Designing a decision support tool for assortment planning

Stienstra, E.J.

Award date:
2014

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
Designing a decision support tool for assortment planning

by

E.J. Stienstra

BSc Industrial Engineering and Management Science — TU/e 2012
Student identity number 0668000

in partial fulfilment of the requirements for the degree of

Master of Science
in Operations Management and Logistics

Supervisors:
Drs. E. Hoppenbrouwer, MBA, Chief Operating Officer, Fagron
BSc. J. Groothuizen, Application Consultant, Fagron
Ir. dr. S.D.P. Flapper, TU/e, OPAC
Dr. D.B.L.P. Honhon, TU/e, OPAC
TUE. School of Industrial Engineering.

Series Master Theses Operations Management and Logistics

Subject headings: assortment planning, product portfolio, basket shopping, one-stop-shop
I. PREFACE

The report you are about to read is the product of six months of hard work on my master project in order to fulfill my master degree in Operations Management and Logistics. I really enjoyed working on this project and I’m really satisfied with the result. I hope you enjoy reading this work and that you can learn something from it. But first, I would like to express my gratitude to some people.

First of all, special thanks goes out to Simme Douw Flapper for being my supervisor. He really helped me in completing this project. During the numerous meetings we had, he always gave very constructive feedback. This was very motivating, inspiring and it always made me walk away with a smile. I would like to thank Dorothee Honhon for the good insights during the project proposal meeting and concept report meeting. This surely helped to improve the project.

Furthermore, I would like to thank the Fagron Group and Spruyt hillen in giving me the opportunity to do this project. Thanks to Erik Hoppenbrouwer who was the initiator of this project. He kept his promise that I would not leave before actually delivering a tool which the company will use. Special thanks goes out to Joost Groothuizen, who really helped me a lot in finding everything out in the company, especially things related with the information system. Also, thanks to everyone within the IM department, who helped me to some extent. From Spruyt hillen I would like to thank Kirsten Dam and Arend-Jan Gerritsen. In the several meetings we had, they always gave good feedback which really helped to improve the tool.

At last, I would like to thank all my fellow students, friends and family who have always been a great support to me. This report is not only the end of my project, but also the closure of a very exciting and interesting period in my life. It brought me through the cities of Tilburg, Eindhoven, Copenhagen and finally Rotterdam, where I met loads of nice people and learned so many interesting things. I really enjoyed it and I can say that it has been an unforgettable period of my life.

Erik Stienstra

Rotterdam, January 2014
II. ABSTRACT

The decision which particular products to sell is of great concern for the Fagron Group. However none of its' subsidiaries had a proper tool which supports their decision making about the assortment. Therefore the assignment of this project was to develop a decision support tool for the assortment planning by the Fagron Group. The project was conducted at Spruyt hillen, one of the companies within the Fagron Group. Spruyt hillen is a one-stop-shop for pharmaceutical products. Consequently, the basket shopping phenomenon was included in the decision support tool which was developed in AIMMS. Furthermore, Spruyt hillen wanted to have some influence on the profitability of the assortment. Therefore, restrictions on profit margins were implemented which can be controlled by the user. The developed tool has been well received by both Spruyt hillen and the Fagron Group as a whole. In the end, it was decided that the tool will actually be implemented in the company.
III. MANAGEMENT SUMMARY

INTRODUCTION

The Fagron Group consists of numerous sales and distribution companies for pharmaceutical products. However, none of the companies has a proper tool which supports their decision making about the assortment. Therefore the assignment of this project is the development of a decision support tool for the assortment planning of the Fagron Group. In first instance, the tool will be validated at Spruyt hillen.

Three main research questions are formulated in order to structure the report.

Main research question 1:

- What are the relevant aspects with respect to the assortment decisions at Spruyt hillen?

Main research question 2:

- How to implement the relevant aspects into a decision support tool for the assortment planning at Spruyt hillen?

Main research question 3:

- What is the performance of the developed decision support tool?

ANALYSIS

Spruyt hillen aims to be a one-stop-shop for pharmaceutical products, therefore the interest in this project is especially in the basket shopping phenomenon. This is supported by a basket analysis, which showed that customers tend to buy more than one product from the company.

It also turned out that better insight and control is desired on the performance of the assortment. Therefore, restrictions with relations to the profit can be included in the tool in order to achieve targets on the assortment. On the other hand, it turned out that for Spruyt hillen inventory and substitution effects were less relevant in relation to assortment decisions and consequently these aspects are out of scope in this project.

DESIGN

In first instance, it was attempted to develop an optimization model which included basket shopping. However, the estimation of customer and product retention was a task which could not be done in a reasonable amount of time. On top of that, it was not possible to find a global optimal solution in a reasonable amount time. Therefore, an optimization model with basket shopping at once was unfeasible for this project. Consequently, a heuristic was designed to solve the problem, see Figure 1 down below for an overview of this heuristic.
It starts with a mathematical model which distinguishes well performing products from candidate products. The mathematical model contains restrictions on the profitability of products, product groups and the total assortment. The values for these restrictions can be set by the user of the tool. For the well performing products it is advised to keep the products in the assortment. The candidate product should however be analyzed in more detail before they can be removed from the assortment. For each candidate product its’ customers and related products should be analyzed. The user of the tool can make assumptions on the customer and product retention. If removing a candidate product results in the loss of a customer, that customer is called a critical customer for that candidate product. Similarly, if removing a candidate product causes a loss in demand for another related product, that product is called a critical product for the candidate product. Furthermore, a customer classification is made to distinguish important customers from less important customers. Taking all these information into account, the tool gives a final advice for the candidate products, according to the decision diagram in Figure 2 on the next page.

Next to the removal of existing products, it is also possible to add new products, which are not yet available in the ERP system. For these new products a cannibalization and an extra sales effect are included.

The heuristic was implemented in AIMMS together with a user-friendly interface. Thereafter, the tool was verified and validated.
The main results are listed below:

- The tool will make the barrier lower to actually be concerned with assortment planning. This will motivate Spruyt hillen to be more concerned with assortment decisions. The expectation is therefore that the assortment is better ‘maintained’, by which is meant that products which are not performing well, are removed earlier from the assortment, resulting in cost savings.
- At Spruyt hillen, sometimes arguments of related products or customers are used, to keep less or unprofitable products in the assortment, without support of data. The tool provides these data about the interaction of products with customers and other products. The expectation is that products will therefore not be kept in the assortment for the wrong reasons.
- On the other hand, the transparency in relation to customers will also prevent that some products will be removed without informing a customer who specifically needs this product.
- The tool gives transparency in the performance on three levels, individual products/customers, product/customers groups and the total assortment. Furthermore, the tool gives the user control to influence the performance of the assortment.
- The expectation of Spruyt hillen is that they will not spend necessarily less time on assortment decisions, but spend time more effectively. They will be more concerned with assortment planning, resulting in a better maintained assortment, where products are removed when they do not generate enough money anymore. This results in costs savings.
- Insight in the cannibalization and extra sales associated with the introduction of a new product is provided. The introduction of a new product will be a more deliberate decision.
The quantification of the relations between products can also be used for product suggestions in the web shop, which may increase sales.

The tool is an extra trigger for Spruyt hillen to further improve their data integrity.

CONCLUSION AND RECOMMENDATIONS

The main assignment of this project was to develop a decision support tool for assortment planning, which has been completed successfully. The tool was well received by both Spruyt hillen and the Fagron Group. It was decided that the tool will be implemented at Spruyt hillen. To facilitate this implementation, a user manual was written and training to the user was provided. Furthermore, the preparation for the roll out to other Fagron Group companies was done as well.

Recommendations

- The tool relies on the data as available in the ERP system. If there are errors in the data, this will influence the results of the tool. Therefore, it is strongly advised not to focus blindly on the results of the tool.
- It is recommended for Spruyt hillen to use a more comprehensive cost price calculation, where also inventory and handling costs are taken into account.

Academic relevance

A lot of research is done within the field of assortment planning. However, the phenomenon of basket shopping is not often studied in combination with assortment planning. Brijs et al. (2000) and Flapper et al. (2010) also did a case study on assortment planning in combination with basket shopping. However, Brijs et al. (2004) ignored the retention of customers. Furthermore, Flapper et al. (2010) assumed that a customer is immediately lost when demand is not satisfied, which does not hold in this project. The heuristic developed in this project takes into account both relations with other products and the retention of customers. The heuristic and the case study are both contributions to science.
## IV. TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Preface</td>
<td>5</td>
</tr>
<tr>
<td>II. Abstract</td>
<td>7</td>
</tr>
<tr>
<td>III. Management Summary</td>
<td>9</td>
</tr>
<tr>
<td>Introduction</td>
<td>9</td>
</tr>
<tr>
<td>Analysis</td>
<td>9</td>
</tr>
<tr>
<td>Design</td>
<td>9</td>
</tr>
<tr>
<td>Results</td>
<td>11</td>
</tr>
<tr>
<td>Conclusion and recommendations</td>
<td>12</td>
</tr>
<tr>
<td>IV. Table of Contents</td>
<td>13</td>
</tr>
<tr>
<td>1. Introduction,</td>
<td>16</td>
</tr>
<tr>
<td>1.1 Introduction to the Project</td>
<td>16</td>
</tr>
<tr>
<td>1.2 Deliverables</td>
<td>16</td>
</tr>
<tr>
<td>1.3 Company Background</td>
<td>16</td>
</tr>
<tr>
<td>1.3.1 Fagron Group</td>
<td>16</td>
</tr>
<tr>
<td>1.3.2 Spruyt Hellen</td>
<td>17</td>
</tr>
<tr>
<td>1.4 Main Research Questions</td>
<td>17</td>
</tr>
<tr>
<td>1.5 Research Design</td>
<td>18</td>
</tr>
<tr>
<td>2. Orientation</td>
<td>19</td>
</tr>
<tr>
<td>2.1 Current Method of Assortment Planning</td>
<td>19</td>
</tr>
<tr>
<td>2.2 Current Assortment</td>
<td>20</td>
</tr>
<tr>
<td>2.3 Customers</td>
<td>21</td>
</tr>
<tr>
<td>2.4 Key Performance Indicators (KPI's)</td>
<td>22</td>
</tr>
<tr>
<td>2.5 Conclusion</td>
<td>22</td>
</tr>
<tr>
<td>3. Analysis</td>
<td>23</td>
</tr>
<tr>
<td>3.1 Basket Shopping</td>
<td>23</td>
</tr>
<tr>
<td>3.1.1 Basket Analysis</td>
<td>23</td>
</tr>
<tr>
<td>3.1.2 Basket Shopping Models</td>
<td>25</td>
</tr>
<tr>
<td>3.2 Other Aspects from Literature</td>
<td>25</td>
</tr>
<tr>
<td>3.2.1 Substitution and Cannibalization</td>
<td>25</td>
</tr>
<tr>
<td>3.2.2 Inventory</td>
<td>26</td>
</tr>
<tr>
<td>3.3 Conclusion</td>
<td>26</td>
</tr>
<tr>
<td>4. Design</td>
<td>28</td>
</tr>
<tr>
<td>4.1 Research Sub-questions</td>
<td>28</td>
</tr>
<tr>
<td>4.2 Optimization Models including Basket Shopping</td>
<td>29</td>
</tr>
<tr>
<td>4.2.1 Starting Point: Agrawal and Smith (2003)</td>
<td>29</td>
</tr>
</tbody>
</table>
4.2.2 Alternative Optimization Model ............................................................ 29
4.3 Heuristic ........................................................................................................... 31
4.3.1 Mathematical Model .................................................................................... 32
4.3.2 Candidate Analysis .................................................................................... 36
4.4 Additional Functionalities ............................................................................. 43
4.5 Conclusion ....................................................................................................... 45
5. Implementation into Software .......................................................................... 46
5.1 Choice of Software ......................................................................................... 46
5.2 Tool Design ..................................................................................................... 46
5.3 Solver ............................................................................................................... 48
5.4 Conclusion ....................................................................................................... 48
6. Verification and Validation ................................................................................ 49
6.1 Verification ...................................................................................................... 49
6.2 Validation ........................................................................................................ 49
6.3 Sensitivity Analysis ....................................................................................... 49
6.4 Conclusion ....................................................................................................... 52
7. Results ................................................................................................................. 53
7.1 Improvement in decisions .............................................................................. 53
7.2 Decision time reduction ................................................................................. 54
7.3 Other implications .......................................................................................... 54
8. Implementation in the Company ...................................................................... 55
8.1 Spruyt hillen .................................................................................................... 55
8.2 Fagron Group .................................................................................................. 56
9. Conclusions and Recommendations ............................................................... 57
9.1 Conclusions ..................................................................................................... 57
9.2 Recommendations for the Company ............................................................. 58
9.3 Contribution to Science and Future Research ............................................... 59
References ............................................................................................................. 60
List of Tables and Figures ..................................................................................... 62
List of Definitions ................................................................................................ 64
List of Abbreviations ............................................................................................ 65
Appendix I: Literature Study ................................................................................ 66
  1.1 Introduction .................................................................................................... 66
  1.2 Basket shopping ............................................................................................. 66
  1.3 Substitution and Cannibalization ................................................................. 67
Appendix II: Mathematical model including basket shopping .......................... 69
Appendix III: User Manual .................................................................................. 73
1. How to use the tool ........................................................................................................73
   1.1 Tool Dashboard ........................................................................................................73
   1.2 Loading Data ...........................................................................................................73
   1.3 Saving and loading cases .......................................................................................73
   1.4 Select Product groups ...........................................................................................74
   1.5 User Inputs ..............................................................................................................74
   1.6 Optimize ..................................................................................................................75
   1.7 View Data ................................................................................................................75
   1.8 Charts .......................................................................................................................76
   1.9 Model input .............................................................................................................77
   1.10 Model output .........................................................................................................79
   1.11 Analyzing candidates ..........................................................................................82
   1.12 Making assumptions ..............................................................................................83
   1.13 Final results ..........................................................................................................84
   1.14 Export to excel ......................................................................................................86

Appendix IV: Mathematical Model ..................................................................................87
Appendix V: List of Formulas ..........................................................................................89
Appendix VI: Training Exercise and Certificate ...............................................................94
Appendix VII: Fagronized Tool .......................................................................................96
1. INTRODUCTION

1.1 INTRODUCTION TO THE PROJECT

This project is conducted for the Fagron Group. The Fagron Group consists of numerous sales and distribution companies for pharmaceutical products. Right now the assortment planning is decentralized. So sales & distribution companies which are part of the Fagron Group are responsible for their assortment planning themselves, where some companies perform better than others. Triggered by this situation Erik Hoppenbrouwer (COO Fagron Group) contacted my graduation supervisor as part of the collaboration between TU/e and Fagron. Also inspired by some academic research, this project was initiated in the expectation that improvement in the assortment planning of the Fagron Group may be possible.

The decision which products to sell is of great concern for the Fagron Group. However none of the companies has a proper tool which supports their decision making about the assortment. Therefore the assignment of this project is the development of a decision support tool for the assortment planning of the Fagron Group. In first instance, the tool will be validated at Spruythillen, which is one of those sales and distribution companies of the Fagron Group. Thereafter, it is intended that the tool will be rolled out across the Fagron Group.

1.2 DELIVERABLES

There are some formal deliverables agreed upon which will result from this project:

- Literature review
- Specification of the tool
- Decision support tool
- Tool validation report
- Training manual
- Training certificate user
- Report with presentation

1.3 COMPANY BACKGROUND

1.3.1 FAGRON GROUP

Fagron is a pharmaceutical company founded in 1990. In only 20 years Fagron has grown from a local player to global market leader in pharmaceutical compounding. At this moment they have about 1,300 employees located in 30 countries. Pharmaceutical compounding is the creation of medicines using appropriate ingredients and tools in order to satisfy the unique needs of a client. Fagron supplies semi-finished goods, high-quality pharmaceutical raw materials and all kind of pharmaceutical equipment and accessories to pharmacies and industry. By developing innovative concepts and solutions Fagron tries to meet the specific and unique wishes of their customers. The consolidation of companies all over the world resulted in market leadership and provides advantages such as the central procurement of raw materials and the international
roll-out of successful and innovative products and concepts. Fagron Group is one of the four subdivisions of the Arseus NV holding.

1.3.2 SPRUYT HILLEN

Spruyt hillen is taken over by Arseus NV in 2007 and also part of the subdivision Fagron Group. Spruyt hillen believes that collaborative innovation is the only way to success in the pharmaceutical industry. In close cooperation they help pharmaceutical and medical professionals and organizations to be prepared for the next step. With the right unique mix of products, services and concepts for every specific target group and need. Spruyt hillen will continuously endeavor to proceed in a professional, enthusiastic and smart way together with their customers.

1.4 MAIN RESEARCH QUESTIONS

The main deliverable of this project is to develop a decision-support tool for the assortment decisions of Spruyt hillen. The tool should give advice on whether products should be in the assortment or not. Before building the tool, it is important to reveal all relevant aspects related to the assortment decisions at Spruyt hillen. Therefore the first research question is as follows

Main research question 1:

- What are the relevant aspects with respect to the assortment decisions at Spruyt hillen?

First, it is considered how assortment decisions are currently made at Spruyt hillen. Furthermore, the products and customers of Spruyt hillen will be regarded. Then, these findings will be related to academic literature on assortment decisions. The literature review is also the first deliverable of the project. After having analyzed both the academic literature and Spruyt hillen, the first research question can be answered. It should be decided which of the relevant aspects should be incorporated into the tool, resulting in the tool specification, which is the second deliverable of the project.

The next question is how to ‘translate’ the aspects from the tool specification into a decision-support tool. Therefore, the next research question is formulated

Main research question 2:

- How to implement the relevant aspects into a decision support tool for the assortment planning at Spruyt hillen?

It should be decided if a mathematical model will be used and how the relevant aspect from the tool specification can be incorporated in a tool. The answer to this question should result into a decision support tool implemented in software, which is the main assignment and third deliverable of this project. The tool should be validated, which is also a formal deliverable from Spruyt hillen.
The last research question considers the actual performance of the tool.

