10-2020

How to effectively use interactivity to improve visual analysis in groups of novices or experts

Maria do Mar Viana

Pedro Cabral

How to effectively use interactivity to improve visual analysis in groups of novices or experts

Maria do Mar Viana, m20180646@novaims.unl.pt

Pedro Cabral, pcabral@novaims.unl.pt

NOVA Information Management School (NOVA IMS), Universidade Nova de Lisboa, Campus de Campolide, Lisbon, 1070-312, Portugal

Abstract

This study contributes with a methodology to evaluate the differences between different groups of individuals, using a prototype with different interactive visualizations, created with the Shiny package, using 6 quantitative and qualitative metrics for the validation. Using an ANOVA single factor test only 1 of the 6 variables showed statistically significant differences between both groups: the engagement. This means that this is the only metric where results can be improved in order to close the gap between the group of experts and novices. The heatmap and the bar chart were considered the best visualizations for both groups, and the worst were the choropleth map and the stacked bar chart. Regarding the interactive component, the select box was a better option for the group of novices and the radio box for the group of experts. Using this study, organizations will be able to create visualizations that are suitable for different audiences.

Keywords: Data visualization; Interactivity; Experts; Novices; Shiny

1. Introduction

Data is beginning to be more and more accessible to the public and is, therefore, important to realize what exactly is the best way for not only the experts but also the general audience to understand a visualization (Blascheck et al., 2019). Data visualization is important in a broad set of fields, for instance, organizations that work in education (Mottus, Kinshuk, Graf & Chen, 2013) or in finance (Zhang, 2015). Therefore, it could be interesting to analyse how different types of individuals discover interactivity and which are the best visualization techniques for each group, since organizations can create visualizations that will be exposed to different types of audience.

Currently, there is a gap in how general individuals understand interactive functionality (Blascheck et al., 2019). Although there are some studies that tried to find differences between how different types of individuals react to interactivity (Blascheck et al, 2019; Oghbaie, Pennock & Rouse, 2016; Rouse, Pennock, Oghbaie & Liu, 2017), none had the specific goal of testing which were the best visualization idioms for each type of group.

This study investigates the differences on how two different types of individuals discover interactivity, through the creation of a prototype with different interactive visualizations and using an ANOVA single factor test to find statistically significant differences between both groups. The prototype will be developed using the Shiny package in the R programming language, with different interactive visualizations techniques to understand which are the most effective visualizations for different types of individuals.

Results will be useful for organizations that want to better understand how to use interactivity and which visualizations and interaction techniques can be used to build interactive data visualizations which convey the information in the most effective manner to different groups of individuals.

Regarding the structure of this article in the first section, an introduction for the project will be done, explaining the background and the project goals.

In the second section a theoretical background research will be done. In the third section, all the methodologies and datasets used to develop the prototype will be explained. In this section, it will also be explained the process of how the prototype was evaluated.

In the fourth section, the discussion of the results will be made and finally, in the last section, the conclusions reached from this study will be explained, as well as the limitations of the project and the future work other researchers may do, applying the knowledge of this study.

2. RELATED WORK

This chapter provides with an overview of what has already been done regarding topics related to this project and is divided into four sections: discover of interactivity in different groups of subjects, suggested interactivity, tools to develop interactive visualizations and validation of interactive visualizations. In each section, it will be discussed the different approaches that have been used in past studies and the corresponding strengths and weaknesses of each method.

2.1. Discovery of interactivity in different groups of subjects

Previous studies have tried to discover how different groups of subjects discover interactivity. However, no differences between these groups have been found (Blascheck et al., 2019). Although it was not possible to find differences between each group of individuals, it was possible to identify different exploration strategies that users may apply when interacting with interactive visualizations. Using the different exploration strategies, it was possible to create six suggestions to improve the functionality of interactive visualizations in different groups of individuals (Blascheck et al., 2019): inviting interaction, combating oscillation, leveraging spatial organization, providing entry points, scaffolding complex interactions and supporting transitions.

Other works have obtained different conclusions regarding the differences in how distinct groups of individuals interact with visualizations, such as the work conducted by Oghbaie, Pennock and Rouse (2016) who used two different types of individuals, experts, and non-experts. In the work conducted by Oghbaie et al. (2016), it was suggested that if the proper visualization methods are applied the gap between experts and non-experts could be closed (**Table 1)**. In a similar work conducted by Rouse, Pennock, Oghbaie and Liu (2017), using the Rasmussen's abstraction-aggregation hierarchy methodology it was possible to reach similar conclusions.

METRICS / AUTHOR	**OGHBAIE ET AL. (2016)**	**ROUSE ET AL. (2017)**
Accuracy	Experts only exceed non-experts in the data with the most complex casual relationships	Experts were more accurate than non-experts
Speed	Same speed for both groups of experts and non-experts	
Data	Experts made full use of all the information displayed whereas non-experts demonstrated a lack of information seeking behavior	

Table 1 - Differences and similarities between the different groups of individuals

2.2. Suggested Interactivity

Understanding how to use interactivity in visualizations is crucial to develop techniques that increase the intuitiveness of online visualizations. Suggested Interactivity (SI) can be defined as: "a set of methods for indicating that a graphical area can be interacted with by subtly directing a user's attention so as not to impede too heavily on this person's focus or on the rest of the interface design" (Boy, Eveillard, Detienne, & Fekete, 2016). According to these authors there are three types of SI cues that can be applied in visualizations:

- SI cues that are present in the object of interest, for example in the visualization itself;
- SI cues that are present in external objects, for example in widgets;
- SI cues that use a mix of the first two SI cues and are therefore present in the object of interest and in external objects. These SI cues can also provide feedforward.

