CHALMERS

Creating documentation of a data visualization tool

- A Study of Process and Contents Levels

Bachelor's thesis in Science in Engineering

JOHANNA BLOMQUIST
SEBASTIAN VALLBO

DEPARTMENT OF TECHNOLOGY MANAGEMENT AND ECONOMICS
DIVISION OF SERVICE MANAGEMENT AND LOGISTICS

CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2023
www.chalmers.se
Report number E2023:051

Report number E2023:051

Creating documentation of a data visualization tool

- A Study of Process and Contents Levels

JOHANNA BLOMQUIST
SEBASTIAN VALLBO

DEPARTMENT OF TECHNOLOGY MANAGEMENT AND ECONOMICS
Division of Service Management and Logistics
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2023

Creating documentation of a data visualization tool

- A Study of Process and Contents Levels

JOHANNA BLOMQUIST
SEBASTIAN VALLBO

Report number E2023:051

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

Gothenburg, Sweden 2023

Gothenburg, Sweden 2023

Creating documentation of a data visualization tool
- A Study of Process and Contents Levels

JOHANNA BLOMQUIST
SEBASTIAN VALLBO

Department of Technology Management and Economics Chalmers University of Technology

Abstract

The aim of the thesis is to create documentation on a data visualization tool for a case company. This documentation is used to allow more employees to learn how to visualize data themselves without having to go through the metrics team. The creation of the documentation is analyzed on both a process and content level, and the data is retrieved through a qualitative research design. On the process level, the focus is placed on factors for maximizing user engagement of the documentation. On the content level, the factors when visualizing the data desired are focused, in order to best document the process of doing so. The factors affecting the process level and the content level of the documentation are determined through analysis of relevant literature, secondary research and through qualitative interviews with relevant employees at the case company. For maximizing user engagement, it was found that the documentation has to primarily address the process of visualizing the data desired, both in regard to relevant literature as well as how to do it using the software, while also presenting itself in a way simple enough to make navigating the documentation more efficient and less time-consuming. For the factors to consider when visualizing the data desired, three criteria of selecting appropriate graphs were deemed most relevant and applied to the desired data accordingly. These criteria are the structure of the data sets, the intended use of the graph and the research question of the graph. Additionally, several data visualization guidelines found in literature were also considered when visualizing the data desired.

Key words: Data Visualization, User engagement, Process level, Content

Acknowledgments

The thesis work was carried out at an anonymous case company. We would like to thank all the people we have met at the company for the warm welcome and the willingness to cooperate and to help us navigate their systems. We would especially like to express our gratitude to Cornelia Richter, Ulf Damberg and Sonja Novakovic, for providing us with the opportunity to carry out the thesis at the case company, while providing great support. Likewise, we would like to thank all the interviewees for their time and participation. They made the outcome of this thesis possible. Lastly, we would like to thank our supervisor and examiner Hendry Raharjo from the Department of Technology Management and Economics, Chalmers University of Technology for raising the quality of the thesis work as well as supporting us with interesting and valuable insights during the thesis work.

Johanna Blomquist & Sebastian Vallbo
Gothenburg, May 2023

Table of Contents

1 Introduction

The following chapter aims to guide the reader into a deeper understanding of the thesis, by providing a background regarding the subject and a description of the case company*. It is followed by the research questions and purpose, and a delimitation and thesis outline.*

1.1 Background

The case company is a major player in the Information and Communication Technology industry, specifically focusing on providing products and services to providers within the service business, such as telecommunications companies. Their goal is to maximize the potential of connectivity by creating innovative and user-friendly technology and services. This technology and services can be easily adopted and scaled by customers, helping them to be successful in a world where everything is connected.

The case company is currently facing challenges related to data visualization in regard to efficiency and simplicity. Many employees need fast and flexible data in their daily work, but the current way of getting data visualizations through the metrics team is considered by many to be unnecessarily time consuming. About a year ago, a new data visualization software, *EazyBI*, was purchased by the company for the purpose of simplifying data visualization processes for more employees. While this software is integrated with the database and technically available for all employees already, many lack the time and resources to learn it, and almost all are even unaware of it. This has led to the tool not yet being integrated in the daily work of employees who would benefit from it.

1.2 Research questions

The following research questions will be central to the work of the thesis:

- What factors to consider for maximizing user engagement?
- What factors to consider when visualizing the data desired?

1.3 Purpose

The purpose of the thesis is to create documentation of a data visualization tool, using both a content and a process perspective, for relevant employees at the case company. We will analyze how to optimize the structure of the documentation and how it is presented in order to further maximize the chances of user engagement. Furthermore, the documentation will provide examples of data visualizations of key metrics as well as information about how to create them using a specific data visualization tool.

1.4 Delimitations

This thesis presupposes that the data visualization tool of choice is a good enough tool to create documentation for and will not discuss whether or not it actually meets the customer needs expected of a data visualization tool. We will assume that it does based on the case company's creation of the thesis, and instead only focus on the creation of the documentation for it. Furthermore, it is important to note that this thesis will not be directly involved in creating dashboards for strategic decision making, nor will it be involved in the development or the improvement of any features of the company's products. The data visualizations performed in the thesis will instead be created for the purpose of acting as examples as part of the documentation. Additionally, due to limited time, the documentation will only be able to address the visualization of some data of interest and not of all customer needs identified on a content level. Furthermore, the terms customer needs and user needs will be used interchangeably throughout the report, due to literature surrounding both being applicable to the employees at the case company. The reader is therefore encouraged to have that in mind while reading the thesis. Additionally, the terms customers and users refer to the employees of the case company whom the documentation is intended for. Lastly, it is important to note that, due to confidentiality restrictions imposed by the case company, the documentation will not be published in this report. Instead, the report aims to describe the processes involved in creating the documentation and the factors to consider. This approach allows for a wider range of readers to find the report useful. However, the documentation itself can be provided upon request.

1.5 Thesis outline

The thesis report starts with an introduction of the aims and background of the thesis and case company, followed by research questions, purpose and delimitations. In the following chapter (Chapter 2), the methodology is discussed so the reader can follow how the information in the thesis is conducted and assess the reliability. It is followed by a theory chapter (Chapter 3) which provides the reader with the required knowledge and lays the foundation of the thesis. The following chapter (Chapter 4) presents the results based on the interviews. The results are connected to the literature study in the discussion chapter (Chapter 5). The thesis is summarized with a conclusion in Chapter 6, including a discussion of limitations and future work.

2 Methodology

In this chapter, the research methods used in this thesis are presented. To start with, the research setting is described to emphasize the problem and the employees' dissatisfaction with the current data visualization process. Subsequently, the research design is outlined, which serves as the foundation of the entire study. Further on, we highlight the various data gathering methods used to collect the necessary information. This was mainly done by literature study but also through qualitative interviews. These methods were carefully selected to ensure that the data collected was comprehensive and reliable. To conclude, we delve into the secondary research process and finally discuss the ethical considerations that were taken into account. This includes the measures taken to maintain the confidentiality and privacy of the participants and ensure that their rights were protected throughout the study.

2.1 Research setting

Many employees at the case company believe that their current data visualization process is lacking and needs improvement. The metrics team, which performs most of the data visualizations currently, primarily does so for upper management on a regular basis. Meanwhile, those in more operational management positions, such as program and project managers, have to put in special requests to the metrics team in order to retrieve data visualizations. This process is reportedly unnecessarily long and time-consuming, even though these types of managers arguably would benefit the most from data visualizations. Many employees at the case company who work within these types of positions all describe a need for data visualizations in their daily work, but complain about the slow process of retrieving them. As described earlier in the report, a new data visualization tool was purchased about a year ago for the purpose of simplifying the data visualization process so that more employees could do it themselves rather than going through the metrics team. The tool itself was deemed sufficient, but its existence and usage were not well-documented, which led to it being forgotten, unintegrated, and unused. When awareness of the tool's benefits and potential were raised to several levels of management at the case company, many supported the idea of creating company-tailored documentation for it. The case company therefore created this thesis, which requires two students to learn the data visualization tool for the purpose of creating company-tailored documentation of it.

2.2 Research design

Creswell (2009) describes three different approaches when selecting and conducting a research design method. These are qualitative, quantitative and mixed methods. A qualitative approach implies exploring, understanding and analyzing more human and social aspects, while quantitative methods imply the usage of mathematical and statistical procedures of measurements to answer more predetermined questions and hypotheses. Lastly, a mixed method implies the use of both qualitative and quantitative methods. This thesis has been conducted through a qualitative method research design. Creswell (2009) describes qualitative research as a method that is concerned with exploring and understanding the subjective experiences, meanings, and perspectives of individuals or groups. This approach emphasizes the importance of context and the researcher's ability to understand and interpret the data in a way that captures the complexity of the participants' experiences.

Furthermore, two different perspectives will be prominent in the thesis report, content and process. The content level refers to the specific information or material being conveyed, while process level refers to how that information is being conveyed and the methods used to communicate it (Pettigrew, 1987). In essence, the thesis report will cover both how to visualize the data of interest for the case company, as well as how the information on how to do so will be conveyed.

In this report, qualitative insights and data are collected from interviews with management and end-users to determine interests regarding data visualization within the case company. Furthermore, the actual data of the key performance indicators of interest that will be visualized is retrieved through preinstalled interoperability between their database and the data visualization tool of choice.

There are two ways to approach facts, either inductive or deductive. Bryman and Bell (2015) refer to the principle of deductivism as a theoretical approach that aims to produce testable hypotheses that will allow explanations of laws to be assessed, while the principle of inductivism is described as a theoretical approach where knowledge is attained by gathering facts which serve as the foundation for establishing laws. Put simply, in the deductive approach, theories serve as the basis for observations while in the inductive approach observations inform the formation of theories. The latter approach was utilized in the extraction of qualitative data through interviews and other qualitative data gathering methods. This approach leveraged observations and both implicit and explicit knowledge to develop a theory that accurately reflects the functioning of the department.

2.3 Literature study

The thesis has mainly been supported by a literature study, which laid the foundation of the work. The literature study is accomplished by a mixed method, meaning both qualitative and quantitative research. The difference between a quantitative and a qualitative literature study is that a qualitative approach divides literature based on themes while the quantitative approach divides the literature based on measurable variables (Creswell, 2009). Creswell (2009) further explains that the qualitative approach acknowledges the research problem without constraining the views of the participants, while the quantitative approach additionally suggests possible questions and hypotheses that need to be addressed, implying a more restricting and predetermined framework.

The literature is mainly found at Chalmers library, Google scholar, Web of science and Scopus. To find relevant articles and literature in databases, there are a few key words which have been used recurrently. These are: *Data Visualization, User engagement, Process level and Content level.* The aim of a literature review is to show awareness and understanding of already published work within the same field as our research project. It will help to address what is already known about the area, select relevant concepts and theories and evaluate any controversies (Bryman and Bell, 2015). According to Backman's (2008), there are multiple elements that should be considered when conducting a literature review. Firstly, it should provide an overview of the previously collected knowledge in the respective field as well as indicate the current research front. It should also highlight potential problems such as knowledge gaps, contradictions and failures along with demonstrating the significance of the research problem at hand. Furthermore, a literature review should aid in problem formulation and the definition of concepts as well as offer methodical ideas, designs and procedures. Lastly, a literature review should also offer a historical perspective to provide a broader foundation for the thesis to build upon. The aspects have all been taken into consideration when conducting the literature review.

2.4 Interviews

To find the underlying need for learning data visualization for the actors who are involved, interviews were conducted with key persons within the case company. Interviews as a qualitative research method is about collecting insight regarding people's social reality and attain an understanding of the problem from the interviewees point of view (Dalen, 2015). Before conducting the interviews, the right persons to answer the questions were selected and asked to participate. To find the people with direct interest in the research questions, a proposive sampling method was used (Bryman and Bell, 2015). The method aims to consciously choose a limited number of individuals from a larger population to participate in the study (Bryman and Bell, 2015). In total, 5 interviews were conducted with managers at different departments within the organization. A descriptive summary of the interviewees is available in Table 1. With this method managers within the area of business intelligence as well as individuals with interests in integrating data visualization in their daily work have been selected. The interviews were semi-structured to encourage an open flow of thoughts and ideas (Bryman and Bell, 2015). To maximize the amount of information gathered, follow-up questions were asked based on the responses received and the variations in answers from other interviewees. To ensure that no information was lost, all interviews were recorded and transcribed.

Table 1: Interviewee's profiles

The five interviewees	Experience with EazyBI	Age	Length of interview	Month of interview
Interviewee X	Very experienced	40-60	30 - 60 min	February
Interviewee Y	Inexperienced	40-60	30 - 60 min	March
Interviewee Z	Inexperienced	40-60	30 - 60 min	March
Interviewee A	Inexperienced	40-60	30 - 60 min	March
Interviewee B	Inexperienced	40-60	30 - 60 min	March

2.5 Secondary Research

Secondary research, also referred to as desk research, involves the collection and analysis of information from existing sources (Bryman and Bell, 2015). Secondary research's primary use includes gaining a general understanding of a subject, identifying existing knowledge gaps, and obtaining data to support research projects. Compared to primary research, desk research is often less time-consuming and less costly as it draws from readily available information without the need for new data collection. However, it is crucial to ensure that the sources used in desk research are both credible and relevant to the research question or problem under investigation (Bryman and Bell, 2015).

Confluence was studied and learned since it is the platform where the documentation was going to be published. Additionally, a crucial aspect of the thesis was to gain a general understanding of the case company´s visualization software of choice, *EazyBI*, as well as how the database used, Jira, is constructed. The secondary research that was primarily used to understand the *Jira* database was conducted within Confluence. The secondary research that was used to understand and learn *EazyBI* was conducted through online research and assistance from experienced professionals.

2.6 Ethical Considerations

The ethical aspect of a research project is important to consider in order to ensure the validity of the more qualitative data collected such as the interviews. In the context of conducting a research project, it is critical to bear in mind four fundamental ethical principles. These principles are: Obtaining informed consent, Protection of privacy through confidentiality, Avoidance of harm and Preventing deception (Bryman and Bell, 2015). Obtaining informed consent is the process of obtaining agreement from participants to participate in the study and to have their data used for research purposes. It is also crucial to avoid any form of harm to participants, be it physical or psychological, during the course of the study. Protection of privacy through confidentiality implies the protection of unauthorized collection or use of personal information, while deception involves withholding information or presenting false information to participants. All the four ethical principles help build trust towards the conducted research, helping to ensure the data provided by the participants to be sincere and honest (Bryman and Bell, 2015).

