

**CAPSTONE PROJECT**

NORTHWIND TRADERS SALES ANALYSIS





dATA ANALYTICS

GODABA VANDANA

**OVERVIEW**

Northwind Traders aims to analyze sales, customer behavior, and employee performance to improve decision-making. This report consolidates data from various sources to provide a clear business overview. Sales trends will highlight revenue patterns, top-selling products, and seasonal demand, helping optimize inventory and pricing strategies. Customer segmentation will categorize buyers based on location and purchase history to enhance marketing efforts. Inventory tracking will ensure efficient stock management, preventing shortages and excess supply. Supplier data will help evaluate vendor reliability and streamline procurement processes. Employee performance metrics will assess productivity, recognizing top contributors and identifying areas for improvement. Interactive visualizations and filters will make data exploration simple and effective. By transforming raw data into meaningful insights, this report will help Northwind Traders stay competitive and drive business growth efficiently.

The Process

**1.Data Acquisition from GitHub:**

Gather the necessary dataset of Northwind database from a designed GitHub repository, containing key business information such as sales transactions, customer details, employee records, product inventory, supplier data, and shipping information.

**2. Data Transformation and Enhancement:**  
Perform data cleaning and transformation to ensure accuracy, consistency, and completeness. Enhance the dataset by deriving new metrics such as total sales, customer segments, and inventory turnover rates for deeper insights.

**3. Connecting with Analytical Tools:**  
Establish seamless integration between the dataset and analytical tools such as Power BI, Excel and SQL. Utilize SQL for data extraction and transformation, while Power BI will serve as the primary tool for visualization and dashboard creation.

**4. Problem Statement Solution in Power BI:**Use Power BI to analyze sales trends, customer segmentation, employee performance, and inventory management. Implement interactive visualizations and dynamic filters to enable stakeholders to explore key business insights efficiently.

**5.Exploratory Data Analysis (EDA):**  
Conduct exploratory data analysis using either Excel or SQL Workbench depending on complexity of the analysis to identify sales patterns, customer preferences, and product performance. Leverage SQL queries and Power BI dashboards to uncover trends, correlations, and business opportunities and to inform subsequent decision-making.

**6.Creation of Visual and Insightful PowerPoint:**  
Develop a comprehensive PowerPoint presentation summarizing the analysis, insights, and key business recommendations. Each section should contain visual representations, data-driven conclusions, and action points for decision-makers.

**7.Detailed Documentation:**  
Compile a structured report documenting the entire process, including data collection, transformation, analysis, Power BI solutions, EDA findings, and visualizations. This report will serve as a reference for future business strategy and optimization.

**Objective:**

Northwind Traders is a renowned company specializing in the sale of food and beverage products. Understanding sales performance, customers behaviour, inventory management, and employee productivity is crucial for optimizing business operations and driving growth. However, challenges such as fluctuating sales trends, diverse customer preferences, inventory stockouts, and employee efficiency variations need to be addressed for sustainable success.

The objective of this project is to conduct an in-depth analysis of the Northwind dataset to uncover key business insights, sales trends, and operational efficiencies. By leveraging data-driven insights, the project aims to enhance decision-making, improve business performance, and optimize resource allocation.

**Project Scope & Key Tasks:**

* Conduct a comprehensive sales analysis, identifying trends, seasonal variations, and customer purchasing behaviors.
* Analyze customer segmentation to understand buying patterns based on demographics, order frequency, and order value.
* Evaluate employee performance, including sales contributions and order handling efficiency.
* Assess inventory management efficiency, including stock levels, product demand, and supplier performance.
* Generate meaningful insights and recommendations for business growth and operational improvements.
* Compile analysis results, insights, and recommendations into a structured report for stakeholders.

**Success Metrics:**

* **Quality of Analysis** – The depth and accuracy of data-driven insights.
* **Relevance of Insights** – The applicability of findings to improve business strategies.
* **Impact of Recommendations** – The effectiveness of proposed solutions in enhancing sales, customer experience, and operational efficiency.

**Significance of Northwind Sales Analysis:**

Sales analysis is essential for businesses like Northwind Traders, providing key insights that drive growth, improve efficiency, and enhance customer satisfaction. By analyzing sales performance, customer behavior, inventory management, and employee productivity, Northwind can make informed business decisions that lead to better profitability and sustainability.

For **business managers**, sales analysis helps identify top-performing products, seasonal trends, and revenue-driving factors. This enables better demand forecasting, optimized pricing strategies, and improved marketing efforts to maximize sales potential.

For **inventory managers**, understanding stock levels and product demand ensures efficient inventory control. Analyzing sales patterns helps prevent overstocking or stockouts, reducing waste and improving supply chain efficiency. This allows Northwind to maintain the right balance between supply and demand.

For **customer relationship management (CRM)**, sales data provides insights into customer preferences, purchasing behavior, and order frequency. By segmenting customers based on demographics and buying patterns, Northwind can personalize promotions, enhance customer engagement, and build long-term relationships.

For **employees and sales teams**, performance analysis helps in recognizing top performers and identifying areas for improvement. Understanding individual and team contributions enables better sales strategies, training programs, and incentive structures, ultimately leading to a more productive workforce.

For **business decision-makers and stakeholders**, sales analysis plays a crucial role in shaping future business strategies. Data-driven insights allow leadership to make informed decisions on expansion, supplier relationships, and investment priorities, ensuring long-term business success.

In conclusion, Northwind sales analysis is a vital tool for optimizing business operations, improving customer experience, and increasing profitability. By leveraging sales data effectively, Northwind Traders can enhance decision-making, strengthen its competitive edge, and achieve sustainable growth in the market.

**Data Dictionary for Northwind Sales Analysis:**

The Northwind database contains sales data for a fictitious company, “Northwind Traders,” which imports and exports specialty foods from around the world. The data is structured into multiple related tables that store information about customers, employees, orders, products, suppliers, shippers, and categories.

**Tables and Fields**

**1. Customers Table**

* **customer\_id**: Unique identifier for each customer.
* **company\_name**: Name of the customer’s company.
* **contact\_name**: Name of the primary contact person.
* **contact\_title**: Job title of the contact person.
* **address**: Street address of the customer.
* **city**: City where the customer is located.
* **region**: Region where the customer is located (if applicable).
* **postal\_code**: Postal code of the customer’s location.
* **country**: Country where the customer is based.
* **phone**: Primary phone number of the customer.
* **fax**: Fax number of the customer (if available).

**2. Employees Table**

* **employee\_id**: Unique identifier for each employee.
* **last\_name**: Last name of the employee.
* **first\_name**: First name of the employee.
* **title**: Job title of the employee.
* **title\_of\_courtesy**: Courtesy title (e.g., Mr., Mrs., Dr.).
* **birth\_date**: Employee’s birth date.
* **hire\_date**: Date when the employee was hired.
* **address**: Street address of the employee.
* **city**: City where the employee resides.
* **region**: Region where the employee resides (if applicable).
* **postal\_code**: Postal code of the employee’s location.
* **country**: Country where the employee is based.
* **home\_phone**: Home phone number of the employee.
* **extension**: Office extension number.
* **photo**: Employee’s photograph.
* **notes**: Additional notes about the employee.
* **reports\_to**: Employee ID of the manager/supervisor.
* **photo\_path**: Path to the employee’s photo.

**3. Orders Table**

* **order\_id**: Unique identifier for each order.
* **customer\_id**: Foreign key referencing the Customers table.
* **employee\_id**: Foreign key referencing the Employees table.
* **order\_date**: Date when the order was placed.
* **required\_date**: Date by which the order is required.
* **shipped\_date**: Date when the order was shipped.
* **ship\_via**: Foreign key referencing the Shippers table.
* **freight**: Shipping cost associated with the order.
* **ship\_name**: Name of the recipient for shipping.
* **ship\_address**: Shipping address.
* **ship\_city**: City of the shipping destination.
* **ship\_region**: Region of the shipping destination (if applicable).
* **ship\_postal\_code**: Postal code of the shipping destination.
* **ship\_country**: Country of the shipping destination.

**4. Order Details Table**

* **order\_id**: Foreign key referencing the Orders table.
* **product\_id**: Foreign key referencing the Products table.
* **unit\_price**: Price per unit of the product at the time of sale.
* **quantity**: Number of units sold in the order.
* **discount**: Discount applied to the order (if any).

**5. Products Table**

* **product\_id**: Unique identifier for each product.
* **product\_name**: Name of the product.
* **supplier\_id**: Foreign key referencing the Suppliers table.
* **category\_id**: Foreign key referencing the Categories table.
* **quantity\_per\_unit**: Quantity of product per unit package.
* **unit\_price**: Price per unit of the product.
* **units\_in\_stock**: Number of units available in stock.
* **units\_on\_order**: Number of units currently on order.
* **reorder\_level**: Minimum stock level before reordering is needed.
* **discontinued**: Indicator of whether the product is still available (1 = Discontinued, 0 = Available).

**6. Suppliers Table**

* **supplier\_id**: Unique identifier for each supplier.
* **company\_name**: Name of the supplier company.
* **contact\_name**: Name of the primary contact person.
* **contact\_title**: Job title of the contact person.
* **address**: Street address of the supplier.
* **city**: City where the supplier is located.
* **region**: Region where the supplier is located (if applicable).
* **postal\_code**: Postal code of the supplier’s location.
* **country**: Country where the supplier is based.
* **phone**: Primary phone number of the supplier.
* **fax**: Fax number of the supplier (if available).
* **home\_page**: Website or online presence of the supplier.

**7. Shippers Table**

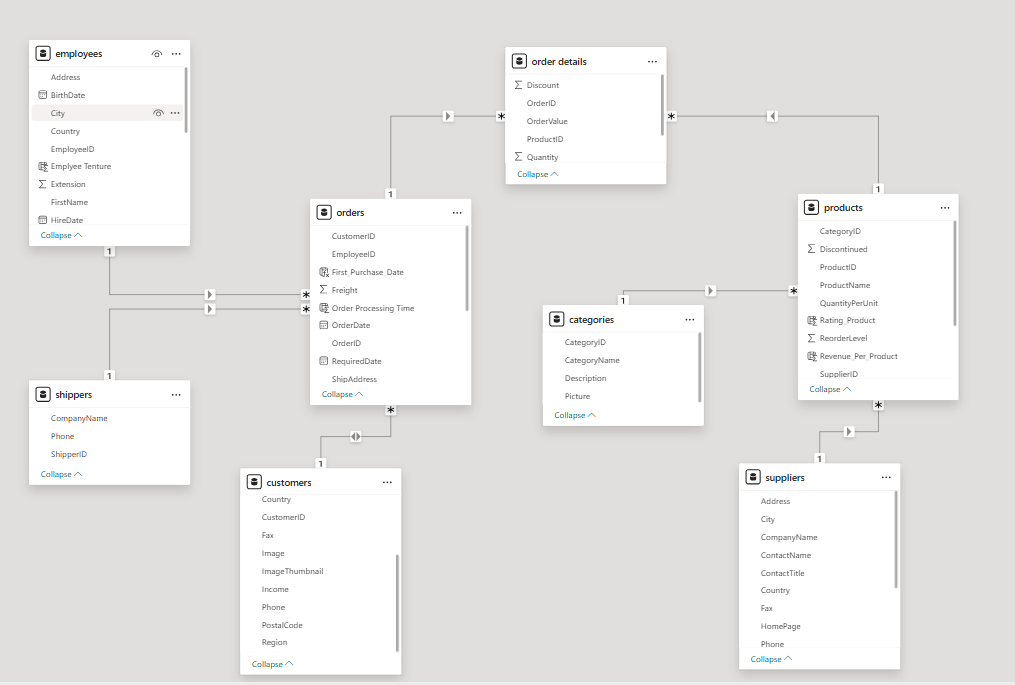
* **shipper\_id**: Unique identifier for each shipping company.
* **company\_name**: Name of the shipping company.
* **phone**: Contact phone number of the shipping company.