2.3. Tools to create interactive visualizations

There are several tools that can be used to create interactive visualizations: ready-to-use tools and tools that require programming skills. One example of these ready-to-use tools is Power BI, which is a business analytics service provided by Microsoft (Microsoft Power BI, 2020). Power BI offers a set of tools to help the users with the manipulation, analysis and visualization of data, which in turn, makes it possible to create visualizations. Another ready-to-use tool is Tableau Desktop, which can also be utilized to create interactive visualizations since it is a data visualization tool that helps the discovery of valuable insights from the data available at a very high speed, and potentials the creation of interactive dashboards. Bhardwaj and Baliyan (2019) mentioned that a few good reasons to use Tableau to create a visualization tool are:

- It has a smaller learning curve than some programming languages like R or Python, and requires less technical expertise;
- It is not as expensive as other tools such as OBE by Oracle or Business objects by SAP.

Other tools require programming skills to create visualizations, such as Python, which is a general-purpose programming language that can also be applied in data visualization. Python provides different libraries for data visualization (Fahad & Yahya, 2018).

R is a programming language and software, which also has some of the same libraries as Python and also has the possibility to use the Shiny package to create interactive visualizations. Shiny is an R package that enables the creation of interactive visualizations that could be displayed online (Ellis & Merdian, 2015). The work conducted by López et al. (2018) tried to understand the best way to use this tool and developed several suggestions that should be applied when creating a Shiny application:

- The application should be developed taking in mind its main goals in order to prioritize the most important information;
- The application should be clear and intuitive and use different visualization and interaction methods;
- The participants that use the application should have an active attitude when interacting with the visualization.

2.4. Validation of interactive visualizations

Regarding the validation of interactive visualizations, it can be interesting to analyse the framework created by Munzner (2015) to design visualizations. That framework consists of four nested levels:

- Domain situation: in this top-level it is necessary to understand which target users and which requirements are necessary for the visualization;
- Data/Task abstraction: in this level, the crucial goal is to discover which data will be used in the visualization;
- Visual encoding/Interaction idiom: in this level, the objective is to realize what types of visual encodings or interactions will be used to display the data in the visualization;
- Algorithm: finally, at the bottom level is necessary to realize if the computer and the code created are effective at displaying the visualization.

An efficient way to validate an interactive visualization can be through the use of quantitative measures, for example questionnaires or surveys and/or through the use of qualitative measures such as interviews (Lu et al., 2018). Both of these measures can be tools to use in the levels of framework created by Munzner (2015).

This strategy was applied in the work conducted by Lu et al. (2018), where an experiment was performed to test an interactive visualization. In the experiment, the participants had time to first get acquainted with the interactive visualization and afterward they were asked to finish a set of tasks, to determine if the visualization was effective to gather the information needed.

Additionally, there are other works that also discuss the use of quantitative and qualitative measures to measure the efficacy of a visual analytic tool in order to improve data comprehension. An example of this is the work conducted by Géryk (2015), that carried out an experiment, where the participants had to complete a set of tasks, with two datasets of different sizes and an interactive visualization.

Through this work, it was possible to discover that when using these quantitative and qualitative measures the results of an experiment can be divided into three sections: accuracy, completion time and the subject's preferences.

3. DATA AND METHODS

In this section, it will be discussed the data sources utilized, the technology used to build the prototype, the justification of the choices applied in the prototype and the validation process.

3.1. Data

The dataset used to create the prototype is public, and available at the Banco de Portugal website (Banco de Portugal, 2019). Even though the website of Banco de Portugal possessed data on several other different topics, considering that the inspiration for the prototype was to analyze the Portuguese economy and its external imbalances, the dataset used only had data regarding the balance of payments (Banco de Portugal, 2019). The dataset had data from 1996 to 2018 and the current and capital account statistics measure was used, reflecting, monthly, net lending/net borrowing of the Portuguese economy vis-à-vis the rest of the world.

An additional dataset from Eurostat was used to create an international comparison of the Portuguese economy with all the countries in the European Union (Eurostat, 2018). This dataset contained the current account balance using the percentage of gross domestic product (a three-year average) for each country in the European Union from 2007 to 2018.

3.2. Methods

The prototype built in this study was developed using the R software and the Shiny package. Regarding the R software in general, one reason to use this type of software is the fact that R is a software that exceeds other tools when it comes to flexibility (Jiang & Carter, 2018). Besides this, the fact that the software is free also proved to be a greater advantage, in comparison to other tools that are used to create interactive visualizations, such as Tableau, since even though this tool is not as expensive as other tools it still has a cost (Bhardwaj & Baliyan, 2019). Another interesting fact is that the R ecosystem offers a wide range of packages that can be used in data visualization.

The fact that R offers the possibility to use the Shiny package was also one of the key reasons to use this type of software. One of the motivations to use this package was the possibility to easily integrate interactivity in the applications created. Besides that, the fact that it makes it possible to host these applications online was also a benefit since it would be feasible, if necessary, to test the prototype online (Ellis & Merdian, 2015).

