|  |  |
| --- | --- |
|  |  |
| CAPSTONE PROJECT: SALES ANALYTICS |  |
|  |  |
|  | DATA ANALYTICS |
|  | RICHA PANDEY |

### SALES ANALYTICS

In this project, I embarked on the task of developing a comprehensive Power BI report aimed at enhancing decision-making processes for Northwind Traders, a wholesale company operating in a competitive market landscape.

The primary objective was to create a visually appealing and user-friendly dashboard that effectively communicated key performance metrics related to sales analysis, customer segmentation, inventory trends, and employee performance.

Leveraging data from multiple tables within the Northwind Traders database, I conducted extensive data analysis and preprocessing to derive actionable insights. The resulting Power BI report featured interactive visualizations, dynamic filters, and drill-down capabilities, enabling stakeholders to explore data in-depth and extract valuable insights. By empowering stakeholders with access to meaningful data, the report aimed to revolutionize how Northwind Traders interacts with its data, facilitating informed decision-making and positioning the company for sustained success in the wholesale market.

Through meticulous design and development, the project culminated in the delivery of a powerful tool that enables Northwind Traders to remain competitive, drive business growth, and adapt to evolving market dynamics.

Top of Form

### OBJECTIVE

The primary objective of this project was to develop an insightful and user-friendly Power BI report tailored specifically for Northwind Traders, a prominent wholesale company operating in a competitive market environment.

With a focus on enhancing decision-making processes, the project aimed to create a comprehensive dashboard that effectively communicated key performance metrics across various aspects of the business, including sales analysis, customer segmentation, inventory trends, and employee performance.

By consolidating and analyzing data from multiple tables within the Northwind Traders database, the objective was to derive actionable insights that would empower stakeholders to make informed decisions. Through interactive visualizations, dynamic filters, and drill-down capabilities, the report sought to provide stakeholders with a deeper understanding of business operations and facilitate data exploration. Ultimately, the overarching goal was to revolutionize how Northwind Traders utilizes its data, enabling the company to remain competitive, drive growth, and navigate the complexities of the wholesale market landscape effectively.

Top of Form

### SIGNIFICANCE

The significance of this project lies in its potential to revolutionize how Northwind Traders harnesses data to drive business success. By creating a tailored Power BI report, the project empowers stakeholders to gain valuable insights into key performance metrics across various aspects of the company's operations.

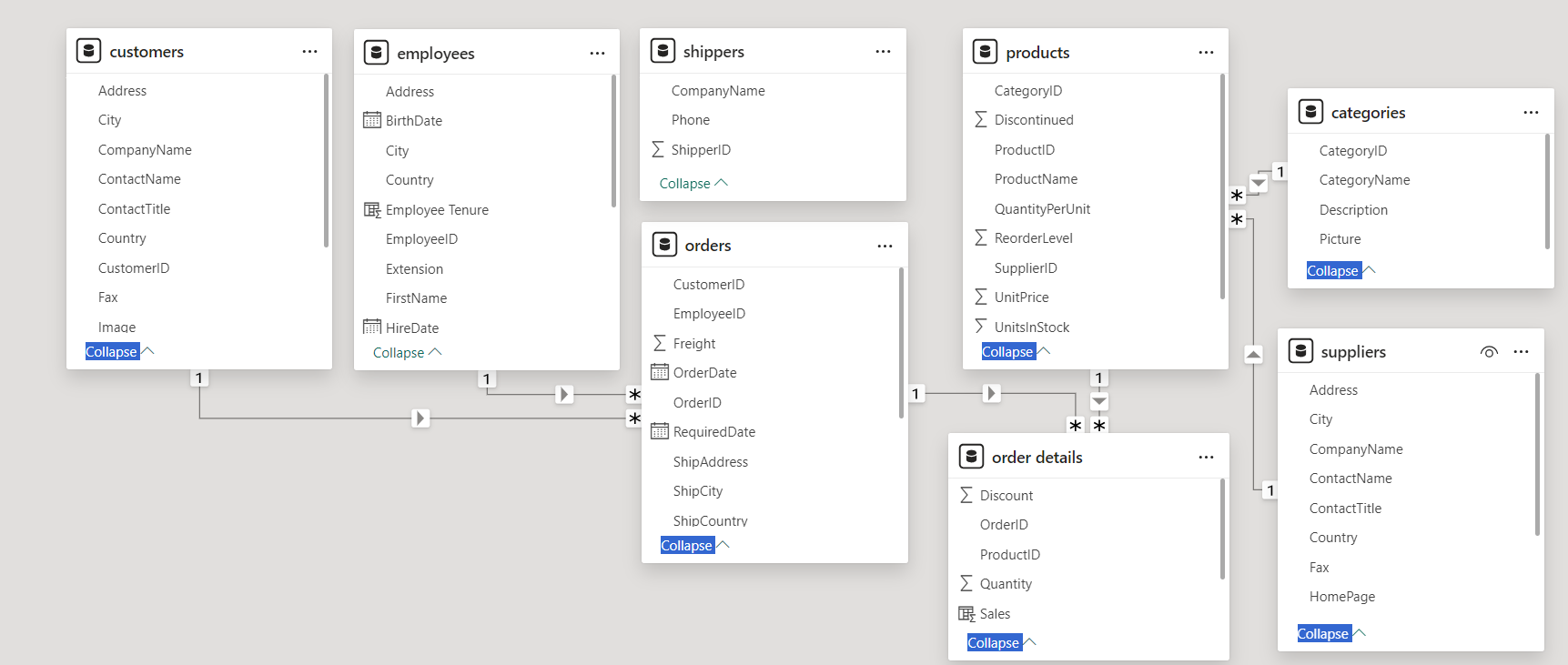
This enhanced understanding enables more informed decision-making processes, ultimately leading to improved strategic planning, resource allocation, and operational efficiency. Furthermore, the interactive nature of the report fosters a culture of data-driven decision-making within the organization, empowering employees at all levels to explore and analyze data independently.

By leveraging data in this manner, Northwind Traders can stay ahead of the curve in a competitive wholesale market landscape, identifying opportunities for growth, optimizing processes, and delivering exceptional value to customers. In essence, the significance of this project lies in its potential to drive meaningful and sustainable improvements in business performance, positioning Northwind Traders for continued success in the future.

### DATA DICTIONARY

1. **Employees Table**:
   * Address: The street address of the employee.
   * Birthdate: The date of birth of the employee.
   * City: The city where the employee resides.
   * Country: The country where the employee resides.
   * Employee ID: Unique identifier for each employee.
   * Extension: Telephone extension number for the employee.
   * First Name: The first name of the employee.
   * Hire Date: The date when the employee was hired.
   * Home Phone: The home telephone number of the employee.
   * Last Name: The last name of the employee.
   * Notes: Additional notes or comments related to the employee.
   * Postal Code: The postal code of the employee's address.
   * Region: The region or state where the employee resides.
   * Reports To: Employee ID of the manager or supervisor to whom the employee reports.
   * Title: The job title or position of the employee.
   * Title of Courtesy: The courtesy title (e.g., Mr., Ms., Dr.) used for addressing the employee.
2. **Customers Table**:
   * Address: The street address of the customer.
   * City: The city where the customer resides.
   * Company Name: The name of the customer's company.
   * Contact Name: The name of the primary contact person at the company.
   * Contact Title: The title or position of the primary contact person.
   * Country: The country where the customer resides.
   * Customer ID: Unique identifier for each customer.
   * Fax: The fax number of the customer.
   * Image: Image associated with the customer.
   * Image Thumbnail: Thumbnail image associated with the customer.
   * Phone: The phone number of the customer.
   * Postal Code: The postal code of the customer's address.
   * Region: The region or state where the customer resides.
3. **Orders Table**:
   * Order ID: Unique identifier for each order.
   * Customer ID: Unique identifier for the customer placing the order.
   * Employee ID: Unique identifier for the employee processing the order.
   * Order Date: The date when the order was placed.
   * Required Date: The date by which the order is required.
   * Shipped Date: The date when the order was shipped.
   * Ship Via: The shipping company or method used for the order.
   * Freight: The freight or shipping cost associated with the order.
   * Ship Name: The name of the recipient for shipping.
   * Ship Address: The shipping address.
   * Ship City: The city for shipping.
   * Ship Region: The region or state for shipping.
   * Ship Postal Code: The postal code for shipping.
   * Ship Country: The country for shipping.
4. **Order Details Table**:
   * Order Detail ID: Unique identifier for each order detail.
   * Order ID: The ID of the order to which the order detail belongs.
   * Product ID: The ID of the product included in the order detail.
   * Quantity: The quantity of the product ordered.
   * Unit Price: The price per unit of the product.
   * Discount: The discount applied to the product.
5. **Products Table**:
   * Product ID: Unique identifier for each product.
   * Product Name: The name of the product.
   * Supplier ID: The ID of the supplier providing the product.
   * Category ID: The ID of the category to which the product belongs.
   * Quantity Per Unit: The quantity of the product per unit (e.g., per case, per dozen).
   * Unit Price: The price per unit of the product.
   * Units In Stock: The number of units of the product in stock.
   * Units On Order: The number of units of the product on order.
   * Reorder Level: The reorder level for the product.
   * Discontinued: Indicates whether the product is discontinued or not.
6. **Categories Table**:
   * Category ID: Unique identifier for each category.
   * Category Name: The name of the category.
   * Description: Description of the category.
7. **Suppliers Table**:
   * Supplier ID: Unique identifier for each supplier.
   * Company Name: The name of the supplier's company.
   * Contact Name: The name of the primary contact person at the company.
   * Contact Title: The title or position of the primary contact person.
   * Address: The street address of the supplier.
   * City: The city where the supplier is located.
   * Region: The region or state where the supplier is located.
   * Postal Code: The postal code of the supplier's address.
   * Country: The country where the supplier is located.
   * Phone: The phone number of the supplier.
   * Fax: The fax number of the supplier.
   * HomePage: The website URL or homepage of the supplier.
8. **Shippers Table**:
   * Shipper ID: Unique identifier for each shipper.
   * Company Name: The name of the shipping company.
   * Phone: The phone number of the shipping company.

