

Elizabeth Earl

DSC680-T301

Project 3 – Happiness

Business Problem:

It is difficult to determine what constitutes as happiness in every individual. But if happiness is taken from a larger scale (a whole country) we can take these finding to see how people react (or do not react to their surroundings). Knowing how people react to situations such as changes in politics can help potential candidates target group happiness and make their way to the top of candidacy.

Background/History:

Exploring the correlation between people's happiness and politics can say a lot about a whole country. If people from a country tend to react positively to politicians will the country be more likely to be successful? Defining success will come from this idea that a country is a first world country.

It is expected that good politics will result in overall people's happiness but how that is reflected in how long people live will also be explored. If people are happier in one country, are they more likely to live longer versus those who reside in a country with less happiness?

Data Explanation (Data Prep/Data Dictionary/etc):

Dataset – Happiness by year:

The following variables were taken from the most recent happiness data dating in 2019

- Overall rank

- Used to rank a county/region on where they fall versus other countries/regions
- Country or region
- Healthy life expectancy
 - Rating of how much a person is expected to live a healthy life
- Freedom to make life choices
 - Rating of how much people per country/region feel they have freedom
- Perceptions of corruption
 - In this variable, the lower the better, as it is an indication of how corrupt said country/region is

Methods:

For this project I will rely on comparison charts. The visualization plots used in matplotlib as well as seaborn will provide a more detailed look on how countries/regions compare against one another. I will also be using K-Means methods to further find trends between variables I deemed essential:

- Healthy life expectancy
- Freedom to make life choices
- Perceptions of corruption

Analysis:

For my initial analysis I created a correlation matrix (as shown in Figure 1.) to get a better understanding of what major/impacting variables I wanted to continue using throughout my project. The variables I decided on were those I felt I would personally associate with happiness (or lack of happiness). These variables were

'Healthy life expectancy', 'Freedom to make life choices' and 'Perceptions of corruption.' Because these variables results with numbers where similar numbers indicated opposite results it was hard to rely on the correlation matrix.

To further investigation I created visualizations using seaborn to give me a better understanding of how different countries ranked when it came down to the chosen variables. Upon further investigation of these visuals, I found only 3 of the countries I deemed as top countries appeared in the charts. There was little to no (only 1 country to be exact) overlapping in countries in the positive charts (figure 2a and figure 2b).

Conclusion:

Based on my findings in Figure 2c and taking the included countries versus those in figures 2a and 2b I concluded there was no direct correlation between a country's level of corruption in politics and overall inhabitant's happiness. One of the countries that stood out to me (using figures 2a and 2b) was Canada; Canada appeared in both graphs. Canada is known for its socially accepted politicians, and I made this assumption that perhaps as it was ranking high in both health life expectancy and freedom to make life choices that this correlated to Canada having a less corrupt perception of politics. To my surprise Canada was not in the list of least perceived corrupt countries.

Because there was no direct mention of Canada, I used this to conclude politics did not fully influence happiness. Nonetheless I kept on with this idea that perhaps in previous years there could have been a correlation. I failed to take into consideration that some countries/regions might not have a system in play that acts as a political party and thus there is no way to be corrupt. This led to more

challenges as it was difficult to understand factors that played into each individual variable and thus influencing results.

Assumptions:

The main assumption that was made throughout my project was based on what makes a country great. I assumed who was a top country based on my own assumption and which countries to be perceived as 'popular' according to the media. This assumption also came from the idea of where people label dreams being achievable. I assumed top countries were as follow:

- United States
- United Kingdom
- Canada
- Spain
- France

Limitations:

The limitations in this project came from the data itself, I believe that there was not enough background on certain variables (corruption) to know what to blame exactly. Political corruption can come in many forms, and it can be categorized differently depending on what political stand a person has. I believe it was a variable that provided a lot of information overall, but it was difficult to exactly know if the corruption was universal or personal corruption that was instilled by an unpopular opinion.

Challenges:

As mentioned under the limitations one of the major challenges was focusing too much on where these indicators of happiness were coming from. I believe it is difficult to simply rely on these variables as each variable can have so many sub variables within them. Because the dataset contains only so much information is it difficult to fully understand where reasonings are coming from. A lot of assumptions are therefore made on why certain countries/regions rank higher/lower than others; there is no clear indication of either side.

Future Uses/Additional Applications:

Going forward it would be interesting to dive deeper into how each variable could be interpreted. This would require a more in-depth survey of the same people with variables that include each question asked. In addition, it would be beneficial to create a variable of political stance of those participating in the survey versus the country's overall political party (if any apply).

Implementation Plan:

Q&A:

1. What countries are the happiest?
2. What countries are the least happiness?
3. How do top/lowest ranking countries compare regarding 'top' countries?
4. How do previous years over all compare to the most recent 2019?
5. How do variables compare against one another by year?
6. Comparing years and same variables, are the trends similar throughout all years?

7. Is there an equal balance of variables by country/region?
8. does healthy life expectancy correlate to any other variables?
9. What is the happiest country/region official political party?
10. Are the happier countries more likely to be 'nice' countries?

