The Northwind Traders sales database is a comprehensive sample database designed to simulate product sales, offering insights into customer behavior, product popularity, and supplier operational efficiency. It supports business intelligence and strategic decision-making by analyzing sales trends, employee performance, customer relationships, and supplier performance.

Northwind Traders sales Analysis

ALOK SARKAR

Abstract

This project provides an extensive analysis of sales data for the fictional company "Northwind Traders." The goal is to extract valuable insights into customer behavior, product performance, sales trends, supplier metrics, and employee performance, enabling data-driven decision-making. The documentation covers critical aspects such as a project overview, data acquisition, dataset table information, data cleaning processes, MECE (Mutually Exclusive, Collectively Exhaustive) breakdown, Power BI usage, Exploratory Data Analysis (EDA) questions, and dashboard analysis.

Through interactive visualizations and dynamic filters, the project delves into key areas including customer distribution, acquisition trends, demographics, order volume, order values, employee productivity, tenure, product performance, supplier metrics, and the geographical distribution of suppliers. These visualizations and analyses reveal significant insights into customer preferences, market segments, product success, and supplier relationships, enabling stakeholders to make informed decisions and identify areas for improvement.

The abstract offers a preview of the detailed analysis and insights derived from the project, highlighting the use of Power BI and EDA to extract valuable information from the dataset. The primary objective of the project is to empower stakeholders with actionable insights and facilitate comprehensive data exploration through interactive visualizations and dynamic filters, ultimately aiding in informed decision-making and driving operational enhancements.

CONTENTS

INTRODUCTION

➢ Project Overview:

The North Wind Traders Capstone Project presents an in-depth analysis of sales data for the fictitious company "North Wind Traders," specializing in the import and export of specialty foods from around the world. This project aims to provide insights into customer behavior, sales patterns, and employee performance to support informed decision-making. It covers a range of areas, including sales analysis, customer segmentation, inventory trends, and employee performance, by consolidating data from multiple tables to offer a comprehensive view of the company's operations.

The primary goal of the project is to empower stakeholders to make data-driven decisions by providing valuable insights and facilitating data exploration through interactive visualizations and dynamic filters. Utilizing MySQL for database management, along with Excel and Power BI for data visualization, the project extracts meaningful information from the dataset. It showcases sample visualizations and analyses to demonstrate the use of interactive visualizations and dynamic filters.

Key functionalities of the project include:

- Sales analysis
- Customer segmentation
- Inventory trends
- Employee performance analysis
- Product performance evaluation
- Supplier metrics visualization
- Interactive visualizations and dynamic filters

The project's documentation includes:

- Project overview
- Data acquisition
- Information about dataset tables
- Data cleaning and MECE breakdown
- Power BI and EDA questions
- Dashboard analysis
- Conclusion

❖ Introduction of data set Table:

The North Wind Traders Capstone Project includes several tables that store information about the company's operations.

Data dictionary for the Northwind database based on the provided tables:

Customers Table

This table stores information about the company's customers.

- CustomerID: Unique identifier for each customer.
- CompanyName: Name of the customer's company.
- ContactName: Name of the contact person at the company.
- ContactTitle: Title of the contact person.
- Country: Country where the customer is located.
- City: City where the customer is located.

Employees Table

This table stores information about the company's employees.

- EmployeeID: Unique identifier for each employee.
- LastName: Last name of the employee.
- FirstName: First name of the employee.
- Title: Job title of the employee.
- TitleOfCourtesy: Title of courtesy (Mr., Mrs., etc.) for the employee.
- BirthDate: Date of birth of the employee.
- HireDate: Date when the employee was hired.
- Address: Street address of the employee.
- City: City where the employee is located.
- Region: Region or state where the employee is located.
- PostalCode: Postal code or ZIP code of the employee's location.
- Country: Country where the employee is located.

- HomePhone: Home phone number of the employee.
- Extension: Extension number for contacting the employee.
- Photo: Binary data of the employee's photograph.
- Notes: Additional notes about the employee.
- ReportsTo: ID of the employee's manager (links to another EmployeeID).
- PhotoPath: File path to the employee's photograph.

Orders Table

This table stores information about the company's orders.

- OrderID: Unique identifier for each order.
- CustomerID: ID of the customer who placed the order (links to CustomerID in Customers table).
- EmployeeID: ID of the employee who processed the order (links to EmployeeID in Employees table).
- OrderDate: Date when the order was placed.
- RequiredDate: Date when the order is required by.
- ShippedDate: Date when the order was shipped.
- ShipVia: ID of the shipping company used for the order (links to ShipperID in Shippers table).
- Freight : Cost of shipping.
- ShipName: Name of the company receiving the shipment.
- ShipAddress: Street address of the shipment destination.
- ShipCity: City of the shipment destination.
- ShipRegion: Region or state of the shipment destination.
- ShipPostalCode: Postal code of the shipment destination.
- ShipCountry: Country of the shipment destination.

Order Details Table

This table stores detailed information about the items within each order.

- OrderID: ID of the order (links to OrderID in Orders table).
- ProductID: ID of the product ordered (links to ProductID in Products table).
- UnitPrice: Price per unit of the product at the time of the order.
- Quantity: Number of units ordered.
- Discount: Discount applied to the product in the order.

Products Table

This table stores information about the company's products.

- ProductID: Unique identifier for each product.
- ProductName: Name of the product.
- SupplierID: ID of the supplier for the product (links to SupplierID in Suppliers table).
- CategoryID: ID of the category to which the product belongs (links to CategoryID in Categories table).
- QuantityPerUnit: Description of the quantity of product per unit.
- UnitPrice: Price per unit of the product.
- UnitsInStock: Number of units currently in stock.
- UnitsOnOrder: Number of units currently on order with suppliers.
- ReorderLevel: Inventory level that triggers a reorder of the product.
- Discontinued: Indicates if the product has been discontinued.

Suppliers Table

This table stores information about the company's suppliers.

- SupplierID: Unique identifier for each supplier.
- CompanyName: Name of the supplier's company.
- ContactName: Name of the contact person at the supplier's company.
- ContactTitle: Title of the contact person.
- Address: Street address of the supplier.
- City: City where the supplier is located.
- Region: Region or state where the supplier is located.
- PostalCode: Postal code or ZIP code of the supplier's location.
- Country: Country where the supplier is located.
- Phone: Contact phone number of the supplier.
- Fax: Fax number of the supplier.
- HomePage: URL for the supplier's website.

Shippers Table

This table stores information about the company's shipping companies.

- ShipperID: Unique identifier for each shipper.
- CompanyName: Name of the shipping company.
- Phone: Contact phone number of the shipping company.

Categories Table

This table stores information about the product categories.

- CategoryID: Unique identifier for each category.
- CategoryName: Name of the category.
- Description: Description of the category.

❖ Reason For the Project:

The North Wind Traders Capstone Project aims to create a visually appealing and user-friendly dashboard to effectively communicate key performance metrics. The primary objective is to generate insights into customer behavior, sales patterns, and employee performance to support decision-making processes. By covering areas such as sales analysis, customer segmentation, order trends, and employee performance, the project consolidates data from multiple tables to provide a comprehensive view of the company's operations. The report is designed to empower stakeholders with valuable insights and facilitate data exploration through interactive visualizations and dynamic filters.

The dataset includes detailed tables for customers, employees, orders, order details, products, suppliers, shippers, and categories, offering a rich source of information for the analysis. The project involves rigorous data cleaning and a MECE (Mutually Exclusive, Collectively Exhaustive) breakdown to ensure the quality and structure of the data. Additionally, a set of Power BI and EDA (Exploratory Data Analysis) questions guide the analysis and visualization process. These questions cover various aspects, including customer distribution, order volume, employee productivity, product performance, supplier metrics, and customer retention.

The expected impact of the project is to revolutionize how North Wind Traders interacts with its data, enabling the company to remain competitive and drive its business forward in the wholesale market landscape. The comprehensive approach and the use of interactive visualizations and dynamic filters are aimed at providing a deep understanding of the company's operations.

❖Key Functionality of the Project:

The key functionalities of this Power BI project for Northwind Traders include:

- **Data Integration and Transformation**
- **Visual Data Exploration**
- **Sales Analysis**
- **Customer Segmentation**
- **Inventory Management**
- **Employee Performance Analysis**
- **Interactive Dashboards and Reports**
- **Report Sharing and Collaboration**

❖ Aim and Objective:

Identify Sales Trends: Examine sales data to uncover patterns, cycles, and shifts in sales performance over time, by product category, geographic region, and customer demographic.

Understand Customer Behavior: Profile customer characteristics to understand purchasing habits, inclinations, and demographics, enabling targeted marketing strategies and enhanced customer relationships.

Assess Employee Effectiveness: Evaluate employee sales performance, productivity, and efficiency to identify high achievers, training needs, and opportunities for growth.

Deliver Actionable Intelligence: Present insights derived from data analysis in a visually engaging and intuitive dashboard format, enabling stakeholders to make informed decisions and take proactive measures.

Foster Data-Driven Culture: Cultivate a culture of data-informed decision-making within the organization by providing access to timely and relevant data, empowering stakeholders at all levels to leverage data for strategic planning and operational enhancements.

Provide Actionable Insights: Present insights derived from data analysis in a visually appealing and intuitive dashboard format, enabling stakeholders to make informed decisions and take proactive actions.

Foster Data-Driven Culture: Promote a culture of data-driven decision-making within the organization by providing access to timely and relevant data, empowering stakeholders at all levels to leverage data for strategic planning and operational improvements.

Boost Business Agility: Utilize data insights to identify competitive advantages, market opportunities, and areas for innovation, enabling Northwind Traders to stay ahead in the wholesale market landscape.

DATA CLEANING

Data cleaning is the process of identifying and correcting or removing errors, inconsistencies, and inaccuracies in data. This step is crucial in data analysis as it ensures that the data is accurate, complete, and consistent. The goal of data cleaning is to improve the quality of data and make it suitable for analysis.

Data cleaning involves several tasks, including:

- Identifying and filling in missing values
- Removing duplicate records
- Correcting spelling and formatting errors
- Standardizing data formats

Some common challenges faced during data cleaning include:

Missing Data: Dealing with missing values in the dataset, which may require imputation or removal of incomplete records.

Inconsistent

Data: Addressing inconsistencies in data formats, such as date formats, currency symbols, or units of measurement.

Duplicate Records: Identifying and removing duplicate entries, which can skew analysis and lead to inaccurate results.

Outliers: Handling outliers that can significantly impact statistical analysis and visualization.

Data Standardization: Ensuring that data is consistent and standardized across different sources or systems.

Data Validation: Verifying the accuracy and integrity of the data, which may involve cross-referencing with external sources or known benchmarks.

Data Transformation: Converting data into a suitable format for analysis, such as aggregating, pivoting, or normalizing data.

Data Quality: Ensuring the overall quality of the data, including accuracy, completeness, and reliability.

These challenges are common in the data cleaning process and require careful attention to detail to ensure the accuracy and integrity of the data for analysis.

Process of Cleaning in North Wind dataset tables:

❖ Customers table:

- In Customer table region column contain many null values that are replaced with N/A.
- After this removed these columns that are not necessary for the visualization. These columns are Phone, Fax, Postalcode , Image and Image Thumbnail

❖Employee table:

From the employee table these columns are removed that are not providing and information for visualization. These columns are Home Phone, Photo, Report To.

❖Supplier table:

Supplier table also contain these columns that are not necessary for visualization. These columns are Phone, Fax, Home Page.

❖**Categories table:** categories table contain one column that are not necessary for visualization**.**

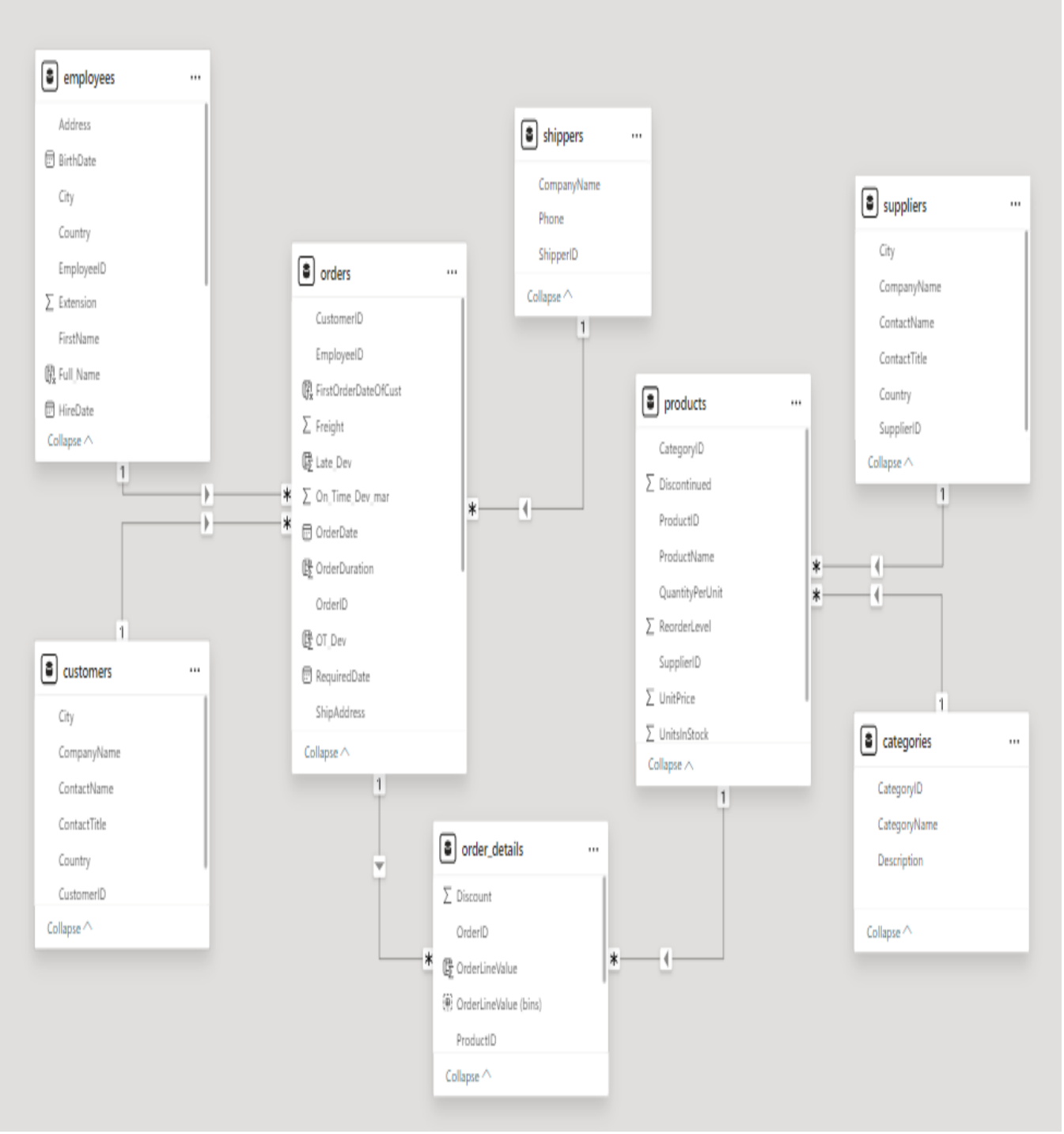
After cleaning all the table are used for the visualization in power BI and build a relationship between the tables :

- Employees table connected to the orders table.
- Customer table connected to the orders table.
- Supplier table connected to the Product table.

- Product table connected to the order details and categories table.
- Order details table connected to the orders table.
- Shipper table also connected to the orders table.

All the tables are connected with each in a one-to-many relationship.

❖ Entity Relationship Diagram of Transformed dataset (North wind traders)

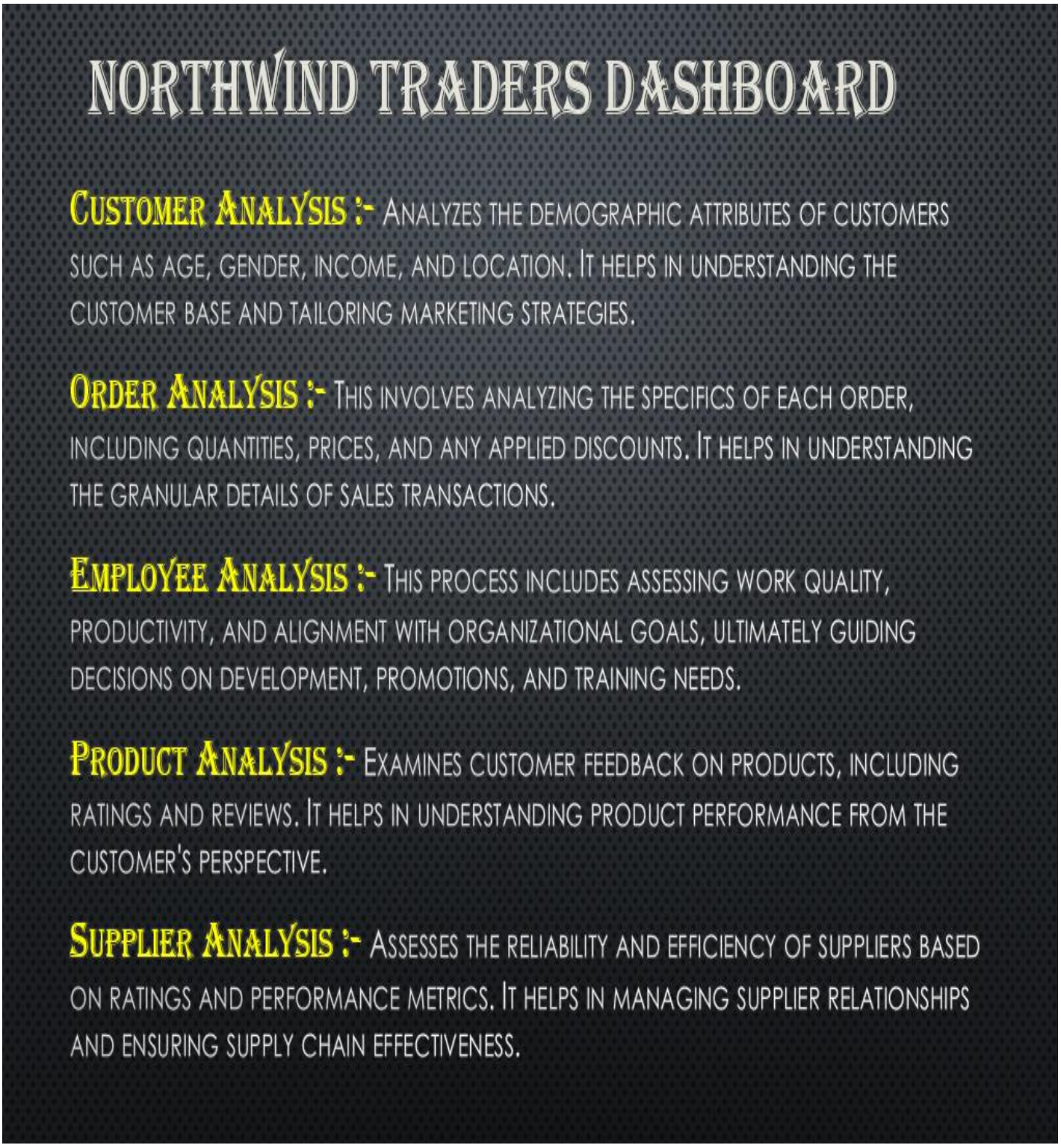


MECE-Breakdown:

MECE stands for 'Mutually Exclusive, Collectively Exhaustive.' This principle is commonly used in problem-solving and structuring information. It ensures that all elements being considered are distinct from one another (mutually exclusive) and, together, they cover all possible options (collectively exhaustive)

Customer Analysis	**Customer Demographics**	**Customer Purchase Behaviour**	**Customer Lifetime Value**
	Segment customers based on demographics (age, gender, location).	Analyze purchasing patterns and preferences. Evaluate frequency and volume of purchases.	Calculate the lifetime value of different customer segments.
Order Analysis	**Order Details Analysis**	**Order Trends**	
	Analyze order quantities, prices, and discounts. Evaluate order completion rates and average order value.	Assess order trends over time (seasonal, monthly, yearly). Identify peak sales periods and slow periods.	
Employee Analysis	**Employee Productivity**	**Employee Tenure**	**Employee Performance**
	Assess employee productivity metrics across various departments and job roles.	Evaluate the length of time employees have been with the company.	Evaluate sales performance by employee. Assess impact of employee interactions on sales.
Product Analysis	**Categories Analysis**	**Product Performance**	**Product Pricing**
	Assess sales performance by product categories. Analyze trends within each category.	Evaluate individual product sales. Identify top-selling and low-performing products.	Evaluate the distribution of product prices to understand pricing strategies and market positioning.
Supplier Analysis	**Supplier Performance**	**Supplier Cost Structure**	**Geographical Distribution**
	Evaluate the ratings or performance metrics of suppliers to understand their reliability and efficiency.	Assess the cost or pricing structures of various suppliers to identify cost-effective and high-cost suppliers.	Map out the locations of suppliers to understand their geographical spread and potential impact on logistics.

