Driven by data – A case study on how to become a more data-driven organization

Kati Robertson

Haaga-Helia University of Applied Sciences
Degree Programme in Business Technologies
Digital Business Opportunities
Master Thesis
2022

Abstract

Author
Kati Robertson

Degree
Master of Business Administration

Report/thesis title
Driven by data – A case study on how to become a more data-driven organization

Number of pages and appendix pages
51 + 8

This thesis' main objective was to investigate how the client organization could become more data-driven by researching what being a data-driven organization means and by demonstrating how data can better support the client organization's decision-making. Another objective was to promote data-driven culture within the client organization. The leaders of the client organization had expressed their interest in utilising data in decision-making which is why this study was undertaken.

The main expected outcome of this research was to understand the client's current and desired data and analytics state and based on these findings provide recommendations for the future. Another expected outcome was that the client organization's leaders' interest towards data-driven decision-making would increase while the benefits could be seen in the form of a demonstration and by participating in this study.

The thesis' theoretical part includes the description of two key concepts; what is meant by data-driven organization and what is modern analytics in a nutshell. These concepts are described in general terms and are based on literature. The purpose of this was to create a solid background and understanding of the topic.

A qualitative research approach was used in this thesis and case study was selected as a research strategy. Demonstration and future recommendations were also part of this study as they provided practical outcomes to the client organization. The data collection methods included questionnaire, workshop, semi-structured interview, document analysis and feedback survey.

This study showed that by understanding data-driven approaches and assessing the client organization's current and desired data and analytics state, recommendations for future steps can be made. In the process of demonstrating the benefits of data-driven approaches and advocating about the data-driven possibilities, the client organization's decision-makers interest towards data-driven journey was elevated. This is important in relation to promoting data-driven culture and it is promising for the future development in the client organization's journey to become more data-driven.

Keywords
Data-driven organization, data culture, analytics

Table of contents

1 Introduction 1
2 Objectives and Research Questions 4
2.1 Research questions 4
2.2 Scope 5
3 Data-driven organization 6
3.1 Data and Analytics maturity 6
3.1.1 Dimensions of the Maturity model 7
3.1.2 Maturity chart 8
3.2 Organization 9
3.2.1 Data-driven culture 10
3.2.2 Data literacy 11
3.2.3 Data strategy 12
3.3 Infrastructure 13
3.3.1 Data Architecture and technologies 14
3.3.2 Analytic platforms 16
3.4 Data Management 18
3.5 Analytics 20
3.5.1 Modern Analytics 20
3.5.2 Analytic techniques 21
3.5.3 Strategic Analytics 22
3.6 Governance 22
3.7 Decision-making 25
3.7.1 Decision-making process 25
3.7.2 Re-engineering decision-making 25
4 Methodology 27
4.1 Qualitative research 27
4.2 Case study 27
4.3 Data collection and analysis 29
4.4 Demonstration and recommendations 31
4.5 Quality of the research 31
4.5.1 Construct validity 32
4.5.2 Internal validity 31
4.5.3 Reliability 32
5 Results 33
5.1 The demonstration 33

5.1.1 Goals 34
5.1.2 Presentation 34
5.1.3 Feedback 35
5.2 Client organization’s maturity 37
5.3 Client organization’s data culture and literacy 38
5.4 Client organization’s infrastructure 39
5.5 Client organization’s data management 40
5.6 Client organization’s decision-making 40
5.7 Recommendations for the client organization for future 41
6 Conclusion 44
7 Discussion 47
References 48
Appendix 1. Questionnaire 51
Appendix 2. Workshop Agenda 56
Appendix 3. Semi-structured interview with CTO 57
Appendix 4. Feedback survey after demo 58

1 Introduction

The objective of the thesis is to look at a way of how to become more data-driven organization by understanding what a data-driven organization means and by demonstrating how the client organization's data can better support the decision-making of the Executive Board. Becoming data-driven can result in better and faster decision-making, create new business opportunities, and offer a competitive edge within the market. This is especially important in the ever-changing world where the market can shift rapidly due to prevailing circumstances.

The leaders of the client organization had expressed their interest in utilising data in their decision-making and that is where the idea for this thesis came from. The commissioning client organization (client) is a Finnish-owned company that operates in the entertainment and e-commerce industry. The client organization is an agile organization with committed leaders who are open for new technological approaches to deliver a high standard of service and optimised operations. As a middle-size company with its owner/CEO working within operations, decisions can be made quicker and new approaches adapted easier, compared to big corporations that are often non-flexible. The decision-makers are looking forward to learning and enjoying the benefits of what data and analytics can provide.

The goal is to increase the interest towards data-driven decision-making within the client organization by engaging the client organization in different activities, advocating about the topic, and showcasing the benefits of data-driven decision-making in the form of a demonstration.

Being a data-driven organization means that the organization has tools, abilities and culture that is driven by data (Anderson 2015, chapter 1). Culture is often named as the biggest contributing factor when talking about data analytics. A data-driven organization trusts data and uses it to inform and influence strategy. There is an evidence-based culture, and the data analysis is highly informative. (Anderson 2015, chapter 1.)

There are multiple benefits of being data-driven and having modern analytics in use. Chambers & Dinsmore (2015, chapter 4) describe the business value of analytics as growing top-line revenues cost efficiently. Modern Analytics can provide you with the tools to help explain what happened in the past but more importantly to help predict the future. Organizations that understand the possibilities of modern analytics and can produce results systematically are the clear winners.

(Chambers & Dinsmore 2015, chapter 4.)

This thesis consists of three parts; theoretical, empirical and conclusion.

The thesis' theoretical part includes the description of two key concepts; what does data-driven organization entail and what is modern analytics in a nutshell. These concepts are being described in general terms and base on literature. This is to create a solid background and understanding of the topic and help showcase things that are beneficial for the client organization.

The empirical part consists of analysing the results of the data-collection and the demonstration. The demonstration showcases how data can be incorporated in the organization's decision-making and creates awareness of data-driven approaches. The information needed is collected by questionnaire, workshop, semi-structured interview as well as document analysis and a feedback survey.

The research investigates the client's current and the desired states; what and how is data currently used for decision-making, and what is wanted from the data in the future. The topic for the demo is chosen by the Executive Board of the company to achieve the most value to the organization. After the demo presentation, feedback from the executive board is collected and analysed. Recommendations for the next steps to take in the client organization's journey to becoming transformationally data-driven organization, are provided in the end of the results chapter.

The thesis is compiled using qualitative research and the research strategy for this study is a case study. Case study strategy was selected as it focuses on real life problems and cases. Case study helps the researcher to understand a particular phenomenon, dig deeper and with this knowledge extend or test theory (Farquhar 2012, chapter 1). The assessment of the quality of the research is presented in end of the methodology chapter.

In the concluding part of the thesis, thoughts of the whole thesis process and the findings are presented. The researcher's comments are presented in the discussion section.

Figure 1 below presents the overall structure of this thesis.

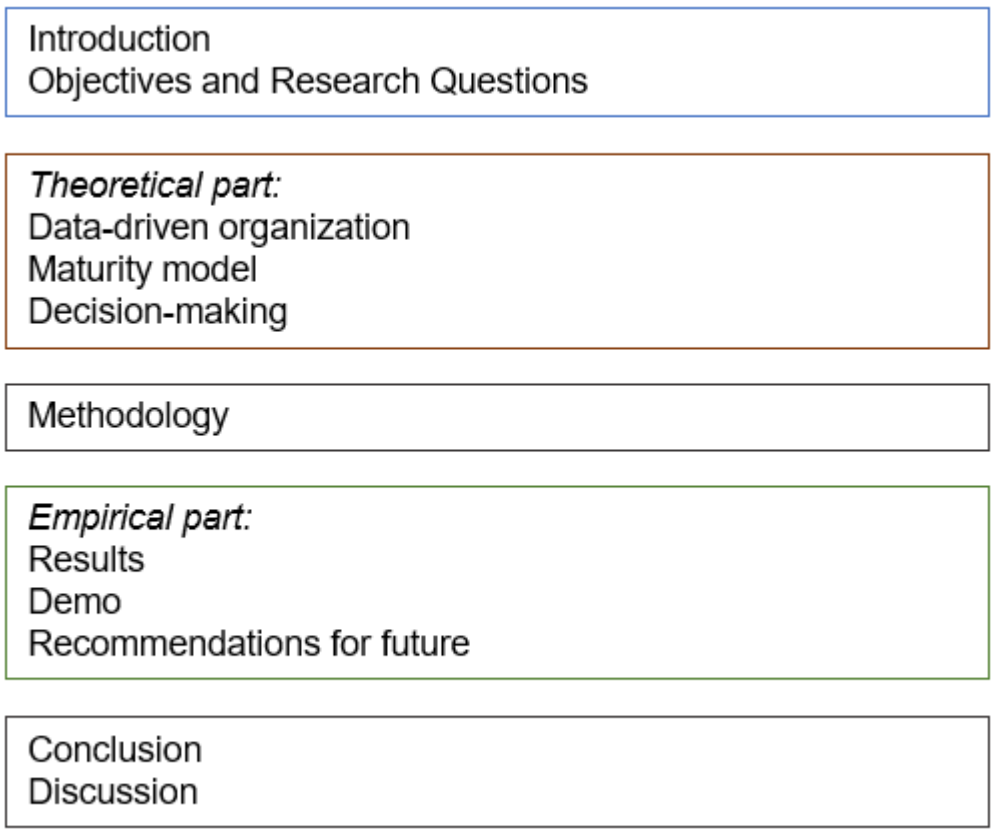


Figure 1. Structure of the thesis

2 Objectives and Research Questions

The objectives of the thesis revolve around identifying data-driven concepts, promoting data-driven culture within the client organization, and communicating the business value of data, particularly in the decision-making part. The objectives are divided into three categories as follows:

Understand what data-driven organization means in general and what it includes
This objective creates a theoretical framework for the thesis and helps in understanding the bigger picture of a data-driven organization. By understanding the bigger picture, it is easier to focus on the areas that are most crucial for the client organization. Also, by knowing the benefits and challenges of being data-driven makes it easier to promote data-driven culture within the client organization and create a push towards becoming more data-driven. The review will be written in general terms, and it is based on literature.

Promote data-driven culture and attitude change in the client organization
The first step in this objective is to measure the client's attitude towards data and analytics and record how the client uses data in the decision-making process currently. The client's wants, needs and expectations for future data use will be collected and kept in mind when promoting the culture change and planning the recommendations for future. In the process of collecting the data and writing this paper, the researcher educates the client organization's executive board with general knowledge of data-driven approaches and decision-making which in turn affects the culture shift positively.

Showcase the business value of data in a form of a demonstration
The final major objective is to demonstrate the business value by conducting a demo of mutually decided topic using Google Data Studio. This will help realise the potential of analytics and open doors for wider use of analytics.

2.1 Research questions

The written research questions below help meet the objectives of the thesis.

RQ1. What is meant by data-driven organization?
Sub-question 1: What dimensions does data-driven organization have?
Sub-question 2: What are the benefits and challenges of data-driven organization?

By answering these questions and understanding the overall concept, makes it easier to analyze and help the client organization. These questions guide the writing process of the theoretical framework for the whole thesis.

RQ2. How can Data and Analytics culture be promoted within the client organization?
This question helps in finding ways to advocate and create a positive push towards adaptive organizational culture for data and analytics.

RQ3. How can the benefits of data-driven decision-making be showcased to the client organization?
This question helps in creating a demo to showcase the benefits for the client organization.

RQ4. What recommendations can be given for future development?
Sub-question 1: What is the client organization's current state in data and analytics?
Sub-question 2: What is the client organization's desired state in data and analytics?
Discovering the current and desired state help in forming recommendations for future.

2.2 Scope

The thesis includes general information of the key concepts of data-driven organization and modern analytics. These concepts offer background information of the topic and create a theoretical framework for the thesis. To understand the concepts is necessary in order to help the client organization in becoming more data-driven.

The thesis does not include a detailed plan for the client to become data-driven nor does it provide a written data strategy. It also doesn't include implementations, excluding the mutually agreed demonstration. The demonstration brings theory to practice. Its main purpose is to showcase the benefits of using data to support decision-making in the area set by the executive board. Recommendations for the future on how to continue in the journey of becoming data-driven will be provided.

3 Data-driven organization

This chapter covers the basics of data-driven organization and modern analytics as well as defines some terms linked to data and analytics. Decision-making is also introduced in the end of this chapter as it is a crucial part of data-driven organizations. This chapter provides answers to research question 1 and its sub-questions, and also provides ideas for research question 2.

Simply put, being data-driven means that the organization has tools, abilities and culture that drives on data. Having data is the foundation of being data-driven. If the organization doesn't have quality data, it cannot make decisions based on facts, instead making them based on gut feeling or experience.

To be able to achieve quality insights or answer given questions the data also needs to be collected in the right manner. This means the data needs to be in the right form, from the right place and taken at the right time. (Anderson 2015, chapter 1-2.)

In the most successful data-driven organizations data is treated and viewed as a core asset and seen as competitive advantage, improving their ability to innovate. The leaders support systematic use of data in strategy and decision-making. They don't treat data just as something to support basic reporting but as a tool to change business models and create new products and services. (Kiron 2017, 9-10.)

The upcoming sub-chapters present different data-driven components that need to be considered when planning a data-driven organization. The components are divided into five categories and are named similarly to the TDWI Analytic maturity model. Each category contains details of chosen components that are selected and put together by the researcher. These categories are as follows: organization, infrastructure, data management, analytics and governance. TDWI Analytic maturity model is presented in more detail in the next sub-chapter.

3.1 Data and Analytics maturity

To create a 360 view on the organization's data and analytics capabilities and current practises, it is good to start with assessing the organization's data and analytics maturity. There are different models for evaluating the organization's analytics maturity. TDWI Analytics Maturity Model was selected for this thesis as it was found comprehensive, and it provides a broad view on different data-driven capabilities within an organization.

3.1.1 Dimensions of the Maturity model

TDWI Analytics maturity model includes assessment criteria for five categories and 35 questions that help determine the organization's analytics maturity. These dimensions are presented in the chart (Figure 2) below (Halper & Stodde 2014, 5).