### ER DIAGRAM



|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |

### THE PROCESS

**1. Data Acquisition from GitHub:**

Obtain the requisite dataset from a designated GitHub repository, containing essential information on university rankings, encompassing various countries and their performance across distinct ranking systems.

**2. Data Transformation and Enhancement:**

If necessary, execute data transformation procedures to ensure data quality and consistency. Additionally, consider augmenting the dataset with new problem statements to enrich the analysis potential.

**3. Connecting with Tools:**

Establish connections between the dataset and various analytical tools. Interface the dataset with Power BI, Excel, and MySQL Workbench, facilitating seamless data integration and processing.

**4. Problem Statement Solution in Power BI:**

Utilize Power BI to delve into the specified problem statements.

Employ its robust features for data visualization, exploration, and analysis, effectively deriving insights and solutions.

**5. Exploratory Data Analysis (EDA):**

Perform exploratory data analysis using either Excel or SQL

Workbench, depending on the complexity of the analysis. Extract meaningful patterns, relationships, and trends from the data to inform subsequent decision-making.

**6. Creation of Visual and Insightful Power Point:**

Develop a comprehensive Power Point presentation that encapsulates the project's objectives, methodologies, problem statement solutions, and kev visualizations. Each problem statement should be accompanied by a dedicated section with pertinent conclusions and insights.

**7. Detailed Documentation:**

Compile a detailed report that meticulously documents the entire project lifecycle. Include sections on data collection, transformation, problem statement formulation, tools integration, Power BI solutions,

EDA insights, and Power Point visualizations.

### EDA PROBLEM STATEMENTS

|  |
| --- |
| Q1. What are the key factors influencing customer retention or loyalty based on the dataset?  Discount & Shipping days influence customer retention. |
|  |
|  |

We conclude that larger discount attract customers to make more orders.

|  |
| --- |
| We conclude that lesser number of shipping days makes the customers order more. |
|  |

Q2. How do customer preferences vary based on their location or demographics? Can we explore this through interactive visualizations?

We found customer preferences based on their location but not demograpics as there are no demographic given to us in in the dataset. We conclude that beverages and dairy products are the most popular and produce is the least popular among customers around the world.

|  |
| --- |
| We can conclude that beverages are the most popular among customers around the world. |
|  |

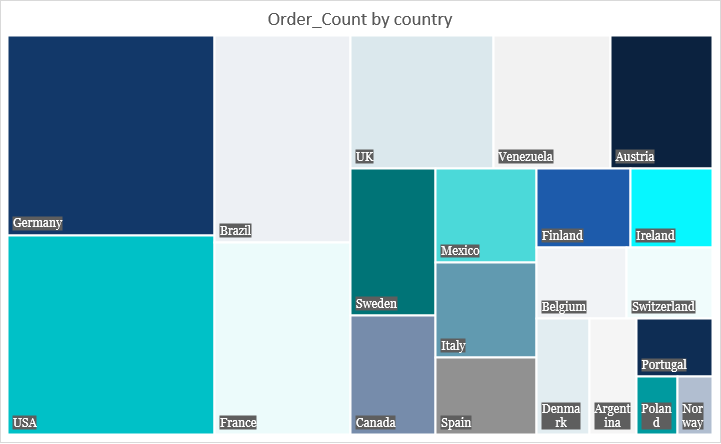
|  |
| --- |
| Q3. Are there any interesting patterns or clusters in customer behavior that can be visualized to identify potential market segments?  Customer behaviour in clusters by countries and categories. USA has the most number of customers and Seafood is bought by most customers. Conversely, Poland and Ireland have the lowest amount of customers and Produce is bought by least number of customers. |
|  |
|  |

|  |
| --- |
| Q4. Are there any specific product categories or SKUs that contribute significantly to order revenue? Can we identify them through visualizations?  Beverages & Dairy Products contribute significantly to the revenue. Produce & Grains/Cereals contribute least revenue. |
|  |

|  |
| --- |
| Q5. Are there any correlations between order size and customer demographics or product categories? Can we explore this visually using scatter plots or heatmaps?  Customer count ranges between 1-38 for different Order sizes and the average order size is is almost the same for all product categories. On an average, Dairy Products have the highest order size and Produce have the lowest. |
|  |

|  |  |
| --- | --- |
| **Category Name** | **Avg Order Size** |
| Dairy Products | 24.9973 |
| Condiments | 24.5278 |
| Meat/Poultry | 24.2717 |
| Confections | 23.6707 |
| Beverages | 23.5941 |
| Seafood | 23.2758 |
| Grains/Cereals | 23.2755 |
| Produce | 21.9853 |

|  |
| --- |
| Q6. How does order frequency vary across different customer segments? Can we visualize this using bar charts or treemaps?  I used location as a customer segment and calculated order count to calculate order frequency approximately as there is nothing given to determine order frequency exactly.  Germany and USA get orders most frequently. Norway and Poland do not get frequent orders. |
|  |



|  |
| --- |
| Q7. Are there any correlations between employee satisfaction levels and key performance indicators? Can we explore this visually through scatter plots or line charts? |
|  |

Tenure gives us the employee satisfaction and and chose Sales as kpi. Employees will only stay with a Company if they are satisfied with the overall perks and tenure and the more tenure an employee has the more products they are likely to sell.

Q8. How does employee turnover vary across different departments or job roles? Can we visualize this using bar charts or heatmaps?

By comparing the order counts for all Job Roles, we can identify whose activity has decreased significantly, potentially indicating turnover. Here Sales Manager indicates turnover.

Q9. Can we identify any patterns or clusters in employee skill sets or qualifications through visualizations? How can this information be used for talent management?

Talent management can give suitable work to employees according to their skillset or qualifications.

Q10. Are there any correlations between product attributes (e.g., size, color, features) and sales performance?

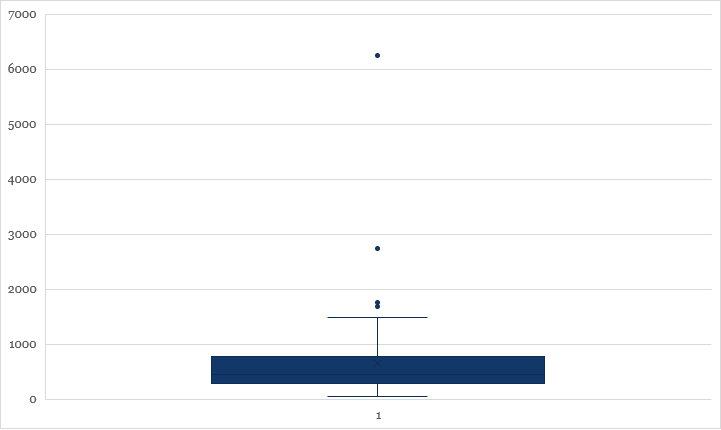
Sales of different products by their categories. Beverages have the highest sales but confections have the most amount of products. Produce have the lowest sales and the least amount of products.

