Business Performance Analytics

Introduction:

This project aims to analyze sales trends, customer behavior, product performance, and the impact of marketing campaigns using Power BI. The goal is to gain actionable insights that can help optimize sales strategies, improve marketing efforts, and manage inventory effectively. The project is based on data provided from various tables including orders, order details, marketing campaigns, suppliers, customers, and inventory.

The insights derived will help in identifying the top-performing products, understanding customer purchasing behavior, and evaluating the effectiveness of marketing campaigns.

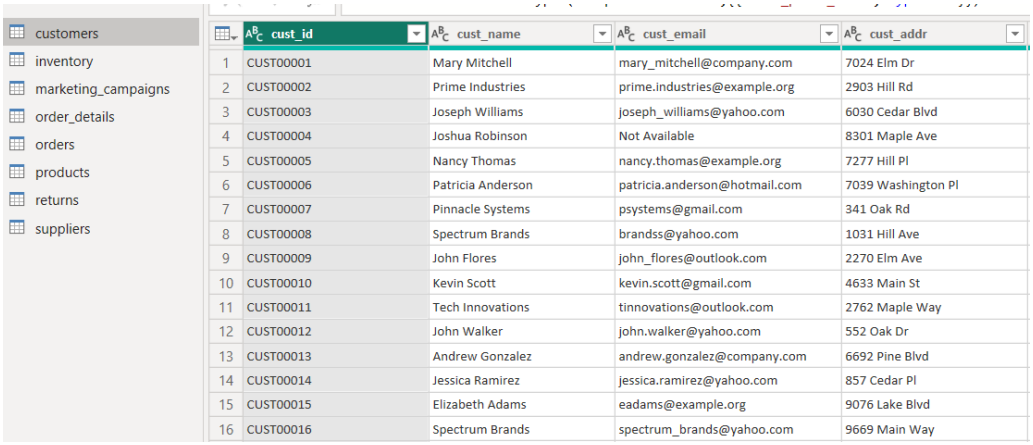
The project utilizes Power BI for data visualization and analysis, while data is sourced from Excel files.

There are few steps before diving deep into analysis:

- Loading the data
- Data Modelling
- Data Preprocessing (data cleaning, removing irrelevant values, handling outliers, standardizing the data)
- Creating new measures
- Data Visualization
- Dashboard creation

Loading the data into Power BI:

The very first step is to get the data and load into the BI tool. There are certain sources like getting it from SQL Server, Excel, Flat Files or Data Warehouses. In our case we have excel files and that can be loaded directly.


customers
inventory
marketing_campaigns
order_details
orders
products
returns
suppliers

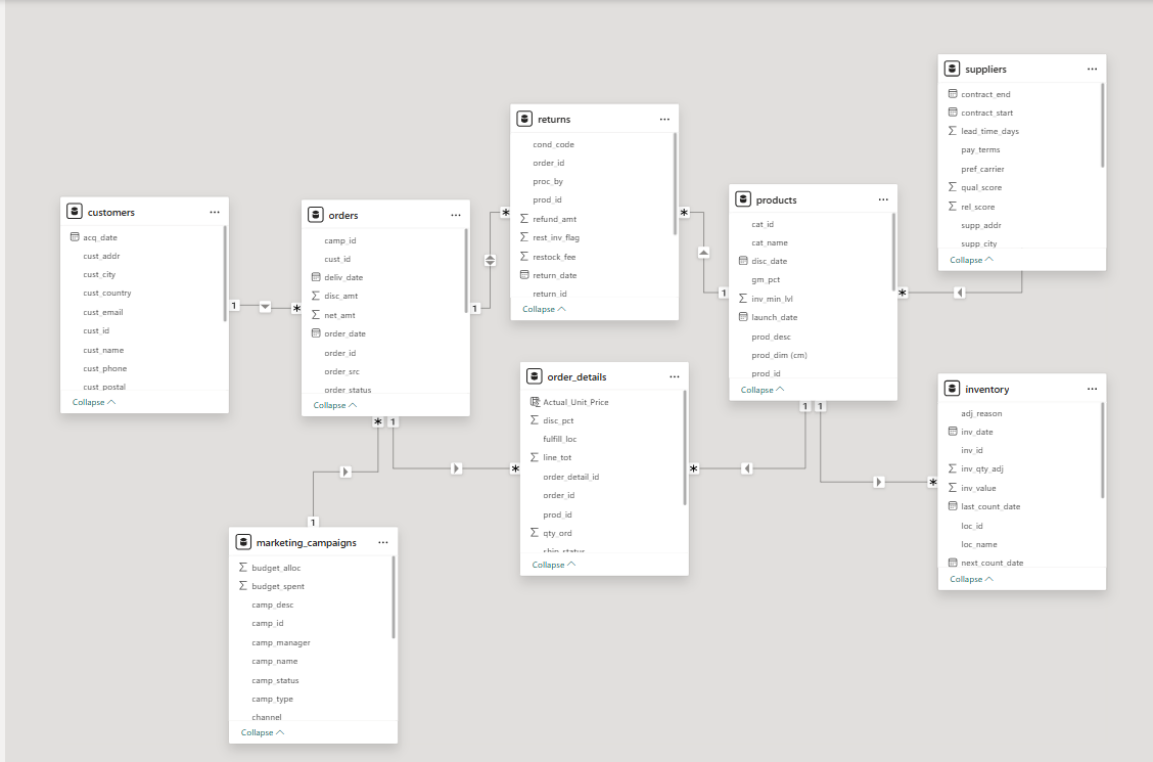
	cust_id	cust_name	cust_email	cust_addr
1	CUST00001	Mary Mitchell	mary_mitchell@company.com	7024 Elm Dr
2	CUST00002	Prime Industries	prime.industries@example.org	2903 Hill Rd
3	CUST00003	Joseph Williams	joseph_williams@yahoo.com	6030 Cedar Blvd
4	CUST00004	Joshua Robinson	Not Available	8301 Maple Ave
5	CUST00005	Nancy Thomas	nancy.thomas@example.org	7277 Hill Pl
6	CUST00006	Patricia Anderson	patricia.anderson@hotmail.com	7039 Washington Pl
7	CUST00007	Pinnacle Systems	psystems@gmail.com	341 Oak Rd
8	CUST00008	Spectrum Brands	brandss@yahoo.com	1031 Hill Ave
9	CUST00009	John Flores	john_flores@outlook.com	2270 Elm Ave
10	CUST00010	Kevin Scott	kevin.scott@gmail.com	4633 Main St
11	CUST00011	Tech Innovations	tinnovations@outlook.com	2762 Maple Way
12	CUST00012	John Walker	john.walker@yahoo.com	552 Oak Dr
13	CUST00013	Andrew Gonzalez	andrew.gonzalez@company.com	6692 Pine Blvd
14	CUST00014	Jessica Ramirez	jessica.ramirez@yahoo.com	857 Cedar Pl
15	CUST00015	Elizabeth Adams	eadams@example.org	9076 Lake Blvd
16	CUST00016	Spectrum Brands	spectrum_brands@yahoo.com	9669 Main Way

