“Power BI- Notes”

✅ Step 1: Power BI Basics

1. What is Power BI?

Power BI is a **Business Intelligence** tool developed by Microsoft that helps convert raw data into meaningful insights through interactive dashboards, visualizations, and reports.

- It's used for **data analysis**, **data visualization**, and **report sharing**.
- Combines **Excel-like familiarity** with powerful BI features.

2. Key Features of Power BI:

- Connects to various data sources (Excel, SQL, web APIs, cloud services)
- Data modelling and transformation (Power Query)
- Powerful calculation engine (DAX)
- Drag-and-drop visualizations
- Cloud-based sharing and collaboration (Power BI Service)

🔷 Power BI Ecosystem (Components)

Component	Description
Power BI Desktop	Main development tool for creating reports and data models. Installed locally.
Power BI Service	Cloud-based platform to publish, share, and schedule reports.
Power BI Mobile	Mobile app to view and interact with dashboards.
Power BI Report Server	On-premise report hosting (for organizations that don’t use the cloud).

🔷 Typical Power BI Workflow:

1. Connect – Import data from Excel, SQL Server, APIs, SharePoint, etc.
2. Transform – Clean and shape the data using Power Query Editor (ETL process).
3. Model – Define relationships, calculated columns, and measures using DAX.
4. Visualize – Create interactive charts, tables, and KPIs.
5. Publish & Share – Upload reports to Power BI Service and share via dashboards or apps.

🔷 **Power BI vs Excel (Key Differences)**

Feature	Power BI	Excel
Visualization	Interactive, modern, customizable	Static, limited
Data Volume	Handles millions of rows efficiently	Struggles with large data
Data Modelling	Star schema, relationships, DAX	Flat tables, basic formulas
Sharing	Web-based dashboards & apps	Files, email-based sharing

---

🔷 **Common Interview Questions on Step 1:**

1. **What is Power BI and what are its main components?**
2. **Explain the Power BI workflow.**
3. **How is Power BI different from Excel?**
4. **What is the use of Power BI Service?**

✅ **Step 2: Power BI Desktop**

Power BI Desktop is a free application that runs on Windows and allows you to connect to data, transform it, model it, and create visual reports.

🔷 **1. Installing Power BI Desktop**

- Download from Microsoft's official website or install via the Microsoft Store.
- Requires Windows OS (not available for macOS).

🔷 **2. Power BI Desktop Interface Overview**

After launching Power BI Desktop, you'll see several key areas:

Area	Purpose
Ribbon (Top)	Access to common features like data import, visuals, and modeling tools.
Report View	Drag-and-drop area for creating visuals.
Data View	View raw data in table format.
Model View	Create and manage relationships between tables.
Fields Pane	Lists your tables, columns, and calculated fields.
Visualizations Pane	Choose visual types and configure their fields.

🔷 3. Connecting to Data

Power BI supports a wide variety of data sources:

✅ **Common Data Sources:**

- Excel
- CSV/Text
- SQL Server
- Web APIs
- SharePoint
- Azure
- Oracle, MySQL, PostgreSQL
- Online Services (Salesforce, Google Analytics, etc.)

💡 **Interview Tip:** Be ready to explain how you connect to SQL Server or Excel and what credentials or steps are involved.

🔷 4. Data Connectivity Modes

Mode	Description
Import	Data is imported and stored in Power BI file (fastest, best for modelling).
DirectQuery	Data stays in source; queries run live every time you view a report.
Live Connection	Used for SSAS (SQL Server Analysis Services) only.

✅ **When to use Import vs DirectQuery?**

- Use **Import** for speed and complex modelling.

- Use **DirectQuery** when working with real-time or large datasets.

🔷 5. Save and Publish

- Save files with .pbix extension.
- Use the **Publish** button to upload your report to Power BI Service for sharing and dashboard creation.

🔷 Common Interview Questions from Step 2:

1. What are the different views in Power BI Desktop?
2. What data sources have you worked with?
3. Difference between Import and DirectQuery mode?
4. Can we change from DirectQuery to Import? (Yes, but only in some cases)

✅ Step 3: Power Query (Data Transformation)

Power Query Editor in Power BI is used for **cleaning**, **transforming**, and **shaping data** before it's loaded into the model.

🔷 1. What is Power Query?

Power Query is a data connection and transformation tool used in Power BI Desktop. It follows **ETL** (Extract, Transform, Load) principles.

- **Extract** data from multiple sources
- **Transform** it using a graphical interface or M code
- **Load** it into Power BI for analysis

You can open Power Query by clicking **"Transform Data"** in the Home tab.

