isec Engenharia

DEFINITIVO

MESTRADO EM INFORMÁTICA E SISTEMAS

Evaluation Methodology for Visual Analytics Software

Autor

Antonio Augusto de Mélo Bessone de Oliveira

Orientadores

Prof. Doutor Jorge Bernardino

Prof. Doutora Marina Tropmann-Frick

INSTITUTO POLITÉCNICO DE COIMBRA

INSTITUTO SUPERIOR DE ENGENHARIA DE COIMBRA

Coimbra, março, 2021

DEPARTAMENTO DE INFORMÁTICA E SISTEMAS

Evaluation Methodology for Visual Analytics Software

Metodologia para Avaliação de Visual Analytics Software

Dissertação para a obtenção do grau de Mestre em Informática e Sistemas

Especialização em Tecnologias da Informação e do Conhecimento.

Autor

Antonio Augusto de Mélo Bessone de Oliveira

Orientadores

Prof. Doutor Jorge Bernardino

(ISEC – Instituto Superior de Engenharia de Coimbra)

Prof. Doutora Marina Tropmann-Frick

(HAW - Universidade de Ciências Aplicadas de Hamburgo)

INSTITUTO POLITÉCNICO
DE COIMBRA

INSTITUTO SUPERIOR
DE ENGENHARIA
DE COIMBRA

Coimbra, março, 2021

ACKNOWLEDGMENTS

I would like to take this opportunity to first and foremost thank God for being my strength and guide in the writing of this thesis. Without Him, I would not have had the wisdom or the physical ability to do so.

Every thesis is a journey, and I would like to thank all the people that accompanied and supported me during mine. The Erasmus protocol between Coimbra Institute of Engineering (ISEC) and Hamburg University of Applied Sciences (HAW) gave me the rich experience not only to develop part of this work abroad but also to work with two supervisors.

So, I leave a special thanks to my supervisors Prof. Dr. Jorge Bernardino and Prof. Dr. Marina Tropmann-Frick for all their kindness, encouragement, counselling, time, consideration, and patience. Also, to Prof. Dr. Jorge Bernardino, for his availability to supervise me and for always encouraging me and supporting me in writing academic publications, which made me more prepared to write this thesis. Still, to Prof. Dr. Marina Tropmann-Frick for kindly accepting my request to work under her supervision at the HAW. This thesis would not have been possible without them.

I also thank Prof. Dr. João Cunha, Prof. Dr. Carlos Pereira, and Prof. Dr. Viriato Marques who did not hesitate to support my idea of developing my thesis abroad. Besides, Ms. Dalia Pires and her team for all the consideration, assistance, professionalism, and efforts to make my Erasmus experience possible. Also, Ms. Maike Lempka and her team for the prompt and kind assistance regarding the arrangements at the HAW's side.

My deepest gratitude goes to my family and friends, especially my parents, godmother Ester, sister Maria Edinah, brother Arthur Jorge, and cousin Sandra who were with me all the time from day zero in this challenge of pursuing a master's degree in another continent. I also give thanks for their understanding and patience regarding my physical absence in the last two years.

I am also grateful for all the people I had the chance to meet, work with and learn from during my studies at ISEC, especially for my colleague Any Keila who was a great friend and study partner up to the end of this challenge, which was the biggest one of my life. Moreover, for my roommate Gonçalo Seco, who, with his friendship and consideration, made me quickly feel integrated in Portugal, also contributing to my successful stay and studies in this country. Last but not least, I thank not only all my professors but also the ISEC's non-teaching staff for the effective and mindful support during the entire program.

This page is intentionally left blank.

RESUMO

O desafio do *Visual Analytics* (VA) é produzir visualizações que ajudem os utilizadores a concentrarem-se no aspecto mais relevante ou mais interessante dos dados apresentados. A sociedade actual enfrenta uma quantidade de dados que aumenta rapidamente. Assim, os utilizadores de informação em todos os domínios acabam por ter mais informação do que aquela com que podem lidar. O software VA deve suportar interacções intuitivas para que os analistas possam concentrar-se na informação que estão a manipular, e não na técnica de manipulação em si. Os ambientes de VA devem procurar minimizar a carga de trabalho cognitivo global dos seus utilizadores, porque se tivermos de pensar menos nas interacções em si, teremos mais tempo para pensar na análise propriamente dita. Tendo em conta os benefícios que as aplicações VA podem trazer e a confusão que ainda existe ao identificar tais aplicações no mercado, propomos neste trabalho uma nova metodologia de avaliação baseada em heurísticas. A nossa metodologia destina-se a avaliar aplicações através de testes de usabilidade considerando as funcionalidades e características desejáveis em sistemas de VA. No entanto, devido à sua natureza quatitativa, pode ser naturalmente utilizada para outros fins, tais como comparação para decisão entre aplicações de VA do mesmo contexto. Além disso, seus critérios poderão servir como fonte de informação para designers e programadores fazerem escolhas apropriadas durante a concepção e desenvolvimento de sistemas de VA.

Palavras-Chave: visual analytics, avaliação heurística, metodologia de avaliação

This page is intentionally left blank.

ABSTRACT

The challenge for Visual Analytics (VA) is to produce visualizations that help users focus on the most relevant or most interesting aspect of the data being presented. Today's society faces a rapidly increasing amount of data. Thus, information users in all domains end up having more information than they can deal with. VA software must support intuitive interactions so that analysts can focus on the information they are manipulating, not the manipulation technique itself. Visual analytic environments should seek to minimize the overall cognitive workload of its users because if we have to think less about the user interactions, we will have more time to think about the analysis itself. In light of the benefits VA applications may bring and the confusion that still exists when identifying such applications on the market, we propose in this work a novel heuristic-based evaluation methodology. Our assessment method is intended to evaluate applications through usability testes considering desirable functionalities and characteristics in VA systems. However, due to its value-driven nature, it may be naturally used for other purposes such as comparison for decision among VA applications from the same context. In addition, its criteria can serve as an information source for designers and programmers to make appropriate choices during the conception and development of VA systems.

Keywords: visual analytics, heuristic evaluation, evaluation methodology

This page is intentionally left blank.

TABLE OF CONTENTS

This page is intentionally left blank.

LIST OF FIGURES

This page is intentionally left blank.

LIST OF TABLES

This page is intentionally left blank.

LIST OF EQUATIONS

This page is intentionally left blank.

ACRONYMS

BI – Business Intelligence

DV – Data Visualization

DVS – Data Visualization Software

HCI – Human-Computer Interaction

HE – Heuristic Evaluation

InfoVis - Information Visualization

IV – Information Visualization

LE - Longitudinal Evaluation

PNNL - Pacific Northwest National Laboratory

RO – Research Objective

RQ – Research Question

SI – Semantic Interaction

UI – User Interface

VA – Visual Analytics

VADS - Visual Analytics for Decision Support

VAE – Visual Analytics Environment

VAIA - Visual Analytics for Investigative Analysis

VAS – Visual Analytics Software

VAs -Visual Analytics Score

VAST - Visual Analytics Science and Technology

VS – Visualization Software

This page is intentionally left blank.

1 INTRODUCTION

Visual Analytics (VA) is "the science of analytical reasoning facilitated by interactive visual interfaces" (Thomas & Cook, 2005). It evolved from information visualization and automatic data analysis (Keim et. al, 2009) and it is a relatively new field. Originally, it was introduced for solving challenging problems that were unsolvable using automatic or visual analysis alone (Keim et. al, 2010). The term visual analytics was coined by Jim Thomas in the research and development agenda "Illuminating the Path" (Thomas & Cook, 2005), which had a strong focus on Homeland Security in the United States.

Meanwhile, the term is used in a wider context, describing a new multidisciplinary field that combines various research areas including visualization, human-computer interaction, data analysis, data management, geospatial and temporal data processing and statistics (Keim et al., 2009). Therefore, it profits from the knowledge and methods of all these areas. VA is not easy to define, considering its multi-disciplinary nature, which comprises multiple processes and the possibility of being applicable to a wide variety of areas (Keim et. al, 2010).

Besides, as a research agenda, it brings together several scientific and technical communities from computer science, information visualization, cognitive and perceptual sciences, interactive design, graphic design, and social sciences (Chopra & Samant, 2011).

The goal of visual analytics is to turn the information overload from nowadays into an opportunity by enabling decision-makers to examine this massive information stream to take effective actions in real-time situations (Keim et al., 2010). It has been applied in many different application domains, such as economics, bioinformatics, health, and social media (Cui, 2019).

Most of the research effort and advances on VA, as well as knowledge about it, are not yet in books. Instead, they have been presented, discussed, and spread through forums, conferences, workshops, symposiums, and scientific publications. For instance, two of the most cited and relevant events is the IEEE Visual Analytics Science and Technology (VAST) Symposium and the annual VAST Challenge. Such challenge is a contest that has taken place since 2006 and aims to provide the research community realistic tasks, scenarios, and data used in analytic work, to help VA researchers evaluate their tools, and to improve and enrich interactive visualization evaluation methods (Plaisant, et al., 2008b; Visual Analytics Community, 2020).

As further examples of important events related to VA, we have the biennial workshop BELIV for Visualization ('BELIV', 2020). Also, the workshops promoted and supported by the Pacific Northwest National Laboratory (PNNL), which has pioneered VA and employs high-impact researchers for the VA field (PNNL, 2020).

1.1 Problem Statement

By analysing the first official definition of Visual Analytics (VA), which is "the science of analytical reasoning facilitated by interactive visual interfaces" (Thomas & Cook, 2005), one can realize that the idea of VA is broad, what may lead to ambiguity and confusion between Data Visualization (DV) and VA Software, and consequently difficulty on identifying whether an application is simply a DV tool or in fact a VA software. Such confusion was reported by (Keim et al., 2008), who associated it to the fact the term VA was still very recent.

We can argue that this ambiguity is still present because we could experience it when searching for VA Solutions for first hands-on experience recently. That became evident through the vast number of curated lists which recommend and evaluate alleged VA tools on the Internet. We used the word "alleged" because several of those lists consider as VA solutions any software capable of delivering data visualization, which is just one of the parts of the VA approach (Keim et al., 2008). Such confusion and misinterpretation have been still found even in recent literature. For instance, we found them in Wasesa et al. (2019), which is a recent research paper. Also, in Adair (2020), which is a webpage that gives a wrong definition of VA.

This ambiguity and confusion not only make harder the task of choosing and looking for VA software on the market, but also downgrades the concept of Visual Analytics, which goes beyond mere data visualization. Moreover, this might give a false idea to designers and developers that they have developed VA systems only because their systems have DV capabilities.

Aiming to address the problems mentioned in the previous paragraph, we propose a heuristic evaluation able to assess qualitatively and quantitively how a VA solution is contributing to the VA field according to its design. This, by integrating guidelines and heuristics related to VA into a coherent structure, which will allow a value-driven evaluation. Thus, our work has the following research objectives (ROs):

- **RO-1:** Building an assessment methodology to calculate a final score for systems according to their alignment to a criteria set which reflects desirable characteristics and functionalities in VA systems.
- **RO-2:** Helping the scientific community and the industry to:
    - Better understand what VA Systems are by presenting an alternative definition of it to minimize the confusion between Data Visualization Software and Visual Analytics Software.
    - Easily infer how much a software solution is in line with the desirable characteristics and functionalities in VA systems, in order to facilitate evaluation, identification, classification, comparison, and design of VA systems.

- **RO-3:** Serving as a clear information source, which helps with making appropriate choices when designing and developing VA software.

1.2 Methodological Approach

In order to build an assessment methodology for VA software, the starting point of this work consisted of acquiring some hands-on experience with popular Visual Analytics Solutions. For that, first, we searched for the top 5 five Self-Service for BI and Analytics solutions to install, use, and do a comparison of them based on their functionalities. This work resulted in the published paper, which is in **Appendix A**. Following this, we selected three of those top 5 for a deeper evaluation and comparison through the OSSpal Methodology by Wasserman et al. (2017). Such assessment work was also accepted for publication and it is in **Appendix B**. Thus, we could build some practical knowledge about VA and find the first directions and inspiration for our research question.

Afterwards, we performed a deep research on concepts about VA, as well as the need, reasons, and forms to assess VA software. The aim was to have a better understanding of the VA field and evaluation opportunities.

With all that information in mind, we have decided on a Heuristic Evaluation regarding VA Software design. Then came the need for deep research about Heuristic Evaluation, namely, concepts and usefulness. Also, the proper way to build, deploy, and test it. Moreover, naturally arose the need to find out the best-suited heuristics for VA software.

This background allowed us to reach the core subject of this work, an evaluation methodology as a contribution to the VA field.

1.3 Main Contributions

The main contributions of this master thesis are:

- A review of the VA approach, explaining its concepts and advantages.
- A highlight of the difference between DV Software and VA Software.
- A comprehensive overview of Heuristic Evaluation and Heuristics for VA.
- A Heuristic-based Approach to assess VA software along with guidelines to deploy it.

These are, in a general way, the main contributions that have been made with the development of this work. We also contributed to the existing literature:

- We published a paper entitled "Evaluating Self-Service BI and Analytics Tools for SMEs" (**Appendix A**) that was presented at the 17^{th} International Conference on e-Business – ICE-B'2020.
- We published a paper entitled "OSSpal Assessment of Self-service BI and Analytics Software" (**Appendix B**) that was presented at the 20th Portuguese Association for Information Systems Conference – CAPSI'2020.

1.4 Outline

The rest of this master's thesis is structured as follows. **Chapter 2** contains the fundamentals of Visual Analytics and Heuristic Evaluation. In **Chapter 3** we analyse related work, explain how the literature was performed, and explain why our evaluation methodology is novel. **Chapter 4** will present the assessment approach for VA systems. **Chapter 5** has the deployment guidelines of the proposed evaluation model and Chapter. As for **Chapter 6**, it contains the methodology use case developed to teste the feasibility of our assessment approach. Finally, in **Chapter 7**, we summarize the main findings of our research, discuss its limitations and point out some future work.

2 VISUAL ANALYTICS AND HEURISTIC EVALUATION FUNDAMENTALS

In this chapter, we present fundamental concepts related to Visual Analytics (VA) and Heuristic Evaluation (HE), so that also non-expert readers in data science can understand it.

2.1 What is Visual Analytics?

After a brief explanation of the VA concept, it is no surprise to have questions about what it actually is and in particular what is new or different about it (in comparison with DV software). Thus, for ease of understanding, we start by giving a list of what **not VA is** according to (Thomas, 2009):

- A large graph structure with no labels.
- A heat map with no labels.
- A chart with no interaction.
- An image with no semantic interpretation.
- A standalone image that does not tell a story.

By observing the list above, we can see that visualization which lacks interaction and does not allow semantic interpretation do not match with VA.

In other words, VA is not simply about presenting information. Rather, it is more a dialogue between the analyst and the data being displayed, where the visual representation is simply the interface or view into the data (Thomas & Cook, 2005). **Figure 1** illustrates such dialogue.

As a science of interaction, it utilizes Interaction as the critical glue to integrate analytics, visualization, and human judgment into algorithmic data-analysis processes (Cui, 2019; Endert et al., 2014).

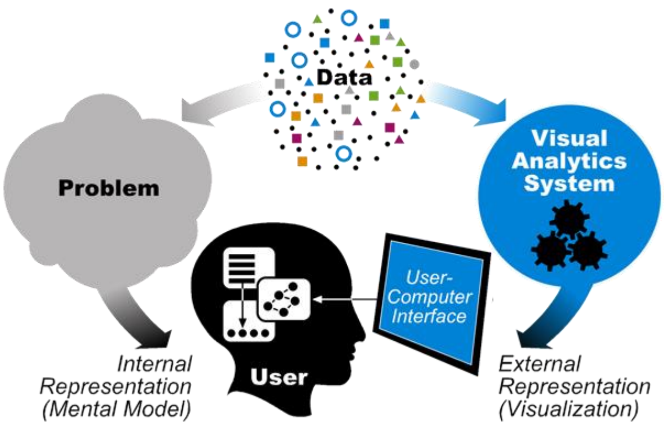

Figure 1 - VA seeks to enhance decision making, knowledge discovery, and insight (source: Visual Analytics Community, 2011).

Visual analytics is a fast-growing field of research combining strengths from visualization, data analysis, knowledge discovery, data management, analytical reasoning, cognition, perception, and human-computer interaction (Cui, 2019). **Figure 2** illustrates the detailed scope of VA.

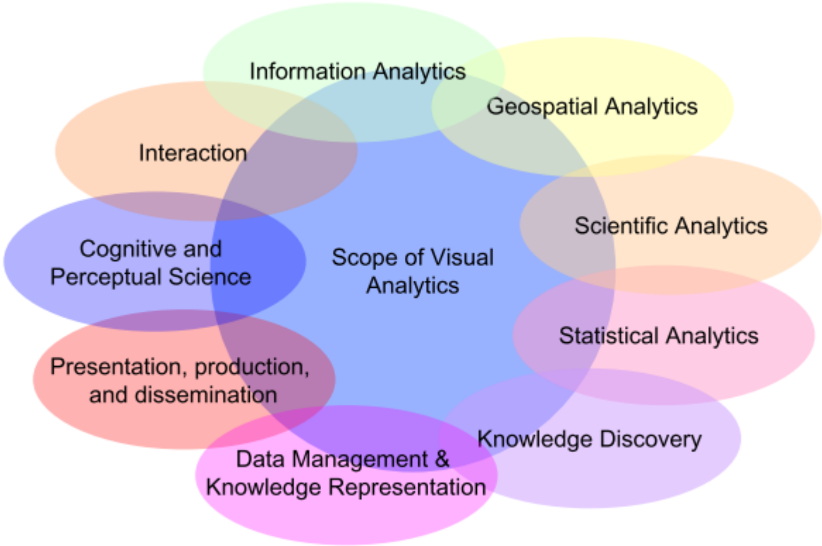

Figure 2 - The Scope of Visual Analytics (source: Mehrotra et al., 2008).

Figure 3 depicts how VA software works. The image contains the stages (colourful ovals) and their transitions (arrows) of the VA process model (Keim et al., 2010).

Everything starts with raw **data** (blue oval). To be pre-processed and transformed, the data has two ways: **data mapping** (methods for visual data exploration) **or data mining** (automatic analysis methods to discover useful patterns).

If it takes the data mining way, it will be built a **model** (green oval) from it with visual findings which will be transformed into **visualization** (pink oval).

Visualization (pink oval) allows the user to interact with the automatic methods by modifying parameters or selecting other analysis algorithms. Besides, it can be used to evaluate the findings of the generated models in **model** (green oval).

Alternating repeatedly between **visualization** (pink oval) and **model** (green oval) is characteristic for the Visual Analytics process and leads to continuous refinement and verification of preliminary results.

This step will be repeated as many times as the user judges necessary, aiming to reveal insightful information. All this process so far will result in **knowledge** (yellow oval), which will serve as support to future analyses (feedback loop). The feedback loop refers to the sensemaking loop which is further explained in **Subsection 2.2.1** of this Chapter.

If the **data mapping** way is performed first, the difference is that the user will have to confirm the generated hypotheses originated first in **visualization** (pink oval) by an automated analysis from **model** (blue oval). But the alternation between **model** and **visualization** will be also repeated as many times the user judges necessary, as it might happen to data coming first from the **data mining** way.

It is worth highlighting that this **user interaction** mentioned in this VA process relates to the concept of *Human in the Loop (HiL)* because the user will have the role of providing continuous feedback, correcting algorithmic approaches and selecting appropriate techniques during the analytical process (Tropmann-Frick & Andersen, 2020). In addition, we should emphasize that visual analytics is especially about the repeating the combination of those stages and their transitions with a focus on interaction and further refinement, as well as the proper feedback or representation of constantly resulting adaptions of the involved models and visualizations (Schütz, Raabe, Bade, & Pietsch, 2017).

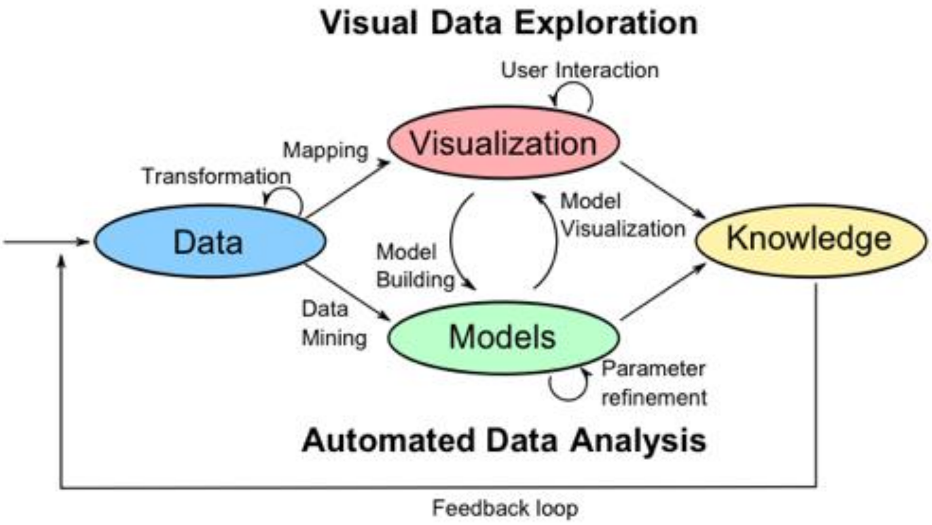

Figure 3 - VA Process model (source: VisMaster, 2013).

2.2 Other Visual Analytics Fundamentals

The two subsections below present two highly relevant approaches for the VA field: Sensemaking and Semantic Interaction. Sensemaking and Interaction have high relevance for VA systems because they support the process of making sense of data by users. That is, they direct to the insights which lead to a successful data analysis. Sensemaking is arguably the most challenging part of any data analysis process considering its explorative and creative characteristics (Fekete et. al, 2019). As for Semantic Interaction, it is a more recent approach for VA systems based on Sensemaking.

2.2.1 Sensemaking

It is an influential concept in VA. Sensemaking is the human process of acquiring an understanding of the world based on our conceptual model of events, actions, and information. In other words, Sensemaking is finding meaning from information to build comprehension. It is the construction, elaboration, and reconciliation of representations which are relevant to explain the information that is received about the world (Endert, 2014; Xu et al., 2015). The Sensemaking process in the human mind is illustrated in **Figure 4**.

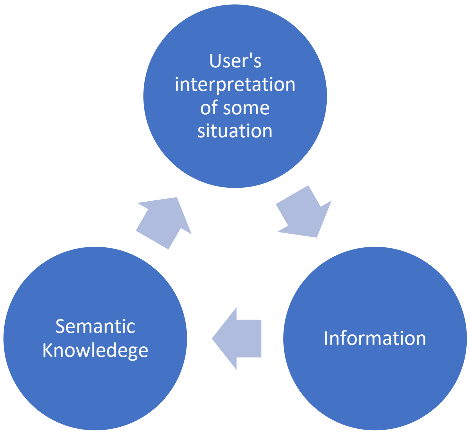

Figure 4 - Sensemaking process in the human mind.

Figure 4 represents the sensemaking process and its three basic elements: user's interpretation of some situation, information, and semantic knowledge. The sensemaking is triggered through the combination the exposure to new information and interpreted considering one's semantic knowledge. Semantic knowledge is generalized knowledge about how the world works, such as previous casual knowledge about a situation which should be interpreted. For example, observing a visualisation showing warmer air rising over colder air can make a meteorologist predict snow, but this is only possible given his previous knowledge regarding meteorology (Attfield et al., 2010).

In the VA domain, Sensemaking implies an iterative reasoning process which is guided by the continuous formulation of new hypotheses by human users of VA systems (Pohl et al., 2012). This iterative reasoning process is also mentioned as sensemaking loop in the literature on Visual Analytics. Such loop is responsible for structuring the whole knowledge discovery process supported through VA (Keim et al., 2008).

Figure 5 illustrates knowledge discovery in VA systems considering the sensemaking loop. The sensemaking loop will happen between the grey "visualization" and "user" areas in **Figure 5**. In the figure, a generic VA system performs an automatic initial analysis of inputted data and will represent it through visualization. The human perception will interpret this visualization based on generalized knowledge about the world and the data domain. This may generate insights that will be transformed knowledge. This knowledge may be submitted again to the human perception until one is able to create hypotheses. These hypotheses will stimulate further exploration and analysis of the data that will result in user interactions. These user interactions will make the system generate a new version of the visualization which will be again submitted to human perception. Then, the sensemaking loop starts again to support the iteratively the knowledge discovery. The term Specification in the figure means the specifications of the new visualization generated after the user interactions.

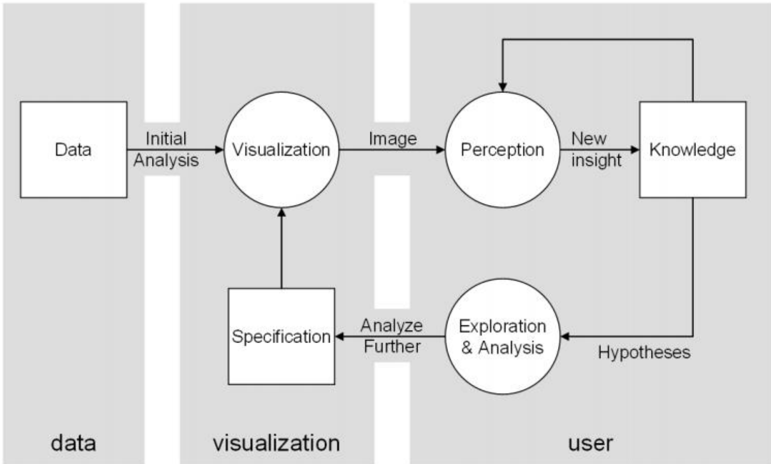

Figure 5 - Sensemaking loop supported by VA systems (source: Keim et al., 2008).

VA systems are supposed to support aspects of the Sensemaking process to better provide knowledge discovery from the data under analysis. This support comes through the systems' user interactions through the visual metaphor (see an example of visual metaphor in **Figure 6**) and underlying models (we are referring to the models in **Figure 3,** which illustrates the VA process).

2.2.2 Semantic Interaction

Semantic Interaction (SI) is an advanced approach for Sensemaking in Visual Analytics and a recent term coined in a PhD dissertation by Endert et al. (2012). SI intends to make co-reasoning between the user and the analytic models (we are referring the models in **Figure 3,** which illustrates the VA process) possible without requiring the user to directly control them (Endert, 2014).

It relates to the intuition that only humans have to put related information together. It puts the users back in the critical role within the science of VA that is essentially to level their intuition and their ability to create informal relationships as opposed to specifying implicit and formal relationships which would not level the users mental model, their understanding of the world, or things that cannot be computationally generated (Endert, 2012).

In Endert (2012), we have a video by the author of SI for VA explaining it in a VA system. According to Endert in his video, SI is supposed to happen within a visual metaphor, which is a space where both the human and the computer can understand and communicate through. See an example of visual metaphor in **Figure 6**. This VA system in **Figure 6** is called IN-Spire. It automatically displays similar documents in clusters as seen in **Figure 6**. If the user thinks that a document is placed in the wrong cluster, he/she should simply move this document to another area (cluster) and the machine learns the right place for that kind of document. When new documents with similar characteristics are submitted to the analyses, the system will already automatically know what its correct cluster. This is especially useful when dealing with huge amounts of data. In other words, the human will have then the role to judge whether relationships among data make sense or not by dragging and moving these relationships to an area in the visual metaphor where there is more similar information, and the system will learn that relationship. The VA system learns the new relationships through mathematical models which are transparent to the user.

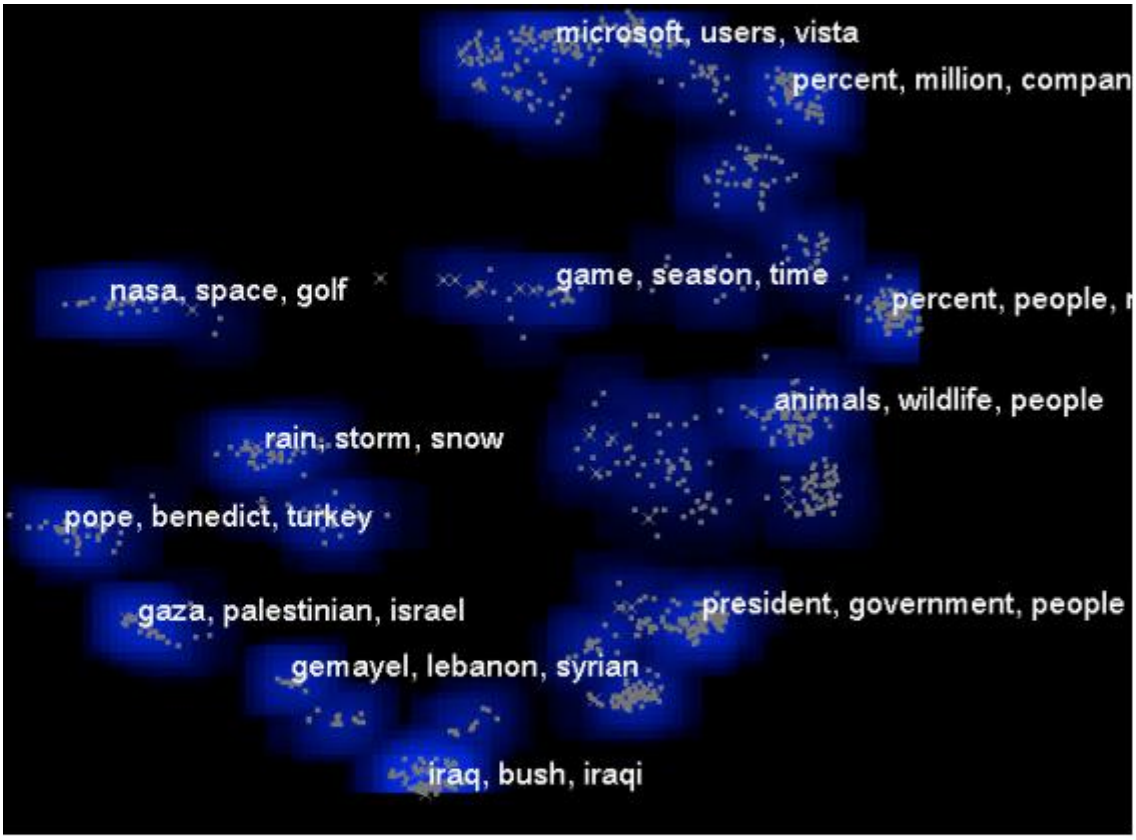

Figure 6 - Visual Metaphor of the VA system IN-Spire shown by Endert in his video (source: Endert & North, 2011).

2.3 Data Visualization Software Versus Visual Analytics Software

The use of analytics is not now restricted to big companies. Also, it has been used in the industry to support decision making. It is now significantly widespread, with 59% of enterprises using analytics in some capacity (Panoho, 2019).

We have currently an increasing and huge number DV and VA software on the market for our benefit and convenience. They are presented as Data Analytics Solutions. However, both types of tools have many aspects in common but are not the same. Given that, confusion and ambiguity on the correct identification and classification of software as DV or VA might easily occur.

As minimizing that confusion is one of the objectives of this master's thesis, we explain further in this section the difference between DV and VA software.

The difficulty mentioned above stems from the fact that both DV and VA represent data in visual interfaces and are concerned with gaining insights from data. There is actually a certain amount of overlap between them. They are two parts of the same coin that aim to make the sea of data at our disposal more understandable and effective (Sharma, 2017).

By definition, DV is a general term that describes any effort to help people understand the significance of data by placing it in a visual context (Islam & Jin, 2019). Differently from VA, DV software is not concerned to heavily integrate intelligent semi-automated data analysis methods. Besides, it is also not developed taking into account that the existence of a human in the sensemaking loop can significantly improve the outcome of an analysis, as VA does (Batch & Elmqvist, 2018).

VA and DV rely heavily on Visualization, but VA goes beyond only DV. It can rather be seen as an integral approach combining visualization, human factors, and data analysis (Keim et al., 2009). That is the major difference between DV and VA. **Figure 7** illustrates where is Visualization regarding the VA approach.

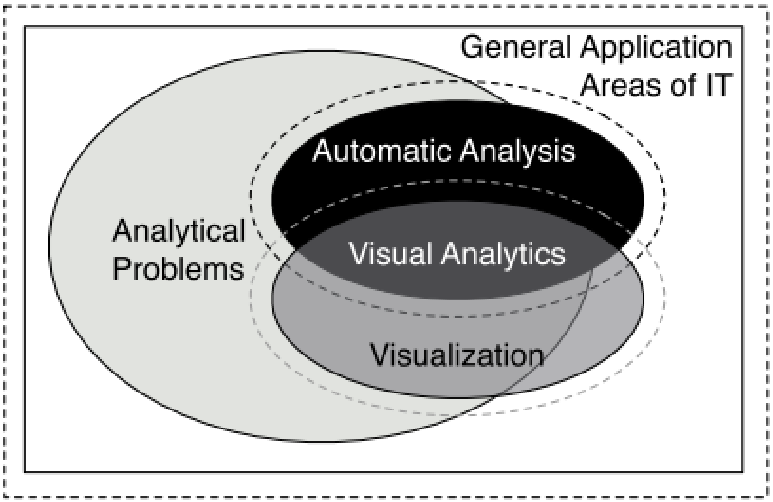

Figure 7 - A visual explanation about where Visualization inside VA is (source: Keim et al., 2009).

VA is reliant on DV, but it goes beyond the visuals to create an end to end experience that focuses on value for the user (Digimasters, 2020).

Thus, we can conclude that Visual Analytics Software is Data Visualization Software, but the opposite is not always true, once Data Visualization Software which does not includes methods for automatic data analysis do not fall in the field of Visual Analytics (Keim et al., 2008).

2.4 Heuristic Evaluation

Heuristic is a word originally from the Greek language, which means “serving to find out or discover.” In 1905, Einstein included this concept in the title of his Nobel prize-winning paper on quantum physics. He used the word to express that the view he presented was incomplete but highly useful (Gigerenzer & Gaissmaier, 2011; Holton, 1988).

Gigerenzer & Gaissmaier (2011) defined it as a strategy that ignores part of the information, intending to make decisions more quickly, frugally, and/or accurately than more complex methods.

Heuristic Evaluation (HE) was developed by usability expert Jakob Nielsen (Nielsen & Molich, 1990). It is a usability testing method developed to identify problems associated with the user interface design. It employs a set of usability design rules of thumb to be used during the usability inspection of a system (Greitzer et al., 2011). The method is a discount method where, in its original form, usability experts review a system and judge how well it meets the goals of some predefined guidelines – heuristics (Forsell & Johansson, 2010).

Heuristics can be more accurate than more complex strategies even though they process less information (less-is-more effects). A heuristic is not good or bad, rational or irrational; its accuracy depends on the structure of the environment (ecological rationality) (Gigerenzer & Gaissmaier, 2011).

HE is considered to save time and to quickly reveal design issues (Greitzer et al., 2011; Tarrell et al., 2014). It is also known to be cost-effective, intuitive, easy to learn and simple to administer. It can be used during all phases of development**,** allowing potential problems to be found and corrected before they become reality (Tarrell et al., 2014). Furthermore, it can be modified in various ways to improve the result by, for example, including domain experts or end-users, including specific user considerations, or by performing specific tasks (Forsell & Johansson, 2010).

HE is well-known and popular within fields related to VA such as Human-Computer Interaction (HCI) (Forsell, 2012; Forsell & Johansson, 2010; Greitzer et al., 2011). In Zuk et al. (2006), its usefulness was already recognized by the Information Visualization Community. Information Visualization (InfoVis) is another field extremely related to VA. In a survey amongst usability practitioners, it was rated as one of the top methods (Vredenburg et. al, 2002).

While HE cannot be used per se to evaluate a visualization application (as it is generally also the case with any other evaluation method), it has several advantages and is generally considered to provide useful results (Santos et. al, 2018).

Another aspect is that heuristics can lead to biased evaluations, which may be a result of personal judgments or experiences about how common things occur and about how representative certain things may be (Snopek, 2013).

Despite the limitations of Heuristics, Zuk and colleagues recognized that several problems would not have been discovered without the visualization specific heuristics during the evaluation of an information visualization system (Zuk et al., 2006).

This page is intentionally left blank.

3 RELATED WORK

Based on the state-of-the-art literature on Visual Analytics Software, this chapter aims to present the reasons and means to evaluate VA Systems, as well as how we performed our research for such literature. Lastly, we explain why the evaluation methodology proposed in this work is innovative.

3.1 Literature Review on Visual Analytics Software Evaluation

In this section, we present our literature review on evaluation of VA software. Our methodology to conduct this review is composed of three steps: definition of research questions, identification of studies, and studies selection.

3.1.1 Definition of Research Questions

As the topic VA was new for us and we needed to present the VA approach before entering in fact into the topic evaluation for VA Software, our research questions were not limited to evaluation methods for VA. Then, we defined the following set of five research questions (RQs) that form the objectives of this literature review.

- **RQ-1:** What is Visual Analytics?
- **RQ-2:** What is the difference between DV and VA Software?
- **RQ-3:** What are the reasons to evaluate VA Software?
- **RQ-4:** What are the advantages and disadvantages of VA Software?
- **RQ-5:** What are the possible approaches to VA evaluation?

3.1.2 Data Sources and Queries

To find the relevant studies to our work, we used five well-known online libraries. That with the intention to reduce possible search bias. For this reason, we adopted the following data sources:

- IEEE Xplore Digital Library (‘IEEE Xplore’, 2020)
- ACM Digital Library (‘ACM Digital Library’, 2020)
- Springer (‘Springer’, 2020)
- DBLP (*DBLP: Computer Science Bibliography*, 2020)
- Google Scholar (‘Google Scholar’, 2020)

Most of the studies related to VA were found on the IEEE Xplore Digital Library, once the Institute of Electrical and Electronics Engineers (IEEE) has supported so far several of the most relevant forums, conferences and symposiums for Visual Analytics, Information Visualization, and Scientific Visualization. For instance, we have the IEEE VIS 2020, which is a forum for advances in theory, methods, and applications of visualization and visual analytics that has happened since 1996 (‘IEEE VIS’, 2020).

The last two data sources work as redirects to literature hosted by the most famous scientific databases, which makes our research naturally even broader and less biased. Google Scholar was used mostly when we could not find information in the three first data sources, once its results include less reliable work.

To perform the search, we used the queries shown in **Table 1**. To exemplify the number of results which had to be analysed, such table also displays the number of results brought by the main database for VA (IEEE Xplore). Along with that, the results from DBLP, which shows results from the most relevant academic databases on computer science that host *peer-reviewed* publications.

Table 1 - Queries for the literature review on VA Software evaluation.

Queries		**Number of results**	
		IEEE Xplore	DBLP
Q1.	*evaluation of visual analytics software*	69	01
Q2.	*visual analytics evaluation*	387	49
Q3.	*visual analytics assessment*	126	18
Q4.	*challenges in visual analytics*	900	18
Q5.	*visual analytics definition*	47	01
Q6.	*visual analytics advantages and disadvantages*	08	00
Q7.	*how to evaluate visual analytics*	361	00
Q8.	*visual analytics and HCI*	30	19
Q9.	*visual analytics and human-computer interaction*	225	03
Q10.	*visual analytics and information visualization*	1448	09
Q11.	*visual analytics and InfoVis*	33	01
Q12.	*visual analytics and visualization*	2635	236
Q13.	*visual analytics and automation*	82	00
Q14.	*visual analytics assessment*	126	18
Q15.	*sensemaking and visual analytics*	69	06
Q16.	*semantic interaction and visual analytics*	47	04
Totals		**6593**	**383**

3.1.3 Studies Selection

The selection of the publications was made in three phases. In the first phase, we analysed the title and abstracts of publications, which is a quite fast method for discarding unsuitable results. In the second phase, for the sake of quality assurance, the second criterion was to adopt only *peer-reviewed* works. Finally, we read the “remainder paragraphs” (inside the introduction section) and findings/conclusions to confirm whether the publication was relevant.

From the works brought by the queries shown in **Table 1**, we collected a set of 105 scientific publications for more in-depth analysis and to extract the knowledge necessary to build and enrich chapters 1, 2, and 3 of this work.

Another key point to remember is that we have found no books focused on evaluation for VA software. Moreover, the books on general VA were collections of lectures and/or research papers (excluded the books that first introduced VA and worked as research agenda, e.g., *Illuminating the Path* (Thomas & Cook, 2005)). This likely because VA is still recent.

The response to **RQ-1** is the content for Section 2.1; to **RQ-2** is Section 2.3; to **RQ-3** is Section 3.3; to **RQ-4** are Subsections 3.3.1 and 3.3.2; to **RQ-5** is Section 3.4.

3.2 Literature Review on Heuristics and Guidelines for VA Systems

In this section, we present our literature review on heuristics and guidelines for VA systems. We used the same methodology presented in **Section 3.1**, i.e., the three steps: definition of research questions, identification of studies, and studies selection.

3.2.1 Definition of Research Questions

During our first literature review in **Section 3.1**, we found most of the studies with guidelines and heuristics that we adopted to develop our new heuristics. However, we decided to search further for studies which could give more support and strength of evidence to the new heuristics to be developed. Thus, we defined the following set of five research questions (RQs) that form the objectives of this second literature review.

- **RQ-1:** What are the Guidelines for Visual Analytics?
- **RQ-2:** What are the Heuristics for Information Visualization?
- **RQ-3:** What are the Heuristics for Visualization?
- **RQ-4:** What the Heuristics for general Interactive Systems?
- **RQ-5**: What are the Heuristics for HCI related to Analytics?

Our idea was creating a new set of heuristics from heuristics for VA systems, once heuristics are widely stated features a system or a user interface is supposed to have in the context of usability inspection. However, after finding no publications presenting standard heuristics for

VA (this lack of standard heuristics were also observed by Tarrell et al. (2014) in 2014). Thus, we collected papers with heuristics for visualization and interactive systems and general guidelines for VA systems to transform them into heuristics for VA **(RQ-1).** Following this, these new heuristics would be augmented by the heuristics found via **RQ-2, RQ-3, RQ-4, and RQ-5**, which would make our evaluation heuristics set as holistic as possible.

