

BASIC FUNCTIONALITY OF POWER BI AND DATA TRANSFORMATION

- Workflow of Power BI
- Get Data Source.
- Data Transformation Process
- Transform Column
- Add Column
- Layout of Power Query Editor
- Understanding Each & Every Features of PQE.
- Create Your Own Data

	Product ID	Product Name	Product Rank	+
1	1	Pen	2	
2	2	Pencil	3	
3	3	Pen Set	1	
4	4	Desk	5	
5	5	Binder	4	
+				



PIVOTING & UNPIVOTING

Pivoting describes the process of turning distinct row values into columns, and unpivoting describes the process of turning distinct columns into rows

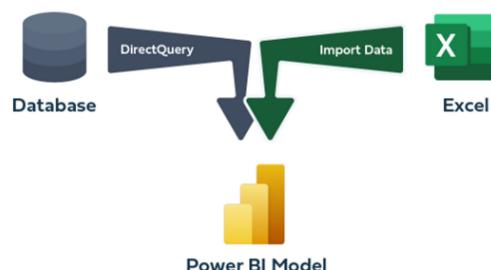
Import vs Direct Query

While connecting to a database, Power BI provides you with options to choose from:

Import:

- Data is stored in Power BI Azure database, with a 1 GB limit.
- It allows you to transform and manipulate data.
- We can schedule refreshes for upto 8 times a day on the normal license and upto 48 times a day on the Premium Pro license.

Composite Model in Power BI



Direct Query:

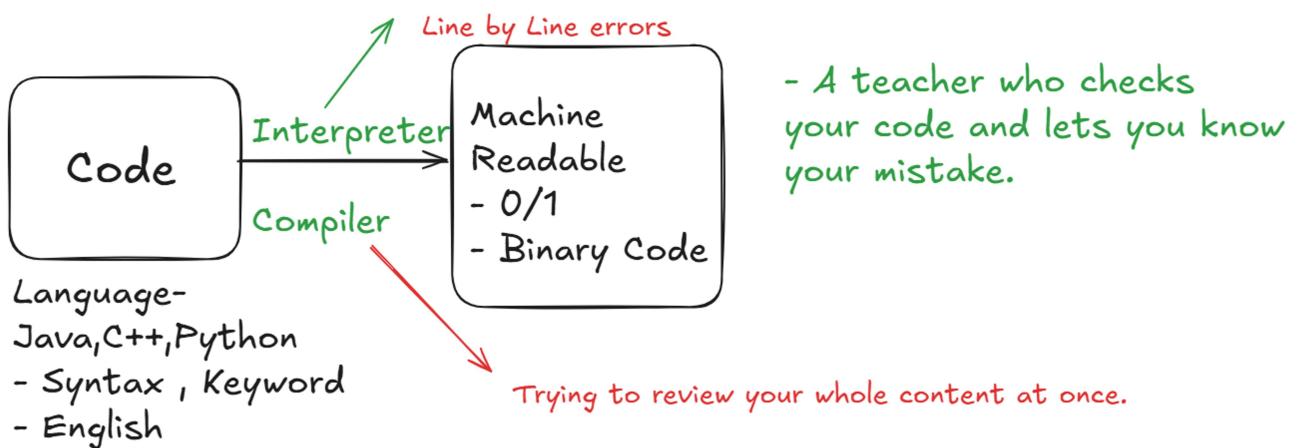
- Data is queried directly from the original database and is not stored in Power BI.
All tables should come from a single dataset.
- It limits the option to transform/manipulate data.
- It is a real time query and there is no need to schedule a refresh.

Live Connection:

- This is similar to Direct Query as it doesn't store any data in Power BI.
- It is different from Direct Query in that it is a direct connection to the analysis services model. Instead of being a connection to a dataset, it is a connection to a model.
- The supported data sources with live connection are limited.
Example: SQL Server Analysis Services (SSAS) Tabular, Power BI Service

When to use which kind of connection:

- Import Mode: is generally preferred for static or slowly changing datasets where high performance and complex data modeling are required within Power BI. It's also suitable when working with data that needs to be transformed or enriched before analysis.
- Direct Query Mode: is beneficial when dealing with very large datasets that cannot be easily imported due to size limitations or when real-time data updates are critical for the analysis. However, it's important to consider the potential impact on query performance and the limitations on data transformations.



SQL As a
Query Language



Its collection of Tables

Workflow Of Power BI

Step1: Get the data - Fetch Data from multiple Source.

Connectors

Step2: ETL: Extract, Transform & Load the data.

