World Happiness Report: An Interactive Visualization

Nychol Bazurto, Raul A. Gutierrez, Carlos F. Torres, John A. Guerra, Tiberio Hernandez, Mario Chamorro and Claire Bulger
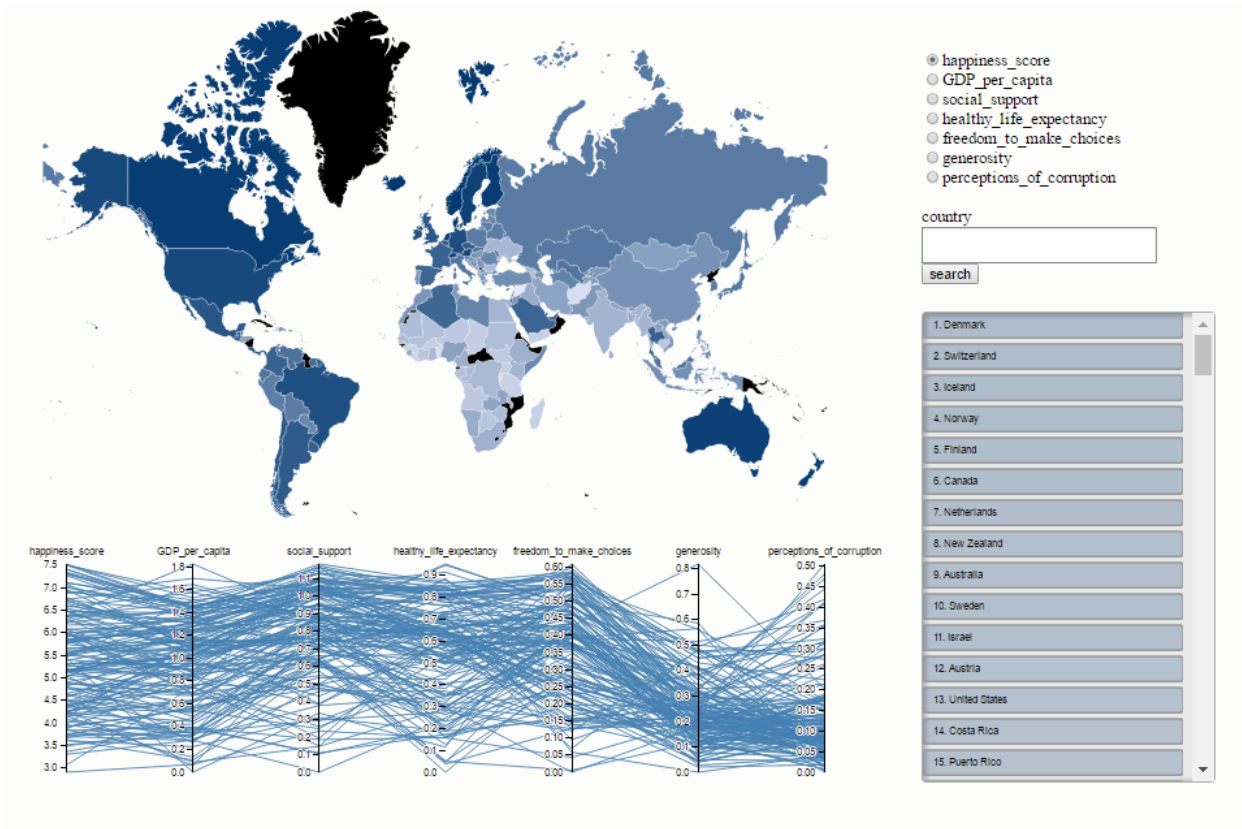
Fig. 1. Initial view of visualization of this project. Authors

Abstract—The world happiness report is a survey conducted in almost all countries of the world, measuring the happiness of people living in those countries. The survey is multitudinous and collects thousands and millions of records which become indexes and scores that finally translate into the measurement of the happiness of people. The visualization and understanding of the final results of this report are overwhelming, and several attempts have been made to realize visualizations that fulfill the objective of being intuitive, understandable and interactive, but have not achieved it. This project intends to carry out a new approach in visualizing the results of the report, in order to raise awareness of these results and that ordinary people, people who do not have knowledge of the report or knowledge in visualization are able to find information interesting and answer questions. Several visualization idioms such as choropleth map, bar chart and parallel coordinates, related in an interactive coordinate view, were used with a specific search filter, obtaining satisfactory results for the work team and for the client as well.

