Eindhoven University of Technology

MASTER

Accurate response to fluctuating demand with regard to promotions in the online retailing industry

Kiemeneij, N.A.

Award date:
2018

Link to publication

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

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

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

Accurate response to fluctuating demand with regard to promotions in the online retailing industry

In partial fulfillment of the requirements for the degree of Master of Science in Operations Management and Logistics

Author: N.A. Kiemeneij - 0775814

Supervisors TU/e: Dr. Z. Atan, OPAC
Dr. N.R. Mutlu, OPAC

Supervisor bol.com: Ing. M. Engelen, Teamlead Supply Chain
Subject headings: promotion, online retailer, promotion forecasting, inventory replenishment, inventory policy, multiple regression analysis, lost sales, excess stock, forecast adjustment
Preface

This report is the result of my master thesis conducted to complete my education at Eindhoven University of Technology. It is the final stage of my master in Operations Management and Logistics and my student life.

Therefore, I want to thank a number of people who have helped me during my master thesis. First, I would like to thank my mentor, Zümbül Atan. You were a great mentor for me and helped me a great deal. You gave me very good feedback with a positive approach. I also really liked our non-thesis related conversations. Thank you very much for everything. Thanks to you I came this far. Second, I would like to thank my second supervisor, Nevin Mutlu. You helped me to think further and explain the things I did in more detail. Thank you for your contributions.

Next, I want to thank my supervisor at bol.com, Mark Engelen. I am glad you became my supervisor. You helped me to get to know many relevant people within bol.com. You made sure that these people made time for me. I am really thankful for that.

Furthermore, I want to thank my family for their support and help. They helped me not to worry too much and enjoy the moment. In addition, my friends helped me a lot. They gave me some textual and substantive advice. With them my student life in Eindhoven became a big experience I would never have missed.

Nina Kiemeneij
Abstract
Companies in the online retailing industry face extensive inventory costs or lost sales due to fluctuating demand during promotions. In this research, a promotion is defined as the temporarily extra-highlighting of the assortment by the retailer. Promotions are planned beforehand, thus the decision to promote does not depend on the level of inventory. Accurate forecasts and replenishments are critical to the success of a promotion (Ettouzani al., 2012). The service value of a product can only be created through a sale if the product is available. In this master thesis, a promotion forecasting model and an adjusted inventory replenishment model are created. The promotion forecasting model makes a prediction of the lift factor by using multiple regression analysis. The predictors of the regression model are identified using literature research and semi-structured interviews. Besides that, an inventory replenishment policy for promotions is created based on the current (R,s,nQ) inventory policy. The initial forecast from the promotion forecasting model is adjusted to actual sales data using time series forecasting techniques. This facilitates an improved response to promotional demand. Finally, based on the analysis, an improved promotion forecasting model and inventory replenishment policy are introduced to accurately respond to the fluctuating demand. The forecasting and inventory model developed in this research shows a substantial difference compared to the old method. The realized mean service levels are 2.6% lower than the realized mean service levels of bol.com. However, the final inventory level shows a substantial difference. The model results in a 72.9% smaller final inventory level, which results in a better trade-off between high inventory costs and lost sales.
Executive summary

This report represents a study concerning the accurate response to fluctuating demand with regard to promotions in the online retailing industry by making use of promotion forecasting and adjusted inventory policy. In this research, a promotion is defined as the temporarily extra-highlighting of the assortment by the retailer. Promotions are planned beforehand, thus the decision to promote does not depend on the level of inventory. The goal of a promotion is to generate traffic on the website of the online retailer, such that products will be bought and profit can be made. In addition, in this research promotion forecasting refers to the prediction of future demand during promotions by using regression techniques. Inventory management includes controlling the level of product availability through holding stock. A retailer who plans a promotion faces a fundamental issue. It is not possible to perfectly predict the demand during the promotion. Overestimation of the demand results in excess inventory and an underestimation results in lost sales (Arminger, 2008; Hopp, 2008). Therefore, it is important to find the balance between lost sales and inventory costs. Concluding, the question researched in this report is:

How can an online retailing company accurately respond to fluctuating demand with regard to promotions by the use of a promotion forecasting model and inventory policy?

In this research, an improved promotion forecasting model was created using multiple regression analysis. A promotion forecast should include particular independent variables that influence the buying behavior of people (van Donselaar et al., 2006). These independent variables were identified by literature research and semi-structured interviews. Table 1 lists the final independent variables and the dependent variable.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Description</th>
<th>Variable Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>$y$</td>
<td>Lift factor</td>
<td>Metric</td>
</tr>
<tr>
<td>$x_1$</td>
<td>Product subgroups</td>
<td>Nonmetric</td>
</tr>
<tr>
<td>$x_2$</td>
<td>Discount</td>
<td>Metric</td>
</tr>
<tr>
<td>$x_3$</td>
<td>Duration promotion</td>
<td>Metric</td>
</tr>
<tr>
<td>$x_4$</td>
<td>Type campaign</td>
<td>Nonmetric</td>
</tr>
<tr>
<td>$x_5$</td>
<td>Action discount type</td>
<td>Nonmetric</td>
</tr>
<tr>
<td>$x_6$</td>
<td>Number of products in action</td>
<td>Metric</td>
</tr>
</tbody>
</table>

Table 1: Overview of the final factors

Due to the data uncertainties in the promotion tool five different models are created and examined, using the stepwise estimation method. These models predict the total lift factor for every unique Stock Keeping Unit (SKU) during the whole promotion period in order to correct for daily fluctuations. These models are validated using different performance measures. Finally, when comparing the different models, model 2 performs the best. This model includes all available data and assigns the daydeals to the corresponding campaigns. Model 2 shows a substantial improvement compared to the current, subjective promotion forecasting technique of the online retailer.

Although the relatively low explanatory power, the model shows a substantial improvement compared to the current, subjective promotion forecasting technique of the online retailer. Besides the promotion forecast model, an inventory replenishment policy for promotion is made. This model is based on the current (R,s,nQ) inventory policy of the company with an accurate response to actual sales data. The initial forecast, which is the preferred promotion forecasting model as described in the previous paragraph, is adjusted using the simple moving average and simple exponential smoothing technique. In addition, the safety stock of the inventory model can be calculated by multiplying the safety factor with the standard deviation of the forecast error. The Root Mean Squared Error (RMSE)
is used for the standard deviation of the forecast error. This method is directly related to the $\sigma$ and can be easily computed. The RMSE is calculated in two ways based on all previous days and based on the review period. To find the safety factor for the discrete demand distribution the fitting procedure of Adan et al. (1995) is used.

The developed promotion forecasting models predict the demand for a complete promotion. To divide the expected demand per day in a promotion, a demand pattern needs to be determined. Due to the limited data, it was not possible to identify a demand pattern for an individual campaign. Therefore, two different demand patterns are introduced based on merged campaigns. First, a Uniform demand pattern is introduced. Second, multiple ratio demand patterns based on different time periods are compared to define the most appropriate demand pattern. A distinction between the different kinds of campaigns is made. The first type includes campaigns which prepare for a special holiday, like Christmas or Valentine’s Day. These kind of days give an uplift to the sales as the day approaches, causing a peak in the sales (Ramanathan, 2012). The other campaigns are unique for the online retailer and are introduced to bring the company more in front of attention. These are named as stable campaigns. To find the second demand pattern the following steps are applied:

1. Normalize daily sales for every promotion against a specific number of time periods (9, 18 or 36 time periods) with a distinction between stable and peak campaigns.
2. Find an empirical distribution based on the normalized sales in which all promotions are merged.
3. Determine possible theoretical distributions.
4. Test the possible theoretical distributions according to a goodness of fit test.
5. Fit the best theoretical distribution to the empirical distribution.

According to above steps, the Logistic Distribution based on 36 time periods seems the most appropriate distribution. Both the stable and peak campaigns fit a Logistic Distribution with different parameters. Therefore, the Logistic Distribution and Uniform Distribution are used as final demand patterns. These distributions divide the expected total demand from the preferred promotion forecasting model over the days.

Finally, based on the different forecast adjustments, forecast errors and demand patterns, eight different models are created. The simple moving average technique and the forecast error based on all previous days do not respond abruptly to sales fluctuations. Therefore, for this research these techniques perform the best. Future research with a clear demand pattern is needed to draw conclusions on the best method. After the first forecast adjustment a clear learning effect can be seen, whereafter no further improvements can be noticed. Therefore, the learning effect does not improve over time. It is not necessary to adjust the forecast continuously, one adjustment is enough to improve the forecast accuracy. The forecasting and inventory model developed in this research shows a substantial difference compared to the old method. The realized mean service levels are a 2.6% lower than the current realized mean service levels of the company. However, the final inventory level shows a substantial difference. The model results in a 72.9% smaller final inventory level, which results in a better trade-off between high inventory costs and lost sales. Following, this research results in the following recommendations:
• **Data collection**
  To draw better conclusions, more data is needed. Important data to create an accurate promotion forecasting model is missing. When more data is collected better relationships and demand patterns can be found.

• **Inform and learn employees about promotion tool**
  The promotion tool is not used appropriately by the employees of the company. To make an appropriate analysis of the promotional data, the employees need to be informed about the purposes of the system and get a proper feeling how to work with the system.

• **Update reference period**
  This research is based on a given reference period which refers to exactly the same days as before the start of the promotion. During this reference period often another promotion takes place. This makes it difficult to withdraw conclusions about the uplift and price discount of the promotion. Therefore, it is necessary to get another reference period which is representative for the non-promotional demand. This would make it easier to compare different promotions. More research is needed to come up with the best reference period.

• **Change the current promotion forecasting method to a statistical method**
  At the moment the company is using a subjective forecasting method based on human judgment to predict the promotional demand. This research shows a substantial increase in the forecast accuracy when making a forecast based on multiple regression analysis. Therefore, it is recommended to change the current promotion forecasting method to a statistical method.

• **Test and implement new inventory policy**
  In this research, a new inventory policy for promotions is introduced. The relative low explanatory power of the promotion forecasting model and the aggregate demand patterns have a big impact on the inventory levels. This makes it difficult to draw conclusions about the different models used for the inventory model. To find the best model this inventory policy first needs to be tested with a better forecasting model and demand pattern. Overall, this inventory policy already showed a big improvement compared to the current situation. Therefore, it is recommended to change the current replenishment strategy for promotions to the inventory model introduced in this research and do more research to identify the best model.

• **Adjust forecast at the beginning of the promotion**
  The introduced inventory policy is based on an accurate response to sales data. It seems that especially in the beginning of the promotion the forecast becomes better after response to the actual demand. Therefore, it is necessary to adjust the forecast as soon as possible on the actual sales. This results in better demand forecasts, less excess of stock and lost sales.
## Table of Contents

Preface ........................................................................................................................................... i
Abstract ........................................................................................................................................ ii
Executive summary ...................................................................................................................... iii
List of figures .................................................................................................................................. viii
List of tables ................................................................................................................................... ix
List of abbreviations .................................................................................................................... x

1 Introduction .................................................................................................................................. 1
   1.1 Problem definition ................................................................................................................ 1
   1.2 Problem analysis .................................................................................................................... 1
   1.3 Research scope ..................................................................................................................... 2
   1.4 Research questions .............................................................................................................. 3
   1.5 Research methodology ...................................................................................................... 3
      1.5.1 Regulative cycle ......................................................................................................... 3
      1.5.2 General approach ....................................................................................................... 4
   1.6 Outline .................................................................................................................................. 4

2 Case setting online retailer .......................................................................................................... 5
   2.1 Company background ......................................................................................................... 5
   2.2 Promotions ......................................................................................................................... 6
      2.2.1 Promotion process ..................................................................................................... 6
   2.3 Demand forecasting and inventory policy ........................................................................... 8
      2.3.1 Inventory policy ......................................................................................................... 8
   2.4 Product groups .................................................................................................................. 9
   2.5 Data analysis current situation ......................................................................................... 10
      2.5.1 Current performance ............................................................................................... 11

3 Promotion Forecasting ............................................................................................................... 14
   3.1 Multiple regression ............................................................................................................ 14
   3.2 Variables forecasting model ............................................................................................. 14
      3.2.1 Dependent variable ................................................................................................. 14
      3.2.2 Independent variable .............................................................................................. 15
      3.2.3 Final factors ............................................................................................................. 17
   3.3 Data preparation ................................................................................................................ 18
      3.3.1 Data split ................................................................................................................ 18
      3.3.2 Outliers .................................................................................................................. 18
   3.4 Assumptions for multiple regression ............................................................................... 19
   3.5 Estimation technique ......................................................................................................... 21
List of figures

Figure 1: The regulative cycle ................................................................. 4
Figure 2: Structure of bol.com ................................................................. 5
Figure 3: Promotions process at bol.com .................................................. 7
Figure 4: Example of the reference period of a promotion................................. 11
Figure 5: Main inventory per product for the campaign: Bulk May ....................... 12
Figure 6: Number of products per campaign ............................................... 13
Figure 7: Influential observations by the Cook’s distance .................................. 19
Figure 8: Timeline of a promotion ................................................................ 27
Figure 9: Example of promotion divided into time periods ................................. 30
Figure 10: Example of promotion divided into many time periods ....................... 31
Figure 11: Cumulative probability plot of the duration of the promotions ............. 31
Figure 12: Empirical distribution for 36 time periods ....................................... 32
Figure 13: Empirical distribution for 9 time periods ....................................... 32
Figure 14: Empirical distribution for 36 time periods ....................................... 32
Figure 15: Product hierarchy for Personal Care & Home Care ............................ 47
Figure 16: Product hierarchy for Televisions .................................................. 47
List of tables

Table 1: Overview of the final factors ................................................................. iii
Table 2: Target service levels per class .............................................................. 8
Table 3: Overview of data filtering ................................................................... 10
Table 4: Overview of the current performance on the KPI’s ............................... 12
Table 5: Overview of important independent variables for promotion forecasting ............................................ 16
Table 6: Relevant factors according to the literature and interviews ................ 17
Table 7: Overview of the final factors ................................................................ 17
Table 8: Summary of the not included factors .................................................... 18
Table 9: Overview of the different models .......................................................... 22
Table 10: Summary of the main characteristics of the models ............................ 24
Table 11: Overview of the performance of the different models .......................... 24
Table 12: Overview of the realized mean service levels ...................................... 25
Table 13: Overview of different campaigns with indication of the corresponding type ........................................... 30
Table 14: Suggested Theoretical Distributions ................................................. 33
Table 15: Corresponding parameters to the different campaigns ....................... 33
Table 16: Overview of the different models ........................................................ 34
Table 17: Achieved mean service levels .............................................................. 34
Table 18: Difference current situation and preferred model ............................... 35
Table 19: Overview of the interviewed employees of bol.com .......................... 48
Table 20: Overview of all campaigns of bol.com ................................................. 50
Table 21: Final promotion forecasting models .................................................... 64
List of abbreviations

F&S  Finance & Support
FinCat  Financial Category
FMCG  Fast Moving Consumer Goods
KPI  Key Performance Indicator
M&S  Merchandising & Sourcing
MAPE  Mean Absolute Percentage Error
O&F  Operations & Fulfillment
OOS  Out-of-stock
OSA  On-shelf availability
RMSE  Root Mean Squared Error
SCM  Supply Chain Management
SKU  Stock Keeping Unit
VIF  Variance Inflation Factor
1 Introduction

This report represents a study concerning the accurate response to fluctuating demand with regard to promotions in the online retailing industry by making use of promotion forecasting and adjusted inventory policy. A promotion is defined as the temporarily extra-highlighting of the assortment by the retailer. Promotions are planned beforehand, thus the decision to promote does not depend on the level of inventory. The goal of a promotion is to generate traffic on the website of the online retailer, such that products will be bought and profit can be made. In addition, promotion forecasting refers to the prediction of future demand during promotions by using regression techniques. Furthermore, inventory management includes controlling the level of product availability through holding stock. This study is the result of the master thesis project conducted at the Eindhoven University of Technology in collaboration with bol.com. The first chapter is an introduction to the rest of the report. In Section 1.1 the problem is explained, followed by the problem analysis in Section 1.2 and definition of the scope in Section 1.3. Then the research questions are formulated in Section 1.4 and Section 1.5 describes the research methodology. Finally, Section 1.6 outlines the remainder of the report.

1.1 Problem definition

Fierce competition for market share has resulted in a significant increase in the number, frequency and depth of promotions by retailers (Ailawadi et al., 2001; IGD Supply Chain Analysis, 2007; McKinnon et al., 2007). Therefore, promotional activity has become a fundamental element of retailer strategy, as well as a key determinant of the performance of a retailer (Ailawadi et al., 2001). According to Kirkup (1999), the main reasons for a retailer to promote is to increase the footfall, i.e. the number of people visiting a retailer store in a particular period of time, develop store loyalty and increase the quantity and frequency of a purchase. Demand during promotions is more subjected to fluctuations than demand for non-promoted stock keeping units (SKU’s), a SKU is a unique product. Therefore, it is difficult to accurately forecast the demand. This often leads to availability problems (Taylor and Fawcett, 2001; Gruen et al., 2002; Dubelaar et al., 2001; McKinnon et al., 2007). During this research the words SKU and product are used interchangeably. Consequently, promotional activity has been observed as one of the root causes of poor on-shelf-availability (OSA) performance (Corsten & Gruen, 2003; Taylor & Fawcett, 2001). Furthermore, the processes for running and replenishing a promotion are critical for its success (Ettouzani et al., 2012). A retailer who plans a promotion faces a fundamental issue. It is not possible to perfectly predict the demand during the promotion. Overestimation of the demand results in excess inventory and underestimation results in lost sales (Arminger, 2008; Hopp, 2008). Therefore, it is important to find the balance between lost sales and inventory costs. Concluding, the problem is defined as follows:

*The existence of fluctuating demand with regard to promotions leads to either high inventory costs or lost sales for online retailers.*

1.2 Problem analysis

The defined problem of high inventory costs in relation to lost sales during promotions is a result of inaccurate promotion forecasting and inventory replenishment. In this section, the importance of promotion forecasting and inventory replenishment to respond to fluctuating demand is explained. Then, an overview of the potential reasons for inaccurate promotion forecasting and inventory replenishment is provided.
According to Ettouzani al. (2012), accurate forecasts and replenishments are critical to the success of a promotion. The service value of a SKU can only be created through a sale if the product is available. The availability of sufficient, but not excessive, stock requires knowledge of the rate at which the SKU will sell, to ensure that deliveries arrive in time. Furthermore, replenishing the SKU’s is vital to OSA. This makes promotion forecasting and inventory replenishment important in order to respond to the fluctuating demand during promotions.