DASHBOARD

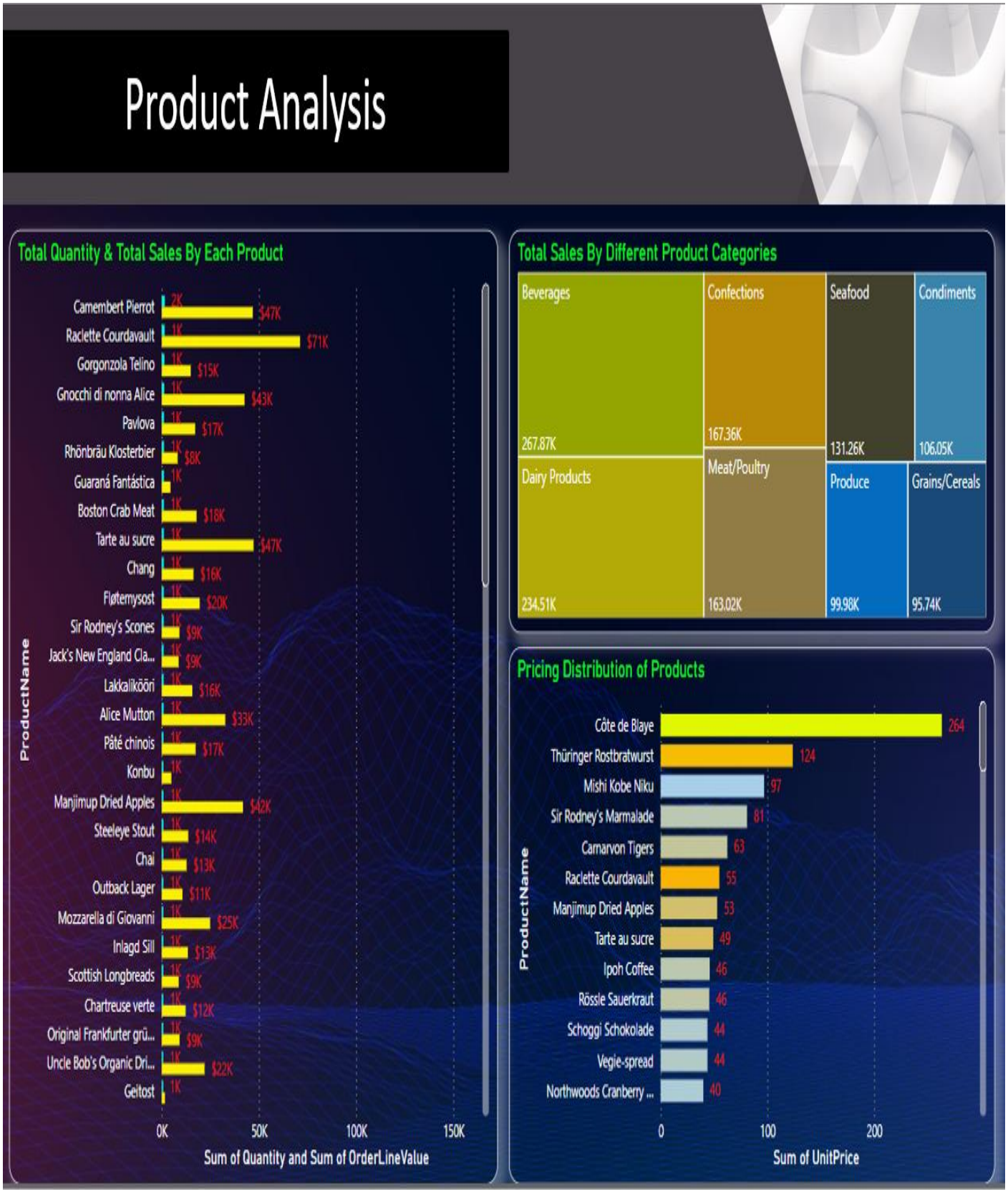

NORTHWIND TRADERS DASHBOARD
CUSTOMER ANALYSIS :- ANALYZES THE DEMOGRAPHIC ATTRIBUTES OF CUSTOMERS SUCH AS AGE, GENDER, INCOME, AND LOCATION. IT HELPS IN UNDERSTANDING THE CUSTOMER BASE AND TAILORING MARKETING STRATEGIES.
ORDER ANALYSIS :- THIS INVOLVES ANALYZING THE SPECIFICS OF EACH ORDER, INCLUDING QUANTITIES, PRICES, AND ANY APPLIED DISCOUNTS. IT HELPS IN UNDERSTANDING THE GRANULAR DETAILS OF SALES TRANSACTIONS.
EMPLOYEE ANALYSIS :- THIS PROCESS INCLUDES ASSESSING WORK QUALITY, PRODUCTIVITY, AND ALIGNMENT WITH ORGANIZATIONAL GOALS, ULTIMATELY GUIDING DECISIONS ON DEVELOPMENT, PROMOTIONS, AND TRAINING NEEDS.
PRODUCT ANALYSIS :- EXAMINES CUSTOMER FEEDBACK ON PRODUCTS, INCLUDING RATINGS AND REVIEWS. IT HELPS IN UNDERSTANDING PRODUCT PERFORMANCE FROM THE CUSTOMER'S PERSPECTIVE.
SUPPLIER ANALYSIS :- ASSESSES THE RELIABILITY AND EFFICIENCY OF SUPPLIERS BASED ON RATINGS AND PERFORMANCE METRICS. IT HELPS IN MANAGING SUPPLIER RELATIONSHIPS AND ENSURING SUPPLY CHAIN EFFECTIVENESS.

CUSTOMER ANALYSIS DASHBOARD

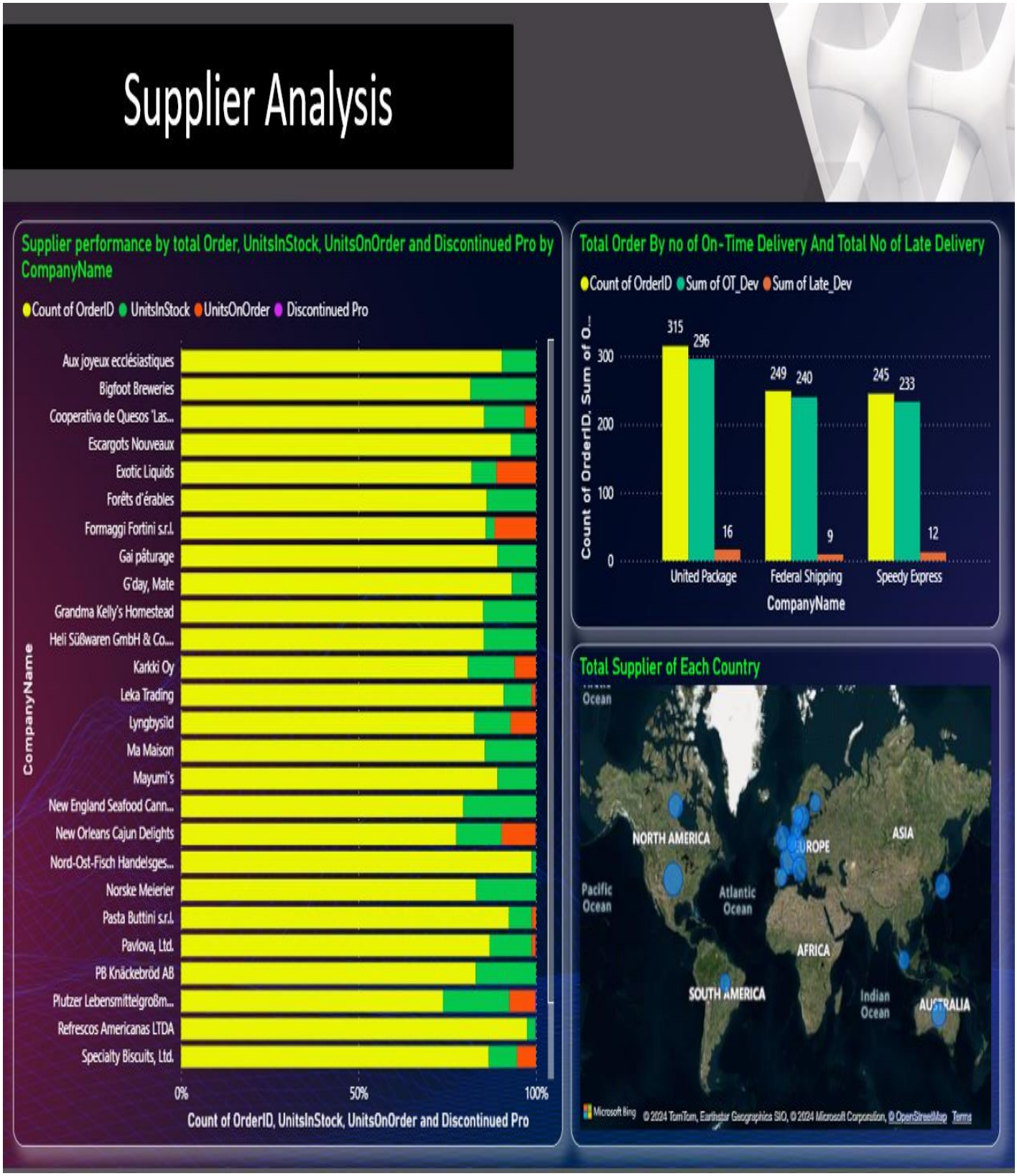
ORDER ANALYSIS DASHBOARD

EMPLOYEE ANALYSIS DASHBOARD

ORDER ANALYSIS DASHBOARD



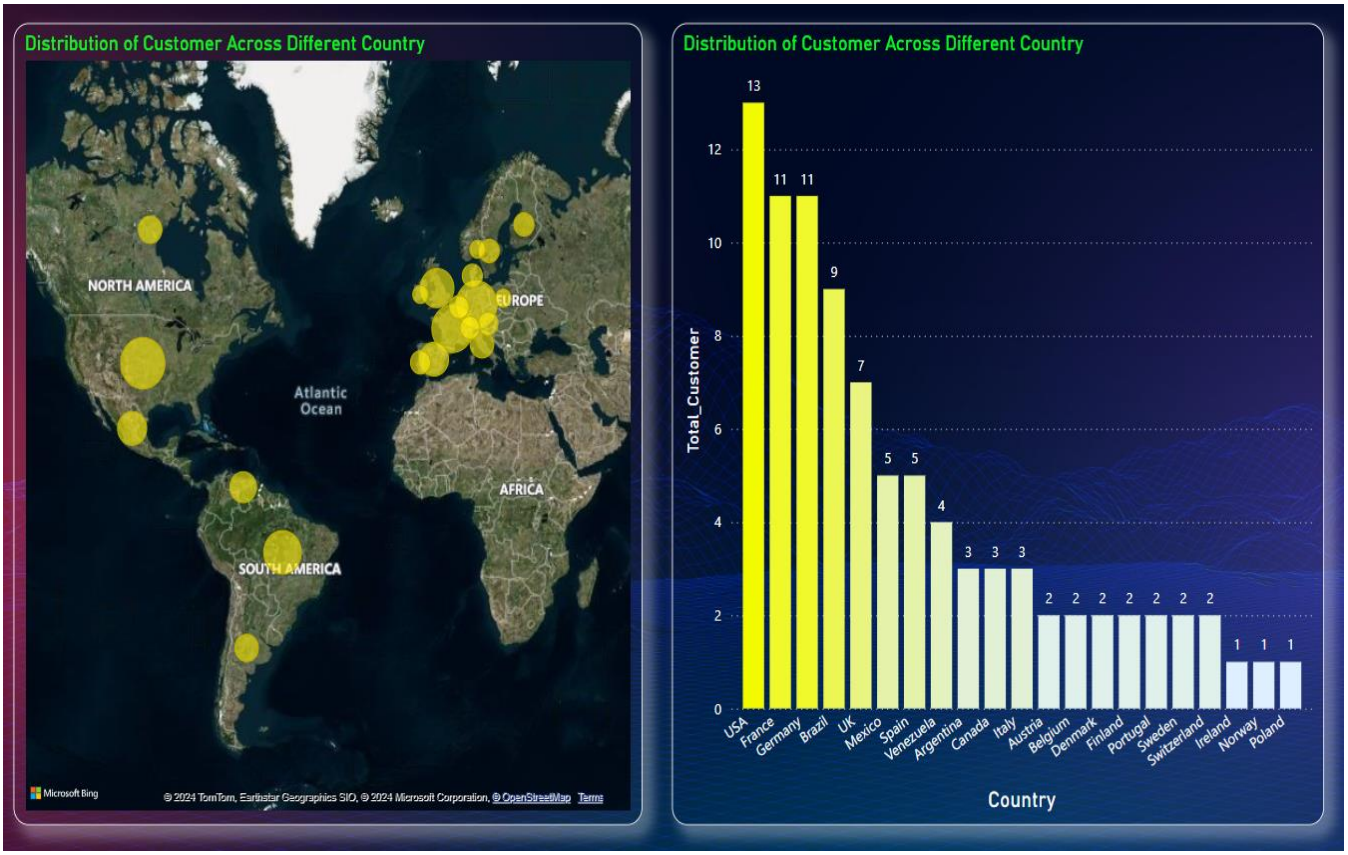
SUPPLIER ANALYSIS DASHBOARD



POWER BI QUESTIONS

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?**
4. **How does order volume change over time? Can we create a time series chart or stacked bar chart to visualize it?**
5. **What is the distribution of order values? Can we create a histogram or box plot to display it?**
6. **Can we visualize the average order processing time or shipping duration using a bar chart or box plot?**
7. **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?**
8. **What is the distribution of employee tenure? Can we create a histogram or box plot to display it?**
9. **Can we visualize employee performance ratings or KPIs using a radar chart or bullet graph?**
10. **What is the distribution of product ratings or reviews? Can we create a histogram or stacked bar chart to visualize it?**
11. **How does the sales volume vary across different product categories? Can we create a bar chart or treemap to display it?**
12. **Can we visualize the pricing distribution of products using a box plot or violin plot?**
13. **What is the distribution of supplier ratings or performance metrics? Can we create a bar chart or radar chart to visualize it?**
14. **How does the cost or pricing structure vary across different suppliers? Can we create a box plot or stacked bar chart to display it?**
15. **Can we visualize the geographical distribution of suppliers using a map or bubble chart?**

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

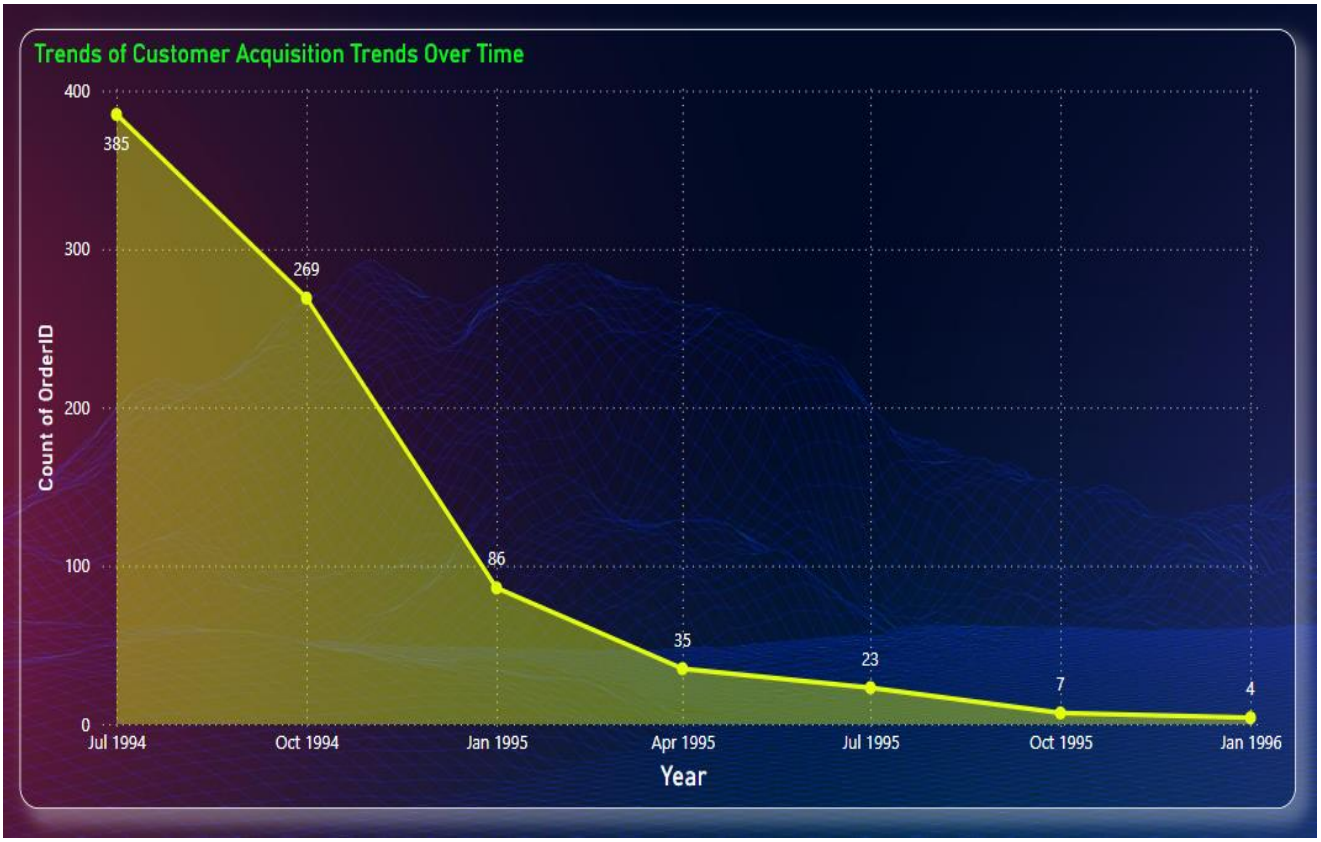


Conclusions:

The customer distribution indicates that the company has a strong presence in Europe, particularly in France, Germany, and the UK. North America, led by the USA, also shows significant customer numbers. There are opportunities to enhance customer engagement and expand the customer base in regions with lower counts, such as some European countries and South American countries.

