

Data Visualization Tools and Software

Aim: To create a dashboard in PowerBI for grocery store dataset.

Task:

1. Introduction to Data Modelling
2. Import Sample Data
3. Configure table and column properties
4. Create calculated columns and tables
5. Create quick measures.
 - a. Use DAX(Data Analysis Expression) for creating various expressions and filters
 - b. Optimize data models
4. Create Visualization

What is data modeling?

Data modeling is the process of analyzing and defining all the different data types your business collects and produces, as well as the relationships between those bits of data. By using text, symbols, and diagrams, data modeling concepts create visual representations of data as it's captured, stored, and used at your business. As your business determines how data is used and when, the data modeling process becomes an exercise in understanding and clarifying your data requirements.

The benefits of data modeling

By modeling your data, you can document what types of data you have, how you use it, and the data management requirements surrounding its usage, protection, and governance. The benefits of data modeling include:

- Creating a structure for collaboration between your IT and business teams.
 - Revealing opportunities for improving business processes by defining data needs and uses.
 - Saving time and money on IT and process investments through appropriate planning.
 - Reducing errors (and error-prone redundant data entry) while improving data integrity.
 - Increasing the speed and performance of data retrieval and analytics by planning for capacity and growth.
 - Setting and tracking target key performance indicators tailored to your business objectives.
- it's not just about the results of data modeling, but how you get those results.

Data modeling concept examples

Now that you know what data modeling is and why it's important, let's look at the three different types of data modeling concepts as examples.

1. Conceptual data modeling

A conceptual data model defines the overall structure of your business and data. Used for organizing business concepts, your conceptual data model is defined by your business stakeholders and data engineers or architects. For instance, you may have customer, employee, and product data and each data bucket, known as entities, has relationships with other entities. Both the entities and the entity relationships are defined within your conceptual data model.

2. Logical data modeling

A logical data model builds upon the conceptual data model with specific attributes of data within each entity and the relationships between those attributes. For instance, Customer A buys Product B from Sales Associate C. This is your technical model of the rules and data structures as defined by data engineers, architects, and business analysts, helping drive decisions about what physical model your data and business require.

3. Physical data modeling

A physical data model is your specific implementation of the logical data model created by database administrators and developers. It is developed for a specific database tool and data storage technology, and with data connectors to serve the data throughout your business systems to users as needed. This is the “thing” the other models have been leading to—the actual implementation of your data estate.

Core modeling features

- Support for different types of data models
- Custom field creation with defined attributes
- User-defined relationships for constraints and entities
- Model subsets and model validation
- Atomic data elements definition
- Business vocabulary definition
- New model development from existing models
- Collaboration features

Usability features

- Object query
- Import and export capabilities
- Management of windows, toolbars and menus
- Easy-to-use graphical interface
- Diagram auto-layout
- Query creation with ordering, grouping, filters and joins

Create data

To create a data model in Power BI, you need to add all data sources in Power BI new report option. To add a data source, go to the Get data option.

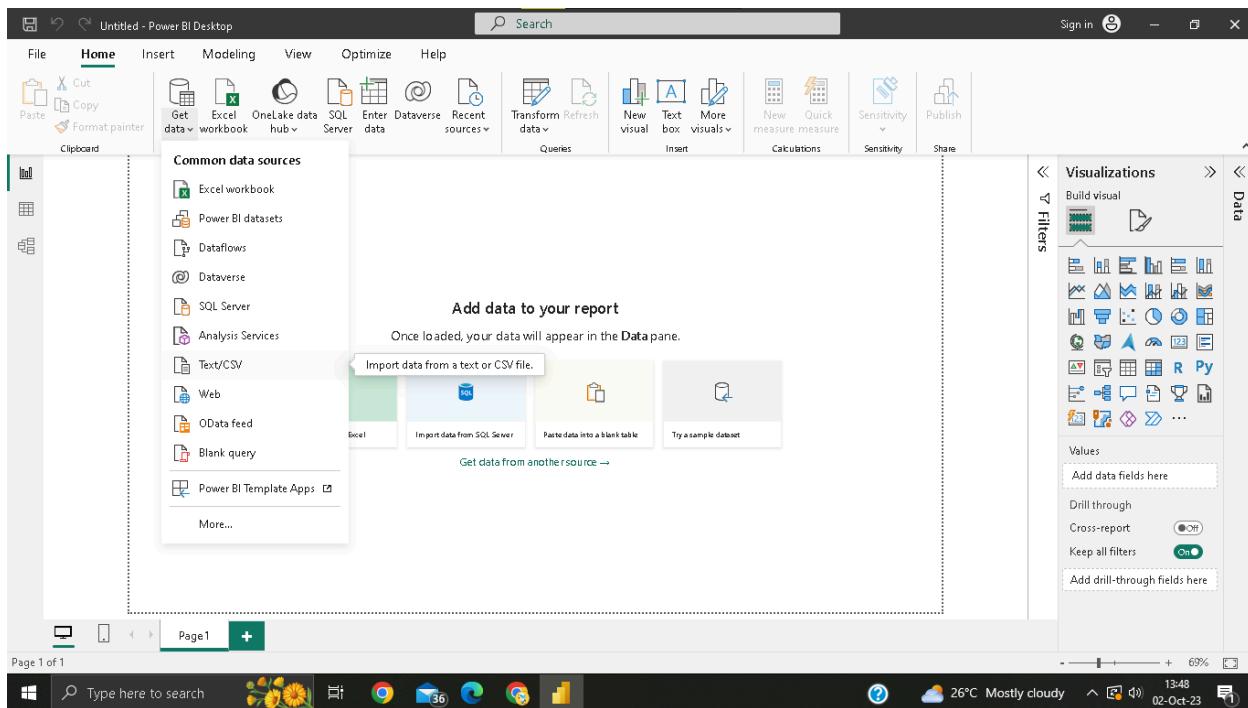


Fig.- Add all data sources in Power BI

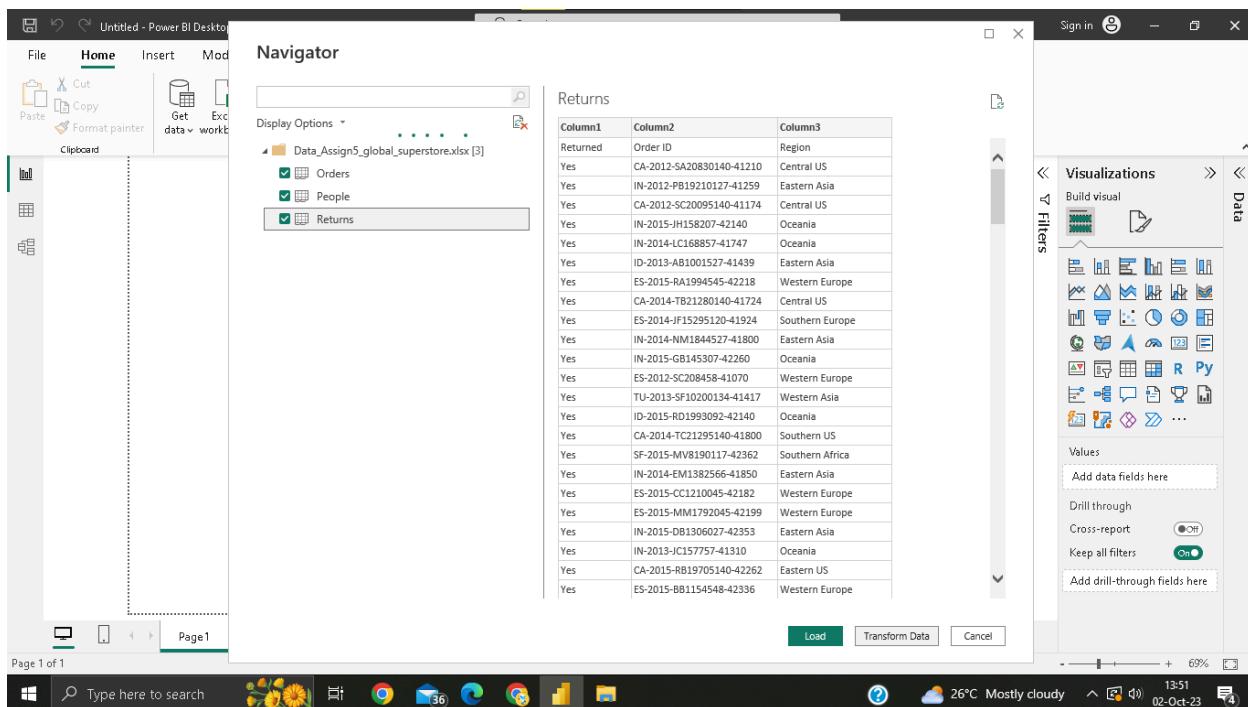


Fig.- Importing data from a CSV file

Remove Empty

When you have blank values in your data table, you can use Power Query transformations to remove the row with blank values.

The screenshot shows the Power Query Editor interface. The main area displays a table with columns: Customer Name, Segment, Postal Code, City, State, and Country. The Postal Code column contains numerous empty rows, indicated by the word 'Empty' and a percentage of 0% in the validation summary at the top. The validation summary also shows 19% of valid entries. The right pane shows the 'APPLIED STEPS' section, which includes a step named 'Changed Type' under the 'Promoted Headers' step. The status bar at the bottom indicates 'PREVIEW DOWNLOADED AT 13:51'.

Fig.-Postal code column with empty rows

This screenshot shows the same Power Query Editor interface as above, but with a different focus. A context menu is open over the 'Postal Code' column, specifically over one of the empty rows. The menu options include 'Copy Quality Metrics', 'Keep Duplicates', 'Keep Errors', 'Remove Duplicates', 'Remove Empty' (which is highlighted in red), 'Replace Errors', 'Show the column quality peek', and 'Show the column distribution peek'. The rest of the table and the right-hand pane remain the same as in the previous screenshot.

Fig.-Remove empty rows from postal code column

Use Headers as First Row

To promote the first row to column headers, select Home > Use First Row As Headers.

The screenshot shows the Power Query Editor interface with the 'People' query selected. The 'Applied Steps' pane indicates a 'Changed Type1' step was applied. The table structure is as follows:

	Person	Region
1	Marielene Rousseau	Caribbean
2	Andile Ihejirika	Central Africa
3	Nicodemo Bautista	Central America
4	Cansu Peynirci	Central Asia
5	Lon Bonher	Central US
6	Wasswa Ahmed	Eastern Africa
7	Hadiia Bousaid	Eastern Asia
8	Lynne Marchand	Eastern Canada
9	Oxana Lagunov	Eastern Europe
10	Dolores Davis	Eastern US
11	Lindwiwe Afoleyan	North Africa
12	Milna Nylund	Northern Europe
13	Kauri Anaru	Oceania
14	Vasco Magalhães	South America
15	Preecha Metharom	Southeastern Asia
16	Nora Culiper	Southern Africa
17	Chandrakant Chaudhri	Southern Asia
18	Gavino Bove	Southern Europe
19	Flannery Newton	Southern US
20	Kattlego Akosua	Western Africa
21	Kaoru Xun	Western Asia

Fig.-Use First Row As Headers in People Table

The screenshot shows the Power Query Editor interface with the 'Returns' query selected. The 'Applied Steps' pane indicates a 'Changed Type1' step was applied. The table structure is as follows:

	Returned	Order ID	Region
1	Yes	CA-2012-SA20830140-41210	Central US
2	Yes	IN-2012-PB19210127-41259	Eastern Asia
3	Yes	CA-2012-SC20095140-41174	Central US
4	Yes	IN-2015-JH158207-42140	Oceania
5	Yes	IN-2014-LC168857-41747	Oceania
6	Yes	ID-2013-AB1001527-41439	Eastern Asia
7	Yes	ES-2015-RA199455-42218	Western Europe
8	Yes	CA-2014-TB21280140-41724	Central US
9	Yes	ES-2014-JF15295120-41924	Southern Europe
10	Yes	IN-2014-NM1844527-41800	Eastern Asia
11	Yes	IN-2015-GB145307-42260	Oceania
12	Yes	ES-2012-SC208458-41070	Western Europe
13	Yes	TU-2013-SF10200134-41417	Western Asia
14	Yes	ID-2015-RD1993092-42140	Oceania
15	Yes	CA-2014-TC21295140-41800	Southern US
16	Yes	SF-2015-MV8190117-42362	Southern Africa
17	Yes	IN-2014-EM1382566-41850	Eastern Asia
18	Yes	ES-2015-CC1210045-42182	Western Europe
19	Yes	ES-2015-MM1792045-42199	Western Europe
20	Yes	IN-2015-DB1306027-42353	Eastern Asia
21	Yes	IN-2013-JC157757-41310	Oceania

Fig.-Use First Row As Headers in Returns Table

Column Quality

Shows the percentage of rows within a column that are valid, empty, or errors.

