

An information visualization application case to understand the World Happiness Report

Nychol Bazuerto-Gomez^{1[0000-0003-0881-6736]}, Carlos Torres J.^{1[0000-0002-5814-6278]}, Raul Gutierrez^{1[0000-0003-1375-8753]}, Mario Chamorro^{3[0000-0002-7247-8236]}, Claire Bulger^{4[0000-0002-3031-2908]}, Tiberio Hernandez^{1[0000-0002-5035-4363]}, and John A. Guerra-Gomez^{1,2[0000-0001-7943-0000]}

¹ Imagine group, Universidad de los Andes, Bogota, Colombia
`{n.bazuerto,cf.torres,ra.gutierrez,jhernand,ja.guerrag}@uniandes.edu.co`

² UC Berkeley, California, Berkeley, United States

³ Make it happy, San Francisco, United States
`marioachamorro@gmail.com`

⁴ World Happiness Report, San Francisco, United States
`clairebulger@gmail.com`

Abstract. Happiness is one of the most important components in life, however, qualifying happiness is not such a happy task. For this purpose, the World Happiness Report was created: An annual survey that measures happiness around the globe. It uses six different metrics: (i) healthy life expectancy, (ii) social support, (iii) freedom to make life choices, (iv) perceptions of corruption, (v) GDP per capita, and (vi) generosity. Those metrics combined in an index rank countries by their happiness. This index has been published by means of a static report that attempts to explain it, however, given the complexity of the index, the creators have felt that the index hasn't been explained properly to the community. In order to propose a more intuitive way to explore the metrics that compose the index, this paper presents an interactive approach. Thanks to our information visualization, we were able to discover insights such as the many changes in the happiness levels of South Africa in 2013, a year with statistics like highest crime rate and country with most public protests in the world. In the same fashion, the application illustrates recent Yemen's history, which in 2011 had a fall in the ranking, as in 2014 and onwards. Looking at the historical facts of this country, we see in 2011 a revolution began, explaining the fall in the happiness level, followed by a period of transition, that explains the peak in following years (2012-2013), finally, in 2014 the revolution continues, resulting in a civil war.

Keywords: World happiness, Coordinated and Multiple Views, Information visualization application

1 Introduction

The world happiness report is a survey which measures the happiness in almost all the countries of the world, this score happiness is sustained by several indexes, such as corruption, or life expectancy. This information, for sure, can be interesting for anyone, however, these reports are large, a little complex and their approaches to show its results had a lack of interactive visualization using a large bar chart to the ranking (which extends by several pages) and in a previous release, an image of a map with a color scale was used to highlight distribution of happiness.

For example, if one wants to determine how happy is a country, one can use a choropleth map and select a country to know the indexes and score associated with a text, but others might want to know how was that country in respect others, extremes for example, which countries are the nearest to his selection in the ranking, which countries are happier with respect your region, or what is the country with higher life expectancy. Also, it exists more than one version of the report, user maybe wants to compare how was his country over the last years, so we can't use a classic and simple visualization. So, how it is going to be exposed, a juxtapose and linked visualization can be used for these tasks, where the choropleth map is refreshed to express score, but also distribution of one index in all the world, allowing the user to explore not just about happiness, but indexes who are related, like corruption.

We evaluated this proposal using a usability test with people of different areas (engineering, math, and school environment), to corroborate advantages with respect last approach and possible future work.

Paper organization. Section 2 briefly describes previous work on similar data such as OECD Better Life Index and the World Happiness Report itself. Section 3 illustrates the Visual Analytics basics applied to the particular problem: characterization of data and tasks. Then, we expose the design of this proposal, emphasizing the idioms used to encode. Finally, we summarize our results and experiments, conclusions and future work.

2 Related work

It's important to take into account that the visualization has different perceptions and preferences depending on the user watching it [1]. For example, a bar chart is not always good for all the users: are not good enough for the scientific community users and the infographics [1] but for new media and government visualization is well received. So, we explored different idioms and tools that could satisfy a wide range of users and that handle similar data.

