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.

The Northwind database for the Power BI project includes essential tables such as Customers (customer details), Employees (employee information), Orders (order details), Order Details (item-level data), Products (product information), Suppliers (supplier details), Shippers (shipping companies), and Categories (product categories). These tables collectively provide insights into customer behaviour, sales patterns, inventory trends, and employee performance, facilitating data-driven decisions for Northwind Traders in the wholesale market landscape.

The data covers multiple years of business operations, including order history and interactions with customers, enabling a comprehensive analysis of customer behaviour and preferences over time.

OBJECTIVE

Objective of this NorthWind Project by exploring data and understanding the matrix. The analysis will give Key performance, sales pattern, order trend ,product analysis and Important Insights which will help to make business growth decision. The Charts will give better visual to understand the business trends with various matrix by which intelligence decision will be possible.

APPROACH

Downloaded this Dataset from GitHub which includes csv and SQL files with problem statements in PowerBI and EDA.

To understand data clearly and thoroughly I have used MySQL Workbench where imported data by running SQL script.

After understanding data very well to gain some insights preferred self-Analysis where made few Questions on visualization and EDA which gives relations among table.

SIGNIFICANCE

The Power BI project for Northwind Traders holds significant value as it provides valuable insights into key performance metrics, sales analysis, customer segmentation, inventory trends, and employee performance. By consolidating data from various tables, the project empowers stakeholders to make informed decisions, enabling the company to stay competitive and make strategic choices in the wholesale market landscape.

Understanding the factors that influence business performance can help Northwind Traders identify areas for improvement and focus efforts to enhance their overall ranking and market position. The project's historical analysis allows the company to track trends, identify strengths, and address weaknesses over time, facilitating benchmarking and strategic planning to ensure sustained growth and competitiveness.

Furthermore, the Power BI project enables data-driven comparisons with other businesses in the industry, allowing Northwind Traders to learn from best practices and adopt measures that can elevate their own performance. By leveraging these insights, the company can optimize operations, enhance customer satisfaction, and drive its business forward, revolutionizing its approach to data analysis and strategic decision-making in the wholesale industry.

Data Dictionary

1. **Customers Table:**
   * Fields: Customer ID, Company Name, Contact Name, Contact Title, Address, City, Region, Postal Code, Country, Phone, and Fax.
   * Purpose: Stores information about the company's customers, their contact details, and location.
2. **Employees Table:**
   * Fields: 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.
   * Purpose: Contains details about the company's employees, including their personal information and job-related details.
3. **Orders Table:**

• Fields: Order ID, Customer ID, Employee ID, Order Date, Required Date, Shipped Date, Ship Via, , Freight, Ship Name ,Freight, Ship Name, Ship Address, Ship City, Ship Region, Ship Postal Code, and Ship.

• Purpose: Stores information about the company's orders, including order details, shipping information, and related customer and employee IDs.

1. **Order Details Table:**

• Fields: Order ID, Product ID, Unit Price, Quantity, and Discount.

• Purpose: Contains detailed information about the items within each order, including product-

specific data and pricing.

1. **Products Table:**

• Fields: Product ID, Product Name, Supplier ID, Category ID, Quantity per Unit, Unit Price, Units in Stock, Units on Order, Reorder Level, and Discontinued.

• Purpose: Stores data related to the company's products, their suppliers, categories, and stock levels.

1. **Suppliers Table:**

• Fields: Supplier ID, Company Name, Contact Name, Contact Title, Address, City, Region, Postal Code, Country, Phone, Fax, and Home Page.

• Purpose: Contains information about the company's suppliers and their contact details.

**7. Shippers Table:**

• Fields: Shipper ID, Company Name, and Phone.

