

IT Skills & Data Analysis I

SYLLABUS

Practical

Unit I : What is Statistics ? (24 hours)

This unit provides an introduction to the fundamentals of datasets, sources of data, frequency distributions and graphical representations of data. The aim is to give students a hands-on experience of initiating data analysis through a spreadsheet.

- Concept of datasets (Variables, Observations)
Reference 1, Chapter 2
- Different types of variables (Quantitative and Qualitative)
Reference 1, Chapter 2
- Distinction between primary and secondary sources of data
Reference 1, Chapter 2
- Basic idea of using questionnaire to collect primary data for analysis
Reference 2, Chapter 1 [Section 1.6]
- How to construct a questionnaire
Reference 1, Chapter 1
- Concept of frequency distribution: cumulative and relative frequencies
Reference 2, Chapter 2
- Introduction to spreadsheet
Reference 2, Chapter 2
- Tabular and graphical presentation of data: data tables, frequency curve, histogram, bar graphs, pie charts (through the use of spreadsheets)
Reference 2, Chapter 2

Unit II: Measures of Central Tendency and Dispersion (36 hours)

The focus of this unit will be to familiarise the student with summary statistics to describe datasets. In particular, two important characteristics of data, viz., central tendency and dispersion, will be used to summarise datasets using a spreadsheet. The concept of the Normal distribution and its characteristics will be discussed to highlight its relevance in modelling real life phenomenon.

- Measures of central tendency: mean, median, mode
Reference 2, Chapter 3
- Examples of situations where it is appropriate to use the mean, median and mode as a measure of central tendency
Reference 2, Chapter 3
- Weighted mean
Reference 2, Chapter 3
- Measures of dispersion: range, variance, standard deviation
Reference 2, Chapter 3

- Quartiles, deciles and percentiles
Reference 2, Chapter 3
- Visualize the measures of central tendency and dispersion through frequency curve and histogram
Reference 2, Chapter 3
- Skewness and kurtosis
Reference 2, Chapter 3
- Normal curve and its basic properties : visual representation of population characteristics (height, weight, IQ etc.)
Reference 2, Chapter 5 [Section 5.6]

References (Readings and Resources)

1. Rowntree, D., Statistics without tears - A primer for non-mathematicians, Allyn and Bacon, 2018.
2. Levin, Rubin, Rastogi and Siddiqui, Statistics for Management, 7th Edn, 2014

Suggested Data Sources

The following data sets are suggested to carry out the activities

1. <https://data.worldbank.org/>
2. <https://www.statista.com/>
3. <https://data.gov.in/>
4. <https://censusindia.gov.in/>
5. <https://www.kaggle.com/>
6. <http://data.un.org/>

