**CAPSTONE PROJECT**

***Classic models – Sales Analysis***

A small company Axon, which is a retailer selling classic cars, is facing issues in managing and analysing their sales data. The sales team is struggling to make sense of the data and they do not have a centralized system to manage and analyse the data. The management is unable to get accurate and up-to-date sales reports, which is affecting the decision-making process.

To address this issue, the company has decided to implement a Business Intelligence (BI) tool that can help them manage and analyse their sales data effectively.

**Database description:**

The company has provided classicmodel business data which contain 8 tables named customers, products, orders, order line, employees, products, product line and offices.

**Tools Required:**

1. **MySQL:** The data provided is in the form of SQL format, so to work on MySQL is needed.
2. **Power BI:** This tool is used to clean and visualize the data to get key insights for making business decisions.

**Description of tables:**

The MySQL sample database schema consists of the following 8 tables:

* Customers: stores customer’s data.
* Products: stores a list of scale model cars.
* ProductLines: stores a list of product line categories.
* Orders: stores sales orders placed by customers.
* OrderDetails: stores sales order line items for each sales order.
* Payments: stores payments made by customers based on their accounts.
* Employees: stores all employee information as well as the organization structure such as who reports to whom.
* Offices: stores sales office data

**Database Download Link:** <https://drive.google.com/file/d/1OB_iGw6vVS5KS7QwiwVChbeTfR4WvUy3/view?usp=share_link>

**Working on database:**

* Firstly, import the data into MySQL database and create a new schema named classicmodels.
* After creating schema, check the tables and columns by using “**select”** function and count no. of records in each table.
* Now load the data into Power BI for cleaning and visualize the data.

**DATA CLEANING: -**

The data cleaning has performed using ‘power bi’ in that power query editor is used.

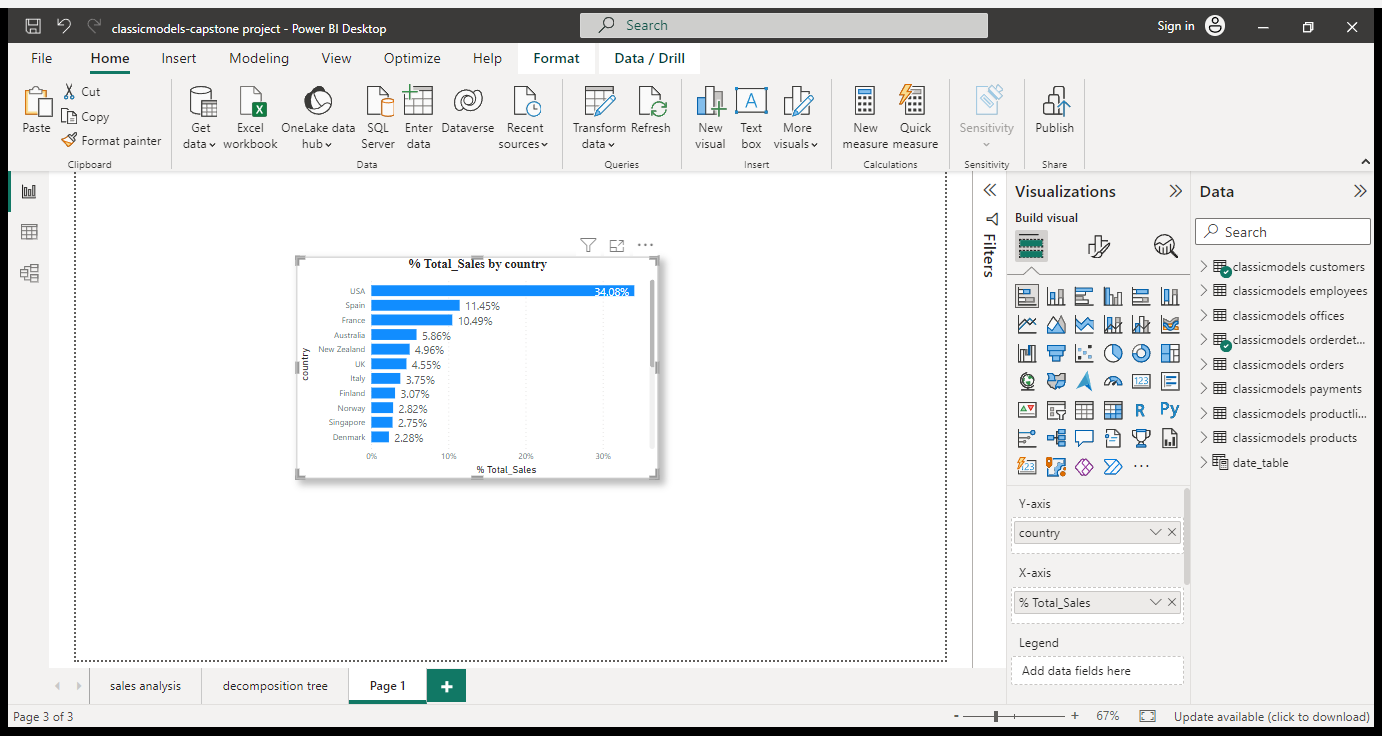
* Missing values: no missing values in all columns.
* Replace value: replace “null” value with “N/A” in comment col in orders table.
* Merge query: “buyprice” column from “products” table merged in “ordersdetails” table.
* Removed column: remove unwanted columns in required tables.