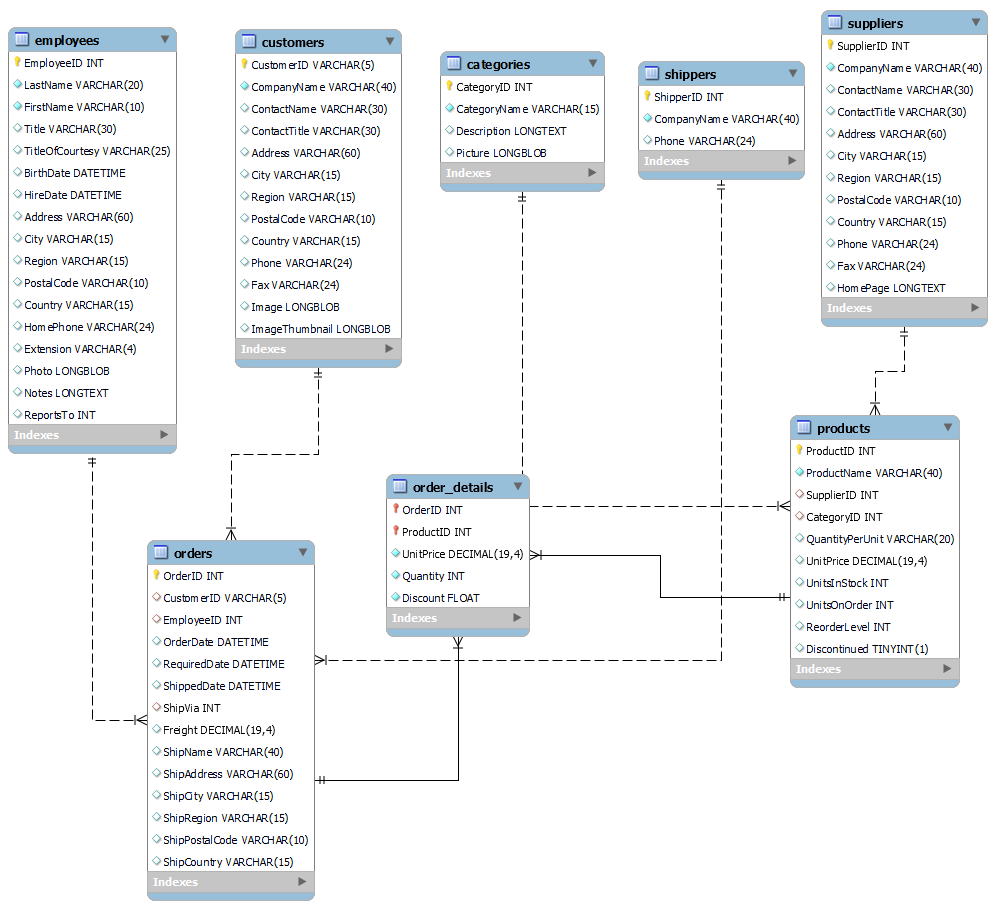
• Purpose: Stores data related to the shipping companies used by Northwind Traders for delivering orders.

**8. Categories Table:**

• Fields: Category ID, Category Name, and Description.

• Purpose: Contains information about the product categories to which the company's products belong.

ER Diagram



Steps to connect data

The given data are in format of csv and SQL files.

**Connect Data to Power BI:**

* Open Power BI Desktop.
* Click on "Home" in the top menu, then select "Get Data."
* Choose "Text/CSV" for CSV files.
* For CSV files, browse and select the CSV files want to import.
* Select the appropriate tables or views from the CSV files and click Transformation to import the data.

**Data Transformation :**

* After importing data, performed data transformations using Power Query Editor tool to clean, filter, and shape the data as needed.

**Visualizations :**

* Once the data is loaded into Power BI, create visualizations such as charts, tables, maps, and more to analyse and explore the data.

**Connect Data to Excel:**

* Open Excel and go to the "Data" tab.
* Select "Get Data" and choose "From Text/CSV" for CSV files or "From SQL Server" for SQL database.
* Follow the prompts to import the data into Excel.
* Perform data manipulations, calculations, and create pivot tables or charts as needed.

**Connect Data to MySQL:**

* In the SQL editor, go to "File" in the top menu, then choose "Open SQL Script."
* Navigate to the location of your SQL script file (.sql) on computer and select it.

**Review the SQL Script:**

* The contents of the SQL script file will be loaded into the SQL editor window. Review the script to ensure it contains the correct commands want to execute.

**Execute the SQL Script:**

* To execute the script
* Use the keyboard shortcut: Ctrl + Shift + Enter .
* Go to "Query" in the top menu and choose "Execute" or "Execute Current Statement."

