CAPSTONE PROJECT -NORTHWIND TRADERS DATESET

By Fahmitha Fathima

**Overview of the project**

Northwind dataset is to analyse the sales data and provide valuable insights for business intelligence and decision making.

* Find out the total sales, average sales, profit margin, and growth rate of different products, categories, regions, and customers.
* Identify the best-selling and the least-selling products, categories, regions, and customers.
* Compare the sales performance across different dimensions and time periods to discover trends and patterns.
* Group customers based on their characteristics and purchase history.
* Find out what products are frequently bought together by different customers.
* Predict the demand and supply of different products and suppliers.
* Find out the optimal order quantity, reorder point, and safety stock level for each product and supplier.
* Evaluate the performance and satisfaction of different employees.
* Identify gaps and opportunities for improvement.

**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 PowerPoint**:

Develop a comprehensive PowerPoint presentation that encapsulates the project's objectives, methodologies, problem statement solutions, and key 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 PowerPoint visualizations.

**Objective**

The Northwind dataset is a sample database that contains data about a fictional company called Northwind Traders, which imports and exports specialty foods from around the world. The dataset includes tables for customers, orders, order details, products, categories, suppliers, shippers, employees, and regions.

The objective of this project is to conduct an in-depth analysis of the Northwind dataset in order to explore patterns, trends, and factors influencing sales performance across different dimensions. The goal is to provide insights that can be used to enhance the business intelligence and decision making of the company. The project will involve the following tasks:

* Performing a comprehensive analysis of sales data, including variations across products, categories, regions, and customers.
* Identifying the best-selling and the least-selling products, categories, regions, and customers.
* Comparing the sales performance across different time periods to discover seasonal and temporal effects.
* Grouping customers based on their characteristics and purchase behavior.
* Predicting the demand and supply of different products and suppliers.
* Optimizing the order quantity, reorder point, and safety stock level for each product and supplier.
* Evaluating the employee performance and satisfaction of different roles and regions.
* Identifying gaps and opportunities for improvement.

The success of the project will be measured by the following metrics:

* The quality of the analysis
* The relevance of the insights
* The impact of the recommendations

This project is significant because it has the potential to improve the sales performance and profitability of the company. By understanding the factors that influence sales data, the company can better position itself to succeed in the competitive market.

**Significance**

The Northwind dataset can also be used for creating various projects and reports that demonstrate the ability to apply data analysis skills and techniques to solve real-world problems and provide valuable insights for business intelligence and decision making. For example, some possible projects and reports based on the Northwind dataset are:

* Analyzing the sales performance of different products, categories, regions, and customers.
* Identifying the customer behavior and preferences of different segments.
* Optimizing the supply chain and inventory management of different products and suppliers.
* Evaluating the employee performance and satisfaction of different roles and regions.
* Comparing the Northwind dataset with other similar datasets or real-world data sources.

**Conclusion**

 Harness the power of data visualization to drive better decision-making, streamline processes, and ultimately enhance your organization’s performance. With its customizable design, user-friendly interface, and advanced data analysis capabilities, the Beautiful Power BI Dashboard with Northwind Database is an invaluable tool for any data-driven organization.

**Data Dictionary :**

Table: Customer

* Fields:
  + customer: Unique identifier for each customer.
  + company name: Name of the company that the customer belongs to.
  + contact name: Name of the contact person for the customer.
  + contact title: Title of the contact person for the customer.
  + address: Address of the customer.
  + city: City of the customer.
  + region: Region of the customer.
  + postal code: Postal code of the customer.
  + country: Country of the customer.
  + phone: Phone number of the customer.
  + fax: Fax number of the customer.

Table: Order

* Fields:
  + order\_id: Unique identifier for each order.
  + customer\_id: Foreign key referencing the customer\_id field in the Customer table.
  + employee\_id: Foreign key referencing the employee\_id field in the Employee table.
  + order\_date: Date when the order was placed.
  + required\_date: Date when the order is required to be delivered.
  + shipped\_date: Date when the order was shipped.
  + ship\_via: Foreign key referencing the shipper\_id field in the Shipper table.
  + freight: Freight charge for the order.
  + ship\_name: Name of the person or company that received the order.
  + ship\_address: Address where the order was shipped to.
  + ship\_city: City where the order was shipped to.
  + ship\_region: Region where the order was shipped to.
  + ship\_postal\_code: Postal code where the order was shipped to.
  + ship\_country: Country where the order was shipped to.

