Happiness Matters: An In-Depth Look into UN's World Happiness Reports through Data Visualizations

Chaw Khin Nyein

S Jonas Nevin

Wang Sheng

Abstract—In this age of globalisation where countries are growing rapidly, we aim to find out whether countries are becoming happier, and what sets happier countries apart from least happier countries. To help with this, database from UN's World's Happiness reports are used and various visualizations such as bar charts, maps, line graphs, boxplots and radar charts are created to facilitate exploration in learning more about the various countries around the world. This paper highlights the various objectives that we aim to achieve with our visualization, the best practices used in each visualization as well as the key findings obtained.

Index Terms—Happiness, UN, World Rankings, visual analytics, GDP.

1 Introduction

Since 2012, United Nation Sustainable Development Solutions Network (UNSDSN), the branch tasked with promoting practical solutions for sustainable development globally, has been producing the World Happiness Report that provides the happiness scores of approximately 156 countries globally. These scores are attained from surveying around 2,000 to 3,000 [1] individuals from each of the participating countries by using 6 main factors as follows:

1. Income of individuals living in a country as measured by GDP per capita
2. Healthy life expectancy
3. Social Support
4. Trust (in Government) and perceived lack of corruption
5. Freedom to make decisions(life choices)
6. Generosity

The reports feature extensive data on the happiness scores, however, few clarification and links are made to these 6 factors than it is comprised of. Moreover, the trends between these factors and the final happiness index are not extensively explained. For example, a country ranked first in the world happiness index might not rank first in one of the 6 factors. Therefore, our analysis explores these reports in-depth and help identify exactly what makes up a happier country as per UN's standards. Furthermore, we aim to see what sets happier countries apart from least happy countries in relation to the 6 factors that make up the happiness index.

This paper, thus, presents the findings, and reports on the reasoning behind the final visualizations selected to provide this in-depth analysis. This paper can be broken down as per the following: Section 1 and 2 provides an introduction and the objectives of this project respectively; Section 3 provides information on our survey of work that inspired our final visualization and some of the features that would be emulated in our visualization; Section 4 shows an in-depth look at our application including the data preparation methodology while Section 5 presents the key findings from our web application. Lastly, in Section 6, the paper ends off by suggesting some future work that can be done to enhance the application.

2 Motivation & Objectives

The motivation behind our data visualization application was mainly due to the lack in analysis of the 6 different factors used to measure happiness in relation to the final happiness index. Currently, there is no efficient way of looking through the various data presented in the World Happiness Report without looking through each page at a time. Thus, our visualization aims to facilitate this in-depth exploration through an easy-to-use and interactive web application.

Overall, the 4 main objectives to be achieved by our visualization can be summarised as follows:

1. Identify the overall trends of happiness ranking index of different countries from 2015 to 2018
2. Identify the rankings of each of the 6 factors that make up the happiness rankings and how they are related to the final happiness rankings.
3. Compare the 6 factors between different countries from 2015 to 2018.
4. Compare different countries by geographical locations to find out the happiness patterns by regions.

3 Survey & Application of Related Work

In order to gain inspiration and help identify which types of visualization would best meet our objectives, research was conducted, and the key takeaways are presented in the following.

3.1 Geographical Data

3.1.1 Map Visualization

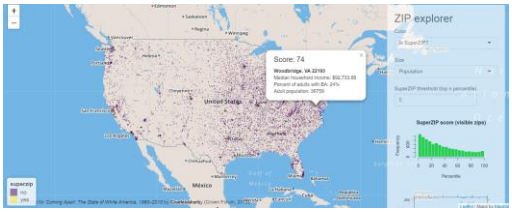

Fig. 1: Symbol Map

The world map was identified as the clearest visualization to help with objective (4) that is related to identify happiness patterns by geographical locations. The map as per Fig. 1 is a good example that can be emulated to present the happiness index of the various countries in our visualization.

