

Questions to answer

1. What is the total revenue generated from completed orders by order date?
2. Which products are generating the most revenue from completed orders?
3. How are sales distributed across different sales territories?
4. How has sales performance grown compared to previous years?
5. Which customers are contributing the most to the company's revenue over time?
6. What is the impact of special offers on sales revenue?

Tables needed

All these tables are taken from Sales scheme

Sales SpecialOfferProduct	Sales CreditCard	Sales SalesOrderHeaderSalesReas
Sales Store	Sales SalesTerritory	Sales SalesOrderHeader
Sales SpecialOffer	Sales SalesReason	Sales PersonCreditCard
Sales ShoppingCartItem	Sales SalesPersonQuotaHistory	Sales SalesOrderDetail
Sales SalesTerritoryHistory	Sales SalesPerson	Sales Customer

I used SQL queries for cleaning tables:

```
--DATA CLEANING--
--Remove Duplicates
--Identify Duplicate Rows:
SELECT SalesOrderID, COUNT(*)
FROM Sales_SalesOrderHeader
GROUP BY SalesOrderID
HAVING COUNT(*) > 1;

--Remove Exact Duplicates:
WITH CTE AS (
    SELECT *, ROW_NUMBER() OVER (PARTITION BY SalesOrderID ORDER BY SalesOrderID) AS row_num
    FROM Sales_SalesOrderHeader
)
DELETE FROM Sales_SalesOrderHeader
WHERE SalesOrderID IN (SELECT SalesOrderID FROM CTE WHERE row_num > 1);

--Handle Missing Values
--Identify Missing Values:
SELECT *
FROM Sales_SalesOrderHeader
WHERE SalesOrderID IS NULL;

--Standardize Data Format
--Trim Whitespaces:
UPDATE Sales_SpecialOffer
SET Category = TRIM(Category);

--Standardize Date Formats:
UPDATE Sales_SalesOrderHeader
SET OrderDate = CONVERT(DATE, OrderDate, 23) -- Style 23 is 'yyyy-mm-dd'
WHERE OrderDate IS NOT NULL;

select * from Sales_SalesOrderHeader

--Edit Missing Values:
UPDATE Sales_SalesOrderHeader
SET SalesOrderID = (SELECT AVG(TerritoryID) FROM Sales_SalesOrderHeader
WHERE TerritoryID IS NOT NULL)
WHERE SalesOrderID IS NULL;
```

We can extract a few KPI values by SQL query, for example:

- Total sales

```
-- Total Sales
SELECT
    SUM(so.LineTotal) AS TotalSales
FROM
    Sales_SalesOrderDetail so
```

- Total Quantity Sold

```
-- Total Quantity Sold
SELECT
    SUM(so.OrderQty) AS TotalQuantitySold
FROM
    Sales_SalesOrderDetail AS so
```

- Average Sales per Order

```
-- Average Sales per Order
SELECT
    AVG(so.LineTotal) AS AvgSalesPerOrder
FROM
    Sales_SalesOrderDetail AS so
```

