UNIVERSITY OF GOTHENBURG
SCHOOL OF BUSINESS, ECONOMICS AND LAW

Master's degree project in Logistics and Transport Management

Implementing data based decision making in logistics processes: Case study at Svenska Mässan

Authors:
Philip Berneblad
Olli Rapanen

Graduate School

Supervisor:
Shahryar Sorooshian

Implementing data based decision making in logistics processes: Case study at Svenska Mässan.

By Philip Berneblad and Olli Rapanen

School of Business, Economics and Law, University of Gothenburg
Vasagatan 1, P.O. Box 610, SE 405 30 Gothenburg, Sweden
Institute of Industrial and Financial Management & Logistics

Contact: philip@berneblad.se; olli.rapanen@gmail.com

Acknowledgement

We would like to thank our supervisor Shahryar Sorooshian for his guidance throughout the research project. We would also like to thank Nicklas Höjer, Director of Logistics and Production at Svenska Mässan for all his help and support as well as all the Svenska Mässan employees who participated in our study and helped and supported us during this project. Lastly we want to thank all the interviewees for taking the time and providing us with their invaluable input.

Gothenburg, May 18, 2020

Philip Berneblad

Olli Rapanen

Abstract

Today utilizing data is crucial for businesses and the logistics sector is no exception. By analyzing and utilizing data, organizations can make better and highly informed decisions. Being able to implement a fluent data collection, analysis and decision-making process, an organization can yield significant competitive advantages. The purpose of this research is to analyze the current traffic data collection system implemented at the case company, Svenska Mässan, to evaluate if it is suitable for decision making in relation to exhibition logistics. Further, the purpose is to find development ideas for the system in order for Svenska Mässan to become more data-driven.

The method used for this research is a qualitative case study utilizing primary data collected through interviews and secondary data mostly from a data set that is carefully analyzed.

This research concludes that the current data collection system should not be utilized for exhibition logistics. The researchers believe that in order for a system to be used, the case company should implement a tracking id that would help them track each vehicle entering and exiting their premises as well as help them to distinguish between different vehicles. Further, it is suggested that the case company should invest in a consolidation center that is solely managed by them, where all cargo would first be handled. This would improve their operations and minimize the current bottlenecks part of their exhibition logistics processes.

Keywords: Data-driven decision making, traffic data collection, exhibition logistics.

Table of Contents

List of tables

List of figures

1. Introduction

1.1 Background

Every operation and process in a company generates data and the data can be sorted into information, both of which are necessary components in decision making. In this day and age utilization of data is crucial for business and for the logistics sector this is no exception. Grabara, Kolcun, & Kot (2014) explains the information exchange and information flows within a company's environment as the "nervous system" and this nervous system is then being referred to as the information system. In logistics and transportation processes, some of the benefits according to Grabara et al. (2014), of using information systems are; improved financial results, better and more efficient coordination in the exchange of information and increased efficiency of transport.

Transportation involves moving goods or people from one place to another, there are different modes of transportation e.g. trains, cars, trucks, vessels etc. all exposed to different rules and regulations. Transportation of goods is important for many companies, whether it is moving goods from a supplier, within a company's internal supply chain or to the end customer. Having a well-functioning transportation system is arguably value adding for companies not only through more efficient transport operations but also through customer satisfaction. Being able to deliver on time increases customer satisfaction which is correlated with customer loyalty and according to Hallowell (1996) these two aspects increase profitability for companies.

Maximizing profits is in the best interest of most companies, therefore, reducing costs without negatively affecting the quality of business would be a satisfactory result. According to Yan and Zhang (2015) costs associated with logistics operations and processes are one of the largest costs for companies, among these maybe the most important is the transportation cost. Transportation costs are heavily correlated with transport management and according to Grabara et al. (2014) transport management is of great importance in logistics. Poorly executed transport management could lead to very high operational costs for companies.

One key element in successful transport management is the use of information technology (IT). The flow of information as one of the main flows in logistics has been part of the definition of logistics since the late 1980s. Easily capturing, analyzing and sharing this information is made possible by IT. The utilization of IT has long been recognized as a way of improving logistics performance both as a resource in logistics and as a competitive tool. Three main tasks of IT are to support business operations, support management decision making and to enable strategic competitive advantages (Closs, Goldsby, & Clinton, 1997; Lewis & Talalayevsky, 2000).

Nowadays it is common knowledge that IT systems are an essential part of any organization and IT has been utilized for decision making for several decades. The same applies for logistics and transport management. For decades modern IT tools have helped organizations to significantly optimize their operations, lower their logistics and transport costs and support managers in making decisions. In modern logistics, the information flow has become as important as the physical flows (Fredholm, 2013). Good IT tools can help in almost all processes of a logistics chain and as an example some of the activities where a transport management system can help are price and transport company selection, booking and booking confirmation, transport documents, transport costs, transport status updates and customer records (Fredholm, 2013).

Transport management systems or any good IT system could also provide an organization with statistics and reports about their performance. As the modern business world is getting increasingly fast phased and managers are required to make more complex decisions faster and almost in real time, presenting and using information derived from data is getting more and more popular. This increases the pressure on businesses to become more agile in their operations and adapt to environmental changes. According to Closs et. al. (1997) IT systems convert data into information to support managers decision making. In the modern world companies are creating vast amounts of data of different sorts in such a great speed that to analyze and turn this data into information requires specific tools and competencies. These tools turn raw data into information by visualizations, reports, alerts and performance measurement indicators that managers and executives can use for making business decisions. This process performed by data analysis professionals with the help of IT tools is called Business Intelligence(BI). By successfully implementing BI in the decision making process an organization can improve its operations and yield significant competitive advantage (Sharda, Delen, & Turban, 2014).

With the optimal use of information technology tools, organizations can not only collect greater amounts of more relevant and real time data, but they can analyze and utilize the data for better and highly informed decision making. Being able to implement this data collection, analysis and decision-making process in an organization can yield significant competitive advantages (Sanders, 2014).

Before data analysis, reports and predictions can be made companies must collect high quality data. There are several methods for collecting data that support operational decision making in logistics and transport, e.g., using sensors or cameras to collect data about vehicles characteristics or general traffic flows. If this kind of data collection methods are used for traffic flow surveillance, the system can capture such data as traffic volumes, vehicle identification data or vehicle speed. This data can then be used for several purposes such as recognizing historical trends, forecasting or planning for future infrastructure investments (Kochlán, Hodon, Cechovic, Kapitulík, & Jurecka, 2014). In transportation, traffic flow data can be used for optimizing transport planning or e.g. real-time adjusting of transport flows according to current traffic situation. These actions can produce competitive advantage by. e.g., increasing service level and making operations more streamlined and profitable.

One example of the above mentioned data collection systems can be found in the case company of this paper Svenska Mässan Gothia Towers AB (from now on referred to as Svenska Mässan). Svenska Mässan has changed and developed as a brand and company since the inauguration just over 100 years ago. Their business strategy has changed and transformed, what once started off as an exhibition company, has today become Europe's largest meeting, exhibition, hotel and conference establishment, everything kept under one roof. As of the beginning of 2020, Svenska Mässan had 1200 hotel rooms including 11 suites, 5 restaurants, 60 meeting and conference rooms, and various exhibition and event halls dedicated for entertainment and exhibitions. The area dedicated for exhibitions and congresses exceeds 41.000 square meters and every year approximately 2 million people visit Svenska Mässan. In 2018 Svenska Mässan organized more than 50 exhibitions with 6572 different companies as exhibitors, attracting approximately 565 000 visitors (Svenska Mässan, 2018). Svenska Mässan has approximately 800 full time employees and during high seasons the number of personnel can reach approximately 1300 (Svenska Mässan, 2020b).

The development of Svenska Mässan has been evident not only for the visual representation and infrastructural changes to the city, the three towers together creating Gothenburg’s skyline, in 2018 Svenska Mässan generated 3,2 billion SEK for the city in revenue associated with tourism (Svenska Mässan, 2018).

1.2 Problem Statement

The years to come, Svenska Mässan are making large infrastructural changes, not only are they building a fourth tower, they are also building a new entrance, and both are expected to be completed in 2025. The fourth tower will be dedicated for office space, hotel rooms, a restaurant and space dedicated for events and conferences (Fastighetsvärlden, 2019). Svenska Mässan also recently decided to decrease their logistic space by 30%, resulting in the same logistics activities performed in less space and this calls for some new innovative solutions. Due to the central location of Svenska Mässan all unloading and loading operations take place in their underground garage. The underground location makes the operations challenging, especially with large items that cannot be handled by a regular forklift but requires a separate lift. According to Höjer, Director Logistics & Production, most of the employees dealing with the planning and carrying out of logistics operations at Svenska Mässan have been working there for many years and therefore most of the work and planning of the operations behind the exhibitions are performed based on experience. However, when losing 30% of the logistic space, working on merit and experience with the same amounts of goods and volumes can be problematic. For Svenska Mässan, this could arguably increase the operational risks, potentially increase costs and decrease customer satisfaction if not adapted to accordingly.

As many companies nowadays, Svenska Mässan has an idea of becoming more data-driven in their operations. As an initial move in this direction, Svenska Mässan in 2019, invested in a new system that collects traffic data to help the decision making process when it comes to the operations. The system is built around magnetic sensors located at the entries and exits, gathering data and information about the inflow and outflow of vehicles arriving at Svenska Mässan for loading, unloading and parking. The logistic operations of Svenska Mässan vary highly, the exhibitions increase the workload and magnitude of the operations significantly and therefore optimizing the logistics operations is required. Without optimization the decrease of space could lead to problems such as delays and queues for which, considering the central location of the exhibition center, there is no space.

1.3 Purpose

The purpose of this study is to analyze the current data collection system implemented at Svenska Mässan, to evaluate if it is suitable for planning of exhibition and logistics operations and to find development ideas for the system based on theory in order for Svenska Mässan to become more data-driven.

1.4 Research question

RQ.1.

Can the traffic data collection system be utilized for exhibition logistics?

RQ.2.

How should the current data collection system and data utilization be developed in order for Svenska Mässan to become more data-driven in logistics operations and planning?

1.5 Delimitations

As the data collection system providing the traffic data analyzed in this research was installed only a year ago, it is very limited in time and there is very little possibility for historical analysis. As some exhibitions only occur every two or three years, some of the exhibitions are not presented in the data at all. This makes finding useful patterns and trends very hard if not impossible. Also, there can be significant differences in the characteristics of cargo between the exhibitions. Some exhibitions receive a lot of very large items, such as the boat exhibition, whereas some receive very small items, such as the book exhibition.

Another delimitation of this research is the irregular cargo flows due to the nature of the exhibition business. Unloading and loading only happens before and after an exhibition and therefore the peak operational volumes are concentrated around the exhibitions. Also, the traffic data collected at Svenska Mässan reflects all vehicles including valet parking for the guests, personal vehicles of the employees and replenishment and maintenance vehicles for the hotel and restaurants. These observations are challenging to exclude from the data set, and it impacts the analysis and results.

2 Literature review

2.1 Logistics in exhibition management

Transportation costs are of great importance within exhibition logistics. According to Wang (2013), compared with a traditional logistics system, the exhibition logistics supply chain is much shorter, since it does not include raw material procurement, processing and production. Exhibition logistics supply chains involve the collection and transportation of the exhibition goods and products, and Wang (2013) argues that the transportation costs of some exhibits exceeds the cost of the exhibition stand as well as the decoration to come with it. In the article by Wang (2013) he discussed the concept of modern international exhibition logistics, a concept that is based on exhibitor firms, both foreign and domestic, who consign their exhibits to a specified transport distributor. This distributor collects, stores and transports the exhibit to the exhibition sites, using economically, fast, safe and convenient modes of transport to make the delivery on time. According to Wang (2013) this concept would compete with the traditional way, where the exhibitors manage the transport themselves.

Other exhibition logistics solutions are highlighted in an article by Zhang (2012) and according to him; there are several design advantages with a well operated exhibition logistics system and its integrated information system. He discussed and evaluated three different logistics systems and supply chains, two of which were focused on exhibition logistics and the most favorable and superior system according to his article was the *Modern Exhibition Logistics Network Model.* What characterized this model was a joint distribution center, where multiple exhibitors can send their goods and a suitable third party logistics company transports the goods to the exhibition site from this distribution center instead of every exhibitor taking care of their own transport. Zhang (2012) argues that this solution would be beneficial for all the involved parties; it would be more cost efficient for the exhibitors, the third party could maximize the transport volumes in the trucks which would decrease the amount of trips, thus reducing emissions and heavy traffic on the roads, and it would also benefit the company organizing the exhibition, not having to deal with as many different individual deliveries which in turn can reduce queues. The second exhibition logistics system and supply chain solution discussed in the article was based around a distribution center owned and organized by the company hosting the exhibition. The exhibition goods and products were consolidated at this distribution center and later transported to the exhibition site. This allowed the exhibition company to be in total control of the delivery flow.

2.2 Becoming data-driven

Data and information are commonly used as a source of support for decision making. Today the term data will guide people to thinking about information technology and the tools it provides for collecting and utilizing data. According to Power (2008) the first system to support data-driven decision making was launched in 1963 for military use in the United States. However, the first system used for the purpose of business was introduced about ten years later in American Airlines. Around the same time in 1971 the term decision support system was introduced and according to Sprague (1980) a decision support system is a system that is aimed at unstructured and unspecified problems that upper level managers face. Further this kind of system combines models and analytics and data access and retrieval functions with features that are easy to use. A decision support system also emphasizes flexibility and adaptability in the user environment (Sprague, 1980).

Since the times of the above definition, decision support systems have become very popular and companies around the world are collecting increasing amounts of data with high levels of details. This is partly due to the development of corporate wide information systems used in modern companies. These systems collect vast amounts of data about an organization's daily operations (Brynjolfsson, Hitt, & Kim, 2011). The increased popularity of these decision support systems has created a widespread expectation in the business world that organizations should be more data-driven and should be taking advantage of vast amounts of data and realizing benefits of this data in their decision making (Bopp, Harmon, & Voida, 2017). Brynjolfsson et. al. (2011) have found in their research that there are real benefits behind the expectations mentioned by Bopp et. al. (2017) and firms utilizing data-driven decision making are performing better in productivity, profitability and market value than those not being data-driven. According to Berndtsson, Forsberg, Stein and Svahn (2018) organizations that have adapted to more data-driven approaches have an increased likelihood to outperform their competitors. Having a data-driven organization will help companies both using and analyzing data in decision making and organization-wide data-driven approach might be a requisite for successful use of analytics (Berndtsson et. al., 2018).

However, it is not easy to become data-driven. Collecting data and using it to create reports and dashboards is not enough. It is important that the data is of the right type, of high quality and that the knowledge of analyzing it exists within an organization to recognize how to utilize the data correctly. The results of the analysis also need to be used in decision making and in concrete actions and in order to make this happen organizations need to have corresponding business processes in place. The key for being successful is having the right organizational culture (Anderson, 2015). Berndtsson et. al. (2018) agree with this view and state that a data-driven culture spanning over an entire organization must be in place to benefit from all the positive aspects related to the data. The authors argue that the solution to a data-driven organizational culture is implementing a decision culture where different possibilities are experimented and tested and where data is trusted more than opinions. Some failures will occur, a fact that is accepted as long as the organization can learn from these failures. The experience will always create some business insight even if real business value is not created (Berndtsson et. al., 2018).

To achieve a data-driven culture and approach Berndtsson et. al. (2018) suggest five important enabling factors; *management, data, tools, organization and decision process*. These are presented below.

Management has a key role in establishing a data-driven culture as they are the initiators and creates the implementation strategy. Organizational changes are normally met with some resistance from employees and it is important that management can handle this accordingly. Employees and middle management can feel threatened by the new working methods and therefore top management needs to be involved in motivating the change to a data-driven culture (Berndtsson et. al., 2018).

It is obvious that *data* is one of the key ingredients in data-driven culture. However, the data cannot just be any data, it must be accessible and of good quality. This is a requirement for performing good analysis of the data which yields insights that can be trusted and utilized. As the amount of data increases and the sources and types of data vary, the right tools for organizing the data become increasingly critical (Berndtsson et. al., 2018).

Data-driven culture needs to be organization wide and therefore all the employees need to have the correct *tools*. Not having the appropriate tools at every level might undermine the acceptance of a new culture and compromise the company's strategy. A company should offer its employees user friendly tools that they can use for making some analysis themselves. This requires users to be trained for using the tools and might put more demand on the availability of appropriate data. However, an organization offering easy access to data for its employees enables ideas to be tested quickly and more and new insights to be created (Berndtsson et. al., 2018).

As an *organization* transforms into a data-driven one it might need to acquire some new IT competence or at least reorganize its current competence. Due to the required data quality and skills needed for analysis the IT department of an organization needs to change its perspectives. The decision on where in an organization to place analytics competence, whether closer to the IT department or closer to actual operations, is every company's individual decision. However, data usage processes might be more effective with an organization wide team making sure that all employees have easy access to data and supporting employees with coaching and education (Berndtsson et. al., 2018).

The last element of the five enablers for data-driven culture is the *decision process*. The decision process is the element that at the end shows how well an organization has become data-driven. When the other four elements are in place the employees are enabled to create valuable experience by testing new insights even if failure sometimes occurs. The new insights created within the organization should never be ignored by senior management. If this happens employees do not feel comfortable with creating and testing new ideas and insights and the company will not achieve a data-driven culture. The same applies when a senior manager has a feeling against the data or newfound insights and decides against what the data is showing. In these scenarios the organization's decision process does not support the data-driven culture and does not trust the data (Berndtsson et. al., 2018).