**Main research question 3:**

- **What is the performance of the developed decision support tool?**

Main research question 3 concerns the results of the tool. It will be evaluated if the tool will actually help Spruyt hillen to make better assortment decisions. It could also be possible that the time needed for making assortment decisions will be reduced. The last main research question is important, since the answer to this question will justify whether the actual implementation in the company is desirable. If it will be decided that the tool will actually be implemented in the company, the user should be trained. This requires the writing of a user manual and providing a training session. At last, the tool and project should both be presented at Spruyt hillen and the TU/e.

### 1.5 RESEARCH DESIGN

The research design, which will structure the report, is displayed in Figure 3. It is the framework of Kempen and Keizer (2000), which is especially developed for graduation projects. It contains four main phases: Orientation, Analysis, Design and Implementation. In Figure 3, the corresponding chapters with each phase are displayed. Furthermore, it is shown to which main research questions the phases relate.

![Figure 3: Research Design](image)

The first step is the Orientation phase in which the current assortment decisions at Spruyt hillen are considered. In the Analysis phase the findings at Spruyt hillen will be related to the academic literature. This analysis will result in the relevant aspects to be incorporated into the tool, the tool specification. The Design phase consists of the development of the decision-support tool based on the tool specification. This phase consists not only of the design of the model itself, but also the implementation into software and validation and verification of the tool. In the implementation phase, the results of the tool will be considered. Based on this, it will be decided if the tool will actually be implemented in the company.
2. ORIENTATION

This chapter contains the Orientation phase of the project. First, the current method of assortment planning at Spruyt hillen is considered. “The goal of assortment planning is to specify an assortment that maximizes sales or gross margin” (Kök et al., 2009). Where, the assortment is defined by the set of products carried in a store at a certain moment (Kök et al., 2009). Next to the current method, also the current assortment and customers are considered.

2.1 CURRENT METHOD OF ASSORTMENT PLANNING

Spruyt hillen has a workgroup for each product group. One of the decisions these workgroups make is the decision about the products to carry in the assortment. Other decisions they make are for example, promotions and discount offers. In each workgroup there is someone from internal sales, field sales, quality, marketing and procurement. These workgroups have regular meetings about once in four to five weeks. Because different divisions of the company are represented, triggers for assortment decisions come from both inside and outside the company. From outside the company these signals can either come from the supply or demand side. If for example a supplier raises the price of a product drastically, the representative from procurement could suggest to consider if it is still beneficial to carry that product. On the other hand, a field salesman could pick up signals from customers who want a specific product, which is not carried in the current assortment. Important triggers from inside the company can be disappointing sales or an unsatisfactory margin. At last, the quality representative ensures that the quality of the products meets company standards and legislation. There are of course more examples to come up with.

When adding a product, they estimate the potential sales that the product will generate associated with a revenue and margin. Based on this they decide whether to add a product to the portfolio. The introduction of a new product can cause cannibalization. However, they do not make calculations associated with this cannibalization.

The removal of products can be because of qualitative reasons like legislation or quantitative reasons such as unsatisfactory sales. For the qualitative reason above they basically have no choice but to remove the product or change the compound of the product. The removal based on business economic reasons is more interesting. Every now and then a report is created by someone of the workgroup which contains a sales overview of products in the product group. This procedure is quite time consuming and therefore they reflect less often than desired on their current assortment. This implicates that products with unsatisfactory sales can be carried longer in the assortment than desired, which can cost money.

In case there is a report on sales created, the products with unsatisfactory sales are discussed in the workgroup meeting. An important measure for these products is – not surprisingly – the profit the products generates. Though, sales people are quite reluctant to remove products. Often they come up with the argument that a product cannot be removed, since a certain customer would be lost or the demand for other products would be lost since they are complementary. These arguments are usually not supported by any data. On the other hand, it also happened once that a product was removed from the assortment without informing the customers for that product, which resulted in a very unsatisfied customer.
2.2 CURRENT ASSORTMENT

Spruyt hillen is a one-stop-shop which means that customers often buy more than one product from the company. Spruyt hillen offers a wide range of products, so that customers can buy almost everything they need at Spruyt hillen. At the time of writing, Spruyt hillen carries as many as 4493 products in the assortment, divided over 14 product groups. In Table 1, the different product groups are displayed with the number of products they contain, the number of products that were actually sold in the past year, the revenues and margin in euro’s and as a percentage. In this example, with past year is meant the period from 5-11-2012 until 5-11-2013.

<table>
<thead>
<tr>
<th>Product group</th>
<th>Number of products</th>
<th>Sold last year</th>
<th>Revenue (€)</th>
<th>Margin (€)</th>
<th>Margin (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

There are three different packaging product groups, ‘BULK’, ‘UA’ and standard. The BULK group contains products which are bought in large amounts. The UA group contains packaging from a specific brand. All other packaging related products are placed in the standard packaging group. Printed matter contains all the labeling, prescription paper and other press-work. The product group equipment contains a variety of equipment needed for pharmaceutical compounding, for example capsule filling machines or weighing scales. There are two compounding product groups, the standard compounding and AD-HOC compounding product group. The standard compounding group contains products which have a standard specific compound, whereas the ad-hoc compounding can be adjusted to requirements of the customer. GAMH contains compounded products from a specific supplier. The Full service group contains a variety of standard ointments and medicines. Examples from this product group are Vaseline and Eye drops. Raw materials speak for itself. These are the raw materials customers need in order to do the pharmaceutical compounding in their own pharmacy. Examples of raw materials are Alcohol.
and Melatonin. The Diabetes product group contains the InsuJet, which is an innovative needleless injector for insulin. Next to, the InsuJet some complementary product of the InsuJet are in this product group such as an extra adapter. All products which cannot be placed in one of the mentioned product groups are placed in the product groups Special and Specials.

2.3 CUSTOMERS

Spruyt hillen has about 4217 customers registered in their ERP system of which 2628 placed an order in the past year. The customers are divided into 12 customer groups, see Table 2. Just as in Table 1, the data is from the past year, so in the period from 5-11-2012 until 5-11-2013.

<table>
<thead>
<tr>
<th>Customer group</th>
<th># Customers in the system</th>
<th>Customers who bought something last year</th>
<th>Revenue (€)</th>
<th>Margin (€)</th>
<th>Margin (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pharmacists</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hospitals</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Industry</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wholesalers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercompany</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

There are a few customer groups to be distinguished, Pharmacists, Hospitals, Industry, Wholesalers, Intercompany and 'Other'. The last group mainly contains schools and general practitioners. Industry is further split up into food, veterinary, distribution, animal feed, pharma, cosmetic and others.

An important thing to notice is that different customers pay a different price for the same product. This is different from the retail industry, where customers pay the same price for one product, not considering discount offers. So, the customers of Spruyt hillen are also divided into different customer price group, which do not coincide with the group classification in Table 2.
2.4 KEY PERFORMANCE INDICATORS (KPI’S)

In order to evaluate the performance of the assortment, some key performance indicators are formulated.

- Absolute profit margin.
- Revenues.
- Costs of goods sold.
- Relative profit margin.
- Number of products.

The absolute profit margin is calculated by the revenues minus the costs of goods sold. The revenues are calculated by multiplying the demand with the sales price. The costs of goods sold are calculated by multiplying the cost price with the demand. The relative profit margin is the absolute profit margin as percentage of the revenues.

These five KPI’s can apply to three different levels: the total assortment, per product group and per individual product. The number of products is of course useless on individual product level. KPI’s are needed in order to evaluate the outcomes of the tool and to make comparisons under different circumstances. Since, it is required to calculate these KPI’s the decision support tool requires a mathematical model.

For the profit measures, Spruyt hillen has specific target values. However, they do not have clear insights in whether these targets are actually achieved. The tool can give transparency on the KPI’s and whether the targets are achieved. Furthermore, it is also possible to include restrictions with relation to the profit measures in order to fulfill the targets on the assortment. With this the user has influence on the profitability of the assortment.

2.5 CONCLUSION

Spruyt hillen is quite reluctant to remove products from their assortment, even though sales can be not that satisfying sometimes. Arguments of related products or customers buying the product, are used to justify for keeping products in the assortment. However, these arguments are often not supported by data. Because of the one-stop-shopping phenomenon, products cannot be considered in isolation of other products and customers. Moreover, Spruyt hillen has an enormous amount of products and customers. Therefore, transparency is desired in the performance of products and their relation with other products and customers. At last, Spruyt hillen has targets on the performance of their assortment. However, better insight in this performance is desired. Above that, it is also desired that the user of the tool can control these performance measures. After having considered the current assortment planning of Spruyt hillen, it can also be concluded that their assortment decisions can be improved with the help of a decision support tool.
3. ANALYSIS

In the Orientation phase, it was considered how assortment decisions are currently made at Spruyt Hillen. In this chapter, these findings will be reflected in relation to the literature on assortment planning. Furthermore, it is also considered why other aspects which are common in the academic literature on assortment planning, do or do not apply to Spruyt Hillen. In the end, it is decided which aspects to incorporate into the tool and which not, resulting in the scope of the project and the answer to main research question 1.

3.1 BASKET SHOPPING

As discussed in chapter 2, the fact that customers buy more than one product, a bundle or basket of products, from Spruyt Hillen is important. Therefore, products cannot be seen in isolation. This phenomenon is called basket shopping. Reasons for this behavior are time-saving convenience and fixed costs associated with a store visit (Messinger et al., 1997). Customers make decisions about multiple categories when they shop (Manchanda et al., 1999). Products can be in the same basket because they are complementary to each other, for example pancake mix and milk or because of similar purchase cycles, for example beer and bread (Manchanda et al., 1999). This first reason is something which give companies some control over the buying behavior of customers, whereas the second reason is coincidentally and less controllable. Furthermore, Bell and Latin (1998) found that store choice is based on the total basket utility. So if a customer wants a specific product and for some reason this product is not carried by the retailer, it is possible that the customer buys the entire basket somewhere else. Borle et al. (2005) showed that a reduction of the assortment leads to a decrease in shopping frequency and purchase quantity. This reduced shopping frequency also leads to a reduction in overall store sales. Cachon and Kök (2007) found that a decentralized assortment planning, where categories are managed independently from each other lead to lower profits than optimal. This is because this decentralized planning ignores cross-category interactions.

3.1.1 BASKET ANALYSIS

Since, the basket shopping phenomenon is important for Spruyt Hillen, it is important to analyze the baskets in more detail. There are two basket definitions to be distinguished. A basket can be defined as the set of products bought during a specific shopping occasion, see also Bell and Lattin (1998), Manchanda et al. (1999), Agrawal and Smith (2003) and Cachon and Kök (2007). In Table 3, the average basket size with the standard deviation per transaction is displayed per customer group.

<table>
<thead>
<tr>
<th>Customer group</th>
<th>Average basket size</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pharmacists</td>
<td>2.12</td>
<td>2.07</td>
</tr>
<tr>
<td>Hospitals</td>
<td>2.25</td>
<td>2.36</td>
</tr>
<tr>
<td>Industry food</td>
<td>1.35</td>
<td>0.49</td>
</tr>
<tr>
<td>Industry veterinary</td>
<td>6.19</td>
<td>6.15</td>
</tr>
<tr>
<td>Industry distribution</td>
<td>8.31</td>
<td>9.78</td>
</tr>
</tbody>
</table>
The average size of the baskets based on transactions is rather low. However, a basket can also be defined as the set of products bought by a customer during a certain period, see also Flapper et al. (2010). The average basket size differs very much, based on which definition is used, see also Table 4.

A customer may demand a set of products, which he does not buy at the same shopping occasion. It can be possible that he stops buying all his products at Spruyt hillen when one product is removed. The transaction based definition cannot take this into account. Notice, that the basket per customer per period is simply built up out of all the transactions of that customer in that period. This is because every transaction is related to one customer. This also means that no valuable information with relation to basket shopping effects is lost, when considering the baskets per period. Therefore, it is desirable to use the basket definition that considers a basket as the products a customer bought during a period. After having analyzed the baskets of Spruyt hillen, it can be concluded that the basket shopping aspect is indeed relevant for Spruyt hillen.
3.1.2 BASKET SHOPPING MODELS

Agrawal and Smith (2003) developed an assortment planning model that is based on a specified structure for sets and substitutes. In this model, Agrawal and Smith (2003) considered multiple categories and the phenomenon that customers buy in sets. Each customer arrives with a demand for a set of products. If the set of products is not available, there are three options. Substitute with a smaller set, substitute with a different set or not buying anything.

Flapper et al. (2010) presented two mathematical models for two different strategies, a product based assortment and a customer based assortment. The customer based assortment accounts for the effects of basket shopping. They assumed that customers with unsatisfied demand are lost, because of the one-stop-shopping requirement. It turned out that the customer-based model is more profitable for the case done. The models of Flapper et al. (2010) are deterministic multi-period models and include inventory costs, setup times, setup costs and lost sales. There is no storage limitation, but there is a limited production capacity.

Another approach of studying assortment planning in combination with basket shopping is the use of data mining. With data mining it is possible to reveal relations between products based on historical sales data, also called association rules. Agrawal R. et al. (1993) introduced the discovery of association rules. An association rule can be a statement like: ‘70% of the people who bought Ointment X also bought Ointment Y’. In this statement 70% stands for the confidence of the rule. In this case Ointment X is called the antecedent and Ointment Y the consequent. Next to the confidence, they also introduce support as an important definition. The support of the rule is the frequency of \( X \cup Y \) in the set of transactions, which can be either expressed in absolute form or as a percentage.

Brijs et al. (2000) developed an assortment model, which makes use of this data mining. The model is called the PROFSET model and uses the cross-sales potential of products instead of the individual product contribution. The optimization model determines which sets of products should be in the assortment. Even though the products are not considered in isolation, but in relation to other products, the relation with customers is completely ignored. In a retail environment this will often do no harm, since individual customer contribution is relatively low. However, in a business to business (B2B) environment there can be customers with an enormous contribution. It is very undesirable to remove a low contributing product which causes a huge customer to leave.

3.2 OTHER ASPECTS FROM LITERATURE

In the academic literature, assortment planning is often studied in combination with substitution effects and/or inventory decisions. In this section will be discussed how these aspects relate to Spruyt hillen.

3.2.1 SUBSTITUTION AND CANNIBALIZATION

Substitution is the switch from the most preferred product to a less preferred product, because the most preferred product was not available. For a more comprehensive explanation of this phenomenon see the literature study in Appendix I. Spruyt hillen offers only one brand per product and therefore substitution between brands is not possible. Furthermore, in the
pharmaceutical industry the substitution to another product with a different substance is not yet that common. For some products in the assortment there are size variants, however it is estimated that the impact of including substitution is not very high. Therefore, in consultation with the company it is decided that substitution is out of scope.

Whereas substitution can occur between existing products in the assortment, the term *cannibalization* is used to refer to the reduction in sales of other products, when introducing a new product. According to Cachon et al. (2005) a retailer has to make a trade-off when adding a new product (variant) to the assortment between the benefit of extra sales and the costs of including the new product. This costs can also include a cannibalization effect, which results in a reduction in demand for other products and thus lowering their profit and operational efficiency (Cachon et al., 2005). At Spruyt hillen a new product is in some cases a successor of an existing product. In that case, Spruyt hillen is interested in the cannibalization effect associated with the introduction of that new product.

### 3.2.2 INVENTORY

In the literature, assortment planning models often include inventory levels of the products. This is because most research focuses on retailers which have limited shelf space in their store. It can also be possible that a company has a very limited capacity in their warehouse. Therefore most assortment planning models are restricted to storage space limitations. For Spruyt hillen these storage space restrictions are not that relevant. Of course, they do not have infinite storage space, but there are never problems with capacity in their warehouse and also for the coming years they expect that this will not be an issue. This has to do with the fact that they have a flexible supply chain, where not every product is held at the warehouse of Spruyt hillen. Some product are stored at the suppliers of Spruyt hillen and can be directly shipped to the customers of Spruyt hillen.

Since, storage space restrictions are not that relevant for Spruyt hillen, it is decided that the decision of how many items to stock of one product is uncoupled from the assortment decision. In this project, the interest will be in what stock keeping units (SKU’s) to carry in the assortment. The decision of how many items to carry for one particular SKU is out of scope.

### 3.3 CONCLUSION

After having considered the assortment planning at Spruyt hillen and the relevant literature on assortment planning, it is possible to define the scope of the project. Since, Spruyt hillen aims to be a one-stop-shop for pharmaceutical products, the interest in this project is especially on the basket shopping phenomenon. This is also supported by the basket analysis in section 3.1.1.

Three assortment models that incorporate basket shopping have been discussed. The model of Brijs et al. (2000) includes product relations but ignores customer retention. The customer-based model of Flapper et al. (2010) does take into account the customer retention. However, they use the one-stop-shop requirement, that assumes that a customer with unsatisfied demand is lost. In this project that is too rigorous. It is not necessarily the case that an entire customer is lost, when not having a certain product in the assortment. The model of Agrawal and Smith (2003) is the closest to the situation we are dealing with in this project. However, it requires knowledge of all possible subsets of products which still satisfy the customer. In the design
phase will be shown to which extend these models can be helpful for the solution design of the project.

On the other hand, it turned out that for Spruyt hillen inventory and substitution effects were less relevant in relation to assortment decisions and therefore these aspects are out of scope in this project. Though, for the introduction of new products cannibalization is considered as relevant. Therefore, this shall be included when new products are to be added to the assortment.

In chapter 2, it turned out that better insight and control is desired on the performance of the assortment. Therefore, restrictions with relations to the profit can be included in the tool in order to achieve targets on the assortment.

Now that all relevant aspects for assortment decisions at Spruyt hillen are revealed, *Main research question 1* is answered. In Table 5 it is summarized how this differs from the assortment planning models in the literature.

