**Power BI Case Studies:**

**Case Study 2: (Power BI)**

**Objective:** To create an Interactive Sales Analysis Report using Power BI

**Source Files:**

1. Sample Superstore data in Excel format

**Problem Statement:**

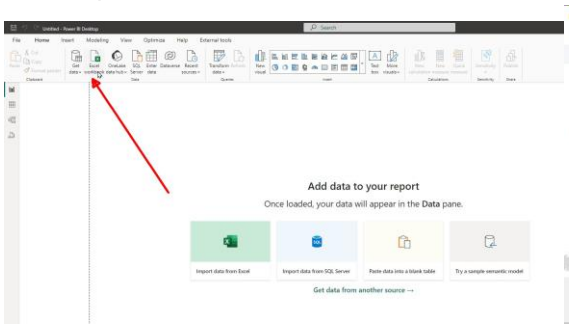
1. The Excel file contains three sheets, Orders, Returns and People.
2. From the flat orders table, dimensions are to be identified and modelled.
3. Calculate various measures and show them in various data visualizations.

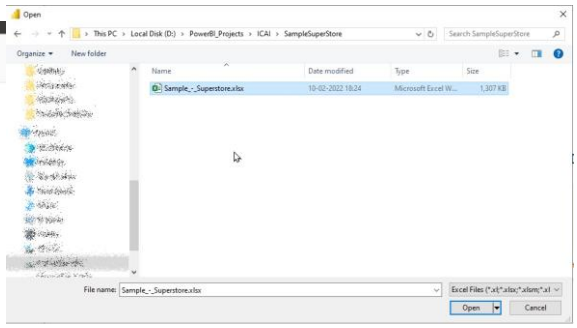
**Overview of steps:**

1. Use Extract and Transform feature of Power Query to get the data from the Excel workbook.
2. Identify the dimension tables and remove duplicates to make them Master data.
3. Remove unwanted columns from Orders and keep it as a fact table.
4. Create a Data Schema keeping Fact table in the center and dimensions table in surrounding.
5. Create Date Table for Time dimension and add that too.
6. Create necessary hierarchies.
7. Create various measures using DAX.
8. Use Power BI Visuals to create the required Interactive Sales Analysis Report.

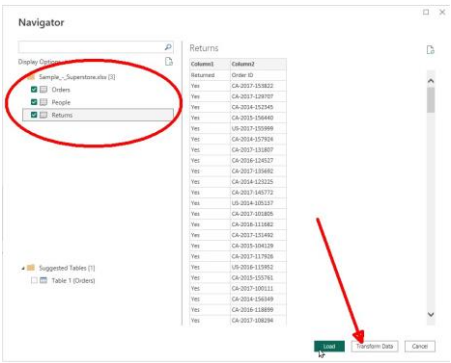
**Detailed Steps:**

1. Open Power BI for a New Project and Get Data from Excel Workbook. Select Sample_-_Superstore.xlsx

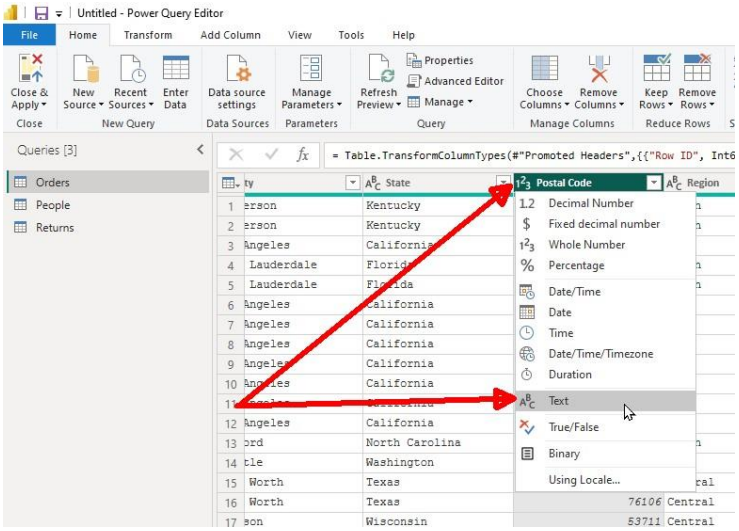


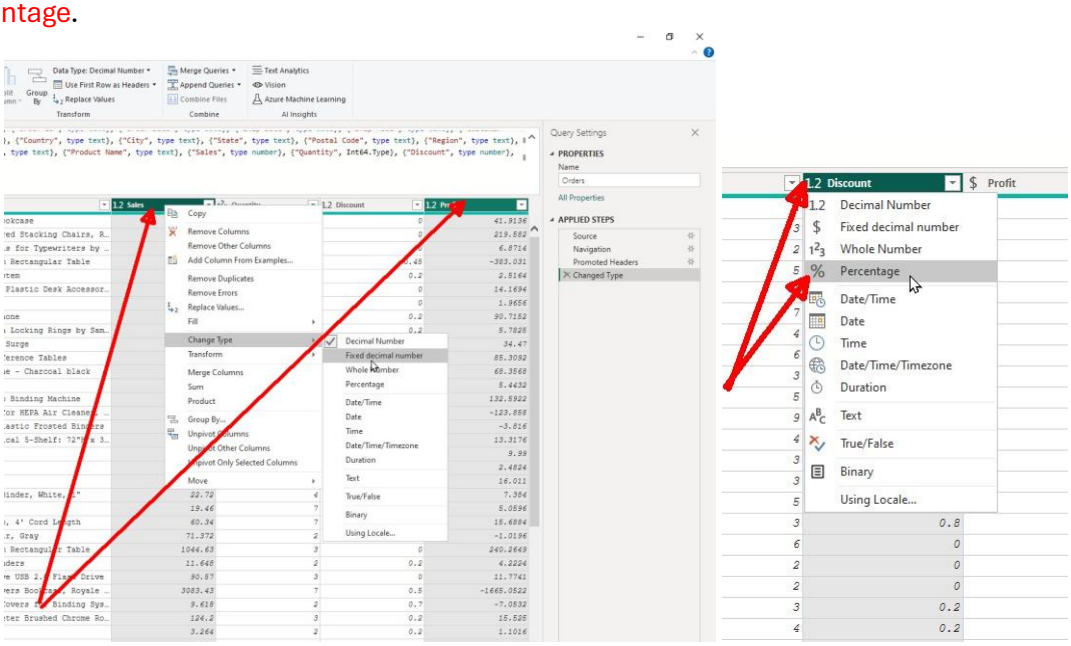

2. Select all the three sheets and click Transform



3. Change the type of Postal Code to Text. Click on '123' button on Postal Code Column Header. Select Text

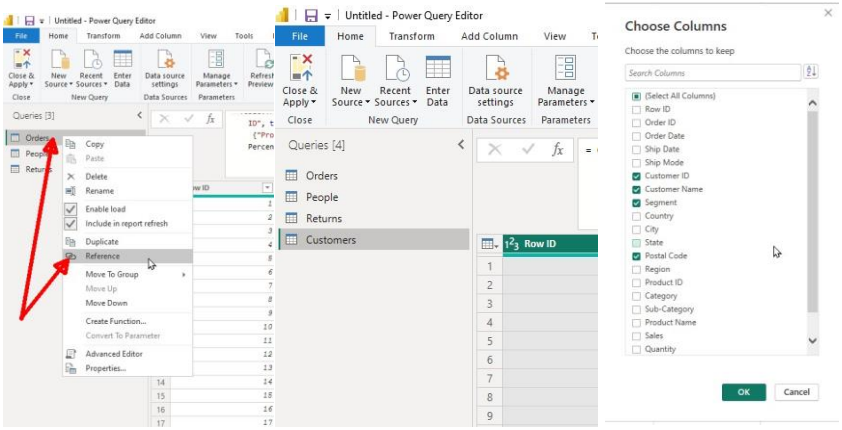



4. Change type of Sales and Profit Columns as Fixed Decimal number, Discount as Percentage.

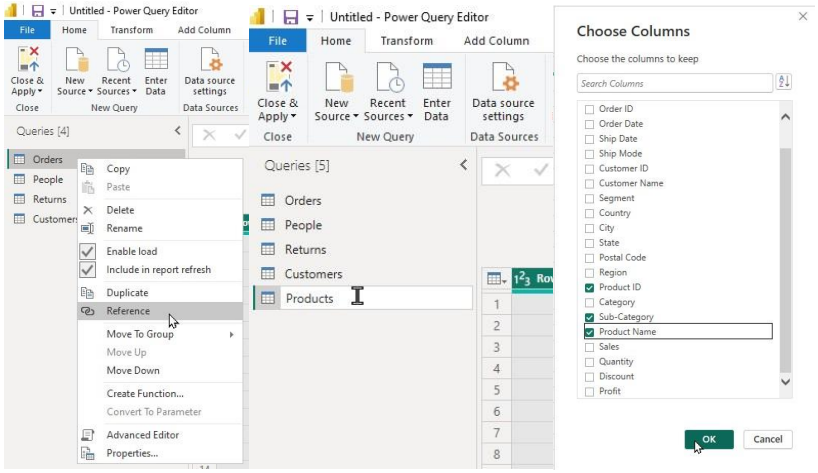



5. Click Replace Current to replace the current type.

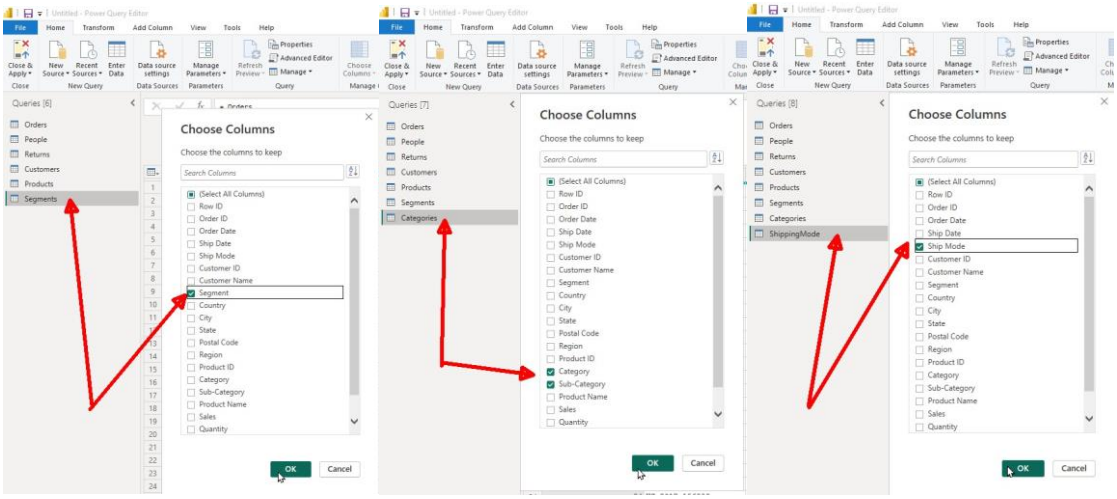

6. Right click on Orders Query and Select Reference. This will create a Query referring to the Orders. We keep Orders Query as the Base Query and create other Queries for Dimension Tables. Rename the resultant query as Customers. Choose columns related to Customers in this query by selecting Choose Columns.



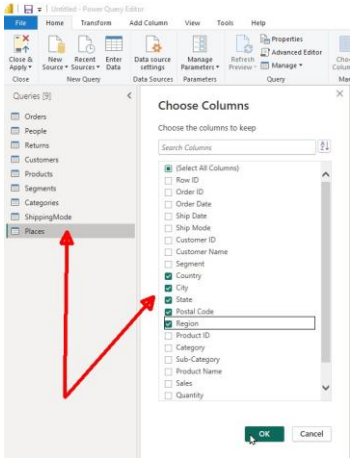

7. Create another Reference to the Orders Query and name it as Products and keep only Product related columns as in the image below.



