jamk.fi

Demand Forecasting of an Industrial Item

A Methodology Study

Zubova Nadezhda

Bachelor's thesis
December 2019
School of Technology, Communication and Transport
Degree Program in International Logistics

Jyväskylän ammattikorkeakoulu
JAMK University of Applied Sciences

jamk.fi

Description

Author(s)	Type of publication	Date
Zubova, Nadezhda	Bachelor's thesis	December 2019
		Language of publication: English
	Number of pages: 81	Permission for web publication: x

Title of publication

Demand Forecasting of an Industrial Item

A Methodology Study

Degree programme

International Logistics

Supervisor(s)

Sipila, Juha

Lähdevaara, Hannu

Assigned by

JAMK University of Applied Sciences

Abstract

The demand forecasting process plays an essential role in any business organization. It leads to smaller inventory costs, better planning of resources and, in general, minimized costs and maximized customer service level. However, surveys among companies show that a very small percentage of them can actually implement the forecasting process successfully. There are several reasons behind this, such as lack of knowledge and tools, skepticism about the quality of the process and other reasons.

The objective of the study was to provide information about the importance of demand forecasting for companies. Moreover, factors, such as demand forecasting errors and those related to achieving a good quality implementation of demand forecasting were studied. Finally, the aim was to create a demand forecasting model for a case company and choose the most suitable method for the process.

During the research process, the author utilized the mixed research approach that included components of quantitative and qualitative research methods. The theoretical part of the study was based on a literature review, which helped to answer some of the research questions. Based on the collected information, the author was able to conduct a case study as well as test and analyze four different methods of demand forecasting.

The results of the case study helped to choose the most suitable model for the case company. Although the proposed model was not originally tested and analyzed, it was possible to make such conclusion based on the results related to the four chosen models for the case study. However, the case study results cannot be generalized because they were highly customized for and adjusted to a specific company and its SKU situation in a certain environment. In addition, proposals for further research and recommendations for the company were given.

Keywords/tags (subjects)

Forecasting, Demand, Statistics, Model

Miscellaneous

Contents

1 Introduction

Nowadays, companies meet many problems and challenges on the market in order to be successful and stay competitive. This is related not only to small and medium-sized organizations, but also to the big market leaders. Companies need to have a good level of customer service, high product availability, precise production planning and excellent inventory management. All of those need to be created and maintained in order to achieve higher profits.

There is one interesting phenomenon which influences all those matters, and it is called demand forecasting (TradeGecko 2018). It is easy to understand what demand forecasting means, but very hard to do with good quality and high accuracy. This is why even when most companies believe that forecasting is important only 13% of organizations implement it effectively. (Silverman 2017).

Moreover, according to Chapman & Arnold (2017, 206) forecasts are almost never right. This is something that create skepticism among some companies and over 30% of organizations do not think forecast is important or should be implemented properly (Silverman 2017). These behaviors are wrong, and even though it is impossible to have a forecast without an error, choosing the right forecast method and implementing the process correctly will have a positive impact on the company's profit and customer satisfaction.

Apart of skepticism, barriers which prevent companies to do an effective forecast are: lack of data, lack of tools and knowledge (ibid.). It is easy to get lost among all available demand forecasting techniques and methods, and the task of not only implementing demand forecast but also choosing a right method becomes almost an impossible. That is why it is important to understand the pros and cons of each method, which groups of demand forecast exist and for which type of data different methods are suitable.

The aim of this study was to provide information about why demand forecasting is crucial for companies and explain the related difficulties as precisely as possible. Furthermore, the purpose was to present sufficient amounts of information about forecasting groups and methods with their advantages and disadvantages. Finally, the

case study in this thesis showcased the situation of how poor done forecast can damage a company and how a proper amount of invested time and efforts in choosing the right forecast method can benefit a company.

2 Research methods

Research is the way to find answers to the questions that appear during the study of a certain phenomenon. It is also about solving problems which are connected to the study of a certain subject. (Kananen 2011, 18.) In order to conduct research, some special techniques, such as research methods are used. They are utilized to collect, sort and analyze data and information which leads to certain conclusions and results. (Walliman 2011, 22.) It is very important to decide the most suitable research method for a particular study before any research project is executed. Choosing the right research method is a very important step in a thesis-process because it can influence the whole process and the validity of the results. (Saukkonen 2017, 7.)

Every research project has requirements that it has to fulfill. The most important requirements that research should strive for are as follow:

- Being objective
- Being valid
- Being reliable
- Being built on previous research (cumulative).

It is a standard that research gives information that is objective and true. Such concept as objectivity can sound quite debatable, and there is debate whether research can be objective and how possible it is to avoid bias. (Dahlberg & McCaig 2010, 14.)

Validity and reliability are quite important requirements when it comes to research, because they provide the possibility to finalize a research project and reach a conclusion. Finally, it is essential that a research project is built on previous researches and that is why researcher should conduct literature reviews. This action will make sure that the research project that an author is carrying out is builds on rather than duplicates previous studies. (Dahlberg & McCaig 2010, 14 – 15.)

In order to meet the requirements above, there are three main approaches that can be utilized in a research project: the quantitative, qualitative and mixed approaches. Undoubtedly, the three approaches are not direct opposites, but rather they represent different ends of a continuum. Usually, researcher choose either a quantitative or qualitative method, depending on the phenomenon of the study. However, quite often a subject cannot be studied and analyzed only from one perspective and in that situation, it is wisely to use the mixed approach. (Creswell 2014, 32.)

2.1 Qualitative research method

The qualitative research approach means using words in order to interpret data which was collected and analyzed (Braun & Clarke 2013, 2). In general, this method provides answers to the *why* questions and give possibility to explain specific phenomena in general or detailed way (Biggam 2008, 86). The most common ways of collecting data include conducting interviews, participant observations procedures, implementation of surveys, focus groups and reading and analyzing different sources of literature. All these processes of collecting information generate data in the forms of notes, audio and video recordings as well as transcripts. (Mack & Woodsong 2005, 2.)

A qualitative research method is more flexible than quantitative, and it allows more interactions between the researcher and study participants. It brings emotional involvement to the researcher and makes it difficult to avoid bias and subjectivity. Moreover, the framework of this method is semi-structured and there are no formulas, which means that the researchers need to decide the structure and algorithm by themselves. Because of this characteristic of the method, it takes more time to complete the research process. (ibid., 3 – 4.)

Advantages and disadvantages of qualitative research

The qualitative research approach has its own benefits and drawbacks. One of the main advantages of this approach is that findings from qualitative data can often be transferred to another setting or to a similar population, which provides an opportunity to gain a rich and complex understanding of a specific issue or subject. This quality allows to cover the subject of the study in depth and in detail. Secondly, one

of the main data collection methods is interviews with open ended questions, which allows high interactions between the researcher and study participants and leads to situations where the participants' responses influence on the choice and manner of the researcher's following questions. This situation gives a possibility to the researcher to redirect the questions, guide the participants and revise the framework of the interviews in real time. Moreover, qualitative research is able to provide the "human" side of the topic, because it is based on human experience and what people feel about a given research subject. According to this, qualitative data can be more influential and compelling in comparison to the quantitative one. (Advantages and Disadvantages of Qualitative Research 2014; Mack & Woodsong 2005, 1 – 4.)

However, in spite of all the advantages of the qualitative approach, it also has some weaknesses. The biggest drawback of this method is that the quality of the research is highly dependable on the skills of the researcher. Furthermore, as it was mentioned above, the researcher is highly involved in the process of data collection, which can influence the responses of the study participants. This situation makes qualitative research very bias-based and impossible to be as objective as quantitative research, and therefore, it can produce wrong results and conclusions. Because of these features, this research approach is sometimes not accepted in scientific communities. Furthermore, as all the collected data is text and as it does not have a structure, the interpretation, analysis and visualization of the data can be time- and labor-consuming. (Advantages and Disadvantages of Qualitative Research 2014.)

2.2 Quantitative research method

The quantitative research approach is differing from the qualitative and it deals with numbers and relationships between them, and it is typically based on variables. Hence, when a researcher uses this method, the interpretation of the data is based on numbers, and the most common way to collect it is by conducting surveys. Not only the data is presented in numbers, but the process also follows statistical rules, and all the stages of using quantitative methods are regulated by precise and exact rules. These characteristics of this approach can ensure that the results are accurate and reliable. Moreover, mistakes are quite expensive in this approach, and if the researcher makes one, it is going to double the work amount. When using qualitative

research methods, it is possible to go back to the beginning of the research process and correct some mistakes. However, if a quantitative survey is carried out in a wrong way, everything should be started again. (Kananen 2011, 73 – 75.)

The study group is often quite large, which gives an opportunity to find general patterns and making the results applicable to a bigger population. Apart from qualitative methodology, a quantitative one does not have emotional engagement of the researcher, which helps to avoid bias and brings objectivity. These qualities are another reason why the results and conclusions from the quantitative approach are reliable and trustworthy. Due to the fact that the data is structured and that the researchers need to work with statistics and formulas quite often, data analysis is done with the help of software or a computer, and all the stages of this research method follow precise algorithms and rules, which leads to the conclusion that the whole process of the quantitative approach can be completed faster than that of the qualitative approach. (Saukkonen 2017, 10.)

Since quantitative approach deals with numbers and statistical analysis two research methods from this approach will be useful for this thesis and specifically for the case study part:

- statistical significance
- confidence intervals

2.2.1 Statistical significance

Statistical significance allows to see whether the results are due to a random chance or whether something influenced this outcome (Taff 2019; Crawley 2014). SPSS Tutorials (n.d.) defines statistical significance as “probability of finding a given deviation from the null hypothesis - or a more extreme one - in a sample”. Hence, it is important to understand two parts of these definitions: what is “unlikely” and what a “null hypothesis” means (Crawley 2014).

Another way to call statistical significance is probability value. In mathematical formulas and papers, it is written as the p-value or just p. It is common to say that if an event happens less frequently than 5% of the time, then it considered unlikely. This

means that $p < 0,05$ when an event is unlikely under a certain null hypothesis. (Crawley 2014; SPSS Tutorials, n.d.)

The second element of statistical significance or probability is a null hypothesis. It is a statement which says that something is happening or not happening. (Crawley 2014.) According to SPSS Tutorials (n.d.), "a null hypothesis is a precise statement about a population that we try to reject with sample data". The idea and purpose of a null hypothesis is that it is not necessarily true, but it is falsifiable and gives an opportunity to run a statistical significance test. (Crawley 2014; SPSS Tutorials, n.d.)

Examples of null hypotheses:

- The average grades of male students in JAMK University are similar to those of female students
- The correlation between stress and depression is zero
- The means of actual demand and forecasting are the same

(SPSS Tutorials, n.d.)

2.2.2 Confidence intervals

Confidence interval is a range of values which is made by using sample statistics in order to calculate unknown population parameters with an established level of confidence (Introduction to Confidence Intervals, n.d.; SPSS Tutorials 2019). As a population mean can never be known exactly, confidence intervals with a stated percentage of confidence can show how far away the population mean can probably be. The most commonly utilized confidence interval is 95%. However, it is possible to use 99% or any other value. (Stewart 2016, 37.)

Since different confidence levels can be used in order to create confidence intervals, their widths can also vary. An interval will be wider if the confidence level is higher. (Crawley 2014.)

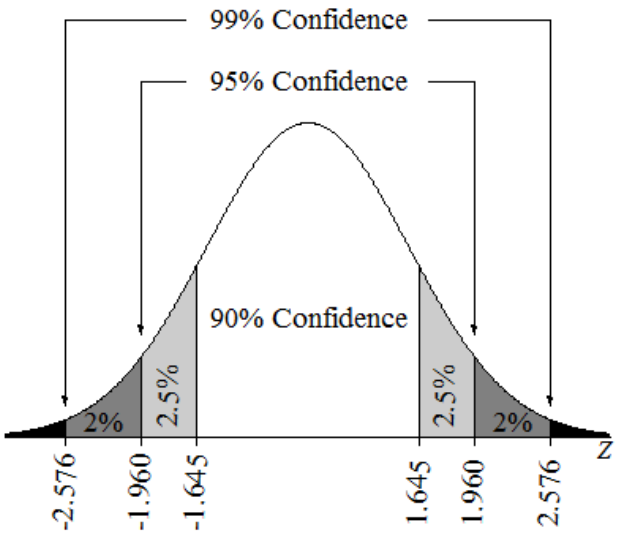

Figure 1. Confidence Intervals (Williams 2018).

Figure 1 shows a graph with confidence intervals. Any confidence interval has a center which is also called the point estimate. Usually, it is a sample mean value which is in the middle of the interval. The margin of error defines the width of the interval. (Introduction to Confidence Intervals, n.d.)

$$Margin\ error = level\ of\ confidence\ \times standard\ error$$

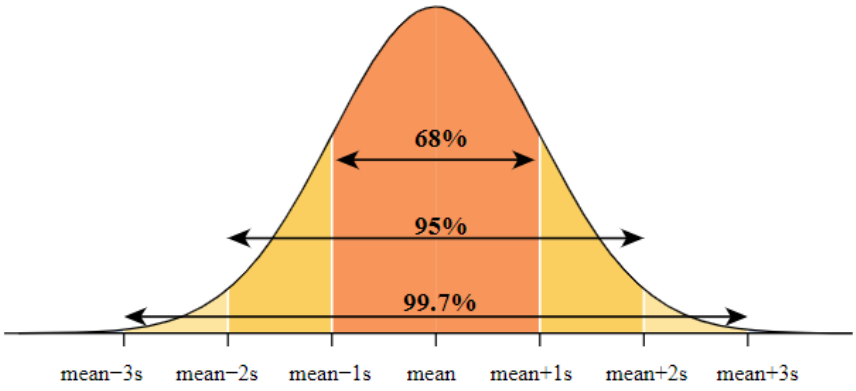

Figure 2. Confidence Intervals (Introduction to Confidence Intervals, n.d.).

By using Figure 2 it is possible to create a general form for the confidence level:

$$Sample\ mean\ \pm\ margin\ of\ error$$

Alternatively, an example form for a 95% confidence interval:

$$Sample\ mean\ \pm 2(standard\ errors)$$

(ibid.)

Advantages and disadvantages of the quantitative approach

As the qualitative approach, quantitative one also has its own benefits and drawbacks. The most important strength of this method is that the generated results are reliable and valid. It is because the process of collecting and analyzing the data is highly structured and controlled, which leads to the opportunity to avoid bias and make the results as objective as possible. Moreover, this method is the best when it comes to dealing with big amounts of data, and it gives the possibility to generalize results to a larger population. (Dowd, n.d.) Finally, the data which is collected and analyzed is quantifiable, which means that computers with software can be used for these processes. This possibility makes the quantitative approach less time- and labor-consuming in comparison to the qualitative approach. (Saukkonen 2017, 11.)

Despite the strengths that quantitative method has it also holds some weaknesses. The main one is that computer software and survey tools are quite vulnerable to errors, especially when it comes to the mistakes in measurements. Such vulnerability can easily affect the results and makes them wrong and irrelevant. (Dowd, n.d.) Secondly, even though the qualitative method means working with a great deal of data, it gives quite common and general results, which does not provide deep and detailed understanding if we compare this to the qualitative approach. Finally, the objectivity of the results is undoubtedly a very good characteristic with this methodology, but sometimes it is important to have a human element and the engagement of the researcher, because not everything can be under strictly controlled conditions. (“DME for peace: Section 1.3 Quantitative Research Module”.)

2.3 Mixed research method

The mixed research method is an approach where quantitative and qualitative methods are combined. This method is recognized as valuable, and it has become quite popular in the last few years because it can highlight the strength of the both traditional methods as well as compensate for their weaknesses. (Ostlund & Kidd 2010, 1.) Moreover, sometimes it is quite complicated to study the issue by only using one method, especially when the research subject is not tied to one discipline. There is

also an assumption about this method that it gives more complete and detailed understanding of a research problem rather than qualitative and quantitative methods separately. (Creswell 2014, 32.) The only drawback of this method is that there is a lack of information in the literature about how to structure a mixed research method and combine qualitative and quantitative findings. (Ostlund & Kidd 2010, 2.)

Table 1. The Differences between Qualitative and Quantitative Research Methods. Adapted from: Dahlberg & McCaig 2010, 22; Mack & Woodsong 2005, 3; Olusegun 2002, 5.

	Qualitative	Quantitative
Data format	Textual/soft data	Numerical/hard data
Cases	Few cases	Many cases
Theory of knowledge	Subjective	Objective
Theory development	Inductive approach	Deductive approach
Framework	Semi structured	Highly structured
Flexibility	Flexible	Stable design, inflexible
Research context	Uncontrolled	Controlled
Question format	Open-ended	Close-ended
Examples of methods	Interviews, observations, focus groups	Questionnaires, observations

2.4 Methods of collecting data

2.4.1 Interview

Interview is a valuable method of data collection. It is a conversation between the researcher and the person who has knowledge or experience needed for the study. Interviews give the opportunity to explore the meaning and negotiate the subject of the study in a natural setting. An in-person or telephone interview is selected as a method when the researcher can assume that opinions and knowledge of another person are meaningful. This can help to create a holistic snapshot and influence the results and success of the research. (Westat 2002, 50; Alshenqeeti 2014, 39.)

There are two types of interviews, which can be utilized by researchers: structured and in-depth. The first type of interview is conducted by following precisely a set of questions, and it has a rigid structure. Interviewers cannot paraphrase or change the questions, and this structure allows to have consistency and sameness in the interview process. An in-depth interview, on the other hand, does not have an inflexible structure, and it gives a chance of open and free responses that help to understand a participant's perspective better. (Westat 2002, 50 – 51.)

Interviews are one of the most enjoyable method of data collection. However, they have advantages and disadvantages, which are presented in the table below.

Table 2. The Advantages and Disadvantages of Interviews as a Data Collection Method.
Adapted from: Westat 2002, 52; Alshenqeeti 2014, 43.

