DATA-DRIVEN DECISION MAKING FRAMEWORK
AND ITS APPLICATION IN ESTONIAN STARTUP
SCENE

Master’s Thesis

Sigita Babarskaite and Kristiina Truuverk

Supervisor: Meelis Kitsing
Second supervisor: Marge Täks

Tallinn 2019
We here declare that we have written the Master’s Thesis independently. References have been indicated for the publications, claim, options and different sources by other authors.

23.05.2019 ............................................../ S. Babarskaite and K. Truusverk
null
2.1. Defining and organising the project ................................................................. 46
2.2. Defining the team ............................................................................................. 46
2.3. Defining the thesis parameters ........................................................................ 48
2.4. Defining the project framework ....................................................................... 49
2.5. Planning of the project ..................................................................................... 52
2.6. Tracking and managing the project ................................................................. 53
2.7. Assessing closing the project ........................................................................... 54
REFERENCES ........................................................................................................... 57
APPENDICES ............................................................................................................. 66
Appendix 1. Questionnaire and responding values ................................................. 66
Appendix 2 Factor analysis in SPSS ....................................................................... 71
Appendix 3. Interview guide .................................................................................. 73
Appendix 4. Interview Transcripts ........................................................................ 74
# LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>3Vs</td>
<td>Volume, Velocity and Variety</td>
</tr>
<tr>
<td>5Vs</td>
<td>Volume, Velocity, Variety, Value and Veracity</td>
</tr>
<tr>
<td>7Vs</td>
<td>Volume, Velocity, Variety, Value, Veracity, Variability and Visualization</td>
</tr>
<tr>
<td>BDA</td>
<td>Big Data Analytics</td>
</tr>
<tr>
<td>BDAC</td>
<td>Big Data Analytic Capabilities</td>
</tr>
<tr>
<td>B-DAD</td>
<td>Big – Data, Analytics, and Decisions</td>
</tr>
<tr>
<td>BDAF</td>
<td>Big Data Analytic Framework</td>
</tr>
<tr>
<td>BDAI</td>
<td>Big Data Analytics Infrastructure</td>
</tr>
<tr>
<td>BDI</td>
<td>Big Data Infrastructure</td>
</tr>
<tr>
<td>BI</td>
<td>Business Intelligence</td>
</tr>
<tr>
<td>CEO</td>
<td>Chief Executive Officer</td>
</tr>
<tr>
<td>DCT</td>
<td>Dynamic Capability Theory</td>
</tr>
<tr>
<td>DDDM</td>
<td>Data-driven decision making</td>
</tr>
<tr>
<td>DDDMI</td>
<td>Data-Driven Decision Making Index</td>
</tr>
<tr>
<td>EBM</td>
<td>Evidence-Based Management</td>
</tr>
<tr>
<td>GDPR</td>
<td>General Data Protection Regulation</td>
</tr>
<tr>
<td>IT</td>
<td>Information Technology</td>
</tr>
<tr>
<td>KBV</td>
<td>Knowledge-Based View</td>
</tr>
<tr>
<td>NoSQL</td>
<td>Not Only SQL Database</td>
</tr>
<tr>
<td>PwC</td>
<td>Price Waterhouse Coopers</td>
</tr>
<tr>
<td>QUAL</td>
<td>Qualitative research methods</td>
</tr>
<tr>
<td>QUAN</td>
<td>Quantitative research methods</td>
</tr>
<tr>
<td>RBV</td>
<td>Resource-Based View</td>
</tr>
<tr>
<td>SQL</td>
<td>Structured Query Language</td>
</tr>
<tr>
<td>US</td>
<td>United States of America</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Figure 1 B-DAD framework. Source: Elgendy and Elragal, 2016 ........................................ 7
Figure 2. Data-driven decision making framework. Source: Compiled by the authors of
this thesis ........................................................................................................................................ 22
Figure 3. Research design process Source: Compiled by the authors ......................... 29
Figure 4. The spread of sample within the indexes. Source: Analysis results compiled
using Tableau software by the authors ....................................................................................... 34
Figure 5. Mean results for leadership capability items by industry. Source: Analysis
results compiled using Tableau software by the authors ......................................................... 35
Figure 6. Mean results for culture capability items by industry. Source: Analysis results
compiled using Tableau software by the authors ......................................................................... 37
Figure 7. Mean results for analytic skills & data literacy capability items by industry.
Source: Analysis results compiled using Tableau software by the authors ....................... 39
Figure 8. Mean results for technology and data capability items by industry. Source: Analysis results compiled using Tableau software by the authors ....................... 40
LIST OF TABLES

Table 1. Technologies and Best Practices for Becoming a Smarter Organization Source: Stodder and Fern, 2017. .......................................................... 14
Table 2. Comparative analysis of frameworks. Source: Compiled by the authors..... 17
Table 3 Overview of interviewees. Source: Compiled by the authors.................... 28
Table 4. Overview of capability indexes by demographics. Source: Analysis results compiled using Tableau software by the authors.................................................... 33
Table 5 The project team and responsibilities. Composed by authors. ............... 47
Table 6 Schedule for project management. Compiled by the authors ..................... 48
Table 7 Project framework for communication. Compiled by the authors........... 50
Table 8 Team Assessment Characteristics. Compiled by authors .................... 55
Table 9 Overview of project effectiveness. Compiled by authors ....................... 55
ABSTRACT


DATA-DRIVEN, DECISION MAKING, STARTUPS, CAPABILITIES, FRAMEWORK, DDDM, STRATEGY.

Increase of data produced and big data have spurred the popularity of data-driven decision making, turning the concept of DDDM into a catchphrase. Companies become aware that in order to gain a competitive advantage, making guesses might not always help and using just experience when it comes to strategic decisions, might not lead to expected outcomes. With the examples of Google, Facebook and Amazon, companies to realise the value and insights of data, when it is gathered, analysed and used as a product, rather than a by-product of information systems. There are several theoretical concepts compiled to understand the different capabilities needed in order to successfully implement data-driven decision making for value creation. This master thesis emphasises on resource-based view, dynamic capabilities theory and evidence-based management. Based on these insights, the authors formed a framework, testing the adoption level of these capabilities in Estonian startups. Then, mixed method research was applied (questionnaire n=33 and semi-structured interviews n=8) to measure the data-driven decision making index.

The key results of this study revealed that data-driven decision making mindset was strongly present through culture and leadership. However, the most improvements needed to become data-driven consisted of technology integration, fostering data literacy and changing the attitude of the management. Together, they shaped the main barriers for organisations growing from aspirational to data transformed. Findings, from this study, are generalizable for IT organisations in growth or maturity stages.
INTRODUCTION

With the increase of technology usage in everyday business and the data produced by that technology, companies are starting to talk more about the need of using data in order to get a competitive advantage over their competitors. This concept is called data-driven decision making (further referred to as DDDM). In DDDM strategic and everyday decisions are relying on proofed concepts by data, rather than on business acumen. DDDM is ought to originate from business management tools like Knowledge Management (Davenport & Prusak, 1998) and Total Quality Management (Deming, 1986). In this thesis Dynamic Capabilities Theory (Teece, Pisano and Shuan, 1997) and Resource-Based View (Barney, 1991) are also used to support the authors’ definition of DDDM.

Furthermore, with the increasing usage of the term “big data”, different companies are not just implementing DDDM in their strategic management, but new companies established are building their products and services either based on data or data-sharing. Due to big data, leaders can measure and get to know much more about their businesses, and translate the insights gained into improved decision making and increase company performance (McAfee and Brynjolfsson, 2012). A study conducted by McAfee and Brynjolfsson (2012) discovered that companies who are data-driven are on average 5% more productive and 6% more profitable than their competitors. Mazzei and Noble (2017) bring out three tiers of big data - data as a tool, data as an industry and data as a strategy. The companies in the highest tier, like Facebook, Google and Alphabet continually develop their platforms that spread through consumers lives building data stocks and data flows, which will be used to later monetised through new products and technology, generating even more data for later use.

This master’s thesis aims to understand the concepts of applying DDDM: what capabilities does a company need in order to define itself as data-driven; and what are the different levels of being data-driven? In the essence of this paper, data-driven decision making is researched in all of the processes of a company. In services and
product development; finance, marketing, human resources related decisions; and using data as a service itself. After defining the basis of the concept, the authors will propose an index of data-driven decision making which will measure the maturity of adoption of those capabilities within a company. This index will be further tested on the Estonian startup companies, to understand how data-driven on average they are. The framework of this research can be later used by companies to have a full overview of what capabilities are important at different data-driven decision making process phases, and the index proposed by the authors can be together with the survey questions used to understand which capabilities need further improvement to become more data-driven in their decision making.

The thesis consists of two parts – a literature review with empirical research and project management of the thesis. In the first section, the authors of this thesis review the literature in order to define DDDM and understand the necessary capabilities for a company to become data-driven, and which are the different levels of being data-driven. They are examined to understand the different phases of being data-driven. At the end of the first part of the thesis, the authors create a framework for applying DDDM in an organisation and defining the required decision-making capabilities and processes.

In the second section, the authors introduce in more detail how the concept is measured, the type of research methods deployed and how the sample for testing for formed. Moreover, the section presents the calculation used for measuring DDDM Index (DDDMI), proposed by the authors. Then, the overview is presented about the survey sent out to the Estonian startup community members, followed by the insights gained from eight semi-structured interviews. The segment then shows the results of the calculation of the DDDMI and where do the companies in the sample stand within the range. The final part of this section introduces conclusions, limitations and proposals for future research in the field of DDDM.

The remaining part of this thesis is an overview of project management. The authors introduce an overview of how this project was organised and later managed. Also, showing the key learnings from the thesis at hand, working together as a team. The information, presented in this section serves as a guide to similar projects for future researchers and business practitioners alike.
1. FRAMEWORK FOR DATA-DRIVEN DECISION MAKING AND ITS ADOPTION WITHIN ESTONIAN STARTUPS

1.1. Literature review

1.1.1. Definition of concepts of data-driven decision making

To start defining data-driven decision making (DDDM) one needs to start by dismantling the concept to pieces. If we look at the first part of this concept, “data” then Davenport and Short (1990, p. 4) describe data as a set of discrete, objective facts about events which in an organisational context is described as structured records of transactions. Rijmenam (2014) in his book “Think bigger” points out that out of all the recorded data, before 2014, 90% had been created within the last two years. Despite that fact that examples of earliest data usage can be tracked back 7000 years to Mesopotamians, who used it for recording growth of crops and herd. Nowadays people do not just talk about data, as the volumes have grown, we mostly talk about “big data”, a term first presented in 2005 by Roger Magoulas (Halevi and Moed, 2012). This term is best described by Lake and Drake (2014) as a location where large, complex data sets are stored and analysed by employing different methods. Özköse and Gencer (2015) have summarized in their work the different definitions for this including data that needs new forms of processing (Gartner, 2014), data that is beyond the technology’s current ability to process (Manyika, Chui, Grown, Bughin, Dobbs, Roxburgh and Byers, 2011) and data as cluster that unfolds hidden values (Hashem, Yaqoob, Anuar, Mokhtar, Gani and Ullah Khan, 2015). When people talk about big data, they mostly refer to its 3Vs – volume, velocity and variety (McAfee and Brynjolfsson, 2012), or 5Vs adding value and veracity to the before mentioned three (Demchenko and Membrey 2014). Most recently even 7Vs – adding variability and visualisation to the 5 mentioned previously (Rijmenam 2014). Bradley, Barbier and Handler (2013) state that as little as half per cent of companies use big data, while more than 50% of the data projects never get
completed. To summarise, then in this master thesis data and big data specifically are referred to as a set of records about transactions which in the case of big data are too large and complex that they need specialised technology to process and analyse it.

Continuing forward with Davenport’s and Prusak’s (1998) concept, data itself is just a raw material and it cannot tell one what to do, but data is essential in order to create information. Davenport and Prusak (1998, p. 3) describe data as “a message, usually in the form of a document or/and audible or visible communication”. Put into simpler words we could say that if data itself is a simple set of numbers and text, then information gives meaning to the data, and the main form for developing meaning from data would be analysis. Without the information, data does not hold any value (Rijmenam 2014). This concept is also discussed by Vollenweider (2017) in his book Mind+Machine, where he describes that data includes four fundamentally different levels. Starting from the lowest: 1) data itself, 2) data evolving into information, 3) insights and 4) finally knowledge. While he agrees that for levels one and two machines can make most of the work, for insights and knowledge, the human mind is required.

Information theory (Blackwell, 1953) and organisational view of information processing (Galbraith, 1974) propose that higher company performance is facilitated by greater use of information in decision making, which should arise from accurate data.

Data analytics can be defined as processes and tools, which include statistics, predictive analytics and data mining, often applied to large datasets in order to create value and insights (Russom, 2011; Ertemel 2015). There are three main types of analytics (Power, 2014) which include 1) retrospective data analysis (historical data used to find patterns), 2) predictive data analysis (scenario analysis on historical to predict future) and 3) prescriptive data analysis (includes the results of two previous types and prescribes how to take advantage of the future). All three types are essential to help make managers informed decisions, but the usage shows how well evolved these companies are in data-driven decision making.

Information itself, according to Davenport and Prusak (1998) does not lead to knowledge, as it is more profound and broader - a mixture of experience, values, expert insights and information. According to Davenport, knowledge evolves from information, which in turn derives from data. In order to make well-informed decisions, one needs to have not only data but also some informative analysis. The analysis is vital.
wider to gain hindsight, insight and foresight. However, Wu, Buyya and Ramamohanarao (2016) argue that more than often these aspects are neglected, and focus is targeted towards data aspects (3-7Vs). Ghasemaghaei, Ebrahimi and Hassanein (2018) have surveyed 151 Information Technology managers and found data analytics competencies to improve a firm’s decision quality and efficiency significantly. The survey analysed the impact of the five most imperative data analytics competency indexes, including data quality, the bigness of data, analytical skills, domain knowledge and sophisticated tools. They argue that without such competency, companies would most probably fail to improve decision making performance despite implementing new analytical tools. The underlying theory taken to improve the importance of the five competency indexes comes from Huber (1990) who in his work concluded that the availability of advanced information technologies leads to increased use of the same technologies, which sequentially increases access to information and changes in organisational design. All of this leads to improvements in the effectiveness of intelligence development and decision making within a company. Data literacy is not only the concern of top managers but something for all parties contributing to the data-driven decision making culture within a company. Data literacy is a set of critical data and information competencies, that enable locating, collection, analysis, integration and communication of information (Bennet, 2004; Mandinach and Gummer 2013; Means, Padilla and Gallagher, 2010).

The last aspect of DDDM is decision making, which is arranged by the decision maker, be it an individual or a group, who use their final value judgment to rank available alternatives in order to identify the best choice (Chankong & Haimes, 2008). Edwards (1954) introduces the theory of riskless choices which he summarises by any man making a decision, which is riskless, is to be considered an economic man. This type of a leader is considered to have three properties – completely informed, infinitely sensitive and rational, ultimate meaning that his choices are based on maximising outcomes or creating value and increasing performance. Organisational decision making at the same time has been defined as information behaviour, which is a result of an individual’s interaction between himself and the environment within what he operates (Taylor, 1991). Furthermore, Taylor’s (1991) theory of Information Use Environments suggests that principal’s data use is primarily affected by three factors including 1) shared assumptions between peers about their work and information usage;
2) problem dimensions used to judge the usefulness of information; and 3) attitude towards the availability and value of information. The concept of environment on the adoption of information usage in decision making is also supported by O’Reilly (1983) who emphasized the importance of organizational context, proposing that information is a commodity, which is more likely used when it is accessible, summarized and interpreted, valid and reliable, and adopted into an operating control system. Other concepts of DDDM discussed in literature originate, in addition to the two aforementioned, on Total Quality Management (Deming, 1986), Knowledge Management (Davenport & Prusak, 1998), Resource-Based View (Barney, 1991), and latest Evidence-Based Management (Barends and Rousseau, 2018). This paper will emphasise on the last two theories in building the framework of the successful data-driven strategy for different value creation levels that will later be tested within the Estonian startup community.

1.1.2. Data usage for the decision-making process

In the heart of data transformation to knowledge is knowledge management theory for this master’s thesis. This theory has evolved from focusing on simple knowledge storing in containers, to then focusing on people and culture. With the evolution of technology, knowledge management is believed to reach the third generation, which focuses on shared context (metadata) (Mašic, Nešic, Nikolic and Dželetovic, 2017).

So far, this paper has looked at data, information, knowledge, data analytics competencies and slightly the importance of advanced analytical tools in the equation to lead to better decision making. Davenport (2014) argues that the ways of analysis for decision making have changed due to the amount of underlying data. While in traditional decision-making situations an analyst has gathered a set of data based on what they have performed analysis and came up with a model which advises the decision makers on the results, then with big data, the data is considered to be an ongoing fast-flowing stream which requires continuous analyses and actions as decisions. Elgendy and Elragal’s (2016) have developed B-DAD (“Big – Data, Analytics, and Decisions”) framework, which is a good source for managers to map which tools, architectures and analytics are needed for different decision-making phases. The framework includes four main phases and their subphases, see Figure 1.
After testing the framework in the retail industry, Elgendy and Elragal (2016) found that the results showing added value when big data analytics (BDA) were integrated into the decision-making process. Furthermore, as BDA differs from traditional by being more unstructured, they argue that not always is it significant to know the decision beforehand as unprecedented insights might lead to new decisions. Several other big data frameworks underlie the successful implementation of data-driven strategy in an organisation to create value, will be discussed later in this chapter.

Most of researches and books about big data and decision making emphasise the importance of value creation and its barriers to the creation of value. A thorough literature review conducted on the topic of value realisation from big data found that current literature is still at nascent stage having more papers focusing on the theoretical perspective rather than empirical (Günter, Rezazade Mehrizi, Huysman and Feldberg, 2017). Out of all examined studies, they found two socio-technical features, portability (remote access and transfer to data in different contexts) and interconnectivity (possibility to combine data from different sources) of data, that most shape the value realisation from big data. Furthermore, they argue that companies need to continually realign their practices, models and stakeholder interests, in order to not fail the creation of value. Other research and papers bring out the importance of data mastering by elaborating data in their business models and strategy (Mazzei and Noble, 2017, McAfee and Brynjolfsson 2012), translating analytics into actions and insights (Marr, 2016), technology (McKinsey, 2014), increasing middle management skills (McKinsey 2014; Barton and Court, 2013; NewVantage Partners, 2017; Jain and Sharma, 2014) and culture (Barens and Rousseau, 2018; Davenport, 1998; Schein, 2010). Despite
having all of the before mentioned in mind, decision-makers also need to consider ethical and legal aspects of big data analysis, in order not to cross any boundaries; as a result, leading to more significant losses than gains.

One should keep in mind that not all data is useful or relevant for analysis and surplus of data might lead to insights unseen. Data is often referred to as the new oil (Monino 2016). With big data, the eagerness to automate processes for decision support increased, bringing along new challenges. These include the complex invisible models, which require knowledge used at the right time within the correct context and approach. Then, more significant issues with data capturing, management and securitisation (Monino 2016). Monino (2016) points out that the basic requirements preceding any analysis are knowing which data is correct and how to tackle it. If the underlying data used for decision making is considered “poor”, it produces weak or even harmful decisions. With no concrete proof of the quality of the data used for analysis, managers are blind to know if the outcomes produced would conclude with positive effects.

A survey conducted on the data quality issues by the Data Warehousing Institute in 2001 estimated that the cost of poor data quality is around $600 billion a year for U.S. businesses (Eckerson, 2002). Such quality may arise from different reasons, manual data input with errors, the same data defined differently in various company departments and similar. Without knowing the business logic behind data, it can be misinterpreted, or other technical issues might impact the entire dataset and turn it to incomplete. To ensure the quality of one’s database, one should use data quality metadata: a description of the quality of the data by its origin, contents, condition and other characteristics. (Cai 2007). Other include simple ratio testing (Ballou and Pazer, 1995) and treating data as a product instead of a by-product of information systems (Wang, 1998). Using metadata in the process of decision-making will potentially reduce the negative impacts of sparse data, but it has a cost and an impact on decision efficiency and time (Cai, 2007; Price and Shanks 2011).

**1.1.3. Business analysis and supportive Information Technology (IT) in value creation**

As mentioned in the first section of this master thesis, we cannot talk about DDDM without going more in-depth into the concepts of business analytics and IT tools, which
support the preceding. Recently, several researches have been published on identifying value creation through business analytics or big data analytics (BDA) (Grover, Chiang, Liang and Zhang, 2018; Elgendy and Elragal, 2016; Wamba, Gunasekaran, Akter, Ren, Dubey and Childe, 2016; Costello and Osborne, 2005; Lavalle, Hopkins, Lesser, Shockley and Kruschwitz, 2010; Power 2016). Even though in the previous sections of this work we spoke about analysis creating meaning to data and producing information, we did not define the concept of business analysis. This short term has many different definitions. According to the International Institute of Business Analysis (IIBA), business analysis is defined as the practice of enabling change in an organisational context by defining needs and recommending solutions that deliver value to stakeholders. Other definitions include: placing emphasis on analysing the operations of the entire business system (Paul, Yeates and Cadle, 2010); business practice to link a company’s capabilities with its objectives; finding business needs and proposing solutions to various problems; focused on integrating IT and business units (Haas, 2017; Clare, 2011) and often confused with the term business architecture (Iyamu, Nehemia-Maletzky and Shaanika, 2016). Chen, Chiang and Storey (2012) bring out that although intelligence was already used in the 1950s and business intelligence (BI) appeared in the 1990s, business analytics was introduced rather late. In 2016, Davenport introduced business analysis as the key component of business analytics (BI).

In general, definitions conclude that business analysis is about investigating something to produce value. In most cases, the analysis is based on processes or data produced by these processes. Lavalle and his colleagues (2010) argue that in most cases companies put the emphasis on gathering the data rather than defining the questions what they want answers to. There are several frameworks for being data-driven, emphasizing the importance of starting decision making with a clear question (Jain and Sharma, 2014; Barends and Rousseau, 2018), strategy (Marr, 2015) or goal (Brownlow, Zaki, Neeley and Urmetzer, 2015). These frameworks will be introduced in the proceeding part of this thesis. Concentrating too much on data, most probably leads to data overload and thus finding a way to recognize the correct data for analysis might become an obstacle for efficient decision making, especially in the light of big data. What makes business analysis so important in the decision-making process? Business analysis is believed to increase company performance, productivity, main financial metrics like return on assets and equity, but also impact positively company’s market value by using data in
order to make more evidence-based business decisions which are less intuition based (Brynjolfsson, Hitt and Kim, 2011; McAfee & Brynjolfsson, 2012; Bange & Janoschek, 2014; Barské-Erdogan, 2014; Barends and Rousseau, 2018). A study conducted by MIT Sloan Management Review and IBM Institute for Business Value found out that top-performing companies are five times more adaptive to use business analytics than lower performers and they are more likely to use analytics in their day-to-day operations and future strategies. Furthermore, companies that include BDA in their daily operations are found to be 5% more productive and 6% more profitable than their peers (Barton and Court, 2012; McAfee & Brynjolfsson, 2012). The findings of research carried out by Niland (2017) within South African businesses, confirm the direct and moderating impact of on BDA and Firm Performance (correlation of 0.632), and emphasise that BDA should not be considered as a silo in an organisation.

A good business analytics setup needs a well thought through IT system, that would support the efficiency of analytics procedures and create competitive advantage. Studies concentrating on investigating the role of IT (resources) in value creation bring out theories like resource-based view (RBV) (Barney, 1991), knowledge-based view (KBV) (Grant, 1996), and dynamic capability theory (DCT) (Teece, et al., 1997). Côrte-Real, Oliveira and Ruivo (2016) have used the combination of these theories to empirically demonstrate the importance of analytics and knowledge management in the creation of agility, which leads to competitive advantage. Furthermore, their study highlights that business analytics can create value at several levels including knowledge, dynamic capabilities, business processes and competitive performance, only if companies understand the needs of proper investments in business analytics technologies.

As several studies emphasise, then investment in underlying technologies is a crucial part of successful business analytics adoption. Davenport and Short (1990) argue that in order to leverage from main IT capabilities (transactional, geographical, automational, analytical, informational, sequential, knowledge management, tracking and disintermediation) companies need to consider the role if IT in the early stages of process design. However, what exactly constitutes to this technology? There are not many research papers on identifying the most valuable foundations of a functional infrastructure for BDA. Some consider a successful infrastructure to include several
data sources; a platform which would integrate, collect, share, process and manage data; and analytics infrastructure which is optimally aggregating, transforming, distributing, analysing and modelling to data for further use (Grover et al., 2018). An ideal infrastructure needs to deliver at least the 3Vs (up to 7Vs) of data in fast speed to users. Poorly performing companies just dump new storage and hardware into existing systems in order to get access to more data, but often this leads to degraded performance. A study conducted by Capgemini found out that 79% of organisation have poorly integrated their data flows and only 35% have built robust systems for data management (Baldwin, 2015). Most of these platforms for data management used to be internally developed by companies but are nowadays more and more cloud-based (storage) and use open-source software (for management, like Hadoop) as the balance between speed and cost, and security is taken into consideration.

Demchenko and Membrey (2014) propose a Big Data Architecture Framework, which comprises of five main components; 1) data models, 2) structures and types, 3) big data management, 4) BDA and tools, 5) big data infrastructure (BDI) and big data security; all of which play an essential role in the big data ecosystem. Their proposed BDI includes comprehensive data management tools, which are typically cloud-based and also BDA infrastructure, which requires high-performance computing clusters and high-performance low-latency network. BDI includes big data management tools, registries and indexing search, security infrastructure and collaborative environment. In contrast, BDAI suggests including cluster services, Hadoop related services and tools, specialist data analytics tools, databases of SQL and NoSQL and massively parallel processing databases in order to provide the best performance. Taking all of this into account there are also several challenges that need to be taken into account while designing a well-performing infrastructure. These challenges include interdependencies of infrastructure elements, unpredictable workloads and data management. While these infrastructures can be built on locally managed hardware solutions, a rising era of cloud computing has recently emerged to provide access to information from everywhere at any time, and maximisation of capacity and performance of the infrastructure (Stergiou and Psannis, 2017).

But it is not IT infrastructure that directly contributes to increased company performance, but rather the expertise of the personnel and relationship infrastructure,
that relates to competitive advantage (Bhatt and Grover, 2005) as the ubiquity of sophisticated infrastructures is available to all companies and thus does not create differentiation (Carr, 2003). Aral, Brynjolfsson and Wu (2010), in their research carried out in 2008, concluded that one of the reasons the US increased productivity, in comparison with Europe, was a correct deployment of the IT systems.

1.1.4. Organisational decision making and company culture

This paper has already brought out the importance of management, personnel and culture for value creation from data and information but has not tackled this topic so far from the perspective of how to manage them in order to create increased value. Several studies emphasize the importance of including data in strategy making, innovation and culture (McAfee and Brynjolfsson, 2012; Mazzei and Noble, 2017; Frisk and Bannister, 2017; Davenport and Harris, 2007; Cao and Duan, 2017; Bhatt and Grover, 2005), but the challenges of adopting new ways of thinking are still significant. One of the challenges brought out by McAfee and Brynjolfsson (2012) is muting the “HiPPO” (highest paid person’s opinion) (Kaushik, 2009), which basically indicates that decisions are made based on experience, intuition and relationships, rather than hard evidence, and as most of the time this term is used to refer to the management. While several companies still dictate which data should be analysed and collected, high performing companies use data collected and analysed to form corporate strategies and gain an advantage based on the changed mentality (Mazzei and Noble, 2017; Bhatt and Grover, 2005).

In order to be effective in the adoption of new ways of thinking, data-based decisions need to be made at all levels of the company and embedded in the system, otherwise it will lead to conflict (Leidner and Kayworth, 2006). The culture for being adaptive to internal innovation and acceptance of new capabilities is created by leaders (Niland, 2017). This is believed to be achieved through clearly stated and communicated vision, the autonomy of personnel, promotion of risk taking and motivation towards adoption (Niland, 2017; Agostini, Nosella and Filippini, 2016). Wohlstetter, Datnow and Park (2008) state that internal stakeholders in the organisations need to each see the value of data within the organisation, using the principle of “build the thirst for it”. In their findings, it is imperative to create a culture where data is regularly deployed among all
layers of the organisations, but also people are valued for using data. They argue, if this principle is not in place, the culture of using data will have temporary effects, as the agents with authority diminish, having the power, but no motivation to use it. In addition to the common culture, creating clear guidelines for everyone in the organisation to be responsible for data ensures that the deployment works.

A strong corporate culture influences the performance of an organisation (Denison & Mishra, 1995). Four characters of organisational culture influence better performance: involvement and adaptability, consistency and mission. Together they are forecasters of effectiveness in terms of quality, employee satisfaction and general performance (Denison & Mishra, 1995). For knowledge management, culture is imperative: by determining the importance of knowledge to be exchanged, shared and valued from singular to business level (Martin, 2008). At the same it, culture is about the context, which takes place in an organisation, thus defining how the information is shared, knowledge created and finally legitimised for further use (Beccera-Fernandez & Sabberwal, 2001; Karlsen & Gottschalk, 2004). McDermott & O’Dell (2001) and Albert and Picq (2004) agree that organisations, which are more open, trusting, encouraging learning and collaboration, tend to be more effective in overall knowledge management (Tian 2017). The work of Aral and his/her colleagues (2010) corresponds with the previous research, conducted by Holmstrom & Milgrom (1994), Milgrom & Roberts (1990, 1995), and Kandel & Lazear (1992), concurs that organisations gain the most significant benefits by working incoherence, not silos.

