# Assignment Solutions – Unit 1 – Introduction to Data Analytics

Model answers:

1. Define Data Analytics and explain its importance in modern business.

Data Analytics refers to the systematic process of examining datasets to uncover patterns, correlations and insights. By applying statistical and computational techniques, analysts turn raw data into useful information. In business contexts, Data Analytics helps organisations make evidence‑based decisions, improve operational efficiency and gain a competitive advantage.

1. Differentiate between Descriptive, Diagnostic, Predictive and Prescriptive analytics.

Descriptive analytics summarises historical data to understand what happened. Diagnostic analytics investigates why something happened by identifying causal relationships. Predictive analytics uses statistical models and machine learning to forecast what could happen in the future. Prescriptive analytics suggests actions to take based on predictive insights to achieve desired outcomes.

1. Outline the Data Analytics life cycle and explain its stages with examples.

A typical Data Analytics life cycle consists of: (a) Problem definition – identify the question to be answered; (b) Data collection – acquire relevant data from surveys, sensors or databases; (c) Data cleaning and preprocessing – handle missing values and outliers; (d) Data analysis – apply statistical methods or models; (e) Visualisation – present insights using charts or dashboards; (f) Decision and deployment – use the insights to guide actions. For example, an e‑commerce company may analyse customer purchase data to determine which products to recommend.

1. Discuss data types and identify at least three measures of central tendency.

Data can be qualitative (categorical, such as gender or colour) or quantitative (numeric, such as height or temperature). Measures of central tendency describe the typical value of a dataset. Three common measures are: (1) the mean (arithmetic average), (2) the median (middle value when data are ordered), and (3) the mode (most frequently occurring value).

1. A sample dataset consists of {10, 15, 20, 20, 25, 30}. Compute the mean, median and mode of the dataset.

Mean = (10 + 15 + 20 + 20 + 25 + 30)/6 = 20. Median: the ordered set is {10, 15, 20, 20, 25, 30}. The two middle values are 20 and 20, so the median is 20. Mode is 20 because it appears twice.