COGS9: Introduction to Data Science

Final Project: Part 1

Submit through Gradescope as a group, do not submit individually

Group Member Information:

First Name	Last Name	PID

Your Group's Data Science Question

Clearly state the specific data science question you're interested in answering. (2 pts)
REPLACE WITH YOUR GROUP'S TEXT HERE AND DELETE INSTRUCTIONS ABOVE

Hypothesis

Background Information

Data

Include a description of the perfect dataset you would need/want to answer this question. How many observations would you need? What variables would you collect? Explain the perfect dataset that you would want to answer this question.

Then, look online for available datasets. Find a dataset that could be used to answer this question. Describe how many observations are included and what variables have been collected. Discuss the datasets limitations and how it differs from your ideal dataset. Alternatively, you can collect your own data. You must explain what information you would collect and how you are going to collect this data. These do not have to be actually collected, but describe what exact data you would want and how it would be collected. (4 pts)

REPLACE WITH YOUR GROUP'S TEXT HERE AND DELETE INSTRUCTIONS ABOVE

Ethical Considerations

Read about [Deon's data science ethics checklist](#) and consider the topics discussed in lecture. Then, discuss what ethical considerations must be made when answering your specific data science question. Brainstorm and explain how you would address these considerations for each of the following categories in your specific project: Data Collection, Data Storage, Analysis, Modeling, Deployment. Feel free to write about additional ethical considerations you would make that aren't included on the checklist. Note that data privacy is NOT the only ethical consideration for a data science project. It is a piece, but there is a lot more that has to be considered. (2pts per section)

(10 pts total for this section)

USE THE FOLLOWING TEXT HEADERS AND DELETE INSTRUCTIONS ABOVE

Data Collection
ADD YOUR GROUP'S TEXT HERE

Data Storage
ADD YOUR GROUP'S TEXT HERE

Data Analysis
ADD YOUR GROUP'S TEXT HERE

Modeling (Statistical or Machine Learning)
ADD YOUR GROUP'S TEXT HERE

Project or Model Deployment
ADD YOUR GROUP'S TEXT HERE

Group Participation

A sentence or two about your contributions to this group project

- __name__ :
- __name__ :
- __name__ :
- __name__ :
- __name__ :