🔷 2. Power Query Interface Overview

Section	Purpose
Queries Pane	Shows all tables (queries) being transformed.
Data Preview Pane	Shows current view of the data table.
Applied Steps Pane	Keeps track of every transformation step (in order).

Section	Purpose
Formula Bar	Shows M code for the selected step (optional, can be enabled).

🔷 3. Common Data Cleaning Tasks

These are the most commonly used transformations in interviews and real projects:

Transformation	Description
Remove Rows	Remove blank rows, duplicates, or rows based on filters.
Split Column	Split text by delimiter (e.g., comma, space).
Change Data Type	Convert text, whole number, decimal, date, etc.
Rename Columns	Make column names clear and consistent.
Trim and Clean	Remove extra spaces or non-printable characters.
Fill Down/Up	Fill nulls with values above or below.
Group By	Aggregate values based on one or more columns.
Pivot/Unpivot	Reshape data from rows to columns or vice versa.

🔷 4. Merge and Append Queries

These are crucial for combining data from multiple sources or tables.

- **Merge Queries** (similar to SQL JOIN): Combine rows from two tables using a key column.
- **Append Queries** (like SQL UNION): Stack data vertically from two tables with similar columns.

💡 **Interview Tip:** Be ready to explain a use case for Merge (e.g., combining Customer data with Sales data).

🔷 5. Custom Columns and Conditional Logic

Use the **Add Column** tab to:

- Create calculated columns using simple logic or M formulas.
- Write conditional logic (like Excel IF statements):

  arduino

CopyEdit

= if [Revenue] > 10000 then "High" else "Low"

🔷 6. M Language (Advanced but useful)

- Power Query uses the **M Language** behind the scenes.
- Each transformation step writes an M function (visible in formula bar).
- You're not expected to write M from scratch in most interviews, but understanding the syntax helps.

🔷 7. Applied Steps & Query Folding

- Each action in Power Query adds an **"Applied Step"**, which is executed in order.
- **Query Folding**: When transformations are pushed back to the source (important for performance in DirectQuery).

🔷 8. Close & Apply

After all transformations:

- Click **Close & Apply** to load cleaned data into Power BI Desktop.

🔶 Common Interview Questions from Step 3:

1. How do you clean messy data in Power BI?
2. What is the difference between Merge and Append?
3. What is Query Folding?
4. Can you explain how you used Group By or Unpivot in your projects?

✅ Step 4: Data Modeling

Data modeling is the process of **structuring your data tables**, **defining relationships**, and **preparing calculated fields** so your visualizations and measures work correctly.

🔷 1. Why is Data Modeling Important?

- Ensures data integrity
- Enables advanced DAX calculations

- Boosts report performance
- Allows scalable and flexible reports

🔷 2. Data View vs Model View

- **Data View**: See individual tables and preview their rows.
- **Model View**: Define and manage relationships between tables (drag-and-drop interface).

🔷 3. Relationships

Power BI automatically detects relationships, but you should understand how to manage them.

Term	Meaning
Primary Table	Table with unique values (usually a dimension).
Foreign Table	Table with repeated values (usually a fact table).
Cardinality	One-to-many (1:*) or many-to-one (*:1) is most common.
Cross Filter Direction	Single or both (usually keep Single for performance and clarity).

💡 **Interview Tip:** Explain that you often work with **one-to-many** relationships (e.g., one customer can have many orders).

🔷 4. Star vs Snowflake Schema

Schema	Description
Star Schema	Central fact table linked directly to dimension tables. Preferred in Power BI.
Snowflake Schema	Dimension tables are normalized (split into sub-dimensions). More complex.

✅ **Best Practice:** Use **Star Schema** for better performance and simpler DAX.

🔷 5. Key Modeling Concepts

✅ **Fact Tables vs Dimension Tables:**

- **Fact Table**: Transactional data (e.g., sales, revenue, orders)
- **Dimension Table**: Descriptive attributes (e.g., customer, product, region)

✅ **Hide Unused Columns**

Hide unnecessary columns (like IDs or unused columns) to keep the model clean.

🔷 6. Calculated Columns vs Measures

Feature	Calculated Column	Measure
Calculated For	Each row in a table	Aggregated over multiple rows
Storage	Stored in the model (uses memory)	Calculated on the fly (efficient)
Use Case	Create new data fields (e.g., Profit = Sales - Cost)	KPIs like Total Sales, Avg Discount

⚠️ **Best Practice:** Prefer **Measures** over Calculated Columns for performance.

🔷 7. Role-playing Dimensions

Sometimes the same dimension table is related to a fact table multiple times (e.g., Date as Order Date and Ship Date). Use **inactive relationships** and USERELATIONSHIP() in DAX.