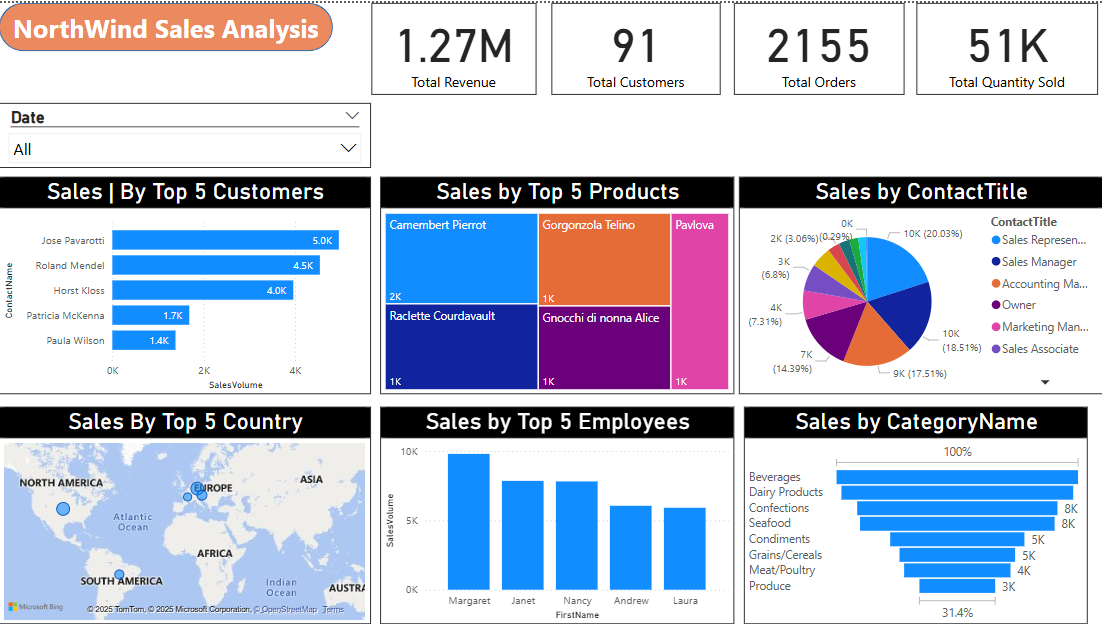
**8. Categories Table**

* **category\_id**: Unique identifier for each product category.
* **category\_name**: Name of the product category.
* **description**: Description of the product category.

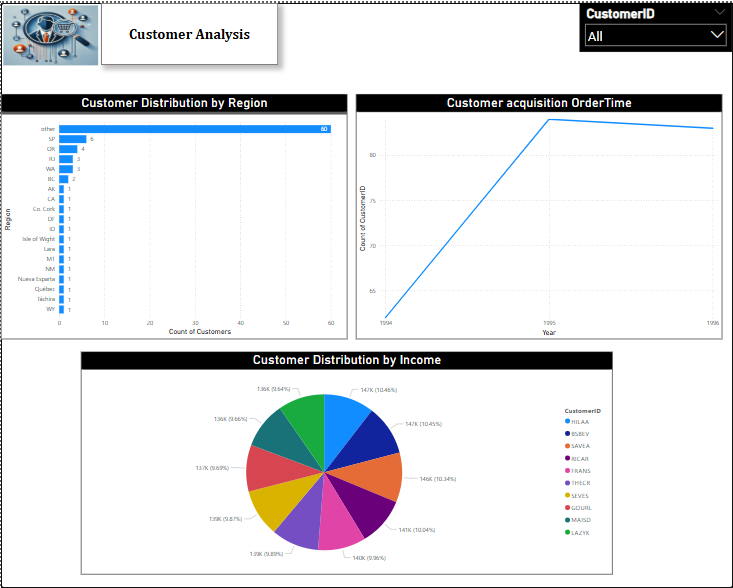
**Conclusion**

This data dictionary provides a comprehensive overview of the Northwind sales dataset. It defines the structure and relationships between tables, facilitating better understanding and seamless integration into analytical tools like Power BI. The well-organized database allows for detailed sales analysis, customer segmentation, inventory tracking, and performance evaluation, helping Northwind Traders make data-driven business decisions.

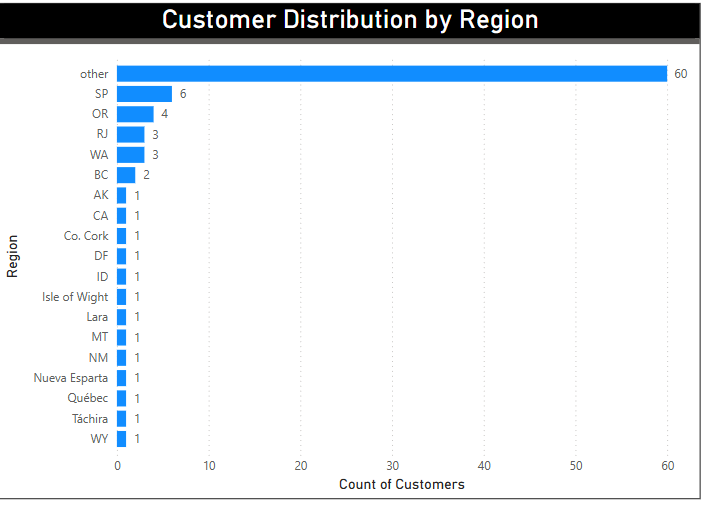
 ER DIAGRAM

POWERBI PROBLEM STATEMENTS

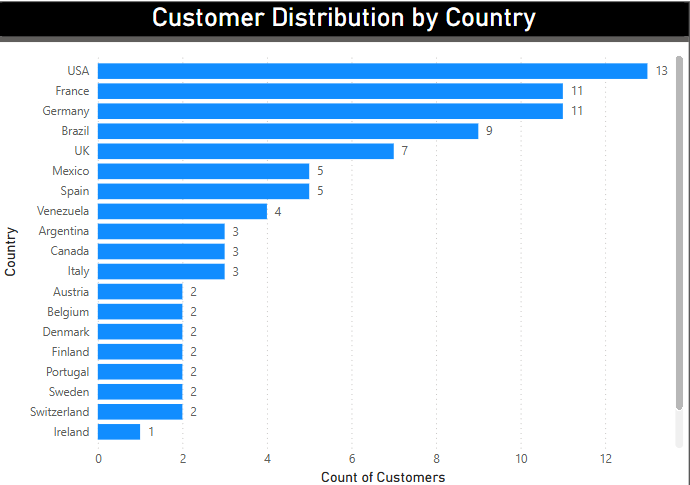
CUSTOMER ANALYSIS SNAPSHOT



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



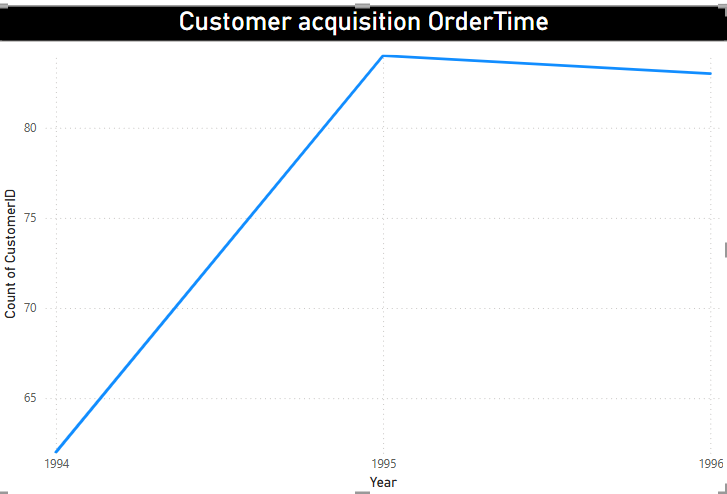
The chart shows that most customers (60) fall under the "Other" category, indicating a high concentration in unspecified regions. São Paulo (SP) has the next highest count (6), followed by Oregon (4), with other regions having very few customers. This suggests a fragmented customer base, and further analysis of the "Other" category could help identify growth opportunities in specific regions. Here in the given dataset, they did not give region names for most of the rows that’s why I have taken other instead of empty region.



So, for getting correct output I took country name and I found Customer Distribution by Country Name (like the second chart). The chart shows that most customers (13) fall under USA and France and Germany has the next highest count (11) and Ireland has a smaller number of Customers.

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

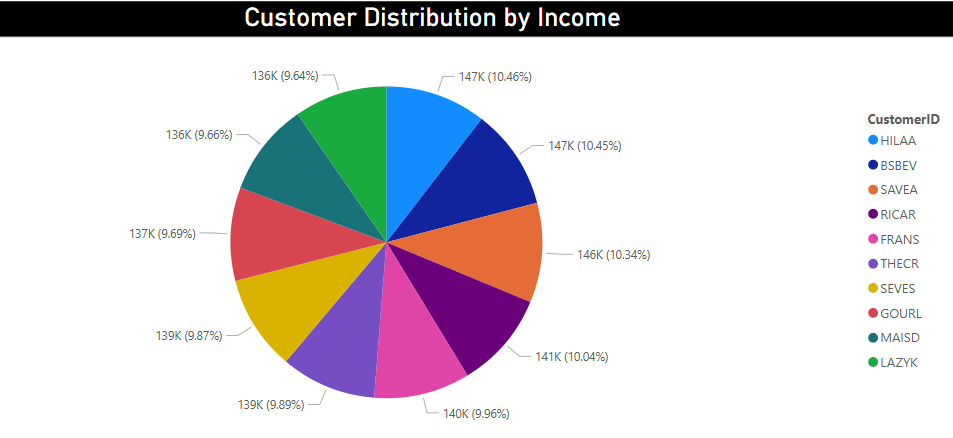
The chart indicates an increase in customer acquisition from 1994 to 1995, followed by a slight decline in 1996. The steep rise in 1995 suggests a strong growth phase, possibly due to successful business strategies or market expansion. However, the slight dip in 1996 may indicate retention challenges or market saturation. Further analysis is needed to understand the reasons behind these trends and identify opportunities for sustained growth.