Supporting illustrations:

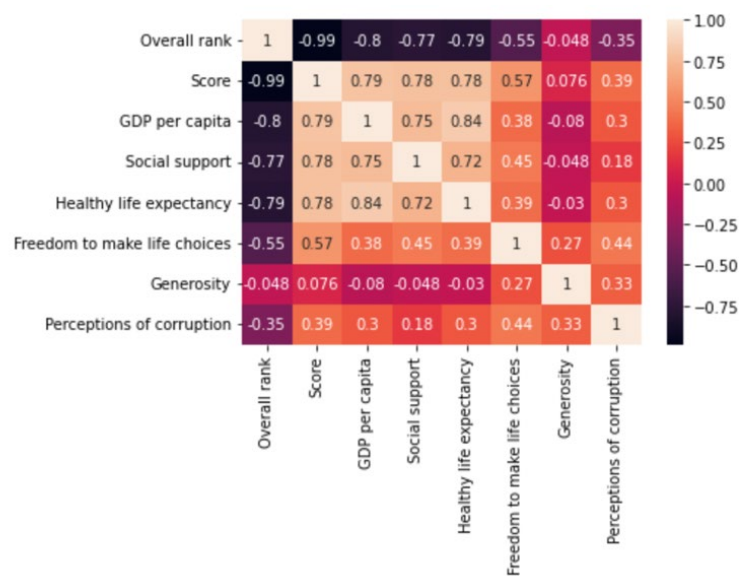


Figure 1:

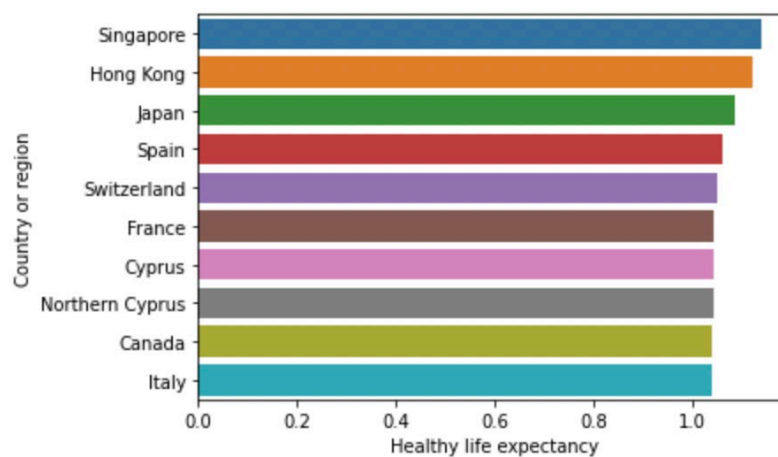


Figure 2a:

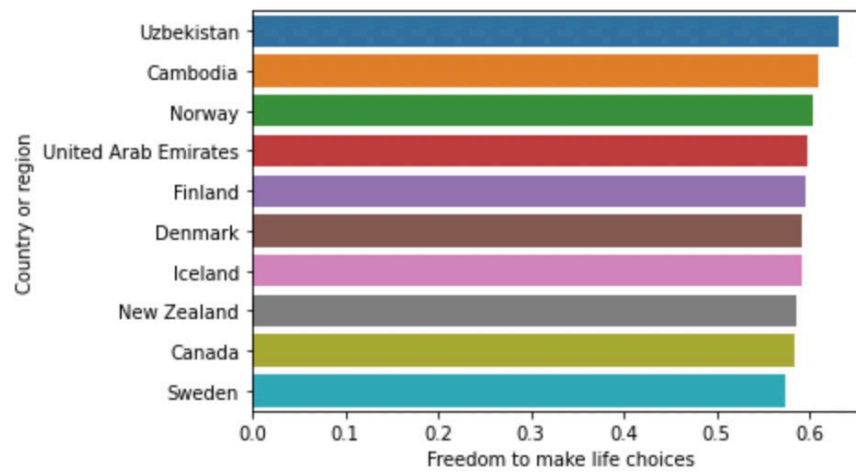


Figure 2b:

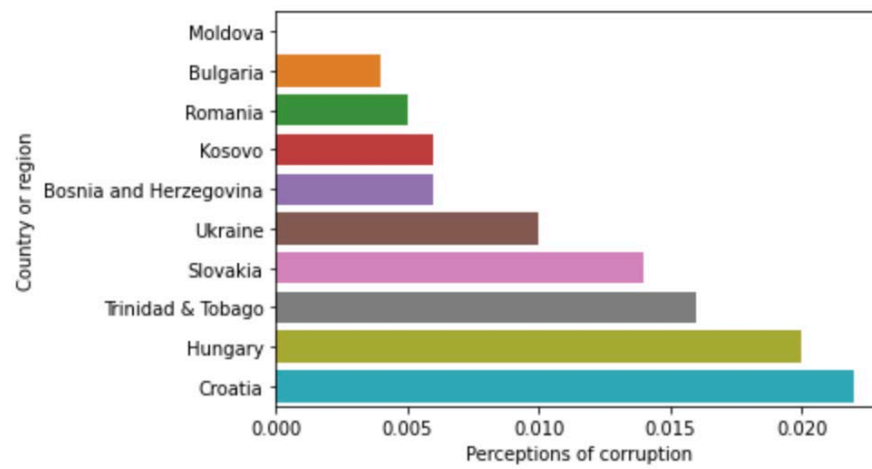


Figure 2c:

Appendix:

World Happiness Report. (2022). Retrieved 18 May 2022, from

<https://www.kaggle.com/datasets/unsdsn/world-happiness>