Module 5: Data Vizualization With Power Bi

1) What is Power BI and how does it differ from Excel? Power BI is a tool by Microsoft to create reports and dashboards. Excel is mainly for spreadsheets and data entry. The main difference is that Power BI can handle very large data and is better for sharing online interactive reports, while Excel is good for small calculations.

2) Explain the concept of data modeling in Power BI. Data modeling basically means connecting different tables together. If we have data in two different files, we create a relationship between them using a common column (like ID) so we can use both in one chart.

3) What are the different types of connections available in Power BI? There are mainly three types: Import mode (brings data inside), Direct Query (keeps data in database), and Live Connection (connects to existing models).

4) How do you handle data transformation in Power BI? We use "Power Query Editor" for this. Before loading data into the dashboard, we can remove errors, split columns, or change data types here.

5) What is DAX (Data Analysis Expressions) and why is it important in Power BI? DAX is a formula language used in Power BI, just like Excel formulas. It is important because it helps us create custom calculations like "Total Sales" or "Year over Year growth" which are not available directly.

6) Can you explain the difference between calculated columns and measures in Power BI? A calculated column saves the answer in every row of the table and takes memory space. A Measure is dynamic; it only calculates the number when we put it in a graph, so it is faster and doesn't take space.

7) How do you handle relationships between tables in Power BI? We go to the "Model View" and drag a line between two tables using a matching key column. Usually, we use a One-to-Many relationship.

8) What is the purpose of a Power BI Gateway? It acts like a bridge. If our data is on a local laptop (on-premise) and the report is online, the Gateway helps to refresh the online data securely.

9) How can you schedule data refresh in Power BI Service? After publishing the report, we go to the dataset settings in Power BI Service. There is an option for "Scheduled refresh" where we can set the time (like 9 AM daily) to update the data automatically.

10) Explain the concept of row-level security in Power BI. RLS is used to restrict data. For example, if I want a manager from Gujarat to see only Gujarat's sales and not Mumbai's, I can set up RLS roles to filter data based on the user.

11) What is the Power BI Desktop and how does it differ from Power BI Service? Desktop is the software we install on the PC to build the reports. Service is the online website (SaaS) where we share and view the reports.

12) Explain the concept of Direct Query in Power BI. In Direct Query, Power BI doesn't copy the data. It queries the database every time we click on a chart. It is good for real-time data but can be slow.

13) What are Power BI templates and how are they useful? Templates are files that have the report design and queries but no actual data. They are useful if we want to share the layout with someone without sending the whole dataset.

14) How do you handle incremental data refresh in Power BI? Incremental refresh means only loading the new data (like yesterday's data) instead of reloading the whole 5 years of data every time. This makes refreshing very fast.

15) What is the role of Power Query in Power BI? Power Query is the tool for ETL (Extract, Transform, Load). It is used to get data from sources and clean it up before we use it for visualization.

16) Explain the difference between calculated columns and calculated tables in Power BI. Calculated column adds a new column to an existing table. Calculated table creates a completely new separate table using DAX formulas.

17) How do you create custom visuals in Power BI? If the default charts are not enough, we can import custom visuals from the "AppSource" marketplace inside Power BI.

18) What are the best practices for optimizing performance in Power BI? We should limit the number of visuals on one page, remove unused columns in Power Query, and use Measures instead of Calculated columns wherever possible.

19) How can you integrate Power BI with other Microsoft products like Azure and Office 365? Since it's a Microsoft tool, it connects easily. We can embed reports in Microsoft Teams, analyze data in Excel, or get data from Azure SQL databases directly.

20) Explain the concept of aggregations in Power BI. Aggregations are like summary tables. Instead of scanning millions of rows, Power BI uses a small summary table to answer questions faster, which improves performance.

21) How do you handle error handling and data quality in Power BI? We can check column quality in Power Query to find errors. We can use options like "Remove Errors" or "Replace Errors" to fix bad data before loading it.

22) What is the purpose of Power BI Embedded and when would you use it? Power BI Embedded is used by developers to put Power BI

reports inside their own custom apps or websites for their customers to see.