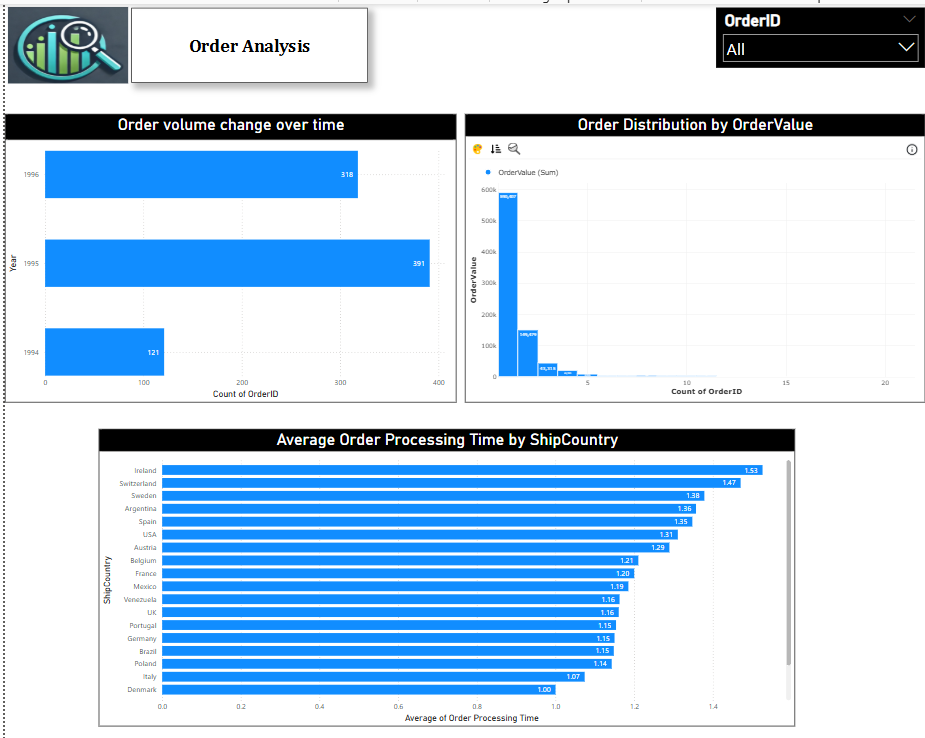
**Can we visualize the distribution of customer demographic such as income?**

The pie chart illustrates the distribution of customers based on income, with income levels ranging from 136K to 147K, each holding a fairly equal share. The variation among income groups is minimal, suggesting a relatively uniform customer base in terms of earnings. This insight can help in targeting marketing strategies and tailoring product offerings based on income segments. Further demographic analysis, including age and gender, can provide deeper insights into customer preferences.

And here the highest income has been recorded for HILAA with 147k and Least with LAZYK with 136k.

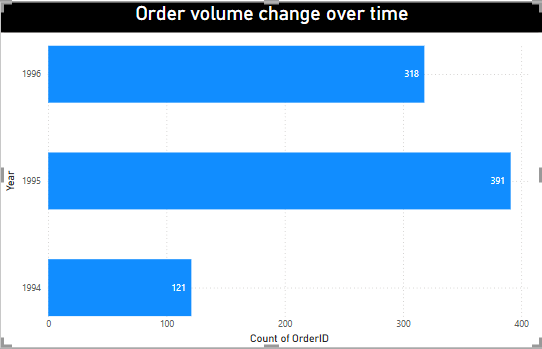


ORDER ANALYSIS SNAPSHOT



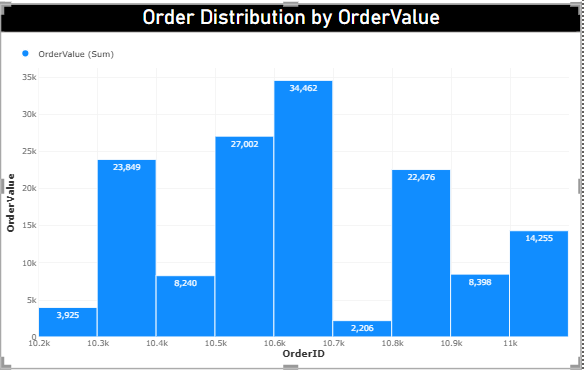
**How does order volume change over time?**

The chart shows a clear upward trend in order volume from **1994 to 1995**, with orders increasing significantly from **121 to 391**. However, in **1996**, there is a decline to **318 orders**, suggesting a potential drop in demand or operational changes. This insight can help in identifying factors influencing sales fluctuations, such as **seasonality, marketing efforts, or supply chain issues**. Further analysis can determine the reasons behind the 1996 decline and opportunities for growth.



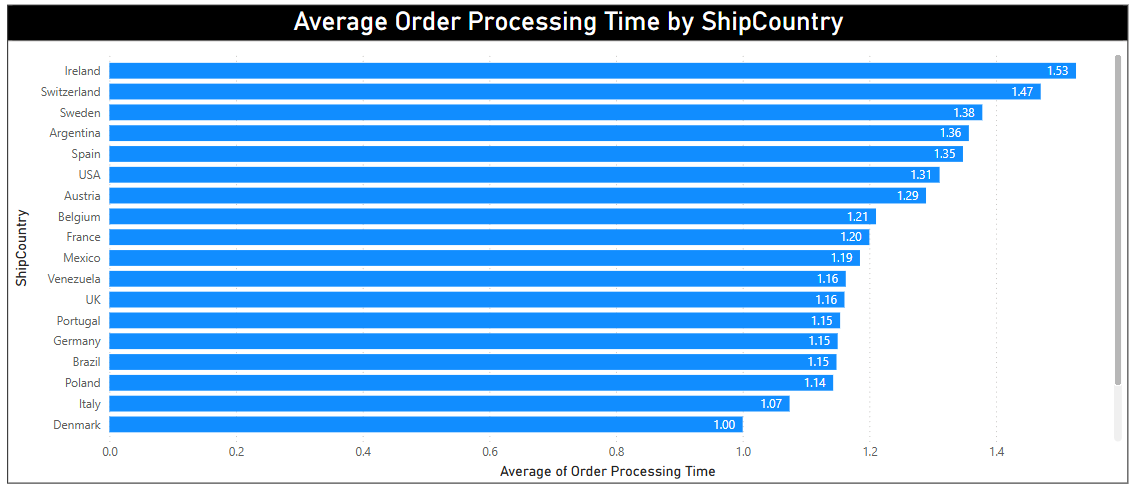
**What is the distribution of order values?**

The order value distribution, as shown in the chart, exhibits significant variation across different Order IDs. The highest order value is 34,462, indicating a peak in sales for a specific order. Other notable high-value orders include 27,002, 23,849, and 22,476, showing consistent large transactions. However, there are also relatively low order values, with the lowest being 2,206, followed by 3,925 and 8,240, highlighting fluctuations in purchasing patterns. This distribution suggests that while some orders contribute significantly to total revenue, others remain relatively small. Such insights can help in identifying high-value orders, optimizing pricing strategies, and improving sales forecasting.



**Can we visualize the average order processing time?**

The bar chart illustrates the average order processing time across different shipping countries. **Ireland (1.53 Weeks), Switzerland (1.47 Weeks), and Sweden (1.38 Weeks)** have the longest processing times, while **Denmark (1.00 days), Italy (1.07 Weeks), and Poland (1.14 Weeks)** are the most efficient. Most countries fall within the **1.10 - 1.35 Weeks** range, indicating relatively consistent processing times. The variations may be due to **logistics efficiency, shipping infrastructure, or operational factors**. To enhance efficiency, it is crucial to identify the causes of delays in high-processing-time regions and optimize logistics accordingly.

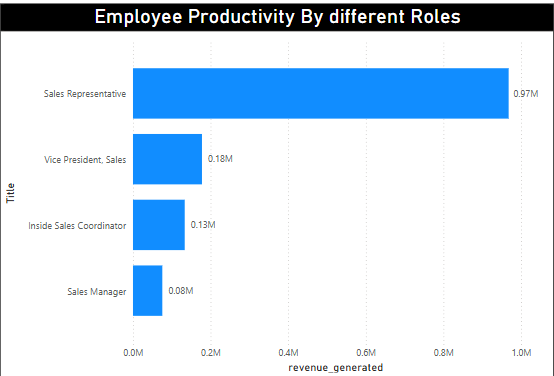


EMPLOYEE ANALYSIS SNAPSHOT



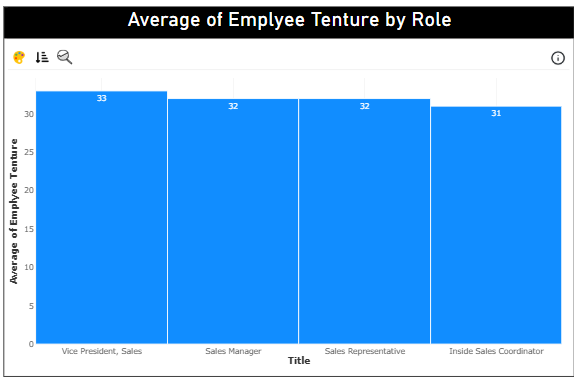
**How does employee productivity vary across different job roles?**

The chart illustrates employee productivity across different job roles based on revenue generated. Sales Representatives exhibit the highest productivity, generating approximately 0.97M in revenue, significantly outperforming other roles. The Vice President of Sales follows with 0.18M, while Inside Sales Coordinators and Sales Managers contribute 0.13M and 0.08M, respectively. This suggests that Sales Representatives play a crucial role in driving revenue, whereas managerial and coordination roles contribute comparatively less.



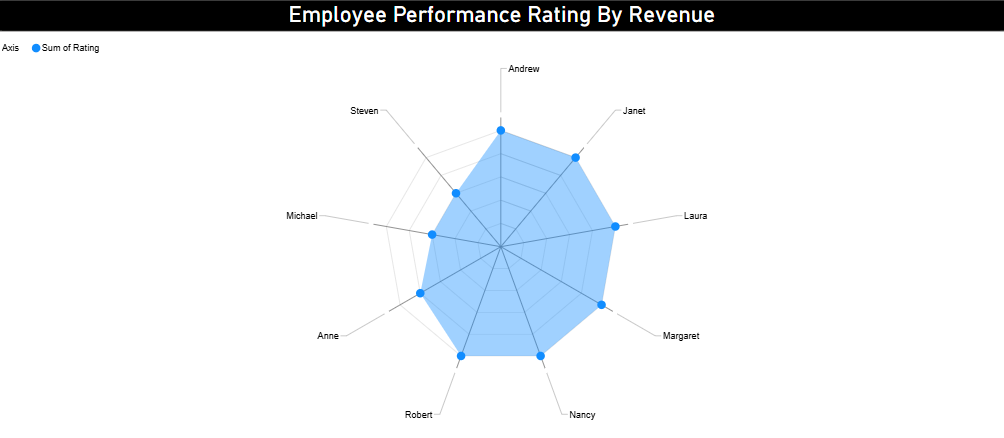
**What is the distribution of employee tenure?**

The chart presents the distribution of employee tenure across different roles, indicating that tenure remains relatively consistent across job titles. The Vice President of Sales has the highest average tenure at 33, followed closely by Sales Managers and Sales Representatives at 32, while Inside Sales Coordinators have the lowest at 31. This suggests that employees, regardless of their role, tend to have a stable tenure within the organization, with minimal variation across positions.



