Case Study Rubric

DS 4002 - Spring 2023 - Lauren Smith
Due: Date TBD
Submission format: GitHub repository (submitted by link to canvas) and hard copy (brought to class)

Individual Assignment

General Description: Complete this case study and submit materials to Canvas and in-class by the due date that will be announced.

Preparatory Assignments

- All past data science assignments and projects

Why am I doing this? The goal of this assignment is to explore a research question and build a solution that is as accurate as possible. Identifying an area of interest and creating a research and analysis plan is an essential process in data science and this case study will give you more preparation as you move forward in your data science education. It is important that you do not think about the study as just another assignment to complete so you have it behind you; use this as an experience to develop your skills and put to use the ones that you have learned in your prior classes. Be curious and explore the topic. Try to build the best solution that you can think of and have fun while doing it.

- Course Learning Objective: create a deliverable based on the prompt given following the data science process

What am I going to do? In this assignment you will read the prompt and deliverable, read through the articles attached to the case study, explore the provided data, and work to build a solution. For your final deliverable, you must create a GitHub repository with all of the code, data, and figures your code produced throughout the process. This is all explained in the rubric below. Make sure to include an extensive README.md with all documentation and information about the project in your repo.

Tips for success:

- Focus, focus, focus – Remove yourself from any distractions that would hinder your progress on this assignment such as phones or talkative friends.
- Do some extra research - Looking at similar projects that others have completed online can give you an idea of what you can do to get started and the best practices for performing a data science project.
- Don't be afraid to ask questions - If you find yourself stuck on a portion of the case study, ask the professor or TA for help on how to move forward.
- Most importantly, have fun!

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

Spec Category	Spec Details

Formatting	· One Github Repository (submitted via link on collab)
	· The top level page should contain
	  ○ A README.md file (which auto displays)
	  ○ A LICENSE.md file (use MIT as default)
	  ○ A SRC folder
	  ○ A DATA folder
	  ○ A FIGURES folder

README.md	· Goal: This file provides information about the project and contents of the repository to anyone that visits it. If someone does not know anything about the project they should be able to recreate it themselves after reading this file.
	· Use markdown headers to divide content
	· Make an H2 (##) section explaining the contents of the repository
	· SRC section
	  ○ Make an H3 section for Installing/Building your code
	  ○ Make an H3 section for Usage of your code
	· DATA section
	  ○ Data Dictionary (use markdown table formatting)
	  ○ Link to data
	  ○ Relevant notes about use of data
	  ○ If all of the data did not fit inside the repository, provide instructions for how to access all of it.
	· FIGURES section
	  ○ Table of contents describing all figures produced and summarizing their takeaways
	  ○ Use markdown table formatting
	· REFERENCES section
	  ○ All references should be listed at the end of the Readme.md file (Use IEEE Documentation style (link))
	  ○ Include any acknowledgements

LICENSE.md	· Goal: This file explains to a visitor the terms under which they may use and cite your repository. · Select an appropriate license from the GitHub options list on repository creation. · Usually, the MIT license is appropriate.
SRC folder	· Goal: This folder contains all the source code for your project. · Include all code files you produce · The high-level documentation for this code lives in the main level README.md file within the SRC section · Make sure that code chunks within source code are commented to provide viewer with knowledge as to what is being accomplished

DATA folder	· Goal: This folder contains all of the data for this project · Make sure to include a link to the data source in the references section · 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.
Figures folder	· Goal: This folder contains all of the figures generated by your project · If you created any figure during the process place it here · Include with every figure relevant notes about the figure
References	· All references should be listed at the end of the document · Use IEEE Documentation style (link)

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