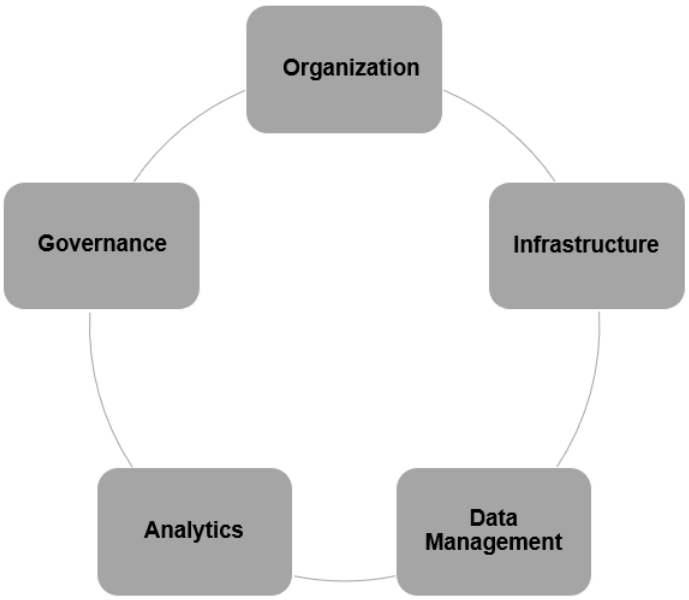


Figure 2. Five criteria categories for Analytics assessment (adapted from Halper & Stodde 2014, 5)

The organization category helps determine whether organizational strategy, culture, leadership, skills and funding successfully support the analytics program. Infrastructure looks at the architecture and technologies and how well they support analytics initiative. Data management dives into data quality, access, integration, processing and of how data is used in analytics. The current use of analytics and its culture are assessed in the analytics section while governance investigates further into the data governance strategy within the organization. (Halper & Stodde 2014, 4.)

Based on these components an overall assessment of the organization's maturity can be done.

3.1.2 Maturity chart

After assessing the individual components, the organization is placed on a level matching its maturity. TDWI Analytics maturity model presents a five-level maturity chart (Figure 3) that is pictured below (Halper & Stodde 2014, 9).

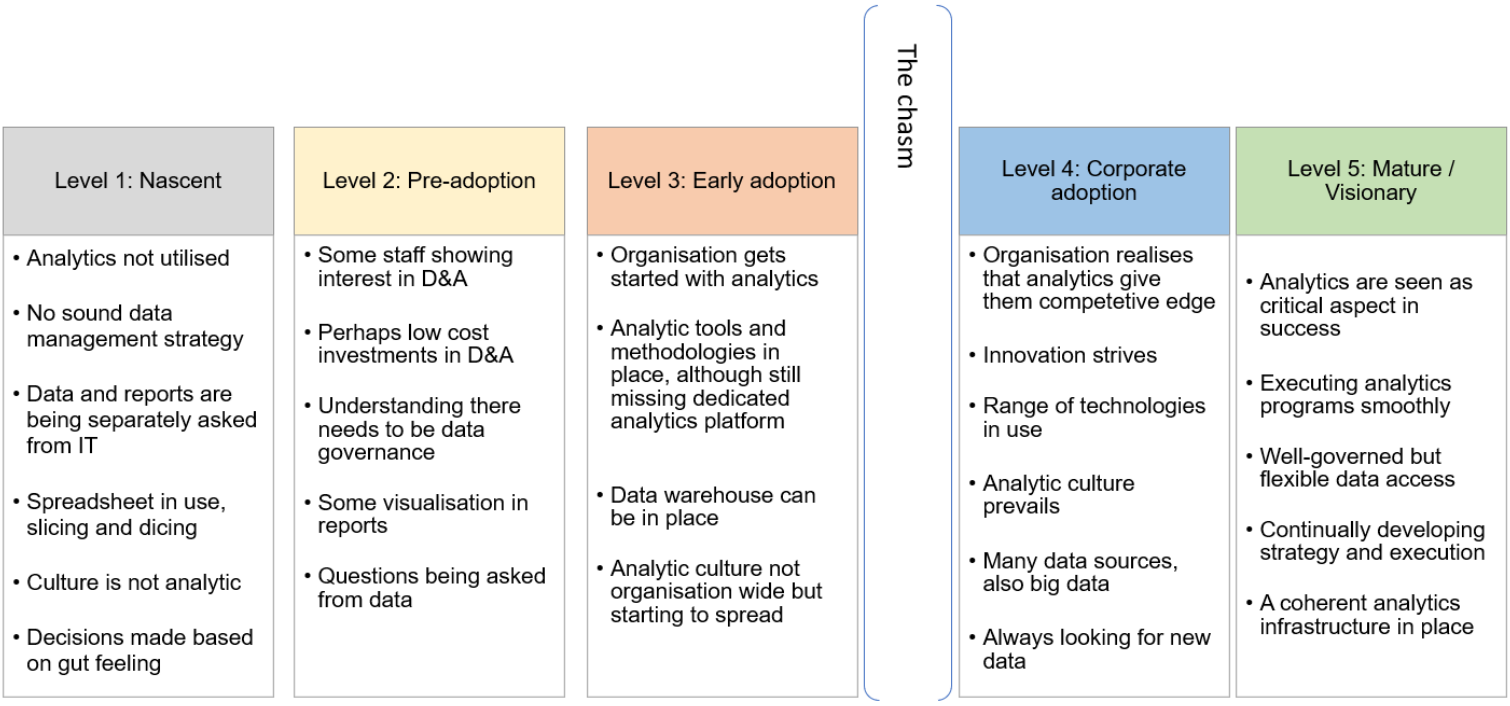


Figure 3. Maturity Model for Data and Analytics (adapted from Halper & Stodde 2014, 9-17)

In the first level of the maturity model, the analytic practises aren't supported, realised or in place. Interest and initiatives grow more as the next levels are reached. The second level already includes some low-cost investments, and the understanding of the D&A requirements has grown. The third level is reached when the organization really gets started with the analytics and the data culture starts to spread among the organization. Between level 3 and 4 there are usually a series of challenges related to funding, data management & governance, cultural & political issues and skills. This gap is called the chasm and many organizations spend large amounts of time solving these issues before moving to the corporate level of maturity. (Halper & Stodde 2014, 10-14.)

To be able to reach transformational stages of data and analytics maturity, you need to first look inward and focus on improving the coordination of processes between the people inside the organization and only then look at how to enhance practises with external partners or acquire new technologies. (McCall & Meulen, 2018.)

To reach the highest level in the maturity chart means increased agility, better integration with partners and easier uses of advanced predictive and prescriptive analytics. All this gives the organization competitive advantage and differentiation. (McCall & Meulen, 2018.)

3.2 Organization

In today's world, technology has enabled easy and inexpensive access to data and analytics. However, technology isn't the only factor in the data and analytics revolution, people within the organization are equally important. On top of having access to data, people also need a process of how to use data and analytics. (Kiron 2017, 3.)

After assessing the organization's overall data and analytics maturity, it is time to look at the organizational components that have been divided into three sub-chapters in this thesis and are as follows: Data-driven culture, data literacy and data strategy.

Figure 4. Organization category for Analytics assessment (adapted from Halper & Stodde 2014, 5)

3.2.1 Data-driven culture

A defining feature described within the literature is the importance of promoting an inspiring data culture within the organizations.

Anderson (2015) states that the most crucial part of being data-driven is to have a culture that drives on data. The data culture among the organization influences what can be shared, who has access to what data, and what investments are made into tools and people. The culture aspect comes into play when talking about sharing data organization-wide instead of keeping it in individual silos or departments. Data sharing should be done without having to comprise security issues and privacy. Employees need to trust that the data is reliable and that it will be used for good cause and the employer needs to trust that the individuals won't use data against them or leak it to others. (Anderson 2015, chapter 10.)

Data sharing extends outside the organization's borders as valuable data can be acquired through cross-organization sharing. In some scenarios businesses and vendors can turn into partners and realise new business opportunities together. (Kiron 2017, 7-8.)

The leaders in the organization should be advocates for data-driven practises and promote towards a data-driven culture. Creating a healthy data culture is not just about giving employees access to data but also to promote an inquisitive culture where everyone is encouraged to challenge, ask questions, dig deeper and learn from failures. It is important to show results and cherish the small wins as that deepens the belief that data is the way to go. (Anderson 2015, chapter 10.)

The change in culture won't happen just by demanding everyone to change their ways as a business culture is often founded on deep-seated beliefs and emotions. Instead, showing how data can influence working practises and decision-making and articulating measurable business outcomes can persuade the organization shifting their ways. (Goasduff 2016.)

Integrating data into decision-making can be hard for managers that are used to act on experience, anecdotal evidence or gut feeling but to be a successful data-driven organization, it is essential to have commitment from the management (Kiron 2017, 4). When employees realise how data has helped them do their job, they become data advocates, which in turn creates more positive attitude towards data usage (Kiron 2017, 8.)

MIT Sloan Management review found that the main obstacle wasn't the actual data management

or its complexity but rather the skill of turning data insights into business actions. To help with this there are ways to support the transformation such as experimental training and formal training. It is also noteworthy to say that the managers need to start trusting the data and people presenting the data as it might feel foreign having made decision based mostly on gut feeling. (Kiron 2017, 12.)

3.2.2 Data literacy

As Saulles (2019) states, it cannot be expected to derive insights from data, unless you are trained to do so. Part of creating a data-culture is to have data literate people in the organization.

White (2019, 9) defines data literacy as presented in the picture (Figure 5) below.

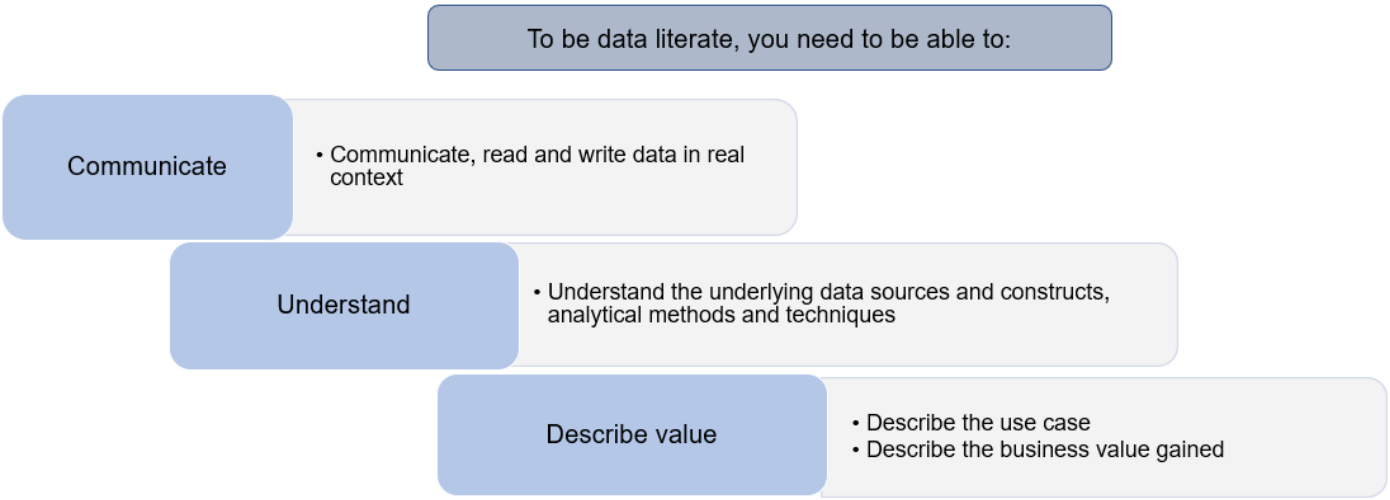


Figure 5. Being data literate (adapted from White 2019, 9)

To be data literate, you need to be able to communicate and understand data as well as describe the business value and use cases (White 2019, 9). And to make good decisions you need be data literate (White 2021, 15).

Data Analysts or Chief Data Officers (CDO) are trained in analytics, but in order for the organization to be data-driven, the decision-makers and managers need to be data literate too. These people don't need know the detailed mechanisms of the whole field but should have a clear understanding of what good experimental design is and basic statistical interference alongside with some basic terminology and metrics. They should also be aware of the potential dangers of

exploitation. (Anderson 2015, chapter 10.)

Data literacy among the workforce can be enhanced with different programs, tailored training or even free online resources. The idea is to start somewhere and awaken the employees' skills and help them get comfortable with data, tools and analysis. (Anderson 2015, chapter 10.) In Gartner survey, data literacy is seen as explicit and necessary driver of business value by 2023 (White 2021, 15).

3.2.3 Data strategy

A clear data and analytics strategy is vital for succeeding with data and analytics investments. The organization should start with a common understanding of the mission and then prioritize the most valuable business outcomes. Keeping business goals in mind and having a priority list, makes it easier to target the organization's data and analytics strategy. (White 2019, 5.)

There are different approaches that can be used to develop the organization's data strategy if there isn't one. Nelson (2018,100) has developed iterative analytics strategy development method that is presented in the picture (Figure 7) below.