**Can we visualize employee performance ratings?**

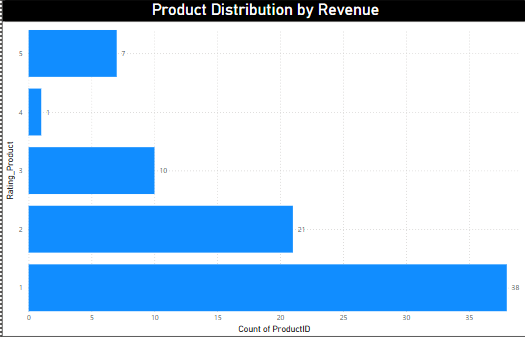
The radar chart effectively visualizes employee performance ratings by revenue, highlighting variations among individuals. Employees like Andrew, Janet, Laura, Margaret, Nancy, Robert appear to have higher ratings, while others, Anne got 4 rating and Michael and Steeven got 3 ratings. The concentration of ratings between 4 and 5 suggests consistently strong performance across the team, with some employees excelling more than others. This visualization helps in identifying top performers and areas where improvement may be needed.



 PRODUCT ANALYSIS SNAPSHOT

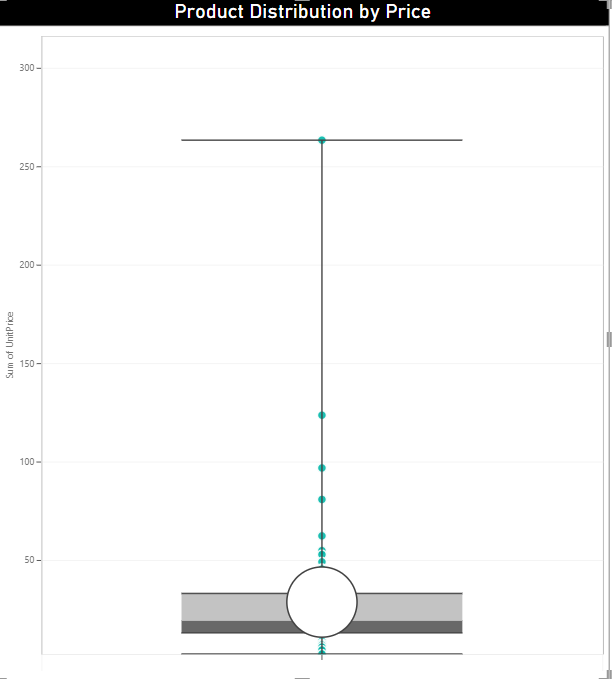
**What is the distribution of product ratings?**

The product distribution chart, based on revenue-based rating conditions, shows that most products fall into the lower revenue categories. Specifically, 38 products have a rating of 1, meaning their revenue per product is below $10,000. Similarly, 21 products are rated 2, indicating revenue between $10,000 and $20,000. The number of products decreases as the revenue threshold increases, with 10 products rated 3 ($20,000–$30,000), 1 product rated 4 ($30,000–$40,000), and only 7 products achieving the highest rating of 5 (above $40,000). This distribution suggests that a significant portion of products generate relatively low revenue, while only a few achieve high performance. It may be beneficial to investigate factors affecting product sales and explore strategies to boost revenue across more products.



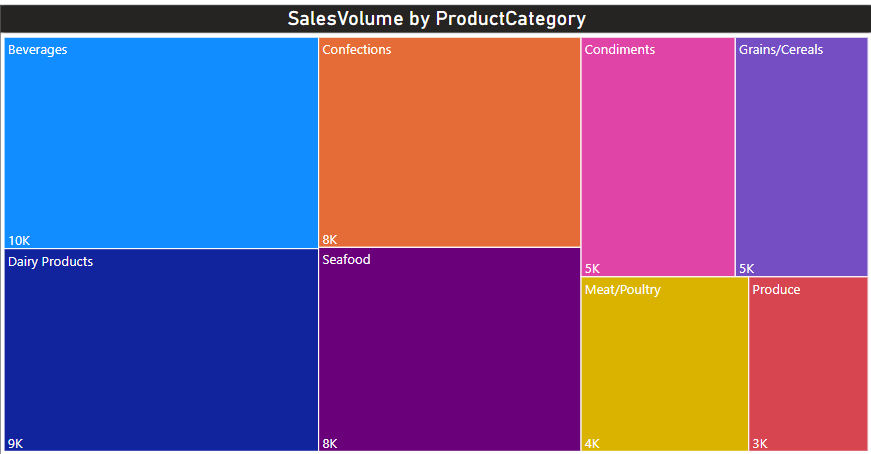
**Can we visualize the pricing distribution of products using a box plot**?

A box plot is an effective way to visualize the pricing distribution of products, as shown in the given chart. The majority of product prices are concentrated within a lower range, indicated by the dense cluster in the box. However, there are significant outliers, such as **CôtedeBlaye (263.5)** and **Thüringer Rostbratwurst (127.9)**, which stand far above the main distribution. The **least-priced product, Geitost (2.5)**, falls within the lower range. The long upper whisker in the box plot suggests that while most products are priced relatively low, a few high-priced items significantly extend the price range. This indicates a **right-skewed distribution**, where premium-priced products create substantial variance in pricing.

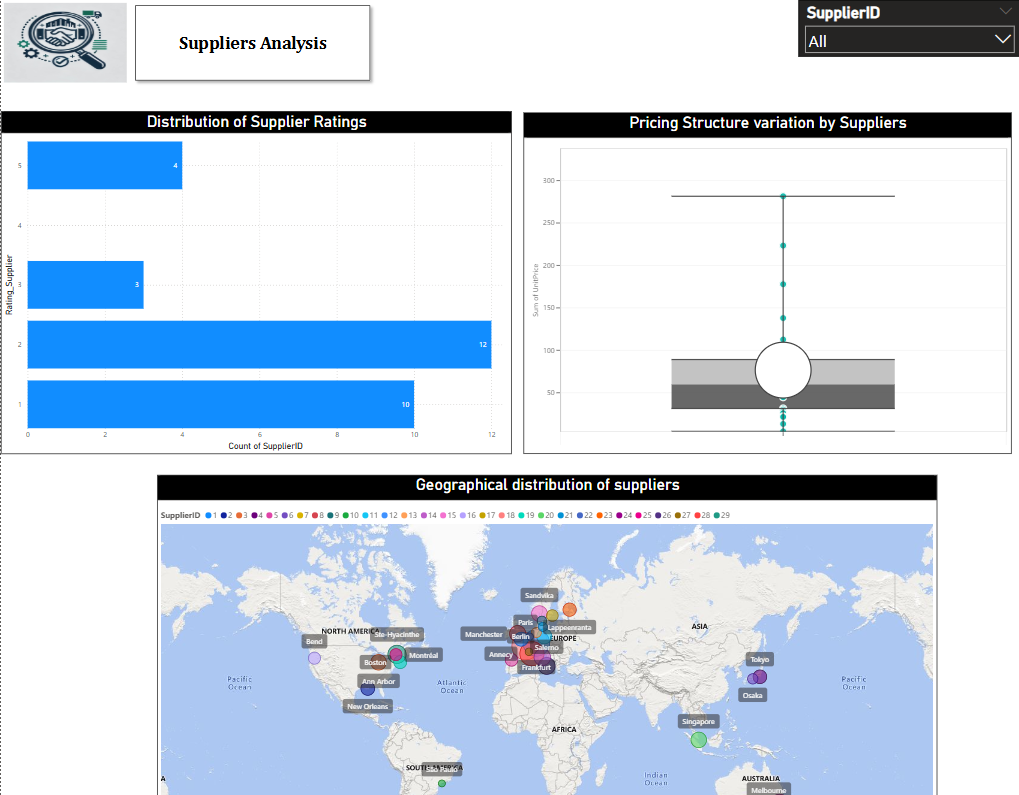


**How does the sales volume vary across different product categories?**

The sales volume varies significantly across different product categories. **Beverages** led with the highest sales volume at **10K**, followed by **Dairy Products (9K)** and **Confections (8K)**, indicating strong consumer demand in these categories. **Seafood** also has a notable sales volume of **8K**, while **Condiments and Grains/Cereals** have moderate sales at **5K** each. However, **Meat/Poultry (4K)** and **Produce (3K)** show the lowest sales volumes, suggesting potential areas for improvement through targeted marketing or inventory optimization. This distribution highlights the need to focus on high-performing categories while strategizing growth for lower-selling products.

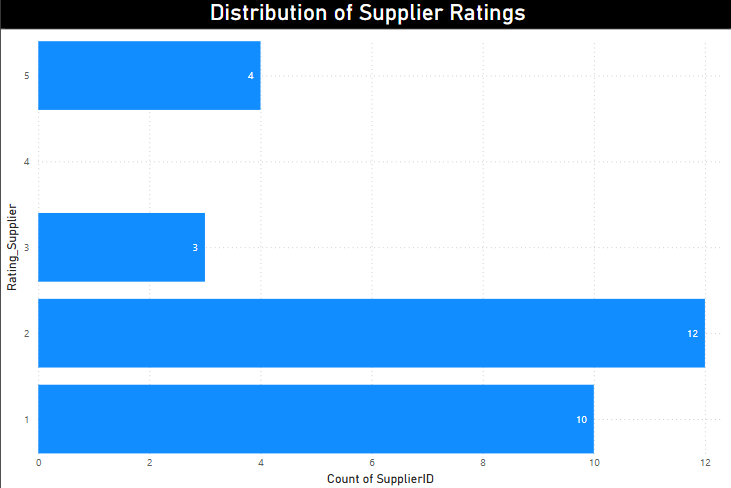


SUPPLIER ANALYSIS SNAPSHOT



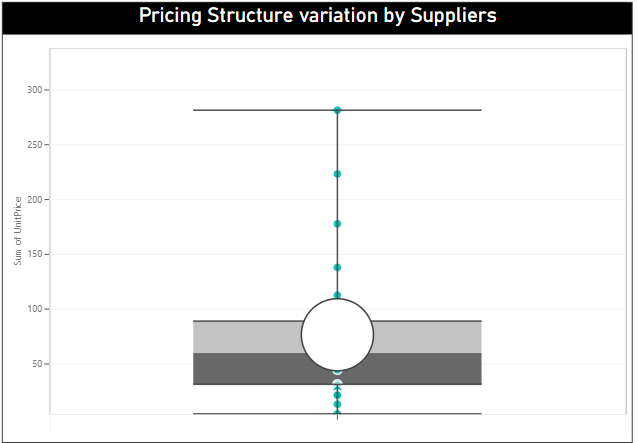
**What is the distribution of supplier ratings?**

The distribution of supplier ratings based on revenue generated, as shown in the chart, indicates that the majority of suppliers have received lower ratings. Rating **2(greater than 25000 revenue and below 50000 revenues generated)** has the highest count with **12 suppliers**, followed by **rating 1** with **10 suppliers**, showing a concentration in the lower rating range. In contrast, higher ratings are less frequent, with only **4 suppliers each receiving ratings of 3 and 5**, and no suppliers rated **4**. This suggests that supplier performance is skewed toward the lower end, indicating potential concerns with supplier quality or consistency. It may be beneficial to investigate the factors contributing to these lower ratings and take steps to improve supplier performance.