Index Terms—World Happiness Report, happiness, visualization, coordinated views

◆

1 INTRODUCTION

- *Nychol Bazurto is a student at the University of the Andes. E-mail: n.bazurto@uniandes.edu.co.*
- *Raul A. Gutierrez is a student at the University of the Andes. E-mail: ra.gutierrez@uniandes.edu.co.*
- *Carlos F. Torres is a student at the University of the Andes. E-mail: cf.torres@uniandes.edu.co.*
- *John A Guerra is a student at the University of the Andes. E-mail: ja.guerrag@uniandes.edu.co.*
- *Tiberio Hernandez is a student at the University of the Andes. E-mail: jhernand@uniandes.edu.co.*
- *Mario Chamorro worked with the World Happiness Report team. E-mail: marioachamorro@gmail.com.*
- *Claire Bulger worked with the World Happiness Report team. E-mail: clairebulger@gmail.com.*

The world happiness report is a survey in almost all the countries of the world, which measures the happiness of people in that country, this score is divided into several indexes that explain it, such as corruption, or life expectancy. The project consists in making a visualization of the data of the world happiness report, understandable and interactive, striking for the users, who are common people who want to explore the report, either for entertainment or for research.

In the past years, the people in charge of the report had made

attempts to show the results, but they did not succeeded since they had a lack of good visualization practices and the idioms used were not attractive to the user nor intuitive. The approaches that we had developed with the stakeholders in this new version of the visualization were very fruitful. The result is a new visualization of the results for report that fulfill the demands of the clients, as well as the target users.

To tackle this project, we carried out a previous research which gave us several ideas which finally inspired us to specify the final solution proposed. As previously mentioned, good information visualization practices were very important at the moment of proposing a new solution to this problem, so we were guided by the reference "Visual Analysis Best Practices" [6] that Tableau in conjunction with its collaborators has written, where they describe each of the idioms that are used for each type of data or dataset. In order to make a coherent and well-made proposal, we also had to see the previous versions of the report [2], and their respective visualizations. With the purpose of not repeating the mistakes made previously. Michelle A. Borkin presents us in her article "What Makes a Visualization Memorable?" [3] as well as Drew Skau, Lane Harrison and Robert Kosara in An Evaluation of the Impact of Visual Embellishments in Bar Charts [1] some good practices on different idioms, which are related to the problem and the proposed solution. These good practices also expose the general perception of a visualization, not only from the technical or objective.

Looking beyond the good practices, and approaching a little more the problem as such, it was found in the previous research that the treatment of the geographic data was very related to the work of the Linkoping University where Mikael Jern, Monica Brezzi and Lars Thygesen worked in a Project called A web-enabled Geovisual Analytics tool applied to OECD Regional Data [4]. This work was the first source of inspiration with which we started to work, since the data handled and the problem raised was very similar, without becoming the same, the World Happiness Report had other considerations. In order to approach this project we used the framework proposed by Tamara Munzner in her book [5], which is composed by What, Why and How and which will be mentioned and used in the characterization section.

Finally, the evaluation method that was used in the project was controlled experiments, since we worked with people without some knowledge in visualization, where the population of the sample was reduced. We measure quantitative and qualitative variables.

This paper is organized as follows: Section 2 describes the previous research done by the group, and exposes the similarities and relevant information with the problem of the project. Section 3 describes the characterization of the data, the What and the Why. Section 4 describes the proposed method for solving the problem, section 5 shows the results and the method of evaluation that we follow to validate the effectiveness of the solution. And finally section 6, conclusions and discussion.

