

CSCE 5320 Section(s) 003,600 (Spring 2024 1)

Scientific Data Visualization

Name: Reddybathuni Jyothirmai

Student ID: 11658930

The screenshot shows the Microsoft Power BI desktop application interface. A modal dialog box titled "transactions.csv" is open in the center. The dialog includes fields for "File Origin" (set to "1252: Western European (Windows)"), "Delimiter" (set to "Comma"), and "Data Type Detection" (set to "Based on first 200 rows"). Below these settings is a preview table with three columns: "date", "store_nbr", and "transactions". The table contains 20 rows of data. A note at the bottom of the preview states, "The data in the preview has been truncated due to size limits." At the bottom of the dialog are buttons for "Extract Table Using Examples", "Load", "Transform Data", and "Cancel". To the right of the dialog, the main Power BI workspace shows the "visualizations" pane with various chart icons and the "Data" pane.

date	store_nbr	transactions
1/1/2013	25	770
1/2/2013	1	2111
1/2/2013	2	2358
1/2/2013	3	3487
1/2/2013	4	1922
1/2/2013	5	1903
1/2/2013	6	2143
1/2/2013	7	1874
1/2/2013	8	3250
1/2/2013	9	2940
1/2/2013	10	1293
1/2/2013	11	3547
1/2/2013	12	1362
1/2/2013	13	1102
1/2/2013	14	2002
1/2/2013	15	1622
1/2/2013	16	1167
1/2/2013	17	1580
1/2/2013	18	1635
1/2/2013	19	1369

Fig.1. Adding a dataset

Queries [3]

	i ² ₃ id	date	i ² ₃ store_nbr	A ² ₃ family	i ² ₃ sales
1	0	1/1/2013	1	AUTOMOTIVE	
2	1	1/1/2013	1	BABY CARE	
3	2	1/1/2013	1	BEAUTY	
4	3	1/1/2013	1	BEVERAGES	
5	4	1/1/2013	1	BOOKS	
6	5	1/1/2013	1	BREAD/BAKERY	
7	6	1/1/2013	1	CELEBRATION	
8	7	1/1/2013	1	CLEANING	
9	8	1/1/2013	1	DAIRY	
10	9	1/1/2013	1	DELI	
11	10	1/1/2013	1	EGGS	
12	11	1/1/2013	1	FROZEN FOODS	
13	12	1/1/2013	1	GROCERY I	
14	13	1/1/2013	1	GROCERY II	
15	14	1/1/2013	1	HARDWARE	
16	15	1/1/2013	1	HOME AND KITCHEN I	
17	16	1/1/2013	1	HOME AND KITCHEN II	
18	17	1/1/2013	1	HOME APPLIANCES	
19	18	1/1/2013	1	HOME CARE	
20	19	1/1/2013	1	LADIESWEAR	
21					

6 COLUMNS, 999+ ROWS Column profiling based on top 1000 rows PREVIEW DOWNLOADED AT 10:23 AM 10:23 AM 3/1/2024

Fig.2. Transforming the data

Queries [4]

	date	i ² ₃ permitDate	i ² ₃ transactions
1	1/1/2013	25	770
2	1/2/2013	1	2111
3	1/2/2013	2	2358
4	1/2/2013	3	3487
5	1/2/2013	4	1922
6	1/2/2013	5	1903
7	1/2/2013	6	2143
8	1/2/2013	7	1874
9	1/2/2013	8	3250
10	1/2/2013	9	2940
11	1/2/2013	10	1293
12	1/2/2013	11	3547
13	1/2/2013	12	1362
14	1/2/2013	13	1102
15	1/2/2013	14	2002
16	1/2/2013	15	1622
17	1/2/2013	16	1167
18	1/2/2013	17	1580
19	1/2/2013	18	1635
20	1/2/2013	19	1369
21	1/2/2013	23	1381

3 COLUMNS, 999+ ROWS Column profiling based on top 1000 rows PREVIEW DOWNLOADED AT 10:31 AM 10:31 AM 3/1/2024

Fig.3. Renaming the column name

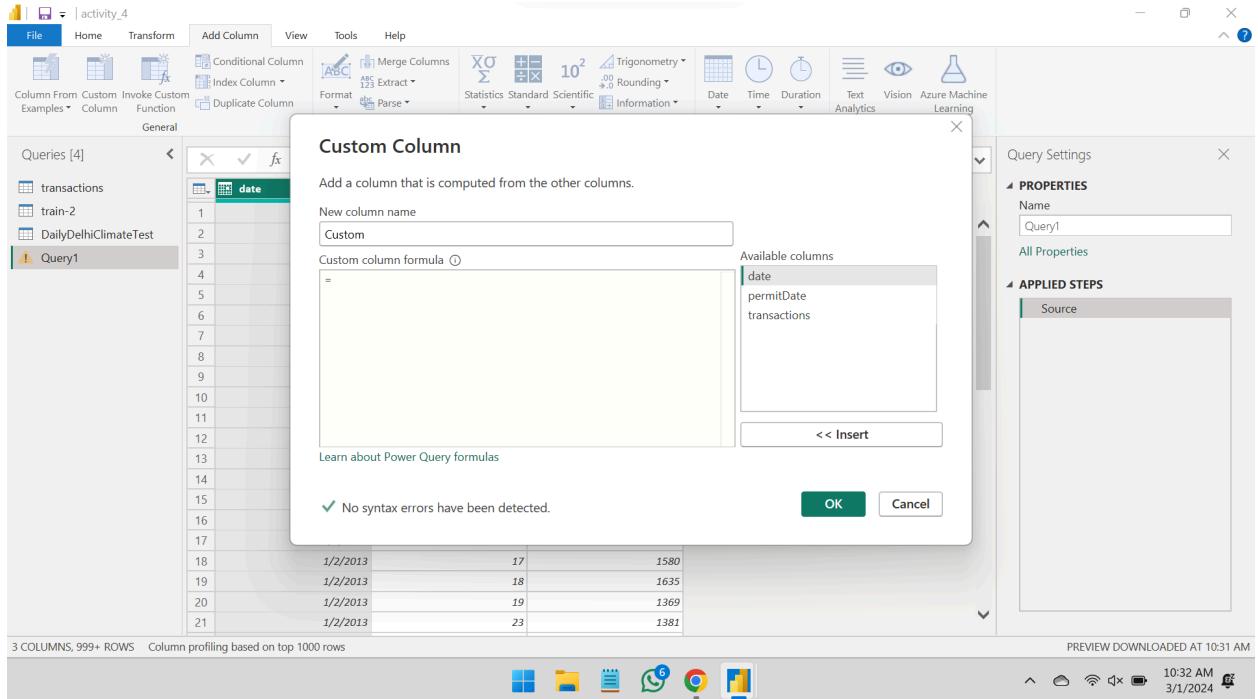


Fig.4. Adding a custom column

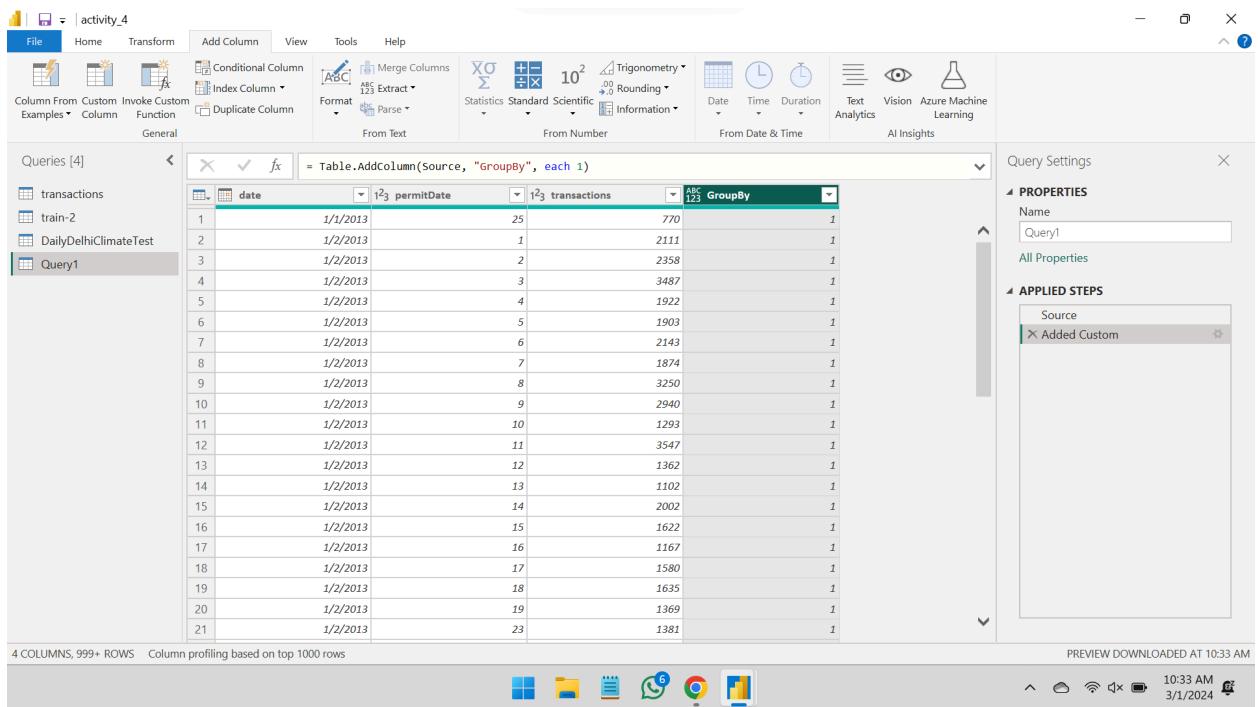


Fig.5. Performing Group By

Group By

Specify the columns to group by and one or more outputs.

Basic Advanced

GroupBy

Add grouping

New column name Operation Column

MinDate	Min	permitDate
MaxDate	Max	permitDate

Add aggregation

OK Cancel

4 COLUMNS, 999+ ROWS Column profiling based on top 1000 rows PREVIEW DOWNLOADED AT 10:33 AM

Fig.6. Attributes to which Group By is applied

Query Settings

PROPERTIES

Name: Query1

All Properties

APPLIED STEPS

- Source
- Added Custom
- Grouped Rows

3 COLUMNS, 1 ROW Column profiling based on top 1000 rows PREVIEW DOWNLOADED AT 10:37 AM

Fig.7. Group By result

activity_4

File Home Transform Add Column View Tools Help

New Recent Enter Data Data source settings Manage Parameters Refresh Preview Advanced Editor Properties Query Choose Columns Remove Columns Keep Rows Remove Rows Sort Split Column Group By Data Type: Whole Number Use First Row as Headers Merge Queries Append Queries Text Analytics Vision Combine Files Azure Machine Learning Combine AI Insights

Queries

= Table.SelectRows(#"Changed Type", each ([sales]>1) and ([onpromotion]>1))

	Id	date	store_nbr	family	sales	onpromotion
1	810784	4/1/2014	9	CLEANING	1752	
2	810789	4/1/2014	9	GROCERY I	7685	
3	812571	4/2/2014	9	GROCERY I	6481	
4	814350	4/3/2014	9	DELI	603	
5	814351	4/3/2014	9	EGGS	136	
6	814365	4/3/2014	9	MEATS	404	
7	814369	4/3/2014	9	POULTRY	685	
8	814373	4/3/2014	9	SEAFOOD	23	
9	816130	4/4/2014	9	CLEANING	1300	
10	817908	4/5/2014	9	BEVERAGES	2217	
11	817933	4/5/2014	9	POULTRY	841	
12	819694	4/6/2014	9	CLEANING	2503	
13	819699	4/6/2014	9	GROCERY I	9917	
14	821481	4/7/2014	9	GROCERY I	6046	
15	823256	4/8/2014	9	BREAD/BAKERY	565	
16	823258	4/8/2014	9	CLEANING	1086	
17	823263	4/8/2014	9	GROCERY I	4387	

28.6 MB FROM TRAIN-2.CSV
10:51 AM 3/1/2024

Fig.8. Executing a query

Question 1

activity_4

File Home Transform Add Column View Tools Help

New Recent Enter Data Data source settings Manage Parameters Refresh Preview Advanced Editor Properties Query Choose Columns Remove Columns Keep Rows Remove Rows Sort Split Column Group By Data Type: Whole Number Use First Row as Headers Merge Queries Append Queries Text Analytics Vision Combine Files Azure Machine Learning Combine AI Insights

Queries [4]

transactions

= Table.RenameColumns(#"Changed Type", {"store_nbr", "store_id"}, {"date", "permitDate"})

	permitDate	store_id	transactions
1	1/1/2013	25	770
2	1/2/2013	1	2111
3	1/2/2013	2	2358
4	1/2/2013	3	3487
5	1/2/2013	4	1922
6	1/2/2013	5	1903
7	1/2/2013	6	2143
8	1/2/2013	7	1874
9	1/2/2013	8	3250
10	1/2/2013	9	2940
11	1/2/2013	10	1293
12	1/2/2013	11	3547
13	1/2/2013	12	1362
14	1/2/2013	13	1102
15	1/2/2013	14	2002
16	1/2/2013	15	1622
17	1/2/2013	16	1167
18	1/2/2013	17	1580
19	1/2/2013	18	1635
20	1/2/2013	19	1369
21	1/2/2013	23	1381

3 COLUMNS, 999+ ROWS Column profiling based on top 1000 rows PREVIEW DOWNLOADED AT 10:53 AM 3/1/2024

Fig.9. Renaming column name

Queries [6]

	i ² ₃ id	date	i ² ₃ store_id	A ² C family	i ² ₃ onpromotion	ABC GroupBy in test
1	3000888	8/16/2017	1	AUTOMOTIVE	0	1
2	3000889	8/16/2017	1	BABY CARE	0	1
3	3000890	8/16/2017	1	BEAUTY	2	1
4	3000891	8/16/2017	1	BEVERAGES	20	1
5	3000892	8/16/2017	1	BOOKS	0	1
6	3000893	8/16/2017	1	BREAD/BAKERY	12	1
7	3000894	8/16/2017	1	CELEBRATION	0	1
8	3000895	8/16/2017	1	CLEANING	25	1
9	3000896	8/16/2017	1	DAIRY	45	1
10	3000897	8/16/2017	1	DELI	18	1
11	3000898	8/16/2017	1	EGGS	1	1
12	3000899	8/16/2017	1	FROZEN FOODS	1	1
13	3000900	8/16/2017	1	GROCERY I	64	1
14	3000901	8/16/2017	1	GROCERY II	0	1
15	3000902	8/16/2017	1	HARDWARE	0	1
16	3000903	8/16/2017	1	HOME AND KITCHEN I	2	1
17	3000904	8/16/2017	1	HOME AND KITCHEN II	6	1
18	3000905	8/16/2017	1	HOME APPLIANCES	0	1
19	3000906	8/16/2017	1	HOME CARE	10	1
20	3000907	8/16/2017	1	LADIESWEAR	0	1
21	3000908	8/16/2017	1	LAWN AND GARDEN	16	1

Fig.10. Applying Group By function

Group By

Specify the columns to group by and one or more outputs.

Basic Advanced

GroupBy in test

Add grouping

New column name	Operation	Column
MinDate	Min	date
MaxDate	Max	date

Add aggregation

OK Cancel

Fig.11. Applying Group By function

The screenshot shows the Power BI desktop interface with the following details:

- File** tab selected.
- Transform** ribbon tab selected.
- Queries [9]** pane on the left shows:
 - Query Errors - 3/1/20...
 - Query Errors - 3/6/20...
 - Other Queries [7]:
 - transactions
 - DailyDelhiClimateTest
 - Query1
 - train-2
 - holidays_events
 - test
 - Query2
- Properties** pane on the right shows:
 - Name: test
 - All Properties
 - APPLIED STEPS:
 - Source
 - Promoted Headers
 - Changed Type
 - Renamed Columns
 - Added Custom
 - Grouped Rows
- Transform** ribbon tab selected.
- Table** view showing the result of the query:

	MinDate	MaxDate
1	8/16/2017	8/31/2017
- System tray at the bottom right shows: 10:40 AM, 3/7/2024.

Fig.12. Result of Group By function

The screenshot shows the Power BI desktop interface with the following details:

- File** tab selected.
- Transform** ribbon tab selected.
- Queries [7]** pane on the left shows:
 - transactions
 - DailyDelhiClimateTest
 - Query1
 - train-2
 - holidays_events
 - test
 - Query2
- Transform** ribbon tab selected.
- Table** view showing the result of the query:

	id	date	store_id	family	onpromotion	GroupBy in test
1	3000888	8/16/2017	1	AUTOMOTIVE	0	1
2	3000889	8/16/2017	1	BABY CARE	0	1
3	3000890	8/16/2017	1	BEAUTY	2	1
4	3000891	8/16/2017	1	BEVERAGES	20	1
5	3000892	8/16/2017	1	BOOKS	0	1
6	3000893	8/16/2017	1	BREAD/BAKERY	12	1
7	3000894	8/16/2017	1	CELEBRATION	0	1
8	3000895	8/16/2017	1	CLEANING	25	1
9	3000896	8/16/2017	1	DAIRY	45	1
10	3000897	8/16/2017	1	DELI	18	1
11	3000898	8/16/2017	1	EGGS	1	1
12	3000899	8/16/2017	1	FROZEN FOODS	1	1
13	3000900	8/16/2017	1	GROCERY I	64	1
14	3000901	8/16/2017	1	GROCERY II	0	1
15	3000902	8/16/2017	1	HARDWARE	0	1
16	3000903	8/16/2017	1	HOME AND KITCHEN I	2	1
17	3000904	8/16/2017	1	HOME AND KITCHEN II	6	1
18	3000905	8/16/2017	1	HOME APPLIANCES	0	1
19	3000906	8/16/2017	1	HOME CARE	10	1
20	3000907	8/16/2017	1	LADIESWEAR	0	1
21	3000908	8/16/2017	1	LAWN AND GARDEN	16	1
- Bottom status bar: 6 COLUMNS, 999+ ROWS, Column profiling based on top 1000 rows, PREVIEW DOWNLOADED AT 11:03 AM, 11:03 AM, 3/1/2024.

