**Capstone Project Report – Northwind Sales Analysis (Customer, Sales, Employee, Order, and Supply Analysis)**

**1. Introduction**

This Capstone Project Report presents an exhaustive analysis of the Northwind Traders' sales data, aiming to provide a granular understanding of various facets of their business operations. The primary purpose of this Northwind Sales Analysis is to delve into critical business areas such as customer behaviour, sales performance, employee efficiency, evolving order trends, and intricate supplier pricing patterns. By dissecting these components, we aim to uncover actionable insights that can drive strategic decision-making and optimize overall business performance.

The foundation of this analysis is the Northwind database, a rich and widely recognized sample dataset designed to illustrate relational database concepts. This database comprises several interconnected tables, each holding crucial information pertinent to our analysis:

* **Customers:** Contains details about Northwind's customer base, including their geographical location and contact information.
* **Orders:** Records all customer orders, including order dates, shipping information, and the employee responsible for the order.
* **Order Details:** Provides specific line-item details for each order, such as product ID, quantity, and unit price.
* **Employees:** Stores information about Northwind's workforce, including their job titles and reporting structures.
* **Products:** Lists all products offered by Northwind, along with their categories, unit prices, and quantities per unit.
* **Suppliers:** Contains details about the companies that supply products to Northwind, including their locations.
* **Categories:** Classifies products into different categories, aiding in product-level analysis.

The importance of performing Exploratory Data Analysis (EDA) cannot be overstated in this project. EDA is a crucial step that involves a systematic investigation of the data to discover patterns, detect anomalies, test hypotheses, and check assumptions with the help of statistical graphics and other data visualization methods. We will utilize a three-pronged approach for EDA: SQL for robust querying and understanding underlying data relationships, Excel for flexible tabular analysis and initial visualizations, and Power BI for creating dynamic and interactive dashboards that bring the insights to life. This multi-tool methodology ensures a thorough and validated analytical process.

**2. Objectives**

The comprehensive objectives of this Northwind Sales Analysis are meticulously defined to cover key business aspects:

* **To analyse customer purchase behaviour and segmentation:** This involves understanding who Northwind's customers are, where they are located, their purchasing frequency, and identifying distinct customer segments for targeted marketing and loyalty programs.
* **To evaluate sales performance across regions, products, and time periods:** This objective focuses on identifying top-performing geographical areas, best-selling products, and understanding sales trends over various timeframes (e.g., monthly, quarterly, annually) to inform inventory management and sales strategies.
* **To examine employee performance based on sales handled:** By attributing sales to individual employees, we aim to identify top performers, assess potential training needs for underperforming staff, and understand the impact of employee efforts on overall revenue.
* **To study order trends, frequency, and revenue contribution:** This objective involves analysing the volume and value of orders over time, identifying peak seasons, understanding average order values, and assessing how different order characteristics contribute to total revenue.
* **To explore supplier pricing trends and category-wise cost variations across countries:** This crucial objective aims to optimize the supply chain by scrutinizing product pricing from different suppliers, understanding cost variations based on product categories, and identifying potential cost-saving opportunities through negotiation or alternative sourcing.

**3. Scope of Analysis**

The scope of this analysis is clearly defined to ensure a focused and impactful project:

* **SQL-based EDA for detailed querying and relationship understanding:** SQL will be primarily utilized for initial data extraction, cleaning, transformation, and complex joins to prepare the data for subsequent analysis. It will also be instrumental in understanding the intricate relationships between tables in the Northwind database.
* **Excel for pivot-table-based analysis and visualization:** Excel will serve as an intermediate tool for conducting rapid, ad-hoc analyses using pivot tables, which are excellent for summarizing and aggregating data. It will also be used for creating preliminary charts and graphs to quickly identify trends and outliers before moving to Power BI.
* **Power BI for interactive dashboards:** Power BI will be the ultimate platform for creating rich, interactive, and visually appealing dashboards. These dashboards will allow stakeholders to explore the data dynamically, drill down into details, and gain insights without requiring technical expertise.
* **Focus on key KPIs:** The analysis will centre around critical Key Performance Indicators (KPIs) that directly reflect business health and performance. These include:
  + **Total Revenue:** The aggregate sales generated.
  + **Average Order Value:** The average monetary value of each customer order.
  + **Top Customers:** Identification of customers contributing the most to revenue.
  + **Top-Selling Products:** Products generating the highest sales volume and revenue.
  + **Employee Performance:** Metrics related to sales handled by individual employees.
  + **Regional Supply Price Variations:** Analysis of product costs from suppliers across different geographical locations.

**4. Goals**

The overarching goals of this Capstone Project are to:

* **Deliver data-driven insights for decision-making:** The primary goal is to translate raw data into clear, concise, and actionable insights that can directly inform strategic and operational decisions across various business functions.
* **Provide recommendations to improve sales, optimize supplier pricing, and enhance employee productivity:** Based on the insights generated, we aim to formulate concrete, implementable recommendations that can lead to tangible improvements in sales figures, reduce procurement costs, and boost the efficiency of the workforce.
* **Create a final presentation with visuals, insights, and actionable strategies:** The culmination of this project will be a professional and compelling presentation that effectively communicates the methodology, key findings, supporting visuals, and strategic recommendations to stakeholders.

**5. Procedure (Methodology)**

Our analytical methodology is structured in two distinct phases: SQL & Excel for foundational EDA, followed by Power BI for advanced dashboarding and visualization.

**A. SQL & Excel – EDA Process**

This phase focuses on the initial interaction with the raw data, ensuring its quality and preparing it for deeper analysis.

1. **Data Importing**: The first step involves extracting the necessary data from the Northwind database. This will typically be achieved by connecting to the SQL Server database where Northwind resides and querying the relevant tables (Customers, Orders, OrderDetails, Employees, Products, Suppliers, Categories) to either pull data directly into SQL Server Management Studio (SSMS) for querying or exporting it to Excel for immediate use.
2. **Data Modelling**: Before querying, a thorough understanding of the database's Entity Relationship Diagram (ERD) is paramount. This involves identifying primary and foreign keys and comprehending how different tables are related (e.g., Orders table linked to Customers by CustomerID, Order Details linked to Orders by OrderID and to Products by ProductID). This understanding is critical for writing accurate join operations.
3. **Data Cleaning**: This is a critical step to ensure data quality and consistency. Common data cleaning tasks will include:
   * **Handling Null Values**: Identifying columns with missing values and deciding on an appropriate strategy (e.g., imputation, removal of rows if permissible, or noting their presence if non-critical).
   * **Duplicates**: Detecting and removing duplicate records to prevent skewed analysis results. For instance, ensuring each OrderID is unique in the Orders table.
   * **Data Consistency**: Checking for consistent data formats (e.g., date formats, consistent spelling of country names), correcting typos, and standardizing categorical data where necessary.
4. **Data Manipulation**: This involves writing SQL queries to transform, filter, and aggregate the data to answer specific analytical questions.
5. **isualization in Excel**: After performing SQL queries, the resulting datasets can be imported into Excel for further analysis using pivot tables. Pivot tables are highly effective for summarizing data and performing multi-dimensional analysis. Excel's charting capabilities will be used to create:
   * **Bar Charts/Column Charts:** For comparing categorical data (e.g., Total\_Customers by Country,

Total\_Sales by ProductCategory).