The last 4 **RQs** were placed deliberately for belonging to the mature fields which compound the VA approach and could therefore provide widely accepted and/or already-validated heuristics that might be highly useful for the novel heuristic set we would build for our evaluation methodology, which must be as feasible as possible.

3.2.2 Data Sources and Queries

For the identification of studies, we used the same data sources presented in **Subsection 3.1.2**. Moreover, we have **Table 2** to exemplify the number of results which had to be analysed from two relevant academic databases, as explained for **Table 1**.

Table 2 - Queries for the literature review on Heuristics and Guidelines for VA

Queries		Number of results	
		IEEE Xplore	DBLP
Q1.	*guidelines visual analytics*	65	03
Q2.	*heuristics information visualization*	102	05
Q3.	*heuristics visualization*	261	16
Q4.	*heuristics interactive systems*	134	02
Q5.	*HCI Analytics*	50	80
Q6.	*human-computer interaction analytics*	260	05
Totals		**872**	**111**

3.2.3 Studies Selection

The selection of publications was similar to the one presented in **Subsection 3.1.3.** From the total of works shown in **Table 1 and Table 2**, we selected publications with heuristics and guidelines applicable to the Visual Analytics field. The response to the research questions (RQs) presented in this section is the scientific works shown in **Table 4**.

3.3 Why Evaluate Visual Analytics Software?

As VA systems come from a relatively new field, aim at knowledge discovery, and require assessment efforts targeted at three different levels (component, system, and work environment), their evaluation is still a big, current, and complex challenge.

By researching on the state-of-the-art of VA evaluation methodologies/guidelines, we can already find invaluable contributions on that sense. However, to the best of our knowledge, solid evaluation guidelines and findings are still lagging (Van Wijk, 2013). Additionally, VA researchers generally highlight that their resulting efforts on building VA assessment methods are just the first steps in that direction. As a result, deeper research on this topic becomes highly important.

As the VA field is recent (2005), the first motivations to evaluate VA software remain current. Then, we found it might be interesting to present them chronologically, from beginning to date.

In 2005, Cook released the first Research and Development Agenda for Visual Analytics, which, stressed the importance of evaluation methodologies for VA, as evaluation had shown to play a critical role in shaping the research and enabling rapid technology transition.

Moreover, he stressed that the benefits of incorporating evaluation as part of the VA research program include: possibility to verify research hypotheses, increased communication among academia, industry, and government, a mean to compare technical approaches, identification of the most promising research approaches, an verification of research hypotheses (Thomas & Cook, 2005).

In the following year, Scholz emphasized the importance of evaluating VA software by announcing it would be advantageous to develop metrics and methodologies to help researchers measure the progress of their work and understand the impact their work would have on the users who would work in VA environments (Scholtz, 2006).

In 2006, to move research into practice, Thomas and Cook recommended that new tools, algorithms, and techniques must be evaluated to ensure that they represent a significant advance over current practice and that they operate correctly (Thomas & Cook, 2006).

In Keim et al. (2009), it was indicated again that evaluation of VA Software was still a challenge to be addressed, once an evaluation as a systematic analysis of usability, worth, and significance of a system would be crucial to the success of visual analytics science and technology. Besides, the development of abstract design guidelines for visual analytics applications would constitute a great contribution.

In 2012, Dill et al. (2012) stated that the VA field already had reached enough maturity that several visual analytics software tools were widely available. At that moment, both private industry and government organizations were exploring the capabilities offered by visual analytics to enable more effective ways of gaining insight from complex information. As it is critical to measure the benefit of the technology to the analyst in producing an improved product (Thomas & Cook, 2005), the more VA software is present on the market, the more important is to evaluate its state-of-the-art versions to encourage improvement and naturally more advances in the VA field.

Yet in that year, Kerren and Schreiber emphasized the lack of new methodologies and, if possible, benchmark data sets to evaluate VA systems. They reiterated that it is even more difficult to assess VA software because of the increased number of disciplines and the diversity of input data sets, potential users, and task to be solved (Kerren & Schreiber, 2012).

Cui (2019) constructed a complete overview of visual analytics based on over 200 publications. In this survey, he mentioned that due to the complexity of the visual-analytics process, there are still no widely accepted evaluation techniques to ensure the trustworthiness of the visual-analytics process. Besides, to address the VA evaluation challenge, he suggested the development of evaluation standards for VA by selecting and combining proper evaluation methods from the fields of visualization and algorithmic data analysis (Andrienk o, Andrienko, Keim, MacEachren, & Wrobel, 2011).

As we can see above, despite the significant progress of the VA field, there is no great advance in developing evaluation methodologies yet. As a result, the VA evaluation challenge remains an open and valuable research opportunity.

3.3.1 VA Benefits and Scientific Evidence

The great benefits VA Software may bring justify the need to develop methodologies to contribute to its continuous improvement. Then, we found to be appropriate presenting VA benefits and scientific evidence of its utility.

Below, we have a non-exhaustive list of VA software benefits (Infinity, 2020). In summary, VA software:

- improves data exploration, minimizes the overall cost and improves the data analysis.
- makes easier the bulk of complex information for better decisions.
- enables users to understand data much more quickly and to make faster, better decisions.
- offers more accurate results for more profitable decisions.
- offers different trends of visualization, so the understandable data presentation modes are guaranteed.

Following this, we present evidence regarding the benefits provided by VA.

In Chinchor et al. (2012), it was reported a case study explaining how VA software was deployed in a media company according to Moore’s Software Adoption Life Cycle (Moore & McKenna, 1999). That had to include managing cultural changes, risk-averse managers, pilot evaluations, training, and gradual deployments. The organization had collected and reported on global traditional media for more than 50 years. By the end of the deployment period, the company recognized that VA software had helped to successfully handle its large amounts of data. That recognition came through an award granted to the responsible team for implementing VA software.

In 2017, Adagha and colleagues mentioned that recent advances show that application of VA tools can facilitate decision making in real-world settings (Adagha et al., 2017).

Dasgupta et al. (2017) in a comparative study of the level-of-trust of domain scientists in visual analytics systems as opposed to more familiar manual analysis methods, demonstrated that for complex tasks, regardless of experience and familiarity, the average level-of-trust in visual analytic systems exceeds the same in manual data analysis methods.

3.4 How to Evaluate Visual Analytics Software?

As explained in **Chapter 2**, the abilities of human and machine combined characterize VA software. The behaviour of VA systems will vary depending on the user who operates them, differently from automated algorithmic solutions which are not supposed to learn from user interactions. This gives to VA systems a hybrid aspect which complicates the development of evaluation methods (Khayat et. al, 2019). Also, the VA multidisciplinary aspect makes the development of evaluations methodologies substantially difficult. According to Thomas & Cook (2005), VA systems are complex and require evaluation efforts targeted at different levels: the component level, the system level, and the work environment level **(Figure 8**).

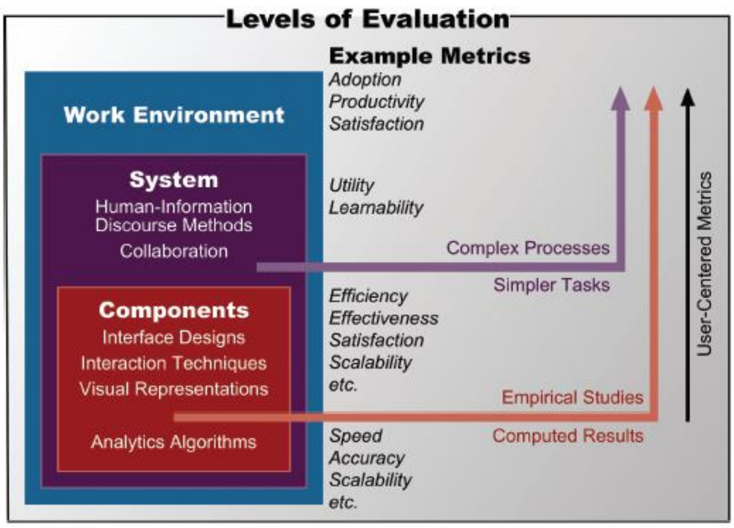

Figure 8 - VA Levels of Evaluation (source: Thomas & Cook, 2005).

As we can see in Figure **5, the component level** consists of individual algorithms, visual representations, interactive techniques, and interface designs. Data analysis algorithms can often be evaluated with metrics that can be observed or computed (for example, speed or accuracy). **The system level** includes interfaces that combine and integrate multiple components and need to be evaluated by comparing them with technology currently used by target users. **The third level** is the work environment level, where evaluation addresses issues related adoption (Plaisant et al. 2008a).

In Munzner (2009), it was suggested the need of evaluating systems which deliver visualizations at four levels: the problem level, the abstraction level, the encoding and interaction level, and the algorithmic level. According to her, each of these levels takes a different type of evaluation technique and different metrics. Munzner notes that the abstraction level needs to be evaluated with real users doing real work in order to obtain information on the utility of the system.

Thus, after being aware of how complex evaluating VA Software could be, we saw we should consider the possible types of methods to evaluate it.

In Khayat et al. (2019), it is presented a survey of evaluation practices used with summative intentions in VA. Thereby, VA assessment methods are divided into 8 categories: Theoretical Methods (THEO), Quantitative User Testing (QUT), Quantitative User Opinion (QUO), Insight-based (INST), Case Studies (CASE), and Inspection Methods (INSP). They concluded that INSP are the most feasible evaluation practice and that they are the same level of quality (considering validity and generalizability) as Insight-based and Case Study methods. It shows as well that the latter ones are less feasible than the INSP method. According to this publication, INSP is the method which uses a set of identified heuristics for researchers to inspect and evaluate a solution to judge whether it satisfies the identified heuristics. The results of Khayat and colleagues' work, come from 182 studies reported in 82 papers. Furthermore, nearly 30% (50) of such studies are INSP type, which is significant evidence that INSP is a well-adopted and useful assessment method for VA.

In Wall et al. (2019), the authors discuss reasons to adopt a Heuristic Evaluation (HE) instead of a Longitudinal Evaluation (LE) for Visualization software (VS). On one hand, they argue that LE seeks to move beyond the limitations of short-term, lab-based evaluations such as HE. Also, the significant power and potential benefit of LE for helping to determine software utility once the use of the system is observed "in the field" as people apply it to real data and problems. On the other hand, they stress that LE may be logistically challenging, very time-consuming, and pragmatically difficult to implement.

As we have mentioned in **Section 2.4**, HE has several advantages, among them the quality of being inexpensive and fast, which are desirable attributes for software evaluation (Scholtz, 2011).

Another advantage of HE is that not only system experts might perform it effectively, but also non-expert system users. In Tarrell et al. (2014), the author reports that system users had been included as part of an evaluation team. He concluded that although system users were not formally trained in usability procedures, they were still able to effectively perform HE. To support his argument, he refers that in Corrao et al. (2010) over 90% of problems identified by novice users of an information system (not usability experts) were accepted as valid, including several system bugs, missing items, or unaccommodated regulatory requirements.

Another point in favour of HE is regarding the participant pool size needed to perform a HE. That is, according to Nielsen (2000) and Nielsen & Landauer (1993) around 5 evaluators are enough to do a HE. This number is based on a statistical formula that claims 5 evaluators will discover around 75% of the overall usability problems. Nielsen and his colleague also mention that we do not gain that much additional information to motivate using larger numbers of evaluators. This was confirmed in the HE developed and validated recently in Wall et al. (2019).

In Scholtz (2011), the author highlights that HE not only can they be done with fewer participants, but also much earlier in the design of the software. Thus, in light of the aforementioned considerations in Corrao et al. (2010), Khayat et al. (2019), Nielsen (2000) Scholtz, 201, Tarrell et al. (2014), and Wall et al. (2019) along with what was presented in **Section 2.4**, we consider that HE is an adequate and highly feasible method do assess Visual Analytics Software.

However, according to Scholtz (as cited in Zuk et al., 2006), while the consensus is that heuristics are useful, there is considerable work that must be done before an agreed-upon set of heuristics exists for VA (Scholtz, 2011). Thus, Scholtz provides a possible process on how to develop guidelines and consequently heuristics for VA. In addition, she highlights that when creating guidelines, we should inform their reference, so that these new guidelines have strength of evidence. The process to develop new guidelines by Scholtz should include:

- Accumulating existing guidelines, from HCI, information visualization, automation, etc. that seem relevant to visual analytics systems.
- Providing a reference for each of these guidelines.
- Updating this list, adding new guidelines as they become available and updating strength of evidence (references) for existing guidelines.

3.4.1 How to structure a Heuristic Evaluation for Visual Analytics?

Searching for an adequate form to structure our heuristic evaluation for VA, we found in Scholtz (2006) the proposal of evaluating VA systems beyond usability. Scholtz suggested in this work a set of possible measures and hypothesis which support the rationale for the development of VA environments. Moreover, these measures and hypotheses were based on famous set of heuristics and guidelines related to usability, interaction, and visualization. She proposed the evaluation of VA software in 5 areas: **utility, situation awareness, collaboration interaction, creativity, and collaboration**. We give next a brief explanation of such areas.

Utility. One of the most important measures of visual analytic environments is the utility of the environment from the user perspective. The environment should allow the user to spend more time on task and less time on the tool or environment being used (Scholtz, 2006). Additionally, to effectively transfer new software into a working environment, it is necessary to ensure that the software has utility for the end-users and that the software can be incorporated into the end user's infrastructure and work practices (Scholtz et. al, 2014).

Situation Awareness (SA). Information analysts seek information used in sensemaking (Endert, 2014). Sensemaking is an understanding of a given situation at a given period of time. Thus, one way to evaluate visualizations is to assess the user's situation awareness as gleaned from the visualization (Scholtz, 2006). To evaluate software, Endsley defines three levels of SA: perception, comprehension, and projection (Endsley, 2000). Perception is achieved if operators can perceive in the user interface the information that is needed to do their job. For comprehension, not only must the information be perceived, but also be combined with other information and interpreted correctly. Projection or the ability to predict what will happen next based on the current situation.

Collaboration. The ability to share and discuss data at a data level while using different views is a necessary feature of visual analytic environments that are collaborative. As systems become more intelligent and act more like human collaborators, analysts will want to know what the system is doing and why the system is making recommendations. Metrics developed for evaluating collaboration should include the typical who, what, when, and where (Scholtz, 2006). VA is much about sensemaking, which is supported by interactive visualization, which in turn should support social interaction (Heer & Agrawala, 2008). Social interaction in the sense of promoting collaboration for the analytic process.

Interaction. VA environments do not rely solely on static displays (Scholtz, 2006). User interaction in visual analytic systems is critical to enabling visual data exploration (Endert, 2016). It lets users test assertions, assumptions, and hypotheses about the information, given their prior knowledge about the world (Endert, 2014). User interaction has customarily been a strategy designed to place a "human in the loop" by augmenting parameters in the system to explore data and gain insights (Endert et. al, 2015). Intuitive and efficient user interactions are a fundamental component which has to be efficiently supported by any Visual Analytics system (Kerren & Schreiber, 2012).

Creativity. Creativity is not a term usually associated with analysis. VA tools should enhance the personal experience of the user(s), improve the products or outcomes, and improve the processes used in creating the product or producing the outcome. It is closely related to satisfaction with the solutions (Scholtz, 2006).

On top of the 5 evaluation areas proposed by (Scholtz, 2006), we also considered important to assess usability, once it was mentioned by her as indispensable.

Usability. It is a critical quality attribute in the information society (Pribeanu, 2017). It is the cornerstone of user-centered design and formative evaluation. It is not only of paramount importance for product engineering but also a powerful tool for researchers. It provides feedback on problems encountered by users and steers designers toward better designs at all three evaluation levels (Plaisant et al., 2008a). Usability is of so high importance that even taken alone is critical to user acceptance (Tarrell et al., 2014). Apart from usefulness, a computer system should be usable, which means to enable a user to accomplish the goals with effectiveness, efficiency, and satisfaction (ISO, 2018).

3.5 Heuristics and Guidelines and for Visual Analytics

First, we would like to clarify that some related works in this section do not belong to the VA field specifically. However, they are useful and referred here because they belong to any of the fields which compound the VA approach.

In the literature, heuristic and guideline are sometimes used interchangeably (Tarrell et al., 2014). Thus, to avoid confusion and ambiguity between them, we make a distinction in their meanings in this work. In our context, heuristic refers to a broad but somewhat summarized concept, while guideline refers to guidance related to the heuristic. That is, guidelines will play the role of longer or alternative versions of the heuristics to provide further information or clarifications about the heuristic. We give an illustration of that in **Figure 9**.

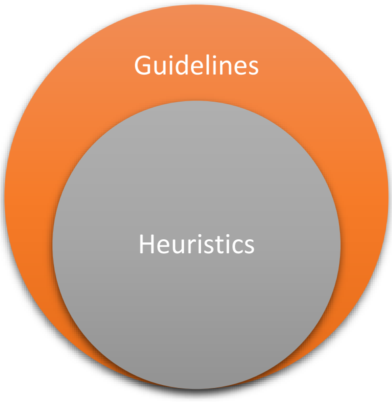

Figure 9 - Guidelines, Heuristics and Metrics in the context of our Evaluation Methodology.

3.5.1 The value of Heuristic Evaluation for VA

As the methodology proposed in this work has the goal of using HE to check whether a software belongs to the VA field according to its design, we have looked for scientific evidence about the value of it for that purpose.

In Pribeanu (2017), the author revised a set of usability heuristics for interactive systems (VA software is interactive) and concluded that they represent valuable design knowledge that could be used to create a user-centered attitude, to incorporate usability into a product, to train novice evaluators, to structure usability guidelines, to explain and document usability problems, and to analyse the ergonomic quality of an application.

In Santos et al. (2018), the authors performed an empirical study of HE for VS (Visualization is part of VA) and their results suggest that using some heuristics may have elicited potential problems that none of the users noticed while using the application. They also reported that those users encountered unpredicted usability issues. Furthermore, it was noticed that a positive effect of using HE is making people more aware of existing principles and guidelines.

We could also find out that some heuristics are so useful and well accepted by the Visualization and VA fields that they were already mentioned as good candidate for standard according to the paper below.

In Tarrell et al. (2014), the authors proposed a framework-based approach to evaluate VS (Visualization is part of VA). Their evaluation framework relies on the heuristics for interface design proposed by Nielsen & Molich (1990). They also highlighted that the afore-mentioned heuristics were so well accepted that they are shown at Services (2020), which is the leading resource for user experience (UX) best practices and guidelines, serving practitioners and students in the government and private sectors of the US. Also, they were used by the National Institute of Standards and Technology (NIST) of the US for use with electronic health records (Lowry et al., 2015).

3.5.2 Heuristics for VA

The choice of which heuristics to use as the basis for an evaluation is no trivial task. It is an open question even in the relatively mature HCI field (Tarrell et al., 2014). This, therefore, would not be different in VA, which newer and strongly tied to the HCI field.

Our work relies primarily on the metrics (indirectly heuristics) suggested in Adagha et al., (2017). Adagha and colleagues' work became a good candidate for base to our work mainly because it synthesizes and presents empirical findings in VADS literature from 2006 to 2012. That is, from the official beginning of the VA field up to 2012. Moreover, it also structures their metrics considering the 5 evaluations areas which were explained in **Section 3.4** and suggested by Scholtz (2006).

The authors performed a systematic review of 470 VA papers. As a result, they provided an overview of application areas, their attributes, and design implications for research and product development of visual analytics decision support (VADS) software.

All of that based predominantly on the evaluation metrics for VA proposed in the research works by Scholtz (2006) and (Wang et al., 2011b), which utilized sets of well-accepted heuristics to be made up. For the convenience of the reader, we present in **Table 3** the metrics proposed in Adagha et al. (2017).

Table 3 - Metrics to evaluate VADS software proposed by (Adagha et al., 2017).

Situation awareness	Collaboration	Interaction
• Can track changes in information • Can provide environment for contextual analysis • Can support future scenario projections • Combination of all	• Ability to share evidence • Can support intuitive communication • Can allow multiple, coordinated views • Can track information flows • Combination of all	• Suitability for the task • Controllability • Self-descriptiveness • Support customization of information • Enable access to information • Combination of all
Creativity	**Utility**	**User-oriented design**
• Support individual tasks • Effective in searching analytical results • Ability to show high quality of analytic • solutions User satisfaction with solutions • Combination of all	• Perceived ease of use • Compatible with the context of use • Perceived usefulness • Enhances effectiveness on the task • Reduction in time • Combination of all	• Analysis of user and context of use • Active involvement of intended users • Iterative design • Evaluation with intended users • Multidisciplinary design input • Use of real-world data • Combination of all

As being Heuristic Evaluation the assessment choice, one needs to decide between lower and higher-level heuristics. The lower the level of the heuristic is, the more specific and less flexible and applicable to other software it is. On the other hand, the higher level of the heuristic is, the more general it becomes and gives naturally more space to misinterpretations by the evaluators.

In Wall et al. (2019), the authors refer that Amar & Stasko (2005) identify heuristics designed to cover the known “gap” in visual analytics processes. However, these heuristics are fairly high level and provide limited guidance on improving specific visual or interactive aspects of a visualization tool. Conversely, Zuk et al. (2006) suggest a set of ten “Cognitive and Perceptual Heuristics” for designing visualizations. But their high specificity in wording leads to less flexibility in interpretation from one visualization to another. However, as Tarrell et al. (2014) point out, by broadly wording such heuristics, they may be misinterpreted by different evaluators.

3.5.3 Guidelines for VA

Guidelines can function as an accelerator for the design process and a replacement for observation of actual users (Scholtz et al., 2014). Even though they do not replace actual user testing, they provide guidance allowing early designs and prototypes to be considered good enough for users to accomplish some tasks and provide more in-depth feedback (Scholtz, 2011). Thus, they play an invaluable role for designers and consequently for evaluation models based on them. Furthermore, they can be turned into heuristics to be used in evaluation, as we explain in the next subsection.

According to Scholtz et al. (2014), we had no agreed-upon of guidelines for the visualizations and interactions of VA systems up to 2014. And the situation seems to remain the same, as we did not find any reference on such agreed-upon in the literature.

3.5.4 Other essential considerations

This subsection addresses other important considerations one should observe when dealing with heuristics and guidelines for the creation of evaluation models.

Different individuals might interpret some heuristics in different ways. For this reason, in Wall et al., (2019), the authors recommended that experts evaluators should be asked to explain what they understood about each heuristic before performing a HE. These authors proposed a value-driven HE for VS and addressed this ambiguity matter by rephrasing the heuristics to diminish the ambiguity of different understandings and increase the HE accuracy.

Evaluators with varying backgrounds and domain knowledge may affect the HE. In Väätäjä et al. (2016), the authors tested the top ten InfoVis heuristics in a HE. They reported that the lack of domain knowledge made the evaluators somewhat uneasy with their capability to carry out the proposed HE in-depth. Thus, they concluded that they may have affected the evaluation findings. As a result, they recommended that training the evaluators for a good understanding of the data and the information system being analysed would help increase the capabilities of the HE.

Existing heuristics and guidelines are an opportunity/source to create evaluation models. In Tarrell et al. (2014), the authors argue that existing heuristics and guidelines for Visualization systems represent latent knowledge of a wide variety of visualization and evaluation experts and can be used by the community as training, design, and evaluation tool. Some of the authors are well-known names in the InfoVis field and suggested the creation of a system for crowdsourcing to define Visualization-specific heuristics aiming to collect contributions from other fellow researchers. They believe so much in the utility of the heuristics and they represent a community-wide snapshot of potential knowledge, which could function essentially as a 'checklist' for designers and a source for HE models.

There are several ways to create new heuristics (based on existing heuristics, literature reviews, usability problems, and guidelines). Authors do it according to their interest and need or sometimes combining more than one technique. In other words, most works do not adopt a standard methodology for heuristic creation (Oliveira & Silva, 2017). A few studies apply a methodology to define, validate and refine the set of heuristics proposed (Quinones & Rusu, 2017). For example, one can transform guidelines into heuristics as in Jaferian et al. (2011), derive heuristics from sets of problems (e.g., (Nielsen & Mack, 1994; Papaloukas et. al, 2009)), and use a specific methodology relying on literature exploration (e.g., (Quinones et al., 2014; Jiménez et al., 2012)). Also, select among existing heuristics coupling it with rules based on theory and practice (e.g., (Forsell, 2010; Mankoff et al., 2003)); or create them from in-deep understanding about a determined kind of system, user needs and indispensable design aspects (Somervell & McCrickard, 2005; Väätäjä et al., 2016). In Tory & Moller (2005), the authors, in a highly cited publication, noted that published heuristics are a good starting point to evaluate Visualizations and recommended using visualization guidelines to develop heuristics.

A minimal set of heuristics very broadly phrased may compromise the evaluation results. While the use of a minimal set of heuristics may seem ideal for its operational employment, focusing on only some minimal number of heuristics almost ensures that they will each be very broadly phrased, perhaps leading to misinterpretation or inconsistent application of them, particularly by evaluators with little experience (Tarrell et al., 2014).

3.6 What is Innovative in the Proposed Methodology?

This subsection aims at showing what is innovative in the evaluation methodology proposed in this work.

Our research revealed that several authors have developed valuable HE studies for VA. However, the one more in line with our proposal is Adagha et al. (2017). In Adagha et al., (2017), it was proposed a product design assessment for VADS based on metrics from well-accepted heuristics. Their work was most influenced by the evaluation frameworks by (Scholtz, 2006) and Wang et al. (2011b).

We have regarded the study presented in Adagha et al. (2017) as so coherent that we have selected it as one of the main studies to rely on and develop our methodology from. Firstly, because the authors performed a literature review on 470 papers of VA publications and applications. Secondly, as the researchers implemented assessment guidelines recommended by Scholtz (2006), who is one of the most well-known and cited researchers on VA to date. Coupled with that, Adagha and colleagues' study relies on the framework for designing and evaluating VADS proposed by Wang et al. (2011b) (some information about it will be given later). Thirdly, their heuristics were developed, refined, and validated in conjunction with more two researchers, which means research efforts of 5 people.

On the other hand, their research reviewed papers from 2006 to 2012 and was published 5 years ago (2015). In other words, it is already not so up-to-date and can be improved considering newer publications from 2013 onwards. Also, they validated their metrics against research VADS systems and not in a real-world environment. Besides, it is not for general VA, rather for VADS specifically.

As for the work of (Wang et al. (2011b), it presented a two-stage framework for informing the design of a VADS systems, which was the outcome of 3 years of iterative design efforts. This paper is very comprehensive and has an educational purpose. Also, it gives well-founded guidance on finding proper criteria to assess VADS.

According to Tarrell et al. (2014), our work can be classified as a framework-based approach because it focuses on selecting and organizing existing heuristics into definite categories. As to the hierarchical format of our proposed evaluation methodology, it is most influenced by the value-driven assessment framework for InfoVis presented by (Wall et al., 2019), which is highly intuitive and was published recently (2019).

In short, our assessment methodology innovates in evaluating in more detail the design of VA systems not considering only on the work by Adagha et al. (2017) but also a variety of more recent literature, which includes papers on evaluation approaches, guidelines and heuristics for VA and its related fields. On top of that, our work is validated using commercial VA software which is already trusted by the growing and competitive Data Analytics market. That is, systems which have already overcome the key issue of product validity. Differently from the study by Adagha et al. (2017), which validated their framework against software that is not yet in the industry. All this envisioning an intuitive tool able to indicate how much a determined a system aligns with the desirable functionalities and characteristics in VA systems considering a novel and more updated criteria set.

4 PROPOSED ASSESSMENT METHODOLOGY

This chapter presents the effort to propose an assessment method based on a novel set of heuristics created and tailored to be as holistic as possible on the evaluation of Visual Analytics Software. The evaluator will have access to a survey with heuristics and guidelines, which should be rated according to a 7-point Likert scale. The resulting survey data will be analysed and turned into a final Visual Analytics Score. In other words, it will be a value-driven assessment. Our proposed value-based methodology includes heuristics and guidelines which are realized in a full methodology.

The evaluation areas proposed in Scholtz (2006) seemed initially to be the best areas to divide our methodology into. However, after classifying all the adopted heuristics and guidelines according to the Scholtz's areas, we found out that several of them belong to more than one of Scholtz's areas at the same time. We illustrated this in **Figure 10**, where each circle represents a group of heuristics and guidelines. Thus, aiming at proposing a more intuitive and objective evaluation framework, we created the new evaluation areas presented in **Section 4.1**.

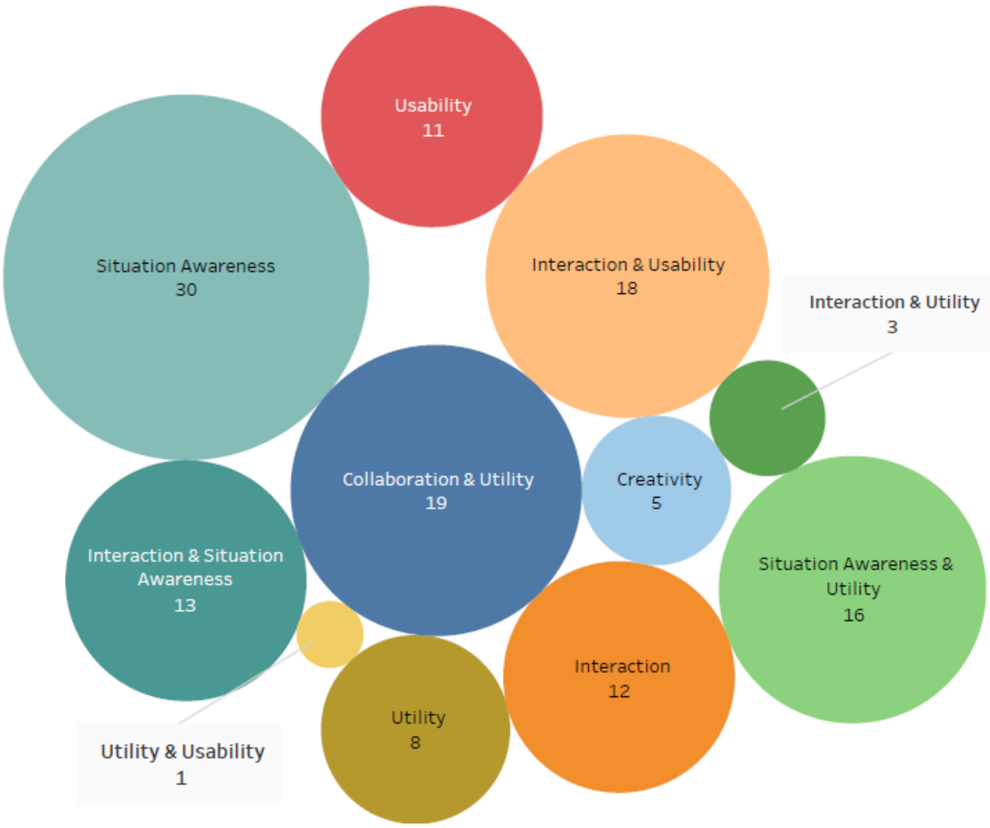

Figure 10 - Number of adopted Heuristics and Guidelines per Scholtz's areas.

Nonetheless, all the referred classification work was useful to illustrate a potential trend regarding the evolution of heuristics and guidelines for general VA software according to Scholtz's areas (see **Figure 11**). The data for the charts in **Figure 10 and Figure 11** comes from **Table 30** of **Appendix E**. Such table shows that all the adopted heuristics could be classified according to the Scholtz's areas, confirming so that our evaluation framework is in line with the proposal by Scholtz, which is to go beyond usability evaluation.

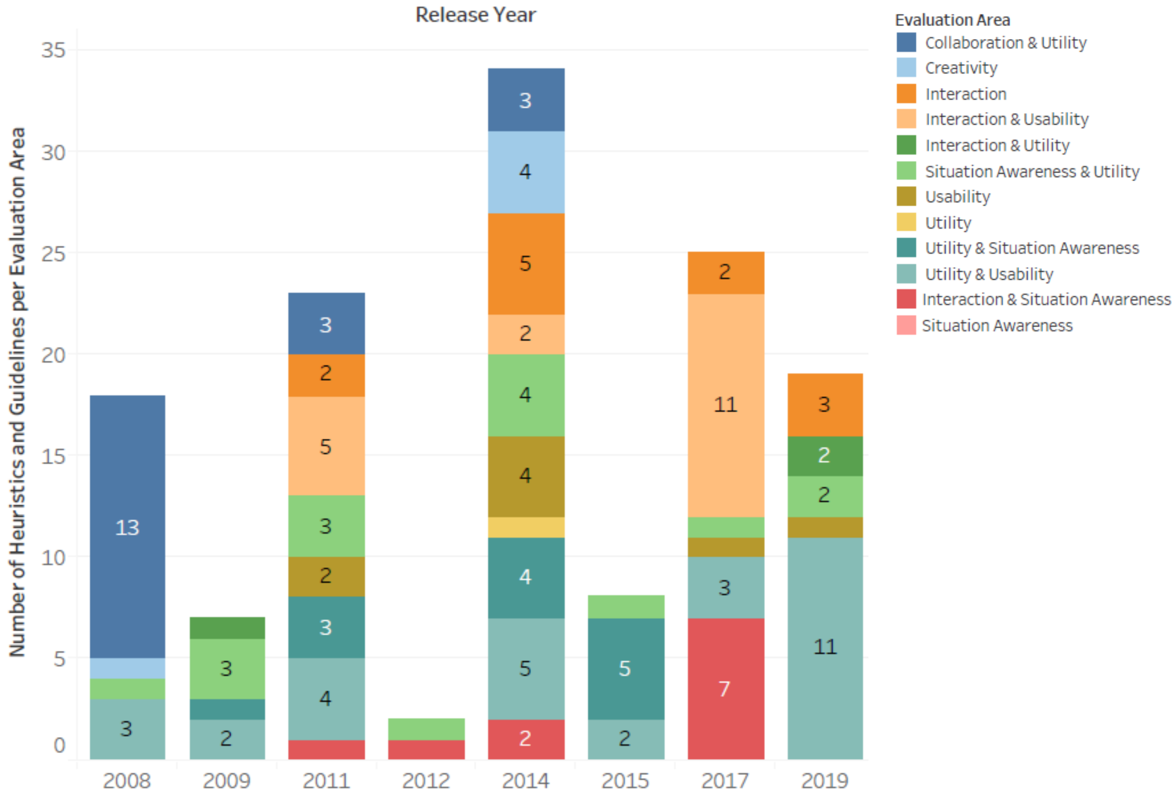

Figure 11 - Timeline on the release of Heuristics and Guidelines for general VA software.

Figure 11 shows a timeline regarding the heuristics and guidelines for general VA software adopted in this work. The values inside the colourful columns represent the number of heuristics and guidelines per evaluation area.

4.1 Visual Analytics Score (VAs)

As our goal in this work is to propose an actionable assessment method with prescriptive capability to evaluate VA Software considering seven areas of high relevance for VA systems. We thought of a value equation below, which will help us calculate a Visual Analytics Score. The maximum score will be 7 (see **Subsection 4.3.3** for instruction regarding scores' aggregation) is given by **Equation 1**:

$$\mathbf{VAs} = \mathbf{AP} + \mathbf{I} + \mathbf{VQ} + \mathbf{UF} + \mathbf{EH} + \mathbf{A} + \mathbf{S} \tag{1}$$

Equation 1 – Visual Analytics Score.

Where the equation components are:

- VAs (Visual Analytics Score) – Final score of an evaluated software considering the 7 evaluations areas proposed in this methodology.
- AP (Analytic Process) – The VA software's ability to deal with automation, data management, data relationships, data characteristics, analytics process flow, and successful meaningful visual schemes that direct to key data characteristics and insight discovery.
- I (Interactivity) - The VA software's ability to provide usability interactions that typically interactive systems are supposed to feature whether being a VA software or not. For instance, capabilities for scrolling, zooming, filtering, interaction menus, and backtracking of actions.
- VQ (Visualization Quality) – The VA software's ability to deliver good-quality visualizations according to the recommended standards to visualizations systems regardless of being VA software. For example, it refers to colour, space, density, animations, and complexity in visualizations.
- UF (User-friendliness) – The VA software's ability to be intuitive and help the user benefit from all its functionalities.
- EH (Error-handling) – The VA software's ability to prevent, inform and correct system errors.
- Satisfaction (S) – The VA software's ability to deliver a solution which provokes satisfaction in the user, including the capability to allow the user to achieve its goals faster.
- Adequacy (A) – The VA software's ability to be appropriate to the kind of analytics it proposes to do.

4.2 Developing Visual Analytics Heuristics and Guidelines

This section explains the development process of the heuristics and guidelines which are used as criteria for the proposed evaluation methodology. **Figure 12** illustrates such process.

Figure 12 - The five-stage process used for developing a new set of heuristics and guidelines.

4.2.1 Methodology to create the new Set of Heuristics and Guidelines

The methodology to develop the new heuristics and guidelines is divided into two parts. Heuristics selection and heuristics reduction. Both parts were made with great attention so that the proposed methodology could be as accurate as possible.

4.2.2 Heuristics and Guidelines Selection

Our initial approach to the challenge of developing the heuristics involved a literature review containing guidelines and heuristics for VA Software and its related fields, which was explained in **Section 3.2**. We were in search of validated studies which could provide heuristics and guidelines to support the evaluation areas proposed in (Scholtz, 2006).

From the collected publications, we selected the 19 in **Table 4** to extract the heuristics and guidelines which would serve as the source for the creation of the novel set of heuristics and guidelines. However, not all the 19 works were adopted, as we explain next.

Likewise, not all the heuristic and guidelines from the adopted publications were utilized, as we explain in **Subsection 4.2.2.2.**

Table 4 - Selected publications with useful guidelines and heuristics for VA

	Paper	Area	Type	Validated?
1.	(Cook et al., 2015)	VA	Guidelines	Yes
2.	(Endert et al., 2015)	VA	Guidelines	No
3.	(Endert, 2014)	VA	Guidelines	No
4.	(Greensmith, et. al, 2009)	VA	Guidelines	Yes
5.	(Heer & Agrawala, 2008)	VA	Guidelines	No
6.	(Scholtz, 2011)	VA	Guidelines	No
7.	(Kang & Stasko, 2012)	VAIA	Guidelines	No
8.	(Kang et al., 2009)	VAIA	Guidelines	No
9.	(Kang et al., 2011)	VAIA	Guidelines	No
10.	(Wang et al., 2011a)	VADS	Guidelines	Yes
11.	(Tarrell et al., 2014)	Visualization and VA	Guidelines/Heuristics	No
12.	(Wall et al., 2019)	Visualization	Guidelines/Heuristics	Yes
13.	(Forsell & Johansson, 2010)	Visualization	Heuristics	Yes
14.	(Oliveira & Silva, 2017)	Visualization	Heuristics	No
15.	(Santos et al., 2018)	Visualization	Heuristics	Yes
16.	(Väätäjä et al., 2016)	Visualization	Heuristics	Yes
17.	(Pribeanu, 2017)	Interaction Systems	Heuristics	No
18.	(Pacheco & Souza-Concilio, 2014)	Interaction Systems	Heuristics	No
19.	(Adagha et al., 2017)	VADS	Metrics*	Yes

*The study calls it “metrics”, however those metrics work actually as higher level heuristics because they do not suggest any quantitative measurement of aspects regarding the evaluation areas by Scholtz (2006).

In **Table 4,** the column "Validated?" refers to the fact whether the guidelines and heuristics were tested in some way. In other words, we had the attention to select papers from where we could extract feasible criteria for our methodology.

Considering that, we would like to justify why publications with "no" in **Table 4** were adopted.

As for **publications 2 and 3**, in Endert (2014) and Endert et al. (2015) respectively, they were considered appropriate for our methodology for being very correlated and by Endert, who introduced the Semantic Interaction approach for VA. The work from 2015 seems to be an extension of the one in 2014. In Endert et al. (2015), the authors proposed guidelines that are an outcome of the 2014 PNNL Workshop about Semantic Interaction for VA Systems. Also, the result of claims and discussions supported by literature during the workshop.

As to **publication 5** by Heer & Agrawala (2008), even though without validation, it is a work of relevance for the VA community considering its number of citations (25 within the IEEE Xplore Digital Library) and comprehensive approach to collaborative VA. In addition, it is very well supported by literature and seems to be so far the only publication that addresses so deeply the collaboration capabilities that a VA software is supposed to offer.

As for **publication 6** by Scholtz (2011), it contains guidelines based on the evaluations of the VA research systems submitted to the 2009 Visual Analytics Science and Technology (VAST) Symposium Challenge. Also, those guidelines were augmented with the heuristics from Forsell & Johansson (2010), which were validated by the study by Väätäjä et al. (2016).

Concerning **publications 7 and 8**, by Kang & Stasko (2012) and Kang et al. (2009) respectively, we have guidelines for Sensemaking derived from findings regarding the evaluation of a VA System for Investigative analyses. The publications address a system which was developed to give higher support to Sensemaking, which is arguably the most challenging part of any data analysis process (Fekete et al., 2019). This made this work valuable for the extraction of guidelines for Sensemaking which could be applied to VA systems in general. Furthermore, the three works are very related to each other. By the way, the studies from 2009 and 2011 suggest the same guidelines to evaluate VA systems for Investigative Analysis.

Concerning **publication 14** by Oliveira & Silva (2017), it was adopted without validation because it merges well-accepted heuristics, namely the heuristics by Shneiderman (1996), Amar & Stasko (2004), Freitas et al. (2002), Scapin & Bastien (1997), and Nielsen & Mack (1994).

With regard to **publication 11** by Tarrell et al. (2014), its heuristics were not validated, but it proposes heuristics and guidelines based on the heuristics by Zuk et al. (2006) in conjunction with the works by Freitas et al. (2002) and Patterson et al. (2014). Those works are widely accepted criteria to evaluate visualizations. Moreover, Scholtz is also authoring it.