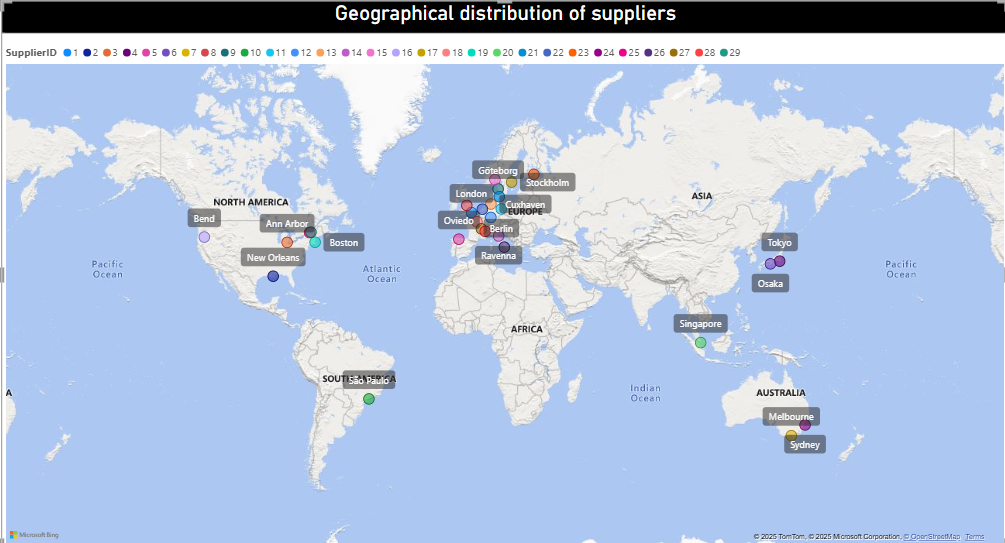
**How does the pricing structure vary across different suppliers?**

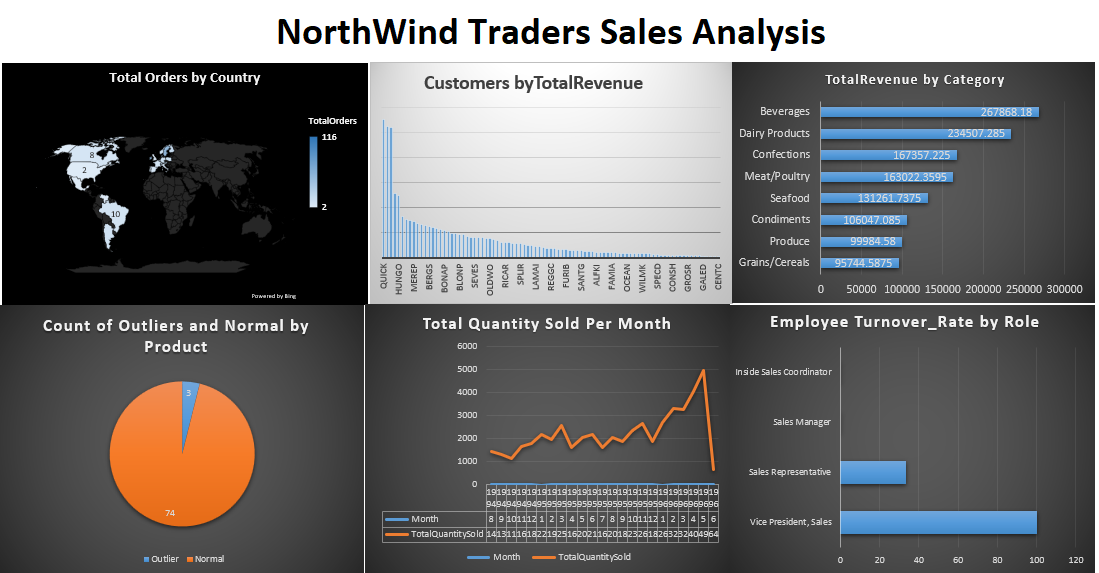
The pricing structure across different suppliers, as depicted in the box plot, shows significant variation. Supplier **ID 12** has the highest sum of unit prices at **223.39**, followed by Supplier **ID 7** with **177.85**, indicating that these suppliers provide higher-priced products. On the other hand, Supplier **ID 10** has the lowest sum of unit prices at **4.50**, highlighting a substantial disparity in pricing among suppliers. The presence of outliers in the chart suggests that certain suppliers offer exceptionally high-priced products, while most suppliers have pricing concentrated within a lower and more consistent range. This variation in pricing could be due to differences in product categories, quality, or market positioning among suppliers.



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

The geographical distribution of suppliers, as visualized on the map, provides valuable insights into their locations and revenue contributions. Supplier ID 18 from France has the highest revenue of 163,135, followed closely by Supplier ID 12 from Germany with 155,956, revenue indicating strong market presence and sales performance in these regions. On the other hand, Supplier ID 10 from Brazil, with the lowest revenue of 4,286.5, suggests limited business activity. The distribution highlights a diverse supplier network spanning North America, Europe, Asia, Australia, and South America, with key hubs in cities like Berlin, London, Tokyo, and São Paulo. This visualization is useful for identifying high-performing regions, optimizing supplier partnerships, and strategizing future business expansion.

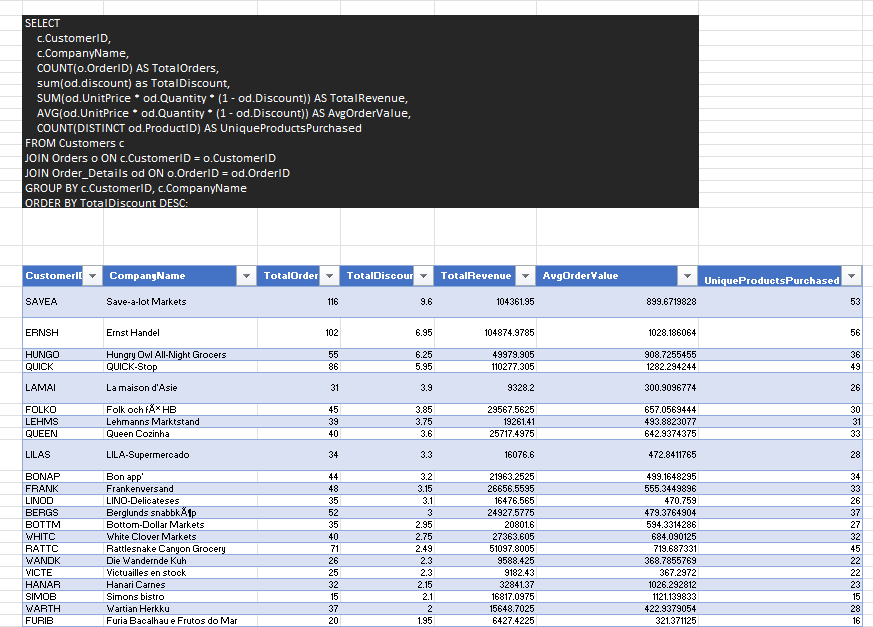


EDA Problem statements

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

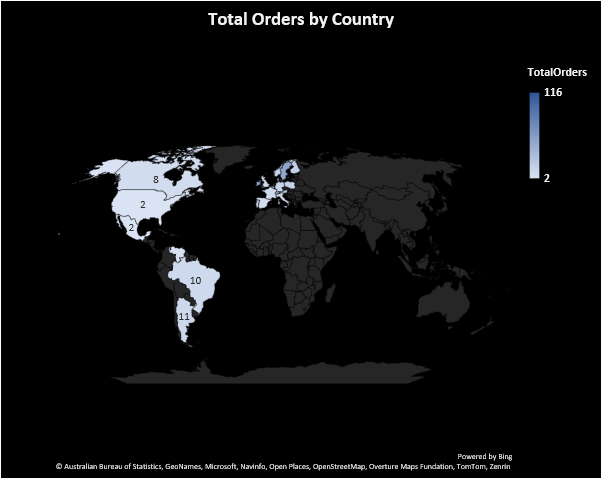
The chart visualizes the relationship between total orders and total discounts received by different customers, highlighting customer retention trends. The **Orange bars** represent total orders, while the **gray bars** indicate total discounts, showing how frequently customers purchase and how much discount they utilize. Customers like **SAVEA (120 orders, low discount)** and **HUNGO (100 orders, low discount)** demonstrate strong retention with minimal discounts, suggesting they value **product quality, availability, and service** over price reductions. Meanwhile, **WELLI (40+ orders, higher discount)** and **RANCH (low orders, high discount)** indicate price-sensitive buyers who may need promotions to stay engaged. Additionally, **GALED and CHOPS (low orders, minimal discount)** suggest one-time buyers who may require targeted retention strategies. Beyond discounts, factors like **delivery speed, brand reputation, customer service, and personalized loyalty programs** also influence retention. Since discounts alone do not guarantee loyalty, Northwind Traders should optimize these aspects to build long-term customer relationships and enhance business sustainability.

