

Individual Final Project Requirements

Stats 140XP Statistical Consulting for Data Scientists

Due Dates Vary, see schedule below

Introduction

You are receiving this assignment because we were not able to place you on a Stats 140XP for a variety of reasons. What follows are the components of your final project and their deadlines. You are excused from class for the remainder of the quarter and if you need to meet with me, please schedule an appointment.

If you need clarification, please reach out to me directly.

Components

- 1. Proposal
 - Due no later than 11:59pm Monday March 3rd. E-mail it directly to me.
 - Please provide a proposal based upon the structure identified at the end of this document (e.g., Project title, introduction, literature review, research questions, methods, potential impact, references). For full credit, at minimum your proposal should have:
 - * Project title
 - * One paragraph literature review
 - * An appropriate data set (no iris, no Titanic, no penguins)
 - * One Research question.
 - * Identification of the Methods needed to test your research question.
- 2. Meeting on Zoom or in person with me to go over your proposal
 - Schedule it (please send an e-mail to me via BruinLearn) for a time between 11am and 7pm on Tuesday March 4th.
 - We just need 15 minutes and we will go over your proposal together.
- 3. Final Report based on the proposal
 - Due no later than 11:59pm Friday March 14th. E-mail it directly to me.
 - There is no minimum or maximum length, it simply should be the fulfillment of your project proposal.
 - There should be an exploratory data analysis (EDA) of your analysis dataset included in the report to make it clear you understood your dataset.
 - You can use any programming language (e.g., Python, R) you want.
 - The final report should be a PDF.
 - The code used to generate the numbers/statistics/visualizations in your report should be made available to me as well as a copy of the data. You can just point me to a link if emailing a copy is difficult (e.g., too large)

- 4. Video Presentation
 - Due no later than 11:59pm Friday March 14th. E-mail it directly to me.
 - This should have the same format as “Your UCLA Story” from week 1. So landscape, .mp4 is a good format but I can use most any, should run between 3-7 minutes, shorter than 3 minutes is OK, longer than 7 is not OK (deductions).
 - I am evaluating your presentation based on correctness and how closely you followed your proposed project.
- 5. Video Review of the other team presentations
 - Review is due no later than 11:59pm Friday March 14th. You will complete it on BruinLearn (link to be made available on Tuesday March 11th)
 - On Tuesday of Week 10 you will receive an assignment to review 7 team presentations chosen at random. These are video presentations and they will be posted to BruinLearn, you can watch them anytime but your review is due on March 14th.
 - You will watch them and answer a few questions about them and evaluate (basically give a grade) to each one. Their presentations are at most 10 minutes long but probably closer to five minutes, so budget one hour for this part.

Grading

The grading/scoring and the composition of your final project grade is the same for you as it is for your classmates (i.e., proposal/meeting/report & code/video presentation/scoring other teams). Given that you are working individually, I am not seeking the same level of complexity in your proposal or resulting report/video as I am of your classmates.

If you have questions or need clarification e-mail me directly from BruinLearn please

The Parts of the Proposal

Project Title

- Make it clear, concise, and informative.
- Example: The Role of Twitter in the Rise of Donald Trump: A Study of Extreme Sentiment

The Introduction (write this last)

- Define the research problem.
- Share why this particular problem is important to you (e.g., theoretical, practical).
- State your research objective or research hypothesis.
- (nice but completely optional) Engage the reader with something compelling (a reason why they should listen to you)

What does it mean to define a research problem?

Start with a general topic that is interesting to you and also relates to stats, data or issues that you care about. Ask yourself:

- What topics do I care about?
- What are the current arguments in my field(s) of interest?
- What real-world issues need statistical intervention/investigation?

At this point, turn to what others have done (e.g., search Google) to help refine the question

Literature Review

- Write a paragraph on what others have said/done
- This demonstrates knowledge of existing arguments
- Tell us how your study fits in, maybe you are replicating another's work (that's fine) maybe you see a gap in knowledge that you aim to fill.

What should a literature review do?

- It should help you narrow the topic. You want it narrow in this class.
- Suggestions, you could narrow it by:
 - Population (e.g., youths)
 - Context (e.g., social media)
 - Time period (e.g., 2015-2016)
 - Geography (e.g., California)
 - Specific Data (e.g., Twitter database)

Research Questions

- Please state a research-type question that helps address the larger research problem of interest
- Focused problems are easier to answer (e.g., young women have lower levels of support)
- The research question should be specific, testable, and consistent with your methods below.

Methods

- Tell us about your research design - your strategy for choosing the data and the analysis of it.
- Tells us a little about the data (surveys, secondary sources, experiments, web scraping, etc.).
- Justify your choice of analysis method (e.g. sentiment analysis)
- Discuss limitations of your choices (you don't need to offer correction)

Potential Impact (if any)

- Explain how your research could advance knowledge.
- Or mention real-world implications (e.g., practice, industry, social impact) of your work.

References

- Use a consistent citation style (APA, MLA, Chicago, etc.).