Fig.13. Executing a Query

The screenshot shows the Power BI desktop interface with a filter dialog open. The dialog is titled "Sort Ascending" and lists a range of dates from 7/28/2017 to 8/6/2017. A message at the bottom left says "List may be incomplete." and "Load more". On the right, the "APPLIED STEPS" pane shows the step "Sorted Rows" has been applied. The main query editor shows a table with columns "permitDate", "store_id", and "transactions".

permitDate	store_id	transactions
8/15/2017	24	2138
8/15/2017	3	2956
8/15/2017	33	919
8/15/2017	10	1010
8/15/2017	47	3581
8/15/2017	52	2255
8/15/2017	6	1589
8/15/2017	54	802
8/15/2017	9	2155
8/15/2017	50	2804
8/15/2017	19	1158
8/15/2017	1	1693
8/15/2017	29	1302
8/15/2017	4	1283
8/15/2017	17	1504
8/15/2017	46	3197
8/15/2017	31	1360
8/15/2017	38	1445
8/15/2017	16	742
8/15/2017	30	825
8/15/2017	40	1392

Fig.14. Filtering the values in a column

The screenshot shows the Power BI desktop interface with the filter applied. The "APPLIED STEPS" pane now shows the step "Filtered Rows". The main query editor shows the same table as before, but only rows where the date is 8/15/2017 are visible. The status bar at the bottom indicates "3 COLUMNS, 999+ ROWS" and "Column profiling based on top 1000 rows".

permitDate	store_id	transactions
8/15/2017	24	2138
8/15/2017	3	2956
8/15/2017	33	919
8/15/2017	10	1010
8/15/2017	47	3581
8/15/2017	52	2255
8/15/2017	6	1589
8/15/2017	54	802
8/15/2017	9	2155
8/15/2017	50	2804
8/15/2017	19	1158
8/15/2017	1	1693
8/15/2017	29	1302
8/15/2017	4	1283
8/15/2017	17	1504
8/15/2017	46	3197
8/15/2017	31	1360
8/15/2017	38	1445
8/15/2017	16	742
8/15/2017	30	825
8/15/2017	40	1392

