Question 1 :- What is Power BI and why is it used in businesses?

Answer :- **Power BI** is a collection of software services, apps, and connectors developed by Microsoft that work together to turn unrelated sources of data into coherent, visually immersive, and interactive insights.
In simpler terms, it is a business analytics tool that lets you visualize your data and share insights across an organization, or embed them in an app or website

Why Businesses Use Power BI

Businesses primarily use Power BI to make data-driven decisions rather than relying on intuition. Here are the specific reasons for its popularity:

- **Data Consolidation:** It connects to hundreds of data sources (like Excel, SQL Server, Salesforce, Google Analytics, and cloud services) and brings all that data into a single view.
- **Interactive Visualizations:** Unlike static spreadsheets, Power BI reports are interactive. Users can click on a specific region, product, or time period in one graph, and all other graphs on the dashboard will automatically filter to show data relevant to that selection.
- **Real-Time Analytics:** It supports real-time data streaming, meaning dashboards can update instantly as data changes (crucial for things like stock market tracking or manufacturing monitoring).
- **Accessibility:** It allows users to access reports and dashboards on the go via robust mobile apps for iOS, Android, and Windows.
- **Microsoft Integration:** Since it is a Microsoft product, it integrates seamlessly with Excel, Azure, and Teams, making it very easy for organizations already using the Microsoft ecosystem to adopt.
- **Self-Service Analytics:** It is designed to be user-friendly, allowing non-technical business users to create their own reports and answer their own questions without constantly relying on IT staff.

Question 2 :- Name and explain the three main components of Power BI.

Answer :- The three main components of Power BI are **Power BI Desktop**, **Power BI Service**, and **Power BI Mobile**.

Here is an explanation of each component and its specific role:

1. Power BI Desktop (The Authoring Tool)

This is a free Windows application where the actual report creation happens. It is the workspace for developers and analysts.

- **Primary Use:** You use it to connect to data sources, clean and transform data (using Power Query), build data models, and design the actual visual reports.
- **Key Feature:** This is the only place where you can calculate complex formulas (DAX) and structure the data relationships. It is effectively the "kitchen" where the meal is cooked.

2. Power BI Service (The Cloud Platform)

This is an online SaaS (Software as a Service) platform where you publish your work after building it in the Desktop version.

- **Primary Use:** It is used for sharing reports with stakeholders, collaborating with team members, and creating "Dashboards" (which are single-page summaries often combining visuals from multiple reports).
- **Key Feature:** It allows for automated data refreshing and setting up email alerts. It is the "dining room" where the meal is served to guests.

3. Power BI Mobile (The Consumption App)

These are native mobile applications available for Windows, iOS, and Android devices.

- **Primary Use:** They allow users to view and interact with reports and dashboards on smartphones or tablets.
- **Key Feature:** The apps are optimized for touchscreens, allowing executives and managers to monitor Key Performance Indicators (KPIs) and business health securely from anywhere, without needing a computer.

Question 3 :- Explain the Power BI workflow.

Answer :- It is generally divided into three main phases: **Preparation** (Desktop), **Publication** (Service), and **Consumption** (Mobile/Web).

Step 1: Connect to Data (Get Data)

Everything begins in **Power BI Desktop**. You use the "Get Data" feature to connect to various data sources.

- **Action:** Connect to databases (like SQL Server), flat files (Excel, CSV), or cloud services (Azure, Salesforce).

Step 2: Transform and Clean Data (The ETL Phase)

Once connected, the data is rarely ready for analysis immediately. You use the **Power Query Editor** to clean it.

- **Action:** Remove duplicates, change data types, split columns, and merge tables.
- **Context:** This is essentially the **ETL** (Extract, Transform, Load) stage of the process where raw data is shaped into a usable format.

Step 3: Model the Data

After cleaning, you load the data into the Power BI model.

- **Action:** Define **Relationships** between different tables (e.g., connecting a "Sales" table to a "Date" table).
- **Calculations:** Create measures and calculated columns using **DAX** (Data Analysis Expressions) for complex analysis, such as "Year-over-Year Growth."

Step 4: Build the Report

This is the visual phase where you design the layout.

- **Action:** Drag and drop fields onto the canvas to create charts, maps, and slicers. You arrange these visuals on multiple pages to tell a data story.

Step 5: Publish to Power BI Service

Once the report is complete in the Desktop app, you publish it to the cloud.