2.1 Previous reports

In the last reports [2–5] we could 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 that users could be interested in other information related.

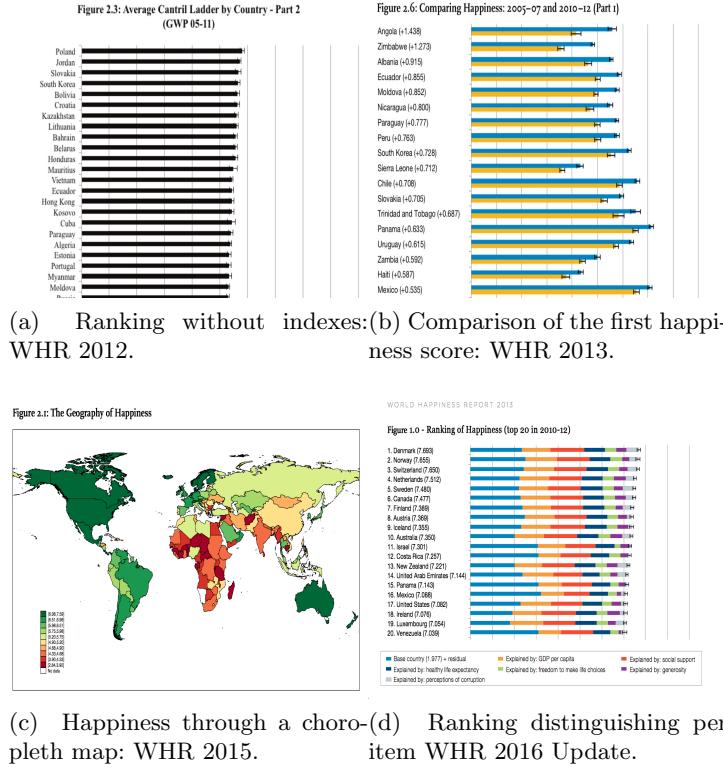


Fig. 1. Main visualizations already made for the World Happiness Report.

In Fig.1a, the graph that represents the level of happiness in all countries, where the survey is taken, shows the data that belongs to 2012. This bar chart is quite basic, no colors are used to represent additional information at the level of happiness. In the following year (Fig.1b) it can be seen that they already include the use of colors to make comparisons with previous years and discriminate by ranking.

The following report is published in 2015, where a map showing the information of each country using colors was incorporated (and is the most common solution for this). 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 chart showing the trends, or even just a crosstab to show the actual data. In this same report, the information for 2015 was not included, so in 2016 an update of the previous report was published, including comparisons with previous years.

It is highlighted that in none of the published reports is possible to interact with the data, they are simply demonstrative graphs that do not allow the

community to investigate more about the data of the report or obtain insights from it.

2.2 OECD Better Life Index

This index is developed by Organization for Economic Co-operation and Development and tries to express how life is based in 11 topics which were identified as essential in quality of the same. Among topics are: Housing, Income, Jobs, Education, Health and Safety. Currently, OECD has a nice visualization for this data [6], letting user select index that is more interesting for him and see another data, like region (a very particular region, for instance Alberta, Canada) and summarize the topics values in this area. It uses right channels and marks for its data and there is sorting options which simplifies determine extremes per topic; however, it's difficult compare entire world, appreciate distribution of one index and rankings are bounded.

2.3 Dashboard and geographic data

In the Whitepaper visual analysis guidebook [7] 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. 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 (and this is our case).

On the other hand, to address the problem of displaying geographic data, the Linkoping University created an interesting solution [8] using coordinated views or dashboard, as we mentioned earlier, 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.

2.4 Ranking

For countries ranking, based on one or a set of criteria we found: a bar chart is great for comparison and ranking because it encodes quantitative values as the length of the same baseline, making it extremely easy to compare values [7]. A sorted list of country names is a different kind of view and it fulfills the main task that consists of getting the max and min of the countries, the choice between ranking as a list or as a bar chart is given by the accuracy of the comparison. According to Robert K. most of the embellishments have an adverse effect on the accuracy of reading values when comparing bars. [9].