3 Theory

This chapter aims to cover theory relevant to creating documentation of a data visualization tool, both on a process and a content level. Theory surrounding the customer need, standardized work process as well as different frameworks for predicting user acceptance, will all be used to support the creation of the documentation on a process level. Furthermore, theory regarding data visualization will be used to support the creation of the content level of the documentation.

3.1 Customer need

The following section aims to delve into the tools and frameworks regarding identification and prioritization of customer needs. The theoretical knowledge provided will be used to subsequently identify the customer needs of the documentation on a process level.

3.1.2 Hierarchies of needs

Olsen (2015) describes the term *hierarchies of needs* to refer to the idea of customer needs having dependencies on one another, and how addressing a certain need is a function of how much another need is met, therefore creating a hierarchy. One example of this that Olsen (2015) refers to is Maslow's hierarchy of human needs which is illustrated in Figure 1. The hierarchy divides the needs into different levels- physiological, safety, love, belonging, esteem and self-actualization. Put simply, the hierarchy implies that higher-level needs may not be as important or meaningful to an individual unless the lower-level needs are also met, so when exploring the problem space for a product, the lower-level needs within similar types of relevant hierarchies should be addressed firstly (Olsen, 2015). Furthermore, when looking into customer needs it is important to understand that needs are taken into account when buying a service or product as well. Customers expect it to be safe and useful but they are usually not able to express the need themselves and (Bayus, B.L, 2008).

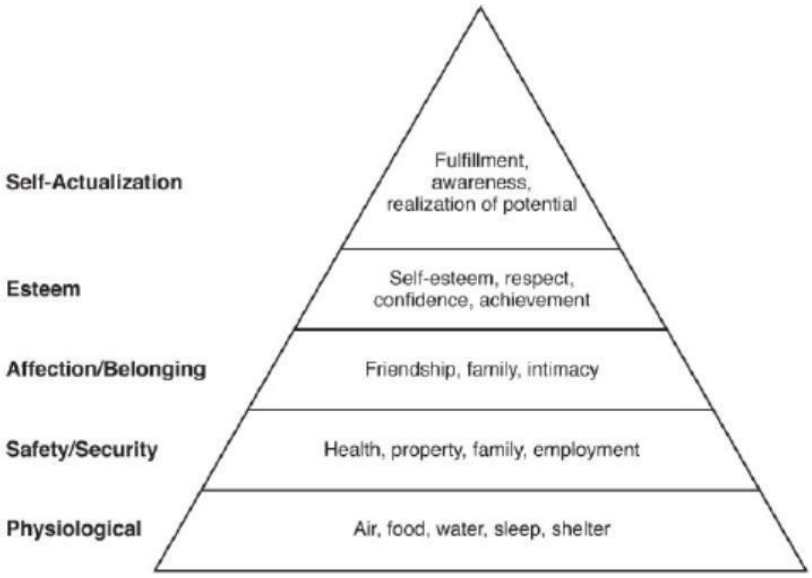

Figure 1: Maslow's hierarchy of human needs (Olsen, 2015)

Developed on the foundation set by Maslow's hierarchy of human needs, Dan Olsen designed a new hierarchy based on the needs of web users (Olsen, 2015). Olsen's hierarchy of web user needs, which is illustrated in Figure 2, aims to delve deeper into explaining customer needs regarding specifically online software.

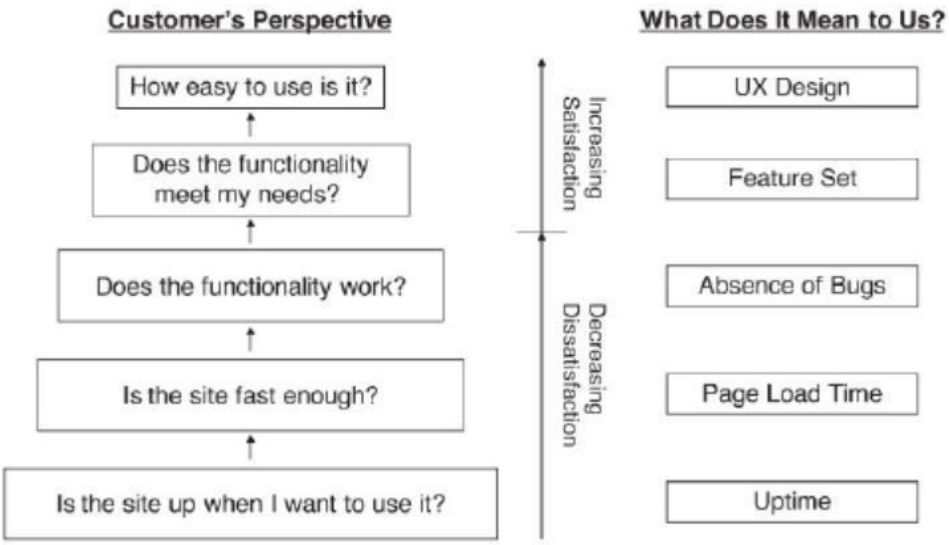

Figure 2: Olsen's hierarchy of web users' needs (Olsen, 2015)

The first level of Olsen's hierarchy is about making sure that the software is available when the customer actually wants it (Olsen, 2015). The second level is about the speed of the software having to meet the requirements of the end-user, and the third level is about the software actually having to work properly. These first three levels need to match the customer's criteria in order to decrease dissatisfaction, and while they are important and often forgotten when perfecting an IT-system, they alone will not increase an end-user's satisfaction with it. What will, however, is the quality of the software, which is divided into the top two levels of Olsen's hierarchy (Olsen, 2015). The first level of

the quality and fourth level of the entire hierarchy is about whether or not the features of the software meet the actual needs of the users. The fifth and final level of Olsen's hierarchy is whether or not the software is easy to use and to what degree it is enjoyable to use.

3.1.5 Identifying customer need through interviews

Olsen (2015) describes a set of questions that help clarify the purpose and benefits of a product to potential customers as well as help identify what and why customers find certain benefits particularly valuable, which he describes is a great way to gain further understanding of how customers think and act. Firstly, the interviewer has to formulate a clear benefit statement regarding the product and then ask the customer what this statement means to them and how it might help them. Subsequently, the interviewer also needs to ask how valuable a product could be for the customer if it actually delivered this benefit, and why that would be the case? As it has been mentioned above, these questions not only help identify customer needs but also clarifies the purpose, benefit and potential need of the product for them. Even if their answer is that it would not be valuable to them, Olsen (2015) explains that asking why still helps you understand what really is important to the customer. He additionally brings up the customer benefit ladder, which can be summarized to repeatedly asking "why is that important to you?" in order to identify the underlying customer needs on a higher level, and therefore also identifying what needs should be met in order to achieve higher satisfaction.

Furthermore, Olsen (2015) explains that the prioritization of customer need could be done through identifying customer value, and he goes on to provide a framework that could be applied to achieve that which can be viewed in Figure 3.

Figure 3: The Importance versus Satisfaction Framework (Olsen, 2015)

Olsen (2015) explains that the importance versus satisfaction framework is based on the idea that the level of importance a customer places on a need, and how satisfied they are with the current solution, are essential factors in determining which customer needs to address to create customer value. As seen above, it is essential that the customer finds the need to be important to them, but depending on how satisfied they are with the current solution the customer need is to be addressed in different ways. For example, Microsoft Excel has set an industry standard in regards to spreadsheet applications (Olsen, 2015). This makes the customer need for spreadsheets much more difficult for new companies to address since the users are already satisfied with the current solution available,

i.e. Excel. On the other hand, Olsen (2015) describes that the company Uber is a perfect example of taking the opportunity to address a customer need where the customers were unsatisfied with the current solution. Customers who needed transportation were unsatisfied with for instance the lack of affordability and convenience with current transportation services, and Uber was able to deliver on this underserved need. This upper left quadrant of the framework, where there are customer needs but also low satisfaction with the current solutions, is described by Olsen (2015) to be offering excellent opportunities to create customer value and therefore are the types of customer needs to prioritize and address.

3.1.6 The Kano Model

Another framework for understanding customer needs and satisfaction that Olsen (2015) describes is *The Kano Model*. It was developed by quality management expert Noriaki Kano in 1984 to measure the importance of product or service features (Olsen, 2015; Rådman & Johansson, 2020). Put simply, it can be used to directly measure customer needs, and it has helped quality management, marketing, product development and research & development extensively with creating offers that align with those customer needs (Rådman & Johansson, 2020). The model plots one parameter for the horizontal and vertical axes respectively (Olsen, 2015). The horizontal axis displays how fully a certain customer need is met while the vertical axis displays the resulting level of customer satisfaction. On this coordinate system, the model breaks down the customer needs into three categories: delighters, performance needs and must- have needs which are all illustrated in Figure 4.

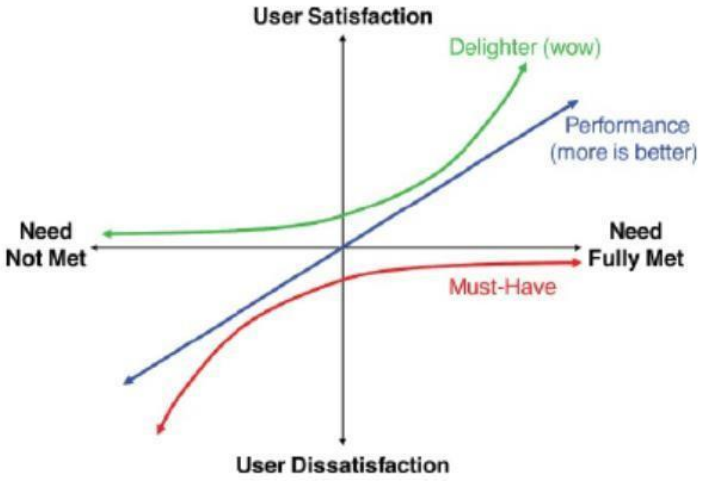

Figure 4: The Kano Model (Olsen, 2015)

Performance needs are needs that increase customer satisfaction the better that they are addressed, while must-have needs only cause dissatisfaction the less that they are addressed (Olsen, 2015). The third category, delighters, refers to benefits that are unexpected. Unlike the performance needs the delighters can only increase satisfaction, but the absence of a delighter can never cause dissatisfaction. Furthermore, the Kano model is not static (Löfgren et al., 2011), and that needs can migrate between the categories over time. For instance, once a delighter has been introduced it will subsequently become a performance need that is somewhat expected. If it continues to be a performance need for sometime it will eventually become completely expected and a requirement in order for the customer to not be dissatisfied, transforming the need into a must-have. Additionally, the Kano model also exhibits traits of a hierarchy, in that the delighters are less likely to provide any satisfaction for the customer if the must-have needs are not addressed first (Olsen, 2015).

3.2 Standardized work process

The success of a business is built upon the competence of its workforce (Misiurek, 2016). Misiurek (2016) claims that the quality of products and services, as well as the efficiency of processes, is greatly influenced by the people involved in every step of the way. While automation has been introduced to reduce the likelihood of for example errors, failures, and downtime, it merely shifts the possibility of human errors to instead occur in different areas such as programming, maintenance, or supervision. Misiurek (2016) further writes that people will continue to be an indispensable part of the manufacturing process, and it is crucial to work on preventing errors from occurring. Misiurek (2016) posits that it is possible to eliminate all human errors from a process, but it requires a fundamental shift in the organizational culture and the management approach. It requires continuous and daily work on the attitudes of the employees, as well as managing their competencies effectively. To achieve this goal, a standardization of work processes is crucial (Misiurek, 2016).

To foster a culture of continuous improvement and quality control, businesses should cultivate attention to detail. Misiurek (2016) suggests providing employees with necessary resources and emphasizing the importance of quality control, safety protocols, and other crucial manufacturing aspects. This reduces error risks and ensures high standards of quality and reliability. Eliminating human errors in manufacturing requires collective efforts towards continuous improvement from all stakeholders (Misiurek, 2016).

3.3 Data visualization

This section aims to describe the prominent literature regarding data visualization, its purpose and goals as well as how to perform it effectively. The theoretical work provided in this section will be used when visualizing the desired data of the case company’s employees.

3.3.1 The Purpose and Goals of graphics

The amount of data available in today's industry is massive and requires high-quality data visualization to enable proper interpretation, with respect for decision-makers without a statistical background (Duke et. al, 2015). Gelman & Unwin (2013) refers to four different aspects making up the purpose of graphic display, written by Tukey (1993 primarily for smaller datasets). The first aspect of them is that they are supposed to be qualitative and descriptive rather than carefully quantitative, meaning that one could find specific values, but a graphic display should convey a larger picture and message. The second aspect is that graphics are made for comparison between amounts, which is similar to the first aspect in that they are not made for accessing specific values, which Gelman & Unwin (2013) is much easier done within a table. The third aspect is that graphics should make an impact on its viewers rather than requiring too much effort and analysis in order to be understandable. Lastly, the fourth aspect further emphasizes the importance of graphics conveying the results of data analysis rather than being descriptive of the process of attaining those results. In other words, the graphics should display results that can be built further upon instead of just offering an alternative process to attaining the same results.

