Nijmegen School of Management
Department of Economics and Business Economics
Master's Thesis Economics (MAN-MTHEC)

Data visualization: a different perspective

By Maks Nas (4742524)

Nijmegen, 30 November 2022

Program: Master's Program in Economics
Specialisation: Corporate Finance and Control
Supervisor: drs. R.H.R.M. Aernoudts

Radboud Universiteit

Abstract

Data visualization is one of many ways to detect potential financial fraud. It can be used to detect outliers by changing the visual representation of a dataset, which is generated from a company's IT systems. This research tries to take a different perspective on data visualization by conducting an experiment at accounting students to test whether graphical and photographical forms are more efficient and effective for detecting potential fraud than tabular forms. To measure efficiency, time is tracked to measure how fast respondents answer each question. To measure effectiveness, the accuracy of answering questions correctly is measured. As the differences between three representations are partly normal distributed, both Wilcoxon paired signed rank tests and paired t-tests are used to find whether the median or mean of differences between the pairs of observations are zero or not. Both significant and insignificant results are found. The most important finding is that the use of both graphical forms and photographical forms lead to quicker fraud identification processes. However, the time it took to create these data representations is not considered. It would highly depend on the skills of the investigator to realize time gain. More research in the field of data visualization is necessary to find differences between different forms of data visualization.

Table of Contents

1 Introduction

Yearly, billions of euros are lost by occupational fraud, according to a recent report by the Association of Certified Fraud Examiners (ACFE, 2022). Occurring in more than one hundred countries in about two thousand cases, fraud by employees is likely to be the most common form of financial crime in the world (ACFE, 2022). The importance of both fraud detection and prevention requires no debate, however, auditors and regulatory forces keep struggling with preventing large scandals from happening. Scientific research can help practitioners by examining new ways of fraud prevention, create new means of fraud detection and make better understanding of trends that occur in the field of financial fraud. Financial fraud can be very complex and has different sources: from banks to manufacturing companies to health care institutions (West & Bhattacharya, 2016). During the past two decades, audit firms have changed rapidly in how they perform their services (Ghasemi et al., 2011). Information technology created new opportunities for quicker audit procedures as larger amounts of data could be processed by computers (West & Bhattacharya, 2016). The combination of IT with fraud detection has been a matter of interest before, however, scientific literature lacks practical implications. This lack of practical awareness has created a wedge between scientists and practitioners. Research think it has proven the usefulness of different IT solutions to fraud detection, to be specific, it has given a great value to the use of data visualization tools in the search for financial fraud. However, from experience, the practical use of these data visualization tools seems limited. Fraud detection is still based on relatively older programmes such as IDEA and SAP, which can be classified as static forms of representation. Outlier detection is mostly done with the help of samples. These samples are selected manually (with smaller populations) or automatically (with larger populations). The research of Lee et al. (2018) is proof for this. Tableau, a programme that is frequently used for data visualization, was the lesser used software tool by accounting professionals. Excel, that is used as a basis for other tools such as IDEA and SAP, was by far the most used software tool. This thesis can help change the perspective on data visualization by showing the practical effectiveness of different data representations in detecting potential fraud.

This thesis aims to answer the following research question:

“Do graphical form representations and photographic representations enable quicker and more accurate fraud identification processes than tabular representations?”

An experiment is conducted among students that have affinity with accounting. As stated in the research by Dilla & Raschke (2015), the effectiveness of research in data visualization is limited, due to the small number of available professionals. Research into data visualization in general is upcoming, but it is difficult to find experiments about data visualization in fraud detection and audit procedures that test specific usages of visualization software. One of very few examples of research in data visualization in fraud detection is the research of Leite et al. (2018), which discuss different approaches to data visualization and compare them in different fraud categories.

As fraud identification is mainly based on judgment, it is important that the experiment is conducted on a representative group for conducting the fraud identification process. However, to ensure this representativeness, only low- and moderate-complexity tasks will be used in the experiment, as it is unlikely that students in accounting are able to perform high complexity fraud identification tasks that require more experience in the field. By using cognitive fit theory by Vessey (1991), the causal relation between problem representation and problem solution can be measured. This thesis uses cognitive fit in a slightly different manner: as it focusses on the judgment of a professional, the focus is on the effectiveness and efficiency of the task. This results in two dependent variables: accuracy and speed. In practice, firms demand professionals to be efficient and effective with their resources, especially regarding time, as the use of a professional for audit procedures is very costly. This directly ensures the practical relevance of this research: it can be a benchmark for further research and experiments in data visualization but can also be a factor of change in the day-to-day procedures of an accounting professional.

The experiment consists of five different fraud predications, about which representations in tabular form, graphical form and photographical form are shown per predication. These three representations form the independent variables of the thesis. The fraud predications are created with a thought experiment (Albrecht et al., 2012). The representations are created from the same dataset, which consists of transactions from a fictive metal company that imports metal from around Europe. Wilcoxon paired signed rank tests and paired t-tests are used to compare the median or mean differences of the dependent variables between the different pairs of representations. As the data is partly non-normal, several choices regarding outliers and form of testing had to be made.

This thesis tries to answer the research question in different chapters. In Chapter 2, the theoretical background of the fraud identification process is discussed. Audit procedures make use of a firm’s information technology systems and understanding how organizations use technology in their business is essential to understanding the fraud identification process that is designed by auditors. When zooming into fraud identification processes, anomaly detection is one of several ways to identify possible fraud. Data visualization can be a useful tool for detecting anomalies. Current literature on anomaly detection and data visualization is explained further. In Chapter 3, the appearance of the experiment is explained in more detail. As it is difficult to find suitable subjects for the experiment, some difficulties had to be solved. On top of this, the variables of interest and two propositions are formed. The results of the experiment and the statistical tests are shown in Chapter 4. The results help answering the research question in Chapter 5. The conclusion is also placed in the perspective of the theoretical perspectives of Chapter 2 to find practical relevance for the results. Finally, in Chapter 6, the thesis is reviewed critically. Different merits and demerits are discussed and a benchmark for further research and experiments is formed. Limitations of this research can help future research improve the realm of data visualization.

2 Theoretical background

Information technology has been changing firms into professional service industries rapidly from the beginning of the century. IT is according to these scholars: "*anything that renders data, information or perceived knowledge in any visual format whatsoever, via any multimedia distribution mechanism"* (Ghasemi et al., 2011, p. 113). Not only the efficiency and accuracy of information improved, but also the lead times of presenting financial information to stakeholders. During the past decade, datasets became larger. IT systems were able to process these larger datasets and a new phenomenon in research raised: big data. Big data is *"generated from an increasing plurality of sources including internet clicks, mobile transactions, user-generated content and social media as well as purposefully generated content through sensor networks or business transactions such as sales queries and purchase transactions"* (George et al., 2014, p. 322). This creates the opportunity for accounting and finance to use big data as input for research as well (Cockcroft & Russell, 2018).

This chapter explains this increasing importance of IT in both business and accounting and accordingly the intertwined existence of Enterprise Resource Planning Systems (hereafter: ERP systems) and Accounting Information Systems (hereafter: AIS). This creates a perspective on the process of fraud detection and how current AIS use ERP systems to aid their search of fraud. Data visualization is a very specific means of detecting anomalies in data and can be used to detect potential fraud. The process of extracting data from an ERP system, turning it into accounting information and using this information with data visualization to detect potential financial fraud is explained in that order, after which a real-life example is used to clarify the use of data visualization.

2.1 ERP and AIS

Ghasemi at al. (2011) mention two different reasons for the rapid change in organizations. Firstly, knowledge-sharing applications have become dominant in business. In literature, knowledge-sharing applications are part of a company's ERP systems. ERP systems were created to integrate all corporate information in one database (Dechow & Mouritsen, 2005). Another definition of ERP is “the attempt to integrate all departments and functions across a company onto a single computer system” (Koch et al., 1999). Robey et al. (2002) mention ERP as a system that integrates a set of software modules, all linked to one database. The takeaway from these definitions is integration. By combining all different functional areas of organizations, ERP systems should improve organizational efficiency, effectiveness and by doing so improve organizational performance (Arnold, 2006). A complete ERP system should automate business processes and provide decision makers with real-time data (Spathis & Constantinides, 2004). Spathis & Constantinides (2004) showed that during the past couple of decades, IT has become an increasingly important factor in business. However, during the time of writing of their research, most organizations were still looking into how to implement changes regarding ERP. As of today, ERP systems are almost completely embedded in both business as the accounting profession and investments in ERP remain large, according to the report ERP Adoption Trends and Customer Experience (2019).

Secondly, audit software has been an important factor of this change. In literature, audit software is more commonly known as Accounting Information Systems (AIS). An AIS is an integrated framework that uses physical resources to transform economic data into financial data (Bodnar & Hopwood, 2004; Wilkinson, 1999). The ERP systems gather and centre all information, after which an AIS transforms the data into financial information. This information can be used for decision making by management (Romney and Steinbart, 2012) and for external stakeholders of organizations (Kieso et al., 2012). Bachmid (2016) created a theoretical framework supported by a large array of literature that better Accounting Information Systems’ quality leads to better accounting information quality. For this reason, it is very important for auditors and other users of accounting information to understand how AIS are embedded in organizations.

The increasing importance of both ERP systems and AIS is emphasized by new International Standards on Auditing. ISA 315 is revised by the IAASB in September 2019. This change in regulatory requirements for auditors magnifies the increasing importance of IT. ISA 315 requires auditors to gain more understanding of the entity’s IT systems and its potential risks of the use of these IT systems. Thus, auditors must identify an organization’s AIS and ERP systems and gain an understanding of how the organization uses these systems to create information for internal and external use. This underlines the above assumptions that businesses have changed rapidly into IT-driven organizations and auditors must adjust their procedures accordingly.

The extent to which IT and big data is used to gather information around financial fraud has been subject of research over the past decade. Different perspectives on this have been given by West & Bhattacharya (2016), Bhattacharyya et al. (2011), Al-Hashedi & Magalingam (2021) and Vasarhelyi et al. (2015). Vasarhelyi et al. (2015) has provided a framework of big data in accounting in which a collection of essays is provided. These essays show the transformation of corporate data into big data and the difficulties that come with this transformation. West & Bhattacharya (2016) specify these issues with big data and research the practicalities of methods to analyse large chunks of data. They research over fifty scientific papers over the period of 2004 to 2014 to present a comprehensive classification on key aspects in fraud detection. Al-Hashedi & Magalingam (2021) present a similar comprehensive review of research in financial fraud from 2009 to 2019. They classify essays on types of fraud and data mining technology. They show that 34 different data mining techniques were used across different financial applications. These financial applications yielded bank fraud, insurance fraud, financial statement fraud and cryptocurrency fraud. Bhattacharyya et al. (2011) focus on credit card fraud. These different authors show the growing interest of big data in accounting.

2.2 Financial fraud detection

The importance of financial fraud detection and prevention can be noticed by every individual, as individuals encounter different governmental institutions that deal with matters of fraud in their lifetime. Further research into how this detection takes place is necessary, as matters of fraud get more complicated. Financial fraud is described by Reurink (2018, p. 1293) as: *"acts and statements through which financial market participants misinform or mislead other participants in the market by deliberately or recklessly providing them with false, incomplete or manipulative information related to financial goods, services or investment opportunities in a way that violates any kind of legal stipulation, be it a regulatory rule, statutory law, civil law, or criminal law".* This is a comprehensive approach to fraud, as it focusses on all forms of fraud and to its lawful form.

Auditing standards described by the IAASB (IAASB, 2021, p. 25) define fraud as *"an intentional act by one or more individuals among management, those charged with governance, employees, or third parties, involving the use of deception to obtain an unjust or illegal advantage"*.

Specific contents of financial fraud vary widely within different industries, which creates a difficulty to assess specific detection and prevention systems. Research on the topic has tried to differentiate between different phases and types of financial fraud detection, for example interviewing, document examination and the use of technology (Albrecht et al., 2012). As shown in section 2.1, the use of technology has become of greater importance in the financial industry. This use of technology has also influenced fraud identification. However, it is important to start with understanding the fraud identification process before going into further detail.

The identification process cannot start without an indication of fraud. Albrecht et al. (2012) explains this as *"circumstances, taken as a whole, that would lead a reasonable, prudent professional to believe a fraud has occurred".* According to theory, the investigator uses data mining techniques to find anomalies in a dataset, whereafter the investigator thinks about theories of how the fraud has occurred and by whom (Albrecht et al., 2012). As auditing firms rely heavily on journal entries as a source of information for data mining, large and complex sheets are produced as output (Gray & Debreceny, 2014). Research on data mining techniques by Ngai et al. (2011) has made a comprehensive overview of the field of financial fraud detection by using large amounts of data. Sharma & Panigrahi (2013) review literature on the application of data mining techniques. They provide a framework based on accounting fraud detection. Data mining can be defined as *"the process that uses statistical, mathematical, AI and machine learning techniques to extract and identify useful information and subsequently gain knowledge from a large database"* (Turban et al., 2007). Data mining can be classified into six different classes: classification, clustering, prediction, outlier detection, regression and visualization (Ngai et al, 2011; Sharma & Panigrahi, 2013).

