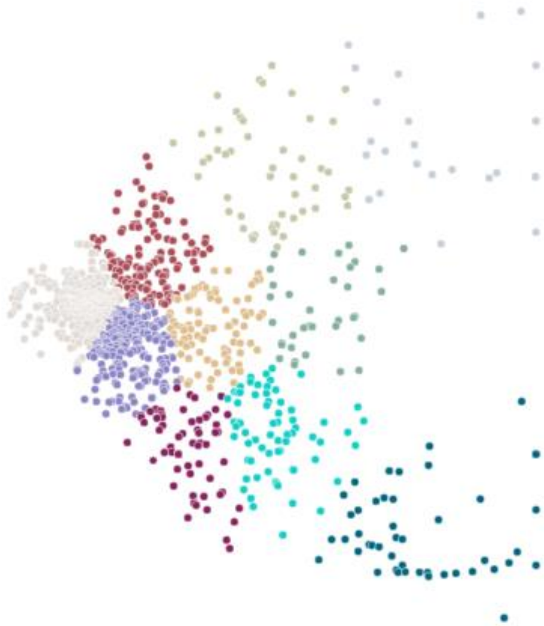
CHALMERS
UNIVERSITY OF TECHNOLOGY



Utilizing Cluster Analysis for Decision Support in Material Planning

Master’s thesis in Supply Chain Management

FABIAN FLAA
WILLIAM NORDGREN

DEPARTMENT OF TECHNOLOGY MANAGEMENT AND ECONOMICS
DIVISION OF SUPPLY AND OPERATIONS MANAGEMENT

CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2024
www.chalmers.se

Utilizing Cluster Analysis for Decision Support in Material Planning

FABIAN FLAA
WILLIAM NORDGREN

Department of Technology Management and Economics
Division of Supply and Operations Management
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2024

Utilizing Cluster Analysis for Decision Support in Material Planning

FABIAN FLAA
WILLIAM NORDGREN

Department of Technology Management and Economics
Chalmers University of Technology
SE-412 96 Gothenburg
Sweden
Telephone + 46 (0)31-772 1000

Cover:
Cluster analysis based on variation versus trend for items in finished goods at the case company.

Gothenburg, Sweden 2024

Utilizing Cluster Analysis for Decision Support in Material Planning

FABIAN FLAA
WILLIAM NORDGREN

Department of Technology Management and Economics
Chalmers University of Technology

SUMMARY
Currently, in the manufacturing industry, planning parameters are infrequently updated. This entails that they over time will become outdated and inaccurate, which in turn results in inefficient operations. To increase the frequency of parameter updates, reducing the amount of work required would be useful. A way of doing this is to group items that have similar demand patterns, or what this study calls item behavior. Grouping the items opens the possibility of handling them in bulk. Therefore, the purpose of this study was to identify the characteristics that make up the item behavior and use these in a cluster analysis to find suitable groups of items. These groups should then be used as decision-support in material planning. This study was carried out in collaboration with Meridion AB which is a consultancy company that specializes in improving organizations' supply chains.

The methodology used to conduct the study included quantitative and qualitative data collection. The main methodology used was cluster analysis, along with supporting methods like Z-score, min-max scaling, and the DeD method. In addition, several other methods were used to conduct the analysis, including an interview, discussions, and a workshop.

The study resulted in five item behavior characteristics that were identified and then combined with each other as well as other parameters to group items based on their behavior. Three cluster analyses were then conducted and validated. The first was regarding safety stock levels, the second was used for ABC-XYZ classification, and the third was used to identify items where production quantities deviated from EOQ. These serve as examples of how cluster analysis can be used to increase the efficiency regarding the process of performing parameter updates, and how it can be applied to material planning issues. All three of these examples successfully provided decision support in material planning, although in different ways. In addition, there is also some more general knowledge regarding the use of cluster analysis for this purpose included that was gathered during the course of the study.

Keywords: Cluster analysis, Item behavior, Material planning, Planning parameters, Safety stock, ABC-XYZ analysis, EOQ

ACKNOWLEDGEMENTS

We would like to extend our sincere gratitude to the following individuals and groups who have been instrumental throughout the work with our thesis:

Firstly, we would like to thank our supervisor and examiner at Chalmers, Patrik Jonsson, for providing continuous support and knowledge throughout this thesis. His academic guidance and expertise has been highly valuable to us.

We would also like to thank our supervisors at Meridion, Johan Bystedt and Henrik Kjellgren, who made this thesis possible thanks to their industry knowledge and guidance. Additionally, we would like to thank Elias Ahlman for his support in the development of the Qlik applications, as this support was crucial in conducting the project successfully.

We are grateful for the respondents at the case company who were generous enough to share their time, insights, and data with us, these contributions were instrumental in shaping this thesis, and it would not have been possible to carry it out without them.

Finally, a sincere thanks to all the colleagues at Meridion who have been very welcoming and interested in our work, their company has had a key role in keeping us motivated throughout the spring.

Fabian Flaa & William Nordgren, Gothenburg, June 2024

LIST OF TERMS

BI - Business Intelligence

Cluster analysis - Grouping a set of objects into different groups, called clusters

Coefficient of variation - A measure of the dispersion of data points around the mean

DeD - Depth Difference method, a method for determining the number of clusters

EOQ - Economic Order Quantity, theoretically optimal quantity an organization should use to keep costs low

ERP - Enterprise Resource Planning, system for integrated management of business processes

Industry 4.0 - The fourth industrial revolution by integration of digital technologies

Item behavior - How the demand or stock level of an item moves over time

Item behavior characteristics - The underlying factors that create the item behavior

k-means - A clustering algorithm

min-max scaling - A normalization method

Planning parameter - Variables used to guide and control the material planning process, e.g. EOQ, safety stock

Z-score - A normalization method

Contents

1 Introduction

In the introduction section, the background surrounding the study is presented and its relevancy is explained. Then, the problem discussion is introduced. The section further establishes the purpose of this research, along with the chosen research questions. Finally, relevant delimitations are made.

1.1 Background

Manufacturing companies use different planning methods for a variety of components in the planning system, there are methods that help with inventory management, determining order sizes, batch sizes, safety stock levels, and much more. These methods often result in planning parameters that will be stored in the companies' enterprise resource planning system (ERP) to simplify the operational planning process. These parameters need to be updated regularly as they are based on certain key performance indicators (KPI) from when the calculations were performed, for example, the demand for an item fluctuates up and down with time so the parameter associated may become inaccurate (Jacobs et al., 2018). Such inaccuracies lead to inefficient operations, which will mean higher costs for the company.

A recent study by Bystedt and Jonsson (2023) found that, in practice, parameters are not updated as often as they should be, for reasons that are not completely known, but a contributing factor is the amount of labor it takes to update the parameters for each stock keeping unit (SKU) an organization has. If it was possible to reduce the amount of work associated with the process of updating these parameters, then companies would be able to update them more frequently, leading to improvements in terms of tied-up capital, service level, and operational costs being achieved, depending on which is prioritized by the company (Jonsson & Mattsson, 2009).

Additionally, the concept of Industry 4.0 has been a hot topic in recent years and promises substantial improvements in operational efficiency through improved analysis and usage of data (Hermann et al., 2015). Within this concept, advanced data analysis methods such as cluster analysis are often cited as decision support in articles investigating how to reach these improvements in operational efficiency (Bordeleau et al., 2018).

The starting point for the study is item behavior, which refers to how the demand of an item moves up and down over time. Item behavior is the basis for many planning methods and their resulting parameters, and one often finds different items that have similar item behavior. This means that when working with large data sets, identifying which items behave similarly and grouping those together may allow treating those in the same way, reducing the amount of work needed to perform parameter updates. However, in order to do so one first needs to identify the factors, or characteristics, that make up the item behavior, as those would be the factors determining what items have similar behavior.

This study was carried out in collaboration with Meridion AB, or Meridion as it will be referred to in this study. Meridion is a consultancy that specializes in supporting companies with their supply chains through the ERP system Infor M3, and the business intelligence (BI) system Qlik (Meridion, n.d.-b). With Infor M3, Meridion has a lot of experience from prior implementation projects and can help customers with most areas of the ERP (Meridion, n.d.-a). Through the BI platform Qlik, Meridion can visualize their customers' business data to deliver useful insights but also make more creative, powerful visualizations due to Qlik actively working to be a leader in data analytics, integration, and quality along with a focus on integrating advanced analytics tools such as certain artificial intelligence and machine learning algorithms (QlikTech, n.d.). Additionally, the study includes another company that will be referred to as the case company to keep them anonymous. The case company provided the underlying data that the analysis in this study stems from.

1.2 Problem Discussion

Update of planning parameters in many companies are carried out infrequently, in most cases only once a year or less frequently (Jonsson & Mattsson, 2014). Some examples of planning parameters are the level of safety stock, order quantities, and ABC classification. In a more recent study, Bystedt and Jonsson (2023) confirmed these results and concluded that there is a need to find a solution to how to get companies to update these parameters more frequently. The reason why it needs to be solved is that outdated parameters do not reflect the current situation in the company and will most likely lead to sub-optimization. If the parameters set are not suited for the strategy the company wants to pursue it will entail complications of finding a balance between service level, operational cost, and tied-up capital. So, to increase the performance of a company in terms of these three, the parameters used must be more up-to-date. To do this is possible with companies' current tools but is time-consuming and hence one of the reasons it is not done today. Therefore, a way of making updates more efficient would most likely result in more frequently updated parameters that in turn result in better results for companies.

To make the update of parameters more efficient a solution could be instead of analyzing how a certain parameter should be decided for each individual item, the method or strategy for determining it can be done for groups of items at a time. A way of doing this could be by handling all items in a product group the same way, however, it may be so that their item behavior is not similar. Hence, to find similarities between items' different behavior and be able to treat them the same way, cluster analysis could be utilized, where items can be grouped in different clusters based on their behavior. To prove the similarities within a cluster there is a need to validate them. This could be done by applying different methods, or by visually examining them. To help with this, there exists different BI systems that can visualize the data and thus provide a visualization of the actual similarities and differences between items. For each of these clusters, a single decision can be made that will be based on how each cluster looks like. Thus, the need to analyze each item on the item level could be reduced, and it could instead be done at an aggregated level, leading to more up-to-date planning parameters that can improve performance in terms of service level, operational cost, or tied-up capital.

1.3 Purpose

The purpose of this study is to identify and utilize different item behavior characteristics in a manufacturing company through cluster analysis. The cluster analysis should provide decision support for operational decision-making in material planning.

1.4 Research Questions

Based on the background and the problem discussion presented, three research questions have been formed as follows:

1. What different kinds of item behavior characteristics exist in material planning?
2. How can these characteristics be used in a cluster analysis to identify and group items based on their behavior?
3. What kind of decision support can such an analysis provide in material planning?

1.5 Delimitations

The time available to conduct this study was limited, and therefore some delimitations needed to be made to make sure that the main goal of the study was fulfilled. This study limits itself to only using data from one of the case company's facilities due both to the time restriction and the experimental nature of the study. Additionally, this study focuses on practical applications of cluster analysis that can be of use to supply chain professionals. This means that it is not relevant to go in-depth on the underlying functionality of certain cluster methods and their supporting methods. When needed, these methods' functionality is instead explained in general terms to be more understandable. This delimitation also means that the conducted cluster analyses are intended to be understandable for persons who have not worked with cluster analysis before. Thus, clustering will only be done in 2 dimensions, as that enables them to be visualized with ease.

2 Theoretical Framework

In this section, a theoretical framework that covers the different topics that are relevant to the study is presented. Firstly, the supply chain management related concepts and methods relevant to this study are explained. This is followed by theory on cluster analysis and associated methods. Finally, theory on the use of IT systems for decision-making is presented.

2.1 Supply Chain Management

Supply chain management is commonly defined as "the management of the flow of goods and the inverse flow of information and cash" (DeSmet, 2017, p. 9). The author further explains that the flow of goods refers to the products that travel from production to stores and consumers. The start of the production of goods is usually initiated by information that travels upstream in the supply chain, telling that there have been sales and that there is a need for additional products. In the definition, cash is also mentioned, which is what the goods are exchanged for (DeSmet, 2017).

2.1.1 Supply Chain Triangle

The supply chain triangle, shown in Figure 1, displays how service, cash, and cost are all related within supply chain management. DeSmet (2021) explains this as organizations can deliver different types of services to their customers that depending on which type of service is delivered is associated with a certain cost. To deliver this type of service it is also required to have a particular amount of inventory, which is referred to as cash in this context. Jonsson and Mattsson (2009) discusses the same concept with performance variables that influence the company's revenues, costs, and assets respectively.

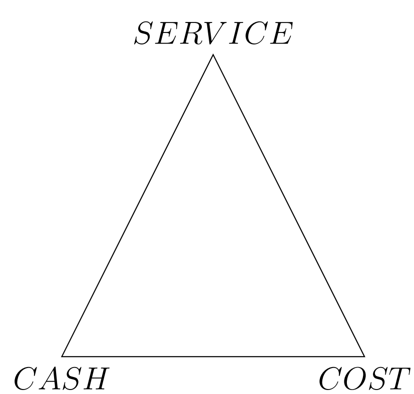


Figure 1: Visualization of the supply chain triangle, adopted from DeSmet (2021).

The service corner in the triangle relates to what service is delivered to the customers and the most common metric to use for this is service level which with basis taken in an agreed target measures the percentage of orders being delivered (DeSmet, 2021). How this should be measured exactly can be done in multiple ways and there exists multiple definitions of service level. For example should the term "on time" be defined as when the order is delivered or when it is shipped? In addition to service level, other things are covered in the service corner, however, those are usually ignored since service level is dominant (DeSmet, 2021). An example of what more is included in service is the order flexibility, which refers to how flexible the supplying company is towards its customers regards the timing of ordering, order quantity, and changes to already placed orders. The more flexible the company is, the higher the customer will value the service since it enables them to be agile and reduces the need of carrying large inventories.

As previously mentioned, to be able to deliver service there have to be some costs. Jonsson and Mattsson (2009) divides logistics costs into two parts, costs associated with material flows and costs associated with production. Costs associated with material flows are for example handling, packaging, and storage costs. On the other hand, costs associated with production are generally less tangible. Some examples from Jonsson and Mattsson (2009) include costs of changing the rate of production and set-up costs, where the former relates to costs accrued when temporarily increasing the production capacity by, for example, increasing overtime. Set-up costs refer to the cost of downtime when changing between manufacturing different products. Altogether, logistics costs commonly equal to around 20-30% of a company's sales (Jonsson & Mattsson, 2009).

The last factor in the triangle, cash, is important to cover the expenses that arise in a company and not to risk cash shortages with the risk of disruptions in the supply chain. DeSmet (2021) uses working capital as a synonym for cash in this context and defines it as the inventory plus accounts receivable minus accounts payable. Jonsson and Mattsson (2009) describes that in current assets, inventories and accounts receivable is the two dominant items. Thus it can be seen as the corresponding term to working capital in this example. There are different ways of measuring the performance of tied-up capital and the simplest way of doing it is to measure the monetary value of inventory (Jonsson & Mattsson, 2009).

The balance between these three corners in the triangle is not an easy task to solve. Companies strive to perform well in all of them and this is where the complexity comes into play. To offer a good service level is easier if the stock level is high to not risk any shortages, but that comes with extra cash required. The same logic could be applied to if one wants to perform well in cash then the inventory should be low. But keeping the inventory low could imply that the production can not manufacture products in the most feasible way from a cost perspective, it could for example be a mismatch between optimal batch size and available material in inventory. In conclusion, there is always a trade-off that has to be made.

2.1.2 Material Planning

Material planning is to a large extent about identifying items for new orders, appropriate quantities for these, when the order should be delivered, and when the order needs to be initiated (Jonsson & Mattsson, 2009). To enhance decision-making efficiency related to this, planning methods with tailored planning parameter settings that are more or less suitable depending on the planning environment are employed.

Planning Methods

Deciding how the planning of material will be performed is an important choice in a manufacturing setting. There are a number of material planning methods that have been developed to be used for this purpose. Jonsson and Mattsson (2014) mentions five different that are used at Swedish companies, namely Re-order point systems (ROP), Run-out time planning, Periodic order system, Material requirements planning (MRP), and Kanban where ROP and MRP are the most common. ROP is according to Jonsson and Mattsson (2009) a method that is based on a comparison between the current stock and a reference quantity, which is called the re-order point. A new order to replenish stocks will be initiated when the stock falls below the re-order point, see Figure 2. The quantity of the re-order point is determined by the demand during the lead time of the replenishment as well as safety stock to handle variation in demand (Jonsson & Mattsson, 2009).

While ROP takes in how much is in stock and makes the decision if a new order will be generated, MRP instead looks at when the stock balance becomes negative in order to plan for new orders. When the balance becomes negative a new order must be scheduled for delivery, and since lead time usually exists the order must be placed earlier than that time period in order to receive it when wanted. Since the order should arrive when the stock becomes negative the order to be placed will be that time minus the lead time. The process of MRP is visualized in Figure 3. How often the stock level should be checked can vary, and can for example be between each stock transaction, once a week, or once a month (Jonsson & Mattsson, 2009).

No single material planning method is, in general, better than another, instead they are better suited for different environments (Newman & Sridharan, 1995). The authors

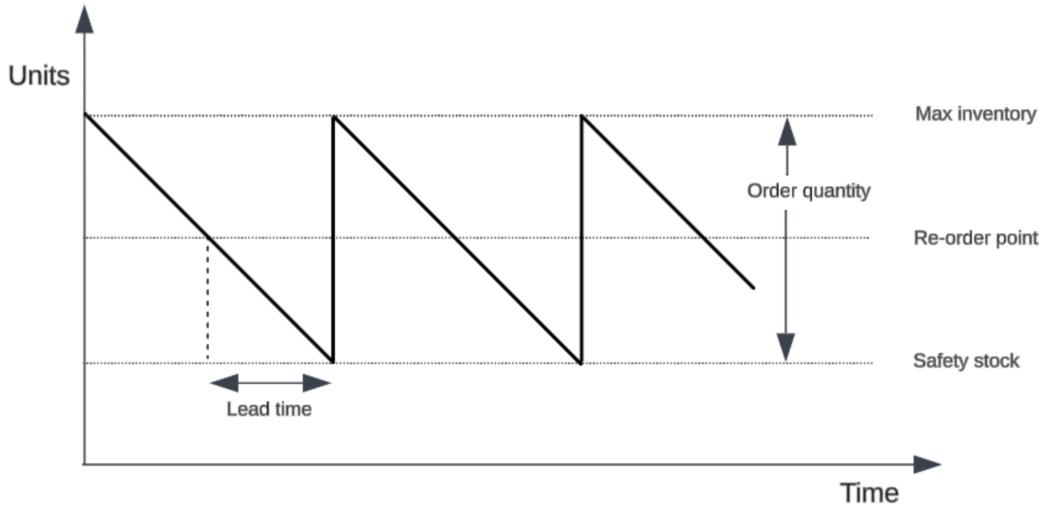


Figure 2: Visualisation of how ROP works.

Week		1	2	3	4	5	6	7
Forecast/requirement		10	10	10	10	10	10	10
Stock on hand	20	10	40	30	20	10	0	-10
Planned order delivery			40					
Planned order start					40			

Figure 3: Visualization of how MRP works, from Jonsson and Mattsson (2009).

found that when the demand is unstable then MRP is a good choice since it can handle complex processes such as volume fluctuations and demand variability. If the environment instead is predictable, ROP could be used and achieve feasible results while still being simple to use (Newman & Sridharan, 1995). Jonsson and Mattsson (2014) explains that each material planning method is only defined as how new orders should be initiated, and how well a method actually performs in the context will rely on how the different planning parameters are determined.

Planning Parameters

As mentioned above, how the planning parameters are determined plays a big role in the performance of the chosen planning method. There exists multiple parameters that can be used for the different methods, but two common are order quantity and safety stock (Jonsson & Mattsson, 2014). Order quantities can be determined in different ways and common ones are economic order quantity (EOQ), lot for lot, and estimated run-out time, where EOQ is the most popular (Jonsson & Mattsson, 2009). EOQ is based on the trade-off between ordering cost and inventory cost and aims at minimizing the total of these costs (Schwarz, 2008). The calculation of it is presented in Equation 1.

$$EOQ = \sqrt{\frac{2 * D * S}{I * C}} \tag{1}$$

where:
D = demand per time unit
S = ordering cost per ordering occasion
I = inventory carrying costs per time unit
C = item value per stock unit

However, this method makes the assumption that the demand rate is constant (Jacobs et al., 2018) which may not correspond to the actual demand. Lot for lot on the other hand is a method that is easy to use and the logic behind it is that the order quantity is the same as the demand in that period (Jonsson & Mattsson, 2009). The authors further state that the implications of this are that tied-up capital and ordering costs are not taken into consideration, that the order quantities can be really small, and there are many orders. Estimated run-out time is based on choosing an order quantity that will cover a determined number of planning periods (Jonsson & Mattsson, 2009). The planning period could for example be a day or a week. The number of periods is chosen based on experience and the order quantity will be recalculated at every ordering occasion. It takes into consideration factors such as estimated demand and price, among others, which forms an assumption of what the best order quantity would be.