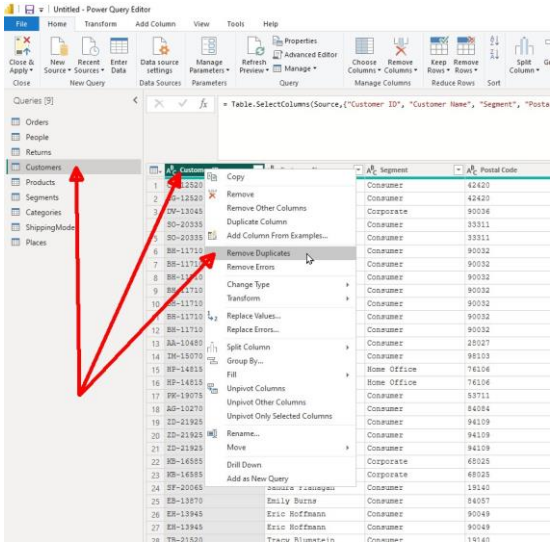

8. Same way, create three reference queries of Orders Query to create Segments, Categories and Shipping Mode. Choose columns accordingly.



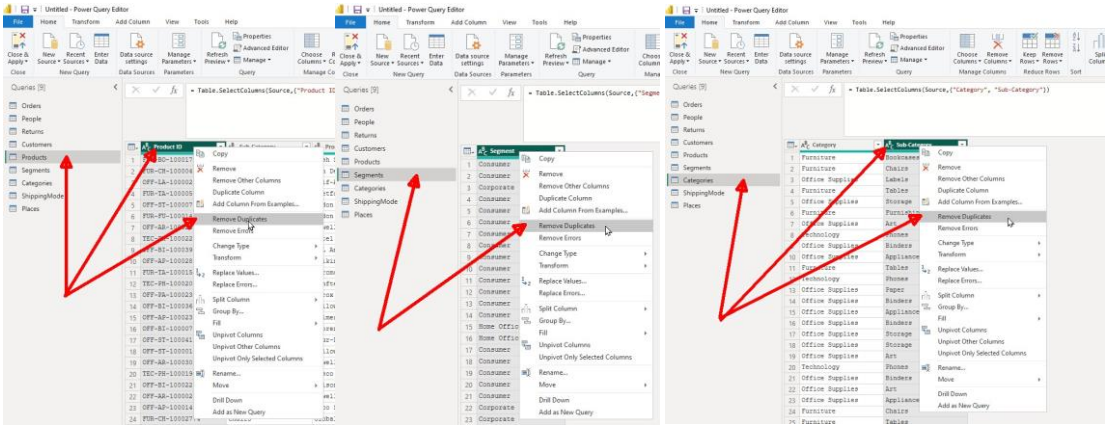

9. Create another reference for Places also.

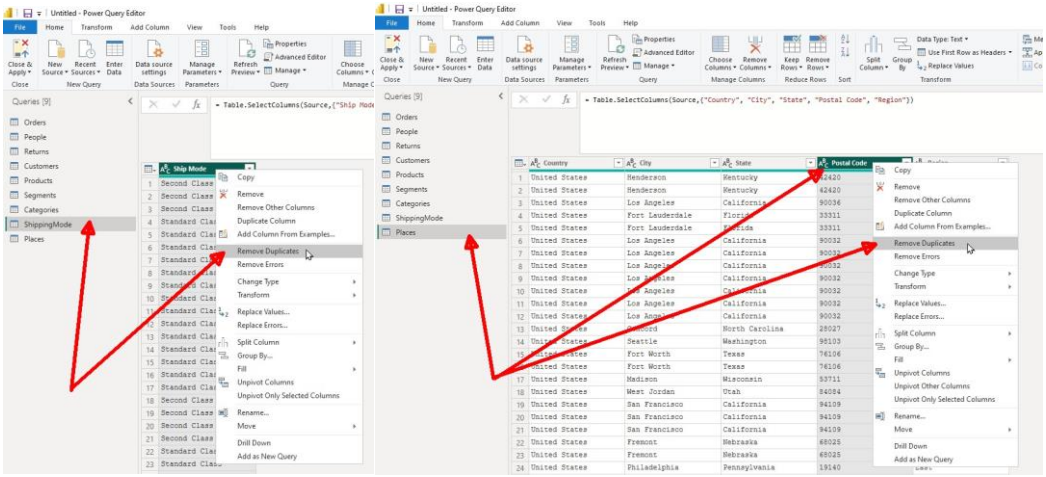



10. We have to remove the duplicates of Customer ID in the customers dimension table. Select Customer ID column, Right Click on the Column Header and Select Remove Duplicates.

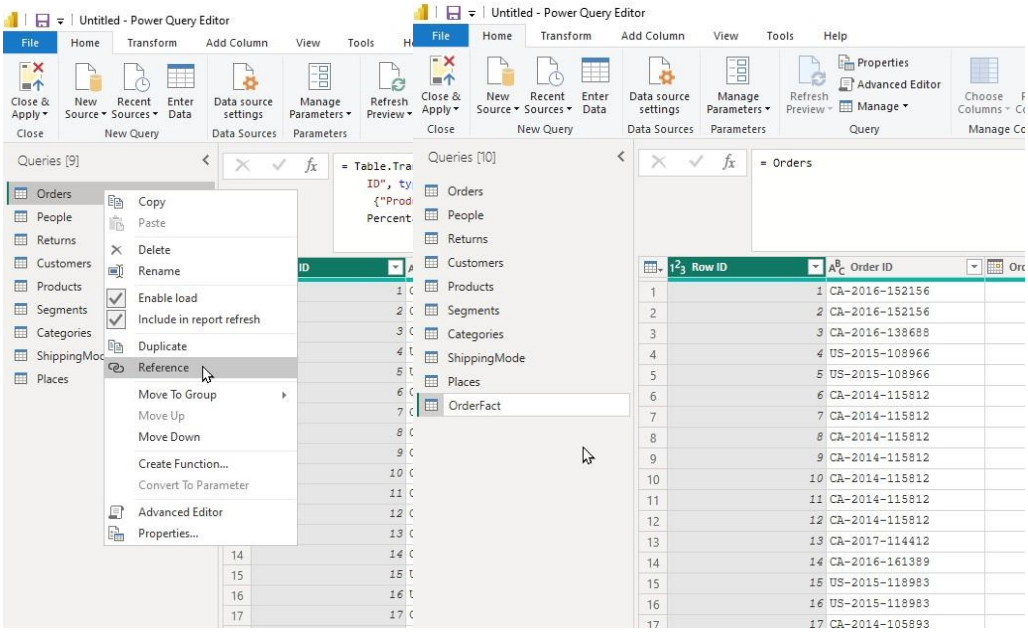


11. Likewise, Remove Duplicates for Products by Product ID, Segments, Categories by Sub-category, Shipping Mode, Places by Postal Code

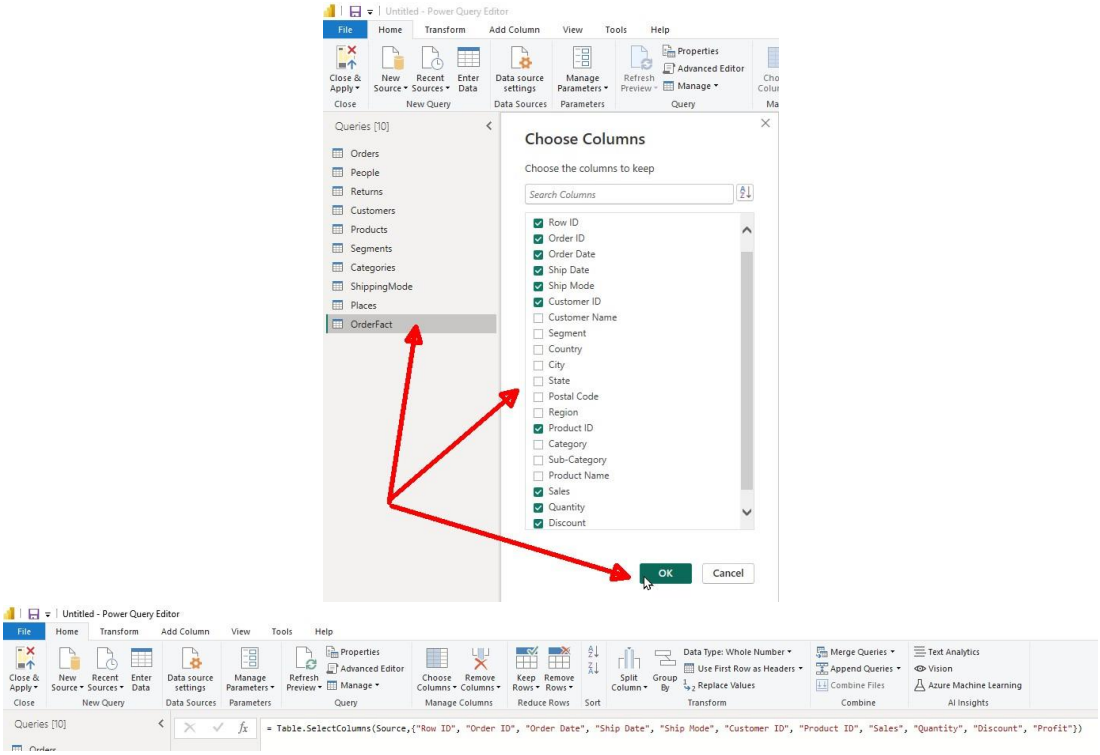




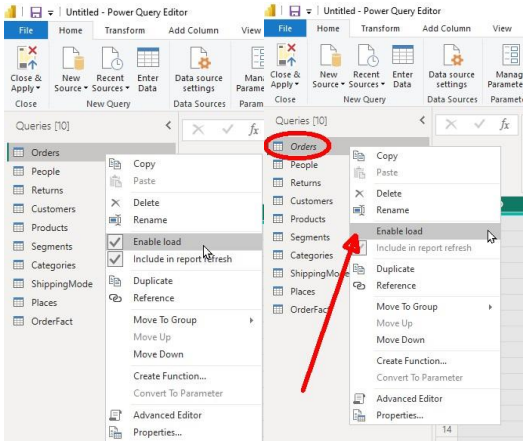
12. Now, we need to create the OrdersFact table. Make again a Reference Query from Orders



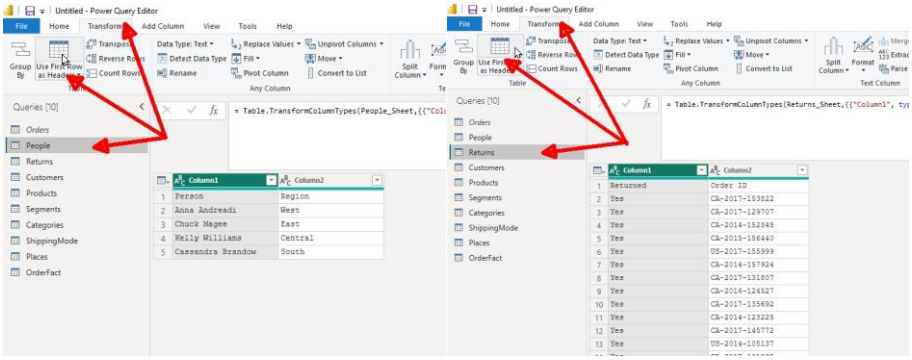

13. Choose only columns related to Facts and References to Dimension Tables. See the image below:



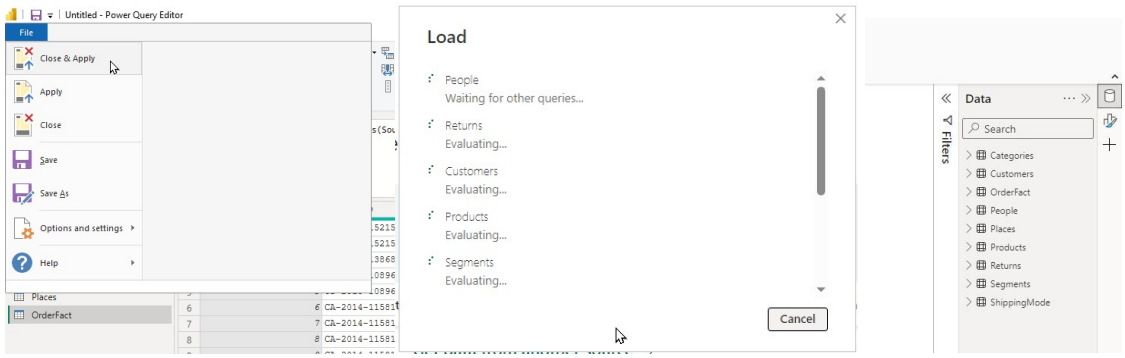

14. We don't require to Load the Orders Base Query. Right Click on the Orders Query and Click on Enable Load and Disable the same. Now the Orders Query is shown in Italics.