The most common way to incorporate DDDM is through the democratisation of analytics. By enabling every person at the company, especially in the executive level, to access data and have reports understandable. Harper and Stodder’s (Fern Halper, 2017, pp. 3-5) survey reported critical attributes of data-driven organisations:
Table 1. Technologies and Best Practices for Becoming a Smarter Organization Source: Stodder and Fern, 2017.

<table>
<thead>
<tr>
<th>Technology Traits</th>
<th>Organisational Traits</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Support of analytics development and deployment through integrated data</td>
<td>• Goal oriented</td>
</tr>
<tr>
<td>management</td>
<td>• Strong leadership</td>
</tr>
<tr>
<td>• Integrated analytics strategy using a variety of tools</td>
<td>• Transparent/ethical</td>
</tr>
<tr>
<td>• Integrated analytics</td>
<td>• Enabling</td>
</tr>
<tr>
<td>• Continuous improvement through consistent feedback about the analytics</td>
<td>• Ability to do it yourself/self-service</td>
</tr>
<tr>
<td></td>
<td>• Focus on skills</td>
</tr>
<tr>
<td></td>
<td>• Governance driven</td>
</tr>
</tbody>
</table>

In the global survey, conducted by PwC in 2016, 2100 executives, top-level leaders and business unit heads say that they would be ready to use data to drive better decisions (61%), but in their organisations, data is somewhat or even rarely used. One of the biggest obstacles in using data is company culture, followed by traditions, which rely on decision-makers using intuition and past experiences (PwC, 2016). Over half of the leaders dismiss data, presented to them because they do not understand the value or the value and use of data are not explained for them clearly enough. Leaders often use data to support their conclusions about the decision, not vice versa (PwC 2016). The study on data usage concludes that 8% of organisations are rarely data driven, 53% are somewhat and 39% are highly driven with the use of data for predictive purposes.

Brownlow and his colleagues (2015) developed a framework to determine six steps a data-driven organisation operates. In their blueprint, they outline that the first step is the goal, which is aimed at achieving with the use of big data. Secondly, an organisation should review and consider the type of the value proposition or offer it wants to provide. Thirdly, the type of data necessary to achieve desired goals, followed by acquisition. For example, considering whether the date is readily available or should be outsourced. Fourthly, follows actions to processing and application of data. To measure the effectiveness of the previous steps, the authors explain that revenue models should be derived from the deployed data. Finally, in order to successfully implement the process of DDDM, it is important to mitigate the risks. In the qualitative research executed by Brownlow, et. al. (2015), both established organisations such as retailer Zara, and startups showed that successful cases depend on clear target setting, while obstacles included data quality issues, accessibility and integrity along with legislative issues such
as GDPR. Setting correct targets with a comparison of readily available information and historical data is also brought out by Brynjolfsson, et al. (2011) as findings of a US-wide study which analysed value added after between early and late adopters of data-driven decision making. Thus, one could conclude that leaders play a crucial role in developing a culture that would emphasise the adoption of organisation-wide data-driven decision making mentality.

1.1.5. Underlying theories for data-driven decision making framework generation

Thus far, this thesis has emphasised on defining the concepts that are behind DDDM and the importance of data, analysis, technology, culture and leadership. Previously introduced in the thesis, Davenport's value creation from data was incremental for unifying data and knowledge. Also, this work will also present Resource-Based View (Barney 1991), Dynamic Capability Theory (Teece et al., 1997) and Evidence-Based Management Theory (Barends and Rousseau, 2018). Together, theories will form the basis for defining a DDDM strategy framework, which will focus on the company's capabilities to create value and increase competitive advantage.

To understand the necessary resources for creating competitive advantage and creating value it is important to start with resource-based view (RBV) theory, the pivotal work of Barney in 1991. The idea of the RBV theory is that companies are a collection of resources, which give them a competitive advantage (Barney, 1991; Teece et al., 1997; Pisano, 2015). The RBV theory classifies resources into three categories: 1) physical capital, 2) human capital, and 3) organizational capital (Barney, 1991). Physical capital constitutes of certain assets, including technology, raw materials (data in the case of this paper) and equipment. The human capital is considered to include expertise, managerial knowledge, judgment and training. The organisational capital comprises of reporting, controlling, coordination and other processes (Barney, 1991). According to Pisano (2015), as RBV gave little definition to the capabilities and guidance on which kind of capabilities need to be either created or further developed in order to achieve or keep a competitive advantage, the theory of dynamic-capabilities (DCT) introduced by Teece, et al. (1997) emerged in 1997. Despite there still exists confusion and different interpretations about the basic definition of dynamic capabilities (Peteraf, Di Stefano and Verona, 2013), this thesis underlies the main concepts outlined by the original
authors being the ability of a company to build and reshape internal and external competencies in order to adapt to the changing environment (Teece, et al., 1997). Furthermore, the authors emphasize that wealth creation is dependent on internal technological, managerial and organisational capabilities, which are difficult-to-imitate. The theory introduces three classes of factors – process (routines and practices), positions (tangible and intangible assets) and paths (strategic alternatives), that help to define internal competencies. For the purpose of this study, the two aforementioned theories are used to map the capabilities necessary for data-driven organizations who are seeking for competitive advantage.

As a third pillar to this research, the authors use evidence-based management (EBM), introduced by Barends and Rousseau (2018) as a practice where decisions are made through meticulous, explicit and careful use of the best available evidence from multiple sources to map the process (first factor of RBV) of a data-driven organization. The framework defines three main phases of an organisational culture that need to be developed (1) have a questioning mindset, (2) make decisions more explicit, and (3) practice and learn on a daily basis. This theory considers information, data, facts and assumptions as evidence. The authors of the strategic management theory pinpoint that confronting people with hard evidence may backfire but having leadership and culture which fosters evidence-based management will increase openness to new beliefs, decrease judgment-based decisions and would increase the quality of decisions made. For the purpose of this thesis, data is considered as evidence for decision making which needs to be acquired from multiple sources. In the heart of evidence-based management (EBM) are six steps that need to be followed in order to increase the likelihood of a favourable outcome from data: ask, acquire, appraise, aggregate, apply and assess. As we previously have mentioned that in order to gain insights from data, we need analytics. EBM can be paralleled with several other publications, which outline frameworks for successful implementation of data-driven decision making, including analytics-based BADIR (Jain and Sharma, 2014) and SMART (Marr, 2015).

BADIR is a highly effective business analytics framework which aims to solve business problems, but to fully benefit from it, both technical and business tracks need to be followed. Jain and Sarma (2014) emphasize that companies need to strive in four main areas for full leverage of analytics: leadership, analytics talent, decision making and
data maturity. These areas can be considered as capabilities of a company essential for effective analytics adoption. BADIR is a five-step framework, which starts with a Business question, Analysis plan and Data collection, and ends with Insights and Recommendations. These steps need to be applied twice – first to gain insights from data, and secondly to gain impact from insights.

The following analytics framework, which will support this paper’s framework is SMART (Marr, 2015). In the heart of this framework is also data (big data) and the underlying is the importance of technology, but like BADIR this framework does not start with data, but Strategy. The other four steps of this framework are Measuring metrics and data, applying analytics, reporting results and transforming business. This way the user will not be lost in “what is possible?”, rather have focus on “what data is important?”. While BADIR AND SMART frameworks are rather business analytics centric, a fourth framework presented by Brownlow et al. (2015) includes six steps similar to evidence-based management. It is more developed to outline how data-driven organisation operates — starting with defining the goal, reviewing value prepositions and offer, determining data requirements, and processing and application of data. The last two steps are monetisation. Revenue models need to be derived from data, and implementation, which also considers risks.

A comparative table aligning different frameworks is visualised in Table 2.

Table 2. Comparative analysis of frameworks. Source: Compiled by the authors

<table>
<thead>
<tr>
<th>Step</th>
<th>EBM</th>
<th>BADIR</th>
<th>SMART</th>
<th>Brownlow</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Ask</td>
<td>Business</td>
<td>Start with strategy</td>
<td>Goal</td>
</tr>
<tr>
<td>2</td>
<td>Acquire</td>
<td>question</td>
<td>Measure metrics and data</td>
<td>Value offering</td>
</tr>
<tr>
<td>3</td>
<td>Appraise</td>
<td>Analysis plan</td>
<td>Determine required data</td>
<td>Determine required data</td>
</tr>
<tr>
<td>4</td>
<td>Aggregate</td>
<td>Data collection</td>
<td>Application</td>
<td>Application</td>
</tr>
<tr>
<td>5</td>
<td>Apply</td>
<td>Insights</td>
<td>Report results</td>
<td>Monetisation</td>
</tr>
<tr>
<td>6</td>
<td>Assess</td>
<td>Recommendation</td>
<td>Transform business</td>
<td>Implementation of obstacles</td>
</tr>
</tbody>
</table>

1.1.6. Mapping of capabilities and value creation

To map capabilities that increase company competitive advantage and increase value, we start with the main three capabilities defined within RBV theory: physical, human
and organisational. These in the context of DCT relate to internal technology, managerial and organisational capabilities. In some of the research also a fourth pillar is introduced – the financial capital (Pyatt and Becker, 1966; Barney and Clark, 2007), excluded from this thesis. This thesis aims to open up these capabilities with the help of other frameworks designed for big data, followed by concepts highlighted in the literature for data-driven organisations (big data, analytics, technology and culture). According to Zeid and Davis (2014) to increase a company’s maturity, it needs to focus four pillars – people, processes, technical infrastructure and culture. While the first three concepts are familiar from RBV and DCT, then a layer of culture (Barney and Clark, 2007) falls under organisational capability. This classification of capabilities into three fundamental categories - people/culture, processes and technology, is also supported by knowledge management theory (Liebowitz, 2009).

To further dismantle and define the main competencies for this paper’s framework, the authors discuss human capital and managerial capabilities. As previously said, this comprises mainly of experience (knowledge), intelligence and judgment (Barney, 1991). In the concept of DDDM this can be viewed from three crucial perspectives – the expertise of personnel, data analytics and data literacy (Lavalle, et al. 2010; Lou, Albrecht and Neill, 2015; Mandinach & Gummer, 2013; Côrte-Real et al., 2016). Big data analytic capabilities (BDAC), which lead to increased firm performance are dominantly described in the literature by already known three dimensions – management, personnel and infrastructure capabilities (McAfee and Brynjolfsson, 2012; Davenport et al., 2012; Barton and Court, 2012). A company’s IT capabilities are considered to also include the three capabilities above, only in the perspective of IT (Kim, Shin and Kwon, 2012). In order to develop BDAC Grover, et al. (2018), supported by Wamba, et al. (2016), state that companies need to invest in big data assets, analytics portfolio and human talent, which by their value creation framework lead to functional value (financial performance) and symbolic value (e.g. positive brand image). They bring out six mechanisms that moderate the linkage between capabilities and value, including transparency and access; discovery and experimentation; prediction and optimization; customization and targeting; learning and crowdsourcing; and continuous monitoring and proactive adoption. While RBV and DCT draw three main capabilities for overall success, BDAC states that these dimensions are also
important on the individual levels of analytics capability and technology capability, and their adoption.

A recent study by Lavalle, et. al (2010) categorises analytics adoption (data literacy) into three stages – aspirational (data used justification based), experienced (data used for guidance) and transformed (data used for prescriptive actions). Each stage posits its challenges, obstacles, management issues and analytics actions. According to his theory, organisations might have all the capabilities needed, but the adoption and sophistication level might lag behind. Organisations at the aspirational stage are relatively far from achieving their analytical goals and only have a few of the underlying necessary building blocks as people, processes and tools in place. Their focus is usually on finding ways to automate and cut costs. The second level of experience is described by Lavalle (2010), where organizations go beyond cost management, creating efficient ways of collecting and incorporating analytics for business optimisation. The last layer of an organisation – transformed – is using analytics as a competitive differentiator. Transformed organisations are less concerned about cutting costs. Instead, they make targeted investments in niche analytics.

An important aspect of human capital is employee knowledge (Grover et al., 2018). Simon (2014) illustrates the concept of a visual organisation as one that comprises of knowledgeable people recognising the importance and power of data. Professionals need to develop skills both on the business and the data side. The skills are the hands-on analytics and testing, working efficiently with data analysts and knowing at least some of the basic analytics tools (Jain and Sharma, 2014). Consequently, companies need to have more technology professionals with business skills (Ohlhorst, 2013). For many enterprises, the biggest challenges in becoming data-driven are the lack of skilled big data experts such as data scientists, developers, programmers and analysts (Grover et al., 2018).

The physical capital capabilities in the context of data-driven decision making can be divided into two – data as an asset and technology needed for managing data and analytics tools. They need to be thought in the context of connectivity, compatibility and modularity (Wamba et al., 2015). Nguyen, Taylor and Franks (2016) bring out that data-driven organisations focus on database performance, data quality and integration
and analytics capabilities. They treat data as a strategic asset, which helps to achieve desired goals.

The data excellence model, published by the Competence Center Corporate Data Quality explains that to achieve data excellence and increase business value, both business capabilities and data management capabilities need to be treated as data strategy. The enablers defined in the perspective of data are data lifecycle, data applications and data architecture. Another perspective of the differentiation of big data usage is presented by Mazzei and Noble (2017), where they introduce a three-tiered framework. As a company progresses within the tiers, data grows from being just a simple tool to industry and finally being a driver of strategy. With each new tier value created from data increases.

While this master thesis previously discussed Demchenko and Membrey’s (2014) BDAF and how effective data-driven organization should build their infrastructure, another perspective introduced by Nguyen, et al. (2016). This view presents organisations differentiation based on the level of data processing and analytics use. They compare companies, which use in-database processing tools against traditional methods practitioners of extracting data from databases for analysis. The latter sets performance constraints, causing duplicate data and increase in chaos around data placement. Those mastering in-database processing, grasp valuable business benefits of streamlined processing and higher systems performance. Similarly, the same differences and similarities arise when comparing traditional to in-memory analytics.

Managing data in a valuable manner does not only include the right technologies and tools, but also broader organisational skills (Tian, 2017).

The organisational capabilities consist of management/corporate decision making (Davenport et al. 2012; McAfee and Brynjolfsson, 2012) and culture. Grant (1996) presents organisational capabilities as the outcome of knowledge integration – both complex and team-based. He stresses that the capabilities of a company depend more on integration mechanisms than specialist knowledge. The notion falls under the concept of culture - the acceptance of innovation and new capabilities, adoption of IT, and strategic outlooks (Niland, 2017). They are valuable, rare and imperfectly imitable (Barney 1986). Organisation’s management plays an important role in shaping its culture. Wamba, et al. (2015) emphasise that the top management is incremental not
only in an effective DDDM implementation but also in optimising decision models (Barton and Court, 2012).

Furthermore, BDAC comprises of all necessary capabilities: planning, decision-making, control, coordination and investment plan (Wamba et al., 2016). To become data-driven, leadership needs to build a culture where data is used in the service of constant improvement. It needs to manage the change for structural and technology improvements. These, in turn, would increase data usage within the company at all levels (Datnow and Park, 2014; Rijmenam, 2014; Harrison and Sullivan, 2011; Zeid and Davis 2014). Enterprises need to develop a questioning mindset, make decisions more explicit, and practice and learn every day (Barends and Rousseau, 2018). The lower the data literacy in the leadership, the fewer data will be included in processes (Mandinach and Gummer, 2013; Luo et al., 2015). It is vital that the management fosters and manages the six main mechanisms presented by Grover et al. (2018) as they mediate the linkage between capabilities and value creation.

In conclusion, this thesis has gone through introducing the main concepts behind data-driven decision making, covering also different aspects of data usage, business analytics and decision-making culture. Furthermore, these concepts were later adopted from the theories of RBV, DCT and EBM. As a result, four main capabilities were drawn out to have significant importance on building a data-driven organisation to create value. These were 1) leadership (human and organisational capability), 2) technology and data (physical capability), 3) analytics skills and data literacy (human capability) and 4) culture (organisational capability). The next part of the thesis will introduce how this thesis aims to measure the adoption of the capabilities in Estonian startups.
1.2. Methodology

The following chapter will introduce the framework, the key concepts of the research design, research aim and question, sampling technique and data gathering process, ethical considerations and analysis methods used. Based on the capabilities and processes a DDDM framework will be developed, and the most suitable research design will be chosen to evaluate the adoption of DDDM within the Estonian startup companies.

1.2.1. Formation of data-driven decision making framework

To define a DDDM framework, the authors of this thesis have analysed and introduced a new DDDM framework. It was built by mapping the five frameworks on the same level to develop a unified approach for deploying DDDM. The new DDDM framework will be the basis for this paper’s research. Considering that evidence-based management is about creating value from data, and one needs analytics for realizing insights which can be used to create value, the authors of this thesis have concluded a framework (Figure 2) which will define a good practice for management to be data-driven in their decision making, including capabilities needed to succeed.

Figure 2. Data-driven decision making framework. Source: Compiled by the authors of this thesis
1.2.2. Research Aim and Questions

The aim of this thesis is to assess the extent Estonian startup companies are data-driven by measuring the capabilities defined in the first part of this research. Based on the measures of adoption levels of different capabilities, the authors will combine the values to an equation, which will compute the data-driven decision making index. Using it as a basis, the authors of the thesis will compare researched companies with each other. The expected outcome of the analysis is to provide an overview of how data driven are Estonian startups, in light of the proposed framework.

1.2.3. Research design and data collection methodology

Mixed methods research has been chosen as the most suiting method to answer the research aim. Johnson, Onwuegbuzie and Turner (2007) define mixed methods as the type of research where researchers combine qualitative and quantitative methods to gain breadth and depth. Besides, mixed methods research draws from the strengths of both qualitative and quantitative research methods. In the study at hand, it allows expanding and strengthening of the study’s conclusions. As well, the authors of the research want to increase knowledge and validity of the research topic at hand (Shoonenboom and Johnson, 2017). By combining an online survey with semi-structured interviews, the authors of the thesis will triangulate results (Greene, 2008, p.14). The online questionnaire will allow to reach a wider audience, gain responses faster. The semi-structured interviews will allow deepening insights about the study at hand. Through development, results from the quantitative (QUAN) study will seek to inform the development of the qualitative (QUAL) part. It will then help to gain completeness of the study, where researchers will aim at bringing a broader view about data-driven decision making in Estonian startup (Bryman, 2006). By combining both methods, it will help to increase the credibility and integrity of the study.

Ethical consideration for the research

It is commonly accepted in the research of dynamic capabilities, that company managers are most reliable in evaluating complex issues on fixed scales (Laaksonen and Peltoniemi, 2018). Even though this is argued by Bertrand and Mullainathan (2001), that managers might distort the results for the means of trying to make their companies
look better. The authors of this thesis, this error is minimised by indicating the respondents that results will be anonymised. The confidentiality of the research participants was protected with the informed consent, gathered in two steps.

For the survey respondents, introduction letter was drafted to inform that the information gathered from the questionnaire will be anonymised and used only for academic purposes. The questionnaire was designed in a manner, where the respondents’ name was not mandatory. For the interview participants, informed consent was recorded during the interviews, where participants were informed that the information collected will be used only for academic purposes, and anonymised.

The Questionnaire

A 26-question survey (Appendix 1) was created on the bases of the created DDDM framework. In addition, some of the questions were inspired by the studies, conducted by Brynjolfsson et al. (2017) and MIT Sloan Management Review (2011). The extent of adoption of different capabilities within the companies is measured using different methodologies. Some questions in the questionnaire are inspired and adapted from the works of Kruschwitz and Shockley (2011), BI-Survey.com and Management and Organizational Practices Survey1. Likert Scale is used in the first part of the questionnaire to understand the attitudes towards a given object of analysis, which cannot be adequately measured by numerical values or certain characteristics - like principles of company culture and leadership, the performance of processes and attitudes towards strategy adoption. The authors of the thesis understand the limitations arising from the usage of this scale, meaning that based on the answers it cannot be concluded how much more or less the respondents are favourable to a topic, it cannot be concluded that the positions on the scale are equally spaced, and there remains the possibility that answers are made based on what they should be rather what they actually are (Kothari, 2004).

The Semi-structured interviews

---

1 BI-Survey.com – online reference to article: https://bi-survey.com/data-driven-decision-making-business
Semi-structured interviews were scheduled comprising of 12 questions (Appendix 3). They mimicked the areas, covered in the questionnaire (leadership, culture, data capabilities and technology). A total of eight interviews were conducted within a duration of two weeks. The goal of the interviews was to expand knowledge and depth, which had been limited by the quantitative part of the study. After both the survey and interviews were completed, integration through planned stopped points was conducted by the authors (illustration of the process below).

**Summary of methodological choices**

While being flexible and rich in information gathered, mixed methods contain certain issues. Primarily, lack of a single method to effectively list single typology of the research design (Shoonenboom and Johnson, 2017). To address that, the authors of this research took a simple research design (consisting of quantitative and qualitative) and continued constructing their design for the research questions to address research framework. As a result, the expected outcome of a hybrid design, which will benefit the research aims.

A concurrent timing method was chosen as the most appropriate for the study, taking into consideration the time and resource limits of the study, as well as allowing researchers to use the time most efficiently. Elements of the complex design were utilised, where both online surveys and semi-structured interviews were conducted at a similar time scope.

To integrate findings from the research, the authors used a framework to bind together the data sets (Creswell and Plano Clark, 2011). For this study, a QUAL + QUAN, the equal-status concurrent design has been chosen (Johnson, et al., 2017).

**1.2.4. The Research Sample**

Sampling methodology used for the purpose of this research is non-probability sampling due to the size of the full population and the characteristics of the available information. Even though usually for this type of sampling, items are deliberately selected by the researcher, then in the case of this thesis, the researchers have selected to exclude certain units from the entire population due to their characteristics. Even though this kind of sampling may give rise to bias and sampling error cannot be estimated exactly, the
conductors of this thesis believe by using elimination method based on relevant limitation of the population, they will not give rise to sampling bias (Kothari, 2004).

For the basis of the population to be measured, the authors have taken a list of startup companies publicly available on the Estonian startups database website, the list selected for basis was selected as sorted by company funding2. The reasons for the choice are twofold: authors’ personal and professional connection to the startup ecosystem as well as them being more open to new technology, such as big data.

There were altogether 587 companies in the list with limited information about their status, contact information, legal name and country of origin as of the research. To obtain contact details and legal status of the company, the authors of this research used the following sources - Estonian e-business register, CrunchBase database, LinkedIn and Google search. Based on finding 197 companies were eliminated from the list due to being bankrupt, removed from the registry, no connections to Estonia or no information available to understand the legal name of the project listed in the Estonian Startup database. In-depth research done on the available list for population formation based on actual activity raises the need for an up to date more detailed list of companies that can we considered as startups in Estonia. Besides, some of the companies contacted on the list stated that their company could not be considered as a startup. Unfortunately, the authors of this work are not familiar with the basis of the database formulation by Startup Estonia. Out of the remaining 390 companies, 152 were contacted. Even though the list consisted of almost four hundred companies, the authors of this research contacted these companies in two stages - during the first, and then the second email marketing campaign. After the second dispatch, it was observed that while the open rate for the email was relatively high (almost 50%), no responses were obtained. Due to limited time to carry out the survey and increase efficiency, it was decided to focus on the semi-structured interviews to gain more insights and not send out additional emails. From the 152 reached companies, 80 opened the link, 65 started filling in the survey and 33 fully completed.

The authors received feedback for nonresponses in-regards-to topic’s non relevance for the researched company (indicating they have no connection to data-based decision

---

2 https://www.startupestonia.ee/startups
making), questionnaire too long and heavy for one person to be filled in, or no time for the managers. Based on this feedback and the analysis of results, the authors are aware of the nonresponse bias in the results of the study, meaning that all generalisations made in this thesis are not relevant to the entire population, but rather reflecting the sample researched (Biometrics Advisory and Support Service to DFID, 2001).

The Questionnaire

As the results of the questionnaire will be anonymised, it is essential to define simple demographic characteristics of the respondents. For that, the authors have decided to measure the size of the company in terms of the number of employees, last financial year revenue (2018), and the industry within which the company operates. For the industry identification, the authors have not used any commonly known separation, but have listed major industries also leaving the option for the respondents to state most suitable for them.

The majority (48%) of respondents work in IT services/product industry, followed by software development (27%), financial services (18%) and other. In 2018 18% of respondent made less than 100 thousand euros for the financial year, 15% made between 100 - 500 thousand euros, most (35%) generating between 1 - 10 million euros, 23% made more than 10 million euros in 2018, and 9% decided not to answer this question. 79% of companies participating had up to 100 employees, the majority between 51-100 (33%), or 11 -50 employees (27%).

The Semi-structured interviews

To triangulate and expand on the questionnaire results, a series of eight (n=8) interviews were conducted during two weeks in March and April 2019. The authors used purposeful sampling technique as it added information-rich cases for study in depth. Patton (2002, p.230) explains the richness of information as “issues of the central importance of the inquiry”, thus allows for deeper insights and understanding. The criteria, by which the interviewees selected were managerial experience and position in a technology or IT company. The interviews were scheduled a week in advance and took place either at the office of the interviewee, a cafe and one of the interviews took
place via Skype due to logistical reasons. The duration of the interviews varied from 40 to 60 minutes.

All of the eight interviewees have an average of 11 years of management experience, leading teams from two to several hundred people. The interviewees represent some of the largest startups in Estonia, and also globally. Three of the interviewed people are CEOs, four are heads of the departments and one specialist, in the field of data science. Majority of the interview participants have business, economics, or finance backgrounds except one CEO, whose background is in social sciences (Company H). The representative of Company D holds a PhD in mathematics and is currently involved in creating a program about machine learning and computer-based statistics curriculum. Company E representative has the education and prior work experience in data science. We have chosen all of the people, who participated in the semi-structured interviews, to further enrich our quantitative (questionnaire) part with their previous education, experience and current work setting.

Table 3 Overview of interviewees. Source: Compiled by the authors

<table>
<thead>
<tr>
<th>Company</th>
<th>Position</th>
<th>Managerial experience (years)</th>
<th>Background</th>
</tr>
</thead>
<tbody>
<tr>
<td>Company A</td>
<td>CEO</td>
<td>10</td>
<td>Finance, economics</td>
</tr>
<tr>
<td>Company B</td>
<td>Head of the department</td>
<td>20</td>
<td>Economics, human resources</td>
</tr>
<tr>
<td>Company C</td>
<td>CEO</td>
<td>3</td>
<td>Business, marketing</td>
</tr>
<tr>
<td>Company D</td>
<td>Head of the department</td>
<td>10</td>
<td>Phd. in mathematics, data science</td>
</tr>
<tr>
<td>Company E</td>
<td>Specialist</td>
<td>1</td>
<td>Economics, cognitive and decision sciences</td>
</tr>
<tr>
<td>Company F</td>
<td>Head of the department</td>
<td>20</td>
<td>IT, business management</td>
</tr>
<tr>
<td>Company G</td>
<td>Head of the department</td>
<td>15</td>
<td>Economics, business administration, interaction design, analytics and data</td>
</tr>
<tr>
<td>Company H</td>
<td>CEO</td>
<td>8</td>
<td>Psychology, economics, management</td>
</tr>
</tbody>
</table>

Average managerial experience 11
1.2.5. The Research Process

To answer the research questions, the authors deployed a four-step research process. (see figure 3).

Figure 3. Research design process Source: Compiled by the authors

During the first step, an online questionnaire, comprising of 26 questions was created. To test the clarity and content, a pilot survey was carried out during three days. During the pilot study, authors received feedback regarding clarity of the questions, completeness of the answers and suggestions for clarity improvement in terms of wording. After applying the adjustments, the survey was sent to the rest of the sample by deploying two data collection methods – personalised emails and email marketing campaign.

During the same time, semi-structured interviews were organised. First of all, by contacting companies in the contact list (created by the authors) and then negotiating dates and times for the meeting. All of the interviews were recorded and later transcribed.

1.2.6. Data Collection

The Questionnaire

The first was the use of an email marketing system, to test the amount of open and click rate. An individual approach was chosen as complementary to ensure response rate, sending personal emails and other means of communication (social media and phone calls) to the recipients in the contact list.

