Utilizing data analytics to forecast demand in the road transportation industry

Lappeenranta–Lahti University of Technology LUT

Master's degree in industrial engineering and management

2022

Aleksi Sahlberg

Examiners: Professor Ville Ojanen

Associate Professor Kalle Elfvengren

ABSTRACT

Lappeenranta–Lahti University of Technology LUT

LUT School of Engineering Science

Industrial Engineering and Management

Aleksi Sahlberg

Utilizing data analytics to forecast demand in the road transportation industry

Master's thesis

2022

65 pages, 10 figures, 1 table and 1 appendix

Examiner: Professor Ville Ojanen

Keywords: Logistics, transportation, data analytics, forecasting

In recent years the transportation industry has started to go through a change at the same time as knowledge management has gained ground and easy access to different tools has increased. The transportation industry is also heavily competitive, so especially small and medium-sized businesses are forced to seek creative solutions to be able to compete with bigger and global corporations. Data analytics has created more changes for smaller companies to increase the utilization rates of the available resources, like personnel and vehicle fleet. Maximizing the utilization of existing resources brings scalability, thus enabling more considerable growth with fewer investments.

In this thesis, two different tools are developed for the case company. These tools are aimed to help the company with the problems mentioned above. The tools are built with close cooperation with the company's representative to match the company's needs. The tools are built following the design research methods. Tools are built using spreadsheet and data visualization tools and they are directly connected to the company's SQL database. As additional research, it's recommended to search for ways to build big data-based methods to utilize the available data even further.

This study showed that building these tools enables direct savings on labor, which allows better utilization of the available resources. This is gained by bringing the existing information to be better available in a format it's needed in. Also, it was found that the available transportation fleet can be utilized better with the created tools, but confirming that requires further tracking of the performance of the tools.

TIIVISTELMÄ

Lappeenrannan–Lahden teknillinen yliopisto LUT

LUT Teknis-luonnontieteellinen

Tuotantotalous

Aleksi Sahlberg

Data-analytiikan hyödyntäminen kysynnän ennustamiseen kuljetusalalla.

Diplomityö. Lappeenrannan-Lahden teknillinen yliopisto LUT, Tuotantotalous.

2022

65 sivua, 10 kuvaa, 1 taulukko ja 1 liite

Tarkastaja: Prof. Ville Ojanen

Avainsanat: Analytiikka, Kuljetusala, Logistiikka, Ennustaminen

Viime vuosina kuljetusala on kokenut murrosta samalla, kun tietojohtaminen on yleistynyt sekä tiedon käsittelyyn helposti saatavilla olevat työkalut ovat lisääntyneet. Kuljetusala on myös erittäin kilpailtu, joten varsinkin pienet sekä keskisuuret yrityksen joutuvat etsimään luovia ratkaisuja pystyäkseen kilpailemaan kansainvälisten yritysten kanssa. Data-analytiikka on luonut myös pienille yrityksille mahdollisuuksia hyödyntää saatavilla olevia resursseja, eli muun muassa kalustoa sekä henkilökuntaa paremmin. Olemassa olevan kaluston maksimaalinen hyödyntäminen vähentää tarvetta lisähankinnoille yrityksen kasvaessa. Tämän vuoksi data-analytiikkaa hyödyntämällä yritykset voivat saavuttaa isompaa kasvua pienemmillä investoinneilla.

Diplomityössä kehitetään yrityksen käyttöön kaksi erilaista työkalua, jotka auttavat löytämään ratkaisut edellä mainittuihin ongelmiin. Työkalut toteutetaan yrityksen tarpeisiin suunniteltuina ja niiden kehitys tapahtuu läheisessä yhteistyössä yrityksen edustajan kanssa. Työkalujen rakentamista lähestytään suunnittelututkimuksen keinoin. Työkalujen rakentamiseen käytetään taulukkolaskenta-, sekä visualisointiohjelmia ja työkalut ovat suoraan yhteydessä yrityksen SQL-tietokantaan. Jatkotutkimuksena suositellaan syvempää perehtymistä big data menetelmien soveltamiseen yrityksen sisällä.

Diplomityössä havaittiin, että työkalujen kehitys mahdollistaa välittömästi säästöt henkilökunnan työajassa, jonka kautta olemassa olevia resursseja pystytään hyödyntämään tehokkaammin. Tämä saavutetaan paremmin saatavilla olevalla tiedolla, joka on työntekijöiden käytettävissä oikeassa muodossa, juuri silloin kun he sitä tarvitsevat. Lisäksi työssä havaittiin, että olemassa oleva kuljetuskalusto voidaan hyödyntää paremmin, mutta rakennettujen työkalujen hyöty tällä alueella pystytään todistamaan vasta, kun työkalut ovat olleet käytössä.

ACKNOWLEDGEMENTS

I would like to thank Transport Jylhä and their representatives for their patience and support during my prolonged research process. Thanks for the on-demand support for Professor Ville Ojanen, especially during the final stretches of the project. Big thank you to my family, who didn’t lose faith at least fully on my graduation. My biggest thanks go to my partner, who supported me through the whole project.

ABBREVIATIONS

NDA	Non-Disclosure Agreement
PCR	Polymerase Chain Reaction
COVID	Covid-19
SDLC	Software Development Life Cycle
CLR	Critical Literature Review
SVM	Support Vector Machine
SWOT	Strength, Weakness, Opportunities and Threats
DCNN	Deep Convolutional Neural Network
TMS	Transport Management System
POS	Point of Sale
DSR	Design Science Research
GPD	Gross Domestic Product
BI	Business Intelligence
PDCA	Plan-Do-Check-Act cycle
EC	E-commerce
PDC	Portable Document Format
XML	Extensible Markup Language
JPG	Joint Photographic Group
ITS	Intelligent Transport System
SAP	System Applications and Products
GPS	Global Positioning System
SQL	Structured Query Language

Table of Contents

1 Introduction

In recent times, it has been noticed that transportation businesses have achieved great lengths of success, as it is a well-known fact that moving goods from one location to another is essential. Thus, the land logistics business is necessary for the economy to satisfy people's needs and maintain communication. This business has many positive sides and is flourishing at a high pace, but it has many challenges. One major challenge is that trucks are sometimes forced to travel distances without cargo. This makes the companies lose profits and is unsuitable for them; thus, they need a solution to mitigate this problem. This study focuses on balancing import and export transportation volumes. The main scope of the study is to find a solution that helps companies to minimize the distance that trucks drive without cargo during their operational lifespan. The goal is to find a data-driven solution to help optimize the sales, planning, and monitoring processes to spot an upcoming dead cargo zone in advance and to avoid traveling without cargo. In this thesis, the study scope focuses on the client company, Transport Jylhä. In this study, the client company is referenced as “Company” or as “Jylhä.”

Truck drivers play a crucial part in distributing goods around the globe. Since the onset of the pandemic, our understanding of the vital role they play in distributing necessities has increased. Over 70% of the nation's freight is moved by truck drivers, often known as OTR drivers (over-the-road), and long-haul truckers (Reddy, 2019).

Working in such a high-risk environment requires extreme maturity and tolerance. As with any profession, a truck driver's working conditions can change depending on where he or she is driving and for which company. Long-distance truck drivers who operate under over-the-road laws are subject to the most stringent standards. Local truck drivers have similar service laws, and their employers may impose further restrictions (Babar, 2018).

The concept of "Big Data" has been widely discussed in numerous sectors. However, what role does big data play within the transportation part of any supply chain? The bottom line is that big data is essential if the aim is to cut costs on freight while expanding business with

data analysis. The power of information makes even the impossible possible, and the distinction between large amounts of data and big data should be recognized. Historically, conventional technologies, like those used in warehouse and distribution systems, have provided analytics on massive amounts of data from such sources. However, big data goes beyond this by enabling businesses to include data from non-traditional sources, like the web, TMS software, and POS systems. This results in an even more extensive data set that needs to be evaluated. Big data stands out from traditional data sources, not only in terms of its sheer size but also because of the range of information it provides. It is now possible to gather not only structured data but also unstructured data in text, photos, and videos, amongst other media (Ford, 2019). Likes, tweets, brand reviews, comments online, blog activity, and other forms of online engagement are being considered in addition to traditional metric analysis. This data can give the vendor insight into their consumers' minds, desires, and timing.

Acceleration is the third key differentiator of big data. Companies used to be content with checking their data once a day, once a week, or once a month, but not so in the present day. Businesses need access to up-to-the-minute information to plan warehouse inventories, select the most suitable carrier, and avoid late delivery. Despite the apparent complexity of all this data, the first step in effectively using big data is identifying the problems that need to be tackled or the parts of the supply network that could benefit from enhancements. After deciding where to apply big data, data collection may begin. Determine what to use; human entry, automated collection boxes, or a hybrid of the two. When sufficient data is collected and analyzed, the next step is to implement the suggested changes and keep an eye on how they are faring.

Since it is established that "big data" is a real thing, next can be discussed how it can help the transportation production chain and save resources. Forecasting using vast amounts of data is what predictive analytics is all about. Big data integrates information from several databases for prediction and instantaneous action.

Delivery and pick-up errors may be reduced, shipment timetables can be optimized, and backhaul routes can be optimized with transportation analytics. In addition to assisting with demand and stockpile forecasting, big data may also aid in inventory management. Just think about how much time and money would be saved if replacement stock was on its way to its new location before the current stock was depleted. It would be easier to answer questions

like, "How many shipments were made from point A to point B in June of last year?" using this data, but it would also be easier to optimize shipping processes for greater efficiency and cost savings. TMS, which monitors and evaluates freight performance, will be the primary source of transportation-related big data. Regarding transportation analytics, TMS is a gold mine of information. On-time pickups, transport mode choice, outward and inbound prices as a percentage of purchasing and selling, and the average time in transit are the five most significant variables to keep an eye on. To recap, big data is a tool for furthering a business's mission and vision. Big data will aid management in making educated decisions, whether the primary goal is to cut costs, boost efficiency, or boost sales. Big data is still most useful when objectives are defined, measurements are understood, and decisive action is taken in a timely manner (Witten, 2016).

1.1 Background of the research

Maximizing the utilization rates of the transportation units has a big impact on the carbon footprint of the company. When a company minimizes the dead freight, it leads to less unnecessary pollution produced by the trucks in the form of greenhouse gasses. Dead freight is the fee charged when a ship is chartered in its entirety but only a portion of its cargo is sold. The fee is based on the percentage of available space. The utilization rate is increased by many factors. The most important is the way the transport planners plan the loads for each trailer. This is in good shape but could be improved by better and more on-demand delivery of the booking data. Another one is the number of bookings. Suppose the company has reserved, for example, four transportation units to be used on a given day that requires enough cargo to transport on that day. This factor is affected by sales, sourcing, and marketing. The existing customer base offers most of the bookings, but in events with less demand than forecasted, remaining orders need to be sourced from somewhere else if the utilization rate is to be maximized. The third aspect is demand forecasting. When that can be done accurately, the supply of cargo units can be adjusted. The type of cargo has a small effect. Most of the cargo that the Company transports is with fixed and short lead times, meaning the goods need to be delivered within the minimum time after pickup. However, if the cargo is bulk, general cargo with an open delivery date can be used to fill the cargo units when there is room left after the most urgent deliveries. The phrase "generic cargo" is used in the logistics industry to describe items that do not require disassembly before shipment.

Pallet, package, barrel, or box—any of these will work. Groupage freight refers to shipments when smaller packages are gathered from various senders and merged to make one larger package (S Ahmed, 2020).

1.2 Research Objectives

The objective of this research is to find ways to reduce the workload by finding the most profitable routes for trucks and increasing the utilization rate as much as possible. Because the company has a lot of data gathered from the year 2016 onwards, it is possible to build comprehensive models and predictions. The data is not yet utilized in a way that would be beneficial for the company, even though it would be possible. This is mostly due to there being no personnel within the company that would see the added value of utilizing the available data or has the resources to act on it timewise.

1.1 Research Questions

Research questions help by paving the way toward the desired aims and objectives. Therefore, before starting any project, a set of research questions are asked. Research questions for this study were selected based on the problems of the case study company and a general need for the transportation industry. This study will answer the following questions:

What are the first financial gains that a road transportation company can have after starting to invest in data analytics?

Is there a need to improve knowledge management in Transport Jylhä?

1.2 Research methods

This study will be conducted using design science methodology. Research on engineering education can gain a great deal from modeling activities carried out in the classroom. Even if the introduction of qualitative research methodologies has not always been met with positive reception, it has been demonstrated that these approaches are beneficial in engineering research. In the discipline of design science, the purpose of qualitative research is to achieve a deeper comprehension of the design procedure. As a result, it provides insights not just into the process of design but also into the finished product. Research in design is another name for the interdisciplinary field of science of design (Hevner, et. al. 2010). Design science research, also known as DSR, is a method for resolving issues that places emphasis on the production of one-of-a-kind goods for the purpose of advancing the common pool of human knowledge. To put it another way, DSR's goal is to enhance the environment in which its solutions are implemented by producing one-of-a-kind artifacts that provide answers to problems that hitherto lacked any solutions that were considered to be satisfactory.

The results of DSR are not limited to the newly designed artifacts themselves; rather, they also include the design knowledge that provides a deeper understanding of why the newly designed artifacts improve or disrupt the relevant application contexts through design theories. This deeper understanding is provided by design knowledge. This chapter will provide a high-level introduction to DSR ideas for the reader to have a better understanding of the DSR case studies that will be provided in subsequent chapters. Over the course of history, empirical studies, particularly positivist research, have garnered an increasingly significant amount of attention. Following a series of conversations concerning the paradigms and techniques of research, it was determined that interpretive research plays an important part in the discipline of the analytics field. An investigation into the subject to a greater depth revealed that the findings of the critical study are extremely applicable to the field of data and analytics as well. The objective of the field of study known as Design Science Research, also known as DSR, is not to gain a deeper comprehension of the world in its current state, as is the case with the Social and Behavioural Sciences, but rather to devise and put into practice innovative approaches to interacting with the world in order to mold and improve it. Because of this, DSR reconstructs the world by creating and

researching items that assist individuals in accomplishing their objectives and finding solutions to their difficulties. The main point of selecting this method for this study was to create a tool that addresses the problem directly since there are no ready solutions available that fit the need fully. For this study, an essential part is to create methods and tools and gain results by analyzing their impact of those.

1.3 Structure of the research

The following steps towards finding answers to the research questions are completed in this study:

- Literature review on the subject
- Gathering data from the personnel of the company
- Making a quantitative analysis on how basic analysis tools help with route optimization and resource utilization.
- Making an analysis on how much the time spent on the analysis tools costs the company and how much it would cost to have someone spend their work time on that.