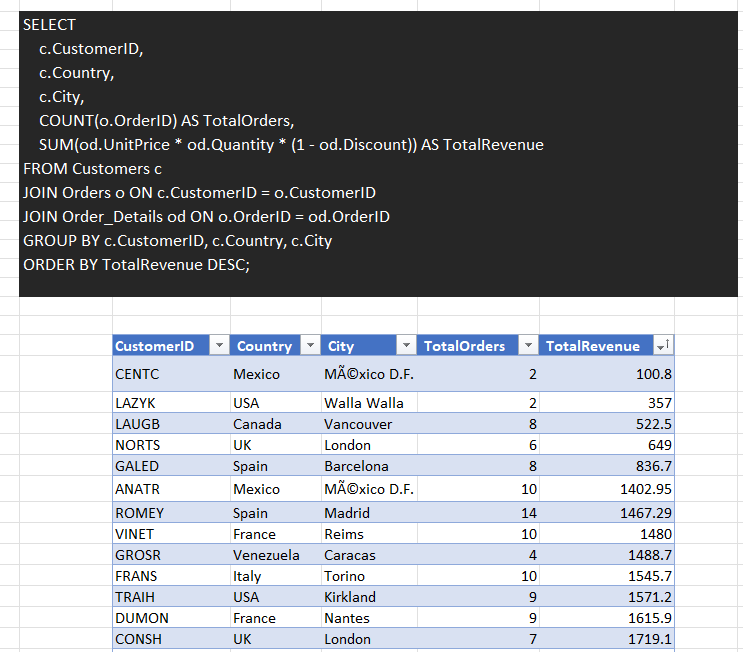
From the below SQL query and from the table we can find Factors based on Customer loyalty (like total orders and discount).



**How do customer preferences vary based on their location or demographics?**

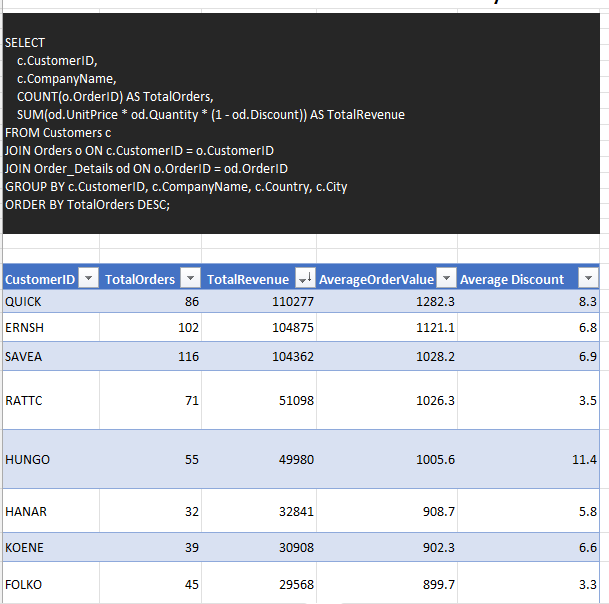
The chart reveals significant variations in customer preferences based on location. **Boise, USA, recorded the highest number of orders (116), making it the top-performing region, while Walla Walla, USA, had the lowest (2 orders).** **Germany generated the highest revenue (110,277.30), with Cunewalde City as a major contributor, whereas Mexico City had the lowest revenue (100.8).** These insights indicate that **customer behavior is influenced by location, order value, and product demand.** By analyzing **product preferences, order value, and purchase quantity across regions, businesses can better understand customer needs.** **Identifying these trends allows for more effective marketing, inventory planning, and sales strategies.** **Tailoring offerings based on location-specific demands can improve customer retention and overall business growth.**



From the below SQL query and from the table we can find customer preferences vary based on their location or demographics.

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

The **Customer and Average Order Value** chart shows two key segments: **high-value, low-frequency buyers** (e.g., QUICK, SAVEA) and **low-value, high-frequency buyers** (e.g., CENTC, GALED). High AOV customers make bulk purchases, while others place smaller, frequent orders. This segmentation helps in designing **targeted marketing strategies**. Businesses can offer personalized incentives to maximize revenue from each group.

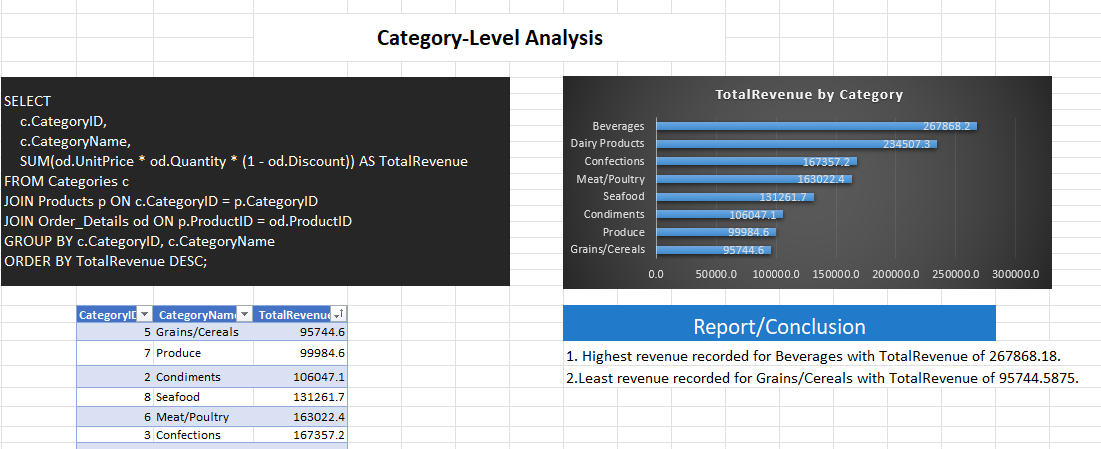


The **Total Orders vs. Average Order Value** chart highlights an **inverse relationship**—more orders generally mean lower AOV. Frequent buyers likely purchase low-cost products, while bulk buyers make larger transactions. Businesses can **offer bulk discounts to high-AOV customers** and **cross-sell or bundle products** to increase revenue from low-AOV buyers. Identifying key drivers like discounts and product preferences can further optimize sales strategies.

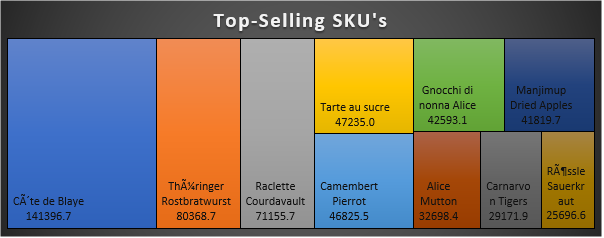
High-value, low-frequency buyers and low-value, high-frequency buyers exhibit distinct purchasing behaviors, allowing businesses to optimize revenue through targeted discounts, bundling, and personalized marketing strategies.

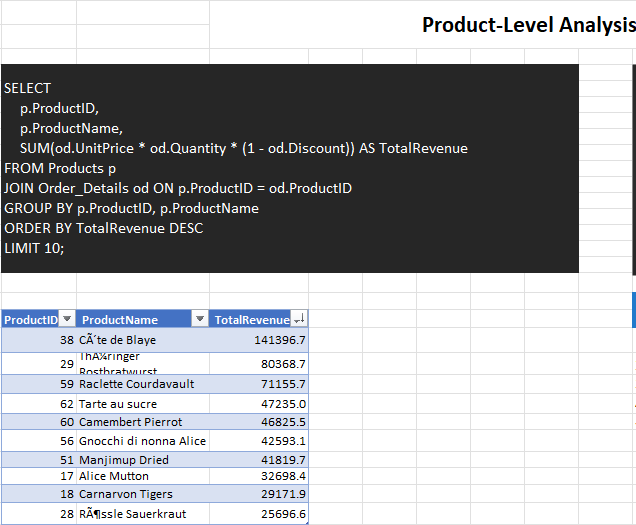
**Are there any specific product categories or SKUs that contribute significantly to order revenue?**

The analysis reveals that the "Beverages" category contributes the highest revenue (267,868.18), while "Grains/Cereals" generates the least (95,744.58). This indicates a significant variance in revenue across different product categories, emphasizing the importance of high-performing categories in driving overall sales.



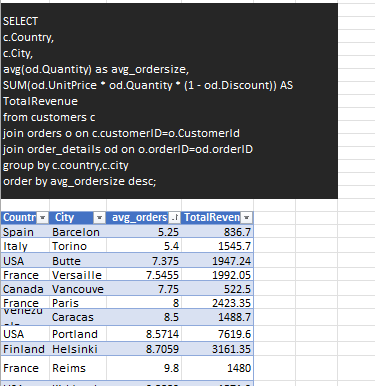
At the SKU level, "CÃ´te de Blaye" is the top-selling product with a revenue of 141,396.73, while "RÃ¤ssle Sauerkraut" records the lowest among the top 10 (25,696). Identifying these high-revenue products helps optimize stock availability, meet demand, and refine pricing strategies for profitability.



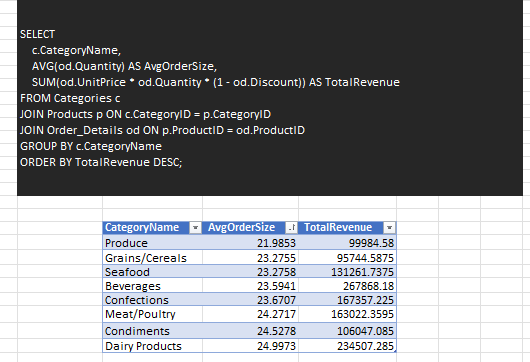


**Are there any correlations between order size and customer demographics or product categories?**

The analysis indicates a positive correlation between average order size and total revenue, with larger orders contributing significantly to revenue growth. Certain cities, such as Portland and Buenos Aires, exhibit higher average order sizes alongside substantial total revenue, suggesting that bulk purchasing plays a key role in driving sales. However, there are exceptions where revenue does not increase proportionally with order size, highlighting the need for further analysis on pricing strategies and customer preferences.

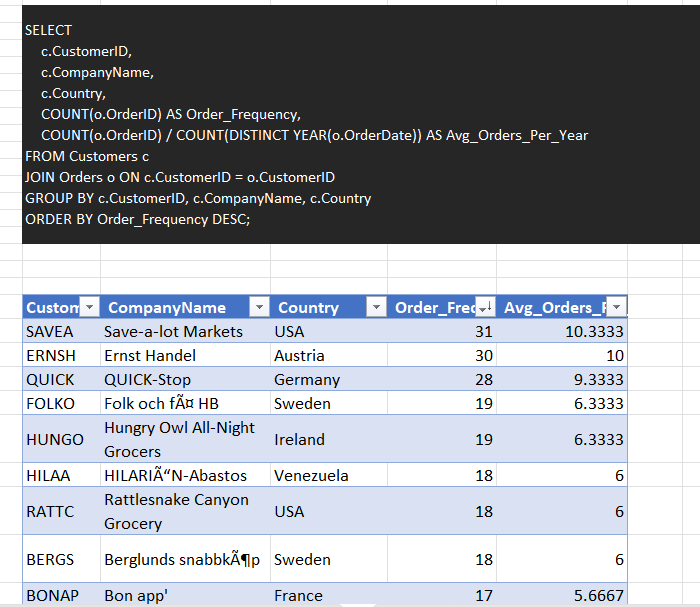


Regarding product categories, the data shows fluctuations in the relationship between average order size and total revenue. Beverages and dairy products generate the highest revenue despite varying order sizes, suggesting that product demand and pricing strategies significantly impact revenue. Some categories, such as condiments and grains/cereals, contribute less to overall revenue despite higher order sizes, indicating that order size alone is not the sole factor influencing profitability.



