EconoVis

Visualizing Economic Trends in a Modern World

Manuel João Pereira Alves

Thesis to obtain the Master of Science Degree in

Information Systems and Computer Engineering

Supervisors: Prof. Sandra Pereira Gama
Prof. Clara Patrícia Costa Raposo

Examination Committee

Chairperson: Prof. Miguel Nuno Dias Alves Pupo Correia
Supervisor: Prof. Sandra Pereira Gama
Member of the Committee: Prof. Daniel Jorge Viegas Gonçalves

November 2018
Acknowledgments

First and foremost, I would like to thank my supervisors Prof. Sandra Gama and Prof. Clara Raposo, for all the support that they have given me, all the advice and guidance through this long journey, for never giving up on me and for always being available. Without them, this work would not be possible, and I will always be grateful to them.

I would also like to express my gratitude to Duncan for all his help with the revision of the English since day one. This document would not be so well written without his support.

To my family that have always supported me when I needed them the most and provided me with everything I needed for my studies which not only allowed me to concentrate on this work but deflected any possible distractions.

A very special thank you to my girlfriend Jessica that gave me all the love and support when I needed the most and had the patience through the days that I was less present. For all the help and encouragement, a special thanks to her.

To all my colleagues and friends that I have made since the beginning of my studies, in particular to Ema for helping me during my research, especially for helping me find one of my advisors and for helping me with the usability tests, I’m so grateful to her. To Adriana, Pedro, and Zé, for helping me and giving me motivation since the first day and to support and help me always when I needed, a big thank you, guys. To Antunes for all the technical help that he gave me and all the advice that he shared with me, I’m grateful for that.

And also to Prof. Victor Barros for helping me with the financial background and with the evaluation part. Thank you also to all the users that helped me test my visualization.

Each of these people made this thesis possible. Thank you all.
Abstract

In financial markets, all users have access to the same data, so to gain benefits over competitors it is essential to present information in a way that allows a rapid analysis. The use of information visualization techniques permits the exploitation of data in an interactive way, the selection of data to analyze, and the choice of different types of representations of the same data. An important source of financial data are the financial statements, and although the advantages of information visualization are already known and being used in finance, regarding financial statements its exploitation is still very limited given its potential. We propose applying information visualization to financial statements and, to accomplish that, we developed an information visualization system called EconoVis that provides an overview of the information contained in financial statements from the companies in the Portuguese Stock Market through a set of different visualizations. We followed an iterative and incremental approach with an emphasis on prototyping using financial experts to help through the process. To validate our research, we performed Usability Tests and Case Studies with financial users that attested that the users favorably accepted EconoVis.

Keywords

Information Visualization, Financial Statements, Financial Data Visualization, Usability.
Resumo

Nos mercados financeiros, todas as pessoas têm acesso aos mesmos tipos de dados, de modo que, para obter benefícios sobre os concorrentes, é essencial apresentar a informação de uma forma que permita uma rápida análise. O uso de técnicas de visualização de informação permite explorar dados de uma forma interactiva, selecionar os dados a analisar e escolher diferentes tipos de representações para os mesmos. Uma fonte importante de dados financeiros são as demonstrações financeiras, e, apesar das vantagens da visualização da informação serem já conhecidas e usadas na área de finanças, em relação às demonstrações financeiras, a sua exploração ainda é muito limitada, tendo em conta o seu potencial. Propusemo-nos a aplicar técnicas de visualização de informação às demonstrações financeiras e, para o conseguir, desenvolvemos um sistema denominado EconoVis, que permite a visualização da informação presente nas demonstrações financeiras das empresas da bolsa portuguesa através de um conjunto de diferentes visualizações. Seguimos uma abordagem iterativa e incremental com ênfase na prototipagem, usando especialistas financeiros para ajudar em todo o processo. Para validar a nossa pesquisa, realizámos Testes de Usabilidade e Casos de Estudo com utilizadores da área que atestaram a aceitação do EconoVis por parte dos utilizadores.

Palavras Chave

Visualização de Informação, Demonstrações Financeiras, Visualização de Dados Financeiros, Usabilidade.
# Contents

1 Introduction ................................................. 1  
   1.1 Goals .................................................. 4  
   1.2 Document Structure ..................................... 5  

2 Related Work .................................................. 7  
   2.1 Stocks Data ............................................. 9  
   2.2 Funds Data ............................................. 14  
   2.3 Transactions Data ....................................... 16  
   2.4 Companies Data ......................................... 19  
   2.5 Discussion ............................................. 21  

3 Visualizing Economic Trends in a Modern World ............. 25  
   3.1 Initial Steps ........................................... 27  
   3.2 Requirements ........................................... 28  
      3.2.1 Tasks and Questions ................................ 28  
      3.2.2 Survey ................................................ 29  
      3.2.3 Interview ............................................ 31  
   3.3 Low Fidelity Prototyping (LFP) .......................... 31  
      3.3.1 The Sketches ......................................... 31  
      3.3.2 Validation ............................................ 33  
   3.4 Architecture ............................................. 35  
      3.4.1 Database Layer ....................................... 36  
      3.4.2 Business Layer ....................................... 37  
      3.4.3 Presentation Layer ................................... 37  
   3.5 Functional Prototype (FP) ............................... 38  
      3.5.1 Initial Version ....................................... 38  
      3.5.2 Final Version ......................................... 43
List of Figures

2.1 Stock Market Analyzer [1] ........................................ 10
2.2 Pixel Bar Charts with time-series data [2] .......................... 11
2.3 Visualization system for analyzing share performance [3] .......... 12
2.4 Bloomberg visualization project [4] ................................ 12
2.5 Stock exchange websites ............................................ 13
2.6 FundExplorer system [5] ............................................. 15
2.7 Portfolio monitoring interface [6] .................................... 15
2.8 Comparison between parallel coordinate plots [7] .................. 16
2.9 Domesticating Bead visualization tool [8] ........................... 17
2.10 WireVis system [9] .................................................. 18
2.11 Self-organizing map of pulp and manufacturing companies [10] .. 19

3.1 Survey Results ......................................................... 29
3.2 Low-Fidelity Prototype (LFP) Dashboard ............................ 32
3.3 Grouped Bar Chart and Scatter Chart options ....................... 32
3.4 Heatmap Sketch ..................................................... 33
3.5 Ring Chart and TreeMap Chart options ............................. 33
3.6 Parallel Coordinates Sketch ......................................... 34
3.7 EconoVis Architecture ............................................... 35
3.8 Add Dialog .......................................................... 40
3.9 Grouped Bar Chart ................................................... 41
3.10 TreeMap .............................................................. 41
3.11 EconoVis Initial Version ............................................. 42
3.12 Color Palettes ....................................................... 43
3.13 Group Bar Chart New Colors ...................................... 44
3.14 Parallel Coordinates Chart New Colors ............................ 44
3.15 TreeMap New Colors ........................................ 45
3.16 Grouped Bar Chart Overlap Problem ..................... 45
3.17 Financial Statements Filter ................................. 46
3.18 Tooltips ......................................................... 46
3.19 Dashboard without companies selected .................... 47
3.20 Years Filter ....................................................... 47
3.21 EconoVis Final Version ....................................... 48
4.1 Duration Results Box Plot .................................... 54
4.2 Number of Errors Results Box Plot ......................... 56
List of Tables

2.1 Related Work Overview ..................................................... 22
4.1 Duration results ............................................................ 54
4.2 Number of Errors results .................................................. 55
4.3 Satisfaction Level results .................................................. 57
4.4 System Usability Scale (SUS) results ................................. 58
Acronyms

API  Application Programming Interface
CSS  Cascading Style Sheets
CSV  Comma-Separated Values
HTML HyperText Markup Language
InfoVis Information Visualisation
ISEG Instituto Superior de Economia e Gestão
JSON JavaScript Object Notation
LFP  Low-Fidelity Prototype
PDF  Portable Document Format
REST Representational State Transfer
SUS  System Usability Scale
URL  Uniform Resource Locator
1

Introduction

Contents

1.1 Goals ................................................................. 4
1.2 Document Structure ........................................... 5
The use of the Internet has increased the exchange of information between investors and companies, allowing information to be available for a more significant number of users, in a more cost-effective way and in a timelier manner, unlike in the past where the usual channels used by firms were paper-based with associated limitations [12].

There are many different sources of financial data (such as news, analysts’ reports, blogs, and tweets) that provide daily information allowing investors and market analysts to make better and informed decisions [13]. Analyzing that kind of information is a hard task due to the large volume, its diversity, and complexity. Adding to the fact that, unlike what happens in other areas, in financial markets all users have access to the same data, so it is essential that the information should be available in an appropriate form, which allows a rapid analysis in order for investors to gain benefits over the competition [14], [13].

One important source of financial data are the financial statements that are intended to communicate an entity’s financial position at a point in time and its results of operations for a period which has ended. They usually include balance sheets, income statements, statements of retained earnings and statements of cash flows. The information is organized in a way that makes it easy to understand the financial performance of a company, so it is an essential tool for many parties. Investors who want to analyze new investments among the thousands of stocks available. Managers who want to compare the performance of their companies with their competitors. Creditors who want to analyze the company's payment ability. Auditors that want to assess the accuracy of a company’s financial statements and financial analysts that want to find financial trends on the markets comparing the performance of different companies [15].

In 2008 the Securities and Exchange Commission (SEC) ordered that all publicly held companies provide financial statement information in extensible business reporting language (XBRL)². That change should promote the development of interactive data viewers that make the information easier to analyze and compare. Although the noticeable increase of incorporate interactive data visualization features on the company’s websites, those features are still very limited such as a hyperlinked table of contents and search engines [16]. The more evolved interactive data visualization tools are mainly used internally by the companies to analyze all the data produced and stored by their information systems and are not available on their websites to simplify the presentation of their financial data.

Information visualization refers to the use of computer-supported, interactive, visual representations of data to amplify cognition or the acquisition and use of knowledge to help individuals carry out tasks more efficiently [17], [18].

We are visual beings who use sight as one of our most important senses for understanding information, and that is one of the main reasons for the use of information visualization. It is already used in

---

everyday life, from a graph in a newspaper to a map of a region with a route on it, providing an alternative or a supplement to textual information [19].

One of the advantages of information visualization is the possibility of using interaction allowing the exploitation of the data to uncover new knowledge. Another benefit is the opportunity of choosing the data we want to display providing a way to exploit only a subset of a large amount of data. We are also able to choose different types of representations to the same data adapting it to what we need to obtain [20].

The use of visualizations in finance is not a new area, techniques such as line graphs and bar charts that have the benefit of being very well known and easy to interpret, have been used for many years [6], [1], [21]. The changes that occurred in 2008 that have already been presented led to rapid growth in the development of new techniques and representations taking advantage of the benefits of information visualization.

According to a recent research [22], technologies that provide an alternative to the presentation of financial data will facilitate investor information gathering, improve disclosure transparency, and influence the investor decision process.

The application of information visualization in finance may bring an added value but remain under-exploited as has been seen previously.

1.1 Goals

The purpose of this work was to study if applying information visualization techniques allow to speed up the analysis of the data from the financial statements, exploring the gaps in this area and allowing a smooth, accurate and complete interpretation of that information.

As we have seen, the application of new forms of information visualization to financial statements is still minimal. Therefore the aim of our work was to present in different ways the data currently contained in tables, mainly in files in a Portable Document Format (PDF), allowing the visualization to help the strategic thinking of investors and other stakeholders by reducing the cognitive load, helping short-term memory and allowing those comparisons and inferences to be facilitated [23], [24].

To meet our goal, we developed an information visualization system called EconoVis that provides an overview of the information contained in financial statements from a large number of companies. This also allows a comparison between the different companies present in the system and each parameter from the data, in an interactive way facilitating the perception of some patterns in the information.

The development of this visualization passed for an incremental and iterative process receiving input from financial specialists to ensure that our research meets their needs, objectives, and improve their financial statement analysis giving them different perspectives of the data.
1.2 Document Structure

This document is structured as follows: in Chapter 2 we start with a survey and analysis of the related work related to this research. After that in Chapter 3 all the implementation process is explained from the initial steps made to the final prototype. The Chapter 4 is where all the evaluation process is detailed including the results and analysis of the Usability Tests and Case Studies performed. The conclusions of this study and future work are reported in Chapter 5.
2

Related Work

Contents

2.1 Stocks Data ................................................................. 9
2.2 Funds Data ................................................................. 14
2.3 Transactions Data .......................................................... 16
2.4 Companies Data ............................................................ 19
2.5 Discussion ................................................................. 21
As it was noted in Chapter 1, the use of information visualization in finance is not something new and, as expected, some studies have already been conducted in this area.

