National COVID Data Rubric

DS 4002 – Spring 2024 – Ramya Tangirala
Due: 5/7/2024
Submission format:

- **Upload link to github repo to canvas**

Individual Assignment

Why am I doing this? This case study looks into the National COVID Data between 2020 and early 2024. As coronavirus is a big topic recently, this is an important case study to review. This case study will use time series data to analyze trends with the number of weekly hospitalizations and weekly deaths.

What am I going to do? You will look at the following repository to get the dataset containing data on this topic. You will look at the dataset and create an ARIMA model to predict the weekly hospitalizations and weekly deaths.

You will turn in:

- Exploratory data analysis graph
- Data dictionary containing all the variables
- All the code used to produce your results
- Results such as graphs exported as a pdf
- Github repository containing the data set, code, and results

Tips for success:

- Start early to ensure you have time to solve any unforeseen problems
- Do the exploratory data analysis thoroughly to ensure that you understand the topic well
- Look into resources for time series forecasting to better understand it

How will I know I have Succeeded? You will meet expectations when you follow the criteria in the rubric below.

Formatting	● Repository – A GitHub repo containing all materials o Contents ▪ Folder containing all the code used ▪ Folder containing all the data ▪ Folder containing all the results ● Results should be in a pdf file in the folder ▪ README.md ▪ LICENSE.md ▪ Folder containing the references used

	○ Use pdf format when possible ○ For code and data products use the appropriate format for whatever it is
README.md	● A brief description about the project ● Includes the research question and hypothesis ● Write a general description for the procedure used to obtain results
Code	● Two scripts for the exploratory data analysis and the ARIMA model ● Code should be well commented with clear sections ● Code should be done in Python but can be written in an application of your choice
Results	● Two files should be in the results folder ○ Exploratory data analysis results ○ ARIMA model results ● Both files should be in pdf format and figures should be labeled properly

Acknowledgements: Special thanks to Professor Javier Rasero and Harsh Anand for this rubric structure