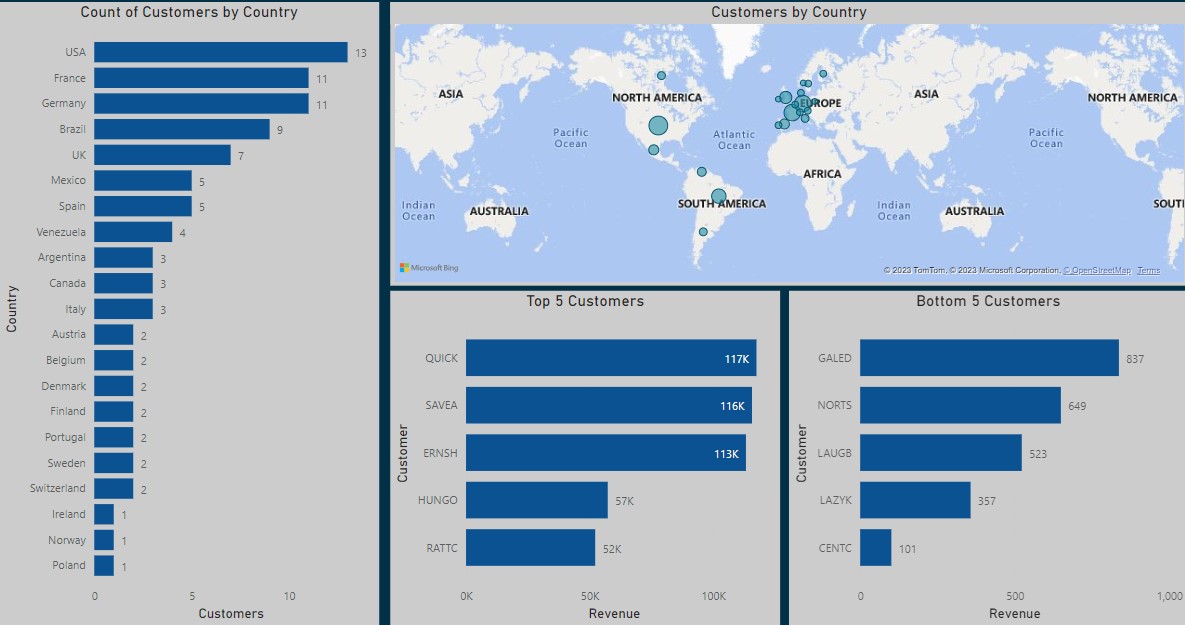
PowerBI Problem Statement

Customers

1) How does customer distribution vary across different regions or customer segments? Can we visualize it on a map or bar chart?

2) What is the trend in customer acquisition over time? Can we create a line chart or area chart to display it?

3) Can we visualize the distribution of customer demographics such as age, gender, or income using histograms or pie charts?

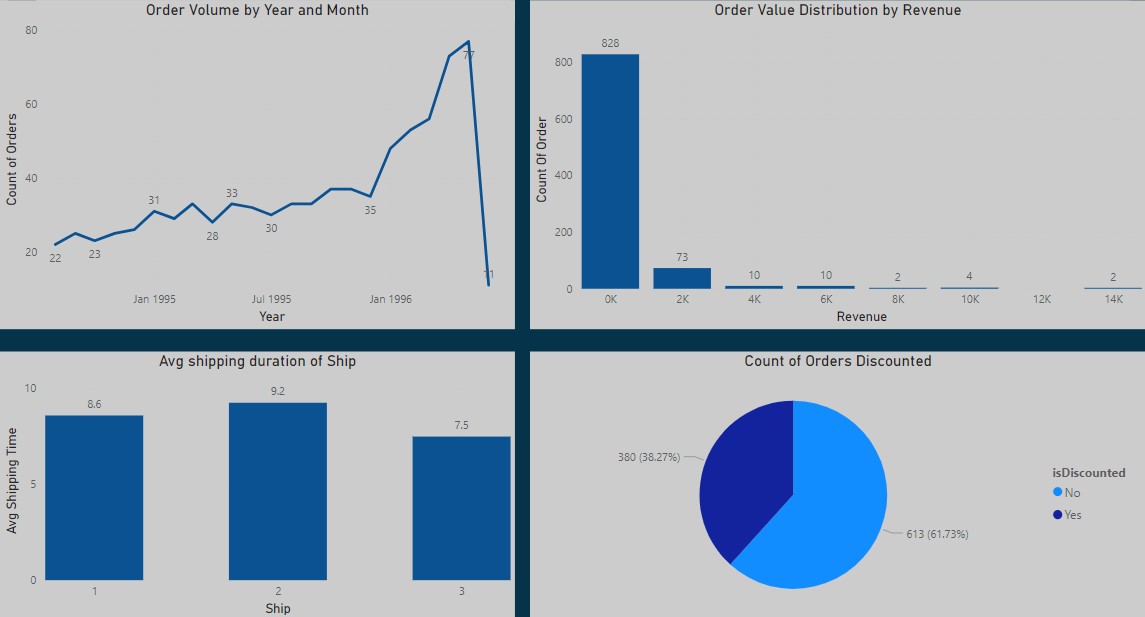


Orders

1) How does order volume change over time? Can we create a time series chart or stacked bar chart to visualize it?