In order to test the adoption of developed framework amongst the sample researched, the authors of this thesis have concluded that high adoption rate and the correlation
between the different capabilities needed for being data-driven (leadership, technology & data, analytical skills and cultures) needs to exist. Adoption of capabilities like the availability of measurement tools, data usage and volume, characteristics of certain processes and objectives are measured using multiple choice questions with both single or multiple answers. This part of the questionnaire is divided based on the capabilities measured into two: 1) strategy, leadership and culture and 2) data, analytics and technology. For single answer questions the authors have given each answer a value representing the adoption level of questioned capability within the company, for multiple answer questions the rate of response gives an overview of the multitude of the questioned processes/capabilities. See Appendix 1 for an overview of questions and the representative measures for answers. As some of the questions were based on Likert scale and multiple selection questions do not have always the same scale, answers gathered need to be set on a comparable level. For this the authors of the work will use standardization. The standardized values will be measured together for specific capability group and each item (question) in the group will be given a weight for summarization. As there is no specific underlying method or research based on what the authors are assigning the questions to the relevant capability, a simple factor analysis through principal component analysis was conducted on the entire set of questions, to understand which questions have the highest correlation and should be considered under the same capability. The outcome of the analysis (see Appendix 2) suggest having seven different components (Total Variance Explained Table in Appendix 2), with last three components including only one item (Rotated Component Matrix in Appendix 2). The authors’ of this thesis believe that the results of factor analysis for grouping the questions do not give expected results, as the groupings of the questions do not match with the initial assignation of capabilities by the authors; thus the grouping of questions for capability measurement is based on the authors’ personal categorisation. In order to group together different items, while keeping their own meaning for the compilation of a measurable index, the weights for each question will be evaluated based on qualitative input from literature and the authors’ own opinions. This method is also used to give the authors a better possibility to integrate the findings from qualitative research together with quantitative.
The Semi-structured interviews

The agenda for the interview guide comprised of 12 questions (Appendix 3), starting with broader questions about the professional background and education, as well as current work responsibilities. The individual interviews aimed to enrich and deepen questionnaire results, primarily focusing on similar structure (leadership, culture analytical skills and data literacy capabilities, technology and data). These questions allowed to ease into the subject. Then, then questions started from more general, simpler to answer and went into more complex. The interviewers used probing techniques to gain deeper insights into the topic.

1.2.7. Data Analysis

For the purpose of this thesis, descriptive analysis is mostly used together with the formulation of an index to measure the adoption of necessary data-driven decision making framework capabilities. Data-Driven Decision Making Index (DDDMI) is proposed by the authors of this thesis to be calculated as follows:

\[
DDDMI = \alpha \frac{\sum_{i=1}^{n} L_i}{n} + \beta \frac{\sum_{i=1}^{n} A_i}{n} + \gamma \frac{\sum_{i=1}^{n} T_i}{n} + \epsilon \frac{\sum_{i=1}^{n} C_i}{n}
\]

where:
- \( \alpha, \beta, \gamma \) and \( \epsilon \) – represent weights assigned to different capabilities
- \( L \) – leadership capabilities index
- \( A \) – analytical skills and data literacy capabilities index
- \( T \) – technology and data capabilities index
- \( C \) – culture capabilities index
- \( n \) – number of questions

As some of the questions have a different scale from Likert 5 point scale, the weights for each capability are calculated based on the maximum results for each capability to make their representation in the equation equal to each other. As the authors have concluded that all of the four capabilities need to be present equally in order to be data-driven, for that a maximum value for DDDMI was defined in this thesis as 4.82. The weights applied to each capability are as follows: \( \alpha = 0.2464 \), \( \beta = 0.2478 \), \( \gamma = 0.2464 \) and \( \epsilon = 0.2594 \).
The rest of the analysis is compiled using descriptive analytics to understand the items within each capability group, and the responses from the sample, mixed with insights gain from semi-structured interviews.

For analysis, all interviews were recorded and later transcribed (Appendix 4). After transcription, all interviews were anonymised using alphabetical order (Companies A, B, C, C, E, F, G and H), removing company sensitive information. After the transcription, data were coded and organised according to the four topics in the framework. To achieve collaboration and correspondence with the quantitative results, a deductive analysis was applied, searching for elaboration on the survey results. Then, the results were integrated at each step after the quantitative results to enrich and expand on the research theme (Bryman, 2006).
1.3. RESULTS OF THE RESEARCH

As the response rate is relatively low (n=33) and based on the demographics of respondents, we conclude that part of the whole startup population, those who are at early stages of business or not operating in IT or similar sector, are not proportionately equally represented in the results of this research. Individuals, who received the survey, shared that they did not answer the questionnaire as it was too detailed; the topic was not relevant to their business activity; or they lacked general motivation or time. Thus, the results of this survey can be used only to describe the sample which in nature mostly reflect those operating within the IT sector and who are on average generating more than 1 million euro per year.

The calculated average scores for each capability and overall DDDM Index based on three different demographics (no. of employees, company revenue and industry) are presented in the following Table 3:

Table 4. Overview of capability indexes by demographics. Source: Analysis results compiled using Tableau software by the authors

<table>
<thead>
<tr>
<th>Mean of capabilities and index by no. of employees</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of capabilities and index by no. of employees</td>
</tr>
<tr>
<td>1 - 10</td>
</tr>
<tr>
<td>11 - 50</td>
</tr>
<tr>
<td>51-100</td>
</tr>
<tr>
<td>101 - 250</td>
</tr>
<tr>
<td>251 - 500</td>
</tr>
<tr>
<td>More than 500</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mean of capabilities and index by revenue in 2018</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than 100k EUR</td>
</tr>
<tr>
<td>100 - 500k EUR</td>
</tr>
<tr>
<td>500k - 1 MEUR</td>
</tr>
<tr>
<td>1 - 10 MEUR</td>
</tr>
<tr>
<td>More than 10 MEUR</td>
</tr>
<tr>
<td>No answer</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mean of capabilities and Index by industry</th>
</tr>
</thead>
<tbody>
<tr>
<td>Financial services</td>
</tr>
<tr>
<td>Other e-services provider</td>
</tr>
<tr>
<td>Other IT services/product provider</td>
</tr>
<tr>
<td>Retail/wholesale</td>
</tr>
<tr>
<td>Software development</td>
</tr>
</tbody>
</table>
Based on the results, we can see that the higher the number of employees, the higher the DDDM Index, except 251 - 500, which is the lowest. This should not be looked separately from the rest of the results, as it only represents the values of one respondent. If we look at the index from the revenue point of view, we could say that companies in their early revenue stages are relatively data-driven, then if they enter into growth stage, the average index is slightly declining, while it crosses a limit of 10 M and starts to grow again. From the industry perspective, we can see that while IT related companies index varies between 3.4-3.85, financial services are in the middle with 3.51 and retail/wholesale company far behind others with 2.11, but as there is only one representative, this cannot be generalised on the whole population. The distribution of companies’ capability adoption and DDDM Index can be viewed from the following Figure 6:

Figure 4. The spread of sample within the indexes. Source: Analysis results compiled using Tableau software by the authors

From the results, it can be seen that while overall DDDMI distribution is quite similar reflecting both culture and leadership adoption, then the spread of results for analytic skill & data literacy and technology & data is quite significant. Based on previous, yet again, the spread is explained by the different industries of the companies, which are more or less adaptive to data-driven decision making concepts from the perspective of all four main capabilities. Due to the low volume of respondents and the similarity of the industry, the analysis of this research will be more descriptive. It will bring light on
the mindsets of Estonian startup community companies’ about data-driven decision making and the adoption of the main four capabilities outlined in the framework of this thesis.

**Leadership**

Figure 5. Mean results for leadership capability items by industry. Source: Analysis results compiled using Tableau software by the authors

<table>
<thead>
<tr>
<th>Industry</th>
<th>E-services provider</th>
<th>Financial services</th>
<th>IT services</th>
<th>Product</th>
<th>Retail</th>
<th>Wholesale</th>
<th>Software development</th>
<th>Overall Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avg. Management team focuses on data when making decisions</td>
<td>5.00</td>
<td>4.37</td>
<td>4.06</td>
<td>2.00</td>
<td>4.67</td>
<td>4.06</td>
<td>4.21</td>
<td></td>
</tr>
<tr>
<td>Avg. Managers encourage the use of data for everyday work</td>
<td>4.00</td>
<td>4.33</td>
<td>4.19</td>
<td>2.00</td>
<td>4.11</td>
<td>4.12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avg. Our company has clearly defined critical success factors</td>
<td>3.00</td>
<td>3.67</td>
<td>4.25</td>
<td>3.00</td>
<td>4.44</td>
<td>4.12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avg. Our company’s targets KPIs follow our company’s strategy</td>
<td>4.00</td>
<td>4.00</td>
<td>4.19</td>
<td>3.00</td>
<td>4.00</td>
<td>4.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avg. Our company has clearly defined strategic aims</td>
<td>4.00</td>
<td>4.00</td>
<td>4.31</td>
<td>4.00</td>
<td>4.22</td>
<td>4.21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avg. How many key KPIs does your company department follow?</td>
<td>5.00</td>
<td>5.00</td>
<td>5.21</td>
<td>5.00</td>
<td>5.22</td>
<td>5.21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avg. How often does your company follow the key performance indicators (KPIs)?</td>
<td>3.00</td>
<td>2.67</td>
<td>2.63</td>
<td>1.00</td>
<td>3.44</td>
<td>2.82</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avg. How does the management in your company value data-based decision making?</td>
<td>3.00</td>
<td>2.67</td>
<td>2.63</td>
<td>1.00</td>
<td>3.44</td>
<td>2.82</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As we can see from Table 5, the score pulling down the overall average for leadership capability is for most companies the number of key performance indicators (KPIs) followed, around 60% of the respondents only have up to 5 KPIs which are monitored, around 27% stated having between 6 - 20 KPIs and only 3% claimed to have more than 20, 9% of respondents stated it is not applicable to them. These KPIs on most cases are monitored on a quarterly (24%), monthly (27%) or weekly (27%) bases. Only 9% of the sample is monitoring KPIs daily, the same amount as those monitoring only once a year. The second item having a significant effect on the overall leadership index is management’s attitude towards DDDM. Around 42% of respondents state that strategic decisions are evaluated based on data from different sources. 21% bring out that data usage is mandatory in all decision making (the highest indicator for this item). For approximately 18% of the companies in the sample DDDM has either never been brought up and decisions are made on experience or management only looks at financial figures.

To get a deeper insight into the beliefs of the companies, respondents were asked to list the most important behaviours business leaders should advance to increase the effectiveness of analytics use within their organisation. Main behaviours brought out by the majority of respondents were fact-based decision making, understanding of the long-term vision, broad access to data and analytics and investment in education and
training. Only around a third believed that rewards and promotion of successful projects will increase the analytics use.

During the interviews, two-thirds of the interviewed managers and business leaders saw high value in evidence-based decision making, referring to an old saying “what gets measured, gets done”. The smaller and younger the company, the fewer decisions were not based on data due to insufficient information.

Effective deployment of analytics rose from the top management having prior experience in analytics or mathematical background. For instance, in one of the companies, one cannot even approach the CEO without having ideas backed up with data, as it would result in a personal disaster for that specific person. No data, no conversation with the manager. In this particular organisation, approximately 200 KPIs are tracked, starting from registration when entering the office to kitchen, product development and human resources. In other cases, efficient usage of analytics has derived from the product development team and learning from mistakes, that could be avoided if more evidence had been deployed. In one of the interviewed organisations “our new head of product, with the background of McKinsey, brings everyone to the table, bringing all knowledge and starting by asking "what is data? Usually, we still think that data is numbers and trends, but also user insights and more. So, bringing everyone to the table and having a healthy discussion encourages more usage of data” However, DDDM is not always spurred by leadership as brought out by the representative of Company D “we measure more not because of the management, but because the organisation is expanding and it would be harder to keep track of the changes”.

The key challenges to a successful deployment of analytics and data are rooted, according to interviews, in the business strategy and measuring the right performance indicators as the measurement is the trickiest part for setting correct KPIs to understand what metrics have the most significant impact on the business and which metrics provide the best company-wide perspective. In the organisations as mentioned earlier, teams are encouraged to learn from their assumptions, thus continually improve.
For cultural capabilities, we can see that the spread between responses to different items is smaller than for leadership. Interestingly, respondents on average measure their organisations as applying data-driven decision making with a score of 3.94 out of 5, which is relatively close to the overall DDDMI score for the entire sample (3.5). Data for 76% of respondents is shared with all stakeholders in the organisation, for 9% only with either with most managers or only senior managers and for the rest with most managers and employees. Openness to data-sharing for the sample is slightly above average. For the majority (52%) of respondents, company-wide figures are shared with everyone while the raw data itself is not accessible. For around quarter of the companies, data was accessible to everyone through back-office tools, the rest of the respondents saw data shared through the company network, based on special requests, or only accessible to specific departments. To gain an understanding of the cultural challenges for analytics adoption then main areas brought out were lack of skills to leverage data, lack of understanding of how to use analytics to improve business, lack of process ownership and management bandwidth to focus on analytics due to competing priorities. Less futile challenges brought out included an unwillingness to accept data or futile effort to deploy data projects.

Deeper insights gained from the interviews are in agreement with the survey results. Majority of the interviewees either had access to all data necessary for decision making or then would have it granted if needed. One of the companies mentioned that all data is available to access for the needed decision making, “everything is accessible by everyone and the data teams centrally support every other team in the company”. The biggest obstacles in deploying data are related to the mentality of employees, where basing decisions on data would empower a professional. Several respondents brought
out that in order to change that, the management teams strongly and actively need to encourage the use of data to establish an environment where data is valued. Interviewees said that data is a support tool to drive customer needs and that it is secondary to knowing how to ask the right questions such as “What do I want to impact or modify? Do we have enough data?” Then they can learn from data and make these actions. Once this cycle is in place, there comes the learning.

To be able to deploy data-driven decision making, a culture of data discovery should be present, where people would be enabled to access information quickly, and where the information would allow for inquiring more. From the practical barriers, is to how to enable people the access to the right information? Finally, when asked to which extent the companies deploy data-driven decision making practices, they agreed that to the extent, which helps answer and drive customer needs “we need to be customer driven, but data informed”.

**Analytic skills and data literacy**

While analytical competencies are considered as some of the critical skills to be data-driven in decision making, survey and interview, participants do not have many analysts employed at the companies they work. For 39% of respondents, there were no analysts’/data stewards’ employed and for 30% only 1-3 analysts. At the same decision making was reported for 48% cases to be either heavily or entirely dependent on data. For 42% of the sample this meant that descriptive (level 1) analysis is applied within in the organisation, and only 15% use prescriptive analysis (level 3), some respondents divided analytics usage based on decision type - for product usage of prescriptive analytics, and internally descriptive or predictive analytics. These were companies
measured above 3.8 on DDDM Index, having on average more than 5 KPIs, following them at least weekly or daily, and employing more than three analysts/data stewards.

Figure 7. Mean results for analytic skills & data literacy capability items by industry. Source: Analysis results compiled using Tableau software by the authors.

Note: Max value for questions 7 and 9 on the list is 4.

To gain a deeper insight into what was measured and for what, the questionnaire included two multiple selection questions. “What are your organisation's primary objectives for leveraging analytics?” showed that in most cases companies were looking at historical performance (76%) and input for changing/improving existing ineffective processes (76%). Not far behind were understanding of tactical actions (64%) and reducing the enterprise cost (52%). Insights gained are further in most cases used for real-time decisions (73%), acceleration of development of new products/services (70%), increasing of customer understanding (64%), and improving resource allocation (19%).

Analytic skills and data literacy were considered highly relevant by the interviewees. The need for business analysis team grew with the size of the company - the larger and more data-oriented organisation, the more prominent data literacy skill. In one instance, data literacy was not at all valued in the organisation, relying only on authority and gut feeling. As a result, it caused a lack of operational efficiencies.

One of the interviewed organisations explained that the business analysis team is ten people, but growing, as it needs to support every department in the organisation to make more evidence-based decisions. For a larger organisation, knowledge and data literacy was preventing from making more informed decisions. In the case of a company, building technology, in order to be hired in the company, basic statistical and mathematical knowledge was required: “We do not hire people who do not have strong analytical and mathematical skills. Our plans are for every employee to have the basic
knowledge of SQL in the coming future”. Controversially, with the exception of the company mentioned above, it was found that data literacy was centred only in segmented business departments. In said company, failures in decision-making were based on “where someone did not take data on into the planning as well as much as they could have”.

**Technology and data**

The last of the DDDM framework is technology and data. It revealed that despite having leadership and culture capabilities, the companies in the sample lacked using different sources of data and the variety of tools used was not that mature for a highly data-driven company, as the average technology index for those items on the questionnaire resulted below 2.5 on a scale of 4.

Figure 8. Mean results for technology and data capability items by industry. Source: Analysis results compiled using Tableau software by the authors

As many as 42% of respondents have up to three different data sources used in their daily business, 21% have between 4-6 and 30% use more than 10 data sources. Data in most cases is received from marketing (67%), finance (64%) and sales (61%) departments. Out of the sample, 45% responded by gaining data from business analysis/business intelligence department, slightly more than the ¼ state having a centralised database accessible to everyone, and only 9% added that they also use third party data sources. To understand which kind of tools are in most cases used for data processing for analytical purposes, it is seen that for Excel is listed 76% of times, followed by internally developed tools (48%) and BI tools (36%), there was no one in the sample who stated that they do not use any analytical tools. Even though one can see that companies also use a lot of more advanced tools than Excel, this study does not give an understanding of their full usage within the company. Meaning, understanding
whether they are used by the entire company or only specific departments. Based on answers we can see that a considerable proportion of the sample state having a significant amount of data which is either stored in a unified structured database (55%) or unstructured databases like Hadoop or other cloud-based providers (36%). Around half of the respondents also showed high confidence in having at least great deal of data available or all data available for decision making, and around 21% believed the data available for decision making was not significant.

In order to understand where companies were lacking in terms of data and technology, they were asked to select as many factors they believed were needed to increase their confidence in using data as the basis for decision making. Major problems brought out were veracity (64%) and visualisation (42%), but also improvements in terms of data volume, velocity, value and variability are expected, all stated around 30% of the time.

The insights, gained from the interviews, highlighted several key aspects, including lack of unified data warehouse and single-use platforms, which each serve a separate function, but are not integrated with each other. For example, in a company, which provides software and hardware, as many as 39 programs are daily used by the 200 employees. When asked how they are merged, it was informed that these are standalone programs, and on average, a professional has to use or log in to approximately ten (similar to survey results). This way, in agreement to the survey results, professionals in organisations use tools which are of personal preference. Another problem brought out by the interviewees was the usage of legacy systems with missing technology and data framework. Lastly, as already mentioned in the questionnaire results, veracity is one of the most recurring issues in applying data-driven decision making. According to Company H “last year, we had a machine learning project where we had more than 40 million data points about something. That was not enough”. With an increase of data received, primary challenges are in managing it, as perfectly put by the representative of Company D:

In quite a few aspects there's just a lot of instrumentation missing or it's a bit lacklustre and not fully reliable. We're also working to have a data warehouse up and a lot of the analysis just due to the size of what the tables are growing is getting quite unmanageable. So it can be quite tricky doing some quick analysis in advance and even post.
A more versatile sample of respondents would most probably change the overall picture of data-driven decision making application within Estonian startup scene. Currently more traditional, less technology driven companies and companies at their very early stages are not represented, and their results for the questionnaire and thus DDDMI could change the average completely as they on average have less resources in terms of investments and knowledge.
1.4. CONCLUSION

The aim of the thesis was to evaluate the extent of Estonian companies listed in the startup database being data-driven. To measure this, the authors of this research proposed a framework highlighting the main capabilities needed for being data-driven in decision making and an index (DDDMI) was compiled to calculate the adoption level of those capabilities. DDDMI was calculated based on responses gathered from a questionnaire. As this is the first research carried out using the proposed index for measuring the data-driven decision making of an organisation, the authors of this master thesis believe, that the index should be further tested on other industries, and further refined to make it more precise. The findings of this study should be considered in light of methodological and researchers’ centric limitations. Methodological limitation occurred during the quantitative research phase; the authors encountered a low response rate (33 answers) from the survey. As stated earlier in the research, it was due to relevance, complexity, language and other work commitments of the survey recipient list that resulted in the non-response. In future, it would be worth considering to localise the survey, based on the country the research takes place. Hence, the low response rate influenced the inability to generalise research findings. From the researchers’ side, limited access to data and time constraints were the most prevailing. Both the quantitative and the qualitative part of the research took place in a relatively short period of three weeks. For future research, it could be considered to extend the research time, to gain more insights and work on improvement of the contact list. Due to the low response rate, no conclusions on the entire population could be drawn. The authors of this thesis decided to include also qualitative research in the form of semi-structured interviews, in order to gain deeper insights about attitudes and processes in regards to data-driven decision making.

The authors of the thesis examined the main concepts behind data-driven decision making, addressing various aspects of data usage, business analytics and decision-making culture. These concepts were integrated using the theories of resource-based management, dynamic capability theory and evidence-based management. Four main capabilities were drawn out to have significant importance on building a data-driven value creation. They comprised of leadership (human and organisational capability),
technology and data (physical capability), analytics skills and data literacy (human capability) and culture (organisational capability).

The results, extrapolated from the research are not generalisable to the entire Estonian startup scene because of the specific respondent profile. The respondents of the survey and interviews did not represent all startup business segments. Instead, they were showing point of view of software developers, IT service/product providers and financial technology companies. The average size of the organisations varied from 11 to 100 employees. The revenue for the last financial year results in more than one million euros. It should be noted that startups, in search of a business model and preliminary service/product offering have a limited amount of data for analytics in comparison to those, which are in growth or maturity stages and have collected analytics from their operations.

The main findings from this research state that even though the companies had on average relatively high positive approach towards data-driven decision making within their leadership and culture, the actual application of analytics, technology and data, could be improved. Aspects of having analysts/data stewards employed or using more than just descriptive analytics to measure were just a few of the areas needing development. Data was mainly used by the respondents to measure historical performance to improve ineffective processes. Despite lower adoption of analytics, the insights gained from interviews still showed that the need for higher data literacy was in required, and fewer decisions should be made on gut feeling. This view goes together with what was argued by Lavalle et al. (2010), that in most cases emphasis is instead put on gathering data than on defining the questions, as more than 50% of respondents have over four different data sources, but less than 40% have more than 1-3 analysts.

Second important block of being highly data-driven derives from access to correct and integrated tools; and support from management for investments into those capabilities. Davenport and Short (1990) state that in order to leverage from IT companies need to consider the role of IT in the early stage of process design. For instance, as information is rather gained from different departments, Excel is commonly used for data handling. Often it is not that widely shared within the company. Only key KPIs are shared with everyone. Similar conclusions are drawn from the interviews. Company A’s data tools
are considered in the later stages of processes. At times, integration of new technology is restricted by legacy systems which are hard to change.

Almost all of the interviewees had the right mindset to be being highly data-driven in decision making. Yet, they also acknowledged that they still have a long way to go in terms of integrating technology, fostering data literacy and mentality. They agree that decisions should be backed with at least some sort of data. The main challenges were the fast pace of technology, growing data volumes and the challenges in changing the mentality of the people in charge of decisions (Jain and Sharma, 2015; McAfee and Brynjolfsson, 2012; Barens and Rousseau, 2018; Monino, 2016).

The study at hand is useful for academic researchers, as it provides insights into the degree of being data-driven and evidence-based in the IT-focused Estonian startup scene. The subject of DDDM had not been previously examined on such level in Estonia. The format of the study allows for further deployment of the same framework in different industries and countries. For example, extended to Latvia and Lithuania, while comparing research findings to test similarities and differences between evidence-based decision making in the startup scene. The empirical part of the study was based on mixed methods research, which in nature is versatile and flexible. From an academic perspective, it could be considered to split the research design into two parts: quantitative and qualitative, and research them separately. For practitioners and businesses, the research provides information and guidelines on how to compare whether the organisation is aspirational, experienced or transformed (Lavalle et al., 2011) in the use of data for decision making. For companies, striving to become data-driven, the framework, questionnaire and interview guide provide an easily applicable value for adaptation.
2. PROJECT MANAGEMENT

2.1. Defining and organising the project

The aim of the project management section is to introduce the reader with the master thesis project team, project scope, targets, risk mitigation and learning outcomes with future improvement suggestions.

2.2. Defining the team

The project team consists of two primary members, Kristiina Truuverk and Sigita Babarskaite. Kristiina Truuverk is the project lead, who is responsible for the structure and the intelligence on the project. Sigita Babarskaite is responsible for supplying information and deliverables, based on Kristiina’s input as well as coordinating the project.

Kristiina Truuverk (project leader)

Kristiina (further project manager) has a 10-year background in finance, with Bachelor degree in International Business Administration, Accounting and Finance. Her previous work experience includes work as an auditor in KPMG, and successful career in one of the most prominent Estonian startups Adcash, where she began her work as a financial analyst, soon changed her position to head of business analysis and then head of finance and business analysis. She is a strong business development professional, with a focus on Business Planning, Entrepreneurship, Venture Capital, Management, and Startups.

The project leader has decision authority in terms of the strategy of the content, literature review and the development of the framework at the project at hand.

Sigita Babarskaite (project team)

Sigita Babarskaite has over 10 years in business to business branding, marketing and communications experience in one of the oldest Baltic technology companies Hansab.
Her Bachelor’s degree is in Public Relations and Communications, the first MBA degree in Business Administration. She has founded the marketing and communications department within Hansab and is currently responsible for establishing strategy and plans for ten organisations within Hansab Group. Her area of focus is primarily on business communication, digital marketing and transformation, content creation, customer research, brand building and awareness, and branding in B2B.

The first step in master thesis preparation was establishing the project organisations, e.g. establishing the thesis team (further referred to as the project team). In this section please describe this process and introduce the team by answering the following questions:

Table 5 The project team and responsibilities. Composed by authors.

<table>
<thead>
<tr>
<th>Name &amp; Title</th>
<th>Main role(s)</th>
<th>Main expertise areas</th>
<th>E-mail</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kristiina Truuverk, Project leader</td>
<td>The strategy of the work, literature review content, framework development</td>
<td>Business analysis, Finance, Startups, Business planning, Entrepreneurship Startups</td>
<td><a href="mailto:ktruuverk@gmail.com">ktruuverk@gmail.com</a></td>
<td>Tallinn</td>
</tr>
<tr>
<td>Sigita Babarskaite, The project team</td>
<td>Editing of content, delivery of information to project leader, methodology, empirical research and presentation of results.</td>
<td>Customer research and market research, B2B branding, Digital marketing, Business strategy, Integrated Marketing Communications, B2B technology integration and service provision</td>
<td><a href="mailto:babarskaite@gmail.com">babarskaite@gmail.com</a></td>
<td>Tallinn</td>
</tr>
<tr>
<td>Meelis Kitsing, PhD. Thesis supervisor</td>
<td>Responsible for guiding and advising the project team. Providing feedback</td>
<td>Economic development, Strategic Planning Management, Consulting, Entrepreneurship, Business strategy.</td>
<td><a href="mailto:meelis.kitsing@gmail.com">meelis.kitsing@gmail.com</a></td>
<td>Tallinn</td>
</tr>
<tr>
<td>Marge Täks, PhD. Second thesis supervisor</td>
<td>Advises regarding academic expectations, rules. Reviews the work</td>
<td>Culture and society Pedagogy and didactics</td>
<td><a href="mailto:marge.taks@ebs.ee">marge.taks@ebs.ee</a></td>
<td>Tallinn</td>
</tr>
</tbody>
</table>
2.3. Defining the thesis parameters

The aim of the section is to show a clear overview, present Project Objective Statement (POS) of the project scope, implementation schedule, critical tasks and deliverables to complete the project on time.

Project scope

The project goal is to create a master thesis, focusing on data-driven decision making and its effects on the organisation. Then, based on the reviewed literature, to create a proposed framework and to test it against the Estonian startup community. The results are expected twofold:

1) To provide academic value about the levels of data-driven decision making and its effects on the organisation.
2) For the startups, which are going to partake in the research, to receive feedback on to which extent they are data-driven. Then, if they are, to which extent have they implemented data practices.

Project Objective Statement (POS): To write and successfully defend the master’s thesis in EBS about data-driven decision making in Estonian startups by the 5th June 2019.