In Fig.2 is possible see a brief comparison between the current approaches for this data (and similar) and its limitations. As illustrates the World Happiness Report's visualizations have many difficulties. This, it is the result that there aren't interactions but several countries, extending the graphic along different

pages. In addition, each idiom was focused on a task. On the other hand, the OECD is better in most of the aspects but the glyph idiom makes difficult some tasks. Its big advantage is the interactions, however, they just handle 38 countries.

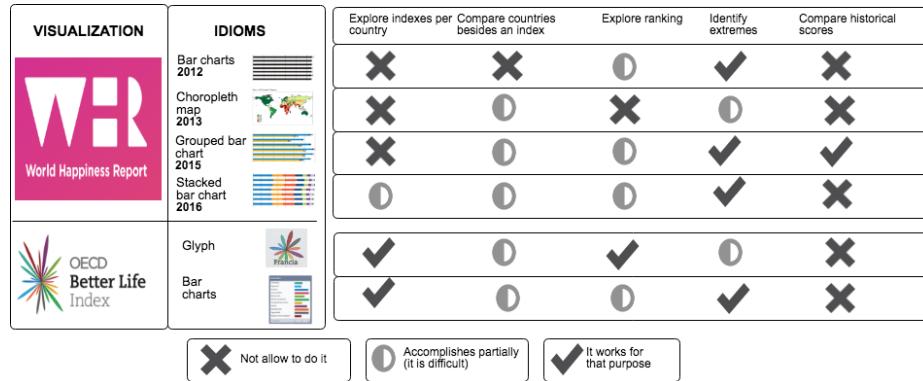


Fig. 2. Comparison between current approaches. Authors.

3 Visual analytics tool abstraction

We did the abstraction based on Tamara Munzner [10] framework which is composed for What, Why and How. This part just describes What (data abstraction) and Why (task abstraction).

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 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. But we thought anyone must to be able to understand and interact with the indexes which are associated, in order to answer different user questions. With this and the data in mind five basic 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*).
6. **Discover** happiness **distribution** in the different world regions. The user can interact with each country in those regions to obtain details on demand.

4 Design and description of the visualization

Continue with the Tamara’s 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.

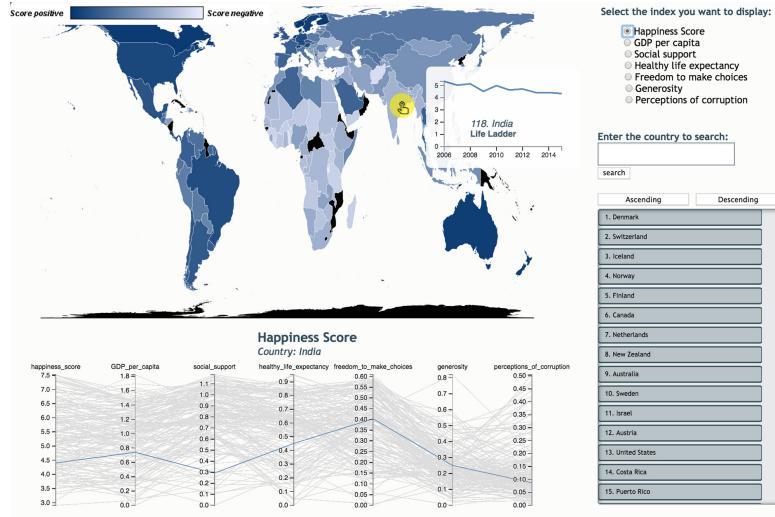


Fig. 3. World Happiness Report visualization. India is selected with happiness score as filter, and its historical scores are shown by a pop up over the country. This let us appreciate that India is descending across the years.