What sets this visualization apart from other forms of geographical visualizations is its use of colours. This visualization uses two different colours to denote two different areas. This is a key takeaway used since our goal is to distinguish "happier" countries from the least "happier" countries. Similarly, one colour can be used to identify happier countries while another could be used to identify least happier countries. Furthermore, this is in line with ColorBrewer's advice for cartography where diverging colours are used to represent two diverging variables [2].

Moreover, the visualization, as per Fig.1, uses symbols to denote the different states. However, in our visualization, entire countries are taken into consideration. Therefore, the choropleth map would be more useful in our case. Likewise, based on axismap.com, we have ensured that the requirements for a choropleth map are met. Firstly, the data being used is attached to enumeration units of countries. Secondly, the values have been standardized for all 6 factors. Lastly, the data has a continuous statistical surface [3].

The last main feature to be included is the tooltip that features specific values and information of the selected countries. This is important as it helps highlight other important information that cannot be seen through the map alone.

3.2 Multi-dimensional Data

3.2.1 Bar Chart

The interactive bar chart, similar to Fig. 2, was identified as a relevant tool in meeting objective (2) where the goal is finding trends to see what exactly sets happier countries apart from least happier countries. Firstly, the different lengths of the bar charts make it easy to discern between the different countries and their values.

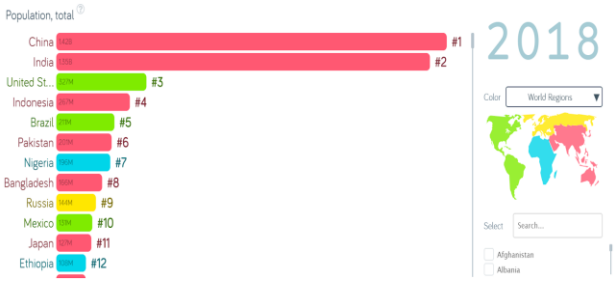

Fig. 2: Bar chart showing country population

Moreover, a drop-down menu filter (Fig.2) allows users to select the specific attributes they want to see in the bar chart. This filter together with scroll bar function is seen to present the visualization in an uncluttered and organized manner. This is a main takeaway especially since the World Happiness Report presents data for around 156 countries and 6 separate factors. A major challenge is to present all these data in a way that does not lead to a mess, thus, these methods and functions used in Fig. 2 serve as good examples to help with the organisation of our visualization application. Specifically, to create a big scale graph, the aim is to reduce as much chart junk as possible and having a drop-down filter would help with that [4].

3.2.2 Radar chart

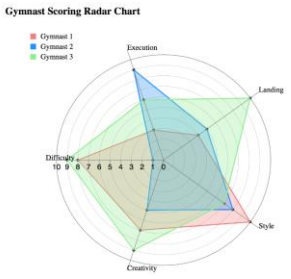

Fig 3: Radar chart

In order to meet objective (3) which requires us to work with the 6 multi-dimensional factors as well to present a comparison between selected countries, the radar chart was selected. This visualization presents data in a consolidated manner and is easy to comprehend. For instance, upon selecting different attributes (countries), users can easily compare the differences in the 6 factors between different countries by observing the size and shapes of the polygons as they overlay on top of each other (Fig. 3).

3.2.3 Box Plot

The boxplot was identified as another useful tool in exploring the 6 factors of the happiness index further. It adds an additional layer to the analysis as it uses descriptive statistics to provide information on the distribution of the data. Tooltips that present the five-number summary, minimum, first quartile, median, third quartile, maximum distribution, can be used to highlight the information on the various countries further.

3.3 Time-Series Data

3.3.1 Line Graph

A time-series visualization is crucial in achieving objective (1) where the goal is to view happiness trends of countries between the years of 2015 to 2018. Therefore, the line graph was the best way to present such data clearly.

4 The Application

4.1.1 Data Preparation

Data Collection

Initially, data was collected from various sources including the official World Happiness website, Kaggle, and the Happy Planet Index website. However, after thorough observations of these data sets, only those from World Happiness websites and Kaggle were identified to be sufficient, correct, well-structured, and relevant for our objectives. These datasets are considered as "Tidy data" since the data structure follow Cord's 3rd Normal Form (3NF).