.

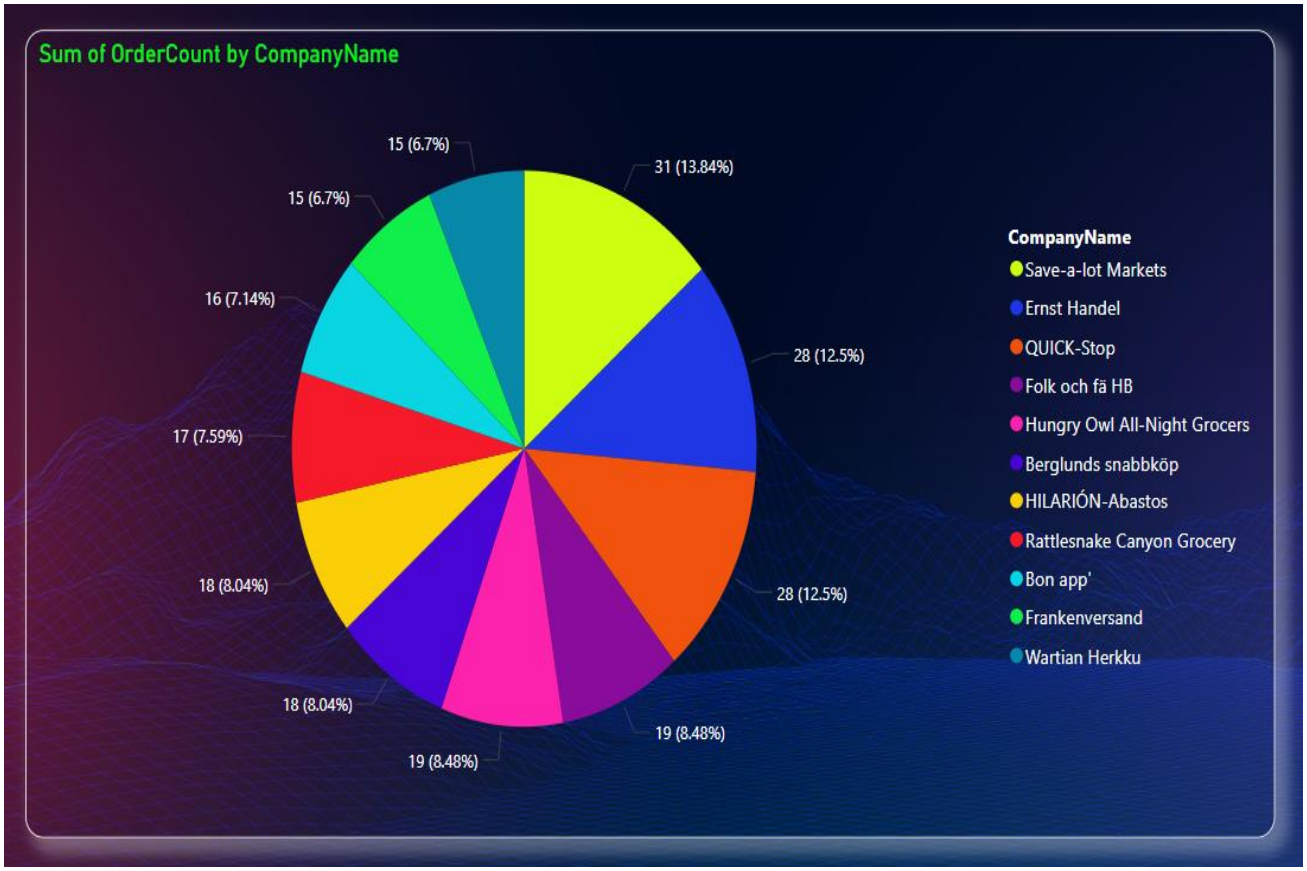
What is the trend in customer acquisition over time? Can we create a line chart or area chart to display it?



Conclusions:

The trend in customer acquisition shows a significant decline from 1994 to early 1996. While the company experienced strong initial growth in 1994, it faced challenges in maintaining this momentum in subsequent quarters. This downward trend suggests a need to investigate and address potential causes, such as changes in market conditions, customer preferences, or internal factors affecting sales and marketing effectiveness.

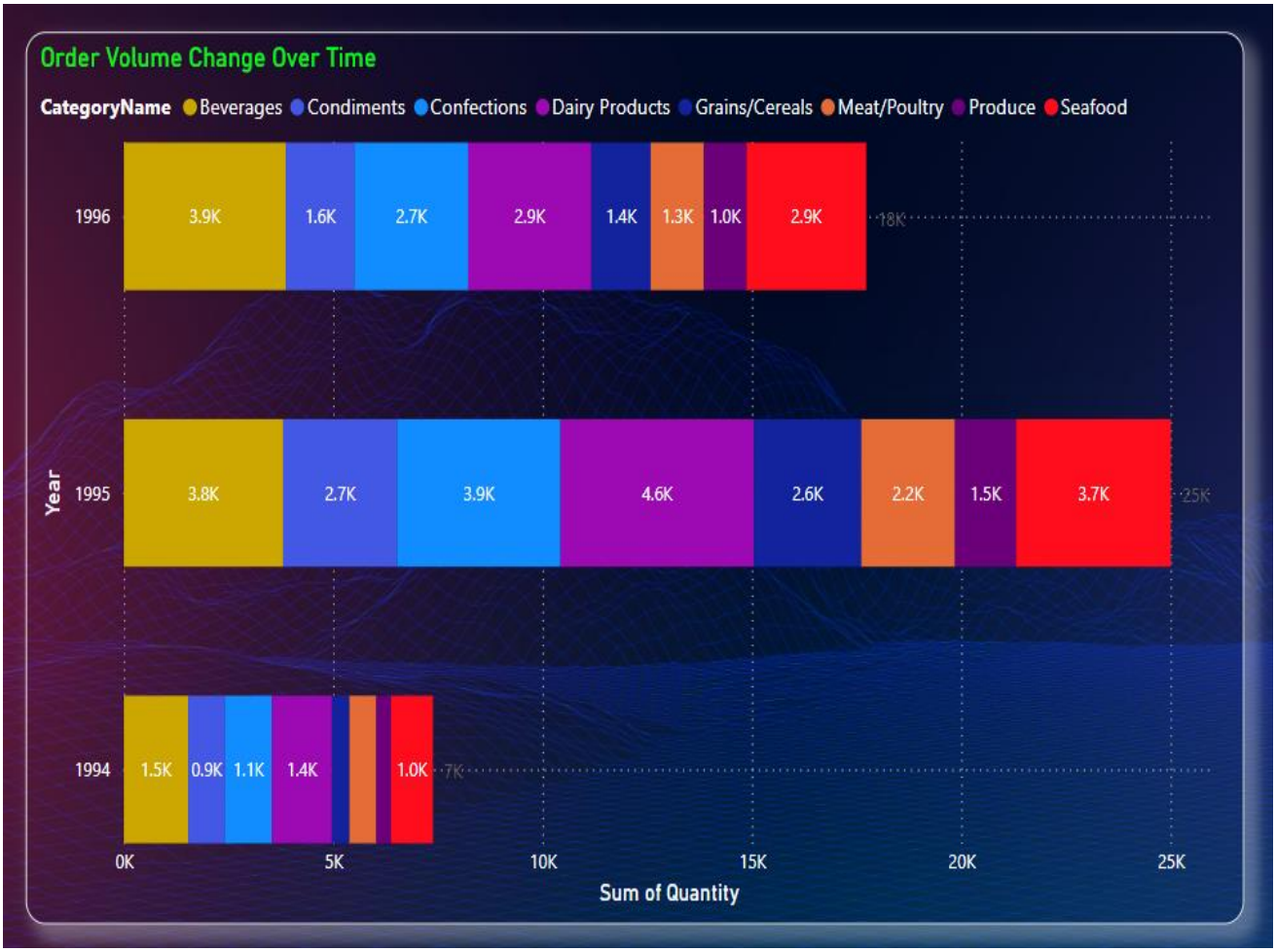
Top 10 customers by order count.



Conclusions:

The top 10 customers by order count show a diverse range of companies with varying levels of engagement. Save-a-lot Markets stands out as the most active customer, followed by Ernst Handel and QUICK-Stop. The order counts reflect strong relationships and repeat business from these key customers, which are crucial for the company's sales stability and growth.

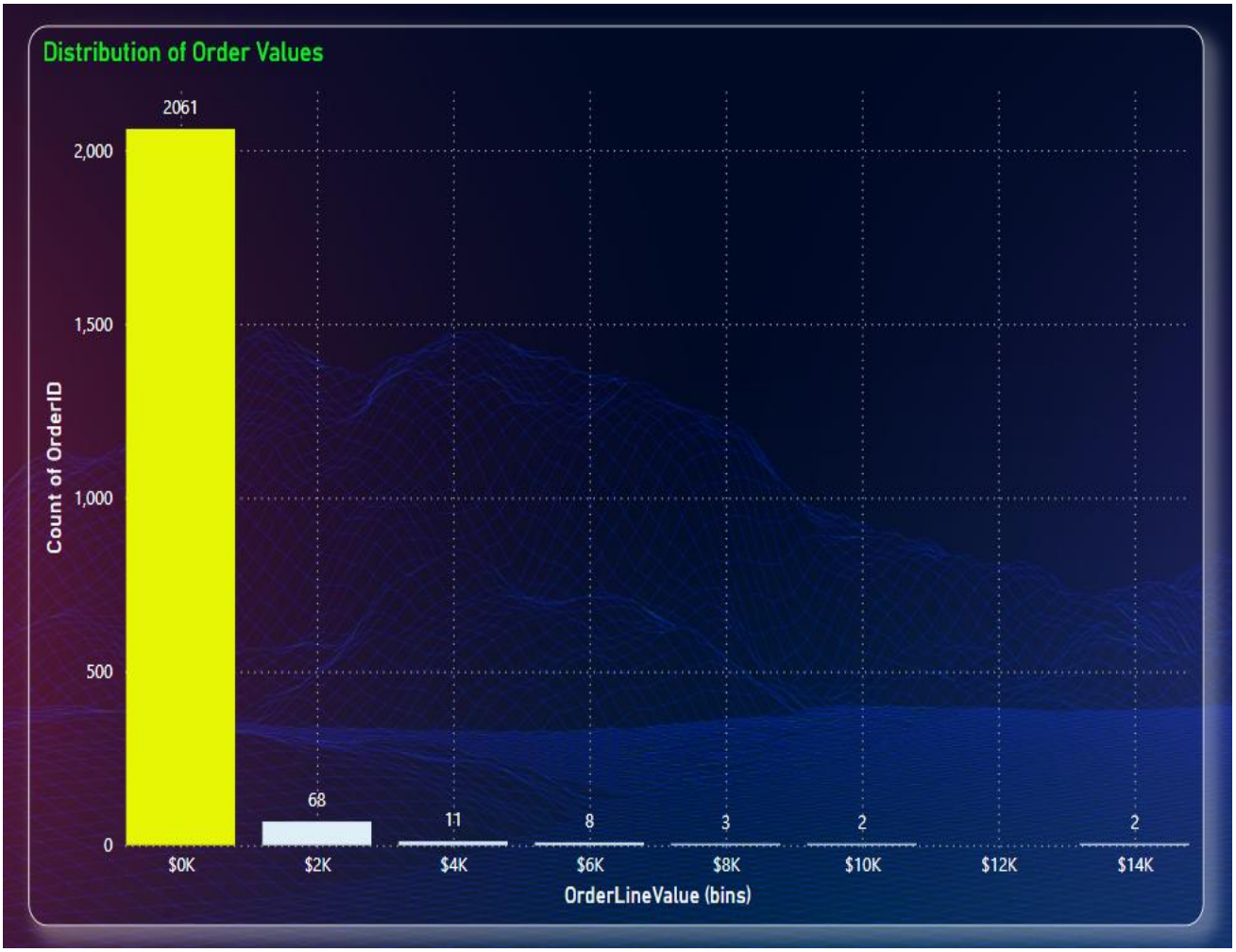
How does order volume change over time? Can we create a time series chart or stacked bar chart to visualize it?



Conclusions:

The overall order volume for the company increased significantly from 1994 to 1995, with 1995 being the peak year for most product categories. However, there is a slight decline in order volume in 1996 compared to 1995, though it remains higher than 1994. This trend suggests that the company experienced rapid growth in 1995, followed by a stabilization or slight reduction in order volume in the following year.

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



Conclusions:

The distribution of order values is highly skewed towards lower value orders, with the majority of orders falling below 2000. There is a sharp decline in the number of orders as the order value increases. High value orders are very rare, suggesting that large purchases are uncommon.

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

Conclusions:

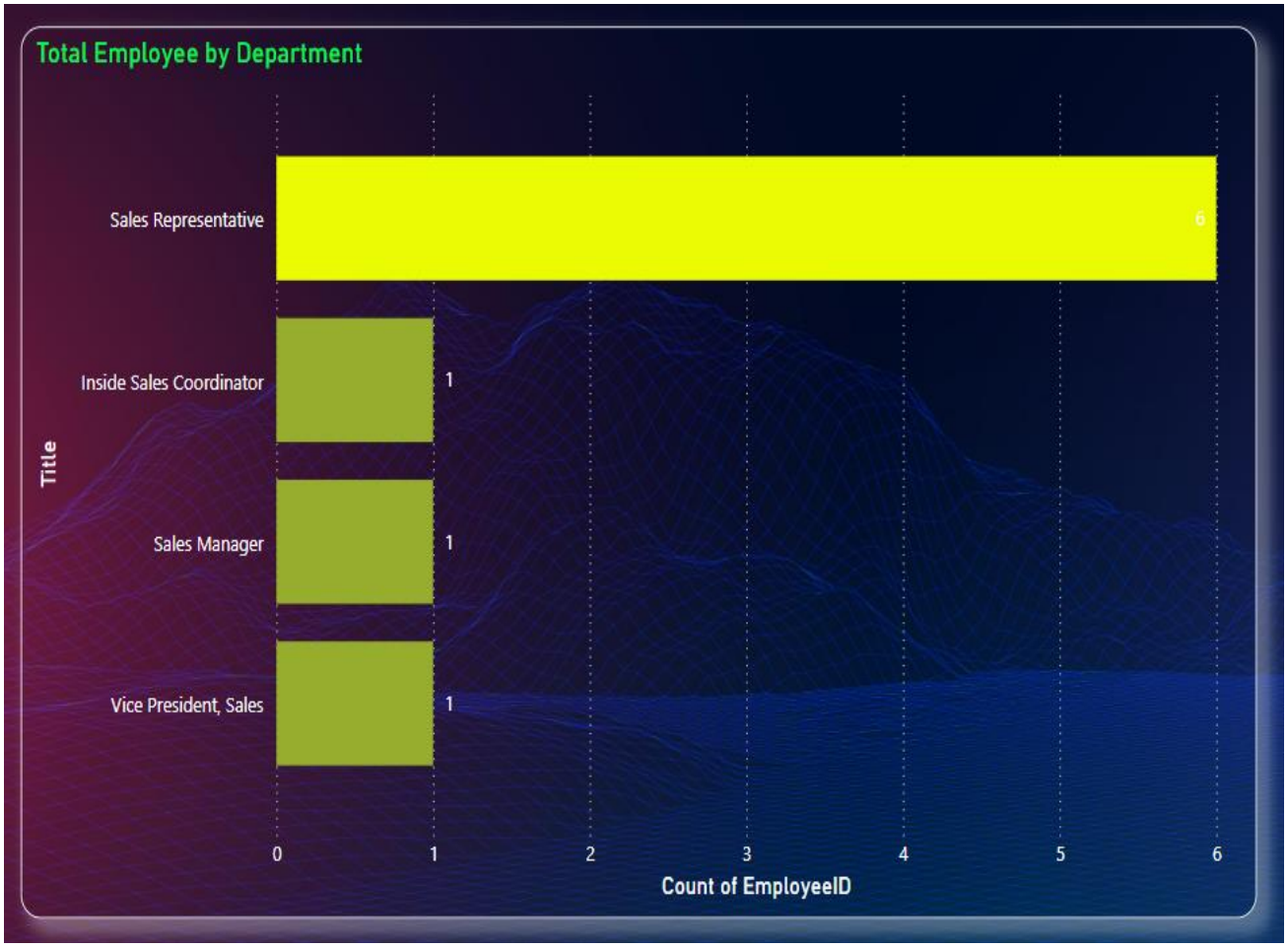
The average shipping duration time of 8.49 days can provide valuable insights into the efficiency of the order processing and shipping process. Using appropriate visualization techniques, stakeholders can easily identify areas for improvement and track changes over time.

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

Conclusions:

Employee productivity varies significantly across different job roles. Sales Representatives are the most productive in terms of both quantity and value, indicating their crucial role in driving sales. The Vice President, Sales handles fewer orders but contributes high-value transactions. Inside Sales Coordinators have balanced productivity, while Sales Managers show the lowest productivity in terms of both quantity and value.

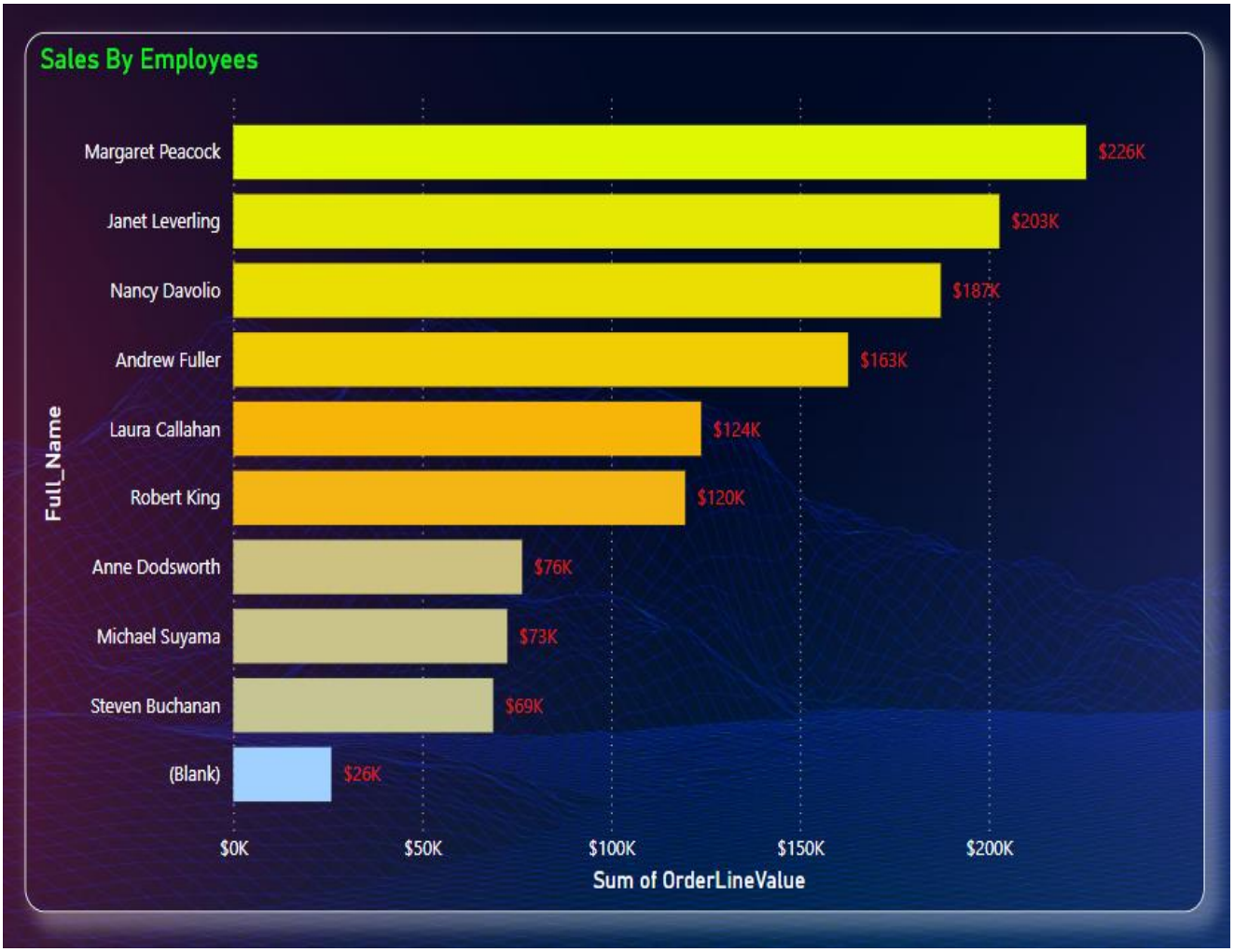
How dose employee distribution by job tittle



Conclusions:

The employee distribution by job title shows that the company employs a significantly higher number of Sales Representatives compared to other roles. With 6 out of 9 employees, Sales Representatives form the core of the sales team. In contrast, the roles of Inside Sales Coordinator, Sales Manager, and Vice President, Sales each have only one employee, indicating specialized roles with unique responsibilities.