### Table 5: Literature overview

<table>
<thead>
<tr>
<th>Case study</th>
<th>Assortment planning</th>
<th>Basket shopping</th>
<th>Customer retention</th>
<th>Product relations</th>
<th>Substitution</th>
<th>Cannibalization</th>
<th>Inventory</th>
<th>Profit restrictions</th>
<th>Case study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flapper et al. (2010)</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>-</td>
<td>-</td>
<td>Yes</td>
<td>-</td>
<td>-</td>
<td>Yes</td>
</tr>
<tr>
<td>Agrawal and Smith (2003)</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Brijs et al. (2000)</td>
<td>Yes</td>
<td>-</td>
<td>Yes</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Yes</td>
<td>-</td>
</tr>
<tr>
<td>Spruyt hillen</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>-</td>
<td>Yes</td>
<td>-</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
</tbody>
</table>

In the next chapter, the implementation of these aspect into a decision-support tool is described.
4. DESIGN

In the last chapter, the relevant aspects with relations to the assortment decisions of Spruyt hillen were revealed. Now, it should be investigated how to implement these aspects into a decision support tool which corresponds to main research question 2. In this chapter the design of the solution will be described.

4.1 RESEARCH SUB-QUESTIONS

In order to structure the design phase of the project some research sub-questions are formulated. Main research question 2 is split up in the following sub-questions:

2.1. How to model the assortment planning optimization problem including basket shopping?
   a) Which assumptions need to be made?
   b) What are the (decision) variables?
   c) How can the objective function be formulated?
   d) What are the relevant constraints?

2.2. How can the created mathematical model be solved?
   a) Can an optimal solution be found in a reasonable time?
   b) If needed, can the problem be solved with a heuristic approximation method?

2.3. How can the mathematical model be implemented in a decision-support tool?
   a) Which software to use?
   b) How to make a user-friendly interface?
   c) How can the tool be verified and validated?
4.2 OPTIMIZATION MODELS INCLUDING BASKET SHOPPING

4.2.1 STARTING POINT: AGRAWAL AND SMITH (2003)

The model of Agrawal and Smith (2003) is the most suitable description of the basket shopping phenomenon and goes as follows. A customer demands a set of products each visit, the customers’ shopping basket. If all these items are available then the customer buys his entire basket and his initial demand is fulfilled. When one of the items is not available there are three alternatives a customer can do. The first option is to buy the set without the missing items. The second option is not to buy anything. In the article of Agrawal and Smith (2003) also a third substitution option is considered. However, as discussed in 3.2.1, in this project substitution is out of scope. To briefly illustrate the situation of Agrawal and Smith (2003) without substitution, consider the following example.

Example 4.1: A customer demands the following set of products \{A, B, D, E, F\}. Assume that product A is removed from the assortment. There are two options a customer can do:

Option 1: The customer would buy \{B, D, E, F\}.

Option 2: The customer buys nothing and all the demand is lost from that customer.

Agrawal and Smith (2003) consider a basket as the set of products during a shopping occasion. As discussed in 3.1.1, it is also possible to consider a basket as the set of products a customer bought in a certain period. The model of Agrawal and Smith (2003) also fits to this basket definition. It is then required to know all subsets of products which are still acceptable for each customer. From section 3.1.1, it follows that the average basket size for customers of Spruyt hillen per year is 37. So, there are potentially \(2^{37} = 137,438,953,472\) subsets for an average customer. For a larger customer with a basket of 100 products, there are even \(2^{100} = 1.267 \cdot 10^{30}\) subsets of products. Of course, the actual number of satisfying subsets is much less, however this gives an idea about the size of the problem. Furthermore, the subsets which are still acceptable for each customer is information which is not available in this case. It is not desirable to estimate all possible satisfying sets of products per customer because it will cost an enormous amount of time, if it would be possible at all.

4.2.2 ALTERNATIVE OPTIMIZATION MODEL

As we do not know the satisfying subsets of products for each customer, another approach is required. Instead of using the satisfying subsets of the different customers, two so-called f-variables are introduced. The \(f^1_{p+p^*}\) variable represents, whether the demand for a complementary product \(p\) is lost when removing product \(p^*\). The \(f^2_{pc}\) variable represents the customer retention of customer \(c\), when product \(p\) is removed from the assortment.

\(f^1_{p+p^*}\): Binary variable, indicating whether the demand for product \(p\) for all customers is lost (1) when \(p^*\) is removed from the assortment, or not (0).
Designing a decision support tool for assortment planning

$f^2_{pc}$: Binary variable, indicating whether customer $c$ is lost (1) when $p$ is removed from the assortment, or not (0).

This deviates from the situation described in 4.2.1, since the $f^1_{pp'}$-variable relates one product to one other product. E.g.: if A is removed from the assortment, then the demand for C is lost, so $f^1_{AC} = 1$. However, it could also be the case that the demand is lost when a set of products is left out of the assortment, so that customers stop buying C, when A and B are both removed from the assortment, and not when either A or B is out of the assortment. The $f^1_{pp'}$-variable does not take this into account. Furthermore, the $f^1_{pp'}$-variable holds for all customers and cannot be specified per customer. The $f^2_{pc}$-variable relates one product to one customer, so a customer can be lost when a certain product is removed. Just as the $f^1_{pp'}$-variable, the $f^2_{pc}$-variable cannot specify that a customer is lost when a certain set of products is removed from the assortment.

Even though this model deviates from the ideal situation of Agrawal and Smith (2003), most basket shopping effects at Spruyt hillen can still be captured with the $f$-variables, therefore this method is still justified. It can for example occur that a product is either a required product for a customer, so when the product is removed the customer is possibly lost. Or on the other hand, a product can be related to another product, due to complementarity which hold for all customers who buy the product. So removing the product can possibly lead to a loss in demand for another product.

A complete mathematical model including these $f$-variables can be found in Appendix II. Although mathematically correct, there are still two reasons why this model cannot be used in this project. First, the $f$-variables should be defined for an enormous amount of combination. The number of combination is approximately $P \times P + P \times C \approx 4500 \times 4500 + 4500 \times 4200 \approx 40,000,000$. This huge matrix might be filled for a huge part with zero’s, but it is still an enormous amount of values to be estimated. On top of that, products and customers change over time, therefore it only holds for a short period of time, since $f$-variables must also be defined for new products and customers. Also taking into account to possible roll out to other Fagron Group companies, it is undesirable to use the model since the $f$-variables must be defined for each company which is a very time-consuming task. Second, it is not easy to solve the model. The model is a MINLP and with using only 3 products and 3 customers, the used solvers KNITRO 8.0 and AIMMS AOA could only find a locally optimal solution. So, it never can be guaranteed that the optimum found is a global optimum. Considering these two issues, it is decided that a heuristic approach is desired. The answer to research question 2.2a is therefore that it is not possible to find a solution in a reasonable amount of time.
4.3 HEURISTIC

In the previous section, it turned out that a mathematical model including the basket shopping effect was not feasible in this project. It is therefore decided to exclude the basket shopping effect from the mathematical model. What is left over, is a product-based assortment model. This model is used to make a distinction between well performing products and so-called candidate products, see Figure 4.

![Figure 4: Heuristic Approach Overview]

The mathematical model contains performance constraints related to the profit on products, product groups and the total assortment. The parameters for these performance constraints are inputs for the tool and can be set by the user. In section 4.3.1 the model and constraints will be described in more detail. For the well performing products, it is advised to keep them in the assortment. Since, they satisfy the performance constraints there is no reason to remove these products from the assortment. The candidate products are the products that do not perform well enough according to the constraints in the mathematical model. However, before removing a candidate product a further analysis is done in order to advise that the product should actually be removed from the assortment. In this analysis the relation with other products is considered. As discussed in chapter 2, the relation with other products is important since products in the assortment of Spruyt Hillen can be complementary. This can cause a loss in demand for other products, when a candidate product is removed. Next to the relation with other products, the customers who bought the candidate product are analyzed. When a customer wants a certain product and this product is removed from the assortment, it is possible that the customer buys his entire basket somewhere else. It is undesirable to lose an important customer, when removing a product from the assortment. Therefore, a product cannot be seen separately from...
the customers who buy the product. So, the analysis of the candidate product is needed in order to account for the basket shopping phenomenon. The details of the mathematical model are described in the next section. In section 4.3.2, the analysis of candidate products is described.

The heuristic starts with a product-based model. One could also argue why not to start with a customer-based model. If the heuristic would start with a customer-based model with for example a constraint with a minimum contribution per customer, the products that are bought by the most profitable customers would be in the assortment. It can be that un- or less profitable products are now in the assortment, because they are bought by a highly profitable customer. But it is not necessarily the case that removing this unprofitable product, leads to a loss of the customer. This leads to products that are held in the assortment unnecessary. So, the whole list of products should be analyzed after all in order to remove these un- or less profitable products. In the case study of Flapper et al. (2010), where the customer-based model originates from, a strict one-stop-shop requirement holds. That implies that a customer with unsatisfied demand is immediately lost, however this is not the case in this project. In the end, Spruyt hillen is interested in which products to remove from the assortment. Therefore, it is best to start with a product-based model, since it distinguishes well performing products from candidate products based on performance constraints. However, before removing a candidate product, the relations to other products and the customers of the candidate product are checked.

4.3.1 MATHEMATICAL MODEL

In this section the mathematical model with which the heuristic starts is described.

4.3.1.1 ASSUMPTIONS

First, the assumptions underlying the mathematical model are given, which are also the answer to research question 2.1a.

Assumptions

1) The length of the period is a function of a year.
2) The demand for the coming year equals demand during last year.
3) Substitution does not occur between products within the assortment.
4) There are no storage space restrictions.
5) The demand can always be directly fulfilled.

Argumentation

1) For some products the demand faces seasonality with cycles of a year. An example is hay fever medication which is sold more often during springtime. Therefore, the period should be as function of a year, otherwise the tool could give a biased output. E.g. if data of half a year is used it could suggest to remove a product which is only sold during a certain period in a year. The period is a variable in the tool and can be changed by the user. The default period is one year.
2) In the ideal situation, forecasts would be used to determine the demand. Spruyt hillen carries approximately 4500 products but forecasts are only available for a small proportion of these products. Furthermore, it would be difficult to even use these forecasts, since the model uses demand per product per customer per transaction,
whereas the forecast of the demand is based on the total demand per product. In consultation with the company it is therefore decided that last year’s sales are used.

An analysis is done for the two most important customer groups to what extend customers tend to buy the same products over the years. The sales data from 2012 and 2013 are used. Table 6 shows that on average 60.95% of the products bought in 2012 are also bought in 2013 by Hospitals. For the Pharmacists this percentage is 64.64%.

<table>
<thead>
<tr>
<th>Customer group</th>
<th>Percentage of products also bought in 2nd year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hospitals</td>
<td>60.95%</td>
</tr>
<tr>
<td>Pharmacists</td>
<td>64.64%</td>
</tr>
</tbody>
</table>

In this analysis all products are taken into account. This also contains for example spot sales and promotional products. It can happen that these products are not available in the second year. Therefore, the percentage is not as high as one would expect, but customers still tend to buy the same products over the years. So, it can be concluded that assumption 2 is acceptable.

3) Spruyt Hillen offers only one brand per product. In a few cases, there are some size variants. Therefore, it is decided that substitution is out of scope, see also 3.2.1.

4) Spruyt Hillen has no storage space restrictions. There are no problems with capacity in their warehouse and also for the coming years they expect that this will not be an issue.

5) No inventory model is included in the model. The interest in this project is especially in the strategic decision which products to carry. The operational decision of how many items to stock of each SKU is out of scope as discussed in 3.2.2. It is therefore assumed that demand can always be directly fulfilled.

4.3.1.2 LIST OF VARIABLES

The variables which are used in the mathematical model are listed below, which is also to answer to research questions 2.1b. The $X_p$-variable is the decision variable of the model, which will be determined by the solver.

$c, c \in \{1 \ldots C\}$: Index denoting the customers.


$d_{tcp}$: Quantity of product $p$ for customer $c$ ordered in transaction $t$.

$g, g \in \{1 \ldots G\}$: Index of product groups under consideration.

$MAP1P$: Minimal absolute profit margin per product.

$MRPPG_g$: Minimal relative profit margin per product group $g$.

$MRPTA$: Minimal relative profit margin on the total assortment.

$p, p \in \{1 \ldots P\}$: Index denoting the products.
Designing a decision support tool for assortment planning

\[ t, t \in \{1 \ldots T\} \text{: Index denoting the transactions.} \]

\[ TC: \text{ Total cost of goods sold over the period.} \]

\[ TP: \text{ Total profit over the period.} \]

\[ TR: \text{ Total revenues over the period.} \]

\[ u_{tcp}: \text{ Adjustment factor for the unit of measure.} \]

\[ w_{tcp}: \text{ Sale price of product } p \text{ for customer } c \text{ in transaction } t. \]

\[ X_p \in \{0,1\}: \text{ Decision variable which indicates if a product } p \text{ is in the assortment (1), or not (0).} \]

### 4.3.1.3 OBJECTIVE FUNCTION

The function that should be maximized is the total profit over the period. Spruyt hillen is interested in what the profit is for the company. The total profit is the total revenue minus the cost of goods sold. The total revenue is calculated by multiplying the sales price of a product with the demand for that product. The cost of goods sold are calculated by multiplying the cost price of a product with the demand. The cost of goods sold should be multiplied by an adjustment factor for the unit of measure. In some transactions the unit of measure can deviate from the Base Unit of Measure (BUM). The BUM is the standard unit of measure for a product, e.g.: units, grams or kilograms. Example 4.2 illustrates why this adjustment is needed. Of course, only products that are in the assortment can contribute to the total profit.

\[
\max TP = TR - TC \tag{4.1} \text{ Total profit over the period}
\]

Where

\[
TR = \sum_{t=1}^{T} \sum_{p=1}^{P} \sum_{c=1}^{C} w_{tcp} \cdot d_{tcp} \cdot X_p \tag{4.2} \text{ Total revenue}
\]

\[
TC = \sum_{t=1}^{T} \sum_{p=1}^{P} \sum_{c=1}^{C} c_p \cdot u_{tcp} \cdot d_{tcp} \cdot X_p \tag{4.3} \text{ Cost of goods}
\]
After having formulated the objective function, research question 2.1c is answered as well.

4.3.1.4 CONSTRAINTS

The decision variable $X_p$ should be a binary variable, since a product should be either in or out the assortment.

$$\forall p, p \in \{1, ..., P\} \quad X_p \in \{0, 1\} \quad (4.4) \text{ Binary bound}$$

In section 2.4, the KPI’s were presented. The reference to the formulas of the KPI’s are displayed in Table 7 below. The formulas can be found in Appendix V. Some of these KPI’s will function as a performance constraint for the mathematical model.

**TABLE 7: KEY PERFORMANCE INDICATORS**

<table>
<thead>
<tr>
<th></th>
<th>Total assortment</th>
<th>Product group</th>
<th>Individual Product</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absolute profit margin</td>
<td>(1.13.13)</td>
<td>(1.10.6)</td>
<td>(1.10.10)</td>
</tr>
<tr>
<td>Revenues</td>
<td>(1.13.11)</td>
<td>(1.10.4)</td>
<td>(1.10.9)</td>
</tr>
<tr>
<td>Costs of goods sold</td>
<td>(1.13.12)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Relative profit margin</td>
<td>(1.13.14)</td>
<td>(1.10.8)</td>
<td>(1.10.11)</td>
</tr>
<tr>
<td>Number of products</td>
<td>(1.13.10)</td>
<td>(1.10.2)</td>
<td>N/A</td>
</tr>
</tbody>
</table>

The performance constraints that are used in the tool are relative profit over the total assortment ($MRPTA$), relative profit per product group $MRPPG_g$ and an absolute profit per product ($MAPIP$). This has been decided in consultation with Spruyt hillen. Notice that there are only performance constraints on the profit measures. Since the profit directly relates to the revenues and costs, this is still acceptable.

---

1 This KPI is also used as a performance constraint.
A relative profit margin on the assortment and per product group are included, since they have target levels for these performance constraints. The total absolute profit over the total assortment (MAPTA) would be a bit superfluous, since it is a constraint on the objective function. The data either satisfies the constraint or not. So, there will be either no candidate product at all or the model will not be solvable, which is undesirable. The same holds for the absolute profit margin per product group (MAPPGg).

An absolute profit margin per product is included (4.7). It could be that a product has a very high relative margin – say 90% - and therefore is not thrown out by either constraint (4.5) or (4.6). However, if the contribution is only 10 euro, it is questionable how much it actually yields, since there are always ‘hidden costs’ associated with having a product in the assortment. The cost price calculation of Spruyt hillen does not include all of these costs. Inventory and handling costs are for example not included in this cost price. So, the performance constraints included are:

\[
\frac{TR-TC}{TR} \geq MRPTA \tag{4.5}
\]

Relative profit margin restriction total assortment.

\[
\forall g, g \in \{1, \ldots, G\} \sum_{t=1}^{T} \frac{\sum_{p \in g} \sum_{c=1}^{C} w_{tcp} d_{tcp} X_p - \sum_{t=1}^{T} \sum_{p \in g} \sum_{c=1}^{C} c_p \cdot d_{tcp} \cdot u_{tcp} \cdot X_p}{\sum_{t=1}^{T} \sum_{p \in g} \sum_{c=1}^{C} w_{tcp} d_{tcp} X_p} \geq MRP\text{PG}_g \tag{4.6}
\]

Relative profit margin restriction product group g.

\[
\forall p, p \in \{1, \ldots, P\} \sum_{t=1}^{T} w_{tcp} \cdot d_{tcp} \cdot X_p - \sum_{t=1}^{T} c_p \cdot d_{tcp} \cdot u_{tcp} \cdot X_p \geq MAP\text{IP} \cdot X_p \tag{4.7}
\]

Absolute profit margin restriction per product p.

The values for the thresholds in these constraints can be set by the user. These are user inputs for the tool, with which the user can influence the performance of the assortment. Now that all the constraints for the model are answered, research question 2.1d is answered.

---

### 4.3.2 CANDIDATE ANALYSIS

After the mathematical model is run, there is a list of candidate products to be removed from the assortment. As explained in 4.3, the candidate products should be analyzed. The following steps are required for this analysis, see Figure 5 down below.