Advantages	**Disadvantages**
Flexibility	Time - consuming
Allows to get new insights, details	Subjective and has bias
Include the reality	High amount of information for analysis
High return ratio	Possible incompatibility and unsteadiness

2.4.2 Observation

Observation is a data collection technique where the researcher acquires firsthand information about the behaviors or processes being studied. This method gives an opportunity to directly observe different activities and operations, and that is how the researcher can obtain a holistic perspective. The best stages of the research project where this method can be used are formative and summative. (Westat 2002, 53, 55.) Observation as one of the methods of data collection has its own advantages and disadvantages which can be seen in the table below.

Table 3. The Advantages and Disadvantages of Observations as Data Collection Method.
Adapted from: Westat 2002, 55; Kananen 2011, 48.

Advantages	Disadvantages
Precise information	Time - consuming
Authentic environment and situation	Requires well qualified people (observers)
Researcher is involved	Can influence the behavior of participants and correctness of the results
Holistic perspective	Possibility of atypical behaviors

2.4.3 Literature review and document studies

Literature reviews and document studies are among the most popular and widely used methods of collecting data for research. Literature reviews or document studies mean a process where the researcher finds already existing information which has been published concerning a specific topic of the research work. Information for a literature review can be easily found nowadays, because of the existence of free information sources with big collections of data. Document studies can be done by reading newspapers, books, different types of reports, scientific papers, articles and many other sources, which can be found in hard copies, but also online on the internet. (Westat 2002, 57 – 58; Robinson & Reed 1998, 58.)

There are many reasons why literature reviews can be conducted:

- To understand what has already been found and discussed about a specific topic
- To allow researchers to find deeper theoretical knowledge
- To see whether there is need for the research
- To prevent duplication of studies
- To identify lack of knowledge and information

- To allow comparisons in order to find connections or contradictions between different research works
- To identify the weaknesses of previous studies

(Aitchison 1998, 58; Bless & Higson-Smith 2000, 20.)

Table 4. The Advantages and Disadvantages of Literature Review as Data Collection Method.
Adapted from: Westat 2002, 59.

Advantages	**Disadvantages**
Cheap	Inaccuracy
Easy access	Require a lot of time to read and analyze
Give possibility to see changes over time	Have chances of incomplete information
Variety of types and forms of information from any time period	

2.4.4 Surveys

A survey basically means using questions in order to collect needed information and data. Surveys provide an opportunity to create questions and receive answers without interaction with every participant or responder. A survey can be both a qualitative and quantitative research method. However, it is mostly used in order to collect quantitative data. It depends on the analytical method which is chosen as well as whether the questions allow open answers or not. Thus, it is in the hands of the researcher how to use a survey, which kind of data needs to be collected and how to shape the questions around this. (Walliman 2011, 73, 97.)

As it was mentioned above there are two types of questions:

- Closed-ended questions

  This type allows the participants to select from a provided set of answers. The advantages of closed-ended questions are that they are quick to answer and that no writing is involved. The researcher benefit is that they are easy to make or code. However, there is a down side of this type, which is the restrictions in obtaining a full range of the potential answers.

- Open-ended questions

  This type provides the participants with the opportunity to answer what they want and feel. This naturally brings such benefits as freedom of content and eliminating the bias of the researcher. On the other hand, open-ended questions are more time consuming to answer and more difficult to create and analyze.

(ibid., 97 – 98.)

In addition to choosing the type of questions for a questionnaire, the researcher can also decide on the method of distribution between three available ones:

- Personal
- By post
- On the Internet

However, irrespective of which method is used, all questionnaires should have an introduction where important and required information is stated, such as who the researcher is, what data will be used for, and how the privacy of this data is protected. Moreover, all respondents should have a possibility to cancel their participation at any time. (ibid., 48, 97.)

The biggest problem of surveys is the response rate. It is very difficult to reach a sufficient number of participants who have actually completed the whole questionnaire. However, it is very important to have a sufficient number of responses because otherwise it is difficult to interpret the results. Moreover, the response rate influences on the reliability and validity of the collected data. (Kananen 2011, 93 – 94.)

Table 5. The Advantages and Disadvantages of Surveys as Data Collection Method. Adapted from: Walliman 2011, 97; Kananen 2011, 93.

Advantages	Disadvantages
Flexible	Time consuming to create
Structured format	Require certain skills to design and develop
Cheap and quick to distribute	Possibility of only short and easy questions
Non-bias	Not all are cable to finish questionnaire
Can cover big geographical areas	Possibility of low response rate

2.4.5 Case study

A case study is an empirical study with an in-depth focus on a certain phenomenon in a natural settings or real-life context. A case study gives the possibility for the researcher to have a holistic view of the important qualities of a real-life event. (Denscombe 2003, 30 – 31; Yin 2009, 4, 18.)

A case study can contain quantitative and qualitative data, and it should be considered a research method or strategy rather than a data collection technique alone. This method includes creating a research design, collecting data and analyzing it. That is why a case study can use multiple sources of data and multiple methods of how to collect it. (Denscombe 2003, 31 -32; Yin 2009, 18 – 19.)

This method of research can be used for a variety reasons, such as conducting an in-depth study of an individual case or willingness to generalize certain findings. Moreover, it can have different purposes, which leads to three main types of case studies:

- Exploratory case study
- Descriptive case study
- Explanatory case study

(Yin 2009, 8, 20.)

Table 6. The Advantages and Disadvantages of a Case Study.
Adapted from: Westat 2002, 61; Denscombe 2003, 38 – 39.

Advantages	**Disadvantages**
Gives a rich picture	Time consuming
Possibility to go into details	Difficult to generalize
Holistic analysis	Requires good skills in the data collection process and reporting
	Needs great deal of resources

2.5 Research data

2.5.1 Primary data

Primary data is information that is collected by a researcher for a certain study problem at first hand (Hox & Boeije 2005, 593; Rabianski 2003, 43). Every time when new primary data is collected, it is going to be added to the existing bank of knowledge. The collection process of primary data has its own advantages and disadvantages. The most important advantage is that the collection of the information is tightly connected to the research problem and helps to solve it. Thus, the purpose and focus of those procedures are finding a solution for the study problem, which makes the research design and collection strategy tailored to the research question. The biggest disadvantage of primary data collection is that it is time consuming and expensive. That is why it brings benefits to reuse already collected data if relevant information about the research question already exists. (Hox & Boeije 2005, 593 – 594.)

The table below presents the different methods of collecting primary data.

Table 7. Methods of Collecting Primary Data from Hox & Boeije (2005, 596).
Adapted from: Hox & Boeije 2005, 596.

Qualitative	Quantitative
Focus group	Different types of surveys
Open interview	Experiment
Observations	Observations

2.5.2 Secondary data

Secondary data is information which is not collected by the researcher but is taken and combined from secondary sources, such as data archives, data storages of universities and other similar sources (Rabianski 2003, 43). When the researcher decides to use secondary data, it means using data which has been collected earlier by other people and for other research purposes. This means that any kind of primary data can become secondary data and be reused. This reusing of information is possible, because most of the secondary data is quantitative and represented by numbers and values, for example, in form of statistics. (Hox & Boeije 2005, 593,594, 596.)

Secondary data has its own advantages and disadvantages. The main problem is that the information has not originally been collected for this specific research but for another one. This means that the collected data is not only coherent with the research problem, but that it is also difficult to explain it without the whole picture of the context and without knowing how the data was collected. On the other hand, the main advantage of secondary data is that it provides plenty of useful information at a lower cost and more quickly when compared to primary data. (ibid., 594.)

3 Research questions and focus

This thesis work consists of two parts: a theoretical literature review and a case study with a real-life data from the company. Both of these parts are necessary in order to fully answer the research questions and form a deeper and bigger picture about such a phenomenon as demand forecast.

The research questions of this study were as follows:

1. *Why do companies need demand forecasting?*
2. *Which methods of demand forecasting exist and how can companies choose the right one?*
3. *What causes the forecast error?*

The hypotheses of this study were as follows:

1. *Correlation does exist between demand and arrivals in the case study data.*

2. *It is possible to create an accurate forecast by using chosen demand forecasting methods based on the case study data.*

The focus of this research work was on demand forecasting and why it would be so important for companies operating in the production sector. This thesis highlights the most important facts about forecasting and how this difficult process plays a big role in the success of companies in the market.

The process of creating demand forecasting is very complex and difficult, and it has its own limitations and disadvantages, which are described and presented in this work. However, it is a crucial for companies to run this process. However, just having this process inside the company is not enough and will not bring any benefits if the chosen method of demand forecasting does not meet the company's plans and data. It is important to know which method to choose from the variety of the existing ones and how to understand which fits the best. In order to do this analysis, a literature review presents most of the demand forecasting methods with their benefits and disadvantages. Moreover, a case study with real-life data was one of the objectives of this thesis. There three different forecasting methods were tested and compared in order to choose the best one.

The final aim of this research work was to add understanding of forecasts and forecasting itself. Furthermore, the aim is to understand why forecasts are often wrong and what are the sources of errors. By understanding those better, companies can also act in order to reduce them and avoid unnecessary inaccuracy.

3.1 Limitations of the research

Since demand forecasting is a very broad phenomenon, this research was limited to production sector companies of the supply chain. Moreover, demand forecasting includes such aspects as safety stock, lot sizing and reorder point. These parts are only briefly explained or not covered at all in this research work as they are not in the main focus of this study.

Furthermore, there are various methods and techniques for demand forecasting. After a long and detailed study of different types of demand forecasting, it was decided

to concentrate on the quantitative ones. This decision was made due to the fact that qualitative ones are subjective and based on the knowledge and judgement of the workers in the company. Since the case study was conducted with pre-collected real-life data, it was not feasible to involve the employees of the company from which this data was collected. Hence, it was logical to concentrate on the quantitative methods of demand forecasting. Moreover, despite the fact that qualitative methods are quite commonly used in companies, their boundaries and limitations clearly over-balance their benefits.

Even though this research work concentrated only on one group of methods of demand forecasting, it was not possible to cover all quantitative methods. Thus, the author decided to focus on a specific category of quantitative methods: time series, more precisely on two methods from this category, namely, exponential smoothing and ARIMA. These two methods were selected due to their practical value for companies and ability to be easily implemented. A close study was conducted in order to understand these two methods and compare them by using the case study data.

Furthermore, this research work utilizes two methods, and three techniques there, as several exponential smoothing approaches exist. After consultations with the thesis supervisors, the author decided to concentrate on a Single and Holt Winter's two-parameter types of exponential smoothing.

Finally, the case study was conducted by utilizing pre-collected data on a specific component of an investment product from an industrial company with a demand pattern that was very unstable and difficult to forecast. The results and calculations may be irrelevant when applied to other processes of creating demand forecasting for different components even with similar demand patterns. This case study process was highly customized and adjusted for a specific company and SKU in a certain environment. Later usage of the results of this work on similar cases needs complementary research and extra analysis.

3.2 Overview of the research structure

Based on the research focus and questions, the mixed research approach was chosen for this study. As it was necessary to collect and analyze data which is presented in

two forms, numerical and categorical, both quantitative and qualitative research methods were utilized. In addition to using two different research methods, the author decided to use two methods of collecting and analyzing data, which were a literature review and a case study.

For answering the first research question both quantitative and qualitative methods were utilized. As the question required understanding of what demand forecasting is and which role it plays in the company, theory was read and studied. Thus, a literature review as a data collection method was utilized as a part of the qualitative research method. However, not only the theoretical background helped to answer this question, but also the case study. Quantitative data were utilized in order to create forecasting with the help of the SPSS program, and in the Conclusion part, the important findings of demand forecasting for the company are presented.

The second research question also required a mixed approach in order to not only list the existing methods of demand forecasting, but also to show how some of them worked in practice. In addition, with the last task the case study with real-life numerical data also helped. Moreover, the case study with graphs and tables from the SPSS program allowed to make a conclusion and justify it concerning the choice of the right forecast method.

The final question was answered by utilizing only a qualitative research method which was presented by the literature review. The theory with a certain amount of formulas, numerical examples and graphs provided the opportunity to explain the nature of a forecast error. Since a literature review and case study with pre-collected data were utilized in this research work, secondary data was used throughout this thesis.

4 Literature review

4.1 Forecasting

Forecasting is a preface to planning. It is important to make assumptions and calculations about the circumstances that will exist in the future before building any strategy and making plans. Almost nothing can be done without any sort of estimation!

(Chapman & Arnold 2017, 202.) That is why forecasting is above all value and price, because not being able to harmonize the available resources with plans and possible future circumstances can ruin any company and business (Kolassa & Siemsen 2014, 6).

Principles of forecasting

In order to use forecasts effectively, it is important to understand their four main characteristics:

1. Forecasts are usually wrong. Forecasts are only assumptions and attempts to predict the unknown future. Errors are unavoidable and must be counted upon.
2. Including a calculation of error in-to forecast is a must. It is not a question if a forecast is wrong or not, because it is mainly guessing and because it is difficult to avoid errors. However, the main question is "by how much?". It is important to know how imprecise the forecast is and that is why it should include a calculation of an error. Most often, the error is presented as a range between maximum and minimum values or as a percentage.
3. Forecasts are more precise for groups or families of the items. In a group with fixed characteristics, the behavior of a particular item is still random. This means that the larger the group of items, the more precise the forecast will be. When a plan is made in industrial production, families of items can be found based on the similarity of the machinery and equipment used and on the process of manufacturing.
4. Forecasts are more correct for the near future. Any future naturally contains some amount of uncertainty, but a period close in time has less uncertainty than the far future. For a firm, it is easier to make short-term forecasts than long-term ones. This is very important for items with long lead times and particularly if their demand is dynamic. The smaller the lead time, the more accurate the forecast will be. (Chapman & Arnold 2017, 206 – 207.)

Forecasting techniques

Nowadays, plenty of forecasting techniques exist, and all of them can be arranged into categories. Forecasting methods in general can be quantitative or qualitative,

and they can have two different factors that they are based on: intrinsic (internal) and extrinsic (external).

- Quantitative techniques

  Quantitative techniques mean using historical or numerical data from inside and outside the company. The results of such forecast are objective, and they can be reached faster than when using qualitative methods. This method is good for inventory and production forecast, because of the need of demand forecast for particular end items.

- Qualitative techniques

  Qualitative techniques are based on opinions, hopes, knowledge and intuition, and that is why the results are quite subjective. Usually, these methods are used by senior managers, because they can forecast possible demand for large groups of items over long periods of time and predict general business tendencies and dynamics. There are many different qualitative methods which are presented later in the demand forecast chapter.

- Intrinsic techniques

  These forecasting methods use historical data which is recorded and stored in the company. The main idea of these techniques is that what happened in the past will be repeated in the future.

- Extrinsic techniques

  This forecasting method means using external indicators that influence and affect the demand of an organization's products. The main idea behind these techniques is that the demand of a specific family of products correspond to the activity in another sphere or field. For example, if the government establishes a new supporting campaign for building houses and gives construction contracts to different companies, it is easy to predict that the demand of housing materials, such bricks, wood, windows etc. is going to increase. Data of these kind of indicators can be found on the internet, in newspapers, magazines and financial papers.

  These indicators are called economic indicators. If these factors exist before those that they are going to affect, then these indicators are called leading indicators.

The main problem is to find the factors that are going to influence the demand of a specific item. It is even more difficult to find leading indicators, and that is why it is sometimes good to find non-leading ones for which other companies or the government has made forecasts. In that case, it means making forecasts on top of another forecast.

Extrinsic forecasting is mostly suitable for production and business planning rather than forecasting for individual end items. It is so because this method gives more reliable and precise data and results for big groups of items and for the total demand of a company's products. (ibid., 207 – 209.)

4.2 Demand

There are many factors that influence demand. Since there are so many of them and since it is difficult to find and consider all of them and their effect on demand, it is beneficial to pay attention to several major factors, such as:

- Market tendency (trends)
- Total business and economic state
- The company's own plans for marketing, costs, investments etc.
- Competition on the market

(Chapman & Arnold 2017, 202.)

Characteristics of demand

Demand has features that affect and influence the forecast. Moreover, based on these characteristics, particular forecasting techniques, which were described in the forecast chapter, are used. (ibid., 204.)

- Demand patterns

  If a person decided to make a graph of demand data over time, the plotted picture would show any shapes or consistent pattern of a time series. Demand is never stable, and the graph in Figure 3 (see below) represents well how it varies from period to period. There are four reasons which can explain why it happens:

  - Cycle

There are wavelike decreases and increases throughout years and decades in the demand, which is influenced and affected by economy. That is why such a task as forecasting of cycles exists and why it is done by economists.

- Seasonality

  Seasonality is a concept or event which happens on a yearly, monthly, weekly and even daily basis. In Figure 3, the graph shows how the demand variation depends on the time of the year. This variation can have different reasons, such as holidays, weather conditions etc.

- Trend

  By looking at Figure 3, it is easy to understand that demand is constantly increasing from year 1 to year 3. This specific graph represents a linear trend. However, there are different possible shapes (geometric and exponential) and types of trend (level, constant, rising and falling).

- Random variation

  Random variation happens when many factors influence demand during a particular time period on a random basis. Even though they are random variations, they also have a pattern which can be measured. There are two types of variations: small and big. A small one is when the real demand follows the pattern closely, and the big variation is when there are different individual data points broadly spread on the graph.

  (ibid., 204 – 205.)

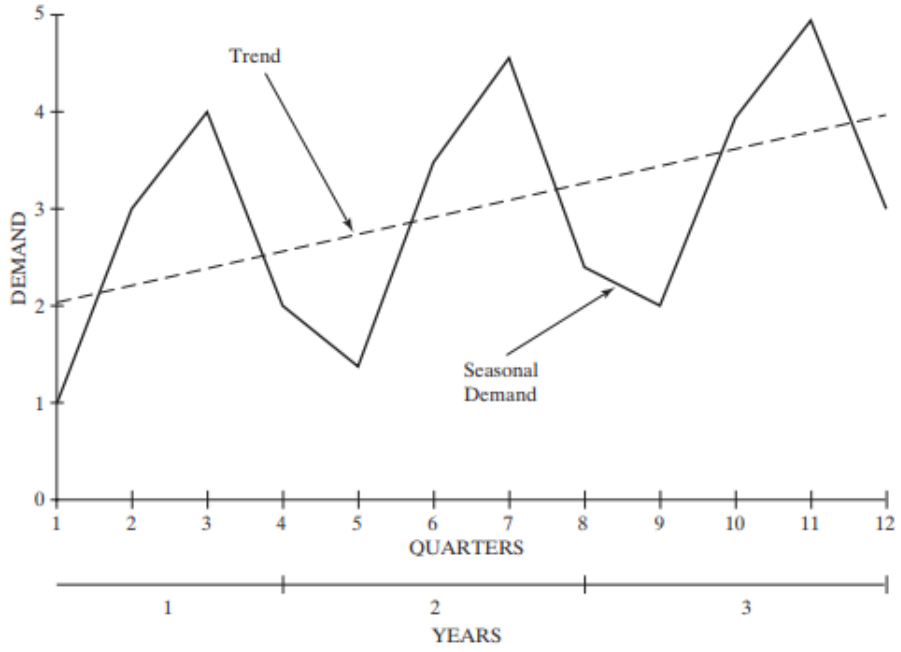

Figure 3. Example of a Graph of Demand Against a Time Scale (Chapman & Arnold 2017, 205).