3.3. Types of visualizations

The prototype consists of a foreword view and five different interactive visualizations all regarding the topic of Balance of Payments, in order to fulfill one of the suggestions when creating a Shiny application, which is to use different visualization methods/idioms (López et al., 2018). Five different types of visualizations were developed:

- In the first visualization (By major items) it was chosen a stacked bar chart to encode the data, since that in the sample of the dataset used for this view, there is one quantitative attribute and two categorical key attributes. There is also a line chart encoded on top of the stacked bar chart;
- In the second visualization (By geographical counterpart), considering that the dataset used had a geographical breakdown, it was possible to create a spatial analysis. In this case, it was used a choropleth map, since the data had one quantitative attribute;
- In the third visualization (By monthly periodicity), it was chosen to apply a heatmap visualization, to compare the differences of the values of the item selected in the different months/years. It is possible to use this type of visualization since the data used had one quantitative attribute and two categorical attributes;
- In the fourth visualization (By type of services), it was chosen to do a scatterplot chart since the dataset had two quantitative value attributes;
- In the fifth and final visualization (By international comparison), the type of idiom used was a bar chart, considering that the data used had one quantitative value attribute. This bar chart was also ordered from the highest value of the attribute to the lowest value, and this order was applied in every year displayed in the visualization.

3.4. Interactive components

López et al. (2018) mentioned that one of the suggestions to apply when creating a Shiny application is to use different interactive methods, and as mentioned previously, considering that this prototype was built using the Shiny package, different interactive methods were applied in each of the visualizations that compose the prototype.

The users had three different widgets to choose from:

- The slider widget was used to help users perform a time analysis on the visualizations and in almost all of them it provides the users with a play button, so it is possible to show an animation of the changes in the data throughout the years. The play button also serves the purpose to invite the users to interact with the visualization, since inviting interaction is one of the suggestions in the work done by Blascheck et al. (2019);
- The select box widget was used to provide the users with more options for variables to display in the visualizations. For example, in the second visualization (By geographical counterpart), using the select box, the users can filter the data by choosing which variable they want to see displayed in the choropleth map;
- The radio buttons widget was used to provide the users with different combinations of variables they may want to display in the visualizations. For example, in the fifth visualization (By international comparison), users can compare the Portuguese economy with different combinations of countries.

The widgets used SI cues applied to an external object, in this case, the widget itself, since the users have to interact with the widget in order to see changes in the visualization but to use the widgets they don't have to use the actual graphs (Boy et al., 2016).

In the main panel, the interaction choice the users had were the tooltips, which display visualization values, by moving the mouse to different sections of the visualization itself. The tooltips used SI cues applied to the object of interest, in this case, the actual graphs, since to interact with the tooltips the users have to click on the graph itself (Boy et al., 2016).

In addition, in the main panel of each visualization, there is always a small text to give a little introduction to the visualization in order to fulfill one of the suggestions in the work done by Blascheck et al. (2019) which is scaffolding complex interactions. There are also suggested actions for the user in the text, which can be seen as SI cues that use feedforward (Boy et al., 2016).

3.5. Experiment procedure

Succeeding the development of the prototype, the validation phase was then implemented. The experiment was in most cases an in-person experience and was conducted with a sample of individuals, from groups of experts and novices. The group of experts are individuals that have knowledge about the topic in the interactive visualizations, in this case, the balance of payments. This type of individuals was recruited by contacting individuals that work in Banco de Portugal, professors and individuals that study or work in the economy field or other similar areas. The group of novices are individuals that have no prior knowledge about the topic of the balance of payments. This group was recruited by contacting students from universities and individuals that worked or studied in different areas, that could participate in the study. It was also ensured that both groups had individuals from different ages from 18 to over 55 and individuals from different genders.

To start the experiment, the sample group, which consisted of a minimum of 30 participants, 15 for each group, had access to the prototype and was able to interact with each one of the interactive visualizations. After a small period for testing had passed, the participants had to follow a specific set of tasks and record their answers in a questionnaire (Géryk, 2015) to fully understand if they were able to use the prototype to the maximum of its potential.

Most of the experiments were in-person, and therefore almost all the participants were observed while performing the experiment. This forced the participants to have an active attitude when interacting with the prototype, fulfilling one of the suggestions to apply when creating a Shiny application (López et al., 2018).

In this part of the experiment, it was possible to collect the quantitative metrics (Lu et al., 2018), such as the accuracy (Zhu, 2007), by comparing the correct answers with the answers of each participant and the completion time (Zhu, 2007), by measuring the time it took each participant to finish the tasks given in each of the interactive visualizations and the total time it took to finish this part of the experiment.

In the next stage of the experiment, the participants had to answer a 5-point Likert scale questionnaire, related to their opinion of the prototype and were able to suggest improvements. In this part of the experiment, each section of this questionnaire measured the qualitative metrics, therefore, the level of usefulness, efficacy, complexity (Zhu, 2007), and engagement (Lu et al., 2018) that the participants classified each one of the visualizations. The participants also classified the level of usefulness and complexity of the components of each visualization. In the final part of the questionnaire, the participants were also asked to rank each one the visualizations that compose the prototype.

In the analysis of the results, the mean values of each one of the metrics (Géryk, 2015) obtained in both questionnaires were analysed to develop the conclusions of this study. An ANOVA single factor test was also conducted to test for statistically significant differences between both groups.

4. RESULTS AND DISCUSSION

4.1. Prototype

The first result of this study was the prototype, that allowed the implementation of the experiments of this project and the results that therefore followed. The visual representation of the prototype can be seen in the **Figures 1** to **6**.

