CS3 Rubric — Case Study Rubric

DS 4002 - Creator: Aryaa Deshpande
Due: TBD

Submission format

- **Upload your PDF report and link to GitHub repo to Canvas**

Individual Assignment

Why am I doing this? This case study is your opportunity to learn a real data science workflow by stepping into the shoes of an analyst or data scientist. You will explore a dataset, use provided scripts, read background material, and produce a final deliverable to submit to Canvas.

What am I going to do? You will complete a structured data science mini-project, using the materials provided in the GitHub repository. You will:

- Read the scenario and mission in the Hook Document
- Analyze the dataset provided using the provided README file as a guide
- Create a GitHub Repo with the outputs of these scripts to support your report
- Write a comprehensive 1-2 page report to present your findings through analyzing the output of the scripts provided to you

Tips for success:

- Make sure you read the Hook Document carefully, since it may reframe your entire mission
- Start early so you have time to explore the data
- Focus on clarity. This is not about advanced methods, but about understanding data and making sense of it.
-

How will I know I have Succeeded? You will meet expectations on CS3 Case Study when you follow the criteria in the rubric below and are able write a 1-2 page report that displays your findings following the criteria in the rubric below.

Spec Category	Spec Details

Formatting	• Your final submission must use the project materials in the GitHub repository.
	  ○ **Scripts Output Directory**: should contain all the outputs (preprocessing, graphs) generated by the scripts
	  ○ **Graphs Output Directory**: should contain the graphs that were produced by the script

	○ **Report PDF**: 1-2 pages, addressing the criteria below.
Scripts Output Directory	● Include all of the scripts output that were generated in this project (such as preprocessed data) ● Name the outputted scripts accordingly The Hook Document will link you to the GitHub repo.
Graphs Output Directory	● You will use the scripts in this GitHub repo to generate exploratory data analysis and graphs in which you are able to analyze the results. ● You should rely on these materials to: ○ Understand the topic ○ Execute the analysis ○ Support your final report
Report PDF	● The final report should clearly document your findings after going through this GitHub repo in a 2 page report. All of the following must be addressed: ○ Data distribution & coverage ■ Are there any notable imbalances, missing years, or states with sparse data? ■ Use graphs to support your answer with references. ○ Trends in Electricity Prices ○ Outliers & Anomalies ○ Model Performance & Results ■ Report quantitative performance metrics ■ Did the model meet the predefined success criteria? ○ Impact of external factors ■ Discuss whether technological growth, post-pandemic recovery, or global events appear to have influenced electricity prices ○ Insights and Future Improvements ■ Suggest how this model could be expanded or applied in real-world decision making Your report must contain a working link to your GitHub repository.

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)