In this chapter, studies which have been done will be shown and towards the conclusion, the most important aspects of each type of visualization will be summarised by comparing each one, observing the tendencies of each group and concluding with the main influences for our study.

The search tools used for this study were surveys (Dumas et al. [25], Ko et al. [13]) and websites from some of the most important stock exchanges from Europe (Euronext\(^3\), London Stock Exchange\(^4\), and Deutsche Börse Group\(^5\). Last accessed on September 26th, 2018.) and the United States of America (USA) (New York Stock Exchange (NYSE)\(^6\) and Nasdaq\(^7\).

Following the division suggested by Ko et al. [13], our study will be categorized by financial data sources analyzing in each, the techniques and methods of information visualization used. Our categories will be Stocks Data, Funds Data, Transactions Data, and Companies Data.

### 2.1 Stocks Data

This category includes data that contains time series of share prices of companies over a period of time, including their trends, patterns, performances, and predictive analysis. The stocks data can also be combined with news media data to extract additional information in order to improve the investor’s knowledge.

To provide research contributions to the design of advanced visual data exploration interface, Merino et al. [1] study different visualizations, which have been employed for stock market data, and apply the results in a new visualization, named Stock Market Analyzer (Figure 2.1).

According to the study’s results, the most appropriate techniques for historical data are Line Charts and Recursive Patterns. To find patterns and explore large time series datasets, pixel techniques can lead to excellent results. For real-time data, Line Charts and Treemaps, have been considered as the two most appropriate techniques. Taking this into account the authors implemented the Stock Market Analyzer in which these techniques were applied.

The study also concluded that it is essential to present to the user an overall picture or overview of the entire data set, which should be possible with the application of interactive zooming and filtering methods, and details of the data should also be offered to analyze the found patterns.

Ziegler et al. [2] presented two applications for the user to analyze large amounts of time-series data interactively in real-time.

---

\(^5\)Deutsche Börse Group - http://deutsche-boerse.com/
Due to the recent crisis, a significant number of assets lost a considerable amount of value. Thus, analyzing and monitoring financial markets has become essential. Taking this into account, the authors implemented two applications that use visualization to explore large amounts of stock market data in order to get an insight into the development of assets, market sectors, countries, and the financial market as a whole.

The first of the two applications (Figure 2.2) allows users to analyze combinations of single assets, market sectors as well as countries, compare them to each other, and to visually discover the periods of time where market sectors and countries find themselves in a period of instability. The second application clusters a selection of large amounts of financial time series data according to their similarity and analyzes the distribution of the assets among market sectors [2]. Both applications are scalable to handle large quantities of data and have real-time functionality for fast and convenient interactive exploration and analysis.

Currently, the traditional line graph with a time axis and a price axis is still the most frequently used visualization technique in the financial domain. The main disadvantage of this graph type is that it only works well with very few time series at the same time [2]. Because of this, in the first application, the authors used a method related to the idea of the Pixel Bar Charts which transforms a two-dimensional line graph into a one-dimensional bar and uses color to code values. Ziegler et al. [2] focus on the analysis of market stability (volatility) rather than the performance (profits/losses) of single assets. To deal with a problem that appears when using colors as coding values (minimal changes in the value result in minimal changes to the color map which are often not perceivable anymore), the authors of this
study coded the relative percentage changes instead of coding the absolute values of an asset.

Because of the nature of financial time series data and the authors aim to cluster them regarding their trajectories, in the second application, they started by normalizing the time series data to make it directly comparable. Ziegler et al. [2] also used the *k-means algorithm* (process for partitioning an N-dimensional population into k sets on the basis of a sample [26]) for generating clusters due to its rapid computation, easy implementation, and ability to specify a desired amount of clusters.

This study concluded that with efficient preprocessing, both applications allow visual analysis and interactive exploration of large amounts of financial time series data in real-time [2].

A further example that uses time series data is the system of Schaefer et al. [3] (Figure 2.3) which introduces a novel visualization system for analyzing share performance from historical stock price time series and sector indices data.

The tool uses a pixel-based view, similar to the previous work from Ziegler et al. [2], that shows the global view of the whole market performance. Each pixel unit represents the performance of one share at a one-time point, and each row represents one share. To facilitate the analysis, the researched system has various user interactions implemented that allows filtering of the data, change time intervals and zoom into a particular period to obtain detailed information.

To visualize the correlation between individual share performance and the sector index, two different visual representations, triangle, and trapezoid polygons, were proposed. Moreover, to help to filter the database on share performance a colormap slider was implemented [3].

Schaefer et al. [3] used cases analyses which demonstrated that the system was able to highlight some interesting patterns in the data successfully, and the visualization techniques scale well to large
The paper of Sorenson et al. [4] describes one important visualization project (Figure 2.4) done by Bloomberg® in 2006-2007 combining the visual representation of continuous time series data (e.g. stock prices) together with associated discrete event data (e.g. news, earnings releases, announcements) [4].

This visualization enabled a single screen to visually organize a large volume of event data, to facilitate inference through visual alignment of related data, and to provide a workflow from the single point of access to a broad range of detailed information [4].

The authors also used a visual mapping technique to describe discrete event data to alphanumeric
pictographs online graphs, as opposed to using conventional abstract glyphs.

As the study was adopted by several users of the Bloomberg platform, the authors conclude that it was a successful study.

For this study, in addition to the searched papers, we also analyzed the information visualization techniques present on the websites of some of the leading stock exchanges in Europe and the USA (Figure 2.5).

Regarding the Euronext website (Figure 3.2(a)), there are visualizations of the stock price and the volume of the operations. It does not allow a comparison between the stocks of different companies but only with pre-defined indexes. The type of chart cannot be changed either.

In the London Stock Exchange website (Figure 3.2(b)), there is a development in the visualization available when compared to the previous case. The website has the option to change the chart type and zoom-into a selected area. The comparison is only possible with defined indexes and defined sectors of activity.

![Image](image1.png)

**Figure 2.5: Stock exchange websites**

The graphics of the Deutsche Börse Group website (Figure 2.5(c)) are similar to the ones in the
Euronext website. While the website allows more functionality, such as the comparison between the value of stocks of different companies, it can also add financial metrics and change the type of chart displayed.

On the USA stock exchanges, the NYSE website (Figure 2.5(e)) allows the use of a greater number of chart types and allows the comparison between company stocks. The Nasdaq website (Figure 2.5(d)) does not have as many chart options as the previous one, but it has some similar features corresponding to comparing stocks.

In common, all websites feature line graphs to represent the value of shares over time, and to the volume of operations, data bar charts are usually used.

### 2.2 Funds Data

Several studies were done using the funds’ data, mostly due to the ease of access to this kind of data. Usually, the study of funds data consists of studying groups of stocks gathered by investment considerations [13].

Mutual funds are a popular investment choice for private investors because they allow people to invest money and have it managed by professional fund managers [5].

Csallner et al. [5] have developed a system called *FundExplorer* that implements a distorted treemap to visualize the amount of money invested in a person’s fund portfolio and the context of remaining market stocks. This study was concentrated on equity mutual funds, a popular form of mutual funds that invests in stocks. The focus was on supporting people with an understanding of the presentation of the diversification of their portfolios so that they can make informed investing decisions.

The authors used a treemap, that is an information visualization technique which uses nested rectangles to create a space-filling representation of hierarchical data because it has been noted that treemaps are well suited to support decision-making processes in hierarchical structures [27], [28].

Visualization is based on the notion of distorting the classic treemap visualization, which allows the efficient use of screen real estate and visualizes all data points. As shown in Figure 2.6, a *Context Treemap* is used to visualize the stocks through a personal fund portfolio.

The research conclusion was that using the Context Treemap visualization in the *FundExplorer* system allows a better indication of how the investments of a person are situated within the overall stock market [5].

Dang [6] also used the treemap along with bullet graphs in their application for supporting the portfolio monitoring of a boutique asset management firm, as can be seen in Figure 2.7.

A further example using mutual funds data is the visualization of Alsakran et al. [7] where it is proposed a *tile-based parallel coordinate plot* where the plotting area is divided into rectangular tiles.
Parallel coordinates is a significant visualization tool for multivariate data representation and correlation analysis but is rarely used in financial visualizations because the visual clutter appears easily for only a few thousands of items, due to the spatial and resolution limit of the physical display devices, as well as the perception limit of the human visual system [7].

In this study the authors improved the parallel coordinates technique by innovating a new tile-based density image assembling the whole plotting area of parallel coordinates, by assigning different colors and opacity levels on the tiles according to the computed density, by providing a real-time human-interaction tool for users to continuously manipulate display results with different tile sizes, and by coupling the tile-based visualization showing global trends with interested outliers to help users study the
In Figure 2.8 it is possible to see the difference between the classic parallel coordinates view (Figure 2.8(a)) and the new tile-based parallel coordinates (Figure 2.8(b)) when using nearly half a million data items.

To demonstrate the benefits of their study, Alsakran et al. [7] provided a case study on a real-world dataset, the mutual fund data of the USA during the year 2006. The authors concluded that the tile-based parallel coordinates technique improves the performance, yields more controllability and promotes visual understanding.

2.3 Transactions Data

Transaction analysis is important because it can reveal hidden transactional patterns, malfunctions in business and evidence or symptoms of security breaches [13]. This category includes the data generated from different types of transactions such as transactions between companies, between countries or between bank customers.

An approach which used this type of data was the Domesticating Bead by Brodbeck et al. [8]. The authors tried to solve the problems described by a bank: gain an overview of the month’s trading activity and set guidelines for the next month’s trading, with a connection between detail and context and an overall relationship in the data. For this, Brodbeck et al. [8] used the Bead visualization system that employs an algorithm for laying out high-dimensional data in a low-dimensional space, and some features for image improvement [8].

Previously to this study, the system had been realized using 3D, but the authors shifted the representation layout to a 2D map display, adjusting all the image features, to make it easier to navigate, to solve problems like occlusion, and to adapt to working practices with which people in the financial world are familiar.

The imageability features which were introduced do not provide comparisons among details. This is an important characteristic because the axes of the layouts do not have an explicit meaning, so the positions of individual objects do not permit quantitative interpretations. To solve this problem, the
authors chose the method of parallel coordinates in combination with brushing and linking. With this technique, n-dimensional data points are visualized as polygonal lines across a set of parallel axes and can directly be compared quantitatively, and allows, for example, the comparison of outliers [8].

In Figure 2.9 we have an example of the Domesticating Bead visualization tool. The main sections are the layout of fixed income trading data, the control panel on the left, the bottom panel is used to show details of a selected object and, using sequential world-in-miniature maps, the display of the histories of searches at the top. The left control panel allows for the zooming in and out, controlling the size of individual objects, searching, pop-up controls, amongst others.

![Figure 2.9: Domesticating Bead visualization tool][8]

Changes that occurred in the bank with whom the authors were working with did not allow the authors to prove if the techniques used solved the problem, and for that reason, there is a lack of validation of this approach.

Chang et al. [9] suggested an approach to help large financial institutions to handle suspicious wire transactions that may occur in the middle of the hundreds of thousands of wire transactions per day that exists.

Hierarchical Interactive visual analysis with multiple linked views can effectively attack this problem because it is adapted toward the visualization and interactive exploration of massive datasets, integrating multiple methods from various disciplines such as information visualization, human-computer interaction, and statistics [9].
The WireVis [9] is a multiview approach that assists analysts in exploring large numbers of categorical, time-varying data containing wire transactions [9]. It uses four coordinated views of transaction activity: the keyword network view to represent the relationships between keywords, the heatmap view to show the relationships between accounts and keywords, the search-by-example tool to help discover accounts of similar activities, and the Strings and Beads to depict transactions over time. The authors also defined four requirements for the system: Interactivity, Filtering, Overview and Detail, and Coordinated Multiple Views.