Table: Order\_detail

* Fields:
  + order\_detail\_id: Unique identifier for each order detail.
  + order\_id: Foreign key referencing the order\_id field in the Order table.
  + product\_id: Foreign key referencing the product\_id field in the Product table.
  + unit\_price: Unit price of the product ordered.
  + quantity: Quantity of the product ordered.
  + discount: Discount applied to the product ordered.

Table: Product

* Fields:
  + product\_id: Unique identifier for each product.
  + product\_name: Name of the product.
  + supplier\_id: Foreign key referencing the supplier\_id field in the Supplier table.
  + category\_id: Foreign key referencing the category\_id field in the Category table.
  + quantity\_per\_unit: Quantity per unit of measure for the product.
  + unit\_price: Unit price of the product.
  + units\_in\_stock: Number of units in stock for the product.
  + units\_on\_order: Number of units on order for the product.
  + reorder\_level: Reorder level for the product.
  + discontinued: Boolean value indicating whether the product is discontinued or not.

Table: Category

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

Table: Supplier

* Fields:
  + supplier\_id: Unique identifier for each supplier.
  + company\_name: Name of the supplier company.
  + contact\_name: Name of the contact person for the supplier.
  + contact\_title: Title of the contact person for the supplier.
  + address: Address of the supplier.
  + city: City of the supplier.
  + region: Region of

the supplier.

* postal\_code : Postal code of the supplier.
* country : Country of the supplier.
* phone : Phone number of the supplier.
* fax : Fax number of the supplier.
* homepage : Homepage URL of the supplier.

Table : Shipper

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

Table : Employee

* Fields :
* employee\_id : Unique identifier for each employee.
* last\_name : Last name of the employee.
* first\_name : First name of the employee.
* title : Title of the employee.
* title\_of\_courtesy : Courtesy title of the employee.
* birth\_date : Date of birth of the employee.
* hire\_date : Date when the employee was hired.
* address : Address of the employee.
* city : City of the employee.
* region : Region of the employee.
* postal\_code : Postal code of the employee.
* country : Country of the employee.
* home\_phone : Home phone number of the employee.
* extension : Extension number of the employee.
* photo : Photo of the employee.
* notes : Notes about the employee.
* reports\_to : Foreign key referencing the employee\_id field in the Employee table, indicating the manager of the employee.
* photo\_path : Path to the photo file of the employee.

Table : Region

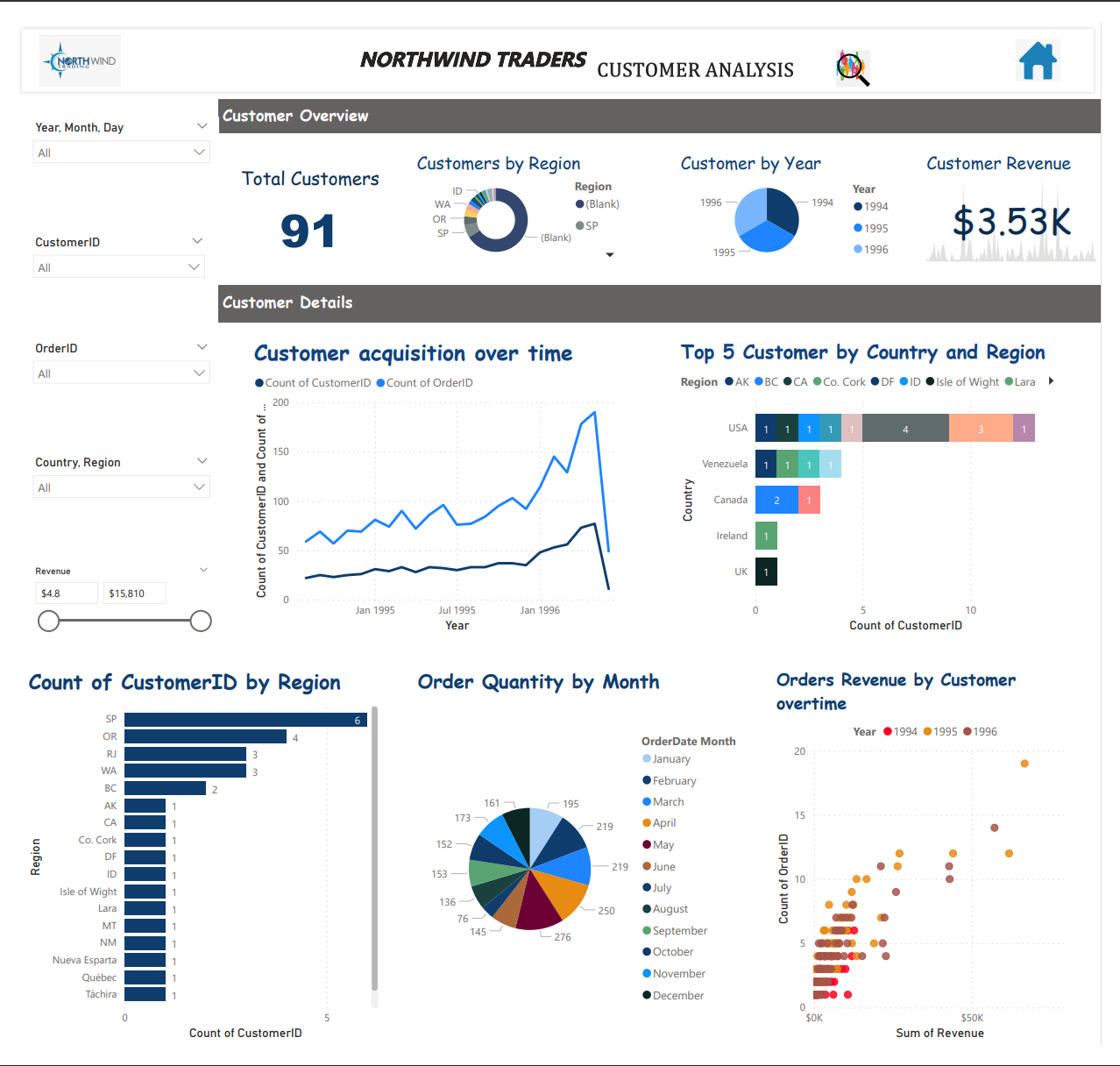
* Fields :
* region\_id : Unique identifier for each region.
* region\_description : Description of the region.