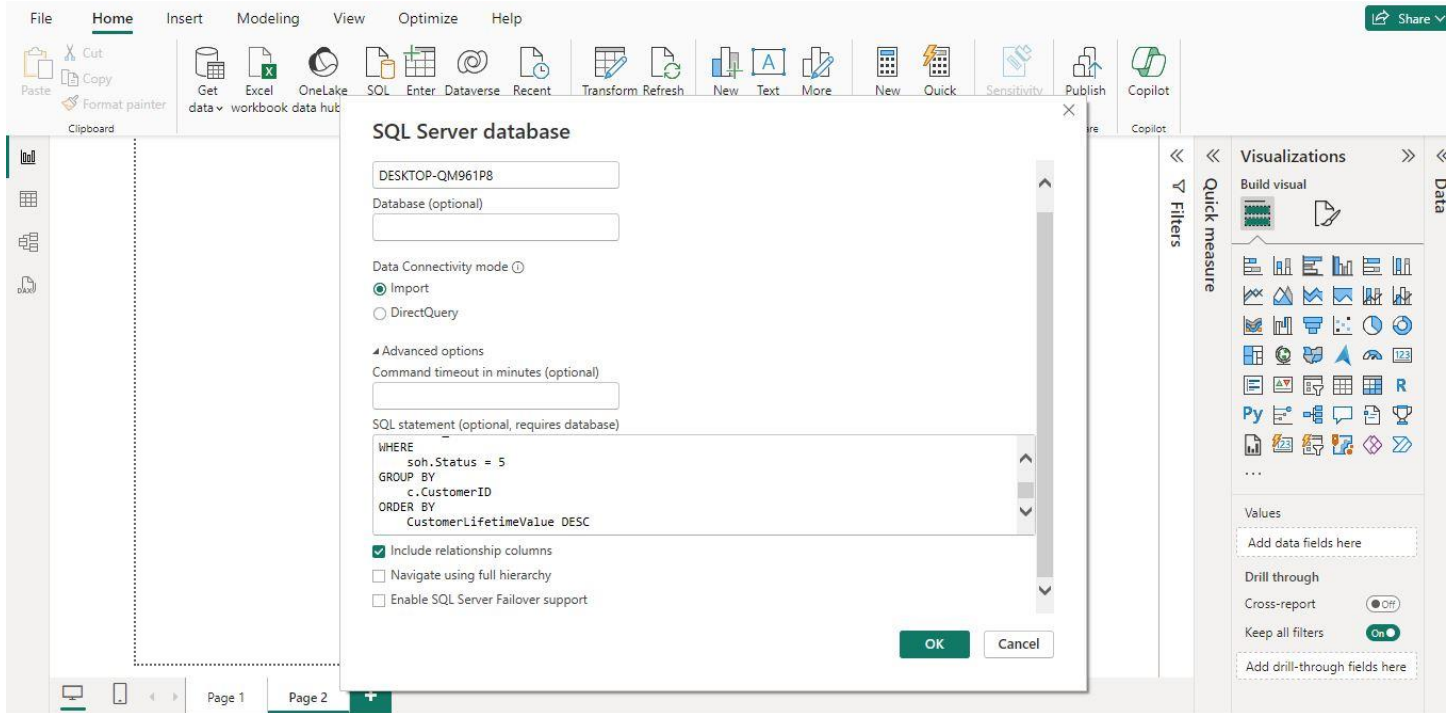
Each of these tables contain some columns that are not necessary for the analysis, so I only selected and cleaned the required columns in the tables.

Load the data to Power BI

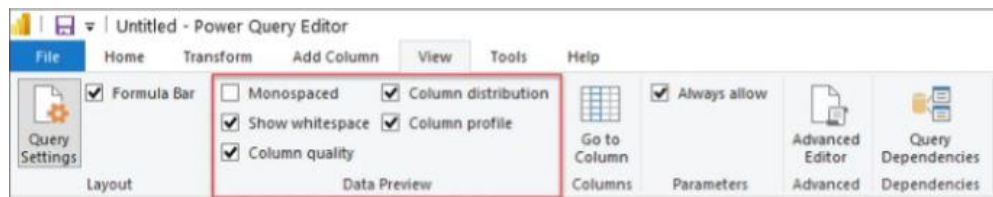
Open Power BI. Get Data. Load the data into Power Query Editor by select (Import data from a SQL server database)

Transform the data in Power Query and In Power Query editor, we can clean and transform each data tables.

- Make sure to rename the tables appropriately with consistency.
- Make sure the data types for all columns are correct and appropriate