Once all the data is gathered according to requirements and needs, a detailed analysis will be performed. This analysis will uncover details such as which tools are better for routes and resource utilization, how much time will be spent on these tools to gather data on them, and how much money the companies will pay to develop these tools. Upon having all this data, it will need to be put in graphs and documentation so it can be presented in theory and practice through software and tools.

2 Transportation industry and examples of utilizing analytics

2.1 Transportation industry in Finland

Finland has a strong tradition of road transportation. Because Finland is a country with long distances and low population density, it is common that road is the only viable transportation mode. That has led to a situation in which Finland is the most transport-intensive country within the European Union, proportional to GPD or population (David, 2021). Finland is heavily reliant on road transport; demand leads to supply, and Finland, as a cause, has many small and medium-sized transportation companies. When Finland joined the European Union, and the borders opened for more free movement of goods, it brought in more competition from countries with cheaper labour, and Finnish companies had to adjust to the changed situation.

2.2 Company cases from Europe

In recent years, data analytics has evolved into a technology that has become an ever-increasingly important part of road transportation-focused European businesses. This is especially true for companies in the United Kingdom. The analysis of vast quantities of data enables these organizations to increase their capacity to forecast demand and run their operations more efficiently (Dokic, 2020). Using data analytics might be advantageous to transportation sector organizations in several ways. It might be used, for instance, to forecast traffic patterns and identify system bottlenecks.

After gathering all this information, one is in a much better position to utilize it successfully to create more time- and resource-efficient schedules and itineraries. Data analytics is a

technology that can benefit businesses responsible for transporting people and goods on European highways. In addition, data analytics may be used to monitor the performance of individual vehicles and drivers, which can help detect and fix any problems that may develop during the race (Shahiri, 2015). Combining the usage of this technology with more traditional ways of the problem-solving process to find solutions to complex problems might allow these firms to improve their operations and give their customers the greatest possible experience.

2.2.1 Fuller Logistics

Fuller Logistics is a firm that uses data analytics to increase its forecasting accuracy and the efficacy of its day-to-day business operations. While making company-affecting choices, the organization placed a great deal of trust in the intuition and judgment of its personnel. Also, as a direct result of using data analytics, they made better-informed choices, which has led to improved outcomes (Reddy, 2019). Demand forecasting is one-way Fuller Logistics has profited from data analytics. These advantages have helped the firm in a variety of ways. Because the organization had a solid understanding of historical data and patterns, it was able to create demand forecasting models that accurately predicted future demand. This was made feasible by the organization's dedication to maintaining a state of continuous progress for its existence. They have been able to boost the effectiveness of their company operations, allowing them to meet their consumers' demands better.

Using data analytics has also been included in the inventory management process to increase its overall efficiency and effectiveness. Because the corporation has access to both sets of data, it can make informed choices about the products to create and keep in its facilities if it is aware of both the amount of development in stock and the quantity expected to be needed shortly (Fiore, 2019). The business may make informed judgments if it knows both numbers. Thus, the company can choose which kind of items it will manufacture and stock at its different facilities. Various things are accessible for purchase, providing customers access to these alternatives.

Consequently, the company's operations have become more cost-effective and efficient, increasing the organization's total efficiency. Fuller Logistics greatly enhanced its overall performance because it used data analytics as a formidable weapon. The company's embrace of a data-driven strategy made this feasible (Papamitsiou, 2014). The company's success was probably a result of its application of this strategy. This step was taken to deliver a higher quality service to the customers. Overall, their results have improved as a direct consequence of this element, which is evident across the board.

Due to basing its choices on the data, the company has provided better outcomes, which has benefited the organization and its customers. The business and its consumers have immediately profited from implementing these enhancements, providing the company with instant value (Amin, 2019). This is an excellent demonstration of how data analytics may be applied to improve the efficiency and effectiveness of corporate operations. It is such a perfect illustration because it is fantastic. Compared to previous cases, it's conceivable this one strike as unusually impressive.

2.2.2 Movivan Removal

Movivan Removals, a firm that offers services related to international road transportation, uses data analytics to meet the demands of its customers better, anticipate their needs, and increase the efficiency of its operations. The firm uses this strategy as part of its attempts to expand its customer base and enhance its total revenue. Due to the organization's use of data analytics, it has identified recurring patterns in the activities and requirements of its customers. As a direct result, they can now better anticipate market changes and adapt their company plan (Zadeh, 2020). The company also has seen more significant levels of customer satisfaction and enhanced levels of efficiency, which have directly led to decreased expenses and increased productivity. Movivan Removals is just one example of how businesses can utilize data analytics to enhance internal processes and the whole customer experience. As the technique of data analytics grows increasingly prominent in today's business environment, organizations of all sizes will have the chance to enjoy the many benefits that data analytics provides.

Customer data analysis is a powerful instrument that firms may use to improve the quality of their internal operations and the consumer experience. This may be achieved using clients' information (Papamitsiou, 2014). Utilizing data analytics, Movivan Removals has effectively identified customer activity patterns and demand for their services. This has allowed them to improve their client service. As a direct result, they have improved their capacity to predict future demand and operate their business following their forecasts.

In addition, customer data analysis has reduced the time required to create these forecasts. As a direct result, the company has seen increased customer satisfaction and efficiency, resulting in decreased expenses and increased production (Forward, 2002). As the technique of data analytics grows increasingly prominent in today's business environment, organizations of all sizes will have the chance to enjoy the many benefits data analytics provides. These skills include the ability to recognize trends in consumer behaviour and demand, the capability to save costs by knowing how to meet customer needs better, and the capacity to forecast demand and optimize operations based on this estimation. In addition, an ability to anticipate client demand and adapt to the company is needed.

Achieve a greater comprehension of the needs and preferences of customers by establishing a heightened degree of awareness. Effectively utilizing all available resources will enable the company to do the most work feasible and maximize the work that can be accomplished. Observing how consumers behave and then using this information to forecast future trends is important (Muslim, 2016). As the technique of data analytics grows increasingly prominent in today's business environment, organizations of all sizes will have the chance to enjoy the many benefits that data analytics provides. Businesses can improve their operations and the customer experience by utilizing data analytics to identify patterns in the behaviour and demand of their customers, reduce costs by understanding how to meet their customer's needs better, increase efficiency by better using their resources, and forecast future trends.

Firms may utilize data analytics to anticipate future trends. Due to enhanced skills, firms can better predict future trends. As a direct consequence, businesses can give their customers the highest possible quality of service, enabling them to retain their market leadership and competitive edge (Fiore, 2019). Data analytics is garnering increasing attention, and as a result, companies of all sizes can make use of its many benefits.

2.2.3 A.P. MOLLER - MAERSK A/S

A.P. MOLLER - MAERSK A/S, later Maersk, is a Danish commercial company engaged in many areas, including shipping, energy, logistics, and retail. The business was established in 1865 and has its headquarters in Copenhagen. The existing site of the company's headquarters dates to its founding in 1865. The organization has used data analytics to assist with demand forecasts and operation improvement. This practice has existed for some time. The container shipping company of Maersk has been made more efficient via data analytics. Data analytics has resulted in constructing models estimating demand for various container types (Papamitsiou, 2014).

The next phase in the planning process is to define how the company will conduct its fleet and cargo operations after gathering this data. In addition, data analytics have been used to improve both the scheduling and routes of containers. This allows Maersk to reduce the expenses connected with transportation while lifting the standard for their production level. Maersk has integrated data analytics as an integral component of the complete plan it has been implementing to improve its container shipping business. Due to using data analytics, the company has improved the planning of its operations and optimized its fleet and cargo routing. There has been a substantial reduction in transportation costs and operational efficiency (Siemens, 2012).

2.3 Goal of data utilization

After performing and following all the desired methods related to the design and methods for doing research on the idea, the collected data will be addressed, in simpler terms, making use of that collected data. The term "data utilization" is used to describe the systematic application of data to business processes to enhance operational effectiveness and productivity. The term "Data-Driven" is popping up in more contexts these days. It's the practice of basing business judgments and solutions on empirical evidence rather than gut feelings or anecdotes. Utilizing data is a crucial step toward being data-driven. About 90% of major enterprises and more than half of smaller and medium-sized organizations report using data in their businesses, per the Ministry of Internal Affairs, Communications'

Research and Study on Measuring this same Economic Value, and Current Status of using Digital Data (2020). Businesses are making greater and greater strides to exploit data.

The goal of every data analysis project should be to derive the answers sought from the data. Data, in its raw state, is just a bunch of numbers and letters; hence, the processing is required before any useful information can be extracted from it. Graphs are a great example of a data visualization tool that makes the data readily apparent and comprehensible. Looking for patterns, anomalies, correlations, and causal linkages gives the data its first real sense of significance. While a firm grounding in statistics and data processing techniques would allow for more in-depth research, using business intelligence (BI) technologies will enable more sophisticated analysis, even for those without such expertise.

2.4 Purpose of Data Utilization

Data usage is the process of putting collected data to work to benefit an organization's operations. Data should be used for a specific, well-defined purpose, which should be established before any work is done with the data. This facilitates analysis of the best courses of action for achieving the objectives. Analysis results will be used to inform strategy and action, and the outcomes will be measured.

If this procedure is used again and again and keeps running the PDCA cycle with data utilization, eventually, strategies and measures can be derived that are more likely to succeed. Having marketing experience and data literacy is a plus when it comes to using data. When such information is collected, it becomes simpler to establish reliable methods of assessment. As was established, the difference between analysis and data use lies in the varying goals served by each.

Data utilization, in contrast to data analysis, seeks to "take essential steps data collection and analysis or use it for business." Putting data to use requires first conducting data analysis. The abilities are also distinguished between the two. Statistics and other data-processing expertise are needed for analysis, while strategy and verification abilities, such as those gained via education and marketing experience, are essential for data utilization. More businesses these days rely on BI tools and other automated methods for data analysis, leaving data usage to humans.

Knowing the market trend of a firm and projecting it into the future is essential for making sound business decisions. Using data, understanding the situation is more comprehensive, and the information can be used to make predictions. It would be simpler to understand the situation, for instance, if sales data by product and store could be visualized in real-time. In addition, when combined with evidence cleaned from data, reliable predictions about the future will be easier to make.

Swift decision-making is crucial in the corporate world. Decisions depending on individual subjectivity, such as experience or intuition, are difficult to communicate to others and might be slow. It is difficult to check whether a decision seems appropriate; therefore, it’s impossible to make decisions quickly. Using data in making decisions not only facilitates evidence-based decision-making but also shortens decision-making cycles by facilitating the easy dissemination of individual perspectives. Another benefit is that it facilitates further validation of the findings.

2.5 Discovering New Business Opportunities

Increasing data consumption could lead to new commercial opportunities. A dashboard can be created from data that has no context on its own; combining it with additional data may shed light on hidden company problems or uncover insights that inspire new product or service ideas. Using data has resulted in the creation of numerous novel commodities and offerings. The company's bottom line might improve if it takes advantage of some of the new business prospects that have presented themselves, applying the large amount of data collected and preserved to any commercial endeavor. Below are examples of different departments and their use of data.

Information gathered from customers, bargaining, and previous purchases are all useful in the sales process. Data on consumers' demographics—such as age, gender, geography, etc.—can be utilized to determine which demographics are most likely to become new buyers after analyzing their purchasing habits. In addition, the dissemination of lessons learned during negotiations can enhance the overall sales capacities of the organization.

Marketers can make use of information gathered from a variety of sources, such as customer surveys, media reports, and sales figures. For example, with EC site advertising, customers'

browsing habits and other data are tracked and analyzed to forecast their likely behaviour, which can then be used to show them more relevant ads.

Sales can be boosted by displaying adverts and recommendations that pique consumers' interest.

Information on products is one form of data that can be used by the manufacturing division. The batteries of an automaker's electric vehicles were giving them trouble, performing inconsistently. They planned to use information from sensors embedded in the battery to track how its performance evolved over time and across varying driving circumstances and usage patterns. A good illustration of how data use can reduce the difficulty of problem prediction and increase quality is presented here.

As an example, from the medical industry, electronic medical records, examination, and imaging data, claim data, and research data are just a few of the many types of medical data processed every day. Numerous efforts and research have been made using these data, and the benefits of data utilization in areas like medical technology advancement and medication discovery have been established. As an illustration, consider the benefits of utilizing data in the medical field, such as in the case of preventive care and early diagnosis of diseases. Finding the cause of an illness and tracking its progression often requires compiling and evaluating data from many people who have experienced similar symptoms. In this way, patients receive care that is more supported by data and experience. Pharmaceutical companies are also making use of data to create innovative new medicines.

Many people believe the time and money needed to create a new drug is excessive due to the complexity of the development processes involved. Using data, researchers have been able to cut back on both the time and money needed to bring a novel medication to market. Further, analysis tools that facilitate the use of a massive volume of healthcare data are contributing to the acceleration of data usage in the medical area. From real medical treatments like surgeries and examinations, complex analysis is made possible by tools that assess the number of prescription days, the prescription amount, etc., based on the medical treatment databases.

Making better use of IoT data is a growing trend in manufacturing. Using IoT data, a factory may create a centralized dashboard that visually represents the current operational state of

all its production lines. The end goal is a more effective method of tracking progress and increasing output. Problems are anticipated by analyzing data on the frequency with which product mistakes and equipment failures occur. Successful businesses and innovative problem-solving hinge on effective data utilization. The information contains numerous indicators that can enhance and grow a company. A company can put data to good use if it analyses it for insights, comes up with strategies in response to those insights, verifies those strategies again and again, and puts them to use in operational processes. A faster decision-making process and developing new business prospects are just two of the many advantages that may be gained through the utilization of data. Many businesses are actively encouraging data use and seeing substantial benefits. As the pace of digitalization continues to quicken, the importance of firms' attempts to make use of data will grow. Consider the following information to grasp the concept fully.

3 Literature Review

Data analytics presents tremendous opportunities for companies to use evidence-based data to make decisions in several domains of business applications. To what extent big data analytics can help the transportation sector, and its various players is an open subject. This section summarizes findings from a recent systematic assessment of the literature on the topic of big data analytics and its usefulness in the transportation sector. This work seeks to comprehend the existing literature, the voids in it, and the difficulties that arise from applying data analytics to the transportation sector. Insights can be gleaned, and transportation system efficiency can be enhanced by applying data analytics, as this article explains. Traffic monitoring devices, crowdsourcing, connected automobiles, and social media are all potential sources for the studied data. The study details several data storage and analytic methods, as well as software and architecture useful to transportation firms. To broaden its appeal, it details the difficulties of putting big data analytics into practice. The review significantly contributes to the field by theorizing how businesses may use information analytics and big information to create transportation systems that are both safer and more environmentally friendly. Due to the limited body of literature on the topic, this study does not offer suggestions for addressing the many issues and knowledge gaps identified in the areas of data analytics and transportation (Willis, 2021).

