

9.4.6.1 Checking assumptions of multiple linear regression analysis

The same assumptions as in analysis 1 for the convenience category will be tested for analysis 1 in the meat products and salads category.

1. Variable types : All variables predictor variables should be quantitative or categorical and the outcome variable must be quantitative, continuous and unbounded
This assumption holds as we have one categorical and 13 quantitative predictors. The dependent variable is quantitative, continuous and unbounded.
2. Non-zero variance: the predictors should have some variation in value

This assumption holds for all 14 predictors.

- No perfect multicollinearity: there should be no perfect linear relationship between two or more of the predictors

This assumption holds for our final model as can be seen in table 26 above. All VIF values are below 10 and the average VIF value is 2,14 which is acceptable.

- Predictors are uncorrelated with 'external variables'

This assumption means that there should be no external variables that correlate with any of the variables included in the regression model. We are not able to test this assumption as about a lot of possible relevant variables for our model no data is available.

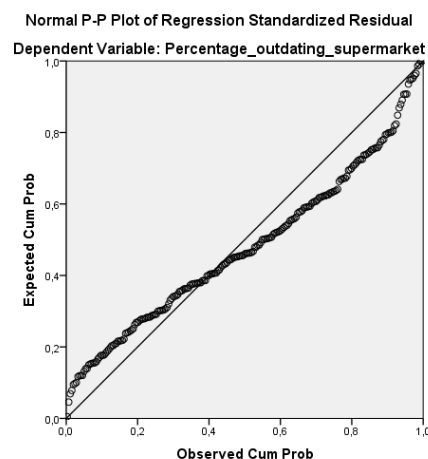
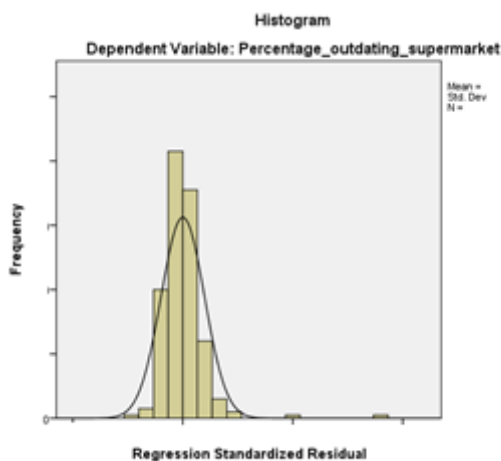
- Homoscedasticity: at each level of the predictor variables, the variance of the residual terms should be constant

As can be seen in the graphs that SPSS provides, most dots are evenly dispersed around zero. In each figure some dots that with a big distance from zero can be found, these are outliers of supermarkets which deviate a lot from similar supermarkets. These points may not be removed as they are coming from real supermarkets. The distribution of the points is acceptable and this means that there is homoscedasticity in the variance of the residual terms of all predictors. So, this assumption is met.

- Independent errors: for any two observations the residual terms should be uncorrelated
Our final model gives an Durbin-Watson statistic of 1,971. This is acceptable according to Field (2009).

- Normally distributed errors

The histogram as displayed below looks like a normal distribution. In the second figure the dots are not far from the line and this indicates a small deviation from normality. From both figures can be concluded that the residuals are normally distributed and our assumption is met.



- Independence

Each of our values come from a separate entity, so all of the values of the outcome variables are independent. So, this assumption is met.

- Linearity

The partial plots for the supermarket QMuM ratio, the average order size, the average coefficient of variation of order size and for the average coefficient of variation of sales show clear linear relationships. The partial plots for the other two predictors show less linear relationships. However, the partial plots of these two predictors mostly follow a straight line and are not showing any other relationships like an exponential or curve linear relationship. The deviations from the straight line are coming from supermarkets whose percentage of outdated differs a lot from similar supermarkets. These points may not be removed as they are coming from real supermarkets. As none of the partial plots show another relationship than a linear relationship, this assumption is met.