2.2.1 Data quality and properties

Besides the culture of an organization there are many aspects about data that should be separately considered when talking about becoming a data-driven organization. As using data is the core idea of becoming data-driven an organization needs to be aware of what kind of data to collect and how data is accessed. Data that is collected should not be random and should instead be relevant to the question an organization is trying to answer or a problem they are trying to solve. The data must also be accurate, clean and unbiased and the organization must be able to trust the data. Often raw data takes a lot of cleaning and organizing before it can be used in analysis. This process can be very time consuming and requires specific competence, but it is a requisite for creating high quality data for successful analysis (Anderson, 2015). Low quality data might be useless for an organization and the level of quality reflects the level of which the data can be used (Hazen, Boone, Ezell, & Jones-Farmer, 2014). Hazen et. al. (2014) also states that the use of low quality data can have a direct negative impact on an organization's performance.

Hazen et. al. (2014) and Turner (2004) present elements that should be considered when evaluating the quality of traffic data. These elements are presented in Table 1 below.

Hazen et. al. (2014)	**Turner (2004)**
Accuracy	Accuracy
Timeliness	Timeliness
Completeness	Completeness
Consistency	Coverage
	Accessibility
	Validity

Table 1. Elements of data quality (Hazen et. al., 2014; Turner, 2004)

Hazen et. al. (2014) present four elements to consider with data quality. The first one *Accuracy* refers to the level of how well the data matches real values that it should represent. These real values could be external values that are known to be correct. The next element *Timeliness* refers to how up to date the data are. It can also refer to how often the data is updated and when the last update took place. The third element *Completeness* is the degree of data that is present and not missing. The last element is *Consistency* which refers to how consistent different observations are in format and structure.

Turner (2004) suggests six elements to consider with data quality. Three elements, *Accuracy, Completeness* and *Timeliness* match with the suggestions from Hazen et. al. (2014). The first additional element is *Coverage* which refers to how well a data collection method covers the observed phenomena in time and in areas such as a part of a road. The second additional element is *Accessibility* which refers to how well a user can access and manipulate the data for answering a need. The last additional element is *Validity* which refers to the level how well the data meets the other requirements.

However just thinking about the quality, even if high, does not make a company data-driven and there are other elements to consider as well. Data from different systems that a company has must be compatible with each other to allow for a more unified analysis. This requires specific tools that can comprehend and connect data from different sources. Data also needs to be shareable. This goes hand in hand with the culture of the company. If employees and different departments do not share their data openly in the organization, it is difficult to make comprehensive analysis. The more data available, the better the results. Lastly an organization must possess correct tools for querying the data. When using data for reporting and analysis a lot of filtering, grouping and aggregating needs to be done in order to have a data set that can be comprehended. Once the dataset is comprehensible it can be used for recognizing patterns and trends (Anderson, 2015).

Even though data-driven might feel less human, people are still in the very core of being data-driven. The organization needs to ask the right questions and people with the required skills and competencies need to find the right data to answer these questions. People are also needed at the end of the process to turn the insights provided by data analysis into decisions and actions. Data by itself will not provide the answer to improve a company's performance, not without the employees (Anderson, 2015).

2.2.2 Importance of strategy in data-driven operations

In any business, strategy provides a company with long term goals and a plan that describes how to reach these goals. Strategy defines the direction of the company. Strategy also describes how a company is going to compete and be successful in its business. Strategy is essential for a company to direct its operations and resources correctly towards reaching its goals. Without a strategy a company can be headed in the wrong direction and be wasting resources (Sanders, 2014). The same basic rules apply to being data-driven and using analytics. The data-driven operations of a company should also be directed by an analytics strategy that is in line with the overall company strategy. If a company-wide analytics strategy does not exist, then data collection, sharing and analysis might be driven by inter-departmental small scale strategies not in line with the overall organizational goal (Sanders, 2014). This can harm the transparency and open culture that is needed in order to make comprehensive company-wide analysis (Anderson, 2015). Having a data analytics team in a company might help with operating under a company-wide analytics strategy and creating processes for handling data and its analysis (Sanders, 2014). Sanders (2014) has recognized some companies using a company-wide analytics strategy and states that the functions should be integrated or at least centrally coordinated. Berndtsson et. al. (2018) have found similar results however they emphasize the role of this function as supporting and training all employees of a company in their data analysis activities.

Within the analytics strategy a company should also consider the amount of investments needed for technology and competence when aiming to become data-driven and to utilize data analytics. Depending on the current hardware, software and analytics competence in a company the investment required might be significant. The amount of investment needed should be calculated based on the current systems and the need in the future. Companies should not copy existing analytics strategy solutions from competitors, even though two companies can produce the same product or services, the company competence, structure and strategy can still differ. One factor that can impact the size of an investment is the decision of outsourcing. As the technologies with analytics are constantly evolving and improving, outsourcing some functions and IT infrastructure might be a wise decision. However, outsourcing can make a company dependent on their external partners. The decision is a trade-off question that should be in line with the overall strategy of the company and consider both the short- and long-term vision and needs (Sanders, 2014).

Analytics strategy needs to be considered from a logistics perspective as well. As mentioned before, data needs to be sharable and the data-driven culture of a company should establish a culture of transparently sharing data (Anderson, 2015). However, logistics operations success is often the results of aggregated efforts of many partners and logistics service providers. To create comprehensive analysis of a wider range of operations, data gathered and analyzed from an organization might not be enough. Therefore, data from external partners could also be needed. Also, it might be beneficial to share the company's own data with its logistics or supply chain partners. According to Sanders (2014) this is one of the key success factors of an integrated logistics and supply chain as well as a successful logistics analytics.

2.2.3 Data ethics and security

When a company is striving to become data-driven and being able to utilize data in analysis, that could lead to improved decision making and an increased competitive advantage they also need to consider some data security issues. Mortier, Haddadi, Henderson, McAuley and Crowcroft (2014) presented some elements how personal data should be collected and handled by companies. The authors state that sometimes data is being collected without the user's knowledge or the knowledge of how the collected data is utilized. They suggest that users should be made aware of the ways data is collected and used as well as be given power to control and modify the data that is collected from them. As companies are collecting and processing increasing amounts of data, including personal data, they need to pay attention to security issues in order to not cause harm to their customers, partners or themselves (Sanders, 2014).

Nowadays companies are enforced by legislation to consider data privacy and security as well as required by their customers to be responsible with the data they possess. For example, in May 2018 European Union (EU) put into effect a General Data Protection Regulation (GDPR) that protects the privacy and security of EU citizens. The regulation requires companies that collect and process data of EU citizens to handle this data in a way that the privacy and security of the people is assured. The regulation e.g. requires companies to handle data securely, ask for consent when collecting personal data and be transparent to the people whose data has been collected (GDPR.EU, 2020).

2.3 Business Intelligence

Business intelligence (BI) utilizes information technologies to turn data into information that support organizations' decision making. Different technologies are used, that can gather and collect historical and real time data, organize and analyze the data and present it in a way that can be utilized for decision making in business (Ain, Vaia, DeLone, & Waheed, 2019). With successful use of BI companies can make well informed decisions and gain competitive advantages (Sharda et. al., 2014).

The term business intelligence became popular in the late 1990s as earlier decision support systems started to offer more visualized reports, different alerts and performance measurements based on the data (Sharda et. al., 2014).

The main idea of BI systems is that they collect large amounts of data of different variety that is catalogued, tagged, analyzed, sorted and filtered in order to be stored in a systematic and easily accessible way. The data is stored in a data warehouse that is organized in a way where the data becomes easy and fast to extract and analyze, and that can give an image of business conditions at a certain moment in time. (Sharda et. al., 2014). Users can manipulate the data by aggregating, filtering and drilling down in order to find useful information and patterns for decision making (Ain et. al., 2019). Dashboards can be created for fast creation of different visualizations like graphs and charts as well as extracting different reports and key performance indicators (Ain et. al., 2019). Data is converted from large quantities to high quality (Yeoh & Koronios, 2010). This provides decision makers the possibility to easily analyze different company data, situation, events and performances in order to make better decisions. The general idea of the process of business intelligence is presented in Figure 1.

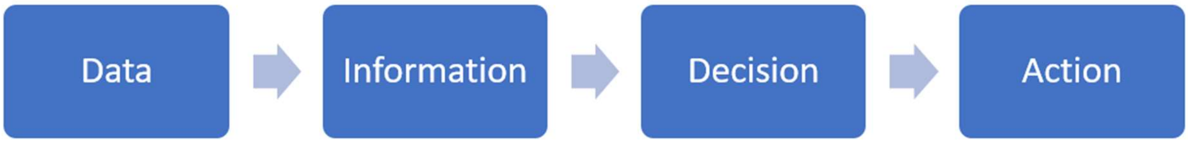

Figure 1. General BI process (Sharda et. al., 2014), developed by the authors.

One essential element of BI that is not visible in Figure 1 is analysis. Analysis is the main step that turns data into decisions and eventually actions. However, analysis with good results is not possible without other supporting steps (Sharda et. al., 2014).

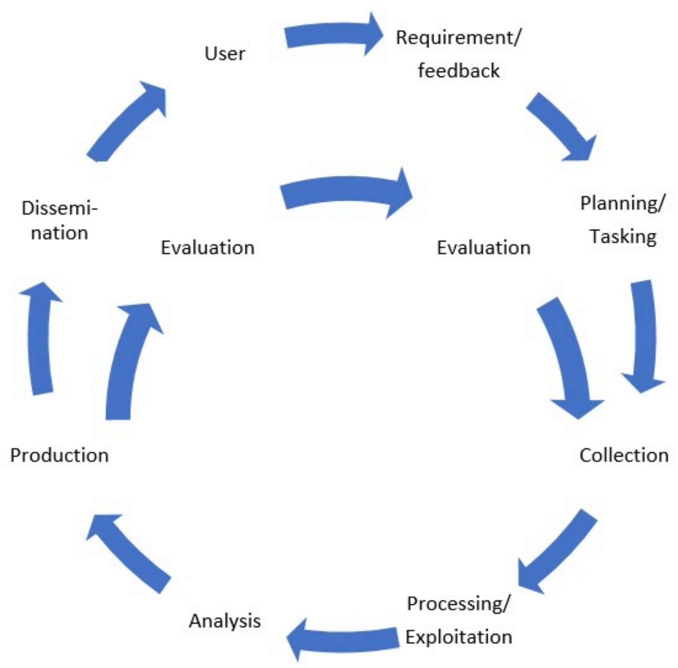

Figure 2. Process of intelligence creation and use. (Krizan, 1999)

As figure 2 presents, the BI process can be quite complex. The importance of a clear process and clear steps becomes even more important as companies increase the amount of data they have. To manage the complexity, companies need to consider a wide range of users for BI and make sure that they possess all the different competencies required for its success. As BI should be planned to benefit the whole organization the importance of having the right people working with it is crucial (Sharda et. al., 2014).

BI investments should only be considered if they align with a company's strategy. If thoroughly implemented it requires a big change in the way a business is operated, and all decision-making processes need to become more data-driven (Sharda et. al., 2014). BI cannot only be an IT endeavor, but it must consider the whole organization, its strategy and culture (Sharda et. al., 2014). Companies and managers must manage both user acceptance and user resistance issues in order to implement BI based processes successfully. With enough training and communication between IT and the user base, BI practices can be implemented as part of the user's routine tasks (Ain et. al., 2019).

There is a difference in adopting BI in a large compared to small and medium sized organizations. Smaller companies have less resources to invest in BI but at the same time they experience lower levels of regulatory influence than larger organizations. Also, the need for BI might depend on the size of the organization and the larger ones might require BI adoption more than small ones due to their size and more complex operations and business environments (Puklavec, Oliveira, & Popovic, 2014).

Puklavec et. al. (2014) have recognized important elements for BI implementation in small and medium sized companies. The most important factors can be divided in two groups: firstly, personnel related factors and secondly organization related factors. In personnel related factors it is important for a successful BI implementation that it has management support. As higher level managers are the ones mostly using the decision support data achieved by BI, they must stand behind the whole project. Also, it is important to have a project champion, a person that drives the usage of BI systems in the organization. As it might be voluntary to use BI and take a long time for results to materialize, it is important to have a project champion who takes initiative and drives the adaptation of BI tools forward. In the organizational side factors such as organizational culture and readiness, organizational data environment and strategic value are important. As small organizations might lack the resources and competencies that BI requires, the factor of having the right culture and the level of readiness in the organization is essential. Related factor is also the data environment of an organization as BI puts a lot of requirements on the IT infrastructure of the company. Further the implemented BI system must have strategic value for the company (Puklavec et. al., 2014).

2.3.1 Data warehousing

One of the core concepts that make business intelligence successful is data warehousing. Traditionally data is scattered around a company in various locations. This makes it very hard and time consuming to use the data for creating reports and useful information to be used in decision making. Data warehouse fixes this problem by aggregating, integrating and organizing data from different sources and forms into one location. The data warehouse will have consistent data that is most relevant and easily available from anywhere in the organization. This way the data can be easily turned into relevant reports, alerts and key performance indicators to support managers and executive’s decision making (Sharda et. al., 2014).

Data warehousing is defined by four characteristics: subject oriented, integrated, time variant and nonvolatile. Subject oriented means that the data in a data warehouse is organized by subject and not by products and transactions. The subject in a data warehouse will only contain data that is relevant for decision making and it can be e.g. sales, customers or products. This way of structuring the data gives the user a better overall view of the whole organization (Sharda et. al., 2014).

The second characteristic, integrated, is related to the subject orientation because the data warehouse must integrate different kinds of data to a consistent form. A data warehouse must make data of different types and formats compatible so that it can be used together under the same subject (Sharda et. al., 2014).

Data warehouse is a time variant. The warehouse includes historical data from different sources and points in time. This data must be organized with the different time points in consideration. The data warehouse aims at identifying different trends, variations and longtime phenomena from the data of different time points. The more real time the system is the more real time information, reports and deviations of events can be extracted for decision making (Sharda et. al., 2014).

Data warehouses are also non-volatile meaning that once data is entered in the data warehouse it cannot be changed or modified. If some data is outdated it is disregarded and changes in the data is recorded as new data (Sharda et. al., 2014).

2.3.2 Analytics

According to Sharda et. al. (2014) the word analytics is sometimes used instead of BI. The core idea is the same, using historical data in generating recommendations to support well informed decisions and actions. Analytics can be divided into three different interconnected levels: descriptive, predictive and prescriptive analytics. Descriptive analytics, also called reporting analytics, is about understanding what is happening in an organization and why by looking at consolidated data in data warehouses. Different techniques and tools are used to turn the data into different reports or alerts. Often the data is presented in a visual form that can give good insight on the operations and performance of an organization (Sharda et. al., 2014). According to Lunsford and Phillips (2018) descriptive analytics looks at current and historical data and company performance and tries to answer questions about what has previously happened in a company and what is happening at the moment.

Predictive analytics is more concentrated in predicting what will happen in the future. According to Lunsford and Phillips (2018) predictive analytics is based in historical data and used for making forecasts about future events. Statistical techniques and data mining are utilized to forecast events that are likely to happen e.g. in customer behavior. As a result, an organization can predict what kind of sales in the future will give them the most profit and attract most customers. Algorithms exist that can predict likely future events based on current and historical customer behavior and suggest actions directly to the customer or to the business serving the customer (Sharda et. al., 2014).

Prescriptive analytics is a combination of descriptive and predictive analytics and it strives to understand the current situation as well as predicting best actions for the future. Prescriptive analytics aims at recognizing the necessary actions to reach a specific result (Lunsford & Phillips, 2018).

2.4 Presenting the data

Different techniques and tools are used in presenting the findings of the data analysis and in turning the data into different reports and alerts. Often the data is presented in a visual form that can give good insight on the operations and performance of an organization. This is an essential part of effective business intelligence and data analytics (Sharda et. al., 2014). Without easily comprehensible data presentation the outcome might not be useful for decision making.

Reports as a tool for presenting business data have been used for a long time. Traditional reports can be printouts communicating information about a certain phenomenon. With the help of modern business intelligence tools presenting data and information has become growingly visual. The goal of data or information visualization is to present data in a form that is easy to make sense of (Sharda et. al., 2014). Comparing visually presented data to a traditional report with numbers and text a visualization can make it much easier to understand a phenomenon and find patterns in it. According to Wexler, Shaffer and Cotgreave (2017) remembering numbers from a report and recognizing patterns in them can be an impossible task. However, visualizing the same data can make this process much easier. The authors say that the core idea of visualizing information is aggregating larger amounts of data and presenting it in order to gain insights.

One very popular way of visualizing information is using dashboards. Wexler et. al. (2017, pp.xiv) define dashboard as "..a visual display of data used to monitor conditions and/or facilitate understanding." According to Sharda et. al., (2014) dashboards are common in probably all modern business intelligence tools. Dashboard is an aggregation of charts and different key performance indicators (KPI) on a single page view that provides integral information of the performance of a certain phenomenon within an organization. The goal of dashboards is to give a quick idea of the state of the organization to support managers and executives in decision making. A dashboard can use any kinds of charts possible and often utilizes colors and maps to improve understanding. A dashboard is also often interactive where the user can click on specific parts of the dashboards to drill down and get more detailed information about a certain part or element of a phenomenon (Sharda et. al., 2014).

According to Karami, Langarizadeh and Fatehi (2017) the goals of a dashboard should be based on users' expectations and needs. The authors state that as dashboards are created to serve the users the user experience should be at the core of designing a dashboard and they should give information about both positive and negative sides of attributes affecting a decision. Therefore, improvements in dashboards can be reached by considering user feedback. Sharda et. al., (2014) agree with the importance of user experience of dashboards. Not considering the user can lead into low levels of utilization and resistance of use. The authors also recommend providing the possibility for real-time user comments while they are using the dashboard. This can help presenting information on the dashboard that improves the users understanding of context and help presenting the correct KPI's. Sharda et. al., (2014) also recommend linking the dashboard with performance alerts. If an abnormal pattern is recognized in real-time data this information can be pushed into a dashboard. This way essential information can reach the user without the user having to look for the information.