Schedule

Table 6 Schedule for project management. Compiled by the authors

<table>
<thead>
<tr>
<th>No.</th>
<th>Task</th>
<th>Deadline</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>To select a research topic</td>
<td>01.11.2018</td>
</tr>
<tr>
<td>2</td>
<td>Select team member/s</td>
<td>05.11.2018</td>
</tr>
<tr>
<td>3</td>
<td>Initial thesis introduction</td>
<td>15.11.2018</td>
</tr>
<tr>
<td>4</td>
<td>Thesis seminar</td>
<td>15.11.2018</td>
</tr>
<tr>
<td>5</td>
<td>Selection of the thesis advisor</td>
<td>13.12.2018</td>
</tr>
<tr>
<td>6</td>
<td>Submission of the official thesis declaration</td>
<td>15.12.2018</td>
</tr>
<tr>
<td>7</td>
<td>Research of selected topic</td>
<td>04.01.2019</td>
</tr>
<tr>
<td>8</td>
<td>The bone structure of the thesis work</td>
<td>28.12.2018</td>
</tr>
<tr>
<td>9</td>
<td>Task distribution between project lead and the team</td>
<td>07.01.2019</td>
</tr>
<tr>
<td>10</td>
<td>Literature review draft completion</td>
<td>18.02.2019</td>
</tr>
<tr>
<td>11</td>
<td>Feedback from the thesis advisor regarding the literature review</td>
<td>20.02.2019</td>
</tr>
<tr>
<td>12</td>
<td>Changes for literature review based on advisor’s feedback</td>
<td>27.02.2019</td>
</tr>
<tr>
<td>13</td>
<td>Research design</td>
<td>27.02.2019</td>
</tr>
<tr>
<td>14</td>
<td>Development of online survey questions</td>
<td>15.03.2019</td>
</tr>
<tr>
<td>15</td>
<td>Testing of the survey</td>
<td>25.02.2019</td>
</tr>
<tr>
<td>16</td>
<td>Development of recipient list (in stages, due to large volume)</td>
<td>05.04.2019</td>
</tr>
<tr>
<td></td>
<td>Event Description</td>
<td>Date</td>
</tr>
<tr>
<td>---</td>
<td>-------------------------------------------------------------------------------------</td>
<td>------------</td>
</tr>
<tr>
<td>17</td>
<td>#1 Dispatch of invitation to participate in the survey</td>
<td>26.03.2019</td>
</tr>
<tr>
<td>18</td>
<td>Personal survey invitations by email, social media and phone</td>
<td>04.04.2019</td>
</tr>
<tr>
<td>19</td>
<td>#2 Dispatch of invitation to participate in the survey</td>
<td>04.04.2019</td>
</tr>
<tr>
<td>20</td>
<td>#1 reminder to the survey sent out on the 26.03</td>
<td>02.04.2019</td>
</tr>
<tr>
<td>21</td>
<td>#2 reminder sent to unresponsive recipients in the list from 04.04</td>
<td>09.04.2019</td>
</tr>
<tr>
<td>22</td>
<td>Development of semi-structured interview guide</td>
<td>05.03.2019</td>
</tr>
<tr>
<td>23</td>
<td>Conduct interviews</td>
<td>14.04.2019</td>
</tr>
<tr>
<td>24</td>
<td>Analysis of surveys and interviews</td>
<td>10.04.2019</td>
</tr>
<tr>
<td>25</td>
<td>Conclusion and discussion of the research project</td>
<td>20.04.2019</td>
</tr>
<tr>
<td>26</td>
<td>Feedback on the teamwork, work process to EBS team</td>
<td>25.04.2019</td>
</tr>
<tr>
<td>27</td>
<td>Pre-defence of the project</td>
<td>15.05.2019</td>
</tr>
<tr>
<td>28</td>
<td>The final defence of the project</td>
<td>05.06.2019</td>
</tr>
</tbody>
</table>

**Resources**

The main resources, required for successful completion of the project at hand are:

1) *Human resources* of the project team. Considering both people in the project team are employed full time, allocating work hours to manage the project on time.

2) *Technology and software*. Due to the nature of the project, the following technology shall be deployed: email marketing software Smaily, survey management system Surveygizmo and analytics platform Tableau to run the SPSS analysis. Google Drive, Docs and Sheets will be used to store information and keep real-time notes and work produced. The main communication channels during the project will be the Facebook messenger (for daily questions, answers and discussions).

3) *Academic literature*. To be able to create a framework for the empirical part of the project and to familiarise with the data-driven decision aspects, access to academic sources is necessary. These include EBSCOhost, Ester, ProQuest, EBS library and Estonian Academic Library.

4) *Space*. The team will be working individually from their selected venues, which include, but are not limited to: home office, work, library and coffees. Status meeting and development meetings will be agreed prior to the meeting.

### 2.4. Defining the project framework

To successfully complete the project on time, team members communicated regularly, openly and in an organised manner. There are two team setups: (1) the primary team,
consisting of Kristiina and Sigita, and (2) the advisor team, consisting of Meelis and Marge. The following table illustrates:

Table 7 Project framework for communication. Compiled by the authors

<table>
<thead>
<tr>
<th>Recipients</th>
<th>Frequency</th>
<th>Medium</th>
<th>Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary team</td>
<td>Bi-weekly</td>
<td>Phone conversations</td>
<td>Status updates, agreements on next plans.</td>
</tr>
<tr>
<td>Primary team</td>
<td>Based on demand</td>
<td>Facebook messenger</td>
<td>Current process discussion, file sharing, agreements on smaller next steps.</td>
</tr>
<tr>
<td>Primary team</td>
<td>Based on demand</td>
<td>Meetings</td>
<td>To agree on project strategy, develop questionnaires and interview guides.</td>
</tr>
<tr>
<td>Thesis advisor</td>
<td>Based on project deadlines</td>
<td>Meetings, email correspondence</td>
<td>Consultation, feedback, guidance for next steps.</td>
</tr>
<tr>
<td>Thesis consultant</td>
<td>Based on project deadlines</td>
<td>Meetings, email correspondence</td>
<td>To review academic integrity, structure and research methods.</td>
</tr>
<tr>
<td>Additional consultants</td>
<td>Based on demand</td>
<td>Meetings</td>
<td>To clarify research methods, to advice on selected qualitative and quantitative methods.</td>
</tr>
<tr>
<td>Testers</td>
<td>1 x per project</td>
<td>Email, social media and phone</td>
<td>To receive feedback regarding the questionnaire and interviews.</td>
</tr>
</tbody>
</table>

**High-level risk mitigation**

The project identifies the following key risks, related to the project:

1) Communication between the primary team

   - *Description:* Primary team members fail to communicate with each other for a mutually agreeable outcome either by not communicating enough or by not finding a common language.

   - *Risk management:* To find the time and to speak through in physical meetings to clarify all unclear points, which might occur. To ask questions about areas which might be misunderstood. To clearly define project parameters and expectations on paper and to regularly review them.

   - *Probability:* medium; Impact: high.
2) Clash of skill and competencies

- **Description:** primary team members (Kristiina and Sigita) come from different academic and professional backgrounds. Due to that, at times, there is a likelihood that one project team member might be the weaker link and not deliver the agreed outcome as expected by the other. Additionally, it might result in different expectations regarding the outcome of the final project.

- **Risk mitigation:** To discuss both of the partners’ capabilities and draw on each of the project members’ strengths and preferences. This would result in a more natural distribution of work as well as balanced results, drawing on each of the member’s strengths.

- **Probability:** high; **Impact:** high

3) Support from the advisors and consultants

- **Description:** lack of timely and sufficient support from the thesis advisors and consultants.

- **Risk mitigation:** to plan consultation times ahead, agree with the thesis advisor on a specific time for feedback. If needed, escalate to the head of the study department. As well, provide enough time for them to find time for response/feedback. In case of non-response email, to call.

- **Probability:** Moderate. **Impact:** high.

4) Survey and interview nonresponse

- **Description:** as one part of the project consists of sending out surveys to the selected amount of startups, there is a likelihood to encounter a nonresponse.

- **Risk mitigation:** To carefully review the list, to communicate value from the research, these organisations would gain and to send timely reminders.

- **Probability:** high; **Impact:** high.

**Project documentation and storage.** All project files are stored in Google Drive to have instant accessibility for all of the team members. Both of the team members maintain the order of information organisation listed below. Once information is updated in the files or comments added, the other team member is notified. Also, to ensure security and that all files are updated regularly, Google Docs are deployed, where:
1) The General Docs file. All links to other files are stored and meeting and agreement minutes are located. In the main file, also all literature review notes with comments are saved.

2) Project Management Docs file. In this file, all information is drafted for the project management of the file.

3) Empirical Research Docs file. The aim of this file is to store all academic literature related to empirical research, notes, suggestions, as well as a questionnaire and interview guide information.

4) Google Sheets. This file consists of the study framework with key aspects to each; a questionnaire with comments and values; startup contact list with corresponding information; interview schedule and completion status.

5) Google Drive. All academic sources, articles, books and other literature stored, so it is available to both team members. Besides, audio recordings and transcripts are saved from the semi-structured interviews.

### 2.5. Planning of the project

The project duration is from the 1st November 2018 and the expected closing of the project on the 12th June 2019.

#### Milestones

1) To confirm the thesis topic and advisor. 15th December 2018
2) To create a literature review draft by 14 February 2019
3) To consolidate survey and interview results by the 20th April 2019
4) To complete the thesis draft by the 22nd April 2019
5) To successfully present the thesis at the pre-defence on the 15 May 2019
6) To successfully pass thesis defence on the 5th June 2019

#### Analysis of resources

During the duration of the project, not all resources were distributed evenly and in the manner that should have taken place. The project had three larger waves of more focused resources: during the literature review, before the deadline; during the empirical research part, when it was time sensitive to collect contact details of the survey recipients and also to conduct interviews.
One of the key challenges to effectively managing resources was the parallel work (main employment) for the project team, and related meetings, reports and business travels. In addition, since the project manager had more experience and primary knowledge on the project, the project team (Sigita) had at times to cross-reference on actions, which resulted in more time on performing a task. Since the POS of the project was not possible to change, the project team decided to review the scope and optimise the workload.

The team applied three trade-offs adjusting the scope of the project, schedule and available resources. The scope of the project has been revised to fit the timeframe and available resources. Initially, the goal was to define the principles of data-driven decision making on a framework to develop, then test it in the Estonian startups, finally comparing the financial results and checking whether deploying DDDMcorrelates with company performance by running financial analysis. Furthermore, the four hypotheses posed at the beginning of the study (DDDM and growth comparison, growth over other startups, how management actions fall under the DDDM framework). After reviewing the scope, all of these elements were crossed out to bring more focus and improve the overall quality of the work.

One of the key learnings from the process was successful management of human resources. To be able to work as efficiently as possible, the team used a proactive approach. Instead of waiting for feedback from the supervisors, the team continued working on the next steps. For example, after sending out a part or a chapter of the thesis, such as literature review for approval and feedback, the team then gathered to discuss next strategic steps and plans for the coming part. This way, we managed to save time and use it efficiently. Similarly, at first, we wanted to conduct all the interviews together. However, after agreeing on specific times and dates, it was physically impossible for both of the team members to be present at the same time. We believe, that it helped us grasp more information and do a more thorough research.

2.6. Tracking and managing the project

The most effective elements of the project were related to clear task distribution and strategy discussion. For example, when programming online surveys and the email campaigns, one of the team members had to experience in doing so, hence, it was more
effective and additional time was saved. Having a clear project lead allowed for better development of the project, not having to spend extra time and effort on arguing about the concepts.

Like every project, this one is no exception and as well could be improved in the future. As the project lead was more knowledgeable about the topic at hand, Kristiina managed to work faster and more efficiently, whereas for the other project member, with less knowledge, at times it was hard to go ahead because of the information bottlenecks. For instance, reviewing the literature and finding the right guidelines.

Another learning curve is from the organisational side of the university part. The information and guidelines provided would slightly alter and be confusing. For the future, it would be best to agree and communicate better about what is expected from students, to avoid confusion. A critical aspect of any cooperation is communication and collaboration. When not receiving timely feedback and consultation from the primary thesis advisor, it causes havoc in the team, thus delaying work progress. In future, the university should carefully select available professors and only let students choose those who have the resources to supervise these projects.

The key learnings and takeaways for future projects are: (1) To choose a topic that is more understandable and the person writing is more familiar with (2) do not rush writing of the thesis project, if parallel commitments are taking a lot of resources, such as work; (3) To not choose a friend to work with and focus on someone who is not as close personally, but with whom he or she could share the same academic habits and values. Moreover, then (4) to communicate openly and timely with the partner, as the success depends on all parties in the team (project lead and the project team). As well as that, (5) to ensure enough time for the empirical part of the research, to be able to execute it with less stress and to obtain more quality results.

2.7. Assessing closing the project

Team assessment (Table 7) is adapted from a Forbes article, which displays 13 traits of a well-performing team. This project’s team have decided to apply a 5-point Likert scale to highlight project and cooperation areas.
Table 8 Team Assessment Characteristics. Compiled by authors

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Team Lead</th>
<th>Team Member</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recognising Individual Strengths</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>A Focus on Hitting Goals</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Alignment</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Open Feedback</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Integrity</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Kept Promises</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Interest in Learning</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Over Communication</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Psychological Safety</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Commitment</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Collaboration</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Trust</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Respect</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td><strong>Index</strong></td>
<td><strong>4.7</strong></td>
<td><strong>4.3</strong></td>
</tr>
</tbody>
</table>

**Project Effectiveness**

Table 9 Overview of project effectiveness. Compiled by authors

<table>
<thead>
<tr>
<th>Practices</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effective practices</td>
<td>• Online document sharing and storing</td>
</tr>
<tr>
<td></td>
<td>• Keeping team partner informed</td>
</tr>
<tr>
<td></td>
<td>• Clear distribution of tasks</td>
</tr>
<tr>
<td></td>
<td>• Ethics and integrity</td>
</tr>
<tr>
<td></td>
<td>• Understanding and support for the partner</td>
</tr>
<tr>
<td>Non effective practices</td>
<td>• Literature review shared between two people, where one was less knowledgeable</td>
</tr>
<tr>
<td></td>
<td>• Lack of communication from one team member at certain stages of the project</td>
</tr>
<tr>
<td></td>
<td>• Trying to apply common project principles from work into academic life.</td>
</tr>
<tr>
<td></td>
<td>• Reaching out to the thesis advisor by email.</td>
</tr>
<tr>
<td>Corrective activities</td>
<td>• At each stage, to inform our team member about the status and share as much as possible.</td>
</tr>
<tr>
<td></td>
<td>• The best way to get in touch with the thesis advisor is through calling and agreeing on specific actions</td>
</tr>
<tr>
<td>Suggestions</td>
<td>• For the project to be effective, to plan parallel work activities in advance as much as possible.</td>
</tr>
<tr>
<td></td>
<td>• To use Agile principles for development of the Master thesis. Work in small sprints, receive and exchange feedback and move ahead. Large deadlines are ineffective and paralysing to complete.</td>
</tr>
</tbody>
</table>
Proposals for future improvement

The master thesis project has been a challenging one. Considering the time constraints and general workload (in parallel work and school), the following proposals for future improvement are:

1) To not have classes in the last semester, when already writing the thesis. It distracts and adds extra workload. Instead, it would be better if classes took place before the actual writing thesis.
2) Focus on Research Methods. It is an important class, which will help to structure the research and understand statistical research methods. The handouts and lecture notes were used to refresh the project team’s memory and to get a better understanding.
3) To start thesis seminars as a project already from the first year of the studies.
4) To provide more support from the advisors and academic staff.
5) If needed, consult with other scholars in EBS. In the project team’s experience, all of the scholars have been glad to consult and assist.

Proposals for future improvement

All participants in the survey and semi-structured interviews were asked orally (during the interviews) and as well in the questionnaire, whether they would like to receive results from the research at hand. Those, who have indicated their interest, will be contacted to provide full access to the master thesis. In addition, transcripts of the interviews, together with the voice recordings have been sent to some of the interviewed professionals. Should the participants of the study require, the primary project team will consult and present results of the study face-to-face.
REFERENCES


Bennet, A. 2004. Alive with the fire of shared understanding: Implementing knowledge management in the Department of the Navy. In M. E. D.


Cai, Y. 2007. The impact of data quality metadata on decision-making. Boston University, USA.


Russom, P. 2011. Big Data Analytics. USA: The Data Warehousing Institute


# APPENDICES

## Appendix 1. Questionnaire and responding values

<table>
<thead>
<tr>
<th>No</th>
<th>Question</th>
<th>Method</th>
<th>Capability</th>
<th>Answers / Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Please insert the name of the company you work for?</td>
<td>Open-end question</td>
<td>Demographic</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>How many employees are currently employed in the company?</td>
<td>Multiple selection: single value</td>
<td>Demographic</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>What was the total company revenue for the last financial year (2018)?</td>
<td>Multiple selection: single value</td>
<td>Demographic</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Please select industry your organisation works in</td>
<td>Multiple selection: single value</td>
<td>Demographic</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Please select to which extent do you agree with the following statements regarding your organisational capital:</td>
<td>Likert scale</td>
<td>Organisational capital</td>
<td>Strongly disagree – 1 Disagree – 2 Neutral – 3 Agree – 4 Strongly agree – 5 Not applicable – 0</td>
</tr>
<tr>
<td></td>
<td>- As an organisation, we have open and honest discussions about data</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- My organisation uses data to uncover problems</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Managers encourage the use of data for everyday work</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Everyone in our organisation has a responsibility and right to make decisions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- We have a dedicated team that is responsible for data collection, management and analysis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Data is the basis for the need for new product/services/actions in our company</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Our company applies data-driven decision making</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Please select to which extent do you agree with the following statements regarding human capital in your organization:</td>
<td>Likert scale</td>
<td>Human capital</td>
<td>Strongly disagree – 1 Disagree – 2 Neutral – 3 Agree – 4 Strongly agree – 5 Not applicable – 0</td>
</tr>
<tr>
<td></td>
<td>- Most of the employees have necessary knowledge and skills to use data</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Using data has improved decisions in my organization</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- If we constantly analyse what we do and adjust, we will improve</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Management team focuses on data when making decisions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Our company has the necessary knowledge and experience to process and analyse data</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Please select to which extent do you agree with the following statements:</td>
<td>Likert scale</td>
<td>Physical capital</td>
<td>Strongly disagree – 1 Disagree – 2 Neutral – 3 Agree – 4 Strongly agree – 5 Not applicable – 0</td>
</tr>
<tr>
<td></td>
<td>- Our company has a well-established reporting system for sharing important data and analytics</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Our company measures and collects the right data in order to make right strategic decisions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Our company has the necessary tools and systems to collect, process and analyse data</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Data is the basis of revenue generation in our company</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Our company uses data to gain a competitive advantage over our competitors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Analysis in our company is used to create value rather than just look at historical patterns</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Data is easily accessible for analysis to everyone in the company</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Question</td>
<td>Type</td>
<td>Scale</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------------------------------------------------------------</td>
<td>-----------------------</td>
<td>----------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>To which degree do you agree with the following statements about leadership in your organisation?</td>
<td>Likert scale</td>
<td>Organisational capital</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Our company has clearly defined strategic aims</td>
<td></td>
<td>Strongly disagree – 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Our company has clearly defined critical success factors</td>
<td></td>
<td>Disagree – 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Our company’s main KPIs follow our company's strategy</td>
<td></td>
<td>Neutral – 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Who is aware of the targets in your organisation?</td>
<td>Multiple selection: multiple value</td>
<td>Organisational capital</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Please select all applicable responses</td>
<td></td>
<td>Only senior managers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>How many key KPIs does your company/department follow?</td>
<td>Multiple selection: single value</td>
<td>Organisational capital</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Most managers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>How often does your company follow the key performance indicators (KPIs)?</td>
<td>Multiple selection: single value</td>
<td>Organisational capital</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Most managers and employees</td>
<td></td>
<td></td>
</tr>
<tr>
<td>In your opinion, what are the most important behaviours business leaders should advance to increase the effectiveness of analytics use within your organisation?</td>
<td>Multiple selection: multiple value</td>
<td>Human capital</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Please select all suitable answers</td>
<td></td>
<td>Fact-based decision making</td>
<td></td>
<td></td>
</tr>
<tr>
<td>How does the management in your company value data-based decision making?</td>
<td>Multiple selection: single value</td>
<td>Human capital</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Understanding of the long-term vision for the organisations</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sharing data across traditional silos</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Distributed decision making</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Broad access to data and analytics</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Articulation of the value of analytics</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Evaluation of data for its enterprise-wide implications</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rewards and promote successful projects</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Investment in education and training</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bias toward experimentation</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Never – 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Yearly – 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Quarterly – 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Monthly – 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Weekly – 4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Daily – 5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>More than 20 – 4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Not applicable – 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Question</td>
<td>Multiplication:</td>
<td>Organisational capital</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>-------------------------------------------------------------------------</td>
<td>-----------------</td>
<td>-------------------------</td>
<td>------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>14 How is data shared within the organisation with other departments?</td>
<td>Multiple selection: single value</td>
<td>Organisational capital</td>
<td>Data is not shared with other departments, each department works on their own data - 1 Company network, where employees can share their documents, information and data based on special requests - 2 Company data is available only to specific departments - 3 Company-wide figures and data is shared with everyone but the raw data itself is not accessible - 4 Company database (back-office data) is accessible to everyone - 5 Other - Write In</td>
<td></td>
</tr>
<tr>
<td>15 What cultural/organisational challenges inhibit the use of analytics? Please select as many options as necessary</td>
<td>Multiple selection: multiple value</td>
<td>Organisational capital</td>
<td>Lack of understanding of how to use analytics to improve the business Lack of management bandwidth to focus on analytics due to competing priorities Lack of skills to interpret and leverage the data Lack of process ownership Lack of executive sponsorship Unwillingness to share data/knowledge internally Perceived costs outweigh the projected benefits Information governance is ineffective Ineffective efforts to deploy data/knowledge projects Unwillingness to accept data from other internal/external sources</td>
<td></td>
</tr>
<tr>
<td>16 How many analysts/data stewards are employed in your company?</td>
<td>Multiple selection: single value</td>
<td>Human capital</td>
<td>0 – 0 1-3 – 1 4-6 – 2 7-10 – 3 More than 10 – 4 Not applicable – 0</td>
<td></td>
</tr>
<tr>
<td>17 Where does data for decision making come from? Please select all suitable answers</td>
<td>Multiple selection: multiple value</td>
<td>Physical capital</td>
<td>Finance department Business Analysis/Intelligence department Sales department Marketing department Other departments Unified database accessible to everyone Third - party sources Other - Write In</td>
<td></td>
</tr>
<tr>
<td>18 How many data sources does your company use in everyday business?</td>
<td>Multiple selection: single value</td>
<td>Physical capital</td>
<td>0 – 0 1-3 – 1 4-6 – 2 7-10 – 3 More than 10 – 4</td>
<td></td>
</tr>
<tr>
<td>19 What best describes the availability of data to support decision making in your organisation?</td>
<td>Multiple selection: single value</td>
<td>Physical capital</td>
<td>Data to support decisions is not available - 1</td>
<td></td>
</tr>
</tbody>
</table>
| 20 | What best describes the use of data to support decision making in your company? | Multiple selection: single value | Human capital | A small amount of data to support decision making is available - 2
A moderate amount of data to support decision making is available - 3
A great deal of data to support decision making is available - 4
All the data we need to support decision making is available - 5
Not applicable - 0 |
| 21 | Which kind of analytics are mostly used in your company everyday business? | Multiple selection: single value | Human capital | Decision making does not use data - 1
Decision making relies slightly on data - 2
Decision making relies moderately on data - 3
Decision making relies heavily on data - 4
Decision making relies entirely on data - 5
Not applicable - 0 |
| 22 | What are your organisation's primary objectives for leveraging analytics? Please select all suitable answers | Multiple selection: multiple value | Organisational capital | Make real-time decisions
Improve resource allocations
Reduce enterprise costs
Increase customer understanding
Increase employee productivity
Accelerate development of new products/services
Identify new markets
Manage enterprise risk resources
Improve supply/demand chain performance |
| 23 | In my organisation, Insights gained from analytics are used for: | Multiple selection: single value | Organisational capital | Changing/improving existing ineffective processes
Measuring historical performance
Gaining an understanding of tactical actions
Evaluation methods for different options
Gained insights are not applied
Other - Write In |
| 24 | What would help you gain more confidence to use data when making decisions? Please select all suitable options | Multiple selection: multiple value | Physical capital | Amount of collected data needs to increase More timely data Diversity of internal data types (structured, text, audio, video, image etc.) Data collected is relevant to business Consistency in the quality of the data More precise data Data should be more structured for further visualisations There is already enough accurate and easily processable data for analysis and visualizations Data from more external sources I am not provided with data to support decisions |
| 25 | Which kind of data processing tools does your company use? Please select all suitable options | Multiple selection: multiple value | Physical capital | Ms. Excel or similar software - 1 BI tools for visualization (Tableau, Qlickview, Power BI etc.) - 2 Statistical analysis tools (SPSS, R, Matlab and similar) - 2 Internally developed tools for data management and processing - 3 All of the above - 4 We do not use any data or analytical tools - 0 |
| 26 | What is the data volume produced in your company? | Multiple selection: multiple value | Physical capital | Minor levels of data mostly from financial aspects - 1 All data can be handled in Excel or similar software - 2 Structured data warehouses with average data volumes - 3 Significant amount of data, which is stored in a unified structured database (SQL or alternative) - 4 Significant amount of data, which is stored in an unstructured database (Hadoop, cloud based solutions or similar) - 4 Not applicable Comment for measuring: value for answer given by highest level selected, if both significant levels selected, value of 5 given for results |
### Appendix 2 Factor analysis in SPSS

<table>
<thead>
<tr>
<th>Component</th>
<th>Total</th>
<th>Initial Eigenvalues</th>
<th>Total Variance Explained</th>
<th>Rotation Sums of Squared Loadings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>% of Variance</td>
<td>Cumulative %</td>
<td>Extraction Sums of Squared Loadings</td>
</tr>
<tr>
<td>1</td>
<td>15.037</td>
<td>45.565</td>
<td>45.565</td>
<td>15.037</td>
</tr>
<tr>
<td>2</td>
<td>3.075</td>
<td>9.135</td>
<td>54.685</td>
<td>3.075</td>
</tr>
<tr>
<td>3</td>
<td>1.981</td>
<td>6.002</td>
<td>60.686</td>
<td>1.981</td>
</tr>
<tr>
<td>4</td>
<td>1.765</td>
<td>5.167</td>
<td>66.853</td>
<td>1.765</td>
</tr>
<tr>
<td>5</td>
<td>1.535</td>
<td>4.552</td>
<td>70.705</td>
<td>1.535</td>
</tr>
<tr>
<td>6</td>
<td>1.264</td>
<td>3.629</td>
<td>74.535</td>
<td>1.264</td>
</tr>
<tr>
<td>7</td>
<td>1.074</td>
<td>3.255</td>
<td>77.789</td>
<td>1.074</td>
</tr>
<tr>
<td>8</td>
<td>.966</td>
<td>2.829</td>
<td>80.717</td>
<td>.966</td>
</tr>
<tr>
<td>9</td>
<td>.933</td>
<td>2.846</td>
<td>83.563</td>
<td>.933</td>
</tr>
<tr>
<td>10</td>
<td>.904</td>
<td>2.437</td>
<td>86.000</td>
<td>.904</td>
</tr>
<tr>
<td>11</td>
<td>.728</td>
<td>2.105</td>
<td>88.205</td>
<td>.728</td>
</tr>
<tr>
<td>12</td>
<td>.583</td>
<td>1.706</td>
<td>93.973</td>
<td>.583</td>
</tr>
<tr>
<td>13</td>
<td>.494</td>
<td>1.497</td>
<td>94.470</td>
<td>.494</td>
</tr>
<tr>
<td>14</td>
<td>.469</td>
<td>1.421</td>
<td>95.891</td>
<td>.469</td>
</tr>
<tr>
<td>15</td>
<td>.464</td>
<td>1.323</td>
<td>94.115</td>
<td>.464</td>
</tr>
<tr>
<td>16</td>
<td>.374</td>
<td>1.132</td>
<td>95.247</td>
<td>.374</td>
</tr>
<tr>
<td>17</td>
<td>.363</td>
<td>.919</td>
<td>96.166</td>
<td>.363</td>
</tr>
<tr>
<td>18</td>
<td>.270</td>
<td>.517</td>
<td>96.683</td>
<td>.270</td>
</tr>
<tr>
<td>19</td>
<td>.154</td>
<td>.502</td>
<td>97.570</td>
<td>.154</td>
</tr>
<tr>
<td>20</td>
<td>.160</td>
<td>.455</td>
<td>98.115</td>
<td>.160</td>
</tr>
<tr>
<td>21</td>
<td>.154</td>
<td>.446</td>
<td>98.561</td>
<td>.154</td>
</tr>
<tr>
<td>22</td>
<td>.112</td>
<td>.309</td>
<td>99.870</td>
<td>.112</td>
</tr>
<tr>
<td>25</td>
<td>.060</td>
<td>.151</td>
<td>99.646</td>
<td>.060</td>
</tr>
<tr>
<td>27</td>
<td>.031</td>
<td>.084</td>
<td>99.991</td>
<td>.031</td>
</tr>
<tr>
<td>28</td>
<td>.024</td>
<td>.074</td>
<td>99.993</td>
<td>.024</td>
</tr>
<tr>
<td>29</td>
<td>.008</td>
<td>.025</td>
<td>99.996</td>
<td>.008</td>
</tr>
<tr>
<td>30</td>
<td>4.193E-16</td>
<td>1.273E-15</td>
<td>100.000</td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>2.78E-16</td>
<td>8.34E-16</td>
<td>100.000</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>1.55E-16</td>
<td>4.72E-16</td>
<td>100.000</td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>-7.34E-16</td>
<td>2.25E-16</td>
<td>100.000</td>
<td></td>
</tr>
</tbody>
</table>

Extraction Method: Principal Component Analysis.
<table>
<thead>
<tr>
<th>Component</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>We have a dedicated team that is responsible for data collection, management, and analysis</td>
<td>.99</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How many analytically skilled employees are employed in your company?</td>
<td>.77</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data is the basis for the need for new product/services/offers in your company</td>
<td>.74</td>
<td>.39</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>What is the data volume produced in your company?</td>
<td>.74</td>
<td>.39</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Which and how analytics are most used in your company's business?</td>
<td>.57</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Our company uses data-driven decision making</td>
<td>.50</td>
<td>.41</td>
<td>.31</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How many data sources does your company use in everyday decision making?</td>
<td>.54</td>
<td>.47</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How does the management in your company value data-based decision making?</td>
<td>.54</td>
<td>.47</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Using data has improved decisions in my organization</td>
<td>.51</td>
<td>.49</td>
<td>.34</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Which kind of data processing tools does your company use?</td>
<td>.75</td>
<td>.33</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Have you completed a data analysis that we did not expect and thus we will improve</td>
<td>.69</td>
<td>.35</td>
<td>.32</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Our companies uses data to gain a competitive advantage over our competitors</td>
<td>.59</td>
<td>.58</td>
<td>.51</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Management team focuses on data when making decisions</td>
<td>.51</td>
<td>.58</td>
<td>.40</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>What best describes the accountability relative to support decision making in your organization?</td>
<td>.51</td>
<td>.40</td>
<td>.34</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Our companies has clearly defined strategic aims</td>
<td>.51</td>
<td>.40</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Our company has clearly defined political success metrics</td>
<td>.72</td>
<td>.40</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>As an organization, we have open and honest discussions about data</td>
<td>.36</td>
<td>.50</td>
<td>.60</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Our company has a well established/existing system for sharing important data among departments</td>
<td>.61</td>
<td>.54</td>
<td>.55</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Everyone in our organization is responsible and right to data</td>
<td>.61</td>
<td>.54</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Our company main KPI's include our company's strategy</td>
<td>.56</td>
<td>.52</td>
<td>.38</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Managers encourage the use of data for everyday decision making</td>
<td>.42</td>
<td>.36</td>
<td>.25</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data integration uses data to uncover problems</td>
<td>.51</td>
<td>.48</td>
<td>.69</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data is easily accessed by all stakeholders in the company</td>
<td>.43</td>
<td></td>
<td>.71</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Analysis in our company is regularly used, rather than just look at historical data.</td>
<td>.32</td>
<td>.37</td>
<td>.37</td>
<td>.34</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Our companies has the necessary tools and systems to collect, process and analyze data</td>
<td>.32</td>
<td>.37</td>
<td>.37</td>
<td>.34</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>What best describes the use of data to support decision making in your company?</td>
<td>.32</td>
<td>.37</td>
<td>.37</td>
<td>.34</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Our company measures and collects the right data at the right time to make decisions</td>
<td>.35</td>
<td>.40</td>
<td>.44</td>
<td>.45</td>
<td>.75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Our company measures the necessary knowledge and experience to process and analyze data</td>
<td>.39</td>
<td>.40</td>
<td>.44</td>
<td>.45</td>
<td>.75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Most of the analytics we use is on the web</td>
<td>.31</td>
<td>.37</td>
<td>.47</td>
<td>.47</td>
<td>.71</td>
<td>.50</td>
<td></td>
</tr>
<tr>
<td>How often does your company use the key performance indicators (KPI)?</td>
<td>.35</td>
<td>.40</td>
<td>.44</td>
<td>.45</td>
<td>.75</td>
<td></td>
<td>.72</td>
</tr>
<tr>
<td>How many key KPI's does your company use?</td>
<td>.35</td>
<td>.40</td>
<td>.44</td>
<td>.45</td>
<td>.75</td>
<td></td>
<td>.72</td>
</tr>
<tr>
<td>How is data shared within the organization among departments?</td>
<td>.35</td>
<td>.40</td>
<td>.44</td>
<td>.45</td>
<td>.75</td>
<td></td>
<td>.72</td>
</tr>
<tr>
<td>Data is the basis for important decisions in our company</td>
<td>.35</td>
<td>.40</td>
<td>.44</td>
<td>.45</td>
<td>.75</td>
<td></td>
<td>.72</td>
</tr>
</tbody>
</table>

Rotation Method: Varimax with Kaiser Normalization.
Rotation converged in 6 iterations.
Appendix 3. Interview guide

Verbal consent
All of the information, gathered from the interview today, will be anonymised and will stay confidential. Are you ok with us recording the interview for the purpose of academic research?