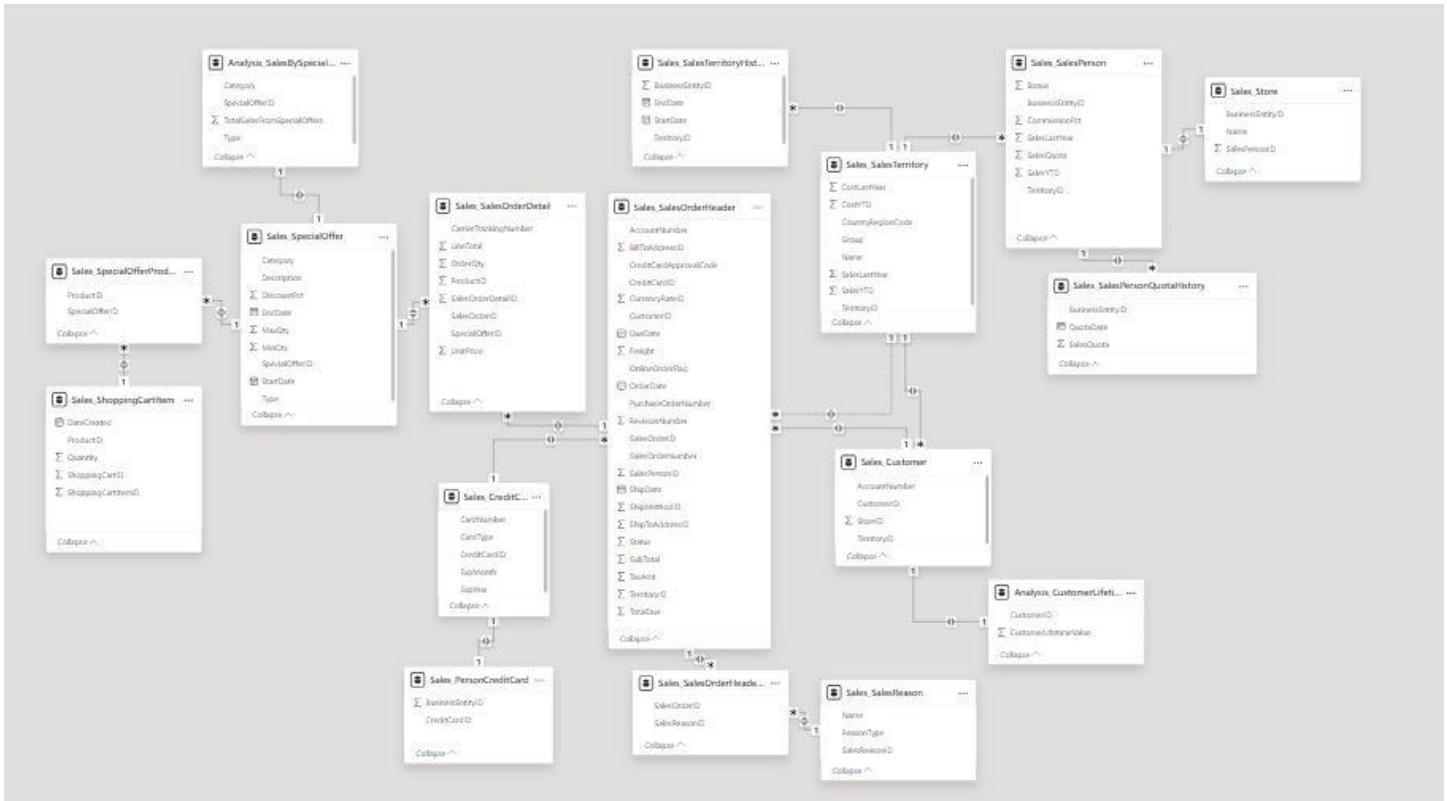


- Use Data profiling tools to help the data transformation. The data profiling tools provide new and intuitive ways to clean, transform, and understand data in Power Query Editor. They include:
 - Column quality
 - Column distribution
 - Column profile



Data Modelling

Once all data are cleaned and loaded into Power BI, we can create a data modelling for the tables. In this project the data model is a Galaxy Schema. A star schema is a type of relational database schema that uses multiple fact tables sharing common dimension tables.