Data Modelling:

After loading the data, the next step is to ensure proper data modelling. We must check that the tables are connected with perfect relationships, ensuring that all relevant tables are properly linked through common fields such as Order_ID,

Product_ID, and Customer_ID. This step is crucial for creating accurate and efficient queries and reports, as it allows data from different tables to be combined seamlessly. It's important to define the relationships between the tables correctly for optimal performance in Power BI.

Though Power BI can automatically detect the relationships and cardinality between the tables but you can't completely rely on it. Like, in this case "marketing_campaigns" table was not connected to any table so, I made sure to connect it properly.



Data Preprocessing:

Though every step is crucial in Data Analysis project but Data Preprocessing is most crucial one. If not done properly, it may lead to inaccurate insights, faulty conclusions, and unreliable results. Poor data quality can cause misinterpretation of trends, and ultimately affect decision-making.

Handling missing values:

First we did basic checks and find out there were some missing values and there are some parameters to handle that. If the missing values are very large we cannot just remove them as it will effect our dataset, rather then we can impute them with mean, median or mode depending on the type of data. If missing values are categorical, replacing them by mode is the best option here. If its numerical, then there are two possibilities, either replace it with mean or median.

If the distribution of data is normal we can replace it mean because in normal distribution mean and median are almost same. If the data is skewed, median will be the best option here.

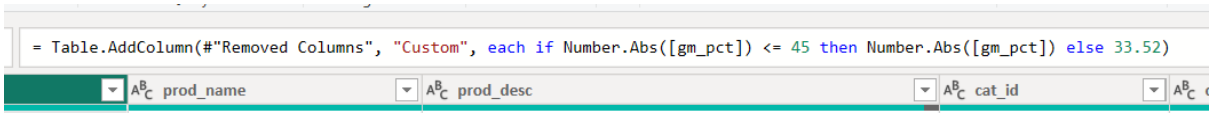
Standardizing the Format and Data types:

There were columns like “cust_phone”, “email” and date, which were not standardized like in email address they were like “umar_11211#yahoo.com”. Phone numbers were like +1-674-303-9928, +1-(674)-303-9928. I standardize them in simple format i.e, 674-303-9928.

Handling outliers:

An outlier is a data point that significantly differs from the rest of the data in a dataset. It is an observation that lies far away from other observations, either much higher or lower than most of the values. Outliers can affect the results of analysis, so it's important to detect them.

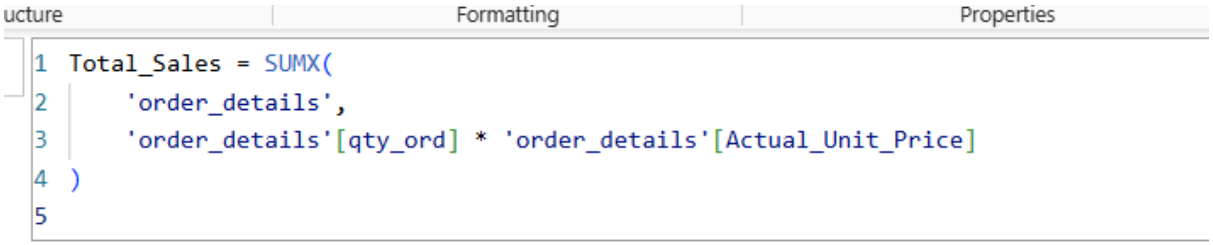
In our case we have some outliers in the “gm_pct” column of “products” table and replaced them with median value.