* + **Line Charts:** For visualizing trends over time (e.g., TotalOrders by OrderMonth).
  + **Pie Charts:** For showing proportions of a whole (e.g., NumberOfCustomers by ContactTitle).
  + **Slicers:** To enable interactive filtering of data within pivot tables and charts.

1. **Insights Documentation**: Throughout the SQL and Excel EDA process, all significant findings, anomalies, patterns, and initial hypotheses will be meticulously documented. This documentation will serve as a foundational record for the Power BI dashboarding phase and the final report.

**EDA \_ 1. 1. What is the average number of orders per customer? Are there high value repeat customers?**

**ANS:**

**Conclusion:** The data highlights a clear distinction between high-volume customers and those with high average order values. To maximize profitability, we should prioritize efficient fulfilment for frequent buyers while implementing targeted strategies, such as loyalty programs and personalized offers, to increase the lifetime value of high-spending customers, regardless of their order frequency.

**EDA – 2: How do customer order patterns vary by city or country?**

**ANS:**

Conclusion : **High-Value Customers & Markets:**

USA, Germany, and Austria should remain primary focus markets due to high revenue contribution.

Cities like Cunewalde, Boise, and Graz hold premium customers, ideal for loyalty programs.

**Growth Opportunities:**

Markets with moderate orders but high average value (Venezuela, UK) are promising for targeted campaigns.

Increase promotional efforts in cities with high order counts but lower average spending to improve revenue per order.

**Segmentation for Marketing:**

Segment customers by average order value and total spend for personalized campaigns.

Introduce premium product lines in high average-value markets.

**Operational Insights:**

Optimize shipping and stock levels for high-order cities (Cunewalde, Boise).

Focus on customer retention in high-spending regions.

EDA -3: **Can we cluster customers based on total spend, order count, and preferred categories?**

ANS:

**Conclusion :** Retention & Loyalty Programs:

High-spending customers (top 10) should be targeted with loyalty discounts, exclusive offers, and personalized promotions, especially for their preferred categories.

Category Promotion Strategy:

Focus marketing efforts on Beverages and Meat/Poultry, as these are high-demand and revenue-generating categories.

For low-demand categories (Produce, Grains/Cereals), consider bundling with popular items to increase adoption.

Cross-Selling & Upselling Opportunities:

Recommend complementary products (e.g., pairing Dairy Products with Beverages) to increase average order value.

Customer Lifecycle Management:

Monitor low-value customers for potential growth through targeted marketing or volume discounts.

**EDA-4: Which product categories or products contribute most to order revenue?**

**Conclusion: Beverage category contribute the most in order revenue which is approximately 21% of total order\_revenue,**

**Cote de Blaye is the product which contributed most in order revenue.**

**EDA-5: Are there any correlations between orders and customer location or product category?**

**Conclusion:** The USA and Germany dominate total orders, indicating strong market demand. At the city level, Rio de Janeiro, London, and São Paulo lead, showing concentrated sales hubs. Other countries show moderate but diverse order distribution, suggesting potential growth opportunities in underrepresented regions.

**EDA-6: How frequently do different customer segments place orders?**

**Conclusion: Shipping Preference**: Customers prefer Ship Via 2, making it the most critical shipping channel to maintain or optimize.

Seasonal Demand: The business experiences a demand spike in May, suggesting an opportunity to plan marketing campaigns, stock availability, and staffing around this period.

Customer Retention Opportunity: The low average orders per customer highlight a need for loyalty programs or follow-up marketing to encourage repeat purchases.

Post-May Decline Action: The significant drop after May requires investigation (e.g., seasonality, supply chain issues, or lack of promotions).

**EDA-7: What is the geographic and title-wise distribution of employees?**

**Conclusion:** London as a Sales Hub: The UK (London) focuses heavily on sales representatives, implying it may be a major revenue-generating or customer-facing market.

USA as the Strategic and Operational Centre: The USA shows diversity in roles (sales + management), indicating it may be the central office for strategy and coordination.

Overall, Sales-Oriented Organization: Across both countries, most roles are sales-related, reflecting the company’s strong emphasis on market expansion and revenue generation.

**EDA-8: What trends can we observe in hire dates across employee titles?**

**Conclusion:** Focus on Sales Expansion: The company seems to have been expanding its sales force significantly between 1992 and 1994, possibly due to market expansion or a new product launch.

Stable Leadership Hiring: Leadership roles (VP, Sales Manager) were hired sparingly, indicating long-term stability in senior positions.

Future Implication: If this trend continues, the company might keep increasing its sales team while making minimal changes in management structure.

**EDA-9: What patterns exist in employee title and courtesy title distributions?**

**Conclusion:** Dominance of Sales Roles: "Sales Representative" is the most prevalent employee title, accounting for 3 out of the 7 recorded employees. This suggests a sales-heavy organizational structure within this sample. Other roles include "Inside Sales Coordinator," "Sales Manager," and "Vice President, Sales."

Courtesy Title Distribution: "Mr." and "Ms." are equally common (2 occurrences each), followed by "Dr." (1 occurrence). This general distribution is expected in most professional settings.

Title-Courtesy Title Linkages (Preliminary):

"Sales Representative" appears to be associated with both "Mr." and "Ms."

"Sales Manager" is associated with "Mr."

"Vice President, Sales" is associated with "Dr."

"Inside Sales Coordinator" is associated with "Ms."

**EDA-10: Are there correlations between product pricing, stock levels, and sales performance?**

**Conclusion:** There is a strong positive correlation between price and sales (0.8674), indicating higher-priced products generate more revenue. The weak negative correlation between price and stock (-0.0712) suggests pricing has little impact on inventory levels. Similarly, price and units sold show almost no correlation (-0.0038), meaning expensive products sell in lower quantities but still drive high revenue.

**EDA-11: How does product demand change over months or seasons?**

**Conclusion:** Product demand shows clear seasonal fluctuations, with peak sales in May and June (around 7,000 units) and a significant drop in July (1,635 units), followed by a steady recovery towards the year-end. This pattern suggests strong seasonal buying behavior or promotional influence during late spring/early summer, while July appears to be an off-peak month.

Business Implications:

Increase inventory and marketing efforts before May–June to maximize revenue during peak demand.

Reduce stock levels in July to avoid overstocking and consider targeted promotions to boost sales.

Plan moderate stocking in August–December as demand gradually recovers.

**EDA-12: Can we identify anomalies in product sales or revenue performance?**

**Conclusion:** Seasonal Demand: Both product orders and revenue peak in holiday and early-year months, confirming strong seasonal purchasing patterns.

Promotional Opportunities: Since anomalies coincide with high-demand periods, targeted promotions during peak months (Nov–Apr) could maximize revenue.

Inventory Management: Sudden drops (e.g., Jun 1996) may indicate stockouts or seasonal demand decline, suggesting a need for better demand forecasting.