Thomopoulos (2015) describes that it is important to have a reliable forecast, which results in better outcomes and decisions. A more accurate forecast can better respond to uncertain situations. The common goal is to minimize the cost of buying and holding the inventory, such that the customer’s demand for its SKU’s can be satisfied. There are multiple potential reasons which explain the difficulty of accurate promotion forecasting. First, demand during promotions is more subjected to fluctuations than non-promotional demand. Especially price promotions, which are short-term incentives tactically designed to push customers to purchase more, result in demand fluctuations (Cox &Brittain, 2000; Gilbert & Jackaria, 2002). Second, accurate promotion forecasting depends on many factors. A forecast should include particular independent variables, which influence the buying behavior of people (van Donselaar et al., 2006). Third, many retailers have problems with the availability of data. Past promotional data and information about potential independent variables are hardly available. This makes it difficult to make an accurate forecast (Fildes & Goodwin, 2007). Finally, promotion forecasting is often a subjective activity; many organizations rely too heavily on unstructured judgment instead of on statistical methods (Fildes & Goodwin, 2007). Overconfidence is a common problem for forecasting. Judgmental overconfidence leads to neglecting decision aids, making predictions contrary to the base rate and groupthink (Armstrong, 2001).

In retail companies inventory is held in anticipation of future demand. It is important to have an optimal level of product availability, such that the customer demand can be served on time without unnecessary high inventory (Chopra & Meindl, 2010). There are many potential reasons which explain the difficulty of inventory replenishment during promotions. Firstly, demand fluctuations during promotions make the bullwhip effect more pronounced and hence inventory management more difficult (O’Donnell et al., 2006). It is challenging for many retailers to guarantee product availability during promotions. Promotional activity has been observed as one of the root causes of poor OSA performance (Corsten & Gruen, 2003; Taylor & Fawcett, 2001). Secondly, promotional SKU’s have a temporary uplift in the sales, this makes it important to minimize the number of leftovers (Blattberg, 1995). Finally, literature about inventory policies specific for promotions is limited. Most researchers assume that promotions arise from leftovers (Zhang et al. 2008; Cooper et al. 1999).

1.3 Research scope
In this section, the scope and assumptions of this research are given. This research will focus on the promotional demand for online retailers in a business to consumer market. The forecasting of promotional demand and the replenishment strategy of products in promotion will be discussed during this research. Here, the replenishment strategy is restricted by the minimum lead time and review period, which are both dependent on the supplier and the internal operations. During this research, it is assumed that promotions are planned beforehand and thus the decision to promote does not depend on the level of inventory. Furthermore, the assumption is made that during a campaign no changes and cancellations can be made and that extra advertisement to uplift the sales is not possible. In Appendix 1 an overview of the assumptions made for this research given.
This research is applied in a case setting at bol.com based on promotional sales data provided in the Retail Promotion Tooling (RPG). Bol.com and the RPG will be discussed in Chapter 2. It is assumed that the sales provided in the RPG are equal to the demand in that period.

1.4 Research questions
In this section, the research question and sub-questions are discussed. This research aims to discover methods to accurately respond to fluctuating demand with regard to promotions by the use of a promotion forecasting model and inventory policy. The main research question is formulated as follows:

*How can an online retailing company accurately respond to fluctuating demand with regard to promotions by the use of a promotion forecasting model and inventory policy?*

This research starts with an analysis of the current performance, where after the report is divided into two phases. First, the promotion forecasting process is analyzed and improved, followed by the inventory policy. To support the research question, the following sub-questions are formulated:

1. **What is the current performance regarding promotion forecasting and inventory replenishment during promotions?**
   In order to come up with a forecasting and replenishment strategy for promotions, it is important to first identify the current processes regarding these tasks. It needs to be clear how the promotion forecasting and replenishment are determined in order to fully understand the problem and find out how these performances can be improved.

2. **Which factors are relevant for promotion forecasting?**
   The demand during a promotion is dependent on different factors, such as the number of SKU’s and the advertisement. To make an accurate promotion forecast it is important to identify these factors from the literature and interviews with relevant employees.

3. **How can the current promotion forecasting model be improved?**
   With the found factors, the current promotion forecasting model can be evaluated and improved.

4. **How should the company modify its inventory replenishment policy during promotions based on the founded forecast?**
   In order to find the best inventory policy during promotions the improved forecast is used. Different scenarios for the promotion demand pattern are taken into account.

1.5 Research methodology
In this chapter, the methodology for this master thesis will be discussed. First, the regulative cycle of van Strien (1997) will be discussed. Then a general approach for this project is given, based on the relevant parts of the regulative cycle.

1.5.1 Regulative cycle
According to van Strien (1997), this research can be defined as a Business Problem Solving (BPS) project. BPS projects aim to improve the performance of a business system, department or a company on one or more criteria. The basis is a set of problems from which a problem definition will be created. The classic problem-solving cycle as elaborated in the regulative cycle can be followed, given in Figure 1. This cycle has three parts: a design part (which consists of the problem definition, analysis and diagnosis and plan of action), a change part (which consists of the intervention) and a learning part (which consists of the evaluation). For this research the bordered part of this regulative cycle is used,
this is given within the square. The steps intervention and evaluation are not within the scope of this research.

![The regulative cycle](image)

**Figure 1: The regulative cycle**

### 1.5.2 General approach

This research will focus on the three steps of the regulative cycle, as can be seen in the square in Figure 1. First, the problem is determined and defined. Then, the problem is analyzed using qualitative and quantitative methods to get specific knowledge on the context and nature of the problem. The qualitative part includes interviews with relevant employees of the company and the quantitative part is the data validation. These will be used to evaluate the current performance. The second step is the analytic step which includes the analysis and diagnosis. To identify the main problem, the causes and consequences are extracted by the use of information gathering. Finally, a solution for the problem and the associated change plan is designed. To solve the business problem the existing literature is used to get a range of solution concepts. Out of this range an appropriate one is chosen, whereupon a specific variant is designed. This variant is adapted to the specific problem and its context. The action plan consists of a solution design to forecast the demand and replenish the inventory during promotions, through a model.

### 1.6 Outline

This report continues by introducing the case setting of bol.com in Chapter 2 and verifying the problem for the company. Furthermore, Chapter 2 identifies the current performance regarding forecasting and inventory replenishment during promotions to answer research question 1. Chapter 3 is about promotion forecasting and describes the important factors for promotion forecasting based on the literature and interviews with employees, answering question 2. In this chapter the optimal promotion forecasting model is determined, addressing question 3. In Chapter 4 the inventory policy during promotions is discussed, an inventory model is provided and tested with actual sales data, referring to question 4. Finally, Chapter 5 contains the main conclusions and recommendations of this report.
2 Case setting online retailer

The research is applied at bol.com, a company operating as an online retailer. Online retailing is a specific kind of retailing, where consumers react differently to stock-outs compared to offline retailing. Consumers have no switching costs, increased options and immediately switching capabilities when shopping (Rigby, 2011; Brynjolfsson et al., 2013). Gruen et al. (2002) found that the biggest loss of a retailer is when a shopper decides to buy the product which is out-of-stock (OOS) from another retailer. To retain the customers it is important to have enough inventory available (Emmelhainz et al., 1991).

In this chapter, the case setting at bol.com is presented, with an introduction of the company in Section 2.1. Then the promotion process of bol.com is explained in Section 2.2, followed by the current promotion forecasting and inventory policy in Section 2.3. Furthermore, the specific problem at the company is explained with an explanation of the chosen product groups. Finally, the available data for this research is explained and examined.

2.1 Company background

This section provides more information about the company, bol.com. In 1999 bol.com (Bertelsmann On-Line) was established by the German media group Bertelsmann AG as an online bookstore. Currently bol.com has become a big online retailer in the Netherlands and Belgium that sells a wide range of products with more than 10 million products in eighteen categories: Digital Reading, Educational Books & Supplies, Entertainment, General Books, Computer & Games, Domestic Appliances, Mobile & Tablets, Sound & Vision, Home Furnishing, Home Improvement & Gardening, Toys, Baby, Beauty & Care, Health & Intimacy, Jewellery, Watches & Accessories, Pet, Sport & Leisure and Cooking, Dining & Houseware. Bol.com provides three main activities via their website. First, their own products are sold online to the customer (B2C). Second, bol.com provides a platform for individuals to sell their second-hand products (C2C). Finally, external business vendors can sell their products on the platform of bol.com (B2C), this concept is called ‘Plaza’.

The company consists of five main departments: Operations & Fulfillment (O&F), Marketing & Platform Innovation, Finance & Support (F&S), Merchandising & Sourcing (M&S) and IT. In Figure 2 the structure of bol.com is given. These main departments consist of multiple other departments. For example, the department Supply Chain Management (SCM), where this research is conducted, is part of the Merchandising & Sourcing department. The M&S department is responsible for the operational and commercial activities in the retail process. The Supply Chain Management Department manages the process from supplier to the warehouse in order to make this process as efficient and effective as possible. The activities of the SCM department include inventory management, supplier management and order management (bol.com, 2017). In Appendix 2 an overview of the relevant job descriptions for this project is given.

![Figure 2: Structure of bol.com](image-url)
2.2 Promotions

Bol.com makes use of a specific strategy concerning promotions. In this section, the different kind of promotions are explained. The goal of promotions is to generate traffic on the website of bol.com, such that products will be bought and profit can be made. First, a distinction must be made between campaigns, promotions and actions. A campaign is the overall theme of multiple actions supported by a fully-integrated marketing mix. An action is a combination of multiple promotions with similar products and the same kind of discount, every campaign includes multiple actions. Promotion is the temporarily extra-highlighting of the assortment with the objective of tempting customers to do more favorable purchases for bol.com. A distinction between two types of promotions can be made: discount promotions and inspiration promotion. This last type is used to get more attention for a specific product group, without discount. This research is focused on the discount promotions.

2.2.1 Promotion process

The promotion process starts with the buyer, who decides together with the supplier which products will be discounted and when. Based on the year planning of the supplier the campaigns of bol.com are planned. The brand and product specialist make a year planning of all the campaigns, this consists of the yearly campaigns of bol.com combined with the agreed campaigns with the supplier. To cover the costs of a promotion the supplier gives a discount to bol.com. In addition to the year planning, new promotions can arise during the year based on deals with the supplier about the price.

Before the start of a campaign, the buyer makes a forecast of the expected demand based on historical data with an invented uplift of the sales. The supply chain specialist converts this forecast to a daily forecast, taking into account the lead time and review period. Most supply chain specialists make sure that two weeks before the campaign starts 70% of the products are available at the warehouse. When a campaign takes only one week or the review period is 7 days or higher, this percentage will be 100 instead of 70.

During a promotion, the supply chain specialist continuously monitors the sales and orders more products if necessary. The order quantity is based on human judgment. If the sales are lower than expected extra actions can be taken, such as a cancellation of the promotion or extra advertisement. However, as described before, this will not be considered during this research. At the end of a campaign the forecast and sales are evaluated. Figure 3 gives an overview of the promotion process at bol.com.
Figure 3: Promotions process at bol.com
2.3 Demand forecasting and inventory policy

In this section, the current demand forecasting and inventory policy of bol.com are explained. First, the inventory replenishment system and forecasting process are explained, followed by the inventory policy.

Slim4, an inventory replenishment system developed by Slimstock B.V., is used for the demand forecasting and inventory replenishment on product level. This system gives an order advice to the supply chain specialist to fulfill the demand. Slim4 is based on exponential smoothing and provides monthly forecasts, which are proportionally divided over the weeks. However, this system does not include promotions. If the sales during a month are four times higher than the other months, Slim4 assumes this month includes a promotion and therefore excludes this month from the forecast. After 2 years of sales data, a seasonal pattern is added to the forecast, by using an F-test. The F-test identifies seasonality in time series by comparing the periods variance with the residual variance (Thomas & Wallis, 1971). A seasonal pattern can also be added manually.

Furthermore, the products are categorized according to the ABC classification. This classification is based on the Pareto-analysis, with demand value and demand volume as the most common ranking criteria (Teunter, Babai, & Syntetos, 2010). The AA-label products are the fast moving items of each product group, these products are responsible for 25% of the actual sales. The A-label products are responsible for the next 25% of the sales and the B-label products for the following 25% of the sales. These three labels have a service level of 96%. C-label products are the slow moving items, which are responsible for the last 25% of the sales. The C-label products have a service level of 85%. The ordering advice of Slim4 is based on the forecast and the safety stock. The safety stock is dependent on the service levels. In Table 2, the target service levels per class are shown.

<table>
<thead>
<tr>
<th>ABC classification</th>
<th>Target service level</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>96%</td>
</tr>
<tr>
<td>A</td>
<td>96%</td>
</tr>
<tr>
<td>B</td>
<td>96%</td>
</tr>
<tr>
<td>C</td>
<td>85%</td>
</tr>
</tbody>
</table>

*Table 2: Target service levels per class*

Promotion forecasting is a manual, time consuming and subjective activity accomplished by employees of the sales department of bol.com. The forecasts are made by human support without a mathematical model. Important factors such as the kind of advertisement and competition are not available and therefore not included. This results in optimistic and inaccurate forecasts. The Retail Promoting Tool (RPG) is a system that receives and stores all the promotional sales.

2.3.1 Inventory policy

Bol.com makes use of the (R,S,nQ) inventory policy, where R is defined as the review period. An order is placed when the inventory position is below a certain S defined as the reorder level, then an ordering amount of nQ is placed where Q is defined as case pack size order (Silver, Pyke & Peterson, 1998). When Q is equal to 1, the (R,S) inventory policy applies. The review period and lead-time are fixed for every supplier. Slim4 calculates the replenishment amount based on the inventory position. The inventory position is defined as the inventory on hand added to the inventory on order (Silver, Pyke & Peterson, 1998). Demand from the customer that cannot be met is lost. If the inventory on hand is zero the offer on the website automatically changes into an offer of an external business vendor if available, otherwise the sale is lost. However, as described before, the external business vendors will not be considered during this research.
Bol.com makes use of a different inventory policy for promotions. This policy is a combination of the alpha-policy, where the items are distributed in two waves, and the \((R,s,nQ)\) inventory policy. First, before the start of a promotion using the alpha policy a specific amount of products from the forecast is pushed to the webshop (Van Donselaar et al., 2005). This amount is based on human judgment and dependent on the duration of the campaign, it is often equal to 70%. During a promotion, the supply chain specialist continually monitors the sales manually and order more products if necessary, based on human judgment without a mathematical model.

Based on interviews it is assumed that the size of the warehouse and inbound planning does not influence the inventory policy of bol.com. Furthermore, based on interviews it is assumed that there are no extra costs for the frequency and number of products that will be ordered.

2.4 Product groups

In this section, the product groups are introduced. Due to the limited size of the data set only two product groups could be included in this research. Based on interviews with relevant employees of bol.com the selection for the product groups Personal Care & Home Care and Televisions was made. In Appendix 3 an overview of the product hierarchy used at bol.com is given.

Personal Care & Home Care is part of the Fast Moving Consumer Goods market. This market is characterized by strong sales peaks caused by promotions in which retailers want to draw attention to the customer (Groot & Musters, 2005). These products have a short rotation time and are heavily promoted. The customer expects a greater frequency of promotions, this results in a reduced consumer reference price (Blattberg, Briesch & Fox, 1995). To get revenue on these fast movers it is important to promote these products very often. Therefore, the strategy of bol.com is to promote these products from the Financial Category (FinCat) Beauty & Care constantly. Currently, 80% of the sales in the Personal & Home Care unit consist of sales from promotions (bol.com, 2017). Earlier internal research at bol.com showed that it is more profitable to include a broad assortment in a campaign, then the demand of every customer can be fulfilled. Therefore, substitution from non-promotional products to products that are on promotion will not appear in this research. Lastly, the FinCat Beauty & Care is unknown for many customers of bol.com, therefore most promotions are shown on the homepage of the website. In Appendix 3 the product hierarchy of this product group is given.

In contrast, Televisions is a highly valued product group with a small number of products in a campaign. Therefore, excess stock is very costly and causes obsolete inventory. Furthermore, the products have a high price elasticity. The goal is to minimize the number of lost sales, excess stock and inventory obsolescence. In Appendix 3, the product hierarchy of televisions is given. Furthermore, the electronics market is a very competitive market, wherein bol.com has a small share. The promotion strategy for televisions is to create traffic to the webshop. Every retailer matches their price automatically to their competitors with web spidering. Therefore, a promotion is never unique. Bol.com automatically matches their price to the main competitors, such as Coolblue and Mediamarkt. If a competitor maintains a price which is not profitable, bol.com asks a fee from the supplier. Otherwise, the loss of the product is compensated by the sales of other products. Most promotions are introduced by the supplier, the main suppliers are Samsung, Sony, LG and Philips. Every retailer provides these promotions at the same time. Therefore, a promotion includes mainly one brand with Televisions from all sizes to fulfill the demand of every customer. Televisions have a relative long purchase cycle, which results in less brand switching caused by promotions (Narasimhan et al., 1996). This makes substitution not important for this research. To get customers it is important to do
advertisement, bol.com makes for example use of google shopping and external banners to create traffic to the shop. Furthermore, bol.com has their own general campaigns. During this time more traffic is already on the website which brings the product group Televisions more in front compared to competitors. Then bol.com will exclude their products from price comparison sites to make price matching more difficult.

2.5 Data analysis current situation

In this section, the available data for this research is explained. First, the data gathering and filtering are explained. Then, an explanation of the reference period is given, followed by a review of the current performance regarding promotions.