Safety stock is used to protect against demand fluctuations, inaccurate forecasts, and uncertain lead times (Jacobs et al., 2018). The logic behind it is to have additional items in stock that are intended to be used if these issues appear. How safety stocks are expressed differs since it can be in either a quantity or a number of days coverage of demand (Jonsson & Mattsson, 2009). By expressing it as time it brings the advantage of coping with demand changes without the need to update the system. Safety stocks can be chosen based on the balance between shortage cost and the cost of tied-up capital, with the same logic as for economic order quantity that aims at minimizing the total cost. For determining the size of safety stock there are two main types of methods according to Jonsson and Mattsson (2009) and they are by manual estimates and simple calculations or by more advanced calculations and information regarding the level of uncertainty. The authors further state that to be able to use the advanced methods requires information about the demand variation during lead time. However, the manual methods come with the benefit that they are easier to use and do not require any software support (Jonsson & Mattsson, 2009). Thus, they may be of interest to use in smaller organizations or when the demand is fairly even. Jonsson and Mattsson (2009) presents one common formula for calculating the safety stock through an advanced method based on demand fill rate (SERV2), which is shown in Equation 2.

$$Safety\ stock = Z * \sigma_{DDLT} \tag{2}$$

where:
Z = safety factor
σ_{DDLT} - standard deviation of demand during lead time

To calculate this, one must first obtain the standard deviation of demand during lead time. This is done with Equation 3. This is then used to determine the service loss function as in Equation 4, which also requires the desired fill rate and the order quantity used as input. Once the service loss function is obtained, the safety factor can be obtained from a safety stock table that matches the value of the service loss function to a value for the safety factor.

$$\sigma_{DDLT} = \sqrt{LT * {\sigma_D}^2 + {\sigma_{LT}}^2 * D^2} \tag{3}$$

where:
LT = average lead time in periods from order to delivery
D - average demand per period
σ_D - standard deviation of demand per period
σ_{LT} - standard deviation of lead time

$$E(z) = \frac{(1 - Fillrate) * Q}{\sigma_{DDLT}} \tag{4}$$

where:
$E(z)$ = service loss function
$Fillrate$ = Share of demand that can be delivered directly from stock
Q = Order size (EOQ)
σ_{DDLT} - standard deviation of demand during lead time

Planning Environment

There are several different ways companies' planning environments can be designed in and Jonsson and Mattsson (2003) defines four main groups that the planning environment of a company can belong to which are based on different levels of product variety and scale of production. These four are complex customer products, configure-to-order products, batch production of standardized products, and repetitive mass production. Rahmani et al. (2022) states that the planning environment can be described by various factors related to the product, market, and production process. Some of these factors are product volume, length of lead time, variability of lead time, and demand variability (Rahmani et al., 2022).

The reason why the planning environment is of interest is due to what Jonsson and Mattsson (2014) states that it affects how competitive the company's delivery capability is as well as the tied-up capital and transaction costs. The authors furthermore mention that the planning environment affects how well the chosen planning methods and planning parameters will work in the specific context of the company. That choosing a suitable method for the company's planning environment is important is also emphasized by Jonsson and Mattsson (2003). This serves to illustrate the importance of adapting the choice of planning methods to the specific planning environment in a given case.

2.1.3 ABC-XYZ Analysis

ABC analysis is a method used for dividing items, customers, suppliers, and more into different classes depending on a criterion, which for example could be volume value per item (Jonsson & Mattsson, 2009). The method is an application of the 80/20 rule, commonly known as the Pareto principle, as it classifies around 20% of the objects to be classified into the A class, which in the example using volume value would mean the top 20% of items in terms of volume value, that contributes to roughly 80% of the volume value, would be A classed (Flores & Whybark, 1986). For the B- and C-class it is generally not specified where the limits should be set, this will instead be quite case-specific. The classification is meant to signify which items are more important for management to pay extra attention to, and sometimes, one criterion may not be enough to describe the managerial needs of the items. In such cases, one might want to consider more than one criteria (Flores & Whybark, 1986). In those cases, a matrix is created which has one criterion per axis, for a total of two. One application of this concept is known as XYZ analysis and can be used as an extension of ABC analysis. The ABC analysis will be used as a criterion on one axis and XYZ analysis will be used as the criterion on the other axis. The XYZ analysis classifies items based on their variation in demand and forecast accuracy, as that is often a consequence of the demand variation (Nowotynska, 2013).

Together, they form what we call ABC-XYZ analysis, an example of which can be seen in Table 1. The idea is that based on the characteristics of each class, different strategies and levels of attention can be applied to the different classes, and using two

Table 1: Table of ABC-XYZ classifications and their characteristics, from ASCM (2020).

	X	Y	Z
A	**AX Class** • High volume value • Even demand • Reliable forecasts	**AY Class** • High volume value • Predictably variable demand • Less reliable forecasts	**AZ Class** • High volume value • Sporadic, variable demand • Forecasting unreliable or impossible
B	**BX Class** • Medium volume value • Even demand • Reliable forecasts	**BY Class** • Medium volume value • Predictably variable demand • Less reliable forecasts	**BZ Class** • Medium volume value • Sporadic, variable demand • Forecasting unreliable or impossible
C	**CX Class** • Low volume value • Even demand • Reliable forecasts	**CY Class** • Low volume value • Predictably variable demand • Less reliable forecasts	**CZ Class** • Low volume value • Sporadic, variable demand • Forecasting unreliable or impossible

criteria makes for a more accurate division than only one would. The value on which XYZ classification is performed is the coefficient of variation (χ). In order to calculate it, the arithmetic mean first needs to be calculated as in Equation 5:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{N} \tag{5}$$

where:
$\bar{x}$ - arithmetic mean
x_i - characteristics of the feature at time i, for i=1,2,...,n
N - population size

The arithmetic mean is then used to calculate the standard deviation (σ) over the given number of time periods as can be seen in Equation 6:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{N}} \tag{6}$$

where:
σ - standard deviation
x_i - characteristics of the feature at time i, for i=1,2,...,n
$\bar{x}$ - arithmetic mean
N - population size

The coefficient of variation (χ) is then calculated by dividing the standard deviation by the arithmetic mean in order to get a value that is decoupled from the size of the sample, as can be seen in Equation 7.

$$\chi = \frac{\sigma}{\bar{x}} \tag{7}$$

where:
χ - coefficient of variation
$\bar{x}$ - arithmetic mean
σ - standard deviation

The coefficient of variation is used to determine which classification an item in an XYZ analysis should have. There exist different guidelines for how these limits between the X, Y, respective Z class should be chosen. One way of doing it is to put the items with a coefficient of variation less than 10% in class X, the ones between 10% and 25% in class Y, and those that are larger than 25% in class Z (Stojanović & Regodić, 2017; Trubchenko et al., 2020). However, when applying XYZ analysis to a real case, Stojanović and Regodić (2017) sets the limits to 50% between X and Y and 100% between Y and Z.

Depending on which class an item ends up in, different inventory management strategies should be employed, and there is sometimes no clear answer to what the best strategy is for a given class. Figure 2 shows the generally applicable strategies and highlights those that are disputed between different sources (ASCM, 2020; Stojanović & Regodić, 2017). Firstly, the AX items are high volume value and consistent demand, meaning forecast accuracy is good and therefore stock levels can be kept low without losing service level. AY items are mostly the same as AX but as the consumption varies more, more attention and care to the planning processes should be devoted. AZ items are complicated as they contribute significant value but are very hard to forecast. Stojanović and Regodić (2017) recommends medium stock levels while ASCM (2020) argues for not stocking these items and instead making them to order. Thus, depending on the specific context and planning environment of the case company, different strategies may be suitable for this class. This also goes for BY and BZ items, as they have largely similar characteristics and are therefore treated in similar ways. BX and CX items have continuous, forecastable demand but make up a smaller share of the volume value, they should have low stock levels while maintaining the desired service levels. Finally, CY and CZ items are infrequently demanded and have negligible impacts on the tide-up capital because of their low volume value. It therefore means not much capital would be tied up by having these items at relatively high stock levels, and that should therefore be preferred in order to maintain the desired service level since forecasting demand is barely possible with these items (Stojanović & Regodić, 2017).

Table 2: Inventory management principles depending on ABCXYZ class. Adapted from Stojanović and Regodić (2017). *Disputed principle

	X	Y	Z
A	**Low inventory**	**Low inventory**	**Medium inventory***
B	**Low inventory**	**Medium inventory***	**Medium inventory***
C	**Low inventory**	**High inventory**	**High inventory**

2.2 Cluster Analysis

Cluster analysis aims to distinguish different groups, called clusters, within a set of data. The clusters should be separated subsets of the data set so that data that belongs to different clusters are noticeably less similar than data belonging to the same clusters (Wierzchoń & Kłopotek, 2018). The authors further explain that therefore, the role of cluster analysis is "to uncover a certain kind of natural structure in the data set" (Wierzchoń & Kłopotek, 2018, p. 9). This leads to two useful qualities of cluster analysis, namely visualization and handling of large datasets. Cluster analysis can provide important cognitive insights through visualization, and as it characterizes certain groups of data together, clusters can often be handled uniformly which allows for large data sets to be analyzed, according to Wierzchoń and Kłopotek (2018).

2.2.1 Feature Selection and Extraction

A feature in this context is a variable in the data set on which cluster analysis is to be performed, meaning that the features will determine which observations are considered similar enough to be in the same cluster. Firstly, feature selection is an important step in the clustering analysis which tends to be performed early in the process. Good feature selection can improve the quality of the resulting clusters, and bad feature selection can have the opposite effect (Aggarwal, 2014). The importance of feature selection has increased in the age of big data as there are huge amounts of high-dimensional data in a wide variety of domains, which leads to the existence of many features (Li et al., 2017). Having many features tends to make the model overfitted, meaning it doesn't generalize well when applied to other data sets than the training data set (Ying, 2019).

Another important aspect is that when it comes to real applications such as this study, the selection of features for clustering should be largely based on the different variables' meanings as they have a substantial impact on what the resulting clusters actually mean (Hennig, 2015). Feature selection entails filtering out features that are irrelevant or redundant (Li et al., 2017). An irrelevant feature is one that is not able to separate two clusters as the data points are all over the place, while a redundant feature is one that can separate two clusters but is strongly correlated with an already selected feature, thereby making it redundant (Li et al., 2017).

Secondly, there is feature extraction, which means to linearly or non-linearly combine two or more features into one in order to reduce the dimensionality, meaning to reduce the number of dimensions (Li et al., 2017). Aside from avoiding overfitting, reducing the dimensionality leads to generally better performance in terms of learning accuracy, computational cost, and model readability (Alelyani et al., 2014).

2.2.2 Normalization Methods

Standardization, also called normalization (Wierzchoń & Kłopotek, 2018), refers to the recasting of the selected features into dimensionless units (Romesburg, 2004). Its purpose is to cast the features into a common scale to ensure that features with larger magnitudes have the same influence as features with smaller magnitudes (C3.ai, n.d.). As such, normalization ensures all features carry the same weight when the clustering algorithm determines the similarities between objects (Romesburg, 2004). Two of the most common normalization methods are Z-score and min-max scaling.

Z-score is a normalization method that scores a measurement based on how many standard deviations above or below the mean of the population (Chubb & Simpson, 2012). This means that if there are outliers in the population, then these will not significantly impact the Z-score of the other measurements, making the method suitable when there are outliers that you do not want to remove from the data set. As an example, Figure 4 shows the relationship between Z-scores and percentiles when the population is distributed according to a normal distribution. The Z-score is calculated by dividing the difference between the observed measurement and the arithmetic mean by the standard deviation of the population, shown in Equation 8.

$$Z = \frac{x - \bar{x}}{\sigma} \tag{8}$$

where:
Z - Z-score of observed measurement
σ - standard deviation
$\bar{x}$ - arithmetic mean
x - observed measurement

On the other hand, min-max scaling uses the minimum and maximum measurements of the population to scale it to the interval [0,1] and is the simplest normalization method available (Sinsomboonthong, 2022). This method works well when the maximum and minimum values of the population are not outliers, as then the population gets scaled to

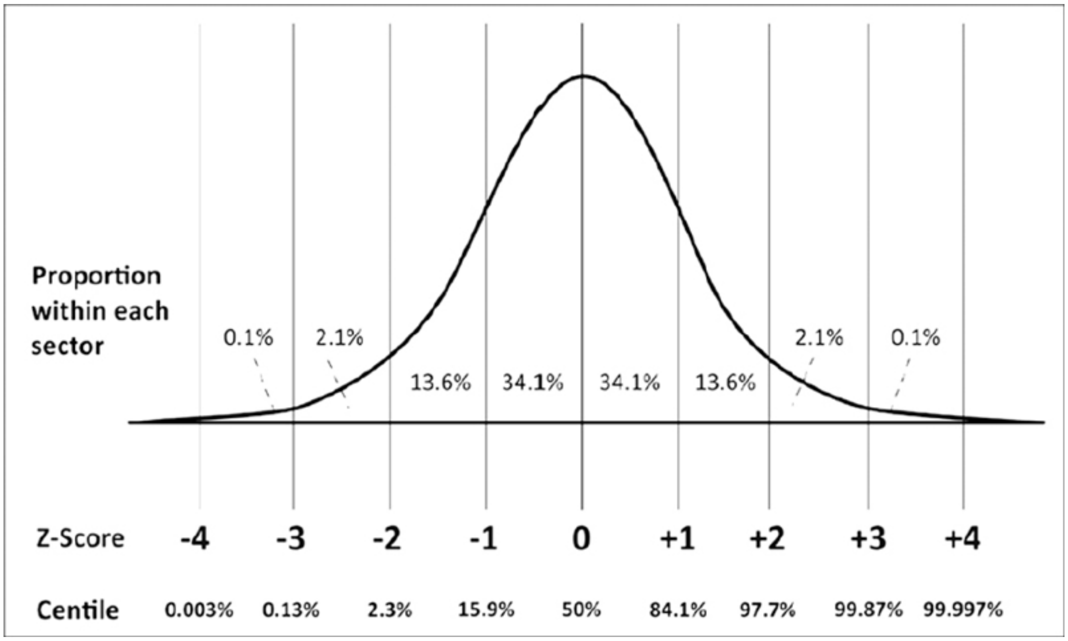


Figure 4: Relationship between Z-scores and centiles when population assumes a normal distribution, from Chubb and Simpson (2012).

reasonable values that represent the underlying measurements well. However, suppose the minimum or maximum values are outliers. In that case, the entire population will be skewed, so the result is not evenly scaled on both axes. To normalize the data with the min-max scaling method, one calculates, as in Equation 9, the difference between the observed measurement and the minimum measurement and divides it by the difference between the maximum measurement and the minimum measurement (Sinsomboonthong, 2022).

$$x' = \frac{x - min(x)}{max(x) - min(x)} \tag{9}$$

where:
x' - normalized value of observed measurement
x - observed measurement

2.2.3 Evaluating Cluster Quality

An important step when performing cluster analysis is to evaluate the quality of the resulting clusters. Two of the most common evaluation criteria are that the distances between members of the same cluster should be small, while the distances between clusters should be large (Hennig, 2015). These are called homogeneity criteria and separation criteria respectively, and will be referred to as such in this study, although there are a myriad of synonyms used by different researchers. There are many different ways of measuring homogeneity and separation (Wierzchoń & Kłopotek, 2018), one way of measuring them when the data is two-dimensional is by visual observation of the resulting clusters (Han et al., 2021). This alone might be seen as inferior to mathematical methods as it only provides an estimation instead of absolute values, but as this clustering is applied on real data, one can further evaluate the quality by seeing if the behavior of the objects in the clusters is as expected. Additionally, human eyes are effective at judging the clustering quality in two dimensions (Han et al., 2021). With real data, it is easy to argue for the meaningfulness of the clusters as previously mentioned, while mathematical performance measuring does not say much since it is only informative if the underlying clusters are meaningful (Hennig, 2015). Thus, arguments for the meaningfulness of the clusters combined with visual observations should yield good measures of the cluster quality.

2.2.4 K-means

There is a vast number of different clustering algorithms to choose from, each with different strengths and weaknesses. For this study, the k-means algorithm was chosen because it is efficient when performing clustering on large data sets (Patel & Mehta, 2011). It is also one of the most popular clustering algorithms. Patel and Mehta (2011) further explains how the algorithm works, which in simple terms is by setting k centroids, which each act as the center point for one cluster. Then the algorithm associates the provided data points with the centroid they are closest to. After that the algorithm goes into a loop where the centroids are moved around slightly and the data points are reassociated until the squared error between a centroid and its data points is minimized. At this point, the algorithm is completed as a local minimum is reached for every centroid (Patel & Mehta, 2011).

Clustering algorithms of k-means are sensitive to differences in size between the different parameters that the clustering is based on (Patel & Mehta, 2011). This means that if one parameter has greater magnitudes in its data than the other parameters, then the parameters will not be weighted equally in the clustering algorithm, which is usually not desirable. Therefore, the use and choice of normalization method is important to consider.

2.2.5 Determining the Number of Clusters

One weakness of the k-means algorithm is that it assumes the number of clusters k to be known beforehand, which is often not the case in real applications such as this study. To deal with this there are a variety of methods of estimating the optimal number of clusters. One may simply run the clustering algorithm several times with different k values each time until the best solution has been found. A more structured solution is to use a method that determines the number of clusters before they are constructed. One such method is the Depth Difference method (DeD), which is a fairly new method that, contrary to several traditional methods, does not need to run any clustering algorithm to determine the value of k (Patil & Baidari, 2019). The algorithm itself is complicated and falls outside the scope of this study, but is explained in detail by Patil and Baidari (2019). The creators of the algorithm tested it on 18 different datasets taken from various repositories that are commonly used for testing such algorithms. In these datasets, the real value of k is known, so the accuracy of the DeD algorithm could be compared to several other methods. The results showed that the DeD method was more accurate and less computationally demanding than all the methods it was compared to (Patil & Baidari, 2019). The DeD method is also integrated in a function in the software used to conduct this study (Qlik, n.d.), which makes it easy to use for the purposes of this study.

2.3 IT-systems for Decision Support

Companies have for a long time used IT systems for effective decision-making to various extents. This has been an increasing trend over the past decades, and the rise of new technologies has made it topical in today's environment.

2.3.1 Business Intelligence

Business Intelligence (BI) has been used in companies since the 1990s (Chen et al., 2012), and according to Lim et al. (2013) it can provide valuable insights from data regarding customers, inventory, products, and transactions. As data collection has expanded and become more accessible, the need to filter it to only look at valuable information has become more important. Lim et al. (2013) and Chen et al. (2012) both describe BI as how technologies, systems, practices, and applications can be developed to analyze data to gain an understanding of businesses and markets. The data to be analyzed are business data and the insights gained aim at achieving better operational efficiency, fostering customer relationships, and improving products and services (Lim et al., 2013).

In manufacturing, it is important to make decisions fast and in an effective manner to correspond to customer needs and in order to achieve this a BI tool can be used (Yusof et al., 2013). The authors further state that in a manufacturing setting data is collected from almost all processes and many companies strive to transform this data into information and knowledge that can help them gain a competitive edge. The state-of-the-art BI from a research perspective is focusing on moving towards real-time or near real-time data processing and the usage of more advanced analysis applications than previously (Bordeleau et al., 2018).

2.3.2 Industry 4.0

Industry 4.0 is a concept that the German government initially introduced in 2011 to lead manufacturing companies into the fourth industrial revolution through the use of system-enabled manufacturing (Lee et al., 2014). Technologies such as communication protocols, cloud computing, business intelligence, and big data are, among others, considered core to Industry 4.0 (Bordeleau et al., 2018). Bordeleau et al. (2018) further emphasize that most of these examples are not novel in themselves, but that the combination of them with business processes and data processing creates novelty. The denotation of an "industrial revolution" suggests that there are significant opportunities for actors in the manufacturing industry, which Hermann et al. (2015) mentions as they say that Industry 4.0 promises substantial increases in operational effectiveness, as well as the possibility for new business models, services, and products.

Following promises of such magnitude, the number of publications surrounding Industry 4.0 has been steadily increasing each year (Souza et al., 2021), and such publications mainly come from manufacturing-centered countries such as Germany, Taiwan, and China (Bordeleau et al., 2018). From these publications, Bordeleau et al. (2018) further presents some trends regarding the content. Firstly, many articles include real-time or near real-time data processing, which has immense technical difficulty. Secondly, about half of the reviewed articles include analytics applications like clustering or decision trees. A third finding from their review was that while the operational value of BI in an Industry 4.0 context is commonly measured, it is rarely extended to include strategic value. This leaves a gap to be filled, considering that Industry 4.0 relies on disruptive innovation which should be where the strategic value is found (Bordeleau et al., 2018).

