**NORTHWIND TRADERS**

**CAPSTONE PROJECT- SALES ANALYSIS REPORT**

**PREPARED BY :**

**HIMANSHU AGRAWAL**

**DATASET OVERVIEW**

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

**Table Explanations**

**Customers Table**

This 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.

**Employees Table**

This table stores information about the company's employees. It includes fields for employee ID, last name, first name, title, title of courtesy, birth date, hire date, address, city, region, postal code, country, home phone, extension, photo, notes, reports to, and photo path.

**Orders Table**

This table stores information about the company's orders. It includes fields for order ID, customer ID, employee ID, order date, required date, shipped date, ship via, freight, ship name, ship address, ship city, ship region, ship postal code, and ship country.

**Order Details Table**

This table stores detailed information about the items within each order. It includes fields for order ID, product ID, unit price, quantity, and discount.

**Products Table**

This table stores information about the company's products. It includes fields for product ID, product name, supplier ID, category ID, quantity per unit, unit price, units in stock, units on order, reorder level, and whether the product is discontinued.

**Suppliers Table**

This table stores information about the company's suppliers. It includes fields for supplier ID, company name, contact name, contact title, address, city, region, postal code, country, phone, fax, and home page.

**Shippers Table**

This table stores information about the company's shipping companies. It includes fields for shipper ID, company name, and phone.

**Categories Table**

This table stores information about the product categories. It includes fields for category ID, category name, and description.

**STEPS INVOLVED**

**1. Data Acquisition from GitHub**:

Obtain the requisite dataset from a designated GitHub repository, containing essential information on “Northwind Traders” in form of Excel , SQL, CSV sheets.

**2. Data Extraction & Transformation :**

To do effective analysis of the obtained data, I have used several data transformation techniques by using MYSQL Workbench, Analyze in Excel & ETL in Power BI to avoid data redundancy and data replication.

**3. Connecting the data with Tools**:

After extraction & transformation of data, I have established connections between the dataset and various analytical tools such as: Power BI, Excel, and MySQL Workbench for facilitating seamless data integration and processing.

**4. Solve the Problem Statement using 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 major objective of this report is to create a visually appealing and user-friendly dashboard that communicates key performance metrics for Northwind Traders effectively.

This report is prepared to connect all the major data connections of the data set to each other to judge their cumulative performance, parameters and to identify and remove the loopholes among them. The report aims to generate insights into customer behavior, sales patterns, and employee performance to aid decision-making processes. It will cover sales analysis, customer segmentation, inventory trends, and employee performance, consolidating data from multiple tables for a comprehensive view of the company's operations. The report will empower stakeholders to make data-driven decisions by offering valuable insights and facilitating data exploration through interactive visualizations and dynamic filters.

The expected impact is to revolutionize how Northwind Traders interacts with its data, enabling the company to remain competitive and drive its business forward in the wholesale market landscape.

The project will involve the following tasks:

* Performing a comprehensive analysis of dataset on the basis of Company’s Orders, Products , Suppliers ,Employees & Customers
* Deriving meaningful conclusions and recommendations for the growth in the number of Orders, quality of Products , inventory management or stock recommendations for Suppliers, improving Employee’s performance and Customer’s loyality.
* 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 boost the number of customer orders & sales revenue, quality of products, performance of employees and trust or loyality among the customers so that the company can improve better position themselves to succeed in the global marketplace.

**Significance**

The Northwind dataset can also be used for creating various projects and reports that demonstrate the ability to apply data analysis skills and techniques to solve real-world problems and provide valuable insights for business intelligence and decision making.

This report can play a significant role in the growth of business as it describes following features:

* Identifcation of Loop Holes
* Solving of Problem
* Valuable Insights
* Comprehensive Feedback
* Trend Identification
* Setting of Benchmarks
* Improvement in Focus
* Data-Driven Decisions

Some possible problem statements of a business that can be identified and solved on the Northwind dataset report can be:

* Analyzing the sales performance of different products, categories, regions, and customers.
* Identifying the customer behavior and preferences of different segments.
* Optimizing the supply chain and inventory management of different products and suppliers.
* Evaluating the employee performance and satisfaction of different roles and regions.
* Comparing the Northwind dataset with other similar datasets or real-world data sources.

**Conclusion**

 Harness the power of data visualization to drive better decision-making, streamline processes, and ultimately enhance your organization’s performance. With its customizable design, user-friendly interface, and advanced data analysis capabilities, the Beautiful Power BI Dashboard with Northwind Database is an invaluable tool for any data-driven organization.

**DATA DICTIONARY**

1. **Table – Categories**

Fields:

* Category ID - Unique id for each category.
* Name - It gives the name of the particular type of food category.
* Description - It gives the food items included in each category id.

1. **Table- Customers**

Fields:

* Customer ID – Unique id of each customer.
* Company Name – Name of the customer’s business or firm.
* Contact Name – Name of the contact person of the customer.
* Contact Title – Title or Designation of customer’s contact person.
* Address - Customer’s business address.
* City – City by which customer belongs.
* Region- Region by which customer belongs.
* Postal Code – Postal Code of customer’s business address.
* Country – Country by which customer belongs.
* Phone - Telephone or Contact number of customer.
* Fax – Fax no. of customer.