2 STATE OF THE ART

This section tries to build an structured classification of the idioms or visualization used to build a solution to the client of the project [6]. Taking in mind that the main task is to make easier to understand the results of the World Happiness Report [2], we initially explored solutions to similar problems as geographic data visualization, and with tool that we thought fit for a solution.

First of all, it is important to describe the whole screen of visualizations where each chart will be placed. In the Whitepaper visual analysis guidebook [6] they suggest a configuration called Dashboards. Basically the dashboard is the main screen where all the detailed visualizations are placed. It gave us the first view of the problem. The main idea of having aggregated visualizations or a dashboard is to present completely all the information to the user, and to be interactive, and coordinated between each visualization
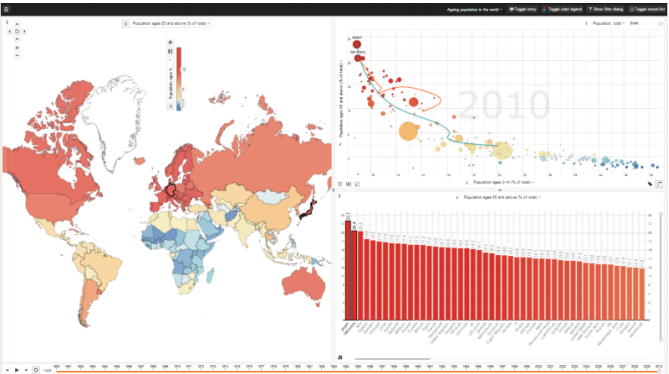

Fig. 2. Coordinated view (OECD Regional eXplorer-Collaboration NCVA and OECD, Linkoping University, Sweden).
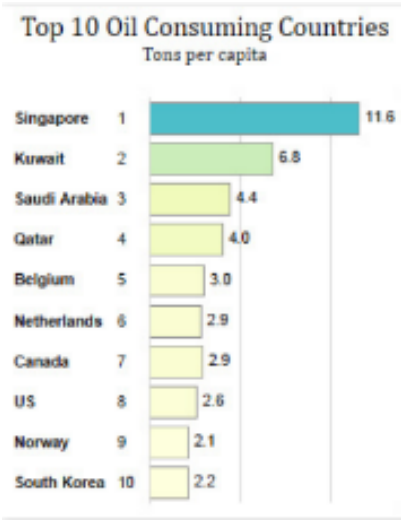

Fig. 3. Example of ranking comparison as a bar chart.

[6]. Dashboards increase the analytical power of the visualization by showing multiple perspectives in the same location. They can also be used to combine multiple types of data in a single location.

To address the problem of displaying geographic data, the Linkoping University created an interesting solution using coordinated views or dashboard, as we mentioned earlier [4], where they showed the geographic data of the world using a choropleth map together with idioms such as bar charts and scatter plots, and showed regional and countries data such as the ratio of the population or the average age, among many others. This was the first approach that inspired us the solution for our project since we have geographic data that can be related by coordinated views.

Taking this first approximation to dashboard as the parent of the hierarchy, now it will be disaggregated into the different visualizations that compose it.

The first visualization analysed is the ranking of countries, it compare and rank countries, regions, based on one or a set of criteria. A bar chart is great for comparison and ranking because it encodes quantitative values as length on the same baseline, making it extremely easy to compare values [6]. The list of country names in Fig 2 are also a visualization of the data, it is a different kind of view and it fulfill the main task that consist of getting the max and min of the countries. there are several consideration about representing a ranking as a list or as a bar chart: one of them is the accuracy of the comparison.

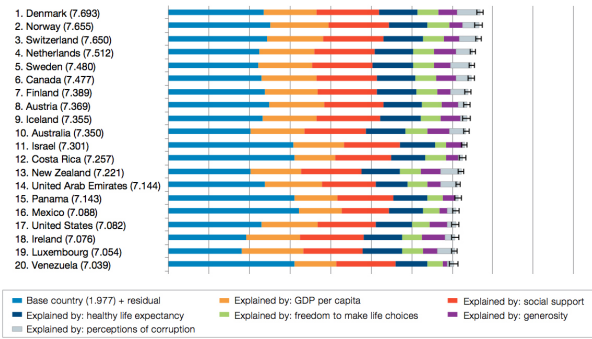**Figure 1.0 - Ranking of Happiness (top 20 in 2010-12)**

