Visualization of World Happiness

Iris Kotsinas

Abstract— Visualization of World Happiness is a web-based application visualizing the world happiness. The purpose was to study how different factors are related to the happiness score of a country.

Index Terms—Multivariate data, Choropleth map, Scatter plot, World happiness score.

1 Introduction

Information visualization can be used to visually represent data, enabling a user to better understand it. Common examples of such visualizations are choropleth maps and scatter plots. Information visualization allows the user to gain deeper insights from abstract data, by displaying relevant connections and depicting an overview. This project report describes the method used for visualization of a given dataset, as well as the result. The data chosen for this project was a dataset of the world happiness. In a choropleth map, as well as a scatter plot, the data was visualized and analyzed in order to draw conclusions about the data. Moreover, the purpose of this study was to examine the following research questions and visualize the data in ways to properly analyze them:

- Which country had the highest happiness score year 2019?
- Is there a correlation between the happiness score and geographical location?

2 Background and Related Work

Data from The World Happiness Reports of 2015-2019 was used in this study, taken from the website *Kaggle*, which in turn was sourced from the original source *Sustainable Development Solutions Network* [2]. The report is a landmark survey of the state of global happiness that ranks 156 countries by how happy their inhabitants perceive themselves to be. The World Happiness Report is a publication of the Sustainable Development Solutions Network, powered by data from the Gallup World Poll [5].

The rankings of global happiness are based on a Cantril ladder survey. Representatives of each country are asked to imagine a ladder, where 10 represents the best possible life and 0 the worst possible life. The respondents are then asked to rank their own life on that 0 to 10 scale. Then the average score of each country is calculated, and the Happiness Score can be retrieved. Several other factors are included in the World Happiness Report, such as GDP per capita and social support. These factors are used to analyze the happiness score of the country and draw connections.

3 Data

The data used in this study is retrieved from the World Happiness Reports of 2015-2019. The data consists of happiness score for over 150 countries along with the factors used to explain the score. The Happiness Score is explained by the following factors:

- GDP per Capita
- Healthy Life Expectancy
- Social Support
- Freedom to make life choices
- Generosity
- Corruption Perception
- Residual Error

Since the report focused on a time period of five years, from 2015 to 2019, five different datasets had to be used. The content and syntax of the datasets differed, therefore in order to enable a working visualization three of the seven factors were chosen, in addition to the happiness score. These factors were GDP per capita, healthy life expectancy and freedom to make life choices.

In order to visualize The World Happiness on a global scale, a map was used. For the visualization of this map, a dataset was used containing information about each country and their coordinates. The two datasets were compared by the country names, and matched.

• *Iris Kotsinas is with Linköping University, Sweden, e-mail: iriko934@student.liu.se.*

4 Method

The World Happiness application is designed to show tendencies of world happiness and the relation to other factors, such as GDP per capita, freedom to make life choices and healthy life expectancy. The application is mainly composed of two views; a choropleth map and a scatter plot. The user can interact with the application and is given the possibility to more specifically choose what information is visualized.

4.1 Choropleth map

To visualize the data globally, e.g. the happiness score, a choropleth map is used. It allows the user to easily examine and analyze the data, and helps the user connect the country with its corresponding data. Choropleth maps are good at utilizing data to easily represent variability of the desired measurement, across a region, in this case the world.

The choropleth map in the application is depicting the world, in which the different countries are colored in proportion to a statistical variable chosen by the user. The default variable is the happiness score, which can be changed by the user in a drop-down menu. In this menu the user is given a choice of four options; the happiness score, GDP per capita, freedom to make life choices and healthy life expectancy. The colors of the choropleth map will change with the option chosen by the user. The colors are gradient, changing from blue to orange, where dark blue represents the lowest value and dark orange represents the highest value. The use of two colors can facilitate the use of the application for color-blind users.

The user can interact with the map by zooming in and out. When the user hovers over a country with the mouse, the country will turn black to indicate that it is selected. As the country is selected, the info box beside the map will change its values and display those of the selected country. The corresponding point of the selected country will also show a black outline in the scatter plot.