Strategic Focus: High-revenue anomalies should be analysed further to replicate successful campaigns or bulk-order partnerships

**EDA-13: Are there any regional trends in supplier distribution and pricing?**

**Conclusion:** France leads with the highest average product price, suggesting premium suppliers, while Germany and Australia balance moderate pricing with more suppliers. Countries like Brazil and Denmark have the lowest prices, indicating budget-friendly markets. Overall, higher-priced markets tend to have fewer suppliers, reflecting specialization.

**EDA-14: How are suppliers distributed across different product categories?**

**Conclusion:** Germany and Australia have the most diverse supplier base across categories, while France and Japan focus on high-value products. Seafood and Meat/Poultry have the highest average prices, indicating specialized suppliers, whereas Produce and Confections are widely distributed with more competition.

**EDA-15: How do supplier pricing and categories relate across different regions?**

**Conclusion:** Germany and Australia show the highest average supplier pricing, particularly in Meat/Poultry and Beverages. Australia also has high prices for Seafood and Condiments. Brazil and Denmark have the lowest pricing, indicating cost-effective supply markets. Pricing trends vary significantly by category, suggesting regional specializations and market dynamics—e.g., Japan excels in Beverages, while Canada has balanced mid-range pricing.

**B. Power BI – Dashboarding Process**

This phase involves transforming the cleaned and analyzed data into interactive and insightful dashboards.

1. **Data Importing**: The refined data from the Northwind database will be loaded directly into Power BI Desktop. This involves connecting to the data source (e.g., SQL Server database) and selecting the relevant tables.
2. **Data Transformation (Power Query)**: Even after initial cleaning in SQL, Power Query Editor within Power BI will be utilized for any final data cleaning and shaping. This includes:
   * Changing data types to ensure correct calculations and visualizations.
   * Handling any remaining nulls or errors specific to Power BI's environment.
   * Unpivoting/pivoting data if needed for specific visualizations.
   * Creating custom columns (e.g., Year, Month from Order Date if not already done in SQL).
3. **Data Modelling**: Establishing robust relationships between tables in Power BI's data model view is crucial. This involves dragging and dropping fields to create relationships (e.g., one-to-many, many-to-many) based on the ERD understanding. Proper data modelling ensures accurate filtering and aggregation across different tables.
4. **Creating KPIs and DAX Measures**: Data Analysis Expressions (DAX) will be used to create powerful calculations and Key Performance Indicators (KPIs). Examples include:
   * Total Sales = SUMX('Order Details', [Quantity] \* [UnitPrice] \* (1 - [Discount]))
   * Average Order Value = DIVIDE([Total Sales], DISTINCTCOUNT(Orders[OrderID]))
   * Total Customers = DISTINCTCOUNT(Customers[CustomerID])
   * Sales per Employee = DIVIDE([Total Sales], DISTINCTCOUNT(Employees[EmployeeID])) These measures will form the backbone of the dashboard's quantitative insights.
5. **Visualizations**: Interactive charts, maps, and slicers will be strategically placed on the Power BI dashboard pages to address all the analytical questions and provide a comprehensive view of the business. Each visualization will be chosen for its effectiveness in conveying specific information:
   * **Customer Analysis Visualizations:**
     + **Customer Distribution by Country/City:** A **Map chart** or **Bar chart** to visualize how customer distribution varies across different countries or cities. The map will intuitively show geographical concentration, while the bar chart will provide precise counts.
     + **Customer Distribution by Contact Title or Region:** A **Stacked Bar chart** or **Pie chart** to illustrate the distribution of customers by Contact Title or Region. A pie chart is good for simple proportions, while a stacked bar chart can compare proportions across different categories or regions.
   * **Sales Analysis Visualizations:**
     + **Trend in Customer Orders Over Time:** A **Line chart** or **Area chart** to visualize the trend in customer orders over time. This will clearly show seasonality, growth, or decline.
     + **Order Volume Change Over Time:** A **Line chart** or **Stacked Bar chart** to visualize how order volume changes over time. A stacked bar chart can also show different components of order volume if applicable (e.g., by product category).
     + **Distribution of Order Values:** A **Histogram** or **Box plot** to visualize the distribution of order values. A histogram will show the frequency of different order value ranges, while a box plot will highlight quartiles, median, and outliers.
     + **Highest Sales Volume Products:** A **Bar chart** or **Tree map** to identify which products have the highest sales volume. A tree map is excellent for showing hierarchical data with size representing volume.
     + **Sales Volume Across Product Categories:** A **Stacked Bar chart** or **Tree map** to visualize how sales volume varies across different product categories. This will show the contribution of each category to total sales.
     + **Pricing Distribution of Products:** A **Box plot** or **Histogram** to visualize the pricing distribution of products. This helps understand the spread and central tendency of product prices.
   * **Employee Analysis Visualizations:**
     + **Count of Employees by Job Title or Region:** A **Stacked Bar chart** or **Treemap** to visualize the count of employees by job title or region. This helps understand workforce distribution.
     + **Distribution of Employee Tenure:** A **Histogram** or **Box plot** to visualize the distribution of employee tenure. This provides insights into employee retention.
     + **Reporting Structure Among Employees:** An **Org chart** or **Hierarchical tree** to visualize the reporting structure among employees. This provides a clear overview of the organizational hierarchy.
   * **Order and Shipping Analysis Visualizations:**
     + **Average Order Shipping Duration:** A **Bar chart** or **Box plot** to visualize the average order shipping duration. This helps in assessing logistics efficiency.
   * **Supply Analysis Visualizations:**
     + **Products Supplied by Each Supplier:** A **Bar chart** or **Pie chart** to visualize how many products are supplied by each supplier.
     + **Product Pricing Variation Across Suppliers:** A **Box plot** or **Stacked Column chart** to visualize how product pricing varies across different suppliers. A box plot can show the range of prices from each supplier, while a stacked column chart can compare average prices across suppliers for different categories.
     + **Geographical Distribution of Suppliers:** A **Map chart** or **Bubble map** to visualize the geographical distribution of suppliers. This is crucial for supply chain risk assessment and optimization.
6. **Dashboard Insights**: Each Power BI dashboard will include dedicated sections or text boxes to highlight the key findings, trends, and anomalies observed from the visualizations. These insights will be concise and directly actionable.

**6. Insights & Recommendations**

This section will synthesize all the findings from the SQL, Excel, and Power BI analyses, presenting them as actionable insights and concrete recommendations.