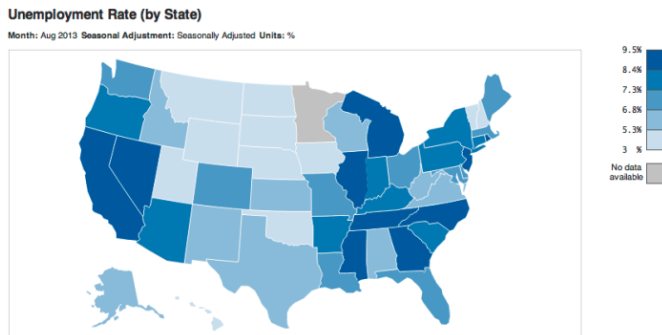Fig. 4. Main visualization already made for the world happiness report.

Fig. 5. Choropleth map of United States of America.

According to Robert K. most of the embellishments have an adverse effect on the accuracy of reading values when comparing bars. [1] The tests they run showed not terribly surprising results, but it is now backed by actual research: most of the embellishments have an adverse effect on the accuracy of reading values when comparing bars. [1]

The older report of happiness was taken into account, because they have done a lot of failed attempts for getting a good visualization. In the project it was taken all that visualizations in a sort of guide for not repeating the history anymore [6]. In the report we can see how there was an effort to represent the results in a way that there was as much data as possible, but did not take into account the tasks they wanted to perform or the user experience. In Fig 3 it is clearly that it is not interactive, and the comparison between indices is impossible, it is simply possible to compare the total, which makes an ordered list as well.

When it is necessary to show a location, the most common solution is to use a map, but maps are often best when paired with another chart that details what the map displays, such as a bar chart sorted from greatest to least, a line charts showing the trends, or even just a crosstab to show the actual data. [6]

It's important to take into account that the visualization have different perceptions depending on the user watching it. [3]There is a basic generic chart ranking but more specifically in the best visualization know, there are different preferences depending on the area or depending on the theme or the interest of the user. For example a bar chart is not always good for all the users. Bar charts are not good enough for the scientific community users and the infographics [3] but for new media and government visualization is the best for users and it is well received. The classification we have found for each of the elements, as well as the whole, is as follows in the Fig 5:
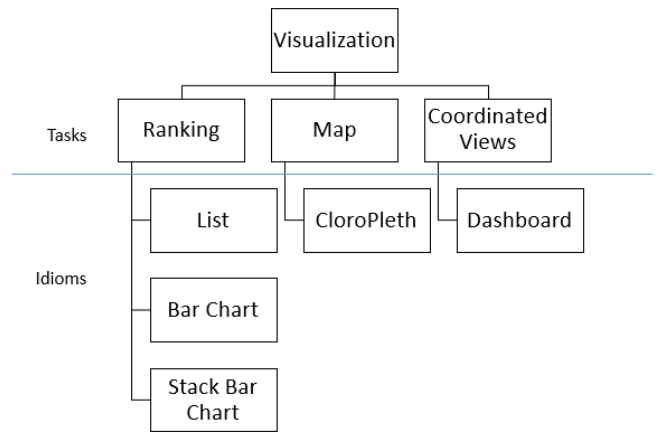

Fig. 6. Final classification of possible solutions for the visualization problem.

3 CHARACTERIZATION

We are going to do the characterization based on Tamara Munzner framework which is composed for What, Why and How. This part just describes What (data abstraction) and Why (task abstraction). [5].

3.1 What

World Happiness Report data has 6 main indexes to define its happiness score ranking by country, these are:

- **Healthy life expectancy:** Calculated based on data from the World Health Organization (WHO), the World Development Indicators (WDI), and statistics published in journal articles.

- **Social support:** Its makes reference to have someone who helps you in hard times, like your family, and its define by an average of answers to this dichotomous question If you were in trouble, do you have relatives or friends you can count on to help you whenever you need them, or not?

