Aalto University
School of Science
Degree programme in Engineering Physics and Mathematics

Clearance price optimization of seasonal products

Bachelor's Thesis
05.05.2019

Emil Nyman

Aalto University
School of Science

Aalto University, P.O. BOX 11000, 00076 AALTO
www.aalto.fi

Abstract of bachelor's thesis

Author Emil Nyman

Title of thesis Clearance price optimization of seasonal products

Degree programme Engineering Physics and Mathematics

Major Mathematics and Systems Sciences **Code of major** SCI3029

Supervisor Prof. Ahti Salo

Thesis advisor(s) M.Sc.(Soc.) Erkka Saarinen

Date 05.05.2019 **Number of pages** 4 + 34 **Language** English

Abstract

In retail, products whose sales occur mainly within a certain period of the year are called seasonal products. Forecasting the sales of these products is hard and thus a retailer may hold a suboptimal inventory level when the selling season is coming to an end. It is important to sell the inventory to make room for new products by the end of the selling season while gaining the maximum revenue possible. This is usually done during clearance periods where the sales quantity of a product is controlled by adjusting its selling price. The objective of this thesis is to review the literature on clearance pricing and formulate a model to optimize the price for the clearance period.

First, we developed a multiple linear regression sales forecasting model which considers the most important parameters from literature in forecasting seasonal products. The performance of the forecast model was tested on transactions data of highly seasonal outdoor activity peripheral products sold by a large European retailer. The optimal price is then obtained by determining the price with which the forecast of our model equals to the quantity which is to be sold by the end of the selling season.

The results of our model were inconsistent and varied from product to product, resulting in lower revenue compared to the revenue from the actual sales quantities and prices. In this thesis, we assumed that the price elasticity of demand decreases exponentially, and we optimized only one price for the clearance period. Future research would benefit from research into the time dependency of the price elasticity of demand as well as optimizing multiple price points for the clearance period.

Keywords clearance sales, price optimization, regression, seasonal products

Aalto-yliopisto
Perustieteiden
korkeakoulu

Aalto-yliopisto, PL 11000, 00076 AALTO
www.aalto.fi

Tekniikan kandidaatintyön tiivistelmä

Tekijä Emil Nyman		
Työn nimi Sesonkituotteiden poistomyynnin hintaoptimointi		
Koulutusohjelma Teknillinen fysiikka ja matematiikka		
Pääaine Matematiikka ja systeemitieteet		**Pääaineen koodi** SCI3029
Vastuuopettaja Prof. Ahti Salo		
Työn ohjaaja(t) VTM Erkka Saarinen		
Päivämäärä 05.05.2019	**Sivumäärä** 4 + 34	**Kieli** englanti

Tiivistelmä

Vähittäiskaupan tuotteita, joiden myynneistä suurin osa keskittyy tiettyyn aikajaksoon, kutsutaan sesonkituotteiksi. Näiden tuotteiden myyntimäärän ennustaminen on vaikeaa. Näin ollen vähittäiskauppiaalla saattaa myyntikauden loppua kohden olla varastossa suboptimaalinen määrä tuotteita. Varasto halutaan yleensä myydä tyhjäksi myyntikauden loppuun mennessä, jolloin uusille tuotteille vapautuu lisää tilaa ja kaikki mahdollinen myyntituotto on saavutettu. Tämä toteutetaan käytännössä poistomyynneillä, joissa myyntimäärää ohjataan säätämällä tuotteen myyntihintaa.

Työn tavoitteena on muodostaa sesonkisten tuotteiden myyntiennustemalli, jonka avulla optimoidaan korkein mahdollinen hinta, jolla tuotteen varasto saadaan myytyä loppuun myyntikauden loppuun mennessä. Mallin rakentamiseen hyödynnetään poistomyynnin hinnoittelua ja hintaoptimointia käsittelevää kirjallisuutta.

Myyntimäärän ennustamiseksi kehitettiin usean muuttujan lineaarinen regressiomalli, joka huomioi kirjallisuudessa mainitut tärkeimmät parametrit sesonkituotteiden myyntimäärää ennustaessa. Ennustemalli sovitettiin ja sen sopivuutta testattiin eräältä suurelta eurooppalaiselta vähittäiskauppiaalta saadulla sesonkisten ulkoliikuntatuotteiden transaktiodatalla. Optimihinta määritettiin puolitusmenetelmällä, jolla myyntimäärän ennuste asetettiin yhtä suureksi halutun myyntimäärän kanssa.

Mallin tulokset eivät olleet johdonmukaisia ja ne vaihtelivat tuotteittain. Optimihinnat johtivat keskimäärin huonompaan myyntitulokseen kuin tapahtuneilla myyntimäärillä ja hinnoilla. Tässä työssä oletimme, että kysynnän hintajousto vähenee eksponentiaalisesti ja sen lisäksi optimoimme ainoastaan yhden hinnan koko poistomyyntijakson ajaksi. Työssä ehdotetaan lisätutkimusten kohteeksi kysynnän hintajouston aikariippuvuutta sekä poistomyyntijakson jaottelua useampaan hinnoittelujaksoon.

Avainsanat poistomyynti, hintaoptimointi, regressio, sesonkiset tuotteet,

Contents

1 Introduction

In retail, products which are sold only during a certain period are called seasonal products. These products have a life cycle that typically ends at a certain time, whereafter they have little to no value as they will be replaced with a newer model in the next season. Typically the inventory of a seasonal product is to be depleted by the end of its selling season. One way to achieve this is to price products optimally during the season so that the whole inventory is sold while gaining maximum profit. However, perfectly optimizing the price and forecasting the demand of a product during its whole life cycle is really hard, which can lead retailers to build excessive inventories to minimize potential loss of sales if the inventory is sold out too quickly [Sachs, 2015]. The end-of-season excessive inventories are usually sold through clearance sales where radical price markdowns are made to ensure no inventory is left over. Optimizing the amount to markdown in a clearance sale can not only help retailers clear the inventory but also increase profits [Smith, 2009].

Because price optimization relies heavily on the forecast, it is important that forecasts are as good as possible. While the historical sales quantities of a product is the most commonly used parameter in forecasting sales quantities, it is not always enough to forecast the sales quantities of discounted seasonal products. Frequently used in literature, the price elasticity of demand is an important factor when forecasting discounted products and it is. In their study, Soysal and Krishnamurthi [2012] also take into consideration the purchase timing while Caro and Gallien [2012] and Smith [2009] like to also consider the quantity of products left in stock.

In this thesis, we formulate a regression model to forecast the sales quantities of seasonal products which combines different parameters from the literature. The model is used on real life transactions data from a European retailer. We determine a single optimal clearance price for each product so that the inventory is cleared at the end of the selling season. The results are evaluated against the actual sales figures, conclusions are made and lastly, we assess future research.

2 Background

Retailers face a more complex and harder to navigate market environment than ever before. Already established in the 1980s, the the portion of markdown sales in overall sales were increasing in a fast pace [Pashigian, 1988].

Multiple factors like ever-expanding choices, rapid changes in consumer behaviour and increasing competitive pressure affects retailer's ability to sell products at full price. According to a recent survey of US retailers, about 40% of non-food grocery sales are made during markdowns. Even if discounting and promotions are getting more frequent and with shopping events such as Black Friday accounting for larger portions of the total promotional sales, clearance sales continues to be one of the most important promotional periods [Weinswig et al., 2019].

2.1 Seasonal products

Products which are sold only during a certain time period are called seasonal products and that time period is called the selling season which can be of varying length. A short selling season is considered to be a few weeks, say, for example Easter themed products, and a long selling season is considered to be for example the whole summer. Some seasonal products such as Easter eggs are sold only during a limited period of time whereafter they are discontinued. Then again, other seasonal products such as biking peripherals are sold throughout the year, but are considered having a selling season from Spring to Fall, as the sales outside this selling season is minimal compared to high-season sales. Seasonal products also have different demand patterns and characteristics compared to non-seasonal or consumer packaged goods. The demand of a seasonal product tends to peak shortly after the season start date and then decreases gradually towards the end of the season [Caro and Gallien, 2012, Smith, 2009].

A seasonal product can provide value only once (e.g. airline ticket or fireworks) or over the whole season (e.g. skis or swimsuits), in which case the customer can use the product more the earlier in the season it is purchased. Durable seasonal products such as summer clothes or winter accessories are typically liquidated with the use of promotions as the season ends to make space for the next season's products and also because the salvage value of seasonal products after their season is typically very low or can even be negative [Soysal and Krishnamurthi, 2012]. Hence, the customer is faced with a trade-off between buying the products at high price while getting the most value out of it and getting the product at a cheap price but then have a risk of being rationed [Nocke and Peitz, 2007]. This means that the product is less sensitive to price changes as the season progresses, meaning that deeper markdowns are needed for the markdowns to be as effective [Gupta et al., 2006, Caro and Gallien, 2012, Soysal and Krishnamurthi, 2012, Smith and

Agrawal, 2017]. This effect is also referred to as time discounting.

2.2 Clearance sales

Clearance sales are a type of promotion used as an inventory management tool by many companies. In clearance sales, the products are sold at a considerable discount relative to their usual retail prices. The clearance period is defined by the start of the aforementioned markdown period and the end date of the product's selling season when the remaining inventory is scrapped and is replaced with new products [Zhao and Zheng, 2000]. For example, clearance sales in winter accessories usually starts when the snow starts melting and summer approaches.

