

BA (H) GEOGRAPHY

SEMESTER IV

SEC: RESEARCH METHODS

Unit 2: – Data Collection: Methods of Collection

Unit 2: Data Collection: **Type and Sources of Data**; Methods of Collection; Input and Editing

Data are characteristics or information, usually numerical, that are collected through observation.^[1] In a more technical sense, data is a set of values of **qualitative** or **quantitative variables** about one or more persons or objects.

Data collection is the process of gathering and measuring information on targeted variables in an established system, which then enables one to answer relevant questions and evaluate outcomes.

Gathering data can be accomplished through a primary source (the researcher is the first person to obtain the data) or a secondary source (the researcher obtains the data that has already been collected by other sources, such as data disseminated in a scientific journal).

Data analysis methodologies vary and include data triangulation and data percolation. The latter offers an articulate method of collecting, classifying and analyzing data using five possible angles of analysis (at least three) in order to maximize the research's objectivity and permit an understanding of the phenomena under investigation as complete as possible: qualitative and quantitative methods, literature reviews (including scholarly articles), interviews with experts, and computer simulation. The data is thereafter "percolated" using a series of pre-determined steps so as to extract the most relevant information.

Importance of Data Collection

Regardless of the field of study or preference for defining data ([quantitative](#) or [qualitative](#)), accurate data collection is essential to maintaining the integrity of research. The selection of appropriate data collection instruments (existing, modified, or newly developed) and clearly delineated instructions for their correct use reduce the likelihood of [errors](#).

A formal data collection process is necessary as it ensures that the data gathered are both defined and accurate. This way, subsequent decisions based on arguments embodied in the findings are made using valid data.^[3] The process provides both a baseline from which to measure and in certain cases an indication of what to improve.

There are 5 common data collection methods; closed-ended surveys and quizzes, open-ended surveys and questionnaires, 1-on-1 interviews, focus groups, and direct observation.

DATA INTEGRITY ISSUES

The main reason for maintaining data integrity is to support the observation of errors in the data collection process. Those errors may be made intentionally (deliberate falsification) or non-intentionally (random or systematic errors).

There are two approaches that may protect data integrity and secure scientific validity of study results invented by Craddick, Crawford, Rhodes, Redican, Rukenbrod and Laws in 2003:

- Quality assurance – all actions carried out before data collection
- Quality control – all actions carried out during and after data collection

Quality assurance

Its main focus is prevention which is primarily a cost-effective activity to protect the integrity of data collection. Standardization of protocol best demonstrates this cost-effective activity, which is developed in a comprehensive and detailed procedures manual for data collection. The risk of failing to identify

problems and errors in the research process is evidently caused by poorly written guidelines. Listed are several examples of such failures:

- Uncertainty of timing, methods and identification of the responsible person
- Partial listing of items needed to be collected
- Vague description of data collection instruments instead of rigorous step-by-step instructions on administering tests
- Failure to recognize exact content and strategies for training and retraining staff members responsible for data collection
- Unclear instructions for using, making adjustments to, and calibrating data collection equipment
- No predetermined mechanism to document changes in procedures that occur during the investigation

Quality control

Since quality control actions occur during or after the data collection all the details are carefully documented. There is a necessity for a clearly defined communication structure as a precondition for establishing monitoring systems. Uncertainty about the flow of information is not recommended as a poorly organized communication structure leads to lax monitoring and can also limit the opportunities for detecting errors. Quality control is also responsible for the identification of actions necessary for correcting faulty data collection practices and also minimizing such future occurrences. A team is more likely to not realize the necessity to perform these actions if their procedures are written vaguely and are not based on feedback or education.

Data collection problems that necessitate prompt action:

- Systematic errors
- Violation of protocol
- Fraud or scientific misconduct

- Errors in individual data items
- Individual staff or site performance problems

TYPES OF DATA

There are two general types of data – quantitative and qualitative

QUANTITATIVE DATA

Quantitative data is information that you can measure.

It's numbers –something you can count. Because it's countable it can be reliable evidence.

Examples include:

How many people took part?

How much did it cost?

How long did it run for?

Average attendance at each programme session?

QUALITATIVE DATA

Qualitative data is information about qualities, you can't count it. That is, it's information about how people feel about something.

Examples include:

Sharing what people like about a programme.

How they think it could be improved.

What difference it has made to their lives.

Whether they would recommend the programme to others.

Quantitative Data vs Quantitative Data

Here are examples of sources of quantitative and qualitative data:

QUANTITATIVE DATA SOURCES

QUALITATIVE DATA SOURCES

Programme information – numbers involved, what happens when, who it is for, cost, participation levels.

Getting feedback on a programme – weekly check-ins, surveys, groups discussions, before-and-after programme questions.

Statistics and statutory data.

Stories of participants' experience and impact.

Environmental monitoring such as bird counts, seeds planted and fish stocks.

Observation of a programme.

Structured surveys involving multiple-choice, rating scales or other closed questions.

Document review such as literature reviews.

Control groups – comparing one group that experienced a programme with another similar group that did not.

Interviews involving open-ended questions with backers, service providers and participants.

Recording transactions, such as website hits.

Visual methods such as photos, DVDs and art or participants' creative works.

Qualitative and Quantitative Data: Scales of Measurement

Nominal – Qualitative Data

Ordinal – Qualitative Data

Interval - Quantitative Data

Ratio - Quantitative Data

CLASSIFICATION OF NUMERIC DATA: **Discrete and Continuous Data**

Discrete Data: These are data that can take only certain specific values rather than a range of values. For example, data on the blood group of a certain population or on their genders is termed as discrete data. A usual way to represent this is by using bar charts.

Continuous Data: These are data that can take values between a certain range with the highest and lowest values. The difference between the highest and lowest value is called the range of data. For example, the age of persons can take values even in decimals or so is the case of the height and weights of the students of your school. These are classified as continuous data. Continuous data can be tabulated in what is called a **frequency distribution**. They can be graphically represented using histograms.