🔷 8. Formatting & Sorting

- Rename fields for clarity (e.g., "cust_id" → "Customer ID").
- Use “Sort by Column” (e.g., sort month names by month number).

🔷 Common Interview Questions from Step 4:

1. What is the difference between a calculated column and a measure?
2. What is the difference between star and snowflake schema?
3. How do you create and manage relationships in Power BI?
4. Explain a scenario where you had to use USERELATIONSHIP().

✅ Step 5: DAX (Data Analysis Expressions)

DAX is the formula language used in Power BI for creating **measures**, **calculated columns**, and **calculated tables**. It allows you to perform custom aggregations, calculations, filters, and time intelligence analysis.

🔷 1. Types of DAX Calculations

Type	Used For	Stored In
Measure	Aggregated results (e.g., Total Sales)	Not stored (calculated on-the-fly)
Calculated Column	Row-level calculations (e.g., Profit per order)	Stored in model table
Calculated Table	Create new tables using logic (e.g., filter data)	Model level

---

🔷 2. Basic DAX Functions

✅ **Aggregation:**

DAX

CopyEdit

SUM(Sales[Amount])

AVERAGE(Orders[Quantity])

MAX(Products[Price])

✅ **Logical:**

DAX

CopyEdit

IF(Sales[Amount] > 1000, "High", "Low")

SWITCH(TRUE(), ...)

✅ **Text:**

DAX

CopyEdit

CONCATENATE(Customer[FirstName], " ", Customer[LastName])

LEFT(Product[Code], 3)

✅ **Date:**

DAX

CopyEdit

YEAR(Orders[OrderDate])

MONTH(Orders[OrderDate])

🔷 3. Filter Functions

These are used to change the context of calculations:

DAX

CopyEdit

CALCULATE(SUM(Sales[Amount]), Region[RegionName] = "West")

FILTER(Sales, Sales[Amount] > 1000)

ALL(Products) -- removes filters

REMOVEFILTERS(Date)

✅ CALCULATE() is **the most powerful function in DAX** — it modifies the filter context.

🔷 4. Time Intelligence Functions

Power BI has built-in support for time-based calculations (if you have a proper date table).

DAX

CopyEdit

TOTALYTD(SUM(Sales[Amount]), 'Date'[Date])

SAMEPERIODLASTYEAR('Date'[Date])

DATESINPERIOD('Date'[Date], TODAY(), -1, MONTH)

⚠️ These functions only work correctly if:

- You have a **Date Table** marked as a **Date Table**
- Your model has a **relationship** between the Date Table and your fact table

---

🔷 5. Row Context vs Filter Context (Advanced Concept)

Context	Description
Row Context	Evaluates values **row-by-row** (used in calculated columns).
Filter Context	Applies filters to a **set of rows** (used in measures).

💡 **CALCULATE** transitions from **row context to filter context** — important for advanced DAX questions.

🔷 6. Common DAX Patterns

✅ **Total Sales**

DAX

CopyEdit

Total Sales = SUM(Sales[Amount])

✅ **Sales for Selected Region**

DAX

CopyEdit

West Sales = CALCULATE([Total Sales], Region[RegionName] = "West")

✅ **YoY Growth**

DAX

CopyEdit

YoY Sales = [Total Sales] - CALCULATE([Total Sales], SAMEPERIODLASTYEAR('Date'[Date]))

🔶 **Common Interview Questions from Step 5:**

1. What is the difference between a measure and a calculated column?
2. What is CALCULATE and why is it important?
3. Explain row context and filter context with examples.
4. How do you calculate Year-over-Year or Month-to-Date in DAX?
5. What is the use of ALL and REMOVEFILTERS in DAX?

✅ **Step 6: Data Visualization in Power BI**

Data visualization in Power BI is about transforming raw data into visuals that help users **understand, explore, and make decisions** effectively.

🔷 **1. Report View: The Design Canvas**

This is where you build your **pages**, add **charts**, and organize your visuals using:

- **Fields Pane** – Drag fields here to build visuals
- **Visualizations Pane** – Select, format, and customize visual types
- **Filters Pane** – Apply page-level, visual-level, or report-level filters

---

🔷 **2. Most Common Visual Types**

Visual Type	Use Case
Bar/Column Charts	Compare values (e.g., Sales by Region)
Line Charts	Show trends over time
Pie/Donut Charts	Show proportions (avoid overuse)
Tables & Matrix	Display detailed or pivoted data
Cards	Display KPIs like total sales or profit
Slicers	Filters with dropdowns, sliders, checkboxes
Maps	Show geographic data (e.g., Sales by State)
Gauge/KPI Visuals	Show progress towards targets

💡 **Interview Tip:** Be prepared to explain **why** you chose a specific visual for a metric.