- Stable versus dynamic.

  Certain products and services have shapes of the demand patterns which change over time, while others do not. The one that has the same pattern over a certain period of time is called stable, and others with changing patterns are called dynamic. According to Chapman and Arnold (2017, 205), it is easier to forecast when demand is stable. Figure 4 represents the graphics of stable and dynamic demand. It is possible to see that the average demand is equal in both cases, and that is why it is usually the one that is forecasted. (ibid., 205.)

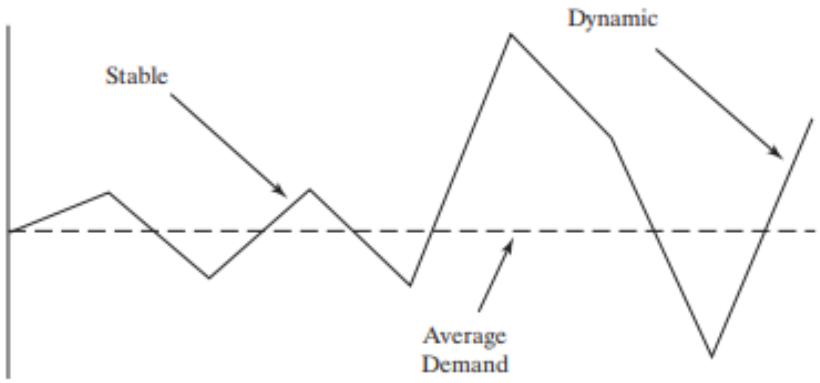

Figure 4. Example of Stable and Dynamic Demand (Chapman & Arnold 2017, 206).

- Dependent versus independent demand.

  Demand is called independent when it is not connected to any other product or service demands, as well as when it is not influenced by the internal activities of the company. When it comes to dependent demand, it is the exactly opposite situation to the independent one, and in that case, the demand of one item can be calculated from another item. That is why items with independent demand need forecasting, while those with dependent demand typically do not need it. (ibid., 206.)

4.3 Demand forecasting

It is fair to ask why companies need demand forecasting. There are many conditions and causes, but forecasting is unavoidable in creating plans to meet future demand. Most companies have “make to stock” production strategy, which means that they cannot wait until orders are placed to start planning and production. Moreover, customers and clients usually want their orders to be delivered as fast as possible, which means that the manufacturers should predict future demand and have capacity and resources to meet it. Even the companies with a “make to order” strategy must make a forecast and plan their resources of equipment, materials and labor to be available to satisfy demand. (Chapman & Arnold 2017, 202.)

Demand forecasting is an important function which affects different industries and companies all over the world. Knowing the future demand helps to decide on and adjust the amount of raw materials, safety stock, number of employees and other factors. In our fast world, companies cannot wait for demand to emerge and then react

in order to meet it. This is because in order to react fast to customer orders, companies need to feel and understand demand signals and create future demand in a way to satisfy customer wishes. (Chase Jr. 2009, 31 – 32.)

Making accurate demand forecasting has never been an easy task, but nowadays it is even more complicated, because the market environment is very inconstant and dynamic. Not only these factors but also globalization and lean manufacturing put a great deal of pressure on the supply chain, so creating effective forecasting in order to react to big swings in demand is becoming an impossible task. (ibid., 32.)

The global market has become very complex and competitive, and in order for a company to be successful and to make more profit, demand forecasting needs to be based on more dynamic signals like advertising, price, sales and many other factors. Focusing on those factors will help to create a more precise demand forecast and be successful on the market. However, companies do not only need to change from using simple demand signals to more complex ones, they also need to be customer obsessed. Lean manufacturing helps to build demand forecasting, but in order to be effective and efficient, companies need to make more accurate demand forecast in order to meet customer needs. (ibid., 32.)

In addition to a very competitive and dynamic market, companies also face problems in terms of creating a demand forecast due to cross-functional communication. Every center or department inside the company makes its own forecast and has its own ideas and goals that are good and suitable for them. However, the overall company is suffering due to these cross-functional forecast conflicts. The main three departments, such as marketing, finance and production, try to maximize their performance and profit, but they do it in very different and sometime opposite ways. (Chapman & Arnold 2008, 9.)

If a company is willing to have large profits, it should concentrate on at least these four main objectives:

- Having production cost as low as possible
- Having the lowest investments in the inventory
- Ensuring the best customer support and service
- Making delivery costs the lowest

Unfortunately, these four objectives create a conflict between the three departments of marketing, finance and production, which can be clearly seen in Figure 5 below:

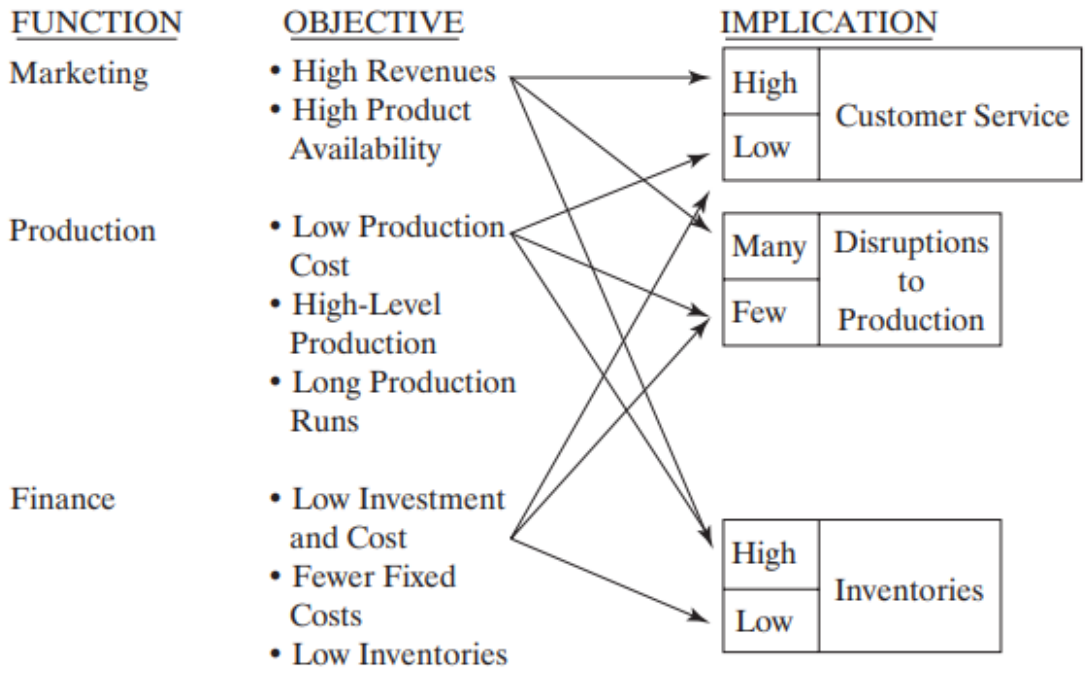

Figure 5. Cross-functional Conflict (Chapman & Arnold 2008, 10).

Based on all mentioned challenges, it can be maintained that making a demand forecast is currently a very difficult task but also a very important if a company wants to stay in the market and be able to compete. It is clear that demand forecasting is unavoidable due to the fact that almost all business decisions are about the future. However, companies can concentrate on the making forecast process as effective and accurate as possible by utilizing all available data, information and knowledge. When a forecast error is great, and the forecast is biased, it is due to the wrong method. Establishing a correct forecasting process will cause quite impressive improvements in the company and save money on investments in other areas. However, there are many challenges because of the dynamic and always changing market, fierce competition and cross-functional communication inside companies. (Chase Jr. 2009, 33; Kolassa & Siemsen 2014, 7 – 8.)

4.4 Methods of demand forecasting

Choosing the right demand forecast method is crucial for companies in order to minimize forecast error and maximize accuracy. Unfortunately, companies nowadays prefer to use very simple and judgmental forecast methods. They tend to choose forecast methods which their forecast makers are feeling comfortable to use, and with which everybody is satisfied with the outcome. However, the problem is that these methods might not be the most competent and suitable for the products which a particular company produces. The most common method is “target setting”. Technically, this method is not even a forecasting method but just a process of setting a mark or goal. Companies use that in order to grow their sales compared to the previous semester or year. It looks more like a target which is set by senior managers, and everybody else needs to check if they can meet this goal, and if not, then they just hope that someone will overachieve to close this gap. In this situation, statistical data usage is very low, or it is not present at all. No need to say that nobody even tries to think how to make demand more accurate by using marketing or sales opportunities in order to close the gap instead of relying on overdo chance. (Chase Jr. 2009, 77 – 78.)

One brilliant quality of all forecast methods that their accuracy and effectiveness can be assessed and judged after the process is done and actual demand happened. Moreover, methods can be compared to see which one is more accurate and effective. When a real demand appears, it is a must to compare forecast to it in order to evaluate forecast accuracy and create a picture over time. This comparison and evaluation are a real evidence if a forecast method is working and has been chosen correctly. Thus, when a demand forecast is far away from reality it is not a matter of belief, but a scientific and mathematical research. (Kolassa & Siemsen 2014, 10.)

Unfortunately, people who do forecast face growing amount of dissatisfaction towards the results of their job. This happens because of forecast makers being unable to precisely or accurately see how future circumstances can affect demand. Thus, people who use forecast, such as managers, become frustrated because of big oscillation in forecast accuracy and growing forecast error. These two factors dramatically affect decision making and planning, making them ineffective and not competent.

However, this environment where things do not go as planned allow people to understand that demand forecast requires even more serious attention and need to be done with a closer study of different levels of the company. (Chase Jr. 2009, 78.)

On the other hand, forecast users need to understand that forecasts with no error does not exist and it is not possible to achieve. There are boundaries to what can be forecasted, and all methods of forecasting are just estimation, calculations and guessing. Moreover, forecast will be more precise and with less error if objective methods will be used, which means that availability of previous (historical) data is crucial. However, it is clear that when sales pattern or market change there is no way to do forecast correctly or accurate. Thus, objective methods or quantitative ones are reliable as long as future has a constant pattern, which is based on the historical data. (ibid., 78.)

It is fair to ask what happens when objective methods are not providing enough accuracy. Then judgment is the only option in order to predict variation in market patterns or other factors. Yet this approach requires proper amount of regulations and observations. It is important to study trends on the market and possible volume or size of the change. This method has its own advantages and disadvantages. The main plus of this method is that it allows to notice change faster and react to it. However, the biggest problem of this method that it is relying on human's ability to be objective, which means that it is very common when people put their willingness to achieve a certain goal or perform better higher than understanding and evaluating real outcome of the forecast. This leads to the fact that forecast become bias and forecast error increasing dramatically. (ibid., 79.)

One common mistake, which happens towards forecast, is that people and companies sometimes think that it is a plan or a goal or a budget. However, forecast is not any of these things! In reality this three matters are often based on the demand forecast, which itself is just an assumption or estimation about future. Thus, companies need to be careful and not mix and misinterpret things, otherwise it will lead to bias and inaccurate demand forecast! (Kolassa & Siemsen 2014, 10.)

4.4.1 Different categories of methods

As mentioned earlier, there are two big groups of demand forecasting methods. Below more extensive and broad information about them is presented.

1. Subjective or qualitative methods
   Methods from qualitative group are based on the judgements or subjective evaluation of one person or group of professionals. This group is also called judgmental.

2. Objective or quantitative methods
   In quantitative group, forecast methods are found on the previous historical data itself or on a connection and correlations between this historical data and several variables. This group of methods is also called mathematical due to the big amount of formulas, graphs etc. which are involved in making process of forecast. That is why these methods can be written as algorithm and inserted into spreadsheets and software.

Even though it is clear that these two categories are different from each other, it is important to understand that often certain amount of subjective assessment is utilized in all methods of demand forecasting. So that quantitative methods are used as foundation and human judgement is based on it. The qualitative methods are done by utilizing knowledge of professionals, who have enough amount of expertise and experience in order to know what is currently happening in the market and what can possibly appear. Due to this, almost all forecast processes are made from the elements of these two groups.
(Chase Jr. 2009, 83 – 84; Kolassa & Siemsen 2014, 9.)

Three mostly common and popular qualitative methods are:

- Delphi method – a head of the project distributes a questionnaire among the team members of the project. This survey consists of different level of questions that are related to the market behaviors, specific company etc. Later on, results from this questionnaire are translated and interpreted into forecast.

- Sales force estimate – this is a collection of the forecasts from salespeople. Each salesperson makes a forecast for each product in the certain sales area, and then all these estimations are combined together.
- Jury of executive opinion – this technique is similar to the Delphi, but instead of questionnaire project team come together and have a meeting.

(Chase Jr. 2009, 84; Martinovic & Damnjanovic 2006, 528.)

Inside the quantitative group of demand forecast methods there are two categories which have own classes and techniques:

1. Time series

   Techniques which are used in time series category are also called reactive or one-dimensional. They are based on the theory or assumption that whatever happened in the past and taken as data from history will happen in the future. That is why these techniques have components, such as seasonality, level, trend and error. Thus, when these components are observed and taken from the historical data, patterns can be created or identified. Many techniques from this category requires stable time series in order to have constant parameters or components, so the forecast will be more accurate. This stability can be achieved by using for example differentiation technique. The challenge is that time series rarely can be stable, and it is important to understand how components will change over time. (Chase Jr. 2009, 84; Kolassa & Siemsen 2014, 40, 47; Chapman & Arnold 2008, 224.)

   As it was mentioned, time series has its own classes and techniques, some of them are:

   - Autoregression
   - Naïve
   - Moving Average
   - Exponential smoothing
     - Single
     - Holt Winter's two – parameter (HWES)
     - Brown's double
     - Winter's three – parameter
   - Seasonal decomposition

- ARIMA (Box-Jenkins)

(Chase Jr. 2009, 84; Burba 2019; Brownlee 2018.)

Time series category has quite many advantages. Below there is a list with main benefits of the time series methods:

- Simple in understanding and use
- Software packages already exist
- Reliable (when historical data represent a wide period of time)
- Good for the demand forecast of a big number of products
- Provide opportunity to measure financial growth
- Systemized easily
- Do not require large data storage
- Good at estimations of trends

However, a part of having quite many advantages, these techniques also having some great disadvantages:

- Cannot smooth out large random fluctuations
- Good only for short-term forecast
- Slow adjustability to changes
- Unstable sales history causes big errors
- Need a big amount of historical data
- Hard to find smoothing weights (alpha, beta, gamma)

(Chase Jr. 2009, 85 – 86; Cross 2017.)

2. Causal Models

These methods are more advanced and complicated compared to the time series. This is because apart of using just historical data there are several other variables that are added in order to create demand forecast. That is why these techniques are also called multidimensional or proactive. Companies often store different kind of information, which could be utilized in order to create a more accurate forecast. Examples of such information are purchase orders, which are already made, or some reservations for some products. Not only information which company have in their databases, but

also changes that are going to be made in sales, price or advertising can be quantified and used to create a demand forecast. (Chase Jr. 2009, 86; Kolassa & Siemsen 2014, 71.)

The most common and used causal model methods are:

- Simple linear regression
- Multiple regression
- ARIMAX

As well as time series methods, causal model ones also have some advantages and disadvantages. The main benefits are:

- More accurate than time series
- Has explanatory power
- Assists in "what – if" forecast
- Can be found in software packages
- Well – known for managers

The main disadvantages are:

- Time and cost consuming
- Not easy to be systematized
- Need deep understanding of statistics from managers
- Require big data storage
- Variables need to be forecasted very accurate

(Chase Jr. 2009, 86 -87; Sillignakis 2016, 4 – 5.)

4.4.2 Exponential smoothing

In 1970s seven experts created a competition where twenty-four time series forecasting methods were compared in order to understand why some of them provide

higher accuracy than others. Their research paper also provides a guide for companies in order to choose a right forecasting method depending on the data availability and other factors. One of the conclusions of this competition was that exponential smoothing is one of the strongest, competent and versatile time series forecasting methods. (Makridakis & Andersen & Carbone & Fildes & Hibon & Lewandowski & Newton & Parzen & Winkler 1982.)

Apart of having these good characteristics exponential smoothing also is easy to learn and interpret. Moreover, it can be utilized for a big amount of time series in real time quite easily, because of a small data storage demand and minimum calculation operations. (Kolassa & Siemsen 2014, 51.)

Exponential smoothing was created in order to improve and upgrade moving average forecasting method. It was necessary because moving average method was no longer useful when it comes to data which has seasonality, trends and other factors. Exponential smoothing technics require a special parameter which is called smoothing constant and in formulas is shown as a Greek letter alpha (α). This parameter allows to give weight to the past demand history and latest actual demand. Alpha is always represented by decimal from 0 to 1.0. (Chase Jr. 2009, 142 - 143; Chapman & Arnold 2008, 227.)

As mentioned earlier in this paper, one of the benefits of the exponential smoothing method is that no large data storage is required. This is because a new forecast is based on the old one and new data. So, in general the exponential smoothing formula for calculating a new forecast is looks as follows:

$$New\ forecast = \alpha \times latest\ demand + (1 - \alpha) \times previous\ forecast \quad (1)$$

The formula allows to give different weights to the most recent demand and to the previous forecast. The problem is to choose the most suitable alpha factor. This decision influences a lot on the forecast accuracy and ability to react fast to changes. When alpha factor is low then old data will be heavier than recent, which means that trends in case of change will not be sensed as soon as possible. However, when alpha is bigger, the new forecast will sense changing trends fast, but will be unstable when facing random oscillations. One of the best methods to choose a good alpha factor is by using computer simulations. (Chapman & Arnold 2008, 227 – 229.)