Q11.How does product demand fluctuate over different seasons or months? Can we visualize this through line charts or area charts?

We used order count to determine the fluctuation of product demands over months.

Q12. Can we identify any outliers or anomalies in product performance or sales using visualizations? How can this information be used for product optimization?

To determine outliers, I used average sales. By identifying outliers, businesses can investigate the underlying reasons for these anomalies and take corrective actions if necessary.



Q13. 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?

Used heatmap and scatter plot to determine the performance of country reorder level. Finland is at the top and Brazil at the bottom.

|  |  |
| --- | --- |
| **SupplierCountry** | **AvgReorderLevel** |
| Finland | 21.6667 |
| Sweden | 20 |
| Italy | 19 |
| Germany | 16.1111 |
| Spain | 15 |
| Canada | 13.75 |
| Singapore | 13.3333 |
| Netherlands | 12.5 |
| UK | 12.1429 |
| Norway | 11.6667 |
| USA | 11.25 |
| Australia | 10 |
| France | 8 |
| Denmark | 7.5 |
| Japan | 2.5 |
| Brazil | 0 |

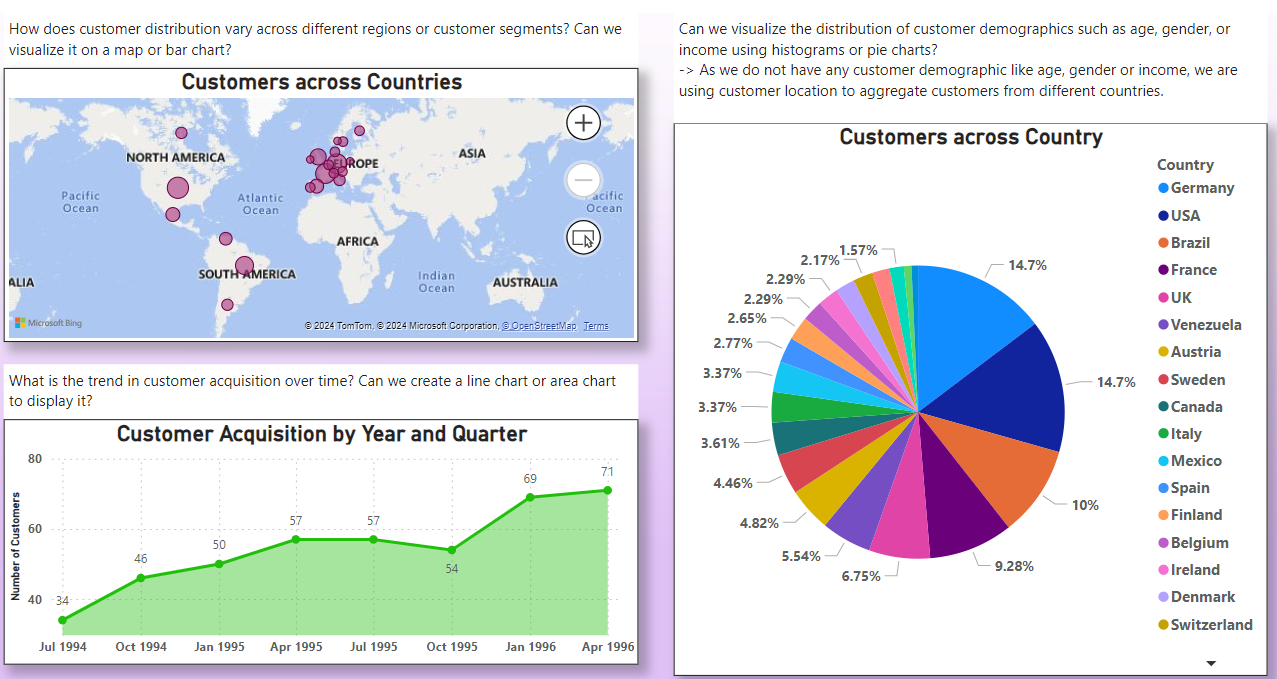
Q14. How does supplier performance vary across different product categories or departments? Can we visualize this using stacked bar charts or grouped column charts?

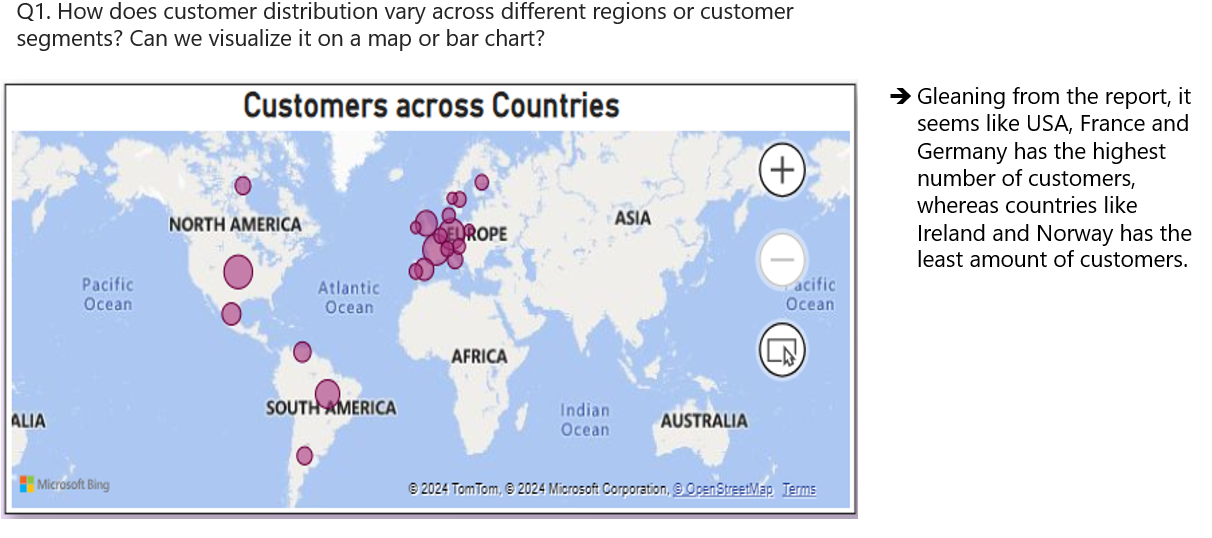
We determine supplier performance by reorder level taken across different categories of products. It indicates that the products a supplier is supplying is of good quality and its sales are decent.

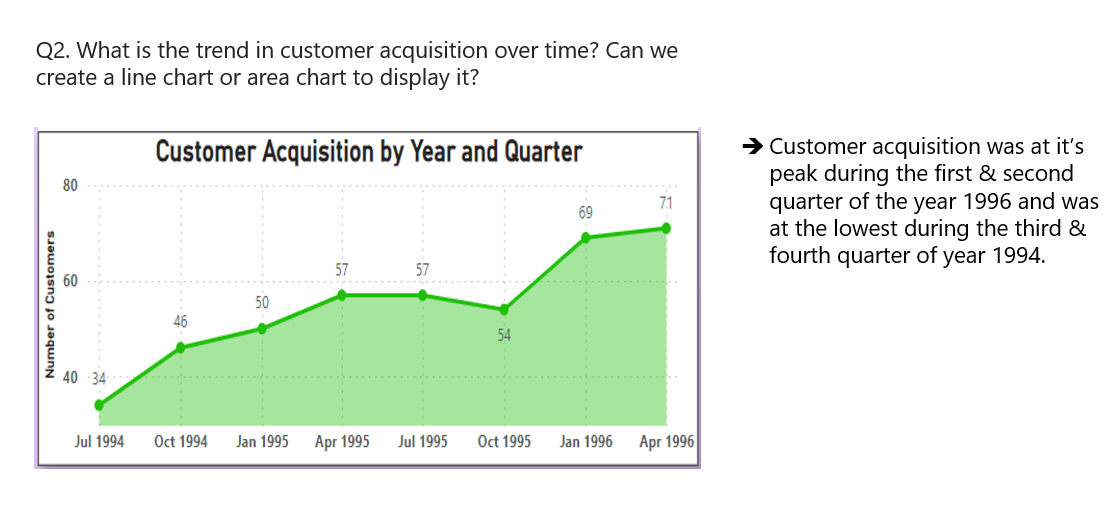
Q15. Can we identify any trends or patterns in supplier costs or pricing structures through visualizations? How can this information be used for procurement optimization?