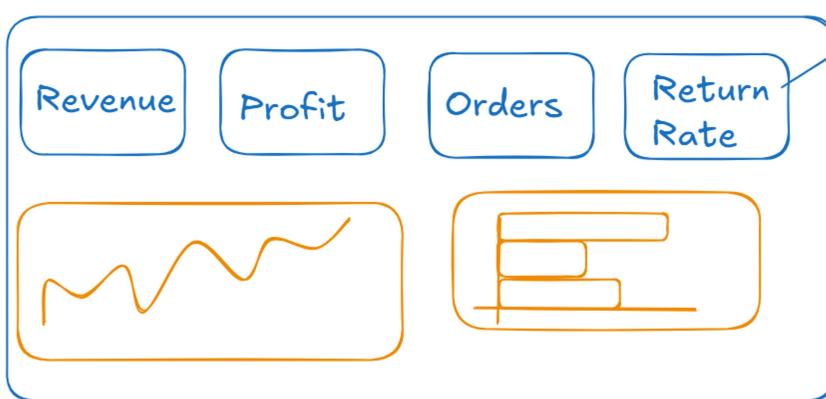
Step2.1: Power Query Editor - It's a place where you clean the data.

Step3: Load the data to the Front End.

Step4: Building a relationship between multiple tables, using P.K {Primary Key} & F.K {Foreign Key}

Data Modelling

Step5: Data View/Table View - DAX[Data Analysis Expression] to write some advance calculation, Measures [KPI, Matrices.]



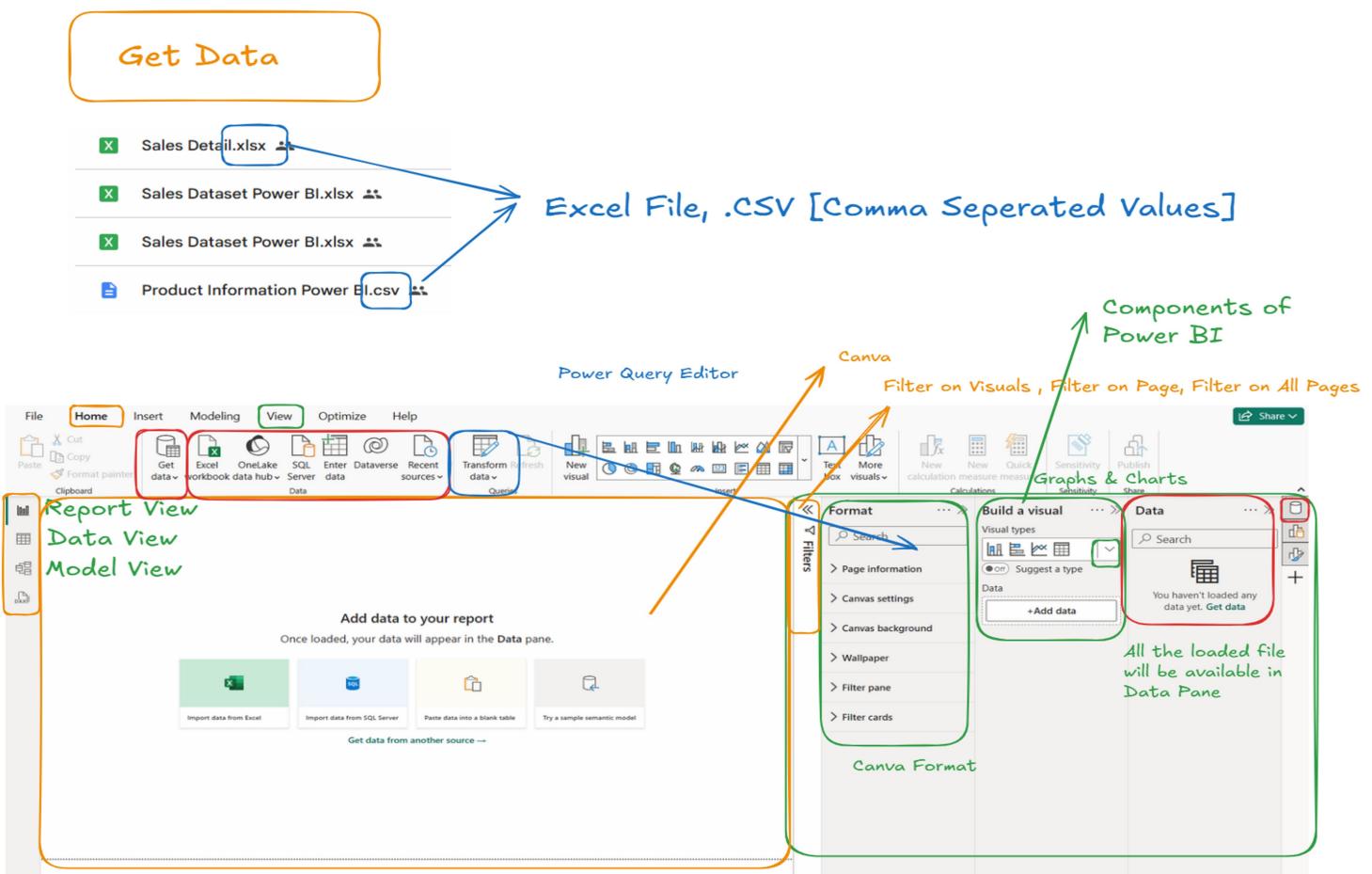
$$\text{Profit} = \text{Sales} - \text{Expenditure}$$