Verbal Consent was obtained from the study participant
Consent was NOT obtained from the study participant

Background information
What is your professional background?
What is the position you currently work in? What is your experience in the field?

Interview Questions
How would you define data-driven decision making?
In your opinion, how important is data in strategy making?
What is the overall feeling about data management and deployment in the organisation?
What are the main organisational challenges when adopting data-based decision making? What are the major challenges do you perceive in implementing data-driven decision making?
Are decisions, made in your company, based more on experience and intuition or more on evidence? Please elaborate.
How would you describe the decision making process in your company?
How is data gathered for decision making? Do you think that if you had considered more options, instead of rushing to a decision, you would have received a better outcome?
How do you share information in the organisation?
How do you analyse your decisions? How do you measure the outcome/results of your decisions?
How do you assess if there is a place for improvement in decision making?
To which extent, do you think, your organisation is data driven?
In your opinion, what does it take to be successful when deploying data-driven decision making?
Appendix 4. Interview Transcripts

Transcript Company A

Author: Right now, we're on tape and I'd like to ask the same question again. Is it okay if I record this? Do I have your consent to later use it for academic purposes?

Yes.

Thank you very much. So, let's start. Perhaps you could share a little bit about your background and what do you do?

My background.

Yes, your background.

And then my background to the company or how far up?

Your professional and academic, as far as you want to go.

Okay. I have a master's degree from Estonian Business School in strategic financial management. I graduated back in 2000 and then uh before that I did my bachelor's also in EBS uh International Business Administration. I did both specialities. Uh, this was the marketing management and accounting and finance. And I also during my study spent a year uh studying in Bocconi in Italy and did uh for a full year there as well. Uh. Different subjects uh professionally started. Uh, I did my professional career so to speak. Although I had some jobs before uh back in 2008 working uh as an analyst uh for a boutique consulting company. Uu and moved to strategy consultant to associate partner. And in the five years, I spent there that I moved more towards finance. Um, I did a year um like freelance consulting, mainly financial M&A and things like that. And then I spent and after that, I spent a year in Deloitte managing their financial advisory services in Estonia. The team there. And then for the last four and a half close to five years, I've been here at Company A. Uh first as the chief financial officer of the uh the whole group, actually. And then uh since last September I took over as the CEO of Company A whereas remaining the CFO of the group.

How many people do you employ at the moment at Company A?

Currently has around 84 people together in the group.

Where are they located?

The headquarters is in Tallinn Estonia. We have about uh 70 people here. And then we have offices in Sofia Bulgaria Barcelona Spain in Sofia we have around 10 and then in Barcelona four at the moment I think.

Could you define what Company A does?

This is a technology company. Focused on Internet advertising. So, what we do is we connect advertisers to their advertisers' campaigns to the target audience globally online anywhere.

So could you say that to an extent you're an I.T. company?

We're a technological company where we have more developers than the salespeople. So we through that also we define ourselves as a technology company. And to me personally, Company A is a company that it's a technology company has chosen to deal with internet advertising at the moment. It could easily be doing something else with the same resources and the same people that we have here today.

As a technology company. So how many developers do you have?

We have, I have to lie again. I think we have around 35 developers. The product development team has 35 people and other support functions. So, it's maybe a bit less. And then the other support functions in finance and marketing.

Do you have a data analysis team?

We have a business analysis team.

How many people do you have?

Currently, we have three people who work in the business analysis team.
I have to ask all of these questions. So.

And so I just hope I give the right answers.

There are no wrong answers actually the point. So how in your opinion is important data in strategy making?

Well, I said I've grown up as a strategy consultant that data is everything. Well, the decisions should be based on data. So, uh. It is practically impossible to make the right decisions time after time without relying on accurate data. That's what differentiates many of the global leading tech companies from other tech companies is that their decisions are based on data.

I think in your experience in your career especially here in Company A, do you feel that decisions are made more on analytics, data or knowledge whatever you receive are still in some cases you have to rely on experience and intuition?

I think it's both. In a way, how it is today for Company A it's been difficult to get the answer because we've come from very much being relying on intuition on this some data too. Ever since I took over I have kind of stated it out loud that I want people to have data-driven decision making and data plays an important part. Now I think the golden point there is that you take that data from the market which shows you a picture of a particular time and place or from a particular angle. You must also put your experience on how you analyse data or how you see that data. And that's kind of where you probably define as the gut feeling comes to come into play as well. So for me, it's kind of the uh in order to make a decision, I've always kind of thought to myself that I will try to make the best decision based on the experience that I have and the data that it has been provided to me at that point. And then I never have to hesitate or never have to regret any decisions because those decisions have always at that moment that I've been the best in the company.

And speaking about you mentioned that historically the company had been relying more on intuition and it brings me to the next question which is the culture. How do you transform how do you reflect on the culture? How do you change it? So people would be more how to say on the same page as you are?

I think it starts with success cases. Well, I think nobody no one can argue that the numbers don't lie. Basically, you have an opinion and then you put the reality of the numbers next to it and then you understand what the actual situation is. So it's the first step to do is, first of all, to start measuring different aspects of the business and setting in place the different KPIs or OKRs or what not. And then collecting that data and actually start then not only about collecting and analysing that data, but you need to also make people use data or move them towards that. And that's uh that's kind of what you need to do. And then you combine that already the work with the first successful cases where you say "okay well that's your opinion. This is what data shows based" on the data we would like to do this and that. And then you're going to see the outcome and then you start implementing this more and more and practice going forward as well.

But then isn't it so that people have to be enabled and they have to know how to use it?

Yeah

How would you assess the situation at the moment and Company A?

Well, first of all, I guess we the business analysis team in that is not that old either. So, we started kind of doing that and started collecting and then kind of putting that thinking into place and somebody who was visualising some parts of it. That's what we've always had in Company A is we've been collecting vast amounts of data. We have more than 1 petabytes of operation and data in different tables because we show on the best days or months we show more than 20 billion ad units per month and we collect for each ad unit that we display, we collect more than 50 parameters of data. So, it's kind of that's how you kind of differentiate this. This is how you differentiate whether it's only operational data or business operation or do we talk about the company operational data. And for me, they're kind of combined with one another still at the end of the day. And we used to collect a lot of data that we used only for maybe operational purpose. Meaning that in our industry we have to have an algorithm that predicts what is the most profitable place or where does the ad convert the best for the user or which type of user. And we use that data for a for prediction partially and this is kind of where we used it more. Most of the time it was still going off based on experience that the decision making and go moving forward. And now what
we're trying to do is to bring all the data that we have visible to the different stakeholders in the organisation allow them to have access to that data.

And because we are now implementing the Agile management principles it also means that the information shouldn't be pushed up to me as a CEO and then I shouldn't be making that decision and forcing the results down.

But what I'm more trying to do is also push the decision making to the level of the people that have the data and have the know-how because they will be able to make faster and better decisions than I.

Because I would all be sharing information through a filter and I should have my task today is to give them the direction where we want to go our goals talk about them company wide and they need to reflect how their heart or work fits into that plan and kind of what decisions do they need to make based on the data that they have as well. For me, I need to create and facilitate an environment where they can flourish the best and have the best tools to achieve the best results.

**How many numerical KPIs do you have in the company?**

It's a good question. So what we have defined is uh there's three companywide KPIs that follow. It's the Active clients. A number of active clients. Average earning spending per client because we have two types of clients and then an average lifetime of a client. So, everything that people do in Company A in on whatever different level needs to through is one of those three cases. So those are the company if that's the thing then we have different cases that go on to them with different departments at the individual level as well. How many do we have? Numerical depends on quite a bit on the manager of that particular department as well. So they can assign their work. For sales targets but they don't have just one target they have a target. What is the acquisition they need to do what is their normal portfolio or total portfolio? And then what is a stretch objective for them as well for different product aspects. We're measuring different things as well. It's very difficult for me to give you a particular kind of one sure no because they're not listed in one file. And I've noticed that people also tend to be their own KPI is in their own overview of things. So, and in addition to that what we have is a bi-weekly meeting where we actually run by the KPI file. We call it the KPI File. different data is put together and kind of we go through the main trends and the main changes of that data. Everybody kind of bi-weekly what has happened in the short term to get an understanding and that file is then shared with everyone in the executive team.

**So if I understand correctly there is one file and all of the KPIs are stored there?**

No, it's there's one file built for the executive team for instance just to understand which way the world is turning etc...So and then every two weeks the executive team in their executive meeting goes through that kind of the main trends and the changes in that file. And then that file helps to summarize and follow some trends and then visualize them etc. as well. But then there are department specific dashboards for instance sales have their own dashboard, both supply and demand side. There are different marketing dashboards that we use. And then those dashboards are, for instance, I have access to all of those dashboards that have been created in Tableau for instance. And then they are shared with the people that could make use of that information as well.

**But there are separate access rights to different levels?**

Yes, there are separate access rights to different levels.

**And do you have one central database where you store all of that all of this information? Let's say marketing has one sales has another database.**

You mean infrastructure wise.

**Yes.**

It's a bit complicated today. Basically, what we have to talk about the database is it's what do you define as a database. What we have is for operational data. We have to have two different databases that are a different level of aggregation than for business analysis purposes. We have a separate copy of a database that's kind of also aggregates again a different type of data on a different aggregation of level as well. And initial two and we're a process of improving those set up as well. Then if you think about it the accounting software that we're using is also a database. It's a financial database and the results etc. as well then whatever we get from Google it is another data source but it's also database because it saves the marketing activity there as well. So in that sense, we have several databases today. They're not all pulled
together in data warehouse today although we're moving more towards that direction to have a bit more unified data set that was what I said we're still in kind of transition from the more or less information based to more decision making.

**So what are the main challenges in this journey?**

There are different challenges.

I think that first and foremost is of course with any change it's changing the mentality of the people involved as well, proving that it makes sense to put effort into this and make it happen.

We are today close to 12 years old. We've been doing this for twelve years we have been for most of the time we've had the start-up mentality which means things need to get done no matter kind of how they get done. Which also means that the infrastructure wise earlier or data wise it can be quite a bit of a mess with different tables, different setups as well.

So there are some technical limitations coming from legacy things that are in place today. So if we haven't been building the organisation with a focus on data from day one you can imagine that it can be quite too big a mess to start sorting it out. And then even on top of that starting visualising it using the visualization tools on top as well. So that's a limitation that we have faced as well but where I think we are almost past that. And then in up to the fact that to find the right people to be part of the team in the business analysis side or then the people building the database for the infrastructure behind.

So far you mentioned here you said first of all mentality then also legacy systems and also throughout the history of how the company has worked and also finding new people who would fit into fit the right profile of the company you're building. Still from all of these what do you feel is the biggest challenge?

It always starts with mentality. Because, well, at the end of the day. We don't need to have the best database in place, we don't need to have the best visualization tools in place. We need to have the right attitude of the people who understand that. This is where the benefit lies and what should be done.

**And in this case at the moment and the transition phase do you feel that people have already shifted towards believing, seeing value?**

Yes, I think people are moving towards it but there's still work to be done. We're definitely taken. We're a long way from the start but there's still a long way to go.

**Where does the information for decision making come from? Is it more internal or do you purchase really parts as well?**

Um, it depends.

Because uh like for salary reports and salary decisions we purchase reports from the local market, wherever offices are situated uh for but mostly I would say 90 per cent of the Internet is generated or it comes from there.

We have also operationally speaking kind of for serving of ads and prediction we have been thinking about also purchasing data from outside about the potential uses for profiles or etc. We have at the moment is managing all internally.

So this usually internal information is customer based in the internal information as kind of customer based information internal information as we have uh maybe we are generating information ourselves and then based on that we are making decisions.

**Do you think at the moment people are enabled with the knowledge and skills to be able to use this data because as you said you produce so much data in the company? It's quite impressive. And then most of it is internal so it's accessible to you already. Do people know how to use it?**

The operational data, definitely. We have a very good of like a database. You can go quite granular to make inquiries. It becomes more difficult if you want to have specific tables or a specific point to point a point of view. For that, we have the business analysis team the dashboards or make sure we don't duplicate anything and then shares the information where necessary. And then, of course, it depends on the need if the person and in which department, as well as the data, needs also differs from quite significantly for sales or operations. It's a bit different for finance it's kind of different although it might be overlapping for support functions in a different way. And what I would say uh people have access to it more and more. Actually, that's one because we have been collecting a lot of data but that doesn't necessarily mean that it's data that we actually need to make use of. Because for instance what we focused
on the first quarter is also putting now place in place data hooks to collect onboarding data customers something that we really haven’t been measuring before. You can measure all the tweaks in terms about performance but we have never really measured about kind of what is the process from registration to spending the first euro or dollar and where it or the drop-offs and what's happening there as well. So we're getting better and better in finding exactly that data that is important for our business and set that in place.

*Do you in your personal experience have a case where you remember because of lack of time or perhaps you didn't have enough data at that given time you had to make a decision and now looking back if you had more data you feel that it could have been a better one?*

I think from my personal perspective from Company A I said I will always try to make the decision that I believe is best for the company. At that time with the data available to me I can't go back and feel sorry about the decisions I've made. Of course, if I would do that I know this is there probably would have been some decisions that could have been better if I had more data or something like that. But uh I couldn't stop the business from developing further just because I am lacking something at that time. I will do whatever I can to make the best decision at that time. I can always change it down the line or most of the time I can change it down the line if I've been wrong and no one is at all right all the time and. But sometimes it's more important to move forward and understand that you've made a mistake or collect data along in the process as well. And then adjust a little and learn for the future as well. And this is also our experience and that data come together in that sense.

*Do you feel that at the moment there is still room for improvement in decision making?*

Yes.

*What kind of data do you as a CEO feel that would be necessary for you to make better decisions?*

Okay. Me as the CEO or being a CEO means that every problem in the organisation is your problem. Which means that all types of data could be more valuable to me. And it starts from the interaction between two people or the cooperation between different departments or some business decisions as well. So, it's very difficult to say exactly kind of what could I use more today because it depends on the problem I am solving at that time. I think that the best thing I can do is establish an environment where data is valued. We know we need to collect it and people understand that and do that in their department. And if I am able to create that environment and the understanding of this, it would start kind of producing or kind of developing further on its own as well. I don't need to make that decision for everyone or push everyone in that direction. End of the day I will have more data to make better decisions. No matter what they make.

*So are you trying to say that you're trying to flatten the organization?*

Uh well, I'm from a Scandinavian business culture background so that is quite flat by itself. This is something we have been working on as well, to make it flatter between the top and the bottom. But at the same time as I mentioned because we're implementing Agile management principles decision making is often pushed to the level where they have the information and they can make that decision. So, for me kind of as a CEO, My role is to create an environment where people can flourish and get the maximum or deliver the maximum from their position. And because I'm taking great care of that on the grassroots level this will have a positive impact also to my decision making with the data that is presented to me as well.

*How far do you think you are on this journey of being data-driven?*

What's the ideal setting? I am a robot and I know all the information in the world? I don't know what's the ideal set up as well, because I work in the technology sector and things might change drastically over the years. I would say that uh we're moving towards the right direction. I can see we're implementing it but I think there is a long way to go still.

*Based on the on already the time since you've been the CEO and you've been transitioning towards that. Do you feel positive effects of that mindset that you've been trying to establish and the environment you've created?*

I think so. I think people are making smarter decisions. Uh I think uh people are understanding the importance of data and uh everyday work as well. I think at the end of the day it comes down to at the moment. I think we're playing the long-term game and it's maybe it's a bit too early to say. I can see the
first signs of it. I can see the success cases that we've had but I'm also thinking that I think I will see the success in the future a lot more than I’m seeing it today. It takes time.

What's stopping from the faster implementation of this that data-driven decision making so in terms of challenges the mentality the mindset but the skills of the people the ability to use the data for the right decisions. Do you think about how have you already thought about training? Are they necessary?

So what we've done is basically introduced the training budget within the organization. So, every person has their own uh personal training budget for a year that they can use for. Whatever type of training that they associate with their position or the development of their position within the organization. Um yes, I think quite often it's important also to teach about how to use the tools that we have to get the information. When onboarding new employees, we do quite a thorough training sessions about going to the back office that we used and how to use that and what information you can see and people learn a lot on their own as well. And we uh tried to introduce also kind of I know that the business analysis team is also when they get in dashboards, they're also introducing me to the people for how to use it etc. as well. And if they sometimes don't, they should be doing it. And that's kind of... that helps definitely with it as well. At the end of the day it's uh I think it's a lot about communication as well. And uh showing how successfully these tools can be used actually. And the managers need to lead by example in doing that as well. They don't pick it up. There's uh little chance that their teams would pick them up as well and use them.

Thank you. How do you analyse the outcome of a project?

For projects or product development. That's what we do is uh during the planning process we say establish KPIs and the benefit hypothesis that we want to reach. And then uh once the uh development is done then after a certain time, be it one month, three months or six months,1 year we should be measuring and uh the outcome. And if it's kind of meeting the outcome or form product development perspective um well we started this Agile framework implementation last year and then we were now past our second planning process. So from a product development perspective, it's built in. Also, the retrospectives and uh. And the constant relentless improvements also in the processes and in kind of what we do. And that part. Uh from a management perspective how do we evaluate it? It's uh it's basically the three KPIs company wise. It's at the end of the day, financial results, revenues, profits they don't like to me that's numbers don't bullshit. So, um I'm using those as well. And then it's um well I guess also the uh uh smiling faces I see around the office because I can make really shitty decisions and people can hate to work here as well but that's what not the goal, the goal is to have a kind of a full package together.

I've asked some questions that we have developed to get together with Kristiina. Is there something else you’d like to share that maybe I haven’t asked regarding data-driven decision making.

I think I've covered most of it.

So I would like to thank you very much. And this concludes our interview. So thank you.
What could you tell me a little bit of what everyone does?

We have collected two very different fields. We have production. We don't define ourselves as a production company but we are developing two different technologies. One of them is software. And they have to work together. So we have different products combined together by hardware and software. We are monitoring our services all over the work. So we have different developing. Departments. And also monitoring and production and whatever we need finance and marketing and sales in which countries are you located.

Where does your company operate?

Yeah. Or our headquarters in Berlin. And we have office in telling us where we are based in this film. We have partners. Different partners. All over there all over the world but not our production or office abroad.

All right. Thank you. And so let's start with the data questions.

So I would like to ask how you understand it and how you define it. What is data driven decision making?

Oh, a really hard question. I guess it's all kind of data you can find, what kind of numbers you can find. If we talk about a HR. And then what kind of money we are paying pronghorn we think what. How much we spend to find people. So everything you have you can analyse.

In your opinion, is data just numbers or more?

Of course, they are connected with decisions later on. What to do. What kind of budget you have to spend to do you have a desired outcome etc.

How important is data and strategy development?

Crucial I guess.

Would you like to elaborate, why?

If you have only the gut feeling then I guess it's not the precise method of analysing something. you have to have some kind of data about I don't know if will explain it to somebody who is not in the field or a manager or a team you have to have the statistics analyse and it can be done only by my numbers, using the others data.

What is the overall feeling in Company B on about deployment of data and the usage of data?

What is Company B on like when it comes to usage of data? We are analysing our candidates in different fields. One of them is we are measuring their self-opinion via test. Their personality and they were uh knowledge. So, this is one kind of data we are measuring. We have defined our values and culture. So, we can measure if the candidate have the same values and same culture we appreciate. So I don't know if I answered correct.

Yes. My my question is more like in Company B. Is it like I am encouraged to use data to make decisions for example or in general to use data to mean something. Or is it more based on OK, we have very little time so we need to make a decision fast let's do it based on our gut feeling experience what not. Or is it more encouraged from the management side in every every department. Let's say to be more data oriented or evidence based?

Yeah I guess it's more and more evidence and analytics based. I can't say it's management encouragement. As organization is growing you have to have some kind of procedures some kind of rules some kind of data gathering, such as budget. You can't have the budget if you haven't analyzed previous year or quarter. So. Yeah I guess we are. Analyzing and gathering statistics and doing decisions based on that. And even if we do quick decisions are not they are based on some kind of data. For example, we gather CVs and we analyze the data out what's there. So education professional skills whatever.

Thank you.

In your personal experience and what you see in the organization what is the what are the biggest challenges to encourage employees to use more data base decisions as opposed to intuition. What are the main obstacles when implementing more data which would help to make better decisions?
Knowledge how to do it must be the first one. Then you have to have some kind of tools like programs and knowledge. Again how to use them and also you have to see that it helps. You may have to actually see the benefits and quite fast to be able to use it.

Do you feel Company B on the moment everyone is enabled in a sense that they have knowledge tools and they see the benefits of using it?

More and more so, but not all the time and not in every field. So we have for example have 39 different platforms. We're using two manage data in everyday life.

How are these programs talking to each other?

They are not. That's the problem.

All right. That's interesting. So you have the thirty nine programs and how many do you use daily yourself?

I have to say I have not counted them but less than 10.

Could you tell me. Okay so you have thirty nine programs. Can you describe what type of platforms are these?

Some of them are just basically information exchange. So it's like Slack, Skype, emails. Then there are warehouse and financial platforms like SAF, Directo and we have Confluence, Solidwworks. We have a class and we have SolidWorks. Information on what we send out for the government. All kinds of statistics. Can't recall all of them, but these are the examples.

Since you have many platforms and lots of data based on what I understand and so you have different roles within the organisation. Is the information really accessible to everyone? Does analytics or are there levels. How how is it organized?

You have access solely where you need to have. For example I have access to emails, Slacks, Skype, Atlassian, Confluence Jira, Telem. Directo, soft, Google Drive. all documents and policies and and uh whatever the general information is based on the. Intranet. And so we have many different levels but it is the depends in which department you are working and what kind of tools you need.

Who makes the decisions to which access someone can have?

Direct manager

Is it effective the way the roles to access to information is distributed?

Not always, but as we are not a very big company yet and we have quite small teams and this is the purpose to keep them quite small. And we would like that manager can answer all questions for team members. So, the first person to turn to is the manager. Who knows every policy, every guideline? Then it's good if your manager is able to answer and give you access to wherever you need it. And also, can groove/approve you. We have also one other desktop tool. I didn't know that there would have. So when somebody is moving. Either position or other. Department then we have in H.R. we are entering the data of that and all other departments can do their own. Things. They have to do to give him or her right tools right. Access. Access. And clothing etc..

What is organizational structure in Company B?

We try to keep it as flat as possible.

Do you feel in your again it's just your opinion for academic purposes but do you feel that decisions are made on the higher level based on gut or based on actually real data?

At the very top I guess they are more like based on gut feeling. But if you go down then you have to analyse the information. If management is telling you that this year we have to have revenue 100 million euros then that's it. We have to think what kind of action we have to take to bring it home.

So where do they get this target? Is that their like opinion that just opinion and experience or is it based on the previous performance of the company, external trends on the market? How do you think is it is it done?
To be honest I don't know exactly but I think that they're all different. It's a mix about history and gut feeling. So some kind of finalize the data they are analyzing and some kind of opinion they had in their stomachs.

*Thank you. How do you gather data for decision making? For example, you mentioned in your field that's mostly based on job applications right. But in general, how is data is it more internal data that is gathered from the customers or do you purchase from databases?*

It depends very much what kind of people we are speaking about. But we are gathering information about how our employees are feeling. We are how surveys ask how people are doing what is well and what is not well and we are doing it once in a year in a big scale. It is very formal and all the time all these same questions. And it is not gathered here but in university of Tartu. And yes, so we have a very big package of information we can get it back but we are doing small surveys every week. Once a week I’m sending out a questionnaire with one question and it can be whatever. How do you feel your week was what was the main reason? What was the main purpose in this week. And then people can give you feedback. What was who then what was bad and you can always step in if something is wrong. This is in the field of HR but if we are talking about for example the quality of our products then we are gathering information from outside from our customers and also from suppliers. How they are doing their work? What was good, what was bad? What kinds of technology or material was better called and what was the pay or what worked well? So whatever information we can gather we will do and we will analyse it.

*And where do you keep it? Is it stored in a separate environment for HR? Or is it integrated with the rest of the data which would then be somehow put together as a consolidated information to make better decisions?*

In H.R. we don't have one consolidated database. They are different databases and for example if we had this big survey what I was talking about then I'm doing come summary just some kinds of fires weather where we are heading. What kind of situation is right now and we can always measure what was last year and how it is today. So it is like a text versus numbers analyzed and combined in one.

*Do you make these recommendations to the department?*

Yes.

*And then how they process? They receive recommendations either presented probably by you or by email or however what happens next?*

Yes both. I will do this summary and send it out and then we’ll have these meetings based on department or small groups and me and the direct manager awaking of all their findings as well.

*And in your feeling how do you feel that this input is actually really taken serious data which then will be used for some changes and improvement or is it more like a soft recommendation and then they still decide what's best*

As we have suggested these things and managers use this information to create actions. For example, last year we found out that we have 6 percent of our employees who don't feel that they are engaged rarely with a company and we were suggesting that we will say that whoever feels so they don't have strong engagement they can switch the department. Or they can have volunteer assignments if they want to. Feel that they belong to the company in that way and if they don't then they have a possibility to get two salaries and go away. And we did it every manager said this is a good idea. And some people left.

*And just a question for you personally in your work is there has been probably situations when there’s a lot of decisions to be made that the same day or in a very short period of time and you make them based like you have to do in hindsight do you feel if you had more options or more information you would have made a different decision and perhaps better?*

If I need more information than I can always say that 'hold on, I will research the subject'. I don't have to do it by minute. I have always some time to analyse.