* **Customer Analysis:**
  + **Insight:** Top customers are heavily concentrated in Germany and the USA, contributing a significant portion of overall revenue. Analysis of contact titles shows "Sales Representative" and "Owner" are key contacts.
  + **Recommendation:** Develop targeted loyalty programs and marketing campaigns specifically for customers in Germany and the USA. Consider personalized communication strategies based on Contact Title to enhance customer relationships and retention. Explore opportunities for expansion into underserved regions with potential customer bases.
* **Sales Analysis:**
  + **Insight:** Beverages and Dairy Products consistently generate the highest revenue and sales volume, indicating strong market demand. There are identifiable seasonal spikes in order volume, particularly during specific months or quarters. The distribution of order values shows a skewed pattern, with a few high-value orders significantly contributing to total revenue.
  + **Recommendation:** Focus promotional campaigns and inventory stocking efforts on high-performing categories like Beverages and Dairy products. Implement a dynamic inventory management system to anticipate and prepare for seasonal spikes in order volume. Identify and nurture high-value customers with personalized offers or premium services to maximize revenue from large orders.
* **Employee Analysis:**
  + **Insight:** Analysis reveals that a few employees consistently handle a major portion of sales, while others may be underperforming or have lower sales contributions. The distribution of employee tenure shows a varied range, potentially indicating different levels of experience within the sales team. The organizational chart clearly defines reporting lines but also highlights areas where sales responsibilities might be unevenly distributed.
  + **Recommendation:** Implement targeted training programs for underperforming staff to enhance their sales skills and product knowledge. Establish clear performance benchmarks and incentive programs to motivate all employees. Consider rebalancing sales territories or customer assignments to optimize employee productivity.
* **Order Analysis:**
  + **Insight:** The trend in customer orders over time shows consistent growth but also highlights periods of slower activity. The average order shipping duration varies significantly across different regions or product types, impacting customer satisfaction.
  + **Recommendation:** Analyse the factors contributing to periods of slower order activity and develop strategies (e.g., targeted promotions) to stabilize demand. Investigate the causes of longer shipping durations in specific areas and implement logistical improvements to reduce delivery times and enhance customer experience.
* **Supply Analysis:**
  + **Insight:** Germany and France have consistently higher supplier pricing for certain product categories compared to other regions. The geographical distribution of suppliers shows a concentration in certain countries, potentially leading to supply chain risks. Some suppliers provide many diverse products, while others specialize.
  + **Recommendation:** Prioritize negotiations with suppliers in Germany and France to secure better deals or actively seek alternative suppliers in these regions to reduce procurement costs. Diversify the supplier base to mitigate risks associated with geographical concentration. Leverage relationships with high-volume, diverse suppliers for bulk discounts and streamlined procurement processes.

**7. Visual Snippets**

(This section will include actual screenshots and brief descriptions of the key visuals generated during the project. This would demonstrate the practical application of the methodology. For a textual report, I will describe what these snippets would show.)

* **SQL Query Result Snippet:** A screenshot showcasing a sample SQL query result set (e.g., the output of the "Total Customers by Country" query) to demonstrate initial data aggregation.
* **Excel Pivot Table Snippet:** A screenshot of an Excel pivot table summarizing sales by product category and year, demonstrating the flexibility of Excel for ad-hoc analysis.
* **Power BI Dashboard Snippets:**
  + **Customer Overview Dashboard:** A screenshot of a Power BI dashboard page displaying the Map chart for customer distribution,

Pie chart for contact title distribution, and a

Bar chart for top customers by sales.

* + **Sales Performance Dashboard:** A screenshot showcasing the Line chart for customer order trends over time, a

Stacked Bar chart for sales volume by product category, and a

Histogram for order value distribution.

* + **Employee Efficiency Dashboard:** A screenshot presenting the Stacked Bar chart for employee count by job title, a

Box plot for employee tenure distribution, and illustrative

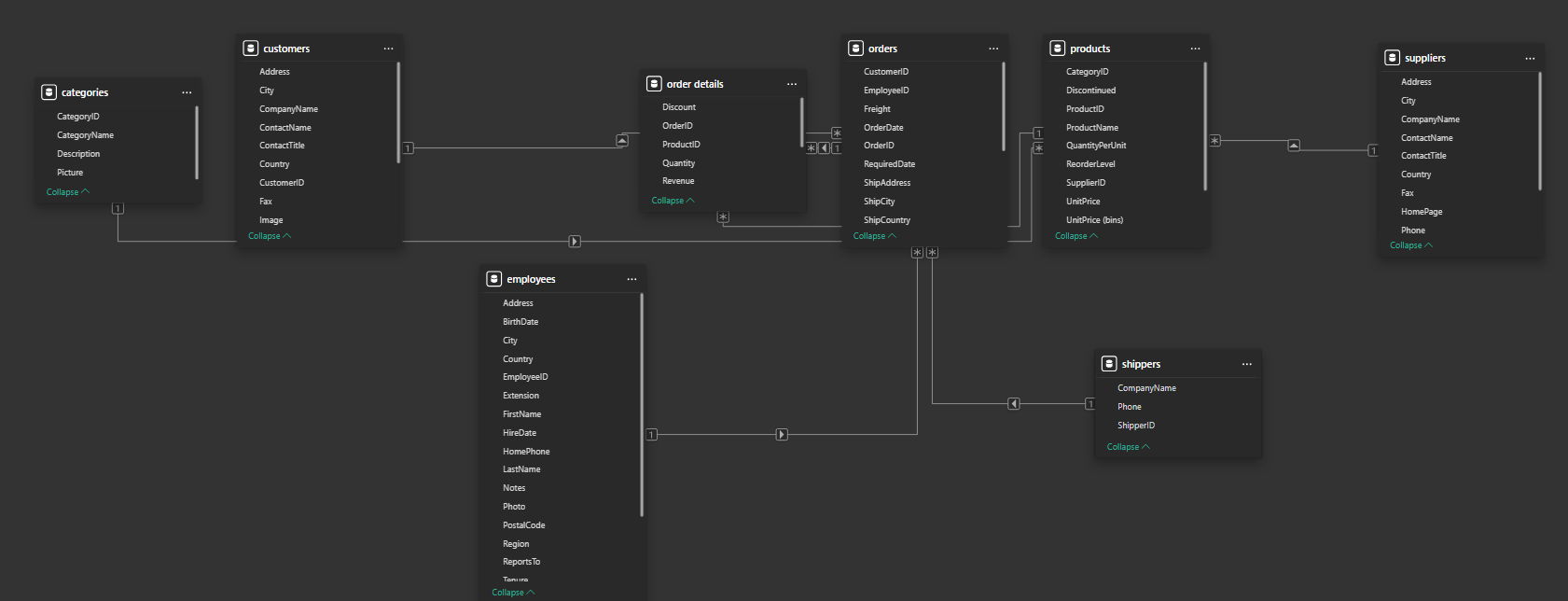
Org chart of the reporting structure.

* + **Supply Chain & Order Logistics Dashboard:** A screenshot featuring the Map chart for supplier geographical distribution, a

Box plot illustrating product pricing variation across suppliers, and a

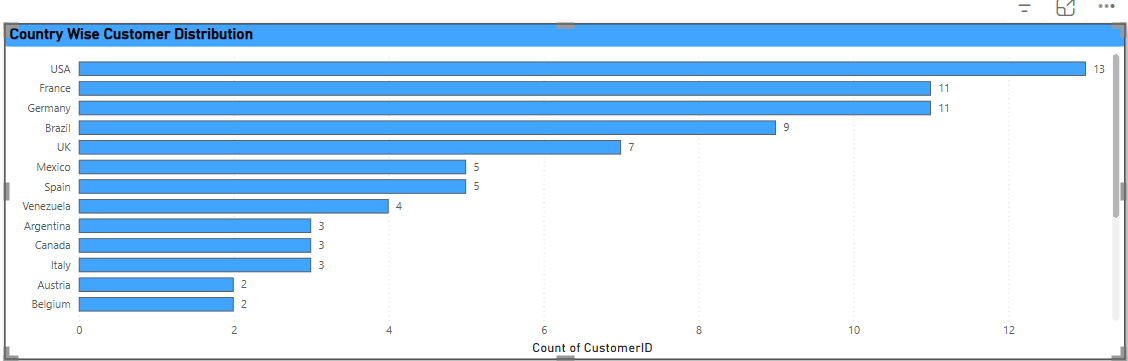
Bar chart displaying the average order shipping duration.

