DATA ANALYTICS

BRANCH:- OPEN ELECTIVE

SEMESTER – FIFTH

These important questions have been prepared using your previous exam papers (PYQs), verified concepts, and additional reference from trusted online academic sources.
For deeper understanding, please refer to your class notes as well.

1️⃣ HIGH & LONG IMPORTANT QUESTIONS

Unit I: Introduction

1. **What is Data Analytics?** Explain its importance and describe the **Life Cycle of Data Analytics** in detail.
2. Explain the different **Types of Data Analytics** (Descriptive, Diagnostic, Predictive, Prescriptive) with examples.
3. Define **Measures of Central Tendency** (Mean, Median, Mode) and **Measures of Dispersion** (Range, Variance, Standard Deviation). Why are they used?

Unit II: Statistical Analysis

4. What is the difference between **Correlation and Regression**? Explain with the help of a **Scatter Diagram**.
5. Explain **Hypothesis Testing**. What are the steps involved in testing a hypothesis?
6. What is **Data Cleaning**? Explain different **Imputation Techniques** used to handle missing data.

Unit III: Data Analytics with Excel

7. What is a **Pivot Table**? Write the step-by-step procedure to **Create, Filter, and Sort** a Pivot Table in Excel.
8. What is an **Excel Dashboard**? Explain its key features like Tables, Data Grids, and Dynamic Filters.

Unit IV: Data Visualization

9. Explain the **Chart Wizard** in Excel. Write the steps to create a simple **Bar Chart** and modify it.
10. Describe the various components of a Chart (Legend, Plot Area, Data Markers, Axes) and how to format them.

Unit V: Data Visualization using Python

11. What is **Matplotlib**? Explain its role in data visualization. Write the steps to **install and set up** Matplotlib in Python.
12. How can you create a **Line Plot** and a **Histogram** using Matplotlib? Explain with basic steps/syntax and how to add titles and labels.

2️⃣ IMPORTANT & SHORT QUESTIONS

Unit I

- Define **Central Limit Theorem (CLT)**.
- What is the **Sampling Funnel**?
- Differentiate between **Quantitative and Qualitative Data**.
- What is a **Confidence Interval**?

Unit II

- Write a short note on **ANOVA**.
- Differentiate between **Skewness and Kurtosis**.
- What is a **Box Plot**? What does it represent?
- Define **Probability Distribution**.

Unit III

- What are **Slicers** in Excel? How are they useful?
- How do you **Group Items** in a Pivot Table?
- Explain **Trend Analysis** in brief.

Unit IV

- How do you **Explode a Slice** of a Pie Chart?

- Write steps to **Move an Embedded Chart** to a new sheet.
- How to change the **Chart Type** (e.g., from Bar to Line) after creating it?

Unit V

- Write the command/steps to **Export and Save** a plot as a PNG or PDF in Matplotlib.
- How to change the **Figure Size** and **Aspect Ratio** in Matplotlib?
- Differentiate between a **Scatter Plot** and a **Bar Chart**.

3️⃣ "AA BHI SAKTA HAI" QUESTIONS (20–30% Probability)

- **Unit I:** Explain **Sampling Variation**. Why does it occur?
- **Unit II:** What is the **Chi-Square Test**? When is it used?
- **Unit III:** How to perform **Forecasting** using Excel Dashboards?
- **Unit IV:** How can you chart **Non-Adjacent Cells** in Excel? (Tricky step-based question).
- **Unit V:** Write a short note on creating **Interactive Visualizations** using Matplotlib widgets.

💡 Analyst's Tip for Students:

Since **Units III, IV, and V** are tool-based (Excel & Python), JUT often asks for **"Steps to..." (Iske steps likho)**.

- Exam me agar computer nahi hai, tab bhi aapko **Steps bullet points me likhne honge** (e.g., Step 1: Select Data -> Step 2: Click Insert Tab...).
- **Draw rough diagrams** of Excel windows or Graphs wherever possible to get full marks!

