Introduction

Data Analytics as a research tool involves collecting both quantitative and qualitative data and allows researchers to make evidence-based conclusions.

Data Analytics consists of the following phases:

1. Gather required data
2. Define data collection
3. Clean and organise data
4. Analyse data
5. Interpret data

Gather required data:

- Understand research objectives, research questions and expected outcomes to ensure:
    - Purpose of analysis
    - Type of data to collect
    - Understand measuring factors

This phase helps in constructing and defining data and variables of interest. Based on this the data model or the data framework is developed accordingly.

Define data Collection

- Appropriate data as defined is collected
- Data collection instruments like surveys, questionnaires and secondary data in the public domain are used.

Clean and Organise data

- Delete duplicate records, white spaces, outliers and irrelevant data to make it error free
- Transform raw data into variables that can be analysed.

Analyse data

- Involves a combination of statistical techniques, visualisation methods and machine learning algorithms.
- Statistical techniques include descriptive statistics and inferential statistics
- Descriptive analysis is the first level of analysis and helps in summarising data to find patterns
- Two types of statistics used for describing data are Measures of Central Tendency (Mean, Median and Mode) and Measures of Dispersion (Range, Variance and Standard Deviation).
- Inferential analysis depicts the relationship between multiple variables to generalise results and make predictions. Some of the techniques include correlation, regression and Analysis of variance.
- Data visualisation methods include charts and graphs. The data shown graphically is easier to understand and process. It helps to discover unknown

facts and trends. Some of the most used visualisation methods include bar chart, pie chart, line chart, scatter diagrams, box plots and bubble charts

Interpret data

- Post analysis results are interpreted and communicated through a report, a Table or a chart.
- R and Python are two of the most used statistical programming languages for analysis. They have libraries for visualisation as well. Tableau has powerful features for visualisation. Microsoft Excel is also popular for both statistical techniques and visualisation methods.