In Figure 2.10 we have a view of the entire system showing the heatmap in the top left corner, the search-by-example in the top right corner, in the lower right corner the keyword graph, and in the lower left corner the strings and beads. The four views together detail the relationships among accounts, time, and keywords within the transactions, and present a global overview of the data, providing the ability to aggregate and organize groups of transactions for better investigation and analysis [9].

In collaboration with the Anti-Money Laundering (AML) division at Bank of America⁹, the authors demonstrate that using interactive visualization techniques coupled with hierarchical analyses in searching for suspicious financial transactions significantly enhances the analysts’ ability to see global trends and that the investigators can detect accounts and transactions that exhibit suspicious behaviors [9]. Moreover, even though the examples and results concentrate on wire transaction data, the authors mention that this approach is general and applicable to any financial transaction data.

⁹Bank of America AML - http://investor.bankofamerica.com/
2.4 Companies Data

The companies’ data includes information related to the performance of a company (e.g., sales, financial statements) during a year. Accessibility to data affects the study [13], and this is something that was noticed in this category where there are fewer research papers compared to the others that we have.

To quickly and accurately analyze the financial data that exists in financial statements, the self-organizing map seems to be a good option. To prove it Eklund et al. [10] evaluated the performance of self-organizing maps by analyzing the financial performance of 77 pulp and paper companies during 1995-2000.

Although many papers on self-organizing maps have been published, this study is different because very few studies have dealt with the use of self-organizing maps for financial analysis.

The authors collected financial data, in the form of seven financial ratios, using the Internet as the primary source of information. For this, Eklund et al. [10] searched for financial statements on the homepage of the companies, as well as in different databases on the Internet. For American, Canadian and Japanese companies the authors found useful databases to complement the financial information that they found, but in Europe, they did not find a good one, so they had to use the annual reports received via email to complement the information.

The self-organizing map technique creates a two-dimensional map from n-dimensional input data. In that map, it is possible to identify borders that define different clusters. These clusters consist of input variables with similar characteristics, i.e., in this report, of companies with similar financial performance [10].

![Figure 2.11: Self-organizing map of pulp and manufacturing companies [10]](image)

In Figure 2.11 the pulp and manufacturing companies were compared with each other. After comparing the results with the information contained in the annual reports Eklund et al. [10]
concluded that the results of the study indicate that self-organizing maps can be appropriate tools for the financial analysis of large amounts of financial data [10]. The authors also demonstrated that the results were easy to visualize and interpret and provided a practical way to compare the financial performance of different companies.

Another example of visualization using the companies data is the MarketAnalyzer [11] which is a visual analytics system for exploring, comparing, analyzing, and predicting trends of the point of sale data. For a business to be able to increase its current market share, it is required to maximize its profits within the market. For this, many companies generate intelligence reports extracting information from a variety of sources using several methods of data collection and analysis to be able to explore, analyze and predict the market share data changes that are relevant. One important key data source is the point of sale data that retailers share with vendors. This data is temporal, multivariate, and spatial in nature.

The MarketAnalyzer, as can be seen in the Figure 2.12, consists of multiple coordinated views linked with interactive filters: (a) Company filter, (b) Store filter, (c) Products filter, (d) Legend view, (e) (Sorted) Matrix view for sales, trends, and growth rates, (f) Stacked bar view, (g) Geographical view, (h) and (i) Line graph small multiples views, (j) and (k) Time slider widgets and aggregation tools for temporal comparison, (l) Tooltip, (m) Filter [11].

In this study, Ko et al. [11] used an enhanced pixel-based visualization approach, similar to the approach followed by Ziegler et al. [2] in its first application, to visualize large store and product information
in a limited screen space. Others series of linked views are also used including, line graphs, stacked bar graphs, and choropleth maps. Additionally, overview and detailed information are provided in a series of multiple coordinated views.

The authors concluded that this system could be easily applied to analysis with any other multivariate spatiotemporal data [11].

2.5 Discussion

In this chapter, we have seen several papers presenting different visualizations using diverse types of data related to Finance.

Most of the studies found used stock data and according to Ko et al. [13] this is expected due to the ease of access to this type of data. Because of this, there are few studies using companies’ data and mainly using financial statements, and this has led us to the choice of our goal.

Regarding the analyzed websites, there are very few visualizations and the ones that exist only use stock data.

In Table 2.1 we have an overview of all the studies and websites that were analyzed by us. In this figure, we pointed out the features of each analysis.

The features analyzed are Comparison, Focus+Detail, Search Mechanisms, Trends Over Time, Avoid Occlusion and Interaction. We divide the Interaction into Select, Reconfigure, Encode, Filter, and Connect. These features are important aspects to take into account in a meaningful visualization of the data that we have. Furthermore, these are the main key aspects for a visualization that meets general user needs in terms of interaction, providing powerful data analysis and inspection tools [18].

Regarding the Comparison feature, we confirmed if there were comparisons between data items of the same type. In the Focus+Detail feature, we looked for zooming features that allow a focus on the data that we wanted to see with more details. The Search Mechanisms, are some features to search for data on the visualization. The Trends Over Time feature refers to being able to find patterns over a period of time. Not to lose information, it is important that there not occur occlusion of information, which happens in the Avoid Occlusion feature. In a good information visualization, it is important to have interactivity; the Select features referred to has to the option of selecting a specific data item in the visualization to keep track of them. The Reconfigure feature is the possibility to view data from a different perspective (e.g., sorting). The Encode feature is when it is possible to change the visual representation of the data (e.g., color, size, different graph). The Filter feature is the possibility of defining specific conditions to the visualization. The last one is the Connect feature and is the association between different data items, being that item highlighted in different visualizations at the same time.

As can be seen in Table 2.1 the features which occur in a number of studies are Filter, Trends
<table>
<thead>
<tr>
<th>Data Source</th>
<th>Comparison</th>
<th>Focus + Detail</th>
<th>Search Mechanisms</th>
<th>Trends Over Time</th>
<th>Avoid Occlusion</th>
<th>Select</th>
<th>Reconfigure</th>
<th>Encode</th>
<th>Filter</th>
<th>Connect</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Stocks Data</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Merino et al.</em> [1]</td>
<td>✓</td>
<td>x</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>x</td>
<td>✓</td>
<td>x</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td><em>Ziegler et al.</em> [2]</td>
<td>✓</td>
<td>x</td>
<td>✓</td>
<td>✓</td>
<td>x</td>
<td>x</td>
<td>✓</td>
<td>x</td>
<td>x</td>
<td>✓</td>
</tr>
<tr>
<td><em>Schaefer et al.</em> [3]</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>x</td>
<td>✓</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td><em>Sorenson et al.</em> [4]</td>
<td>✓</td>
<td>x</td>
<td>✓</td>
<td></td>
<td></td>
<td>x</td>
<td>✓</td>
<td>x</td>
<td>x</td>
<td>✓</td>
</tr>
<tr>
<td><em>Euronext</em> 3</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>x</td>
<td>✓</td>
<td>x</td>
<td>x</td>
<td>✓</td>
</tr>
<tr>
<td><em>London Stock Exchange</em> 4</td>
<td></td>
<td></td>
<td>✓</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>✓</td>
<td>x</td>
<td>x</td>
<td>✓</td>
</tr>
<tr>
<td><em>Deutsche Börse Group</em> 5</td>
<td>✓</td>
<td>x</td>
<td>✓</td>
<td></td>
<td>x</td>
<td>x</td>
<td>✓</td>
<td>x</td>
<td>x</td>
<td>✓</td>
</tr>
<tr>
<td><em>New York Stock Exchange</em> 6 <em>(NYSE)</em></td>
<td>✓</td>
<td>x</td>
<td>✓</td>
<td>x</td>
<td>x</td>
<td>✓</td>
<td>x</td>
<td>x</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td><em>Nasdaq</em> 7</td>
<td>✓</td>
<td>x</td>
<td>✓</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>✓</td>
<td>x</td>
<td>x</td>
<td>✓</td>
</tr>
<tr>
<td><strong>Funds Data</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Fundexplorer</em> [5]</td>
<td>✓</td>
<td>x</td>
<td>x</td>
<td></td>
<td>x</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>x</td>
<td>✓</td>
</tr>
<tr>
<td><em>Dang</em> [6]</td>
<td>✓</td>
<td>✓</td>
<td>x</td>
<td></td>
<td>x</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>x</td>
<td>✓</td>
</tr>
<tr>
<td><em>Alsakran et al.</em> [7]</td>
<td>✓</td>
<td>x</td>
<td>x</td>
<td></td>
<td>x</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>x</td>
<td>✓</td>
</tr>
<tr>
<td><strong>Transactions Data</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Domesticating bead</em> [8]</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>✓</td>
<td>x</td>
<td>x</td>
<td>✓</td>
</tr>
<tr>
<td><em>Wirevis</em> [9]</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>✓</td>
<td>x</td>
<td>x</td>
<td>✓</td>
</tr>
<tr>
<td><em>Ekdund et al.</em> [10]</td>
<td>✓</td>
<td>x</td>
<td>✓</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td><em>Marketanalyzer</em> [11]</td>
<td>✓</td>
<td>x</td>
<td>✓</td>
<td></td>
<td>x</td>
<td>✓</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>✓</td>
</tr>
</tbody>
</table>
**Over Time, Connect and Comparision.** It is normal that the *Trends Over Time* feature happens in so many studies because the data used is obtained over time, which visualizations take into account. The *Connect* feature occurs in the visualizations that show different graphs, and there was a concern by the authors to interlink all the graphics making the view more fluid and using the *Filter* feature is something important since a large amount of data is used. To analyze multiple data items, the *Comparison* feature is important for searching for variation among those items, therefore, it is an important feature.

On the other hand, the features that appear less often are the Reconfigure and the Encode.

The studies that include more features are the *Wirevis* [9], the *Dang* [6], and the Domesticating bead [8], which have seven of ten features. All are useful studies that implement a visualization similar to that we intend to do in our study. Even though there are more studies using stock data, the ones that have more features are mostly those that use transaction data.

Our objective is to have all the features in our visualization and because of that we will base ourselves on applications with a large number of features since these are representative aspects that are desirable in a view that is effectively able to evidence the data that we intend to show and allows us to find aspects that are not immediately noticeable. We will inspire our research on those studies, but we will also present new ideas taking into account our goal.

We based our work not only on studies that have already been done using companies data but also with the other studies, because the data to be used is similar to the latter.
Visualizing Economic Trends in a Modern World

Contents

3.1 Initial Steps ......................................................... 27
3.2 Requirements ...................................................... 28
3.3 Low Fidelity Prototyping (LFP) .................................. 31
3.4 Architecture ......................................................... 35
3.5 Functional Prototype (FP) ......................................... 38
Taking into account the goals defined in Chapter 1 and the Related Work described in the last chapter, we developed a visualization that uses the information from the financial reports in order to give the users a new and improved way of using and analyzing information made available regularly by the companies.

The followed approach applies the techniques and methods of Information Visualisation (InfoVis) to implement a visualization capable of providing the user with the possibility of exploring the data in an interactively way, finding information more quickly and finding new trends. The user benefits from our visualization because it will simplify some everyday tasks and will provide new possibilities to find patterns and information which would be harder to find without it.

This chapter starts with the description of the first steps made by us in order to find the best approach to follow to give value to an area where more accurate information is essential.

### 3.1 Initial Steps

As we described in Chapter 2, this work began with previous context research, aiming to understand what had already done in the area of finance and what gaps existed. It was also essential to realize what would and would not be relevant to visualize.

From the Related Work and the conversations with Prof. Clara Raposo, an advisor in this study and also a specialist in the financial area, we noticed the importance of the analysis of the financial statements and also the gap in these kinds of visualizations.

Having taken notice of the above themes, we then tried to identify the interest of our plan for a financial institution. On September 16, 2016, we had a meeting in Euronext Lisbon with Dra. Isabel Ucha, an administrator of that institution. In that meeting, we realized that in Euronext Lisbon, they only manage the content present on their website and all the financial data are centralized and made available by their headquarters in Paris. For that reason, a partnership with them would not be possible in an early phase where there was nothing concrete to show. We decided to continue our research and would re-evaluate it at a later stage of the study if a new contact were seemed relevant.