Gelman & Unwin (2013) offers two sets of goals for graphics regarding larger datasets specifically: discovery goals and communication goals. There are three different discovery goals. The first one is to give an overview, meaning that graphics should allow for the viewer to get a qualitative understanding of what is in the datasets, as well as confirm the results and allow for simpler detection of patterns. The second discovery goal is to convey a sense of the scale and complexity of the dataset to the viewer, and the third discovery goal is to allow for further exploration and discovery of unexpected aspects of the data. Gelman & Unwin (2013) also describes the communication goals for graphics, such as displaying the information in a way that is understandable and comprehensible to both the creator of the graphic as well as the audience of it. The graphics should attract the attention and stimulate the interest of its viewers, and it is supposed to tell a story according to Gelman & Unwin (2013). Furthermore, put simply, communication goals refer to displaying a convincing pattern, while discovery goals refer to observing deviations from expectations, and achieving both sets of goals simultaneously can be challenging (Gelman & Unwin, 2013)

3.3.2 Dashboards - One way of displaying data

A dashboard is a type of data visualization that visualizes data regarding key performance indicators (KPIs) and metrics, specifically (Goldmeier & Duggirala, 2015). According to Pauwels et al. (2009), there are several different benefits that dashboards generate for an organization such as sharing of metrics, framework for recognizing excellent performance, source of organizational learning, tool for increased profitability, and decision making. Dashboards are described as beneficial, powerful management tools to decision-makers when faced with the diversity and complexity of market data that is available now in the information age. They can reduce all of the different metrics and this diverse and complex data into a single visual display, making the data comprehensible for a wider range of actors. Furthermore, Pauwels et al. (2009) explain that the purpose of dashboards is to enforce consistency in measures and measurement procedures across the different departments and units of the business. They go on to say that they are useful in monitoring performance, planning resources and also communicating the progression and current status regarding different measurements to the stakeholders. Another benefit of dashboards is that they allow for more integrated goals within organizations by highlighting inconsistencies between different departments. They help ensure that all employees, stakeholders and other actors involved are in alignment of the current status of different measurements as well as the targets and benchmarks that are aspired.

3.3.2.1 Risks with dashboards

The use of dashboards in data analysis also comes with various risks and challenges, such as lack of transparency, incomplete information and misunderstandings as well as being too expensive and resource-intensive (Matheus et al., 2020). Evidently, displaying incomplete or inaccurate information can lead to misunderstandings and faulty conclusions. Also, largely dependent on the design of the dashboard, displaying an overload of information might actually lead to less transparency. There are other factors being the cause of incomplete transparency as well, such as information asymmetry and anonymization of data due to legal requirements. However, Matheus et al. (2020) also explains how correct data can be misinterpreted as well. For example, if data has been retrieved from various sources or there are too many owners of a dashboard, the quality of it may decrease. Also, being able to detect fraud or to identify strange patterns does not serve a purpose if there are not any department or workforce qualified to further investigate these patterns, no matter how well-designed the dashboard might be.

The development of dashboards also risks being too expensive. Developing dashboards, operating them and analyzing their data require a lot of specialized human resources, and data scientists are rare and expensive (Matheus et al., 2020). Matheus et al. (2020) goes on to describe that the high costs of employing data scientists and analysts can result in a lack of new development, as well as a limited, standardized, and preconceived perspective in the long run.

3.3.3 Guidelines for data visualization

According to Tufte (2001), the organization of data visualizations should have a distinct visual hierarchy, with the most crucial information standing out prominently. The design should place emphasis on the most important information and make it noticeable, while deemphasizing or removing less significant information. To establish the visual hierarchy, tools like font size, color, and placement are utilized. The visualization should be organized in a manner that enables the viewer to quickly compare different aspects of the data, such as contrasting the lengths of bars in a bar graph, or evaluating the placement of data points in a scatter plot. Furthermore, Tufte (2001) advocates for the layering of information in data visualizations, enabling viewers to perceive multiple facets of the data at once. To achieve that, Tufte (2001) suggests breaking down a larger data set into smaller, more manageable components, and then presenting each component in its own visualization. This way, the viewer can compare and understand the relationships between the different pieces of information with ease (Tufte, 2001). Additionally, visualization should be structured in a way that enables the viewer to smoothly zoom in and out to explore various levels of detail in the data, and not just provide a broad overview (Tufte, 2001).

Expanding upon the ideas and concepts introduced by Tufte (2001), Kelleher & Wagener (2011) describes ten different guidelines to follow when creating effective data visualizations. The first guideline refers to conveying the intended information while avoiding redundancy and excess ink. This is done by simplifying the graph through minimizing the data-ink ratio, as well as prioritizing simplicity over impressiveness. In regards to the first guideline, Tufte (2001) explains that the key to presenting data effectively and efficiently is to use the least amount of ink possible, while still effectively conveying the information. By emphasizing the data-ink ratio, designers can produce visualizations that are simple, straightforward, and easily understood (Tufte, 2001). The first guideline of Kelleher and Wagener (2011) and its importance is illustrated in Figure 5.

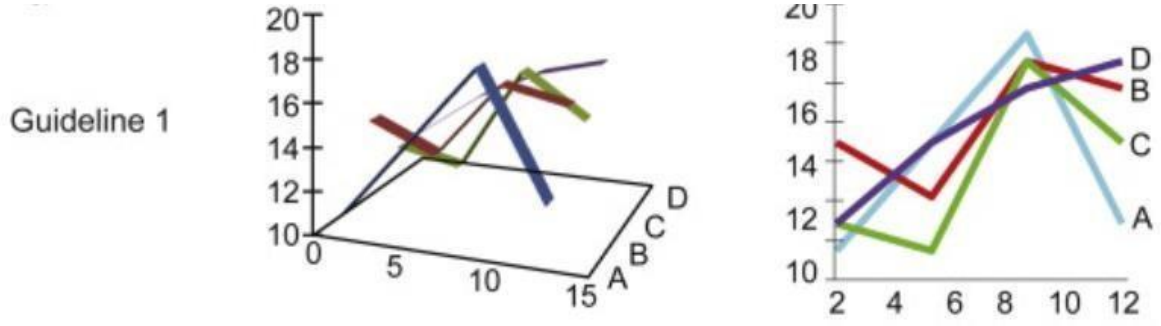

Figure 5: Guideline 1 (Kelleher & Wagener, 2011)

The second guideline provided by Kelleher & Wagener refers to carefully selecting graphical encoding objects and their value-encoding attributes when creating plots to effectively display information from a dataset. For instance, attributes like the length and the position of objects are suggested to convey quantitative information and display actual values, while attributes like width, color tint or the marker area are suggested to display relative comparisons or general patterns

(Kelleher & Wagener, 2011). The third guideline suggests that when creating a plot, one should determine whether the aim is to visualize patterns or details based on the purpose of the plot (Kelleher & Wagener, 2011). For instance, bar or line graphs are better at highlighting individual values while bubble plots or heatmaps are better at efficiently communicating patterns, as visualized in Figure 6.

Guideline 3

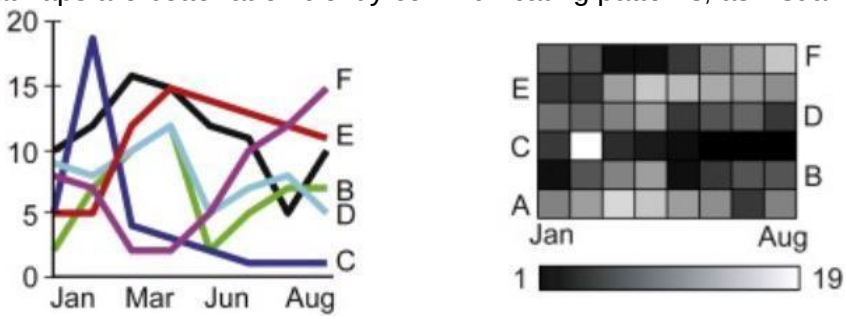

Figure 6: Guideline 3 (Kelleher & Wagener, 2011)

The fourth guideline refers to the importance of choosing the right axes for the graph (Kelleher & Wagener, 2011). It is suggested that the vertical axis should begin at zero when absolute magnitudes are important in order to avoid exaggerations of the relative difference between values, as Figure 7 exemplifies. However, if the relative difference between values are important and wished to be highlighted, Kelleher & Wagener (2011) suggest the limits of the plot to be set as closely as possible to the dataset range.

Guideline 4

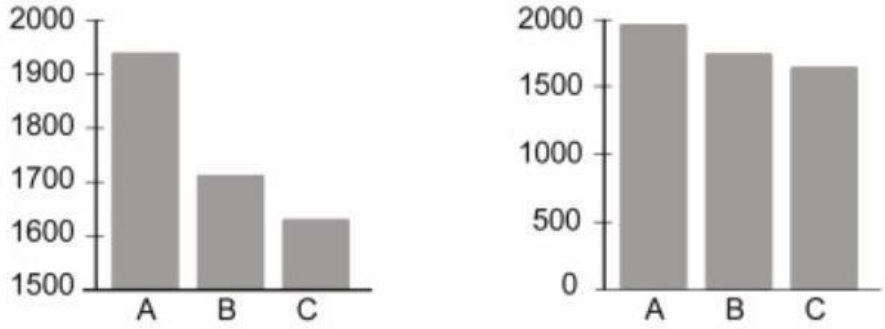

Figure 7: Guideline 4 (Kelleher & Wagener, 2011)

The fifth guideline provided by Kelleher & Wagener (2011) emphasizes the importance of selecting the right graph aspect ratios to emphasize rates of change for time-series data. For instance, as it is illustrated in Figure 8, logarithmic scales on the vertical axis can remove skewness in datasets as well as normalize the rate of change to an initial value. However, it is important to note that the decision to use a transformation should depend on the dataset and the intent of the plot, rather than being a universal recommendation (Kelleher & Wagener (2011).

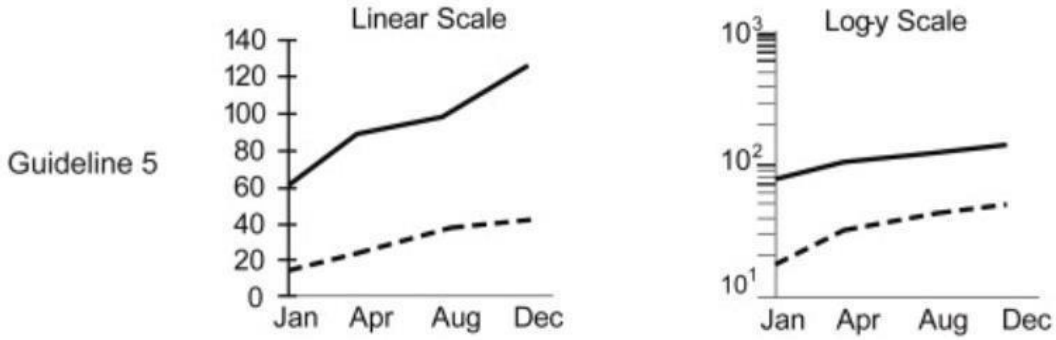

Figure 8: Guideline 5 (Kelleher & Wagener, 2011)

The sixth guidelines refers to Scatter plots specifically, where the points that are plotted out might overlap. Kelleher & Wagener (2011) suggest that when encountering datasets like this, unfilled or transparent points should be used in order to visualize density differences once they overlap. This is illustrated in Figure 9.

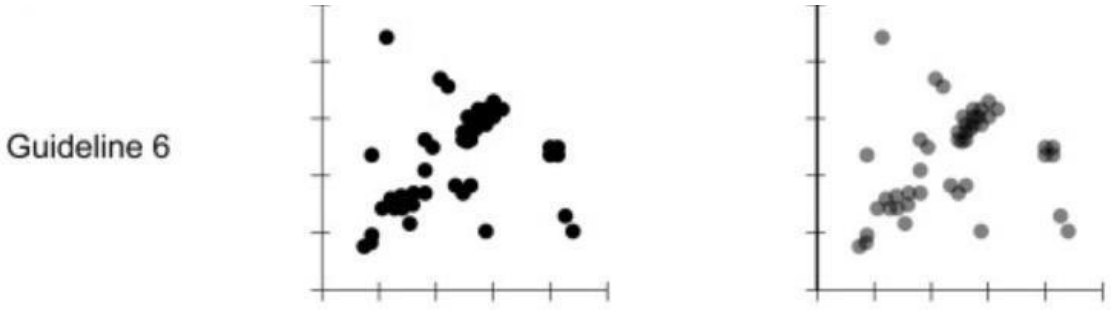

Figure 9: Guideline 6 (Kelleher & Wagener, 2011)

The seventh guideline recommends that when visualizing sequential data in time-series plots, lines should be used where there is a period of missing data. Kelleher & Wagener (2011) explain this by saying that when data is missing between two points in time, the best assumption is that the change between the two points has been linear. An example of a line being used in such an occasion is illustrated in Figure 10, however it is important to not apply this method to points in non-sequential data because of irrelevancy (Kelleher & Wagener, 2011).

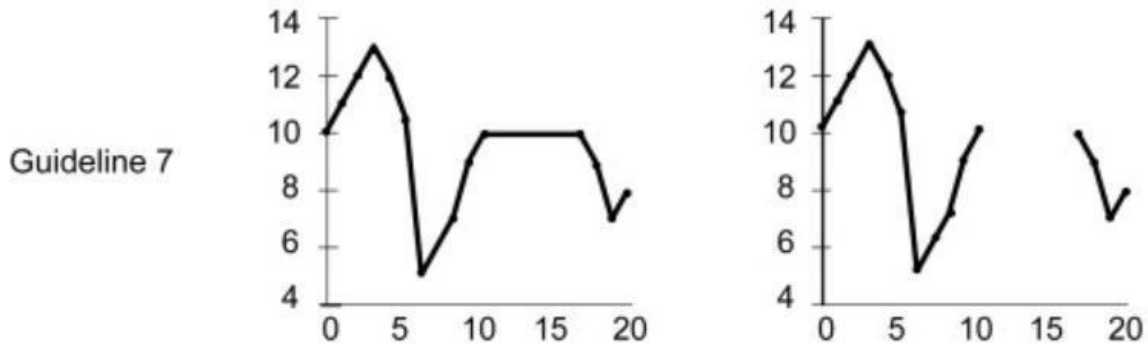

Figure 10: Guideline 7 (Kelleher & Wagener, 2011)

The eight guideline provided by Kelleher & Wagener (2011) offers a way to simplify graphs when containing a lot of information. For instance, for quantitative information box-and-whisker plots could be used to smoothen and summarize the information. Furthermore, when dealing with both numbers and categories, e.g., the number of types of fruit, Kelleher & Wagener (2011) suggest using a Cleveland dot plot or a linked micromap plot in order to better illustrate how the numbers are related to different categories. Lastly regarding the eighth guideline, Kelleher & Wagener (2011) advise avoidance of pie charts because of their risks regarding misrepresentation.