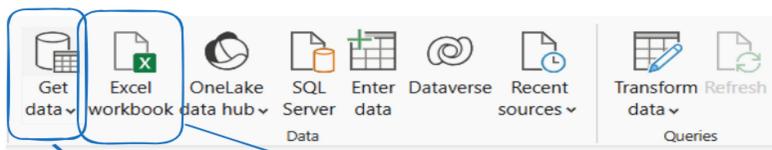
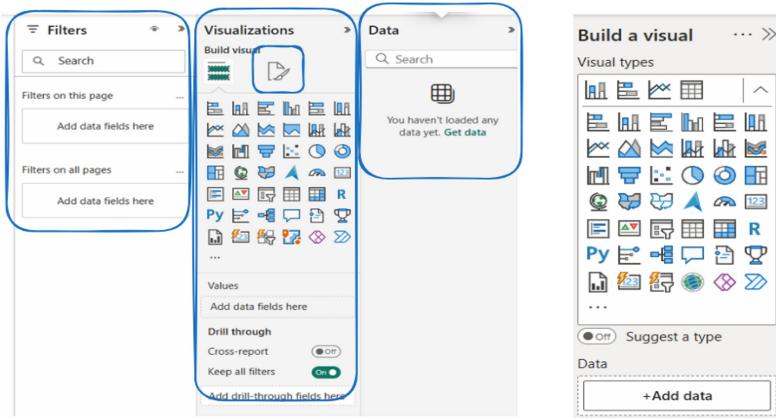
$$5510 - 3075 \\ = 2435$$

Quantity	Product Price	Product Cost	Sales	Expenditure	Revenue
10	100	50	1000	500	Sum of Sales 5510
20	50	20	100	400	
30	20	10	60	300	
10	10	5	100	50	
5	10	5	50	25	
10	100	40	1000	400	
100	25	10	2500	1000	
20	35	20	700	400	
					Qty * P.Cost 3075

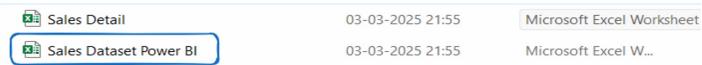
Step6: [Report View] Using Visuals, you have to build the dashboard.
 & Multiple dashboards is nothing but a report

Step7: Publish the Dashboard [Netflix Project] [Sign In]

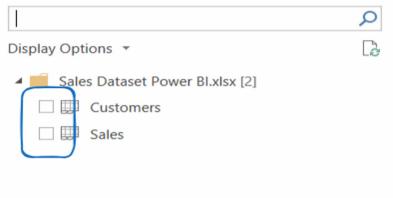




Import The Data



Navigator



Navigator

The Navigator displays a preview of the 'Sales' dataset. The columns are Customer ID, Customer Region, Customer Age, and Sales. The data includes rows such as North, 49, 309.38; North, 10, 57.71; North, 47, 686.95; South, 11, 1619.19; East, 17, 1005.9; null, 23, 174.65; null, 21, 825; null, 27, 250; null, 44, 9.03; null, 10, 255.84; North, abc, 151.24; South, 29, 251.72; South, 42, 1139.43; North, 12, 575.36; null, 45, 18.06; null, 30, 299.85; null, 35, 54.89; West, 50, 479.04; null, 42, 1879.06; null, 44, 86.43.

P.Q.E [Power Query Editor]
- Cleaning Process.

Load Transform Data Cancel

P.Q.E



Bring the data to Front End



Power Query Editor

Properties

Name
Customers

Applied Steps

Source
Navigation

Tracking

File Home Transform Add Column View Tools Help

Transform

in place changes - Manipulation on existing data

File Home Transform Add Column View Tools Help

Add Column

Adding a new Column.