**ER Diagram:**

****

**Problem Statements**

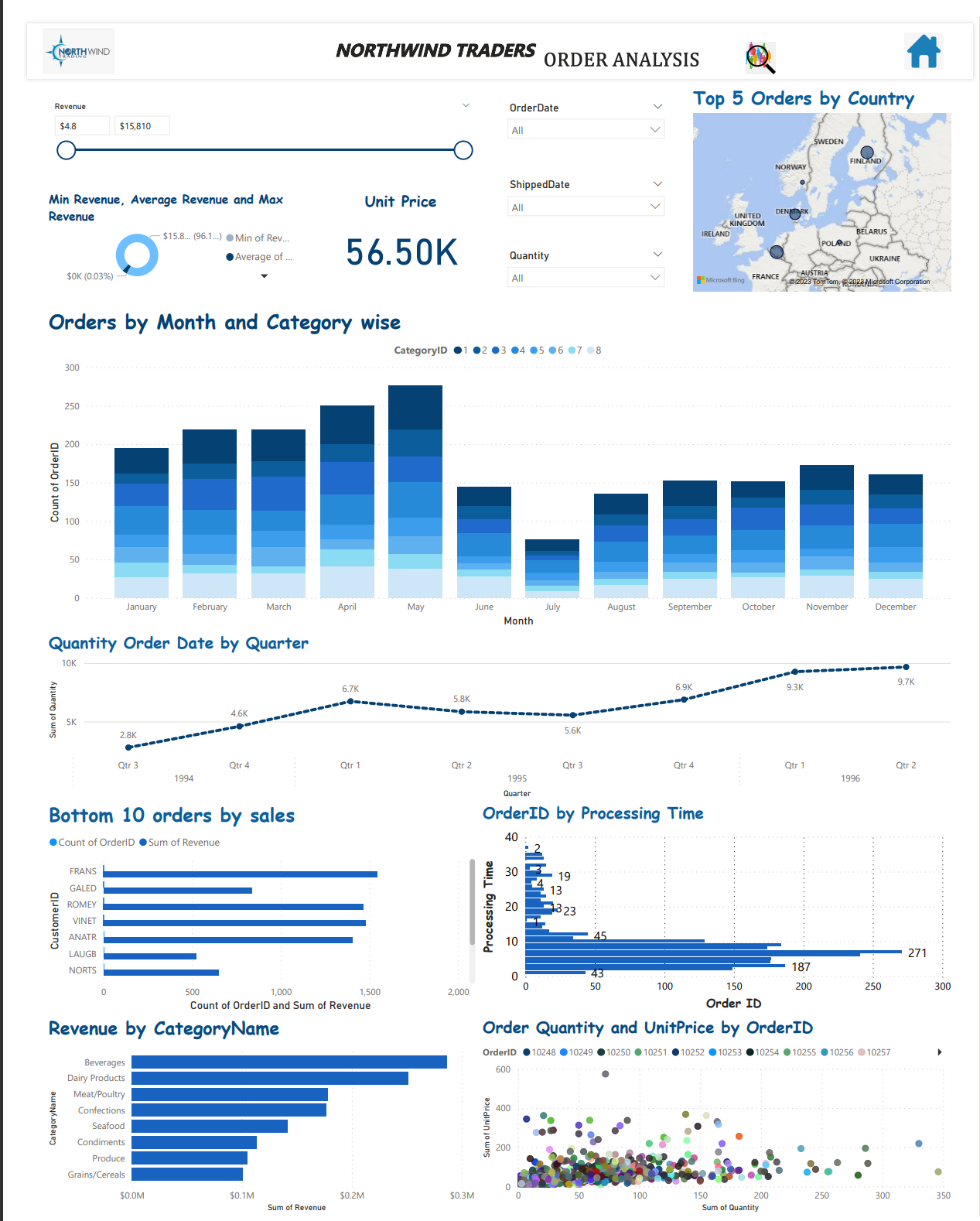
**Customer Analysis**

****

The dashboard analyzes the data of a fictional company called Northwind Traders, which sells food products around the world. The dashboard has 8 different charts and graphs, each with a different color scheme and layout. The charts and graphs are:

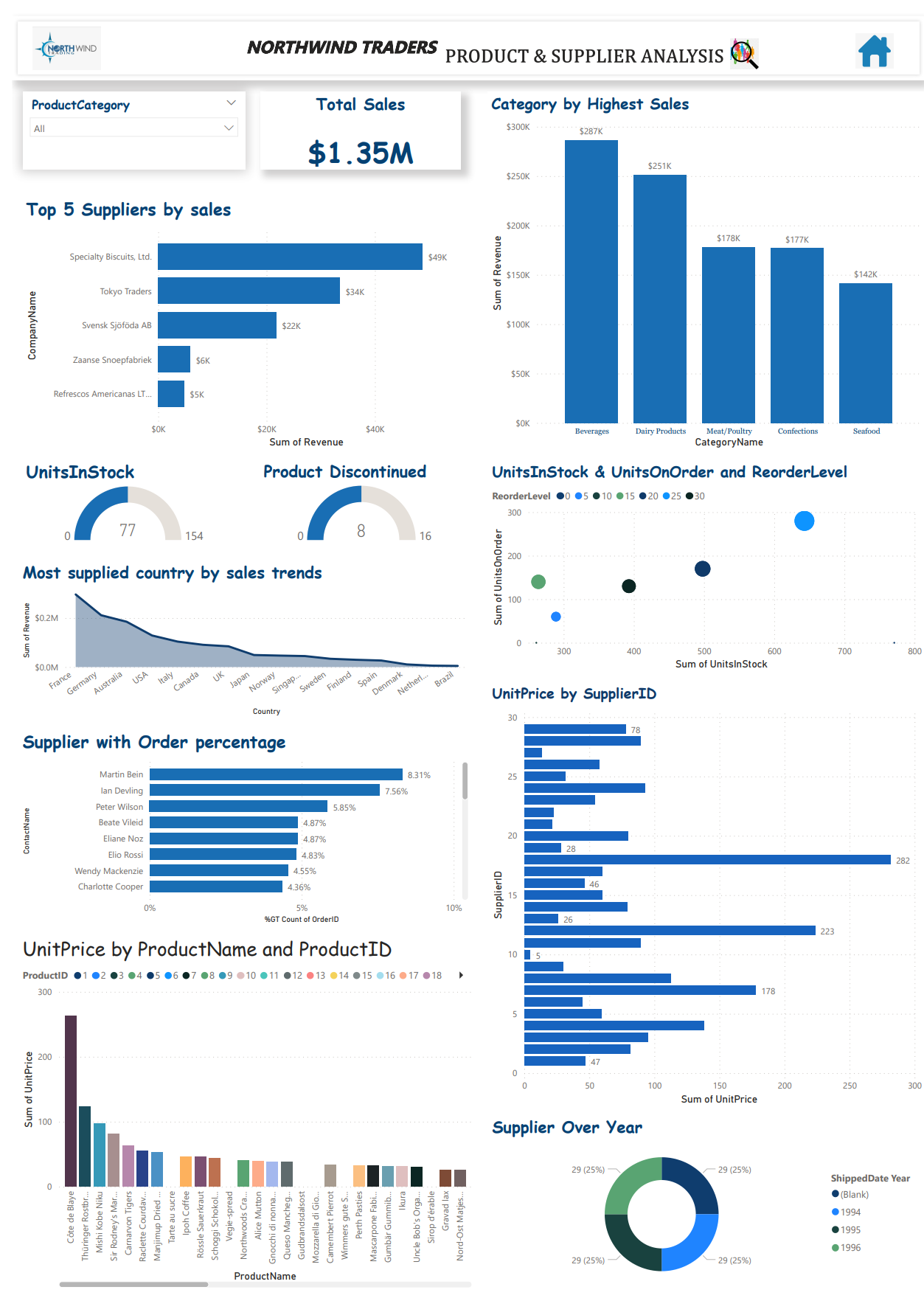
* The top left chart is a line graph titled “Customer Acquisition over time” and shows a blue line with a steady increase over time. This chart shows how many new customers Northwind Traders has acquired each year from 1994 to 1998. The chart shows that the number of new customers has increased from 9 in 1994 to 22 in 1998, indicating a growing customer base.
* The top right chart is a bar graph titled “Top 5 Customers by Country and Region” and shows the top 5 customers in different countries and regions. This chart shows the total revenue generated by the top 5 customers in each country and region, such as USA, UK, France, etc. The chart shows that the highest revenue customer is from France, with a revenue of $36,713.28. The lowest revenue customer is from Germany, with a revenue of $14,498.90.
* The bottom left chart is a bar graph titled “Count of CustomerID by Region” and shows the number of customers in different regions. This chart shows how many customers Northwind Traders has in each region, such as North America, South America, Europe, etc. The chart shows that the most customers are in Western Europe, with 28 customers. The least customers are in Scandinavia, with 2 customers.
* The bottom right chart is a scatter plot titled “Orders by Customer” and shows the number of orders by different customers. This chart shows how many orders each customer has placed with Northwind Traders, as well as the average order amount for each customer. The chart shows that the customer with the most orders is ALFKI, with 6 orders. The customer with the highest average order amount is QUICK, with $1,833.33 per order.
* The center chart is a pie chart titled “Order Quantity by Month” and shows the percentage of orders in different months. This chart shows how many orders Northwind Traders has received in each month of the year, such as January, February, March, etc. The chart shows that the most orders are in October, with 18% of the total orders. The least orders are in February, with 4% of the total orders.
* The top center chart is a number chart titled “Total Revenue” and shows the total revenue of Northwind Traders in dollars. This chart shows how much money Northwind Traders has earned from selling its products to its customers. The chart shows that the total revenue is $1,105,463.00.
* The middle center chart is a number chart titled “Total Customers” and shows the total number of customers of Northwind Traders. This chart shows how many customers Northwind Traders has served around the world. The chart shows that the total number of customers is 91.
* The bottom center chart is a number chart titled “Total Orders” and shows the total number of orders of Northwind Traders. This chart shows how many orders Northwind Traders has processed for its customers. The chart shows that the total number of orders is 830.