4.1 Idioms: encoding

This proposal is composed for 4 main idioms; Fig.3 shows a view using the initial setup (score happiness as filter).

1. A choropleth map: This idiom supports the second and fourth tasks, 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 or an index (something quantitative and sequential). User can **manipulate** the choropleth, in this case, map allows **select** and **navigate** by all the map. The idea of this design was let the user picks up on the map any country and apply a geometric zoom of that country. Moreover, it's presents a reduce action,when 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).

On the other hand, 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

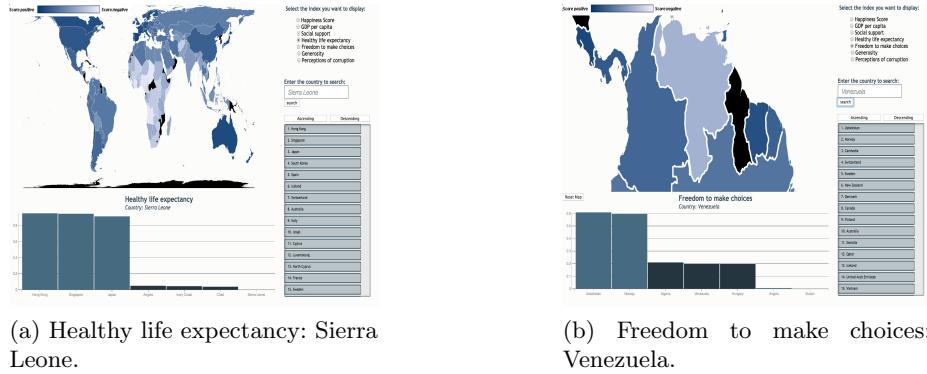


Fig. 4. Bar Chart for different index and countries. Authors

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. User can select and highlight one line which corresponds to a particular country and also filter items to get not only the information of a country but a set of countries with similar values in certain indexes. This idiom supports the third task in conjunction with the map.

3. **Bar chart:** This idiom let show and compare one index, visualizing the highest, lowest countries and the selected country by the user (there is a link between choropleth map and bars). Respect to selection on the map, highlights and selects a country.
4. **Ranking list:** We thought just the map isn't 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). This idiom let us support the first and fifth tasks. User can manipulate changing order, from lowest to highest position or vice versa and by a filter associated reduces shown items to the user.
5. **Bubble chart:** This idiom allows appreciating a different way to see the happiness, avoiding the possible bias produced for the choropleth map and the countries' size (user can understand that Russia is happier than Germany due to the area each country covers). This bias is handled using dots (bubbles) which its size is related with the happiness same as its color, and each one has a repulsion force to avoid overlapping. Additionally, the implemented filters for this idiom show a classification by happiness levels and by regions, supporting the sixth task (through arrange-separating the data).

