Term Project

Preliminary Phase

There are five steps in the preliminary phase of term project.

First, find your teammates (for undergraduate students only).   Each team consists of at most three members.

Second, select a dataset of your interests, and name the source of the dataset.   Except for special cases, the minimum requirements of the dataset are 10 or more data features (columns) and 1,000 or more records (rows).

Bonus (for term project) will be given to:

1.   (5 points) Download the dataset by a program.

2.   (10 points) Keep updating the dataset with a program automatically run daily or weekly.

Third, select at least one of the four major categories of analysis methods that you would like to experiment on.   You might need to think about some auxiliary utilities that you need to develop for the project.   The four major categories include:

1.   Regression Analysis

2.   Clustering

3.   Classification

4.   Time Series Analysis and Forecasting

Hint: You may validate the selection of features with correlation analysis.

Four, list the tools and programming languages (Python and R are suggested, but not limited) that you plan to use for the project.

Five, give a goal to your project, and list the work items.   For clarity, give some example charts or data tables that you expect when the project is done.

If you complete the five steps, congratulations, you may start working on your term project

Presentation Phase

For each team, you need to provide 3 items for the term project.

1. A presentation note (a pdf edition of a powerpoint file),
2. A presentation with all members on stage, and
3. A project report (a pdf edition of a word file).

1. A Presentation Note

Save the note in pdf format and upload to elearn by the end of the day for your presentation. These items should be included in the note:

- The selected dataset, with numbers of features (columns) and records (rows) specified,
- The goals of analysis, with one or more major categories of analysis methods specified,
- The programming languages and utility tools used in the analysis,
- A high-level framework or architecture of the analysis,
- Examples of test runs of analysis programs,
- The analysis result from the programs, and
- The contribution of the analysis results.

2. Presentation of Term Project

The presentation should include three parts:

- (7 minutes) Verbal description of the term project, using the presentation note,
- (3 minutes) Discussion of the analysis programs, with high-lights on parts of the source code.
- (5 minutes) Demonstration of the execution of analysis programs, either live or recorded, and
- (5 minutes) Q&A.

3. Project Report

The report needs to include the following items.

I. A cover page.
II. The problem description.
III. A high-level framework or architecture diagram.
IV. Highlight of the way you write each of the programs.
V. The program listing.
VI. Test run results.
VII. Contribution of the Analysis Results and Discussion.

Note: "No-show" in the selected time will get <u>no chance</u> to make up the verbal presentation. The score of term project will be 70% of the report score (30% penalty).