Capstone Project

NORTHWIND TRADERS

**PARVEEN BANU.K**

# **OVERVIEW**

In this project, we are working on developing a polished Power BI report and EDA. The central objective is to effectively convey critical performance metrics for Northwind Traders.

The Northwind Traders database typically contains tables and data related to a wholesale distribution company. It includes information about products, customers, orders, employees, suppliers, and other aspects of the business.

Imbued with the context of a global food supply company, the project delves into various operational facets, encompassing inventory control, customer orders, employee supervision, and sales performance evaluation. Participants engage in a series of tasks that entail querying the database, synthesizing diverse reports, and extracting actionable insights to inform strategic decision-making

To make the report even more engaging, we will incorporate interactive visuals and dynamic filters. These features will allow decision-makers to easily extract actionable insights and make well-informed choices. Ultimately, the anticipated impact of this endeavour is a significant transformation in how Northwind Traders leverages data.

**THE PROCESS**

1. **Data Acquisition from GitHub:**

Obtain the requisite dataset from a designated GitHub repository, Contains essential information on Northwind Traders

**2.Data Transformation and Enhancement:**

Execute data transformation procedures to ensure data quality and consistency. The Dataset were cleaned and validated that to ensure Accurate and reliable results for subsequent analysis. Additionally, consider augmenting the dataset with new problem statements to enrich the analysis potential

**3. Connecting with Tools:**

Establish connections between the dataset and various analytical tools. Interface the dataset with Power BI, Excel, and MySQL Workbench, facilitating seamless data integration and processing.

**4. Problem Statement Solution in Power BI:**

Utilize Power BI to delve into the specified problem statements. Employ its robust features for data visualization, exploration, and analysis, effectively deriving insights and solutions*.*

**5. Exploratory Data Analysis (EDA):**

Perform exploratory data analysis using either Excel or SQL Workbench, depending on the complexity of the analysis. Extract meaningful patterns, relationships, and trends from the data to inform subsequent decision-making.

**6. Creation of Visual and Insightful PowerPoint:**

Develop a comprehensive PowerPoint presentation that encapsulates the project's objectives, methodologies, problem statement solutions, and key visualizations. Each problem statement should be accompanied by a dedicated section with pertinent conclusions and insights.

**7. Detailed Documentation:**

Compile a detailed report that meticulously documents the entire project lifecycle. Include sections on data collection, transformation, problem statement formulation, tools integration, Power BI solutions, EDA insights, and PowerPoint visualizations.

# **OBJECTIVE**

The Northwind database contains sales data for a fictitious company called Northwind Traders, which imports and exports specialty foods from around the world. The Northwind Traders project seeks to transform the company by embracing technology and optimizing operations to create a more efficient, customer-focused, and profitable business.

The objective of the Northwind Traders project is to implement a comprehensive business solution that optimizes various aspects of the company's operations, enhances customer satisfaction, and improves overall profitability. The project aims to achieve this by streamlining internal processes, leveraging technology, and delivering a seamless experience for both customers and employees.

The project will involve the following tasks:

* Analysing the Customer acquisition over time, Segmenting the Customer Based on their region and country.
* Examine Employees Performance, Productivity and Employee tenure
* Analysing the orders, order frequency of customer, and finding orders discontinued by the customer
* Performing deep analyse on suppliers and shippers, and inspecting what are the orders were not shipped yet
* Inspecting product quantity, unit price, orders in stock, units on order and finding the reorder level and grouping the products to their categories.
* Deriving meaningful conclusions and recommendations for improving ranking methodologies
* Compiling analysis results, conclusions, and recommendations for stakeholders.

The success of the project will be measured by the following metrics:

* The quality of the analysis
* The relevance of the insights
* The impact of the recommendations

This project is significant because it has the potential to improve the Customer loyalty, profit growth to company and increase the sales. By understanding the factors that influence Customer and Employee Satisfaction can better position themselves to succeed in the global marketplace.

# **Significance**

Northwind Traders, an illustrative example, plays a significant role in the business landscape, offering valuable insights for potential investors, business professionals, and economic analysts. By delving into the factors that influence a company's performance and success, individuals and Organisations can make more informed decisions about Sales, Customer Loyalty, and Profit Growth.

While Northwind Traders is fictional, its attributes and significance offer useful takeaways:

* Northwind Traders' ability to thrive across industries suggests a strong customer-centric approach. Real-world companies can prioritize understanding customer needs and preferences to tailor their products and services effectively.
* The global presence of Northwind Traders underscores the potential of international collaboration. Real-world companies can learn from this by expanding their networks, understanding diverse markets, and leveraging global partnerships.
* Analysing the Sales performance, Identifying Target Customers by grouping them by their Demographics and Geographics.
* We can learn import and Export and what are the things are happened in real world import and export companies.
* Analyse that the season also consider under the orders sold of food products.
* While Northwind Traders is a fictional entity, the lessons derived from its attributes can guide real-world businesses toward growth, innovation, sustainability, and responsible practices.