There are many reasons why retailers use clearance sales, but the main reason is to get rid of excessive stock. Having an inventory buffer is a way to reduce uncertainty, but there is rarely only one reason that leads to having excess inventory and often it is also caused by several reasons [Zaarour et al., 2016]. For instance, the actual sales may have been lower than anticipated or that a manufacturer had released a newer substitute product which renders the old product obsolete. Overstock could also originate from weather conditions, competition or a shift in trends.

The combination of lowering prices radically and the fact that forward looking customers are expecting end-of-season sales increases store traffic. Forward looking customers anticipate clearance sales at the risk of the particular product being sold out quickly and therefore, some customers prefer to buy the product at a higher price before the sales start [Nocke and Peitz, 2007, Soysal and Krishnamurthi, 2012]. Discounted prices also offer access to goods at a lower price which can attract new customers and could consequently even increase the customer base.

The ultimate goal in clearance sales is to clear the inventory to make room for newer products while generating the maximum profits. There are two main decisions to make when planning clearance sales: when to start the markdown period and how to set the correct price. During the clearance period, there is no restocking and the markdowns are permanent, which means that the price cannot be raised from its current price. The initial markdowns should be deeper than the customers are accustomed to having while avoiding excessive markdowns at the end of the clearance period [Smith, 2009, Soysal and Krishnamurthi, 2012].

2.3 Price elasticity of demand

Price elasticity of demand is commonly used to evaluate the sensitivity of customer behaviour when prices are changed. It is expressed as a dimensionless value of the ratio of the percentage of change in demand (Q) with respect to the percentage of change in price (P): $\beta = \frac{\Delta Q/Q}{\Delta P/P}$ [Zaarour et al., 2016]. The price elasticity can either have a linear or an exponential relationship with demand. Linear effect of the price elasticity is calculated by multiplying the price change ΔP with the elasticity value E: $(\Delta P - 1) \cdot \beta$. Consequently, the exponential effect of price elasticity is calculated by raising the price change to the power of the elasticity value so that ΔP^{β} . The main difference of these two relationships is the rate the effect of price elasticity grows. For example, when the price elasticity value β is -2.0 and a 30% decrease is made in the price, the linear effect would suggest a 60% increase in sales whereas the exponential effect would suggest a 100% increase in sales.

The sign of the elasticity value indicates the direction of change. It is typically negative due to the inverse relationship between price and quantity demand. When the price elasticity is negative, lowering the prices will increase sales and vice versa. However, if the price elasticity is assumed to have a linear relationship with demand and the elasticity value is larger than -1, then decreasing the price will lead to a lower profit even when selling a lower quantity with the original price. A product with such a price elasticity is said to have a relatively inelastic demand. When a product has an elasticity lower than -1, the profit growth is positive compared to the original price and it is thus said to have an elastic demand. See Table 1 for the full interpretation of the price elasticity coefficient. When the relationship of the elasticity value and demand is exponential, the threshold value of the elasticity for positive profit growth depends on the discount and the elasticity value. Situations where the elasticity is positive and a higher price leads to increased demand are exceptional and are disregarded.

Table 1: Interpretation of the price elasticity coefficient β when the relationship with demand is assumed to be linear.

$\beta = 0$	Perfectly inelastic demand
$0 > \beta > -1$	Relatively inelastic demand
$\beta = -1$	Unitarily elastic demand
$\beta < -1$	Elastic demand

Estimating the value of the price elasticity is important in forecasting the demand of products with discounts. It is known that the elasticity of a product is not constant and changes during the selling season, becoming increasingly inelastic when the selling season approaches its end [Gupta et al., 2006, Caro and Gallien, 2012, Soysal and Krishnamurthi, 2012, Smith and Agrawal, 2017]. Thus, the estimation of the price elasticity can be a challenge especially when there is not much data on earlier price changes resulting in corresponding changes in demand. Both Caro and Gallien [2012] and Smith et al. [1994] used clothing in their pricing studies, establishing that using a constant price elasticity value for product groups yielded better values than using elasticities for every product or price-category. However, clearance pricing studies including cosmetic items [Zaarour et al., 2016] and typical retail products with inventory dependent demand [Smith and Agrawal, 2017] suggested the use of different price elasticities for every product-location combination.

2.4 Inventory effect

The level of remaining inventory has a significant influence on demand both positively and negatively [Smith, 2009, Smith and Agrawal, 2017]. A large inventory that is visible to the customer can appear more attractive to the customer and thus boost demand. Adequate representation or, in other words, having a complete selection of sizes and colours in store is needed in order to not have a negative inventory effect on the demand. When the level of inventory becomes lower than the adequate representation level, the demand usually declines significantly and is called the broken inventory effect [Caro and Gallien, 2012]. The inventory level can vary from location to location and thus the demand of the same product can differ depending on the location. Relocating inventory from locations with plenty of inventory to those with low inventory can be the most efficient approach when trying to clear company wide inventories, but this idea is usually discarded due to high transportation costs.

3 Literature review

Zaarour et al. [2016] present a model using a multi-period nonlinear programming that maximizes profits from the discounted items:

$$\text{Maximize} \sum_{t=1}^{T} \alpha P_t^{\beta+1} \tag{1}$$

$$\text{Subject to} \sum_{t=1}^{T} \alpha P_t^{\beta} \leq I \ , \tag{2}$$

where α is an estimated positive constant, P_t is the price of the product, β is the price elasticity, I is the initial inventory and T is the phase-out time horizon.

Zaarour et al. used data on obsolete cosmetic items sold by a large US retailer which has over 6000 stores over the United States. The data contained weekly sales data including price from a 54-week period. After clustering the data with a k-means cluster method, Zaarour et al. used regression analysis to find out the best fitting model which was a power regression model. The demand-price model (2) is heavily dependent on the price elasticity which, in lack of price sensitive data, is hard to estimate. However, every SKU in this research had price fluctuations during the review period, which enabled good estimation of the price elasticities. The improvement in profits by using the developed model ranged from 8.6 to 22.6 percent, a considerable increase compared to the retailer's old strategy.

In an empirical analysis of the demand dynamics of seasonal goods, Soysal and Krishnamurthi [2012] developed a dynamic demand model where consumers are strategic and heterogeneous. The model describes the consumers' decision process and considers the consumers' responsiveness to price changes, decreasing demand as well as the changing market composition during the season. The optimal purchase timing of a customer is determined by solving a dynamic programming problem in which the total utility of the product at a certain purchasing point is defined considering the purchase price, time of season, price elasticity, seasonality, availability as well as demand shocks. The model was estimated with sales data on women's coats category from an apparel retailer for 105 different SKUs. Soysal and Krishnamurthi use counterfactual experiments rather than solving the optimization problem to investigate implications of changes in the pricing policy. The results show that not accounting for the changing utility or ignoring customers' expectation of the future availability may lead to incorrect demand forecasts and lower price elasticities. Soysal and Krishnamurthi also establish that small and early markdowns yield better profits than later markdowns or deep early ones. They also emphasize that the correct timing of early markdowns is critical as they can have significant negative profit implications because the market composition is more sensitive to price changes.

Caro and Gallien [2012] developed and integrated a clearance pricing optimization for Zara, a fashion retailer which, before the implementation, had relied on manual and informal decision-making process for determining markdowns. Zara is a fast-fashion retailer which sells articles with short life-cycles that are almost never discounted before the clearance period. Thus, there is almost no price sensitive data available at the beginning of the clearance period which renders price elasticity estimation useless in optimizing the first markdown price. This is completely the opposite of the study by Zaarour et al. [2016] where the most important factor was to estimate the price elasticity parameter.

Through lots of experimentation, Caro constructed a forecasting model which is a particular instance of the model presented by Smith and Achabal [1998]. To overcome the lack of price sensitive data, a two-stage estimation procedure was used. The model consists of the past sales quantity, the age of the article, the broken assortment effect and the price discount. Caro did not consider cross-product dependencies, impact of competition nor strategic customers because of the additional complexity they impose. A price optimization model was built alongside the forecast model with multiple constraints to comply with the complex clearance price policies of Zara. The models were tested in a live pilot in two entire countries and measured with multiple performance metrics, the most important ones being the realized income and the percentage of stock sold. The pilot resulted in a significant financial impact of a 6 % increase in clearance sales revenue, which company-wide corresponded to $90 million in 2008.

Smith [2009] formulates a deterministic model for clearance pricing which assumes that the sales rate depends explicitly on price, seasonal variation and like in [Caro and Gallien, 2012] also inventory level. Price accounts for the price elasticity of demand and seasonal variation captures sales spikes during special season specific periods such as Christmas and declining demand towards the end of the selling season. The inventory level adds the broken assortment effect to the model which means a too low inventory may decrease sales. Here, the minimum inventory level of a product is defined as the threshold from where the effects steps into place. Smith does not consider competition nor demand uncertainty, because these would complicate the problem in a great extent. As the clearance period is rather short, Smith also overlooks inventory costs and time discounting.

