

1. Define PivotTable. List any three advantages of using PivotTables in Excel.

Answer:

- A PivotTable is a data analysis tool in Excel that helps to **summarize, analyze, and reorganize** large datasets.

Advantages:

1. Summarizes large datasets quickly.
2. Allows filtering, sorting, and grouping.
3. Supports calculated fields and items for deeper analysis.

2. Explain the steps involved in creating a PivotChart from a PivotTable.

Answer:

Steps:

1. Select the PivotTable.
2. Go to **Insert** → **PivotChart**.
3. Choose the chart type (Column, Line, etc.).
4. Click **OK** → PivotChart is created and linked to the PivotTable.

3. Differentiate between Calculated Fields and Calculated Items in PivotTables with examples.

Answer:

- **Calculated Field** – Uses existing fields across the dataset.
Example: Profit Margin = (Sales – Cost) / Sales.
- **Calculated Item** – Performs calculations on items within the same field.
Example: New Region = North + South.

4. What are the different types of grouping in PivotTables? Explain any two with examples.

Answer:

Types of grouping:

1. **By Date** – Group sales by months/quarters.
2. **By Numeric Ranges** – Group ages into 0–18, 19–30, etc.
3. **By Text** – Group products by first letter.

5. Write the steps to create a PivotTable showing total sales by region and category.

Answer:

1. Select dataset → Insert → PivotTable.
2. Drag **Region** → Rows, **Category** → Columns.
3. Drag **Sales Amount** → Values.
4. PivotTable will show total sales by region and category.

6. State the importance of data cleaning (any three points).

Answer:

- Ensures accurate analysis and decision-making.
- Enhances data integrity and reliability.
- Saves time and prevents reporting errors.

7. Explain the types of missing data: MCAR, MAR, MNAR.

Answer:

1. **MCAR** – Missing completely at random, e.g., accidental skips.

2. **MAR** – Missing at random, depends on observed data, e.g., older patients missing visits.
3. **MNAR** – Missing not at random, depends on unobserved data, e.g., high-income people not disclosing salary.

8. Write Excel formulas for handling missing values.

Answer:

- a) Identify blank cells → `=COUNTBLANK(range)`
- b) Mean imputation → `=AVERAGE(range)`
- c) Mode imputation → `=MODE.SNGL(range)`
- d) Last Observation Carried Forward → `=IF(ISBLANK(A2), A1, A2)`

9. Deletion vs. Imputation methods for missing values.

Answer:

- **Deletion:** Removes missing data.
Advantage: Simple; *Disadvantage:* Loss of data.
- **Imputation:** Replaces missing values.
Advantage: Preserves dataset size; *Disadvantage:* May reduce accuracy.

10. Case Study – BookMart Sales Analysis.

Answer:

- **Seasonal Trends:** Group dates → months/quarters.
- **Store Performance:** Compare total sales by store.
- **Genre Popularity:** Analyze which genres sell most.
- **Sales Targets:** Use calculated fields (e.g., Target = Sales × 1.1).