EconoVis

Visualizing Economic Trends in a Modern World

Manuel Alves
Department of Computer Engineering
Instituto Superior Técnico
Lisboa, Portugal
manuelalves@tecnico.ulisboa.pt

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

1 INTRODUCTION

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 [4], [14].

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 [8].

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.

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 [20], [21].

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.

2 RELATED WORK

Following the division suggested by Ko et al. [11], 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

To provide research contributions to the design of advanced visual data exploration interface, Merino et al. [13] study different visualizations, which have been employed for stock market data, and apply the results in a new visualization, named *Stock Market Analyzer*.

According to the study's results, the most appropriate techniques for historical data are *Line Charts* and *Recursive Patterns*.

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

The first of the two applications (Figure 1) 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 [22]. Both applications are scalable to handle large quantities of data and have real-time functionality for fast and convenient interactive exploration and analysis.

Figure 1: *Pixel Bar Charts* with time-series data [22]

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 [22].

A further example that uses time series data is the system of Schaefer et al. [15] 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. [22], that shows the global view of the whole market performance.

Schaefer et al. [15] 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 datasets.

Figure 2: *Bloomberg* visualization project [18]

The paper of Sorenson et al. [18] describes one important visualization project (Figure 2) done by *Bloomberg*[1] 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) [18].

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 [18].

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.

[1]Bloomberg - https://www.bloomberg.com/.

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.

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 [11].

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 [6].

Csallner et al. [6] 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 [2], [16].

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 [6].

Dang [7] 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 3.

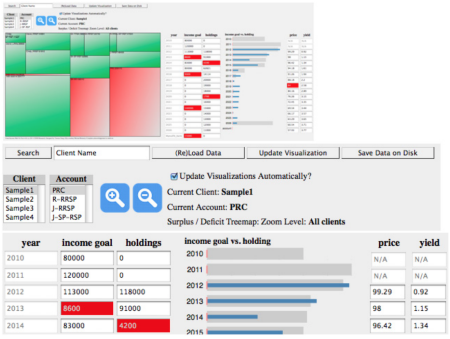

Figure 3: Portfolio monitoring interface [7]

A further example using mutual funds data is the visualization of Alsakran et al. [1] 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 [1].

To demonstrate the benefits of their study, Alsakran et al. [1] 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 [11]. 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. [3]. 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. [3] 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 [3].

In Figure 4 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.

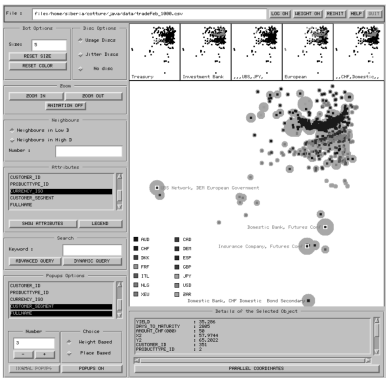

Figure 4: *Domesticating Bead* visualization tool [3]

Chang et al. [5] 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 [5].

The *WireVis* [5] is a multiview approach that assists analysts in exploring large numbers of categorical, time-varying data containing wire transactions [5]. 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 helps discover accounts of similar activities, and the Strings and Beads to depicts transactions over time. The authors also defined four requirements for the system: Interactivity, Filtering, Overview and Detail, and Coordinated Multiple Views.

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 [11], 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. [9] evaluated the performance of self-organizing maps by analyzing the financial performance of 77 pulp and paper companies during 1995-2000.

Another example of visualization using the companies data is the *MarketAnalyzer* [12] 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.

In this study, Ko et al. [12] used an enhanced pixel-based visualization approach, similar to the approach followed by Ziegler et al. [22] 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 [12].

2.5 Discussion

Most of the studies found used stock data and according to Ko et al. [11] this is expected due to the ease of access to this type of data.