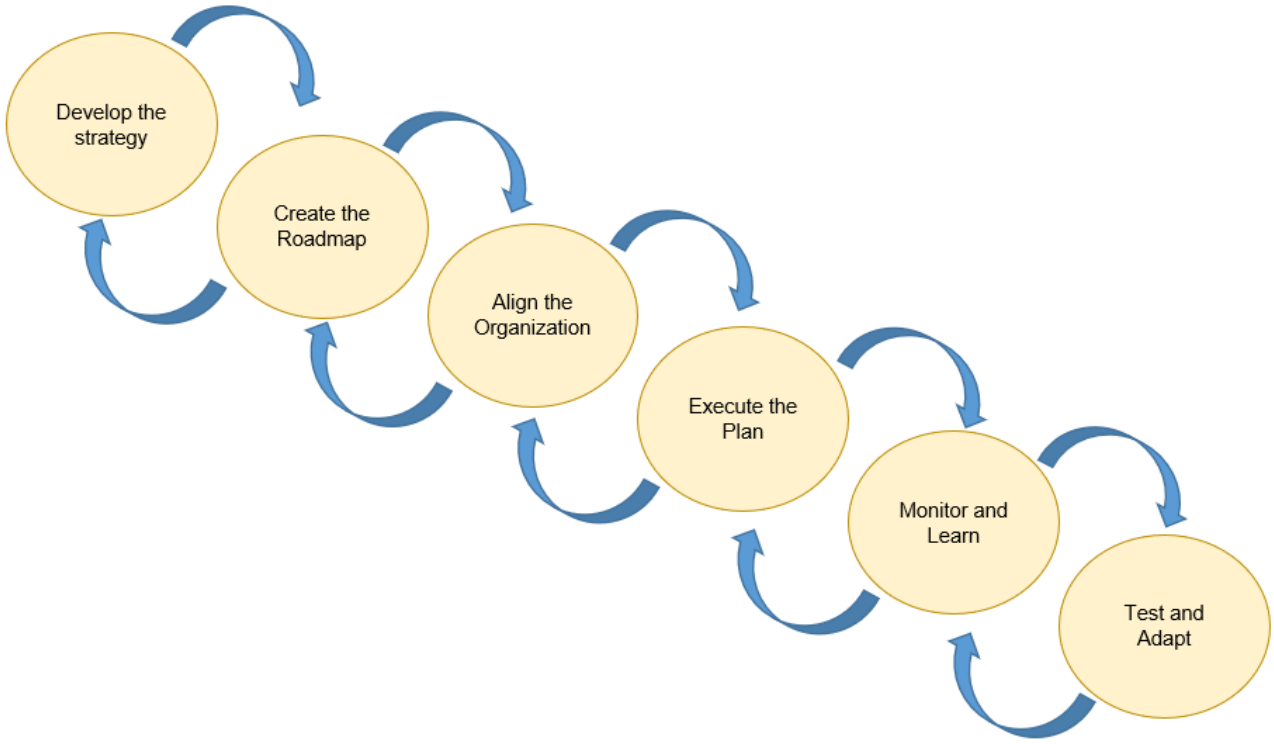


Figure 7. Analytics strategy development process (adapted from Nelson 2018, 100)

This strategy development process can help organizations with drafting a strategy and a suitable roadmap. The first step is to develop a data strategy and state exactly what business they are in, identify key issues and how to reach best possible outcomes. Vision and mission need to be clear. In the second step the vision, the mission and strategy need to be translated into execution and measurement plan. The roadmap will guide action and resource allocation.

To optimize the strategy execution in step three, you need strong management methods to motivate, engage and align the organization. Once it comes to executing the plan in step four, the day-to-day operations need to be linked with the long-term strategy and operational impacts need to be built into the workflows. In step five, once the strategy has been implemented, monitoring and learning about challenges will help you determine if the strategy has been executed correctly. The final step in the cycle – test and adapt- will help determine if the strategy has been right using internal operational data with new external data. Then the whole cycle can start again. (Nelson 2018, 100-103.)

3.3 Infrastructure

As data becomes more distributed and diverse and the organizations more complex, data management infrastructure needs to be planned to meet these requirements (Friedman & Heudecker 2021). The following two sub-chapters cover the main aspects of the infrastructure; architecture and analytic platforms.

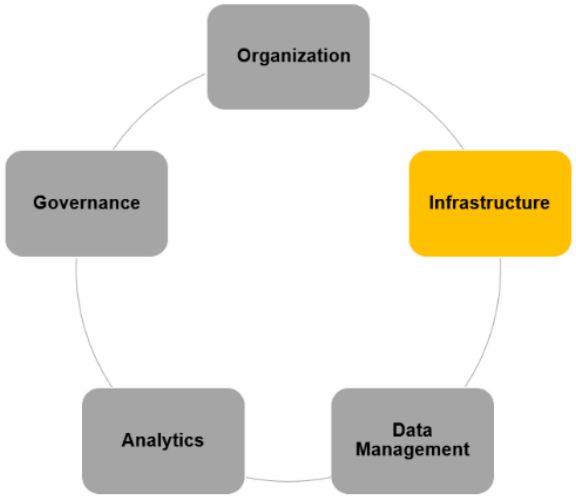


Figure 8. Infrastructure category for Analytics assessment (adapted from Halper & Stodde 2014, 5)

3.3.1 Data Architecture and technologies

When planning the overall analytics architecture, it is good to keep asking oneself questions revolving storage, access, data format and archiving. It is often not essential to keep all data available in one place and in real time. Storing data can be done on premises or in the cloud. (Nelson 2018, 106-107.)

It is important to understand the difference in architectural patterns and terminology as they create a lot of confusion. The best results can often be found in combining different data organization and processing approaches. Gartner defines data lake, data warehouse and data hub as pictured in image (Figure 9) below. (Friedman & Heudecker 2021.)

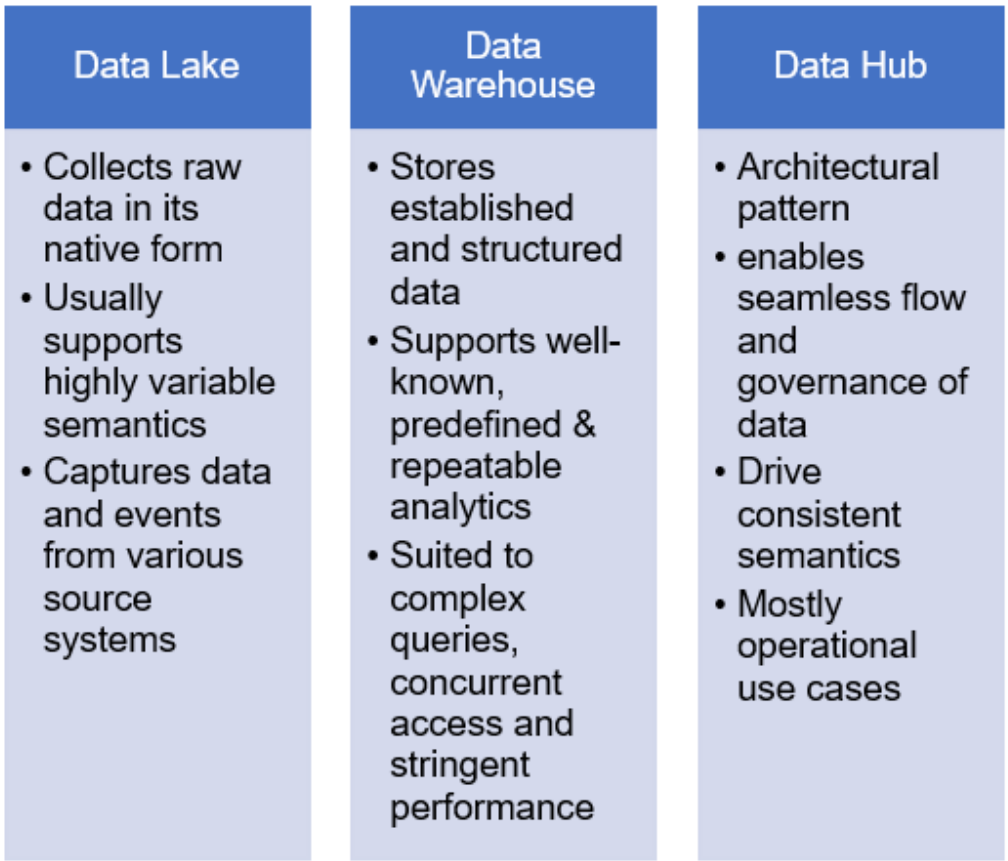


Figure 9. Data organization and processing approaches (adapted from Friedman & Heudecker 2021)

Technologies evolve constantly and new services appear to support or replace existing ones. What used to be novelty 10 years ago can be a common part of the infrastructure these days. The architectural offerings and their names can differ depending on the service provider. There are multiple

reference architectures available but the one from Amazon Web Services (AWS) was chosen to be displayed in this thesis as it was the most familiar to the researcher.

The graph (Figure 10) below presents data flow in the Amazon Web Services reference architecture for modern data analytics.

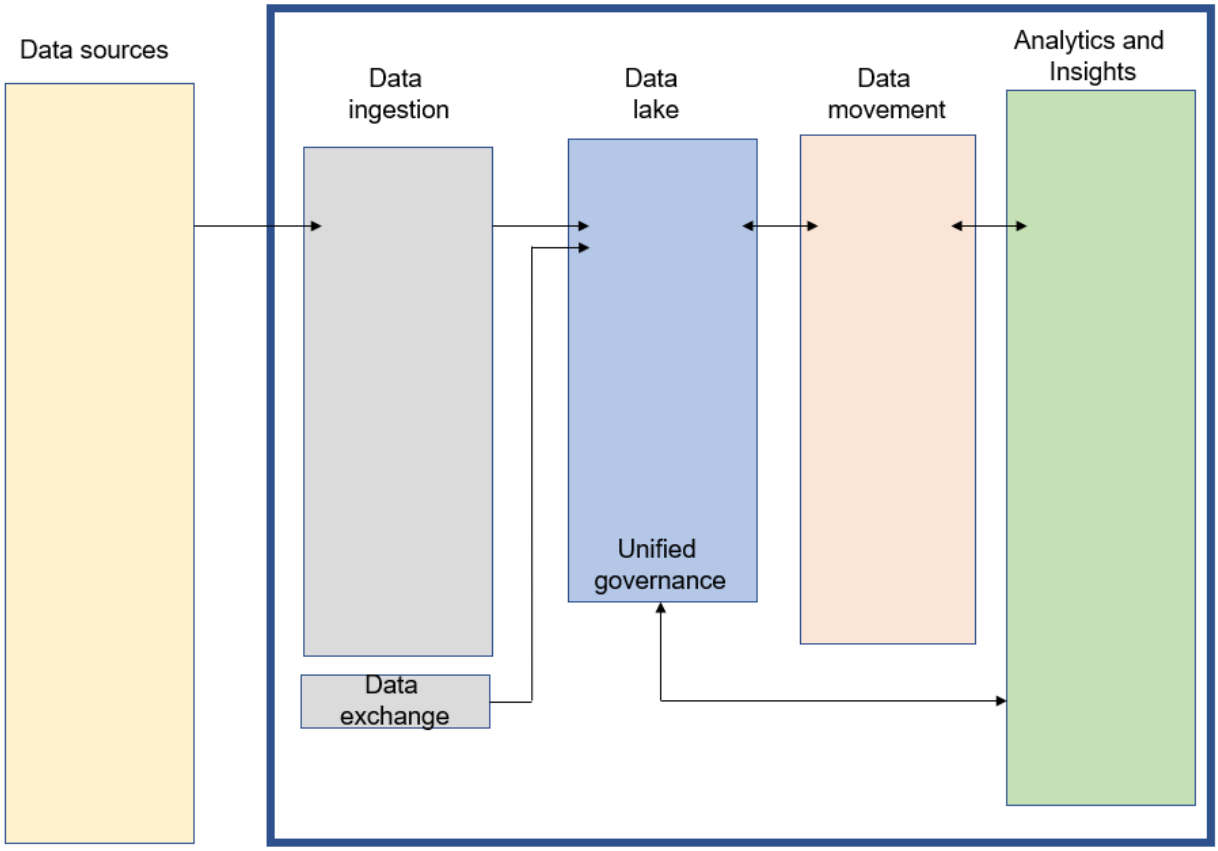


Figure 10. AWS reference architecture (adapted from AWS 2021)

In the AWS example, data is being collected from different sources across the organization including ERP, SaaS platforms, logs, social networks and streaming media. Using different services based on the data source, the data is then being ingested into a data lake. If third party data is needed, then data exchange is used for the integration. Data is then stored within the scalable data lake. AWS offers unified governance to audit trails and centrally manage security and access control. AWS provides glue service for seamless data movement. This can be used to extract, transform, catalog and ingest data from various data stores. Analytics and insights can then be drawn using different Amazon services. For example Amazon Redshift can be used as data warehouse and Amazon OpenSearch for operational analytics. (AWS 2021.)

3.3.2 Analytic platforms

To choose the best fitting analytic platform depends on the organization's needs, budget and overall architecture. There are a myriad of software and hardware vendors offering technologies that enable analytics. There are solutions for nearly every need and step in the analytics ecosystem from data storage to analysis and presentation. (Nelson 2018, 115-116.)

Advanced analytics used to require command line code and were complicated to use, however in current times the interfaces have been simplified so they are easier to use and report visualizations are quicker to construct. Even blending, preparing and integrating data has been made simple. Ease of use also helps the organizations gain higher level of maturity as it can aid organizations to gain successes early and scale analytics for wider use within the organization (Halper & Stodde 2014, 6.)

On January 31st 2022 alone there were 21 572 start-up companies under the analytic industry on Crunchbase which is a leading platform for finding company insights from start-up companies (Crunchbase 2022).

According to Gartner's Analytics and Business Intelligence Platforms Reviews and Ratings on 2nd February 2022, the top four analytic platforms ranked by the highest number of ratings and the highest score by peers are presented in the chart (Figure 11) on the next page. (Gartner Peer Insights 2022.)

