Team: Scrambled Legs

1/12/19

Rutgers Data Science Bootcamp

World Happiness & its Correlates: Analysis and Findings

Why study happiness? As a team, we collectively felt that this topic was rather intriguing and wanted to learn more about the factors that affects a country’s level of happiness, as a whole. This is important because governments, organizations and civil society are increasingly using happiness indicators to inform their policy-making decisions. While pondering on where to begin, we stumbled upon the World Happiness reports that were published in 2015, 2016, and 2017. Six factors were examined each year: economic production, social support, life expectancy, freedom, perception of corruption, and generosity. While collaborating, we felt that weather and religion may have some correlation with ones’ perceived evaluation of life. We used datasets from Gallup Poll (2017) for the religion data and the World Bank Weather Data for trends in weather (1961-1999).

After importing, reading, and cleaning the data, we were able to clearly observe the trends and correlations. In our initial proposal, we aimed to rank the countries in terms of the highest and lowest happiness scores. For the countries that ranked the highest in terms of happiness, we saw that the contribution of the six factors seem to be more consistent for each of the “happiest countries.” We also observed that each of the “happiest countries” scored at least a seven. Those countries are: 1.) Finland 2.) Norway 3.) Denmark 4.) Iceland and 5.) Switzerland. For the lowest ranked countries, we did the same graph and observed the countries were: 1.) Lesotho 2.) Guinea 3.) Togo 4.) Ukraine and 5.) Sudan. We also decided to rank the happiness indicators in how much they contribute to the average happiness score. The factors that affected the ranking the most are shown from highest to lowest: Social Support, GDP per capita, healthy life expectancy at birth, freedom to make life choices, generosity and then perceptions of corruption. All in all, we can conclude that social support had the highest relative contribution to explaining the average life evaluation score and perception of corruption had the lowest relative contribution to explaining the average life evaluation score.

The next topic we honed in on was the correlation between temperature and life evaluation scores. We used the life evaluation scores because the happiness scores are actually based on the life evaluation scores, which indicate how much the respondents value their lives on a scale from 0 to 10. In regards to the correlation between temperature and life correlation scores, we created a scatterplot and a hexbin to see the relationship between annual temperature and life evaluation scores. There is a moderate correlation (-0.45) between annual temperature and life evaluation scores. As annual temperature increases, life evaluation scores seem to decrease. The hexbin allows us to more easily see the trend of the data. The lighter the color, the higher density of data points in that area. As for the correlation between precipitation and life evaluation scores, we used the same process. Except, for these graphs, we obtained different results. There is a very weak correlation (0.04) between annual precipitation and life evaluation scores. There is a lot of variation between life evaluation scores for most levels of precipitation.

Lastly, we focused on how the varying levels of religiousness and how having specific religious beliefs correlated with life evaluation scores. To do this, we did a Pearson correlation analysis and ranked the correlation coefficients for each category: Religious, not religious, confirmed atheist, and individuals who were non-responsive or simply undecided. We found that the people who were undecided or unresponsive in regards to religion had the highest positive correlation with life evaluation, though we found it to be a moderation correlation. People who were religious had the highest negative correlation with life evaluation. This is also a moderate correlation. The confirmed atheists and non-religious people had positive correlations with life evaluation, but they are weaker. The next factor we wanted to focus on was specific religious beliefs and how they correlate with life evaluation scores. We also did the Pearson correlation analysis for this and found that all specific beliefs were negatively correlated with life evaluation scores. Specifically, people who believe in life after death had the highest negative correlation with life evaluation scores. It is a moderately strong correlation because it is under -.5 benchmark. The other categories are moderate correlations. This data seems to be consistent with the findings that religiousness is negatively correlated with life evaluation scores.

All in all, we understand that our findings are based on certain limitations. First and foremost, correlation does not imply causation. Even though one variable may correlate with another, there could be an unknown confounding variable contributing to the relationship between the two variables. Second, the analyses are subject to volunteer bias. The people who responded to the surveys may be different from the general populations of the corresponding countries. They may respond to the surveys in a more biased way. Lastly, we used old datasets regarding weather. However, it was the only source we found that listed the average annual temperature for whole countries.