- Classification: research literature describes classification as the process of identifying a set of common features or patterns in data and by doing so, describe and distinguish data in classes or concepts (Zhang & Zhou, 2004). This way, possible risky transactions can be spotted through predefined, discrete and unordered labels (Han & Kamber, 2006).

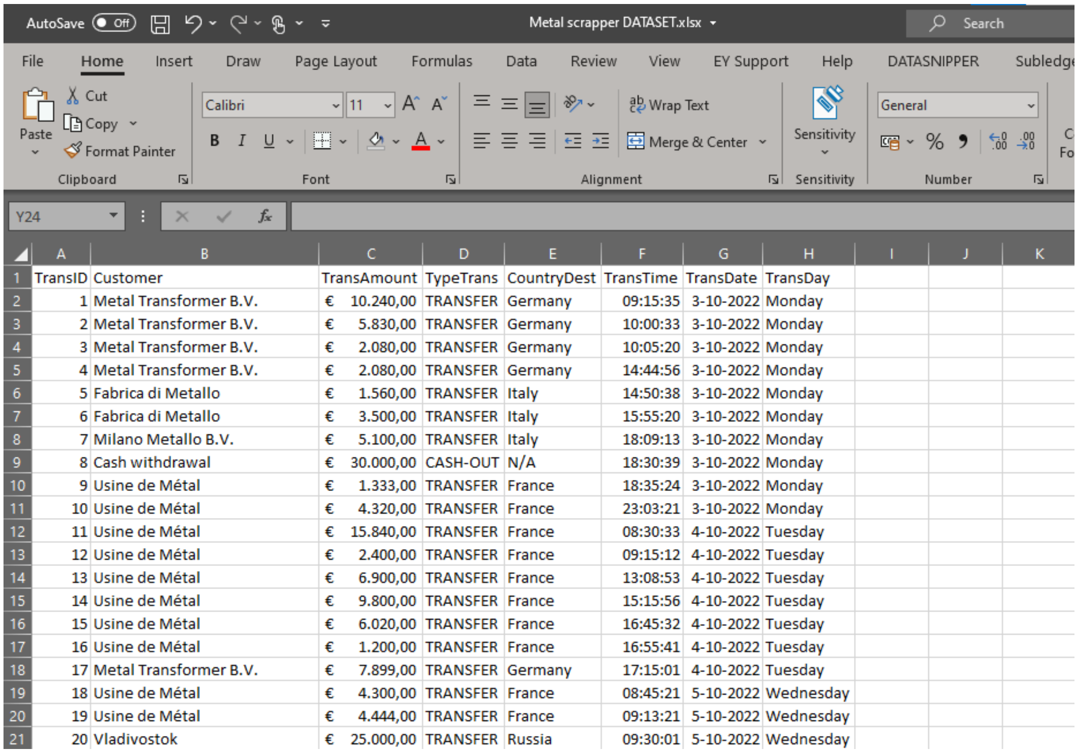
- Clustering: clustering is used to allocate objects into conceptually meaningful groups (called clusters). In other words, data is segmented into related clusters, but segmented data has low similarity with other clusters (Han & Kamber, 2006).
- Prediction: prediction estimates future values by using current patterns in the dataset (Han & Kamber, 2006).
- Outlier detection: outlier detection measures the distance between objects, to detect the objects that are very different or inconsistent with the other data (Han & Kamber, 2006). According to Hilal et al. (2022), outlier detection is interchangeably used with anomaly detection. This can be confirmed by Hawkins (1980), who defines an outlier as being an observation which deviates much from other observations.
- Regression: regression is a statistical methodology that uses causal relationships between independent and dependent variables (Han & Kamber, 2006).
- Visualization: visualization refers to presentations of data that changes difficult data characteristics into clear patterns. This eases the detection of outliers and recognizing certain patterns in data (Turban et al., 2007).

As the theory of these data mining techniques is relatively old, some notes have to be made. Data visualization today has become a combination of outlier detection and visualization (Leite et al., 2018; Dilla & Raschke, 2015). But it can also be used to combine clustering or classification and visualization (Ali et al., 2016), as neural networks can be used as input for visual representations as well. The differentiations made by Han & Kamber (2006) are not as distinct today as they have been described then and even after Ngai et al. (2011) and Sharma & Panigrahi (2013), data visualization has changed into a broader field of data analysis. Therefore, it is very important look for opportunities that data visualization could be able to fill.

2.3 Data visualization

Probable fraudulent activities could be easier to identify by using interactive data visualization, which is also proposed by Deloitte (2012). As Few & Edge (2007, p. 4) describe in their paper: *"Visual analysis software allows us to not only represent data graphically, but to also interact with* *those visual representations to change the nature of the display, filter out what's not relevant, drill into lower levels of detail, and highlight subsets of data across multiple graphs simultaneously. This makes good use of our eyes and assists our brains, resulting in insights that cannot be matched by traditional approaches."* Dilla & Raschke (2015) mention the difference between interactive data visualization and static interactive data visualization. This was elaborated on in the paper by Dilla et al. (2010), which explains this as an "on demand" visualization process that allows its users to navigate through and change the format on different levels of detail. This interactive data visualization allows more variability in the way data is analysed. Yi et al. (2007) created a theoretical framework in which information visualization systems consist of two components: representation and interaction. Yi et al. (2007) explain these components by providing examples using visualization tools and show not only the importance of visualizing data, but also show that interaction can create better understanding of the data. When zooming into the different representations that data visualization tools can create, Vessey (1991) made some clear statements about the table versus graph discussion. Most importantly, he emphasized that it is important to consider the type of task that is under investigation. However, more detailed research that specifies the differences between different forms of representation in the field of possible fraud detection is not present.

Many different data visualization tools exist (Caldarola & Rinaldi, 2017). This research uses Tableau, as it is very common in research (Ali et al., 2016; Batt et al., 2020; Lee et al., 2018). Tableau is easy to use in general and has lots of features from different older visualization tools (Batt et al., 2020). On top of this, it is interactive and different visualizations can be used. In this research, Tableau is used as a data mining tool. From the ERP system of a company, accounting information is generated in Excel (see figure 1). This information is generated from the purchases department and consists of cash withdrawals and payments to companies in different countries. Only the first twenty transactions are shown in the picture. As larger companies have many more daily transactions, it is difficult to just analyze data using Excel. The Excel-sheet is loaded into Tableau. Tableau can then be used to create the variables, which are shown in rectangle A. These variables can be dragged in rectangles B to automatically create different tables, graphs and photographs. In this example, the data is sorted on days of the week, to show how many transactions occurred each day and what the total size of the transactions was per day.



TransID	Customer	TransAmount	TypeTrans	CountryDest	TransTime	TransDate	TransDay
1	Metal Transformer B.V.	€ 10.240,00	TRANSFER	Germany	09:15:35	3-10-2022	Monday
2	Metal Transformer B.V.	€ 5.830,00	TRANSFER	Germany	10:00:33	3-10-2022	Monday
3	Metal Transformer B.V.	€ 2.080,00	TRANSFER	Germany	10:05:20	3-10-2022	Monday
4	Metal Transformer B.V.	€ 2.080,00	TRANSFER	Germany	14:44:56	3-10-2022	Monday
5	Fabrica di Metallo	€ 1.560,00	TRANSFER	Italy	14:50:38	3-10-2022	Monday
6	Fabrica di Metallo	€ 3.500,00	TRANSFER	Italy	15:55:20	3-10-2022	Monday
7	Milano Metallo B.V.	€ 5.100,00	TRANSFER	Italy	18:09:13	3-10-2022	Monday
8	Cash withdrawal	€ 30.000,00	CASH-OUT	N/A	18:30:39	3-10-2022	Monday
9	Usine de Métal	€ 1.333,00	TRANSFER	France	18:35:24	3-10-2022	Monday
10	Usine de Métal	€ 4.320,00	TRANSFER	France	23:03:21	3-10-2022	Monday
11	Usine de Métal	€ 15.840,00	TRANSFER	France	08:30:33	4-10-2022	Tuesday
12	Usine de Métal	€ 2.400,00	TRANSFER	France	09:15:12	4-10-2022	Tuesday
13	Usine de Métal	€ 6.900,00	TRANSFER	France	13:08:53	4-10-2022	Tuesday
14	Usine de Métal	€ 9.800,00	TRANSFER	France	15:15:56	4-10-2022	Tuesday
15	Usine de Métal	€ 6.020,00	TRANSFER	France	16:45:32	4-10-2022	Tuesday
16	Usine de Métal	€ 1.200,00	TRANSFER	France	16:55:41	4-10-2022	Tuesday
17	Metal Transformer B.V.	€ 7.899,00	TRANSFER	Germany	17:15:01	4-10-2022	Tuesday
18	Usine de Métal	€ 4.300,00	TRANSFER	France	08:45:21	5-10-2022	Wednesday
19	Usine de Métal	€ 4.444,00	TRANSFER	France	09:13:21	5-10-2022	Wednesday
20	Vladivostok	€ 25.000,00	TRANSFER	Russia	09:30:01	5-10-2022	Wednesday

Figure 1: Accounting information in Excel

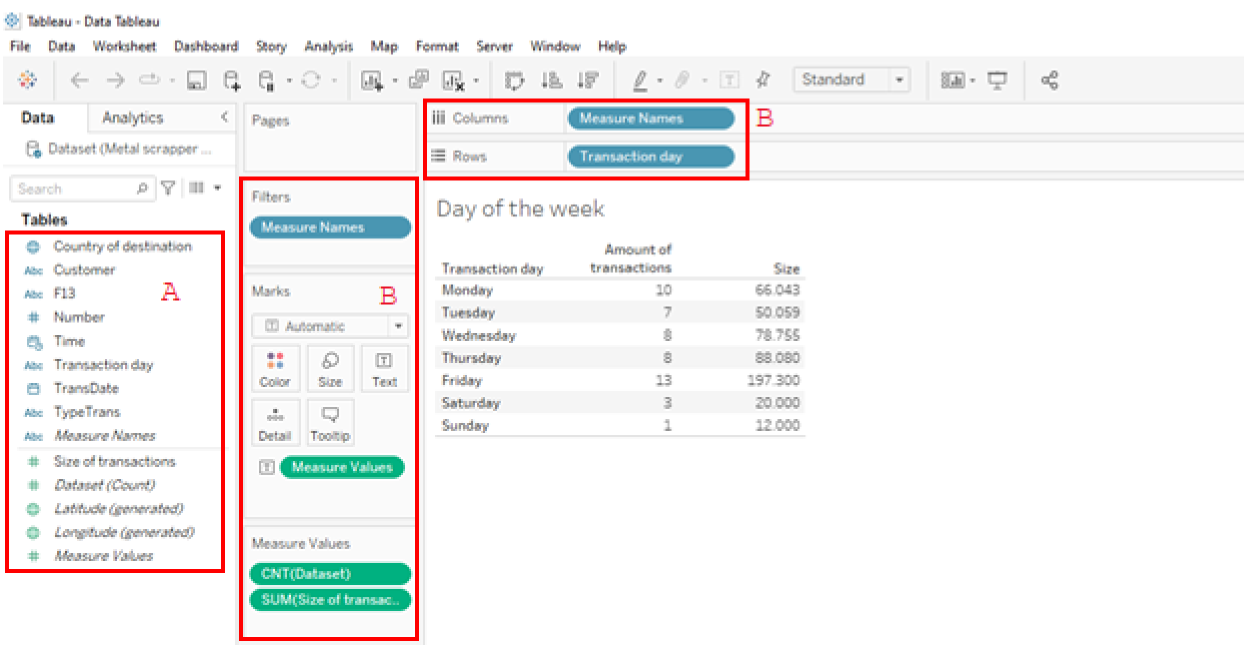


Figure 2: Representation created in Tableau

This chapter showed that as IT systems have risen in importance over the past decade, fraud identification has changed accordingly. Possible fraud identification is about analysing accounting information that is created by ERP systems within organizations. Data mining techniques are used to handle these large chunks of data in different manners. Data visualization can be used to identify possible fraudulent transactions by visualizing anomalies in data. Data visualization tools enable the user to make representations, but can also be used to interact with data. This thesis focuses merely on the representation part of data visualization and how different representations could improve the fraud identification process by detecting anomalies.