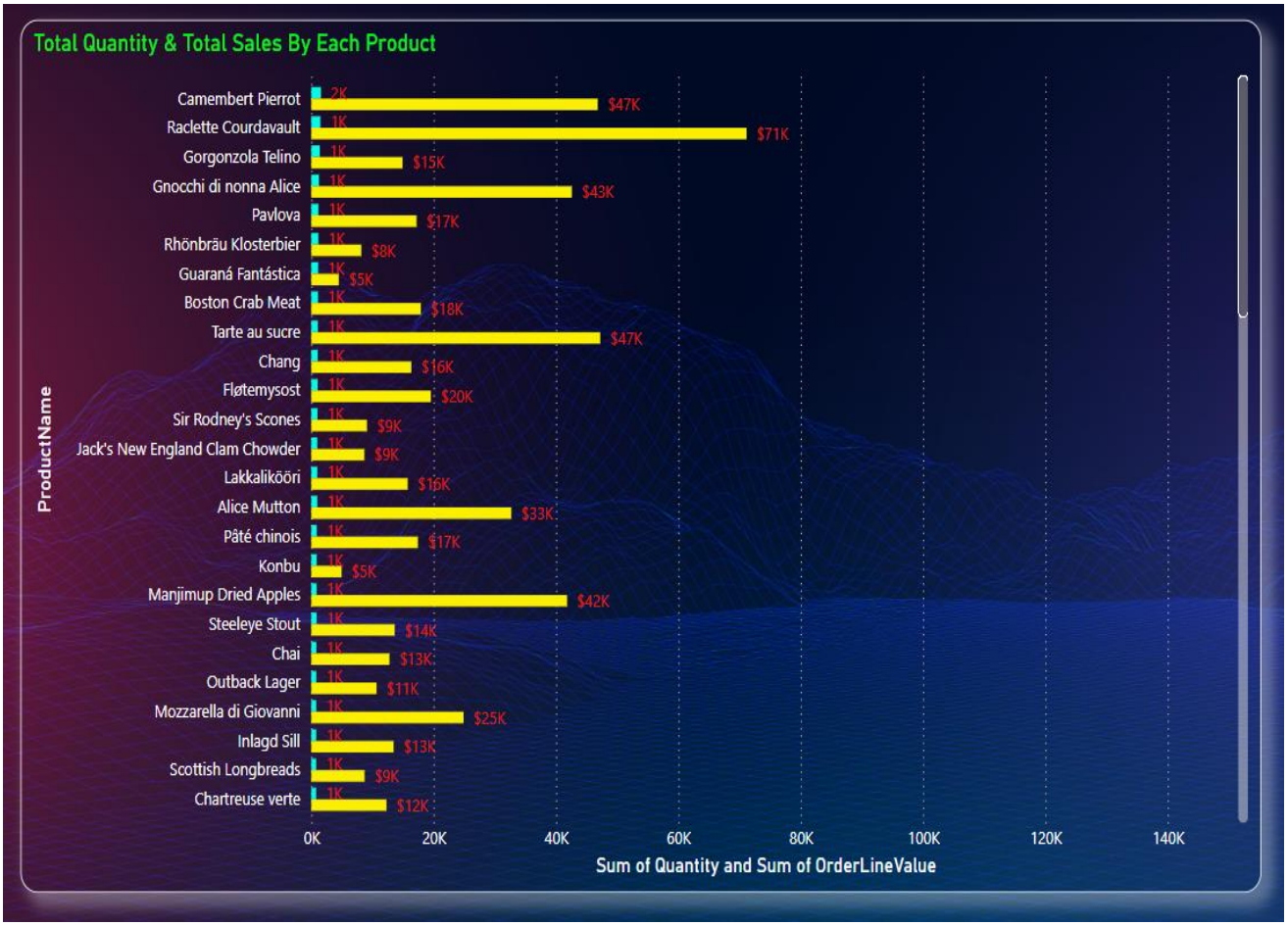
How many sales has each employee?



Conclusions:

The distribution of sales among employees varies significantly, with some employees contributing substantially more to the total sales value than others. Margaret Peacock stands out as the highest contributor, followed by Janet Leverling and Nancy Davolio. The data indicates that sales performance is not evenly distributed among employees.

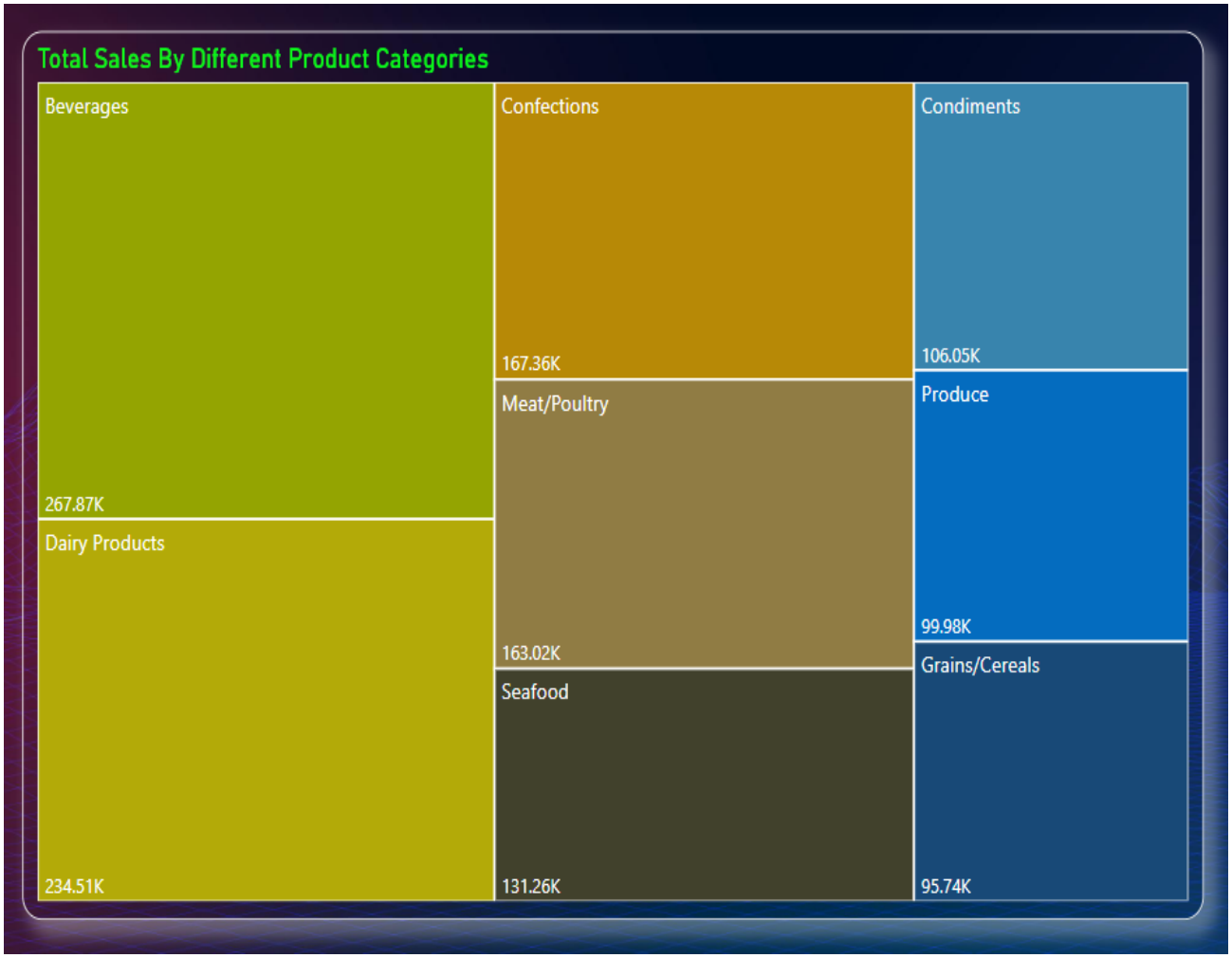
How much each product quantity and sales



Conclusions:

The data shows that **Côte de Blaye** is the top-selling product by sales value, while **Camembert Pierrot** has the highest quantity sold. Most products maintain a consistent quantity of 1k, but their sales values vary significantly, indicating differences in pricing or demand.

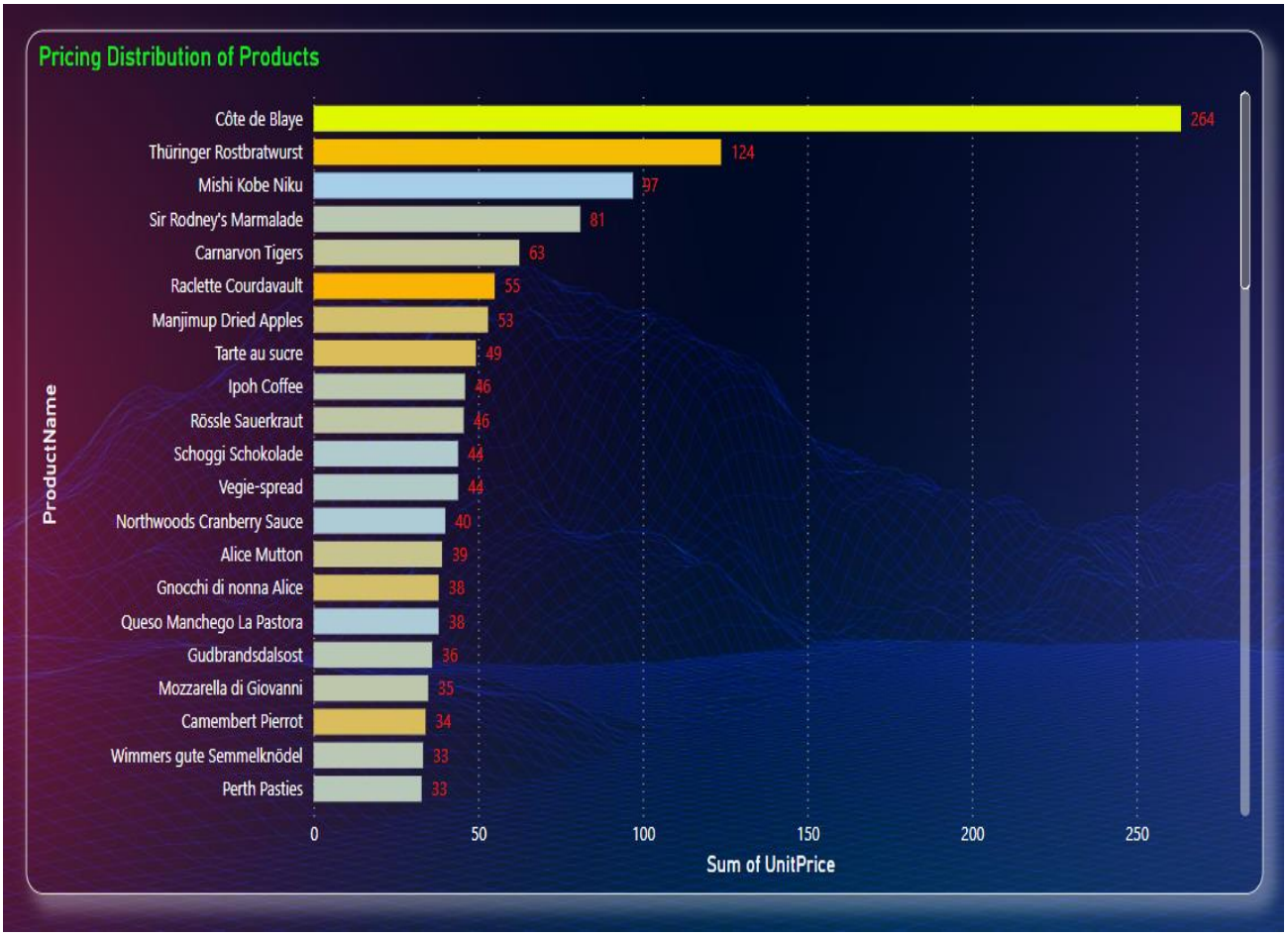
How does the sales volume vary across different product categories?



Conclusions:

The sales volume varies significantly across different product categories, with **Beverages** leading in total order value. **Dairy Products**, **Confections**, and **Meat/Poultry** are also strong contributors to the overall sales. In contrast, **Grains/Cereals** and **Produce** have lower sales volumes compared to other categories.

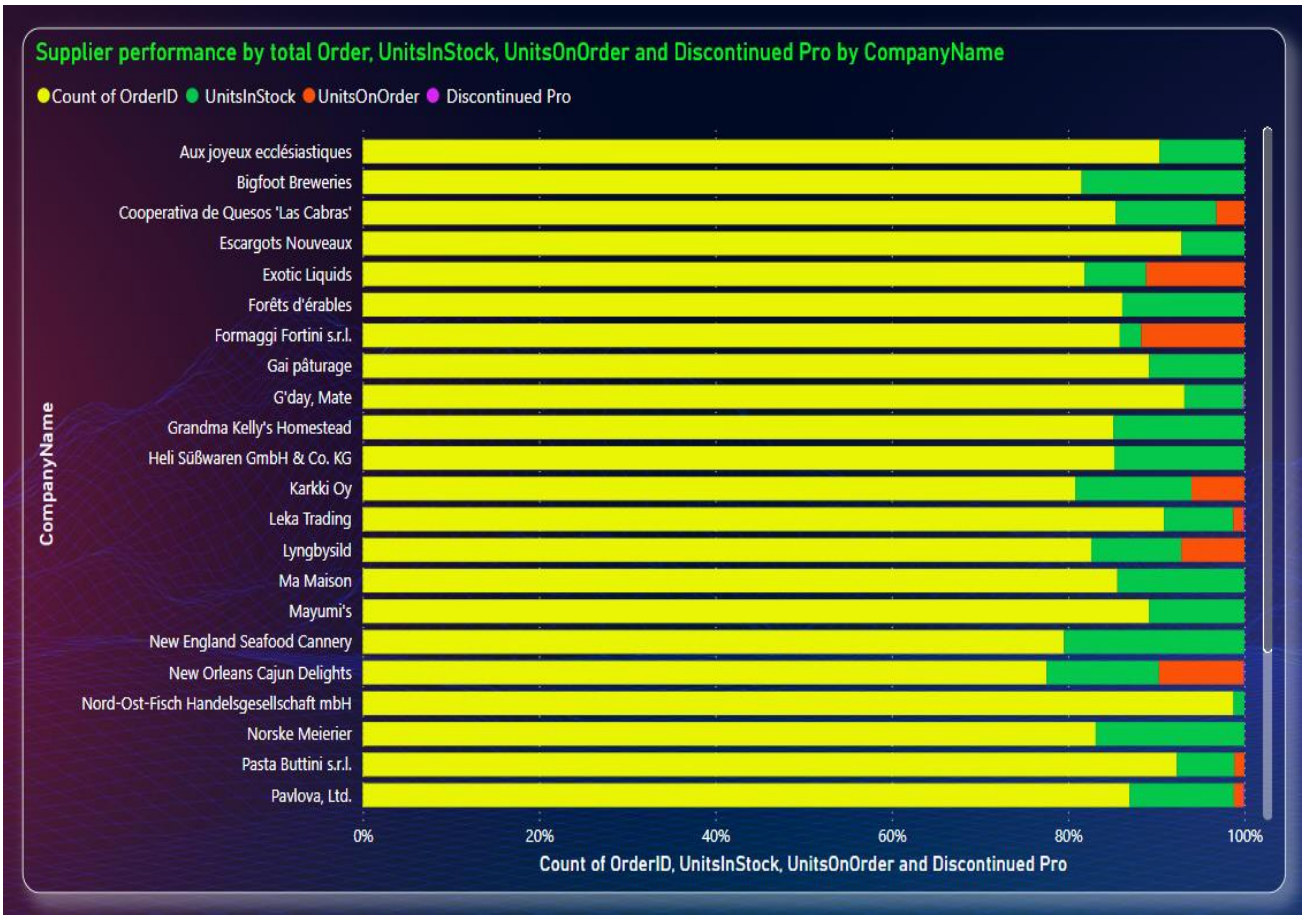
Can we visualize the pricing distribution of products using a box plot or violin plot?



Conclusions:

The pricing distribution of products shows significant variability across the range. The prices vary widely, with a minimum of $2.50 and a maximum of $263.50, indicating a broad price range of $261.00. The mean price is approximately $33.75, but this is skewed upwards by a few high-priced items. The median price is about $21.05, suggesting that more than half of the products are priced below this amount, with many clustered in the lower to mid-price range. The high standard deviation of approximately $45.85 reflects considerable variability in prices. Overall, while most products are priced affordably, the presence of a few premium items results in a diverse pricing structure.