Fig.15. Result of filter

Tutorial- 2

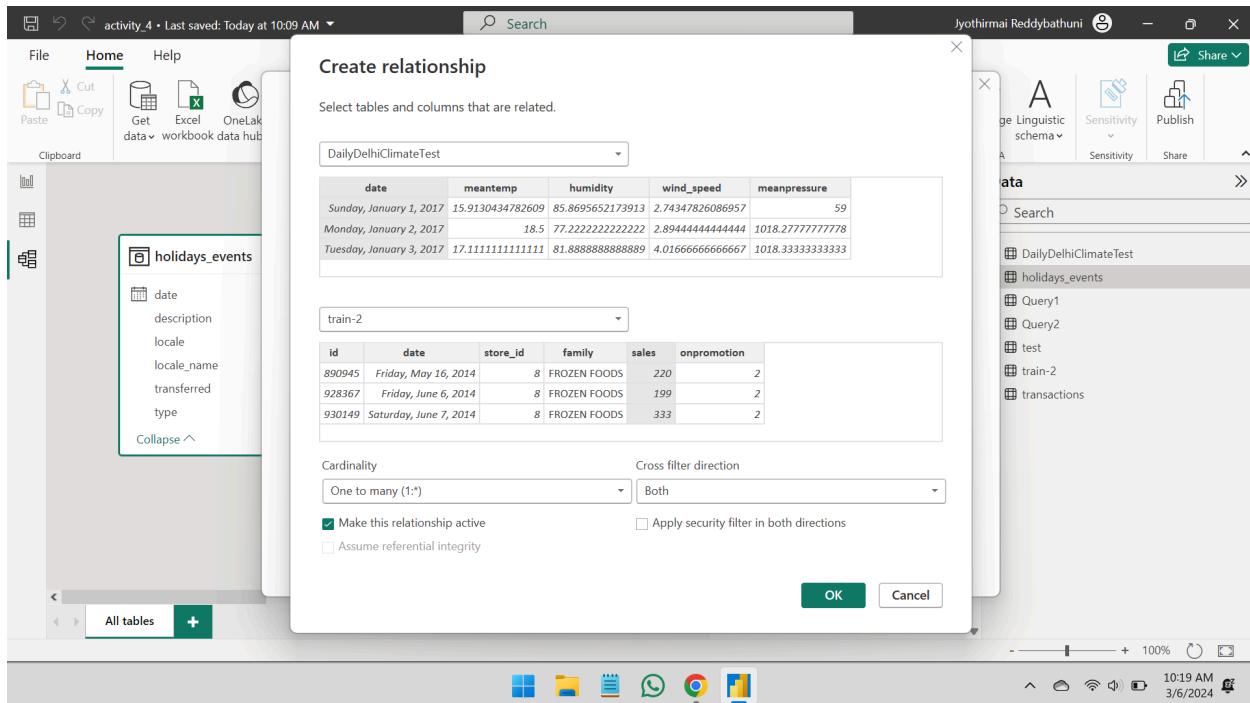


Fig.16. Creating a relationship

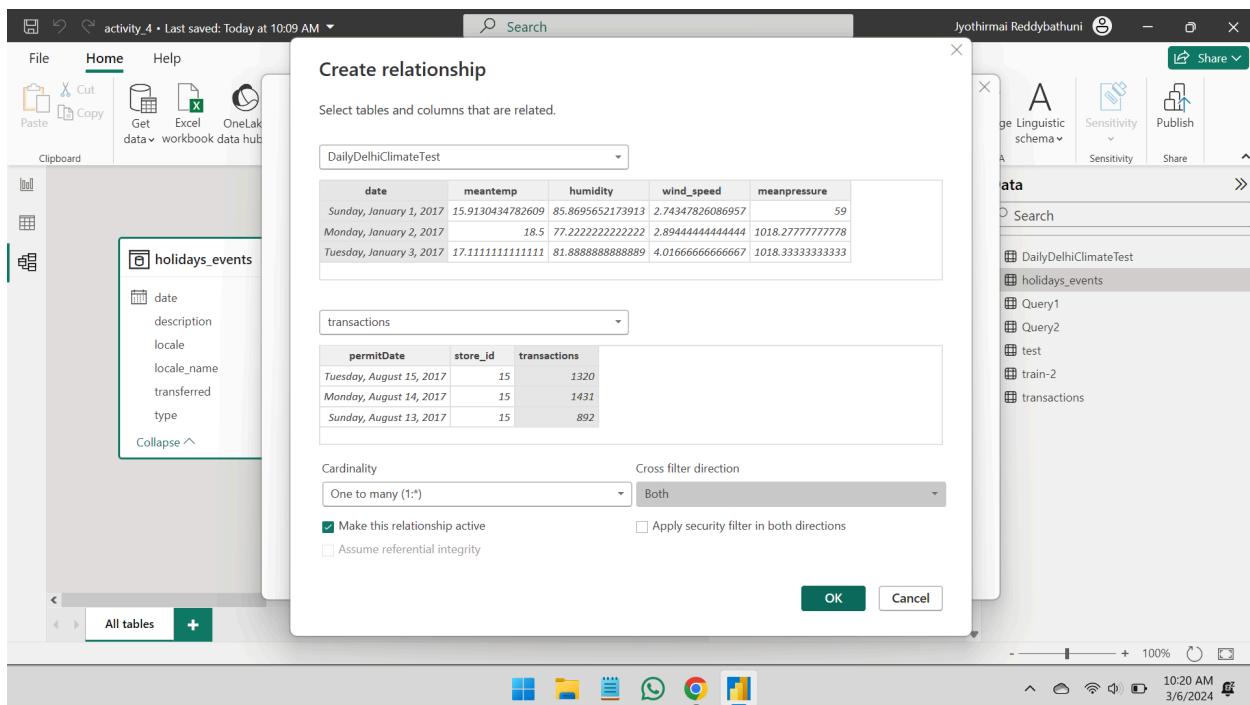


Fig.17.Creating a relationship

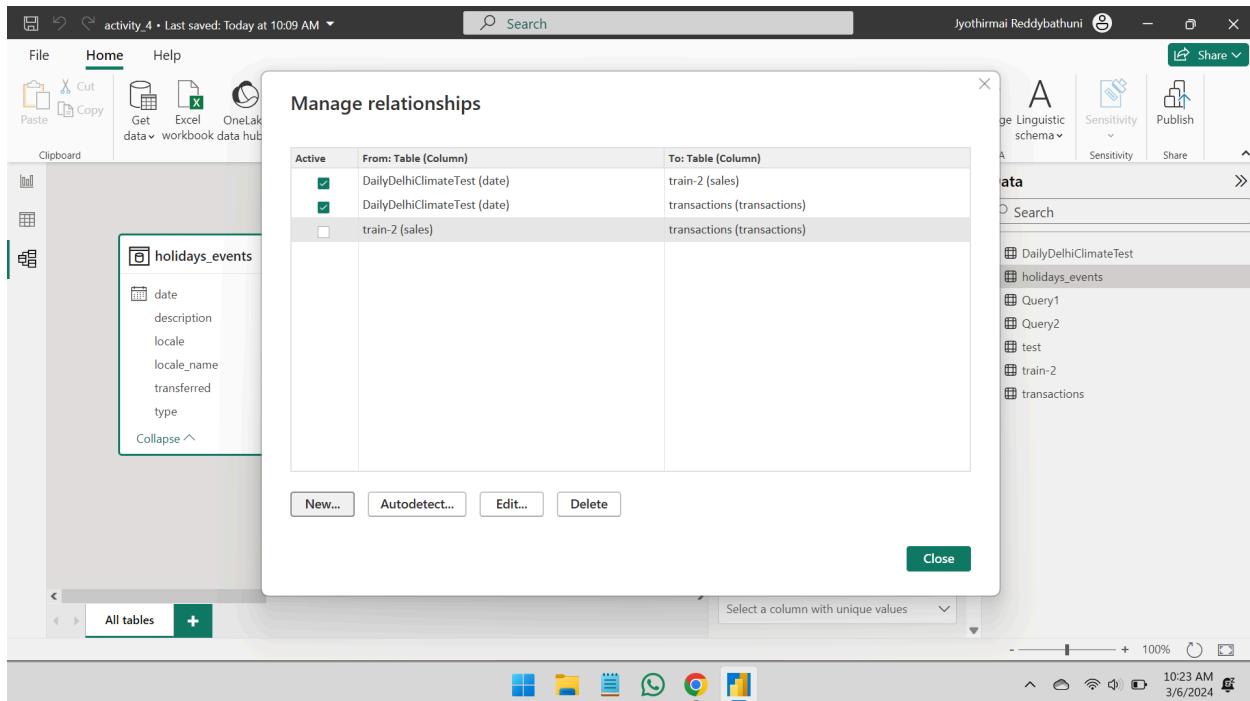


Fig.18. List of relations in the workbook

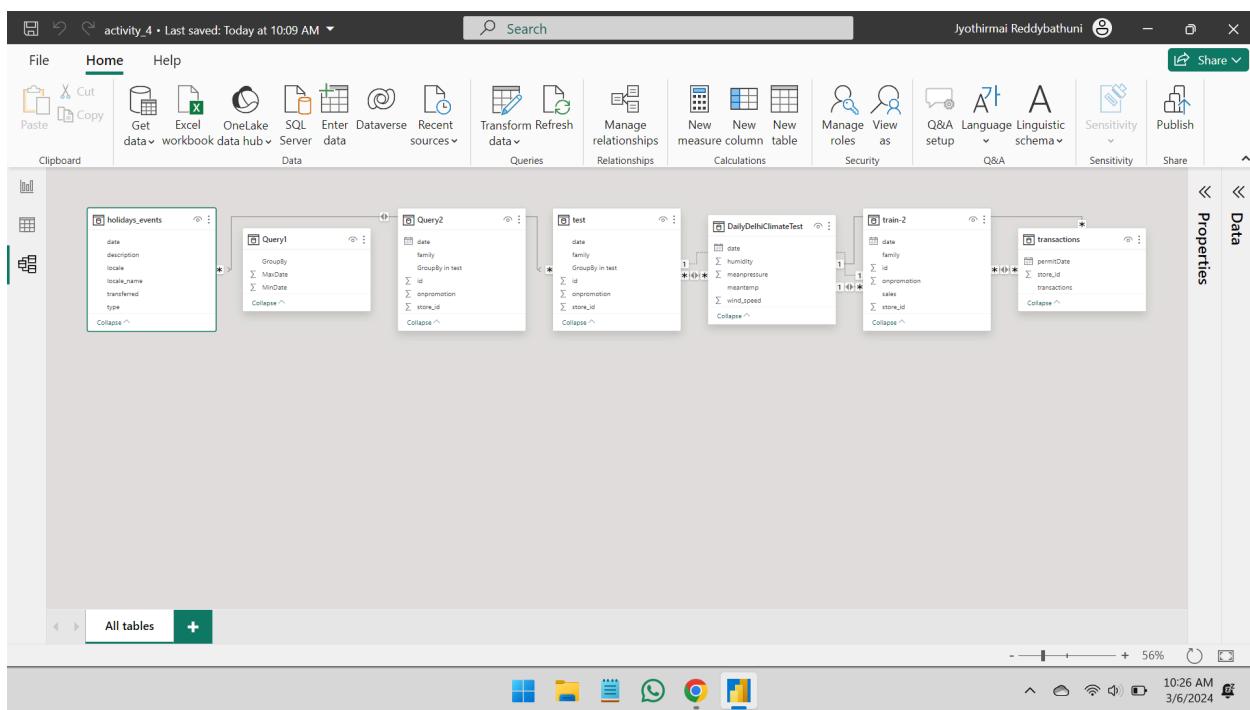


Fig.19. Model View

Question- 2

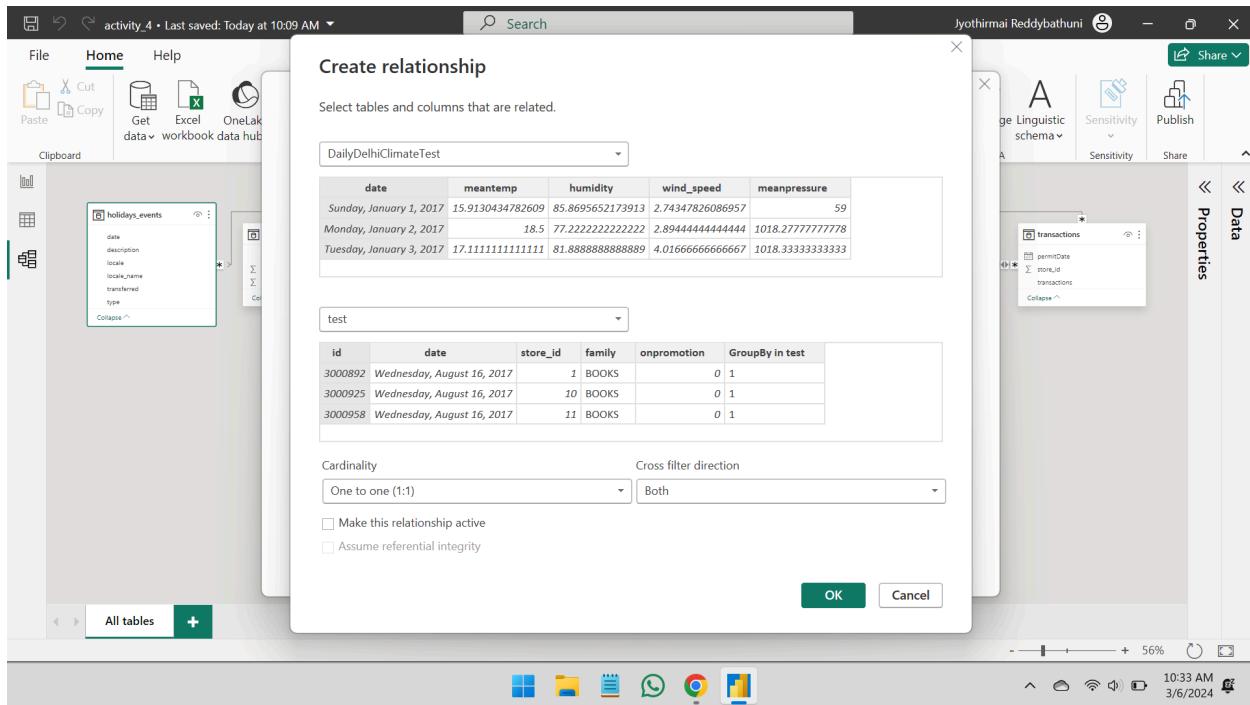


Fig.20. One to One relationship

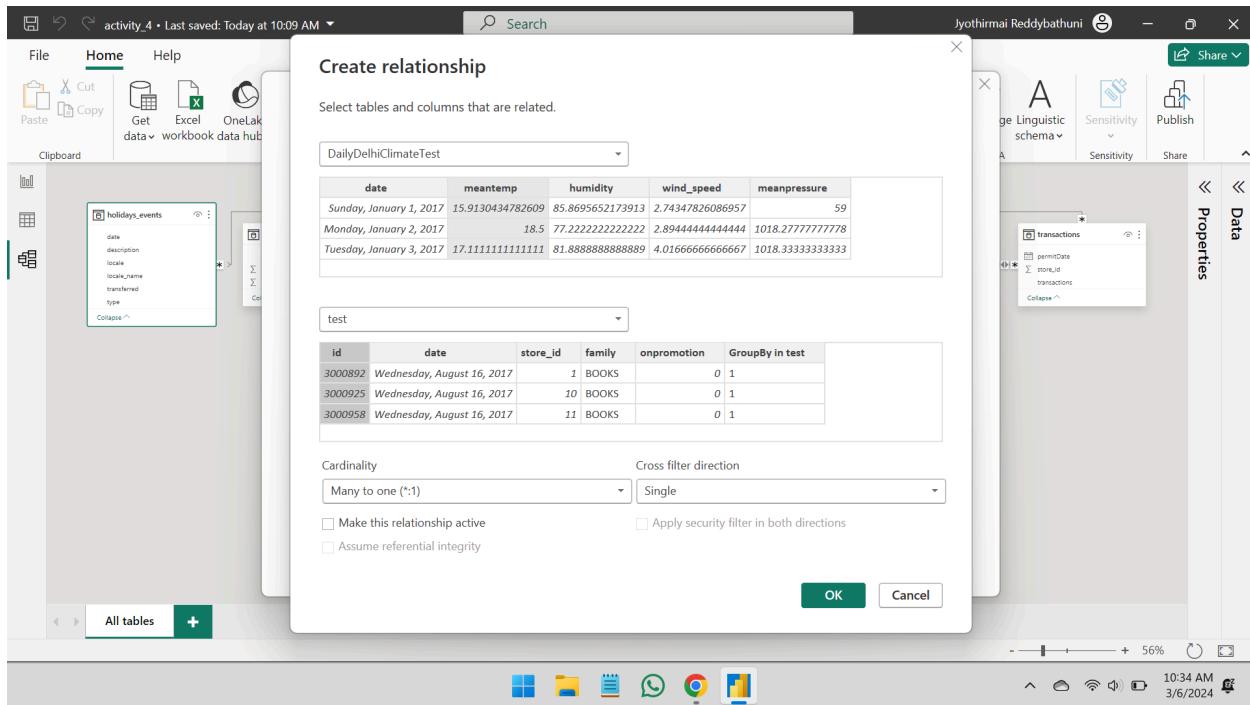


Fig.21. Many to one relationship

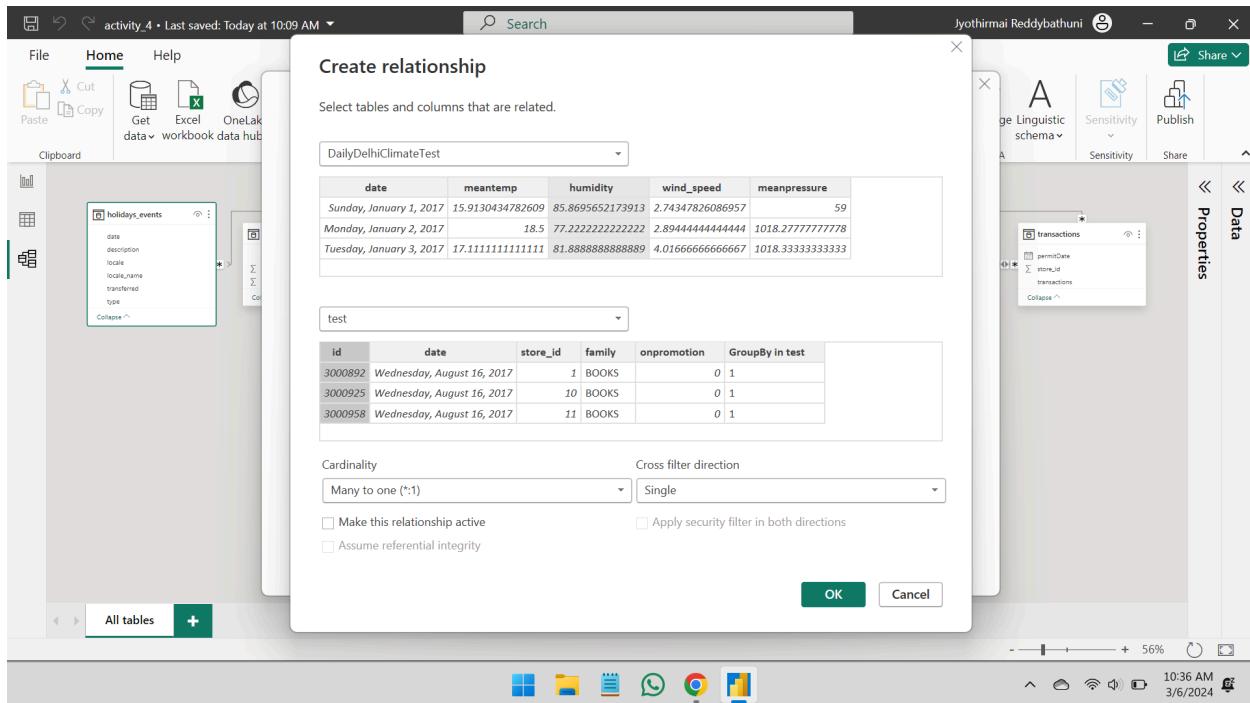


Fig.22. Many to one relationship

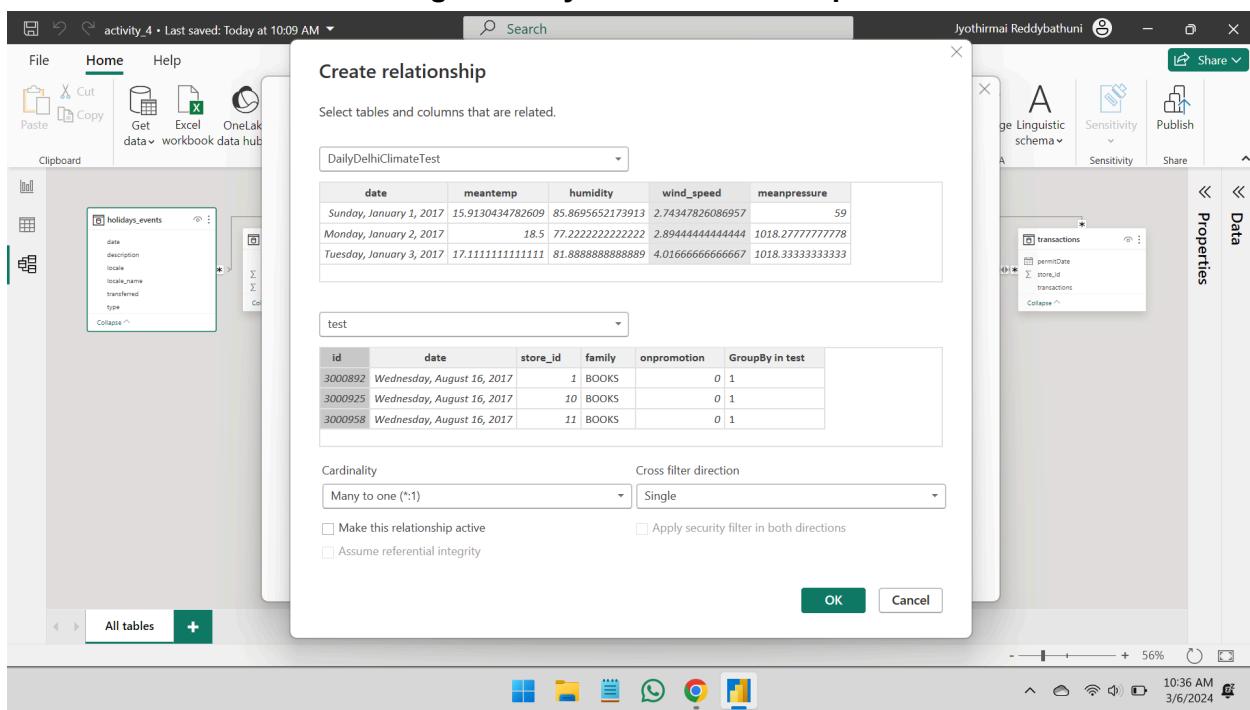


Fig.23. Many to one relationship

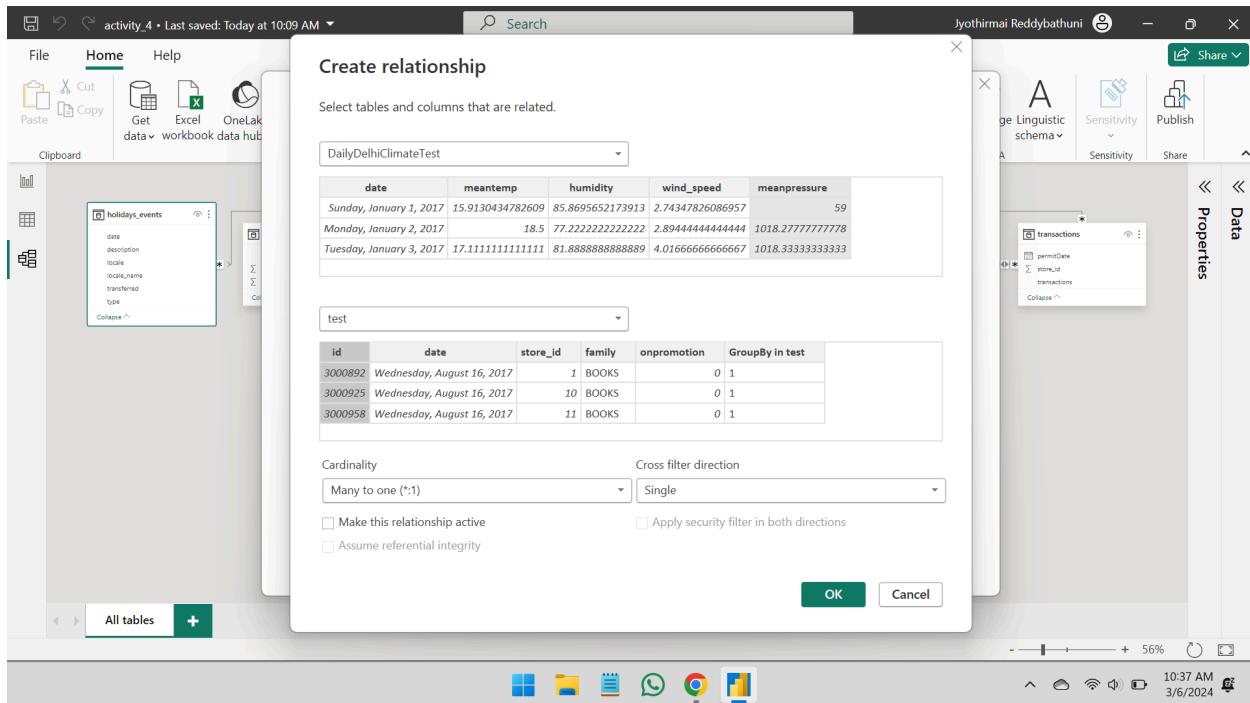


Fig.24. Many to one relationship

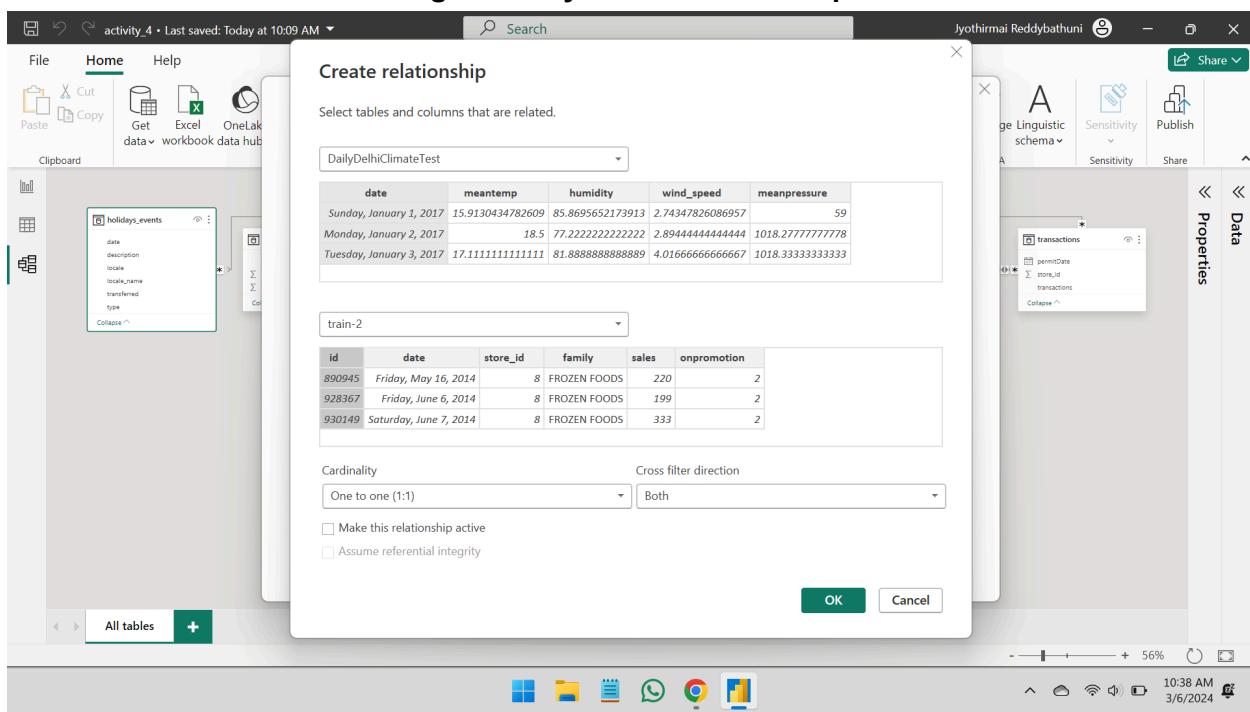