QUICK REVISE

UNIT I: Introduction to Data Analytics

- **Data Analytics:** The process of examining raw data to find trends, draw conclusions, and support decision-making.
- **4 Types of Analytics (Most Imp):**
    1. **Descriptive:** "What happened?" (Past data).
    2. **Diagnostic:** "Why did it happen?" (Root cause).
    3. **Predictive:** "What is likely to happen?" (Forecasting).
    4. **Prescriptive:** "What should we do?" (Recommendation).
- **Life Cycle:** Problem Definition > Data Collection > Data Cleaning > Analysis > Interpretation.
- **Central Tendency (Averages):**
    - **Mean:** Average value.
    - **Median:** Middle value.
    - **Mode:** Most repeating value.
- **Dispersion:** How spread out the data is (Range, Variance, Standard Deviation).
- **Central Limit Theorem (CLT):** If you take large samples, their mean will look like a **Normal Distribution** (Bell Curve), regardless of the original data shape.

UNIT II: Statistical Analysis

- **Box Plot:** A chart showing 5 things: Minimum, Q1 (25%), Median, Q3 (75%), Maximum. Good for finding outliers.
- **Skewness vs. Kurtosis:**
    - **Skewness:** Measure of asymmetry (Is data leaning left or right?).
    - **Kurtosis:** Measure of "tailedness" or peak (Is data flat or sharp?).
- **Correlation:** Relationship between two variables (e.g., Study time vs. Marks).
    - *Positive:* Both go up.
    - *Negative:* One goes up, other goes down.
- **Regression:** Used to **predict** the value of a dependent variable based on an independent variable.
- **Data Cleaning:** Removing errors, duplicates, or fixing missing values.
- **Imputation:** Filling missing data (e.g., replacing empty cells with the Mean).
- **Hypothesis Testing:** Testing an assumption.
    - *Null Hypothesis* No difference exists.
    - *Alternate Hypothesis :* A difference exists.

UNIT III: Data Analytics with Excel (Tools)

- **Excel Dashboard:** A single screen showing key data points, charts, and tables for quick analysis.

- **Pivot Table:** A tool to **summarize** large datasets without using formulas. You can drag-and-drop rows and columns.
- **Steps to Create Pivot Table:**
    1. Select Data.
    2. Insert Tab > Click **PivotTable**.
    3. Choose destination (New Sheet).
    4. Drag fields to Rows, Columns, and Values areas.
- **Slicers:** Visual buttons used to **filter** Pivot Tables quickly.
- **Trend Analysis:** Looking at past data to predict future trends (linear, growth).

UNIT IV: Data Visualization (Charts)

- **Chart Wizard:** A step-by-step guide in older Excel versions to create graphs.
- **Common Chart Types:**
    - **Column/Bar Chart:** Comparing categories.
    - **Line Chart:** Showing trends over time (e.g., Sales per month).
    - **Pie Chart:** Showing parts of a whole (Percentage).
- **Chart Components:**
    - **Legend:** Explains what colors represent.
    - **Plot Area:** The area where the graph is drawn.
    - **Data Labels:** The actual numbers on the bars/lines.
- **Exploding a Pie Chart:** Separating one slice from the pie to highlight it.
- **Embedded Chart:** A chart that floats on the worksheet (can be moved or resized).

UNIT V: Data Visualization using Python

- **Matplotlib:** A Python library used for 2D plotting and visualization.
- **Installation:** Command is pip install matplotlib.
- **Key Plotting Functions:**
    - plt.plot() > Line plot.
    - plt.bar() > Bar chart.
    - plt.scatter() > Scatter plot (dots).
    - plt.hist() > Histogram.
- **Customization:**
    - plt.title("Name") > Adds Title.
    - plt.xlabel("X name") > Labels X-axis.
    - plt.show() > Displays the graph.
- **Exporting:** Saving plots using plt.savefig('filename.png').

📝 **Exam-Writing Tips for JUT (To Score High)**

1. **Definitions First:** Start every answer with a clear definition (Data Analytics is...).
2. **Write Steps:** For Unit 3, 4, & 5, always write steps in bullet points (Step 1, Step 2...).

3. **Draw Diagrams:**
    - Draw a rough **Bar Chart** or **Pie Chart** with pencil.
    - Draw the **Data Analytics Lifecycle** (Circle diagram).
    - Draw **Skewness curves** (Bell shape).
4. **Difference Between:** Use a **Table** format for "Correlation vs Regression" or "Mean vs Median".

Study Hard! All the best! 🚀

DIPLOMA WALLAH (SWANGAM 💗)