9.4.7 Meat products and salads: analysis 2: analysis on product level

Table 11: Model summary of analysis 2 in meat products and salads category

Model Summary				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,837 ^a	,701	,692	6,80785755516 2716
2	,837 ^b	,701	,693	6,79514626302 7042
3	,837 ^c	,700	,694	6,78775268554 6454
4	,836 ^d	,699	,693	6,79269297300 6307

a. Predictors: (Constant), Q_Mu_M_product, Q_product, Avg_sl_product, Avg_coefvar_o_product, Mu_product, Avg_coefvar_sales_product, Avg_o_product, Ln_avg_sl_product

b. Predictors: (Constant), Q_Mu_M_product, Q_product, Avg_sl_product, Avg_coefvar_o_product, Mu_product, Avg_coefvar_sales_product, Avg_o_product

c. Predictors: (Constant), Q_Mu_M_product, Q_product, Avg_coefvar_o_product, Mu_product, Avg_coefvar_sales_product, Avg_o_product

d. Predictors: (Constant), Q_Mu_M_product, Q_product, Mu_product, Avg_coefvar_sales_product, Avg_o_product

Table 12: Regression coefficients of analysis 2 in meat products and salads category

Model		Coefficients ^a						
		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
4	(Constant)				-4,625	,000		
	Avg_o_product	,434	,239	,153	1,813	,071	,156	6,390
	Q_product	-1,006	,250	-,331	-4,031	,000	,166	6,030
	Mu_product	2,484	,482	,271	5,154	,000	,405	2,472
	Avg_coefvar_sales_product	4,455	1,371	,166	3,249	,001	,425	2,353
	Q_Mu_M_product	43,742	2,112	,846	20,710	,000	,668	1,496

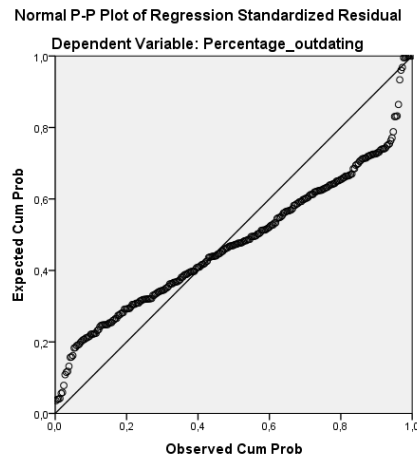
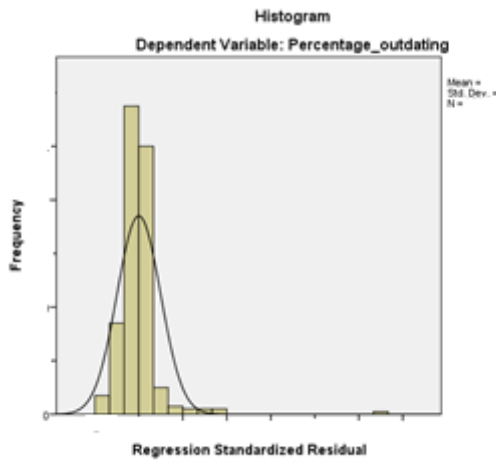
a. Dependent Variable: Percentage_outdating

9.4.7.1 Checking assumptions of multiple linear regression analysis

The same assumptions as for analysis 1 will be tested for analysis 2.

1. Variable types : All variables predictor variables should be quantitative or categorical and the outcome variable must be quantitative, continuous and unbounded
This assumption holds as we have 8 quantitative predictors. The dependent variable is quantitative, continuous and unbounded.
2. Non-zero variance: the predictors should have some variation in value
This assumption holds for all 8 predictors.
3. No perfect multicollinearity: there should be no perfect linear relationship between two or more of the predictors
This assumptions holds for our final model as can be seen in table 28 above. All VIF values are below 10 and the average VIF value is 3,7482 which is relatively high, but acceptable.
4. Predictors are uncorrelated with 'external variables'
This assumption means that there should be no external variables that correlate with any of the variables included in the regression model. We are not able to test this assumption as about a lot of possible relevant variables for our model no data is available.
5. Homoscedasticity: at each level of the predictor variables, the variance of the residual terms should be constant
As can be seen in the graphs that SPSS provides, most dots are evenly dispersed around zero. In each figure some dots that with a big distance from zero can be found, these are outliers of products which deviate a lot from similar products. These points may not be removed as they are coming from real products. The distribution of the points is acceptable and this means that there is homoscedasticity in the variance of the residual terms of all predictors. So, this assumption is met.
6. Independent errors: for any two observations the residual terms should be uncorrelated
Our final model gives an Durbin-Watson statistic of 1,747. This is acceptable according to Field (2009).
7. Normally distributed errors