Fig.25. One to One relationship

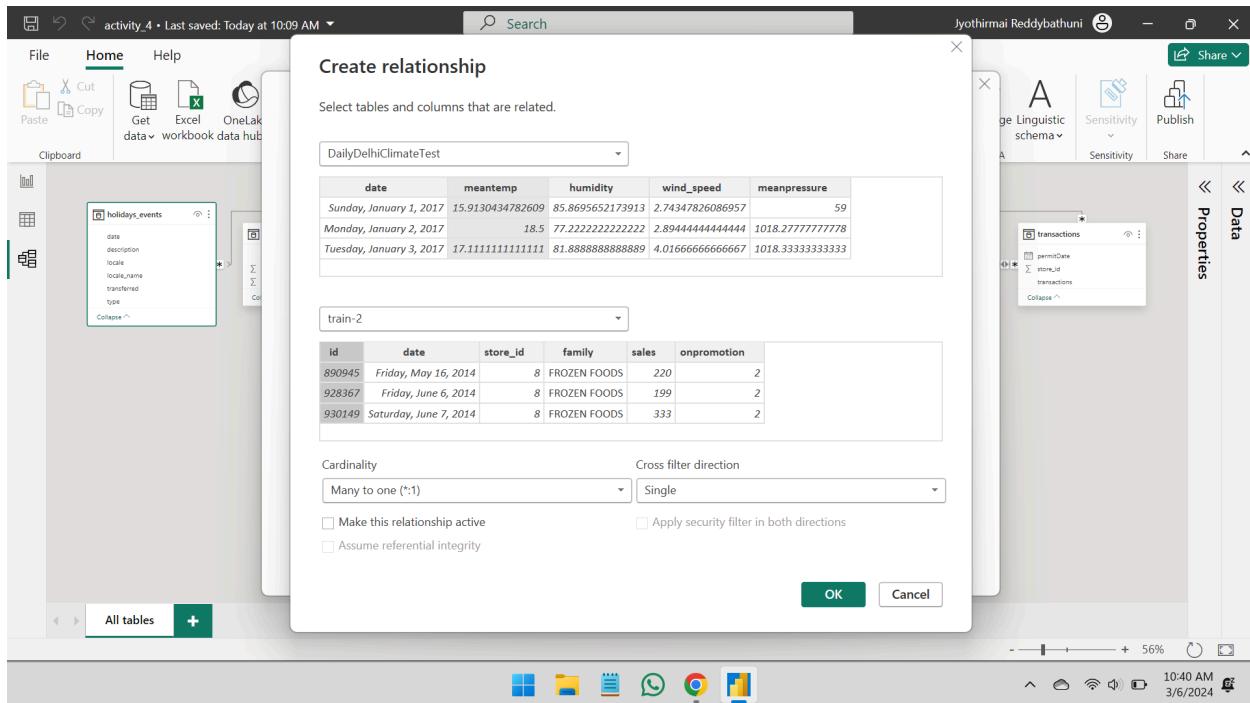


Fig.26. Many to one relationship

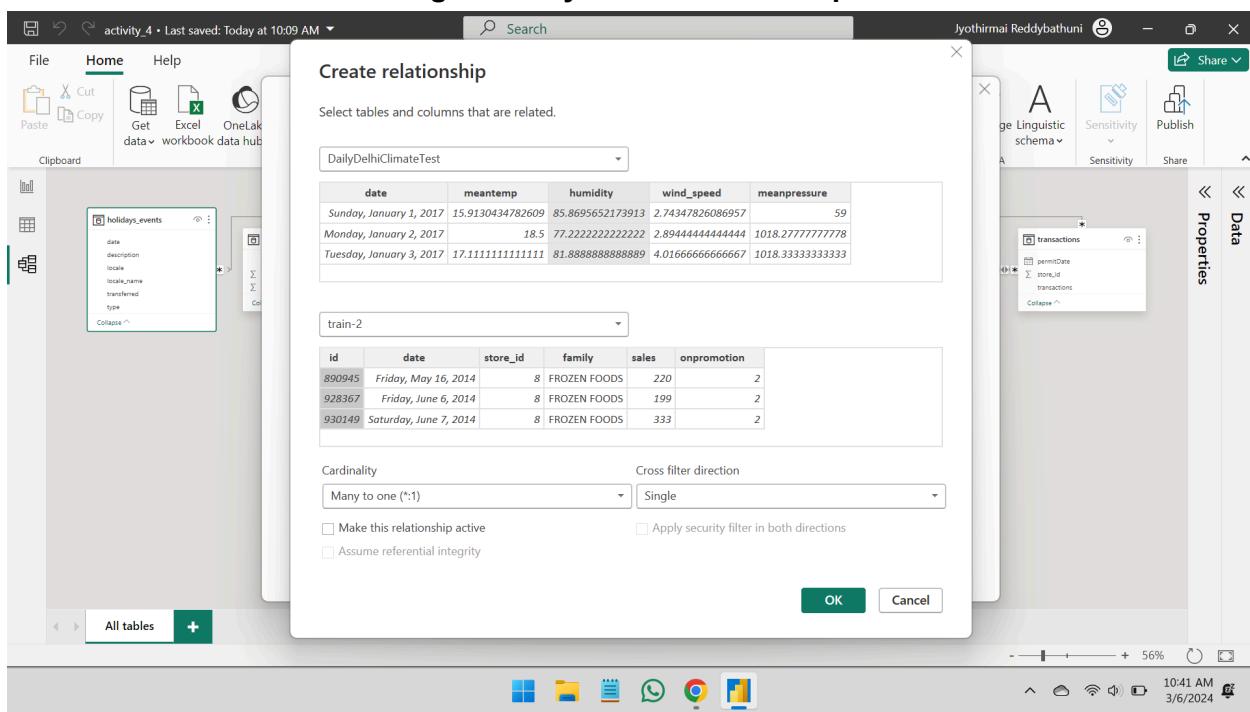


Fig.27. Many to one relationship

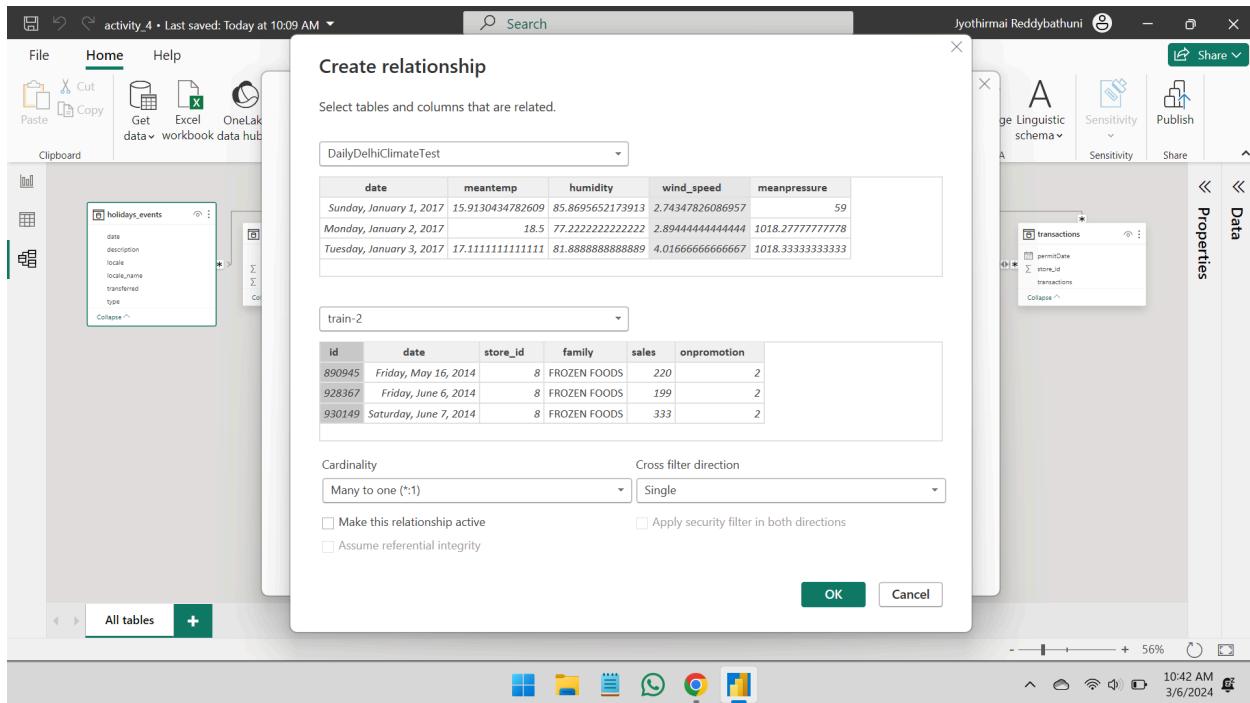


Fig.28. Many to one relationship

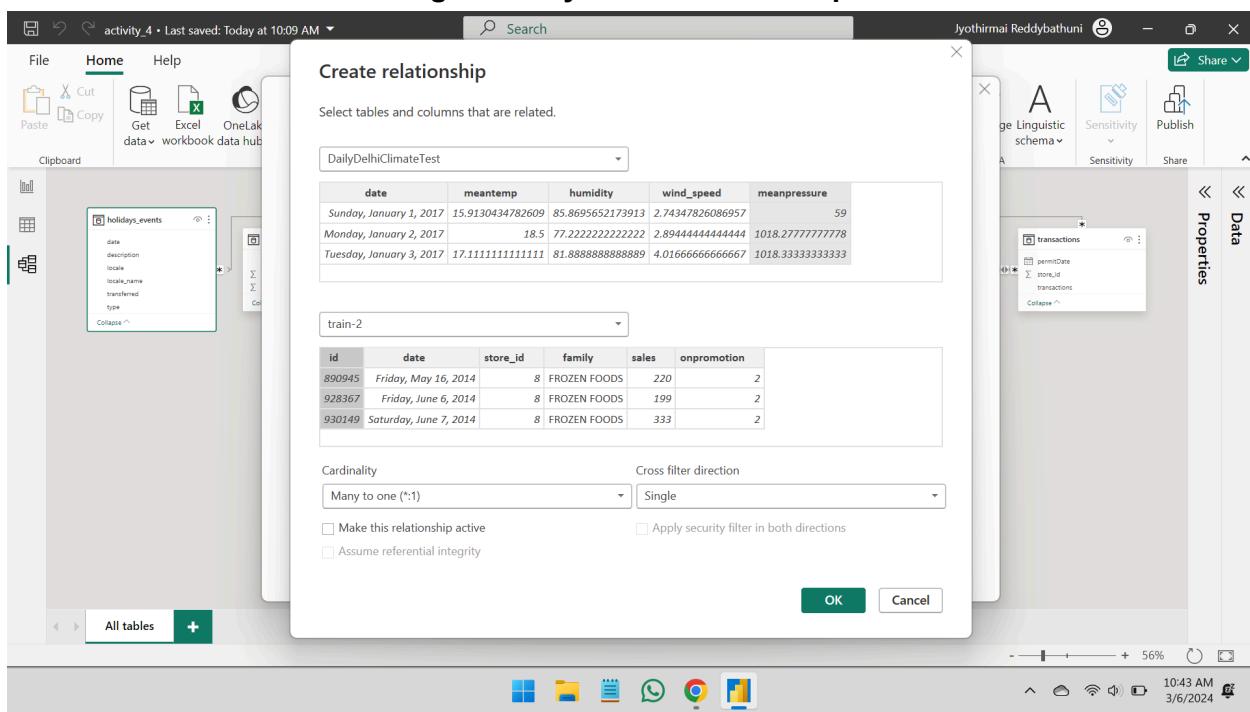


Fig.29. Many-to-one relationship

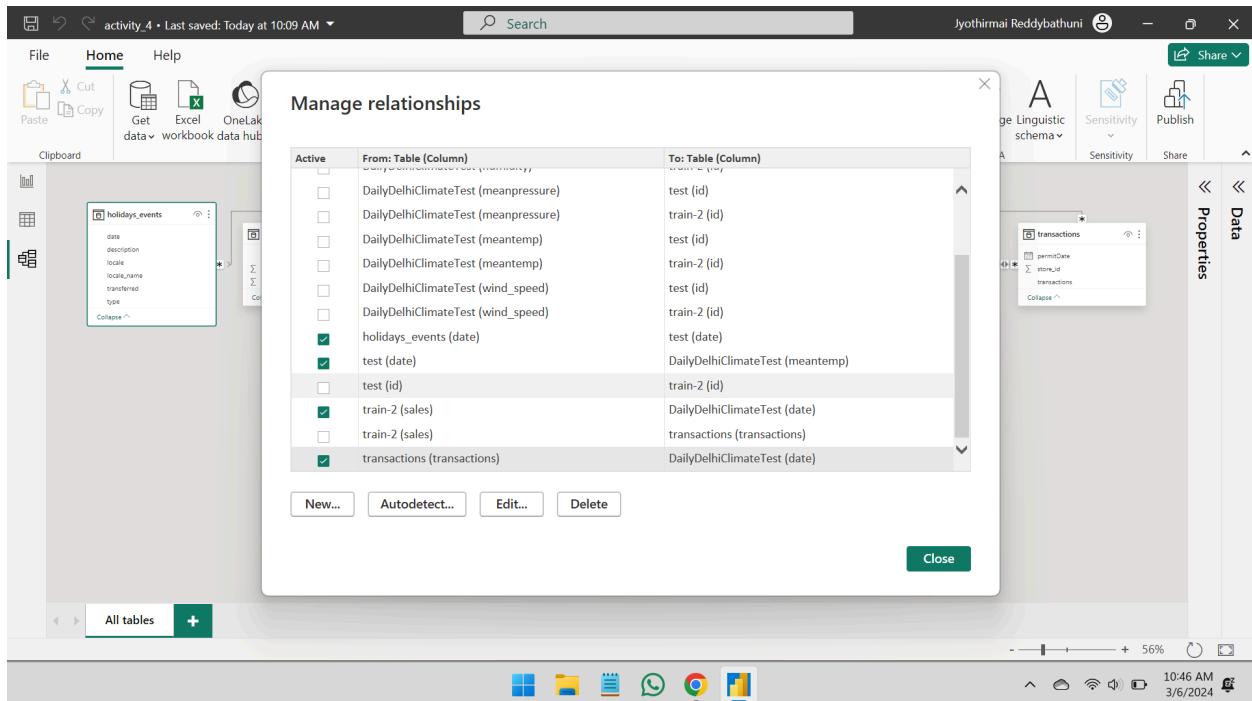


Fig.30. Managing relationships

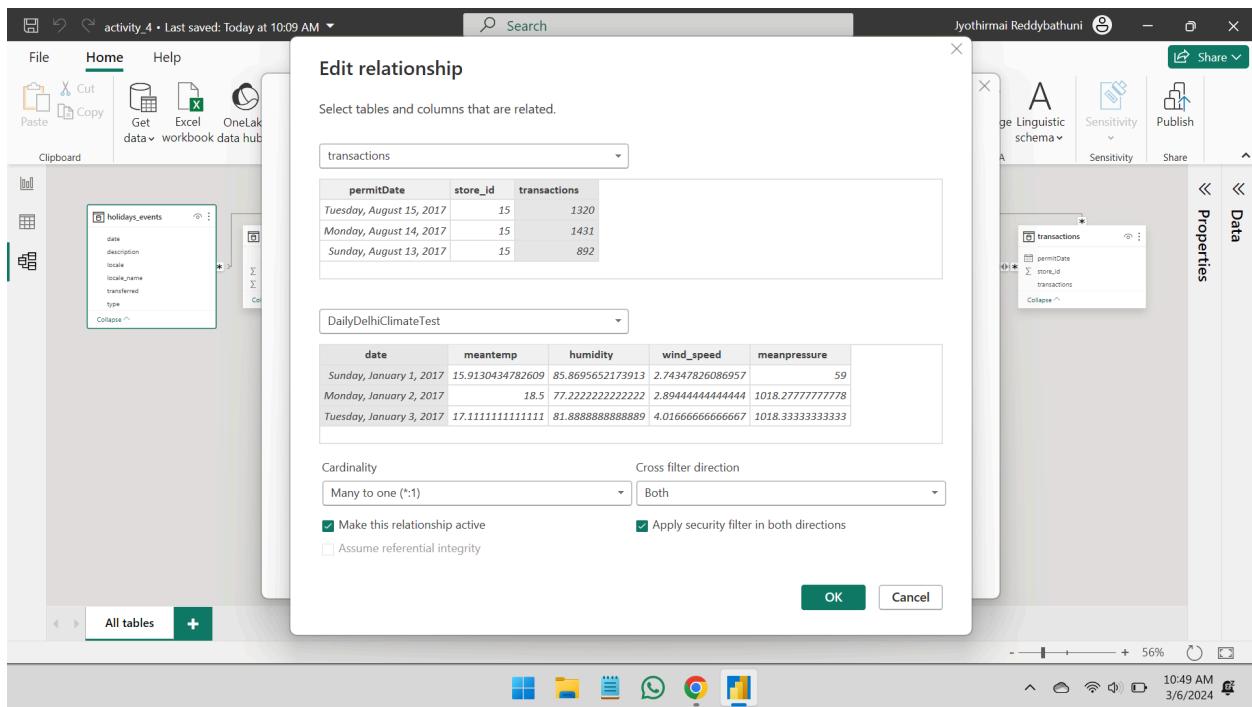


Fig.31. Cross filter Direction

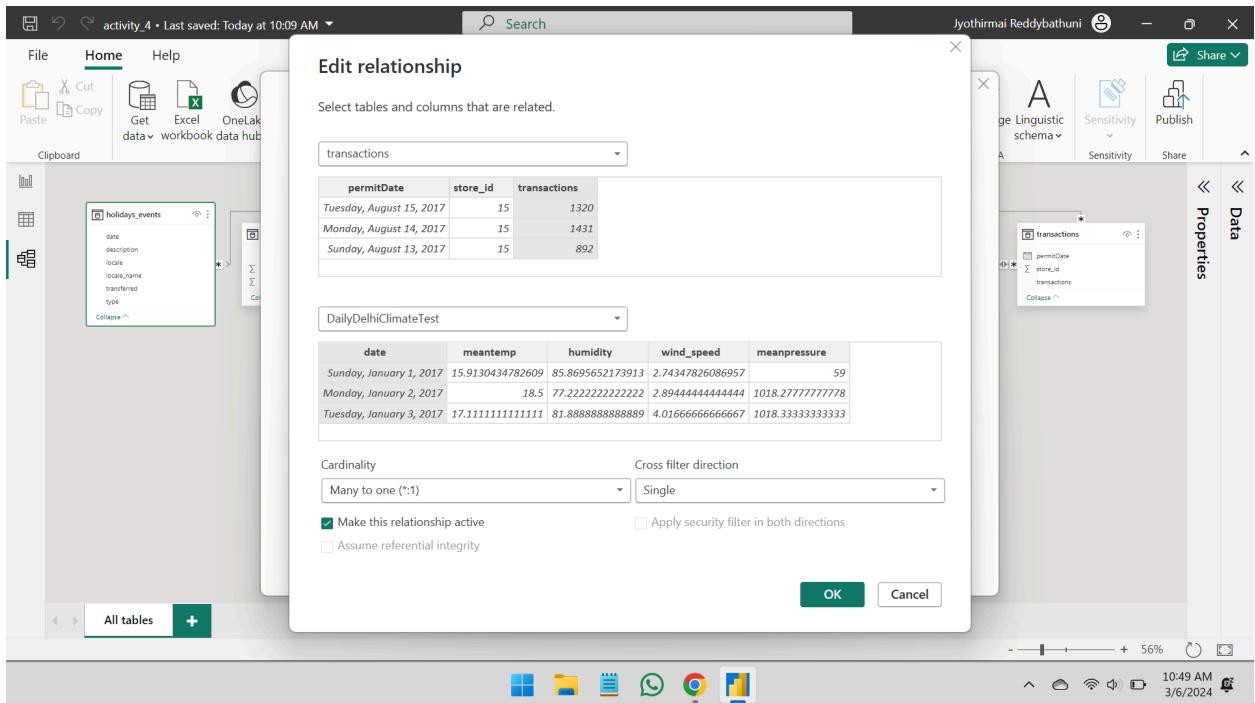


Fig.32. Cross filter direction

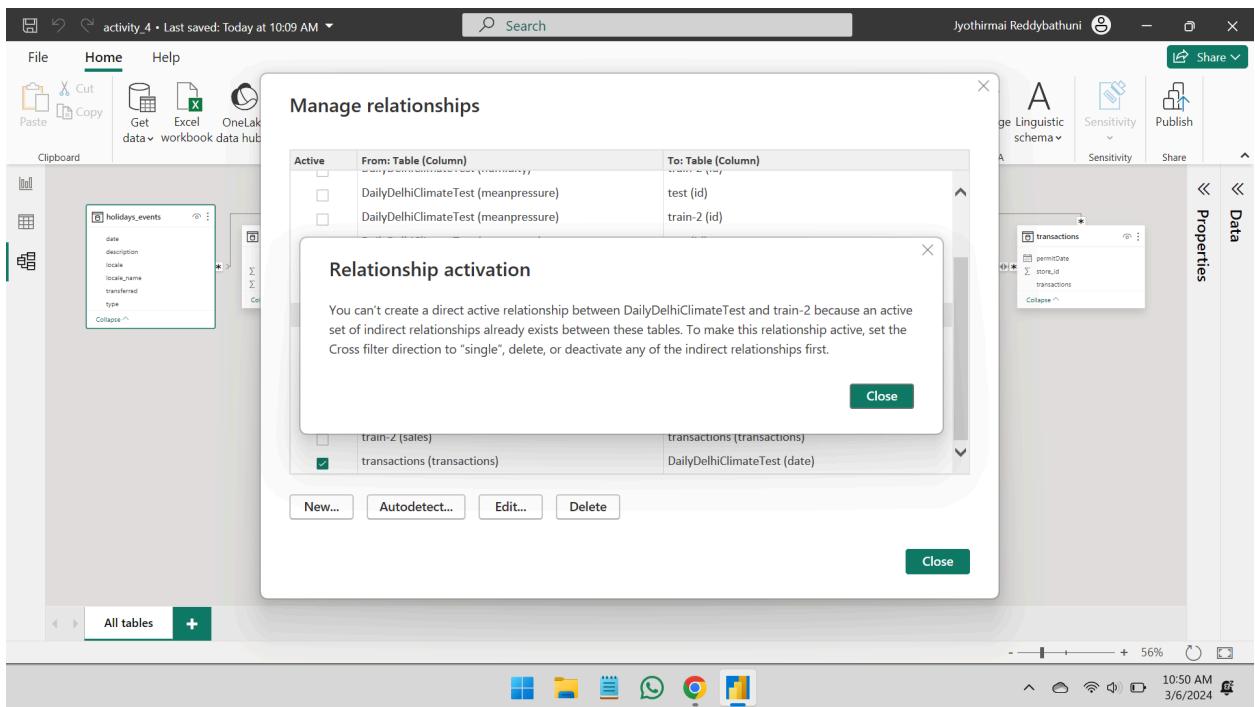


Fig.33. Relationship Activation

Tutorial 3

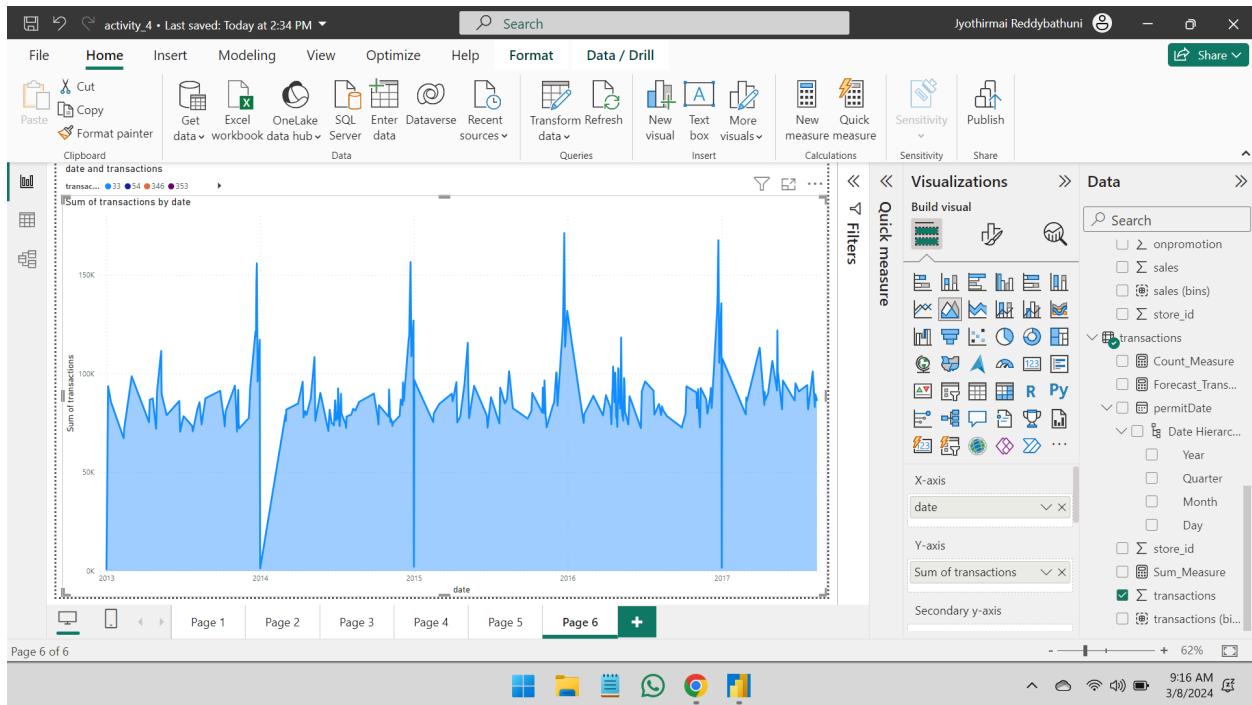


Fig.34. Stacked Area chart

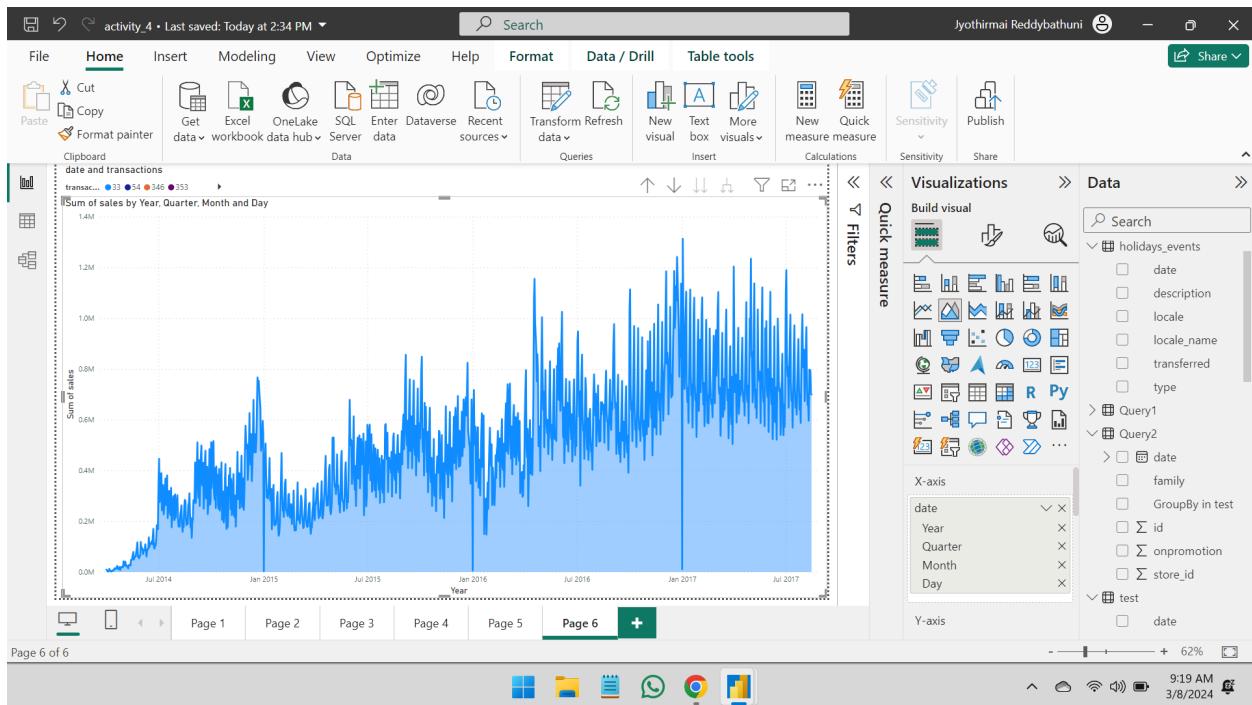


Fig.35. Stacked Area chart

