CS – Choosing a Career Rubric

DS 4002 – Spring 2023
Due: Wednesday, May 10
Submission format: Upload link to github repo to canvas

Individual Assignment

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

Preparatory Assignments – CS1

Why am I doing this? This case study will test your ability to act like a data scientist. This case study provides real data to answer commonly asked questions. Throughout your data science courses, you have developed your analysis, interpretation, and presentation skills. You will now put those skills to the test by completing this case study. By completing this case study, you will have demonstrated your ability to think and perform like a data scientist.

- Course Learning Objective: prepare findings for presentation to your peers

What am I going to do? You will begin by reading the one-page hook document, which will give a generalized description of the case study. Then you will read over the provided materials and data set. You will use these given articles as motivation for your analysis. You will then perform an analysis on the data. Your analysis will then be packaged as a slideshow presentation.

Tips for success:
- Look at different ways of breaking up the data. For example, you don't only have to look at gender, age of death, and manner of death
- You can create your own general career paths, but I would recommend sorting them as STEM, Business, Art, Education, and Other (feel free to add more if needed).
- Use different modeling techniques – linear regression isn't the only option!
- Your presentation is pretty open-ended, just make sure you touch on the main ideas outlined in the rubric below

How will I know I have Succeeded? You will meet expectations on CS – Choosing a Career when you follow the criteria in the rubric below.

Spec Category	Spec Details
Formatting	- Repository – A github repo (and cloud storage folder if necessary) containing all materials - Submit a link to the repo - Everything is contained in the repo or linked to it if appropriate

	- o Contents ▪ A README.md ▪ A LICENSE.md ▪ A DATA folder ▪ A FIGURES folder ▪ A SRC folder ▪ A PRESENTATION folder
README.md	- Goal: This file will provide a general view on everything contained in your repository - Use markdown headers to divide content - Make an H2 section explaining the contents of the repository - DATA section - o Link to data - Figures section - o Table of contents describing figures produced with their summaries - PRESENTATION section - o Explanation of hypothesis to be explored in presentation - REFERENCES section - o List any outside references using IEEE formatting - o Include any acknowledgments
LICENSE.md	- Goal: This file explains the rights a visitor has when using or siting your repository - Select the best option (likely MIT) from the GitHub license options list
DATA folder	- Goal: This folder contains all of your data for this case study - Upload your data as a CSV file - If your data does not fit, write a single file explaining how to find your data
FIGURES folder	- Goal: This folder contains all of your relevant figures for this case study - Each figure must be labeled - Include a description under each figure with a 1-2 sentence takeaway about the figure - All figures used in your final presentation must be in this folder
SRC folder	- Goal: This folder contains all of the source code for this case study - Include all code files you produce through your analysis
PRESENTATION folder	- Goal: This folder contains your presentation on the case study - Include your presentation in pdf format - Slides: - o Title ▪ Name ▪ Motivating question ▪ Overview of topics - o Motivation ▪ Hypothesis ▪ Modeling approach ▪ Any relevant background information

	○ Data Acquisition ▪ Explain how you acquired the data ▪ Summarize your data set ▪ Explain any cleaning you performed on the dataset ▪ State the format of the data ○ Analysis Plan ▪ Explain your analysis method ▪ Include an explanation on why you chose the modeling method you did ○ Results ▪ Answer the hypothesis ▪ Explain the hypothesis in the context of the dataset ▪ Include at least 1 relevant figure ○ Acknowledgements ▪ List your references in IEEE formatting ▪ Include any acknowledgments

Acknowledgements: Thank you to Professor Alonzi for his help on this project and for the rubric outline. Thank you to my group mates Avery Goldberg and Olivia Kaczmarskyj for all their hard work this semester.