Guideline 8

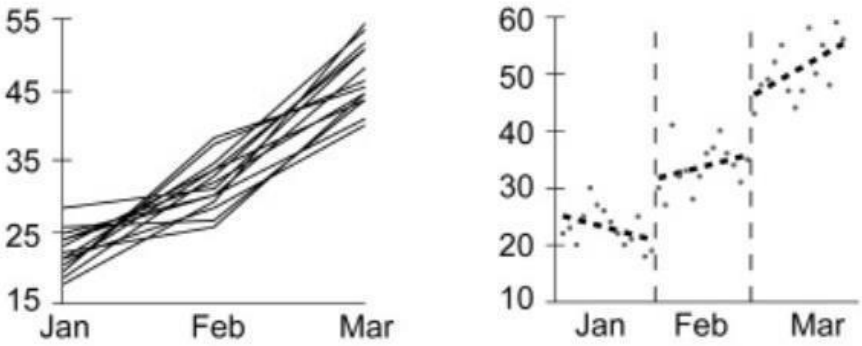

Figure 11: Guideline 8 (Kelleher & Wagener, 2011)

The ninth guideline suggests separating datasets that differ a lot in range in order to better illustrate and highlight the differences within each dataset (Kelleher & Wagener, 2011), as illustrated in Figure 12.

Guideline 9

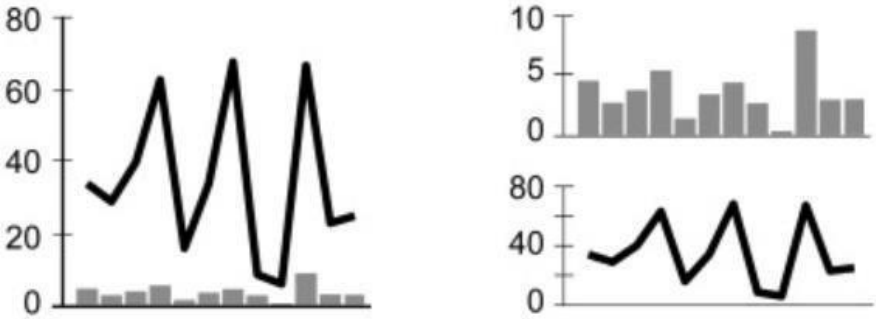

Figure 12: Guideline 9 (Kelleher & Wagener, 2011)

The last guideline, Guideline 10, emphasizes the importance of selecting appropriate color and color scales, with the lighter shades representing the lower values and the darker shades representing the higher value (Kelleher & Wagener, 2011). Guideline 10 is illustrated in Figure 13.

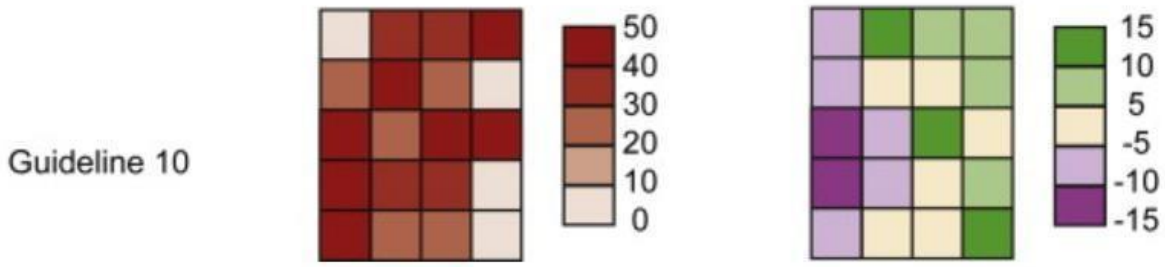

Figure 13: Guideline 10 (Kelleher & Wagener, 2011)

3.3.4 Choosing appropriate Graphics

According to Tufte (2001) there are various techniques for encoding data in visualizations, such as bar charts, line charts, and scatter plots. Different encoding methods should be applied to different types of data, and the designer should carefully select the most appropriate technique for the information being displayed. Doumont & Vandenbroeck (2015) present three different criteria to consider when choosing the right graph for data representation: the structure of the dataset, the intended use of the graph and the research question. The first criteria, the structure of the dataset, is referred to as the most important one, and implies the quantity and the quality of the data meaning the number of variables as well as the type of variables. Furthermore, variables can either be continuous or discrete, implying the dataset can either represent a range of values, often along a scale, or they can be a set of specific values. Doumont & Vandenbroeck (2015) exemplify this by explaining the temperatures 30 degrees and 80 degrees are considered continuous variables if the data implies range of temperatures between the two, while the same two temperatures are considered discrete variables if they are to be looked at as two individual labels.

The second criteria to consider when selecting the optimal graph is the intended use (Doumont & Vandenbroeck, 2015). The optimal graph is explained to not necessarily have to be perfect, but instead to be one best suited for its audience. For instance, a graph made for personal use should allow for analysis or answering questions about oneself, while a graph made for peers should allow for a discussion and a graph made for publication should simply just convey the message. Furthermore, Doumont & Vandenbroeck (2015) emphasizes the importance of noting that the graphs that are well-suited for analysis may not necessarily be good for communication, and vice versa.

The third criteria to consider when selecting the optimal graph is the research question or the intended message at the communication end (Duomont & Vandenbroeck, 2015). There are four generic categories of research questions: comparing among individual data, distribution of data along a scale, correlation between variables and evolution over time of a variable. Additionally, Duomont & Vandenbroeck (2015)- also mentions a fifth meta category, comparing groups of data, meaning the data grouped together by the four categories above can also be compared to each other.

Furthermore, Duomont & Vandenbroeck (2015) emphasize that no graph type is perfect or absolute, meaning each will make answering some questions easier while answering other questions more difficult. Additionally, Duomont & Vandenbroeck (2015) goes on to explain several graph types and what type of data they are best suited for. For instance, bar charts and dot charts are explained to be the most common graphic types to use for representing quantities along a numerical scale (Duomont & Vandenbroeck, 2015). In order for a bar chart to allow for actual meaningful comparisons, Duomont & Vandenbroeck (2015) advise that they should be drawn along a linear ratio scale starting from zero, as opposed to along a logarithmic scale for instance. Dot charts can instead be used for any scale

and are therefore better at resolving closely grouped data as well as accommodating additional information, such as subsets or whiskers for instance. Duomont & Vandenbroeck (2015) also describe line plots. They are explained to be a graph type best used to show the evolution of one variable over time, using connected dots. Furthermore, multiline plots compare the evolution of several variables expressed in the same units, while multipanel plots show the evolution of several variables along different scales.

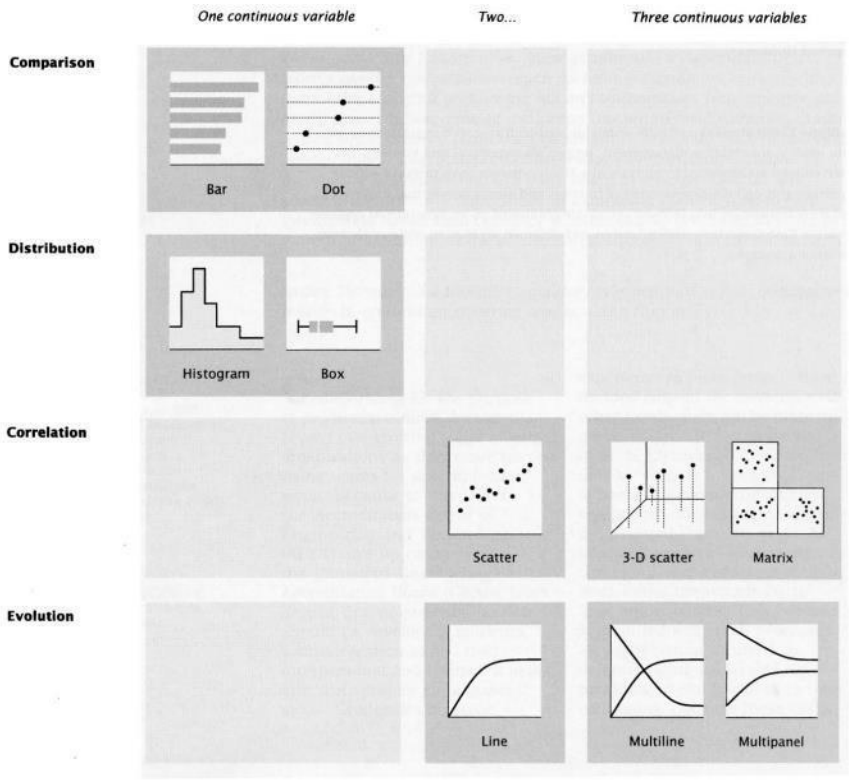

Figure 14: A table of what graph type to select (Duomont & Vandenbroeck, 2015)

3.5 User acceptance of technology

This section aims to delve into the theories regarding user acceptance and engagement when introduced to new technology, to subsequently ensure the actual usage of it. Dillon & Morris (1996) explain three different frameworks that can be used to determine the user acceptance of new technology: the theory of reasoned action, the theory of planned behaviour and the technology acceptance model. The frameworks described will not only highlight what factors are in favour of EazyBI usage at the case company, but also allow for insights in how the documentation could be organized to maximize user engagement.

3.5.1 Theory of reasoned action

The first model that can be used to determine user acceptance is the *Theory of reasoned action*, also known as TRA (Dillon & Morris, 1996). TRA suggests that an individual's behaviour is determined by their intention to perform that behaviour. Bharathy (2021) describes three different boundary circumstances that can affect the strength of the link between intentions and behaviour. The first

boundary circumstance is referred to as the degree of specificity between the measure of intention and the behavioural criteria. For instance, the intention “I want to be healthier” is not as specific as the intention “I want to go to the gym one hour a day” in regard to the behaviour of exercising. The second boundary circumstance is the consistency of intentions between measurements and behaviour performance. This boundary simply claims that the more consistent the intentions of someone matches their behaviour, the stronger is the link between the two. Lastly, the third boundary circumstance is the degree of volitional control over carrying out the intention. This refers to the level of control or lack thereof an individual has over their behaviour. For instance, if someone has trouble quitting smoking even though they intend to, it suggests that the intention of that individual is not strongly linked to their behaviour.

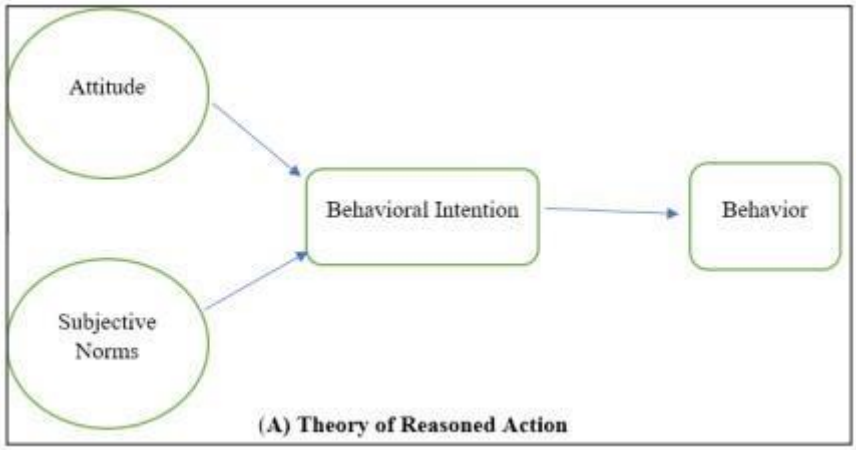

Figure 15: Theory of Reasoned Action (Dillon & Morris, 1996)

Furthermore, Dillon & Morris explains that the intention to perform a certain behaviour is influenced by two factors: the individual’s attitude towards the behaviour as well as the subjective norm. Firstly, Dillon & Morris (1996) explains that an individual’s attitude towards a behaviour is a result of their beliefs regarding what consequences a decision or act may result in, as well as an evaluation of those believed consequences. Secondly, Dillon & Morris (1996) defines the subjective norm as the individual’s perception of how people who are important to them think they should behave and explains that to be the second major factor influencing the intention to perform a certain behaviour. Out of the three frameworks discussed in this section of the report, TRA is arguably the one that has been applied successfully in the widest range of industries and situations beyond its original scope, such as voting in elections and consumption of alcoholic beverages (Dillon & Morris, 1996).

3.5.2 Theory of planned behavior

A second model that is proposed by Dillon & Morris (1996) as a theoretical perspective to determine user acceptance is the *Theory of planned behavior*, also known as TPB. TPB is a descendent from TRA, but instead adds a third factor besides attitudes and subjective norms to take into consideration when predicting an individual’s user acceptance. This third factor is the user’s perceived behavioral control (Dillon & Morris, 1996). This is explained to be determined by the availability of skills, resources and opportunities for the individual, in regards to a certain behavior, determining their intention towards that very behavior which in turn influences it.