Conclusion:

Northwind Traders analysis provides valuable insights for prospective students, researchers, and Organisations. It enables individuals to make informed decisions about their Sales and Customer loyalty and Company growth paths. It helps Companies to identify areas for improvement, and promotes data-driven decision-making for enhancing High Sales and competitiveness.

# Data Dictionary

**Table: Customers**

Fields:

* CustomerID: Unique identifier for each customer.
* CompanyName: The name of the customer's company.
* ContactName: The name of the primary contact person.
* ContactTitle: The title or role of the contact person.
* Address: The address of the customer.
* City: The city where the customer is located.
* Region: The region or state where the customer is located.
* PostalCode: The postal code or ZIP code of the customer.
* Country: The country where the customer is located.
* Phone: The phone number to contact the customer.
* Fax: The fax number to contact the customer.

**Table: Employees**

**Fields:**

* EmployeeID: A unique identifier for each employee.
* LastName: The surname or family name of the employee.
* FirstName: The first name of the employee.
* Title: The job title or position held by the employee.
* TitleofCourtesy: An honorific title or form of address used for the employee.
* Birthdate: The date of birth of the employee.
* Hire Date: The date when the employee was hired.
* Address: The residential address of the employee.
* City: The city of the Employee
* Region: The region of the Employee
* PostalCode: The postal associated with the employee's address.
* Country: The country of the Employee
* HomePhone: The personal phone number of the employee.
* Extension: An extension number to reach the employee within the organization.
* ReportsTo: The Employee ID of the supervisor or manager to whom this employee reports.
* Notes: The Education and Qualification details of the Employee

**Table: Orders**

Fields:

* OrderID: A unique identifier for each order.
* CustomerID: A unique identifier for each customer who placed the order. This links the order to a specific customer in Customer Table.
* EmployeeID: A unique identifier for the employee responsible for processing or managing the order. It is a primary key in the employee table.
* OrderDate: The date on which the order was placed by the customer. This helps in tracking when the order was initiated.
* Required Date: The date by which the customer expects the order to be delivered. This provides a deadline for fulfilling the order.
* Shipped Date: The date on which the order was shipped to the customer. This indicates when the order left the warehouse or facility.
* Ship Via: It indicates Ship ID.
* Freight: The cost of shipping and handling the order. This might include charges for packaging, transportation, and other related expenses.
* Ship Name: The name of the person or company who receive the order.
* Ship Address: The address where the order should be delivered.
* Ship City: The city where the order should be delivered
* Ship Region: The region or state where the order should be delivered. This could be a state, province, or other geographical subdivision.
* Postal Code: The postal code of the shipping address.
* Ship Country: The country to which the order should be shipped.

**Table: Orderdetails**

Fields:

* OrderID: Identifier for the order containing the product. It is a Primary key in the Orders Table
* ProductID: Unique Identifier for the product in the order. This links the ProductID in the Product Table
* UnitPrice: Price of a single unit of the product.
* Quantity: Number of units of the product in the order
* Discount: Discount of the Ordered Product

**Table: Products**

Fields:

* ProductID: Unique identifier for each product.
* ProductName: Name of the product.
* SupplierID: Identifier of the supplier providing the product. It is the Primary key in the Supplier Table
* CategoryID: Identifier of the category the product belongs to. It is a Primary key in the Category table
* QuantityPerUnit: Quantity measurement associated with each unit
* UnitPrice: Price of a single unit of the products
* UnitsInStock: Number of units in stock
* UnitsOnOrder: Number of units on order
* Reorderlevel: Reorder level for the products.
* Discontinued: Status indicating if the product is discontinued.

**Table: Categories**

Fields:

* CategoryID: Unique identifier of the Each Category.
* CategoryName: Name of the Category
* Description: Description of the Category

**Table: Shippers**

Fields:

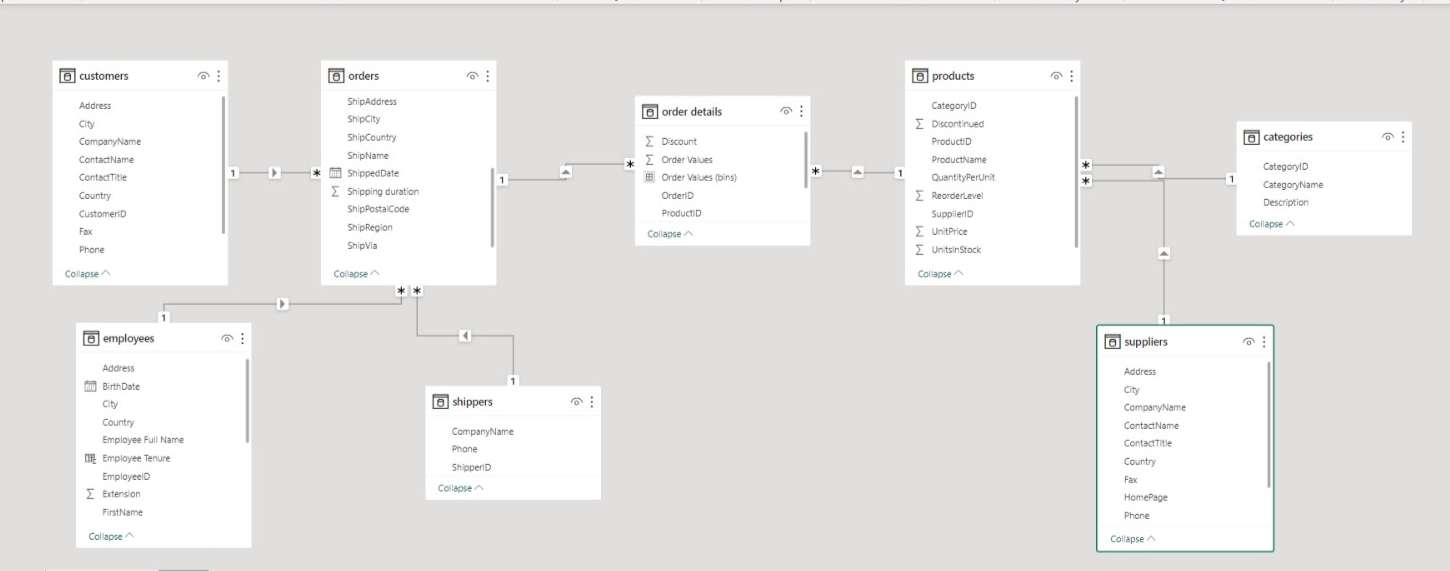
* ShipID: Unique identifier for each shipper.
* CompanyName: Name of the shipping company.
* Phone: Contact phone number of the Shipper.

**Table: Suppliers**

Fields:

* SupplierID: Unique identifier for each supplier.
* CompanyName: Name of the supplier company.
* ContactName: Name of the primary contact person.
* ContactTitle: Title or role of the primary contact person.
* Address: Address of the Supplier.
* City: City where the supplier is located
* Region: Region of the Supplier.
* PostalCode: Postal Code of the Supplier’s Location.
* Country: Country of the Supplier.
* Phone: Contact phone number for the supplier.
* Fax: Fax number of the Supplier.