1. **Table - Employees**

Fields :

* Employee Id – Unique Id of each employee.
* Last Name - Last name of employee.
* First Name – First name of employee.
* Title - Job title of employee.
* Title of Courtesy - Courtesy of each employee.
* Birth date - Date of birth of each employee.
* Hire date - Date on which employee hired.
* Address – Address of each employee.
* City – City by which employee belongs.
* Region - Region by which employee belongs.
* Postal Code – Postal Code of each employee.
* Country - Country by which employee belongs.
* Home Phone - Phone or contact number of each employee.
* Notes - Educational qualifications of each employee.
* Report The employee for which the given employee reports.

1. **Table - Order Details**

Fields:

* Order ID– Unique id for each order.
* Product ID- Unique id of product.
* Unit Price- Unit price of product.
* Quantity – Number of products ordered .
* Discount – Discount given to the order.

1. **Table - Orders**

Fields :

* Order ID– Unique Id for each order
* Customer ID - Unique Id for each customer.
* Employee ID - Unique Id for each employee.
* Order Date – Date on which order has been done by the customer.
* Required Date - Date till which customer required a product.
* Shipped Date - Date on which order has been shipped by employee.
* Shipped Via – The Id of shipper who shipped the order
* Freight – The weight of the order shipped
* Ship Name – The name of the company by which the order has been shipped.
* Ship Address- The address which the order has been shipped.
* Ship City - The city which the order has been shipped.
* Ship Region - The region which the order has been shipped.
* Ship Postal Code – The postal code for which the order has been shipped.
* Ship Country - The country for which the order has been shipped.

1. **Table - Products**

Fields :

* Product ID - Unique ID of product.
* Name - Name of the product
* Supplier ID - Unique id of supplier for which the product belongs.
* Category ID – Category id for which the product belongs.
* Quantity per unit – Quantity involved in each unit.
* Unit Price – Unit price of product.
* Units in stock- Number of units of product available ion stock.
* Units on Order – Number of units of product which are under order.
* Reorder Level – Number of times for which the product which are ordered again.
* Discontinued – The product which get removed by the supplier.

1. **Shippers**

Fields

* ID - The unique ID of shipper.
* Company Name – The name of the shipper company.
* Phone- Contact or phone number of shipper.

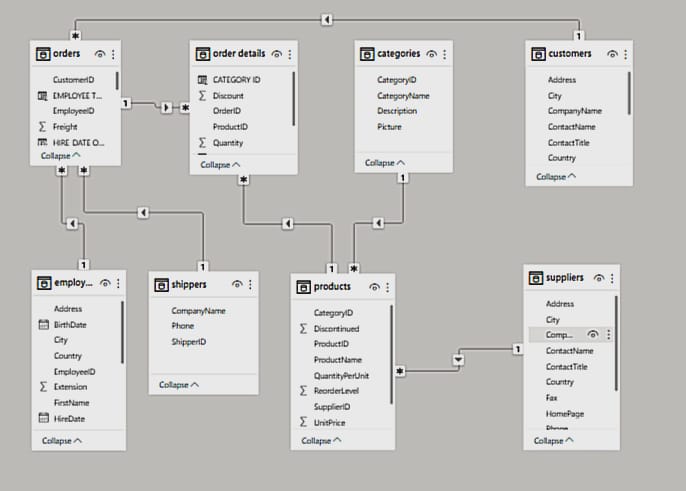
1. **Suppliers**

Fields

* Supplier ID - Unique ID of supplier for which the product belongs.
* Company Name – The name of the supplier’s company.
* Contact Name - Contact or Phone number of supplier.
* Contact Title – Designation of the supplier.
* Address - Address of the supplier.
* City - City by which supplier belongs.
* Region - Region by which supplier belongs.
* Postal Code – Postal code of the supplier.
* Country – Country by which supplier belongs.
* Phone - Contact or phone number of supplier.
* Fax- Fax number of employee.
* Home page - Website or domain address of supplier.

The data dictionary provides a comprehensive overview of the tables and their respective fields in the dataset. It outlines the relationships between the tables, allowing for a better understanding of the data structure and facilitating the design and implementation of the Power BI Dashboard.