When it comes to the challenges of data or information visualization the starting point is the same as with any data utilization, the quality of the data. According to Sharda et. al., (2014) it is very important to consider the quality of the data used in creating the visualizations. If the data is not reliable, if some parts are missing or it is outdated, a very good visualization will not save it. Another factor that is important to consider, especially when creating dashboards, is the use of correct presentation methods. According to Wexler et. al. (2017) different information can be either easy or very hard to capture and understand depending on the choice of chart. As there are various types of charts that can be used to present the same information, it is important to consider not only the graphical look of the chart but also the ability of that specific chart to clearly present the information and trends that are meant to be presented.

2.5 Traffic data collection systems

Today globalization and diversified product ranges are creating new challenges for logistics and transport systems and requires these systems to become increasingly smart and intelligent. This system intelligence can be achieved by more fluent, digitalized and automated processes and information flows. This requires automated collection of data, sometimes in real time, from the logistics environments and its elements. Data can be collected automatically by the use of sensors or cameras attached on logistics infrastructure or on logistics objects moving through and within the system (Borstell, 2018). However, the need for automated data collection is not limited to the logistics environment, the general increase of traffic numbers is creating a need for traffic management. This management can be made efficient and real time controlled by data collection via automated methods (Jain, Saini, & Preeti, 2019).

According to Jain et. al. (2019) traffic monitoring or data collection systems are categorized into intrusive and non-intrusive solutions. Intrusive solutions are physically located on or under the road surface and installing these solutions requires disrupting the traffic flow. Also, in the case of maintenance the traffic would once again have to be disrupted. Non-intrusive solutions are located above or on the side of the road surface on existing infrastructure. The installation and maintenance of these solutions can be made with none or very little disturbance of traffic. Bottero, Dalla Chiara and Deflorio (2013) also present a third category, off-road solutions such as mobile devices used for monitoring traffic and collecting data. These solutions do not require solid installation and pose no disturbance to traffic.

This section will discuss two common automatic traffic data collection systems. The first covered method is data collection by sensors used e.g. for calculating traffic flows. The second method is image capturing systems, used for vehicle recognition, often utilizing cameras. The distinction between the two methods is the following; sensors are solutions using e.g., magnetism, instead of images and cameras are solutions capturing and processing images for data collection and analysis.

2.5.1 Sensor systems for traffic data collection

Sensors are electronic devices used for monitoring and measuring changes in their environment and turning the observations into usable data (Pandey, & Mishra, 2019). Using sensors for traffic data collection is a good way of collecting such data as quantities of vehicles, classification of vehicles and vehicle speed (Taghvaeeyan, & Rajamani, 2014). Sensor technologies might require installation under the roadway and might be disrupting regular traffic flows. According to Jain et. al. (2019) magnetic sensors are one example of such intrusive technology. Other examples of sensor technologies include ultrasonic sensors, temperature sensors, acceleration sensors and radar sensors. According to Borstell and Reggelin (2019) these are commonly used in logistics solutions.

Typically, magnetic sensors are placed under the road pavement of the lane that will be observed. The sensor observes changes in the earth's magnetic field when a vehicle drives over it and obtains a magnetic profile of the vehicle. Based on the observation the system calculates and categorizes the vehicle based on previously set characteristics of each different variable collected. This data is then transmitted to a traffic management system that can be local or remote. When using multiple sensors such variables as speed and length of a vehicle can be observed. The lower the speed of the observed vehicles the closer together the two sensors can be installed (Bottero et. al., 2013).

Taghvaeeyan and Rajamani (2014) have proposed an alternative non-intrusive system to reduce traffic disturbance during installation and maintenance. This system is based on the same technology, but the magnetic sensors are placed on the side of the road and observe the changes in magnetic waves horizontally rather than vertically. This makes the system significantly cheaper to install and maintain, however it possesses challenges such as disruptive magnetic waves detected from heavier vehicles driving on a non-adjacent lane. Similar research has been conducted by Wang, Zheng, Xu, Xu and Chen (2018) with a slightly different sensor set-up due to their research goal being only to detect vehicles and not measure speed. However, they find similar results with lower costs due to cheaper installation and maintenance and disruptive observations from heavy traffic on non-adjacent lanes. They also conclude that the performance of their sensor system might be lower during traffic congestion due to very low distance between single vehicles, which might make it hard to separate vehicles from one another by a magnetic image.

When it comes to the accuracy of magnetic sensor systems, different data collection methods researched have reached varying results. Cheung, Coleri, Dundar, Ganesh, Tan and Varaiya (2005) found in their research from a single sensor that vehicle count accuracy can be as high as 99% and vehicle speed and length measurements can have accuracy over 90%. Their research suggests that a single sensor can classify vehicles with 60% accuracy and two sensors with 80% accuracy. Taghvaeeyan and Rajamani (2014) tested non-intrusive portable sensors in their research and concluded that the system managed to measure vehicle speed with an accuracy of 97,5% and counted vehicles with an accuracy of 95%. Bottero et. al. (2013) tested magnetic sensors in a logistics village and both vehicle classification and vehicle counting reached high accuracy levels, 91% and 94% respectively.

All traffic data collection methods come with some characteristic advantages and disadvantages. According to Taghvaeeyan and Rajamani (2014) the general weakness of a magnetic sensor is its inability to separate small and medium sized vehicles from each other. These vehicles are often put in the same category when the data collection system characterizes the vehicles on the magnetic observations of the sensor. Further Bottero et. al. (2013) observed that vehicle classification results with a magnetic sensor might vary especially with heavy cargo vehicles where type and quantity of cargo might affect the magnetic image of the vehicle. Jain et. al. (2019) have presented a list of advantages and disadvantages of magnetic sensor technology used in traffic surveillance systems. These are presented in Table 2.

Advantage	Disadvantage
Not very sensitive towards strong traffic	Installation under pavement required
Not affected by weather changes	Installation and maintenance possess disturbance to traffic
Possibility for wireless data transfer	Small units might require several devices for detecting the whole area of observation
Provides limited amount of information	Might not detect vehicles standing still

Table 2. Advantages and disadvantages of a magnetic sensor based traffic surveillance system. (Jain et. al., 2019)

2.5.2 Application of magnetic sensor based methods in logistics

Magnetic sensors are widely used in different applications of traffic management and surveillance. As traffic numbers grow they create issues such as congestion, accidents and environmental problems. To manage these problems an intelligent transport system can be utilized (Karpis, 2013; Wang et. al., 2018). The use of these systems can lead to both economic and environmental savings as transport becomes more efficient, safe and environmentally friendly (Karpis, 2013). One way of collecting required traffic data for intelligent transport systems is using magnetic sensors. Bottero et. al. (2013) suggest that data collected by the sensors can be used for signal control and traffic monitoring which could include applications such as monitoring traffic flows, adaptive traffic control and supervising traffic on a wider area and could help in solving issues such as congestion.

Data that a magnetic sensor often collects is related to quantities of passing vehicles, vehicle characteristics and vehicle movement characteristics like speed (Taghvaeeyan, & Rajamani, 2014). This sort of information can be used for future planning of infrastructure or short term transportation decisions and could be utilized in a SmartCity concept (Zarnescu, Ungurelu, Iordache, Secere, & Spoiala, 2017).

Sifuentes, Casas and Pallas-Areny (2011) suggest a method using magnetic sensors for detecting idle cars for several purposes including detecting empty parking places, parking meters, automatically opening a door or a gate, traffic control and railroad crossing control.

2.5.3 Camera systems for traffic data collection

Video cameras can be a great source of information for traffic surveillance both in public road networks as well as private areas such as logistics facilities, and the popularity of using them has grown strongly in recent years. The development of processing and analysis of camera collected data has further increased the popularity of camera based traffic surveillance (Al-Smadi, Abdulrahim, & Abdul Salam, 2016).

The environment where the cameras operate poses a lot of challenges to the data collection. The different specifications of vehicles such as size make vehicle recognition complicated.

Especially in urban areas, where infrastructure is more tense and traffic solutions are not necessarily so straight forward, the surroundings and the positioning and orientation of the vehicle bring in further challenges (Al-Smadi et. al., 2016).

The starting point of the technology behind video camera based traffic data collection is detecting the vehicle from the video camera image. There are two major techniques to detect vehicles from the image, *motion based* technique and *appearance based* techniques, both of which have several different solutions for vehicle detection. The *Motion based* technique has its origin within frame differencing where consecutive frames of the camera image are compared, on background subtraction where current frames are compared with a default frame or on a more complicated method optical flow where flow vectors are used to recognize moving regions in a video image. In the second major vehicle detection technique *appearance based* technique, historical vehicle characteristics are used in vehicle detection. This can mean feature based technique where previously coded descriptions of vehicles, such as symmetry, are used to characterize vehicles. Other appearance based techniques are; part based detection where the relation between different vehicle parts e.g. windows or tires are compared or three dimensional modeling technique where current image is compared with previously generated 3D models of different vehicles (Al-Smadi et. al., 2016).

After a moving object has been detected from a video camera image it needs to be recognized in detail and classified to separate different types of vehicles. Vehicle recognition software uses different vehicle characteristics such as color, logo, type of vehicle or license plate. Significant challenge for color recognition is lighting. Lighting can distract the recognition of differences in colors and therefore most color recognition systems only consider main color groups to recognize. Detecting and recognizing vehicle logos can give important information about a vehicle model, which can have a significant role in classifying and identifying different types of vehicles. Vehicle type recognition is based on vehicle shape and appearance. Different solutions can recognize a variety of different vehicle types however, the high number of different types is a great challenge for a successful classification (Al-Smadi et. al., 2016).

As mentioned above there are factors and challenges affecting the collection of traffic data by camera application and all the different types and methods possess their own strengths and weaknesses. Jain et. al. (2019) have presented a list of advantages and disadvantages of camera and image capturing based technology used in traffic surveillance. These are presented in Table 3.

Advantage	Disadvantage
Possible to monitor multiple lanes/zones.	Installation and maintenance might require traffic to be stopped or partially disrupted.
Easy to add and modify detection zones.	Performance affected by weather conditions by decreasing or blocking visibility. Vehicle shadows disturbing detection.
Large variety of data available	Requires 9 to 15 meter camera mounting height for optimal data collection.
Provides wide-area detection when information gathered by one camera can be linked to another	Wind or vibration can cause camera motion that decreases accuracy.
Cost-effective when many detection zones within the field-of-view of the camera or specialized data are required.	High level of accuracy during nighttime requires lighting.

Table 3. Advantages and disadvantages of a camera based traffic surveillance system. (Jain et. al., 2019)

2.5.4 Application of camera based methods in logistics

Modern logistics systems are faced with many challenges such as instability, customer service requirement, short throughput times and need for flexibility. One solution to face these issues is an image-based real time state detection system for logistics operations. This system collects data on the current state of the logistics operations with the help of image capturing e.g. a camera. There are many types of logistic and transport operations that could benefit from these types of systems and methods. The information it can obtain can be used for e.g. inventory planning and warehouse management. Many different logistics solutions during the last decade have utilized camera and image capturing based solutions and examples of these solutions are presented below (Borstell & Reggelin, 2019).

One method of tracking cargo by utilizing image capturing that is very common is reading optical codes such as a barcode. Barcodes and other optical codes are used widely in logistics processes in tracking inbound and outbound cargo flows. This method is utilized all the way down to a single unit level and gives information that can be used for numerous logistical applications. However, barcode reading is based on one dimensional image. More sophisticated image capturing methods are capable of recognizing two dimension codes or character codes with such technology as OCR. This method is utilized in tracking full trucks by reading their license plates. This way e.g. a delivery truck location can be found and used for tracking the cargo it is carrying. This of course requires information about the cargo loaded on the truck (Borstell, 2018).

Another method utilizing image capturing and OCR is automatic container number recognition. Container numbers are used to track the container movement at several steps during container transport, and often the number is still recognized and recorded manually (Verma, Sharma, Hebbalaguppe, Hassan, & Vig, 2016; Roeksukrungrueang, Kusonthammrat, Kunapronsujarit, Nunjaruwong, & Chivapreecha, 2018). This manual method has several disadvantages, e.g. the possibility of errors and high time consumption (Roeksukrungrueang et. al., 2018). Implementing a functional camera system for automatic container number recognition is a good solution for this issue but possesses several challenges. The container number is controlled by a standard, but it can still be in different formats and located on different parts of the container. Also, the container surface being uneven, and the possibility of dirt, rust or other disturbances pose further challenges for optical character recognition (Verma et. al., 2016). Especially the variety of locations for the container number, differentiates the automatic container number recognition technologies from e.g. license plate recognition. Several different methods to overcome the challenges have been proposed. One of the methods, presented by Wu, Guo and Li (2019), uses a process where it captures an image of every side of the container to position the container correctly. It preprocesses the image to decrease disturbance, locates the container number and identifies the number.

Besides being able to track cargo by barcodes, there are other beneficial image based methods that could support inventory planning and warehouse management. An additional method is using a camera based real-time positioning system that can track forklifts and other moving objects in a warehouse to improve safety and efficiency. The object recognition can be based on the object's characteristics or separate markers and when built with quality, the system can be very accurate. However, similar challenges as with any camera based system apply, lighting must be good and the camera cannot be blocked or displaced (Halawa, Dauod, Lee, Li, & Yoon, 2019). Another application of camera system in a warehouse is detecting free storage areas or slots. This could help in detecting critical situations where an out of stock situation might occur or help in optimizing operations. Similar technology could be used in other space detection tasks as well. Parking lot or loading space availability could be monitored with image capturing. Detecting when a loading dock is available can help in logistics management and to further optimize operations (Borstell, 2018).

Camera based systems are widely used in traffic surveillance e.g. detecting road incidents, analyzing vehicle behavior and understanding traffic patterns (Al-Smadi et. al., 2016). The technology can also be used for minimizing congestion by using real time data in traffic management. Being able to do this leads to a decreasing number of road accidents (Jain et. al., 2019). Payments like parking and road tolls can be automated with a camera based traffic surveillance system or access to certain locations can be based on a camera surveillance system (Tiwari, Sharma, Singh, & Rathi, 2016).

When an organization is trying to optimize its logistics operations, simulating a real situation can be beneficial. If the data as well as the collection method represents reality it will allow results from scenario tests to be more accurate and give better support for logistics planning and decision making (Borstell, 2018).

2.5.5 License plate recognition

License plate recognition (LPR) is one applicable technique for collecting traffic data by a camera system. Having access to license plate information, the data collection method can move from collecting general anonymous traffic flow information to gathering individual data assigned to each and every vehicle. According to Bakhtan, Abdullah and Rahman (2016) LPR is a four step process including capturing and preprocessing an image, license plate localization, character segmentation and license plate number recognition. The same process is recognized by multiple authors (Gupta, Tayal, Gupta, Goyal, & Goyal, 2017; Puranic, Deepak, & Umadevi, 2016; Mani & Panicker, 2019). In the first step of the process, an image of the vehicle is captured and preprocessed for making the next steps easier. The image is then turned into a greyscale image and further to binary image. In the second step, license plate localization, the license plate is located from the picture and the background, such as the car frame, is excluded from the image leaving only the license plate for the further steps. In the third step the characters of the license plate are segmented, separating different characters from each other. A common technique to use in character segmentation is optical character recognition (OCR). This step is very important because it highly affects the outcome of the LPR process. The final step in the LPR process is character recognition where the final results are achieved, the license plate numbers and letters are turned into a usable string of characters. A variety of techniques for LPR exist but they generally follow the same process. For example, edge detection has been suggested for the preprocessing step of the process in order to isolate the license plate area from the background of the image (Bakhtan et. al., 2016; Gupta et. al., 2017; Puranic et. al., 2016; Mani & Panicker, 2019).

There are several challenges associated with LPR that can be due to technological challenges, changes in environment and changes in the license plate. One technological challenge is the resolution used for the captured image. According to Mani and Panicker (2019) more details on a license plate can be detected from an image of higher resolution. However, this is a trade-off with costs as the computing of images of higher resolution is more demanding. Environmental challenges are mostly due to lighting and background. Lighting may vary due to the time of day, infrastructure e.g. street lights or a vehicle's own lights. Background may contain characters or shapes similar to the license plate making the recognition process harder. Challenges due to the changes in the license plate can include issues such as the location of the plate, size or color of the plate, font used in the plate, damage on the plate, unreadability due to dirt or other elements such as screws or frames. Many of the differences in license plates have evolved through country specific standards and formats (Bakhtan et. al., 2016). As the accuracy of different LPR methods vary, so do the challenges. According to Mani and Panicker (2019) the strengths and weaknesses vary depending on the environment they were tested in. Research is often done under certain conditions and might therefore lack support for changes in lighting or background. More recent methods have increased complexity to overcome different challenges but are still struggling to process complex images. This emphasizes the significance of environment and external effects on LPR. Bakhtan et. al. (2016) present a review of different methods used for LPR from 1999 to 2015 where the accuracy of the methods varies between 66% and 99% averaging at 89%.

3 Methodology

3.1 Research Strategy

According to Bryman and Bell (2015) if one seeks to perform research in a flexible and adaptive manner, that allows for new and unpredicted information to appear that could possibly affect the research, a qualitative approach is preferable. Also stated by Bryman and Bell (2015) is the importance when using a mixed method design to motivate and discuss the decision for this choice of method. The method used for the case study of the paper is primarily qualitative, however, evaluating and analyzing the secondary data is done through a mix of qualitative and quantitative method to potentially fill any gaps that could be overlooked using only a qualitative method. The qualitative results will mainly be used to gain a better understanding of the situation of the case study company's operations, to better be able to utilize and develop the quantitative data with the purpose of making it a cornerstone in future decision-making processes.