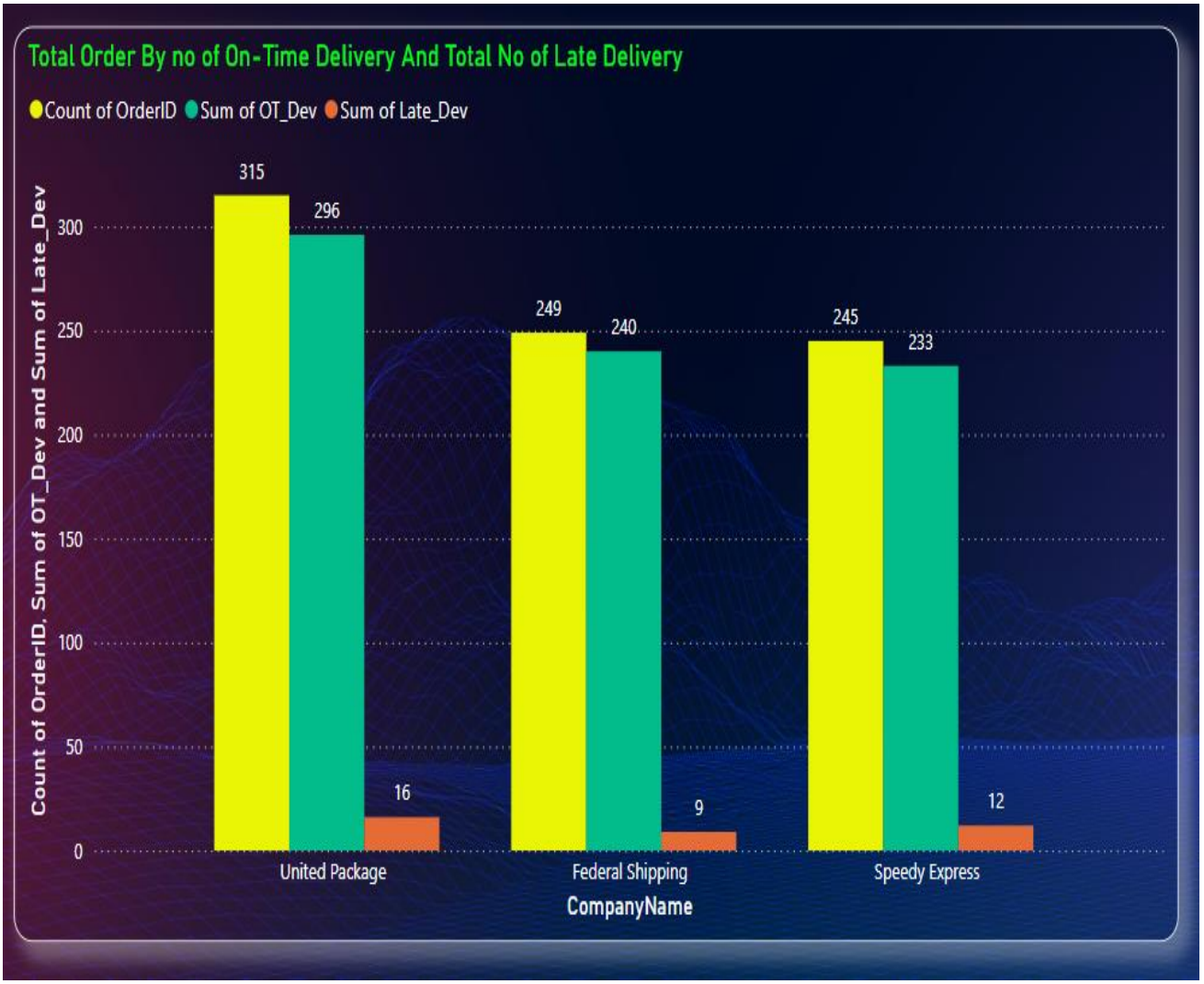
How dose the performance of various supplier company very?



Conclusions:

The pricing distribution shows significant variation in inventory and ordering across companies. **Svensk Sjöföda AB** has the highest stock with 224 units, while **Nord-Ost-Fisch Handelsgesellschaft mbH** has the lowest with 10 units. **Exotic Liquids** has the largest order volume of 110 units, whereas several companies have none. Discontinued products are minimal, affecting only 8 items in total.

What is the distribution of supplier delivery performance ?



Conclusions:

The pricing distribution of products indicates a notable variance in delivery metrics across companies. **United Package** has the highest **Sum of OT_Dev** with 296, suggesting frequent pricing deviations, while **Federal Shipping** has the lowest with 240. **United Package** also shows the highest **Sum of Late_Dev** at 16, indicating more late deliveries compared to other companies. Overall, **Federal Shipping** has the lowest deviation and late delivery figures, suggesting more consistent pricing and delivery performance.

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

Conclusions:

The pricing distribution of products across countries shows that the **USA** has the highest number of suppliers with 4, indicating a significant presence in the market. **France** and **Germany** each have 3 suppliers, reflecting strong representation in those regions. Other countries like **Australia**, **Canada**, **Italy**, **Japan**, and **Sweden** each have 2 suppliers. Several countries, including **Brazil**, **Denmark**, **Finland**, **Netherlands**, **Norway**, **Singapore**, **Spain**, and the **UK**, have 1 supplier each. Overall, the distribution is varied, with the USA leading in supplier count.

EDA QUESTIONS

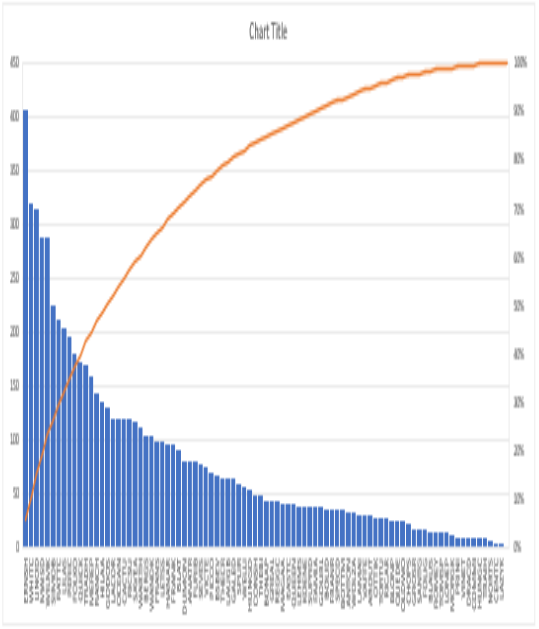
1. **What are the key factors influencing customer retention or loyalty based on the dataset?**
2. **How do customer preferences vary based on their location or demographics? Can we explore this through interactive visualizations?**
3. **Are there any interesting patterns or clusters in customer behavior that can be visualized to identify potential market segments?**
4. **Are there any specific product categories or SKUs that contribute significantly to order revenue? Can we identify them through visualizations?**
5. **Are there any correlations between order size and customer demographics or product categories? Can we explore this visually using scatter plots or heatmaps?**
6. **How does order frequency vary across different customer segments? Can we visualize this using bar charts or treemaps?**
7. **Are there any correlations between employee satisfaction levels and key performance indicators? Can we explore this visually through scatter plots or line charts?**
8. **How does employee turnover vary across different departments or job roles? Can we visualize this using bar charts or heatmaps?**
9. **Can we identify any patterns or clusters in employee skill sets or qualifications through visualizations? How can this information be used for talent management?**
10. **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?**
11. **How does product demand fluctuate over different seasons or months? Can we visualize this through line charts or area charts?**
12. **Can we identify any outliers or anomalies in product performance or sales using visualizations? How can this information be used for product optimization?**
13. **Are there any correlations between supplier attributes (e.g., location, size, industry) and performance metrics (e.g., on-time delivery, product quality)? Can we explore this visually through scatter plots or heatmaps?**
14. **How does supplier performance vary across different product categories or departments? Can we visualize this using stacked bar charts or grouped column charts?**
15. **Can we identify any trends or patterns in supplier costs or pricing structures through visualizations? How can this information be used for procurement optimization?**

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

```
WITH
table_x AS (
  SELECT o.CustomerID,
    TIMESTAMPDIFF(DAY, o.ShippedDate, o.RequiredDate)
AS Delivery_Time,
    o.OrderDate,
    od.ProductID,
    COUNT(o.OrderID) OVER(PARTITION BY o.CustomerID) AS
No_Of_Orders
  FROM orders o
  JOIN order_details od ON o.OrderID = od.OrderID
  JOIN Products p ON p.ProductID = od.ProductID
),
table_2 AS (
  SELECT DISTINCT o.CustomerID, cs.CompanyName ,
od.ProductID , p.ProductName ,
      COUNT(od.ProductID) OVER(PARTITION BY
o.CustomerID, od.ProductID) AS NoOfTimesPurchased
  FROM orders o
  JOIN order_details od ON o.OrderID = od.OrderID
  JOIN Products p ON p.ProductID = od.ProductID
  JOIN customers cs ON cs.CustomerID=o.CustomerID
),
table_3 AS (
  SELECT CustomerID, CompanyName , ProductID ,
ProductName ,
    DENSE_RANK() OVER(PARTITION BY CustomerID
ORDER BY NoOfTimesPurchased DESC) AS Preference_Rank
  FROM table_2
),
table_y AS (
  SELECT CustomerID,
    SUM(CASE WHEN Delivery_Time < 0 THEN 1 ELSE 0
END) AS Number_Of_Late_Deliveries,
    MIN(OrderDate) AS Earliest_Order_Date ,
    MAX(OrderDate) AS Latest_Order_Date ,
    No_Of_Orders
  FROM table_x
  GROUP BY CustomerID,No_Of_Orders
),
table_5 AS (
  SELECT CustomerID, CompanyName , ProductID ,
ProductName
  FROM table_3
  WHERE Preference_Rank = 1
)

SELECT table_y.* , table_5.CompanyName ,
  TIMESTAMPDIFF(MONTH, Earliest_Order_Date,
Latest_Order_Date) AS Retention_Period_In_Months,
  table_5.ProductName AS 'Most_Preferred_Product'
FROM table_y
JOIN table_5 ON table_y.CustomerID = table_5.CustomerID;
```

CustomerID	Number_Of_Late_Deliveries	Earliest_Order_Date	Latest_Order_Date	No_Of_Orders	CompanyName	Retention_Period_In_Months	Most_Preferred_Product
ALFKI	0	25-08-1995	09-05-1996	12	Alfreds Futterkist	7	RÃ/Ã¶gsås Sauerkraut
ANATR	0	19-10-1994	03-04-1996	10	Ana Trujillo Emp	17	Mozzarella di Giovanni
ANATR	0	19-10-1994	03-04-1996	10	Ana Trujillo Emp	17	Outback Lager
ANATR	0	19-10-1994	03-04-1996	10	Ana Trujillo Emp	17	Gudbrandsdalsost
ANATR	0	19-10-1994	03-04-1996	10	Ana Trujillo Emp	17	Camembert Pierrot
ANATR	0	19-10-1994	03-04-1996	10	Ana Trujillo Emp	17	Singaporean Hokkien Fried Mee
ANATR	0	19-10-1994	03-04-1996	10	Ana Trujillo Emp	17	Mascarpone Fabioli
ANATR	0	19-10-1994	03-04-1996	10	Ana Trujillo Emp	17	Teatime Chocolate Biscuits
ANATR	0	19-10-1994	03-04-1996	10	Ana Trujillo Emp	17	Tofu
ANATR	0	19-10-1994	03-04-1996	10	Ana Trujillo Emp	17	Konbu
ANATR	0	19-10-1994	03-04-1996	10	Ana Trujillo Emp	17	Queso Cabrales
ANTON	0	28-12-1994	28-02-1996	17	Antonio Moreno`	14	Geitost
ANTON	0	28-12-1994	28-02-1996	17	Antonio Moreno`	14	Queso Cabrales
AROUT	0	16-12-1994	10-05-1996	30	Around the Horn	16	Gorgonzola Telino
BERGS	6	12-09-1994	03-04-1996	52	Berglunds snabb	18	RÃ/Ã¶dkaviar/Ã¤u Kloskebier
BERGS	6	12-09-1994	03-04-1996	52	Berglunds snabb	18	TourtiÃ/Ã¨re
BLAUS	0	10-05-1995	29-05-1996	14	Blauer See Delik	12	Sir Rodney's Scones
BLONP	0	25-08-1994	12-02-1996	26	Blondel pÃ/Ã¨re	17	Gorgonzola Telino
BOLID	1	10-11-1994	23-04-1996	6	BÃ/Ã¶lido Comi	17	RÃ/Ã¶dkaviar/Ã¤u Kloskebier
BOLID	1	10-11-1994	23-04-1996	6	BÃ/Ã¶lido Comi	17	Ravioli Angelo
BOLID	1	10-11-1994	23-04-1996	6	BÃ/Ã¶lido Comi	17	Filo Mix
BOLID	1	10-11-1994	23-04-1996	6	BÃ/Ã¶lido Comi	17	TMÃ/Ã¼nger Rostbratwurst
BOLID	1	10-11-1994	23-04-1996	6	BÃ/Ã¶lido Comi	17	Alice Mutton
BOLID	1	10-11-1994	23-04-1996	6	BÃ/Ã¶lido Comi	17	Chef Anton's Cajun Seasoning
BONAP	5	16-11-1994	05-06-1996	44	Bon app'	18	Pavlova
BOTTM	0	20-01-1995	24-05-1996	35	Bottom-Dollar Ma	16	Torta au sucre
BSBEV	2	26-09-1994	14-05-1996	22	B's Beverages	19	Sir Rodney's Scones
BSBEV	2	26-09-1994	14-05-1996	22	B's Beverages	19	Konbu
BSBEV	2	26-09-1994	14-05-1996	22	B's Beverages	19	Uncle Bob's Organic Dried Pears
CACTU	0	30-05-1995	28-05-1996	11	Cactus Comidas	11	RÃ/Ã¶dkaviar/Ã¤u Kloskebier
CACTU	0	30-05-1995	28-05-1996	11	Cactus Comidas	11	RÃ/Ã¶d Kaviar
CACTU	0	30-05-1995	28-05-1996	11	Cactus Comidas	11	Scottish Longbreads
CACTU	0	30-05-1995	28-05-1996	11	Cactus Comidas	11	Laughing Lumberjack Lager
CACTU	0	30-05-1995	28-05-1996	11	Cactus Comidas	11	Ipoh Coffee
CACTU	0	30-05-1995	28-05-1996	11	Cactus Comidas	11	Jack's New England Clam Chowder
CACTU	0	30-05-1995	28-05-1996	11	Cactus Comidas	11	Steeleye Stout
CACTU	0	30-05-1995	28-05-1996	11	Cactus Comidas	11	Sasquatch Ale
CACTU	0	30-05-1995	28-05-1996	11	Cactus Comidas	11	Geitost
CACTU	0	30-05-1995	28-05-1996	11	Cactus Comidas	11	Gorgonzola Telino
CACTU	0	30-05-1995	28-05-1996	11	Cactus Comidas	11	RÃ/Ã¶gsås Sauerkraut
CENTC	0	18-08-1994	18-08-1994	2	Centro comercial	0	Gravad lax
CENTC	0	18-08-1994	18-08-1994	2	Centro comercial	0	Sir Rodney's Scones
CHOPS	0	11-08-1994	22-05-1996	22	Chop-suey Chine	21	Gnocchi di nonna Alice
COMMI	0	27-09-1994	22-05-1996	10	ComÃ/Ãércio M	19	Spegesild
CONSH	0	07-03-1995	23-02-1996	7	Consolidated Ho	11	Mozzarella di Giovanni
CONSH	0	07-03-1995	23-02-1996	7	Consolidated Ho	11	TunnbrÃ/Ã¶d
CONSH	0	07-03-1995	23-02-1996	7	Consolidated Ho	11	GustafÃ¶s KÃ/Ã¤ckebrÃ/Ã¶d
CONSH	0	07-03-1995	23-02-1996	7	Consolidated Ho	11	Konbu
CONSH	0	07-03-1995	23-02-1996	7	Consolidated Ho	11	Mishi Kobe Niku
CONSH	0	07-03-1995	23-02-1996	7	Consolidated Ho	11	Chef Anton's Gumbo Mix
CONSH	0	07-03-1995	23-02-1996	7	Consolidated Ho	11	Chang
DRACD	0	27-12-1994	03-06-1996	10	Drachenblut Deli	17	Konbu
DUMON	0	21-10-1994	18-03-1996	9	Du monde entier	16	Gudbrandsdalsost
DUMON	0	21-10-1994	18-03-1996	9	Du monde entier	16	Filo Mix
DUMON	0	21-10-1994	18-03-1996	9	Du monde entier	16	Singaporean Hokkien Fried Mee
DUMON	0	21-10-1994	18-03-1996	9	Du monde entier	16	Jack's New England Clam Chowder
DUMON	0	21-10-1994	18-03-1996	9	Du monde entier	16	Sasquatch Ale
DUMON	0	21-10-1994	18-03-1996	9	Du monde entier	16	Sir Rodney's Scones
DUMON	0	21-10-1994	18-03-1996	9	Du monde entier	16	Alice Mutton
DUMON	0	21-10-1994	18-03-1996	9	Du monde entier	16	Ikura
DUMON	0	21-10-1994	18-03-1996	9	Du monde entier	16	Chai
EASTC	2	27-12-1994	28-05-1996	21	Eastern Connect	17	FÃ/Ã¤rnyost
EASTC	2	27-12-1994	28-05-1996	21	Eastern Connect	17	Uncle Bob's Organic Dried Pears
ERNSH	0	17-08-1994	04-06-1996	102	Ernst Handel	21	Wimmers gute SemmelknÃ/Ã¶del



CONCLUSION:

Key Findings:

Customer Retention Period: The retention period for customers varies widely, ranging from 0 months to 34 months.

Number of Late Deliveries: Only a small percentage of customers (1.6%) have experienced late deliveries.

Most Preferred Products: Customers have a diverse range of most preferred products, with no single product standing out as a clear favorite.

Number of Orders: The number of orders placed by customers also varies widely, with some customers placing only one order and others placing up to 102 orders.

Insights:

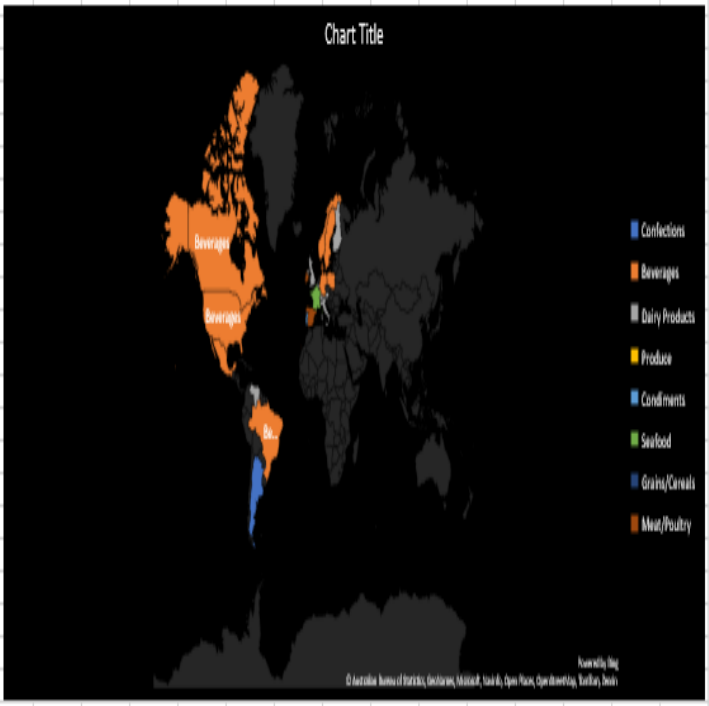
Focus on On-Time Delivery: While only a small percentage of customers have experienced late deliveries, it is essential to maintain a high level of on-time delivery to ensure customer satisfaction and retention.

Personalized Product Recommendations: Given the wide variety of most preferred products, it is important to offer personalized product recommendations based on each customer's purchase history.

Customer Segmentation: Given the wide range in the number of orders placed by customers, it may be beneficial to segment customers based on their purchasing behavior and develop targeted marketing strategies for each segment.

Monitor Customer Retention: Given the wide range in the retention period for customers, it is important to monitor customer retention and develop strategies to retain customers for longer periods.