Using SQL queries (Views) for analysis:

- Total Revenue from All Orders

```
SELECT
    soh.OrderDate, SUM(soh.TotalDue) AS TotalSales
FROM
    Sales_SalesOrderHeader soh
WHERE
    soh.Status = 5 -- Completed Orders (Status = 5)
GROUP BY soh.OrderDate
```

Results		Messages	
	OrderDate	TotalSales	
1	2011-06-27	21862.8284000000	
2	2011-09-06	12634.4688000000	
3	2012-05-24	19572.9426000000	
4	2012-08-03	11540.4817000000	
5	2013-04-21	23824.3309000000	
6	2014-03-19	76549.1422000000	
7	2011-06-04	16588.4572000000	
8	2011-09-29	24299.4346000000	
9	2012-05-01	44069.3766000000	
10	2012-08-26	5540.0842000000	

- Sales by Product (Top 10 Products Total Sales)

```
SELECT
    top 10 p.Name,
    SUM(od.LineTotal) AS TotalSales
FROM
    Sales_SalesOrderDetail od
JOIN
    AdventureWorks2022.Production.Product p ON od.ProductID = p.ProductID
JOIN
    Sales_SalesOrderHeader soh ON od.SalesOrderID = soh.SalesOrderID
WHERE
    soh.Status = 5 -- Completed Orders (Status = 5)
GROUP BY
    p.Name
ORDER BY
    TotalSales DESC;
```

133 %

Results Messages		
	Name	TotalSales
1	Mountain-200 Black, 38	4400592.76953125
2	Mountain-200 Black, 42	4009494.73156738
3	Mountain-200 Silver, 38	3693678.04431152
4	Mountain-200 Silver, 42	3438478.87963867
5	Mountain-200 Silver, 46	3434256.9666748
6	Mountain-200 Black, 46	3309673.19067383
7	Road-250 Black, 44	2516857.31494141
8	Road-250 Black, 48	2347655.95751953
9	Road-250 Black, 52	2012447.7824707
10	Road-150 Red, 56	1847818.61669922

- Sales by Store (Top 10 Stores Total Sales)

```
SELECT
    top 10 s.Name,
    SUM(od.LineTotal) AS TotalSales
FROM
    Sales_SalesOrderDetail od
JOIN
    Sales_SalesOrderHeader soh ON od.SalesOrderID = soh.SalesOrderID
JOIN
    Sales_Store s ON soh.SalesPersonID = s.SalesPersonID
WHERE
    soh.Status = 5 -- Completed Orders (Status = 5)
GROUP BY
    s.Name
ORDER BY
    TotalSales DESC;
```

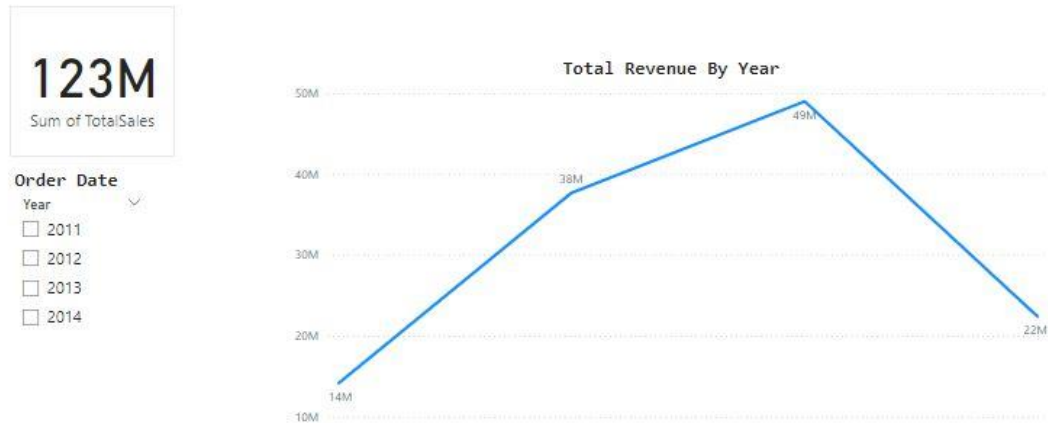
Results Messages		
	Name	TotalSales
1	Sports Products Store	14876896.3335912
2	Friendly Bike Shop	13023848.3349835
3	Aerobic Exercise Company	10367007.4067898
4	Alpine Ski House	10367007.4067898
5	Another Sporting Goods Company	10367007.4067898
6	Basic Sports Equipment	10367007.4067898
7	Bold Bike Accessories	10367007.4067898
8	Brakes and Gears	10367007.4067898
9	Countryside Company	10367007.4067898
10	Eastside Department Store	10367007.4067898

We also have the additional queries data delivered from the dataset.

