CS3 Rubric – E-Commerce Case Study

DS 4002 – Spring 2024 - Instructor: Rebekah Weaver
Due: June 10, 2024
Submission format: Github Repository

Individual Assignment

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

Preparatory Assignments – Hook Document and attached supplemental materials

Why am I doing this? This is your chance as a 2nd year in university to showcase some of your reading comprehension and coding skills. It is important to be able to read and synthesize information to generate different ideas. It is also important to get familiar with the data science process of creating models and analyzing results. This assignment uses a real world example to analyze and create models for.

- Course Learning Objective: Follow steps of case study and create a model approach
- Course Learning Objective: communicate findings in presentation form for your peers

What am I going to do?
The GitHub repository for this case study can be found at https://github.com/rgw3wgs/CaseStudy3 . You will follow the case study provided and read through the supplementary materials about e-commerce and linear regression. You will conduct a thorough analysis of e-commerce review data using the dataset provided. You will use your technical and conceptual skills to analyze this case study and create your own model to analyze your topic. You will provide a deliverable that covers all requirements, models, results, and conclusions. You will then make a presentation for your findings.

- GitHub repository: this should contain code and necessary data
    - Data dictionary
    - Source code
    - Model images
- Presentation: discussing steps taken, model approach, and results

All of this will be submitted electronically via a link to a github repository

Tips for success:
- Read through all of your materials
- Stay organized and focused
- Take your time and ask questions if you need to
- Talk to your fellow students.
- Have fun and get creative!

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

Formatting	● One Github Repository (submitted via link in canvas) o A README.md file (auto displays) o A LICENSE.md file (use MIT as default) o A DATA folder o A SRC folder o A FIGURES folder o A REFERENCES folder ● Presentation
Github Repository	● Goal of this repository is to organize all of your code, data, materials, and images into one organized platform ● Use markdown headers to divide content ● README.md o Make a section explaining the contents of the repository o SRC section - Installing/Building your code and usage o DATA section - data dictionary, data links ● LICENSE.md o Goal: This file explains to a visitor the terms under which they may use and cite your repository. o Select an appropriate license from the GitHub options list on repository creation. o Usually, the MIT license is appropriate. ● SRC folder - contains annotated markdown file o This folder contains all the source code for your project. o Include all the scripts you used. Try to name each script according to the order it needs to be executed to reproduce the results. o 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. ● Data folder o contains all data sets including cleaned and uncleaned versions. Also contains a data dictionary and EDA for your variables. Hypothesis Test also included. ● Figures folder o contains all figures produced in your project with takeaways ● References Folder o All references should be listed at the end of the document o Use IEEE Documentation style (link)
Presentation	● About 8 slides ● PDF format for submission to canvas ● Generate the slides through the program of your choice ● Slide numbers (except for title slide) ● Order o Title & Outline - 1 slide o Case Study summary and goal (context, hypothesis, research question, modeling approach, goal) - 1-2 slides o Data Explanation (data dictionary information) - 1 slide o Analysis Plan and Justification - 1 slide o Any uncertainty or problems that resulted - 1 slide o Results and Conclusions - 1-2 slides o What did you learn - 1 slide o References at the end