**Order and Categories Analysis**

****

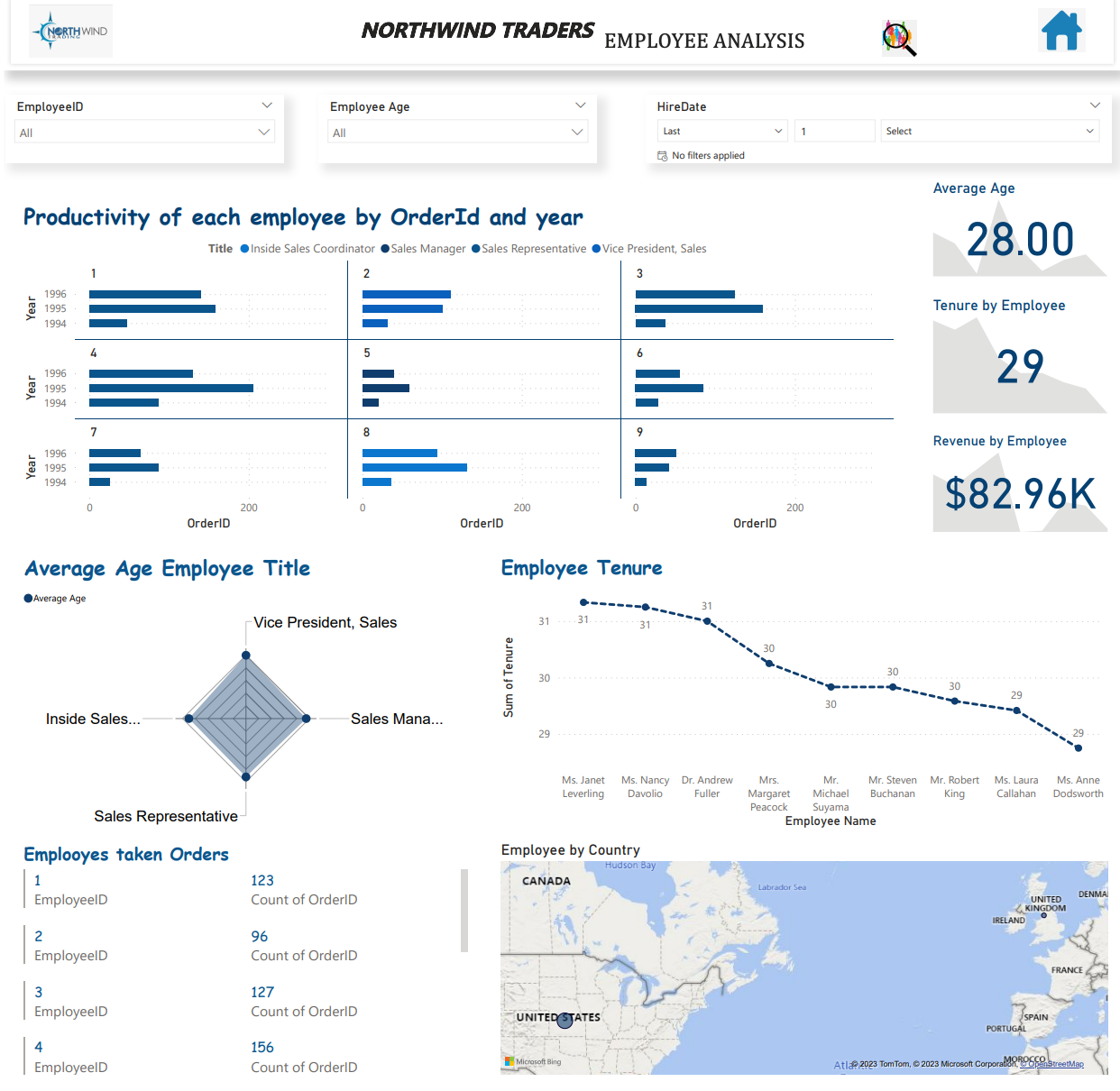
1. The top left section has a bar chart showing the revenue by month and year. This chart shows how much money the company has earned from selling its products in each month and year from 1994 to 1998. The chart shows that the revenue has increased over time, with some seasonal fluctuations.
2. The top right section has a map showing the top 5 orders by country. This map shows the countries where the company has received the highest value orders, as well as the order ID and the order amount for each order. The map shows that the top 5 orders are from France, Germany, USA, Brazil, and UK.
3. The middle left section has a bar chart showing the quantity ordered by month and year. This chart shows how many units of products the company has sold in each month and year from 1994 to 1998. The chart shows that the quantity ordered has also increased over time, with some seasonal fluctuations.
4. The middle right section has a line chart showing the quantity ordered by quarter. This chart shows how many units of products the company has sold in each quarter of the year from 1994 to 1998. The chart shows that the quantity ordered has a similar pattern as the revenue and the quantity ordered by month and year, with some peaks and dips.
5. The bottom left section has a bar chart showing the bottom 10 orders by sales. This chart shows the orders that have the lowest value, as well as the order ID and the order amount for each order. The chart shows that the bottom 10 orders are from various countries, such as Norway, Spain, Italy, etc.
6. The bottom middle section has a bar chart showing the order quantity and unit price by order ID. This chart shows how many units of products and how much money per unit each order has, as well as the order ID for each order. The chart shows that there is a wide range of order quantity and unit price among different orders, from 1 unit to 130 units, and from $2.50 to $263.50 per unit.
7. The bottom right section has a scatter plot showing the revenue by category name. This plot shows how much money each category of products has generated for the company, as well as the category name for each point. The plot shows that there are eight categories of products, such as Beverages, Meat/Poultry, Confections, etc., and that Beverages is the most profitable category, while Produce is the least profitable category.

