Power Bi Assignment

1. What is Power BI and how does it differ from Excel?

Ans:

Power BI is a business intelligence (BI) tool by Microsoft that helps users visualize, analyze, and share data through interactive dashboards and reports.

- **Difference between Power BI and Excel**

Feature	Excel	Power BI
Purpose	Mainly used for data entry, calculation, and analysis	Used for advanced data visualization and business intelligence
Data Handling	Handles smaller datasets (limited by rows/columns)	Can handle very large datasets efficiently
Visualization	Charts, graphs, pivot tables	Interactive dashboards, advanced visuals, real-time updates
Ease of Use	Easier for beginners (spreadsheet-based)	Requires learning but more powerful for analysis
Collaboration	Sharing mostly via files (email, cloud)	Easy online sharing & collaboration through Power BI Service
Best For	Basic analysis & calculations	Business dashboards & decision-making insights

2. Explain the concept of data modeling in Power BI.

Ans:

Data modeling in Power BI is the process of organizing and connecting different data tables to create a structured model that supports accurate analysis and reporting.

Instead of keeping all data in one big table, Power BI allows you to create relationships between multiple tables (similar to a database) It involves:

• Defining fact and dimension tables.

• Creating relationships (one-to-many, many-to-many, bi-directional when needed).

• Setting filter directions and cardinality.

• Creating calculated columns, measures, hierarchies, and role-playing dimensions (e.g., multiple date relationships). Good modeling yields faster queries, simpler DAX, and more intuitive visuals.

3. What are the different types of connections available in Power BI?

Ans:

Connection Type	Description	Use Case
Import Mode	Data is copied and stored inside Power BI. Reports run on the imported data.	Best for small–medium datasets, faster performance.
direct Query Mode	Data is not imported. Power BI queries the source directly each time you view the report.	Useful for large datasets (e.g., SQL, Azure), real-time data.

Composite Mode	Mix of Import + Direct Query in the same model.	When some data should be real-time, and some imported for speed.
Live Connection	Connects directly to Analysis Services / Power BI datasets without storing data in PBIX.	For centralized data models and shared datasets.

4. How do you handle data transformation in Power BI?

Ans:

Data transformation in Power BI is the process of cleaning, shaping, and preparing raw data so it becomes suitable for analysis and reporting. This is mainly done using Power Query Editor.

Steps to Handle Data Transformation in Power BI:

1. **Load Data** → Import data from files, databases, or online sources.
2. **Use Power Query Editor** → Apply transformations without changing the original source data.
3. **Common Transformations:**

- Remove duplicates / blanks
- Rename columns & change data types
- Split or merge columns
- Filter rows (keep/remove specific records)
- Create calculated columns
- Unpivot/Transpose data
- Replace values (e.g., NULL → 0)

4. **Apply Steps** → Each transformation step is recorded and can be modified anytime.
5. **Load to Model** → Final cleaned data is loaded into Power BI for analysis.

5. What is DAX (Data Analysis Expressions) and why is it important in Power BI?

Ans:

DAX (Data Analysis Expressions) is a formula and expression language used in Power BI, Excel, and Analysis Services to perform calculations and data analysis.

It is similar to Excel formulas but more powerful, designed for working with relational data and large datasets.

Why is DAX Important in Power BI?

1. **Advanced Calculations** – Create custom metrics like Total Sales, Profit Margin, Year-to-Date Sales.
2. **Aggregations** – Summarize data (SUM, AVERAGE, COUNT, etc.) across different dimensions.
3. **Time Intelligence** – Perform date-based analysis (YTD, MTD, Growth vs. Previous Year).
4. **Dynamic Analysis** – Results change automatically with filters and slicers.
5. **Data Modeling** – Helps build calculated columns, measures, and tables for deeper insights.