The data for this research is gathered from the RPG and consists of all promotions per day for each product of the chosen product groups in the period October 15 in 2016 until June 19 in 2017. The promotions from March 2017 are missing due to an error in the data. The available data includes 404172 observations, with 29 campaigns and 46 daydeals. However, the RPG is occasionally used for product adjustments unrelated to promotions. This is improper use of the RPG. Subsequently, before analyzing the data, it first needs to be filtered. In Table 3 an overview of the types of data that are deleted is given with the corresponding reason. Finally, the remaining data consists of 57076 observations with 12 campaigns and 38 daydeals. Due to the broad promotional assortment of Personal Care & Home Care, the remaining dataset entails mainly Personal Care & Home Care products, namely 97% of the available observations.

<table>
<thead>
<tr>
<th>Types of data</th>
<th>Reason</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duplicated data</td>
<td>Duplicated data is removed.</td>
</tr>
<tr>
<td>Duration &gt; 90 days</td>
<td>The maximum number of days of a promotion is 90 days. Promotions that last longer than 90 days are deleted.</td>
</tr>
<tr>
<td>Other campaigns</td>
<td>‘Other campaigns’ are the remaining campaigns which cannot be classified into a general campaign of bol.com. These campaigns include non-repetitive promotions and non-promotion related product adjustments. Due to the non-repetitive and irregular nature of these ‘other campaigns’, they are deleted from the dataset.</td>
</tr>
<tr>
<td>Outlet</td>
<td>Outlet consists of outdated products, which will not be ordered again. These type of products are excluded from the scope.</td>
</tr>
<tr>
<td>New products</td>
<td>New products are excluded from the scope, since it is not possible to measure the uplift in sales. Products which have no sales during the reference period and do have sales during the promotion are classified as new products.</td>
</tr>
<tr>
<td>Action type: Other discount promotion</td>
<td>These actions are awareness promotions and are not part of the scope.</td>
</tr>
<tr>
<td>Plaza partners in promotion period</td>
<td>External business vendors (Plaza partners) are not part of the scope and therefore classified as OOS. An assumption is made that the reference period is representative and when there are plaza partners in the reference period these are taken into account.</td>
</tr>
<tr>
<td>The last action is taken</td>
<td>The data consists of promotions with multiple start and end dates, caused by adaptations in the RPG. Since each promotion is unique there is chosen to take the promotions with the latest dates. An assumption is made that these are the revised dates.</td>
</tr>
<tr>
<td>Not available products</td>
<td>Promotions with products which are not available at the supplier during the promotion are deleted. Products with no sales during the promotion and no stock data are classified as not available products.</td>
</tr>
</tbody>
</table>

Table 3: Overview of data filtering
Furthermore, the available data from the RPG contains a reference period. This reference period has three characteristics. First, the reference period has exactly the same number of days as the duration of the promotion. Second, it starts on the same day of the week as the promotion starts. Third, the reference period is prior to the promotion and both periods do not interfere with each other. However, this reference period frequently contains another promotion. Figure 4 gives an example of two promotions with their reference period. This figure shows that the reference period of ‘Moederdag’ is during the promotion period of ‘Elektronica Deals’. Due to limited data, the assumption is made that the reference period is a representative for the non-promotional demand.

![Figure 4: Example of the reference period of a promotion](image)

2.5.1 Current performance

In this section the current performance regarding promotions for Personal Care & Home Care and Television is reviewed, using newly introduced Key Performance Indicators (KPI’s). This analysis is based on the available data. At this moment, bol.com has no KPI’s regarding promotions. Therefore, the following KPI’s are introduced:

- **Mean service level**: the mean probability of no stockout per replenishment cycle (Silver, Pyke & Peterson, 1998). The mean service level is calculated by dividing the number of days a product is not OOS by the total number of days.
- **Sale**: the percentage of the products that are actually sold during the promotion. The sale is calculated by dividing the number of products that are not sold during the promotion by the total number of products. This KPI is calculated by 1 minus the number of products with zero sales during a promotion divided by the total amount of products.
- **Uplift**: the extra consumer sales generated during the promotion compared to the available reference period. This KPI is calculated by dividing the total consumer sales of the promotions by the total consumer sales in the reference period.
Table 4 gives an overview of the current performance of bol.com regarding promotions using the KPI’s. The second column, mentioned as ‘Both Product Groups’, includes the performance of all available data. The data consists mainly of products from the product group Personal Care & Home Care. To review the performance of each of the product groups a separate analysis is added to the table. First, the mean service level is given. Here can be seen that 98.6% of the days the products were available for the customer. This percentage is lower for Personal Care & Home Care. This can be explained by the broad assortment and the excess stock for televisions. The promotion forecast for Televisions is often overconfident and results in bad forecast accuracies. For example, the forecast error of one campaign has resulted in a Mean Absolute Percentage Error of 153% for Televisions and 91.9% for Personal Care & Home Care. The forecast of Televisions gave a too high forecast for every product in the campaign. This resulted in the following mean inventory levels per product, given in Figure 5. This figure shows a higher inventory level for Televisions compared to Personal Care & Home Care. However, Personal Care & Home Care contains FMCG products, with a shorter rotation time. This implies that Personal Care & Home Care needs more products on inventory compared to Televisions. Due to a poor connection between the product number from the RPG and the other data systems of bol.com, it was not possible to determine the current inventory levels from all promotions.

![Mean inventory per product, Bulk May](image)

The second KPI is the sale. The column ‘Personal Care & Home Care’ shows that 10.5% of the products are not sold during the whole promotion period. This can be explained by the broad assortment of Personal Care & Home Care, which consists of many slow moving products. Figure 6 gives an overview of the number of products in a campaign for the two product groups. This figure shows that Personal Care & Home Care includes a broad assortment in the campaigns compared to the product group Televisions. Furthermore, the product group Personal Care & Home Care is included in more campaigns compared to Televisions. The last KPI is the uplift. Table 4 shows that Personal Care & Home Care has a higher uplift and is more dependent on promotions compared to Televisions.
In conclusion, Personal Care & Home Care has a broad assortment and is included in more campaigns compared to Televisions. This results in more unsold products during a promotion, which leads to a lower score on the Sale KPI. Furthermore, the uplift of Personal Care & Home Care is higher compared to Televisions, caused by the strong sales peaks during promotions for FMCG products. On the other hand, the promotion forecast for Televisions is often overconfident, which results in a high score on the mean service level, but also in excess of stock. Due to a poor connection between the product number from the RPG and the other data systems of bol.com, it was not possible to determine the current inventory levels from all promotions. Therefore, it was not possible to validate a KPI based on the excess of stock.
3. Promotion Forecasting

In this section, the promotion forecasting model is introduced. The literature about promotion forecasting is discussed, followed by the important factors according to the literature and interviews. Then, the split of the data and the detection of outliers is described, followed by the performance measures. Last, multiple regression analysis is used to find the best promotion forecasting models in different situations, these models are explained and validated.

3.1 Multiple regression

Mathematical forecasting techniques are divided into two types: time series techniques and explanatory models (Makridakis, 1988). Cooper et al. (1999) say that extrapolation of time series is not appropriate for demand forecasting for promotions. Demand forecasting should be based on the impact of promotions and advertising on customer buying and/or marketing intelligence and should take price-elasticities into account. An explanatory model takes care of the causal relationship between the dependent and independent variables, where it predicts the dependent variable based on the independent variables (Makridakis, 1988). An important explanatory method is multiple regression analysis. Multiple regression is a statistical technique which can be used to analyze this causal relationship. The regression consists of multiple independent variables, whose values predict the dependent variable. Each independent variable is weighted by the regression analysis procedure to ensure maximal prediction from the set of independent variables. These weights denote the relative contribution of the independent variables to the overall prediction. The set of weighted independent variables forms the regression variate, which leads to a linear combination of the independent variables to predict the dependent variable (Hair et al. 2014). Multiple regression analysis is a simple, understandable technique and can therefore easily be used in practice at bol.com. This method is chosen as the promotion forecasting method.

The most frequently used approach to promotional analysis compares the average baseline sales for a category to the lift, which is the multiple of baseline sales expected when any item in a category is promoted in a particular way. This method is called the base-times-lift. One model based on this base-times-lift approach is the model by Falck & Småros (2005). This model uses a promotion coefficient to increase the forecast in a promotion period from its baseline sales level. The increase in the sales can be captured using the sales uplift, also called lift factor. This is defined as the ratio of the sales volume during a promotional event compared to the baseline sales, given in Formula 1 (van Donselaar et al., 2016).

\[
\text{lift factor}_x = \frac{\text{Units sold of product } x \text{ during the promotion}}{\text{Baseline sales of product } x \text{ in units}}
\] (1)

3.2 Variables forecasting model

3.2.1 Dependent variable

To measure the success of a promotion the lift factor is used as dependent variable. To enable comparison of different SKU’s, the lift factor is based on the relative sales, given in Formula 1. The baseline sales are based on the current forecasting system of bol.com, slim4. Therefore, only the lift factor needs to be determined to find a forecast for the SKU during a promotion.
3.2.2 Independent variable

The sales uplift during a promotion depends on different factors. The relevant factors are identified based on the literature and interviews with relevant employees of bol.com.

3.2.2.1 Literature

There are two perspectives, when it comes to understanding what information is relevant to any promotional event. The first perspective is the promotional mix. Different results for different combinations of price-cuts, ads and displays are expected regardless of what item is involved. These combinations influence the buying behavior of the customer. Therefore, each of these combinations must be separately included as an independent variable in the promotion forecasting model. The second perspective takes the item itself as a base point; how well did the item perform during promotions in the past, by looking at the historic baseline sales. With these perspectives, the forecasting variables in the regression model can be found (Cooper et al., 1999). Table 5 lists important independent variables found in the literature.

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ability to stockpile</td>
<td>The ability to stock-pile is the degree to which a product category is easy to stockpile. This relates directly to the mechanism of purchase acceleration. Easy to stockpile products have a higher promotional elasticity (Narasimhan et al., 1996).</td>
</tr>
<tr>
<td>Category factors</td>
<td>Van Donselaar et al. (2005) describe that category factors should be considered as an exploratory variable. Every product is part of a general category, where complementary products co-exist.</td>
</tr>
<tr>
<td>Competition</td>
<td>Competitive marketing activities such as price competition are also important for the product sales. The buying behavior of a consumer is influenced by price competition. One reason can be that the sales of a retailer is below average if for example, a competitor has a specific item on discount (De Leeuw, 1996).</td>
</tr>
<tr>
<td>Duration</td>
<td>According to Cooper et al. (1999) is duration another independent variable for promotion forecasting. In their study, they found that longer promotions (three or four weeks) will sell on average fewer units in total than promotions that take two weeks (these include fast movers).</td>
</tr>
<tr>
<td>Effects of advertisement</td>
<td>Research of Ailawadi et al. (2001) investigated that there are two types of advertisement namely in-store and out-of-store promotion. The first type includes displays and in-store specials, these are encountered in the store to attract customers to promoted articles and customers buy the products often impulsively. For online retailers, this includes advertisement on the website. Out-store advertisement is on the other hand actively considered before the consumer goes shopping, this can include advertisement in newspapers, television, coupons and store flyers.</td>
</tr>
<tr>
<td>Holidays</td>
<td>Cooper et al. (1999) included dummy-variables for the main holidays or also mentioned as special days in their regression model. These special seasonal influences highlight specific weeks when special events tend to boost sales, like Easter.</td>
</tr>
<tr>
<td>Interpurchase time</td>
<td>Narasimhan et al. (1996) also found that shorter interpurchase times, i.e. shorter purchase cycles, result in more brand switching caused by promotions. The customer makes the trade-off between lower costs against substitution costs of buying the less preferred brand. With a shorter interpurchase time the impact of buying another brand is not that severe, therefore the customer will switch easier.</td>
</tr>
</tbody>
</table>
The number of items in a category that are on promotion is another important independent variable. If a large percentage of items in a category is on promotion at the same time, consumers will perceive that they have more choice and are more likely to switch to that particular store (Ailawadi et al., 2006).

Ailawadi et al. (2006) defined promotion frequency as the % of weeks in the year the brand or category is on promotion. Consumers expect to find frequently promoted brands often on promotion and therefore adapt their purchase behavior to these finding.

Blattberg et al. (1995) and van Heerde et al. (2002) show that the relative price discount is important. Temporary retail price reductions increase sales substantially.

The weather influence the overall sales and the specific products that will be sold (De Leeuw, 1996). When it is rainy people will rather stay inside and buy their products online compared to a sunny day where people go outside for shopping (Chiang & Dholakia, 2003).

### Table 5: Overview of important independent variables for promotion forecasting

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of items in a category</td>
<td>The number of items in a category that are on promotion is another important independent variable. If a large percentage of items in a category is on promotion at the same time, consumers will perceive that they have more choice and are more likely to switch to that particular store (Ailawadi et al., 2006).</td>
</tr>
<tr>
<td>Promotion frequency</td>
<td>Ailawadi et al. (2006) defined promotion frequency as the % of weeks in the year the brand or category is on promotion. Consumers expect to find frequently promoted brands often on promotion and therefore adapt their purchase behavior to these finding.</td>
</tr>
<tr>
<td>Relative price discount</td>
<td>Blattberg et al. (1995) and van Heerde et al. (2002) show that the relative price discount is important. Temporary retail price reductions increase sales substantially.</td>
</tr>
<tr>
<td>Weather</td>
<td>The weather influence the overall sales and the specific products that will be sold (De Leeuw, 1996). When it is rainy people will rather stay inside and buy their products online compared to a sunny day where people go outside for shopping (Chiang &amp; Dholakia, 2003).</td>
</tr>
</tbody>
</table>

### 3.2.2.2 Interviews

In addition to the literature, the important factors for promotion forecasting are identified according to semi-structured interviews with relevant employees. Semi-structured interviews include predetermined questions, but also offer participants the chance to explore issues they feel are important (Longhurst, 2003). Appendix 4 includes an overview of the 19 respondents. First, relevant employees are interviewed to identify more important variables for bol.com. Then, validation interviews were performed to evaluate the independent variables from the literature and the first interviews. In Table 6 an overview of the results of the interviews is given.
### 3.2.3 Final factors

In this section, the final factors for the promotion forecasting model are selected. Table 7 lists the final independent variables and dependent variable. In Appendix 5, the final variables are explained in detail. Due to limited data and other reasons, given in Table 8, not all factors could be included.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Description</th>
<th>Variable Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>$y$</td>
<td>Lift factor</td>
<td>Metric</td>
</tr>
<tr>
<td>$x_1$</td>
<td>Product subgroups</td>
<td>Nonmetric</td>
</tr>
<tr>
<td>$x_2$</td>
<td>Discount</td>
<td>Metric</td>
</tr>
<tr>
<td>$x_3$</td>
<td>Duration promotion</td>
<td>Metric</td>
</tr>
<tr>
<td>$x_4$</td>
<td>Type campaign</td>
<td>Nonmetric</td>
</tr>
<tr>
<td>$x_5$</td>
<td>Action discount type</td>
<td>Nonmetric</td>
</tr>
<tr>
<td>$x_6$</td>
<td>Number of products in action</td>
<td>Metric</td>
</tr>
</tbody>
</table>

Table 7: Overview of the final factors

The factors ability to stockpile, effects of advertisement, google analytics, interpurchase time, list price, price competition, the promotion frequency and the visitors of the website were omitted because no data was available about these factors. The ability to stockpile and the interpurchase time are related to the product subgroup. Each product subgroup includes the same kind of products. No
further information about the ability to stockpile and interpurchase time was available. Furthermore, due to the improper use of the RPG the available data includes many non-promotion related product adjustments. It was not possible to distinguish the non-promotion and promotional data, which makes it difficult to determine the promotion frequency.

In addition, the factors period of the year and holidays were left out of the promotion forecasting model because they are already taken into account in the baseline forecast. This baseline forecast already includes seasonality. Furthermore, the weather is not included due to the time range of the promotion process. Six weeks before the start of the promotion the forecast needs to be complete, at that moment no detailed information about the weather is available. Finally, product group is not included in the model. This variable correlates too heavily with product subgroup. Product subgroup has the highest correlation with the lift factor and is therefore included instead of product group. A summary of the not included factors is provided in Table 8.

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>Reason</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ability to stockpile</td>
<td>Limited data</td>
</tr>
<tr>
<td>Competition</td>
<td>Limited data</td>
</tr>
<tr>
<td>Effects of advertisement</td>
<td>Limited data</td>
</tr>
<tr>
<td>Holidays</td>
<td>In normal forecast and in campaigns</td>
</tr>
<tr>
<td>Interpurchase time</td>
<td>Limited data</td>
</tr>
<tr>
<td>List price (advice selling price from the supplier)</td>
<td>Limited data</td>
</tr>
<tr>
<td>Period of the year</td>
<td>Seasons in normal forecast</td>
</tr>
<tr>
<td>Product group</td>
<td>Heavily correlated with product subgroup</td>
</tr>
<tr>
<td>Promotion frequency</td>
<td>Limited data</td>
</tr>
<tr>
<td>Search results on Google</td>
<td>Limited data</td>
</tr>
<tr>
<td>Visitors of the website</td>
<td>Limited data</td>
</tr>
<tr>
<td>Weather</td>
<td>Time range</td>
</tr>
</tbody>
</table>

*Table 8: Summary of the not included factors*

3.3 Data preparation
Before continuing to the data analysis, the filtered data introduced in Section 2.5 must be prepared. In this section, data preparation is described, including data split and outlier detection.

3.3.1 Data split
First, the data is split into a train and validation set. The train set is used to develop the forecasting models. The validation set is used for the validation of the results. In total 57076 observations are available. A random sample split of 70-30 is applied, taking into account that each category is represented in both samples. To avoid overfitting, 20 observations per estimated parameter must be available. The two datasets meet these conditions (Hair et al. 2014).