To decide if certain data can be utilized for decision making processes within a specific company one must understand the current situation of the company, their operations, processes and needs. To further investigate the aspect of data utilization in decision making a case study was conducted at Svenska Mässan, with the purpose of analyzing their data collection method to further develop, utilize and integrate the data in their operative decision processes and making them more data-driven. This paper is based on primary and secondary data, hence there are two different methods used gathering, evaluating and analyzing the data. As mentioned above, for the primary data a qualitative method was used and for the secondary data a mixed qualitative method with aspects of quantitative measures was used. The latter mentioned method allows for more flexibility, since also having aspects of quantitative measures. Using the same qualitative method when dealing with two different aspects of a case study could affect the results negatively. Therefore, it is important that an appropriate method is used for each aspect of the case study.

The first research question (RQ1); *Can the traffic data collection system be utilized for exhibition logistics?* will primarily be answered with the qualitative method, since qualitative methods reflect a deeper and more detailed understanding of thoughts and feelings expressed by the people. The mixed method with the quantitative aspects entangled in it will be used to answer the second research question (RQ2); *How should the current data collection system and data utilization be developed in order for Svenska Mässan to become more data-driven in logistics operations and planning?* The quantitative results from evaluating and analyzing the data and data collection method will present potential areas of improvements as well as current level of knowledge and utilization. Having qualitative aspects part of the mixed method results in the approach to share some of the issues with the qualitative method, for instance, the inability to generalize results. However, when having embedded a quantitative method in the mixed method, the ability to generalize results increases (Bryman & Bell, 2015).

According to Patel and Davidsson (2011) there are three different types of research approaches; deductive, inductive and abductive. The deductive approach is conditioned on hypothesis testing and the hypothesis have been developed from an underlying theory that the research is based on. The hypotheses are tested on e.g. observations where the test either confirms or rejects the stated research hypothesis. An inductive approach is not based on any theory, however the research questions could be. The approach is being used when observations and tests are conducted for the purpose of identifying or searching for patterns in data to eventually arrive at a theory. Lastly, the abductive approach, using either a qualitative or quantitative method, seeks to find the best and most likely explanation or conclusion of a certain effect. According to Bryman and Bell (2015) abductive reasoning has become more common, due to the fact that it surmounts certain limitations or aspects that could hinder deductive or inductive reasoning. Unlike inductive, abductive reasoning permits the researcher to use empirical data to build theories. Also, the mandatory hypothesis testing in the deductive approach is not required for reaching a conclusion in abductive reasoning. An abductive approach has been used in this paper and hence the results of the case study is not necessarily applicable for other companies than Svenska Mässan. The results and conclusions of the case study should not be seen as anything more than the best explanation when seeking out to implement data based decision making in Svenska Mässans logistic processes.

3.2 Research Design

As previously mentioned, a case study research design has been adopted for this paper and according to Yin (2018) a case study is to prefer when research aims to answer questions regarding why and how. Also, stated by Yin (2018) is the relevance of a case study research design when investigating in-depth complexities of an event while being anchored in a real-world element. The purpose of the paper is to analyze the current data collection system implemented at Svenska Mässan. To evaluate if it is suitable for planning of exhibition and logistics operations and to find development ideas for the system based on theory in order for Svenska Mässan to become more data-driven. Given the purpose of the paper a case study research design is suitable for the occasion.

The research questions in this paper aims to answer the question *how,* and like Yin (2018) states when conducting a case study, one of the most important elements is the research question. Other important aspects are; the overall case, the proposition and linkage between the proposition and the data as well as the criteria for interpretation of the findings. Interviews were held with employees at Svenska Mässan and for clarification of certain variables and functionalities found in the data set, questions were asked to the company providing Svenska Mässan with the data. This creates the linkage between the data and the propositions stated in the introduction and the problem statement of this paper. The purpose of the interviews was to get a better understanding of Svenska Mässans current logistics operations and activities, trying to identify potential weaknesses where the implementation of data could be value adding.

3.3 Description of case study company

In 1918 Svenska Mässan, Sweden's oldest exhibition, was inaugurated. In Sweden at that time there was a lack of a business community between industry and trade and the purpose of the exhibition was to introduce a gathering and network place for the two. Svenska Mässan is owned by the Svenska Mässan Stiftelse (Eng. Swedish Exhibition & Congress Center Foundation) and there are no external owners or investors involved or influencing the foundation, which makes it financially independent. Therefore, the foundation makes all financial decisions based on their own merits and all profits are reinvested into Svenska Mässan (Svenska Mässan, 2020b).

Svenska Mässan has come far since the inauguration and has now become a central hub in Gothenburg. Their vision is to become Europe's most attractive meeting place by offering the best possible overall experience. The way they will do this is by promoting business. Their business model is based on a comprehensive offering that allows the visitors to under the same roof, both enjoy and engage in many different experiences. However, the overall objective of Svenska Mässan is to have a prosperous and profitable growth that provides the conditions that enable them to reach and strive towards the vision and the fulfilment of the mission (Svenska Mässan, 2020a).

3.4 Data collection

The data utilized in this thesis is both of qualitative and quantitative nature. According to Collis and Hussey (2014) qualitative data can only be understood within a context and usually accurately reflects the phenomenon that is studied. The primary data consist of only qualitative data whereas the secondary data consists both qualitative and quantitative data. Both the primary and the secondary data are discussed in the following sections as well as the interview process and framework.

3.4.1 Primary data

According to Collis and Hussey (2014) primary data is data generated by the researchers such as experiments, surveys or interviews. The main method of collecting the primary data for this research were interviews. The purpose of the interviews was to get a better understanding about Svenska Mässans operations and what the interviewed employees think and how they feel about certain topics that were discussed in the interviews. Also, their attitudes, opinions, understanding and memories of certain phenomena and events are discussed (Collis & Hussey, 2014).

The interviews conducted are of semi-structured format. As Collis and Hussey (2014) present, semi-structured interviews have beforehand prepared questions in order to guide the interview towards the topics of interest. Interview questions may also develop as the interview is taking place and these questions can be used. In a semi-structured interview, the order of the questions is flexible and some questions might not be necessary to ask if they have been covered in the answer of another stated question. Semi-structured interviews were used to make sure that the main subject was covered and at the same time allowing for the conversation to evolve and cover other topics. For the beforehand prepared questions an interview guide was developed. Using an interview guide helps to reach successful implementation of the interviews and to stay on the intended topic (Kallio, Pietilä, Johnson, & Kangasniemi, 2016). The interview guide is presented in Appendix 1.

Interviews were audio recorded for further analysis. The analysis mostly focuses on the verbal communication but as suggested by Onwuegbuzie, Leech, & Collins (2010) non-verbal communication is also considered. However, as the interviews mostly cover topics about business processes and professional tasks a lot of meaningful non-verbal communication was not expected. When the interview questions might reflect feelings of the interview the emotional non-verbal reactions will be more carefully observed by the interviewer. Also, the recordings of the interviews will be analyzed with the volume, pitch and quality of the interviewees voice kept in mind in order to recognize possible non-verbal aspects (Onwuegbuzie et. al., 2010).

As the case study only focuses on the logistics operations of Svenska Mässan and only personnel working with operations planning are interviewed, a random sample cannot be used to select the interviewees. Collis and Hussey (2014) suggest that in the lack of statistical analysis the results of the sample cannot be generalized to a population and therefore a random sample is not necessary. This approach correlates with the sample selection of this paper. The sample is a specifically selected natural sample within the personnel whose tasks are in the area of the research topic. The researchers received help in selecting the interviewees from Höjer. There was a total of seven interviews conducted with four different people. The details of the interviewees and timetable can be found below in Table 4.

Interview	Interviewee	Position	Time
1	Nicklas Höjer	Director Logistics & Production	07/01/2020
2	Nicklas Höjer	Director Logistics & Production	03/02/2020
3	Nicklas Höjer	Director Logistics & Production	27/02/2020
4	System Provider	Regional Manager	03/04/2020
5	Henrik Martinsson	Logistic Property & Facilities	07/04/2020
6	Martin Godhe	Supervisor Logistics	17/04/2020
7	System Provider	Regional Manager	21/04/2020

Table 4 Interview information

As the research evolved, one interviewee was added to the sample by a recommendation from another interviewee thus also including snowball sampling in the sample selection methods.

In addition to the interviews, observations as a primary data collection approach were used. Both non-participant observations and participant observations were planned. For the non-participant observations, the researchers observe what people do and say without being involved. The researcher must avoid causing disruption that might affect the results of the observation. For the participant observations, the researchers are involved in what is being researched together with the participants. This might give a more detailed understanding of the researched phenomenon (Collis & Hussey, 2014). The non-participant observations in this research were planned to involve observing the logistics operations in the premises of the case company. Unfortunately, due to covid-19 situation and certain restrictions, the researchers were not able to carry out the non-participant observations and they had to be cancelled. The participant observations involved a tour in the case company premises to understand the space where the operations and data collection of the case company takes place. Further, participant observations of daily tasks of the personnel working with the planning and data usage at the case company were planned but these were also forced to be cancelled.

3.4.2 The interview process

A short description of the scope of the thesis had been given to the participants before the interviews, enlightening them as why they have been selected to participate in the interviews. This could potentially have affected the interviewees thoughts and ideas about certain questions and therefore impacted the answers. The interviews were scheduled via email communication and the interviews schedules and execution were affected by the Covid-19 pandemic. The interviews were postponed for several weeks due to this pandemic. The majority of interviews were held with members of staff at Svenska Mässan and one respondent came from the company providing the data collection system used by Svenska Mässan. Due to Covid-19 some of the interviews were rescheduled from being face-to-face into interviews conducted over Teams. Before the interview started, the interviewees were asked if the interview could be recorded, and the state of personal anonymity when used in the thesis, whether the interviewees name or job title could be mentioned. Recording the interviews helped when later transcribing the interviews. In accordance with the nature of semi-structured interviews the questions of the interviews were flexible, and the interviews were adjusted to suit every interview separately. Certain questions were the same for every interview, however some more task and role specific questions were adjusted for each interviewee. The interview guide presents all the questions, even though not all of them were used in every interview. Some of the answers were also irrelevant and were excluded from the findings and analysis.

3.4.3 Interview framework

The interviews were mainly conducted in English even though all the interviewees had Swedish as their mother tongue. This decision was based on two factors; firstly, one of the authors of the thesis does not have Swedish as mother tongue and secondly, to avoid language barriers and translation errors if having to translate the interviews from Swedish to English. However, the phrasing of the interviewees might be affected by their level of English, since some expressions and words came out in Swedish.

3.4.4 Secondary data

According to Collis and Hussey (2014) secondary data is data collected by the researchers from an existing source e.g. company internal records, publications or databases. As mentioned earlier secondary data was also utilized in this research. To better understand and describe the context and present the case company, a PowerPoint presentation received from Höjer was used. This qualitative data helped the researchers to better understand the case company's operations, structure and strategy.

A very important piece of secondary data was utilized, the traffic data collected by the case company's existing sensor system. The sensor system includes six sensors in the entrances and exits of Svenska Mässans logistics operation premises. The first sensors have been collecting traffic data since February 2019 and the last one of the six being installed in May 2019. The data presents e.g. observations of vehicles of different length entering and exiting the premises as well as speed and other data. The researchers extracted the data from the case company's existing database. The data is presented with more details in the following chapter.

A major part of carrying out this research and being able to understand the topic at hand was through the literature review that was conducted. The literature review is based on secondary data sources such as scientific journals, articles, books, web pages, news articles and online publications. The researchers are aware that some of the sources used are over ten years old, but the topics and conclusions are still up to date. These sources were used since they felt more related to the specific topic of the paper. This part of the secondary data was essential for finding theoretical background and best practices for analyzing the case company and answering the research questions. The topics covered in the literature review were developed based on the problem description and information of the current operations of Svenska Mässan. The information of the problem and current state was gathered during initial meetings with Höjer at Svenska Mässan as well as from Svenska Mässan online publications. The tools used in finding the theoretical data for the literature review included Gothenburg University Library literature search engine Supersearch. This source provided the researchers with a lot of useful material. Also, Google Scholar was highly utilized in finding the material relevant to the topic. General Google searches were also conducted to better find material outside of the academic world.

3.5 Empirical writing and data analysis

In this paper thematic analysis was used when conducting the data analysis of the interviews, one of the most common tools to analyze qualitative studies (Bryman & Bell, 2015). To achieve this type of analysis the interviews were first transcribed and afterwards analyzed using a coding system. The coding followed certain themes and areas of the literature framework, identifying important aspects to connect the research with previous literature. This method also ensures that the findings were more accurately represented. However, the method of analysis can be questioned and criticized, one cannot exclude or be completely certain of the authors' abilities of staying objective when coding and analyzing the interviews. Hence, the authors' own assumptions about the replies from the interviewees not being present (Alvesson, 2003).

To answer both research questions of the paper a framework of analysis was created, where the empirical findings were evaluated, and certain connections were made to the theoretical framework.

3.6 Research quality

According to Collis and Hussey (2017) there are two concepts that help define the quality of a research, these are validity and reliability. The validity of a study is determining and making sure that the correct components are measured and corresponds with the framework of the research. Reliability of a study is dependent on the consistency of the results when the study is replicated. However, if the research is of qualitative nature it has lower measurability compared to quantitative research and thus according to Bryman and Bell (2015), qualitative research should be examined based on trustworthiness rather than validity and reliability. There are four benchmarks used when evaluating the trustworthiness of a research; *dependability*, *conformability*, *credibility* and *transferability*.

Dependability can be interpreted in a similar way as reliability, i.e. if the research were to be performed again how accurate and precise would the results be when compared to the previous attempt (Bryman & Bell, 2015). In the case of this thesis, the research is very company specific, thus if the same kind of theory and research were to be used and conducted in another company, the results would arguably become different.

According to Bryman and Bell (2015) the concept *conformability* measures the biased level of the participants towards the results of the study or research. Bryman and Bell (2015) also states that; within qualitative research it is more desirable to have a low level of conformability, since qualitative research is normally more subjective than quantitative research. This paper covers both a qualitative and quantitative aspect and the majority of all the interview respondents, which relates to the qualitative aspect of the study, are employees at Svenska Mässan. Further, the observations that took place were also related to Svenska Mässan's operations. However, the conformability of this study could be considered to be on a low level, due to the outside perspective of the researchers conducting the research.

According to Bryman and Bell (2015) *credibility* is equivalent to the concept, internal validity, which tries to measure if a research was conducted correctly and describes the extent of how sources of systematic errors can be excluded from the results. Systematic errors are considered as consistent errors, and thus when repeating the same research or experiment, the same errors will be present. The majority of all the data presented in the paper is either collected through interviews or observation connected to Svenska Mässan and therefore systematic errors cannot be ruled out of the research. However, the transparency of the interview guide as well as the selection of interviewees increases the credibility of the paper.

Like credibility, *transferability* also has an equivalent concept; external validity, which outlines the extent of how well the results of a research can be generalized (Bryman & Bell, 2015). The results of this research are not applicable to all companies and organizations, not even necessarily within the same industry as the case study company Svenska Mässan. Even though companies are part of the same industry, that does not necessarily mean their operations and processes are the same. There are numerous variables affecting the transferability of a research, in this case; e.g. IT systems, customers, knowledge, experience, logistics space etc. However, some findings from this paper can probably be applied to other exhibition companies with similar exhibitors and exhibitions as Svenska Mässan.

3.7 Research ethics

Ethics of a research refers to the way a research is conducted and to the way findings and results are reported and presented. There are a lot of ethics related issues that must be considered when planning a research. The most relevant ethical issues for this research are voluntary participation, dignity, privacy, confidentiality, anonymity, honesty and transparency as suggested by Collis and Hussey (2014). The way this research deals with the previously mentioned ethical issues is presented in Table 5.

Ethical issue	Approach of this paper
Voluntary participation	All participants are informed that participation is voluntary and that they have the right to withdraw from taking part in the research at any time.
Dignity	The dignity of the participants is respected by staying in related topics and acting neutral on matters of emotions, attitudes or opinions
Privacy	The research respects the privacy of all participants by asking for consent of the participants
Confidentiality	The information collected during the research can be made confidential and untraceable to single persons if the participants so desire
Anonymity	All participants are given the possibility to stay anonymous in the research if they so desire
Honesty and transparency	The researchers act honestly and transparently when communicating with all stakeholders

Table 5. Approach to ethical issues (Developed by the researchers based on Collis and Hussey, 2014)

4 Empirical findings

4.1 Exhibition Planning and Execution

The planning of an exhibition starts about 1 year prior to the actual date of the exhibition. It starts with an information package that is put together and sent to the exhibitors containing e.g. where the exhibitors are supposed to send their goods, who they should contact, what the safety regulations are etc. One of these information packages is created for every exhibition. About 2-3 months before the exhibition more operative questions are being discussed between the exhibitors and Svenska Mässan, mainly if there are any special requirements e.g. transports of large goods that need a special forklift in the venue (Martinsson, 2020).

Closer to the date of the exhibition the exhibitors start planning the transport of the goods. Svenska Mässan provides transporting service to the exhibitors, using their 3PL (third party logistics) provider. However, in many of the cases the exhibitors use their own transport service and therefore only need to book an unloading slot at Svenska Mässan. When booking an unloading slot Svenska Mässan requires the estimated time of arrival, delivery quantity and to which exhibition space the exhibitor has been assigned. The exhibitors provide this information in a digital document connected to an IT system used by Svenska Mässan. The assigned time slots are approximately 40 min. First, they must make sure that the truck is actually supposed to deliver to Svenska Mässan, according to Martinsson, it is not uncommon that there are trucks arriving at Svenska Mässan that are not supposed to make the delivery at their facility. When a truck arrives with goods to Svenska Mässan the delivery is registered in the IT system by scanning the shipping note and barcodes are then applied to all the packages. The barcode contains information about when and where the package was delivered, who the customer is and the assigned location in the exhibition hall (Martinsson, 2020).