Some new columns are created such as “profit and Sales” to measure the “total profit and total sales.”

Another table called “Date” is created to extract Year, Month from given dates and linked to “orders date” in “orders” table.

**DATA VISULIZATION:**

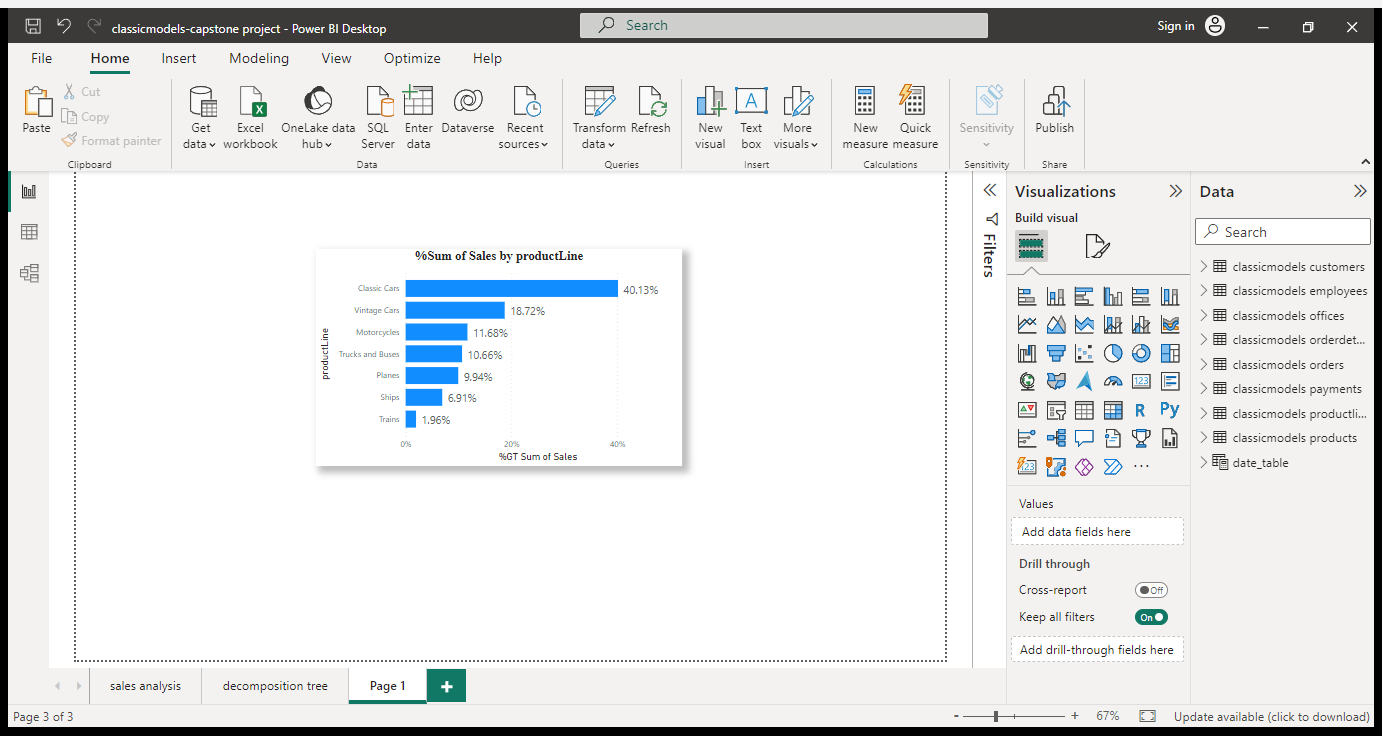
Visualisation would enable us to understand the data better and turns complex data into simple graphs to extract key insights from the visualization and become a storyteller when we speak about insights with data.