Data analytics studies provide openings for using an evidence-based decision-making framework in pivotal organizational operations. Data analysis has the potential to increase transportation system safety and sustainability (Channa, 2020). To keep an eye on traffic conditions and make sure everyone is safe, many transportation agencies have placed monitoring systems, including sensors and cameras. There is a wealth of information gleaned from the devices that can be used by transportation firms to improve areas like traffic flow (Babar, 2018). Due to the availability of vast amounts of data, it is necessary to conduct both real-time and retrospective analyses. Important details, including traffic patterns, congestion, and opportunities to improve transportation efficiency, can be uncovered through the investigation, making it crucial. The numerous data analytics used in comprehensive data

analysis, such as machine learning, can sift through the mountain of data to unearth actionable insights for transportation stakeholders.

Values exist in traffic data that could make transportation systems more secure and environmentally friendly. For example, transport companies can detect vehicle speed and analyze it to identify traffic congestion patterns when using roadside sensors. With the design, the companies can provide travel alerts to drivers, which can help them identify alternative routes to reduce their delivery time (Babar, 2018). Analysis of the times vehicles take when waiting for traffic also helps search for better ways of improving traffic policies and transportation efficiency. By analyzing video footage, it's easier to spot objects and pedestrians on the road and identify critical traffic incidents, like collisions and sudden stops. Stakeholders in the transportation sector can improve traffic safety and then save lives by identifying the aforementioned factors.

The properties of big data can be seen in traffic data. Big data is classified under five dimensions: velocity, volume, value, variety, and integrity. The monitoring devices installed on roads generate a large amount of data (Dokic, 2020). This data grows exponentially, especially when connected equipment communicates within itself and with other nearby vehicles. This can generate over 20 gigabytes of data a day, which could rise to terabytes in a month (Fiore, 2019). The variety of traffic data comes in PDF, XML, and JPG. Third, the data velocity in transport systems is obvious since new data is constantly being produced by various sources. Veracity caters to typical uncertainties with transportation data, such as data inconsistency.

The procedure will aid in creating a summary of the established body of knowledge in this field. The protocol outlines the research methodology used to locate, synthesize, and filter literature on big data relevant to the study's topics. The study employs a systematic review to concisely summarize the existing literature on information analytics and big data inside the transportation sector. Included in the protocol for this review is a statement of the survey questions, a list of databases to search, inclusion and exclusion criteria, and search terms.

3.1 Databases and search terms

The following databases were used: Google Scholar and Science Direct, The Xplore Digital Library, Association for Computing Machinery Digital Library, Scopus, and the Directory of Open Access Journals. The databases were selected due to their depth in content covering the domain of big data and computer science applications.

Possible relevant articles were retrieved using the following search terms.

- Big data/ applications/ transportation
- Smart transportation/safety on the roads
- Transportation based on big data and cars that are connected

Criteria for leaving an article out or putting it in:

- It must have been published within the last five years.
- At least a single research question must be answered.
- There should be several citations from reliable sources in the papers.
- The article's publisher needs to have a good name.
- The article's full text must be easy to find and read online.

A recent review of the literature served as a framework for the paper's five-year timeline. Due to big data's status as a relatively new technology, a recent five-year timeframe was relevant, which should be examined according to the current trends. Other articles older than five years were assumed not to have credible data that could be used to print a picture of the research on the topic. Other criteria were used to show that academic publications and peer-reviewed articles were the primary focus of the review. Most of the studies analyzed were originally published by respected publishers like Elsevier. The organizations were selected because they provide a detailed wide range of quality technology content accessible to professionals and researchers.

The following results elucidate snapshots of the search results on applying the search tags on the different data-based returned 183 articles. The abstracts of the articles were reviewed to determine how well they fit with the research questions. After evaluating abstracts, the pool of possible articles is narrowed down to 56. To a large extent, the papers were disregarded because they didn't address the issues being investigated. Some papers were excused since their complete versions weren't available online. From among the 56 articles

that made it past the preliminary screening, only 28 were reviewed for their relevance to the reach questions.

3.2 Big data systems, architecture, and their utilization in transportation

This section summarizes the findings from the literature review, which show how data analytics and big data systems handle some of the impending difficulties in transportation and how big data may improve the industry (Shahiri, 2015). Questions like, "What is the current level of scientific research on the application of data analytics in transportation? " "What types of data repositories are being employed in the sector?" and "How are transport companies leveraging the insights produced from analytics of big data?" are answered here. How can businesses make use of insights from big data analytics to boost productivity and long-term viability?

Establishing an intelligent transportation system in the transportation setting using big data platforms or methodologies discussed in the literature is feasible (ITS). To improve the environmental friendliness and operational efficacy of transportation networks, this management system integrates diverse technologies, including sensors and information communications networks with communication infrastructures (Hariri, 2019). The innovations in transit enabled by this technology are more user-friendly, safer for the environment, and more accessible. Smart city BID data design is one type of ITS. The seven components include gathering raw data, cleaning it up, brokering it, storing it, analyzing it, and finally, visualizing it. Data gathering and interpretation, as well as data analytics and visualization, are two overarching categories used to describe the layers (Lepenioti, 2020). Data from a variety of sources is gathered, cleaned, and loaded into databases by the normalization data sourcing components. Data broking solves the heterogeneity of data challenges. It deals with various formats, frequency of data updates, and integration of data. Such a category extracts meaningful knowledge and important data patterns (Duran-Lopez, 2020). The results of the analysis are then displayed graphically to assist users in making

educated decisions during the visualization phase. These systems result in an efficient and sustainable transportation strategy for transport companies.

This section records how extensive data analysis has promoted the creation of safe and sustainable transportation systems. The significant ideas highlighted include sharing real-time information, the use of big data in transportation and city planning, and the improvement of analysis of close shaves and collisions, which are two ways traffic safety can be enhanced. Incorporating big data into existing systems allows businesses to optimize their processes, enhance their customer service, develop more targeted marketing strategies, and other similar initiatives that have the potential to boost revenue and profitability. Companies that implement its use may gain an advantage over rivals that do not because the former can make more timely and well-informed business decisions.

Big data, for example, can help businesses improve customer service, marketing, and promotions by revealing previously unknown aspects of their target audience and uncovering new avenues to connect with them. Businesses may also benefit by analyzing past and present information to better understand their client's tastes and preferences (Aiello, 2020).

Healthcare organizations and government parties have access to current data about outbreaks of infectious diseases or epidemics from a variety of sources, including health records, social media, the web, and others (Batko, 2022).

Here are a few more practical applications of big data in businesses:

- Big data is used by energy companies to locate prominent locations for gathering resources (Laplante, 2015).
- Big data systems are used by finance and investment companies for risk assessments and on-demand market analysis.
- Big data is used by manufacturing and logistics firms for supply chain management and route optimization (Paliszkiewicz, 2020).
- Uses in the government extend beyond emergencies to include things like crime prevention and "smart city" efforts.

Accounting information systems and programs, customer information, documents, messaging, medical data, mobile app usage, and social media usage logs are just a couple of examples of the many places where big data can be gathered. Data provided by different sensors on machinery, industrial equipment, and different kind of Internet of Things devices are also part of this category (Wang, 2017).

Ways of finding big data on individual and big groups of consumers, economic markets, traffic conditions, geospatial data, and more are commonly incorporated into big data settings in addition to the data from internal systems. A lot of the big data applications contain continuously flowing data that is analyzed and gathered, and this includes not only textual data but also images, videos, and audio recordings.

The term "big data" is typically used to describe large amounts of data. The nature of something like the data being gathered and kept inside a big data environment typically necessitates a huge volume of data to be present; however, this is not always the case. Massive amounts of data are continuously generated from a variety of sources, including clickstreams, system logs, and stream processing systems.

Big data also includes the following categories of information (Ramesh, 2019):

- Semi-structured data which includes:
    - web server logs
    - data from sensors
- Unstructured data which includes:
    - text, documents
    - multimedia files like social media posts

Different data kinds might have to be stored and managed jointly within big data systems. Numerous data sets, not necessarily interconnected at the outset, are typically a part of big data applications. Such big data analytics programs might, for instance, analyze sales history, returns, online reviews, and customer support call to make predictions about future sales of a product. The rate at which new information is being created and must be analyzed and

interpreted is what's called data velocity. Rather than the daily, weekly, or monthly updates found in many typical data warehouses, massive data sets are often updated in near real-time. As big data analysis evolves into artificial intelligence and machine learning (AI), where analytical algorithms automatically discover patterns in the data and use those patterns to develop insights, managing data velocity becomes increasingly crucial (Hariri, 2019).

3.3 Distribution of real-time traffic information

The sustainability and security of transportation networks can be greatly enhanced by the widespread dissemination of real-time traffic data. When transportation companies are equipped with relevant traffic information, the companies can make informed transportation decisions (Wang, 2017). For instance, accurate time information may aid in the avoidance of traffic congestion as people can use the information to avoid congested routes. The choice of alternative routes may help reduce losses that vehicles succumb to in traffic jams. Data analytics makes it possible to evaluate traffic data and relay relevant and valuable information to the drivers. Also, real-time information may help emergency vehicles reach their destinations in good time, thereby saving lives (Zadeh, 2020). The information is usually shared using various communication channels. The channels include mobile applications, billboards, and traffic radios. The information is in the form of alerts that brief drivers on the situation on the roads. These analytics are enhanced by systems such as ITs, which give a predictive analysis of traffic conditions. With the enhancement of sharing traffic information, data analytics ensures transportation companies attain safety and sustainability in their areas of operations.

Optimizing urban freight distribution routes often occurs offline, using information about past network states. Intelligent Transportation Systems (ITSs) collect data in real time, allowing for immediate re-optimization in response to changing traffic conditions. This research is being conducted with an eye on implementing its findings into the city of Rome's perishable freight distribution system (Italy). Unlike the online problem, where it may not skip any clients or rearrange the deliveries, Jylhä’s issue is framed as a vehicle route planning with short time windows. These variants require distinct sets of data and algorithms to solve.

The findings could be used to estimate how much money can be made by investing in various traffic data sets. Results may also be of use to information providers in determining how much to charge for their offline and online resources and how much of a piece of the market they may expect to earn from those resources (Ltifi, 2013).

3.4 Transportation and urban planning.

Evidence-based decision-making in urban and transportation planning is aided by analytics of historic transportation data. Plans for transportation systems can be developed without relying exclusively on gut feelings or unproven hypotheses by industry decision-makers. Present vehicle movement patterns can predict future infrastructure requirements by applying traffic analysis. Data analytics helps simulate traffic flow even before construction commences (Zadeh, 2020). This helps to avoid wastage in transport planning, which helps develop many means of transportation. Traffic flow can also be improved by looking at how long cars wait at different traffic signals and stops, leading to operating transport companies efficiently. Transport companies use data analytics to improve their transportation networks by analyzing public transportation usage. The purpose of urban transportation management is to map out a comprehensive strategy for moving people about a city, so it benefits everyone who lives there. While there have been efforts to standardize the procedure, each study still needs to collect and handle a huge quantity of data unique to the area of interest. Estimates of projected travel demand and facility use are derived from this data, with computerized models of transportation networks and travel patterns. Because of this, computers play a crucial role in providing transportation planners with the ability to evaluate various proposed transport networks to recommend the allocation of government resources and drive transportation policy.

Statistics on accidents and near-misses can help make the transportation sector safer. It helps monitor and visualize places and locations to identify accident-prone areas and causes of collisions (Zeroual, 2019). Data attributes, such as frequency, causes of incidents, and types of vehicles involved, are aggregated during data analysis. To aid transportation stakeholders

in evaluating traffic and policy, the compiled data is shown on a map. Analyses, such as proactive safety analysis of vehicles, can distinguish between vehicles and pedestrians by analyzing metrics such as speed, directions, and characteristics of movement (Zhu et al., 2019). Such analysis may reveal potential risks on the road and the road infrastructure situation. The analysis of traffic helps transportation companies to improve the security of commodities in transit by advising on vehicles that need the installation of trackers and video cameras. This helps in the natural management of situations and visualization of goods in transit.

3.5 Opportunities that data analytics and big data present to the transport sector

One significant opportunity that big data presents to the sector is a chance for transportation companies to use connected vehicles. Connected vehicles are a technology that proposes using GPs, sensors, and other communication technologies to enhance information sharing within a transport network (Babar, 2018). The data shared in this process can be employed for internal operation and review. The technology is effective because it enables the personalization of information provided to the driver. The information may include to-do lists, delivery instructions, and appointments.

To further minimize delivery times and fuel costs, transport companies can have their connected trucks perform engine diagnostics using data stored in the cloud (Ltifi, 2013). The above are examples of how connected vehicles used by transport companies use data obtained from analytical techniques available in data analytics and big data literature. Further, data analytics reveals a future where autonomous vehicles can be adapted for transportation (Hariri, 2019). The vehicles can eliminate the need for human-driving vehicles. Although this idea seems futuristic, recent literature evidence that advances in technology will undoubtedly result in efficient autonomous vehicles. The vehicles can help reduce the number of accidents on the road, eliminating human errors that occur on the roads due to fatigue. However, the literature is still scanty and requires more research.

Nowadays, transportation and mobility data are produced at a rate and volume that outpaces their ability to be recorded, processed, and analyzed around the turn of the century. Insight into the integrated transportation reality has been greatly enhanced by the convergence of new digitalization paradigms, including the Internet of Things (IoT), the pervasiveness of Smart Cities, the sharp decrease in data storage costs inside the silicon industry, the latest updates in wireless technologies, and the low-cost but widespread usage of detectors or personal devices (Hariri, 2019).

Data-intensive tasks, like integration, visualization, querying, and analysis for large-scale real-time systems, are becoming increasingly important in the design and implementation of Intelligent Transportation Systems because of the volume, variety (source, type, and format), and variability (frequent changes of data that complicate to decipher their exact meaning within their context) of transport and mobility data (ITS). It is generally agreed that the data analysis and monitoring capabilities of existing ITS applications are inadequate. The difficulties necessitate practical, efficient monitoring, decision-making, and data-management capabilities, the needs of which rarely are satisfied by the ITS now in use. In addition, the proliferation of actual data streaming sources has the potential to yield novel insights and policy-relevant enhancements to transportation services, activities, and operations with sophisticated aggregation, fusion, and incremental learning strategies.