Similar to [Smith, 2009, Smith and Achabal, 1998, Caro and Gallien, 2012] inventory levels are taken into consideration in a study on markdown optimization by Vakhutinsky et al. [2012]. In contrast to these studies, Vakhutinsky considers a revenue maximizing SCAN*PRO demand model for which price recommendations and inventory trajectories are obtained. No time discounting is considered due to the short clearance period and the limited set of price values are considered as usually the price is selected from a pre-defined price ladder. The demand model is assumed to be a multiplicative and separable function with three components: price- and inventory-dependent components as well as a time-dependent seasonality coefficient. Vakhutinsky found that a constant price elasticity value and a power law function for the inventory effect gave the best fit for the data. The seasonality coefficient is estimated from historical sales data and adjusted for specific dates where higher sales are expected.

Vakhutinsky introduces an optimal continuous price control model and a more realistic model for markdowns when price and time are discrete. The latter model was tested on two-year weekly sales history of one hundred products from a large fashion retailer. The sales volumes of the products were aggregated over different sizes and colours. The study did not entail financial results of the testing but, instead, they compared the regret, optimality and calculation speeds to the current markdown approach used by the retailer.

An interesting approach is taken by Meijer and Bhulai [2013] who applied survival analysis to determine the optimal pricing for a period. Cox regression is used to estimate the proportional hazard which depicts the instantaneous rate of failure at a given time which in retail context means the rate of selling the product at the given time assuming that the product is available. The survival model estimates the probability of surviving longer than a specified time which again in retail means the probability that the product has not yet been sold at the given time. The model can then be used to calculate the estimated price difference needed to achieve the desirable sell-through at a given time. The model was tested on a data set of almost 6000 different T-shirts sold in a half years' time in a department store in the Netherlands. The survival rates at different time points were estimated for each size and colour of every product and the optimal prices were estimated. The average return with markdown was 3 percent higher compared with the returns of no markdown and write-off.

4 Research question and methods

Pricing during clearance sales raises many questions ranging from the date to start the clearance markdown period to the amount of markdown adjustments allowed during the clearance period. In this thesis, we approach the clearance pricing problem in the context of the following rather simple situation. The clearance period starts from a predefined date, which is well after the peak selling point of the season. We then define one price for each product for the rest of the selling season from the said date onward so that the remaining inventory is depleted by the end of the selling season.

According to Smith [2009], a successful clearance pricing optimization system has three components: a sales forecasting model, a clearance price optimization algorithm and financial performance metrics for measuring the effectiveness of the system. In this thesis, we mostly focus on the first two components. First, we formulate a regression model for predicting future sales and, second, we optimize the clearance selling price so that the predicted sales totals to the inventory level on the starting day of the clearance period.

4.1 Data

The data consists of sales transactions data from highly seasonal products sold by a European retailer. The products are peripherals for outdoor activities which have visible selling seasons but are nonetheless sold during the whole year and for each product we get two full seasons of sales. A selling season is defined as a whole year which starts from week 12. The data is split into two parts: the training data from week 11 of year one to week 48 of year two and the data for validation purposed from week 49 of year two to week 10 of year three. We assume that the present time is at the end of week 49 of year two. An example of the sales quantity and average sales prices of a product used can be seen in Figure 1 where the green line represents the point from where we forecast the sales.

The transactions have been aggregated product-wise to the weekly level and they contain the sales quantities as well as the sales prices of the products which have both been rescaled to a scale from 0 to 1 for privacy reasons. Negative transaction rows are not considered and have been filtered out from the data. The sales prices of the products are attached to the sales quantities meaning that the average prices are only calculated on products which are

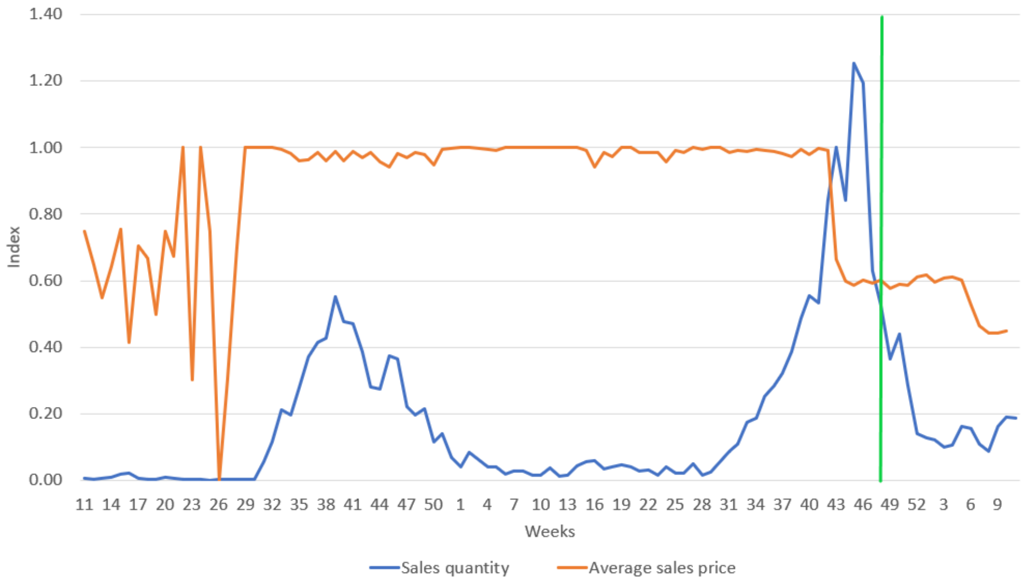

Figure 1: Rescaled sales quantity and average sales price over a period of two 52-week seasons

sold, which does not fully represent the average in-store sales prices at the aggregate-level. In addition, the retailer has multiple promotions during each product's selling season which often overlap and are only active in few locations. Differing markdown promotions at various locations can skew the average sales prices at the aggregate-level. This can clearly be seen in Figure 1. However, the average sales price is used for clarity.

4.2 Fitting the regression model

To predict future sales, we formulate a linear regression model based on the data and the literature review.

The general multiple linear regression model can be written as

$$Y = \beta_0 + \beta_1 X_1 + \cdots + \beta_p X_p + \varepsilon = \beta_0 + \sum_{i=1}^{p} \beta_i X_i + \varepsilon \ , \tag{3}$$

where Y is the response, $X_1, \ldots, X_p$ are regressors variables and $\beta_1, \ldots, \beta_p$ are the corresponding regression coefficients to be estimated and ε is a zero-mean random error component with unknown variance[Montgomery et al., 2012]. A regression problem starts with a collection of predictors which are derived from the available data. The regressors can among others be simple predictors, transformed predictors, polynomials or interactions and combinations of predictors.

- Simple predictors are used to directly explain the data where the model is fitted.
- Predictors are transformed when the original predictors do not give a reasonable approximation for the data to which the model is fitted. A common example of a predictor transformation is to use the predictor in log scale, but also other methods such as differentiation or normalization is used. Transformations greatly expand the range of problems that can be summarized with linear regression.
- Polynomial regressors are used when the problem includes curved functions. For example, to fit a quadratic polynomial both predictors X_p and its square X_p^2 are used as regressors.
- Interactions and combinations are often useful for explaining joint effects of the predictors which can not be explained by separate predictors.

Regression coefficients of a multiple linear regression model are commonly estimated with the least squares method. That is, estimating the regressor coefficients $\beta_1, \ldots, \beta_p$ so that the sum of the squares of the differences between the observed responses and the corresponding fitted values called residuals $e_i = Y_i - \hat{Y}_i$ are minimized. This method can be generalized as

$$\min_{\beta_0 \ldots \beta_k} \sum_{i=1}^{n} \varepsilon_i^2 = \min_{\beta_0 \ldots \beta_k} \sum_{i=1}^{n} \Big(y_i - \beta_0 - \sum_{j=1}^{k} \beta_j x_{ij} \Big)^2 \ . \tag{4}$$

The results from this equation are least-square estimators $\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_k$. The fitted regression model corresponding to the levels of regressor variables is

$$\hat{Y} = \hat{\beta}_0 + \sum_{j=1}^{k} \hat{\beta}_j X_j \ . \tag{5}$$

To fit our regression model, we used RStudio where the lm-function uses the method of least squares to estimate the best regressor variables.

A manual stepwise variable selection process was used to find the best model fitted on the training data. In this process, regressor variables are added to the model one-by-one and after each step all candidate variables in the model are evaluated to see if they still provide improvement in the model fit, or in other words, are still significant. If a non-significant variable is found it is removed from the model [Yan and Su, 2009].

At each step, the candidate variables in the formulated regression models were evaluated based on the following methods:

- **Graphical evaluation of the fit**
  An important factor is how well the estimated regression model fits the observed data with the given regressors. Any errors in the data or regressors can easily be seen when plotting the fit on top of the observed data.

- **The coefficient of determination**
  More commonly referred as the R-squared value is the proportion of variation explained by the regressors. In other words, it measures the "goodness" of the fit. The value of the R-squared is $0 \leq R^2 \leq 1$ and values closer to 1 imply that most of the variability of the response is explained by the regression model. This parameter should not be the only parameter to observe when fitting a model as it is possible to make R^2 larger just by adding more variables to the model, which usually leads to overfitting.

- **Significance of regression coefficients**
  The p-value of an estimated regression coefficient indicates whether the coefficient is significant in the model. In other words, a low p-value indicates that changes in the regressors value is related to changes in the response variable meaning that the regressor is a meaningful addition to the model. A significance level of 0.05 was used to determine if the regressor is significant or not.