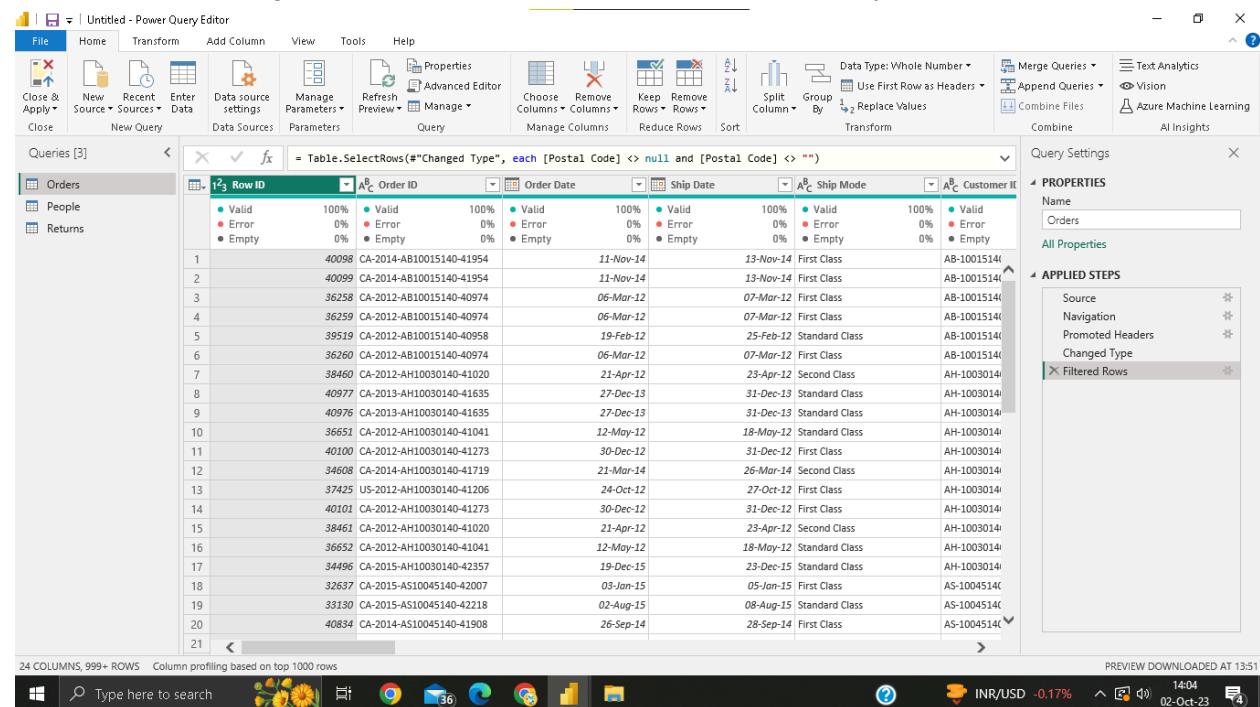


Fig.-Column quality details of Orders Table

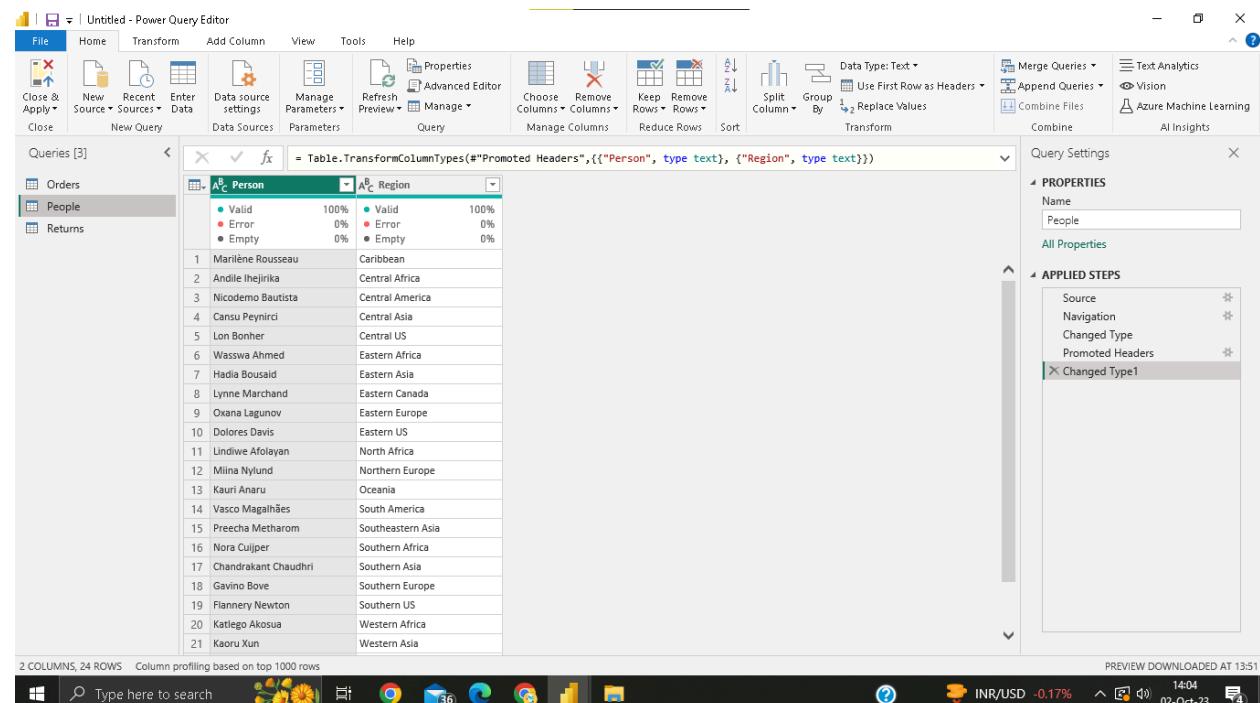


Fig.-Column quality details of People Table

The screenshot shows the Power Query Editor interface with the 'Returns' query selected. The main area displays a table with columns: 'Returned', 'Order ID', and 'Region'. A 'Column Quality' summary table is overlaid on the top row, showing counts for 'Valid', 'Error', and 'Empty' status across three categories. The 'APPLIED STEPS' pane on the right shows the 'Promoted Headers' step. The bottom status bar indicates 'PREVIEW DOWNLOADED AT 13:51'.

	Returned	Order ID	Region
● Valid	100%	● Valid	100%
● Error	0%	● Error	0%
● Empty	0%	● Empty	0%
1 Yes	CA-2012-SA20830140-41210	Central US	
2 Yes	IN-2012-PB19210127-41259	Eastern Asia	
3 Yes	CA-2012-SC20095140-41174	Central US	
4 Yes	IN-2015-JH158207-42140	Oceania	
5 Yes	IN-2014-LC168857-41747	Oceania	
6 Yes	ID-2013-AB1001527-41439	Eastern Asia	
7 Yes	ES-2015-RA1994545-42218	Western Europe	
8 Yes	CA-2014-TB21280140-41724	Central US	
9 Yes	ES-2014-JF15295120-41924	Southern Europe	
10 Yes	IN-2014-NM1844527-41800	Eastern Asia	
11 Yes	IN-2015-GB145307-42260	Oceania	
12 Yes	ES-2012-SC208458-41070	Western Europe	
13 Yes	TU-2013-SF10200134-41417	Western Asia	
14 Yes	ID-2015-RD1993092-42140	Oceania	
15 Yes	CA-2014-TC21295140-41800	Southern US	
16 Yes	SF-2015-MV8190117-42362	Southern Africa	
17 Yes	IN-2014-EM1382566-41850	Eastern Asia	
18 Yes	ES-2015-CC1210045-42182	Western Europe	
19 Yes	ES-2015-MM1792045-42199	Western Europe	
20 Yes	IN-2015-DB1306027-42353	Eastern Asia	
21 Yes	IN-2013-JC157757-41310	Oceania	

Fig.-Column quality details of Returns Table

Save your work

When our query is where we want , select Close & Apply from Power Query Editor's File menu.
This action applies the changes and closes the editor.

The screenshot shows the Power Query Editor interface with the 'Returns' query selected. The main area displays the same table as before. The bottom status bar indicates 'PREVIEW DOWNLOADED AT 13:51'.

Fig.-Save and Close the Power Query editor

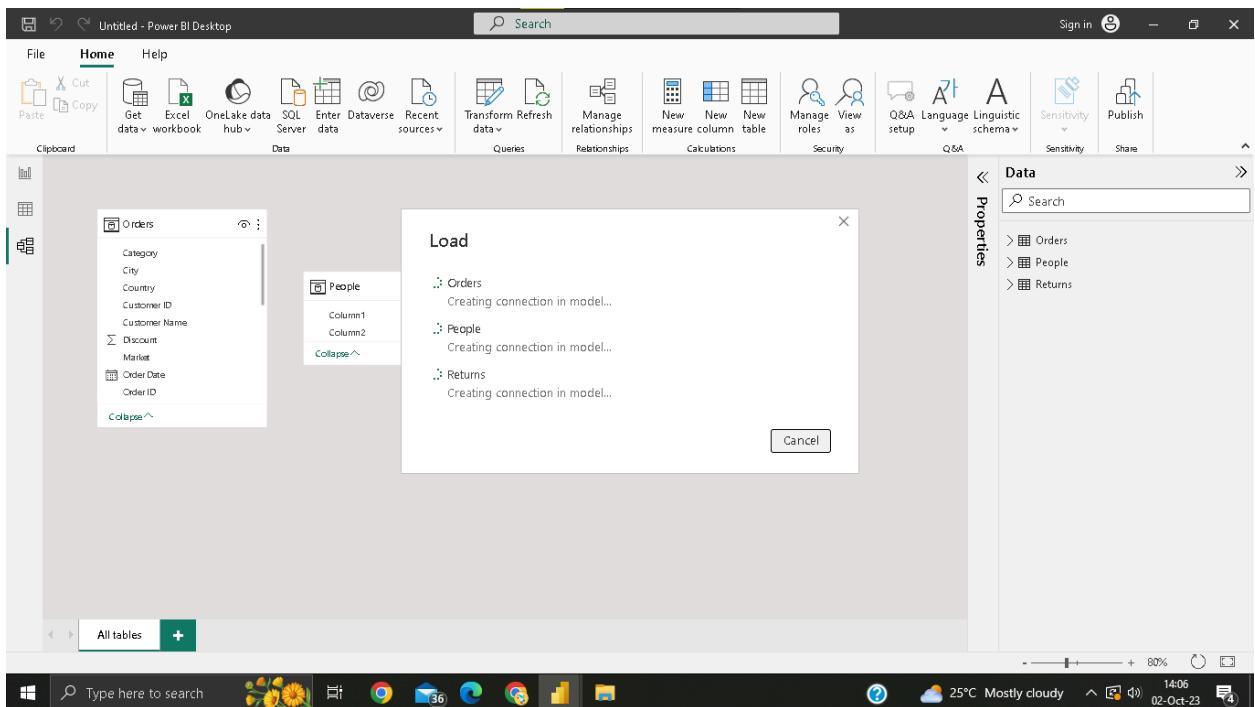


Fig.-Creating model

Report tab

When you navigate to the Report tab, you can see a dashboard and a chart selected for data visualization.