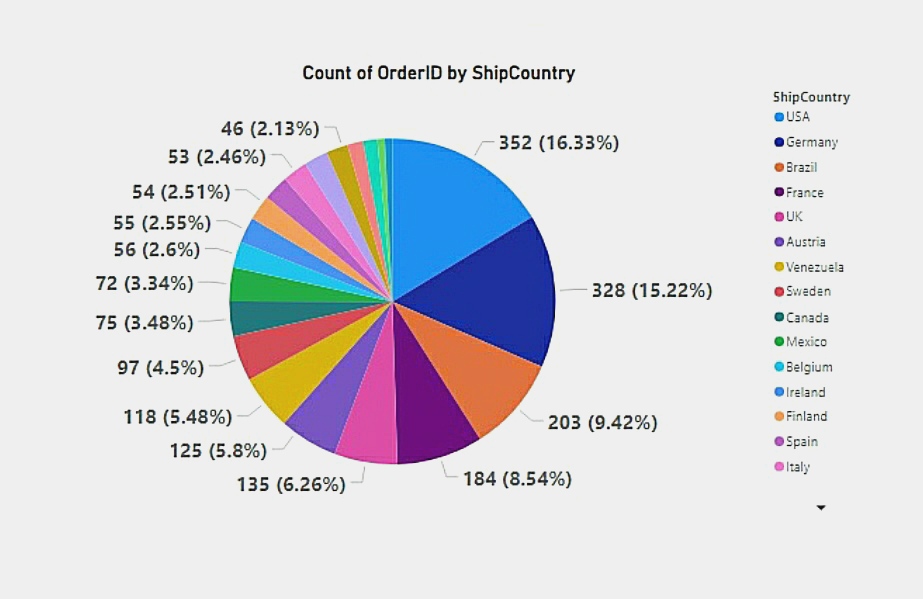
**ER DIAGRAM**



**POWER BI PROBLEM STATEMENTS**

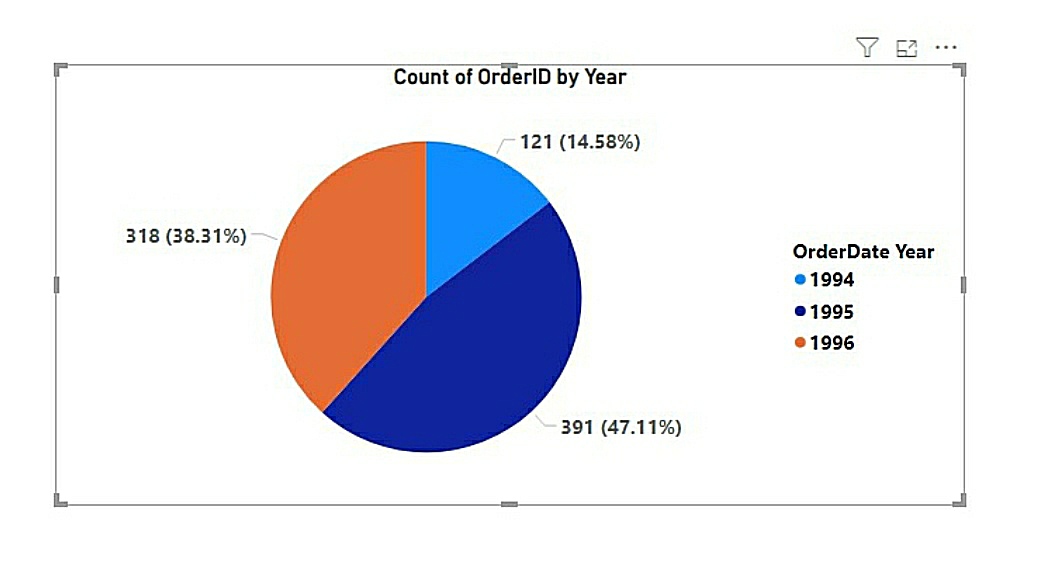
**Customer Distribution by Countries**

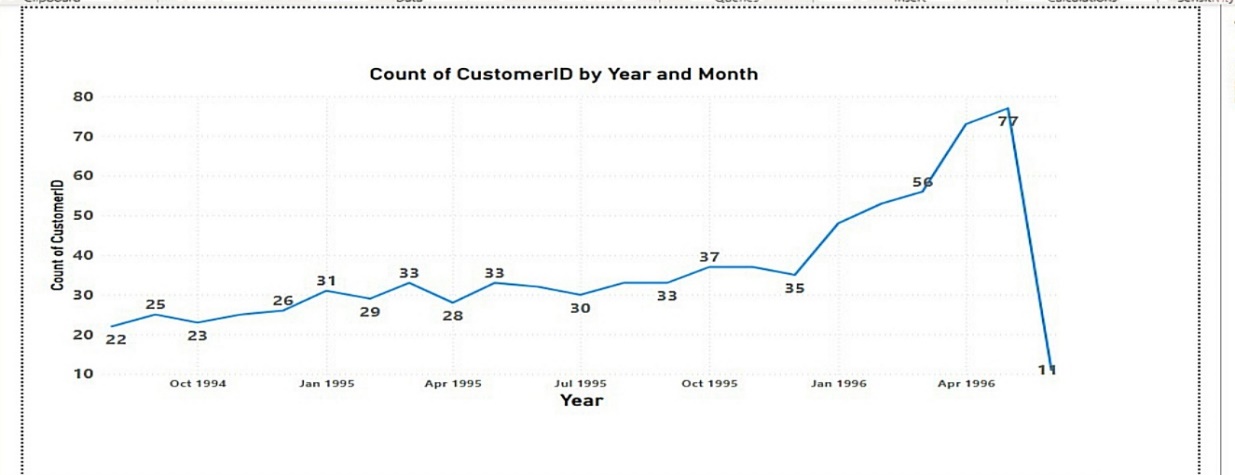
By analyzing the count of orders as per the country, we can generate meaningful insights that are as follows:



1. On the basis of above graph, Top 5 Countries with maximum number of orders are the countries that comes under top 10 GDP’S of the world. It shows that your count of orders are mainly depends on the per capital income of the country as these 5 countries have highest number of per capital income because they are leading economies of the world.
2. Out of the total of 21 countries in the dataset, most of the orders are come from the countries belong to the EUROPEAN REGION as 16 countries among the given countries are belong to the European Region. This shows that company has firm grip & loyality in the market of European Region.
3. Among the total countries of the dataset, USA has highest number of count of orders. This shows that in the Region Of The Americas, USA is the dominant market for the company followed by Brazil. It means that major count of orders for the company belongs Regions Of The America as America & Brazil are among the top 3 countries with highest number of orders.