Data Cleaning

For some datasets, certain columns were missing. For example, "Happiness Ranking" was absent in the 2016 dataset and they had to be calculated accordingly to their happiness scores. Our final data sets of 2015, 2016, 2017, 2018 carry the headers and sequences as follows:

"Year, Country, Region, HappinessRank, HappinessScore, Economy, Family, Health, Freedom, Trust, Generosity and DystopiaResidual."

Furthermore, some of the data visualizations required further transformations for the visualizations to be viable. For instance, as the 6 factors are given in different scales, normalization of data ((x - min(x)) / (max(x) - min(x))) was conducted for the 6 factors to be presented in the radar chart and choropleth map.

Data Source Management

Each data set is uploaded to Github repository for easy data management that includes accessing the data from the source code, updating and checking the data.

4.1.2 System Architecture

Data are retrieved via source codes written in html, css and js development languages. Two main external libraries, D3 and plotly, are used for creating the visualization. Users can access the application via the link through chrome or firefox browser. The application is deployed via AWS web service at http://34.218.239.243/HappinessMatters/.

Fig.4: Tools and Technologies Used

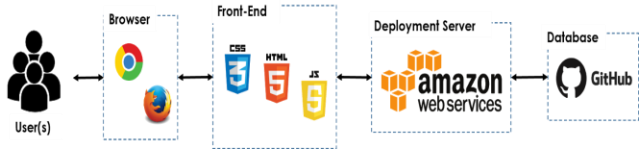

Fig.5: System Architecture

4.1.3 Data Visualization Application

The design of our visualization application begins by providing a general overview before drilling down to graphs that analyses the 6 supporting factors in various ways. All these visualizations are tied together using a web application. The process of navigating through the different visualizations has been made easier by the navigation tab located at the side of the application in which each of these tabs contain a different visualization. Altogether, there are five components (tabs) that help address the issues as per our objectives and each of these components incorporates features identified in section 3 of this report.

Choropleth Map

The first visual component in the application, under the "Regional Happiness Overview" tab, presents the general happiness index overview using a choropleth map, specifically an unclassed choropleth map where each data value is given a unique colour [3]. This approach to choropleth map allows even the subtle difference to be highlighted greatly and helps distinguish between values in-depth. Furthermore, the visualization uses a blue, brown and red color palette to represent happy, neutral and sad countries respectively. Using color palette instead of color gradient allow the viewer to spot the happier countries and least happier countries at first glance as the difference in color is more obvious. It also enhances viewers' ability to differentiate the outlier quickly.

Line Graph

The second component of the application, the line graph, displays the time-series data of happiness ranking of all countries from 2015 to 2018. Different colors represent different countries and there is a filter to choose countries. Thus, the viewers can compare the happiness rank between countries consequently. For instance, in Fig. 6, it seems to show Singapore happiness rankings to be dropping from 2016 onwards.

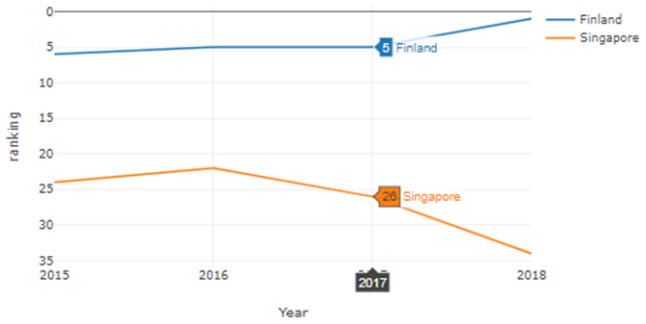

Fig. 6: Line Chart from Application

Bar chart

The third component of the application, under the "Rank by Factors" tab, shows the bar chart (Fig.7) for each of the 6 factors of the happiness index. Initially, the user would be introduced by the graph for GDP for the year of 2018. These factors and the year can be changed through the filter which is in the form of a drop-down menu.