🔷 3. Best Practices for Report Design

✅ Keep it Clean & Simple

- Use 4–6 visuals per page (avoid clutter)
- Use consistent font styles and colors

✅ Use Titles and Labels Clearly

- Make sure every chart has a clear title and axis labels

✅ Use Tooltips Wisely

- Add extra info without cluttering the chart

✅ Use Slicers/Filters for Interactivity

- Use slicers for Date, Region, Product, etc.
- Sync slicers across pages if needed

✅ Use Drill Down & Drill Through

- Allow users to click and explore more detailed levels
- Example: Click on a Region to drill into State, then City

✅ Use Bookmarks and Buttons (Advanced)

- For navigation, hiding/showing visuals, pop-ups, etc.

🔷 4. Filters & Interactions

Filter Type	Scope
Visual-level	Affects only one visual
Page-level	Affects all visuals on that page
Report-level	Affects the entire report

You can also control **visual interactions** using **"Edit Interactions"** under the Format tab.

🔷 5. Themes & Formatting

- Apply consistent **themes** for colors and fonts
- Use conditional formatting for KPIs (e.g., red/green status)
- Use **custom tooltips** and **background images** (optional)

🔷 6. Custom Visuals

Power BI allows you to import visuals from the marketplace:

- Examples: Word Cloud, Sankey Chart, Chiclet Slicer
- Use custom visuals **only if built-in ones are insufficient**

🔶 Common Interview Questions from Step 6:

1. How do you choose which visual to use for a given metric?
2. What are drill-down and drill-through? How have you used them?
3. What are slicers and how do you use them for interactivity?
4. How do you make your reports user-friendly and readable?

✅ Step 7: Dashboard Publishing & Sharing (Power BI Service)

Once your report is built in Power BI Desktop, you need to **publish it**, **share it**, and **manage its access, refresh, and usage** using the Power BI Service (cloud platform).

🔷 1. What is Power BI Service?

Power BI Service (app.powerbi.com) is the **online platform** for:

- Hosting and sharing reports

- Setting up data refresh schedules
- Creating dashboards
- Managing access and permissions
- Collaborating across teams

🔷 2. Publishing Reports to Power BI Service

From Power BI Desktop:

- Click **"Publish"** > Select Workspace
- Report is uploaded to **Power BI Service** in that workspace

📌 A **workspace** is a container for datasets, reports, dashboards, and dataflows.

🔷 3. Dashboards vs Reports

Feature	Report	Dashboard
Pages	Can have multiple pages	Single-page summary
Data Source	Single dataset	Can pin tiles from multiple reports
Interactivity	Full interactivity	Limited (tile-based navigation)

✅ Dashboards are used to **monitor KPIs** at a glance; reports are for deeper analysis.

🔷 4. Sharing Options

You can share reports in multiple ways:

Method	Notes
Share report	Directly with users via email (Pro license required)
Publish to Web	Generates public link (⚠️ Not secure for confidential data)
Workspace Access	Add users to a workspace with Viewer/Contributor/Admin roles
Power BI Apps	Package multiple reports into an app for easy distribution

🔐 Make sure you manage **Row-Level Security (RLS)** if users should only see their own data.

🔷 5. Scheduled Data Refresh

- Setup via **Power BI Service > Dataset > Schedule Refresh**

- Configure **refresh frequency** (daily, hourly, etc.)
- Set up **data gateway** if data is on-premises

✅ Always monitor **refresh failures** — they're common interview questions.

🔷 6. Row-Level Security (RLS)

Control **which data a user can see** using filters:

1. Define roles in Power BI Desktop (Manage Roles)
2. Apply filters (e.g., [Region] = "West")
3. Test with "View as Role"
4. Publish and assign users to roles in Power BI Service

💡 **Interview Tip:** RLS is often asked in scenarios — like restricting sales reps to see only their region's data.

🔷 7. Power BI License Types (Basic Overview)

License Type	Key Features
Free	Build reports locally (can't share or collaborate)
Pro	Share reports, dashboards, use workspaces
Premium	Higher capacity, paginated reports, advanced AI, and RLS at scale

---

🔶 Common Interview Questions from Step 7:

1. How do you share your Power BI reports with management?
2. What is the difference between a report and a dashboard?
3. What is Row-Level Security and how do you implement it?
4. How do you schedule data refresh in Power BI Service?
5. What's the difference between Power BI Pro and Premium?
