Assignment 3: Data Analytics (Fall 2025) Oral

Due: February 21st, 2025

Presentation method: Presentations during the class time on February 18th and 21st

Note: Your presentation for this assignment should be the result of your own individual work.

Take care to avoid plagiarism ("copying"), and include references to all web resources, texts, and class presentations. You may discuss the project with other students, but do not take written notes during these discussions, and do not share your presentations before class.

General assignment: Term project proposal. This is a chance to get feedback on what project (questions, data, methods, etc.) you currently plan for your term project. The grade is NOT based on presentation quality. The grade is based on covering content detailed in #1a-f below.

Guidance: Your term projects should fall within the scope of a data analytics problem of the type you have worked with in class/ labs or know of yourself – the bigger the data the better.

This means that the work must go beyond just making lots of figures. You should develop the project to indicate you are thinking of and exploring the relationships and distributions within your data. Start with a hypothesis, think of a way to model and use the hypothesis, find or collect the necessary data, and do both preliminary analysis, detailed modeling and summary (interpretation).

Details of what will be required will be in Assignment 6.

Note: What you present in this assignment does NOT have to be what you eventually conduct your project on. This is to get you to start thinking about how an end-to-end project would look.

You can propose to use datasets made available by data-challenge competitions, but you

CANNOT use their hypotheses or challenge questions. You must develop your own hypotheses/questions.

4000 Level

Note: There may be methods you use that we have not covered – that is okay.

1. Oral presentation (5 mins)

a). Title (with your name)

b). Problem area – why it is of interest (in general or to you), what you might want to predict? This could be a hypothesis.

c). The data – where it might come from, why it may be applicable, and any preliminary assessment you've made.

d). How you plan to conduct your analysis: distribution, pattern/ relationship and model construction. What techniques do you think you will use?

e). How do you plan to apply the model? What are the possible uncertainties?

f). What do you want to predict and what decisions (prescriptions) may be possible? What would a good outcome be?

6000 Level

Note: There may be methods you use that we have not covered – that is okay.

1. Oral presentation (6 mins).

a). Title (with your name)

b). Problem area – why it is of interest (in general or to you), what you might want to predict? This could be a hypothesis.

c). The data – 6000 Level students must have minimum of 2 datasets* for the project.

where it might come from, why it may be applicable, and any preliminary assessment you've made.

d) How you plan to merge (combine) two or more datasets.

e). How you plan to conduct your analysis: distribution, pattern/ relationship and model construction. What techniques do you think you will use?

f). How do you plan to apply the model? What are the possible uncertainties?

g). What do you want to predict and what decisions (prescriptions) may be possible? What would a good outcome be?

NOTE*: 6000 Level Students must have a minimum of two datasets (or more datasets) and must use them to combine and create the project dataset that you are planning to conduct the analysis.

Two datasets should have a meaningful relationship when you combine them. Please come and talk to the instructor if you have any questions or doubts during the One-on-One sessions or during office hours.