Module 5) Creating Dashboard with Visualization Tool

Assignment

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

Power BI is a business analytics service provided by Microsoft that lets you visualize your data and share insights. It converts data from different sources to build interactive dashboards and Business Intelligence reports.

Excel is an excellent tool for small to medium-sized data sets with a lower budget. Whereas, Power BI caters to large data sets, real-time action, and cloud capabilities.

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

Data modeling is the process of creating visual representations of multiple tables or dataset connections. These datasets have attributes and fields with relevant information.

A data model is an organized visual representation of different data elements, their interconnections, and their relation with business needs or events. It helps in extracting data, transforming the data, and loading the data in the form of visuals which helps in making important business decisions.

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

In Power BI there are 3 main data Connection Types which are:

1. Import
2. Direct Query
3. Live

1.import data connection: The Import Data connection type pulls the whole dataset from the data source and stores it in the memory of the PBI file either on-prem or in the cloud. In this case, the loaded data needs to be refreshed at a scheduled time to avoid being obsolete.

2.Direct Quary: When using the DirectQuery connection type, the Data schema is pulled directly from the data source in real time and stored in Power BI Desktop Model. While the Dataset remains in the data source. All data transformations and data modeling are done at the source i.e. outside of Power BI. It is usually preferred when the data source is really large.

Unlike Import Data Connection Type, DirectQuery does not support all data sources but mostly supports Data sources with Relational Database Models (RDMs). Below is a list of data sources DirectQuery Supports.

3.Live: When using the DirectQuery connection type, the Data schema is pulled directly from the data source in real time and stored in Power BI Desktop Model. While the Dataset remains in the data source. All data transformations and data modeling are done at the source i.e. outside of Power BI. It is usually preferred when the data source is really large.

Unlike Import Data Connection Type, DirectQuery does not support all data sources but mostly supports Data sources with Relational Database Models (RDMs). Below is a list of data sources DirectQuery Supports.

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

Step 1: Accessing Power Query Editor

Open Power BI Desktop.

Navigate to ‘Home’ > ‘Transform Data’ to open Power Query Editor.

Step 2: Executing Transformations

**Cleaning Data: Cleaning data often involves multiple steps, such as removing duplicates, filtering rows, or handling missing values.**

* **Remove Duplicates**: Select the column(s) from which you want to remove duplicates, then go to the "Home" tab and click "Remove Duplicates":

Remove Duplicates in Power BI

* **Remove Columns:** Right-click the column you want to remove and select "Remove":

Remove Columns in Power BI

* **Replacing Values**: Replace specific values within your data, such as substituting "N/A" with null values or standardizing names.

Replacing Missing Values with N/A

* **Type Conversion**: Click on the column you want to convert. In the "Transform" tab, choose "Data Type" and select the appropriate type (e.g., Text, Whole Number, Decimal Number, Date, etc.):
* **Pivoting/Unpivoting Data**: Under the ‘Transform’ tab, use ‘Pivot Column’ for pivoting or ‘Unpivot Columns’ for unpivoting. To pivot, select the column that contains the values which will become new column headers. Go to "Transform" > "Pivot Column". Choose the column to pivot on and the values column:
* Pivoting a Column
* **Merging Data**:
* Open Power Query Editor: Go to "Home" > "Edit Queries".
* New Query: Click "Merge Queries" in the "Home" tab.
* Select the Main Table: Choose the first table from the dropdown.
* Select the Table to Merge: Choose the second table from the dropdown.
* Choose Columns to Match: Select the matching columns in both tables.
* Specify Join Kind: Choose the type of join (e.g., Inner, Full Outer, Left Outer, Right Outer).
* OK: Click OK to merge the tables.
* Merging tables with a common column
* **Adding Conditional Columns**: From the ‘Add Column’ tab, choose ‘Conditional Column’ and define the conditions and outputs for each scenario:
* Adding Conditional Columns
* **Grouping Data:** In the "Home" tab, click "Group By". Choose the column to group by, and specify the aggregation function (e.g., Count, Sum, Average, Max, Min) and the column to perform the function on.
* Grouping Data
* Step 3: Applying and Loading Transformations
* After performing the desired transformations, click ‘Close & Apply’ to load the transformed data into Power BI for reporting and visualization.

5) What is DAX (Data Analysis Expressions) and why is it important in

Power BI?

Data Analysis Expressions (DAX) is a formula expression language used in Analysis Services, Power BI, and Power Pivot in Excel. DAX formulas include functions, operators, and values to perform advanced calculations and queries on data in related tables and columns in tabular data models.

6) Can you explain the difference between calculated columns and

measures in Power BI?

Calculated Columns: Computed once during the process refresh, stored in the model, and occupy space. Measures: Computed at query time, exist as source code in the model, and do not consume additional space

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

when more than one table is loaded, Power BI automatically detects the relationship between tables by looking at the column names. If Power BI can’t determine any potential relationship, it doesn’t create any. The table relationships can be viewed from the model view option present on the left side of the screen.

If after loading all the tables, still no relationship can be seen between tables, then the **Autodetect**tool comes in handy. To use Autodetect, on the **Modeling tab** select **Manage Relationships** > **Autodetect.**

But sometimes Power BI doesn’t create the correct relationships so there’s also an option of creating relationships between table manually. For this follow the steps below –

1. On the **Modeling tab**, click **Manage Relationships** > **New**..
2. Select the tables and columns that has a relation from the tables that were loaded. After select both tables and columns, certain options will be configured automatically such as Cardinality and Cross filter direction. It can be changed.
3. Click**OK**.

