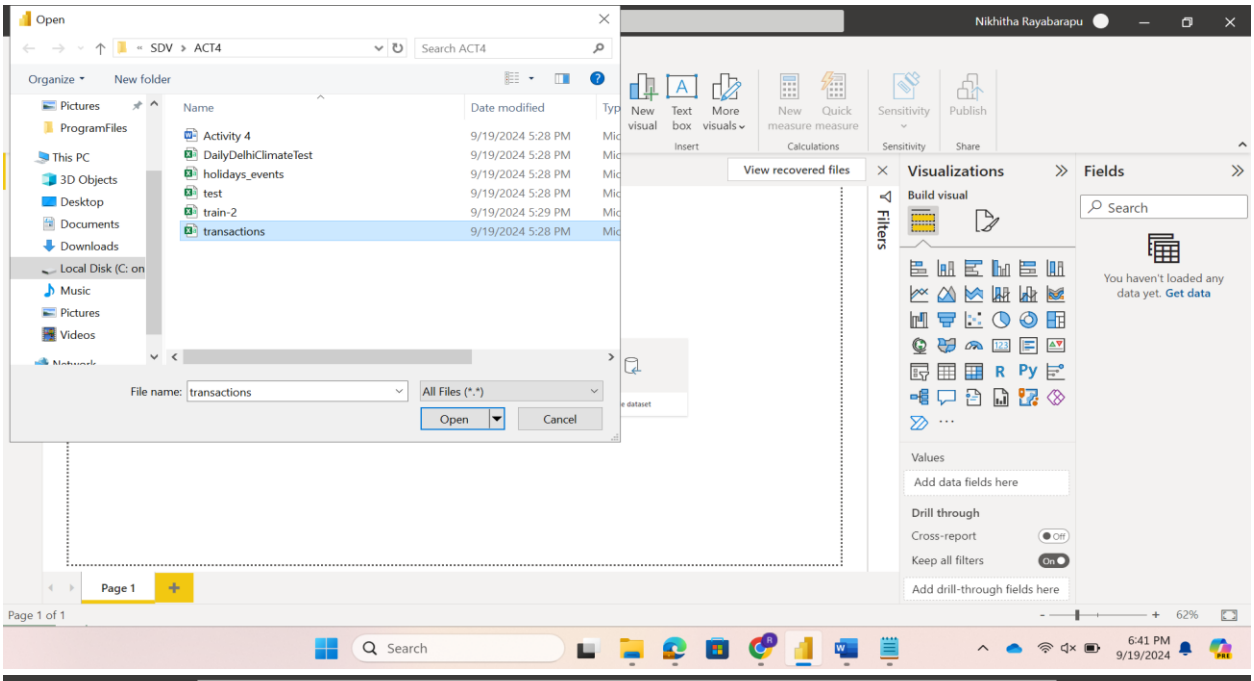
CSCE 5320: SCIENTIFIC DATA VISUALIZATION ACTIVITY-4 TUTORIAL

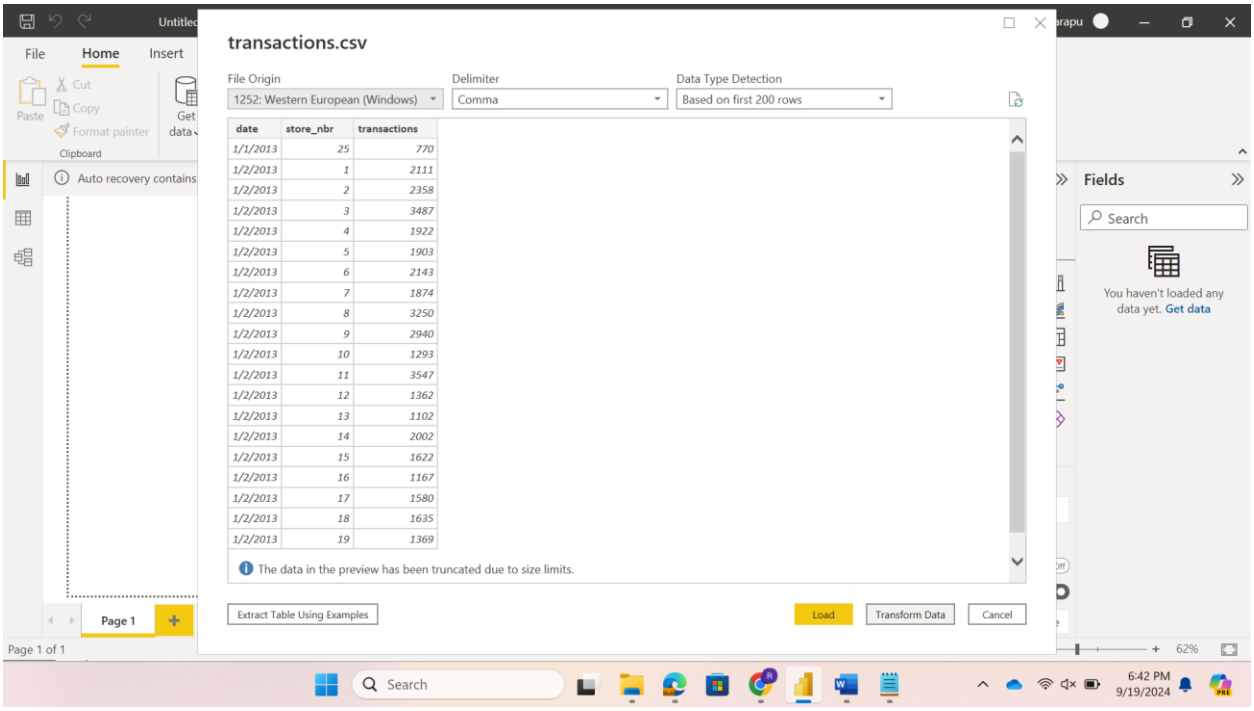
R. Nikhitha

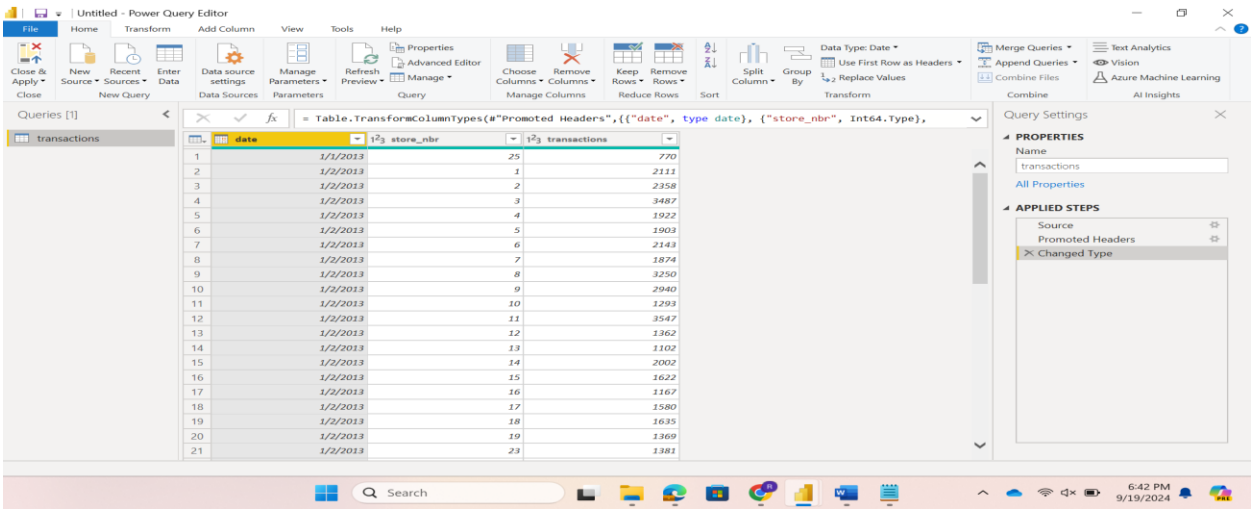
Date: 09/23/2024

Tutorial 1: Power Query(Data Transformation and Group By).

Step 1:- Load the 1st Dataset Named Transactions.csv and click on Transform Data.

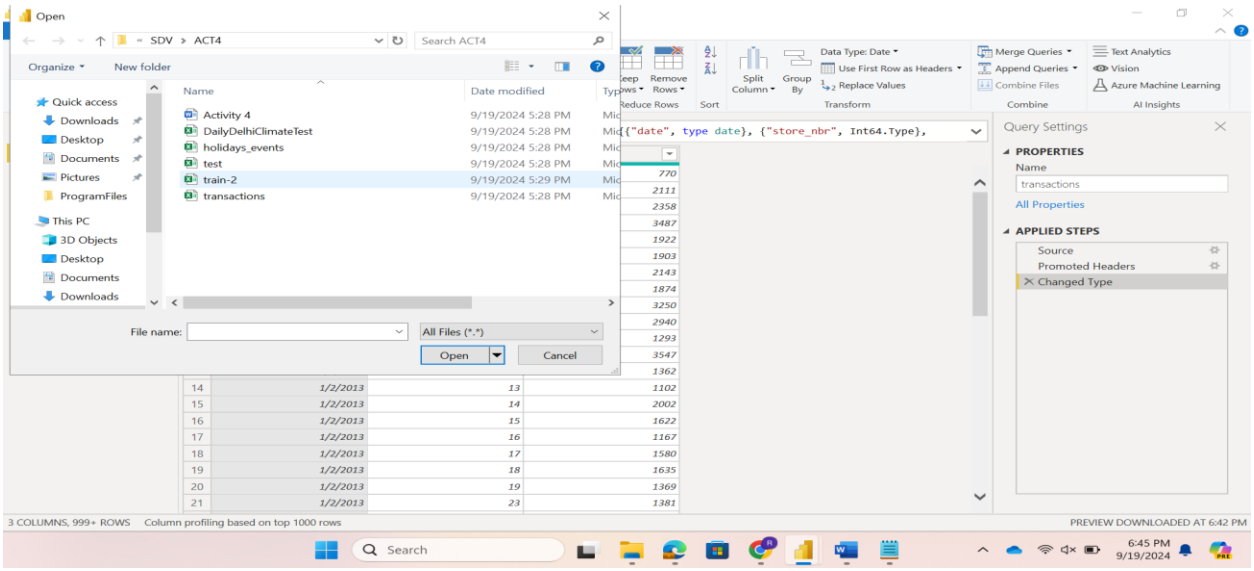




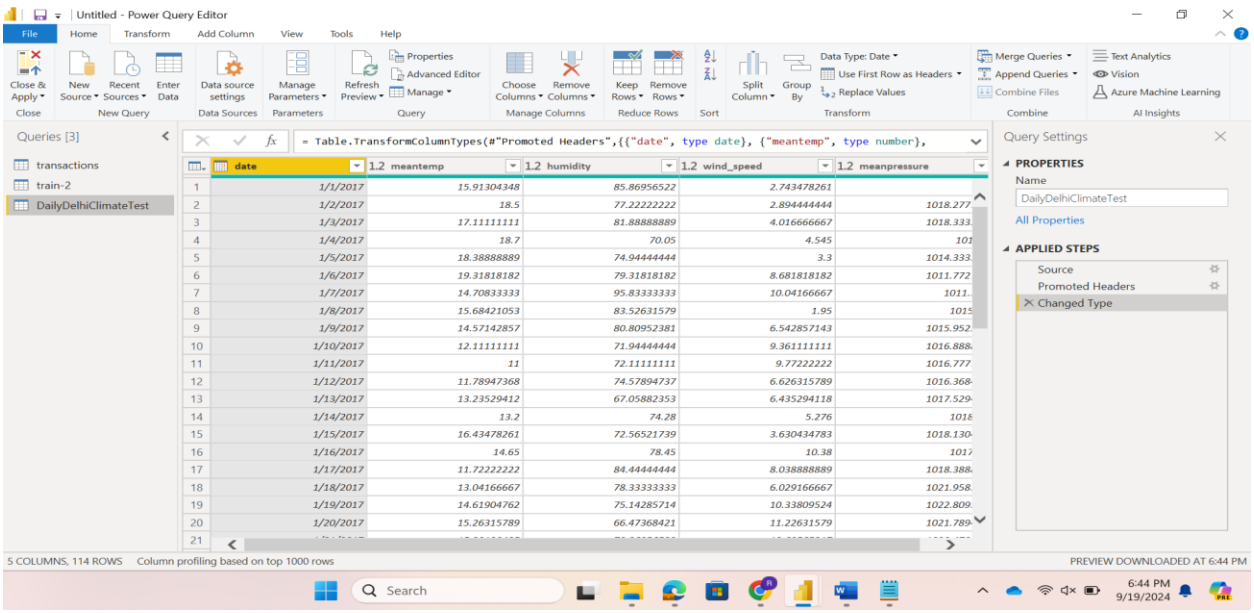


Step 2:- Once you click on Transform Data it takes you to the Power Query Editor where you can make changes to the dataset.

