CS Rubric – Ideal LinkedIn Skillset

DS 4002 – Fall 2023 – Professor Alonzi – Abby Zorn
Due: December 2023
Submission format: Upload link to Github repo to canvas and upload slide deck in PDF format.
Individual Assignment

General Description: Submit to canvas a link to your Github repository and a slide deck in PDF format.

Preparatory Assignments: Case study description and attached resources.

Why am I doing this? As data scientists, we often face the challenge of communicating complex questions and analyses in an efficient and digestible way to those who may not be in the data science field. This particular assignment focuses on completing an analysis and creating a deliverable to discover particular skills that will inevitably benefit you to learn or showcase down the line when engaging in the job search. Finally, placing yourself in the shoes of a true data scientist and addressing a way to give yourself and fellow soon-to-be enlightened peers a competitive advantage will allow you to model what being a data scientist is really like as well as engage in critical thinking that will *personally* benefit you!

- Course Learning Objective: use analysis and critical thinking skills
- Course Learning Objective: seamless and digestible communication/presentation methods

What am I going to do? You will begin by reading the one-page prompt for this case study. In that prompt, you will be asked to take on the role of a true data scientist! You will also be introduced to the requirements needed to complete this case study. Next, read the resources available to you on how to begin and which packages you may want to use/install. You will then obtain and clean the data, perform your analysis, and prepare a presentation to communicate your findings to your fellow students. All code and resources should be uploaded to a singular Github repository. Deliverables include:

- Github repository - to provide all annotated code, figures, and resources so that your results may be replicated
- Presentation - to guide the audience through the work done by the student effectively

Tips for success:
- Take the bull by the horns—you've prepared all of college for this! This is your chance to pick something you learned and share it in a creative way.
- Keep it simple-- don't overthink it. You are creating something for peers and fellow almost-college-grads—not Steve Jobs.
- Don't give up and don't be afraid to do your own research. It's okay to fail and find solutions after the fact… failure will ultimately lead to success.

How will I know I have succeeded? You will meet expectations on this assignment when you meet the criteria outlined in the rubric below.

Spec Category	Spec Details

Formatting	• Repository – A GitHub repo containing all materials
	   o Submit a link to the repo in Canvas
	   o Everything is contained in the repo or linked to it if appropriate, including the presentation slides
	   o Should contain a source code folder, data folder, figures folder, license declaration, README.MD. and a copy of your slide deck
	• Presentation – should contain an 10-12 slide deck in PDF format

Github Repository	*Goal:* This repo is your credible method of explaining your results. Someone should be able to download this repo and produce the same findings and easily understand how to do so—keep in mind you're working with peers your own age. • **Source code folder:** contains annotated markdown file, including packages used. No unused packages are necessary in final file. • **Data folder:** contains data set (both before and after cleaning) • **Figures folder:** contains downloadable version of all used figures and captions. • **License:** in the form of LICENSE.MD file (use the MIT specifically) • **README.MD:** explain how to use repo, sections designated using "##". It should include: o Table of contents, purpose/context statement, description of source code folder, description of data, link, & data dictionary, figures section including figures and captions, and a references section • **References:** attach any new resources used for completion of study • **Copy of slides in PDF format**
Presentation	*Goal:* This presentation is how you will primarily be communicating with your peers and fellow workforce enthusiasts! It will allow you to communicate your findings clearly and efficiently. The following slides should be included: • **Title (1 slide):** list your primary finding as the title in a creative way (hook); include name, course, and date of this presentation. • **Motivation/context (1-2 slides):** using the provided resources, explain the motivation/context for your research. • **Hypothesis (1 slide):** state your hypothesis. • **Data (2 slides):** introduce data set and link it, provide data dictionary, and explain data set cleaning process. • **Analysis plan (1-2 slides):** explain to audience how you reached your conclusion—avoid too much data science-ey jargon! Explain to your common peer. • **Analysis bias & justification (1 slide):** not everyone may 100% believe your findings, so it is your job to remove that uncertainty. Explain any tricky analysis decisions and potential biases that might exist in your analysis.

	Results (max. 3 slides): show your peers your results… what traits did you discover?! Be sure to include any necessary figures and interpret graphs. **Conclusions (1 slide):** this is the slide where you present what your peers should be taking away from this analysis. List important findings and offer a concluding sentence of recommendation. **References/resources/acknowledgements (1 slide)**: cite your resources and any contributors