

SOC 4930/5050: Final Project Instructions

Christopher Prener, Ph.D.

October 9th, 2017

Overview

All students will select a set of variables from the 2012 General Social Survey and perform an original data analysis culminating in a series of linear regression models. Each student will be responsible for selecting a research question, establishing a hypothesis, conducting an analysis to test that hypothesis, and presenting the results. This process and project mirrors the steps taken to author a quantitative conference presentation (for all students) and a quantitative journal article (for students enrolled in SOC 5050).

Data

Data for the final project is available in the [Final-Project-Data](#) repository that is part of our [SOC 4930 & SOC 5050 organization](#). The data have been obtained from the [Inter-university Consortium for Political and Social Research \(ICPSR\)](#). Their license precludes making the data publicly available, and so these data are only available to current students.

Memo



Due by September 25th



Submit as a GitHub issue

All students should submit a memo as a GitHub issue that identifies a topic and key variables you are interested in exploring. Make sure these variables are present in the 2012 edition of the GSS - not all variables are available for every year. Ideally, you should also identify a continuous dependent variable. All memos should be accompanied

by two or three key citations that describe the social science research on your topic.

Annotated Bibliography



SOC 5050 students only



Due by October 16th



Submit as a Markdown (.md) file

An annotated bibliography with twenty to twenty-five scholarly, peer-reviewed sources should be submitted as part of the preparation of the final project. Peer-reviewed sources include articles from reputable academic journals and books from University presses and other scholarly publishing houses.

Before you begin, you should select either the [American Sociological Association \(ASA\)](#), the [American Psychological Association \(APA\)](#), or the [Chicago Manual of Style \(CMS\)](#) style guidelines. If you use Chicago, please use parenthetical citations in the Author Date in-text citation format. Sociology masters students are strongly encouraged to use ASA, and Public and Social Policy doctoral students are strongly encouraged to use CMS. **Do not mix citation styles - you must select one and apply it consistently.**

Your annotated bibliography should include a properly formatted bibliography entry for each source as well as a paragraph length description of the source and its relevance to your project.

Conducting a Literature Search

I recommend following the steps below as part of a comprehensive strategy for conducting a literature search.

1. **Define your research topic and question:** What is it that you are trying to identify sources in the literature for?

2. **Identify appropriate search terms:** What are the keywords that might help you find relevant articles?
3. **Identify the “state of the art”:** Start by searching Annual Review journals (or similar, such as Sociology Compass) for recent articles that summarize the state of the literature in your area. You may not find an exact match, but you may find reviews of related topics. Use these articles to get a sense of what the “state of the art” is for your topic area(s).
4. **Citations from review articles:** Use the review articles you find to identify both seminal publications that define a field of study (may not be recent) as well as more recent empirical studies. Find and download these articles.
5. **Search other social science publications:** Whether or not you find helpful review articles, you will want to also search other key journals. Start with generalist journals and then move to more specific journals focused on particular types of research. Use the same search terms described above. In sociology, for example, the major generalist journals are *American Sociological Review*, *American Journal of Sociology*, *Social Forces*, and *Social Problems*. In one of my sub-fields, medical sociology, *Social Science and Medicine*, *Journal of Health and Social Behavior*, and *Health and Place* are all key journals.
6. **Search related journals:** Are there journals from other disciplines that are related? If so, search those as well using your search terms. In my research, the *American Journal of Public Health* and the *Annals of the American Association of Geographers* are both important journals from related disciplines.
7. **Search Sociological Abstracts:** If you are not finding many citations, you may want to search the Sociological Abstracts database, which catalogs many more sociological journals.
8. **Iterative Processing of Articles:** As you read through the articles you find and download in steps 3 through 7, seek out the studies that they cite that seem related. Download them, read them, and identify more related studies that they cite, then start the process over. Also use this process to identify books.

As you find sources, I recommend keeping track of them in a spreadsheet. This makes it easier to figure out if you have already found and downloaded an article that you may come across multiple times. Once you identify an article you might think is relevant, read the abstract then read the introduction and the conclusion. Does the article still seem related? If not, put the article to the side and mark it

as not likely relevant in your spreadsheet. Do not spend a lot of time reading or taking notes on articles that are not likely to be useful to you.

For articles that are relevant, focus on trying to connect articles to each other - who is citing who? Who is working on similar topics or making similar arguments? Who is contradicting or refuting other scholars? These connections will be useful to develop as you think about *synthesizing* the literature for your paper.

Reviews that contain the requested number of citations (20 to 25) at a minimum will be eligible for a check-plus. To earn the check-plus, each of these citations should be from peer reviewed sources. These sources should largely be drawn from sociology journals or journals that have a clear relationship to your topic. The reviews you write should be rich but succinct summaries of each source and note connections. If not having enough sources, having sources that are not peer reviewed or not drawn from sociology sources, or having reviews that are not clear, concise, and informative are all reasons for having a check or check-minus grade.

Analysis



Complete using an RNotebook (.Rmd) file

Your analysis should be completed in a well-organized RNotebook (.Rmd) file that uses literate programming techniques and RMarkdown text syntax. It should include the following elements:

1. A well-constructed hypothesis: what are you trying to test with your final regression models?
2. Descriptive statistics for all variables, as appropriate for their level of measurement.
3. A histogram for your dependent variable.
4. A full normality assessment of your dependent variable, including q-q plots.
5. Independent t-tests of your main independent variable of interest and your dependent variable.
6. Correlation analysis of all of your variables.

7. A scatter plot showing your main independent variable of interest and your dependent variable.
8. An initial set of three OLS regression models, each specified both as a \LaTeX equation and the appropriate R functions.
9. A full assessment of the OLS regression assumptions on your third and final model.
10. If adjustments are warranted, re-fit your initial three models (make sure to update your equations as necessary) on the adjusted set of variables and/or observations.

Draft Presentation



Due by November 13th



Submit your slides as a .pdf file

Use presentation software (either Microsoft PowerPoint or Apple Keynote) to put together a short, six to eight minute presentation of your findings. Please refer to my super-meta [presentation on presentations](#) for guidance on constructing a good presentation. Your draft will be graded using the “check” system outlined in the syllabus. At the draft stage, presentations should include:

1. A title slide (presentation title, name, date, etc.)
2. An agenda slide with the main topics you want to cover
3. A literature slide (SOC 4930 students should include a few key citations; SOC 5050 students should avoid the temptation to summarize each of their sources)
4. Plots, including your histogram, q-q plot, and scatter plot described in the *Analysis* section above.

Draft Handout**Due by November 13th****Submit your handout as a .pdf file**

Use \LaTeX to produce a handout that includes your title slide information (presentation title, name, date, etc.), a summary of your research question and hypothesis, a descriptive statistics table, and a few key citations. These should all fit on one page. Your draft will be graded using the “check” system outlined in the syllabus.

Draft Paper**SOC 5050 students only****Due by November 13th****Submit your paper as a .pdf file**

Use \LaTeX to produce a paper that, at a minimum, includes a draft “Introduction”, “Literature”, and “Data and Methods” sections. Keep in mind that your final paper should be approximately 5000 words long, and will also contain “Results”, “Discussion”, and “Conclusion” sections. Your initial sections should be written with this final length and content mind. Your draft paper should include a well-formatted bibliography. Your draft will be graded using the “check” system outlined in the syllabus.