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

The Power BI Gateway tool is a software application that helps users access the required data residing in an on-premises network. The tool acts as a gatekeeper for the source of data, and any requests made by users to access the data from a cloud or web-based application go through this gateway.

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

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

Row-level security (RLS) is a type of data governance that allows developers and admins to limit the data an end-user has visibility to within a Power BI report and/or dataset based on the logic applied at the row level. You can think of RLS as a horizontal limitation applied to the dataset, whereas object-level security is applied to columns or tables.

11) What is the Power BI Desktop and how does it differ from Power

BI Service?

The biggest advantage of Power BI Service is its ability to provide real-time data updates and enhance collaboration. While Power BI Desktop excels at crafting individual reports, opting for Power BI online promotes teamwork by enabling multiple users to collaborate on the same dataset utilizing workspaces.

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

DirectQuery is a connectivity mode in Power BI that allows you to extract data directly from database sources like SQL Server and Snowflake. When you use this mode, Power BI creates schema (metadata) that defines the model structure of the connected source.

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

Power BI templates are pre-designed report layouts that include the report pages, visuals, data model schema, and queries. They ensure consistency and save time by allowing users to recreate reports without including the actual data.

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

1. In Data view, right-click a table in the Data pane and select Incremental refresh.
2. In Incremental refresh and real-time data > Select table, verify or select the table. ...
3. Specify required settings: ...
4. Specify optional settings: ...
5. Review your settings and then select Apply to complete the refresh policy.

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

 Power Query Editor in Power BI is a powerful and versatile tool that enables users to shape, clean, and transform data from various sources, making it suitable for analysis and visualization, ultimately enhancing the effectiveness and accuracy of business decision-making .

16) Explain the difference between calculated columns and

calculated tables in Power BI.

**Calculated Columns**: **Purpose**: Calculated columns are additional columns that you create within a table in your data model.

**Calculated Tables**: **Purpose**: Calculated tables are new tables that you create in your data model based on calculations or filters applied to existing tables.

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

**Power BI** comes with core **visuals** readily available on the **visualization** pane. You can also import **visuals** from Microsoft AppSource or **Power BI**.

18) What are the best practices for optimizing performance in

Power BI?

* Use STAR Schema. Power BI is built to perform most efficiently when using a STAR schema. ...
* Limit Data Model. ...
* Push Data Transformations Upstream. ...
* Move Row Level Logic to Power Query​ ...
* Use Measures Instead of Calculated Columns. ...
* Convert Multiple Measures to Variables. ...
* Amend Dates and Column Types. ...
* Minimize Visuals.

19) How can you integrate Power BI with other Microsoft

products like Azure and Office 365?

1. Power BI App for Teams: Install the Power BI app for Microsoft Teams. ...
2. Data Sharing: Share Power BI reports and dashboards directly within Teams conversations. ...
3. Q&A: Utilize the Q&A feature within Power BI reports embedded in Teams.

20) Explain the concept of aggregations in Power BI.

When you combine values in your data, it's called aggregating. The result of that mathematical operation is an aggregate. When you create visualizations in Power BI Desktop and the Power BI service, they may aggregate your data.

21) How do you handle error handling and data quality in Power

BI?

There are two main types of error in the[Power Query](https://datascientest.com/en/power-query-what-is-it-whats-it-for) editor: step and column errors.

**Stage level**

If an error occurs in the step, the request cannot be loaded. You will then see an error message appear in a yellow window.

This provides information on the various components of the error: reason, message and details

And if you need more information, you can also click on the “Access error” button. It’s all this data that enables you to adopt effective error management in **Power Query.**

Here are the most common errors:

**Source not found:** this error appears if you can’t access the data source, if you don’t have identification information or if the source has been moved. In such cases, you should modify the access path.

**Column not found:** this is what happens if you’ve specified a column that doesn’t exist in the query. There are several possible solutions here, including deleting the step referring to the column in question.

**Formula.firewall:** this error occurs when you combine or merge data from several sources. Before applying a solution, you need to understand the exact cause of the problem by accessing the data confidentiality firewall.

**Column level**

In this case, the **query loads, but the value appears as an error in the corresponding cell.** To understand the origin of the error, you can click on the white space of the cell concerned (a new yellow window will appear) or use the data profiling tools.

These errors can be linked to a data type conversion problem, unsupported operations, nested structured values, etc.

Whatever the problem, there are a number of error management solutions available with**Power Query.** Here’s what you need to know:

**Deleting errors:** select the column concerned, then go to the “Reduce rows” group under the “Home” tab. Click on “Delete rows”, then on “Delete errors”.

**Replacing errors:** select the column concerned, then go to the “Any column” group under the “Transform” tab. Click on “Replace values”, then on “Replace errors”. Then enter the new value (as an integer, custom text, etc.).

**Keeping errors:** this makes it easier to identify lines containing errors. The procedure is as follows: Home > Reduce lines > Keep lines > Keep errors.

22) What is the purpose of Power BI Embedded and when would

you use it?

Microsoft Power BI Embedded (PBIE) allows application developers to embed stunning, fully interactive reports into their applications without having to build their own data visualizations and controls from scratch.