- **Action:** Click the "Publish" button. The report is uploaded to your workspace in the **Power BI Service** (app.powerbi.com).

Step 6: Create Dashboards and Share

In the cloud service, you refine how others see the data.

- **Action:** Pin key visuals from your report to a **Dashboard** (a single-page executive view).
- **Sharing:** Share the dashboard with colleagues or set up scheduled data refreshes so the numbers stay current.

Question 4 :- List any four data cleaning tasks that can be performed in Power Query.

Answer :- Here are four common data cleaning tasks performed in Power Query Editor:

1. **Promoting Headers (Use First Row as Headers)**

- ○ **What it is:** Often, when data is imported from Excel or CSV files, the actual column names appear in the first row of the data rather than the header section.
- ○ **The Task:** You use the "Use First Row as Headers" command to move that first row up, ensuring the columns are correctly labeled for analysis.

2. **Changing Data Types**
   - ○ **What it is:** Data might load as "Text" when it should be "Date" or "Currency."
   - ○ **The Task:** You explicitly define the data type for each column (e.g., changing a "Sales" column from *Text* to *Decimal Number*). This is critical because you cannot perform math calculations on text fields.

3. **Splitting Columns**
   - ○ **What it is:** You might receive a column like "Full Name" that contains both the first and last names, or an "Address" column that combines city and state.
   - ○ **The Task:** You can split one column into two or more based on a **delimiter** (like a comma, space, or hyphen). For example, splitting "John Doe" into "John" and "Doe" columns.

4. **Removing Rows (Top/Bottom/Alternate)**
   - ○ **What it is:** Raw reports often contain title information in the first few lines or summary disclaimers in the footer.
   - ○ **The Task:** You can specify a number of rows to remove from the top or bottom of the dataset to get rid of this irrelevant metadata, leaving only the clean data table.

Question 5 :- Write step by step instructions to load the above dataset.

Answer :- Here are the step-by-step instructions to load the `Real_Estate_Data_100_Rows.csv` dataset into Microsoft Power BI:

1. **Open Power BI Desktop**: Launch the application on your computer.
2. **Get Data**:
   - ○ On the **Home** ribbon tab, look for the **Data** section.
   - ○ Click on the **Get Data** icon. A drop-down menu or a dialog box will appear.
3. **Select Source Type**:
   - ○ In the list of common data sources, select **Text/CSV**.
   - ○ If you don't see it immediately, you can click "More..." at the bottom, select "File" from the left sidebar, and then choose "Text/CSV". Click **Connect**.
4. **Locate the File**:
   - ○ A file explorer window will open. Navigate to the folder where you have saved the `Real_Estate_Data_100_Rows.csv` file.
   - ○ Select the file and click **Open**.
5. **Preview and Configure**:
   - ○ Power BI will open a preview window showing the first few rows of your data.
   - ○ **File Origin**: Usually defaults to `1252: Western European (Windows)` or `65001: Unicode (UTF-8)`. You can leave this as is.

- ○ **Delimiter**: Ensure this is set to **Comma**, as your file is a CSV.
- ○ **Data Type Detection**: Leave as "Based on first 200 rows".
6. **Load or Transform**:
   - ○ **Load**: Click this if the data looks correct in the preview and you want to import it directly into your report.
   - ○ **Transform Data**: Click this (recommended) to open the Power Query Editor. This allows you to fix any issues (like the missing values in the `Price` or `Area_sqft` columns observed in your data) before loading it.
7. **Finalize**:
   - ○ If you chose "Transform Data", click **Close & Apply** in the top-left corner of the Power Query Editor to finish loading the data into Power BI.

Question 6 :- Define Data View, Report View, and Model View. Explain the purpose of each view.

Answer :- **1. Report View**

- ● **Definition:** This is the default view when you open Power BI Desktop. It is the canvas where you design and build your reports.
- ● **Purpose:**
  - ○ **Visualization:** It is used to create visual representations of your data (charts, graphs, maps, cards).
  - ○ **Layout:** You arrange and format visuals on multiple pages to create interactive dashboards and stories.
  - ○ **Interaction:** This is where you configure slicers, filters, and bookmarks to control how end-users will interact with the data.