3 Methodology

This section presents the methodology used to carry out the study and is divided into six different subsections. To start, the research strategy used is described, followed by the process and general design of the research. After that, some reasoning around how the theoretical framework was developed is introduced. Then the data collection that was carried out is described, after which the data analysis that was conducted is explained. The section is wrapped up with a discussion surrounding research quality.

3.1 Research Strategy

One of the most important decisions in research is the choice of research strategy which is an action plan designed to achieve a specific goal of the research (Denscombe, 2018). There are many different strategies and they are not in general good or bad, it is rather a question of how they are used that determines if they are right or wrong based on how suitable they are for the research in question (Denscombe, 2018). For this study, a case study is used to explore applications of cluster analysis in material planning. A case study allows for a deeper understanding of the potentially complex relationships in a certain context (Denscombe, 2018), which is useful understanding to have when applying an advanced analytics tool such as cluster analysis with the purpose of getting decision support in different areas. In short, the deep understanding that a case study gives is needed to ensure accuracy in the subsequent analysis in this study.

Another way of distinguishing between different types of research is to classify it as either qualitative or quantitative (Bell et al., 2022). Bell et al. (2022) distinguishes between these two by whether quantification exists or not while Denscombe (2018) explains qualitative research as words and pictures and quantitative as numbers. Even though it is common to classify research into one of these, they can be combined in what is called mixed methods research and is popular in business research (Bell et al., 2022). This allows the study to use methods of both quantitative and qualitative nature and can according to Denscombe (2018) make it possible to get a better picture of the subject being studied. This study has used a mixed methods research approach, leveraging the strengths and mitigating the limitations associated with each method. The quantitative data was the main data used for conducting the analysis, as it is mostly surrounding cluster analysis, while the qualitative data had a more supporting role through achieving a deep understanding of the case company and its context, which would be hard to do with only quantitative data, as numbers in a table alone do not necessarily show the real meaning. Using both quantitative and qualitative data can be seen as a necessity when performing a case study as otherwise one would not be able to get the level of understanding that is needed in such a study (Denscombe, 2018).

Bell et al. (2022) mentions that there are different ways of designing a mixed methods research and there are two issues that need to be taken into consideration while choosing which type, namely the priority decision and the sequence decision. The priority decision refers to how the two methods should be weighted, while the sequence decision refers to which method occurs first or if they occur during the same time (Bell et al., 2022). Connecting to this study, quantitative data will be the main data sources for answering research question 1 and 2, and the results from those are then combined with qualitative data in order to answer research question 3. The sequence decision is thus that the quantitative comes first and the qualitative second, while the priority decision allowed the authors in this study to weigh the quantitative and qualitative findings equally. The choice of weighing them equally is because they aim to answer different questions. This way, the qualitative data can provide a new perspective on the already analyzed quantitative data, helping to confirm some insights and reach new ones.

The data has been analyzed in order to develop some applications regarding how cluster analysis can be used as decision support in different aspects of material planning, as well as to contribute to the authors' understanding of the contextual factors that impact the material planning of the company. A case study is also advantageous here because they are suitable for exploring less studied phenomena, and for developing, testing, and refining theories (Säfsten & Gustavsson, 2020), all of which can be applicable to this study. To improve the quality of a case study, one often uses several methods of collecting data from the case, a concept that is called triangulation (Säfsten & Gustavsson, 2020). These methods could include interviews, observations, enterprise data, and more. The reason triangulation improves study quality is because if there is a weakness in one data collection method then that can be offset by another method that covers said weakness (Säfsten & Gustavsson, 2020).

3.2 Process and Design of the Research

The overarching process that this study was based upon is depicted in Figure 5, and is a 6-step model developed by the authors as a general guiding tool throughout the study work. The first step was to get familiar with the area of research. This included, among other things, understanding the context of material planning, the consequences that has on the study, and getting comfortable with the program where the data analysis was to be performed. The second step was to determine the research questions, purpose, and scope of the research. At this point, the context and research area were familiar to the authors. Therefore it was possible to set a reasonable scope that captured the relevant aspects of this research while not making it too broad. The third step was to plan the study and was about investigating which methods of reaching the desired result could be beneficial, and how they should be used in the study.

The fourth step was data collection, which consisted of collecting quantitative data in the form of enterprise data and qualitative data through an interview with a case company representative. The fifth step was to carry out the data analysis, where the knowledge and data gathered from the previous steps were applied to the problem at hand. This was mainly done by performing a cluster analysis followed by analyzing its output to gain insights that are of significance for decision-making in material planning. The final step was to validate, verify, and discuss the results. This step included verification of the models' functionality, validation, and a workshop with employees at Meridion and the case company to provide substance regarding what decision support the analyses could provide.

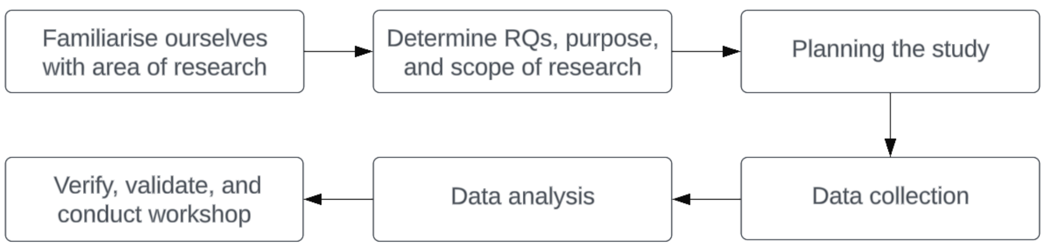


Figure 5: The general steps included in the research process plan.

3.3 Theoretical Framework

To deepen the authors' knowledge of the relevant subjects and to form a basis for the data collection and analysis, a literature review was conducted which is the foundation for this subsection. Most contributions to the theoretical framework came from the early stages of the study but it was also revisited and expanded as new insights emerged over the course of the study. By conducting the literature review early in the study, knowledge regarding the relevant subjects as well as previous research in the field was gathered. This knowledge could then be used to refine the scope of the study to ensure that it provides some form of novelty to the field.

To find relevant literature, search words such as *Business Intelligence, Cluster Analysis, Material Planning,* and *Planning Parameters* were used. The main databases that were used were Chalmers Library, Emerald, Google Scholar, and Scopus. Forward snowballing and backward snowballing are concepts that were used in the study to find relevant literature as well. Wohlin (2014) explains backward snowballing as the process of searching for additional literature by looking into what the current literature refers to. Forward snowballing is the process based on relevant literature searching for literature that refers to the aforementioned literature (Wohlin, 2014). Using these methods made the creation of the theoretical framework more systematic, and proved efficient for identifying relevant literature. Combining this with documentation of what searches already have been done assisted with making the process structured. Additionally, most initial literature searches result in many articles, and therefore a deliberate reading strategy needed to be employed to guide the literature review (Snyder, 2019). Snyder (2019) provides some examples of such strategies, where one of them is to read all articles in the search in full, which would be very useful, but inefficient. Instead, the strategy employed in this study was to divide the review into stages (see Figure 6). This strategy involves firstly the initial search for articles using wide search terms that would give many results, then selecting the titles that seemed interesting for the study and reading the abstracts of those articles. The next step was selecting articles that seemed to have relevant information, and reading those more thoroughly before making the final selection regarding which were included in the theoretical framework (Snyder, 2019). In order to reduce the risk of missing literature that may have meaningful contributions to the study, literature search words were used with both their British and American spelling when applicable.

In total, the abstracts in 40 articles were read. Of those, 26 moved on to the next stage of a more thorough reading. Most of these are used in the study, but because of later changes and revisions, some might not have been included in the final report. In addition, more articles were gathered during the course of the study, as additional topics of interest emerged and needed to be researched and cited.

Figure 6: The overall process used to develop the theoretical framework.

3.4 Data Collection

In this subsection, the methods used to collect the data are presented. The data collection was divided into quantitative and qualitative data collection. The collected data can also be divided into primary or secondary data. Primary data is such that the researchers themselves gather it firsthand, while secondary data comes from secondary sources, which includes previously published research or other data that is not directly compiled by the researcher (Rabianski, 2003).

3.4.1 Quantitative Data Collection

The quantitative data collection only consists of primary data collection, as it was limited to enterprise data provided by the case company. This data was mainly transaction data, which served as the main platform for the analysis performed. The transaction data included data of each transaction such as a customer order or production order, item number, quantity, date, and available stock after the transaction. Two sets of transaction data were collected, the first one had transactional data over all items in the analyzed facility from January 2022 to December 2022. The second set of data only included data from a specific production step and covered the time period from January 2018 to April 2024. In addition, other data was collected that supported the transaction data where needed. This was, for example, article values and current (as of April 2024) parameter settings, such as ABC classification, EOQ, and planning method. Such data was used in the data analysis not only to create the cluster analysis but later also to further analyze and validate some parts of the analysis to a certain extent. These data sources were confidential information of the case company and are therefore anonymized in this study. As there were large amounts of data, it was important to create visualizations of it in terms of graphs, tables, or charts as those make it possible to see the big picture (Säfsten & Gustavsson, 2020). In addition to helping the researchers' own understanding, using tables and charts to visualize quantitative data is also an effective way of sharing findings with others, according to Denscombe (2018). He further states that large volumes of quantitative data can be analyzed relatively quickly. However, there is a risk that these volumes could be too large and thus make the analysis too complex (Denscombe, 2018). Thus, the importance of the previously made delimitations becomes evident to ensure the analysis remains tied to the purpose of the study throughout the entire process.

3.4.2 Qualitative Data Collection

In addition to the quantitative data that was collected, there was also a need to collect some qualitative data. According to Denscombe (2018) interviews are best suited when there are complex things that need to be understood and a need to understand how different factors are connected, which is the case in this study. Säfsten and Gustavsson (2020) explains that interviews help the researcher get both deep and detailed information about the subject and the interviewee. Therefore, an interview with a representative at the case company that has a senior role in logistics was conducted. The questions that were asked to the interviewee were chosen to give an increased understanding of the processes and to also help in understanding the quantitative data that was collected. For example, insights regarding the as-is state of the material planning processes at the case company were gathered through the interview. Data collected here was for example regarding variation, item behavior, ABC classification, use of BI, and planning parameters. The complete interview guide can be found in Appendix A. The questions asked also had the aim of increasing the authors' understanding of the case company's specific planning context. Although the respondent could answer most of the questions asked, they were unable to answer some of them so a second interview with another respondent, who also worked within logistics but in a more operational role and handled these issues more often, was planned. However, there ended up being no opportunity to conduct another interview so instead a set of questions was answered in writing, unfortunately limiting the possibility for follow-up questions. This was only the case for a small number of mostly closed questions regarding planning parameters.

There are different ways an interview could be conducted, according to Säfsten and Gustavsson (2020), and the one chosen in this study is the semi-structured format. The reason for this is that it allows questions on specific topics while still allowing for reflections and discussions that have not been prepared in advance (Bell et al., 2022). This can be suitable if an interviewee says something interesting that the interviewers want to ask further questions upon. The questions asked were a mix of closed and open ones, which allowed the interviewee to speak freely and give answers that could broaden the understanding of the interviewers. To make sure the main questions and areas of interest were covered and the interview stayed on topic, an interview guide was created that served as a foundation for the questions asked. But as mentioned, it allowed for follow-up questions if certain answers invoked interest. When creating the interview guide the authors took ethical aspects into consideration, which is described more thoroughly in Section 3.6.1.

According to Denscombe (2018) it is good to use some sort of method of recording interviews since humans tend to remember things differently than they were said. Therefore, the interviewers took notes during the interview in order to gather the most crucial details. The interview was also recorded, with the interviewees' consent, making it possible to go back and ensure that the answer perceived by the interviewers was the one given by the interviewee, as this is of importance in the analysis and reporting to ensure the accuracy of the study (Bell et al., 2022). This procedure also aligns with how Denscombe (2018) recommends recordings to be carried out, that notes should be used as a compliment to sound recordings.

In addition to interviews, qualitative data was also collected during a workshop with representatives from both Meridion and the case company, where the representatives from the case company came from senior positions within logistics and analytics. Ørngreen and Levinsen (2017) mention that workshops can inspire new insights into the research and is beneficial to combine with other methods. The data collected during the workshop aimed at identifying areas where the cluster analyses could be used as decision support, which is in accordance with Säfsten and Gustavsson (2020) findings that a workshop could encourage a common use of knowledge to uncover solutions in practical scenarios. The workshop started with the study's authors presenting the cluster analyses that have been developed and their thoughts on potential application areas. The other people present during the workshop then gave feedback and suggested additional areas where they could provide decision support. All presented potential application areas were then discussed regarding usefulness and added value.

3.5 Data analysis

This section describes the methodology surrounding the data analysis that was performed throughout the study. It describes from start to finish which general steps were involved in the process, including how the data was processed and which methods and techniques were employed to reach the final results.

3.5.1 Preprocessing

The first step of the data analysis process was some initial preprocessing of the data. This involved restructuring data into forms more suitable for the analysis and handling of null-values. The restructuring of data involved was done differently depending on which cluster analysis it was for. For two of the cluster analyses, the transaction data was aggregated on item level and time period, both monthly and weekly. For the final cluster analysis, another level of aggregation was added, namely transaction type, as it was important to differentiate between different types of orders, such as manufacturing orders receipt, and issued when conducting this analysis. In some cases, the occurence of certain items was really low, these were then put in an own category and removed before the cluster analysis was performed. These items could be analyzed on their own as including them in the cluster analysis would most likely have made it worse. However, these items were not analyzed in this study as they have a minimal impact.

Regarding the handling of null- or missing values, a common technique is the removal of the attributes or cases where missing data occurs, but a more favored method is imputation (Aljuaid & Sasi, 2016). Removal of null-values was not preferred as in this study it would mean that an entire item is removed from the analysis, and one wants to keep as much of the data as possible intact. Imputation refers to using a sensible approach to predict what value would have been in place of a missing value and is the preferred technique if there exists a sensible approach for the specific case. There are a wide variety of different methods of imputation that each are sensible to use in different cases (Donders et al., 2006; Engels & Diehr, 2003; Rässler et al., 2012). There are often several imputation methods that could be reasonable to use and it is often difficult to compare the accuracy of these (Engels & Diehr, 2003). The choice of imputation method is important as the wrong choice can lead to severe biases in the analysis (Donders et al., 2006). A sensible approach in the context of this study is a method that provides values that are logical and can be soundly argued for. In this study, some null values were generated when the aggregation was performed. During aggregation, a value for the average on-hand balance over the aggregated time period was generated, but if the item in question didn't have any transactions for one of the time periods then a null-value would be generated. In this case, imputation of the previous time period's average on-hand balance makes sense since without any transactions the on-hand balance can not have changed. The aggregation also caused a similar issue with the aggregated demand per time period, which would be null if no transactions existed in the given time period. Here, the value of 0 was imputed instead as if no transactions exist then there has been no demand during that time period.

3.5.2 Item Behavior Characteristics

The second step of the data analysis was to determine which item behavior characteristics that exist. Item behavior characteristics in this context is defined as in Section 1.1 and relates to the components that affect the demand. This was done by having some initial discussion with employees at Meridion that resulted in some valuable insights. These insights were further built upon during an interview with a representative at the case company where the previous insights were validated and also some new were gained. Generating ideas for potential characteristics in this way enabled the authors to get ideas from practitioners working with material planning. In addition, the authors discussed between themselves throughout the whole study to find further potential characteristics.

The insights gained from the conversations with employees at Meridion as well as the case company served as a basis when it was time to analyze the data set from the case company. The suggested characteristics were first analyzed, hoping to see the connection between those and the item behavior. After that first analysis, the authors own potential characteristics were also analyzed. To be able to analyze the large amount of data it was visualized in graphs, in Qlik, and then these visualizations were analyzed. This was done by visually inspecting the graphs and looking for connections between the characteristics chosen and the item behavior. By identifying connections between the graphs and the proposed item behavior characteristics, the characteristics were able to be verified.

3.5.3 Cluster Analysis

The next step in the data analysis is the cluster analysis, which was conducted using the k-means algorithm. Here it was firstly important to determine which parameters the cluster analysis should be based on, an important decision considering the parameters determine what meaning the cluster analysis has. The evaluation of the cluster parameters were based on several sources, including an interview with a representative of the case company, a set of questions sent to a representative of the case company, findings from the literature review regarding how good cluster quality is visualized, and the authors' reasoning regarding which parameters could yield the most satisfactory results. Since the cluster parameters determine the meaning of the cluster analysis, a hypothesis of what the analysis should strive to show is good to have before selecting cluster parameters. This is not to say that one should know what the analysis will look like, rather there should be an idea of what purpose the cluster analysis will have that can guide the selection of features. Several combinations of parameters were tested and evaluated as features until satisfactory solutions were reached. The second step was to prepare the data for k-means clustering analysis and primarily involved normalization of the data due to the k-means algorithm being sensitive to differing variable sizes. As such, two different normalization methods were tested, namely Z-score normalization and min-max scaling. The main difference between these normalization methods is that min-max scaling equally scales all values in the range relative to the minimum and maximum value, meaning an outlier could have a large impact on the result of the cluster analysis, while the Z-score considers the arithmetic mean and standard deviation to better account for outliers that remain in the data set.

One of the most popular clustering algorithms is k-means clustering, because it is efficient and simple when performing clustering on large data sets (Patel & Mehta, 2011). Aside from being efficient on large data sets, k-means clustering was chosen because the BI system used in this study, Qlik, has built-in functions for k-means clustering that allow the users to relatively quickly build and test different clustering models (Qlik, n.d.). This simplified the cluster analysis and allowed the focus of the study to remain closely tied to the research questions. In addition, using tools available in the BI system allows the authors to deliver solutions that with ease can be adapted by practitioners.

Determining the number of clusters was first done by the built-in method DeD in Qlik. However, when evaluating the clusters' quality by the homogeneity criteria and the separation criteria, they were sometimes not very good. In those cases, additional attempts of cluster analysis were conducted, with the number of clusters set manually, starting at two, and incrementing by one at each try. Upon identifying a number of clusters that performed well in terms of the two above mentioned criteria, the iteration continued for a couple more rounds and evaluating if the results became better or not. If the results was better the iteration continued, if not the number of clusters was determined by the one attempt that yielded the most satisfactory result based on the two criteria.

Regarding outliers, they can be defined as a data point which does not fit the overall pattern of the clusters, and that perhaps should be removed to make the clustering more reliable (Hautamäki et al., 2005). However, in this context one also needs to consider that an outlier may still have significant meaning in practice, so it should not be removed simply because it does not fit the overall pattern. Perceived outliers occur in this study, as values that reside far away from most other values were at times encountered. These values were often legitimate ones that had significant meaning in practice, so they could not simply be removed.

After satisfactory clusters were created, an evaluation regarding what insights the clusters provided in relation to the hypothesis from before was conducted. These insights were then used to connect to known theories in material planning, which could motivate implementing some action on the items analyzed depending on which cluster they belong to, and in that way get decision support from the cluster analysis.

Finally, the finished clustering model needed to be validated. Validation is an interesting issue as the clustering process does not have clear examples of what a desirable solution looks like since it is an unsupervised process (Halkidi et al., 2001). This is done with the same homogeneity and separation criteria mentioned when discussing the number of clusters, but done in a more objective way as these criteria were only used to determine which solution had the best quality, but not if that entails that the quality could be considered good, or valid. Since the cluster analyses in this study only have two parameters, and therefore can be visualized easily, one can inspect how homogeneous and how separated the clusters are. This allows for a simple way of validating any given cluster analysis, and can also with ease be used by practitioners.

3.5.4 Decision Support

The final step in the data analysis was to investigate what decision support the cluster analyses could provide in material planning. This was done by combining the authors' own reasoning with findings from a workshop that was conducted.

The authors first discussed and reasoned regarding what the cluster analyses meant and what decision support could be provided by them. This was done through visual inspections of the clusters along with supporting graphs, for example, stock levels and demand, and then discussed with each other. These insights were then used as the starting point for a workshop with knowledgeable staff at both Meridion and the case company. The goal of this workshop was to have professionals give their opinions on the cluster analysis results and what kinds of decision support each cluster analysis could provide. The authors' own reasoning was used as a starting point in this workshop in order to provide additional context and a starting point for the discussions that were to be held. The results of this workshop were combined with the authors' reasoning to find the most likely applications for this type of analysis, as if several sources believed in the same type of decision support then that would be more likely to be accurate.

3.6 Research Quality

This section covers the research ethics considerations that apply to this study and how they were adhered to. Then, the reliability and validity of the study is explained. Finally, relevant method criticisms are brought up and addressed.

3.6.1 Research Ethics

Research ethics is an important subject to treat in a study such as this one. Research ethics is essentially the norms and values that researchers are expected to adhere to in order to ensure their research is accepted by the scientific community as well as that it benefits society as a whole (Säfsten & Gustavsson, 2020). The authors further present several reasons why scientists should consider research ethics in their studies. One is that it protects the reputation of scientific research as a whole, as if society doesn't trust that research is presented correctly and honestly, then it loses its purpose. A second reason mentioned is that participants and consumers of the research need to be protected (Säfsten & Gustavsson, 2020).