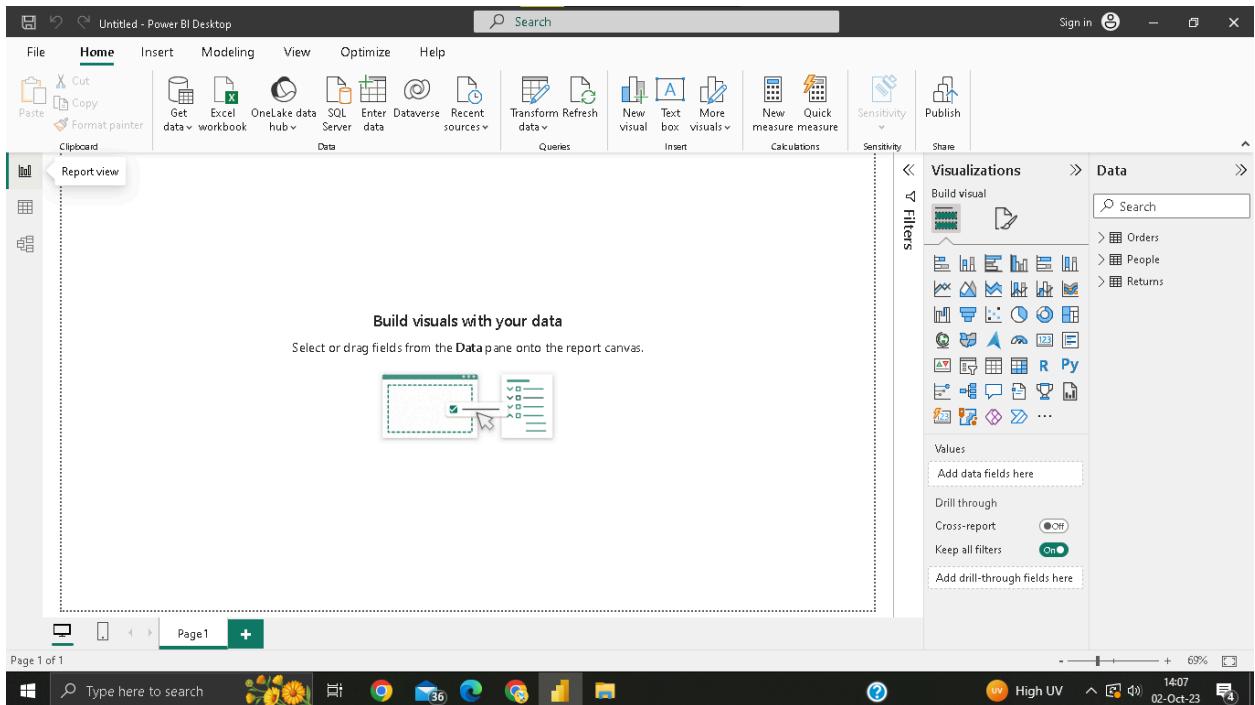


Fig.-Report view of model

Data tab

When you go to the Data tab, you can see all the data as per the defined Relationship from the data sources.

Table: Orders (9,994 rows)

14:08 02-Oct-23

Fig.-Table view of Orders model

Table: People (24 rows)

14:08 02-Oct-23

Fig.-Table view of People model

Untitled - Power BI Desktop

Table tools

Name: Returns

Structure

Calendars Relationships Calculations

Returned Order ID Region

Returned	Order ID	Region
Yes	ES-2015-RA194545-42218	Western Europe
Yes	ES-2012-SC200458-41070	Western Europe
Yes	ES-2015-CC1210045-42182	Western Europe
Yes	ES-2015-MM1792045-42199	Western Europe
Yes	ES-2015-BE1154548-42336	Western Europe
Yes	ES-2014-E61390045-41644	Western Europe
Yes	ES-2013-DB1297045-41458	Western Europe
Yes	IT-2012-MP1817545-41040	Western Europe
Yes	IT-2015-AR1082545-42290	Western Europe
Yes	ES-2015-BN1151545-42304	Western Europe
Yes	ES-2012-AG1049545-41002	Western Europe
Yes	ES-2015-LR1691545-42237	Western Europe
Yes	ES-2012-TT2146045-41014	Western Europe
Yes	ES-2015-RA1945458-42249	Western Europe
Yes	IT-2015-KD1634545-42238	Western Europe
Yes	ES-2014-RM1937545-41643	Western Europe
Yes	ES-2013-BF1117048-41527	Western Europe
Yes	ES-2013-AJ1094548-41444	Western Europe
Yes	ES-2015-BM1165045-42178	Western Europe
Yes	IT-2013-BF1102045-41418	Western Europe
Yes	ES-2012-TB21624545-41262	Western Europe
Yes	ES-2013-CM1244550-41375	Western Europe
Yes	ES-2015-KA1652545-42353	Western Europe

Table: Returns (1,079 rows)

File Home Help Table tools

Search

Sign in

Data

Order ID
Region
Returned

Fig.-Table view of Returns model

Model tab

In the Model tab, you can see the relationship between data sources.

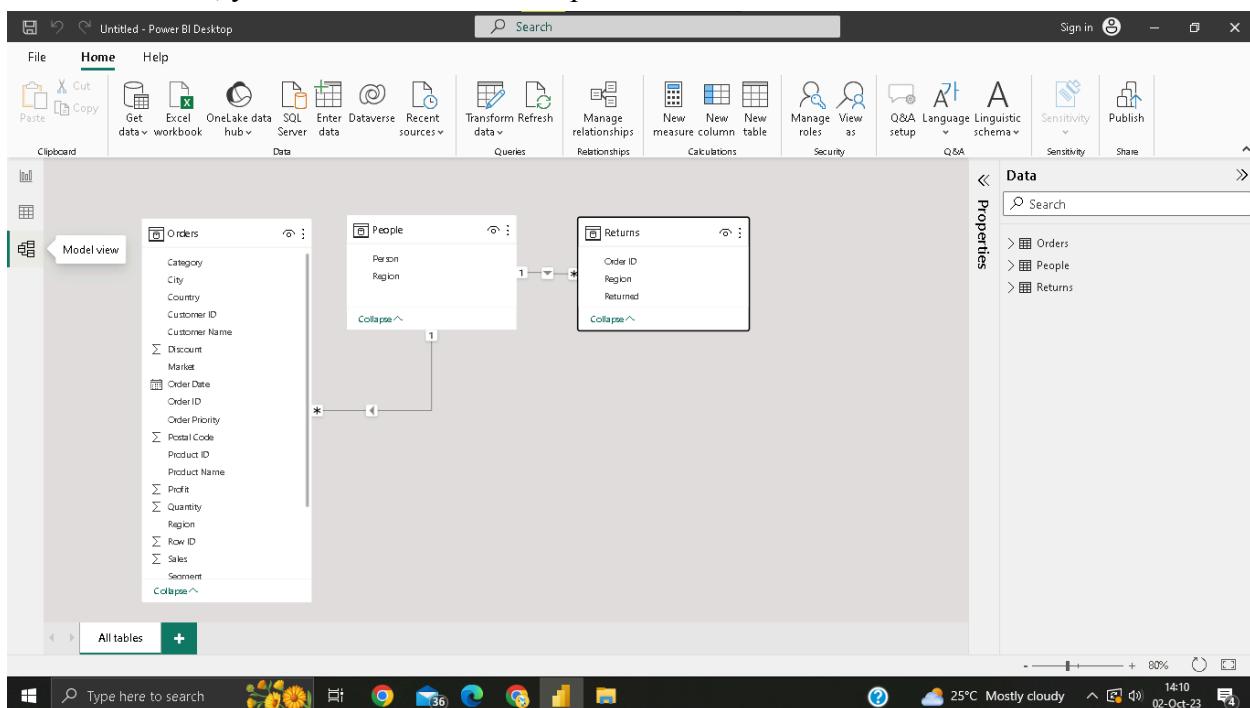


Fig.- Model view of model

Add New Columns

let us create a new column Delivery Days in Orders Table

```
Delivery Days = IF([Ship Date].[Date]-[Order Date].[Date] > 27, 0, [Ship Date].[Date]-[Order Date].[Date])
```

The screenshot shows the Power BI Desktop interface with the 'Table tools' tab selected. A new column 'Delivery Days' is being created with the formula: `IF([Ship Date].[Date]-[Order Date].[Date] > 27, 0, [Ship Date].[Date]-[Order Date].[Date])`. The column properties pane shows the name is 'Delivery Days', data type is 'Decimal number', summarization is 'Sum', and data category is 'Uncategorized'. The table view shows 9,994 rows of order data. The 'Data' pane on the right lists various measures and columns, with 'Delivery Days' highlighted.

Fig.-Create new column Delivery Days in Orders Table

Add New Table

Select "Transform Data." In the Query Editor, select the data you want to use in your new table. Click on the "Add Column" tab and select "Table."

let us create a new Table State_Total_Orders and Total sales

```
State_Total_Orders = SUMMARIZE(Orders, Orders[State], Orders[Region], "Total Quantity Sold", SUM(Orders[Quantity]))
```

```
Total sales = SUMMARIZE(Orders, Orders[Product ID], Orders[ProductName], Orders[Region], "Total Quantities Sold", SUM(Orders[Sales]))
```

Table: State_Total_Orders (49 rows)

Fig.-Create a new table State_Total_Orders

Table: Total sales (5,221 rows)

Fig.-Create a new table Total sales

Creating Calculated Tables

To create a new table, navigate to the Data View tab on the left side of the screen, and then go to the Modeling option at the top of the screen.