Figure 1 - Prototype "Foreword" tab

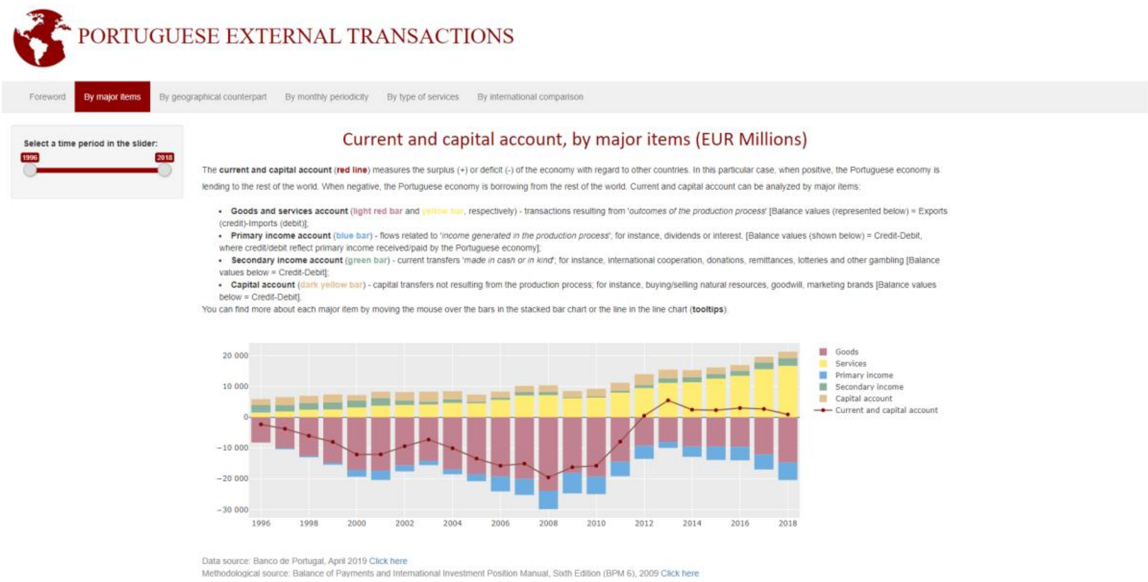

Figure 2 - Prototype "By major items" tab

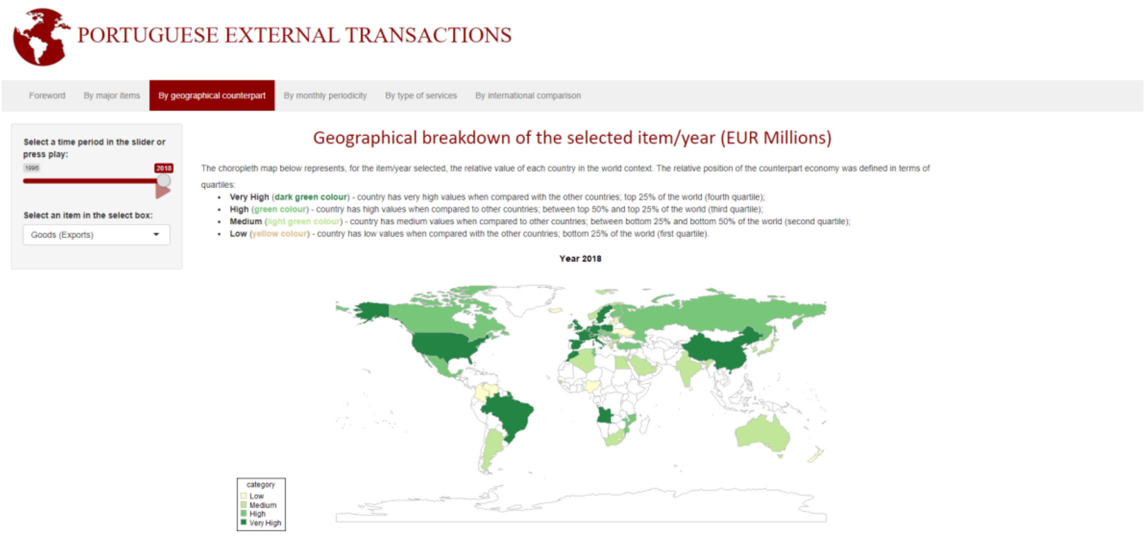

Figure 3 - Prototype "By geographical counterpart" tab

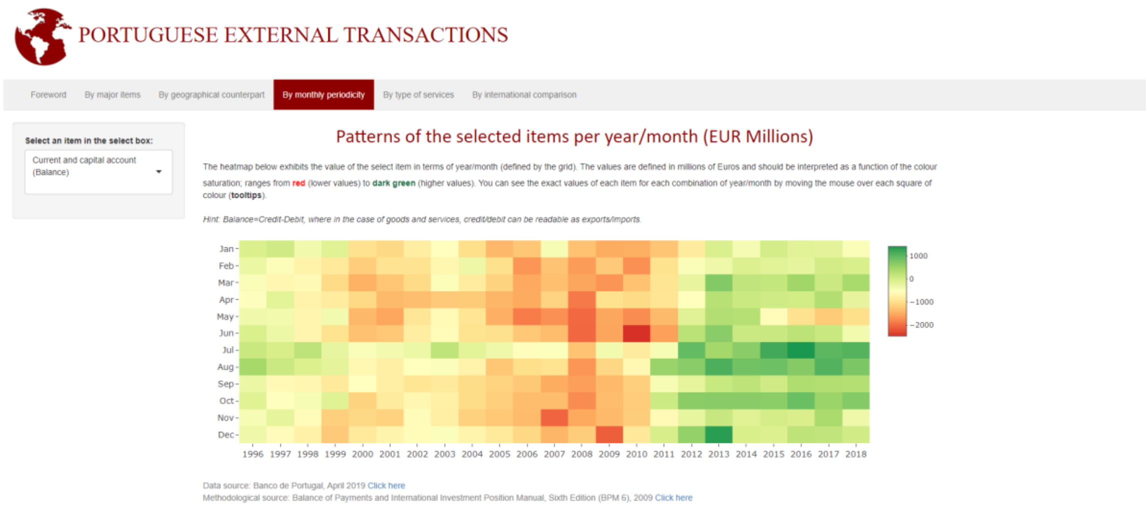

Figure 4 - Prototype "By monthly periodicity" tab

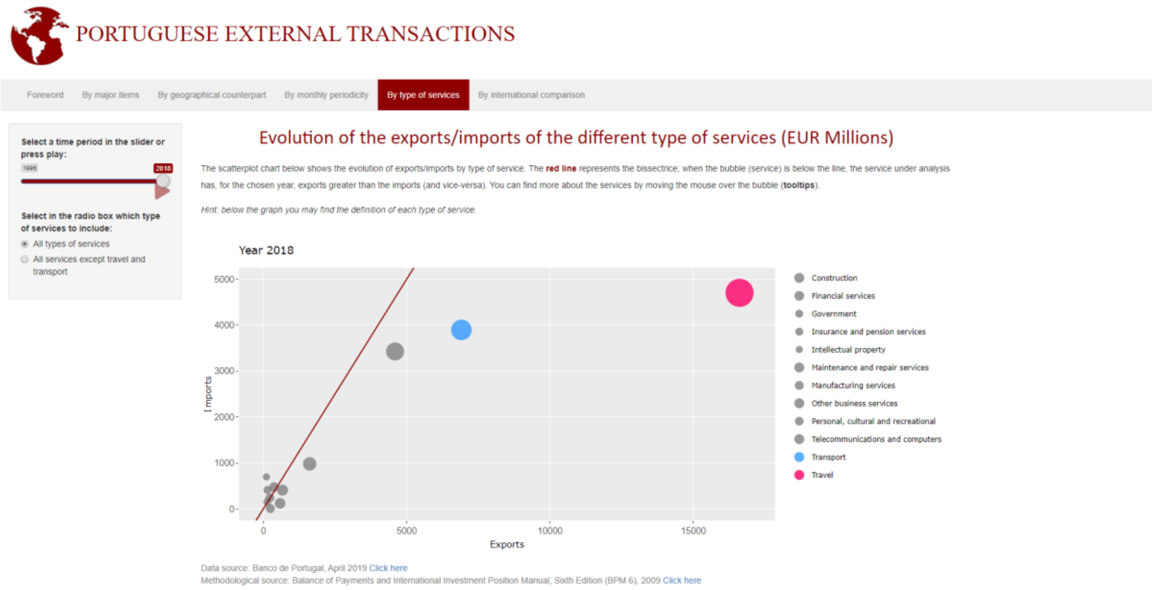

Figure 5 - Prototype "By types of services" tab

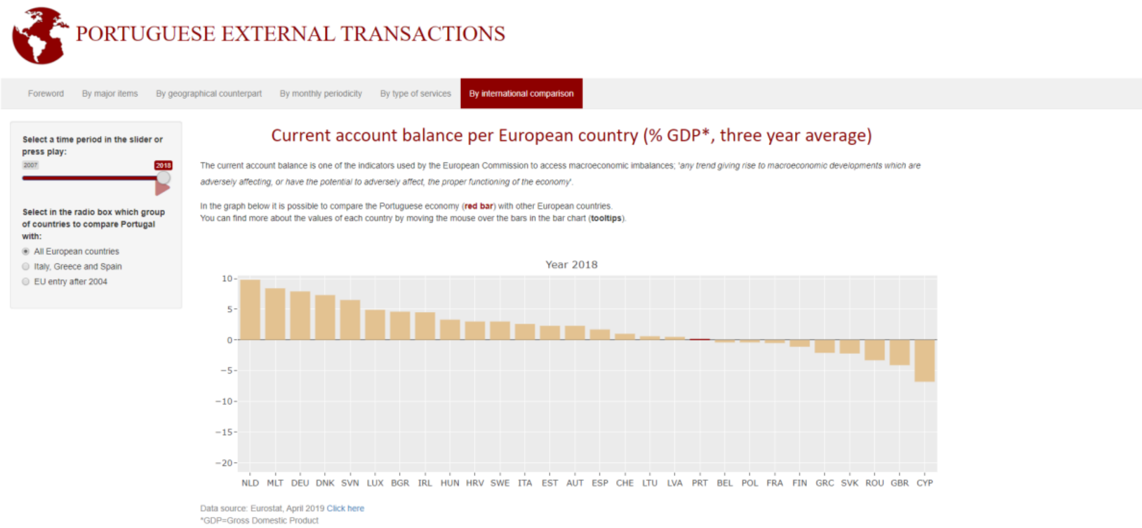

Figure 6 - Prototype "By international comparison" tab

The source code repository for the original prototype can be accessed at the following link:

https://github.com/m20180646/MasterWorkProject

4.2. Metrics analysed

30 individuals participated in this experiment, 15 in the group of experts and 15 in the group of novices. The metrics obtained were quantitative (accuracy and the completion time) and qualitative (usefulness, efficacy, engagement and complexity) (**Table 2**). Regarding the accuracy, this metric was divided into four intervals, very low (0-25%), low (25%-50%), moderate (50-75%) and high (75%-100%). As for the qualitative metrics, they were measured using a 5-point Likert scale and they were also classified by dividing the results into four intervals, very low (1-1.99), low (2-2.99), moderate (3-3.99) and high (4-5) .