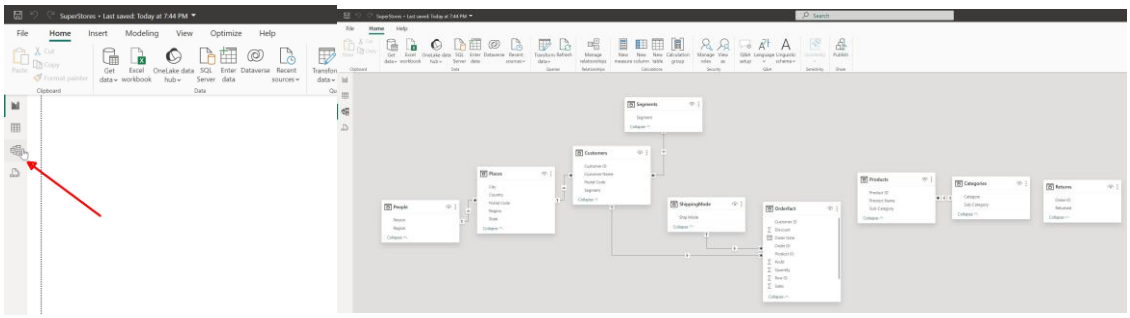

15. Select the People and Returns Queries and promote the headers by clicking Transform Tab-> Use First Row as Headers.



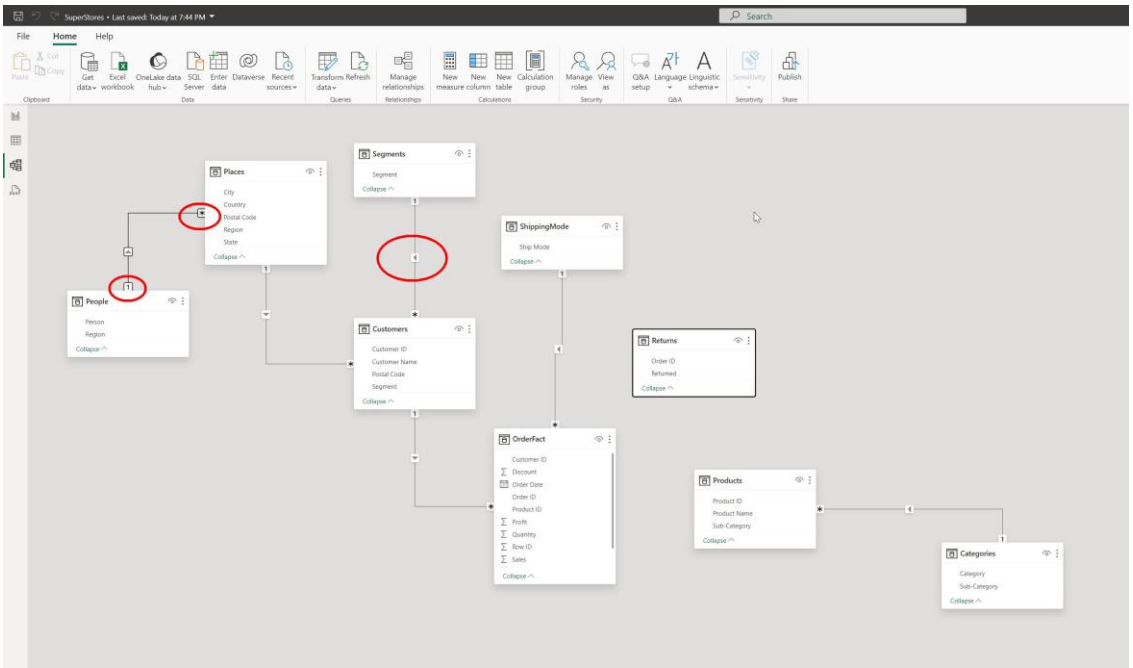

16. All transformations are over and now we can close Power Query and load the data. Click File->Close and Apply. System will load the data.



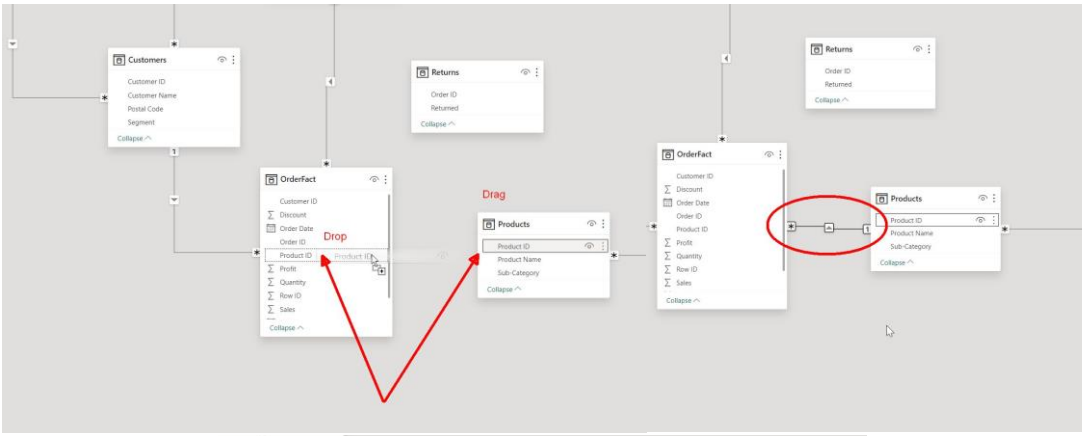

17. Click on Data Model view and we can see some Tables are already connected by default.

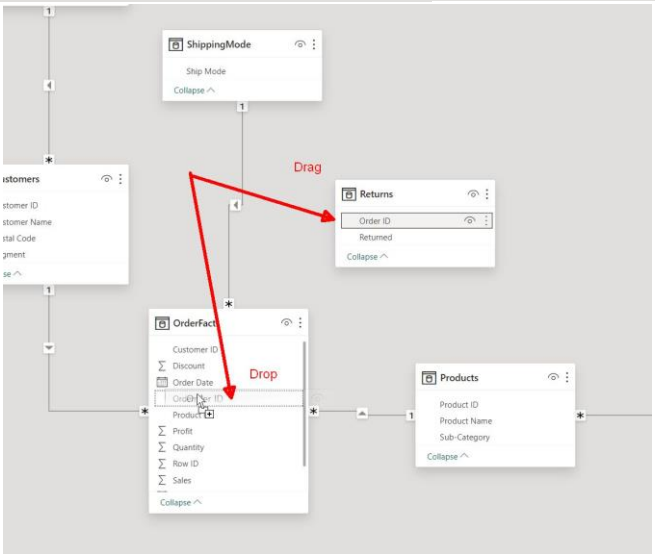


18. Place the OrderFact in the Center and Dimension Tables in the surrounding. Note how the filter flow is shown from One -> Many.

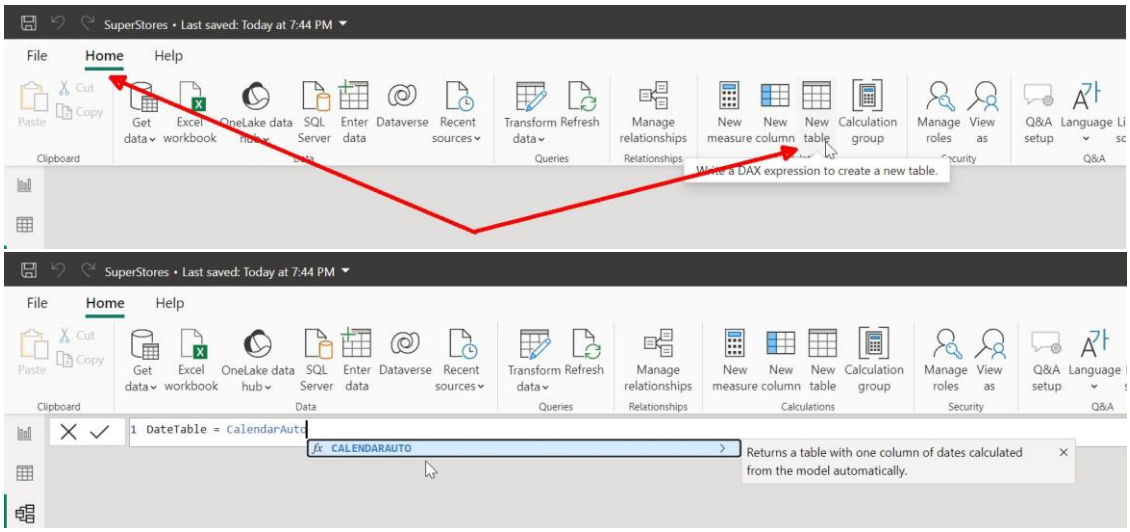



19. Connect the other dimensions manually by dragging and dropping the Common Key Fields eg,. Products -> Product ID, Returns -> Order ID etc.,