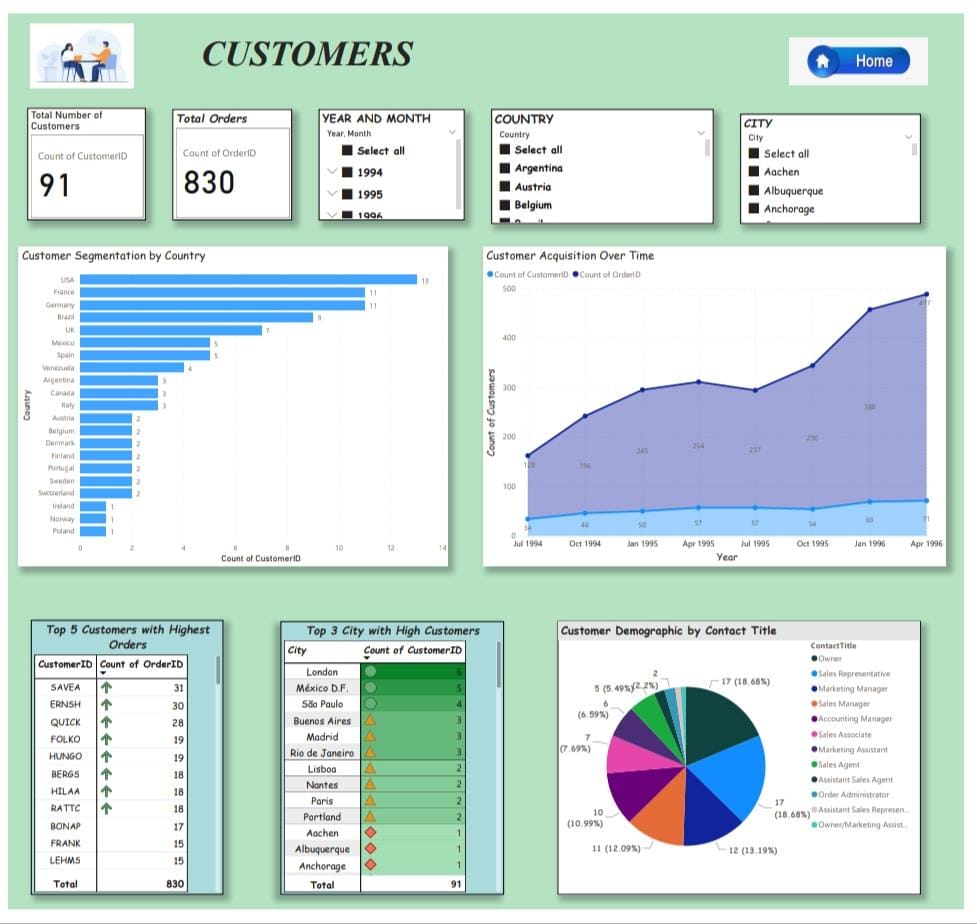
ER Diagram



***Power Bi Problem Statements***

The Northwind database contains the sales data for a fictitious company called “Northwind Traders,” which imports and exports specialty foods from around the world.

***CUSTOMERS:***



This Customer Table stores information about the company's customers. It includes fields for customer ID, company name, contact name, contact title, address, city, region, postal code, country, phone, and fax.

* There are totally 91 Customers.
* The Count of Orders did by the Customers are 830 orders over 3 years.

**How does customer distribution vary across different regions or customer segments?**

Customer segmentation is the tagging and grouping of customers with shared characteristics like age, Industry, gender, Location and interests. Companies employing customer segmentation know that each customer is different and that targeting all their customers with one approach isn't effective. Instead, the organization's marketing efforts would be better served if they target specific, smaller groups with messages those consumers find relevant. By Examining the Customer table, we can make Customer Segmentation by Country. With this we also calculate the Top 3 Countries which has high number of customers. The top 3 countries who have most customers are USA-13 customer, France and Germany – 11 customers and Brazil as 9 customers. With Customer Segmentation we can easily personalise the marketing, service and sales efforts needs to specific Group.

**What is the trend in customer acquisition over time?**

Analysing how many orders were placed by customer for each year mean we can find our growth market. By inspecting customer finally, we conclude The Customers count and orders counts are increased from year 1994 to year 1996.

Customer acquisition is a fundamental aspect of business growth and success. By implementing effective acquisition strategies, businesses can expand their customer base, increase revenue, and establish a competitive advantage.

**Can we visualize the distribution of customer demographics such as age, gender, or income?**

Analysing the customer table there is no data related to age, gender o income. We can take Customer’s Contact title as demographic. By analysing customer’s demographic, we can find that most of the customers belongs to Owner, Sales representative and Marketing manager.

The Customer Demographic data is used by businesses to help them understand the characteristics of the people who buy their products and services.

**Top Five Customers**

Calculating the Count of Orders for each customer id give the Top 5 Customers who Orders the most. This indicates Customers loyalty towards the Company.

|  |  |
| --- | --- |
| CUSTOMERID | ORDER COUNT |
| SAVEA | 31 |
| ERNSH | 30 |
| QUICK | 28 |
| FOLKO, HUNGO | 19 |
| BERGS, HILAA and RATTC | 18 |

**Top 3 City**

For this we can take the Customers count for each City with this we can find Top 3 City with high customers. London, Mexico and São Paulo are the top 3 cities .

***EMPLOYEES***

****

**How does employee productivity vary across different departments or job roles?**

The Employee Productivity can be defined as the amount of work or output produced by am Employee in a Specific Period of time. Workplace productivity is the value each team brings to the success of the overall business.

By analysing we can conclude that Employee’s productivity is vary across the different departments or job roles. Because Sales Representative job role Employee’s handles the max orders comparing with Sales manager, Inside Sales Coordinator and Vice President, Sales. The total orders handle by the Sales representatives were 701.

**What is the distribution of employee tenure?**

Employee tenure means the period of a time a person works at a particular job title o office. Employees who have been with a firm for over five years are long-tenured, while those who have been with the company for less than five years are short-tenured.

We can calculate employee tenure with hire date and there is no relevant data when the employee switch to another organisation. Is there any data available related to dismissal then the employee still working there only. We can use Histogram chart to visualizing the distribution of Employee tenure.