Exponential smoothing method performs well with stable items and short-time forecasting. It is not recommended to use when demand is not high and faltering. Moreover, as it was said before this method can sense trends, but in case of existence of a certain trend forecast will linger real demand. In the Figure 6 it can be clearly seen how different alpha influence on the ability of the forecast to follow actual demand with a trend. (ibid., 228.)

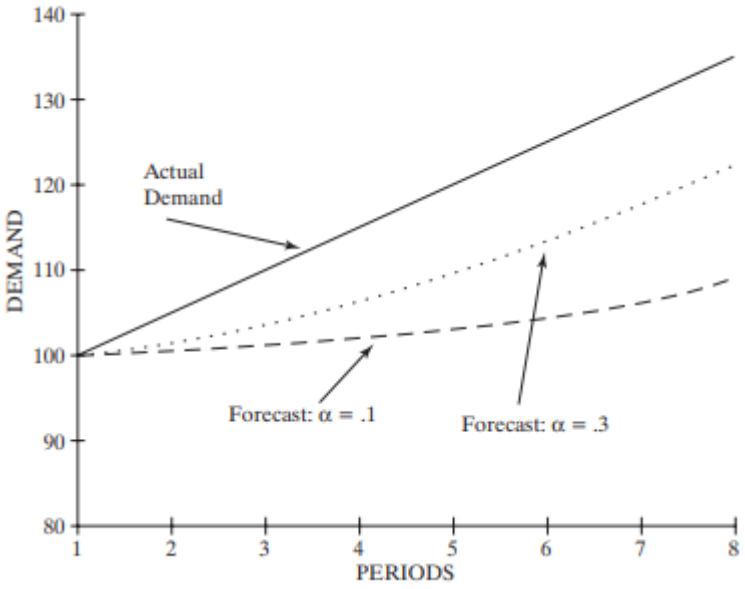

Figure 6. Influence of the Alpha Factor on Ability of the Forecast to Follow Actual Demand (Chapman & Arnold 2008, 228).

As it was mentioned earlier, in this research work two types of exponential smoothing will be used in the case study in order to test them on real data. The first method is single exponential smoothing, which is called like that because of only one formula (Formula 1) usage and existence of only one parameter (α) for overall smoothing. The second type is more complex and described in the sub-chapter below.

4.4.2.1 Holt Winter's two – parameter

This method is utilized when demand which should be forecasted has a trend. It is called two-parameter or double smoothing because forecast is based on two different values/parameters which are level and trend. This method has three separate formulas which are used together in order to create a final forecast. (Kluwer Academic Publishers 2000; Kalekar 2004, 3 – 4.)

Level – is a number which is a smoothed value of data estimation at the end of each period. Trend – is an average growth which is calculated as variation between the last two smoothed values. (ibid., ibid, 4.)

And as it was mentioned before this type of exponential smoothing has three different equations, one for trend another for level and last one for the forecast itself. In these equations $0 \leq \alpha \leq 1$ is a smoothed parameter for level and $0 \leq \gamma \leq 1$ is smoothed parameter for trend. (Hyndman & Athanasopoulos 2018, chapter 7.2; Kluwer Academic Publishers 2000.)

$$Level = \alpha \times latest\ demand + (1 - \alpha) \times (last\ period\ level + last\ period\ trend)$$

$$Trend = \gamma \times (\ latest\ level - last\ period\ level) + (1 - \gamma) \times last\ period\ trend$$

$$New\ forecast = latest\ level + latest\ trend$$

4.4.3 ARIMA Models

ARIMA stands for "Autoregressive Integrated Moving Average" models which can also be called Box-Jenkins models. In 1960 – 1970s George Box and Gwilym Jenkins provided sufficient amount of information in order to understand and use ARIMA, which became popular after that (Chase Jr. 2009, 203). This method of forecasting for instance demand, works well for different types of time series. However, it requires several calculations and estimations. That is why it is advisable to use computer software to find useful solutions. (SPSS 1994, 59.)

By the name of this method it is easy to understand that it consists of three types of processes:

- Autoregression (AR)
  It is an autoregressive process where dependent variable is a linear function of the past values of itself.
- Integration (I)
  It is a process of making data stationary which is very important in order to successfully do AR and MA. Stationary is achieved by predicting differences of the series from one time period to another.
- Moving averages (MA)

This process is similar to the AR, but instead of considering how previous values are affecting current it estimates how previous observations errors influence on current observations. Basically, it gives analysis of how wrong were predicted values of previous time steps in order to do better estimation for the current time step.

(SPSS 1994, 59 – 60; Rajbhoj 2019.)

As ARIMA consists of three processes, each of them is described by an integer and there they can get values from 0 to *n*. Typically the highest value for *n* is between 1…3. So general ARIMA model is traditionally written as ARIMA (*p,d,q*), where

- *p* – the order of autoregression or lag order, which shows the number of previous observations that are included into the model
- *d* – degree or order of differencing (important parameter which influence the success of the model)
- *q* – order of moving averaged involved or size of MA window

(SPSS 1994, 59; Rajbhoj 2019.)

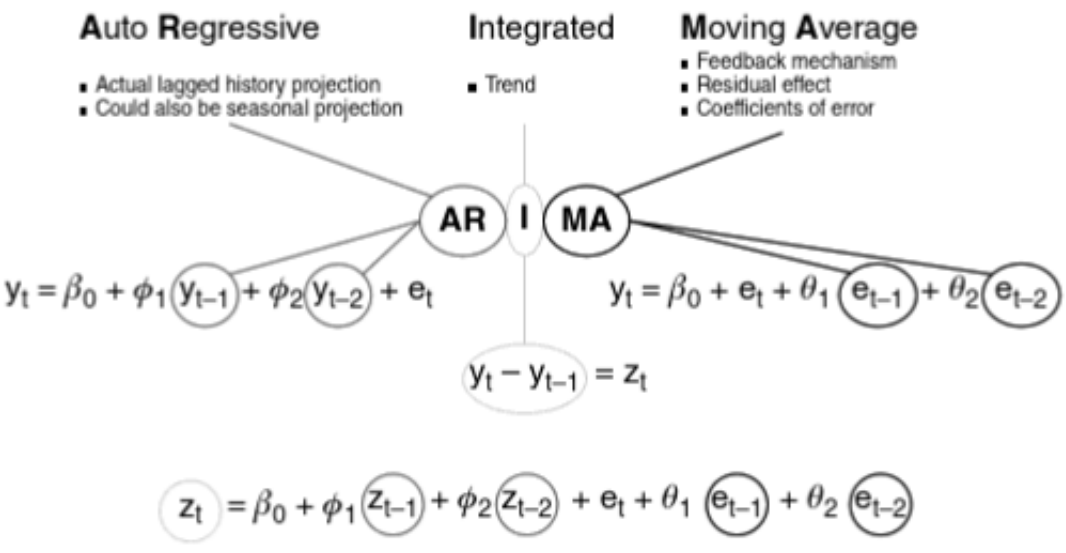

Figure 7. ARIMA (Chase Jr. 2009, 205).

For further details, see Chase Jr. 2009 page 205.

4.5 Forecast error

As mentioned, forecasts are not always right. The nature of forecasts includes errors as one natural factors of the forecasting processes. Thus, typically there are differences between forecasted demand and actual one, but the magnitude and regularity of those errors are based on these four situations:

1. Unstable patterns and random variations

   By using averages in statistical analysis people aim to find patterns and relationships, but almost every time there are inescapable variations around these averages. It leads to inaccurate and inexact results. Minimizing errors is the purpose of statistical analysis by sensing relationships and patterns, but it cannot provide total assurance that in the future errors will be systematic or random, or that they are going to stay inside the variation level which was identified. In some cases, demand will be stable with small variations, but in another unstable demand with big variations.

2. Inability to avoid bias

   Bias appears when the cumulative demand forecast is different from the actual one. This situation means that the forecast of the average demand was done incorrectly. When this happens, actual demand is constantly below or above the forecasted demand. In order to change the situation evaluation and changes should be done. It is highly important to track the forecast in order to be capable to react, plan and reduce errors. Moreover, when big errors happen it is significant to find and understand the cause.

   Cumulative demand should be tracked in order to avoid two most common situations which lead to errors: one-time events and timing. Examples of one-time events could be customer closedown, plans and operations of competitive firm. These events should be recorded and data about them collected and stored in the company. These actions will help to adjust the demand history and pay attention in the future when making forecasts.

3. Connections and patterns can vary from time to time

   With the time past by any relationship or pattern will change and sometimes it is quite hard to predict it. Continuous errors happened exactly because of these variations which are impossible to consider before they appear. How

big the error will be strictly connected to the volume and lasting period of the change.

4. Incorrect identification of relationships and patterns

   Sometimes a false pattern can be identified because of the deviation that actually does not exist. It happens due to different reasons and the most common one are:

   - Managers neglect the obvious factors
   - Managers concentrate on specific correlation
   - Little number of observations
   - Limited amount of data
   - Incompetent and wrong information
   - Too complex situation for modeling and forecasting

   Due to all these reasons which were listed it is advised to use domain knowledge and not just judgment. Analyzing and explaining results in a correct way where everything is making sense and no data is incorrectly utilized can be done with a help of domain knowledge. Moreover, this technique will take into consideration the marketplace dynamics.

   (Chapman & Arnold 2017, 216 – 217; Chase Jr. 2009, 91 – 93.)

4.5.1 Bias and accuracy

In this chapter such phenomena as bias and accuracy in forecast error calculations will be presented. By using several formulas, calculations and examples these two phenomena will be introduced.

Imagine a researcher calculated a single point forecast and after that real demand was observed. In this case formula for forecast error will look as follows:

$$e = \hat{x} - x, \text{ where}$$

$\hat{x}$ – single point of forecast

x – actual demand

e – forecast error

This formula allows easily understand that positive errors belong to overforecast ($\hat{x} > x$) and negative errors to underforecast ($\hat{x} < x$). Moreover, by restructuring this formula into:

$x = \hat{x} - e$, where

$\hat{x}$ – single point of forecast

x – actual demand

e – forecast error

such conclusion as real demand is equal the forecasted one minus error can be done. This error explanation is quite common, but not universal. There are a lot of different formulas which are used in different companies and businesses. It does not matter which definition of error company will choose to use as long as calculations are correct and everybody inside the company use the same one. (Kolassa & Siemsen 2014, 102.)

If a company wants to have a high quality forecast, then it should be based on a competent and large enough number of forecasts which were taken. That is why it is not worthwhile to calculate an error between only one forecast and real demand. Calculation of the error should be done based on a sufficient number of forecasts made with the same method. If researcher has v demand forecasts $\widehat{x_1}$ … $\widehat{x_v}$ and v real demand observations x_1 … x_v it allows to create a formula for v errors:

$e_1 = \widehat{x_1} - x_1 \ldots, e_v = \widehat{x_v} - x_v$, where

e_1 – error of the first forecast

e_v – error of the vth forecast

$\widehat{x_1}$ – first forecast

$\widehat{x_v}$ – vth forecast

x_1 – first actual demand

x_v – vth actual demand

The next step is formulating an equation which will represent a summary of those large amount of errors. The easiest way to create this formula is to take summary average – mean error (ME).

$ME = \frac{1}{v} \sum e_i$, where

v – number of observations

∑ - sign of summation

e_i – sum of errors

This is a main metric in estimation of a bias in a forecasting method. Bias can tell how on average a forecast is "on target". Forecast in unbiased if ME = 0, if ME > 0 then demand is overforecasting and if ME < 0 then it is underforecasting of demand. The bigger distance between ME and 0 the more bias forecast is! (ibid., 103.)

Such phenomenon as bias is quite important in forecasting errors, however companies are more interested in accuracy of the forecasting method. Accuracy represent how close forecasted demand is to the actual one. In order to measure accuracy often absolute difference between forecast and actual demands is used:

$$|e| = |\hat{x} - x|, \text{ where}$$

e – forecast error

$\hat{x}$ – single point of forecast

x – actual demand

$|\ldots|$ - absolute value brackets

Absolute error eliminates any plus or minus signs and shows number of units between forecasted and actual demands in either direction. Absolute error also has a simple way of calculating an average – MAE (mean absolute error) or MAD (mean absolute deviation). (ibid., 103.)

$$MAE = MAD = \frac{1}{v}\sum|e_i|, \text{ where}$$

v – number of observations

$\sum$ - sign of summation

$|\ldots|$ - absolute value brackets

e_i – sum of errors

Mean absolute error gives a chance to see if forecast method is on average accurate, so either forecast demand close to or far away from actual, and not taking in consideration the sign of the error (ibid., 104).

It is fair to ask what is more important: having an unbiased forecast method or the most accurate one. Moreover, is it possible that unbiased method can be not the most accurate. Well, to answer on some of those questions an example from the book "Demand Forecasting for Managers" will be presented:

Say, a company has a constant forecast of 11 for 10 months. However actual demand is $x = (10, 12, 10, 12, 10, 12, 10, 12, 10, 12)$ What would be ME and MAD?

$\hat{x} = (11, 11 \dots 11)$

$x = (10, 12, 10, 12, 10, 12, 10, 12, 10, 12)$

$v = 10$

$$ME = \frac{1}{v}\sum e_i$$
$$= \frac{1}{10} \times ((11-10) + (11-12) + (11-10) + (11-12) + (11-10) + (11-12) + (11-10) + (11-12) + (11-10) + (11-12) = 0$$

$$MAD = \frac{1}{v}\sum |e_i| = \frac{1}{10} \times 10 = 1$$

In this case forecast method is unbiased because ME = 0, but not accurate since MAD = 1. But what happens if forecast demand will be changed to $x = (9.5, 11.5, 9.5, 11.5, 9.5, 11.5, 9.5, 11.5, 9.5, 11.5)$

$$ME = \frac{1}{v}\sum e_i$$
$$= \frac{1}{10} \times ((9.5-10) + (11.5-12) + (9.5-10) + (11.5-12) + (9.5-10) + (11.5-12) + (9.5-10) + (11.5-12) + (9.5-10) + (11.5-12) = -0.5$$

$$MAD = \frac{1}{v}\sum |e_i| = \frac{1}{10} \times 5 = 0.5$$

In this case forecast method is bias and demand is constantly underforecasted because error is -0.5, but it is more accurate due to the fact that MAD = 0.5. In the Figure 8 below 2 graphs represents these two cases:

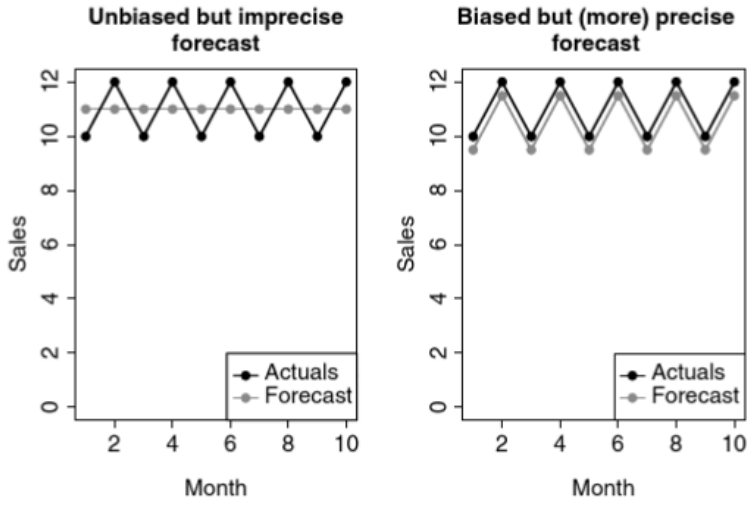

Figure 8. Bias vs. Accuracy Cases (Kolassa & Siemsen 2014, 104).

This is a dilemma which forecasters need to solve. It is a hard decision which every person who deal with forecasting need to make and choose whether use unbiased but imprecise method or bias and more accurate! It is not possible to tell, which of those is more correct or better, since it requires taking into account other things and plans, such as safety stock, lot sizes, production time. (ibid., 105.)

5 Case Study

The research data was collected from an industrial company. It is real-life data of one component (SKU) of an end-product, which the company is manufacturing. The purchasing price of the component is high, and its lead time is long. Therefore, demand forecasting is a necessity. As the SKU is a component, the nature of its demand of is dependent. It was mentioned earlier in sub-chapter 4.2 that typically, dependent demand does not require forecasts. However, this case is not typical and thus not following that theory, particularly as the end-product is an investment product.

Most often, when textbooks, journals and articles discuss forecasting of dependent and independent demand, they refer to consumer products. However, this research work and especially the case study data concern a production unit. Therefore, in that case forecasting of dependent demand is necessarily for the industrial component of an investment product. Otherwise, overstock or late production and not satisfying customer orders can easily happen.

In this case study, three different types of demand forecast were utilized in order to process the data. The results were compared, and the best method for this specific case was chosen. In order to process the data, graphs and different analyses with the IBM SPSS Statistics Data Editor program were utilized.

5.1 Case study data

The research data which was pre-collected consisted of 108 months of actual demand and the number of arrivals in stock. The last one´s numbers for 108 months are based on the previously done forecast by the company. The table below shows a descriptive statistical analysis of the case study data.

Table 8. Descriptive Statistical Analysis of the Case Study Data.

	N (number of months)	Minimum	Maximum	Mean
Demand	108	0	47	5,69
Arrivals	108	0	64	6,30
Valid N (listwise)	108			

Table 8 shows that both demand and arrivals had months with zero orders or arrivals. However, when it comes to the highest numbers it is noticeable that the difference is 17 SKUs, which is in this case quite significant taken into consideration that the price of the component is high, which leads to a high stock value.

In the collected data, we had one outliner which was month number 27. It was an outliner because it was so exceptional, and based on the data analysis, it turned out to be an inventory error. That is why, a new descriptive statistical analysis of the case study data needed to be created. The data is presented in Table 9 below.