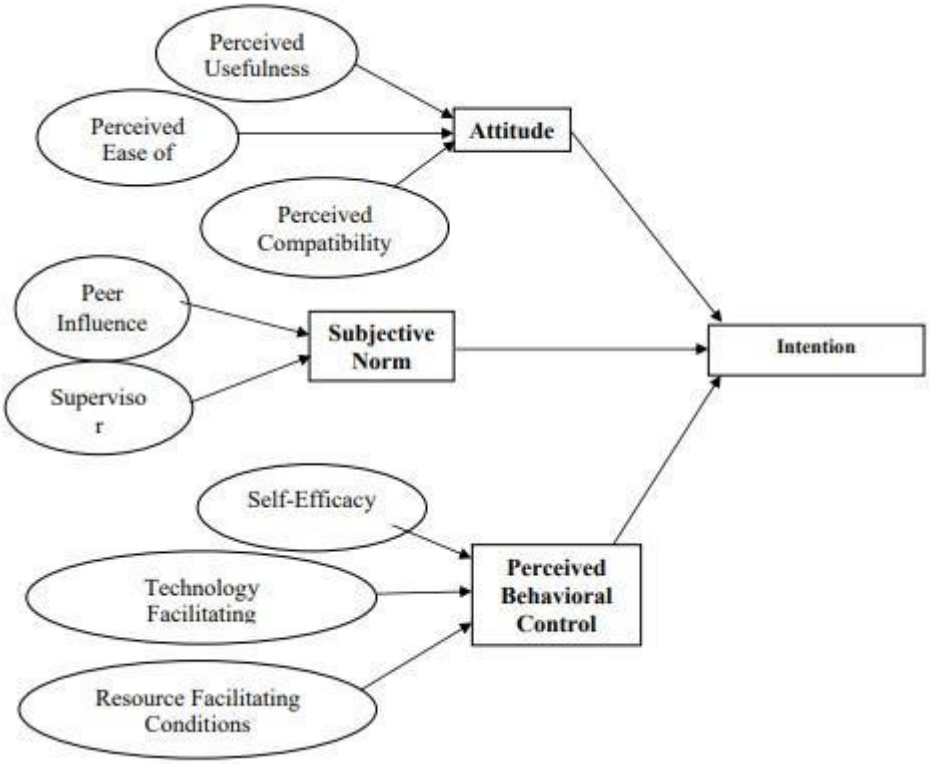

Figure 16: Taylor and Todds Decomposed theory of Planned behavior (Keong & Husin, 2019)

Dillon & Morris (1996) also describes a specific version of TPB called DTPB, a *Decomposed theory of planned behavior*, which aims to identify and model the specific antecedents to attitude, subjective norm and behavioral control relevant to IT usage specifically. Zaman et al. (2021) explain that by breaking down these factors into their respective antecedents, a better understanding of an individual's intention to adopt technology can be achieved. In DTPB, Taylor & Todd (1995) refers to determinants of attitude as self-efficiency, resource facilitating conditions and technology facilitating conditions. Additionally, Taylor & Todd (1995) suggest that the subjective norm is determined by peer pressure and the influence of superiors, while the attitude is suggested to be determined by perceived usefulness, perceived ease of use as well as perceived compatibility, which are determinants largely consistent with those of another framework discussed later in the report, TAM. Furthermore, this approach of DTPB has been applied successfully when investigating the adoption of many different technologies such as big data analytics and artificial intelligence (Zaman et al., 2021). However, according to Bharathy (2021) the TPB serves more as the foundational framework within green marketing, as it has been used predominantly to investigate pro-environmental behavior such as recycling energy use and mode of transportation.

3.5.3 Technology acceptance model

A third model that can be used to determine the user acceptance of new technology is the *Technology acceptance model* (Davis,1987). The technology acceptance model, also known as TAM, is another theoretical model derived from TRA that aims to explain why the intended user of introduced technology might accept or reject it simply based on the design features of the system (Dillon & Morris, 1996). It can predict acceptance of an information system as well as diagnose design problems even before users actually experience that system. Furthermore, Davis (1987) goes on to describe the benefits of using TAM when designing a dashboard, emphasizing its usefulness in both the development and the evaluation stages before implementing it.

As it can be seen in the illustration of the TAM below, users' motivation of using new technology is primarily affected by two factors regarding the system's design: its perceived ease of use as well as its perceived usefulness. Dillon & Morris (1996) describes the perceived usefulness to be to what degree a user believes a system will enhance their performance, while the perceived ease of use refers to the degree to which a user believes they can use a system effortlessly. Both these factors are explained to have a significant impact on the user's attitude and behavioral intentions towards using a system.

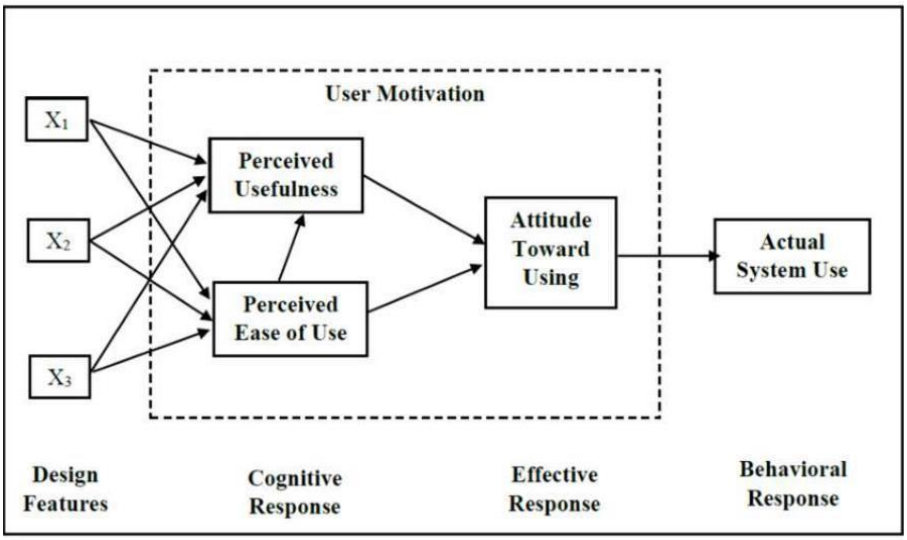

Figure 17: Technology Acceptance Model (Davis, 1987)

As illustrated in Figure 17, the perceived ease of use of a system has an effect on the perceived usefulness. Additionally, the perceived usefulness of a system in and of itself, as it has been referred to earlier in this report, was statistically found to be 50% more influential when determining usage than simply the perceived ease of usage (Davis 1987), suggesting that to be a crucial aspect when designing a system. However, the exact interplay between all characteristics of actual system usage in practice was overall concluded to not yet be defined well enough and therefore need further research.

Dillon & Morris (1996) identifies several key differences between TAM and TRA. One major difference identified between the two is that the empirical work validating TAM does not find social factors and subjective norms to be any major predictors of user intentions. The technology studied within the empirical work behind TAM was very personal, and its usage was found to be unrelated to others' use of that same technology, hence the conclusion was drawn that an individual's usage of a system is likely unrelated to social influences. Another major difference between TAM and TRA, is that TAM suggests that a person's general beliefs are not important to whether or not they will use a certain technology (Dillon & Morris, 1996). In other words, an employee's dislike towards new technology is claimed to be insignificant if they believe it will make a positive impact on their job performance. Furthermore, while the aforementioned TPB model provides a better understanding of specific determinants of intention, TAM is believed to be the more suitable model when predicting IT usage specifically.

4. Result

This chapter aims to present the results of the interviews with the potential end users. Qualitative data affecting the creation of the documentation has been divided into a process level and a content level. The qualitative data on a process level perspective will offer insights addressed in the discussion regarding how to structure and present the documentation, while the qualitative data on a content level perspective will display what data the documentation needs to address and how.

4.1 Process level

This section will present how the results that were deemed could potentially offer insights into what factors to consider for maximizing user engagement of the documentation.

4.1.1 Frameworks for predicting EazyBI usage

By using the TRA framework, as seen in Table 2, it shows that the interviewees look positively towards EazyBI and therefore also the documentation on it.

Table 2: Theory of reasoned action

	Theory of reasoned action	
Interviewee	**Subjective norm**	**Attitude**
Y	Do not know if they're affected	Has a positive attitude
Z	Says it does not affect them	Has a positive attitude
A	Do not know if they're affected	Has a positive attitude
B	Says it does not affect them	Has a positive attitude

Interviewee B expressed enthusiasm towards data visualization, highlighting that anything that improves the ability to visualize data is positive. While this particular individual had not explored eazyBI extensively, he held a positive inclination towards the concept of data visualization in general. "*I'm sure it brings with it positive consequences. Everything that improves the possibility to visualize data is positive to me. In general I'm very positive towards data visualization. I have not looked deeply enough at Easy BI to have a firm opinion yet." - Interviewee B*

Another interviewee, interviewee A, voiced their appreciation for guidelines and general queries that could be used directly. He believed that having such resources would have a positive impact on their experience with EazyBI. The sentiment here implies that easily applicable tools and guidance would be valuable for their data analysis needs. "*If we had more general guidelines, it would probably be positive. I am positive about the guide, but what I am hesitant about is... I would appreciate more general queries that we can use directly. It's great to have general tools that we can use directly." - Interviewee A*

For interviewee Z, a strong emphasis was placed on the importance of excelling in data visualization. The interviewee acknowledged the abundance of data available within the organization, as well as

emphasized the need to extract meaningful information from it. He believed that EazyBI had the potential to be helpful in improving their understanding of data through visualization. *"In general, I believe it's an incredibly important area that we need to excel at. At [the case company], we have an abundance of data, but data is worthless if we can't extract information from it. To extract information, we need to be able to visualize it in order to understand it. However, we're not very good at this at Ericsson. We clearly need to improve… It sounds like it could be really helpful for me." - Interviewee Z*

In the rapidly evolving software industry, interviewees stressed the importance of data visualization for tracking progress. Visualizing software progress was deemed crucial, as it allowed teams to effectively monitor and comprehend their advancements. *"In software it changes so fast and in that context data visualization is very important because we need to visualize our progress." - Interviewee Y*

Furthermore, overall none of the interviewees seemed particularly affected by their colleagues' perception of EazyBI nor data visualization in general. Judging by this sentiment only, this implies that creating documentation on EazyBI to further enable its usage will not require the subjective norm to be addressed.

One interviewee, interviewee B stated that he had no knowledge of how their colleagues perceived EazyBI, suggesting that the interviewee's perception and utilization of the tool would not be influenced by their colleagues' views. *"I have no idea how my colleagues perceive Easy BI" - Interviewee B*

Similarly, another interviewee, interviewee Y, highlighted that their perception of eazyBI and data visualization would remain unchanged, regardless of their colleagues' feelings. The alignment in reviewing key performance indicators (KPIs) with colleagues was seen as a necessity for the job rather than specifically being influenced by colleagues' opinions. *"In terms of how my colleagues review KPIs we work very aligned because it cannot be made in so many other ways. I don't know how my colleagues feel about it but I don't think it would change my perception." - Interviewee Y*

Interviewee A emphasized that the interest in data and the need for data information varied among different roles within the organization. He specifically mentioned that while leaders and managers found the data valuable, employees such as coders had no interest in it. This further supports the notion that the perception of eazyBI and data visualization is not significantly influenced by the subjective norm within the organization. *"First, it should be understood that only a small number of people are interested in this data. For example, our coders have no interest. We leaders and managers are interested in this data." - Interviewee A*

Interviewee Z acknowledged being more analytically inclined compared to their colleagues, however he still believed that their colleagues had similar needs for data information. Interviewee Z implies that he consider the perceived value and utilization of eazyBI to be driven more by individual needs rather than the influence of any subjective norm. *"I have three colleagues who have the same role as me, so they are my three closest colleagues. I think they have about the same needs as me and roughly the same view on data information. Though it is nothing I have reflected about. Then I think that by nature, I am a very analytical person. Perhaps a little more than the others, so I may find it a bit more interesting than they do. But they definitely have the same needs as I do." - Interviewee Z*

Overall, the quotes suggest that the interviewees' perception and usage of eazyBI are primarily driven by personal needs and interests, rather than being influenced by their colleagues' opinions or any subjective norm. Therefore, when creating documentation to enable the usage of eazyBI, it may not be necessary to address the subjective norm extensively. Instead, focusing on providing valuable resources and meeting the specific needs of individual users would likely be more impactful.

When asking the interviewees about their perceived behavioral control when learning EazyBI, it can be determined that this was lacking, as illustrated in Table 3. According to the interviewees, this was primarily because of a lack of time to learn this tool, i.e a lack of opportunity, and also because of a lack of adequate documentation on the tool that provides information about how to use it to perform the data visualizations of interest, i.e a lack of resources. Using the TPB, it can be assessed that the documentation on EazyBI has to address these two factors: Providing documentation that is easy to follow quickly as well as documentation that provides the information that the end-users are interested in.

Interviewee B, for instance, mentioned missing the introduction to EazyBI and expressed a desire for a "get started" type of presentation. He acknowledged their lacking EazyBI proficiency and the need to catch up before being able to use EazyBI efficiently. The interviewee attributed their lack of behavioral control to both a lack of time and missing the initial introduction, which may in and of itself have been caused by time constraints as well. *"Since I totally missed the introduction I must say I miss a get started type of presentation… I'm at a medium level perhaps… I definitely have a catch up to do before I will be able to use Eazy BI in an efficient way… I have a lack of time, for sure, but the main issue for me is probably that I missed the introduction completely. Maybe, that was because of a lack of time? " - Interviewee B*

Intervewee Y also emphasize the lack of relevant documentation for EazyBI as the foremost obstacle that is hindering them from learning the tool. *"I would definitely say that there is a lack of relevant documentation for EazyBI" - Interviewee Y*

Interviewee A speculated their lack of involvement in the learning process of EazyBI to be because of several factors, however primarily attributing it to their limited time. Interviewee A also acknowledged that different individuals have unique ways of visualizing things, suggesting the importance of tailoring the documentation to cater to various visualization preferences. *"Can I have the opportunity to receive an introduction to this tool? It would make things much easier for me. Personally, I don't have the time to learn by experimenting on my own… I think all of these things have an impact [Skills, resources and opportunity]. I am wondering why I haven't been involved yet, but it is probably mostly due to my limited time. [Interviewee X] is now doing a lot of other things and cannot support us in the same way as before. Everyone has their own way of visualizing things." - Interviewee A*

Furthermore, Interviewee Z emphasized the lack of time as the primary reason for hindering the process of learning new IT systems, such as EazyBI. The lack of time, i.e the lack of opportunity, is therefore what negatively affects interviewee Z's perceived behavioral control the most. "*That's an easy question. It's time. Tools and resources can always be solved. Time is the difficult part." - Interviewee Z*

Table 3: Perceived behavioral control (TPB)

	Perceived behavioral control (TPB)		
Interviewee	**Skills**	**Resources**	**Opportunity**
Y	Not addressed	Lacking	Not addressed
Z	Not addressed	Not addressed	Lacking
A	Lacking	Lacking	Lacking
B	Not addressed	Lacking	Lacking

Using the Technology acceptance model, it can be determined that EazyBI will likely be used if the documentation on how to use it is provided. However, as it can be assessed from Table 4, the perceived ease of use of the interviewees regarding EazyBI is slightly insufficient. This means that the documentation on EazyBI also has to address this by simplifying its usage and being easy to follow. Furthermore, interviewee Y was slightly skeptical regarding the usefulness of EazyBI. To address this, the documentation has to provide knowledge that is useful to that interviewee, i.e. how to visualize data of his interest.

