Investigating Vaccine Efficacy Through COVID-19 Mortality Trends

Task	Criteria
Formatting	- Final repository: github repo containing all materials - Submit a link to the repository - Contents is present and correctly organized - Contents of repo - README.md - LICENSE.md - DATA folder - SCRIPTS folder - OUTPUT folder - Results
README.md file	- Goal: Orient reader with the contents of the github repository - Software and Platform: - Explain software and packages used for analysis. In this case it is python, refer to README file in provided repo for specific packages used - Documentation Map: - Explain contents within DATA, SCRIPTS, and OUTPUT folders - Reproducing Results: - Explain how results can be replicated using data within repo - References: - Include references used to complete analysis
LICENSE.md file	- Goal: Explain to a visitor the terms under which they can use and cite your repository - Use the MIT license
DATA folder	- Goal: Provide original and cleaned data necessary for analysis - Vaccination-data-CDC.csv - Daily vaccine administration data in US, reported by CDC - COVID-19_Death_Counts.csv - Weekly data on deaths involving Covid-19 in US, reported by NCHS - Cleaned Covid-19 deaths data set (csv format) - Cleaned vaccination rate data set (csv format) - Merged data set (csv format)
SCRIPTS folder	Goal: provide code with commentary as necessary. File names should align with the content of the code.

Investigating Vaccine Efficacy Through COVID-19 Mortality Trends

	- Cleaning Covid Death Data: Contains steps taken to clean original NCHS data to prepare for merging, as well as EDA.
	   - Refer to the 'Cleaning Covid Death Data' file located in the materials folder
	- Cleaning Vaccine Data: Contains steps taken to clean original CDC data to prepare for merging, as well as EDA.
	   - Refer to the 'Cleaning vaccine data' file located in the materials folder
	- Combining Data: Contains steps taken to combine cleaned data for analysis.
	   - Refer to the 'Combining data' file located in the materials folder
	- Final Data Analysis: Contains final analysis.
	   - Identify time-lagged correlations using a cross-correlation function (CCF).
	   - Identify relationship strength using Pearson's correlation
	      - Ensure p-value < 0.05 so correlation is statistically significant.
	   - Conduct Granger causality to determine whether vaccine rates can predict future mortality rates.
	   - Refer to the 'Final Data Analysis' file located in the materials folder
	- Master Script: Single code file integrating all steps (cleaning, merging, and analysis) to ensure reproducibility.
	   - Refer to "Master Script" file located in the materials folder

OUTPUT folder	- Goal: provide appropriate visualizations of data and outputs from analysis - Hint: focus on relevant trends in vaccination rates and Covid-19 deaths overtime - Examples included in materials folder - Include CCF graph

Results	- Goal: Clearly communicate analysis goals and findings in single page
	- Correctly interpret results from
	   - CCF output
	   - Pearson correlation output
	   - Granger causality output