Finally, we would like to clarify the presence of the interactive systems study, which is **publication 17** from **Table 4**. The main idea of our work was to evaluate VA software beyond usability evaluation. However, without discarding the usability criteria, which is invaluable. Thus, we selected the study by Pribeanu (2017), which is a revised set of usability heuristics for the evaluation of interactive systems. It is the result of the integration between the ergonomic criteria proposed by Scapin & Bastien, (1997) and the usability criteria of Nielsen (1995a). Both criteria sets are well-known sources of design knowledge that have been already validated and widely used for more than two decades (Pribeanu, 2017). Even though other of our adopted studies (e.g. 14 by Oliveira & Silva (2017) and 18 by Tarrell et al. (2014)) considered Nielsen's widely accepted usability heuristics, the use of Pribeanu's work would give us further guarantee that our framework would not be overlooking any indispensable usability heuristic.

Following this, we would like to present some clarifications regarding the "validated" publications.

As to **publication 1** by Cook et al. (2015), the authors showed the feasibility of their guidelines through a prototype of a spatial workspace which supports the analytic process via task recommendations. In **study 5** by Greensmith et al. (2009), the authors assessed their guidelines through the evaluation of several visualization examples.

In regard to **publication 19** by Adagha et al. (2017), the heuristics proposed were tested by external researchers who applied them to several VA research systems. As for study 12 by Wang et al., (2011a), its guidelines were used to design a system which was validated with positive feedback within an organizational environment. Concerning **publication 17** by Wall et al. (2019), its heuristics and guidelines were tested against 3 visualizations by 15 evaluators.

In short, we used for the extraction of heuristics and guidelines publications 1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 14, 17 and 19.

4.2.2.1 Non-adopted Studies for Extraction of Heuristics and Guidelines

There were the works from **Table 4** which were not adopted to extract heuristics and guidelines. Thus, **publication 9** by Kang et al. (2011) was not adopted, once it suggests the same guidelines we found in publication 9 by Kang et al. (2009). Moreover, we did not adopt the heuristics from **publication 13** by Forsell & Johansson (2010) because it is included in the work of **publication 7** by Scholtz (2011). Also, the ones from **publication 16** by Väätäjä et al. (2016), as it is a study which validates the good coverage of the heuristics in **publication 13** by Forsell & Johansson (2010). Moreover, we did not use **publication 15** by Santos et al. (2018) because the authors compare the performance of the heuristics sets proposed in Forsell & Johansson (2010), Nielsen (1993), and Zuk et al. (2006), instead of proposing new heuristics. Lastly, **publication 18** by Pacheco & Souza-Concilio (2014) was not used because their heuristics neither validated nor based in other heuristics which were validated, as it was the case of the study by Pribeanu (2017).

To sum up, studies 9, 13, 15, 16 and 18 were not adopted for the extraction of heuristics and guidelines.

4.2.2.2 Non-adopted Heuristics and Guidelines

From the 19 studies from **Table 4**, 5 were not used, as justified in the previous section. As a result, we had 14 publications from **Table 4** to extract heuristics and guidelines. However, not all heuristics and guidelines from these 14 works were used in the proposed evaluation framework. **Table 5** shows the non-adopted Heuristics and Guidelines. For the convenience of the reader, we decided to maintain the publication numbers from **Table 4.** Next, we have the justifications for the referred non-adoptions.

Table 5 - Non-adopted Heuristics and Guidelines.

	Publication	Non-adopted Heuristics and Guidelines
3.	(Endert, 2014)	• A visual "near = similar" metaphor supports analysts' spatial cognition and is generated by statistical models and similarity metrics. • Interpret and map the semantic interactions to the model' underlying parameters, by updating weights and adding information. • Models should learn incrementally by taking into account interaction during the entire analytic process, supporting analysts' process of incremental formalism.
5.	(Heer & Agrawala, 2008)	• Allows appropriate within-group diversity. • Support nuanced pointing through selection techniques and visual effects.
7.	(Kang & Stasko, 2012)	• Supplement automatic entity identification. • Allow flexible data (document) management.
8.	(Kang et al., 2009)	• Facilitate further exploration.
10.	(Wang et al., 2011a)	• Incorporate elements from the organizational knowledge base. • Manually update knowledge base systems.
11.	(Tarrell et al., 2014)	• Encourage implicit learning - develop training regimes for implicitly learning about statistical regularities within a visualization.
17.	(Pribeanu, 2017)	• Provide a clear structure of the application.
19.	(Adagha et al., 2017)	• Support individual tasks. • Multidisciplinary design input. • Evaluation with intended users. • Iterative design. • Active involvement of intended users. • Analysis of user and context of use.

For being implicit in the guidelines from **publication 2** by Endert et al. (2015), the guidelines from **publication 3** by Endert (2014) "a visual "near = similar" metaphor supports analysts' spatial cognition and is generated by statistical models and similarity metrics", "interpret and map the semantic interactions to the model' underlying parameters, by updating weights and adding information", and "models should learn incrementally by taking into account interaction during the entire analytic process, supporting analysts' process of incremental formalism", were not utilized.

We have not also adopted heuristics and guidelines with unclear meanings. Thus, we have not utilized "support nuanced pointing through selection techniques and visual effects from **publication 5** by Heer & Agrawala (2008); "facilitate further exploration" from **publication 8** by Kang et al., (2009). Furthermore, "encourage implicit learning - develop training regimes for implicitly learning about statistical regularities within a visualization" from **publication 11** by Tarrell et al. (2014). Lastly, "support individual tasks" from **publication 19** by Adagha et al. (2017).

There were also heuristics discarded for not being applicable to general VA software. It was the case of "supplement automatic entity identification" and "allow flexible document management" from **publication 7** by (Kang & Stasko, 2012) because they refer specifically to systems which deal with documents. Also, "incorporate elements from organizational knowledge base" and "manually update knowledge base systems" in **publication 10 by** (Wang et al., 2011a), once they are stated as optional in the study and specific to VA systems organizational environments.

Moreover, for being specific to the evaluation of web pages, we did not use "provide a clear structure of the application" from **publication 17** by Pribeanu (2017).

Finally, we did not keep the user design guidelines proposed by Adagha et al. (2017) in **publication 19**, namely "multidisciplinary design input", "evaluation with intended users", "iterative design", "active involvement of intended users", and "analysis of user and context of use", as they do not belong to the evaluation areas proposed in Scholtz, (2006).

4.2.3 Heuristics and Guidelines Reduction

We followed the hypotheses that the amount of heuristics should not be large, in order to demand less cognitive effort from the evaluator when applying HE (Oliveira & Silva, 2017). We also had the intention to create new heuristics set which should be small enough to be efficient to apply and general enough to cover and explain important aspects that a general VA software is supposed to have.

As explained in **Subsection 3.5.4,** there are several ways to create new heuristics (based on existing heuristics, literature reviews, usability problems, and guidelines). With that in mind, we reduced **136** heuristics and guidelines in **Appendix E** into **59** with attention to redundancy and conflict among them. Besides, looking for the best organization of the new set, dividing it into categories to ease the understand of our evaluation model. As a reference, we could follow the examples of merge and reduction of pre-existing heuristics sets into new ones made in Oliveira & Silva (2017) and Pribeanu (2017).

Following this, to ease the merge and reduction of the **136** heuristics and guidelines into a smaller set, first, we classified them one by one according to the areas proposed by Scholtz (2006). Thus, we could have an initial sorting of the heuristic and guidelines in a way that the ones which had something in common would be closer to each other. This was helpful for the comparison phase of our adopted grouping method illustrated in **Figure 13**.

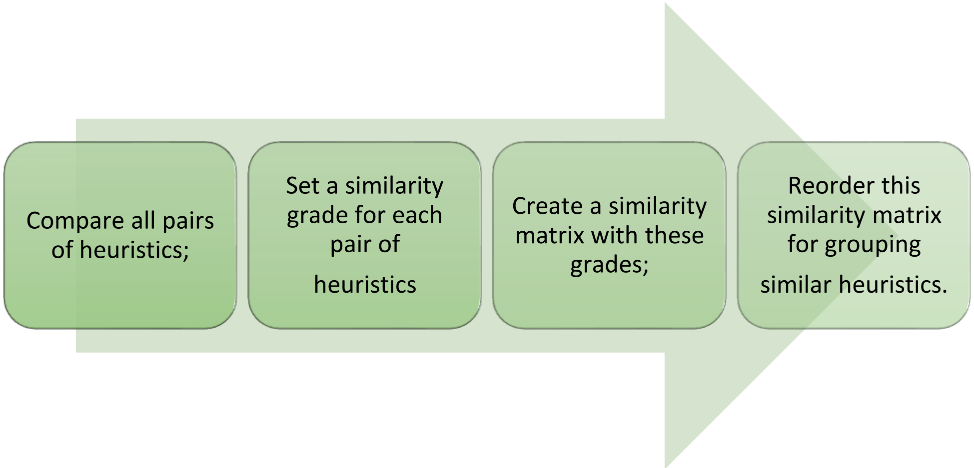

Figure 13 - Heuristics and guidelines grouping method in Oliveira & Silva (2017).

The comparison phase of the grouping method showed in **Figure 13** was performed in an Excel sheet where the selected heuristics and guidelines were placed in both rows and columns for comparison against each other. Intending to be as precise as possible, the first step (compare all pairs of heuristics) was overcome only after comparing all the **136** heuristics and guidelines against each other. That is, we had to perform **9248** manual comparisons.

The goal of such comparisons was to build a similarity matrix to be used as a dataset for grouping and clustering the heuristics with help of the programming language Python. We chose then for similarity grade one of the following four possible values: 0 = not similar; 0.33 = partially similar; 0.66 = quasi-equal; 1 = equal.

After creating the similarity matrix with the heuristics and guidelines, we needed to reorder and cluster them to facilitate identification of similarities and so to proceed with the reduction itself. For that, we used Python to develop heatmaps with clusters. **Figure 14** shows the developed Python code.

```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt


df = pd.read_excel('heatmap_i3.xlsx')
dataset = df.T
sns.clustermap(dataset, method='single', metric='chebyshev', figsize=(10, 10))
plt.savefig('heatmap_i3.png', figsize=(10, 10))
```

Figure 14 - Python code for heatmaps with clustering on Jupyter Notebook.

From **Figure 14**, we would like to clarify the need to set up the parameters “method” and “metric” to be able to visualize better the clusters on our heatmap. For “method”, we made use of “single”, which refers to the Nearest Point Algorithm within the several tested linkage options for clustering that we found in Scipy Community (2020a). For “metric”, the Chebyshev distance was chosen, instead of the Euclidian distance, which is the default distance value for the cluster map function from the Seaborn library. We tested all the possible distances available in Scipy Community (2020b). **Figure 15** shows our resulting heatmap with cluster, which would be not readable regardless of its size inside this document due to the amount and extension of the heuristics and guidelines used for the analysis.

Figure 16 shows a more readable image of the cluster map used as a resource to group heuristics and guidelines. It refers to the final cluster map regarding the Interactivity Evaluation area of the assessment framework proposed in this work. Each number represent a heuristic or guideline from **Table 6**.

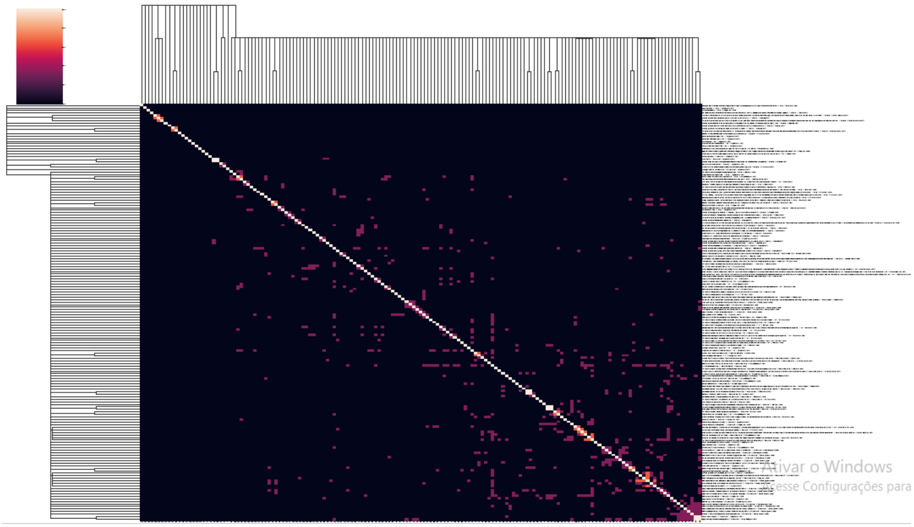

Figure 15 - Resulting Heuristics and Guidelines Heatmap with Clustering developed in Python.

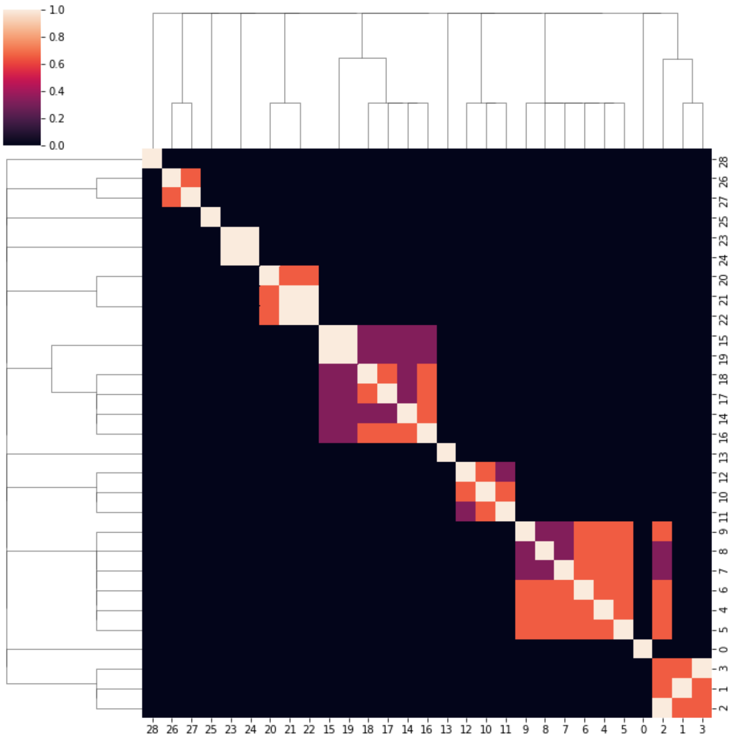

Figure 16 - Heatmap Interactivity Evaluation Area.

Table 6 - Heuristics and guidelines from the cluster map in Figure 16.

1.	Data set reduction (including filtering, clustering, and pruning).
2.	Data Manipulation: provide tools for data manipulation, such as filters and detailed view.
3.	Enable facet filtering for information personalization.
4.	Enable access to information.
5.	Navigation and Querying.
6.	The visualization provides useful interactive capabilities to help investigate the data in multiple ways.
7.	The visualization supports smooth transitions between different levels of detail in viewing the data.
8.	Spatial Organization and Perspective: care the visualization overall layout, as well as provide change of perspective.
9.	Interactive content exploration and filtering.
10.	Support customization of information.
11.	The interface supports using different attributes of the data to reorganize the visualization's appearance.
12.	Multidimensionality: allow users to visualize three or more dimensions simultaneously.
13.	The visualization avoids complex commands and textual queries by providing direct interaction with the data representation.
14.	Orientation and Help.
15.	User Control: enable full system control by the user.
16.	Provide a way to backtrack or undo actions.
17.	Can track changes in information
18.	Provide histories of actions performed on artefacts (representations/visualizations).
19.	Controllability.
20.	Reduction in time.
21.	Flexibility and Efficiency: provide accelerators and customization features.
22.	Flexibility: provide means to customize the interface and select the preferred way to accomplish a goal.
23.	Provide means for analysts to explore visualizations that do not require repetitive interactions on the part of the analyst.
24.	Minimal actions: Minimize the number of actions needed to accomplish a task's goal.
25.	Feedback: Provide appropriate feedback as a response to the user's actions within reasonable time.
26.	Prompting: Guide users towards making specific actions.
27.	Visible Actions: make all possible actions visible.
28.	Grouping/distinction: Provide means to group similar objects and distinguish between different classes of objects.

The work to classify, group and reduce was time-consuming and difficult. Firstly, due to the need to ensure our understanding of all the selected heuristics and guidelines and its related concepts. Secondly, considering the significant number of selected heuristics and guidelines from different areas (most related heuristic works worked on average with a third of our number of heuristics and guidelines). Thirdly, we had a great mix of heuristics (lower and higher level) and guidelines to work on, coming from a diversity of authors, who sometimes used related technical terms interchangeably. Along, we had the attention to compare all the selected 136 heuristics and guidelines one by one against each other in search of similarities to obtain so a more accurate and coherent reduction.

As for the grouping method proposed by Oliveira & Silva (2017), it was useful because it allowed us to have several related heuristics and guidelines closer to each other for comparison. However, in our case, several heuristics and guidelines showed similarity to more than one heuristic and/or guideline at the same time. For example, sometimes a heuristic or guideline was related to a pair of unrelated heuristics and guidelines. As a result, that single heuristic and guideline was allocated in only one of its related clusters. That happened also in Oliveira & Silva (2017). However, considering we had over two times more heuristics guidelines to handle than Oliveira & Silva (2017), our analytical work of grouping and refining became longer and more complicated. In **Figure 16**, we have an example of such kind of situation: **heuristic 2** from **Figure 16, which refers to heuristic 2 from Table 6**, "data manipulation: provide tools for data manipulation, such as filters and detailed view" belongs to more than one cluster at the same time.

4.2.4 The new Set of Heuristics and Guidelines per Evaluation Area

After applying the reduction method described in **Subsection 4.2.3**, the result was a set of **59** heuristics and guidelines which is presented per evaluation area in **Subsections 4.2.4.1 to 4.2.4.7**. The new heuristic set was ordered in a way that allows a more logical and therefore less time-consuming evaluation. We will explain every heuristic through its respective guideline so that the heuristics meanings are clearer, and misinterpretations are minimized when applying them. The heuristics are divided into the seven evaluation areas presented in **Section 4.1** and illustrated in **Figure 17**.

Figure 17 - The VA Evaluation Methodology and its 7 Areas.

4.2.4.1 Analytic Process

Table 7 shows the 32 heuristics along with their respective guidelines for assessment regarding the Analytic Process area shown in **Figure 17.** These heuristics and guidelines come from the original ones presented in **Table 31** of **Appendix F**.

Table 7 - Heuristics Analytic Process Area.

Analytic Process			
	Heuristic	**Guideline**	**Source**
AP1	VAS shows key characteristics of data at a glance.	VAS should feature visualizations that allow identifying key characteristics of data in a quick short look.	(Wall et al., 2019)
AP2	VAS makes data relationships noticeable.	VAS should facilitate answering questions about the data by making relationships in it noticeable. That is, by making visible, for instance, distribution of variables, correlations, and clusters.	(Heer & Agrawala, 2008; Oliveira & Silva, 2017; Wall et al., 2019; Wang et al., 2011a)
AP3	VAS provides a new or better understanding of the data.	VAS should provide a new or better understanding of the data. This through helping identify unexpected, duplicate, missing, or invalid data. Also, dependent, independent, and important dimensions.	(Oliveira & Silva, 2017; Tarrell et al., 2014; Wall et al., 2019)
AP4	VAS helps generate data-driven questions.	VAS should help the user generate data-driven questions from its analytical outcomes.	(Wall et al., 2019)
AP5	VAS suggests relevant information beyond dataset information.	VAS should not only suggest relevant information about the dataset itself and its attributes, but also, for instance, about related views, comments, and data to current points of interest, as well as notification subscriptions for views, artefacts (reports, dashboards, and datasets), people, and groups.	(Heer & Agrawala, 2008)
AP6	VAS features visualization which provides comprehensive data overview with meaningful visual schema.	VAS should feature visualization that provides a big picture/perspective of data through an accessible data overview and meaningful visual schema.	(Wall et al., 2019; Wang et al., 2011a)
AP7	VAS provides coordinated views for linked information.	Visualizations on VAS should be coordinated together in such a way that action performed in one view affects all other views.	(Wang et al., 2011a)
AP8	VAS displays related information nearby.	VAS should show related information in close proximity.	(Scholtz, 2011)
AP9	VAS minimizes distractions for the analyst.	VAS should minimize distractions for the analyst. That is, minimize aesthetics or interactions that take the user outside of the frame of the task. Minimizing distractions assists endogenous attention and reduction in time.	(Adagha et al., 2017; Greensmith et al., 2009; Kang et al., 2009; Tarrell et al., 2014)
AP10	VAS provides opportunities for serendipitous discoveries.	VAS should provide opportunities for serendipitous discoveries by displaying information from multiple aspects, as well as related and partially related data points.	(Wall et al., 2019; Wang et al., 2011a)

AP11	VAS allows flexibility in the organization of the visual metaphor.	VAS should allow flexibility in the organization of the visual metaphor (visual space where the user interacts).	(Kang et al., 2009, 2011)
AP12	VAS facilitates finding starting points or clues.	VAS should provide an environment in which the user can capture information to find starting points or clues. That is, it should direct attention to the most critical information.	(Cook et al., 2015; Kang et al., 2009, 2011)
AP13	VAS provides strong retrieval cues for mental models.	VAS should structure information in a way which provides strong retrieval cues for mental models* aiding in reasoning**.	(Tarrell et al., 2014)
AP14	VAS allows share evidence and hypothesis.	VAS should have the ability to share evidence and hypothesis so that users can create hypothesis regarding their analysis, collect them, and share them with other users. Likewise, it should be possible for the collected evidence. That being feasible, for instance through shared, editable representations; in-app collaborative editing; embedding of annotated views in external media (e.g., email, blogs, and reports); or sharing of views across media (e.g., URLs).	(Adagha et al., 2017; Heer & Agrawala, 2008; Wang et al., 2011a)
AP15	VAS supports collection of evidence and annotations in a beneficial organization to sensemaking.	VAS should allow collecting and grouping evidence and annotations, as well as to register the need for more evidence or other future actions, preferably through storytelling, in a beneficial scheme to the sensemaking process.	(Heer & Agrawala, 2008; Kang et al., 2011; Tarrell et al., 2014; Wang et al., 2011a)
AP16	VAS allows registering need for more evidence or other future actions.	VAS should allow registering need for more evidence or other future actions regarding the analytic process.	(Heer & Agrawala, 2008)
AP17	VAS supports sensemaking by recommending relevant information.	VAS should support sensemaking by presenting semantically meaningful recommendations that enrich the current analytic process based on the user's current activity and potential next step.	(Cook et al., 2015)
AP18	VAS displays statistics and measures about data sources, datasets, and/or records.	VAS should support evidence discovery by displaying statistics and measures regarding data sources, datasets, and/or records.	(Kang & Stasko, 2012)
AP19	VAS features a visual display of the analytic process.	VAS should feature a visual display of the process so that there is no need to keep external notes.	(Scholtz, 2011)
AP20	VAS provides an easy-to-interpret environment for contextual analysis with relevant information.	VAS should provide an easy-to-interpret environment for contextual analysis composed by relevant information for the analysis and suggestions about what may have been overlooked.	(Adagha et al., 2017; Cook et al., 2015; Greensmith et al., 2009; Heer & Agrawala, 2008; Wall et al., 2019)
AP21	VAS provides transparent automation regarding the underlying mathematical models and parameters.	VAS should contain automation, which is transparent to the user, shielding users from the complexity of underlying mathematical models and parameters.	(Endert, 2014; Scholtz, 2011)

AP22	VAS captures and understands user interactions.	VAS should be able to capture and understand (the kind of action) user interactions in spatial analytic processes such as searching, highlighting, annotating, and repositioning documents for future automation.	(Endert, 2014; Endert et al., 2015; Wang et al., 2011a)
AP23	VAS makes inferences from user interactions.	VAS should be able to make inferences (deductions) from user interactions. For example, for suggestions of recommendations regarding the analysis.	(Cook et al., 2015; Endert et al., 2015)
AP24	VAS reacts and takes initiative based on inferences from user interactions.	VAS should be able to react and take initiative based on those inferences at three levels: interface, computation, and cognitive.	(Cook et al., 2015)
AP25	VAS provides visual feedback regarding the updated model.	VAS should also provide visual feedback of the updated model and learned parameters within the visual metaphor (visual space where the user interacts).	(Endert, 2014; Oliveira & Silva, 2017)
AP26	VAS features teamwork management.	VAS should feature group creation and teamwork management, including division of labour among participants.	(Heer & Agrawala, 2008)
AP27	VAS features activity indicators per collaborator increasing so engagement.	VAS should provide a history of past contributions, to create activity indicators, as well as to aid reputation and visibility of contributions, so that engagement increases.	(Heer & Agrawala, 2008)
AP28	VAS supports intuitive communication among collaborators.	VAS should support intuitive communication to support discussions on common ground. In other words, it should provide intuitive means to share understanding among collaborators to facilitate consensus and decision making.	(Adagha et al., 2017; Heer & Agrawala, 2008)
AP29	VAS allows tracking the update of collaborative threads.	VAS should allow following collaborative threads' updates regarding the analytic process.	(Adagha et al., 2017; Wang et al., 2011a)
AP30	VAS supports future scenario projections.	VAS should support users in making future scenario projections such as forecasting.	(Adagha et al., 2017)
AP31	VAS integrates multiple information channels.	VAS should provide means to integrate multiple information sources, forming a single unified content collection.	(Wang et al., 2011a)
AP32	VAS increases engagement and attention with game design elements.	VAS could use game design elements to reframe tedious data entry tasks as actions within online games for increasing engagement. For instance, a team-oriented 'scavenger hunt' analysis would allocate more attention.	(Heer & Agrawala, 2008)

*Mental models are internal representations mirroring the structure of the external world.

**Reasoning is the mental process of drawing a conclusion from a set of premises.

4.2.4.2 Visualization Quality

Table 8 shows the 5 heuristics and guidelines for assessment regarding the Visualization Quality area shown in **Figure 17.** These heuristics and guidelines come from the original ones presented in **Table 32** of **Appendix F**.

Table 8 - Heuristics for the Visualization Quality Area.

Visualization Quality			
	Heuristic	**Guideline**	**Source**
VQ1	VAS facilitates perception via Gestalt principles.	VAS should guide and maximize perception via Gestalt principles (proximity, similarity, enclosure, closure, continuity, and connection) in its visualizations.	(Oliveira & Silva, 2017; Tarrell et al., 2014)
VQ2	VAS provides visualizations with meaningful spatial organization.	VAS should care about the visualization overall layout, displaying a meaningful spatial organization of the data.	(Oliveira & Silva, 2017; Tarrell et al., 2014; Wall et al., 2019)
VQ3	VAS avoids dense visualizations by featuring properties for size and distance.	VAS should offer appropriate and easy to interpret representations for properties such as size and distance in visualizations, avoiding so dense visualizations.	(Adagha et al., 2017; Scholtz, 2011)
VQ4	VAS uses animation only to show an effect that moves over time.	VAS should use animations only to show an effect that moves over time. Give analysts control to manipulate the speed of the animation.	(Scholtz, 2011)
VQ5	VAS displays only relevant information and elements in a straightforward fashion.	VAS should avoid misleading and complex representations by displaying only relevant elements to the analytic process in a straightforward fashion.	(Oliveira & Silva, 2017; Wall et al., 2019; Wang et al., 2011a)

4.2.4.3 Interactivity

Table 9 shows the 11 heuristics and guidelines for assessment regarding the Analytic Process area shown in **Figure 17.** These heuristics and guidelines come from the original ones presented in **Table 33** of **Appendix F**.

Table 9 - Heuristics for the Interactivity Area.

Interactivity			
	Heuristic	**Guideline**	**Source**
I1	VAS features self-descriptive interactions.	VAS should feature self-descriptive interactions. That is, it should allow users intuitively understand what they can do with the interaction and how they can do it.	(Oliveira & Silva, 2017; Pribeanu, 2017)
I2	VAS provides tools for data manipulation.	VAS should provide tools to help users in data manipulation. For instance, tools for filtering, clustering, pruning.	(Adagha et al., 2017; Heer & Agrawala, 2008; Oliveira & Silva, 2017; Scholtz, 2011; Tarrell et al., 2014)

I3	VAS provides capabilities for data exploration.	VAS should feature useful interactive capabilities to help investigate data in multiple ways. For example, zooming; navigation and querying (including the selection of objects, viewpoint manipulation, geometric manipulation, and searching).	(Adagha et al., 2017; Oliveira & Silva, 2017; Tarrell et al., 2014; Wall et al., 2019)
I4	VAS allows interactive visualization customization.	VAS should support customization of the visualization. For instance, by using different attributes of the data to reorganize its appearance and supporting several dimensions simultaneously in it.	(Adagha et al., 2017; Oliveira & Silva, 2017; Wall et al., 2019)
15	VAS avoids complex commands and queries in visualizations.	VAS should avoid complex commands and textual queries in visualizations by providing direct interaction with the data representation.	(Pribeanu, 2017; Scholtz, 2011)
I6	VAS provides a way to backtrack and to undo actions.	VAS should provide a way to track changes in information and undo actions (a history with all user actions may be used).	(Oliveira & Silva, 2017; Tarrell et al., 2014; Wang et al., 2011a)
I7	VAS provides alternative ways to perform a task.	VAS should provide alternative ways to perform a task, such as shortcuts for experienced users, to increase the interaction speed and to reduce time.	(Pribeanu, 2017)
I8	VAS provides interaction with minimal need for repetitive actions.	VAS should provide means to explore visualizations and overall system functions avoiding as much as possible repetitive actions on the part of the end-user.	(Scholtz, 2011; Wall et al., 2019)
I9	VAS gives proper feedback to user actions within reasonable time.	VAS should provide appropriate feedback as a response to user actions within reasonable time.	(Pribeanu, 2017)
I10	VAS guides users towards making specific actions.	VAS should guide users on specific actions by showing selectable option, windows titles, system status, data fields with labels and acceptable values and formats.	(Oliveira & Silva, 2017; Pribeanu, 2017)
I11	VAS facilitates the understanding of relationships between the various user interface items	VAS should provide means to understand the relationships among items by grouping similar objects according to formats and graphical features; screen areas; and between different classes of objects.	(Pribeanu, 2017)

4.2.4.4 User-friendliness

Table 10 shows the 05 heuristics and guidelines for assessment regarding the Analytic Process area shown in **Figure 17.** These heuristics and guidelines come from the original ones presented in **Table 34** of **Appendix F**.

Table 10 - Heuristics for the User-friendliness Area

User-friendliness			
	Heuristic	**Guideline**	**Source**
UF1	VAS provides coherent UI elements	VAS should follow similar meaning and design choices in similar contexts. That is, the interface elements should be coherent.	(Adagha et al., 2017; Oliveira & Silva, 2017; Pribeanu, 2017)
UF2	VAS features a UI with familiar signs to the user.	VAS should feature a UI where all signs (codes, names, texts, figures, and icons) in the UI are familiar to the user and have an expected meaning.	(Adagha et al., 2017; Oliveira & Silva, 2017)
UF3	VAS matches user characteristics with the UI characteristics	VAS should feature a UI compatible with the user characteristics, (language, measurement units, calendar, and accessibility capabilities).	(Adagha et al., 2017; Pribeanu, 2017; Wang et al., 2011a)
UF4	VAS provides proper help and documentation to guide the user	VAS should provide help and documentation to guide the user. All actions that the user can realize in the system should be easily identified/visible. Also, should be available easy-to-understand tutorials (especially for the not-easily-identified functionalities).	(Adagha et al., 2017; Kang & Stasko, 2012; Oliveira & Silva, 2017; Pribeanu, 2017)
UF5	VAS makes easily visible all possible actions for the user	VAS should make easily visible all possible actions the user can perform, be intuitively or through help, documentation, and tutorials.	(Adagha et al., 2017; Kang & Stasko, 2012; Oliveira & Silva, 2017; Pribeanu, 2017)

4.2.4.5 Satisfaction

Table 11 shows the 4 heuristics and guidelines for assessment regarding the Analytic Process area shown in **Figure 17.** These heuristics and guidelines come from the original ones presented in **Table 35** of **Appendix F**.

Table 11 - Heuristics for the Satisfaction Area.

Satisfaction			
	Heuristic	**Guideline**	**Source**
S1	VAS is effective in representing high-quality of analytical outcomes.	VAS should show the high quality of analytic outcomes on the visual interfaces. That is the system should be effective in depicting through the visualization the outcomes from the automated data analysis.	(Adagha et al., 2017; Tarrell et al., 2014)
S2	VAS ensures the end-user's subjective assessment is overall positive.	VAS should ensure an overall positive end-user subjective assessment. This involves overall subjective satisfaction about the system, which relates to how pleasant and easy-to-use it is, as well as frustrating experiences, and productivity through it.	(Adagha et al., 2017; Tarrell et al., 2014)
S3	VAS is considered highly useful.	VAS should be considered highly useful. In other words, it refers to whether the system provides the features the user needs.	(Adagha et al., 2017)

S4	VAS minimizes the needed resources to achieve the goal.	VAS should maximize efficiency by minimizing the necessary resources to achieve the goal. It should maximize the speed and minimize the number of steps to achieve an objective.	(Tarrell et al., 2014)

4.2.4.6 Error-handling

Table 12 shows the heuristics and for assessment regarding the Error-handling area shown in **Figure 17.** These heuristics and guidelines come from the original ones presented in **Table 36** of **Appendix F**.

Table 12 - Heuristics for the Error-handling Area.

Error-handling			
	Heuristic	**Guideline**	**Source**
EH1	VAS prevents, diagnoses, and correct errors.	VAS should prevent, diagnose, correct, and recover from errors with clear and informative messages, giving reasons, as well as the means to correct them.	(Oliveira & Silva, 2017; Pribeanu, 2017)

4.2.4.7 Adequacy

Table 13 shows the heuristics and guidelines for assessment regarding the Adequacy area shown in **Figure 17.** These heuristics and guidelines come from the original ones presented in **Table 37** of **Appendix F**.

Table 13 - Heuristics for the Adequacy Area.

Adequacy			
	Heuristic	**Guideline**	**Source**
A1	VAS is adequate for its context of use.	VAS should be compatible with the context for which it was designed. That is should be suitable to facilitate the analytical goals of its context of use.	(Adagha et al., 2017)

4.3 Value-driven Assessment Methodology

Our idea was to develop a sustainable but at the same time flexible evaluation methodology. Thus, we decided on a framework-based approach because it would provide the needed flexibility to present a comprehensive, well-organized, and generalizable structure for our methodology (Tarrell et al., 2014). For a sustainable evaluation methodology, we mean an assessment method capable to adapt and improve in view of better understanding, future changes, and advances in the VA field.

4.3.1 Methodology Structure

We intended to include the new set of heuristics and guidelines into a hierarchical organization which would facilitate the understanding and use of the proposed methodology by both designers and evaluators.

Our proposed framework contains three hierarchical levels, which are distinguished by the colours: orange, blue, and green in **Figure 18**. The orange level contains 7 components, i.e., our 7 proposed evaluation areas. The blue level contains the heuristics for the respective evaluation area in the orange level. As for the green level, it contains the guidelines for the respective heuristics of the blue level. The number of heuristics and guidelines of each evaluation area are shown inside the squares from the blue and green levels. **Figure 18** illustrates the structure of the proposed evaluation methodology.

We had the intention to develop easy-to-rate heuristics. Thus, the blue level contains guidelines which were intended to work as longer or alternative versions of the heuristics in the blue level. In other words, the guidelines are to provide further clarification when the heuristics are being rated. Hence, for each component, we have the same number of guidelines and heuristics (the blue and green levels).

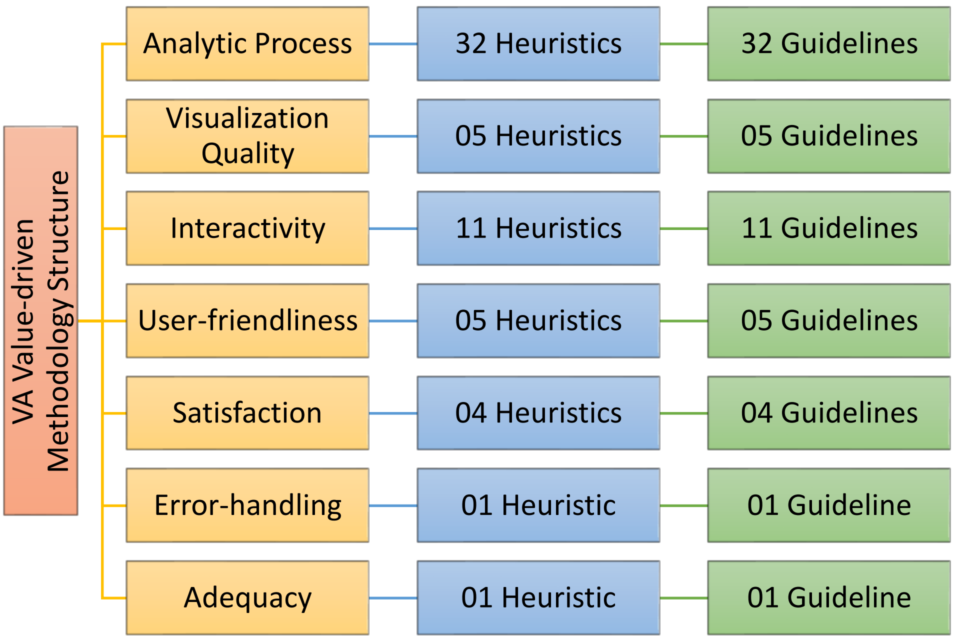

Figure 18 - The framework structure of our assessment methodology.

Considering that the VA field are likely to evolve and change over time, higher-level heuristics become more interesting than lower-level ones, as the higher-level ones are more general and therefore more time-invariant. Thus, we propose mid-level and higher-level heuristics, which give to our methodology less need to change and hence more sustainability.

4.3.2 Implementation through a Survey

This methodology was thought to be administered using a survey (**Appendix C**). Each heuristic should be individually rated along a 7-point scale ranging from 1-strongly disagree to 7-strongly agree, or N/A-not applicable. All the heuristics within the survey will be rated with positive values according to **Table 14**. The highest score (7) means a strong agreement regarding the satisfaction of the heuristic by the software under evaluation. On the other hand, the lowest score (1) indicates the rater strongly disagrees that the program meets the heuristic.

Table 14 - Scores meanings for the 7-point Likert scale of the proposed methodology.

Score	Meaning
1	Strongly Disagree
2	Disagree
3	Somewhat Disagree
4	Neither Agree nor Disagree
5	Somewhat Agree
6	Agree
7	Strongly Agree

The Likert scale dates from 1923 and it is one of the most frequently used evaluation tools in educational and social sciences research. Such scale has typically two versions a 5- and 7-point ones. It was chosen the 7-point one because it provides a wider range of degrees to express opinion, making it naturally more accurate (Joshi et al., 2015).

Besides that, it has the advantage of allowing us to easily obtain quantitative data from the evaluator's qualitative assessments. Moreover, this data can be analysed with relative ease (McLeod, 2019).

The survey should be preferably used by evaluators who are VA domain experts with experience using VA solutions. If not so, the raters should have at least some familiarity with concepts from the VA field. The recommended number of raters for HE is five, especially if the raters are domain experts (Nielsen, 2000). Wall et al. (2019) confirm this number after conducting a power analysis in their study which evaluates visualizations through heuristics. As mentioned in **Section 3.4**, five raters will be enough to discover 75% of the overall detectable problems through usability tests with heuristics. More than that will not increase significantly the problem detection rate (Nielsen, 2000).

4.3.3 Aggregating Scores to calculate the Visual Analytics Score (VAs)

To calculate the Visual Analytics Score, first, all the heuristics should be rated with a score from 1-7 through the survey. We believe evaluation areas should have different weights, totalling the sum of weights 100%. Thus, we propose initially aggregating scores at each level of the hierarchy by making use of a simple average. Then should be calculated a weighted average considering the average means obtained for each evaluation area, as **Equation 2** illustrates. The highest Visual Analytics score will be 7.

$$\mathrm{VAs}=\sum s_i w_i \tag{2}$$

Equation 2 - Aggregation Formula to calculate the Visual Analytics Score.

$\mathrm{VA}s$ stands for Visual Analytics Score. s_i refers to the scores regarding each area under evaluation. Namely, *Sap* (score for analytic process), *Svq* (score for visualization quality), *Si* (score for interactivity), *Suf* (score for user-friendliness), S*s* (score for satisfaction), *She* (score for error-handling), and *Sa* (score for adequacy). w_i refers to weight. That is, the weights attributed to each evaluation area.

Considering the areas which most characterizes and differentiates VA software from other software types, we suggest the weights from **Table 15** for an initial and fairer aggregation approach. Next, we give the justifications these weights.

Table 15 - Weighs per Evaluation Area.

Evaluation Area	Suggested Weight
Analytic Process	40%
Visualization	20%
Interactivity	20%
User-friendliness	5%
Satisfaction	5%
Error-handling	5%
Adequacy	5%

To support the generation of knowledge from data, a VA system is supposed to combine automatic and visual analysis methods with a tight coupling through human interaction. In other words, a VA should combine visualization generated by automatic analysis and human interaction to support data analysis (Keim et al., 2009). Considering this, we suggested the highest weights to the Analytic Process, Visualization, and Interactivity evaluations areas of the proposed methodology.

However, among these three areas, Analytic Process received the highest weight (40%) because all its criteria (heuristics and guidelines), except for criteria **AP26** from **Table 7**, relate to the provision of information which will trigger the Sensemaking loop during the analytic process. The indispensable role information plays in the sensemaking process is described in **Subsection 2.2.1.** Along with that, as explained in **Subsection 2.2.1**, the sensemaking loop is responsible for structuring the whole knowledge discovery process supported through VA (Keim et al., 2008). Thus, as knowledge discovery is the main goal of VA systems, we theorize that the evaluation which considers criteria related to such provision of information should be the most relevant one for Visual Analytics Systems. The only criteria from the Analytic Process Area which does not provide information for knowledge gain from data through the sensemaking loop is **AP26** (VAS features teamwork management), once it is about the division of labour among participants of the same data analysis.