**How does order frequency vary across different customer segments?**

The **bar chart** shows that a few customers place the majority of orders, with **SAVEA, HILAA, and FRANK** being the most frequent buyers. The order frequency decreases significantly for other customers, indicating that a **small group of high-value customers** contributes heavily to sales. These customers can be targeted for loyalty programs or exclusive offers to encourage repeat purchases.

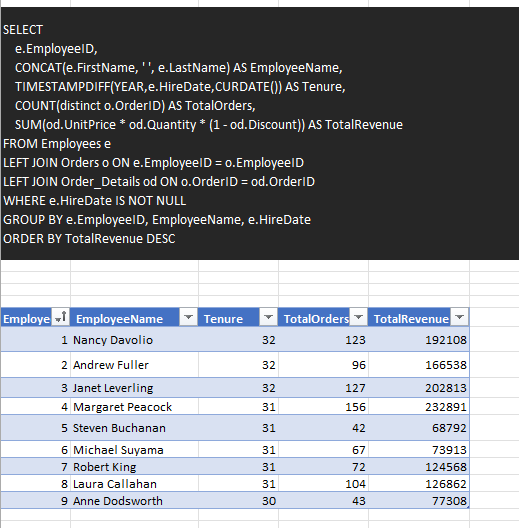


The **Treemap** highlights order frequency by country, with **Germany and the USA** leading at **122 orders each**, followed by **Brazil (83) and France (77)**. Countries like the **UK, Venezuela, and Austria** also contribute significantly, while **Poland and Norway** have much lower order volumes. This suggests that businesses should focus marketing efforts on top-performing regions while exploring ways to increase engagement in lower-order regions through localized promotions or better distribution.



**Are there any correlations between employee satisfaction levels and key performance indicators?**

The **line chart** shows the relationship between employee **tenure and total orders** processed. The trend fluctuates significantly, with no clear upward or downward pattern, indicating that tenure alone may not directly correlate with order processing efficiency. Some employees with longer tenure handle a high volume of orders, while others show lower productivity, suggesting that other factors like job satisfaction, motivation, or workload distribution might be influencing performance.



To explore these further, additional factors like **employee satisfaction scores, workload, and incentives** should be included in the analysis. A **scatter plot** comparing satisfaction levels with total orders could help identify whether happier employees tend to perform better. If a strong correlation exists, businesses can invest in improving workplace conditions to enhance productivity.

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

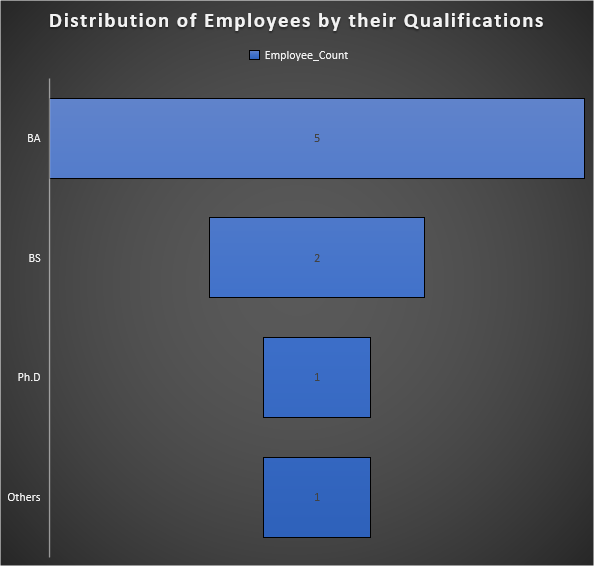
The **bar chart** illustrates the variation in **turnover rates** across different job roles. The highest turnover is observed among **Vice Presidents of Sales**, followed by **Sales Representatives**, while other roles like **Inside Sales Coordinators and Sales Managers** show little to no turnover. This suggests that higher-level sales positions may experience greater pressure, responsibility, or dissatisfaction, leading to higher attrition.

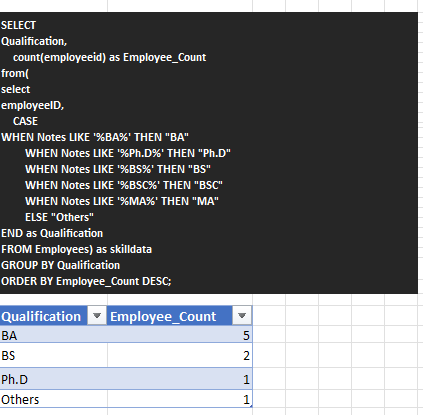


This can help identify problem areas where intervention—such as better incentives, work-life balance initiatives, or career growth opportunities—may reduce turnover and improve employee retention.

**Can we identify any patterns or clusters in employee skill sets or qualifications through visualizations?**

The funnel chart visually represents the distribution of employee qualifications. The largest segment consists of employees with a **BA degree**, followed by a smaller number with **BS degrees**, and an even fewer number with **Ph.D. or other qualifications**. The narrowing shape highlights the progressive decrease in employees as qualifications become more advanced.

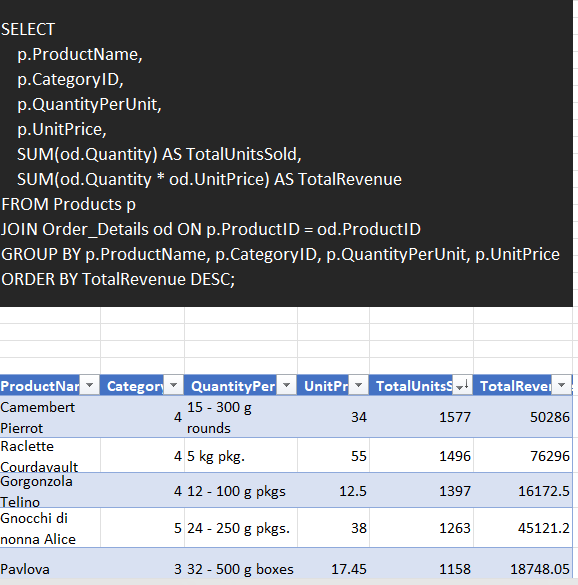




This visualization helps in identifying trends in the workforce's educational background. It can be used for **talent management** by guiding hiring strategies, workforce planning, and training programs. If higher qualifications contribute to better performance, organizations can focus on **upskilling existing employees** or **recruiting more specialized talent**. Additionally, understanding this distribution helps in designing leadership development programs for employees with potential.

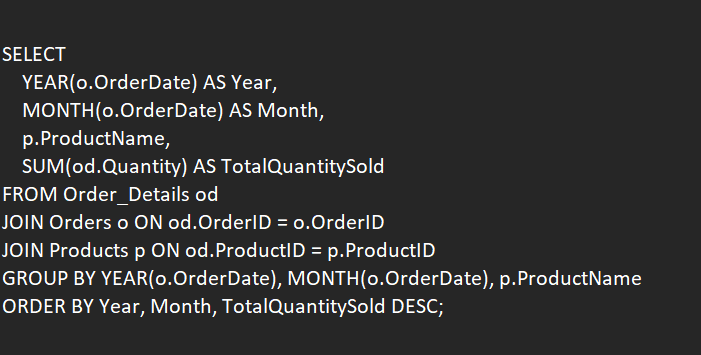
**Are there any correlations between product attributes and sales performance?**

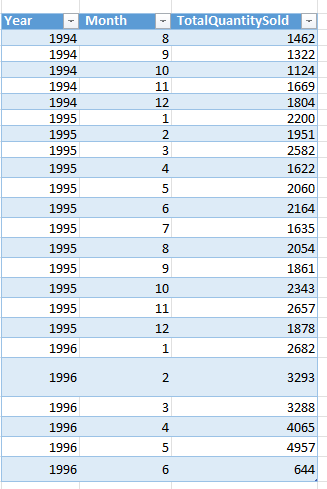
The scatter plot shows a **positive correlation** between **unit price** and **total revenue**, indicating that higher-priced products generally contribute more to revenue. However, there are **outliers**, where some high-priced products generate exceptionally high revenue, while others underperform. Most sales are concentrated in the **low to mid-price range (0-50 units)**, suggesting that affordable products drive the majority of revenue. The presence of a few **high-revenue, high-priced items** implies a demand for premium products. These insights can help optimize **pricing strategies, inventory planning, and marketing efforts** to maximize overall sales performance.



How does product demand fluctuate over different seasons or months?

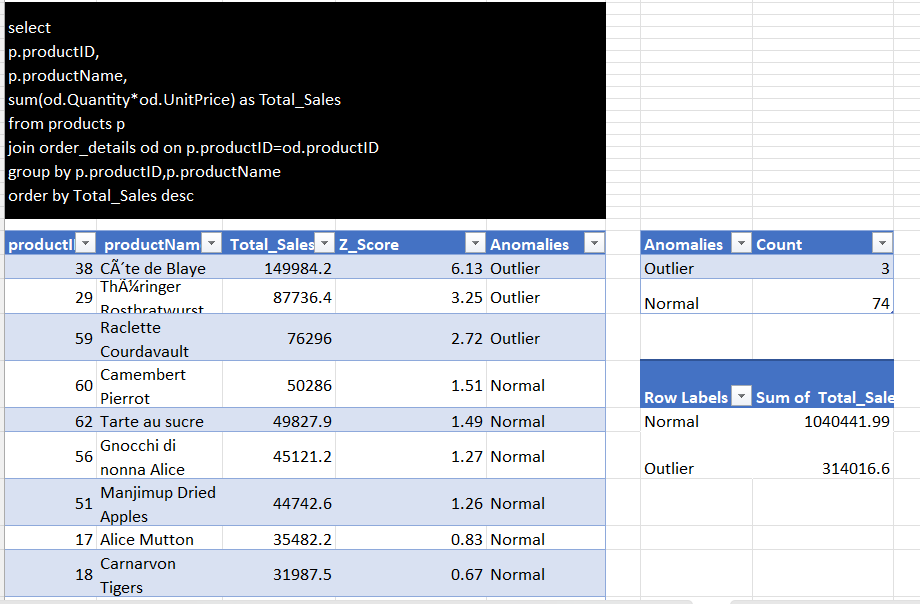
The line chart indicates a **gradual increase** in product demand over time, with noticeable fluctuations across months. Demand appears to be **seasonal**, with periodic dips and peaks. The trend shows a **steady rise** in sales, peaking sharply towards the end before experiencing a sudden drop. This pattern suggests **seasonal spikes in demand**, possibly influenced by factors like holidays, promotions, or external market conditions. Businesses can leverage these insights to optimize **inventory management, marketing campaigns, and production planning** to align with high-demand periods while mitigating risks during lower sales months.