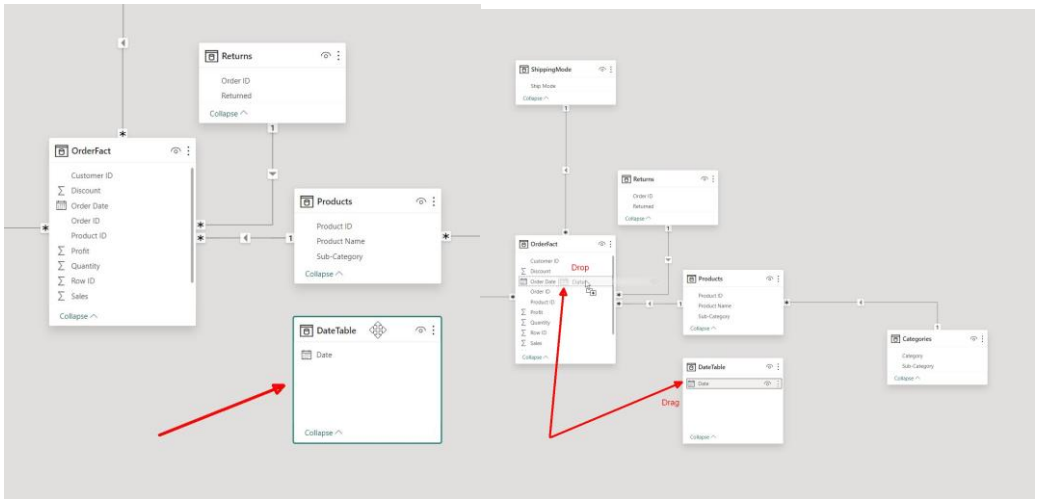


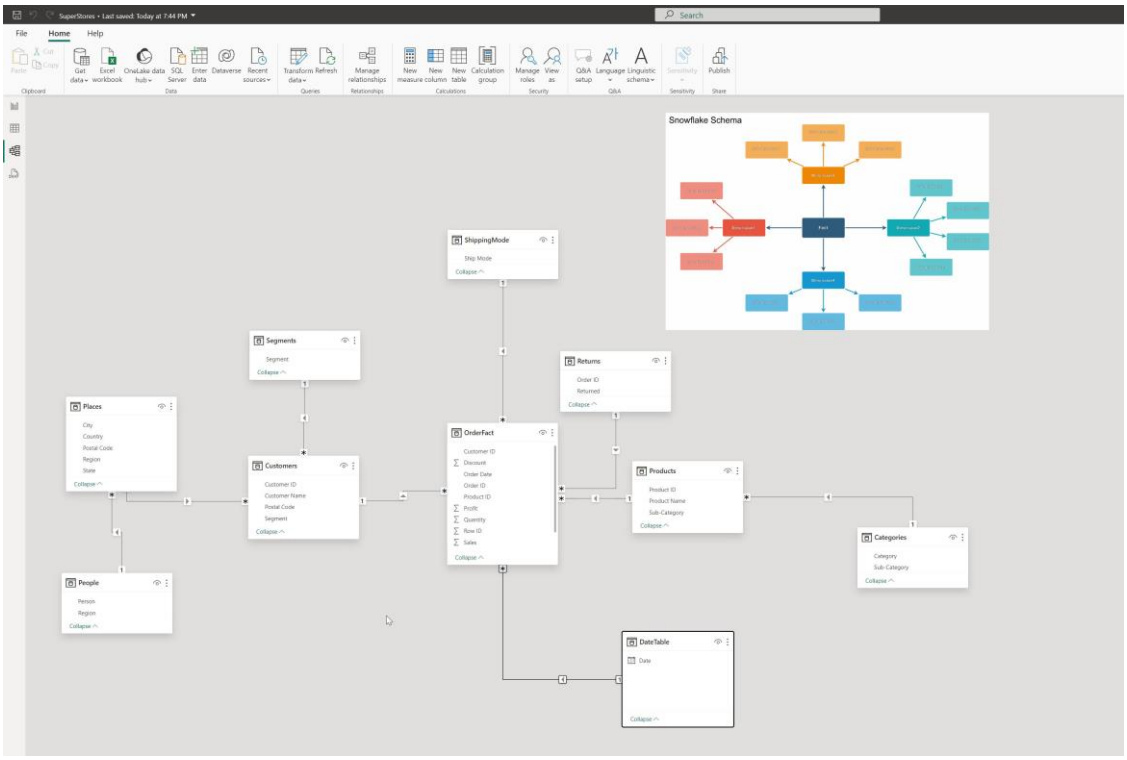

20. We have to create DateTable which is used for Time Dimension and Time Intelligence Calculations. Click on Home Tab and New Table. We use CalendarAuto() DAX Function to do so. We can choose 3 (for March) as Financial Year End. Since it is US data and we choose Calendar year.



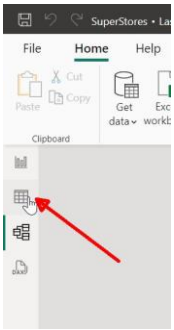

21. We can see a new table called Date Table is added to the Data Model. Drag and Drop the Date field to OrderDate of OrderFact.



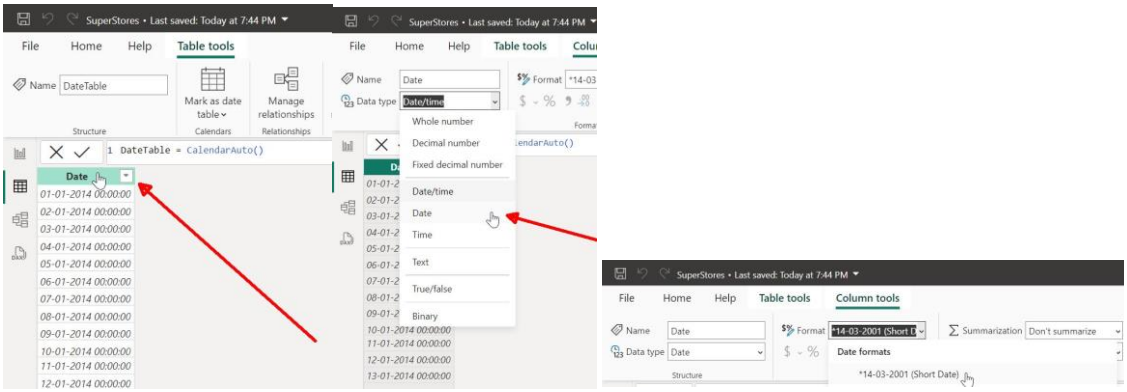

22. We can see a Snowflake Schema, having OrdersFact in the center and other dimensions and sub-dimensions in the surrounding.



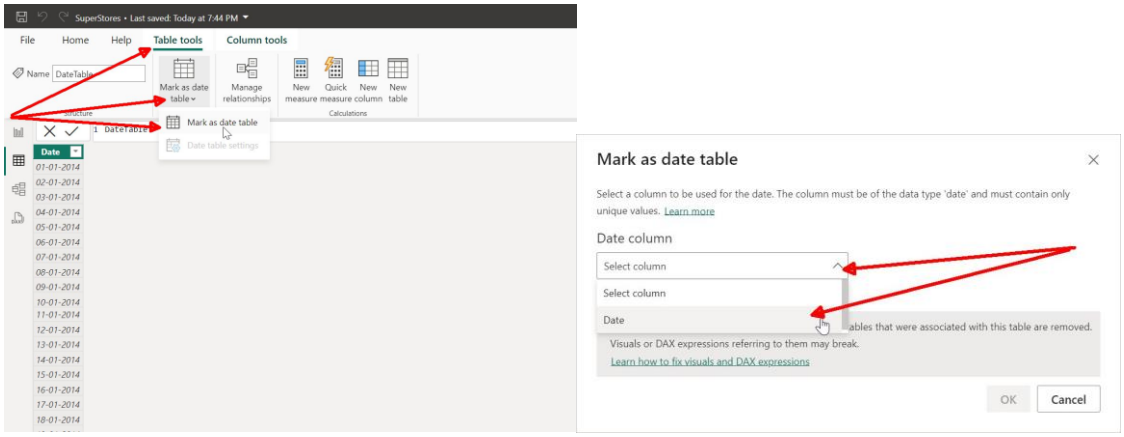

23. Select Table View



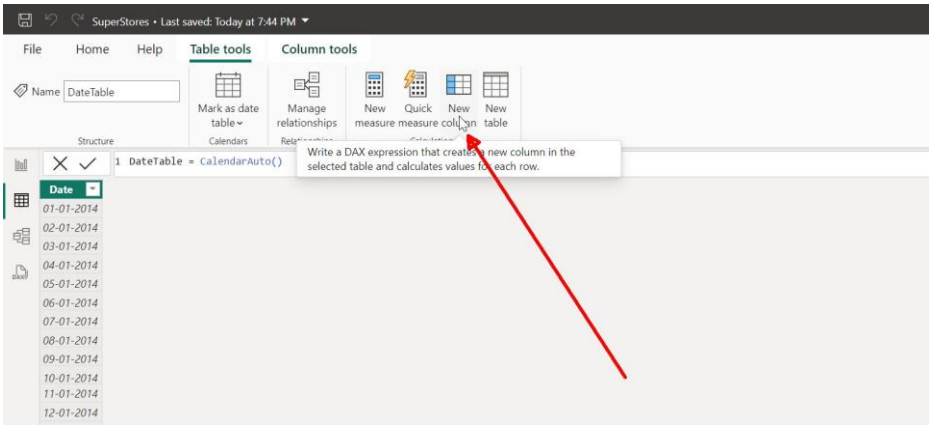

24. Click on the DateTable Date column and convert to Short Date as detailed below:



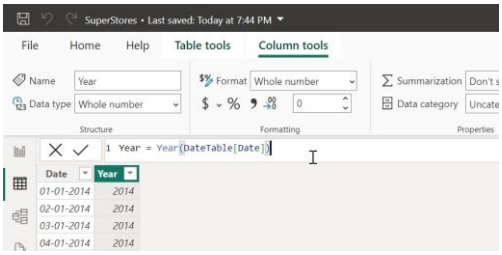
25. Click on Table Tools and Click Mark as Date Table. Select the Date Column in the drop-down. Click Ok to mark the Date table as the Date Table for time intelligence calculations.



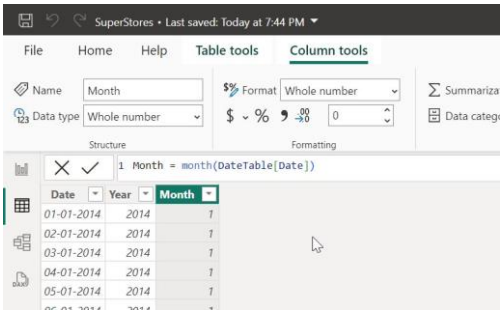
26. Let us create some hierarchies for time dimension by adding calculated columns using DAX. Click Table Tools and Select New Column.