One interviewee, interviewee Y, recognized the potential user-friendliness of EazyBI due to the organization's vast amount of data. He believed that learning the tools within EazyBI would not be overly difficult, but their main challenge lies in locating the desired information within the data. However, he expressed a need for the tool framework to support their specific data visualization requirements. This highlights the importance of the documentation addressing the customization and adaptability of EazyBI to cater to different user needs. *"EazyBI is probably very user friendly. We have an enormous amount of data so the main problem is where do I need to go to find what I want. I think it wouldn't be too difficult to learn the tools in eazyBI… I am not overly familiar with EazyBI. I don't need another framework. I may need some visualization of the data that we already have but the problem is that the tool framework does not support that… Data Visualization is absolutely useful, picking the right things to look at and how you manage the data visualization is more critical." - Interviewee Y*

Another interviewee, interviewee A admitted to having limited familiarity with EazyBI but acknowledged its usefulness, particularly in terms of visualizing existing data. He noted that although the creation of queries within EazyBI required some effort, he considered the end results to be excellent. This underscores the need for the documentation to provide clear guidance on creating queries effectively. *"I have looked into it a little bit and it is indeed very useful, but it does require some work to create these queries… I have never done anything myself, just used the results, so I don't know how easy it is to learn, but the result is excellent." - Interviewee A*

Interviewee B finds it difficult to determine whether EazyBI seems easy to use nor if it seems useful, however he still perceive data visualization in general to be positive. "*Don't know yet [if it is easy to learn]... In general I'm very positive towards data visualization. I have not looked deeply enough at Easy BI to have a firm opinion yet."- Interviewee B*

Interviewee Z had a positive impression of EazyBI based on a brief 10-minute demonstration, where he observed the ease of building queries by dragging and dropping building blocks. While acknowledging his limited experience, he expressed interest in exploring the tool further, and emphasized both perceived usefulness as well as ease of use of the tool. He additionally mention EazyBI being designed for Jira, highlighting the potential advantages of the tool already being integrated with existing systems. *"I have also seen a roughly 10-minute demo of how to build queries with different building blocks. I know that it seems to be relatively easy in that sense, you don't need to write any SQL code, you can just drag and drop building blocks and put them together. It looked interesting and nice to work with. But it was only 10 minutes, I haven't had a chance to test it myself yet. I am a rookie at it… Yes, the advantage of that is that a lot of the data I work with is in Jira. And EazyBI is built for Jira." - Interviewee Z*

Table 4: Technology acceptance model

	Technology acceptance model	

Interviewee	Perceived Usefulness	Perceived Ease of Use
Y	Don't know if EazyBI seems useful	Think EazybI seems easy
Z	Think EazyBI seems useful	Think EazybI seems easy
A	Think EazyBI seems useful	Think EazyBi seems difficult
B	Think EazyBI seems useful	Don't know if EazyBI seems easy

4.1.2 Customer needs of the documentation

The interviewees were also asked questions regarding customer need specifically on a process level of the documentation, additionally to the process-level insights derived from the usage prediction frameworks. Just like it can be assessed from Table 3, there is a need for simple, efficient documentation that can be read, understood and followed quickly in order to deal with the lack of time as an obstacle for the interviewees. Also, similarly to the lack of resources identified in the Table 3, there is naturally a process level customer need for the documentation on EazyBI to be informative, as well as for it to contain an introduction of the tool to begin with. Furthermore, from the interviews the customer need for a standardized work process of visualizing data can also be identified. All process level insights in regard to customer needs are illustrated in Table 5.

One interviewee, interviewee A, expressed the need for a guide or training sessions, as he found it challenging to fully comprehend the graphs presented by his colleague. He emphasized the importance of short, clear videos targeted towards beginners. "*I think we need a guide because when [interviewee X] presents, the graphs are very nice, but almost only she can fully understand them. Those of us who need this should have some training sessions. I want short, clear videos for "dummies". - Interviewee A*

Another interviewee, interviewee Y, highlighted the lack of consistency in data visualization practices across different departments within the organization. He suggested that different teams used different tools, resulting in a lack of standardization. This observation points to a customer need for establishing a standardized work process of visualizing data at the case company. "*I would say every part of the organization has their own part of data visualization. If you go to another department they work differently but they are all doing the same thing. Different tools and at that point of view we have no consistency." - Interviewee Y*

Similarly to interviewee A, interviewee Z acknowledged the need for training in EazyBI but expressed uncertainty about the available help resources. For interviewee Z, it is important that documentation on EazyBI is informative and easy to follow and understand, but he also especially emphasize the need for the documentation to lead to the data visualization process being the same across all different departments at the case company. *"I will definitely need to have a training in EazyBI. I don't know today what kind of help is available. What built-in help is there, where can I go online, etc. I don't know what the options are. But we definitely need something. I assume that you are doing this because there is a need for it." - Interviewee Z*

Interviewee B also emphasizes several of the customer needs described above to be important for the documentation to consider. For interviewee B, the two biggest customer needs that documentation on EazyBI has to meet are being informative as well as being simple and efficient to

use. *"For me it's important that there are resources such as your documentation I guess that are informative and simple to use" - Interviewee B*

Based on the interviews, it is evident that the interviewees perceive a need for documentation that facilitates quick comprehension and usage due to time constraints. There is also a demand for informative documentation that introduces the tool and establishes a standardized work process for data visualization. Additionally, there is also a need for the documentation to standardize the data visualization process across all different departments within the case company.

Table 5: Customer need of a documentation of a visualization tool

	Customer need of a documentation on a data visualization tool		
Interviewee	**Informative**	**Simplicity, Efficiency**	**Standardized work process**
Y	Strongly agree	Agree	Strongly agree
Z	Agree	Agree	Strongly agree
A	Strongly agree	Strongly agree	Strongly agree
B	Strongly agree	Strongly agree	Indifferent

4.2 Content level

This section presents the results of the data that the employees of the case company regularly need to visualize in order to lay the groundwork for the content to be included when creating documentation. Four main areas of data were identified as being of most interest, as shown in Table 6.

Table 6: Data desired (Content level of the documentation)

	Data desired (Content level of the documentation)			
Interviewee	**Trouble reports**	**Number of Features with F4 date**	**F4 precision**	**Capability**
Y	High interest			High interest
Z	High interest	High interest	High interest	High interest
A	High interest			
B	High interest	High interest	High interest	

The main type of data that all interviewees emphasized to be of interest in their daily work was data regarding TR, trouble reports - i.e. quality related data. Furthermore, data regarding the number of features (issues) with F4 date, F4 precision and capability was also described to be of interest.

Features with F4 date per month refers to how many features were completed in a specific month, F4 precision refers to how well the end date for the issue predicted matches its actual end date, and capability refers to the number of features that are planned and ongoing at the same time.

5. Discussion

This chapter will discuss the creation of the documentation from the perspective of a content level and a process level, by analyzing the results in regards to the theory. Usage prediction frameworks as well as theory of customer needs and the kano model will be discussed in regards to how they can affect the process level of the documentation and maximize the user engagement of it. Furthermore, theory on data visualization will be discussed in order to subsequently determine what factors to consider when visualizing the data desired.

5.1 Process level

This section will discuss the process level factors to consider when creating the documentation on the data visualization tool, in regard to maximizing user engagement.

5.1.1 Applying usage prediction frameworks to EazyBI

In Section 3.5.1 the theoretical framework Theory of reasoned action is described, that can be used to determine user acceptance (Dillon & Morris, 1996). TRA suggests that an individual's behaviour is determined by their intention to perform that behaviour, and that their intention is determined by two factors: the subjective norm and the individual's attitude towards that behaviour. Using the TRA framework, it can be assessed from the result that all interviewees perceive the consequences of integrating EazyBI in their daily work to be positive, hence all interviewees therefore share a positive attitude towards. Many of them emphasize how dashboards and data visualizations help them make better and more strategic decisions in their daily work. This assessment from the results align particularly well with the literature provided by Pauwels et al. (2009) in Section 3.3.2 of the report, and is therefore not considered very surprising.

Regarding the subjective norm however, it can at first glance appear to be non-existent regarding EazyBI and data visualization in general. All interviewees were asked specifically if their own perception aligned with their colleagues, as well as whether they believed it had been influenced by them or not. While their perceptions of EazyBI and data visualization in general appeared to align, none of them claimed their own perception to have been influenced by others. However, even if the colleagues deny being influenced by each other, their shared agreement on a particular topic may still create a perception of social pressure to conform to that agreement. As a result, this shared agreement can still influence an individual's behavioural intentions and subsequent behaviour. In essence, subjective norms reflect an individual's perception of the social pressure to conform to the expectations of others, regardless of whether or not they are aware of it, which can be assessed to be the case in this situation.

As described in Section 3.5.2 the Theory of Planned Behaviour, TPB, adds a third factor, perceived behavioural control, to the attitudes and subjective norms considered by TRA.
Drawing from the perspectives of the interviewees, TPB provides insights into their perceived behavioural control in relation to adopting the new data visualization tool EazyBI. In essence, the three factors affecting the perceived behavioural control towards EazyBI were lacking for all interviewees. Many of them refer to lack of time as the primary obstacle when learning and integrating a new IT-system in their daily work. Additionally, the lack of relevant documentation was mentioned as another obstacle preventing them from learning it, meaning the documentation that is available is considered too general and not applicable to the work of the case company specifically. Using the

TPB model, the lack of time can be categorized as a lack of opportunity, and the lack of relevant documentation can be categorized as a lack of resources. In conclusion, the challenges highlighted by the interviewees underscore the importance of addressing perceived behavioural control when creating the documentation, i.e., making the documentation easily accessible and simple and efficient to read and follow, to potentially increase the likeness of user engagement.

The technology acceptance model presented in Section 3.5.3, also known as TAM, is another theoretical model derived from TRA that aims to explain why the intended user of introduced technology might accept or reject it simply based on the design features of the system (Dillon & Morris, 1996). The design features are explained to affect the intended user's perceived usefulness and perceived ease of use of a certain technology, which in turn is explained to indicate the likeness of that technology's usage. As it can be seen in the result, all the interviewees except interviewee Y perceive EazyBI as useful, suggesting a higher likeliness of adoption. Interviewee Y, however, seems to be more indifferent to the perception of EazyBI being useful. This could potentially be because of a lack of knowledge. To encourage the perception of EazyBI being useful for interviewee Y specifically, the documentation should consider the visualization process of data relevant to him, in order to increase the likeness of him perceiving the software as useful. Regarding the ease of use however, the results differed slightly more between the interviewees. Since many have very limited knowledge of EazyBI and data visualization tools in general, the difficulties of learning a new tool may appear off-putting. This insight suggests that the documentation on EazyBI should aim to simplify the usage as much as possible. For instance, several interviewees described a video tutorial in the beginning of the documentation as a potential way to simplify the documentation.

5.1.2 Customer needs of the documentation - Process level

In essence, there is a lack of information for the case company's employees regarding EazyBI overall, implying a customer need for information regarding its existence and usage for the case company's employees to view it as a viable option to begin with. None of the interviewees claim to be particularly familiar with Eazy BI, except for interviewee X who initiated the thesis of course. Out of the other four, interviewee A and Z have had some interaction with the software. This has however been very brief and experimental, and they do not know enough to use it in their daily work. By introducing EazyBI and explaining its purpose and usage in the interviews as well as in the actual guide being produced, the highest level of Maslow's hierarchy of human needs will be addressed in the sense that the interviewees realize the potential of such a tool being learnt.

As mentioned above, it can also be assessed from the interviewees' answers and statements that there is a lack of time for learning a new visualization tool, suggesting there is a customer need for an efficient way to do so. For instance, both interviewees B and Z state clearly that the lack of time is the foremost obstacle when learning a new system. Interviewee Z goes on to explain that since the case company has an abundance of data, it is very important for the company to be able to extract any information from that data in order for it to be of any value. However, he assesses the case company to be rather poor at this and needs to improve. Regarding learning EazyBI specifically, interviewee Z explains that he is not familiar with what help is available today but claims a need for at least some type of instructional documentation to exist. Furthermore, Interviewee A further supports the customer need for efficiency, speculating when asked that the lack of time is a likely reason for not having learned the tool yet, even though they admit they themselves not having thought too much about it prior to the interview. Interviewee A also emphasizes the need for the guide specifically, explaining that they do not have the time to experiment on their own and would therefore like an introduction of the tool in order to learn it much quicker and easier.

The customer need for a more efficient way to learn how to visualize data that's of interest using EazyBI, can be derived from the two highest levels of Olsen's hierarchy of web users' needs, described in Section 3.1.2 of the report. In essence, the guide and documentation of EazyBI needs to be easy to follow, use and navigate around with as well as actually providing documentation that's relevant to the end-users. Furthermore, meeting the needs of this level presupposes that the needs of the level below, i.e whether or not the functionally and feature set meets the needs of the user, are already addressed and fulfilled. To put it succinctly, the documentation firstly needs to provide the information needed by its end users, and secondly it has to be structured efficiently and be easy to understand.

Additionally, all interviewees explain that they currently do not do any data visualizations themselves, and instead only use the data visualizations that they have requested from others. The assumption can therefore be made that if they were to learn more about data visualization, their confidence and selfesteem in the area would increase, implying the need for learning data visualization efficiently to also be derived from the second to highest level of Maslow's hierarchy of human needs.

Interviewee Z states that they're overall positive about the guide. They go on to explain that all the different departments, including theirs, have created their own templates for visualizing data. Interviewee Y further supports this statement, claiming all departments to work with data visualization differently and there to be an overall lack of consistency in the area. Because of this, interviewee Z believes more general guidelines on data visualization would have a positive effect on the overall work at the case company, since it would allow for a better understanding between these different departments. In Section 3.2, Misiurek (2016) posits that the key to eliminating all human errors from a process is to implement a standardized work process, which this documentation would lead to for EazyBI usage. However, achieving this requires a significant transformation in organizational culture and management approach, which involves consistent efforts to improve people's attitudes and effectively manage their competencies on a daily basis.