File Home Transform Add Column View Tools Help

Conditional Column

Monospaced Column distribution

Show whitespace Column profile

Column quality

Data Preview

Monospaced Column distribution Always allow Advanced Editor

Show whitespace Column profile Go to Column Columns Parameters Advanced

Column quality

Data Preview Advanced Editor

View

= Table.TransformColumnTypes(#"Promoted Headers", {{"Customer ID", Int64.Type}, {"Customer Name", Text.Type}})

	Customer ID	Customer Name
1	1	Elvin Catron
2	2	Gustano Primmin
3	3	Lacresha Whitty
4	4	Octavio Ricci
5	5	Jacinda Moffett
6	6	Betty scruggs
7	7	Era Gassner
8	8	Laticia Crimi
9	9	Alberta Fabela
10	10	Taylor Pogue
11	11	Grisede Mariscal
12	12	Sylvester Willbourn
13	13	Ricky Cheshire
14	14	Karisa Oquwndo
15	15	Debbie Orone
16	16	Loralee Widell
17	17	Season Viers
18	18	Juan Luker
19	19	Shan texeria
20	20	Yee Bordeau

Transform Add Column View Tools Help

Monospaced Column distribution Always allow Advanced Editor Query Dependencies

Show whitespace Column profile Go to Column Columns Parameters Advanced Dependencies

Column quality

Data Preview Advanced Editor

View

= Table.TransformColumnTypes(#"Promoted Headers", {{"Customer ID", Int64.Type}, {"Customer Name", Text.Type}})

	Customer ID	Customer Name
1	1	Elvin Catron
2	2	Gustano Primmin
3	3	Lacresha Whitty
4	4	Octavio Ricci
5	5	Jacinda Moffett
6	6	Betty scruggs
7	7	Era Gassner
8	8	Laticia Crimi
9	9	Alberta Fabela
10	10	Taylor Pogue
11	11	Grisede Mariscal
12	12	Sylvester Willbourn
13	13	Ricky Cheshire
14	14	Karisa Oquwndo
15	15	Debbie Orone
16	16	Loralee Widell
17	17	Season Viers
18	18	Juan Luker
19	19	Shan texeria
20	20	Yee Bordeau

Transform Add Column View Tools Help

Monospaced Column distribution Always allow Advanced Editor Query Dependencies

Show whitespace Column profile Go to Column Columns Parameters Advanced Dependencies

Column quality

Data Preview Advanced Editor

View

= Table.TransformColumnTypes(#"Promoted Headers", {{"Customer ID", Int64.Type}, {"Customer Name", Text.Type}})

	Customer ID	Customer Name
1	1	Elvin Catron
2	2	Gustano Primmin
3	3	Lacresha Whitty
4	4	Octavio Ricci
5	5	Jacinda Moffett
6	6	Betty scruggs

Column statistics

Count	20
Error	0
Empty	0
Distinct	20
Unique	20
Nan	0
Zero	0
Min	1

Value distribution

* Select the particular column to get the column statistics, Value Distribution under Column Profile Tool.

Distinct VS Unique

Cricket	Football
Cricket	Football
Football	Badminton
Swimming	Cricket
Swimming	Tennis
Volleyball	Tennis
Volleyball	Tennis

Distinct - 6

- Any data that comes one or more than that will be distinct.

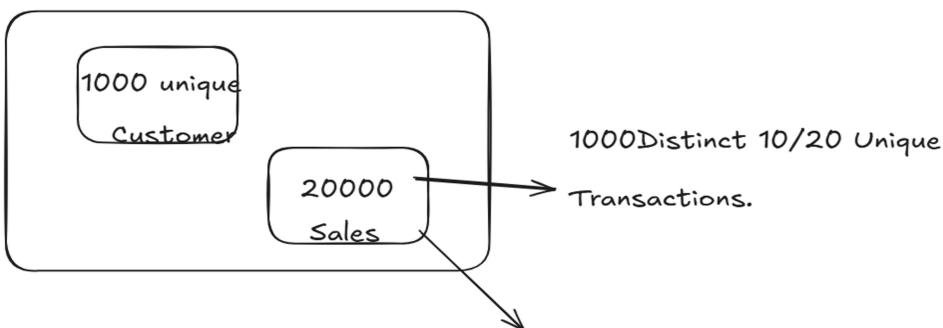
Cricket - 3
Football - 3
Badminton - 1
Tennis - 4
Volleyball - 2
Swimming - 2

Unique - 1

- Only appears once.

Badminton

Amazon Sales Dashboard



Group BY

Grouping being done on Categorical Data and all the other numerical column will be taken as a reference to perform Aggregation like doing- SUM, Avg, Min, Max, etc.

The screenshot shows the Power BI interface with two tables: "Customers" and "Sales". The "Sales" table is currently selected. The top ribbon shows various data transformation tools like Transpose, Reverse Rows, and Group By. The "Customer Region" column is highlighted, and a "Group By" operation is being performed. The resulting table on the right shows the total sales for each region, with a breakdown of valid, error, and empty values. The table includes columns for Customer Region, Total Sales, and a detailed breakdown of the 1.2 Total Sales.

Customer Region	Total Sales	Valid	Error	Empty
North	1722.93	83%	0%	17%
South	3010.34	83%	0%	17%
East	1005.9	83%	0%	17%
West	479.04	83%	0%	17%
Unknown	3910.52	83%	0%	17%

Identity Property = 0 when you do sum.

Identity Property is 1 when you do multiplication.