# POWER BI ASSIGNMENT

1. **What is power bi and how does it differ from excel?**

**ANS:** Power bi is a business intelligence (BI) and data visualization tool developed by Microsoft. It helps users connect to data source, clean and transform the data, create interactive dashboards and reports, and share insights across organizations.

**How power bi differ from excle**

| **Key point** | **Excel** | **Power Bi** |
| --- | --- | --- |
| **Primary purpose** | Spredsheet tool for calculations, analysis and reporting. | Business Intelligence tool for interactive reporting, dashboards and sharing. |
| **Data Handling** | Best for smaller datasets | Handle very large datasets |
| **Data sources** | Mostly local files, databases, and manual entry. | Connects to a wide range of sources(databases, cloud services, APIs, live data streams) |
| **Visualization** | Charts, pivot tables, conditional formating | Rich, interactive visuals with drill-downs, slicers, filters, maps, etc. |
| **Collaboration** | Shared as files (email,onedrive,share point) | Reports and dashboards can be published and shared online via power bi service. |
| **Analysis** | Strong for ad-hoc analysis, calculations, financial modeling | Strong for dashboards, KPIs, trend analysis and big-picture insights |
| **Automation** | Manual refresh (unless connected with macros or power query) | Automatic refresh (scheduled updates, real time dashboards) |

**In short:**

* Use Excle when we want to detailed calculations, modeling, or working with small datasets.
* Use power bi when you we want to create interactive dashboards, to analyze large datasets, and to share insights at scale across an organization.