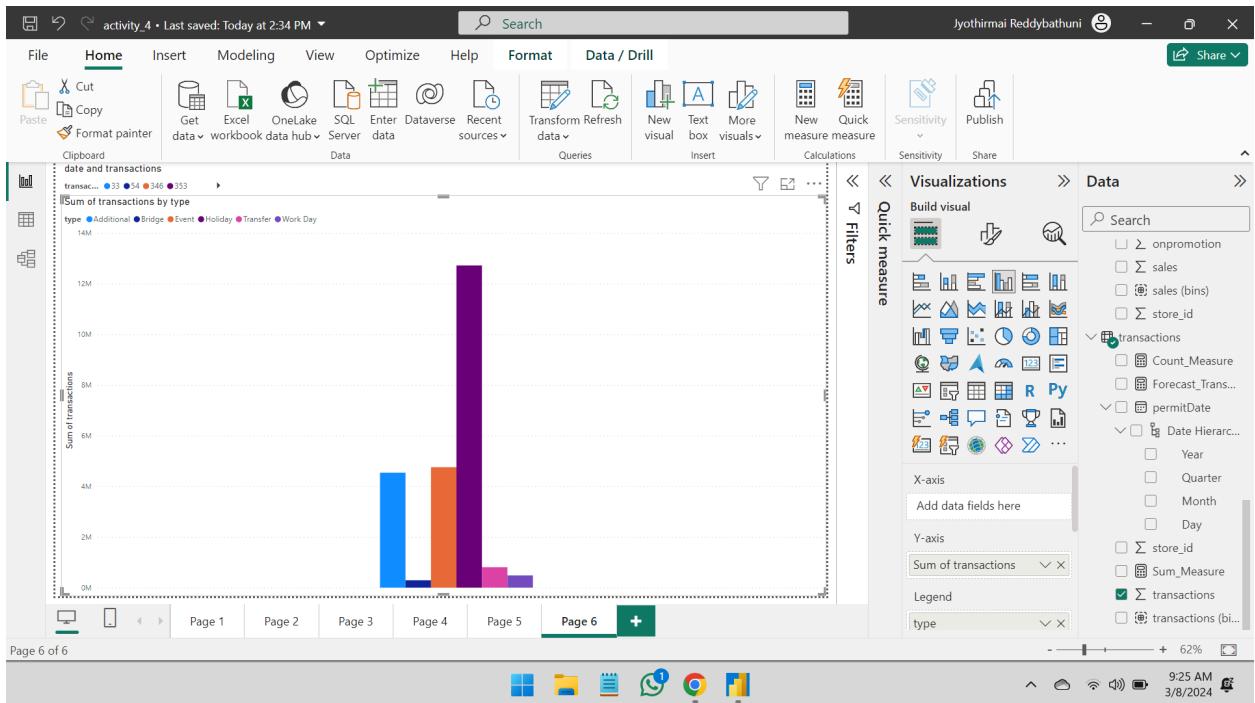


Fig.36. Clustered Column chart

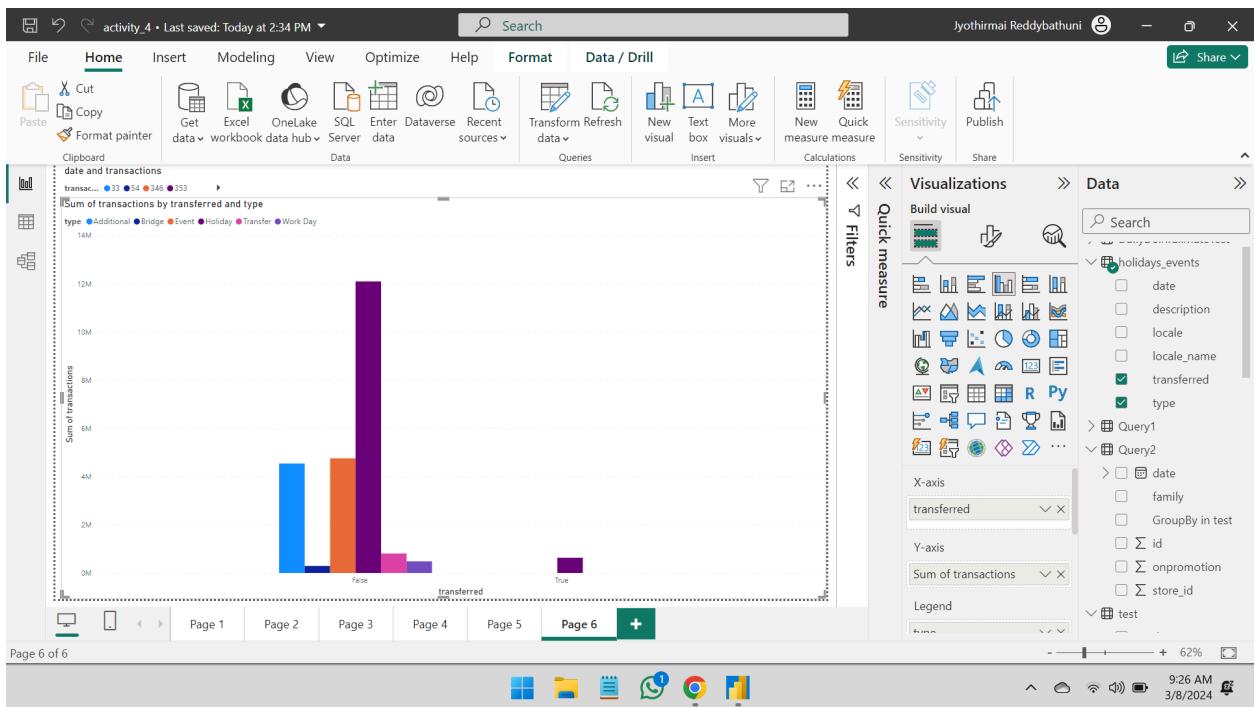


Fig.37. Clustered Column chart

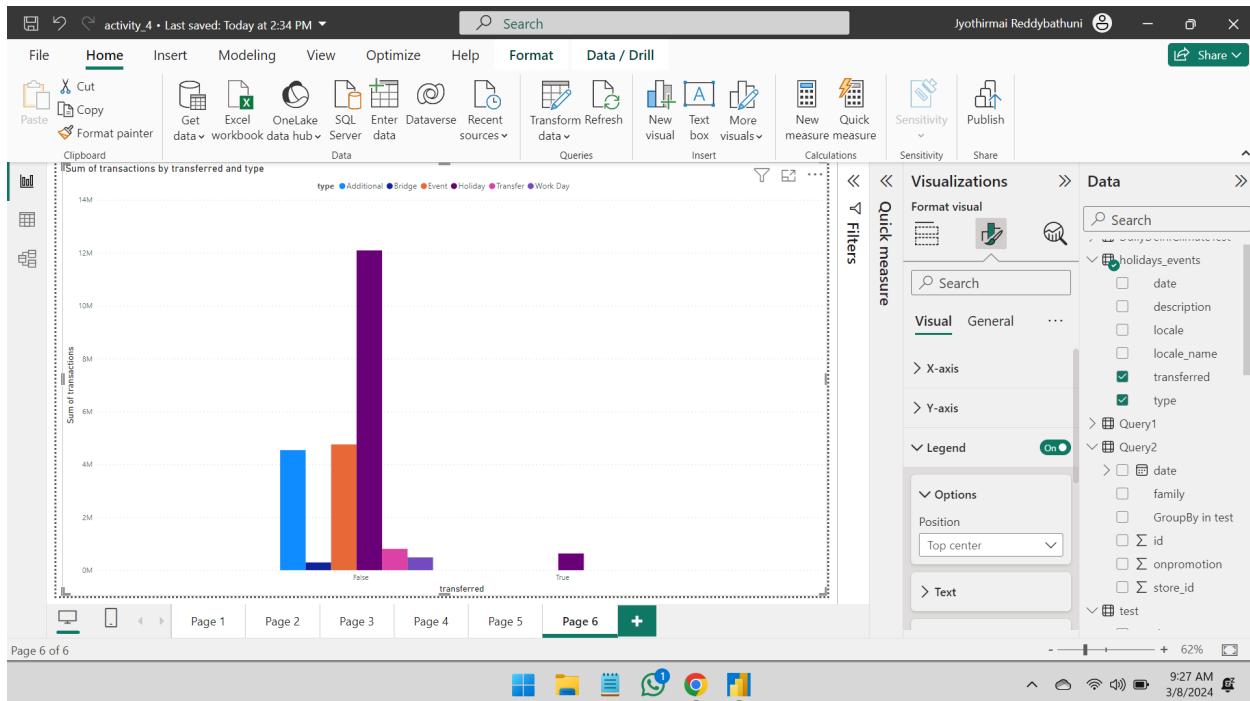


Fig.38.Clustered Column chart

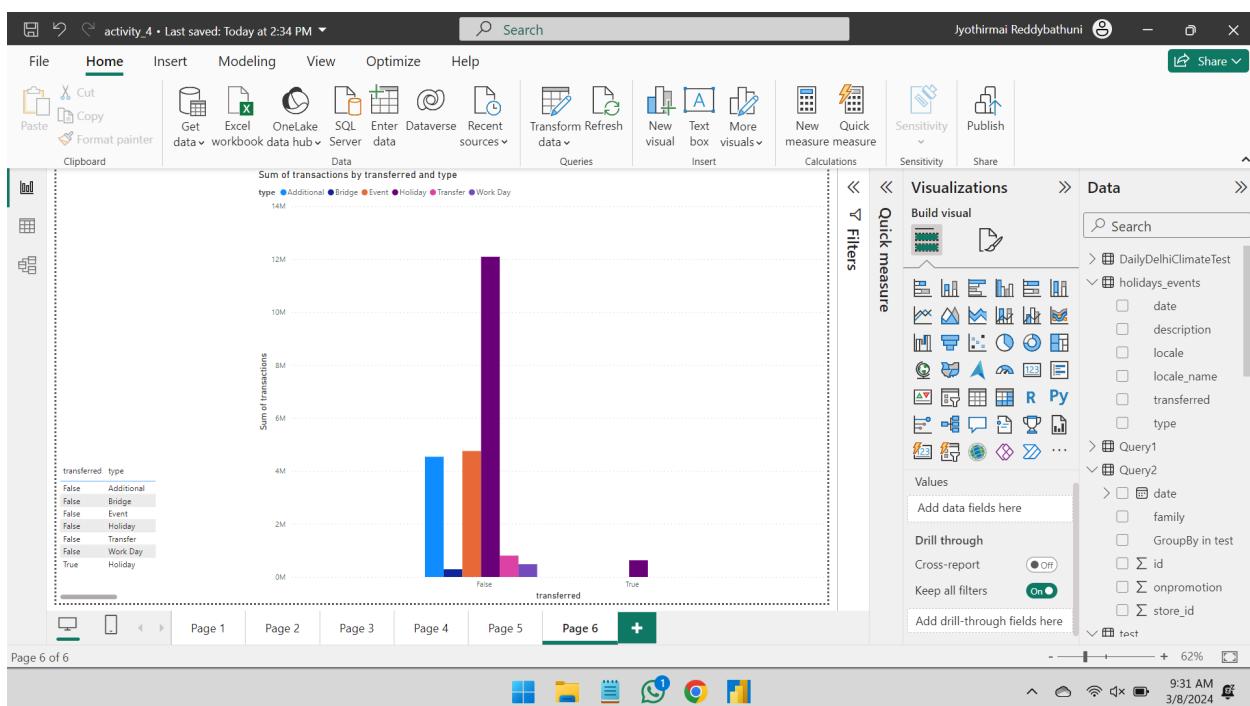


Fig.39.Clustered Column chart with visual table

Question 3

The screenshot shows the Microsoft Power BI interface. On the left, there is a visual table titled "Clipboard" with columns "type", "transferred", "date", and "Sum of transactions". The table contains numerous rows of data, with a total sum of 141478945. On the right, the "Data" pane is open, displaying various data sources and measures. The "transactions" measure is selected. The bottom status bar shows the date as 3/8/2024 and the time as 9:32 AM.