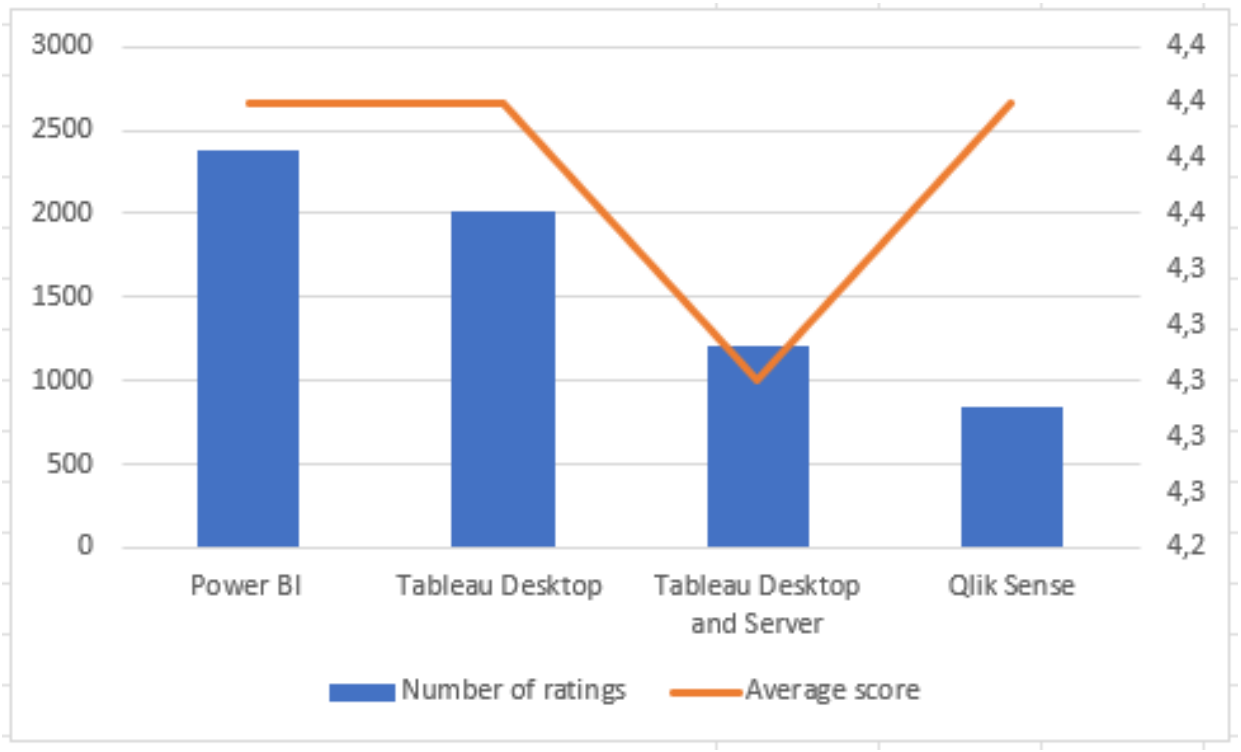


Figure 11. The best D&A platforms by peer ranking (adapted from Gartner Peer Insights 2022)

On the website peers can rate 17 product features ranging from Interactive Visual Exploration to Security and Use Administration. The highest ranking by average score (4,4) was received by Microsoft Power BI followed by Tableau Desktop and Qlik Sense. Power BI had received the highest number of ratings.

On top of modern analytic tools, there are data visualisation tools such as Google Data studio that can be connected to various data sources and are compatible to a number of platforms and sources (Google Data Studio 2022).

There is not one simple way of determining the best analytics platform for your organization but evaluating scope, complexity of decisions, urgency and impact might be useful when considering the technology landscape and platform required for analytics (Nelson 2018, 117).

3.4 Data Management

The organization needs to be collecting data to use it for their benefit. This sub-chapter examines the importance of data quality.

Figure 12. Data management category for Analytics assessment (adapted from Halper & Stodde 2014, 5)

Data can come in hundreds of formats and can be digitally transferred. Raw data, also called primary data, is something that the analyst has collected themselves and it is in its original form. Raw data is often preferred in analytics as it offers more flexibility and can be examined, tested for quality and summarised easier. Raw data turns into secondary data when it's been aggregated or summarized. (Herzog 2016, chapter 1.)

To get support from data, it needs to be the right type of data relevant to the issue. The data has to be, first and foremost, trustworthy alongside with timely, accurate and unbiased. It needs to be cleaned and prepared, which in most cases take the most of time of the data analyst's job. (Anderson 2015, chapter 1.)

The picture (Figure 13) on the next page represents what good quality data consists of according to Anderson (2015, chapter 2).

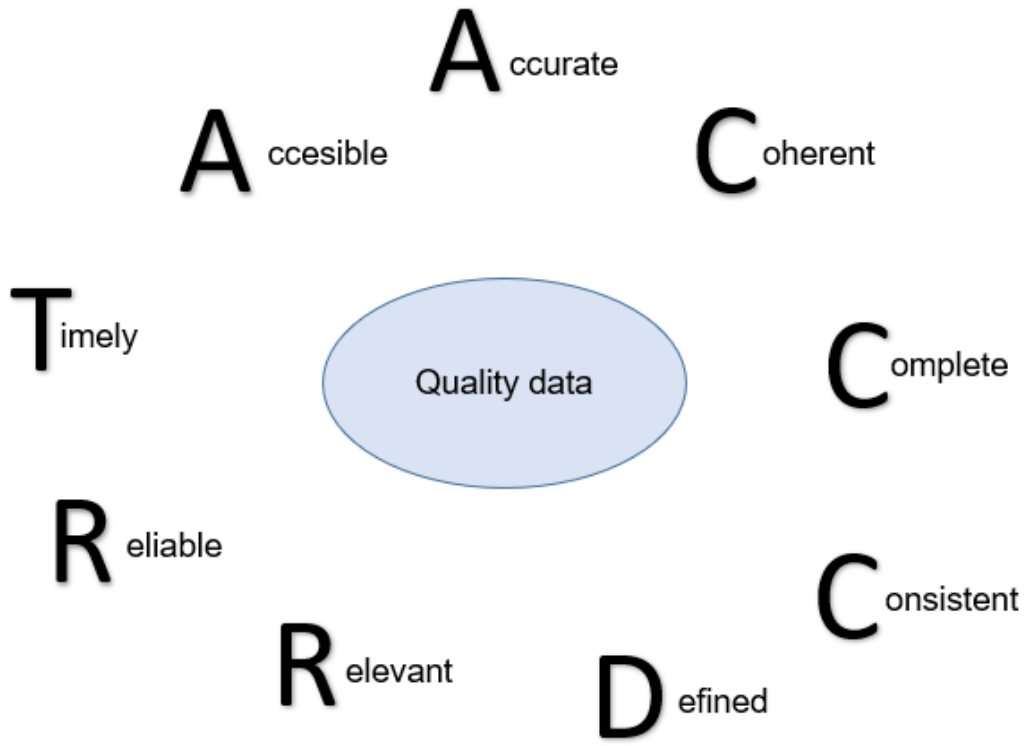


Figure 13. Quality data features (Anderson 2015, chapter 2)

For the data to be useful it needs to be accessible, accurate, coherent, complete and consistent. It also needs to be defined, relevant to the issue, reliable and timely. Data can be useless or misleading even if one of these facets fail. (Anderson 2015, chapter 2.)

To get the best business value out of the data, it's not enough for the right person to have the right information at the right time but the data needs to have the right importance to the person using it. This is especially important in the times of data overload. The person needs to be able to choose the right type of data from the vast selection of data. (Kiron 2017, 12.)

There are lots of terms related to data and below are examples of a few of them.

In today's world data is everywhere and the amount will keep increasing the more people get connected. People are already connected to multiple appliances that are collecting, transferring and communicating data with each other. This phenomenon is also called "internet of things" (IoT). Monino & Sedkaoui 2016, introduction.)

The term Big data refers to generating massive amount of data and making it accessible.

In order to combine and process these massive amounts of data, new processing tools are necessary. (Monino & Sedkaoui 2016, chapter 1.)

3.5 Analytics

To be data-driven organization, there needs to be analytics in use. In the upcoming subchapters modern analytics along with techniques and strategic analytics are explained.

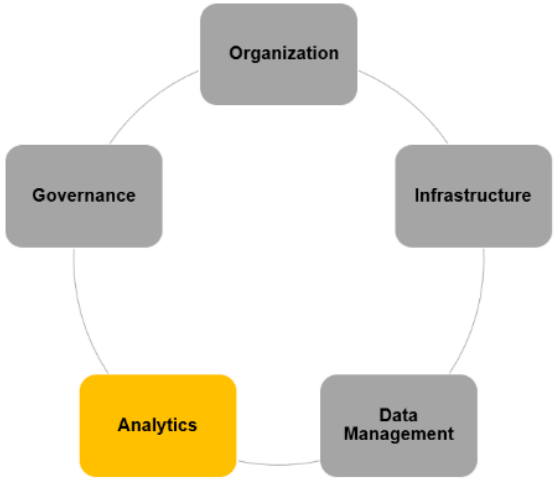


Figure 14. Analytics category for Analytics assessment (adapted from Halper & Stodde 2014, 5)

3.5.1 Modern Analytics

Nelson (2018, 6) defines analytics as *"a comprehensive, data-driven strategy for problem solving"*. Chambers & Dinsmore (2015) point out that one of the principles of Modern Analytics is to concentrate on the projects that are the most potential in bringing value to the organization.

Combination of data, logic, inductive and deductive reasoning, critical thinking and quantative methods make analytics. These result in examining a phenomena and determining it's essential features. Analytics can be used to solve operational or strategic problems and to help understand, predict and optimize issues. Analytics can utilize any type of data from simple, complex, structured, unstructured, quantative to qualitative data. (Nelson 2018, 8-9.)

Identifying potential analytic opportunities can be done by thinking outside the box. Organizations

may have faced entrenchment problems within the industry or company, that have been around for a long time. The organization might think these barriers are everlasting but by looking at them in a new light might help them discover these barriers no longer exist. (Chambers 2015, chapter 1.)

Data analytics aren't just for corporates or large companies. Modern computing technology has made data analytics available for organizations of all sizes. In fact, larger companies are generally not as flexible and data culture might not be as strong as at fellow start-ups or new entrants. (Saulles, 2019.)

Advanced analytics incorporate a range of highly developed and often automated or semi-automated techniques including data mining, machine learning, forecasting, predictive modelling, simulations and text mining (Nelson 2018, 8).

3.5.2 Analytic techniques

Data analytics can help you explain what has already happened but more importantly it can help you explore and discover what may be possible for the future. Analytics can aid with complicated decision-making, discover insights and highlight opportunities that are often overlooked. (Chambers & Dinsmore 2015, chapter 4.)

There are different analytic techniques that can be used for different types of purposes. Four analytic techniques that can help solve different types of problems are listed below:

- *Descriptive analytics* describes what has happened in the past and usually answers to questions like what and why did something happen.
- *Predictive analytics* use past data to create connections and insights and use them to make predictions for future. Predictive analytics often answer questions such as what, when and why will something most likely happen.
- *Simulation* simulates models and likelihoods of various outcomes. Simulations often answer questions like what would happen if something was changed or what else could happen.
- *Prescriptive* analytics evaluate potential scenarios to find out the best possible outcome and to achieve given goals. Prescriptive analytics often answer questions like what would be the best option from given choices, how could you increase or better certain actions, and who is most likely to do something. (Chambers & Dinsmore 2015, chapter 5.)

3.5.3 Strategic Analytics

Chambers & Dinsmore (2015, chapter 6) categorizes enterprise analytics into five main categories as per the image (Figure 15) below.

Figure 15. Enterprise Analytics categories adapted from Chambers & Dinsmore (2015, chapter 6)

The strategic analytics was chosen as an example category of the enterprise analytics as it can be especially useful for the executive board of the client organization.

Strategic Analytics, as the name suggests addresses strategic questions and issues. It also serves the organization's top management and decision-makers by providing support for the decision-making. Strategic issues usually have the following four characteristics:

- high stakes
- they fall outside existing policy
- nonrepeatable
- no clear consensus

In addition to the listed characteristics, strategic issues are often urgent and important, so they need a speedy resolution. (Chambers & Dinsmore 2015, chapter 6.)

3.6 Governance

Quality data can be useless without governance. In short, data governance means the correct management of information assets by using authority combined with policy (Ladley, 2012). The importance of data governance is explained in this sub-chapter in more detail.

Figure 16. Governance category for Analytics assessment (adapted from Halper & Stodde 2014, 5)

It is recommended that the organization nominates someone who is taking charge of the data governance. Some companies have nominated data councils that consist of multiple people from across the organization. Those in charge of data governance should create general rules on who has access to the data, how and with whom the data can be shared, and how the data can be used organization wide. The first task is to create a common language around data that ensures not just clear communication but also organizational cohesion. (Kiron 2017, 5-6.)

Kiron (2017) also identifies that “The work of data councils involves mapping workflows, assigning ownership and accountability for key deliverables, defining and enforcing the cultural norms for the use of data in decision-making, and measuring what data and analytics contribute to organizational performance” (p.5).

It is important to differentiate the governance of information with the management of information. Those who manage information, often the executives of the company, should not perform the governance and vice versa. (Ladley 2012, chapter 2.)

Ladley (2012, chapter 2) has introduced a concept (Figure 17) called Governance V that is presented below.

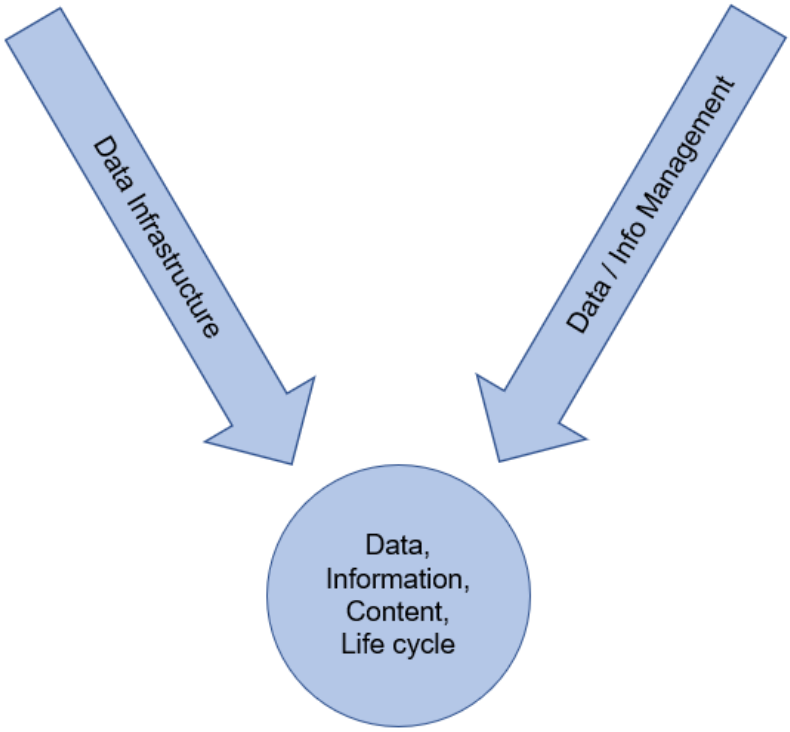


Figure 17. Governance V (adapted from Ladley 2012, chapter 2)

On the right side of the figure are the decision makers of the organization, the ones that manage information within the organization and make sure that the goals are achieved. On the left side of the figure are the people in charge of data governance that are making sure that the data is doing what it's supposed to and ensuring that the rules and policies are in place. On the bottom of the figure are the activities which maintain content, information and it's life cycle. (Ladley 2012, chapter 2.)