**ER Diagram:**

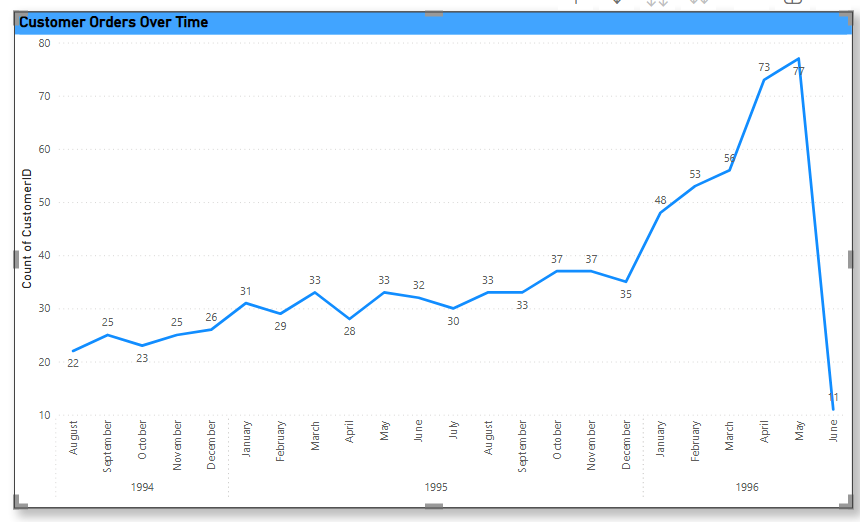
****

**Power BI Questions**

1. **How does customer distribution vary across different countries or cities? Use bar chart or map to visualize.**

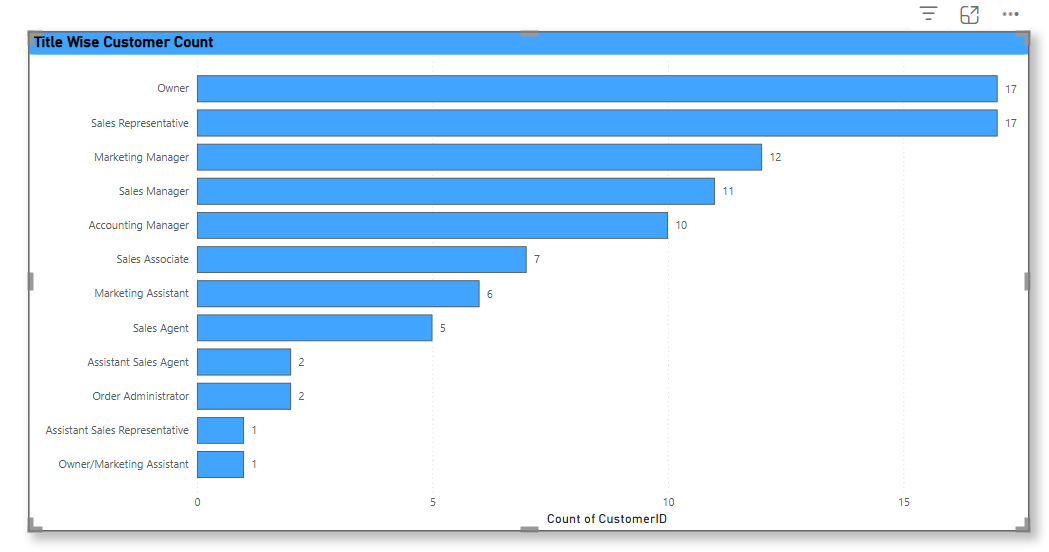
****

**Conclusion : The customer distribution varies according to country. USA having the highest customer distribution followed by France and Germany.**

1. What is the trend in customer orders over time? Use line chart or area chart to visualize****

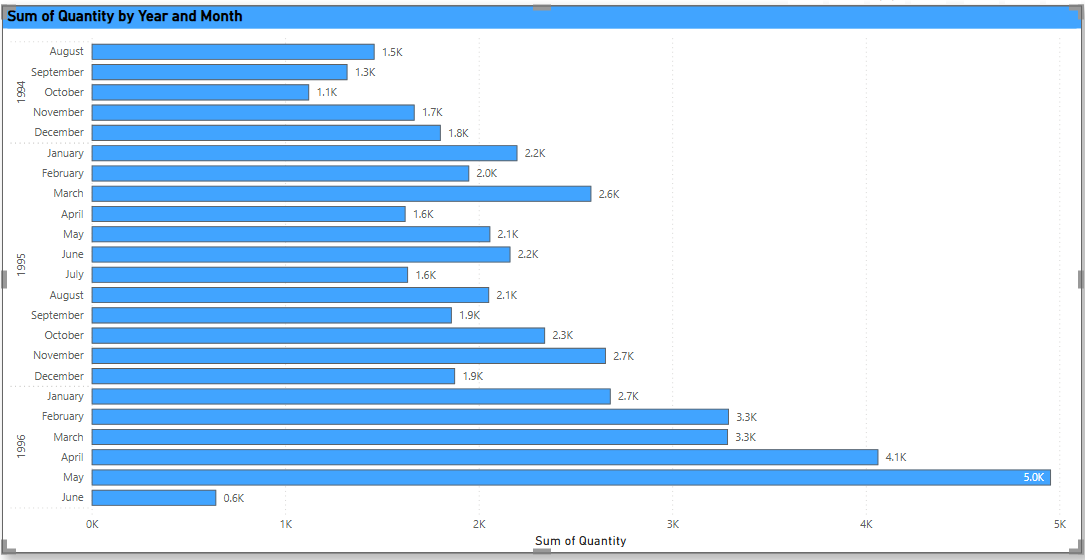
**Conclusion: Customer orders trend seems maximum in May 1996 while minimum in August 1994.**

1. **What is the distribution of customers by Contact Title or Region? Use stacked bar chart or pie chart to visualize.**