3.3.2 Outliers
Outliers are observations that deviate strongly from the other observations. These can influence the outcome of the model and therefore need to be detected (Field, 2009). It is unknown beforehand whether a promotion is an outlier. Therefore, the outliers are only deleted in the analysis sample. There are different kind of outlier detection, namely univariate, bivariate and multivariate detection. Univariate methods examine all metric variables to identify unique or extreme observations. Bivariate methods focus on the relation between two variables. Multivariate methods are suited for multiple of variables. A researcher should utilize as many of these perspectives as possible (Hair et al., 2006). Therefore, first univariate outlier detection using the standardized scores is applied. For large sample
sizes, outliers are defined as cases with an absolute standard score of 4 or higher. These cases are analyzed and deleted. In total, only 6 cases from the 39953 observations are deleted.

Second, multivariate detection is applied to measure the multidimensional position of each observation relative to a common point, this makes bivariate methods unnecessary. A method to check for multivariate outliers is the Cooks Distance. In Figure 7 a graph with the influential observations by the Cook’s distance is given. Cook and Weisberg (1982) suggested that values above 1 may be cause for concern. The potential outliers are investigated manually whether the cases are unusual combinations. One potential outlier is deleted since it has a very high lift factor in comparison with the discount and it has a Cook’s distance greater than 1. The other outliers are not influential and will be kept in the dataset.

3.4 Assumptions for multiple regression
To apply multivariate regression analysis, several assumptions must be true. It needs to be checked if the individual variables meet the following assumptions (Hair et al., 2014):

- Homoscedasticity: it is assumed that at each level of the independent variables, the variance of the residuals is constant.
- Normality: it is assumed that the residuals in the model are random, normally distributed with a mean of zero.
- Linearity: it is assumed that there is a linear relationship between the dependent and each of the independent variables.
- Independence of the error terms: it is assumed that the explanatory variables are independent of the error term.
- Multicollinearity: it is assumed that there is no perfect linear relationship between two or more predictors.

First, the relationship between the variables should be homoscedastic. If the dispersion of the dependent values is unequal across values of the independent variable, the relationship is mentioned as heteroscedastic. To test for homoscedasticity a scatterplot of the residuals against the values is made (Hair et al. 2014). In Appendix 6 the scatterplot is given, here a clear structure can be seen which is an indication of heteroscedasticity. Furthermore, the Breusch-Pagan test also shows that there is heteroscedasticity. This test is significant therefore the null-hypothesis that there is homoscedasticity
can be rejected. Heteroscedasticity can be remedied by data transformations of the dependent variable. In Appendix 6 can be seen that the lift factor is heavily left skewed, therefore data transformation using natural logarithm is necessary. After the data transformation, the scatterplot looks better. However, the Breusch-Pagan test is still significant, indicating that the homoscedasticity assumption was not met. Violation of homoscedasticity can lead to under- or overestimated standard errors. This can lead to too high or too low p-values. Therefore, the standard errors of the estimated regression parameters are corrected for the presence of heteroscedasticity. The adjustment of robust standard errors does not change the estimate of the coefficients, but it changes the (robust) standard error and the p-values (Hair et al. 2014).

Second, normality is the most fundamental assumption for multivariate analysis. Normality is required to use the F and t statistics. The severity of the nonnormality depends on the shape of the distribution and the sample size. When the sample size is large, larger than 30 observations, the normality of the distributed error term does not need to be checked. According to the central limit theorem, the sample mean of large samples tends to a normal distribution, regardless of the underlying distribution of the data. Furthermore, a goodness-of-fit test to test this assumption is not appropriate here. When having a large number of observations the null hypothesis of having a distribution will almost always be rejected. Therefore, a graphical representation is made to check the normality, given in Appendix 6. This Appendix shows that the residuals are normally distributed. Therefore, the normality assumption is met. If this assumption is not met, data transformations are needed (Hair et al. 2014).

Third, the concept of correlation is based on a linear relationship, therefore it is critical to have a linear relationship between the dependent and each of the independent variables. Linearity can be checked with partial regression plots, which show the relationship of a single independent variable to the dependent variable, controlling the effects of all other independent variables. Linearity only needs to be checked between two metric values. Therefore, the partial regression plots of the metric values are given in Appendix 6, these all show no curvilinear patterns which indicates that the assumption linearity is met. When the assumption is not met, corrective actions need to be taken. Examples of corrective action are transforming the data values of one or more independent variables, directly including the nonlinear relationship in the model or using specialized models such as nonlinear regression to accommodate this curvilinear effect of the independent variable (Hair et al. 2014).

Fourth, the last assumption for multivariate regression analysis is the absence of correlated errors. Field (2009) uses the Durbin-Watson test to check whether serial correlations between the errors occur. The closer the value is to 2 the better the value. The final model for this research gives a value of 1.9589. This is very close to 2, which means that the residuals are uncorrelated. Indicates that the assumption is met. When the assumption is not met, omitted causal factor need to be included in the multivariate analysis (Hair et al., 2014).

Finally, multicollinearity is a key issue in interpreting the regression variate. This includes the correlation between the independent variables. Multicollinearity is the extent to which a variable can be explained by the other variables. If this collinearity increases, the shared prediction will also increase and the unique variance of each independent variable decreases. To maximize the prediction of the independent variables, it is necessary to look at independent variables that have a high correlation with the dependent variable and low multicollinearity with other independent variables. Multicollinearity can be examined by the variance inflation factor (VIF). This value should be lower than 10 (Hair et al., 2014). The VIF values vary between 1.03 and 8.72, with an average VIF value of 4.62. This concludes that the multicollinearity assumption is satisfied.
3.5 Estimation technique

In this section, the estimation technique is explained. First, the different approaches are explained, where after one approach is chosen and further explained. Finally, the different models are explained.

There are three different approaches to construct a multiple regression. The first one is the simultaneous regression, this approach includes all the variables at the same time. Here the independent variables are exactly specified, the regression model is used for confirmation. The second approach is the sequential estimation, in which the independent variables must be examined. From the defined set of variables, the variables are selectively added or deleted based on their correlation with the outcome until the highest $R^2_{adj}$ is achieved. The third type of estimation technique is the combinational approach. The combinational approach combines multiple of search processes. This method got criticism of its a-theoretical nature and the lack of consideration of factors as multicollinearity, identification of outliers and influencers and interpretability of the results. Furthermore, there are many possible regressions which makes it a complex process (Hair et al., 2014). For this research the sequential approach is the most appropriate. With this method the added value of each independent variables is examined, in order to come up with a subset that maximizes the predictive accuracy (Hair et al., 2014).

There are two different sequential methods: the stepwise estimation and forward/backward estimation. The stepwise estimation is the most popular method and differs from the forward/backward method in the ability to add or delete variables at each stage. Once a variable is deleted in the forward or backward elimination method it cannot be reversed at a later stage (Hair et al., 2014). Finally, the stepwise method is applied.

3.5.1 Stepwise estimation

The stepwise estimation method starts by selecting the best metric predictor that has the highest simple correlation with the dependent variable. If this value significantly improves the ability of the model to predict the outcome, it will be added. The second predictor is the variable that explains the largest statistically significant portion of the unexplained variance remaining from the first equation. This is the variable with the largest semi-partial correlation with the outcome. More metric variables are added as long as their correlation coefficients are statistically significant (Hair et al., 2014). The non-metric explanatory variables with more than 2 groups are tested for joint significance using the Wald-test. This method helps to define the non-metric variables that help to explain the variation in the dependent variable. When the non-metric explanatory variable has 2 groups, the t-test is used (Hair et al., 2014). Finally, interaction effects between the different independent variables are analyzed. The interactions that improve the ability of the model to predict the independent variables are added.

A restriction of the stepwise method is that it is an automated approach with almost no control over the final model specification (Hair et al., 2014). Therefore, this method is combined with the trial and error method to see if important variables are missing or variables need to be deleted. The model with the highest $R^2_{adj}$ is assigned as the best model. The reference group for the dummy variables is chosen in order that the most significant dummy variables can be found.

3.5.2 Promotion forecasting models

Due to the data uncertainties in the RPG, five different models based on different datasets are created and examined, using the stepwise estimation method. These models predict the total lift factor for every SKU during the whole promotion period. The forecasts are made for the whole promotion period in order to correct for daily fluctuations. Previous procedures have been followed for all models in
order to confirm all the assumptions, given in Appendix 6. In this section, the differences between the five promotion forecasting models are explained.

The first model consists of all the available data from the RPG. This includes all the variables described in Appendix 5. The following four models are all based on this first model with an adjustment to the data. Model 2 uses the same data as model 1. Although, the campaign ‘Deals’ is deleted and the daydeals in this campaign are assigned to the corresponding general campaign. In the available data the daydeals are not part of a specific campaign, but are all assigned to the campaign ‘Deals’. A daydeal can take place during a general campaign or as a separate promotion. The mean lift factor for daydeals during a general campaign is 50.29 and the lift factors for separate daydeals is 37.23. Therefore, the assumption is made that a daydeal during a general campaign gets a higher lift factor due to the extra attention from the general campaign. In this model all daydeals that are not part of a campaign are deleted. In total, five daydeals are not included.

Model 3 corrects the available data for the revised promotion dates. As mentioned in Section 2.5, some promotions have multiple start and end dates caused by adaptations in the RPG. Therefore, it is hard to determine which promotion is correct. The assumption is made that the last action is the revised promotion and thus the correct promotion. The available data from model 1 consists of this revised promotion. However, the revised promotion sometimes lasts longer than the advertisement of bol.com. For example, a promotion that continues after Valentine’s Day has ended. To examine this assumption model 3 is created. This model contains only the dates during the general advertisement period of bol.com. Furthermore, this model also deletes the campaign ‘Deals’ and assign the daydeals to the general campaigns, which makes it easy to compare model 2 and 3.

Finally, model 4 and 5 differ from the first model by correcting the available data for the reference period. Due to limited data, an assumption is made that the available reference period is correct. However, this period often includes another campaign as shown in Figure 4. This makes the reference period unreliable. Therefore, two models with a corrected reference period are made. Model 4 is based on the last two days of the campaign ‘Bulk Mei 2017’, these are the only days that do not overlap with another campaign. The rest of the available data is not included. Model 5 consists of the same campaign and includes only the SKU’s that were not promoted in the reference period. For model 5, the assumption is made that the sales of the SKU’s in the reference period are not affected by the campaign in the reference period. An overview of the different models is given in Table 9.

<table>
<thead>
<tr>
<th>Model</th>
<th>Contains</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>All available data</td>
</tr>
<tr>
<td>Model 2</td>
<td>Delete campaign Deals and assign the daydeals to the other campaigns.</td>
</tr>
<tr>
<td>Model 3</td>
<td>Delete campaign Deals, assign the daydeals to the other campaigns and include only the dates during the general advertisement period of bol.com.</td>
</tr>
<tr>
<td>Model 4</td>
<td>Corrected for reference period using 2 days that do not overlap with another campaign.</td>
</tr>
<tr>
<td>Model 5</td>
<td>Corrected for reference period using only the SKU’s that were not promoted in the campaign in the reference period.</td>
</tr>
</tbody>
</table>

Table 9: Overview of the different models
3.6 Performance measures

It is important to evaluate the performance of the promotion forecasting models in an appropriate way. There are many different methods to assess the different regression models. For this research, three different evaluation techniques are used in order to compare the differences between the models. First, for the model selection the $R^2_{adj}$ is used, given in Formula 2. The $R^2_{adj}$ is useful for comparison between equations with different numbers of independent variables and different sample sizes. The higher the value of the $R^2_{adj}$ the greater the explanatory power of the regression equation (Hair et al. 2014). Below the formula for the $R^2_{adj}$ is given, with the number of observations $N$, the number of explanatory variables $p$ and the $R^2$. This $R^2$ is calculated by the explained variance divided by the total variance. This measures the proportion of the variance of the dependent variable about its mean that is explained by the independent variable (Hair et al., 2014).

$$R^2_{adj} = 1 - (1 - R^2) \frac{N-1}{N-p-1} \tag{2}$$

The second evaluation method is the MAPE, this method is currently used by bol.com. The Mean Absolute Percentage Error (MAPE) is the most widely used method according to Chockalingam (2009). It assumes that the absolute error on each item is equally important. The following formula shows how to calculate the MAPE:

$$MAPE = \frac{1}{N} \sum_{t=1}^{N} \frac{|Y_t - \hat{Y}_t|}{Y_t} \tag{3}$$

$\hat{Y}_t$ = Forecast demand value at time $t$

$Y_t$ = Actual sales demand value at time $t$

$N$ = Number of observations

The last method to assess the different forecasting methods is the RMSE. According to Sanders (1997), the most beneficial method to measure the forecast accuracy in situations when the cost function increases with the square of error is the MSE. The MSE is a forecasting error measurement method, in which the variability of the forecast errors will be estimated. Within the MSE method each residual is squared, therefore the larger forecast errors, the heavier the penalizing of the error. The RMSE is the square root of the MSE. This method is easier to understand and makes it easier to compare the results. The following formula shows how to calculate the RMSE:

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^{N} (\hat{Y}_t - Y_t)^2} \tag{4}$$

$\hat{Y}_t$ = Forecast demand value at time $t$

$Y_t$ = Actual sales demand value at time $t$

$N$ = Number of observations

3.7 Model results

In this section, the results of the different models are presented. First, a short overview of the results based on the train set is given. Second, the overall performance of the different models based on the validation set is described.
3.7.1 General performance

In this section, an overview of the results based on the train dataset is provided. The final variables of the models with the estimates and significance levels are given in Appendix 7. In this section, the differences between the variables and estimates in the models are discussed.

Table 10 gives an overview of the main characteristics of the different models. The $R^2_{\text{adjusted}}$ is used to compare the different models. As can be seen in Table 10, the number of explanatory variables differentiate between the first three models and model 4 and 5. This can be explained by the predictor product subgroup, which is not included in model 4 and 5, and the many interaction effects in the first three models. Due to the limited data of one campaign it was not possible to include the predictor product subgroup in model 4 and 5. Furthermore, the different sample sizes are also given in Table 10. These satisfy the minimum sample size of 50+8$k$, where $k$ is the number of predictors (Hair et al., 2014).  

<table>
<thead>
<tr>
<th>Model</th>
<th>$R^2_{\text{adjusted}}$</th>
<th># explanatory variables</th>
<th># observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>0.4531</td>
<td>11</td>
<td>2431</td>
</tr>
<tr>
<td>Model 2</td>
<td>0.4617</td>
<td>12</td>
<td>2426</td>
</tr>
<tr>
<td>Model 3</td>
<td>0.4464</td>
<td>14</td>
<td>2370</td>
</tr>
<tr>
<td>Model 4</td>
<td>0.3927</td>
<td>7</td>
<td>286</td>
</tr>
<tr>
<td>Model 5</td>
<td>0.3695</td>
<td>8</td>
<td>358</td>
</tr>
</tbody>
</table>

*Table 10: Summary of the main characteristics of the models*

3.7.2 Validation

In this section, the forecasting models will be validated according to the performance measures, based on the validation sample set. The validation sample set is used to test the robustness of the models. With these results, the performance of the models can be compared. The outliers are kept in the validation sample since it is unknown beforehand whether a promotion is an outlier. Finally, the preferred model is given. In Appendix 8, this model is further explained with the independent variables, significance and magnitude of the effect size.

In Table 11 the values of the performance indicators for the different models are given. The $R^2$ values are relative low. This low $R^2$ can be explained by the lack of important predictors due to limited data. The most important variable according to existing literature and employees of bol.com, being advertisement (Ailawadi et al., 2009; Cooper et al., 1999; Donselaar et al., 2015 & Gijsbrechts et al., 2003), is namely not included in the models. Another important variable which is not included in the model is the competition. Competitive marketing activities including prices and promotions of the competitive products are important driving factors of product sales (Huang et al., 2014; Curry et al., 1995; Divakar et al., 2005; de Leeuw, 1996).

Furthermore, when comparing the $R^2_{\text{adjusted}}$ values in Table 11 to the $R^2$ the predictive power reveals little loss. This indicates a lack of overfitting that would be shown by more marked difference between the two values (Hair et al., 2014).

<table>
<thead>
<tr>
<th>R²</th>
<th>$R^2_{\text{adjusted}}$</th>
<th>MAPE</th>
<th>RMSE</th>
<th>MAD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>0.4689</td>
<td>0.71</td>
<td>26.30</td>
<td>6.37</td>
</tr>
<tr>
<td>Model 2</td>
<td>0.4793</td>
<td>0.60</td>
<td>9.35</td>
<td>3.63</td>
</tr>
<tr>
<td>Model 3</td>
<td>0.4679</td>
<td>0.67</td>
<td>19.25</td>
<td>4.63</td>
</tr>
<tr>
<td>Model 4</td>
<td>0.4142</td>
<td>0.64</td>
<td>16.72</td>
<td>7.22</td>
</tr>
<tr>
<td>Model 5</td>
<td>0.4031</td>
<td>0.60</td>
<td>13.86</td>
<td>7.21</td>
</tr>
</tbody>
</table>

*Table 11: Overview of the performance of the different models*
Table 11 shows that model 2 performs the best, since it has the best score on the performance measures. This indicates that it is better to assign the daydeals to the corresponding campaigns than to combine the daydeals into one separate campaign named ‘Deals’. In addition, the results show that it is not better to correct the promotion dates to the general advertisement of bol.com, as done in Model 3. A full promotion must be included to create the best model. This can indicate that the promotion still continues after the general advertisement of bol.com has stopped or that the product has an uplift in the sales due to other reasons. A possible explanation can be the place on the product page. A product with a high number of views gets a higher position on the product page and thus is easier to find for the customers. This results in more sales. When a promotion has finished, the product can have a high position on this page. Therefore, it is difficult to notice if the extra uplift after the advertisement is the result of the promotion that still continues or because the product has risen on the product page compared to the reference period. Future research with more accurate promotion dates is needed to draw conclusions.

When comparing model 4 and 5, the models which are corrected for the reference period, it can be concluded that model 5 performs the best. This indicates that it is better to include the whole campaign than including only a subset of a couple of days. Due to the limited amount of data included in model 4 and 5, the $R^2$ is relative low. According to Hair et al. (2014), large samples of 1000 observations or more make the statistical significance tests overly sensitive. These often indicate that almost any relationship is statistically significant. This can be an explanation of the low performance of model 4 and 5 compared to the other models. Furthermore, due to the limited data of one campaign it was not possible to include the predictor product subgroup in model 4 and 5. For the remainder of this research model 2 is used as preferred model, because of the best performance according to Table 11.