Svenska Mässan has a consolidation center in proximity to the city center, all deliveries with less than 8 pallets get assigned to deliver to the consolidation center instead of Svenska Mässan. This is due to Svenska Mässan exceeding their capacity limit during exhibitions. Therefore, they reroute the smaller deliveries to the consolidation center where they consolidate the shipment to increase the fill rate of the trucks and from here they have more control over the delivery and can decide when to deliver to Svenska Mässan. However, the process of scanning the shipping note and applying barcodes to the packages is still the same at the consolidation center as it is at Svenska Mässan. The IT system, handling the delivery information, customer information and the track and trace of the packages, is connected to Svenska Mässans ERP system (Enterprise Resource Planning). After an exhibition, all the data and information that is customer related from the IT system is moved over to the ERP system which creates the invoice to each individual customer. The current set up of Svenska Mässans exhibition logistics transport chain, having access to a consolidation center, is similar to the Exhibition Logistics Supply Chain System discussed by Zhang (2012).

4.2 Current data collection system

The current data collection system consists of six intrusive sensors strategically placed in the ground in proximity and inside Svenska Mässans underground garage, four of which are placed at the four different entries and exits to the underground garage approximately 1,5 meters from the gates. By using magnetism that senses movement in earth's magnetic field, the sensors collect and register data from passing vehicles. However, in connection with the underground space used mainly for loading and unloading, is a valet parking garage and a separate parking space dedicated to the employees. Therefore, it is important to separate the data that comes from both valet- and employee parking from the data set in order to get a more accurate and representable data set.

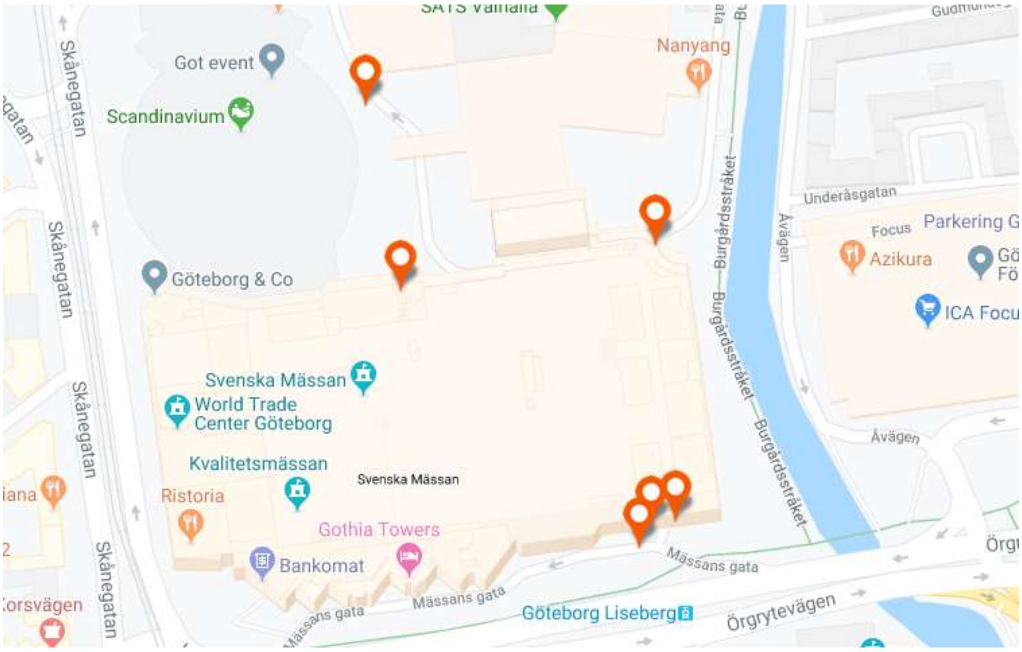

Figure 3. Sensor locations on a map (Höjer, N. 2019, personal communication, 27 December)

The system was bought and installed for the purpose of gathering data and information about the traffic movements in connection with Svenska Mässan, information that could be utilized for strategic decision making. According to Martinsson, head of Logistic Property and Facilities at Svenska Mässan, the work around the exhibitions are mostly done through experience. As exhibitions develop through the years and only a few of them are annual events, the installment of a system like this could help them to make decisions based on facts and data instead of being made based on experience and best guesses.

4.3 Data description

Svenska Mässan acquired the data collection system from a company specialized in collection and analysis of traffic data. The company provides the collected data to Svenska Mässan through their online platform. In the platform the data is summarized in a few different tables and graphs that give an overview of the data. The same data can also be exported in two formats; in PDF with the graphs and tables provided from the online platform or in Excel format which consists of all collected data for the selected time period. The data is reported through ten different variables, presented in Table 6 below.

1	2	3	4	5	6	7	8	9	10
Measuring Point	Time	ID	Amount	Type	Length	Speed	Direction	Lane	Extra info

Table 6. Data variables

The first variable of the data set is called *Measuring Point* and it represents the location where the observation was measured. All six different sensors have been named with an appropriate name to represent each location.

The second data variable, *Time*, represents the exact time of each individual observation. This is one of the essential variables as nearly all analysis performed, and all application of the data needs a reference in time. The level of detail is one thousandth of a second which is more than enough for utilizing the variable. The date format is year-month-day which follows ISO 8601 standard (ISO, 2020).

The third variable, *ID*, is an individual ID number for each observation. This number is not connected to a specific vehicle in a way that would enable it to act as a tracking ID. Therefore, the data does not specify when a specific single vehicle entered or exited the Svenska Mässan premises. The ID numbers mostly run in logical consecutive order, but sometimes certain values are skipped. The ID is not a very useful variable for the analysis, the only purpose it might serve is using the ID to refer to a single observation. However, as most analysis is done in more of an overall perspective, referring to single observations is not necessary. Therefore, the variable will not be part of the analysis.

The fourth variable, *Amount*, represents the number of observations the sensor registered at a specific date and time. This variable only takes a value equal to 1 for all the observations, because the system only measures and registers one observation at a time. The variable does not provide the research with any useful information and is therefore irrelevant for the data analysis.

The fifth variable, *Type*, represents the vehicle type, either if it is a “lighter” or a “heavier” vehicle. All vehicles under 3,5 tonnes are classified as light vehicles and all vehicles above 3,5 tonnes are classified as heavy vehicles. However, if there is a measurement error in the variable *Type*, the observation is classified as “Unknown”.

Variable six of the data set, *Length*, represent the length of the passing vehicle, measured in meters. *Speed*, the seventh variable of the data set, is only measured at one specific point inside Svenska Mässan’s underground garage and it is presented in km/h. If these two variables are experiencing a measurement error the observation will be blank.

Variable number eight, *Direction*, shows the travel direction of the vehicle, if it is going south, north, east or west. This variable is being used to decide if the vehicle is moving in or out of Svenska Mässan’s underground garage at each assigned entry and exit point. If the eighth variable *Direction* is exposed to a measuring error the variable will report the value "Unknown".

Variable number nine, *Lane*, presented in the data set represents the lane in which the vehicle is driving, and this variable only takes a value of 1 for all the observations. The variable does not provide the research with any useful information and is therefore irrelevant for the data analysis.

Lastly, the tenth variable is called *Extra Info* which is either left blank for an observation or an observation is labelled with a text that translates into "lowered accuracy". The observations labelled with "lowered accuracy" have experienced a signal interference, affecting one of the variables *Type*, *Length*, *Speed* or *Direction*. The observations are still part of the data set since the signal is good enough for the system to be sure that a vehicle has passed the sensor and therefore it is not excluded as it will be part of the total number of observations. However, there is a high chance that a measurement error has occurred in one of the variables *Type*, *Length, Speed* or Direction and it will therefore be analyzed together with these four variables.

5 Analysis

5.1 Data collection system analysis

This section will analyze the current data collection system. Firstly, all the attributes of the system are analyzed followed by an analysis of the quality of the data collected by the system.

As explained earlier, the data from the sensors are collected at 6 different measuring points in the Svenska Mässan area. For analysis purposes the measuring points are number coded and the numbers as well as their inbound and outbound geographical directions are presented in Table 7. As measuring points 4, 5 and 6 are one way points the opposite directions are presented in *italics*. The date each measuring point started recording traffic data is presented in the last column.

	Measuring point	**In**	**Out**	**Started**
1	In och Utfart Glenns kurva	West	East	11/02/2019
2	In och Utfart Scandinaviumsidan nedre plan	West	East	11/02/2019
3	In och Utfart Scandinaviumsidan Ramp till övre plan	South	North	05/04/2019
4	Infart Lisebergssidan	North	*South*	11/02/2019
5	Ramp till Hotell och Mässentré	West	*East*	10/05/2019
6	Utfart Lisebergssidan	*North*	South	11/02/2019

Table 7. Number coding of data measuring points.

Further, measuring points 2 and 5 are seen as irrelevant. Measuring point 5 is located inside the underground premises of Svenska Mässan and all the vehicles passing this specific sensor have already been registered by one of the sensors located at one of the entries or exits. Using the data from sensor five would result in every one of these observations being used twice. Measuring point 2 is located on a ramp leading up to the hotel entrance, there are mainly taxis and private cars passing this sensor on their way to the hotel lobby of the Svenska Mässan establishment. This ramp is not used for deliveries and it was therefore decided to be excluded from the analysis. This decision was confirmed by Martinsson who expressed in his interview that those measuring points should be ignored.

5.1.1 Main Variable analysis

As presented in Table 6 there are ten different variables in the data set. However, only six of them will be part of the analysis according to statements made in the Empirical findings in the variable description section. The first data variable part of the analysis presents the location of each *Measuring point*. The names of each point are selected logically based on their true physical location in the Svenska Mässan premises. However, to really understand the measuring points one needs to be familiar with the premises and the observation locations. Also, according to Martinsson the name of measuring point three is misleading as the name includes both entrance and exit but in real life it is only an exit. This is also visible in the data where approximately 94% of the observations at this location are outbound moves. But, as long as the person utilizing the data knows the actual locations of the sensors and is familiar with the true nature of measuring point three the variable can be fully utilized as it is.

The second variable, *Type*, describes the type of the vehicle. As mentioned before vehicles under 3,5 tonnes are classified as *Light* and above 3,5 tonnes *Heavy*. This study is interested in cargo vehicles only and the *Type* variable could be of great value in separating cargo vehicles from personal vehicles. However, as a cargo delivery vehicle can have a weight of under 3,5 tonnes the research cannot exclude the type *Light*. This is a big weakness for the variable *Type* since it cannot be determined which of the vehicles with the classification *Light* is personal cars being parked inside the Svenska Mässan premises. As the valet parking of the hotel is in the same premises all those vehicle observations are visible in the data. With the current setting of variable *Type* personal vehicles cannot be excluded for proper analysis of cargo vehicles only.

Looking at the relevant data (data from measuring points 1, 3, 4 and 6) for variable *Type* from the first measurements of the system until the end of March 2020 there are a total of 145 452 observations. Different vehicle types and the percentage of total observations are presented in Table 8 below.

Type	**Total**	**Light**	**Heavy**	**Unknown**
Observation	145425	129102	11969	4354
% of observations	100,00	88,78	8,23	**2,99**

Table 8. Division of data for variable Type.

In the above table the share of "Unknown" observations of the total amount is bolded, giving us an error level of 2,99%. This means that for approximately 97% of all observations have been classified as either a light or a heavy vehicle. The remaining 3% are unknown observations where the system has detected a vehicle passing but cannot detect the type of the vehicle. These values are still valuable for overall analysis but do not give any indication of what type of vehicle.

The third variable in our analysis is *Length* and it describes the length of the observed vehicle in meters with a precision of one centimeter. The observations are divided in five different groups. The length values in the data set vary from 0,84 meters up to a staggering 98,14 meters with an average of 5,14 meters. The median of the variable is 4,76 meters. As mentioned above, Type cannot be used to exclude personal vehicles from the data. *Length* could be used in making distinctions between different kinds of vehicles but as the length of cargo transporting vehicles can vary a lot, it cannot be surely stated that a vehicle of certain length is a cargo vehicle. Table 9 shows a division of lengths that might help in separating personal vehicles from larger vehicles. From the table it can be concluded that there are 4,79% of the total observations that have been left blank. In the data set a total of 6 970 observations of vehicles were measured where *Length* could not be determined.

Length m	Total	-2,99	3-5,09	5,1-7,29	7,3-25,25	25,26-	Blank
Observation	145425	28905	53064	35707	20449	330	6970
% of observations	100,00	19,88	36,49	24,55	14,06	0,23	**4,79**

Table 9. Division of data for variable Length.

The different length intervals in the table could represent different types of vehicles. The length of a regular personal vehicle is approximately between 3 and 5 meters (Mechanic Base, 2020). A van or a small truck is approximately between 5 meters and around 7,3 meters (Sportsmobile, 2020) and a regular or a big truck between 7,3 and 25,25 meters. As the EU regulation (European Parliamentary Research Service Blog, 2020) does not allow heavy cargo vehicles to exceed 25,25 meters, all observations with lengths above 25,25 meters are considered measurement errors. Observations with lengths below 3 meters are too small for being a regular car and are therefore also considered as measurement errors. According to Martinsson some of the sensors are installed in a way that they might detect passing cyclists. These might also be included in the observations of lengths under 3 meters. Considering the above, observations under 3 meter and over 25,25 meter will be added together with the blank observations to create the measurement error. All the blank observations, observations below 3 meters and all above 25,25 meters together sums up to a total of 36 205 observations. This results in a measurement error of 24,9% which is significantly worse than the previously mentioned 4,79% that only considered the observations where the length could not be determined. Looking at Table 9 the observations under 3 meters represent almost 20% of the data. This high number of less than 3 meter observations is also reflected in the median which at 4,76 meters is lower than the average of 5,14 meters.

In Table 10 below, the fourth variable *Speed* is presented in kilometers per hour (km/h) and all observations are divided into 5 different intervals. The speeds vary from 1 km/h to 199,1 km/h with an average of 16,24 km/h. The number of blank observations add up to approximately 3% of the total amount of observations.

Speed km/h	Total	-10	10,0-20	20,1-30	30,1-40	40,1-	Blank
Observation	145425	71339	64753	3376	409	1194	4354
% of observations	100,00	49,06	44,53	2,32	0,28	0,82	**2,99**

Table 10. Division of data for variable Speed.

Most observations, approximately 94%, are found under the speed of 20 km/h. The location of the sensors, as previously mentioned, are located approximately 1,5 meters from the gates of the entries and exits of the establishment and therefore low speeds are expected of the vehicles. Because of this, observations exceeding speeds of 40 km/h are considered measurement errors and result in 0,82% of the total observations. Adding these observations together with the observations left blank results in a total measurement error of 3,82% of total observations. Speed as a variable from this system does not provide a lot of value adding information other than potentially from a safety perspective. The speed is measured at each sensor and there is no connection between the different sensors. Therefore, the variable cannot be used in identifying traffic situations, such as queues within Svenska Mässan premises.

The fifth variable *Direction* has a total of 3 187 observations that are categorized as "Unknown" which represents 2,19% of the total amount of observations. As mentioned earlier direction is the variable that defines if a vehicle is inbound or outbound. This is very useful as two out of the four relevant measuring points are observing both the in- and outflow. The direction helps to analyze the overall traffic balance in the Svenska Mässan premises and can also help determine the accuracy of the data collection system by aggregating the total amount of inbound and outbound. The number of total vehicles going into Svenska Mässans underground garage should equal the same amount going out of the establishment for the system to have high accuracy in the variable. This will be analyzed at the end of the next section.

Direction	**Total**	**South**	**North**	**East**	**West**	**Unknown**
Observation	145425	22431	71593	19589	28625	3187
% of observations	100,00	15,42	49,23	13,47	19,68	**2,19**

Table 11. Division of data for variable Direction.

Lastly, the sixth variable is called *Extra Info* which is either left blank for an observation or an observation is labelled with a text that translates into "lowered accuracy". This variable is used to determine the accuracy and potential measurement errors of the systems. As mentioned above, observations labelled with "lowered accuracy" have experienced a signal interference, affecting one of the variables *Type*, *Length, Speed* or *Direction*. The observations are still part of the data set since the signal is good enough for the system to be sure that a vehicle has passed the sensor and therefore it is not excluded as it will be part of the total number of observations. However, there is a high chance that a measurement error has occurred in one of the variables and therefore it will be analyzed together with these four variables.

5.1.2 Data quality

First part of this section will cover the analysis of the measurement errors for each measuring point based on observations labelled with "lowered accuracy" as well as the "Unknown" and blank observations. In the second part the unknown data and how it is represented in some individual attributes is analyzed in comparison between different measuring points. Thirdly the data quality analysis looks at how the global Covid-19 pandemic has affected the traffic data during early 2020. This is done as an additional quality measurement to see if the data reflects the Covid-19 situation that has had a great impact on Svenska Mässans operations. Lastly the balance of inbound and outbound traffic volumes is analyzed.

Observations with lowered accuracy and Unknown or blank observation

Discussed above were the measurement errors for the variables *Type*, *Length*, *Speed* and *Direction*. Overview of the percentages of the measurement errors is presented in Table 12 below.

Variable	Error %
Type	2,99
Length	24,90
Speed	3,82
Direction	2,19

Table 12. Measurement error per variable.

