GDELT

We found the Global Database of Events, Language and Tone(GDELT), which is an open source database available through Amazon Web Services that tracks events which impact the stability of every country around the world. These events are placed on a scale from -10 to +10 that quantifies how intense the impact could have to help or hurt the stability of the country. The tone is scored on a -100 to +100 scale from extremely negative to extremely positive [8]. Since each media mention is also recorded individually, the number of mentions could also lead to insight into how important the event was to the population. The expectation is an event that triggers increased media mentions along with a strong tone will be the events with the highest impact.

The yearly average Goldstein score and tone score of each country will be used to evaluate the connection between world happiness, the Goldstein score and the tone of the media. The overall average Goldstein score is approximately 0.898, which indicates that the world as a whole is generally stable as well as the standard deviation of 1.02, showing a low variance in the stability of different countries. The average tone of the media score is -2.13 indicating that despite the strong sense of stability from the Goldstein score, the media tends to portray these events negatively. The standard deviation is 1.44, indicating there is a larger discrepancy between the portrayal of world events than compared to the Goldstein scale. So even though there is a statistically significant trend between the Goldstein scale and world happiness scores, very little of the variation in world happiness is explained by the Goldstein scale.

Data Information

We also found several datasets on Kaggle for analyzing the factors that influence world happiness. The happiness score and ranking used data from Gallup World Poll, which is the poll that tracks the most important issues worldwide, such as food access, employment, leadership performance, and well-being.

The first world happiness report was published in 2012, and the happiness report from 2015-2020 were available for analysis. Each data table shows the ranking of the happiness score of each country and indicates those countries’ GDP per capita, social support, freedom score, and some other features that may influence people’s happiness in a single year.

The scores for features are calculated based on the poll that asked each respondent to select a level from 0-10 as the score that they estimate their happiness to be and give a score on economic production, freedom, and so on… for how much they think this factor contributes to their happiness. All of the data in each category is numerical.

Data cleaning

In order to compare the happiness score and the factors for different years, we have to make the variables consistent. First, we make sure that all the country names are written in the same style so it is easy for selecting and filtering in further steps, and we delete some features and countries that were not measured every year. After deleting the unnecessary columns, six factors remained that will be used for further analysis, which are “GDP”, “Health”, “Social support”, “Freedom”, “Generosity”, and “Corruption”. Adding the basic variables like “country name”, we have 9 columns and 146 rows for each table.

Does covid-19 influence on world happiness index?

In our hypothesis, covid-19 will slow down the increase of GDP or even make a negative impact for most of the countries, and it will also cause the decrease in the measurement in health, freedom, and social support. For testing this hypothesis, we made a linear graph based on the world happiness and social factors scores during 2015-2020 to find out the trends of world happiness index and the factors. In the observation, all of the measurement scores do not have any remarkable changes during those years. This result may be caused by some measurements like the government’s subsidies and tax reduction policy which make people’s lives kept stabilized in 2020.

Then, we want to have a look that does covid-19 influence people’s happiness in those countries that are most influenced by covid in 2020, like China, United State, and India. And we still not found any remarkable differences. The only thing that mentionable is that the happiness score ranking of most of those countries have gotten lower through 6 years.