**Week 3 – Data Visualization, PowerBI**

**1.How would you define Power BI as an effective solution?**

Power BI is a strong business analytical tool that creates useful insights and reports by collating data from unrelated sources. This data can be extracted from any source like Microsoft Excel or hybrid data warehouses. Power BI drives an extreme level of utility and purpose using interactive graphical interface and visualizations. we can create reports using the Excel BI toolkit and share them on-cloud with the co-workers.

**2. What are the major components of Power BI?**

Power BI has these major components:

Power Query : It is used for data mash-up and transformation and we can use this to extract data from various databases like SQL Server, MySql, and many others  and to delete a chunk of data from various sources.

Power Pivot : It is a tabular data modeling engine that uses a functional language called Data Analysis Expression (DAX) to perform the calculations. Also, creates a relationship between various tables to be viewed as pivot tables.

Power View : It is used for viewing data visualizations that provides an interactive display of various data sources to extract metadata for proper data analysis.

Power BI Desktop : Power Desktop is an aggregated companion development tool of Power Query, Power View, and Power Pivot. Create advanced queries, models, and reports using the desktop tool.

**3. What are the various refresh options available?**

Four main refresh options are available in Power BI:

Package/OneDrive refresh: This synchronizes Power BI desktop or Excel file between the Power BI service and OneDrive

Data/Model refresh: This means scheduling the data import from all the sources based on either refresh schedule or on-demand.

Tile refresh: Refresh the tiles cache on the dashboard every time the data changes.

Visual container refresh: Update the reports visuals and visual container once the data changes.

**4. What are the different connectivity modes in Power BI?**

The three major connectivity modes in Power BI are:

Direct Query: The method allows direct connection to the Power BI model. The data doesn’t get stored in Power BI. Interestingly, Power BI will only store the metadata of the data tables involved and not the actual data. The supported sources of data query are:

Amazon Redshift

Azure HDInsight Spark (Beta)

Azure SQL Database

Azure SQL Data Warehouse

IBM Netezza (Beta)

Impala (version 2.x)

Oracle Database (version 12 and above)

SAP Business Warehouse (Beta)

SAP HANA

Snowflake

Spark (Beta) (version 0.9 and above)

SQL Server

Teradata Database

Live Connection: Live connection is analogous to the direct query method as it doesn’t store any data in Power BI either. But opposed to the direct query method, it is a direct connection to the analysis services model. Also, the supported data sources with live connection method are limited:

SQL Server Analysis Services (SSAS) Tabular

SQL Server Analysis Services (SSAS) Multi-Dimensional

Power BI Service

Import Data (Scheduled Refresh): we upload the data into Power BI. Uploading data on Power BI means consuming the memory space of your Power BI desktop. If it is on the website, it consumes the space of the Power BI cloud machine. Even though it is the fastest method, the maximum size of the file to be uploaded cannot exceed 1 GB until and unless you have Power BI premium.

**5.What are the different kinds of views?**

The different  kinds of views are:

Data View: Curating, exploring, and viewing data tables in the data set. Unlike, Power Query editor, with data view, you are looking at the data after it has been fed to the model.

Model View: This view shows you all the tables along with their complex relationships. With this, you can break these complex models into simplified diagrams or set properties for them at once.

Report View: The report view displays the tables in an interactive format to simplify data analysis. You can create n number of reports, provide visualizations, merge them, or apply any such functionality.

**6.What are the data sources Power BI can connect?**

The data source is the point from which the data has been retrieved. It can be files in various formats like .xlsx, .csv, .pbix, .xml, .txt etc, and databases like SQL database, SQL Data Warehouse, Spark on Azure HDInsight, or form content packets like Google Analytics etc.,

**7.What are the building blocks of Power BI?**

The major building blocks of Power BI are:

Datasets: Dataset is a collection of data gathered from various sources like SQL Server, Azure, Text, Oracle, XML, JSON, and many more. With the GetData feature in Power BI, we can easily fetch data from any data source.

Visualizations: Visualization is the visual aesthetic representation of data in the form of maps, charts, or tables.

Reports: Reports are a structured representation of datasets that consists of multiple pages. Reports help to extract important information and insights from datasets to take major business decisions.

Dashboards: A dashboard is a single-page representation of reports made of various datasets. Each element is termed a tile.

Tiles: Tiles are single-block containing visualizations of a report. Tiles help to differentiate each report.

**8. What is DAX?**

Data Analysis Expression (DAX) is a library of formulas used for calculations and data analysis. This library comprises functions, constants, and operators to perform calculations and give results. DAX lets you use the data sets to their full potential and provide insightful reports.DAX is based on different nested filters which magnificently improves the performance of data merging, modeling, and filtering tables.

DAX is a functional language containing conditional statements, nested functions, value references, and much more. The formulas are either numeric (integers, decimals, etc.) or non-numeric (string, binary). A DAX formula always starts with an equal sign and Name of the project,Start of the DAX formula,DAX function (to add),Parentheses defining arguments,Name of the table,Name of the field,Operator

**9. What is Power Pivot?**

Power Pivot enables you to import millions of rows from heterogeneous sources of data into a single excel sheet. It lets us create relationships between the various tables, create columns, calculate using formulas, and create PivotCharts and PivotTables.At a time there can be only one active relationship between the tables which is represented by a continuous line.