Creating new measures:

After cleaning the data, the next step is to create new measures from existing columns to dive deep into analysis. I created new measures like “Total_Sales”, “Total_Orders”, “Gross_Profit”, “Total_Products_Sold” ,“Sales before and after Marketing Campaigns” and “Actual_Unit_Price”(due to discounts in some products) etc using DAX expressions.

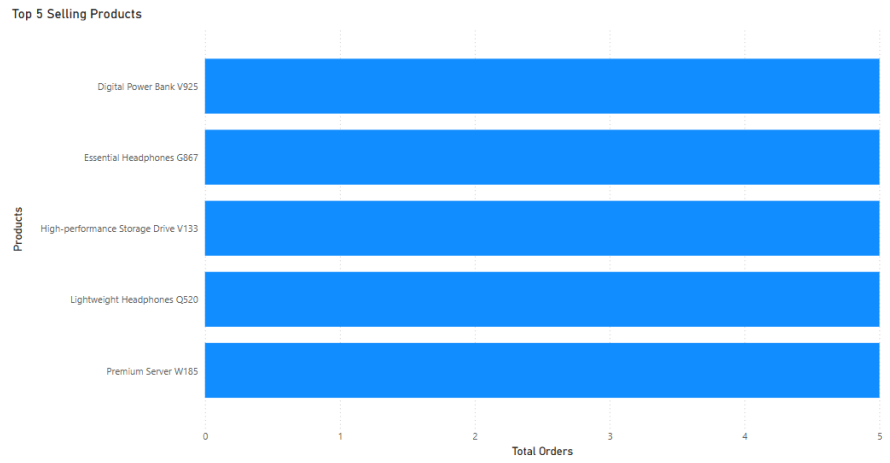
We should go with creating “New Measure” option rather than creating new columns because creating new columns will increase the complexity of the dataset.



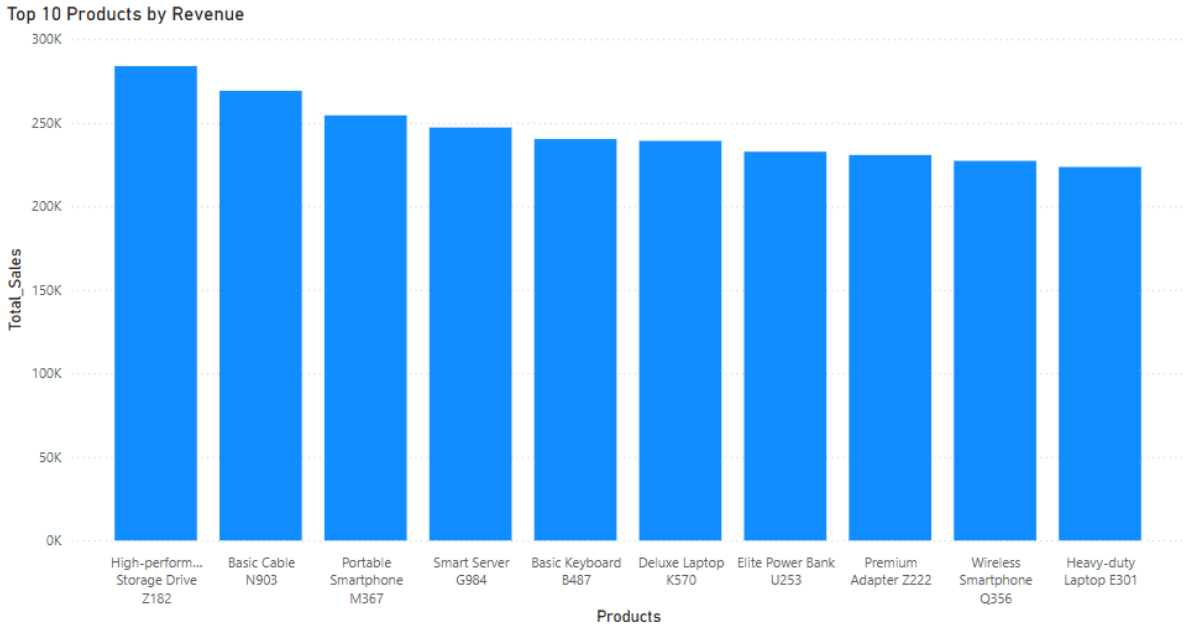
Data Visualization:

Created different kinds of graphs and charts to get insights.