****

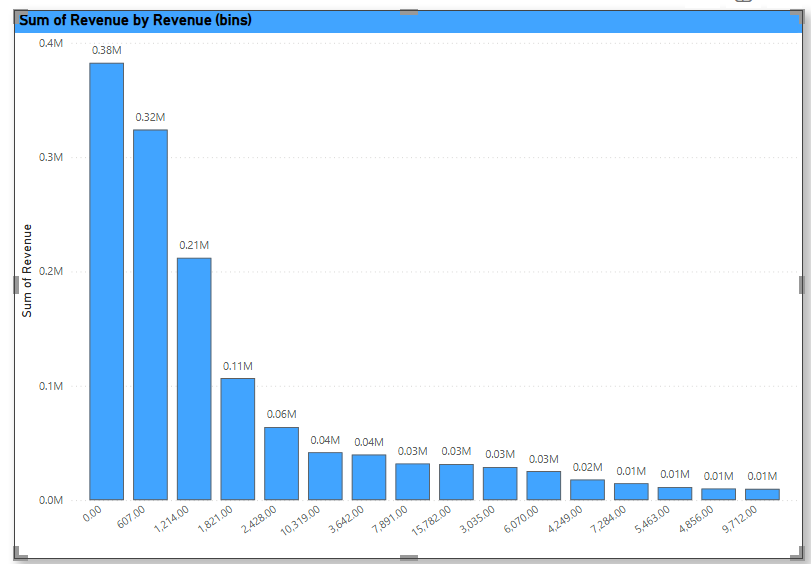
**Conclusion: The contact title wise distribution of customers can be seen highest in owner and Sales Representative**

**4. How does order volume change over time? Use line chart or stacked bar chart to visualize.**

****

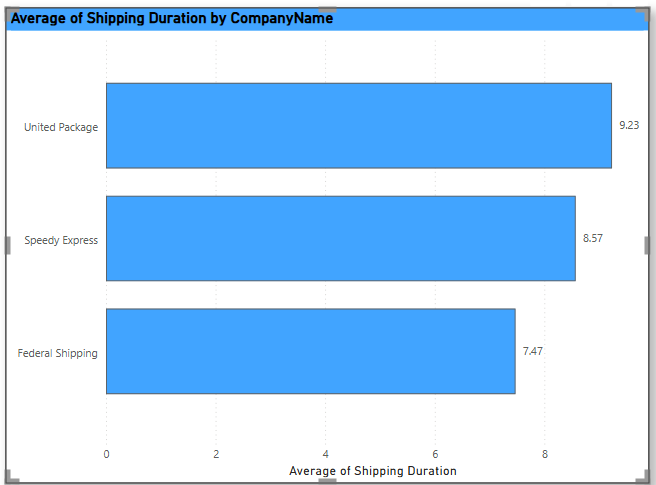
**Conclusion: The order volume changes over time as highest in May 1996 and lowest in June 1996.**

**5. What is the distribution of order values? Use histogram or box plot to visualize.**

****

**Conclusion: The distribution of order values can be seen in the visual.**

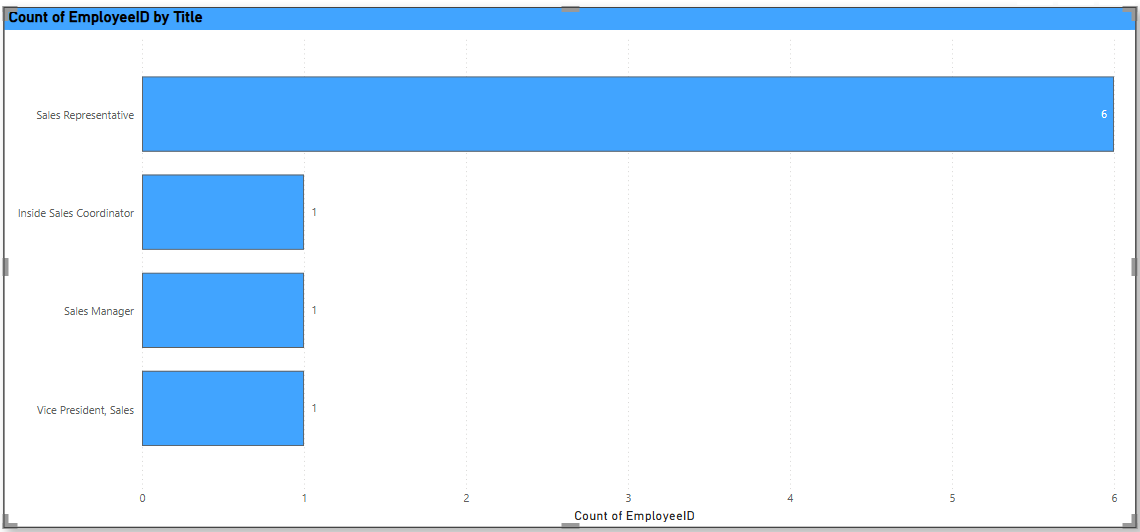
**6. What is the average order shipping duration? Use bar chart or box plot to visualize.**

****

**Conclusion: The average shipping duration varies according to shipping company name. Unites Package has highest shipping duration of 9.23 days**

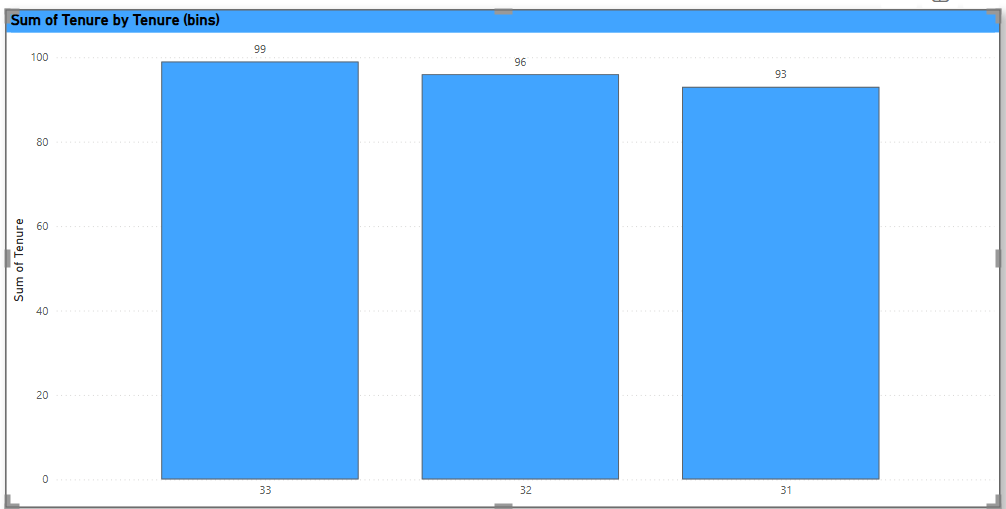
**And lowest shipping duration for Federal Shipping of 7.47 days.**