3 Methodology

3.1 Research design

This thesis uses the framework presented by Dilla & Raschke (2015), as it makes clear considerations about what research in the topic of data visualization should look like. This fits the purpose of this study best, as graphic representations are used as a contrary of tabular representations. Firstly, it is important to derive different values from the research question: *"Do graphical form representations and photographic representations enable quicker and more accurate fraud identification processes than tabular representations?"* The most significant issue for the experiment is its internal validity. The measurement of efficient and effective fraud identification processes is difficult. By using the predictive validity framework of Libby et al. (2002) this research uses a general theory to construct measurable variables from the research question. Cognitive fit theory (Vessey, 1991) has several important aspects, but its roots lie in the interaction between problem representation and problem-solving tasks that results in a problem solution. The outcome of the problem solution is the dependent variable, as it measures the performance of the solution. The performance of the solution is measured in two measurements: speed and accuracy. These are the dependent variables. The representations that are used in the experiment are the independent variables.

This thesis uses propositions proposed by Dilla & Rashke (2015) as a basis, however, numerous alterations are made. Firstly, "task complexity" is difficult to test for, as this experiment uses students as respondents. Students have limited knowledge of accounting problems and are presented with low- and/or moderate complexity tasks. As testing for task complexity would require a larger diversity of task difficulty, this is not possible to test for.

Secondly, the propositions of Dilla & Rashke (2015) test for certainty. As the propositions test about the *"presence or absence of fraudulent transactions"* a clear definition of what is fraudulent and what not is needed. However, it is difficult to determine whether a transaction is fraudulent or not, as the determination of this classification goes through many different phases. It is legally binding, so not only is fraud classified by auditors, but it also needs to be supported by law. As this research merely focusses on the detection of *possible* fraud, it investigates the detection of anomalies.

Lastly, Dilla & Rashke (2015) use interactive representation tools in their hypotheses. As explained in Section 2, these tools enable the user to alter the representation of the data. However, this is not possible within the limitations of this experiment. The respondents are provided with visual representations that are produced beforehand and no alterations can be made by the respondent. Due to these three reasons, the propositions this thesis tries to test look like the following:

Proposition 1:

Investigators will require less time to find anomalies in the presented data when they use graphical form representations and photographic representations as opposed to tabular representations.

Proposition 2:

Investigators will draw more accurate conclusions about finding anomalies in the presented data when they use graphical form representations and photographic representations as opposed to tabular representations.

As this research does not use a real treatment, there is no control group and test group created. This experiment benefits from many respondents and as the experiment can only be conducted at accounting students, it is difficult to find enough respondents. Therefore, the choice to do a quasi-experiment is made. All respondents are provided with the same questions and to ensure there is no bias involved by answering the questions in a particular order, there is always a different anomaly that is correct each question, but the origin of the anomaly is very similar. The respondent is asked to complete the experiment as quickly and accurate as possible. Time is tracked during the experiment. In the beginning of the experiment, a simple instruction is provided about what questions are asked and how the respondent can use the provided representations to find the correct answer. However, the explanation is short to ensure the respondents only use the provided representations to come to their answer. On top of this, the explanation is about a fraud predication that is not used in the experiment, to make sure the respondents are not biased before the experiment starts.

3.2 Experiment

The experiment consists of five fraud predications. These predications are based on the work of Albrecht et al. (2012) by thinking of possible ways fraud can occur. Each fraud predication comes with three questions, one with a tabular representation (figure 3, left), one with a graph representation (figure 3, right) and one with a photographic representation (figure 4). Tableau is used to create these representations, as explained in Section 2. The following questions are asked per fraud predication:

1. Day of the week: which day would you investigate further? This question has seven possible answers, knowingly: Monday, Tuesday, Wednesday, Thursday, Friday, Saturday, Sunday.
2. Size of the transaction: what transaction would you investigate into further detail? This question has four different answers, where each question has different answer possibilities to reduce the chance of bias.
3. Timing of the transaction: what transaction would you investigate into more detail? This question has four different answers, where each question again has different answer possibilities.
4. Country of destination: what transaction would you investigate into more depth? Again, this question has four different answers, differing per question.
5. Purpose of the transaction: what transactions would you investigate into more detail? This question has three different answers, again differing per question.

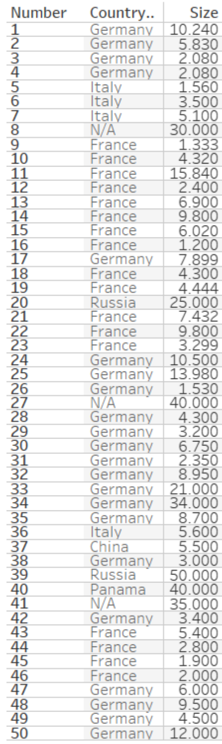
By creating more answer possibilities, the risk of guessing correctly or bias by looking at the answers first is reduced. Qualtrics is used to track time between arriving on the page of the representation and submitting the answer to the question. Answer codification is used to track whether the respondent has answered the question correctly. One data representation and one question is shown on each page and it is not possible to go back and forth between questions, as this would influence the timing aspect.

The availability of public datasets with financial transactions is limited, due to the intrinsically private roots of financial transactions. Because this research is not particularly about detecting fraud, but rather about using different representations to detect anomalies, there should be no issue in using datasets without particular private information. The presence or absence of information that is not real does not influence the knowledge or actions of the respondents doing the experiment, as they base their choices on the way the data is presented rather than the data itself that is presented. Therefor a dataset is created completely from scratch. The dataset consists of bank transactions from a fictive metal company that imports metal scrap from different countries in Europe to the Netherlands. These bank transactions have different sizes, dates, timestamps and target countries. A part of the dataset can be found in figure 1 in Section 2.

The experiment is conducted at students that have had experience with accounting courses during their Bachelor and/or Master degree. As the level of difficulty varies from easy to intermediate, they are a suitable subject for this experiment (Dilla & Raschke, 2015). Students that have affinity with accounting followed different accounting related topics in their study. Courses like Financial Accounting, Accounting, (Advanced) Bookkeeping or related studies are sufficient to understand how transactions within organizations work. They are assumed to have basic knowledge about journal entries, financial statements and have basic understanding of organizations that use financial transactions. However, as financial fraud detection is often not a subject for university degrees, a small explanation is given to guide the student through the experiment.

In the beginning of the experiment, several questions regarding basic characteristics are asked. The respondents are asked about whether they did an accounting related study, their age and their gender. These questions can help characterize the results and can help filtering out students that do not have any accounting background. It is likely that these respondents cannot answer the questions accordingly and are therefore filtered out of the results. It is not possible to go back and forth between questions and every person can only enter one survey per person. The survey can be found in the Appendix.

It is important to discuss what are considered correct and wrong answers. As explained in section 2, an anomaly is an object that is very different or inconsistent with the other data. An auditor approaches a dataset by creating a thought experiment about where fraud could potentially take place (Albrecht et al., 2012). For example, a manufacturing company has a normal operating cycle from Monday to Friday. Therefore, it would be out of the ordinary if transactions were done on Saturday or Sunday, as nobody would be in the office on that day. This is the first potential fraud indicator that was tested with the respondents. The second potential fraud indicator was size of the transaction. The metal scrap company has an average transaction size of around 10.000 EUR. A transaction much larger would be different than the normal course of business. The third potential fraud indicator is the timing of the transaction. As the normal operating cycle on a business day is from 8 AM to 6 PM, it would be very different from the ordinary if transactions took place late in the evening. The fourth potential risk indicator is the country of destination. The metal company imports metal scrap from different countries in Europe. Therefor it would be considered an anomaly if the company has financial transactions to Russia (that is heavily sanctioned by European Law) and Panama. Finally, the fifth risk indicator is the purpose of the transaction. All transactions are done via wire transfer, however, some out of the ordinary businesses show up in the descriptions. On top of that, large cash withdrawals are not usual for companies reliant on import.



Number	Country..	Size
1	Germany	10.240
2	Germany	5.830
3	Germany	2.080
4	Germany	2.080
5	Italy	1.560
6	Italy	3.500
7	Italy	5.100
8	N/A	30.000
9	France	1.333
10	France	4.320
11	France	15.840
12	France	2.400
13	France	6.900
14	France	9.800
15	France	6.020
16	France	1.200
17	Germany	7.899
18	France	4.300
19	France	4.444
20	Russia	25.000
21	France	7.432
22	France	9.800
23	France	3.299
24	Germany	10.500
25	Germany	13.980
26	Germany	1.530
27	N/A	40.000
28	Germany	4.300
29	Germany	3.200
30	Germany	6.750
31	Germany	2.350
32	Germany	8.950
33	Germany	21.000
34	Germany	34.000
35	Germany	8.700
36	Italy	5.600
37	China	5.500
38	Germany	3.000
39	Russia	50.000
40	Panama	40.000
41	N/A	35.000
42	Germany	3.400
43	France	5.400
44	France	2.800
45	France	1.900
46	France	2.000
47	Germany	6.000
48	Germany	9.500
49	Germany	4.500
50	Germany	12.000

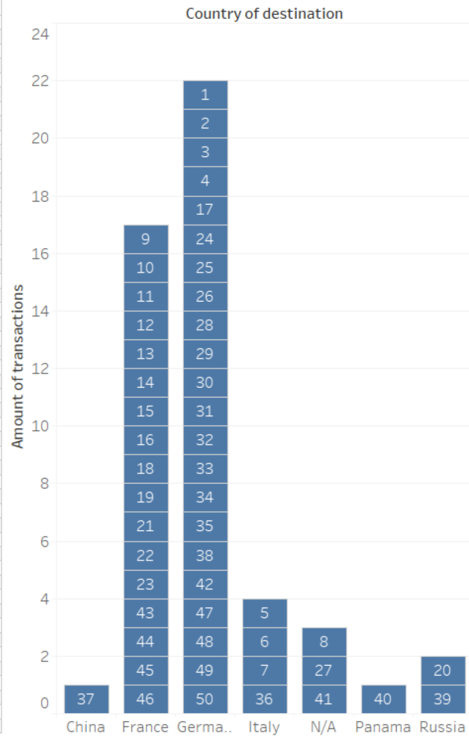


Figure 3: Examples of a table (left) and a graph (right)

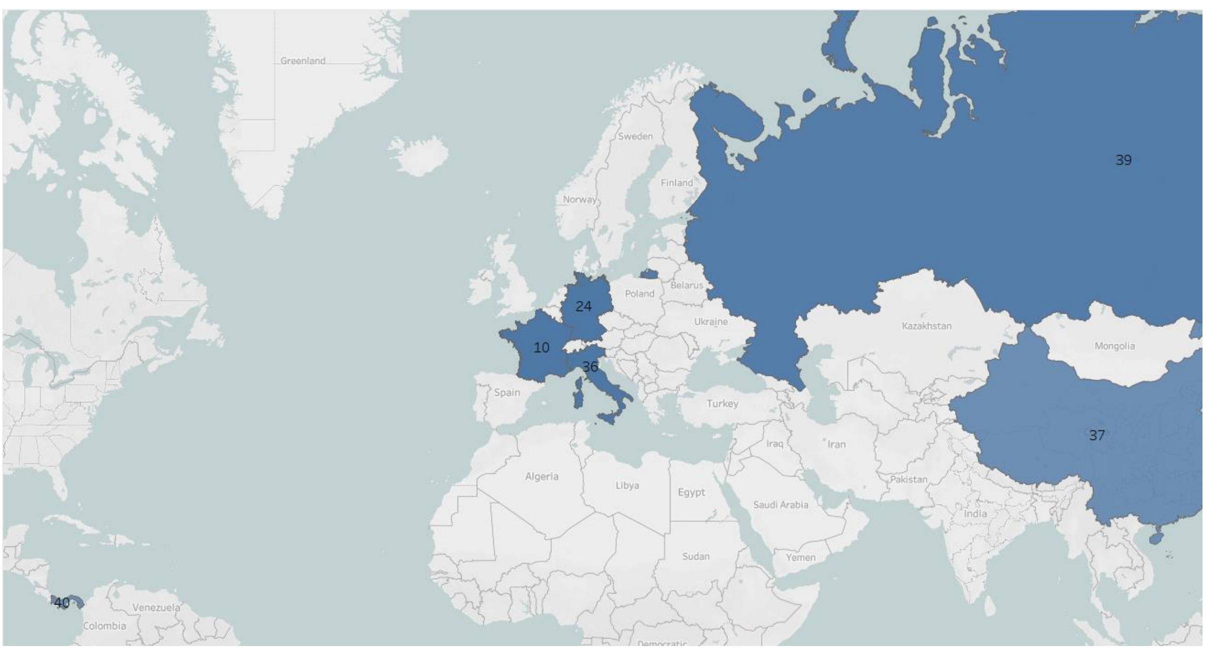


Figure 4: Example of a photographical representation

4 Results

In this chapter, the results of the experiment are discussed. A total of 51 responses were received of which 31 were male and 20 were female. Most of the respondents were doing their bachelor's degree in Business Administration. The youngest participant was 17 years old, the oldest 33 years old. People that did not have any affinity with accounting were filtered out of the experiment, as their response is not reliable enough for the experiment that is performed. It can be assumed that they do not have enough knowledge about financial transactions and do not know what possible fraudulent transactions can be. The risk of them guessing the answer and potentially affecting the results is large. This reduces the number of responses with 8, leading to a total of 43 observations. Initially, performing paired t-tests would suit the purpose of this thesis best, as this experiment creates just one sample and results in the sample create pairs (table vs graph or table vs photograph). To conduct a paired t-test, several assumptions must be met (Kim, 2015):

1. The subjects are independent;
2. The paired measurements are obtained from the same subject;
3. The distribution of differences is normally distributed with no severe outliers.