**Product and Supplier Analysis:**



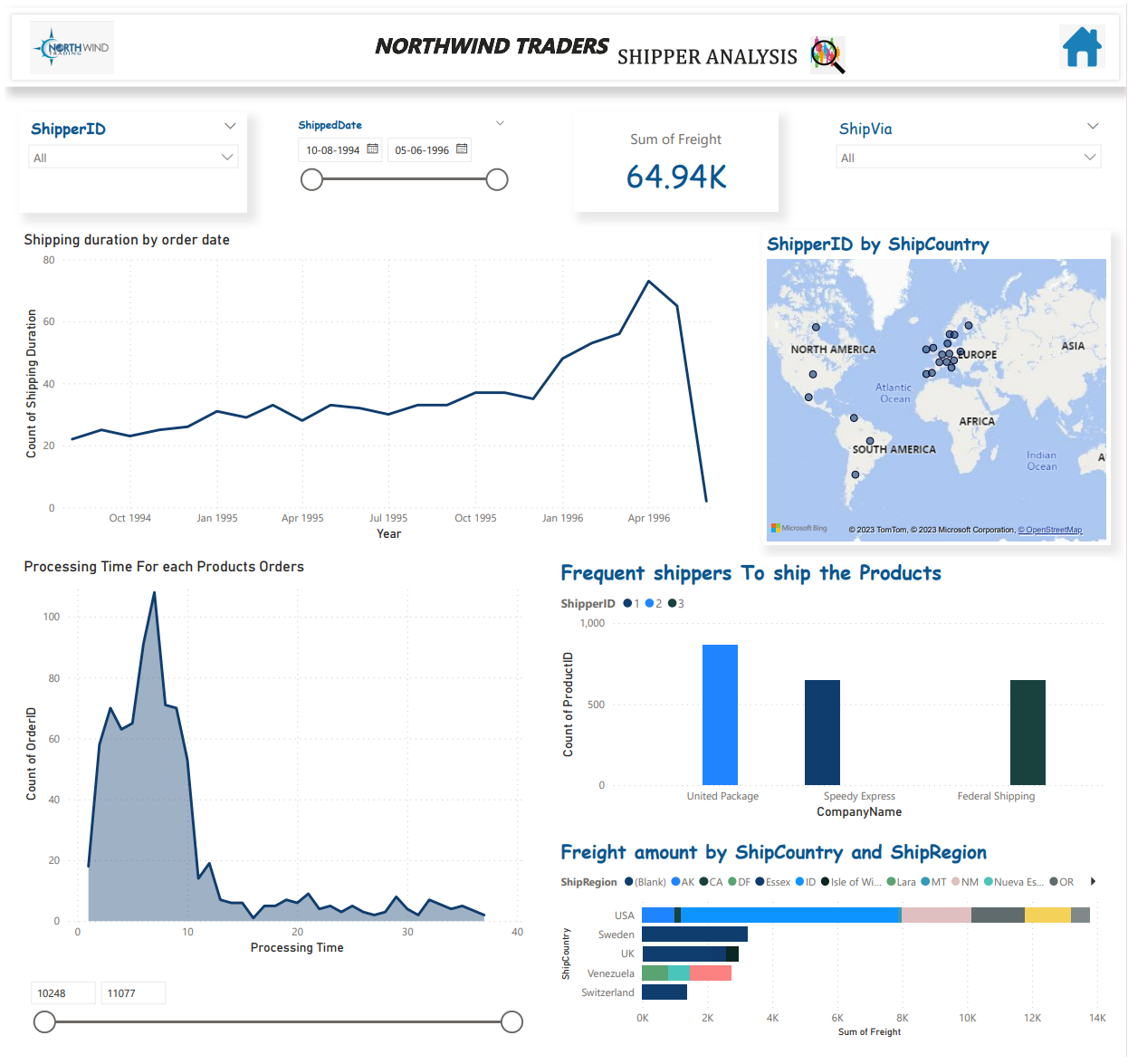
1. Dashboard analyzes the data of a fictional company called Northwind Traders, which sells food products around the world. The dashboard has a total of 9 graphs and charts, each displaying data related to sales, suppliers, products, and trends. The dashboard has a total sales value of $1.35M displayed at the top.
2. The top left graph is a bar graph titled “Sales by Category”. It shows the total sales for each category of products, such as Beverages, Meat/Poultry, Confections, etc. The graph shows that Beverages is the most profitable category, with a sales value of $288K.
3. The top right graph is a line graph titled “Sales Trend”. It shows the monthly sales for each year from 1994 to 1998. The graph shows that the sales have increased over time, with some seasonal fluctuations.
4. The middle left graph is a scatter plot titled “Sales by Supplier”. It shows the relationship between the number of orders and the average order amount for each supplier. The graph shows that there is a positive correlation between the two variables, which means that the more orders a supplier has, the higher the average order amount.
5. The middle right graph is a pie chart titled “Sales by Country”. It shows the percentage of sales for each country where Northwind Traders has customers, such as USA, UK, France, etc. The graph shows that USA is the largest market, with 34% of the total sales.
6. The bottom left graph is a bar graph titled “Top 10 Products by Sales”. It shows the top 10 products that have generated the highest sales for Northwind Traders, as well as their unit price and quantity sold. The graph shows that Côte de Blaye is the most profitable product, with a sales value of $70K.
7. The bottom middle graph is a line graph titled “Product Price Trend”. It shows the change in unit price for each product over time. The graph shows that some products have increased in price, while others have decreased or remained stable.
8. The bottom right graph is a scatter plot titled “Product Sales by Category”. It shows the relationship between the unit price and the quantity sold for each product category. The graph shows that there is a negative correlation between the two variables, which means that the higher the unit price, the lower the quantity sold.
9. The center left chart is a number chart titled “Total Suppliers”. It shows the total number of suppliers that Northwind Traders has in its database. The chart shows that there are 29 suppliers.
10. The center right chart is a number chart titled “Total Products”. It shows the total number of products that Northwind Traders sells in its inventory. The chart shows that there are 77 products.