**Customer Acquisition Over Time**

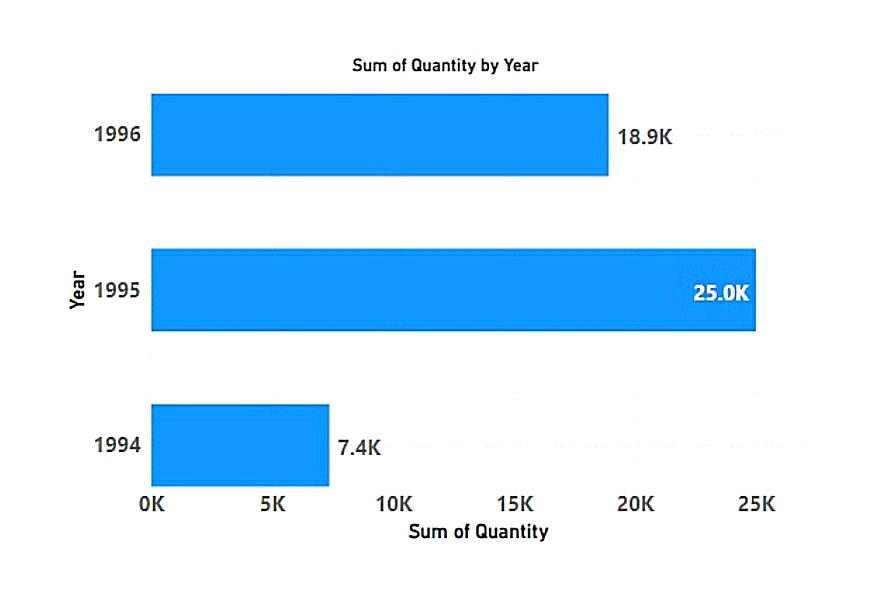




By the above analysis, we can generate following insights which are as follows:

1. From the above graphs, we can generate insights that company’s count of orders are increasing year to year. This shows that with the span of time company is became successful in acquiring more and more customers.
2. By the above graphical representation, we can generate insights that in 1994, company is getting average of approx. 24 orders per month while in 1995, this average count of order changed into approx. 33 orders per month and in the year of 1996, company has achieved a tremendous increase in orders count which has changed into a figure of average of 53 orders per month which is more than 200% of the monthly average orders in 1994. Thus , company is acquiring more and more customers with the time period.

**Volume Change Over Time**



By the above analysis, we can generate following insights which are as follows:

1- By the above graphical representation, we can say that company has received highest number of orders volume in the year 1995 among all the three years.

In year 1995, company has got total order volume of more than 25000 which is 3 times of the total orders volume received in 1994.

By this above representation, we can say that company has expanding its business with a exent of time-period.

2- In year 1996, company has got total order quantities of about 19000 in just 4 months which shows tremendous incrase in company order volume year by year.

**Distribution Of Order Values**

****

By the above analysis, we can generate following insights which are as follows:

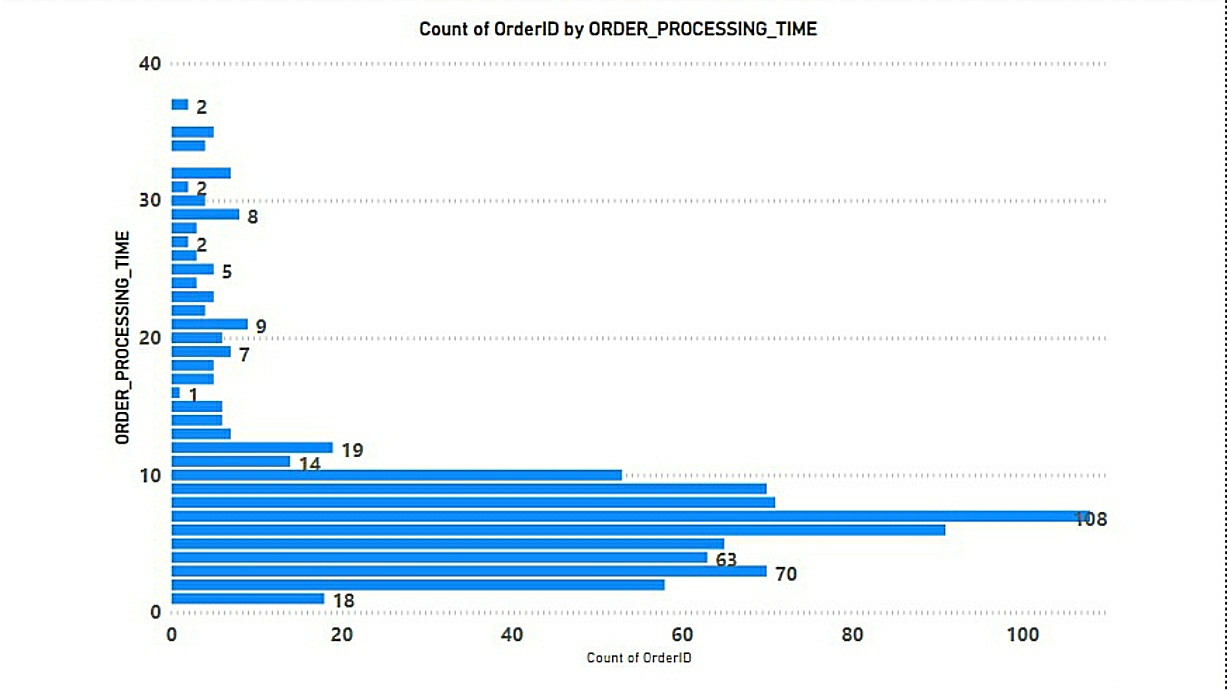
1- By the above graphical representation, we can say that company has received highest number of sales value in the year 1995 among all the three years.