Table 9. Descriptive Statistical Analysis of the Case Study Data without an Outliner.

	N (number of months)	Minimum	Maximum	Mean
Demand	107	0	24	5,31
Arrivals	107	0	43	5,76
Valid N (listwise)	107			

Even though the minimum number did not change in demand and arrivals when the outliner was taken away, the maximum number changed quite significantly in both

cases, which allowed to lower variations throughout the 107 months. However, the difference between demand and arrivals increased by two units and became 19 SKUs. In order to understand the demand patter, a histogram was created in the SPSS program which can be seen below.

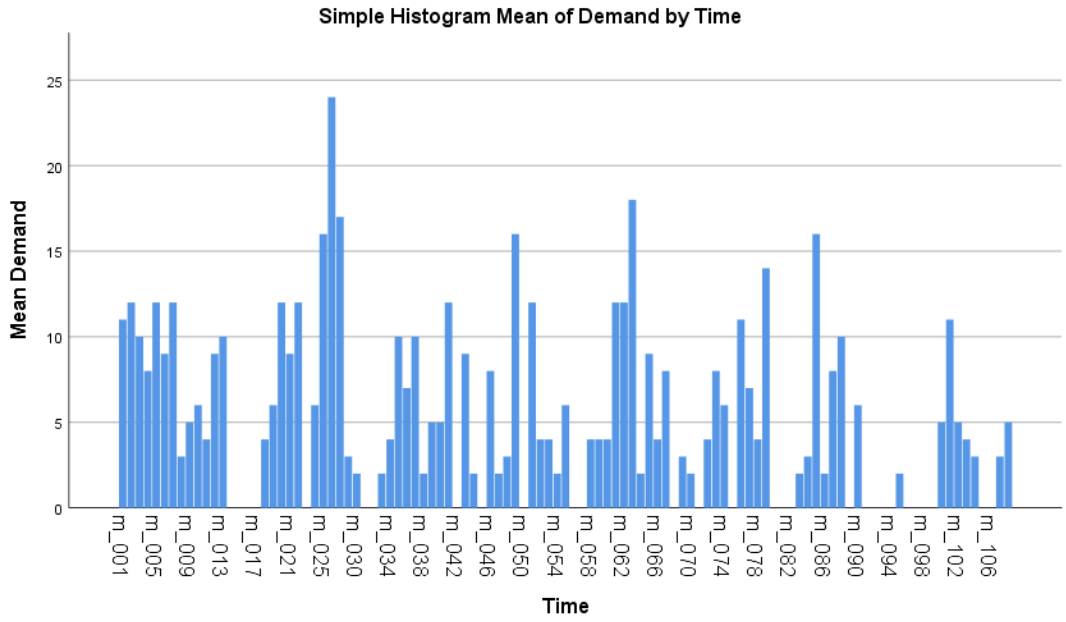

Figure 9. Histogram of the demand.

Figure 9 with the histogram represents the demand pattern throughout the 107 months, and it is clear how unstable demand is and how great the variations are. This leads to the conclusion that creating forecast for such demand is challenging. Not only demand is inconstant, but also arrivals which are based on the previously done forecast as well. This means that the forecast was made quite poorly. In order to deeply understand variations of the demand and arrivals, frequency tables of demand and arrivals were created, and they are presented below in Tables 10 and 11.

Table 10. Frequency Table of Demand.

Demand

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	29	27,1	27,1	27,1
	2	11	10,3	10,3	37,4
	3	7	6,5	6,5	43,9
	4	12	11,2	11,2	55,1
	5	6	5,6	5,6	60,7
	6	6	5,6	5,6	66,4
	7	2	1,9	1,9	68,2
	8	5	4,7	4,7	72,9
	9	5	4,7	4,7	77,6
	10	5	4,7	4,7	82,2
	11	3	2,8	2,8	85,0
	12	9	8,4	8,4	93,5
	14	1	,9	,9	94,4
	16	3	2,8	2,8	97,2
	17	1	,9	,9	98,1
	18	1	,9	,9	99,1
	24	1	,9	,9	100,0
	Total	107	100,0	100,0	

Table 11. Frequency Table of Arrivals.

Arrivals

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	59	55,1	55,1	55,1
	1	1	,9	,9	56,1
	2	1	,9	,9	57,0
	5	2	1,9	1,9	58,9
	6	1	,9	,9	59,8
	7	3	2,8	2,8	62,6
	8	4	3,7	3,7	66,4
	9	2	1,9	1,9	68,2
	11	1	,9	,9	69,2
	12	24	22,4	22,4	91,6
	14	1	,9	,9	92,5
	18	1	,9	,9	93,5
	20	1	,9	,9	94,4
	24	4	3,7	3,7	98,1
	36	1	,9	,9	99,1
	43	1	,9	,9	100,0
	Total	107	100,0	100,0	

Frequency tables of demand and arrivals show how many times each number of orders or arrivals were stated trough out 107 months of data. So, for example, from the table 10 it can be seen that 5 orders per months were done only in 6 months out of 107. Table 11 present that 20 SKUs should have been arrived only in one month among 107.

In this case, the challenges and problems are a result of a very unstable demand and poorly created previous forecast. The last one can be identified not only by variations in the arrival numbers but also by running a correlation analysis between demand and arrivals. The results of this analysis are presented in Table 12 below.

Table 12. Correlations Between Demand and Arrivals.

Correlations

		Demand	Arrivals
Demand	Pearson Correlation	1	,113
	Sig. (2-tailed)		,244
	N (number of months)	107	107
Arrivals	Pearson Correlation	,113	1
	Sig. (2-tailed)	,244	
	N (number of months)	107	107

Table 12 presents results, which shows that the previous forecast does not have a high quality. It can be seen that there is practically no correlation between demand and arrivals. The number of correlations is significantly low, as it is 0,113. Moreover, significant level from this table shows how likely the correlation is different from zero. In this case since probability is more than 5% (it is 24,4%) correlation is insignificantly different from 0. Thus, it can be stated that there is almost no correlation. However, if the forecast had been done properly, the number of correlations could have been higher and closer to 1 while significance lower and closer to 0, which allows to have less lags.

5.2 Single exponential smoothing

In order to create demand forecasting with the use of single exponential smoothing method the SPSS program was utilized. Pre-collected data from the company was download into the SPSS software. Then by utilizing tab “forecasting” exponential smoothing was chosen in order to create tables with statistical analysis of the method and estimated parameters of demand forecasting for this data. Moreover, tab “graphs” was used to create a graphical representation of forecasted values. In the tables and graph below results of the model can be seen.

Table 13. Single Exponential Smoothing Model Statistics.

Model Statistics

Model	MAE
Single Exponential Smoothing	4,182

Table 13 shows mean absolute error which is 4,18. This mean that every forecasted value will be ±4 SKUs away from actual demand value. It allows to create upper and

lower limits of the error for the forecasting line. This error can be considered quite significant for this study data since the mean value of the demand throughout 107 months is 5,31 from the table 9. Thus, the error on average is 73%. However, for each month where demand is higher than 10 SKUs the error will be less than 50%.

Table 14. Single Exponential Smoothing Model Parameters.

Exponential Smoothing Model Parameters

Model		Estimate	SE	Sig.
Single Exponential Smoothing	Alpha (Level)	,052	,030	,084

Table 14 shows smoothing constant – alpha, which was automatically estimated by SPSS program and calculated to be the best fit to this model using the provided data. Alpha was estimated to be 0,052 which is a quite small number, and knowing formula of single exponential smoothing, which is:

$$New\ forecast = \alpha \times latest\ demand + (1 - \alpha) \times previous\ forecast$$

it can be clearly seen that latest demand is having a small weight and does not play a big role in creating a new forecast. Moreover, by looking at significance level it allows to see that alpha is not significantly different from 0. It leads to the conclusion that even though the estimated alpha is the best fit, there is a probability of 8,4% that alpha is 0. Furthermore, standard error of estimated alpha is 0,30. It is quite large number since scope of data can be considered small for statistical analysis.

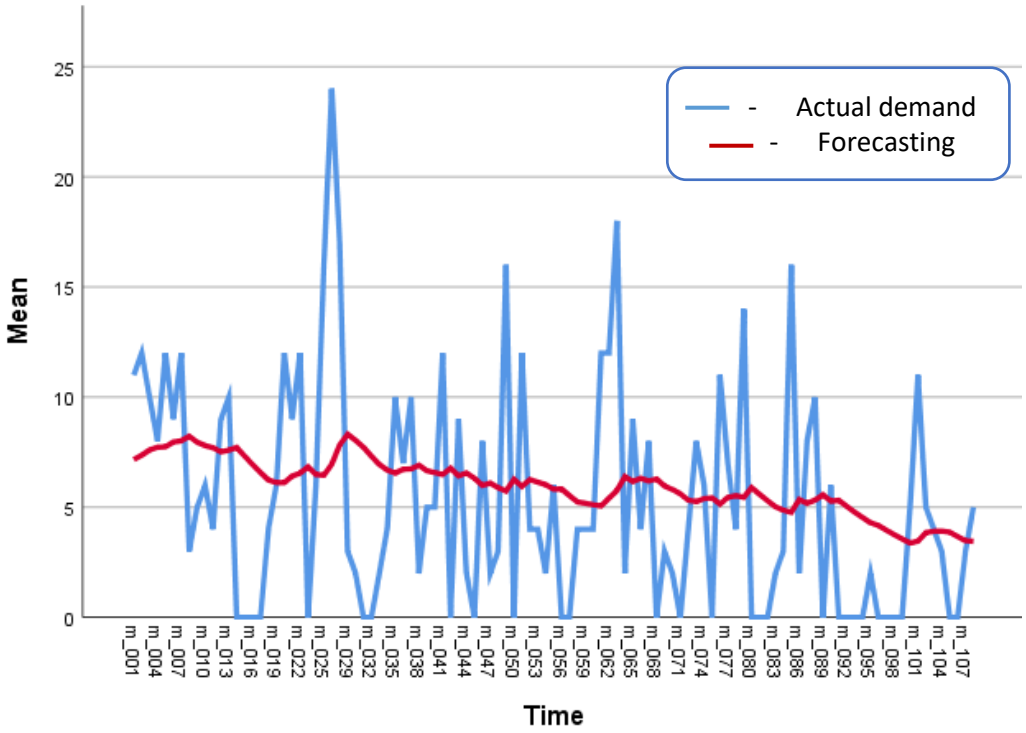

Figure 10. Demand Forecast Using Single Exponential Model.

Graph on Figure 10 shows two lines, where blue one represents demand of 107 month of pre-collected data and red – created demand forecast. One can notice that single exponential smoothing method tried to create more stable demand with less variations and smooth everything down. This method was able to sense a trend in data and notice drop at the end of the period. However, this method does not allow forecast to follow peaks of the actual demand well. During month 101 – 104 actual demand having a peak and forecast having a drop. Even though, exponential smoothing is a lagging model, utilizing this demand forecasting method in the situation like this will lead to big lags and inability to produce necessary amount of the investment product.

5.3 Holt Winter's two – parameter exponential smoothing

Process of creation demand forecast by utilizing two – parameter exponential smoothing with the help of SPSS software was similar to single exponential method. Tables and graph show results below.

Table 15. Holt Winter's two - parameter Exponential Smoothing Model Statistics.

Model Statistics

Model	MAE
Two – parameter exponential smoothing	4,034

Table 15 shows the mean absolute error of the double exponential smoothing model. It is lower compared to single exponential smoothing which means that this model is a little bit more precise than previous one.

Table 16. Holt Winter's two - parameter Exponential Smoothing Model Parameters.

Holt Winter's Exponential Smoothing Model Parameters

Model		Estimate	SE	Sig.
Two – parameter exponential smoothing	Alpha (Level)	,100	,052	,060
	Gamma (Trend)	4,466E-5	,047	,999

By looking at Table 16, it is noticeable that alpha parameter in this model of exponential smoothing is almost double the size of the alpha from single exponential smoothing. However, smoothing constant with this low value still makes latest demand quite irrelevant in creating a new forecast. Moreover, significance level or probability of alpha dropped compared to single exponential smoothing, however it is still insignificantly different from 0.

Since this model is double exponential smoothing, one extra parameter can be found in the table above. It represents trend of the given data. Estimated value of the trend is highly small. Furthermore, there is a probability of 99,9% that it is 0, so the conclusion that this data does not have trend can be made. Standard errors for level and trend are remarkable.

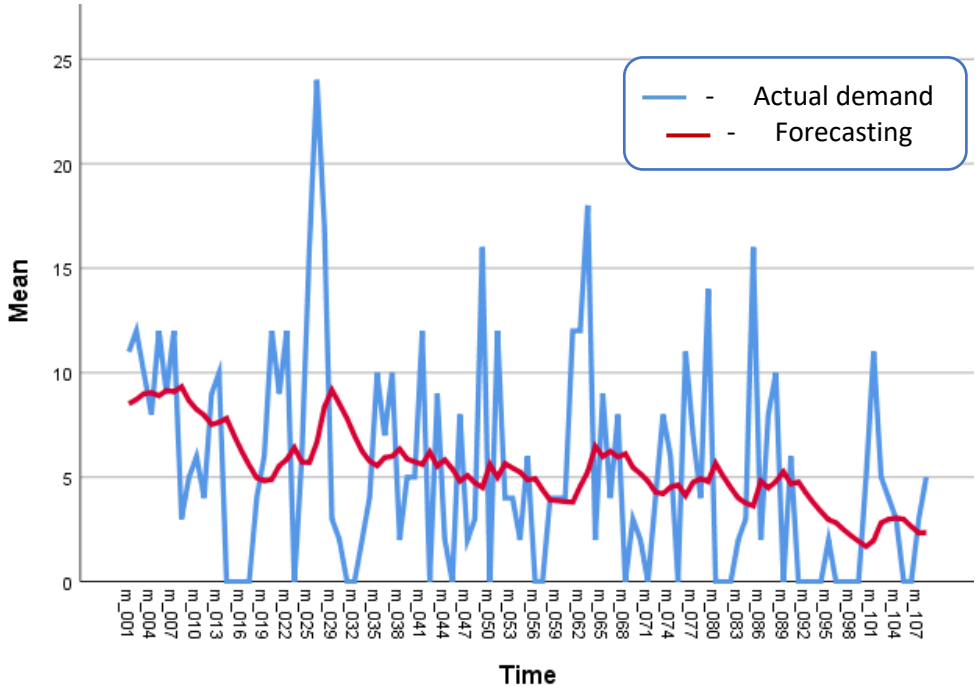

Figure 11. Demand Forecast Using Double Exponential Smoothing Model.

Graph on Figure 11 presents demand forecasting which was created in the SPSS program utilizing double exponential smoothing model. It is obvious that this model allows forecast to follow peaks of actual demand better compared to single smoothing. It leads to the fact that lags are smaller, and fulfillment of production plan is better.

5.4 ARIMA

In order to create demand forecasting by utilizing ARIMA method same steps in the SPSS software were done as in exponential smoothing methods. However, ARIMA allows to have different combinations of three parameters which represent three different processes. Since the most typical value for parameters is the range from 0 to 3, 11 different ARIMA models were created in the SPSS program in order to compare and choose the one with the most accurate forecast. In this chapter two ARIMA models are presented as they showed out to be the most suitable for this case study data. Other 9 rejected models can be found in Appendixes.

5.4.1 ARIMA (0,0,1)

Table 17. ARIMA (0,0,1) Model Statistics.

Model Statistics

Model	MAE
ARIMA (0,0,1)	4,074

Mean absolute error of the ARIMA (0,0,1) model is 4,074. This means that every forecasted value for demand is going to be with an accuracy of ±4 SKUs. It is lower than with single exponential smoothing, but more than with double exponential smoothing.

Table 18. ARIMA (0,0,1) Model Parameters.

ARIMA Model Parameters

			Estimate	SE	Sig.
ARIMA (0,0,1)	Constant		5,317	,583	,000
	MA	Lag 1	-,209	,096	,031

Table 18 presents estimated constant by SPSS which is the best fit for the ARIMA (0,0,1) model when it is applied to this case study data. Also, significant level shows that it is significantly different from 0 with a probability of 100%. This constant is almost equal to the mean of the demand data, which is 5,31. It can be explained as there are almost no correlations between months of demand and all of them act like individual ones. However, it is not possible to say that there are no correlations at all between months. This can be seen by analyzing moving average lag 1 value.

Moving Average Lag 1 here is a small number, however it is significantly different from 0. It allows to increase the accuracy of the model and it can be seen in MAE which is smaller than MAE in single exponential smoothing. Moreover, MA Lag 1 allows to understand that previous month demand has a small influence on the next month demand. For instance, when production has delays in the current month it will affect next month. Standard error for both parameters in this model is significantly lower than in previous models and is less than 50%. It allows to understand how reliable this model is and accurate. Furthermore, all the parameters add value to this model since both of them significantly different from 0.

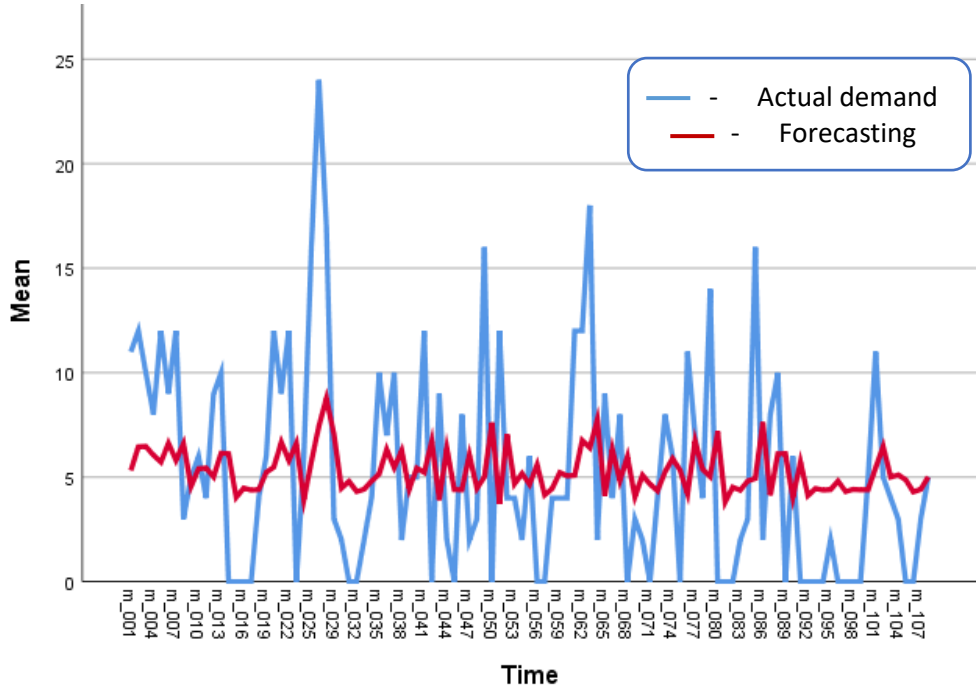

Figure 12. Demand Forecast Using ARIMA (0,0,1) Model.