Now, Click on New Source Located at the Top and load the rest of the Datasets.



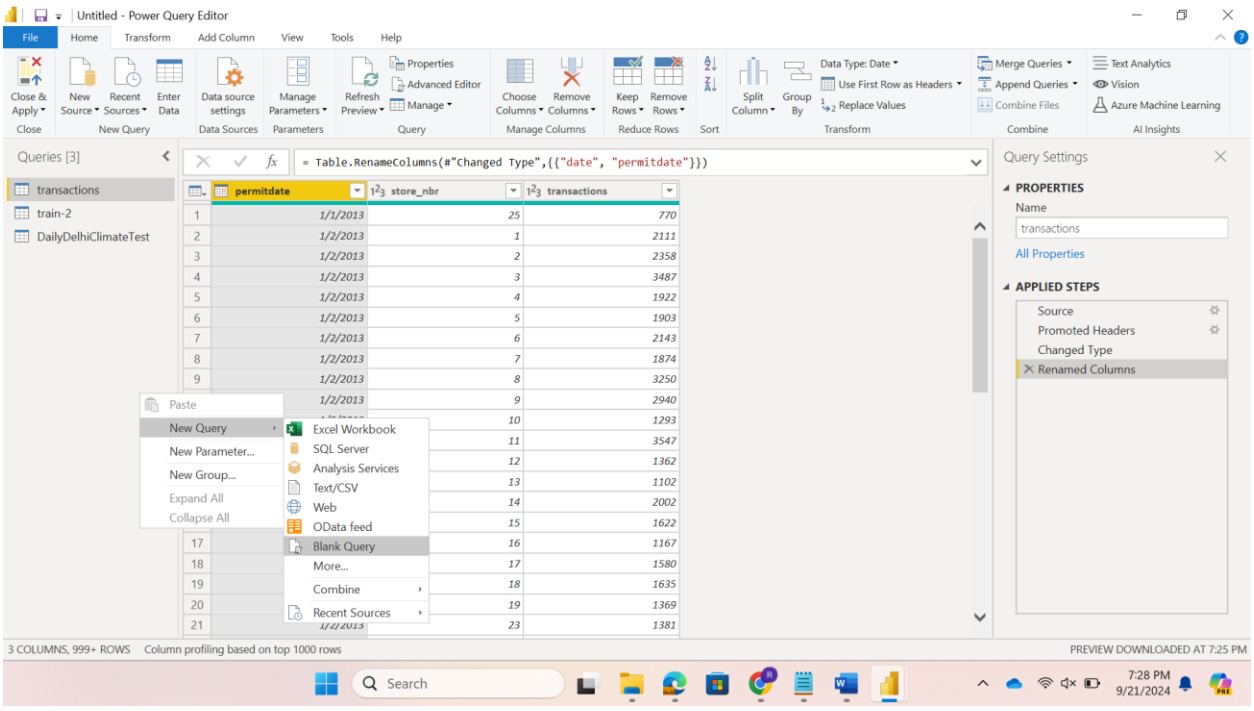
Step 3:- I Have loaded 3 Datasets or Tables in total where there is no common attributes except the Date Column in the DailyDelhiClimateTest.csv



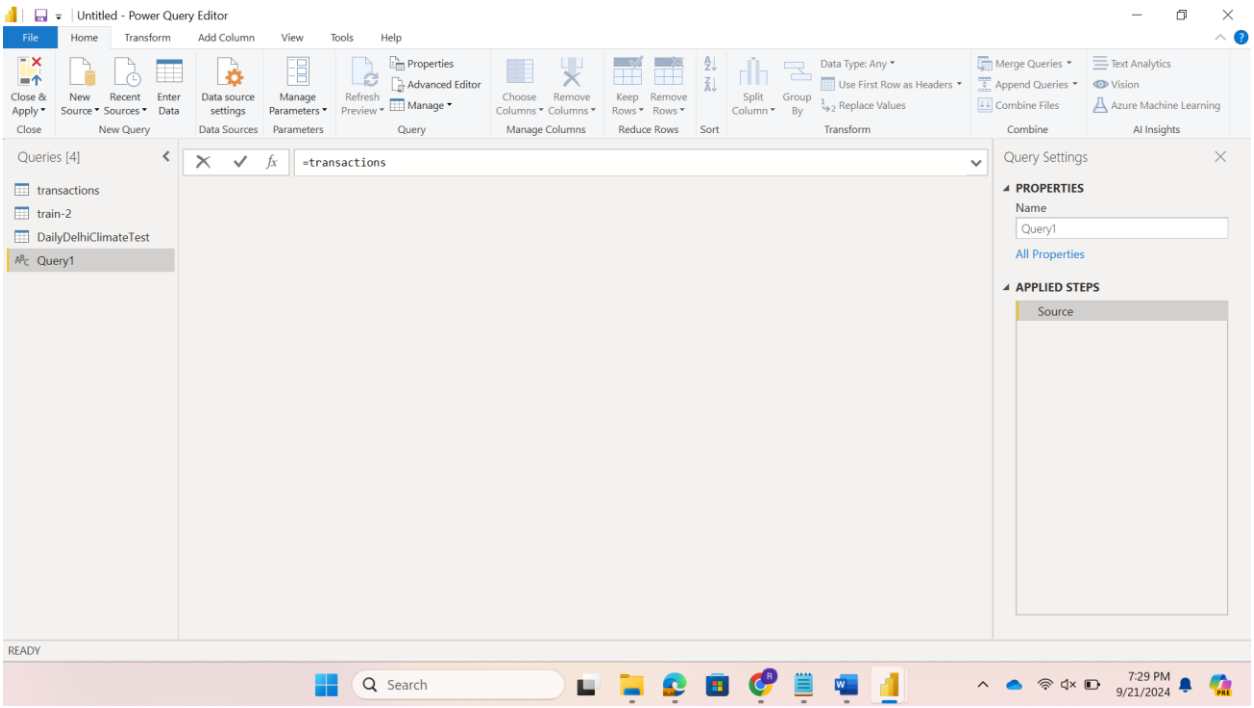
Step 4:- Now I want to get the columns min and max values. You can do that by using a query.

Right click and click on New Query -> Blank Query.

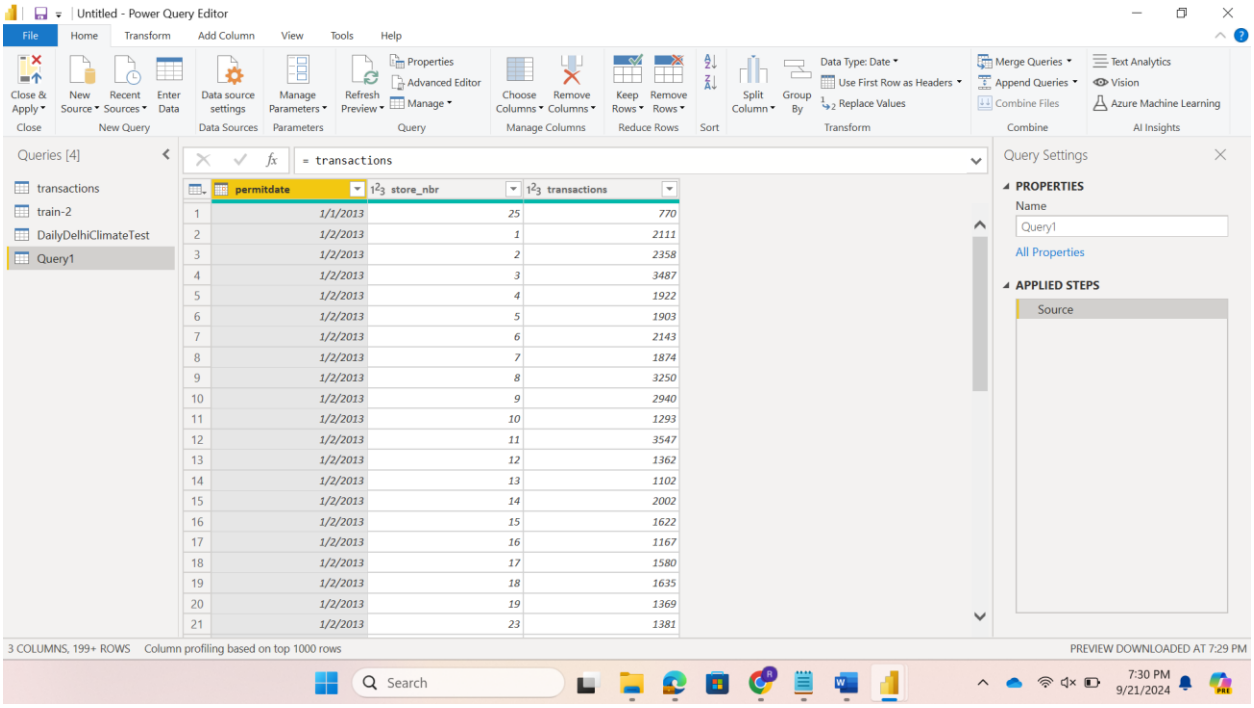
I have renamed the Date to permitdate. You can also rename the attributes name just by clicking on the attribute and renaming it with the other name.