**10. What is Power Query?**

Power query is a function that filters transforms, and combines the data extracted from various sources. It helps to import data from databases, files, etc and append data

**11. Difference between Power BI and Tableau?**

The major differences between Power BI and Tableau are:

While Power BI uses DAX for calculating columns of a table, Tableau uses MDX (Multidimensional Expressions).

Tableau is more efficient as it can handle a large chunk of data while Power BI can handle only a limited amount.

Tableau is more challenging to use than Power BI.

**12. What is GetData in Power BI?**

GetData offers data connectivity to various data sources. Connect data files on your local system. The supported data sources are:

File: Excel, Text/CSV, XML, PDF, JSON, Folder, SharePoint.

Database: SQL Server database, Access database, Oracle database, SAP HANA database, IBM, MySQL, Teradata, Impala, Amazon Redshift, Google BigQuery, etc.

Power BI: Power BI datasets, Power BI dataflows.

Azure: Azure SQL, Azure SQL Data Warehouse, Azure Analysis Services, Azure Data Lake, Azure Cosmos DB, etc.

Online Services: Salesforce, Azure DevOps, Google Analytics, Adobe Analytics, Dynamics 365, Facebook, GitHub, etc.

Others: Python script, R script, Web, Spark, Hadoop File (HDFS), ODBC, OLE DB, Active Directory, etc.

**13. What are filters in Power BI?**

Filters sort data based on the condition applied to it. Filters enable us to select particular fields and extract information in a page/visualization/report level. For example, filters can provide sales reports from the year 2022 for the US region.  Power BI can make changes based on the filters and create graphs or visuals accordingly.

The types of filters are:

Page-level filters: These are applied on a particular page from various pages available within a report.

Visualization-level filters: These are applied to both data and calculation conditions for particular visualizations.

Report-level filters: These are applied to the entire report

**14. What are the types of visualizations in Power BI?**

Visualization is a graphical representation of data. We can use visualizations to create reports and dashboards. The kinds of visualizations available in Power BI are Bar charts, Column charts, Line chart, Area chart, Stacked area chart, Ribbon chart, Waterfall chart, Scatter chart, Pie chart, Donut chart, Treemap chart, Map, Funnel chart, Gauge chart, Cards, KPI, Slicer, Table, Matrix, R script visual, Python visual, etc.

**15. What do we understand by Power BI services?**

Power BI provides services for its cloud-based business analytics. With these services, you can view and share reports via the Power BI website. Power BI is a web-based service for sharing reports. Power BI service can be best referred to as PowerBI.com, PowerBI workspace, PowerBI site, or PowerBI portal.

**16. What is the comprehensive working system of Power BI?**

Power BI’s working system mainly comprises three steps:

Data Integration: The first step is to extract and integrate the data from heterogeneous data sources. After integration, the data is converted into a standard format and stored in a common area called the staging area.

Data Processing: Once the data is assembled and integrated, it requires some cleaning up. Raw data is not so useful therefore, a few transformation and cleaning operations are performed on the data to remove redundant values, etc. After the data is transformed, it is stored in data warehouses.

Data Presentation: Now that the data is transformed and cleaned, it is visually presented on the Power BI desktop as reports, dashboards, or scorecards. These reports can be shared via mobile apps or web to various business users.

**17. What are custom visuals in Power BI?**

Using Power BI visualizations, you can apply customized visualizations like charts, KPIs, etc. from the rich library of PowerBI’s custom visuals. It refrains the developers from creating it from scratch using JQuery or Javascript SDK. Once the custom visual is ready, it is tested thoroughly. Post testing, they are packaged in .pbiviz file format and shared within the organization.

Types of visuals available in Power BI are:

Custom visual files.

Organizational files.

Marketplace files.

**18. What are the various type of users who can use Power BI?**

Anyone and everyone can use PowerBI to their advantage. But even then a specific set of users are more likely to use it:

Business Users: Business users are the ones who constantly keep an eye on the reports to make important business decisions based on the insights.

Business Analysts: Analysts are the ones who create dashboards, reports, and visual representations of data to study the dataset properly. Studying data needs an analytical eye to capture important trends within the reports.

Developers: Developers are involved while creating custom visuals to create Power BI, integrating Power BI with other applications, etc.

Professionals: They use Power BI to check the data scalability, security, and availability of data.

**19.What are the advantages of Power BI?**

The advantages of Power BI that make it an excellent business intelligence software:

- It’s easy to use, even for non-technical people.

- It has a powerful toolkit for conducting ETL (extraction, transformation, and loading the data).

- It helps share the insights from the data with data consumers.

- It accommodates fast updates of the data in use from the data sources.

- It is equipped with template dashboards and SaaS solution reports.

- It allows real-time dashboard and report updates.