After the context research analyzes made, we then decided to create a visualization that accurately shows the different types of financial statements data, and for that, we decided to follow an iterative and incremental approach with an emphasis on prototyping where the users had a crucial role in the process of obtaining the final prototype.

---

3.2 Requirements

When we decided that the best approach would be to use the data from the financial statements, we had to know the needs of the users and know better the data that we were working with. For that, we developed some tasks and questions and validated them with users in order to know which would be the most important for them and also to receive some feedback which would take advantage of their experience. To validate the questions and tasks and also to receive some feedback we designed a questionnaire and also organized interviews with financial specialists.

3.2.1 Tasks and Questions

The set of tasks and questions we want our visualization to respond to were:

1. **Compare multiple financial indicators throughout years for a specific company.**
   
   (a) **Description:** This task allows users to compare multiple financial indicators for a particular company during a defined interval of years.
   
   (b) **Questions:** Comparing EDP’s gross margin %, operating margin %, and net margin % key ratios, in 2012 and 2015, which margins improved every year? Was Jeronimo Martins’ (JMT) increase in revenues between 2012 and 2015 accompanied by an increase in its net margin %? Between 2014-2016, taking into account the Dupont Identity, which ratio (Net Margin %, Asset Turnover (Average), Financial Leverage (Average)) caused the decline in 2015 on Mota-Engil’s (EGL) Return on Equity %?

2. **Identify the value of a financial indicator for a particular company.**
   
   (a) **Description:** The user can find the value of a specific indicator for a company.
   
   (b) **Questions:** How much of Sonae’s (SON) total liabilities in 2016 was long-term debt? In 2016, what were Sonae’s (SON) total liabilities? What were NOS’s diluted earnings per share in the fiscal year 2013?

3. **Compare different financial indicators for a group of companies.**
   
   (a) **Description:** This task allows a comparison between different companies and between different financial indicators.
   
   (b) **Questions:** Comparing Jeronimo Martins’ (JMT) and Sonae’s (SON) gross margin % and net margin %, which company was more profitable in 2012? In 2015, comparing GALP and EDP current ratio and quick ratio, which company has higher asset liquidity?

4. **Realize the changes in a specific financial indicator throughout years.**
(a) **Description**: The user can notice the changes in an indicator during a defined interval of years.

(b) **Questions**: Did GALP’s Interest Coverage Ratio ever fall below 4 from 2008 to 2011? How did The Navigator’s (NVG) debt-equity ratio change from 2010 to 2012? From 2012 to 2015, in what year did GALP experience a decrease in its Investments in property, plant, and equipment? In the period 2010-2014, in which year did NOS experience an increase in its return on equity %? What was the change in CTT’s book value per share from 2013 to 2014? In the period 2011-2015, in which year did Sonae’s (SON) increase its Cash and cash equivalents? From 2012 to 2016 did GALP experiences in some period an increase in its basic earnings per share?

### 3.2.2 Survey

As previously stated, to choose the most relevant questions that we wanted to support in our application we surveyed a group of users to discern their opinion. The group was composed of six people: three finance professors, two investors, and one financial student. Their ages range from 18 to 54, and their experience in the profession spans from less than one month to more than nine years. They are divided into 67% male users and 33% female users.

The survey had four different parts, and it begins with an introduction to the work that we were developing and to the aims of the survey. We began with some questions to profile the user who was responding to the questionnaire. In the third part, we had 15 questions, and the user had to choose their importance from not important to very important on a scale from 1 to 5. The last part of the survey was the part of the thanks where we thanked the individuals for their collaboration and also an open question where the users could give some suggestions and new questions.

![Figure 3.1: Survey Results](image_url)
As can be seen in Figure 3.1, there are no questions with an average ratio lower than 2.50, and for that reason, we use all questions in our study giving priority to the questions that had an average ratio equal to or greater than 4.00 out of 5.00.

The 15 questions that we used ordered by the average answer rating of each of them was:

1. Comparing EDP’s gross margin %, operating margin %, and net margin % key ratios, in 2012 and 2015, which margins improved every year? **(4.83)**

2. How much of Sonae’s (SON) total liabilities in 2016 was long-term debt? **(4.50)**

3. Was Jeronimo Martins’ (JMT) increase in revenues between 2012 and 2015 accompanied by an increase in its net margin %? **(4.33)**

4. Between 2014-2016, taking into account the Dupont Identity, which ratio (Net Margin %, Asset Turnover (Average), Financial Leverage (Average)) caused the decline in 2015 on Mota-Engil’s (EGL) Return on Equity % **(4.33)**


7. From 2012 to 2015, in what year did GALP experience a decrease in its Investments in property, plant, and equipment? **(4.00)**

8. Comparing Jeronimo Martins’ (JMT) and Sonae’s (SON) gross margin % and net margin %, which company was more profitable in 2012? **(4.00)**

9. In 2015, comparing GALP and EDP current ratio and quick ratio, which company has higher asset liquidity? **(4.00)**

10. In 2016, what were Sonae’s (SON) total liabilities? **(3.83)**

11. What were NOS’s diluted earnings per share in the fiscal year 2013? **(3.83)**

12. In the period 2010-2014, in which year did NOS experience an increase in its return on equity %? **(3.83)**

13. What was the change in CTT’s book value per share from 2013 to 2014? **(3.33)**


15. From 2012 to 2016 did GALP experiences in some period an increase in its basic earnings per share? **(3.17)**
3.2.3 Interview

From the interview with two finance professors of Instituto Superior de Economia e Gestão (ISEG), we can conclude that this area is not an exact science, and the use of financial data will always be analyzed for each person depending on the type of use that each one does. In our case, one lower rating does not mean that there is a lack of interest for that particular question but only for some depending on if they are a research user or an investor.

There were some significant comments that they made which reinforced the importance of our study, for example, one of the users said:

“For investors, it is not enough to see the debt as it is, it is interesting to see the debt is one side of the equation, and the other side is the growth of the company, to combine the two.”

As our visualization will support a comparison between different financial indicators, the problem described will be addressed with our study.

Also, the other user strengthens the utility of our approach when he said:

“The work is interesting, and it is a friendlier way to acquire information and receive information.”

3.3 Low Fidelity Prototyping (LFP)

As we are following an incremental and iterative approach, and after validating the tasks and questions, which will be answered, it was during this phase we used a Low-Fidelity Prototype (LFP) to sketch out our visualization. The use of these types of prototypes is useful because they tend to be simple, cheap and quick to produce and test [29].

A financial specialist performed the validation of the prototypes in order to help us find the best visualization for this kind of data.

3.3.1 The Sketches

Our initial goal was to show a design for our dashboard, what it should contain and how each idiom would combine there. For that, we sketched our proposal for the dashboard that we proposed.

As it can be noticed in Figure 3.2, the dashboard should contain an area where all the idioms are available, an option to add new idioms easily, some option to see the companies in use at each moment, and a button that would allow for displaying a menu with more options. Those options could be, for example, a filter for the dates or a search mechanism to quickly find a company.

Each idiom can also contain options to minimize, close and a help button to allow the user to customize the dashboard in any way that suits it.
In Figure 3.3 we had two different approaches to task 1, task 2 and task 4. We had a Grouped Bar Chart and a Scatter Chart where the x-axis represents the different years, the y-axis expresses the values, and in the Bar Chart the bars represents each company, and in the Scatter Chart, the companies are represented using a different geometric figure. Both of them allow the user to see the same data, but in this phase, we wanted to have multiple options before choosing and validate the final one. For that reason, we also sketched a different option where we did a heatmap with axis allowing the heatmap to show the value and the axes will be for the companies (y-axis) and the years (x-axis) as can be figured in Figure 3.4.

Some parameters are a combination of other ones, for example, the Total Assets is a sum of the Total Liabilities and the Total Stockholders’ Equity, and for those cases, we also sketched two different visualizations. This visualizations also deals with task 2. As can be noticed in Figure 3.5 we suggested the Ring Chart and the TreeMap to deal with these cases where each part of the idiom is the value of one of the parameters.
3.3.2 Validation

After sketching the different alternatives and as was mentioned before we had to perform the validation of each sketch and in some cases choose the better and more completed visualization, for that we relied on the help of a financial specialist.

Regarding the dashboard sketch we figured out that would be a good starting point, and some modifications would only be needed if we found something for a more advanced prototype.
We then analyzed the Figure 3.3 where we had a Grouped Bar Chart and the Scatter Chart. The initial analyses showed that with a large number of companies the geometric figures are not good and the colors would work better. A different approach to that problem could be, use the same geometric figure in the Scatter Chart and change the color depending on the company. Another problem found was regarding the close values wherein the Scatter Chart the geometric figures would be very close to each other and in some cases overlapping, in those cases the Grouped Bar Chart would work better. The fact that bar charts are a more common visualization would also be referred to by the specialist as an advantage. Taking into account the problems referred, we decided to choose the Grouped Bar Chart because we wanted to avoid overlapping of the information and a more common visualization for the users to be more natural and quicker to obtain information.

Also from the Heatmap present in Figure 3.4 we concluded that were an excellent idiom to be used as a complement to the Grouped Bar Chart as it would show the same data but display it differently. In that way, it would allow the user to choose the one that they prefer and in some cases could even show new patterns not visible in the other visualization.

The next sketches that we examined were the presents in Figure 3.5 where the analysis focused on two idioms, the Ring Chart, and the TreeMap. The biggest problem found with the Ring Chart was related to the fact that is not easy to interpret which one is the most significant value when the values are very similar. Also when there are a large number of parameters, and even though in the case that we are testing we only needed three different parameters, we expected that our approach could be scalable. For the previous reasons, in that case, we choose the TreeMap as one of the idioms that we use in our visualization.

![Parallel Coordinates Sketch](image)

**Figure 3.6: Parallel Coordinates Sketch**

After analyzing the sketches that we had prepared, we noticed that we needed one more case to deal with the task 3. For that case, and taking into account that the line charts were one of the most used charts in finances, as we explained in the Related Work chapter, we sketched a Parallel Coordinates
Chart as can be seen in Figure 3.6. The idea was that each axis would represent each parameter with its value, and each line represents one different company, differentiating them by a different color.

3.4 Architecture

Our visualization was developed following a layered architecture because it supports the incremental development of systems, is changeable so an equivalent layer can replace another one. Moreover, when some layer is changed or updated, only the adjacent layer is affected [30]. Also, every layer of the EconoVis application can be used individually with other similar applications or can be easily changed without compromising the other layers.

![EconoVis Architecture Diagram]

Figure 3.7: EconoVis Architecture

The three main layers that compose our system are the Presentation, the Business and the Database Layer, as can be seen in Figure 3.7.
3.4.1 Database Layer

The Database Layer is responsible for managing all the data that we use in EconoVis. It receives data from the Import Script and answers the requested information from the Business Layer.

In Portugal, in contrast to the USA, there is no legislation specifying that companies should disclose financial statement information in a language like XBRL, and for that reason, that information is usually found in tables inside PDF files.

We decided to use the financial statement data of the companies in the Portuguese Stock Market\textsuperscript{11} to serve as a sample for our research work. These data came from the Income Statement, Balance Sheet, and Cash Flow from the Financial Statements of each company and also there are some Key Ratios using the data from the other statements.

The data come from the Morningstar database\textsuperscript{12}, and to use it we developed a script based on an open-source project, the Morningstar API\textsuperscript{13}.

The Import Script allows for the download of all the financial statements for all the companies that we requested. The updated version is set to download the Income Statement, Balance Sheet, Cash Flow, and Key Ratios, all together or each one individually, and all the companies on the Portuguese Stock Market can also be downloaded individually. The script will download the chosen files in a Comma-Separated Values (CSV) format and will put them inside a folder. The Business Layer will update the Database with the new files added by the script.

The Import Script is easily customizable, and it is possible to change, for example, to download the stocks of a different stock market.

For the database we used MongoDB that is an open-source document database which was designed to be robust, flexible and scalable [31]. It uses an ordered set of keys with associated values that in JavaScript is represented as objects. Our decision for this type of database was because it is open-source that allows scalability in the future and it uses objects that allow for the storing all kinds of data quickly.

In the EconoVis we store the sets of financial reports for each company and also a collection of colors that we use in our visualizations.

The system is ready to be easily customized, and to receive new values in order to improve and display more information in the visualization.