Considering the first VA definition by Thomas & Cook (2005), which is “Visual Analytics is the science of analytical reasoning facilitated by interactive visual interfaces”. One can conclude interactions and visualizations are indispensable to VA, but they work as the tools to facilitate the analytical reasoning. That is, they come in the secondary position of working as support to the data analytical purpose VA systems have. Thus, the Visualization Quality and Interactivity categories received the second highest weights, 20% each.

The User-friendliness evaluation area contains the criteria related to ease-of-use, which is a basic concept that describes how easily a system can be used (IDF, 2019). They are so important and desirable to VA systems as it is to any kind of system which aims higher acceptance, once we naturally tend to avoid software which is hard to be understood and be used. Concerning the Satisfaction category, it has criteria about the capability of VA systems to provoke user satisfaction. Satisfaction has its roots in psychology, and it involves the attitudes and feelings of the users towards something, in our context a software (Bailey & Pearson, 1983). According to Antonopoulou & Kotsilieris (2019), user satisfaction has a significant impact on information systems success. Hence, satisfaction is needed for VA system, as it strives for such success as any information system.

The Error-handling category has criteria regarding the ability of VA systems to deal with software bugs. Software errors can crash a program, cause data loss and limit productivity (Mott, 2020). Thus, the ability to deal with errors is so necessary to VA software as it is to any system which strives for quality. As to the Adequacy area, it aims to measure the overall impression a user has regarding the suitability of a VA system to its context of use, which is also another significant and desirable criterion for VA systems, once that information systems, in general, intend to meet is to achieve successful adoption.

As explained in the previous paragraphs, the User-friendliness, Satisfaction, Error-handling, and Adequacy evaluation areas contain criteria which is important and desirable to any information system and therefore also to VA systems. However, they do not contain criteria related to data analysis itself, which is the VA systems’ goal. Considering this, each of them received the lowest weight, 5%.

This page is intentionally left blank.

5 DEPLOYMENT GUIDELINES OF THE PROPOSED ASSESSMENT METHODOLOGY

In this chapter, we present information regarding the use of our evaluation methodology and its potential applications.

The methodology proposed in this work can be used counting on one or more evaluators. In case of more than one evaluator, the whole **Subsections 5.2 to 5.6** should be followed.

However, if it is to be used with only one evaluator, some adaptions apply. From **Section 5.2,** consider only information regarding documents which should be emailed to evaluators. From **Section 5.3,** Subsections 5.3.1, 5.3.2, and 5.3.5 should not be considered. **Subsections 5.3.3 and 5.3.4** are indispensable because they describe the use of the main tools for the implementation of the methodology. Moreover, they can be intuitively adapted for the use by only one reviewer. **Section 5.6** should also be adapted by disregarding the recommendation "list of disagreements among evaluators". The remaining sections should be properly observed and followed.

5.1 Potential Applications

In the use case of the proposed methodology (**Chapter 6**), the novel set of heuristics and guidelines was used to rate and compare two applications. However, the methodology usage is not limited to comparative scenarios. Since the methodology results in quantitative measures of an application, it can be used to evaluate a single application in isolation by several evaluators as following the standards regarding the inspection of software through heuristic evaluation.

Our assessment can be very interesting to developers seeking to achieve a particular score level. For instance, the levels could be the ones present in **Table 17** (unacceptable, poor, acceptable, good, and excellent).

Potential uses of our methodology include early evaluations of the design of research or commercial systems in order to find strengths and weaknesses. Furthermore, support to decision among systems used for the same context or purpose.

5.2 Recruiting Evaluators

According to Nielsen (1995b), the number of evaluators should be at least 3. However, the reasonable recommended number is about 5, once more than that will not significantly increase the results' precision. That was concluded by Nielsen after performing several evaluation studies. Nonetheless, more raters should naturally be used when higher precision in the results is required.

As done by (Wall et al., 2019), as good practice to recruit evaluators, we recommend to email them with:

- An electronic consent form to sign.
- An evaluator's background questionnaire to collect data for the columns of **Table 16.**
- The Survey for Evaluation of VA software (**Appendix C**) or a link to it.

Table 16 - Background of the invited evaluators (Väätäjä et al., 2016)**.**

ID	Age	Gender	Nationality	Experience in VA or related fields	Work experience in data visualization and analytics	Previous expert evaluations for interactive software

Table 16 will serve to provide a clear picture of the invited reviewers' background. If there is some lack of domain knowledge, it is recommended to train domain experts with understanding of the data and application domain, so that is possible to get insightful feedback beyond usability issues. That was concluded by Väätäjä et al. (2016) in a study using the top ten heuristics for Information Visualization with 5 evaluators with different backgrounds. The raters reported that the lack of domain knowledge made them somewhat uneasy with their capability to perform the heuristic evaluation in-depth.

Access to the Survey for Evaluation of VA software beforehand will allow the evaluators to know exactly what they should cover for the evaluation. That might give them time to refresh some domain knowledge which may be needed to perform the evaluation.

5.3 Applying the Criteria of the Proposed Methodology

In this section, we explain how to apply the criteria (heuristics and guidelines) of the proposed assessment methodology. The criteria application is divided into 5 parts: Brief Section, First Evaluation Phase, Second Evaluation Phase, Detailed Score Justifications, and Debriefing Section.

5.3.1 Brief Session

Before the evaluation itself, all the evaluators should receive instructions in a brief session. This will ensure the evaluators have the same instructions, avoiding bias in their evaluation. To also reduce bias, it is also recommended to give the evaluators a set of tasks or ask them to create and agree on a set of tasks based on their experience and expertise (Wong, 2020).

5.3.2 First Evaluation Phase

In this phase, evaluators should freely use the software which will be evaluated, so that they can gain a feel for the methods of interaction and the scope of the software. In other words, so that the evaluators get familiarized with the application (Wong, 2020).

5.3.3 Second Evaluation Phase

In this phase, it should be done a heuristic evaluation completing the Survey for Evaluation of VAS (**Appendix C)** and the Sheet for Overview and Calculation of the Visual Analytics Score (**Appendix D**). The Survey contains the novel set of heuristics and guidelines, which is divided into the 7 evaluation areas proposed in this work. Each heuristic should be rated from 1 to 7 as we did in the use case of the methodology in **Chapter 6**. Each score has a different meaning (**see Table 20**). Concerning the Sheet for Overview and Calculation of the Visual Analytics Score, it is to facilitate the calculation of the VAs (Visual Analytics Score) and the analysis of the results.

Evaluators should be asked to work independently so that they can form and express their own opinions. Furthermore, to help the generation of more detailed feedback from them, we may ask evaluators to think aloud or probe them with questions while they use the software (Tory & Moller, 2005). This phase may be recorded for notes later.

5.3.4 Detailed Score Justifications

Each score given to each heuristic should be justified by the evaluators while they carry out their various tasks so that we can identify more precisely what is missing for the application to be more in line with the criteria (heuristics and guidelines) set for the VA systems. Evaluators should be asked to be as detailed and specific as possible when registering their justifications beside each heuristic (Wong, 2020).

5.3.5 Debriefing Session

This session involves collaboration between the evaluators and the responsible person for applying the methodology. All the reviews provided by the evaluators should be brought together so that similarities and differences can be seen, discussed, and registered in detail (Wong, 2020). Also, to identify common themes and areas of disagreement among evaluators. If needed, further questions about the assigned scores and justifications should be placed to encourage evaluators to discuss their opinions in more detail (Tory & Moller, 2005). All this effort has the aim to establish a complete list of weak points of the application regarding its VAs (Visual Analytics Score). On top of that, evaluators should be encouraged to suggest potential solutions for the identified weak points. This phase may be recorded for notes later about what expressed verbally by evaluators. The notes from this session together with the ones from the second evaluation phase should be used to enrich the expected outcomes of the evaluation (see **Section 5.6** for information regarding expected outcomes).

5.4 Determining the Visual Analytics Score (VAs)

Once evaluators have completed their ratings of the application using the survey (**Appendix C**), the scores per evaluation area and Visual Analytics Score are to be calculated. **Subsection 4.3.3** explains how to aggregate the scores. The aforementioned subsection suggests an initial approach for the weights to be used in the VAs formula. Nonetheless, different weights can be adopted if the responsible person for the application of our methodology understands that the evaluations areas should have different weights. However, the sum of weights must always total 100%.

5.5 Interpreting the Visual Analytics Score (VAs)

In general, the highest is the VAs (Visual Analytics Score) obtained through the application of the proposed methodology, the more the application under evaluation meets the criteria set of desirable capabilities in VA Systems. **Table 17** presents the scale for interpretation of the VAs.

Table 17 - Scale for the interpretation of the Visual Analytics Score.

VA Score	Meaning
6.7 – 7.0	Excellent
6.1 – 6.6	Good
5.1 – 6.0	Acceptable
3.0 – 5.0	Poor
Under 3.0	Unacceptable

Table 17 presents the scale for the interpretation of the Visual Analytics Score (VAs). These scores meanings aim to translate the overall opinion of raters regarding the satisfaction of the heuristics and guidelines proposed in this work. To reach the excellent score a software needs a score over 6.7. Such excellent score will be achieved if an application obtains a score 7 in all evaluation areas along with the maximum score (7) in at least one of the criteria (heuristics and guidelines) related to Semantic Interaction (**AP22 to AP25** from **Table 7**). As explained for AP22 in **Table 21**, Semantic Interaction for VA is not available in commercial software yet. We decided to include SI criteria in view of the advances of the VA field so that only applications able to deliver this most advanced approach for Sensemaking in VA can be differentiated through our evaluation methodology.

5.6 Expected Outcomes

Lastly, as the main outcomes of the proposed evaluation framework, we suggest a report containing a list of the features of the evaluated application needs to better its VAs (Visual Analytics Score), which will be a reflex of the scores under 6 obtained by each heuristic. Moreover, a list of common themes and areas of disagreement among evaluators, a completed sheet for overview and calculation of the Visual Analytics Score (**Appendix D**), and the Survey (**Appendix C**) with justification from evaluators of the assigned scores to the criteria (heuristics and guidelines).

All this information mentioned above along with the guidelines from **Sections 5.2, 5.4 and 5.5** should be regarded, interpreted, and compiled into the report. This document should have the following sections: brief methodology description, participants (for number and background of them), materials, results, limitations, discussion (when applicable), and conclusion. The report needs to be presented in sufficient detail to allow the reader a clear understanding of its value and contribution (Forsell, 2010).

This page is intentionally left blank.

6 METHODOLOGY USE CASE

Aiming to show the feasibility of our assessment framework and our novel set of heuristics and guidelines, we validated the soundness of our methodology against two top commercial Business Intelligence Analytics tools according to the Gartner Magic Quadrant, namely, Tableau and Power BI (Gartner, 2019). Thus, we could assess how much both tools satisfy the criteria set for evaluating VA systems proposed in this work. On top of that, the results would be an indication of how much the VA approach is already present among top Business Analytics applications.

Tableau is especially interesting for our evaluation test because it is officially entitled as Visual Analytics Software in the literature by (Scholtz, 2017). As for Power BI, it was a relevant option for offering some advanced analytic features and being one of the main Tableau competitors (Gartner, 2019). The Microsoft Research team has put efforts to make Power BI feature VA functionalities. As evidence of these efforts, we found a recent study about bringing Artificial Intelligence to Power BI by such team in Edge et al. (2018). Besides, we could perform a fairer evaluation comparison, in the sense of comparing software developed for the same purpose and already established on the market.

For this evaluation test, due to financial constraints, we had only access to the licenses for students of Tableau Desktop 2020.3 and Power BI Desktop 2.86, which are basic but do not restrict the main analytics features.

6.1 Evaluated Tools and Methodology Application

In the next two subsections, we briefly describe the evaluated applications. Following this, we have the last section, which presents the methodology application along with its results, limitations, and conclusions.

6.1.1 Power BI Desktop

Power BI is a desktop proprietary Microsoft platform released in 2011 (Microsoft, 2020d). It works in association with a cloud application that makes possible to publish reports throughout the business. Power BI can only be installed on Windows OS and is updated every month. It is intended for small to midsize organizations.

Its free license has the same rich visualizations and filters as the paid one, including a natural language question and answering functionality. Additionally, it saves, uploads, and publishes reports to the Web with a limit of 10 GB per user. Its other two types of license, Pro and

Premium, which are paid, allow report collaboration, direct query, more advanced analytics features, and the use of the Power BI Report Server (Folio3, 2019).

Advantages: Inexpensive upgrade; Large custom visualizations range; Easy integration Excel; Quick learning curve for basic use.

Limitations: Bulky user interface; Online reports must be public to the whole Internet.

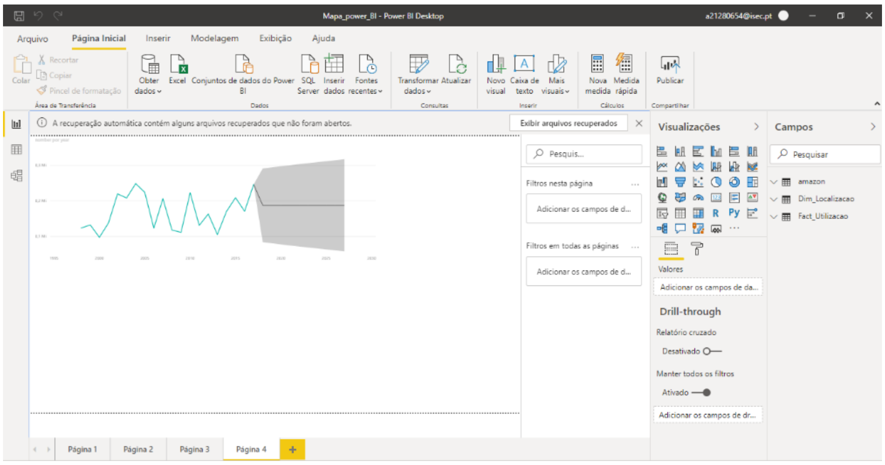

Figure 19 - Power BI Free user interface.

Power BI has available a huge number of learning resources on the web, an active community, and over 70 numerous integrations and data sources (Microsoft, 2020a).

6.1.2 Tableau Desktop

Tableau is one of the most famous "self-service" visualization and analytics tools on the market. Its desktop version was first released in 2004 (Tableau, 2004) and was designed for companies of all sizes. It is not open-source, but it has a commercially free platform, which is updated frequently.

Tableau runs on Windows or Mac OS and can be used in association with its free web repository for publication of visualizations. The same powerful visualization capabilities which its paid desktop and server versions feature are available at no cost in its free license. Data Analyses is possible from sources such as Excel sheets for geographical visualizations, Gantt charts, treemaps, and other templates. It can connect to over 80 different types of data sources (Tableau, 2020c).

Advantages: Quick responsiveness; Extensive training resources available for free; Very intuitive user interface; Dashboards can be viewed on multiple devices.

Limitations: To keep workbooks private, a paid subscription is required; Complex visualizations require time and cost-intensive training.

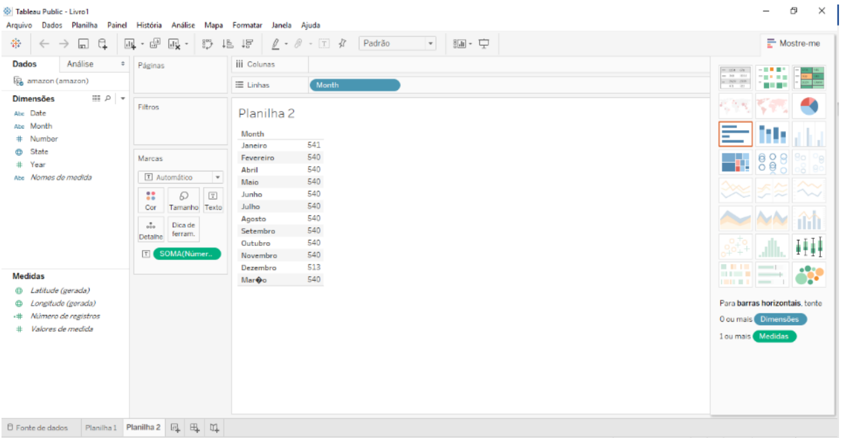

Figure 20 - Tableau Public Desktop user interface.

Tableau surpasses other tools mostly in data visualization. It provides an all-inclusive and user-friendly data visualization experience (Tableau, 2020d).

6.1.3 Methodology Application

We started our case study by creating the same interactive views with Tableau (**Figure 22**) and Power BI (**Figure 23**). We used sales dataset provided by DataFlair (2020), which is composed of 5 excel sheets, namely, customer, order, sales, region, and product. Every sheet has a column in common with the sales sheet so that we could establish relationships among them to form a star scheme as illustrated in **Figure 21** (GeeksforGeeks, 2018). This dataset has neither missing values nor outliers. **Table 18** shows the dataset attributes and their number of entries. We chose a dataset supposed to be easy to understand by most readers of our work regardless of them being from the Data Science field or not. By "number of entries", we mean the number of lines in the dataset.

Table 18 - Datasets which compound our sales dataset.

Dataset	Attributes	Number of entries
Customer	ID and Name	15
Order	ID and Date	15
Sales	Order ID, Customer ID, Place ID, Product ID, Sales ID, Sales, Quantity, Discount, and Cost	15
Region	ID, City and State	15
Product	ID, Category, Sub-category, and Product Name	15

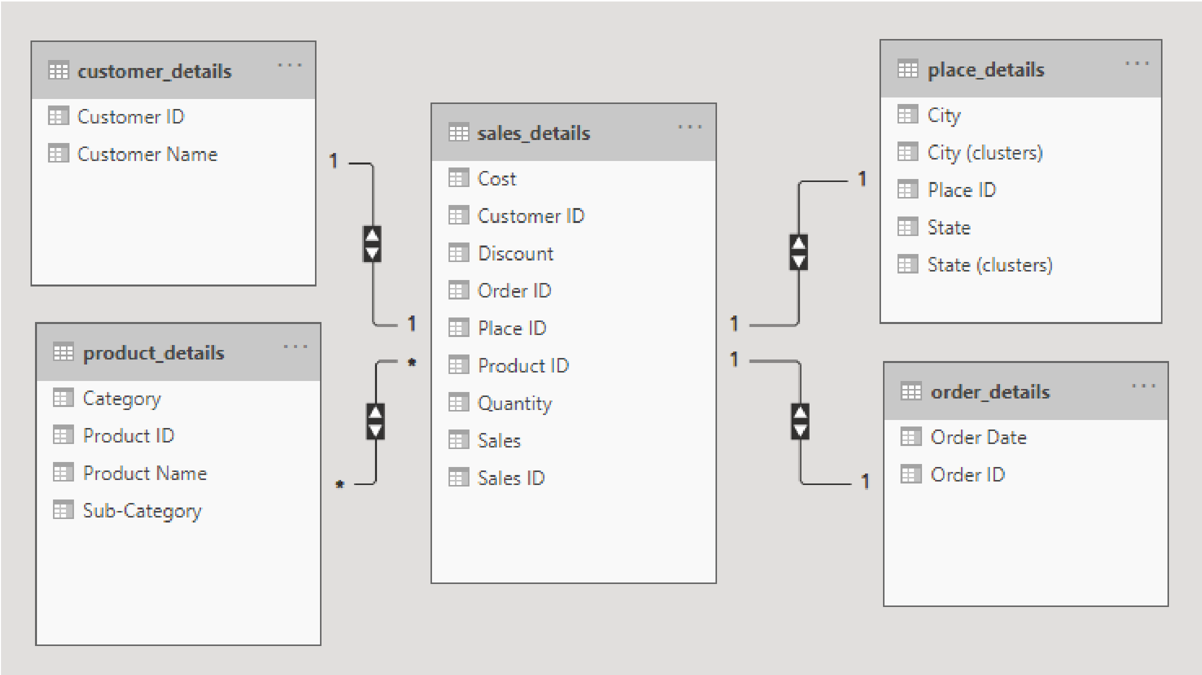

Figure 21 - Sales dataset used in our validation study after importing it to Power BI.

Figure 21 shows the star scheme of the dataset used in our validation study. This is the data model displayed by Power BI Desktop 2.86.

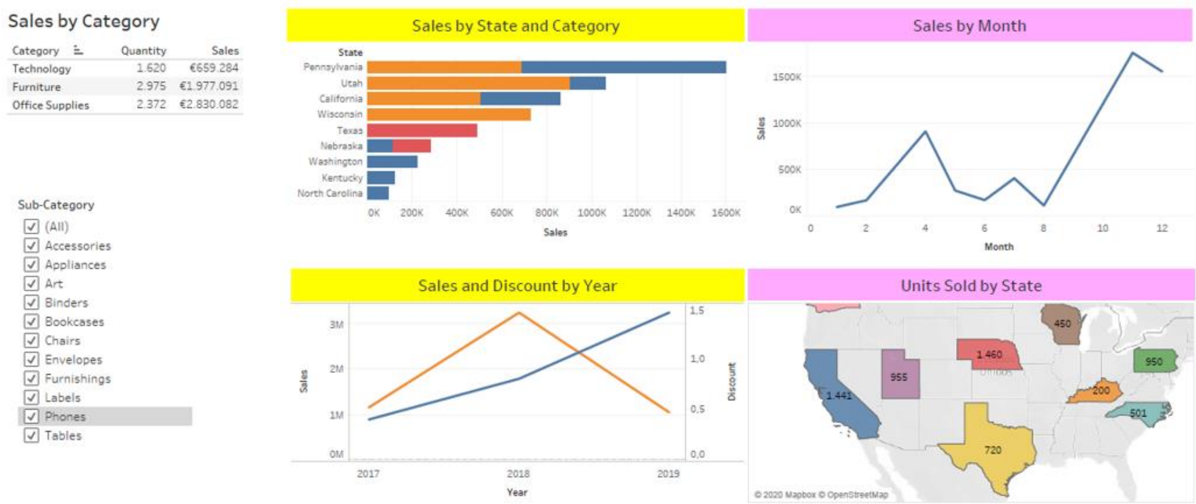

Figure 22 - Tableau Interactive Visualizations.

Figure 22 shows the dashboard build with Tableau to test our novel set of heuristic and guidelines.

Figure 23 - Power BI Interactive Visualizations.

Figure 23 shows the dashboard built with Power BI to test our novel set heuristics and guidelines.

Tableau and Power BI were evaluated under the seven areas proposed in this work, namely Analytic Process, Visualization, Interactivity, User-friendliness, Satisfaction, Error-handling, and Adequacy. Each evaluation area is under a different Subsection (**from 6.1.3.1 to 6.1.3.7**). Moreover, each subsection has a table with 4 columns. The first one contains the novel heuristic set, the next two ones the scores we assigned to each tool regarding the heuristics from the first column. Finally, the last column explains the reason for the assigned scores, which range from 1 to 7 as explained through **Table 19**.

Table 19 - Scores and their meanings in our evaluation.

Heuristic Score	Meaning
1	Strongly Disagree
2	Disagree
3	Somewhat Disagree
4	Neither Agree nor Disagree
5	Somewhat Agree
6	Agree
7	Strongly Agree

Table 19 contains the Likert scale which should be adopted when using our evaluation approach. Each score expresses the evaluator's degree of agreement regarding the satisfaction of a heuristic against the software under evaluation.

Table 20 - Scale for the interpretation of the Visual Analytics Score.

VA Score	Meaning
6.7 – 7.0	Excellent
6.1 – 6.6	Good
5.1 – 6.0	Acceptable
3.0 – 5.0	Poor
Under 3.0	Unacceptable

6.1.3.1 Analytic Process

Table 21 presents the evaluation of Tableau and Power BI using the heuristics and guidelines for the Analytic Process area of the proposed methodology.

Table 21 - Evaluation with heuristics and guidelines for the Analytic Process area.

Heuristic	Score Tableau	Score Power BI	Justification
AP1 - VAS shows key characteristics of data at a glance. VAS should feature visualizations that allow identifying key characteristics of data in a quick short look.	7	7	As we can see in **Figures 22 and 23, Tableau and Power BI** can show key characteristics of data in a quick short look. Both solutions feature visual representations which allowed us to quickly interpret data regarding "sales". We can find in those tools not only the charts and cards presented in **Figures 22 and 23**, but also, for instance, bubble, pie, scatter, and treemap charts. Additionally, KPI and Gauge visuals. **Tableau and Power BI** offer such a wide variety of visuals, that the capability to show key characteristics of data at glance will be also dependent on the end-user expertise on choosing the most adequate visuals according to its type of data and purpose. Considering this, **we strongly agree that both applications meet this heuristic.**
AP2 - VAS makes data relationships noticeable. VAS should facilitate answering questions about the data by making relationships in it noticeable. That is, by making visible, for instance: distribution of variables, correlations, and clusters.	7	7	**Tableau.** It features an analytics feature based on K-Means algorithm which allowed us to uncover patterns by creating clusters regarding sales by subcategory. It was also possible to build a scatter plot to analyse the correlation between discount and costs by state. Besides, we could create a histogram to observe the distribution of sales with bins. Moreover, it was possible to add to both charts, for example, box plots to ease the understanding regarding the distribution of variables. We could also add references (lines or bands) regarding, for instance, average, median, maximum, minimum, and sum to the plots. **Power BI**. It offers a similar intuitive feature for clustering. It allowed us to see patterns regarding the quantities of sold products and sales by city. As for the distribution of variables. Power BI also features box plots charts among other statistic tools. The variety of chart options (visuals) for Statistic purposes is comparable to Tableau's, considering the downloadable third-party visuals. However, the application of Box Plots and other Statistical

			references over histograms and scatter plots are not so easy to apply as in Tableau, which allowed us to do that through "drag and drop". Considering that both tools provide similar means to make data relationships noticeable, **we strongly agree that they meet this heuristic.**
AP3 - VAS provides a new or better understanding of the data. VAS should provide a new or better understanding of the data. This through helping identify unexpected, duplicate, missing, or invalid data. Also, dependent, independent, and important dimensions.	6	6	**Tableau.** It would allow the identification of outliers through charts with the application of box plots. As for duplicates, we found on it no visual features which at least count or highlight them. That would be possible through the creation of calculated fields (Tableau, 2020c). Regarding missing data, we found a feature in visualizations to optionally show for which items in the chart we have no data available. Finally, dependent, and independent dimensions could be identified through the available correlation chart. **Power BI.** Power BI offers similar capabilities regarding this heuristic, the difference is that duplicates would have to be identified through DAX queries (Microsoft, 2020c). Considering that both applications do not provide intuitive visual means only for the identification of duplicates, **we agree that they meet this heuristic.**
AP4 - VAS helps generate data-driven questions. VAS should help the user generate data-driven questions from its analytical outcomes.	7	7	In the case of BI applications, the help for the generation of data-driven questions comes from the variety of visualizations, reports, and dashboards available to analyse the data relationships (as explained for heuristic **AP2**). Considering that Tableau and Power BI offer a similar variety of representations to analyse data, **we strongly agree that they meet this heuristic.**
AP5 - VAS suggests relevant information beyond dataset information. VAS should not only suggest relevant information about the dataset itself and its attributes, but also, for instance, about related views, comments, and data to current points of interest, as well as notification subscriptions for views, artefacts (reports, dashboards, and datasets), people, and groups.	1	1	We found no features for the suggestion of relevant information about related views, comments, and data to current points of interest, as well as notification subscriptions for artefacts, people, and groups on both applications. **Tableau.** It features advanced collaboration functionalities such as teamwork management and subscription to views and artefacts, and suggestions of views based on popular artefacts and user preferences, but not in the license type here under evaluation (Tableau, 2020b). **Power BI**. It has an online application called workspaces for teamwork and content sharing, including also view subscriptions, but not for the license type used in this study (Microsoft, 2020b). Considering the absence of some features and the restriction of collaboration capabilities, **we strongly disagree that Tableau and Power BI meet this criterion.**

AP6 - VAS features visualization which provides a comprehensive data overview with a meaningful visual schema. VAS should feature visualization that provides a big picture/perspective of the data through an accessible data overview and meaningful visual schema.	7	7	**Tableau and Power BI** offer a significant variety of meaningful and accessible charts that allows us to visualize data from several different perspectives. For instance, **Figures 22 and 23** give different perspectives about the measure sales in the charts "sales and discounts by year" and "sales by month". In view of this, it becomes clear that both tools give the necessary means to obtain a comprehensive data overview with a meaningful scheme. This being also dependent on the end-user expertise on choosing the most adequate available artefacts to build such data overview. Considering this, **we strongly agree that Tableau and Power BI meet this heuristic.**
AP7 - VAS provides coordinated views for linked information. Visualizations on VAS should be coordinated together in such a way that action performed in one view affects all other views.	7	7	The dashboards in **Figures 22 and 23** were built with global filters. That is, for instance, the act of selecting a sub-category (card on the lower left side), will make all the charts show only information regarding such sub-category. Likewise, if we select areas from any of the charts in the dashboard. Considering this, **we strongly agree that Tableau and Power BI meet this heuristic.**
AP8 - VAS displays related information nearby. VAS should show related information in close proximity.	7	7	**On Tableau and Power BI**, we found the capability to create charts which automatically place related information in proximity. This is possible due to clustering based on the Algorithm K-Means, as explained for the heuristic **AP2.** Considering this, **we strongly agree that Tableau and Power BI meet this heuristic.**
AP9 - VAS minimizes distractions for the analyst. VAS should minimize distractions for the analyst. That is, minimize aesthetics or interactions that take the user outside of the frame of the task. Minimizing distractions assists endogenous attention and reduction in time.	7	7	**On Tableau and Power BI**, we could progress naturally through the browsing of both applications to build charts and dashboards without distractions. Excess of colours or colours which could cause distractions were found nowhere, as well as fonts which we are not normally used to see in high-quality commercial software. Visual effects, especially in excess, can be a source of distraction, but we did not even realize them. The ways for the main tasks were easily and quickly found. Considering this, **we strongly agree that both tools meet this heuristic.**
AP10 - VAS provides opportunities for serendipitous discoveries.	6	7	Serendipity happens when we find unexpected information while engaged in any information activity (André, Schraefel, Teevan, & Dumais, 2009). People discover and acquire information during the interaction with a data space through search (Khalili, Van Andel, Van Den Besselaar, & Andries De Graaf, 2017). Therefore, any means that allows

VAS should provide opportunities for serendipitous discoveries by displaying information from multiple aspects, as well as related and partially related data points.			information search might be relevant to facilitate serendipity. **On Tableau and Power BI**, it is possible to add search boxes to visualizations, see information through hovering over or clicking on artefacts. Additionally, they even feature a natural language query tool to ask questions about data. Considering that both applications provide similar opportunities for serendipity**, we strongly agree that Power BI meet this heuristic**. Nonetheless, **as for Tableau, we only agree instead**, because the natural language query tool is restricted in our license type (Tableau, 2020c).
AP11 - VAS allows flexibility in the organization of the visual metaphor. VAS should allow flexibility in the organization of the visual metaphor schemes.	7	7	Visual metaphors provide multiple complimentary views of information, which assist analysts regarding reasoning and perception (Thomas & Cook, 2005). The visual metaphors for BI rely primarily on dashboards that are normally composed by a set of charts and cards. These can be customized and organized inside the dashboard in a way that makes the most sense for the analyst to facilitate reasoning and perception. **On Tableau and Power BI**, we can through drag-and-drop actions to easily perform such (re)organization. Considering this, **we strongly agree that both tools meet this heuristic.**
AP12 - VAS facilitates finding starting points or clues. VAS should provide an environment in which the user can capture information to find starting points or clues. That is, it should direct attention to the most critical information.	7	7	As we are dealing with BI software in this study, the starting point for data analysis tends to be naturally first regarding sales, costs, profit, and sold-quantity dimensions of a dataset. Assuming this and that the end-user intuitively built charts including those dimensions, both applications will feature functionality which suggests explanation about data points in a chart just by right-clicking them. Those explanations may work as starting points to lead to the next steps in the analysis. **Tableau.** This functionality is called "Explain Data" on Tableau and is based on Bayesian models (Tableau, 2020a). By the right-clicking the bars from chart "Sales by State and Category" in **Figure 22,** Tableau was informing if the sum of sales for the category was lower or higher than the average, stating weather this category was important to the global increase of sales. **Power BI.** We did likewise in the Power BI chart from **Figure 23.** Through the option "Analyse", the application performed an automatic clustering analysis grouping the states which have similar sales amounts and highlighting so the amounts which most affected the data distribution. Considering both tools make available means that suggest and might direct the analyst to highly relevant information or clues, **we strongly agree that the two applications meet this heuristic.**

AP13 - VAS provides strong retrieval cues for mental models. VAS should structure information in a way which provides strong retrieval cues for mental models* aiding in reasoning**. *Mental models are internal representations mirroring the structure of the external world **Reasoning is the mental process of drawing a conclusion from a set of premises.	7	7	VAS relies on visualizations and therefore visual cues. Those cues are encoded in charts through shapes, volumes, symbols, position, directions, areas, colours and size which represent data (Yau, 2013). Both tools can deliver the same types of visualizations for BI and they apply these visual cues coherently. This can be seen in Figures 22 and 23. As retrieval cues are cues which will be used to activate acquired memories (Brooks, 2012), coherent visual cues work as a source for the strong retrieval cues which facilitate the creation of mental models and drawing of conclusions during a longer data analysis process. Considering that **Tableau and Power BI** can deliver data visualization with similar quality, **we strongly agree that both tools meet this heuristic equally.**
AP14 - VAS allows sharing of evidence and hypothesis. VAS should have the ability to share evidence and hypotheses so that users can create hypotheses regarding their analysis, collect them, and share them with other users. Likewise, it should be possible for the collected evidence. That being feasible, for instance through shared, editable representations; in-app collaborative editing; embedding of annotated views in external media (e.g., email, blogs, and reports); or sharing of views across media (e.g. URLs).	5	5	**Tableau.** Our Tableau license only allows sharing of representations but no edition in the shared visualization. Commenting and Editing, including collaborative, is available only to other license types. We could share our dashboard (**Figure 22**) by uploading it to a repository which assigns a public URL to each of them. Annotations, which can be used to register evidence and hypotheses in views remained visible even after public sharing. The public repository page made available the HTML code for embedding the dashboard within webpages, as well as a button to email the visualization. **Power BI.** Sharing of content in our Power BI license is similar to our Tableau's. However, Power BI does not have a free public website to share and host our representations as Tableau. After using the "Publish" button to upload our dashboard (**Figure 23**), our visual went to our personal online Power BI workspace. There, we could obtain HTML code to embed our representation into websites to then share them publicly. There was no formal annotation feature to apply to our charts, however, Power BI allows us to add text boxes, which can be used likewise to register evidence and hypotheses which will be also part of the visualization and therefore shared automatically together with it. Collaboration is strongly linked with the capability to share work, evidence, and hypotheses, but it is restricted on both applications. Considering that Tableau and Power BI restricts its collaborations capabilities which allow a practical share of representations and annotation from the desktop application, but they still offer means to share them, **we somewhat agree** that they meet this heuristic.

AP15 - VAS supports collection of evidence and annotations in a beneficial organization to sensemaking. VAS should allow collecting and grouping evidence and annotations, as well as to register the need for more evidence or other future actions, preferably through storytelling, in a beneficial scheme to the sensemaking process.	7	7	**Tableau.** It supports the collection of evidence and annotations through data storytelling. It has a separate intuitive interface which fully eases that. We intended to create a data story starting with a big picture of "sales by month" (chart from **Figure 22**) and then focus on the most relevant data points about the chart's topic. As tableau works based on story points, we could add in the first story point the referred chart and in the next ones the same chart highlighting only the relevant data points for our story. Additionally, we could add a title to each story point and text boxes with relevant notes about the chart. Finally, we could sequence them in the way the made the most sense for us to then explain through the story presentation feature. **Power BI.** Collection of evidence and annotations is supported by data storytelling as on Tableau. However, our experience was not as intuitive as in Tableau because Power BI does not offer a separated interface for that. Thus, to build the same data story, we selected the "sales per month" chart (**Figure 23**) and used a feature called "bookmark" to highlight the relevant data points for our story by clicking on the list of bookmarked data points. Text boxes for relevant notes could also be added. There was no functionality to create story points and rearrange them for presentation in sequence as on Tableau. On the other hand, Power BI offers some visuals which alone can build a data story, for instance, the timeline story visualization, which is only adequate to data analysis that regards time. Despite the difference regarding intuitiveness**, Tableau and Power BI** provide means which supports collection of evidence and annotations, as well as its organization in a beneficial scheme to the sensemaking process. Therefore, **we strongly agree that both tools meet this heuristic.**
AP16 - VAS allows registering need for more evidence or other future actions. VAS should allow registering need for more evidence or other future actions regarding the analytic process.	3	3	**On Tableau and Power BI**, we found no feature for the specific purpose of registering the need for more evidence or future actions. However, for instance, text boxes with annotations can be easily added to artefacts containing any kind of information. Therefore, **we somewhat disagree that both tools meet this heuristic.**
AP17 – VAS supports sensemaking by recommending relevant information based on the user's current activity and potential next step. VAS should support sensemaking by presenting	1	1	Sensemaking is the process of understanding and making sense out of data so that a user gradually builds up a mental representation of the information to achieve its analytic goal (Barbulescu, Stoica, & Stoica, 2016). It is an iterative process which includes finding information and extracting its meaning. Besides, a mental model which is refined in loops as the user identifies more supporting evidence, new relations in the information, or even more basic information itself (Blum, Cetin, & Stuerzlinger, 2019). Hence, the recommendation of relevant information during the data

semantically meaningful recommendations that enrich the current analytic process based on the user's current activity and potential next step.			analysis process can play an invaluable role for sensemaking and therefore for the analysis goals. However, we experienced no recommendation of information during our data analysis using **Tableau and Power BI**. We also found no reference to that on the documentation of both tools (Microsoft, 2020c; Tableau, 2020c). This would be a feature to be supported by the interpretation of our interactions with data representations, which seems not to be available on commercial VAS yet. We only found a reference to that kind of recommendation feature on research software, e.g., the VA application present in the cases study by (Cook et al., 2015; Kang & Stasko, 2012). Considering this, **we strongly disagree that Tableau and Power BI meet this heuristic.**
AP18 - VAS displays statistics and measures about data sources, datasets, and/or records. VAS should support evidence discovery by displaying statistics and measures regarding data sources, datasets, and/or records.	5	5	**Tableau.** We found no intuitive display of summaries regarding the number of data sources, datasets, and/or records. Except for counts of dataset rows which could be seen by right-clicking the dataset columns. However, the capability to display statistics and measures about number and types of records is significantly extended if we use the feature "calculated field" to create measures. These measures could be easily applied to the BI charts to display and summarize the desired statistics about de the data sets. **Power BI.** As on Tableau, no intuitive display of statistic summaries about datasets and data sources were found. Also, except for the number of rows regarding the data set being manipulated. To create measures to calculate the desired statistics, we had the DAX expressions instead of the "calculated fields" from Tableau. The measures results could be displayed through the BI charts. To sum up, both tools are very similar regarding this heuristic. Nonetheless, considering we found neither a display regarding the number of data sources nor a mean to calculate this number, both tools did not fully meet this criterion. **Therefore, we somewhat agree that they meet this criterion.**
AP19 - VAS features a visual display of the analytic process. VAS should feature a visual display of the process, so that there is no need to keep external notes.	1	1	As presented in section 2.1, the iterative visual analytics process is basically composed of data transformation and mapping; display of interactive visualizations which iteratively feed underlying models; and knowledge discovery. **On Tableau and Power BI**, it was found no visual display to show the status of our data analysis. Moreover, the data transformation and mapping stage must be performed in a separated user interface from the stage for building and displaying interactive visualizations. In the case of our tools, the interface for interactive visualizations is the one for creating BI charts. In other words, on both software, the visual analytics process will be a self-guided process performed in segregated user interfaces with no display to indicate how far we went in our analytic process, making us

			keep external notes about it. Considering that, **we strongly disagree that both applications meet this heuristic.**
AP20 - VAS provides an easy-to-interpret environment for contextual analysis with relevant information. VAS should provide an easy-to-interpret environment for contextual analysis composed by relevant information for the analysis and suggestions about what may have been overlooked.	6	6	**Tableau.** It features an intuitive environment which eases interpretation about the data being imported and analyzed. In this case study, for instance, we adopted a data source composed by a set of interrelated excel sheets which had columns with the same name. Then, Tableau automatically detected that and suggested the creation of relationships. During the analysis of charts, we could experience suggestion of relevant information, as explained for heuristic **AP12.** However, no suggestions concerning data which might have been overlooked were presented. Besides, no functionality to search for that was found. **Power BI.** We had a similar experience on Power BI, the difference was that Power BI automatically identified and created relationships among our sheets instead of only suggesting them. However, not all the relationships were automatically detected, and we had to add the missing ones manually. Both tools were able to present our imported data in a way that eased the understanding and manipulation of it. They were also able to analyze our charts and based on them to suggest information that might extend our understanding of data. On both, no feature was found to spot overlooked data. Because of that, **we agree that the two applications meet this criterion, but not strongly.**
AP21 - VAS provides transparent automation to the user regarding the underlying mathematical models and parameters. VAS should contain automation, which is transparent to the user, shielding users from the complexity of the underlying mathematical models and parameters.	7	7	**On Tableau and Power BI**, we identified automated data analysis features to make data relationships noticeable, as explained for heuristic **AP2**. Those features are drag and drop and require no parameters to deliver analytical results. This means the user is shielded from the complexity of the underlying mathematical models and parameters. On the other hand, if the analytical results are not satisfactory, the analyst can still use a few parameters. For instance, the user can set the number of clusters for the clustering feature. Considering this, **we strongly agree that Tableau and Power BI meet this heuristic.**
AP22 - VAS captures and understands user interactions. VAS should be able to capture and understand (the kind of action) user interactions in spatial analytic processes such as searching, highlighting,	1	1	Heuristics AP22 to AP25 relate to Semantic Interaction (SI), which seems to be the most challenging approach related to Visual Analytics to support sensemaking to date. We came to this conclusion because we only found a reference to the practical use of it in the literature and among studies which discussed its implementation in non-commercial VA software, as in (Bian, Dowling, & North, 2020; Cook et al., 2015; Endert, Fiaux, & North, 2012). SI for VA systems is a recent concept - from 2012 (Endert, 2014). Furthermore, it is a complex approach to be implemented as it gives to the VA system the responsibility to tune

annotating, and repositioning documents for future automation.			underlying Machine Learning models by capturing user interactions and inferring the analyst's intent (Self, Vinayagam, Fry, & North, 2016). In other words, the system should capture and try to understand the user's cognitive intents as they directly manipulate data projections during sensemaking activity (Bian et al., 2020). Besides, this whole process should be transparent to the user (Endert, 2014; Scholtz, 2011). Since a prerequisite for SI would be the capture and store of user interaction logs from the analytic process itself to benefit the sensemaking process (Endert et al., 2015) and we found no reference to this kind of implementation in Tableau's and Power BI's documentation (Microsoft, 2020c; Tableau, 2020c)**, we strongly disagree that Tableau and Power BI meet this heuristic.** As the heuristic AP21 works a prerequisite for heuristics AP22 to 25, **we consequently also strongly disagree that both applications meet heuristics AP22 to 25**.
AP23 - VAS makes inferences from user interactions. VAS should be able to make inferences (deductions) from user interactions. For example, for suggestions of recommendations regarding the analysis.	1	1	**We strongly disagree that Tableau and Power BI meet this heuristic.** See AP22's justification for clarification about our disagreement.
AP24 - VAS reacts and takes initiative based on inferences from user interactions. VAS should be able to react and take initiative based on those inferences at three levels: interface, computation, and cognitive**	1	1	**We strongly disagree that Tableau and Power BI meet this heuristic.** See AP22's justification for clarification about our disagreement.
AP25 - VAS provides visual feedback regarding the updated model. VAS should also provide visual feedback of the updated model and	1	1	**We strongly disagree that Tableau and Power BI meet this heuristic.** See AP22's justification for clarification about our disagreement.