- **Freedom to make life choices:** National average related to the answer to Are you satisfied or dissatisfied with your freedom to choose what you do with your life?.

- **Perceptions of corruption:** Average of answers to this two question: Is corruption widespread throughout the government or not and Is corruption widespread within businesses or not?

- **GDP per capita:** Its the purchasing power parity (PPP) value of all final goods and services produced within a country in a given year, divided by the average (or mid-year) population for the same year. For the report was adjusted to constant 2011 international dollars.

- **Generosity:** Its the residual of regressing national average of response to the GWP question Have you donated money to a charity in the past month? on GDP per capita.

All these, with the happiness score are ordered quantitative sequential attributes.

As categorical attribute has countries which can be consider a geometric dataset too because have an implicit geometry and the quantitative attributes are associated with each one (table about each country).

3.2 Why

The main target of World Happiness Report is shows the ranking, that everyone can understand it. Anyone must to be able to understand and interact with the indexes which are associated, in order to get interest information. With this and the data in mind five main task were identified to achieve:

1. **Present** the ranking, countries and their respective score (**feature**).

2. **Discover** happiness **distribution** in the world. The user can interact with each country to obtain details on demand.

3. **Identify** the happiness score and indexes by country (**features**).

4. **Locate** (knowing the country that I want to find (e.g. my country) a country and query how happy is it and how indexes are in this. (**features**).

5. **Identify** which countries are happier and which ones less happy (**extremes**).

Tasks are ordered by priority.

4 PROPOSED METHOD

Continue with the Tamaras framework once what and why were established its possible think in the visual encoding of these task, its mean find the best options (channels and marks) for our tasks and data. So, considering the most important tasks and the state of art, we decided to use coordinate views (juxtaposed), which support the data and let the user discover interesting relationship between the different indexes. This proposal is composed for 4 main idioms:

1. A choropleth map: To support the first task. This idiom preserves something familiar to the user and reuses their common knowledge of the world, in order to achieve everyone can understand easily the main idea. For this idiom its necessary choose the right channel, so, saturation its perfect to express the happiness score (something quantitative and sequential).

   - Manipulate: In this case, map allows select and navigate. The idea of this design was let the user picks up on the map any country and apply a geometric zoom of that country.

   - Reduce: The proposed tooltip associated to this idiom, provides to the user a embedded chart, with information about the ranking in the last years in a particular country (with mouse over action).

To express the indexes data, we decided use two different idioms, which are affected by independent filters: parallel coordinates to show all indexes at the same time and bar chart per index.

2. Parallel coordinates: Considering indexes have a different scale between them, it was necessary find a suitable idiom to show all together and avoid breaks expressiveness, selecting parallel coordinates. This idiom encodes expressing in different axis, and if there are interactions like selecting and highlight, linked with the map, makes possible establish relation between the general score in a country and its indexes. To avoid occlusion between different countries, saturation could be use.

   - Manipulate:User can select and highlight one line which corresponds to a particular country.

   - Reduce: Filter items is an main interaction to offer to user not only the information of a country but a set of countries with similar values in certain indexes.

3. Bar chart: To show and compare one index, visualizing the highest, lowest countries and the selected country by the user (is a link between map and bars,described in the table X.)

   - Manipulate: Select and highlight country respect selection on the map.

   - Reduce: Filter items, showing just a few countries per index.

4. Ranking list: Just the map isnt enough to express all the data, e.g , ranking needs to be express it with a better encoding, more directly:ordering, so a list with the values it's a simple alternative (spatially distributed text).

   - Manipulate: Change order, from lowest to highest position or vice versa.

   - Reduce: By a filter associated reduces shown items to the user.

In general, all the views are going to be linked with shared data, with the interactions shown in Table 1, between them.

Idioms	Map	Parallel Coordinates	Bar chart
Map	X	A country selected highlights a line.	Selected country bar is highlights and shows the neighbour countries bars in respect of that index.
Ranking list	Searched country in the filter list, highlights on the map that country.	Searched country in the filter list, highlights the parallel coordinate	X

Table 1. Interactions between juxtaposed views.