Big data has been proposed as an innovative tool for the transportation sector. To address these problems and advance the transport area, this modern technological paradigm offers solutions and tools that can acquire, handle, and analyze vast amounts of organized and unstructured data. Extracting value and knowledge from the newly available massive amount of data requires a new generation of intelligent transportation systems (ITS) and mobility services built on the ideas and technologies of something like the Big Data paradigm. The goal is to make sure transportation companies actually use the data they collect, which means creating better, cleaner, and more efficient transportation systems and services that passengers actually enjoy using. In addition, predictive analytics, a popular technological alternative to Big Data that is an aggregating part of the latter, allows for the collection and mining of data to reveal previously unseen trends in transportation issues, paving the way for proactive applications and services with evident ties to safety and security. Infrastructure monitoring, the creation of mobility-enhanced services, a deeper understanding of user

needs, and the visualization of person flows, and their growth across cities are just a few examples of applications that use Big Data in the transportation sector (S Ahmed, 2020).

Even before the emergence of Big Data, data has played a vital part in scenarios involving transportation and mobility, and traffic management is one such example. Research has advanced thanks to incorporating new sources of data and implementing Big Data technology, resulting in novel approaches and ground-breaking findings that go well beyond the capabilities of traditional theoretical frameworks and analytical tools. In this context, Big Data has made numerous innovations to traffic capacity management, traffic situation forecasts, and traffic demand management. Overall, there is consensus inside this community on the possibilities of Big Data as a crucial enabler towards attaining a smarter transportation industry, although Big Data has focused on select application verticals (Duran-Lopez, 2020).

By reviewing and analyzing the most recent research efforts connected to this highlighted synergy, this study hopes to emphasize the momentum obtained by Big Data again for the transport and mobility business. The survey specifically investigates the applications and scenarios treated in the current papers from a clear data-based perspective, analyzing them in terms of how Big Data functionality (modelling, processing, analysis, and visualization) is resolved. A list of legal frameworks that govern the collecting and use of data and zero in on aspects imperative for developing Big Data in this field, such as the privacy statement for transportation data, will be presented. This paper serves as a resource for both seasoned professionals in transportation and those just beginning their journey into this exciting and promising new frontier of data science. Briefly, this survey can be said to have made the following contributions:

With an eye toward the practical needs, data sources, and use cases that support the deployment of Big Data technologies in transportation and mobility, this paper present a contextually relevant definition of Big Data within the context of these domains.

The study looks closely at the situation of Big Data in transportation and mobility, detecting trends and pointing out the obstacles that stand in the way of its widespread adoption. To do this, categorizing the recent literature on the topic according to the stage of something like the Big Information life cycle that each reported advancement has helped; to the authors' knowledge, this systematic approach has not yet been applied in any other study.

Additionally, the survey digs into each stage of the abovementioned Big Data life cycle, analyzing research niches and areas meriting additional study directly related to such cycle stages.

The literature review is completed with an overview of ITS applications and projects that have used Big Data technology, establishing a solid foundation upon which to dictate unanswered paradigms in the field (Duran-Lopez, 2020).

3.6 Contribution of data analytics to transportation

Historical data analysis of transportation data helps reveal trends and patterns that can be used to make long-term decisions in the transport sector. For instance, data analysis can review the traffic flow and usage of the infrastructure before plans are initiated for building transportation infrastructure. The analyzed transportation data can be utilized by policymakers to adjust the rules governing the operation of transport companies. Using data analytics, businesses can do everything from creating personalized marketing campaigns for individual customers to locating and eliminating threats to their operations. Analyzing data can be as simple as performing repetitive chores or as complicated as analyzing behaviour and making predictions. Consolidating Marketing and Sales Efforts Data analytics helps businesses see patterns in client behaviour and tailor their marketing to certain demographics (Neilson, 2019).

Besides assisting with marketing tactics, data analytics may reveal which products and services are most likely to be purchased by clients. Analytics and data collection can also gauge consumer happiness and pinpoint places for development in the business's offerings. Optimizing Current Procedures Analytics of data can help businesses cut costs and increase profits.

Ineffective procedures and infrastructure can be recognized with the help of data analytics. In addition, analytics and data can be utilized to track and enhance worker output. Eliminating Danger Analyzing data can help businesses spot dangers and prepare for them.

Analytics of data can help pinpoint threats to a company's image, finances, and operations. The goal of big data analytics is to examine massive amounts of data to unearth useful

insights, such as previously unknown patterns, correlations, market dynamics, and customer preferences, that may then inform and improve business choices.

Using data analytics to anticipate consumer trends and actions, boost competitiveness, and make data-informed decisions is a core strategic focus for all large corporations.

The importance of data analytics as a tool for enhancing business processes, choices, and outcomes is growing inside organizations. Because of its usefulness in predicting consumer trends and behaviours, boosting business performance, and making data-informed decisions, data analytics has become an integral part of strategy development at all major corporations. Data analytics, or the study of how to use data to improve an organization's operations and bottom line, is a growing field in the scientific community.

Data mining, scientific techniques, and predictive modelling are just a few of the data analytics approaches used in business analytics. Business analysts can get jobs without a business degree in analytics, but those with one have a leg up on the competition. Usually, employers prefer candidates with at least a bachelor's degree in business or a related quantitative discipline. However, master's degree holders in business intelligence or data science are the first choice of many businesses.

Data mining, statistical analysis, and predictive modelling expertise are also highly valued by employers. In most organizations, coordinators keep things running smoothly daily. They manage the workforce, making sure organizational objectives are realized through the establishment and enforcement of policies and procedures and more. An undergraduate degree in business or a similar subject is usually required for employment. Having excellent organizing and communication abilities is also crucial.

Predictive analytics provides useful insights for making data-driven choices. By analyzing large amounts of data, businesses can gain useful information about product viability, manufacturing decisions, assessing progress, and directing changes.

Insights collected from data analytics can drive better decisions and improve project management efficiency by providing a more comprehensive understanding of the company's client base and its own productivity. Profits will rise, or prices will drop as a result. Computing and data analytics can be utilized to enhance project management, but the final decision should still be made by humans (Siegel, 2014).

If data analytics is used for project management, it’s essential to make sure the data used is up-to-date, accurate, and easy to access. The success rate of projects within an organization can be increased by using data analytics to manage projects better. Organizations can improve their project leadership and decision-making using data analytics. Effective and efficient project management can be enhanced by utilizing data analytics.

Using data analytics, marketing experts can learn more about customers' wants and needs from a variety of data sources, including high-quality, previously collected data. Data science is useful for every firm, no matter the industry, because it provides statistics and insights into the entire work process, helps with hiring new individuals, and assists senior executives in making more educated decisions.

Thanks to developments in parallel computing and more affordable processing power, businesses may now use big data analytics to draw useful conclusions from complex, multiple data sources. Organizations can benefit from the insights that enterprise customers gain from the available data to act more swiftly and intelligently.

Effective analytics procedures and tools can help businesses guarantee ready access to high-quality data whenever it's needed. Using analytics programs, relevant data may be extracted, sorted, and analyzed to reveal patterns and numerical data that can guide decision-making and lead to monetary gains (Lepenioti, 2020).

Since these same data analysts are tasked with reviewing the most important metrics for efficiency, interpreting the data, and using it to generate successful methods for driving the organization's productivity, data analytics is critical for the success of modern enterprises. As we've seen, trained data analysts are an invaluable resource for modern businesses since they assist upper management in making vital strategic decisions affecting the company's primary lines of business.

3.7 Real-Time Planning

Analytics refers to the processing and distribution of information. The analysis provides immediate information for decision-makers for faster business decisions (Neilson, 2019). For example, transport companies can use real-time analytics to control traffic control systems, such as timing advisory and speed. Research in real-time data analytics is crucial

to big data. For instance, the analysis can help assess traffic flow in the transportation domain to manage road conditions, accidents, and transportation emergencies. The term "real-time analytics" describes the practice of analyzing and evaluating data as it is being entered into a database. The data is processed so quickly that users can immediately make sense of it and draw conclusions.

Companies can move thanks to real-time analytics quickly. They're able to grasp chances and head off potential issues. The results of batch-style analytics may not be available for several hours or even days. As a result, "after the fact" insights are typically all that can be gleaned from batch analytical applications (lagging indicators). Businesses can get a leg up on the competition with BI Insights derived from real-time analytics.

Applications that perform real-time analytics should have both high availability and quick reaction times if they are to be of immediate utility. They also need to be able to process data sets that are many terabytes in size. However, they must still respond to questions in a few seconds. For "real-time" to work, it must be able to deal with new data sources that appear because of shifting market conditions and company priorities. They should be able to deal with "big data.” Financial trading now uses big data analytics in real time. Instantaneously, they make investments based on information gleaned from social media, financial databases, and even satellite weather sensors. Using technology in business nowadays is ubiquitous. There is a rising need for real-time big data analytics, so the volume and variety of data being collected must expand. To accommodate these needs, various technologies are available.

Some depend on high-tech home equipment (hardware and software systems). As an alternative, some organizations rely on unique configurations of processing and memory chips or on analytics performed within the database (the database has analytics capabilities embedded in it). However, standard computing infrastructures and any data source can also be utilized. The genuine analytics app needs to be programmed to make the most of the resources provided by regular computers' processors and memory. Because of this, the cost of real-time analytics decreases. When the program is also simple to use, it provides business users unprecedented access to actionable, real-time business intelligence. Likewise, this is the optimal location for maximum commercial gain.

3.8 Challenges in using data analytics in transportation

The transportation sector is unique, presenting several problems and knowledge gaps. Data gathering, storage, quality evaluation, and safety are four areas where the alterations can be clarified (Lepenioti, 2020). Data analytics faces the challenge of multiple data sources in terms of data collection. Some sources require analytical processing before meaningful information can be extracted from them (Lepenioti, 2020). It's also possible that some of these sources are under the jurisdiction of parties other than those involved in transportation, rendering the data unavailable to those who need it. Various quality challenges need to be resolved before the transportation industry can benefit from the abundance of data. Issues, such as conflicting data, necessitate cross-validation from multiple sources, an aspect that is both time-consuming and expensive. Data analytics also faces the challenge of the absence of tools for managing data integrity, an aspect that lowers the stability of data. Therefore, researchers should find a way in which the changes in data quality and unavailability of tools can be solved.

Transportation advancements have contributed to a greater global network. But there are several obstacles in the transportation industry. In 2016, fatal car accidents claimed the lives of about 37,461 individuals in the United States. This is the greatest number of traffic deaths since 2008, and it represents a 5.6% rise from 2015. The price of transportation includes more than human lives lost. Commuters in the United States wasted 3.1 billion barrels of oil and $160 billion in time sitting in traffic in 2014. As the last point, 27% of America's greenhouse gas emissions come from automobiles. These negative effects of transport networks are felt worldwide, from rising vehicle movement in the United States to a major rise in privately owned cars in underdeveloped countries.

Freeways, buses, trains, traffic signals, cycling lanes, and more all have sensors and monitoring installed on them, making transportation networks relatively instrumented. Besides these permanently installed sensors, automobiles are also highly instrumented, and private organizations collect transportation data by tracking the whereabouts of people and their vehicles using GPS. Automated and linked car technology is finally being aggressively developed by the government, private automakers, and aftermarket companies. Some estimates put the rate of data created from this technology at over 20 petabytes per second. Data on transportation is collected and stored routinely, though methods may vary. By better

using datasets obtained and archives, it’s possible to optimize the timing of traffic signals in real-time, provide more tailored routes, better respond to traffic accidents by coordinating responses across transportation agencies, and gain a deeper knowledge of traffic bottlenecks. These measures would make the streets safer, lessen traffic, and cut down on greenhouse gas production.

There is a window of opportunity provided by the abundance of data generated by transportation systems, which might address several pressing problems. Diverse new data management tools are needed to realize the potential of transportation data for uses like enhancing safety and lowering congestion and pollution. The complicated relationships between various forms of transportation may necessitate the use of novel data models, such as multi-graphs and semantic ontologies. Integrating heterogeneous, multi-source transportation information in real-time requires novel data integration strategies. However, the existing designs of physical systems are not optimal for the analysis of the huge range of dynamic datasets obtained that must be collected. To process more complex declarative questions that arbitrarily choose and mix parts of the transportation graph, new query processing algorithms are required. New updates and query semantics are needed due to the rapidly evolving nature of transportation data.

Finally, knowledge about how to systematically repurpose data obtained for one purpose (transportation system operation) or another (research or planning) is necessary. Based on the author's 20 years of experience in the transportation industry, this paper describes the domain's prospects and problems (Papamitsiou, 2014).

4 Case study

4.1 Transport Jylhä

Kuljetus ja muutto O. Jylhä oy, or Transport Jylhä as their brand name since 2021, is a family-owned transportation company. The company was founded in 1961 by Olavi, the father of the current owner Pekka. The company is still family-owned. The company began with only one truck, transporting bakery goods domestically. Eventually, the company started importing flowers and plants at the request of a local garden shop. That was the catalyst for starting the whole new business of becoming the biggest family-owned transporter of flowers and plants in Finland. The company operates within Finland, Sweden, and Norway and in mainland Europe. Their trucks go into the British Isles also on a weekly basis. The company has a long history, but within recent years, growth has been exponential. The revenue has grown 155% from 2016 to 2020, while the profit has stayed at around three to three and a half million over that time. The amount of personnel has grown from 74 to 115 over that timeframe (Figures 1-4). The growth has led to situations in which it is more difficult to keep the utilization rate high all the time. The transportation business is a rather low profit-margin business (Remes, 2015), so the revenue must be drawn from all possible sources. Due to those factors, among others, the Company has developed a strategy that involves being more flexible than the competition. The strategy has paid off when looking at the growth rate, but it also leads to difficulties in planning the most profitable routes.

