Case Study Rubric – COVID-19 Cases and Influenza Vaccine Rates Across Virginia

DS 4002 – Fall 2024 - Instructors: Michelle Kim
Submission format: Upload link to Github repo to Canvas

Individual Assignment

General Description: Submit to Canvas a link to your case study repository and pdf of presentation

Preparatory Assignments – Everything in this course, but CS1 and CS2 should help you better understand how this case study should be completed.

Why am I doing this? This is an opportunity to use what you've learned and explore the relationship between COVID-19 cases and influenza vaccination rates across Virginia. This case study connects data science and public health, aiming to deepen your understanding of ARIMA modeling. Through this work, you will see how data-driven approaches can highlight relationships and meaningful insights.

- Course Learning Objective: Prepare findings for presentation to your peers.

What am I going to do? To begin, go to the Github link here and read the rubric and hook documents: https://github.com/michellehkim280/DS4002_CaseStudy . Use the dataset, Vaccination_Cases.csv and to conduct EDA and ARIMA modeling, while analyzing the relationship between influenza vaccination rates with COVID-19 cases in Virginia. Then, you will create a model to forecast influenza vaccination rates. Finally, compile your results into a presentation. Your deliverable should include:

- Github Repository - containing scripts, data, and output
- Presentation PDF - containing reflection of the case study

Tips for success:

- Do your research: Get familiar with both Python and R. Teach yourself to understand ARIMA.
- Ask for help: Talk to the instructors or other students. It can save a lot of time and energy!

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

Formatting	● Repository – a Gitbhub repository containing all materials ○ To ensure **reproducibility**, the repository will adapt parts of the TIER Protocol 4.0. In a nutshell, the top level page of the repository should contain: ■ A README.md file (which auto displays) ■ A LICENSE.md file (use MIT as default) ■ A SCRIPTS folder ■ A DATA folder ■ AN OUTPUT folder ● Presentation ○ About 7 slides ○ PDF format for submission to collab ○ Generate the slides through the program of your choice
Github Repository	● Goal: Showcase all findings and contents of analysis ● Contents: ○ README.md: discusses project overview, background research, links, etc ■ Section 1: Software and platform section: The type(s) of software you used for the project. ■ Section 2: A Map of your documentation ■ Section 3: Instructions for reproducing your results. ○ DATA: this folder includes original source data ○ SCRIPTS: This folder contains all the source code for your project. ○ OUTPUT: This folder contains all of the output generated by your project, e.g. figures, tables, etc. ○ LICENSE: This file explains to a visitor the terms under which they may use and cite your repository.
Presentation	● Goal: Create a visual slideshow of findings that is concise and easy to understand. ● Presentation ○ Order ○ Title & Outline ○ Motivation/Context/Hypothesis/Research Question/Modeling Approach/Goal/Etc.

	○ Data Explanation/Acquisition ○ Analysis Plan ○ Tricky Analysis Decision ○ Bias and Uncertainty Validation ○ Results/Conclusions ○ Next Steps ○ References/Resources/Acknowledgements ○ Closing Slide
References	● All references should be listed at the end of the document ● Use IEEE Documentation style (link)

Acknowledgements: Special thanks to Professor Alonzi for providing this rubric structure and Jess Taggart from UVA CTE for coaching on making this rubric. This structure is pulled from Streifer & Palmer (2020).