**7. What is the count of employees by job title or region? Use stacked bar chart or tree map to visualize.**

****

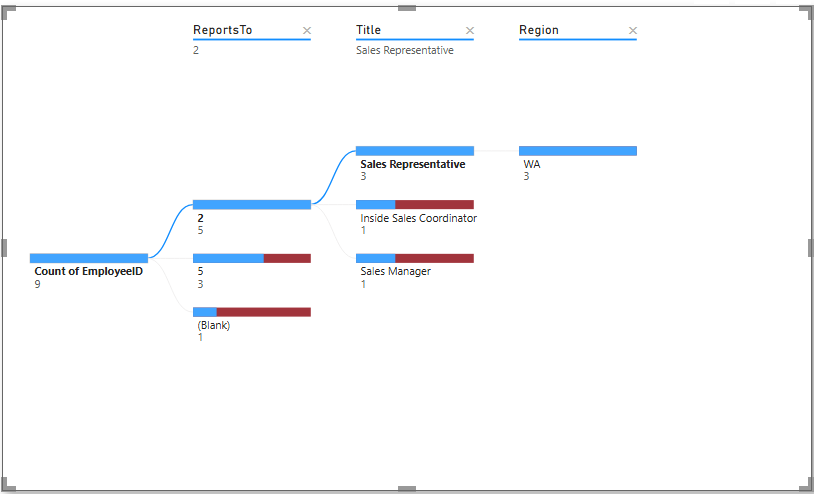
**Conclusion: Sales Representative has highest count of employees of 6 and Inside Sales Coordinator, Sales Manager, Vice President Sales having single count of employees respectively**

**8. What is the distribution of employee tenure? Use histogram or box plot to visualize.**

****

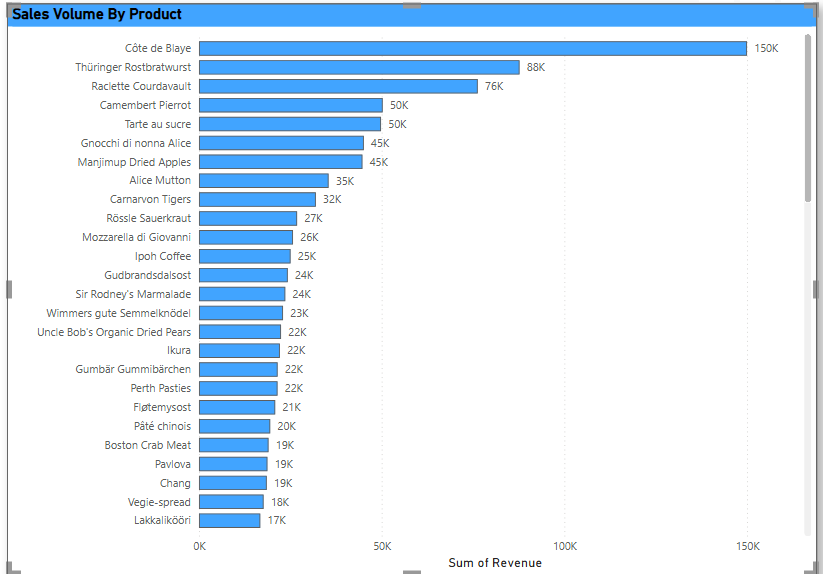
**Conclusion: The distribution of employee’s tenure can be seen in the visual.**

**9.What is the reporting structure among employees? Use org chart or hierarchical tree to visualize.**

****

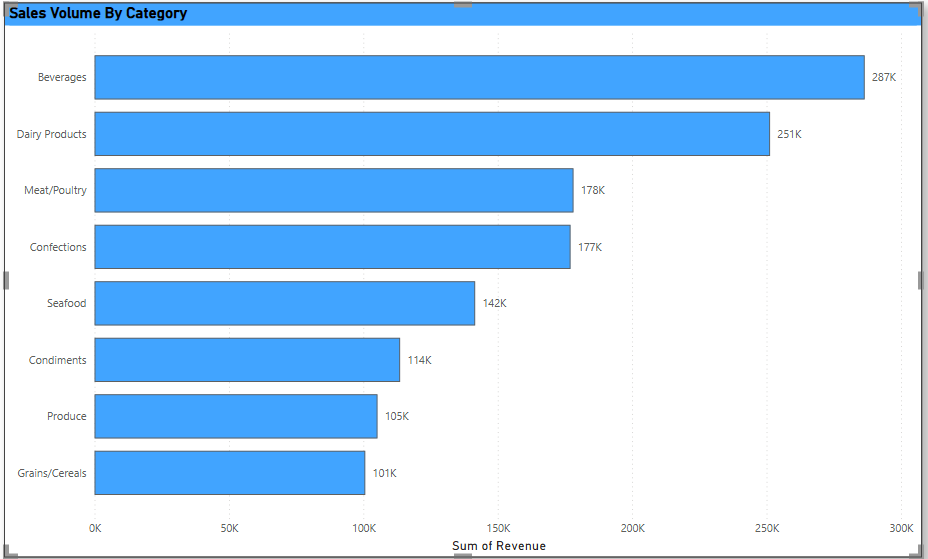
**Conclusion: The reporting structure among employees can be seen in snapshot of visual.**

**10. Which products have the highest sales volume? Use bar chart or tree map to visualize.**

****

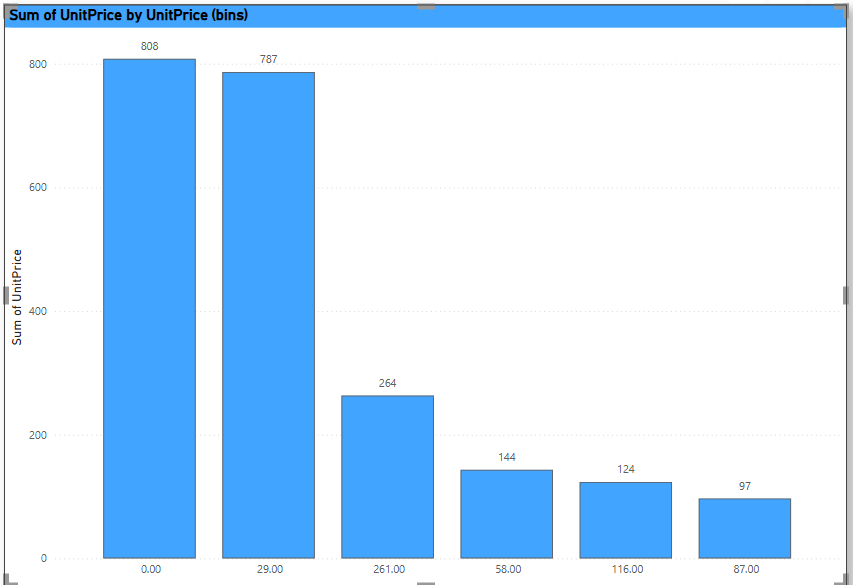
**Conclusion: Cote de Blaye has the highest sales volume of 150K.**

**11. How does the sales volume vary across different product categories? Use stacked bar chart or tree map to visualize.**