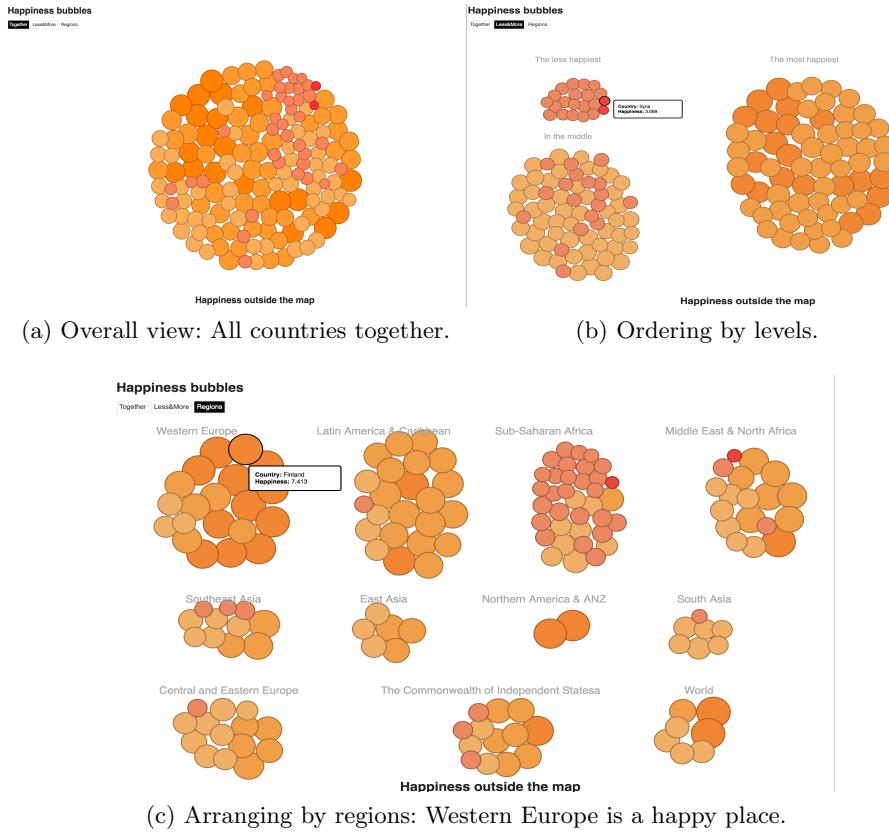


Fig. 5. Bubble chart through the different data arranges. Authors

4.2 Interactions

All the views are linked with shared data with purpose to give the user a richer experience with the data, letting him analyze and relation indexes between different countries and other characteristics. User can select in radio buttons which information he wants to see distributed on the map: about score happiness or a single index. If user selects score happiness as filter, he's going to see the parallel coordinates chart, which relations indexes of all countries. If user decides see information about an index, like generosity, a bar chart, instead parallel coordinates, is going to appear, as is possible see in Fig.4, when sub Fig.4a shows healthy life expectancy with Sierra Leone bar in center (it is the selected country by user) and let compares with his neighbors in values and extremes. Sub Fig.4b shows how Venezuela has a freedom to make choices 4 units behind the highest, and Angola is one of the worst in this aspect (country with several problems with human rights).

The country selected could be updated when user picks up on the map other countries. Moreover, as soon as the user moves the mouse to through a country, embedded chart appears which is a chart line about the country positions in the ranking over the last years, even if user has an index selected, this chart appears to give the user a summarize about happiness.

On the other hand, there are several filters to help the users to find information quickly: there is a filter by country name which makes automatic selection and the country is zoomed in. Also, there is filters incorporated to parallel coordinates, in order to allow the users to combine indexes and ranges of his interest, in Fig.6 appreciates an example, when countries with freedom to choices is between 0,10 and 0,15 and also have GDP per capita between 0,8 to 1,0 are highlighted.

Also, in order to make easier to find extremes, an ordering mechanism was integrated, listing in ascending and descending way, this works too when an index is selected affecting the bar chart.

Finally, the user can switch to a bubble chart visualization. This visualization has three presentations: a general view, a happiness level classification, and region classification. In Fig.5 is possible to see the three data arranges and the detail offer when the user hovers a bubble. In Fig.5a is possible see all the countries where the red color highlights the less happiest countries. Then, in Fig.5b the bubbles are split by three levels: the highest, average and lowest happiness. In addition, Fig.5c shows the bubbles classified by regions.