For calculating Employee tenure:

Employee tenure = datediff([hiredate], today (), year)

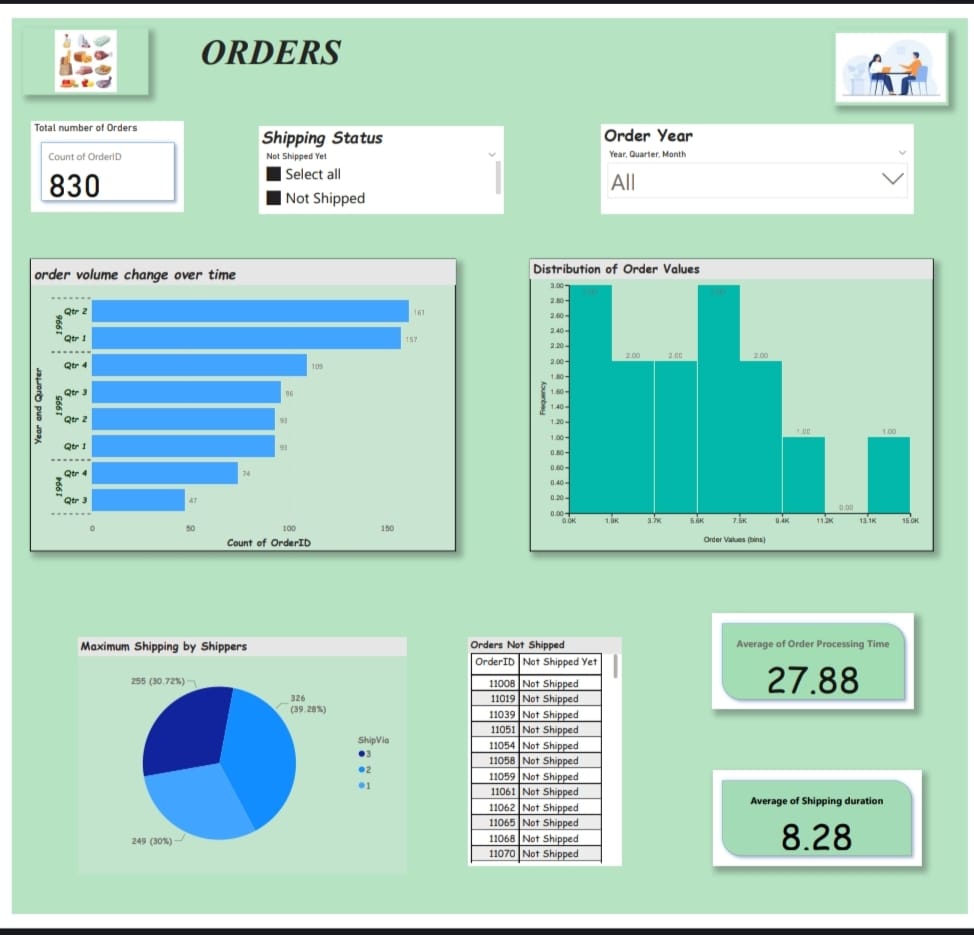
**How many sales were done by each employee by Each year?**

Analyse the Each employee with orders handled by them for each year mean we can find Sales of Employees. By analysing the employee table, the orders handled by the Employee increase and decrease. The Employee Id 4 handles the maximum number of orders compared to other employees.

**Can we visualize employee performance ratings?**

Examining the count of orders handled by the employee, we can give performance ratings. We Can give five-star ratings for the employee Margaret Peacock because he/she handled he maximum number of orders during three years. In 1994, 1995 and 1996 all the three Years the employee ID 4 that is Margaret Peacock Handled the maximum number of orders. The remaining Employees Performance may vary year to year

***Analysing Of Orders***

******

**What is the distribution of order values?**

To determine the distribution of order values from the table, we need o calculate the total value of each order first.

Order value= (unit price \* Quantity)-Discount.

We can Visualize it with the help of histogram for better understanding.

**How to calculate how many orders are shipped by each shipper?**

It is important to analyse what are the products and counts of the products are shipped by shipping company. With the help of shipped date, we can find the order delay and does It cause Customers discontinuation. Customers satisfaction is very important for business growth. Late shipping may cause customers dissatisfaction. The shipper ID 3 ships the maximum number of orders that is 326 orders.

**How to fid what are the orders are not shipped yet?**

The order table contain shipping date for each order. To find the count of orders not shipped yet, we can take null values of the Shipped date. Because the order not yet shipped means they leave the record empty. And also, we have to analyse the reason for not shipping. Which shipping company ships that order and find the order shipping was delay or that order have time to ship.20 orders are not shipped yet.