Furthermore, when toggling between the factors and the year, the countries in the x-axis of the chart would be sorted by overall happiness rank and not by the factor that is selected. This is essential as it helps identify exactly what sets happier countries apart from least happier countries, as per UN's standards. For instance, users can identify what the trend is like for each of the factors as the Happiness scores reduces. Nevertheless, there is also a sort function that allows the user to sort the countries in descending values of the selected factor. Through this, it can be seen exactly what countries rank higher in the selected factors. When clicking the sort button, a animation occurs which shows the bars moving until all the bars are sorted in descending order. This animation is used to bring attention towards the sorting of figures and to highlight how significant the changes are as countries transition when being sorted from their overall happiness scores to their factor rankings [6].

Lastly, the tooltip is another important feature of this visualization. Upon hovering over the different bars, the name of the country, the overall happiness rank as well as the individual country's value for the selected factor are presented.

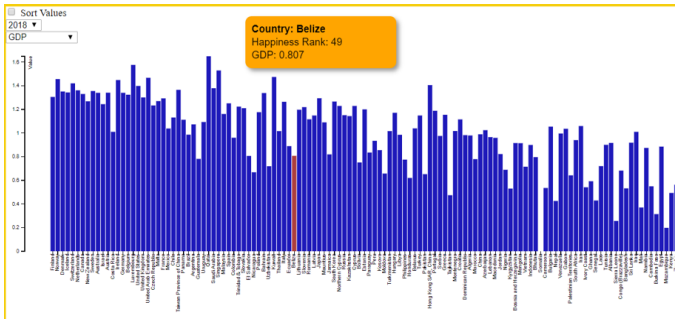

Fig. 7: Bar Chart from Application

Radar Chart

The fourth component to be included is the radar chart (Fig.8) that allows users to compare the 6 factors for selected countries. The overall design of the radar chart obeys proper design principle of radar charts as it is kept simple such that it "be seen and understood at a glance" [5].

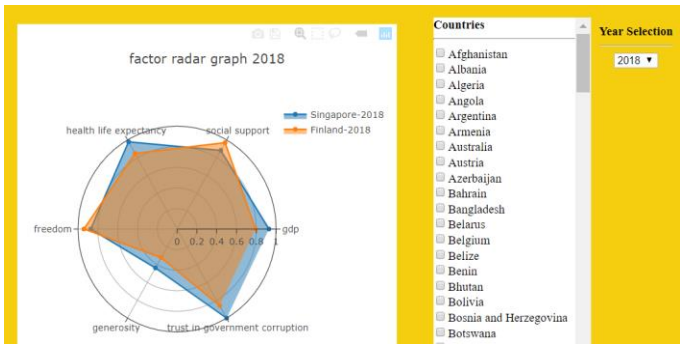

Fig. 8: Radar Chart from Application

The main features in this component include filters which allows users to select the countries and years they are interested in making comparisons with. The countries in the filter are sorted by alphabetical order so that it is easy to locate countries of interest.

Also, as this visualization allows filtering by both country and year, users can compare not only among countries but also among years. For example, the user could identify difference in ranking of all 6 factors for perhaps, India, between years 2017 and 2018.

Boxplot

The fifth and final component is the box plot where distribution of countries among the different 6 factors can be seen. The box-plot displays the minimum, average and max values of each factor by year. All countries are displayed beside each box according to their position in regards with the specific factor. The design which combines the boxplot as well as a scatterplot allows viewers to pay attention to even the smallest detail (Fig.9). Furthermore, it allows the user to compare the value of their point of interest (country) with minimum and maximum values specific to the chosen factor. For

example, viewer can compare GDP value of their point of interest (Guinea) with min and max GDP values in 2018(Fig.9).