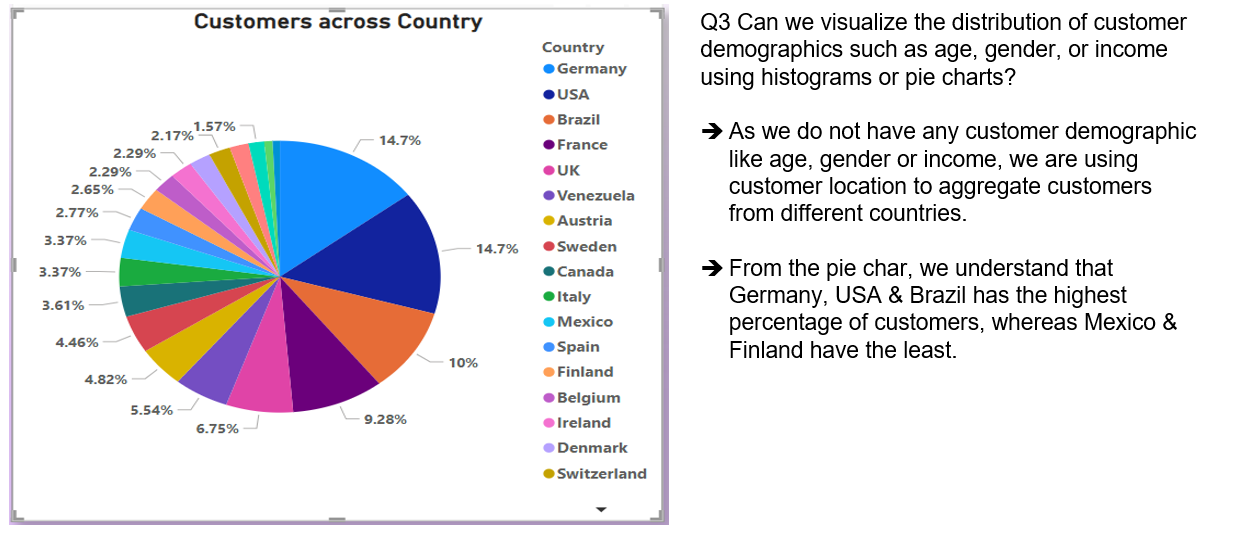
Pricing varies for suppliers of different products between range of 4-282. Aux Joyeux ecdes iastiques has the highest pricing whereas Rrefrescos Americanas LTDA has the lowest.