2) What is the distribution of order values? Can we create a histogram or box plot to display it?

3) Can we visualize the average order processing time or shipping duration using a bar chart or box plot?

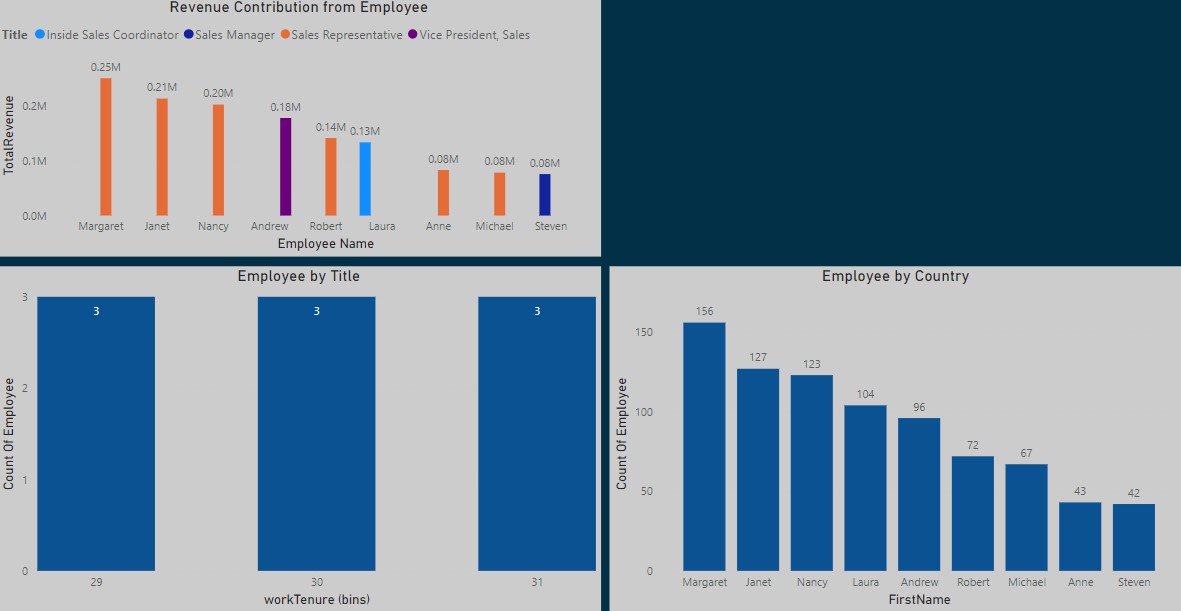


Employee

1) How does employee productivity vary across different departments or job roles? Can we create a stacked bar chart or grouped column chart to visualize it?

2) What is the distribution of employee tenure? Can we create a histogram or box plot to display it?

3) Can we visualize employee performance ratings or KPIs using a radar chart or bullet graph?