learned parameters within the visual metaphor.			
AP26 - VAS features teamwork management. VAS should feature group creation and teamwork management, including division of labour among participants.	1	1	**Tableau.** It allows the creation of users and groups with permission rights regarding shared artefacts, such as visualizations. Team members can edit and share data, as well as make follow up queries. We found no reference to feature for the division of labour among users. However, these collaboration functionalities are not available in the tableau license we are evaluating (Tableau, 2020c). **Power BI.** Collaboration features are restricted in our Power BI basic license, as in our Tableau's. Also, Power BI offers similar collaboration capabilities in its other license types. Moreover, there is another Microsoft application called Teams which can be integrated into Power BI to extend collaboration capabilities (Microsoft, 2020c). Nonetheless, considering that our Tableau and Power BI licenses do not provide features to support teamwork management, **we strongly disagree that they meet this heuristic.**
AP27 - VAS features activity indicators per collaborator increasing so engagement. VAS should provide a history of past contributions, to create activity indicators, as well as to aid reputation and visibility of contributions, so that engagement increases.	1	1	As explained for heuristic **AP26**, **Tableau and Power BI** restrict collaboration features for the license types used in this study. Moreover, in their documentation, we found neither reference to history nor activity indicators based on past user contributions. Considering this, **we strongly disagree that both tools meet this criterion.**
AP28 - VAS supports intuitive communication among collaborators. VAS should support intuitive communication to support discussions on common ground. In other words, it should provide intuitive means to share understanding among collaborators to facilitate consensus and decision making.	1	1	**Tableau.** For communication, Tableau provides a comments feature on views to share a conversation about data discoveries with other users. But this feature is not available for our license type (Tableau, 2020c). **Power BI.** It also has communication capabilities similar to Tableau's. Furthermore, additional communication features among team members can be used if we integrate Power BI with another application called Microsoft Teams (Microsoft, 2020c). However, our license type does not allow the use of communication features. Considering that out Tableau and Power BI license types restrict communication among collaborators, **we strongly disagree that they meet this heuristic.**

AP29 - VAS allows to track update of collaborative threads. VAS should allow tracking update of collaborative threads regarding the analysis.	1	1	Collaborative threads depend on the collaborator's comments regarding the analysis**.** As mentioned for heuristic for **AP28,** our **Tableau and Power BI** licenses restrict communication among participants of the analysis. Hence, consequently, it would not be possible to track collaborative threads. Moreover, we found in Tableau and Power BI's documentation no information regarding functionality which alerts or presents information about updated threads. Considering that out Tableau and Power BI license types restrict communication among collaborators, **we strongly disagree that they meet this heuristic.**
AP30 - VAS supports future scenario projections. VAS should support users in making future scenario projections such as forecasting.	7	6	**Tableau.** It offers an intuitive forecast feature which we could easily apply to our "sales and discount by year" chart from **Figure 22.** We chose to forecast the number of sales and discounts for the next two quarters, and just by right-clicking our chart, we enabled the forecasting functionality, which added lines representing the forecast for sales and discounts to the right side of our chart. This feature has several customization preferences, such as forecast length and aggregation options from years to seconds. Also, a functionality which describes the prediction and even mentions how well the forecast fits the actual data. We could also visualize the predictions in table format. **Power BI.** It offers a forecast feature similar to Tableau's. However, it forecast only line charts and with only one measure. Thus, we could not test this feature with the "sales and discount by year" from **Figure 23** as on Tableau. We created then a chart for "sales by month" and activated the forecast function. It was less intuitive because it is not possible to do it by right-clicking the chart as on Tableau. Nonetheless, the main customization options such as forecast length and aggregation options were also present. There was no function to describe the prediction as on Tableau. Considering that Tableau presents more visualization and customization options on forecasting than Power BI, **we <u>agree</u> that Power BI meets this criterion and <u>strongly agree</u> concerning Tableau.**
AP31 - VAS integrates multiple information channels. VAS should provide means to integrate multiple information sources, forming a single unified content collection.	7	7	**Tableau.** As shown in **Table 18,** our data source is composed of 5 datasets (excel sheets) which should be linked to the main one (sales) to form a start scheme. After loading the datasets in Tableau, they were collected into a single user interface and displayed as individual tables, facilitating data management and creation of relationships among datasets. **Power BI.** We had an identical experience in Power BI when integrating data to our data analysis. Considering this, **we strongly agree that both applications meet this heuristic.**

AP32 - VAS increases engagement and attention by using game design elements. VAS should use game design elements to reframe tedious data entry tasks as actions within online games for increasing engagement. For instance, a team-oriented 'scavenger hunt' analysis would allocate more attention.	1	1	The use of game design elements aimed at driving user engagement and motivation in non-game systems is a recent concept called Gamification (Marache-Francisco & Brangier, 2016). Among the typical game design elements, we have points, badges, leaderboards, performance graphs, meaningful stories, avatars, and teammates (Sailer, Hense, Mayr, & Mandl, 2017). They can work as collaboration features. **We found those game design elements neither on Tableau nor Power BI**. Something closer to the "teammates" element would be the identification of users who take part in the same analytic process for content sharing and communication, but as explained for heuristic **AP26 and others related to collaboration,** both applications restrict collaboration features**.** Therefore, **we strongly disagree that Tableau and Power BI meet this heuristic.**
Overall score (simple average)	**4.38**	**4.34**	

6.1.3.2 Visualization Quality

Table 22 presents the evaluation of Tableau and Power BI using the heuristics and guidelines for the Visualization Quality area of the proposed methodology.

Table 22 - Evaluation with heuristics and guidelines for the Visualization Quality area.

Heuristic	Score Tableau	Score Power BI	Justification
VQ1 - VAS facilitates perception via Gestalt principles. VAS should guide and maximize perception via Gestalt principles (proximity, similarity, enclosure, closure, continuity, and connection) in its visualizations.	7	7	To make the information we want to communicate in visualizations identifiable, we should consider the Gestalt Principles (Knaflic, 2015). **Tableau.** By analysing the types of charts present on Tableau, we could observe their alignment with the Gestalt Principles. For instance, good use of **proximity** was observed in scatter plots and packed bubble charts for clustering, once the related data was grouped nearby in the chart. At the same time, the related data had similar colour, shape and orientation assigned. This fulfils the **similarity** principle. Proper use of **enclosure** was found when using the forecast feature, once the chart area concerning the forecasted data was shaded to be distinguished. As for **closure**, if we observe the charts from **Figure 22,** we see no unnecessary elements such as borders and full background shading. This helps data stand out more, revealing so their alignment of them with the closure principle. A clear example of the good application of the **continuity** principle is seen in the chart "Sales by State and Category" from **Figure 23** because the y-axis line is not present at all and our eyes actually still see that the bars are lined up at the same point because of the consistent white space

			between the labels on the left and the data on the right. The absence of an unnecessary element made data stand out more. Regarding connection, we could experience its adequate application at the line chart “Sales by year” from Figure **23** where all the data points are connected in line to help us see the order in the data regarding sales along the years. **Power BI.** As seen in Figure **23,** we could build the same charts we had on Tableau. Hence, Power BI is as able as Tableau to properly apply the Gestalt principles. Considering this, **we strongly agree that both tools meet this heuristic.**
VQ2 VAS provides visualizations with meaningful spatial organization. VAS should care the visualization overall layout, displaying a meaningful spatial organization of the data.	7	7	Spatial organization relates to the overall layout of a visual representation. This includes analysing how easy it is to locate an information element in the display. Also, being aware of the overall distribution of information elements in the representation (Freitas et al., 2002). Those elements are points, lines, areas, and volumes or a composition of them in visualizations (Ignatius & Senay, 1996). **Tableau.** By analysing the charts in **Figure 22,** we see that the distribution of lines, points and areas are coherent and ease our understanding about data. For instance, in the chart for sales and discounts by year, the position and direction of lines show us clearly when they were increasing, decreasing or stable. The lines were placed automatically in that layout and so it was easy to spot along the years when the sales and discounts were higher and even compare amounts of sales and discounts by seeing their distribution along the lines. **Power BI.** As we can see in **Figure 23,** we could build in Power BI similar charts with identical quality to Tableau’s. Therefore, it is as good as Tableau in providing visualization with meaningful organization. Considering this, **we strongly agree that both tools meet this heuristic.**
VQ3 - VAS avoids dense visualizations by featuring properties for size and distance to avoid dense visualizations. VAS should offer appropriate and easy to interpret representations for properties such as size and distance in visualizations,	7	7	**Tableau.** When building the dashboard presented in **Figure 23,** we could freely resize charts with the mouse pointer and relocate them by dragging. On the other hand, we found no properties to edit size and distance inside the charts, which by default are not dense visualizations. In case of Tableau, as we intuitively can alter size and distance among visualizations (charts) when building dashboard-like visualizations, the task of avoiding dense panels will be more on the side of the user than on Tableau’s. **Power BI.** It offers the same easy-to-interpret features for size and distance in visualizations as Tableau and generates chart not dense by default.

avoiding so dense visualizations.			Considering that both tools can generate charts which by default are not dense and allows us to intuitively place, resize and reorganize them inside dashboards, **we strongly agree that they meet this heuristic.**
VQ4 - VAS uses animation only to show an effect that moves over time. VAS should use animations only to show an effect that moves over time. Give analysts control to manipulate the speed of the animation.	7	7	**Tableau.** It has a feature for animations in visualizations which can be used in charts. To tested it, we applied it to the chart sales by month. As a result, we could watch an animated sequence of sales results over time plotted in our chart. The animation feature was disabled by default. Also, it was possible to custom its speed and select a transition style. **Power BI.** It also allows animations in charts. However, this functionality is not native as on Tableau. It should be downloaded as third-party visual and added to the workspace of the chart you wish to animate. We teste it in our chart sales by month and could watch the sales amount evolving over the months. The third-party visual looks like a small control panel with play, pause, stop, reverse and forward buttons. Within its configuration options, it is animation speed and visual aspects of the control panel, such as colour and size. Considering this heuristic does not encourage the use of animations and that both applications optionally feature animations which can be used only to present behaviour over time, **we strongly agree that the two tools meet this heuristic.**
VQ5 – VAS displays only relevant information and elements in a straightforward fashion. VAS should avoid misleading and complex representations by displaying only relevant elements to the analytic process in a straightforward fashion.	7	7	The display of relevant information depends logically on the alignment with heuristics **V1 to V4**. Moreover, on heuristic **AP9 (**reduction of distractions during the analytic process). Furthermore, through the BI charts generated by Tableau and Power BI, we could understand well and without distractions the data regarding the dimensions we considered relevant. Considering that both applications fully met the referred heuristics and generated charts which displayed straightforwardly the kind of information we judge to be relevant for a sales dataset, **we strongly agree that the two BI tools meet this heuristic.**
Overall score (simple average)	**7.00**	**7.00**	

6.1.3.3 Interactivity

Table 23 presents the evaluation of Tableau and Power BI using the heuristics and guidelines for the Interactivity area of the proposed methodology.

Table 23 - Evaluation with heuristics and guidelines for the Interactivity area.

Heuristic	Score Tableau	Score Power BI	Justification
I1 - VAS features self-descriptive interactions. VAS should feature self-descriptive interactions. That is, it should allow users intuitively understand what they can do with the interaction and how they can do it (Usability.de, 2020).	7	7	**Tableau.** Every visualization created on Tableau features interaction capabilities which were normally located at the visualizations' edges and could be easily interpreted after hovering over them (a short description showed up), e.g., sorting. Moreover, there were always interactions options properly named, e.g., selection and annotation, which were available after right-clicking inside the visualization. **Power BI.** Our experience with interaction descriptions in Power BI was similar to ours in Tableau. Considering that hovering over and right-clicking were in most cases enough to understand which kinds of interactions were available on Tableau and Power BI, **we strongly agree that both tools meet this heuristic.**
I2 - VAS provides tools for data manipulation. VAS should provide tools to help users in data manipulation. For example, tools for filtering, clustering, pruning.	6	7	**Tableau.** After loading our datasets, we could easily edit data in Tableau to prepare it to be used in our visualizations. We could rename and create columns. Also, change their value types. Besides, it was possible to copy, sort and filter values. There was also a "hide data" functionality to make Tableau ignore data without deleting it from the dataset. **Power BI.** In Power BI, we found more data manipulation options. Besides the options from Tableau (except for the hiding feature), we found features for transposing, removing, splitting, and merging columns. Also, for replacing values and extracting data parts. Considering that Power BI offers a more complete solution regarding data manipulation than Tableau, **we strongly agree that Power BI meets this heuristic.** As for Tableau, **we agree that it meets this criterion.**
I3 - VAS provides capabilities for data exploration. VAS should feature useful interactive capabilities to help investigate data in multiple ways. For example, zooming; navigation and querying (including selection of objects; viewpoint and geometric	7	5	**Tableau.** In the charts from **Figure 22**, we found zooming, selection (radial, rectangular, and lasso), viewpoint manipulation, and geometric manipulation capabilities. We could also enable filters based on one of the dataset dimensions. Those filters were available in forms such as list and wildcard match. The implementation of a general search box that would filter all the dashboards charts at once would be possible through the resource "calculated fields" (Tableau, 2020c). Furthermore, we also tested the interaction capability presented in the justification for **AP7**. **Power BI.** We found zooming and viewpoint manipulation capabilities only for geo map charts, e.g., the "Units sold by State" chart from **Figure 23.** Moreover, geometric manipulation and selection (by clicking data points). As for

manipulation; geometric and searching)			searching, it was possible only through the cards for filtering and slicing, e.g., the “Sub-category” card (left lower side from **Figure 23**). Considering Tableau provided more interaction capabilities for data exploration than Power BI and that those capabilities are present in a larger variety of charts, **we strongly agree that Tableau meet this heuristic.** As for Power BI, for being somewhat limited regarding interactions for data exploration, not allowing a common capability such as zooming in its most charts, **we somewhat agree it meets this heuristic.**
I4 - VAS allows interactive visualization customization. VAS should support customization of the visualization. For instance, by using different attributes of the data to reorganize its appearance and supporting several dimensions simultaneously in it.	7	7	**Tableau.** We could create charts freely by picking up dimensions to be plotted simultaneously in our chart. For instance, we have the chart “Sales and Discounts by year” (**Figure 22),** where the dimensions sales and discounts are shown in the same chart simultaneously per year. Besides that, we found resources on Tableau to swap rows and columns. In the chart “Sales by State and Category” (**Figure 22)**, we could sort its bars by sales and state. Moreover, there was always customization options regarding colours, sizes, fonts, alignments, labels, lines, and borders for all charts. **Power BI.** We found on Power BI similar customization capabilities for interactive visualizations. Considering that both applications feature similar types of visualizations with similar customizations option, **we strongly agree that they meet this heuristic.**
I5 - VAS avoids complex commands and queries in visualizations. VAS should avoid complex commands and textual queries in visualizations by providing direct interaction with the data representation.	7	7	**On Tableau and Power BI,** it was not necessary the use complex commands or queries to filter data in the representations built in **Figures 22 and 23**. The filtering possibilities were explained for heuristic **I3.** Therefore, **we strongly agree that both tools meet this heuristic.**
I6 - VAS provides a way to backtrack or undo actions. VAS should provide a way to track changes in information and undo actions (a history with all user actions may be used).	7	7	**On Tableau and Power BI,** we found the undo feature. Considering that, **we agree that both applications meet this heuristic.**

I7 - VAS provides alternative ways to perform a task. VAS should provide alternative ways to perform a task, such as shortcuts for experienced users, to increase the interaction speed and to reduce time.	7	6	**Tableau.** Most of the tasks regarding visualizations we could perform by right-clicking on them or through menus, e.g., formatting options. There were also tasks such as sorting, which could be done through the taskbar or buttons in the visualization (Beside the axis names from the chart “Sales by State and Category” chart from **Figure 22**. As for keyboard shortcuts, Tableau‘s documentation informs 54 (Tableau, 2020c). **Power BI.** Differently from Tableau, we did not have the option to perform the same tasks through menus and right-clicking the visualizations. Menus on Power BI are to display taskbars for functionalities which are not available by right-clicking. For the “Sales by State and Category” chart from **Figure 23,** the sorting functionality was only available through a three-dot menu located at the visualization’s upper corner**.** As for keyboard shortcuts, Power BI documentation presents 38 (Microsoft, 2020c). Considering that **Tableau** allows us to perform several of the same tasks regarding visualizations through menus and right-clicking. Also, it offers more keyboard shortcut options in comparison with Power BI, **we strongly agree that it meets this heuristic.** Concerning **Power BI**, as it offers less keyboard shortcut options, and right-clicking is generally not an alternative way for tasks available in menus and taskbars, **we only agree it meets this heuristic, instead of strongly agreeing.**
I8 - VAS provides interaction with minimal need for repetitive actions**.** VAS should provide means to explore visualizations and overall system functions avoiding as much as possible repetitive actions on the part of the end-user.	7	7	Overall speaking, **on Tableau and Power BI,** we realized unnecessary repetitive actions neither while loading and exploring data nor building visualizations. Nonetheless, when we needed to undo actions, we always had to repeatedly click on undo or redo buttons until because both tools do not feature backtrack of actions with a history. Considering that in most cases we did not have to perform unnecessary repetitive actions to achieve a task, **we agree that both applications meet this heuristic.**
I9 - VAS gives proper feedback to user actions within reasonable time. VAS should provide appropriate feedback as a response to user actions within reasonable time.	7	7	Delays up to 0.1 seconds are not noticeable (instantaneous response), therefore no feedback is necessary. However, some feedback is necessary for delays between 0.1 and 1.0. Furthermore, as a rule of thumb, percent-done progress indicators should be used for operations taking longer than 10 seconds (Nielsen, 1993). **Tableau.** While building our dashboard from **Figure 22,** most actions had an instantaneous response and we never experienced a delay longer than 10 seconds. We observed longer processing times for data loading and manipulation operations, which had their delay represented by a dynamic loading indicator. As our main tasks were related to building visualizations and loading data by mostly using buttons and

			drag and drop capabilities no text outcome to inform success or failure of actions would make sense. As to feedbacks for data inputs, while trying the resource "calculated fields" to create a correlation chart, we obtained instantaneous messages telling whether the inserted formula was right. **Power BI.** While building our dashboard from **Figure 23,** our experience on Power BI was comparable to Tableau's. However, the data manipulation and loading actions were somewhat longer than in Tableau but not longer than 10 seconds. Concerning feedbacks for data inputs, we experienced them in the natural language query feature called "Q&A", being the outcomes regarding the typed queries always immediate regardless of the query's coherence. Considering that both applications always provided satisfactory feedback to our actions and within the times considered reasonable by (Nielsen, 1993), **we strongly agree that both applications meet this heuristic.**
I10 - VAS guides users towards making specific actions VAS should guide users on specific actions by showing selectable options; windows titles; system or task status; data fields with labels informing acceptable values and formats	7	7	**Tableau.** When building the visualizations from **Figure 22,** there was a taskbar on the right side displaying all the chart options. Besides that, selectable options were present in menus and some cases through right-clicking visualizations or empty workspaces. Moreover, windows, menus and tabs were named with intuitive titles. Furthermore, while Tableau was busy performing a longer task, it displayed a dynamic loading indicator to tell the system was still busy with that. Finally, most of the data fields were properly labelled and to filter or search for information. Nonetheless, they did not inform acceptable values and formats because they accepted any kind of character to ease filtering or searching. We tried also to manually add data to dataset fields (cells), but it was not possible. **Power BI.** When building the visualizations from **Figure 23,** our experience with Power BI was similar to Tableau's. The difference was that Power BI allowed us to edit cells (fields) from the imported data by replacing values for similar value types. For that, the application pops up a data field which does mention acceptable values and formats for accepting any value type or format. Considering that both applications show when they are still busy with a task. Also, that they indicate through windows titles, labels, data fields, and selectable options which ways the end-user may take to perform tasks, **we strongly agree that both tools meet this heuristic.**
I11 - VAS facilitates the understanding of relationships between the various user interface items. Should provide means to understand the	7	7	**Tableau.** While building the dashboard from **Figure 22**, we had no difficulty to interpret the relationship between the user interface items on Tableau because related items were somehow grouped. For instance, all the chart options were grouped on the left side of the screen, had a similar format, and were represented by different colourful mini charts. Thus, we could easily understand the purpose of that group of items. Another example would be the menu the items placed at the top of the screen, with similar formats and

relationships among items by grouping similar objects according to formats and graphical features; screen areas; and between different classes (groups) of objects.			simple graphical design as the most standard menus we normally find in user interfaces. **Power BI.** While building the dashboard from **Figure 23,** our experience was comparable to Tableau's. A difference would be, for instance, the group of items for building visualization located in a different screen area. Besides that, in a format of little colourful squares format with simpler graphical features. For being easy to understand the relationships between the user interface items considering their location, format, and similarity on both tools, **we strongly agree that Tableau and Power BI meet this heuristic**.
Overall score (simple average)	**6.73**	**6.55**	

6.1.3.4 User-friendliness

Table 24 presents the evaluation of Tableau and Power BI using the heuristics and guidelines for the User-friendliness area of the proposed methodology.

Table 24 - Evaluation with heuristics and guidelines for the User-friendliness area.

Heuristic	Score Tableau	Score Power BI	Justification
UF1 - VAS provides coherent UI elements. VAS should follow similar meaning and design choices in similar contexts. That is, the interface elements should be coherent.	7	7	**Tableau.** While using Tableau to build our dashboard (**Figure 22),** we observed the use of similar colour, phrasing, text justification and punctuation in UI elements such as buttons, text fields, icons, notifications, and windows. Also, similar objects such as windows buttons (exit, close, and minimize/maximize) were displayed in the same way and at the same location. Moreover, similar functions, e.g. building of different visualizations types overall followed the same procedure. We could always understand the individual interface elements and find them in the places where similar elements usually are in other applications, which means alignment to platform conventions (Nielsen, 1999). **Power BI.** While using Power BI to build our dashboard (**Figure 23**), our experience in Power BI was comparable to Tableau's. Considering we always found coherent UI elements while operating both tools, **we strongly agree that hey meet this heuristic.**
OF	7	7	**Tableau.** Most of the signs we had to deal with while operating Tableau were somehow familiar. In cases, for instance, where icons were unfamiliar, we could easily obtain textual hints about their meanings by hovering over them. **Power BI.** Our experience on Power BI as to familiarity with UI signs was similar to Tableau's.

			Considering that most signs on both applications had an expected meaning, **we strongly agree that both tools meet this heuristic.**
UF3 - VAS matches user characteristics with the UI characteristics. VAS should feature a UI compatible with the user characteristics, (language, measurement units, calendar, and accessibility capabilities).	6	7	**Tableau.** As to preferences, it can work in 11 different languages and it respects cultural-related requirements by adopting automatically calendar and measurement units used from the operating system settings. When it comes to accessibility, we found in the documentation best practice guides on designing accessible views. That is, guidance on how to edit visualizations regarding titles, filters, legends, captions, marks, and labels conform to the Web Content Accessibility guidelines by the US Government (W3, 2008). Besides that, it was possible to interact and explore the system and visualizations using commonly supported WAI-ARIA standards for keyboard navigation, such as ESC to clear mark selections in views (W3, 2016). **Power BI.** It can work in 44 different languages. It obtains its default calendar and measurement units from the operating system settings as Tableau. Concerning accessibility, it also has a guide on how to build accessible visualizations based on Web Content Accessibility guidelines by the US Government (W3, 2008). Moreover, it features comparable keyboard navigation for accessibility. We found a built-in accessibility feature for titles, labels, markers, themes, and alternative texts descriptions for visuals. Also, accessibility features for colours, focus, high contrast, and display of data table. Both applications offer similar UI compatibility with the user characteristics, but Power BI features more explicit capabilities (built-in features) for accessibility than Tableau. Thus, **we agree that Tableau and Power BI meet this heuristic, being the agreement in regard to Power BI stronger.**
UF4 - VAS provides customizable workspaces VAS should provide means to customize the UI and alternative ways to perform a task.	6	3	**Tableau.** Tableau workspace consists of menus, a toolbar, data pane, cards, shelves, and a bar with tabs for worksheets. Except for menus and worksheet bar, all can be hidden or moved to other screen areas. **Power BI.** Power BI workspace consists of menus which reveals different toolbars, panes, and a worksheet bar. All of them cannot be moved. The panes can be hidden, and the toolbars could be switched to a smaller format so that we could gain free workspace. Tableau is much more flexible regarding its workspace layout than Power BI. Thus, **we agree that Tableau meets this heuristic and somewhat disagree that Power BI does it.**
UF5 - VAS makes easily visible all possible actions for the user. VAS should make easily visible all	6	6	**Tableau.** While building our dashboard from **Figure 22,** the basic actions needed to build the visualizations which would compound it were intuitively performed because, as explained for heuristic **UF2**, most signs were familiar. We needed help to know how to assemble our dashboard and for that, we found a tutorial inside the documentation with video. Tableau offers extensive documentation, a

possible actions the user can perform, be intuitively or through help, documentation, and tutorials.			comprehensive collection of free training videos, online help, and an active community. **Power BI.** Our experience when building the dashboard from **Figure 23** was comparable to our Tableau's. Power BI also offers extensive documentation, a comprehensive collection of free training videos, online help, and an active community. Considering our experience when building visualizations, **we agree that Tableau and Power BI meet this heuristic,** being our agreement not strong because we cannot assure that they make **all** the possible actions easily visible.
Overall score (simple average)	**6.00**	**5.80**	

6.1.3.5 Satisfaction

Table 25 presents the evaluation of Tableau and Power BI using the heuristics and guidelines for the Satisfaction area of the proposed methodology.

Table 25 - Evaluation with heuristics and guidelines for the Satisfaction area.

Heuristic	Score Tableau	Score Power BI	Justification
S1 - VAS is effective in representing high-quality of analytical outcomes. VAS should show the high quality of analytic outcomes on the visual interfaces. That is the system should be effective in depicting through the visualization the outcomes from the automated data analysis.	7	7	Considering both applications fully meet heuristics **AP1, AP2, AP3, and AP6,** which are related to the system capability to reveal data aspects which might lead to analytical insights. Moreover, heuristics **V1** to **V5,** which refer to the visualizations' capability to display well aspects of our data (analytical outcomes), we strongly agree that **Tableau and Power BI** meet this criterion.
S2 - VAS ensures the end user's subjective assessment is overall positive VAS should ensure an overall positive end-user subjective assessment. This involves overall subjective satisfaction about the system, which relates to how pleasant and easy-to-use it is, as well as	7	7	Subjective satisfaction is about how pleasant it is to use a system. To score Tableau and Power BI under this heuristic, we asked ourselves the typical questions to measure subjective satisfaction according to (Nielsen, 2010): • How easy was to learn the system? • Was using this system a frustrating experience? • Did I have the feeling this system allows me to achieve very high productivity? • Did I get worried that many of the things I did with this system may have been wrong? • Can this system do all the things I think I would need? • Is this system very pleasant to work with?

frustrating experiences, and productivity through it.			Power BI was easier than Tableau to learn because we were already familiar with Microsoft products. However, both were easy to learn, especially through all the available good documentation and free training resources. We had no frustrating experience, as we were able to build our dashboards (Figures 14 and 15) relatively quickly. Thus, we obtained the feeling of being productive and doing everything right. Also, both systems meet our task goals and expectations. Finally, we found pleasant to work with both applications considering their responsiveness and absence of errors during our experience. Our overall positive experience with both tools made us believe that many efforts to assure the end-user satisfaction were done. Thus, **we agree that Tableau and Power BI meet this criterion.**
S3 - VAS is considered highly useful. VAS should be considered highly useful. In other words, it refers to whether the system provides the features the user needs.	6	6	Usefulness refers to whether the system can be used to accomplish some desired goal. In other words, it refers to whether it provides the features you need, and how easy and pleasant these features are to use (Nielsen, 1993). We considered Tableau and Power BI to be highly useful, once they feature what we needed to build the dashboards from **Figures 22 and 23.** Besides that, we had an overall pleasant experience because all the used features were whether intuitive or easy to learn through their documentation. On top of that, both fully met 16 out of the 32 heuristics regarding the Analytical Process (AP1 to AP32), which an expressive result because the features related to 5 heuristics (AP22 to AP25) are not yet available in the industry as explained for **AP22**. Thus, we agree that both applications meet this heuristic.
S4 - VAS minimizes the needed resources to achieve the goal VAS should maximize efficiency by minimizing the necessary resources to achieve the goal. It should maximize the speed and minimize the number of steps to achieve an objective.	7	7	Efficiency refers to the speed and the number of steps to achieve an objective. That is, how fast the user can finish its job (Jordan, 2020). The number of steps to build our dashboards from **Figures 22 and 23** in Tableau was equivalent. Also, we never got the feeling that we had to perform redundant steps, which made us understand that the number of necessary steps for our goal was optimized and therefore minimized. **Thus, we strongly agree that both meet this heuristic.**
Overall score (simple average)	**6.75**	**6.75**	

6.1.3.6 Error-handling

Table 26 presents the evaluation of Tableau and Power BI using the heuristics and guidelines for the Error-handling area of the proposed methodology.

Table 26 - Evaluation with heuristics and guidelines for the Error-handling area.

Heuristic	Score Tableau	Score Power BI	Justification
EH1 - VAS prevents, diagnoses, and correct errors. VAS should prevent, diagnose, correct, and recover from errors with clear and informative messages, giving reasons, as well as the means to correct them.	7	7	According to the interaction design conventions, to prevent errors the system should match target users' expectations. Also, feature helpful constraints, good defaults, forgiving formatting, communicate affordances, warn before errors are made, preview functionalities, undo functionality, and confirm before destructive actions. Finally, remove memory burdens (Laubheimer, 2015a, 2015b). **Tableau.** Our experience while building the dashboard from **Figure 22** with Tableau was almost error-free. As to its error prevention capability, we found that Tableau in most cases corresponded to our expectation about possible interactions. We found helpful constraints, for instance, when creating charts because only the charts compatible with the amount and kinds of measures and dimensions became available to use in the chart bar. Also, good defaults in regard to charts formatting. In search bars, we could input any kind of characters, which means forgiving format. As for the communication of affordances, it was always easy to identify them through icons, hovering over, or right-clicking. We obtained a warning when creating calculated fields because we inserted an invalid formula. We found no explicit preview functionality, but visualizations could be built to fit different screen sizes such as smartphones. We experienced confirmation before destructive actions such as deletion of workbooks or the act of closing workbooks without saving changes. Elimination of data tables was only possible through right-clicking, reducing so the likelihood of accidental deletion and the need for such confirmation. There were no memory burdens once there was no need to keep information in our own memory while moving from one step to another to build our visuals. We obtained an error when trying to connect to Tableau's public repository to publish our dashboard. The error happened because we inserted an invalid URL. As a result, a window popped up stating that our action could not be completed along with a code (Internet communication error), suggestion to check the server name, and an explanation that the URL's hostname could not be resolved. Besides, there was a button to copy the error message and a link to Tableau's support page where we could paste the error message. This explains how Tableau tries to diagnose and correct errors. **Power BI.** Our experience with Tableau was similar to our Tableau's. However, helpful constraints regarding chart building were less straightforward because we could pick any chart type up and the constraints about the supported measures and dimensions would be revealed only after chart

			selection. There were also warning when inserting wrong DAX formulas to try to build a measure. Concerning confirmation before destructive action, we were asked to confirm deletion even before eliminating. On the other hand, we could perform deletion not only through right-clicking. We obtained an error when trying to connect to an invalid SQL database. The error windows had the same elements that Tableau's windows. However, there was neither a bottom to copy the error nor a link to Power BI's page support. Considering that we had an almost error-free experience with Tableau and Power BI. Also, we identified the alignment of both with the interaction design conventions to prevent errors. Moreover, the way they inform, diagnose, justify errors, and suggest corrections for them. **We strongly agree that both tools meet this criterion.**
Overall score (simple average)	**7.00**	**7.00**	

6.1.3.7 Adequacy

Table 27 presents the evaluation of Tableau and Power BI using the heuristics and guidelines for the Adequacy area of the proposed methodology.

Table 27 - Evaluation with heuristics and guidelines for the Adequacy area.

Heuristic	Score Tableau	Score Power BI	Justification
A1 - VAS is adequate to its context of use. VAS should be compatible with the context for which it was designed. That is should be suitable to facilitate the analytical goals of its context of use.	7	7	Considering the good scores **Tableau and Power BI** obtained in the previous categories. Also, that these good scores are a result of our user experience exploring relevant analytical features for VA through the building of typical BI charts to compound dashboards and reports, which are the core of Business Analytics tools, **we strongly agree that both tools meet this heuristic.** That confirms the high popularity of both applications on the BI market.
Overall score (simple average)	**7.00**	**7.00**	

6.1.3.8 Results

This subsection presents the results regarding the scores Tableau and Power BI obtained through the application of the evaluation methodology proposed in this work. **Table 28** presents the scores that Tableau and Power BI obtained for each evaluation area. The values per area were obtained by doing the average mean of the scores (from 1 to 7) given to each heuristic during the evaluation. As to the final score, it was calculated using the weighted mean once the evaluation areas have different weights as follows: Analytic Process (40%), Visualization Quality (20%), Interactivity (20%), User-friendliness (5%), Satisfaction (5%), Error-handling (5%), and Adequacy (5%). The justification for such weights was explained in **Subsection 4.3.3.**

Table 28 - Tableau and Power BI's scores per evaluation area and their final scores.

Evaluation Area	Weight	Tableau	Power BI
Analytic Process	40%	4.41	4.41
Visualization Quality	20%	7.00	7.00
Interactivity	20%	6.91	6.73
User-friendliness	5%	6.40	6.00
Satisfaction	5%	6.75	6.75
Error-handling	5%	7.00	7.00
Adequacy	5%	7.00	7.00
Visual Analytics Score		**5.90**	**5.85**

As for Tableau, it obtained its lowest score in the category Analytic Process (**4.41 out of 7**), which has the highest weight according to our methodology **(40%)**. However, this score is still high considering that we probably would not find in the industry any software able to meet heuristics **AP22 to AP25**, which refers to Semantic Interaction and is the most challenge sensemaking approach in Visual Analytics as explained for heuristic **AP22** in **Table 21**. Therefore, we consider that we would not find any software able to reach a scored higher than **(6.06)**, which would be the case of an application being rated with 7 in all heuristics except for heuristics **AP22 to AP25**. Furthermore, Tableau did not score higher because it was weaker in the heuristics related to register and share of evidence and hypothesis. Also, it presented no game design elements to increase engagement. Along with that, it featured no functionality regarding the recommendation of relevant information for the current analysis, display of the analytic process, and collaboration. Lastly, it scored over 6 in the rest of the evaluation categories, which indicates a full alignment of Tableau with these other relevant areas to Visual Analytics. Concerning, Power BI, it also obtained the lowest score **(4.41 out of 7)** for the most important evaluation area: Analytical Process. Being that also a good score considering what was explained above for Tableau's score. Furthermore, it was weaker or unable in the same heuristics mentioned above for Tableau, except for being weaker than Tableau in view manipulation capabilities and lack of flexibility to personalize workspaces. Finally, it also scored over 6 in the remaining categories, obtaining likewise an acceptable final VAS (Visual Analytics)

Table 29 presents the recommended features for Tableau and Power BI to achieve a higher score. The table lists all the evaluation criteria where both tools obtained scores under 6. In other words, it points out Tableau and Power BI weaknesses. It is important to highlight that we recommended features regarding AP16, AP26, AP28, but that those features are already available in Tableau and Power BI's licenses different from ours.

Table 29 - Recommended Feature for Tableau and Power BI.

Heuristic	Evaluation Area	Features	Recommended for:
AP5	Analytic Process	Features for the suggestion of relevant information about related views, comments, and data to current points of interest, as well as notification subscriptions for artefacts, people, and groups on both applications.	Tableau and Power BI
AP14*	Analytic Process	Features for sharing of visualizations, annotations, and reports directly from the application.	Tableau and Power BI
AP16	Analytic Process	Feature for registering the need for more evidence or future actions about the data analysis process.	Tableau and Power BI
AP18	Analytic Process	Summary display with statistics regarding the number of data sources, datasets, and records.	Tableau and Power BI
AP19	Analytic Process	Feature which displays the status of the analytical process.	Tableau and Power BI
AP26*	Analytic Process	Features for teamwork management.	Tableau and Power BI
AP27*	Analytic Process	Features which displays activity indicators per collaborator.	Tableau and Power BI
AP28*	Analytic Process	Features for communication among collaborators.	Tableau and Power BI
AP29*	Analytic Process	Features which alerts or displays information about the collaborative threads.	Tableau and Power BI
AP32	Analytic Process	Implementation of game design elements to turn the analytic process in a kind of game which engages collaborators.	Tableau and Power BI
I3	Interactivity	Zooming and view manipulation capabilities in other visualizations other than geo maps. For instance, line and bar charts.	Power BI
UF4	User-friendliness	Personal workspace which allows adding and removing the needed tools and layout personalization.	Power BI

6.1.3.9 Study Limitations

The study had a few limitations that likely affected our findings. First, as we do not have the background possessed by experts from the VA field, our assigned scores may not be so accurate as of the scores by VA experts. On the other hand, due to the research done to justify the assigned scores and the acquired knowledge about VA to develop this work, we consider that our evaluation results are at least somewhat representative of the ones we might expect from experts.

Secondly, our scores were mostly based on our experience on importing a small dataset and building visualizations (**Figure 22 and Figure 23**), which surely did not require the use of all the capabilities from both applications. The small data set size was chosen to make the experiment feasible in a reasonable amount of time. We speculate that some of the findings would only be amplified when working with larger datasets and more complex visualizations. All things considered, we would like to highlight that our main goal with this case study was not to check how good Tableau and Power BI in respect to their features are, instead to detect the presence of functionalities which might be missing considering the criteria set for VA System proposed in this work. And this was possible even with the visualizations we build and the small dataset.

Lastly, Heuristic Evaluation which aims at more precise results should be performed by at least 3 reviewers (Nielsen, 2000). To compensate for our lack of evaluators, we assessed two applications, so that the same criteria might be met and checked in more than one system.

6.1.3.10 Conclusions

We tested our novel evaluation methodology for VA against Tableau and Power BI through a usability test. Both applications obtained similar final scores, **5.9 and 5.85 out of 7**, respectively (see **Table 28** for results). These scores mean an acceptable Visual Analytics Score (see **Table 20** for interpretation of final scores). As we can see, Tableau scored slightly better than Power BI, which means Tableau was comparatively more successful to satisfy the criteria (heuristics and guidelines) set for VA solutions proposed in this work. Tableau and Power BI's capabilities to import, transform data, and to build visualizations, showed us that there are already leading data analytics solutions able to satisfy most of our evaluation criteria (heuristics and guidelines) on the market. This indicates that our evaluation criteria and methodology have feasibility, giving us then positive feedback about the main goal of this use case.

Nonetheless, it is important to mention that some heuristics from the Analytical Process area and one could not have their feasibility checked. First, heuristics **AP22 to AP25** could not have their feasibility verified for not having their related functionalities available in the industry yet (see the justification for heuristic **AP22** in **Table 21**). Secondly, heuristics **AP14, AP26, and AP28,** which relates to collaboration capabilities**,** because we had to perform this study using license versions of Tableau and Power BI that restrict collaboration features. However, such heuristics are highly likely feasible, considering that we found good quality official documentation and tutorials about collaboration features which matches these heuristics for other license types of Tableau and Power BI. Furthermore, criteria **AP27** and **AP29,** which also support collaboration, and **AP32** that relates to creativity, because none of the evaluated tools met these heuristics. Finally, we would like to highlight that Tableau and Power BI would probably have their final scores easily raised to over 6, which means a good VA Score if we had evaluated their licenses which do not limit collaboration capabilities.

7 CONCLUSIONS

In this chapter, we present the conclusions and future work, providing an overview of the problem statement. We also describe our main research contributions and a summarized analysis of the results obtained. We end by discussing the main limitations of our work and directions for further research.

7.1 Study's Overview

We presented a novel value-driven methodology for evaluation of systems considering criteria related to desirable functionalities in VA systems, capturing theoretical and practical state-of-the-art in the VA field. This included the research on heuristics and guidelines for VA and its related fields for the creation of a novel set of heuristics and guidelines. Our value-driven evaluation framework addresses seven evaluation areas, namely Analytic Process, Visualization, Interactivity, User-friendliness, Satisfaction, Error-handling, and Adequacy.