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




```
WITH CustomerPurchases AS (
    SELECT
        c.CustomerID,
        c.CompanyName,
        c.Country,
        cat.CategoryName,
        SUM(od.Quantity) AS TotalQuantity,
        SUM(od.UnitPrice * od.Quantity * (1 -
od.Discount)) AS TotalValue
    FROM
        customers c
    JOIN
        orders o ON c.CustomerID = o.CustomerID
    JOIN
        order_details od ON o.OrderID = od.OrderID
    JOIN
        products p ON od.ProductID = p.ProductID
    JOIN
        categories cat ON p.CategoryID =
cat.CategoryID
    GROUP BY
        c.CustomerID, c.CompanyName, c.Country,
cat.CategoryName
),
CountryCategoryPreferences AS (
    SELECT
        Country,
        CategoryName,
        SUM(TotalQuantity) AS QuantityPurchased,
        SUM(TotalValue) AS ValuePurchased
    FROM
        CustomerPurchases
    GROUP BY
        Country, CategoryName
    ORDER BY
        Country, ValuePurchased DESC
)

SELECT
    Country,
    CategoryName,
    QuantityPurchased,
    ValuePurchased
FROM
    CountryCategoryPreferences
ORDER BY
```

Count	CategoryNa	QuantityPurch	ValuePurcha
Argentina	Confections	57	2135.1
Argentina	Beverages	82	1796
Argentina	Dairy Products	54	1143.5
Argentina	Produce	33	1139
Argentina	Condiments	45	907
Argentina	Seafood	48	606.5
Argentina	Grains/Cereals	20	390
Austria	Dairy Products	1027	28106.74
Austria	Beverages	982	23317.3
Austria	Condiments	720	15237.2425
Austria	Grains/Cereals	580	14527.7
Austria	Confections	575	13534.4075
Austria	Produce	388	13120.6675
Austria	Meat/Poultry	362	10821.556
Austria	Seafood	533	9338.225
Belgium	Dairy Products	295	8680
Belgium	Confections	270	7479.58
Belgium	Beverages	272	5428.68
Belgium	Grains/Cereals	145	3226
Belgium	Produce	98	3223.2
Belgium	Condiments	147	2454.795
Belgium	Meat/Poultry	89	2138.5
Belgium	Seafood	78	1214.1
Brazil	Beverages	968	37193.445
Brazil	Dairy Products	683	15680.45
Brazil	Seafood	635	12887.135
Brazil	Confections	722	11436.1025
Brazil	Condiments	568	11180.76
Brazil	Meat/Poultry	223	7563.114
Brazil	Grains/Cereals	315	6121.55
Brazil	Produce	133	4883.22
Canada	Beverages	303	11276.06
Canada	Dairy Products	381	9557.175
Canada	Confections	418	8842.48
Canada	Grains/Cereals	207	5557.65
Canada	Condiments	256	4760.075
Canada	Seafood	204	4324.05
Canada	Meat/Poultry	141	3768.8
Canada	Produce	74	2112
Denmark	Beverages	195	12025.7
Denmark	Condiments	210	4325.15
Denmark	Produce	100	3954.9
Denmark	Seafood	230	3857.2625
Denmark	Meat/Poultry	146	3162.5
Denmark	Dairy Products	89	2844.3
Denmark	Confections	185	2591.46
Denmark	Grains/Cereals	15	99.75
Finland	Dairy Products	230	5558.11
Finland	Meat/Poultry	93	3340.78
Finland	Grains/Cereals	100	2477
Finland	Beverages	107	2182.4

CONCLUSION:-

Beverages are popular globally: Beverages are among the top 3 categories in terms of value purchased in most countries, including Argentina, Austria, Brazil, Canada, Denmark, France, Germany, Ireland, Italy, Mexico, Norway, Poland, Portugal, Spain, Sweden, Switzerland, UK, and USA.

Dairy Products are favored in Europe: Dairy Products are a top category in many European countries, including Austria, Belgium, Denmark, Finland, France, Germany, Ireland, Italy, Norway, Poland, Portugal, Spain, Sweden, and Switzerland.

Confections are popular in the Americas: Confections are a top category in several countries in the Americas, including Argentina, Brazil, Canada, Mexico, USA, and Venezuela.

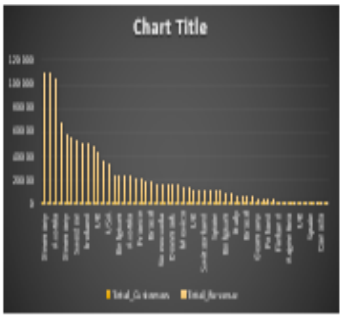
Seafood is popular in coastal countries: Seafood is a top category in countries with a significant coastline, such as Argentina, Brazil, Canada, Denmark, France, Germany, Ireland, Norway, Portugal, Spain, Sweden, and USA.

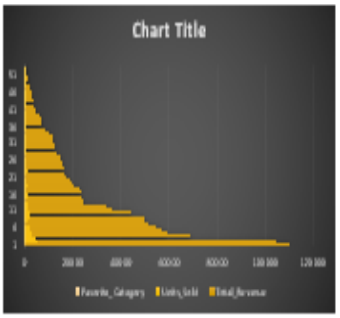
Meat/Poultry and Grains/Cereals: show significant value in some countries, such as Spain and Portugal.

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

```
WITH customer_purchases AS (
  SELECT
    ca.Country,
    ca.ContactName AS Customer_Name,
    SUM(od.UnitPrice * od.Quantity * (1 - Discount)) AS Purchased_Value,
    SUM(od.Quantity) AS Units_Purchased,
    p.CategoryID,
    COUNT(ca.ContactName) OVER (PARTITION BY ca.Country) AS Total_Customers
  FROM
    orders o
  JOIN
    order_details od ON od.OrderID = o.OrderID
  JOIN
    products p ON p.ProductID = od.ProductID
  JOIN
    customers ca ON ca.CustomerID = o.CustomerID
  JOIN
    categories cg ON cg.CategoryID = p.CategoryID
),
favorite_categories AS (
  SELECT
    Country,
    Customer_Name,
    CategoryID,
    SUM(Units_Purchased) AS Total_Units
  FROM
    customer_purchases
  GROUP BY
    Country,
    Customer_Name,
    CategoryID
),
customer_favorites AS (
  SELECT
    fc.Country,
    fc.Customer_Name,
    cg.CategoryName AS Favorite_Category,
    ROW_NUMBER() OVER (PARTITION BY fc.Country, fc.Customer_Name ORDER BY fc.Total_Units DESC) AS rn
  FROM
    favorite_categories fc
  JOIN
    categories cg ON cg.CategoryID = fc.CategoryID
)
SELECT
  cp.Country,
  cp.Total_Customers,
  SUM(cp.Purchased_Value) AS Total_Revenue,
  SUM(cp.Units_Purchased) AS Units_Sold,
  cf.Favorite_Category
FROM
  customer_purchases cp
JOIN
  (SELECT Country, Customer_Name, Favorite_Category FROM customer_favorites WHERE rn = 1) cf
ON
  cp.Country = cf.Country AND cp.Customer_Name = cf.Customer_Name
GROUP BY
  cp.Country,
  cp.Total_Customers,
  cf.Favorite_Category
ORDER BY
  Total_Revenue DESC, cp.Country asc ;
```

Country	Total_Customers	Total_Revenue	Units_Sold	Favorite_Category
Germany	328	110277.305	2661	Beverages
USA	352	109729.39	5159	Seafood
Austria	125	104874.9785	4543	Dairy Products
Brazil	203	68734.6175	2494	Beverages
Germany	328	59269.6045	2671	Dairy Products
Germany	328	56464.724	2107	Confections
Sweden	97	54495.14	2235	Beverages
USA	352	51097.8005	1383	Dairy Products
Ireland	55	49979.905	1684	Seafood
Canada	75	49673.79	1922	Confections
UK	135	44367.01	2037	Dairy Products
USA	352	36554.105	1332	Beverages
USA	352	33684.9125	948	Confections
Venezuela	116	31257.464	1130	Beverages
Belgium	56	31088.78	1072	Beverages
France	184	30579.1525	1060	Seafood
Austria	125	23128.66	624	Beverages
France	184	22437.43	812	Dairy Products
France	184	20502.68	959	Beverages
Switzerlar	52	19043.779	810	Dairy Products
Brazil	203	18669.685	842	Condiments
Denmark	46	16817.0975	378	Confections
Venezuela	116	16476.565	970	Confections
Venezuela	116	16076.6	836	Dairy Products
Denmark	46	15843.925	792	Seafood
Finland	54	15648.7025	737	Dairy Products
Mexico	72	15054.35	592	Beverages
France	184	14639.06	423	Confections
UK	135	12665.2	618	Beverages
Brazil	203	12450.8	660	Confections
Switzerlar	52	12348.88	465	Grains/Cereals
Portugal	30	11472.3625	533	Condiments
Spain	54	11446.36	395	Confections
USA	352	11441.63	327	Meat/Poultry
Belgium	56	9736.075	320	Confections
Mexico	72	8426.9275	422	Dairy Products
Italy	53	7176.215	433	Beverages
Italy	53	7048.24	305	Dairy Products
Brazil	203	6850.664	251	Dairy Products
Norway	16	5735.15	161	Beverages
Germany	328	4273	174	Condiments
Spain	54	4232.85	190	Meat/Poultry
Poland	16	3531.95	205	Beverages
Argentina	34	3460.2	132	Dairy Products
Finland	54	3161.35	149	Confections
USA	352	3076.4725	181	Condiments
Argentina	34	2644.1	92	Confections
Argentina	34	1614.8	115	Beverages
UK	135	1719.1	87	Grains/Cereals
Italy	53	1545.7	54	Produce
Spain	54	1467.29	91	Seafood
Spain	54	836.7	42	Grains/Cereals
Canada	75	522.5	62	Seafood
Mexico	72	100.8	11	Confections





CONCLUSION:-

High Revenue Generating Countries and Favorite Categories:

Germany: Generates the highest revenue with 110277.305 in Beverages. Also notable for Dairy Products and Confections.
USA: Close second in revenue with 109729.39 in Seafood. Has significant contributions from Dairy Products, Beverages, and Confections.
Austria: High revenue with 104874.9785, predominantly in Dairy Products.
Brazil: 68734.6175 in Beverages but also contributes to Condiments.

Diverse Category Preferences:

Germany: Not limited to one category, significant revenue from Beverages, Dairy Products, and Confections.
USA: Similarly diverse, with high contributions from Seafood, Dairy Products, Beverages, Confections, and Meat/Poultry.
France: Revenue spread across Seafood, Dairy Products, Beverages, and Confections.

Specialized Markets:

Sweden: High revenue in Beverages with 54495.14.
Ireland: Significant in Seafood with 49979.905.
Venezuela: Predominantly Beverages and Confections.
Finland: Focused on Dairy Products.

Smaller Markets with Unique Preferences:

Portugal: Significant revenue in Condiments.
Norway: Focus on Beverages.
Argentina: Revenue from Dairy Products and Confections.

Low Revenue Markets:

Poland, Argentina, Finland, Canada, and Mexico: Lower revenue figures with specific category preferences.

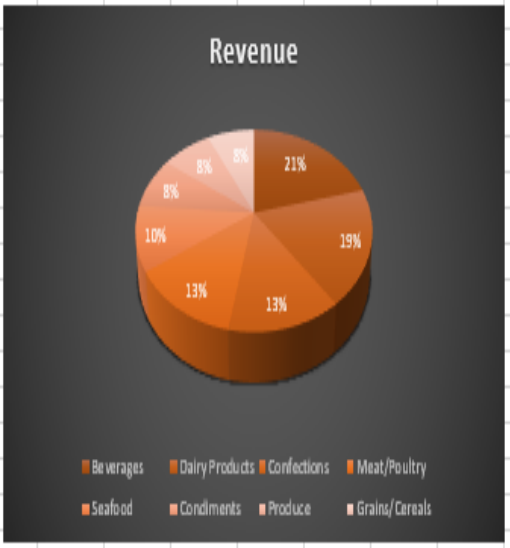
Potential Market Segments:

High Revenue, Diverse Preferences: Germany, USA, France
High Revenue, Specialized Preferences: Austria (Dairy Products), Brazil (Beverages)
Moderate Revenue, Specialized Preferences: Sweden (Beverages), Ireland (Seafood), Venezuela (Beverages, Confections)
Small Markets, Unique Preferences: Portugal (Condiments), Norway (Beverages)
Low Revenue Markets: Poland, Argentina, Finland, Canada, Mexico

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

```
WITH order_prices AS (
    SELECT
        cg.CategoryName AS Product_Category,
        (od.UnitPrice * od.Quantity * (1 - Discount)) AS
Order_Price
    FROM
        orders o
    JOIN
        order_details od ON o.OrderID = od.OrderID
    JOIN
        products p ON p.ProductID = od.ProductID
    JOIN
        categories cg ON cg.CategoryID = p.CategoryID
),
category_revenue AS (
    SELECT
        Product_Category,
        SUM(Order_Price) AS Revenue
    FROM
        order_prices
    GROUP BY
        Product_Category
),
total_revenue AS (
    SELECT
        SUM(Revenue) AS Total_Revenue
    FROM
        category_revenue
)
SELECT
    cr.Product_Category,
    cr.Revenue,
    ((cr.Revenue * 100) / tr.Total_Revenue) AS
Revenue_in_per
FROM
    category_revenue cr
CROSS JOIN
    total_revenue tr
ORDER BY
    cr.Revenue DESC;
```

Product_Categ	Revenu	Revenue_in_
Beverages	267868	21.1620835
Dairy Products	234507	18.52651087
Confections	167357	13.22153147
Meat/Poultry	163022	12.87906904
Seafood	131262	10.36992094
Condiments	106047	8.377916586
Produce	99984.6	7.898987436
Grains/Cereals	95744.6	7.564000157



CONCLUSION:-

Key findings:

Top 3 categories: Beverages, Dairy Products, and Confections are the top 3 categories that contribute the most to order revenue, accounting for approximately 52.9% of the total revenue.

1. Beverages lead with the highest revenue, totaling $267,868.18, which constitutes about 21.16% of the total revenue.
2. Dairy Products follow with $234,507.29, making up 18.53% of the total revenue.
3. Meat/Poultry and Seafood are also significant contributors to order revenue, accounting for 12.9% and 10.4% of the total revenue, respectively.

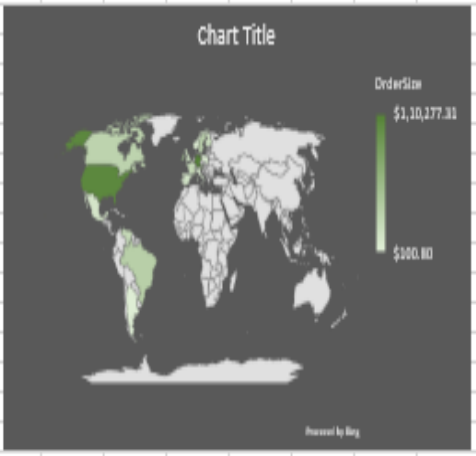
Business insights:

1. Focus on top categories: The company should focus on the top 3 categories (Beverages, Dairy Products, and Confections) to maximize revenue growth.
2. Optimize product offerings: The company should optimize its product offerings in the top categories to meet customer demand and increase revenue.
3. Monitor and adjust: The company should continuously monitor the revenue contribution of each product category and adjust its strategies accordingly to ensure maximum revenue growth.

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

```
SELECT
    C.CustomerID,
    C.CompanyName,
    C.Country,
    SUM(OD.UnitPrice * OD.Quantity * (1 - Discount)) AS OrderSize
FROM
    customers C
JOIN
    orders O ON C.CustomerID = O.CustomerID
JOIN
    order_details OD ON O.OrderID = OD.OrderID
GROUP BY
    C.CustomerID, C.CompanyName, C.City, C.Country
ORDER BY
    OrderSize DESC;
```

Custo	Comp	Coun	OrderSize
QUICK	QUICK-S	Germany	$ 1,10,277.31
ERNSH	Ernst Han	Austria	$ 1,04,874.98
SAVEA	Save-a-lo	USA	$ 1,04,361.95
RATTC	Rattlesnal	USA	$ 51,097.80
HUNGO	Hungry C	Ireland	$ 49,979.91
HANAR	Hanari Ca	Brazil	$ 32,841.37
KOENE	KÃƒÂ¶nig	Germany	$ 30,908.38
FOLKO	Folk och f	Sweden	$ 29,567.56
MEREP	MÃƒÂ¨re	Canada	$ 28,872.19
WHITC	White Clo	USA	$ 27,363.61
FRANK	Frankenv	Germany	$ 26,656.56
QUEEN	Queen Co	Brazil	$ 25,717.50
BERGS	Berglunds	Sweden	$ 24,927.58
SUPRD	SuprÃƒÂ	Belgium	$ 24,088.78
PICCO	Piccolo u	Austria	$ 23,128.66
HILAA	HILARIÃ	Venezuel	$ 22,768.76
BONAP	Bon app'	France	$ 21,963.25
BOTTM	Bottom-D	Canada	$ 20,801.60
RICSU	Richter S	Switzerla	$ 19,343.78
LEHMS	Lehmann	Germany	$ 19,261.41
BLONP	Blondel p	France	$ 18,534.08
GREAL	Great Lak	USA	$ 18,507.45
SIMOB	Simons b	Denmark	$ 16,817.10
LINOD	LINO-Del	Venezuel	$ 16,476.57
SEVES	Seven Se	UK	$ 16,215.33
LILAS	LILA-Sup	Venezuel	$ 16,076.60
VAFFE	Vaffeljern	Denmark	$ 15,843.93
WARTH	Wartian H	Finland	$ 15,648.70
OLDWO	Old Worl	USA	$ 15,177.46
EASTC	Eastern C	UK	$ 14,761.04
AROUT	Around th	UK	$ 13,390.65
OTTIK	Ottilies K	Germany	$ 12,496.20
RICAR	Ricardo A	Brazil	$ 12,450.80
CHOPS	Chop-suc	Switzerla	$ 12,348.88
FOLIG	Folies go	France	$ 11,868.90
GODOS	Godos Co	Spain	$ 11,446.36
SPLIR	Split Rail	USA	$ 11,441.63
TORTU	Tortuga R	Mexico	$ 10,812.15
MAISD	Maison D	Belgium	$ 9,736.07
WANDK	Die Wand	Germany	$ 9,588.43
LAMAI	La maison	France	$ 9,328.20
VICTE	Victuailles	France	$ 9,182.43
GOURL	Gourmet	Brazil	$ 8,414.14
MAGAA	Magazzin	Italy	$ 7,178.22
REGGC	Reggiani	Italy	$ 7,048.24
ANTON	Antonio M	Mexico	$ 7,023.98
TRADH	TradiÃƒÂ	Brazil	$ 6,850.66
QUEDE	Que DelÃ	Brazil	$ 6,664.81
FURIB	Furia Bac	Portugal	$ 6,427.42
ISLAT	Island Tra	UK	$ 6,146.30
BSBEV	B's Bever	UK	$ 6,089.90
WELLI	Wellingto	Brazil	$ 6,068.20
SANTG	SantÃƒÂ	Norway	$ 5,735.15
PRINI	Princesa	Portugal	$ 5,044.94
MORGK	Morgenst	Germany	$ 5,042.20
TOMSP	Toms Spe	Germany	$ 4,778.14



CONCLUSION:-

Based on the return data we have to conclude that

Key findings:

1. Order size distribution: The order size distribution is skewed, with a few customers having very large order sizes (above $100,000) and many customers having smaller order sizes (below $10,000).
2. Country-wise order size: Customers from Germany, USA, and Austria have the largest average order sizes, while customers from Argentina, Mexico, and Venezuela have smaller average order sizes.
3. Company-wise order size: Some companies, such as QUICK-Stop, Ernst Handel, and Save-a-lot Markets, have very large order sizes, while others, such as Lazy K Kountry Store and Centro comercial Moctezuma, have very small order sizes.