Appendix 8 gives a detailed explanation of the independent variables, significance and magnitude of the effect size of model 2, the preferred model. It seems that the variable discount type has the biggest impact on the lift factor. This variable includes the dummy variable daydeals, which has the highest lift factor. This indicates that during a campaign the daydeal has the biggest impact on the lift factor. Additionally, the price discount has a small impact on the lift factor. Future research with a representative reference period is needed to draw conclusions on the impact of the price discount. The independent variable ‘number of products’ seems to have the least impact on the lift factor. The effect varies between the different models. Model 2 shows a positive effect of the number of products on the lift factor, while the other models show a negative effect. Future research is needed to draw conclusions on the effect of the variable ‘number of products’.

In conclusion, model 2, the preferred model, shows a substantial improvement compared to the current, subjective promotion forecasting technique used at bol.com. In Section 2.5.1 the current performance of one campaign is given. Table 12 gives an overview of the performance, using the MAPE, of the current forecast of bol.com and the final model for two product groups, being the Personal Care & Home Care and the Televisions product group. Table 12 shows that model 2 results in a substantial improvement in the forecast accuracy compared to the current forecast strategy applied at bol.com.

<table>
<thead>
<tr>
<th></th>
<th>Personal Care &amp; Home Care</th>
<th>Televisions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current forecast bol.com</td>
<td>91.9%</td>
<td>152.6%</td>
</tr>
<tr>
<td>Final model</td>
<td>60.3%</td>
<td>73.4%</td>
</tr>
</tbody>
</table>

*Table 12: Overview of the realized mean service levels*
4. Replenishment model

In this section, the inventory model for promotions is introduced. First, the inventory policies according to the literature are explained. Then, the inventory model is explained and the demand patterns are identified. Finally, the inventory model is tested using actual data of bol.com. For this chapter, the validation set is used in order to apply the forecasting model. First, the outliers and influential numbers are deleted using the same procedure as explained in Section 3.3.2.

4.1 Theoretical background

Limited literature is available about an applicable inventory policy for promotions. Most researchers assume that promotions arise from leftovers. Zhang et al. (2008) show that once the promotion is determined the optimal pricing/inventory policy is the base-stock list-price-policy. In this policy, a price discount is only given when the initial inventory is larger than the base-stock level, during this time nothing is ordered. This policy is not appropriate for bol.com, because the promotions of bol.com are independent of the inventory level. Van Donselaar et al. (2005) and De Leeuw (1996) use the alpha-policy as inventory replenishment policy during promotions, where the SKU’s are distributed in two waves. In the first wave α%, optimally between 70% and 80%, is pushed to the stores. The second wave takes place a few days after the promotion started. Then, the remaining 20% to 30% is distributed based on the early sales data. The alpha policy does not cover the main problem of promotions, namely the fluctuations in demand (Taylor and Fawcett, 2001; Gruen et al., 2002; Dubelaar et al., 2001; McKinnon et al., 2007).

According to Fisher & Raman (1996), the forecast accuracy for SKU’s with an unpredictable demand, can be significantly improved by adapting forecasts based on early sales data. They apply a response-based production planning in the fashion industry, with two ordering moments. After observing 20% of the initial demand, a big improvement in forecast accuracy can be made by adjusting the forecast.

The probability of being out of stock depends directly on the forecast accuracy. Therefore, the forecasting performance should be an important determinant of the effectiveness of the inventory rules. Furthermore, the forecast error is also a primary determinant of the safety stock component of inventory investment (Gardner, 1990). There must be a consistent story between how we forecast and how we plan safety stock (Graves, 1999). This can be realized by the use of forecast errors in order to estimate the variance in the demand process. According to Graves (1999), the safety stock should be larger in a situation with a nonstationary demand, since the demand deviation over the lead-time is larger. When combining the sources of Fisher & Raman (1996), Gardner (1990) and Graves (1999) a new inventory replenishment strategy for promotions can be created. This policy suggests that the safety stock decreases over time by accurately responding to actual sales data. This research will test if an increase in forecast adaptation moments results in a better forecast accuracy and smaller safety stock over time.

4.2 Inventory model

In this section the inventory replenishment model during promotions is determined. In order to anticipate on fluctuations in the demand during promotions, an adapted inventory policy is used. This model is based on the current (R,s,nQ) inventory policy. The model is aggregated on product level with an order advice for every day, noted as t. The promotion starts on the first day (t=1) and can end on different days depending on the duration of the campaign. Before the start of the promotion, an initial order is placed, which will arrive and stocked two days before the start of the campaign. An assumption is made that the inventory position at the start of day 1 is equal to the forecast plus safety
stock of day 1. The lead time and review period are equal to three days, which is the mean lead time and review period of bol.com. The second order arrives on day 2. An order is placed in the morning before the demand arrives. Figure 8 gives an example of a promotion with the corresponding timeline and the ordering moments. In this figure the promotion starts at day $t$, which is equal to $t=1$. The replenishment decision of day $t+4$ is out of the scope of this research. This order arrives after the promotion has ended.

![Timeline of a promotion](image)

**Figure 8: Timeline of a promotion**

An order is placed if at a review moment the inventory position is below the reorder level. Then an order of size $nQ$ is placed, which is calculated according to Formula 5. This order is needed to bring the inventory position after ordering back to or above the reorder level. Based on interviews, the casepack size for most products of bol.com is equal to one, for this model this assumption is also made. Therefore, the number of casepacks to order will only be based on the reorder level and inventory position. The sales during a stock out is lost, therefore no backorders are taken into account. The inventory position is equal to the expected inventory on hand plus the inventory in-transit, given in Formula 6 (Silver, Pyke & Peterson, 1998). In this research, the assumption is made that the order moment takes place directly after the inventory has arrived. Therefore no inventory in transit has to be considered.

\[
\begin{align*}
\text{if } IP_{x,t} < s_{x,t} \text{ then } n_{x,t} &= \left\lfloor \frac{s_{x,t}-IP_{x,t}}{Q_x} \right\rfloor = s_{x,t} - IP_{x,t} \\
IP_{x,t} &= E[I_{x,t+1}] + IT_{x,t} \\
IP_{x,t} &= \text{inventory position before potential replenishment has been placed for product } x \text{ for day } t \\
s_{x,t} &= \text{reorder level for product } x \text{ for day } t \\
n_{x,t} &= \text{number of casepack sizes to order for product } x \text{ for day } t \\
Q_x &= \text{casepack size for product } x \\
I_{x,t} &= \text{inventory on hand for product } x \text{ for day } t \\
IT_{x,t} &= \text{inventory in transit for product } x \text{ for day } t \\
L &= \text{lead time} \\
R &= \text{Review period}
\end{align*}
\]
The demand during a promotion is more subjected to fluctuations than the demand for non-promoted products (Taylor and Fawcett, 2001; Gruen et al., 2002; Dubelaar et al., 2001; McKinnon et al., 2007). Through accurate response to early sales this demand uncertainty can be reduced (Fisher & Raman, 1996). Therefore an adjusted forecast based on the realized sales is used to determine the reorder level. The adjusted forecast is calculated using realized sales and the initial forecast. Time series forecasting is applied to calculate the adjusted forecast, this method assumes that past demand history is a good indicator of future demand (Chopra & Meindl, 2010). Two methods to adjust the forecast are introduced, based on the simple moving average and the simple exponential smoothing technique (Silver, Pyke & Peterson, 1998). These methods are relatively simple, understandable techniques and can therefore easily be used in practice at bol.com. The initial forecast is based on the forecasting model as described in Chapter 3. This forecast is divided over the days using the different demand patterns, explained in section 4.3. First, the relative deviation of the initial forecast is determined, in order to evaluate the forecast. This is done using the simple moving average technique. The actual sales and the initial forecast are weighted according to the same N-periods. Therefore, these values cancel each other out, which results in the sum of both values. To compensate the new order for the actual deviation, this ratio is multiplied by the initial forecast of the future order.

\[
Y'_{x,t} = \frac{\sum_{i=1}^{t-1} D_{x,i} / N}{\sum_{i=1}^{t-1} Y_{x,i} / N} \cdot Y_{x,t} = \frac{\sum_{i=1}^{t-1} D_{x,i} / N}{\sum_{i=1}^{t-1} Y_{x,i} / N} \cdot Y_{x,t} \tag{7}
\]

\[
D_{x,t} = \text{realized demand for product } x \text{ for day } t
\]

\[
Y_{x,t} = \text{initial forecast for product } x \text{ for day } t
\]

\[
Y'_{x,t} = \text{adjusted forecast for product } x \text{ for day } t
\]

\[
N = \text{number of days}
\]

Another method to adjust the forecast is using simple exponential smoothing. This method uses a weighting to the historical data which decreases geometrically when going back in time (Silver, Pyke & Peterson, 1998). For this research, a weighting factor \( \alpha \) of 0.3 is used. This \( \alpha \) is widely used in practice (Gardner, 2006). Using the SES technique the relative deviation of the initial forecast is determined and multiplied with the initial forecast of the future order, given in Formula 8.

\[
Y'_{x,t} = \frac{\alpha^*D_{x,t-1}+(1-\alpha)^*Y'_{x,t-1}}{Y_{x,t-1}} \cdot Y_{x,t} \tag{8}
\]

\[
\alpha = \text{weight factor}
\]

The reorder level consists of the adjusted forecast over a replenishment lead time and review period plus an extra safety stock. A safety stock is kept to meet unexpected fluctuations in demand (Silver, Pyke & Peterson, 1998). The formula for the reorder level is given in Formula 9. An order has to be an integer, therefore the reorder level is rounded up.

\[
s_{x,t} = \left[\sum_{i=t+L}^{t+L+R-1} Y'_{x,i} \right] + SS_{x,t} \tag{9}
\]

\[
s_{x,t} = \text{reorder level for product } x \text{ for day } t
\]

\[
SS_{x,t} = \text{safety stock for product } x \text{ for day } t
\]
4.2.1 Safety stock
The safety stock is defined as the average level of the net stock just before a replenishment arrives (Silver, Pyke & Peterson, 1998). The safety stock can be calculated by multiplying the safety factor with the standard deviation of the forecast error, given in Formula 10. The standard deviation of the forecast error is determined based on the Root Mean Squared Error (RMSE). The RMSE is often used in fitting of squared errors of a straight line to historical data. Furthermore, the RMSE is directly related to the \( \sigma \) and can be easily computed. The safety factor depends on the service level \( P_1 \), which is defined as the probability of no stockout per replenishment cycle (Silver, Pyke & Peterson, 1998). To find the safety factor for the discrete demand distribution the fitting procedure of Adan et al. (1995) is used, which includes the Binomial, Negative Binomial, Poisson and Geometric Distribution. The demand fluctuates and therefore the standard deviation is bigger than the mean demand, which results in a mixtures of Negative Binomial and Geometric Distributions with balanced means. To make a realistic safety factor these distributions are calculated for every campaign separately. The service levels are determined using the ABC classification, as explained in Section 2.3. The current classification is not appropriate for promotions, due to the large uplifts. Therefore, a new ABC classification is made based on the promotional products. To make it applicable for bol.com one ABC classification for all the promotional products for each of the product groups is made.

\[
SS_{x,t} = k \times \sigma_{x,t}
\]

(10)
\[
k = \text{safety factor}
\]
\[
\sigma_{x,t} = \text{standard deviation of errors of forecasts for product } x \text{ for day } t
\]

To determine the RMSE the used forecast needs to be evaluated with the actual sales. There are multiple of possibilities to determine the RMSE. In the general formula of RMSE all values from previous days are taken into account and equally weighted, shown in Formula 11. To correct for the equally weighted days another formula is introduced, Formula 12. Here the focus is on the days during the review period, since these forecasts are all based on the same forecast adjustment.

\[
RMSE_{x,t} = \sqrt{\frac{1}{t-1} \sum_{i=1}^{t-1} (Y'_{x,i} - D_{x,i})^2}
\]

(11)
\[
RMSE_{x,t} = \sqrt{\frac{1}{R-1} \sum_{i=1}^{R-1} (Y'_{x,i} - D_{x,i})^2}
\]

(12)
\[
RMSE_{x,t} = \text{Root mean squared error for product } x \text{ for day } t
\]

4.3 Demand pattern
This section describes the demand pattern used for the inventory policy. The promotion forecasting models developed in Chapter 3 predict the sale for a complete promotion. To divide the expected sales per day in a promotion, a demand pattern needs to be determined. This section explains the two different demand patterns that are applied. It is assumed that the sales provided in the RPG are equal to the demand in that period.

First of all, multiple campaigns are merged to identify one demand pattern. Due to the limited data of only nine months, it was not possible to identify a demand pattern for an individual campaign. To gain sufficient data the campaigns are divided into two different types. The first type includes campaigns which prepare for a special holiday, like Christmas or Valentine’s Day. This type of campaign shows a peak in the sales as the special holiday approaches and will be referred to as Peak campaigns (Ramanathan, 2012). The other type of campaign is not related to any special holiday. An assumption
is made that these campaigns have a relatively stable demand pattern and are therefore called ‘stable campaigns’. The division between peak and stable campaigns can also be noticed when comparing the demand patterns in Appendix 9. Table 13 gives an overview of the different campaigns with an indication to which type the campaign belongs. The campaign ‘Black Friday | Cyber Monday’ is classified as a stable campaign since it does not have a run-up period.

<table>
<thead>
<tr>
<th>Campaign</th>
<th>Type of campaign</th>
</tr>
</thead>
<tbody>
<tr>
<td>Black Friday</td>
<td>Cyber Monday</td>
</tr>
<tr>
<td>Bulk 10-daagse May 2017</td>
<td>Stable campaign</td>
</tr>
<tr>
<td>Electronica deals</td>
<td>Stable campaign</td>
</tr>
<tr>
<td>Huishoudweken februari/oktober</td>
<td>Stable campaign</td>
</tr>
<tr>
<td>Kerstvoordeel</td>
<td>Peak campaign</td>
</tr>
<tr>
<td>Mid Season April</td>
<td>Stable campaign</td>
</tr>
<tr>
<td>Moederdag</td>
<td>Peak campaign</td>
</tr>
<tr>
<td>Sale/Solden januari 2017</td>
<td>Stable campaign</td>
</tr>
<tr>
<td>Sinterklaas</td>
<td>Peak campaign</td>
</tr>
<tr>
<td>Vaderdag</td>
<td>Peak campaign</td>
</tr>
<tr>
<td>Valentijn</td>
<td>Peak campaign</td>
</tr>
</tbody>
</table>

Table 13: Overview of different campaigns with indication of the corresponding type

A demand pattern is identified based on ratios in sales. In the available data only the daily sales are provided. Therefore, an assumption is made that the demand throughout the day is uniformly distributed. The duration of a promotion varies between 1 and 36 days. To compare these days each promotion is split to the same number of time periods, such that a demand pattern based on the duration in percentage can be determined. Figure 9 gives an example of a promotion which lasts 4 days and is split into three time periods.

![Figure 9: Example of promotion divided into time periods](image)

Preferably, a promotion is split into many time periods. More data points result in better demand patterns (Law & Kelton, 2000). However, due to the Uniform Distribution throughout the day this is not always useful, as can be seen in Figure 10. This figure shows that the first three time periods fall in the same day and are therefore the same. This results in a bias in the final demand pattern. Therefore, in this research it is decided to exclude the campaigns with a small duration.
To determine which promotions must be excluded a cumulative probability plot is made, given in Figure 11. This figure shows that there is a small probability of 15% that a promotion lasts 8 days or shorter. While, the percentage of promotions that take 9 days or shorter is 30%, which is substantially higher. Therefore, the promotions that last shorter than 8 days are excluded from the pattern and the rest of the inventory analysis.

Multiple ratio demand patterns based on different time periods are compared to define the most appropriate demand pattern. The first ratio based demand pattern is based on 9 time periods. This prevents a bias in the demand pattern through the Uniform Distribution throughout the day, as shown in Figure 10. The second ratio based demand pattern is based on the maximum duration of a promotion, which is equal to 36 days. This is the highest level of detail, because of the granularity of one day. Lastly, the third ratio based demand pattern that is evaluated is based on 18 time periods.

To come up with a demand pattern based on the different time periods, the following steps are applied:

1. Normalize daily sales for every promotion against a specific number of time periods (9, 18 or 36 time periods) with a distinction between stable and peak campaigns.
2. Find an empirical distribution based on the normalized sales in which all promotions are merged.
3. Determine possible theoretical distributions.
4. Test the possible theoretical distributions according to a goodness of fit test.
5. Fit the best theoretical distribution to the empirical distribution.
The steps described above results in the following empirical distributions given in Figure 12, 13 and 14. These figures show a very small difference when comparing the peak and stable campaign. A disadvantage of an empirical distribution is that it has certain “irregularities”, particularly if only a small number of data values are added. To smooth out the data and provide information on the underlying distribution a theoretical distribution can be used. The possible theoretical distributions are tested according to a goodness of fit test. The goodness-of-fit test can be used to formally assess whether the observations are an independent sample from a particular distribution. This test is not very powerful for small sample sizes (Law & Kelton, 2000). Therefore, to limit the fluctuating output first the potential distributions are determined. When looking at Figure 12, Figure 13 and Figure 14 three theoretical distributions can be appropriate, namely the Logistic Distribution, Normal Distribution or Uniform Distribution. These three theoretical distributions are tested according to a goodness of fit test. The @risk add-in for Microsoft Excel is used to establish which theoretical distribution fits the data the best. This software estimates the parameters for each of the theoretical distributions using the maximum likelihood estimator.