To show the practical relevance of our framework and the feasibility of the new heuristics and guidelines set, we validated our assessment method against two top commercial applications from the Business Intelligence (BI) context. The proposed evaluation model hopes to be as holistic as possible regarding the VA Software. For this reason, we proposed criteria intended to go beyond usability. Moreover, it gives more weight to heuristics on sensemaking, interactions and visualization, which are the three core elements for VA systems.

7.2 Research Contributions

First, in search of hands-on experience and initial knowledge regarding Visual Analytics applications, we developed two works about Self-service Business Intelligence and Analytics solutions which resulted in two published papers (**Appendix A and B**).

Secondly, this work provides a practical framework to assess VA solutions based on a novel and organized set of heuristics and guidelines along with deployment instructions. Also, this set of heuristics and guidelines can be used as an information source when designing and developing VA Software. Moreover, our work can help to better understand the Visual Analytics Approach, Heuristic Evaluation, and how to create new heuristics. Finally, through an alternative definition about VA systems, it assists in easily differentiating VA and DV solutions to minimize confusion in their identification.

7.3 General Findings

Evaluation of VA systems is a challenge due to its multidisciplinarity. Despite the extensive research, it was found no heuristic set specific for a holistic evaluation of VA systems. This reiterates a situation reported in 2011 when the VA field had already five years of existence.

Heuristic Evaluation (HE) is one of the most popular ways to evaluate systems, once it is considered a cost-effective, intuitive, and easy to learn method which quickly reviews design issues. There is no standard procedure to create new heuristics. They may be generated from different types of sources such as other heuristics, guidelines, usability problems, and literature review. Therefore, we chose to create heuristics from sets of guidelines and already-validated heuristics from fields related to VA, which initially might seem to be a simple task. However, it turned out to be a very time-consuming and complex task due to the number of heuristics we had to deal with. Along with that, all the research done to properly understand, cluster and merge low and high-levels heuristics from a significant number of studies. Also, from authors who sometimes use terms interchangeably when phrasing their heuristics.

As for the adopted heuristics and guidelines which go beyond usability according to the areas proposed by Scholtz (2006), our study adopted 59 relate to Situation Awareness, 47 to Utility, 46 to Interaction, 30 to Usability, 19 to Collaboration, and 5 to Creativity. Their sum surpasses 136 because several of them belong to more than one area as was explained in **Chapter 4.** This result indicates that our novel heuristic set is likely to cover best the Situation Awareness, Utility, and Interaction areas. That is the areas more related to sensemaking, utility of the environment from the user perspective, and interaction capabilities, respectively. Furthermore, we just found 5 heuristics related to Creativity. This may indicate the need for more criteria to assure creativity in VA systems. We are aware that generalizability of these analyses is limited to the publications from where we extracted heuristics and guidelines as detailed in **Subsection 4.2.2**.

Concerning our proposal of an alternative definition for Visual Analytics systems to help differentiate them from Data Visualization (DV) Software, we could say that in general terms VA Systems are systems able to automatically analyse data and present its analysis in form of interactive visualizations. The interactive capabilities of VA systems should allow the human analyst to see the automated data analysis outcome (visualization) from different perspectives, so that he/she can make sense of data and change parameters in the automated data analysis method. This will allow the analyst to refine the analytic process and obtain further insights from the data being analysed. In short, if a DV software does not employ automatic data analysis with interactive visualizations, this software does not fit into the Visual Analytics field.

From the results and conclusions of our methodology use case and presented in **Chapter 6**, which evaluated two top Business Intelligence systems, we can suggest, considering the final scores obtained by both evaluated solutions, that we already have in the industry Business Analytics software able to satisfy most of the criteria which reflect desirable functionalities and characteristics of VA systems. The methodology use case also resulted in a positive feedback about the heuristics and guidelines feasibility, once nearly 83% of them could be satisfied by the evaluated applications.

7.4 Study Limitations

There are limitations to this research that must be addressed. First, due to the diverse publication venues available to VA researchers, extracting all the papers in the field would have been something difficult. There is also a risk that relevant papers may have been omitted due to our choice of keywords and search strings. However, we are confident that our framework is relevant because we did an exhaustive search for heuristics and guidelines and adopted the ones which were validated, a result of merge from validated ones, or from studies of relevance for the VA field, giving so strength of evidence to the criteria used in the proposed methodology.

The hierarchical structure of the proposed methodology is segmented into 7 evaluation areas with different weights. But the allocation of the guidelines and heuristics to the areas has some subjectivity in it. However, we believe that the subjectivity is inherent to evaluating the overall value of an application and is therefore a part of this methodology.

We followed the hypotheses that the number of heuristics should not be large, to require less cognitive effort from the reviewers when applying HE. Hence, we reduced 136 heuristics and guidelines into a smaller set of 59 and for that, we used a method which attributes a similarity grade among the elements being compared. Thus, this similarity may be biased because only the author of this work defined them.

The way as the new heuristics and guidelines were phrased may give space to different interpretations. To mitigate that, we would need to survey their understanding through expert evaluators. On the other hand, each proposed heuristic has its respective guideline, which is supposed to work as a longer or alternative version of its heuristic, clarifying so further the heuristic meaning.

As for the weights attributed to each evaluation area in the evaluated methodology, they are an initial proposal which favours the areas Analytic Process, Interactivity, and Visualization because they characterize best the VA approach. However, we are aware that the evaluation methodology might give slightly different results if not only areas had different weights, but also individual heuristics.

7.5 Directions for Future Research

As we future work, we intend to submit our evaluation framework and its criteria (heuristics and guidelines) to the critical review of Visual Analytics Experts, in order to refine the heuristics. Experts have knowledge (personal or professional) about the area, and/or usability, and/or system interfaces. Thus, we can survey the understanding regarding the phrasing of our heuristics and guidelines. Moreover, we can further assure the reliability of the heuristics and guidelines by asking experts to rate reliability on the heuristics and guidelines. Also, we plan to mitigate bias regarding the similarity grades used for the heuristic reduction through brainstorm sessions with the experts. These validations studies will take place with at least 3 expert evaluators.

We also intend to apply the evaluation methodology in a larger number of commercial applications from diverse contexts to confirm further the feasibility of the new heuristics and guidelines of the evaluation model. Also, to observe whether VA Applications from other contexts are so aligned to the VA approach as BI applications. Thus, we can have a sharper picture of the presence of VA in the industry.

As this evaluation framework is intended to VA systems in general, we plan to create additional sets of heuristics and guidelines which apply only to specific contexts of VA systems such as Predictive Analysis or and Healthcare. This will likely allow better evaluation support to such types of systems.

Furthermore, we plan to keep updating this evaluation framework with new heuristics and guidelines extracted from further advances in the VA field and related ones.

To conclude, our evaluation framework is a model whose future will be surely characterized by additions, corrections, and further evaluations. It might be considered a first assessment approach to holistically evaluate VA systems in more detail, which will be updated according to the advances for Visual Analytics. We cannot affirm it is complete, but we believe to be on the right path because it features criteria to inspect more carefully VA systems in their sensemaking, visualization, and interaction capabilities. Finally, we expect our work will lead to an impact on significantly reducing ambiguity concerning the VA Approach and on improving the design and development of VA systems.

REFERENCES

ACM Digital Library. (2020). Retrieved 4 October 2020, from https://dl.acm.org/

Adagha, O., Levy, R. M., & Carpendale, S. (2017). Towards a product design assessment of visual analytics in decision support applications: a systematic review. *Journal of Intelligent Manufacturing*, *28*(7), 1623–1633. https://doi.org/10.1007/s10845-015-1118-5

Adair, B. (2020). Best Visual Analytics Tools For Business Intelligence In 2020. Retrieved 12 October 2020, from https://www.selecthub.com/business-intelligence/visual-analytics-tools/

Amar, R., & Stasko, J. (2004). BEST PAPER: A Knowledge Task-Based Framework for Design and Evaluation of Information Visualizations. *IEEE Symposium on Information Visualization*, 143–150. IEEE. https://doi.org/10.1109/INFVIS.2004.10

Amar, R., & Stasko, J. (2005). Knowledge precepts for design and evaluation of information visualizations. *IEEE Transactions on Visualization and Computer Graphics*, *11*(4), 432–442. https://doi.org/10.1109/TVCG.2005.63

André, P., Schraefel, C., Teevan, J., & Dumais, S. T. (2009). *Discovery Is Never by Chance: Designing for (Un)Serendipity*.

Andrienko, G., Andrienko, N., Keim, D., MacEachren, A., & Wrobel, S. (2011). Challenging problems of geospatial visual analytics. *J. Vis. Lang. Comput.*, *22*, 251–256. https://doi.org/10.1016/j.jvlc.2011.04.001

Antonopoulou, M., & Kotsilieris, T. (2019). A Literature Review of User Satisfaction Models Towards Information System Success. *International Journal of E-Services and Mobile Applications*, *11*, 71. https://doi.org/10.4018/IJESMA.2019040105

Attfield, S., Hara, S., & Wong, B. L. (2010). *Sensemaking in Visual Analytics: Processes and Challenges*.

Bailey, J. E., & Pearson, S. W. (1983). Development of a Tool for Measuring and Analyzing Computer User Satisfaction. *Management Science*, *29*(5), 530–545. Retrieved from http://www.jstor.org/stable/2631354

Barbulescu, A., Stoica, F., & Stoica, L. (2016). *Sensemaking with Interactive Data Visualization*.

Batch, A., & Elmqvist, N. (2018). The Interactive Visualization Gap in Initial Exploratory Data Analysis. *IEEE Transactions on Visualization and Computer Graphics*, *24*(1), 278–287. https://doi.org/10.1109/TVCG.2017.2743990

BELIV. (2020). Retrieved 20 June 2020, from https://beliv-workshop.github.io/

Bian, Y., Dowling, M., & North, C. (2020). *Evaluating Semantic Interaction on Word Embeddings via Simulation*. Retrieved from http://arxiv.org/abs/2007.15824

Blum, S., Cetin, G., & Stuerzlinger, W. (2019). Immersive analytics sensemaking on different platforms. *Journal of WSCG*, *2019*(WSCG2019CSRN2901), 69–80. https://doi.org/10.24132/CSRN.2019.2902.2.9

Brooks, D. C. (2012). Retrieval Cues and Learning. In N. M. Seel (Ed.), *Encyclopedia of the Sciences of Learning* (pp. 2859–2862). Boston, MA: Springer US. https://doi.org/10.1007/978-1-4419-1428-6_1032

Chinchor, N., Cook, K. A., & Scholtz, J. (2012). *Building Adoption of Visual Analytics Software*. https://doi.org/10.1007/978-1-4471-2804-5_29

Chopra, A., & Samant, R. M. (2011). Overview of visual analytics. *Proceedings of the International Conference & Workshop on Emerging Trends in Technology - ICWET '11*, 1358. New York, New York, USA: Association for Computing Machinery (ACM). https://doi.org/10.1145/1980022.1980329

Cook, K. A., Cramer, N., Israel, D., Wolverton, M., Bruce, J., Burtner, R., & Endert, A. (2015). Mixed-initiative visual analytics using task-driven recommendations. *2015 IEEE Conference on Visual Analytics Science and Technology (VAST)*, 9–16. IEEE. https://doi.org/10.1109/VAST.2015.7347625

Corrao, N. J., Robinson, A. G., Swiernik, M. A., & Naeim, A. (2010). Importance of Testing for Usability When Selecting and Implementing an Electronic Health or Medical Record System. *Journal of Oncology Practice*, *6*(3), 120–124. https://doi.org/10.1200/JOP.200017

Cui, W. (2019). Visual Analytics: A Comprehensive Overview. *IEEE Access*, *7*(June), 81555–81573. https://doi.org/10.1109/ACCESS.2019.2923736

Dasgupta, A., Lee, J.-Y., Wilson, R., Lafrance, R. A., Cramer, N., Cook, K. A., & Payne, S. (2017). Familiarity Vs Trust: A Comparative Study of Domain Scientists' Trust in Visual Analytics and Conventional Analysis Methods. *IEEE Transactions on Visualization and Computer Graphics*, *23*(1), 271–280. https://doi.org/10.1109/TVCG.2016.2598544

DataFlair. (2020). Incredibly Easy Method to Create Power BI Dashboard - Best Tutorial Ever! - DataFlair. Retrieved 17 October 2020, from https://data-flair.training/blogs/power-bi-dashboard/

dblp: computer science bibliography. (2020). Retrieved 4 October 2020, from https://dblp.uni-trier.de/

Digimasters. (2020). Digimasters | Visual Analytics, Virtual Reality and Digital Consulting. Retrieved 4 May 2020, from http://digimasters.co.uk/blog_details.php?ID=2

Dill, J., Earnshaw, R., Kasik, D., Vince, J., & Wong, P. (2012). *Expanding the Frontiers of Visual Analytics and Visualization*. https://doi.org/10.1007/978-1-4471-2804-5

Edge, D., Larson, J., & White, C. (2018). Bringing ai to bi: Enabling visual analytics of unstructured data in a modern business intelligence platform. *Conference on Human Factors in Computing Systems - Proceedings*, *2018-April*. https://doi.org/10.1145/3170427.3174367

Endert, A. (2012). Semantic Interaction for Sense-making: Alex Endert at TEDxVirginiaTech - YouTube. Retrieved 29 August 2020, from https://www.youtube.com/watch?v=bHNAEeXP-jQ

Endert, A. (2014). Semantic Interaction for Visual Analytics: Toward Coupling Cognition and Computation. *IEEE Computer Graphics and Applications*, *34*(4), 8–15. https://doi.org/10.1109/MCG.2014.73

Endert, A. (2016). Semantic Interaction for Visual Analytics: Inferring Analytical Reasoning for Model Steering. *Synthesis Lectures on Visualization*, *4*(2), 1–99. https://doi.org/10.2200/S00730ED1V01Y201608VIS007

Endert, A., Chang, R., North, C., & Zhou, M. (2015). Semantic Interaction: Coupling Cognition and Computation through Usable Interactive Analytics. *IEEE Computer Graphics and Applications, 35*(4), 94–99. https://doi.org/10.1109/MCG.2015.91

Endert, A., Fiaux, P., & North, C. (2012). Semantic Interaction for Sensemaking: Inferring Analytical Reasoning for Model Steering. *IEEE Transactions on Visualization and Computer Graphics*, *18*(12), 2879–2888. https://doi.org/10.1109/TVCG.2012.260

Endert, A., Hossain, M. S., Ramakrishnan, N., North, C., Fiaux, P., & Andrews, C. (2014). The human is the loop: new directions for visual analytics. *Journal of Intelligent Information Systems*, *43*(3), 411–435. https://doi.org/10.1007/s10844-014-0304-9

Endert, A., & North, C. (2011). *Spatial Interaction: Exploring Interaction within the Spatial Metaphor*. Retrieved from http://www.cc.gatech.edu/~aendert3/resources/HCIC2011 Boaster Endert.pdf

Endert, A., North, C. L., Bowman, D. A., Leman, S. C., May, R. A., & Quek, F. (2012). *Semantic Interaction for Visual Analytics: Inferring Analytical Reasoning for Model Steering*.

Endsley, M. (2000). Theoretical underpinnings of situation awareness: A critical review. *Situation Awareness Analysis and Measurement*, 3–32.

Fekete, J., Jankun-Kelly, T. J., Tory, M., & Xu, K. (2019). Provenance Analysis for Sensemaking. *IEEE Computer Graphics and Applications*, *39*(6), 27–29. https://doi.org/10.1109/MCG.2019.2945378

Folio3. (2019). Difference Between Power BI Pro vs. Power BI Free vs. Power BI Premium. Retrieved 23 January 2020, from https://dynamics.folio3.com/blog/difference-between-power-bi-pro-vs-free-vs-premium/

Forsell, C. (2010). A Guide to Scientific Evaluation in Information Visualization. *2010 14th International Conference Information Visualisation*, 162–169. IEEE. https://doi.org/10.1109/IV.2010.33

Forsell, C. (2012). Evaluation in Information Visualization: Heuristic Evaluation. *2012 16th International Conference on Information Visualisation*, 136–142. https://doi.org/10.1109/IV.2012.33

Forsell, C., & Johansson, J. (2010). An Heuristic Set for Evaluation in Information Visualization. *Proceedings of the International Conference on Advanced Visual Interfaces*, 199–206. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/1842993.1843029

Freitas, C. M. D. S., Luzzardi, P. R. G., Cava, R. A., Winckler, M. A. A., Pimenta, M. S., & Nedel, L. P. (2002). Evaluating Usability of Information Visualization Techniques. *Proceedings of 5th Symposium on Human Factors in Computer Systems (IHC 2002)*, 40–51. Retrieved from http://www.irit.fr/recherches/IHCS/winckler/projects/spider/publications/IHC2002-FreitasEtAl.pdf

Gartner. (2019). *Magic Quadrant for Analytics and Business Intelligence Platforms*.

Gartner. (2020). Self-service Analytics. Retrieved 11 January 2020, from https://www.gartner.com/en/information-technology/glossary/self-service-analytics

GeeksforGeeks. (2018). Star Schema in Data Warehouse modeling - GeeksforGeeks. Retrieved 25 October 2020, from https://www.geeksforgeeks.org/star-schema-in-data-warehouse-modeling/

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic Decision Making. *Annual Review of Psychology*, *62*, 451–482. https://doi.org/10.1146/annurev-psych-120709-145346

Google Scholar. (2020). Retrieved 4 October 2020, from https://scholar.google.com/

Greensmith, T., Ribarsky, W., & Fisher, B. (2009). Building and Applying a Human Cognition Model for Visual Analytics. *Information Visualization*, *8*, 1–13. https://doi.org/10.1057/ivs.2008.28

Greitzer, F. L., Noonan, C., & Franklin, L. (2011). *Cognitive Foundations for Visual Analytics*. (February).

Heer, J., & Agrawala, M. (2008). Design Considerations for Collaborative Visual Analytics. *Information Visualization*, *7*(1), 171–178. https://doi.org/10.1057/palgrave.ivs.9500167

Holton, G. (1988). *Thematic Origins of Scientific Thought* (2nd ed.). Cambridge, MA: Harvard University Press.

Howson, A. C., Richardson, J., Sallam, R., & Kronz, A. (2019). Magic Quadrant for Analytics and Business Intelligence Platforms. Retrieved 19 February 2020, from https://www.gartner.com/doc/reprints?id=1-68720FP&ct=190213&st=sb

IDF. (2019). What is Accessibility? | Interaction Design Foundation. Retrieved 7 December 2020, from Interaction Design Foundation website: https://www.interaction-design.org/literature/topics/ease-of-use

IEEE VIS. (2020). Retrieved 4 October 2020, from http://ieeevis.org/year/2020/welcome

IEEE Xplore. (2020). Retrieved 4 October 2020, from https://ieeexplore.ieee.org/Xplore/home.jsp

Ignatius, E., & Senay, H. (1996). *Visualization Concepts*. 1–4. Retrieved from https://www.cs.rit.edu/usr/local/pub/ncs/hypervis/concepts/marks.htm

Infinity. (2020). Visual Analytics – Key Attributes, Scope and Advantages. Retrieved 16 May 2020, from https://infinitylimited.co.uk/visual-analytics-key-attributes-scope-and-advantages/

Insider, B. D. (2017). Was ist Self-Service BI (Business Intelligence)? Retrieved 18 March 2020, from https://www.bigdata-insider.de/was-ist-self-service-bi-business-intelligence-a-659722/

Islam, M., & Jin, S. (2019). An Overview of Data Visualization. *2019 International Conference on Information Science and Communications Technologies (ICISCT)*, 1–7. https://doi.org/10.1109/ICISCT47635.2019.9012031

ISO. (2018). Ergonomics of human-system interaction - Part 11: Usability: Definitions and Concepts. *Iso 9241-11:2018(E)*, p. 2. Retrieved from https://www.iso.org/standard/63500.html

Jaferian, P., Hawkey, K., Sotirakopoulos, A., Velez-Rojas, M., & Beznosov, K. (2011). Heuristics for Evaluating IT Security Management Tools. *Proceedings of the Seventh Symposium on Usable Privacy and Security*. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/2078827.2078837

Jiménez, C., Rusu, C., Roncagliolo, S., Inostroza, R., & Rusu, V. (2012). Evaluating a methodology to establish usability heuristics. *Proceedings - International Conference of the Chilean Computer Science Society, SCCC*, 51–59. https://doi.org/10.1109/SCCC.2012.14

Jordan, P. W. (2020). An Introduction to Usability. In *An Introduction to Usability*. https://doi.org/10.1201/9781003062769

Joshi, A., Kale, S., Chandel, S., & Pal, D. (2015). Likert Scale: Explored and Explained. *British Journal of Applied Science & Technology*, *7*(4), 396–403. https://doi.org/10.9734/bjast/2015/14975

Kang, Y., Gorg, C., & Stasko, J. (2009). Evaluating visual analytics systems for investigative analysis: Deriving design principles from a case study. *2009 IEEE Symposium on Visual Analytics Science and Technology*, 139–146. IEEE. https://doi.org/10.1109/VAST.2009.5333878

Kang, Y., Görg, C., & Stasko, J. (2011). How Can Visual Analytics Assist Investigative Analysis? Design Implications from an Evaluation. *IEEE Transactions on Visualization and Computer Graphics*, *17*(5), 570–583. https://doi.org/10.1109/TVCG.2010.84

Kang, Y., & Stasko, J. (2012). Examining the Use of a Visual Analytics System for Sensemaking Tasks: Case Studies with Domain Experts. *IEEE Transactions on Visualization and Computer Graphics*, *18*(12), 2869–2878. https://doi.org/10.1109/TVCG.2012.224

Keim, D., Andrienko, G., Fekete, J.-D., Görg, C., Kohlhammer, J., & Melançon, G. (2008). Visual Analytics: Definition, Process, and Challenges. In *Information Visualization* (pp. 154–175). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-70956-5_7

Keim, D., Kohlhammer, J., & Ellis, G. (2010). *Mastering the information age : solving problems with visual analytics* (D. Keim, J. Kohlhammer, G. Ellis, & F. Mansmann, Eds.). Goslar : Eurographics Association. Retrieved from https://diglib.eg.org/handle/10.2312/14803

Keim, D., Mansmann, F., Schneidewind, J., Thomas, J. J., & Ziegler, H. (2008). Visual Analytics: Scope and Challenges. *First Publ. in: Lecture Notes in Computer Science, No. 4404 (2008), Pp. 76-90, 1*. https://doi.org/10.1007/978-3-540-71080-6_6

Keim, D., Mansmann, F., Stoffel, A., & Ziegler, H. (2009). Visual Analytics Visual Analysis ; Visual Data Analysis ; Visual Data Mining. *Iwaw*, 27–37.

Keim, D., Mansmann, F., & Thomas, J. J. (2010). Visual Analytics : How Much Visualization and How Much Analytics ? *ACM SIGKDD Explorations Newsletter*, *11*(2), 5–8. https://doi.org/10.1145/1809400.1809403

Kerren, A., & Schreiber, F. (2012). Toward the role of interaction in Visual Analytics. *Proceedings of the 2012 Winter Simulation Conference (WSC)*, 1–13. https://doi.org/10.1109/WSC.2012.6465208

Khalili, A., Van Andel, P., Van Den Besselaar, P., & Andries De Graaf, K. (2017). *Fostering Serendipitous Knowledge Discovery using an Adaptive Multigraph-based Faceted Browser*. https://doi.org/10.1145/nnnnnnn.nnnnnnn

Khayat, M., Karimzadeh, M., Ebert, D. S., & Ghafoor, A. (2019). The Validity, Generalizability and Feasibility of Summative Evaluation Methods in Visual Analytics. *IEEE Transactions on Visualization and Computer Graphics*, *26*(1), 1–1. https://doi.org/10.1109/TVCG.2019.2934264

Knaflic, C. N. (2015). *Storytelling with data : a data visualization guide for business professionals /*. John Wiley & Sons, Inc.

Kumar, A. (2019). Comparison of Tableau Vs QlikView | QlikView Tableau Differences. Retrieved 22 January 2020, from https://www.knowledgehut.com/blog/business-intelligence-and-visualization/tableau-vs-qlikview

Laubheimer, P. (2015a). Preventing User Errors: Avoiding Conscious Mistakes. Retrieved 7 November 2020, from https://www.nngroup.com/articles/user-mistakes/

Laubheimer, P. (2015b). Preventing User Errors: Avoiding Unconscious Slips. Retrieved 7 November 2020, from https://www.nngroup.com/articles/slips/

Lowry, S. Z., Ramaiah, M., Taylor, S., Patterson, E. S., Prettyman, S. S., Simmons, D., … Gibbons, M. C. (2015). *Technical Evaluation, Testing, and Validation of the Usability of Electronic Health Records: Empirically Based Use Cases for Validating Safety-Enhanced Usability and Guidelines for Standardization NISTIR 7804-1*. https://doi.org/10.6028/NIST.IR.7804-1

Mankoff, J., Dey, A. K., Hsieh, G., Kientz, J., Lederer, S., & Ames, M. (2003). Heuristic Evaluation of Ambient Displays. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 169–176. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/642611.642642

Marache-Francisco, C., & Brangier, E. (2016). *Validation of a Gamification Design Guide: Does a Gamification Booklet Help UX Designers to Be More Creative? 9747*, 284–293. https://doi.org/10.1007/978-3-319-40355-7_27

McLeod, S. (2019). Likert Scale Definition, Examples and Analysis | Simply Psychology. *Simply Psychology*. Retrieved from https://www.simplypsychology.org/likert-scale.html

Mehrotra, C., Chitransh, N., & Singh, A. (2008). Visual Analytics: Scope and Challenges Daniel. *Proceeding - IEEE International Conference on Computing, Communication and Automation, ICCCA 2017*, Vol. 2017-Janua, pp. 1229–1234. https://doi.org/10.1109/CCAA.2017.8229987

Microsoft. (2020a). Data sources in Power BI Desktop - Power BI | Microsoft Docs. Retrieved 22 January 2020, from https://docs.microsoft.com/en-us/power-bi/desktop-data-sources

Microsoft. (2020b). Feature availability for users with free licenses - Power BI | Microsoft Docs. Retrieved 15 October 2020, from https://docs.microsoft.com/en-us/power-bi/consumer/end-user-features

Microsoft. (2020c). Power BI documentation - Power BI | Microsoft Docs. Retrieved 23 October 2020, from https://docs.microsoft.com/en-us/power-bi/

Microsoft. (2020d). Visualização de Dados | Microsoft Power BI. Retrieved 5 December 2020, from https://powerbi.microsoft.com/pt-pt/

Moore, G. A., & McKenna, R. (1999). *Crossing the Chasm: Marketing and Selling High-tech Products to Mainstream Customers*. HarperBusiness. Retrieved from https://books.google.de/books?id=cj7EQgAACAAJ

Mott, E. (2020). Why Is it Not Feasible for Companies to Produce an Error-Free Software? Retrieved 7 December 2020, from https://smallbusiness.chron.com/not-feasible-companies-produce-errorfree-software-59309.html

Munzner, T. (2009). A Nested Model for Visualization Design and Validation. *IEEE Transactions on Visualization and Computer Graphics*, *15*(6), 921–928. https://doi.org/10.1109/TVCG.2009.111

Nielsen, J. (1993). *Usability Engineering*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.

Nielsen, J. (1995a). 10 Heuristics for User Interface Design. Retrieved 18 July 2020, from Nielsen Norman Group website: https://www.nngroup.com/articles/ten-usability-heuristics/

Nielsen, J. (1995b). Heuristic Evaluation: How-To: Article by Jakob Nielsen. *Nielson Norman GroupNorman*, (Nielsen 1992), 1–11. Retrieved from https://www.nngroup.com/articles/how-to-conduct-a-heuristic-evaluation/

Nielsen, J. (1999). Do Interface Standards Stifle Design Creativity. *NN Group*, 1–3. Retrieved from https://www.nngroup.com/articles/do-interface-standards-stifle-design-creativity/

Nielsen, J. (2000). Why You Only Need to Test with 5 Users. https://doi.org/http://www.useit.com/alertbox/20000319.html

Nielsen, J. (2010). CHAPTER 1 - What Is Usability? In C. Wilson (Ed.), *User Experience Re-Mastered* (pp. 3–22). Boston: Morgan Kaufmann. https://doi.org/https://doi.org/10.1016/B978-0-12-375114-0.00004-9

Nielsen, J., & Landauer, T. K. (1993). A Mathematical Model of the Finding of Usability Problems. *Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems*, 206–213. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/169059.169166

Nielsen, J., & Mack, R. L. (1994). *Usability Inspection Methods*. New York, NY: John Wiley & Sons.

Nielsen, J., & Molich, R. (1990). Heuristic Evaluation of User Interfaces. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 249–256. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/97243.97281

Oliveira, M. R. de, & Silva, C. G. da. (2017). Adapting Heuristic Evaluation to Information Visualization - A Method for Defining a Heuristic Set by Heuristic Grouping. *Proceedings of the 12th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, 225–232. SCITEPRESS - Science and Technology Publications. https://doi.org/10.5220/0006133202250232

Pacheco, B. de A., & Souza-Concilio, I. de A. (2014). Externalist philosophy and interaction: Proposal of heuristics for the design of interactive systems. *2014 Science and Information Conference*, 702–708. https://doi.org/10.1109/SAI.2014.6918264

Panoho, K. (2019). Council Post: The Age Of Analytics And The Importance Of Data Quality. Retrieved 3 May 2020, from https://www.forbes.com/sites/forbesagencycouncil/2019/10/01/the-age-of-analytics-and-the-importance-of-data-quality/#205abd1b5c3c

Papaloukas, S., Patriarcheas, K., & Xenos, M. (2009). Usability Assessment Heuristics in New Genre Videogames. *2009 13th Panhellenic Conference on Informatics*, 202–206. https://doi.org/10.1109/PCI.2009.14

Pat Research. (2019a). About PAT RESEARCH - PAT RESEARCH: B2B Reviews, Buying Guides & Best Practices. Retrieved 14 December 2019, from https://www.predictiveanalyticstoday.com/about/

Pat Research. (2019b). How to Select the Best Self Service Analytics Software for Your Business - Compare Reviews, Features, Pricing in 2019 - PAT RESEARCH: B2B Reviews, Buying Guides & Best Practices. Retrieved 25 November 2019, from https://www.predictiveanalyticstoday.com/self-service-analytics-software/

Patterson, R. E., Blaha, L. M., Grinstein, G. G., Liggett, K. K., Kaveney, D. E., Sheldon, K. C., … Moore, J. A. (2014). A human cognition framework for information visualization. *Computers & Graphics*, *42*, 42–58. https://doi.org/https://doi.org/10.1016/j.cag.2014.03.002

Plaisant, C., Fekete, J., & Grinstein, G. (2008a). Promoting Insight-Based Evaluation of Visualizations: From Contest to Benchmark Repository. *IEEE Transactions on Visualization and Computer Graphics*, *14*(1), 120–134. https://doi.org/10.1109/TVCG.2007.70412

Plaisant, C., Grinstein, G., Scholtz, J., Whiting, M., O’Connell, T., & Laskowski, S. (2008b). Evaluating visual analytics at the 2007 VAST symposium contest. *IEEE on Computer Graphics and Applications*, *28*. https://doi.org/10.1109/MCG.2008.27

PNNL. (2020). Visual Analytics. Retrieved 4 October 2020, from https://www.pnnl.gov/computing/visual_analytics/

Pohl, M., Smuc, M., & Mayr, E. (2012). The User Puzzle—Explaining the Interaction with Visual Analytics Systems. *IEEE Transactions on Visualization and Computer Graphics*, *18*(12), 2908–2916. https://doi.org/10.1109/TVCG.2012.273

Pribeanu, C. (2017). A Revised Set of Usability Heuristics for the Evaluation of Interactive Systems. *Informatica Economica*, *21*, 31–38. https://doi.org/10.12948/issn14531305/21.3.2017.03

Qlik. (2020a). Released QlikView versions including build number. Retrieved 13 February 2020, from https://support.qlik.com/articles/000003216

Qlik. (2020b). System Requirements for QlikView – QlikView. Retrieved 5 April 2020, from https://help.qlik.com/en-US/qlikview/April2019/Content/QV_HelpSites/System-requirements.htm

Quinones, D., & Rusu, C. (2017). How to develop usability heuristics: A systematic literature review. *Computer Standards and Interfaces*, *53*, 89–122. https://doi.org/10.1016/j.csi.2017.03.009

Quinones, D., Rusu, C., & Roncagliolo, S. (2014). Redefining Usability Heuristics for

Transactional Web Applications. *2014 11th International Conference on Information Technology: New Generations*, 260–265. https://doi.org/10.1109/ITNG.2014.46

Sailer, M., Hense, J. U., Mayr, S. K., & Mandl, H. (2017). How gamification motivates: An experimental study of the effects of specific game design elements on psychological need satisfaction. *Computers in Human Behavior*, *69*, 371–380. https://doi.org/https://doi.org/10.1016/j.chb.2016.12.033

Santos, B. S., Silva, S., & Dias, P. (2018). Heuristic Evaluation in Visualization: An Empirical Study : Position paper. *2018 IEEE Evaluation and Beyond - Methodological Approaches for Visualization (BELIV)*, 78–85. https://doi.org/10.1109/BELIV.2018.8634108

Scapin, D. L., & Bastien, J. M. C. (1997). Ergonomic criteria for evaluating the ergonomic quality of interactive systems. *Behaviour & Information Technology*, *16*(4–5), 220–231. https://doi.org/10.1080/014492997119806

Scholtz, J. (2006). Beyond usability: Evaluation aspects of visual analytic environments. *IEEE Symposium on Visual Analytics Science and Technology 2006, VAST 2006 - Proceedings*, 145–150. https://doi.org/10.1109/VAST.2006.261416

Scholtz, J. (2011). Developing guidelines for assessing visual analytics environments. *Information Visualization*, *10*(3), 212–231. https://doi.org/10.1177/1473871611407399

Scholtz, J. (2017). User-Centered Evaluation of Visual Analytics. *Synthesis Lectures on Visualization*, *5*(1), i–71. https://doi.org/10.2200/s00797ed1v01y201709vis009

Scholtz, J., Plaisant, C., Whiting, M., & Grinstein, G. (2014). Evaluation of visual analytics environments: The road to the Visual Analytics Science and Technology challenge evaluation methodology. *Information Visualization*, *13*(4), 326–335. https://doi.org/10.1177/1473871613490290

Schütz, L., Raabe, S., Bade, K., & Pietsch, M. (2017). Using Visual Analytics for Decision Making. *Journal of Digital Landscape Architecture*, *2*, 94–101. https://doi.org/10.14627/537629010

Scipy Community. (2020a). scipy.cluster.hierarchy.linkage — SciPy v1.5.2 Reference Guide. Retrieved 15 September 2020, from https://docs.scipy.org/doc/scipy/reference/generated/scipy.cluster.hierarchy.linkage.html

Scipy Community. (2020b). scipy.spatial.distance.pdist — SciPy v1.5.2 Reference Guide. Retrieved 15 September 2020, from https://docs.scipy.org/doc/scipy/reference/generated/scipy.spatial.distance.pdist.html

Self, J. Z., Vinayagam, R. K., Fry, J. T., & North, C. (2016). Bridging the gap between user intention and model parameters for human-in-the-loop data analytics. *HILDA 2016 - Proceedings of the Workshop on Human-In-the-Loop Data Analytics*. https://doi.org/10.1145/2939502.2939505

Services, U. S. D. of H. & H. (2020). About Us | Usability.gov. Retrieved 7 June 2020, from https://www.usability.gov/about-us/index.html

Sharma, K. (2017). Data Visualization and Visual Analytics Are Not the Same - Know the Difference - Ascentt. Retrieved 3 May 2020, from http://www.ascentt.com/data-visualization-and-visual-analytics-are-not-the-same-know-the-difference/

Shneiderman, B. (1996). Eyes have it: a task by data type taxonomy for information visualizations. *IEEE Symposium on Visual Languages, Proceedings*, 336–343. https://doi.org/10.1109/vl.1996.545307

Snopek, L. (2013). Heuristics and Cognitive Biases. In *The Complete Guide to Portfolio Construction and Management* (pp. 153–159). https://doi.org/10.1002/9781118467220.ch21

Somervell, J., & McCrickard, D. S. (2005). Better discount evaluation: illustrating how critical parameters support heuristic creation. *Interacting with Computers*, *17*(5), 592–612. https://doi.org/10.1016/j.intcom.2005.03.007

Springer. (2020). Retrieved 4 October 2020, from https://link.springer.com/

Tableau. (2004). 1.0 | Tableau Software. Retrieved 18 March 2020, from https://www.tableau.com/fast-pace-innovation/1.0

Tableau. (2016). Tableau Public Frequently Asked Questions |Tableau Community Forums. Retrieved 5 January 2020, from https://community.tableau.com/docs/DOC-9135#subhead20

Tableau. (2020a). How Explain Data Works - Tableau. Retrieved 16 October 2020, from https://help.tableau.com/current/pro/desktop/en-us/explain_data_explained.htm

Tableau. (2020b). Now available in Tableau: View Recommendations, table improvements, Webhooks support, and more. Retrieved 15 October 2020, from https://www.tableau.com/about/blog/2019/11/available-today-table-improvements-view-recommendations-webhooks

Tableau. (2020c). Tableau Help | Tableau Software. Retrieved 19 October 2020, from https://www.tableau.com/support/help

Tableau. (2020d). What is data visualization? A definition, examples, and resources. Retrieved 22 January 2020, from https://www.tableau.com/learn/articles/data-visualization

Tarrell, A., Fruhling, A., Borgo, R., Forsell, C., Grinstein, G., & Scholtz, J. (2014). Toward Visualization-Specific Heuristic Evaluation. *Proceedings of the Fifth Workshop on Beyond Time and Errors: Novel Evaluation Methods for Visualization*, 110–117. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/2669557.2669580

Techtarget. (2016). What is self-service business intelligence (BI)? - Definition from WhatIs.com. Retrieved 10 January 2020, from https://searchbusinessanalytics.techtarget.com/definition/self-service-business-intelligence-BI

Thomas, J. J. (2009). *Visual Analytics Techniques that Enable Knowledge Discovery: Detect the Expected and Discover the Unexpected*. Retrieved from http://nvac.pnl.gov

Thomas, J. J., & Cook, K. A. (2005). Illuminating the Path: The Research and Development Agenda for Visual Analytics. In *IEEE Computer Society*. Retrieved from http://vis.pnnl.gov/pdf/RD_Agenda_VisualAnalytics.pdf

Thomas, J. J., & Cook, K. A. (2006). A visual analytics agenda. *IEEE Computer Graphics and Applications*, *26*(1), 10–13. https://doi.org/10.1109/MCG.2006.5

Tory, M., & Moller, T. (2005). Evaluating visualizations: do expert reviews work? *IEEE*

Computer Graphics and Applications, *25*(5), 8–11. https://doi.org/10.1109/MCG.2005.102

Tropmann-Frick, M., & Andersen, J. S. (2020). Towards Visual Data Science - An Exploration. In T. Ahram, R. Taiar, S. Colson, & A. Choplin (Eds.), *Human Interaction and Emerging Technologies* (pp. 371–377). Cham: Springer International Publishing.

Usability.de. (2020). Self-Descriptiveness in the usability.de Glossary. Retrieved 15 September 2020, from https://www.usability.de/en/usability-user-experience/glossary/self-descriptiveness.html

Väätäjä, H., Varsaluoma, J., Heimonen, T., Tiitinen, K., Hakulinen, J., Turunen, M., … Ihantola, P. (2016). Information Visualization Heuristics in Practical Expert Evaluation. *Proceedings of the Beyond Time and Errors on Novel Evaluation Methods for Visualization - BELIV '16*, 36–43. New York, New York, USA: ACM Press. https://doi.org/10.1145/2993901.2993918

Van Wijk, J. J. (2013). Evaluation: A challenge for visual analytics. *Computer*, *46*(7), 56–60. https://doi.org/10.1109/MC.2013.151

VisMaster. (2013). The Visual Analytics Process - VisMaster. Retrieved 30 April 2020, from https://www.vismaster.eu/faq/the-visual-analytics-process/

Visual Analytics Community. (2011). Cognitive Foundations. Retrieved 30 April 2020, from http://vacommunity.org/tiki-index.php

Visual Analytics Community. (2020). About the VAST Challenge. Retrieved 4 October 2020, from http://www.vacommunity.org/About+the+VAST+Challenge

Vredenburg, K., Mao, J.-Y., Smith, P. W., & Carey, T. (2002). A Survey of User-Centered Design Practice. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 471–478. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/503376.503460

W3. (2008). Web Content Accessibility Guidelines (WCAG) 2.0. Retrieved 5 November 2020, from https://www.w3.org/TR/2008/REC-WCAG20-20081211/

W3. (2016). WAI-ARIA Overview | Web Accessibility Initiative (WAI) | W3C. Retrieved 5 November 2020, from https://www.w3.org/WAI/standards-guidelines/aria/

Wall, E., Agnihotri, M., Matzen, L., Divis, K., Haass, M., Endert, A., & Stasko, J. (2019). A Heuristic Approach to Value-Driven Evaluation of Visualizations. *IEEE Transactions on Visualization and Computer Graphics*, *25*(1), 491–500. https://doi.org/10.1109/TVCG.2018.2865146

Wang, X., Bier, E., Butkiewicz, T., Ribarsky, W., & Dou, W. (2011a). Designing Knowledge-assisted Visual Analytics Systems for Organizational Environments. *Proceedings of the 2011 Visual Information Communication - International Symposium on - VINCI '11*, 1–9. New York, New York, USA: ACM Press. https://doi.org/10.1145/2016656.2016657

Wang, X., Dou, W., Butkiewicz, T., Bier, E., & Ribarsky, W. (2011b). A two-stage framework for designing visual analytics system in organizational environments. *2011 IEEE Conference on Visual Analytics Science and Technology (VAST)*, 251–260. IEEE. https://doi.org/10.1109/VAST.2011.6102463

Wasesa, M., Mashuri, M., Handayani, P., & Putro, U. S. (2019). Open Data Visual Analytics to Support Decisions on Physical Investments. *Procedia Computer Science*, *161*, 797–804. https://doi.org/https://doi.org/10.1016/j.procs.2019.11.185

Wasserman, A. I., Guo, X., Mcmillian, B., Qian, K., Wei, M.-Y., & Xu, Q. (2017). *OSSpal: Finding and Evaluating Open Source Software*. https://doi.org/10.1007/978-3-319-57735-7_18

Wikipedia. (2019). Power BI – Wikipédia, a enciclopédia livre. Retrieved 13 February 2020, from https://pt.wikipedia.org/wiki/Power_BI

Wong, E. (2020). Heuristic evaluation: How to conduct a heuristic evaluation. Retrieved 23 November 2020, from Interaction Design Foundation website: https://www.interaction-design.org/literature/article/heuristic-evaluation-how-to-conduct-a-heuristic-evaluation

Wu, A., Zhang, X., & Cai, G. (2010). An interactive sensemaking framework for mobile visual analytics. *VINCI 2010: 3rd Visual Information Communication - International Symposium*. https://doi.org/10.1145/1865841.1865870

Xu, K., Attfield, S., Jankun-Kelly, T. J., Wheat, A., Nguyen, P. H., & Selvaraj, N. (2015). Analytic Provenance for Sensemaking: A Research Agenda. *IEEE Computer Graphics and Applications*, *35*(3), 56–64. https://doi.org/10.1109/MCG.2015.50

Yau, N. (2013). *Data Points: Visualization That Means Something*.