3.7 Decision-making

Concentrating purely on data and analytics is not enough as the decision-making method and style of the organization needs to be considered too. How the organization is making decisions and what they base their decisions on contributes to how well they are able to adapt to data-driven decision-making.

The complexity of decision-making has increased and according to a survey made by Gartner in 2021, over half of the respondents, report that decisions are more complex compared to 2 years prior and there are higher expectations for them being able to justify their decisions (White & Rollings 2021, 2).

It is also good to keep in mind that not every need is met by use of the same set of methods or tools. The main question that should drive the analytic decision-making is what the organization is going to do with the results. (Chambers & Dinsmore 2015, chapter 6.) If the analytical findings are ignored, unread or never acted upon or taken seriously the organization cannot be called data-driven. To be data-driven the analyses need to have a direct impact on the business and decisions. (Anderson 2015, preface.)

3.7.1 Decision-making process

Organization's decision-making generally follows an internal process. There are multiple formal processes how decision-making can happen within the organization.

A process map (Figure 18) on the next page presents a five-step decision-making process.

Figure 18. Five-step decision-making process (adapted from Eby 2018)

First it is necessary to identify the end goal and what is wanted from the decision. The next step is to gather all relevant information and then evaluate any risks or consequences that could follow the decision. Fourth step is to make and execute the decision. Finally comes the evaluation of the outcome which determines how successful the decision has been. (Eby 2018.)

3.7.2 Re-engineering decision-making

To make decisions future-proof and optimal they must be connected, contextual and continuous. This means there should be a symbiotic relationship between humans and machines. (White & Rollings 2021, 1.)

White & Rollings (2021, 11) have created an iterative process map (Figure 19) below to help design re-engineering.

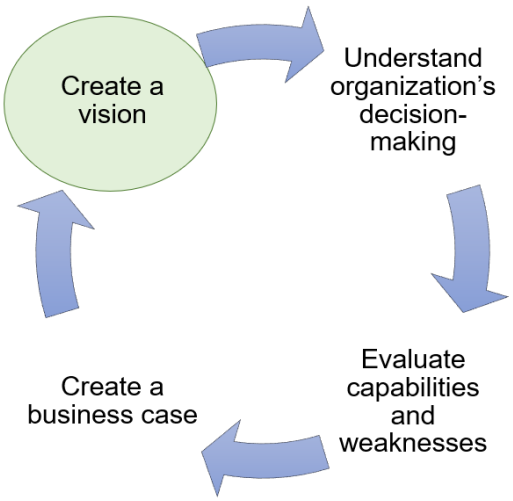


Figure 19. Re-engineering decision-making (adapted from White 2021, 11)

It all begins by creating a vision from where it is considered how re-engineering the decision-making process helps and in which direction does it take the organization. Once the vision is clear, the second step is to understand the current decision-making process. Where are the decision made and by who and which are the components, rules and habits of decision-making within the organization. It is good to evaluate which decisions are unmanageable or short of insight. When coming to a final decision, it is important to assess the capabilities and weaknesses of each component. This is to create an understanding for example if skills, technological capabilities, augmentation or more data is needed. (White 2021, 11.)

4 Methodology

In this chapter I present and justify the chosen research approaches and describe the data collection plan and methods. I also share why the demonstration and recommendations were chosen to be part of this study. The chapter ends with me assessing the quality of the research.

4.1 Qualitative research

I chose to do this thesis using qualitative research as I find that the empirical approach is the most suitable as it is close to practise. Data collection methods such as questionnaire, semi-structured interview, workshops, document analysis and feedback survey were used to help me gather as much information as possible to create a 360 view of the topic.

Flick (2007, 2) defines that instead of finding statistical dependencies or looking at numbers, the qualitative research is interested in participants' perspectives, everyday practices, and knowledge of the issue at hand. Flick (2007,6) continues that often the qualitative research's aim is to generate relevant practical information about the issue of the study rather than just for a scientific purpose.

Qualitative research aims to explain a phenomenon at hand and understand behaviour related to that phenomenon rather than creating statistical generalisations. Therefore, it is important to choose subject matter experts that hold as much information about the matter as possible. (Tuomi & Sarajärvi 2018, chapter 3.)

The practicality and understanding the client organization's behaviour in the data-driven context are some of the key reasons for me choosing qualitative research for this thesis. The client organization's executive board members are all subject matter experts in their domain and fit the target group of qualitative research well.

4.2 Case study

I selected a case study research approach for this study as it supports investigating a real-life phenomenon in-context and therefore can provide the best support in finding answers to the topic of this thesis. Case study research can give an understanding or insights into a chosen research topic, which is usually a contemporary phenomenon. In business research, the evidence collection happens in its actual environment and real-life circumstances. Case study is an empirical in-

context investigation that can use known research methods for data collection like interviews and surveys. (Farquhar 2012, chapter 1.)

According to Tight (2017, 9) case study includes features that are presented in the chart (Figure 20) below:

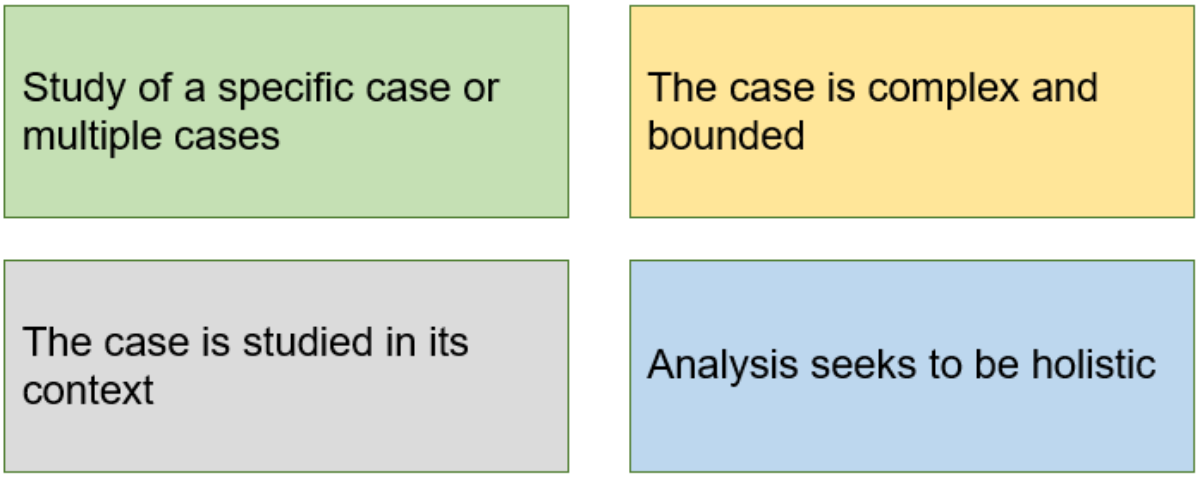


Figure 20. Features of a case study (adapted from Tight 2017, 9)

The above features can be found in this research work. It is a study of a specific case inspired by a real-life phenomenon within the client organization. The case is studied in its context and the nature of the case is complex and bounded. The analysis of the case seeks to be holistic rather than particularistic.

Either qualitative or quantative data can be used for case study and data should be collected using different techniques and data sources (Yin 2009, 261.)

In the next sub-chapter I present the qualitative data collection methods and analysis for this case study.

4.3 Data collection and analysis

Below is a flowchart (Figure 21) that represents the data collection plan for this case study.

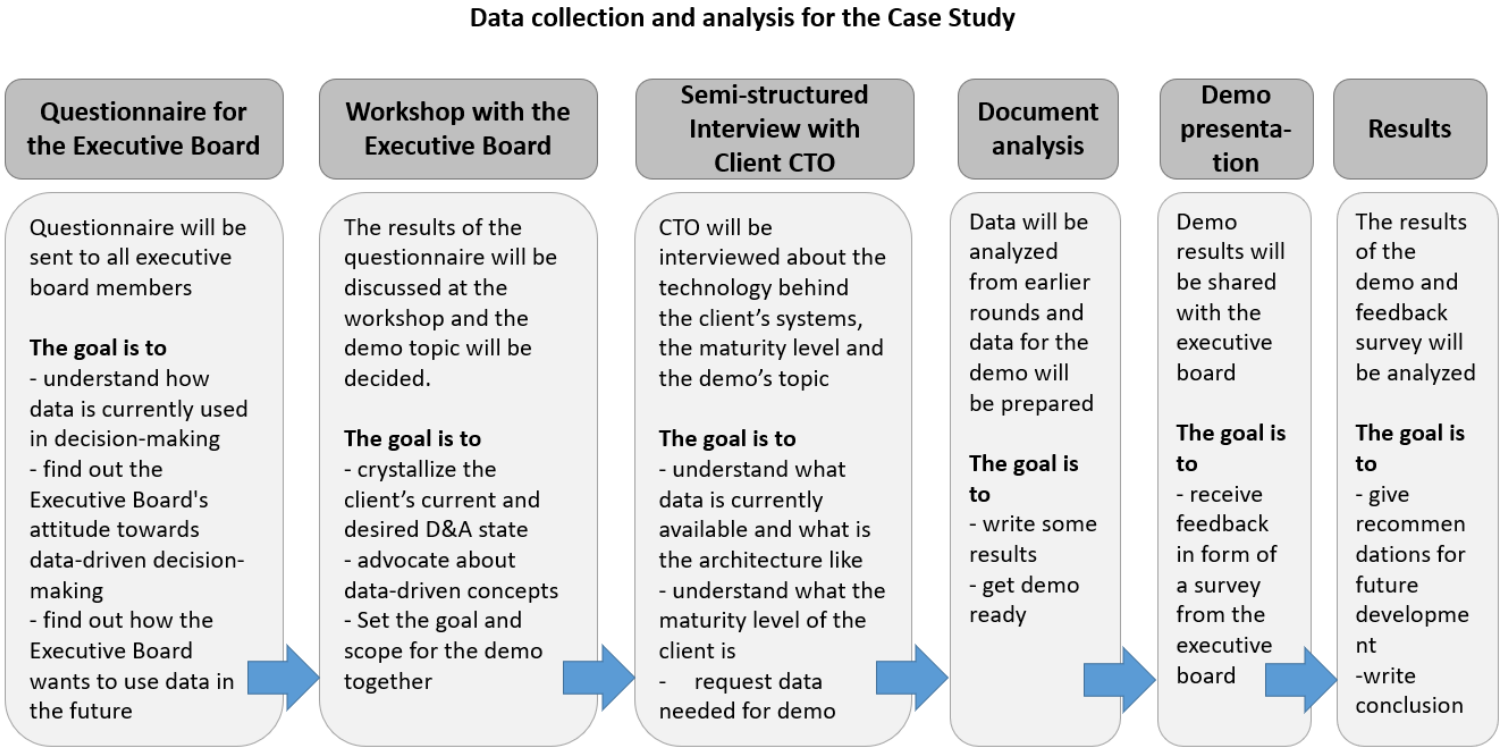


Figure 21. Data collection plan

As a first step, I created a questionnaire with relevant questions recording the organization's current and desired data analytics state and measuring the organization's leaders' attitude towards data-driven decision-making. The questions in the questionnaire can be found in appendix 1. The questionnaire was sent to each member of the client's executive board as a Webropol-link. That was in total of 5 people. The Executive board was chosen as the target group as it represents the organization's decision-makers and includes subject matter experts from every department. There were 5 members at the time and this number includes me as I am part of the executive board. The questionnaire was responded by everyone by early January 2022.

The second step was to create a deeper understanding of the matter. Therefore, a workshop was chosen as it allowed me to further interview and observe the members. Tuomi & Sarajärvi (2018, chapter 3) state that the benefits of an interview are the overall flexibility, a chance to add and make follow-up questions and observe not only what respondents say but how they say it. I found this true while I was facilitating the workshop.

The workshop's intention was to refine the topic for the demo, as well as go through the summary of the answers. The meeting was held late January 2022 online via Teams and had 7 participants. This included one new member of the Executive Board and the CTO of the organization. CTO was included in this stage already as it was important for me to know, if the data that was requested, was accessible. The agenda for the workshop can be found in appendix 2.

Berndtsson, Ericsson & Svahn (2020) name bad quality data or no access to data as key barriers during a data-driven pilot. Organizations tend to think they have either better access to data or better quality data than they actually do (Berndtsson et al. 2020). To prevent this barrier, as a third step, I held a semi-structured interview with the client organization's CTO, which allowed me to gain a deeper understanding of the quality and availability of the data. We also discussed the topics of the back-end architecture and the data maturity. The interview took place via Teams in February 2022 and the themes of the interview can be found in appendix 3.

In step four, once the files I needed for the demo were received, I carried out an analysis and started preparing the demo. At this stage, I found that the data extracted didn't include certain columns that would have helped providing insights and answer the questions that were presented.

In step five, I presented the demonstration to the executive board with findings and thoughts. This happened virtually in mid-April 2022. After the demo was presented, feedback was collected from each 6 members of the executive board in form of an online feedback survey. The survey can be found in appendix 4. The survey only received 5 responses and was missing one.

Finally, once all the data was acquired, I reflected on the data collected and prepared recommendations for the future development.

4.4 Demonstration and recommendations

I chose a demonstration and recommendations for future to be part of this study to produce something practical for the client organization. The biggest reason for the demo in this thesis was to bring theory to practise. The demo worked as a showcase of how data can be used and therefore provided valuable content for research question 2.

The demonstration in this thesis was built using a pilot approach. A pilot is often seen as a way to test and evaluate new technology and remove uncertainties before scaling it up in the organization (Berndtsson et al. 2020). White (2019) recommends selecting a business area that needs

improvement but also demonstrates an openness to change.
To make organizations realise the business value, it is better to deliver results quickly rather than waiting and building a perfect analytic model. The value can be shown by creating a proof of concept that is limited in scope but can quickly mature into a production deployment. (Chambers & Dinsmore 2015, chapter 1.)