In the Fig 7 , a mockup about the proposal is showed.

(a) General view with the happiness score filter.
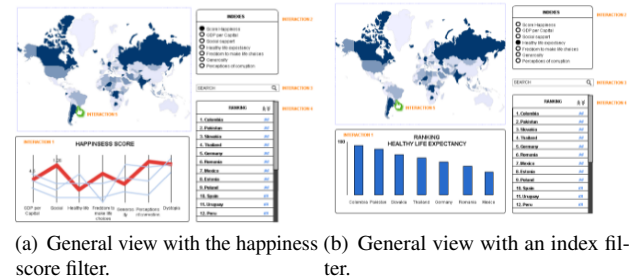
(b) General view with an index filter.

Fig. 7. Mockup of proposed method.Authors.

5 RESULTS AND EVALUATION

Previous proposal was implemented using a server in Amazon, this solution is web and the main technologies used were: D3, jQuery, Js, Topojson and HTML. In the next section could be aprecciated the final visualization.

5.1 Results

The general view with happiness score filter selected is shown in the Fig 8. As seen, design was keeped. When the user mouse hover over a country, throws an embedded chart, like in the Fig 9.

If the user picks up a country, indexes information of that country is highlighted on the parallel coordinates, located below the map. In the Fig 10, this behavior is appreciated. Additionally, user can play with the parallel coordinates by his own account, using a area selectors in each axis, as seen in Fig 11.

By other hand, when the user selects an index filter, parallel coordinates is changed by bars chart which reflects index distribution in the world with the most representative countries (Fig 12). When user
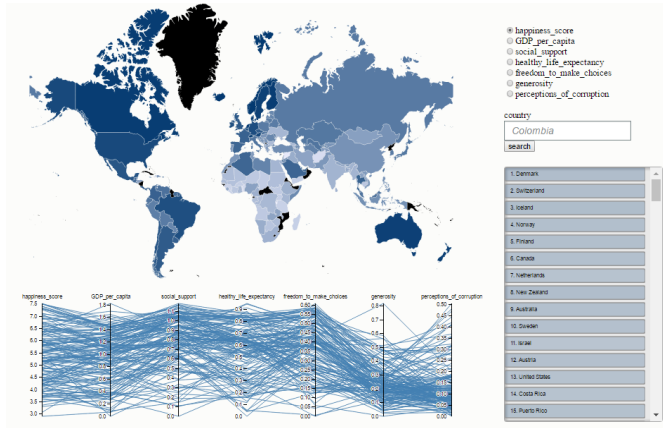
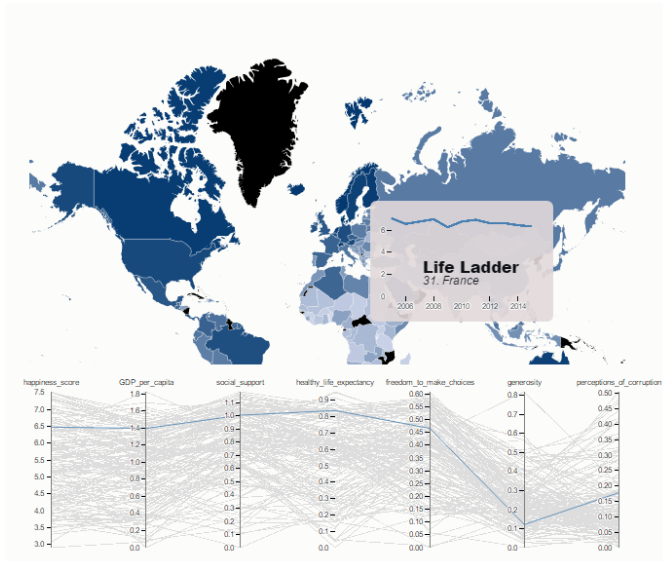Fig. 8. Initial view for the user, happiness score filter applied. Authors.
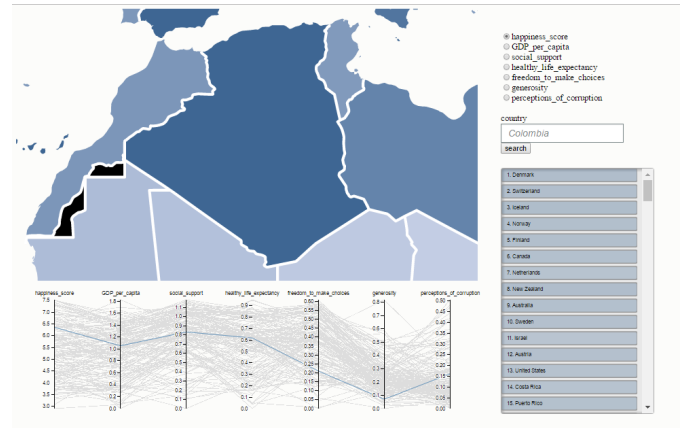