Graph on Figure 12 represents forecasting which was created by utilizing ARIMA (0,0,1) model. It is noticeable that this method allows forecasting to follow peaks of the actual demand better than two previous exponential smoothing models. Thus, smaller lags can be achieved. Moreover, ARIMA (0,0,1) helps to smoothed down the variation and create more stable pattern.

5.4.2 ARIMA (0,0,2)

Table 19. ARIMA (0,0,2) Model Statistics.

Model Statistics

Model	MAE
ARIMA (0,0,2)	3,810

ARIMA (0,0,2) model has the lowest mean absolute error among all four chosen methods for creating demand forecast. It means that for every forecasted value of demand accuracy will be ±3 or 4 SKUs.

Table 20. ARIMA (0,0,2) Model Parameters.

ARIMA Model Parameters

			Estimate	SE	Sig.
ARIMA (0,0,2)	Constant		5,346	,705	,000
	MA	Lag 1	-,202	,093	,033
		Lag 2	-,310	,094	,001

Table 20 shows that all three parameters of the ARIMA (0,0,2) model are significantly different from 0. Therefore, they bring value for the model.

This model has one more lag in moving average parameter since the MA window became 2. This model helps to prove that even though there are almost no correlations between demand months, there is small influence of previous two months on the current one. Moreover, ARIMA (0,0,2) has small standard error, which is lower than 50% from the estimated parameters.

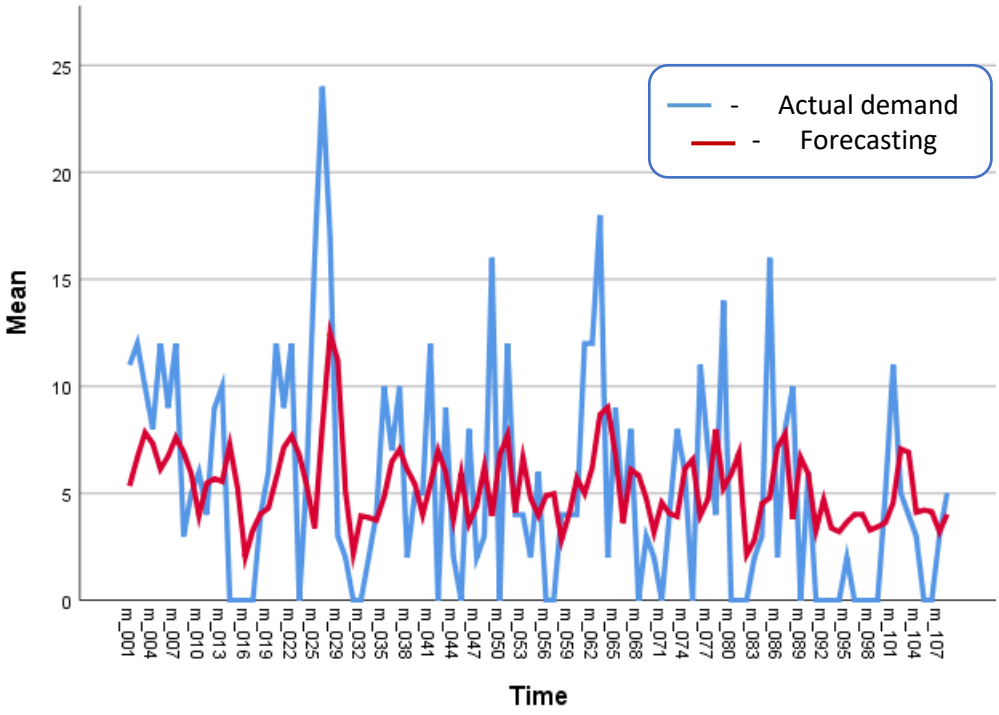

Figure 13. Demand Forecast Using ARIMA (0,0,2) Model.

Figure 13 shows demand forecast of ARIMA (0,0,2) model. Out of these selected four forecasting models, this model follows the peaks of the actual demand best. However, as a consequence, there is more variation between the lowest and highest forecasted values limits. From the graph it can be seen that the lowest forecasted value

will be around 2 SKUs and the highest around 13 SKUs which brings variation of 11 SKUs.

6 Analysis of the results and conclusion

6.1 Analysis of case study results

In this case study four different demand forecasting models were presented. Tables and graphs helped to analyze every single model and understand their flows. The first model of this case study, which was single exponential smoothing, showed out not to be the best fit for this case study data. This conclusion is based on following findings:

- The biggest MAE (mean absolute error) among all four different models
- Insignificance of estimated alpha parameter
- Big standard error
- Inability of the forecast to follow and sense peaks of the actual demand, which leads to big lags

By taking into consideration all the listed above disadvantages of the single exponential smoothing, this model was rejected.

The second tested model of the case study was double exponential smoothing or so-called Holt Winter's two-parameter exponential smoothing. This model as well cannot be considered the best fit for this specific case study, since two parameters of the model: level and trend are not significantly different from zero. It is not reasonable to use this model which takes into consideration trend of the data as trend actually does not exist in this specific demand data. Moreover, standard errors of the estimated parameters are bigger than 50%. Thus, this model was rejected as well.

The last two tested models represent ARIMA method for demand forecasting. Both of them can be considered as a best fit for this case study data. All estimated parameters of ARIMA (0,0,1) and ARIMA (0,0,2) are significantly different from zero and add values to the models. Moreover, standard errors of parameters are lower than 50%. Furthermore, created forecast sense and follow actual demand peaks more precisely, which leads to smaller lags.

6.2 Conclusions

Case study results show that two models of ARIMA fit the best this case study demand pattern. However, ARIMA (0,0,2) creates quite significant variations of 11 SKUs in the forecasted demand. Moreover, the aim of this thesis is to has a practical value for the companies and that is why ARIMA (0,0,1) is more useful rather than (0,0,2).

ARIMA (0,0,1) model is more useful and practically applicable for companies since it basically means moving average process. As it was mentioned earlier, ARIMA consist of three different processes: Autoregression, Integration and Moving Averages. Thus, model (0,0,1) means that two first processes were not included in the model, as only moving average process was.

Thus, the conclusion of this case study is that moving average forecasting method can be used in order to create better forecast for the provided demand pattern. It will help to lower variations and lags between demand and arrivals. Moreover, creating correlation between future demand and arrivals since they will be based on more accurate forecast.

Moving average is very applicable method in practice due to its simplicity and possibility to use Excel software for modeling. Utilizing ARIMA may be too complicated or demanding for practitioners.

7 Discussion

7.1 Discussion of research questions

This research work concentrated on two parts of the research: theoretical and case study (practical). The goal of theoretical part of this thesis, which was presented with literature review, was to study and understand the phenomenon of demand forecasting. Moreover, understand the value of forecasting and difficulties of creation quality and accurate one. Also, analyze and compare existing forecasting methods and techniques. On the other hand, case study objective was to use gained knowledge to the real life company situation in order to create the best demand

forecast. The main purpose of this thesis was to answer on three research questions. The first research question was:

1. *Why do companies need demand forecasting?*

In order to find the answer on this question the author studied the scientific literature about demand forecasting starting from the general concept of demand and its features and going into more details about usefulness and benefits of implementation of demand forecasting process in the company. Textbooks and articles about demand forecasting gave expressive understanding about a key role of this process and provided an opportunity to understand benefits for the company. Particularly, the author found out that forecast of future demand helps to plan and decide amount of materials, number of needed equipment and machinery, labor availability and size of safety stock. Therefore, no matter which production strategy company has, demand forecasting is needed in order to have capacity and resources to meet the upcoming demand.

Demand forecasting helps companies to stay on the market, compete and be relevant for the customers and clients. Today, as never before, customer service level and satisfaction need to be high in order to gain profit. Demand forecasting provide with possibility to react fast to orders and deliver product or service fast.

Nowadays, such question as “Does company need demand forecasting process?” should not exist anymore, since it is clear that demand forecasting in unavoidable when it comes to almost all business decisions. Companies need to concentrate on making the process as accurate and effective as possible. It can be done by smartly utilizing all available data, information, tools and knowledge. It is crucial for companies to choose a correct forecasting method and technique in order to avoid big errors and bias. Therefore, second research question was:

2. *Which methods of demand forecasting exist and how can companies choose the right one?*

To answer this question, the author studied significant amount of existing scientific literature about forecasting methods, categories and techniques. Moreover, literature review helped to understand on what companies need to focus in order to choose the best forecasting method. Furthermore, case study allowed the author to

practically show comparison of two different methods and four different models in order to choose the best fit for the particular demand pattern.

Theoretical study made the author understand that forecasting methods can be divided into two big groups of quantitative methods and qualitative methods. Both of the groups have advantages and disadvantages. However, the main focus of this research work was quantitative methods and more specifically time series category of forecasting techniques.

Existing scientific literature provides expressive amount of information about different forecasting methods, their limitations, advantages and application. All the available information allows to study each method in details and make a choice towards the right and most suitable one.

In order for companies to choose the right forecasting method and avoid unnecessary inaccuracy and errors, they should pay attention on what kind of data, tools and knowledge is available for use. Since these resources have big influence on the choice of the forecasting method. For example, in order to successfully implement forecasting method from qualitative group significant amount of knowledge is required from forecaster. The person who will create the forecast based on judgment need to understand such thing as market trend and possible changes. Moreover, that kind of forecast will be more accurate if company has team of professionals who are qualified enough to create predictions of the state of the future market.

On the other hand, quantitative methods require availability of the historical data of previous years' actual demands and previously done forecasts. Moreover, these methods usually include high amount of calculations and statistical analysis which can be done easily in different software. Thus, company need to have forecaster who can successfully use software and has sufficient amount of knowledge in statistics. In some cases, apart of historical data several other variables are needed to create demand forecast, for example, in causal models of forecasting.

Finally, company can use one very important quality of all forecast methods in order to choose the right one. This quality is an opportunity to assess and judge the accuracy and effectiveness of the forecast after the process is done and actual demand happened. Forecast must be compared to the actual demand in order to understand

how accurate it was. Moreover, different methods can be tested in order to compare and see which one is more accurate and effective. This comparison and evaluation are the proof if a forecast method is working and was chosen correctly.

In this thesis, case study allowed to test four different models on the same data in order to compare them and choose the most suitable one. In the chapter Case study and Analysis of the case study results, the author provided statistical analysis of each model and shows why some models are a good fit for this specific demand patter and why others are not.

However, even when companies successfully implement forecasting process and method is chosen correctly errors still occur. It is possible to minimize errors by choosing the right method and it was clearly seen in the case study, but unfortunately free error forecast does not exist. Therefore, it was crucial to understand where forecasting errors are coming from, and thus the third research question was:

3. *What causes the forecast error?*

Going through multiple informational sources helped to understand that certain reasons which are causing errors in a forecast can be avoided, but not all. This is exactly why forecasts cannot be done without error.

When it comes to the human mistakes in creating forecasts such as wrong sensing of the pattern, utilizing wrong information, adding bias and others. These things can be avoided by hiring more knowledgeable forecasters and paying closer attention to the actions.

However, there are certain things that cannot be sensed in advance, before they happen, and thus their influence on the accuracy of the forecast cannot be prevented. Such things as random variations or pattern change happened in demand. Also, one-time events such as close down of the market or new strategy of the competitive company. All of those affect demand accuracy and unfortunately cannot be predicted early in advance. Moreover, fast changing trends in customer preferences and behaviors and fast growing market as well does not help with minimizing forecast errors.

7.2 Discussion of study hypotheses

In the beginning of this research work two hypotheses were formed:

1. *Correlation does exist between demand and arrivals in the case study data.*
2. *It is possible to create an accurate forecast by using chosen demand forecasting methods based on the case study data.*

However, both of the hypotheses turned out to be wrong and were rejected in the process of writing this thesis.

First hypothesis was rejected by utilizing the SPSS software where correlation table was created. In the Case study chapter table was presented and explained where it can be seen that demand and arrivals actually do not have correlations which leads to the situation of having large lags and inability to fulfil the production plan.

Second hypothesis was rejected since in the results of case study it was stated that Moving Average method will be the most suitable and suggested demand forecasting method for this particular case study data. Since MA method was not chosen one for the case study, this hypothesis was wrong. However, ARIMA (0,0,1,) which was one of the best fitting forecasting models from the chosen one, basically mean MA process was not recommended since it is more complicated and less practical in use. By applying ARIMA models for the case study data it was possible to find out that MA method will be the best for creating forecast for such unstable demand.

7.3 Suggestions for future research

The aim of the case study was to create a better forecast for the company and improve the demand forecasting process by selecting the most suitable method. However, in order to improve the accuracy of the forecast the suggestion is to smoother the actual demand.

During case study, graph of actual demand was created and it was clear that it is highly unstable and variations are large. In that kind of case with such challenging demand patter creation of the accurate forecast become a very hard task. However, it is possible to make demand more smooth by creation better cross-functional collaboration inside the company. As the reference to Figure 5 with the cross-functional

conflict the suggestion is to create better communication between production and sales/marketing teams. It will allow deeper understanding of strategies which different teams have and possibility to explain which challenges can arise from the made decisions.

Furthermore, for future research Moving Average method for creating demand forecast can be discussed more into details by utilizing available scientific literature. Moreover, in the practical part of the case study MA method can be applied in order to create forecast. By utilizing Excel software tables and graphs can be created for this method.

Finally, theoretical part of this research work can be used as foundation for future researches in the similar field. Conducted literature review is reliable since sufficient amount of scientific resources of good quality was utilized. Results of the case study are very reliable as well, since all the parameters, tables and graphs were created in the SPSS software and human mistakes in calculations are minimized. Moreover, statistical analysis of the tables was written with close guidance of two professional university teachers. However, this reliability of case study results is only concerning this particular data. It is not a case for generalization.

Acknowledgements

Lastly, I would like to write how grateful I am for the support and help which I received during the process of writing this thesis. First, I would like to express my huge amount of respect towards my main thesis supervisor and an amazing teacher from JAMK Mr. Juha Sipilä who made my thesis possible. For his constant support, guidance and assistance I am forever grateful! Secondly, I would like to appreciate the help from Mr. Hannu Lähdevaara, who was my second thesis supervisor, for important inputs and advices. Thirdly, I want to express my gratefulness towards Mr. Toni Nieminen for essential assistance in the case study part with the SPSS program. Last, but not least, I would like to thank Mr. Hannu Ryynänen whom I appreciate for his huge efforts which he put into the English grammar and reference style of this research work!

Finally, I would like to express my gratefulness to all my friends and family members who supported me during this difficult period of my life. All of you motivated and empowered me to move forward no matter how hard it was!

All of you have contributed to the success of this thesis and my graduation and for this I am forever grateful.
Thank you all,
Nadezhda Zubova.

References

Advantages and Disadvantages of Qualitative Research. N.d. OCCUPYTHEORY 21 April 2014. Accessed on 27.03.2019. Retrieved from http://occupytheory.org/advantages-and-disadvantages-of-qualitativeresearch/

Aitchison, J. (1998). *Access to books and journal articles by post-graduate students on a course-work master's programme in information studies at the University of Natal, Pietermaritzburg*. MIS Thesis, University of Natal, Pietermaritzburg.

Alshenqeeti, H. (2014). *Interviewing as a Data Collection Method: A Critical Review*. Newcastle upon Tyne, United Kingdom.

Biggam, J. (2008). *Succeeding with Your Master's Dissertation.* London: Open University Press.

Bless, C., Higson-Smith, C. (2000). *Fundamentals of social research: An African perspective. 3rd ed.* Cape Town: Juta.

Braun, V., Clarke, V. (2013). *Successful Qualitative Research: A practical guide for beginners.* London: SAGE Publications Ltd.

Brownlee, J. (August 6th, 2018). *11 Classical Time Series Forecasting Methods in Python (Cheat Sheet).* Accessed on 29.10.2019. Retrieved from https://machinelearningmastery.com/time-series-forecasting-methods-in-python-cheat-sheet/

Burba, D. (October 3rd, 2019). *An overview of time series forecasting models.* Accessed on 29.10.2019. Retrieved from https://towardsdatascience.com/an-overview-of-time-series-forecasting-models-a2fa7a358fcb

Chapman, S., Arnold, T. (2017). *Introduction to Materials Management.* 8th ed. London: Pearson Education.

Chapman, S., Arnold, T. (2008). *Introduction to Materials Management.* 6th ed. London: Pearson Education.

Chase Jr., C. (2009). *Demand – driven Forecasting: a structured approach to forecasting.* 2nd ed. New Jersey: John Wiley & Sons, Inc.

Crawley, Michael J. (2014). *Statistics: An Introduction Using R.* 2nd ed. John Wiley & Sons.

Creswell, J. (2014). *Qualitative, Quantitative and Mixed Methods Approaches, Research Design*. London: SAGE Publications Ltd.

Cross, V. (September 26th, 2017). *The Advantages of the Time Series Method of Forecasting.* Accessed on 29.10.2019. Retrieved from https://bizfluent.com/info-8610535-advantages-time-series-method-forecasting.html

Dahlberg, L., McCaig, C. (2010). *Practical Research and Evaluation*. London: SAGE Publications Ltd. Accessed on 26.03.2019. Retrieved from

https://books.google.fi/books?hl=en&lr=&id=kdXiJ1ZSF1MC&oi=fnd&pg=PP1&dq=+Practical+Research+and+Evaluation&ots=N5U5QLhUyU&sig=ZH779BFhMFYDqv0m0AMoqMu95rg&redir_esc=y#v=onepage&q=Practical%20Research%20and%20Evaluation&f=false

Denscombe, M. (2003). *The Good Research Guide: for small-scale social research projects.* 2nd ed. Open University Press.