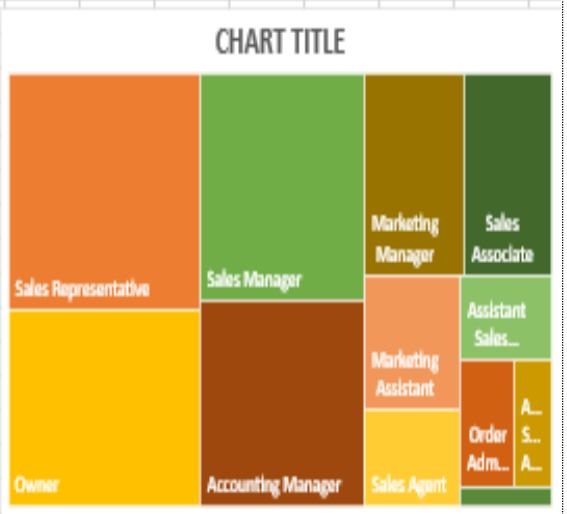
Business insights:

1. Target high-value customers: The company should focus on targeting customers with large order sizes, such as those from Germany, USA, and Austria, to maximize revenue growth.
2. Optimize product offerings: The company should optimize its product offerings to meet the needs of customers with large order sizes, such as
QUICK-Stop and Ernst Handel.
3. Improve customer relationships: The company should improve its customer relationships with companies that have smaller order sizes, such as Lazy K Kountry Store and Centro comercial Moctezuma, to increase revenue growth.

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

```
SELECT
    cs.ContactTitle AS Customer_Segment,
    COUNT(od.OrderID) AS Order_Frequency
FROM
    order_details od
JOIN
    orders o ON o.OrderID = od.OrderID
JOIN
    customers cs ON cs.CustomerID = o.CustomerID
GROUP BY
    cs.ContactTitle
    order by Order_Frequency desc;
```

Customer_Segme	Order_Frequen
Sales Representative	414
Owner	347
Sales Manager	342
Accounting Manager	311
Marketing Manager	185
Sales Associate	160
Marketing Assistant	118
Sales Agent	85
Assistant Sales Repres	71
Order Administrator	62
Assistant Sales Agent	43
Owner/Marketing Assi	17



CONCLUSION:-

Key findings:

1. Order frequency distribution: The order frequency is highest for Sales Representatives, followed by Owners, Sales Managers, and Accounting Managers.
2. Customer segment hierarchy: The customer segments can be grouped into three categories: Sales-related (Sales Representative, Sales Manager, Sales Associate, Sales Agent, Assistant Sales Representative, Assistant Sales Agent), Management-related (Owner, Accounting Manager), and Marketing-related (Marketing Manager, Marketing Assistant, Owner/Marketing Assistant).
3. Order frequency variation: The order frequency varies significantly across different customer segments, with Sales Representatives having more than twice the order frequency of Owners, and Owners having more than twice the order frequency of Marketing Managers.

Business insights:

1. Target high-frequency segments: The company should focus on targeting customer segments with high order frequencies, such as Sales Representatives and Owners, to maximize revenue growth.
2. Optimize marketing strategies: The company should optimize its marketing strategies to target customer segments with lower order frequencies, such as Marketing Managers and Marketing Assistants, to increase revenue growth.
3. Improve customer relationships: The company should improve its customer relationships with customer segments that have lower order frequencies, such as Order Administrators and Assistant Sales Agents, to increase revenue growth.

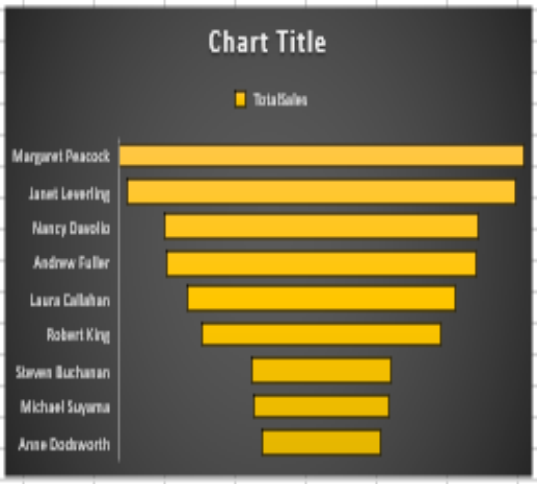
Are there any correlations between employee satisfaction levels and key performance indicators? Can we explore this visually through scatter plots or line charts?

We don't have data for employee satisfaction. For this question we need employees satisfaction review or feedback report then we can slove this.

For unavailable data we could't slove the question that's why I create a question by own that is. Que- How does order frequency vary across different employee

```
SELECT
  e.EmployeeID,
  CONCAT(e.FirstName, ' ', e.LastName) AS
EmployeeName,
  SUM(o.Freight) AS TotalSales
FROM
  orders o
  JOIN employees e ON o.EmployeeID = e.EmployeeID
GROUP BY
  e.EmployeeID, CONCAT(e.FirstName, ' ', e.LastName)
ORDER BY
  TotalSales DESC;
```

EmployeeI	EmployeeNam	TotalSales
4	Margaret Peacock	$ 11,346.14
3	Janet Leverling	$ 10,884.74
1	Nancy Davolio	$ 8,836.64
2	Andrew Fuller	$ 8,696.41
8	Laura Callahan	$ 7,487.88
7	Robert King	$ 6,665.44
5	Steven Buchanan	$ 3,918.71
6	Michael Suyama	$ 3,780.47
9	Anne Dodsworth	$ 3,326.26



Conclusion:
The analysis of order frequency across different employees reveals significant variations in their sales performance, highlighting opportunities for improvement and growth.

Key Findings:

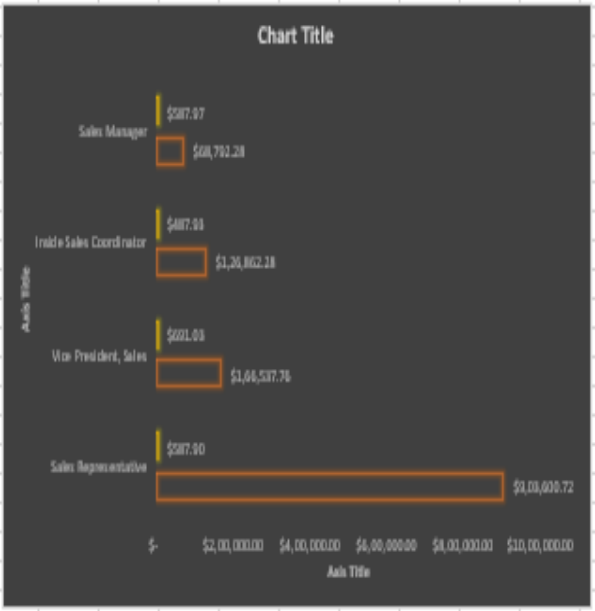
Top Performing Employees: The results show that certain employees generate significantly higher total sales than others.
Employee Performance Variance: The data indicates substantial differences in sales performance among employees, suggesting differences in their sales strategies, customer engagement, or product knowledge.
Insights:

1. Identify top-performing employees and recognize their achievements.
2. Analyze performance gaps between employees to target training or coaching for improvement.
3. Develop strategies to enhance sales performance and customer engagement

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

```
SELECT
    a.Title as Job Role,
    Sum(ap.UnitPrice*ap.Quantity*(1-ap.Discount)) AS Turnover,
    avg(ap.UnitPrice*ap.Quantity*(1-ap.Discount)) AS
Average Turnover
  FROM
    order details ap
    JOIN orders p ON p.OrderID = ap.OrderID
    JOIN employees a ON a.EmployeeID=p.EmployeeID
  group by
     Job Role
  order by
  Turnover desc
```

Job Role	Turnover	Average Turnover
Sales Representative	#########	$ 587.90
Vice President, Sales	#########	$ 691.03
Inside Sales Coordinat	#########	$ 487.93
Sales Manager	$ 68,792.28	$ 587.97



CONCLUSION:-

Key findings:

Employee Turnover Distribution: The total turnover is highest for Sales Representatives, followed by Vice President, Sales, Inside Sales Coordinators, and Sales Managers.

Job Role Hierarchy: The job roles can be grouped into two categories: Sales-related (Sales Representative, Vice President, Sales, Inside Sales Coordinator, Sales Manager) and Management-related (Sales Manager).

Turnover Variation: The average turnover per employee varies significantly across different job roles, with Vice President, Sales having the highest average turnover, and Inside Sales Coordinators having the lowest average turnover.

Business insights:

Focus on High-Turnover Job Roles: The company should focus on supporting and retaining employees in high-turnover job roles, such as Sales Representatives and Vice President, Sales, to maximize revenue growth.

Optimize Sales Strategies: The company should optimize its sales strategies to target job roles with lower average turnovers, such as Inside Sales Coordinators and Sales Managers, to increase revenue growth.

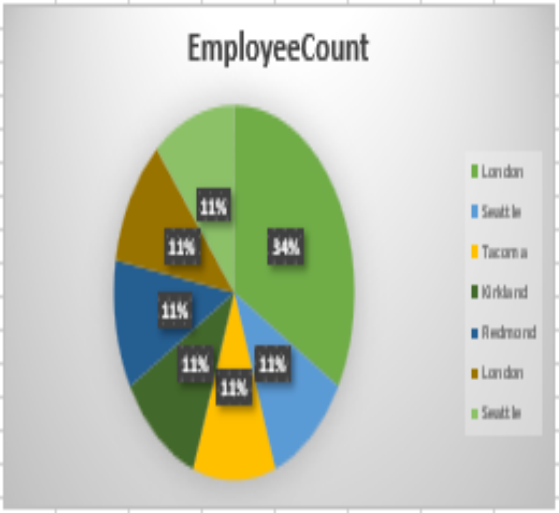
Improve Employee Performance: The company should improve employee performance and provide training and development opportunities to employees in job roles with lower average turnovers to increase revenue growth.

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

We don't have data for this question. If we want to solve this question so we need employee skill sets or qualifications data

So we create a question by own

Can we identify any patterns or clusters in employee locations (cities or regions) through visualizations? How can this information be used for resource allocation or team management?




```
SELECT
Title as Job_Rol,
Country,
City,
COUNT(*) as EmployeeCount
FROM employees
GROUP BY
  Title,
  Country,
  City
ORDER BY EmployeeCount DESC;
```

Job_R	Count	City	EmployeeCou
Sales Rep	UK	London	3
Sales Rep	USA	Seattle	1
Vice Presi	USA	Tacoma	1
Sales Rep	USA	Kirkland	1
Sales Rep	USA	Redmond	1
Sales Mar	UK	London	1
Inside Sale	USA	Seattle	1

CONCLUSION:

Key Findings:

Job Role Distribution: Sales Representatives are the most common job role, accounting for 71.4% of employees.
Country-wise Distribution: The USA has the highest number of employees, followed by the UK.
City-wise Distribution: London is the city with the highest number of employees, followed by Seattle.
Job Role Concentration: Sales Representatives are concentrated in London and the USA.

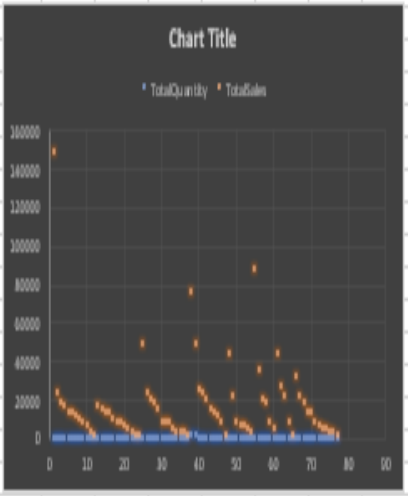
Business Insights:

Focus on Key Locations: The company should focus on London and Seattle, which are the key locations with the highest number of employees.
Develop Job Role-Specific Strategies: The company should develop job role-specific strategies to target high-performing job roles, such as Sales Representatives, and improve employee performance within each role.
Optimize Resource Allocation: The company should optimize resource allocation to ensure that the right resources are allocated to the right job roles and locations to maximize efficiency and productivity.
Consider Decentralization: The company should consider decentralizing certain job roles, such as Sales Representatives, to improve coverage and reach in different locations.

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?

```
SELECT
    c.CategoryName,
    p.ProductName,
    SUM(od.Quantity) AS TotalQuantity,
    SUM(od.UnitPrice * od.Quantity) AS TotalSales
FROM
    order details od
    INNER JOIN products p ON od.ProductID = p.ProductID
    INNER JOIN categories c ON p.CategoryID = c.CategoryID
GROUP BY
    c.CategoryName,
    p.ProductName
ORDER BY
    c.CategoryName,
    TotalSales DESC;
```

CategoryNam	ProductNam	TotalQuantity	TotalSale
Beverages	CÃƒÂ´te de Bla	623	149984.2
Beverages	Ipoh Coffee	580	25079.2
Beverages	Chang	1057	18559.2
Beverages	LakkalikÃƒÂ¶	981	16794
Beverages	Steeleye Stout	883	14536.8
Beverages	Chai	828	14277.6
Beverages	Chartreuse vert	793	13150.8
Beverages	Outback Lager	817	11472
Beverages	RhÃƒÂ¶nbrÃƒÂ	1155	8650.55
Beverages	Sasquatch Ale	506	6678
Beverages	GuaranÃƒÂ¡ Fa	1125	4782.6
Beverages	Laughing Lumb	184	2562
Condiments	Vegie-spread	445	17696.3
Condiments	Sirop d'Ã‚rabi	603	16438.8
Condiments	Louisiana Fiery	745	14607
Condiments	Northwoods Cra	372	13760
Condiments	Gula Malacca	601	10524.2
Condiments	Original Frankf	791	9685
Condiments	Chef Anton's Ca	453	9424.8
Condiments	Grandma's Boys	301	7345
Condiments	Chef Anton's Gu	298	5801.15
Condiments	Louisiana Hot S	239	3519
Condiments	Aniseed Syrup	328	3080
Condiments	Genen Shouyu	122	1813.5
Confections	Tarte au sucre	1083	49827.9
Confections	Sir Rodney's M	313	23635.8
Confections	GumbÃƒÂ¤r Gu	753	21534.9
Confections	Pavlova	1158	18748.05
Confections	Schoggi Schok	365	15231.5
Confections	Sir Rodney's Sc	1016	9636
Confections	Maxilaku	520	9500
Confections	Scottish Longbr	799	9362.5
Confections	Teatime Chocol	723	6159.5
Confections	Zaanse koeken	485	4358.6
Confections	NuNuCa NuÃƒÂ	318	4051.6
Confections	Valkoinen sukl	235	3510
Confections	Chocolade	138	1542.75
Dairy Products	Raclette Courd	1496	76296
Dairy Products	Camembert Pie	1577	50286
Dairy Products	Mozzarella di G	806	25738.8
Dairy Products	Gudbrandsdals	714	24307.2
Dairy Products	FÃƒÂ¶rsyso	1057	20876.5
Dairy Products	Gorgonzola Tel	1397	16172.5
Dairy Products	Queso Cabralo	706	13902
Dairy Products	Queso Manche	344	12866.8
Dairy Products	Mascarpone Fa	297	9171.2
Dairy Products	Geitost	755	1713.5



CONCLUSION:-

Key findings:

Product Category Performance: Beverages, Confections, and Dairy Products are the top-performing categories in terms of total sales, with Beverages having the highest total sales.

Product Attribute Correlations: There is a correlation between product attributes such as size, color, and features and sales performance, with certain product categories and attributes performing better than others.

Category-Specific Insights: Within each category, there are specific products that perform better than others, such as Côte de Blaye in Beverages and Raclette Courdavault in Dairy Products.

Business insights:

Focus on High-Performing Categories: The company should focus on the top-performing categories, such as Beverages, Confections, and Dairy Products, to maximize revenue growth.

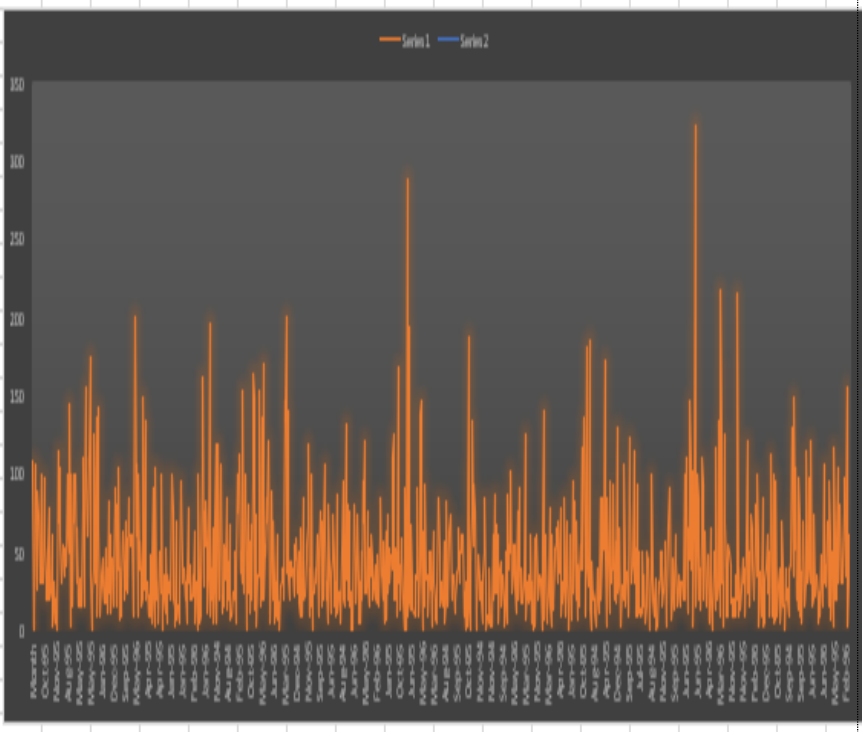
Optimize Product Attributes: The company should optimize product attributes such as size, color, and features to improve sales performance, particularly in categories with lower sales.

Category-Specific Strategies: The company should develop category-specific strategies to target high-performing products and improve sales performance within each category.

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

```
SELECT
  p.ProductName,
  DATE_FORMAT(o.OrderDate, '%m-%Y') AS Month,
  SUM(od.Quantity * od.UnitPrice * (1 - od.Discount)) AS TotalSales,
  SUM(od.Quantity) AS TotalQuantity
FROM
  orders o
  INNER JOIN order_details od ON o.OrderID = od.OrderID
  INNER JOIN products p ON od.ProductID = p.ProductID
GROUP BY
  p.ProductName, DATE_FORMAT(o.OrderDate, '%m-%Y')
ORDER BY
  p.ProductName, Month;
```

ProductName	Month	Total Sale	TotalQuantity
Alice Mutton	Jan-96	4212	108
Alice Mutton	Feb-95	62.4	2
Alice Mutton	Feb-96	234	6
Alice Mutton	Mar-95	3276	105
Alice Mutton	Mar-96	3705	95
Alice Mutton	Apr-96	1053	27
Alice Mutton	May-96	3471	89
Alice Mutton	Jun-95	2847	73
Alice Mutton	Jul-95	1326	34
Alice Mutton	Aug-94	936	30
Alice Mutton	Aug-95	3900	100
Alice Mutton	Sep-94	936	30
Alice Mutton	Oct-94	1248	40
Alice Mutton	Oct-95	1170	30
Alice Mutton	Nov-94	3057.6	98
Alice Mutton	Nov-95	2145	55
Alice Mutton	Dec-94	1123.2	36
Alice Mutton	Dec-95	780	20
Aniseed Syrup	Jan-96	200	20
Aniseed Syrup	Feb-95	400	50
Aniseed Syrup	Feb-96	790	79
Aniseed Syrup	Apr-95	160	20
Aniseed Syrup	May-96	250	25
Aniseed Syrup	Jun-95	600	60
Aniseed Syrup	Jun-96	40	4
Aniseed Syrup	Aug-95	140	14
Aniseed Syrup	Sep-94	240	30
Aniseed Syrup	Nov-95	60	6
Aniseed Syrup	Dec-95	200	20
Boston Crab Me	Jan-96	18.4	1
Boston Crab Me	Feb-95	29.4	2
Boston Crab Me	Feb-96	2116	115
Boston Crab Me	Mar-95	1029	70
Boston Crab Me	Mar-96	1895.2	103
Boston Crab Me	Apr-95	896.7	61
Boston Crab Me	Apr-96	552	30
Boston Crab Me	May-95	552	30
Boston Crab Me	May-96	1012	55
Boston Crab Me	Jun-95	644	35
Boston Crab Me	Jul-95	956.8	52
Boston Crab Me	Aug-94	735	50
Boston Crab Me	Aug-95	772.8	42
Boston Crab Me	Sep-94	1470	100
Boston Crab Me	Sep-95	1104	60
Boston Crab Me	Oct-94	735	50
Boston Crab Me	Oct-95	2649.6	144



CONCLUSION:-

Key findings:

Product Demand Fluctuation: Product demand fluctuates significantly over different months, with some products experiencing high demand in certain months and low demand in others.
Seasonal Trends: There are seasonal trends in product demand, with some products experiencing higher demand in the winter months (e.g., Chai, Glühwein) and others in the summer months (e.g., Sasquatch Ale, Spegesild).
Product-Specific Insights: Certain products, such as Alice Mutton and Côte de Blaye, have consistent demand throughout the year, while others, such as Raclette Courdavault and Manjimup Dried Apples, have spikes in demand during specific months.

