

Projects

Materials and outlines for the Final Project for COGS108.

Carefully read through all the details you'll need to know for your COGS108 Final Project.

- Option 1: Group (pdf here)
- Option 2: Individual (pdf here)

Project Schedule

Project Proposal

Due Friday, Oct 23 (11:59 PM)

Final Project

Due Wed, Dec 16 (11:59 PM) - Finals Week

Project Templates

Option 1: Group Project

- Proposal
- Final Project

Option 2: Individual Project

- Proposal
- Final Project

Final Project Checklist

Students often ask for a rubric. 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.

Overview, Question & Background

Overview: - ☐ Write a clear summary of what you did - ☐ Briefly describe the results of your project - ☐
Limit overview to 3-4 sentences

Research Question: - ☐ Include a specific, clear data science question - ☐ Make sure what you're measuring (variables) to answer the question is clear

Background & Prior Work: - ☐ Include a general introduction to your topic - ☐ Include explanation of what work has been done previously - ☐ Include citations or links to previous work

Hypothesis: - ☐ 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):

- ☐ 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:

<<<<<<< **HEAD Data Cleaning & Pre-processing:** - ☐ 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 is actually clean and usable after data wrangling steps carried out
===== **Data Cleaning & Pre-processing** - ☐ Perform Data Cleaning and explain steps taken OR include an 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 >>>>>>>
542161c7326dcf0862e28bf4adb1630bdfbc0ce5

Data Visualization: - ☐ 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 the text

Data Analysis & Results: - ☐ 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:

- ☐ Thoughtful discussion of ethical concerns included
- ☐ Ethical concerns consider the whole data science process (question asked, data collected, data being used, the bias in data, analysis, post-analysis, etc.)

Conclusion & Discussion:

- ☐ Clear conclusion (answer to the question being asked) and discussion of results
- ☐ Limitations of analysis discussed
- ☐ Does not ramble on beyond providing necessary information

Final Checks:

- ☐ Edit all text for clarity
- ☐ Remove all instructions
- ☐ Check to make sure all text and images are visible
- ☐ Names and IDs included
- ☐ Renamed file :
 - Option 1: `FinalProject_groupXXX.ipynb`, where 'XXX' is replaced by your group's group number
 - Option 2: `FinalProject_GH.ipynb`, where 'GH' is replaced by your GitHub Username