In year 1995, company has got total order sales value of about 0.63 million which is more than 3 times of the total sales value received in 1994.

By this above representation, we can say that company has expanding its business with a exent of time-period.

2- In year 1996, company has got total sales value of about 0.55 in just 4 months which is 87 % of the total sales value of 12 months of 1995. It shows that company’s revenue is increasing year by year at a very high pace.

**Average Order Processing Time**



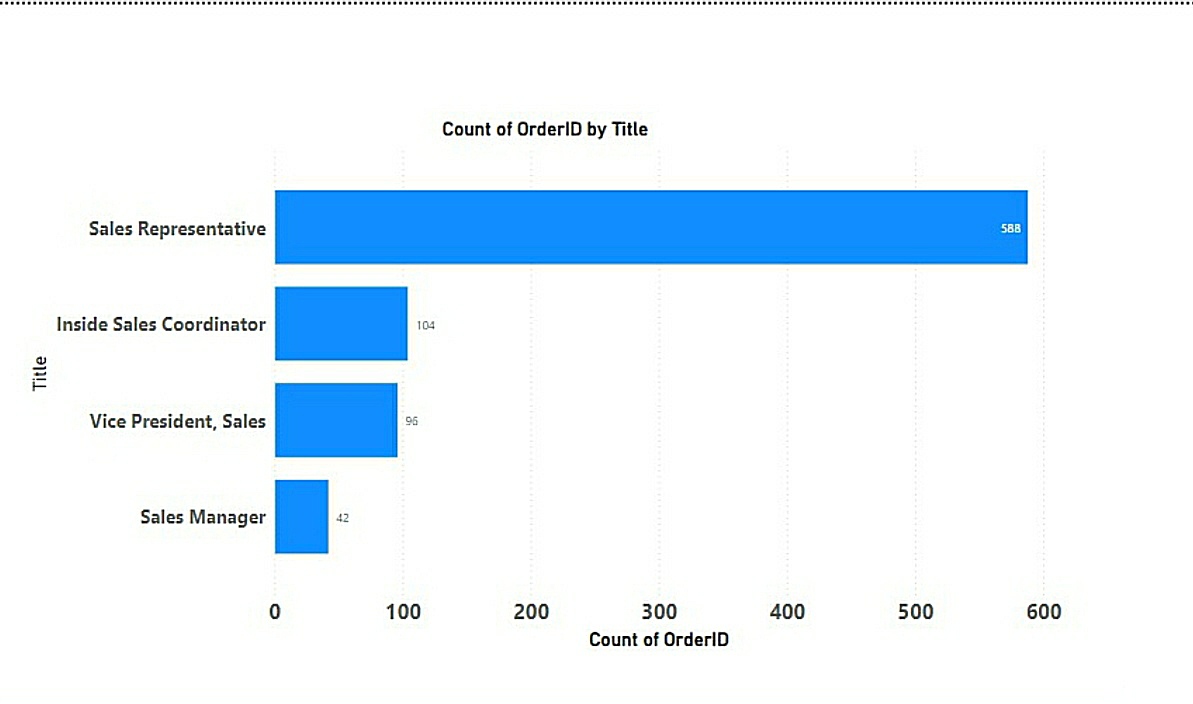
By the above analysis, we can generate following insights which are as follows:

1- By the above graphical representation, we can say that as major part of the orders are being shipped with in a time period of 10 days. So, we can say that company is taking an average time period of 10 days for dispatching the customer orders.

2- Maximum time taken by the company for the shipment of order is 37 days and Minimum time taken by the company for the shipment of order is 1 day.

3- After analyzing the time taken in shipment of orders, we have analyzed that 55% of total order counts are being shipped by the company within 3-7 days.

