Sephora Review Sentiment Analysis Rubric

DS 4002 – Fall 2024 - Randa Ampah
Due: Monday, December 16th
Submission format: Upload link to GitHub repository to Canvas
Individual Assignment

General Description: Submit to Canvas a link to your GitHub repository for this project.

Why am I doing this? This assignment is meant to allow you to leverage the data science skills you have accumulated thus far, while also pushing you to expose yourself to new techniques, namely sentiment analysis. In completing this project, you will both reinforce your previous knowledge and teach yourself to work through unfamiliar concepts, as you often will in the future.

- Course Learning Objective: complete a sentiment analysis
- Course Learning Objective: practice translating information to a new audience

What am I going to do? You will go to the GitHub repository link and complete the sentiment analysis. You will then visualize your results, and summarize your findings in a one page paper. Deliverables include:

- GitHub repository containing your analysis - to provide the contents of your analysis
- One page summary paper - to summarize and contextualize your findings

Tips for success:

- Don't be afraid to ask for help. Your instructors are here to help you.
- It is normal to feel confused, this should feel new to you

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

Formatting	● One Github Repository (submitted via link on Canvas) ● 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 ○ A summary file
README.md	● Goal: This file serves as an orientation to everyone who comes to your repository, it should enable them to get their bearings. ● Use markdown headers to divide content. ● Make an H2 (##) section explaining the contents of the repository ● Section 1: Software and platform section ○ The type(s) of software you used for the project. ○ The names of any add-on packages that need to be installed with the software. ○ The platform (e.g., Windows, Mac, or Linux) you used. ● Section 2: A Map of your documentation ○ In this section, you should provide an outline or tree illustrating the hierarchy of folders and subfolders contained in your Project Folder, and listing the files stored in each folder or subfolder.
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
SCRIPTS folder	● Goal: This folder contains all the source code for your project. ○ Your code should contain some sort of visualization, whether it be for exploratory data analysis or reporting your results ● Include all the scripts you used. Try to name each script according to the order it needs to be executed to reproduce the results. ● All script files should include header comments at the beginning of a script to provide information that anyone working with or executing the script should be aware of. Throughout all your

	scripts, you should include copious comments explaining what each command or sequence of commands accomplishes and what the purpose is.
OUTPUT folder	● Goal: This folder contains all of the output generated by your project, e.g. figures, tables, etc. ● The content here can be in progress when MI3 is complete. It should be finished during MI4 though. ● Importantly, any information like tables, figures shown in your presentation should be here. ● Use informative names for your files.
Summary file	● Goal: This file summarizes your findings from your analysis in a clear and concise way (as if you are making your argument for a new product) ● One page ● PDF format ● 3 paragraphs ○ What was your analysis process like? ○ Summarize the results (what did your analysis tell you?) ○ What steps should the company take?
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 from Streifer & Palmer (2020).