Zuk, T., Schlesier, L., Neumann, P., Hancock, M., & Carpendale, S. (2006). Heuristics for Information Visualization Evaluation. *Proceedings of the 2006 AVI Workshop on BEyond Time and Errors: Novel Evaluation Methods for Information Visualization*, 1–6. https://doi.org/10.1145/1168149.1168162

APPENDIX A – 17TH INTERNATIONAL CONFERENCE ON E-BUSINESS PAPER

Link for the publication **Evaluating Self-Service BI and Analytics Tools for SMEs**:

https://doi.org/10.5220/0009820400890097

Evaluating Self-Service BI and Analytics Tools for SMEs

Antonio Oliveira[1a] and Jorge Bernardino[1,2b]
[1]Polytechnic of Coimbra – ISEC, Rua Pedro Nunes, Quinta da Nora, 3030-199 Coimbra, Portugal
[2]CISUC - Centre of Informatics and Systems of University of Coimbra, Pinhal de Marrocos, 3030-290 Coimbra, Portugal

Keywords: Self-Service Business Intelligence and Analytics, Business, Business Intelligence, Analytics, Metabase, Pentaho Community, Power BI Free, QlikView, Tableau Public.

Abstract: In the vast Data Science domain, Business Intelligence (BI) and Data Analytics are two of the most relevant topics nowadays. Regardless of the data type an end-user needs to work on, data visualization and/or analytics can be a valuable support for a successful decision-making process in management. For that, free and open-source business intelligence and analytics solutions are on the market as an indispensable opportunity for companies to start benefiting from data analytics at no cost. In addition, that task has been currently eased by a group of BI and Analytics tools named as "Self-service", which is an Advanced Analytics topic and designed to enable users with no IT background to perform analyses of data and find business opportunities themselves with minimal or no assistance from IT technicians. Considering that, to help Small and Medium-sized Enterprises (SMEs) decide on a free self-service data tool according to their needs, we compare in this paper, on a functionality basis, 5 popular Self-Service BI and Analytics tools: Metabase, Pentaho Community, Power BI Free, QlikView, and Tableau Public.

1 INTRODUCTION

Self-Service Analytics Software allows business users to analyse data to find business opportunities without an IT background. These technological applications are an approach to advanced analytics. They ensure that users can easily benefit a lot from their business data without necessarily possessing statistical or technological background. Many organizations have recognized the importance of Self-Service Analytics Software, as they have been using these computer tools for their processes (Pat Research, 2019b).

Self-service Analytics belongs to the Business Intelligence (BI) field and empowers line-of-business professionals to build reports and queries on their own with minimal support from IT specialists. Additionally, Self-Service Analytics is represented by BI tools, which have a slight learning curve and offer uncomplicated data access through basic Analytics and simplified underlying data models (Gartner, 2020).

There are many features that makes Self-Service Analytics Software important to organizations, and some of them are: Data Gathering, Filters, Visualizations, Reporting, Collaboration, Data Analysis, Dashboards, Predictive and Real-time Analytics, high ease-of-use for lower-skilled users, Integration of Data, Natural Language Processing, and Security. All that naturally in addition to the software capabilities designed for the high-skilled or more technical users (Pat Research, 2019b).

The most visible benefits of the features mentioned above are predictive power for project future trends and events to plan as soon as possible for their effects. Thus, it is possible to discover business opportunities which stimulates insights that are invisible in not-yet-analysed data. Additionally, urgent issues can be addressed with help of real-time-analysis, and easier access to data about customers, as the business will possess and analyse the data of their customers itself, with no need to wait for industry reports or other third-party sources of data (Bernardino, 2011).

The more end-users can do BI and Analytics on their own by performing, for instance, drill-

[a] https://orcid.org/0000-0002-0078-1033
[b] https://orcid.org/0000-0001-9660-2011

89

This page is intentionally left blank.

APPENDIX B – 20TH PORTUGUESE ASSOCIATION FOR INFORMATION SYSTEMS CONFERENCE PAPER

OSSPal Assessment of Self-Service BI and Analytics Software

Antonio Oliveira, Polytechnic Institute of Coimbra, Coimbra Institute of Engineering – ISEC, Portugal, a21280654@alunos.isec.pt

Jorge Bernardino, Polytechnic Institute of Coimbra, Coimbra Institute of Engineering – ISEC, Applied Research Institute of Polytechnic of Coimbra – i2A, Portugal, jorge@isec.pt

Abstract

Business Intelligence (BI) and Data Analytics are among the top Data Science topics nowadays. They are available as Self-Service solutions of valuable utility when business professionals need to perform data visualization and/or analytics. In addition to that, a great opportunity for companies to start exploring their data with minimal or no assistance from IT technicians. In other words, a shortcut to business opportunities. In this paper, through the OSSPal methodology, we assess the free versions of three popular Self-Service BI and Analytics tools: Power BI, QlikView, and Tableau Public. In conclusion, we could see that Power BI offers more features at no cost, being so highly recommended for Small and Medium-Sized Enterprises (SMEs). On the other hand, QlikView and Tableau Public were considered almost as powerful as Power BI and might also naturally be a more suitable choice according to the requirements of a company.

Keywords: Self-Service Business Intelligence and Analytics; Power BI Free; QlikView; Tableau Public; OSSPal.

1. INTRODUCTION

In the past, when modern companies had not yet so large volumes of data to process and analyse, it was still manageable to make use only of the classic Business Intelligence and Analytics tools to support the decision-making process efficiently. These solutions were efficient but required typically close cooperation with IT specialists to be operated.

Nowadays the amount of data is growing more and more and must be handled in ever shorter times by companies, which may naturally overload their IT personal. From this context emerged then the Self-Service BI and Analytics solutions, allowing business professionals themselves to work, evaluate and visualize data.

All of that through intuitive user interfaces, drag-and-drop menus and low-code environments, to allow user-friendly access to all relevant data and most common analytics capabilities with little or no assistance from Data Science experts. We are interested in open-source software that is freely distributed without any fees related to use. Although the use of the open-source software does not have to involve license costs, the cost of use of any software should be always expressed by sum of all cost related to software implementation, configuration, maintenance and support. Besides the zero licenses cost, open-source software has the following qualities: reliability, customizability, freedom of choice, support and scalability (Bernardino, 2011). Furthermore, freeing up consequently the IT department for more strategic activities (Insider, 2017).

Ideally, training for business professionals would be still recommended but mostly to help them understand what data is available and how to query information to make data-driven decisions to solve business problems (Techtarget, 2016).

Self-Service Analytics is a form of Business Intelligence (BI) in which line-of-business professionals are enabled and encouraged to perform queries and generate reports on their own, with minimal IT support. It is often characterized by simple-to-use BI tools with basic analytic capabilities and an underlying data model that has been simplified or scaled down for ease of understanding and straightforward data access (Gartner, 2020).

Self-service BI and Analytics bring many gains for companies, as it enables Small and Medium-Sized Enterprises (SMEs) to start making data-driven decisions, without acquisition costs (Bernardino & Neves, 2016). Considering that, we have decided to search for the top trend free and open-source analytics tools and evaluate them according to the OSSPal methodology. This, to help SMEs find and adopt the best solution according to their needs. Using OSSPal, quantitative and qualitative measures are combined for evaluating open-source software in several categories, resulting in a quantitative value that allows the comparison between the tools (Wasserman et al., 2017).

By using that methodology, three popular BI and Analytics tools are evaluated: Power BI Free, QlikView, and Tableau Public. These tools will be scored considering the features we considered as fundamental in Self-Service solutions.

The rest of this paper is structured as follows. Section 2 describes the BI and Analytics tools under evaluation. Section 3 explains the OSSPal methodology. Section 4 presents the evaluation through the methodology. Finally, Section 5 presents the conclusions and some future work.

2. BI AND ANALYTICS TOOLS

First, to decide on the Self-Service tools for evaluation, we did an extensive search for the most adopted and best assessed free and open-source tools on websites which ranks Self-Service BI and Analytics Tools. Surprisingly, despite the significant number of tools available on the market, we could observe that the most relevant tools currently are free but not open-source.

To confirm that, we checked the tool rankings provided by renowned Research Companies such as Gartner and Predictive Analytics Today, which base their reviews not only on customers opinions but also on an unbiased methodology (Pat Research, 2019a). From Gartner, we considered its yearly renowned software ranking called Magic Quadrant (Howson, Richardson, Sallam, & Kronz, 2019), which ranks solutions based on a set of critical functionalities and trends on solutions for BI/Analytics tools.

With all this information, we could build our Top 3 of Self-service solutions. As a result, we had then the free versions of Power BI, QlikView, and Tableau for evaluation. It is also worth highlighting that we regarded the functionalities utilized by Gartner on (Howson et al., 2019) and the ones indicated by Predictive Analytics Today on (Pat Research, 2019b), to build our set of essential features to evaluate the tools addressed in this paper.

In the following sections, we describe the main characteristics of each Self-Service BI and Analytics tool. Besides that, some major advantages and limitations of each tool are outlined.

2.1 *Power BI Free*

Power BI is a free desktop proprietary Microsoft platform released in 2011 (Wikipedia, 2019). It works in conjunction with a cloud application that makes possible to publish reports throughout the business. Power BI can only be installed on Windows OS and is updated every month. It is intended for small to midsize organizations.

It has the same rich visualizations and filters as the paid version, including a natural language question and answering functionality. Additionally, it saves, uploads and publishes reports to the web with a limit of 10 GB per user. Its other two types of license, Pro and Premium, which are paid, allow report collaboration, direct query, more advanced analytics features, and the use of the Power BI Report Server (Folio3, 2019).

Advantages: Inexpensive upgrade; Large custom visualizations range; Easy integration Excel; Quick learning curve for basic use.

Limitations: Bulky user interface; Online reports must be public to the whole Internet.

Figure 1 shows the Power BI Free user interface.

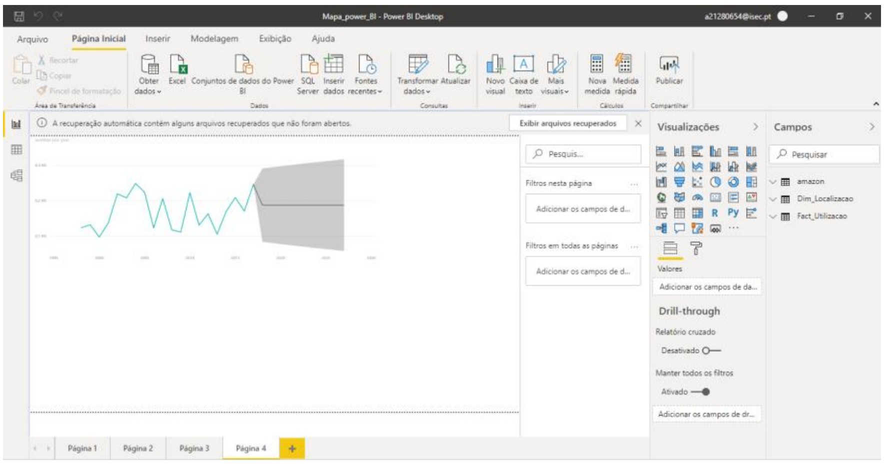

Figure 1: Power BI Free user interface.

Power BI Free has available a huge number of learning resources available on the web, an active community, and over 70 numerous integrations and data sources (Microsoft, 2020a).

2.2 *QlikView*

QlikView is a robust proprietary desktop platform for business discovery which offers a powerful free version in terms of features. It was first released in 2012 (Qlik, 2020a) with frequent updates since then. It can be installed only on Windows OS (Qlik, 2020b), being suitable to companies of all sizes.

Its free version has no limitations in terms of time or functionality compared with its paid edition. However, the files/documents created by a free-license user cannot be opened on another computer or shared with a user who has a paid license.

Advantages: Fast user experience for being a memory-resident application; Fast implementation.

Limitations: It does not allow "write back " to the database; Reloading can take a significant amount of time as it loads most data into the system RAM; The user interface is not intuitive and looks unfriendly.

Figure 2 illustrates the QlikView user interface.

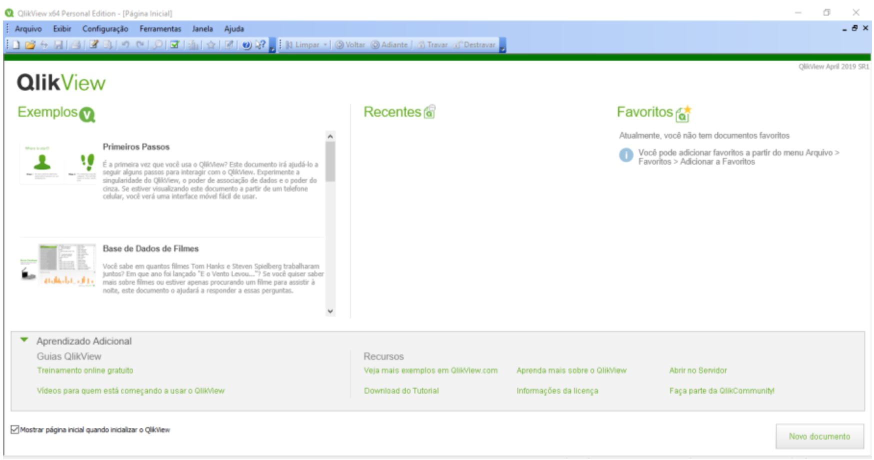

Figure 2: QlikView user interface.

Its in-memory engine recognizes patterns in data that we are not normally able to do it by using SQL alone (Kumar, 2019).

2.3 *Tableau Public*

Tableau is one of the most famous “self-service” visualization and analytics tools on the market. Its desktop version was first released in 2004 (Tableau, 2004) and was designed for companies of all sizes. It is not open-source but it has a commercially free platform, which is updated frequently.

It runs on Windows or Mac OS to be used in conjunction with the web free version. Many of the same powerful visualization capabilities its paid desktop and server versions features are available at no cost. Data Analyses is possible from sources such as Excel sheets for geographical visualizations, Gantt charts, treemaps, and other templates.

However, it is possible only to connect to Excel sheets, text file formats, statistical files, Google sheets, and web data connectors, which must be uploaded to the cloud. The free version has a limitation of 15.000.000 data rows per workbook (Tableau, 2016).

Advantages: Quick responsiveness; Extensive training resources available for free; Very intuitive user interface; Dashboards can be viewed on multiple devices;

Limitations: To keep workbooks private, a paid subscription is required; Complex visualizations require time and cost-intensive training.

Figure 3 represents the Tableau Public Desktop user interface.

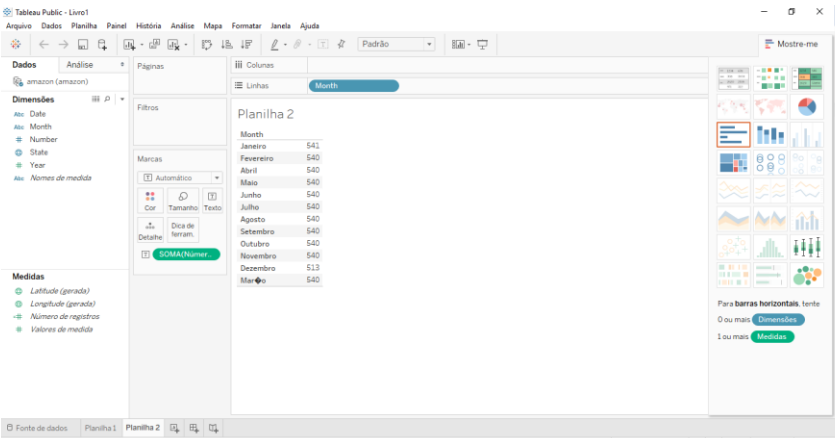

Figure 3: Tableau Public Desktop user interface.

Tableau Public is a very sophisticated and advanced system. It surpasses other tools mostly in data visualization. It provides an all-inclusive and user-friendly data visualization experience (Tableau, 2020d).

3. OSSPAL METHODOLOGY

OSSPal methodology is an evolution of OpenBRR methodology (Marinheiro & Bernardino, 2015). The OSSPal methodology uses metrics to identify software quality level in seven categories (Wasserman et al., 2017):

- **Functionality:** How well will the software meet the average user's requirements?
- **Operational Software Characteristics:** How secure is the software? How well does the software perform? How well does the software scale to a large environment? How good is the

UI? How easy to use is the software for end-users? How easy is the software to install, configure, deploy and maintain?

- **Support and Services:** How well is the software component supported? Is there commercial and/or community support? Are there people and organizations that can provide training and consulting services?
- **Documentation:** Is there adequate tutorials and reference documentation for the software?
- **Software Technology Attributes:** How well is the software architected? How modular, portable, flexible, extensible, open, and easy to integrate is it? Are the design, the code, and the tests of high quality? How complete and error-free are they?
- **Community and Adoption:** How well is the component adopted by community, market, and industry? How active and lively is the community for the software?
- **Development Process:** What is the level of the professionalism of the development process and the project organization as a whole?

The evaluation process is divided into four phases:

1. Construction of a capabilities list we consider to be essential in a determined software type for analyses and measurement.
2. Weight attribution for categories and measures by assigning a percentage for each category according to its importance, which should total 100%.
3. Data gathering for each measure of each category to calculate its weight from 1 to 5 (1 - Unacceptable, 2 - Poor, 3 - Acceptable, 4 - Very Good, 5 - Excellent).
4. Finally, OSSPal final score calculation based on 2).

As the category ‘Functionality’ is composed of the features mentioned in 1), it must be calculated separately, as follows:

- Score each feature from 1 to 3 (less important to very important);
- Use weighted average to scale the scores given in a range from 1 to 5.

The functionality category will have the following scale:

- Under 65%, Score = 1 (Unacceptable);
- 65% - 80%, Score = 2 (Poor);
- 80% - 90%, Score = 3 (Acceptable);
- 90% - 96%, Score = 4 (Good);
- Over 96%, Score = 5 (Excellent).

4. EVALUATION PROCESS

First, we determined a weight for each category of this methodology in order of importance (see Table 1).

CATEGORY	WEIGHT
Functionality	30%
Operational Software Characteristics	15%
Documentation	15%
Community and Adoption	15%
Software Technology Attributes	10%
Support and Service	10%
Development Process	5%

Table 1: Assigned weights to the categories

The software's functionalities set is the most relevant aspect, as it reveals the software utility. For this reason, the category "Functionality" received the highest weight, 30%. The next three categories had the second most relevant weight, 15%.

"Operational Software Characteristics" involves aspects such as security, performance, usability, and implementation. It had attributed this weight because without the referred aspects no software can be useful regardless of the functionalities it may offer. Moreover, we have "Documentation", once they are essential for software implementation and troubleshooting. "Community and Adoption" is at the same weight level because it is where users can obtain support, especially in case of free software. Furthermore, it allows us to measure the tool's acceptance in its market.

Following this, with 10% of weight, is "Software Technology Attributes", as it considers how error-free the tool is, which is indeed important. However, it also includes aspects that normally self-service and BI and Analytics end-users are not interested in, e.g. code and test quality. "Support and Service" has a similar weight because end-users of free tools are generally aware that they cannot require commercial support, training or consulting services, unless they pay for it.

"Development Process" had the lowest weight, 5%, as it concerns the quality level of the software's project organization, professionalism and development fashion. These are generally irrelevant aspects for the Self-service BI and Analytics software's target users since they have normally little IT technical knowledge and will not likely consider this category when deciding on a software.

Next, we have Table 2, where weights were assigned to each functionality category according to its relevance (1 - slightly important, 2 - important and 3 - very important).

FUNCTIONALITIES / CRITERIA	WEIGHT
Access control and security	3
Ad-hoc reporting	3
Ad-hoc query	3
Cloud Services	2
Data visualization variety	3
Data Integration	3
Dashboard Designer	3
Interactive Visualization	3
Mobile capabilities	2
Natural Language Query	1
OLAP	3
Predictive Analytics	3
Real-time Analytics	3
Real-time Collaboration	3
Report Customization and Scheduling	3

Table 2: Weights for each functionality category.

Now, after collecting data, we calculate a score for all measures of each category in a range between 1 to 5 (see Table 3).

CATEGORY	SCORE		
	POWER BI FREE	QLIKVIEW	TABLEAU PUBLIC
Functionality	3.83	3.32	2.78
Operational Software Characteristics	4	3	3
Software Technology Attributes	5	5	5
Documentation	5	5	5
Community and Adoption	5	5	5
Support and Service	4	3	4
Development Process	5	5	5

Table 3: OSSPal score by category.

As we can see in Table 3, Power BI Free obtained the highest score, in a range from 0 to 5, for the "Functionality" category. That is justified by the fact it lacks only "Real-time Collaboration" among all the referred functionalities in Table 2. Besides that, its score is not even higher because it had attributed low punctuation in 3 categories as follows: "Access Control and Security", once published workbooks must be public to the Internet, "Cloud Services", as its cloud application just allow dashboard/report visualization and small editions, and "Mobile Capabilities" because the mobile app just allows dashboard/report visualizations.

QlikView for the Functionality criteria obtained the second-best score, as it lacks “Cloud Services”, “Natural Language Query”, “Real-time Analytics”, and “Real-time Collaboration”. In addition to that, low punctuation was given to “Mobile Capabilities” because we just found an outdated iOS application available for installation.

Tableau Public occupies the third position in the functionalities category for not offering “Natural Language Query”, “Real-time Analytics”, “Real-time Collaboration”, and “Report Customization and Scheduling” capabilities. Moreover, it had a low score for the “Cloud Services” capability, once its cloud application allows data visualization, but with no dashboard editing possibilities, differently from Power BI Free, which allows some basic editing for visualizations.

Concerning the remaining categories, the three evaluated tools had very similar scores, once they are already mature software solutions on the market. However, it is important to mention that in “Operational Software Characteristics” Power BI Free stood out for having a more intuitive user interface compared to the other two solutions. As for “Support and Service” for the three tools, the end-users can count on support from an active community and extensive online official training resources, such as user guides and videos. Despite this, QlikView had a penalization, since its official training videos are paid.

After scoring Functionalities in Table 2 and Categories in Table 3, we have calculated a final score for every tool. By multiplying each score from Table 3 by the category weights from Table 1, we have obtained the scores in Table 4.

	SCORE		
	POWER BI FREE	**QLIKVIEW**	**TABLEAU PUBLIC**
Total	4.40	4.00	3.93

Table 4: OSSPal final score.

Overall, as we can see in Table 4, Power BI Free has the best final score of 4.4 (out of 5) through the application of the OSSPal methodology. QlikView has the next best score of 4.00, and Tableau Public the lowest score of 3.93.

5. Conclusions and Future Work

In this paper, we evaluated three of the most popular Self-Service BI and Analytics tools in their free versions. To perform this evaluation, we tested the tools. Besides, we considered our experience, official documentation, and third-party websites which publish reviews and rankings about the tools, so that we could also choose some tests to perform based on the pros and cons those webpages published about each tool.

Through OSSPal, we could classify Power BI as "Good", due to its best final score. This explains the high acceptance the solution has on the market. Its costless version offers almost everything we considered to be essential for a software, restricting only a feature related to collaboration and refresh for Real-time Analytics just every 30 minutes. Furthermore, it stands out in features such as Ad-hoc Reporting and Predictive Analytics due to its higher easy-of-use compared to the other tools, being the only one to feature Natural Language Query at no cost. It relies more on drag-and-drop and intuitive features.

With the second-best score is QlikView, which was also classified as "Good". It is also a powerful and solid BI and Analytics tool. However, it obtained a lower score mostly because it offers a few less free features than Power BI. Moreover, its interface is not as intuitive as Power BI's interface, which is highly relevant for a Self-service solution. QlikView seems to be more a traditional, technical tool for users with already some experience with BI, data analytics, and reporting and has the very strong point of offering in its free version all functionalities of its paid license, locking just workbook's sharing capabilities.

Tableau Public had the lowest final punctuation and was then categorized as "Acceptable". However, with almost the same score as QlikView. It was more penalized than the previous tool because it offers fewer functionalities from Table 2 for free. Also, as it is not as intuitive as Power BI, new users may have to learn some Data Science before starting to benefit from it. However, it is the quickest in responsiveness with overall higher quality for visualizations. Furthermore, it also stands out from the other two tools for its extensive community and free training resources.

We would like to emphasize that our assessment study considered only the free functionalities of the evaluated tools. Thus, if we had considered paid capabilities, the score results would naturally be different, once the "Functionality" category has a 30% weight.

As future work, we intend to perform a comparative analysis of other relevant Self-Service BI and Analytics solutions to make available a wider set of them for choice by the SMEs according to their requirements.

REFERENCES

Bernardino, J. (2011). Open Source Business Intelligence Platforms for Engineering Education. In Proceedings of the 1st World Engineering Education Flash Week, WEE 2011, pp. 27–30.

Bernardino, J. & Neves P. C. (2016). Decision-Making with Big Data Using Open Source Business Intelligence Systems. Human Development and Interaction in the Age of Ubiquitous Technology, IGI Global, pp. 120-147.

Folio3. (2019). Difference Between Power BI Pro vs. Power BI Free vs. Power BI Premium. Retrieved January 23, 2020, from https://dynamics.folio3.com/blog/difference-between-power-bi-pro-vs-free-vs-premium/

Gartner. (2020). Self-service Analytics. Retrieved January 11, 2020, from https://www.gartner.com/en/information-technology/glossary/self-service-analytics

Howson, A. C., Richardson, J., Sallam, R., & Kronz, A. (2019). Magic Quadrant for Analytics and Business Intelligence Platforms. Retrieved February 19, 2020, from https://www.gartner.com/doc/reprints?id=1-68720FP&ct=190213&st=sb

Insider, B. D. (2017). Was ist Self-Service BI (Business Intelligence)? Retrieved March 18, 2020, from https://www.bigdata-insider.de/was-ist-self-service-bi-business-intelligence-a-659722/

Kumar, A. (2019). Comparison of Tableau Vs QlikView | QlikView Tableau Differences. Retrieved January 22, 2020, from https://www.knowledgehut.com/blog/business-intelligence-and-visualization/tableau-vs-qlikview

Marinheiro, A., & Bernardino, J. (2015). Experimental Evaluation of Open Source Business Intelligence Suites using OpenBRR, IEEE Latin America Transactions 13 (3), 810-817.

Microsoft. (2020). Data sources in Power BI Desktop - Power BI | Microsoft Docs. Retrieved January 22, 2020, from https://docs.microsoft.com/en-us/power-bi/desktop-data-sources

Pat Research. (2019a). About Pat Research - Pat Research: B2B Reviews, Buying Guides & Best Practices. Retrieved December 14, 2019, from https://www.predictiveanalyticstoday.com/about/

Pat Research. (2019b). How to Select the Best Self Service Analytics Software for Your Business - Compare Reviews, Features, Pricing in 2019 - Pat Research: B2B Reviews, Buying Guides & Best Practices. Retrieved November 25, 2019, from https://www.predictiveanalyticstoday.com/self-service-analytics-software/

Qlik. (2020a). Released QlikView versions including build number. Retrieved February 13, 2020, from https://support.qlik.com/articles/000003216

Qlik. (2020b). System Requirements for QlikView – QlikView. Retrieved April 5, 2020, from https://help.qlik.com/en-US/qlikview/April2019/Content/QV_HelpSites/System-requirements.htm

Tableau. (2004). 1.0 | Tableau Software. Retrieved March 18, 2020, from https://www.tableau.com/fast-pace-innovation/1.0

Tableau. (2016). Tableau Public Frequently Asked Questions |Tableau Community Forums. Retrieved January 5, 2020, from https://community.tableau.com/docs/DOC-9135#subhead20

Tableau. (2020). What is data visualization? A definition, examples, and resources. Retrieved January 22, 2020, from https://www.tableau.com/learn/articles/data-visualization

Techtarget. (2016). What is self-service business intelligence (BI)? - Definition from WhatIs.com. Retrieved January 10, 2020, from https://searchbusinessanalytics.techtarget.com/definition/self-service-business-intelligence-BI

Wasserman, A. I., Guo, X., Mcmillian, B., Qian, K., Wei, M.-Y., & Xu, Q. (2017). OSSpal: Finding and Evaluating Open Source Software. https://doi.org/10.1007/978-3-319-57735-7_18

Wikipedia. (2019). Power BI – Wikipédia, a enciclopédia livre. Retrieved February 13, 2020, from https://pt.wikipedia.org/wiki/Power_BI

This page is intentionally left blank.

APPENDIX C – SURVEY FOR EVALUATION OF VISUAL ANALYTICS SOFTWARE

Evaluation Criteria for the Analytic Process Area (1/4)										
			How would you rate your agreement with the following statements?							
	Heuristic	Guideline	Strongly Disagree (1)	Disagree (2)	Somewhat Disagree (3)	Neither Agree nor Disagree (4)	Somewhat Agree (5)	Agree (6)	Strongly Agree (7)	N/A
AP1	VAS shows key characteristics of data at a glance.	VAS should feature visualizations that allow identifying key characteristics of data in a quick short look.	○	○	○	○	○	○	○	○
AP2	VAS makes data relationships noticeable.	VAS should facilitate answering questions about the data by making relationships in it noticeable. That is, by making visible, for instance: distribution of variables, correlations, and clusters.	○	○	○	○	○	○	○	○
AP3	VAS provides a new or better understanding of the data.	VAS should provide a new or better understanding of the data. This through helping identify unexpected, duplicate, missing, or invalid data. Also, dependent, independent, and important dimensions.	○	○	○	○	○	○	○	○
AP4	VAS helps generate data-driven questions.	VAS should help the user generate data-driven questions from its analytical outcomes.	○	○	○	○	○	○	○	○
AP5	VAS suggests relevant information beyond dataset information.	VAS should not only suggest relevant information about the dataset itself and its attributes, but also, for instance, about related views, comments, and data to current points of interest, as well as notification subscriptions for views, artefacts (reports, dashboards, and datasets), people, and groups.	○	○	○	○	○	○	○	○
AP6	VAS features visualization which provides a comprehensive data overview with a meaningful visual schema.	VAS should feature visualization that provides a big picture/perspective of the data through an accessible data overview and meaningful visual schema.	○	○	○	○	○	○	○	○
AP7	VAS provides coordinated views for linked information.	Visualizations on VAS should be coordinated together in such a way that action performed in one view affects all other views.	○	○	○	○	○	○	○	○
AP8	VAS displays related information nearby.	VAS should show related information in close proximity.	○	○	○	○	○	○	○	○
AP9	VAS minimizes distractions for the analyst.	VAS should minimize distractions for the analyst. That is, minimize aesthetics or interactions that take the user outside of the frame of the task. Minimizing distractions assists endogenous attention and reduction in time.	○	○	○	○	○	○	○	○

Evaluation Criteria for the Analytic Process Area (Continued - 2/4)										
			How would you rate your agreement with the following statements?							
	Heuristic	Guideline	Strongly Disagree (1)	Disagree (2)	Somewhat Disagree (3)	Neither Agree nor Disagree (4)	Somewhat Agree (5)	Agree (6)	Strongly Agree (7)	N/A
AP10	VAS provides opportunities for serendipitous discoveries.	VAS should provide opportunities for serendipitous discoveries by displaying information from multiple aspects, as well as related and partially related data points.	○	○	○	○	○	○	○	○
AP11	VAS allows flexibility in the organization of the visual metaphor.	VAS should allow flexibility in the organization of the visual metaphor.	○	○	○	○	○	○	○	○
AP12	VAS facilitates finding starting points or clues.	VAS should provide an environment in which the user can capture information to find starting points or clues. That is, it should direct attention to the most critical information.	○	○	○	○	○	○	○	○
AP13	VAS provides strong retrieval cues for mental models.	VAS should structure information in a way which provides strong retrieval cues for mental models* aiding in reasoning**.	○	○	○	○	○	○	○	○
AP14	VAS allows share evidence and hypothesis.	VAS should have the ability to share evidence and hypotheses so that users can create hypotheses regarding their analysis, collect them, and share them with other users. Likewise, it should be possible for the collected evidence. That being feasible, for instance through shared, editable representations; in-app collaborative editing; embedding of annotated views in external media (e.g., email, blogs, and reports); or sharing of views across media (e.g., URLs).	○	○	○	○	○	○	○	○
AP15	VAS supports collection of evidence and annotations in a beneficial organization to sensemaking.	VAS should allow collecting and grouping evidence and annotations, as well as to register the need for more evidence or other future actions, preferably through storytelling, in a beneficial scheme to the sensemaking process.	○	○	○	○	○	○	○	○
AP16	VAS should allow registering need for more evidence or other future actions regarding the analytic process.	VAS should allow registering need for more evidence or other future actions regarding the analytic process.	○	○	○	○	○	○	○	○
AP17	VAS supports sensemaking by recommending relevant information.	VAS should support sensemaking by presenting semantically meaningful recommendations that enrich the current analytic process based on the user’s current activity and potential next step.	○	○	○	○	○	○	○	○
AP18	VAS displays statistics and measures about data sources, datasets, and/or records.	VAS should support evidence discovery by displaying statistics and measures regarding data sources, datasets, and/or records.	○	○	○	○	○	○	○	○

Evaluation Criteria for the Analytic Process Area (Continued – 3/4)										
			How would you rate your agreement with the following statements?							
	Heuristic	Guideline	Strongly Disagree (1)	Disagree (2)	Somewhat Disagree (3)	Neither Agree nor Disagree (4)	Somewhat Agree (5)	Agree (6)	Strongly Agree (7)	N/A
AP19	VAS features a visual display of the analytic process.	VAS should feature a visual display of the process so that there is no need to keep external notes.	○	○	○	○	○	○	○	○
AP20	VAS provides an easy-to-interpret environment for contextual analysis with relevant information.	VAS should provide an easy-to-interpret environment for contextual analysis composed by relevant information for the analysis and suggestions about what may have been overlooked.	○	○	○	○	○	○	○	○
AP21	VAS provides transparent automation to the user regarding the underlying mathematical models and parameters.	VAS should contain automation, which is transparent to the user, shielding users from the complexity of underlying mathematical models and parameters.	○	○	○	○	○	○	○	○
AP22	VAS captures and understands user interactions.	VAS should be able to capture and understand (the kind of action) user interactions in spatial analytic processes such as searching, highlighting, annotating, and repositioning documents for future automation.	○	○	○	○	○	○	○	○
AP23	VAS makes inferences from user interactions.	VAS should be able to make inferences (deductions) from user interactions. For example, for suggestions of recommendations regarding the analysis.	○	○	○	○	○	○	○	○
AP24	VAS reacts and takes initiative based on inferences from user interactions.	VAS should be able to react and take initiative based on those inferences at three levels: interface, computation, and cognitive.	○	○	○	○	○	○	○	○
AP25	VAS provides visual feedback regarding the updated model.	VAS should also provide visual feedback of the updated model and learned parameters within the visual metaphor.	○	○	○	○	○	○	○	○
AP26	VAS features teamwork management.	VAS should feature group creation and teamwork management, including division of labour among participants.	○	○	○	○	○	○	○	○
AP27	VAS features activity indicators per collaborator increasing so engagement.	VAS should provide a history of past contributions, to create activity indicators, as well as to aid reputation and visibility of contributions, so that engagement increases.	○	○	○	○	○	○	○	○
AP28	VAS supports intuitive communication among collaborators.	VAS should support intuitive communication to support discussions on common ground. In other words, it should provide intuitive means to share understanding among collaborators to facilitate consensus and decision making.	○	○	○	○	○	○	○	○

Evaluation Criteria for the Analytic Process Area (Continued - 4/4)										
			How would you rate your agreement with the following statements?							
	Heuristic	Guideline	Strongly Disagree (1)	Disagree (2)	Somewhat Disagree (3)	Neither Agree nor Disagree (4)	Somewhat Agree (5)	Agree (6)	Strongly Agree (7)	N/A
AP29	VAS allows to track update of collaborative threads.	VAS should allow tracking update of collaborative threads regarding the analysis.	○	○	○	○	○	○	○	○
AP30	VAS supports future scenario projections.	VAS should support users in making future scenario projections such as forecasting.	○	○	○	○	○	○	○	○
AP31	VAS integrates multiple information channels.	VAS should provide means to integrate multiple information sources, forming a single unified content collection.	○	○	○	○	○	○	○	○
AP32	VAS increases engagement and attention with game design elements.	VAS could use game design elements to reframe tedious data entry tasks as actions within online games for increasing engagement. For instance, a team-oriented 'scavenger hunt' analysis would allocate more attention.	○	○	○	○	○	○	○	○

Evaluation Criteria for the Visualization Quality Area										
			How would you rate your agreement with the following statements?							
	Heuristic	Guideline	Strongly Disagree (1)	Disagree (2)	Somewhat Disagree (3)	Neither Agree nor Disagree (4)	Somewhat Agree (5)	Agree (6)	Strongly Agree (7)	N/A
VQ1	VAS facilitates perception via Gestalt principles.	VAS should guide and maximize perception via Gestalt principles (proximity, similarity, enclosure, closure, continuity, and connection) in its visualizations.	○	○	○	○	○	○	○	○
VQ2	VAS provides visualizations with meaningful spatial organization.	VAS should care about the visualization overall layout, displaying a meaningful spatial organization of the data.	○	○	○	○	○	○	○	○
VQ3	VAS avoids dense visualizations by featuring properties for size and distance.	VAS should offer appropriate and easy to interpret representations for properties such as size and distance in visualizations, avoiding so dense visualizations.	○	○	○	○	○	○	○	○
VQ4	VAS uses animation only to show an effect that moves over time.	VAS should use animations only to show an effect that moves over time. Give analysts control to manipulate the speed of the animation.	○	○	○	○	○	○	○	○
VQ5	VAS displays only relevant information and elements in a straightforward fashion.	VAS should avoid misleading and complex representations by displaying only relevant elements to the analytic process in a straightforward fashion.	○	○	○	○	○	○	○	○

Evaluation Criteria for the Interactivity Area (1/1)										
			How would you rate your agreement with the following statements?							
	Heuristic	Guideline	Strongly Disagree (1)	Disagree (2)	Somewhat Disagree (3)	Neither Agree nor Disagree (4)	Somewhat Agree (5)	Agree (6)	Strongly Agree (7)	N/A
I1	VAS features self-descriptive interactions.	VAS should feature self-descriptive interactions. That is, it should allow users intuitively understand what they can do with the interaction and how they can do it (Usability.de, 2020).	○	○	○	○	○	○	○	○
I2	VAS provides tools for data manipulation.	VAS should provide tools to help users in data manipulation. For instance, tools for filtering, clustering, pruning.	○	○	○	○	○	○	○	○
I3	VAS provides capabilities for data exploration.	VAS should feature useful interactive capabilities to help investigate data in multiple ways. For example, zooming; navigation and querying (including selection of objects, viewpoint manipulation, geometric manipulation, and searching).	○	○	○	○	○	○	○	○

Evaluation Criteria for the Interactivity Area (Continued - 2/2)										
			How would you rate your agreement with the following statements?							
	Heuristic	Guideline	Strongly Disagree (1)	Disagree (2)	Somewhat Disagree (3)	Neither Agree nor Disagree (4)	Somewhat Agree (5)	Agree (6)	Strongly Agree (7)	N/A
I4	VAS allows interactive visualization customization.	VAS should support customization of the visualization. For instance, by using different attributes of the data to reorganize its appearance and supporting several dimensions simultaneously in it.	○	○	○	○	○	○	○	○
I5	VAS avoids complex commands and queries in visualizations.	VAS should avoid complex commands and textual queries in visualizations by providing direct interaction with the data representation.	○	○	○	○	○	○	○	○
I6	VAS provides a way to backtrack and to undo actions.	VAS should provide a way to track changes in information and undo actions (a history with all user actions may be used).	○	○	○	○	○	○	○	○
I7	VAS provides alternative ways to perform a task.	VAS should provide alternative ways to perform a task, such as shortcuts for experienced users, to increase the interaction speed and to reduce time.	○	○	○	○	○	○	○	○
I8	VAS provides interaction with minimal need for repetitive actions.	VAS should provide means to explore visualizations and overall system functions avoiding as much as possible repetitive actions on the part of the end-user.	○	○	○	○	○	○	○	○
I9	VAS gives proper feedback to user actions within reasonable time.	VAS should provide appropriate feedback as a response to user actions within reasonable time.	○	○	○	○	○	○	○	○
I10	VAS guides users towards making specific actions.	VAS should guide users on specific actions by showing selectable option, windows titles, system status, data fields with labels and acceptable values and formats.	○	○	○	○	○	○	○	○
I11	VAS facilitates the understanding of relationships between the various user interface items	VAS should provide means to understand the relationships among items by grouping similar objects according to formats and graphical features; screen areas; and between different classes of objects.	○	○	○	○	○	○	○	○

Evaluation Criteria for the User-friendliness Area										
			How would you rate your agreement with the following statements?							
	Heuristic	Guideline	Strongly Disagree (1)	Disagree (2)	Somewhat Disagree (3)	Neither Agree nor Disagree (4)	Somewhat Agree (5)	Agree (6)	Strongly Agree (7)	N/A
UF1	VAS provides coherent UI elements.	VAS should follow similar meaning and design choices in similar contexts. That is, the interface elements should be coherent.	○	○	○	○	○	○	○	○
UF2	VAS features a UI with familiar signs to the user.	VAS should feature a UI where all signs (codes, names, texts, figures, and icons) in the UI are familiar to the user and have an expected meaning.	○	○	○	○	○	○	○	○
UF3	VAS matches user characteristics with the UI characteristics	VAS should feature a UI compatible with the user characteristics, (language, measurement units, calendar, and accessibility capabilities).	○	○	○	○	○	○	○	○
UF4	VAS provides proper help and documentation to guide the user	VAS should provide help and documentation to guide the user. All actions that the user can realize in the system should be easily identified/visible. Also, should be available easy-to-understand tutorials (especially for the not-easily-identified functionalities).	○	○	○	○	○	○	○	○
UF5	VAS makes easily visible all possible actions for the user	VAS should make easily visible all possible actions the user can perform, be intuitively or through help, documentation, and tutorials.	○	○	○	○	○	○	○	○

Evaluation Criteria for the Satisfaction Area										
			How would you rate your agreement with the following statements?							
	Heuristic	Guideline	Strongly Disagree (1)	Disagree (2)	Somewhat Disagree (3)	Neither Agree nor Disagree (4)	Somewhat Agree (5)	Agree (6)	Strongly Agree (7)	N/A
S1	VAS is effective in representing high-quality of analytical outcomes.	VAS should show the high quality of analytic outcomes on the visual interfaces. That is the system should be effective in depicting through the visualization the outcomes from the automated data analysis.	○	○	○	○	○	○	○	○
S2	VAS ensures the end-user's subjective assessment is overall positive.	VAS should ensure an overall positive end-user subjective assessment. This involves overall subjective satisfaction about the system, which relates to how pleasant and easy-to-use it is, as well as frustrating experiences, and productivity through it.	○	○	○	○	○	○	○	○
S3	VAS is considered highly useful.	VAS should be considered highly useful. In other words, it refers to whether the system provides the features the user needs.	○	○	○	○	○	○	○	○
S4	VAS minimizes the needed resources to achieve the goal.	VAS should maximize efficiency by minimizing the necessary resources to achieve the goal. It should maximize the speed and minimize the number of steps to achieve an objective.	○	○	○	○	○	○	○	○

Evaluation Criteria for the Error-Handling Area										
			How would you rate your agreement with the following statements?							
	Heuristic	Guideline	Strongly Disagree (1)	Disagree (2)	Somewhat Disagree (3)	Neither Agree nor Disagree (4)	Somewhat Agree (5)	Agree (6)	Strongly Agree (7)	N/A
EH1	VAS prevents, diagnoses, and correct errors.	VAS should prevent, diagnose, correct, and recover from errors with clear and informative messages, giving reasons, as well as the means to correct them.	○	○	○	○	○	○	○	○

valuation Criteria for the Adequacy Area										
			How would you rate your agreement with the following statements?							
	Heuristic	Guideline	Strongly Disagree (1)	Disagree (2)	Somewhat Disagree (3)	Neither Agree nor Disagree (4)	Somewhat Agree (5)	Agree (6)	Strongly Agree (7)	N/A
A1	VAS is adequate for its context of use.	VAS should be compatible with the context for which it was designed. That is should be suitable to facilitate the analytical goals of its context of use.	○	○	○	○	○	○	○	○

This page is intentionally left blank.