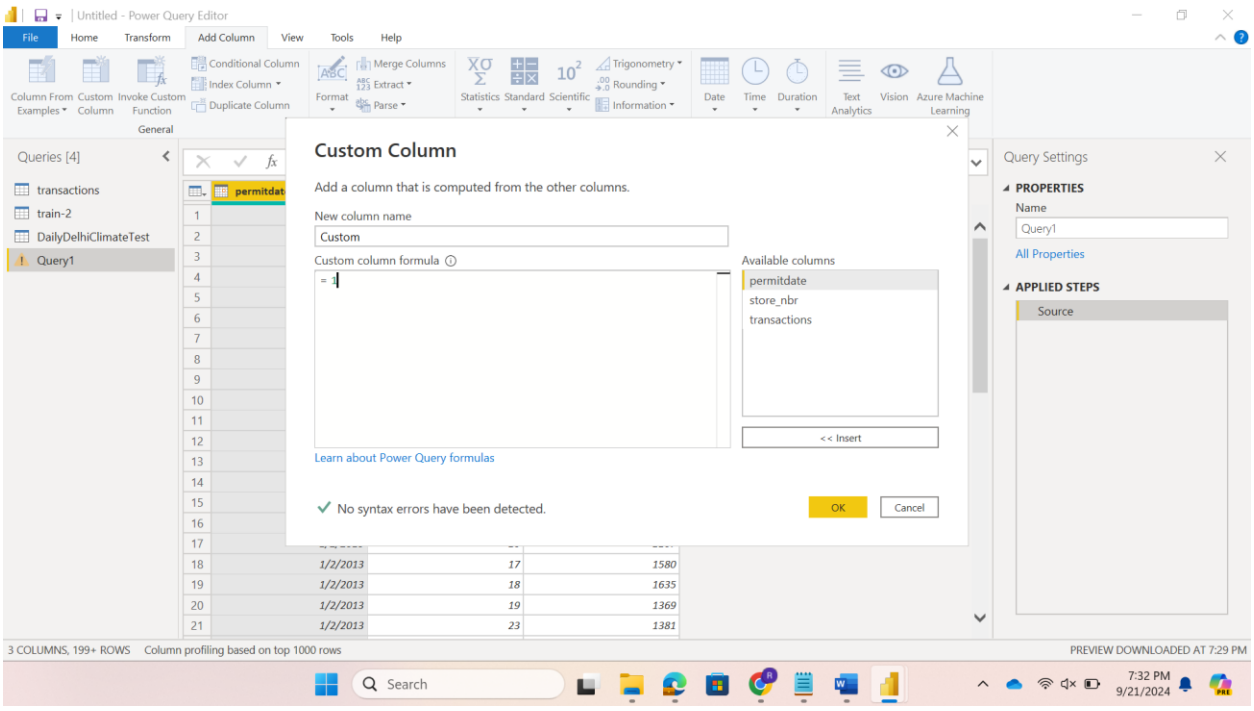
Step 5:- You can observe a new Query1 added to the Queries tab where I have selected the transactions table to perform modifications.

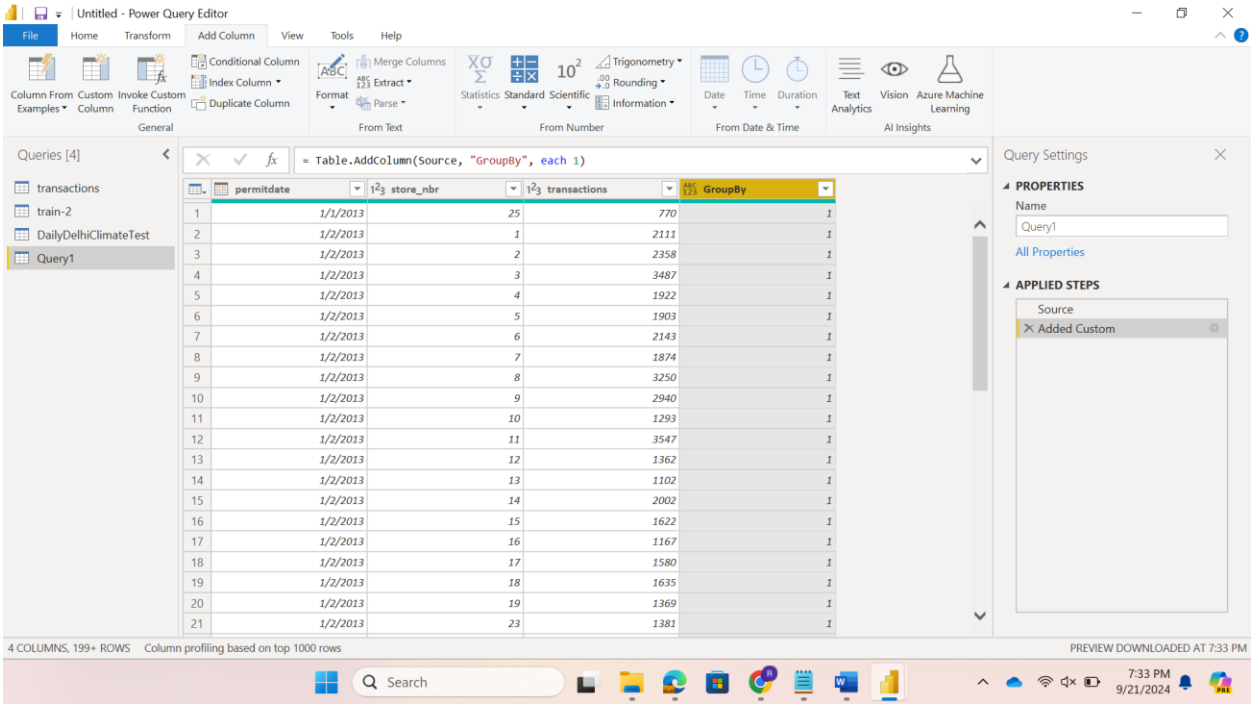


Step 6:- The “=transactions” display the Transactions table. You will be able to find the Applied steps at the right side under Query settings where the steps performed in that table will be displayed. You can click on the X mark on that step if you want to remove that step from the changes made.

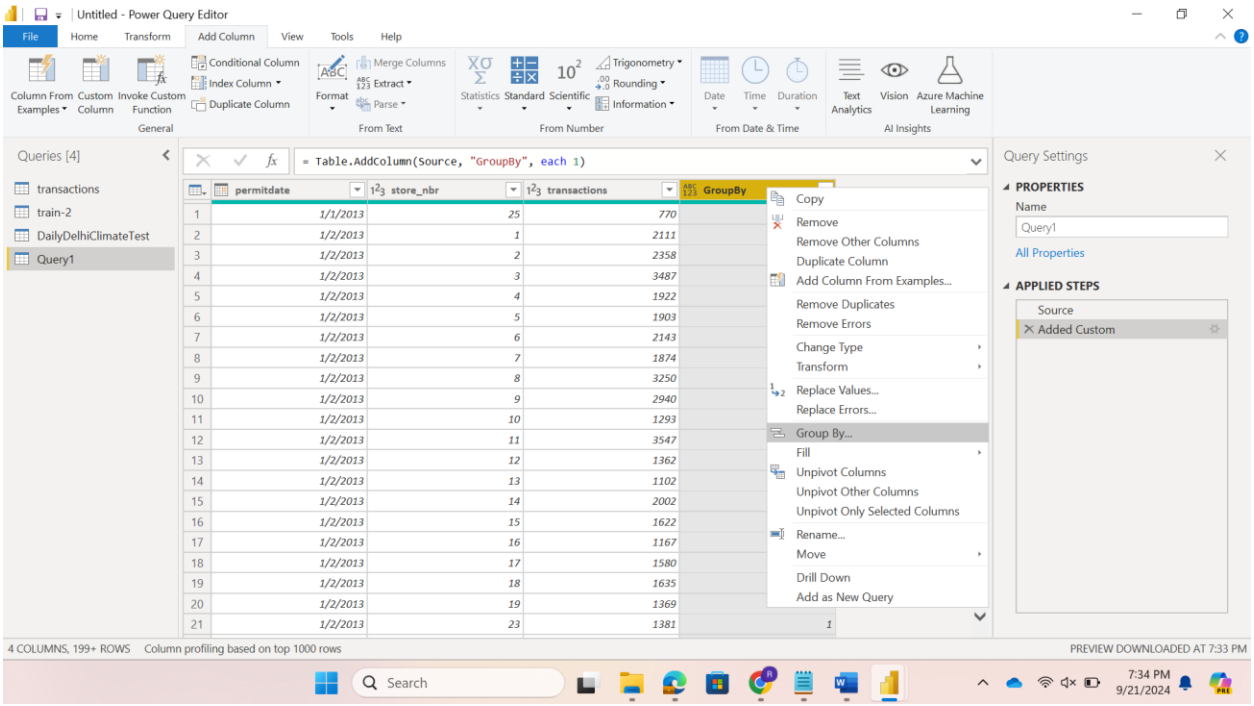


Step 7:- In add column, Click on Custom column to create a new column. But our main purpose is to perform Group by. The cells will be inserted with a 1.

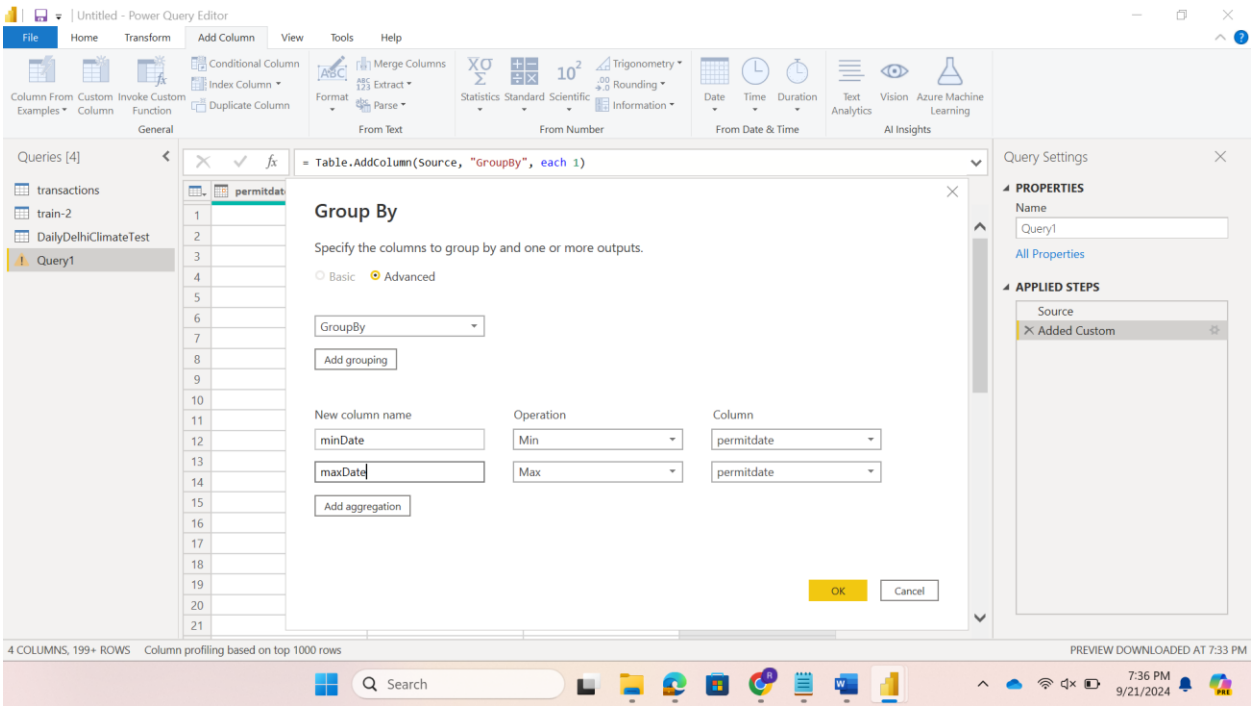




Step 8:- Right click on the attribute and click on group By.