**Employee Productivity Variation across different departments or job roles?**



By the above analysis, we can generate following insights which are as follows:

1- Among all the total order counts, 70% of the orders are being received, managed and shipped by the sales representatives.

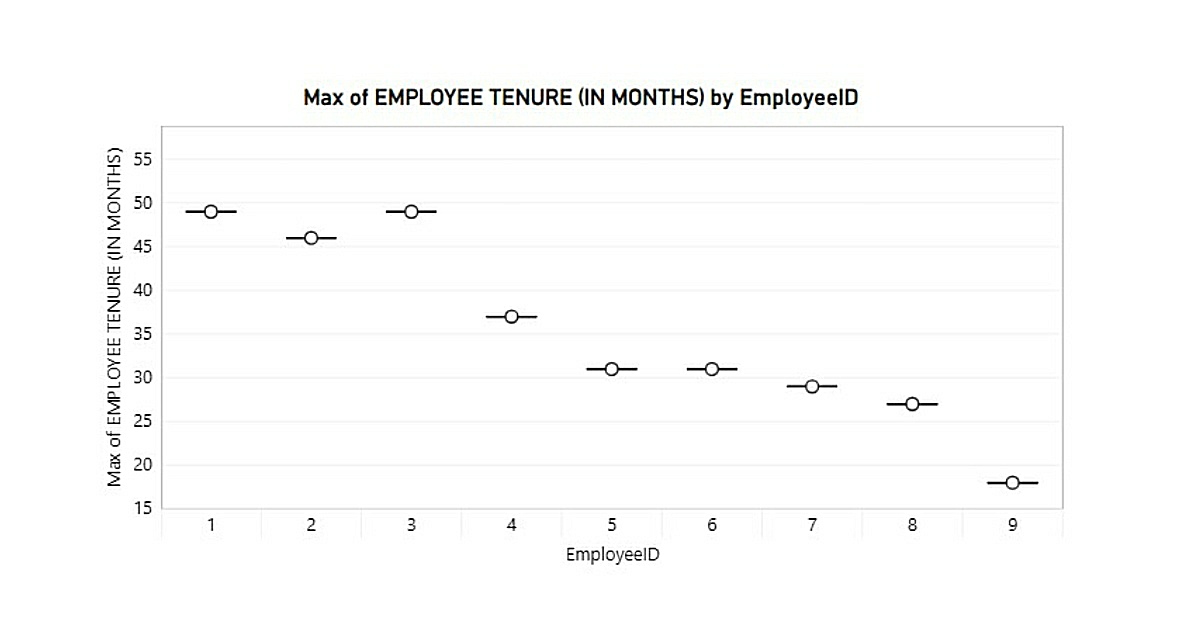
So, we can say that Sales representatives are the major players among all the employees of company that plays a vital role in company’s orders.

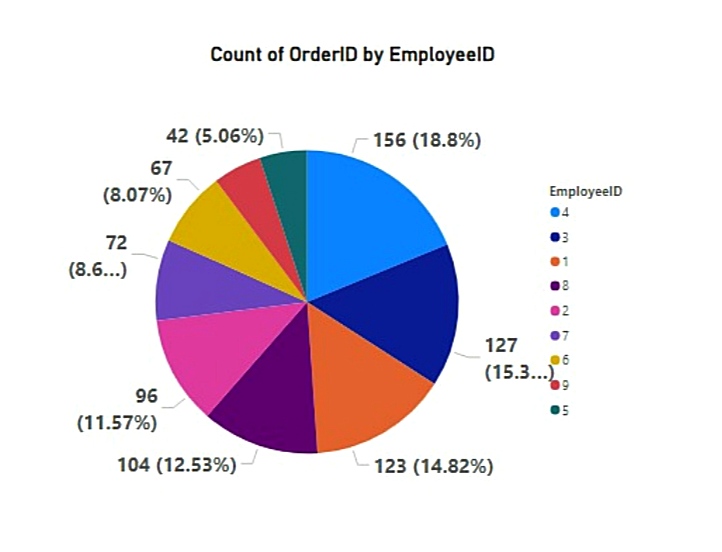
2- On the basis of orders received by the employee, we can notice that sales representatives are the mutual connections of contact between company and customer as most of the orders are being managed by them.

3- As per the above representation, we can say that in Company’s Sales Hierarchy table, Sales Representative is the lowest position and Vice-President Sales is the highest position among all the positions.

Sales representative are more productive which are followed by Sales Coordinator and Vice-President sales on the basis of Customer Orders Management.

**Distribution of Employee Tenure**



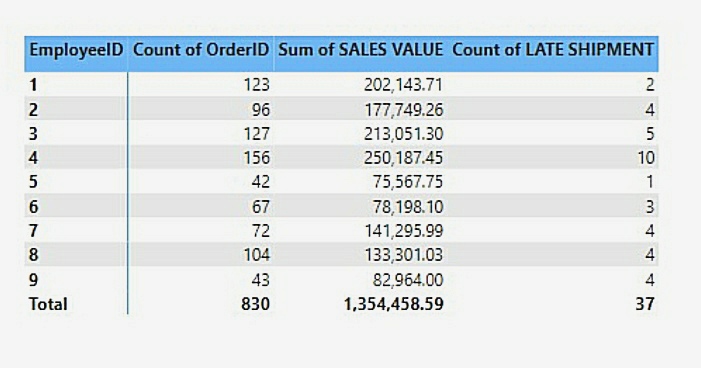


By the above analysis, we can generate following insights which are as follows:

1- Among all the employees, Employee with id 3 is having highest tenure period while Employee with id 9 is having lowest tenure period having all the employees.