- **Variance inflation factors**
  The variance inflation factor or VIF is used to measure multicollinearity in the model. Multicollinearity is defined as the existence of strong correlations between the regressor variables which can arise from the way of collecting data or by using regressor which are naturally correlated. This means it is hard to vary one variable by holding other variables constant which renders the individual coefficients to be less useful. A model which exhibits multicollinearity can fit a dataset very well, but it reduces the effectiveness of regression analysis. The variance inflation factor quantifies the effect of multicollinearity in a model by a value that is relative to the overall fit R^2 of the fit. VIF values

larger than $\frac{1}{1-R^2}$ imply stronger relations among the regressors than their relation to the response [Freund et al., 2006].

- **Residual analysis**
  There are major assumptions in linear regression analysis: The error terms ε have a zero mean, a constant variance σ^2, they are uncorrelated and they are normally distributed. Standard summary statistics rely heavily on these assumptions and do not ensure adequacy on their own. Instead, it is necessary to consider the validity of these assumptions by diagnostic methods. There are multiple methods to perform residual analysis, but we will use graphical analysis [Montgomery et al., 2012].

In this thesis, the regression involves seasonal time series data which means the regressors are time-oriented. Time series data and especially seasonal data sets exhibit some type of autocorrelated structure at different time periods. Thus, the usual assumption of uncorrelated or independent errors which is made for not time-dependent regression data is usually inappropriate for time series data. Autocorrelation can be detected by the graphical residual analysis or by using various statistical tests for example the Durbin-Watson test [Montgomery et al., 2012]. The data used has a season of 52 weeks; however, the data used in fitting is only roughly one and a half times the season length and the data used for predicting is less than a third of the season length. Therefore, we will not fully consider autocorrelation, nor will we fully rely on traditional regression analysis as much as we would in a time-independent case. Rather we use these methods to guide us in the right direction.

4.3 The sales forecast model

We constructed a sales forecast model based on the literature review, available data and regression analysis. This model considers two important characteristics of seasonal products demand. First, consumers' responsiveness to prices change during the season and as noted in the literature review, the price sensitivity of demand decreases over time, diminishing the effect of price markdowns as the selling season comes to end. Second, as the selling season is usually also the season when the product is used, the highest sales in seasonal products are seen in the beginning of the selling season which then decrease exponentially towards the end of the season as can be seen in 1. Despite several past studies using the broken inventory effect as a component in their forecast model, we chose to not use it, because the sales volumes were really high compared to the minimum inventory and the impact would have been minimal.

The model constructed for the sales quantity S_w of week w is

$$S_w = \beta_0 + \beta_{sales} \times S_{w-52} + \beta_{price} \times S_{w-52} \times \left(\frac{p_w}{\bar{p}}\right)^{E_w} + \beta_{distance} \times d_w \qquad (6)$$
$$+ \beta_{sales,distance} \times S_{w-52} \times d_w + \varepsilon_w \ ,$$

where S_{w-52} is the sales quantity for week w last year, p_w is the price of the product on week w, $\bar{p}$ is the normal retail price of the product, E_w is the price elasticity of demand for week w and d_w is the distance of week w from the week of peak sales. The β parameters are estimated with the model: β_0 is a constant which measures the average level of sales without any any other component, β_{sales} is the coefficient for last year's sales, β_{price} is the coefficient of the estimated price elasticity, $\beta_{distance}$ is the coefficient for the distance from the peak sales, $\beta_{sales,distance}$ is the coefficient for the relation of distance and sales from the peak week sales of last year and ε_w is the error variable of week w.

The first regressor in model (6) is a predictor of the sales last year at the same time point as the current week S_{w-52}. We assume that the season sales profile as well as the start and end date of the season stay the same every year. Using last years' sales is the simplest way to get a reasonable estimate of the forthcoming sales which is why it is chosen as the first regressor when forming this model. This way of forecasting is more commonly known as "naïve forecasting" where the next period's level of sales will be the same as that one of the preceding period [McLaughlin, 1983].

The second regressor $S_{w-52} \times \left(\frac{p_w}{\bar{p}}\right)^{E_w}$ is an interaction of last years' sales quantity and the effect from the price elasticity of demand. The relation between the current selling price and the original selling price raised to the power of the price elasticity gives us the coefficient for the effect of price elasticity. We assume that the price elasticity has an exponential relationship with demand. Last years' sales are multiplied by this coefficient, which enables this regressor to capture the effect of price change to the sales quantity. The price elasticity is calculated from the ratio of change in demand of this year and last year with respect to the ratio of change in price of this year and last year which can be defined as

$$E_w = \frac{(S_{w-52} - S_w)/S_{w-52}}{(p_{w-52} - p_w)/p_{w-52}} \ . \qquad (7)$$

This equation is not defined when there has not been price changes between this and last year or if there has not been any sales last year. We will, however, discard this small issue by inserting a zero where the price elasticity can not be calculated. With this equation, we get the actual price elasticity for every week of the training data. There are many studies which suggest that the price elasticity decreases over time towards the end of the season, but none of them explicitly defines how the elasticity behaves. We choose an exponential approach where the calculated price elasticity decreases exponentially to reflect the less price sensitive customer base towards the season end. This is done by multiplying the mean of the calculated price elasticities by a rescaled and adjusted quadratic curve fitted to last years' sales data from its peak sales point onward. The quadratic fitting is done with a simple regression model

$$C_t = \beta_0 + \beta_1 t + \beta_2 t^2 \ , \tag{8}$$

where t is the time in weeks from the highest point of sales [Freund et al., 2006]. Thereafter the quadratic fit is rescaled to a scale of 0 to 1 and adjusted so that after multiplication the price elasticity doesn't go under a fourth of its starting value.

The third regressor d_w represents the effect of distance of week w from the peak sales week of last year. This is because as well as with the last years sale quantities we assume that the week of the peak sales is the same every year. The effect of distance is the same rescaled fitted quadratic curve used in the price elasticity estimation, but it is not adjusted to a minimum limit. The fourth regressor $S_{w-52} \times d_w$ is an interaction between the last years' sales and the effect of distance. This regressor captures the effect of exponentially diminishing sales towards the end of the season which is an important characteristic of seasonal products demand.

4.4 Predicting sales and optimizing the selling price

After fitting the regression model (6) to the training data it is ready to be used to predict sales. Towards this end, we need to know the future values of the regressors of which only the price p_w is unknown. The price is the parameter that is to be optimized so that the forecasted sales from the model are equal to the inventory at the start of the forecasting period. The aim is to optimize one price for the rest of the selling season which makes the parameter p_w a constant during all the weeks w during the forecasting period and therefore it is from now on referred as p.

The forecasting period length is 15 weeks denoted by $w = 1, \ldots, 15$ where w is the period to be forecasted. The optimization problem gives an equation where we minimize the absolute value of difference between the starting inventory level I and the forecast. As we do not account for storage costs and salvage prices and because of behaviour of the price elasticity of demand the resulting price from this optimization will also be the price which yields to most revenue possible. The optimization problem is

$$\min_{p} \left| I - \sum_{w=1}^{15} S_w \right| , \tag{9}$$

which can be expanded to

$$\begin{aligned} \min_{p} \Bigg| I - \Bigg(& \sum_{w=1}^{15} \beta_0 + \sum_{w=1}^{15} \beta_{sales} \times S_{w-52} + \sum_{w=1}^{15} \beta_{price} \times S_{w-52} \left(\frac{p}{\bar{p}} \right)^{E_w} \\ & + \sum_{w=1}^{15} \beta_{distance} \times d_w + \sum_{w=1}^{15} \beta_{sales,distance} \times S_{w-52} \times d_t \Bigg) \Bigg| . \end{aligned} \tag{10}$$

We use the actual sales quantity during the forecast period as our starting inventory level I which is to be depleted. In other words, we minimize the forecast error by finding the optimal the selling price p. This is because the acquired data did not withhold the desirable level of ending inventory. This also simulates the situation where the inventory needs to be completely depleted, but instead of the objective being zero inventory we have an inventory level to achieve.

To measure the quality of our forecast, we use two commonly used metrics: forecast bias percentage and mean absolute deviation. The forecast bias percentage is the ratio between the sum of the forecasted sales quantity and the sum of the actual sales so the closer the value is to 100% the better the forecast is in terms of the total forecast quantity. The mean absolute percentage error (MAPE) is the average percentage error in the forecast compared to the actual sales so the closer it is to zero the better the forecast is [Relex Oy, 2017].

To optimize the selling price for the forecasted period we use the bisection method using a looping function in Rstudio. The bisection method works so that we have an interval $[a, b]$ in where the optimal price p is. The algorithm starts with a candidate price value p_c which is in the range of $[a, b]$. The difference between the forecasted sales quantity calculated with the candidate price p_c and the inventory to be sold is evaluated. If the forecast is larger than the amount that needs to be sold, the price needs to be raised and, vice versa the price needs to be lowered if the forecast is lower than the amount to be sold. Thus, the interval for the next loop is chosen to be either $[a, p_c]$ or $[p_c, b]$ depending on in which way the price needs to be adjusted so that the forecast equals the inventory to be sold. The algorithm then chooses the next candidate price from the middle of the new interval $p_c = \frac{a+p_c}{2}$ or $p_c = \frac{p_c+b}{2}$. The loop is repeated until the difference between the sales forecast and the inventory to be sold is less than 1.0.