**FIGURE 5: CANDIDATE ANALYSIS**
For each candidate product the user can see which customers bought the candidate product, but also what other products are related to the candidate product. Products are related if they occur in the same basket. For each candidate product, the user can make assumptions about the retention of demand of related products and customers. These assumptions correspond with the determination of the f-variables displayed below, also previously mentioned in 4.2.2.

\[ f^1_{p*p} : \text{Binary variable, indicating whether the demand for product } p \text{ for all customers is lost (1) when } p^* \text{ is removed from the assortment, or not (0).} \]

\[ f^2_{pc} : \text{Binary variable, indicating whether customer } c \text{ is lost (1) when } p \text{ is removed from the assortment, or not (0).} \]

In this heuristic, the f-variables only need to be specified for the candidate products. In contrast to the optimization model in 4.2.2, where potentially about 40,000,000 f-values needed to be determined, it is now a task which can be done in a relatively short time.

**Customers**

An overview of the customers who bought the candidate product can easily be generated. These are simply the customers who bought the candidate product in the last year. The customers are listed and sorted on the total absolute profit of a customer from high to low. For each candidate product, the user can specify which customers are possibly lost. The customers who are lost when removing the candidate product are from now on called the *critical customers* for a candidate product, see also Figure 6.

**Related products**

The relation to other products requires a more extensive explanation. A product is related to a candidate product when the product has been in the same basket in the period. There are two basket definitions to be distinguished. The set of products a customer bought during the period (I) or the set of products in one transaction (II).
Spruyt hillen is a one-stop-shop, but complementary products are not necessarily bought at the same shopping occasion. So if only the transaction based relation is used, product relations could be missed. Hence, the relations based on a basket per customer per year is used as well. One could argue, not to display the transaction based relations anymore, since the basket per customer per year is built up out of the transactions. However, when there is also a strong relation based on transactions, the probability of complementarity is higher. So, displaying the relations based on transactions will provide additional useful information.

In this project three measures are used to classify the related products. This classification is desired, because products that are very often bought together are more likely to be complementary, than products which are less often bought together.

In section 3.1.2, association rules were introduced, which can be a statement like: ‘70% of the customers who bought Ointment X also bought Ointment Y’. In this statement 70% stands for the confidence of the rule Agrawal R. et al. (1993). The confidence of the rule, which is the first measure, is used in order to sort the related products, from highest confidence to lowest confidence. The confidence is calculated on the basis of the support of a product, which is the proportion of customers who buy the product, or the proportion of transactions that contain the product. An example of the measures can be found in example 4.3. In a formula it can be written as follows

- Confidence of the association rule
  
  \[ \text{confidence}(X \rightarrow Y) = \frac{\text{support}(X \cup Y)}{\text{support}(X)} \]

Next to the confidence, the confidence of the reverse association rule is displayed. E.g. ‘50% of the customers who bought Ointment Y also bought Ointment X’ or ‘50% of the transactions containing Ointment Y also contains Ointment X’. This measure is introduced, because of the following reason. An association rule might have a very high confidence, due to the fact that a related product is simply a very often sold product. So the chance that a customer buys the related product anyway is already very high. The confidence of the reverse association rule is therefore used as well, to check to what extend the reverse relation holds. If the confidence of the association rule is high and the confidence of the reverse association rule is low, it is possible that the confidence is high because the related product is simply sold very often.

- Confidence of the reverse association rule
  
  \[ \text{confidence}(Y \rightarrow X) = \frac{\text{support}(X \cup Y)}{\text{support}(Y)} \]

The third and last measure takes into account the frequency of both products into account directly, whereas the confidence only takes into account the frequency of the combination \((X \cup Y)\) and one other product \((X)\) or \((Y)\). This measure is called the interest measure (Silverstein et al., 1998) or often referred to as the lift measure. The lift measure can be used as a guidance to identify complementarity (Brijs et al., 2004). In a formula this lift measure can be written as follows

- Lift of the association rule

\[ \text{lift}(X \rightarrow Y) = \frac{\text{support}(X \cup Y)}{\text{support}(X) \times \text{support}(Y)} \]
Designing a decision support tool for assortment planning

\[ \text{lift}(X \rightarrow Y) = \text{lift}(Y \rightarrow X) = \frac{\text{support}(X \cup Y)}{\text{support}(X) \times \text{support}(Y)} \]

If the value of the lift is above 1, this indicates positive dependency, which means that this combination occurs more often in the data than expected. A value below 1, indicates negative dependency and thus the combination occurs less often than expected. Furthermore, the value of the lift in absolute way is rather meaningless, however it can be very well used for comparing relations (Silverstein et al., 1998).

**Example 4.3:** Consider the following transaction history, with 7 transactions, 3 customers and 3 products.

<table>
<thead>
<tr>
<th>Transaction</th>
<th>Customer</th>
<th>Product</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>K</td>
<td>{A,B}</td>
</tr>
<tr>
<td>2</td>
<td>K</td>
<td>{A,C}</td>
</tr>
<tr>
<td>3</td>
<td>K</td>
<td>{A}</td>
</tr>
<tr>
<td>4</td>
<td>L</td>
<td>{A,B}</td>
</tr>
<tr>
<td>5</td>
<td>M</td>
<td>{B}</td>
</tr>
<tr>
<td>6</td>
<td>M</td>
<td>{C}</td>
</tr>
<tr>
<td>7</td>
<td>L</td>
<td>{A}</td>
</tr>
</tbody>
</table>

Product A is bought 5 times, product B 3 times and product C twice. Product A and B are both twice together, A and C occurred once together in a basket and product B and C are not bought together. The relations based on transactions are as follows.

<table>
<thead>
<tr>
<th>Products</th>
<th>Support</th>
<th>Support</th>
<th>Support</th>
<th>Confidence</th>
<th>Confidence</th>
<th>Lift</th>
</tr>
</thead>
<tbody>
<tr>
<td>(X,Y)</td>
<td>X</td>
<td>Y</td>
<td>(X∪Y)</td>
<td>(X→Y)</td>
<td>(Y→X)</td>
<td>(X→Y)</td>
</tr>
<tr>
<td>(A,B)</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>0.40</td>
<td>0.67</td>
<td>0.93</td>
</tr>
<tr>
<td>(A,C)</td>
<td>5</td>
<td>2</td>
<td>1</td>
<td>0.20</td>
<td>0.50</td>
<td>1.4</td>
</tr>
</tbody>
</table>

The relations based on customer per period are as follows.

<table>
<thead>
<tr>
<th>Products</th>
<th>Support</th>
<th>Support</th>
<th>Support</th>
<th>Confidence</th>
<th>Confidence</th>
<th>Lift</th>
</tr>
</thead>
<tbody>
<tr>
<td>(X,Y)</td>
<td>X</td>
<td>Y</td>
<td>(X∪Y)</td>
<td>(X→Y)</td>
<td>(Y→X)</td>
<td>(X→Y)</td>
</tr>
<tr>
<td>(A,B)</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>1.00</td>
<td>0.67</td>
<td>1</td>
</tr>
<tr>
<td>(A,C)</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0.50</td>
<td>0.50</td>
<td>0.75</td>
</tr>
</tbody>
</table>

The measures for determining the strength of the relation are important. To recall, in the current method of assortment planning at Spruyt hillen, a product is sometimes held in the assortment, because its removal could cause a loss in demand of the complementary product. However, this is not supported by any data. In the tool, the user can see what the relation is with other products, expressed in the measures mentioned. In some cases it will confirm the user that products are indeed complementary. However, it is also possible that products are not related,
whereas the user thought they were complementary and thus strongly related. In that case, it
would be questionable to keep a candidate product in the assortment for this weakly related
product.

Just as for the customers, the user can specify the products for which the demand is lost when
the candidate product is removed. These related products are from now on called the critical
products for a candidate product.

To summarize, a candidate product is bought by customers, for which the user can specify
whether a customer is critical or not. Next to customers, a candidate product can have related
products. For each related product, the user can specify whether that product is critical or not,
see also Figure 6.

**Default value of f-variables**

The f-variables are by default 0. It could be argued to set the f-variable to 1 for a certain level of
confidence. E.g. two products A and B, with confidence (A→B) 70% and (B→A) 80%. This
relation is pretty strong and the default could for this case be set to 1. However, the fact that they
are often bought together does not necessarily say that removing one of the two products has an
impact on the other product when the product is removed. If this is indeed not the case, the user
should manually set the value back to zero for this relation. It can be confusing for the user that
on the one hand values should be changed from 0 to 1 and on the other hand from 1 to 0. This is
undesirable and it is therefore in consultation with the company decided that the f-variables
are by default always zero.

**Classification of customers**

Whereas the performance of products is determined by the mathematical model, there is no
classification for customers. The reason why this is desired at this point can be explained as
follows. If a candidate product is bought by a customer that bought a few products incidentally
in the last year with a low contribution, it is questionable to leave the candidate product in the
assortment. Therefore a customer classification will be made, which distinguishes between
important and less important customers, based on the total amount of profit a customer
contributes in a year. The threshold value – the minimal customer margin - is an input that can
be set by the user of the tool. The customers who satisfy the following restriction are classified
as important customers.

\[
\sum_{t=1}^{T} \sum_{p=1}^{P} w_{tcp} \cdot d_{tcp} \cdot X_p = \sum_{t=1}^{T} \sum_{p=1}^{P} c_p \cdot d_{tcp} \cdot u_{tcp} \cdot X_p \geq \text{Minimal customer margin}
\]

(4.8) Customer classification

*Note: this inequality is not a restriction in the mathematical model in 4.3.1. It is used for
distinguishing important customers from less important customers.*

**Final decision on a candidate product**

The decision on whether a candidate product should be removed or should be kept in the
assortment goes as follows. The procedure starts on the top left, after which several questions
should be answered. Depending on the answers on these questions, there are two possible
outcomes, either removing or keeping the candidate product in the assortment.
For each candidate product it can be questioned whether the product is bought by at least one important customer. If this is the case, it should be questioned whether removing the candidate product causes important customers to leave the company. When at least one important customer will start to buy all of his products somewhere else - and thus considered to be a critical customer - the candidate product should be kept in the assortment. When, removing the candidate product has no influence on the important customers or has no important customers at all, the related products should be considered.

The first thing to consider is if the related products are well-performing. If the related products are also candidate products or if there are no related products, it is advised to remove the candidate product. However, if the related products of the candidate product are well-performing, the impact of removing the candidate product should be questioned. If the removal of the candidate product leads to a loss in demand for at least one of these well performing products, the candidate product should be kept in the assortment. On the other hand, if there is no impact on the related products which are well performing, the candidate product may be removed from the assortment.

**Sequence of processing the candidate products**

Since, a heuristic approach instead of an optimization model that includes basket shopping is used, there are some implications. Each candidate product is considered separately and assumptions are made on the product and customer retention. Making these assumptions has impacts on the results. If the numbers the assumptions are based on, are updated, each time an assumption is made, the sequence through the candidate products is important. A different order can lead to different results. Example 4.4 gives an illustration of this effect.
So, when the numbers where the decisions are based on, are updated after each candidate product is processed, it is important to use a fixed sequence. The probability that products are kept in the assortment ‘wrongly’, is the highest in the beginning, when no candidate products have been processed so far. This probability continues to decrease more when more assumptions on product and customer retention are made. It is better to keep the candidate products with the higher margins in the assortment, than candidate products with lower margins. Therefore, the best sequence is to start with the products with the highest absolute margin.

To force the user to go through the fixed order, the candidate products should follow up in consecutive screens. If there are a lot of candidate products, the user should click through a lot of screens. It is possible that he does not need or want to make assumptions on customer and product retention for a lot of these candidate products. Therefore, the candidate products will not follow up in consecutive screens, but are listed in a table. In that case it cannot be guaranteed that the user sticks to the sequence. It is therefore decided that the values will not be updated and thus the order of processing the candidate products is not important anymore. The impact of this decision is that the model is less rigorous. It can happen that it will be advised that some products should be in the assortment, while they might have been left out when updating values. However, it will never be the case that a product is left out the assortment which should have been in the assortment. The idea is that bad performing products will be removed in the end.

Now that the heuristic is completely described, research question 2.2b has been answered as well. It can be concluded that the problem can indeed be approached with a heuristic.
4.4 ADDITIONAL FUNCTIONALITIES

Fixing Products

An option which is incorporated in the tool is the ability to ‘fix products’, i.e. the user can specify products which should be in the assortment anyway. The optimization model will not decide over these products. The user can have several reasons for this. E.g.: the product is still quite new and sales are disappointing until now, but the user has indications from the market that in the nearby future the demand will start to rise. The formal description of the functionality of fixing products looks as follows. First the set of fixed products should be defined.

\[ FP: \text{ Set of fixed products, where } FP \subseteq \{1, \ldots, P\} \]

Then, the following constraint should be added to the mathematical model.

\[ \forall p \in FP \quad X_p = 1 \]

(4.9) Fixed products

Furthermore two constraints should be modified. It can be possible that the user wants to fix a product which has a contribution less than the required contribution per product. In that case (4.7) will not hold and therefore this constraint should be changed into (4.7*).

\[ \forall p, p \in \{1, \ldots, P\} \setminus FP \quad \sum_{t=1}^{T} \sum_{c=1}^{C} (w_{tcp} - c_p) \cdot d_{tcp} \cdot X_p \geq MAPIP \cdot X_p \]

(4.7*) Absolute profit margin restriction per product \( p \).

Adding New Products

New products are not yet entered in the ERP-system. On request of the company an option is added in order to add products to the dataset. For these products, it is possible to enter the data manually. With the entered data, the tool will incorporate the data of this new product and the user can see what the forecasted performance is of the new product. For this new product not only the sale price, cost price and forecasted demand can be entered. It is also possible to select products that are cannibalized and to specify the cannibalization rate. It is also possible that a new product will have a positive impact on the existing assortment. The user can specify the products that will generate extra sales when introducing the new product. For these products it is required to fill in the percentage of increase in demand. The addition of new products is summarized in Figure 8 down below.
Designing a decision support tool for assortment planning

The calculation of the net gain of a new product requires some calculations. First, the following two variables are introduced.

\( q_{p^*p} \in [0,100] \): Cannibalization rate from product \( p^* \) to \( p \). If new product \( p^* \) is added to the assortment, \( \frac{q_{p^*p}}{100} \) of the demand for product \( p \) is lost because of cannibalization related to \( p^* \).

\( s_{p^*p} \in [0,100] \): Percentage of increase in sales for product \( p \) when introducing new product \( p^* \). If new product \( p^* \) is added to the assortment, the demand for product \( p \) increases with \( \frac{s_{p^*p}}{100} \).

**NP:** Set of new products, where \( NP \subseteq P \)

The cannibalization loss associated with the introduction of product \( p^* \) can be specified as follows.

\[ \forall p^* \in NP \quad \sum_{p=1}^{P} q_{p^*p} \cdot \sum_{t=1}^{T} \sum_{c=1}^{C} (w_{tcp} - c_p) \cdot d_{tcp} \cdot X_p \]

(4.10) Cannibalization loss \((p^*)\)

The increase in demand associated with the introduction of product \( p^* \) can be specified as follows.

\[ \forall p^* \in NP \quad \sum_{p=1}^{P} s_{p^*p} \cdot \sum_{t=1}^{T} \sum_{c=1}^{C} (w_{tcp} - c_p) \cdot d_{tcp} \cdot X_p \]

(4.11) Gains of extra sales \((p^*)\)
The net gain of the introduction of a new product is calculated by the profit of the product minus the cannibalization loss plus the gains of extra sales.

$$\forall p^* \in NP \sum_{t=1}^{T} \sum_{c=1}^{C} (w_{tcp^*} - c_{p^*}) \cdot d_{tcp^*} \cdot X_{p^*} - \text{Cannibalization loss } (p^*) + \text{Gains of extra sales } (p^*)$$

(4.12) Net gain \( (p^*) \)

The decision whether it is advised to actually introduce the new product, depends on the net gain of the product. If this is larger than zero, it is advised to introduce the product. If not, then it is recommended not to introduce the product at this time.

4.5 CONCLUSION

After having evaluated two mathematical optimization models including basket shopping it was concluded that a heuristic approach is required in this project. The heuristic starts with a product-based model in order to distinguish between well performing products from candidate products. However, before removing candidate products, the relations to other products and their customers are examined. Two additional functionalities are included. First, the ability to fix products, which excludes them from the decision making process and let them be in the assortment anyhow. Second, it is possible to add new products and see how they may perform in relation to the existing assortment. Now that the design of the heuristic is completed, the implementation of the heuristic into software can be done in the next chapter.
5. IMPLEMENTATION INTO SOFTWARE

In the last chapter the design of the heuristic approach is described. In this chapter the implementation into software is described.

5.1 CHOICE OF SOFTWARE

It is decided that tool will be developed in the software package called *Advanced Interactive Multidimensional Modeling System* (AIMMS). AIMMS combines an integrated development environment for optimization models and the design of a graphical user interface (GUI). It is relatively easy to create a nice-looking tool without knowing any programming languages. The tool is developed in version 3.13 with a educational stand-alone license. AIMMS supports a variety of optimization solvers which can solve the following problem types: LP, MIP, NLP, QP, MIQP, MIQCP, MCP, MINLP and COP. Furthermore, it supports the database connectivity ‘Open DataBase Connectivity’ (ODBC). With this ODBC connection it is possible to connect directly to the database of the ERP system and to automatically load the data into the tool. This has the advantage that the user of the tool does not need to prepare datasets for the use of the tool. It also makes the roll out over Fagron Group companies easier. The tool only needs to be connected to another database. Now that the decision on the choice of software is made, research question 2.3a is answered.

5.2 TOOL DESIGN

In Figure 9 a schematic overview of the tool is displayed. A square stands for a page or screen in the tool, whereas the curved rectangles are actions that can be performed. The starting page is the ‘Tool dashboard’, which is the page that opens when the tool is started up.

![FIGURE 9: SCHEMATIC OVERVIEW OF THE TOOL](image-url)
In Figure 10 a screenshot of the tool dashboard page is displayed. On the tool dashboard page it is possible to load the data from the ERP-system. It is also possible to manually add (new) products, which are not entered in the ERP system yet. Furthermore, the user can clear the data if he wants to do a new analysis or an analysis on a different data set. On the tool dashboard page, the user can also select which product groups to include or exclude in the analysis. From the dashboard it is possible to start an optimization run. Via the navigation menu the different pages of the tool can be reached.