The error levels can be considered low for all the variables except *Length.* However, the above numbers have not taken into consideration the variable *Extra info*. This is an issue that needs to be considered since according to the interview with the System Provider, observations labelled with “lowered accuracy” have experienced a signal interference affecting one of the variables *Type*, *Length, Speed* or *Direction* and there is a high chance this will result in a measuring error. The data set of the studied time period consists of 145 425 observations in total and out of these, 59 339 observations have experienced a signal interference and therefore been labelled with “lowered accuracy”. Considering this, then 40,8% of the total observations have a high chance of a measurement error in one of the variables *Type*, *Length, Speed* or *Direction*. To see which variable out of these four that was affected by the signal interference is not easily determined.

Seen in Table 13 below is each measuring point analyzed according to total observations as well as observations labelled with “lowered accuracy” and how many percentage “lowered accuracy” is of the total observations at each point.

Measuring point	Total observations	Lowered accuracy	% of observations
1	48608	16569	34,09
3	18379	3881	21,12
4	57168	35906	62,81
6	21270	2983	14,02

Table 13. Observations with Lowered accuracy per measuring points.

From Table 13 it can be seen that the measuring point with the highest percentage of “lowered accuracy” observations is measuring point 4 “*Infart Liseberg*” with a total of 62,8%. This means that approximately six out of ten observations could have experienced a measurement error. The same conclusion about the high measurement error potential of measuring point 4 can be made when observing Table 14 below. The percentage of a measuring point is calculated as the average of ”Unknown” and blank observations of the four variables. Measuring point 4 has the highest share of “Unknown” and blank observations out of total observations.

Measuring point	%
1	1,29
3	1,40
4	6,27
6	1,17

Table 14. Unknown and blank observations for the variables per Measuring point

The above findings might be due to the location of the sensor, measuring point 4 is located outside the Svenska Mässan garage and according to the System Provider it is approximately 1,5 meters from the entrance. This sensor is the most likely to have disturbance in the quality of data due to external effects e.g. bicycles crossing the sensor, causing a change in the magnetic field and potentially affecting one of the variables *Type*, *Length, Speed* or *Direction.* Even though the sensor is outside of the building, the weather should not influence the intrusive magnetic sensor. According to Martinsson, it is not uncommon that delivery vehicles accidentally arrive at Svenska Mässan instead of their correct destination, a factor that could potentially affect the sensor and its generated observation if e.g. a vehicle were to reverse or make a U-turn over the sensor. The placement on the road has a large impact on the sensors capability of registering and collecting observations.

Comparison between measuring points and variables Type, Length, Speed and Direction

This part will present specific data divided per measuring point and presenting all categories of the variables *Type*, *Length*, *Speed* and *Direction*. It will start by looking at measuring point 1 "*In och Utfart Glenns kurva*" which measures both in- and outbound moves. The data from measuring point 1 seen in Table 15 below, shows the percentage of the "lowered accuracy" observations regarding the total amount of observations at this measuring point, resulting in 34,09%. A pattern seen with all the four variables are the blank or "Unknown" observations having a significantly higher percentage of observations being affected by a signal interference and thus resulting in being labelled as "lowered accuracy". This indicates that when the measuring point was not able to record a detailed observation ("Unknown" or blank) for the passing vehicle it was more likely that accuracy was lowered. For the variable *Type,* the percentage lowered accuracy is the highest for those vehicles categorized as heavy. This makes 62,38% of the observation classified as heavy to have been affected by a signal interference and could therefore be exposed to a measurement error. This could indicate that the sensor is having issues measuring larger vehicles. As discussed previously, the observations under 3 meters and above 25,25 meters for the variable *Length* are considered measurement errors and this can clearly be seen in the high share of "lowered accuracy" observations for these groups. According to the observations the most common vehicle going in and out of Svenska Mässan could be personal cars since the amount of observations for the group with a length between 3-5,09m is the highest and this group also has the lowest percent of observations with "lowered accuracy". Similar trends can be seen with variable *Speed* where observations with a speed exceeding 40 km/h have a very high share of "lowered accuracy" observations as well as blank observations. The lowest values can be found in the group 10,1-20 km/h which is the most common speed group as seen in Table 10 earlier in the analysis. For the variable *Direction,* the highest share of “lowered accuracy” is clearly within “Unknown” observations. However, observations going outbound east, have a significantly higher share of “lowered accuracy” then the inbound west observations. This suggests that the sensor at measuring point 1 has fewer measurement errors for observations going inbound. This might create imbalance with out- and inbound moves when looking at the overall traffic flow. This will be discussed later in the analysis.

	Total observations	Lowered accuracy	% of lowered accuracy
Type	48608	16569	34,09
Light	44875	14247	31,75
Heavy	3174	1980	62,38
Unknown	559	342	61,18
Error			1,15

	Total observations	Lowered accuracy	% of lowered accuracy
Direction	48608	16569	34,09
East	19589	9841	50,24
West	28625	6476	22,62
Unknown	394	252	63,96
Error			0,81

	Total observations	Lowered accuracy	% of lowered accuracy
Speed	48608	16569	34,09
-10	12772	5652	44,25
10,1-20	32592	9571	29,37
20,1-30	2121	641	30,22
30,1-40	237	123	51,90
40,1-	327	240	73,39
Blank	559	342	61,18
Error			1,15

	Total observations	Lowered accuracy	% of lowered accuracy
Length	48608	16569	34,09
-2,99	5355	3526	65,85
3-5,09	24372	5928	24,32
5,1-7,29	11940	3774	31,61
7,3-25,25	5861	2558	43,64
25,26-	84	73	86,90
Blank	996	710	71,29
Error			2,05

Table 15. Measuring point 1 divided by variables Type, Direction, Length and Speed

The next table looks at measuring point 3 “*In och Utfart Scandinaviumsidan Ramp till övre plan”* which as explained before only measures outbound movements despite the name translating into both in- and outbound vehicle movement. The overall percentage of “lowered accuracy” observations for measuring point 3 is 21,12% as seen below in Table 16, a lot lower than for measuring point 1. Unknown or blank moves for all variables have a significantly higher share of “lowered accuracy” following the same trend as measuring point 1. The percentage difference between “lowered accuracy” for the light and heavy groups in the variable *Type* is much lower than for measuring point 1, showing only a few percentage points difference. This suggests that the sensor at measuring point 3 does not have the similar potential issue as the sensor at measuring point 1, having many signal interferences for the group heavy vehicle. The same trends as for measuring point 1, discussed above, can also be seen for the variables *Length* and *Speed* measuring point 3. The blank and “Unknown” observations have high percentages observations with “lowered accuracy”. However, for this measuring point there is a lower percentage of “lowered accuracy” observations with a length above 25,25 meters than for the blank observations. At this measuring point for the variable *Speed* are high percentages of signal interference for both the blank and speeds exceeding 40 km/h observations. Were once again the group of observations having a speed that exceeds 40 km/h have the highest percentage of “lowered accuracy” observations. This measuring point should only cover vehicles exiting the Svenska Mässan establishment since it is only an exit and should only have outbound north vehicle movement. Therefore, it is expected that the inbound observations have a significantly higher share of “lowered accuracy” observations and the southbound vehicle movements for this measuring point have almost as much “lowered accuracy” observations as the “unknown” observations.

	Total observations	Lowered accuracy	% of lowered accuracy
Type	18379	3881	21,12
Light	16563	3406	20,56
Heavy	1616	374	23,14
Unknown	200	101	50,50
Error			1,09

	Total observations	Lowered accuracy	% of lowered accuracy
Direction	18379	3881	21,12
North	17424	3452	19,81
South	817	363	44,43
Unknown	138	66	47,83
Error			0,75

	Total observations	Lowered accuracy	% of lowered accuracy
Speed	18379	3881	21,12
-10	7106	2605	36,66
10,1-20	10022	1026	10,24
20,1-30	904	74	8,19
30,1-40	51	11	21,57
40,1-	96	64	66,67
Blank	200	101	50,50
Error			1,09

	Total observations	Lowered accuracy	% of lowered accuracy
Length	18379	3881	21,12
-2,99	2837	1612	56,82
3-5,09	3707	969	26,14
5,1-7,29	6324	463	7,32
7,3-25,25	4996	469	9,39
25,26-	24	15	62,50
Blank	491	353	71,89
Error			2,67

Table 16. Measuring point 3 divided by variables Type, Direction, Length and Speed

Measuring point 4, “*Infart Lisebergssidan*” only measures inbound north moves. As discovered earlier in the analysis, measuring point 4 has significantly higher error rates compared with the other measuring points. This is clearly visible also in Table 17. The percentage of “lowered accuracy” observations in comparison to the total amount of observations for this measuring point is 62,81% which is higher than the percentages of “lowered accuracy” observations regarding the unknown or blank observations for all three variables except *Length.* All the four variables have a higher percentage of “lowered accuracy” observations compared to at least one of the groups the observations have been divided in. The variable *Type* has a high percentage of “lowered accuracy” observations in the group heavy 77,50% of all the observations classified as heavy. For the variable *Length* 95,37% of the total observations classified as having a length above 25,25 meters have been exposed to a signal interference that could potentially cause the high percentage of undisclosed measurement errors for this group. Observations for the vehicles that had a length exceeding 25,25 meters were classified as measuring errors. Also, for the variable Speed, there is a high percentage (89,17%) of observations in the group where speeds exceed 40 km/h that have been experiencing signal interference. A similar variable trend for *Type*, *Length* and *Speed* was seen for measuring point 1. The variable Direction does not follow the same pattern as it did for measuring point 1. This was expected since measuring point 4 is only used for vehicles entering the establishment (North) and therefore the observations from group “South” have a very high percentage of "lowered accuracy". Overall reliability of measuring point 4 seems to be quite low as discussed before in the analysis.

	Total observations	Lowered accuracy	% of lowered accuracy
Type	57168	35906	62,81
Light	47897	29606	61,81
Heavy	5862	4543	77,50
Unknown	3409	1757	51,54
Error			5,96

	Total observations	Lowered accuracy	% of lowered accuracy
Direction	57168	35906	62,81
North	52911	33339	63,01
South	1756	1434	81,66
Unknown	2501	1133	45,30
Error			4,37

	Total observations	Lowered accuracy	% of lowered accuracy
Speed	57168	35906	62,81
-10	41031	27990	68,22
10,1-20	11631	5280	45,40
20,1-30	274	159	58,03
30,1-40	112	86	76,79
40,1-	711	634	89,17
Blank	3409	1757	51,54
Error			5,96

	Total observations	Lowered accuracy	% of lowered accuracy
Length	57168	35906	62,81
-2,99	18440	14573	79,03
3-5,09	15208	10022	65,90
5,1-7,29	11017	4778	43,37
7,3-25,25	7276	3069	42,18
25,26-	216	206	95,37
Blank	5011	3258	65,02
Error			8,77

Table 17. Measuring point 4 divided by variables Type, Direction, Length and Speed

Lastly, measuring point 6 "*Utfart Lisebergssidan*" which only measures vehicle movement going out from Svenska Mässan, direction south. The percentage of "lowered accuracy" observations in comparison to the total amount of observations for this measuring point is 14,02%, which is the lowest for all four measuring points. All the variables follow similar trends as with the previous measuring points, *Type* suffering from higher percentage of "lowered accuracy" in the group "Heavy", *Length* and *Speed* having a large portion of signal interferences for the groups of vehicles with a length above 25,25 meters and speeds exceeding 40 km/h. Both of these groups were classified as measurement errors. According to the variable *Direction*, there are 1 258 vehicles being registered as moving north, which would mean that they entered the Svenska Mässan establishment through an exit lane. The reliability of these observations is therefore questionable and could be one reason that the "lowered accuracy" percentage is above 50%.

	Total observations	Lowered accuracy	% of lowered accuracy
Type	21270	2983	14,02
Light	19767	2542	12,86
Heavy	1317	343	26,04
Unknown	186	98	52,69
Error			0,87

	Total observations	Lowered accuracy	% of lowered accuracy
Direction	21270	2983	14,02
North	1258	632	50,24
South	19858	2267	11,42
Unknown	154	84	54,55
Error			0,72

	Total observations	Lowered accuracy	% of lowered accuracy
Speed	21270	2983	14,02
-10	10430	2308	22,13
10,1-20	10508	492	4,68
20,1-30	77	24	31,17
30,1-40	9	6	66,67
40,1-	60	55	91,67
Blank	186	98	52,69
Error			0,87

	Total observations	Lowered accuracy	% of lowered accuracy
Length	21270	2983	14,02
-2,99	2273	1216	53,50
3-5,09	9777	846	8,65
5,1-7,29	6426	335	5,21
7,3-25,25	2316	227	9,80
25,26-	6	5	83,33
Blank	472	354	75,00
Error			2,22

Table 18. Measuring point 6 divided by variables Type, Direction, Length and Speed

Generally, the above tables show that the total percentages of "lowered accuracy" is high for the "Unknown" or blank observations and for the observations that have been deemed as measurement errors due to extremely low or high values for lengths and speeds. This shows that when the system has made an observation with a length that is too long or too short or with a speed that is too fast there is a higher chance that the observation suffered from lowered accuracy. Based on this it can be concluded that when the system is affected by "lowered accuracy" it seems to make more extreme, unknown and blank observations, e.g. a vehicle with length of 98 meters.

All measuring points have the highest percentage of measurement error rate in *Length* which is the percentage of unknown or blank observations of total observations. Similar results were also seen in Table 12 with adjusted measuring error per variable. This suggests that the system has most trouble recognizing the correct length of an observed vehicle. This might be due to several reasons that research has often concluded about the problems with magnetic sensors. Wang et. al. (2018) suggests, based on their research, that the performance of a sensor is lower when vehicles observed are very close to each other. As the magnetic images of the vehicles blend together it becomes more difficult for the system to make individual accurate observations. This might be the situation if there is a queue of vehicles entering or exiting the Svenska Mässan premises. A similar queue situation might lead into vehicles standing still on top of the sensor. According to Jain et. al. (2019) magnetic sensors have problems recognizing vehicles standing still and this might further increase the measurement errors for the variable representing length.

The variable *Type* has more signal interferences for those observations classified as heavy vehicles. This could suggest that the system does not observe larger vehicles as fluently as smaller ones. It might be as Bottero et. al. (2013) stated in their article that vehicle classification results vary more with heavy vehicles as the cargo they carry might affect the magnetic image. Also, according to Taghvaeeyan, & Rajamani (2014) heavy vehicles on non-adjacent lanes might disturb the magnetic sensor. This might cause higher levels of signal interference for heavy observations in the Svenska Mässan data collection system. These factors suggest that larger vehicles create more disturbance for the sensor as the magnetic fields might vary more than with smaller vehicles.

The impact of Covid-19 on traffic volumes from January 2020 to March 2020

This part of analyzing the quality of the data strives to recognize the impact of the Covid-19 situation by comparing traffic data from January 2020 to March 2020. The result will give an indication if the collected data is accurate or not. The second part of this analysis will cover the "Unknown" data and how some individual attributes of it are represented. Lastly the balance of inbound and outbound traffic volumes is analyzed.

On March 11th, 2020 the Swedish government prohibited all events with over 500 participants (Regeringskansliet, 2020) leading to the cancellation of all exhibitions taking place at Svenska Mässan starting from 12th of March 2020 (Svenska Mässan, 2020c). Due to this cancellation of exhibitions and other events for over 500 participants as well as general government recommendations for avoiding unnecessary travel etc. the number of visitors and traffic activity at Svenska Mässan has dramatically declined. This decline in activity is also visible in the data collected by the sensors at Svenska Mässan and is therefore also affecting the data analysis.

The traffic volumes from the first three months of 2020 are presented Table 19. The color coding presents the volume difference per month, from lowest to highest. Seen in the table is the decline of traffic volume going to Svenska Mässan during March, following the above mentioned exhibition cancellations due to covid-19. This suggests that the data collection system accurately measures the true traffic volumes in Svenska Mässan and reflects the general covid-19 situation in Sweden in March 2020. In four out of six measuring points, declines can already be seen from January to February. This might be due to the global covid-19 situation and actions taken in countries outside of Sweden. It is very likely that the decline, in especially international visitors, could already be seen in February due to the covid-19 precautions.

	January	February	March	Total	%	Coding
1	4556	4047	3526	12129	9,58	Lowest
2	4401	5754	3588	13743	10,86	Mid
3	3692	2689	1508	7889	6,23	Highest
4	7159	4956	3755	15870	12,54	
5	29300	27909	15077	72286	57,12	
6	1796	1859	986	4641	3,67	
				126558	100	

Table 19. All traffic from January 2020 to March 2020.

As mentioned earlier the relevant measuring points for this research are number 1, 3, 4 and 6. Table 20 presents the data for only the relevant measuring points from January 2020 to March 2020.

	January	February	March	Total	%
1	4556	4047	3526	12129	29,93
3	3692	2689	1508	7889	19,47
4	7159	4956	3755	15870	39,16
6	1796	1859	986	4641	11,45
	17203	13551	9775	40529	100,00

Table 20. Relevant traffic from January 2020 to March 2020.

This clearly shows the decline of vehicle flow at Svenska Mässan. Not unexpectedly March is the month with the fewest observations across all four measuring points. The declining trend is visible already between January and February for measuring points 1, 3 and 4. However, the decline between January and February could also be due to traffic volumes in relation to exhibitions since the motorcycle exhibition was in the end of January. As a specific example, looking at Table 20 for measuring point 4 the number of observations from January to February have decreased with approximately 30,8% and from February to March with another approximately 24,2%. This means that the difference of the number of vehicle observations between January and March for measuring point 4 is 47,5%. Table 21 shows the number of vehicle observations for the month of March 2019. The visible difference in the number of observations between March 2019 and 2020 further supports the effect of covid-19 on the month of March 2020. A comparison for the months January and February of 2019 cannot be done since the data collection system was installed between February and May in 2019.

