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.

- 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. (Do not use Kaggle.)

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 to submit an informal proposal).

2. Data collection and data cleaning

Depends on your project and data that you can find.

3. EDA

- 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).

4. Main part

- What methods/techniques/algorithms can you use? Here, you can use any of the techniques studied in the second part of the 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 and analysis of the results. Complete the report before preparing your presentation. Include each member's contribution. Upload to Grdescope.

6. Presentation

- Prepare slides for your presentation.

7. Writing questions

- Write three questions 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 were not very familiar with the topic before).

8. Recording

- To receive credit for this part, you need to record your presentation. The presentation should be 7 – 10 min long and should cover all parts of your project.

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

- Watch 10 presentations prepared by other groups and answer at least one question 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/Project Report)
5. Video recording and PPT (Google Drive, your team folder)
6. Answered questions (Gradescope/Answering Project Questions).

Grading:

Project Proposal, Project description – 15 points

Data preparation and EDA – 10 points

Main part (including PPT) – 60 points (15p./60p. for difficulty/creativity)

Recorded presentation (use PPT) – 10 points

Questions (answering) – 5 points

Penalties: for the first two days, for each day you are late, you are losing 15% of the score per day, for the next two days – 20% per day – no exceptions will be made.

The Project is due Thursday, March 23, 8am.

Answering questions part is due Thursday, March 23, 11:59pm.