Table 1: Related Work Overview

		Comparison	Focus + Detail	Search Mechanisms	Trends Over Time	Avoid Occlusion	Interacton				
							Select	Reconfigure	Encode	Filter	Connect
Stocks Data	*Merino et al.* [13]	✔	✘	✘	✔	✔	✔	✘	✘	✔	✔
	Ziegler et al. [22]	✔	✘	✘	✔	✘	✔	✘	✘	✔	✔
	Schaefer et al. [15]	✔	✔	✘	✔	✘	✔	✘	✘	✔	✘
	Sorenson et al. [18]	✔	✘	✘	✔	✘	✔	✘	✘	✔	✔
	Euronext	✘	✘	✘	✔	✔	✘	✘	✘	✔	✔
	London Stock Exchange	✘	✔	✘	✔	✔	✘	✘	✔	✔	✘
	Deutsche Böerse Group	✔	✘	✔	✔	✔	✘	✘	✔	✔	✔
	New York Stock Exchange (NYSE)	✔	✔	✘	✔	✔	✘	✘	✔	✔	✔
	Nasdaq	✔	✔	✔	✔	✔	✘	✘	✘	✔	✔
Funds Data	*Fundexplorer* [6]	✔	✘	✘	✘	✔	✔	✘	✔	✔	✔
	Dang [7]	✔	✔	✔	✘	✔	✔	✘	✘	✔	✔
	Alsakran et al. [1]	✔	✘	✘	✘	✘	✘	✔	✔	✔	✘
Transactions Data	*Domesticating bead* [3]	✔	✔	✔	✔	✘	✔	✘	✘	✔	✔
	Wirevis [5]	✔	✔	✘	✔	✔	✔	✘	✘	✔	✔
Companies Data	*Eklund et al.* [9]	✔	✘	✔	✔	✘	✔	✘	✘	✘	✘
	Marketanalyzer [12]	✔	✔	✘	✔	✘	✔	✘	✘	✔	✔

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.

In Table 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.

3 VISUALIZING ECONOMIC TRENDS IN A MODERN WORLD

Taking into account the goals defined in the Introduction and in 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 Visualization 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.

3.1 Initial Steps

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[2] 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.

[2]Euronext Lisbon - http://www.bolsadelisboa.com.pt. Last accessed on September 26th, 2018.

3.2 Requirements

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.

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.**
(2) **Identify the value of a financial indicator for a particular company.**
(3) **Compare different financial indicators for a group of companies.**
(4) **Realize the changes in a specific financial indicator throughout years.**

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.

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)**
(5) Did GALP's Interest Coverage Ratio ever fall below 4 from 2008 to 2011? **(4.17)**
(6) How did The Navigator's (NVG) debt-equity ratio change from 2010 to 2012? **(4.17)**
(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)**
(14) In the period 2011-2015, in what year did Sonae's (SON) increase its Cash and cash equivalents? **(3.33)**
(15) From 2012 to 2016 did GALP experiences in some period an increase in its basic earnings per share? **(3.17)**

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.

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 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 [10].

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.

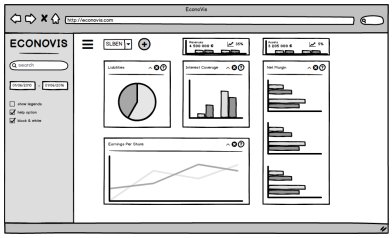

Figure 5: ***LFP Dashboard***

As it can be noticed in Figure 5, 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.

In Figure 6 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.

Some parameters are a combination of other ones, for example, the Total Assets is a sum of the Total Liabilities and the Total

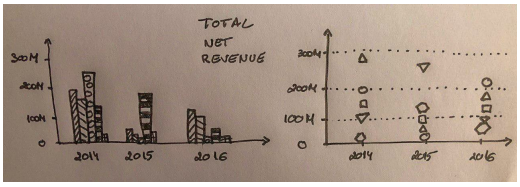

Figure 6: ***Grouped Bar Chart and Scatter Chart options***

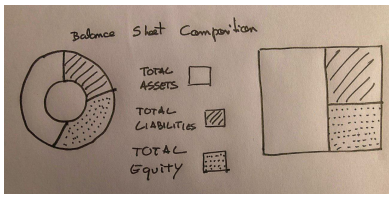

Figure 7: ***Ring Chart and TreeMap Chart options***

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 7 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.