	January	February	March	Total	%
1	N/A	N/A	4382	4382	41,28
3	N/A	N/A	N/A	0	0,00
4	N/A	N/A	4777	4777	45,00
6	N/A	N/A	1457	1457	13,72
				10616	100

Table 21. Relevant traffic from January 2019 to March 2019.

The current global pandemic and the impact it has had on travel as well as daily business is clearly visible in the sensor data from Svenska Mässan, as total traffic has decreased from January to March by approximately 43,2 %.

<u>Balance of inbound and outbound moves</u>

As the measuring system should measure observations going in and out of the Svenska Mässan premises, the same amount of observations going inbound should also go outbound. In other words, every vehicle that enters the premises should also exit. In Table 22 the data is categorized according to inbound and outbound observations by comparing each measuring point with the variable *Direction*. Total volumes are then compared.

	1	3	4	6	total	balance %
in	28625	817	52911	1258	83611	100
out	19589	17424	1756	19858	58627	**70,12**
unknown	394	138	2501	154	3187	3,81

Table 22. Balance of inbound and outbound moves

Looking at the total number of observations in Table 22 it is obvious that there is a large discrepancy between inbound and outbound observations. There are 24 984 fewer observations going out of Svenska Mässan than there are going in. The outbound observations represent 70,12% of the total inbound observations. In other words, the data is either missing completely or then the observations are incorrect for almost 30% of traffic in Svenska Mässan premises. Adding all the "Unknown" observations to the outbound observations would only increase the outbound observations with 3,8%.

This shows that there is an issue with variable *Direction*. This might be due to reasons that are related to the location of the sensors. As presented before Jain et. al. (2019) suggested that magnetic sensors potentially have trouble observing vehicles standing still. Considering the detailed locations of the sensors as described in an interview by the System Provider the sensor for measuring points 3, 4 and 6 are outside the gates of the entrances or exits and the sensor for measuring point 1 is on the inside of the gate. As measuring point 3 and 6 are only for exiting vehicles, and as the sensors are outside the gate it is likely that vehicles passing these sensors are in continuous motion and do not stop on top of the sensor. However, as measuring point 1 is both exit and entrance and the sensor is located inside the gate the vehicles exiting through this point might have to stand still before getting through the gate. This might cause them to stand still on top of the sensor. Similar scenario could happen for vehicles entering through measuring point 4 where the sensor is located before the gate which might force vehicles to stop on top of it. Table 13 presents the percentages of signal interference for each measuring point and the findings correlate with what has been stated above, that measuring points 1 and 4 potentially suffers from problematic observations due to vehicles stopping on top of the sensors.

Also, it must be noted that as measuring points 3, 4 and 6 are only supposed to go one way they should not have any observations going the opposite direction. However, according to the data set, 4,45% for measuring point 3, 3,07% for measuring point 4 and 5,91% for measuring point 6 are going the opposite direction. These percentages can be combined with the share of "Unknown" observations that are classified as errors presented in Table 23.

	1	3	4	6
total	48608	18379	57168	21270
in	28625	817	52911	1258
in % of total	58,89	4,45	92,55	5,91
out	19589	17424	1756	19858
out % of total	40,30	94,80	3,07	93,36
unknown	394	138	2501	154
unknown % of total	0,81	0,75	4,37	0,72
% unknown + incorrect direction	0,81	5,20	7,45	6,64

Table 23. Percentages of errors for variable Direction *per measuring points.*

5.1.3 Elements of data quality

To make a further analysis of the above presented different percentages of measurement errors per both variable and measuring point, this section will evaluate the data based on the data quality elements suggested by Hazen et. al. (2014) and Turner (2004). Both authors present a total of four elements of analysis, three of which are matching elements; accuracy, timeliness and completeness. These three are complimented by an additional element from each author, consistency and coverage respectively. This analysis will start by looking at each of these elements in relation to the variables *Type, Length, Speed* and *Direction,* that has been analyzed above.

Accuracy per variable

Type: The variable is divided into two groups, light and heavy, based on a weight limit of 3,5 tonnes. Cargo carrying vehicles can be found in each group. This might be due to similar reasons that Taghvaeeyan, & Rajamani (2014) found in their research about magnetic sensor systems struggling to recognize small vehicles from medium vehicles and categorizes them in the same group.

Length: As previously discussed *Length* has an overall measurement error of almost 25%. These are either observations that do not have a value registered for length or have a length over or under what is considered a reasonable vehicle length. Due to the high percentage of measurement error*,* one fourth of the observations are potentially incorrect.

Speed: Speed has an overall measurement error of 3,81% combining blank observations and observations exceeding 40km/h. Further, approximately 94% of observations have registered speeds of less than 20km/h which is considered a reasonable speed driving over the sensors given their location. Based on the above mentioned, it could be argued that the variable *Speed* is a good representation of the observed vehicles true speeds.

Direction: Only looking at the “Unknown” observations, *Direction* has a measurement error of 2,19% which can be considered very low. However, according to the analysis made about this variable and the specifications of the different measuring points presented in Table 23, the reliability of the variable cannot be judged only based on the number of observations that has not been able to identify a direction.

Extra info: Another way to consider the accuracy of the data set can be done by looking at the percentage of "lowered accuracy" observations. According to the data set 40,8% of the observations have experienced a signal interference, and therefore the accuracy and reliability of the data can and should be questioned.

Timeliness per variables

The Data: All observations in the data set have a very precise timestamp presenting the individual date and time when an observation was registered by one of the sensors. The data set containing all observations is updated every hour. This is sufficient for the current use of the data however, this does not enable real-time usage of the data.

Completeness per variables

The Data: Completeness only looks at the degree of data available, therefore by looking at the observations that are either blank or "Unknown" are the ones tampering with the completeness of the data. For the variables *Type, Length, Speed* and *Directions* the percentage of "Unknown" compared to total observations are the following in respective order; 2,99%, 4,79%, 2,99%, 2,19%. Based on these numbers, the degree of completeness for the data set and variables is considered good.

Consistency per variable

The Data: All the observations and variables across the data set are very consistent in their structure and formatting and a clear consistency is visible.

Coverage per variable

The Data: The data is constantly collected without any unplanned disruptions. The area and position that the sensors observe does not change. However, the analysis has revealed some inconsistencies that might be due to signal interference especially related to the variable *Direction* which suggests that magnetic signals from other areas than the one being observed might interfere with the sensor. As this hypothesis cannot be supported studying the data in this data analysis , the conclusion is made that the coverage of the variables is good.

	Type	**Length**	**Speed**	**Direction**
Accuracy	X	X	O	O
Timeliness	O	O	O	O
Completeness	O	O	O	O
Consistency	O	O	O	O
Coverage	O	O	O	O

Table 24. Quality of analyzed variables and measuring points by quality elements.

5.2 Analysis for Svenska Mässan to become more data-driven

According to Höjer, Svenska Mässan wants to better understand how to utilize the collected traffic data. To understand how the data could better fit in their current ways of operations. This section will analyze data based decision making at Svenska Mässan in comparison with theoretical findings. However, the desired insights into the organization's daily operations and its culture was limited due to the covid-19 pandemic. This lack of insight was compensated by a far more extensive data set than the one initially anticipated and therefore this part of the analysis will also have a data based perspective.

The following section will look at the five enabling factors of becoming data-driven suggested by Berndtsson et. al. (2018) and see how these are visible at Svenska Mässan.

Management

“Our wish is to be able to use this data to move away from experience and instead use hard facts.” (Martinsson, 2020)

According to Martinsson and Höjer both of whom would like to see Svenska Mässan becoming more data-driven in certain aspects and processes. Mainly for the purpose of decision making being based on data rather than experience. When it comes to changes, Berndtsson et. al. (2018) states that employees and middle management should be included in this process and be well informed. Looking at the implementation of the current data collection system, Godhe stated in his interview that he knows that the system was installed and its basic function, but he was not involved in the planning. Since the system is not widely used in the organization, this could be one of the reasons why some of the employees and middle management were not included in the process. However, they could have been used as a source of ideas, knowledge and guidance for future development.

Data

The next element suggested by Berndtsson et. al. (2018) is data. The quality of the current data collected has been analyzed earlier in this paper and many of its issues have been pointed out. To truly find benefits and new insights of the data, it should be more reliable and measurement errors minimized. Having data easily accessible could be beneficial for an organization in spreading new insight. These elements should be considered with any future data collection initiative.

Tools

According to the Godhe, Svenska Mässan is currently not utilizing any KPI’s or dashboards. Their operative processes are based on two connected systems, one managing operations and cargo flows and the other managing customer information and relations. For the employees to utilize the current or any future data to possibly create insights they should have access to user friendly tools (Berndtsson et. al., 2018). As a starting point this could mean creating a dashboard, as presented by Sharda et. al. (2014), that would visualize the data in an easy and accessible way to the employees. The dashboard could present KPIs based on the data and thus give the employees a possibility to be involved in using the data. According to Karami et. al. (2017) employees should also be involved in planning the dashboards. Their needs and expectations should be considered so that the information presented through the dashboard would serve them better and support them making decisions. However, the current situation with the data collection system is not developed far enough to be used for creating reliable dashboards. Despite this, the planning for the data collection system and its usage should consider how the data will be used and consider what kind of visualizations they want to create from the data.

Organization

According to Anderson (2015) and Berndtsson et. al. (2018) one of the cornerstones for an organization to become data-driven is their culture. Based on the interviews with Svenska Mässan employees, their culture seems to be open for new ideas and improvements which supports the ideas of the previously mentioned authors.

According to Berndtsson et. al. (2018) being data-driven would require creating a culture of trust where employees can test their ideas freely. This, even if failed, would create some new business insights. This research did not find any elements opposing this sort of culture. One indication of a trusting and open culture can be seen in the interview with Martinsson who described currently ongoing work to analyze different scenarios which includes a wider perspective than just Svenska Mässan premises. This was one of the reasons that the current data collection system was installed. This shows that the organizational culture is supportive and open to new initiatives and changes that are based on data.

Decision process

The final enabling factor for becoming data-driven is the decision process. When the previous elements are in place the decision process investigates how well management is ready to implement the findings based on data in their decision making. If management does not trust the new insights and chooses to ignore both results and findings they undermine the whole process of becoming data-driven and instead give the impression to employees that the insights they have gathered are not taken seriously.

Outside the above mentioned factors when becoming data-driven one should consider the context and business where an organization operates. As mentioned before the exhibition logistics of Svenska Mässan is a very fluctuating logistics operation as some exhibitions only take place every second or third year. This fact is also recognized by Martinsson who acknowledges that they might have to collect data for several years before being able to use it for analyzing specific exhibitions and related cargo flows. Another fact affecting the exhibition logistic operations is; exhibitioners supplying the goods for an exhibition are not only suppliers but also the customers of Svenska Mässan. Due to this, the cargo handling operations need to maintain a certain level of flexibility and Svenska Mässan cannot e.g. invoice for early or late deliveries as stated by Martinsson. Therefore, no matter how data-driven Svenska Mässan becomes, their operations will always have this perspective of flexibility and ad hoc adjustment of procedures given the nature of exhibition logistics. This is also recognized by Godhe who states that they will always try to fix problems when they arise. Further supporting this statement, Martinsson says, as long as the cargo arrives at the scheduled date it is all right for them. The procedure of being flexible seems to be a requirement as Svenska Mässan must be able to get the cargo unloaded and delivered to the exhibition hall in order to properly serve their customers.

Becoming data-driven, including having the correct culture, organizational structure, competencies, tools and knowledge might be a very complicated process as e.g. presented in Figure 2 showing the circular nature of Business intelligence practices. However, after this process has been initiated the organization should reach a state where they can find simple solutions based on the data they collect and analyze and make well informed decisions to gain competitive advantages as suggested by Sharda et. al. (2014). Finding simple solutions is something that, according to Martinsson, Svenska Mässan wants to do. He believes that simple solutions are often the best ones and these solutions could be reached by data utilization in decision making. It can also be seen in his interview that he believes in investing resources into projects that can yield profit and enable growth in the future and connecting this to the data collection system, after the groundwork for becoming data-driven has been made the organization should be able to create new insights to use for simple ideas and solutions that will in the long run benefit the operations and the entire organization.

6 Discussion

6.1 Data collection method

After having analyzed the data the researchers believe that it should not be used for exhibition logistics purposes. However, the data is found qualified for being used for decision making regarding the facilities and other planning affecting movement of vehicles circulating the premises. The reason why it is not believed that the data is good enough for the purpose of exhibition logistics is the high amount of signal interference and measurement errors in the data. When planning an exhibition, it is important to know e.g. the type of vehicles arriving at the exhibition center in order to utilize all the loading and unloading space effectively. According to Godhe and Martinsson one of the main bottlenecks is the loading and unloading space of Svenska Mässan. Not being able to determine from the data if "light" vehicles are personal vehicles from either guests of the hotel, employees working at Svenska Mässan or if they are light cargo vehicles there to make a delivery, is problematic when planning an exhibition. Also, a signal interference being identified in 40,8% of the total observations is affecting one of the four main variables of the data set. Since the System Provider stated that, if an observation has experienced a signal interference there is a high chance of a measurement error in one of the variables *Type, Length, Speed* or *Direction*. There is no correlation between the variables, meaning that a vehicle that is measured to a length of 16 meter can still be classified as a "light" vehicle (a weight under 3,5 tonnes). This also jeopardizes the amount of trust that should be put into the system while planning and making decisions in regard to larger operations heavily impacted by vehicle flow and movement, like exhibitions.

6.2 Potential Improvements

The placement of the sensors has a significant impact on the quality of the data. In the case of Svenska Mässan, measuring point 4, "*Infart Lisebergssidan*" is located outside the entrance and according to Martinsson, there is a possibility that some of the observations from this location are bicycles passing over the sensor. Also based on findings from literature it is recommended to locate the sensors so that vehicles do not need to stop on top of them and thus avoid possible measuring errors. Therefore, really studying and putting thought into the placement of the sensors is vital for a well-functioning collection system.

One variable that could be useful for Svenska Mässan to be able to utilize traffic data for exhibition logistics operations is a tracking id. Having a tracking id for the vehicles entering the premises would make it possible to know when and from where they entered, how long they have been inside and which way they exit. This data could be used for tracking primary cargo flows as well as estimating loading and unloading times. Our suggestion for this is to implement an LPR system as mentioned and explained above in the literature section. This could potentially improve the efficiency of their operations, allowing for better data to base decisions on. As mentioned above, one of the identified bottlenecks is the loading and unloading area, and as of now according to Martinsson, they have an estimated time for the loading and unloading process for each vehicle entering the establishment. However, Martinsson also mentioned that the goods in regard to exhibition logistics can differ for each customer. Therefore, using tracking id, not only to determine the number of vehicles going inbound and outbound of the establishment, which is an issue with the current system, it could also gather data to create a more accurate estimate of the actual time for loading and unloading. A tracking id could also help in streamlining the other bottleneck mentioned by Martinsson, the entrance. If an LPR system would recognize the vehicle at the entrance the current registration of incoming deliveries could possibly be automated and thus made more efficient. This connected with more detailed data about the cargo carried by a delivery vehicle could be used for faster and more fluent labelling of pallets and thus also contribute in eliminating the other bottleneck.

One solution that could really improve the logistics and exhibition management of Svenska Mässan, would be to have a consolidation center that they could fully control and manage. The distribution and consolidation center they use today at Gullbergsvass is managed and owned by a third party logistics company. If Svenska Mässan could completely manage the cargo flow to their exhibition center in the city, they could make all the incoming deliveries to this consolidation center and the goods could be registered, labelled and better consolidated. This would be a solution that could eliminate the bottlenecks at Svenska Mässan, for the loading and unloading process, and instead steer the data collection system towards other purposes.

The question it all comes down to is cost. System development, implementation, integration etc. all comes at a cost. There is a tradeoff between the potential benefits of the data gathered and cost. As mentioned in the introduction, companies are profit maximizing, and keeping costs low is of great importance. The three potential improvements and solutions mentioned for the purpose of exhibition logistics; 1. Rethinking the placement of the sensors for the current system to improve the accuracy as well as reevaluate the variables e.g. being able to exclude valet parking and personal vehicles from light cargo vehicles. 2. Changing the current system into a camera based system using LPR technique. 3. Having access to a fully controlled and managed consolidation center in proximity to their main establishment that is located in the city center. They all come at a cost, therefore it is necessary to make a detailed plan and an extensive cost calculation before making any larger change to the current system and way of thinking.

Data can be hard to understand, and the way businesses have developed, data and information flows becoming more important both internally and externally with other businesses. The complexity of having too much data is a real issue, if not having the knowledge or competence of knowing how to utilize it. “Try to make it as simple as possible, that is what is so hard.” (Martinsson, 2020)

7 Conclusion

Research Question 1.

Can the traffic data collection system be utilized for exhibition logistics?

After analyzing the data from the traffic data system, this paper concludes that it should not be utilized for exhibition logistics. Reasons for this being; the high percentage of identified and specified measurement errors, the data's lack of distinguishing between different vehicle types, the placement of the sensors and the debatable reliability of the variable data output from specifically; *Type*, *Length* and *Direction*.

However, the researchers believe that the data can be utilized for decision making related to planning where vehicle movement in proximity to Svenska Mässan's facilities are one of the main factors to consider, e.g. infrastructural planning.

Research Question 2.

How should the current data collection system and data utilization be developed in order for Svenska Mässan to become more data-driven in logistics operations and planning?