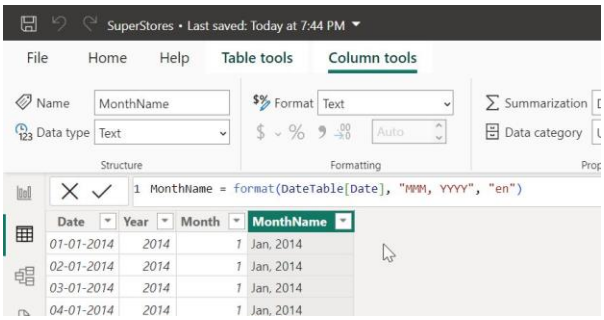
27. Add Year Column using Year = Year( DateTable[Date])



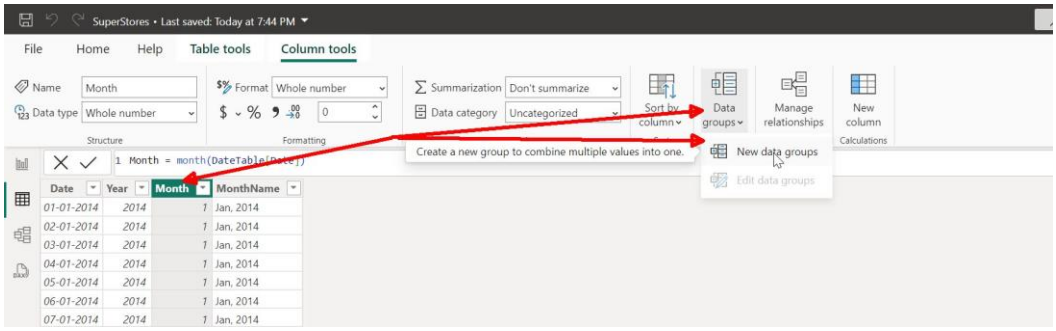
28. Add Month column using Month = Month(DateTable[Date])



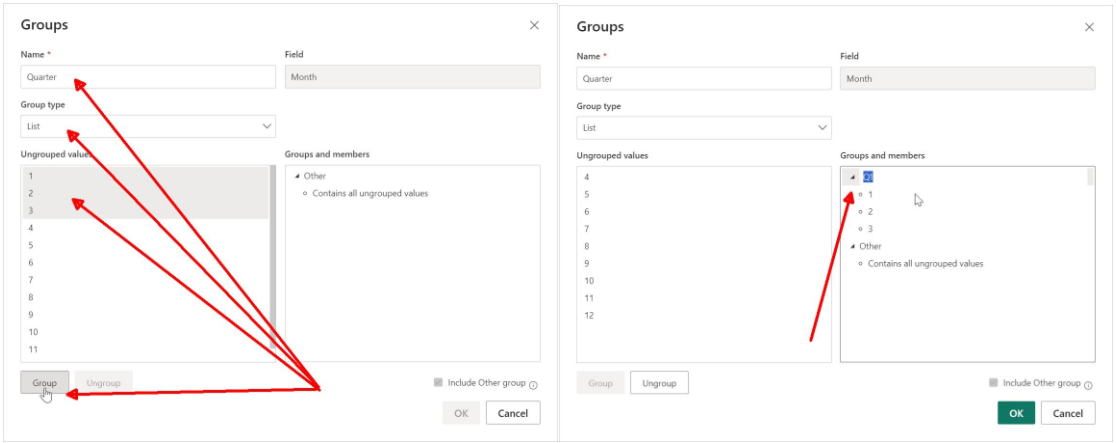

29. We can add MonthName for Reporting purpose by MonthName = Format(DateTable[Date], “MMM, YYYY”, “en”)



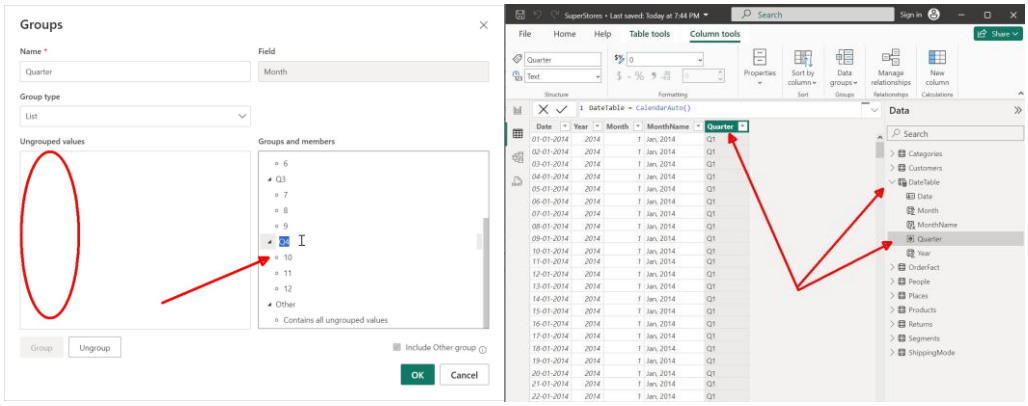

30. We can create Quarters by grouping the month by the Power BI Feature Data Groups. Click on Column Tools->Data Groups->New Data Groups




31. Give the Group name as Quarter. Select List in Group Type. Select 1,2,3 Months and Group them as Q1.

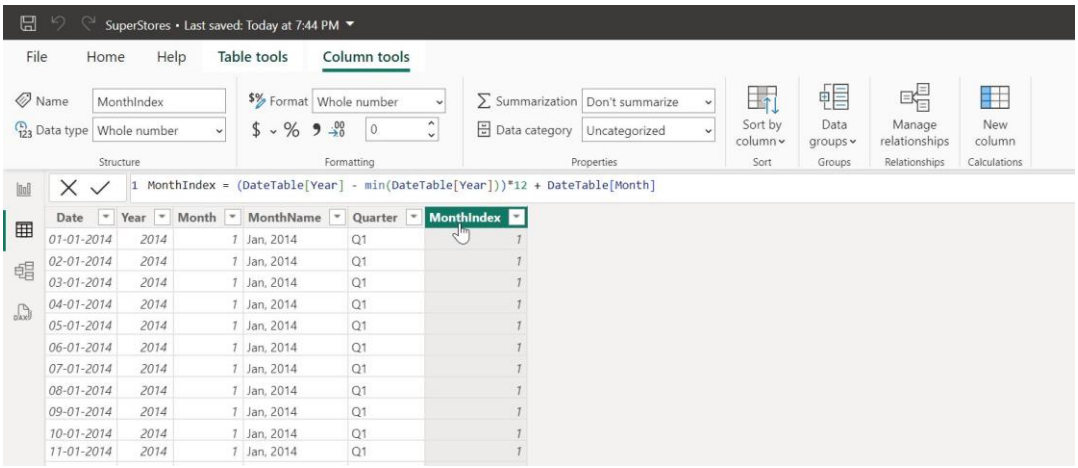



32. Likewise, group all the balance months into Q2, Q3 and Q4.

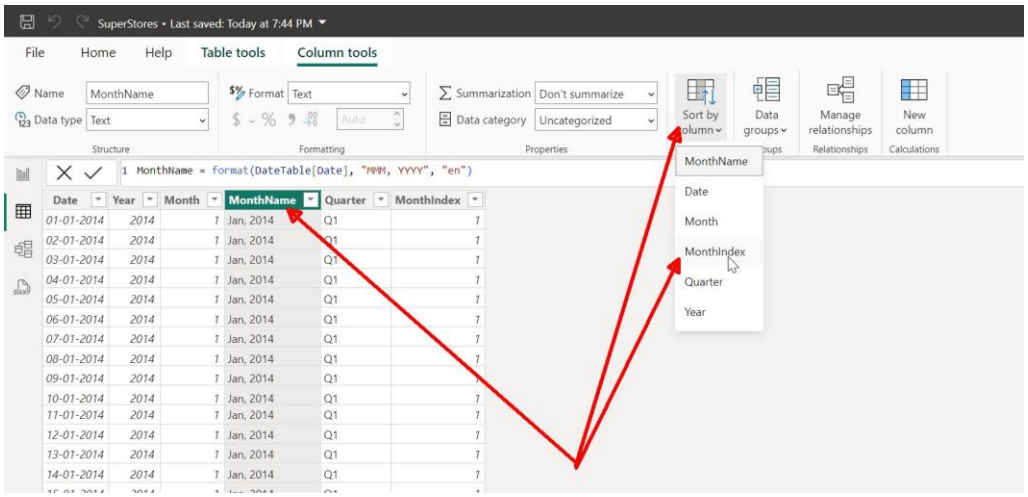



33. We need to create a MonthIndex column by using the following DAX Expression.

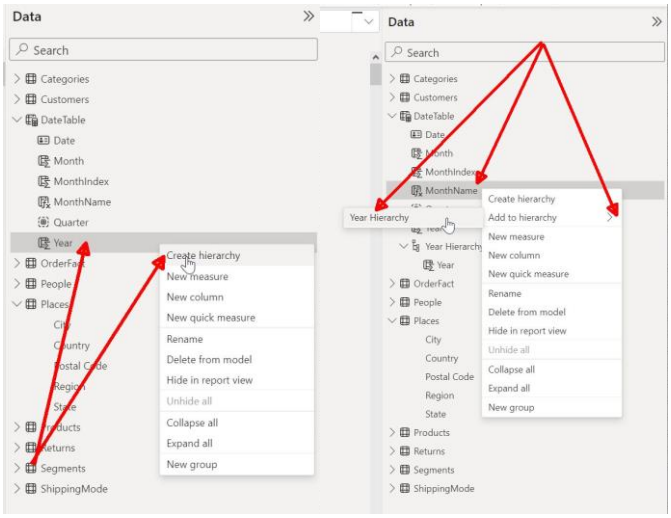
MonthIndex = (DateTable[Year] – Min(DateTable[Year])*12 + DateTable[Month]



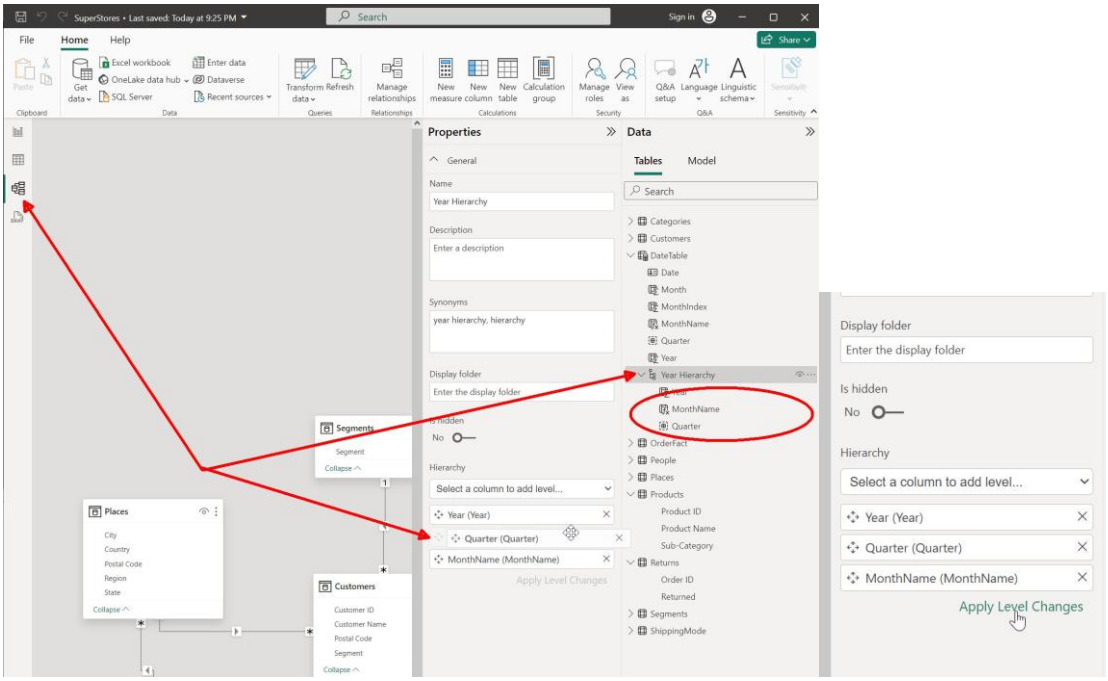

34. Select MonthName and Click on Sort by column. Select MonthIndex in the list. This is one of the nice features Power BI to sort the MonthName column by MonthIndex always. This will avoid sorting the months as April, August, December etc., instead of January, February, March.



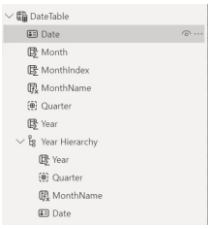

35. We can now move on to creating Hierarchies. First select and Expand Date Table in the right hand side Data area. Right click on the Year and Select Create Hierarchy. Add MonthName, Quarter also to the Hierarchy.