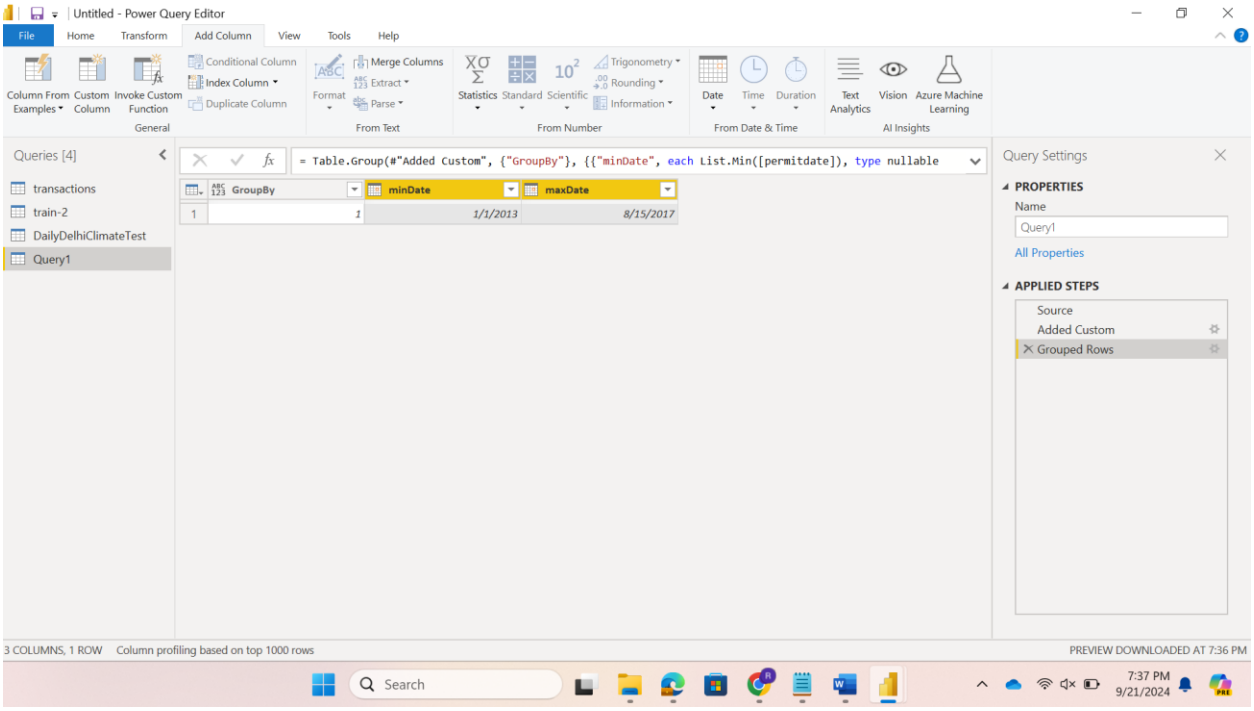
Step 9:- Go the advanced options and type the same min and max dates as shown in the fig below and select the attribute for which you want the group by function to be applied.



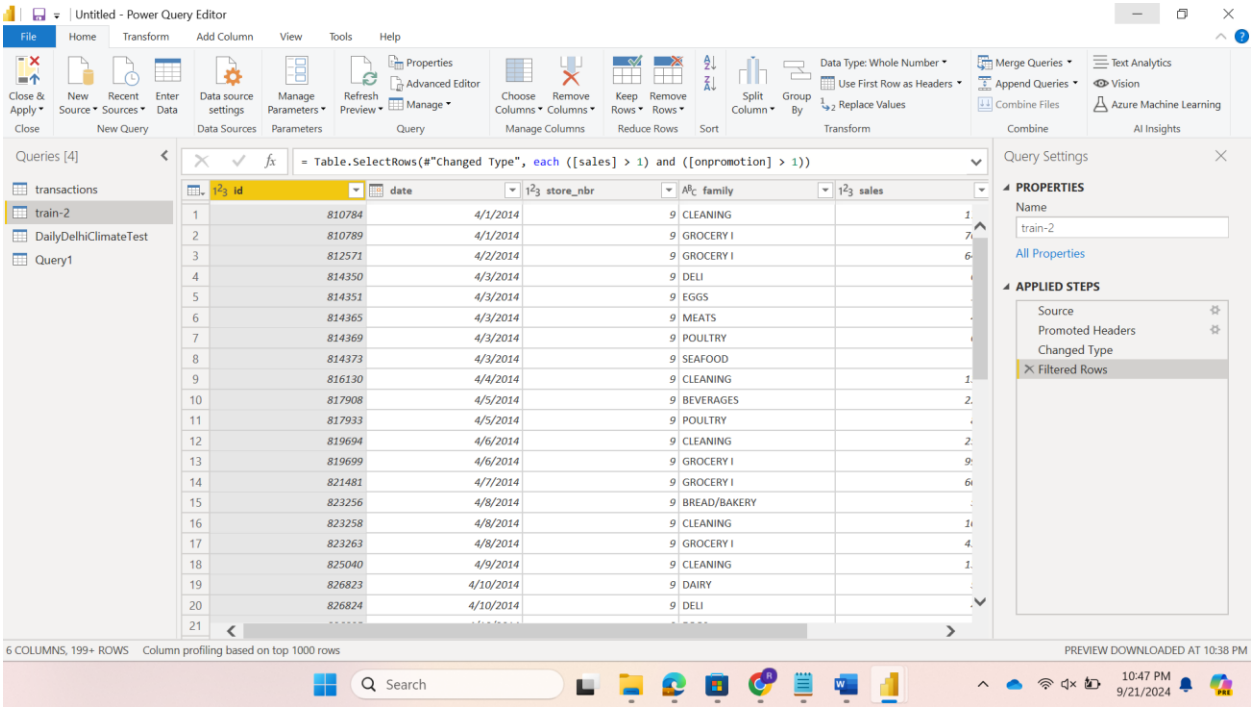
Step 10:- Now the min date and the max date are displayed. From 1/1/2013 and 8/15/2017.

You can even round the dates to the nearest day, month, and year. For ex:- 8/15/2017 can be rounded until 8/31/2017.

As my dates are already in order I don't want to sort them again based on the years starting from 2013 to 2017.

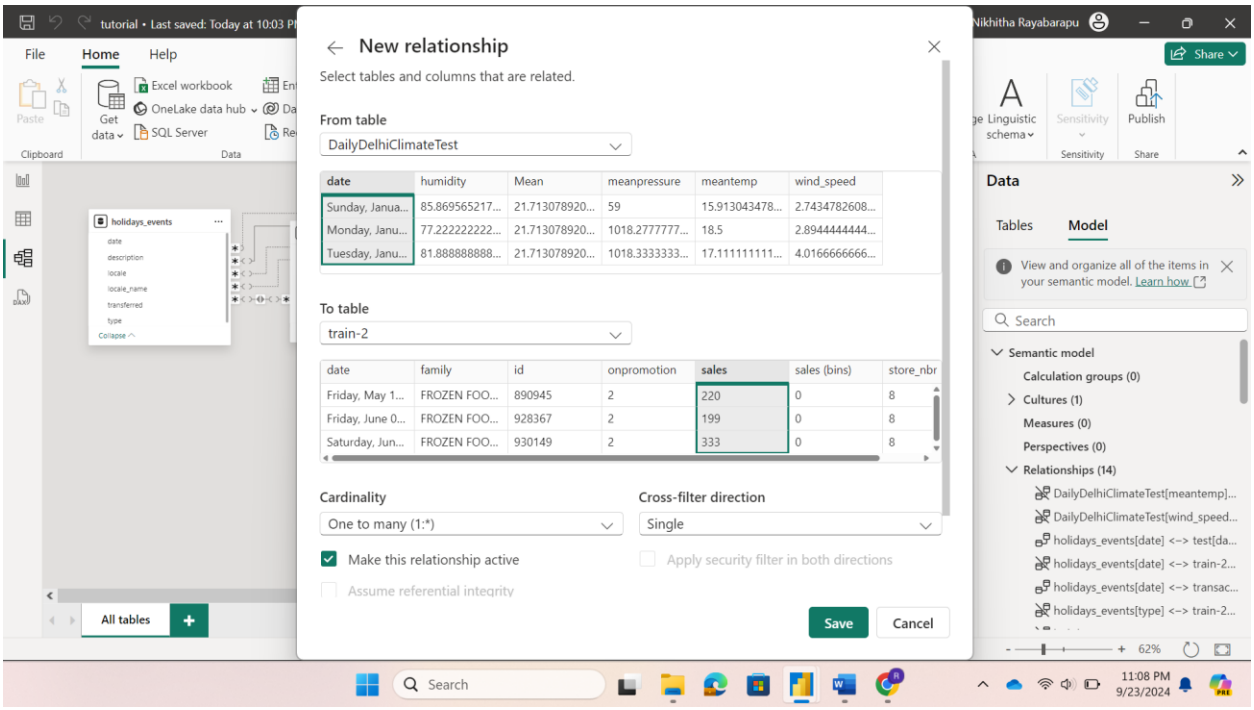


Step 11:- Change the query where the sales and onpromotion is greater than 1.