Dashboards

1. What is the total revenue generated from completed orders by order date?

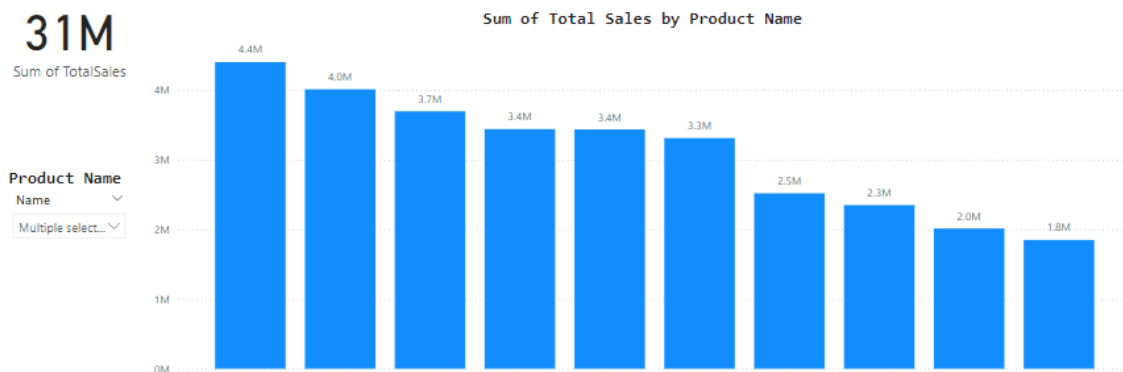
What is the total revenue generated from completed orders by order date ?



Total sales Revenue across all orders over year shows that 2013 is the highest with revenue about 49 M S .

2. Which products are generating the most revenue from completed orders?

Which products are generating the most revenue from completed orders?

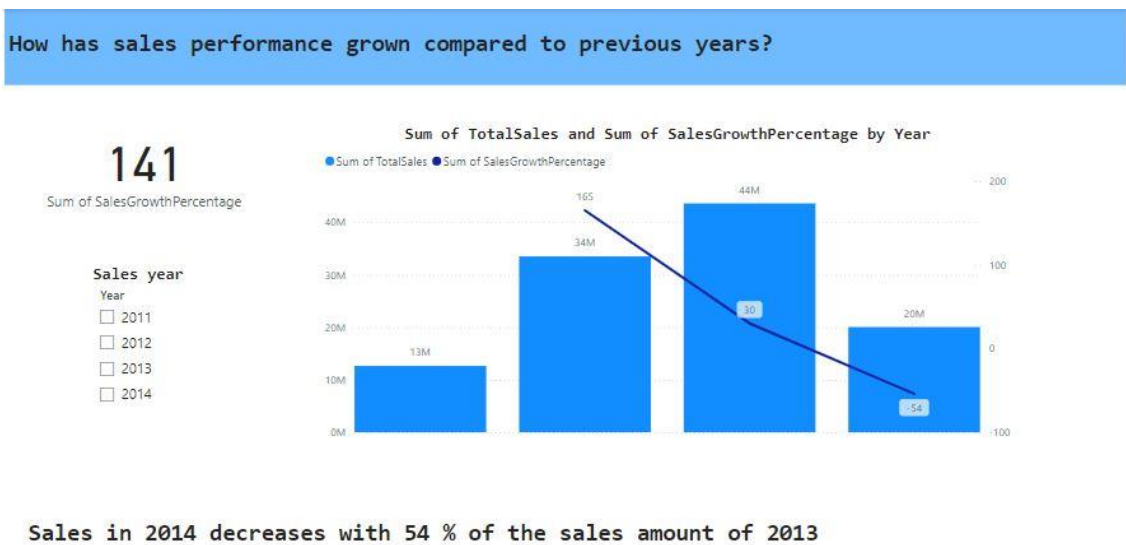


Product (Mountain-200 Black, 38) is the top-performing product by 4.5 M S total sales.