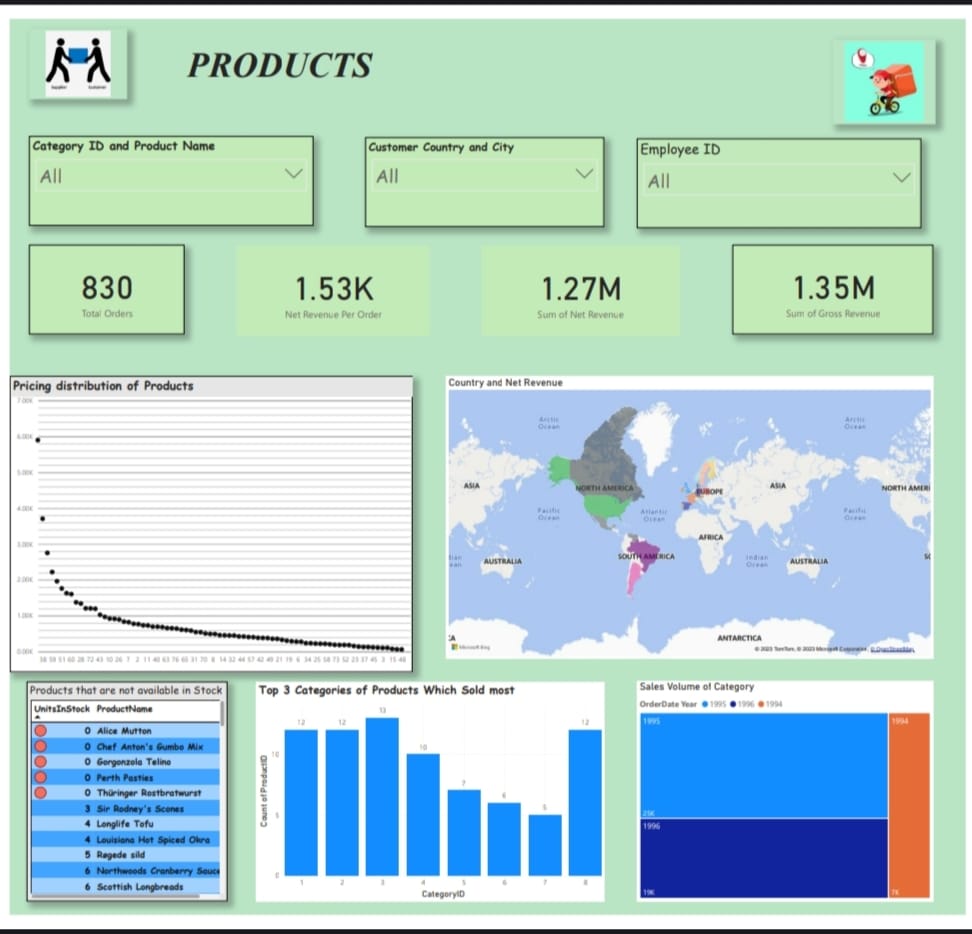
**Can we visualize the average order processing time or shipping duration**

Order Processing Time = Required date – Order date

Average Shipping Duration = Shipped date – Order Date

With this calculation we can find Average order processing time and average shipping duration.

**Analysing Of Products:**

****

**How does the sales volume vary across different product categories?**

Analysing each product Category. Every year the sales volume of the product categories was increase. Sales volume refers to the number of units your company sells during a specific reporting period. This period could be a month, a quarter, or a year depending on what level of sales volume you're seeking to analyse. Investors frequently look at sales volume to assess the health of a growing or contracting company.

**Revenue generated by the products**

Total Orders = 830

Net Revenue Per Order = DIVIDE (sum ('order details'[Net Revenue]), [Total Orders])

**Net Revenue:**

Net revenue, or net income, is equal to a company's gross revenue minus all of its expenses, including fixed expenses.

Sum of Net Revenue = 'order details'[Gross Revenue] - 'order details'[Discount]

**Gross Revenue**:

Gross revenue, also known as gross income, is the sum of all money generated by a business, without taking into account any part of that total that has been or will be used for expenses.

Gross Revenue = ' 'order details'[Unit Price] \* 'order details'[Quantity]

We can display this revenue with the help of cards.

**Can we visualize the pricing distribution of products**

To visualise the pricing distribution of products we have t take product id and unit price. We can Visualise it by using plot box chart.

Pricing is an important decision-making aspect after the product is manufactured. Price determines the future of the product, acceptability of the product to the customers and return and profitability from the product. It is a tool of competition.

**Analyse what are the products are not available in the stock**

It is necessary to know the availability of the products in the stock. The unavailability of the products may lead to delayed delivery and customer dissatisfaction. We can analyse the Products not available in the stock with the help of tables.

Some ways to measure product availability:

Weekly gap scan score – what the associates can see as out-of-stock on the shelf.

Online first-time pick rate – personal shopper pick performance.

Image capture – video and photo of what's available either continuously or at set times of the day.