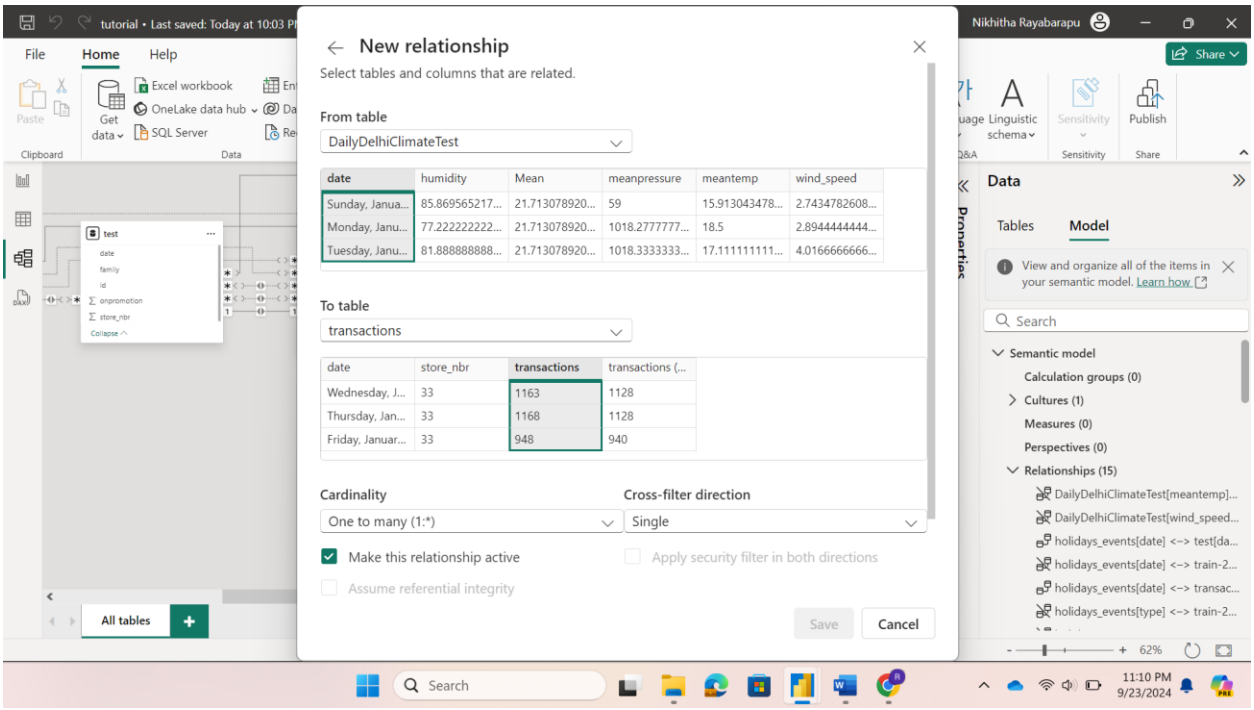


Tutorial 2: Creating Relationships

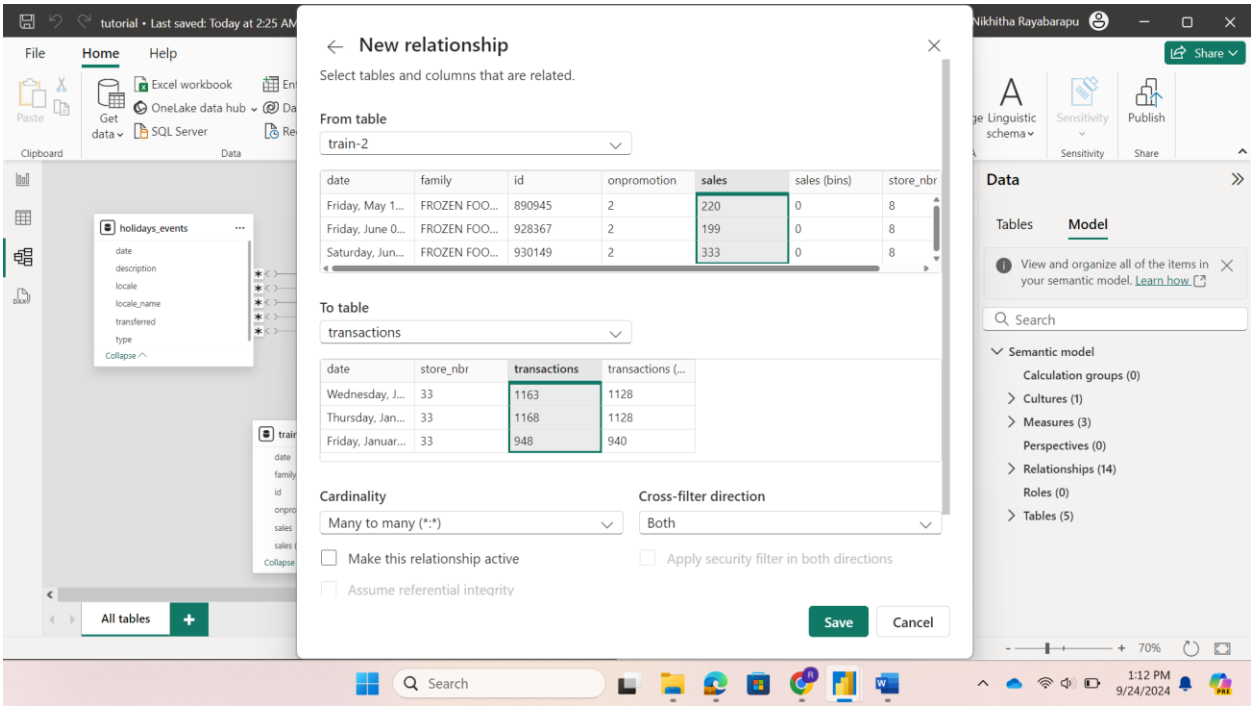
Step 1:- Go to the model view and right click on any table -> Create Relationship. Create all the possible and necessary relations which can be one to one, one to many, etc.....



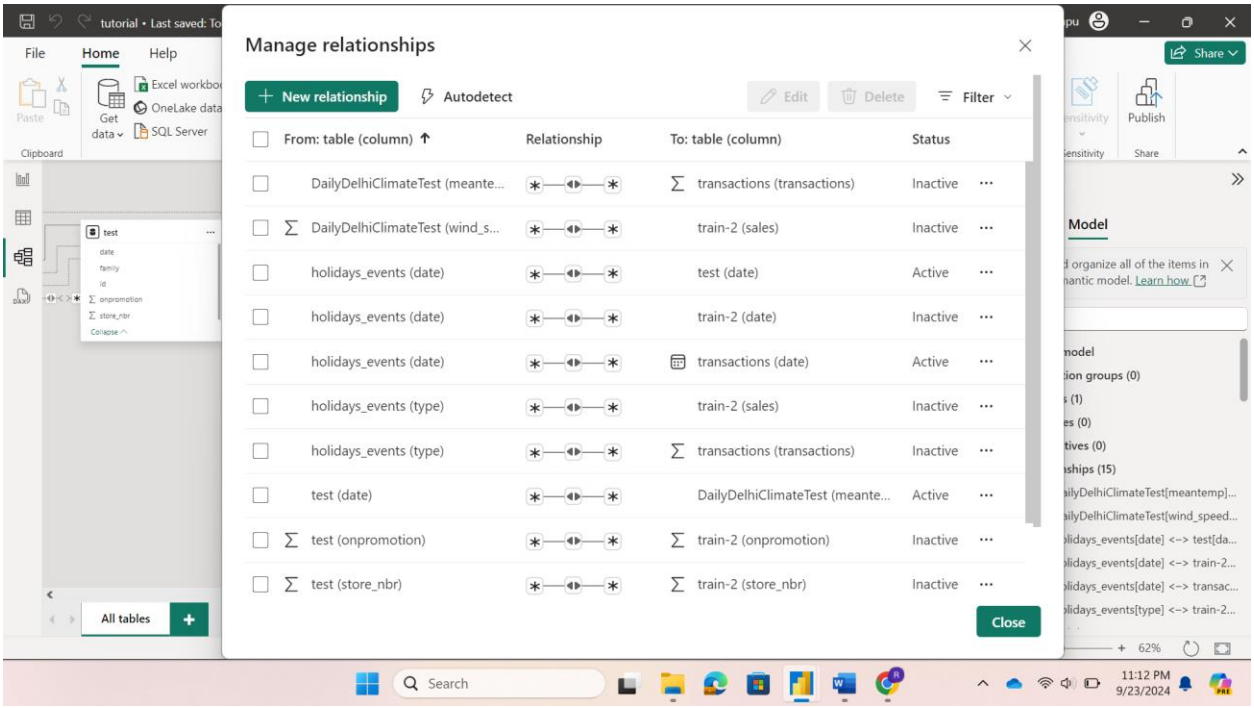
Step 2:- I am creating a connection between DailyDelhiClimateTest table and transactions table.



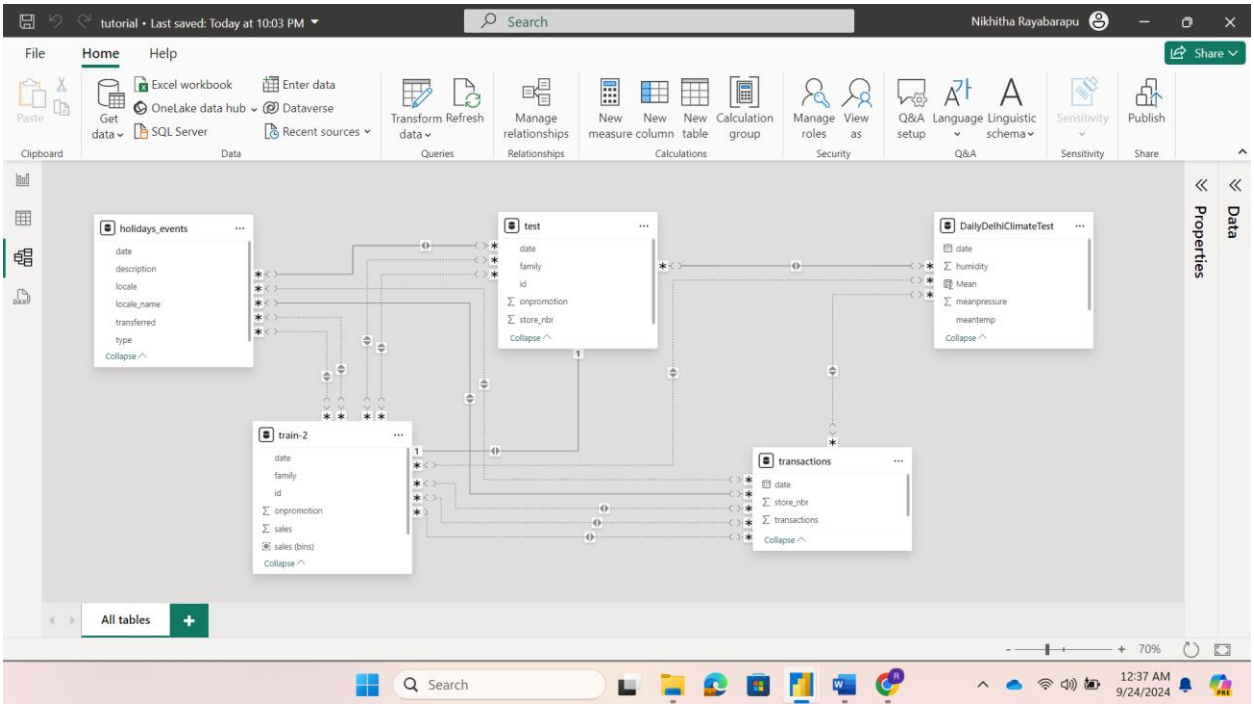
Step 3:- We are creating a relationship between train and transaction table. I have selected 1 attribute from each table to link the tables. The cardinality is automatically populated based on the data and attributes selected.



Step 4:- The created relationships can be viewed using the manage relationships where the active relationships give us that a connection between a table is active. Only a single connection between tables can be active.

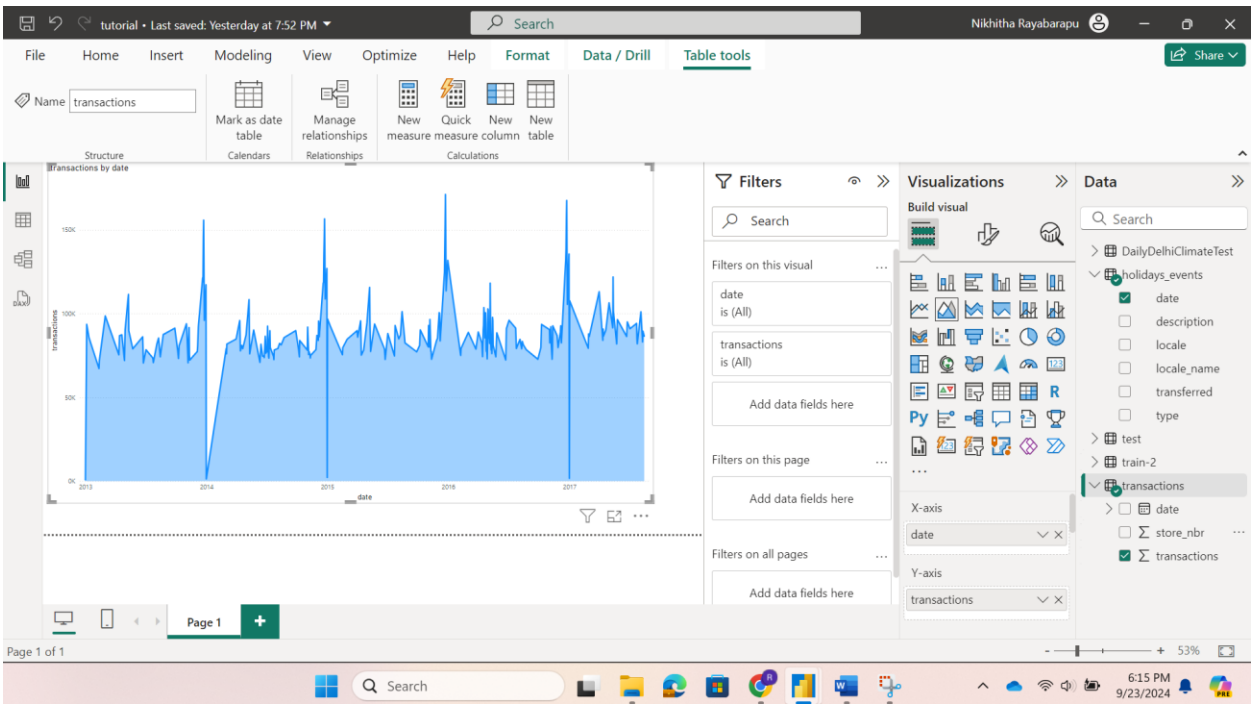


Step 5:- The model view of the tables can be viewed after the relationships between the tables are established.

