

Qualifying Exam in Methods, Summer, 2021

The questions below constitute the qualifying exam for **methodologies**. To complete the exam, you must answer one question from each of the three sections below in an essay format. Each answer should be a self-contained essay approximately 8–10 pages long (i.e., 2,000–2,500 words).

Although you must cite to referenced work in each essay, one comprehensive bibliography for all three questions will be sufficient. You may consult materials from the reading list you've been provided, as well as any other scholarly sources you've read during the exam period or you may read for answering the questions. This qualifying exam is to be completed in two weeks. The evaluation of the exam will take two weeks. You will be notified of your performance by email.

Make sure that the office has your up-to-date contact information. Please use the identifying student number that you have been assigned at the top of your exam. The exams are graded confidentially, and numbers only will be used to identify the author of the exam. See the *Graduate Handbook* for any further details.

Answer one of the following:

1. Over the past half a century, panel surveys have become an indispensable method for sociological data collection. A forerunner in this regard is the Panel Study of Income Dynamics in the U.S. The three sub-questions all pertain to panel surveys.
 - a. A panel survey is a type of longitudinal survey. In comparison to a typical cross-sectional survey, what are some major advantages and limitations of panel surveys? Discuss.
 - b. Following the lead of the Panel Study of Income Dynamics, researchers in numerous countries have also adopted this method of data collection. Select one example of such panel surveys from another country and describe its major features.
 - c. The Panel Study of Income Dynamics has contributed to many sociological breakthroughs. Discuss at least three such major breakthroughs in our sociological understanding in the past five decades.
2. Panel data may require different statistical models for analysis. In recent decades, a lot of important research has been conducted to advance such models.
 - a. What are some of the most important statistical models for analyzing panel data? How do they differ from the standard regression model for cross-sectional data in terms of their major features?
 - b. When panel data are analyzed, what is the most serious challenge to statistical estimation? How do some of the models you discussed in the previous question solve such a challenge? Discuss.
 - c. When one analyzes a collection of cross-sectional datasets over time, the famous age-period-cohort conundrum becomes an issue. If you use a panel statistical model for analyzing panel data, how would you deal with this issue?

Answer one of the following:

3. Suppose you have collected tens of thousands of digital documents that you want to analyze. Some scholars would suggest that you look into computational methods. What

are your central options here? Avoiding jargon, explain what the relevant computational methods do, discuss their pros and cons, and reflect on which method(s) you would most likely pursue.

4. What is machine learning? What are the key differences between supervised and unsupervised approaches? How have sociologists used machine learning thus far? What is some promising potential uses of machine learning for scholars who want to advance social theory?

Answer one of the following:

5. One of the biggest changes in the past ten years in social science research methodology is the arrival of “big data” and applications for big data. Please answer the following questions:
 - a. What, exactly, is “big data” and how does it differ from conventionally collected social science data?
 - b. What are the strengths and weaknesses of big data in comparison to conventionally collected social science data?
6. Lieberman and others talk extensively about sample selection bias and sampling contamination as two problems that plague social science data analysis. If you answer:
 - a. Define what sample selection bias and sampling contamination is and provide a real-world example of each; and then
 - b. discuss potential solutions in the abstract and with reference to your examples.