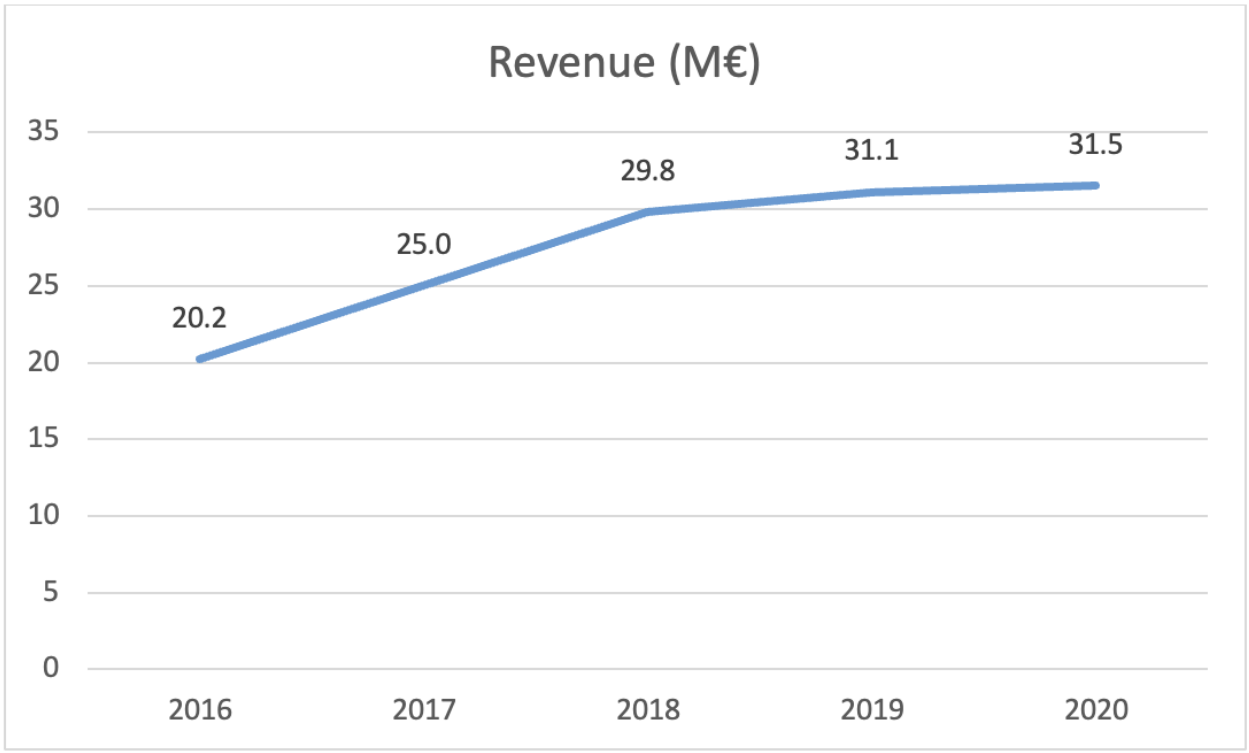


Figure 1, revenue 2016 - 2020

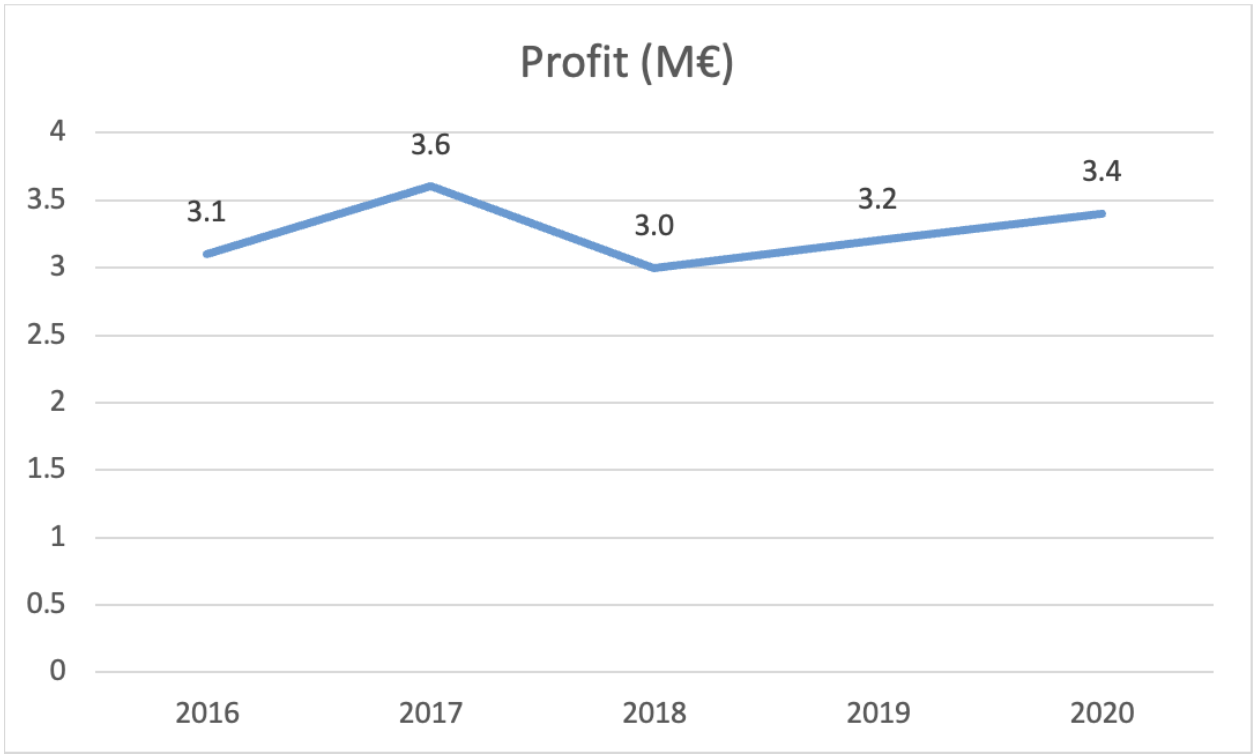


Figure 2, profit 2016 - 2020

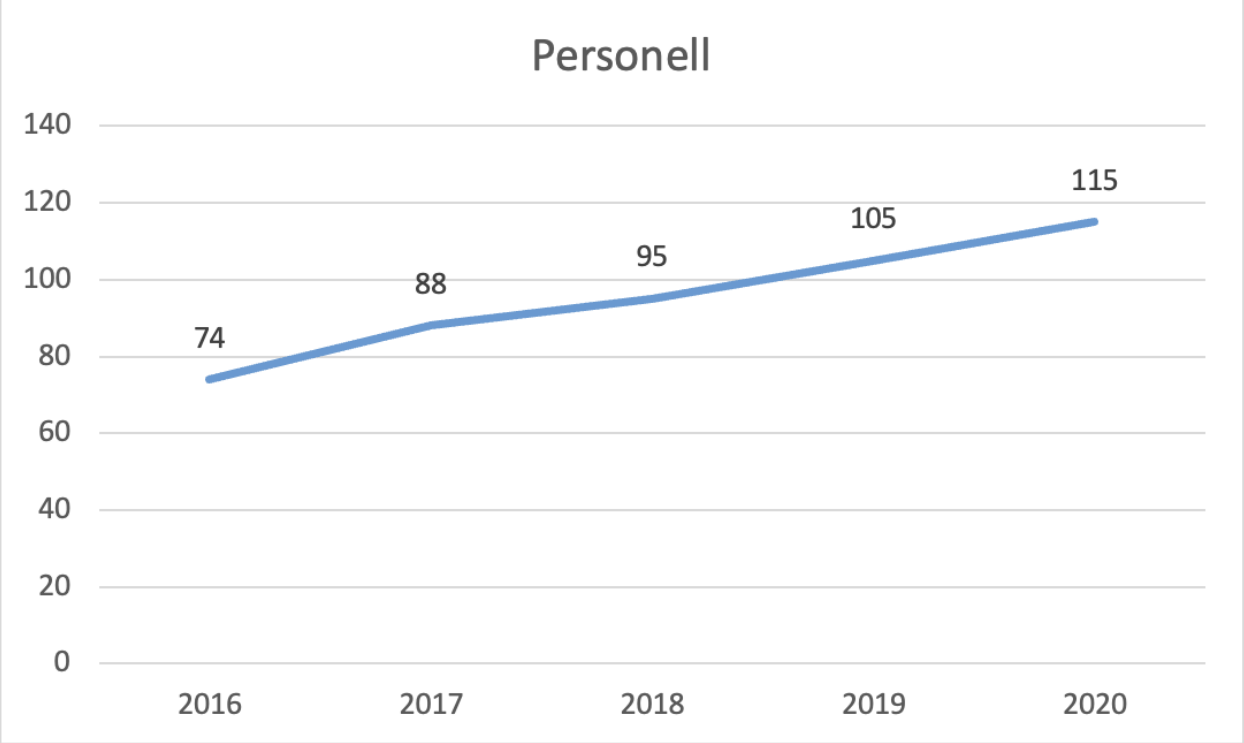


Figure 3, personnel 2016 - 2020

Main operating areas in Europe include the south and southeast Denmark, Northwest Germany, Netherlands, and Belgium. Weekly operations include the rest of Germany, Italy, Northeast France, South England, and Spain. On the map (figure 4), the red area illustrates the rough area of daily operations (Jylhä, 2022).

Figure 4, map of daily operations

Year	Total driven distance(m. Km)	Distance driven empty(m. Km)
2018	5.3	0.7
2019	10.8	1.7
2020	11.2	1.5
2021	13.3	1.6

Table 1, total distance driven and distance drove without cargo

4.2 Information gathering

Another part of the data collection for this study was a set of questions for the office personnel of the company. The gathering was conducted using a Microsoft forms questionnaire that was sent out to the key personnel of the company. A questionnaire form

was chosen because the personnel is based in different locations, and the coverage was aimed to maximize. Also, personnel work in three shifts, so scheduling face-to-face interviews with everyone would've been difficult. The personnel consisted of transport planners, Sales, Booking, and Management. The questionnaire was sent out to 18 people and received 16 responses. That equals an 88% response rate, which is surprisingly good from the author's point of view. The questions asked were:

- Role in the company
- Information is easily accessible. 1 = Fully disagree / 5= Fully agree
- I spend time looking for info on a daily basis because the info is not easily available?
- I often have to consult my colleagues since I don't have all the needed info available to do my own tasks?
- How much time you would save on a daily basis, if information would be more easily accessible?
- How much of the decisions made in the company are made by feeling, not based on data?
- Do you think that data-based decision making would lead into a better result?
- Do you think that data-based decision making would allow the company to grow faster or easier?
- Select the biggest reasons you think are behind the lack of information on your own work.
    - Poor information from the customers
    - Flawed tools
    - Poor utilization of existing tools
    - Internal communication

The results gained from the interviews are presented in the appendix.

The interviews were designed based on a previous understanding that the company lacks information or the information doesn't travel within the company as it should. This was confirmed by the interviews. Question number two was, "The information needed in my work is easily accessible." Answer options were given on a scale from one to five, where one equals fully disagree, and five equals fully agree. The average score on this question was 3.5, which indicates that the information is available but could be more easily accessed. "Information" in this questionnaire means everything from booking data to customer addresses, phone numbers, pricing, and other agreements.

56% of the interviewees answered that they are using extra time searching for information daily. 100% answered that they need to ask their colleagues daily for extra information. 100% answered that they'd save some amount of working time if the information they need were more easily accessible. 56% estimated that they'd save 30-60min daily and 44% answered that they'd save 10-30min. These answers indicate that there is potential to reduce everyone's workload with more easily accessible information. Let's assume that every interviewee would save 10min daily with better tools. That would lead to 90 working days a year in total.

88% of the interviewees answered that 0%-50% of all the decisions made in the company are made without data backing them up. 94% answered that basing the decision making more on data would lead to a better result. 100% answered that better accessible information would make it easier or faster for the company to grow. From these answers, it can be derived that there is a general opinion within the company that the decision-making should be more backed up by concrete information. Being more transparent with some decisions could possibly decrease that feeling.

Question number nine is an interesting one. The question was, "What do you think is the biggest reason that you're not getting enough information to complete your work the best way possible?" The main reasons were internal communication, with 37% of the answers, and poor information from the customers, with 40% of the answers. This is something that should be investigated further since it's out of this study's scope, but very apparent that improving those two fields would improve the information flow to the employees.

4.1 Problem Background Information

The first problem to fix was on the transport planner side. Based on the interviews, the common thought was that the planning work is slowed by the lack of predictability or at least the hard-to-get information about it. An example of this was from the next day's planning shift. In order to succeed in that shift, the transport planner must distribute the fleet of vehicles throughout the main areas of operation and to the unique locations if there is cargo planned for loading in those locations. To succeed on that shift, the planner needs to have a clear idea of where the fleet is needed the most on a given day. For example, there can be a customer that always orders a few pallets of goods to be picked up on Thursdays but always sends in the orders on Thursday morning. That's why the transport planner needs to make a decision to locate one truck near that place, so when the orders come in, the nearest truck is not 200 kilometers and would miss the pick-up of the goods.

Another problem is predicting how many trailers are needed per day. As the Company has its terminal building near the Lübeck port, they can have their own trailers parked in the terminal yard waiting for cargo. But a trailer is also a transportation unit, and even though it doesn't have running costs while standing parked, it does cost something to purchase; the company needs to pay for insurance and taxes on the unit. That's why optimization of the number of trailers is also needed. The basic cycle of one trailer goes as follows. It is loaded into a vessel in the port of Vuosaari in Finland. The domestic transport planners inform the international transport planners of the trailers, trucks, and what's loaded on each one. The next day, the trailers are unloaded from the vessel in the port of Lübeck, Germany. From there, they might leave for the unloading site immediately, the cargo might be unloaded at the Lübeck terminal, or if the trailer is empty or unloading later in the week, it might be left parked in the terminal yard. If the trailer leaves for the unloading location immediately, a driver takes it and drives towards the location, and unloading usually happens the next morning. After unloading, the trailer collects goods in the nearby areas. After the trailer is full, it is taken back north to the port of Lübeck, where it is loaded into a vessel again towards Finland.

Two big reasons can interrupt this cycle, thus causing bottlenecks. One is the available space on the vessels from and to Finland. Especially during the planting season in spring, Finland is getting a lot of seedlings, greenhouse plants, and garden plants on top of normal daily

goods, like vegetables and fruits. Since the daily number of trailers can be 20x the slowest times, it does cause capacity issues on the vessels between Germany and Finland. Another bottleneck is the actually available trailers. Since it's not economically beneficial to own the maximum number of trailers the company needs on the busiest day of the year, during high season, relying on subcontractors and rental trailers more than usual is needed. It also requires the planning of the daily amounts to be on point.

In the past, the problem has been the lack of exports. That led to a lot of dead freight driven when empty trailers were sent from Finland to be loaded in Europe. During the covid pandemic, that changed, and the exports raised a lot. The general feeling at the Company is that the competition has better tools to track their capacity and knows better when to say no to export bookings. After the major completion refuses to take more bookings, the remaining bookings will be offered to the companies willing to take them, often to the Company. The profit made from exports is lower than from imports, but that is some profit in comparison to the trailers being empty. Trailers needed for exports need to be used on imports as soon as they get empty since there is no financial room for small companies to have transport units standing idle.

The number of bookings received for imports and the time of receiving the bookings depend on the day of the week. For example, the majority of Monday bookings come in on Friday, but most of the Thursday bookings come in on Tuesday. Most bookings for Friday come on Thursday. That causes issues in the forecasting of the need for transport units per day per location. Since the lead time from the export cargo loading place in Finland to the import cargo loading place in, for example, the Netherlands is two days, the number of additional trailers needed for Friday loadings needs to be sent out from Finland on Wednesday. That, combined with the fact that most Friday bookings come in on Thursday, causes an issue with fleet optimization. But if the bookings for Friday could be predicted on Tuesday or even on Monday, that would allow better capacity availability on the vessels on Wednesday and a more accurate estimate of the actual number of units needed.

