

CS3 Rubric – Healthcare Trend Analysis

DS 4002 – Fall 2024 - Instructor: Loreto Alonzi

Due: TBD

Submission format: Upload link to GitHub repo to Canvas

Individual Assignment

General Description: Submit to canvas a link to your case study repository.

Why am I doing this? This case study provides an opportunity to apply your data science expertise by utilizing trend analysis techniques to examine healthcare consumption patterns in the U.S. Throughout this assignment, you will gain insights into how data analysis can be employed in real-world scenarios, with significant implications for governance and the healthcare sector.

What am I going to do? The GitHub repository for this case study can be found at <https://github.com/Sabrina-Hendricks/DS4002-CS3>. You will use the healthcare subsection of the “National Data - National Income and Product Accounts” dataset from the Bureau of Economic Analysis to conduct trend analysis on healthcare consumption across different service types. The GitHub repository contains the original dataset, as well as cleaned healthcare subset dataset, for you to conduct analysis. You will run the trend analysis scripts using moving averages to draw conclusions and observations based on the graphs generated. Deliverables include:

- Data folder
- Scripts folder
- Written portion
- References
- A GitHub repository containing all materials used

Tips for success:

- Think creatively. Why are you observing certain trends? What major events might have occurred at that time to influence the outcome of your analysis?
- Use the cleaned dataset if you are having a hard time shaping the original dataset. It took us some time (and manual effort) to wrangle the data into a usable format. Make this case study about analysis and critical thinking, not frustrating cleaning.
- Familiarize yourself with the ARIMA model. Although this project only uses the “MA” portion – moving averages – it is useful to understand where the idea comes from. You could also expand further and predict future consumption if you feel so inclined.

How will I know I have Succeeded? You will meet expectations on this case study when you successfully follow and complete the criteria in the rubric below:

Spec Category	Spec Details
Formatting	<ul style="list-style-type: none"> Repository – A GitHub repo containing all materials <ul style="list-style-type: none"> Submit a link to the repo Contents <ul style="list-style-type: none"> Data folder Scripts folder Written portion References
Data folder	<ul style="list-style-type: none"> <u>Goal</u>: This folder contains all of the data for this project. Include original and cleaned datasets If your data fits in GitHub, place all of it here. If your data does not fit in GitHub use a single file explaining the process to obtain the dataset.
Scripts folder	<ul style="list-style-type: none"> <u>Goal</u>: This folder contains all the source code for the case study. Try to name each script according to the order it needs to be executed. Throughout all your scripts, you should include copious comments explaining what each command or sequence of commands accomplishes and what the purpose is.
Written portion	<ul style="list-style-type: none"> <u>Goal</u>: Analyze the results of the case study. Write a small paragraph summarizing the project. Discuss 1-2 findings or observations. Provide graphics from the analysis to support your findings. Think critically and defend your reasoning.
References	<ul style="list-style-type: none"> Include all references in a PDF document in IEEE Documentation style.

Acknowledgements: Special thanks to Professor Alonzi for providing the rubric format.