Säfsten and Gustavsson (2020) additionally point out that students carrying out degree projects also need to adhere to research ethics as they will be exposed to many ethical considerations during the degree project. This is indeed true for this study, one reason being that the research involves other people through an interview being a part of the data collection performed. Blomkvist and Hallin (2015) present four requirements for ethical research within social sciences, originally adapted from the Swedish Research Council ("Good Research Practice", 2017) (originally from 2011, revised 2017). These are as follows:

- **The information requirement** - The people who are studied (e.g. interviewed) have to be informed about the purpose of the study.

- **The consent requirement** - The people who are studied have to agree to being studied.

- **The confidentiality requirement** - Material collected or created during the study has to be treated confidentially.

- **The good use requirement** - Material collected during the study may only be used for the purpose that was stated when collecting the material.

Regarding how these requirements have been fulfilled in this study, the first two were fulfilled by informing participants about the purpose of the study, how it would be performed, what university and which individuals were responsible, and that participation is voluntary. This is in accordance with Säfsten and Gustavsson (2020). These two requirements are in conjunction also known as informed consent (Säfsten & Gustavsson, 2020). The confidentiality requirement has been fulfilled by asking the providers of data about what is okay to show, and through letting them inspect the paper and request changes in the anonymity level of their data. Additionally, persons who participated through being interviewed have been anonymized and instead identified by a key, in accordance with Säfsten and Gustavsson (2020). Finally, the good use requirement was fulfilled through establishing a clear purpose early in the study so that the purpose could be communicated to the providers of data, and then maintaining that purpose for the duration of the study.

The study also adhered to ethical aspects highlighted by Denscombe (2018) including that the data collected should be stored securely, for as long as needed only, be used for its original purpose, and be anonymized. Another aspect that was important to take into consideration when creating the interview guide was to not ask any leading questions (Bell et al., 2022). The importance of ethics in the written report should also be taken into account. This does mainly consider plagiarism and the use of other researchers' works. To give credit when its due, to support something, or using a model, the reference system APA were used.

3.6.2 Reliability and Validity

Two concepts that are crucial in research methodology is reliability and validity since they contribute to the study findings' credibility and trustworthiness. Reliability refers to how well a study can be done by another researcher following the same method and result in the same findings, which could be described as how well the study can be replicated (Denscombe, 2018). Validity is about if the data in the study is accurate and precise (Denscombe, 2018). This can be seen as if it actually looks into what it is intended to do and covers relevant aspects.

For a study to be able to be replicated, and have a strong reliability, it is important to thoroughly explain what has been done and make sure that the reader can comprehend it. This study had this in mind throughout the work and has therefore tried to explain what has been done to the greatest extent possible. The visualizations from which the analysis is based on is shown as well to help the reader better understand the thought process of the authors. During the study the choice of methods for the data analysis have mostly been already existing methods in Qlik in order to make it easier for others to replicate the study. Since the methods in Qlik is well documented on their website, it enables others that does not use Qlik to replicate the methods employed as well after some adaptation to their system. When it was a need to deviate from already existing methods the calculations done is presented, making it easier to follow the process. To further understand what questions were asked in the interview that served as a foundation, the interview guide is included as an appendix to the report (see Appendix A).

In terms of validity, the thorough theoretical framework has ensured that all choices of methods have taken into consideration the strengths and weaknesses of each. That way, methods were chosen that best fit the specific context of the study. For example, the choice of normalization method is different for different data sets since they are each good at handling different types of data sets, as explained in chapter 2. Additionally, triangulation was used in several areas of the study to add validity to the research. For example, some claims from the interview could be either verified or discredited through analyzing quantitative data. Another example is how the authors' knowledge was combined with the knowledge of other employees at Meridion to increase the likelihood that the interpretations made were correct. Another aspect is that in the cases where parameter settings are proposed, the suggestions can be compared against the current parameter settings as a form of validation. The expected result would be that many of them are the same, but with distinct differences that can be traced back to the methods employed to conduct the analysis in this study. However, one can not be sure that the current parameters are a good reference for the validity of the analysis since the authors are aware that they can be infrequently updated. However, it is still a fairly good reference point. Additionally, the usage of imputation can have an impact on the validity of the result, but since the need for imputation has been limited to cases that have obvious, correct imputation methods, this should not be a concern. Finally, both authors were present at the interview in order to reduce the risk of misinterpretations occurring.

3.6.3 Method Criticism

The data used in the study can also be criticized to some degree, specifically that only one interview was conducted, so the sample size is very small and therefore, potentially unreliable. Because of this, the need for triangulation was even larger to ensure the validity of the study. The findings from the quantitative data were not contradictory to the answers given in the interview and set of questions, therefore that data is perceived as valid. Although, the person interviewed has been in the company for quite a long time and has great insight into the organization. Since the person was not completely sure of the answers to some of the questions and recommended to ask someone with better insight, one could assume that the answers the person gave were in fact true.

The quantitative data that has been used in the study has only been analyzed using a time period of one year, which could be seen as a too short timeframe. As a consequence of this the seasonality in demand for the items is not included in the analysis in order to not wrongfully make assumptions. However, if the data covered a wider timeframe seasonality would have been good to analyze as well when analyzing the variation in general. The current parameters used in the company may not be correct and there have been no good way of verifying this. Therefore, it is important to be a bit skeptical towards them when performing the validation. However, one could assume that they should be pretty good since they play an important part for the company.

Another valid criticism is that the cluster analyses that have been performed are only using two features each, while cluster analysis often uses more. There is definitely a case to be made that there are more factors that could have been included as features to get a more complete cluster analysis. Including more features that are not correlated with any other features should in theory result in more accurate clusters. However, this is a trade-off that has been deliberately made in this study in order to be able to visualize the clusters in two dimensions which vastly increases the readability. This is especially important when the person using the cluster analysis is not an expert in it, which is believed to be quite common for the applications these cluster analyses will be used in.

4 Empirical Findings and Analysis

This section presents the empirical findings and analysis conducted. It begins by presenting some information about the case company that mainly are findings from the interview and set of questions to provide context for the rest of the analysis. Then, the concept of item behavior is broken down into its underlying characteristics. These characteristics are then combined in three different cluster analyses for three different applications. Finally, analysis is conducted on these cluster analyses to discover what decision support they can provide to material planning.

4.1 Case Company

The case company in this study is a manufacturing company that has multiple production facilities and a presence all around the world. They produce components that their customers in turn are using when producing their products. In Sweden, the company would be referred to as a large company and they have a history that stretches more than 100 years back in time. They sell to many big international companies, especially in the automotive industry, and according to themselves they are trying to be innovative as well as sustainable.

From the interview and the set of questions, respondents from the case company gave valuable insights that laid a foundation for the authors' understanding of the company's processes. Through an explanation of the manufacturing processes performed as well as which production steps are problematic, the authors could identify which production step the analysis would prioritize. There was also other valuable information that was brought up during the interview that was of importance for the analysis. The case company said that maintaining a good service for their customers is the most important part, in some situations whatever the cost. Relating this to the supply chain triangle it will mean that the only dimension it can vary between is cash and cost, as decreasing the service is not an option. Additionally, the planning environment is of the type "batch production of standardized products". The planning environment is characterized by that there is quite a bit of variability in the demand, which is dependent on their customers' forecast quality as they produce according to those forecasts. In order to increase the accuracy in forecasting when customers sometimes only have forecasts for a couple of months, the case company has recently started to forecast many items for 18 months ahead, in order to be able to plan according to their own forecast when the customers' forecast is no longer considered reliable enough. Because of this variability, or instability, in demand, MRP is a suitable method for material planning and that is the method used by the case company.

During the study, it was brought up that the case company has a current goal of lowering their tied-up capital since it has previously been very high. Probably, the reason for doing this is because of the present economic situation with high interest rates. Another possible explanation is that many companies built up tied-up capital during the Covid-19 pandemic and the disruptions that followed, as the supply uncertainty was high, and now need to reduce it. During the interview, it was also said when they are manufacturing products, they should not make it in batch sizes exceeding the yearly demand for that product. This is a rule that corresponds to the goal of lowering the tied-up capital.

An issue, which according to the case company is the largest, is to even out the production. Currently, there are issues where during some periods the total demand is higher, and in other it is lower, creating an uneven production rate. If the production instead could be more even it could yield satisfactory results for the company. One thing that is needed to be able to make this happen is to have system support, according to the interviewee. By having this kind of system support it could help them plan the production in a more efficient way which will result in an even production rate. From previous analyses, it has been concluded that one reason for this issue is that the machine capacity has been too low as a consequence of frequent changes of items produced resulting in a large amount of the available time going to setup time instead.

Table 3: Summary of main characteristics of the case company's planning environment.

Main Characteristics of Planning Environment
High service level
Batch production of standardized products
Quite much variability
Mainly producing according to customers' forecasts
MRP is used

4.2 Research Question 1

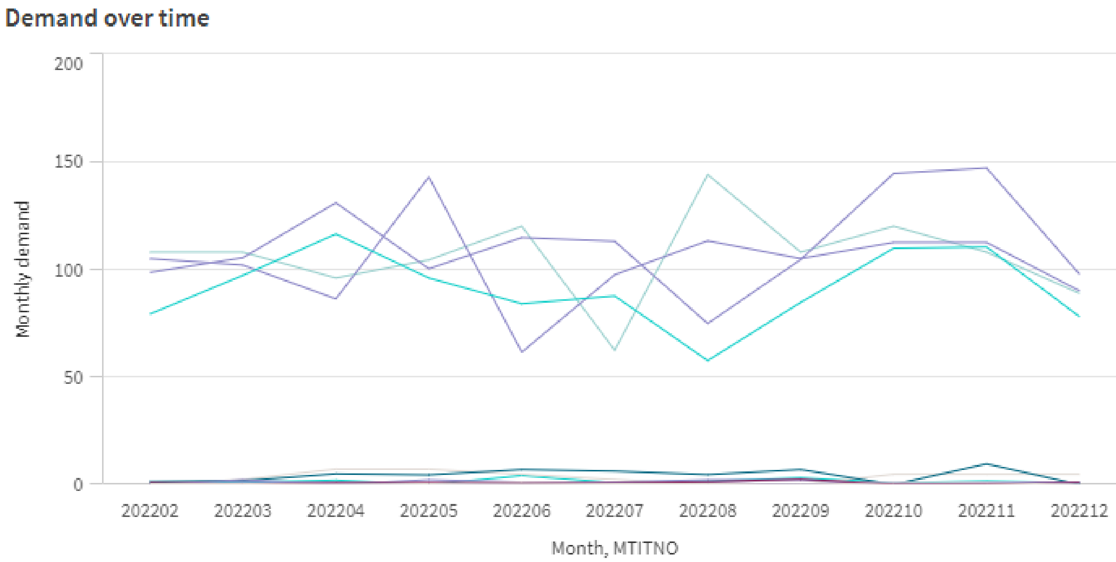
The first research question to be answered is what different kinds of item behavior characteristics exist in material planning. In order to answer this question one first needs to define item behavior and item behavior characteristic. Thus, the previously mentioned definition is repeated here.

> Item behavior refers to how an item moves up and down over time in terms of demand. Item behavior characteristics are the underlying factors that create the item behavior and are what this question aims to identify.

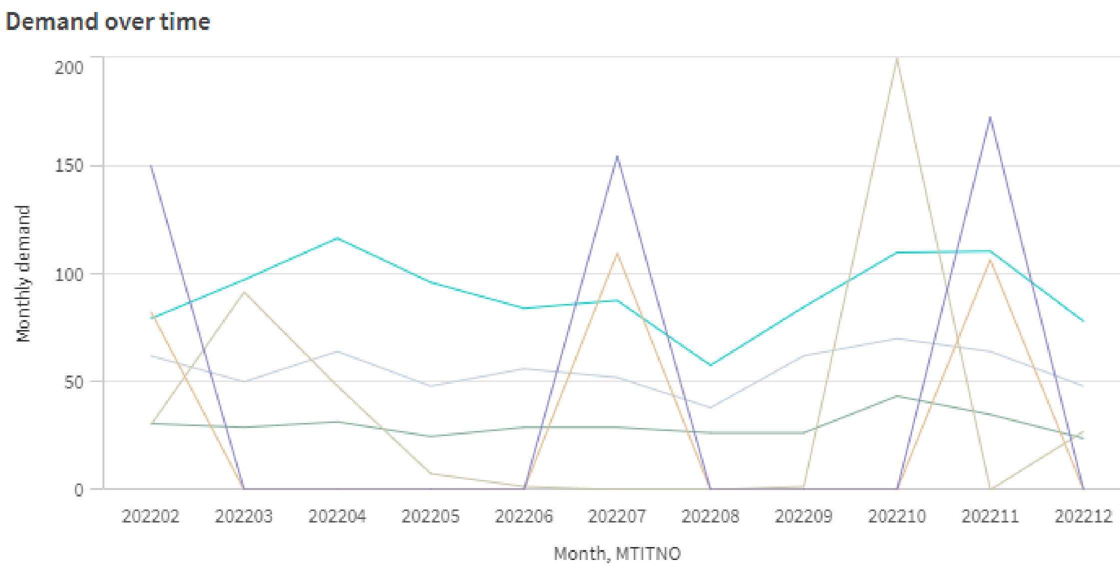
Research question 1 is closely tied to the quantitative data that was collected from the case company, and understanding the structure and nature of the aforementioned data may be needed in order to understand the specific results that are presented. Therefore, the structure and nature of the data are shared here. The raw data was a list of stock transactions and was limited to one of the case company's facilities. These transactions had a field identifying the type so that one or more types could be filtered out when needed, and otherwise contained a wide range of information regarding these stock transactions. For use in answering this research question, the data was transformed in order to be grouped by item number and date, as it is a necessity to produce the demand graphs that are analyzed in this research question.

Demand graphs were analyzed because it was a good starting point, as the item behavior characteristics are supposed to be visible in those graphs. Additionally, from the interview and the set of questions with the case company it was concluded that the most important characteristics for the case company are the customer demand and variation of the demand, as they have a large impact on their business. The authors also reason for the importance of demand as an item behavior characteristic as it is crucial in understanding what magnitudes one is working with. Customer demand is important to consider when performing material planning since it in the long run is what determines what needs to be produced. It will be important to produce quantities of items that corresponds to the actual customer demand. So of importance is both the time of customer demand as well as the quantities demanded, otherwise one will find oneself in a situation where either the tied-up capital is high or the service one offers is not satisfactory, because of shortages. But to just consider each customer's order on its own and working towards that will not be efficient since a large portion of the time available will be lost in setup times, thereby increasing operational costs. Hence, it is important to consider the total demand for each item when planning for production. A conclusion that thus can be drawn regarding what affects the planning is if the item is demanded often or seldom and if the volumes demanded are large or small. Depending on how these two characteristics behave, it will imply different approaches for the material planning. A graph showing the difference between articles' volume of demand is presented in Figure 7a and a graph showing differences in lumpiness in demand is shown in Figure 7b.

The variation in demand could refer to different types of variation. This was not elaborated on during the interview, but instead, the authors reason that variation can come in multiple forms, each of which should be separated. One of these is random demand variation, where there can be naturally occurring changes up and down each period. Figure 8 aims at showing the difference between two articles where one has a large



(a)



(b)

Figure 7: Graphs each visualizing articles' demand each month during 2022. 7a aims at visualizing the difference in demand quantities and 7b aims at visualizing the difference in lumpiness in demand.

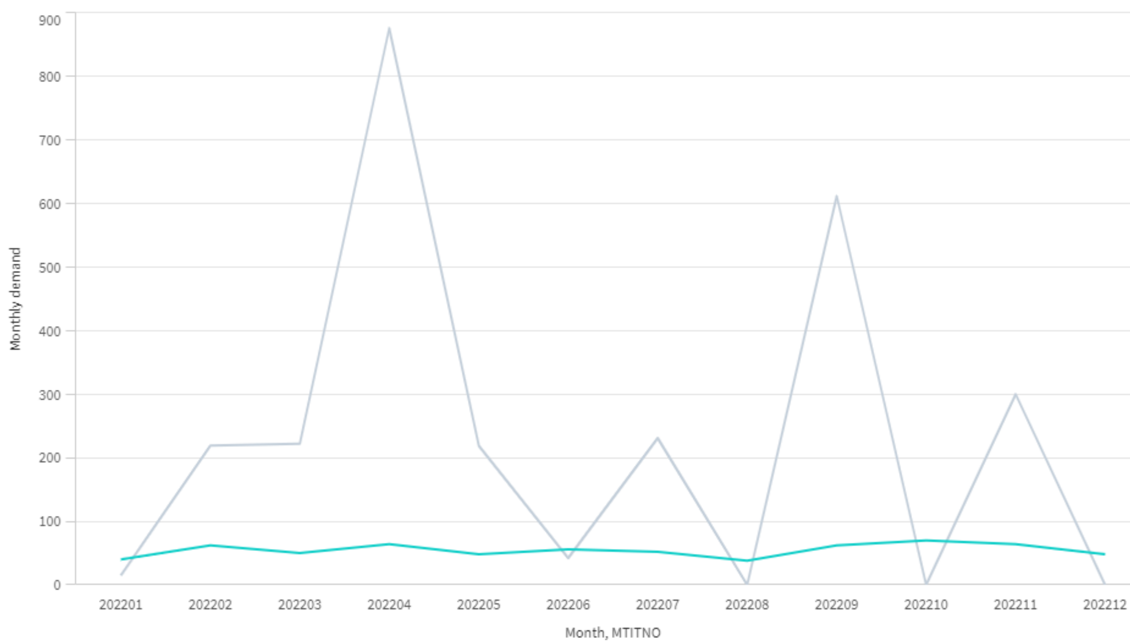


Figure 8: This figure aims at visualizing the difference in random variation of demand.

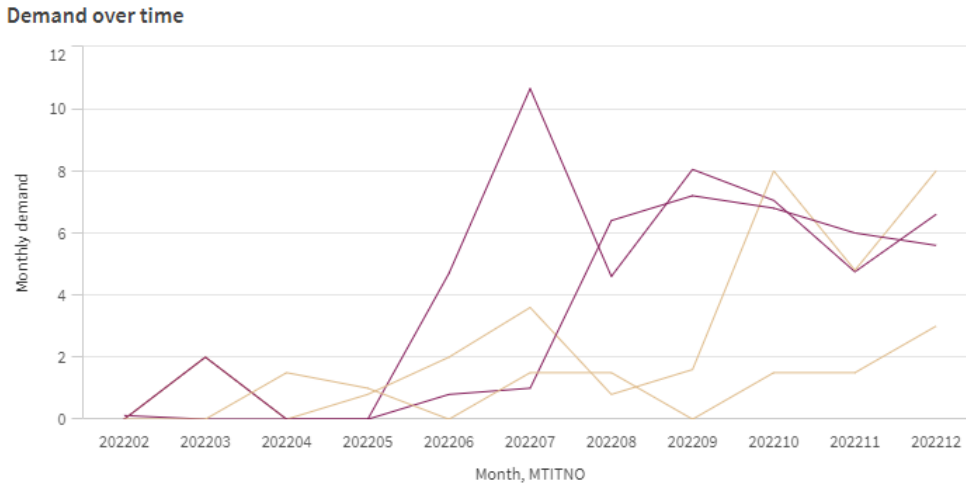
random variation in demand while the other does not. This type of variation could be analyzed without the need to have many time periods, although the answer given from having more periods will be more accurate because of the larger sample size. The higher the variation is, the harder it will be to make a good forecast. Without a good forecast it will be challenging to perform the material planning in a good way. Deviations between actual and forecasted volumes may result in shortages or excessive inventory, depending on which direction the deviation has. The variation described here is commonly calculated by taking the standard deviation divided by the arithmetic mean as in Equation 7 and is referred to as the coefficient of variation.

Another type of variation in demand that was identified for some items were multiple periods in a row that all had an increasing or decreasing demand. This would be referred to as a trend in demand and examples of this is shown in Figure 9. In simple terms, the trend could be described as how much the demand increases or decreases each month in either absolute or relative terms. A trivial approach to determine the trend is to calculate the m-value in the function y=mx+b over the analyzed time periods. This serves as a convenient way to obtain a measure of the trend in a single number, making it suitable for use in a cluster analysis. If there is a trend it will imply that the material planning should take it into consideration when planning and the action they will take will depend on if it is positive or negative. If it is positive it will probably mean that the demand will continue to grow, hence there will be a larger need for items and the number of produced items should be increased. The safety stock may need to be increased as well, as with larger volumes, the amount of random variation should also increase. The reversed logic can be applied to an item with a negative trend, there it may be of interest to instead reduce the batch sizes and safety stock. When analyzing the trend it is important to consider the number of analyzed time periods. If only a couple of periods are analyzed, there are too few data points to say that a trend is present, while if too many periods are analyzed then the formula might not pick up on any recent trends as they may not be large enough to change the value of m. Therefore, it is of great importance to determine the number of periods the trend should be based on, otherwise it would result in too frequent updates of planning parameters.