The first two assumptions are met for both dependent variables, as the respondents all took the same experiment and as the experiment was conducted via an online survey program, the respondents are likely not to have influenced each other. The respondents are also sampled randomly, as the experiment was spread in different large social media groups for accounting students. However, the normality assumption requires more analysis for both variables, which is done with descriptive analyses in section 4.2. First, some general comments on the data have to be made.

4.1 Descriptive analysis of variables

The descriptive analyses of the dependent variables accurateness and speed were done in STATA. Questions 1 to 3 are questions regarding the first fraud predication, questions 4 to 6 about the second fraud predication, questions 7 to 9 about the third fraud predication, questions 10 to

12 about the fourth fraud predication and 13 to 15 about the fifth fraud predication. If the respondent answered the question correctly, value 1 was used. If answered incorrectly, the value 0 was used. By calculating the mean of all question's answers, the overall understanding of the questions by respondents can be analyzed.

Table 1. Descriptive Statistics of accurateness

Variable	Obs	Mean	Std. Dev.	Min	Max
Question 1	43	.256	.441	0	1
Question 2	43	.14	.351	0	1
Question 3	43	.186	.394	0	1
Question 4	43	.674	.474	0	1
Question 5	43	.628	.489	0	1
Question 6	43	.837	.374	0	1
Question 7	43	.651	.482	0	1
Question 8	43	.744	.441	0	1
Question 9	43	.698	.465	0	1
Question 10	43	.884	.324	0	1
Question 11	43	.721	.454	0	1
Question 12	43	.837	.374	0	1
Question 13	43	.349	.482	0	1
Question 14	43	.488	.506	0	1
Question 15	43	.767	.427	0	1

Question 1 has a mean of 0.256, which means roughly 26% of the respondents answered the question correctly. Questions 2 and 3 are answered correctly by less than 20% of the respondents. In comparison with the other questions, the number of correct answers to the questions about the first fraud predication are answered much worse. Reason for this could be that the respondents did not understand the fraud predication correctly. However, as there are no large differences between the answering of the questions 1, 2 and 3, it would not hurt the trustworthiness of the results regarding speed. People identified the wrong outlier, however, the speed by which they come to a decision by using different representations remains the same.

Table 2. Statistics of speed

Variable	Obs	Mean	Std. Dev.	Min	Max
Question 1	43	25.933	15.04	4.85	65.26
Question 2	43	19.293	10.968	3.24	58.37
Question 3	43	17.982	15.373	5.25	95.3
Question 4	43	38.3	17.533	3.09	87.88
Question 5	43	28.246	13.803	2.55	70.56
Question 6	43	15.845	13.381	4.08	68.18

Question 7	43	38.866	51.886	3.07	335.68
Question 8	43	25.781	17.959	2.5	89.2
Question 9	43	30.625	36.403	1.43	234.32
Question 10	43	29.61	20.215	1.9	102.97
Question 11	43	32.482	53.006	2.09	357.42
Question 12	43	20.633	10.966	1.3	45.34
Question 13	43	21.964	10.275	2.07	47.57
Question 14	43	18.048	10.923	1.66	46.74
Question 15	43	11.969	10.102	1.75	54.45

Regarding the means of speed, there are no outliers. All questions are answered within averagely 40 seconds and as the expectation of the experiment was that it would take around five minutes, this is in line of expectation. However, the maximum observation of questions 7 and 11 are very high. It took one respondent more than 300 seconds to answer one question, which is a severe outlier taking the means into account. This could influence the results. On top of this, it violates one of the assumptions of the paired t-tests approach, as severe outliers could hurt the normality assumption. For this reason, the specific respondent is removed from the data. Another outlier is detected at question 9, with 234 seconds. However, the respondent has answered the other questions around the mean and overall, the total duration of all questions is not influenced severely by this outlier. As this experiment does not have many respondents, this observation is not removed from the data.

Table 3. Descriptive Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Correct answers	42	8.881	2.707	2	15
Total duration	42	361.085	148.526	51.32	742

42 respondents remain. When summarizing the number of correct answers of all respondents and the total duration that it took for the respondents to take the experiment, no severe outliers are detected. By average, people needed around 6 minutes to take the experiment, with a maximum of 12 minutes. Besides, the respondents answered almost 9 out of 15 questions correctly and no one failed all questions. Based on these numbers, no further action is taken.

4.2 Tests for normality

As mentioned previously, some issues with outliers had to be resolved. One of the assumptions of the paired t-test is that the paired differences of the means must be normally distributed. This is tested using the Shapiro-Wilk test at first. Four different variables are created to measure the mean differences between the three different forms of data representations:

Table 4. Shapiro-Wilk W test for normal data

Variable	Obs	W	V	z	Prob>z
Table - graph (speed)	42	0.813	7.668	4.299	0.000
Table - photo (speed)	42	0.891	4.486	3.168	0.001
Table - graph (accurateness)	42	0.985	0.631	-0.973	0.835
Table - photo (accurateness)	42	0.958	1.723	1.148	0.125
Graph - photo (speed)	42	0.894	4.370	3.113	0.001
Graph - photo (accurateness)	42	0.904	3.957	2.903	0.002

The Shapiro-Wilk W test shows that the two mean differences for variable speed are highly significant (p-value: 0.001). This means the H_0 hypothesis of the Shapiro-Wilk test is rejected and therefore the data is non-normal. This creates a problem for the paired t-test approach. The reason for the fact that the data is non-normal, seems to stem from more outliers in the data. This is shown in a QQ-plot (figure 9). The differences for the variable accurateness are normally distributed, which allows us to do a paired t-test.

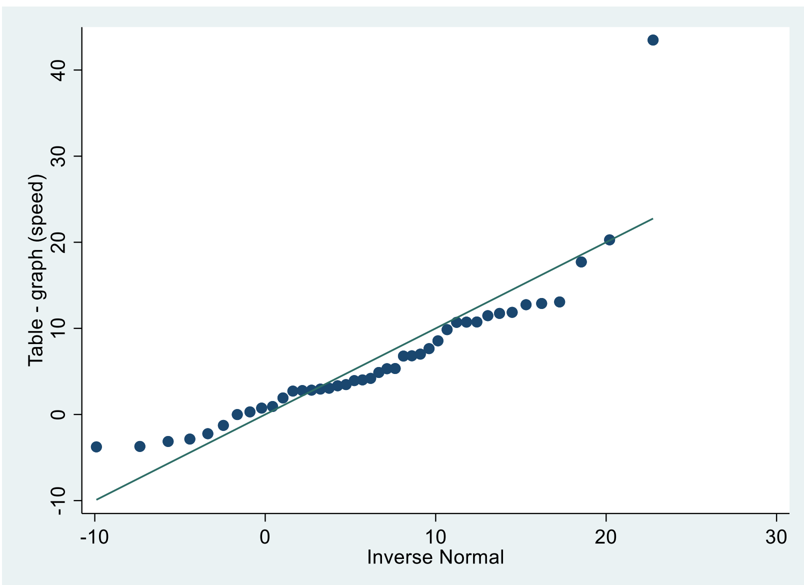


Figure 5: QQ-plot mean differences between table and graph (speed)

The QQ-plot is used to visualize the structure of the data. The differences between table and graph seem normally distributed as most observations are around the normal line, however, it is hard to tell how much the outliers influence the results of a paired t-test. Some literature argue that normality does not influence the results of a paired t-test if the sample size is not very small. Another option is to transform the data or delete outliers. However, transforming data makes it more difficult to analyze the results and deleting more observations would decrease the usefulness of the research (Stevens, 2012). Other research opts to do a Wilcoxon paired signed rank test instead (McCrum-Gardner, 2008). The Wilcoxon paired signed rank test is less powerful than the paired t-test. However, it is a non-parametric test whereby the distribution of the differences does not have to be normal (Pratt & Gibbons, 2012; Woolson, 2007). This trade-off between power of the test and achieving the assumptions is difficult to answer, but for the sake of this research the Wilcoxon paired signed rank test will do just fine in answering the hypotheses. The Wilcoxon paired singed rank tests take the medians instead of the means of the differences, however, as the data seems to be almost normally distributed, the mean and median are not expected to differ much.

4.3 Testing propositions

Using the two propositions formed in section 3.1 of this thesis, several specific hypotheses that are testable with the Wilcoxon paired signed rank test are formed to answer the research question. There are four different hypotheses, as the differences between tables and graphs, and tables and photographs are tested. The dependent variable speed is constructed by taking the mean of the number of seconds it took for respondents to answer questions about each representation. Thus, variable “tabtime” is the mean of seconds it took to answer questions with a tabular form (questions 1, 4, 7, 10, 13). Variable “graphtime” is the mean of seconds it took to answer questions with a graph (questions 2, 5, 8, 11, 14). The variable “phtime” is the mean of seconds it took to answer questions with a photograph (questions 3, 6, 9, 12, 15).

Table 5. Wilcoxon signed-rank tests regarding speed

H_0*: The median differences between tables and graphs are zero regarding speed.*

H_1*: The median differences between tables and graphs are not zero regarding speed.*

Sign	Obs	Sum ranks	Expected
Positive	7	73	451.500
Negative	35	830	451.500

Zero	0	0	0
All	42	903	903

Unadjusted variance 6396.25
Adjustment for ties 0.00
Adjustment for zeros 0.00
Adjusted variance 6396.25

H0: graphtime = tabtime
z = -4.733
Prob > z = 0.0000
Exact prob = 0.0000

H_0: The median differences between tables and photographical forms are zero regarding speed.

H_1: The median differences between tables and photographical forms are not zero regarding speed.

Sign	Obs	Sum ranks	Expected
Positive	5	92	451.500
Negative	37	811	451.500
Zero	0	0	0
All	42	903	903

Unadjusted variance 6396.25
Adjustment for ties 0.00
Adjustment for zeros 0.00
Adjusted variance 6396.25

H0: phtime = tabtime
z = -4.495
Prob > z = 0.0000
Exact prob = 0.0000

The results for both tests are highly significant (p = 0.000). This means the H_0 can be rejected and H_1 is true in both cases. The median differences between tables and graphs, and tables and photographs are not zero. The z-values (-4.377 and -4.495) show that the time it took to answer the question is less for both graphs and photographs in comparison with tables.

Table 6. Paired t-tests for accuracy

H_0: The mean differences between tables and graphs are zero regarding accurateness.

H_1: The mean differences between tables and graphs are not zero regarding accurateness.

	obs	Mean1	Mean2	dif	St Err	t value	p value
graphanswer - tabanswer	42	.547	0.567	-.019	.045	-.4	.678

H_0: The mean differences between tables and photographical forms are zero regarding accurateness.

H_1: The mean differences between tables and photographical formsares not zero regarding accurateness.

	obs	Mean1	Mean2	dif	St Err	t value	p value

phanswer - tabanswer	42	.662	0.567	.095	.043	2.25	.031

Regarding accuracy, the variables are created in the same manner as with speed. Thus, "tabanswer" is created by taking the mean of correct/incorrect answers for questions with a table. "Graphanswer" is the mean for questions with a graph and "phanswer" is the mean for questions with a photograph. However, as the differences were normally distributed (table 4), all the assumptions for the paired t-test are met. The paired t-test in figure 10 is insignificant (p = 0.678). This means the H_0 cannot be rejected and therefore the median difference between tables and graphs is zero. The second test is significant (p = 0.031), so the H_0 is rejected. The median difference is not zero regarding accurateness between tables and photographical forms. The z-value is 2.090, which is a positive number. This indicates that questions with photographical forms outperform questions with tables.