2. Data View (or Table View)

- ● **Definition:** This view displays the data in a row-and-column grid format, similar to an Excel spreadsheet or a database table.
- ● **Purpose:**
  - ○ **Inspection:** It allows you to inspect the actual data loaded into the model to verify values, data types, and formatting.
  - ○ **Calculation:** It is the primary place for creating and viewing **Calculated Columns** using DAX (Data Analysis Expressions).
  - ○ **Validation:** It helps in troubleshooting data issues by allowing you to see exactly what data is present in each table after the transformation steps.

3. Model View

- ● **Definition:** This view provides a diagrammatic representation of your data model, showing tables as blocks and the connections between them as lines.

- **Purpose:**
  - **Relationship Management:** It is used to define, edit, and manage relationships (one-to-one, one-to-many, etc.) between different tables.
  - **Schema Design:** It helps you visualize and organize complex schemas (like Star or Snowflake schemas).
  - **Properties Configuration:** You can bulk-edit column properties, create hierarchies, and set display folders to keep the data model organized.

Question 7 :- Discuss the different data sources that Power BI supports.

Answer :- **1. File**

This category includes data stored in flat files, which are often local or stored on a network drive.

- **Examples:** Excel Workbooks (`.xlsx`, `.xlsm`), Text/CSV, XML, JSON, PDF, and Parquet.
- **Folder:** You can also connect to a "Folder" source, which allows you to import and combine multiple files (e.g., daily sales logs) from a single directory automatically.

2. Database

Power BI supports direct connections to a wide variety of relational and non-relational database management systems. This is ideal for enterprise environments where data is centralized.

- **On-Premises Databases:** SQL Server, Oracle, IBM DB2, MySQL, PostgreSQL, Sybase, and Teradata.
- **Cloud Warehouses:** Amazon Redshift, Google BigQuery, and Snowflake.
- **Capabilities:** Many of these sources support **DirectQuery**, allowing Power BI to query the database in real-time without importing the data into the file.

3. Microsoft Fabric (formerly Power Platform components)

This category connects to data artifacts within the Microsoft ecosystem, allowing you to reuse existing logic and datasets.

- **Power BI Semantic Models:** Connect to datasets already published to the Power BI Service to avoid duplicating data models.
- **Dataflows:** Connect to ETL (Extract, Transform, Load) logic created in Power BI Service or Power Apps.
- **Datamarts, Lakehouses, and Warehouses:** Connect to modern data architecture components within Microsoft Fabric.

4. Azure

As a Microsoft product, Power BI has extensive native integration with Azure cloud services.

- **SQL & Storage:** Azure SQL Database, Azure Synapse Analytics (formerly SQL Data Warehouse), Azure Blob Storage, and Azure Table Storage.
- **Big Data:** Azure HDInsight (Spark, HDFS) and Azure Cosmos DB.

5. Online Services

This category includes connectors for popular Software-as-a-Service (SaaS) platforms and web-based tools.

- **Productivity & Sales:** SharePoint Online Lists, Microsoft Exchange Online, Dynamics 365, and Salesforce (Objects and Reports).
- **Analytics:** Google Analytics, Adobe Analytics, and LinkedIn Sales Navigator.

6. Other

This catch-all category contains flexible connectors for sources that don't fit into the standard buckets.

- **Web:** Allows scraping data from HTML tables on websites.
- **Scripts:** R Scripts and Python Scripts can be used to generate data frames or import data from sources not natively supported.
- **Generic Interfaces:** ODBC and OLE DB connectors allow you to connect to virtually any database that has a standard driver, even if a specific named connector doesn't exist in Power BI.

Question 8 :- Split Owner Name to create two new columns as First Name and Last Name.

Answer :- **Method 1: Using Power Query (Recommended)**

1. **Open Power Query Editor**: Click on **Transform Data** in the Home ribbon.
2. **Select the Column**: Click on the `Owner_Name` header to select it.
3. **Split Column**:
   - Go to the **Home** tab (or **Transform** tab).
   - Click on **Split Column** > **By Delimiter**.
4. **Configure Split**:
   - **Select or enter delimiter**: Choose **Space**.
   - **Split at**: Select **Left-most delimiter** (this ensures that if a name has multiple parts like "Van der Waal", "Van" becomes the First Name and "der Waal" becomes the Last Name).

- ○ Click **OK**.
5. **Rename Columns**: The columns will likely be named `Owner_Name.1` and `Owner_Name.2`. Double-click the headers and rename them to `First Name` and `Last Name`.
6. **Close & Apply**: Click **Close & Apply** to save changes.