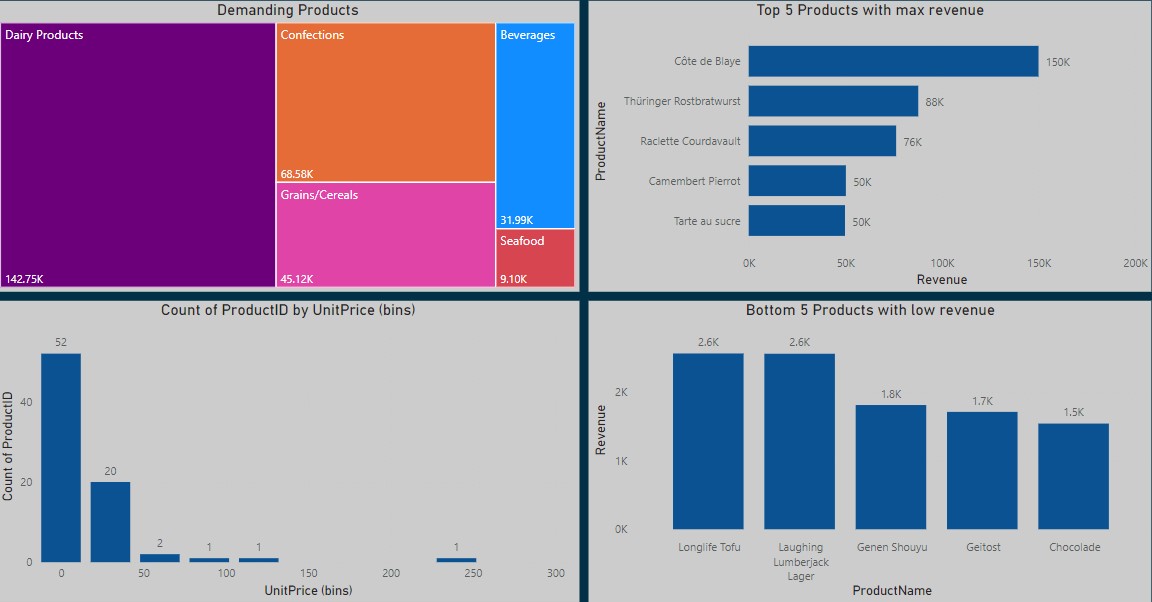


Product:

1) What is the distribution of product ratings or reviews? Can we create a histogram or stacked bar chart to visualize it?

2) How does the sales volume vary across different product categories? Can we create a bar chart or treemap to display it?

3) Can we visualize the pricing distribution of products using a box plot or violin plot?

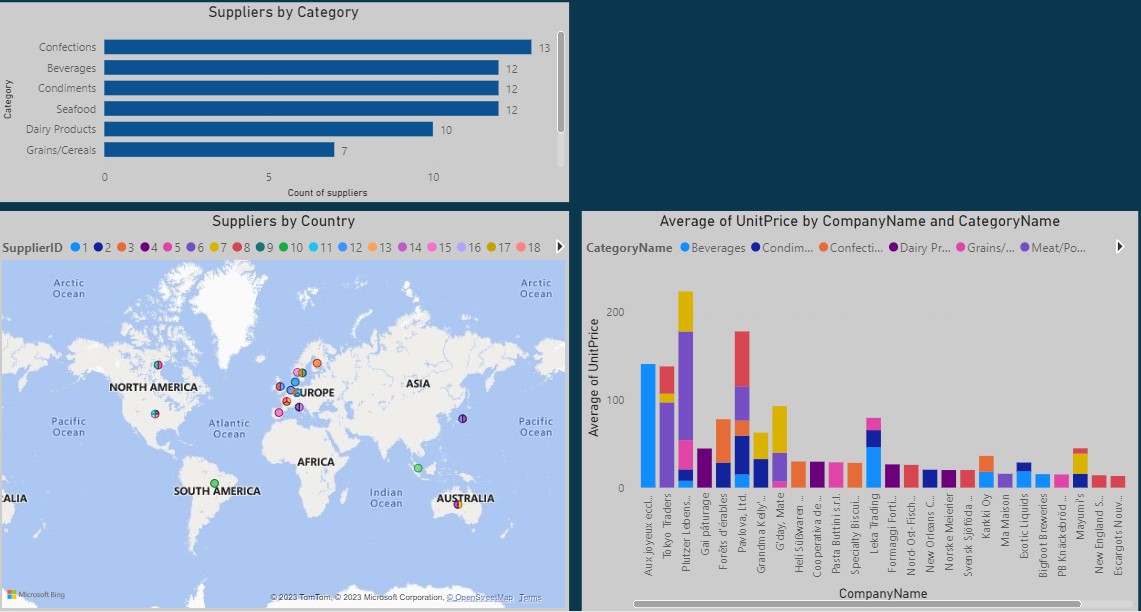


Suppliers:

1) What is the distribution of supplier ratings or performance metrics? Can we create a bar chart or radar chart to visualize it?

2) How does the cost or pricing structure vary across different suppliers? Can we create a box plot or stacked bar chart to display it?

3) Can we visualize the geographical distribution of suppliers using a map or bubble chart?