5 Results

5.1 Model analysis

The regression analysis assumed the errors ε are independent, normally distributed and have equal variances. We start by graphically analyzing the residuals of the fitted model (6) to see if these assumptions are valid. Then, we analyze the model by evaluating the fit on the data and evaluating the model with few other statistical tests. We summarize the regression analysis of only on the model fitted on product 1, because the results for other products were more or less the same. Plots used in the graphical analysis of other products are in the appendix A.

Substantial departures from the straight line in the normal probability plot indicates that the distribution is not normal. The normal probability plot in Figure 2 the residuals are heavy-tailed, meaning that they have extreme positive and negative residuals. This implies that the residuals are not totally normally distributed, as shown in the residual histogram in Figure 3. The extremities originate from sales spikes from the lowered sales price which can be seen in Figure 1. This tells us that our model is not fully able to estimate the sales spikes.

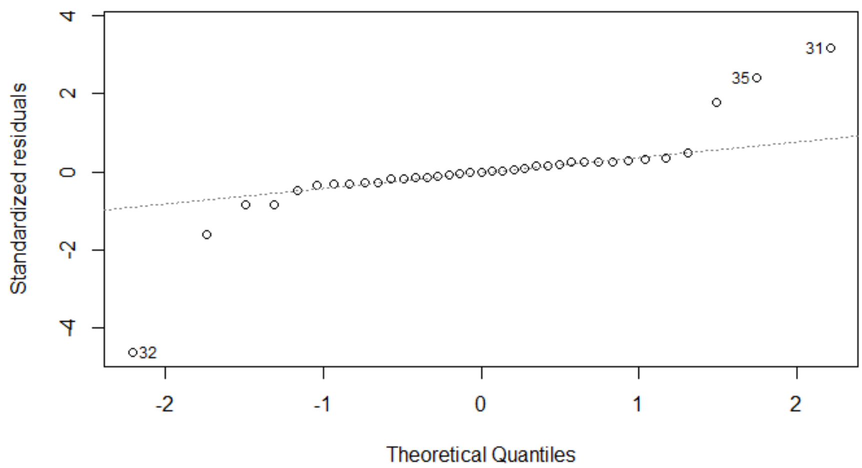

Figure 2: Normal probability plot of residuals in the model for product 1.

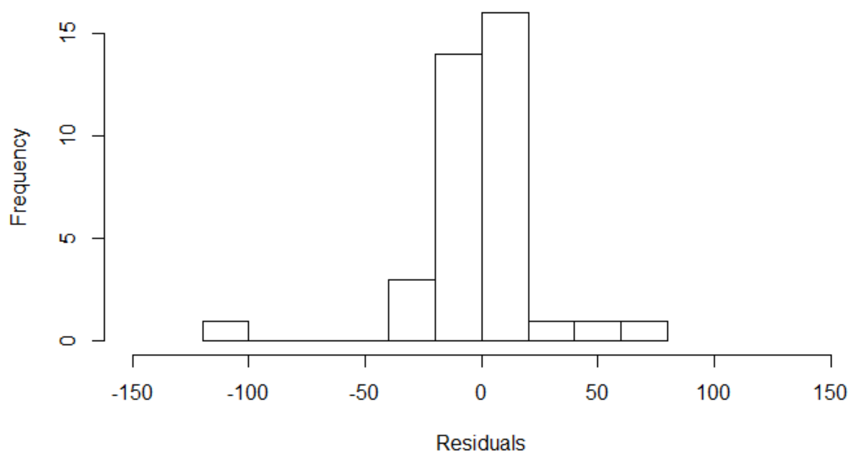

Figure 3: Histogram of residuals of the fitted model for product 1.

A residual versus fit plot is used to detect non-linearity, unequal error variances and possible outliers. This plot is created by taking a scatter plot of residuals of the fitted model versus the fitted values of the model. In the plot created from our model in Figure 4 we can see a slight curvature or an outward-opening funnel pattern in the larger fitted values. This suggests minor nonlinearity or inequality in variances. An impression of a horizontal band where the points lie helps confirm the assumptions made in regression analysis. The data points in the larger fitted values have much larger residual errors compared to the smaller fitted values and can be seen as outliers. The data points under the fitted values of 300 are relatively randomly distributed around the horizontal line which implies that we do not have non-linear relations. Most of the fitted values are really small compared to the largest ones which is a result of the seasonal character of the sales of the product.

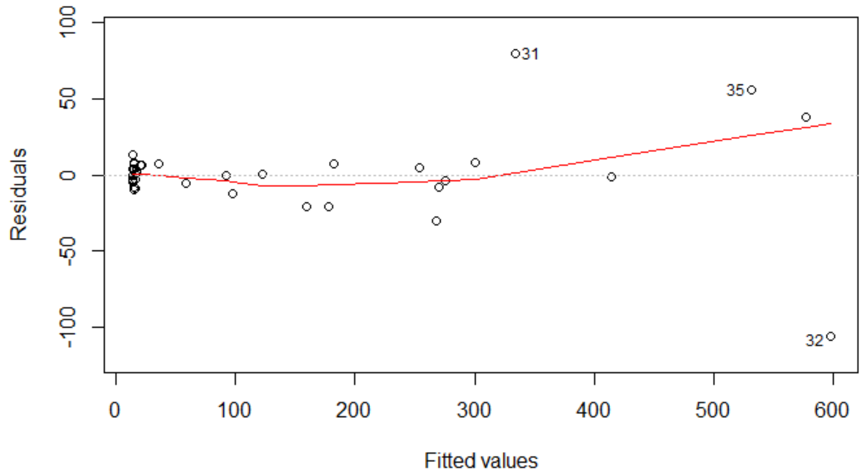

Figure 4: Plot of residuals versus fitted values of the model for product 1.

To inspect the slight curvature or the outward-opening funnel observed in Figure 4 in detail, we plot the residuals against the corresponding values of the regressors. This is done in Figure 5 where we see that most values in every plot seem to be cramped near the origin. This is because the extreme values inflate the axes. Nonetheless, when inspecting the data points more closely, we see that they are randomly distributed on a horizontal band which do not reveal any clear indication of a problem with either the regressors or the inequality of variance.

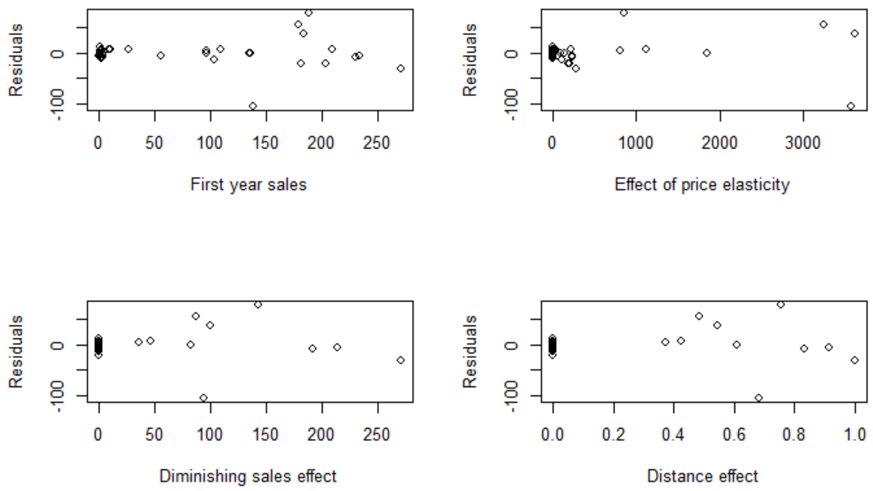

Figure 5: Plots of the residuals against the regressors of the model for product 1.

Thus far, we have found some extreme values in our data which can sometimes have a crucial influence on the model. To find the influential data points we can examine the plot of residuals against leverage in Figure 6. In this plot the pattern of the data points does not matter but instead the two dashed lines representing the Cook's distance scores are important. The Cook's distance measures the influence of a data point on the model if it is deleted from the sample data. That is, the more a data point is behind the Cook's distance lines, the more influence it has to the model. Leaving one or more of these data points out of the model can result to the model being changed radically. In Figure 6 we can see that we have five influential data points of which the most significant ones are numbered. These data points are the weeks where the highest sales quantities were made in Figure 1. This means that the model relies heavily on the high sales after the the first price discount.

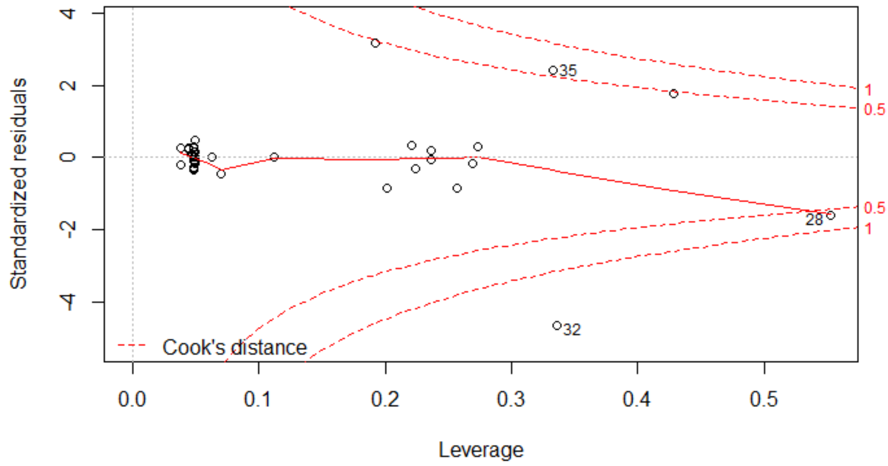

Figure 6: Plot of residuals against leverage of the model for product 1.

