**Data Science Blog Post - Covid-19, Time To Go Out?**

**Table of Contents**

* [Installation](http://localhost:8888/notebooks/OneDrive/Data%20Science/covid/final/ReadMe.ipynb#installation)
* [Project Motivation](http://localhost:8888/notebooks/OneDrive/Data%20Science/covid/final/ReadMe.ipynb#motivation)
* [File Descriptions](http://localhost:8888/notebooks/OneDrive/Data%20Science/covid/final/ReadMe.ipynb#files)
* [Results](http://localhost:8888/notebooks/OneDrive/Data%20Science/covid/final/ReadMe.ipynb#results)
* [Deploy](http://localhost:8888/notebooks/OneDrive/Data%20Science/covid/final/ReadMe.ipynb#deploy)
* [Comments](http://localhost:8888/notebooks/OneDrive/Data%20Science/covid/final/ReadMe.ipynb#comments)
* [Licensing, Authors, and Acknowledgements](http://localhost:8888/notebooks/OneDrive/Data%20Science/covid/final/ReadMe.ipynb#licensing)

**Installation**

The code runs with Python version 3.6.3. There are a large number of libraries required to run this notebook. These can be found near the top of notebook in the # import libraries cell.

**Project Motivation**

It can be difficult to get accurate and time sensitive information on how countries are doing at present during the SARS-2 pandemic (COVID-19 is the illness, the pandemic is the virus). Exploring factors potentially related to outcomes is important as a means to consider further outcome prediction. Key factors are likely related to government and public health leadership and the potential for actual public adoption of mitigation strategies for the virus. In this post, I answer 4 questions that should provide some insight into this very important issue.

* Q1: Who are the best and worst countries dealing with Covid-19?
* Q2: Which countries have flattened or are flattening the curve?
* Q3: Can I see a global geographic representation of infections?
* Q4: What is the projected global mortality by December 31, 2020?

**File Descriptions**

The following data files (available in this repository) are required.

* I am using data from John Hopkins University. There is a GitHub repository that holds this data. To refresh it just go to <https://github.com/CSSEGISandData/COVID-19> and hit the Clone of download button. It will ask to launch GitHub Desktop and will update it automatically.
* The data for median income and population comes from <https://worldpopulationreview.com/countries/median-income-by-country/>
* World Bank Country Codes <https://wits.worldbank.org/wits/wits/witshelp/content/codes/country_codes.htm>
* Transparency International Dataset <https://www.transparency.org/cpi2019>
* This is the source for the geopandas choropleth map <https://towardsdatascience.com/a-complete-guide-to-an-interactive-geographical-map-using-python-f4c5197e23e0>

**Results**

The interpretation of the results are published in this blog [post](https://medium.com/@mlataibrahim/how-data-people-are-using-stackoverflow-f8799f8e0e9e)

**Deploy**

Update the John Hopkins, Covid-19 data <https://github.com/CSSEGISandData/COVID-19> by refreshing your local copy of the GitHub repository. Please point this notebook to your local repository of that data. For your convenience this GitHub repository does include the data frozen as of a certain date. The ARIMA model has been hyper parameter tuned, stored on disk and is ready to go. There is a forecasting module that can be ran at any time. It remains to be seen whether or not its predictions will be accurate.

All of the code and data required to run this notebook is in included in the GitHub repository.

**Comments**

There are a large number of functions in this notebook. Usually I would put those in a helper.py file. However, the specific instructions were to have all code in a single Jupyter Notebook.

**Licensing, Authors, Acknowledgements**

The John Hopkins data is available under the Creative Commons License. For the particulars of that license please go to <https://data.humdata.org/event/covid-19> and search for Creative Commons.

The ARIMA model code was largely taken from this book which I bought "Time Series Forecasting With Python" By Jason Brownlee of Machine Learning Mastery. <https://machinelearningmastery.com/make-sample-forecasts-arima-python/> The code was updated to reflect the unique requirements of this covid analysis. Jason provides code with this book that he explicitly tells you to use and make your own.

My thanks to all of the contributors of the data to this project.