Fig.40. Visual Table

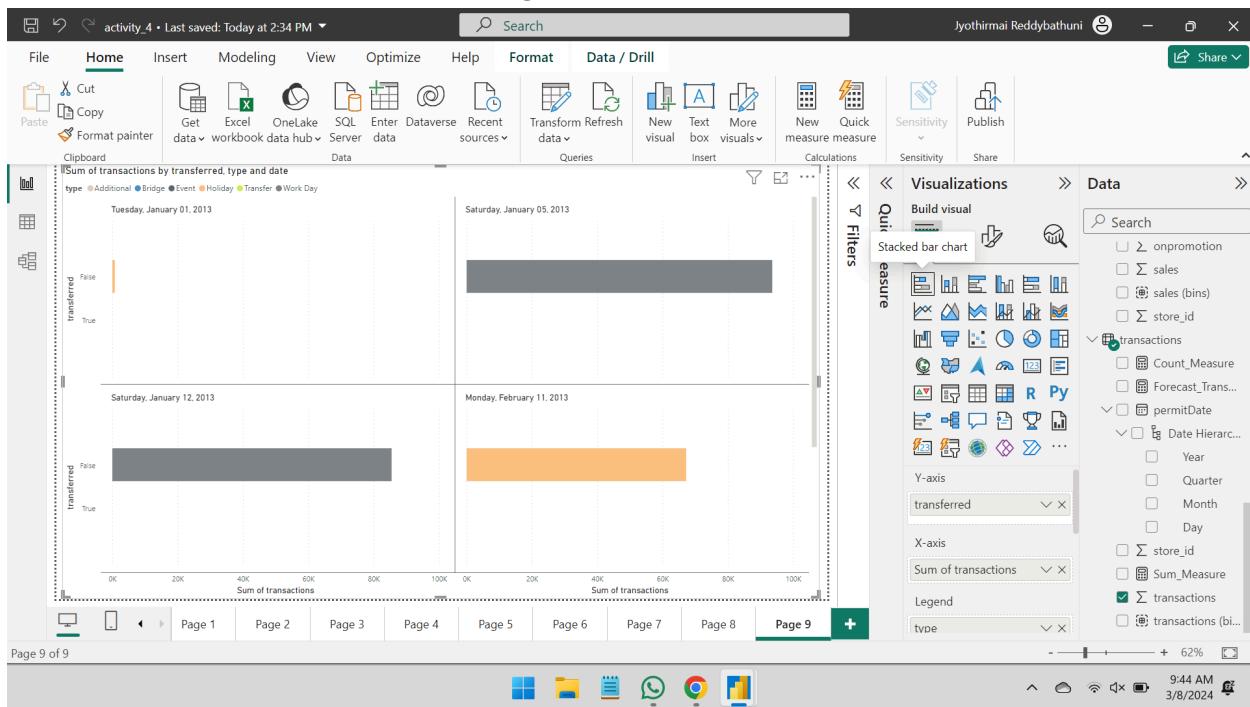


Fig.41. Stacked bar chart

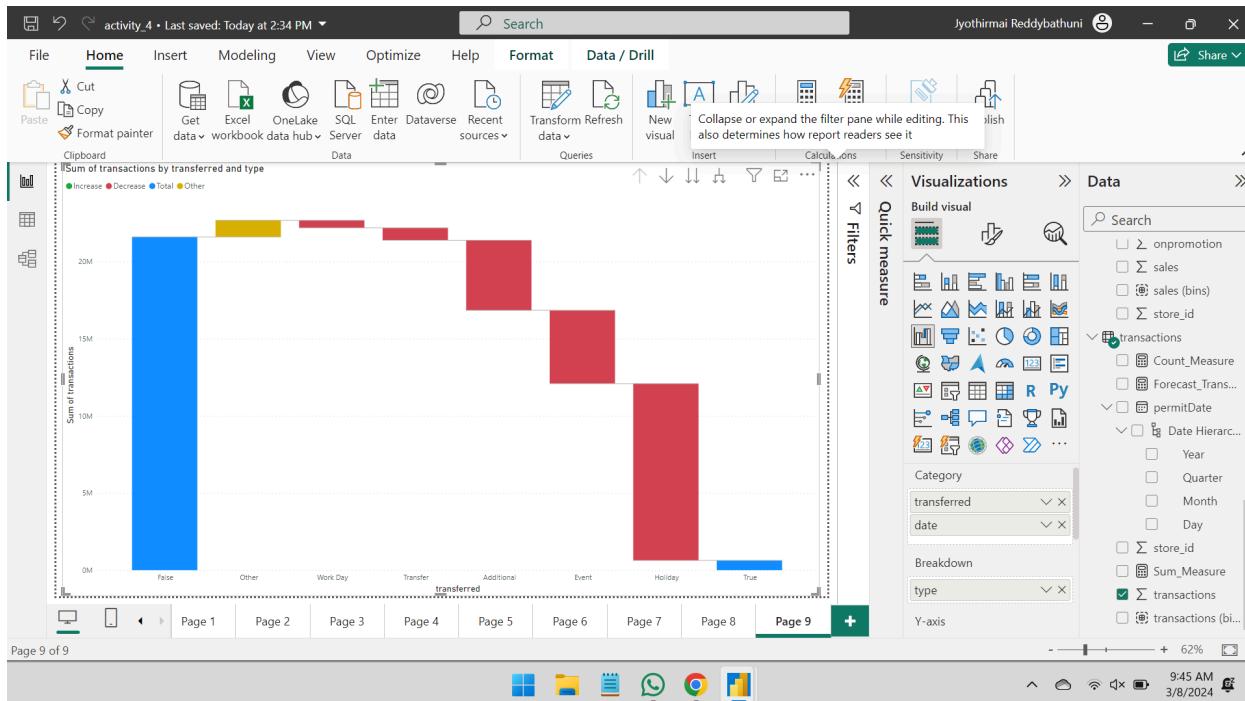


Fig.42. Waterfall model

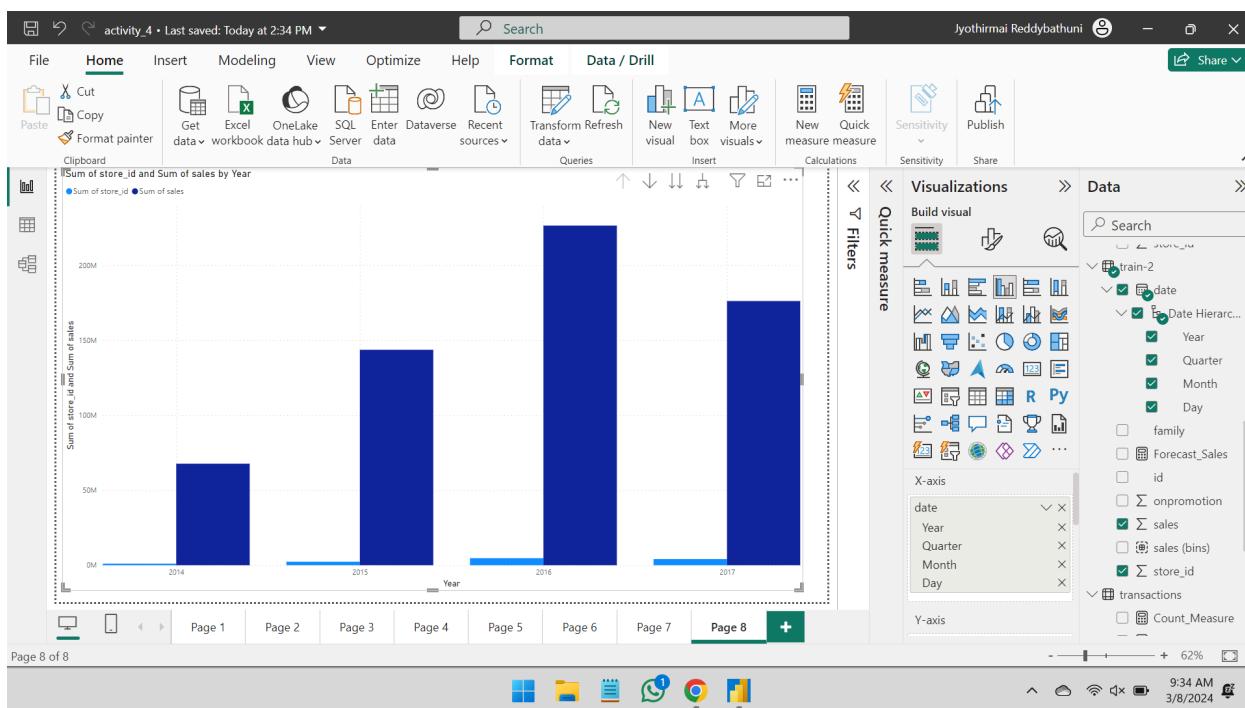


Fig.43. Clustered Column chart

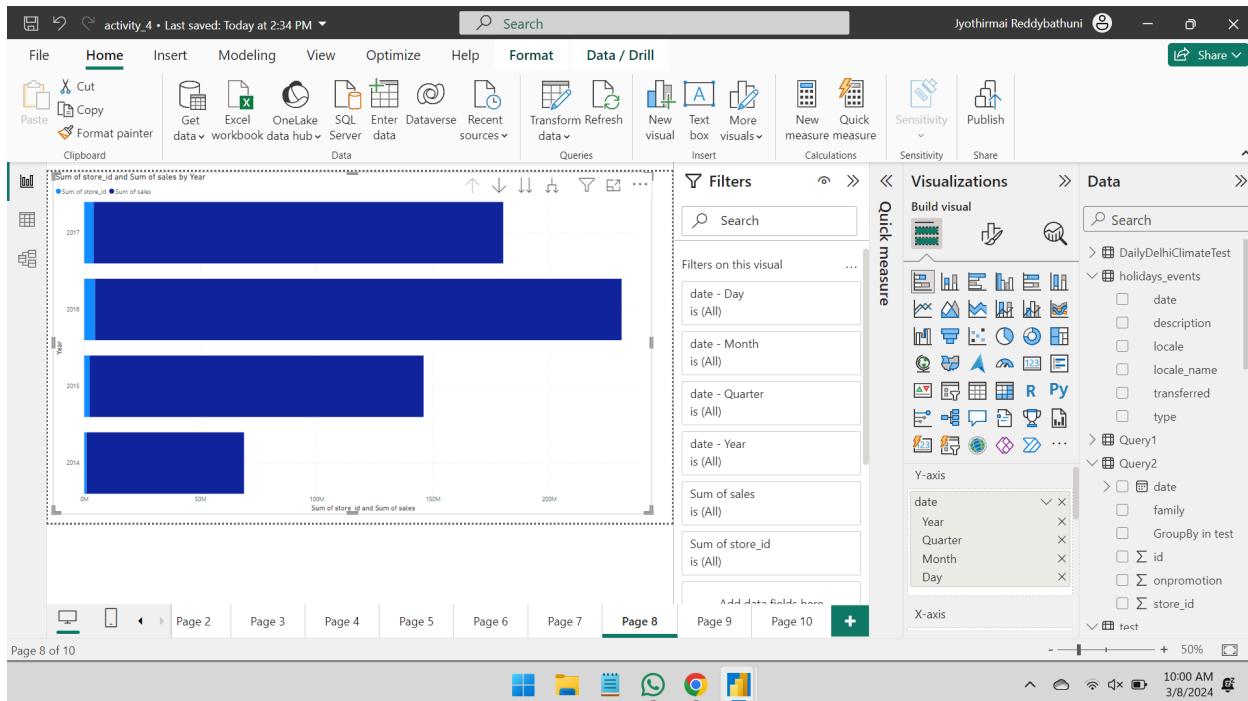


Fig.44. Stacked Bar Chart

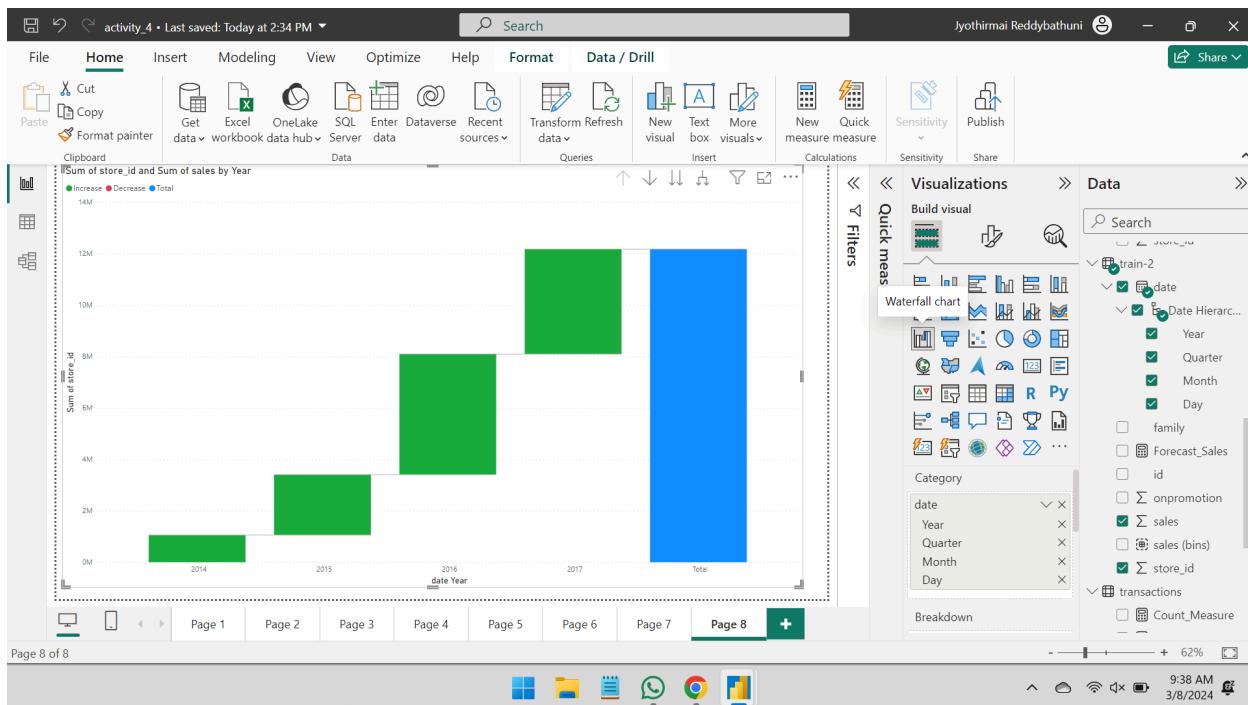


Fig.45. Waterfall model

The screenshot shows the Power BI desktop interface with a visual table. The table has columns: Sum of sales, Year, Quarter, Month, Day, and Sum of store_id. The data includes rows such as 9437, 2014, Qtr 2, April, 1, 18; 6481, 2014, Qtr 2, April, 2, 9; etc. To the right, the 'Data' pane shows filters for date (Year, Quarter, Month, Day) and store_id. The bottom status bar shows the date as 3/8/2024 and time as 9:54 AM.

Fig.46. Visual Table

The screenshot shows the Power BI desktop interface with a visual table. The table has columns: sales, Year, and Sum of store_id. The data shows sales for 2016 across various store IDs. A filter for 'date - Year' is applied to show only 2016 data. The 'Data' pane shows filters for date (Year, Quarter, Month, Day) and store_id. The bottom status bar shows the date as 3/8/2024 and time as 12:02 PM.

Fig.47. Result for Question

Tutorial 4:

Queries [9]

- Query Errors - 3/1/20...
- Query Errors - 3/6/20...
- Other Queries [7]
- DailyDelhiClimateTest**
- Query1
- train-2
- holidays_events
- test
- Query2

Table.AddColumn(#"Changed Type", "Mean", each List.Average(DailyDelhiClimateTest[meantemp]))

	date	1.2 meantemp	1.2 humidity	1.2 wind_speed	1.2 meanpressure
1	1/1/2017	15.91304348	85.86956522	2.743478261	1018.277
2	1/2/2017	18.5	77.22222222	2.894444444	1018.333
3	1/3/2017	17.11111111	81.88888889	4.016666667	1018.333
4	1/4/2017	18.7	70.05	4.545	101
5	1/5/2017	18.38888889	74.94444444	3.3	1014.333
6	1/6/2017	19.31818182	79.31818182	8.681818182	1011.772
7	1/7/2017	14.70833333	95.83333333	10.04166667	1011.
8	1/8/2017	15.68421053	83.52631579	1.95	1015.952
9	1/9/2017	14.57142857	80.80952381	6.542857143	1015.952
10	1/10/2017	12.11111111	71.94444444	9.361111111	1016.888
11	1/11/2017	11	72.11111111	9.77222222	1016.777
12	1/12/2017	11.78947368	74.57894737	6.626315789	1016.368
13	1/13/2017	13.23529412	67.05882353	6.435294118	1017.529
14	1/14/2017	13.2	74.28	5.276	1016
15	1/15/2017	16.43478261	72.56521739	3.630434783	1018.130
16	1/16/2017	14.65	78.45	10.38	1017
17	1/17/2017	11.72222222	84.44444444	8.03888889	1018.388
18	1/18/2017	13.04166667	78.33333333	6.029166667	1021.958
19	1/19/2017	14.61904762	75.14285714	10.33809524	1022.809
20	1/20/2017	15.26315789	66.47368421	11.22631579	1021.789
21					

6 COLUMNS, 114 ROWS Column profiling based on top 1000 rows

PREVIEW DOWNLOADED AT 12:14 PM

12:15 PM 3/7/2024

Fig.48. Dataset

Groups

Name *: sales (bins)