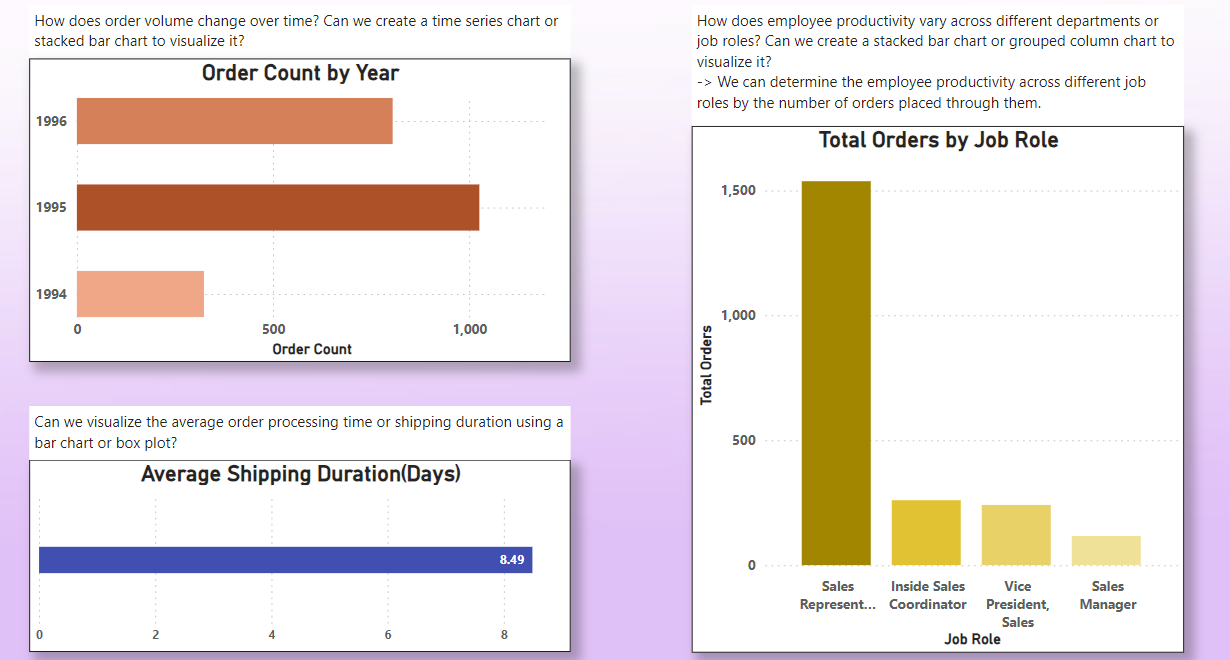
### POWER BI PROBLEM STATEMENTS

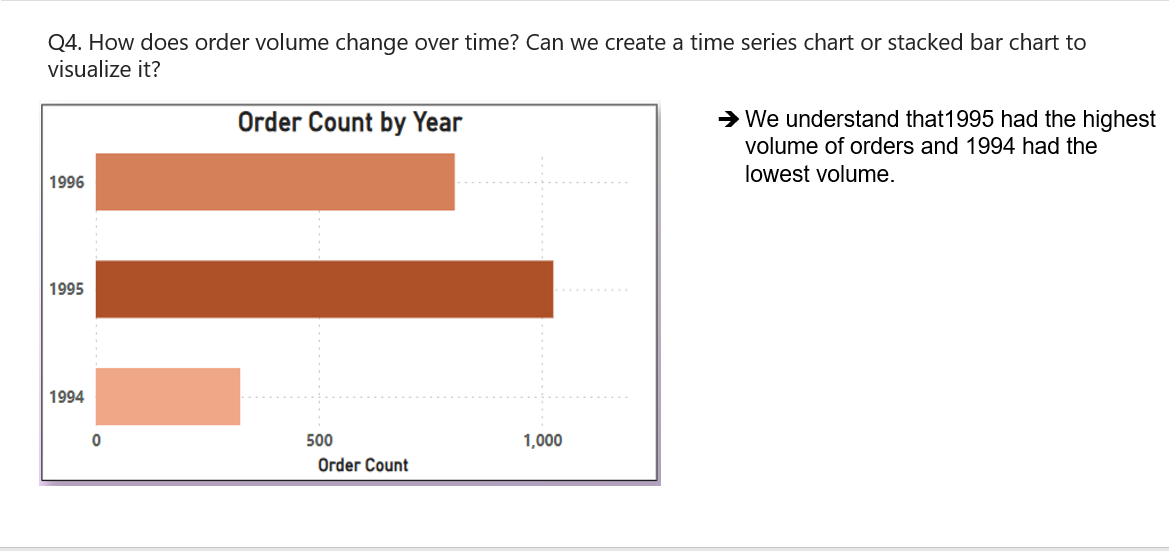


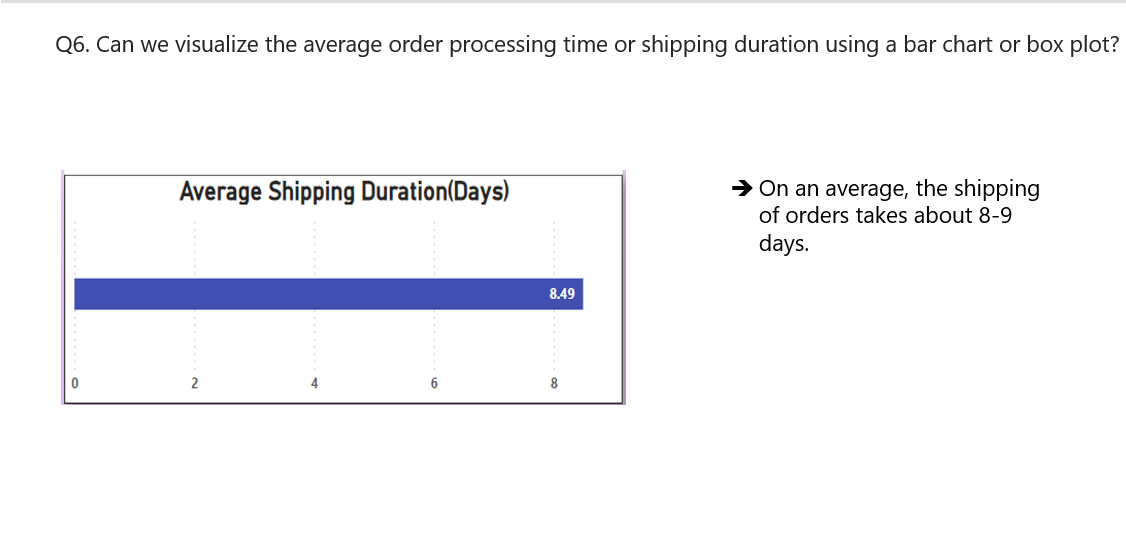


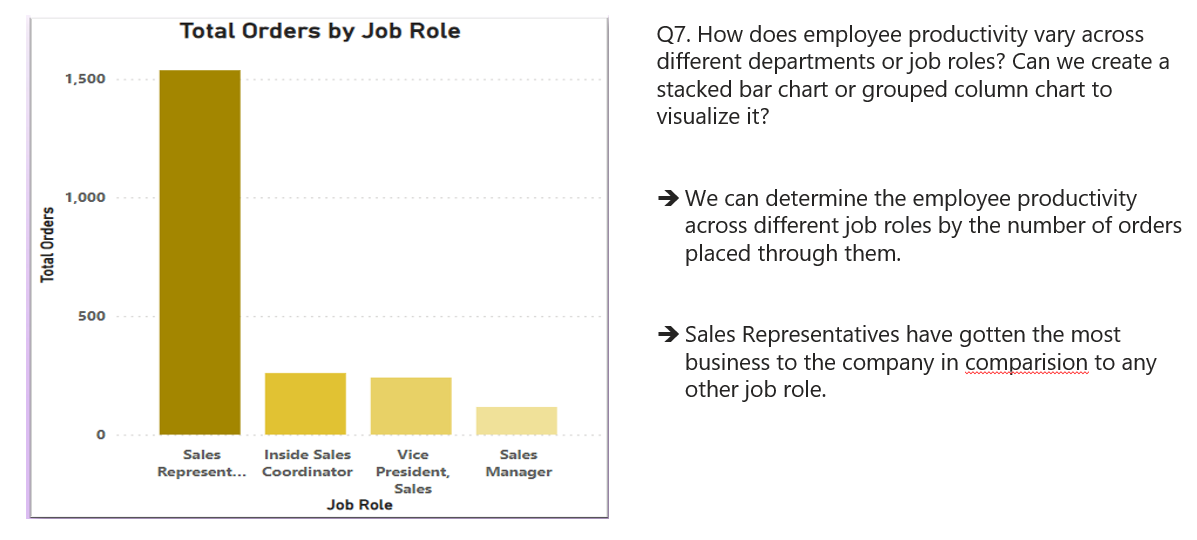


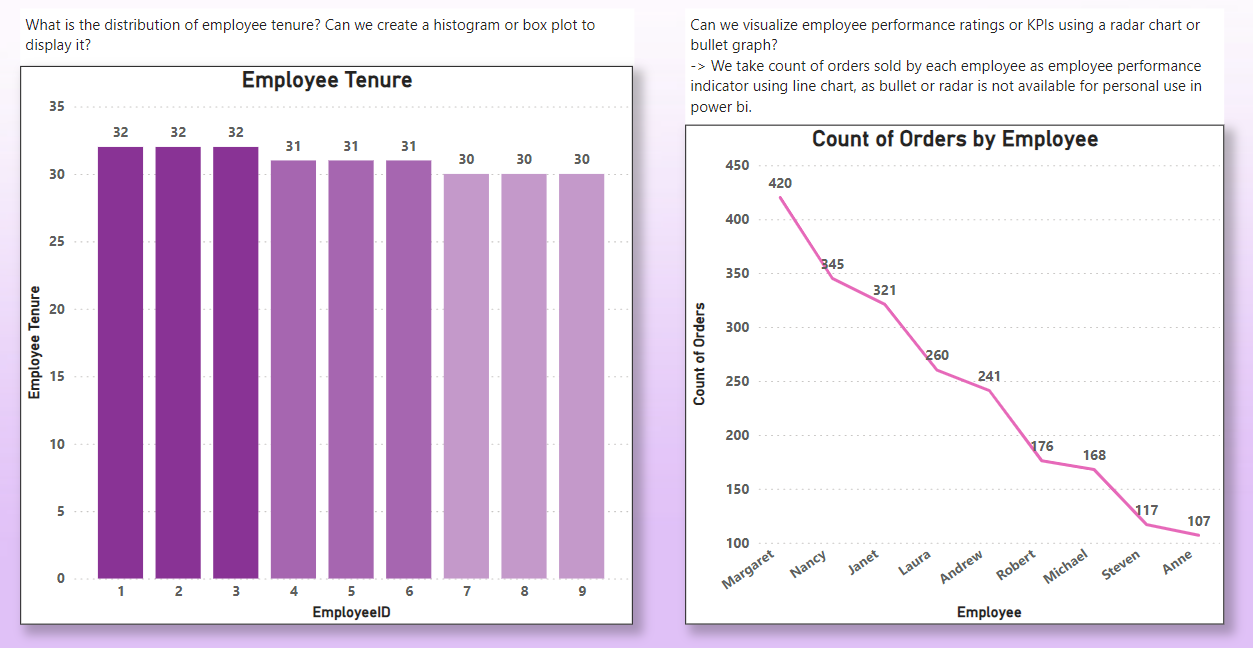


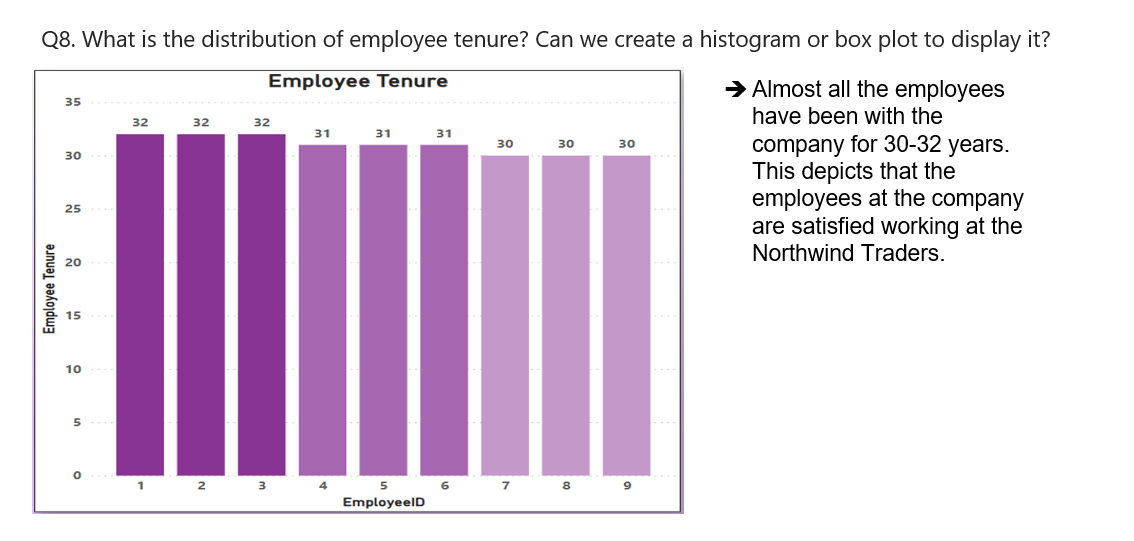


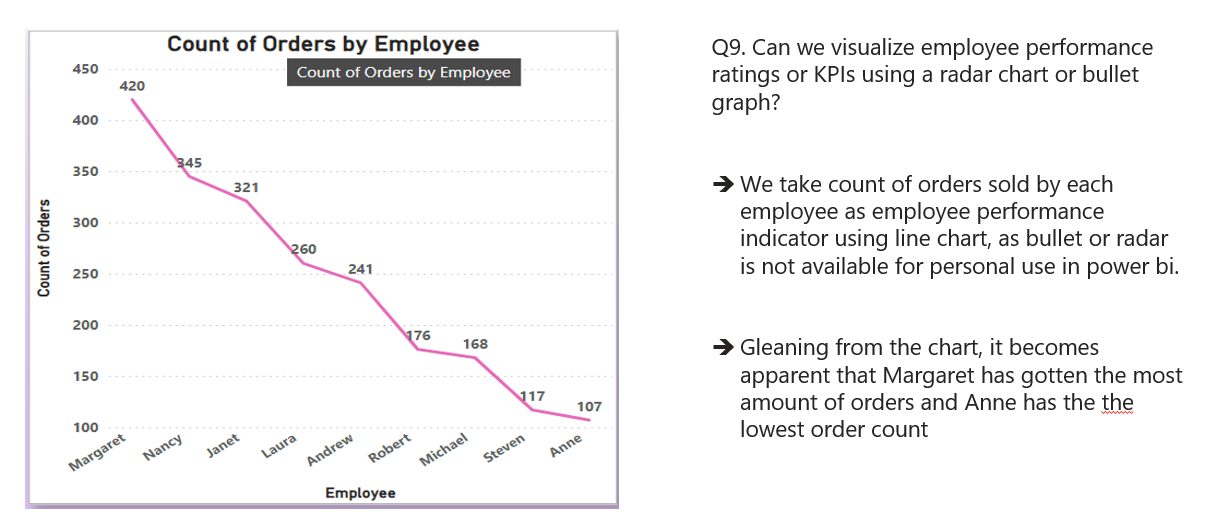


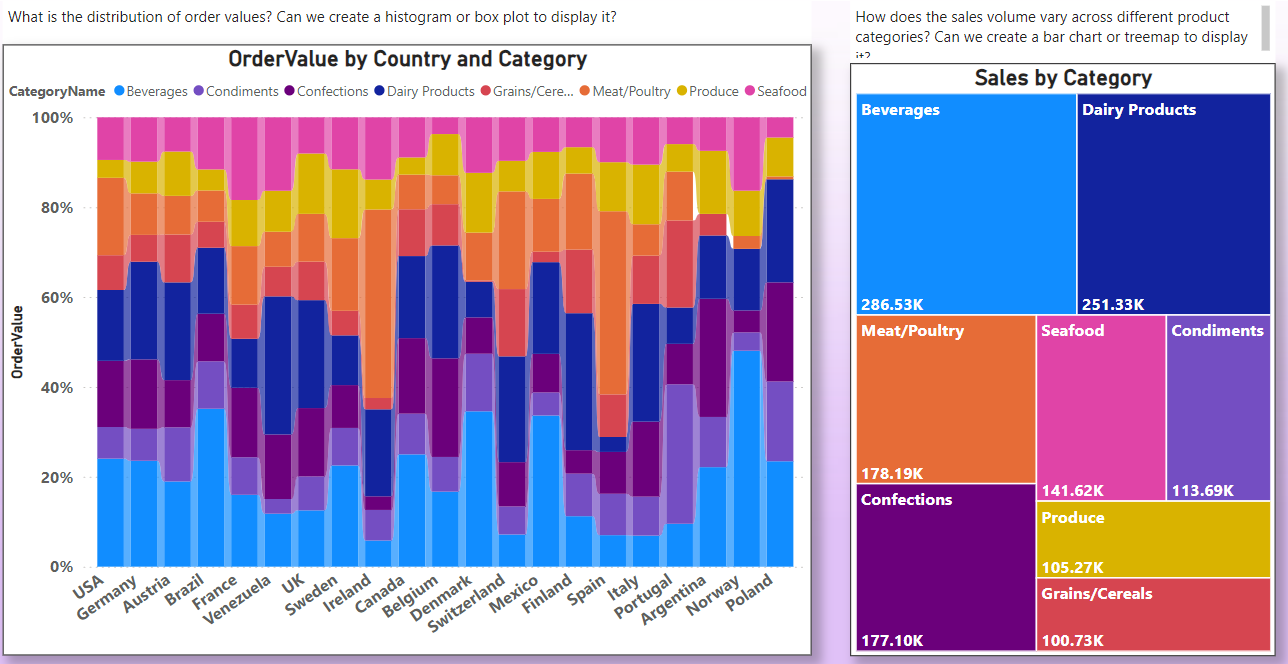


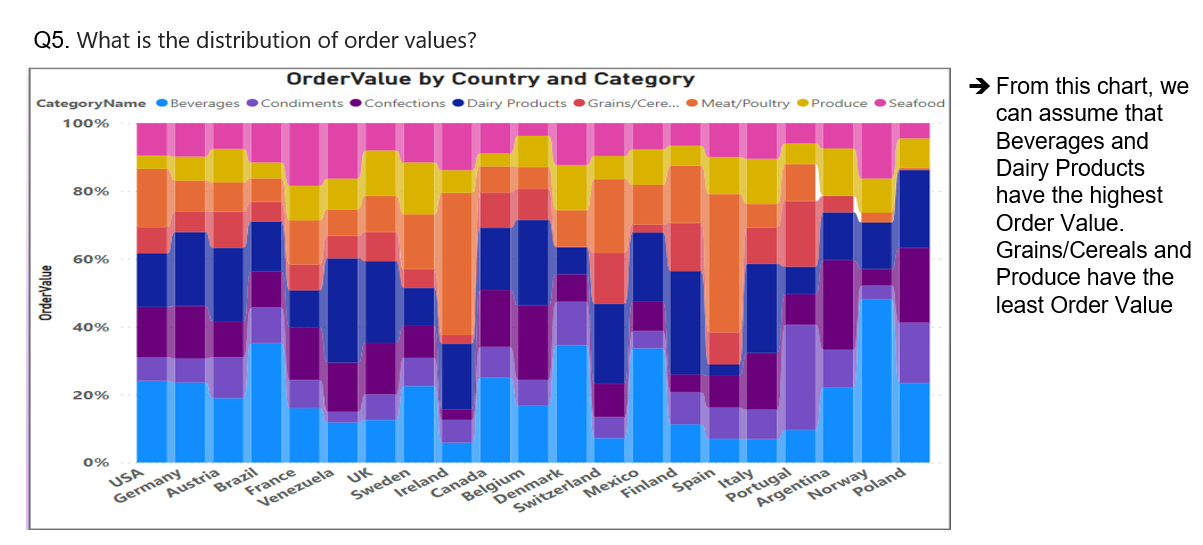


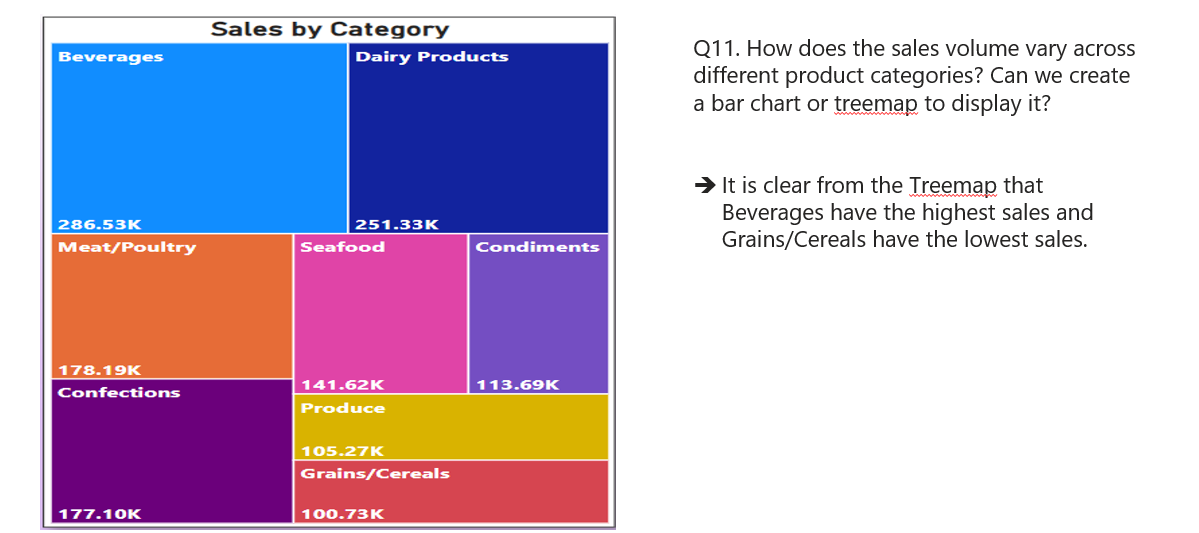


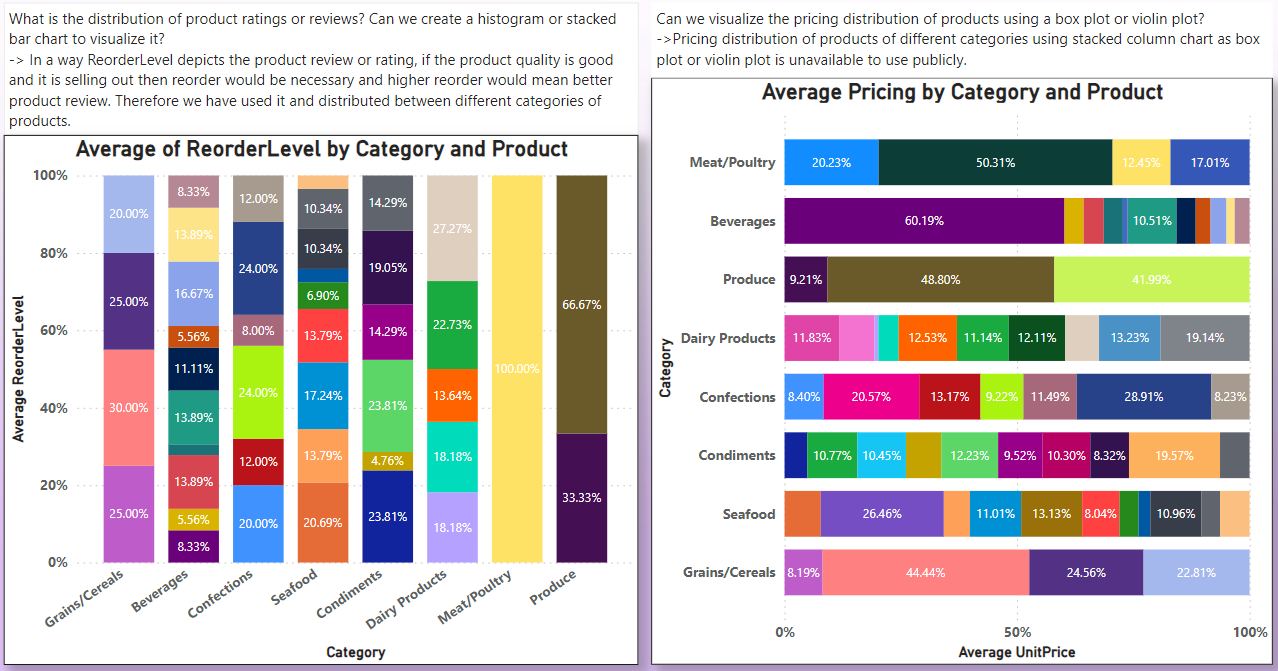