The estimates of the coefficients and their p-values of product 1 are in Table 2. When our model is fitted to the data of product 1 all coefficients have a significance level under the cutoff level of 0.05. This means that the coefficients in the model are all meaningful and a change in any of the regressor values will have an impact on the response. The estimate of the distance coefficient $\beta_{distance}$ is significantly larger than other estimates, which originates from the way the regressor is built. The effect of distance regressor is rescaled to have values between 1 and 0 which is a magnitude smaller than for example the values of the last years' sales regressor. The p-values of the coefficients of all products were under the cutoff level of 0.05 and the estimates had values of similar magnitude.

Table 2: Estimates and their p-value of our model fitted to the product 1.

Coefficient	Estimate	p-value
β_0	14.2	0.030
β_{sales}	0.7	$7.6 \cdot 10^{-10}$
$\beta_{distance}$	363.7	$3.3 \cdot 10^{-14}$
β_{price}	0.1	$5.6 \cdot 10^{-5}$
$\beta_{sales,distance}$	-1.2	0.001

As it has been established earlier regressors which get variance inflation factor (VIF) values larger than $\frac{1}{1-R^2}$ can be said to be heavily involved in the multicollinearity of the model. The R^2 of the model fitted on product 1 is 0.974 which gives us a VIF threshold of 38.46. The calculated VIF values of our model can be seen from the Table 3. As we can see from the table, the regressors last years sales and effect of price elasticity have a relatively low value indicating that they have relatively low correlation with other regressors. However, the effect of distance and the effect of diminishing sales have clearly higher VIF values. While they do not exceed the calculated VIF threshold they are still quite high, suggesting some level of multicollinearity in the model and correlation between the two and other regressors. The correlation between the effect of distance and the effect of diminishing sales is obvious when looking at how the two regressors are built. They both use the same fitted quadratic model (8), which immediately adds positive correlation between them. Thus, minor multicollinearity is inevitable in our situation, but because it does not exceed the threshold of serious multicollinearity, we accept it.

Table 3: Regressors and their VIF values.

Regressor	VIF
Last years sales	2.42
Effect of distance	28.18
Effect of price elasticity	2.61
Effect of diminishing sales	24.43

Next, we assess how the fitted model compares with the actual sales data. The coefficients of determination or R^2 are in Table 4, indicating that the fit of the model depends on the product.

Table 4: Coefficient of determinations for each of the products.

Product	R^2
1	0.974
2	0.733
3	0.971
4	0.414

The high R^2 of our model fitted on product 1 is 0.974, which indicates our model is a good fit on the data. This can be confirmed when evaluating it graphically from the Figure 7. The R^2 of the fit on product 4 is 0.414 which indicates a rather poor fit as seen in Figure 8. The start of the selling season of product 1 and 4 are slightly different which explains the differing lengths of data in the figures below.

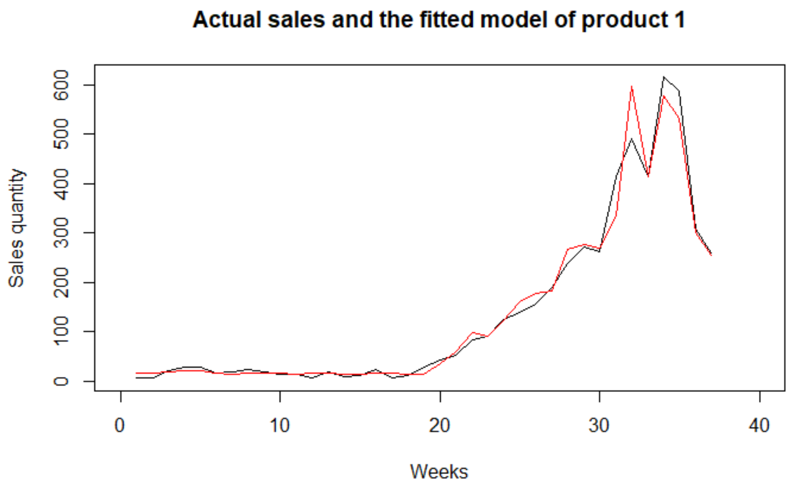

Figure 7: Plot of the actual sales of the training data on product 1 in black and the fitted model in red.

Figure 8: Plot of the actual sales of the training data on product 4 in black and the fitted model in red.

5.2 Forecasting sales and optimizing the selling price

We used our model to forecast the sales quantity of each product for approximately 15 weeks ahead. The forecast for product 1 and 3 can be seen in Figures 9 and 10 respectively. The figures for the forecasts of products 2 and 4 can be found the Appendix A. The prices used for forecasting were the actual selling prices which enables us to measure the performance of the model. The prices used for forecasting the products are in Table 6.

The performance of the forecasts were measured with the metrics in Table 5. The forecasts for product 1 and 2 seem reasonably good but slightly underestimating. The forecast has a MAPE of 37.9% which is relatively high but, in our situation, where we do not replenish our inventory MAPE is not that important and we focused more on the difference in total forecast. The forecast bias of the forecast is 92.7% meaning that total sales quantity is 7.3% less than the actual sales or in other words our model forecasts that there would still be 7.3% of the inventory left at the end of the forecast period. The same conclusions can be made for product 2, which has a forecast bias of 88.1% resulting in a total forecast amount error of −11.9%. The forecast error for product 4 is small, which indicates a good total forecast quantity but as its MAPE is 49% we can presume the forecast error being more of a coincidence than good forecasting capabilities of the model.

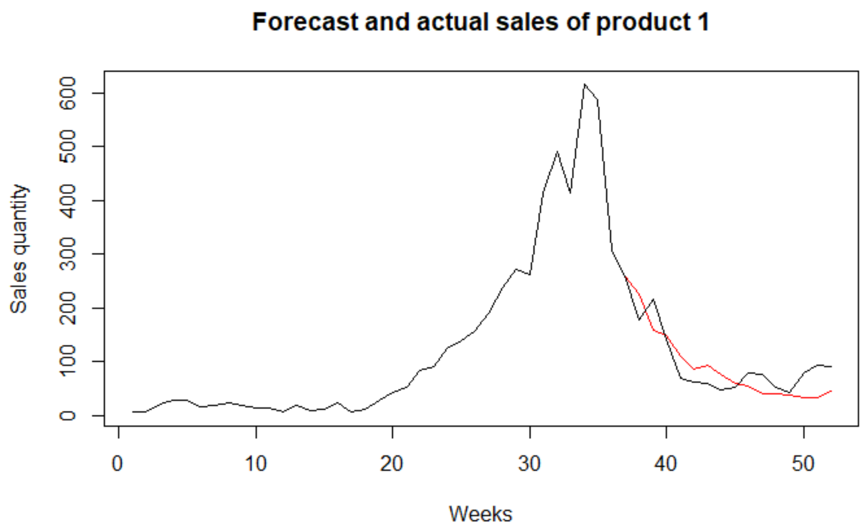

Figure 9: Plot of the actual sales for the whole year for product 1 in black and the forecast in red.

The results for products 1, 2 and 4 were approximately similar in terms of forecast accuracy but for product 3 the model had poor forecast results. This is an interesting result as the model fitted the data quite well and did not have any clear signs of not being suitable for this product.

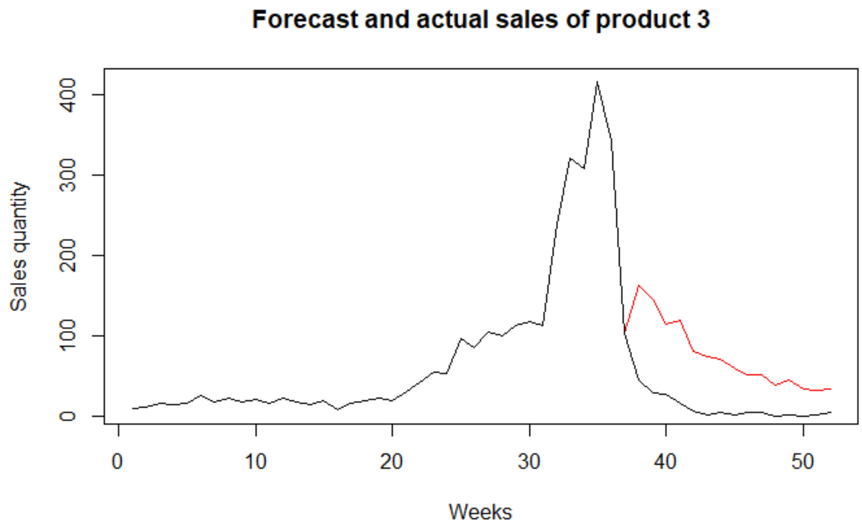

Figure 10: Plot of the actual sales in black and the forecast in red for the product 3.

Table 5: The forecast metrics of forecasts for each product.

	Forecast bias	MAPE	Total forecast error
Product 1	92.7%	37.9%	−7.3%
Product 2	88.1%	34.5%	−11.9%
Product 3	4877.5%	724.6%	656.0%
Product 4	101.0%	49.0%	1.0%

The selling prices used in the data and the optimal selling price in percentages of their retail prices are in Table 6. When optimizing the selling prices, we got results for products 1, 2 and 4 but no optimal price could be calculated for product 3 due to the poor forecasting capabilities of our model for this particular product.