To understand the expectations beforehand is important as sometimes just the knowledge of using analytics and its known potential can give false expectations to the senior management. For the same reason, it must be stated when the pilot is supposed to be completed and what defines the pilot's success or failure. (Berndtsson 2020.)

Recommendations for future action were created to support the client organization in the journey on becoming data-driven. These recommendations stem from matching the theoretical framework with the client organization's current and desired data and analytic state.

4.5 Quality of the research

To demonstrate the quality of my research, I will assess the construct validity, internal validity and reliability of my thesis in the next subchapters.

4.5.1 Construct validity

Construct validity measures how well the research studies what it claims to study. Construct validity is often rejected by interpretivist researchers as it assumes an objective reality. Construct validity can be raised by using multiple data sources to minimize bias and creating a clear chain of evidence showing how the research is being constructed. (Farquhar 2012, chapter 7.)

In my research I have used multiple data sources and data collection methods. I have also showcased the structure of the study and how the data collection proceeds from one step to another. The concluding part of the thesis provides answers to the research questions, and in my opinion meet the objectives and goals that were set for the whole study.

4.5.2 Internal validity

Internal validity looks at the causal relationships between results and variables and aims to persuade the reader that the findings of the research are the cause of critical investigation of the data.

Critical investigation can be proved by providing details of how the research data was analysed. (Farquhar 2012, chapter 7.) Eisenhardt (1989 in Farquhar 2012, chapter 7) argues that the internal validity can be proved by closely examining the theory of the research case study with existing literature.

I have built my research in a way that I first created the theoretical framework based on existing literature and then examined the case study with the existing information. There was no conflicting information when examining the theory of the case study with existing literature and this increased the internal validity of the research. I also analyzed the questionnaire results, the contents of the workshop and the interview, and feedback survey thoroughly.

4.5.3 Reliability

Reliability means that if the research was repeated, the researchers would derive same insights. This means there are no random errors and that the evidence is consistent and stable. Transparency and replication increase the reliability. (Farquhar 2012, chapter 7.)

I facilitated the workshop and made sure the goal and the scope for the demo were set together. The workshop was held online and recorded and the memo and all the answers to the questionnaire and feedback survey are kept in record to increase the reliability of the research and that they can be further investigated or repeated if needed.

Also the semi-structured interview was held online and the interview was recorded for the reliability purposes. I tried to keep the research as transparent as possible.

5 Results

This chapter goes through findings and insights from the data collection. The primary data sources were the questionnaire, the workshop, the demonstration, and the demonstration's feedback survey. The maturity, infrastructure and data management sub-chapters also have an input from the semi-structured interview with the organization's CTO. The questionnaire can be found in Appendix 1, the workshop structure in Appendix 2, the interview questions in Appendix 3, and the demo's feedback survey in Appendix 4.

The demo's topic, goal, results and feedback are introduced first as they provide input for the sub-chapters after. The results chapter's other sub-chapters are mostly providing answers to research question 4 and its sub-questions. The sub-chapters are constructed with a similar structure than the theoretical part of this thesis with the following exceptions: The data strategy part was left out as there is currently no data strategy at the client organization. The data literacy and the data culture were combined into one sub-chapter as they go hand in hand. The analytics was incorporated in other sub-chapters and the data governance section was left out as there are some comments about it in the infrastructure section.

In the end of this chapter, recommendations are provided for future development for the client organization based on the knowledge acquired while writing the current state analysis and the theoretical part of this thesis, and by conducting the demo.

5.1 The demonstration

The researcher and the executive board went through the demo's scope, timeframe and expectations together during the workshop.

In question 9 in the questionnaire, the executive board picked top five areas that were seen as the most desired areas to receive support from data. Those five areas were chosen from the questionnaire's 20 broad options plus an opportunity to suggest your own topic. The following five domains were selected:

- Human resources and working time monitoring
- Review of business profitability
- Reorganization of work or working methods
- Strategic decision-making
- Information from client organization's services (event attendees, message automation readings, accreditation utilization rate, sub-services information)

During the workshop, the executive board combined two of the priority domains (working time monitoring and business profitability) into one domain, which was seen to bring the most value to the organization within the scope of the pilot. As discussed in chapter 4.4 it is important to concentrate on business areas that need improvement and demonstrates an openness to change.

The main topic chosen was to look at expenses versus income at three different service levels. Comparison was done between events pre-covid19 in 2019 and in 2022 when the pandemic was still on but most of the restrictions were loosened. Insights were also provided about the client organization's services by comparing different stages of the process, number of changes and number of people involved. The timeframe for the demo was from February to April 2022.

5.1.1 Goals

The goals of the demonstration were introduced to the Executive Board during the workshop and were agreed on together.

The first goal was to wake up the executive board's interest towards a thriving data culture and data-driven approaches. Another goal was to articulate the benefits and showcase measurable business outcomes of data-driven decision-making by demonstrating it to the executive board. The last goal was that by working to create a demo, the researcher gets valuable information about the client's current and desired data and analytics states and can therefore give concrete recommendations for steps and actions to take in the future. All the above goals coincided with the overall goals of the thesis.

5.1.2 Presentation

The completed demo with different graphs and insights were virtually presented to the executive board in April 2022. The demo showcased four graphs with different potential decisions that could be derived from them. This was to showcase how different data can support the decision-making and therefore meet one of the objectives of this study. Actual decisions based on these graphs have not been made yet.

The graphs were as follows:
1st Graph showed the amount of time it takes to create an event in different categories. It also showed the number of employees being involved in the process. Looking at these numbers,

decisions could be made regarding allocating resources and streamlining the internal processes.

2nd Graph presented the total revenue versus the total expenses in the example categories.
This graph answered the main question that was set for the demo and decisions could be made regarding categorizing different clients and investigating revenues and expenses in detail.

3rd Graph included the client organization's total share in the example categories.
This graph displayed the client organization's total share and decisions could be made regarding the pricing and altering the business model.

4th Graph showcased the time (in hours) between creation and publishing (an internal process of the client organization) in the example events.
By investigating these times, decisions could be made regarding streamlining internal processes and service delivery.

The presentation provided the insights made by the researcher from the results and also comments on what affected the results. These comments were that there should have been an additional timestamp on the events for calculating the exact times. Also, there should have been more events per category for reliable comparison.

5.1.3 Feedback

After the presentation, members of the board were asked to answer an online feedback survey to help measure the outcomes of the demo. Answers were received from 5 members. The feedback questions can be found in appendix 4.

The survey's questions were designed to measure how the demo's goals were met. The first goal was to increase the executive board's interest towards a thriving data-driven culture. In question 2, when asked about the respondent's interest level towards data-driven decision-making compared to four months prior, 4 out of 5 respondents replied saying their interest level had increased by agreeing or strongly agreeing to the statement. One respondent didn't agree or disagree.

100 % of the respondents also either agreed or strongly agreed to the statement "I would like to know more about data-driven approaches". All five respondents felt like they had received valuable information about data-driven approaches in the process of participating in this thesis and they all also agreed that they received valuable information about the quality of the data. 100% of the

respondents agreed or strongly agreed that they would like the executive board to make more decisions based on data. This shows the desire to be more data literate which in turn creates more knowledge and enables healthier data culture.

Based on these answers, it can be concluded that the knowledge and interest have increased among the decision-makers while participating in this demo and the desire to use data-driven approaches is strong.

Another goal was to portray the benefits and showcase measurable business outcomes of data-driven decision-making by demonstrating it to the executive board. When asked how well the demonstration provided answers to given questions, 40% of the respondents felt like it didn't by disagreeing to the statement. Another 40% agreed. 20% of the respondents didn't agree or disagree.

It became evident while building the demo that the data used for the demonstration wasn't sufficient and the questions were a bit too complicated for this particular demo. Also, the number of events wasn't sufficient for reliable comparison and this feedback was also received from the executive board. However, when asked about whether the demonstration provided an idea of how analytics can be used or built, 80 % agreed or strongly agreed that it did. Therefore, even though the demonstration didn't provide sufficient answers to the questions that were presented, the knowledge and understanding of how data can be utilized increased among the target group.

The last goal was that by working to create a demo, the researcher gains valuable information about the client's current and desired data and analytics state and can give concrete recommendations for actions to take in the future. The researcher found a lot of information about the client organization's data-related approaches and could therefore provide recommendations for the organization to take to become more data-driven. This was the researcher's opinion but when asked from the board members, if those recommendations derived from the demo were sufficient, 80% of the respondents agreed.

Based on the feedback and comparing it to the goals that were set for the demo, I would conclude that the first and the third goal were successfully met while the second goal was only half met.

5.2 Client organization's maturity

When analyzing the overall results of the questionnaire, the workshop, and the semi-structured interview and comparing it to the TDWI Maturity model presented in chapter 3.1.2, I would position the client organization currently in level 2 and reaching level 3 within the next two to six months. Levels 2 and 3 are presented in the chart (Figure 22) below. This evaluation was backed up by the organization's CTO during the interview.

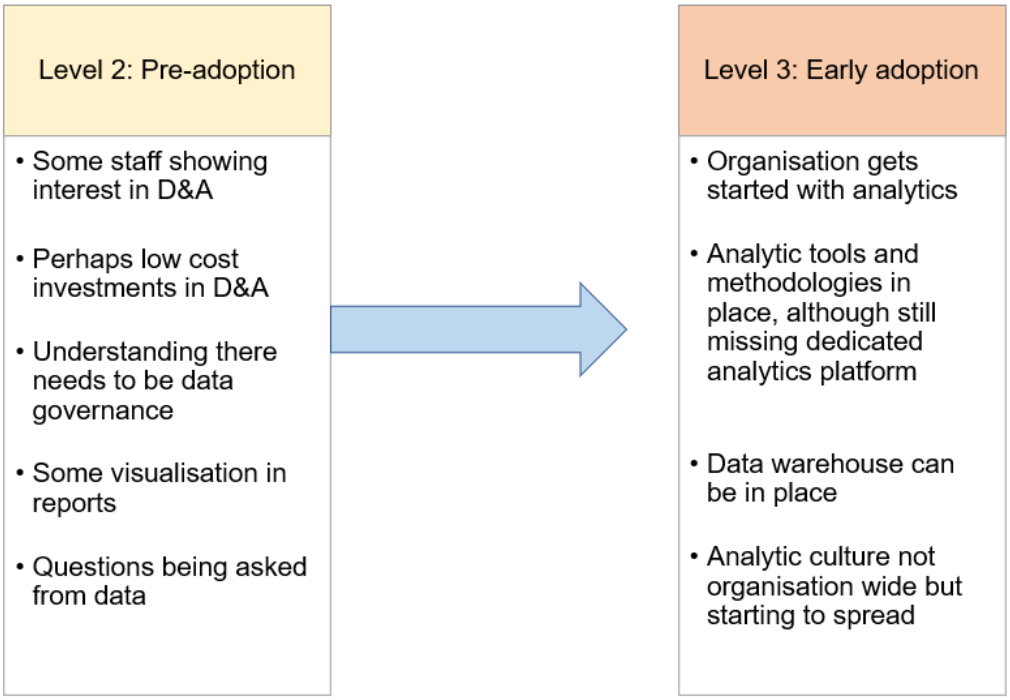


Figure 22. Level 2 and 3 of TDWI Maturity model (adapted from Halper & Stodde 2014, 9-17)

The client organization's architecture is currently being modified to better enable data and analytics. The staff is showing interest and the executive board has been advocated about the topic. Marketing department has started asking more specific questions from the data and are requesting more access to data. Also, the workshop and demo included in the thesis, have showcased how to ask questions from data. The participation of the decision-makers in the pilot was active which is promising for both data culture and the maturity.

Data warehouse in the form of Big Query has been installed however, there is no data strategy being discussed yet. The importance of data governance has been realised but the process is still ongoing. Transition to level 3 is possible because of the current changes in the back-end system and the growing interest towards D&A within the staff. However, to reach level 3 on the maturity chart,

the data literacy needs to increase and spread as well as data governance and methodologies need to be discussed more in detail. It also helps to get some experience from using Biq Query and Google Data Studio.

Question number 4 focused on measuring the current use of analytic tool at the client organization. The client's sales system produces some reports but only Microsoft Excel is being used to handle imported data and make customised reports. Some employees are learning to use Google Data Studio. Perhaps another tool such as modern analytics software, should be taken into consideration especially at later maturity levels.

5.3 Client organization's data culture and literacy

The Executive Board's overall attitude towards data-driven decision-making is promising. Each member either strongly disagreed or disagreed to the statement *"I am happy with the current use of data analytics in the work of the executive board".* This shows that there is a need and want for change when it comes to using data.

When asked about the potential benefits of Data and Analytics in question number 11, the respondents see them like presented in the chart (Figure 23) below.

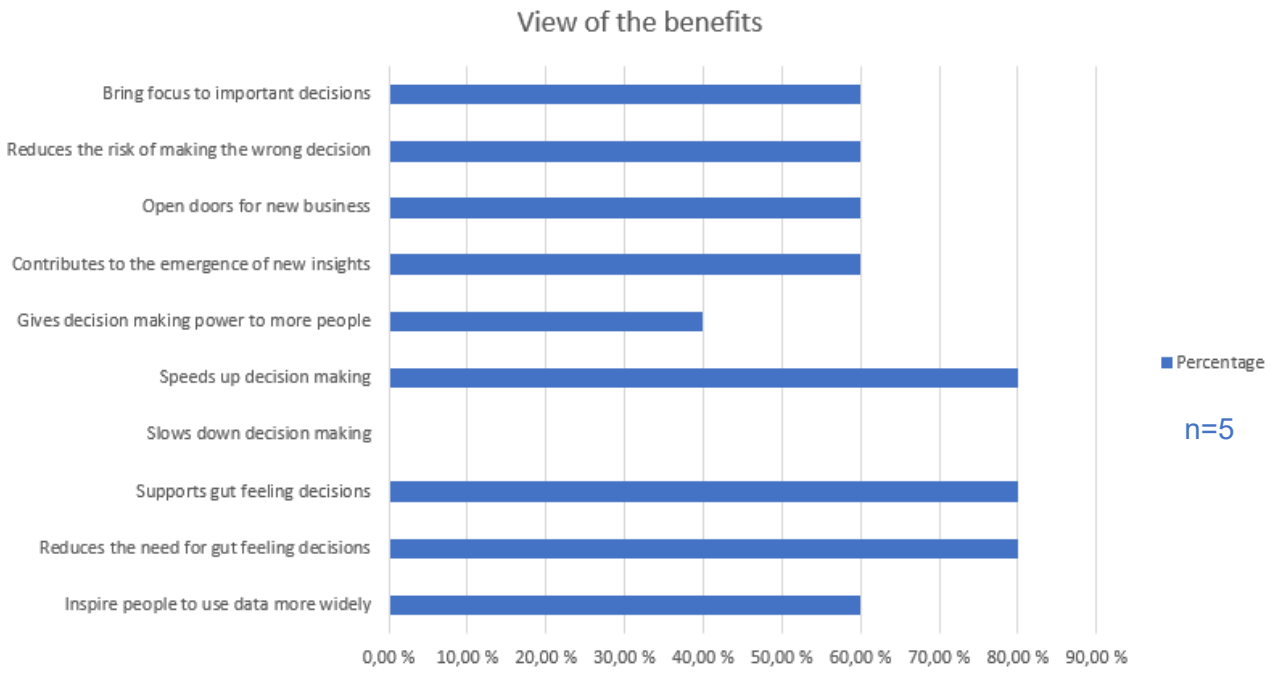


Figure 23. Respondents' views of potential benefits

The biggest benefits were seen as speeding up the decision-making and having an impact on gut feeling decisions. Both “supporting gut feeling decisions” and “reducing the need for gut feeling decisions” received the most votes. The answers come to show that data-driven decision-making can be seen to support decision-making but also that the decision-making needs remodelling.