6. Can you explain the difference between calculated columns and measures in Power BI?

Ans:

Feature	Calculated Column	Measure
Definition	A new column created in a table using DAX.	A calculation used in visuals, not stored as a column.
Storage	Physically stored in the data model (increases size).	Not stored; calculated on the fly when used in visuals.

Row Context	Works at row level (adds value for each row).	Works at aggregate level (summary across rows).
Use Case	When you need new fields for filtering, grouping, or relationships (e.g., Profit = Sales – Cost per row).	When you need calculations in visuals like totals, averages, percentages (e.g., Total Sales = SUM(Sales[Amount])).
Performance	Slower for large data (since it increases model size).	Faster and more efficient (calculated only when needed).

7. How do you handle relationships between tables in Power BI?

Ans:

- **Load Tables** → Import all related tables into Power BI.
- **Open Model View** → See tables as boxes and their fields.
- **Create Relationship** → Drag and drop a field (e.g., *CustomerID* in Sales → *CustomerID* in Customers).
- **Set Cardinality (type of relationship):**
    - One-to-Many (1:*) → Most common (e.g., One Customer → Many Sales).
    - One-to-One (1:1) → Rare, when both tables have unique values.
    - Many-to-Many (*:*) → When both tables have duplicates (requires bridge tables sometimes).
- **Cross Filter Direction** → Choose how filters flow (Single or Both directions).
- **Validate** → Ensure correct fields are linked and no duplicate/incorrect joins.

8. What is the purpose of a Power BI Gateway?

Ans:

A Power BI Gateway is a bridge that connects on-premises data sources (like SQL Server, Oracle, Excel files, local databases) with Power BI cloud services.

It allows organizations to keep their data securely stored on-premises while still being able to use it for dashboards and reports in Power BI Service.

- **Secure Connection** – Transfers data between on-premises sources and the Power BI cloud without moving the actual database.
- **Data Refresh** – Enables automatic/scheduled refresh of reports so users always see the latest data.
- **Hybrid Use** – Supports both cloud data (like Azure, SharePoint Online) and on-premises data together.
- **Centralized Management** – Admins can control access, security, and performance of data connections.
- **Other Uses** – Also works with Power Apps, Power Automate, and Azure Logic Apps.

9. How can you schedule data refresh in Power BI Service?

Ans:

- **Publish Report** → Upload your Power BI report (PBIX file) to Power BI Service.
- **Go to Dataset Settings** → In the workspace, find your dataset and click on Settings.
- **Configure Gateway (if needed)** → If using on-premises data, set up a Power BI Gateway for connection.
- **Set Refresh Schedule** →
    - Turn on Scheduled Refresh.
    - Choose Refresh Frequency → Daily, Weekly, or Multiple times per day.
    - Set Time Zone & Time Slots.

- **Apply Settings** → Save the schedule.
- **Monitor Refresh** → Check refresh history to confirm success or troubleshoot errors.

10. Explain the concept of row-level security in Power BI.

Ans:

Row-Level Security (RLS) in Power BI is a feature that restricts data access for specific users based on filters you define. It ensures that each user sees only the data they are allowed to view, even though everyone is using the same report.

How RLS Works:

1. **Define Roles** → Create roles (e.g., Sales Manager, Regional User).
2. **Apply Filters** → Assign DAX filters on tables (e.g., [Region] = "North").
3. **Assign Users** → Map Power BI Service users/groups to these roles.
4. **Result** → When a user views the report, Power BI applies the filter and shows only their permitted data.

11. What is the Power BI Desktop and how does it differ from Power BI Service?

Ans:

Feature	Power BI Desktop	Power BI Service
Type	Desktop application (installed on PC)	Cloud-based service (web)
Purpose	Used for creating, modeling, and designing reports	Used for sharing, viewing, and collaborating on reports
Data Preparation	Can clean, transform, and model data (Power Query, DAX)	Limited data modeling, mainly for viewing & sharing

Collaboration	Mostly individual work	Team collaboration, sharing, and access control
Data Refresh	Manual refresh	Automatic/scheduled refresh available
Usage	Developers/Analysts	End-users, managers, and teams

12. Explain the concept of Direct Query in Power BI.

Ans:

Direct Query is a data connection mode in Power BI where data is not imported into Power BI. Instead, Power BI queries the underlying data source in real-time whenever you interact with a report.