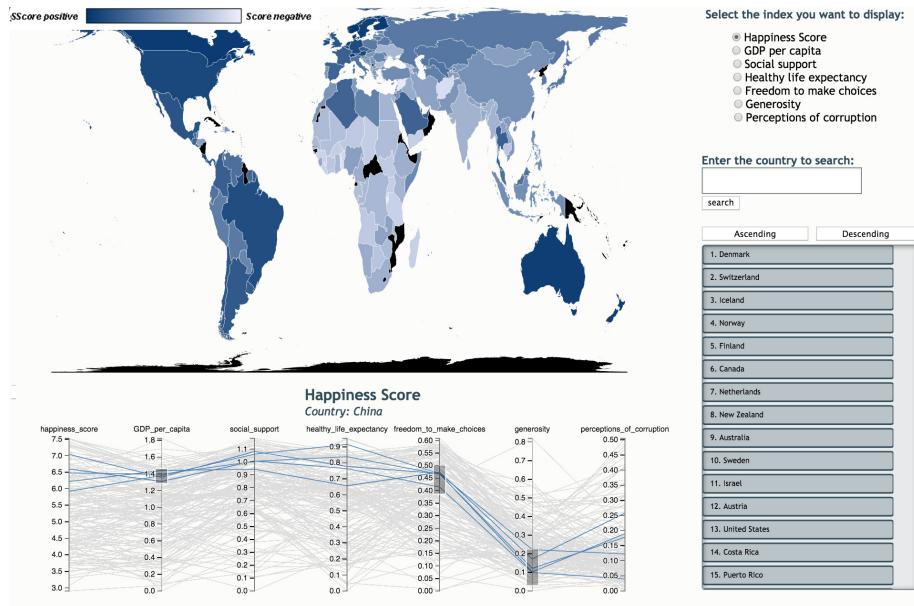


Fig. 6. Filter on parallel coordinates. Authors.

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.

5 Usability test and results

We ran a usability test with people from different areas and age: Music, math and Chemical students, systems engineers and math professionals, covering demographic diversity. In the first step, the subject had to explore current report, understanding about what it is. This step was repeated with our visualization. Usually, the first interaction that they had was with the map and then with radio buttons, trying to figure out relations between different components.

In general, all subjects agree with some aspects when interact with the report: It was a little hard to find a particular country, if they were lucky country searched was near to page opened. Also, know the values about indexes was complicated an they estimated.

On other hand, with our visualization they agreed on one thing: It easiest and didactic understand and learn about world happiness. Some suggestions were made about other interactions as select over the list and update all views, and a user told us about a different color for countries without data, because can be confusing, as darker color could be interpreted as "happiest color".

After exploring it free both scenarios, users had to solve the some tasks:

1. Can you identify what country is the least happy in the world?
2. Could you tell us how your country is and its indexes?
3. Can you tell us what country has highest life expectancy /generosity index?
4. How is your country respect of country with highest and lowest generosity index?

In the report, all of them had problems looking for its country and they highlighted it isn't friendly have to slide down by many pages to find the least happy country. The third and fourth tasks couldn't be determined with precision because is so complicate compares indexes between countries. By other hand, with visualization most of them could do all the tasks and understand some interactions like bar chart was more complicated than other like a filter. In final feedback, they highlight the next things that they could do and the couldn't with report and was so interesting for them:

- It's possible find and interesting geographic correlation. You can see some clusters, countries who their culture and economics are related, for example places in occidental Europe presents a similar behavior and could be justified by his acquisitive power with GDP index, as Fig.7 shows. This is lost with the report because barely you can compare indexes between countries.
- Historic positions in the ranking give you a panorama of growth of the country or let establish a relationship with possible conflicts in some areas, so this perspective brings an extra information about report who exposes separately these results.

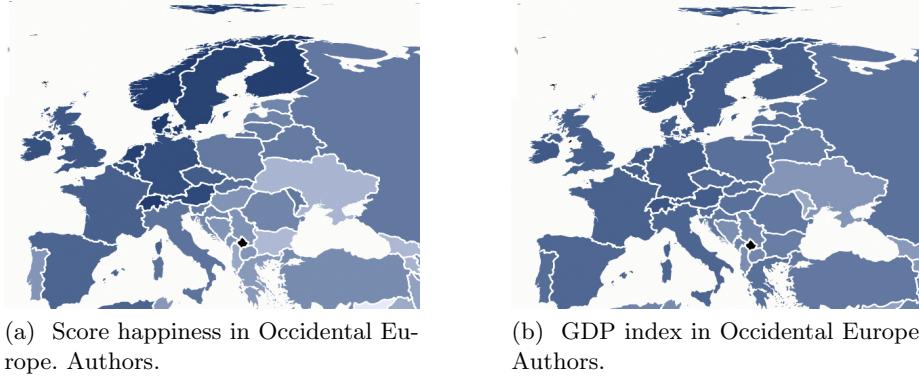


Fig. 7. Interested case found by user: A geographic relationship can be appreciate. Authors

- It's possible see value for each index, so user doesn't need to estimate this.
- Allows evaluation from a single index over all countries, so it's interesting analyze how is perception of corruption is in the whole world.