5.1.3 The Kano model applied to the documentation

In Section 3.1.6, Olsen (2015) describes The Kano Model, which can be used to prioritize customer needs and preferences. In Figure 18 The Kano Model is applied on the created documentation. It is divided in three categories: Must-haves, Performances and Delighters. Since documentation on EazyBI does not currently exist internally for the case company, it can be assessed that there are not many must-haves attached to it. One must-have that appears likely however, is the one of documenting the software, since it can be assessed that many would find this to be the documentation's functional purpose. In essence, this category within the Kano model will not be as difficult to address in this thesis as the latters.

Considering the performance needs, the documentation on the software also has to provide guidelines on how to visualize the data that's specifically of interest to the end-users. Instead of just simply providing guidelines within the documentation on how to use the tool, providing guidelines on how to visualize specifically the data that the interviewees are interested in will help meet the performance needs as well. In essence, the documentation has to be tailored specifically after the data desired, see the results Section 4.2, in order to increase the likelihood of meeting the performance needs.

The third category, delighters, refers to benefits that are unexpected. Unlike the performance needs the delighters can only increase satisfaction, but the absence of a delighter can never cause dissatisfaction. For the documentation on EazyBI, delighters would be the tools and knowledge provided that the end-users of it would not expect. One need or preference that no interviewee

mentioned, was the need for learning which type of data visualization and graph is appropriate for what type of data. All of them wanted to know how to visualize data using EazyBI in regards to technical aspects, but none referred to theoretical aspects regarding it. Yet this too is an important aspect to visualizing data, according to Tufte (2001). Considering some of the interviewees currently do not visualize data themselves they therefore might lack this theoretical knowledge. By additionally documenting theory of data visualization and not just theory on EazyBI usage, a delighter will be provided.

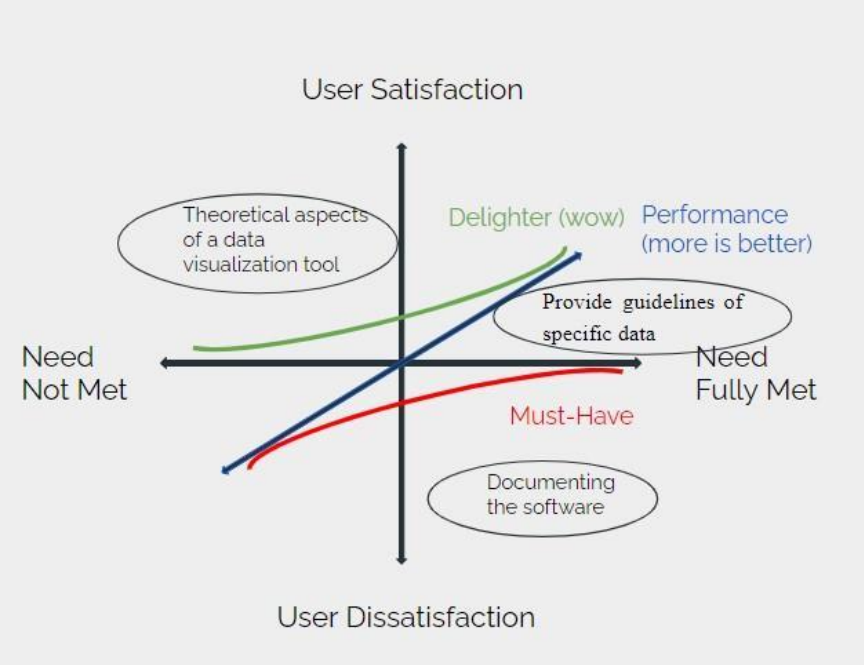

Figure 18: The Kano model applied on the created documentation

5.2 Content level

The following section aims to discuss the factors to consider when visualizing the desired data. These factors will be taken into consideration when creating the data visualizations that will be used as examples throughout the documentation.

5.2.1 The desired data

As presented in the results, the four areas of data that the interviewees presented to be of interest were Trouble reports, Number of features with F4 date, F4 precision and capability. As explained in section 3.3.4, Doumont & Vandenbroeck (2015) describe three criterias to selecting the appropriate graph when visualizing data. According to the first criteria the quality and the quantity of the variables, i.e the structure of the dataset needs to be considered when visualizing data.

5.2.2 The structures of the data sets

Regarding trouble reports, there were several attributes described by the interviewees that make up the variables relevant to trouble reports. What was of most interest was how many trouble reports are created in a period of time and how many trouble reports are resolved in a period of time.

Additionally, the interviewees also emphasized that they wanted to filter the trouble, for example how many trouble reports are created during a period of time for a specific product owner area. Lastly, they also emphasized that they generally wanted to filter the data per month, in this case how many trouble reports are created per month for a specific product owner area. Studying the interviewees statements regarding trouble reports through the perspective of Doumont & Vandenbroeck's (2015) first criteria, see section 3.3.4, the variables relevant to trouble reports are *the number of TRs created*, *the number of TRs resolved*, *the product owner area* and the time period *monthly*. Additionally, based on secondary research, the number of TRs created for instance consists of two variables: the *number of issues created* and the *issue type trouble reports*.
The other data which was shown interest in by the interviews is *Features with F4 date*. The interviewees emphasized that they mainly wanted to filter the Features with F4 date per month, which refers to how many features were completed a specific month to be able to compare. Furthermore, they also showed interest in filtering on different teams, projects, and PO-areas, while keeping the filter on month. With knowledge collected from secondary research, the number of Features with F4 date per month consists of the variable *Issues with F4 Date* and the time period *monthly.*

Further on, the result showed a need for visualizing *F4 precision,* where the interviewees expressed a need for visualizing and comparing how good they are at estimating F4 Date at different times of the ongoing process. They call the different stages of an ongoing project for F1, F2, F3 and when it is finished it is called F4. It can be assessed that the F4 precision gets more accurate the closer the prediction is made to the actual F4 Date. With knowledge collected from secondary research, the *F4 precision* per month consists of the variables *Issues with F4 Date*, *Issues with F4@F2, Issues with F4@F3, Issues with F4@F1* and the time period *monthly.* The actual F4 precision is calculated as a division, with the variables "*Issues with F4@F1*", *"Issues with F4@F2" etc.* divided by the actual F4 date: *"Issues F4 Date".*

The last area of interest presented in the result is *Capability.* With capability they are referring to how many features are planned and ongoing during a period of time and with the help of that data they will be able to do more accurate planning. The variables that are needed to calculate this are again the variable of time measured monthly, the variable of *issues with an F0 date* (meaning they are planned), and the variables of *issues that are in the stages F1, F2 or F3* (meaning they are ongoing). Irrespective of the stage, the total count of features will be displayed. However, if there is a requirement to showcase the number of planned features at the initial stage, a filtering option can be applied specifically for issues tagged with F1.

Doumont & Vandenbroeck (2015) describe two different structures of data sets: discrete and continuous. Almost all of the variables identified in all areas of data are considered discrete, since they all take on a finite or countable value. This includes variables such as *the number of TRs created, Issues with F4 Date, Issues with F4@F3* and *issues with an F0 date.* However, there were also two continuous variables identified in the areas of data studied in the thesis, meaning that these variables could take on any value within a specific range or interval. The first one, that all of the areas of data were measured in regard to on a monthly basis, was the variable of *time*. The other continuous variable identified was the *F4 precision values*, as they are calculated through divisions, which also can take on an infinite number of values depending on how many decimals are used.

5.2.3 The Intended Use of the Graph

The intended use of the graph, also referred to the intended audience of the graph, refers to taking into consideration whether the graph is intended for personal analysis, colleagues or the public (Doumont & Vandenbroeck, 2015). What Doumont & Vandenbroeck mean by this is that certain

graph types present data better for personal analysis, while a graph towards an audience such as colleagues or even the public might want to convey a different message. All of the graphs in this thesis are created for the purpose of their data being visualized to the interviewees. Considering this only, the intended audience of the graphs would be the interviewees themselves. However, since the purpose of the documentation created is to enable the interviewees to visualize data themselves, the graphs included in the documentation should have a graph type that is suitable for the intended use of personal analysis instead. To put it simply, the graphs created in this thesis will serve as examples of how to create graphs for personal analysis. Therefore, the graph types should be chosen accordingly for *personal use* as the audience.

5.2.4 The Research Questions for the areas of data

The third criteria, the *research question*, also known as the intended message, refers to the graph not only acting as data storage or decoration but also answering questions for the intended audience identified above (Doumont & Vandenbroeck, 2015). The research questions are explained to be categorized in four different groups: comparison among individual data, distribution of data along a scale, correlation between variables, and evolution over time of a variable. Additionally, Duomont & Vandenbroeck (2015) also explain in section 3.3.4 that different research questions can be combined with one another, which is the case with several of the areas of data discussed in the thesis.

Regarding the trouble reports, the interviewees show interest in *comparison among individual data* since they express the need for looking at different data i.e issues created vs issues resolved. They are also interested in *evolution over time of a variable*, where the trend is valuable feedback for them to see how they progress. Lastly, the interviewees also expressed an interest in comparing data regarding trouble reports for different product owner areas, which would also be considered Duomont & Vandenbroeck's (2015) first research question of *comparison among individual data*. For calculating how many issues with an F4 date exists per month, the research question identified is simply the evolution of this variable over time, in other words the fourth research question described by Duomont & Vandenbroeck (2015). For the F4 precision, the interviewees described an interest in both how the F4 forecasts predicted at the different stages of an issue or project compared to each other, and also if the F4 forecasts made at certain stages had become more or less accurate over time. Using the four research question categories provided by Duomont & Vandenbroeck (2015), the interviewees' interests in F4 precision would be categorized as a combination between *comparison among individual data* and the *evolution over time of a variable*. Regarding Capability, the interviewees find the *evolution over time of a variable* the most interesting, where again the trend of the capability is valuable feedback for them. By following the development of planned and ongoing issues, it is possible to make better predictions of the capability and plan their resources.

5.2.5 Choosing appropriate graphs and applying the data visualization guidelines

When using all the theoretical frameworks provided by Duomont & Vandenbroeck (2015) as it can be seen above, the following graph types were identified as the most appropriate ones:

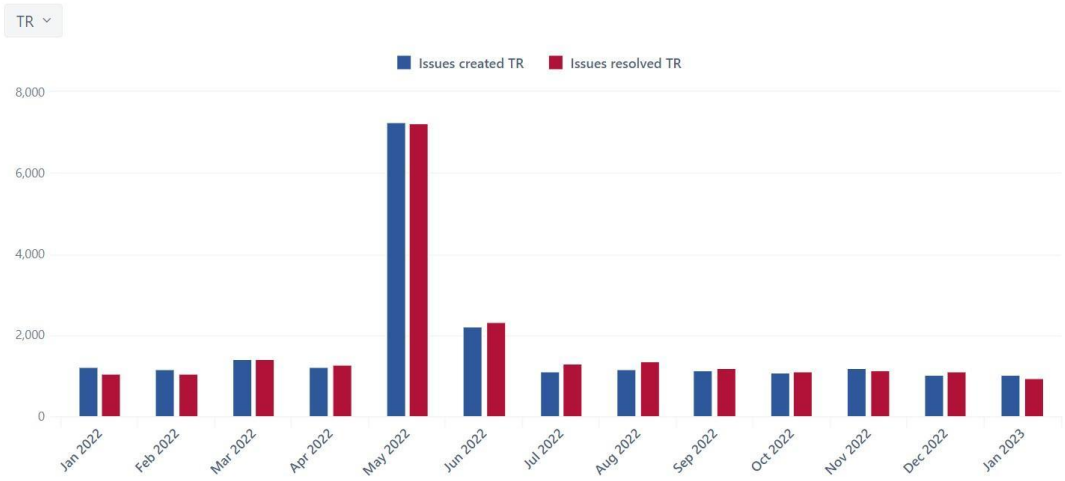

*Figure 18: **Bar graphs** chosen for "Number of trouble reports created vs resolved"*

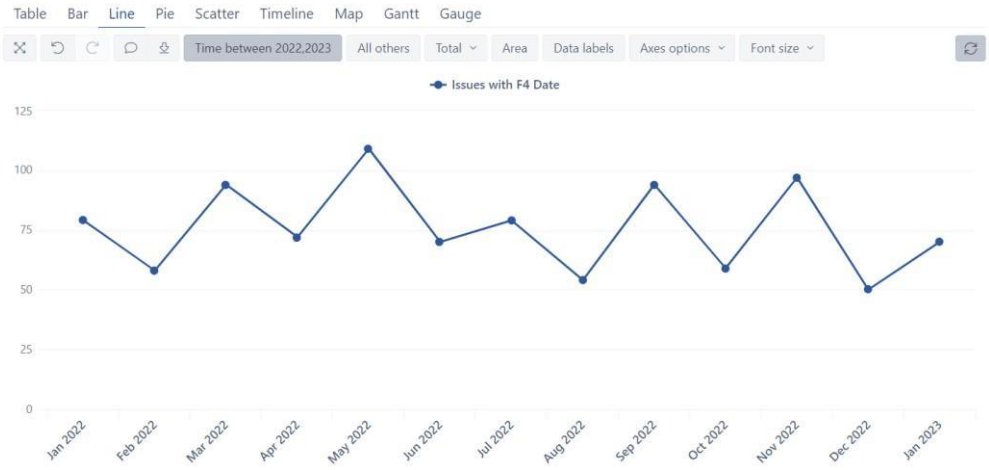

*Figure 19: **Line graph** chosen for "Number of issues with an F4 Date over time"*

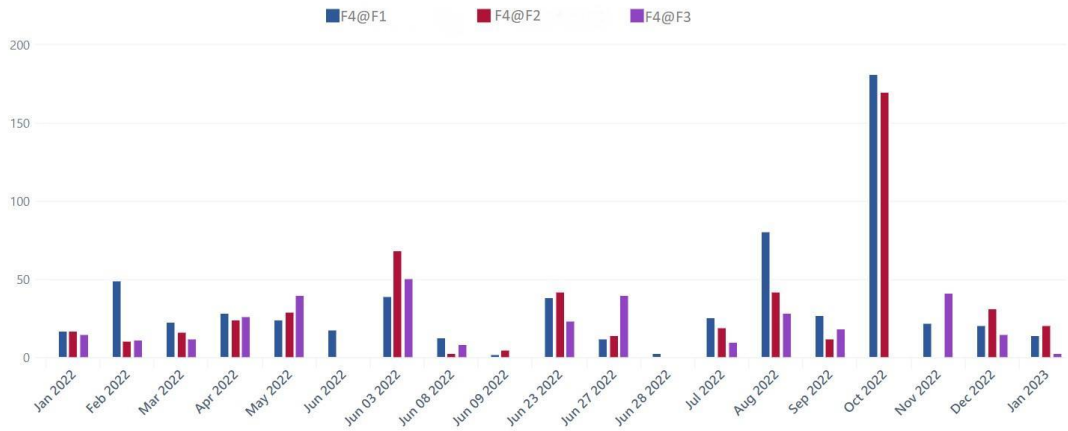

*Figure 20: **Bar graph** chosen for "F4 precision"*

*Figure 21: **Line graph** chosen for "Capability"*

Selecting appropriate graphs based on theoretical frameworks provided by Duomont & Vandenbroeck (2015) is a preferred initial step. Using the literature provided by Duomont & Vandenbroeck (2015) in Section 3.3.4, as it can be seen in the graphs selected above, line graphs were chosen when the trend over time for a variable was specifically of interest, but when comparing different sets of data to each other bar graphs were deemed more suitable instead. However, it is worth noting that practical situations can present exceptions where alternative graph types prove more suitable. For instance, since trouble reports do not only refer to the comparison between resolved and created but also trend over time, one could add a linear trend to each set of data within the graph as well. This could however be considered too messy and would contradict Kelleher & Wageners (2011) first guideline of the data to ink ratio, see the picture below.