The histogram as displayed below looks like a normal distribution. In the second figure the dots are not far from the line and this indicates a small deviation from normality. From both figures can be concluded that the residuals are normally distributed and our assumption is met.



8. Independence

Each of our values come from a separate entity, so all of the values of the outcome variables are independent. So, this assumption is met.

9. Linearity

The partial plot for the product QMuM ratio shows a clear linear relationship. The partial plots for the other predictors show less linear relationships. However, the partial plots mostly follow a straight line and are not showing any other relationships like an exponential or curve linear relationship. The deviations from the straight line are coming from products whose percentage of outdated differs a lot from similar products. These points may not be removed as they are coming from real products. For the partial plot of the case pack size it is clear that no clear linear relationship can be found as the case pack size is always an integer number and for some integer values like two there are no products. As none of the partial plots show another relationship than a linear relationship, this assumption is met.

9.4.8 Meat products and salads: analysis 3: analysis on supermarket – product combination

The total dataset of supermarket – product combinations is split per article. So, we run per article a multiple regression analysis on each supermarket product combination using the backward method. For five articles the dependent variable is constant and no models can be formed. For the remaining 276 articles the adjusted R square of the models varies from 0,000 till 1,000. Articles with an adjusted R square of below 0,10 for their best regression model were deleted from the dataset. This results in 253 remaining articles. We run a new multiple regression analysis and find a model with all independent variables. Some of them are highly correlated. The adjusted R square of this model is 0,221. This is only a small improvement of the base model, so we continue our deep analysis.

We split the total dataset of remaining supermarket – product combinations per supermarket. So, we run per supermarket a multiple regression analysis on each supermarket product combination. For

the 233 supermarkets the adjusted R square of the models varies from 0,000 till 0,934. Supermarkets with an adjusted R square of below 0,10 for their best regression model were deleted from the dataset. This results in 223 remaining supermarkets.

Table 13: Model summary of final model in analysis 3 in meat products and salads category

Model Summary				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,490 ^a	,240	,240	30,0493768711 90578
2	,490 ^b	,240	,240	30,0491135715 96580
3	,490 ^c	,240	,240	30,0494864454 07110

- a. Predictors: (Constant), Ln_avg_shelflife_combination, Mu_combination, Coefvar_o_combination, Q_Mu_M_combination, Q_combination, Coefvar_sales_combination, Avg_o_combination
- b. Predictors: (Constant), Ln_avg_shelflife_combination, Mu_combination, Coefvar_o_combination, Q_Mu_M_combination, Coefvar_sales_combination, Avg_o_combination
- c. Predictors: (Constant), Ln_avg_shelflife_combination, Coefvar_o_combination, Q_Mu_M_combination, Coefvar_sales_combination, Avg_o_combination

Table 14: Regression coefficients of final model in analysis 3 in meat products and salads category