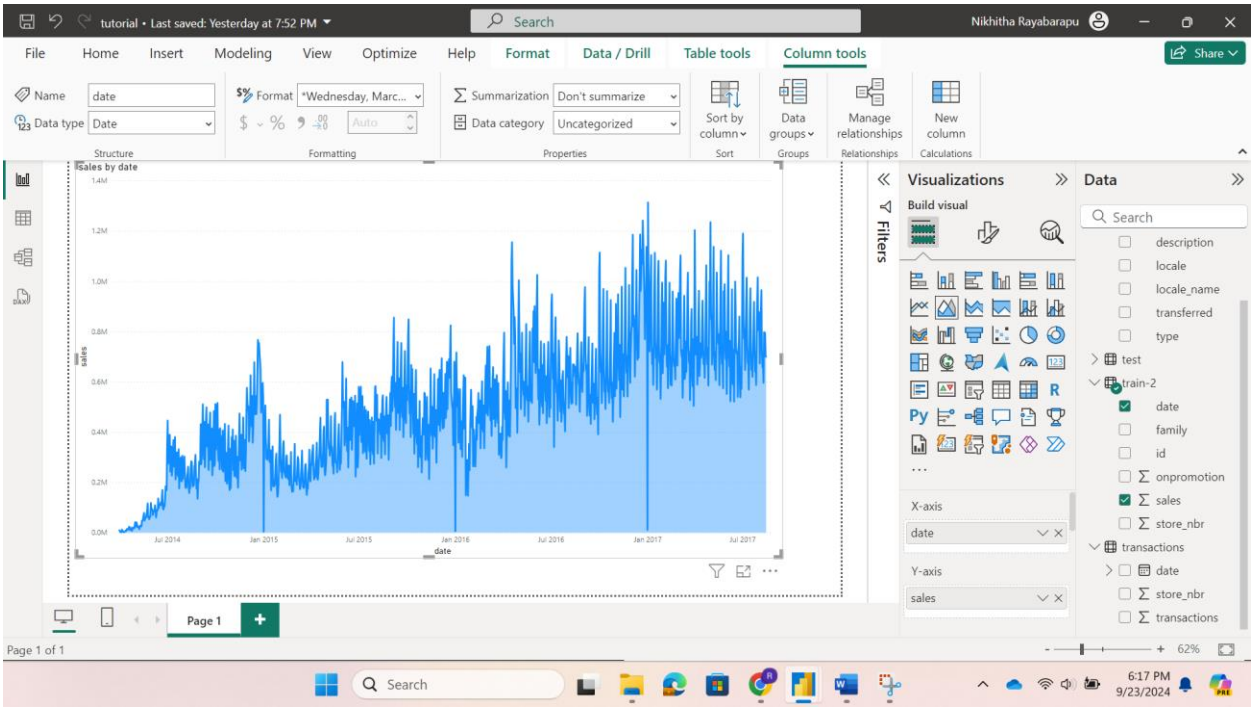


Tutorial 3: Visualizing using different tables

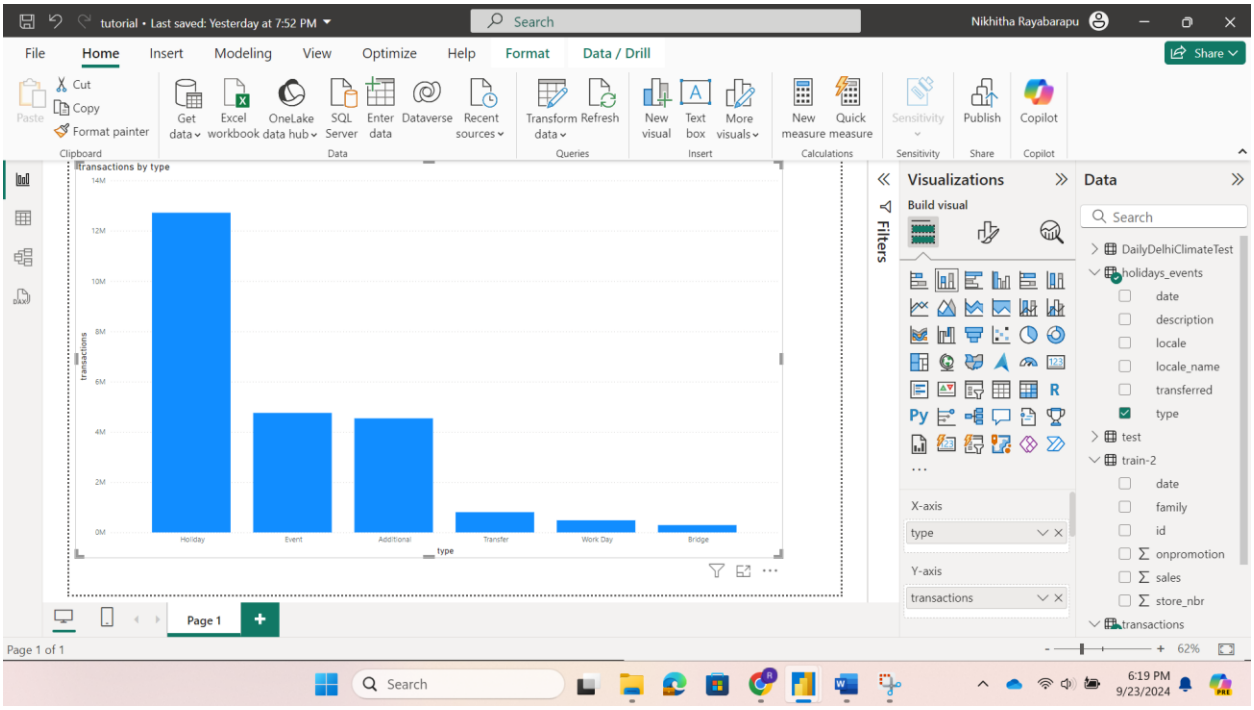
Step 1:- Select the date from the holidays_events table and transactions from the transaction table and click on the Area chart to visualize the sum of transactions with respect to years.



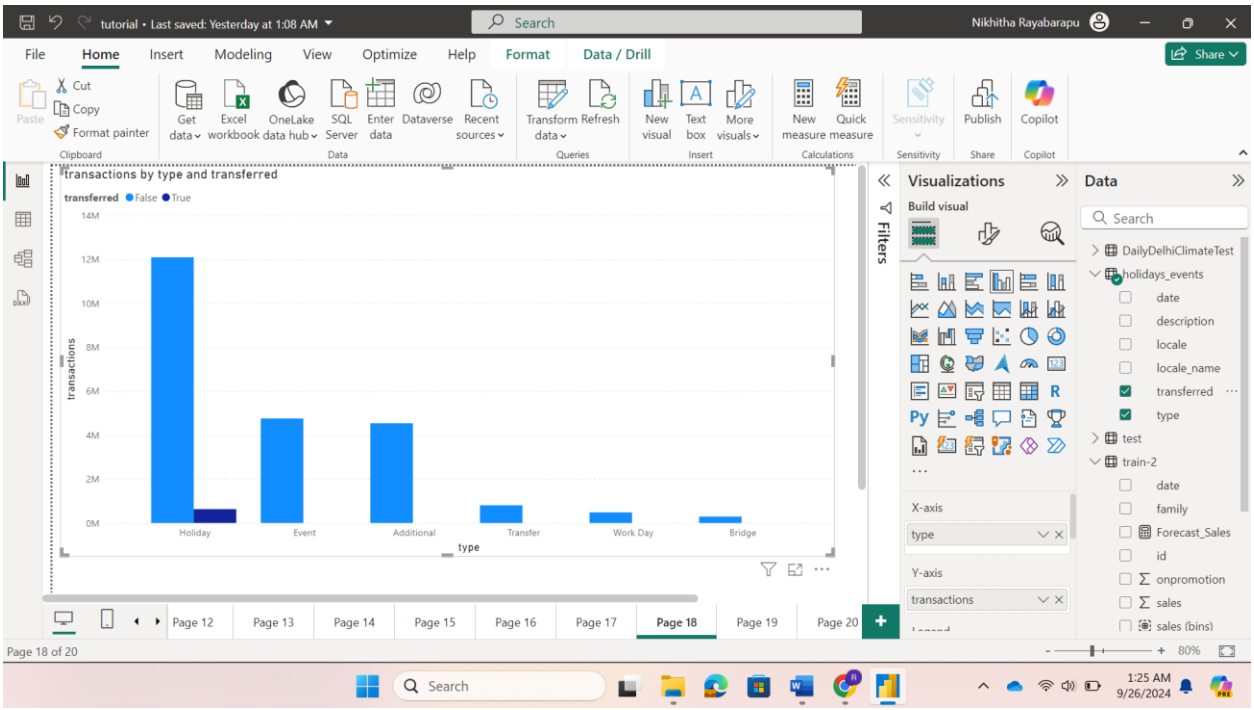
Step 2:- Now I am selecting the sales and date from the same train table and use the area chart to visualize the sales across years.