Another pattern of variation in demand could be that every so often a demand pattern repeats itself, which is referred to as seasonality. This type of variation was encountered in this study when studying graphs containing several years of data. However, for all three of the cluster analyses that are to be presented, the time period analyzed is only one year as the most recent data is what matters for those cluster analyses. Historical data that is more than one year old is likely to be outdated already, and therefore unfit to use as a basis for updating parameters. Therefore, seasonality as an item behavior can exist, but can not be used in the coming cluster analyses. Seasonality is important to consider for items when planning production since the demand for a certain item can vary significantly between different periods of time. An example of this is the sales of winter clothes which is not distributed evenly over all months in a year, rather much higher in some than others. To be sure that it is an actual seasonality that exists the number of time periods the data is analyzed over should be over at least two years as one year is one of the longest, commonly occurring demand cycles. The data needs to cover that many periods in order to see that the pattern repeats itself for the same periods in each new cycle. If not accounting for this it could be another form of variation that is mistaken for seasonality. To summarize the findings from research question 1, the identified item behavior characteristics can be seen in Table 4.

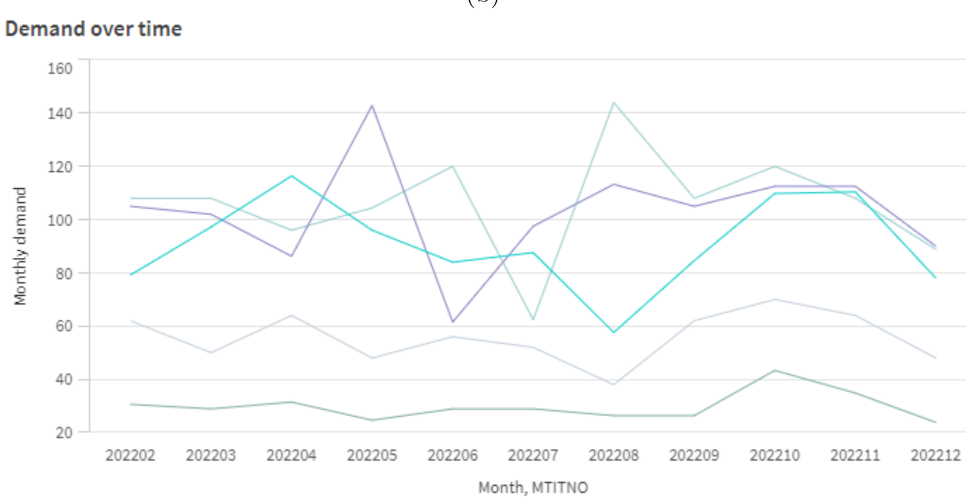
Table 4: Identified item behavior characteristics.

Item behavior characteristic
Volume of demand
Lumpiness in demand
Random variation
Trend in demand
Seasonality (not applicable for this study)



(a)

(b)



(c)

Figure 9: Graphs visualizing articles demand in terms of trend. 9a aims at visualizing items with an increasing trend, 9b aims at visualizing items with a decreasing trend, and 9c aims at visualizing items with no strong trend.

4.3 Research Question 2

The second research question is also closely tied to the quantitative data that was collected from the case company, as the purpose of this question is to use the item behavior characteristics from research question 1 in a cluster analysis to group items with similar item behavior. This can be based on many different combinations of features from the item behavior characteristics, therefore three different cluster analyses are presented in this section.

The choice of which issues to apply the cluster analysis to was based on common planning scenarios where the authors saw potential benefits of applying cluster analysis, and to some extent, problem areas of the case company. The first cluster analysis concerns safety stocks, and how to update the associated parameters. The second cluster analysis aims to perform an ABC-XYZ analysis and apply relevant strategies to each class. The final cluster analysis concerns batch sizing and updating of the associated parameters. As mentioned in 3.4.1 the study used two different data sets, where the one that covered all items from 2022 was used for the first cluster analysis, and the second data set was used for the other two. The reasoning is that when investigating safety stocks, the specific production step that the second data set contains is not of particular interest. For the second cluster analysis, this may also be the case, but data for item value was only available for the second data set so that one was chosen, as item values are needed to obtain the volume values that the ABC-XYZ analysis uses. For the third cluster analysis, the second data set was chosen since the production step it contains is particularly interesting to perform that analysis on, as it is a production step that has had issues and is critical for the case company.

The cluster analyses that were to be performed were also combined with findings from the literature and the interview to understand the given context at the case company and more generally in material planning. Understanding the context is important in order to be able to conduct the cluster analysis well. If not, the findings from it may not be relevant at all. Therefore the findings from literature and the interview served as a foundation that the cluster analysis further built upon. Those findings were important in order to know what the goal of the cluster analysis should be. That is, what it should aim at achieving in the form of grouping and identifying items based on their characteristics and behavior was crucial to consider here.

In a cluster analysis, the features need to be chosen so that the clusters created actually represent what they are intended to. Hence, it was important to evaluate different features and think through if the result would show what was desired. This becomes especially important when working with real data, where the chosen features entirely determine the meaning of the cluster analysis. Since the cluster analysis groups the items based on the chosen features, the features needed to correspond to what characterizes the similarities and distinctions between clusters.

Another aspect to consider when performing the cluster analysis was to normalize the data since the raw data was not normalized. Both normalization methods Z-score and min-max scaling were considered for each analysis and the one chosen for each was the one that was best suited for the data structure seen in each cluster analysis. In Qlik, switching between normalization methods was a simple task as only one parameter in the k-means function needed to be changed. Another positive outcome from the normalization was that the two chosen features were weighted equally.

4.3.1 Cluster Analysis 1

The first cluster analysis developed concerned setting the safety stock level for items and was an application that the authors had in mind from early in the study as a promising application for cluster analysis. As previously mentioned, the choice of clustering features is strongly tied to the meaning that the cluster analysis is intended to have. Therefore, factors that impact how a safety stock level is determined were evaluated as potential clustering parameters. A standard formula (Equation 2) for determining the safety stock, qualitative information from the conducted interview, and other theoretical evidence served as a base for selecting these potential clustering parameters. From the standard formula, the lead time, demand during lead time, standard deviation of the demand during lead time, and the standard deviation of the lead time can be obtained, but not all of these are of interest in this analysis. Firstly, lead time information was not available to the authors, so it was not possible to use that as a parameter. Secondly, it was preferred to avoid using features that correlate to the absolute size of the demand, otherwise, it would not be possible to analyze similarities in behavior among items with small demand and items with large demand. Therefore, relative measures were chosen as features going forward with this cluster analysis. It should however be noted that one could have included the demand as a third feature, in order to differentiate between items that have more or less demand, as they would have differing levels of importance to the company. Since the authors have made delimitations to only use two parameters, this was not pursued.

The first chosen features was the trend, as an increase or decrease in demand will affect the safety stock needed to maintain the same service level, and therefore a trend may be cause for updating the parameters associated with safety stocks. This feature was measured as the m-value from the trend calculation ($y = mx + b$), but it was needed to be divided by the average demand in order to decouple it from the absolute size of the demand. The number of periods was set as 12 months, as that was reasoned to be short enough to pick up on emerging trends and long enough not to misinterpret random variation as a trend. However, the desirable time period to use may differ between industries. For example, in the fashion industry one would want to use significantly less than a 12 month horizon. Thus, it needs to be adapted to the specific context. The second chosen feature was the coefficient of variation (χ), from Equation 7. This feature is by definition already relative to the volume of demand, and therefore required no more feature extraction. The motivation behind using the coefficient of variation as a feature is since variation in demand is the only component of the safety stock formula for which data exists and that a relative measure can be calculated from. These were then used as the features in the first cluster analysis, as seen in Figure 10.

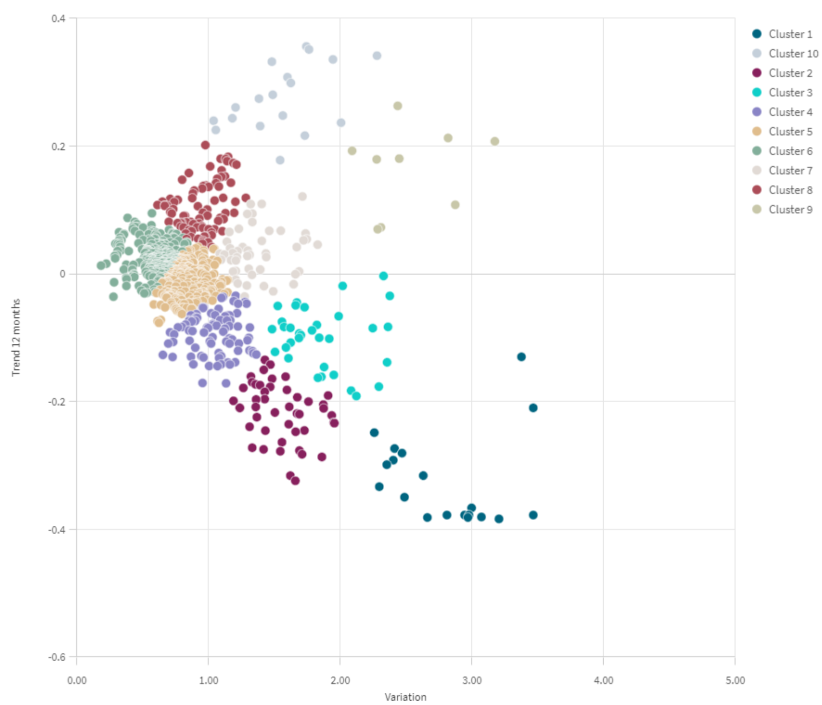


Figure 10: Cluster analysis for relative trend versus coefficient of variation.

The DeD method gave a result of 4 clusters for this cluster analysis, but those clusters had no clear separation, and they were very large so the homogeneity was also low. Instead, the number of clusters was in the end set at 10, as that was found to give a good balance by having enough clusters to recommend fairly specific actions on each cluster while not sacrificing the validity, as the homogeneity and separation was better. However, the separation was still not the best. Additionally, min-max scaling was used since all relevant values fall within a given range. A maximum value for the coefficient

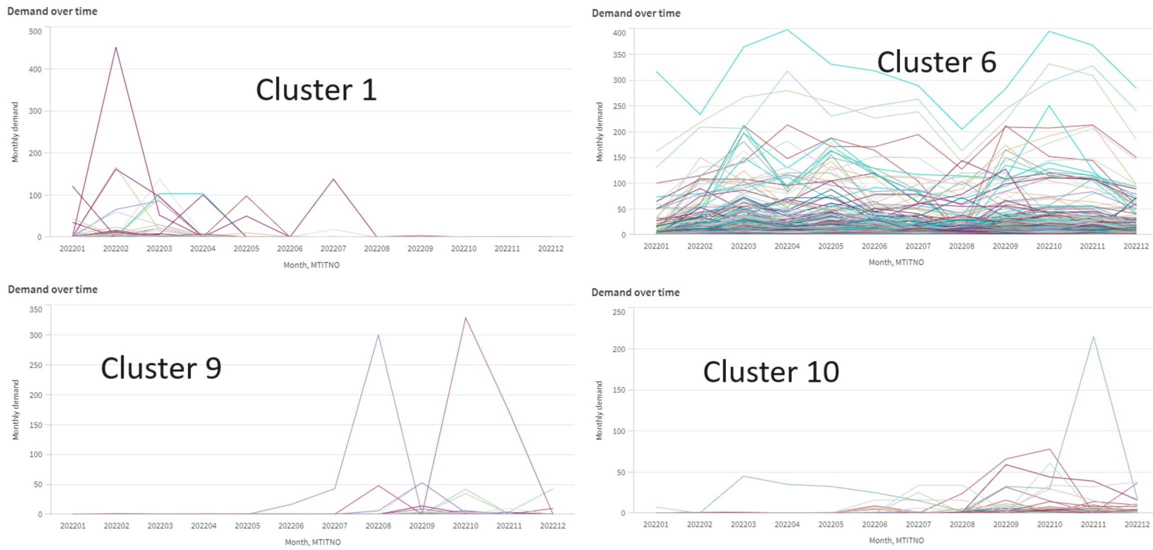


Figure 11: Demand graphs of the sufficiently homogeneous clusters, from cluster analysis 1.

of variation was reached when there was exactly one time period with non-zero demand. Any values outside this range should not be possible with correct data and should be removed, but no such values were encountered. For the trend there was also a range wherein the values possibly could be. The maximum value was when there was only one time period with a value and it was the latest period, while the minimum was when there only existed a value in the earliest time period.

The validity of this cluster analysis is deemed to be neither poor nor good. This conclusion was reached since upon a visual inspection, there is at times no clear separation between the clusters. Additionally, some of them, for example, cluster 5 have items with negative and positive trends which means the homogeneity is quite low for this cluster. This all indicates poor cluster validity. However, upon inspection of the item behavior of all items in one cluster, several of the clusters were homogeneous in terms of that they had similar behavior (see Figure 11). As such, it seems the data is at times not suitable for conducting cluster analysis, as there is a lack of an underlying structure to make good clusters from, but on the other hand, some of the clusters do exhibit good cluster quality.

4.3.2 Cluster Analysis 2

An ABC analysis, as previously described, is usually based on one parameter, most commonly volume value, and has the purpose of allowing differentiation in inventory management strategy depending on how important different articles are to the organization. This has the benefit that more resources can be devoted to the products that are more important to the organization. Another analysis method that has a similar purpose is the XYZ analysis, which groups articles based on their demand variability, or in other terms, ease or difficulty of forecasting demand. If both these methods are combined in a cluster analysis, a classification model with 9 groups instead of 3 can be obtained. This can provide a more complete picture of what items are important to focus on for the organization. This was an idea the authors thought of that could be interesting to try for a cluster analysis.

To conduct such a cluster analysis, the selection of features is more straightforward than in the first cluster analysis, since the visualization resulting from the cluster analysis should serve as the basis for an ABC-XYZ classification. As such, the preferred measure for ABC-XYZ analysis is chosen as the features for the respective axis. The chosen features for this analysis are the volume value on one axis, and the coefficient of variation (χ) on the other axis. The coefficient of variation is given since the previous cluster analysis, but the volume value is not. Feature extraction was conducted by linearly multiplying the volume of an article with its value to obtain the volume value. The resulting cluster analysis is seen in Figure 12.

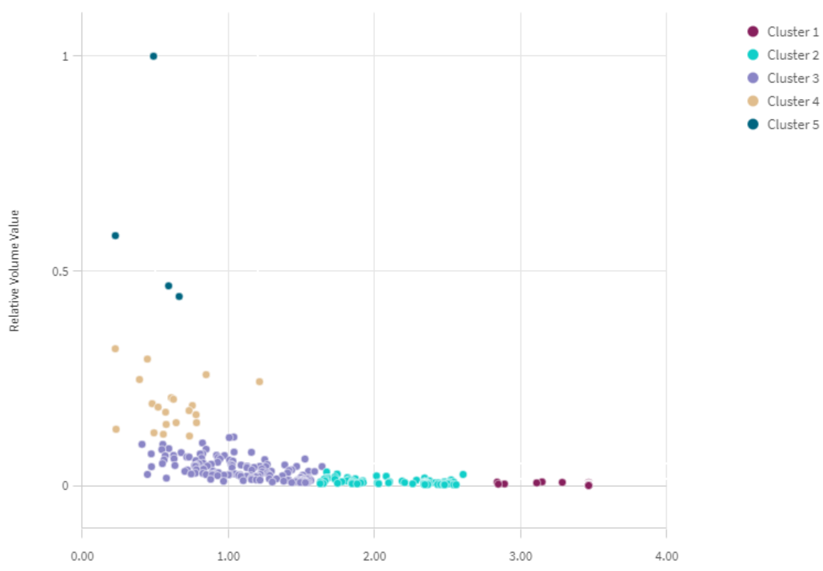


Figure 12: Cluster analysis for ABC-XYZ analysis, volume value versus coefficient of variation. Volume value has been normalized for anonymity reasons.

The number of clusters determined by the DeD method was 2, which is a poor result and probably means the DeD method failed to find a proper solution. Therefore, manual testing of setting different numbers of clusters ensued, before finding the solution with seemingly the best homogeneity and separation out of the solutions tested. This solution was to have 5 clusters, as seen in Figure 12. Further, testing with 4 clusters resulted in clusters 3 and 4 being combined, leading to lower homogeneity while clustering with 6 clusters resulted in the top-most data point being its' own cluster, which could be considered overfitting. Additionally, it should be noted that at this point in the study, the authors noticed that using any set of clusters for an ABC-XYZ analysis would result in a rough-cut classification, as the split between clusters did not align to where theory suggests to split between the classes. Therefore, it was determined to not use the clusters to classify items, and that is also the reason why the solution is not to have 9 clusters. This will be elaborated on in Section 4.4.2.

The Z-score normalization method was used since there are some values that lie far away from where most of the values lie, so they could have been considered outliers. However, since the application for this cluster analysis is an ABC-XYZ analysis, these points should not be removed. Especially those that are far up on the y-axis since they provide a significant amount of the total volume value. This type of situation is when the Z-score method does a good job of normalizing the data. Regarding the validity of this cluster analysis, visual observations say that the clusters seem to be separated well, but that some of the clusters are less homogeneous, notably cluster 5 which covers a wide range of variation amounts, especially the top-most data point is far away from the others. Additionally, upon inspection of an entire cluster's behavior, it was seen that the clusters were to a large extent homogeneous, aside from cluster 5 as previously mentioned. Therefore, this cluster analysis is overall considered to be valid.

4.3.3 Cluster Analysis 3

The third and final cluster analysis has the purpose of visualizing if the company is producing its items in reasonable batch sizes. In general, there is a sweet spot where the associated costs are minimized, known as EOQ, and relies on balancing order costs versus the costs of tied-up capital. In practice, it is often not possible to use the EOQ as machine time is limited, leading to the need to prioritize orders and adjust batch sizes to ensure customer orders are delivered on time. To handle these problems, companies want to produce in batch sizes that are as close to EOQ as possible, while still fully fulfilling the customer demand.

This is also a relevant issue in the case company, especially in a specific production step that will remain anonymized as naming it would possibly reveal the identity of the case company. The production step in question is identified by the case company as being important and is one of the first production steps in their manufacturing process, making it particularly interesting to use in the cluster analysis.

The features chosen for said cluster analysis were the average batch size relative to EOQ, and the EOQ. Note that this is the EOQ as specified in the case company's ERP, and is therefore not necessarily up to date with the most recent information. The motivation behind the average batch size relative to EOQ is relevant as it is generally desirable to have order sizes close to the EOQ for reasons of minimizing costs. The average batch size serves as a good representation of how large the batches usually are. This feature is calculated by taking the average order size of the production orders in this production step for each item and dividing it by each item's EOQ, thus gaining the relative value. Using the EOQ as the other feature was done because it includes an absolute value that gives some indication of magnitude, which can be useful for the analysis. Additionally, the EOQ provides some interesting insights that are elaborated on in Section 4.4.3. The resulting cluster analysis can be seen in Figure 13. Of note is that the Y-axis is in a logarithmic scale, meaning that a y-value of 0 means the average batch size is as large as the EOQ. A y-value of 1 means that the average batch size is 10 times larger than EOQ and a y-value of -1 means the average batch size is 10 times smaller than EOQ. This scale was chosen since it showed both the larger and smaller batches in the same relation to the EOQ. Using a linear scale placed the values that were a third of the size of the EOQ closer to the EOQ line than values that were three times as large.

Another measure that could have an impact on how batch sizes should be determined in production is the set-up time associated with switching products. The set-up time means that some of the machine time available will be unproductive, and encourages producing larger batches as to reduce the amount of unproductive time. Unfortunately, data regarding set-up times was not able to be obtained by the authors, and thus not considered in this study.

To determine an appropriate number of clusters, the DeD method gave 4 clusters, which resulted in clusters that were not homogeneous enough to provide concrete recommendations for decision support. Therefore, the number of clusters was manually increased until 10 clusters was reached, as that seemed to strike a balance between having homogeneous enough clusters to provide concrete recommendations, while not having so many clusters that the validity of the analysis could be in question. For normalization, Z-score was used since there were a couple of values that lie far from the others, so it was suitable for this cluster analysis.