3. How are sales distributed across different sales territories?

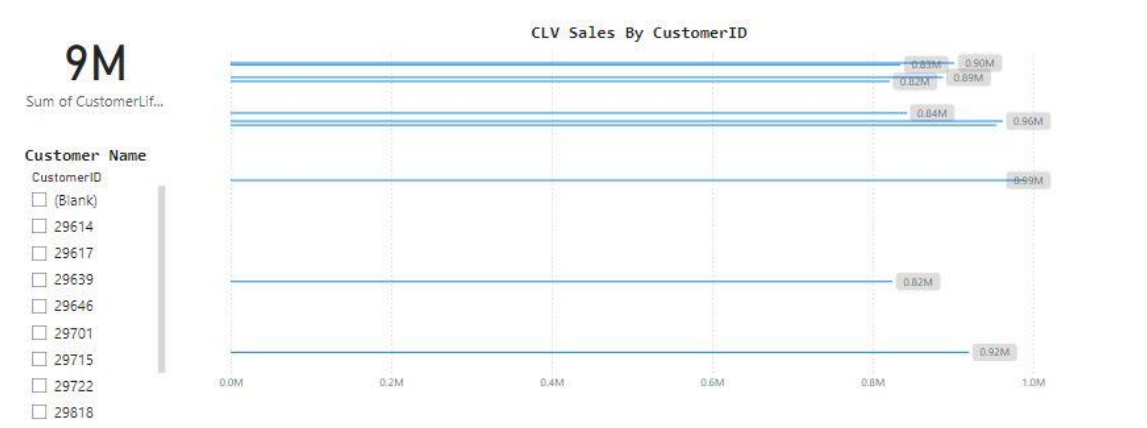


4. How has sales performance grown compared to previous years?



5. Which customers are contributing the most to the company's revenue over time?

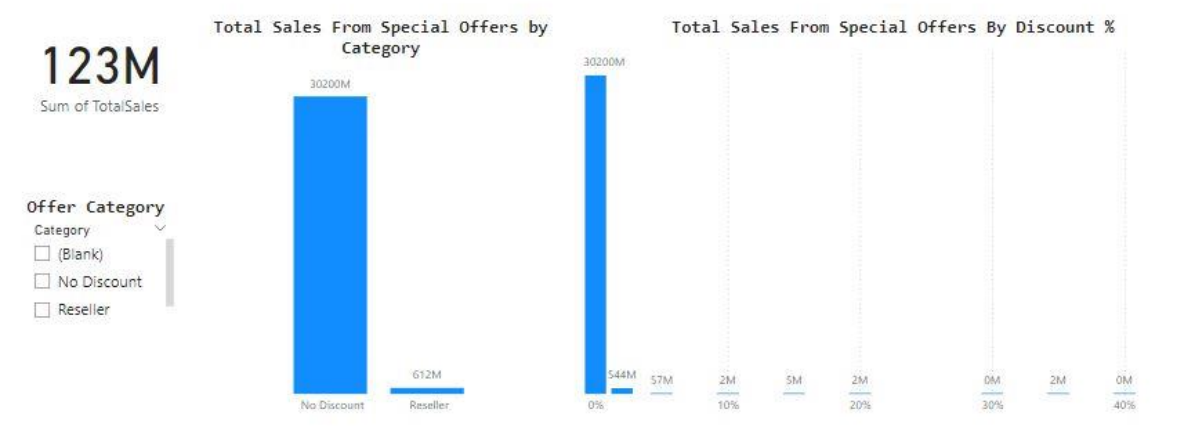
Which customers are contributing the most to the company's revenue over time?



(Customer Name = ID 29818) is the most valuable customer with total sales across all orders equal 0.99 M S.

6. What is the impact of special offers on sales revenue?

What is the impact of special offers on sales revenue?



Total sales Revenue across all orders can not affected by offers. As previewed that the highest total sales come from NO DISCOUNT by 3 B S.