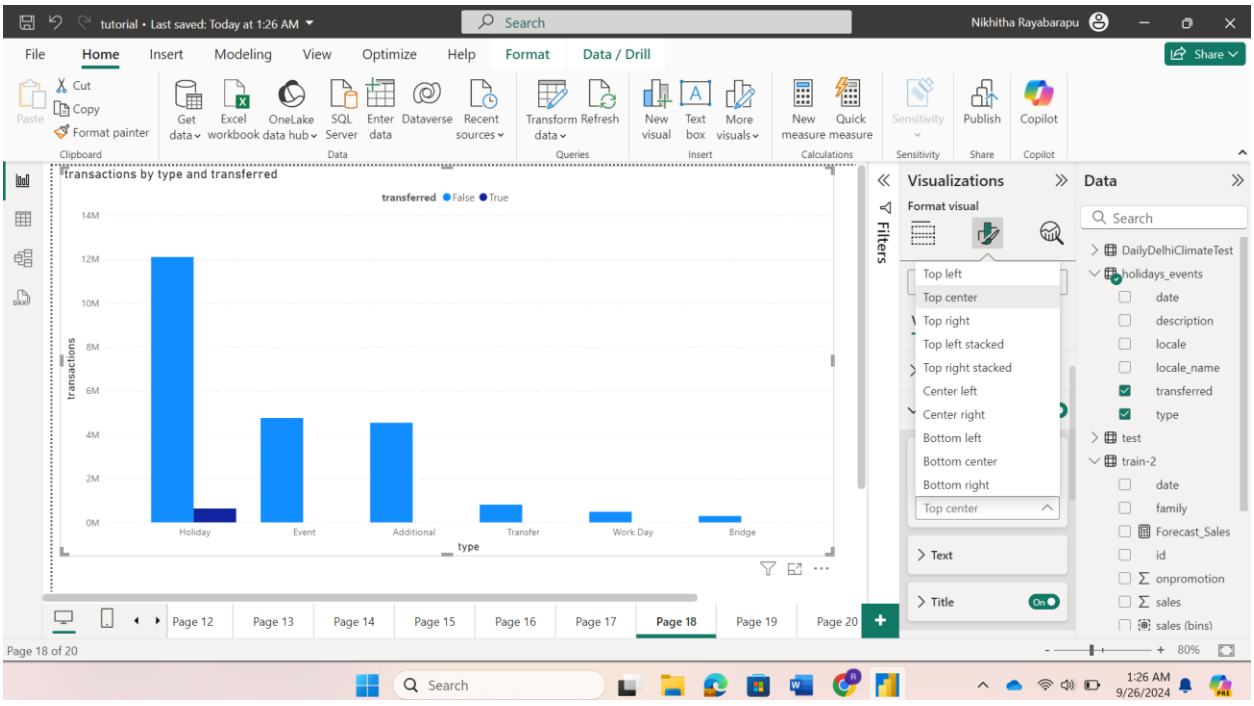
Step 3:- Select the Type from holidays_events and transactions and select the visualization.



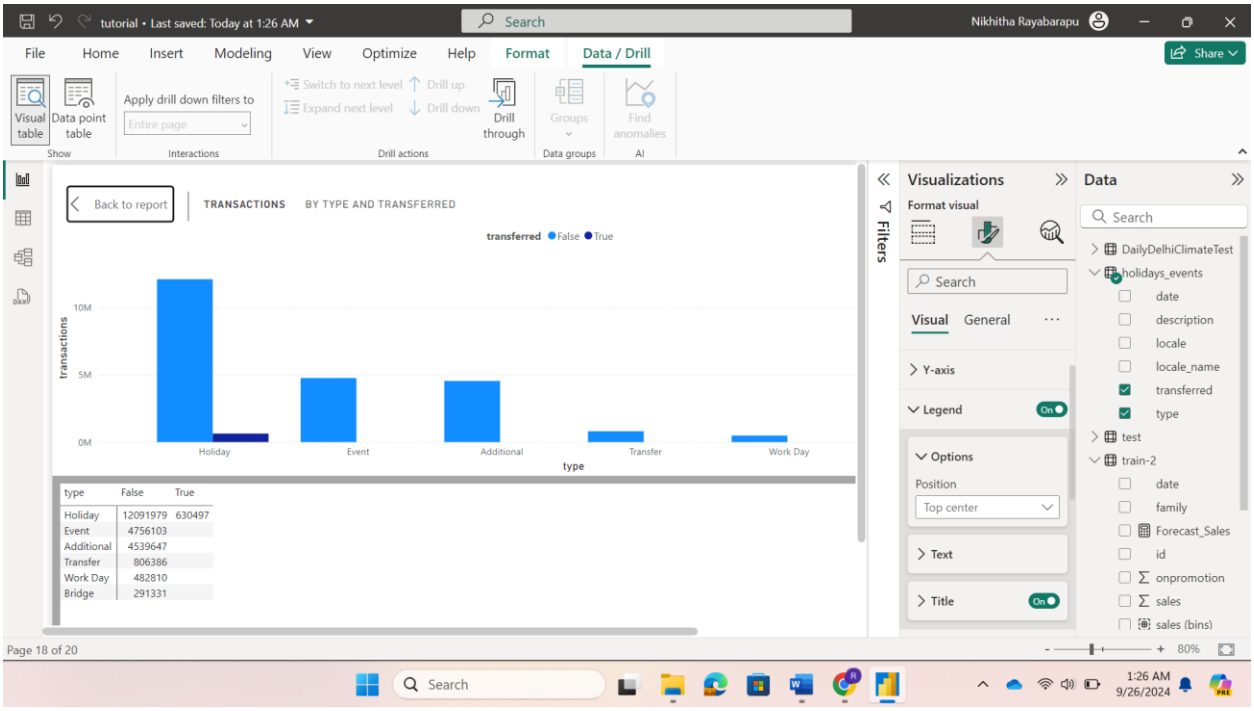
Step 4:- Now add another attribute i.e. Transferred



Step 5:- place the legend at the top center

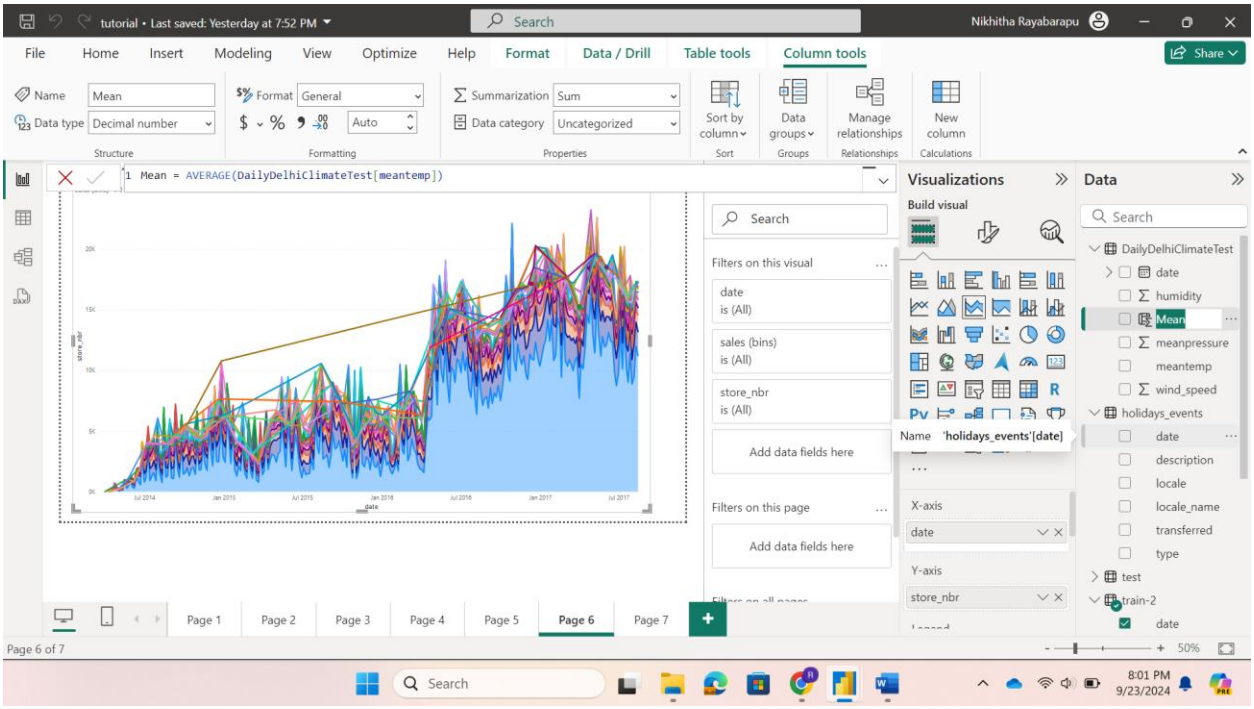


Step 6:- Click on visual table in Data/Drill to get the information about the graph in tabular format.