4.4 Additional testing

Although it is not part of the propositions tested, the graphical representations can be compared with the photographical representations. As shown in table 4, the differences between graphical and photographical representations for both speed and accuracy are non-normally distributed. This means the Wilcoxon paired signed rank test is used to analyze the differences.

Table 7. Wilcoxon signed-rank test for speed and accurateness (graph vs photo)

H_0: The median differences between graphs and photographical is zero regarding speed.

H_1: The median differences between graphs and photographical forms is not zero regarding speed.

Sign	Obs	Sum ranks	Expected
Positive	12	199	451.500
Negative	30	704	451.500
Zero	0	0	0
All	42	903	903

Unadjusted variance	6396.25
Adjustment for ties	0.00
Adjustment for zeros	0.00
Adjusted variance	6396.25

H0: phtime = graphtime
z = -3.157
Prob > z = 0.0016
Exact prob = 0.0012

H_0: The median differences between graphs and photographical forms is zero regarding accurateness.

H_1: The median differences between graphs and photographical forms is not zero regarding accurateness.

Sign	Obs	Sum ranks	Expected
Positive	24	704.500	406
Negative	5	107.500	406
Zero	13	91	91
All	42	903	903

Unadjusted variance	6396.25
Adjustment for ties	-88.88
Adjustment for zeros	-204.75
Adjusted variance	6102.63

H0: phanswer = graphanswer
z = 3.821
Prob > z = 0.0001
Exact prob = 0.0001

When comparing photographical to graphical forms, both speed and accurateness are significant (both p = 0.00). The H_0 can be rejected, so the median difference between photographs and graphs is not zero. Questions with photographical forms are answered quicker (z = -3.157) than questions with graphical forms. On top of this, questions with photographical forms are answered more accurately (z = 3.821) than graphical forms.

5 Conclusion

This thesis investigated whether different data representations can lead to quicker and more accurate fraud identification processes. The theoretical background was formed by showing the significance of research in fraud identification and how fraud identification has become more reliant on IT, as the business environment has changed significantly over the past couple of years. ERP systems have unified the internal processes within organizations and created AIS by which financial information is shared with third parties. The importance of IT is amplified by the revised standards of ISA 315. To answer the research question, an experiment was created by using a thought experiment, which led to five different fraud predications. Each fraud predication had three different forms of data representations: a table, a graph and a photograph. These data representations were created with Tableau using a dataset from a fictive metal company. Due to the lack of many respondents with accounting experience, the choice was made to focus the sample on accounting students. To make sure as many respondents were used as possible, the quasi-experiment created pairs of observations on the speed and accuracy of the three different representations. By using a Wilcoxon paired signed rank test for the variable speed and the paired t-test for the variable accuracy, the differences between two different representations are investigated. The tests had significant and insignificant results. A clear answer to the research question can be formed.

Especially regarding time, the evidence clearly states that it took less time for the respondents to answer the question when they had both a graph and a photographical form than it took to answer the question with a tabular form. The experiment had always different correct answers and different representations, so the order of the questions would not matter on the speed of the respondent's action. When looking into the accurateness of each question, there was no significant difference between tables and graphs. Respondents do not answer the question better when they are provided with a graph in comparison to a table. However, they took more time to answer the question when using a table. Respondents answer the question more accurately when provided with photographical representations. This could be because photographical forms make it easier to detect an outlier and see causal relations. Photographical representations could have a "self-explanatory" effect. It could be easier for respondents to detect some red flag if the red flag is clearly visioned. In tables and graphs, people need to understand the table and graph itself first, while a picture explains itself. As Few & Edge (2007) described, a graphical approach better assists our brains, as it is something that matches real life relationships. Concluding this thesis, it is more efficient and effective to use graphs and photographical forms to detect possible fraud than it is to use tabular forms.

6 Discussion

This thesis takes a different perspective on data visualization. Not only by looking into the theoretical side of data visualization, but also by looking to the practical perspective by using an experiment to find proof for the relevance of data visualization in the realm of detecting possible fraud.

6.1 Internal validity

By conducting an experiment, the internal validity remains a risk at hand. This experiment used accurateness and speed as measurables for effectiveness and efficiency. Effectiveness is difficult to measure. Accurateness is subjective, especially in the ranks of fraud detection. What is an anomaly and what is not? The discussion about what anomalies are and when it could be a predication of fraud is one that is not answered by literature. Research into the perception of respondents regarding anomalies could help clarifying what an anomaly is and what not but can also enlighten the different perceptions of respondents regarding different anomalies. However, this research did not use difficult tasks to measure the effectiveness. The outliers that were correct were very different as opposed to the incorrect answers and by doing so, it tried to stay away from this discussion.

6.2 External validity

As promising as data visualization might look in theory (as explained in section 2), it has several issues in its practical relevance. It is time consuming to learn the basics of data visualization tools like Tableau and even when these are known, becoming experienced is necessary to achieve real time gain. In this experiment, the time it takes to learn and master data visualization techniques is not considered. This might shed a different light on the results. In practise, it is easier to just create a pivot table with Excel than it is to create a graph in Tableau and ordering data accordingly. Time gain can however be achieved if education would focus more on IT developments and on different subjects of data mining with IT applications, especially in accounting related studies. The more experienced auditors are with data visualization, the more they can use all its advantages for day-to-day practises. On the contrary, the larger datasets become, the more difficult it gets to analyse raw data. This data needs to be transformed anyway, which is when data visualization should be an option in the skillset of any investigator.

On top of this, many different forms of data representations exist. Future research could look into different visualizations and how perception of the respondent matters. Some people could be working better with photographic representations, while others would prefer graphs. This personal bias could affect results as well. On top of this, future research could test whether high complexity tasks benefit more from data visualization than lower complexity tasks. Respondents for high complexity tasks should come from experienced external auditors, which is are people difficult to address. Doing so would raise the bar for future research in the field of possible fraud detection.

7 Literature

ACFE. (2022). *Occupational Fraud 2022: A Report to the Nations*.

Al-Hashedi, K. G., & Magalingam, P. (2021). Financial fraud detection applying data mining techniques: A comprehensive review from 2009 to 2019. *Computer Science Review*, *40*, 100402.

Albrecht, W. S., Albrecht, C. O., Albrecht, C. C., & Zimbelman, M. F. (2012). Fraud Examination. South-Western Cengage Learning. *Mason, OH*.

Ali, S. M., Gupta, N., Nayak, G. K., & Lenka, R. K. (2016). Big data visualization: Tools and challenges. 2016 2nd International Conference on Contemporary Computing and Informatics (IC3I),

Apparao, G., Singh, A., Rao, G., Bhavani, B. L., Eswar, K., & Rajani, D. (2009). Financial statement fraud detection by data mining. *Corporate governance*, *3*(1), 159-163.

Arnold, V. (2006). Behavioral research opportunities: Understanding the impact of enterprise systems. *International Journal of Accounting Information Systems*, *7*(1), 7-17.

Bachmid, F. S. (2016). The effect of accounting information system quality on accounting information quality. *Research Journal of Finance and Accounting*, *6*.

Batt, S., Grealis, T., Harmon, O., & Tomolonis, P. (2020). Learning Tableau: A data visualization tool. *The Journal of Economic Education*, *51*(3-4), 317-328.

Bhattacharyya, S., Jha, S., Tharakunnel, K., & Westland, J. C. (2011). Data mining for credit card fraud: A comparative study. *Decision support systems*, *50*(3), 602-613.

Bodnar, G. H., & Hopwood, W. S. (2004). *Accounting information systems* (9th ed.). Pearson Prentice Hall.

Bolton, R. J., & Hand, D. J. (2002). Statistical fraud detection: A review. *Statistical science*, *17*(3), 235-255.

Bose, I., & Mahapatra, R. K. (2001). Business data mining—a machine learning perspective. *Information & management*, *39*(3), 211-225.

Caldarola, E. G., & Rinaldi, A. M. (2017). Big data visualization tools: a survey. *Research Gate*.

Chen, H. J., Huang, S. Y., Chiu, A. A., & Pai, F. C. (2012). The ERP system impact on the role of accountants. *Industrial Management & Data Systems*.

Cockcroft, S., & Russell, M. (2018). Big data opportunities for accounting and finance practice and research. *Australian Accounting Review*, *28*(3), 323-333.

Coderre, D. (2009). *Computer Aided Fraud Prevention and Detection: A Step by Step Guide*. John Wiley & Sons.

Dechow, N., & Mouritsen, J. (2005). Enterprise resource planning systems, management control and the quest for integration. *Accounting, organizations and society*, *30*(7-8), 691-733.

Deloitte. (2012). *Visual Analytics: Revealing Corruption, Fraud, Waste, and Abuse*. https://www.slideshare.net/DeloitteForensicCenter/visual-analytics-revealing-corruption-fraud-waste-and-abuse-13958016

Dilla, W., Janvrin, D. J., & Raschke, R. (2010). Interactive data visualization: New directions for accounting information systems research. *Journal of Information Systems*, *24*(2), 1-37.

Dilla, W. N., & Raschke, R. L. (2015). Data visualization for fraud detection: Practice implications and a call for future research. *International Journal of Accounting Information Systems*, *16*, 1-22.

Drnevich, P. L., & Croson, D. C. (2013). Information technology and business-level strategy: Toward an integrated theoretical perspective. *MIS quarterly*, 483-509.

Economics, C. (2019). *ERP Adoption Trends and Customer Experience*.

Few, S., & Edge, P. (2007). Data visualization: past, present, and future. *IBM Cognos Innovation Center*.

George, G., Haas, M. R., & Pentland, A. (2014). Big data and management. In (Vol. 57, pp. 321-326): Academy of Management Briarcliff Manor, NY.

Ghasemi, M., Shafeiepour, V., Aslani, M., & Barvayeh, E. (2011). The impact of Information Technology (IT) on modern accounting systems. *Procedia-Social and Behavioral Sciences*, *28*, 112-116.

Grabski, S., Leech, S., & Sangster, A. (2009). *Management accounting in enterprise resource planning systems*. Butterworth-Heinemann.

Grabski, S. V., Leech, S. A., & Schmidt, P. J. (2011). A review of ERP research: A future agenda for accounting information systems. *Journal of Information Systems*, *25*(1), 37-78.

Gray, G. L., & Debreceny, R. S. (2014). A taxonomy to guide research on the application of data mining to fraud detection in financial statement audits. *International Journal of Accounting Information Systems*, *15*(4), 357-380.

Gurbaxani, V., & Whang, S. (1991). The impact of information systems on organizations and markets. *Communications of the ACM*, *34*(1), 59-73.

Han, J., & Kamber, M. (2006). Data Mining: Concepts and Techniques Second Edition. Morgan Kauffman. *San Fransisco*.

Hawkins, D. M. (1980). *Identification of outliers* (Vol. 11). Springer.

Hilal, W., Gadsden, S. A., & Yawney, J. (2022). Financial Fraud:: A Review of Anomaly Detection Techniques and Recent Advances.

IAASB. (2021). *2020 Handbook of International Quality Control, Auditing, Review, Other Assurance, and Related Services Pronouncements*. The International Federation of Accountants (IFAC). https://www.iaasb.org/publications/2020-handbook-international-quality-control-auditing-review-other-assurance-and-related-services

Iskandar, D. (2015). Analysis of factors affecting the success of the application of accounting information system. *International Journal of scientific & Technology research*, *4*(2), 155-162.

Janvrin, D., Bierstaker, J., & Lowe, D. J. (2008). An examination of audit information technology use and perceived importance. *Accounting horizons*, *22*(1), 1-21.

Kieso, D. E., Weygandt, J. J., Warfield, T. D., Wiecek, I. M., & McConomy, B. J. (2019). *Intermediate Accounting, Volume 2*. John Wiley & Sons.

Kilpatrick, T. (2000). Auditing manufacturing costs. *Internal Auditor*, *57*(3), 25-25.

Kim, T. K. (2015). T test as a parametric statistic. *Korean journal of anesthesiology*, *68*(6), 540-546.

Koch, C., Slater, D., & Baatz, E. (1999). the ABCs of ERP. *CIO magazine*, *22*.

Lee, L., Kerler, W., & Ivancevich, D. (2018). Beyond Excel: Software tools and the accounting curriculum. *AIS Educator Journal*, *13*(1), 44-61.

Leite, R. A., Gschwandtner, T., Miksch, S., Gstrein, E., & Kuntner, J. (2018). Visual analytics for event detection: Focusing on fraud. *Visual informatics*, *2*(4), 198-212.