1. Stacked Bar Chart:

The above stacked bar chart shows the “%Total sales by country” in which x-axis represents % total\_sales and y-axis represent country.

**Conclusion-1:**

The maximum no. of sales is generated in “USA” with “34.08%” followed by Spain with “11.45%” and France “10.49%” in 2003, 2004 and 2005 respectively.



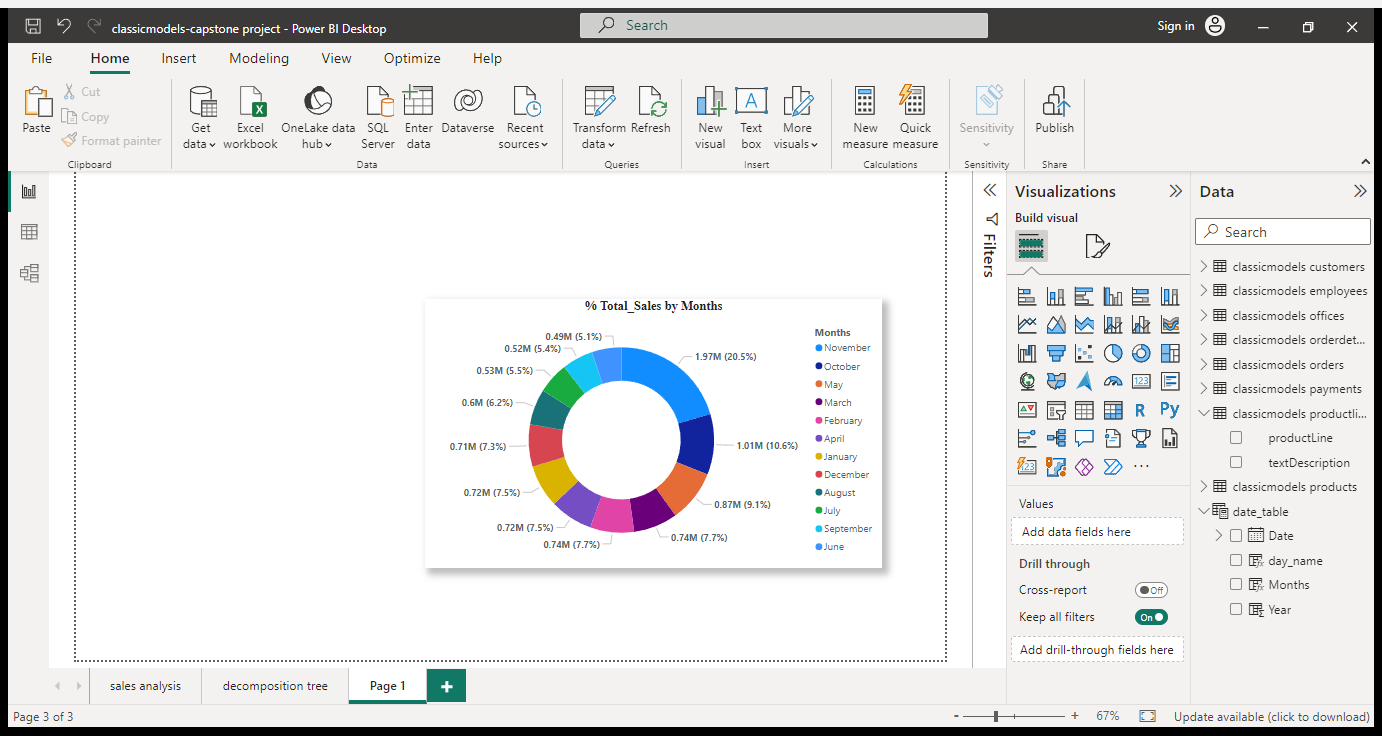
The above stacked bar chart shows the “% total sales by product line” in which x-axis represents % total\_sales and y-axis represent product line.