![FIGURE 10: TOOL DASHBOARD PAGE](image)

On the data page all products and customers are displayed that are loaded into the tool. For each product the relations with other products can be viewed. For the customers, the basket overview of products bought last year can be obtained.

The charts page displays two graphs on the assortment. In one graph the cumulative absolute profit and cumulative revenues are presented. In the second graph the cumulative relative profit is displayed. In both graphs the products are sorted on the basis of absolute profit. In section 1.8 of Appendix III, these figures are shown. These graphs are included in the tool, because Spruyt hillen explicitly requested this. It gives them insight in how their assortment is built up.

The model input page, is the page where the user can set the parameters for the mathematical model. The mathematical model contains three performance constraints, for which parameters must be set. As discussed in the design of the heuristic, there is also a threshold needed in order to distinguish important from less important customers. This threshold value can also be set on the model input page.

On the fix products page, the products can obliviously be fixed in order to exclude them from the optimization model.

When an optimization run is done, the user can view the model output. The model output is first displayed on product and customer group level. Thereafter, the user is able to click through in order to view the output on individual product and customer level. For each individual customer, a basket overview can be obtained. In the model output per SKU, two tables are
displayed. The well performing products, that should be in the assortment and the candidate products. Before removing the candidate products can be analyzed. On the candidate page, the relations with other products and the customers who bought the candidate product are displayed. The user can then proceed to the assumptions page, where he is able to make assumptions on customer and product retention, by specifying the critical products and critical customers for the candidate product.

After the candidate analyses is done, the final results can be obtained on the results page. Here, an overview is presented of which product should be in the assortment and which not. The user also has the option to export the results to an Excel document. In Appendix III a more comprehensive user manual is attached.

5.3 SOLVER

The solver that is used to solve the mathematical model in the heuristic is CBC 2.7.5. This is an open source solver which is provided in each AIMMS license. The mathematical model is diagnosed as a mixed integer program (MIP). The model was solved on a laptop with 4GB RAM and a Pentium Dual-Core processor of 2.10GHz. In Figure 11, an example progress window of AIMMS is displayed, where the 'total time' displays the time required for solving the problem in a particular run. The computation time of the model varied from just a few seconds to 15 seconds.

FIGURE 11: SOLVER PROGRESS WINDOW

5.4 CONCLUSION

In this chapter it is described why AIMMS is used for building the tool, which is the answer to research question 2.3a. Furthermore, the implementation of the heuristic into AIMMS is described. In the tool a user-friendly interface is built, which is an answer to research question 2.3b.
6. VERIFICATION AND VALIDATION

There exists a lot of literature about validation and verification. Sojda (2007) made a review about this field of study with the focus on decision support systems. “Verification is ensuring that the system is internally complete, coherent, and logical from a modeling and programming perspective”(Sojda, 2007). “Validation is examining whether the system is realistic and useful to the user or decision maker, and should answer the question: “Was the system successful at addressing its intended purpose” (Sojda, 2007)?

6.1 VERIFICATION

AIMMS provides a debugger and a math program inspector. The source debugger checks for misformulations in assignments and definitions of the parameters in the tool. The math program inspector checks for error in the formulation of the mathematical model. Next to the tools of AIMMS itself, for a small problem instance, a recalculation is made by hand.

6.2 VALIDATION

The validation concerns whether the tool achieved its’ purpose, helping the user make decisions about the assortment. Sojda (2007) stated that this can involve making better decisions, avoiding poor decisions or helping the user to make decisions with less time, data, information or knowledge.

The tool was presented both to the management team of Spruyt and to the actual user of the tool. They were satisfied with the results and confirmed that the end result definitely achieved its’ purpose. The demonstrations of the tool during the presentations showed that the tool is achieving its’ purpose. Some of the candidate products suggested by the tool, were indeed products they were considering to remove. In chapter 7, it will be elaborated more on how assortment decisions at Spruyt hillen are affected by the tool in terms of time needed for and the quality of the decisions.

6.3 SENSITIVITY ANALYSIS

“Sensitivity analysis can be a validation tool, especially for heuristic-based systems, and for systems where few or no test cases are available for comparison” (Sojda, 2007). Therefore, a sensitivity analysis is conducted in this section.

Sensitivity analysis is difficult to conduct on decision support systems, because of the complexity and the difficulty to isolate individual inputs (Sojda, 2007). Though, it is often possible to do a sensitivity analysis on the individual underlying models (Sojda, 2007). In this project, the heuristic contains a mathematical model on which a sensitivity analysis can be performed. The input parameters for the different performance constraints are analyzed below. It will be explained if the behavior of the model can be interpreted.
**Minimal absolute profit per product (MAPIP)**

The first performance constraint that is considered is the minimal absolute profit per individual product (MAPIP). In Table 8, the results are displayed for two product groups Raw materials and Compounding. The first column represents the value for the MAPIP performance constraint. The second column represents the number of products in the assortment. In the fourth, sixth and eighth column, the revenue, profit in euro's and profit in percentage are displayed respectively. The MAPIP is varied from 0 to 51,200 in steps of 100%.

<table>
<thead>
<tr>
<th>MAPIP (€)</th>
<th>#SKU</th>
<th>Δ</th>
<th>Revenue (€)</th>
<th>Δ</th>
<th>Profit (€)</th>
<th>Δ</th>
<th>Profit (%)</th>
<th>Δ</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The higher the value of the input parameter MAPIP, the more sensitive the output to this parameter becomes. The number of SKU's in the assortment declines quite drastic between 200 and 12,800. Whereas, the revenue, absolute profit and relative profit start to decline really quickly after 6,400. This can be explained by the way the assortment is built up. In section 1.8 of Appendix III, a cumulative profit graph is displayed. This shows that about 20% of the product is responsible for 80% of the profit. Therefore, the revenue and profit measure start to decline very quickly only after 80% of the products is already removed. A minimal absolute profit per product of 51,200 is very high, and thus it is very unlikely that a user will specify such a high value. However, by doing this analysis, it is shown that the model behaves as can be expected for this performance constraint.

**Minimal relative profit per product group (MRPPG_g)**

The next performance constraint to be analyzed is the minimal relative profit per product group (MRPPG_g). This is done for the product group Raw materials. The MRPPG_g is varied from 60.5 to 62.5, see also Table 9. The third column represent the number of products in the group. The fifth, seventh and ninth column display the revenues, profit in euro’s and profit as a percentage.
When raising the MRPPGₙ, the number of products, revenues and absolute profit decline. This is logical, since tightening the minimal relative profit results in leaving products out the assortment with a lower relative profit than the threshold. However, from 62.0% to 62.5% the number of products is increasing. This can be explained as follows. Suppose there is a product A, with a low relative profit, e.g. 30% but a high absolute profit. Since, the model maximizes over absolute profit, the model will try to keep this product in the assortment. However, at a certain level, the products with a high relative profit cannot compensate for this product A anymore. Product A will then be left out the assortment, which causes the relative profit margin of the total assortment to rise way above the minimal relative profit margin. Then, there is room to again include less profitable products who have been left out before.

**Minimal relative profit total assortment (MRPTA)**

The third and last performance constraint is the minimal relative profit on the total assortment (MRPTA). The MRPTA is varied from 52.5 to 54.5, see also Table 10.

This constraint behaves similar as the minimal relative profit per product group. In essence, it is also the same constraint, but it applies to all products instead of a subgroup of products. It should be noted that there is a high number of products which are unprofitable or products which have not been last year. It also illustrates that with a reduction in the number of products...
of almost 38% a higher absolute profit can be obtained potentially. This indicates that there are a lot of products which will be candidate for removal and thus really illustrates the need of the tool.

6.4 CONCLUSION

In this chapter the decision support tool has been verified and validated, which corresponds with research question 2.3c. Main research question 2, ‘How to implement the relevant aspects into a decision support tool?’, is now completely answered. In chapter 4, the design of the heuristic was described, in chapter 5 the implementation of the heuristic into software and finally the verification and validation were done in this chapter. In the next chapter, the results of the tool will be evaluated.
7. RESULTS

Now that the decision support tool has been developed and validated, the performance of the tool should be evaluated (Main research question 3) in order to be able to decide whether implementation is desired. In chapter 2.1, the current method of assortment planning was described. The question is how Spruyt hillen will perform if they would use the developed tool in comparison with the current situation.

The tool can be beneficial in two ways. First, the tool can lead to better assortment decisions, by which is meant that the right products are kept in the assortment or removed from the assortment. Secondly, the tool can potentially save time for employees when they have to make decisions about the assortment, since all information is within easy reach. As Benjamin Franklin once stated "Time is money".

7.1 IMPROVEMENT IN DECISIONS

First, the fact that Spruyt hillen will have the assortment tool will make the barrier lower to actually be more concerned with assortment planning. In the current situation, assortment planning is a more time-consuming task, which causes that they reflect less often on their assortment than desired. The tool will motivate Spruyt hillen to be more concerned with assortment decisions. The expectation is therefore that the assortment is better ‘maintained’, by which is meant that products which are not performing well are removed earlier from the assortment. This saves money, since with carrying a product in the assortment are costs associated, e.g. inventory costs.

In section 2.1, it was mentioned that Spruyt hillen is quite reluctant to remove products from the assortment. Arguments of related products and customers to leave were used in order to keep un- or less profitable products in the assortment, but this was not supported by data. In the developed tool it is possible to see all related products and customers who buy the product within one mouse-click. Therefore, products are not considered isolated from its’ customers and related products. In some cases, the thoughts of the user that some products are related, may indeed be confirmed by the data. However, it could also be very much possible that these products are not related and that the argument for keeping a product in the assortment is not valid. The expectation is therefore that bad performing products will not be kept in the assortment for the wrong reasons.

The transparency in relation to customers will also prevent that a product will be removed without informing the customer. In section 2.1, it was mentioned that it happened that a product, which was important for a very important customer, was removed without informing that customer, resulting in a dissatisfied customer. The probability that this will happen will be much less when using the tool.

The tool also gives transparency in the performance of products, customers, products groups, customer groups and the total assortment. Furthermore, the tool gives the user control to influence the performance of the assortment by varying the values for the performance measures on individual products, product groups and for the total assortment. Spruyt hillen has target levels for some of the performance measures, for which the tool can help them in achieving these targets.
7.2 DECISION TIME REDUCTION

As discussed in chapter 2.1, Spruyt hillen reflects less often than desired on their assortment, due to the time consuming procedures of making reports on sales of products. The expectation of Spruyt hillen is that they will not spend necessarily less time on assortment decisions, but spend time more effectively. The tool gives them the opportunity to be more concerned with assortment planning where previously the barriers were often too high to be concerned with these kind of decisions. So in an absolute sense, the tool will not save time for Spruyt hillen. However, when spending time on assortment decisions, this is done much more efficient.

7.3 OTHER IMPLICATIONS

When Spruyt hillen wants to add a new product to the assortment, a cannibalization effect and an extra sales effect is considered. The tool quantifies these effects, which gives an indication how the new product will perform in relation to the existing assortment. The decision for adding a new product can now be made more deliberately.

The quantification of the relations between products can also be used for other purposes. Spruyt hillen stated that they are thinking about using the information about product relations for the web shop. E.g. if it is known that a certain percentage of customers who buy product A also buy product B, a suggestion on product B can be made when customers are looking on the page of product A, under the heading: “Customers also bought:”. This could increase sales for some products.

The tool heavily relies on the data as available in the ERP system. Spruyt hillen indicated that therefore the tool is an extra trigger for the company to further improve their data integrity. This is an implication which stands apart from the design of the tool itself.

Now that the performance of the tool has been evaluated, main research question 3 is answered. The next chapter will describe the implementation process in the company.
8. IMPLEMENTATION IN THE COMPANY

In this chapter the process of implementation in the company will be described. A distinction is made between the implementation at Spruyt Hillen and the implementation for the Fagron Group as a whole.

8.1 SPRUYT HILLEN

After having finished the development of the tool, the result was presented to the General Manager, the Marketing Manager, the Controller and the Logistics Manager of Spruyt Hillen. They reacted positively and believed very much in the potential of the tool. It was therefore decided that the tool should be actually implemented. They evaluated that the benefits of the tool outweigh the costs associated with the purchase of AIMMS.

With this decision, an implementation process was started. First, a presentation of the tool to the actual user was given. The user of the tool also saw the potential of the tool, after which the decision on implementation became definite. He was confident that the tool will help him in making decisions on the assortment. It will also lower the barrier to actually do an assortment analysis, which was previously a very time-consuming task.

The next thing was to plan a training session for the user of the tool. In Appendix VI the training exercises can be found. After the training was completed, the user had to sign a training certificate as proof that he masters the usage of the tool. This certificate can also be found in Appendix VI.

In order to make the system available to the user, the software should be purchased and the system should be set up. This process is still going on at this moment of writing. The final handover of the tool will also follow up shortly after the end of this project.

FIGURE 1: IMPLEMENTATION STEPS SPRUYT HILLEN
8.2 FAGRON GROUP

For the roll out to different Fagron Group companies, the tool had to be adapted in some ways. First, the layout of the tool had to be changed from Spruyt hillen to Fagron. An example of the ‘Fagronized’ tool is displayed in Appendix VII. Second, some changes in the database assignments had to be made in order to be able to load the correct data into the tool.

After the tool was prepared for usage at other Fagron Group companies, the tool had to be presented to the management. First the tool was presented to the CIO of the Fagron Group and the Area General Manager of Fagron responsible for the areas Belgium, Poland, Germany, France and Scandinavia. Although, the reaction was positive, it was decided not to implement it immediately. The tool depends heavily on the data from the ERP system and they evaluated that they were not ready for it at this moment. However, in the nearby future they are willing to use it.

Later on, the tool was also presented to the management of Fagron Netherlands, who responded positively as well. However, since they are already quite overloaded with tools of which some are not even used, they first want to await the experiences at Spruyt hillen.

FIGURE 2: IMPLEMENTATION STEPS FAGRON GROUP
9. CONCLUSIONS AND RECOMMENDATIONS

This is the final chapter of this report. Here, the conclusion of the project, recommendations for the company and the contribution to science will be described.

9.1 CONCLUSIONS

The main assignment of this tool was to develop a decision support for the assortment planning at Spruyt hillen. Three main research questions were formulated in order to complete this project.

Main research question 1:

- What are the relevant aspects with respect to the assortment decisions of Spruyt hillen?

Main research question 2:

- How to implement the relevant aspects into a decision support tool for the assortment planning at Spruyt hillen?

Main research question 3:

- What is the performance of the developed decision support tool?

In the first phase of the project, assortment planning at Spruyt hillen and the literature with relation to assortment planning were considered. It was concluded that the phenomenon of basket shopping was an important aspect for Spruyt hillen. Furthermore, the desire on influencing the profitability of the assortment was important for Spruyt hillen.

After having revealed the relevant aspects, the next step was to implement these aspects into a decision support tool. It turned out that a mathematical model that included basket shopping at once was unfeasible for this project. A heuristic was designed, which first uses a product-based model to distinguish between well-performing products and candidate products to be removed. Before removal, the customers and related products of the candidate products are analyzed. Depending on the impact of the removal on the related products and customers, it is decided if the candidate product may be removed or not. The heuristic was implemented into a decision support tool in the software package AIMMS. Furthermore, a GUI was built for the tool.

In chapter 7, the results of the tool were presented. The main results are listed below:

- The tool will make the barrier lower to actually be concerned with assortment planning. This will motivate Spruyt hillen to be more concerned with assortment decisions. The expectation is therefore that the assortment is better 'maintained'. This means that products which are not performing well are removed earlier from the assortment, resulting in cost savings.
- At Spruyt hillen, sometimes arguments of related products or customers are used, to keep less or unprofitable products in the assortment, without support of data. The tool provides these data about the interaction of products with customers and other
products. The expectation is that products will therefore not be kept in the assortment for the wrong reasons.

- On the other hand, the transparency in relation to customers will also prevent that some products will be removed without informing a customer who specifically needs this product.
- The tool gives transparency in the performance on three levels, individual products/customers, product/customers groups and the total assortment. Furthermore, the tool gives the user control to influence the performance of the assortment.
- The expectation of Spruyt hillen is that they will not spend necessarily less time on assortment decisions, but spend time more effectively. They will be more concerned with assortment planning, resulting in a better maintained assortment, where products are removed when they do not generate enough money anymore. This results in costs savings.
- Insight in the cannibalization and extra sales associated with the introduction of a new product is provided. The introduction of a new product will be a more deliberate decision.
- The quantification of the relations between products can also be used for product suggestions in the web shop, which may increase sales.
- The tool is an extra trigger for Spruyt hillen to further improve their data integrity.

The tool was well received by both Spruyt hillen and the Fagron Group. Therefore, it was decided that the tool will be implemented at Spruyt hillen. To facilitate this implementation, a user manual was written and a training to the user was provided. Furthermore, the preparation for the roll out to other Fagron Group companies was done as well. The implementation at Fagron Netherlands was not done immediately. Though, it is nice to mention that a week after the presentation to Fagron NL, it was instructed that a lot of products which were not sold in the last year should be removed from the assortment. Apart from the power of the tool itself, the fact that there is a project concerning assortment optimization creates awareness for the product assortment throughout the organization.

9.2 RECOMMENDATIONS FOR THE COMPANY

The tool relies on the data as available in the ERP system. If there are errors in the data this will also influence the results of the tool. As a result, it is very important that the data in the ERP system is correct. Therefore, it is strongly advised not to focus blindly on the results of the tool. If in the data a decimal point would be wrong, it can be that a well performing is advised to be removed from the product. However, the user of the tool is supposed to have knowledge and experience with the products. So if the user uses his common sense, he should detect these possible errors in data.