3.4.2 Business Layer

This layer is principally composed for the Node.js server that is an open-source server-side JavaScript environment that supports long-running server processes [32]. The main reason for the server chosen was because it is open-source and that works with JavaScript that is the primary programming language of our visualization so in that way we keep a consistency in the languages used.

With Node.js we also used Mongoose that provides a straight-forward, schema-based to model the application data and that works with Node.js, MongoDB, and Express that is a minimal and flexible Node.js web application framework that provides a robust set of features for web applications.

With Mongoose we structured our MongoDB database using schemas and connected the Business Layer with the Database Layer. With Express we built a Representational State Transfer (REST) Application Programming Interface (API) that allowed the interaction between the Presentation Layer and the other two databases.

The first thing that the Business Layer is responsible for is to import from the database all the files that the Import Script has already downloaded, structuring all the data using javascript objects. This layer is also responsible for the definition of the REST API, defining the specific format of each Uniform Resource Locator (URL) that when a call is made it will return a JavaScript Object Notation (JSON) file with the information requested.

This REST API allows that the data we used can be accessed and used by other applications that only need to make a specific call.

3.4.3 Presentation Layer

The third layer is the Presentation Layer that is responsible for presenting the visualizations to the user and where the user can interact with the dashboard and obtain the visualization that wants. This layer will interact with the Business Layer through calls to the REST API.

For development we used HyperText Markup Language (HTML), Cascading Style Sheets (CSS), Bootstrap, JavaScript, jQuery and D3.js. The HTML is used to structure the entire dashboard, the CSS and the Bootstrap are responsible for the style of the application. All the logic was made using the JavaScript and its libraries the JQuery and the D3.js. This last one will be responsible for building the visualizations that we have.

---

To access the EconoVis application the user only needs to access the application through the Internet using a browser. With this option, the application is always available. There is also the option of using the application in a local environment but, for that, it is necessary to install the database and the server in order for the dashboard be accessed.

### 3.5 Functional Prototype (FP)

After having validated the LFP and having a set of tasks that we had to support, we proceeded to the creation of the functional prototype. Following the iterative and incremental process, an evaluation of the first version had been made by financial experts. This evaluation was made through direct observation and with the think-aloud technique.

With the list of changes and the new features that we collected from this interaction, we made the recommended changes and added new features. Once again a validation was made, and finally, the final version was reached, which was then evaluated as we will see in the next chapter.

We started by developing the Database and the Business Layer and then imported all the data. By doing this in the first place, it allowed us to have real data to be able to use with the idioms and also allowed us to focus on the Presentation Layer that would be the priority for this study.

To develop these first two layers we had to define and structure the data that we had and define how best to store them. We defined that each company was an object that had a list of financial reports and that could have up to four objects that will be each of the statements that we are importing (Balance Sheet, Cash Flow, Income Statement and Key Performance Indicators). In addition to the statements, each company also had the company name and its abbreviation.

After defining the database structure, it was also essential to develop the REST API structuring all the different calls that would be supported by the business layer.

When the development of these two layers and the tests were finished, we began to develop the first version of the Presentation Layer.

#### 3.5.1 Initial Version

We decided to start by developing the dashboard where each of the visualizations would work in order to have a base where we could then try out each of the idioms we had developed.

Following our LFP we split our dashboard into three main areas, a navigation bar, a sidebar, and a main central area.

The navigation bar would be located at the top of the page and would contain the logo, the search bar and other menus that could be added.
The sidebar would be located on the left side and would contain the system settings. Since we wanted the focus of the user to be the visualizations we decided that this sidebar would not always be visible and that the user could hide it, for that, a button would be added and when the bar was hidden that button would be displayed.

The main central zone would be the core of our dashboard, and it would be where all the visualizations would appear and where the user would interact with them. We wanted to make this area as clean as possible so that the user focus could be on the views.

As more and more devices with different screen sizes exist, we wanted the dashboard to adapt automatically to each of them. For that, we developed our dashboard to be responsive so the dashboard could adapt to the size of the user's screen which facilitates its use and analysis of the information.

Since we wanted the user to have the freedom to customize the location of the idioms, we decided to create a grid structure by dividing the main area into small containers where each view would be and for which the user could move.

These containers would have each idiom, a descriptive title of it, and three buttons: a minimize, a help button and also a button to erase the view. The container would also have the possibility to be resized and to be moved to another grid element. With these features, the user could increase the graph they were seeing or move closer to another in order to facilitate their analysis. Being able to be minimized allows the user to keep the visualization on his dashboard but minimize it when it does not need it in order to gain more space for the others.

Another required functionality in the dashboard was the option to add new visualizations, and for that, a button was added. Clicking in this button would pop up a dialog component where the user could choose the required data for each visualization (Figure 3.8). The user could then choose the company they needed, and the parameters they wanted. It was in this phase where our REST API was necessary because the options that appeared to the user were those that were stored in the database and to access them, a call was used to obtain for example the list of companies. Also, statements types and parameters were obtained in this way, so when changes were made to the companies in the database, these changes were automatically updated in the visualizations.

After selecting all the values for each field, a new call was made to our server obtaining all the data taking into account the choices made, and then each idiom would use it.

Although many data are being used, the performance of the system has never been affected, so the required data could be obtained quickly without any performance problems.

As we wanted the comparison between different companies and also different parameters, it was necessary to allow more than one option to be selected both in the selection of companies and in the selection of the parameters. In addition to the multi-selection feature, we also needed a searching mechanism to found companies and parameters which would drastically decrease the time it takes to
add new idioms to the dashboard. To address this, we used Select2\(^{22}\) that has everything that had been requested. The user was then able to select more than one company or parameter, and it was also possible to search by name. These functionalities were also available in the navbar, in the search option, where the choice of the selected companies was possible and where the information for the user of the companies selected at any moment is available.

After these developments, it was then possible to begin to develop each of the views and test them using the dashboard structure previously made.

The first chart we developed was the Grouped Bar Chart (Figure 3.9). We intended that each group of columns represents a different year, that each column represents a company and that the value of the column was the value of the parameter to be analyzed.

Given the LFP, we developed a similar graph using D3.js where data used by this visualization would be obtained through the JSON file that the Business Layer call returned. About the colors used we decided to assign a different color to each company and in this way we ensured that there was always the same color for the same company.

For the user to identify the color of each company was also added in this chart a caption containing the color and stock symbol of the company. When we already had a graph similar to the one we had previously drawn, we moved on to the next view, the TreeMap.

\(^{22}\text{Select2 - https://select2.org/}. \text{Last accessed on September 26th, 2018.}\)
For the TreeMap we followed the steps we had done for the last graph, that is, we used the LFP as a starting point and developed a chart similar to what we had planned (Figure 3.10). The idea of this graph was to show how a parameter was subdivided into others allowing to perceive the composition of a specific parameter. Concerning the colors we used a function of D3.js, the d3.schemeCategory in combination with an ordinal scale giving thus different colors for each one of the presented parameters. As in the previous one, the data used in this visualization are those received from the call to the EconoVis REST API with the selected parameters and companies.

Subsequently, the development starts of the graph of the parallel coordinates that allowed answering
the task of comparing different financial indicators for a group of companies.

Each vertical axis would be a different parameter with its scale, and each line that crosses the axes would be a different company.

The most significant difficulty in this chart was the question of each axis had a scale independent of the others and to solve this problem we had to define the maximum and minimum value of each parameter individually and not for the whole set as we did for example with the Grouped Bar Chart.

To promote the association between the color of the line and the company we added, as we previously had done in the Grouped Bar Chart, a legend with the color of the line and the stock symbol of the company.

We also have added some interactivity features, such as the ability to switch the axes by moving to one side and switching it with what is on their side. Also, a selection tool that allows the selection of a set of lines on the axis and it will be highlighted while the others were colorless in order to lose their prominence. This kind of interactivity helps the user when filtering and analyzing the data they use.

Once again we validated the chart we developed with the one that we had outlined in LFP in order to see if we were meeting what we had validity and defined.

After we finished developing this visualization, and following the iterative and incremental process, we tested the first version of EconoVis (Figure 3.11) with a financial expert. With they feedback, it was possible to make a list of changes and improvements for the final solution.

![EconoVis Initial Version](image_url)

**Figure 3.11:** *EconoVis Initial Version*

In the tests we did on the first prototype, we identified many problems and improvements that could be made which we addressed in the final version of the prototype. The biggest problem was about the
color, having a color per company meant that the colors might not work well together when choosing companies with entirely different colors. Also, the colors used in the TreeMap had been criticized. Solve this problem would have to be our priority in the next iteration.

Another problem was regarding the format of the values in the graphs, there would have to be a format for the ratios in which there would be no units and another one for the monetary values, and in these we would have to present different alternatives depending on whether it is millions, hundreds, and so on.

In the Grouped Bar Chart when there were negative values, the columns overlapped the value of the axis something that was also commented on and taken into account for the next version.

Finally, the lack of a possibility of filtering the graphs by the data types was something pointed out because in finance often use only data from one of the reports and as such an option would be beneficial.

In general, the user liked this first version, having seen its vast potential. The comments we received were taken into account in the next iteration in order to meet the needs of the financial users.

3.5.2 Final Version

We started by solving the problem we had concerning the color scheme. Since we had a large number of companies, having a color per company, it implied that we also had a large number of colors and that would turn out to be entirely different between them.

We needed a palette of colors that vary in brightness, in which the colors could be distinguished from each other, that we had enough different colors and that integrated well with the visualization.

However, before defining a color palette, we had to make a decision, keep one color per company or use a set of colors per visualization, in which in the same graph the same company had the same color but in another graphic already had a pallet color set to that.

![Color Palettes](image)

Since we had a large number of colors and considering that we wanted to leave the system ready for more companies from other markets, the solution of having one color per company would be complicated.

We then decided to define one set of colors per chart and to maintain consistency by always associating the same color with the same company in two graphs of the same type.
We found three palettes of colors\(^\text{23}\) (Figure 3.12) that went against what we wanted, each one with 12 colors and, we assigned each palette to each of the three charts. The blue tones were for the Grouped Bar Chart (Figure 3.13), the pink tones for the parallel coordinates (Figure 3.14), and the orange tones for the TreeMap (Figure 3.15).

With these colors, each view will have a set of colors that will be added as new companies are added.

![Initial Version](image1)

![Final Version](image2)

(a) Initial Version  (b) Final Version

**Figure 3.13:** Group Bar Chart New Colors

- The second problem which needed to be solved was the format of the axes. We started to solve the cases where they were ratios, in these cases no unit appears.
- For the other units what we did was add a "B" when they are Billions, an "M" when there are Millions and "k" when there are Hundreds. In this way, the user would have a smaller number presented, but it allowed him a much more natural reading and interpretation of the information, something that met our objective for this study.

Regarding the problem of the Grouped Bar Chart, to solve it we had to make the x-axis always positioned relative to the minimum value, so even with negative numbers the axis will not be on top of the graph, something we wanted to avoid (Figure 3.16).

To filter by financial reports, which had also been requested, an option was added to the sidebar (Figure 3.17), for this the names of each of the reports were placed, and each time the user clicks on them, it will act as a button (showing or hiding). This filter only affects the display of the visualization because it remains in the system being affected by edits such as remove or add new companies. There is also an ALL button that shows or hides all graphs.
After we resolved the reported issues, we started adding new features to our dashboard.

The first of these was adding a tooltip in the Grouped Bar Chart and the Parallel Coordinates Chart. The first contains the company name and value (Figure 3.18(a)), while in Parallel Coordinates it contains the name of the selected company (Figure 3.18(b)). These tooltips allow the user to interpret the data and get the information they need more rapidly.

We also adapted the behavior of the dashboard to the case where the user removes all the companies, in that case, all containers are minimized and only return to their normal state when a new company is added (Figure 3.19).

In order to filter the view for years, a bar has been added at the top of the dashboard (Figure 3.20) that allows the user to select the year or range of years that they want to filter, updating all views whenever there is a change in those values.

The charts are also updated when a new company is added by adding this new company to existing charts.

Finally, we had to make a selection of our development priorities and since the heatmap would only complement the information we already had with the other visualizations, we decided not to do so.
Therefore, it is an option that could be done as future work.