Leite, R. A., Gschwandtner, T., Miksch, S., Kriglstein, S., Pohl, M., Gstrein, E., & Kuntner, J. (2017). Eva: Visual analytics to identify fraudulent events. *IEEE transactions on visualization and computer graphics*, *24*(1), 330-339.

Libby, R., Bloomfield, R., & Nelson, M. W. (2002). Experimental research in financial accounting. *Accounting, organizations and society*, *27*(8), 775-810.

McCrum-Gardner, E. (2008). Which is the correct statistical test to use? *British Journal of Oral and Maxillofacial Surgery*, *46*(1), 38-41.

Ngai, E. W., Hu, Y., Wong, Y. H., Chen, Y., & Sun, X. (2011). The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature. *Decision support systems*, *50*(3), 559-569.

Nuamah, J. K., Seong, Y., Jiang, S., Park, E., & Mountjoy, D. (2020). Evaluating effectiveness of information visualizations using cognitive fit theory: A neuroergonomics approach. *Applied Ergonomics*, *88*, 103173.

Pratt, J. W., & Gibbons, J. D. (2012). *Concepts of nonparametric theory*. Springer Science & Business Media.

Ramos Montesdeoca, M., Sánchez Medina, A. J., & Blázquez Santana, F. (2019). Research topics in accounting fraud in the 21st century: A state of the art. *Sustainability*, *11*(6), 1570.

Reurink, A. (2018). Financial fraud: a literature review. *Journal of Economic Surveys*, *32*(5), 1292-1325.

Robey, D., Ross, J. W., & Boudreau, M.-C. (2002). Learning to implement enterprise systems: An exploratory study of the dialectics of change. *Journal of management information systems*, *19*(1), 17-46.

Romney, M. B., & Steinbart, P. J. (2006). *Accounting information systems* (10th ed.). Pearson Prentice Hall. http://catdir.loc.gov/catdir/toc/ecip057/2005003053.html

Salijeni, G., Samsonova-Taddei, A., & Turley, S. (2019). Big Data and changes in audit technology: contemplating a research agenda. *Accounting and business research*, *49*(1), 95-119.

Sharda, R., Delen, D., Turban, E., Aronson, J., & Liang, T. (2014). Business intelligence and analytics. *System for Decesion Support*.

Sharma, A., & Panigrahi, P. K. (2013). A review of financial accounting fraud detection based on data mining techniques. *arXiv preprint arXiv:1309.3944*.

Spathis, C., & Constantinides, S. (2004). Enterprise resource planning systems' impact on accounting processes. *Business Process management journal*.

Stevens, J. P. (2012). *Applied multivariate statistics for the social sciences*. Routledge.

Turban, E., Sharda, R., & Delen, D. (2010). Decision support and business intelligence systems (required). *Google Scholar*.

Vasarhelyi, M. A., Kogan, A., & Tuttle, B. M. (2015). Big data in accounting: An overview. *Accounting horizons*, *29*(2), 381-396.

Vessey, I. (1991). Cognitive fit: A theory-based analysis of the graphs versus tables literature. *Decision sciences*, *22*(2), 219-240.

Vessey, I., Zhang, P., & Galletta, D. (2006). The theory of cognitive fit. *Human-computer interaction and management information systems: Foundations*, *5*, 141-183.

Wells, J. T. (2014). *Principles of fraud examination*. John Wiley & Sons.

West, J., & Bhattacharya, M. (2016). Intelligent financial fraud detection: a comprehensive review. *Computers & security*, *57*, 47-66.

Wilkinson, J. W. (1999). *Accounting Information Systems: Essential Concepts and Applications*. John Wiley & Sons Australia, Limited. https://books.google.nl/books?id=Fr2LNAEACAAJ

Woolson, R. F. (2007). Wilcoxon signed-rank test. *Wiley encyclopedia of clinical trials*, 1-3.

Yi, J. S., ah Kang, Y., Stasko, J., & Jacko, J. A. (2007). Toward a deeper understanding of the role of interaction in information visualization. *IEEE transactions on visualization and computer graphics*, *13*(6), 1224-1231.

Zhang, D., & Zhou, L. (2004). Discovering golden nuggets: data mining in financial application. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, *34*(4), 513-522.

8 Appendix

8.1 Survey exported from Qualtrics

Data visualization

Start of Block: Introduction

Dear participant

Thank you for participating in this experiment. This experiment is part of a master's thesis for Corporate Finance and Control. It should only take about 5 minutes to complete the experiment. This experiment tests whether different data representations lead to quicker and more accurate fraud identification processes.

Firstly, some questions about personal characteristics will be asked. The data arising from these questions will be used for this thesis and this thesis only, and will be deleted when the experiment is completed. Information is processed anonymously.

You will be presented with 15 different data representations of 5 different subjects of potential fraud. I ask you to take the seat of an external auditor and review the data critically to spot potential fraudulent transactions. Time will be tracked and only one answer is correct per question.

Thanks in advance for taking the experiment!

End of Block: Introduction

Start of Block: Personal characteristics

Education

What educational programme are you currently following?

- ○ Bachelor Business Administration/Economics (1)
- ○ Bachelor Economics (2)
- ○ Master's degree in Economics (3)
- ○ Master's degree in Business Administration (4)
- ○ None of the above, however, I have affinity with Accounting. (5)
- ○ None of the above and I have no affinity with Accounting at all. (6)

Age

How old are you?

____

Gender

What is your gender?

- Male (1)
- Female (2)
- Non-binary / third gender (3)
- Prefer not to say (4)

End of Block: Personal characteristics

Start of Block: Questions

Introduction

The data that is presented is from a fictive Metal Company that imports metal scrap from different countries in Europe to the Netherlands. In the Netherlands, it creates different constructions that are used for building bridges. Its purchasing department works from Mondays to Fridays, within normal business hours. Most payments are done with wire transfers.

However, you are looking for suspicious activity within the company. You use different data representations to look for anomalies. An anomaly is something that is out of the ordinary or not in the usual course of business. For example: you can encounter transactions to private bank accounts. This is not in the usual course of business, as the metal scrap company imports from different businesses throughout Europe. This could indicate potential fraud and should raise a red flag.

This experiment consists of five different subjects of potential fraud, knowingly:

1. Large transactions;

2. Transactions on certain days of the week;
3. Transactions on different hours in the day;
4. Transactions to different countries;
5. Transactions with different purposes.

Try to answer the questions as quickly and accurate as possible. Only one answer can be given per question. Good luck!

Page Break

Representation 1 Day of the week of the transaction

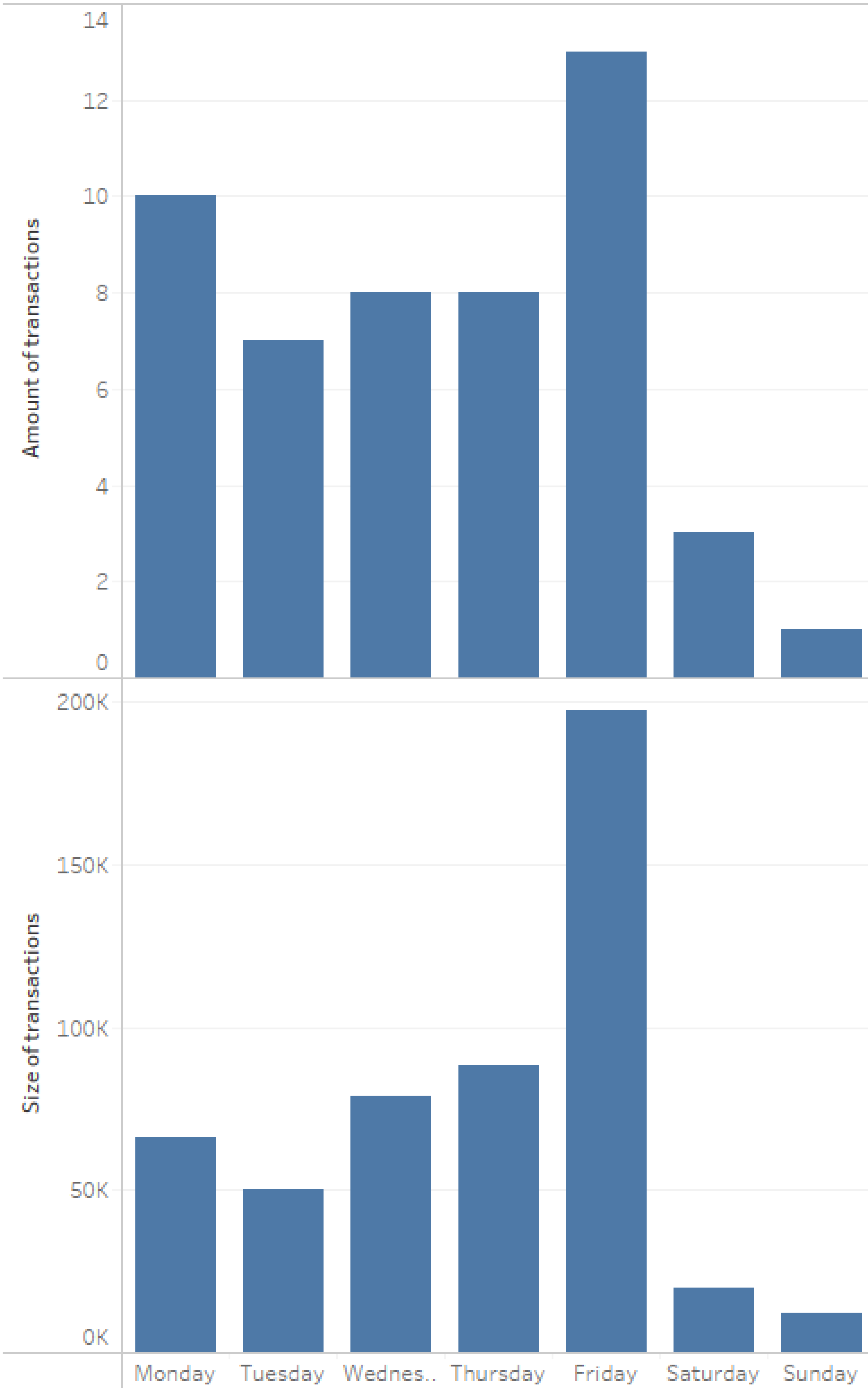
Transaction day	Amount of transactions	Size of transactions
Monday	10	66.043
Tuesday	7	50.059
Wednesday	8	78.755
Thursday	8	88.080
Friday	13	197.300
Saturday	3	20.000
Sunday	1	12.000

Q1 Which day would you investigate in more depth?

- ○ Monday (1)
- ○ Tuesday (2)
- ○ Wednesday (3)
- ○ Thursday (4)
- ○ Friday (5)
- ○ Saturday (6)
- ○ Sunday (7)

Page Break

Representation 2 Day of the week of the transaction

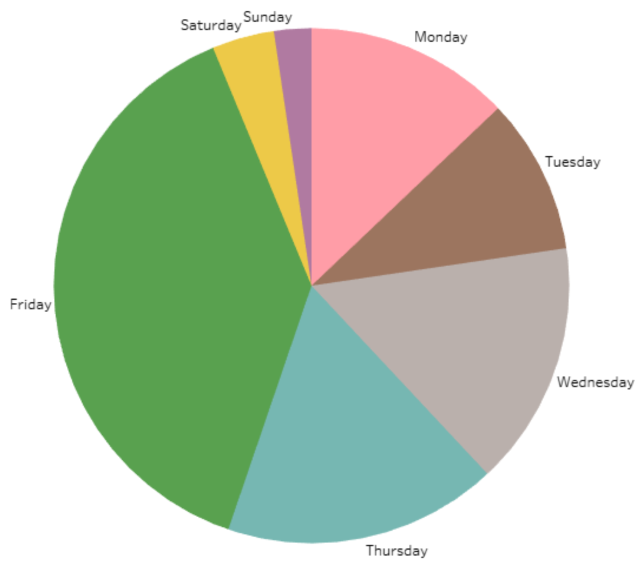


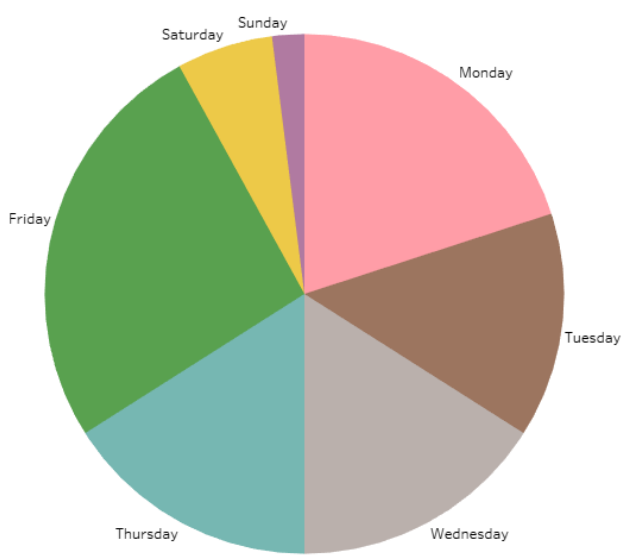
Q2 Which day would you investigate in more depth?