The cost price at Spruyt hillen does not include all indirect costs like inventory and handling costs. If Spruyt hillen will also use a more comprehensive cost price calculation, the results of the tool are even more valuable, since then also the ‘hidden’ costs are taken into account.
9.3 CONTRIBUTION TO SCIENCE AND FUTURE RESEARCH

A lot of research is done within the field of assortment planning. However, the phenomenon of basket shopping is not often studied in combination with assortment planning. Brijs et al. (2000) and Flapper et al. (2010) also did a case study on assortment planning in combination with basket shopping. However, Brijs et al. (2004) ignored the retention of customers. Furthermore, Flapper et al. (2010) assumed that a customer is immediately lost when demand is not satisfied, which does not hold in this project.

The heuristic developed in this project takes into account both relations with other products and the retention of customers. The user is able to specify if there is indeed an effect for the customers or related products when a product is removed from the assortment. Thereafter, the tool suggests which products should be kept in and which products should be removed from the assortment. The heuristic and the case study are both contributions to science.

I experienced that people are reluctant to remove products from the assortment, fearing consequences for customers or loss in demand for other products. Insights in the behavior of customers are therefore very important. However this is information which is often not available. It would be very interesting if anyone can invent widely applicable methods to extract information on customer retention and impacts on related products from sales data. In my opinion that could cause a ‘revolution’ in assortment planning.
REFERENCES


LIST OF TABLES AND FIGURES

Tables
Table 1: Product groups ......................................................... 20
Table 2: Customer groups ..................................................... 21
Table 3: Basket size per transaction per customer group ................................................ 23
Table 4: Basket size per customer per year per customer group ........................................ 24
Table 5: Literature overview .................................................................................. 27
Table 6: Buying behavior over two years ..................................................................... 33
Table 7: Key performance indicators ........................................................................... 35
Table 8: Sensitivity analysis (MAPIP) .......................................................................... 50
Table 9: Sensitivity analysis (MRPPGg) ...................................................................... 51
Table 10: Sensitivity analysis (MRPTA) ..................................................................... 51
Table 11: Buying behavior over two years ..................................................................... 69
Table 12: Example 1 ................................................................................................. 77
Table 13: Example 1 (II) .......................................................................................... 78
Table 14: Example 2 .................................................................................................. 78
Figures

Figure 1: Heuristic approach overview ................................................................. 10
Figure 2: Decision diagram for a candidate product ........................................... 11
Figure 3: Research design .................................................................................... 18
Figure 4: Heuristic approach overview ............................................................... 31
Figure 5: Candidate analysis ................................................................................ 36
Figure 6: Relations of a Candidate product ....................................................... 37
Figure 7: Decision diagram for a candidate product ........................................... 41
Figure 8: Flow diagram of adding a product ..................................................... 44
Figure 9: Schematic overview of the tool ............................................................ 46
Figure 10: Tool dashboard page ......................................................................... 47
Figure 11: Solver progress Window ..................................................................... 48
Figure 12: Implementation steps Spruyt hille .................................................................. 55
Figure 13: Implementation steps Fagron Group ................................................. 56
Figure 14: Tool Dashboard ................................................................................. 73
Figure 15: Loaded data ......................................................................................... 75
Figure 16: Product relations ............................................................................... 75
Figure 17: Basket overview .............................................................................. 76
Figure 18: Cumulative revenue and margin graph ............................................ 76
Figure 19: Cumulative margin (%) graph ............................................................ 77
Figure 20: Performance constraints ................................................................. 79
Figure 21: Output on product group level ......................................................... 80
Figure 22: Output on individual product level .................................................. 80
Figure 23: Output on customer group level ....................................................... 81
Figure 24: Output on individual customer level ................................................. 81
Figure 25: Basket overview .............................................................................. 82
Figure 26: Analyzing candidate product on customers .................................... 82
Figure 27: Analyzing candidate product in relation to other products .......... 83
Figure 28: Assumptions page .......................................................................... 83
Figure 29: Table of new products ..................................................................... 84
Figure 30: Table of candidate products ............................................................ 85
Figure 31: Final impact on assortment level ..................................................... 85
Figure 32: Final impact assortment level graphs .............................................. 86
LIST OF DEFINITIONS

**Association rule**: Relations between products based on historical sales data.

**Assortment**: Set of products that are available for sale at a certain moment in time.

**Base unit of measure**: The standard unit of measure of product.

**Basket [customer based]**: A customer based basket is defined as the set of products that a customer buys during the period under consideration.

**Basket [transaction based]**: A transaction based basket is defined as the set of products that occurs in a certain transaction.

**Basket shopping**: The phenomenon that customers buy more products than one at a shop.

**Candidate product**: A product which does not satisfy the performance constraints and is marked as a candidate.

**Cannibalization**: The reduction in sales of other products, when introducing a new product.

**Confidence**: A percentage to which extent one product implicates another product.

**Cost price**: The cost price of a product as displayed in the ERP system of Spruyt hillen.

**Critical customer**: A customer of a candidate product which is lost when the candidate product is removed.

**Critical product**: A related product to a candidate product for which the demand is lost when the candidate product is removed.

**Customer group**: Existing classification of customers, as shown in Table 2.

**ERP system**: Enterprise Resource Planning system. Integrates all information from a company. Fagron and Spruyt hillen use Microsoft Dynamics Navision.

**Fixed product**: A product, selected by the user which is excluded from the optimization model. The tool will advise to keep the product in the assortment anyhow.

**Pharmaceutical compounding**: The creation of medicines using appropriate ingredients and tools in order to satisfy the unique needs of a client.

**Product group**: Existing classification of products, as shown in Table 1.

**Sales price**: The price a customer pays to the company for a particular product. Different customers can have different sale prices for the same product.

**Support**: The support of a product or set of products is the frequency of that set in the set of all baskets.

**Unit of measure**: The units in which a product is measured. E.g.: grams or kilograms.
LIST OF ABBREVIATIONS

**AIMMS**: Advanced Interactive Multidimensional Modeling System. The name of the software in which the tool is developed.

**B2B**: Business to business.

**B2C**: Business to customer.

**BUM**: Base unit of measure.

**ERP system**: Enterprise Resource Planning system.

**GUI**: Graphical User Interface

**KPI**: Key performance indicator.

**MINLP**: Mixed integer non linear problem.

**SKU**: Stock keeping unit.
APPENDIX I: LITERATURE STUDY

1.1 INTRODUCTION

The assortment is defined by the set of products carried in a store at a certain moment (Kök et al., 2009). "The goal of assortment planning is to specify an assortment that maximizes sales or gross margin subject to various constraints, such as a limited budget for purchases of products, limited shelf space for displaying products, and a variety of miscellaneous constraints such as a desire to have at least two vendors for each type of product" (Kök et al., 2009).

Product assortment planning seeks to find the right balance among variety, depth and service level (Mantrala et al., 2009). Variety means the number of product categories, depth means the number of stock keeping units (SKUs) within a category and service level implies the amount of a specific SKU to keep in stock. Where a category is defined a group of SKUs with similar attributes (Kök et al., 2009). If a company offers a higher service level, the variety and/or depth will probably be less. A retailer tries to find the right trade-off in order to satisfy the customer as good as possible and thus get the highest profit.

The assortment has a big influence on both sales and gross margin and therefore the decision is of high priority (Kök et al., 2009). Consequently, a lot of academic research is quite recently done about assortment planning. Nonetheless, there is no dominant solution yet available.

1.2 BASKET SHOPPING

A lot of academic research about assortment planning focuses on a single product category. However customers often buy more products at once, a bundle or basket of products. This phenomenon is called basket shopping. Reasons for this behavior are time-saving convenience and fixed costs associated with a store visit (Messinger et al., 1997). Customers make decisions about multiple categories when they shop (Manchanda et al., 1999). Products can be in the same basket because they are complementary to each other, for example pancake mix and milk or because of similar purchase cycles, for example beer and bread (Manchanda et al., 1999). This first reason is something which give retailers some control over the buying behavior of customers, whereas the second reason is coincidentally and less controllable. Furthermore Bell and Latin (1998) found that store choice is based on the total basket utility. So if a customer wants a specific product and for some reason this product is not carried by the retailer, it is possible that the customer buys the entire basket somewhere else. Borle et al. (2005) showed that a reduction of the assortment leads to a decrease in shopping frequency and purchase quantity. This reduced shopping frequency also leads to a reduction in overall store sales. Cachon and Kök (2007) found that a decentralized assortment planning, where categories are managed independently from each other lead to lower profits than optimal. This is because this decentralized planning ignores cross-category interactions.

Agrawal and Smith (2003) developed an assortment planning model that is based on a specified structure for sets and substitutes. In this model, Agrawal and Smith (2003) considered multiple categories and the phenomenon that customers buy in sets. Each customer arrives with a demand for a set of products. If the set of products is not available, there are three options.
Substitute with a smaller set, substitute with a different set or not buying anything. So this model requires the knowledge of all subsets of products to which a customer can substitute.

Flapper et al. (2010) presented two mathematical models for two different strategies, a product based assortment and a customer based assortment. The customer based assortment accounts for the effects of basket shopping. They assumed that customers with unsatisfied demand are lost, because of the one-stop-shopping requirement. It turned out that the customer-based model is more profitable for the case done. The models of Flapper et al. (2010) are deterministic multi-period models and include inventory costs, setup times, setup costs and lost sales. There is no storage limitation, but there is a limited production capacity.

Another approach of studying assortment planning in combination with basket shopping is the use of data mining. With data mining it is possible to reveal relations between products based on historical sales data, also called association rules. Agrawal R. et al. (1993) introduced the discovery of association rules. An association rule can be a statement like: ‘70% of the people who bought Ointment X also bought Ointment Y’. In this statement 70% stands for the confidence of the rule. In this case Ointment X is called the antecedent and Ointment Y the consequent. Next to the confidence, they also introduce support as an important definition. The support of the rule is the frequency of \( X \cup Y \) in the set of transactions, which can be either expressed in absolute form or as a percentage.

Brijs et al. (2000) developed an assortment model, which makes use of this data mining. The model is called the PROFSET model and uses the cross-sales potential of products instead of the individual product contribution. The optimization model determines which sets of products should be in the assortment. Even though the products are not considered in isolation, but in relation to other products, the relation with customers is completely ignored. In a retail environment this will often do no harm, since individual customer contribution is relatively low. However, in a business to business (B2B) environment there can be customers with an enormous contribution. It is very undesirable to remove a low contributing product which causes a huge customer to leave.

### 1.3 Substitution and Cannibalization

Another aspect which is often studied in combination with product assortment planning is substitution. In the literature substitution is classified into three groups (Kök and Fisher, 2007). The first is stock-out based substitution. In this case a customer goes to a shop repetitively for buying a daily consumable. When he or she finds the product to be sold out one day, he or she buys another product. The second group of substitution is assortment-based substitution. In this case, a customer has a preference for a certain product. At one day he or she visits a store where this product is not carried in the assortment and he or she will switch to another product. The third and last group of substitution can be described as follows. A customer goes to a store and chooses the favorite product from what is carried on the shelf. She buys it when the utility is higher than her no-purchase option. It could be the case that she would have preferred another product, but the retailer did not have the product in the assortment or the product was stocked out. This third group can be an example of both substitution types depending on the situation. The first two substitution types are very common for repeated purchases whereas the third group fits to one-time purchases.
From a pharmaceutical perspective, substitution can be classified into *generic substitution* and *therapeutic substitution* (AlAmeri et al., 2010). Generic substitution is the substitution to a product with exactly the same compound but from a different manufacturer. Generic products arise when the patent expires from the manufacturer that initially invented the product. Therapeutic substitution is the switch to another medicine with another active substance (AlAmeri et al., 2010). Where generic substitution is already a bit controversial, this holds even more for therapeutic substitution (AlAmeri et al., 2010). This is due to the fact that substitution can affect clinical outcomes. In the Netherlands, generic substitution is already quite common, due to the preference policy of health insurance companies (Van de Steeg-van Gompel et al., 2011). This preference policy means that some insurance companies only reimburse a specific - often generic - product. Therapeutic substitution is still in its infancy. At this moment, only a doctor can prescribe another medication. However, in the future pharmacies and insurances companies might have an influence on this (Van de Steeg-van Gompel et al., 2011).

Whereas substitution can occur between existing products in the assortment, the term *cannibalization* is used to refer to the reduction in sales of other products, when introducing a new product. According to Cachon et al. (2005) a retailer has to make a trade-off when adding a new product (variant) to the assortment between the benefit of extra sales and the costs of including the new product. This costs can also include a cannibalization effect, which results in a reduction in demand for other products and thus lowering their profit and operational efficiency (Cachon et al., 2005).
APPENDIX II: MATHEMATICAL MODEL INCLUDING BASKET SHOPPING

Assumptions

1) The length of the period is a function of a year.
2) The demand for the coming year equals demand during last year.
3) Substitution does not occur between products within the assortment.
4) There are no storage space restrictions.
5) The demand can always be directly fulfilled.

Argumentation

1) For some products the demand faces seasonality with cycles of a year. An example is hay fever medication which is sold more often during Springtime. Therefore, the period should be as function of a year, otherwise the tool could give a biased output. E.g. if data of half a year is used it could suggest to remove a product which is only sold during a certain period in a year. The period is a variable in the tool and can be changed by the user. The default period is one year.
2) In the ideal situation, forecasts would be used to determine the demand. Spruyt hillen carries approximately 4500 products but forecasts are only available for a small proportion of these products. Furthermore, it would be difficult to even use these forecasts, since the model uses demand per product per customer per transaction, whereas the forecast of the demand is based on the total demand per product. In consultation with the company it is therefore decided that last year's sales are used.

An analysis is done for the two most important customer groups to what extent customers tend to buy the same products over the years. The sales data from 2012 and 2013 are used. Table 6 shows that on average 60.95% of the products bought in 2012 are also bought in 2013 by Hospitals. For the Pharmacists this percentage is 64.64%.

<table>
<thead>
<tr>
<th>Customer group</th>
<th>Percentage of products also bought in 2nd year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hospitals</td>
<td>60.95%</td>
</tr>
<tr>
<td>Pharmacists</td>
<td>64.64%</td>
</tr>
</tbody>
</table>

In this analysis all products are taken into account. This also contains for example spot sales and promotional products. It can happen that these products are not available in the second year. Therefore, the percentage is not as high as one would expect, but customers still tend to buy the same products over the years. So, it can be concluded that assumption 2 is acceptable.

3) Spruyt hillen offers only one brand per product. In a few cases, there are some size variants. Therefore, it is decided that substitution is out of scope, see also 3.2.1.
4) Spruyt hillen has no storage space restrictions. There are no problems with capacity in their warehouse and also for the coming years they expect that this will not be an issue.
5) No inventory model is included in the model. The interest in this project is especially in the strategic decision which products to carry. The operational decision of how many items to stock of each SKU is out of scope as discussed in 3.2.2. It is therefore assumed that demand can always be directly fulfilled.

**List of variables**

- \( c, c \in \{1 \ldots C\} \): Index denoting the customers.
- \( c_p \): Cost price of product \( p \).
- \( d_{tcp} \): Demand of product \( p \) for customer \( c \) in transaction \( t \).
- \( dl_{tcp} \): Demand for product \( p \) by customer \( c \) in transaction \( t \) lost through basket shopping.
- \( f^1_{p^*p} \): Binary variable, indicating whether the demand for product \( p \) for all customers is lost (1) when \( p^* \) is removed from the assortment, or not (0).
- \( f^2_{pc} \): Binary variable, indicating whether customer \( c \) is lost (1) when \( p \) is removed from the assortment, or not (0).
- \( g, g \in \{1 \ldots G\} \): Index of product groups under consideration.
- \( MAPIP \): Minimal absolute profit margin per product over the period.
- \( MAPP_{Pg} \): Minimal absolute profit margin per product group \( g \) over the period.
- \( MAPTA \): Minimal absolute profit margin on the total assortment over the period.
- \( MRIP \): Minimal relative profit margin per product over the period.
- \( MRPP_{Pg} \): Minimal relative profit margin per product group \( g \) over the period.
- \( MRPTA \): Minimal relative profit margin on the total assortment over the period.
- \( p, p \in \{1 \ldots P\} \): Index denoting the products.
- \( TC \): Total cost of goods sold over the period.
- \( TP \): Total profit over the period.
- \( TR \): Total revenues over the period.
- \( u_{tcp} \): Adjustment factor for the unit of measure.
- \( w_{tcp} \): Sales price of product \( p \) to customer \( c \) in transaction \( t \).
- \( X_p \in \{0,1\} \): Decision variable which indicates if a product is in the assortment (1) or not (0).
Objective function

\[
\max TP = TR - TC
\]  \hspace{0.5cm} (3.1) Absolute profit over the period over the total assortment.

Where

\[
TR = \sum_{t=1}^{\tau} \sum_{p=1}^{P} \sum_{c=1}^{C} w_{tcp} \cdot (d_{tcp} - d_{l_{tcp}}) \cdot X_p
\]  \hspace{0.5cm} (3.2) Total revenues over the period over the total assortment.

\[
TC = \sum_{t=1}^{\tau} \sum_{p=1}^{P} c_p \sum_{c=1}^{C} u_{tcp} \cdot (d_{tcp} - d_{l_{tcp}}) \cdot X_p
\]  \hspace{0.5cm} (3.3) Costs of goods sold over the period over the total assortment.

Constraints

Next to the objective function, constraints should be formulated. The demand lost because of basket shopping is given in formula (3.4). The decision whether a product is in the assortment is simply a yes-no question and therefore this decision variable should be a binary variable.

For all \((c, p, p'), c \in \{1, \ldots, C\}, \ p \in \{1, \ldots, P\}, p' \in \{1, \ldots, P\}\)

\[
dl_{tcp} = d_{tcp} \cdot \max_{p' \neq p} \left(1 - X_{p'}\right) \cdot f_{p'pc}
\]  \hspace{0.5cm} (3.4) Demand lost because of basket shopping for customer \(c\) for product \(p\).

\[
\forall p, p \in \{1, \ldots, P\} \quad X_p \in \{0,1\}
\]  \hspace{0.5cm} (3.5) Binary bound

Except from the constraints mentioned above, there are some additional constraints. Spruyt hillen want to have influence on the profitability of the products which are in the assortment. Therefore, some performance constraints are formulated on the profit. In the KPI section, six measures on profit were presented, which results in the following constraints.

\[
TR - TC \geq MAPTA
\]  \hspace{0.5cm} (3.6) Absolute profit margin restriction total assortment.

\[
\frac{TR - TC}{TR} \geq MRPTA
\]  \hspace{0.5cm} (3.7) Relative profit margin restriction total assortment.