Create a new table **Measures (DAX)** from **Enter data**

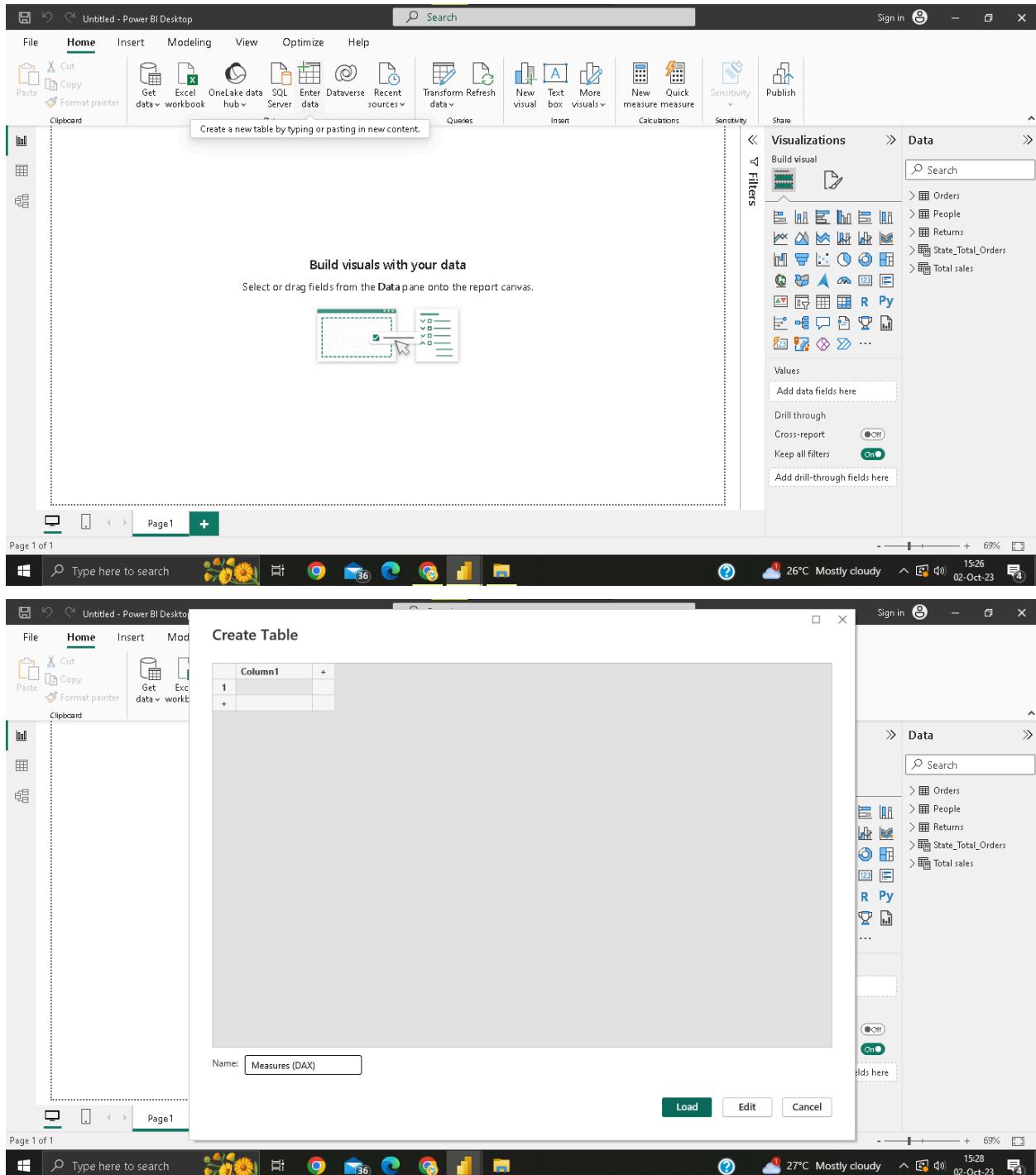


Fig.-Create a new table Measures (DAX)

Create a new measure

1. Bulk Orders

```
Bulk Orders = CALCULATE([Total_Order], 'Orders'[Quantity]>3)
```

2. Total Profit

```
Total Profit = SUM(Orders[Profit])
```

3. Total_Order

```
Total_Order = SUM(Orders[Quantity])
```

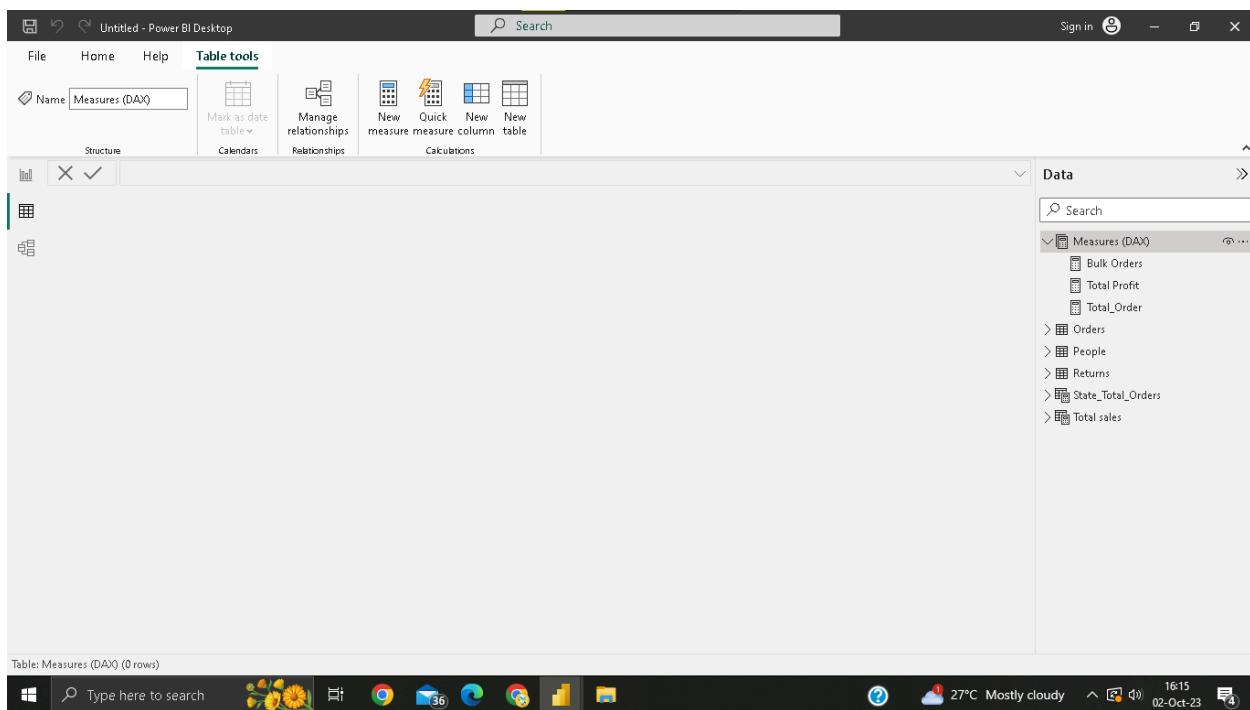


Fig.-Created Measures in Table Measures (DAX)

Add relationships

A relationship can be created in one of two ways firstly, by selecting 'Manage Relationships' from the 'Home'

Secondly, by clicking and dragging the column that you want the relationship to be formed from one table onto the column of another table.

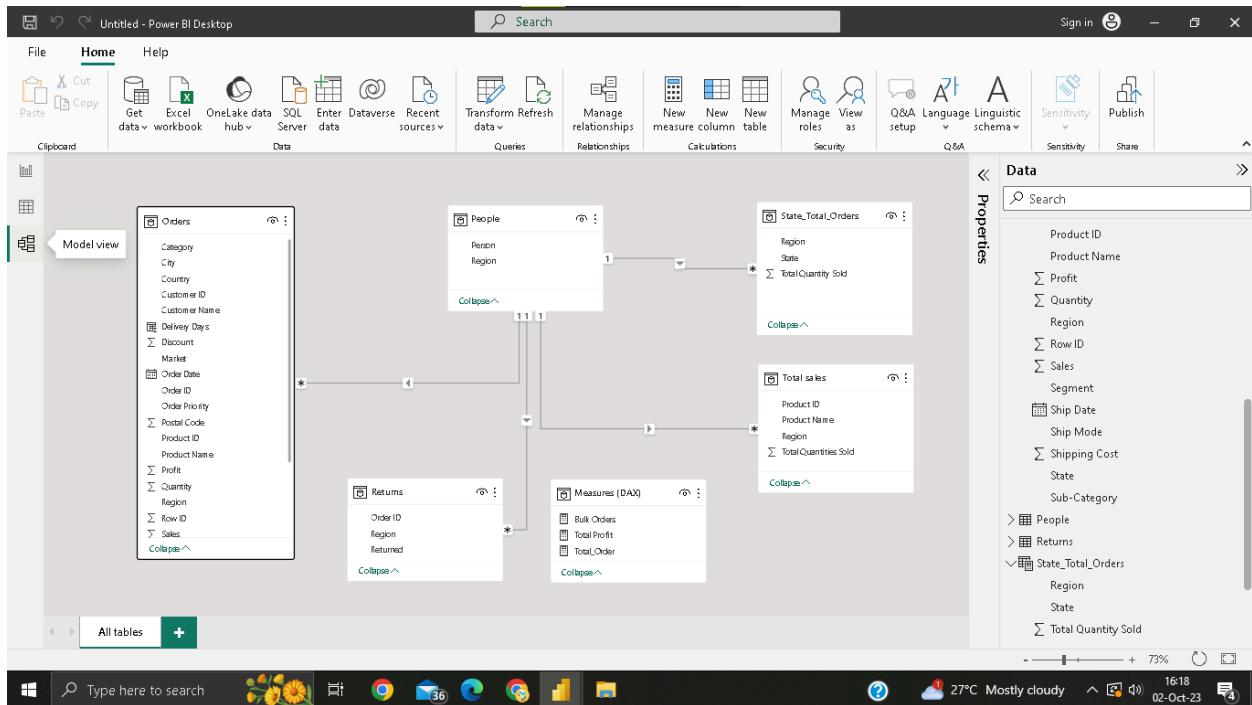


Fig.- Model view of model after Adding relationships

Add Image

In Power BI, you can insert an image directly into a report page, matrix, or slicer visual.

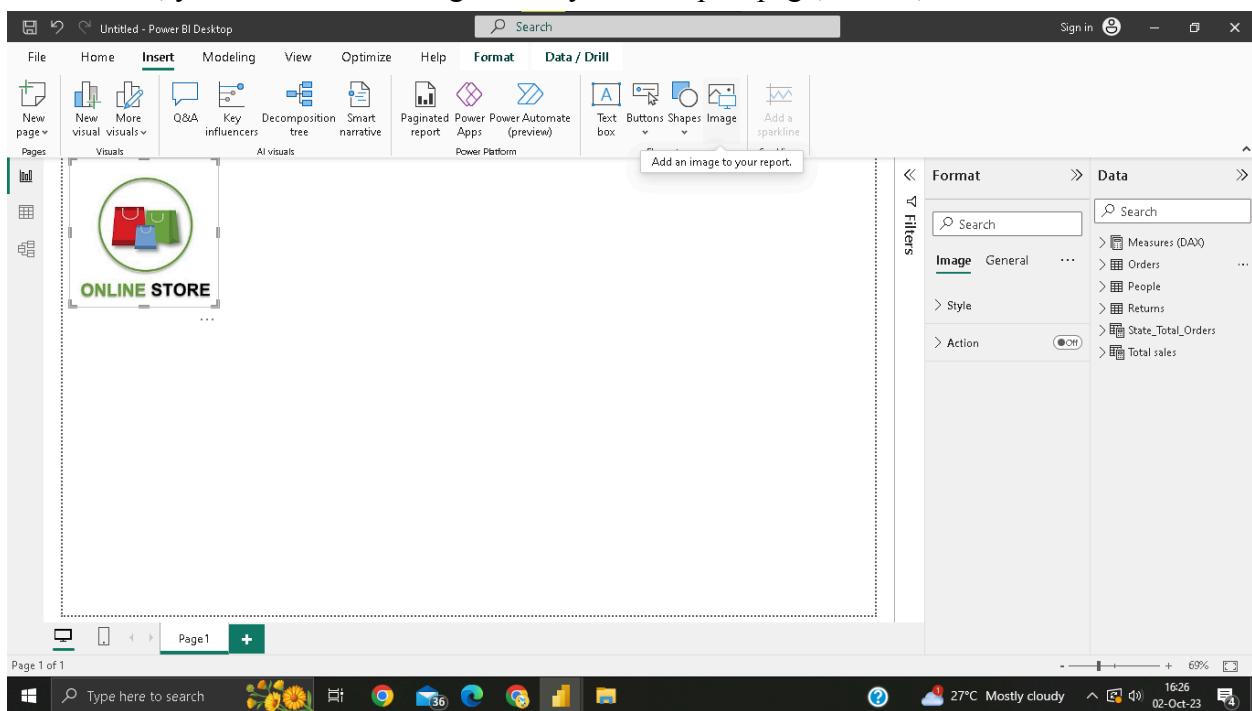


Fig.- Add Image in Report tab

Add Shapes and Cards

We can add text boxes, shapes, and smart narrative visuals to reports in Power BI Desktop. A card is a type of visual element in Power BI that displays a single value or metric.

