

Toolkit 2: DATA COLLECTION

Data collection is one of the first stages in any research project. The researcher might use data collected by others or collect them directly.

Where do I find data?

There are several data sources. Statistical public organisations, such as the [Office for National Statistics](#), provide datasets. Interactive tools, as [Nomis \(ONS\)](#), allows the user to choose the variable and characteristics through personalised datasets on the webpage. Our list of data sources and tools gives suggestions of repositories, directories, and other resources that might help you to find the data you need.

Using secondary data is the most common procedure, as it is less time-consuming. Nonetheless, previous datasets might not have all the information the researcher/policymakers need. In most cases, it is necessary to organise the dataset and adapt it to investigate the research questions. When we work with statistical software or programming languages such as R and Python, we refer to data collection as an “import” as we transfer the existing dataset to the system.

Data can also be collected directly from first-hand sources. For example, when researchers design and conduct questionnaires, they are collecting primary data.

Types of data and datasets

Regarding data collection, there are two types of data:

- Primary: when the researcher collects data directly from first-hand sources. Example: interviews and surveys with a group/individuals.
- Secondary: when the researcher works with data collected by other researchers or institutions (most usual). Example: databases made available by the university, think-tanks, or public organisations.

Database files might have different formats: XLSX, CSV, and DB, among others. CSV (Comma-Separated List) is the simplest file type. Large datasets might require other formats and tools.

References and further readings:

Statistics How To. [Primary Data & Secondary Data: Definition & Example - Statistics How To](#). *Primary Data & Secondary Data: Definition & Example*.

ONS: [Home - Office for National Statistics \(ons.gov.uk\)](#)

Nomis: [Nomis - Official Census and Labour Market Statistics \(nomisweb.co.uk\)](#)