![Figure 12: Empirical distribution for 36 time periods](image1)

![Figure 13: Empirical distribution for 9 time periods](image2)

![Figure 14: Empirical distribution for 36 time periods](image3)
The Kolmogorov-Smirnov test (K-S test) is used as goodness-of-fit hypothesis test. This test tends to be more powerful than the chi-square test (Stephens, 1974). The K-S test tests the different models by comparing the largest distance between the empirical distribution function and the fitted distribution function for all values, named as $D_n$. Small values of $D_n$ indicate a better fit compared to larger values (Law & Kelton, 2000). When performing the K-S test different theoretical distributions are suggested, given in Table 14. The results are tested against the null hypothesis and all approve the hypothesis. The demand pattern based on 36 time periods shows smaller distances compared to the demand pattern based on 9 and 18 time periods. Therefore, it can be concluded that the demand pattern based on 36 time periods show a better fit. This confirms the theory of Law & Kelton (2000) which suggest that an empirical distribution with more data values has less “irregulations” and is therefore more accurate. This is more important than the division of days in smaller parts.

<table>
<thead>
<tr>
<th>Time periods</th>
<th>Peak</th>
<th>Stable</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>Uniform distribution with $D_n = 0.1353$</td>
<td>Normal distribution with $D_n = 0.1568$</td>
</tr>
<tr>
<td>18</td>
<td>Uniform distribution with $D_n = 0.1310$</td>
<td>Normal distribution with $D_n = 0.0996$</td>
</tr>
<tr>
<td>36</td>
<td>Logistic distribution with $D_n = 0.1199$</td>
<td>Logistic distribution with $D_n = 0.0664$</td>
</tr>
</tbody>
</table>

*Table 14: Suggested Theoretical Distributions*

In conclusion, the ratio based demand pattern based on 36 time periods fit the best to a Logistic distribution for both the peak and stable campaigns. The corresponding parameters based on the @risk add-in for Microsoft Excel are given in Table 15.

<table>
<thead>
<tr>
<th></th>
<th>Alpha</th>
<th>Beta</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peak</td>
<td>0.0276541</td>
<td>0.0032784</td>
</tr>
<tr>
<td>Stable</td>
<td>0.0277778</td>
<td>0.0050392</td>
</tr>
</tbody>
</table>

*Table 15: Corresponding parameters to the different campaigns*

In addition, to compare the ratio based demand pattern another demand pattern is introduced. This is a linear demand pattern, which is an easy applicable demand pattern. At the moment, bol.com makes use of such demand pattern. The demand pattern can be classified as a Uniform Distribution with the same parameters at the start and end of the promotion.

The following sections of this report continue with two different demand patterns based on a theoretical distribution, being a Logistic Distribution with a distinction between peak and stable campaigns and a Uniform Distribution. The initial total forecast for every promotion from Chapter 3 based on the preferred model is divided over the days according to these distributions.

### 4.4 Results

In this section, an overview of the results of the realized sales is given after running the inventory model on the validation set. The preferred forecast from Chapter 3 is used as initial forecast. The first replenishment is based on this forecast. First, the different models are explained, where after the results are described.

Table 16 gives an overview of the different models that are tested. A distinction is made between the demand pattern, adjusted forecast and the RMSE. First, the demand pattern is calculated using a Logistic Distribution and Uniform Distribution as described in Section 4.3. Second, the initial forecast is adjusted using the simple moving average and the simple exponential smoothing technique. Third, the forecast error is calculated using two different methods to determine the RMSE. The first is based on all previous forecast errors, the second takes only the days during the review period into account. Combining these three possibilities results in 8 different models shown in Table 16.
### Table 16: Overview of the different models

<table>
<thead>
<tr>
<th>Model</th>
<th>Demand pattern</th>
<th>Adjusted forecasting method</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>Uniform Distribution</td>
<td>Simple moving average</td>
<td>Including all previous days</td>
</tr>
<tr>
<td>Model 2</td>
<td>Logistic Distribution</td>
<td>Simple moving average</td>
<td>Including all previous days</td>
</tr>
<tr>
<td>Model 3</td>
<td>Uniform Distribution</td>
<td>Simple exponential smoothing</td>
<td>Including all previous days</td>
</tr>
<tr>
<td>Model 4</td>
<td>Logistic Distribution</td>
<td>Simple exponential smoothing</td>
<td>Including all previous days</td>
</tr>
<tr>
<td>Model 5</td>
<td>Uniform Distribution</td>
<td>Simple moving average</td>
<td>Focus on days of review period</td>
</tr>
<tr>
<td>Model 6</td>
<td>Logistic Distribution</td>
<td>Simple moving average</td>
<td>Focus on days of review period</td>
</tr>
<tr>
<td>Model 7</td>
<td>Uniform Distribution</td>
<td>Simple exponential smoothing</td>
<td>Focus on days of review period</td>
</tr>
<tr>
<td>Model 8</td>
<td>Logistic Distribution</td>
<td>Simple exponential smoothing</td>
<td>Focus on days of review period</td>
</tr>
</tbody>
</table>

In Appendix 10 an overview of the safety stock levels over time is given. For every promotion with a different duration a separate figure is made, including all the models from Table 16. Most figures have a peak in the beginning, whereafter the safety stock decreases. Appendix 11 gives an overview of the adjusted forecasts and the average sales over time. Overall, a higher forecast is given compared to the average sales.

### 4.5 Validation

In this section, the different models are validated according to the achieved mean service levels. Furthermore, an explanation for the differences between the models is given.

Table 17 gives an overview of the achieved mean service levels. The mean service level is calculated using the KPI introduced in Section 2.5.1. The last column shows the desired service levels for bol.com. The C-label products, slow movers, are the only products that achieve the desired service level. This can be explained by the assumption that for every product on promotion at least one product needs to be on stock, in order to fulfill the possible demand. The mean sales of these slow movers is equal to 1.25 per day. Therefore, almost always enough inventory is available, which results in high service levels. Furthermore, Table 17 shows that it is more difficult to achieve the desired service levels for the fast moving products than for the less popular products. When comparing the different models based on the achieved service levels, model 8 seems the best. Although, the differences are very small.

### Table 17: Achieved mean service levels

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
<th>Model 5</th>
<th>Model 6</th>
<th>Model 7</th>
<th>Model 8</th>
<th>Goal</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>79.2%</td>
<td>79.5%</td>
<td>83.9%</td>
<td>83.9%</td>
<td>79.3%</td>
<td>79.3%</td>
<td>83.8%</td>
<td>84.5%</td>
<td>96.0%</td>
</tr>
<tr>
<td>A</td>
<td>85.1%</td>
<td>84.3%</td>
<td>87.2%</td>
<td>87.0%</td>
<td>81.8%</td>
<td>84.1%</td>
<td>85.7%</td>
<td>86.8%</td>
<td>96.0%</td>
</tr>
<tr>
<td>B</td>
<td>90.6%</td>
<td>90.6%</td>
<td>90.2%</td>
<td>91.0%</td>
<td>90.0%</td>
<td>90.7%</td>
<td>89.5%</td>
<td>91.2%</td>
<td>96.0%</td>
</tr>
<tr>
<td>C</td>
<td>98.4%</td>
<td>99.0%</td>
<td>98.0%</td>
<td>98.1%</td>
<td>98.3%</td>
<td>98.6%</td>
<td>97.9%</td>
<td>98.2%</td>
<td>85.0%</td>
</tr>
</tbody>
</table>

In Appendix 10 the different safety stock levels over time are given. After 5 days, the first time a forecast is adjusted, a clear learning effect is visible, whereafter the safety stock level fluctuates around a fixed point. However, some graphs show a peak in the safety stock after these 5 days. This is an anticipation on a peak in the sales, which was not included in the demand pattern. Especially, the models with a forecast error based on the review period strongly react to these sales peaks. These models anticipate quicker on the sales and therefore show more fluctuations in the safety stock. In contrast, models with a forecast error based on all previous days are more stable. These models show a safety stock that slowly decreases over time, due to the big forecast error in the beginning. Furthermore, these models do not react abruptly to sales fluctuations. Therefore, the forecast error based on all previous days seems the best. This research confirms the results of Fisher & Raman (1996), who show that the forecast accuracy improves by accurate response to early sales data. The safety stock decreases after responding to early sales data. Although, this learning effect can only be seen
once. Due to the fluctuating demand, no further improvements can be made after the promotion continues.

Subsequently, in Appendix 11 the adjusted forecasts over time are given for each of the different durations. These figures show the differences between the forecasting adjustment methods. Overall can be concluded that all models are very optimistic in predicting the demand. This can be explained by the rounding of the forecasts, each time the ceiling has been taken. When comparing the different forecasting adjustment methods, it seems that the simple moving average technique is the best for bol.com. This method shows the closest fit to the realized sales.

Moreover, a first conclusion can be drawn about the different demand patterns that are used. The figures from Appendix 10 and Appendix 11 does not show a clear difference between the different demand patterns. Both distributions can be used to forecast the demand. Furthermore, a small effect can be found between the different type of campaigns, the campaigns with a peak and the stable campaigns. This effect is based on the different parameter values. More data is needed to draw conclusions on the different demand patterns.

Next, when comparing the different models no final model can be determined. However, for the comparison with the current situation of bol.com model 2 is assigned as preferred model. This model includes the simple moving average technique and the forecast error based on all previous days, such that the sales does not respond abruptly to demand fluctuations. When comparing the achieved mean service levels in Table 17, the Logistic Distribution seems to perform the best. Therefore, model 2 is used as preferred model. Table 18 compares model 2 with the current situation of bol.com. This table gives an overview of the achieved mean service levels and the final mean inventory levels per product. The table shows a lower service level for the preferred model compared to the current situation. This can be explained by the high amount of leftovers. When comparing the original number of leftovers with the number of leftovers based on this promotion forecasting and inventory model, the final inventory level decreases with 72.9%. Although, the desired service level of 98% is not achieved the inventory model shows a big improvement when comparing the final inventory levels. Future research can improve this inventory model, such that the desired service levels can be achieved.

<table>
<thead>
<tr>
<th>Both product groups</th>
<th>Personal Care &amp; Home Care</th>
<th>Televisions</th>
<th>Final mean inventory levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current situation</td>
<td>98.6%</td>
<td>98.5%</td>
<td>99.6%</td>
</tr>
<tr>
<td>Preferred model</td>
<td>96.0%</td>
<td>95.9%</td>
<td>97.3%</td>
</tr>
</tbody>
</table>

Table 18: Difference current situation and preferred model

In conclusion, the forecasting and inventory models developed in this research show a substantial difference compared to the old method. The realized mean service levels are 2.6% lower than the mean realized service levels of bol.com. However, the final inventory level shows a substantial positive difference. The final model results in a 72.9% smaller final inventory level. Therefore, the final model results in a better trade-off between high inventory costs and lost sales. Furthermore, it is unclear which inventory model performs the best. The simple moving average technique and the forecast error based on all previous days do not respond abruptly to sales fluctuations. Therefore, in this case these techniques perform the best. Future research with a clear demand pattern is needed to draw conclusions on the best method. Furthermore, the learning effect from the forecast adjustment can especially be seen in the beginning after correcting the initial forecast for the realized sales, whereafter no clear future improvements can be seen. Therefore, the learning effect does not improve over time.
5. Conclusion and reflection

5.1 Conclusion

This report studies the difficulties of promotion forecasting and inventory replenishment in the online retailing industry due to fluctuating demand. This leads either to high inventory costs or lost sales. Although enough inventory can cover the promotional demand, a balance between costs and lost sales is preferred. To achieve this, a data driven approach was followed where a promotion forecasting model was created and then tested in an adapted inventory policy for promotions. This answers the main research question of this report:

How can an online retailing company accurately respond to fluctuating demand with regard to promotions by the use of a promotion forecasting model and inventory policy?

The promotion forecasting model was created using multiple regression analysis. Currently, subjective forecasts based on human judgment are used to predict promotional demand. First, before the promotion forecasting model was created, the relevant factors related to the promotion forecasting model were identified using literature research and semi-structured interviews. Finally, five different models were developed. Due to the lack of important predictors, the explanatory power of the models was relatively low. The most important variable according to many researchers and the employees of bol.com could not be included, namely advertisement. Another important variable which was not included in the model is the competition. Finally, when comparing the different models, model 2 performs the best. This preferred model is made from all available data. Although, the campaign ‘Deals’ is deleted and the daydeals in this campaign are assigned to the corresponding general campaign. For this model, the predictor discount type has the biggest impact on the lift factor. This variable includes the dummy variable daydeals, which has the highest lift factor. This preferred model shows a substantial improvement compared to the current, subjective promotion forecasting technique of bol.com. For the remainder of this research model 2 was used.

The inventory replenishment policy for promotions is based on the current (R,s,nQ) inventory policy with an accurate response to actual sales data. The forecast is adjusted using the simple moving average and simple exponential smoothing technique. These techniques are corrected for the demand pattern. The applied demand patterns are the Uniform Distribution and the Logistic Distribution. Within the Logistic Distribution, a distinction between stable and peak campaigns is made. A small difference between these campaigns is noticed. More research is needed to draw conclusions on the appropriateness of these different campaigns. Furthermore, the forecast error is determined based on the RMSE. Two different methods are applied including all values from previous days and the days during the review period. The simple moving average technique and the forecast error based on all previous days perform the best. These methods do not respond abruptly to sales fluctuations. After the first forecast adjustment a clear learning effect can be seen, whereafter no further improvements can be noticed. In conclusion, it is not necessary to adjust the forecast continuously, one adjustment is enough to improve the forecast accuracy. The forecasting and inventory model developed in this research shows a substantial difference compared to the old method. The realized mean service levels are 2.6% lower than the realized mean service levels of bol.com. However, the final inventory level shows a substantial difference. The model results in a 72.9% smaller final inventory level, which results in a better trade-off between high inventory costs and lost sales.
5.2 Recommendations

Several recommendations for bol.com follow from this research. Some of these recommendations can be implemented directly, while for others more research is needed. In this section, the recommendations following from this research are listed and explained.

Recommendation 1: Data collection
To draw better conclusions, more data is needed. Important data to create an accurate promotion forecasting model is missing, for example the advertisement during promotion and information about the competition. Furthermore, it is necessary to save all the coming promotions. For this research limited historical data was available. Therefore, the predictive power of the promotion forecasting model was relatively low. More historical data and more factors, like advertisement and competition, would result in a substantial improvement, with a representative demand pattern.

Recommendation 2: Inform and learn employees about RPG
The RPG is not used appropriately by the employees of bol.com. The system is occasionally used for other purposes instead of promotions. This results in data troubles, for example in promotions that take 365 days a year and duplicated promotions. It was hard to distinguish the non-promotional data and the promotional data, which makes it difficult to draw conclusions. Therefore, assumptions were made. To make an appropriate analysis of the promotional data, the employees need to be informed about the purposes of the system and get a proper feeling how to work with the system.

Recommendation 3: Update reference period
This research is based on a given reference period which refers to exactly the same days as before the start of the promotion. During this reference period often another promotion takes place. This makes it difficult to draw conclusions about the uplift and price discount of the promotion. Therefore, it is necessary to get another reference period which is representative for the non-promotional demand. This would make it easier to compare different promotions. More research is needed to come up with the best reference period.

Recommendation 4: Change the current promotion forecasting method to a statistical method
At the moment bol.com is using a subjective forecasting method based on human judgment to predict the promotional demand. This research shows a substantial increase in the forecast accuracy when making a forecast based on multiple regression analysis. Therefore, it is recommended to change the current promotion forecasting method to a statistical method.

Recommendation 5: Test and implement new inventory policy
In this research, a new inventory policy for promotions is introduced. The relative low explanatory power of the promotion forecasting model and the aggregate demand patterns have a big impact on the inventory levels. This makes it difficult to draw conclusions about the different models used for the inventory model. To find the best model this inventory policy first needs to be tested with a better forecasting model and demand pattern. Overall, this inventory policy already showed a big improvement compared to the current situation. Therefore, it is recommended to change the current replenishment strategy for promotions to the inventory model introduced in this research and do more research to identify the best model.

Recommendation 6: Adjust forecast at the beginning of the promotion
The introduced inventory policy is based on an accurate response to sales data. It seems that especially in the beginning of the promotion the forecast becomes better after response to the realized sales. Therefore, it is necessary to adjust the forecast as soon as possible on the actual sales. This results in better demand forecasts, less excess of stock and lost sales.
5.3 Reflection

In this section, a reflection of the performed research is given. First, the academic contribution of this research is described, followed by the limitations. Finally, the ideas for future research are presented.

5.3.1 Academic contribution

Much research is performed on the topic promotion forecasting. The focus of these studies is on the traditional retailers instead of the e-commerce sector. This is notable since the online market is different from the traditional market. There are lower switching costs in the e-commerce sector, it is easier to compare prices of different retailers and switch to another retailer. Furthermore, an online retailer has a bigger assortment and can promote more different products. The inventory replenishment strategy is also different. The storage capacity is bigger for the online retailer and the products are bought directly from an external supplier. This makes it necessary to anticipate on time to promotions and consult the supplier. In conclusion, it is important to take the differences between the traditional and online retailer into account for the demand forecasting and inventory replenishment strategy for promotions. This research will contribute to this topic with a promotion forecasting model and an inventory replenishment strategy specific for the e-commerce sector.

In addition, there is limited literature available about an applicable inventory policy for promotions. Most researchers assume that promotions arise from leftovers. This research introduces a new combination of the push and pull replenishment strategy with multiple of techniques to find the best inventory replenishment policy for promotions. First, the promotion is pushed to the customer based on an explanatory forecast. Where after accurate response to the actual sales is shown. Every review moment the forecast is adjusted for the actual sales, using time series techniques. This ensures an optimal response to the fluctuating demand based on historical and actual data. Furthermore, the safety stock is also adjusted for the actual sales using the forecast error. This research established a new combination of multiple of methods in the e-commerce sector to find the best inventory replenishment policy for promotions.

Furthermore, this research has introduced a new method to determine the demand pattern. This method is determined using normalization of daily sales data in order to create an empirical distribution, where after this distribution is fit to a theoretical distribution. This distribution was based on different time periods. The empirical distribution with more data values shows a better fit. The number of data points seems more important than the division of the days in smaller parts. Furthermore, a distinction between two kinds of campaigns is analyzed, namely campaigns with a peak because of a special holiday and the general campaigns of bol.com, stable campaigns. A small difference between these campaigns is noticed. More research is needed to draw conclusions on the appropriateness of this split.