Regarding the cluster analysis' validity, the homogeneity of the clusters is generally good, as they are all fairly small and compact, meaning they are similar in regard to the characteristics relevant for this analysis. However, the separation is at times not as good. For example, between cluster 5 and 6 in Figure 13 there can be seen a couple of

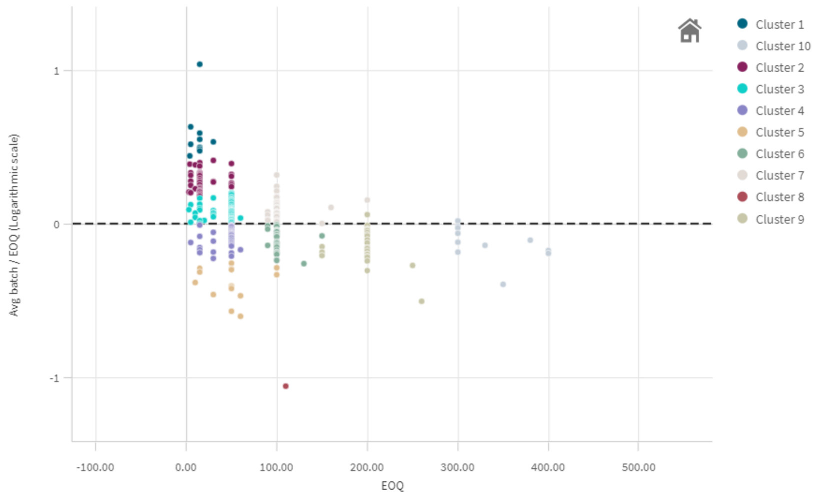


Figure 13: Cluster analysis regarding batch sizes. Average batch size in relation to EOQ versus EOQ.

points at x-value of 100 that belong to cluster 5 but seem more like they should belong to cluster 6. This is a clear sign of poor separation. This problem originates from that the values along the x-axis are consolidated on certain values, for example, 15, 50, 100, and 200. A data set with such characteristics is not very suitable to cluster analysis, and it is understandable that the k-means algorithm can struggle with clustering such data. With that in mind, this cluster analysis has some clusters with questionable validity, while others are valid, meaning that some care needs to be taken when using these clusters as decision support.

4.3.4 Summary of Findings

To summarize the findings from research question 2, a starting point when clustering based on item behavior characteristics should be to have an idea of what behavior is desirable to identify and group items based on, as the features that should be chosen depend on this. One issue to reflect upon is that for any feature that one wants to include in a cluster analysis, it needs to be boiled down into a single number, which is sometimes difficult to do. For example, accurately representing the trend in a single number can be quite hard to do, which is why this study uses quite a trivial measure for it. Cluster analysis can also be applied to a lot of other areas or issues in the supply chain, the ones presented here are simply examples that show how they should generally be performed. It should also be noted that some applications are more suitable for cluster analysis than others. For example, using cluster analysis for ABC-XYZ classification proved not to contribute much as the clusters were not possible to use for classification, so there should be simpler methods of conducting this kind of analysis. It was also found that cluster analysis should be accompanied by a normalization method such as min-max scaling or the Z-score method to ensure equal weight between the features.

The DeD method can be used to select the number of clusters, but it may not produce a desirable result, in which case it needs to be selected manually. The reason why it did not select appropriate number of clusters in these applications were most likely due to the fact that the underlying data was not suitable to perform cluster analysis on, which is common with real data. This is important to keep in mind when carrying out analyses like these, it may not always be the case that the data has a natural structure suitable for cluster analysis. When deciding on the number of clusters it is important to keep the decision support it should represent in mind. If the decision support is to apply specific strategies to certain clusters, then the homogeneity in a cluster needs to be high, which generally indicates that a quite high number of clusters should be chosen. The risk of overfitting should not be discounted, as it increases when using many clusters. However, it is only an issue when the number of clusters approaches the number of items in the clusters analysis.

4.4 Research Question 3

For the third research question to be answered the cluster analysis needed to be performed in a good way to give a valuable basis for decision making. Since an important aspect when conducting the choice of features to the cluster analysis was to consider what the result should show, part of the decision support the cluster analyses provided were speculated beforehand. However, even though there are preconceived suspicions regarding what decision support the cluster analyses should be able to provide, it could go much deeper and potentially to other applications than already suspected. Therefore, the decision support it could provide was analyzed in a qualitative way from the authors own thoughts that were grounded in theory as well as additional findings from the conducted workshop with representatives from Meridion and the case company.

4.4.1 Cluster Analysis 1

The first application was created with the idea that it could serve as a basis for decisions regarding the safety stock for items. What the decision support is based on is where in Figure 10 the respective clusters are. Clusters in different parts of the figure have different characteristics when it comes to the chosen features, trend and coefficient of variation.

In general terms, for the ones with low variation, it entails that the demand is fairly predictable, and hence the safety stock could possibly be reduced. For an item with high variation, it is the other way around, it is hard to predict the demand, so to maintain the service level one solution could be to increase the safety stock. For the items with a strong increasing trend a conclusion is that the trend will most likely continue in the future, therefore it would be good to increase the safety stock to better cope with this. If the item is of strategic nature, this is especially important. For an item with a strong declining trend it could instead be an idea to lower the safety stock to not stock additional items and in that way reduce the risk of obsolescence as the item likely should be phased out. Regarding the level of variation the safety stock formula already covers this aspect, but it does not consider an emerging trend to the same extent. Therefore, if using the safety stock formula for calculating level of safety stock, this cluster analysis could still serve as a basis for determining which clusters need to be analyzed in more depth, where the parameters for safety stock could cope with the trend.

How this cluster analysis can be utilized in practice as decision support entails recommendations regarding safety stock level. Here safety stock level refers to both safety stock as a quantity as well as safety time. An assumption here is that the safety stock formula from Equation 2 is used when calculating the level of safety stock. This also works if using safety time as the only real difference is that the safety stock determined is transformed into days of consumption. For clusters with a strong increasing trend, regardless of level of variation, the recommendation is to update the safety stock parameter to a value that is greater than the proposed level obtained from the formula, thus making sure that it will cover future increases of demand as well. For clusters that have a strong decreasing trend, regardless of level of variation, the safety stock parameter should instead be updated to a lower value than the proposed, in order to not risk excessive inventory of items that would result in large tied-up capital costs. Clusters that have a high variation and no distinct trend, the recommendation is to use the proposed level obtained from the formula as the parameter. It is however important to make sure that the parameter is updated frequently, to ensure that the level of safety stock is enough to cover these variations in demand. For clusters that have low variation alongside no distinct trend, the recommendation is to use the value obtained from the formula. However, if the organization aims at minimizing the tied-up capital, the chosen level for this parameter could be lowered since the demand is fairly predictable. Important to notice, if this is done it is crucial to keep the parameter updated frequently.

In addition, when evaluating what decision-support the cluster analysis can provide, one needs to consider the cluster quality in terms of whether the underlying item behavior of the items in a cluster is homogeneous. If the homogeneity is low in this regard, then decisions based on what the cluster should say can be misleading and potentially harmful since some items will get the wrong treatment. From Figure 10, clusters 1, 6, 9, and 10 fit the criteria of being homogeneous enough to make decisions upon, and their respective demand graphs can be seen in Figure 11. The rest of the clusters either had a lot of varying behaviors or did not show any convincing common trends.

Cluster 1, having a sharp negative trend along with high variation can be interpreted as if its items should be phased out, as all of them have some demand in the beginning of the year, and no demand at all in the latter part of the year. Having only a year of data means that this insight may not be an accurate representation of reality, because of the low sample size. Nonetheless, with the data this study has access to, these items should be phased out. Regarding cluster 6, it has the smallest variation and essentially no trend. This should mean that the items within the cluster can be reliably forecasted with high accuracy, and therefore may be room to reduce the safety stock for those items without impacting the service level. For cluster 9, a slight positive trend along with high variation doesn't say much on its own, but the demand curves of the items within shows large increases and decreases in demand towards the end of the year. This is very high and recent variation, and indicates that the items probably are trending upward, but the recent variation creates some uncertainty around that. Therefore, the recommendation is to increase safety stocks on these items by a little, as if the trend does not continue you do not want to have a lot of extra stock. However, there is quite a bit of uncertainty surrounding the recommendation for this cluster, which can serve as an example of how it can often look in a real case. Finally, with cluster 10 one can see a very high trend along with medium variation. A clear picture of increasing demand can be seen and therefore the items in the clusters should have increased safety stocks for now in anticipation for increasing demand.

To support the decision-making in this phase it would be beneficial to display the current parameter setting (safety stock level) in the Qlik application for the items, as then it is possible to see if adjustments have been made already. This was however not applicable in this study since this parameter did not exist in the data set.

During the workshop, it was brought up how a good way of validating the proposed changes and what effect they would have. One idea was, since this application analyzed data from 2022, to compare the proposed changes to the real demand after the analyzed period and see how correct it would have been. It is not a question of extreme accuracy, but rather if the items in the clusters with proposed actions belonged in the correct cluster. This could for example be measured by if one item with an increasing trend continued to experience increased demand, or if an item that was thought to be stable remained stable in the near future from the analyzed period. However, since the period the data was collected during was relatively short, it was not feasible to divide that data into two different sets to confirm the validity. Additional data were not collected to confirm the validity either. Another point brought up during the workshop was to increase the length of the analyzed time period in order to make sure that the findings were more reliable since the analysis then would have been carried out on a larger data set. Increasing the time period in this way could have made it possible to spot seasonality in the data, that otherwise may have been interpreted as variation or trend. The analyzed period was not extended to cover this due to that the analysis had already been carried out already having been carried out on all available dates in the used data set.

A way to improve this cluster analysis to serve as a better decision support was also brought up during the workshop. The way to do it would be to include an additional variable that measured the volumes or the volume value of the items. Thus, it could be more comprehensible and provide a more accurate representation. The idea of including this variable was founded in that some clusters may have been of so small volumes that their impact on the company would have been negligible. However, due to the application being developed in Qlik, it was not possible to visualize a cluster analysis in 3D. To still take this feature into account in the cluster analysis was tried by having three dimensions in the cluster analysis, while still only visualizing two, but the clusters that appeared were widely spread, resulting in bad quality of the clusters regarding the separation and homogeneity criteria. Due to this, and the fact that the application showed the demand of the articles in respective cluster after choosing it, the original cluster analysis was retained.

This cluster analysis used historical data and another idea, that was raised at the workshop, to further develop the cluster analysis is to in addition use planned orders up to some point. This would allow the planners to base their decisions on future orders that are to be delivered, instead of only analyzing historical data. Combining historical data and future orders would allow the use of the most up-to-date information possible, while retaining data over a longer time period to increase the sample size. Even though the historical data is assessed to be good to make decisions on, including the future orders is deemed a better alternative. Future data was not available in the study although, which is why it is not used.

To summarize the recommendations for cluster analysis 1, Table 5 has been developed. This table shows the recommended actions that can be taken in different scenarios.

Table 5: Observations from cluster analysis 1 and recommended actions.

Observation	Recommended action
For all	Check which clusters are homogeneous enough to make decisions on
Negative trend	Reduce safety stock
Positive trend	Increase safety stock
Low variation	Check if possible to reduce safety stock
High variation	Check if needed to increase safety stock

Firstly, one should always check which clusters are homogeneous enough to make decisions on, as if they are not homogeneous enough then any decision might be misleading and damaging. Then, the other four observations along with recommended actions are presented. These are generalized in the table, and depending on how exactly the analyzed cluster is positioned in the cluster analysis, they should be adapted. For example, if there is a slight negative trend, then the recommendation would be adjusted to reducing the safety stock by a smaller amount.

4.4.2 Cluster Analysis 2

The goal with the second application was to use cluster analysis as an aid in performing an ABC-XYZ analysis. The choice of features were made to display this in the best possible way, and therefore were volume value and coefficient of variation the chosen ones.

The first axis in the cluster analysis represents the volume value, which is the most common decision basis for an ABC analysis. As has been mentioned earlier, the resulting classification commonly follows the Pareto principle, which refers to that roughly 80% of outcomes result from 20% of causes. In this setting, it means that 80% of the volume value comes from 20% of the items, and these are then classified as the A-class. In this study, this was not the case, possibly because not the entire assortment was included in the classification due to the data set only had items in a certain production step, but it was decided to include 20% of the items in the A-class, which made up 66% of the volume value. The B-class included 24% of the items and made up 25% of the volume value, while the C-class made up the remaining 56% of the items and 9% of the volume value.

The other axis in this cluster analysis is the decision basis for a XYZ analysis. To determine the items that belong to X, Y, and Z class respectively the items' coefficient of variation were analyzed. As mentioned in Section 2.1.3 there are two different common ways of doing this. One is by assigning all items with coefficient of variation less or equal to 10% to X, items larger than 10% and lower or equal to 25% to Y, and items larger than 25% to Z. The other is by the same logic but the values between the different classes is instead 50% and 100% or 0.5 and 1, with 0.5 as the limit between X and Y and 1 between Y and Z. In this particular set of data, the first way was not feasible since there were only two values that had a coefficient of variation less than 25%. Hence, if this was chosen the outcome of the XYZ analysis would not be helpful to any large extent. Therefore, 0.5 and 1 was chosen as the limits between classes.

In the developed application with the cluster analysis, there are two supporting variables included that show what share of the items are currently selected and what their share of the total volume value sums up to. This was helpful in determining where the limits between the different classes should be. In order to increase the readability of the cluster analysis and what it actually means reference lines were plotted after the appropriate limits were determined, showing the limits in the cluster analysis, and resulting in nine different "boxes" as seen in Figure 14. The "boxes" are named after which combination of ABC and XYZ classification they show, for example, AX and BZ. These "boxes" are used for visualizing the different combinations of the ABC-XYZ classification. The initial ambition was that the data would be distributed in a way so that the clusters could be directly used to classify the items. This would also satisfy the criteria of homogeneity and separation, which would make the cluster quality good. However, it was quickly realized that this was not the case for the data set. This was because the clusters resulted in a rough-cut classification when in reality one would want to be quite precise with the classification rules. This was the case because the clusters spanned over multiple "boxes", and in one "box" there could be items from more than one cluster. Therefore, it does not make sense to classify the items in this analysis after which cluster they belong to. Instead, they should be classified according to which "box" they belong to.

However, even though the clusters are not suitable for the ABC-XYZ classification, there can be some observations from them. First, cluster 5 comprises the items that have the highest volume value among all and is only spread out over the X and Y classifications in terms of variation. As a consequence of the high volume value, it can be considered that these items are very important. Cluster 4 contains the items that after the previous ones, have the highest volume value and generally low variation with only one item classified as Z. These items are also pretty important since they also have high volume value. Cluster 3 is more spread out and has items in almost all 9 "boxes". Therefore it is hard to say anything for this cluster, except that it is not useful for this classification. Cluster 2 compromises items that have a pretty low volume value, either B- or C-classes, and high variation, all classed as Z. Cluster 1 in turn only has items belonging to the CZ-class, and accordingly, the only cluster that suited the initial ambition of being able to classify all items in a cluster in the same way. These items are however very hard to predict the demand for and do not contribute to any large extent to the total volume value.

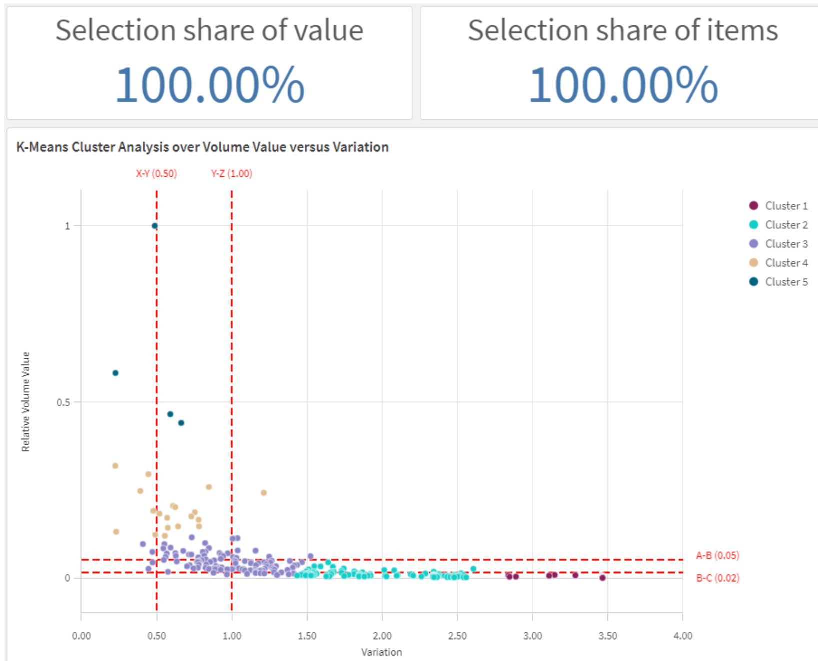


Figure 14: Cluster analysis for ABC-XYZ analysis, with limits dividing the proposed classes, and supporting variables. Volume value has been normalized for anonymity reasons.

Grouping items based on which "box" they belonged to resulted in more manual work than the initial idea but was pretty efficient once it was known how it should be done. Qlik has a built-in tool that lets the user make a selection of the data based on values on the axes. Utilization of this tool made it possible to set limits on the axes corresponding to the reference lines that were already plotted. When these limits were set the selection of items was thus only the items inside the set limits, resulting in items inside one of those nine "boxes". To select the next "box" only one limit was needed to be changed since one of the axes should remain the same while the other only changed one lower limit to instead be the upper limit. For example to change from the selection of AX to BX the limits for the coefficient of variation can remain to still display X while the upper limit for A needs to be changed so it instead corresponds to the lower limit of B. The previously lower limit of A does correspond to the same value as the upper limit of B, hence, it can remain unchanged. The same logic then applies to changing between all nine different "boxes" or combinations, resulting in some manual work needed but once it is understood how it is performed it becomes relatively efficient. For each selection it was possible to extract the data points currently showing, resulting in an Excel file with item numbers that belong in that "box".

Table 6: The resulting number of items in each class obtained from the ABC-XYZ analysis.

	X	Y	Z
A	**AX: 10 items**	**AY: 39 items**	**AZ: 9 items**
B	**BX: 2 items**	**BY: 24 items**	**BZ: 59 items**
C	**CX: 0 items**	**CY: 2 items**	**CZ: 141 items**

After performing this step, the resulting ABC-XYZ classification recommendations can be seen in Table 6. Of note is that there are no items in class CX and only two items in class CY, while class CZ has 141 items. This explains that out of the items that have low enough volume values to be C-items, almost all of them have very intermittent demand that is hard to anticipate. Moving up in the figure, the B-items generally have less variation in its demand compared to C-items, so we see a slight shift toward more items in class BY, but class BZ still has more items. When looking at the A-items, most of them fall into class AY, with an essentially equal number of products in AX as in AZ. This means A-items generally have less variation than the other items. This table also illustrates the quite large amount of variation in demand that the case company has to deal with in general.

The different strategies presented in Section 2.1.3 should be applied to each class of items. Most of them can be taken directly, but some strategies were disputed and need revisiting. AX, AY, and BX items should have low stock levels associated with them as they contribute a large amount of tied-up capital and are possible to accurately forecast. CY and CZ items should have high stock levels associated as they are nearly impossible to forecast, but do not cause too much tied-up capital, so maintaining a high service level is preferred here. The most complex classes from a planning perspective are AZ, BY, and BZ. The strategies applied to these are highly case dependent, and since the case company generally plans their production in accordance with their customers' forecasts, AZ, BY, and BZ items can be produced to orders and not stocked. The reason behind this is that they have fairly high volume value, which could lead to high tied-up capital when considering the low forecast accuracy that accompanies them.

To achieve the differentiation in stock levels desired for the different classes, a couple of recommendations are made with regard to safety stocks and production order sizing. First of all, as the data that is analyzed in this cluster analysis is from a production step early in the production, the resulting items are usually fed into the next production step and are not finished goods. For this reason, the case company has no defined safety stock or safety time on about 75% of the analyzed items. Therefore, recommending them to implement a safety stock on these items would result in a long backlog of production orders that would disrupt the production planning system due to the fact that each item needs to be produced to satisfy the requirement of this safety stock. With this reason, and that the case company has not expressed safety stocks in this production steps as an issue, it is recommended to maintain the safety stocks as they are and not make changes to it.

However, when it comes to production order sizing, recommendations are made regarding the method used, how it is used, the maximum order size, and the order multiple. To ensure that the classes AX, AY, and BX have low stock levels, the lot-for-lot method should be employed, and when planning for individual orders the planner should aim for order sizes slightly lower than the calculated EOQ, as then you increase the operational costs while decreasing the tied-up capital, which is exactly what you want to do with these classes. To help keep the order sizes on the smaller side, the maximum order size should be at EOQ. Regarding the CY and CZ classes which should have high inventory, the recommendations are to use the calculated EOQ to determine order sizes, and only deviate from it in special circumstances. In relation to the previously mentioned strategy, this strategy is considered to lead to higher inventory levels since it leads to more tied up capital, but less costs. The max order size should be larger than the EOQ, as there could be situations where a larger order is needed. Then, regarding the AZ, BY, and BZ classes, the lot-for-lot method should be used, and the strategy should be to have order sizes based strictly on customer orders. For that reason, one might consider removing the maximum order size for these items, but the authors recommend that it is kept, but placed at a fairly high level, as to not risk a massive customer order creating scheduling issues.