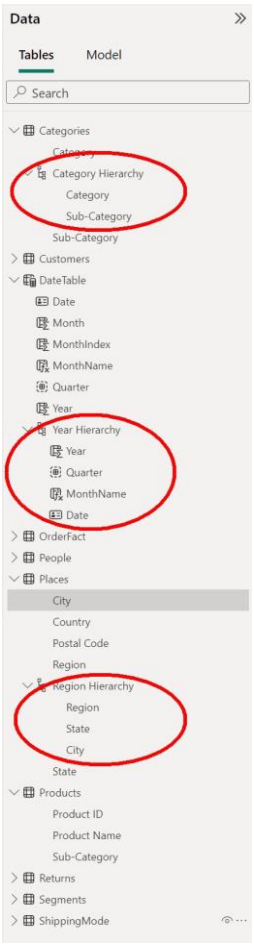

36. To re-organize the Hierarchy like Year->Quarter->MonthName->Date, go to Model view and drag and drop the Hierarchy as below:



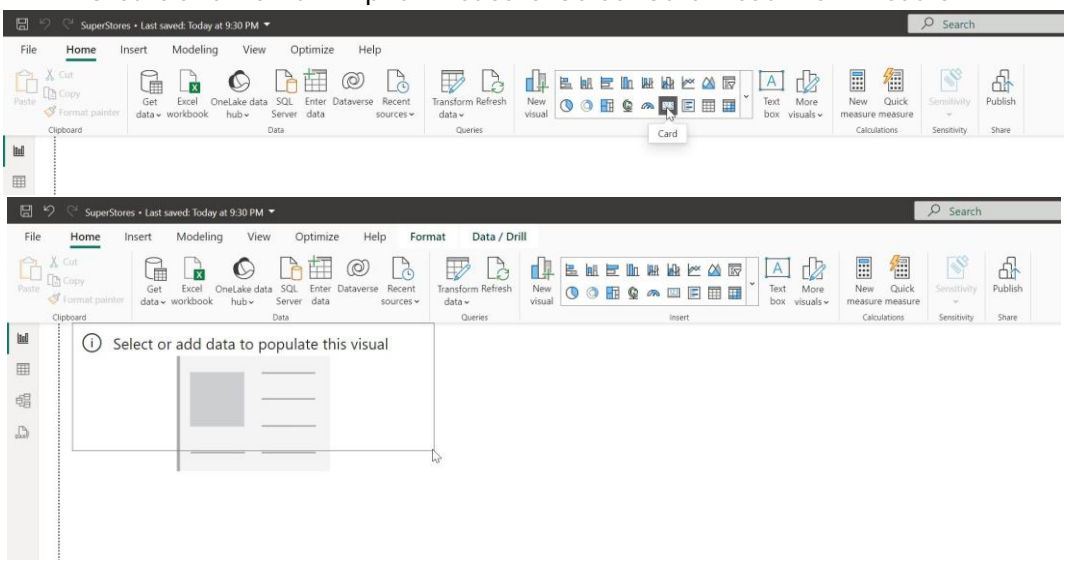

37. See the Year Hierarchy in order as we require.



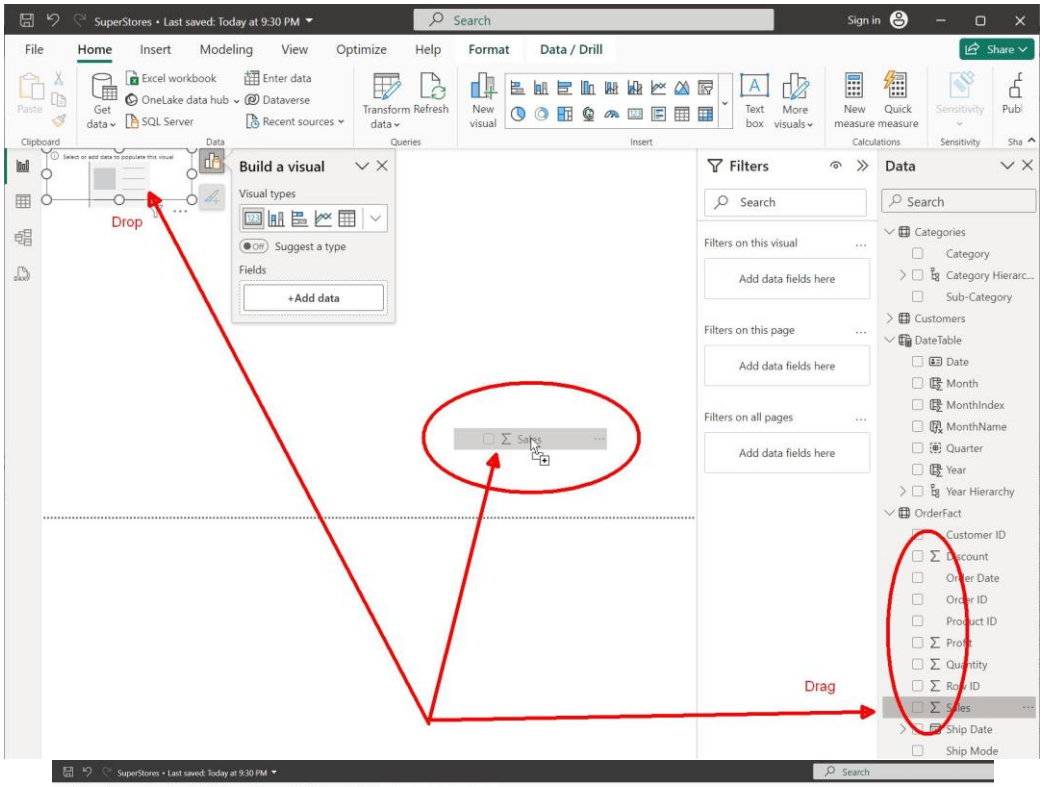

38. Likewise, create hierarchies for other dimensions also as below:

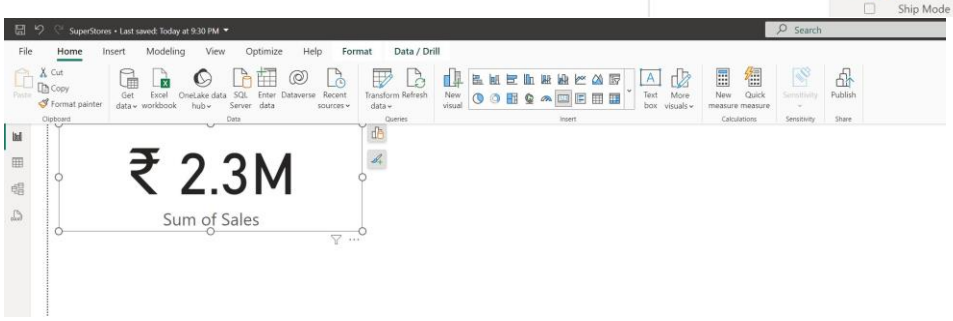



39. Create and view an Implicit Measure. Select Card Visual from Visuals.

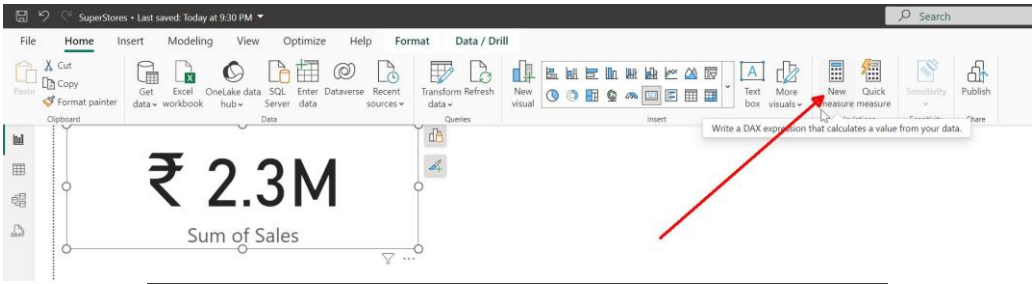


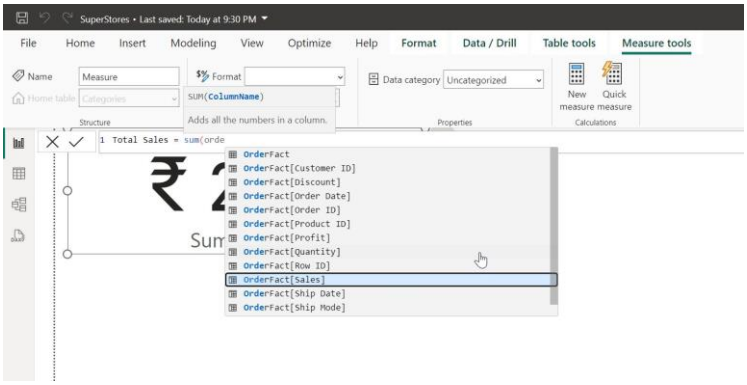

40. Drag and Drop Sales field from OrderFact Table to the Card Visual. See the Sigma symbol for all the numeric fields. This is called an Implicit Measure. Power BI automatically creates these measures when we drag and drop any value field to the visuals.



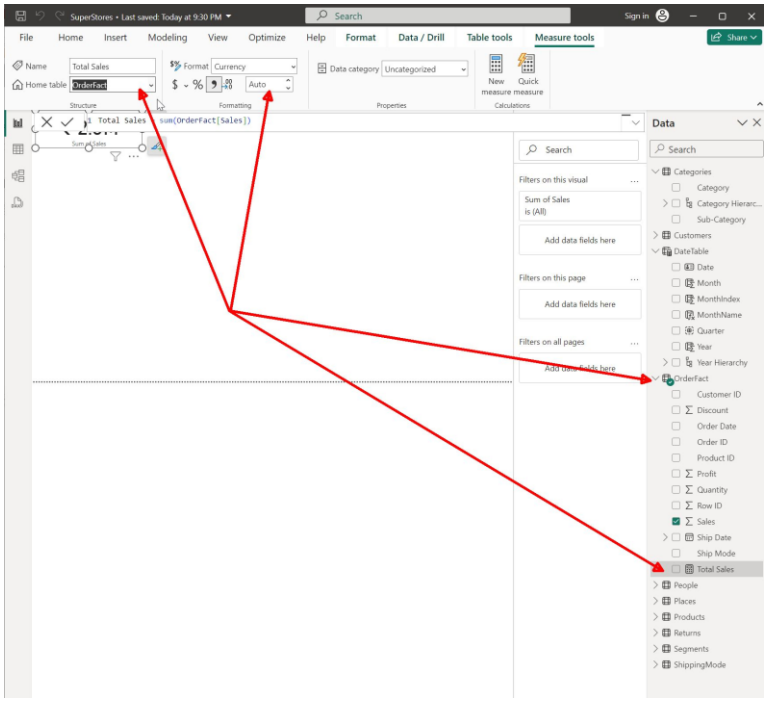




41. From now onwards, we shall use Explicit Measures which are calculated using DAX formulas. Click on New Measure and Enter the following Formula: Total Sales = Sum(OrderFact[Sales])