2 – On the basis of above pie chart, quantity of ordered processed by the employees are proportional to their tenure period as we can notice that employee id 9 who has lowest tenure period among all the employees, has processed lesser number of orders while employee id 3 who has highest tenure period has processed second highest count of orders.

**Employee Performance Ratings**



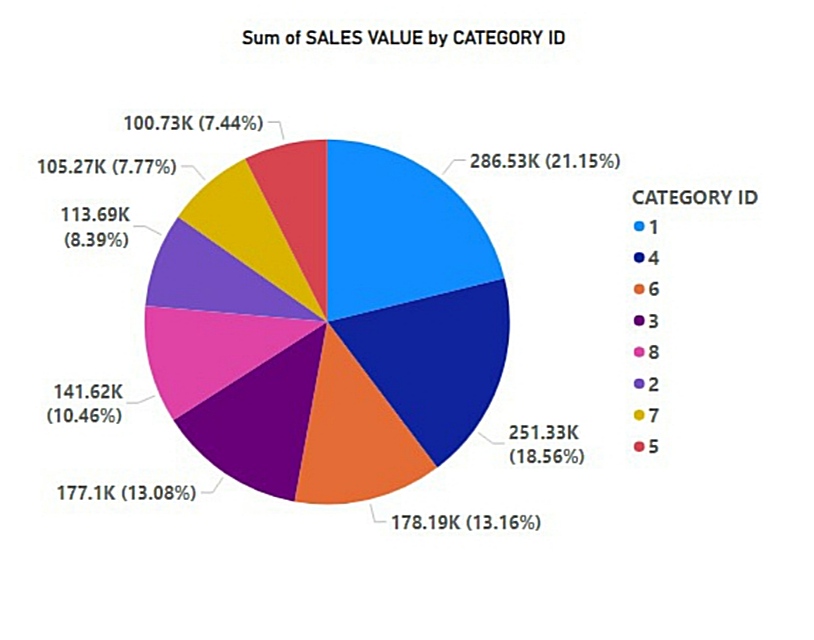
By the above analysis, we can generate following insights which are as follows :

1- Employee with id 1 is among the most productive employees of the company because he has lowest late shipment rate among all the employees and he is also among the top 3 employees in terms of order processing and sales generation.

2- Chances of late shipment by the employees are also depends on the count of the orders processed by them. .

As, employee who has processed more number of orders has highest count of late shipments while the employee who has processed lesser number of orders has lowest count of late shipments.

**Sales Volume Variation Across Different Product Categories**



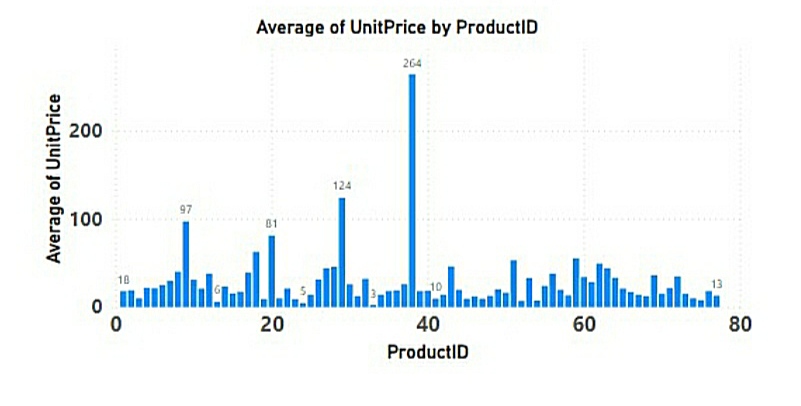
By the above analysis, we can generate following insights which are as follows :

1- In total sales volume of the company, more than 50% of the sales volume come from only three categories having category id 1,4 & 6 in which category id 1 is the leading one which is followed by 4 and 6 respectively.

Thus, we can say that customers spend most of their money in beverages, dairy products and meat products among all the food categories.

2- Category id 2, 5 and 7 are the lowest categories in terms of sales value which shows that customers spend lesser amount of their money in condiments , dry fruits and cereals among all the food categories.

**Pricing Distribution Of Products**

****

By the above analysis, we can generate following insights which are as follows :

1- product id 38 has highest price among all the products while the product id 33 has the lowest price.

2- After analyzing other data tables, we can say that the product id 38plays a vital role in sales values. As major amount of sales value belongs to the product id 38.