A slider component allows the user to change the visualized year.

4.2 Scatter plot

To visualize the relationships between the data, a scatter plot is used. Scatter plots are one of the better known methods in information visualization, familiar to many people outside the field, and can easily

be implemented. Scatter plots are primarily used to observe and show relationships between two numeric variables. The dots in a scatter plot not only represents the values of individual data points, but can also be useful for identifying patterns in the data [3].

The scatter plot visualizes the correlation between GDP per capita (x-axis) and the happiness score (y-axis), with the scatter points representing each country. In addition to this, the radius of the scatter points is used to visualize the healthy life expectancy. The bigger the radius, the higher the life expectancy of the country. The colors of the scatter plot are connected to the choropleth map, with the same color scheme and gradient. When the user hovers over a point with the mouse, the point will increase in size to indicate that it is selected. As the point is selected, the info box beside the map will change its values and display those of the selected country. The selected country will also turn black in the map.

The plot is also connected to the slider component, which allows the user to change the year visualized.

4.3 Minor components

As mentioned, the application contains a slider and a drop-down menu, allowing the user to interact with the visualization.

4.3.1 Drop-down menu

The drop-down menu presents the user with four options, and allows for a change of visualization of the map. The user can choose from visualizing the happiness score, GDP per capita, freedom to make life choices and healthy life expectancy. The default option is happiness score.

4.3.2 Slider

The slider presents the user with five options and allows the user to change the year visualized in the choropleth map and scatter plot. By changing the year, the visualizations are changed. The default year is 2015.

5 Implementation

The data was pre-processed directly in the json-files, in order to make it more manageable. In some cases the data in the different data sets did not match, e.g. some country names, and therefore had to be changed.

The application is web-based and optimized to work on computers. The framework used for the development of the application was React, a JavaScript library for building user interfaces. The slider and the drop-down menu were created with the installation of React components. The D3.js library was used to load and display the data.

6 Results

The World Happiness application was designed to work as an intuitive tool for the user to easily explore the world happiness score and its factors. The website is constructed with the choropleth map in the upper left corner, and the scatter plot directly below it. The choropleth map and the scatter plot were given the most space in the application, since these are the two main components of the visualization.

In Figure 1 the total application can be seen. The first half consists of the two main visualization components, and the other half contains the information box, drop-down menu and slider component, as well as some informative text for the user. If a user hovers over the countries in the map, the country turns black, indicating it is selected. The correlating point's outline in the scatter plot will then turn black. If the user hovers over the points in the scatter plot, the point will increase in size, and the correlating country in the map turns black. As the user hovers over the map or the scatter plot the information box on the right of the map will change its information.

In Figure 2 the choropleth map can be seen in detail. A legend can be found in the lower left corner, which indicates the meaning of each color. There is also a legend for the countries with no data, displayed in the color gray.

In Figure 3 the scatter plot can be seen in greater detail. The x-axis represents the GDP per capita, and the y-axis represents the happiness

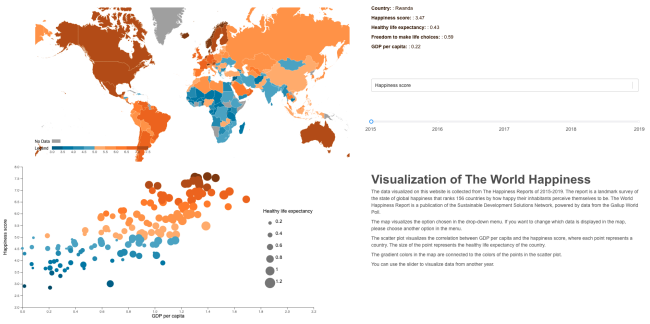

Fig. 1. Visualization of The World Happiness application in full.