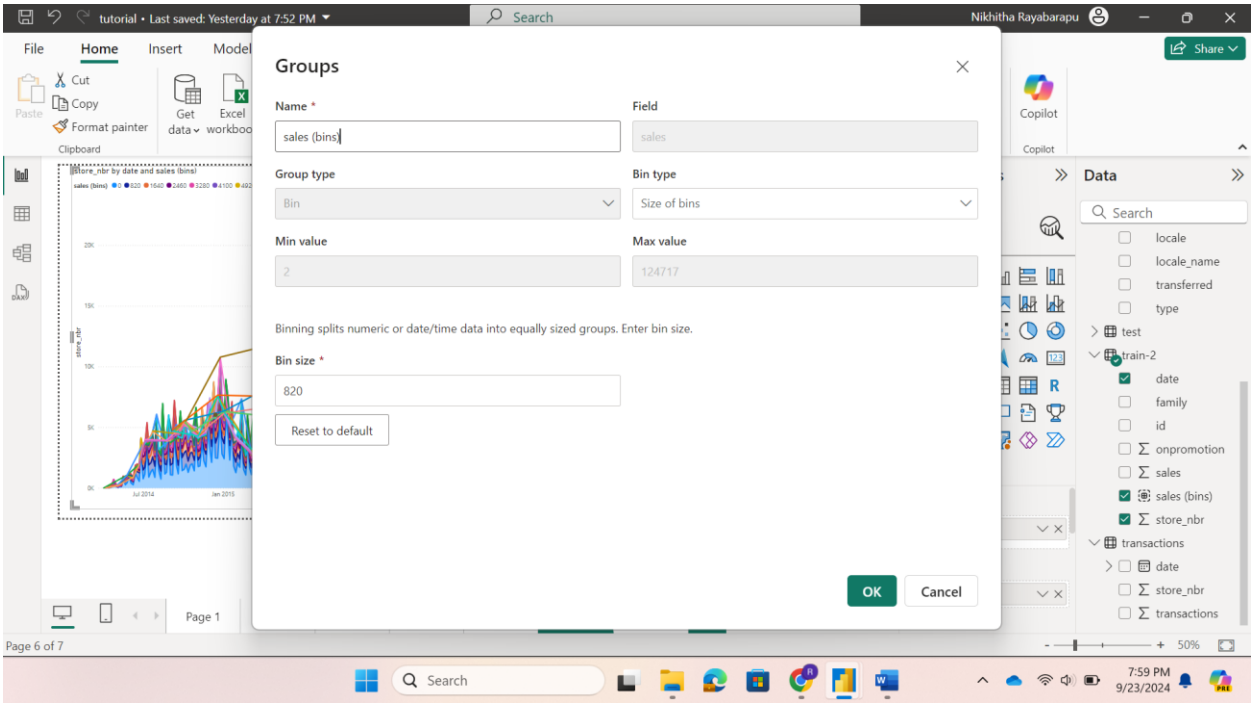


Tutorial 4: Creating an attribute to group the attributes

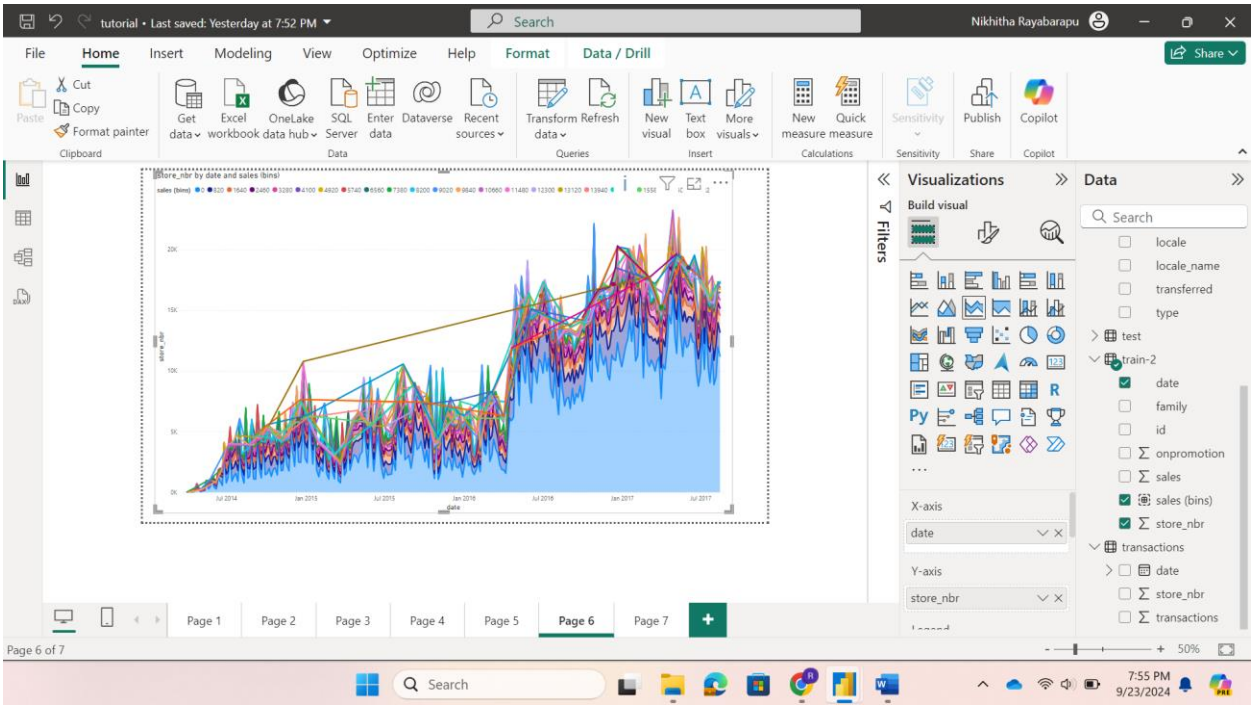
Step 1: Click on Custom Column and give the below query to get the mean. If you want to insert the attributes then just click on the attributes and click on insert to use them in the query.



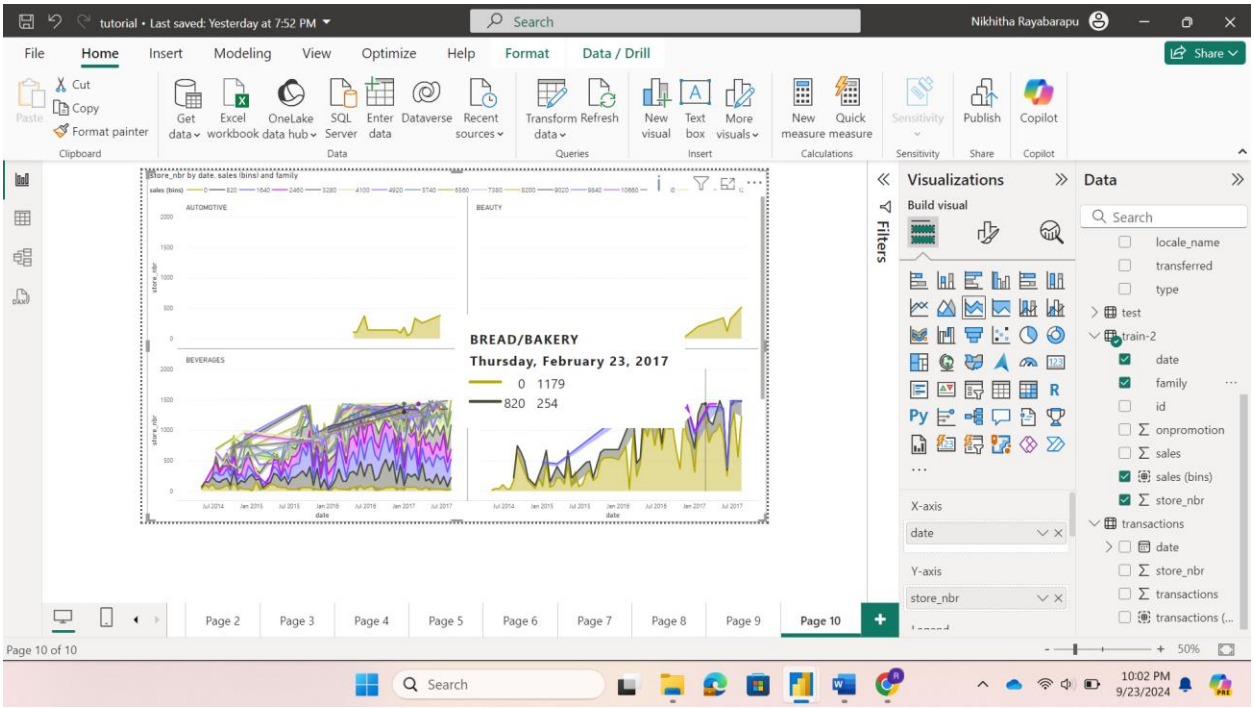
Step 2:- Grouping based on a sales attribute.



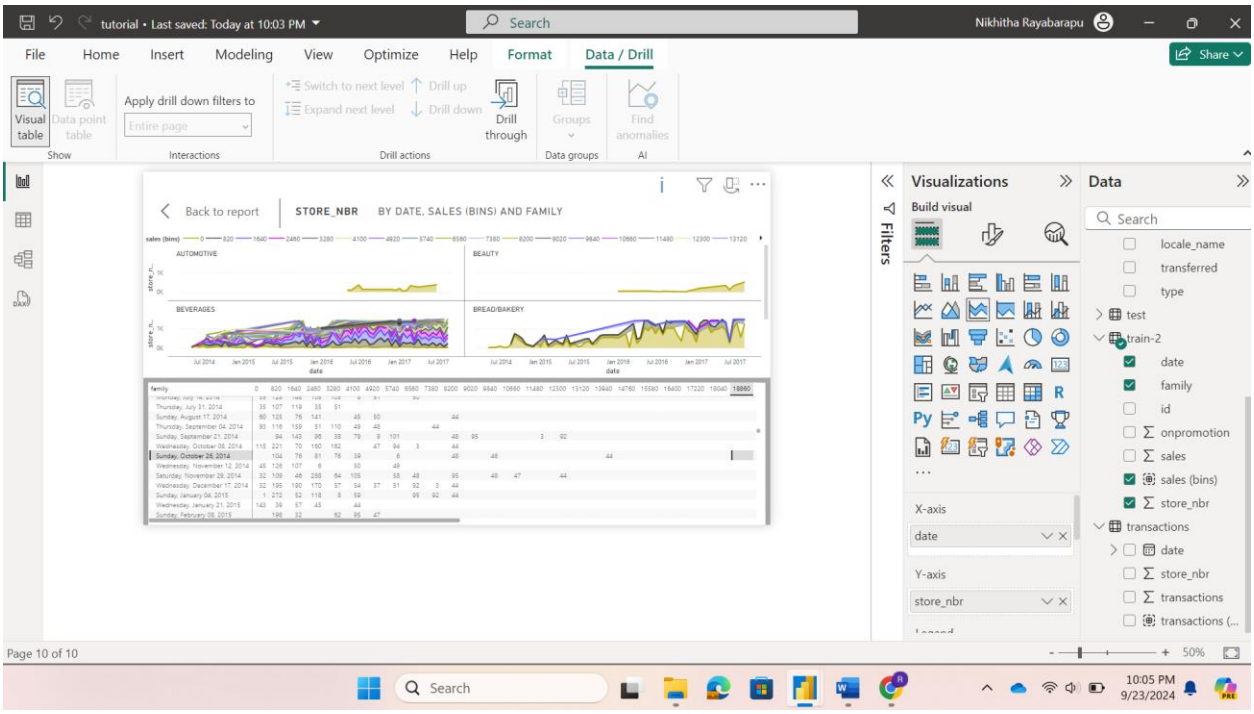
Step 3:- Select the attributes Date, sales(bins) and store_nbr and use the stacked area chart.



Step 4:- Select attribute Family and we can observe that the sales have been divided based on the family.

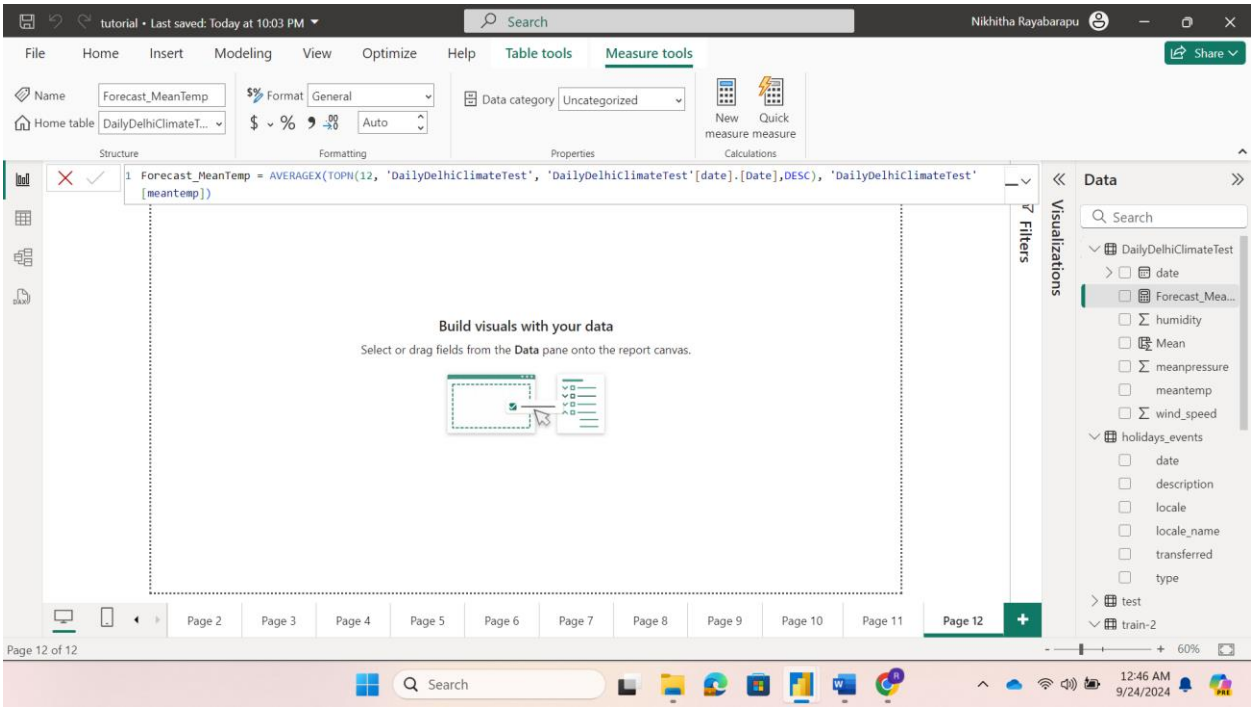


Step 5:- Get the visual table for the above Stacked area chart.



Tutorial 5: Forecasting

Step 1:- Using the New measure by naming it Forecast_MeanTemp and choosing the home table as DailyDelhiClimateTest. Use the below Query and you can see that an attribute named by Forecast_MeanTemp will be created under DailyDelhiClimateTest.



Step 2:- Select date, Forecast_Sales attributes and select the Line chart.