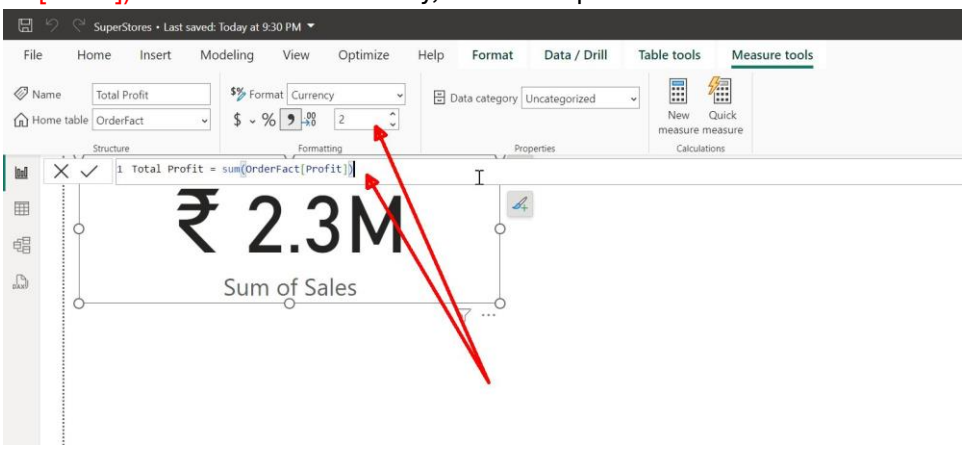


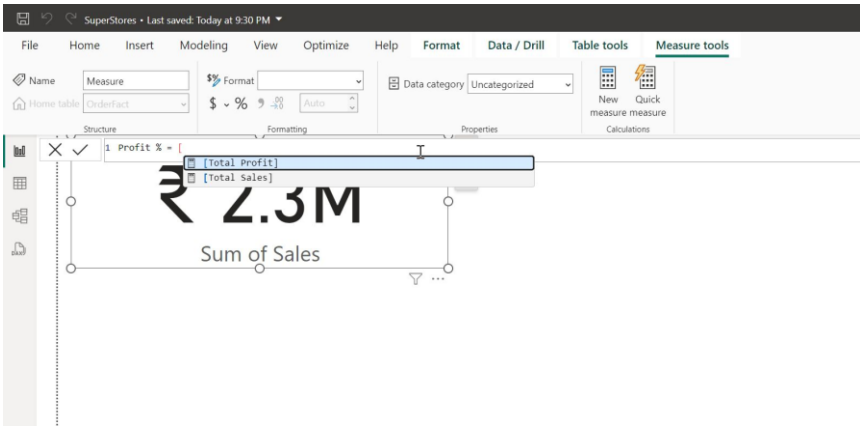
42. Keep the Home Table of the measure in OrderFact Table and format the measure as Currency with 2 Decimals.



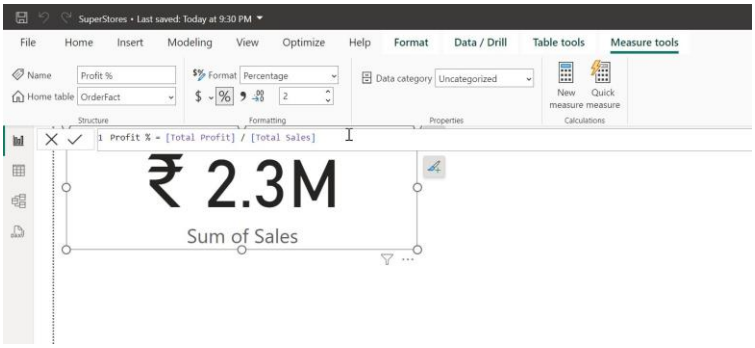
43. Likewise create another Measure for Total Profit using the Formula Total Profit = Sum(OrderFact[Profit]) and format as Currency, 2 decimal places.



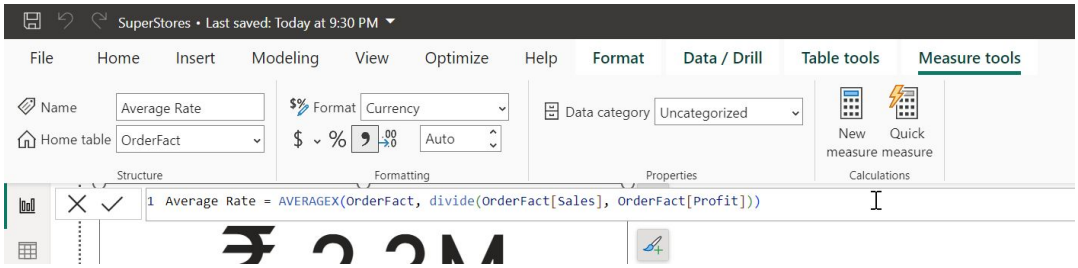

44. Profit Rate can be calculated by using the above two measures. We can re-use only Explicit Measures as this way. This is one of the major advantages of Explicit Measures. For referring the Measures, we can just use '[' bracket without Table Name as below:



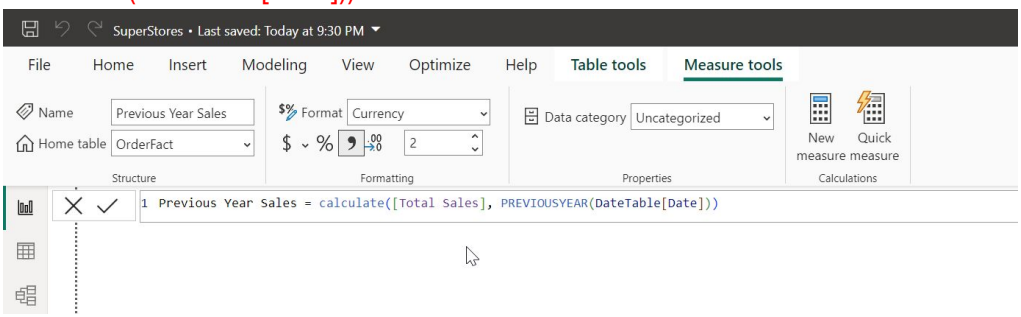
45. Format the measure with % symbol and 2 decimal places.



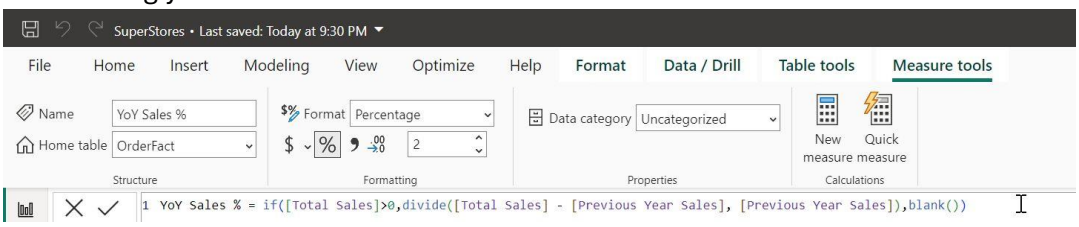

46. For Average Rate we are going to use AverageX DAX function. All functions ending with 'X' are used to evaluate any expression in row context (here we want to divide sales by quantity to get the rate) and then do the summary. This will avoid preparing a separate calculated column for Rate. DAX Formula: Average Rate = AVERAGEX(OrderFact, Divide( OrderFact[Sales], OrderFact[Quantity])) Divide Function in DAX is used for safe dividing the numbers excluding zero division errors.



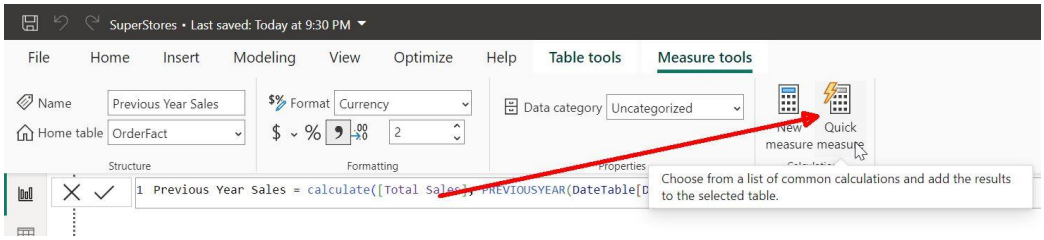

47. Calculate() function is used in DAX for calculating any expression for a modified filter context. This is one of the most used functions in Power BI. We used Calculate function along with PreviousYear() for calculating Previous Year Sales for comparing Sales with Previous Year. This is possible because of DateTable. Formula: Previous Year Sales = calculate([Total Sales], PreviousYear(DateTable[Date]))



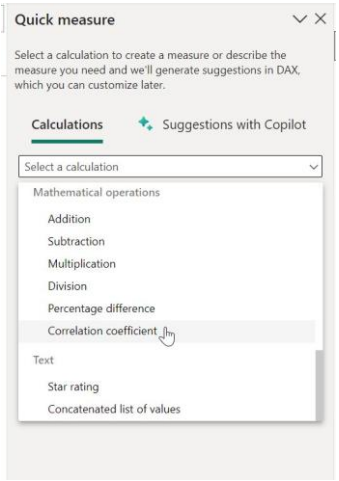

48. Calculate YoY Sales % with the following Measure. YoY Sales % = if([Total Sales] > 0, divide([Total Sales] – [Previous Year Sales], [Previous Year Sales]), blank()) Format the measures accordingly.