Create Card for
Total Order,
Sum of Sales,
Total Profit,
Count of Returns

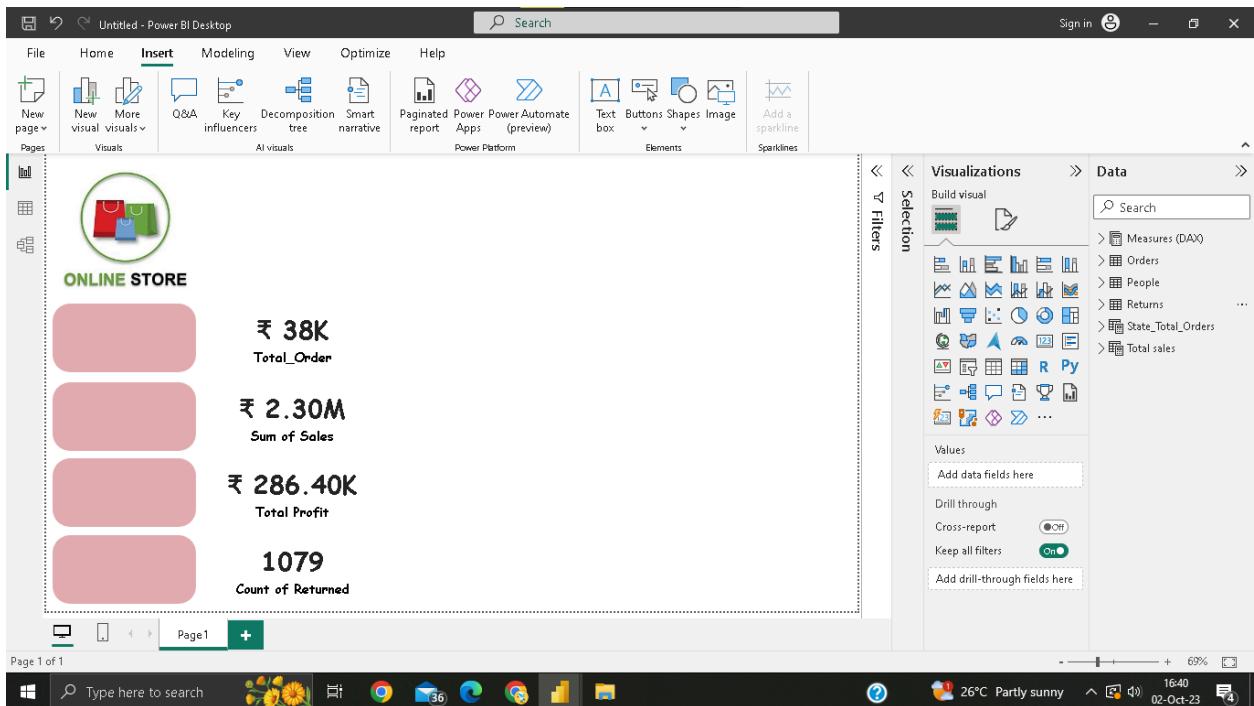


Fig.- Add Shapes And Cards in Report tab

Create Group

To create a group of visuals in Power BI Desktop, select the first visual from the canvas, then holding the CTRL button, click one or more additional visuals that you want in the group. Create Group of all Shapes Rename as Shapes Group and Create Group of All cards and rename as Cards Group

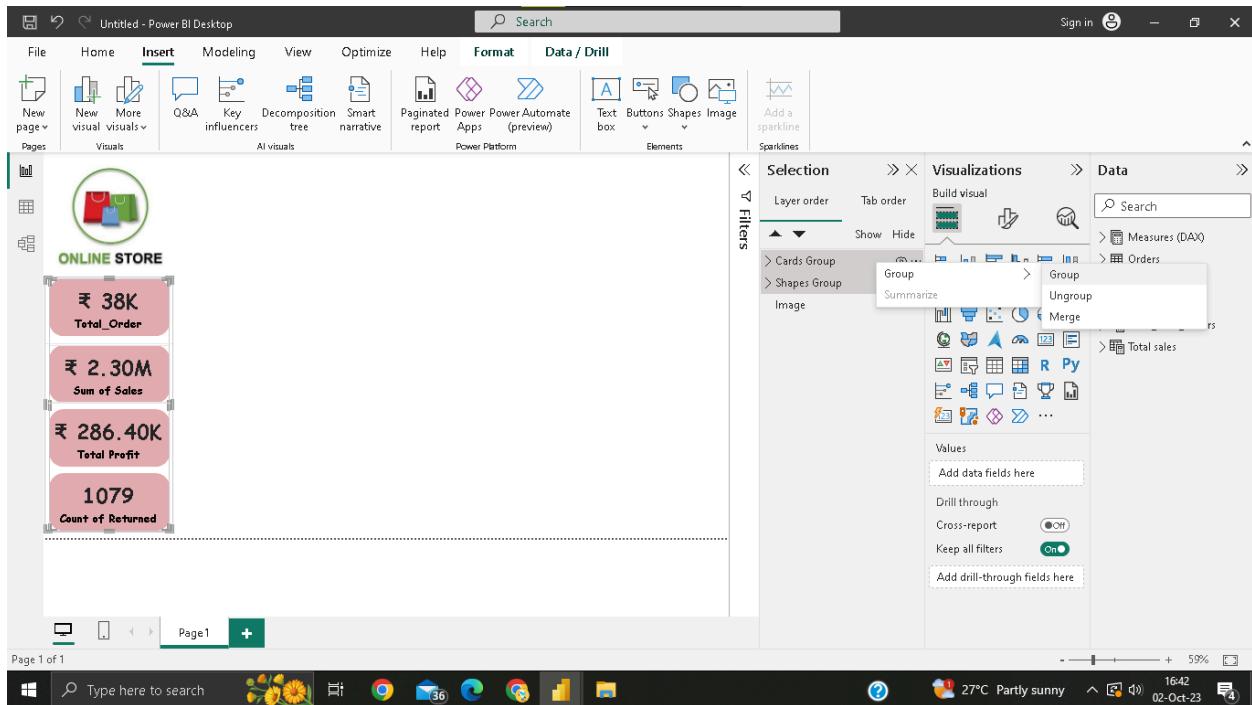


Fig.- Create Group Of Cards And Shape in Report tab

Slicer

Slicers are another way of filtering. They're displayed on the report page, and narrow the portion of the dataset that's shown in the other report visualizations.

Create Slicer With Product Year

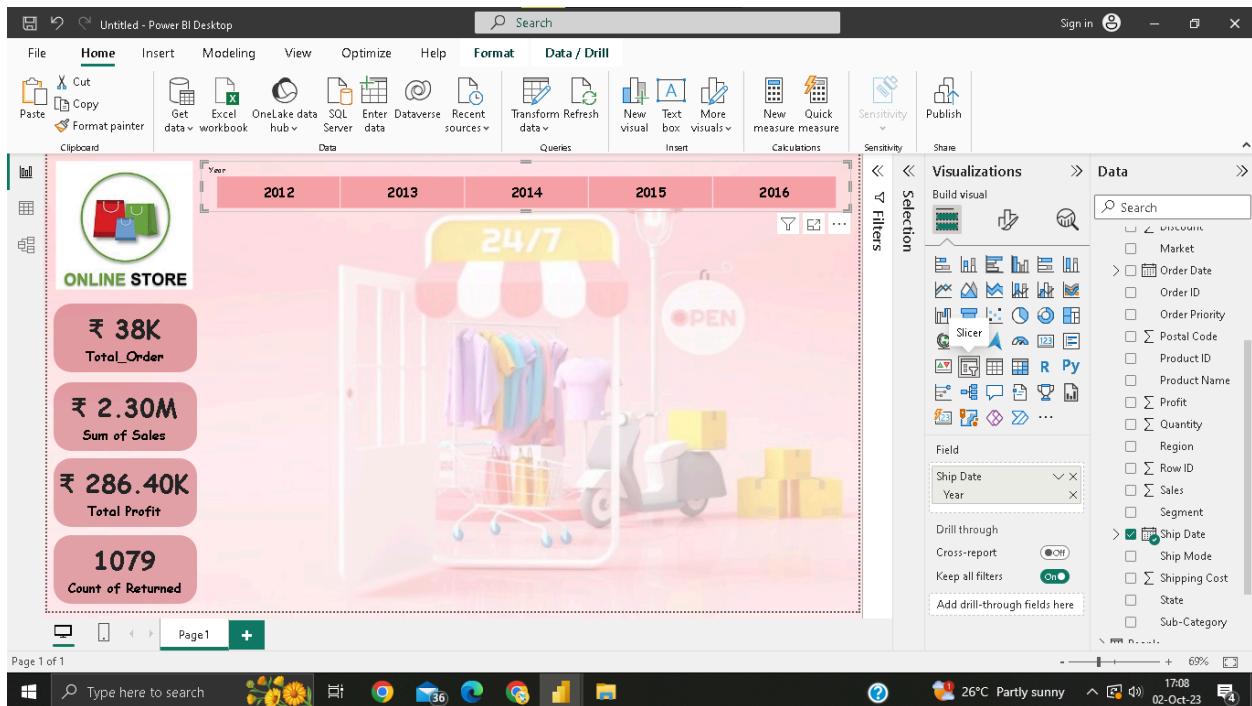


Fig.-Slicer of Year in Report tab

Doughnut Chart

A doughnut chart is similar to a pie chart in that it shows the relationship of parts to a whole.