To test for significant effects, an ANOVA single factor test was conducted regarding the quantitative metrics. Regarding the accuracy metric, no statistically significant differences were found ($F(1,8) = 0.122605$, $p = 0.735266$). The same results were obtained regarding the completion time metric, ($F(1,8) = 0.358601$, $p = 0.565851$). This means that although there seemed to be differences between both groups, these differences are not statistically significant, regarding the quantitative metrics.

Regarding the qualitative metrics an ANOVA single factor test was also conducted. For the usefulness metric the results did not have differences that were statistically significant ($F(1,8) = 2.119055$, $p = 0.183568$). The same applied for the usefulness metric regarding the components ($F(1,8) = 1.717874$, $p = 0.22634$). For the complexity metric the same results were also obtained, where no statistically significant differences were found, neither for the complexity in general ($F(1,8) = 0.0975$, $p = 0.762835$), or for the complexity of the components ($F(1,8) = 0.001779$, $p = 0.967388$).

Finally, regarding the efficacy metric, again no statistically significant differences ($F(1,8) = 0.333662$, $p = 0.579401$) were found between both groups. However, as for the engagement metric, this was the only metric where statistically significant differences ($F(1,8) = 10.58511$, $p = 0.01164$) were found, considering the p value was under 0.05. Although it appeared that there were differences between groups on all qualitative metrics, the engagement was the only metric where these differences were statistically significant. This means that this is the only metric where results can be improved in order to close the gap between the group of experts and novices, and perhaps if only the best visualizations for each group were used, this gap could have been closed.

METRICS	NOVICES	EXPERTS	ANOVA RESULTS
Accuracy	91%	93%	(F(1,8) = 0.122605, p = 0.735266)
Time	21 minutes and 57 seconds	24 minutes and 1 second	(F(1,8) = 0.358601, p = 0.565851)
Usefulness	4.86	4.75	(F(1,8) = 2.119055, p = 0.183568)
Usefulness of the components	4.78	4.71	(F(1,8) = 1.717874, p = 0.22634)
Efficacy	4.84	4.81	(F(1,8) = 0.333662, p = 0.579401)
Engagement	4.83	4.44	(F(1,8) = 10.58511, p = 0.01164)
Complexity	1.87	1.96	(F(1,8) = 0.0975, p = 0.762835)
Complexity of the components	1.37	1.38	(F(1,8) = 0.001779, p = 0.967388)

Table 2 – Results of the metrics analysed for the different groups

4.3. Visualizations

Regarding the best visualization for each group (**Table 3**), it is possible to conclude that for the novice's group the best visualizations were the third (**Figure 4),** where the idiom was an heatmap, which was the visualization that ranked in first place, had the highest values of accuracy and also had one of the lowest completion times, and the fifth visualization (**Figure 6**), where the idiom was a bar chart, since it was the visualization with the lowest completion time, the highest average of usefulness, the highest average of efficacy and the lowest average of complexity either in general or regarding the components.

As for the expert's group, the best visualization was also the third visualization (**Figure 4**), for the same reason as the novice's group and since it had the highest average of usefulness and the lowest completion time. The fifth visualization (**Figure 6**), can also be considered one of the best visualizations since it had the best averages in two qualitative metrics, the complexity in general and of the components, and it had some of the highest averages of accuracy and some of the lowest averages of completion times, meaning that there were no differences between the two groups in which were the best visualizations.

MEASURE ANALYSED	VISUALIZATION WITH THE BEST VALUES IN THE GROUP OF EXPERTS	VISUALIZATION WITH THE BEST VALUES IN THE GROUP OF NOVICES	VISUALIZATION WITH THE WORST VALUES IN THE GROUP OF EXPERTS	VISUALIZATION WITH THE WORST VALUES IN THE GROUP OF NOVICES
Accuracy (Highest better)	Heatmap (3rd visualization) and Bar Chart (5th visualization)	Heatmap (3rd visualization)	Stacked bar chart (1st visualization)	Choropleth map (2nd visualization)
Completion time (Lowest better)	Heatmap (3rd visualization)	Bar Chart (5th visualization)	Choropleth map (2nd visualization)	Stacked bar chart (1st visualization)
Complexity (Lowest better)	Bar Chart (5th visualization)	Bar Chart (5th visualization)	Scatterplot chart (4th visualization)	Stacked bar chart (1st visualization)
Complexity of components (Lowest better)	Bar Chart (5th visualization)	Bar Chart (5th visualization)	Stacked bar chart (1st visualization)	Stacked bar chart (1st visualization) and Heatmap (3rd visualization)
Engagement (Highest better)	Choropleth map (2nd visualization)	Scatterplot chart (4th visualization)	Stacked bar chart (1st visualization) and Bar Chart (5th visualization)	Stacked bar chart (1st visualization) and Heatmap (3rd visualization)
Usefulness (Highest better)	Heatmap (3rd visualization)	Bar Chart (5th visualization)	Bar Chart (5th visualization)	Stacked bar chart (1st visualization)
Usefulness of components (Highest better)	Stacked bar chart (1st visualization)	Choropleth map (2nd visualization)	Heatmap (3rd visualization)	Stacked bar chart (1st visualization)
Efficacy (Highest better)	Scatterplot chart (4th visualization)	Bar Chart (5th visualization)	All the same besides the Scatterplot chart (4th visualization)	Stacked bar chart (1st visualization)
Average rankings (Highest better)	Heatmap (3rd visualization)	Heatmap (3rd visualization)	Choropleth map (2nd visualization)	Choropleth map (2nd visualization)