We did another validation with an expert, and we realized that we already had a stable and final version (Figure 3.21) that could answer the questions and tasks that we had defined and that would allow for verification of the objective of our study.

For that verification, it would be necessary to make tests with the users, which will be described in the next chapter.
(a) Close Sidebar

(b) Open Sidebar

**Figure 3.21**: EconoVis Final Version
4 Evaluation

Contents

4.1 Usability Tests .................................................. 51
4.2 Case Studies ..................................................... 59
4.3 Discussion ....................................................... 59
After finalizing the iterative and incremental development part of our visualization, we tested it with a group of users to gather a set of quantitative and qualitative usability metrics. For that, we divided the evaluation into two main parts: Usability Tests and Case Studies. In both cases, users with experience in the finance area were used, but in the usability tests, the objective was to evaluate the interactivity and usability of EconoVis, whereas the primary goal of the case studies was to realize the utility and functionality of our system using for that, users with more experience.

4.1 Usability Tests

The first evaluation we made were the usability tests that using a summative evaluation we wanted to evaluate the success of the final version of the visualization and to ensure that it met precise, and measurable performance goals [33].

We performed the tests with a group of fifteen users who had a financial background. Each user received a survey and a list with the questions that we defined in 3.2.1. Their performance was measured quantitatively; estimating the time it took the user to do each question, the number of errors made and its satisfaction level in each one.

To perform these tests we had to previously prepare all the necessary material used, after that we had to do each test individually and in the end, the analysis of the results was made.

4.1.1 Protocol

To ensure that all the users evaluated the system under the same conditions we had to prepare the protocol which needed to be followed to have a controlled environment where all the users perform the same. For this, we created a survey (Appendix B) which was divided into three parts, the demographic survey, the questions to be answered using EconoVis and finally the System Usability Scale (SUS).

The test started with a contextualization about the study that we were doing explaining the goals of the research and the type of evaluation that we were performing. We also clarified that what we were testing was the system and not the users in order to remove some unnecessary pressure and nervousness in each user.

After that, a demographic survey had been performed allowing us to have a profile of the study, and for that, we collected demographic information such as the age, the gender and their relationship with the finance area.

The user then had five minutes to familiarize themselves with EconoVis and also where we explained the system, presenting all visualizations and the features that EconoVis had. No overpass the established time in this phase was crucial for all users to start with a similar experience with the system to have better test results.
After the previous phase, the users started to go through each question informing us when they were ready to begin, for us to start recording the duration and the number of errors. Each task was considered not to have any error when the user correctly answered the question without any assistance.

After each task, the users had to classify their satisfaction grade to each question. The order of the questions that each user followed was:

1. How much of Sonae’s (SON) total liabilities in 2016 was long-term debt?
2. In 2016, what were Sonae’s (SON) total liabilities?
3. In the period 2011-2015, in what year did Sonae’s (SON) increase its Cash and cash equivalents?
4. Was Jeronimo Martins’ (JMT) increase in revenues between 2012 and 2015 accompanied by an increase in its net margin %?
5. Comparing Jeronimo Martins’ (JMT) and Sonae’s (SON) gross margin % and net margin %, which company was more profitable in 2012?
6. What were NOS’s diluted earnings per share in the fiscal year 2013?
7. In the period 2010-2014, in which year did NOS experience an increase in its return on equity %?
8. What was the change in CTT’s book value per share from 2013 to 2014?
9. How did The Navigator’s (NVG) debt-equity ratio change from 2010 to 2012?
10. Between 2014-2016, taking into account the Dupont Identity, which ratio (Net Margin %, Asset Turnover (Average), Financial Leverage (Average)) caused the decline in 2015 on Mota-Engil’s (EGL) Return on Equity %?
11. Did GALP’s Interest Coverage Ratio ever fall below 4 from 2008 to 2011?
12. From 2012 to 2015, in what year did GALP experience a decrease in its Investments in property, plant, and equipment?
13. From 2012 to 2016 did GALP experiences in some period an increase in its basic earnings per share?
14. Comparing EDP’s gross margin %, operating margin %, and net margin % key ratios, in 2012 and 2015, which margins improved every year?
15. In 2015, comparing GALP and EDP current ratio and quick ratio, which company has higher asset liquidity?
The chosen order allowed the user to start with more straightforward questions to become more familiar with the system and to avoid frustration in the initial phase.

After finishing the fifteen questions, the users then had to answer ten more questions related to the SUS. These questions allowed a global view of subjective assessments of usability [34] about the user experience with EconoVis. For each question, we had a five scale option from Strongly Disagree (1) to Strongly Agree (5) following what had been established by Brooke et al. [34].

In the end, we thanked everyone for their contribution and also to took note of their opinions.

4.1.2 Results

The usability tests were performed by fifteen users where 53.3% (8 users) were female, and 46.7% (7 users) were male. Their age group for 86.7% (13 users) was from 18 to 24 years old and for 13.3% (2 users) from 25 to 34 years old. All of them are either current financial students (73.3%), or have already graduated (26.7%), and their experience in the economic area spans from less than one month to 9 years.

In the next sections, the results from the Response Time (4.1.2.A), the number of Errors (4.1.2.B), the Satisfaction Level (4.1.2.C) and the SUS (4.1.2.D) will be analyzed.

4.1.2.A Response Time

The duration of each question performed by each user is presented in Table 4.2 with also the statistics about those values, as the minimum, the maximum, the average time (Mean), the standard deviation (SD) and the confidence interval (CI) with a confidence level of 95%. To complement the table, we also created a box-plot that is shown in Figure 4.1 that includes the median, the minimum, and maximum values and even the lower and upper quartiles.

After analyzing Table 4.2 and the corresponding box-plot some conclusions could be made.

The questions with a longer duration were Q04, Q07, and Q10 with an average time higher than one minute (60 seconds). Q10 was the one where users took longer, and this was expected because of two main reasons. Firstly, this question had a financial concept that for some users they were not familiar with (the Dupont Identity). The second reason was that this question involved using at least three parameters and for those reasons, the users took longer, first for understanding the financial concept and afterward to look for all the requested information.

In Q04 the main reason for the higher duration time was because the user needed parameters from two different financial statements and this took longer because they had to combine two different kinds of information.

Q07 was a question where we noticed that some users had more problems understanding the financial parameters that were requested and for that reason, we can see that some users took 35 seconds.
### Table 4.1: Duration results

<table>
<thead>
<tr>
<th>Users</th>
<th>Q01</th>
<th>Q02</th>
<th>Q03</th>
<th>Q04</th>
<th>Q05</th>
<th>Q06</th>
<th>Q07</th>
<th>Q08</th>
<th>Q09</th>
<th>Q10</th>
<th>Q11</th>
<th>Q12</th>
<th>Q13</th>
<th>Q14</th>
<th>Q15</th>
</tr>
</thead>
<tbody>
<tr>
<td>U01</td>
<td>16.33</td>
<td>8.00</td>
<td>20.48</td>
<td>58.02</td>
<td>16.26</td>
<td>20.08</td>
<td>35.45</td>
<td>23.72</td>
<td>36.75</td>
<td>114.38</td>
<td>38.61</td>
<td>37.97</td>
<td>26.83</td>
<td>68.35</td>
<td>27.91</td>
</tr>
<tr>
<td>U02</td>
<td>28.80</td>
<td>3.00</td>
<td>17.32</td>
<td>37.65</td>
<td>37.02</td>
<td>21.94</td>
<td>73.98</td>
<td>41.25</td>
<td>41.66</td>
<td>56.63</td>
<td>28.60</td>
<td>21.11</td>
<td>19.73</td>
<td>34.36</td>
<td>52.09</td>
</tr>
<tr>
<td>U03</td>
<td>31.29</td>
<td>14.00</td>
<td>19.22</td>
<td>89.18</td>
<td>48.65</td>
<td>16.60</td>
<td>75.37</td>
<td>70.24</td>
<td>56.58</td>
<td>87.80</td>
<td>42.66</td>
<td>26.02</td>
<td>17.80</td>
<td>51.51</td>
<td>47.61</td>
</tr>
<tr>
<td>U04</td>
<td>18.83</td>
<td>7.52</td>
<td>22.31</td>
<td>96.26</td>
<td>28.32</td>
<td>41.18</td>
<td>82.68</td>
<td>29.16</td>
<td>40.67</td>
<td>77.28</td>
<td>34.34</td>
<td>43.04</td>
<td>42.37</td>
<td>61.95</td>
<td>43.83</td>
</tr>
<tr>
<td>U05</td>
<td>12.56</td>
<td>14.83</td>
<td>64.04</td>
<td>62.82</td>
<td>54.13</td>
<td>62.45</td>
<td>63.80</td>
<td>52.44</td>
<td>90.03</td>
<td>94.90</td>
<td>43.66</td>
<td>51.05</td>
<td>64.39</td>
<td>86.43</td>
<td>44.91</td>
</tr>
<tr>
<td>U06</td>
<td>43.12</td>
<td>6.22</td>
<td>14.23</td>
<td>71.87</td>
<td>64.83</td>
<td>71.66</td>
<td>64.07</td>
<td>61.53</td>
<td>34.52</td>
<td>67.78</td>
<td>47.30</td>
<td>44.23</td>
<td>41.03</td>
<td>48.69</td>
<td>56.92</td>
</tr>
<tr>
<td>U07</td>
<td>52.41</td>
<td>6.83</td>
<td>61.78</td>
<td>109.93</td>
<td>54.71</td>
<td>43.69</td>
<td>48.33</td>
<td>37.85</td>
<td>42.21</td>
<td>86.37</td>
<td>44.04</td>
<td>50.17</td>
<td>49.15</td>
<td>52.19</td>
<td>51.62</td>
</tr>
<tr>
<td>U08</td>
<td>110.20</td>
<td>5.47</td>
<td>19.20</td>
<td>66.17</td>
<td>39.81</td>
<td>33.94</td>
<td>67.22</td>
<td>37.83</td>
<td>43.86</td>
<td>95.21</td>
<td>54.17</td>
<td>46.38</td>
<td>44.70</td>
<td>25.82</td>
<td>46.98</td>
</tr>
<tr>
<td>U09</td>
<td>70.00</td>
<td>34.51</td>
<td>68.44</td>
<td>100.62</td>
<td>76.53</td>
<td>73.10</td>
<td>78.84</td>
<td>39.97</td>
<td>52.59</td>
<td>81.96</td>
<td>53.60</td>
<td>47.80</td>
<td>63.73</td>
<td>73.43</td>
<td>67.50</td>
</tr>
<tr>
<td>U10</td>
<td>22.00</td>
<td>35.00</td>
<td>19.56</td>
<td>53.93</td>
<td>53.17</td>
<td>59.70</td>
<td>57.19</td>
<td>50.75</td>
<td>34.18</td>
<td>80.31</td>
<td>45.89</td>
<td>32.82</td>
<td>26.23</td>
<td>55.87</td>
<td>42.88</td>
</tr>
<tr>
<td>U11</td>
<td>45.68</td>
<td>10.55</td>
<td>21.91</td>
<td>56.54</td>
<td>34.95</td>
<td>38.88</td>
<td>65.78</td>
<td>47.01</td>
<td>39.02</td>
<td>113.38</td>
<td>46.92</td>
<td>45.01</td>
<td>26.68</td>
<td>39.26</td>
<td>45.90</td>
</tr>
<tr>
<td>U12</td>
<td>35.06</td>
<td>24.76</td>
<td>15.82</td>
<td>91.41</td>
<td>17.06</td>
<td>50.72</td>
<td>41.09</td>
<td>34.26</td>
<td>46.10</td>
<td>64.49</td>
<td>62.36</td>
<td>31.90</td>
<td>53.40</td>
<td>59.09</td>
<td>33.89</td>
</tr>
<tr>
<td>U13</td>
<td>20.31</td>
<td>36.14</td>
<td>51.17</td>
<td>94.20</td>
<td>55.06</td>
<td>33.29</td>
<td>118.48</td>
<td>60.91</td>
<td>77.00</td>
<td>90.36</td>
<td>54.85</td>
<td>40.45</td>
<td>26.23</td>
<td>67.16</td>
<td>68.23</td>
</tr>
<tr>
<td>U14</td>
<td>10.10</td>
<td>8.11</td>
<td>44.45</td>
<td>35.18</td>
<td>28.45</td>
<td>46.22</td>
<td>95.67</td>
<td>45.63</td>
<td>67.01</td>
<td>106.46</td>
<td>81.92</td>
<td>77.93</td>
<td>33.43</td>
<td>22.03</td>
<td>66.85</td>
</tr>
<tr>
<td>U15</td>
<td>41.89</td>
<td>15.27</td>
<td>34.35</td>
<td>105.24</td>
<td>32.96</td>
<td>62.14</td>
<td>76.91</td>
<td>40.70</td>
<td>41.35</td>
<td>62.19</td>
<td>66.55</td>
<td>49.72</td>
<td>43.68</td>
<td>37.20</td>
<td>39.50</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Statistics

|       | Min  | 10   | 3    | 14   | 35   | 16   | 17   | 35   | 24   | 51   | 29   | 21   | 18   | 22   | 28   | 488  |
|-------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| U01   | 110.36 | 68 | 110.36 | 77 | 82 | 118 | 70 | 90 | 114 | 89 | 78 | 64 | 86 | 68 | 983 |    |
| U02   | 37.24 | 15.43 | 32.95 | 75.67 | 42.73 | 46.40 | 70.99 | 44.89 | 49.54 | 85.03 | 51.23 | 43.03 | 38.63 | 52.22 | 47.11 | 734.08 |
| U04   | 13.19 | 5.82 | 9.90 | 12.53 | 8.74 | 10.25 | 10.68 | 6.39 | 8.34 | 9.46 | 8.18 | 6.70 | 7.55 | 9.17 | 6.74 | 64.81 |