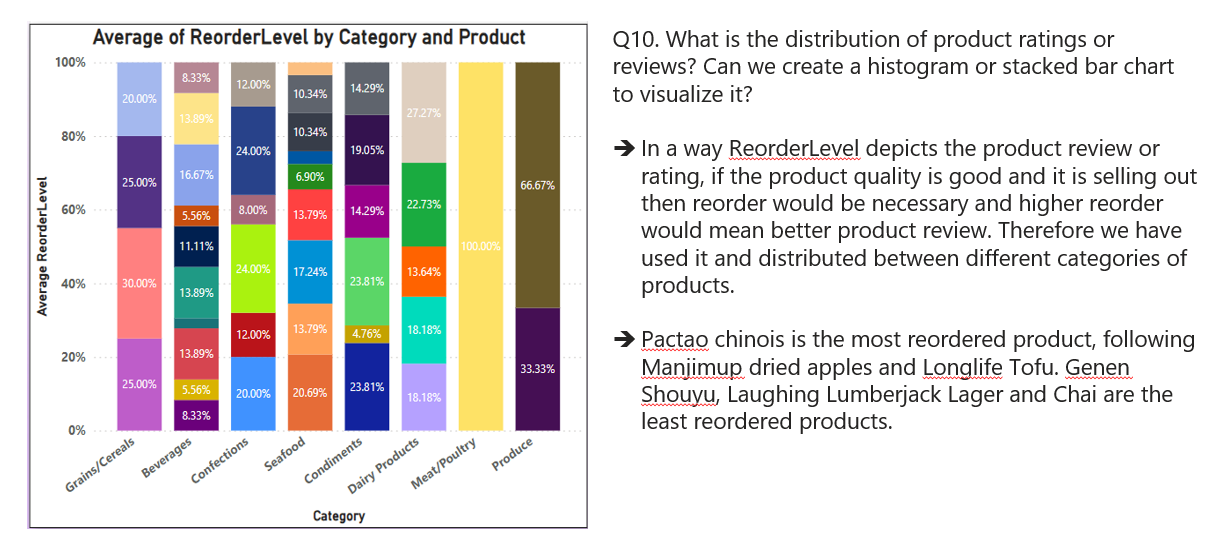


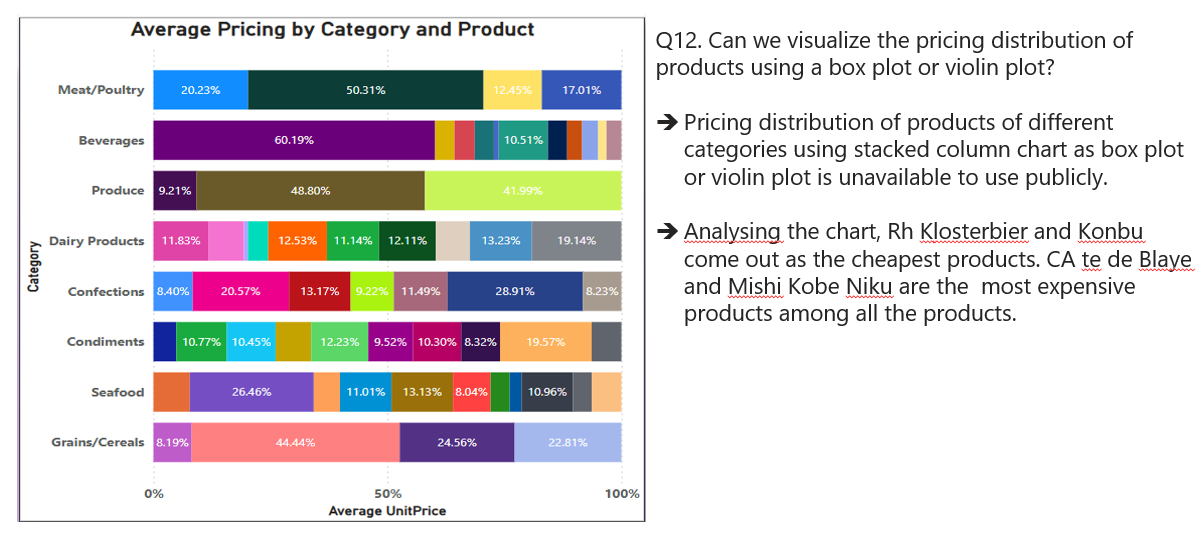


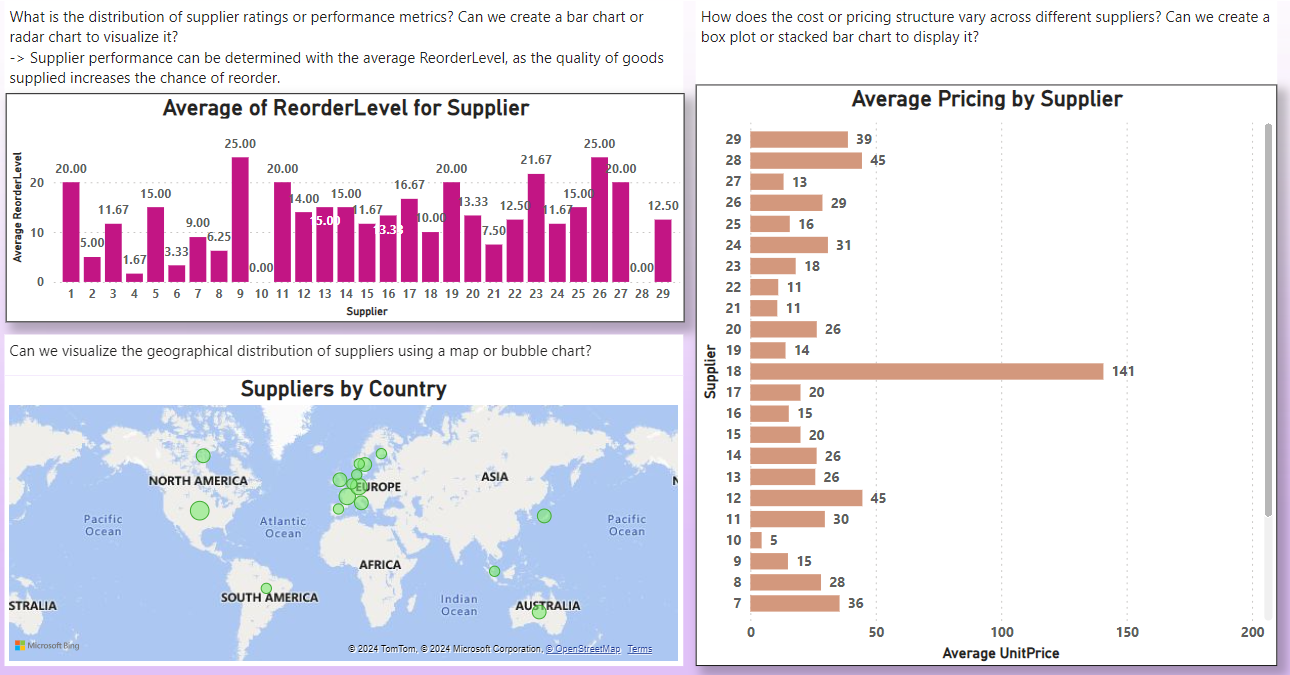


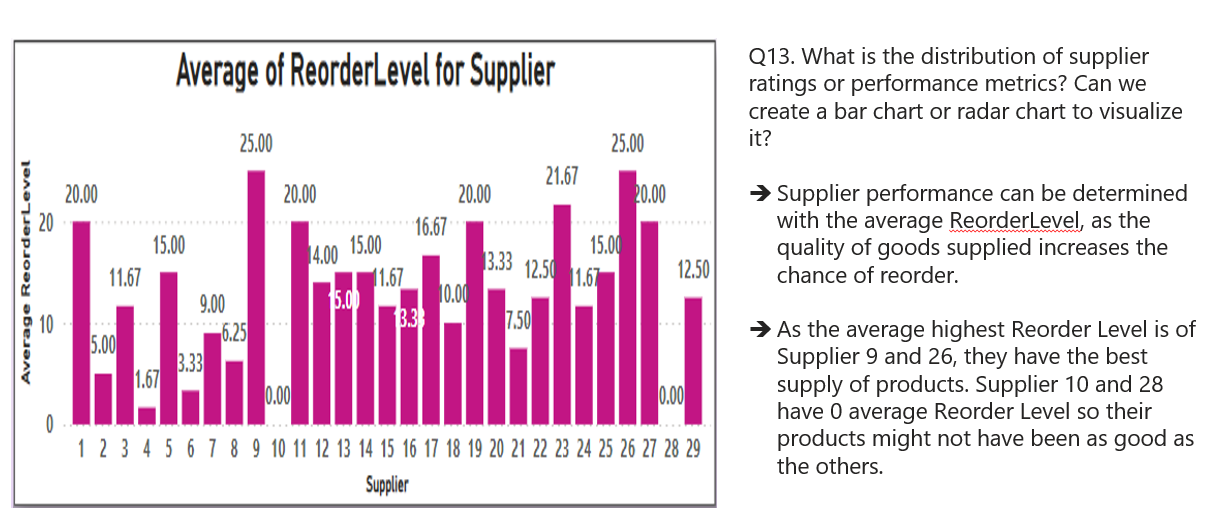


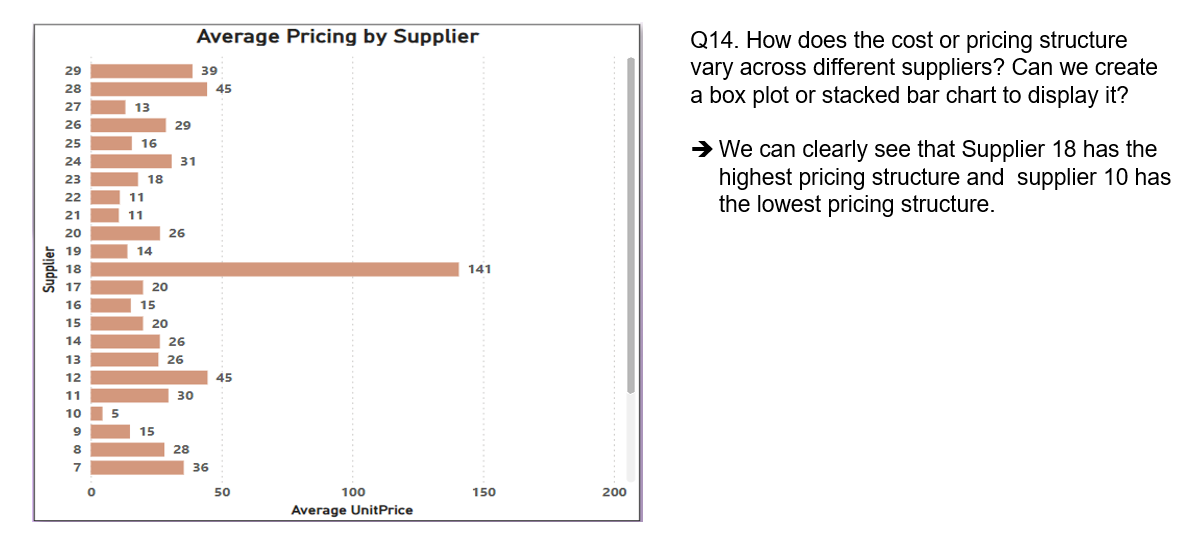


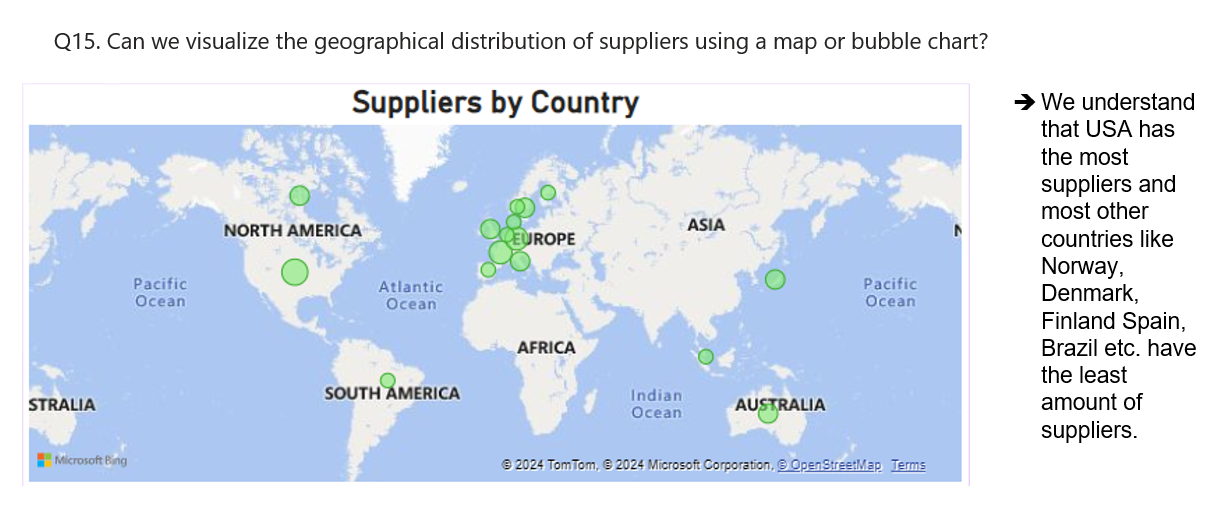












|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  | | |  |
|  | **CONCLUSION**  In conclusion, the development of the Power BI report for Northwind Traders marks a significant milestone towards my learning and in the company's journey towards data-driven decision-making and operational excellence. Through meticulous data analysis and visualization, the report provides stakeholders with actionable insights into key performance metrics, spanning sales analysis, customer segmentation, inventory trends, and employee performance. By empowering decision-makers with comprehensive and interactive data exploration tools, the report enables Northwind Traders to make informed decisions, optimize processes, and capitalize on opportunities for growth and efficiency. The project underscores the importance of leveraging data as a strategic asset, driving continuous improvement and innovation within the organization. Moving forward, the insights gleaned from this report will serve as a foundation for future initiatives, propelling Northwind Traders towards sustained success and competitiveness in the dynamic wholesale market landscape. | | |  |  |
|  |  | | | |  |
|  | | |  | | | |

|  |  |  |
| --- | --- | --- |
|  |  |  |
| Thankyou |
|  |