Table 3 – Best and worst visualizations for each metric for the different groups

As for the worst visualizations for each group (**Table 3**), the conclusions were that for the novice's group it was the second visualization (**Figure 3**), where the idiom was a choropleth map, considering that it was the visualization that ranked in last and had the lowest accuracy average, and the first visualization (**Figure 2**), where the idiom was an stacked bar chart, considering it had the highest completion time, the lowest averages of usefulness in general and of the components, efficacy and engagement and also had one of the highest averages of complexity in general and of the components.

Regarding the experts group the worst visualization was also the second visualization (**Figure 3**), for the same reason as the novice's, since it was also the visualization that ranked in last and since it was one of the visualizations that had one of the lowest average of usefulness, lowest average of efficacy and the highest completion time. It is also possible to consider that the first visualization (**Figure 2**) is also one of the worst for the expert's group, considering it had the lowest averages of engagement and the highest averages of complexity regarding the components, and also the lowest accuracy average. This means, that similar to the novice's group the worst visualizations were also the second and the first visualizations, meaning that no differences were found between the two groups regarding this aspect and that both groups had difficulties with this type of idioms, stacked bar chart and choropleth map.

4.4. Components

Regarding the components (**Table 4**), it is possible to conclude that as for the interactive features, besides the actual visualizations, for the novice's group, the radio box and the text can be one of the most complex components, whereas for the expert's group is the select box, meaning that for the expert's group the radio box may be a more suitable component and for the novice's group it might be the select box. The text, depends on the use, considering it is for both groups sometimes one of the most complex components and one of the least complex. The slider and the tooltips are considered in both groups to be one of the least complex components, which means that both groups are comfortable interacting with these features, and they might be adequate for both groups.

As for the most useful components, the most useful component for the novice's group is the slider and for the expert's group is the radio box, the slider and the tooltips. The least useful component, was in both groups and in all of the visualizations the text, meaning that the participants of both groups found that the visualizations and their components were more useful, likely due to the fact that they were interactive and the text was the only component that wasn't, concluding that both groups prefer features that are interactive and associate interactivity with usefulness.

COMPONENTS	NOVICES	EXPERTS
Most Complex	Radio box and text	Select box
Least Complex	Slider and tooltips	Slider and tooltips
Most Useful	Slider	Radio box, the slider and the tooltips
Least Useful	Text	Text

Table 4 – Complexity and usefulness of the components for the different groups

4.5. Additional findings

Another additional finding that was found in this study was that it was proved that the Shiny package is a tool that is capable of creating interactive visualizations for different types of individuals, since both groups obtained high averages of accuracy in all of the visualizations and in general, and all of the average values of usefulness, efficacy and engagement of all the visualizations were ranked as high. Moreover, the average values of complexity of the prototype in general ranked as very low, meaning the visualizations and the prototype in general were easy to analyse.

4.6. Comparison of the results with the existing literature

Comparing this results with the existing literature, using the example in the study conducted by Blascheck et al. (2019), it is possible to see a major difference between this study and the project in question, considering that, although this study also used different types of individuals in their experiences, it did not find any differences between those groups, which in turn, is not true for this project, since there were differences identified in the way both groups perceived the different components/interactive features and regarding the qualitative measure engagement.

Besides it can also differ from the work conducted by Blascheck et al. (2019) since it provided different suggestions for companies and organizations to apply, where the work conducted by Blascheck et al. (2019) developed several exploration strategies and suggestions for the users to apply in the visualizations and this work provided which were the best and worst visualizations users can utilize and the best types of interactive components for each type of individuals.

Finally, when comparing the results of this study, with the results of the study conducted by Oghbaie et al. (2016) and Rouse et al. (2017) it is possible to observe that there are some differences and similarities. Regarding the completion time, this work concluded that the speed was the same for both groups, as it was in the case of both studies being compared. Regarding the accuracy, the results were different than the results in the work conducted by Rouse et al. (2017), since the experts group did not have any statistically significant differences than the novice's group regarding the accuracy measure, and they were also different than the results in the work conducted by Oghbaie et al. (2016), that concluded that the experts only exceed non-experts in the data with the most complex casual relationships.

5. CONCLUSIONS

Through the ANOVA single factor test, we found that the only metric that had statistically significant differences between both groups was the engagement metric. Regarding the visualizations, both groups agreed that the best visualizations were the heatmap and the bar chart and the worst visualizations were the choropleth map and the stacked bar chart. As for the components, the select box was a better option for the novice's group, while the radio box was the best for the expert's group, and the tooltips and the slider are adequate for both types of individuals. With this study, we conclude that, although there are some similarities in how the different types of individuals perceive the interactive visualizations, there are also differences between the two groups and so it is possible for companies and organizations to use these suggestions and adapt their visualizations for the different types of individuals in order to create visualizations that are effective for different types of audience. We also found that the Shiny package is a powerful tool that makes possible the creation of effective interactive visualizations that are suitable for different types of individuals, which can be extremely useful and may provide them with a competitive advantage.