Next, Fisher and Raman (1996) found that the forecast accuracy can be improved through accurate response to early sales data. Their research is applied in the fashion industry with a one-time response to early sales data. This master thesis provides an extension on their work by accurate responding multiple of times to the actual sales data during the whole promotion period in the online retailing industry. Promotions imply short time periods compared to the fashion industry. Furthermore, this research shows the impact of multiple response moments to the realized sales. Also, the impact on the safety stock is reviewed in order to find a learning effect. In conclusion, no further improvements can be found after the first time the forecast is adjusted.
Finally, online retailing companies can benefit from the research by adjusting their promotion forecasting and inventory replenishment strategy to the findings in this report. Although the results would differ, as long as replenishment during the promotion is possible, methods can still be applied. Small lead times and reviews periods make it possible to respond to the actual sales. If the lead time and review period are longer than the promotion takes, it is not possible to respond to the actual sales. Then, the promotion is dependent on the initial forecast. To draw accurate initial forecasts, data collection is very important. Furthermore, this research can also be applied to other situations with uncertain demand, for example new products. Then comparable situations need to be analyzed to come up with an initial forecast, where after the inventory policy will respond to early sales data.

5.3.2 Limitations
In this section, the limitations of this research are discussed. The most important limitation comprises the data availability and quality. Limited data was available and important data about predictors was missing. Therefore, many assumptions are made during this research. These assumptions do not necessarily hold. Due to the inappropriate use of the RPG the quality of the data was limited. Therefore, it was not possible to identify the promotion frequency and an assumption about the correct promotional dates is made. In addition, it was not possible to access all the data sources of bol.com. Therefore, many assumptions based on interviews with employees are made. Furthermore, there was a poor connection between the product numbers from the RPG and the other data systems of bol.com. This makes it impossible to determine the current inventory levels from all promotions. Another limitation based on the limited data evolves the statistical significance of the demand pattern. Due to the limited data of only nine months, it was not possible to identify a demand pattern for an individual campaign.

Furthermore, the available reference period is not representative. The reference period is often referred to a period with another promotion. This makes it hard to compare the promotion with the regular demand. The initial forecast is dependent on the uplift and price discount which are determined using this inappropriate reference period. This makes the initial forecast and also the expected demand as input for the inventory policy not representative.

In addition, the data was not proportionally distributed over the product groups. The product group Personal Care & Home Care included more data than the product group Televisions. Within Personal Care & Home Care, 12 product subgroups were available, compared to one product subgroup for Televisions. Therefore, the results are not necessarily generalizable to the product group Televisions.

Finally, in this research certain techniques for the promotion forecasting and inventory model are used to come up to the final results. There are also other methods to forecast the promotional demand, adjust the forecast, calculate the safety stock and determine the demand pattern. These can result in different outcomes. Future research could investigate which technique performs the best.

5.3.3 Future research
Several directions for future research follow from this research. In the recommendations, several directions for future research are already introduced. In this section, several other directions based on the limitations are further explained.
First, more research on important factors for promotion forecasting and the demand pattern during a promotion in the e-commerce sector is needed. It seems that the position on the product page is an important factor, which needs to be tested in future research. The current research provides some exploratory insights, but establishing an accurate forecast and replenishment strategy is not possible due to lack of data. More insights on this topic will help online retailers to investigate these drivers for themselves. They can use these insights as a guideline and address them.

Second, more research on the customer behavior is needed in order to improve the profitability of promotions. This includes the promotional assortment, the position of the products on the product page and the number of products in promotion. At the moment bol.com includes a broad assortment in the campaigns. This results in many C-label products, which are barely sold. The number of products has a small effect on the lift factor, in some models this effect is positive and in other models this effect is negative. Therefore, in order to introduce more profitable promotions more research on the customer behavior is needed.

Third, future research on the best reference period can contribute to an accurate promotion forecast. The reference period must be representative for the non-promotional demand, including the seasons and general growth of the company. This results in a better promotion forecast and inventory levels.

Fourth, more research about an appropriate deviation of the products can help to improve promotion forecasting. When including all promotional products a cluster analysis can be used to classify the products in groups based on their characteristics (Hair et al., 2014). This classification can be included as a predictor in the promotion forecasting model.

Lastly, this research can be extended by including the financial, logistics, and marketing research fields. More insight on these topics would help the online retailing industry to have more profitable promotions.
Reference list


Chockalingam, M. (2010). Response to MAPE and MPE calculations. USA.


Thomopoulos, N.T. (2015). Demand forecasting for inventory control. In *Demand Forecasting for Inventory Control* (pp. 1-10). Springer International Publishing.


Appendix 1 Assumptions

In this section, an overview of the assumptions made for this research is given.

1. Most suppliers of bol.com deliver on time. Furthermore, when a supplier does not have enough inventory available, bol.com will be informed on time. Therefore, the assumption is made that each supplier has infinite stock and delivers always on time.

2. It is assumed that the sales provided in the RPG are equal to the demand in that period.

3. It is assumed that promotions are planned beforehand and thus the decision to promote does not depend on the level of inventory, during a campaign no changes and cancellations. A cancellation and extra advertisement involve extra costs, which must be prevented.

4. Based on interviews it is assumed that the size of the warehouse and inbound planning does not influence the inventory policy of bol.com.

5. Based on interviews it is assumed that there are no extra costs for the frequency and number of products that will be ordered.

6. The assumption is made that the last action in the RPG is the revised promotion and thus the correct promotion.

7. Due to limited data, the assumption is made that the reference period is a representative for the non-promotional demand.

8. The assumption is made that a daydeal during a general campaign gets a higher lift factor due to the extra attention from the general campaign. The mean lift factor for daydeals during a general campaign is 50.29 and the lift factors for separate daydeals is 37.23.

9. For model 5 from the promotion forecasting models, the assumption is made that the sales of the SKU’s in the reference period are not affected by the campaign in the reference period.

10. The Mean Absolute Percentage Error (MAPE) assumes that the absolute error on each item is equally important.

11. An assumption is made that the inventory position at the start of day 1 is equal to the forecast plus safety stock of day 1.

12. The casepack size for most products of bol.com is equal to one, therefore this is also assumed during this research.

13. In this research, the assumption is made that the order moment takes place directly after the inventory has arrived.

14. Based on the demand patterns in Appendix 9, the assumption is made that the general campaigns, thus the campaigns not related to any special holiday, have a relatively stable demand pattern and are therefore called ‘stable campaigns’.

15. In the available data only the daily sales are provided. Therefore, an assumption is made that the demand throughout the day is uniformly distributed.

16. The assumption is made that for every product on promotion at least one product needs to be on stock, in order to fulfill the possible demand.
Appendix 2  Functions

In this part, an introduction to the relevant job descriptions for this project is given. This includes the supply chain specialist, product specialist, brand specialist, promotion specialist and buyer. First, the supply chain specialist is part of the Supply Chain Management department. The supply chain specialist is responsible for the operations in the supply chain, this includes the ordering of products, supplier management and inventory management. Second, the product specialist is responsible for a product group of a specific category. Each category is responsible for a range of products within bol.com. The product hierarchy of bol.com is explained in Appendix 3. The product specialist wants to provide an optimal consumer experience, this includes making promotions. Third, the brand specialist is responsible for a specific brand within a category. Fourth, the promotion specialist takes care of the promotions of a specific category. Lastly, the buyer will contract suppliers and is responsible for the purchase prices of the products.
Appendix 3   Product hierarchy

The products of bol.com are structured according to the GPC product classification. First, a global classification in Units is made: Unit Books & Entertainment, Unit Electronics, Unit Home, Garden & Toys and Unit Lifestyle & Daily Needs. Within the department of Supply Chain Management, a distinction is made between these units. Secondly, the units are further classified in clusters. In Figure 15 the product hierarchy of the FinCat Beauty & Care is given as an example. Each Unit/Cluster takes care of different FinCat’s or also called Categories. Daily Needs consist of the following FinCat’s: Pet, Beauty & Care and Health & Intimacy. Each FinCat consist of multiple Product Groups. Furthermore, each Product Group consists of different Bricks and each Brick of different Chunks, these are not given in the figure. Figure 16 gives the product hierarchy of the FinCat Sound & Vision.

![Figure 15: Product hierarchy for Personal Care & Home Care](image1)

![Figure 16: Product hierarchy for Televisions](image2)
### Appendix 4  Interviewed employees

<table>
<thead>
<tr>
<th>Respondent</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Respondent 1</td>
<td>Supply Chain Specialist unit lead Electronics</td>
</tr>
<tr>
<td>Respondent 2</td>
<td>Supply Chain Specialist unit lead Lifestyle &amp; Daily Needs</td>
</tr>
<tr>
<td>Respondent 3</td>
<td>Supply Chain Specialist unit lead Kitchen Table &amp; Household (validation)</td>
</tr>
<tr>
<td>Respondent 4</td>
<td>Supply Chain Specialist Beauty &amp; Care</td>
</tr>
<tr>
<td>Respondent 6</td>
<td>Supply Chain Specialist Sound &amp; Vision</td>
</tr>
<tr>
<td>Respondent 7</td>
<td>Former supply chain Specialist Beauty &amp; Care (validation)</td>
</tr>
<tr>
<td>Respondent 8</td>
<td>Former supply chain specialist Television and slim4 expert (validation)</td>
</tr>
<tr>
<td>Respondent 9</td>
<td>Supply Chain specialist and slim4 expert</td>
</tr>
<tr>
<td>Respondent 10</td>
<td>Trainee promotion forecasting at slim4</td>
</tr>
<tr>
<td>Respondent 11</td>
<td>Business Analyst (former Supply Chain specialist)</td>
</tr>
<tr>
<td>Respondent 12</td>
<td>Business Analyst (manager forecasting team) (validation)</td>
</tr>
<tr>
<td>Respondent 13</td>
<td>Big data analyst promotions (validation)</td>
</tr>
<tr>
<td>Respondent 14</td>
<td>Promotion specialist Beauty &amp; Care</td>
</tr>
<tr>
<td>Respondent 15</td>
<td>Promotion specialist Televisions</td>
</tr>
<tr>
<td>Respondent 16</td>
<td>Buyer Beauty &amp; care</td>
</tr>
<tr>
<td>Respondent 17</td>
<td>Buyer Beauty &amp; care (validation)</td>
</tr>
<tr>
<td>Respondent 18</td>
<td>Buyer Television</td>
</tr>
<tr>
<td>Respondent 19</td>
<td>Commercial specialist</td>
</tr>
</tbody>
</table>

*Table 19: Overview of the interviewed employees of bol.com*
Appendix 5  Variables forecasting model

Independent variable

**Lift Factor (min 0, max 576)**
The Lift Factor is defined as the multiple of the baseline sales in units expected when any is promoted in a particular way (van Donselaar, 2016). The Lift Factor is measured by dividing the sales during a promotion in units divided by the baseline sales in units. For the baseline sales, a reference period is used. These will be the same days before the start of the promotion. It is assumed that these days will be the same as the weeks beforehand, this can be seen in Figure 4. Since the daily sales strongly fluctuate there is chosen to take the total lift factor of the whole promotion.

**Dependent variable**

**TotalPriceDiscount (min -246, max 75)**
The TotalPriceDiscount is defined as the total price discount for a product during the whole promotion period. The promotion price and reference price differ over time. When the total price discount is negative the product is more expensive in the promotion period than in the reference period.

**NumberOfProducts (min 1, max 754)**
The NumberOfProducts is defined as the total number of products in an action. Actions are unique for a specific amount of products from the same product subgroup that are promoted together and also have the same action discount type. These are communicated against the same action so with the same advertisement and kind of products. The actions are not taken into account, since these every action is unique. Furthermore, has an action too little observations.

**DurationGlobalInPromotion (min 1 day, max 36 days)**
The DurationGlobalInPromotion is the number of days a specific product in an action is on promotion.

**campaignName (nominal)**
The name of the specific campaign is defined with the variable campaignName. There are 13 different campaigns in the given dataset. Table 20 gives an overview of all the campaigns with an explanation.
<table>
<thead>
<tr>
<th>Campaign</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Black Friday</td>
<td>Cyber Monday</td>
</tr>
<tr>
<td>Bulk 10-daagse May 2017</td>
<td>Bol.com broad campaign applied 4 times a year, with a lot of promotions for all categories of bol.com</td>
</tr>
<tr>
<td>Deals</td>
<td>Daydeals are not part of a specific campaign and therefore mentioned as a separate campaign. Since these do not get the same advertisement and only takes one day.</td>
</tr>
<tr>
<td>Electronika deals</td>
<td>Campaign for electronica products that takes place once a year.</td>
</tr>
<tr>
<td>Huishoudweken februari/oktober</td>
<td>Campaign that takes place twice a year with especially cleaning products. The available data contains the huishoudweken of October and February.</td>
</tr>
<tr>
<td>Kerstvoordeel</td>
<td>Campaign around Christmas.</td>
</tr>
<tr>
<td>Mid Season April</td>
<td>A campaign from fashion, last year Personal Care &amp; Home Care responded to this by the promotion of sunscreen.</td>
</tr>
<tr>
<td>Moederdag</td>
<td>Special campaign around Mother’s Day.</td>
</tr>
<tr>
<td>Sale/Solden januari 2017</td>
<td>Bol.com broad campaign applied 2 times a year, with especially outlet products from all categories. Some categories use this campaign also as an extra promotion for their products, such as personal care &amp; home care. The outlet products are excluded from this research.</td>
</tr>
<tr>
<td>Sinterklaas</td>
<td>Campaign around special day Sinterklaas.</td>
</tr>
<tr>
<td>Vaderdag</td>
<td>Special campaign around special day Father’s Day.</td>
</tr>
<tr>
<td>Valentijn</td>
<td>Special campaign around special day Valentine.</td>
</tr>
</tbody>
</table>

Table 20: Overview of all campaigns of bol.com

**actionDiscountTypeName (nominal)**

Specific kind of discount, there are 6 different types:

- Cheapest product free
- Daydeal
- Item discount, amount off
- Item discount, percentage off
- Price off
- X for fixed
ProductSubgroup (nominal)
The categorization within a product group. For this research the following product subgroups are used:

- Aromatherapie
- Gezichtsverzorging
- Haarverzorging
- Hand & Voetverzorging
- Lichaamsverzorging BPH
- Mondverzorging
- Persoonlijke Hygiene BPH
- Persoonlijke Verzorging Geschenksets
- Scheren & Ontharen
- Schoonmaakartikelen
- Televisies (the only product subgroup of Televisions)
- Zelfbruiners
- Zonnebrandcreme
Appendix 6 Assumptions

Transformation

Histogram with Normal Curve

Normal Q-Q Plot

Histogram with Normal Curve

Normal Q-Q Plot

Normality

Dependent Variable: LLI_Liftfactor

Expected Cum Prob

Observed Cum Prob
Model 1

Homoscedasticity

Before data transformation:

![Scatterplot to check heteroskedasticity](image1.png)

**studentized Breusch-Pagan test**

_data: reg1_

BP = 99.028, df = 32, p-value = 8.963e-09

After data transformation:

![Scatterplot to check heteroskedasticity](image2.png)

**studentized Breusch-Pagan test**

_data: reg1_

BP = 147.50, df = 32, p-value < 2.2e-16
Linearity

Model 2
Homoscedasticity

studentized Breusch-Pagan test

data: Reg1SplitSample
BP = 200.43, df = 79, p-value = 1.61e-12
Linearity

Model 3
Homoscedasticity
studentized Breusch-Pagan test

data: Reg1SplitSample
BP = 322.84, df = 92, p-value < 2.2e-16

Linearity

Model 4
Homoscedasticity
studentized Breusch-Pagan test
data:  reg1
BP = 25.567, df = 10, p-value = 0.004369

Linearity

Model 5
Homoscedasticity

studentized Breusch-Pagan test
data:  Reg1SplitsSample
BP = 22.577, df = 19, p-value = 0.2565
Linearity

Added-Variable Plots

- TotalPriceDiscount vs. others
- DurationGloballPromotion vs. others
- NumberOfProducts vs. others
### Appendix 7  Final models