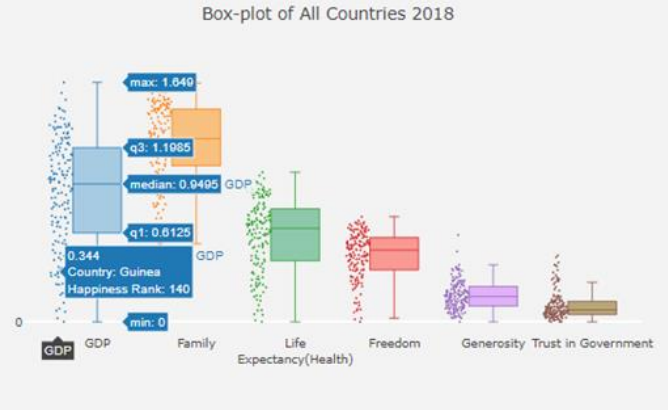

Fig.9: Box Plot highlight Guinea from Application

5. FINDINGS

Our analysis using the 5 different visualization components explained in section 4 revealed a few notable trends.

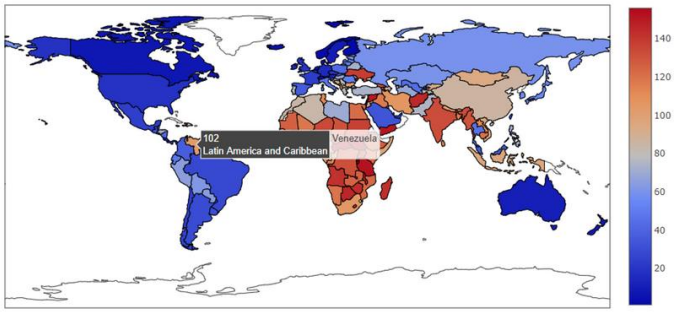

Fig.10: Choropleth map for 2018 from the application with outlier highlighted

Finding 1:

Through the general view presented by the choropleth map in tab 1, obvious distinctive patterns can be observed with regards to global happiness. As per Fig.10, it is noted that in 2018, the Sub-Saharan African region, which consists of countries like Nigeria and Zimbabwe, are shown to be the least-happiest region among all regions. This is identified by the clear distinctive red area seen on the map (Fig.10). Furthermore, a striking outlier can be observed for the Latin America and Caribbean region whereby Venezuela is shown to be the only region with low happiness scores (in red colour) while all other neighbouring countries are shown to be much happier as denoted by darker hue intensities of blue. This is further analysed by the line chart (Fig.11) where it can be observed that Venezuela happiness rank keeps dropping throughout the year.

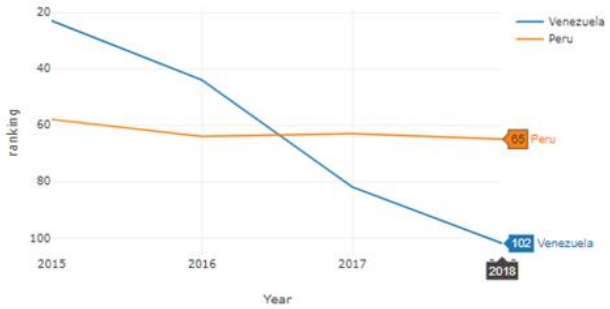

Fig.11: Line chart showing Venezuela's overall happiness rank

To see the reason behind, the radar graph (Fig.12) is used, and it is shown that 'freedom to make life decisions' factor of the country has significantly decreased over the year. This could be possibly attributed to the country's volatile political situation it is facing in current times [10].

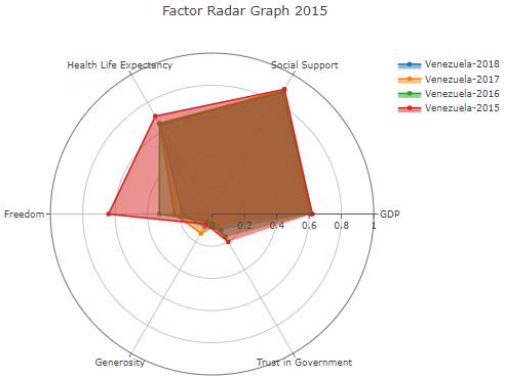

Fig.12: Radar Chart comparing Venezuela from 2015 to 2018

Finding 2:

The third tab which features the interactive bar chart enables the user to get an overview of the happiness scores of different countries in relation to the 6 factors that it is comprised of. It could be assumed that as overall happiness scores decrease, values of the 6 factors will reduce as well. That is exactly what we've found as in Figure 8, it is seen there is a downward sloping trend in GDP for 2018 as happiness scores decreases (Fig.13).