**Top 3 Product Category Sold most:**

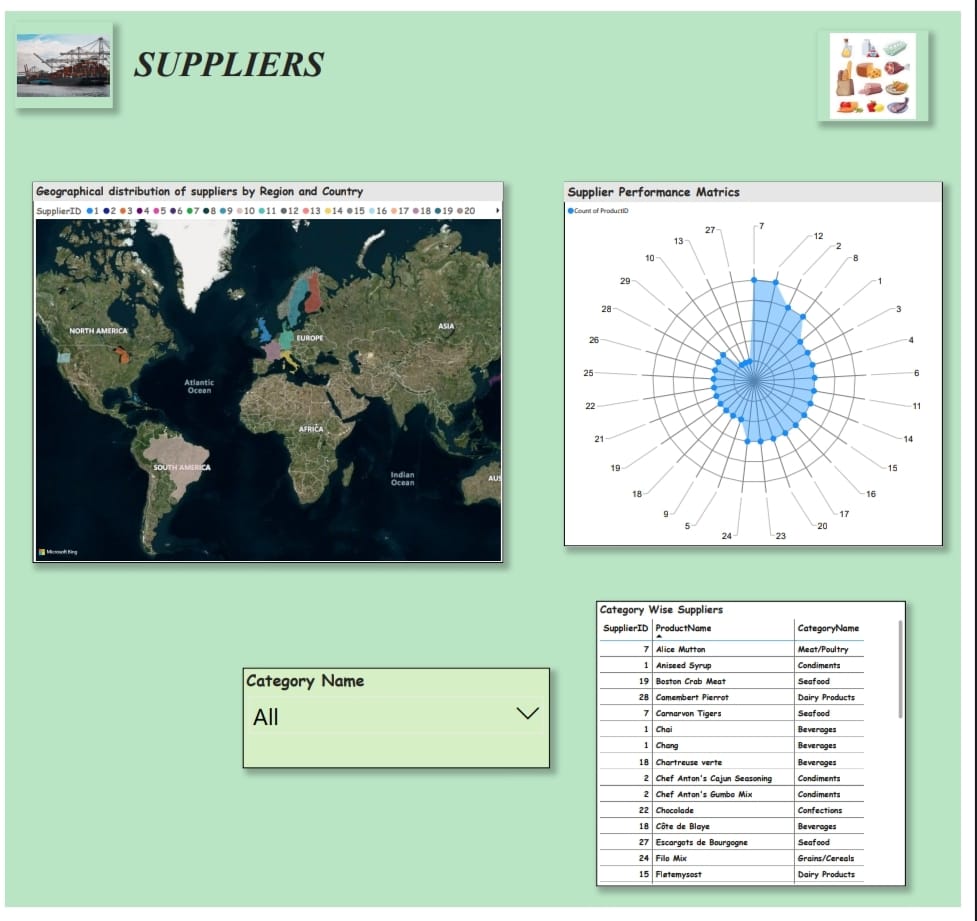
Analysing the Most Product categories sold indicates Customer’s Interest or need of that product. With the help of this we can increase sales by targeting customers.

Confections, beverages, Condiments, sea food and Dairy products are the top 3 high sold product categories.

**Country and Net Revenue:**

The map chart indicates the country wise net revenue generated by the products**.**

**Supplier Analyse**

****

27 Suppliers are there in the supplier table.

**What is the distribution of supplier ratings or performance metrics?**

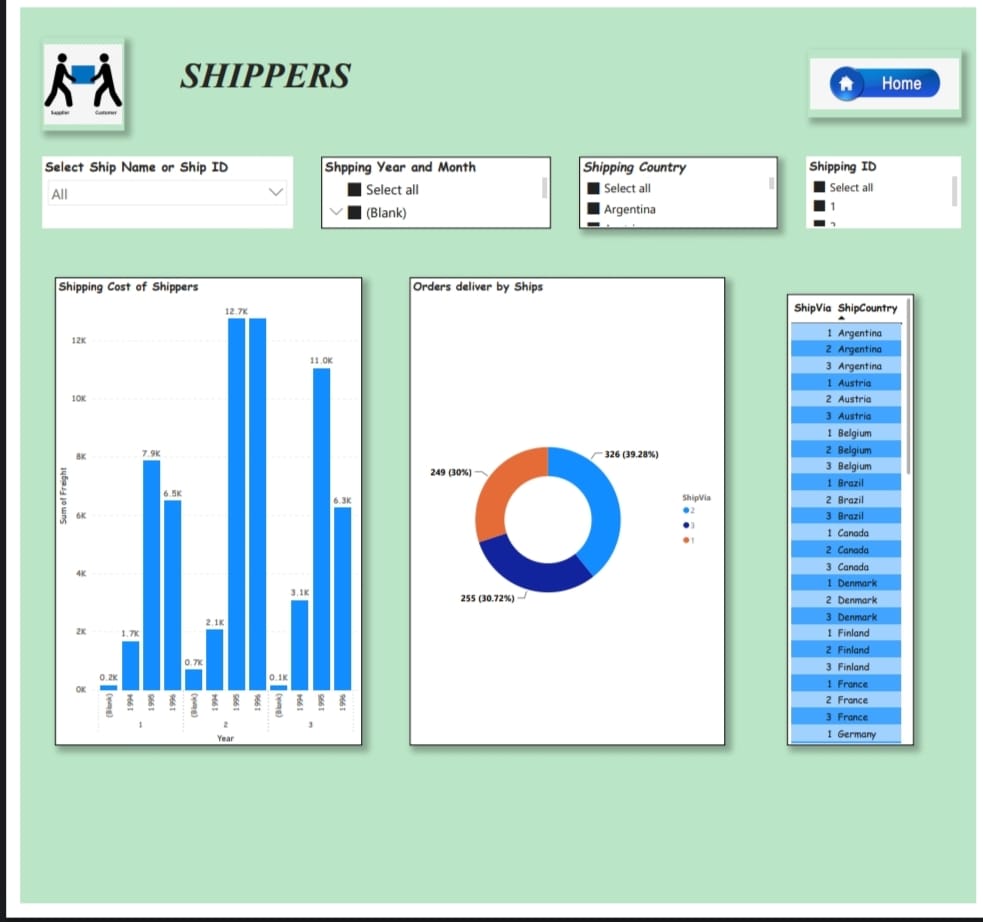
Analyse the supplier Performance Matrix with the help of count of supplies made by the Suppliers. To visualise this, we can use Radar Chart.