*You’re not a person who has made quick decisions because there was no time. You’d rather you would analyze different options?*

Even if they are big decisions, then I can always analyse. But based on experience you can do smaller and analyse it's pre-done. You already know so you can get a more precise answer. And you don't have to wait to analyse it depends. I didn't have a rush with decisions.
So you don't have experience when you rushed into a decision and if you had had more information it might have been a better decision. It doesn't apply to you?

Not now, but in my previous work life I have felt and and uh.

Is there any specific example that you could share?

Yeah. Mm hmm. It was connected to the legislation and it was not in Company B but in my previous employment and uh I think that case was pretty simple. We had to dismiss one person, but the information was not shared with me. And I uh I made my decision based on what I had. And uh it was not the best decision. So um if I make more inquiries than I could have made a decision. What kind of legislation to use.

What kind of information, what kind of data analytics or knowledge are you currently missing the most in your work to make decisions or to make better decisions?

I guess I would like to speak more with our owner and top manager to see how he's making his decisions and what kind of data he has. I don't think that my work quality will wise from that. But that I know better how to respond to his needs.

I understand. Okay but in general in other departments you don't feel that if you had more you know trends on the market or any kind of type of data internal or external it would make your decision making easier? Or perhaps even more effective.

I have to say I have quite a lot of information. What kind of statistics which comes from outside. Well we are buying salary surveys and the trends on the market. We know pretty well our customers.

When you say you know pretty well your customers are. How do you know them? How do you get information about the customers?

As they visit us quite often you learn about our machines and technology. Also, we have surveys. And feedback from the end users of our products.

And that information is stored also in a platform or is it also in different places.

Do you feel it would benefit if it was somehow consolidated or organized and displayed an easy to understand way?

Yes, definitely. But I have to say that maybe somebody already consolidates all of this information but I don't know about that because this is is one of our weaknesses that I don't know exactly what everybody's doing though.

So my question is then how do you share information with within inside of the organization?

As I said it's a different survey. And I also said I think so based on light that graphics and verbal texts in the meetings.

What about Slack and other channels? Is all about. It's like only for certain departments or is like the Universal Channel that you have.

Universal channels as well as meetings where different departments gather together and all these protocols are put in internet so everybody can see what was the subject and what kind of information was shared. Yeah. Of course there are many things we could do a better share.

Could you just give a couple maybe examples from your experience now what could be the ways to improve information sharing? Is it based on technology or is it based on human capital?

Absolutely human capital. Those we have as I said so many so many ways how to gather the information and share information. And we are working towards make agreement that we. Don't use so much different platforms. We're gathering all this information something like 10 platforms, so we have tie all of these platforms, so you could find information in one platfrom, from the otehr.

So who is this working and whose decision is to make that initiative to improve information sharing?

We have shared responsibility I have to say the management different departments heads together. As well as we are planning to go to the stock market. Then different needs come from there. So work in progress in this field.
How do you analyze your decisions. Usually I don’t know if you could only speak about each organization and in general and how do you measure the outcomes of these decisions?

Once again you have different platforms. For example we ask from our candidates, who have applied for the position, after four months we ask how was then all of the procedures about how they feel, what was good. What was good what was bad. And we ask to score us, so we would know which field we need to improve.

Then as I said these personality tests we have data from there. We can see if the person was a team player, as this test said he or she will be then we will follow our candidates’ outcome work wise as they do their homework before they come to the company. Then we can see was is done by them or somebody else. Maybe. And as I said survey is a great service.

Is there any room for improvement in the decision making across all of the organization?

And of course, all the time of the time.

And in which direction do you think that this decision making should be improved? from the top or from the down or?

I haven’t thought about it. But as you asked previously and now I think that we have quite a lot of information already but we haven’t consolidated it together. So we have to do this one first and then gather more information.

Do you mean do you think that employees in your organization have the skills and the knowledge how to use this data for their benefit?

I guess they have. Of course, they can’t go to every platform we have. They don’t need to do that. But it can be done better. Always.

How data driven do you think your organisation is from scale one to 10?

To be honest, I really don’t know. I don’t know how to answer because I would need to know what information others are using. OK but if I compared it with my previous employer who was big company a corporate company then the information they gathered was on a very large scale. And in many cases it was unnecessary.

Why?

Because we gave the same answers all the time via different platforms. In many cases we didn’t see the end users and what kind of information they really used. So it was data gathering, but in many cases never analysed.

What does it take to be successful to be data driven or to use evidence based decision making in an organization?

Probably have to ask "What do you want to establish? What kind of problem we would like to solve? What’s the answer to my question must be answered? Based on that, you can ask the right questions and you get right answers.

Thank you. Is there anything else you would like to share?

As the subject in these words and this kind of angle is new to me, I have to say I had to process everything you asked and I think it made me think that we have a lot of work to do with our data.

In your opinion is data so necessary to be effective or evidence?

Yes, of course it is.

How quickly are decisions made?

It’s fast.

What enables fast decision making?

Uh. We are not afraid of failing it even if it has a cost.

But is that a feeling across only the management because you’re still in the management position or is it across the whole organization across the board is that. How is it communicated to the employees?
That failures are allowed. If it brings us one step forward. And do something what nobody hasn't so far. Then it's necessary to fail. We can't go further if you don't fail.I guess financially people they are more driven by than that they are they are giving the answers that's why because so they have to think numbers wise all the time.

*Do you have any data analysts in their organization?*

Yes, we have them in the Finance department.

*How many do you have?*

Two right now. And we have some as I said we have a quality control people in every department. They are very much data based as well. And as I said we have platforms like solidworks. They are also administrators who are validating the data and we have you a document secretary. Let's say who is responsible for the documents are they in the right format and are they adequate all the time. So they are like quality and data people as well.

*Is it okay that I'm using this interview for academic purposes and anonymized and do I have your content that I've recorded this and I can use this?*

Yeah. Thank you.

So perfectly ok that you are recording.

*Thank you so much. I got an I wanted to ask.*

---

**Transcript Company C**

*Let's start. I would like to ask a couple of questions before I do I have to follow the protocol and inform you that all of the information will be anonymized and stay confidential. Are you OK with us recording this interview and later using this information for academic purposes?*

Yes.

*Great. Could you tell me a little bit about your company and what do you do there and what the company does.*

Ok. I'm CEO and founder of Company C. We are changing the recruitment. We're innovating the recruitment process the way that you don't look for a job. So as, as a one pager I said job I'd like one picture but you. You as a company representative you are posting or looking for a job with a video job offer.

Let's say you state like we need let's say as a company needs a Russian speaking person in this for disposition and you are saying in your resume that you are Your Russian is excellent. But in the video you can really show it off. You know. you can ask like let's like let's say company posts to you like three questions to answer and they say that for one question please answer in Russian and you as a company representative or recruiter you can directly see if they really speak Russian or not. You know this kind of samples. So you can understand. You can see the person who this person is you can understand.

Who they are as a personality I think you are there from their body language like let's say if they are an energetic person or they are not so energetic person so on. There are so many things that you can see from the video so and why this video is so important.

Instead we have concentrated on talk with each our managers recruiters from I think around 10 different countries and then very very wide range of companies say things like this one person company ten person company are like a hundred thousand person company it is like in the U.S. one company had president of human resources we were in contact and we are still in contact they are hanged as an exciting thirty five thousand people a year thirty five thousand people you had only hiring so it means they have like over 100000 employees in all over the world.

And why-why it's so important is that even all of these recruiters and H.R. managers we have met and the top two or not met but in we escape or calls or whatever and that everybody is saying that they don't make a decision if to hire a person or not before interview if it's Skype in three or it's really interviewing
you are meeting in the office so and and what they are looking from there or from many maintains had these how this person looks like who how they are talking what is their personality what their attitude you know what how energetic they have and one actually few recruiters or H.R. manager or business owners leaders have said.

A good word I'm using in my meetings as well is they are looking for sparkling people's eyes. You know if they have this part manpower or not and this you cannot read from the resume.

So that's the main lake purpose we receive in this recruitment field and we see that it's totally changing you know five years ago maybe you only checked if they had this degree if they have this very key story.

Okay I'm on page one but now it's moving more and more the way that the companies want to hire you know this good person as this personality and everything around this person and they are teaching you know specific skills so knowledge in-house. So it's it's totally you know this way you is changing like totally in for the opposite side you know they don't grow so much anymore like liking it so now we have a.

I just so like you know it's so so well and bringing out what we are thinking and what we are seeing in the future is that you definitely know what I noticed. It's like a brewer company. They are pretty they're like a mix exploding right now you know we're on the way that they're approving different drinks alcohol and non-alcoholic drinks like a small Brewer and they are very. They have a very good brand and very strong life and brand communication everything and they hire like assets for its CEO position 21-year-old person. No. Yeah.

So it's so good except when you know they are just looking more and even for a CEO position you know they are looking for this park and this power of this positivity. You know what. Whatever this company needs to do to get things done. And they they owners our shareholders who are like experienced more like all the people they are ditching the CEO you know they just need this power or the sparkle in the eyes and they just itching you know do this do this to this like that.

Use these tools or use this I don't know Excel files like that you know to do and to use the data to use the feedback or whatever is this special task you know and this is a very good example that's showing that it's moving felt that this were in five years ago with we couldn't even imagine that somebody hires for a CEO position like a 21 year old person with no experience.

**How can you tell me how many people work in your company?**

Yeah we are like basically we are like three people. I and I'm CEO and founder then we have two small investors. Because we are in a startup we are still in the process that we don't have this you know specific like goal that in one year we will be doing this in five years we will be doing this and kind of be there because we are we see the main thing I thought about this that we want to we see this changing and we want to be the leader of this change in the market.

But how it's exactly changing it. We cannot say right now you know we have to adopt like on a monthly basis what we hear from our clients from new clients from the feedback on the market from different conferences meeting different level of each our managers or recruiters. So we have to adopt all the time to be like really what the market needs. If we choose just the way okay we want to be this we want to achieve you know only these goals it's it's gonna be fair if we can adopt on the feedback all the time like in monthly basis, then we can be it can be like a market leaders one day. So that's why we are not we are still in the in this very like very very young stage that we are learning every day you know from the feedback from the feedback from the market see what's going around what's going around in this market. And of course we see that in Estonian or Baltics market. (...) So we know that the market this kindness is same maybe like small difference but actually quite exactly the same. You have job boards your like you post the job like one picture one paper you posted in a job board or social media you gather all your candidates and then you are doing the phone screening process and then you are inviting people to the meeting through Skype or real meeting.

**How important in your strategy development is data or how do you build your strategy?**

it's two things. One is like in business development data you the other thing I have enough yet talked about this. What we are doing with this video of data or candidate data there's like two different fields. From candidate data, we can analyse from the video their speech, their emotions. We also analysed their speech but what they are talking and how they are talking comparing this with emotions and we are already we already built our own algorithms the way that we can say like if you are more like introvert
or extrovert. How high is your energy level how high is your positivity level and how stressful was this process for you. And also, language fluency right now it's only from Estonian and English but like we can easily adopt with new languages. And the other part is sales and feedback. But we have talked with so many companies from so many different markets but we still don't believe or think that we have enough data to make good decisions.

You still don't have enough data?

Exactly.

Can you elaborate on that?

You know we think that actually it's it's pretty interesting for me as well as business neither. What was this the point that the what. How much data should I have to make decisions. To be confident that now I have enough data. okay. Based on the state that could choose. What are the next steps or hope that they these steps. And it's pretty interesting for me that as well. Okay now I have enough things that now I can make a decision based on data only. Okay but of course we are thinking like every day on that that based on this feedback or data that we get from the market. What are the next steps. We are thinking all the way all the time about that but to say right now that we are confident that okay we have this amount of data now we can make a decision. No. But we are. We are like collecting everything like how our clients behave how we get to clients from how many from how many like sales calls, we get how many meetings from how many meetings do we get there how big sales and so on. It's at the end of course it's because of that. As we are pretty small we cannot put like a thousand goals a day or something like that. So every company is so different. You know you at one point they can say that from hundred calls I got 30 meetings and from 30 meetings I got let's say like 100 sales it means like a hundred pieces we are selling like pieces you know a piece of the recruitment process to use our tool for one recruitment process. So let's say like from these 30 meeting, I don't know I got 10 clients and every client took the 10 job offer or job the recruitment process using in ten recruiting process our tool. So we got like 100 sales but at the other point I could think like also make 100 calls got the 20 meetings. Things also got like five clients, but every client took to hundred process, you know. So And then I have like 1000 sales. So it's so like you. It's for me as well like what this this and the same time like you know I could compare liking to say OK in this companies where is like office or decision making team is it can be same big in a company where they have a thousand employees and like let's say they have like a five decision makers in let's say they like Tallink as an example they have like OK maybe five decision makers or five people who are making decision to use our service to buy our shirt but they have like over 10000 employees. And so they would they would like to use our service for like 400 to 500 recruitments a year. Other at other point there is like a 30 people company maybe they have like also like five people who are making decisions and they are hiring 10 people a year.

How are decisions made in your company because you said you feel that still you don't have enough data to make the right decision and you have relatively small team. As you said you have a lawyer who is very hands on probably with a lot of experience. Are decisions still more based on experience or on evidence?

I would I would say that both. It depends on what kind of decision. If it's if it's sales you know it's... As an example I could I can tell you example big shopping malls, grocery stores like Selver, Prisma, Rimi and so on. I have a call to Rimi I have a call. I have talked to Coop. called. Selver everybody not using it. No we don't use it. And I was like okay Prisma is the same. .Mmm hmmm. And they I even did didn't think about like what why should they call them if I see that thought this same exact same field except the same companies just a different brand name. Why why they should use you know I got three likes strong No's. Why should they call it the fourth one. And they a forgot I went. The other companies to call and talk. And now six months around or four months later I was like I was still like OK I'm going to make this call and they were they are super interested in us. So it's you know if I if I say that there is like six grocery stores it's like a supermarket so grocery store chains and I talked to four and they've got strong no's not just like Oh No let's start like in six months not like strong no we're never gonna stop like hiring like very low level customer support people with this kind of tool. So if I make my decision based on this data. So, like 80 percent I got strong No I'm 100 percent no's oh and from 80 percent of this market you know this this one field of clients.

So the data says that I should it's pointless to call it the last one. But I made the decision that Okay I'm going to call anyway. And I got a yes. When do you say that data is enough? That was a simple example from a sales process.
If you have a relatively small team but growing. Do you feel that there are some specific challenges or problems when adopting more data indecision making or in everyday operations in your company?

No, data is the main thing. The question it's just when there is enough data to make a decision.

**How could you describe the decision-making process in your company in general? Now you gave me an example how intuition or experience won over data but in general, how are decisions made?**

Basically. I am looking for more more like thoughts from the other. Um. Like I could rip apart like development or technical side and like business development. We have CTO and we talk everything through with him and he's like a kind that... I or we with the shareholders make a decision. But based on his insight and we as as he's like a big part in our company we really like highly use his thoughts and to make a decision. We are never late he's saying one thing and we are saying no let's do the other thing. But we can adjust, because these technical people they are often like very too technocal; you know and sometimes you have to make like technical decision but from the business perspective. So this way we are combining. But business development is totally like I'm asking from shareholders about their thoughts. And as I said we have like a monthly roundtable with everybody but also like a CTO and then other employees and we gather feedback from them and their thoughts and then make a decision. But the like last decision making is always like mine. **So it's basically authority based.**

Yeah okay.

But I'm always like talking never like let's do this that I'm thinking like let's do this. Never. Like if I have some thoughts or idea I'm always like asking feedback from others and then I make a decision.

Okay.

So I'm always like listening what others have to say. I am listening to them and never say it's a stupid idea. Because that's not good leadership.

**But how do you gather this data for decision making? You mentioned that there's your business development data and there's video data that for the video data you made it very clear what kind of information you can extract. So you have that information and then business development data. So how do you gather it? You have the CTO and the CTO gathers all of the information?**

I gather all business development data and it's not very well organised. I mean for the sales data we use Piperdrive. So we gather all the sales data that they're there for the business development as we are like still I'm going to call like a very young company that we have like very so many decisions. So sometimes you're never you don't even have time together this information you have to make a decision in five minutes. So you just take three calls, ask ask for feedback and then make a decision. So you don't gather, but usually like in in a record all of this. We have Slack. We gather data, but it's not structured or integrated.

Not yet.

Okay.

But it's the this is the future you know when we see that we have to do it just one day when we understand that we have like this focus we know what we are doing in the big picture because right now we don't even know what we are doing in the book.I mean we know what we are doing right now but we don't see what this big fishery for the future. So that's why it's kind of hard to make like the frame or like work structure or how to collect this data. **Do you feel that when you said that sometimes there's like hundreds of decision or 10 decisions need to be made very fast? You don't have the privilege of time. Do you feel that sometimes if you had taken more time and considered more options or more facts evidence analytics it would have provided a better outcome?**

Of course.

Yes I think. Mm hmm. I think it says it's sometimes it's hard to say I know how to track it. How are you that how you track it. You know how you cannot compare like. if you know those who make this decision. Yeah and this is the outcome. and you had like let's say. In the perfect world you had 2 similar decisions, had led them to same similar decisions. not similar but you had two decisions not like three or four or five. In an easy and ideal work. But you cannot. No. Still if you had only two the and you can never say
that. What would have happened with this other decision. You know like let's say like you made a decision. Outcome is like fail. But how you can say that with the other decision you would have made the outcome wouldn't be two times fail?

*If you had enough time and data was readily available maybe to you?*

Yeah. Yeah. That way. Yes definitely. But then if you had this enough data that you are confident about and you are confident to make that this data shows you this decision or this data says that this is the right decision. Then you will never think that this was a wrong decision.

*Do you trust data?*

Yeah definitely. OK.

*How do you share information?*

Yeah. Mostly like Pipedrive or Slack and of course like sometimes in Slack there are like files. In Slack, it's all the information.

*Do you feel that in your situation you have a smaller team yet still but you're growing? Do you feel that information is and data is available to everyone in the company or is there some kind of levels what people can access to?*

Right now, it's very flat.

*And would you like to keep it that way for the future?*

I mean. Mostly yes. Let's say sales team or salesman in the sales team can see something that's good for let's say marketing and marketing can think that Okay. How is Sales people never know what is good but they can hear from the market something back or doing sales something that can be helpful? And that's it should be quite flat. But of course like high level discussions like let's say investments are what markets to move or sell something like these big decisions. This is not always like you know it's just. Of course you shouldn't make in you know very like that's a lower level or easier level you should talk with the on the flat but still the decision making cannot be so flat because there is like this problem that at some different perspectives and thoughts are good. But at one point if there are too many thoughts and ideas.

*In this case you have facts versus opinions?*

Yeah yeah yeah but let's say like which market to move. Let's say you have like five similar markets based on data. Then you have to choose one. So which do one you choose?

*And is that decision then based on data?*

Yeah, exactly.

And that that it's like let's say there is a new employee who has been in the company for 3 months. They don't have this you know they have their own thoughts and ideas maybe and maybe it's very strong ideas but they don't have this experience with our company. So they have only in there their respect from ideas from their own perspective.

*How do you gather data for decision making?*

In sales process it's like two things that we first this we are doing sales on the meetings are in communication with the possible client. The other part this we are asking questions and this data we are using like. We have a lake specific ten questions. And we ask all these questions from every company. And these are the questions, based on which we base our next decisions.

And based on this let's say that right now we are in and we are thinking if we should also move to job offer, job posting and advertising field because we see that it's moving from the job post to social media. So right now we are asking like very specifically how they are advertising right now their job offers are job postings in what channels how much they pay for that. What's the budget for their job posting? How many people they want to reach for? How many CVs they want to get? And so on. So very is specific. And then there is the question of when is the point that okay. Now we have like 100 answers. Now we can make a decision, as we have 200 answers or 1000 answers. And now we can make a decision.

*In the future, do you see a role for business analyst?*
Yeah definitely, definitely. Of course we have. There is. I'm very. Let's say data driven a day. I really believe in data because its data actually comes from experience or data states experience.

You know what they mean?

*No not quite.*

If we ask the same question from hundred companies. This process asking these questions and getting answers is experience;

Exactly. And we can call it the experience you know. What we get out of that or these answers is data. But how how we got these answers is experience. So I'm very like experience driven. So it's it means actually data driven.

*So how do you analyze your decisions? How do you measure the outcomes The results of your decisions?*

Mm hhm. Of course, it's like a money perspective is the easiest thing here. If this new feature or new way or new feature brings more money out of let's say you using less time or money on sales then it was a good decision.

Okay.

And because right now we are very like easily... It's in this small team it's so easy to say that. How much money we spent on that? How much we got out of that?

*So return on investment. And then do you say OK do you go back and you say based on the data now we gathered from this decision we will apply these and these changes*

Exactly.

*How do you feel that, is there room for decision making improvement in your company?*

I think not so much in decision making but more like I would... I will take in consideration in this question like how. Like how long we keep this decision you know? Okay now we have this amount of data. We make this decision, let's say in one month it's not and second month it's not working. If we now change or do we keep it? You know maybe this is too fast right have been too fast but same time week we have like and that that why it's a problem. Actually, it's a problem. This decision was wrong. But if and we haven't you know tried many things but then maybe this decision was right. But the timing or this amount of time we keep this decision in the process was still too small. So in the future when we have a similar decision to make we will elaborate on this decision and then we say it was that decision. But actually, maybe it was just like I said wrong time or too small of time we keep the decision and then we make the wrong decision again because we had this experience that this decision didn't work.

And that's a bad thing. And that needs like that. can be much improved. I think so. We are still so young I cannot say.

*How old is your organisation?*

We have been from the beginning of last year so it's one year and three-four months. Yeah, three months.

*To which extent do you think your organization is data driven?*

Mm-hmm.

Do you mean as a company?

*Yeah. How much is it driven by data you perceive today?*

Like a percentage or level?

*You can define yourself. Or do you consider your company a data-driven company?*

Not yet.

Yeah. If you if you ask like that then no. But if you asked like how much data which I think maybe 30 per cent 20 per cent 20-30 percent something like that. And of course, if-if we talk about this business side of business development if we talk about this standardisation part it's like it's data drper centiven.
All right. Ok. So what does it take for you in your opinion. Now you have enough experience to say to be data how to see, when deploying data into your organisation? How to kind of deploy it successfully so it actually would work?

We should. We can put it work and be data driven. We should make a framework for that data collection and also analysis. It cannot build some like a pre-made frame. It has to be very specific for our company.

What kind of culture should it be in your organization to be data driven?

I think the culture is not a problem. If we can make a good framework, you know it’s because...Because we in like employee level as well everybody talks about data you know. Like everybody are interested in this data and they want to...Because people don't want to be decision makers actually.

How so?

Because they're afraid. You know I made a decision and it was wrong. I made a mistake. But if you can elaborate to this data that you can say that they made the decision yes. But based on this data. So if this decision was wrong, then data was wrong or data they gave me wrong analysis.

So data encourages people to make more decisions?

Exactly.

That's an interesting idea

And if we add this frame to follow this frame and the framework and the and always put the data they are collecting. If it sales, if it's marketing if it’s technical then this is not the problem. You know. It's easy to make it like to automate it the way you know to automate this frame the way it works well and so on. So, it's this. I see that it’s very like you can you bring decision making if the very low level. And then that way the company will grow faster.

So, you feel that decisions should be every person should be empowered to take a decision?

Exactly. Because if I don't know the good example if the salesman like come the CEO to ask if he can make this discount for this plan it's just way too... Let's say. Very small decision for the CEO. And if we have this data, that you know they can use this data then if it’s such a big company it's good to make first the like this discount and then they will return and the return or the profit we make it will be much bigger here.

So if I understand correctly you’re saying that if every person is enabled with enough knowledge you would say that they can make the decisions faster and improve efficiency? Thank you. Was there something else you would like to share?

I don't know.

About data. No. Yeah that's that. I think that’s the main thing about data that. In-house everybody issued should be able to use it. And there has to be like a framework designed for exactly for this or for this company you are talking about. It cannot be like some basic framework.

Ok. Every company and then for your information at the moment that you have the data is internal mostly or external as well? Do you purchase any external data?

We don't purchase. But we use some in technical side data we can get for free. And because it's just available there. Otherwise you would like buy it but and also like let's say we're analyzing new markets or current like Estonian market. How many job posting was in last month. What kind of companies have, how many live job postings. Right now we the active job postings so when you get this this kind of thing but this is not like paid.

Do you feel that there's one piece of data if you could have that would improve the overall decision making?

Yes. know we're like let's say in whatever market we are talking about. If I can get a list of companies, all of the companies and how much they are hiring a year and what kind of positions? Even not like exactly like I don't know Tallin, Viru shop customer so specialist but like you know we're a lot like customer service person. It's like field in field.

Yeah. Thank you very much. We will now close our interview.
Transcript Interview D

*Could we start by you perhaps introducing yourself a little bit. What do you do in the company?*

My name is X, I’m here for three and a half years. I used to build robot software. I was leading three teams making robots smart. And now I’m a head of data science and leading a team, supporting every aspect of the company development with data dashboards and KPIs and so on.

*And how long how big is your team by the way?*

Last one was about 200. At the moment we're 10 people, but I am hiring at the moment. I studied business but I dropped out of uni, then I studied mathematics bachelor’s masters and I did a piece in computer science.

*How important do you think data is in strategy building?*

That's a complex question. Because strategy is a very long term; strategy is very much related to vision or whatever you kind of believe will happen in the world or whatever you want to achieve. I guess it's dependent on the stage of the company or if it's like a very early stage start-up then probably there is almost no data. And even if You mean in a company like Company D then even half a year ago we didn't have hundreds or thousands of customers, we just had robots. And the strategy. Wasn't too much filled with data. Of course, you stay rational and honest. Yeah I didn't know it a bit definitely, you cannot be out of the world completely but I guess there is a lot of creative thinking and some arbitrary vision or whatever you want to achieve; doesn’t have to be fully data driven but usually must coincide with some big trends in the society in your company so unless it's a strategy for 50 years. But if it's like a 10-year strategy or something then I guess you're looking at some vague important or you must be looking at some trends in society.

*How are decisions made in your company?*

Yes. I mean it is an organisation as well, but a start-up definitely. In the sense that we are still improving our business model then we are still not fully sure of what will be our most important thing. We are just scaling. We're definitely not profitable as a company. We might be profitable operationally somewhere but with the scaling list in the most important aspect of not being profitable.

*How are decisions made in the company?*

Probably tens of thousands of decisions are made every week and in different ways. But we definitely hope to be data driven or our objectives are defined numerically for every team. Sometimes for a person as well. Updated regularly, tracked weekly. So yeah, it's everything is measured and done in a really data driven way.

*And at the very top level it does still the same rule apply?*

Yes, definitely. Company level metrics for every quarter for long term ones definitely, business models everything is numerical and based on data and.

*If we speak about two extremes one is intuition based decision making vs data driven so that even at the top level you could say that?*

What type of decisions most decisions are done in a data driven way not intuition based. So for example if we want to make robots cheaper we don’t do anything intuition. We list all of the components we see how often they break and so on. So it's purely data based think. And then fulfilling company goals. It's also that the data based on whatever you want to improve or decrease or when you are looking only at the data. But the very very first company level key objectives. This is more complex like this is maybe half and hald, but it's only for their general direction. Like for example should we be expanding the business at the moment or it should be improving. Should we be improving the quality or should we be improving the unit economics. This decision is much more difficult to do a purely data driven way.

*So still intuition or experience plays a role?*

Yes, couple of decisions must be done like this but most of the decisions are done data driven way.

*How is data gathered for decision making?*
Data is gathered by everything we do. So robots are producing tens of thousands of data points per second. All of this is saved. Every application is gathering data by customers by kitchen by by people repairing robots producing them. Every interaction with H.R. are hiring people and so on. Everything is kind of predictable.

**And how do you share information within the organization? Is it still segmented?**

No, we have a central data, data team and data warehouse team. And all the data and all the dashboards and then KPI and metrics and everything is accessible by everyone and the data teams centrally supports every other team in the company.

**Thank you. So how would you describe the culture in your organisation?**

Related data then yes, it is data driven company. But we are definitely technically very heavy company as well so a lot of people are highly skilled in math and statistics. So, I guess that makes it easier.

**In your decisions when you make decisions you have an outcome. For example whether I don’t know building your robot or improve existing and do you assess based on the data?**

Yes, definitely. Everything is recorded regularly (weekly), not only at the end of the project.

**Do you feel that considering your solid data heavy driven and so highly skilled do you feel that there’s still room for improvement for data, for making decisions in general?**

Yeah definitely there is. But. That's I think fine. It's always possible to improve everything. But but in general I think we are operating well at the moment.

**In your opinion and based on also your personal experience what does it take for organization to be successfully deploying data?**

People need to be skilled, trained and good. I mean if a random person and company does not understand what is the distribution or whatever then you cannot make it a data company. So everyone needs to understand what is distribution or something like this are; or expected value or whatever. People need to be skilled I guess people can be trained to be skilled or when it's easier to hire people who are skilled. And then Like definitely the accessibility or putting resources like how important or powerful is that the team responsible for dealing with data and sharing it and how much impact do they have to the company. Like I as a data team. That the team really good and analyse what inner workings of a random team and I can show what they are doing wrong and they cannot argue against it. And so, I mean even if they don't think in terms of data in the beginning then someone else will find mistakes in their operation. So, I guess empowering the team is important.