Field: sales

Group type: Bin

Bin type: Size of bins

Min value: 2

Max value: 124717

Bin size *: 820

Reset to default

OK Cancel

Data

store_id

test

- date
- family
- GroupBy in test
- id
- onpromotion
- sales
- store_id

train-2

- date
- family
- Forecast_Sales
- id
- onpromotion
- sales
- store_id

transactions

Fig.49. Adding group

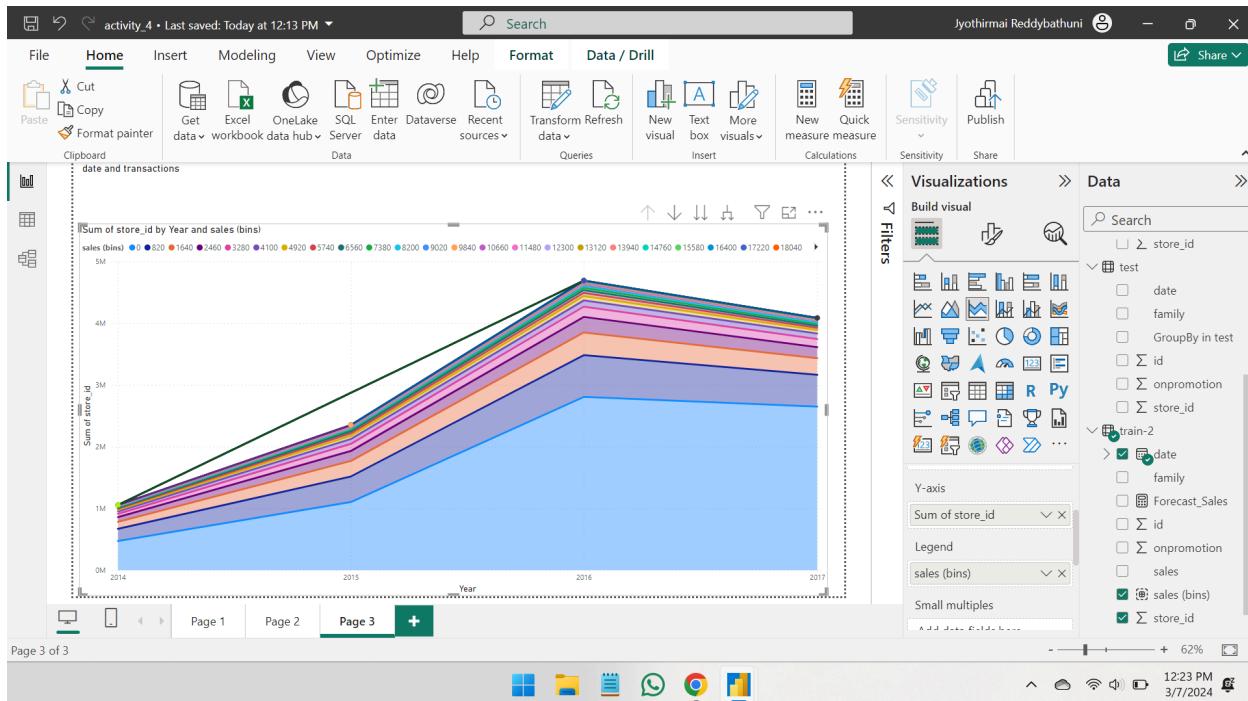


Fig.50. Stacked Area Chart

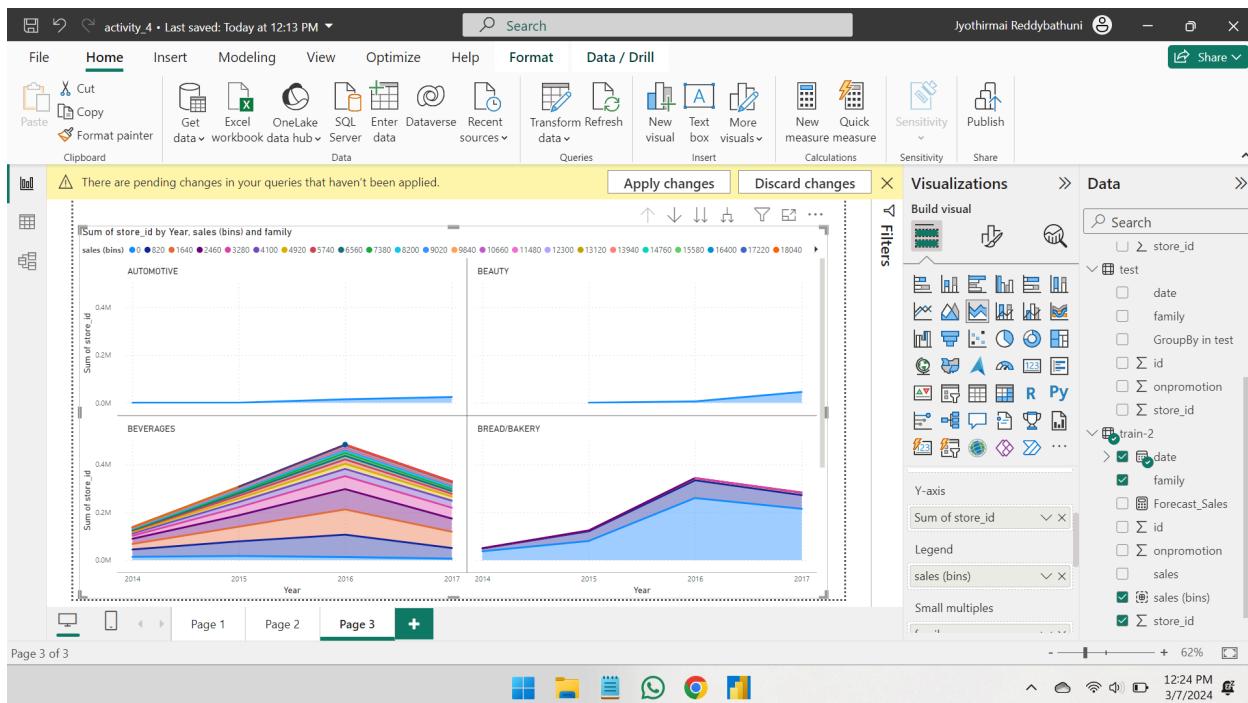


Fig.51. Stacked Area Chart

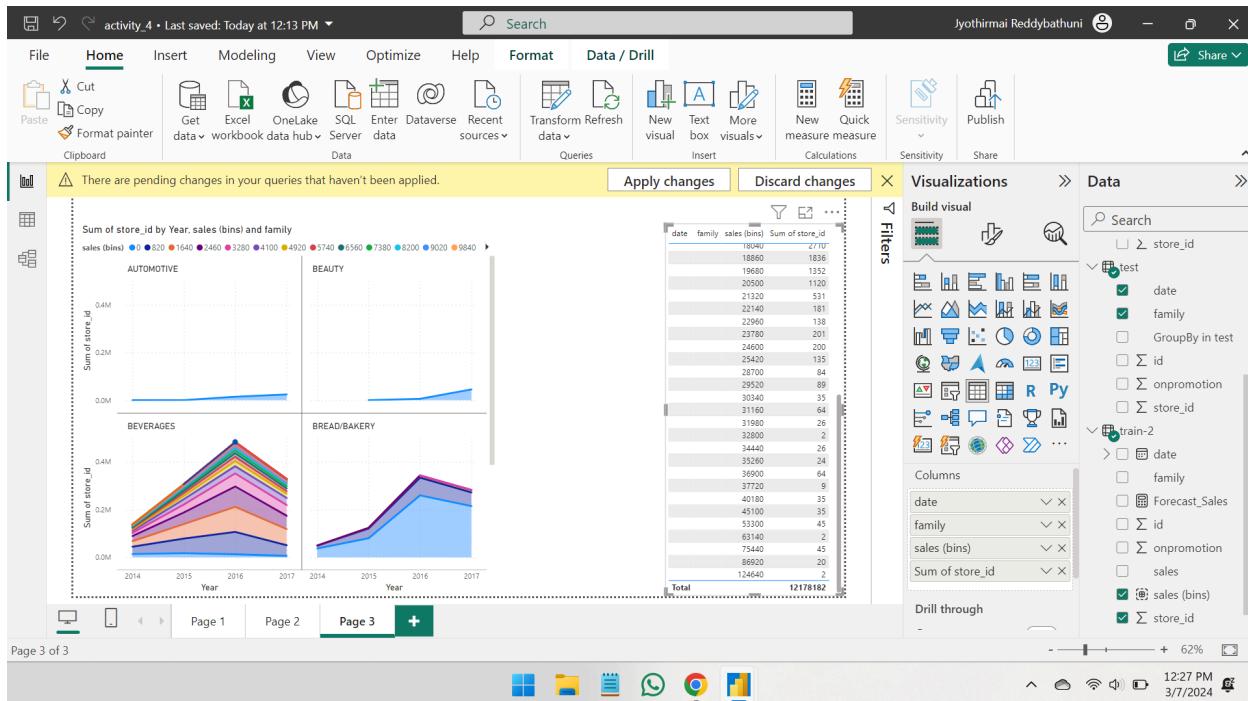


Fig.52. Stacked Area Chart and Visual Table

Question 4

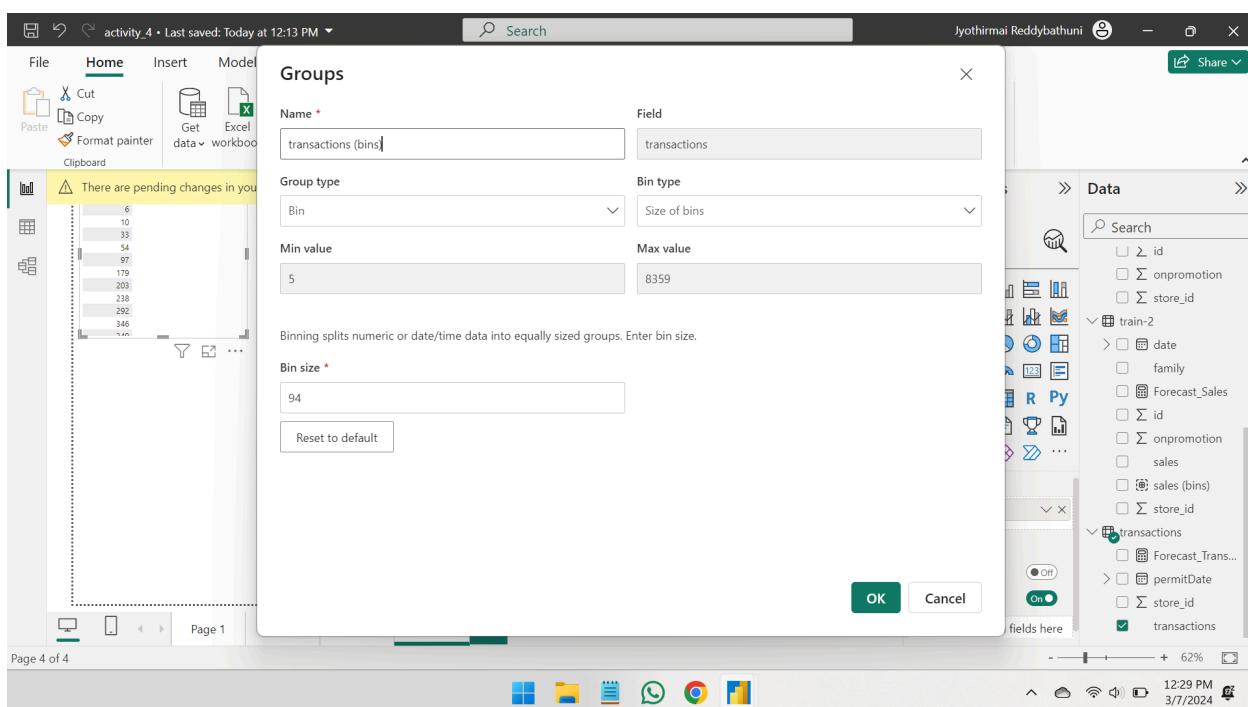


Fig.53. Adding a new group

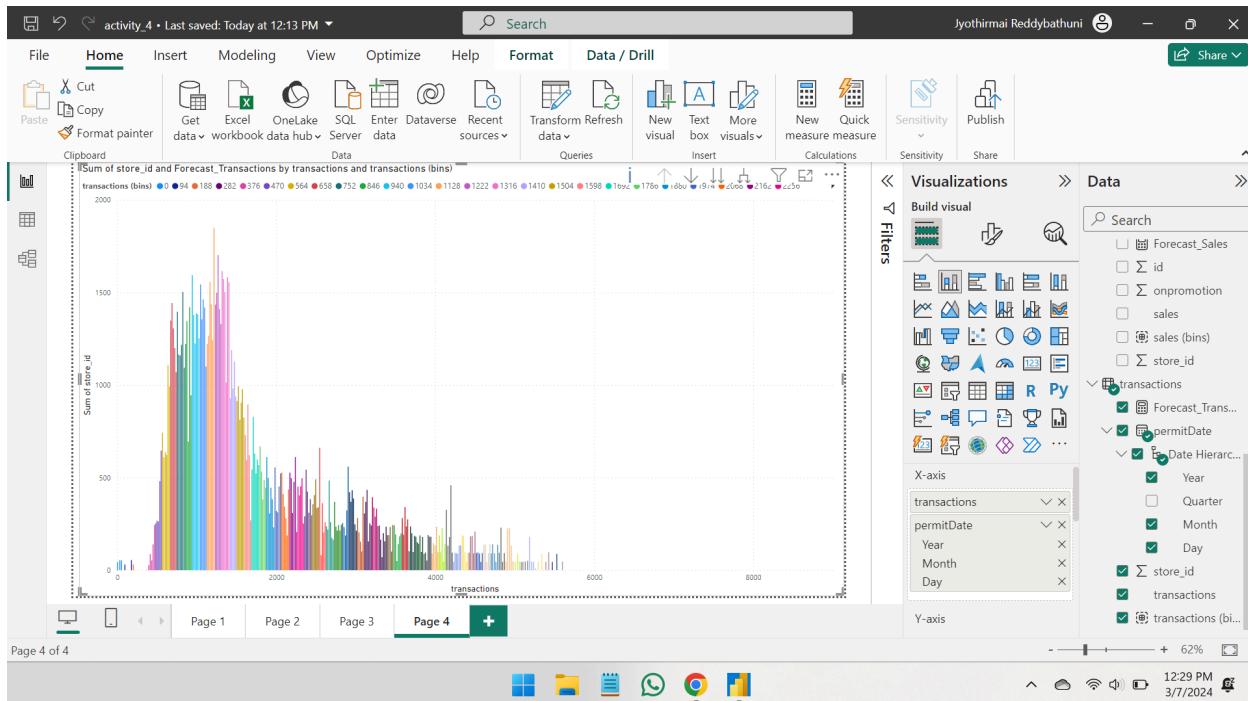


Fig.54. Stacked column chart

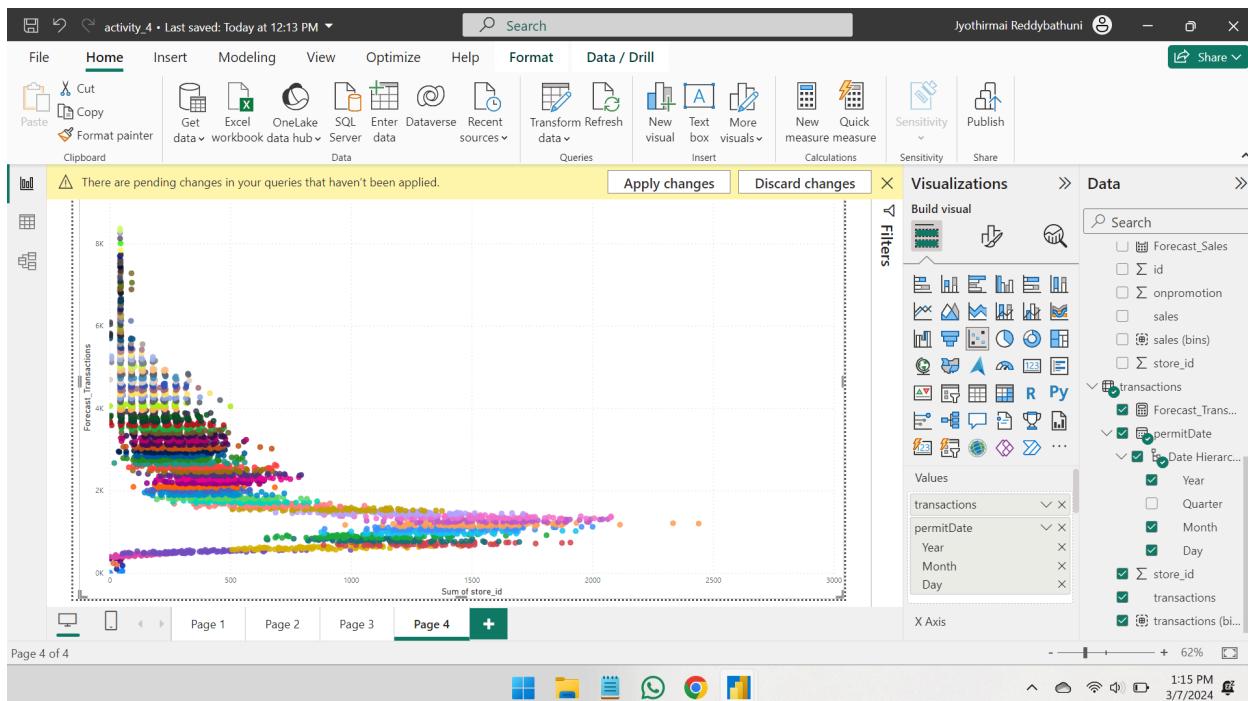


Fig.55. Scatter plot

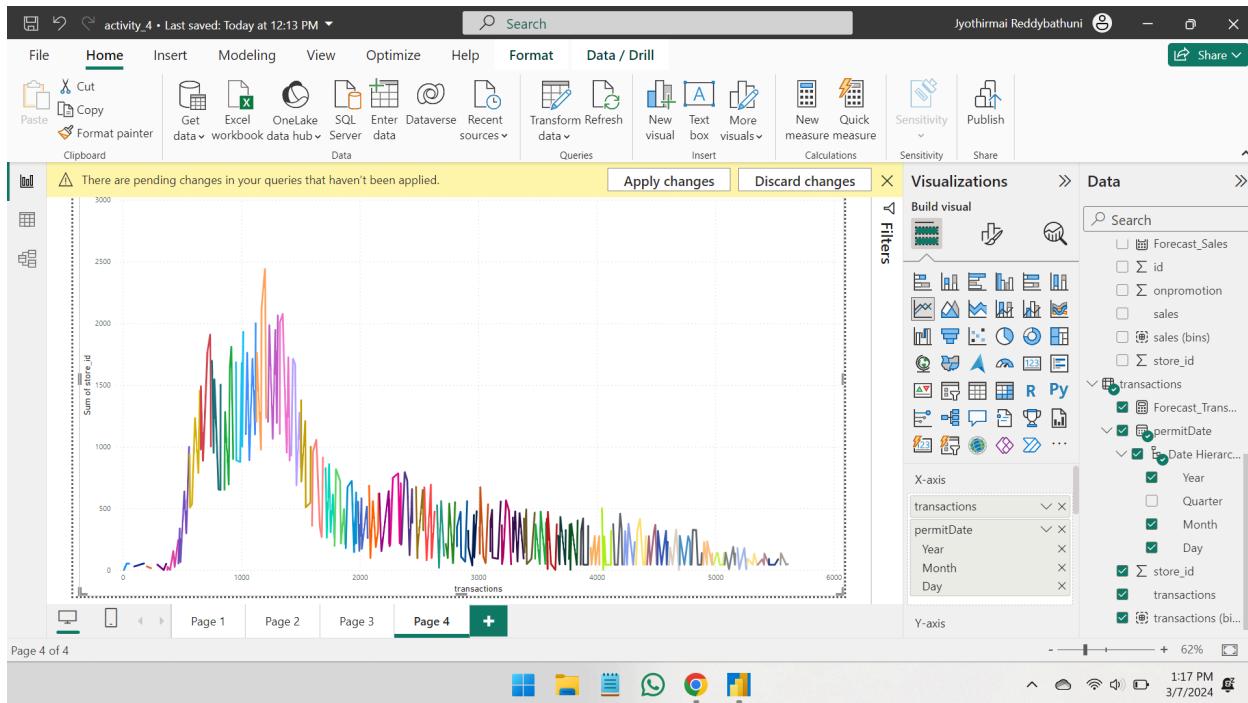


Fig.56. Line graph

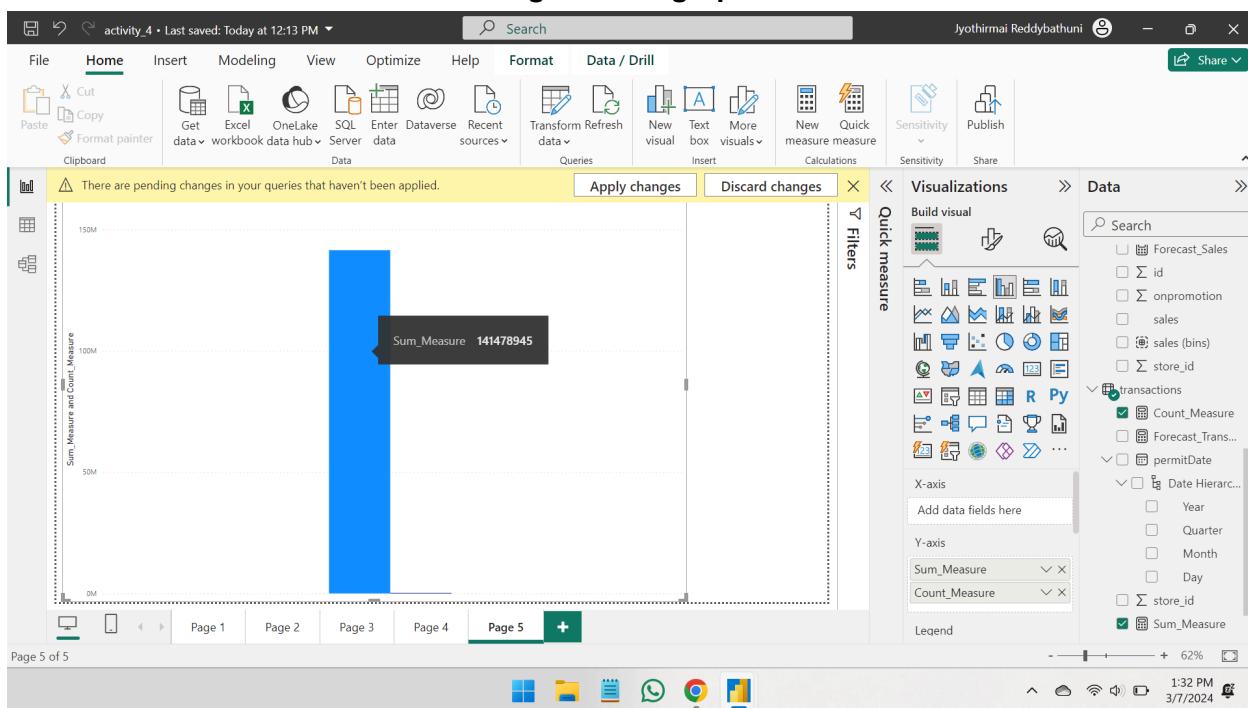


Fig.57. Sum measure

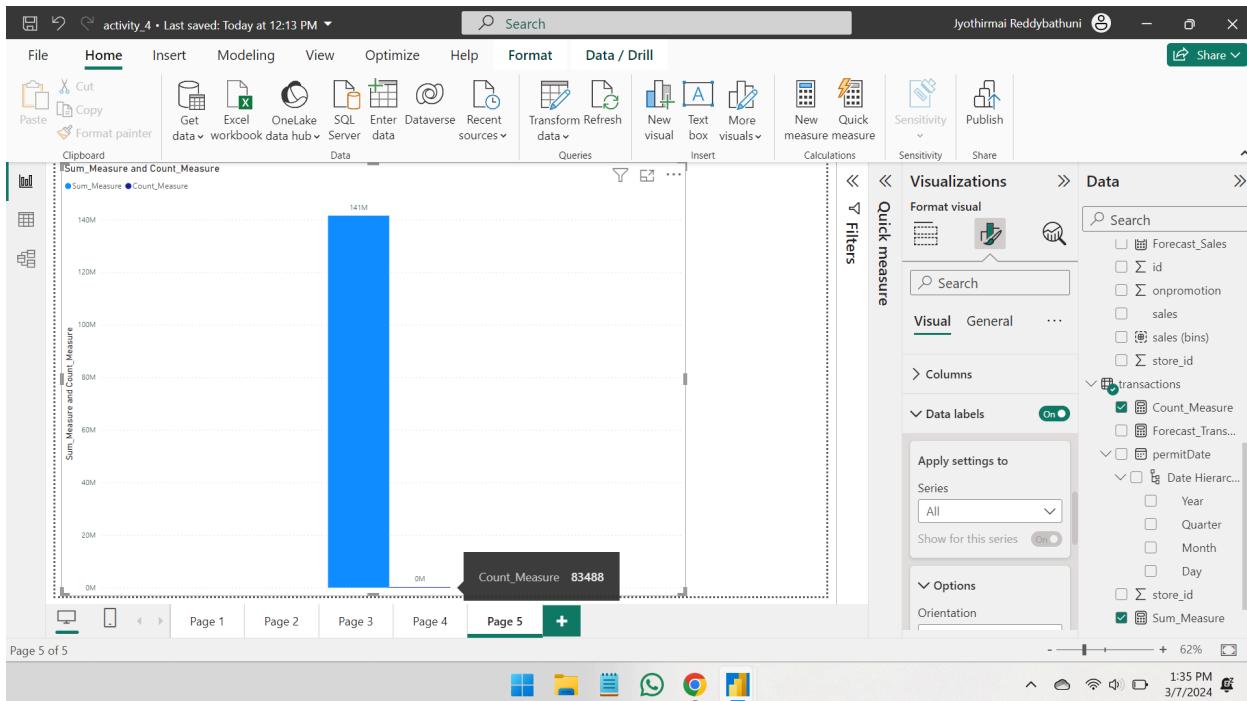