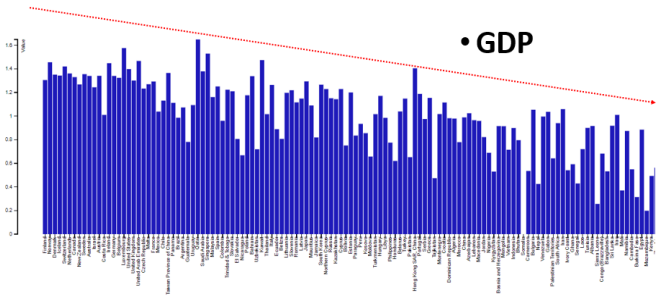

Fig. 13: Bar Chart for GDP in 2018 (X-axis sorted from most happy to least happy country)

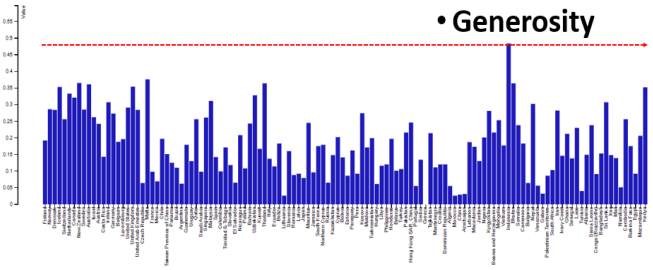

Fig. 14: Bar Chart for Generosity in 2018 (X-axis sorted from most happy to least happy country)

Nevertheless, when looking at the factor of Generosity (Fig.14), the distribution graph does not seem to follow a downward trend. In fact, there are several countries that are considered less happy but have exceedingly high values of Generosity as compared to the happier countries. For instance, Myanmar that is ranked 130th out of 156 countries for happiness ranking has the highest value of Generosity. Thus, this shows that there is very low collinearity between happiness and generosity among countries. Moreover, it could be showing that GDP plays a bigger role in UN's model of calculating the happiness score.

These results are further complemented by the boxplot (Fig.15) which shows GDP and social support to have higher average scores. These 2 factors can be deduced to be the main components in

calculating the overall happiness score since the overall score is calculated by adding different components together, higher averages in each factor would indicate higher impact on the overall score.