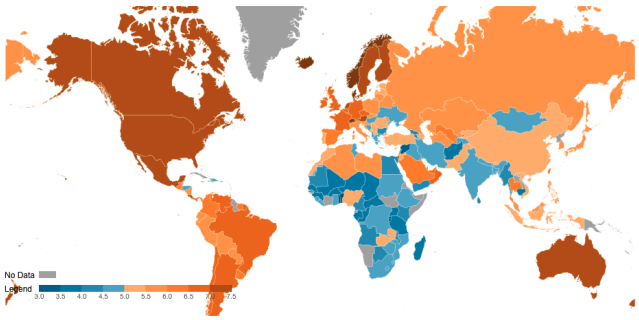

Fig. 2. Choropleth map visualization of happiness score.

score of the country. The points represents the countries. The colors are connected to the colors of the map. There is a legend on the right informing the user that the size of the scatter points represents the healthy life expectancy of the country.

7 Evaluation

An evaluation was carried out on The Happiness Score application in order to test if it was sufficient and understandable enough for a user. The usability attributes investigated were efficiency, learnability and satisfaction. The development of the application was inspired by Interaction Design by Arvola and the process of the evaluation was inspired by Usability Engineering by Rosson and Carroll [1] [4]. Three students were asked to take part in the evaluation. All participants were in the age span 17-23 years, but one of them was in high school and two of them studied at university level, but different programs.

A qualitative evaluation method was applied, *think aloud*, where the participant had to speak their mind while using the application. The evaluation method was a qualitative way of observing the participants in order to measure emotion and thinking process.

In order to measure the usability attributes a number of factors were examined. Learnability was measured with the time spent on each task by the user, in order to examine how easy it was for users to accomplish basic tasks the first time. Accuracy was also used to measure learnability in order to audit how well the users could achieve specific goals. The attitude of the user in the beginning of the test was observed in order to measure the frustration level.

Efficiency was measured with time spent by the user on each task, as well as the mental workload of the user. Time spent was measured for efficiency to give a sense of how quickly the user could perform the tasks, once they had learned the system. The accuracy of more advanced tasks was measured, in relation to the resources expended and completeness of goals achieved.

User satisfaction was measured by observing the attitude of the user, as well as the mood, when performing tasks. Questions were asked

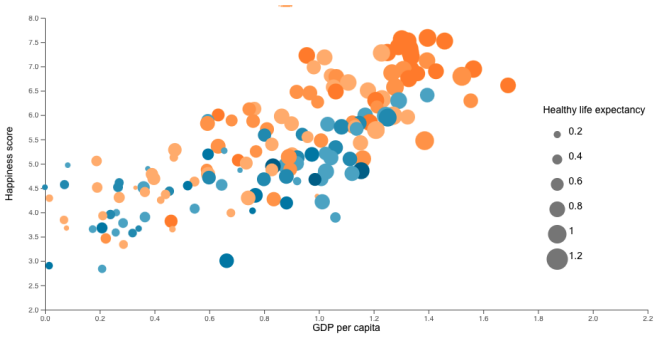

Fig. 3. Scatter plot visualization in detail.

concerning frustration level, acceptability of the system and mood. This gave the moderator a sense of the user satisfaction, what the user thought of the system and what increased their frustration level.

The test was carried out remotely via Zoom, with the use of web cameras in order for the moderator to study the participant's facial expressions. The test was designed in three major parts. First there was a brief introduction to the background of the project and the purpose of the evaluation. Thereafter, the participant was given three tasks to perform.

- Find a country with a high happiness score from the year 2019.
- Find which country has the highest happiness score as well as GDP per capita from the year 2016.
- Find a country with a low freedom to make life choices from the year 2018, both in the map and the scatter plot.

The participants were encouraged to speak aloud when performing the three tasks, allowing the moderator to follow the person's thinking process. As the participant finished a task, the next one was given. When all three tasks had been completed, the third step of the test was held. This step consisted of a short interview with some questions to give the moderator information about the user satisfaction, learnability and efficiency of the application. Lastly, the participants were encouraged to speak freely about the application. The questions asked were:

- What did the map represent and display?
- What did the scatter plot represent and display?
- How did you feel about the application?
- Could you draw any conclusions from the visualizations in the application?