Businesses generate tons of data daily that organizations must try to conceptualize efficiently. With the revolution technology brought forward, there is software available that stores and manipulates the data. Some applications utilized by companies encompass excel and Microsoft office suite. However, a business intelligence tool, such as Microsoft Power BI, is commonly used by many companies. For instance, transportation companies maximize

the use of Power BI as it takes in numerous pieces of information and processes the information into useful and easily comprehendible data. Statistically, there are records documenting 36039 companies using Power BI, and some of these companies include transportation companies. This section will detail how transportation companies will leverage Power BI capabilities to enhance profitability.

Microsoft Power BI is a business intelligence tool that takes in tons of information and manipulates the data into understandable information. Power Bi presents the final information visually, which makes it easier to conceptualize. The features of Power BI bear a resemblance to that of spreadsheets, but the capabilities of Power BI are unmatched compared to those of spreadsheets. One reason Power BI is superior to excel is that Power BI has a user-friendly interface. Additionally, the Power BI dashboard is customizable depending on the user's wants, while excel isn't customizable. Similarly, when contrasting power in reports and tables, excel lags behind as Power BI has augmented capabilities, such as taking in numerous information, which excel cannot. Interestingly, Power BI is convenient in businesses since it's mobile, such that one can access its services from a mobile phone.

Transportation companies upload data to Power BI for processing to enhance profitability. Compared to other software available, Power BI meets the expectation of transportation companies as it manipulates data effectively. Some of the data uploaded include fuel costs, type of goods in transit, expected and departure times, and routes available for faster delivery. With this information, transportation companies can effectively plan their operations and assess which options are beneficial to the company. The information is in visual form, making it easy to understand and comprehend.

Information uploaded to Power BI aids in establishing trends that transport companies use to increase profits. Trends are crucial in business as it aids in predicting the most likely scenario the business will take based on present and past data available on Power BI (Zadeh, 2020). Also, the statistics uploaded are unique to the firm, which makes the trends reliable. With this information, a transportation company can make necessary adjustments to its budget and operations to mirror the expected trends. The company would find itself in a position to experience increased profits.

The Power BI platform allows customizations that transport companies use to develop dashboards that make data access easy. Having quick access to information is crucial to a firm as it aids in making an informed decision that will steer the firm toward profitability. For instance, transportation companies, such as Uber, greatly benefit from Power BI since the drivers can look up information to know when customers use their services. The visualizations can indicate that the numbers are high during Friday, Saturday, and Sunday nights, contrasted to other days of the week. With this information, the driver knows the peak times that will earn more income.

It is important for transport companies to identify key performance indicators correctly, as these indicators aid in augmenting profits immensely. Here, the key performance indicators are shown in Power BI, which users can easily access and use to implement approaches that push the company to realize more revenues. For transportation companies, the most common KPIs encompass transportation time, overheads relating to freight, and costs of fuel. When transporting companies update this information to Power BI, the output will outline what the company must focus on to enhance net profits and the end of a financial year.

Companies focusing on transportation use Power BI for its ease of integration with other applications, such as Google Analytics. For a company to affect profit-oriented schemas, all information must be available. For transportation companies, information from an application such as SAP, SharePoint, and Hadoop play a crucial role in identifying areas that impact the profitability of transport companies. Reports detail that Uber utilizes Hadoop to store huge amounts of data that is later processed to make useful information (Willis, 2021). Information collected encompasses GPS information, which is essential to predict the patterns of a user and make appropriate suggestions when requesting a ride. The processing of such data into useful information and issuing suggestions to clients makes Uber the people`s favorite, thus increasing the profits of Uber (Fiore, 2019).

Data protection and security are pivotal to guaranteeing the success of a transportation firm. Most businesses strive to have reinforced data protection measures as data loss has detrimental effects. Some of these effects include the leaking of client information, which can lead to extortion if the information falls into the wrong hands. That is why transportation companies invest in having impenetrable systems to secure the safety of clients' information. For instance, a transport company like Uber cannot risk losing its information. This is because people`s addresses will be available, risking their lives and property loss. Similarly,

data breaches destroy the reputation of a company, which will result in fewer clients. Consequently, fewer clients mean the business will run at a loss. It’s imperative to adopt a secure system, such as Power BI, which adequately protects the company`s data against breaches and loss.

Power BI is a business intelligence tool used by firms to manipulate data into visual, usable data. Power BI is secure and keeps away intruders and prevents data loss that leads to the loss of a company`s reputation. Some of the importance of Power BI in transportation companies that ensures profitability encompasses the establishment of trends and customization of dashboards, which prompts easy access (Reddy, 2019). Power BI integrates with other applications, such as SAP, which makes it easier to combine all the information in one application to aid in the data processing. This promotes effective and informed decision-making. Likewise, Power BI offers data security, which is essential in guarding the reputation of transportation companies. The incorporation of Power BI by transportation companies will augment profit margins.

4.2 Description of Created Tools

The first tool to be created was a real-time visualization of the daily booking amounts from different countries and how much cargo will be brought to the Lübeck terminal by subcontractors. The main problem this tool was to solve was the actual problem of slowness in getting the information straight from the TMS. The reason this tool needed to be on real-time was that the orders come in throughout the day, and, for example, if a salesperson sells a full truck to be loaded on the same day at 8 am, the transport planners immediately see a large spike on the amounts for that day, not just an additional row in the TMS. That was also the reason the tool needed to be a graph. The need for real-time data led to the utilization of Microsoft Power BI. It allows creating direct queries into the SQL database, and it allows smooth updating for every second if desired. Developing this model took about 48 hours. This tool is on display at the transport planner’s office and makes it easy for the planners to check the booked amounts in the upcoming days. Transport planners are not able to find the cargo. That’s the responsibility of the salespeople in Jylhä’s organization. But what they can do is maximize the utilization rate. Figure 2 and Table 1 represent the profits and distances driven both with and without cargo. Distance driven empty has remained almost the same

from 2018 to 2021, even though the total distance driven along with the profit has increased. This is where the transport planners have a direct impact. Their job is to optimize the routes that trucks travel, and with good planning of the available resources, they can maximize the utilization rate and minimize the distance traveled without cargo.

Figure 5, Real-time cargo info screen

Another, and probably the more essential, tool was for the whole company. It needed to be directly connected to the database, easy and clear to use, and include forecasting. Since this tool is meant for the whole company as a daily tool that can also be used to copy data, it was decided this would be built with Excel. This allows for easy scalability if, for example, some employee wants to create new views for new markets. The tool was built, so the spreadsheet queries data directly from the SQL database. It has four queries: one for loading amounts and locations, one for unloading amounts and locations, and two related to forecasting. Those two are the times when bookings are created, and the info about the bookings is listed. The user interface of the tool works, so different tabs on the workbook are named differently. There's a front page that includes the main markets of the company, loading amounts, unloading amounts, and the forecasted loading amount for each day for the upcoming ten days per market. For the first five days, the forecasting logic checks the current date, time, current week number, and weekday. It then checks the previous year's bookings for the same

week number and weekday. Instead of just using the exact date, i.e., 2.10.2021, it was decided that the week number and weekday combination is more accurate. Since the booking amounts differ radically on different weekdays, it was decided that a more accurate forecast would be obtained when looking at the same weekday on the same week from the previous year. When scaling the tool, a whole country row can be copied and pasted below the last one, and the formulas will work without modifications.

Figures seven, eight, and nine show the user interface of the tool. Column headers are dates, and every country has three rows assigned to it. These rows have different sets of values, and because this tool was built for the whole company, the language was decided to be Finnish. In the column two, the row names are:

- Lastaukset = Loadings
- Purut = Unloadings
- Ennuste = Forecast

Maa	Päivä	7.10.2022	8.10.2022	9.10.2022	10.10.2022	11.10.2022	12.10.2022	13.10.2022	14.10.2022	15.10.2022	16.10.2022	17.10.2022	18.10.2022	19.10.2022	20.10.2022
NL	Lastaukset	16,4	4,2	0,1	5,2	1,7	1,1	0,8	1,1	0,0	0,0	0,0	1,1	0,0	0,0
	Purut	12,6	5,0	5,4	11,7	3,2	4,4	2,3	2,1	1,0	0,0	1,2	1,0	1,0	0,0
	Ennuste	16,6	4,2		9,3	9,6	0,0	0,7	1,2	0,0	0,0	0,0	0,0	0,0	0,0
DE	Lastaukset	5,0	5,5	0,0	1,0	0,0	0,0	0,6	1,0	0,0	0,0	0,0	0,0	0,0	0,0
	Purut	6,3	6,2	5,3	3,3	2,1	6,1	1,4	0,2	0,0	0,0	1,0	0,0	0,0	0,0
	Ennuste	5,0	5,7		1,1	0,0	0,0	0,7	1,2	0,0	0,0	0,0	0,0	0,0	0,0
DK	Lastaukset	0,7	0,0	0,0	0,1	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
	Purut	0,0	0,0	1,0	2,2	0,0	0,0	0,0	0,8	0,0	0,0	0,0	0,0	0,0	0,0
	Ennuste	0,7			0,2		0,0	0,7	1,2	0,0	0,0	0,0	0,0	0,0	0,0
BE	Lastaukset	1,0	0,8	0,0	2,0	1,9	2,0	0,0	0,5	0,0	0,0	0,0	0,0	0,0	0,1
	Purut	5,1	3,0	1,0	0,0	0,3	0,5	0,3	0,0	0,0	0,0	0,0	0,0	0,0	4,1
	Ennuste	1,1	0,8		2,4	3,0	0,0	0,7	1,2	0,0	0,0	0,0	0,0	0,0	0,0
FI	Lastaukset	32,7	7,9	2,1	7,8	3,7	2,8	1,4	5,1	0,0	0,6	3,0	2,0	3,0	3,1
	Purut	32,0	31,6	17,2	21,5	2,5	9,2	8,6	3,5	0,8	3,4	0,0	0,0	0,5	1,1
	Ennuste	32,8	11,0	2,7	12,4	9,3	0,0	0,7	1,2	0,0	0,0	0,0	0,0	0,0	0,0
GB	Lastaukset	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
	Purut	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,2	0,0	0,0
	Ennuste						0,0	0,7	1,2	0,0	0,0	0,0	0,0	0,0	0,0

Figure 6, main page of the excel tool

The second page needed a similar view for the different loading regions. Loading regions are determined by the company, and they are categories of different regions clustered under four loading regions. Those regions are “Kotimaa” for domestic transports, “Vienti” for exports, “INT Runko” for the regions on the daily operating route and “Muut maat” for all others. Loading regions are something the company has started utilising only recently, so in the loading regions field, there are two empty rows for future additions.

Alue	Päivä	7.10.2022	8.10.2022	9.10.2022	10.10.2022	11.10.2022	12.10.2022	13.10.2022	14.10.2022	15.10.2022	16.10.2022	17.10.2022	18.10.2022	19.10.2022	20.10.2022
Kotimaa	Lastaukset	12,3	2,9	0,2	0,7	0,3	0,7	0,2	0,0	0,0	0,0	0,0	0,0	0,0	0,0
	Purut	16,6	7,4	1,0	1,6	0,3	0,7	0,9	0,1	0,0	0,0	0,0	0,0	0,0	0,0
	Ennuste	12,3		0,3	2,4	0,9	3,7	1,0	3,1	0,0	0,0	0,0	1,3	0,0	0,1
INT Runko	Lastaukset	22,5	10,5	0,1	8,1	3,6	3,1	0,8	2,6	0,0	0,0	0,0	1,1	0,0	0,1
	Purut	17,0	23,8	13,8	18,1	1,4	8,5	7,7	3,1	0,2	3,2	0,0	0,0	0,5	1,1
	Ennuste	22,8	10,7		12,8	12,3	3,7	1,0	3,1	0,0	0,0	0,0	1,3	0,0	0,1
Vienti	Lastaukset	20,4	5,1	1,9	7,1	3,4	2,1	1,2	5,1	0,0	0,6	3,0	2,0	3,0	3,1
	Purut	21,6	14,3	12,8	19,6	8,6	11,5	4,0	3,1	1,0	0,0	2,2	1,2	1,0	4,1
	Ennuste	20,5	7,0	2,4	10,8	8,4	3,7	1,0	3,1	0,0	0,0	0,0	1,3	0,0	0,1
Muut maat	Lastaukset	1,1	0,0	0,0	0,2	0,3	0,0	0,8	0,0	0,0	0,0	0,0	0,0	0,0	0,5
	Purut	0,9	0,4	2,4	1,7	0,8	0,2	0,0	0,3	0,6	0,2	0,0	0,0	0,0	0,0
	Ennuste	1,1					3,7	1,0	3,1	0,0	0,0	0,0	1,3	0,0	0,1
	Lastaukset	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
	Purut	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
	Ennuste	0,0					3,7	1,0	3,1	0,0	0,0	0,0	1,3	0,0	0,1
	Lastaukset	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
	Purut	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
	Ennuste	0,0					3,7	1,0	3,1	0,0	0,0	0,0	1,3	0,0	0,1

Figure 7, Loading regions view

The third views are the country specific views. They are all from a single country, divided by postal code regions. For example, the Netherlands, the main market of the company, is split by the postal code areas in a way that 1 represents the area in which the postal codes start with 1. See figure 10.