Business insights:

Stock Management: The company should manage its inventory and stock levels based on the fluctuation in product demand, ensuring that high-demand products are always in stock and low-demand products are not overstocked.
Seasonal Promotions: The company should consider running seasonal promotions and marketing campaigns to capitalize on the increased demand for certain products during specific months.
Product-Specific Strategies: The company should develop product-specific strategies to target high-demand products and improve sales performance, such as offering discounts or bundling high-demand products with low-demand products.

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

```
WITH ProductSales AS (
  SELECT
    p.ProductID,
    p.ProductName,
    SUM(od.Quantity * od.UnitPrice) AS TotalSales
  FROM
    order_details od
    INNER JOIN products p ON od.ProductID =
p.ProductID
  GROUP BY
    p.ProductID, p.ProductName
),
ProductStats AS (
  SELECT
    AVG(TotalSales) AS AvgSales,
    STDDEV(TotalSales) AS StdDevSales
  FROM
    ProductSales
)
SELECT
  ps.ProductID,
  ps.ProductName,
  ps.TotalSales
FROM
  ProductSales ps
  CROSS JOIN ProductStats ps2
WHERE
  ps.TotalSales > ps2.AvgSales + 2 * ps2.StdDevSales
  OR ps.TotalSales < ps2.AvgSales - 2 *
ps2.StdDevSales;
```

ProductID	ProductName	TotalSales
59	Raclette Courdavault	76296
29	ThÃƒÃ¼ringer Rostbratwur	87736.4
38	CÃƒÃ´te de Blaye	149984.2

CONCLUSION:

Key Findings:

Highest Total Sales: Côte de Blaye, with a total sales of $149,984.20 Highest Outperforming Product: Côte de Blaye, with sales $73,247.80 above the average Lowest Outperforming Product: Raclette Courdavault, with sales $23,688.80 below the average

Ranking of Products by Total Sales:

Côte de Blaye: $149,984.20 Thüringer Rostbratwurst: $87,736.40 Raclette Courdavault: $76,296.00

Insights:

Côte de Blaye is the top-performing product in terms of sales. Raclette Courdavault is the lowest-performing product, indicating potential areas for improvement. This data can inform product optimization strategies, identify high-performing products, and potentially address any performance gaps in other products.

By analyzing these results, you can:

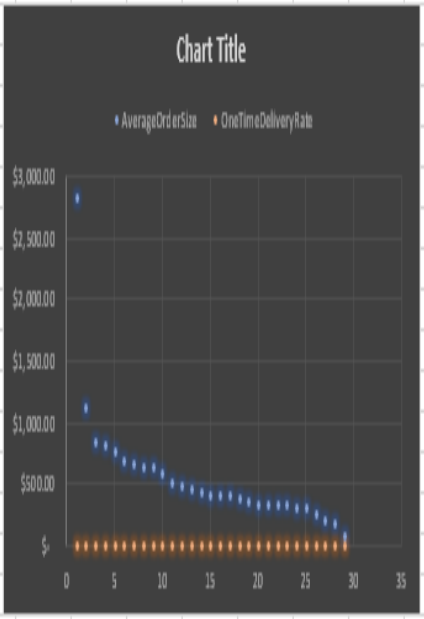
Identify opportunities to improve the performance of underperforming products, such as Raclette Courdavault.
Develop targeted marketing strategies to promote high-performing products, such as Côte de Blaye.

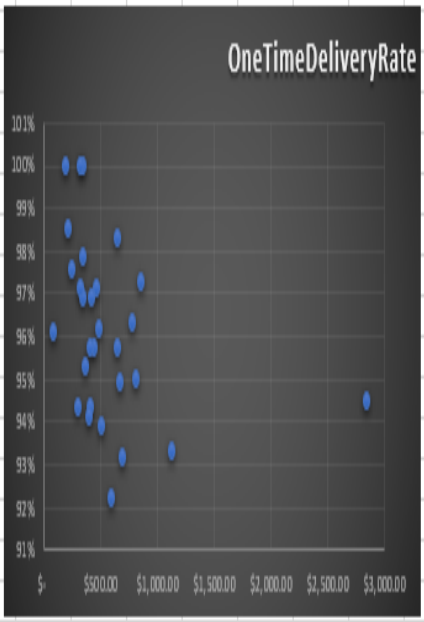
Are there any correlations between supplier attributes (e.g., location, size, industry) and performance metrics (e.g., on-time delivery,
product quality)? Can we explore this visually through scatter plots or heatmaps?

I don't get all the data for this question. So I did the answer by location as cuntry, size by order size and On-time delivery

```
SELECT
  s.CompanyName as Supplier_Name,
  s.Country,
  AVG(od.Quantity * od.UnitPrice * (1 - od.Discount)) AS
AverageOrderSize,
  SUM(CASE WHEN o.ShippedDate <= o.RequiredDate THEN 1
ELSE 0 END) / COUNT(o.OrderID) AS OneTimeDeliveryRate
FROM
  suppliers s
  JOIN products p ON s.SupplierID = p.SupplierID
  JOIN order_details od ON p.ProductID = od.ProductID
  JOIN orders o ON od.OrderID = o.OrderID
GROUP BY
  Supplier_Name,
  s.Country
ORDER BY
  AverageOrderSize DESC,
  OneTimeDeliveryRate DESC;
```

Supplier_Name	Country	AverageOrderS	OneTimeDeliveryR
Aux joyeux eccl**Â**ƒÂ©siastiques	France	$ 2,846.13	94%
Gai pÂƒÂ¢turage	France	$ 1,123.83	93%
For**Â**ƒÂªts d'Â**Â**ƒÂ©rables	Canada	$ 855.38	97%
Plutzer Lebensmittelgro**Â**ƒÂ mÂƒÂ¤rkte	Germany	$ 812.14	95%
Grandma Kelly's Homestead	USA	$ 776.91	96%
Pasta Buttini s.r.l.	Italy	$ 688.42	93%
G'day, Mate	Australia	$ 669.66	95%
Heli S**Â**ƒÂ¤**Â**ƒÂ waren GmbH & Co. K	Germany	$ 655.14	98%
Pavlova, Ltd.	Australia	$ 653.13	96%
Tokyo Traders	Japan	$ 598.56	92%
Leka Trading	Singapore	$ 512.41	94%
Cooperativa de Quesos 'Las Cabras'	Spain	$ 483.84	96%
Formaggi Fortini s.r.l.	Italy	$ 463.70	97%
New Orleans Cajun Delights	USA	$ 445.26	96%
Nord-Ost-Fisch Handelsgesellschaft mb	Germany	$ 419.51	97%
Norske Meierier	Norway	$ 410.87	94%
Karkki Oy	Finland	$ 408.32	96%
Svensk Sj**Â**ƒ**Â**ƒ**Â**ƒ**Â**¶da AB	Sweden	$ 394.98	94%
Specialty Biscuits, Ltd.	UK	$ 367.02	95%
PB Kn**Â**ƒÂ¤ckebr**Â**ƒ**Â**¶d AB	Sweden	$ 344.83	100%
Bigfoot Breweries	USA	$ 344.48	97%
Exotic Liquids	UK	$ 342.43	98%
Escargots Nouveaux	France	$ 328.76	100%
Ma Maison	Canada	$ 321.08	97%
New England Seafood Cannery	USA	$ 302.17	94%
Lyngbysild	Denmark	$ 249.30	98%
Mayumi's	Japan	$ 216.72	99%
Zaanse Snoepfabriek	Netherland	$ 197.29	100%
Refrescos Americanas LTDA	Brazil	$ 88.32	96%

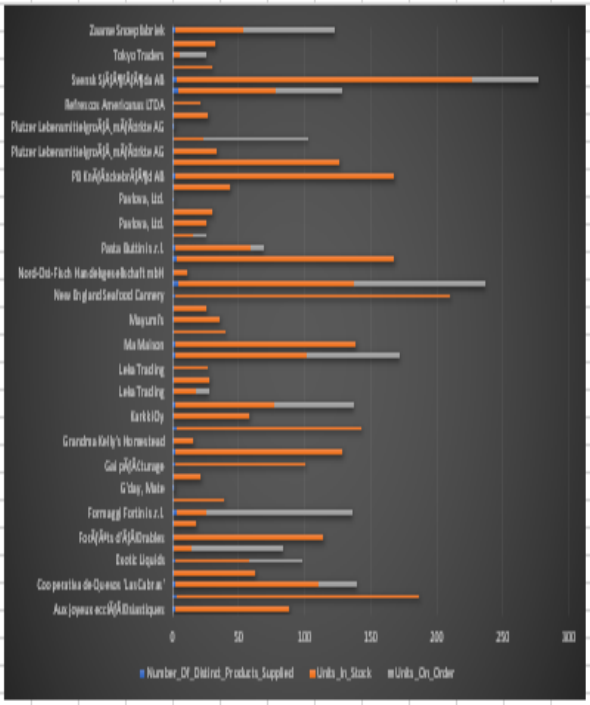




How does supplier performance vary across different product categories or departments? Can we visualize this using stacked bar charts or grouped column charts?

```
SELECT
  s.CompanyName AS Supplier Name,
  c.CategoryName AS Product Category,
  COUNT(DISTINCT p.ProductID) AS
Number Of Distinct Products Supplied,
  SUM(p.UnitsInStock) AS Units In Stock,
  SUM(p.UnitsOnOrder) AS Units On Order
FROM
  suppliers s
  JOIN products p ON s.SupplierID = p.SupplierID
  JOIN categories c ON p.CategoryID = c.CategoryID
GROUP BY
  Supplier Name, Product Category
ORDER BY
  Supplier Name ASC
```

Supplier_Nam	Product_Catego	Number_Of_Distinct_Products_Supp	Units_In_Stoc	Units_On_Ord
Aux joyeux eccl	Beverages	2	86	0
Bigfoot Brewerie	Beverages	3	183	0
Cooperativa de C	Dairy Products	2	108	30
Escargots Nouve	Seafood	1	62	0
Exotic Liquids	Beverages	2	56	40
Exotic Liquids	Condiments	1	13	70
ForÃƒÂªts dÃƒÂ	Condiments	1	113	0
ForÃƒÂªts dÃƒÂ	Confections	1	17	0
Formaggi Fortini	Dairy Products	3	23	110
Gday, Mate	Grains/Cereals	1	38	0
Gday, Mate	Meat/Poultry	1	0	0
Gday, Mate	Produce	1	20	0
Gai pÃƒÂ¢turage	Dairy Products	2	98	0
Grandma Kelly's	Condiments	2	126	0
Grandma Kelly's	Produce	1	15	0
Heli SÃƒÂ¼ÃƒÂ	Confections	3	140	0
Karkki Oy	Beverages	1	57	0
Karkki Oy	Confections	2	75	60
Leka Trading	Beverages	1	17	10
Leka Trading	Condiments	1	27	0
Leka Trading	Grains/Cereals	1	26	0
Lyngbysild	Seafood	2	100	70
Ma Maison	Meat/Poultry	2	136	0
Mayumi's	Condiments	1	39	0
Mayumi's	Produce	1	35	0
Mayumi's	Seafood	1	24	0
New England Se	Seafood	2	208	0
New Orleans Caj	Condiments	4	133	100
Nord-Ost-Fisch H	Seafood	1	10	0
Norske Meierier	Dairy Products	3	164	0
Pasta Buttini s.r.l	Grains/Cereals	2	57	10
Pavlova, Ltd.	Beverages	1	15	10
Pavlova, Ltd.	Condiments	1	24	0
Pavlova, Ltd.	Confections	1	29	0
Pavlova, Ltd.	Meat/Poultry	1	0	0
Pavlova, Ltd.	Seafood	1	42	0
PB KnÃƒÂ¤ckeb	Grains/Cereals	2	165	0
Plutzer Lebensm	Beverages	1	125	0
Plutzer Lebensm	Condiments	1	32	0
Plutzer Lebensm	Grains/Cereals	1	22	80
Plutzer Lebensm	Meat/Poultry	1	0	0
Plutzer Lebensm	Produce	1	26	0
Refrescos Ameri	Beverages	1	20	0
Specialty Biscuit	Confections	4	74	50
Svensk SjÃƒÂ¶f	Seafood	3	224	50
Tokyo Traders	Meat/Poultry	1	29	0
Tokyo Traders	Produce	1	4	20
Tokyo Traders	Seafood	1	31	0
Zaanse Snoepfat	Confections	2	51	70



CONCLUSION:-

Key findings:

Supplier Performance: Supplier performance varies significantly across different product categories, with some suppliers specializing in specific categories and others supplying a wider range of products.

Product Category Distribution: The distribution of product categories varies among suppliers, with some suppliers focusing on a particular category (e.g., Forêts d'érables and Condiments, New Orleans Cajun Delights and Condiments) and others supplying a more diverse range of products.

Units in Stock and On Order: The number of units in stock and on order also varies among suppliers and product categories, indicating potential differences in demand and supply chain management.

Business insights:

Supplier Selection: The company should consider supplier performance and product category distribution when selecting suppliers, ensuring that they have a diverse range of products and are reliable in supplying those products.

Supply Chain Management: The company should monitor the number of units in stock and on order for each supplier and product category, ensuring that they have sufficient inventory to meet demand and that they are not overstocking low-demand products.

Product Category Strategy: The company should develop a product category strategy that takes into account supplier performance and distribution, ensuring that they are focusing on high-demand and high-margin categories.

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

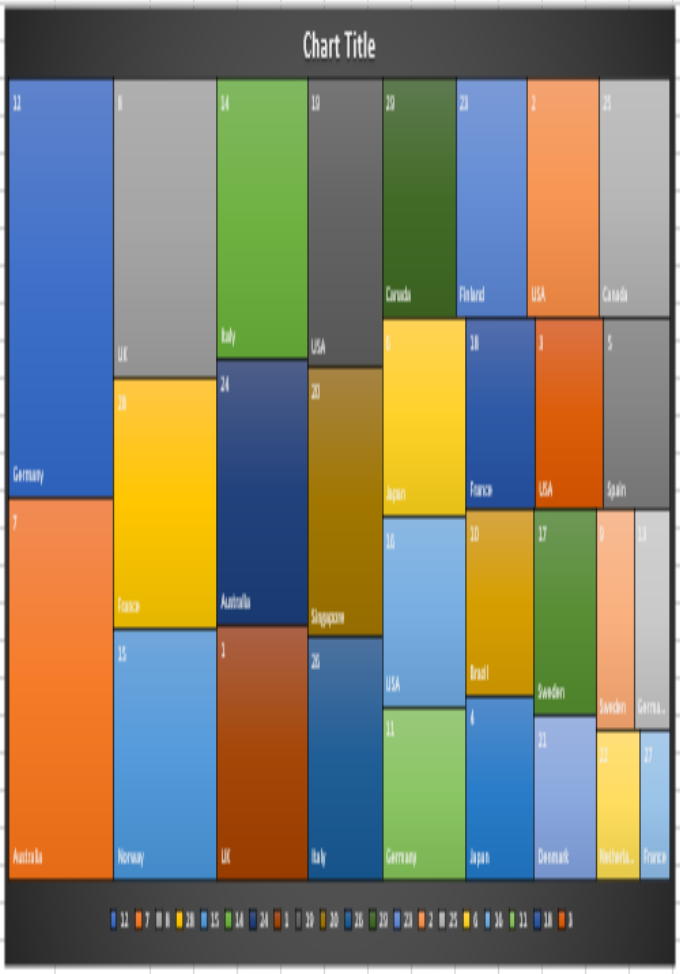
We don't have data for this question. If we want to solve this question so we need supplier coust for product supplay

So we create a question by own

Supplier with the Most Orders by Cuntry?

```
SELECT
  s.SupplierID,
  s.CompanyName,
  s.Country,
  COUNT(od.OrderID) AS OrderCount
FROM
  suppliers s
    JOIN
  products p ON s.SupplierID = p.SupplierID
    JOIN
  order_details  od ON p.ProductID = od.ProductID
    JOIN
  orders o ON od.OrderID = o.OrderID
    JOIN
  customers c ON o.CustomerID = c.CustomerID
GROUP BY s.SupplierID , s.Country, s.CompanyName
ORDER BY OrderCount DESC , s.Country;
```

SupplierID	CompanyName	Country	OrderCount
12	Plutzer Lebensmittelgroßmärkte	Germany	179
7	Pavlova, Ltd.	Australia	163
8	Specialty Biscuits, Ltd.	UK	126
28	Gai pâturage	France	105
15	Norske Meierier	Norway	105
14	Formaggi Fortini s.r.l.	Italy	104
24	G'day, Mate	Australia	98
1	Exotic Liquids	UK	94
19	New England Seafood Cannery	USA	88
20	Leka Trading	Singapore	82
26	Pasta Buttini s.r.l.	Italy	73
29	Forêts d'érables	Canada	72
23	Karkki Oy	Finland	70
2	New Orleans Cajun Delights	USA	70
25	Ma Maison	Canada	69
6	Mayumi's	Japan	68
16	Bigfoot Breweries	USA	65
11	Heli Süßwaren GmbH & Co	Germany	59
18	Aux joyeux ecclésiastiques	France	54
3	Grandma Kelly's Homestead	USA	54
5	Cooperativa de Quesos 'Las Cabras'	Spain	52
10	Refrescos Americanas LTDA	Brazil	51
4	Tokyo Traders	Japan	51
17	Svensk Sjöföda AB	Sweden	51
21	Lyngbysild	Denmark	41
9	PB Knäckebröd AB	Sweden	34
13	Nord-Ost-Fisch Handelsgesellschaft	Germany	32
22	Zaanse Snoepfabriek	Netherland	27



Conclusion:

By providing a visually compelling and user-friendly dashboard with interactive features, this Power BI report empowers stakeholders at Northwind Traders to make data-driven decisions. This comprehensive analysis of customers, sales, inventory, and employees offers valuable insights into business operations and facilitates data exploration. The expected impact is a transformation in how Northwind interacts with its data, ultimately enhancing competitiveness and driving the company forward in the wholesale market.

THANK YOU