**Conclusion-2:**

The maximum no. of sales is generated by product line named “Classic cars” with “40.13%” followed by “Vintage cars with “18.72%” and “Motorcycles “with 11.68%” and “Ships and Trains” are having least no. of sales in 2003, 2004 and 2005 respectively.

.

1. Pie Chart:

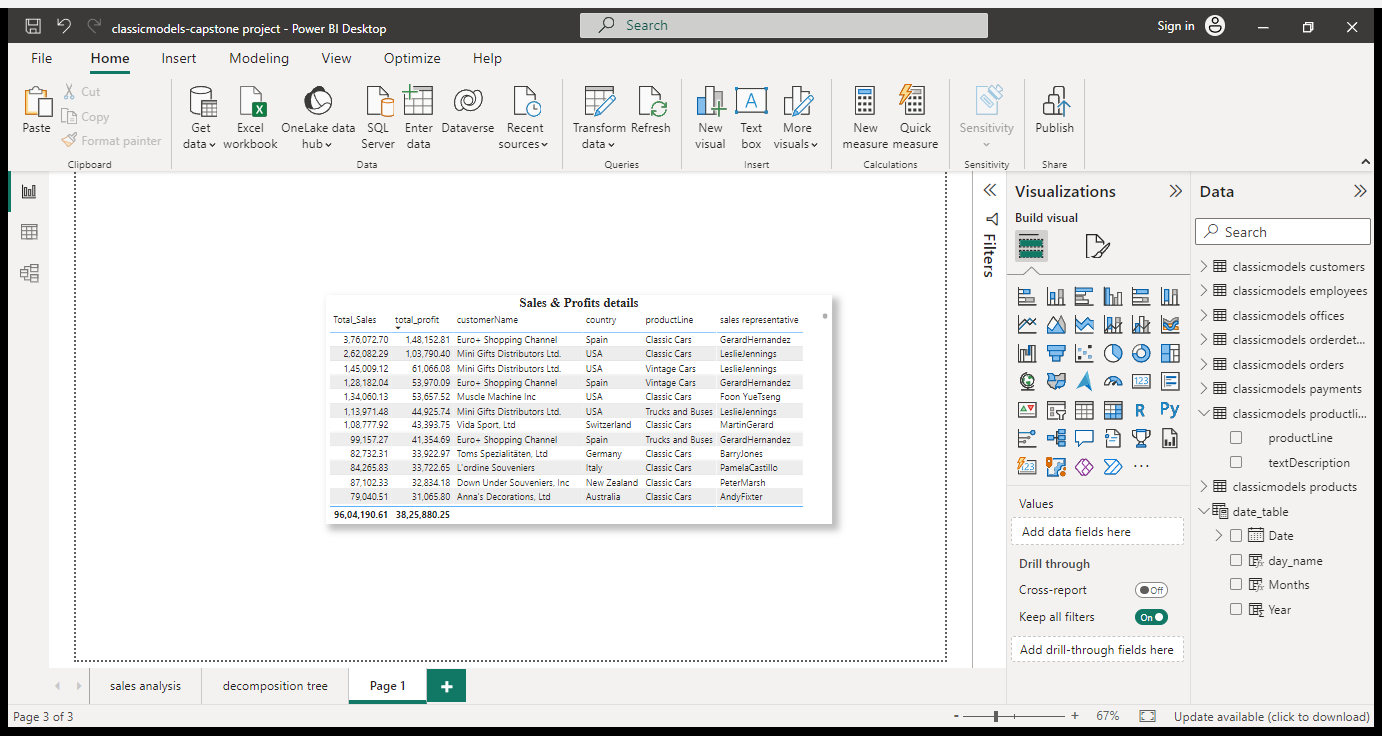


The above Pie chart shows the “% total sales by months” in which value represents % total\_sales and legend represent product line.