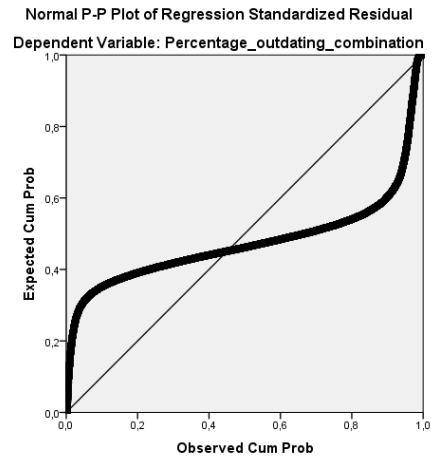
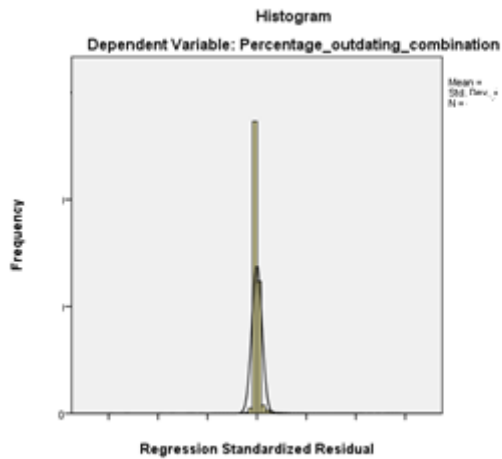
Coefficients ^a							
Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
	B	Std. Error	Beta			Tolerance	VIF
3 (Constant)				16,407	,000		
Avg_o_combination	,184	,032	,027	5,812	,000	,849	1,177
Coefvar_o_combination	-5,843	,539	-,048	-	,000	,909	1,101
Coefvar_sales_combination	6,538	,317	,127	20,598	,000	,478	2,093
Q_Mu_M_combination	22,212	,357	,374	62,244	,000	,502	1,994
Ln_avg_shelflife_combination	-7,880	,336	-,116	-	,000	,744	1,345
				23,432			

- a. Dependent Variable: Percentage_outdating_combination

9.4.8.1 *Checking assumptions of multiple linear regression analysis*

The same assumptions as for analysis 1 and 2 will be tested for analysis 3.

1. Variable types : All variables predictor variables should be quantitative or categorical and the outcome variable must be quantitative, continuous and unbounded
This assumption holds as we have 8 quantitative predictors. The dependent variable is quantitative, continuous and unbounded.
2. Non-zero variance: the predictors should have some variation in value
This assumption holds for all 8 predictors.
3. No perfect multicollinearity: there should be no perfect linear relationship between two or more of the predictors
This assumption holds for our final model as can be seen in table 30 above. All VIF values are below 10 and the average VIF value is 1,542 which is acceptable.
4. Predictors are uncorrelated with 'external variables'
This assumption means that there should be no external variables that correlate with any of the variables included in the regression model. We are not able to test this assumption as about a lot of possible relevant variables for our model no data is available.
5. Homoscedasticity: at each level of the predictor variables, the variance of the residual terms should be constant
As can be seen in the graphs that SPSS provides, most dots are evenly dispersed around zero. In each figure some dots that with a big distance from zero can be found, these are outliers of combinations which deviate a lot from similar combinations. These points may not be removed as they are coming from real combinations. The distribution of the points is acceptable and this means that there is homoscedasticity in the variance of the residual terms of all predictors. So, this assumption is met.
6. Independent errors: for any two observations the residual terms should be uncorrelated
Our final model gives an Durbin-Watson statistic of 1,862. This is acceptable according to Field (2009).
7. Normally distributed errors
The histogram as displayed below shows that the residuals are normally distributed. However, the normal probability plot shows up quite a lot deviation from normality. This means that our assumption is not fully met.



8. Independence

Each of our values come from a separate entity, so all of the values of the outcome variables are independent. So, this assumption is met.

9. Linearity

Most partial plots don't show clear linear relationships. However, the partial plots mostly follow a straight line and are not showing any other relationships like an exponential or curve linear relationship. The deviations from the straight line are coming from combinations whose percentage of outdated differs a lot from similar combinations. These points may not be removed as they are coming from real combinations. As none of the partial plots show another relationship than a linear relationship, this assumption is met.

9.4.9 Meat products and salads: analysis 4: extra analysis on supermarket level

Table 15: Model summary of analysis 4 in meat products and salads category

Model Summary				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,782 ^a	,612	,599	,867978950789 323

a. Predictors: (Constant), analyse1