**Is it just about empowering the data team or enabling everyone the company to have access to information?**

Yes but in order to have access someone needs to create this access, it needs to make data which is simple, correct, up to date, nice. Have tools for playing with it. Like teaching people how to use it. So supporting them and so on if you don't have resources doing it then...

**Is it challenging as the company is growing?**

Yes, definitely, but not as challenging as building robots.

**So is there any information or data that is missing to build better robots at this point?**

I guess at the moment we are bit bad in doing like customer service and doing analysis with interviews and things with customers. So we are getting the quantitative data but we are lacking sometimes qualitative data. One reason also is that for the data team, for example the customers are in a different time zone and different continent, so most of our customers in America. So it's easier to gather in numerical data, but harder to get qualitative data. So I guess maybe that would be something that they are lacking at the moment.

**So if I understand correctly you’re so quantitative driven that the qualitative, the soft part is missing?**

It's not missing but we could have more of it. But we still have it.

**Absolutely. Ok I will ask a bit of a provocative question. Do you think that it's very industry specific or in general that it's actually more companies deployed data when making decisions, they would be better?**
I don't know or see that any industry would be an exception. Or how can it be? You don't have to stay by every decision or every aspect but in some aspects and in some decisions, data can always help. For every company. So. Definitely is.

**Have there been any major challenges that you feel comfortable in sharing regarding deploying data?**

Yeah. This phrasing of a question deploying data itself is confusing for me because they died self actually is rather meaningless, just terabytes of numbers that no one understands it, the matter is the insight and action from it. Or. I don't think that there is anything major. It's just takes time and effort but I don't think that there is anything stopping using it. Yeah there are some things which I haven't experienced myself but I know from theory from books and from my mentors and things that data team can lose trust in the eyes of other teams too if their data is sometimes wrong or not up to date if it happens a couple of times. And even if you show correct data later then they don't trust it. So. So, like the data must be trusted by others and the trust can only be earned not forced. So, I guess that is a bit of a risk that if you have a crappier data team and you start forcing it you might ruin it lot more. But I haven't had this experience myself so I don't know. And of course, I mean, I guess it still comes from the very top as well. The CEO is still the person who says what is important in the company. And if he doesn't constantly and constantly say that data is important for years, every week, every month. If it doesn't come to it then it's not going to happen either I guess.

**And in your case that supports from the top is present?**

Yes. I think there was also by previous, by founders and and by the CEO now and just so it has been always there.

**So you even say to the extent of not even monthly not yearly but actually weekly it has to be emphasized as the core?**

Yes. Yes you cannot go and talk to the CEO and saying something without backing it up my data. It can be a disaster for the person.

**Really? So basically if you want to come to this you know you have to have the insights the information the knowledge?**

That definitely, he is not going to listen to you then if you don't have it.

**And that has been from the beginning of the company?**

Yes.

**And may I ask what did you do before myself?**

I was working in London in a small startup doing some machine learning.

**Was it also the same data driven?**

It was very early. So it was a bit more difficult. I think it was one of the founders was from banking industry and finance. Other one had a scientific background. So I think it was but we were just too tiny. Like if you have four people there. So everyone knows everything. Then you don't need to prove so much and you don't have much data because you don't have customers and stuff. Before that I was in the university of Tartu, building some curriculums and doing my Ph.D. I mean I haven't worked in these large organizations, usually I work in small ones.

**Would you like to share anything else that I might have left out from the interview?**

I guess there is one conflict. Which will always be inherent. Is that the better the person is like numerically, mathematically scientifically, the worse he or she quite often in explaining, visualizing it, making it understandable by the other person. So there is this conflict of understanding a common person. Being empathic versus being extreme in understanding math. And the trade-off which is I guess important. That can be supported you can You can have internal, for example, I know that now that in Air BNB they have they have internal scientific journals so that when before you contribute to the company you need to publish it their internal data team distribute them and see whether it was understandable by a random person and only then they publish to text or team because you are getting kind of immersed with that data and you think like everything becomes clear for you. But it is not clear by the reader who reads it for five minutes.

**Have you experienced it here?**
Yes, definitely yes.

**And how do you handle it?**

The same thing. Easy thing. Someone needs to read it before, who hasn't been working with the same analysis for a week. And just force him to standard things like on graph at the time axis they need to be labeled and so I would say it's like you can go over and see the basic mistakes very easily, but there needs to be forced still because the more technically heavy the person the more automatically everything comes from him or her to understand everything related numerically, but not what people want that.

**So everyone even the head of HR is knowledgeable or trained well enough to understand it?**

Yes, I guess for me for things like KPIs and distributions and some things yes I think every team heads can manage. There are even companies where before they train absolutely everyone to understand database language SQL so that there is not a single person in the company who cannot do. That you can keep a quarry there that they compare it to the times that now is common that you would assume that everyone can use Microsoft Excel, that it's okay if you put excel file into your e-mail you don't have to ask "Have you seen it"? You assume that the person has seen it or understands it. They assume in their company everyone can write basic database language.

**Is that the future of Company D?**

Yes, probably yes I am yes. Yes, it is even in my roll up. But maybe I can start lightly can be on the team needs something.

**So do you believe that the people should be already hired with that mind-set understanding or can they be trained?**

They can be trained but the hiring is easier than training but uh but I mean it's not that difficult if you have a decent education then I mean these things that you need to understand basic things that are complex like I mean and median and stuff I would like to see'.

All right.

So but you need to be able to read distribution and scatter plot and stuff like this. It seems impossible.

**Well what about companies like if let's consider a scenario where CEO is not to the extent knowledgeable?**

Then it's bad, it's very difficult. I guess there are probably companies where the CEO is very inspirational down and out outward facing them and going giving interviews and having some vision and then there are very strong. There is a very strong team of actually like organizing the company. But if you look at the best companies in the world especially from Modern Times like all the CEOs are extremely good in numbers and data. And there are no exceptions. Even Steve Jobs. Who was maybe most visionary and most of the work space goes out most maybe didn't really care about optimization, optimizing things but instead he like I think everyone every company like major good new companies like Uber and Lyft and Spotify and whatever. Definitely extremely data driven.

**Thank you. When speaking about data you mentioned that one theoretical bigger mistakes would be that people might lose trust in data people. How are for example considering small mistakes. Or. Mistakes or errors treated in Company D?**

Mistakes in data and mistakes in general? No not big think correcting and correcting that learning. Yeah. Yeah. I don't think that our players have ever been punished by mistakes Maybe lack of trying is worse. If you lose motivation and you don't care anymore to try better. That can be punished but I don't think that mistakes are punished.

**Thank you very much. So if you're any other ideas you'd like to share that would be very helpful for us.**

Maybe some more as the research is qualitative research nearing I would personally be interested in knowing some kind of more quantitative things as one for example even simple things. Do they have a data team? Also, how large is the data team?

Do they have like numerical objectives that are regularly tracked yearly quarterly monthly whatever but are regularly daily or weekly or monthly tracked or not. All of the companies I know are very advanced.
Oh the reason we're doing the study is a mixed method. So the first part was the quantitative. And now we're going again with the qualitative and the sciences from the first one you can get this thing right. We already did get from Company D, since it's anonymous I just know the company name I don't know who exactly answered You did so. So that's what I did I go back. I didn't want to go back with the same question. It makes sense.

Good. Oh yeah. Maybe then it's there. But maybe it's something we're still missing from there are the same thing that I'm not sure about. But there must be some put that numerically measurable objectives and regularly tracking them numerically.

Because the goal of a company or what's the aim of all the people doing something you are as you should be going somewhere or there should be some aim. And if you are measuring and tracking it then you will be better than you even know. Are you getting better faster like we're doing quick things last mile. If you don't have formal kind of KPI or objectives which are numerical, and if you don't track them numerically. For me it seems so impossible to do anything.

For me that comes from the fact that how can a business operate? Could you do little to speak about the KPIs?

We are using the method Objectives and Key Results used by many companies such as Intel and Google and so on and they are being IBM and these are objectives which are all qualitative, very aspirational and then every objective has some key results. 3 to 5 which are precisely numerical and then the company has three objectives every quarter for us and every objective has three key results which are fully numerical which are correlated with this objective. And then every team has three objectives and every objective has three key results. So in Starship, case numerically tracked early track results for the company is about two hundred. Uh uh. Company level. And these things should be tied up the company level and they are aligned as well with each other. There are a lot of alignment meetings and things.

How often do these alignment meetings happen?

Each quarter.

What about the strategy. How is it built?

I don't know.

The strategy has been fishy because sometimes it's because it's for a start-up everything changes so rapidly and sometimes very stretched but next month the strategy is different. So there is a strategy and there is for every timeframe. So, there is something for the 20 years, there is something for five years something. I could phrase it and probably know it a bit. But yeah it's uh there is not a huge amount of effort put into it because we are small. Two hundred people. I think I mean small to Air BnB with twenty thousand people. Soon we will have tens of thousands of people.

Right and then do you feel that it's going to be as a company will be growing again I'd like to come back to it. Will it be more challenging or will be more natural to measure and create these KPIs?

I guess it's gonna be even more meaningful to do because how would you know that all the different teams are aligned to the same thing. I think it's gonna be more important to do it. It might be more annoying but uh but I guess it's more important to.

You said that you have a central data warehouse. Did you create it yourself or did you purchase it?

Ourselves everything. But it s there is there are our cases are usually more complex because there we have special cases for robotic data because it's so large and huge and then actually want to do things that are kind of together. So Yeah…

Thank you very much for your time.

Transcript Interview E

Thank you. Uh first of all I need your consent. You say in your full name and saying that it's okay that I recorded you for the purpose of my master's thesis.
Yeah. So my name is O.S. Is Company E’s name gonna be mentioned anywhere?

No

Cool.

Everything is anonymised. So a little background. Uh what we're doing we're basically uh we were developing a framework how to understand if a company is data driven. It has several aspects. What I can’t currently introduce to you. Yeah. But my aim is to get a little bit more insights of how this organization breathes uh when it comes to data and decision making. Mm hmm. So I would like to start with. Your personal opinion. You and you can projected that on the basis of uh what you do for work but how important do you think that data is in decision making?

Important I think it should be or how important I think it is here?

It should be first of all and then how important it is here. Uh

Yeah, I think I mean especially in the context of um what we do. It's probably one of the or should be one of the most important um parts of decision making. Um whether it’s product design or we're due to um like the the manual fallback in there just being a lot of people um there's a necessity to understand how people perform in general. And there's many or without getting into specifics there's many kinds of sides from the business of the financial side as well. Um whether to oh kind of what parts of the company to invest more into building up for. And none of this none of those decisions can really be made. Um but any assumptions and they need to be data driven. So yeah, I'd say it's among the most important things we could have for decisions.

And how is the reality? If this was how it should be?

I think that's somewhat hard to say. I think it used to be in a pretty bad spot um even a couple of months ago. And we've been working quite hard to introduce what we changed the BI platform we use into something that could be a lot more self-serving, and that people could use without having to rely on kind of loads of ad hoc requests from the um from the data team.

Mm hmm.

And that has helped a lot I think especially kind of the middle management level um who were able to do quite in-depth analysis by using that self-service layer themselves. I am not sure how much data is actually used at the top level what kind of a managerial level to make very high level decisions. But what I do know is that uh for example a CEO CFO here are all really keen on um just kind of in principle data driven decision making. So I would expect and hope that most of the decisions are included as data driven as they can be given our current infrastructure and that they're quite keen on improving that as well.

Are they somehow also fostering this mentality to be data driven in all of the levels in this company?

Yes, and how?

You mean the managerial?

Yes, the C level.

Yeah definitely. So for example every Friday um there's kind of a weekly recap um so kind of the highs and lows of the week. Mm hmm. Um and then there'll be a specific slide which focus on um how the um kind of of the few core KPIs as we measure how. Whether we hit the goals they set on Monday for Friday and um kind of any lessons or improvements learned from that going into the next week. It's usually every Monday there'll be some data driven goals. And every Friday there'll be reviewed. And then towards the end of every quarter. Now as of last quarter we're also introducing the ideas of the idea of OKRs.

Mm hmm.

And that's definitely a top level push um to be very specific and data focused when defining the OKRs.

Okay. But when it comes to your I don’t know how much you know about product development? How much do you make the decisions, which product improvement? Based on the data.

I think...

Or is it still more like somebody has a gut feeling that this is something that should be done?
As it currently is. Um. Although it is improving. I think the current status is kind of people have a hunch about what should be made; it's made and then retroactively data is used to kind of try and examine whether there was an effect or a positive effect. But we are yet to move to a stage where there's one where you might need pause it.

So where were we? We were talking about uh product development and data usage and you said that most of the time you still think that it's based on gut feeling or a hunch.

Yeah that any data use is kind of retroactive rather than contested in advance on a subset of our users and then the decision is based on that.

What do you think is missing from your company to also for the product to become more data driven?

I think. In quite a few aspects there's just a lot of instrumentation missing or it's a bit lacklustre and not fully reliable. We're also working to have a data warehouse up and a lot of the analysis just due to the size of what the tables are growing, is getting quite unmanageable. So it can be quite difficult doing some quick analysis in advance and even post hoc. I think, I think kind of the mindset is that most the people working on the product actually really want to test. But so far it's just due to the lack of people who can devote their time to building up some kind of I don't know AP testing frameworks or, or any kind of well defined testing pipelines. It's yeah I'd say it's mostly just like a lack of people who could actually devote their time to setting this frameworks up rather than any kind of company level mindset missing.

Okay so the mindset is there but is there also let's say publicly discussed that yes we will make an investment in those areas get that personnel on board and to specifically deal with those areas or no?

There are... we are quite well aware of how many more people that could be in the analytics team to cover a lot of the work because currently we're essentially two and we're just about able to like fight the fires. But there's very little time left over to actually kind of build anything going forward and it's definitely the need for more analysts has been recognized at the company level and where we're putting good spec out in the next few days but I'm not sure whether this is true specifically for the kind of shift in how product is developed. So, I don't know whether there's been a chat about you know product decisions being more proactive and data driven before implementation. I don't think like that the people we're looking to hire are being hired without specifically mind. So it's a bit I think it's fair to say that currently there's a bit of... I think it's the kind of thing that's maybe kind of up in the air in every everyone would like for it to change. But it's a bit unclear who the person responsible for it should be? Whether it's whether the initiative should come more from the product management side or whether the initiative for building up the infrastructure to be able to test but it should come from the analytics or data side.

Who is currently responsible in this company for our data and the storage of data? And then the entire infrastructure behind it not just only the last part of analysis.

So that's also a bit hazy. There's a well-defined dev ops team who manages the various database and instant as well. The I guess everything in the end reports to the V.P. of engineering corps or the CTO. But as it is currently the analytics or data team doesn't have a team lead per say. So that's definitely I guess one of the positions that needs to be filled for it to be kind of a more. So that for there to be a person like with a data vision and can have clear ideas of where we should be headed rather than just someone kind of underground fighting for us on a day to day basis.

How much, help me understand. I don't know that much about Company E, is data itself used to provide the service or is it just accompanying it more?

Yeah well the service is provided using well data so photos, videos that the users upload. So in that sense yes. But the vast majority of data is just a company.

Okay. So. So the data is used to verify the decision of verifying is still in the hands of the human person.

It's mixed. So it's a mix of automation and manual fall back on people.

So it's still half of the parties kind of like automated and using machine learning and so. So in that sense data itself is very much used in the company but not just in everyday business decisions that much.?

Yeah yeah.

Okay. So. How would you evaluate the fundamentals of Company E being able to be data driven. What are the areas that need a lot of improvement? You already mentioned some, but I would like you to take it again.
Okay. Um I remember seeing I think one of the big things that kind of is starting to click in place is just the awareness of a lot of data existing. I think currently a lot of teams and maybe even team leads aren't fully aware of the potential metrics that could be drawn out from the data we already store and have stood for a while. And all it needs is just someone to go into the BI platform use put together a dashboard and you'd actually have something really informative. And I guess there's a there's a bit of a vacuum where the data team doesn't know quite what the right questions are to ask because they're not as familiar with I don't know the end user flows or the product or design or whatever and the respective teams aren't aware of the data that's there. So I think one of things for just more company Y data driven decisions is to bridge this gap between one set of people not knowing what to ask and the other so people not knowing that those potential questions could be answered if they did so I think and we're trying to work in this to a lot of people just also needs a bit more trading in kind of what makes a good KPI how to use the BI platform that we have maybe some kind of core statistical understandings of you know distributions and sample sizes and differences between various averages like meetings and beans and whatnot.

And I'd probably say those are the most important I think everyone here agrees that the data should be kind of a bottom up rather than top down thing that the teams themselves should have their own KPIs is defined by themselves that they are genuinely interested in looking at that actually are informative to their decision making and that it kind of it's pushed upwards from those teams actually having a vested interest in it themselves.

But technology wise? How much improvement does the company need to be more flexible in providing all of this data?

The I think the main thing stopping us currently is just a lack of a properly set up data warehouse. I think the BI platform is sufficient for making fairly well-informed proactive decisions and most teams. We definitely need to think about how and what needs to be improved once we want to start getting into predictive analytics also in outside of the core product automated automation. But I think the main hurdle currently is the warehousing or the data storage and the data cleaning side.

Okay, but do you or does anybody in this company have a clear overview of all of the data that you store?

What do you mean?

Like a map a data map like this is the information that we store. Or is it very hectic every every department only knows their side? And if you start asking then you basically have to go through all of the 10 people to get your overview.

I think the database level it's quite team specific and like there are some fields that I don't know where a developer has put in a year ago and it's just by now like a legacy field that's never used ,the Looker, the BI platform we use.

What was it?

Looker. I guess it kind of does that job for you because essentially you, the people on the developing side of it essentially right the or choose which columns from each table get shown and you kind of write the join like joining syntax and that's under the hood and people can just kind of click around and look at us the joining underneath so they can go through different tables without actually having an idea of how the or the database architecture.

But the BI is connected to all of the data?

Yes.

Okay.

But it doesn't display all of it

OK

It only displays what the people writing the Looker a kind of logic want to display outwards. So in that sense I think there's quite a good overview of the things people actually need. But I wouldn't say there's anyone maybe bar a few core developers who actually have like a complete understanding of what each table means and what it is there.

Okay. So let's say a marketing manager would like to get an overview of all of the client segments in quite specific details. Does anybody in the company know that. Okay. The information that I provide
this is the full information that there is no aspects missing that we actually somewhere in the system store about them.

Yeah. Well marketing is kind of a bad example because we none the marketing or advertising analytics is stored in our database. It's all on Google Analytics.

So, that's kind of siloed.

Yeah. I mean we're looking to join them at one point.

But currently the database itself is more kind of product and end user data oriented. And if the raw questions from that side then yeah I think they can be exhaustively answered at any specificity I'm using to be there too put your marketing is a good example of how I think they're probably current in the worst spot for like an easy access to data that would be relevant to them especially because often some of it will be in google analytics but then you'd like to also include the a lot of the stuff that's in a lot more detail on the database.

Okay. And currently there is no link in between there. In your opinion, who should be responsible of establishing this link in this company? Or is it this person that you said that is missing that should be the one?

Yeah... Given I think there should be like ideally a data kind of a head of data. But in lieu of that probably the V.P. of engineering. I guess that's that I mean that kind of draws out the size of the vacuum between you know data engineers' analysts and the V.P. of engineering. And I think a lot of those of decisions probably they will probably like the improvements we'll get made but they will get made as a reaction to something missing rather than something having been proactively built up.

So basically it's easing to fire again. OK. Do you know any good examples of that you have observed or been part of that. That you know that if data would have been used in the first place for the decision making then let's say a not a bad word but the outcome that came would have been more favourable?

Mm hmm. Yeah I think given a couple of weeks ago there was uh there was like an alternative product decision made. And there was quite a clear assumption they would cut the time these specialists have to use. In reality it turned or I just kind of I think at best kind of inconclusive at the moment. I haven't thought like the strong positive effects that were expected. It didn't really seem to be there. I mean could have yet to be fully confirmed. But if that is the case then that's a lot of development time and kind of product design time misused for something that could have been tested at a smaller scale beforehand rather than pushing live and then seeing how it works.

Do you see it as a as your kind of like a learning point that would make a change in the product team? In that specific instance?

Yes. Like what would you give the grounds of OK that you you plan to product and then you discover that or in any other examples that that you put a lot of effort time invested money and then you look back at it and decide you understand that okay 'Why did we choose this path? Like if we would have evaluated more or made a thorough analysis we'd go there and then the next project the mentalities changed then.

And I think that's what the reaction to the case I just mentioned has been. I think a lot of people have realised that the product gets more complex. Some kind of intuitions and hunches about what's going to work or not are becoming less and less valid. Also probably because the product changes they are becoming kind of more nuanced and incremental. And although I haven't been too actively involved in this myself I get the sense that people realise that this should have been a lot more kind of data driven thinking applied beforehand. Even in that case thought I would assume and hope that that's at least in the context of that team in a learning lesson.

In this company, how is it would be with you know this is the analysts always invite it behind the table for discussions or is it more that the analysts afterwards learns that OK. Like if you would have taken me in the beginning then I could have provided you this information or is it turned to the analyst like the last thing.

I think it varies a bit from team to team. But in general so far I'd say it's more the latter that decisions are made and the analysts are involved retroactively to look at the past effective something rather than having to move from step one and kind of keeping some kind of data awareness in mind when actually deciding.
Maybe not necessarily how to design the product but how to measure its effect. That's definitely something that could be improved.

*But what is the part of analysts in this company is it just somebody that provides let's say a report and then the insights and everything is basically driven by the person receiving the report or is the analyst also providing the insights to the analysis? What's the overall role for the analysts here?*

Well it’s kind of so far it's been up to the analysts themselves to define. I think it's been more or less a mix of that. So there's been a lot of things that the teams have specifically requested. They need some kind of metrics built up and dashboards built up on top of that. But from the data so we've also kind of tried to sit down and put ourselves in the shoes of other teams and kind of help come up with metrics and funds or whatever. But haven't been requested but that we think will be useful. If they were aware of it and then obviously having chat and spending some time working on it.

*How do you feel that analysis is going to waste? That is just like you send it out and just goes into this black box somewhere. No response.*

Not much but at least so far. When I say there have been a lot of direct requests it's more to create the kind of self-serving capacity for that specific metric or KPI. And it's up to the teams themselves to use it. So it's quite hard to actually gauge from the analytical side whether those tools are being used in or inside those teams. But I haven't felt like anything's going to waste because like everything we build up we essentially build up as the self service kind of stage logic and that logic needs to be there for things that eventually people are looking to anyway.

*How much do you feel that some requests are very one problem specific that if you if you get it sorted out there's no use of that data anymore or it is still more like big picture views that that doesn't change that rapidly when you build them up?*

I think there used to be a lot more of that kind of one of our customers would request some information about like a segment of the data or about the users or the user behaviour forms and it would just be like an ad hoc one-off thing with quite a lot of time going into getting the data prepared and presented and then probably being used. But from the data side not being correctly stored in logic so that could be kind of recyclable or reusable the next time it comes around. I think that's improved a lot since we changed the BI platform just due to the nature of how it's set up. And yes since, since we swapped which is a couple of months ago I haven’t really felt like I’ve had to do a lot of redundant one-off work that's already been done before and I'm just kind of going through the motions.

*I hear a lot of this upwards progress going on in the company towards actually being very data driven. What are the... We've talked about let's say issues and blockers. Yeah. What is something that you would want to bring out as a very good thing? Is it only this BI tool or is it some kind of mentality or a specific person joining the company that has changed something?*

I mean one of the thing I think, I think the main thing probably is what I mentioned earlier which is how like visibly interested the very top-level management is in kind of having you having a few well-defined KPIs isn't going over them at a company level. Literally every two I mean twice a week. So, kind of the beginning the week goals and then the week recap. I think another dimension kind of seeing the leaders of the company take metrics that involve and kind of centered around the performance of the company and everyone listening to them as a whole is really useful for fostering like a data driven approach in a company so I definitely highlight that and a lot of people have been also quite proactive in learning the various tools we provide, which has been quite nice and potentially also motivated by that last factor.

*Anything else?*

Um...not from top of my head I think the main ones.

*If you would have to evaluate how far is Company E becoming very data driven company? On a scale of zero to 10 for example. How far are you?*

Probably 5, 6. Kind of like in the middle. Where like the there's this very clear steps that need to be done to improve it and there is active work being done to improve that which is quite reassuring. There's definitely quite a long way to go still.

*Okay. Would you say that the last miles are the heaviest ones?*

Yeah I mean I imagine they'd be quiet if we stay on the ten point scale right. I'd probably imagine it’s quite easy to get the kind of 8 ish area but it's the last few steps which I think they'll be quite big problems
with kind of ensuring great data quality and and really making sure that data is used in a timely manner for all decisions. I think that’s something that’s quite difficult to implement but I think we can get very near there quite fast.

As a last question. I'd like you to do a little imagination exercise if Company E could have all of the money, What would you do to get to that ten? Like how would it ideally work. And then please to try to describe the aspects of all of the aspects of decision making in all of the departments in this company?

Yeah. Given an infinite kind of pool of people I think the first I think probably what what would or what’s missing the most and would probably have the biggest impact is a kind of purely data focused top level management people so like head of data over a data lead. Who because I mentioned earlier would be able to devote their time to actually kind of mapping out a data roadmap and having like a proactive plan rather than kind of shooting from the hip as we as we currently do and kind of having ideas of what to solve before they become problems. Then I think that. What we're also trying to achieve in the next few months and which is quite necessary is to have analytics embedded in all teams within the company. So rather than have you know just a few analysts who kind of or data scientists to kind of jump from project to project I want to change context and having to kind of readjust to the mind-set and the metrics of that team that each team would have you know their own least person who would kind of be there permanently and who have a very deep understanding of the context of that team which probably is never going to be achieved if there's just one central team jumping from issue to issue. And I mean that at a more kind of abstract level I think you just have to be kind of a general philosophy that no decision gets made without there being or if possible there not being any kind of analysis having done beforehand so that all decisions whether it's about the financials or the product design or whatever are made based on other things based on uh on some kind of analysis that lies on data. And yeah I would probably kind of I think it also be very useful to have workshops for members of the company to better understand some kind of statistical or mathematical concepts so that they can better understand what what can be measured and how things would have to be measured. Once they start testing things kind of proactively before they make any changes or with any developmental or resources kind of based on a hunch. I think this to be the main ones that I can think of now kind of.

You're covering quite all of the aspects. Very last question I know I probably every last one but. Do you personally think that all start-ups should be data driven to succeed?

I mean I think it depends on how you define a start-up. But. I'm sure there'll be a few exceptions but in general I think.

Or does this data driven mentality. Should it grow when a company when you reach a certain level for example?

I think that should be the right from the very start. Obviously it doesn't make sense you know overdoing it when you're just a company of three people or something. I think some kind of awareness of the eEven like the future value of past data having been collected. Even if the resources weren't there to analyze it at the time. Just having awareness of how valuable that historic data once the company has grown and the analytical capability and capabilities have grown is something that's probably too much and undervalued and should be done more. But in general yes.

Okay. So they should just recap. They should have data in mind. Yeah but it depends on the capabilities of whether they can use it or not. Okay, very interesting. Thank you.

I think I have got all of my insights that I was trying to get.

Transcript for Company H

Background

It & It Service management, more than 20 years in the field of IT
How would you describe decision making in your company? The structure?
Management team listens to every person who is involved in this decision. In most cases, it’s a consensus.

Are decisions made on someone opinion or do you bring facts?
We try to make decisions based on facts. But often it’s not possible. It’s quite hard to predict the future I might say. We base it on experience and personal feelings. When it seems like a good decision.

How important is data in strategy-making of the company?
Data is not so important. History is much more important. Why we’re here where we are. Seeing what were the right decision and wrong decisions in the past.

What do you use?
We use financial data. That’s all. We don’t have any other data. We can compare it to market trends or sector trends. There is a lot of studies on the internet. But every story predicts something, which one is the correct prediction no one knows.

What are the challenges of using data? Main problems?
The biggest problem is having no data. Because we don’t collect data. We can improve measuring work efficiency. For example, in customer service, we have tickets (service request, etc.) - based on that we could make decisions, plus financial data.

Where do you keep this customer data?
Ticketing system. It’s a standard Jira service desk.

Where do you get most of your data?
What data do you mean? I work with customer support. I can get most of the input data from Jira, because we measure how many incident, how quickly they are resolved. We can make some kind of decision. We can see who we need to hire, improve the system.

Do you think you have enough data to make better decisions? What are you missing?
No. To be more effective we need to know what to do. At the moment we don’t record, we don’t track the time. We don’t know where our developers spend their time and how. It’s not clear how much time they spend on each customer, how much time they spend on learning. Based on this data we could compare personal performance.

You know the lean concept? In our case, I can see we can do a lot to reduce time wasted. Do you buy any external data?
Sometimes we buy financial reports from companies. Sometimes we buy real-time maps.