And we reaffirm the possibility to do this:

- Make easier to find a country and know its location.
- Find faster the extremes in indexes and compares with a particular selected country.
- The user doesn't need to see all ranking, he can use descending order if is looking for the worst 10 countries in kind aspect (score happiness and index).
- Get countries which have a similar behavior or are close in some indexes intervals.

General feedback was positive, visualization was seen as an easier way to handle the report results.

6 Conclusion and future work

The usability testing demonstrates visualization enables the regular users to take a more active role in the discovery process of exploring world, regional and country indicators.

Its main purpose was accomplished, improvement the way to the user solves main tasks and expressing the data in the correct way. Moreover, we have shown how a user can deduce and relate external information, looking for effects of politician regimens, conflicts, economic crisis and even events like realization of soccer world cup, this was a titanic labor with the report, because there isn't a perspective can give enough information without incurs in saturate users' mind.

Nevertheless, we identify two possible enhances. First, improve the visualization by regions through linking the bubble chart to others idioms, this as future work, when the work of Hans Rosling “Gapminder”¹ can serve as inspiration to incorporate a similar solution in the work already done, as it allows users see details on demand very intuitively.

Secondly, develop a version with texts more informative for the user, if the user doesn’t know what is each index and its explanation, it could be a little confusing, maybe incorporate a menu with help or an animation about world happiness report.

World Happiness visualization was designed to work 100% from the browser, developed in D3 (javaScript) with a backend that hosts the data. It’s an open source project available at <http://nycholbazurto.me/worldHappinessReport/viz/>. Ultimately, we hope that it enables a new way for interested people in understanding the report.

References

1. Borkin, M.A., Vo, A.A., Bylinskii, Z., Isola, P., Sunkavalli, S., Oliva, A., Pfister, H.: What makes a visualization memorable? IEEE (2013). <https://doi.org/10.1109/TVCG.2013.234>
2. Helliwell, J., Layard, R., Sachs, J.: World happiness report, update. Tech. rep., World happiness report (2012)
3. Helliwell, J., Layard, R., Sachs, J.: World happiness report 2013. Tech. rep., World happiness report (2013)
4. Helliwell, J., Layard, R., Sachs, J.: World happiness report 2015. Tech. rep., World happiness report (2015)
5. Helliwell, J., Layard, R., Sachs, J.: World happiness report 2016, update. Tech. rep., World happiness report (2016)
6. Durand, M.: The oecd better life initiative: How’s life? and the measurement of well-being. *Review of Income and Wealth* **61**(1), 4–17 (2015). <https://doi.org/10.1111/roiw.12156>, <http://dx.doi.org/10.1111/roiw.12156>
7. Tableau, Software, T.: Visual Analysis Best Practices, Simple Techniques for Making Every Data Visualization Useful and Beautiful. Tableau and Tableau Software (2016)
8. Jern, M., Brezzi, M., Thygesen, L.: 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). <https://doi.org/10.2312/ega.20091004>
9. Skau, D., Harrison, L., Kosara, R.: An evaluation of the impact of visual embellishments in bar charts. *Computer Graphics Forum (Proceedings EuroVis)* **34** (2015). <https://doi.org/10.1111/cgf.12634>
10. Munzner, T.: *Visualization Analysis and Design*. AK Peters Visualization Series, CRC Press (2014), <https://books.google.com.co/books?id=dznSBQAAQBAJ>

¹ His work can be found at www.gapminder.org