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 6 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.

The next sketches that we examined were the presents in Figure 7 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.

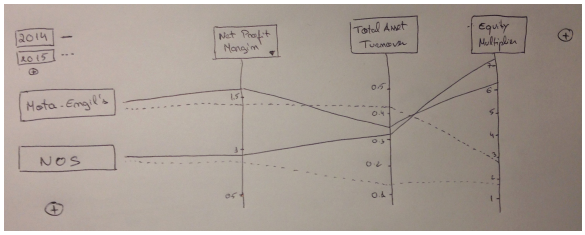

Figure 8: ***Parallel Coordinates Sketch***

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 8. 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 [17]. 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.

The three main layers that compose our system are the Presentation, the Business and the Database Layer, as can be seen in Figure 9.

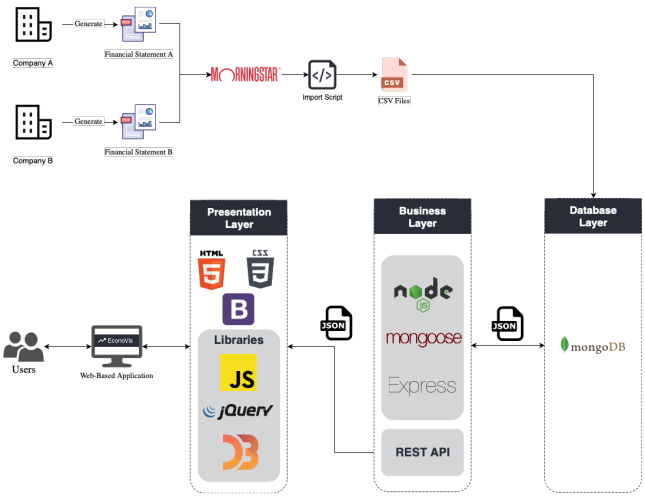

Figure 9: ***EconoVis Architecture***

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.

The Business layer is principally composed for the Node.js server that is an open-source server-side JavaScript environment that supports long-running server processes [19]. 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.

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 Representational State Transfer (REST) Application Programming Interface (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.

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.

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)

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.

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.

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 [3] 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 10). 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.

[3]Select2 - https://select2.org/. Last accessed on September 26th, 2018.

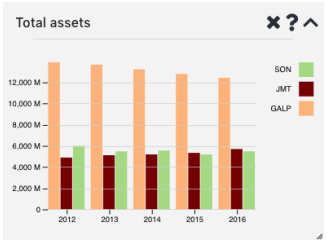

Figure 10: ***Grouped Bar Chart***

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 11). 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.

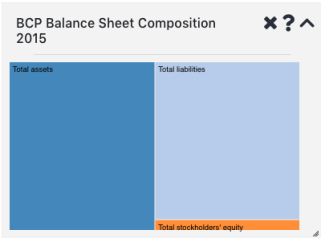

Figure 11: ***TreeMap***

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 12) with a financial expert. With they feedback, it was possible to make a list of changes and improvements for the final solution.

Figure 12: ***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.

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.

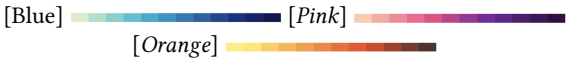

Figure 13: Color Palettes

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 (Figure 13) 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 14), the pink tones for the parallel coordinates (Figure 15), and the orange tones for the TreeMap (Figure 16).