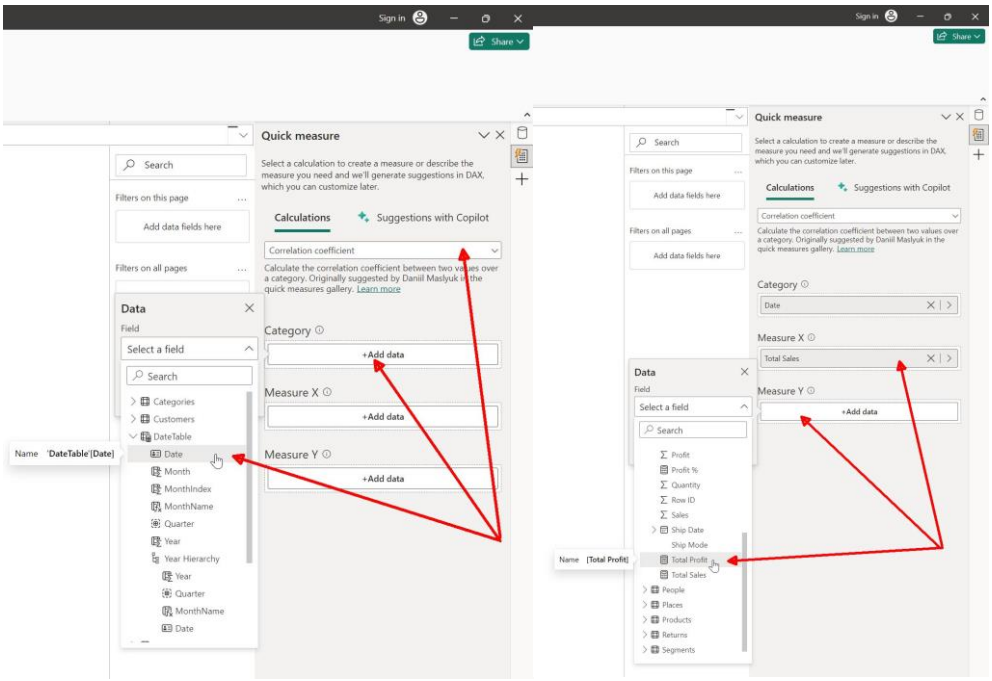

49. Click on Quick Measure to add a pre-defined measure by Microsoft inside Power BI.



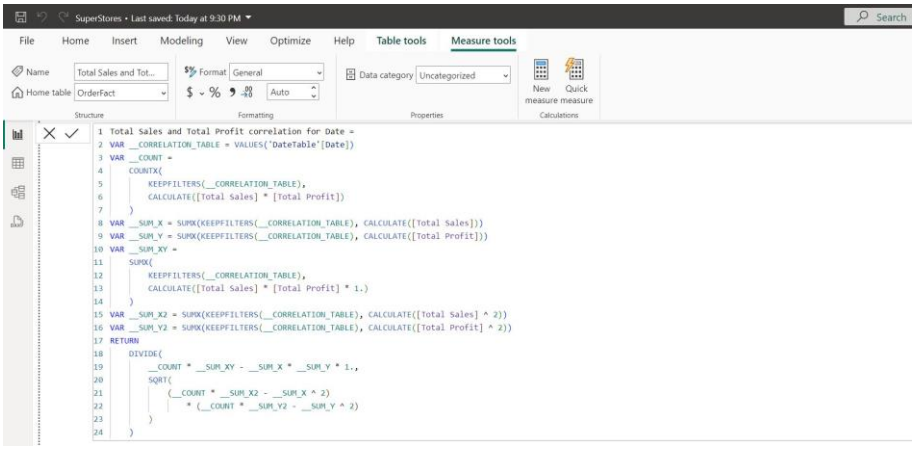

50. Click on Correlation Co-efficient



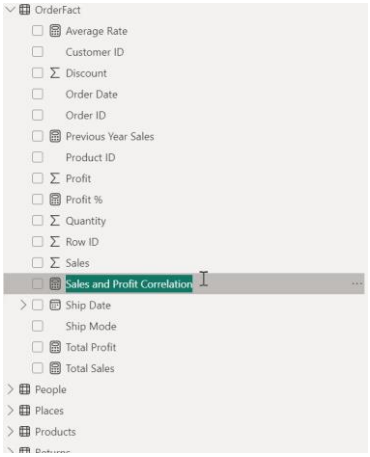

51. Add DateTable[Date] in the Category, [Total Sales] in Measure X and [Total Profit] in Measure Y



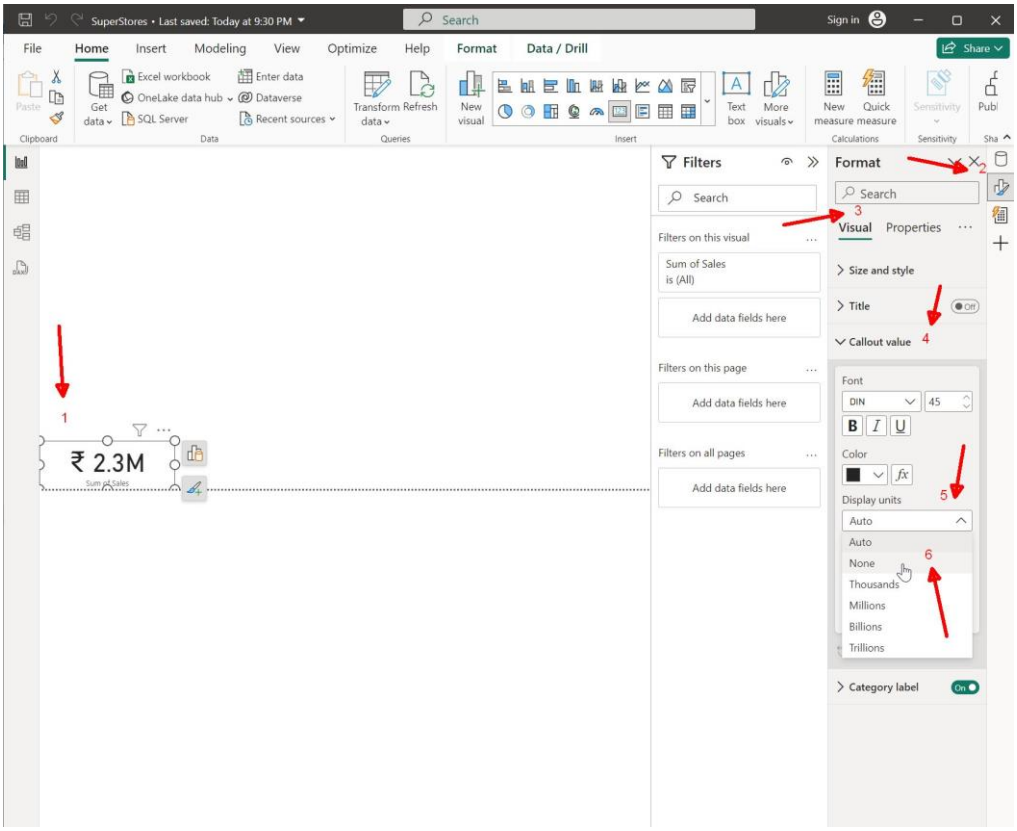

52. A new measure for calculating Correlation co-efficient of Daily Sales and Profit are created by a lengthy DAX formula. We can learn how to use VAR (variables) and RETURN commands inside DAX formulas.



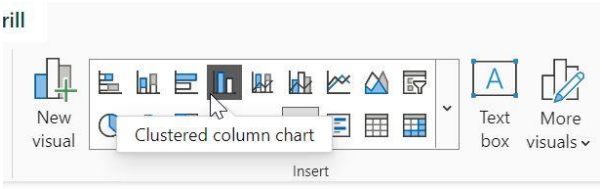

53. Rename the Correlation measure as Sales and Profit Correlation.



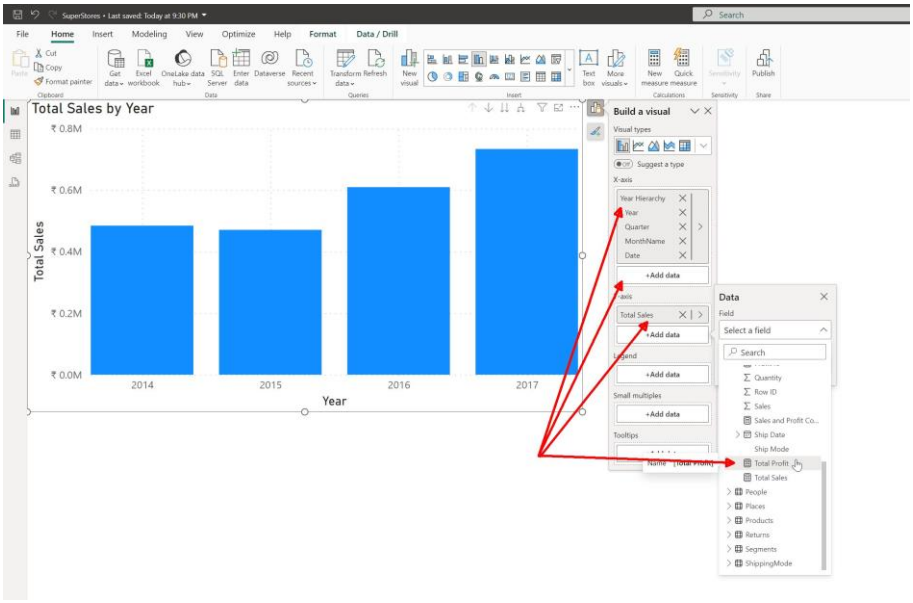

54. If we want to show the full amount on any visual instead of showing in Thousands or Millions, Select the Visual-> Format Pane-> Visual->Call out value->Display Units->None



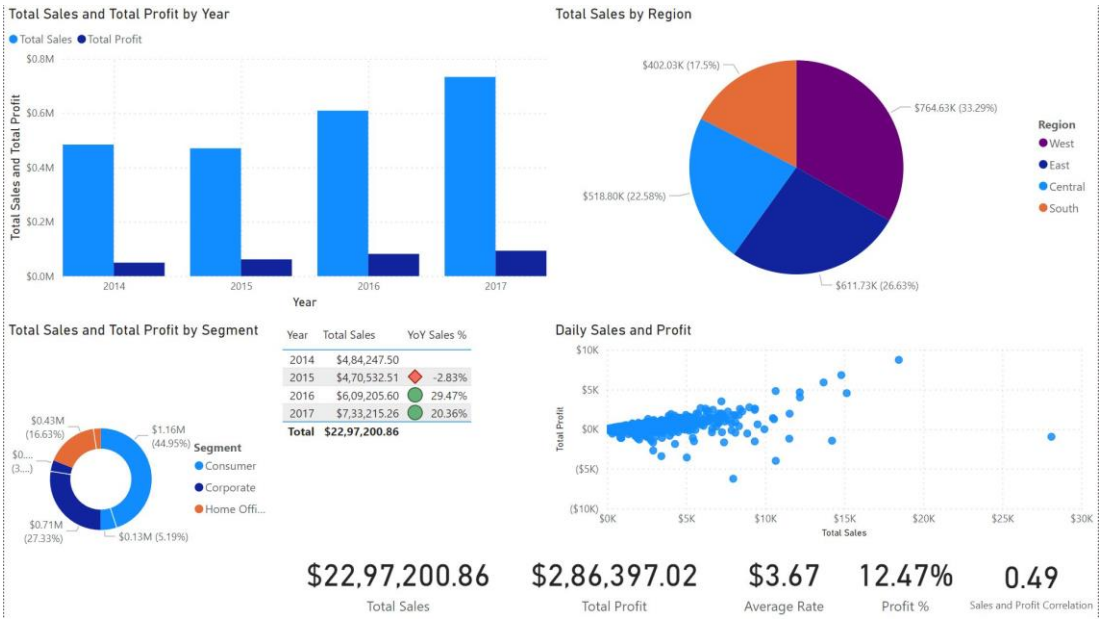
55. Let us add a Chart. Click on empty area of the Report View and Select Clustered column chart.




56. Resize the chart and add fields Year Hierarchy to X axis, Total Sales and Total Profit to Y Axis.

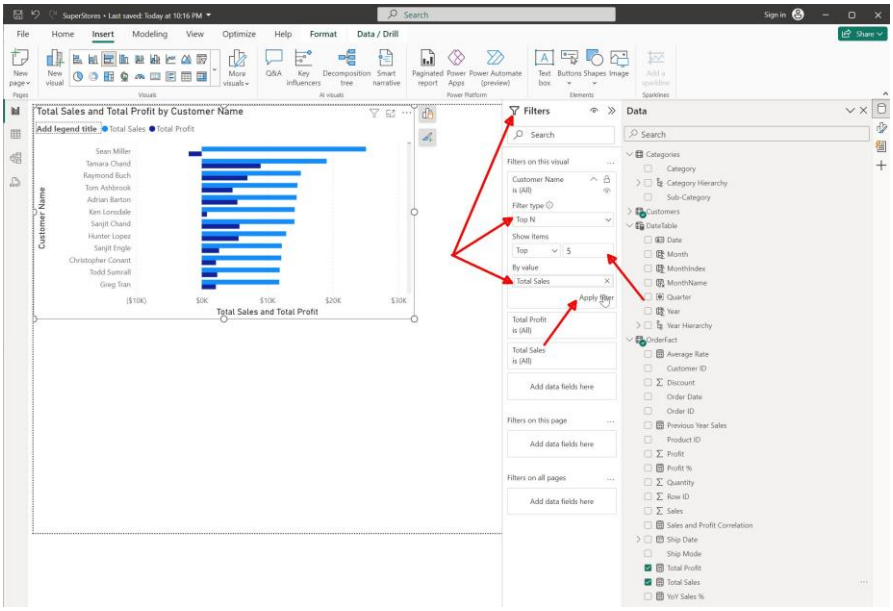



57. Try to create a Report Page like the one below:




58. Click on any visual and see the impact on other visuals. All the visuals are by default interactive to each other and all the measures are recalculated according to the current filter context (except for measures calculated by calculate() which uses a modified filter context).

59. We can use Filter Area to create Top 5 Customers like below:




60. Try to create another Report Page like the one below with Top 5 Customers.