Firstly, for them to be able to utilize the current data collection system to become more data-driven they need to improve the quality of the data. One possible solution would be to add a tracking id, that would enable them to track individual vehicles in the establishment and to trace the cargo flow through the operations. A tracking id could potentially improve the data through actually knowing at what time a specific vehicle entered and exited the area and therefore eliminating the imbalance between inbound and outbound observations. Additionally, the tracking id could be used to increase the efficiency of loading and unloading by studying throughput time and improving time slot planning, thus eliminating the current bottleneck.

Once they have improved the data collection quality and considered additional variables, they should include their employees in planning how to utilize data. Then use it together with the employee's inputs and visions to create alerts, KPI's and visualization for future decision making and to improve logistics processes and planning. They should also take into consideration that becoming data-driven is an organization-wide commitment where everyone should be included and provided with the right tools for actually enabling data usage in decision making.

No matter how exceptional the data collection method is, the most important variable for an organization to become data-driven and achieve a competitive advantage is the personnel.

7.1 Future Research

Further research of implementing and analyzing traffic data collection systems within exhibition logistics is necessary to determine which type of traffic data collection systems are most suitable for exhibition logistics. Mentioned in this paper are two types of traffic data collection systems, sensor based and camera based systems. To establish which of the two systems is the superior one can be very individual and company specific depending on the desired outcome, however both methods and their accuracies should be studied to better conclude which method is more suitable for the exhibition logistics.

Further research on different exhibition centers for benchmarking purposes would also be useful. The operations between different exhibition organizations are bound to differ and therefore investigating other establishments in the industry might bring valuable insights into the topic of using traffic data in exhibition logistics.

Regarding this paper and the discussed potential improvements of the traffic system, from a perspective of Svenska Mässans exhibition and logistics operations, further research is needed to find empirical evidence supporting other variables that could be beneficial for utilizing traffic data for these kinds of operations. Mentioned in the discussion above was tracking id, and how using a camera based system could allow for an LPR system to act as a tracking id for vehicles.

8 References

Ain, N., Vaia, G., DeLone, W. H., & Waheed, M. (2019). Two decades of research on business intelligence system adoption, utilization and success – A systematic literature review. Decision Support Systems. Vol. 125.

Alvesson, M. (2003). Beyond neopositivists, romantics, and localists: A reflexive approach to interviews in organizational research. Academy Of Management Review, 28(1), pp. 13-33.

Al-Smadi, M., Abdulrahim, K., & Abdul Salam, R. (2016). Traffic Surveillance: A Review of Vision Based Vehicle Detection, Recognition and Tracking. International Journal of Applied Engineering Research, Vol. 11, No. 1, pp.713-726.

Anderson, C. (2015). *Creating a Data-Driven Organization.* O'Reilly Media.

Bakhtan, M., Abdullah, M., & Rahman, A. (2016). A Review on License Plate Recognition System Algorithms. International Conference on Information and Communication Technology (ICICTM). Kuala Lumpur.

Berndtsson, M., Forsberg, D., Stein, D., & Svahn, T. (2018). Becoming a Data-Driven Organisation. *Research-in-Progress Papers. No. 43*.

Bopp, C., Harmon, E., & Voida, A. (2017). Disempowered by Data: Nonprofits, Social Enterprises, and the Consequences of Data-Driven Work. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp. 3608–3619). Denver: Association for Computing Machinery.

Borstell, H. (2018). *A Short Survey of Image Processing in Logistics.* Magdeburgo. Institute of Logistics and Material: In: Schmidtke, N., ed. 11th International Doctoral Students Workshop on Logistics.

Borstell, H., & Reggelin, T. (2019). Towards Virtual Commissioning of Image-based Information Systems for State Detection in Logistics. *IFAC-PapersOnLine*, Vol. 52, No. 13, pp. 2463-2470.

Bottero, M., Dalla Chiara, B., & Deflorio, F. P. (2013). Wireless sensor networks for traffic monitoring in a logistic centre. *Transportation Research Part C*, Vol. 26, pp. 99-124.

Bryman, A., Bell, E. (2015). Business Research methods. 4th ed.

Brynjolfsson, E., Hitt, L. M., & Kim, H. H. (2011). *Strength in Numbers: How Does Data-Driven Decision making Affect Firm Performance?* Last accessed February 21st 2020: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1819486

Cheung, S., Coleri, S., Dundar, B., Ganesh, S., Tan, C.-W., & Varaiya, P. (2005). Traffic Measurement and Vehicle Classification with Single Magnetic Sensor. *Journal of the Transportation Research Board*, No. 1917, pp. 173–181.

Closs, D. J., Goldsby, T. J., & Clinton, S. R. (1997). Information technology influences on world class logistics capability. *International Journal of Physical Distribution & Logistics Management*, Vol. 27 No. 1, pp.4-17.

Collis, J. & Hussey, R. (2014). Business Research: A Practical Guide for Undergraduate and Postgraduate Students.

Collis, J. & Hussey, R. (2017). Business Research: A Practical Guide for Undergraduate and Postgraduate Students. 4th ed.

European Parliamentary Research Service Blog. (2014, April 10). Weights And Dimensions Of Road Vehicles In The EU. Last accessed May 4th 2020: https://epthinktank.eu/2014/04/10/weights-and-dimensions-of-road-vehicles-in-the-eu/

Fastighetsvärlden. (2019, October 8th). Så blir fjärde tornet vid Gothia – högst av alla. Last accessed May 4th 2020: https://www.fastighetsvarlden.se/notiser/bildextra-sa-blir-fjarde-tornet-vid-gothia-hogst-av-alla/

Fredholm, P. (2013). *Logistik & IT: [för effektivare varuflöden].* Lund: Studentlitteratur.

GDPR.EU. (2020). What is GDPR, the EU's new data protection law? Last accessed May 4th 2020: https://gdpr.eu/what-is-gdpr/?cn-reloaded=1

Grabara, J., Kolcun, M., & Kot, S. (2014) The Role of Information Systems in Transport Logistics.

Gupta, N., Tayal, S., Gupta, P., Goyal, D., & Goyal, M. (2017). A Review: Recognition of Automatic License Plate in Image Processing. Advances in Computational Sciences and Technology, Vol. 10, No. 5, pp. 771-779.

Halawa, F., Dauod, H., Lee, I., Li, Y., & Yoon, S. W. (2019). Introduction of a real time location system to enhance the warehouse safety and operational efficiency. *International Journal of Production Economics*. Vol. 224.

Hallowell, R. (1996). The relationships of customer satisfaction, customer loyalty, and profitability: an empirical Study.

Hazen, B., Boone, C., Ezell, J., & Jones-Farmer, A. (2014). Data quality for data science, predictive analytics, and big data in supply chain management: An introduction to the problem and suggestions for research and applications. *International Journal of Production Economies*, Vol. 154, pp. 72-80.

Höjer, N. (2019). Personal communication, 27 December.

ISO. (2020). *Date and time format*. Last accessed April 27th 2020: https://www.iso.org/iso-8601-date-and-time-format.html

Jain, N. K., Saini, R. K., & Preeti, M. (2019). A Review on Traffic Monitoring System Techniques. In S. Rawat, K. Ray, R. Saini, & T. Sharma, *Soft Computing: Theories and Applications. Advances in Intelligent Systems and Computing* (pp. 569-577). Singapore: Springer.

Kallio, H., Pietilä, A.-M., Johnson, M., & Kangasniemi, M. (2016). Systematic methodological review: developing a framework for a qualitative semi-structured interview guide. *Journal of Advanced Nursing*, Vol. 72, No. 12, pp. 2954–2965.

Karami, M., Langarizadeh, M., & Fatehi, M. (2017). Evaluation of Effective Dashboards: Key Concepts and Criteria. *The Open Medical Informatics Journal*, Vol. 11, pp. 52-57.

Karpis, O. (2013). Wireless Sensor Networks in Intelligent Transportation Systems. International Journal of Modern Engineering Research, Vol. 3, No. 2, pp. 611-617.

Kochlán, M., Hodon, M., Cechovic, L., Kapitulík, J., & Jurecka, M. (2014). WSN for Traffic Monitoring using Raspberry Pi Board. Computer Science and Information Systems, Vol. 2, pp. 1023–1026.

Krizan, L. (1999). Intelligence Essentials for Everyone. Joint Military Intelligence College.

Lewis, I., & Talalayevsky, A. (2000). Third-Party Logistics: Leveraging Information Technology. *Journal Of Business Logistics*, Vol. 21, No. 2, pp.173-185.

Lunsford, D., & Phillips, P. (2018). Tools Used by Organizations to Support Human Capital Analytics. *Performance Improvement,* Vol. 57, No. 3, pp. 6-15.

Mani, A., & Panicker, N. J. (2019). An Efficient Method for License Plate Detection and Recognition using OCR. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, Vol. 5, No. 3, pp. 540-547.

Mechanic Base. (2019, August 23rd). *Average Car Length – List of Car Lengths*. Last accessed April 27th 2020: https://mechanicbase.com/cars/average-car-length/

Mortier, R., Haddadi, H., Henderson, T., McAuley, D., & Crowcroft, J. (2014, October 11). Human-Data Interaction: The Human Face of the Data-Driven Society. Last accessed May 4th 2020: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2508051

Onwuegbuzie, A. J., Leech, N. L., & Collins, K. M. (2010). Innovative Data Collection Strategies in Qualitative Research. The Qualitative Report, Vol. 15, No. 3, pp. 696-726.

Pandey, M., & Mishra, G. (2019). Types of Sensor and Their Applications, Advantages, and Disadvantages. In A. Abraham, Emerging Technologies in Data Mining and Information. Springer Nature Singapore.

Patel, R. & Davidson, B. (2011). Forskningsmetodikens grunder: att planera, genomföra och rapportera en undersökning, 4 Edition.

Power, D. J. (2008). Understanding Data-Driven Decision Support Systems. Information Systems Management, Vol. 25, No. 2, pp. 149–154.

Puklavec, B., Oliveira, T., & Popovic, A. (2014). Unpacking Business Intelligence Systems Adoption Determinants: An Exploratory Study of Small and Medium Enterprises. Economic and Business Review. Vol. 16, No. 2, pp. 185-213.

Puranic, A., Deepak, K. T., & Umadevi, V. (2016). Vehicle Number Plate Recognition System: A Literature Review and Implementation using Template Matching. *International Journal of Computer Applications*, Vol. 134, No.1, pp.12-16.

Regeringskansliet. (2020, March 11th). Förordning om förbud mot att hålla allmänna sammankomster och offentliga tillställningar. Last accessed May 4th 2020: https://www.regeringen.se/artiklar/2020/03/forordning-om-forbud-mot-att-halla-allmanna-sammankomster-och-offentliga-tillstallningar/

Roeksukrungrueang, C., Kusonthammrat, T., Kunapronsujarit, N., Nunjaruwong, T., & Chivapreecha, S. (2018). An Implementation of Automatic Container Number Recognition System. 2018 International Workshop on Advanced Image Technology (pp. 1-4). Chiang Mai: IEEE.

Sanders, N. R. (2014). Big Data Driven Supply Chain Management. Pearson Education.

Sharda, R., Delen, D., & Turban, E. (2014). Business Intelligence: A Managerial Perspective on Analytics. Pearson Education.

Sifuentes, E., Casas, O., & Pallas-Areny, R. (2011). Wireless Magnetic Sensor Node for Vehicle Detection With Optical Wake-Up. *IEEE Sensors Journal*, Vol. 11, No. 8, pp. 1669-1676.

Sportsmobile. (2020). *Van Dimensions & Comparisons*. Last Accessed April 27th 2020: https://sportsmobile.com/van-dimensions/

Sprague, R. H. (1980). A Framework for the Development of Decision Support Systems. *MIS Quarterly*, Vol. 4, No. 4, pp. 1-26.

Svenska Mässan, (2020a). Om oss, utforska oss. Last accessed May 4th 2020: https://svenskamassan.se/utforska-oss/om-oss/oversikt/

Svenska Mässan, (2020b). Om oss, organisation. Last accessed May 4th 2020: https://svenskamassan.se/utforska-oss/om-oss/organisation/

Svenska Mässan. (2020c, March 11th). *Meddelande från Svenska Mässan Gothia Towers rörande Folkhälsomyndighetens förslag*. Last accessed May 4th 2020: https://svenskamassan.se/2020/03/meddelande-angaende-covid-19-coronaviruset/

Svenska Mässan (2018) Årsredovisning. Last accessed May 4th 2020: https://svenskamassan.se/uploads/sites/46/2019/04/arsredovisning-svenska-massan-stiftelse-2018-webb.pdf

Taghvaeeyan, S., & Rajamani, R. (2014). Portable Roadside Sensors for Vehicle Counting, Classification, and Speed Measurement. *IEEE Transactions on Intelligent Transportation Systems*, Vol 15, No. 1, pp. 73-83.

Tiwari, B., Sharma, A., Singh, M. G., & Rathi, B. (2016). Automatic Vehicle Number Plate Recognition System using Matlab. *Journal of Electronics and Communication Engineering*, Vol. 11, No. 4, pp. 10-16.

Turner, S. (2004). Defining and Measuring Traffic Data Quality. *Journal of the Transportation Research Board*, No. 1870, pp. 62-69.

Verma, A., Sharma, M., Hebbalaguppe, R., Hassan, E., & Vig, L. (2016). Automatic Container Code Recognition via Spatial Transformer Networks and Connected Component Region Proposals. *15th IEEE International Conference on Machine Learning and Applications* (pp. 728-733). Anaheim: IEEE.

Wang, Q., Zheng, J., Xu, H., Xu, B., & Chen, R. (2018). Roadside Magnetic Sensor System for Vehicle Detection in Urban Environments. *IEEE Transactions on Intelligent Transportation Systems*, Vol. 19, No. 5, pp. 1365-1374.

Wang, X,. (2013). Model Innovation of Exhibition Logistic Supply Chain. Volume 397-400, pp. 2672-2676.

Wexler, S., Shaffer, J., & Cotgreave, A. (2017). *The Big Book of Dashboards.* Wiley.

Wu, J., Guo, J., & Li, X. (2019). Key Methods of Recognizing Container Number Automatically Using Video Stream in Intelligent Tally. *Proceedings of 2019 Chinese Intelligent Systems Conference* (pp. 10-17). Springer.

Yan, Q., & Zhang, Q. (2015) The Optimization of Transportation Costs in Logistics Enterprises with Time-Window Constraints. Advanced Dynamic Simulations in Transportation. Vol. 2015.

Yeoh , W., & Koronios, A. (2010). Critical Success Factors for Business Intelligence Systems. Journal of Computer Information Systems, Vol.50, No.3, pp. 23-32.

Yin, R. (2018). Case Study Research and Applications, 6th edition.

Zarnescu, A., Ungurelu, R., Iordache, A. G., Secere, M., & Spoiala, M. (2017). Crossroad Traffic Monitoring Using Magnetic Sensors. IEEE 23rd International Symposium for Design and Technology in Electronic Packaging (pp. 413-418). Constanța: IEEE.

Zhang, S. (2012). Research on Exhibition Logistics Information System Designing. Advanced Designs and Researches for Manufacturing. Vol. 605-607, pp. 479-482.

Appendix I: Interview guide

Info:

-The purpose of the research and the interview
-Participation is voluntary, and the participant has the right to withdraw at any time
-Participant has the right to confidentiality and anonymity

Can we record the interview?
Can we use your name and job title in the thesis?

Short description of your job and title.
How long have you worked for Svenska Mässan?

Questions:

Current operations:
An exhibition is coming up in a few weeks:
What is the first thing you do? (Operational wise)
Please go through the different tasks associated with an exhibition from your end.

Please describe the current delivery system.
Consolidated deliveries and customer deliveries.

Job Specific questions:
How do you plan the time slots?
Time slots give for both unloading and loading?
What elements are considered when planning the time slots? (cargo characteristics, position in the exhibition…)
What kind of issues are there with time slots/deliveries? (delays, queues)
How much flexibility is in the time slot plan?
What happens if a delivery is early or late?

The timeline of the delivery operations?
Does unloading and loading happen at the same time?
Is there real time optimization of unloading/loading? (in case of queue, based on vehicle size for example)
What kind of issues are there with the delivery operations?
How do they organize/consolidate the cargo?
How are the fill rates with consolidated cargo?
Do customers deliver cargo whenever?

What kind of agreements on deliveries do you have with the exhibitioners?
How do you communicate with the delivery party?
Who does the deliveries? (familiar people)

What kind of differences do you see between domestic deliveries or deliveries from abroad?
What kind of issues are there with communication with the delivering party/exhibitor?

Usage of data and IT:
What systems do you use in your work?
Where do you get data that decisions are based on?
    How do you use the data?
    Data availability, usability, accuracy.
Do you feel that something is missing?
    Some system for a certain operations?
    Some information for decision making?
Do you see some elements that could be done with a system instead of manually?
    What benefits would this give you?

Current operating systems:
    When did you start using this?
    How good is it?
    Do you remember how it was at the beginning?
    What do you do with it?
    What data does it give you? Information?
    Do you extract reports from the system?
    Do you insert data in the system?

The current sensor system.
    Please shortly describe the current camera system.
    For what purpose did you get the system?
    What was the original plan/requirements for the system?
    Was it purchased externally?
    How much competence with the system is there in house?
    How does it work?
    What it can/cannot do?
    What kind of data comes out?
How is it analyzed?
What is the resolution of the cameras?
    How does it recognize vehicle size?
        Does it recognize regular car with a trailer as heavy traffic?
    How does it recognize speed?
    How accurate is the recognition?
Could it recognize registration plates?
    The 10% gab in inbound and outbound.
How do you separate in- and outbound from the data?

Extras:
What is the biggest operational risk associated with your work?
What is bigger issue, not respecting the time slot or the unloading/loading taking too long?

Current strategy.

How is IT/Data usage/(business intelligence) included in the strategy?

Is the current strategy data usage friendly?

Future plans/future vision for Svenska Mässan