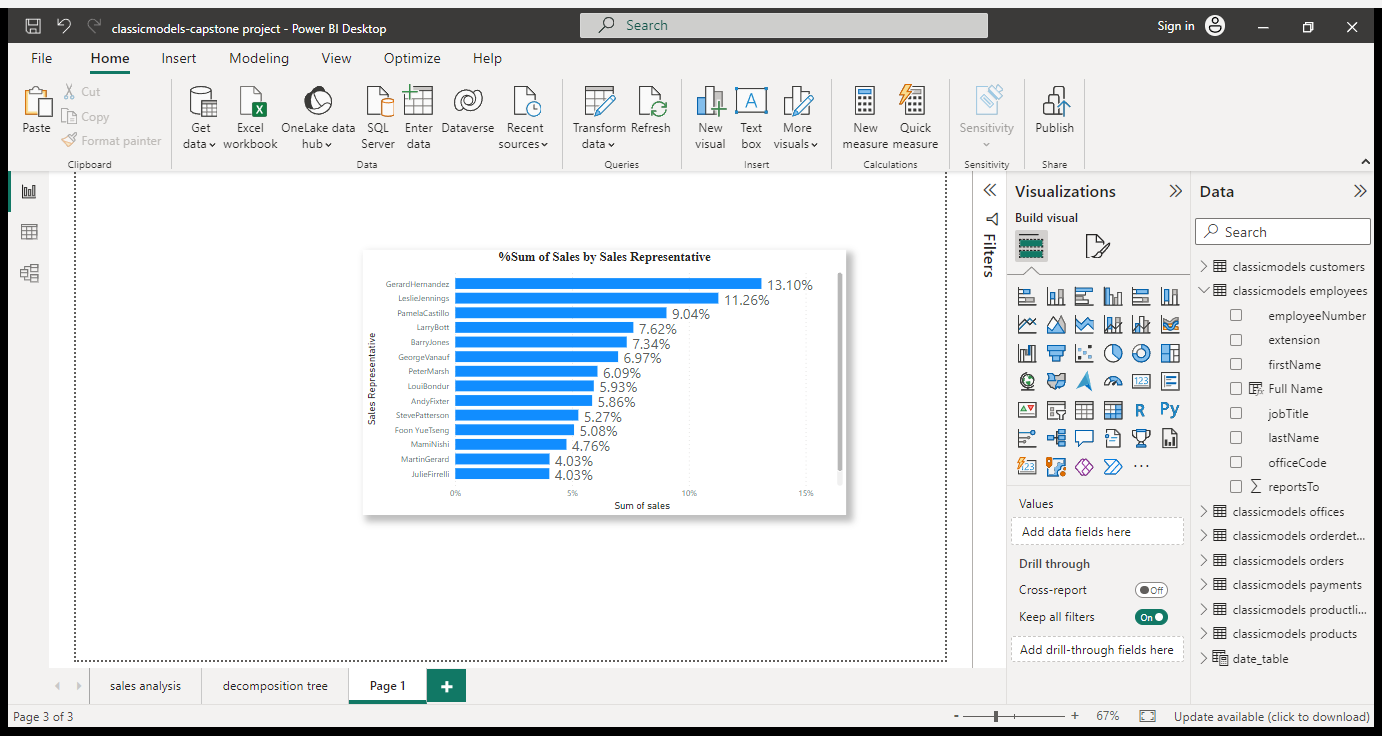
**Conclusion-3:**

The maximum no. of sales is generated in month of “November” with “20.5% (1.97M) sales” followed by “October” with “10.65% (1.01M)” sales and least sales are seen in “June” with “5.1% (0.49M) sales” in 2003 and 2004 respectively.