Figure 22: "Number of trouble reports created vs resolved" illustrated with both a bar graph and a linear trend

Furthermore, when the data desired was visualized in this thesis, the guidelines and recommendations of both Tufte (2001) and Kelleher & Wagener (2011) were considered, see Section 3.3.3. It is important to emphasize however that EazyBI in and of itself is highly intuitive, and many of the visualization guidelines provided by Tufte and Kelleher & Wagener are adhered to as default. For instance, Tufte (2001) emphasizes that well executed data visualizations should remove insignificant information, which is an assessment that aligns with the concerns addressed by Matheus et al. (2020) in their examination of dashboard risks (Section 3.3.2.1). EazyBI was deemed to be intuitively very good at this, as well as at efficiently establishing a visual hierarchy with font size and color, even though attributes such as those also could be manually altered. Tufte (2001) also suggests the ability to quickly "zoom in and out of data" to be important, meaning that the data visualization should allow for both an overview as well as for more detailed analysis. While EazyBI does not offer any option to visually zoom in and out of graphs, there is the option of "drill through issue", which allows the user to quickly and easily identify for instance the data that a bar in a bar chart consists of. Lastly, almost all of the 10 guidelines provided by Kelleher & Wagener (2011) are already adhered to intuitively by EazyBI except for one: the ninth guideline. When two sets of data are visualized simultaneously but also differ from each other a lot in their averages, EazyBI will still only create one graph. In situations where this occurred, the ninth guideline was considered by separating the two data sets in order to allow for better analysis. Additionally, the ninth guideline was also explained specifically within the documentation, as advice for the end users to consider when encountering similar situations.

6. Conclusion

In this chapter, the research questions will be answered. Also, the aim in this chapter is to encompass a discussion of the thesis' limitations and recommendations to overcome them. Additionally, the chapter will propose potential avenues for the future development of the research presented in this thesis.

6.1 Research conclusion

This thesis aimed to address the challenges faced by the case company in relation to data visualization by providing documentation of a data visualization tool, EazyBI, to relevant employees. By adopting a content and process perspective, the goal was to optimize the structure and presentation of the documentation in order to maximize user engagement, as well as to visualize the desired data in order for the documentation to present the process of doing so.

RQ1: What factors to consider for maximizing user engagement?

In regards to the first research question of what factors to consider for maximizing user engagement, several factors were identified to be important. The process level customer needs of the documentation were identified to be simplicity and efficiency, informative and descriptive of a standardized work process of how to visualize data using EazyBI. Using the Kano model, the performance needs of the documentation were identified to be to what degree the documentation included visualization processes of the data desired. This further emphasized the foundation for the second research question in that the desired data had to be identified. Furthermore, the delighters of the Kano model for the documentation were identified to be theoretical knowledge of data visualization, as none of the interviewees mentioned this as a request or expectation of the documentation and would therefore be surprised if it was included.

By applying the usage prediction models to the answers of the interviewees, it can be assessed that the TRA framework in this case can be neglected, while the TPB framework and TAM offer valuable insights in how to maximize user engagement for the documentation, additionally to the ones presented above. By using TPB, it can be assessed that several of the interviewees primarily lacked the opportunity and resources to learn EazyBI, which the documentation could offer a solution to by being easy to navigate around quickly and efficiently as well as relevant to what data they are interested in. By assuring the data visualized within the documentation to be relevant to the interviewees, as well as presenting the documentation and EazyBI in a simple and comprehensible way, the perceived usefulness and the perceived ease of use of the technology acceptance model will also increase.

RQ2: What factors to consider when visualizing the data desired?

When visualizing the desired data, four areas of data were firstly identified to be of most interest: trouble reports, the number of issues with an end date per month, end date forecast precision and capability. When visualizing this data, the three criteria for choosing an appropriate graph by Doumont & Vandenbroeck (2015) were primarily considered. These three criteria are the structure of the data

sets, the intended use of the graphs and the intended research questions of the graphs. The structures of the four data sets identified to be of interest were all very similar, with all of the variables included being discrete, except for the variable of time and the division between the F4 forecast and the actual F4 date, i.e the variable of *F4 precision*. Additionally, since the aim of the documentation was to enable managers such as the interviewees to visualize data for personal analysis, the personal use was therefore considered as the intended use and audience of the graphs. Lastly, there were two main intended messages of the data sets identified to consider when visualizing: showcasing trends over time for Capability and issues with F4 date, as well as facilitating dataset comparisons and trends over time for Trouble reports and F4 precision. Using these three criterias, bar graphs were chosen for the visualization of trouble reports and F4 precision, while line graphs were chosen for visualization of capability and the number of issues with an F4 date. Besides the three criterias of choosing an appropriate graph presented by Doumont & Vandenbroeck (2015), the data visualization guidelines presented by Tufte (2001) and Kelleher & Wagener (2011) were also identified as factors to consider when visualizing the desired data.

6.2 Limitation

One limitation of the documentation created in this thesis, is that the desired data and the visualization process of it is not validated by a large and diverse group of users. When the documentation is available to its intended users, there is no guarantee that its content will be adopted by the greater mass. Early adopters may provide valuable feedback on the usability and effectiveness of the documentation, but it is difficult to determine how well the documentation will perform in a wider range of scenarios. The early adopters initiating this thesis may be biased since they are described as an “enthusiast” in the area. Even if the documentation is well-designed and effective, there is no guarantee that it will be widely adopted or used within an organization. This can be due to a variety of factors, including resistance to change or lack of awareness of the documentation. Furthermore, there is no guarantee that the documentation alone will actually lead to increased EazyBI usage. There may still be a lack of resources or capacity to actually integrate the tool effectively in the daily work of the employees. One risk is that the intended users do not see the value of using the documentation to create their own data visualizations, perhaps because they believe it still requires too much effort to learn.

6.3 Future work

One area of the thesis which could benefit from future work is the areas of different learning methods. Besides the content level of the documentation, the thesis on a process level primarily focuses on how to increase usage of the documentation and on how to make it more appealing, while the dimension of how different individuals learn softwares effectively is still very much relevant to their satisfaction with the documentation. By incorporating even more interactive and user-generated elements into learning materials, with the support of theory, we may be able to create more engaging and effective learning experiences that are better tailored to the needs and preferences of individual users. One example of another learning method is video tutorials instead of only written documentation. This approach could be particularly valuable in a world where remote learning and digital education are becoming increasingly prevalent.

Additionally, it is important to recognize that the existence of documentation of a tool is just one component that leads to increased uses of that tool. Both the tool and the documentation of it need

to be promoted as well. Future work should also delve into organizational change and consider how these materials are implemented and integrated into existing workflows and processes. Successful adoption and usage will require a broader shift in organizational culture and processes as well as the development of clear goals to create a culture of continuous learning and improvement.

References

Backman, J. (2008) *Rapporter och uppsatser* (2. uppl.). Studentlitteratur AB.

Bayus, B.L (2008). Understanding Customer Need. Scott, S: *The Handbook of Technology and Innovation Management (*ss. 115-130). Wiley-Blackwell.

Bharathy, M. S. (2021). *Relationship between Theory of Reasoned Action and Theory of Planned Behaviour.* XIBA Business Review, 4(1), 34–37.

Bryman, A., & Bell, E. (2015). *Business research methods*. Oxford University Press.

Creswell, J. W. (2009). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches (3rd ed.)*. Thousand Oaks, CA: Sage Publications.

Dalen, M. (2015). *Intervju som metod* (2. uppl.). Gleerups.

Davis, F. (1987). *User Acceptance of Information Systems: The Technology Acceptance Model (TAM)*. 1st ed. [ebook] Ann Arbor: University of Michigan.

Dillon, A., & Morris, M. G. (1996). *User acceptance of new information technology: theories and models.*

Doumont, J. L., & Vandenbroeck, P. (2015). Choosing the right graph. *Writing and Speaking in the Technology Professions: A Practical Guide,* 117-122.

Gelman, A., & Unwin, A. (2013). Infovis and statistical graphics: different goals, different looks. *Journal of Computational and Graphical Statistics*, 22(1), 2-28.

Goldmeier, J., & Duggirala, P. (2015). *Dashboards for Excel*. Apress.

Kelleher, C., & Wagener, T. (2011). Ten guidelines for effective data visualization in scientific publications. *Environmental Modelling & Software*, 26(6), 822-827.

Keong, Y. W., & Husin, M. M. (2019). Should Past Experience be Ignored? An Insight from the Decomposed Theory of Planned Behavior. Global Business and Management Research, 11(3), 1418.

Löfgren, M., Witell, L., & Gustafsson, A. (2011). Theory of attractive quality and life cycles of quality attributes. *The TQM Journal*, 23(2), 235-246.

Matheus, R., Janssen, M., & Maheshwari, D. (2020). *Data science empowering the public: Datadriven dashboards for transparent and accountable decision-making in smart cities.* Government Information Quarterly, 37(3). https://doi.org/10.1016/j.giq.2018.01.006

Misiurek, B. (2016). *Standardized work with TWI: eliminating human errors in production and service processes.* CRC Press.

Olsen, D. (2015). *The lean product playbook: How to innovate with minimum viable products and rapid customer feedback*. John Wiley & Sons.

Pauwels, K., Ambler, T., Clark, B. H., LaPointe, P., Reibstein, D., Skiera, B., ... & Wiesel, T. (2009). *Dashboards as a service: why, what, how, and what research is needed?*. Journal of service research, 12(2), 175-189.

Pettigrew, A. M. (1987). Context and action in the transformation of the firm. Faulkner, D: *Strategy: Critical Perspectives on Business and Management* (ss. 419-457). Routledge.

Rådman, E., & Johansson, E. (2020). *What is Quality in Coworking Spaces?: Identification and Classification of Customer Needs*. [Master thesis, Chalmers university of technology]. Källa. https://odr.chalmers.se/server/api/core/bitstreams/b7d29165-9109-4ad1-868397298e29c01c/content

Taylor, S., & Todd, P. (1995). *Assessing IT usage: The role of prior experience.* MIS quarterly, 561570. https://doi.org/10.2307/249633.

Tufte, E. R. (2001). *The visual display of quantitative information*. Graphics Press.

Tukey, J. W. (1993). *Graphic comparisons of several linked aspects: Alternatives and suggested principles.* Journal of Computational and Graphical Statistics, 2(1), 1-33.

Zaman, U., Zahid, H., Habibullah, M. S., & Din, B. H. (2021). *Adoption of big data analytics (BDA) technologies in disaster management: a decomposed theory of planned behavior (DTPB) approach*.

Cogent Business & Management, 8(1), 1880253.

Appendix

Interview questions

What is your occupation/title?

What is your perception of data visualization in general and at [the case company]?

When speaking to Interviewee X, she believes that the main advantage of introducing the data visualization software Eazy BI to more Ericsson employees is that they will be able to visualize and track data regarding planning, capability and quality within only two minutes instead of having to write a ticket to someone else like the metrics team first. What does this statement mean to you? How valuable would you consider the documentation to be if it led to this?

What data are you interested in (visualizing)? What measurements are interesting to you in your work?

Regarding what data you are interested in having visualized, is the purpose of this data to help you to view trends or rather track yourself or is it to compare to others? How many other people are you interested in comparing to in that case? In other words, how specific/tailored or how general do you need the data and the metrics to be?

What do you believe your colleagues' perception of data visualization and the need for it is? What do you think your co-workers would think of this guide? Do their perceptions align with yours, and do you believe their opinions on data visualization have influenced yours?

What do you believe documentation on how to use Eazy BI and visualize data would lead to for you? How could it affect you? Do you consider these consequences positive or negative? Would you say that there is a demand for this documentation?

What is your perception of challenges when learning data visualization?

- Are there available tools/resources out there for learning it?
- How much do you know about data visualization already? What is your perception of how much you would need to learn? Do you feel like the knowledge gap is large or small? Do you feel like you have the opportunity to learn data visualization? Or is something making it more challenging, like for example if you feel like you don't have enough time to learn it?

Do you perceive Eazy BI and being able to visualize data to be useful?

Do you perceive data visualization in general and Eazy BI to be difficult, or easy to use?

Wich needs would you consider most important for the documentation to meet in order to maximize its usage?

DEPARTMENT OF TECHNOLOGY MANAGEMENT AND ECONOMICS
DIVISION OF SERVICE MANAGEMENT AND LOGISTICS
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2023
www.chalmers.se

CHALMERS