Table 6: The selling prices used and the optimal selling prices in percentage of retail price.

	Used prices	Optimal prices	Impact on profit
Product 1	60% and 45%	50%	-8.5%
Product 2	40%	15%	-63.0%
Product 3	60%	no result	-
Product 4	60%, 55% and 50%	60%	0.7%

When comparing the forecast metric Table 5 with the prices in Table 6 we can see that products 1 and 2 had negative forecast error and their optimal price is below the actual price. We also see that for product 4 the forecast error is positive while the optimal price is slightly higher than the mean of the used prices. This can be explained by the price elasticity of demand component which has a positive effect on sales when the price is lowered and vice versa. We can see a significant drop in price for product 2 even when the forecast bias is not that much lower than the one of product 1 which only received a slight reduction in the optimal price. This suggests that the effect of price elasticity of the fitted model doesn't have as large of an effect in product 2 than it has on product 1.

The optimal prices would have resulted in an 8.5% drop in profit for product 1 and a substantial 63.0% profit drop for product 2 in comparison to the profit generated from the used prices. A slight increase of 0.7% in the profit of product 4 would have happened with the optimized price.

6 Conclusions

In this thesis we used the findings from the literature review of related research papers to first formulate a sales forecasting model for seasonal products using a linear multiple regression model. We then obtained data from a European retailer and fitted and analyzed the model on four different highly seasonal products for outdoor activities. Finally, we calculated the forecasts and optimized the selling price for the clearance period for each of the four products using the bisection method.

There are numerous studies of successful price optimization implementations resulting in positive revenue growth and better capabilities of depleting the inventories by the end of season. All such past studies started by formulating a demand model of some kind and then optimizing the price. According to our literature review the most important factors in forecasting the demand of a seasonal product are the past sales, price discount, the time of season and the inventory level. The price discount is closely related to the price elasticity of demand, which is used in most past studies and is the simple most important factor in optimizing pricing. In fact, it is so important that in some past studies where there was a lack of price sensitive data, two-phase estimation methods are used to calculate the price elasticities. Also, a noteworthy remark from past studies is that the price elasticity of demand is not constant and tends to become less elastic towards the end of the selling season.

We constructed a forecast model which considers last years' sales, the effect of price elasticity and the effect of diminishing sales towards the end of the season. The price elasticity of demand as well as the effect of diminishing sales were approximated to decrease exponentially towards the end of the season. Our regression model did not quite fulfill the conditions of a good model in terms of regression analysis which is partially explained by our data being time series that is usually correlated in some way. However, we still used regression analysis as a method to guide us in the right direction. The model fitted the given training data quite well, while the accuracies of the forecasts were varying.

We used two years' worth of transactions data which we aggregated to weekly product-level information. We constructed the model to work with the data provided, which is not the best approach if the model is intended to work generally for all seasonal products. As this data was used as a time series and the data was split into a training and validating series, we only practically had roughly one and a half 52-week-seasons worth of data and one more season of data could have potentially resulted in a better forecast. One more season worth of data would, however, change the formatting of the model to also take the additional season into consideration.

Using the actual prices, gave reasonably good forecasts for products 1, 2 and 4 but the model could not get an appropriate forecast for product 3. We measured the performance of the forecasts using two commonly used metrics: mean absolute percentage error (MAPE) and forecast bias which gave good results for products 1 and 2 with a MAPE of 37.9% and 34.5% as well as a forecast bias of 92.7% and 88.1% respectively. Product 4 got a MAPE of 49.0% which indicates that its forecast bias of 101.0% is rather a result from good coincidence than the outcome of a good forecast model. Considering this and the fact that product 3 did not get a good forecast at all, the forecasting model constructed would not be suitable for general use without further modifications.

By optimizing the prices with the bisection model, we got an optimal price of 50%, 15% and 60% for products 1, 2 and 4 we respectively as a percentage of the normal retail price. For products 1 and 2 the optimal price was lower than the actual price or prices used. As the forecast model underestimated the sales when forecasting with the actual prices, the lower optimal prices resulted in a negative impact of -8.5% and -62.0% on profit respectively. The optimal price for product 4 resulted in a slight positive impact on profits with a 0.7% increase. However, we optimized the price only so that the model would empty the initial inventory by the end of the clearance period and we did not account for the possible impacts on profit.

7 Future prospects

Our model did not perform as well as expected. Further modifying of the model and testing it with a much larger sample of products would be needed to assess the full potential of the model. This can be challenging, because seasonal products tend not to have that many selling seasons, however, more data could result in better forecasts. For that purpose, reference products with similar sales patterns could be used.

Past research points out that the price elasticity decreases over time [Gupta et al., 2006, Caro and Gallien, 2012, Soysal and Krishnamurthi, 2012, Smith and Agrawal, 2017] but no explicit behaviour is mentioned. We assumed that the price elasticity value decreases in an exponential manner to a certain minimum value. The decreasing of the price elasticity of demand over the season would benefit from more research especially when the price elasticity is considered to be the most important factor in defining optimal price discounts.

We optimized only one price for the whole clearance period which, according to past research [Caro and Gallien, 2012, Soysal and Krishnamurthi, 2012], is suboptimal at least in the terms of generating profit. Optimizing the price so that profit is also considered would require products to have a salvage value as well as possible inventory costs. Also, formulating a dynamic model with a few, say, two or three points at which the price is optimized could improve profits. An interesting approach would be to apply the two-stage estimation system used by Caro and Gallien [2012]. In addition, according to Soysal and Krishnamurthi [2012], optimizing the starting time of the clearance period can further improve the results.

A highly useful addition to the model would be to account for the inventory levels. As Smith [2009], Smith and Achabal [1998], Caro and Gallien [2012] and Vakhutinsky et al. [2012] stated, the inventory level can have either negative or positive impact on sales. A large sales volume compared to the reference shelf level renders the inventory effect nonexistent. Thus, the benefit of considering the inventory effect the products would be highest when the sales volume is low compared to the minimum shelf level.

References

Felipe Caro and Jérémie Gallien. Clearance pricing optimization for a fast-fashion retailer. *Operations Research*, 60(6):1404–1422, 2012.

Rudolf J Freund, William J Wilson, and Ping Sa. *Regression Analysis*. Elsevier, 2nd edition, 2006.

Diwakar Gupta, Arthur V Hill, and Tatiana Bouzdine-Chameeva. A pricing model for clearing end-of-season retail inventory. *European Journal of Operational Research*, 170(2):518–540, 2006.

Robert L McLaughlin. Forecasting models: Sophisticated or naive? *Journal of Forecasting (pre-1986)*, 2(3):274, 1983.

Rudi Meijer and Sandjai Bhulai. Optimal pricing in retail: a Cox regression approach. *International Journal of Retail & Distribution Management*, 41 (4):311–320, 2013.

Douglas C Montgomery, Elizabeth A Peck, and G Geoffrey Vining. *Introduction to Linear Regression Analysis*. John Wiley & Sons, 3rd edition, 2012.

Volker Nocke and Martin Peitz. A theory of clearance sales. *The Economic Journal*, 117(522):964–990, 2007.

B Peter Pashigian. Demand uncertainty and sales: A study of fashion and markdown pricing. *The American Economic Review*, 78(5):936–953, 1988.

Relex Oy. Measuring forecast accuracy: The complete guide, 2017. URL https://www.relexsolutions.com/measuring-forecast-accuracy/.

Anna-Lena Sachs. Safety stock planning under causal demand forecasting. 680(1):13–33, 2015.

Stephen Smith, Shelby McIntyre, and Dale Achabal. A two-stage sales forecasting procedure using discounted least squares. *Journal of Marketing Research*, 31(1):44–56, 1994.

Stephen A Smith. *Retail Supply Chain Management: Quantitative Models and Empirical Studies*. Springer, 2nd edition, 2009.

Stephen A Smith and Dale D Achabal. Clearance pricing and inventory policies for retail chains. *Management Science*, 44(3):285–300, 1998.

Stephen A Smith and Narendra Agrawal. Optimal markdown pricing and inventory allocation for retail chains with inventory dependent demand. *Manufacturing & Service Operations Management*, 19(2):290–304, 2017.

Gonca P. Soysal and Lakshman Krishnamurthi. Demand dynamics in the seasonal goods industry: An empirical analysis. *Marketing Science*, 31(2): 293–316, 2012.

Andrew Vakhutinsky, Alex Kushkuley, and Manish Gupte. Markdown optimization with an inventory-depletion effect. *Journal of Revenue and Pricing Management*, 11(6):632–644, 2012.

Deborah Weinswig, John Mercer, and Filippo Battaini. Us retailer survey: Revealing the hidden costs of poor inventory management, 2019. URL https://coresight.com/research/us-retailer-survey-revealing-the-hidden-costs-of-poor-inventory-management-2/.

Xin Yan and Xiaogang Su. *Linear Regression Analysis: Theory and Computing*. World Scientific, 1st edition, 2009.

Nizar Zaarour, Emanuel Melachrinoudis, and Marius M Solomon. Maximizing revenue of end of life items in retail stores. *European Journal of Operational Research*, 255(1):133–141, 2016.

Wen Zhao and Yu-Sheng Zheng. Optimal dynamic pricing for perishable assets with nonhomogeneous demand. *Management science*, 46(3):375–388, 2000.