Some limitations of this study include, for example, the fact the data could have been richer if it was used a sample of a bigger size, or for example if more groups were created in this study. Also, the fact that the groups were only separated by novices and experts can be a limitation considering the knowledge they had on the topic of the visualization was the only distinguishable variable between both groups, and no other characteristics were taken into consideration, such as age or the knowledge they had on data visualization. Another limitation could be the fact that the prototype only used five different types of idioms, but considering the time used for the experiences, it would not be feasible to have more visualizations to test.

As for future work, it would be interesting to try this experience with other different groups, for example following the work conducted by Blascheck et al. (2019), using three different types of individuals taking in consideration whether or not they have knowledge on the subject of data visualization and not only on the topic of the visualization. It is also recommended to recreate the same prototype with different tools to discover which could be more efficient and compare them, as well as using different visualization idioms in the prototype. Another suggestion is to use a larger sample size to understand if the same results would be obtained or if they could change. Lastly, it would be useful to continue to understand why the engagement metric was the only metric with statistically significant differences between both groups and to try to discover methods to decrease this difference and completely close the gap between experts and novices.

REFERENCES

Banco de Portugal. (2019). BP Stat Estatisticas Online. Retrieved from Banco de Portugal Eurosistema: https://www.bportugal.pt/PAS/sem/src/(S(gg1qej45tcbdlvzut2sbrf3n))/selecAnalise.aspx?Token=B542AB22-FA09-4822-AE4C-CFEF9937B23

Bhardwaj, P., & Baliyan, N. (2019). Hadoop based Analysis and Visualization of Diabetes Data through Tableau. 2019 Twelfth International Conference on Contemporary Computing (IC3), (pp.1–5). Noida, India

Blascheck, T., Vermeulen, L. M., Vermeulen, J., Perin, C., Willett, W., Ertl, T., & Carpendale, S. (2019). Exploration strategies for discovery of interactivity in visualizations. IEEE Transactions on Visualization and Computer Graphics, 25(2) (pp. 1407–1420).

Boy, J., Eveillard, L., Detienne, F., & Fekete, J. D. (2016). Suggested Interactivity: Seeking Perceived Affordances for Information Visualization. IEEE Transactions on Visualization and Computer Graphics, 22(1) (pp. 639–648). https://doi.org/10.1109/TVCG.2015.2467201

Ellis, D. A., & Merdian, H. L. (2015). Thinking outside the box: Developing dynamic data visualizations for psychology with Shiny. Frontiers in Psychology, 6 (pp. 1–6).

Eurostat. (2018). The Macroeconomic Imbalance Procedure (MIP) introduced. Retrieved from Eurostat Statistics Explained: https://ec.europa.eu/eurostat/statistics-explained/index.php/The_Macroeconomic_Imbalance_Procedure_%28MIP%29_introduced

Fahad, A., & Yahya, A. E. (2018). Big Data Visualization: Allotting by R and Python with GUI Tools. 2018 International Conference on Smart Computing and Electronic Enterprise (ICSCEE), (pp. 1–8). Kuala Lumpur, Malaysia.

Géryk, J. (2015). Using Visual Analytics Tool for Improving Data Comprehension. Proceedings of the 8th International Conference on Educational Data Mining (pp. 327–334). Madrid, Spain.

Jiang, Z., & Carter, R. (2018). Visualizing library data interactively: two demonstrations using R language. Library Hi Tech News, 35(5) (pp. 14–17).

López, M., Cobo, E., & Cortés, J. (2018). Assessing Shiny apps through student feedback : Recommendations from a qualitative study. Computer Applications in Engineering Education, 26(5), (pp. 1061–3773). https://doi.org/10.1002/cae.21932

Lu, F., Yu, D., Liang, H. N., Chen, W., Papangelis, K., & Ali, N. M. (2018). Evaluating Engagement Level and Analytical Support of Interactive Visualizations in Virtual Reality Environments. Proceedings of the 2018 IEEE International Symposium on Mixed and Augmented Reality, ISMAR, (pp. 143–152). Munich, Germany.

Microsoft Power BI. (2020). Turn data into opportunity. Retrieved from Microsoft: https://powerbi.microsoft.com/en-us/

Mottus, A., Kinshuk, Graf, S., & Chen, N. S. (2013). Visualization and interactivity in the teacher decision support system. Proceedings - 2013 IEEE 13th International Conference on Advanced Learning Technologies, ICALT 2013 (pp. 502–503). Beijing.

Munzner, T. (2015). Visualization Analysis and Design. CRC Press.

Oghbaie, M., Pennock, M. J., & Rouse, W. B. (2016). Understanding the Efficacy of Interactive Visualization for Decision Making for Complex Systems. 2016 Annual IEEE Systems Conference (SysCon) (pp. 1–6). Orlando, FL.

Rouse, W. B., Pennock, M. J., Oghbaie, M., & Liu, C. (2017). Interactive visualizations for decision support: Application of Rasmussen' s abstraction-aggregation hierarchy. Applied Ergonomics, 59(Part B) (pp. 541–553). https://doi.org/10.1016/j.apergo.2016.03.006

Zhang, C. (2015). Using Excel's Data Table and Chart Tools Effectively in Finance Courses. Journal of Accounting and Finance, 15(7), (pp. 79–93).

Zhu Y. (2007) Measuring Effective Data Visualization. Advances in Visual Computing. ISVC 2007. Lecture Notes in Computer Science, 4842. Springer, Berlin, Heidelberg