Finally, the order multiple has not yet been discussed. The reason is because determining it is the same regardless of inventory strategy. It should simply be divisible from the EOQ of the item. That is, by following the order multiple, it should be possible to land at an order size the same as EOQ. A complicating factor can be if the machines in production have limitations that only allow them to produce batches in certain incrementing sizes. Such data has not been available to the authors, but it is believed that the EOQ should then be adapted to suit the order multiple. The important thing is that the order multiple is divisible from the EOQ.

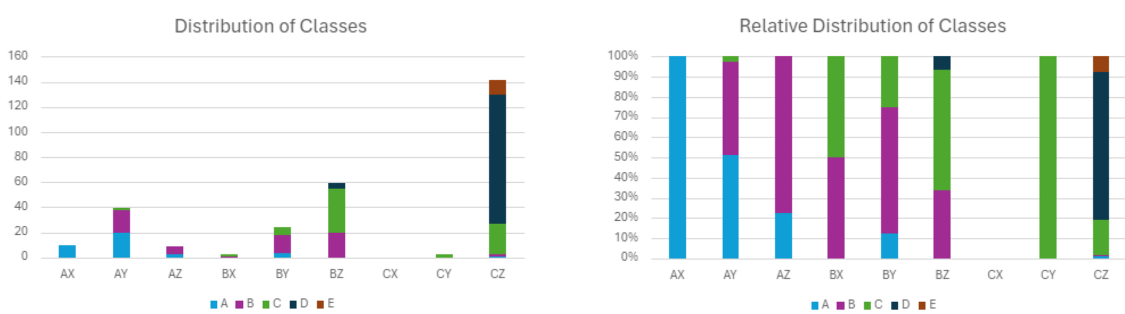
The previous paragraphs described the last step in the process of achieving suitable inventory management strategies for the different ABC-XYZ classes and a summary of the overall process can be seen in Table 7.

A comparison was done between the proposed ABC-XYZ classification and the current

Table 7: Summary of the process for achieving suitable inventory management strategies for ABC-XYZ classification.

Overview of the process for determining inventory management strategies based on ABC-XYZ class
1. Determine limits for ABC & XYZ.
2. Extract classes from the determined limits.
3. Determine if medium or no inventory should be used for the disputed classes based on the specific case.
4. Determine how the level of inventory should be achieved, for example determining order sizes.

classification at the case company. The current classification is an ABCDE classification, and works like a regular ABC classification for volume value, just with two additional categories that make up some share of the total volume value. This comparison can be seen in Figure 15. The left graph shows that the three largest classes by number of items are AY, BZ, and CZ, once again showing that the case company in general has high variation in demand, and that the variation in demand is generally higher for lower volume value items. Additionally, almost all items that currently are in D or E class would be in the class CZ according to the ABC-XYZ classification. The graph on the right of Figure 15 aims to show the reasonability of the proposed classification, as one would expect that classification to largely reflect the current classification, aside from if the current classification is outdated. The figure shows a clear flow from left to right how A items show up first, then mostly B items, further moving to mostly C items before finally the D and E items show up. This provides some indication that the proposed classification could be valid.



(a) (b)

Figure 15: Comparison between the proposed ABC-XYZ classification and the items' current ABCDE classification. The absolute number of items per class is shown in Figure 15a, while Figure 15b shows the relative share of different classes from the current classification.

From the conducted workshop, there was a limited amount of feedback for this cluster analysis, the consensus seemed to be that the ideas presented were mostly reasonable. This makes sense in terms of the choice of features in the cluster analysis as there was not a lot of choices to be made. The method is fairly standard and so are the features used. However, when applying different inventory management principles to the different classes, there was some feedback. Namely, for the classes BY, BZ, and AZ, the recommendation to produce these items lot for lot was received as a good strategy, but should be accompanied by a minimum batch size, as otherwise you risk very small orders being produced, which would lead to a large proportion of production time being lost to setup times, and in turn, higher costs.

4.4.3 Cluster Analysis 3

In the third cluster analysis, the aim was to be able to detect which items are produced in inefficient batch sizes, and to be able to provide recommendations regarding what actions should be taken to improve the efficiency of their production. Therefore, most of the decision support originates from the variable on the y-axis in Figure 13, considering that this variable measures the deviation from the determined EOQ in the system that exists for each item on average. As was mentioned in the previous section, deviating from EOQ will entail lower operational costs or tied-up capital, at the expense of increasing the cost of the other. This can be related to the supply chain triangle, through that if the average batch size is significantly higher than the EOQ, the amount of tied-up capital is high, leading to a focus on lowering operational costs. On the other hand, if the average batch size is significantly lower than the EOQ, then the operational costs are increased in order to focus on reducing the tied-up capital. This way one can use differentiation in batch sizes to establish different strategies based on lowering tied-up capital or operational costs. However, another possibility is that deviating from EOQ is not an intentional strategy, and instead indicates poor production performance, which should be corrected.

To use this cluster analysis for decision support, one would select a cluster that has generally high or low values on the y-axis and then further analyze each of the items in that cluster in more detail before arriving at concrete actions to be implemented. The reason that further analysis is needed is that there are a myriad of potential reasons why the items deviate from EOQ like they do. These reasons could be strategic choices that justify that an item is produced in a certain way, meaning no changes should be made. But there are also other reasons that point to different actions that should be taken. In this section, six reasons along with recommended actions are identified and listed in Table 8.
The first potential reason for deviation from EOQ is that the EOQ in the system is not updated, and therefore does not provide an accurate representation of what batch size is the economic optimum. This is an explanation that should be the first check when analyzing potential reasons for deviations, as if the EOQ is inaccurate then it will not be possible to determine the root cause for the deviation, or if a deviation even exists. The EOQ should then be recalculated and updated in the system. Additionally, it is

Table 8: Reasons for deviation from EOQ in production and recommended actions for each.

Reason for deviation from EOQ	**Recommended action**
EOQ in system is not updated	Recalculate EOQ, match order multiple to new EOQ
Order multiple does not match up with EOQ in system	Recalculate EOQ, match order multiple to new EOQ
Intentionally producing above EOQ to reduce operational costs	No action required
Intentionally producing below EOQ to reduce tied-up capital	No action required
Lack of materials leading to small batches	Investigate root cause
Difficulties with production scheduling leading to deviating batch sizes	Investigate root cause

important to ensure that the order multiple can add up to the EOQ in the system, as if it is not possible to order a batch the size of the EOQ, then the batches will always be inefficient.

The second reason is that the order multiple does not match up with the EOQ, which as just stated means that it will not be possible to order a batch of EOQ size. This reason has the same recommended action as the previous one, so assuming that action has already been taken there is no further action required here.

The third reason is that the company is intentionally producing in batches larger than EOQ in order to reduce operational costs. This could be the case for certain items where the capital requirements are low, but has not been recommended for the case company in this study. For such cases where the batches are larger than EOQ, it is reasonable to verify that the items in question have such a strategy applied to them, for example, if it is a CZ class item such a strategy could be used, but otherwise no action is required.

Similarly, the fourth reason is that the company is intentionally producing in batches smaller than EOQ in order to reduce the amount of tied-up capital. This could be the case for certain items, as is recommended by the authors for some of the classes in Section 4.4.2. For these cases, it is once again reasonable to verify that the items in question have such a strategy applied to them before moving on, but otherwise no action is required.

The fifth reason, identified from the workshop, could be a lack of materials that leads to smaller batches than planned. In such cases, the underlying issue is on a larger scale as it has to do with inbound deliveries, and requires a wider investigation to determine the root cause. Such an investigation will have a wide scope, and the specifics of it also fall outside the scope of this study, and thus no further recommendation is made here.

Finally, the sixth reason is difficulties with production scheduling, which leads to deviating batch sizes. The meaning of this is that sometimes, there are tight deadlines for orders to be delivered which lead to the prioritization of filling orders rather than producing efficient batches. Such issues are deemed to need solving on a higher level than is covered here, as it may involve increasing the available capacity by investing in new machinery, among other things. Therefore it is simply stated that the root cause should be investigated. This reason for deviations from EOQ thus shares similarities with the fifth reason regarding the scope of the solutions needed.

Aside from these six reasons for deviation from EOQ, there were also more general findings from the workshop. Firstly, it was mentioned that the case company had felt like they were often producing in batches that were smaller than EOQ, as opposed to larger than EOQ. But for the data analyzed here, there were more items that on average were produced in batches larger than EOQ than items that on average were produced in batches smaller than EOQ. One potential reason for this that was discussed is that most of the items that had high relative batch sizes were ones with small EOQs, and therefore might not get noticed as much as the items with large batch sizes, thereby being underestimated. Another finding was that some EOQ values occur with high frequency in the data, which is not what one would expect if all the items had calculated EOQs. Coupled with that there seems to be a lot of deviation from the EOQs in general, it seems that the case company does not maintain this parameter enough.

4.4.4 Summary of Findings

In summary, the first cluster analysis could be used to suggest increasing or decreasing the safety stock of items depending on the item behavior. These suggestions were possible to make for some of the clusters, but not all of them since the cluster quality was varying. The second cluster analysis proved useful for performing an ABC-XYZ analysis to classify the items and apply suitable inventory management strategies to all the items analyzed. However, the clusters themselves were not suitable to use for the classification, although the visualization it provided was still insightful. The third cluster analysis showed how large the actual batch sizes were in comparison to EOQ, and could be used to identify groups of items that deviated from EOQ. Six potential reasons for why the deviation from EOQ exists were provided that should be analyzed to determine the action that should be performed.

After conducting these three cluster analyses and investigating the decision support they can provide, it became clear that in the context of performing an ABC-XYZ analysis, cluster analysis itself does not provide much value as decision support, as the clusters were not possible to utilize in a good way. However, in the other two applications, cluster analysis was much more useful. This leads to a more general conclusion about using cluster analysis for these kinds of material planning issues, which is that it is believed that they will fall into two main types of analyses.

The first type is similar to the first cluster analysis presented in this study, where variables that impact the issue at hand are compared in a cluster analysis in order to provide recommendations for parameter updates. Determining the parameter settings is then based on strategies suitable for the characteristics of each cluster. This type of cluster analysis requires a lot of knowledge about the specific case context in order to be able to provide a reliable recommendation. The developed cluster analysis is not fully reliable in its current state but does still serve as good decision support in the form of indicating what different actions should be taken.

The second type is more like the third cluster analysis presented, where there is a comparison between observed performance and parameter settings in the system. This type of analysis has the objective of making parameter updates easier by identifying the items that should be prioritized to be updated. This type of cluster analysis also requires case-specific knowledge, but is more feasible to implement in its current state. This is due to that the parameter updates themselves are determined manually, but the amount of work is reduced through that the clusters show where the largest improvements are likely to be made. With these two types of analyses defined, it should be easier to develop more applications of cluster analysis for different material planning issues in the future.

5 Discussion

In this chapter, a discussion based on the findings and analysis is presented. The discussion relates the findings and analysis to the theoretical framework and tries to bring additional perspectives to light. The disposition of the chapter is that each research question is discussed in order, and the chapter ends with a discussion regarding some ecological and societal aspects that are relevant to this study.

5.1 Research Question 1

The first research question aimed at identifying different characteristics in item behavior. This is something that is not that well anchored in literature, rather, these characteristics were found by discussion with representatives at Meridion and an interview with a representative at the case company, which was then analyzed with the data set from the case company. Hence the characteristics found in this study may not cover all potential item behavior characteristics, as that could differ between companies and industries. However, it is to the authors' uttermost belief that the major ones are brought up in this study. A way of potentially validating this would be to analyze another facility to examine if the same patterns appear there. This facility could either be one of the company's own to validate the patterns for the specific items or it could be another company's in order to validate the patterns themselves.

One characteristic that was mentioned in the study but not included to the same extent as others is seasonality. It is an important characteristic of item behavior and would have been relevant to include in the study. If analyzing data that covered a longer time period it would have been helpful to gain insights if there exists seasonality for items or not. If seasonality exists it could entail that different decisions should be taken regarding parameter settings for specific items during certain periods in time. However, including this longer period of time could also pose a risk that the data that is analyzed is outdated and thus affecting the results in a way that is not desirable. So it would be good to analyze a longer period of time in order to be able to take seasonality into consideration, but the period that should form the basis for the cluster analysis should remain the same to avoid outdated data.

5.2 Research Question 2

As a starting point for the discussion surrounding research question 2, the quality of the cluster analyses is an interesting topic, as a recurring theme through all three cluster analyses performed was that there existed reasons to doubt the validity of said analysis. In the first cluster analysis, the separation between different clusters was at times poor, while in the second cluster analysis, the homogeneity of the clusters was questioned for some clusters. In the third one, the separation was once again poor in some cases. One potential explanation could be that the quality measure is at fault, as although visual observations of the clusters are considered a good quality measure for two dimensional clustering (Han et al., 2021), it should be considered that this method has some level of ambiguity. For this reason, one could argue that a mathematical method for measuring homogeneity and separation could have provided a more accurate representation of the validity due to it definitively being void of any subjectivity.

However, as Hennig (2015) states, when applying cluster analysis to real data, such mathematical methods are only effective as long as the underlying clusters are meaningful. Since only two features are used in this study, it can be visualized to determine the clusters' meaningfulness. Consider cluster 3 from Figure 12, showing the cluster analysis for ABC-XYZ analysis, it looks quite compact, and is decently separated from its nearby clusters, especially cluster 4. Therefore, a mathematical method would likely say that this cluster is of fairly good quality. However, that is not the whole truth, since upon inspecting Figure 14 which shows how the ABC-XYZ classification should be done, it can be seen that items belonging to cluster 3 exist in eight out of nine classifications, where the ninth one had no items in it anyway. It would simply not be possible to use this cluster for the classification, and therefore this cluster is not meaningful and it would be wrong to only have a mathematical method that deems it to be of good quality. Perhaps, using visual observations to determine the meaningfulness before using a mathematical method to objectively determine the quality is the best of both worlds.

Another potential explanation behind the shaky cluster quality is simply that the underlying data is not that suitable for cluster analysis. This is believed to be a more probable explanation because the fundamental purpose of cluster analysis is to uncover a kind of natural structure in the data set (Wierzchoń & Kłopotek, 2018). Similarly, the DeD method is supposed to estimate how many clusters this natural structure has. For example, in Figure 12, the DeD method suggested there to be 2 clusters. This essentially means that it was not able to find a natural structure in the data set. So if there is no natural structure, one could not expect the cluster algorithm to provide a satisfactory result. In such cases where there is no underlying natural structure, the better option was to make the algorithm create many clusters, so that the items within would have similar enough behavior to be used as groups for decision support.

An additional interesting topic is the limitation of only using two features for the cluster analyses. The main argument for this choice has been that it allows for simple visualization of the clusters, which makes it easy for the observer to interpret and understand the results. It is believed that such an approach makes the analysis more suitable for the application it is intended for, which is for material planners to use for parameter updates. However, including additional features that do not correlate with existing features, can improve the quality of the cluster analysis (Aggarwal, 2014). One such additional feature was including the volume value in the first cluster analysis. This was suggested during the workshop in order to identify how important the different items are to the case company. Including that feature would in theory be an improvement for the cluster analysis, as with three features, one could technically still visualize the entire model with a 3D graph. However, this functionality does not exist in Qlik, possibly because 3D graphs have a reputation for being hard to interpret. An additional issue with including more parameters is that the built-in function in Qlik for k-means cluster analysis with more than two parameters does not have normalization included (Qlik, n.d.), unlike the function for two-dimensional cluster analysis. This unfortunately makes such an analysis less accessible as the normalization must be performed manually on the data set beforehand, especially when it is uncertain which normalization method is the most suitable.

In the first cluster analysis, the chosen features of trend and variation could be questioned. This is because they are somewhat correlated to each other, as when there is an increasing or decreasing trend that will to some extent also count as variation. The solution to this issue is to calculate the parameters in a more sophisticated way. This was not done in this study due to time being limited, and that it is not core to the study. The coefficient of variation that was calculated would be a better feature if the trend and seasonality factors were isolated and removed from it so that only the random variation remained. This would get rid of the correlation issue. The trend parameter could also be improved, as in the current state it is quite trivial, and therefore may not be the most accurate representation available. A better trend parameter could for example weigh more recent periods higher than older ones and have a smoothing function. One of the main difficulties of implementing this is that to be used in a cluster analysis, all these factors have to be boiled down to a single value that characterizes the trend.

5.3 Research Question 3

What kind of decision support the cluster analyses can provide can be discussed to a large extent. To start, the developed cluster analyses use historical data and thus the decisions taken are based on past data. A way of improving the analysis could be to instead use future planned orders, which in a manufacturing setting usually is available for the near future. Therefore, the actions taken from the analysis could actually reflect how the future looks which is in accordance with what Bordeleau et al. (2018) mentions that moving towards the use of real-time data characterizes state-of-the-art BI, and is a common topic in the discussion of Industry 4.0. In the application this would be easy to change since the code needed to be changed is only which files should serve as input data, and potentially a couple of variables. Thus it would likely represent different data points and the clusters may look a bit different. Another way of utilizing the developed cluster analysis is to analyze another step in the production process. The same logic behind the analyses could be used, and it could give valuable insights into other parts of the supply chain.

For decision support regarding ABC-XYZ analysis, it can be discussed if cluster analysis is the best available method to determine those classifications. As explained in the study the clusters did not correspond well to the different classes and hence it was needed to perform a manual selection of each class. As it was done in this cluster analysis with the limits between different classes being strict values, it could have been performed in other ways as well and obtain the same results. In this specific case, the method to classify the classes could be more trivial as it for example would be possible to carry it out in Excel. However, as Wierzchoń and Kłopotek (2018) argues cluster analysis is a good way of visualizing and analyzing large data sets, so the visualization aspect will be lost if choosing a more trivial method.

In the second cluster analysis, the choice of limits between the XYZ classes were chosen to be 0.5 and 1. However, both Stojanović and Regodić (2017) and Trubchenko et al. (2020) mention that the limits can be chosen as 0.1 and 0.25 instead. Which one is the correct way is something that can be discussed, but the authors of this study believe that it is dependent on the underlying data and therefore is case-specific. In this case, if the lower classes were chosen there would only be three items that would not have gotten a Z-class. This would mean that the XYZ classification would not add anything of value to the analysis since the items have not been differentiated at all. According to the same logic, one could argue for why the limits not is increased then for a more well-spread distribution between the classes. This is not recommended for this data set, or other data sets, since the items with a coefficient of variation over 1 is already very high, since a coefficient of variation of 1 corresponds to a standard deviation that is equally large as the average demand. Such items are then very hard to forecast so they should be treated as such, and not be put in a class that should be easier to forecast. Doing that would entail difficulties for planning.

Based on the classifications the items get in cluster analysis two, it entails different strategies regarding inventory as discussed by ASCM (2020) and Stojanović and Regodić (2017), visualized in Table 1. Those strategies differentiate between high and low stocks, and relate to the inventory strategy in general and not specifically to safety stocks. As this analysis focuses on a step in production and the demand here is derived from the actual customer orders, everything that is produced should correspond to an actual customer order. Therefore, in this step that comes early in the production process, it is not needed to have safety stocks for many items. In addition, introducing safety stocks here would yield many production orders to earlier steps to produce quantities corresponding to the desired safety stock which would entail a large problem for planning. As the case company has set a target to lower the tied-up capital it is furthermore not recommended to have any safety stock here. However, if the same analysis is done for finished goods storage then safety stocks will be desired to use to handle the variations that could for example arise during the production process. Thus, it is recommended to have safety stocks in the finished goods instead of in each production step since that safety stock could handle variations regardless of in which step the deviation occurs.

In the last cluster analysis, the EOQ obtained from the system plays a vital role. However, when carrying out the analysis in a case where the variation of demand is as high as here it may be problematic since EOQ is based on the assumption that the demand rate is constant, according to Jacobs et al. (2018). Therefore it can be questioned if the EOQ is the best measure to use in this specific case, considering that the high variation of demand can lead to that the real optimal batch size is either lower or higher than the EOQ at a given time. This can cause an issue when comparing the average batch size to the EOQ in a specific case, namely when some months have zero demand. For example, if all the demand for an item occurs in the first six months, then the optimal batch size would be significantly larger than EOQ, and nothing would be produced in the second half of the year. In such a case, the analysis from the last cluster analysis would indicate that the item is produced in poor batch sizes since it would be twice the size of the EOQ. However, this is an edge case, and there is no better alternative measure than comparing the average batch size to EOQ, so this should not have a big impact on the results. Additionally, it is interesting to see that there were signs in this cluster analysis that the EOQs had not been updated frequently enough, which is in line with the problem formulation of this study and Jonsson and Mattsson (2014).

