

Tips for Collecting, Reviewing, and Analyzing Secondary Data

Secondary data analysis is the analysis of data or information that was gathered by someone else (researchers, other organizations, etc.), collected for some other purpose, or a combination of the two. Secondary data analysis is different than *primary data analysis* where the same individual or team designs, collects, and analyzes data for their own purposes. Below are advantages and disadvantages of secondary data.

Advantages	Disadvantages
<ul style="list-style-type: none">• Saves time and money because the study design and data collection are completed• Facilitates access to both large and small samples (in some cases)• Provides access to data that may be of high quality• Provides access to many different variables	<ul style="list-style-type: none">• Covers only the topics and samples it was originally designed for and not your question of interest• Risks being out of date for your needs or potentially inaccurate

The first step in secondary data analysis is establishing a purpose for the analysis. A clear understanding of why you are collecting the data and of what kind of data you want to collect, analyze, and better understand will help you remain focused and prevent you from becoming overwhelmed with the volume of data.

- **Purpose (Why):** To help you stay focused, your first step should be to state the overall goal of your research. Potential goals for your research might include:
 - Supporting Grant Applications and Other Funding Opportunities
 - Monitoring Trends and Identifying Emerging Problems and Needs
 - Disseminating Information to Engage Community and Policymakers
 - Evaluating Progress in Meeting Goals
 - Establishing Priorities and Planning Programs
 - Characterizing Disparities Across Sub Populations or Communities
- **Population (Who):** What demographic groups are you interested in?
- **Place (Where):** What is your geographic area of focus? The state? Your community?
- **Priority Areas (What):** What topic areas are you interested in exploring? Are you interested in issues around education? Health? Poverty?
- **Period (When):** What is the specific time frame of interest?

There are a variety of places to locate secondary data that might be relevant to your study. Below is a list of common sources.

- **Government Statistics:** Official statistics are statistics collected by governments and their various agencies, bureaus, and departments. These data are often easily obtainable and comprehensive sources of information. They often cover long periods of time. Nonetheless, it is important to critically analyze official statistics for accuracy and validity.

- **Private organizations:** Think tanks, independent oversight organizations, and research organizations all collect data and distribute information on various topics that may be relevant for your analysis. Some of these organizations, however, are not politically neutral. Nonetheless, many organizations with a political stance try to make their studies as objective as possible. It's important to be aware of any potential bias within the data you use as well as how its credibility may be perceived by your intended audience.
- **Technical Reports:** Technical reports are accounts of work done on research or evaluation projects. They are written to provide research and evaluation results to stakeholders, funders, community members, the research community, governments, and other interested parties. A report may reflect completed or on-going projects.
- **Scholarly Journals and literature reviews:** Scholarly journals generally contain reports of original research or experimentation written by experts in specific fields. Articles in scholarly journals usually undergo a peer review where other experts in the same field review the content of the article for accuracy, originality, and relevance. Literature reviews assemble and review original research dealing with a specific topic.

Make Use of Experts and Specialists When Possible!

When searching for secondary data or questioning the quality of a source that you have already collected, seek advice from the people most familiar with these and other types of secondary data. This may be an expert at a university, a specialist in a local government department, or a staff person in a research organization. They are valuable sources of information and expertise!

A key part of the secondary data analysis process is assessing the quality of the data that you intend to use. The following key questions will help you assess the quality of the data and its relevance for your analysis.

1. Who collected the data?
2. What was the original purpose for the data collected?
3. When was the data collected?
4. How was the data collected?
5. How were the variables defined?
6. Have the data collection methods and variable definitions changed over time?
7. What is the size of the sample?
8. Was the data collection process itself based on previous research?
9. Are the data collection instruments available?
10. In what form is the data available?
11. What are the geographic and demographic limitations of the data?
12. Were there any anomalies or challenges during data collection and, if so, what? How might they impact data quality?
13. What restrictions, including ethical concerns, are there in the use of the data?
14. Is the data in raw form or summarized into higher aggregates?
15. What are the costs, if any, of purchasing the data?