For all \(g, g \in \{1, \ldots, G\}\)

\[
\sum_{p \in g} \sum_{t=1}^{\tau} \sum_{c=1}^{C} w_{tcp} \cdot (d_{tcp} - d_{l_{tcp}}) \cdot X_p - \sum_{p \in g} \sum_{t=1}^{\tau} \sum_{c=1}^{C} c_p \cdot (d_{tcp} - d_{l_{tcp}}) \cdot u_{tcp} \cdot X_p \geq MAPPG_g
\]  \hspace{0.5cm} (3.8) Absolute profit margin restriction product group \(g\).

\[
\forall g, g \in \{1, \ldots, G\} \quad \frac{\sum_{p \in g} \sum_{t=1}^{\tau} \sum_{c=1}^{C} w_{tcp} (d_{tcp} - d_{l_{tcp}}) X_p - \sum_{p \in g} \sum_{t=1}^{\tau} \sum_{c=1}^{C} c_p (d_{tcp} - d_{l_{tcp}}) u_{tcp} X_p}{\sum_{p \in g} \sum_{t=1}^{\tau} \sum_{c=1}^{C} w_{tcp} (d_{tcp} - d_{l_{tcp}}) X_p} \geq MRPPG_g
\]  \hspace{0.5cm} (3.9) Relative profit margin restriction product group \(g\).
\[ \forall p, p \in \{1, ..., P\} \]

\[ \sum_{t=1}^{T} \sum_{c=1}^{C} w_{tcp} \cdot (d_{tcp} - d_{l_{tcp}}) \cdot X_p - \sum_{t=1}^{T} \sum_{c=1}^{C} c_p \cdot (d_{tcp} - d_{l_{tcp}}) \cdot u_{tcp} \cdot X_p \geq MAPIP \cdot X_p \]  

(3.10) Absolute profit margin restriction per product \( p \).

\[ \forall p, p \in \{1, ..., P\} \]

\[ \frac{\sum_{t=1}^{T} \sum_{c=1}^{C} w_{tcp} \cdot (d_{tcp} - d_{l_{tcp}}) \cdot X_p - \sum_{t=1}^{T} \sum_{c=1}^{C} c_p \cdot (d_{tcp} - d_{l_{tcp}}) \cdot u_{tcp} \cdot X_p}{\sum_{t=1}^{T} \sum_{c=1}^{C} w_{tcp} \cdot (d_{tcp} - d_{l_{tcp}}) \cdot X_p} \geq MRPIP \cdot X_p \]  

(3.11) Relative profit margin restriction per product \( p \).
APPENDIX III: USER MANUAL

1. HOW TO USE THE TOOL

1.1 TOOL DASHBOARD

In Figure 14, the tool dashboard is displayed. This is the starting screen, when the tool is opened. On the left side the navigation menu is shown, from there it possible to get to the various tool pages.

![Figure 14: Tool Dashboard](image)

1.2 LOADING DATA

**Data via ODS Database**

It is possible to directly load the data from the ODS Database. To make a connection with the database press the ‘Check Connection’ button. Fill in the required login to make a connection. When you are connected with the database, you can press ‘Load data from DB’ in order to load the data into the tool. Please be patient, this can take up to 10 minutes.

By pressing ‘Clear Data’, all data is removed. This can be used if another dataset is required. **Warning**: this cannot be undone. To check if the data is indeed loaded into the model, see if the product groups are listed on the bottom left.

You can also vary the period of the sales history that will be loaded into the model. The period is by default one year. By using the drop down menu another period can be chosen.

1.3 SAVING AND LOADING CASES

The loading of data via the ODS database connection can require some time. It is therefore advised to save the loaded data into a so-called ‘Case’. If you have saved the data in a case, it only takes a few seconds to load data. This option can be useful if you want to do multiple analyses in
a short period of time. Then you do not need to load the data each time out of the ODS database, which can cost quite some time. When there is a longer period between the analyses it is advised to reload the data from the ODS database in order to get the most recent data.

This functionality is an option provided by the software AIMMS itself. It can be reached via the menu bar of the software itself on top of the screen.

Saving a case: ‘Data’ -> ‘Save Case as…’

After which you can give the case a name and press ‘Save’.

Loading a case: ‘Data’ -> ‘Load Case’ -> ‘into Active…’

When loading a case, just click the name of the case you want to load in your model and press ‘Load’.

1.4 SELECT PRODUCT GROUPS

After having loaded the data into the tool, you can choose which product groups to include in your decision making process. To select multiple groups, hold the SHIFT or CTRL button. **Important:** Press the button ‘CONFIRM’ in order to confirm your selection. It can take up to about twenty seconds for the tool to load the selection.

1.5 USER INPUTS

**Fix Products**

The ‘Fix Products’ option gives the option to select products which are excluded from the decision-making process and should be in the assortment anyway. First select a product group from which you want to fix a product, then select the actual product by selecting the checkbox.

**Remark:** The more products are fixed, the less power the tool has. When there is no justified reason to fix a product, do not use this option.

**Add New Product**

In order to add a new product to the dataset, press the ‘Add New Product’ button. Some pop-up’s will arise with the request to give information about the new product. The following information is required: product number, product description, product group, cost price, sales price, and forecasted yearly demand. It is also asked if demand for existing products will be cannibalized and if extra sales are generated by introducing the new product. In this case, you should select the products which be cannibalized and/or the products which will generate extra sales. When this option is chosen, it is also requires a cannibalization percentage and/or a percentage of extra sales.

**Initialize Parameters**

By pressing this button, the tool redirects to the ‘Parameters’ page. See section 1.9 for more information.
1.6 OPTIMIZE

When all options and parameter values are set correct, it is possible to start optimizing by pressing that button. The optimization will take approximately 20 seconds. When the calculation is finished, the output can be seen by clicking the ‘View Output’ button. It is also displayed how many optimization runs there are ran so far.

1.7 VIEW DATA

When the data is loaded from either the Excel file or the ODS database, it is possible to see an overview of the loaded data by pressing ‘Data’ in the navigation menu, see Figure 15. There are two tabbed pages, products and customers. The products page gives an overview of all products loaded into the model with their product number, description, base unit of measure, class, group and cost price.

![FIGURE 15: LOADED DATA](image)

By typing keywords, in the keyword box it is possible to search for particular products. Furthermore, it is possible to select a product and view their relations with other products, see Figure 16. In the left table, the relations based on a customer basket per period are given. In the right table, the product relations based on transactions are displayed.

![FIGURE 16: PRODUCT RELATIONS](image)
Formulas (1.7.1 - 1.7.8)

On the customers page, the customers are displayed with their customer number, description, city and customer group. A customer can be selected in order to view the basket of this customer which is an overview of the products he ordered the past year.

![Customer basket overview](image)

**FIGURE 17: BASKET OVERVIEW**

Formulas (1.7.9 - 1.7.10)

1.8 CHARTS

The chart gives an indication about how the assortment is built up. The first graph gives the cumulative absolute margin (€) and cumulative revenue (€). The products on the x-axis are sorted based on absolute profit margin (€).

![Revenue and margin chart](image)

**FIGURE 18: CUMULATIVE REVENUE AND MARGIN GRAPH**
The second graph displays the minimal profit margin as a percentage, the products on the x-axis are again sorted on highest absolute margin.

![FIGURE 19: CUMULATIVE MARGIN (%) GRAPH](image)

Formulas (1.8.1) – (1.8.5)

### 1.9 MODEL INPUT

There are three different performance constraints for the mathematical model. A minimal margin (%) for the total assortment, a minimal margin per product group (%) and a minimal absolute margin per product. For these restriction variables should be initialized, which specify the value that the constraint should fulfill.

**Minimal profit margin (%) for the total assortment**

In this case, the assortment in total should fulfill the minimal profit margin restriction. If the parameter is set on 30% and the example in Table 12 is considered, no products are candidate for removal, since the margin (%) over the total assortment is 31.6% in this example. With this restriction, products with a high margin can compensate for products with a low margin.

**TABLE 12: EXAMPLE 1**

<table>
<thead>
<tr>
<th>Product</th>
<th>Revenue</th>
<th>Margin</th>
<th>Margin (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1000</td>
<td>500</td>
<td>50</td>
</tr>
<tr>
<td>B</td>
<td>1000</td>
<td>400</td>
<td>40</td>
</tr>
<tr>
<td>C</td>
<td>100</td>
<td>35</td>
<td>35</td>
</tr>
<tr>
<td>D</td>
<td>10000</td>
<td>2900</td>
<td>29</td>
</tr>
<tr>
<td>E</td>
<td>100</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>12200</td>
<td>3855</td>
<td>31.6%</td>
</tr>
</tbody>
</table>

If the parameter would be raised to 40%, product D would be removed, see Table 13. Since, the absolute margin of product D is quite high comparing to the products with a higher margin (%), these products cannot compensate enough anymore.
Minimal profit margin (%) per product group

This restriction works almost the same as the previous restriction. However the restriction should hold for a subset of products, namely a certain product group instead of the total assortment.

Minimal absolute profit margin per product

In this case a minimal margin is set in an absolute way. Idea behind this that there might be products with a very high margin (%), but with a low absolute margin. For example a product with cost price €1 and sale price €10 and only sold once a year. This product has a very high margin as a percentage (90%), but does not contribute much to the company (just €9/year).

Consider the example in Table 14 and let the minimal profit margin in absolute terms be 100. In this case product C and E will be candidates to be removed from the assortment.

The minimal margin for the assortment (%) and absolute minimal margin per product can be set by the slider or by pressing the ‘Specify Value’ button. A pop-up will arise and it is possible to give a value. For the minimal margin per product group (%), the value can be entered by pointing the cursor in the table and typing the desired value.
After having set all the values for the performance constraints, it is possible to do an optimization run.

**Minimal customer margin**

Next to the three performance constraints, there is a fourth parameter which should be initialized, the minimal customer margin. This margin distinguishes important customers from less important customers. If customers fulfill this minimal customer margin, they are considered to be important. This distinction is required, because of the following reason. After assumptions on customer retention have been made, it can be possible that customers are lost when removing a candidate product. However, if the contribution of this customer is very low, it is questionable to leave the candidate product in the assortment for this customer. This minimal customer margin determines to which extent customers should be taken into account.

Say you want to remove a candidate product with a yearly profit of only €50, which is bought by customer X. Assume that customer X is lost when removing the candidate product. The contribution of this customer in total is only €100 a year. If you would leave the minimal customer margin to zero, the tool will advise to keep this product in the assortment for this customer. However, if you would raise the minimal customer margin to for example €500, the tool will advise to remove the candidate product from the assortment.

1.10 **MODEL OUTPUT**

The model output screen can be reached both through the menu navigation or via 'View Output' in the tool dashboard. The model output is split in product and customer output.

1.10.1 **PRODUCT OUTPUT**

The output is first displayed at product group level. In the most left column the product groups are displayed corresponding to the values in the following columns. There are 4 performance indicators both displayed in red and green for each product group. The numbers in red indicate the current state of assortment as shown in Microsoft Navision. The numbers in green display the values after the optimization run. To see the output on product level, select the product group for which you want to see the products and click 'Zoom in on group'.
Formulas (1.10.1) – (1.10.8)

**FIGURE 21: OUTPUT ON PRODUCT GROUP LEVEL**

The output on product level is divided into two tables. Products that should be in and products that should be out the assortment according to the mathematical model. The products in the second table are the so-called candidate products. The candidate products can be further analyzed by selecting the product in the dropdown menu and pressing the ‘candidate button’.

**FIGURE 22: OUTPUT ON INDIVIDUAL PRODUCT LEVEL**

Formulas (1.10.9) - (1.10.14)

**1.10.2 CUSTOMER OUTPUT**

The output on customer group level looks similar to the output on product group level. However there is not distinction between size of the group. Here, it is also possible to see the output on individual customer level, by clicking ‘Zoom in on group’.
Designing a decision support tool for assortment planning

FIGURE 23: OUTPUT ON CUSTOMER GROUP LEVEL

Formulas (1.10.15) – (1.10.20)

On individual customer level, the revenue, margin and margin (%) are again both displayed in red and green, to distinguish between the current situation and the model output. In this window it is also possible to select a customer and to observe their basket. Via the basket overview it is displayed how much units the selected customer ordered from what product and how much profit it has yielded.

FIGURE 24: OUTPUT ON INDIVIDUAL CUSTOMER LEVEL

Formulas (1.10.21) – (1.10.26)
Formulas (1.10.27) – (1.10.28)

1.11 ANALYZING CANDIDATES

Customers
If a candidate product is selected, the user can see various information. For each candidate product it is displayed which customer bought the product and how much revenue and margin both in absolute as in percentage that customer contributed. Furthermore, it is displayed what other products those customers bought and how much that contributed to the profit.

Product relations
Next to the information of the customers who bought the product, the relation with other products is given. The relation with other products can be based on two shopping basket definitions. The first one is that customers buy multiple products during the period, these products then form the shopping basket of that customer. The second definition of is a shopping basket is the set of products in one specific order/transaction. The relations based on definition one are displayed in the left table of Figure 27, whereas the relations based on definition two are displayed in the right table of Figure 27.
FIGURE 27: ANALYZING CANDIDATE PRODUCT IN RELATION TO OTHER PRODUCTS

Product relations where a basket is the set of products ordered during the period.
The 'relation (%) (X→Y)' column, indicates whether the selected candidate product implicates another product. E.g. there is a product in the table with a relation 50% in the left table, this implicates that 50% of the customers that buy the candidate product also buy the product displayed in the table. This measure is also known as the confidence. The column ‘Support Y’ indicates by how many customers the product listed in the table is bought. The last column displays the opposite relations. So how many customers in terms of percentage buy the candidate product, if the product listed in the table is bought.

Formulas (1.7.1) – (1.7.4)

Product relations where a basket is the set of products in one transaction.
Again, the 'relation (%) (X→Y)' column, indicates whether the selected candidate product implicates another product. E.g. there is a product in the table with a relation 30% in the left table, this implicates that when the candidate product is ordered, in 30% of the cases the product displayed in the table is bought as well in that transaction. As in the other definition, it is also shown in how many transactions the other product occurred and the opposite relations.

In both tables it is possible to filter by specifying a minimum confidence. Furthermore it is displayed how many customers bought the candidate product and in how many transaction the candidate product occurred.

Formulas (1.7.5) – (1.7.8)

To make assumptions on customer and/or product retention press the 'Assumptions' button, in order to go to the Assumptions page.

1.12 MAKING ASSUMPTIONS

On the Assumptions page there are two lists displayed. The set of customers that bought the product on the left and the set of related products. The set of related products are determined by the minimum confidence that is specified in the product relations screen. It is possible to display only related products based on transactions and/or total customer basket.
The user can make a selection of customers and/or products, which are lost or for which the demand is lost when removing the candidate product. To make this selection hold the SHIFT or CTRL button and point the customers and/or products you want to select. To confirm the selection, press 'Make assumptions'. An overview of the assumptions made, can be seen by clicking on 'View assumptions'.

1.13 Final Results

Product Level

After the assumptions are made, a final result is calculated. There are three tables: the suggested assortment, a list of new products and the list of candidate products. For the new products the forecasted revenue, margin and margin as a percentage are displayed. Above that, the cannibalization and extra sales are displayed. At last, there is a forecasted net gain associated with the introduction of the new product displayed and an advise whether the product should be introduced or not, see also Figure 29.
The table of candidate products displays all the candidate products. For each candidate product the revenue, margin, margin (%), introduction date, inventory, durability, minimum order quantity and reorder quantity are displayed. Furthermore, it is advised whether the candidate product should be kept in or removed from the assortment. This advise is based on the procedure as described in Figure 7 located in chapter 4.3.2.

FIGURE 30: TABLE OF CANDIDATE PRODUCTS

Formulas (1.13.7) – (1.13.9)

Assortment level
The final impacts can also be viewed at assortment level, see Figure 31 and Figure 32. There are three situations distinguished. The current state of the assortment as in the data, the assortment after an optimization run and finally the state after the assumptions are made. For each state the following key performance indicators are displayed: Assortment size, total revenue, total costs, total profit and profit margin as a percentage.

FIGURE 31: FINAL IMPACT ON ASSORTMENT LEVEL
On the Results page, there is also an option to export the analysis you have done to Microsoft Excel. Just press the button 'Export to Excel'. When you have done this, an Excel file is created in the folder of the tool, called 'Output.xlsx'.

In the Excel document an overview is given of the results. The first sheet displays the suggested assortment. The second sheet gives the products that should be removed from the assortment and the third sheet gives an overview of the KPI's as in Figure 31.

**Warning:** The next time you export an analysis to Excel, the data is overwritten. If you want to use or show the results in, for example, a meeting, copy the file to another location. Do not use 'Cut' instead of 'Copy'! The original Excel document must remain in its location, since it has named ranges.
APPENDIX IV: MATHEMATICAL MODEL

Assumptions

1) The length of the period is a function of a year.
2) The demand for the coming year equals demand during last year.
3) Substitution does not occur between products within the assortment.
4) There are no storage space restrictions.
5) The demand can always be directly fulfilled.

List of variables

c, c ∈ {1 ... C}: Index denoting the customers.
d_{tcp}: Quantity of product p for customer c ordered in transaction t.
g, g ∈ {1 ... G}: Index of product groups under consideration.
MAPIP: Minimal absolute profit margin per product.
MRP_{PG_g}: Minimal relative profit margin per product group g.
MRPTA: Minimal relative profit margin on the total assortment.
p, p ∈ {1 ... P}: Index denoting the products.
t, t ∈ {1 ... T}: Index denoting the transactions.
TC: Total cost of goods sold over the period.
TP: Total profit over the period.
TR: Total revenues over the period.
u_{tcp}: Adjustment factor of the unit of measure.
w_{tcp}: Sale price of product p for customer c in transaction t.
X_p ∈ {0,1}: Decision variable which indicates if a product p is in the assortment (1), or not (0).