Fig. 10. Visualization when an user selects a country and the happiness score filter is applied. Authors.

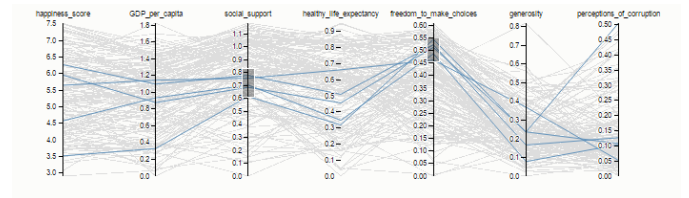Fig. 11. Possible filters in parallel coordinates view. Authors.

Fig. 9. Mouse-over on a country. Authors.

obtain certain tasks, whether it took long or short time. Tasks such as obtaining the happiest country in the world, obtaining the score of a specific country, or finding the statistics of a particular country comparing it with others.

Finally, we ask a question about what else could be intuited, obtaining answers such as: Some country has a ranking score that probably could be affected by its freedom of choice or vote.

6 DISCUSSION AND CONCLUSION

Authors (or clients) expect visualization to improve the use and understanding of the results of the World Happiness Report.

It will enable the regular users to take a more active role in the discovery process of exploring world, regional and country indicators, for example, to identify those countries with higher corruption or better life expectancy.

selects a particular country its bar is highlighted and show its heighbors in that aspect (Fig 13).

To help the user locanting a country, lateral filter was implemented to highlight on the map and filter ranking list (Fig 14).

5.2 Evaluation

The method to perform the evaluation of the visualization developed for the solution of the problem was considered taking into account the target population to which it was directed. The target population is people who do not have any knowledge in visualization and do not know good practices of data visualization, people who are of different levels of education but that their common interest is to know and to explore the results of the world happiness report. A sample of the population with these characteristics was chosen to make an initial control group.

Controlled experiments were the method chosen to make the respective evaluation, with a controlled group of people, different variables were evaluated in a controlled environment. Initially interactivity was measured and how intuitive it was for the user. These measurements were performed qualitatively, evidencing by observation whether the subjects could perform their tasks successfully or were frustrated by not being able to develop its task. Likewise we quantitatively evidence some other variables, such as the ability to
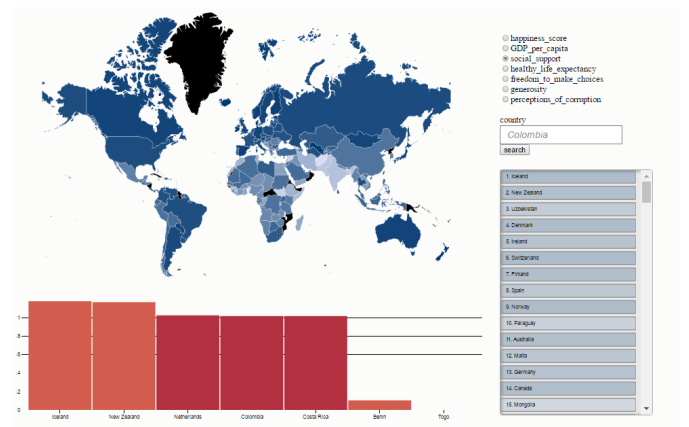