The supplier ID 7,12,2 and 8 are the best supplier according to count of supplies.

**Can we visualize the geographical distribution of suppliers**

The map chart distributes the supplier by country and region.

**Shippers Analyse:**

****

**Shipping Cost of the shippers:**

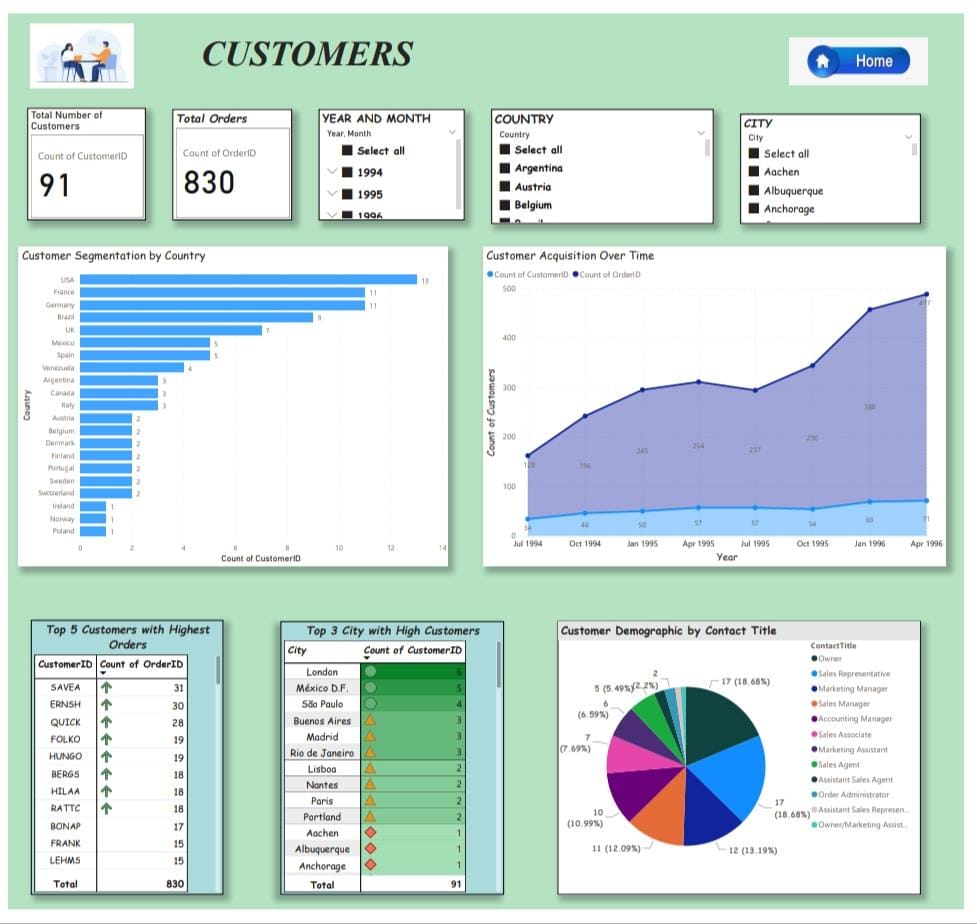
One of the key reasons is that shipping cost has a direct impact on profits of any business. If an exporter does not know how much they have to pay to ship goods to an importer, they may face challenges while determining profit margins

We can calculate the sum of Freight of each Shippers. The Shipping Company 2 ships the maximum order.

**How does Shipper’s distribution vary across different regions or customer segments**?

There are totally 3 shipping Companies are there.These three companies only shipped the all orders of the Customers.

**Power BI Report Snap shots**

****