APPENDIX D – SHEET FOR OVERVIEW AND CALCULATION OF THE VISUAL ANALYTICS SCORE

Heuristics per Evaluation Area and Assigned Scores (SCs)													
Analytical Process		**Visualization Quality**		**Interactivity**		**User-friendliness**		**Satisfaction**		**Error-Handling**		**Adequacy**	
AP	**SC**	**V**	**SC**	**I**	**SC**	**UF**	**SC**	**S**	**SC**	**EH**	**SC**	**A**	**SC**
AP1		V1		I1		UF1		S1		EH1		A1	
AP2		V2		I2		UF2		S2					
AP3		V3		I3		UF3		S3					
AP4		V4		I4		UF4		S4					
AP5		V5		I5		UF5							
AP6				I6									
AP7				I7									
AP8				I8									
AP9				I9									
AP10				I10									
AP11				I11									
AP12													
AP13													
AP14													
AP15													
AP16													
AP17													
AP18													
AP19													
AP20													
AP21													
AP22													
AP23													
AP24													
AP25													
AP26													
AP27													
AP28													
AP29													
AP30													
AP31													
AP32													
Sap		**Svq**		**Si**		**Suf**		**Ss**		**Seh**		**Sa**	
Visual Analytics Score (VAs)													

Equation for the calculation of the Visual Analytics Score: $\text{VAs}=\sum s_i w_i$.

S_i refers to Sap, Sv, Si, Suf, Ss, Seh, and Sa, which are simple averages. W_i refers to the weights assigned to each the evaluation area. See **Subsection 4.3.3** for a suggestion regarding weights.

This page is intentionally left blank.

APPENDIX E – ORIGINAL HEURISTICS AND GUIDELINES CLASSIFIED BY SCHOLTZ'S AREAS

Table 30 shows, in alphabetical order, the original heuristics and guidelines adopted to create the new heuristics and guidelines set by Scholtz's Evaluation Areas.

Table 30 - Original Heuristics and Guidelines per Scholtz's Areas.

Heuristic	Scholtz's Areas	Release Year	Reference
A visual display of the process used is excellent as it relieves the analyst from having to keep external notes.	Utility & Situation Awareness	2011	(Scholtz, 2011)
Ability to share evidence	Collaboration & Utility	2014	(Adagha et al., 2017)
Ability to show high quality of analytic solutions	Creativity	2014	(Adagha et al., 2017)
Adaptive personal workspaces and preferences	Interaction & Usability	2011	(Wang et al., 2011a)
Aggregate information and show its patterns.	Utility & Situation Awareness	2011	(Wang et al., 2011a)
Aid analogical reasoning - structure information so as to provide strong retrieval cues for knowledge structures (mental models) to aid in analogical reasoning.	Situation Awareness & Utility	2014	(Tarrell et al., 2014)
Aid reasoning with mental models - organize information based on mental models to provide strong retrieval cues for knowledge structures in long-term memory to aid reasoning.	Situation Awareness & Utility	2014	(Tarrell et al., 2014)
Allow evidence collection and annotation	Utility & Situation Awareness	2011	(Wang et al., 2011a)
Allow flexibility in the organization of schemes	Interaction & Situation Awareness	2009	(Kang et al., 2009)
Allow view relations among data.	Utility & Situation Awareness	2017	(Oliveira & Silva, 2017)
Allows optimal group size determination (improve efficiency of analysis).	Collaboration & Utility	2008	(Heer & Agrawala, 2008)
Can provide an environment for contextual analysis	Situation Awareness	2014	(Adagha et al., 2017)
Can support future scenario projections	Utility & Situation Awareness	2014	(Adagha et al., 2017)
Can support intuitive communication	Collaboration & Utility	2014	(Adagha et al., 2017)
Can track changes in information	Interaction & Situation Awareness	2014	(Adagha et al., 2017)
Can track information flows	Collaboration & Utility	2014	(Adagha et al., 2017)
Capture exogenous attention - alert users to important attributes of a visualization.	Situation Awareness & Utility	2014	(Tarrell et al., 2014)
Compatibility with the user: Provide means to match the users'	Interaction & Usability	2017	(Pribeanu, 2017)

characteristics with the characteristics of the user interface.			
Compatible with the context of use	Utility	2014	(Adagha et al., 2017)
Consistency: Provide similar meanings and design choices in similar contexts.	Usability	2017	(Pribeanu, 2017)
Consistency: the interface elements must be coherent.	Usability	2017	(Oliveira & Silva, 2017)
Construct coordinated views for linked information	Situation Awareness	2011	(Wang et al., 2011a)
Controllability	Interaction & Usability	2014	(Adagha et al., 2017)
Data Characterization: assist data understanding.	Situation Awareness	2017	(Oliveira & Silva, 2017)
Data Manipulation: provide tools for data manipulation, such as filters and detailed view	Interaction	2017	(Oliveira & Silva, 2017)
Data set reduction (including filtering, clustering, and pruning)	Interaction	2014	(Tarrell et al., 2014)
Deliver contents in straightforward representation	Utility	2011	(Wang et al., 2011a)
Display coordinating information in close proximity.	Situation Awareness	2011	(Scholtz, 2011)
Display information in consistent format	Situation Awareness	2011	(Wang et al., 2011a)
Division of labour among participants.	Collaboration & Utility	2008	(Heer & Agrawala, 2008)
Empower with Numbers	Utility & Situation Awareness	2012	(Kang & Stasko, 2012)
Enable access to information	Interaction	2014	(Adagha et al., 2017)
Enable facet filtering for information personalization	Interaction	2011	(Wang et al., 2011a)
Enable identification of collaborators in a contextually appropriate manner.	Collaboration & Utility	2008	(Heer & Agrawala, 2008)
Enable in-app collaborative editing	Collaboration & Utility	2011	(Wang et al., 2011a)
Enable lightweight sharing of views across media with bookmarks (e.g., URLs)	Collaboration & Utility	2008	(Heer & Agrawala, 2008)
Enable rapid visual interpretation of recommendations.	Situation Awareness	2015	(Cook et al., 2015)
Error Correction: inform users about errors that occurred with clear messages and present means to correct these errors.	Interaction & Usability	2017	(Oliveira & Silva, 2017)
Error management: Provide means to prevent, diagnose, correct, and recover from errors.	Interaction & Usability	2017	(Pribeanu, 2017)
Error Prevention: prevent error occurrence, eliminating error-prone conditions.	Interaction & Usability	2017	(Oliveira & Silva, 2017)
Facilitate chunking - choose visualization parameters that provide strong grouping cues to facilitate the chunking of information, which will minimize the effects of working-memory capacity limitations.	Interaction & Situation Awareness	2014	(Tarrell et al., 2014)
Feedback: Provide appropriate feedback as a response to user's actions within reasonable time.	Interaction & Usability	2017	(Pribeanu, 2017)

Flexibility and Efficiency: provide accelerators and customization features.	Interaction & Usability	2017	(Oliveira & Silva, 2017)
Flexibility: provide means to customize the interface and select the preferred way to accomplish a goal.	Interaction & Usability	2017	(Pribeanu, 2017)
Group creation and management mechanisms.	Collaboration & Utility	2008	(Heer & Agrawala, 2008)
Grouping/distinction: Provide means to group similar objects and distinguish between different classes of objects.	Interaction & Usability	2017	(Pribeanu, 2017)
Guide endogenous attention - provide appropriate organization of material or interaction options to assist endogenous attention and minimize distracting information.	Utility	2014	(Tarrell et al., 2014)
Guide perception using pre-attentive attributes such as spatial position.	Situation Awareness	2014	(Tarrell et al., 2014)
Guide perception via Gestalt principles.	Situation Awareness	2014	(Tarrell et al., 2014)
Guide the analyst to follow the right trail, without distraction	Situation Awareness & Utility	2009	(Kang et al., 2009)
Has activity indicators or summaries (aid reputation and visibility of contributions).	Collaboration & Utility	2008	(Heer & Agrawala, 2008)
Help and documentation: Provide online help and documentation.	Usability	2017	(Pribeanu, 2017)
Help to find appropriate next steps when encountering a dead-end	Situation Awareness	2009	(Kang et al., 2009)
Help to provide information scent appropriately, thus helping to find initial clues	Situation Awareness	2009	(Kang et al., 2009)
If the system contains automation, it should be well balanced and transparent to the user.	Utility	2011	(Scholtz, 2011)
If there were data issues like unexpected, duplicate, missing, or invalid data, the visualization would highlight those issues	Situation Awareness & Utility	2019	(Wall et al., 2019)
Increase engagement by increasing personal relevance of data sets.	Situation Awareness	2008	(Heer & Agrawala, 2008)
Infer users' tasks based on their activities	Interaction & Situation Awareness	2015	(Cook et al., 2015)
Interactive content exploration and filtering	Interaction	2011	(Wang et al., 2011a)
Intimate interaction is both transparent and supportive of holistic cognition.	Situation Awareness & Utility	2009	(Greensmith et al., 2009)
Invest in tutorials	Usability	2012	(Kang & Stasko, 2012)
Make activity patterns visible, determine popular and neglected data regions.	Situation Awareness	2008	(Heer & Agrawala, 2008)
Mark needed future actions: unanswered questions, need for evidence, etc.	Utility & Situation Awareness	2008	(Heer & Agrawala, 2008)
Maxime effectiveness: maximize the extent to which the goals of the users are achieved	Usability	2014	(Tarrell et al., 2014)

Maximize efficiency minimize the resources necessary to achieve the goal.	Usability	2014	(Tarrell et al., 2014)
Minimal actions: Minimize the number of actions needed to accomplish a task's goal.	Interaction & Usability	2017	(Pribeanu, 2017)
Mixed initiative sensemaking systems should provide an environment in which users can capture information relevant to their ongoing task and refine their conceptual models	Situation Awareness	2015	(Cook et al., 2015)
Mixed-initiative iterative sensemaking environments should recommend relevant data based on the user's current activity and potential next step	Interaction & Situation Awareness	2015	(Cook et al., 2015)
Multidimensionality: allow users to visualize three or more dimensions simultaneously	Interaction	2017	(Oliveira & Silva, 2017)
Navigation and Querying	Interaction	2014	(Tarrell et al., 2014)
Orientation and Help (control of level of details and support for undo)	Interaction & Usability	2014	(Tarrell et al., 2014)
Perceived ease of use	Creativity	2014	(Adagha et al., 2017)
Perceived usefulness	Utility	2014	(Adagha et al., 2017)
Personal action histories allow past contributions to be assessed	Collaboration & Utility	2008	(Heer & Agrawala, 2008)
Present status update of collaborative threads	Collaboration & Utility	2011	(Wang et al., 2011a)
Prompting: Guide users towards making specific actions.	Interaction & Usability	2017	(Pribeanu, 2017)
Provide a way to backtrack or undo actions.	Interaction & Usability	2011	(Scholtz, 2011)
Provide histories of actions performed on artifacts (representations/visualizations)	Situation Awareness	2008	(Heer & Agrawala, 2008)
Provide means for analysts to explore visualizations that do not require repetitive interactions on the part of the analyst.	Interaction & Usability	2011	(Scholtz, 2011)
Provide recommendations in context	Utility & Situation Awareness	2015	(Cook et al., 2015)
Provide satisfaction ensure the user s subjective assessment is generally positive.	Creativity	2014	(Tarrell et al., 2014)
Provide visual feedback of the updated model and learned parameters within the visual metaphor.	Interaction & Situation Awareness	2014	(Endert, 2014)
Real World Equivalency: use familiar signs to the user.	Usability	2017	(Oliveira & Silva, 2017)
Reduction in time	Utility & Usability	2014	(Adagha et al., 2017)
Self-descriptiveness	Interaction	2014	(Adagha et al., 2017)
Shield users from the complexity of the underlying mathematical models and parameters.	Utility	2014	(Endert, 2014)
Should integrate multiple information channels	Interaction & Usability	2011	(Wang et al., 2011a)
Spatial Organization and Perspective: care the visualization overall layout, as well as provide change of perspective.	Situation Awareness	2017	(Oliveira & Silva, 2017)

Spatial Organization and Perspective: care the visualization overall layout, as well as provide change of perspective.	Situation Awareness	2017	(Oliveira & Silva, 2017)
Structure collaboration through shared, editable representations.	Collaboration & Utility	2008	(Heer & Agrawala, 2008)
Suggest related views, comments, and data to current points of interest.	Collaboration & Utility	2008	(Heer & Agrawala, 2008)
Suggesting what may have been overlooked and keeping relevant information present.	Situation Awareness & Utility	2009	(Greensmith et al., 2009)
Support commentary; consider implications of discussion model on common ground.	Collaboration & Utility	2008	(Heer & Agrawala, 2008)
Support creation and export of presentations for telling analysis stories.	Collaboration & Utility	2008	(Heer & Agrawala, 2008)
Support customization of information	Interaction	2014	(Adagha et al., 2017)
Support embedding of annotated views in external media (e.g., email, blogs, and reports)	Collaboration & Utility	2008	(Heer & Agrawala, 2008)
Support notification subscriptions for views, artefacts, people, and groups.	Collaboration & Utility	2008	(Heer & Agrawala, 2008)
Support sharing of evidence and hypothesis	Collaboration & Utility	2011	(Wang et al., 2011a)
Support storytelling and enable interactive grouping of the evidence with users reasoning logic	Interaction & Situation Awareness	2011	(Wang et al., 2011a)
System Status and Feedback: notify users about the system status, and always provide quick and proper feedback	Usability	2017	(Oliveira & Silva, 2017)
Systems must be able to capture and understand user actions	Interaction & Situation Awareness	2015	(Endert et al., 2015)
Systems must be able to make inferences based on users' interactions.	Interaction & Situation Awareness	2015	(Endert et al., 2015)
Systems must be able to react and take initiative based on the inferences at three levels: interface, computation, and cognitive.	Interaction & Situation Awareness	2015	(Endert et al., 2015)
Task guidance and support: Provide the user with the procedure and associated support (forms, documents, etc.) needed to perform specific tasks.	Usability	2017	(Pribeanu, 2017)
The interface supports using different attributes of the data to reorganize the visualization's appearance	Interaction	2019	(Wall et al., 2019)
The visualization avoids complex commands and textual queries by providing direct interaction with the data representation.	Interaction & Utility	2019	(Wall et al., 2019)
The visualization avoids using misleading representations	Utility	2019	(Wall et al., 2019)
The visualization exposes individual data cases and their attributes	Situation Awareness	2019	(Wall et al., 2019)
The visualization facilitates perceiving relationships	Situation Awareness	2019	(Wall et al., 2019)
The visualization helps generate data-driven questions	Situation Awareness	2019	(Wall et al., 2019)

The visualization helps identify unusual or unexpected, yet valid, data characteristics or values	Situation Awareness	2019	(Wall et al., 2019)
The visualization presents the data by providing a meaningful visual schema	Situation Awareness	2019	(Wall et al., 2019)
The visualization promotes exploring relationships between individual data cases as well as different groupings of data cases	Interaction & Utility	2019	(Wall et al., 2019)
The visualization promotes understanding data domain characteristics beyond the individual data cases and attributes	Situation Awareness	2019	(Wall et al., 2019)
The visualization provides a comprehensive and accessible overview of the data	Situation Awareness	2019	(Wall et al., 2019)
The visualization provides a meaningful spatial organization of the data	Situation Awareness	2019	(Wall et al., 2019)
The visualization provides useful interactive capabilities to help investigate the data in multiple ways	Interaction	2019	(Wall et al., 2019)
The visualization shows key characteristics of the data at a glance	Situation Awareness	2019	(Wall et al., 2019)
The visualization shows multiple perspectives about the data	Situation Awareness	2019	(Wall et al., 2019)
The visualization supports smooth transitions between different levels of detail in viewing the data	Interaction	2019	(Wall et al., 2019)
The visualization uses an effective representation of the data that shows related and partially related data cases	Situation Awareness	2019	(Wall et al., 2019)
The visualization uses meaningful and accurate visual encodings to represent the data	Situation Awareness & Utility	2019	(Wall et al., 2019)
They must display only relevant information and elements to the user	Utility	2017	(Oliveira & Silva, 2017)
Trace interactions and system usage for future automation	Interaction & Situation Awareness	2011	(Wang et al., 2011a)
Unified content interface	Interaction & Situation Awareness	2011	(Wang et al., 2011a)
Use animations only to show an effect that moves over time. Give analysts control to manipulate the speed of the animation.	Interaction & Usability	2011	(Scholtz, 2011)
Use appropriate and easy to interpret representations for properties such as size and distance in visualizations.	Usability	2011	(Scholtz, 2011)
Use color to maximize perceptive effects.	Situation Awareness	2014	(Tarrell et al., 2014)
Use game design elements to provide incentives and to direct effort.	Creativity	2008	(Heer & Agrawala, 2008)
Use good aesthetics to minimize distractions and maximize perceptive effects Other perceptive aspects not represented above	Situation Awareness	2014	(Tarrell et al., 2014)
Use semantic interactions within the visual metaphor, based on common interactions occurring in spatial	Interaction & Situation Awareness	2014	(Endert, 2014)

analytic processes such as searching, highlighting, annotating, and repositioning documents.			
User Control: enable full system control by the user.	Interaction & Usability	2017	(Oliveira & Silva, 2017)
User satisfaction with solutions	Creativity	2014	(Adagha et al., 2017)
Visible Actions: make all possible actions visible.	Usability	2017	(Oliveira & Silva, 2017)
Visual Properties: perform data mapping correctly, considering pre-attentive properties and Gestalt principles	Situation Awareness	2017	(Oliveira & Silva, 2017)
Visualization design should avoid, as much as possible, menus or other actions that take the user outside of the frame of the task.	Interaction & Utility	2009	(Greensmith et al., 2009)
Visualize information from multiple aspects	Situation Awareness	2011	(Wang et al., 2011a)

This page is intentionally left blank.

APPENDIX F – LIST OF DERIVED HEURISTICS ALONG WITH ORIGINAL HEURISTICS AND GUIDELINES

Table 31 shows the heuristics for the **Analytic Process** evaluation area and their original heuristics and guidelines.

Table 31 - Heuristics for the Analytic Process evaluation area and their original heuristics and guidelines.

AP1 - VAS shows key characteristics of data at a glance.	• The visualization shows key characteristics of the data at a glance ** SA ** (Wall et al., 2019)
AP2 - VAS makes data relationships noticeable.	• Make activity patterns visible, determine popular and neglected data regions. ** SA ** (Heer & Agrawala, 2008) • Aggregate information and show its patterns ** Utility and SA ** (Wang et al., 2011a) • The visualization facilitates perceiving relationships ** SA ** (Wall et al., 2019) • The visualization exposes individual data cases and their attributes ** SA ** (Wall et al., 2019) • The visualization promotes exploring relationships between individual data cases as well as different groupings of data cases ** Interaction & Utility ** (Wall et al., 2019) • Allow view relations among data. ** Utility & SA ** (Oliveira & da Silva, 2017)
AP3 - VAS provides a new or better understanding of the data.	• Data Characterization: assist data understanding. ** SA ** (Oliveira & da Silva, 2017) • The visualization promotes understanding data domain characteristics beyond the individual data cases and attributes ** SA ** (Wall et al., 2019) • The visualization helps identify unusual or unexpected, yet valid, data characteristics or values ** SA ** (Wall et al., 2019) • If there were data issues like unexpected, duplicate, missing, or invalid data, the visualization would highlight those issues ** SA & Utility ** (Wall et al., 2019) • Capture exogenous attention - alert users to important attributes of a visualization. ** SA & Utility ** (Tarrell et al., 2014)
AP4 - VAS helps generate data-driven questions.	• The visualization helps generate data-driven questions ** SA ** (Wall et al., 2019)
AP5 - VAS suggests relevant information beyond dataset information.	• Suggest related views, comments, and data to current points of interest. ** Collaboration & Utility ** (Heer & Agrawala, 2008) • Support notification subscriptions for views, artefacts, people, and groups. ** Collaboration & Utility ** (Heer & Agrawala, 2008)
AP6 - VAS features visualization which provides a comprehensive data overview with a meaningful visual schema.	• The visualization provides a comprehensive and accessible overview of the data ** SA ** (Wall et al., 2019) • The visualization presents the data by providing a meaningful visual schema ** SA ** (Wall et al., 2019) • Display information in consistent format ** SA ** (Wang et al., 2011a)
AP7 - VAS provides coordinated views for linked information.	• Construct coordinated views for linked information ** SA ** (Wang et al., 2011a).
AP8 - VAS displays related information nearby.	• Display coordinating information in close proximity. ** SA ** (Scholtz, 2011)

AP9 - VAS minimizes distractions for the analyst.	• Visualization design should avoid, as much as possible, menus or other actions that take the user outside of the frame of the task. ** Interaction & Utility ** (Greensmith et al., 2009) • Intimate interaction is both transparent and supportive of holistic cognition. ** SA & Utility ** (Greensmith et al., 2009) • Guide the analyst to follow the right trail, without distraction ** SA & Utility ** (Kang et al., 2009) • Use good aesthetics to minimize distractions and maximize perceptive effects Other perceptive aspects not represented above ** SA ** (Tarrell et al., 2014) • Guide endogenous attention - provide appropriate organization of material or interaction options to assist endogenous attention and minimize distracting information. ** Utility ** (Tarrell et al., 2014)
AP10 - VAS provides opportunities for serendipitous discoveries.	• The visualization shows multiple perspectives about the data ** SA ** (Wall et al., 2019) • The visualization uses an effective representation of the data that shows related and partially related data cases ** SA ** (Wall et al., 2019) • Visualize information from multiple aspects ** SA ** (Wang et al., 2011a)
AP11 - VAS allows flexibility in the organization of the visual metaphor.	• Allow flexibility in organization of schemes ** SA & Interaction ** (Kang et al., 2009)
AP12 - VAS facilitates finding starting points or clues.	• Help to provide information scent appropriately, thus helping to find initial clues ** SA ** (Kang et al., 2009) • Help to find appropriate next steps when encountering a dead-end ** SA ** (Kang et al., 2009) • Mixed initiative sensemaking systems should provide an environment in which users can capture information relevant to their ongoing task and refine their conceptual models ** SA ** (Cook et al., 2015)
AP13 - VAS provides strong retrieval cues for mental models.	• Aid analogical reasoning - structure information so as to provide strong retrieval cues for knowledge structures (mental models) to aid in analogical reasoning. ** SA & Utility ** (Tarrell et al., 2014) • Aid reasoning with mental models - organize information based on mental models so as to provide strong retrieval cues for knowledge structures in long-term memory to aid reasoning. ** SA & Utility ** (Tarrell et al., 2014). •
AP14 - VAS allows sharing of evidence and hypothesis.	• Ability to share evidence ** Collaboration ** (Adagha et al., 2017) • Support sharing of evidence and hypothesis ** Collaboration ** (Wang et al., 2011a) • Enable in-app collaborative editing ** Collaboration ** (Wang et al., 2011a) • Structure collaboration through shared, editable representations. ** Collaboration ** (Heer & Agrawala, 2008) • Support creation and export of presentations for telling analysis stories. ** Collaboration ** (Heer & Agrawala, 2008) • Support embedding of annotated views in external media (e.g., email, blogs, and reports) ** Collaboration ** (Heer & Agrawala, 2008) • Enable lightweight sharing of views across media with bookmarks (e.g., URLs) ** Collaboration ** (Heer & Agrawala, 2008)
AP15 - VAS supports collection of evidence and annotations in a beneficial organization to sensemaking.	• Support storytelling and enable interactive grouping of the evidence with users reasoning logic ** Interaction & SA ** (Wang et al., 2011a)

	• Facilitate chunking - choose visualization parameters that provide strong grouping cues to facilitate the chunking of information, which will minimize the effects of working-memory capacity limitations. ** Interaction & SA ** (Tarrell et al., 2014) • Allow evidence collection and annotation ** Utility & SA ** (Wang et al., 2011a) • Support creation and export of presentations for telling analysis stories. ** Collaboration ** (Heer & Agrawala, 2008)
AP16 - VAS allows registering need for more evidence or other future actions.	• Mark needed future actions: unanswered questions, need for evidence, etc. ** Utility & SA ** (Heer & Agrawala, 2008)
AP17 - VAS supports sensemaking by recommending relevant information.	• Mixed-initiative iterative sensemaking environments should recommend relevant data based on the user's current activity and potential next step ** Interaction & SA ** (Cook et al., 2015) • Provide recommendations in context ** Utility & SA ** (Cook et al., 2015)
AP18 - VAS displays statistics and measures about data sources, datasets, and/or records.	• Empower with Numbers ** Utility & SA ** (Kang & Stasko, 2012)
AP19 - VAS features a visual display of the analytic process.	• A visual display of the process used is excellent as it relieves the analyst from having to keep external notes. ** Utility & SA ** (Scholtz, 2011)
AP20 - VAS provides an easy-to-interpret environment for contextual analysis with relevant information.	• Can provide an environment for contextual analysis ** SA ** (Adagha et al., 2017) • Enable rapid visual interpretation of recommendations. ** SA ** (Cook et al., 2015) • The visualization uses meaningful and accurate visual encodings to represent the data ** SA & Utility** (Wall et al., 2019) • Suggesting what may have been overlooked and keeping relevant information present. ** SA & Utility ** (Greensmith et al., 2009) • Mixed initiative sensemaking systems should provide an environment in which users can capture information relevant to their ongoing task and refine their conceptual models ** SA ** (Cook et al., 2015) • Increase engagement by increasing personal relevance of data sets. ** SA ** (Heer & Agrawala, 2008)
AP21 - VAS provides transparent automation to the user regarding the underlying mathematical models and parameters.	• If the system contains automation, it should be well balanced and transparent to the user. ** Utility ** (Scholtz, 2011) • Shield users from the complexity of the underlying mathematical models and parameters. ** Utility ** (Endert, 2014)
AP22 - VAS captures and understands user interactions.	• Systems must be able to capture and understand user actions ** Interaction & SA ** (Endert et al., 2015) • Trace interactions and system usage for future automation ** Interaction & SA ** (Wang et al., 2011a) • Use semantic interactions within the visual metaphor, based on common interactions occurring in spatial analytic processes such as searching, highlighting, annotating, and repositioning documents. ** Interaction & SA ** (Endert, 2014). •
AP23 - VAS makes inferences from user interactions.	• Systems must be able to make inferences based on users' interactions. ** Interaction & SA ** (Endert et al., 2015) • Infer users' tasks based on their activities ** Interaction & SA ** (Cook et al., 2015).

AP24 - VAS reacts and takes initiative based on inferences from user interactions.	• Systems must be able to react and take initiative based on the inferences at three levels: interface, computation, and cognitive. ** Interaction & SA ** (Endert et al., 2015)
AP25 - VAS provides visual feedback regarding the updated model.	• Provide visual feedback of the updated model and learned parameters within the visual metaphor. ** Interaction & SA ** (Endert, 2014) • System Status and Feedback: notify users about the system status, and always provide quick and proper feedback ** Usability ** (Oliveira & da Silva, 2017).
AP26 - VAS features teamwork management.	• Allows optimal group size determination (improve efficiency of analysis). ** Collaboration ** (Heer & Agrawala, 2008) • Division of labour among participants. ** Collaboration ** (Heer & Agrawala, 2008) • Group creation and management mechanisms. ** Collaboration ** (Heer & Agrawala, 2008).
AP27 - VAS features activity indicators per collaborator increasing so engagement.	• Has activity indicators or summaries (aid reputation and visibility of contributions). ** Collaboration ** (Heer & Agrawala, 2008) • Personal action histories allow past contributions to be assessed ** Collaboration ** (Heer & Agrawala, 2008) • Enable identification of collaborators in a contextually appropriate manner. ** Collaboration ** (Heer & Agrawala, 2008)
AP28 - VAS supports intuitive communication among collaborators.	• Support commentary; consider implications of discussion model on common ground. ** Collaboration ** (Heer & Agrawala, 2008) • Can support intuitive communication ** Collaboration ** (Adagha et al., 2017)
AP29 - VAS supports future scenario projections.	• Can support future scenario projections ** Utility & SA ** (Adagha et al., 2017)
AP30 - VAS integrates multiple information channels.	• Unified content interface ** Interaction & SA ** (Wang et al., 2011a) • Should integrate multiple information channels ** Interaction & Usability ** (Wang et al., 2011a)
AP31 - VAS increases engagement and attention by using game design elements.	• Use game design elements to provide incentives and to direct effort. ** Creativity ** (Heer & Agrawala, 2008)
AP32 - VAS allows to track the status of collaborative threads.	• Present status update of collaborative threads ** Collaboration ** (Wang et al., 2011a) • Can track information flows ** Collaboration ** (Adagha et al., 2017)

Table 32 shows the heuristics for the **Visualization Quality** evaluation area and their original heuristics and guidelines.

Table 32 - Heuristics for the Visualization Quality evaluation area and their original heuristics and guidelines.

VQ1 - VAS facilitates perception via Gestalt principles.	• Guide perception via Gestalt principles. ** SA ** (Tarrell et al., 2014) • Use color to maximize perceptive effects ** SA ** (Tarrell et al., 2014) • Visual Properties: perform data mapping correctly, considering pre-attentive properties and Gestalt principles ** SA ** (Oliveira & da Silva, 2017)
VQ2 VAS provides visualizations with meaningful spatial organization.	• The visualization provides a meaningful spatial organization of the data ** SA ** (Wall et al., 2019) • Spatial Organization and Perspective: care the visualization overall layout, as well as provide change of perspective. ** SA ** (Oliveira & da Silva, 2017) • Guide perception using pre-attentive attributes such as spatial position. (Tarrell et al., 2014)
VQ3 - VAS avoids dense visualizations by featuring properties for size and distance to avoid dense visualizations	• Use appropriate and easy to interpret representations for properties such as size and distance in visualizations. ** Usability ** (Scholtz, 2011) • Perceived ease of use ** Creativity ** (Adagha et al., 2017)
VQ4 - VAS uses animation only to show an effect that moves over time.	• Use animations only to show an effect that moves over time. Give analysts control to manipulate the speed of the animation. ** Interaction & Usability ** (Scholtz, 2011)
VQ5 – VAS displays only relevant information and elements in a straightforward fashion.	• Deliver contents in straightforward representation ** Utility ** (Wang et al., 2011a) • They must display only relevant information and elements to the user ** Utility ** (Oliveira & da Silva, 2017) • The visualization avoids using misleading representations ** Utility ** (Wall et al., 2019)

Table 33 shows the heuristics for the **Interactivity** evaluation area and their original heuristics and guidelines.

Table 33 - Heuristics for the Interactivity evaluation area and their original heuristics and guidelines.

I1 - VAS features self-descriptive interactions.	• Self-descriptiveness ** Interaction ** (Adagha et al., 2017)
I2 - VAS provides tools for data manipulation.	• Data set reduction (including filtering, clustering, and pruning) ** Interaction ** (Tarrell et al., 2014) • Data Manipulation: provide tools for data manipulation, such as filters and detailed view ** Interaction ** (Oliveira & da Silva, 2017) • Enable facet filtering for information personalization ** Interaction ** (Wang et al., 2011a)
I3 - VAS provides capabilities for data exploration.	• Enable access to information ** Interaction ** (Adagha et al., 2017) • Navigation and Querying ** Interaction ** (Tarrell et al., 2014)

	• The visualization provides useful interactive capabilities to help investigate the data in multiple ways ** Interaction ** (Wall et al., 2019) • The visualization supports smooth transitions between different levels of detail in viewing the data ** Interaction ** (Wall et al., 2019) • Spatial Organization and Perspective: care the visualization overall layout, as well as provide change of perspective. ** SA ** (Oliveira & da Silva, 2017) • Data Manipulation: provide tools for data manipulation, such as filters and detailed view ** Interaction ** (Oliveira & da Silva, 2017) • Interactive content exploration and filtering ** Interaction ** (Wang et al., 2011a)
I4 - VAS allows interactive visualization customization.	• Support customization of information ** Interaction ** (Adagha et al., 2017) • The interface supports using different attributes of the data to reorganize the visualization's appearance ** Interaction ** (Wall et al., 2019) • Multidimensionality: allow users to visualize three or more dimensions simultaneously ** Interaction ** (Oliveira & da Silva, 2017)
I5 - VAS avoids complex commands and queries in visualizations.	• The visualization avoids complex commands and textual queries by providing direct interaction with the data representation. ** Interaction & Utility ** (Wall et al., 2019)
I6 - VAS provides a way to backtrack and undo actions.	• Orientation and Help (control of level of details and support for undo) ** Interaction & Usability ** (Tarrell et al., 2014) • User Control: enable full system control by user. ** Interaction & Usability ** (Oliveira & da Silva, 2017) • Provide a way to backtrack or undo actions. ** Interaction & Usability ** (Scholtz, 2011) • Can track changes in information ** Interaction & SA ** (Adagha et al., 2017) • Provide histories of actions performed on artifacts (representations/visualizations) ** SA ** (Heer & Agrawala, 2008) • Controllability ** Interaction & Usability ** (Adagha et al., 2017)
I7 - VAS provides alternative ways to perform a task.	• Reduction in time ** Utility & Usability ** (Adagha et al., 2017) • Flexibility and Efficiency: provide accelerators and customization features. ** Interaction & Usability ** (Oliveira & da Silva, 2017) • Flexibility: provide means to customize the interface and select the preferred way to accomplish a goal. ** Interaction & Usability ** (Pribeanu, 2017)
I8 - VAS provides interaction with minimal need for repetitive actions**.**	• Provide means for analysts to explore visualizations that do not require repetitive interactions on the part of the analyst. ** Interaction & Usability ** (Scholtz, 2011) • Minimal actions: Minimize the number of actions needed to accomplish a task's goal. ** Interaction & Usability ** (Pribeanu, 2017)
I9 - VAS gives proper feedback to user actions within reasonable time.	• Feedback: Provide appropriate feedback as a response to user's actions within reasonable time. ** Interaction & Usability ** (Pribeanu, 2017)
I10 - VAS guides users towards making specific actions	• Prompting: Guide users towards making specific actions. ** Interaction & Usability ** (Pribeanu, 2017) • Visible Actions: make all possible actions visible. ** Usability ** (Oliveira & da Silva, 2017)

I11 - VAS facilitates the understanding of relationships between the various user interface items.	• Grouping/distinction: Provide means to group similar objects and distinguish between different classes of objects. ** Interaction & Usability ** (Pribeanu, 2017)

Table 34 shows the heuristics for the **User-friendliness** evaluation area and their original heuristics and guidelines.

Table 34 - Heuristics for the User-friendliness evaluation area and their original heuristics and guidelines.

UF1 - VAS provides coherent UI elements	• Consistency: Provide similar meanings and design choices in similar contexts. ** Usability ** (Pribeanu, 2017) • Consistency: the interface elements must be coherent. ** Usability ** (Oliveira & da Silva, 2017) • Perceived ease of use ** Creativity ** (Adagha et al., 2017)
UF2 - VAS features a UI with familiar signs to the user.	• Real World Equivalency: use familiar signs to the user. ** Usability ** (Oliveira & da Silva, 2017) • Perceived ease of use ** Creativity ** (Adagha et al., 2017)
UF3 - VAS matches user characteristics with the UI characteristics	• Compatibility with the user: Provide means to match the users' characteristics with the characteristics of the user interface. ** Interaction & Usability ** (Pribeanu, 2017) • Perceived ease of use ** Creativity ** (Adagha et al., 2017)
UF4 - VAS provides customizable workspaces:	• Flexibility: provide means to customize the interface and select the preferred way to accomplish a goal. ** Interaction & Usability ** (Pribeanu, 2017) • Adaptive personal workspaces and preferences ** Interaction & Usability ** (Wang et al., 2011a)
UF5 - VAS makes easily visible all possible actions for the user	• Help and documentation: Provide online help and documentation. ** Usability ** (Pribeanu, 2017) • Task guidance and support: Provide the user with the procedure and associated support (forms, documents, etc.) needed to perform specific tasks. ** Usability ** (Pribeanu, 2017) • Visible Actions: make all possible actions visible. ** Usability ** (Oliveira & da Silva, 2017) • Perceived ease of use ** Creativity ** (Adagha et al., 2017) • Invest in tutorials ** Usability ** (Kang & Stasko, 2012)

Table 35 shows the heuristics for the **Satisfaction** evaluation area and their original heuristics and guidelines.

Table 35 - Heuristics for the Satisfaction evaluation area and their original heuristics and guidelines.

S1 - VAS is effective in representing high-quality of analytical outcomes.	• Ability to show high quality of analytic solutions ** Creativity ** (Adagha et al., 2017) • Maximize effectiveness: maximize the extent to which the goals of the users are achieved ** Usability ** (Tarrell et al., 2014)
S2 - VAS ensures the end user's subjective assessment is overall positive.	• User satisfaction with solutions ** Creativity ** (Adagha et al., 2017) • Provide satisfaction ensure the user s subjective assessment is generally positive. ** Creativity ** (Tarrell et al., 2014)
S3 - VAS is considered highly useful.	• Perceived usefulness ** Utility ** (Adagha et al., 2017)
S4 - VAS minimizes the needed resources to achieve the goal.	• Maximize efficiency: minimize the resources necessary to achieve the goal. ** Usability ** (Tarrell et al., 2014)

Table 36 shows the heuristics for the **Error-handling** evaluation area and their original heuristics and guidelines.

Table 36 - Heuristics for the Error-handling evaluation area and their original heuristics and guidelines.

EH1 - VAS prevents, diagnoses, and correct errors.	• Error Prevention: prevent error occurrence, eliminating error-prone conditions. ** Interaction and Usability ** (Oliveira & da Silva, 2017) • Error Correction: inform users about errors occurred with clear messages and present means to correct these errors. ** Interaction and Usability ** (Oliveira & da Silva, 2017) • Error management: Provide means to prevent, diagnose, correct, and recover from errors. ** Interaction and Usability ** (Pribeanu, 2017)

Table 37 shows the heuristics for the **Adequacy** evaluation area and their original heuristics and guidelines.

Table 37 - Heuristics for the Adequacy evaluation area and their original heuristics and guidelines.

A1 - VAS is adequate to its context of use.	• Compatible with the context of use ** Utility ** (Adagha et al., 2017)