Alue	Päivä	7.10.2022	8.10.2022	9.10.2022	10.10.2022	11.10.2022	12.10.2022	13.10.2022	14.10.2022	15.10.2022	16.10.2022	17.10.2022	18.10.2022	19.10.2022	20.10.2022
	Viikonpäivä	perjantai	lauantai	sunnuntai	maanantai	tiistai	keskiviikko	torstai	perjantai	lauantai	sunnuntai	maanantai	tiistai	keskiviikko	torstai
1	Lastaukset	4,5	2,5	0,0	3,2	1,3	0,7	0,4	1,0	0,0	0,0	0,0	0,0	0,0	0,0
	Purut	4,1	0,1	1,0	0,0	0,1	0,3	0,0	0,0	1,0	0,0	0,0	0,0	0,0	0,0
	Ennuste	4,5	2,5		4,4	4,7	0,9	0,5	1,2	0,0	0,0	0,0	0,0	0,0	0,0
2	Lastaukset	3,9	1,4	0,0	0,8	0,3	0,3	0,0	0,0	0,0	0,0	0,0	0,1	0,0	0,0
	Purut	0,0	0,3	0,0	1,6	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
	Ennuste	4,0	1,4		1,8	3,1	0,9	0,5	1,2	0,0	0,0	0,0	0,0	0,0	0,0
3	Lastaukset	1,5	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
	Purut	5,2	3,9	0,3	2,1	1,0	0,0	0,1	1,0	0,0	0,0	1,2	1,0	0,0	0,0
	Ennuste	1,5	0,0		0,0		0,9	0,5	1,2	0,0	0,0	0,0	0,0	0,0	0,0
4	Lastaukset	3,1	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	1,0	0,0	0,0
	Purut	3,0	0,0	3,0	3,0	0,0	3,0	1,2	0,1	0,0	0,0	0,0	0,0	0,0	0,0
	Ennuste	3,1			0,0	0,0	0,9	0,5	1,2	0,0	0,0	0,0	0,0	0,0	0,0
5	Lastaukset	1,4	0,2	0,1	0,6	0,1	0,0	0,4	0,0	0,0	0,0	0,0	0,0	0,0	0,0
	Purut	0,2	0,1	0,1	1,0	1,0	1,0	1,0	1,0	0,0	0,0	0,0	0,0	1,0	0,0
	Ennuste	1,4	0,2		1,1		0,9	0,5	1,2	0,0	0,0	0,0	0,0	0,0	0,0
6	Lastaukset	0,7	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
	Purut	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
	Ennuste	0,7				0,0	0,9	0,5	1,2	0,0	0,0	0,0	0,0	0,0	0,0
7	Lastaukset	0,3	0,0	0,0	0,1	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
	Purut	0,0	0,4	1,0	1,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
	Ennuste	0,3				0,0	0,9	0,5	1,2	0,0	0,0	0,0	0,0	0,0	0,0
8	Lastaukset	1,1	0,0	0,0	0,5	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
	Purut	0,0	0,0	0,0	3,0	1,1	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
	Ennuste	1,2	0,0		0,5		0,9	0,5	1,2	0,0	0,0	0,0	0,0	0,0	0,0
9	Lastaukset	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
	Purut	0,0	0,3	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
	Ennuste	0,0					0,9	0,5	1,2	0,0	0,0	0,0	0,0	0,0	0,0

Figure 8, Country specific view

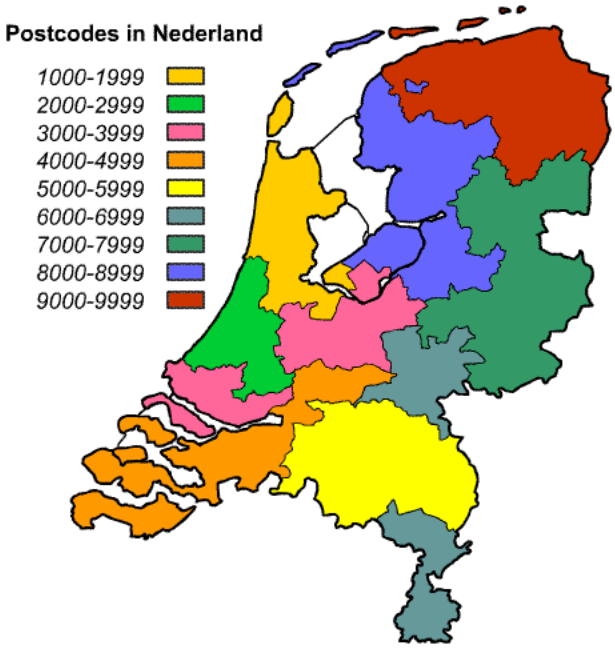


Figure 9, Netherlands postal code areas

5 Discussion

Using data analytics by European enterprises involved in delivering goods by road to improve consumer demand forecasting and operational efficiency is gaining traction rapidly. These firms expect this approach to help them become more nationally and internationally competitive (Duran-Lopez, 2020). This is true for several reasons; the most significant of which is that firms may enhance the quality of their strategic and operational decisions using data analytics. This is one of the most critical factors contributing to the current situation. In today's highly competitive business environment, businesses adept at making decisions under time pressure will have a greater chance of attaining economic success.

Utilizing data analytics across various application domains and contexts can enhance decision-making. That might be, for example, used to forecast demand and choose the most popular courses based on predictions of which methods would be the most popular. It might be used to identify the most common routes (Povoda, 2015). It might, for example, be utilized to forecast demand and identify the most popular courses to supply appropriate resources. This information may then be used to plan the allocation of resources to guarantee those resources are used in the most productive manner feasible to meet consumer demands.

Implementing this objective may be performed in various ways, one of which is by using the available information to organize the distribution of the available resources. In addition, organizations can increase the quality of their operational choices by employing data analytics to identify patterns within their data and then implementing the learned knowledge. This helps firms to use their data more effectively (Purnus and Bodea 2017, pp.747). Consider the situation of a corporation that noticed that the number of customers who visited their facilities on certain days of the week increased. In such a situation, they may use data analytics to identify why this was occurring and what measures they may take to resolve the problem by understanding why it was happening.

Comprehensive data analytics is an essential tool that, when used by European road transportation businesses, may assist these companies in making better decisions about the

techniques they use to manage their operations (Shahiri, 2015). It seems possible this is the case, given these organizations potentially have access to more data. Imagine that a company uses data to forecast future demand and meticulously plan its delivery routes. It can therefore meet the needs of its consumers in a timely and effective manner, in line with their expectations. This ensures the company can uphold its end of the bargain and fulfill its responsibilities.

When firms use data analytics for demand forecasting and enhancing operational efficiency, they can make more informed decisions about their functional operations. This is one of the most significant benefits that may be acquired via the usage of data analytics. Data analytics might be used, for instance, to forecast which routes would be the most popular and to manage the allocation of resources in the most effective way to meet the needs of consumers (Reddy, 2019). This would provide clients with a more gratifying experience. Besides predicting which routes will be the most popular, data analysis may also be used to predict which routes will be the trendiest. Data analytics may also be advantageous to organizations since it helps them to see patterns in their data and utilize the resulting information to improve their operational choices because of the enhancements made possible by these enhancements. Data analytics helps organizations to see trends in their data, which may be advantageous for enterprises.

Using big data in their systems, businesses may enhance their operations, customer service, and marketing efforts to boost revenue and profits. Businesses that use it can get an advantage over rivals that don't because the former can act with greater speed and accuracy when making strategic judgments. Big data can help businesses improve their marketing, advertising, and promotions, for instance, by revealing previously unknown facets of their target audiences, increasing the likelihood those audiences will become paying customers. By analyzing past and present data, firms can better cater to their clients' wants and needs (Aiello, 2020). Medical researchers use big data to find symptoms and risk factors for diseases, and clinicians use it to better diagnose their patients. Healthcare institutions and government agencies have access to the most recent data on infectious disease threats or outbreaks by combining data from electronic health records, social media sites, the web, and other sources (Laplante, 2015).

Some other ways businesses make use of big data are as follows:

- Big data is used by oil and gas businesses to locate promising drilling sites and pipeline operations and by utilities to monitor and maintain their power grids.
- The financial sector increasingly relies on big data technologies for risk management and real-time market research.
- Big data is used by manufacturing and transportation firms for supply chain management and route optimization.
- Emergency services, crime prevention, and "smart city" programmes are among some of the government's other applications (Wang, 2017).

There are a plethora of places where one can find big data, such as transaction processing systems, customer data, documents, emails, medical records, clickstream logs, mobile apps, and social media platforms. Data from sensors on manufacturing machinery, industrial equipment, and internet-of-things devices also fall under this category.

External data on consumers, financial markets, traffic and weather conditions, geographical information, scientific research, and more are commonly incorporated into big data settings in addition to data from internal systems. Also included in the category of "big data" are non-textual assets, such as images, videos, and audio recordings, the processing and collection of which is a central feature of many big data applications.

One of the most defining features of big data is its sheer volume. The nature of the information gathered and kept in a big data environment typically necessitates a huge volume of data to be present; however, this is not always the case. Continuously huge amounts of data are generally produced by sources like clickstreams, system logs, and stream processing systems.

As a whole, big data includes many kinds of information: There is a wide variety of data types, including structured data like financial records and transaction logs, unstructured data like text, documents, and multimedia files, and semi-structured data like web server logs and streaming data from sensors.

Several data kinds might need to be kept and managed jointly in large data systems. The many data sets used in big data applications are not always integrated at the outset. One application of big data analytics is the prediction of future product demand by analyzing historical transactional data, such as purchases, returns, reviews, and support requests. The

rate at which information is created, processed, and analyzed is called its "velocity." Many traditional data warehouses only update their information once a day, once a week, or once a month, but big data sets are often updated in real-time. As big data analysis evolves into artificial intelligence and machine learning (AI) that can automatically identify patterns in the data and employ those patterns to generate insights, it will become increasingly important to manage the speed at which data is generated.

Businesses in Europe that provide road transportation services may increase their operational efficiency by using data analytics methods for forecasting and optimization. Companies are now in a far better position to meet the demands of their consumers in a fast and effective manner. This development is beneficial for both sides. Data analytics may be used for several objectives, such as forecasting, optimization, and assessing whether a firm is likely to encounter problems soon (Papamitsiou, 2014). In addition, it may be used to determine if a technique has the potential to result in possible advantages by analyzing whether an edge can be acquired. Suppose European enterprises that provide road transportation-related services use data analytics with other information, such as logs and traffic reports. In that case, they can detect issues more rapidly and take corrective action steps. If businesses operate in this way, they can boost the efficiency of their operations and reduce the negative impact of disruptions on the customers they serve (Amin, 2019).

Overall, data analytics is a formidable instrument that, when used by European road transportation corporations, can aid these organizations in making better operational choices. This is one of the possible advantages of integrating data analytics in European road transportation firms (Reddy, 2019). It seems probable this is the case, given these organizations potentially have access to more data. When businesses employ data analytics for goals such as forecasting and optimization, they can increase their operations' efficiency, lowering the negative impact of disruptions on their consumers. This circumstance benefits all parties concerned. In addition, data analytics may be used in a process to identify any possible problems. This may be done to prevent future difficulties.

The primary objective of the literature review was to examine the existing literature and draw conclusions about the benefits of data analytics and big data for transportation businesses. This analysis looks at how big data and related analytics are now being used to enhance transportation systems' effectiveness and longevity. Data scientists in the transport industry are still in infancy, as made clear throughout the evaluation period. Significant

projects on the topic are academic research documents done in universities; however, the projects are not yet deployed to scale in the practical world. The review aims to identify the steps made in data analytics rather than prescribe solutions to challenges experienced in the sector. The review identifies numerous possible advantages of data analytics. It's important to keep in mind, however, these advantages have not fully manifested or been confirmed. Data analytics and big data are emerging technologies that will take time to mature and apply on a large scale.

In the transportation environment, an intelligent transportation system can be built using the big data platforms or approaches discussed in the literature (ITS). To improve the environmental friendliness and operational efficacy of transportation networks, this management system makes use of several technologies, including sensors and information communication systems and communication infrastructures (Muslim, 2016). This innovation in transportation ensures a cleaner, safer, and more user-friendly experience. The BIDdata architecture utilised by modern "smart cities" is one type of ITS system. It has seven distinct parts: data collection, data cleaning, data wrangling, data warehousing, data analysis, and data presentation. Data gathering and interpretation, as well as data analytics and visualisation, are two overarching categories that can be used to describe the layers (Lepenioti, 2020). One element of normalising data sourcing gathers data from various sources, transforms it, and loads it into databases. Data broking solves the heterogeneity of data challenges. It deals with various formats, frequency of data updates, and integration of data. Such a category extracts meaningful knowledge and important data patterns (Lepenioti, 2020). The visualization stage then organizes the results in graphical forms to help users make informed choices. These systems result in an efficient and sustainable transportation strategy for transport companies. The significant ideas highlighted include sharing real-time information, improving traffic safety by enhancing analysis of near misses and collisions, and using bid data for transportation and urban planning.

Sharing real-time traffic information is a crucial way of improving the safety and sustainability of traffic systems. When transportation companies are equipped with relevant traffic information, the companies can make informed transportation decisions (Wang, 2017). For instance, accurate time information may aid in the avoidance of traffic congestion as people can use the information to avoid congested routes. The choice of alternative routes

may help reduce losses that vehicles succumb to in traffic jams. Data analytics makes it possible to evaluate traffic data and relay relevant and valuable information to the drivers. Also, real-time information may help emergency vehicles reach their destinations in good time, saving lives (Zadeh, 2020). The information is usually shared using various communication channels. The channels include mobile applications, billboards, and traffic radios. The information is laid out in the form of alerts that brief drivers on the situation on the roads. These analytics are enhanced by systems such as ITs, which analyze traffic conditions. With the enhancement of sharing traffic information, data analytics ensures that transportation companies attain safety and sustainability in their areas of operations.

Historical analytics of transportation data helps in supporting evidence-based decision-making in urban and transportation planning. Planning for transportation infrastructure no longer requires decision-makers to rely exclusively on gut feelings or unproven hypotheses (Wilson, 2003). Vehicle movement patterns can be used to determine future infrastructure requirements by analyzing traffic flows and trends. Data analytics helps simulate traffic flow even before construction commences (Willis, 2021). This helps to avoid wastage in transport planning, which helps develop many means of transport. Analysis of vehicle wait times at different traffic lights and stops can also help optimize traffic flow, leading to efficiently operating transport companies. Transport companies use data analytics to improve their transportation networks by analyzing public transportation usage.

The transportation industry can benefit from a deeper understanding of the causes and consequences of accidents and near accidents. It helps monitor and visualize places and locations to identify accident-prone areas and causes of collisions (Lepenioti, 2020). Data attributes, such as frequencies, causes of collisions, and types of vehicles involved, are aggregated during data analysis. Transportation stakeholders can use the map to evaluate the current state of traffic and policy. Analyses such as proactive safety analysis of vehicles can distinguish between vehicles and pedestrians by analyzing metrics such as speed, directions, and characteristics of movement. Such analysis may reveal potential risks on the road and the situation of the road infrastructure. The analysis of traffic helps transportation companies to improve the security of commodities in transit by advising on vehicles that need the installation of trackers and video cameras. This helps in management of situations and visualization of goods in transit.

One significant opportunity that big data presents to the sector is a chance for transportation companies to use connected vehicles. Connected vehicles are a technology that proposes using GPS, sensors, and other communication technologies to enhance information sharing within a transport network. The data shared in this process can be employed for internal operation and review. The technology is effective because it enables the personalization of information provided to the driver. The information may include to-do lists, delivery instructions, and appointments.

Connected vehicles can check the documentation on the cloud platform to perform engine diagnostics and calculate optimal routes that the transport companies can use to reduce their delivery time and save fuel. The above are examples of how connected vehicles used by transport companies use data obtained from analytical techniques available in data analytics and big data literature. Further, data analytics reveals a future where autonomous vehicles can be adapted for transportation. The vehicles can eliminate the need for human-driving vehicles. Although this idea seems futuristic, recent literature provides evidence that advances in technology will undoubtedly result in efficient autonomous vehicles. The vehicles can help reduce the number of accidents on the road, eliminating human errors that occur on the roads due to fatigue. However, the literature is still scanty and requires more research.

