

Excel Assignment 1

Name: Parth Desai

Q1) What is data and why is it important in the modern world?

→ Data is raw information that helps in making decision. It can be numbers, text, images or even videos.

The data is important in the modern world because -

- i. Helps in decision making: Companies use data to identify trends and make strategic decision.
- ii. Identifies Pattern and Trends: Helps in understanding customers behaviour and market trends.
- iii. Automates Processes: Business use data to streamline operation and reduce manual effort.

Q2) What is the difference qualitative and quantitative data?

→ Qualitative Data is the data that represents non-numeric information. It describe categories or label.

Eg: Colors of products → Red, Blue, Green etc.

Customer Feedback: Positive, Negative, Neutral.

Types of Department: HR, IT, Sales

Whereas Quantitative Data consists of numbers and can be measured or counted.

Eg:

Number of Employees in a company.

Height of students in a Class.

Q3 what is data analysis and how does it differ from data science?

→ Data Analysis is the process of inspecting, cleaning, transforming and modeling data to discover useful information, identify patterns and support decision making.

Key Differences

i. Scope: Data analysis is highly focused and question driven, providing actionable insights for existing queries; data science is broader and often deals with structured data and unknown questions.

ii. Technique: Data analysis relies primarily on traditional statistics and visualization, while data science uses ML, predictive analysis, and programming to build advanced models.

Q4 what are the key steps involved in a typical data analysis process.

→ Data Analysis means converting raw data into useful knowledge through structured steps.

Steps Involved:

1. Collecting Data: From spreadsheet, surveys, databases etc.

2. Cleaning Data: Fixing errors and filling missing values.

3. Analyzing Data: Using basic statistics to discover trends and pattern.

4. Visualizing Data: Presenting findings with chart and dashboards.

Q5 What Tools and Techniques are commonly used in data science?

→ Common Tools used in Data Science.

i. Programming Languages: Python, R, SQL are foundations for data handling, analysis and building models.

ii. Visualization Tools: Tableau, Power BI, Matplotlib, Seaborn enables to creates charts and dashboards.

iii. ML Libraries: SciKit-Learn, TensorFlow, Keras provides frameworks for building and training predictive models.

iv. Big Data Tools: Hadoop, Apache Spark and Snowflake handle vast volume of structured and unstructured data.

v. Statistical Tools: Pandas, Numpy and Excel support statistical analysis, regression and modeling.

Techniques in Data Science.

i. Data Cleaning and preprocessing: Removing errors, handling missing values and preparing data for analysis.

ii. EDA: Using descriptive statistics and visualization to understand dataset and uncover initial pattern.

iii. Machine Learning Modeling: Creating predictive algorithm using libraries.

iv. Statistical Analysis: Probability, regression, hypothesis testing and inferential statistics help draw conclusion.

v. Big data Analysis: Leveraging distributed computing and storage solution for analyzing large scale datasets.