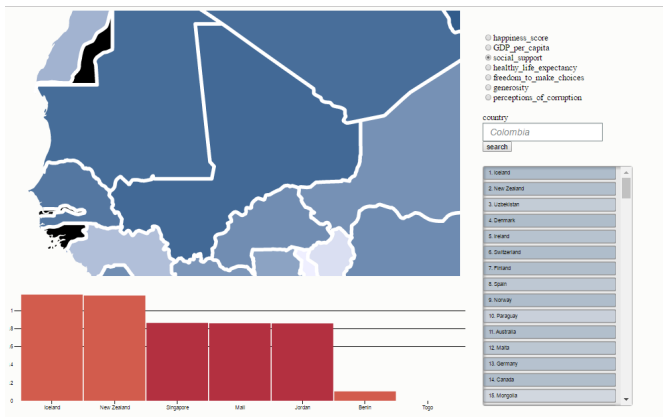Fig. 12. General view with an index filter applied. Authors.

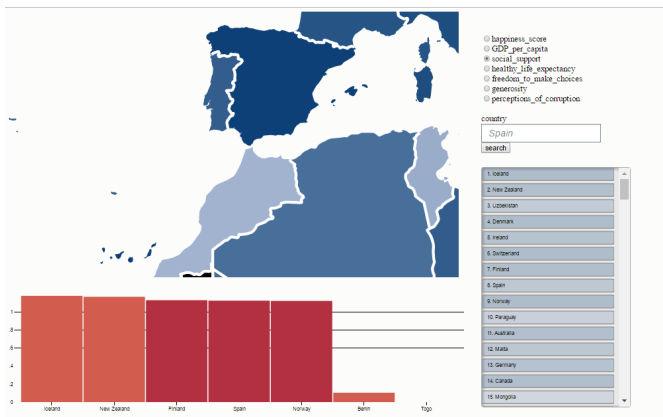Fig. 13. Bars chart when a country is selected. Authors.

Fig. 14. Ranking and lateral filter (text entered: Spain). Authors.

Major achievements that can be summarized:

-The architecture was designed to work 100% from the browser, developed in D3 (javaScript) with a backend that hosts the data. This architecture facilitates the interaction with the information, through the use of web services that allow the communication with the server.

- The use of a successful assessment taking into account the focus population. The result of this project was to answer certain questions such as: What is the relation of the index of corruption with the score of happiness? Is there a high correlation between life expectancy and people's happiness? What is the happiest country in the world? Where is my country at this moment in the ranking? How are the indexes in my country? How is my country respect to others countries? How is my country over time ranking?

As future work, according to the recommendations of the class tutors and previous research of the project, the work of Hans Rosling "Gapminder"[1] can be incorporated in the work already done, as it allows users to easily disaggregate information and see details on demand very intuitively.

REFERENCES

[1] R. K. Drew Skau, Lane Harrison. An evaluation of the impact of visual embellishments in bar charts. *Computer Graphics Forum (Proceedings EuroVis)*, 34, 2015. To appear.
[2] R. L. Helliwell, John F. and J. Sachs. World happiness report 2015. *IEEE Transactions on Visualization and Computer Graphics*, 2015. To appear.
[3] I. A. A. V. Z. B. P. I. S. M. I. S. S. A. O. Michelle A. Borkin, Student Member and S. M. Hanspeter Pfister. What makes a visualization memorable? *IEEE*, 2013. To appear.
[4] M. B. Mikael Jern and L. Thygesen. A web-enabled geovisual analytics tool applied to oecd regional data. *NCVA National Center for Visual Analytics, ITN, Linkoping University, Sweden, OECD, Paris, France*, 2015. To appear.
[5] T. Munzner. *Visualization Analysis and Design*. AK Peters Visualization Series. CRC Press, 2014.
[6] Tableau and T. Software. *Visual Analysis Best Practices, Simple Techniques for Making Every Data Visualization Useful and Beautiful*. Tableau and Tableau Software. 2016.

[1] His work can be find in www.gapminder.org