DME for peace: Section 1.3 Quantitative Research Module. N.d. Search for Common Ground. Accessed on 28.03.2019. Retrieved from http://dmeforpeace.org/sites/default/files/1.3%20Quantitative%20Research.pdf

Dr. Mary Dowd. *Advantages & Disadvantages of Qualitative & Quantitative Research.* Accessed on 28.03.2019. Retrieved from https://www.theclassroom.com/advantages-disadvantages-of-qualitative-quantitative-research-12082716.html

Hox, Joop J., Boeije, Hennie, R. (2005). *Data Collection, Primary vs. Secondary.* Elsevier Inc.

Hyndman, R. J., Athanasopoulos, G. (2018). *Forecasting: principles and practice, 2nd edition*. OTexts: Melbourne, Australia. Accessed on 18.11.2019. Retrieved from https://otexts.com/fpp2/

Introduction to Confidence Intervals (n.d.). Page on PennState Eberly College of Science's website. Accessed on 24.11.2019. Retrieved from https://onlinecourses.science.psu.edu/stat200/lesson/4/4.2

Kalekar, P. (2004). *Time series Forecasting using Holt-Winters Exponential Smoothing.* Kanwal Rekhi School of Information Technology.

Kananen, J. (2011). *Rafting Through the Thesis Process.* JAMK University of Applied Sciences.

Kluwer Academic Publishers (2000). *HOLT'S FORECASTING MODEL*. In: Swamidass P.M. (eds) Encyclopedia of Production and Manufacturing Management. Springer, Boston, MA. Accessed on 18.11.2019. Retrieved from https://link.springer.com/referenceworkentry/10.1007%2F1-4020-0612-8_409

Kolassa, S., Siemsen, E. (2014). *Demand Forecasting for Managers.* Business Expert Press, LCC, 2016.

Mack, N., Woodsong, C. (2005). *Qualitative Research Methods: A Data Collector's Field Guide.* Family Health International 2005. Accessed on 27.03.2019. Retrieved from https://batukarinfo.com/system/files/Qualitative%20research%20-%20datacollector-guideenrh.pdf

Makridakis, S., Andersen, A., Carbone, R., Fildes, R., Hibon, M., Lewandowski, R., Newton, J., Parzen, E., Winkler, R. (1982). *The Accuracy of Extrapolation (Time Series) Methods: Results of a Forecasting Competition.* Journal of Forecasting, Vol. 1. John Wiley & Sons. Ltd.

Martinovic, J., Damnjanovic, V. (2006). *The Sales Forecasting Techniques.* Faculty of Economic and Management SAU in Nitra.

Olusegun, Sogunro A. (September 2002). *Selecting a Quantitative or Qualitative Research Methodology: An Experience.* Educational Research Quaterly, Vol. 26.1. EBSCO Publishing.

Ostlund, U., Kidd, L. (2010). *Combining qualitative and quantitative research within mixed method research designs: A methodological review.* Elsevier Ltd: International Journal of Nursing Studies. Access on 27.03.2019. Retrieved from https://repository.edulll.gr/edulll/bitstream/10795/2235/2/2235_2%20%20QUANTITY%20QUALITY.pdf

Rabianski, Joseph S. (January 2003). *Primary and secondary data: Concepts, concerns, errors, and issues.* The Appraisal Journal, Chicago. Access on 24.09.2019. Retrieved from https://search.proquest.com/openview/68bfe6580297494936f0126ed5d3d017/1?pq-origsite=gscholar&cbl=35147

Rajbhoj, A. (September 26th, 2019). *ARIMA simplified. A simplistic explanation to the most popular Time Series Forecasting model out there.* Page on Towards Data Science website. Accessed on 20.11.2019. Retrieved from https://towardsdatascience.com/arima-simplified-b63315f27cbc

Robinson, D., Reed, V. (1988). *The A – Z of social research jargon*. England: Ashgate.

Saukkonen, A. (2017). *Development of Materials Management System: Case Black Bruin Inc.* Bachelor's thesis. JAMK University of Applied Sciences, School of Technology, Communication and Transport, Degree Programme in Logistics Engeering. Accessed on 24.03.2019. Retrieved from https://www.theseus.fi/bitstream/handle/10024/135696/Development%20of%20Materials%20Management%20System.pdf?sequence=1&isAllowed=y

Sillignakis, Kostas E. (2016). *The relative advantages and disadvantages of the causal and non-causal approaches to tourism demand forecasting.* Accessed on 29.10.2019. Retrieved from https://docplayer.net/15116047-The-relative-advantages-and-disadvantages-of-the-causal-and-non-causal-approaches-to-tourism-demand-forecasting.html

Silverman, G. (December 1st, 2017). *Survey Finds Only 13% of Companies Are Effective At Forecasting.* Page on Concentric. Accessed on 2.12.2019. Retrieved from https://concentricmarket.com/blog/survey-finds-only-13-of-companies-are-effective-at-forecasting

SPSS (1994). *SPSS Trends 6.1.* SPSS Inc.

SPSS Tutorial (July 2019). *Simple Introduction to Confidence Intervals.* SPSS Inc. Accessed on 24.11.2019. Retrieved from https://www.spss-tutorials.com/confidence-intervals/

SPSS Tutorials (n.d.). *What Does "Statistical Significance" Mean?* SPSS Inc. Accessed on 24.11.2019. Retrieved from https://www.spss-tutorials.com/statistical-significance/

Stewart, A. (2016). *Basic Statistics and Epidemiology: A Practical Guide.* 4th ed. Boca Raton: CRC Press.

Taff, A. (2019). *Hypothesis Testing: The Ultimate Beginner's Guide to Statistical Significance.* Venture Ink. Accessed on 24.11.2019. Retrieved from https://books.google.fi/books?id=T17FxwEACAAJ&dq=what+is+statistical+significance&hl=en&sa=X&ved=0ahUKEwihicGvr4LmAhUQyKYKHWKRAUYQ6AEIRjAE

TradeGecko (September 13th, 2018). *What is demand forecasting and how can it help your business?* Accessed on 2.12.2019. Retrieved from https://www.tradegecko.com/blog/inventory-management/what-is-demand-forecasting-and-how-can-it-help-your-business

Walliman, N. (2011). *Research Methods the basics*. Routledge 2011. Accessed on 26.03.2019. Retrieved from http://uni.delf.pro/uploads/7/1/0/7/7107980/research_methods_the_basics.pdf

Westat, F. J. 2002. The User-Friendly Handbook for Project Evaluation. The National Science Foundation Directorate for Education & Human Resources Division of Research, Evaluation, and Communication, January 2002. Accessed on 09.06.2019. Retrieved from https://www.nsf.gov/pubs/2002/nsf02057/nsf02057.pdf

Williams, K. P. (October 22nd, 2019). *Confidence interval for the financial sector: …a measurement paradox.* Accessed on 24.11.2019. Retrieved from https://thebftonline.com/2018/features/confidence-interval-for-the-financial-sector-a-measurement-paradox/

Yin, Robert K. (2009). *Case Study Research: Design and Methods.* 4th ed. SAGE Gecko

Appendices

These nine ARIMA models were rejected since either all or some parameters of the models are insignificantly different from 0. This means that kind of parameters do not add new information and value to the model. Moreover, in some of the models standard error of the parameters are higher than at selected ones. Even though more complicated ARIMA models with more parameters allows forecast to follow demand more precisely, it looks more like overfitting. In that kind of case generalization cannot be done and predictions will be wrong since all the parameters are random and insignificantly different from zero.

Appendix 1. ARIMA (0,1,0)

Table 21. ARIMA (0,1,0) Model Statistics.

Model Statistics

Model	MAE
ARIMA (0,1,0)	4,504

Table 22. ARIMA (0,1,0) Model Parameters.

ARIMA Model Parameters

		Estimate	SE	Sig.
ARIMA (0,1,0)	Constant	-,057	,595	,924
	Difference	1		

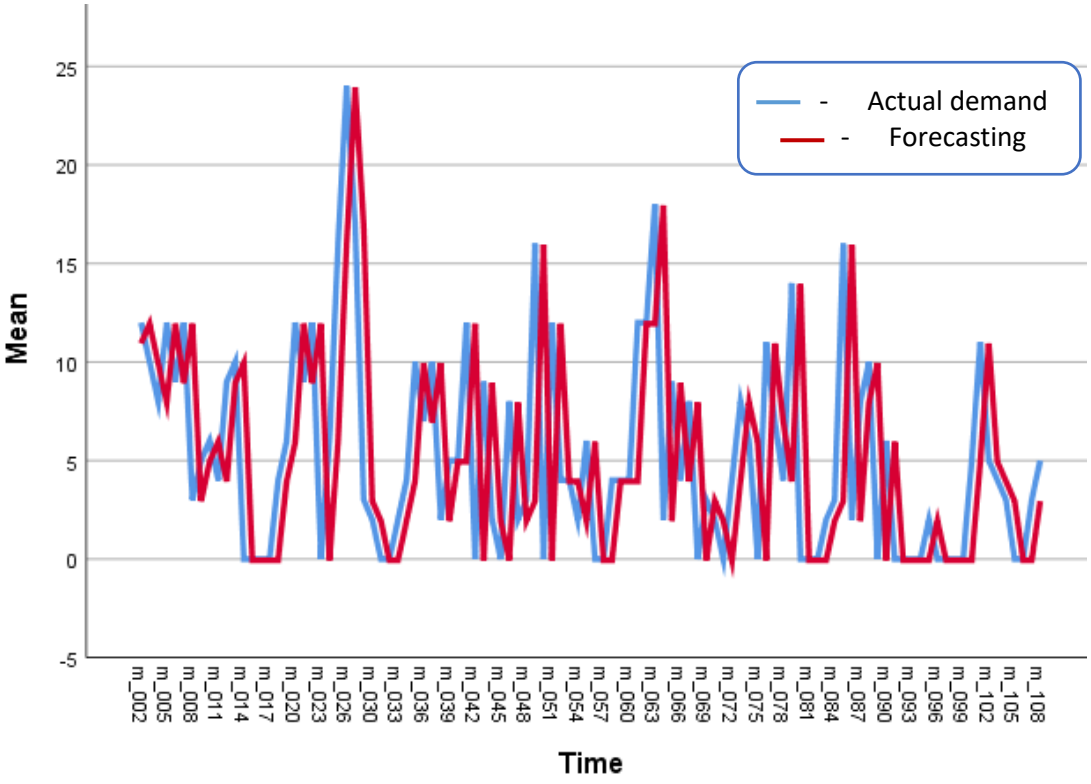

Figure 14. Demand Forecast Using ARIMA (0,1,0) Model.

Appendix 2. ARIMA (1,0,0)

Table 23. ARIMA (1,0,0) Model Statistics.

Model Statistics

Model	MAE
ARIMA (1,0,0)	3,987

Table 24. ARIMA (1,0,0) Model Parameters.

ARIMA Model Parameters

			Estimate	SE	Sig.
ARIMA (1,0,0)	Constant		5,328	,666	,000
	AR	Lag 1	,287	,093	,003

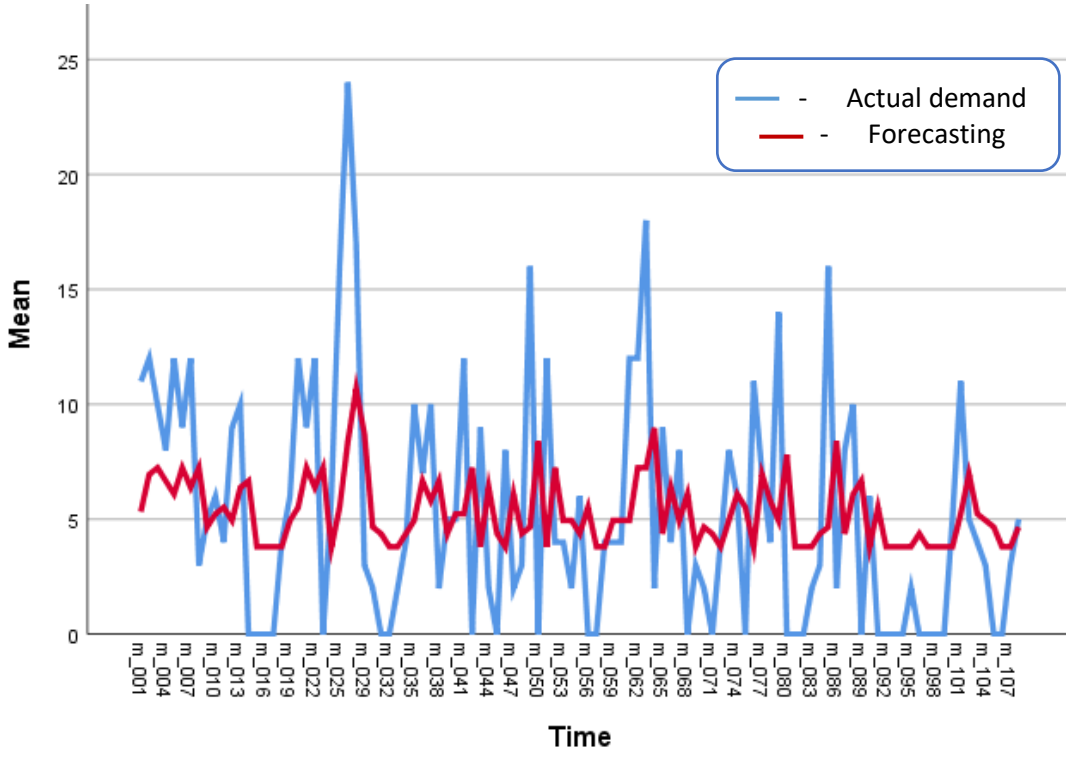

Figure 15. Demand Forecast Using ARIMA (1,0,0) Model.

Appendix 3. ARIMA (1,1,1)

Table 25. ARIMA (1,1,1) Model Statistics.

Model Statistics

Model	MAE
ARIMA (1,1,1)	3,857

Table 26. ARIMA (1,1,1) Model Parameters.

ARIMA Model Parameters

			Estimate	SE	Sig.
ARIMA (1,1,1)	Constant		-,046	,020	,025
	AR	Lag 1	,251	,104	,018
	Difference		1		
	MA	Lag 1	,998	,995	,318

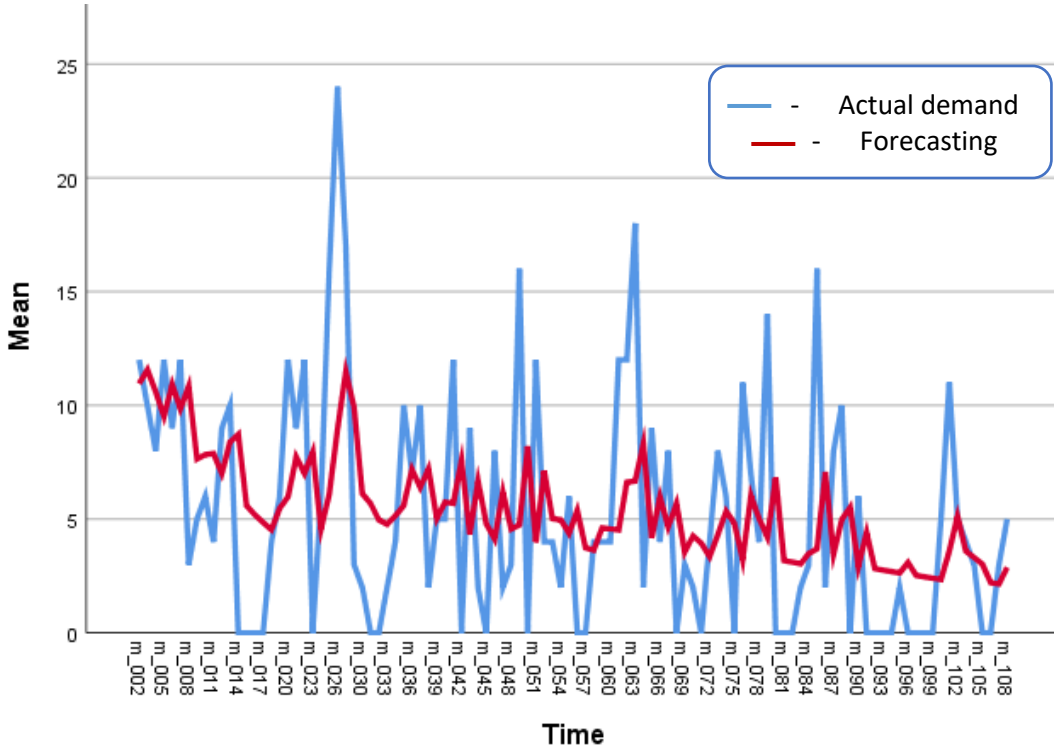

Figure 16. Demand Forecast Using ARIMA (1,1,1) Model.

Appendix 4. ARIMA (1,0,1)

Table 27. ARIMA (1,0,1) Model Statistics.

Model Statistics

Model	MAE
ARIMA (1,0,1)	3,940

Table 28. ARIMA (1,0,1) Model Parameters.