### Figure 4.1: Duration Results Box Plot
to complete the task and others 118 seconds, which is three times longer.

The question that took less time was Q02 with an average of 15.43 seconds, proving that the question was very intuitive for users. To answer this question, the users had only to visualize the value that was already present in a visualization, and consequently, this question is the one with the lowest dispersion, duration time, and standard deviation.

From the results, we can also conclude that the mean time to complete the 15 questions was of 734.08 seconds and using the confidence interval, we are 95% confident that the average time to answer all the questions will be between approximately 11.1 minutes (669.27 seconds) and 13.4 minutes (798.89 seconds). That is less than one minute per question which shows that even for users unfamiliar with the system they can perform a set of questions rapidly and if they become more familiar with EconoVis these numbers could be lowered.

### 4.1.2.B Errors

In addition to the measured time, the number of errors that each user committed in each task was also counted. An error occurred when the user could not reach the result without help, when they missed the answer or when they repeated processes they had already done previously.

The results recorded are present in Table 4.2 as well as in Figure 4.2 where they are in the form of box-plots.

<table>
<thead>
<tr>
<th>Users</th>
<th>Number of Errors</th>
<th>Total Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>U01</td>
<td>0 0 0 0 0 0 0 0 4 2 2 0 1 0 9</td>
<td>Q01 Q02 Q03 Q04 Q05 Q06 Q07 Q08 Q09 Q10 Q11 Q12 Q13 Q14 Q15</td>
</tr>
<tr>
<td>U02</td>
<td>2 0 0 2 0 1 2 0 1 2 2 1 1 0 0 14</td>
<td></td>
</tr>
<tr>
<td>U03</td>
<td>3 1 0 3 0 0 2 1 1 1 1 1 0 1 1 16</td>
<td></td>
</tr>
<tr>
<td>U04</td>
<td>1 0 0 1 0 1 2 0 0 1 0 0 0 0 1 8</td>
<td></td>
</tr>
<tr>
<td>U05</td>
<td>0 0 1 2 0 1 2 2 1 1 0 0 1 0 0 11</td>
<td></td>
</tr>
<tr>
<td>U06</td>
<td>1 0 0 1 0 1 0 2 0 1 0 0 0 0 0 6</td>
<td></td>
</tr>
<tr>
<td>U07</td>
<td>1 0 1 2 0 0 0 0 0 0 0 1 1 0 0 6</td>
<td></td>
</tr>
<tr>
<td>U08</td>
<td>1 0 0 1 0 0 1 0 1 1 0 0 0 0 0 5</td>
<td></td>
</tr>
<tr>
<td>U09</td>
<td>1 0 1 2 2 1 1 0 3 1 0 0 1 1 0 14</td>
<td></td>
</tr>
<tr>
<td>U10</td>
<td>1 1 0 1 1 0 0 1 0 1 0 1 0 0 0 7</td>
<td></td>
</tr>
<tr>
<td>U11</td>
<td>0 0 0 0 0 0 1 0 0 1 1 1 0 0 0 4</td>
<td></td>
</tr>
<tr>
<td>U12</td>
<td>1 0 0 1 0 0 0 0 0 1 1 0 1 1 0 6</td>
<td></td>
</tr>
<tr>
<td>U13</td>
<td>0 1 0 0 0 0 0 1 0 1 1 0 0 0 1 1 6</td>
<td></td>
</tr>
<tr>
<td>U14</td>
<td>0 0 0 1 0 0 0 1 1 1 0 0 0 0 2 6</td>
<td></td>
</tr>
<tr>
<td>U15</td>
<td>0 0 0 2 0 1 0 0 0 1 0 0 0 0 0 4</td>
<td></td>
</tr>
</tbody>
</table>

| Statistics | Min | 0 0 0 0 0 0 0 0 0 0 0 0 0 0 | 4 |
| Max       | 4   | 3 1 1 3 2 1 2 2 3 4 2 2 1 1 2 10 |
| Mean      | 0.80 | 0.20 0.20 1.27 0.20 0.40 0.80 0.47 0.60 1.13 0.53 0.47 0.33 0.40 0.33 8.13 |
| SD        | 0.86 | 0.41 0.41 0.88 0.56 0.51 0.86 0.74 0.83 0.92 0.74 0.64 0.49 0.51 0.62 3.85 |
| CI (95%)  | 0.44 | 0.21 0.21 0.45 0.28 0.26 0.44 0.38 0.42 0.46 0.38 0.32 0.25 0.26 0.31 1.95 |
From the results, we can verify that the questions with a greater average number of errors were questions Q01, Q04, Q07, and Q10, with mean values greater than or equal to 0.8 errors per question. Of the four questions with a higher average number of errors, three of them coincide as was expected with the three questions with a longer average duration as we saw in the last section. The reasons mentioned in the previous point also justify a higher number of errors in these questions.

Regarding Q01 the number of errors can be justified as it was the first question that the users had to answer and although they had tried the system for 5 min previously, this being the first question led to some errors until the users understood how each question would be and how the system worked.

From the data, we can also deduce that on average the number of errors in the 15 questions was less than one error per question and since it is the first use of the system and taking into account the different types of financial concepts mentioned in the questions, the value was not too high.

In the questions with a smaller mean number of errors, Q02, Q03 and Q05, with an average number of errors of 0.2 per question, using the confidence interval, we are 95% confident that the average number of errors in these questions is between 0.00 and 0.48 errors per question that reveals that with more experience the user would eventually have a number of errors close to 0.

4.1.2.C Satisfaction Level

After finishing each question, the users gave their satisfaction rate to each one. The results of the satisfaction of each question can be seen in Table 4.3.

The mean in every question was always higher than 4.3 in 5, so we can conclude that the overall satisfaction from the users with the questions was high.
We also noticed that the three questions with lower satisfaction rating are the Q4, Q7, and Q10, that they are also the ones that the users took more time to obtain the answer and where the mean was higher when talking about the number of errors.

Taking more time to perform a question and also making more errors could lead to a bad experience, and the satisfaction rating could be lower for that reason even so the satisfaction of those questions remained above 4.3 which is a good indicator.

Table 4.3: Satisfaction Level results

<table>
<thead>
<tr>
<th>Users</th>
<th>Q01</th>
<th>Q02</th>
<th>Q03</th>
<th>Q04</th>
<th>Q05</th>
<th>Q06</th>
<th>Q07</th>
<th>Q08</th>
<th>Q09</th>
<th>Q10</th>
<th>Q11</th>
<th>Q12</th>
<th>Q13</th>
<th>Q14</th>
<th>Q15</th>
</tr>
</thead>
<tbody>
<tr>
<td>U01</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>U02</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>U03</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>3</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>U04</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>U05</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>3</td>
<td>5</td>
<td>5</td>
<td>3</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>U06</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>U07</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>U08</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>U09</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>U10</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>U11</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>U12</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>U13</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>U14</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>3</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>U15</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>2</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>

Statistics

| Min   | 5   | 5   | 4   | 5   | 3   | 5   | 3   | 5   | 2   | 4   | 4   | 4   | 2   | 4   |
| Max   | 5   | 5   | 5   | 5   | 5   | 5   | 5   | 5   | 5   | 5   | 5   | 5   | 5   | 5   |
| Mean  | 4.73 | 4.73 | 4.80 | 4.33 | 4.73 | 4.87 | 4.53 | 4.73 | 4.73 | 4.40 | 4.80 | 4.80 | 4.73 | 4.80 |
| SD    | 0.46 | 0.46 | 0.41 | 0.82 | 0.59 | 0.35 | 0.64 | 0.46 | 0.46 | 0.91 | 0.41 | 0.41 | 0.35 | 0.80 |
| CI    | 0.23 | 0.23 | 0.21 | 0.41 | 0.30 | 0.18 | 0.32 | 0.23 | 0.23 | 0.46 | 0.21 | 0.21 | 0.18 | 0.40 |

4.1.2.D System Usability Scale (SUS)

The last stage of the usability tests was the SUS and in Table 4.4 we can analyze the results of each user. As well as calculating for each question the minimal value, the maximum and the mean we also calculated for each user the SUS score following the work of Brooke [34] where he provides scoring guidelines.

From the Score column, we can see that the minimum value was 75 points and the maximum amount was 95 points on a range from 0 to 100 points. To analyze these results, we followed the study of Bangor et al. [35] and using the average of the SUS score we could establish if the usability of our system were considered "Worst Imaginable", "Awful", "Poor", "OK", "Good", "Excellent" or "Best Imaginable".

As our mean score was 85.83 points, and considering the last metrics, we can conclude that our
Table 4.4: System Usability Scale (SUS) results

<table>
<thead>
<tr>
<th>Users</th>
<th>Rating (1 to 5)</th>
<th>Questions</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Q01</td>
<td>Q02</td>
<td>Q03</td>
</tr>
<tr>
<td>U01</td>
<td>5</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>U02</td>
<td>5</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>U03</td>
<td>5</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>U04</td>
<td>5</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>U05</td>
<td>4</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>U06</td>
<td>4</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>U07</td>
<td>4</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>U08</td>
<td>5</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>U09</td>
<td>5</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>U10</td>
<td>5</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>U11</td>
<td>4</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>U12</td>
<td>5</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>U13</td>
<td>5</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>U14</td>
<td>3</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>U15</td>
<td>4</td>
<td>2</td>
<td>5</td>
</tr>
</tbody>
</table>

Statistics

<table>
<thead>
<tr>
<th></th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3</td>
<td>5</td>
<td>4.53</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>5</td>
<td>1.80</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>3</td>
<td>4.60</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>5</td>
<td>1.73</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>5</td>
<td>1.47</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>5</td>
<td>1.47</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>5</td>
<td>4.40</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>5</td>
<td>4.47</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>5</td>
<td>1.80</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>5</td>
<td>85.83</td>
</tr>
</tbody>
</table>

system is considered “Excellent”.

4.1.3 Discussion

From the results of the usability tests, we can conclude that EconoVis had excellent usability results which indicate that the information visualization techniques used allowed an effective and efficient analysis of the data of the financial statements. The users show that the system favors a quick perception of the financial data, that the system is intuitive since the number of errors is low and that the overall satisfaction with each question was high, so the users enjoyed answering the questions using EconoVis.

Regarding the usability of the system, the results of the SUS concluded that EconoVis has excellent usability.

It is important to notice that this was the first time each user interacted with the system and that we have users with various levels of financial experience so the results of the tests can be considered very good and with more experience, each user would improve their numbers and outcomes would be even better.
4.2 Case Studies

The second moment of the evaluation was done through case studies. For this, we counted on the help of a financial expert who used our system and gave us feedback on it.