- ○ Monday (1)
- ○ Tuesday (2)
- ○ Wednesday (3)
- ○ Thursday (4)
- ○ Friday (5)
- ○ Saturday (6)
- ○ Sunday (7)

Page Break

Representation 3 Day of the week of the transaction





Q32

Q3 Which day would you investigate further?

○ Monday (1)

○ Tuesday (2)

○ Wednesday (3)

○ Thursday (4)

○ Friday (5)

○ Saturday (6)

○ Sunday (7)

Page Break

Representation 4 Size of the transaction

Number	Size of transactions
1	10.240
2	5.830
3	2.080
4	2.080
5	1.560
6	3.500
7	5.100
8	30.000
9	1.333
10	4.320
11	15.840
12	2.400
13	6.900
14	9.800
15	6.020
16	1.200
17	7.899
18	4.300
19	4.444
20	25.000
21	7.432
22	9.800
23	3.299
24	10.500
25	13.980
26	1.530
27	40.000
28	4.300
29	3.200
30	6.750
31	2.350
32	8.950
33	21.000
34	34.000
35	8.700
36	5.600
37	5.500
38	3.000
39	50.000
40	40.000
41	35.000
42	3.400
43	5.400
44	2.800
45	1.900
46	2.000
47	6.000
48	9.500
49	4.500
50	12.000

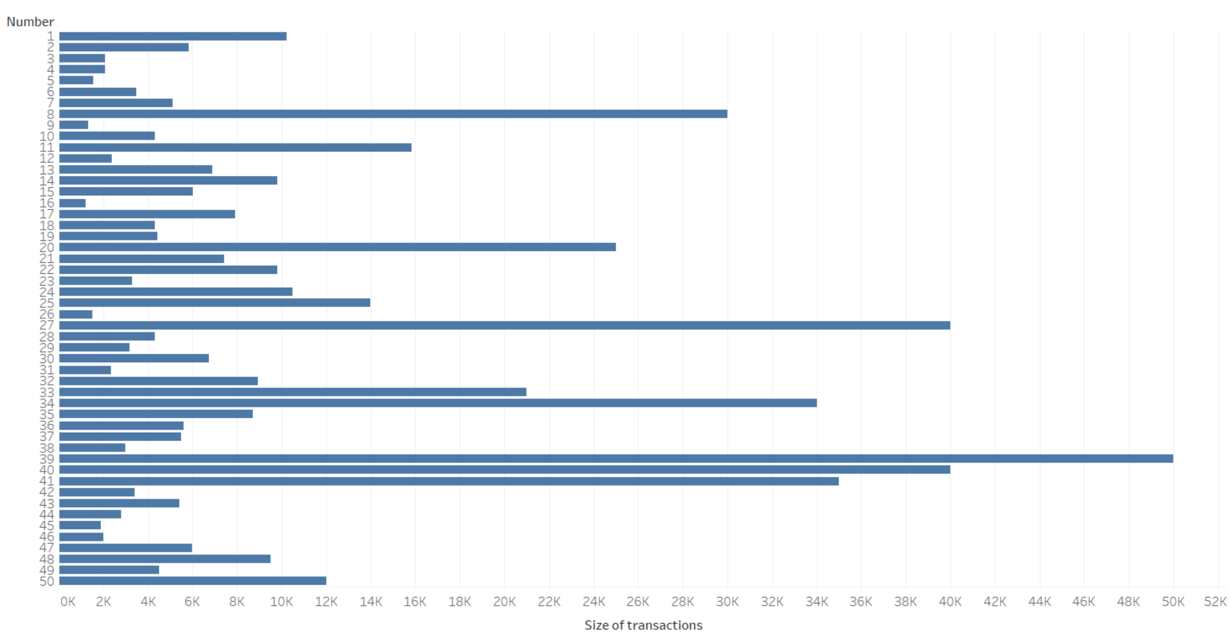
Q4 What transaction would you investigate into further detail?

- 26 (1)
- 41 (2)
- 10 (3)
- 45 (4)

Page Break

Representation 5 Size of the transaction

Pinch you finger to zoom in.

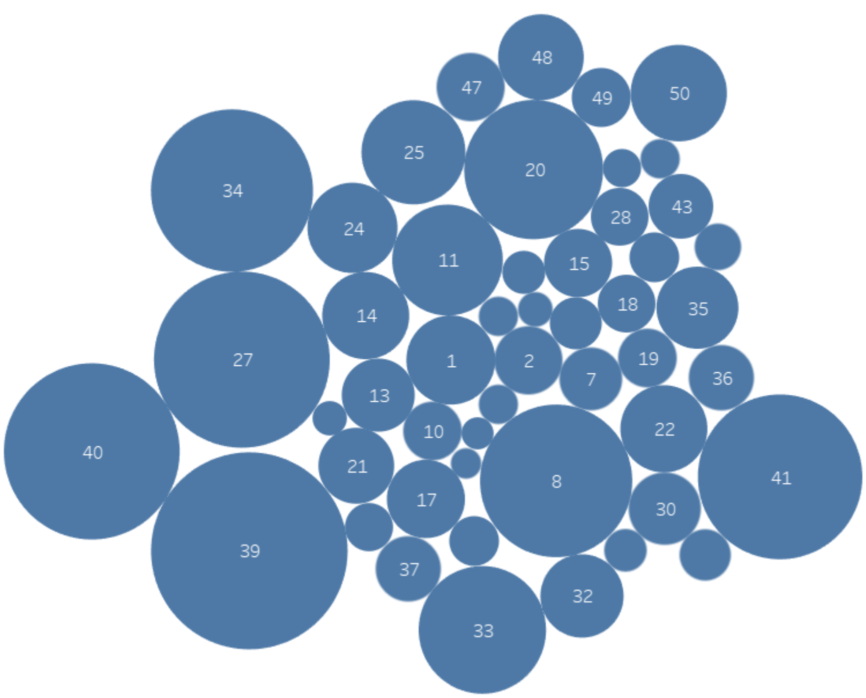


Q5 What transaction would you like to investigate into further detail?

- 16 (1)
- 49 (2)
- 38 (3)
- 27 (4)

Page Break

Representation 6 Size of the transaction



Q6 What transaction would you investigate in more depth?

- 39 (1)
- 20 (2)
- 2 (3)
- 13 (4)

Page Break

Representation 7 Time when the transaction is done

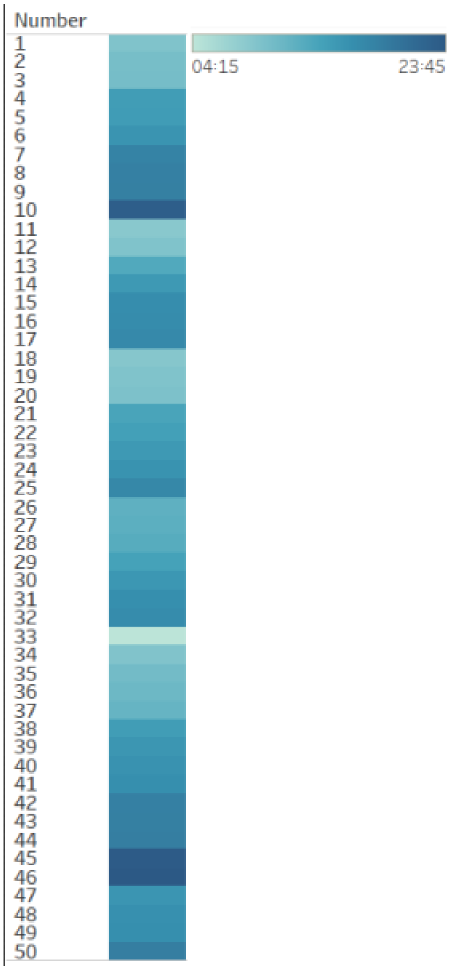
Number	Minute ..	Size
1	09:15:00	10.240
2	10:00:00	5.830
3	10:05:00	2.080
4	14:44:00	2.080
5	14:50:00	1.560
6	15:55:00	3.500
7	18:09:00	5.100
8	18:30:00	30.000
9	18:35:00	1.333
10	23:03:00	4.320
11	08:30:00	15.840
12	09:15:00	2.400
13	13:08:00	6.900
14	15:15:00	9.800
15	16:45:00	6.020
16	16:55:00	1.200
17	17:15:00	7.899
18	08:45:00	4.300
19	09:13:00	4.444
20	09:30:00	25.000
21	13:45:00	7.432
22	14:23:00	9.800
23	15:12:00	3.299
24	16:00:00	10.500
25	17:30:00	13.980
26	12:03:00	1.530
27	12:15:00	40.000
28	12:42:00	4.300
29	14:01:00	3.200
30	15:30:00	6.750
31	16:33:00	2.350
32	16:55:00	8.950
33	04:15:00	21.000
34	09:10:00	34.000
35	10:23:00	8.700
36	10:55:00	5.600
37	11:32:00	5.500
38	14:45:00	3.000
39	15:35:00	50.000
40	16:05:00	40.000
41	16:30:00	35.000
42	18:30:00	3.400
43	18:35:00	5.400
44	18:45:00	2.800
45	23:30:00	1.900
46	23:45:00	2.000
47	15:44:00	6.000
48	16:23:00	9.500
49	16:33:00	4.500
50	18:45:00	12.000

Q7 What transaction would you investigate into more detail?

- 26 (1)
- 9 (2)
- 46 (3)
- 22 (4)

Page Break

Representation 8 Time when the transaction is done



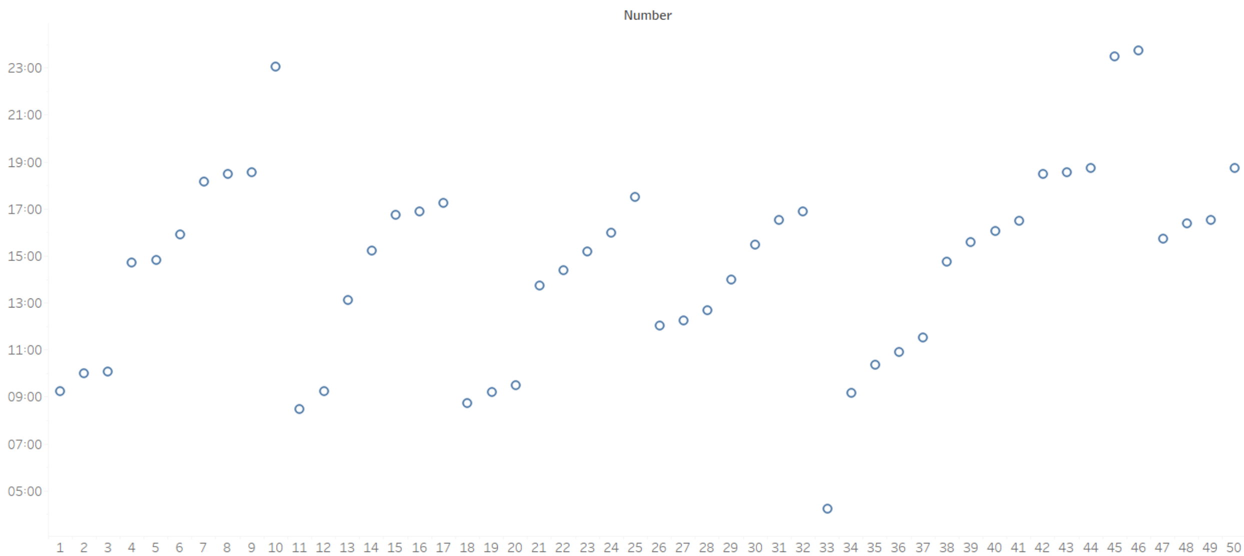
Q8 What transaction would you investigate in more depth?

- 33 (1)
- 29 (2)
- 25 (3)
- 21 (4)

Page Break

Representation 9 Time when the transaction is done

Pinch you finger to zoom in.



Q9 What transaction would you investigate into more detail?