Another issue to discuss briefly is how to take the insights gained through the analysis and use them to update planning parameters in practice. The foundation is that from Qlik, an Excel sheet containing the data from a table can be easily extracted, so the user would extract a sheet containing either recommended ABC-XYZ classifications or a list of items that should be looked over and analyzed a bit further for some purpose. If needed, the additional analysis would be carried out, followed by manually updating the Excel sheet accordingly. Then, one option would be to simply transfer the information from the Excel sheet manually into the company's ERP system, but that would entail a fair amount of manual work, especially for larger volumes. A better option would be to use an external program that allows the user to open tables from the ERP in Excel, make changes, and push them back to the ERP. If the Excel output from Qlik is structured to match up with the format of the ERP table, then updating parameters will be quite simple with this method as they can be copy-pasted in bulk.

5.4 Ecological and Societal Aspects

There are also a few ecological and societal aspects to consider in relation to this study. Firstly, if planning parameters are more frequently updated and maintained, then parameters such as batch sizes will be more accurate, which in turn can reduce the risk of obsolescence from overproduction. This can have a positive effect on the environment as less waste is generated. In addition, having more updated planning parameters can in the bigger picture provide planners with a better overview resulting in fewer situations where, for example, urgent deliveries, are needed. Reducing the number of urgent deliveries, which depending on the severity may otherwise have been delivered by airplane, could provide a reduction in emissions. Such an environment could also be less stressful for the planners, improving their working environment.

A societal factor to consider is the working conditions for employees working the production lines. If parameters are maintained and updated more frequently, the planning environment could likely become more stable in terms of evening out the production requirements throughout the year. Such a stabilization would entail that the amount of man-hours required would remain the same throughout most of the year, so there would be less overtime in some periods, and less need for temporary workers that need to be trained, straining the regular workers. Thereby, working conditions should possibly improve. Another societal factor is one that has been considered throughout the study, and is to what extent the cluster algorithm can be trusted. An initial idea early in the study was to have the cluster analysis feed automatic updates to the planning parameters. However, it was determined that an algorithm should not be trusted to directly make changes without human intervention, as an error could cause obscene amounts of damage, so that even if it is generally more reliable than a human, it still could not be trusted. Instead the proposed solutions are more akin to indicating which parameters should be looked over, and a future development could be to suggest parameter changes.

6 Conclusion and Further Research

The purpose of this study was to identify and utilize different item behavior characteristics in a manufacturing company through cluster analysis. The cluster analysis should provide decision support for operational decision-making in material planning. Based on this purpose, the three research questions were established as follows:

1. What different kinds of item behavior characteristics exist in material planning?
2. How can these characteristics be used in a cluster analysis to identify and group items based on their behavior?
3. What kind of decision support can such an analysis provide in material planning?

The methodology used to answer these questions involved both quantitative and qualitative data collection. The analysis firstly involved preprocessing of the quantitative data, which was then combined with an interview and discussions to identify item behavior characteristics. Three cluster analyses were then developed using k-means clustering, with supporting methods like Z-score or min-max scaling, the DeD method, and a validation method. Finally, a workshop was conducted to combine the authors' own reasoning with input from professionals regarding how the cluster analyses could be used as decision support.

6.1 Conclusion

Initially, transactional data was preprocessed to be used for identification of item behavior characteristics and subsequently cluster analysis. This included mainly aggregating the data per item, and using imputation to handle missing values that appeared after the aggregation. Then, initial discussions were had at Meridion surrounding item behavior and potential item behavior characteristics. The ideas from these discussions were further built upon in an interview with a representative from the case company, where the existing insights were validated and some new ones emerged. Subsequently, demand graphs were visualized in order to see that the proposed item behavior characteristics exist in a real dataset. This resulted in the answer to research question 1, that there are five item behavior characteristics: volume of demand, lumpiness in demand, random variation, trend in demand, and seasonality in demand.

These item behavior characteristics were then combined with each other, or with other parameters, such as item values, in three different cluster analyses in order to group items together based on their similar behavior in certain aspects. Ideas for what parameters to combine as features were thought out by the authors, but for the third cluster analysis, an issue particularly relevant to the case company was chosen. Feature extraction was conducted to establish the desired features, and then they were normalized using either min-max scaling or the Z-score method, depending on which one was more suitable in each case. The number of clusters was suggested by the DeD method, but it did not give satisfactory results in any of the applications developed, so the number of clusters was determined by hand. The number of clusters chosen should generally be quite high in these types of applications, as that allows for more homogeneous clusters to which specific decision support can be applied. To validate the cluster analyses, visual approximation of its homogeneity and separation was used. This methodology can answer research question 2 regarding how item behavior characteristics can be used in cluster analysis. The resulting cluster analyses were: One comparing the variation in demand versus the trend in demand, a second one comparing volume value versus variation in demand, and the final one comparing the average batch size in relation to EOQ versus the EOQ.

To answer the third research question, there was an initial thought behind the features chosen for the developed cluster analyses regarding what decision support they could provide. When they had been developed, further reasoning and analysis was done by the authors, resulting in a concrete idea for each cluster analysis. These ideas were then presented in a workshop with employees from both Meridion and the case company in order to obtain the perspectives of professionals from both consulting and the industry, which brought both new ideas and validated the existing ideas. The resulting answer was that the first cluster analysis could be used to suggest increasing or decreasing the safety stock of items depending on the item behavior. These suggestions were possible to make for some of the clusters, but not all of them due to the varying quality of the clusters. The second cluster analysis proved useful to perform an ABC-XYZ analysis to classify the items, in order to then apply suitable inventory management strategies to all the items analyzed. However, the clusters themselves were not possible to use for said classification, but the visualization it provided was still useful. The third cluster analysis showed how large the actual batch sizes used were in comparison to EOQ, and could be used to identify groups of items that deviated from EOQ, and therefore potentially needed parameter updates or were being produced inefficiently. In total, six reasons for such deviations were provided along with recommended actions.

After conducting these three cluster analyses and finding that in one of them, cluster analysis did not provide much value, it was reasoned that further cluster analyses in material planning would fall into two types that correspond to the first and the third cluster analysis. The first type of analysis is where variables that impact the issue at hand are compared in a cluster analysis to provide recommendations for parameter updates. The second type is where observed performance in an area, for example, batch sizes, is compared to parameter settings in the system. While analyses corresponding to both these types are deemed feasible to develop, the first type requires more work to reach reliable recommendations than the second type of analysis does, but both of them can serve to guide persons who want to conduct these types of cluster analyses in material planning.

6.2 Limitations

There were some limitations that applied to the study which had some effects on the results. Firstly, the available time was limited, so prioritization of central aspects of the study had to be done, which in the end resulted in that some avenues were not explored to their full potential. For one, the trend parameter used in the first cluster analysis was quite simple, and there exists several different ways of calculating a trend variable that were not tested due to this time limitation. Another aspect that could be evaluated with more time is exactly what features should be used for each cluster analysis, as, for example, in cluster analysis 3, the feature on the Y-axis (comparing actual batches to EOQ) contributes more than the one on the X-axis (EOQ). Even though both of them have reasonable contributions, the latter feels like there could be a better parameter to use, if more time was available.

Another limitation was the transactional data available to the authors, as it occasionally took a long time to obtain data from the case company, which could partly be explained by that they were very protective of their data, which is understandable. However, this led to that in the second cluster analysis, the underlying data was not what was desired for that analysis. What would have been desired was a data set with the entire assortment, along with data regarding current classification and item value. Unfortunately, the only data set the authors managed to obtain that had the current classification and item value was only for the specific production step mentioned.

Finally, Qlik proved to be a limitation in a couple of aspects. Firstly, the software only allows for the visualization of two-dimensional graphs, which makes it much harder to visualize a cluster analysis with more than two dimensions in any convincing manner. The reason for not allowing three-dimensional graphs is believed to be that they are considered hard to understand. This essentially made it unfeasible for the study to investigate clustering with three or more dimensions, which is believed could improve how precisely the clusters represent reality. Additionally, clustering in more than two dimensions using Qlik is quite more time-consuming than using two dimensions. Qlik has built-in functions for both options, but unlike the one for two dimensions, the one for more than two dimensions does not allow users to simply pass an argument that determines which normalization method should be used. Instead, the normalization formulas must be written by hand for each of the dimensions, making it even less feasible to study clustering in more than two dimensions.

6.3 Recommendations for Further Research

This study analyzed data from only one facility within the case company. This resulted in that the findings may be specific to the case company, or even the facility, and might not be applicable to other organizations. To make the findings more generalizable it would be beneficial to conduct further studies that analyzes other organizations' facilities, or by analyzing multiple organizations. By analyzing multiple organizations, the possibility to conduct a comparative analysis arises. This broader perspective could contribute to additional findings or strengthen the findings of this study.

Another aspect that would be interesting to consider in further research would be to validate how well the proposed changes actually perform. This could be done by changing the parameters and evaluating how it would have performed given the actual demand. To perform this it would be possible to either look into historical data, propose changes, and compare against more recent data to see if the proposed changes would have been suitable. Alternatively, carrying out the changes in real-time and seeing how the system responds to future demand. Another way of testing this would be to simulate how the system would have performed if the proposed changes were made, which may be a great way to do this since different scenarios can be tested and it will not affect how the business currently operates.

In the study, the developed cluster analyses only had two features that the data was clustered on. An issue was that sometimes the validity, homogeneity, and separation of the clusters were not the best. To improve this it could be of interest to conduct further studies with an increased number of features in a cluster analysis, as there may at times be more that two features that are of interest, and including those should provide a more accurate representation of reality, such as including the volume of demand as an additional dimension in cluster analysis 1. A recommendation would be to utilize a different tool than Qlik for such analyses, as the support for multidimensional clustering leaves some to be desired. A general recommendation for anyone looking to perform clustering with three or more dimensions would be to constantly evaluate how well the reality is represented by the cluster analysis. If there is a missing aspect then it is likely to be good to add another feature, as long as it does not correlate with already existing features.

References

Aggarwal, C. C. (2014). An introduction to cluster analysis. In C. C. Aggarwal & C. K. Reddy (Eds.), *Data clustering: Algorithms and applications*. Chapman; Hall/CRC.

Alelyani, S., Tang, J., & Liu, H. (2014). Feature selection for clustering: A review. In C. C. Aggarwal & C. K. Reddy (Eds.), *Data clustering: Algorithms and applications*. Chapman; Hall/CRC.

Aljuaid, T., & Sasi, S. (2016). Proper imputation techniques for missing values in data sets. *Proceedings of the 2016 International Conference on Data Science and Engineering (ICDSE)*.

ASCM. (2020, October 28). *The xyzs of inventory management*. Association for Supply Chain Management. Retrieved April 4, 2024, from https://www.ascm.org/ascm-insights/the-xyzs-of-inventory-management/

Bell, E., Bryman, A., & Harley, B. (2022). *Business research methods* (6th ed.). Oxford Univeristy Press.

Blomkvist, P., & Hallin, A. (2015). *Method for engineering students: Degree projects using the 4-phase model* (1st ed.). Studentlitteratur AB.

Bordeleau, F.-E., Mosconi, E., & Santa-Eulalia, L. A. D. (2018). Business intelligence in industry 4.0: State of the art and research opportunities. *Proceedings of the 51st Hawaii International Conference on System Sciences*.

Bystedt, J., & Jonsson, P. (2023). State of sweden - en komparativ studie av svenska tillverkande och distribuerande bolag abseende affärssystemsanvändning och best practice för lagerstyrnings och planeringsparametrar. *PLAN-konferensen 2023*, 1–21.

C3.ai. (n.d.). *Glossary - normalization*. C3.ai Inc. Retrieved February 26, 2024, from https://c3.ai/glossary/data-science/normalization/

Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS Quarterly*, *36*(4).

Chubb, H., & Simpson, J. M. (2012). The use of z-scores in paediatric cardiology. *Annals of Pediatric Cardiology*, *5*(2), 179–184.

Denscombe, M. (2018). *Forskningshandboken - för småskaliga forskningsprojekt inom samhällsvetenskaperna*. Studentlitteratur AB.

DeSmet, B. (2017). *Supply chain strategy and financial metrics: The supply chain triangle of service, cost and cash*. Kogan Page.

DeSmet, B. (2021). *The strategy-driven supply chain: Integrating strategy, finance and supply chain for a competitive edge*. Kogan Page.

Donders, A. R. T., van der Heijden, G. J., Stijnen, T., & Moons, K. G. (2006). Review: A gentle introduction to imputation of missing values. *Journal of Clinical Epidemiology*, *59*(10), 1087–1091.

Engels, J. M., & Diehr, P. (2003). Imputation of missing longitudinal data: A comparison of methods. *Journal of Clinical Epidemiology*, *56*(10), 968–976.

Flores, B. E., & Whybark, D. C. (1986). Multiple criteria abc analysis. *International Journal of Operations & Production Management*, *6*(3), 38–46.

Good research practice. (2017). The Swedish Research Council. Retrieved April 5, 2024, from https://www.vr.se/english/analysis/reports/our-reports/2017-08-31-good-research-practice.html

Halkidi, M., Batistakis, Y., & Vazirgiannis, M. (2001). On clustering validation techniques. *Journal of Intelligent Information Systems*, *17*(2), 107–145.

Han, J., Kamber, M., & Pei, J. (2021). *Data mining: Concepts and techniques*. Morgan Kaupmann.

Hautamäki, V., Cherednichenko, S., Kärkkäinen, I., Kinnunen, T., & Fränti, P. (2005). Improving k-means by outlier removal. *Proceedings of the Scandinavian Conference on Image Analysis (SCIA 2005)*.

Hennig, C. (2015). What are the true clusters? *Pattern Recognition Letters*, *64*, 53–62. https://doi.org/https://doi.org/10.1016/j.patrec.2015.04.009

Hermann, M., Pentek, T., & Otto, B. (2015). *Design principles for industrie 4.0 scenarios: A literature review* (Working Paper No. 1). Technische universität dortmund. Dortmund.

Jacobs, F. R., Berry, W. L., Whybark, D. C., & Vollmann, T. E. (2018). *Manufacturing planning and control for supply chain management - the cpim reference* (2nd ed.). McGraw-Hill Education.

Jonsson, P., & Mattsson, S.-A. (2003). The implications of fit between planning environments and manufacturing planning and control methods. *International Journal of Operations & Production Management*.

Jonsson, P., & Mattsson, S.-A. (2009). *Manufacturing planning and control*. McGraw-Hill Education.

Jonsson, P., & Mattsson, S.-A. (2014). Best practice vid lagerstyrning i svensk industri, 5–51.

Lee, J., Kao, H.-a., & Yang, S. (2014). Service innovation and smart analytics for industry 4.0 and big data environment. *NSF I/UCRC Center for Intelligent Maintenance Systems (IMS), University of Cincinnati, Cincinnati*.

Li, J., Cheng, K., Wang, S., Morstatter, F., Trevino, R. P., Tang, J., & Liu, H. (2017). Feature selection: A data perspective. *ACM Computing Surveys*, *50*(6), 1–45.

Lim, E.-P., Chen, H., & Chen, G. (2013). Business intelligence and analytics: Research directions. *ACM Trans. Manage. Inf. Syst.*, *3*(4).

Meridion. (n.d.-a). *Infor m3*. Meridion AB. Retrieved January 31, 2024, from https://meridion.se/infor-m3/

Meridion. (n.d.-b). *Upptäck meridion*. Meridion AB. Retrieved January 31, 2024, from https://meridion.se/om-oss/

Newman, W. R., & Sridharan, V. (1995). Linking manufacturing planning and control to the manufacturing environment. *Integrated Manufacturing Systems*.

Nowotynska, I. (2013). An application of XYZ analysis in company stock management. *Modern Management Review*, *18*(20), 77–86.

Ørngreen, R., & Levinsen, K. (2017). Workshops as a research methodology. *Electronic Journal of E-Learning*, *15*(1), 70–81.

Patel, V. R., & Mehta, R. G. (2011). Impact of outlier removal and normalization approach in modified k-means clustering algorithm. *International Journal of Computer Science Issues*, *8*(2), 331–336.

Patil, C., & Baidari, I. (2019). Estimating the optimal number of klusters k in a dataset using data depth. *Data Science and Engineering*, *4*, 132–140. https://doi.org/https://doi.org/10.1007/s41019-019-0091-y

Qlik. (n.d.). *Kmeansnd - chart function*. QlikTech International AB. Retrieved February 7, 2024, from https://help.qlik.com/en-US/cloud-services/Subsystems/Hub/Content/Sense_Hub/ChartFunctions/RelationalFunctions/KMeansND.htm

QlikTech. (n.d.). *About us*. QlikTech International AB. Retrieved January 31, 2024, from https://www.qlik.com/us

Rabianski, J. S. (2003). Primary and secondary data: Concepts, concerns, errors, and issues. *The Appraisal Journal*, *71*(1), 43–55.

Rahmani, M., Romsdal, A., Sgarbossa, F., Strandhagen, J. O., & Holm, M. (2022). Towards smart production planning and control; a conceptual framework linking planning environment characteristics with the need for smart production planning and control. *Annual Reviews in Control*, *53*.

Rässler, S., Rubin, D. B., & Zell, E. R. (2012). Imputation. *WIREs Computational Statistics*, *5*(1), 20–29.

Romesburg, H. C. (2004). *Cluster analysis for researchers*. Lulu Press.

Säfsten, K., & Gustavsson, M. (2020). *Research methodology for engineers and other problem-solvers* (1st ed.). Studentlitteratur AB.

Schwarz, L. B. (2008). The economic order-quantity (EOQ) model. In D. Chhajed & T. J. Lowe (Eds.), *Building intuition*. Springer.

Sinsomboonthong, S. (2022). Performance comparison of new adjuster min-max with decimal scaling and statistical column normalization methods for artificial neural network classification. *International Journal of Mathematics and Mathematical Sciences*, *2022*, 1–9.

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, *104*(1), 333–339.

Souza, M., Bezerra, J., Bresciani, L., & Silva, E. (2021). Industry 4.0: An analysis of the production of scientific papers based on indicators. *Research, Society and Development*, *10*. https://doi.org/10.33448/rsd-v10i11.18923

Stojanović, M., & Regodić, D. (2017). The significance of the integrated multicriteria abc-xyz method for the inventory management process. *Acta Polytechnica Hungarica*, *14*(5), 29–48.

Trubchenko, T., Kiseleva, E., Loshchilova, M., Dreval, A., Ryzhakina, T., & Shaftelskaya, N. (2020). Application of abc and xyz analysis to inventory optimization at a commercial enterprise. *SHS Web of Conferences*, *80*.

Wierzchoń, S., & Kłopotek, M. (2018). *Modern algorithms of cluster analysis*. Springer International Publishing AG.

Wohlin, C. (2014). Guidelines for snowballing in systematic literature studies and a replication in software engineering. *Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering*.

Ying, X. (2019). An overview of overfitting and its solutions. *Journal of Physics: Conference Series, 1168*, 1–7.

Yusof, E. M. M., Othman, M. S., Omar, Y., & Yusof, A. R. M. (2013). The study on the application of business intelligence in manufacturing: A review. *International Journal of Computer Science Issues*, *10*.

A Interview Guide

	Questions
Introduction	• Presentation of the authors • Presentation of the purpose of the study • Why we have chosen to interview the respondent • Asking for permission to record the interview • Asking the respondent to present themselves
Current State & Context	• How does the production environment look like at your company? • Do you have a way to classify your articles/items? – **Based on product groups or individual products?** • Do you see different item behaviors in your production or are they all similar? • How does the process flow look like in your facility? • How does the product structure look like generally for your products? • We understand different products have varying amount of steps in the production, why is that? • What would you say are the biggest problems in your planning? – **Do you have any idea how they could be solved?** • There are trade-offs between tied-up capital, operational costs and service levels. How does your company relate to those? • What other challenges are there with your current planning environment and processes? • What would you say is the desired to-be state for your planning?
Parameters	• How often do you update planning paramers? – **Why not more frequent/infrequent?** – **How often do you believe is optimal?** • Are all planning parameters updated equally frequent or does it vary? • What planning parameters do you update? • Would you say it is complicated/hard to perform updates? • What information do you need to make an informed decision about updating planning parameters? • What methods do you use to calculate planning parameters? • What has the biggest impact on how you determine what values specific planning parameters should have?
Business Intelligence	• Do you believe you have good enough data to make informed decisions based on? • Do you use a BI-tool in some way? – **How do you use it and for what purpose?** • We described cluster analysis in the introduction, do you have any spontaneous applications you suspect it might be useful for? • Do you have anything you wish that you could use BI for but currently are unable to?
Ethical Considerations	• Would you be able to trust a program to automatically determine and update planning parameters? Without human intervention • What would make such a concept more trustable? • If such updates were possible to completely automate, how would that affect your role?
Conclusion	• Thanking the respondent for their participation • Is there anything that we have not touched upon that you would like to add? • Can we contact you again should more questions arise?

DEPARTMENT OF TECHNOLOGY MANAGEMENT AND ECONOMICS
DIVISION OF SUPPLY AND OPERATIONS MANAGEMENT
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden
www.chalmers.se

CHALMERS
UNIVERSITY OF TECHNOLOGY