Final Project
P8106: Data Science II

Deliverables and deadlines

1. Your proposal is due on **Sunday April 22** at 11:59pm.
2. Your write-up is due on **Sunday May 6** at 11:59pm.
3. Your peer assessment is due on **Sunday May 6** at 11:59pm.

Each team only has to submit one proposal and one write-up.

Team registration and proposal

Define your teams and propose a project. This proposal should include:

- The group members (names and UNIs)
- The tentative project title
- The anticipated data sources
- The planned analyses/visualizations
- The planned timeline

Write-up

The write-up should be a polished report. Your report should have the following sections (you can of course add other sections if you want), and should be no more than 8 pages.

Introduction

Describe your data set. Provide proper motivation for your work.

- What questions are you trying to answer?
- How did you prepare and clean the data?

Unsupervised analysis/Exploratory data analysis

- Is there any interesting structure present in the data?
- What were your findings? What insights into the data can you make?

Here you can use any techniques as long as they are adequately explained. If you don't find anything interesting, then describe what you tried, and show that there isn't much visible structure. Data science is NOT manipulating the data in some way until you get an answer.

Supervised analysis

- What predictor variables did you include?
- What technique did you use, and why did you choose it? What assumptions, if any, are being made by using this technique?
- If there were tuning parameters, how did you pick their values?
- How did you make your predictions?
- Discuss the training/test performance if you have a test data set (or you could split the data into two parts).
- Can you explain anything about the nature of the relationship between the predictors in your model and the predictions themselves?

- What were your findings? Are they what you expect? What insights into the data can you make?

You can use any of the classification/regression techniques that we learned in the course, or any other techniques as long as they are adequately described.

Peer assessment

We ask you to provide an honest assessment of the contributions of the members of your team, including yourself. The feedback you provide should reflect your judgment of each team member:

- Preparation - were they prepared during team meetings?
- Contribution - did they contribute productively to the team discussion and work? What are their contributions?
- Respect - did they encourage others to contribute their ideas, and provide feedback in a constructive way?
- Flexibility - were they flexible when disagreements occurred?