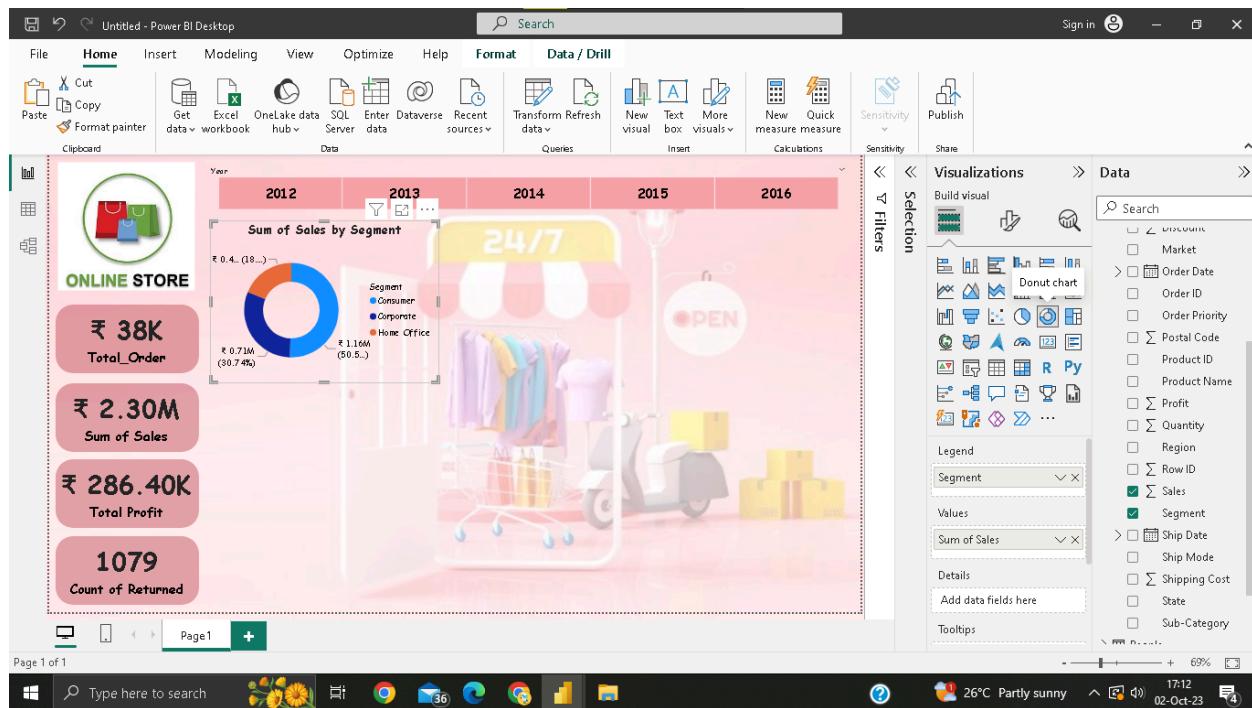


Fig.- Doughnut Chart of Sum of sales by Segment

Pie Chart

The pie chart is a round-shaped circle chart where each category data set is shown in a pie shape.

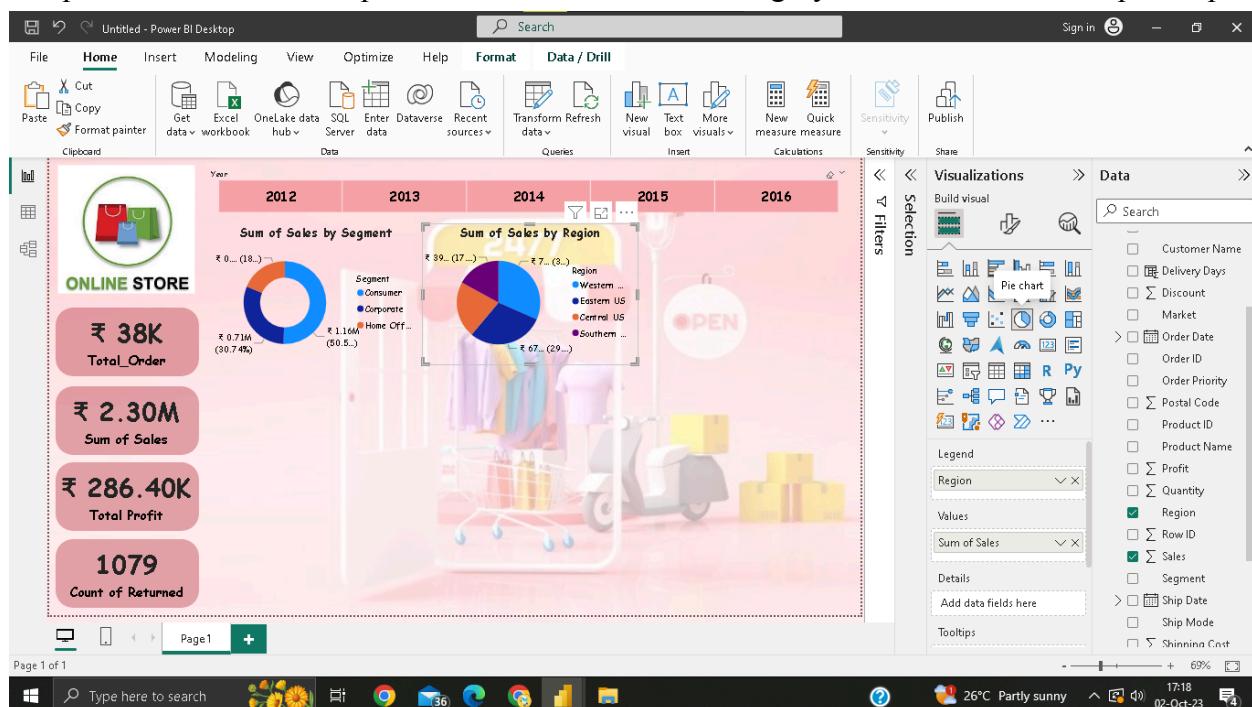


Fig.- Pi Chart of Sum of sales by Region

Clustered Column Chart

The Line and Clustered Column Chart is a combo chart that combines the Line chart and Column chart together in one visual.

Filters

Filters in Power BI sort data and information based on some selected criteria. That is, you can select particular fields or values within fields and view only the information related to that.

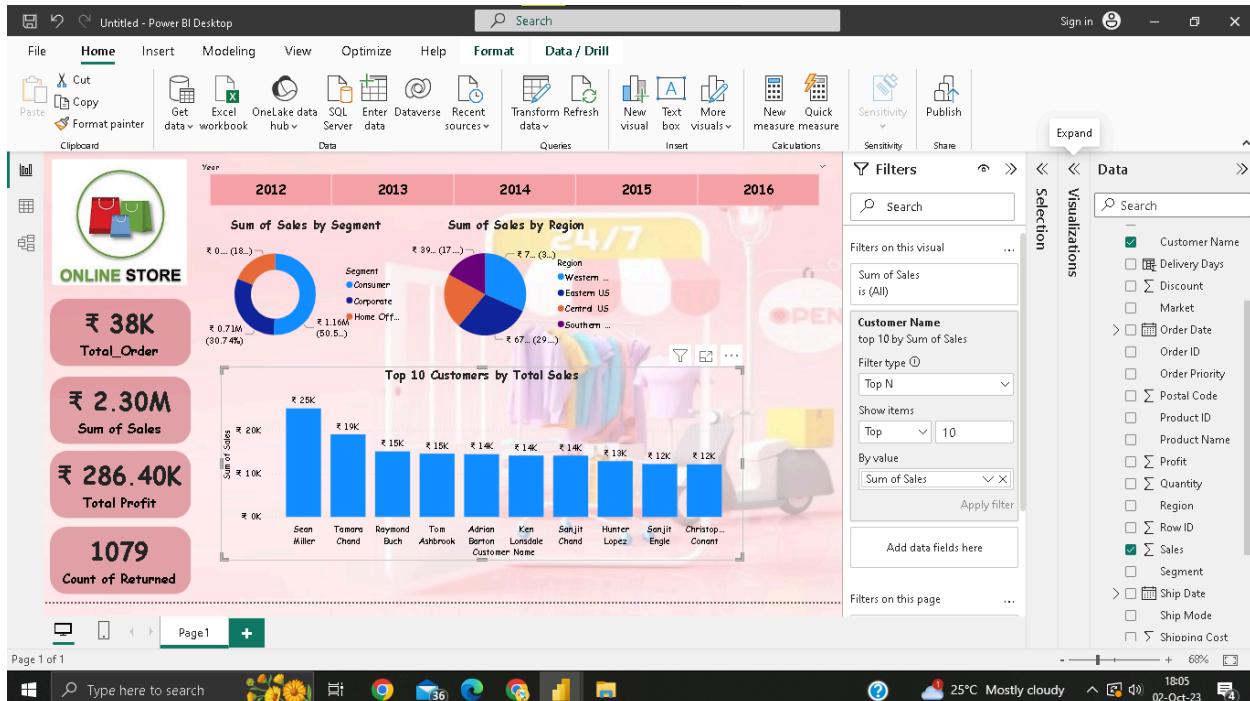


Fig.- Clustered Column Chart of Top 10 Customers by Total Sales

Clustered bar chart

A clustered bar chart is a horizontal chart, which could present multiple bars in the form of a cluster.