The questionnaire also measured the board’s thoughts about learning in questions 15 and 16. 80% of the Executive Board members would like to know more about data-driven decision-making at executive level. It was seen as an important part of the decision-making process and as a way to support the culture change within the organization. Also 80 % of the respondents saw that it would be important for either the whole company or a specified group to learn more about data analytics. The executive board sees the time span for any changes from immediate to within a year.

All of the above results show the organization’s decision-makers significant interest in heading towards data-driven decision-making. Wanting to learn more is crucial in creating and promoting data culture and increasing the data literacy within the organization. This is also one of the ways to promote the culture within the organization as per research question 2.

5.4 Client organization’s infrastructure

There were altogether 30 sources listed for question 3 that asked where the data for the decisions currently come from. There were an additional four sources listed during the workshop. This is to show that there is not one centralised place to get information from.

According to the organization’s CTO some changes are currently being done in the back-end architecture. This is for overall optimised performance and to decrease costs as well as allowing better data and analytics structure. This in turn will enable better access and better utilisation of data for more people. It will also free up the developers’ time as certain reports don’t need to be generated by the developers like they currently are.

The next step is to create queries and decide which data needs to be moved to Big Query as it’s not necessary or cost efficient to move all data. The sorting and deciding still needs to be done and it requires co-operation between departments. Currently the order data is being transferred as it provides valuable information for the marketing department.

Big Query was chosen as the data warehouse as it is cheaper than the Amazon’s equivalent Redshift. Big Query is also elastic which Redshift is not currently and being elastic brings benefits such

as cost efficiency as you only pay for what you use. Google Data studio was chosen as the software because the Marketing department already uses it for the website metrics and tracking campaigns.

The executive board didn't prefer any analytic platform over other when asked in question 14 nor it was brought up in the workshop. Over half of the respondents felt like the software didn't matter, as long as it is easy to use and there will be training provided.

When asked about the preferred qualities of the future analytics software in question 13, every respondent chose the option "simple to use". Other attributes that gained the majority's opinion were that the analysis must be understood, it can be exported to Excel, it can be built by oneself, and that the analysis is real-time. Ease of use also helps the organizations gain higher level of maturity as discussed in chapter 3.3.2.

The organization's marketing department has recently started learning the use of Google Data studio. Data Studio was used for the demo also as it is better to have only one platform at this stage.

5.5 Client organization's data management

The data used for the demo was extracted from the system by the CTO and provided to the researcher in the CSV-format. The client's back-end architecture is currently in the process of changing so that other users can generate queries to the database also. That will make the company data more accessible, and free up the developers' time for other tasks. This desire was also brought up during the workshop.

The data rows needed for this demo were accurate however, to serve the demo's purpose, some additional time stamps and history rows would have increased the completeness and therefore the quality of the data. The data was relevant to the issue, reliable and consistent with other reports extracted from the system. Sales reports could be generated by the researcher and the data was reliable, complete and timely however, not having the option to extract certain reports as csv-format caused a bigger workload in the data preparation phase.

5.6 Client organization's decision-making

The questionnaire investigated the current decision-making method of the Executive Board by asking questions about how decisions are made, currently based on each members' opinion.

There was discrepancy in answers when asked about what the current decision-making is based on. The six following options were given; Gut feeling, Recommendations of others, Experience, Data and hard facts, Specialist consultation or Something else.

It turned out the most decisions currently were made based on experience and the least was professional consultation however, in the workshop it was agreed that the professional consultation option in the questionnaire was seen as consultation from outside the organization rather than subject matter expert's consultation within the organization. The discrepancy in the answers suggests that the decision-making is not unified or follow a certain style.

There were also conflicting answers when asked in which business areas does the executive board use reliable data to support decision-making currently. 15 out of 18 business options presented in question 2 received at least one vote but no option received more than three votes. This proves to show that the respondents see differently, and most likely objectively, which decisions are based on data, and which are not. Also the executive board consists of leaders from different departments that follow different data.

When asked about how often the members use reliable data for decision-making, 60% of respondents say daily but 40% of respondents couldn't evaluate the frequency. The desired frequency was asked, and each respondent agreed that it should be daily in the future.

When analysing the answers to question number 5 in the questionnaire, there was a strong desire to use data and analytics in the future with every respondent wanting to use data to back up decision-making either more or significantly more compared to now.

5.7 Recommendations for the client organization for future

Based on what I learned about data-driven organizations from writing the theoretical framework for this thesis, conducting the demo and gaining understanding of the client organization's current and desired data maturity and other data-driven attributes, I am making the following practical recommendations for the client organization. On top of addressing the first objective of the thesis, this also helps in promoting the data culture within the client organization, which is the second objective of this thesis.

Create data & analytics strategy and roadmap

As the client organization has no data or analytics strategy or roadmap written out, I recommend this as one of the first things to do. This will give the organization guidance and help with establishing how, in which scope and timeline to move on with data analytics. This will also ensure the development goes ahead as it is written out. Chapter 3.2.3 showcases an iterative analytics strategy development method that could be used for developing a data strategy.

Create a healthy data culture

One of the objectives was to promote data-driven culture in the client organization and while I think this happened, to keep it consistent and continuing requires a lot of work. The executive board members should be advocates for data-driven practises and promote the data culture within the organization. To create an inquisitive environment among the workforce, where data is seen as a core asset, I recommend the leaders to show example and encourage the staff to ask questions and utilize the insights in their work. It is also important to celebrate the victories and gained benefits that show the value of data analytics. A good way to bring data and analytics to be a part of everyday working life is to communicate and speak about data in a positive way and show example whenever possible.

Train and learn

It is not possible to expect anyone to utilize data or create insights if they are not trained to do so. This obviously goes to both leaders and staff. The desire to learn more and become more data literate is high according to the questionnaire and feedback survey. I would recommend training the whole staff with basic data management concepts and more specific training for the decision-makers and leaders first. Good ways to achieve understanding of the topic is to order a tailored training program from a specialist and hold a workshop to innovate and discuss different ways to use data within the client organization. Training and raising analytic skillsets of the employees can keep them engaged and motivated and that will benefit the whole organization.

Nominate someone in charge

I recommend the organization to nominate a data officer or someone in charge of pushing the data and analytics development further. This will make sure that D&A won’t be dismissed but incorporated in the decision-making and the organization stays on top of their data strategy. Some organizations have even created data councils that consist of leader from different departments.

Consider using Modern Analytics platform

To reach higher levels in the data maturity, I would recommend considering the use of modern analytics platforms like Microsoft Power BI or Tableau that incorporates the use of machine learning

and other advanced technologies that contributes to predictive and descriptive analytics. These can offer future modelling and give suggestions for further action. Another thing to look into, when reaching a higher maturity level, is the use of big data and the advantages and insights it can provide for the business development.

Remodel decision-making processes

As per the decision-making chapter 3.7 and the results from the questionnaire presented in chapter 5.6, it would be beneficial to remodel certain decision-making processes at the client organization. When the organization is used to making gut feeling decisions, any change won't be expected to happen overnight. An iterative process map presented in figure 19 in chapter 3.7.2 can aid with the decision remodelling. It is important to primarily focus on decision-making domains that benefit the organization the most.

6 Conclusion

This chapter presents answers to the research questions that were presented in the beginning of the study. It will also assess how the objectives were met. The thesis concludes with discussion.

The research questions were to support the objectives of the thesis. They were presented in the beginning of the paper and are as follows:

RQ1. What is meant by data-driven organization?
Data-driven organization means that the organization has tools, abilities and culture that drives on data. A data-driven organization trusts data and uses it to inform and influence strategy. There is an evidence-based culture, and the data analysis is highly informative. Data-driven organization treats and views data as a core asset and see it as a competitive advantage that improves their ability to innovate. The leaders support systematic use of data in strategy and decision-making.

Sub-question 1: What dimensions does data-driven organization have?
There are five different dimensions according to TDWI Analytics maturity model. These include organization, infrastructure, data management, analytics and governance. The dimensions can be found in figure 2 in chapter 3.1.1. These five dimensions create a picture of a data-driven organization and these can be used to assess the overall analytics maturity within the organization.

Sub-question 2: What are the benefits and challenges of data-driven organization?
There are multiple benefits of being data-driven. It contributes to better and faster decision-making and provides insights for a wider audience. Being data-driven creates innovation which enhances or creates new business. Analytics can help explain what happened in the past but the most benefit is often gained from predicting the future. Overall being data-driven can help streamline processes by providing insights and grow top-line revenue cost efficiently.

The challenges often have to do with the data-driven culture. Trusting data is often harder for leaders that are used to making decisions based on their gut feeling. Also, organizations often have problems with data quality and governance. Often, organizations have the wrong idea about the quality and type of data they have. The executives or leaders might have unrealistic expectations when it comes to the results or scope of analytics projects simply because of the rumours and hype around what modern analytics can do. Sometimes the main obstacle isn’t the actual data management or its complexity but rather lacking the skill of turning data insights into business actions.

RQ2. How can Data and Analytics culture be promoted within the client organization?
As data-driven culture has been named as one of the biggest contributing factors on becoming a data-driven organization, it is important to look at ways that can promote the data-driven culture within the client organization. Firstly, I think that understanding what being data-driven means and realizing the benefits and power of analytics, is the key to promoting data-driven culture. Without having the knowledge, it is unrealistic to expect results. Leaders are the ones that show an example to the rest of the organization so keeping them motivated is important. Overall celebrating wins and communicating the benefits across the organization help create a positive atmosphere and drive for wanting more. Everything starts with training and having everyone involved. Bringing data and analytic related topics to conversations and meetings is one practical example of making D&A a part of everyday working life.

RQ3. How can the benefits of data-driven decision-making be showcased to the client organization?
It is often the best way to show something in a form of a pilot before scaling up or jumping straight ahead into something. A demo can showcase how the organization reacts to a specific scenario for example. A great starting point is to pick a topic that will benefit the organization the most as when the analytics provide insights to something that is important, it has a greater effect. The organization is encouraged to ask questions from data to realise the true potential of D&A. To share gained results, it is recommended to generate visualisations that provide insights in a simple and understandable way.

RQ4. What recommendations can be given for future development?
Sub-question 1: What is the client organization's current state in data and analytics?
Sub-question 2: What is the client organization's desired state in data and analytics?
For me to give recommendations, it was crucial to understand the current and desired states, so I will first answer the sub-questions 1 and 2.

Chapters from 5.2 to 5.6 present the current and desired capabilities in different data-driven dimensions such as maturity, organization, infrastructure, data management and decision-making. All in all, I would place the client organization on level 2 soon to reach 3 on the TDWI maturity chart presented in figure 3 in chapter 3.1.2. This means that the organization is lacking certain capabilities but the desire and motivation is in place so there is high potential to reach level 3 and beyond.

Recommendations stemming from the current and desired states, guided by the theoretical

framework, are presented in chapter 5.7. The first thing in my opinion is to set a clear data strategy and plan for future action. Establishing and promoting a data-driven culture and data literacy also plays a major role. It is also a good idea to remodel certain decision-making processes and expand the analytics trialling into modern analytics tool that can provide more advanced insights with the use of machine learning and artificial intelligence.

The objectives of the thesis revolved around identifying data-driven concepts and communicating the business value of data particularly in the decision-making process and the theoretical framework was built around this.

I found that the objectives were met as a clear understanding of the data-driven concepts was achieved by the researcher and communicated to the client organization. The client organization's executives were involved in the process and the interest towards data-driven journey was elevated.

The finalised thesis paper will be sent to the executive board members and the recommendations for future development will be presented and discussed together. I can see that the development will go ahead as the client organization's measured attitude is promising and active. The demo did not fully showcase how particular questions asked from data can be answered, but it showcased how data can be utilised.

7 Discussion

Several points became apparent while writing this thesis that I will reflect on this chapter.

I find that the tight schedule and timeline for doing this thesis contributed to the coherence of the whole work. As this thesis was done for a client organization and specially targeted for leading decision-makers, I had to consider others' timetables and couldn't proceed solely at my own pace.

I created and collected answers from the questionnaire before I had fully researched the topic and written the theoretical part of my thesis. This affected forming of some questions, and I think that if I had written the questionnaire later in the process, certain questions would have been more targeted or straight-forward.

TDWI has updated their Analytics Maturity Model and Assessment for 2022 where Data management dimension has been replaced with resources. This model looks more modern, however it was too late to incorporate it in this thesis.

The client organization had chosen Google Data studio as the visualisation tool to be used. In the process of learning how to use the tool I assessed how easy it is to learn by self-teaching materials. In this aspect Google Data studio is a great option for the organization as there is plenty of material online for everyone to study and the software is easy to use.