◯ 10 (1)

◯ 28 (2)

◯ 3 (3)

◯ 20 (4)

Page Break

Representation 10 Country of destination of the transaction

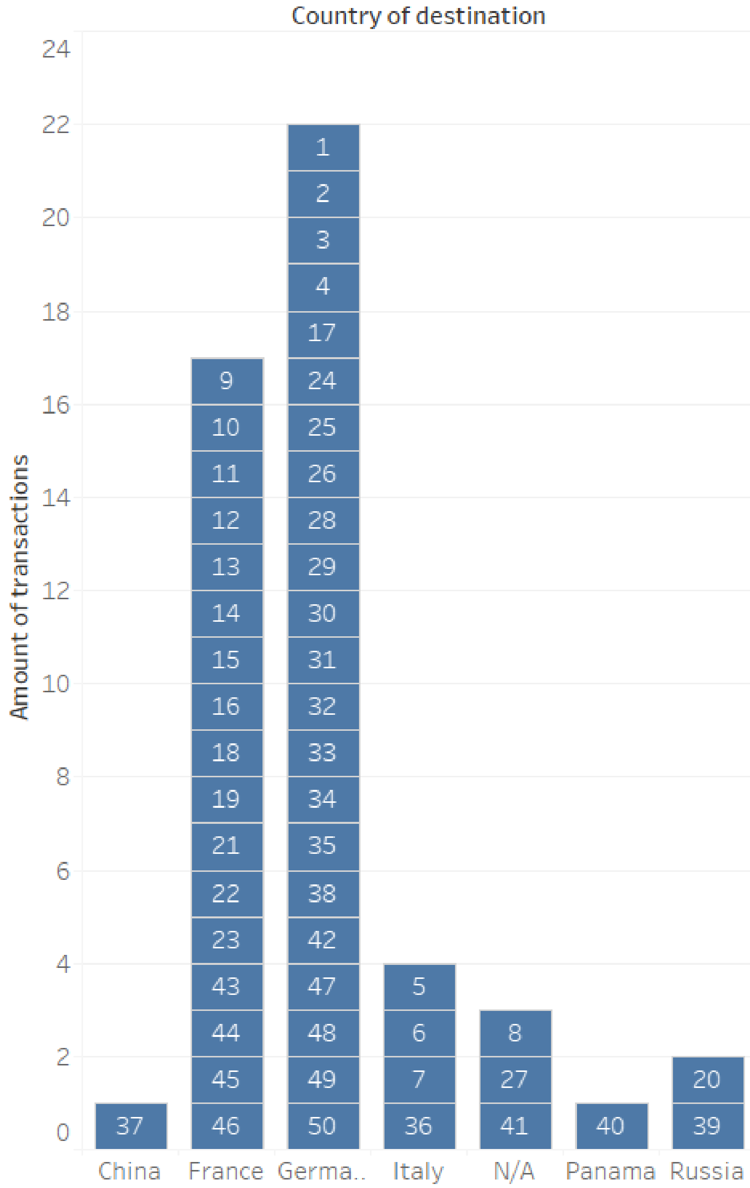
Number	Country..	Size
1	Germany	10.240
2	Germany	5.830
3	Germany	2.080
4	Germany	2.080
5	Italy	1.560
6	Italy	3.500
7	Italy	5.100
8	N/A	30.000
9	France	1.333
10	France	4.320
11	France	15.840
12	France	2.400
13	France	6.900
14	France	9.800
15	France	6.020
16	France	1.200
17	Germany	7.899
18	France	4.300
19	France	4.444
20	Russia	25.000
21	France	7.432
22	France	9.800
23	France	3.299
24	Germany	10.500
25	Germany	13.980
26	Germany	1.530
27	N/A	40.000
28	Germany	4.300
29	Germany	3.200
30	Germany	6.750
31	Germany	2.350
32	Germany	8.950
33	Germany	21.000
34	Germany	34.000
35	Germany	8.700
36	Italy	5.600
37	China	5.500
38	Germany	3.000
39	Russia	50.000
40	Panama	40.000
41	N/A	35.000
42	Germany	3.400
43	France	5.400
44	France	2.800
45	France	1.900
46	France	2.000
47	Germany	6.000
48	Germany	9.500
49	Germany	4.500
50	Germany	12.000

Q10 What transaction would you investigate into more depth?

- 38 (1)
- 20 (2)
- 31 (3)
- 44 (4)

Page Break

Representation 11 Country of destination of the transaction



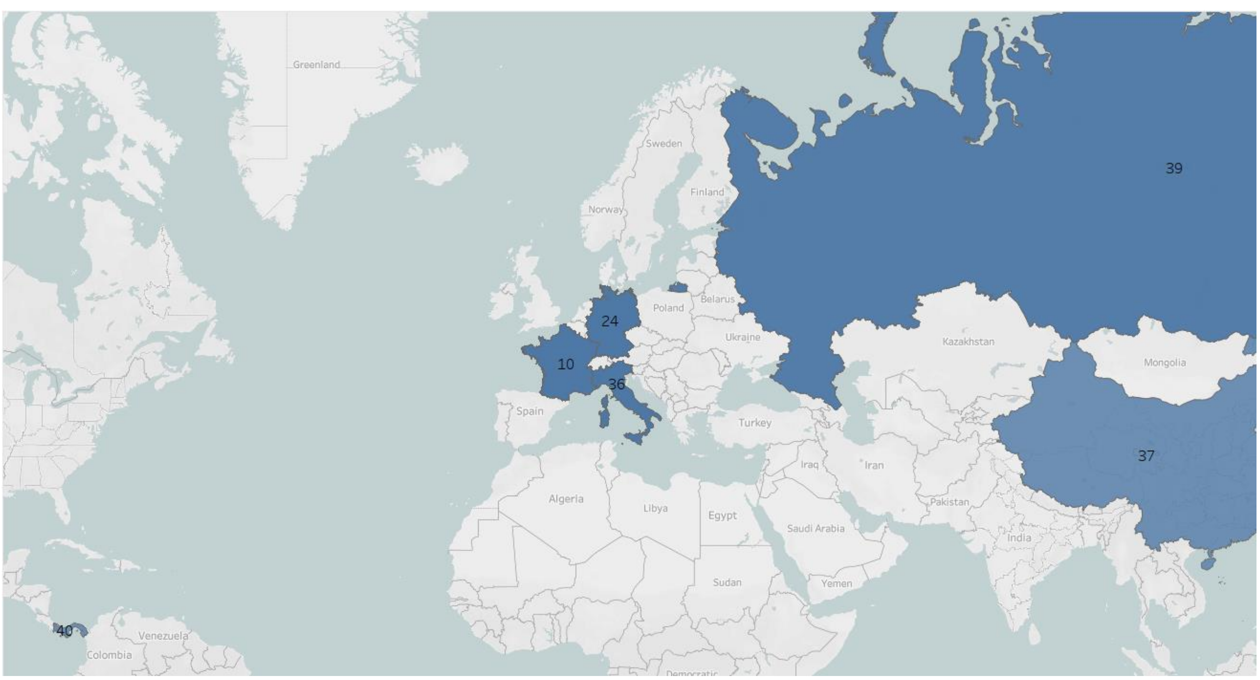
Q11 What transaction would you investigate into more depth?

- 19 (1)
- 5 (2)
- 48 (3)
- 37 (4)

Page Break

Representation 12 Country of destination of the transaction

Pinch you finger to zoom in.



Q12 What transaction would you investigate into further detail?

- 10 (1)
- 40 (2)
- 36 (3)
- 24 (4)

Page Break

Representation 13 Customer / Purpose of the transaction

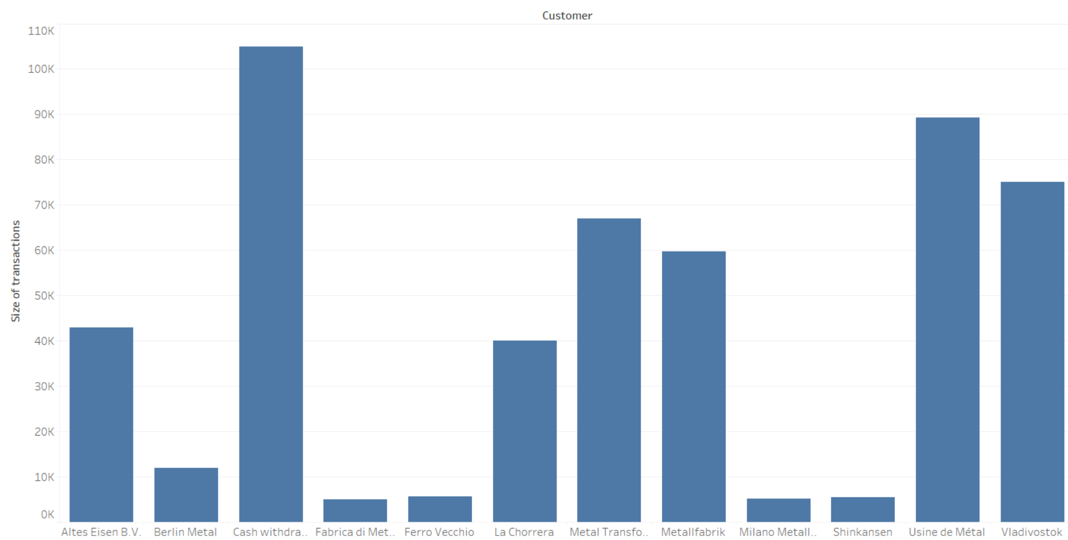
Customer	Size
Altes Eisen B.V.	43.000
Berlin Metal	12.000
Cash withdrawal	105.000
Fabrica di Metallo	5.060
Ferro Vecchio	5.600
La Chorrera	40.000
Metal Transformer ..	67.059
Metallfabrik	59.730
Milano Metallo B.V.	5.100
Shinkansen	5.500
Usine de Métal	89.188
Vladivostok	75.000

Q13 What transactions would you investigate into more detail?

- Metallfabrik (1)
- La Chorrera (2)
- Usine de Métal (3)

Page Break

Representation 14 Customer / Purpose of the transaction

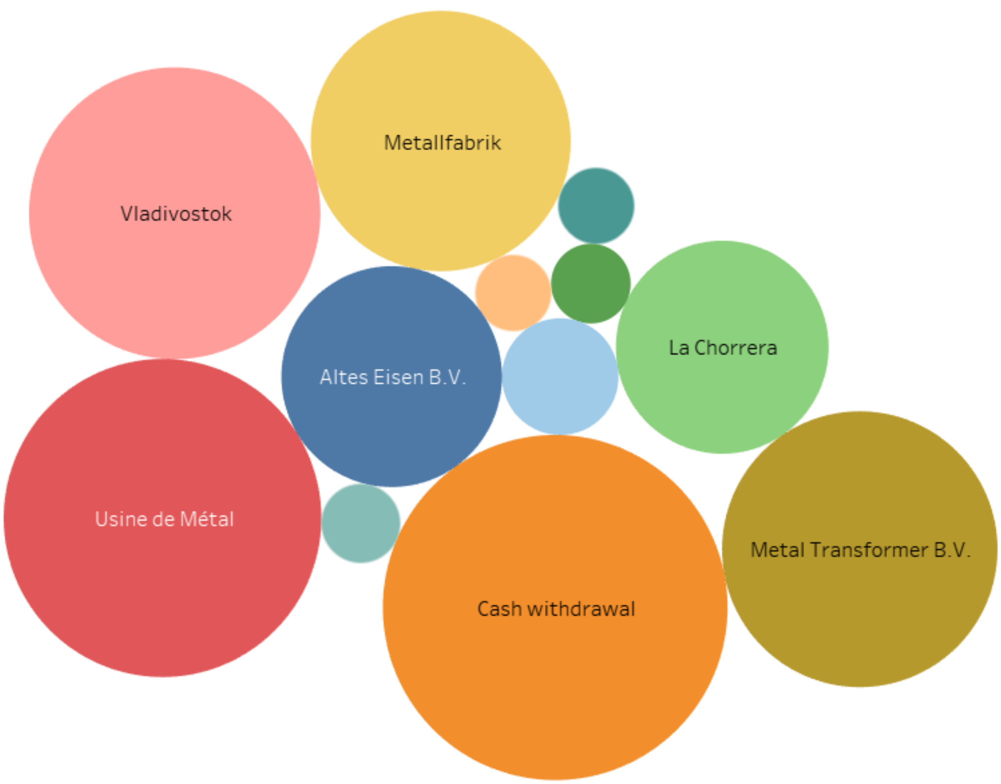


Q14 What transactions would you investigate into more detail?

- ○ Metal Transformer (1)
- ○ Altes Eisen B.V. (2)
- ○ Shinkansen (3)

Page Break

Representation 15 Customer / Purpose of the transaction



Q15 What transactions would you investigate into more detail?

- ○ Usine de Métal (1)
- ○ Cash withdrawal (2)
- ○ Metallfabrik (3)

Page Break

End of Block: Questions