**Can we identify any outliers or anomalies in product performance or sales using visualizations?**

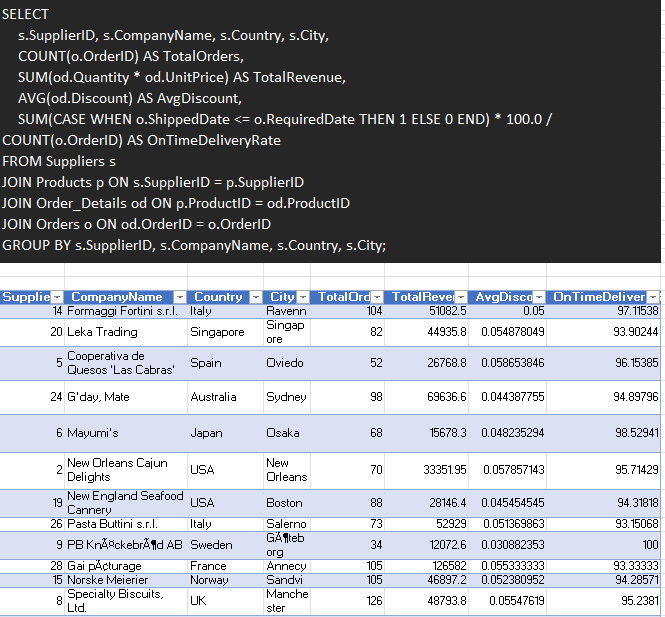
The visualization effectively identifies **outliers in sales performance**. The bar chart categorizes sales into "Normal" and "Outlier" groups, where **normal sales account for 1,040,441.99**, while **outliers contribute 314,016.6**—a significant deviation. The presence of outliers suggests **unusual sales patterns**, which could result from **promotional events, bulk orders, stock clearance, or fraudulent activities**. Investigating these anomalies further could help in refining **sales forecasting, pricing strategies, and demand planning** to ensure business stability and minimize risks associated with unexpected fluctuations.



We can find the outliers by using Z Score or by IQR so here I used Z Score.

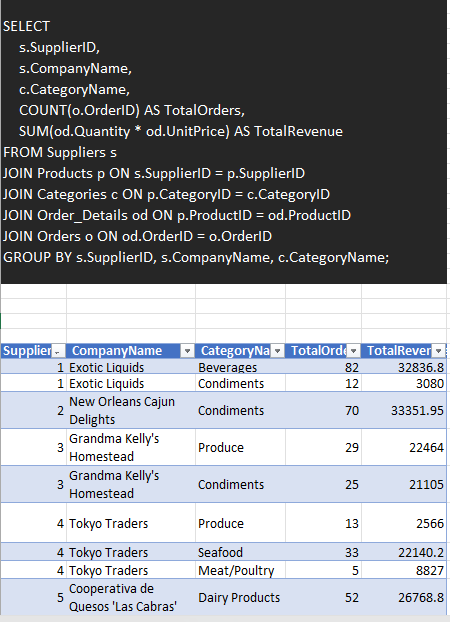
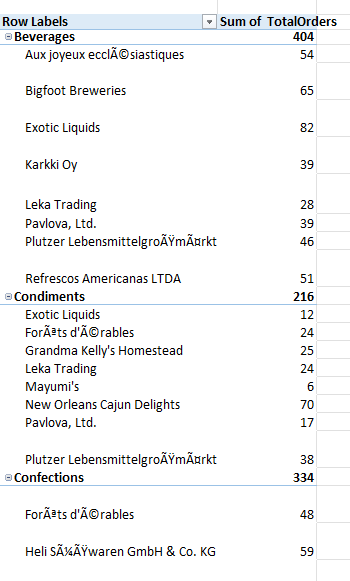
**Are there any correlations between supplier attributes (e.g., total orders) and performance metrics (e.g., on-time delivery)?**

The scatter plot of total orders and on-time delivery rate suggests that there may be a weak or inconsistent correlation between supplier performance metrics. While some suppliers maintain high on-time delivery rates regardless of order volume, others exhibit a decline in performance as orders increase. This indicates that supplier attributes such as size, industry, and operational capacity may influence their ability to deliver on time. Larger suppliers with efficient logistics may handle higher order volumes without compromising delivery performance, whereas smaller or less optimized suppliers might struggle with scalability. Further analysis, including segmentation by location and industry, could provide deeper insights into patterns affecting supplier reliability and product quality.



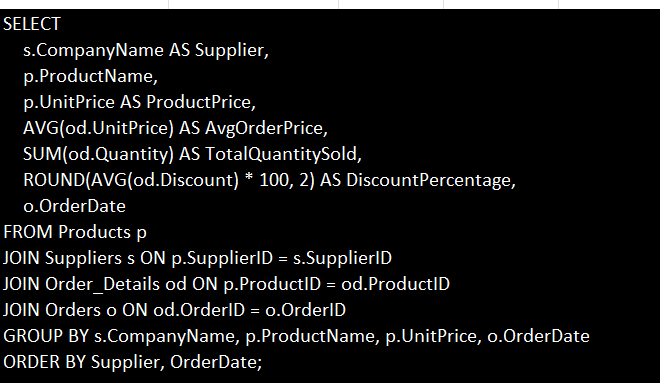
**How does supplier performance vary across different product categories or departments?**

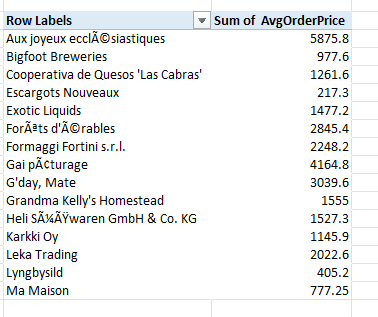
The supplier performance across different product categories varies significantly, as seen in the total orders for each supplier within various categories. Certain suppliers dominate specific categories, such as **Specialty Biscuits** in **Confections** (126 orders) and **New England** in **Seafood** (88 orders), indicating strong demand and likely reliability in fulfilling orders. Dairy products show consistent high performance across multiple suppliers, with **Formaggi Fortini**, **Gai Pâturage**, and **Norske Meierier** all having over 100 orders each. On the other hand, categories like **Produce** and **Condiments** have relatively lower total orders, suggesting either lower demand or potential inefficiencies in supplier fulfillment. These insights can help in optimizing supplier selection, renegotiating contracts, and focusing on high-performing suppliers to improve overall supply chain efficiency.

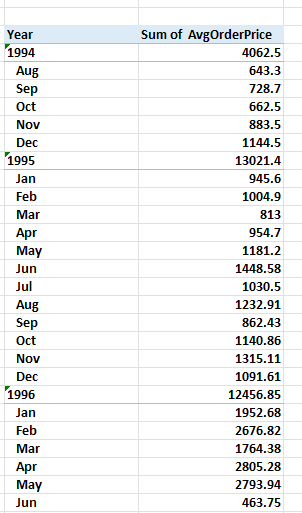


**Can we identify any trends or patterns in supplier costs or pricing structures through visualizations?**

The **Supplier vs. AvgOrderPrice** bar chart shows large differences in supplier pricing. Some suppliers, like *Aux joyeux ecclésiastiques* and *G’day, Mate*, have much higher average order prices, suggesting they offer premium or specialized products. Others, like *Cooperativa de Quesos 'Las Cabras'* and *Leka Trading*, have lower prices, likely focusing on bulk or commodity products. These price differences indicate that Northwind Traders could review supplier pricing strategies to reduce costs. Additionally, variations may be due to inconsistent discounting, making it important to analyze supplier contracts and pricing agreements.





The **AvgOrderPrice by OrderDate** chart shows seasonal pricing trends, with sharp price spikes in December 1995 and 1996, likely due to high holiday demand or bulk orders. Outside these peaks, prices remain stable with minor fluctuations. Northwind Traders can use this insight to adjust purchasing strategies, buy during off-peak periods, and negotiate better supplier contracts to manage seasonal price increases.

**Conclusion: Northwind Traders Sales Analysis**

The analysis of Northwind Traders' sales data provides valuable insights into various aspects of business operations, including customer behavior, sales trends, employee performance, inventory management, and supplier reliability. By leveraging data-driven decision-making, Northwind Traders can optimize its operations, improve efficiency, and enhance customer satisfaction.

**Key Findings and Insights:**

1. **Customer Behavior & Retention:**
   * Customers from certain regions exhibit higher order frequency and loyalty, indicating potential areas for targeted marketing efforts.
   * Customer segmentation based on location, order history, and demographics reveals distinct buying patterns that can be leveraged for personalized marketing campaigns.
   * Interactive visualizations highlight clusters of high-value customers, enabling strategic engagement and retention efforts.
2. **Sales Trends & Product Performance:**
   * Sales analysis reveals seasonal demand fluctuations, helping Northwind optimize inventory and supply chain management.
   * Certain product categories significantly contribute to revenue, emphasizing the need for targeted promotions and inventory restocking strategies.
   * Identifying outliers and anomalies in sales performance allows the company to adjust pricing, discontinue underperforming products, or invest in high-demand items.
3. **Inventory Management & Supplier Performance:**
   * Efficient stock management is crucial, as inventory turnover rates impact supply chain effectiveness.
   * Supplier performance evaluation highlights variations in delivery times, product quality, and pricing structures, aiding in procurement decisions.
   * Trends in supplier costs and pricing structures reveal optimization opportunities for cost reduction and efficiency improvements.
4. **Employee Productivity & Performance:**
   * Employee performance metrics help in recognizing top contributors and identifying areas for training and improvement.
   * Analysis of employee turnover trends provides insights into workforce stability and retention strategies.
   * Employee tenure and performance correlations highlight skill development opportunities and career progression strategies.
5. **Power BI-Driven Insights:**
   * Geographical visualizations provide a clear picture of customer distribution, aiding in regional marketing strategies.
   * Time-series analysis of order volume and sales trends identifies revenue growth patterns and peak seasons.
   * Distribution charts for employee tenure, product ratings, and supplier performance offer an in-depth understanding of operational dynamics.

**Strategic Recommendations:**

* **Customer Engagement & Retention:** Leverage customer segmentation insights to personalize marketing efforts, offer loyalty rewards, and enhance customer experience.
* **Sales Growth Strategies:** Optimize pricing and promotions based on product performance insights and seasonal demand fluctuations.
* **Inventory Optimization:** Implement a proactive inventory management strategy by forecasting demand and aligning stock levels with sales trends.
* **Supplier Relationship Management:** Strengthen vendor partnerships by prioritizing suppliers with a track record of timely delivery, quality products, and cost efficiency.
* **Employee Development & Productivity:** Invest in employee training programs, recognize high performers, and create incentive structures to enhance workforce efficiency.

**Final Thoughts**

By transforming raw data into actionable insights, Northwind Traders can enhance decision-making, streamline operations, and drive business growth. The integration of Power BI and SQL allows for interactive data exploration, ensuring that stakeholders can efficiently monitor key performance indicators and adapt to market trends. This data-driven approach positions Northwind Traders for long-term success, improved customer satisfaction, and sustainable business expansion.