

A3: Final Project Proposal

For this assignment, you will write up a study plan for your final class project. The plan will cover a variety of details about your final project, including what data you will use, what you will do with the data (e.g., descriptive analysis, statistical analysis, train a model, artwork), what results/outcomes you expect or intend, and most importantly, why your project is interesting or important (and to whom, besides yourself). It should include these components:

- Motivation/problem statement: Why are you planning to do this analysis? Why is it potentially interesting and useful, from a scientific, practical, and/or human-centered perspective? What do you hope to learn? (3-point)
- Research questions and/or hypotheses: These describe what you hope to discover or determine in the course of your research. (1-point)
 - Example research question: What is the impact of an MS degree on data scientist salaries over the course of their careers?
 - Example hypothesis: Earning an MS degree is associated with an increase of x% in career data scientist salaries compared to similar data scientists who do not earn a degree.
- Background and/or Related Work: What is already known about the phenomenon you are investigating? How does previous research or background info inform your decision to perform this study, the way you designed the study or your specific research questions? Make sure to include references (endnotes and/or inline hyperlinks) to the sources of background information—whether they are websites, news articles, or peer-reviewed research. (2-point)
- Data used: What dataset do you plan to use, and why? Summarize what is represented in the dataset; Link to the dataset, and specify license/terms of use; Briefly justify why this dataset is relevant to your problem statement; Highlight any possible ethical considerations to using this dataset (and say why or why not). (1-point)
- Unknowns and dependencies: Are there any factors outside of your control that might impact your ability to complete this project by the end of the semester? The purpose of this section is to get you thinking, in a practical sense, about your ability to complete this project within the time allotted. (1-point)
- Methodology: Describe how you plan to investigate this phenomenon. Don't just describe what your analytical methods are (e.g., "ordinary least squares", "student's t-test", "heatmap visualization", or "recurrent neural network"), it's critical to justify why these are appropriate methods for gathering and analyzing your data or presenting your findings. You are expected to be thorough here: please describe to the best of your ability the entire series of gathering, analysis, and presentation methods you plan to use. (1-point)
- Timeline: What are the components of your project and how are those scheduled? You need to set a realistic scope to make sure you can complete the project. (1-point)

The deliverable for this assignment will be a document in PDF format. You will turn the assignment in by submitting your PDF before the deadline of **9PM on November 1**. Please make sure to make **ONE SUBMISSION PER TEAM** and **list all the NAMES and NETIDs of team members** at the top of the first page of your proposal.