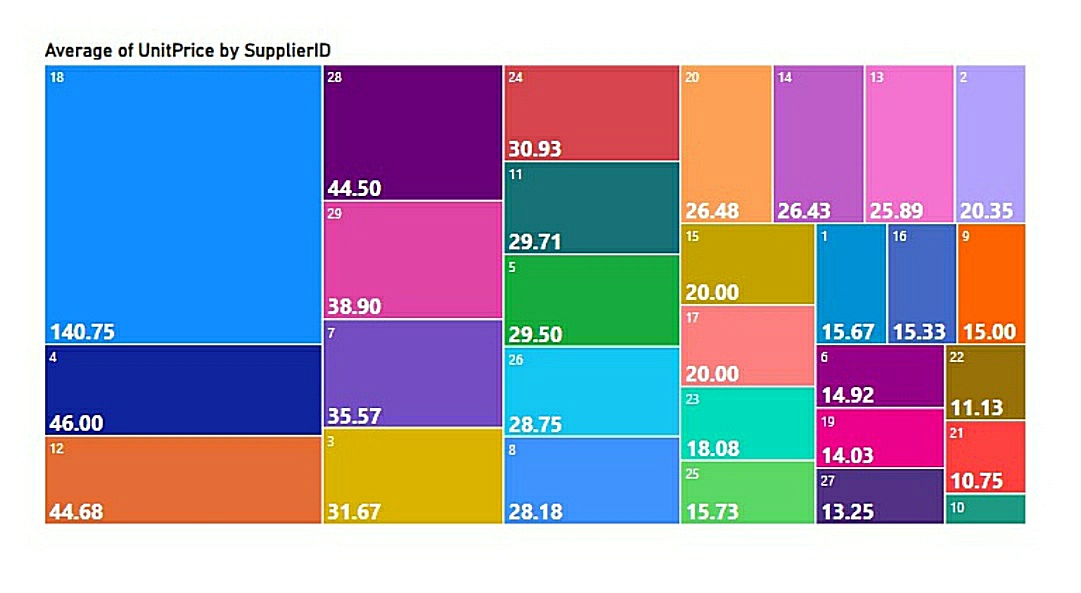
**Supplier Performance Metrics**



By the above analysis, we can generate following insights which are as follows :

1- supplier’s performance can be monitor by its average unit price, number of category id’s in which he deals and total stock available in its inventory.

2- On the above parameters, we can say that supplier with id 12 should be the best among all the suppliers as it has largest number of reorder levels, highest number of stock in its inventory and maximum number of categories in which he deals among all the suppliers.

**Price Structure Variation Across Different Suppliers**

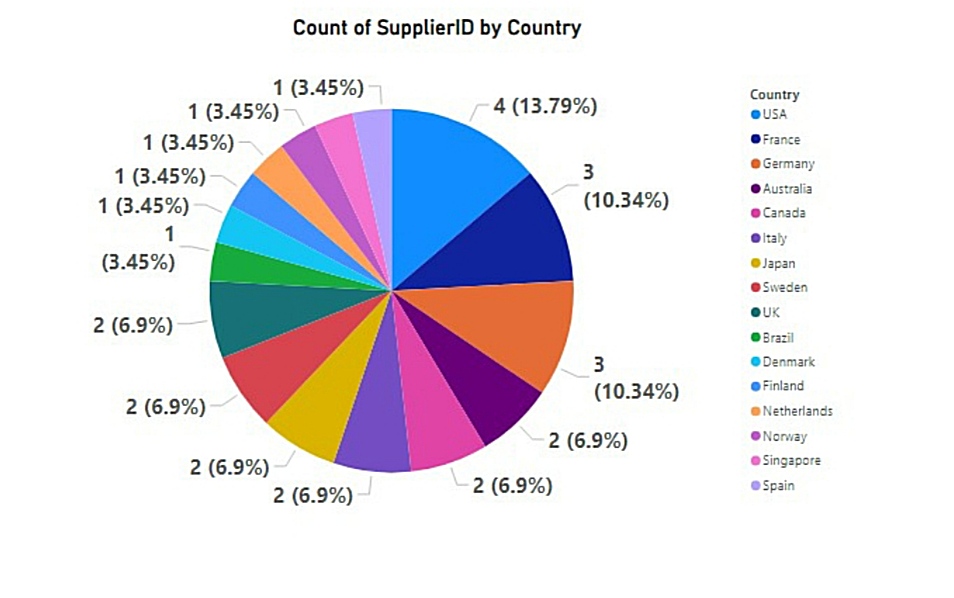
By the above analysis, we can generate following insights which are as follows :

1- We can say that supplier with id 18 has highest average unit price of the products while supplier id 10 has lowest average unit price of the products among all the suppliers.

2- Maximum number of the suppliers have their average unit price is in the range of 30-50.

**Geographical Representation Of Suppliers**





1- Company’s suppliers belong to three regions.

* European Region (Norway, France, Finland, Netherland, UK, Denmark, Spain, Germany, Sweden, Italy)
* Region of the Americas ( USA , Brazil, Canada)
* Western Pacific Region (Singapore, Japan, Australia)

2- After analysis, we can say that more than 50 % of suppliers of the company belong to the European Regions which is followed by Region of the Americas and Western Pacific Regions.

3- On the basis of country, highest number of suppliers belongs to USA.