EDA Problem Statement

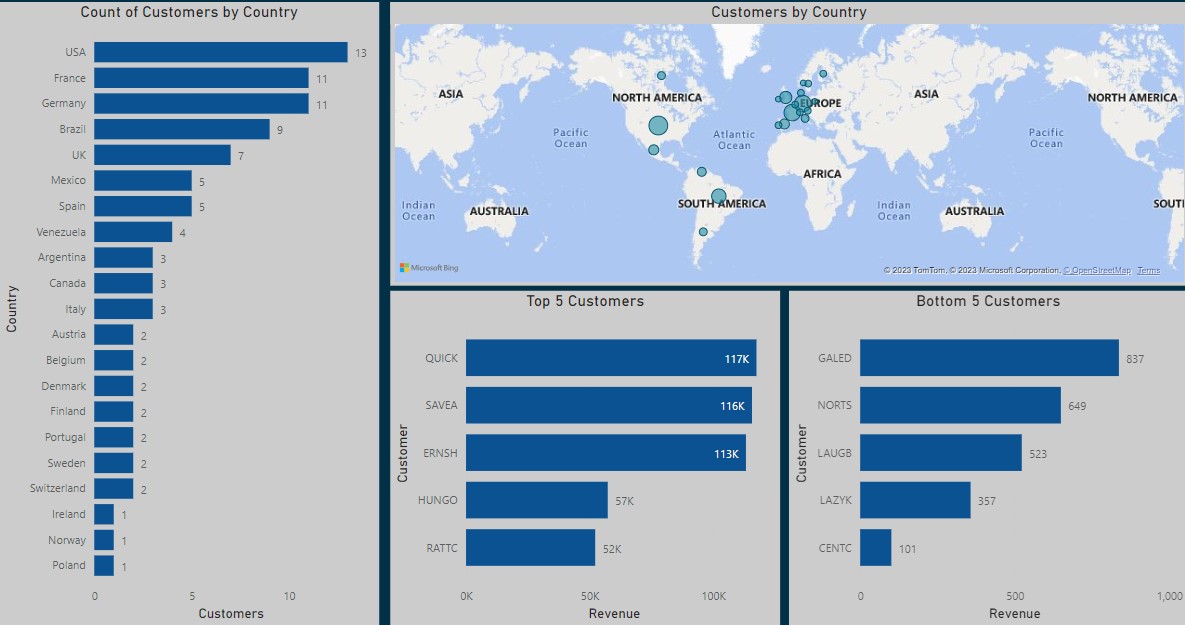
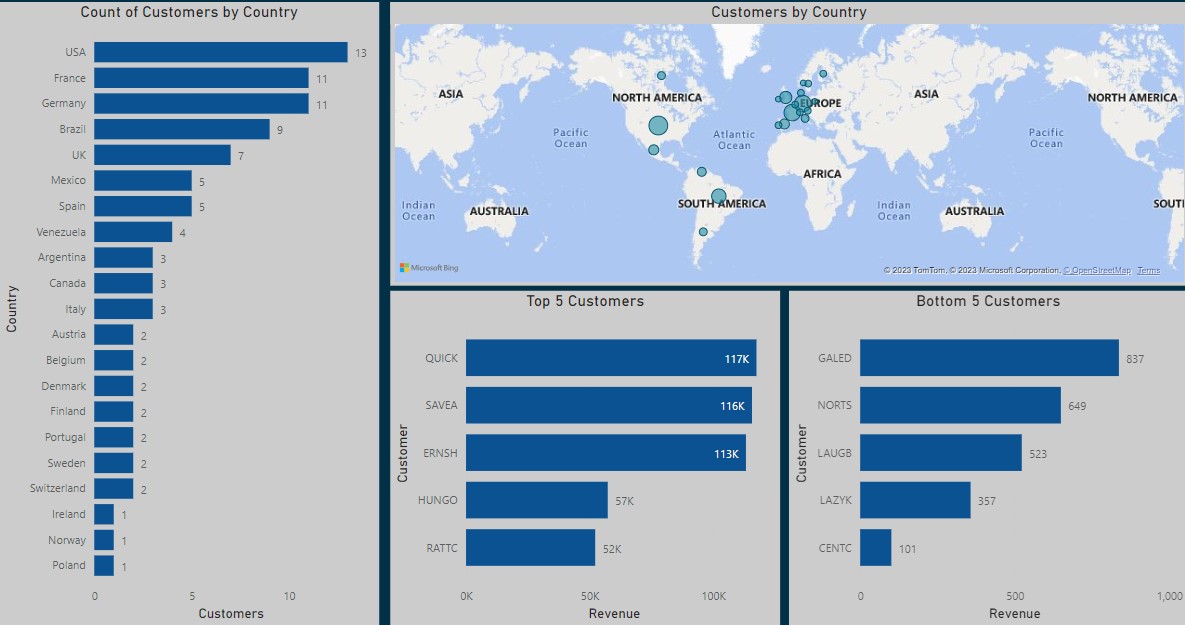
Customers

1) What are the key factors influencing customer retention or loyalty based on the dataset?