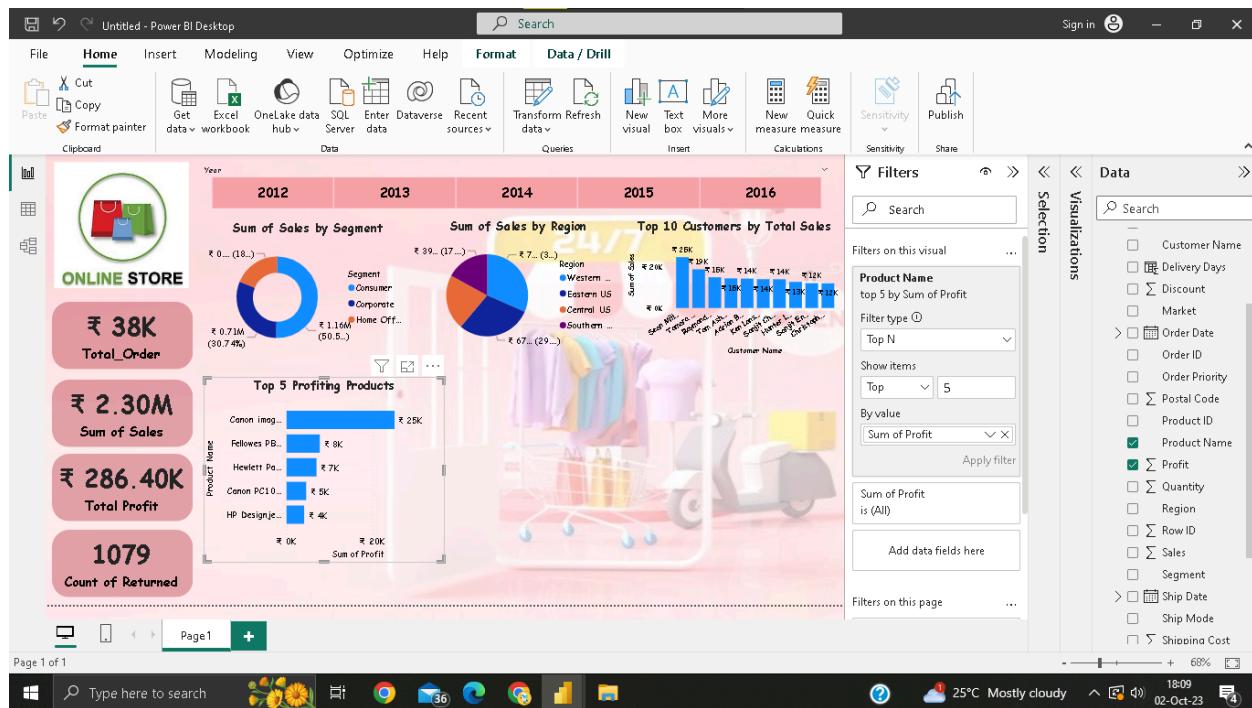


Fig.- Clustered bar chart of Top 5 Profiting Products

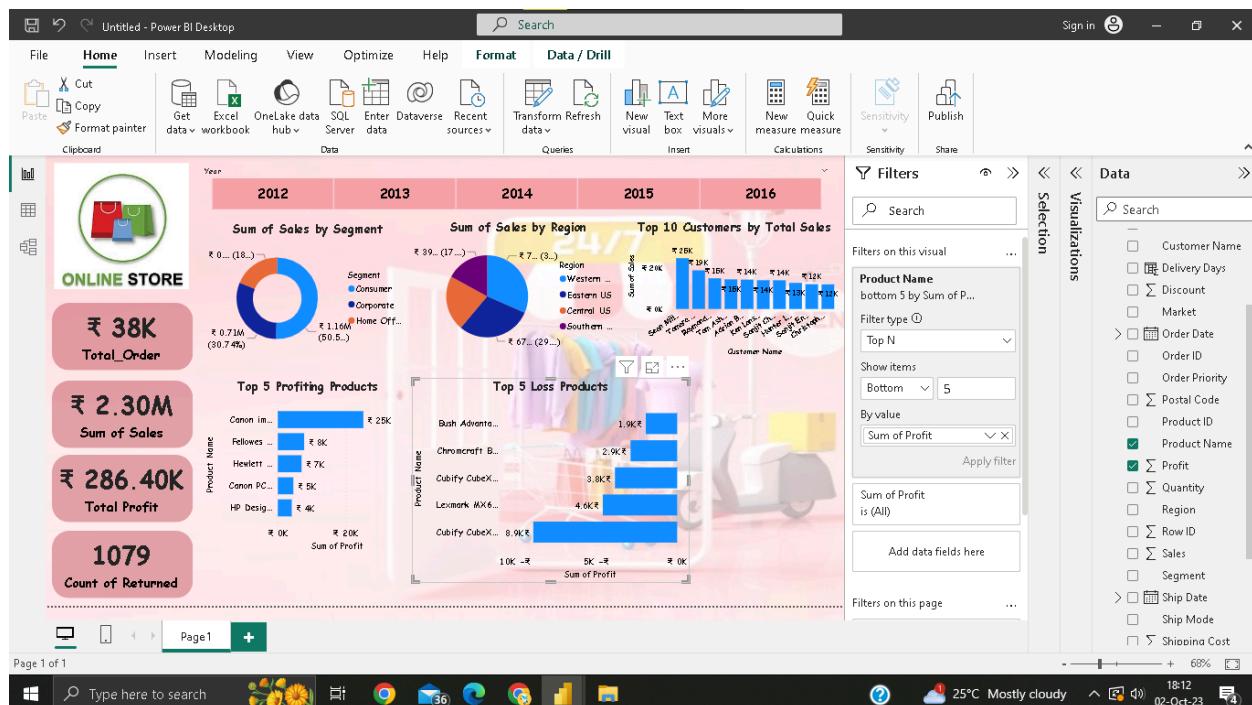


Fig.- Clustered bar chart of Top 5 Loss Products

Map Visualisations

The map is a kind of vision part of Power BI visualization software.

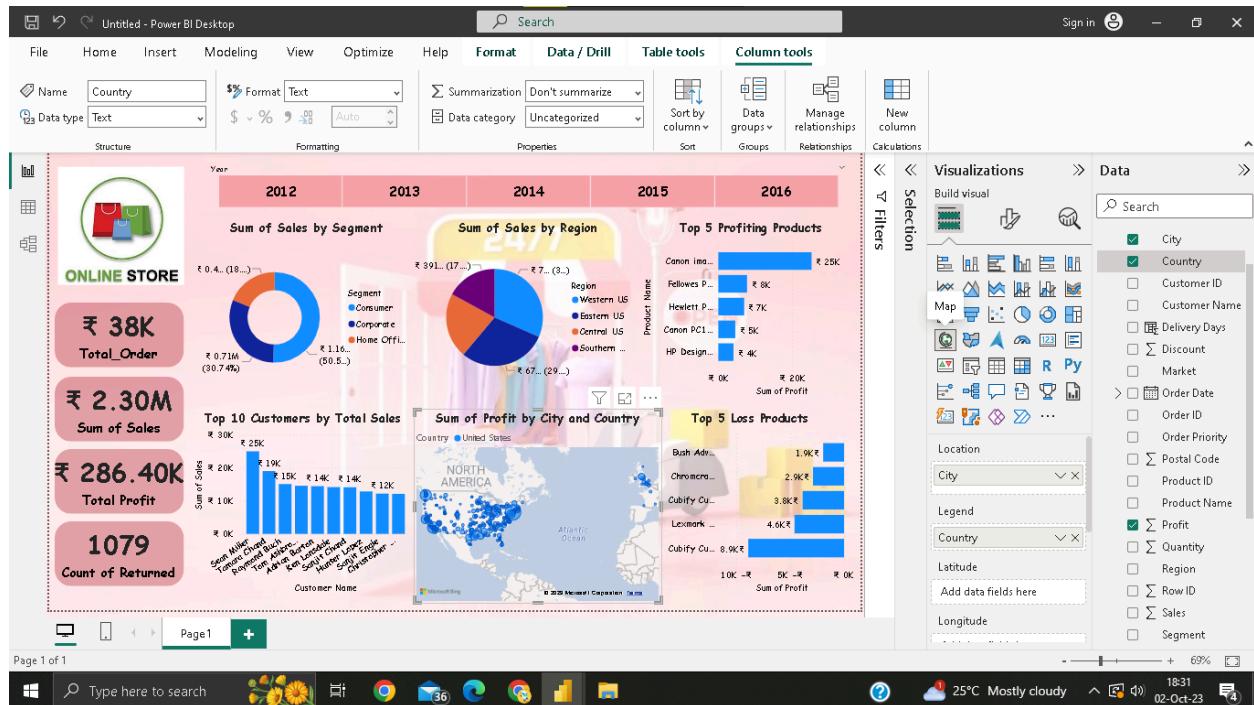


Fig.-Bubble maps (Sum of profit by City and Country)

Report tab

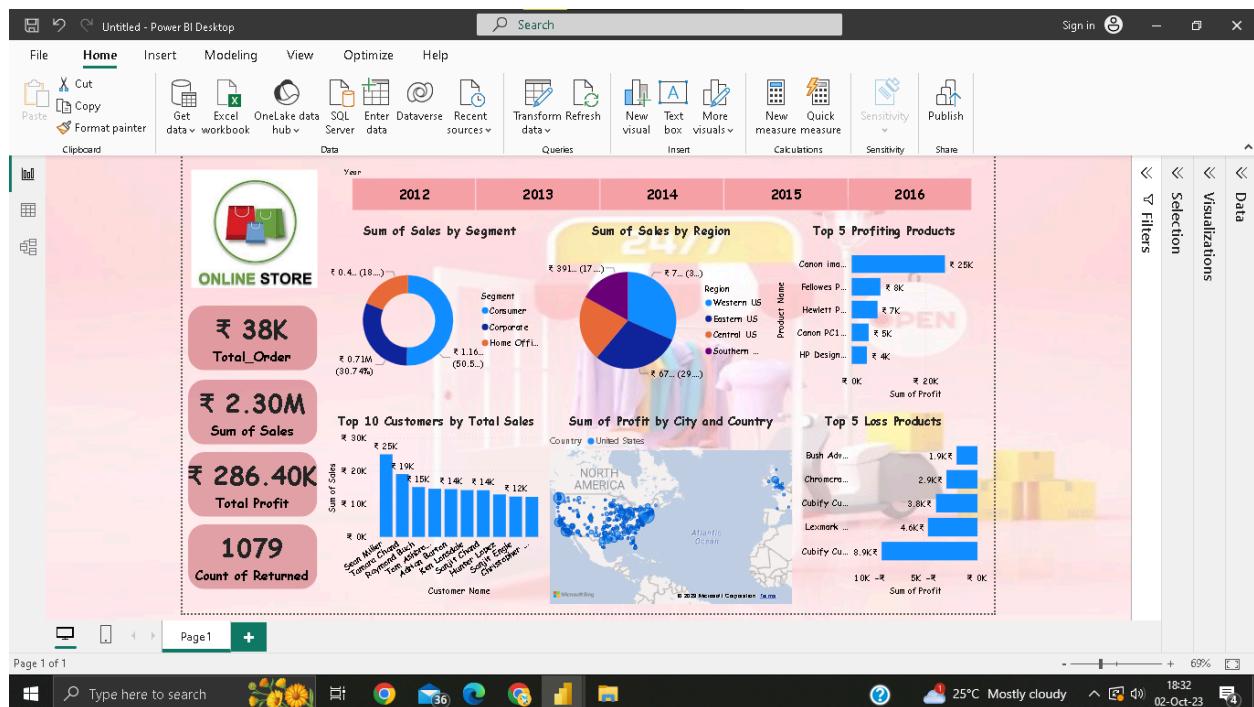
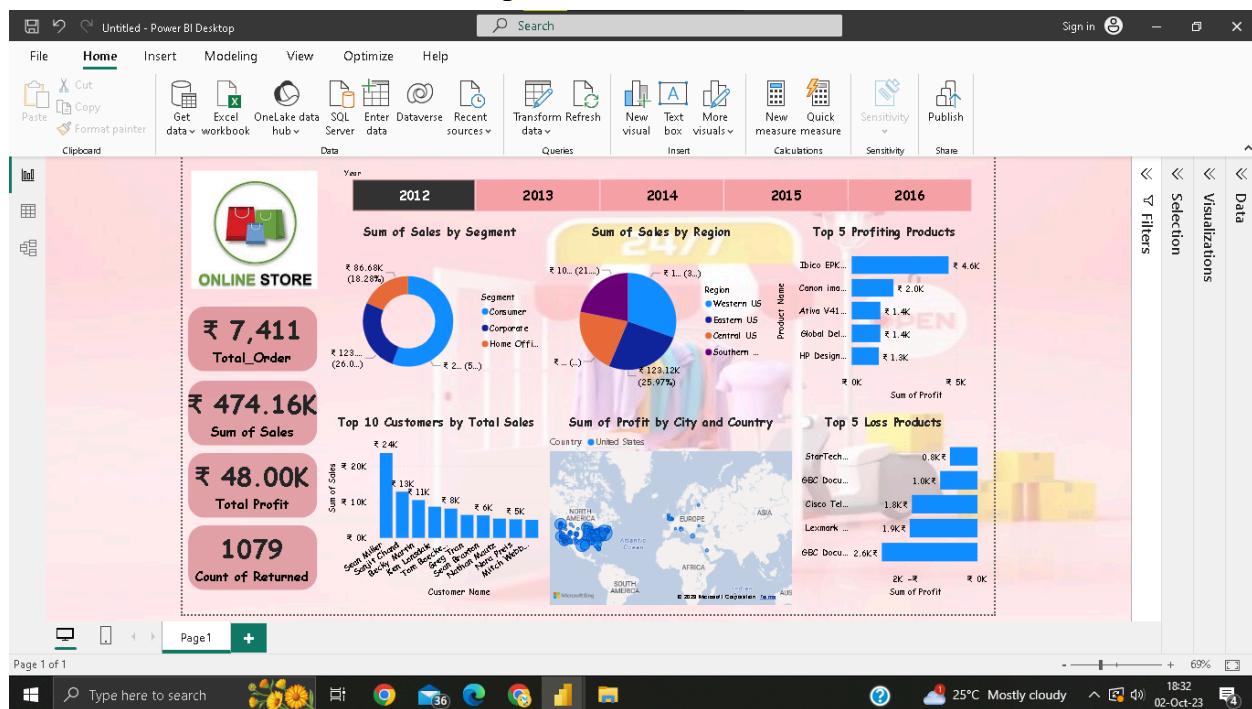
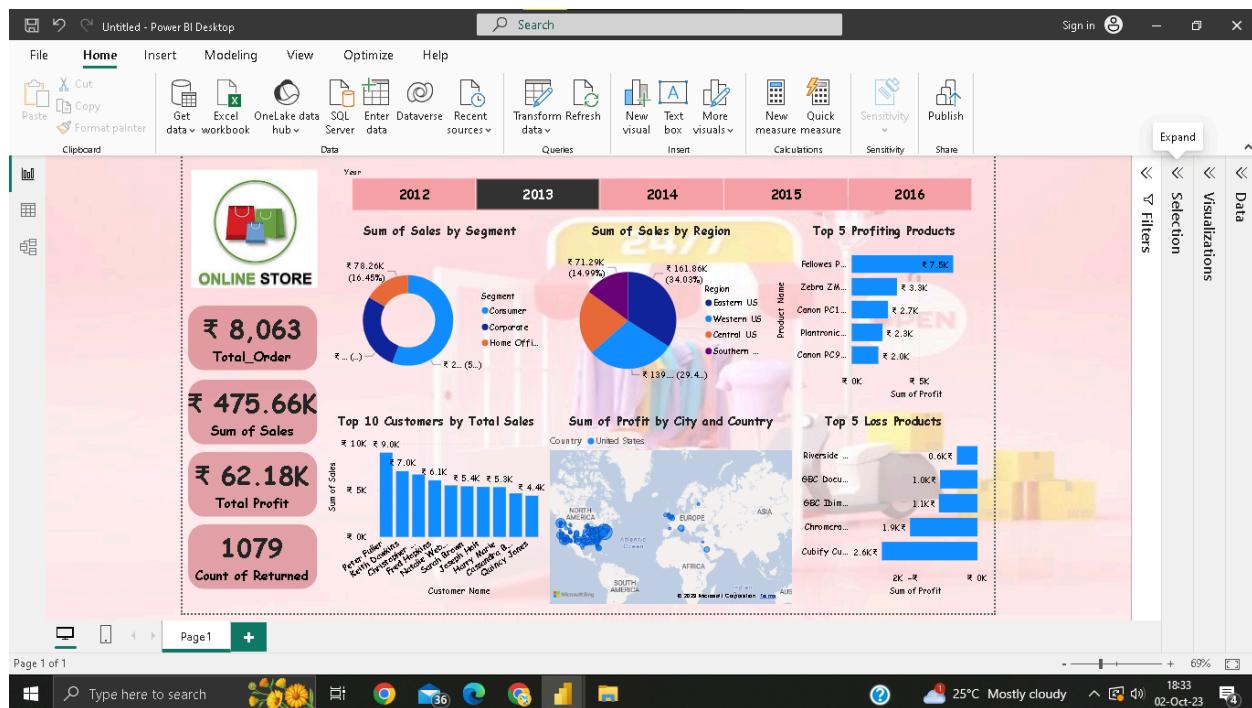