The created tools are more or less buildable with slightly advanced level of Power BI and Excel knowledge, but only require entry level understanding of SQL databases. In order to be able to build similar and better tools in the future and improve the existing ones, it’s ony a matter of putting someone’s time into building them. When done inhouse, similar tools can be built in 40-80 hours. Coming back to the previous assumption, that these tools can save 90 days of working time annually, investing 5-10 days of time to save 90 days will turn out profitable. If thinking about building better, more big data based toolsets in the future, they are most likely cheaper to be outsourced at least in the beginning. If the results from them seem promising, a more comprehensive return on investment analysis should be made with the current set of pricing from consulting companies, different variable costs to be saved and if the production of tools should be in-housed or not. First things to develop more big data based tools would be route optimization, demand forecasting based on variables like vegetable growing season times or raising megatrends that could have an impact on the cargo volumes for Jylhä.

Based on the results of the interviews, it's worth mentioning that the information flow from customers to the company is not flawless. 40% of the personnel answered, that the lack of complete information from customers to the company in the order phase is the biggest reason for they are not getting enough information to complete their own work. That is an indicator that there is a lot of issues to be solved also in the booking phase that would solve some of the communication issues within the company. Internal communication is also one of the bigger reasons that personnel think they cannot complete their own work in the best way possible. That is a problem that's been tried to solve for a while now, for example on the thesis completed by Reetta Jylhä in 2018. (Jylhä, 2018)

6 Conclusions

This study has shown that transportation companies need to find ways to maintain profitability. The transportation industry must move cargo from one location to another by using trucks, but often they are forces to move from location A to B without any cargo. By minimising those distances, the utilisation rate of the vehicles will go up, which will result into higher profit margins. So, there may be a way through which can shorten the travelling distances if they are without cargo. The main scope of the study was to find a solution that helps the company to minimize the distance that trucks drive without cargo during their operational lifespan. The goal is to find a data-driven solution to help optimize the sales, planning, and monitoring processes to spot upcoming dead cargo zone in advance and to avoid travelling without cargo.

The two research questions are both answered in this study. There is room for bigger amount of knowledge management in the company, which was backed by the answers on the interviews. Lack of knowledge on the reasons behind decisions made in the company was apparent and could be fixed with more transparency. If a company starts to invest more into utilising its data, it’s easier to bring up the different reasons behind the decisions and get micro level details right more often. Also, scalability will be easier with better set of tools available. Direct savings that a company can gain when starting to invest into analytics is a better flow of information, more accurate decision making and saving working time when the data searched for is accessible and available. Also, if the company can be more scalable after investing more into the data-analytics they can save on the labour costs.

The goal is to help the client company become more nationally and internationally competitive. This is true for several reasons; the most significant of which is that firms may enhance the quality of their strategic and operational decisions using data analytics. This is one of the most critical factors contributing to the current situation. In today's highly competitive business environment, businesses adept at making decisions under time pressure will have a greater chance of attaining economic success. Utilizing data analytics across various application domains and contexts can enhance decision-making. It might, for example, be used to forecast demand and choose the most popular courses based on

predictions of which methods would be the most popular. Additionally, it might be used to identify the most common routes. It might, for example, be utilized to forecast demand and identify the most popular courses to supply resources. This information may then be used to plan the allocation of resources to guarantee those resources are used in the most productive manner feasible to meet consumer demands. Implementing this objective may be performed in various ways, one of which is by using the available information to organize the distribution of the available resources. In addition, organizations can increase the quality of their operational choices by employing data analytics to identify patterns within their data and then implementing the learned knowledge. This helps firms use their data more effectively. Consider the situation of a corporation that noticed the number of customers who visited their facilities on certain days of the week increased. In such a situation, they may use data analytics to identify why this was occurring and what measures they may take to resolve the problem by understanding why it was happening.

Utilising big data and better data analytics will bring better results according to the findings in this study. For example, historical analytics of transportation data helps in supporting evidence-based decision-making in urban and transportation planning. Planning for transportation infrastructure no longer requires decision-makers to rely exclusively on gut feelings or unproven hypotheses. Vehicle movement patterns can be used to determine future infrastructure requirements by analyzing traffic flows and trends. Data analytics helps simulate traffic flow even before construction commences. This helps to avoid wastage in transport planning, which helps develop many means of transport. Analysis of vehicle wait times at different traffic lights and stops can also help optimize traffic flow, leading to operating transport companies efficiently. Transport companies use data analytics to improve their transportation networks by analyzing public transportation usage. Data analysis of near misses and collisions helps improve safety in the transportation industry. It helps monitor and visualize places and locations to identify accident-prone areas and causes of collisions. Data attributes, such as frequencies, causes of collisions, and types of vehicles involved, are aggregated during data analysis. Transportation stakeholders can use the map to evaluate the current state of traffic and policy. Analyses such as proactive safety analysis of vehicles can distinguish between vehicles and pedestrians by analyzing metrics such as speed, directions, and characteristics of movement. Such analysis may reveal potential risks on the road and the situation of the road infrastructure. The analysis of traffic helps transportation companies to improve the security of commodities in transit by advising on

vehicles that need the installation of trackers and video cameras. This helps in management of situations and visualization of goods in transit.

Another significant opportunity that big data presents to the sector is a chance for transportation companies to use connected vehicles. Connected vehicles are a technology that proposes using GPS, sensors, and other communication technologies to enhance information sharing within a transport network. The data shared in this process can be employed for internal operation and review. The technology is effective because it enables the personalization of information provided to the driver. The information may include to-do lists, delivery instructions, and appointments.

Data analytics is a formidable instrument that, if used by European road transportation corporations, can aid these organizations in making better operational choices. This is one of the possible advantages of integrating data analytics in European road transportation firms. It seems probable this is the case, given these organizations potentially have access to more data. When businesses employ data analytics for goals such as forecasting and optimization, they can increase their operations' efficiency, lowering the negative impact of disruptions on their consumers. Data analytics can be used for a wide variety of purposes, including but not limited to enhancing vehicle performance, decreasing costs, enhancing procedures, establishing strategies, optimising routes and timings, and anticipating and identifying problems. Using advanced analytical methods including machine learning, Big Data, and geospatial data, transportation analytics assists industry executives in optimising business strategy. Enterprises can use predictive analytics to discover things like "What is the best probable outcome?" and "What is the most efficient route to successful distribution?" Companies can identify an alternate project that won't impede mobility by considering potential transportation-impacting events like weather, road closures, strikes, maintenance, traffic, and risk zones.

7 References

Ahmed, S. 2020. Edge computed NILM: a phone-based implementation using MobileNet compressed by AI. *Proceedings of the 5th International Workshop on Non-Intrusive Load MonitoringNovember, pp. 44–48*

Aiello, A, E. 2020. Social media- and internet based disease surveillance for public health. *National Library of Medicine.*

Amin, A. 2019. Big Data Role in Improving Intelligent Transportation Systems Safety: A Survey. *Advances In Internet, Data And Web Technologies,* pp. 187-199.

Babar, M. 2018. Real-time data processing scheme using big data analytics in the internet of things-based intelligent transportation environment.

Batko, K. 2022. The Use of Big Data Analytics in Healthcare. *National Library of Medicine.*

Channa, A. 2020. Retrieved from Robust technique to detect big data.

David, R. 2021. Tensorflow lite micro: Embedded machine learning for tinyml systems. *Proceedings of Machine Learning and Systems 3.*

Dokic, K. 2020. Inference speed and quantization of neural networks with Big data for Microcontrollers framework. *5th South-East Europe Design Automation, Computer Engineering, Computer Networks and Social Media Conference*

Duran-Lopez, L. 2020. A custom deep learning system to diagnose and locate important data. *Applied Sciences.*

Forward, A. 2002. The relevance of software documentation, tools and technologies: a survey. *Humanized Computing,* 10(10), pp. 4167-4177.

Finlay, R. 2022. 3 reasons why you need data to scale your company. Entrepeneur.

Fiore, S. 2019. An Integrated Big and Fast Data Analytics Platform for Smart Urban Transportation Management. *IEEE Access, 7,* pp. 117652-117677.

Ford, H. 2019. Guest Editorial: Big Data Analytics and Artificial Intelligence (AI) Applications for Smart Transportation - Selected Papers from World Transportation Congress (WTC) 2018. *IET Intelligent Transport Systems,* 13(3), pp. 425-426.

Hariri, R, H. 2019. Uncertainty in big data analytics: survey, opportunities, and challenges. *Journal Of Big Data,* 6(1).

Hevner, A. and Chatterjee, S., 2010. Introduction to design science research. In *Design Research in Information Systems* (pp. 1-8). Springer, Boston, MA.

IBM Cloud Education. (2021) *Structured vs. unstructured data: What's the difference?*, *Blog post.* Available at: https://www.ibm.com/cloud/blog/structured-vs-unstructured-data (Accessed: September 2022).

Jylhä, R. 2018. *Muutosjohtamien digitaalisen tehtävälistan käyttöönotossa.* Bachelor's thesis. Jyvsäskylä university of applied sciences.

Jylhä, T. 2022. *General interview about tranport jylhä*. Interview by Aleksi Sahlberg (over the phone), 19 September

Laplante, P, A. 2015. Encyclopedia of Information Systems and Technology. *CRC Press, pp. 102*

Lepenioti, K. 2020. Prescriptive analytics: Literature review and research challenges. *International Journal Of Information Management,* pp. 57-70.

Ltifi, H. 2013. A human-centered design approach for developing a dynamic decision support system based on knowledge discovery. *Journal of Decision Systems,* 22(2), pp. 69-96.

Muslim, A. 2016. *A rule-based indicator definition tool for personalized learning analytics. ACM,* p. 264.

Neilson, A. 2019. Systematic Review of the Literature on Big Data in the Transportation Domain: Concepts and Applications. *Big Data Research,* pp. 35-44.

Paliszkiewicz, J. 2020. Management in the Era of Big Data: Issues and Challenges. *CRC Press, pp. 77*

Papamitsiou, Z. 2014. Learning Analytics and Educational Data Mining in Practice: A Systematic Literature Review of Empirical Evidence. *Journal of Educational Technology & Society,* 17(4), pp. 49-64.

Ramesh, C. 2019. Security, Privacy and Forensics Issues in Big Data. *IGI Global, pp. 98-99*

Reddy, C. 2019. A survey on business intelligence tools for marketing, financial, and transportation services. *Intelligent computing and applications,* pp. 495-504.

Shahiri, A, M. 2015. A Review on Predicting Student's Performance Using Data Mining Techniques. *Procedia Computer Science,* 72(6), pp. 414-422.

Siegel, K. 2014. Organisational Models for Big Data and Analytics. *Journal of Organisation and Design.*

Siemens, G. 2012. Learning analytics and educational data mining: towards communication and collaboration. *ACM,* 6(8), pp. 252-300.

vom Brocke, J., Hevner, A. and Maedche, A., 2020. Introduction to design science research. In *Design science research. Cases* (pp. 1-13). Springer, Cham.

Wang, H. 2017. A big data analytics framework for border crossing transportation. *International Journal Of Big Data Intelligence,* 4(4), p. 227.

Willis, G. 2021. Using 'Big Data to understand the impacts of Uber on taxis in New York City. *Travel Behaviour and Society,* pp. 94-107.

Wilson, J. 2003. Gantt charts: A centenary appreciation. *European Journal of Operational Research,* 149(2), pp. 430-437.

Witten, I, H. 2016. Data Mining: Practical machine learning tools and techniques. *Morgan Kaufmann.*

Zadeh, A, H. 2020. Enhancing ERP learning outcomes through Microsoft dynamics. *Journal of Information Systems Education,* pp. 83-95.

Zeroual, A. 2019. Using a Fine-tuning method for a Deep authentication in Mobile Cloud Computing based on Tensorflow lite framework. *Journal of Information and Technology.*

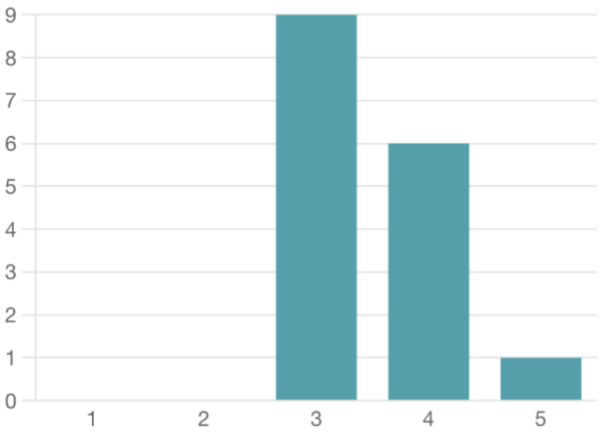
Appendix I

1. Työnkuvasi yrityksessä (Jos usea, valitse se mitä teet enimmäkseen)

Ajojärjestelijä	10
Myynti	1
Johto	1
Tilausten vastaanotto ja käsittely	4

2. Työssäni tarvittava tieto on helposti saatavilla. 1 = täysin eri mieltä / 5=täysin samaa mieltä

3.50
Average Rating

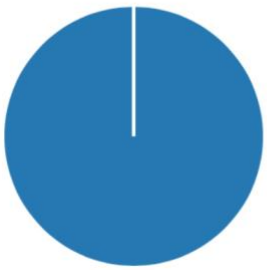


3. Käytän päivittäin ylimääräistä aikaa tiedon etsimiseen, koska en tiedä mistä tarvittavat tiedot löytäisin.

Kyllä	9
En	7

4. Joudun usein kysymään kollegoiltani lisätietoja, jotta voin hoitaa omat työtehtäväni.

Kyllä	16
En	0



5. Arvioi paljonko työaikaa säästäisit päivittäin, jos päivittäisissä tehtävissäsi tarvittava tieto olisi välittömästi saatavilla?

En yhtään	0
0-10min	0
10-30min	7
30-60min	9
Enemmän kuin 60min	0

6. Arvioi kuinka suuri osa yrityksen päivittäistoimintojen päätöksistä tehdään mielipiteiden perusteella, ilman konkreettista faktatietoa?

Ei ollenkaan	1
0-25%	8
25-50%	6
50-75%	1
75-100%	0

7. Koetko, että päätöksenteko perustuen faktatietoon johtaisi parempaan tulokseen?

Kyllä	15
Ei	1

8. Koetko, että paremmin saatavilla oleva tieto helpottaisi tai nopeuttaisi yrityksen kasvua?

Kyllä	16
Ei	0

9. Mikä seuraavista on mielestäsi suurin syy sille, ettet saa työssäsi tarpeeksi tietoa? Voit valita useamman.