A Attachment

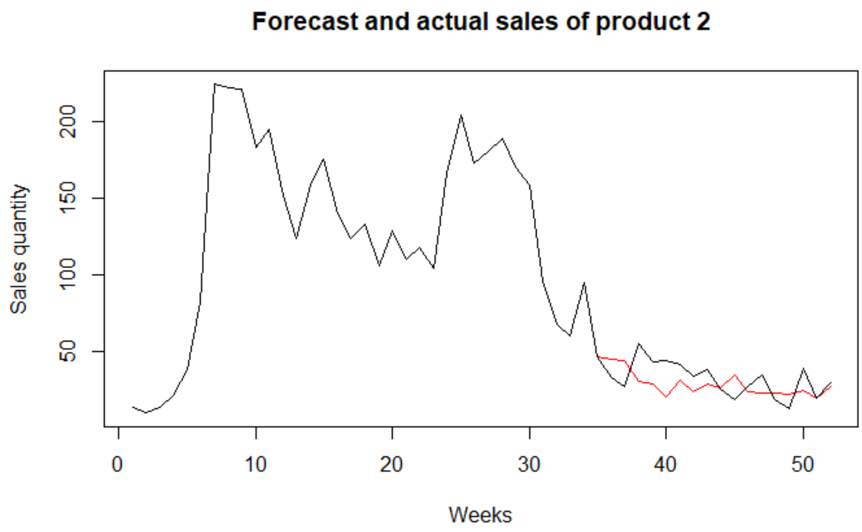

Figure 11: Plot of the actual sales and the forecast for product 2.

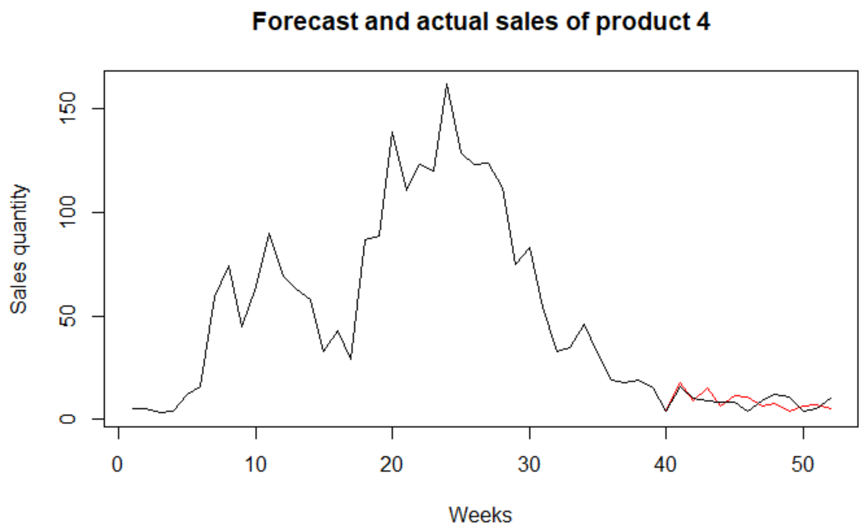

Figure 12: Plot of the actual sales and the forecast for product 4.

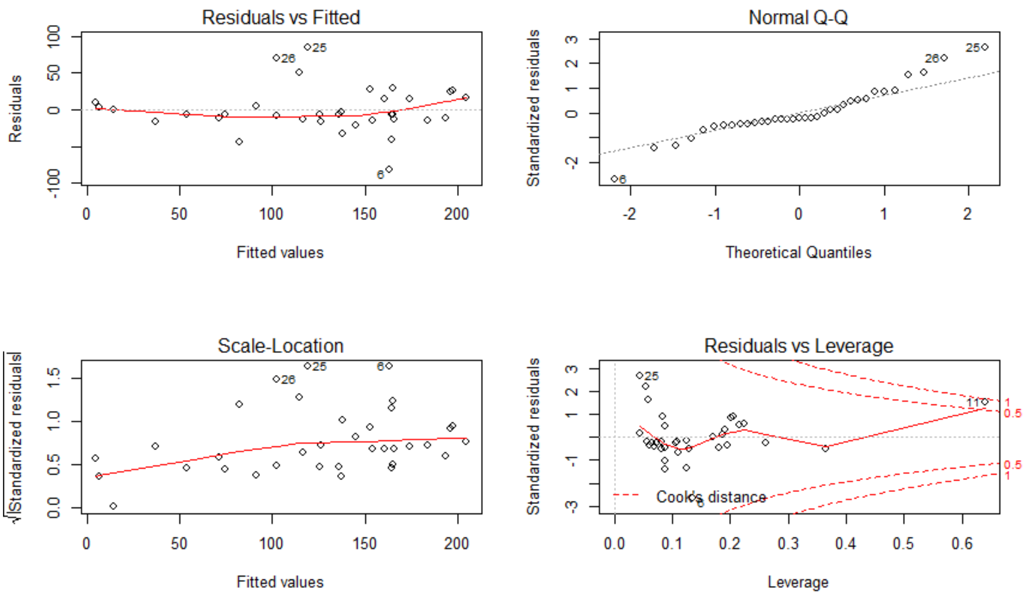

Figure 13: Plots used in regression analysis on the fit on product 2.

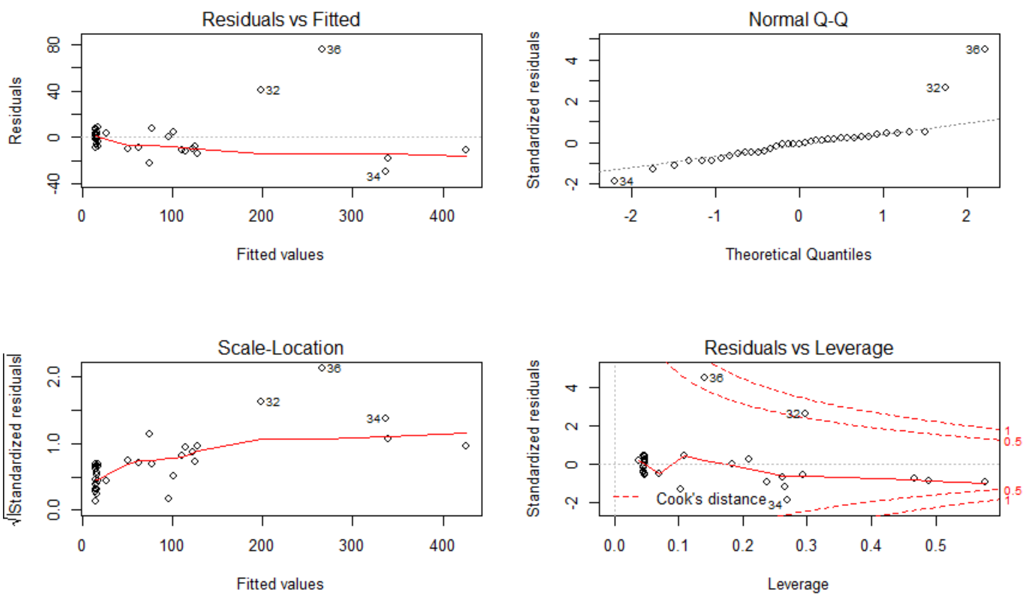

Figure 14: Plots used in regression analysis on the fit on product 2.

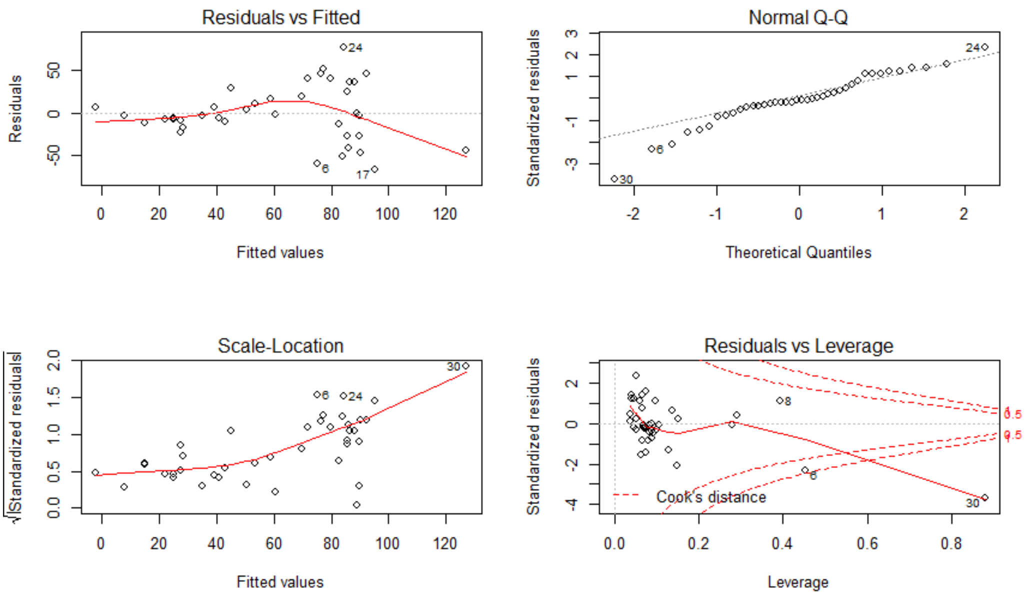

Figure 15: Plots in regression analysis on the fit on product 2.