CS 105 Final Project

Steps:

1. **Project proposal**
2. **Data collection and data cleaning**
3. **EDA**
4. **Completing your project**
5. **Writing a report**
6. **Presentation**
7. **Writing Questions**
8. **Answering Questions**

Description:

Find a topic that interests you. Be creative, do not use popular ideas. Choose something that you care about.

1. Project proposal

A short introduction to the topic, description of your project, technique(s) that you are planning on using (use our shared doc in google drive).

2. Data collection and data cleaning

Depends on your project and data that you can find.

3. EDA

(see below)

4. Completing your project

- State a question(s) or set a goal. How can you answer your questions or achieve your goal?
- What data do you need? Can you find it?
- Find a sufficiently large dataset(s) online (or use your dataset from Lab5&6).
- Do you need to clean it?
- Perform EDA on your data to better understand it. Decide, what needs to be done to capture interesting (related to your topic) information about the dataset. Report the results (use visualizations).

This is the main part:

- What methods/techniques/algorithms can you use and why? Here, you can use any of the techniques studied in class (but not anything that was done in the labs).

Use these techniques on your data, analyze the results.

5. Project report

Describe all work that you performed on the project in detail. Show all results, analyze the results. Complete the report before preparing your presentation. Please include each member's contribution. Upload to Grdescope.

6. Presentation

To receive credit for this part, you need to record your presentation.

7. Writing questions

Write three questions (with answers) about your project and include them into your ppt. All students should be able to answer your questions after watching your presentation (even if they are not very familiar with the topic before).

8. Answering questions (to be completed individually and uploaded to Gradescope).

Watch 10 presentations prepared by other groups and answer at least two questions from each presentation. Upload your answers to Gradescope.

What to submit:

1. Project proposal and techniques that you are planning to use – (Google Drive/Project teams)
2. Clean dataset (include the source) (Google Drive, your team folder)
3. Jupyter notebook (Google Drive, your team folder)
4. Report, pdf (Gradescope)
5. Video recording and PPT (Google Drive, your team folder)
6. Answered questions (Gradescope).

Grading:

Project proposal - -10 points off if submitted late

Project description – 10 points

Data preparation and EDA – 10 points

Main part (including PPT) – 50 points (15p./50p. for difficulty/creativity)

Recorded presentation (use PPT) – 10 points

Questions (answering) – 10 points

Penalties: for the first two days, for each day you are late, you are losing 20% of the score per day, for the next two days – 30% per day (for the part that was due) – no exceptions will be made.

Timeline

Final project proposal: March 4 (it is considered to be final if approved).

Uploading the project report, slides, videos – March 14, 11:59pm.

Answering questions – March 15, 11:59pm.