

# DATA 23700 Autumn 2025

## Assignment 2: Project Proposal

*Due October 13, 2025*

For the project, students will produce an original data analysis and a short written report. The first steps involved in the project are to **(1) motivate a problem or question** the student will address using skills acquired in this course and **(2) find, collate, or collect a dataset to analyze** in order to address their motivating problem or question. These first steps are the emphasis of the project proposal.

Students should submit a two-page PDF document both: (1) motivating a problem or question they wish to address in their project; and (2) describing the dataset they plan to use in terms of what variables are included, provenance, and why the data are suitable to address the motivated question or problem. It is common for students to collate multiple data sources in order to enable a satisfactory analysis. Proposals should be single spaced in 11pt Times New Roman (or similar font) with one-inch margins. Students should include at least one visualization of their chosen dataset to demonstrate its value for the purpose they have identified.

It is important that students identify a project motivation. Proposing to do open-ended exploration of a topic is not satisfactory. Instead students should pose a narrative problem framing, research question, and/or hypothesis that can drive their work. This will make for an interesting and logically coherent project, and it provides students an opportunity to develop data science skills around a topic of interest.

Similarly, it is important to choose a dataset that can support analysis of sufficient depth to demonstrate skills acquired in the course and of sufficient interest to support a narrative about the analysis akin to a technical report or a piece of data journalism. Choosing a dataset that has no interesting structure or which cannot support the intended analysis would set students on shaky footing for the project.

The purpose of this proposal is to help students start early and find a good project idea. Students are encouraged to visit office hours to discuss ideas for their projects.

**Students must work alone.** Recall that the project is intended to be individual work. Collaborating in any way (e.g., choosing the same dataset as a peer in order to reduce effort) is an academic honesty violation.

**Students should turn in Assignment 2 by submitting a PDF document to Gradescope.** The project proposal has the same grading, late work, and resubmission policies as other assignments.