

Data Visualization – Power BI

1. Answer the following questions (15 Marks)

a. Explain qualitative and quantitative data in brief with examples.

Qualitative and quantitative data are two primary types of data used in research and analysis. Qualitative data describes qualities or characteristics and is typically non-numeric. It focuses on understanding concepts, opinions, or experiences and is often collected through interviews, observations, or open-ended surveys. For example, customer reviews describing a product as “excellent” or “poor” represent qualitative data. On the other hand, quantitative data deals with numbers and measurable forms. It answers questions like how much, how many, or how often, and is collected through structured methods such as surveys or experiments. For instance, recording the number of products sold in a month or measuring a person’s height in centimeters are examples of quantitative data. Both types of data are important, with qualitative data providing depth and context, while quantitative data offers measurable and comparable results.

b. Explain descriptive analytics in brief with examples.

Descriptive analytics help answer questions about what has happened based on historical data. Descriptive analytics techniques summarize large semantic models to describe outcomes to stakeholders.

By developing key performance indicators (KPIs), these strategies can help track the success or failure of key objectives. Metrics such as return on investment (ROI) are used in many industries, and specialized metrics are developed to track performance in specific industries.

An example of descriptive analytics is generating reports to provide a view of an organization's sales and financial data.

c. When you decide to work with Power BI for report or dashboard creation what is the typical workflow that needs to be followed?

To create reports with Power BI, you must first understand the necessary tools. There are three primary components to Power BI:

- Power BI Desktop (desktop application)
- Power BI service (online platform)
- Power BI Mobile (cross-platform mobile app)

Power BI Desktop is the development tool available to data analysts and other report creators. While the Power BI service allows you to organize, manage, and distribute your reports and other Power BI items. Power BI Desktop is available to download for free either through the Windows store or directly online.

There's a common flow when creating reports with Power BI. First, you start with Power BI Desktop to connect to data and create the report. Then you publish the report to the Power BI service and distribute it to consumers.

The flow of Power BI is:

- Connect to data with Power BI Desktop.
- Transform and model data with Power BI Desktop.
- Create visualizations and reports with Power BI Desktop.
- Publish report to Power BI service.
- Distribute and manage reports in the Power BI service.

The Power BI service also allows you to create high-level dashboards that drill down to reports, and apps to easily group related reports to users in a simple format.

2. Answer the following questions (15 Marks)

a. State and explain any 3 sources of data that you can connect to using Power BI

- Data from CSV files
Organizations often export and store data in files. One possible file format is a flat file. A flat file is a type of file that has only one data table and every row of data is in the same structure. The file doesn't contain hierarchies. Likely, you're familiar with the most common types of flat files, which are comma-separated values (.csv) files, delimited text (.txt) files, and fixed width files.
- Data from Excel files
Another type of file would be the output files from different applications, like Microsoft Excel workbooks (xlsx).
Your Excel files might exist in one of the following locations:
 - o Local
 - o OneDrive for Business
 - o OneDrive – Personal
 - o SharePoint – Team Sites
- Data from SQL server
If your organization uses a relational database for sales, you can use Power BI Desktop to connect directly to the database instead of using exported flat files.
Connecting Power BI to your database will help you to monitor the progress of your business and identify trends, so you can forecast sales figures, plan budgets and set performance indicators and targets. Power BI Desktop can connect to many relational databases that are either in the cloud or on-premises.

b. Differentiate between appending and merging a query in Power Query.

Appending and merging are two different methods used in Power Query to combine data from multiple queries or tables. Here's the difference between them:

Appending a query means stacking data from two or more queries on top of each other. It is used when you have datasets with the same structure (same columns) and you want to combine their rows into a single table. For example, if you have sales data from January and February in separate tables, you can append them to create one combined table with all the sales data.

Merging a query means joining two queries side by side based on a common column, like a JOIN operation in databases. It is used when you need to bring in related data from different tables. For example, if you have a table with customer orders and another with customer details, you can merge them based on the Customer ID to add customer information to the orders table.

In short, append = combine rows, while merge = combine columns based on a key.

- c. Assign data types for the following variables/fields in Power Query.
 - i. Invoice ID - Text
 - ii. Country - Text
 - iii. Product Category - Text
 - iv. Order Date - Date
 - v. Quantity Sold – Whole Number

3. Answer the following questions (15 Marks)

- a. Explain the concept of granularity while working with any BI tool.

Granularity in Business Intelligence (BI) refers to the level of detail or depth of data stored and analyzed in a dataset or report. It determines how detailed or summarized the data is. High granularity means the data is very detailed (e.g., sales data recorded for each transaction or each minute), while low granularity means the data is more summarized (e.g., total sales per month or per region).

Choosing the right granularity is important in BI tools like Power BI or Tableau because it affects the accuracy and usefulness of the analysis. For example, if a company analyzes daily sales data, the granularity is per day. If they summarize it by month, the granularity is per month. Proper granularity ensures that users get the right insights at the right level, balancing detail with simplicity and performance.

- b. Differentiate between calculated columns and measures.

Regarding similarities between calculated columns and measures, both are:

- Calculations that you can add to your semantic model.
- Defined by using a DAX formula.
- Referenced in DAX formulas by enclosing their names within square brackets.

The areas where calculated columns and measures differ include:

- Purpose - Calculated columns extend a table with a new column, while measures define how to summarize model data.
- Evaluation - Calculated columns are evaluated by using row context at data refresh time, while measures are evaluated by using filter context at query time. Filter context is introduced in a later module; it's an important topic to understand and master so that you can achieve complex summarizations.
- Storage - Calculated columns (in Import storage mode tables) store a value for each row in the table, but a measure never stores values in the model.
- Visual use - Calculated columns (like any column) can be used to filter, group, or summarize (as an implicit measure), whereas measures are designed to summarize.

c. Explain with an example what simple measures are.

Simple measures are basic calculations or aggregations performed on data fields in Business Intelligence (BI) tools like Power BI or Tableau. These measures are typically created using standard aggregation functions such as SUM, AVERAGE, COUNT, MIN, and MAX. They help summarize data to provide key insights without complex formulas or logic.

For example, in a sales dataset, a simple measure could be Total Sales, which calculates the sum of all sales amounts using the SUM function. If there's a column called "Sales Amount," a simple measure would be:

Total Sales = SUM(Sales Amount)

Another example is calculating the average order value, which uses the AVERAGE function on the sales amounts to find the mean value of all orders.

Simple measures provide quick, straightforward insights into data and are essential for reporting and basic analysis.

4. Case Study