Fig.-Report tab

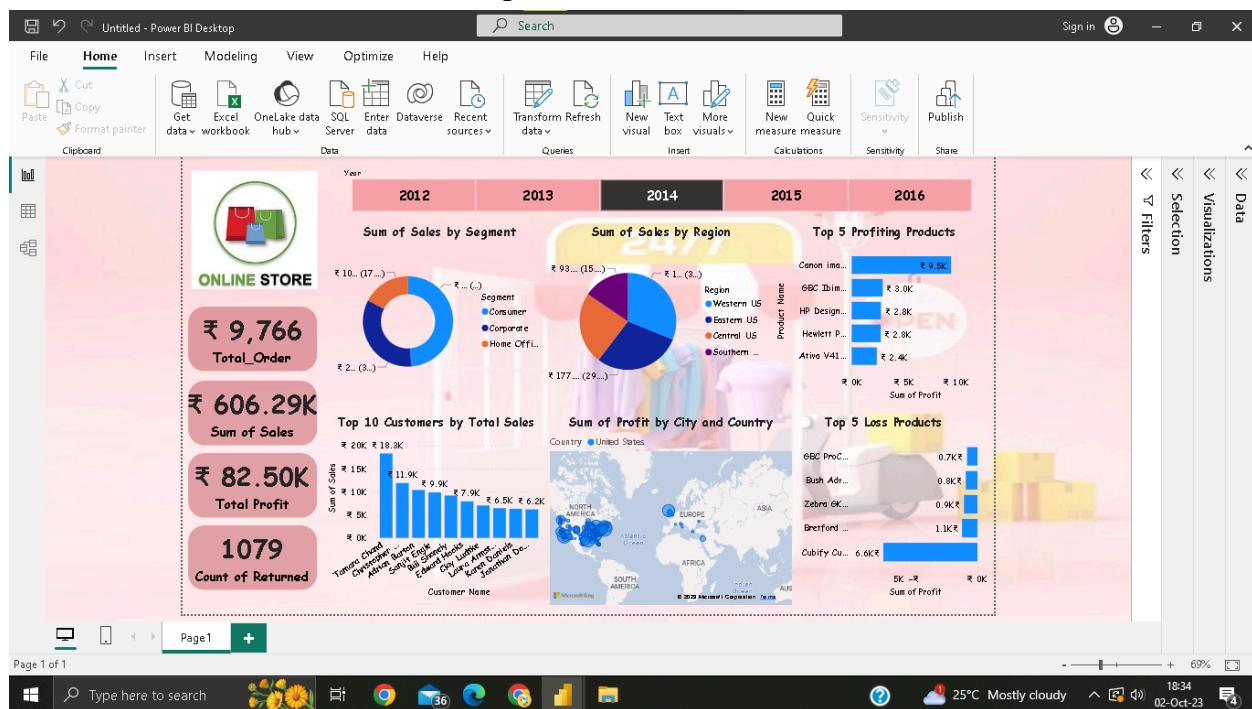
Report tab of Year 2012



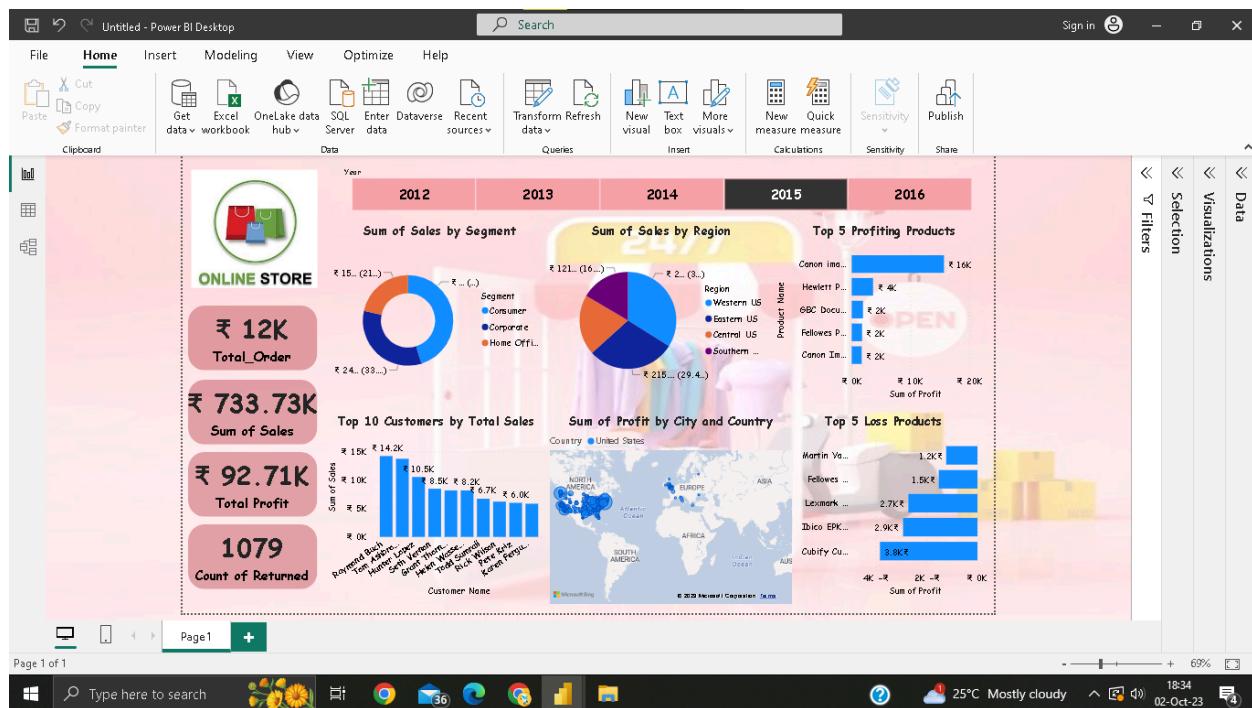
Report tab of Year 2013



Report tab of Year 2014



Report tab of Year 2015



Report tab of Year 2016

Untitled - Power BI Desktop

File Home Insert Modeling View Optimize Help

Cut Copy Format painter Paste Get data from workbook OneLake data hub SQL Server data Enter Dataverse Recent sources Transform Refresh data New visual Text box More visuals Insert Calculations New measure Sensitivity Quick measure Share

Clipboard Data Queries Insert Calculations Sensitivity Share

ONLINE STORE

₹ 175 Total Order

₹ 7.36K Sum of Sales

₹ 1.01K Total Profit

1079 Count of Returned

Sum of Sales by Segment

Segment	Value	Percentage
Consumer	₹ 1.18K	(15.07%)
Corporate	₹ 4.05K	(53.11%)
Home Office	₹ 1.15K	(15.07%)
Total	₹ 7.36K	(100%)

Sum of Sales by Region

Region	Value	Percentage
Western US	₹ 1.89K	(25.53%)
Southern US	₹ 3.97K	(53.99%)
Central US	₹ 0.90K	(12.2%)
Eastern US	₹ 0.90K	(12.2%)
Total	₹ 7.36K	(100%)

Top 5 Profiting Products

Product	Profit
DMI Art...	₹ 314.04
Adjustabl...	₹ 210.49
Global Le...	₹ 105.30
Xerox 19...	₹ 100.66
Global Le...	₹ 87.28

Top 10 Customers by Total Sales

Customer Name	Sum of Sales
Kotteri Han...	₹ 2,000
Alfreds F...	₹ 1,814.68
Jonathan ...	₹ 1,500
Anne ...	₹ 1,003.62
Francisco ...	₹ 725.84
Mark L...	₹ 466.84
Roland ...	₹ 460.84
Patricia ...	₹ 209.70
Donna ...	₹ 100.00
Michael ...	₹ 100.00

Sum of Profit by City and Country

Top 5 Loss Products

Product	Loss
Carina Do...	12.96 - ₹
Novimex ...	34.07 - ₹
Wilson J...	39.58 - ₹
Global Hi...	44.28 - ₹
Wilson J...	65.57 - ₹

Page 1 of 1

Page 1 +

Type here to search

25°C Mostly cloudy 18:34 02-Oct-23