Are you using this data enough, why not?
No. Because customers don’t know how to use it, so they don’t want it. We know how our customers are moving. Based on that they can make lots of changes in public transportations.

Who are your customers?
Cities. Transportation organizations. Those who organize public transport in the city.

Do you think if you had more time you’d make better decisions?
I believe that quicker options are better or equal to data-driven decisions. Because there is always so many open topics. You can’t predict them. You have to feel things based on your experience.

In your company decisions are made on data or feeling/intuition?
Intuition. Based on historical data and the feeling that “market is growing” - based on that we make the decision. We can’t be 100% sure, but we feel that negotiation is going well. Hopefully, we will get it. There is never a guarantee.

How do you share information in your organisation? Is it open/layers?
Financial information is accessible only for management. Any information related to products is public information internally.
We use two platforms: for product related is Confluence and for other Hr&marketing stuff is Sharepoint. 

Are they integrated in any way?
No.

Do you feel that your data is together or in some ways segmented?
Segmented by business units. I don’t see the problem at the moment.

You only use historical data. And based on the historical data you make decisions. You don’t use predictive analytics. What other platforms for information sharing do you use?
We also chat. Fleek for example. Somebody asks a question, somebody answers. We can refer to that.

How do you understand data-driven decision making?
To make some kind of decision before that you collect data. Then you analyse the data. And based on that analysis you make a decision. Then after some period you repeat the data collection and analyze again to check if there has been any improvements.

Do you have an analyst in the team?
No. In our ticketing system, we have a reporting module. We have specialists who report based on transport data, but not for our internal system.

How do you analyse decisions & assess them?
The business decision in the management meeting we do a status update and we see if we’re going in the right direction. There are no such KPI systems in our company. We predict that turnover will be this number, a number of new customers is that kind of number but the department don’t have KPIs. We check progress. Is it going in the positive or negative way and we steer in the right direction.

Is there room for improvement in the way decisions are made?
Of course, but it’s hard to change it.

What does it take to be successful when implementing data-driven decision making?
Management must lead the project. Can’t be that lower level employees will start doing something and it will be successful. Has to be a management decision.

It is not present in our company right now because of the company culture and management style.

How is your company culture?
It is open and sharing information, helping each other. Positive atmosphere. It’s quite a flat company. Everybody the same, everybody can speak to each other without fear.

Is it easy for your company to achieve targets?
One target is the sales target. Sales targets are fulfilled. Our challenge is implementing projects because of lack of resources, such as IT programmers or developers.

How is strategy created in your company?
We have a management meeting. We review historical data, previous period results and activities during the last 2,3,5 years based on that we check what was good, what was wrong. Based on that we make a new strategy. Every year we review. In the strategy, we don’t have KPIs. We have our main market, main customers, topics related to our products, do we need to add new sectors.

Do you think that if you had more information you could work more efficiently?
Process-wise, yes. We could work more efficiently. At the moment we only have feelings.

We have customer service (ticketing system), passenger and public transport usage & data about the busses, where they are and how many passengers are on the bus. We have Financial information, Sales & Marketing data. For external data, we have company reports, which help us decide if the partner/customer is reliable.

Where are your customers located?
Sweden, Finland, Lithuania, Ukraine,
**How do you decide where your customers are located?**

We participate in public tenders. To decide to work there we must have some kind of partner there. Because we can’t provide local on-site service, we need to use partners.

**Transcript for Company G**

**Data + decision making**

_DATA is a broad term. Decisions made in different settings. Some kind of data. What's your data source? Where do you get it? Where did you get the data?_

Gut feeling + decision making

Early days more on beliefs, as not enough customer data. About the belief the team has. Not individual belief, but rather what team as a group thinks. Now decisions are based on data, because we have real customer data. Trend is more and more. Since we have many target groups, do these assumptions hold through? Because we see a lot of different behaviors there. For me it’s not about decisions you make but it’s more about questions you ask about data. Everyone at the management level speaks up their assumptions, about what they think and then we try to prove/dissprove with data. More about rejection than validation. How do you measure if nothing is happening? Complex issue. Company decisions are based on analytics and data science, because the company is getting more mature. We have stability and historical data to base it on.

The faster the product roll out, the faster it is to build the gut feeling. Even if information is 80% correct, and 20% not, it is still enough

**Challenges for data adoption**

_Which number to trust? 1. What are the metrics you should be guiding? If they are fluctuating, then which number to trust? (2) what are the key metrics that have impact for the business? (3) What are the key metrics that have impact on the customer behavior?_

Understanding those is a tricky one. Because you are constantly developing your product, bringing in new customers etc. Some are leaving, some are staying. So understanding that space is the most difficult part. SECONDLY, basing your decisions on one person's opinion (due to his previous contracting experience) 3#CHALLENGE: Flat organisation, where everyone is using their own tools. Based on operational tasks. Once the tool has enriched the data, it does not get back to the central data warehouse and we cannot compare. DATA COMPLETENESS ISSUE. Accessing data on the third party tool, but not owning the data. Different team uses a different tool. Do we have the same numbers? Do we have the same definitions?

You never have complete understanding of data. We have all access

Challenges is to teach people and give the right tools so it would help to understand customer behavior and its impact on the overall business. Trickiest part is to understand which data will help customers be more profitable

Challenges: what does the product release mean for the user experience? How many new customers did we receive with the new release? How to grasp correct data from unbiased historical information?

Using correct data to measure the impact. For example, when a new app release comes our and users don't update the apps

_How to help manage all tools? Should there be 1 person?_

One person should be there to establish the connections, help own the data. Each team should have the authority to understand which tool works the best for them. They are doing a task that is important to them, so the tools should correspond. Allowing people choices for data/tools.

**Leadership (encouraging learning through analytics)**

#1 Data and analytics usage is about the people in the management, with analytical skills #2 Using data starts from leadership style and culture. Not only about someone saying 'we have to make decisions based
on data” EXAMPLE: our new head of product, with the background of McKinsey, bringing everyone on the table, bringing all knowledge and starting by asking "what is data?”. Usually we still think that data is numbers and trends, but also user insights and more. So, bringing everyone to the table and having a healthy discussion.

Using data starts from leadership. Education for data (mentality). It’s about educating people and giving them easy tools rather than to consider them incapable or not having an infrastructure in place.

**Culture**

Seeing questions, not answers. Being curious about data. He considers it a data discovery. It’s not about showing a presentation with dashboards and numbers, but more taking real time data and playing with it. Seeing questions, not answers. Being curious about data. He considers it a data discovery. It’s not about showing a presentation with dashboards and numbers, but more taking real time data and playing with it. MENTALITY: everyone could be at the table (referring to who is responsible for bringing analysts to the decision table). Also, it is important to participate and speak up during the time when the decision is being made.

We have a culture, where you don’t blame where you get it wrong. But you learn from it.

All departments need to be data informed. Decisions have to be justified (hard facts from data, market assumptions or discussions from the customers, comparisons from competitive offerings). Justified/customer driven/data informed.

**Revision of outcome**

(REGARDING NEW PRODUCT DEVELOPMENT AND REVISION OF OUTCOME) Product teams do retros and lots of learnings. In a week’s time, you see short term effects. We monitor if it had an impact. Proper AB testing to measure the impact on the new product updates

It’s not about measuring, it’s about the holistic view on the company, so that the whole company understands the impacts.

**Does sample size matter?** Yes. Good thinking from the data science team. Even one customer request can be impactful.

**KPIs**

Measurement is the trickiest part for setting correct KPIs. What matters is the biggest business impact. How could we measure company-wide perspective? Once the team knows they are measuring their own assumptions, they can learn about what was correct or what was incorrect. Outcome - learning for the next time.

Quarterly planning, annual overview

**Journey - DDDM**

Yes, probably we won't ever be able to reach the perfection, because it won't be possible to measure with such precision. However, it is possible to find out the impact of the product. What are the things that product features don't impact?

**Analytics**

Performing analytics for new products and modeling and forecasting each month. 6/7 on being data driven. We have weekly hands on. Every function presents data about the data to the entire company. Public TVS (already analysed higher level info). Every product team has access to already ready available information. Based on request, they can ask for additional information in Slack. If the data is not yet running, then we can update it and roll it out within the same day. Quick reaction to internal data needs.

#1 CULTURE Proper analysis team, so that every product lead has a dedicated product analyst. #2 CULTURE OF DATA DISCOVERY. You have nice and easy to #3 culture that after discovering these insights, they could ask more questions.

How data driven should the company be? We should be data informed, but customer driven. Customer should drive actions. Then the customer behavior could be measured. Customer driven, data informed. Building knowledge inside product teams, so that everyone can have analyst support. Steps 2+3 Letting people easier access to use that discovery and then building up the knowledge. How do you
share knowledge/experiments/decisions to 300 people? How did they learn and how to share this knowledge to all teams?

What do you collect? What don't you collect but are unaware of it? What decisions you make and what actions do you make? Everyone should have a data discovery. Each person in the company should have key persons showing how to do it in the company. This would enable data usage for everyone in the company. How do you cut down these data analysis people spend time on but not use for decisions? Wrong approach. They just do because they have to do, but don't start from a business question. WHAT DO I WANT TO IMPACT, CHANGE or MODIFY? Do we have enough data? Then they can learn from data and make these actions. Once this cycle is in place, there comes the learning.

Why do you exist in the company? To provide customer needs and satisfied customer needs. Everyone does not have all access to data. Each dept has access to information they need. If there is a need, access is granted.

HOW DATA INFORMED IS YOUR COMPANY? How complex is to measure and evaluate the impact? Data pool: for example, marketing person uses third party tools or has some metrics, then they have access.

DDDM can't be measured on a scale. We are well informed about where we attract the customers, who convert and sign up. We are not well informed about the intents f use. Because we have so many of these. Classifying and segmenting properly. 50% have access, but 50% have only public access. They don't have the need. If they do have the need, it is possible to grant it. Access to area specific information.

Transcript for Company H

Anyways I need to ask for your consent. Are you OK that I'm recording you for the purpose of my Master's research?

I have to say my full name, I name last name, I give you my full consent to record me in this interview.

Thank you. It's just formality. Anyways as I was telling you before that we're doing a research based on data driven decision making. And I thank you again for answering the questionnaire. and I would like to get a little more insight about how your business uses data in the decision making process. So first of all I would like to get your understanding of how important do you think is using data in decision making.

In theory immensely important, in practice. Not that much, quite often, unfortunately. Not that much of them used or important in practice?

It's one of those things which is good on paper but in practice does not apply. Perhaps it's connected to the stage of the company or whatever challenges you have. Or or whatever other excuses. Filling out the questionnaire I realized as well that we're not using data per say in so many decision points that we should.

How often do you use data?

What's the definition of data?

Data is any numerical, let's say facts.

So it's not only numerical?

No, it also can it can be... Even by or in this context apply to. H R?

Yes it's everything. It's like when you're making H R decisions then you know. I don’t know what are the best personality traits to have for a person to apply, so any kind of information that is kind of proven that he's not experience. That is not gut feeling is considered data in the purpose of this research. So even if the person. And employee is sick. That's a data fact. sick for X number of days. I have to take that into consideration.

Yes.
OK. If we expand like that then there's a news article data? News article about my market or my competitors?

Yes, it is.

Good. news articles affect my gut feeling, there's a fine line.

But there's like based on the news feedings you make a kind of like you and you feel like you don't do it like a per say a paper analysis but if you're let's say you read through different articles about your competitors you get the perception of what happens and then based on that making a decision is already database decision.

Now what was the question?

Yes. Why did you say that using data is not that important in decision making?

You know what I rather meant is that I consider data in the beginning before the interview that something numerical something big data like numbers more than larfer than 100. If we now narrow it down to just the facts and these kind of things as well, then I can say that yes we are using data every day. Hopefully as many business decisions as possible.

Basically if you make a decision, you get some kind of different information from outside, not just on your own experience?

Hopefully. I would. Yeah. I would at least hope that my other senior leaders are doing that. Hopefully. In theory, with a lot of things I can not check all of it. But yes, data being bases of scientific approach or in the start-up world, lean methodology or whatever. Yes it has its place. It should be used as much as possible.

Do you foster for your employees? That whenever you are deciding something get a second opinion or a set and second information source?

Depends on the seniority the employee. With some people I actually encourage them to make the decision based on their let's say 15 years of experience. I actually encourage to make faster decisions because I need to trust them. Even the younger ones who were kind of training at the moment. The more junior employee, the younger the employee, the more I would ask for more data backing of any decision or action there is a. Makes sense?

Yes. So basically, the more senior, you still encourage also gut feeling and expertise?

No encourage gut feeling but I've just seen that more seniors... They don't have to kind of defend and prove all of their points just based on their seniority they tend to make decisions based on data anyway.

OK. For your company, you're developing a product.

Yes.

When you do this. Just help me to remind. Is it one product and then you make let's say additions and modifications to it? Or is it? Or was it just. different products?

We have a very simple one product for all customers for essentially one price point as well. We steer away from customization or custom features. So... All the customers are getting almost the same product.

How do you make decisions based on what to? What modifications to do for the product?

In our company, we focus more on customer success. Their all job purpose is to constantly be in contact with our clients. To figure out what our clients' biggest needs, problems, questions. All of that. Then we try to gather that these two not into the spreadsheets always but be somewhere together. so that when we we have a decision point. Should we do it or not? We would have backing: "How or how many clients does this implementation this new thing apply? So that's almost like a requirement for us to do anything. There is no way that someone will have an idea based on about feeling and we would do that.

Ok so it doesn't happen like let's say product manager says and we might think that this is good and then you start doing?

Yes.

So the initial input still comes from the clients?

Yes, data from the clients basically.
But it's not us filling their requests but us looking at them looking at their requests and trying to figure out what is behind that. What is the meaning behind the ask?

**OK. so... I. could say that basically your product development is based on client data?**

Yes, but in all honesty I know quite a few let's say features or projects where were the initial data for the decision was in hindsight very minimal or misleading. Let's say a feature of that one of our clients asked that made a ton of sense to all of our team members. After implementation, we learned that it doesn't apply to the 99 out of 100 of other plans. So there was there was data, there was proof in the beginning, but I don't know. what was the word for scaling back on that was wrong it did not apply to all of us. But in general our product is based on data point.

**How often do you monitor the effects of certain products for your entire company revenues?** Saying that. "Okay now we did this. Did it actually have any effect on the big picture or not? Did it improve our selling points?" For example?

Weekly. Uh weekly in the context of if we look at data weekly, at least for marketing.

**No, I mean like do you put together product development with the impact on data? Specific product development. Or you don’t go in that deep?**

You mean how fast will the developers get feedback if the feature they built had an effect?

**Saying... Let's say you developed an extra feature that I don’t know gives you a unique selling point. And then would you see that after you launched the product, then the revenues went up or you are not looking at a product launch and then the impact on the performance of the company?**

In our company at this stage revenues are actually key performance indicator. We're looking for all sorts of other engagement data points. So we're looking at engagement.

**Ok... so engagement effort. So my question again. Let's put it instead of revenue, engagement.**

**Okay. How often? Do you?**

Yeah. Yeah. Of course...

**Basically after each product update, you see what is the impact on the engagement?**

Definitely.

**And based on that do you make modifications to evaluate what went wrong what's didn’t or no? How is that process going in the company? How do you evaluate it? What was the impact of those decisions that were made?**

On the bigger picture our first question is like did clients even who requested did they engage with the new feature or product? Yes or no. Then we do numbers. It can be anything from session duration to yes or no do they use it? Or even do they know about it? For example do they open the newsletter that talks about that new launch? Then then then then, their hardest part is to find the key performance indicators breach for this product. But, for example to set what is the threshold when something is successful or not. Got to say that's not very scientific like someone just figures out guesses like 'Hey, 80 percent of our customers are using it then it's a success'. Do we make any changes based on this? Yes. For example, simplify. If numbers are low, we have to change something. If we have we have implemented, and we have for some products, projects more in-depth engagement; Then I think how many times is Then we can look into deeper and specifics on the granular stuff like. Button working years this workflow. Does it make something simpler? Slower? Faster? Or Does it even work? We haven't had cases when we have made stuff works.

**Last question before we have lunch. If you discover that a decision that you've made didn’t have the effect you wanted based one the data. Do you change the decision-making process?**

Yes. That's a quick guess because I already know that most of our semi-failures failures or stuff that were very much criticizing ourselves has been the cases where someone did not take data on into the planning as well as much as they could have.

**So you asked them to widen the scope of data? Or how do they change? Just ask them to use data?**

So, the biggest failures are when you didn't use data. And that's why you asked them to change. For sure, for sure.
Let's have lunch. now. Thank you.

Anyway so let's carry on. And. I liked the small discussion online. I would like to ask you what is data driven decision making for you?

When I first got the questionnaire from you guys, I would have said that my company is not data driven, because when I hear data driven I think about big data. I think about Google or Amazon running tens of thousands of this based on millions of data points in real time. Or. any other big company doing A B test with significant volume. But. let's say my company or any other smaller start-up making decisions based on one or two facts about the market. I wouldn't even consider that data driven. Just... because...data means a lot of data

What is a lot? Please define what is a lot in your head.

When I think about my own clients and we have like with a right now more than a hundred business clients I just said the other day to one of my one of my people that before any number is at least a hundred, do not talk to me in percentages. Like growing from one to two is not 100 percent growth. You just got one client more; so what is what amount is data? At minimum 100 points of data. But now as I say it out when we are gathering customer feedback we definitely do not have the milestone threshold or benchmark that it always has to be at least 100. So it depends if we’re doing these kind of interviews like you're doing with me with clients and where do you get all the time. Then we rely quite a lot on the smaller data sets. Let's say five interviews with five different clients. And that's completely fine. When we are talking about analyzing our marketing traffic on the website, we’re talking about tens of thousands and this is a small number. This is not a lot of page visits.

Does that answer your question?

Yes that actually leads me to... you kind of answered my next question...I wanted to ask, if you consider your company as data driven? You said that OK before you didn’t it. Now, based on the discussions?

Well it's kind of like so flattering that "Oh...it turns out we are". But I wouldn't consider us that data driven like I know how much improvement there could be. And especially as half of my people are developers. They would laugh me out, because, let's say even last year we had a machine learning project. Where we had more than.... God dammit, what was it? More than 40 million data points about something. That wasn't enough. That's that's not big data, at least. So very few years back there was a trend about startups doing something something big data. Serious developer laughed out most of them as currently the trend is about let's say machine learning or A.I. Most of the companies saying that they’re doing machine learning or A.I. are not for scientific like a heavy scientific person. They’re not actually.

But if you now exclude this big data from the formula. If you would need to evaluate the maturity of being data driven, for your company. How would you evaluate it based on a scale of one to zero to 10.

Still five. Meaning that I know we can do twice as much. Definitely more. But now with the specification that's just a fact or just the number can be data as well. Then we at least are incorporating that into all of the decision making processes.

Is it only data volumes that keeps you from saying that you are a 10 or are there any other aspects of the entire process. Do you see that prevent you from being very well established?

It's the classical balance between quality versus quantity. We have tens of thousands of end users using our software, answering to a simple question. And we can make some decision based on that. At the same time, like an in-depth interview with that decision maker can be as a reliable or even more insightful at times. Something else?

I was more hoping that you would talk about that you see something about processes that could be improved or underlying technologies. That you see that you need to become excelling or data driven.

To understand correctly, you're asking where should we incorporate data more?

No. My question was that...you said that you give yourself five points out of ten. What else besides that you don’t have big data volumes, You don’t want to give yourself ten.

I see.

That was the question.
Mhm...

Or is it only you due to the data volume that you don't want to evaluate yourself as ten?

Ahh...OK. Quality versus quantity. That was my point that for some decisions we just don't yet have enough data points. On other decision, the data we have, the data integrity they might be misleading. So the problem being that we, meaning let's say a lot of startups have. That's the biggest challenge for me as a leader to make decisions based on very minimal data at all. I would say that's one of the hardest things though being a startup. If you would have a perfect business plan with data points, then everyone can do it. But that's not the case. And then you have to make decisions based on...I don't want to say gut feeling because it's not gut feeling, but in many cases you might...you might be looking at...you might be searching for data, facts, to prove your point and you might be disregarding something that still disproves it.

I still haven't answered your question why don't we consider us 10 out of 10. Because of speed perhaps as well. A lot of the projects to get some kind of data we're not agile enough. So, for example, developer is focused on one project for two months. I would say that there is a lot more than we could to where they could agily test ideas, features on the go. We're still in some sense doing stuff waterfall. Meaning, we have a plan, we do use data to come up with a plan. But then it kind of results positive or negative will only come after the development after the launch. Then after gathering data so we can talk about months of the time. If we were faster, people would be more agile. We could in theory incorporate maybe even more data as well. Why aren't we 10 out of 10? The HR and human stuff is hard to put in numbers. Although we do incorporate for example number based evaluation scores for employees as well.

But then...But then you have the problem like how well you...can you evaluate the extrovert marketing person the self-assessing themselves versus a very introverted developer assessing themselves and thus the data is not comparable.

Well you mentioned that you gather data from interviews to make decisions.

Yes.

Are there any other aspects that you are not currently doing that could help you get more data for the bases of your decisions?

Definitely. By game theory in a perfect world we would have technically connected to every action, blow, page whatever we have in our software product. In practice we have seen that we're getting faster results by what they call interviews rather than implementing the heavy duty measure everything. Because that requires a lot of development work and someone has to check the data. And even if the data is there we have a lot of raw data about everything. I don't know...No one is responsible for that. No one has a hypothesis to prove or disprove. Then the data is worthless. You have to do something. I think you don't have a clear goal if you don't have a clear thing what the measure then it's worthless.

Do you think that in such cases as you described there should be only a certain specific person responsible for gathering the data and evaluating the data and making edible or for all of the others?

Or should it be something that all of the members of the team should be responsible of doing and using and interpreting?

We work with data data scientists and data analysts. I don't see how one person can be like miles ahead of everyone else. Like. Lets say...these these kind of people showed me what what should be done what can be done. Which was mind blowing. But I would still say that one person responsible for this would create the bottleneck. Rather to incorporate the mindset of "try to find data for your decisions" to everyone. We are definitely doing that even if their beginnings are negative, stuff like that. That would be my approach. To try everyone to get if not data driven, at least to get them data informed.

You as a leader how you can make your team like that? That they are data informed, data driven?

Definitely by example. We in our company would definitely incorporate a lot of transparency. We're sharing everything from revenue numbers to clients numbers to marketing numbers to product and development and the engagement numbers to everyone. I've seen that it helps a lot. People value and care about that even if they don't focus on. Someone has to set an example for others to take lead. So if you see someone making a decision and it's data based, they are less inclined to have a guess. The next day then they definitely want to do that as well.
How is data managed in your company?

Yeah sure. If I could only answer that. Depends. Depends on the department. I would say. Marketing department has their own set of tools that actually other departments and people have access to as well. We have not the policy but just that tradition that all of the departments so to speak although we're very small or the people responsible for some things share their numbers every week of it to everyone. And that does a lot. It shows transparency, it informs others. The question was how do we manage our data. Well of course we have different databases. I think one of the questionnaire a question was what kind of database we use. Don't even know.

What form of data is available to all of the company? Is it the raw data itself is basically accessible to each member of the company and then they can use that data for whatever decisions they need to use it or they need to go through certain people to get their own data? Or do you already only share the analyzed data?

We've got a lot of top level stuff like perhaps how many clients? What is the revenue? This kind of thing is automated then it all automatically feeds into our Slack channel, our communication platform. Then a lot of top level stuff like specific product features. Then I would describe the process that non-technical person perhaps from the customer success myself or marketing will ask the developer "could you please give us the last 30 days clients engagement with this button?" And then the developer would make the query based on the complexity it would take X amount of time. And furthermore, I would describe that when the developer gets annoyed how many times he has been asked something, he will not commit it and then we're going to have another automation. No, not all of our employees have access to the raw data

But say based on inquiries they can have, if they need it for their work?

Yes. Yes yes. There is no locks there are no limitations. Of course, as our company is dealing also with our own clients' private data. We have some security forces there but i.e. engagement, that's fine.

Okay. And in terms of interpreting numbers...If. you as a manager are given a full or already done analysis? How often do you question the underlying data?

I would hope that every time. Like...

And how do you do it?

I definitely want to see the logic behind the number...just to be sure that it's valid. Or that the first assumption for the data was correct or at least something that I would agree with. Only then can I be certain though about the number. Like I need to have data integrity and data validity. And that's definitely again something that I would encourage my own people to do, to have the understanding where the numbers are coming from, just not to take them as point blank accurate but ok if the marketing department says we have X number of traffic or leads. We have seen many cases where they'd say the more mathematically inclined developers attack question, the data. And because of that we have found weaknesses or we have improved the quality of the data for sure. Definitely definitely people should be certain where the numbers are coming from or are they legitimate?

I don't jump a little bit away from your company. Technology driven company is quite easy to be data driven because you gather data points all of the time. If you would need to suggest to let's say more traditional company, which wants to become data driven. Where and with what should the start? In your opinion?

So that's actually something that our own product does do. That is one of our value propositions were essentially summarized one of our key value propositions and human error is that we make physical face to face service digitalized and thus measurable. Thus our own value proposition for any non-technical company is "You get a measure, you've got to measure everything" or I think everything is too bad to say, but as much as possible. One of my main challenges with my own clients for non-technical companies is that they think they can do business by a gut feeling. I got a lot of contact with people who have been let's say industry experts in retail for 10 years. They claim that they know the seasonality that trends of the market do everything by gut feeling. A lot of times then to know a lot. But now...I'm getting more and more feedback to my own product that as our product showing them the data has revealed a lot of surprises. So I would claim that. I would say that almost everything can be put into numbers. Of course some things are harder to put into numbers and some things can be misleading. But... as the saying goes what gets measured gets managed. So whether it's something to do with people, which stereotypically is thought to be hard to measure... or something bigger where I say there's too many moving parts of it here
at all. You know everything can be better and it's becoming easier and everything. The challenge I would say for most companies is how to interpret the data. They don't know what to do with it. That's why all of the big players such as Google are giving their own assumptions and suggestions based on what they're learning and people are bad with numbers.

**What is your opinion on how to improve this matter?**

The easiest thing is to start holding people responsible for an X number, like an X KPI. That's the simplest thing. But then the very next problem right away is that people get too focused on their number. If that's a specially connected with their salary or performance reviews and then they disregard or forget all of the other aspects that are important in their business or in their specific jobs. So they will start to cheat as well. They will start to focus too much on the one number and forget all of the others. How to get people more data driven? Of course positive feedback is always good when you see that some kind of positive impact to an organization has been done based on the data driven decision. That's a big point. Or if they see like their competitor or another similar company... something that they can relate to...is actually DD...

I would guess that then they are more inclined to test it out at the. And. I would still say that everything can be can be and should be measured as much as possible.

**Last question I would jump back again to your company. If you had have to evaluate is the leadership of your company data driven?**

I would hope so. but trying to be honest transparent and self criticizing. I feel that leadership is very different. They're the ones accountable for the numbers. In many, in many cases many times I actually found myself in a position where I have had to say to a manager. That yeah. Okay. Your numbers are good but something that we're currently not measuring let's say your team is a distraction or something like that is perhaps could need some input. So I would say yes the management is there.

**Would you say that the culture of the company fosters database decision making?**

I would say so. The main reason the main driver behind it being that. more than half of our people are developers does they have either scientific, physics, math, informatics background. These people just tend to be more numbers orientated than everyone else. Even our even our marketing lead is from a very mathematical background. Yeah I guess these kind of people just brought it in and now it's there to stay. And it works, definitely.

**And the technology?**

What do you mean by technology?

**Your own technology that you work with. Is it something that fosters being data driven or... are there several aspects that you would need to let's say update or change, that would enable your company to be more data driven?**

So our own product that we're selling to clients' is a value proposition is we are giving you the product.

**Not the product, the systems of the company... the tools that you use...**

Are they data driven? Do they provide us data

**I mean, do they enable you to be very data driven? Or is there room for growth?**

There is always room for growth but I would say that tools that we do use. The trend in the market with every tool whatever you do is like a customer, customer support tool we use or our marketing tool we use all of the other companies are showing more and more data. I just coincidentally just before the meeting was looking at the keynotes from Google Cloud someone important saying that that's the trend that they're seeing in the short term market as well. All of the companies that used to used to just digitise something let's say payments with the mobile phone or. whatever. The simple analogy being they put menus in iPods. All of these companies now are going into to big data analysis business intelligence and giving recommendations, insights into the data that they have already gathered and thus they can make some kind of prediction based on that. As we use a lot of these systems. Yes, a lot of the systems aren't telling me how to run my business... which is helpful.

**Is there something that I did and that's what you would like to tell me about this topic, DDDM?**

I guess the first part was the most helpful and would be helpful for everyone else. That data driven does not mean base data per say right away. But even if you measure a number somewhere every day consistently that's a good place to start. It's connected to the one of the questions you have "How can
other companies be more data driven? They just have to realize that they might already have the numbers they need. Look beyond the revenue numbers. Don't rely only on your accountant. You can do a lot already with the numbers you have. You don't have to be technical for that. You don't have to do math. You don't have to sum it up. I'd say you just have to search her patterns and trends in the numbers. And that's pretty easy to do.

*Thank you. I think they got quite nice information.*