* The key factors for customers retention are mainly service, product quality and packaging. But according to dataset we can conclude the number of times customer visited which defines customers loyalty and order delivery status also influence customer retention.

2) How do customer preferences vary based on their location or demographics? Can we explore this through interactive visualizations?

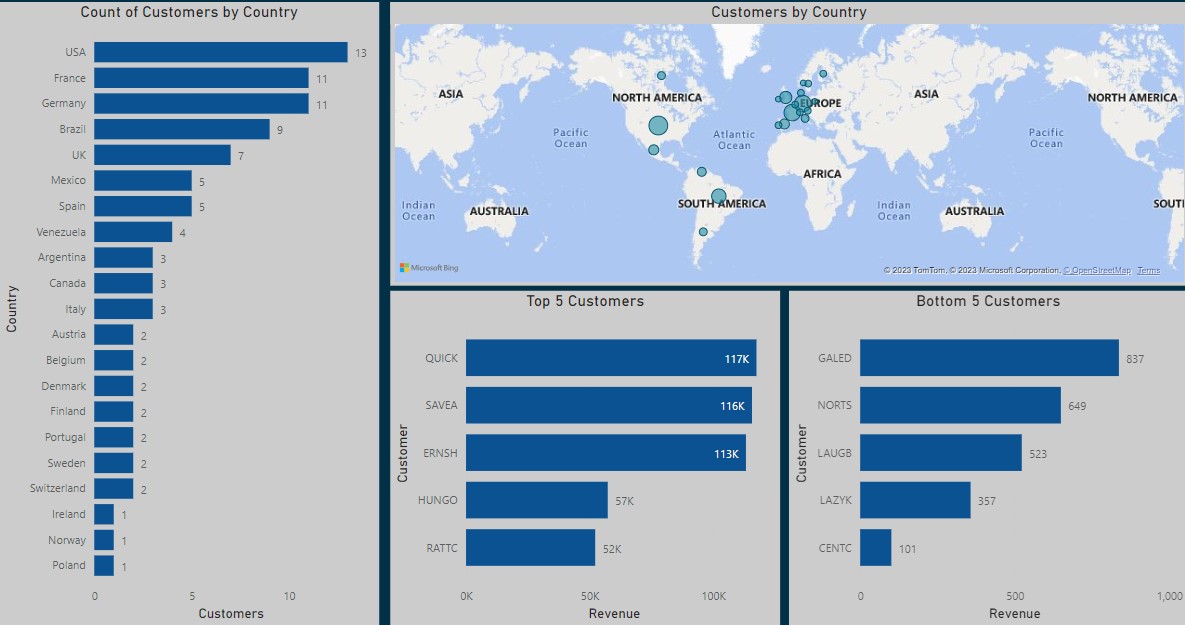
* The customers preference based on the near location of market first within country if not available then nearby country or region where product quality and packaging with service matters.



3) Are there any interesting patterns or clusters in customer behavior that can be visualized to identify potential market segments?

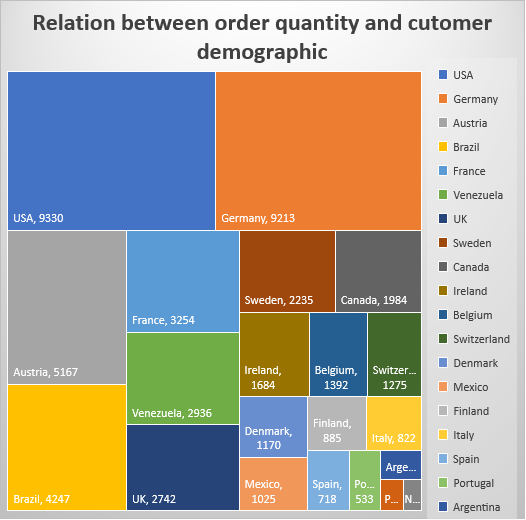
* Interesting Patterns can be conclude from dataset like customers who deals in every category , those who orders more than 10 times within given time span, top and bottom customers by sales.