The result of the evaluation was that the application was easy to understand and visually pleasing. The scatter plot did, however, not represent the data as well, and could be improved. The users thought the information box could have been closer to the scatter plot to clearly show the connection between the two. Apart from that, the users appreciated the application, and especially the hover-effect, as it contributed to the user feedback.

The participants all agreed that the map was easy to use and understand. The hover-effect increased the efficiency, since it distinctly made the location of the data obvious for the user. The zoom effect allowed the user to study the map in detail, which was appreciated.

Visually, the application was aesthetically pleasing and intuitive. The participants liked the instructions given and stated that it increased the user satisfaction as well as the learnability of the application. The feedback on the colors used in the application was positive, and the participants thought it contributed to a better understanding of the data. The visualizations allowed the user to draw interesting conclusions.

Time wise, the tasks took less time to complete once the user had gotten familiar with the system. This indicated that the learnability of the application was good. The time spent on each task was quite small which indicated that the efficiency was good as well. The attitudes of the users were positive, and none of the participants expressed or showed signs of having a negative attitude while using the application. Overall, the evaluation resulted in mostly positive feedback, but as previously stated also some negative.

8 Conclusions and Future Work

In regards to the research questions mentioned in the introduction and from the result, a conclusion can be drawn that Finland had the highest happiness score 2019. This could be obtained from both the choropleth map, and the scatter plot. It could also be found that Nordic countries top the ranking, and overall had a higher happiness score than countries and regions such as Central African Republic, South Sudan, Tanzania, Rwanda and Haiti. As can be concluded, self-reported life satisfaction correlates with other measures of well-being. Richer and healthier countries tend to have higher average happiness scores, and therefore one can state that the geographical location correlates to the happiness score.

A conclusion that can be drawn from the evaluation was that the application was easy to use. It provided good instructions and user-friendly tools. The use of two different colors to represent the data made a positive contribution to the user experience and made the visualization easier to understand.

Another conclusion that can be drawn from the result and evaluation is that while the application was easy to use, the data could have been represented in more complex ways. For example, instead of the use of a scatter plot, parallel coordinates could have been used to represent and visualize the factors and their correlations. This would have contributed to a more accurate result, and allowed the user to notice abnormalities in the data, and therefore helped display the complexity of the connections between the factors. The use of a regression line could also have improved the visualization of the scatter plot. Another future work would be to allow the user to compare the data throughout the years, perhaps presented in a scatter plot, instead of having to use the slider.

In conclusion, the application met the requirements of the research questions and provided the user with useful tools to help draw conclusions from the data. The visualizations could however had been more advanced to sufficiently visualize the complexity of the data. Apart from that, the present visualizations did contribute to the purpose of the study and quite accurately represented the data of world happiness.

References

[1] M. Arvola. *Interaktionsdesign och UX - om att skapa en god användarupplevelse*. Studentlitteratur AB, Lund, 2014. edition: 16.

[2] Helliwell, John F., Richard Layard, Jeffrey Sachs, and Jan-Emmanuel De Neve. *World Happiness Report 2020*. Addison-Wesley Longman Publishing Co., Inc. `https://www.kaggle.com/mathurinache/world-happiness-report`, 2020. Last accessed: 2021-03-10.

[3] R. Kosara, H. Hauser, and D. L. Gresh. *An Interaction View on Information Visualization*, 2003.

[4] M. B. Rosson and J. M. Carroll. *Usability Engineering: Scenario-Based Development of Human-Computer Interaction*. Morgan Kaufmann Publishers, San Fancisco, CA, USA, 2002.

[5] Sustainable Development Solutions Network. *The World Happiness Report*. Addison-Wesley Longman Publishing Co., Inc. `https://worldhappiness.report/`, 2021. Last accessed: 2021-03-08.