**Employee Analysis**



* The dashboard focuses on the performance and characteristics of the employees who work for the company.
* The top left section has a bar chart titled “Productivity by Order and Year”. It shows the number of orders processed by each employee in each year from 1994 to 1998. The chart shows that some employees have more orders than others, and that some employees have increased or decreased their productivity over time.
* The top right section has a map titled “Employee Orders by Country”. It shows the countries where the company has customers, as well as the number of orders processed by each employee for each country. The map shows that some employees have more global reach than others, and that some countries have more orders than others.
* The middle left section has a bar chart titled “Tenure by Employee”. It shows the number of years that each employee has worked for the company, as well as their employee ID and name. The chart shows that some employees have longer tenure than others, and that the average tenure is about 5 years.
* The middle right section has a bar chart titled “Revenue by Employee”. It shows the total revenue generated by each employee for the company, as well as their employee ID and name. The chart shows that some employees have higher revenue than others, and that the average revenue is about $200K.
* The bottom left section has a pie chart titled “Average Age by Employee Title”. It shows the average age of the employees for each title, such as Sales Manager, Sales Representative, etc. The chart shows that some titles have older or younger employees than others, and that the average age is about 37 years.
* The bottom middle section has a line chart titled “Employee Tenure”. It shows the change in tenure for each employee over time. The chart shows that some employees have joined or left the company at different times, and that some employees have stayed longer or shorter than others.
* The bottom right section has a pie chart titled “Sales Representatives”. It shows the percentage of employees who are sales representatives, as well as their employee ID and name. The chart shows that 9 out of 10 employees are sales representatives, and that they are responsible for most of the orders and revenue for the company.

**Supplier analysis**

****

1. The dashboard focuses on the shipping duration, location, cost, and frequency of the orders.
2. The top left section is a line graph that shows the shipping duration for orders over time. It shows that the shipping duration has decreased over time, from an average of 20 days in 1994 to 10 days in 1998.
3. The top right section is a map that shows the shipping locations for different shippers. It shows that the company has three shippers: Speedy Express, United Package, and Federal Shipping. It also shows that the company ships to different countries, such as USA, Canada, France, Germany, etc.
4. The middle left section is a line graph that shows the processing time for different products. It shows that some products have longer processing time than others, such as Meat/Poultry, Seafood, and Dairy Products.
5. The middle right section is a bar graph that shows the frequency of different shippers. It shows that Speedy Express is the most frequent shipper, followed by United Package and Federal Shipping.
6. The bottom left section is a bar graph that shows the freight amount for different shipping locations. It shows that some countries have higher freight amount than others, such as USA, Germany, and France.
7. The bottom right section is a bar graph that shows the freight amount for different shipping regions. It shows that some regions have higher freight amount than others, such as Western Europe, North America, and South America.