

POLI381 FINAL PROJECT: DATA-DRIVEN RED TEAMING

This version: January 8, 2020

A major objective of Poli381 has been to develop your ability to be part of data analytical conversations: about what data to use, how to analyze them, and what conclusions to draw. The final project for this class asks you to put those skills to use to answer a substantive question in political science.

We will be engaged in what I call “red teaming.” (I think the term comes from the CIA.) In a nutshell, a red team is a group that works independently from a main analysis team to provide a check on procedures and conclusions. For instance, if CIA analysts concluded that a foreign adversary had a secret weapons program, CIA leaders might form a red team to look at all the intelligence material again, and see if they reach the same conclusions. The procedure can help confirm that the analysis was done properly, and guard against motivated reasoning (e.g. an analyst who wants to reach a particular conclusion, because it will help his or her career).

For the final project, there will not be a “main” analysis team. Rather, each of several groups will be given a substantive question, and a dataset that might be used to answer it. You are to write a research paper that answers the question and explains how you answered it. But you are expressly prohibited from discussing your approach or conclusions with other groups during the project. At the end of the semester, all groups will share their procedures and conclusions, and we will see how much of a consensus was reached.

It bears emphasis that it’s best not think of this as a *competitive* assignment. The point of having multiple groups answer the same question is not to see if one group can do better than the others. The assignments will be graded in absolute—not relative—terms. Instead, compartmentalizing the labor, as this exercise does, helps to identify different approaches to answering questions—and then we can have a conversation about what the pros and cons of each procedure might be. Working independently also helps guard against blind spots. But as with a CIA red team, we should all ultimately think of ourselves as being on the same “side.”

The Final Project is due in three parts, on the dates noted on the syllabus:

- **First draft of Red Team Report.** This is a document that lays out your procedures and conclusions as concerns one of the substantive questions below. The document should follow all the structural guidelines—laid out below—for the final draft, except it does not need to have the “Response to Comments” section. Note that “first draft” is not a euphemism for “incomplete” or “sloppy.” You should think of this as a formal, polished document. “First draft” merely signifies that, because the inferences have not been subjected to external scrutiny, they should be considered more tentative than in a final draft.
- **Comment on other teams’ reports.** For this part of the final project, your group should read the reports of the other teams assigned to the same substantive question. Then, you should write a 1,200 word comment on them. (One 1,200-word document covering all the other reports.) This document will be shared with the report authors, for them to consider in revising their report. You can comment on any aspect of the other reports. However, the best comments 1) flag analytical choices that might lead a group to erroneous conclusions, explaining why, and 2) articulate any additional analytical steps to take (e.g. conducting an additional test to assess whether a problem exists, or proposing a correction to the underlying procedure).
- **Final draft of Red Team Report.** This document lays out your final analysis and conclusions. We will discuss expectations at greater length in class, but it should be organized as follows. (Word counts are rough guidelines, not ironclad rules.)

- The report should start with an introduction, which should briefly discuss the substantive question you are answering, why it is important, and provide an overview of your procedures and conclusions. (500 words)
- Next should be a section that lays out, in full detail, the analyses you conducted and what you found. You should include all the information that an outsider would need to reproduce your results (even without having access to your R code). This section will be the longest. I expect it to include several tables and figures, and it very well might be organized into a series of subsections. (E.g. “Coding of the dependent variable,” “Justification for regression model,” “Considering alternative models,” and “Conditionality in the main effect.” The exact subsections will vary by group.) (3,000 words)
- For the final draft only, a response to the comments you received from the other red teams working on the same question as you. Here you should discuss what steps you took in response to the critiques you received, or explain why you disagree with one or more critiques. (800 words)
- A conclusion section that summarizes your steps, inferences, and discusses limitations. You should include a discussion of what additional information would be the highest priority to collect, and how you would incorporate this information into your analysis if it became available. (600 words)
- Separately, each team member must upload analytical files that produce all of the quantitative results in the paper. You can discuss the steps to be taken and discuss how to accomplish them in R. But you cannot copy-and-paste each other’s code.

Citations in the paper should be parenthetical format and should look like this (Ryan 2019). There should be a works cited section at the end of the document that provides details on all the references.

All teams will focus on the 2019 [Quality of Government Standard Dataset](#), which is an agglomeration of information from several data sources. Here are the tentative substantive questions we will Red Team in Poli381 (approximately three teams per question).

- To what extent is the political equality of women (as of 2015) affected by public support for education? To what extent do significant armed conflicts thwart this progress?
- To what extent are life expectancy and infant mortality (as of 2015) affected by the receipt of foreign aid? To what extent does this relationship depend on the quality of government functions (fh_fog)?
- To what extent is governmental support for public services—especially health care (as of 2015)—affected by the number of women in government? How does this relationship depend on the extent of religious fractionalization in a country?

Although those are the core questions we will examine, they should be considered a starting point. I am open to discussing refinements and elaborations, as time allows.