ARIMA Model Parameters

			Estimate	SE	Sig.
ARIMA (1,0,1)	Constant		5,342	,745	,000
	AR	Lag 1	,523	,266	,052
	MA	Lag 1	,249	,302	,412

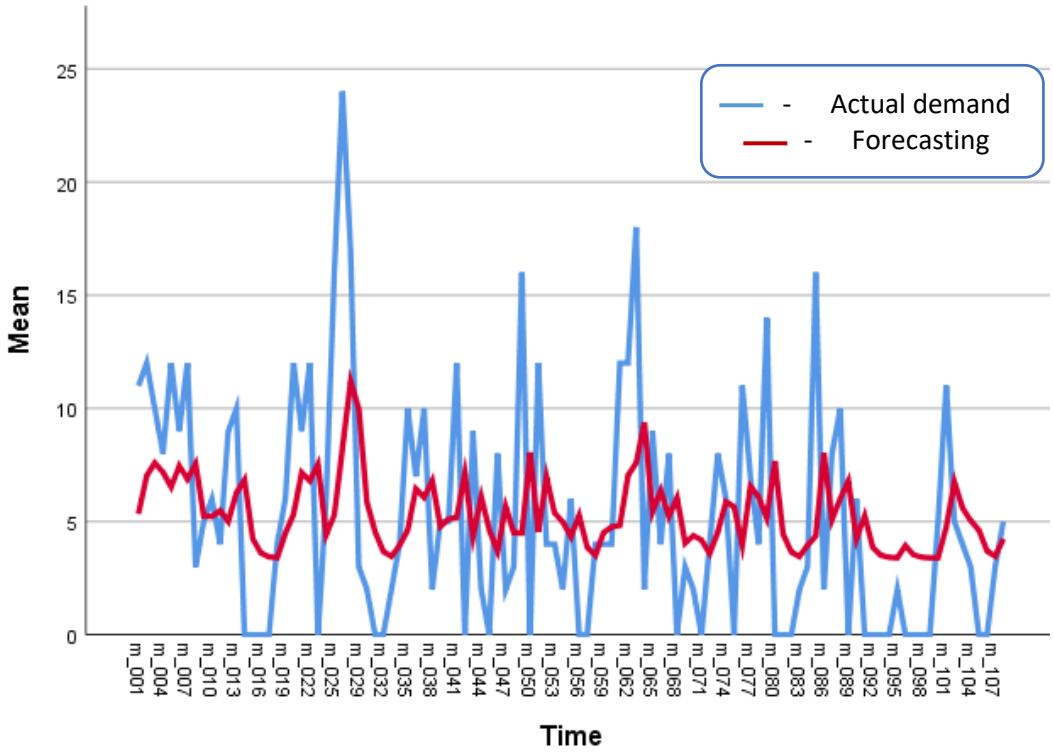

Figure 17. Demand Forecast Using ARIMA (1,0,1) Model.

Appendix 5. ARIMA (2,1,1)

Table 29. ARIMA (2,1,1) Model Statistics.

Model Statistics

Model	MAE
ARIMA (2,1,1)	3,773

Table 30. ARIMA (2,1,1) Model Parameters.

ARIMA Model Parameters

			Estimate	SE	Sig.
ARIMA (2,1,1)	Constant		-,048	,024	,044
	AR	Lag 1	,223	,108	,042
		Lag 2	,134	,106	,207
	Difference		1		
	MA	Lag 1	,999	1,795	,579

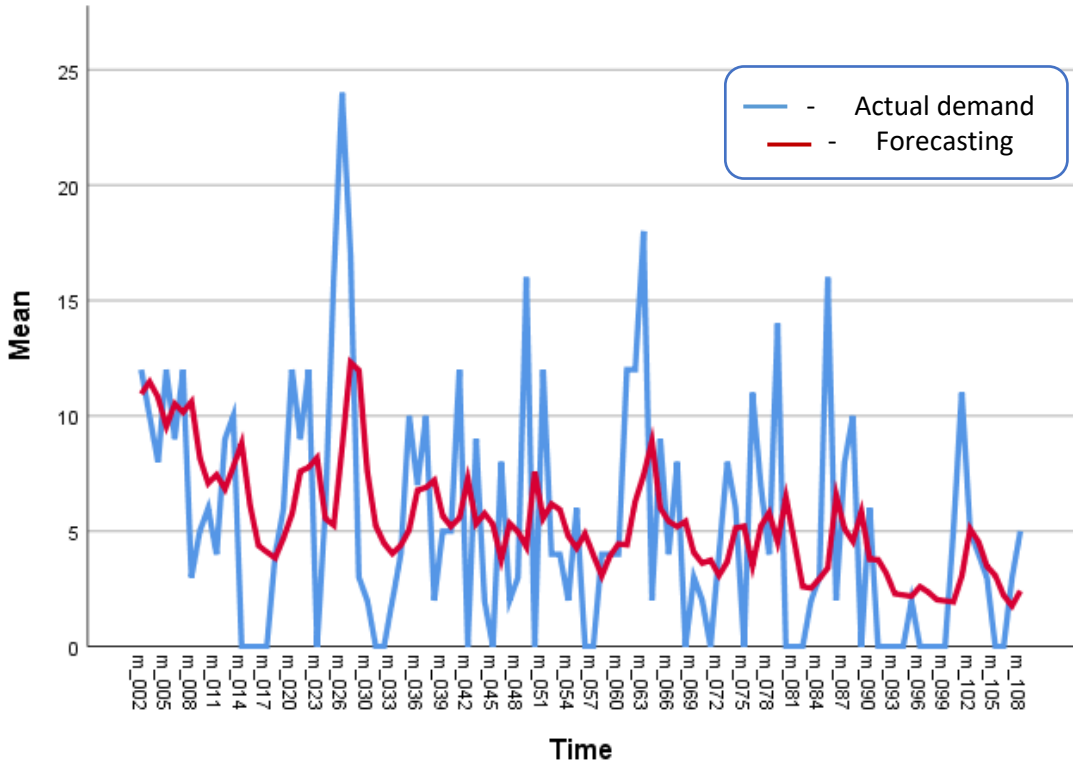

Figure 18. Demand Forecast Using ARIMA (2,1,1) Model.

Appendix 6. ARIMA (1,1,0)

Table 31. ARIMA (1,1,0) Model Statistics.

Model Statistics

Model	MAE
ARIMA (1,1,0)	4,026

Table 32. ARIMA (1,1,0) Model Parameters.

ARIMA Model Parameters

			Estimate	SE	Sig.
ARIMA (1,1,0)	Constant		-,066	,366	,857
	AR	Lag 1	-,456	,087	,000
	Difference		1		

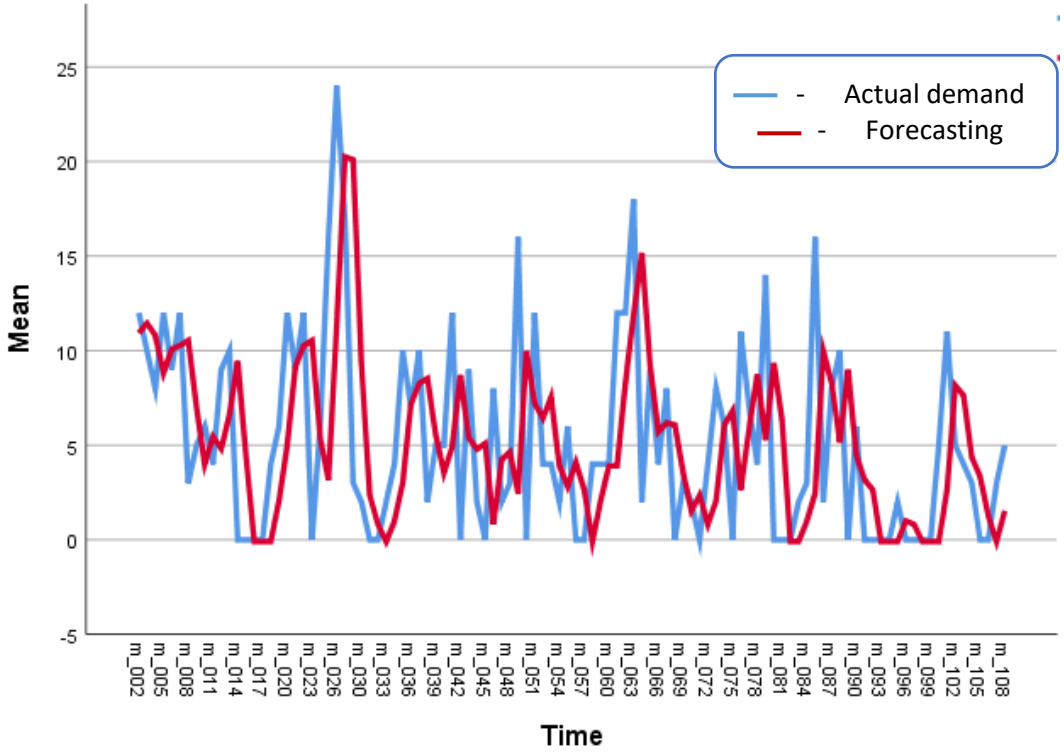

Figure 19. Demand Forecast Using ARIMA (1,1,0) Model.

Appendix 7. ARIMA (3,2,1)

Table 33. ARIMA (3,2,1) Model Statistics.

Model Statistics

Model	MAE
ARIMA (3,2,1)	4,164

Table 34. ARIMA (3,2,1) Model Parameters.

ARIMA Model Parameters

			Estimate	SE	Sig.
ARIMA (3,2,1)	Constant		,001	,011	,917
	AR	Lag 1	-,496	,110	,000
		Lag 2	-,131	,127	,302
		Lag 3	-,048	,108	,662
	Difference		2		
	MA	Lag 1	,993	,426	,022

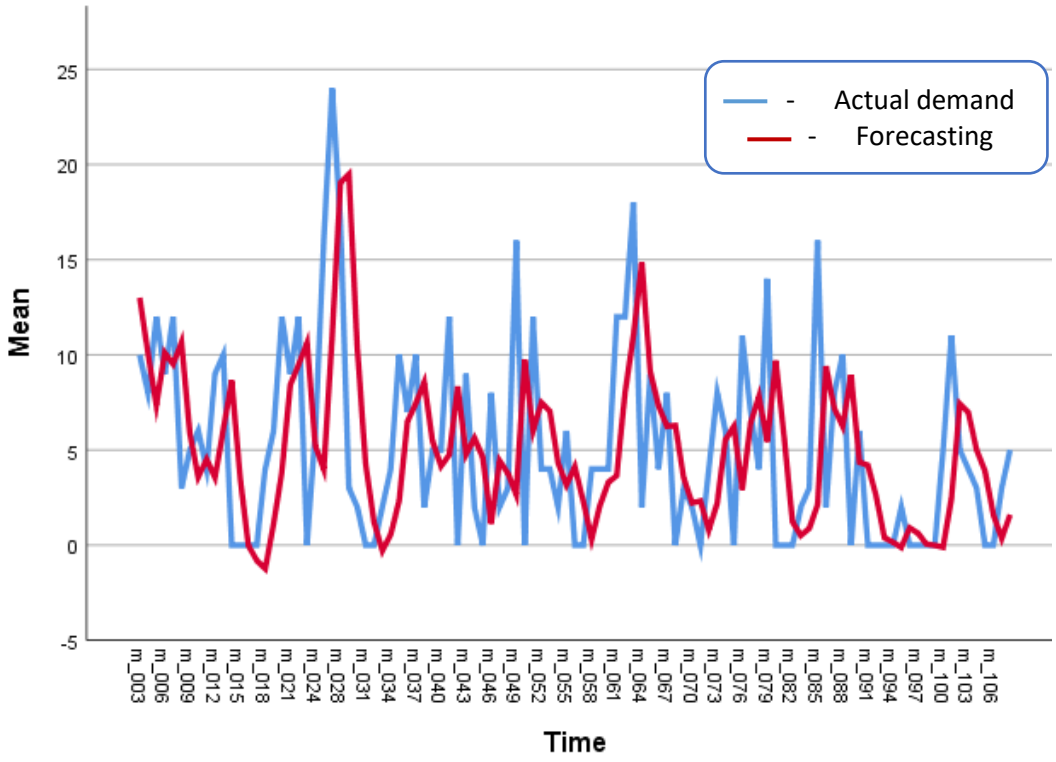

Figure 20. Demand Forecast Using ARIMA (3,2,1) Model.

Appendix 8. ARIMA (2,1,3)

Table 35. ARIMA (2,1,3) Model Statistics.

Model Statistics

Model	MAE
ARIMA (2,1,3)	3,653

Table 36. ARIMA (2,1,3) Model Parameters.

ARIMA Model Parameters

			Estimate	SE	Sig.
ARIMA (2,1,3)	Constant		-,042	,023	,068
	AR	Lag 1	,606	,223	,008
		Lag 2	-,475	,208	,024
	Difference		1		
	MA	Lag 1	1,433	26,125	,956
		Lag 2	-1,122	11,269	,921
		Lag 3	,689	17,982	,970

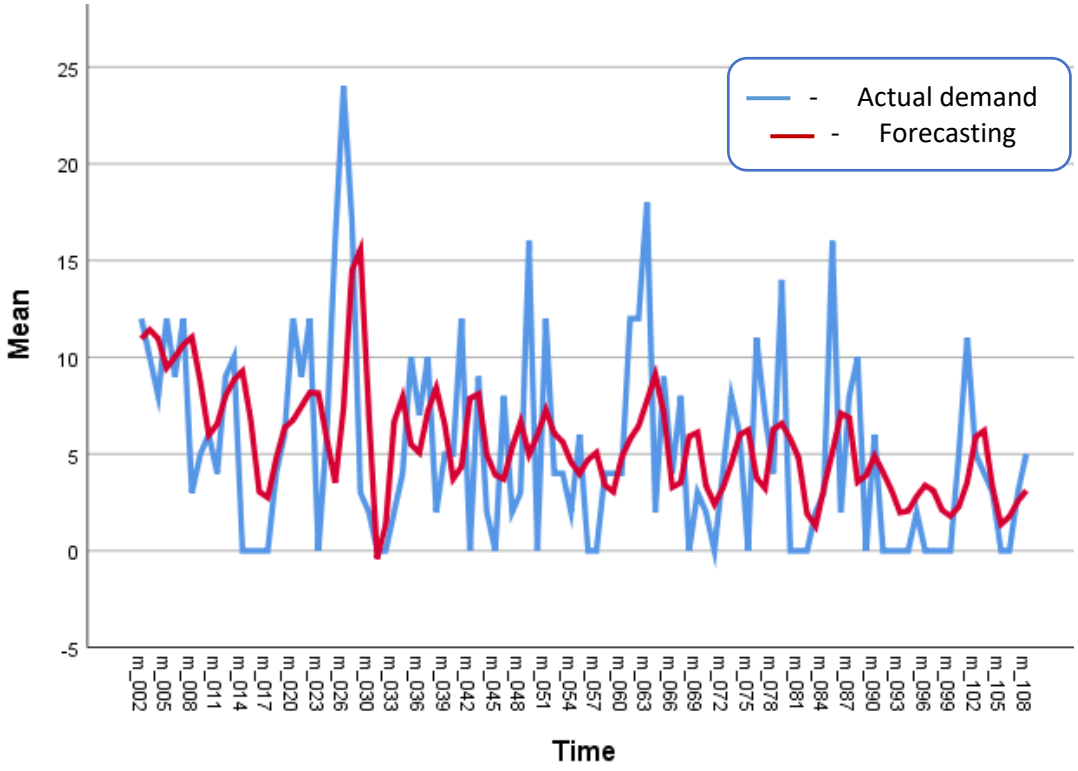

Figure 21. Demand Forecast Using ARIMA (2,1,3) Model.

Appendix 9. ARIMA (2,0,1)

Table 37. ARIMA (2,0,1) Model Statistics.

Model Statistics

Model	MAE
ARIMA (2,0,1)	3,888

Table 38. ARIMA (2,0,1) Model Parameters.

ARIMA Model Parameters

			Estimate	SE	Sig.
ARIMA (2,0,1)	Constant		5,349	,768	,000
	AR	Lag 1	,033	,557	,953
		Lag 2	,222	,169	,194
	MA	Lag 1	-,214	,570	,708

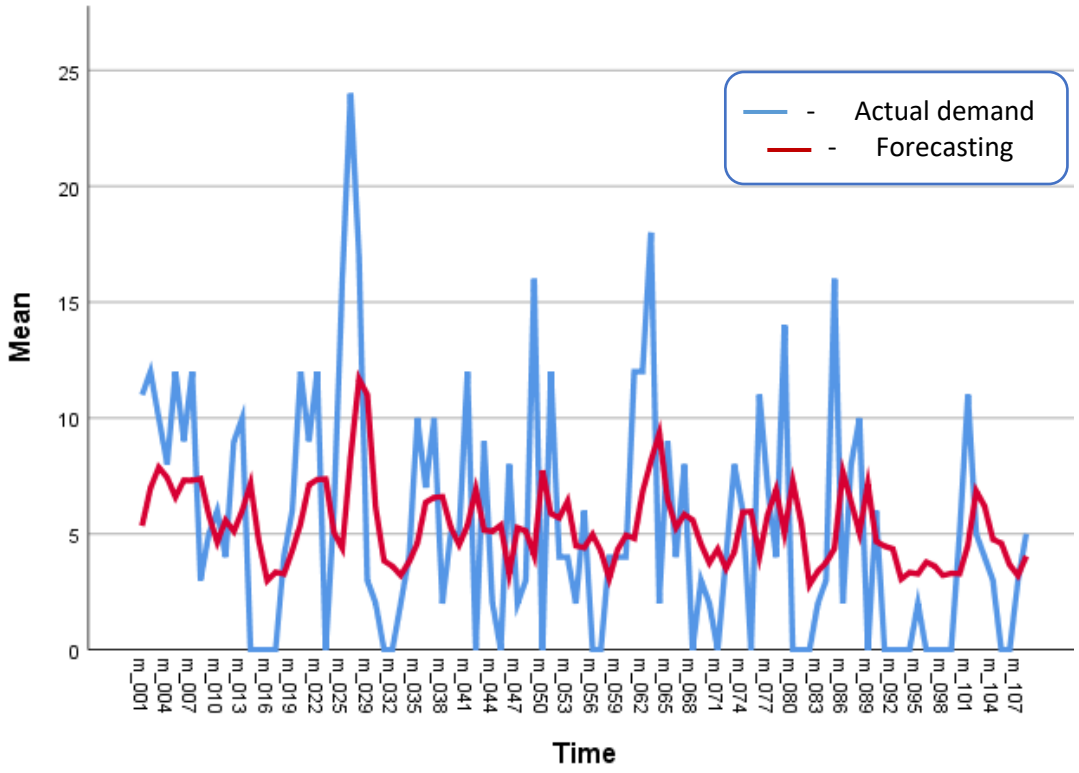

Figure 22. Demand Forecast Using ARIMA (2,0,1) Model.