- It allows results displays on various devices (computers, tablets, and mobile phones).

- It ensures quick and safe connection to the data sources in the cloud or locally.

- It enables data querying using natural language processing.

- It provides hybrid configuration and smart deployment.

**20.What are some disadvantages of Power BI?**

The disadvantages of Power BI are:

The software is not very intuitive for the beginners.

Dashboard and report sharing is limited: only users with the same email domain can access the results.

The majority of data sources don't support real-time connections to Power BI interactive dashboards and reports.

Power BI for free users can't process datasets larger than 1 GB.

We can't store an adjusted filter in the saved Power BI visual report filter. In addition, the filter is always displayed on the report, which isn’t always convenient

However, Power Bi is in the process of constant development and improvement, so we can expect the software to overcome some or all of its limitations.

**21.What is a common workflow in Power BI?**

A standard Power BI workflow includes the following four steps:

- Fetch the data to the Power BI Desktop, clean and manipulate the data, and create a report.

- Publish the report to the Power BI Service and build dashboards.

- Share the dashboards with your colleagues, managers, or shareholders.

- Interact with the final dashboards and reports in Power BI Mobile apps to extract business insights.

**22. What are the main business applications of Power BI?**

Since Power BI is a business intelligence application, we can apply it to a range of business spheres. Its most crucial applications include the following:

- Extracting meaningful business insights from the available raw data

- Creating compelling live reports and insightful interactive dashboards

- Identifying the current state of different departments or projects

- Tracking progress and KPIs of different departments or projects

- Detecting the strong and weak sides of a project from the standpoint of its performance

- Distributing the roles inside the team

- Granting access to the dashboards and reports to the relevant group of team members

- Displaying various statistics of a certain business on many different applications and websites in a   favorable light for a potential customer

**23. What kind of specialists typically use Power BI?**

Project managers

Business analysts

Data analysts

Data scientists

IT specialists

Data administrators

Developers

Report consumers

**24.Where is the data stored in Power BI?**

The data in Power BI is stored in the form of either fact tables which are quantitative, usually non-normalized data or dimension tables like the attributes and dimensions related to the data in a fact table  and in one of the two cloud repositories:

Microsoft Azure Blob Storage: contains the data uploaded by the users

Microsoft Azure SQL Database: contains all the metadata and the artifacts of the system

For both, encryption and passwords protect the data.

**25.What does self-service business intelligence (SSBI) mean?**

SSBI is a set of approaches and tools that enable end users — even those without any background in BI (e.g., sales or marketing teams, product developers, etc.) — to access, manipulate, analyze, and visualize the data in an intuitive way to make strategic, data-driven business decisions.

**26.What are content packs in Power BI?**

A content pack is a package of Power BI interrelated documents, such as dashboards, reports, and datasets, that are stored as a group. In Power BI, there are two types of such packages: service content packs from services providers like Google Analytics, Marketo, MailChimp, or Twilio that we can access by typing our account data, and organizational content packs created by the users of our company and shared with the entire organization or a selected group of people.

**27.How can we define the relationships between two tables in a data model in the Power BI Desktop?**

There are two approaches:

Manual: by using primary and foreign keys

Automatic: the relationships are identified automatically if the autodetect feature is switched on

To define the relationships between two tables, there shouldn't be any null values or duplicate rows in the data. Also, it's possible to have multiple relationships between tables (represented by dotted lines), but only one of them can be active (represented by a continuous line).

The following roles comprise the majority of Power BI users:

Project managers

Business analysts

Data analysts

Data scientists

IT specialists

Data administrators

Developers

Report consumers

**24.Where is the data stored in Power BI?**

The data in Power BI is stored in the form of either fact tables which are quantitative, usually non-normalized data or dimension tables like the attributes and dimensions related to the data in a fact table  and in one of the two cloud repositories:

Microsoft Azure Blob Storage: contains the data uploaded by the users

Microsoft Azure SQL Database: contains all the metadata and the artifacts of the system

For both, encryption and passwords protect the data.

**25.What does self-service business intelligence (SSBI) mean?**

SSBI is a set of approaches and tools that enable end users — even those without any background in BI (e.g., sales or marketing teams, product developers, etc.) — to access, manipulate, analyze, and visualize the data in an intuitive way to make strategic, data-driven business decisions.

**26.What are content packs in Power BI?**

A content pack is a package of Power BI interrelated documents, such as dashboards, reports, and datasets, that are stored as a group. In Power BI, there are two types of such packages: service content packs from services providers like Google Analytics, Marketo, MailChimp, or Twilio that we can access by typing our account data, and organizational content packs created by the users of our company and shared with the entire organization or a selected group of people.

**27.How can we define the relationships between two tables in a data model in the Power BI Desktop?**

There are two approaches:

Manual: by using primary and foreign keys

Automatic: the relationships are identified automatically if the autodetect feature is switched on

To define the relationships between two tables, there shouldn't be any null values or duplicate rows in the data. Also, it's possible to have multiple relationships between tables (represented by dotted lines), but only one of them can be active (represented by a continuous line).