What surfaced in the discussion with the executive board was that even if certain decisions are not smart or logical in the light of data, there can still be different type of value that is harder to measure. These things include for example having a customer as a reference or developing a service as a prerequisite to enter a bid.

References

Anderson, C.2015. Creating a Data-Driven Organization. O'Reilly Media, Inc. Retrieved from https://learning.oreilly.com/library/view/creating-a-data-driven/9781491916902/ch01.html#idp3690112. Accessed 1st February 2022.

AWS. 2021. AWS Architecture Center. Modern Data Analytics Reference Architecture on AWS. Retrieved from https://d1.awsstatic.com/architecture-diagrams/ArchitectureDiagrams/modern-data-analytics-using-lake-house-ra.pdf?did=wp_card&trk=wp_card. Accessed 4 February 2022.

Berndtsson, M., Ericsson, A., & Svahn, T. 2020. Scaling-up data-driven pilot projects. AI Magazine, 41(3), 94-102. Retrieved from https://www.proquest.com/scholarly-journals/scaling-up-data-driven-pilot-projects/docview/2451174524/se-2?accountid=27436. Accessed 23 January 2022.

Chambers, M. & Dinsmore, T.W. 2015. Advanced Analytics Methodologies: Driving Business Value with Analytics. Pearson. Retrieved from https://learning.oreilly.com/library/view/advanced-analytics-methodologies/9780133498653/title.html. Accessed 31st January 2022.

Crunchbase. 2022. Search companies. Retrieved from https://www.crunchbase.com/. Accessed 31st January 2022.

Eby, K. 2018. The Definitive Guide to Business Decision-Making. Smartsheet Inc. Retrieved from https://www.smartsheet.com/business-decision-making-process. Accessed 1st March 2022.

Farquhar, J. 2012. What is case study research? In *Case study research for business*. SAGE Publications Ltd. Retrieved from https://dx.doi.org/10.4135/9781446287910. Accessed 2 January 2022.

Flick, U. 2007. *Designing qualitative research*. SAGE Publications, Ltd. Retrieved from https://dx.doi.org/10.4135/9781849208826. Accessed 2nd January 2022.

Friedman, T & Heudecker, N. 2021. Data Hubs, Data Lakes and Data Warehouses: How They Are Different and Why They Are Better Together. Gartner. Retrieved from https://www.gartner.com/doc/reprints?id=1-24IZJZ2F&ct=201103&st=sb. Accessed 4th February 2022.

Gartner Peer Insights. 2022. Analytics and Business Intelligence Platforms Reviews and Ratings. Retrieved from https://www.gartner.com/reviews/mafrket/analytics-business-intelligence-platforms. Accessed 2 February 2022.

Goasduff, L. 2016. 3 Key Steps to a Data-Driven Business. Gartner. Retrieved from https://www.gartner.com/smarterwithgartner/3-key-steps-to-a-data-driven-business. Accessed 13 January 2022.

Google Data Studio. 2022. Data Studio Help. Retrieved from https://support.google.com/datastudio/answer/6283323?hl=en&ref_topic=6267740. Accessed 15 February 2022.

Halper, F. & Stodde, D. 2014. TDWI Analytics Maturity Model Guide. The Data Warehousing Institute. Retrieved from https://www3.microstrategy.com/getmedia/9b914607-084f-4869-ae64-e0b3f9e003de/TDWI_Analytics-Maturity-Guide_2014-2015.pdf. Accessed 1st February 2022.

Herzog, D. 2016. *Data literacy: A user's guide.* SAGE Publications, Inc. Retrieved from https://www-doi-org.ezproxy.haaga-helia.fi/10.4135/9781483399966. Accessed 19 January 2022.

Kiron, D. 2017. Lessons from becoming a data-driven organization. MIT Sloan Management Review. Retrieved from https://www.proquest.com/scholarly-journals/lessons-becoming-data-driven-organization/docview/1875399253/se-2?accountid=27436. Accessed 19 January 2022.

Ladley, J. 2012. Data Governance. Morgan Kaufmann. Retrieved from https://learning.oreilly.com/library/view/data-governance/9780124158290/. Accessed 1st February 2022.

McCall, T. Meulen, R. 2018. Gartner Survey Shows Organizations Are Slow to Advance in Data and Analytics. Gartner. Stamford. Retrieved from https://www.gartner.com/en/newsroom/press-releases/2018-02-05-gartner-survey-shows-organizations-are-slow-to-advance-in-data-and-analytics. Accessed 18 January 2022.

Monino, J-L & Sedkaoui, S. 2016. Big Data, Open Data and Data Development. Wiley-ISTE. Retrieved from https://learning.oreilly.com/library/view/big-data-open/9781848218802/.

Nelson, G. S. 2018. The Analytics Lifecycle Toolkit : A Practical Guide for an Effective Analytics Capability. John Wiley & Sons, Incorporated. Retrieved from https://www.proquest.com/legacydocview/EBC/5317892?accountid=27436. Accessed 2 February 2022.

Saulles, M. D. 2019. What exactly is a data-driven organization? CIO. Framingham. Retrieved from https://www.proquest.com/trade-journals/what-exactly-is-data-driven-organization/docview/2309566094/se-2?accountid=27436. Accessed 2 February 2022.

Tight, M. 2017. *Understanding case study research*. SAGE Publications Ltd. Retrieved from https://dx.doi.org/10.4135/9781473920118. Accessed 2nd January 2022.

Tuomi, J. & Sarajärvi, A. 2018. Laadullinen tutkimus ja sisällönanalyysi. Kustannusosakeyhtiö Tammi.

White, A. 2019. Design a Data and Analytics Strategy. Gartner. Retrieved from https://emtemp.gcom.cloud/ngw/globalassets/en/publications/documents/data-analytics-strategy-ebook.pdf. Accessed 18 January 2022.

White, A. & Rolling, M. 2021. 5 Key Actions for IT Leaders for Better Decisions. Gartner. Retrieved from https://emtemp.gcom.cloud/ngw/globalassets/en/publications/documents/5-key-actions-for-it-leaders-for-better-decisions.pdf. Accessed 26 January 2022.

Yin, R. 2009. How to do better case studies: (with illustrations from 20 exemplary case studies). In *The SAGE handbook of applied social research methods*. SAGE Publications, Inc. Retrieved from https://dx.doi.org/10.4135/9781483348858. Accessed 2 January 2022.

Appendix 1. Questionnaire

1 Contact details

2. In which of the following areas does the executive board (i.e. not just yourself) us reliable data to support decision-making currently?

You can choose as many options as needed.

- [ ] Service pricing
- [ ] Human resources and working time monitoring
- [ ] Recruitment
- [ ] Review of business profitability
- [ ] Review of operating expenses
- [ ] Budgeting
- [ ] Acquisition of new customers
- [ ] reorganization of work or working methods
- [ ] Strategic decision making
- [ ] Selection of service providers
- [ ] Office matters
- [ ] Knowledge / observation of the work Monitoring atmosphere
- [ ] Technical performance of the system (sales peaks, queues, etc.)
- [ ] Ecommerce Performance
- [ ] Information about the various services of the ticket (eg visitors to the events, effects of message automation, utilization rate of accreditation, information on streams and ticketless)
- [ ] Customer relationship management
- [ ] Measuring customer satisfaction
- [ ] Else, what _____

3. From what sources do you get data for decision making currently?

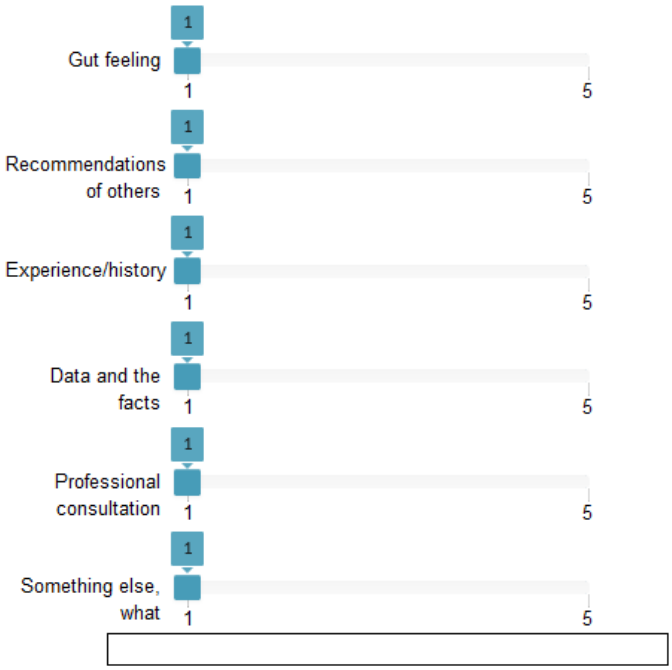
4. What analytics platforms are you currently using to view and edit data?

5. Would you like to use data analytics to support the decisions of the executive board to what extent compared to the current situation?

- ◯ Significantly less
- ◯ Slightly less
- ◯ Saman verran kuin nyt
- ◯ Slightly more
- ◯ Significantly more

6. In my opinion, the current decision-making of the executive board is based on

1 = Strongly disagree
2= Disagree
3= Neither agree nor disagree
4= Agree
5= Stronly agree



7. I am happy with the current use of data analytics in the work of the executive board

1 = Strongly disagree
2= Disagree
3= Neither agree nor disagree
4= Agree
5= Stronly agree

8. How often do you use reliable data for decision-making currently?

- [ ] Never
- [ ] Daily
- [ ] Once a week
- [ ] Less than once a week
- [ ] I can not estimate

9. For which of the following things would you like the executive board (not just yourself) to get the most support from the data?

Choose top 5.

- [ ] Service pricing
- [ ] Human resources and working time monitoring
- [ ] Recruitment
- [ ] To review the profitability of the business
- [ ] To review business expenses
- [ ] To gain new customers
- [ ] Reorganization of work or working methods
- [ ] Strategic decision making
- [ ] Selecting service providers
- [ ] To observe the work atmosphere
- [ ] Technical performance of the system (sales peaks, queues, etc.)
- [ ] Website performance
- [ ] Customer behavior on the site
- [ ] Information from [redacted] services [redacted] message automation readings, accreditation utilization rate, [redacted]
- [ ] Creating new business
- [ ] Customer relationship management
- [ ] To measure customer satisfaction
- [ ] Predicting future
- [ ] Creating and monitoring business metrics
- [ ] Marketing planning
- [ ] Else, what? _____

10. How often would you like to use the data to support decision-making in the future?

- ◯ Never
- ◯ Daily
- ◯ Once a week
- ◯ Less than once a week
- ◯ I can not estimate

11. What opportunities do you see for using data analytics to support d

You can choose as many options as needed.

- ☐ Inspire people to use data more widely
- ☐ Reduces the need for gut feeling decisions
- ☐ Supports gut feeling decisions
- ☐ Slows down decision making
- ☐ Speeds up decision making
- ☐ Gives decision making power to more people
- ☐ Contributes to the emergence of new insights
- ☐ Open doors for new business
- ☐ Reduces the risk of making the wrong decision
- ☐ Bring focus to important decisions
- ☐ Else, what

12. How would you like to be able to use data and analytics in future decision making?

13. The most important things in the analysis for me personally?

You can choose as many options as needed.

- ☐ The analysis must be understood
- ☐ The analysis is visual
- ☐ The analysis can be exported to Excel
- ☐ The analysis is self-constructive
- ☐ The analysis tool is easy to use
- ☐ The analysis is real-time
- ☐ The analysis can be shared outside the company with a link
- ☐ The analysis can be extracted as a PDF file, for example
- ☐ Else, what

14. What softwares or platforms would you like to use to analyze the data and why?

15. Would you be interested in hearing more about knowledge management, especially in relation to executive board's work?

- ◯ Yes, why ☐
- ◯ No, why ☐

16. Do you see the training of the whole company in data analytics as necessary?

- ◯ Yes
- ◯ No
- ◯ Yes, carefully selected group

17. If you see data analytics as necessary to support decision making, in which time span do you see this happening?

- ◯ As soon as possible
- ◯ Within the next six months
- ◯ Within the next year
- ◯ Within 2-5 years
- ◯ I do not see this as necessary

Appendix 2. Workshop Agenda

Workshop agenda 21.1.2022

- What does it mean to be data-driven
- Kati's thesis in a nutshell
- Presenting questionnaire results (as is , to be)
- Discussion of the topic and defining the topic of the demo (which are the question we want to find answers to)
- Discussion about the goal for the demo

Appendix 3. Semi-structured interview with CTO

CTO Semi-structured interview with the CTO
Date:10th February 2022
Platform: Teams

Themes:

Architecture and Data flow

Has the architecture been designed analytics in mind?
What is our analytics architecture like?
Could you describe the features in more detail (data warehouse, data lake etc)
What kind of changes are coming?
Is it possible to get the current data architecture model for further inspection?
Where is most of our data stored and where is it archived?
What is the data quality?

Analytics and platforms

Why was Google Analytics chosen instead of other software?
What kind of things does GA and Redshift allow?
What needs have the organization expressed for the tool?
What analytics platforms you have experienced with?
Do you have recommendations?

Maturity

Where would you place [redacted] on the following maturity model based on your opinion?
Why is that?
How do you find the data culture and attitude at the client organization?

Demonstration

What data is available for the demonstration?
In which format and in what schedule is it possible to get?

Any other topics

Appendix 4. Feedback survey after demo

1. How would you evaluate the following statements

1 = Strongly disagree
2= Disagree
3= Neither agree nor disagree
4= Agree
5= Strongly agree

	1	2	3	4	5
The demonstration answered the questions that were set for it	○	○	○	○	○
The demonstration gave me an idea of how data analytics can be built	○	○	○	○	○
The demonstration gave me valuable information about the system's data quality	○	○	○	○	○
The questionnaire, workshop and the virtual demonstration gave me valuable information about the data-driven approaches	○	○	○	○	○
I am more interested in data-driven decision-making now than four months ago	○	○	○	○	○
I would like to know more about data-driven approaches	○	○	○	○	○
I would like to build my own analytics and graphs	○	○	○	○	○
I would like the Executive Board to make decisions based more on data	○	○	○	○	○
The recommendations stemmed from the demo for the next steps were sufficient	○	○	○	○	○

2. Any other comments you would like to let the researcher know

You can give feedback also in Finnish