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
<th>Model 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>5.574 ***</td>
<td>5.423 ***</td>
<td>3.045 ***</td>
<td>0.629</td>
<td>8.604 ***</td>
</tr>
<tr>
<td>Total price discount</td>
<td>0.052 ***</td>
<td>0.046 ***</td>
<td>-0.016</td>
<td>0.095 ***</td>
<td>-0.743 **</td>
</tr>
<tr>
<td>Duration global in promotion</td>
<td>0.024 ***</td>
<td>-0.200 ***</td>
<td>0.080 ***</td>
<td>0.197 *</td>
<td>0.155 **</td>
</tr>
<tr>
<td>Number of products</td>
<td>-0.004 ***</td>
<td>0.003 *</td>
<td>-0.002 **</td>
<td>-0.003 *</td>
<td>-0.558 **</td>
</tr>
<tr>
<td>Campaign: Black Friday</td>
<td>1.719 ***</td>
<td>1.466</td>
<td>3.129 ***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Campaign: Bulk 10-daagse mei</td>
<td>0.093 ***</td>
<td>-0.932</td>
<td>1.353 ***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Campaign: Deals</td>
<td>-1.758</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Campaign: Elektronica deals april</td>
<td>1.033 *</td>
<td>0.154</td>
<td>1.733 *</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Campaign: Huishoudweken februari</td>
<td>0.099</td>
<td>-1.985 ***</td>
<td>0.677</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Campaign: Huishoudweken oktober</td>
<td>REF</td>
<td>REF</td>
<td>2.545 **</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Campaign: Kerstwinkel</td>
<td>-0.803 **</td>
<td>-1.892 **</td>
<td>1.301 **</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Campaign: Mid Season Sale</td>
<td>-0.016</td>
<td>3.059</td>
<td>4.626</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Campaign: Moederdag</td>
<td>0.082</td>
<td>-0.544</td>
<td>1.629 **</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Campaign: Sale/solden januari</td>
<td>-0.480</td>
<td>-2.793 ***</td>
<td>REF</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Campaign: Sintwinkel</td>
<td>-0.746 ***</td>
<td>-2.073 ***</td>
<td>0.603</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Campaign: Vaderdag</td>
<td>-0.577 *</td>
<td>-1.988 ***</td>
<td>-0.102</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Campaign: Valentijn</td>
<td>-0.549 *</td>
<td>-2.329 ***</td>
<td>-0.127</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Discount type: Amount off</td>
<td>-4.724 ***</td>
<td>-3.385 ***</td>
<td>-2.977 ***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Discount type: Cheapest Product free</td>
<td>-2.530 **</td>
<td>-1.424 ***</td>
<td>-2.128 ***</td>
<td>5.857</td>
<td>REF</td>
</tr>
<tr>
<td>Discount type: Daydeal</td>
<td>REF</td>
<td>REF</td>
<td>REF</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Discount type: Percentage off</td>
<td>-5.223 ***</td>
<td>-3.897 ***</td>
<td>-3.541 ***</td>
<td>-1.625 ***</td>
<td>-5.778 ***</td>
</tr>
<tr>
<td>Discount type: X for fixed</td>
<td>-4.093 **</td>
<td>-3.009 ***</td>
<td>-3.329 ***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Product subgroup:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aromatherapie</td>
<td>-1.940</td>
<td>-2.510</td>
<td>-1.522</td>
<td>-1.317</td>
<td></td>
</tr>
<tr>
<td>Gezichtsverzorging</td>
<td>-1.402 ***</td>
<td>-1.075 ***</td>
<td>-0.865 ***</td>
<td>-3.350 **</td>
<td></td>
</tr>
<tr>
<td>Product subgroup:</td>
<td>-1.973 ***</td>
<td>-1.491 ***</td>
<td>-1.320 ***</td>
<td>-3.870 **</td>
<td></td>
</tr>
<tr>
<td>Haarverzorging</td>
<td>-2.014 ***</td>
<td>-1.930 ***</td>
<td>-1.495 **</td>
<td>-2.100 *</td>
<td></td>
</tr>
<tr>
<td>Product subgroup:</td>
<td>-1.358 ***</td>
<td>-1.267 ***</td>
<td>-1.034 ***</td>
<td>-3.949 **</td>
<td></td>
</tr>
<tr>
<td>Lichaamsverzorging</td>
<td>-1.776 ***</td>
<td>-0.924 *</td>
<td>-0.685</td>
<td>-3.709 **</td>
<td></td>
</tr>
<tr>
<td>Product subgroup:</td>
<td>-4.095</td>
<td>-0.521</td>
<td>-0.376</td>
<td>-3.789 **</td>
<td></td>
</tr>
<tr>
<td>Persoonlijke hygiene BPH Product subgroup: Persoonlijke Verzorging Geschenksets</td>
<td>REF</td>
<td>-0.284</td>
<td>REF</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Product subgroup: Scheren &amp; Ontharen</td>
<td>1.612 ***</td>
<td>-0.801 **</td>
<td>-0.414</td>
<td>-2.118</td>
<td></td>
</tr>
<tr>
<td>---------------------------------------</td>
<td>-----------</td>
<td>-----------</td>
<td>--------</td>
<td>--------</td>
<td></td>
</tr>
<tr>
<td>Product subgroup: Schoonmaakartikelen</td>
<td>-1.552 ***</td>
<td>-0.975 ***</td>
<td>-0.594 *</td>
<td>REF</td>
<td></td>
</tr>
<tr>
<td>Product subgroup: Televisions</td>
<td>-1.045 **</td>
<td>-0.043</td>
<td>-0.488</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Product subgroup: Zelfbruiners</td>
<td>-2.404</td>
<td>-3.618 ***</td>
<td>-3.201 **</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Product subgroup: Zonnebrandcreme</td>
<td>-0.673 *</td>
<td>REF</td>
<td>0.108</td>
<td>-4.077 **</td>
<td></td>
</tr>
</tbody>
</table>

<p>| Totalprice discount * Campaign: Black Friday | -0.039 *** | -0.030 ** | 0.036 ** |
| Totalprice discount * Campaign: Bulk 10-daagse mei | -0.021 * | -0.025 *** | 0.042 *** |
| Totalprice discount * Campaign: Deals | -0.009    | -0.086 *  | -0.002 |
| Totalprice discount * Campaign: Elektronica deals april | -0.118 * | -0.024 *** | 0.045 *** |
| Totalprice discount * Campaign: Huishoudweken februari | -0.026 *** | REF | 0.064 *** |
| Totalprice discount * Campaign: Huishoudweken oktober | REF | -0.052 *** | 0.035 *** |
| Totalprice discount * Campaign: Kerstwinkel | -0.050 *** | -0.030   | 0.087 * |
| Totalprice discount * Campaign: Mid Season Sale | 0.005    | -0.032 *** | 0.047 *** |
| Totalprice discount * Campaign: Moederdag | -0.033 *** | -0.036 *** | REF |
| Totalprice discount * Campaign: Sale/solden januari | -0.028 *** | -0.042 *** | 0.026 ** |
| Totalprice discount * Campaign: Sintwinkel | -0.044 *** | -0.026 *** | 0.041 *** |
| Totalprice discount * Campaign: Vaderdag | -0.025 ** | -0.029 *** | 0.036 *** |
| Totalprice discount * Campaign: Valentijn | -0.027 *** | 0.339    | 0.217 |
| Number * Product subgroup: Aromatherapie | -0.029    | -0.001   | 0.004 *** |
| Number * Product subgroup: Gezichtsverzorging | 0.005 *** | -0.001   | 0.005 *** |
| Number * Product subgroup: Haarverzorging | 0.006 *** | -0.002   | 0.003 ** |
| Number * Product subgroup: Hand &amp; Voetverzorging | 0.005 *** | -0.002   | 0.004 *** |</p>
<table>
<thead>
<tr>
<th>Number * Product subgroup: Lichaamsverzorging</th>
<th>0.005 ***</th>
<th>-0.003</th>
<th>0.007 **</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number * Product subgroup: Mondverzorging</td>
<td>0.005 ***</td>
<td>-0.006 *</td>
<td>-0.000</td>
</tr>
<tr>
<td>Number * Product subgroup: Persoonlijke hygiene BPH</td>
<td>0.006</td>
<td>-0.006 ***</td>
<td>REF</td>
</tr>
<tr>
<td>Number * Product subgroup: Persoonlijke Verzorging Geschenksets</td>
<td>REF</td>
<td>0.000</td>
<td>0.006 ***</td>
</tr>
<tr>
<td>Number * Product subgroup: Scheren &amp; Ontharen</td>
<td>0.005 ***</td>
<td>-0.001</td>
<td>0.004 ***</td>
</tr>
<tr>
<td>Number * Product subgroup: Schoonmaakartikelen</td>
<td>0.005 ***</td>
<td>-0.014</td>
<td>-0.035</td>
</tr>
<tr>
<td>Number * Product subgroup: Televisions</td>
<td>-0.016</td>
<td>0.005</td>
<td>0.008 *</td>
</tr>
<tr>
<td>Number * Product subgroup: Zelfbruiners</td>
<td>0.009</td>
<td>REF</td>
<td>0.004 **</td>
</tr>
<tr>
<td>Number * Product subgroup: Zonnebrandcreme</td>
<td>0.002 *</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duration * Campaign: Black Friday</td>
<td>-0.204</td>
<td>-0.068</td>
<td></td>
</tr>
<tr>
<td>Duration * Campaign: Bulk 10-daagse mei</td>
<td>0.140 **</td>
<td>0.002</td>
<td></td>
</tr>
<tr>
<td>Duration * Campaign: Deals</td>
<td>-0.057</td>
<td>0.028</td>
<td></td>
</tr>
<tr>
<td>Duration * Campaign: Elektronica deals april</td>
<td>0.166 ***</td>
<td>0.007</td>
<td></td>
</tr>
<tr>
<td>Duration * Campaign: REF</td>
<td>-0.165 *</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duration * Campaign: Huishoudweken februari</td>
<td>0.094 *</td>
<td>-0.039 *</td>
<td></td>
</tr>
<tr>
<td>Duration * Campaign: Huishoudweken oktober</td>
<td>-0.145</td>
<td>-0.269</td>
<td></td>
</tr>
<tr>
<td>Duration * Campaign: Kerstwinkel</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duration * Campaign: Mid Season Sale</td>
<td>0.051</td>
<td>-0.065 *</td>
<td></td>
</tr>
<tr>
<td>Duration * Campaign: Moederdag</td>
<td>0.146 **</td>
<td>REF</td>
<td></td>
</tr>
<tr>
<td>Duration * Campaign: Sale/solden januari</td>
<td>0.122 *</td>
<td>-0.024</td>
<td></td>
</tr>
<tr>
<td>Duration * Campaign: Sintwinkel</td>
<td>0.129 **</td>
<td>0.022</td>
<td></td>
</tr>
<tr>
<td>Duration * Campaign: Vaderdag</td>
<td>0.137 *</td>
<td>0.027</td>
<td></td>
</tr>
<tr>
<td>Duration * Campaign: Valentijn</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Product subgroup</td>
<td>Total Price discount *</td>
<td>Duration * Number of products</td>
<td>Total Price discount *</td>
</tr>
<tr>
<td>----------------------------------------</td>
<td>------------------------</td>
<td>-------------------------------</td>
<td>------------------------</td>
</tr>
<tr>
<td>Aromatherapie</td>
<td>-0.029</td>
<td>-0.000 **</td>
<td>-0.000 **</td>
</tr>
<tr>
<td>Gezichtsverzorging</td>
<td>-0.013</td>
<td>-0.000 **</td>
<td>-0.006</td>
</tr>
<tr>
<td>Haarverzorging</td>
<td>-0.015</td>
<td></td>
<td>0.040</td>
</tr>
<tr>
<td>Hand &amp; Voetverzorging</td>
<td>-0.008</td>
<td></td>
<td>0.040</td>
</tr>
<tr>
<td>Lichaamsverzorging</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mondverzorging</td>
<td>0.120 *</td>
<td></td>
<td>0.046 ***</td>
</tr>
<tr>
<td>Persoonlijke hygiene BPH</td>
<td>REF</td>
<td></td>
<td>REF</td>
</tr>
<tr>
<td>Persoonlijke Verzorging Geschenksets</td>
<td>0.005</td>
<td></td>
<td>-0.011</td>
</tr>
<tr>
<td>Scheren &amp; Ontharen</td>
<td>-0.006</td>
<td></td>
<td>-0.000 **</td>
</tr>
<tr>
<td>Schoonmaakartikelen</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Televisions</td>
<td>-0.040</td>
<td></td>
<td>0.009</td>
</tr>
<tr>
<td>Zelfbruiners</td>
<td>-0.046 ***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zonnebrandcreme</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duration * Product subgroup</td>
<td>0.102 ***</td>
<td>-0.017</td>
<td></td>
</tr>
<tr>
<td>----------------------------------------</td>
<td>-----------</td>
<td>--------</td>
<td></td>
</tr>
<tr>
<td>Duration * Product subgroup</td>
<td>0.013 ***</td>
<td>0.005</td>
<td></td>
</tr>
<tr>
<td>Duration * Product subgroup: Haarverzorging</td>
<td>0.117 ***</td>
<td>-0.018</td>
<td></td>
</tr>
<tr>
<td>Duration * Product subgroup: Hand &amp; Voetverzorging</td>
<td>0.127 **</td>
<td>-0.122 *</td>
<td></td>
</tr>
<tr>
<td>Duration * Product subgroup: Mondverzorging</td>
<td>0.177</td>
<td>-0.130</td>
<td></td>
</tr>
<tr>
<td>Duration * Product subgroup: Lichaamsverzorging</td>
<td>0.146 ***</td>
<td>REF</td>
<td></td>
</tr>
<tr>
<td>Duration * Product subgroup: Persoonlijke hygiene BPH</td>
<td>0.098 ***</td>
<td>-0.020</td>
<td></td>
</tr>
<tr>
<td>Duration * Product subgroup: Persoonlijke Verzorging Geschenksets</td>
<td>0.053 *</td>
<td>-0.076 ***</td>
<td></td>
</tr>
<tr>
<td>Duration * Product subgroup: Scheren &amp; Ontharen</td>
<td>0.086 **</td>
<td>-0.013</td>
<td></td>
</tr>
<tr>
<td>Duration * Product subgroup: Schoonmaakartikelen</td>
<td>0.169 ***</td>
<td>0.056</td>
<td></td>
</tr>
<tr>
<td>Duration * Product subgroup: Zonnebrandcreme</td>
<td>REF</td>
<td>-0.084 ***</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Total Price Discount * action type: amount off</th>
<th>-0.166</th>
<th>REF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total price discount * action type: Cheapest product free</td>
<td>-0.068 ***</td>
<td>0.787 **</td>
</tr>
<tr>
<td>Total price discount * discount type: daydeal</td>
<td>REF</td>
<td>0.809 **</td>
</tr>
<tr>
<td>Total price discount * discount type: percentage off</td>
<td>-1.495</td>
<td>REF</td>
</tr>
<tr>
<td>Total price discount * discount type: price off</td>
<td>0.012 **</td>
<td>0.559 **</td>
</tr>
<tr>
<td>Total Price discount * discount type: X for fixed Number</td>
<td>discount type: Amount off</td>
<td></td>
</tr>
<tr>
<td>Number * Discount type:</td>
<td>REF</td>
<td>0.566 **</td>
</tr>
<tr>
<td>-------------------------</td>
<td>-----</td>
<td>----------</td>
</tr>
<tr>
<td>Price off</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number * Discount type: X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>for fixed</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Table 21: Final promotion forecasting models*

REF = Reference period

Significance codes:

* 0.05
** 0.01
*** 0.001
Appendix 8  

Explanation of the chosen model

In this section, the chosen model will be further explained with the independent variables, significance and magnitude of the effect size. In the previous part is explained that model 2 is the best model according to the performance measures. In the explanation, a distinction according to the different kind of variables is made.

Metric variables

The price discount of the product during the promotion period positively affect the lift factor. This indicates that a higher price discount increases the lift factor. This result corresponds to the result of Blattberg (1995), who found that temporary retail price reductions substantially increase sales. The duration of a global in promotion has a negative impact on the lift factor. This corresponds to the research of Cooper et al. (1999), although here a different promotion policy is applied which keeps an item on promotion until inventory runs out. Another independent variable is the number of products in an action is positively affecting the lift factor, indicating that a higher number of products decreases the lift factor. This result corresponds with the results of Ailawadi (2006), who found a positive relationship between the number of products in a category that is on promotion and the lift factor.

Dummy variables

Variables which contain different groups are divided into dummy variables to measure the effect to the lift factor. Compared to metric independent variables, insignificant dummy variables cannot be deleted, since these variables are dependent on each other. It is therefore to choose the reference period in order that the most significant dummy variables can be found.

The kind of campaign, which is divided into different dummy variables, also affects the lift factor. All the dummy variables are measured against a reference variable, which in this case is the Huishoudweken oktober. Therefore, all conclusions can be drawn compared to this campaign. The lift factor is expected to increase when the campaign Black Friday, Electronica Deals April, or Mid Season Sale are held compared to the campaign Huishoudweken oktober. The other campaigns have a negative effect on the lift factor compared to the Huishoudweken oktober. Not all the campaigns are a significant predictor. Due to the other significant predictors these insignificant variables must be held in the model. Remarkable is the negative effect of the Sinterklaas and Christmas campaign. These campaigns show a less impact on the lift factor than the huishoudweken oktober. A possible explanation for this result can be the reference period, since the reference period of the Huishoudweken oktober corresponds with a less popular campaign than the other reference periods. Further research is needed with a corrected reference period to obtain more robust insights.

The action discount type has a big impact on the lift factor. Especially, the action discount type daydeal shows a big positive effect on the lift factor. Therefore, can be concluded that the daydeal is the best action discount type, which generates the highest lift factor. The last dummy variable is the product subgroup, which also impacts the lift factor. All the product subgroups will negatively affect the lift factor compared to the reference variable, Sunscreen. A possible explanation for the big effect of sunscreen on the lift factor can be by the weather. Due to the time range is this factor not included in the final model. Therefore real conclusions compared to weather cannot be drawn. Many variables have a significant impact. The product subgroup Televisions is not significant, therefore it is not possible to compare the results of the different product groups. This can be explained by the relatively low number of observations for the product group Televisions compared to the product group Personal Care & Home Care. Further research is needed to obtain more robust insights about the product group Televisions.
Interaction effect
In order to improve the performance of the model, interactions between variables were included. An interaction between the kind of campaign and the price discount was added, indicating that the price discount is dependent on the kind of campaign. Every campaign has their own price discount also depending on the reference period. Another interaction effect between the duration of a global in promotion and the kind of campaign is added. This indicates that every campaign has a specific duration. The third interaction effect that is added is the duration of a global in promotion and the product subgroup. This can be explained by the duration a global is in promotion which is measured by the duration an action takes. Each action contains a specific product group. The number of products in a campaign is also dependent on the product subgroup. This can be explained by the number of products, which is also measured by the number of products within an action. Another interaction effect is the number of products with the duration of a global in promotion, this has a really small significant negative effect on the lift factor. The effect is really small, since both duration and number of products has both another effect on the lift factor. Furthermore, this interaction can be distracted from the interaction effects productsubgroup * duration and productsubgroup * number, which makes the impact low. The last interaction effect is between the price discount and the duration of a global in promotion. This is the only interaction effect which is not significant and therefore is this interaction not included in the final model.
Appendix 9  Demand patterns

[Graphs showing demand patterns for different events and categories such as Black Friday | Cyber Monday 2016 (Personal Care & Home Care), Black Friday | Cyber Monday 2016 (Televisions), Bulk 10-daagse mei 2017 (Personal Care & Home Care), Bulk 10-daagse mei 2017 (Televisions), Elektronica Deals april 2017 (Televisions), Huishoudweken februari 2017 (Personal Care & Home Care).]
Appendix 10  Safety stock levels

Safety stock over 9 days

Safety stock over 10 days

Safety stock over 11 days

Safety stock over 12 days

Safety stock over 13 days

Safety stock over 14 days
Appendix 11  Adjusted forecasts