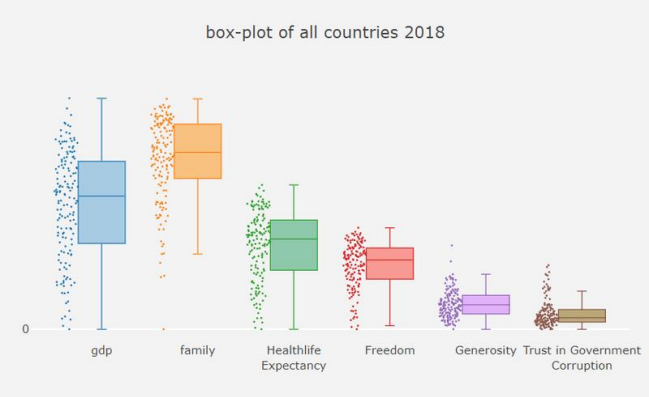

Fig.15: Boxplot showing the distribution of the 6 factors for 2018

Finding 3

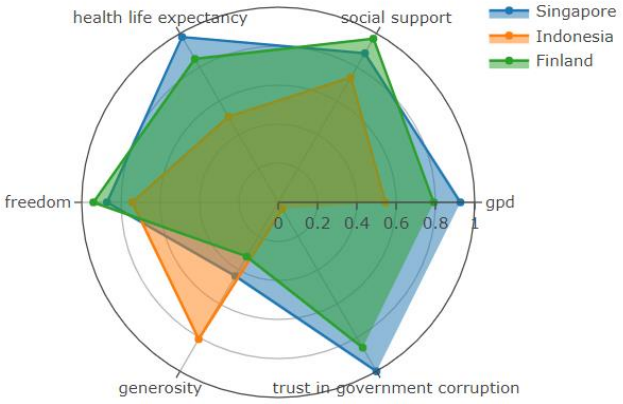

Fig.16: Radar Chart comparing Singapore, Indonesia and Finland

To help understand the findings from the radar graph, Singapore and Indonesia are compared against the "happiest" country of 2018 - Finland. Though major differences can be seen between Indonesia and both Singapore and Finland, the differences between Singapore and Finland are very subtle (Fig.16). More specifically, Singapore's score in terms of healthy-life expectancy, trust in government, generosity and GDP are higher while Finland is higher in only social support and freedom in making life decisions. These shows that for Singapore to be truly happy as per UN's standards, it must improve on both these factors.

6. Conclusion

Our web application offers various types of visualizations to gain an in-depth analysis on the World Happiness reports. The results obtained from this could be crucial as it can help understand what attributes specifically make a happier country.

Nevertheless, the application still has a lot of room to be enhanced and scaled further. For one, the visualizations could be possibly extended to include more factors that may not be considered by UN but by other organisations. For example, environmental conditions of countries have been raised as a possible crucial factor in determining happiness and thus, it could also be considered as for further studies [8]. The World Bank, as well, has a plethora of information that could be used. Therefore, including new measures as such could reveal further interesting insights about countries' happiness.

Acknowledgments

The authors wish to thank Professor Kam Tin Seong for assisting in this project and offering valuable guidance.

References

[1] World Happiness Report. (2018). Retrieved from https://s3.amazonaws.com/happiness-report/2018/WHR_web.pdf

[2] ColorBrewer.org Colorbrewer: Color Advice for Maps. Available at http://colorbrewer2.org/. Accessed 29th October 2018

[3] Choropleth Maps. (2017). Available at https://www.axismaps.com/guide/univariate/choropleth/. Accessed 29th October 2018

[4] Few, S. (2011). The Chart Junk Debate https://www.perceptualedge.com/articles/visual_business_intelligence/the_chartjunk_debate.pdf. Accessed 15th October 2018

[5] Few, S. (2006). An Introduction to Visual Multivariate Analysis. Retrieved from http://www.perceptualedge.com/articles/b-eye/visual_multivariate_analysis.pdf. Accessed 3rd November 2018

[6] Heer, J. & Robertson, G, G. (2007). Animated Transitions in Statistical Data Graphics. Retrieved from http://vis.berkeley.edu/papers/animated_transitions/2007-AnimatedTransitions-InfoVis.pdf

[7] Wickham, H. (2014). Tidy Data. Retrieved from https://vita.had.co.nz/papers/tidy-data.pdf. Accessed 29th October 2018

[8] Bonini, Astra N. 2008. "Cross-National Variation in Individual Life Satisfaction: Effects of National Wealth, Human Development, and Environmental Conditions". Social Indicators Research, 87(2): 223-236.

[9] R.J. Vidmar, "On the Use of Atmospheric Plasmas as Electromagnetic Reflectors," *IEEE Trans. Plasma Science*, vol. 21, no. 3, pp. 876-880, available at http://www.halcyon.com/pub/journals/21ps03-vidmar, Aug. 1992. (URL for Transaction, journal, or magazine)

[10] Al Jazeera. (2018). Venezuela's crisis explained from the beginning. Retrieved November 10, 2018, from https://www.aljazeera.com/indepth/features/2017/04/venezuela-happening-170412114045595.html