1. Table:



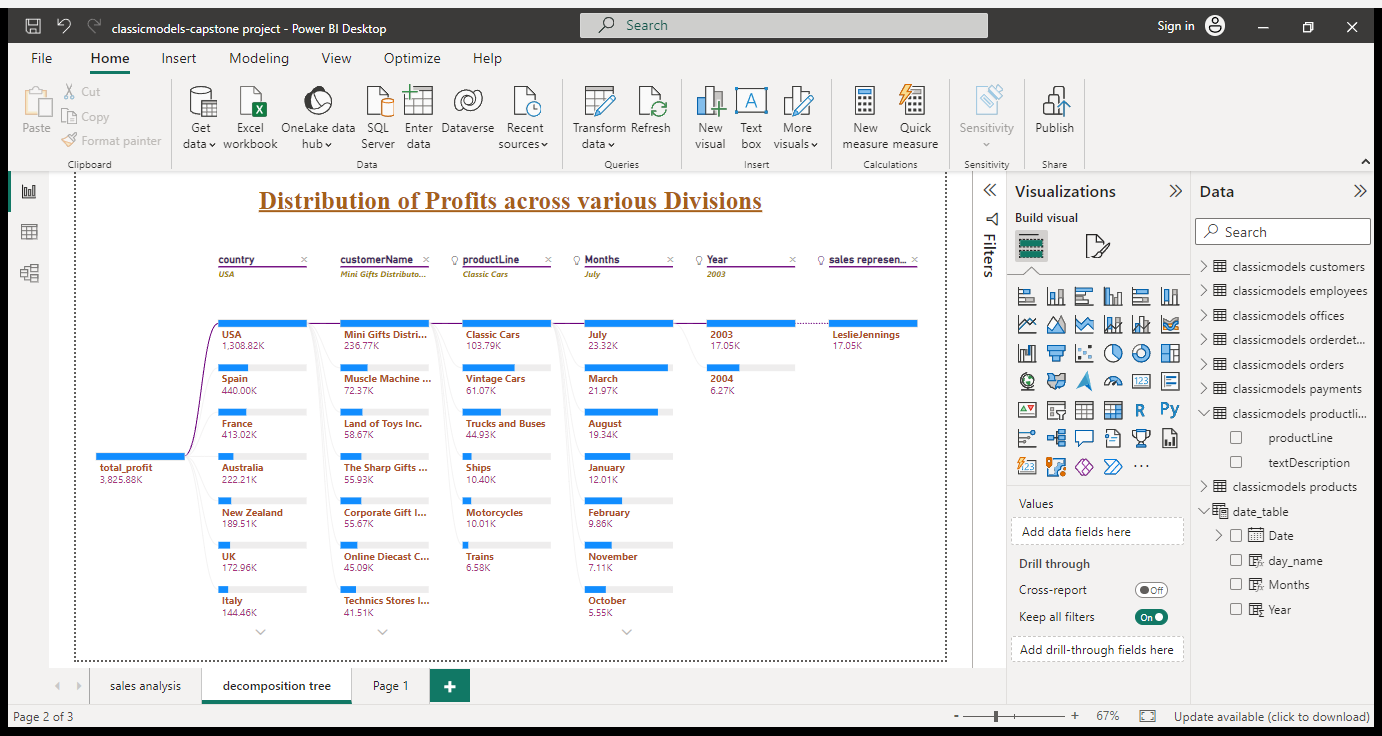
The table represents “Sales & Profit details “of various categories such as customers, country, product line and sales representative.



The table represents “Sum of Sales by Sales Representative “in which value represents % sum of sales and legend represent Sales Representative.