This type of analysis served to realize the utility of the system and its functionality with someone more experienced. Unlike the usability tests in these tests, we did not measure times or numbers of errors, only recorded the comments that the user said while freely testing EconoVis.

In this type of evaluation, the user does tasks that he would do if he had a system of these available in his day to day work and would say aloud everything he is doing, thinking and his opinion about what is happening in the system.

Through these types of tests, we can perceive what the user thinks about each task he performs and what he finds most relevant and useful in the system.

The user who took this test was Prof. Victor Barros Assistant Professor at ISEG who having experience in the area was the ideal person to do it.

We began by explaining the goal of our study as well as the purpose of the evaluation we were doing.

Then the user started interacting with the system trying to understand the visualizations that were already on the dashboard. He analyzed each parameter and then started adding new as well as new financial parameters.

He began praising the system in general saying "The visualization it's very intuitive, and the idioms work well allowing you to see the variations over the years of various companies quickly."

He also said that the treemap was very useful because it allows you to understand how to divide the parameter "Total Assets" immediately. He said that perhaps it would be better to change the order of the other two parameters that appear because it is currently the American format and that in Europe are usually reversed and would be more natural for a European user.

Regarding the parallel coordinates idiom, he said that it worked quite well with ratios and that it allowed a quick perception of their changes.

The user continued to make several comparisons using the application and was always trying to use several parameters, the use and the search that he was making was becoming faster and more intuitive for him.

Ultimately, he reinforced the idea that EconoVis was very intuitive and that he liked what he saw and used.

4.3 Discussion

After having done the usability tests as well as the case studies we can conclude that our study fulfilled our objective that was to study if applying information visualization techniques allow to speed up the
analysis of the data from the financial statements, exploring the gaps in this area and allowing a smooth, accurate and complete interpretation of that information.

Concerning the usability, the score obtained in the SUS that classified our system as excellent, as well as the results of the other measures collected in the usability tests allow us to consider that our objective regarding usability was accomplished.

Regarding the usefulness of our visualization, the feedback received in the Case Studies also allows us to realize that EconoVis is indeed useful for a financial user. Moreover, during the usability tests, some users said: "At the moment I'm doing a task that was asked, I wish this system were already working to help me in this commitment" which strongly demonstrates the acceptance and usefulness of this study.

With the usability tests, we show that EconoVis allow an effective and efficient analysis of the data of the financial statements; with the Case Studies we show that our visualization is useful for the financial area; so we show that with the information visualizations mechanisms we can, in fact, improve the analysis of the data from the financial statements.
Conclusions and Future Work
There is a large amount of financial data, and with the increased use of the Internet, that number has grown. The existence of different data sources, its large volume, diversity, and complexity make difficult to analyze financial information. This obviously leads to the need to find an approach to allow a quick analysis of that information. That leads to an increase in the use of information visualization with financial data, taking into account its benefits in an area where it is essential to be able to make accurate and quick analysis to gain some benefit. Although there are already several research works in this area related to financial statements, there was a lack regarding this type of data which was not as available as it is in the case of stocks and funds data.

EconoVis is an information visualization system that uses the data from the financial statements of the listed Portuguese companies in the market to provides that information using the information visualization techniques. The users can interact with a dashboard that has multiple visualizations and with that they can quickly analyze the data, compare, and perceive patterns in that information.

Through the development phase, an iterative and incremental approach was followed in order to understand with the help of financial users their needs and adapt the prototype to them. After the requirements analysis and validation, some low fidelity prototypes have been outlined before reach the functional prototype.

EconoVis has been structured following a layered architecture where there exists a separation between the data, the server, and the presentation, and where each of these layers can be used and upgraded independently. The access to EconoVis can be made locally or through the internet as it is also available online.

After the functional prototype had been finished, we started the evaluation of EconoVis. The first phase was to do some Usability Tests with fifteen users from the financial area. The duration that each user took to do each of the fifteen questions was recorded as also the number of errors and the satisfaction grade. From the analysis of the results of the SUS score, we can conclude that our system has excellent usability and with all the results we can conclude that answers all the tasks that we defined. A user case test was also performed with a specialist that confirmed that our research met the goals that we established being intuitive, easy to interpret and that is an advantage for the financial area.

Regarding the future work, we think that adding market features could be something that improves the present system as it would allow the users to have in the same place all the financial statement information and also the market information, as the stocks value and some news. For that, we would add a new section to EconoVis and with a new idiom will show the market value of the companies, in the same section would also add some key news for the selected companies and we could also add some relevant social media information.

Adding some profiling features could also be something interesting for EconoVis, with that a user could define is profile, adapting the system to its necessities and making possible that each time the user
returns it would have all the previous settings saved. We would add a login mechanism to authenticate each user, and after each logout, we would save the actual state of the system in the database. In each login in the EconoVis the settings of each user would be loaded.

Allowing the system to take a snapshot of the system at any given point could also be interesting as the user could come back later to analyze that and also send it to someone in a format that the other person opens in EconoVis and the system shows what it had saved. For that, we could use the database and the profiling feature which we mentioned earlier to keep a record of the system in the point defined by the user. To send it would be necessary to encrypt and convert the database object to a proper format to send and develop a feature to transform that object in the visualizations in the other user environment.

Many features can be added to our study, and is architecture allows that even for a system not related to information visualization some layers of EconoVis could also be reused, extending its usefulness.
Bibliography


Glossary of Accounting Terminology

**Asset** - An economic resource that is expected to be of benefit in the future.

**Balance Sheets** - The basic financial statement, usually accompanied by appropriate disclosures, that describes the basis of accounting used in its preparation and presentation of a specified date the entity's assets, liabilities and the equity of its owners.

**Capital Stock** - Ownership shares of a corporation authorized by its articles of incorporation.

**Cash Equivalents** - Short-term (generally less than three months), highly liquid investments that are convertible to known amounts of cash.

**Dividends** - Distribution of earnings to owners of a corporation in cash, other assets of the corporation, or the corporation's capital stock.

**Equity** - Residual interest in the assets of an entity that remains after deducting its liabilities.

**Financial Statements** - Presentation of financial data including balance sheets, income statements...
and statements of cash flows, or any supporting statement that is intended to communicate an entity’s financial position at a point in time and its results of operations for a period then ended.

**Income Statement** - A summary of the effect of revenues and expenses over a period of time.

**Liability** - Debts or obligations owed by one entity (Debtor) to another entity (Creditor) payable in money, goods, or services.

**Net** - Figure remaining after all relevant deductions have been made from the gross amount.

**Retained Earnings** - Accumulated undistributed earnings of a company retained for future needs or for future distribution to its owners.

**Revenues** - Sales of products, merchandise, and services; and earnings from interest, dividend, rents.

**Statement of Cash Flows** - It categorizes net cash provided or used during a period as operating, investing and financing activities, and reconciles beginning and ending cash and cash equivalents.
Usability Test Survey
Introduction

I am an MSc Finalist in Engineering Systems and Computer Engineering at Técnico Lisboa developing a thesis in collaboration with ISEG - Lisbon School of Economics & Management in which I pretend to use the visualization of information to present the financial data included in the financial statements to be possible to find patterns in the information, to discover new metadata in the data, to have a general understanding of the present information and to easily find data items.

This test aims to evaluate the application and will takes about 25 minutes to complete.

Your answers will be held in strict confidentiality and will be used only for the purposes of this study.

1 About You

a. What is your age range? *

- Under 18
- 18 - 24
- 25 - 34
- 35 - 44
- 45 - 54
- 55 - 64
- 65 - 74
- Over 74

b. What is your gender? *

- Male
- Female

c. What is your relation with Finance/Economics? *

- Finance and/or Economics Professor
- Finance and/or Economics Student
- Finance and/or Economics Graduate
- Investor
- Other
d. How long have you been in the Finance and/or Economics area? *

- Less than 1 month
- 1 - 3 months
- 4 - 6 months
- 7 - 11 months
- 1 - 3 years
- 4 - 6 years
- 6 - 9 years
- More than 9 years

---

2. Try The System

Now you can try the system for about 5 min

a. Are you ready?

- Yes
- No

---

3. Tasks

You will do each task, and after each one, you will have to answer a satisfaction question.

We are testing the system and not the person so don’t be afraid of wrong answers and don’t worry about some errors that may occur.

a. How much of Sonae’s (SON) total liabilities in 2016 was long-term debt?

**Tip:** Balance Sheet Report

b. What was your satisfaction grade with the last task?

- 1
- 2
- 3
- 4
- 5

c. In 2016, what were Sonae’s (SON) total liabilities?

**Tip:** Balance Sheet Report

d. What was your satisfaction grade with the last task?

- 1
- 2
- 3
- 4
- 5
e. In the period 2011-2015, in which year did Sonae’s (SON) increase its Cash and cash equivalents?

Tip: Balance Sheet Report

f. What was your satisfaction grade with the last task?

\[ \begin{array}{ccccc}
1 & 2 & 3 & 4 & 5 \\
\end{array} \]

g. Was Jeronimo Martins’ (JMT) increase in revenues between 2012 and 2015 accompanied by an increase in its net margin %?

Tip: Income Statement Report and Key Ratios

h. What was your satisfaction grade with the last task?

\[ \begin{array}{ccccc}
1 & 2 & 3 & 4 & 5 \\
\end{array} \]

i. Comparing Jeronimo Martins’ (JMT) and Sonae’s (SON) gross margin % and net margin %, which company was more profitable in 2012?

Tip: Key Ratios

j. What was your satisfaction grade with the last task?

\[ \begin{array}{ccccc}
1 & 2 & 3 & 4 & 5 \\
\end{array} \]

k. What were NOS’s diluted earnings per share in the fiscal year 2013?

Tip: Income Statement Report

l. What was your satisfaction grade with the last task?

\[ \begin{array}{ccccc}
1 & 2 & 3 & 4 & 5 \\
\end{array} \]

m. In the period 2010-2014, in which year did NOS experience an increase in its return on equity %?
n. What was your satisfaction grade with the last task?

o. What was the change in CTT's book value per share from 2013 to 2014?

p. What was your satisfaction grade with the last task?

q. How did The Navigator's (NVC) debt-equity ratio change from 2010 to 2012?

r. What was your satisfaction grade with the last task?

s. Between 2014-2016, taking into account the DuPont Identity, which ratio (Net Margin %, Asset Turnover (Average), Financial Leverage (Average)) caused the decline in 2015 on Mota-Engil's (EGL) Return on Equity %?

t. What was your satisfaction grade with the last task?

u. Did GALP's Interest Coverage Ratio ever fall below 4 from 2008 to 2011?
v. What was your satisfaction grade with the last task?

w. From 2012 to 2015, in what year did GALP experience a decrease in its investments in property, plant, and equipment?

Tip: Cash Flow Report

x. What was your satisfaction grade with the last task?

y. From 2012 to 2016 did GALP experience an increase in its basic earnings per share?

Tip: Income Statement Report

z. What was your satisfaction grade with the last task?

aa. Comparing EDP's gross margin %, operating margin %, and net margin % key ratios, in 2012 and 2015, which margins improved every year?

Tip: Key Ratios

ab. What was your satisfaction grade with the last task?

ac. In 2015, comparing GALP and EDP current ratio and quick ratio, which company has higher asset liquidity?

Tip: Key Ratios

ad. What was your satisfaction grade with the last task?
System Usability Scale

How much do you agree with the next questions?

I think that I would like to use this system frequently

1 2 3 4 5

Strongly disagree Strongly agree

I found the system unnecessarily complex

1 2 3 4 5

Strongly disagree Strongly agree

I thought the system was easy to use

1 2 3 4 5

Strongly disagree Strongly agree

I think that I would need the support of a technical person to be able to use this system

1 2 3 4 5

Strongly disagree Strongly agree

I found the various functions in this system were well integrated

1 2 3 4 5

Strongly disagree Strongly agree

I thought there was too much inconsistency in this system

1 2 3 4 5

Strongly disagree Strongly agree

I would imagine that most people would learn to use this system very quickly

1 2 3 4 5
I found the system very cumbersome to use

1 2 3 4 5

I felt very confident using the system

1 2 3 4 5

I needed to learn a lot of things before I could get going with this system

1 2 3 4 5

Thank you for taking the time to complete this test. Your responses will contribute to the success of this project.