COGS108 Final Project Checklist

Students have asked for a rubric. While the points for each of these checkboxes is not included, you can use this checklist to help guide your thinking on the final project. If you check off all the boxes below, you should be in good shape to get a perfect score on your final project.

- ❏ **Introduction & Background (3.5 pts)**
    - ❏ <u>Overview:</u>
        - ❏ Write a clear summary of what you did
        - ❏ Limit overview to 3-4 sentences
    - ❏ <u>Research Question:</u>
        - ❏ Include your *specific* data science question
        - ❏ Make sure what you're measuring (variables) to answer question is clear
    - ❏ <u>Background & Prior Work</u>
        - ❏ Include explanation of what work has been done previously
        - ❏ Include citations or links to previous work
    - ❏ <u>Hypothesis</u>
        - ❏ Include your team's hypothesis
        - ❏ Ensure that this hypothesis is clear to readers
        - ❏ Explain why you think this will be the outcome (what was your thinking?)
- ❏ **Dataset(s) (3.5 pts)**
    - ❏ Include an explanation of dataset(s) used (i.e. features/variables included, number of observations, information in dataset)
    - ❏ Source included (if outside dataset(s) being used)
- ❏ **Data Analysis**
    - ❏ <u>Data Cleaning & Pre-processing (3.5 pts)</u>
        - ❏ Perform Data Cleaning and explain steps taken OR include explanation as to why data cleaning was unnecessary (how did you determine your dataset was ready to go)
        - ❏ Dataset actually clean and usable after data wrangling steps carried out
    - ❏ <u>Data Visualization (5.25 pts)</u>
        - ❏ Include at least three visualizations
        - ❏ Clearly label all axes on plots
        - ❏ Type of all plots appropriate given data displayed
        - ❏ Interpretation of each visualization included in text
    - ❏ <u>Data Analysis & Results (8.75 pts)</u>
        - ❏ EDA carried out with explanations of what was done and interpretations of output included
        - ❏ Appropriate analysis performed
        - ❏ Output of analysis interpreted and interpretation included in notebook
- ❏ **Privacy/Ethics Considerations (5.25 pts)**
    - ❏ Thoughtful discussion of ethical concerns included
    - ❏ Ethical concerns consider the whole data science process (question asked, data collected, data being used, bias in data, analysis, post-analysis, etc.)
- ❏ **Conclusion & Discussion (5.25 pts)**
    - ❏ Clear conclusion (answer to the question being asked) and discussion of results
    - ❏ Limitations of analysis discussed
    - ❏ Does not ramble on beyond providing necessary information

- ❏ **Submission**
  - ❏ Edit all text for clarity
  - ❏ Name file **FinalProject_groupXXX.ipynb** (where XXX is replaced with your group number)
  - ❏ Convert file to PDF
    - ❏ Check to make sure all text and images are visible
  - ❏ Submit PDF to gradescope
    - ❏ Mark which questions are on which pages
    - ❏ Add all group members to submission