Fig.58. Count measure

Tutorial-5

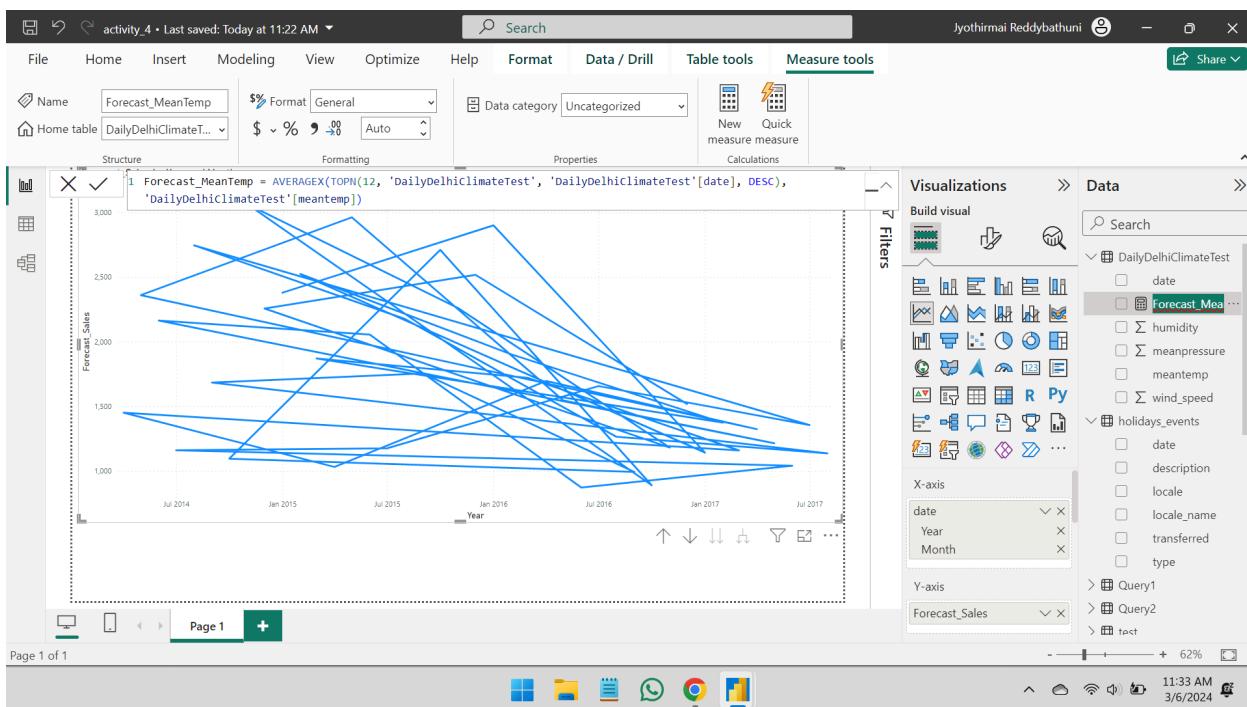


Fig.59. Query Execution

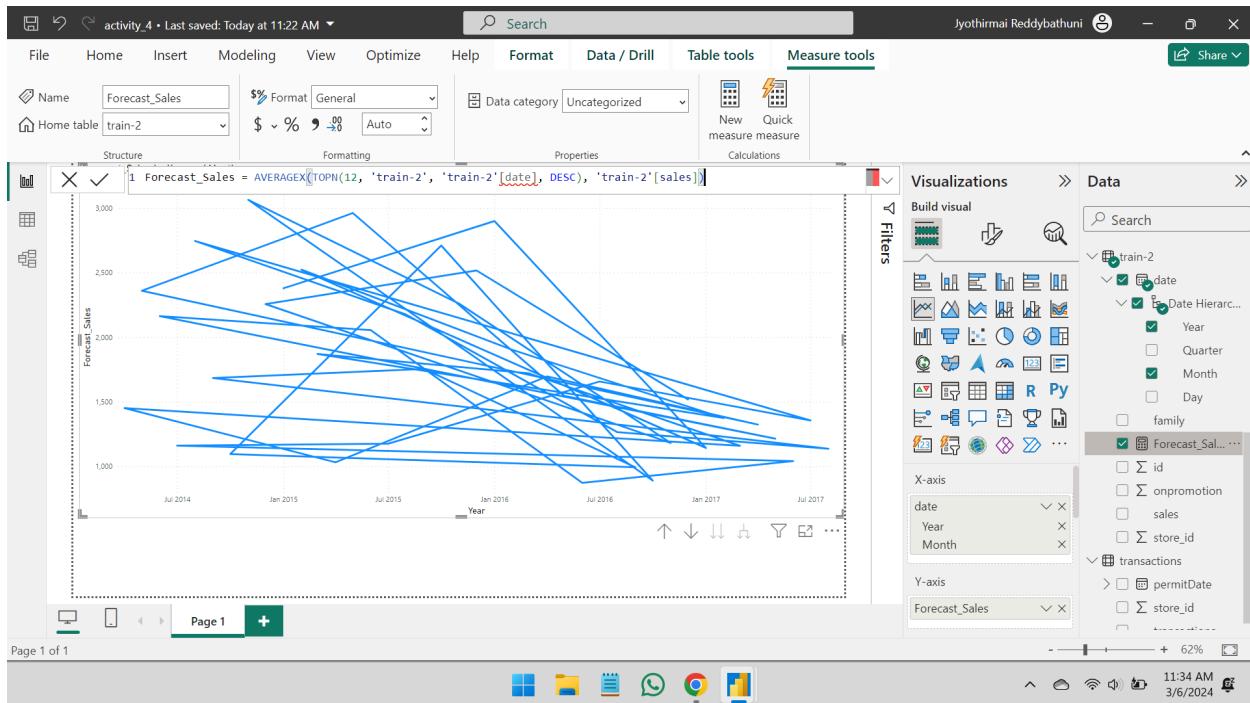


Fig.60. Query Execution

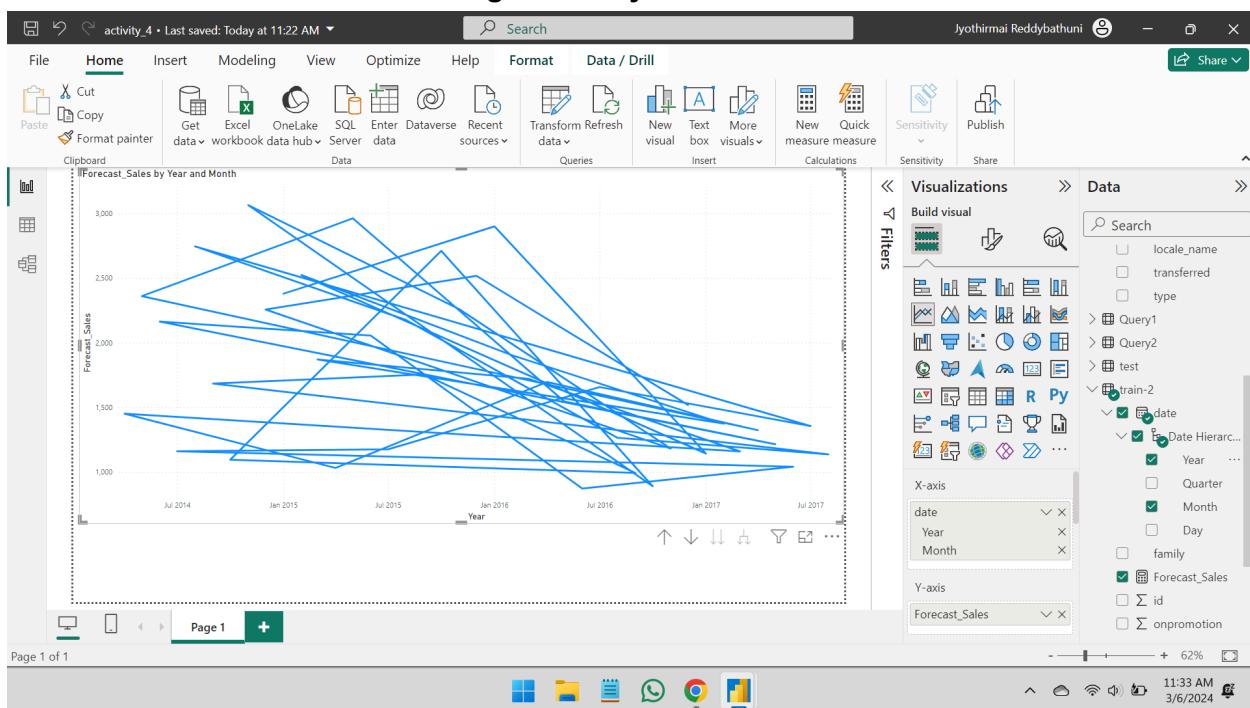


Fig.61. Line Chart

Question-5

The screenshot shows the Power BI desktop interface with the 'Measure tools' tab selected. A new measure is being defined with the following DAX formula:

```
Forecast_Transactions = AVERAGEX(TOPN(12, 'transactions', 'transactions'[permitDate], DESC), transactions[transactions])
```

The 'Formatting' section shows the measure is set to '\$' with a percentage symbol (%) and a decimal separator of ','.

The 'Properties' section shows the 'Data category' is set to 'Uncategorized'.

The 'Calculated Fields' pane on the right lists the measure 'Forecast_Transactions' under the 'transactions' table.

Fig.62. Visualisation 1

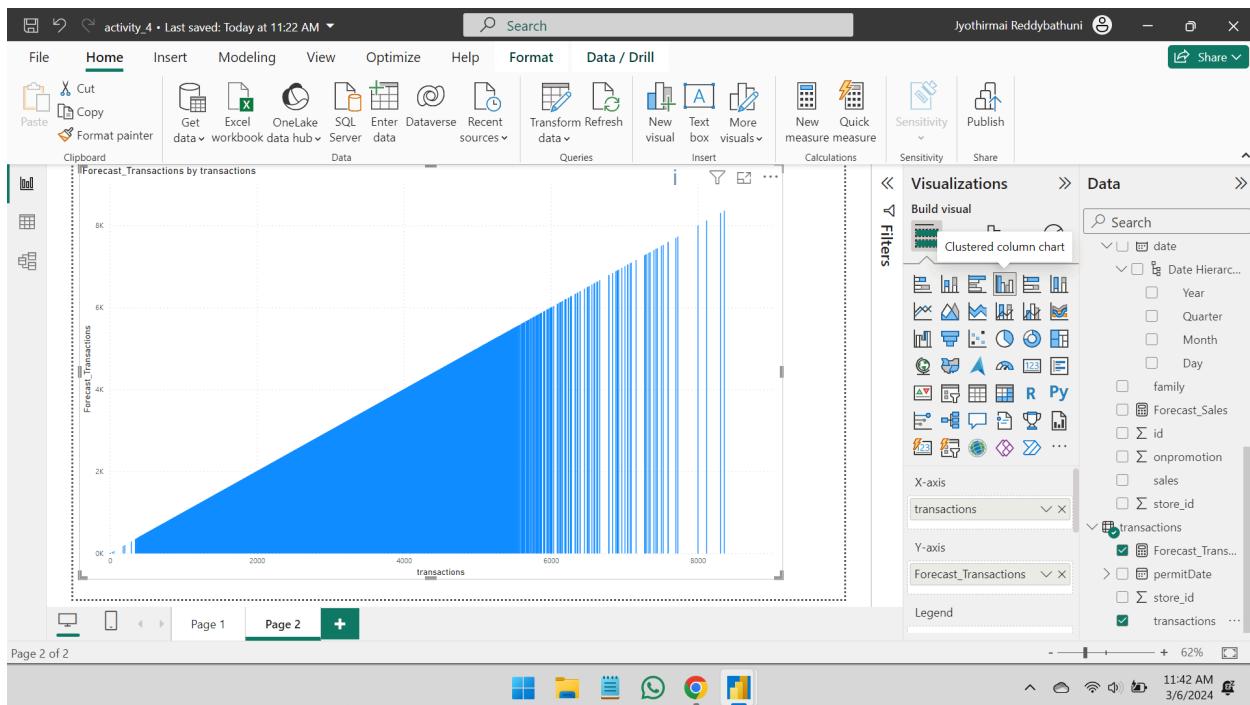


Fig.63. Visualisation 2

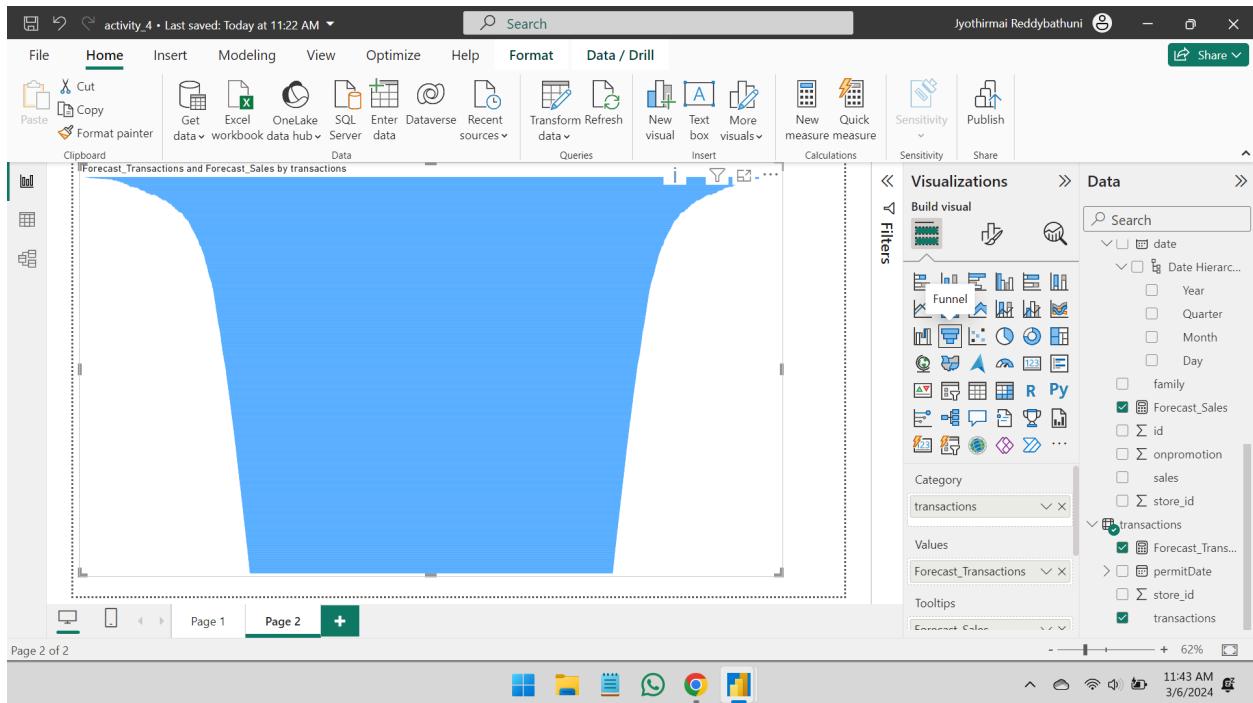


Fig.64.