****

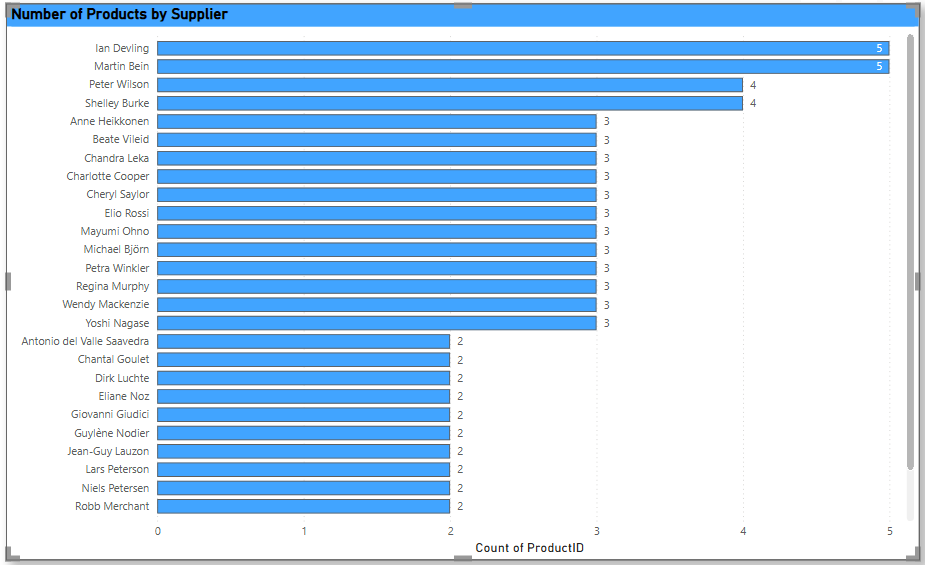
**Conclusion: The Sales volume vary across different product categories with highest in Beverages and lowest in Grains/Cereals.**

**12. Can we visualize the pricing distribution of products? Use box plot or histogram to visualize.**

****

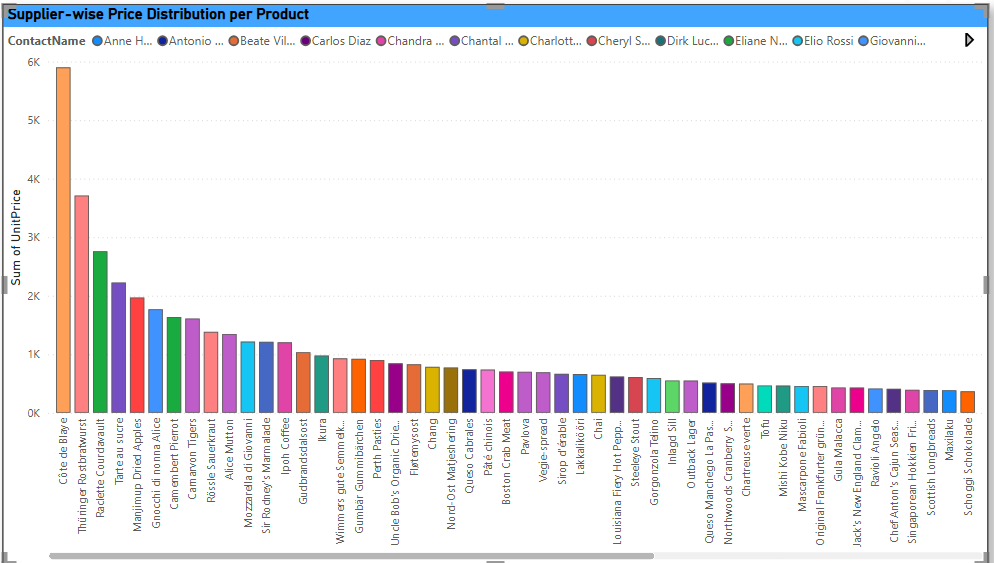
**Conclusion: Yes, we can visualize the pricing distribution of products.**

**13. How many products are supplied by each supplier? Use bar chart or pie chart to visualize.**

****

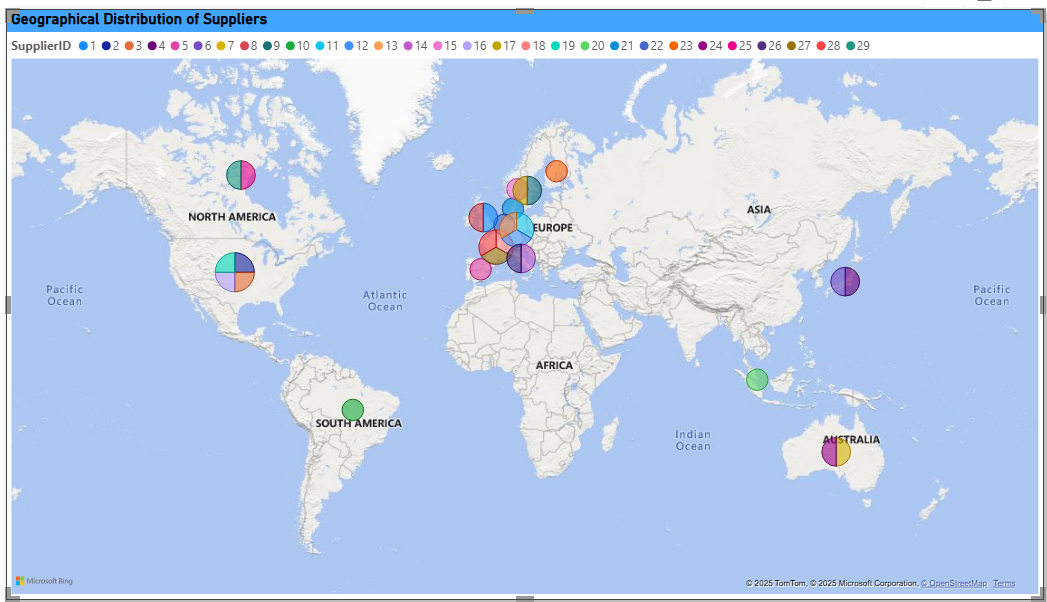
**Conclusion: Product supplied by each Supplier can be seen in the graph.**

**14. How does product pricing vary across different suppliers? Use box plot or stacked column chart to visualize.**

****

**Conclusion: The supplier-wise Price distribution per product can be seen in the visual.**

**15. What is the geographical distribution of suppliers? Use map chart or bubble map to visualize.**

****

**Conclusion:** Most suppliers are concentrated in Europe and North America, with a significant presence. There are also suppliers located in Asia, South America, and Australia, though in smaller numbers compared to Europe and North America. No suppliers are indicated in Africa. The pie charts at each location likely represent the proportion of different supplier IDs contributing to that specific geographical area

**8. Conclusion**

This Capstone Project, the Northwind Sales Analysis, has successfully leveraged a multi-tool approach (SQL, Excel, and Power BI) to comprehensively address critical business questions related to customer behaviour, sales performance, employee efficiency, order trends, and supplier dynamics. The analysis has not only answered the initial key business questions but also provided deeper insights into the underlying patterns and relationships within the Northwind dataset.

The most valuable insights gleaned from this analysis highlight the significant impact of geographical customer concentration on sales, the consistent performance of specific product categories, the need for optimized employee training and sales allocation, and crucial opportunities for cost reduction within the supply chain by strategically managing supplier relationships and pricing.

Ultimately, this analysis provides a robust framework for strategic decision-making. By acting on the data-driven insights and implementing the recommended strategies, Northwind Traders can expect to enhance sales performance, optimize operational costs, improve customer satisfaction, and foster a more efficient and productive workforce. This project underscores the power of data analytics in transforming raw data into actionable intelligence for sustainable business growth.