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

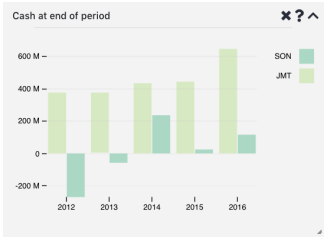

Figure 14: ***Group Bar Chart New Colors***

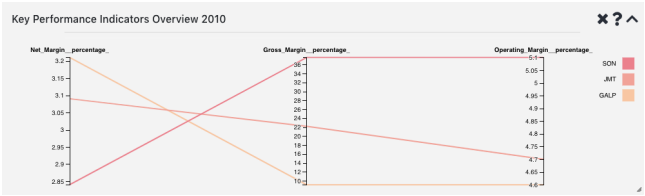

Figure 15: ***Parallel Coordinates Chart New Colors***

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

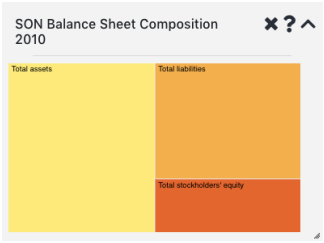

Figure 16: ***TreeMap New Colors***

value, so even with negative numbers the axis will not be on top of the graph, something we wanted to avoid.

To filter by financial reports, which had also been requested, an option was added to the sidebar, 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 17), while in Parallel Coordinates it contains the name of the selected company. These tooltips allow the user to interpret the data and get the information they need more rapidly.

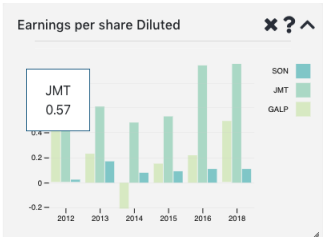

Figure 17: ***Grouped Bar Chart Tooltip***

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 18).

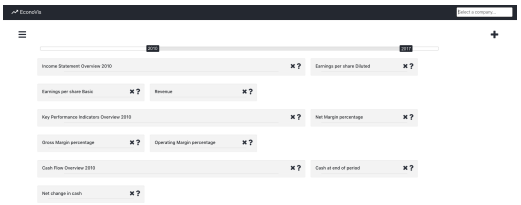

Figure 18: ***Dashboard without companies selected***

In order to filter the view for years, a bar has been added at the top of the dashboard (Figure 19) 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.

Figure 19: ***Years Filter***

We did another validation with an expert, and we realized that we already had a stable and final version (Figure 20) that could answer the questions and tasks that we had defined and that would allow for verification of the objective of our study.

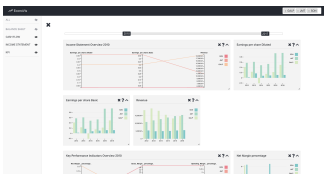

Figure 20: ***EconoVis Final Version***

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

4 EVALUATION

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.

4.1 Usability Tests

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. 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.

From the results of the usability tests, we can conclude that the 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.

Regarding the usability of the system, the results of the System Usability Scale (SUS) concluded that EconoVis has excellent usability.

4.2 Case Studies

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.

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.

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

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.

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.

5 CONCLUSIONS AND FUTURE WORK

There is a large amount of financial data, and with the increased use of the Internet, that number has grown. 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.

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.

After the functional prototype had been finished, we started the evaluation of EconoVis. 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.

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.

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.

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.

REFERENCES

[1] Jamal Alsakran, Ye Zhao, and Xinlei Zhao. 2010. Tile-based parallel coordinates and its application in financial visualization. In *Visualization and Data Analysis 2010*, Vol. 7530. International Society for Optics and Photonics, 753003.

[2] Toshiyuki Asahi, David Turo, and Ben Shneiderman. 1995. Using treemaps to visualize the analytic hierarchy process. *Information Systems Research* 6, 4 (1995), 357–375.

[3] Dominique Brodbeck, Matthew Chalmers, Aran Lunzer, and Pamela Cotture. 1997. Domesticating bead: adapting an information visualization system to a financial institution. In *Information Visualization, 1997. Proceedings., IEEE Symposium on*. IEEE, 73–80.

[4] Stuart K Card, Ben Shneiderman, and Jock D MacKinlay. 1999. *Readings in information visualization: using vision to think*. Morgan Kaufmann.