1. Decomposition tree:

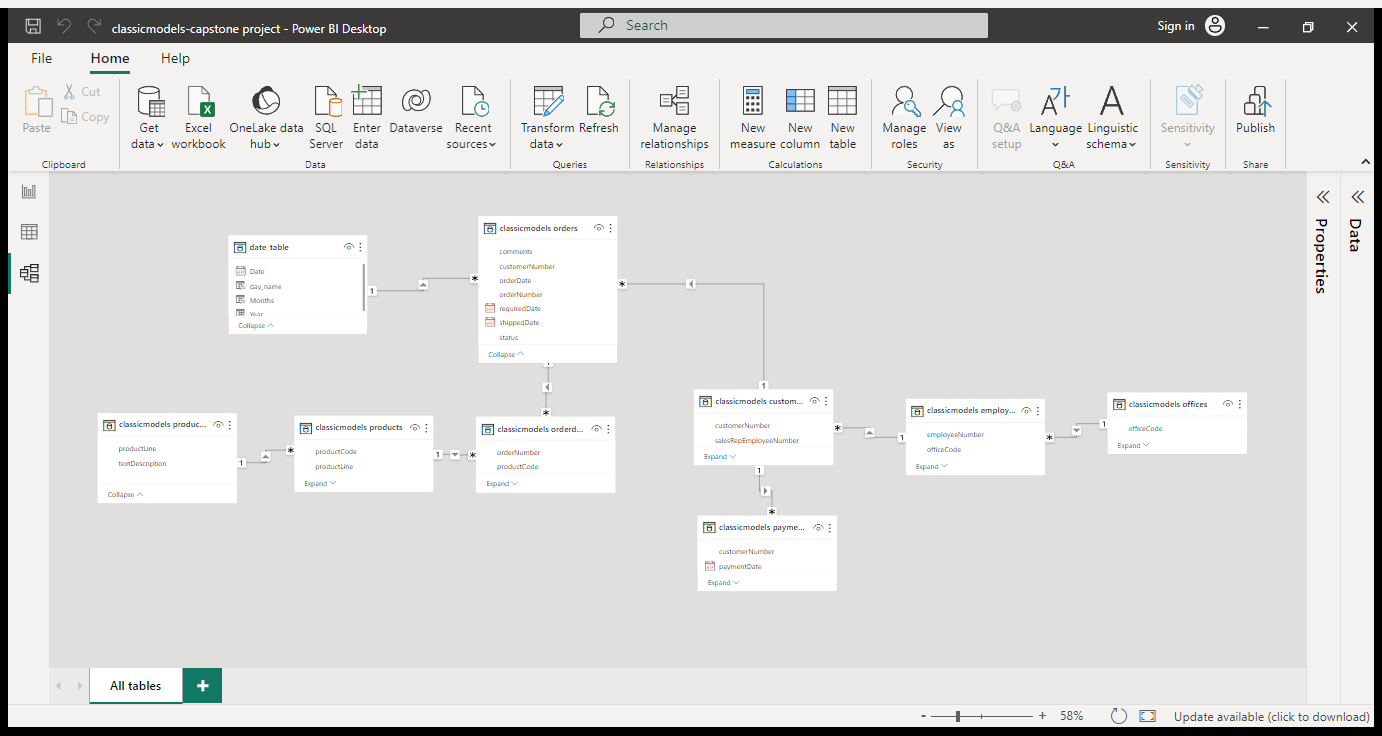
This visual can be decomposed using one metric of the data to analyse the other multiple dimensions. As far as I was concerned, this visual enabled us to see the customer’s journey better. This is a great way to see customer’s details, product’s breakdown in a detailed way.



**Model View in Power BI:**

In model view, it shows the entity-relationship between the dimension tables and fact table.

Here, fact table is “orders table” and dimension tables are date table, orderdetails and customers tables. Again, orderdetails are connected to product table and productline table as well as customers table connected to payment table and employees table followed by offices. The E-R diagram shown in below:



By observing the model view, its concludes that the given database has “Snowflake schema” and the relationship is “many-to-one relations.”

**Key Insights:**

1. Overall, 2003, 2004, and 2005 years, the customer “Euro+ Shopping channel” generates the greatest number of sales of 3.76M with profits of 1.48M belongs to country Spain followed by “Mini Gifts Distributors ltd” with sales of 2.32M and profits 1.02M belongs to country USA.
2. The “Gerard Hernandez” is a sales representative has the greatest number of sales i.e., 13.10% of overall sales followed by “Leslie Jennings” with 11.26% of sales accomplished by considering overall 2003 to 2005 years
3. The month of November was an exception to otherwise relatively consistent orders from 2003 to 2005. This is because of Thanksgiving or early Christmas shopping affecting the sales positively.
4. USA has highest sales of 34.08% followed by Spain and France in which sales are 11.45% and 10.49% seen in all those 2003, 2004 and 2005 years.
5. Out of all product Line, Classic Cars has highest sales of 40.13% followed by Vintage cars sales are 18.72% and Motorcycle sales of 11.68% seen in all those 2003, 2004 and 2005 years. Least product line are ships and trains of 6.91% and 1.96% sales.