Key Features of direct Query:

1. **No Data Stored in PBIX** – Only the connection is stored, not the data itself.
2. **Real-Time Data** – Always shows the latest data from the source.
3. **Source Dependent** – Performance depends on the speed of the database.
4. **Limited Transformations** – Some modeling and DAX functions are restricted compared to Import mode

13. What are Power BI templates and how are they useful?

Ans:

A Power BI Template (.PBIT file) is a reusable file that contains the report structure (layouts, visuals, queries, data model, and measures) but does not include the actual data.

Users can open the template, connect to their own data sources, and reuse the same report design.

How Power BI Templates Are Useful:

1. **Reusability** – Save time by reusing the same report structure for different datasets.
2. **Consistency** – Ensures all users/reports follow the same design and KPIs.
3. **Sharing** – Easy to share report designs without sharing sensitive data.
4. **Customization** – Users can plug in their own data and still use the prepared dashboards.
5. **Standardization** – Useful in organizations where multiple teams need the same type of report.

14. How do you handle incremental data refresh in Power BI?

Ans:

How to Handle Incremental Refresh:

1. **Prepare Data with Date Column** → The table must have a date/time field (e.g., OrderDate).
2. **Define Parameters** → Create RangeStart and RangeEnd parameters in Power Query.
3. **Filter Data** → Apply filters using these parameters on the date column.
4. **Publish to Power BI Service** → Upload the dataset/report.
5. **Set Incremental Refresh Policy** →
    - Go to Dataset Settings → Incremental Refresh.
    - Define:
        - Historical Range (e.g., last 5 years kept in full).
        - Refresh Period (e.g., refresh only last 1 month).
6. **Automatic Refresh** → Power BI will refresh only the recent period while keeping old data unchanged.

15. What is the role of Power Query in Power BI?

Ans:

Data Import – Connects to various sources (Excel, SQL, Web, APIs, etc.).

Data Cleaning – Remove duplicates, blanks, and errors.

Data Transformation – Rename, split, merge, pivot/unpivot, and change data types.

Automation – Every transformation step is recorded and can be reused automatically.

Combine Data – Append or merge queries from multiple tables/sources.

Prepare Data Model – Ensures the dataset is clean and structured before loading into Power BI for analysis.

16. What is the role of Power Query in Power BI?

Ans:

Feature	Calculated Column	Calculated Table
Definition	A new column created inside an existing table using DAX.	A completely new table created using DAX.
Where It Exists	Added to an existing table.	Created as a separate table in the model.
Row Context	Works at row level (calculated for each row in the table).	Works at table level (can use data from one or more tables).
Storage	Physically stored in the table (increases model size).	Stored as a new table in the model.

Use Case	To add new fields for filtering, grouping, or simple row-based calculations (e.g., Profit = Sales – Cost).	To create reference/summary tables, aggregations, or role-playing dimensions (e.g., Calendar Table, Sales Summary Table).

17. How do you create custom visuals in Power BI?

Ans:

- **Import from AppSource**
  - Go to Visualizations Pane → More Options (…) → Get More Visuals.
  - Download visuals from Microsoft AppSource (e.g., Heatmaps, Gantt charts, KPI cards).
  - Imported visuals appear in the visualization pane.
- **Import from File (.pbiviz)**
  - Developers can create visuals using Power BI Custom Visual SDK (built with TypeScript, D3.js).
  - Exported as a .pbiviz file.
  - Import into Power BI using Import a Visual from File.
- **Develop Your Own Visual (Advanced)**
  - Install Power BI Visuals SDK (Node.js + PowerShell).
  - Use programming (TypeScript, JavaScript, D3.js) to design the visual.
  - Package and test the custom visual in Power BI Desktop.