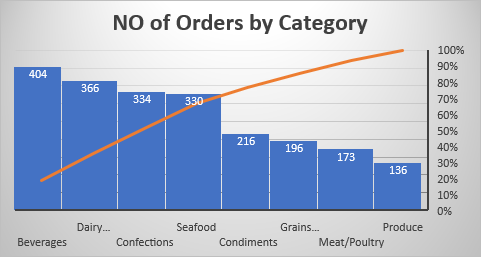


Order

1) Are there any correlations between order size and customer demographics or product categories? Can we explore this visually using scatter plots or heatmaps?



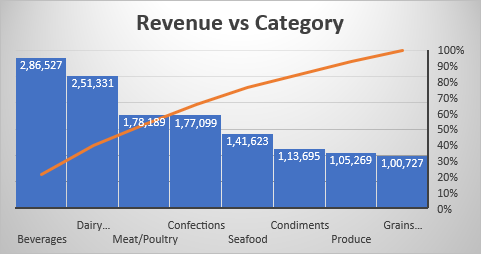
2)How does order frequency vary across different customer segments? Can we visualize this using bar charts or treemaps ?



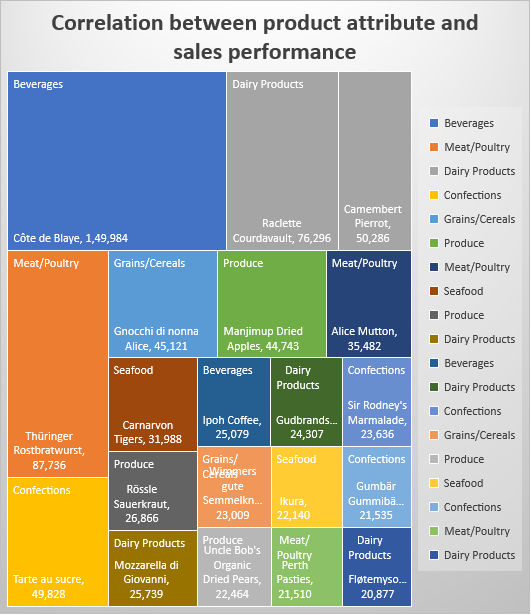


Products

1) Are there any specific product categories or SKUs that contribute significantly to order revenue? Can we identify them through visualizations?



2) Are there any correlations between product attributes (e.g., size, color, features) and sales performance? Can we explore this visually using scatter plots or heatmaps?



3) How does product demand fluctuate over different seasons or months? Can we visualize this through line charts or area charts?

4) Can we identify any outliers or anomalies in product performance or sales using visualizations? How can this information be used for product optimization?



Employee

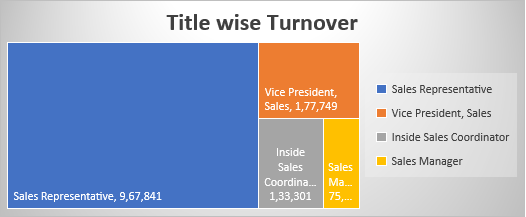
1) Are there any correlations between employee satisfaction levels and key performance indicators? Can we explore this visually through scatter plots or line charts?

* According to Dataset there is insufficient information about employee satisfaction levels like incentive, bonus, commission etc.
* Though there are Key performance indicators but wont be able create correlations.

2) Can we identify any patterns or clusters in employee skill sets or qualifications through visualizations? How can this information be used for talent management?

* From the qualification of employees I can conclude that almost all sales representative have completed their degree in BA or MA with few certification.
* Inside sales coordinator has also BA degree but with business certification course.
* Sales Manager has BSC degree but he is having experience in sales field.
* Vice President in sales has PHD degree due to his skill and hard work he reach this position from sales representative.

3)How does employee turnover vary across different departments or job roles? Can we visualize this using bar charts or heatmaps?



Supplier

1) Are there any correlations between supplier attributes (e.g., location, size, industry) and performance metrics (e.g., on-time delivery, product quality)?

* There is not sufficient information about suppliers side product quality or delivery.

2) How does supplier performance vary across different product categories or departments? Can we visualize this using stacked bar charts or grouped column charts?



3) Can we identify any trends or patterns in supplier costs or pricing structures through visualizations? How can this information be used for procurement optimization?

* From cost structure of dataset we can identify that minimum value for single product is cost up to $2.50 and it gone up to $263.50. Average value cost us for around $28.86