|  |  |  |
| --- | --- | --- |
| 1. **Explain the concept of data modeling in power bi.**   **ANS:** Data modeling in power bi is the process of structuring, connecting, and organizing data tables so that they can work together to provide meaningful insights.  When you bring data from multiple sources ( Excel sheets, SQL, etc..),those datasets are usually raw and unrelated.  A data model tells power bi how these tables are related and how calculations should be performed.  **Key Element of data modeling:**   * **Tables:**   You load data into power bi in the form of tables (like a excel) and each table represents a dataset (e.g., sales, customer, products, profit, and etc).   * **Relationship:**   Just like in databases, tables are connected through relationships  Ex.,   * Sales table has customer id * Customer table has customer id   Linking them lets you analyze sales by customer details.   * **Calculated columns & Measures:** * Calculated columns: new fields created row by row (e.g., Profit= sales-cost). * Measures: Calculations made on aggregated data using Dax (e.g., total sales = sum(sales[amount]). * **Data types and & Hierarchies:** * Each field must have the correct data type(text, number, date). * You can create hierarchies (e.g.,Date, year, quarter, month, day)for drill-down analysis.  1. **What are the different types of connections available in power bi.**   **ANS:** Types of connection in power bi   1. Import mode:  * Data is copied into Power bi’s in memory storage. * Fast performance because all data is loaded in RAM. * Best for small to medium datasets that don’t change frequently. * Supports full power BI features (modeling, Dax, transformations)   E.x., Importing an EXCEL or SQL table  2) Direct query mode:   * Data is not imported. Instead, queries are sent live to the source whenever visuals are refreshed. * Good for very large Datasets (that can’t fit in memory). * Ensures up to date data (real time or near time). * Limited transformation and some dax functions may not work. * Performance depends on the speed of the source system.  1. **How do you handle data transformation in power bi?**   **ANS:** In power bi, data transformation means cleaning, reshaping and preparing raw data so it becomes usable for reporting and analysis. This is mainly done in power query editor.   1. Power query editor:  * Connect to data sources * Clean the data (remove duplicates, handle nulls, filter rows) * Reshape the data (pivot/unpivot, merge) * Apply transformation without changing the original source  1. Common data transformation tasks: 2. Data cleaning:  * Remove nulls or blank value. * Replace values(e.g., N/A to blank)  1. Column operations :  * Split columns * Merge columns * Change data types  1. Row operations:  * Filter rows * Remove duplicates * Sort data  1. Combine queries:  * Merge : join two table  1. Adding calculations:  * Add custom column * Create conditional logic  1. Advanced Transformation:  * Group by: aggregate data(e.g., total sales by region)   **In short:** we handle data transformation in power bi using power query editor, where we clean, shape, and combine raw data into a structured formate ready for modeling and visualization.   1. **What is dax ( data analysis expressions) and why is it important in power bi?**   **ANS:** Dax is a formula language used in power bi to perform calculation and data analysis on our data model. It looks similar to excel formulas but is more powerful because it can work with relationship between tables and large datasets.  Dax is mainly used to create:   * Calculated columns Add new data at the row level. * Measures aggregate data dynamically (e.g., total sales, average profit and etc..) * Tables Create calculated tables for modeling. * **Dax is important because it allows:** * Perform advanced calculations:   Beyond simple sums or averages, DAX helps we to calculate running totals, year to date, month over month growth, percentage changes, etc.   * Enable Dynamic Reporting:   Measures written in dax automatically adjust based on filters, slicers, or visuals in power bi. For e.g., A “Total sales” measures will change when we filter by region or month and year.   * Enhance data modeling :   You can create new business logic (e.g., profit margine, customer lifetime value)directly in the data model, which is not available in raw data.   * Work with time intelligence:   Dax has built in functions for date and time analysis, which are essential for trend analysis.   * Dax is the backbone of calculations in power bi. Without it, we can only create basic visuals.  1. **Can you explain difference between calculated columns and measures in power bi?**   **ANS:**   1. **Calculated columns:**  * Definition: A calculated column is a new column we create in our data table using dax formulas. * Row context: it is calculated row by row when the data is loded or refreshed. * Storage: values are physically strored in the data model, which increases file size. * Use: Useful when we need a value for each row (e.g., Profit= sales – cost).   For e.g., Total sales = Quantity\*price   1. **Measures:**  * Definition: A measures is a dax calculation that is evaluated only when used in a report. * Filter context: it is calculated on the fly based on filters/slicers applied. * Storage: Results are not stored in the model, making it more efficient. * Use : Useful for aggregations like sum, average, count, percentages, ratios.   For e.g., Total sales = sum(Table[sales])   1. **How do you handle relationship between tables in power bi?**   **ANS:**   * Identifying Keys : unique key in one table, repeated key in another table. * Creating relationships : In model view or manage relationships. * Choosing cardinality: One-to-many, one-to-one, many-to-many. * Managing active/inactive relationships : power bi allows maltipal relationship between two tables, but only one can be active. * Use a date/lookup table: always create a proper date table and link alol date fields to it. Use lookup/dimension tables (customer, product, regions) linked to fact tables (sales , orders). * Following best practice: use star schema (fact + dimension tables) for clarity and performance.  1. **What is the purpose of power bi gateway?**   **ANS:**  A power bi gateway acts as a bridge between on premises data ( databases, files, servers) and the power bi service(cloud).  It allows us to:   * Securely connect on-premises data to the cloud without moving the data permanently. * Refresh datasets automatically so reports always show the latest data. * Enable live queries from cloud reports to your on premises sources. * Maintain security by keeping data within our network until a query is made   **In short:** a power bi gateway lets power bi services access and refresh our on premises data sources securely and in real time.   1. **How can you schedule data refresh in power bi service?**   **ANS:**   * Publish our report: upload the power bi report (.pbix) to the power bi service. * Go to the dataset: in the workspace, find our datasets. * Open settings : click on settings of the dataset. * Gateway connection: configure a gateway if the data source in on premises. * Schedule refresh: Under scheduled refresh:   Turn it on  Set frequency( daily , weekly)  Choose time for refresh   * Save settings   In short: upload report> dataset setting> configure gateway ( if required)> enable & set scheduled refresh.   1. **Explain the concept of row level security in power bi.**   **ANS:** Row-level security is a feature in power bi that restricts data access at the row level for different users, based on filters you define.  It ensures that users only see data they are authorized to see in reports and dashboards.   * **How it works** * Define roles & rules : in power bi desktop, we create roles with dax filters.   **E.g.,** sales[region] = “east” that only shows data for east region.   * Assign uses : in power bi services, assign users/groups to those roles. * Filter context applied: when a user o-pens the report, the filter is a automatically applied, hiding other data. * **Typs of RLS:** * Static RLS : filters are fixed (e.g., east region always sees east data). * Dynamic RLS : filters depenfd on logged in user( e.g., sales[region] = USERPRINCIPALNAME() to show data only for the user’s region). * **In short:**   Row level security ensures different users see only the data they’re allowed to, by applying filters at the row level.   1. **What is power bi desktop and how it differ from power bi service?**   **ANS:**  **Power BI desktop:**   * A free windows application used to connect, transform, and model data, and build reports. * works offline on our computer. * Best for data preparation, modeling and report creation.   **Power Bi services**   * A cloud based platform to publish, share, and collaborate on reports. * Work online in the browser. * Best for sharing ,dashboards, scheduled refresh and collaboration.  1. **Explain the concept of direct query in power bi.**   **ANS:**   * Data is not imported; queries run directly on the source. * Ensure real time/ up to date data. * Pros: no data storage in PBIX, handles large datasets. * Slower performance, limited dax/modeling features, depends on sources availability. * **In short:** direct query keeps reports live and real time, but with some performance and features trade-offs.  1. **What are power bi templates and how are they useful?**   **ANS:**  A .PBIT file that saves the report structure, visuals, queries, and model, but not the data. When opened, users can connect to their own data sources.   * **Usefulness:** * Reusable for standardized report formats. * Saves time and effort by avoiding redesign. * Ensures consistency across reports for different teams or projects. * **In short:** power bi templates let we reuse report designs with new data, ensuring speed and consistency.  1. **How do you handle incremental data refresh in power bi?**   **ANS:** In power bi, incremental data refresh is handled by refreshing only the new or updated data instead of the full dataset.  **Steps:**   * Create two parameters: rangestart and rangeEnd * Apply a filter on a date column using these parameters. * In power bi desktop, set up the incremental refresh policy(e.g., store last 5 year, refresh last 1 month). * Publish the dataset to power bi services. * Power bi then refreshes only the recent partition while keeping historical data unchanged. * **Advantage:** saves time, reduces resource usage and improves performance.  1. **What is the role of power query in power bi ?**   **ANS:** The role of power query in power bi is to connect, clean, transform and shape raw data before loading it into the data model.   * Data connection: connects to multiple sources ( excel, SQL,etc..) * Data transformation: cleans, filters, merges, splits, pivots/unpivots data. * Automation: steps are recorded as queries, so refresh automatically applies the same transformations. * Preparation for modeling: ensures data is structured and ready for relationships, DAX , and visuals. * **In short:** power query = ETL tool (Extract, Transform, load) inside power bi.  1. **Explain the difference between calculated columns and calculated tables in power bi.**   **ANS:**   * **Calculated column** * Adds a new column to an existing table. * Formula is evaluated row by row. * Stored in the data model and increase model size. * **E.g.,** Profit = sales[revenue] – sales[cost] * **Calculated Table:** * Creates an entire new table using DAX. * Based on existing tables/columns (filtered, aggregated and joined) * Helpful for creating summary tables, disconnected table. * **E.g.,** Topcustomers = TOPN(10,Sales, Sales[Revenue])   **In short:**   * Column = adds field in a table * Table = creates a new table from DAX.  1. **How do you create custom visuals in power bi?**   **ANS:**  In power BI, we can create custom visuals in two main ways:   * **Import from Appsource**: download ready-made custom visuals from the marketplace and add them to our report. * **Develop our own:**  use power bi visual tools with typescript, react to code and package a visual, then import it into power bi.  1. **What are the best practices for optimizing performance in power bi?**   **ANS :**   * Use star schema for data modeling. * Import only needed columns/rows. * Prefer measures over calculated columns. * Apply filters/transformation in power query. * Use incremental data refresh. * Limit visuals and slicers on report pages.   **In short:** keep data light, model simple, dax efficient, and visuals minimal.   1. **How can you integrate Power bi with others Microsoft products like a azure and office 365?**   **ANS:**   * **With azure:** connect to azure SQL, synapse, data lake for data, use azure ML for predictive analytics, and manage security with azure AD. * **With office 365:**  import from excel, embed reports in teams and sharepoint and share insights via outlook/powerpoint. * Power bi + Azure = advanced data + AI * Power Bi + office 365 = easy sharing + collaboration.  1. **Explain the concept of aggregations in power bi.**   **ANS:**  In power bi, aggregations are a way to improve performance by summarizing large datasets into smaller,pre-calculated tables.   1. **Key points:**  * Instead of querying billions of rows, power bi queries an aggregated table. * The detailed table stays in storage, but most reports can run off the lightweight aggregated table. * If a user needs detail not in the aggregation, power bi automatically switches to the detailed data.  1. **Example:**  * We have sales data with 1 billion rows. * Create an aggregation table with sales by year, region and product. * Reports run faster because they query the small aggregated table.   **In short:** aggregation = summary tables that reduce data volume and improve query performance in power bi.   1. **How do you handle error handling and data quality in power bi?**   **ANS:**   * **How to handle errors & ensure quality:**  1. In power query (ETL stage):  * Use remove errors or keep errors options to handle problematic rows. * Apply replace errors to substitute missing/invalid values with defaults. * Perform cleaning (trim, split, merge, deduplicate, handle nulls).  1. Data validation & Quality:  * Create conditional columns to flag invalid or unexpected values. * Apply constraints through relationships and dax (e.g., IFERROR, ISBLANK).  1. Monitoring in power BI services:  * Set up data refresh error alerts. * Track Data quality issues with dashboard KPIs or audit reports.   **In short:** Power bi ensures data quality by cleaning & validating in power query, using Dax for error handling, and monitoring refresh/report errors in the service.   1. **What is the purpose of power bi embedded and when would you use it?**   **ANS:**   * **Purpose:** * Provide interactive analytics inside custom apps without requiring users to open power bi separately. * Allow non power bi users to view reports securely. * When to use: * If we organization wants to share reports with external users (customers, partners) without giving them power bi licenses. * When we need scalable, embedded reporting inweb or mobile apps.   **In short:** Use power bi embedded when we need to integrate interactive power bi reports into external apps or portals for customers or partners.                    **** |  |  |