Objective function:

\[
\max TP = TR - TC \quad (4.1) \text{ Total profit over the period}
\]

Where

\[
TR = \sum_{t=1}^{T} \sum_{p=1}^{P} \sum_{c=1}^{C} w_{tcp} \cdot d_{tcp} \cdot X_p \quad (4.2) \text{ Total revenue}
\]
Designing a decision support tool for assortment planning

\[ TC = \sum_{t=1}^{T} \sum_{p=1}^{P} \sum_{c=1}^{C} c_p \cdot u_{tcp} \cdot d_{tcp} \cdot X_p \]  
\hspace{1cm} (4.3) Cost of goods

**Subject to:**

\[ \forall p, p \in \{1, ..., P\} \quad X_p \in \{0,1\} \]  
\hspace{1cm} (4.4) Binary bound

\[ \frac{TR - TC}{TR} \geq MRPTA \]  
\hspace{1cm} (4.5) Relative profit margin restriction total assortment.

\[ \forall g, g \in \{1, ..., G\} \quad \frac{\sum_{p \in g} \sum_{t=1}^{T} \sum_{c=1}^{C} w_{tcp} \cdot d_{tcp} \cdot X_p - \sum_{p \in g} \sum_{t=1}^{T} \sum_{c=1}^{C} c_p \cdot d_{tcp} \cdot u_{tcp} \cdot X_p}{\sum_{p \in g} \sum_{t=1}^{T} \sum_{c=1}^{C} w_{tcp} \cdot d_{tcp} \cdot X_p} \geq MRPPC_g \]  
\hspace{1cm} (4.6) Relative profit margin restriction product group \( g \).

\[ \forall p, p \in \{1, ..., P\} \quad \sum_{t=1}^{T} \sum_{c=1}^{C} w_{tcp} \cdot d_{tcp} \cdot X_p - \sum_{t=1}^{T} \sum_{c=1}^{C} c_p \cdot d_{tcp} \cdot u_{tcp} \cdot X_p \geq MAPIP \cdot X_p \]  
\hspace{1cm} (4.7) Absolute profit margin restriction per product \( p \).
APPENDIX V: LIST OF FORMULAS

1.7: View Data

Product relations where a basket is the set of products ordered during the period.

Set of customers that bought the candidate product: $S$, where $s \subseteq \{1, ..., C\}$

(1.7.1a) $\sum_{t=1}^{T} d_{tcp}$

Demand$(c, p)$: Total Demand for product $p$ by customer $c$.

(1.7.1b) $f(c, p) = \begin{cases} 1, & \text{Demand}(c, p) > 0 \\ 0, & \text{Demand}(c, p) = 0 \end{cases}$

Binary indicating if customer $c$ bought product $p$.

(1.7.1c) $\frac{\sum_{c \in s} f(c, p)}{|s|} \times 100$

Relation (%) $(X \rightarrow Y) (p)$

(1.7.2) $\sum_{c=1}^{C} f(c, p)$

Support $Y (p)$

(1.7.3) $\frac{\sum_{c \in s} f(c, p)}{\sum_{t=1}^{T} f(t, c, p)} \times 100$

Relation (%) $(Y \rightarrow X) (p)$

(1.7.4) $\frac{\sum_{c \in s} f(c, p)}{|s|} \times \frac{\sum_{t=1}^{T} f(t, c, p)}{|T|}$

Lift $(p)$

Product relations where a basket is the set of products in one transaction.

Set of transactions containing the candidate product: $ST$, where $ST \subseteq T$

(1.7.5a) $g(t, c, p) = \begin{cases} 1, & d_{tcp} > 0 \\ 0, & d_{tcp} = 0 \end{cases}$

Binary indicating if product $p$ is ordered in transaction $t$ of customer $c$.

(1.7.5b) $\frac{\sum_{t \in ST} \sum_{c=1}^{C} g(t, c, p)}{|ST|} \times 100$

Relation (%) $(X \rightarrow Y) (p)$

(1.7.6) $\sum_{t=1}^{T} \sum_{c=1}^{C} g(t, c, p)$

Support $Y (p)$

(1.7.7) $\frac{\sum_{t \in ST} \sum_{c=1}^{C} g(t, c, p)}{\sum_{t \in ST} \sum_{c=1}^{C} g(t, c, p)}$

Relation (%) $(Y \rightarrow X) (p)$

(1.7.8) $\frac{\sum_{t \in ST} \sum_{c=1}^{C} g(t, c, p)}{|ST|} \times \frac{\sum_{t \in ST} \sum_{c=1}^{C} g(t, c, p)}{|T|}$

Lift $(p)$

Basket overview for customer $c$

(1.7.9) $\sum_{t=1}^{T} d_{tcp}$

Quantity ordered for product $p$ by customer $c$.

(1.7.10) $\sum_{t=1}^{T} (w_{tcp} - c_{p}) \cdot d_{tcp} \cdot X_{p}$

Margin for product $p$ for customer $c$. 
1.8: Charts

(1.8.1) \[ y(p) = \sum_{t=1}^{T} \sum_{c=1}^{C} (w_{tcp} - c_p) \cdot d_{tcp} \]

Absolute margin per product \((p)\)

Order \(P\) descending by \(y(p)\). So \(p = 1 \rightarrow \max_{p \in \{1,...,P\}} y(p)\)

(1.8.2) \[ \sum_{p=1}^{P} y(p) \]

Cumulative margin \((P)\)

(1.8.3) \[ r(p) = \sum_{t=1}^{T} \sum_{c=1}^{C} w_{tcp} \cdot d_{tcp} \]

Revenue \((p)\)

(1.8.4) \[ \sum_{p=1}^{P} r(p) \]

Cumulative revenue \((P)\)

(1.8.5) \[ \frac{\text{cumulative margin (}P\text{)}}{\text{cumulative revenue (}P\text{)}} \times 100 \]

Cumulative margin percentage \((P)\)

1.10: Model Output

1.10.1: Product output

\(G: \text{Product group}\)

(1.10.1) \(|G|\)

Current size \((g)\)

(1.10.2) \[ \sum_{p \in G} X_p \]

Optimal size \((g)\)

(1.10.3) \[ \sum_{p \in G} \sum_{t=1}^{T} \sum_{c=1}^{C} w_{tcp} \cdot d_{tcp} \]

Current revenue \((g)\)

(1.10.4) \[ \sum_{p \in G} \sum_{t=1}^{T} \sum_{c=1}^{C} w_{tcp} \cdot d_{tcp} \cdot X_p \]

Optimal revenue \((g)\)

(1.10.5) \[ \sum_{p \in G} \sum_{t=1}^{T} \sum_{c=1}^{C} (w_{tcp} - c_p) \cdot d_{tcp} \]

Current margin \((g)\)

(1.10.6) \[ \sum_{p \in G} \sum_{t=1}^{T} \sum_{c=1}^{C} (w_{tcp} - c_p) \cdot d_{tcp} \cdot X_p \]

Optimal margin \((g)\)

(1.10.7) \[ \frac{\text{current margin (}\ g\text{)}}{\text{current revenue (}\ g\text{)}} \times 100 \]

Current margin \((\%)\) \((g)\)

(1.10.8) \[ \frac{\text{optimal margin (}\ g\text{)}}{\text{optimal revenue (}\ g\text{)}} \times 100 \]

Optimal margin \((\%)\) \((g)\)

Table IN Assortment:

(1.10.9) \[ \sum_{t=1}^{T} \sum_{c=1}^{C} w_{tcp} \cdot d_{tcp} \cdot X_p \]

Revenue \((p)\)

(1.10.10) \[ \sum_{t=1}^{T} \sum_{c=1}^{C} (w_{tcp} - c_p) \cdot d_{tcp} \cdot X_p \]

Margin \((p)\)

(1.10.11) \[ \frac{\text{margin (}\ p\text{)}}{\text{revenue (}\ p\text{)}} \times 100 \]

Margin \((\%)\) \((p)\)

Table OUT Assortment:

(1.10.12) \[ \sum_{t=1}^{T} \sum_{c=1}^{C} w_{tcp} \cdot d_{tcp} \]

Revenue \((p)\)

(1.10.13) \[ \sum_{t=1}^{T} \sum_{c=1}^{C} (w_{tcp} - c_p) \cdot d_{tcp} \]

Margin \((p)\)
Designing a decision support tool for assortment planning

\[ \text{Margin (p)} \times 100 \]

1.10.2: Customer output

**CG: Customer group**

\[ \text{Number of customers (cg)} \]
\[ \sum_{t=1}^{T} \sum_{p=1}^{P} w_{tcp} \cdot d_{tcp} \]
\[ \text{Current revenue (cg)} \]
\[ \sum_{t=1}^{T} \sum_{p=1}^{P} w_{tcp} \cdot d_{tcp} \cdot X_{p} \]
\[ \text{Optimal revenue (cg)} \]
\[ \sum_{t=1}^{T} \sum_{p=1}^{P} (w_{tcp} - c_{p}) \cdot d_{tcp} \]
\[ \text{Current margin (cg)} \]
\[ \sum_{t=1}^{T} \sum_{p=1}^{P} (w_{tcp} - c_{p}) \cdot d_{tcp} \cdot X_{p} \]
\[ \text{Optimal margin (cg)} \]

\[ \left( \frac{\text{Current margin (cg)}}{\text{Current revenue (cg)}} \right) \times 100 \]
\[ \text{Current margin (cg)} \]

\[ \left( \frac{\text{Optimal margin (cg)}}{\text{Optimal revenue (cg)}} \right) \times 100 \]
\[ \text{Optimal margin (cg)} \]

**Customer level**

\[ \text{Current revenue (c)} \]
\[ \sum_{t=1}^{T} \sum_{p=1}^{P} w_{tcp} \cdot d_{tcp} \]
\[ \text{Optimal revenue (c)} \]
\[ \sum_{t=1}^{T} \sum_{p=1}^{P} w_{tcp} \cdot d_{tcp} \cdot X_{p} \]
\[ \text{Current margin (c)} \]
\[ \sum_{t=1}^{T} \sum_{p=1}^{P} (w_{tcp} - c_{p}) \cdot d_{tcp} \]
\[ \text{Optimal margin (c)} \]
\[ \sum_{t=1}^{T} \sum_{p=1}^{P} (w_{tcp} - c_{p}) \cdot d_{tcp} \cdot X_{p} \]
\[ \left( \frac{\text{Current margin (c)}}{\text{Current revenue (c)}} \right) \times 100 \]
\[ \text{Current margin (c)} \]
\[ \left( \frac{\text{Optimal margin (c)}}{\text{Optimal revenue (c)}} \right) \times 100 \]
\[ \text{Optimal margin (c)} \]

**Basket overview for customer c**

\[ \text{Quantity ordered for product p by customer c} \]
\[ \sum_{t=1}^{T} d_{tcp} \cdot X_{p} \]

\[ \text{Margin for product p for customer c} \]
\[ \sum_{t=1}^{T} (w_{tcp} - c_{p}) \cdot d_{tcp} \cdot X_{p} \]

1.11: Analyzing candidates

Customers c that bought product p

\[ \sum_{t=1}^{T} \sum_{p=1}^{P} w_{tcp} \cdot d_{tcp} \cdot X_{p} \]
\[ \text{Customer revenue (c)} \]
\[ \sum_{t=1}^{T} \sum_{p=1}^{P} (w_{tcp} - c_{p}) \cdot d_{tcp} \cdot X_{p} \]
\[ \text{Customer margin (c)} \]
Designing a decision support tool for assortment planning

1.13: Final impacts

Product level

New products

Set of new products: \( NP, \text{where } NP \subseteq \{1, \ldots, P\} \)

\[
\begin{align*}
(1.13.1) & \quad \sum_{t=1}^{T} \sum_{c=1}^{C} w_{tcp} \cdot d_{tcp} \cdot X_p & \text{Revenue (p)} \\
(1.13.2) & \quad \sum_{t=1}^{T} \sum_{c=1}^{C} (w_{tcp} - c_p) \cdot d_{tcp} \cdot X_p & \text{Margin (p)} \\
(1.13.3) & \quad \frac{\text{Margin (p)}}{\text{Revenue (p)}} \times 100 & \text{Margin (\%)(p)} \\
\end{align*}
\]

Cannibalization rate from new product \( p^* \) to product \( p \): \( q_{p^*p} \)

\[
\begin{align*}
(1.13.4) & \quad \sum_{p=1}^{P} q_{p^*p} \cdot \sum_{t=1}^{T} \sum_{c=1}^{C} (w_{tcp} - c_p) \cdot d_{tcp} \cdot X_p & \text{Cannibalization loss associated with introduction of product } p^*. \\
\end{align*}
\]

Percentage of increase in sales for product \( p \) when introducing new product \( p^* \): \( s_{p^*p} \)

\[
\begin{align*}
(1.13.5) & \quad \sum_{p=1}^{P} s_{p^*p} \cdot \sum_{t=1}^{T} \sum_{c=1}^{C} (w_{tcp} - c_p) \cdot d_{tcp} \cdot X_p & \text{Gains of extra sales associated with the introduction of product } p^*. \\
\end{align*}
\]

\[
(1.13.6) \quad \text{Margin}(p) - \text{cannibalization loss}(p) + \text{gains of extra sales (p)}
\]

Net gain (p)

Candidate products:

\[
\begin{align*}
(1.13.7) & \quad \sum_{t=1}^{T} \sum_{c=1}^{C} w_{tcp} \cdot d_{tcp} & \text{Revenue (p)} \\
(1.13.8) & \quad \sum_{t=1}^{T} \sum_{c=1}^{C} (w_{tcp} - c_p) \cdot d_{tcp} & \text{Margin (p)} \\
(1.13.9) & \quad \frac{\text{Margin (p)}}{\text{Revenue (p)}} \times 100 & \text{Margin (\%)(p)} \\
\end{align*}
\]

Assortment level

Current situation

\[
\begin{align*}
(1.13.5) & \quad |P| & \text{Current assortment size} \\
(1.13.6) & \quad \sum_{p=1}^{P} \sum_{t=1}^{T} \sum_{c=1}^{C} w_{tcp} \cdot d_{tcp} & \text{Current revenue} \\
(1.13.7) & \quad \sum_{p=1}^{P} \sum_{t=1}^{T} \sum_{c=1}^{C} c_p \cdot d_{tcp} & \text{Current costs} \\
\end{align*}
\]
Designing a decision support tool for assortment planning

\[ (1.13.8) \sum_{p=1}^{P} \sum_{t=1}^{T} \sum_{c=1}^{C} (w_{tcp} - c_p) \cdot d_{tcp} \]

Current margin

\[ (1.13.9) \frac{\text{Current margin}}{\text{Current revenue}} \times 100 \]

Current margin (%)

Optimal before assumptions

\[ (1.13.10) \sum_{p=1}^{P} X_p \]

Optimal assortment size

\[ (1.13.11) \sum_{p=1}^{P} \sum_{t=1}^{T} \sum_{c=1}^{C} w_{tcp} \cdot d_{tcp} \cdot X_p \]

Optimal revenue

\[ (1.13.12) \sum_{p=1}^{P} \sum_{t=1}^{T} \sum_{c=1}^{C} c_p \cdot d_{tcp} \cdot X_p \]

Optimal costs

\[ (1.13.13) \sum_{p=1}^{P} \sum_{t=1}^{T} \sum_{c=1}^{C} (w_{tcp} - c_p) \cdot d_{tcp} \cdot X_p \]

Optimal margin

\[ (1.13.14) \frac{\text{Optimal margin}}{\text{Optimal revenue}} \times 100 \]

Optimal margin (%)

Optimal after assumptions

\[ |\text{ASS}| \]

Optimal assortment size

\[ (1.13.16) \sum_{p \in \text{ASS}} \sum_{t=1}^{T} \sum_{c=1}^{C} w_{tcp} \cdot d_{tcp} \]

Optimal revenue

\[ (1.13.17) \sum_{p \in \text{ASS}} \sum_{t=1}^{T} \sum_{c=1}^{C} c_p \cdot d_{tcp} \]

Optimal costs

\[ (1.13.18) \sum_{p \in \text{ASS}} \sum_{t=1}^{T} \sum_{c=1}^{C} (w_{tcp} - c_p) \cdot d_{tcp} \]

Optimal margin

\[ (1.13.19) \frac{\text{Optimal margin}}{\text{Optimal revenue}} \times 100 \]

Optimal margin (%)
Training Assortment Tool

Exercises

1) Load the data from the ODS database with a period of 1 year.

2) Save the data into a case.

3) Select all product groups and inspect the data on the ‘Data’ page.

4) Do an assortment analysis on the product group ‘Grondstof’.

   *In this first analysis, you do not necessarily need to set the parameters.*

5) Clear the data and load the case you saved earlier.

6) Do an assortment analysis on the product group ‘Collegiale’ and ‘Grondstof’ and add a new product.

   *For the required data for the new product, you may choose any.*

   *For the model input use the following values:*
   
   a. Minimal margin per product: 150
   
   b. Minimal margin per product group ‘Collegiale’: 54%
   
   c. Customer margin: 1500

7) Clear the data and load the case you saved earlier.

8) Do an assortment analysis on the product group ‘Grondstof’ and also fix a product

   *For the fixed product, you may choose any.*

   *For the model input use the following values:*
   
   a. Minimal margin per product: 500
   
   b. Minimal margin for the product group ‘Grondstof’: 60.50%
   
   c. Customer margin: 1000

9) Export your last analysis to Excel and also inspect the Excel document.

10) You can now perform an assortment analysis, where you are totally free.
## Training certificate user

### Assortment tool

By signing this, I confirm that I have participated in the training for the assortment tool.

<table>
<thead>
<tr>
<th>Name</th>
<th>Signature</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

**Confirmation project owner**

Erik Hoppenbrouwer

---

**Confirmation supervisor**

Joost Groothuizen

---

**Confirmation project leader**

Erik Stienstra

---
APPENDIX VII: FAGRONIZED TOOL