Conclusions

Project Goal

This project is just a part of the learning project that serves the task to illustrate how the creator can apply the tools of Data Analysis and interpret their results. The author did not pursue a goal of making a thorough and detailed analysis; it is just the first approach to the problem.

Problem

The objective of the project is to find is there any correlation between the number of COVID-19 cases and the Happiness/Wealth Index of the population. We suppose that there is a reverse relationship between the number of COVID-19 cases and the Happiness/Wealth Index of the population (i.e. the wealthier country is the less number of COVID-19 cases should be found).

Target Audience of this project

Anybody who is interested in the problem and in Data Analysis in general.

Data

To achieve our goal, we decided to use the following data:

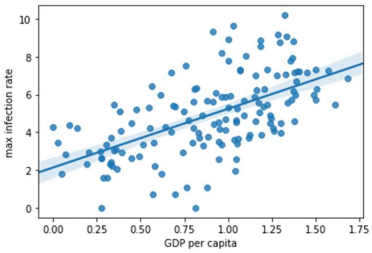
- Covid-19 Confirmed Number of cases
- Worldwide Happiness Report

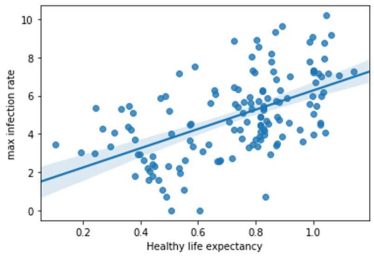
Methodology

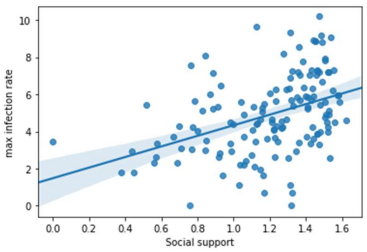
- Data will be collected, cleaned, and processed into a dataframe.
- Two datasets will be join together for further analysis.
- Correlation Matrix will be built for the dataset.
- Finally, the data be will be visually assessed using graphing from Python libraries.

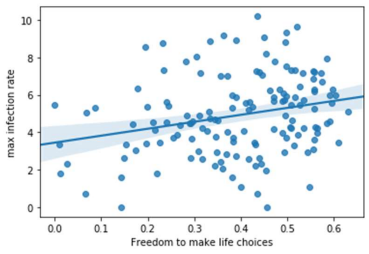
Results

The results of our analysis showed below:









Discussion Section

Based on the results of our analysis, we should come to a paradoxical conclusion that the wealthier country is, the more chances is to be infected by COVID-19. That is opposite to our initial guess. However, some limitations of our research make the above-mentioned conclusion wrong.

Limitations and suggestions for future researches

All of the above analysis depends on the accuracy of the Data we are using. And here is the biggest problem of our research, the poorest countries do not have enough statistical data to provide us the correct and trustworthy information. For better results, future research and analysis could consider using the number of excess deaths or applying a special Index of Trustworthy to the data of each country.

Conclusions

This project illustrates how important it is to use only sufficient data you can trust. Otherwise, even the statistical tools and application of all needed methods can lead to totally wrong results. When there is no possibility to collect relevant and trustworthy data, it is important to use additional data or apply indices that can help the data scientist to eliminate possible errors.