[5] Remco Chang, Mohammad Ghoniem, Robert Kosara, William Ribarsky, Jing Yang, Evan Suma, Caroline Ziemkiewicz, Daniel Kern, and Agus Sudjianto. 2007. Wirevis: Visualization of categorical, time-varying data from financial transactions. In *Visual Analytics Science and Technology, 2007. VAST 2007. IEEE Symposium on*. IEEE, 155–162.

[6] Christoph Csallner, Marcus Handte, Othmar Lehmann, and John Stasko. 2003. Fundexplorer: Supporting the diversification of mutual fund portfolios using context treemaps. In *Information Visualization, 2003. INFOVIS 2003. IEEE Symposium on*. IEEE, 203–208.

[7] Luan Quang Dang. 2012. *A functional framework for evaluating visualization applications, with a focus on financial analysis problems*. Ph.D. Dissertation. University of British Columbia.

[8] William Dilla, Diane J Janvrin, and Robyn Raschke. 2010. Interactive data visualization: New directions for accounting information systems research. *Journal of Information Systems* 24, 2 (2010), 1–37.

[9] Tomas Eklund, Barbro Back, Hannu Vanharanta, and Ari Visa. 2002. Assessing the feasibility of self organizing maps for data mining financial information. *ECIS 2002 Proceedings* (2002), 140.

[10] Manuel J Fonseca, Pedro Campos, and Daniel Gonçalves. 2012. Introdução ao design de interfaces. *FCA Editora. ISBN* (2012), 978–972.

[11] Sungahn Ko, Isaac Cho, Shehzad Afzal, Calvin Yau, Junghoon Chae, Abish Malik, Kaethe Beck, Yun Jang, William Ribarsky, and David S Ebert. 2016. A survey on visual analysis approaches for financial data. In *Computer Graphics Forum*, Vol. 35. Wiley Online Library, 599–617.

[12] Sungahn Ko, Ross Maciejewski, Yun Jang, and David S Ebert. 2012. Marketanalyzer: an interactive visual analytics system for analyzing competitive advantage using point of sale data. In *Computer Graphics Forum*, Vol. 31. Wiley Online Library, 1245–1254.

[13] Carmen Sanz Merino, Mike Sips, Daniel A Keim, Christian Panse, and Robert Spence. 2006. Task-at-hand interface for change detection in stock market data. In *Proceedings of the working conference on Advanced visual interfaces*. ACM, 420–427.

[14] Tamara Munzner. 2014. *Visualization analysis and design*. AK Peters/CRC Press.

[15] Matthias Schaefer, Franz Wanner, Roman Kahl, Leishi Zhang, Tobias Schreck, and Daniel Keim. 2011. A novel explorative visualization tool for financial time series data analysis. In *VAW2 011: The Third International UKVAC Workshop on Visual Analytics*.

[16] Ben Shneiderman. 1992. Tree visualization with tree-maps: 2-d space-filling approach. *ACM Transactions on graphics (TOG)* 11, 1 (1992), 92–99.

[17] Ian Sommerville et al. 2007. *Software engineering*. Addison-wesley.

[18] Eugene Sorenson and Richard Brath. 2013. Financial visualization case study: Correlating financial timeseries and discrete events to support investment decisions. In *Information Visualisation (IV), 2013 17th International Conference*. IEEE, 232–238.

[19] Stefan Tilkov and Steve Vinoski. 2010. Node. js: Using JavaScript to build high-performance network programs. *IEEE Internet Computing* 14, 6 (2010), 80–83.

[20] Edward R Tufte, Nora Hillman Goeler, and Richard Benson. 1990. *Envisioning information*. Vol. 126. Graphics press Cheshire, CT.

[21] Colin Ware. 2012. *Information visualization: perception for design*. Elsevier.

[22] Hartmut Ziegler, Marco Jenny, Tino Gruse, and Daniel A Keim. 2010. Visual market sector analysis for financial time series data. In *Visual Analytics Science and Technology (VAST), 2010 IEEE Symposium on*. IEEE, 83–90.