18. What are the best practices for optimizing performance in Power BI?

Ans:

- Use Import mode whenever possible
- Follow star schema modeling
- Reduce columns & cardinality
- Use measures, not calculated columns
- Filter data in Power Query, not visuals
- Disable unnecessary auto-date/time
- Use summary tables and aggregations
- Optimize DAX using variables, avoid CALCULATE misuse
- Enable Performance Analyzer and DAX Studio to detect bottlenecks
- Apply incremental refresh for large datasets

19. How can you integrate Power BI with other Microsoft products like Azure and Office 365?

Ans:

Integration with Azure:

1. **Azure SQL Database / Data Lake / Synapse** → Connect Power BI directly to cloud databases for reporting.
2. **Azure Analysis Services** → Use as a semantic layer and connect Power BI via Live Connection.
3. **Azure Machine Learning** → Import ML models into Power BI for predictive analytics.
4. **Azure Active Directory (AAD)** → Manage authentication, user security, and Row-Level Security (RLS).

Integration with Office 365:

1. **Excel** → Import PivotTables, Power Query, and PowerPivot models into Power BI; export visuals back to Excel.
2. **SharePoint Online** → Publish and embed Power BI reports directly in SharePoint pages.
3. **Teams** → Embed Power BI dashboards inside Microsoft Teams for collaboration.
4. **Outlook** → Share links or snapshots of Power BI reports via email.
5. **OneDrive** → Keep Excel/CSV files in OneDrive; Power BI auto-refreshes when the file updates.

20. Explain the concept of aggregations in Power BI.

Ans:

Aggregations in Power BI are a performance optimization technique where summary tables (aggregated data) are created to speed up queries on large datasets. Instead of scanning millions of rows, Power BI uses pre-aggregated data for faster reporting.

How Aggregations Work:

1. Create an aggregation table (e.g., total sales by Region, Year, Product).
2. Define it in the Data Model and map it to the detailed fact table.
3. Power BI automatically decides whether to use the aggregation table or the detailed fact table when a query runs.
    - If the required data is in the aggregation table → It uses that (faster).
    - If not → It goes to the detailed fact table.

21. How do you handle error handling and data quality in Power BI?

Ans:

Power BI provides tools to ensure that the data used in reports is clean, accurate, and reliable. This is mainly done in Power Query Editor before loading the data into the model.

1. Error Handling in Power BI

- **Detect Errors** → Use *Power Query* to check for errors in columns (e.g., invalid data types, null values).
- **Replace Errors** → Replace error values with defaults (e.g., null → 0).
- **Remove Errors** → Delete rows with errors if they are not needed.
- **Conditional Handling** → Use *if…then…else* logic in Power Query or DAX to manage problematic data.

2. Ensuring Data Quality

- **Remove Duplicates** → Eliminate duplicate rows/records.
- **Data Type Correction** → Assign correct data types (Date, Number, Text).
- **Standardization** → Clean column names, unify formats (e.g., DD/MM/YYYY).
- **Fill Missing Values** → Replace blanks with default values, averages, or carry forward values.
- **Profiling Tools** → Use *Column Quality, Column Distribution, Column Profile* options in Power Query to check data validity.
- **Validation Rules** → Apply business logic checks (e.g., Sales Amount should not be negative).

22. What is the purpose of Power BI Embedded and when would you use it?

Ans:

Power BI Embedded is a PaaS (Platform-as-a-Service) offering for ISVs and developers to embed Power BI reports/dashboards into their own applications or websites.

Use cases:

- Customer-facing portals (e.g., SaaS apps).
- No need for users to have Power BI licenses.
- Full control over branding, UI, and navigation.
- Pay-per-use or capacity-based pricing.

Used when you want interactive analytics inside your own product, not the Power BI platform