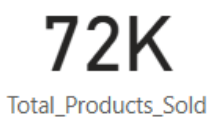
Top 5 Selling Products:



Top 10 Products by Revenue:



Total Products Sold:



Total Sales:

49.94M

Total_Sales

Total Orders:

8000

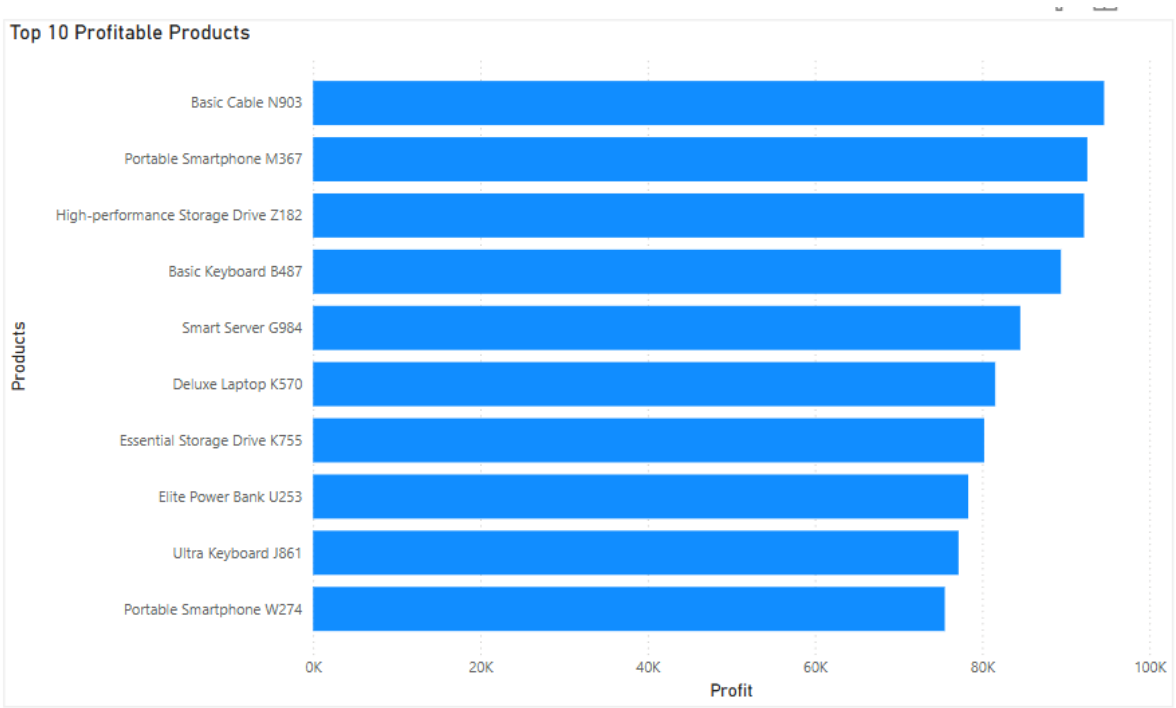
Total_Orders

Gross Profit:

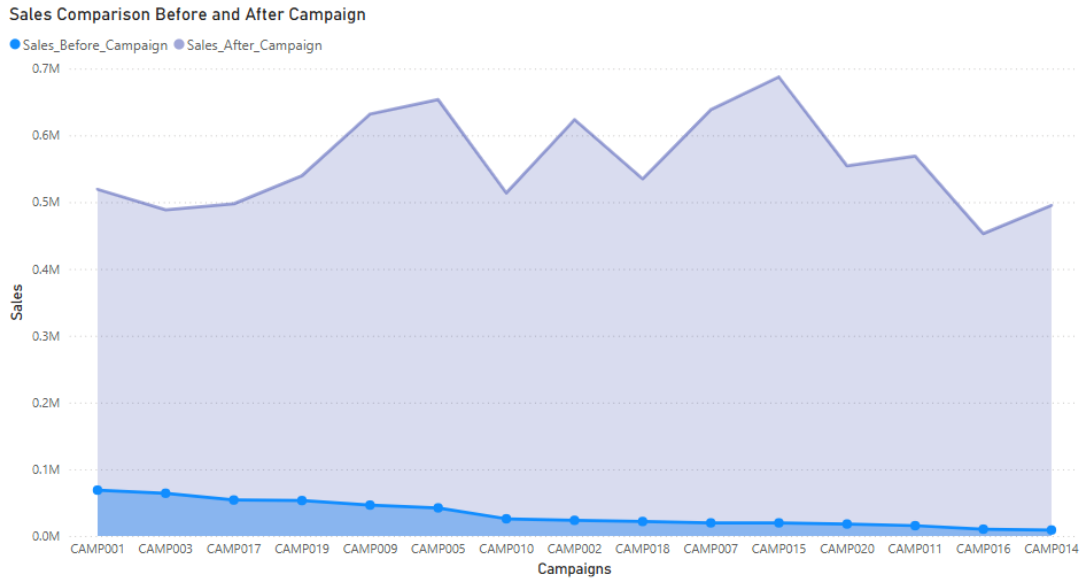
13.82M

Gross_Profit

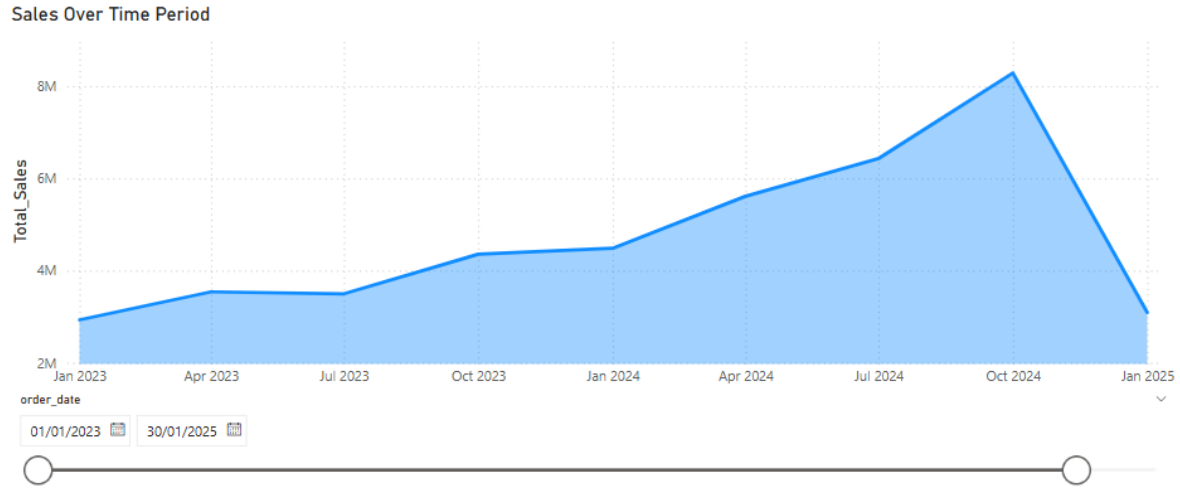
Top 10 Profitable Products:



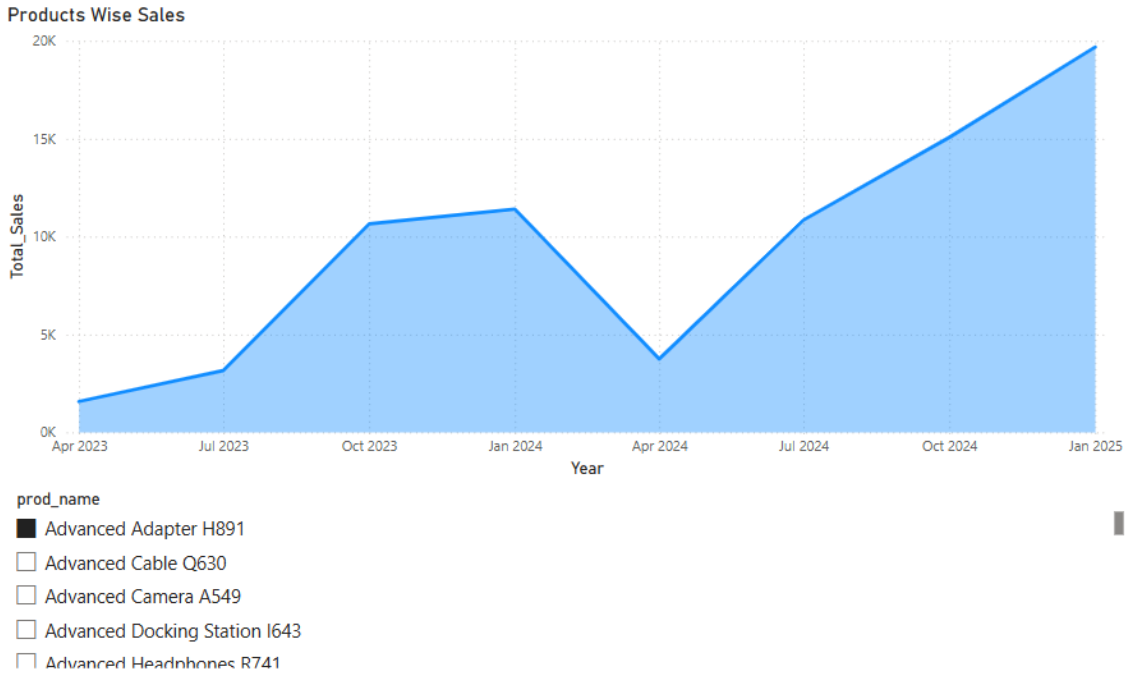
Sales Comparison before and after Campaign:



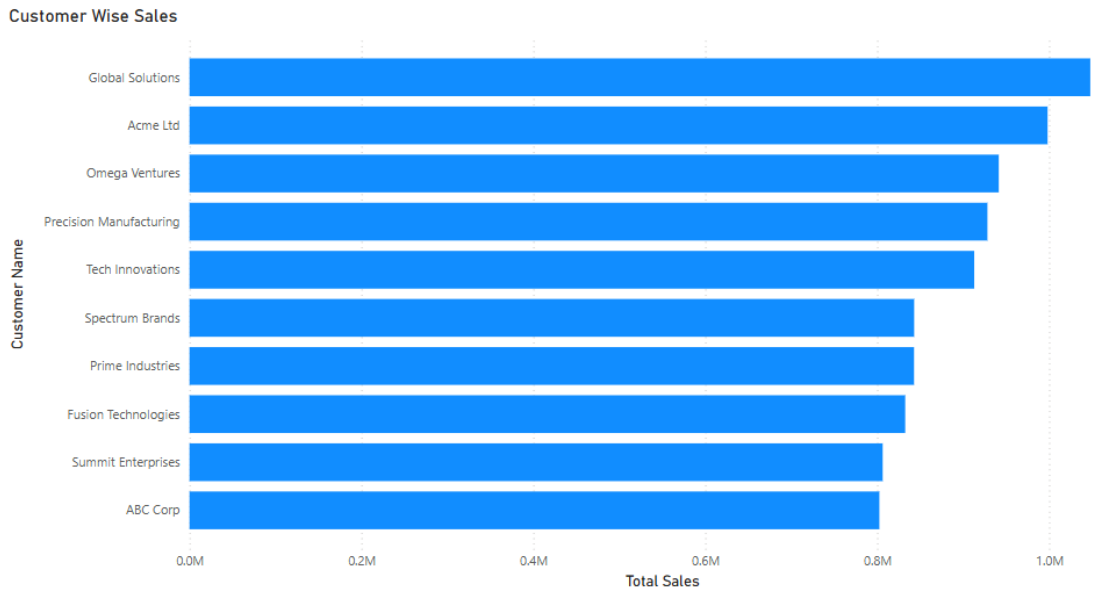
Sales over Time Period:



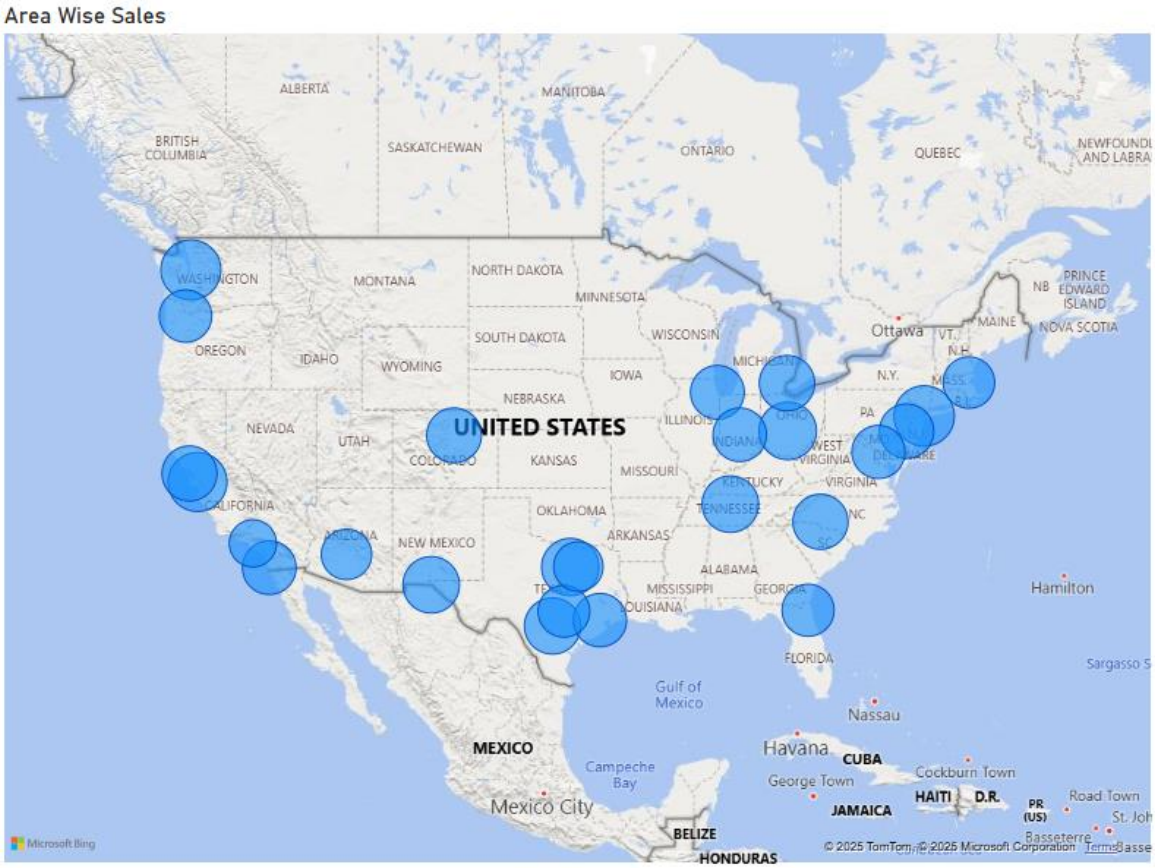
Products Wise Sales:



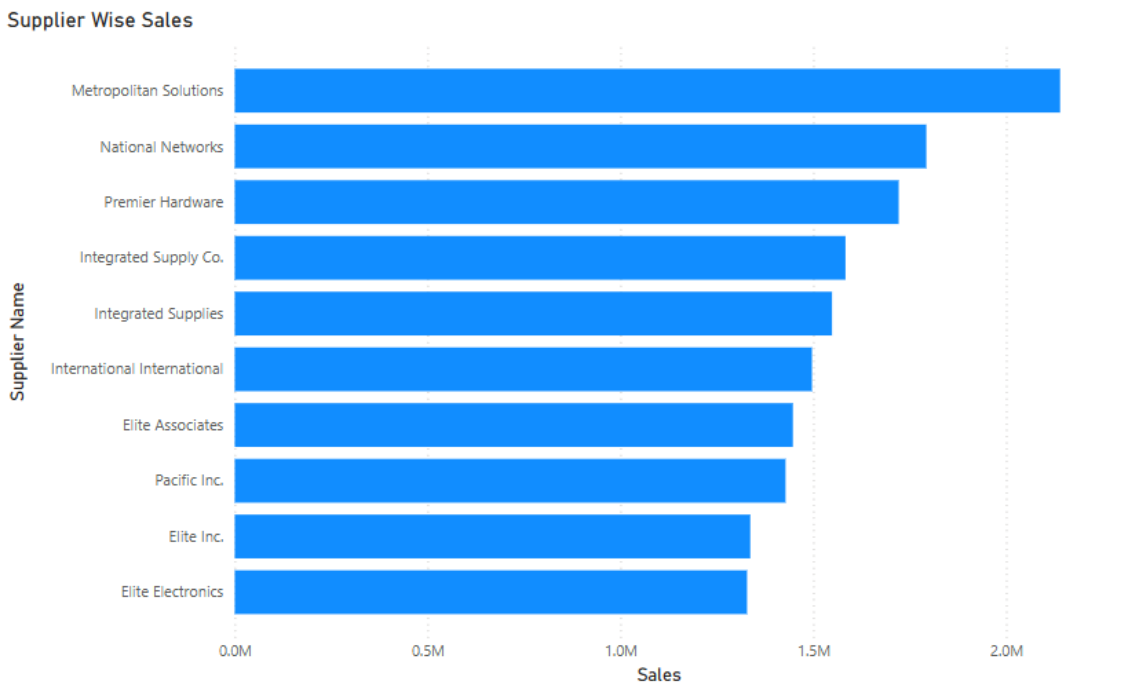
Customer Wise Sales:



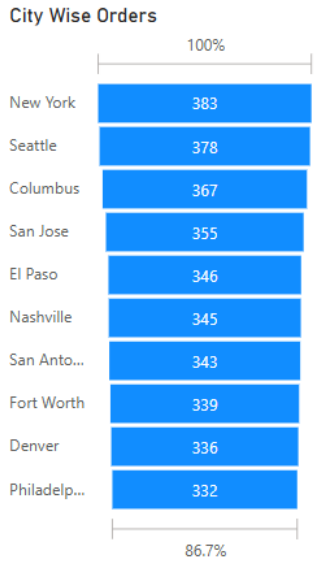
Area Wise Sales:



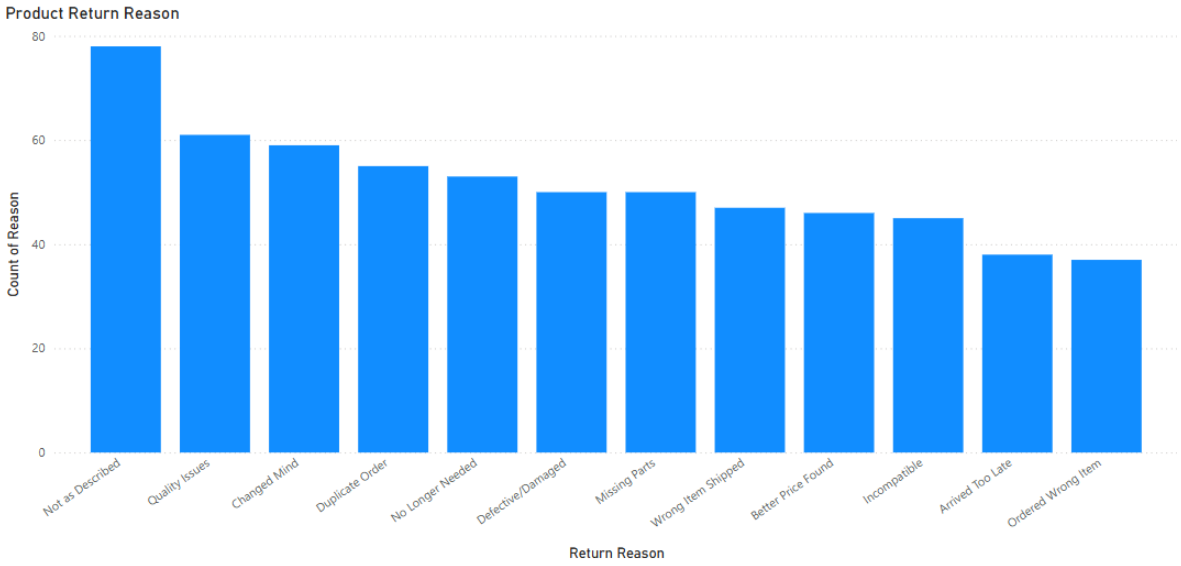
Supplier Wise Sales:



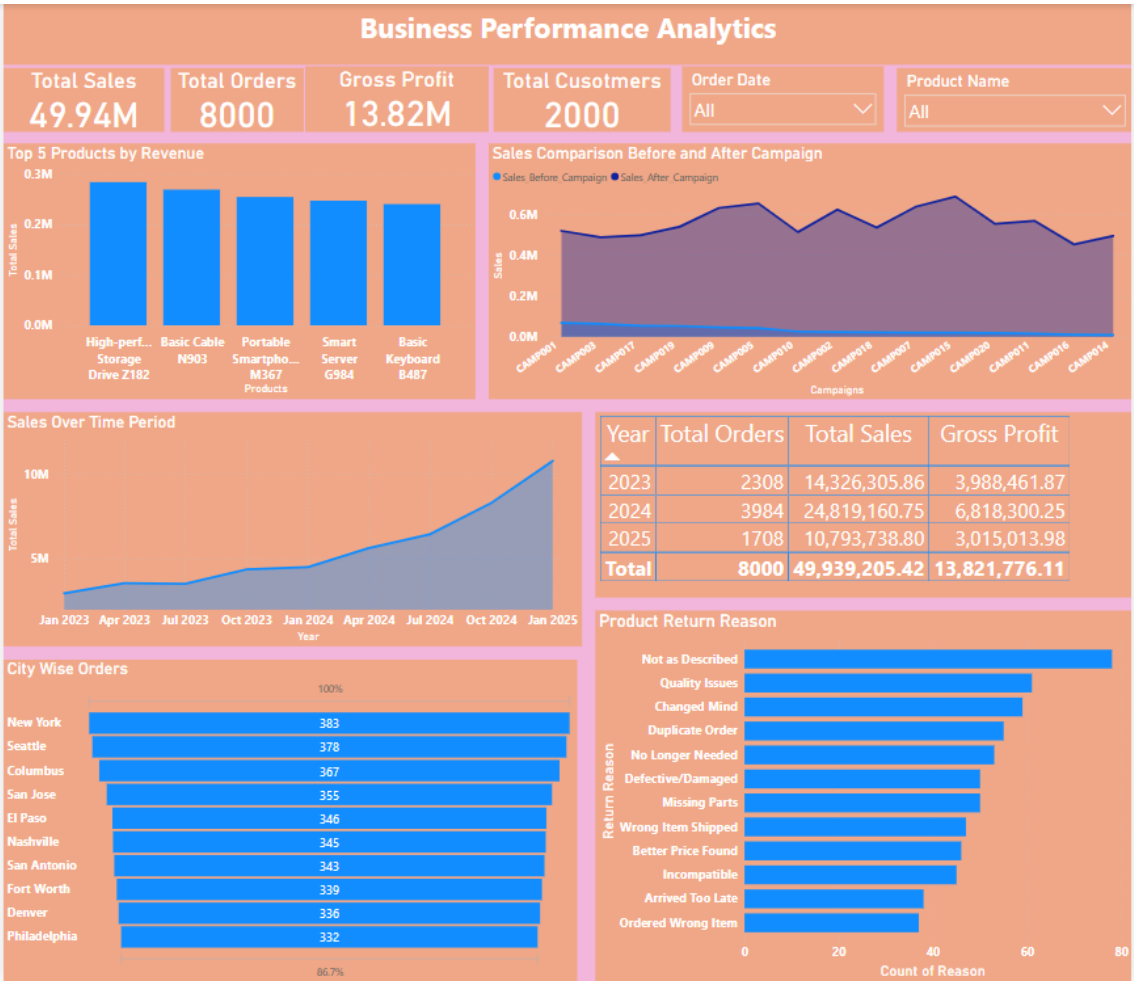
City Wise Orders:



Product Return Reason:



Dashboard:


Business Performance Analytics
Total Sales
49.94M
Total Orders
8000
Gross Profit
13.82M
Total Cusotmers
2000
Order Date
All
Product Name
All
Top 5 Products by Revenue
0.3M
0.2M
0.1M
0.0M
Total Sales
High-perf... Storage Drive Z182
Basic Cable N903
Portable Smartpho... M367
Smart Server G984
Basic Keyboard B487
Products
Sales Comparison Before and After Campaign
Sales_Before_Campaign
Sales_After_Campaign
0.6M
0.4M
0.2M
0.0M
Sales
CAMP001
CAMP003
CAMP017
CAMP019
CAMP009
CAMP005
CAMP010
CAMP002
CAMP018
CAMP007
CAMP015
CAMP020
CAMP011
CAMP016
CAMP014
Campaigns
Sales Over Time Period
10M
5M
Total Sales
Jan 2023
Apr 2023
Jul 2023
Oct 2023
Jan 2024
Apr 2024
Jul 2024
Oct 2024
Jan 2025
Year
Year
Total Orders
Total Sales
Gross Profit
2023
2308
14,326,305.86
3,988,461.87
2024
3984
24,819,160.75
6,818,300.25
2025
1708
10,793,738.80
3,015,013.98
Total
8000
49,939,205.42
13,821,776.11
City Wise Orders
100%
New York
383
Seattle
378
Columbus
367
San Jose
355
El Paso
346
Nashville
345
San Antonio
343
Fort Worth
339
Denver
336
Philadelphia
332
86.7%
Product Return Reason
Not as Described
Quality Issues
Changed Mind
Duplicate Order
No Longer Needed
Defective/Damaged
Missing Parts
Wrong Item Shipped
Better Price Found
Incompatible
Arrived Too Late
Ordered Wrong Item
Return Reason
0
20
40
60
80
Count of Reason