

NorthWind Sales Analysis

By Preety Paramanick



Overview

Of Northwind Sales Analysis

The aim of this report is to design an engaging and intuitive dashboard that effectively communicates Northwind Traders' key performance indicators. The report will provide insights into customer behavior, sales trends, and employee performance, aiding in decision-making processes. It will encompass sales analysis, customer segmentation, inventory trends, and employee performance, amalgamating data from various tables for a holistic view of the company's operations. The report will enable stakeholders to make informed decisions by providing valuable insights and facilitating data exploration through interactive visualizations and dynamic filters. The anticipated impact is a transformation in how Northwind Traders interacts with its data, equipping the company to stay competitive and propel its business forward in the wholesale market landscape. Interacts with its data, equipping the company to stay competitive and propel its business forward in the wholesale market landscape.

The Northwind database houses the sales data for a fictitious company named "Northwind Traders," which specializes in the import and export of unique foods globally.

Objective

The objective of this Power BI report is to create a visually appealing and user-friendly dashboard that effectively communicates key performance metrics for Northwind Traders. The report aims to provide insights into customer behavior, sales patterns, and employee performance, thereby aiding in the decision-making processes. It will cover areas such as sales analysis, customer segmentation, inventory trends, and employee performance, consolidating data from multiple tables for a comprehensive view of the company's operations. The report is designed to empower stakeholders to make data-driven decisions by offering valuable insights and facilitating data exploration through interactive visualizations and dynamic filters. The expected impact is to revolutionize how Northwind Traders interacts with its data, enabling the company to remain competitive and drive its business forward in the wholesale market landscape.

Data Dictionary

This data dictionary provides a detailed explanation of each field in the tables, making it easier to understand the data.

Customers Table

- **Customer ID:** Unique identifier for each customer
- **Company Name:** Name of the customer's company
- **Contact Name:** Name of the contact person at the customer's company
- **Contact Title:** Job title of the contact person
- **Address:** Address of the customer's company
- **City:** City where the customer's company is located
- **Region:** Region where the customer's company is located
- **Postal Code:** Postal code of the customer's company
- **Country:** Country where the customer's company is located
- **Phone:** Phone number of the customer's company
- **Fax:** Fax number of the customer's company

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 for the employee (e.g., Mr., Ms., Dr.)
- **Birth Date:** Birth date of the employee
- **Hire Date:** Date when the employee was hired
- **Address:** Address of the employee
- **City:** City where the employee lives
- **Region:** Region where the employee lives
- **Postal Code:** Postal code of the employee's address
- **Country:** Country where the employee lives
- **Home Phone:** Home phone number of the employee
- **Extension:** Extension number at the office for the employee
- **Photo:** Photo of the employee
- **Notes:** Additional notes about the employee
- **Reports To:** The superior/manager to whom the employee reports
- **Photo Path:** File path of the employee's photo
- **NumOrder:** Sum of NumOrder of Employee
- **Tenure:** Number of Working days of Employees

-
- **Total OrderValue:** Total Order Value of Employees

Orders Table

- **Order ID:** Unique identifier for each order
- **Customer ID:** Unique identifier for the customer who placed the order
- **Employee ID:** Unique identifier for the employee who handled the order
- **Order Date:** Date when the order was placed
- **Required Date:** Date when the order needs to be delivered
- **Shipped Date:** Date when the order was shipped
- **Ship Via:** Method of shipping
- **Freight:** Freight charges for the order
- **Ship Name:** Name of the ship/vehicle used for shipping
- **Ship Address:** Address where the order needs to be shipped
- **Ship City:** City where the order needs to be shipped
- **Ship Region:** Region where the order needs to be shipped
- **Ship Postal Code:** Postal code of the shipping address
- **Ship Country:** Country where the order needs to be shipped
- **Processing Time:** Processing time of orders
- **NumOrders:** NumOrder of orders

Order Details Table

- **Order ID:** Unique identifier for the order
- **Product ID:** Unique identifier for the product ordered
- **Unit Price:** Price per unit of the product
- **Quantity:** Quantity of the product ordered
- **Discount:** Discount applied to the product
- **OrderValue:** Ordervalue of order details
- **TotalValue:** Totalvalue of order details

Products Table

- **Product ID:** Unique identifier for each product
- **Product Name:** Name of the product
- **Supplier ID:** Unique identifier for the supplier of the product
- **Category ID:** Unique identifier for the category of the product
- **Quantity Per Unit:** Quantity of the product per unit
- **Unit Price:** Price per unit of the product
- **Units In Stock:** Number of units of the product in stock

-
- **Units On Order:** Number of units of the product on order
 - **Reorder Level:** Level of stock at which the product should be reordered
 - **Discontinued:** Whether the product is discontinued
 - **Average_Unit_Price:** Average value of unit price of the product
 - **Revenue:** Total sales of product
 - **Total Quantity Sold:** Number of total quantity sold
 - **Total Sales Volume:** Total Number of sales

Suppliers Table

- **Supplier ID:** Unique identifier for each supplier
- **Company Name:** Name of the supplier company
- **Contact Name:** Name of the contact person at the supplier company
- **Contact Title:** Job title of the contact person
- **Address:** Address of the supplier company
- **City:** City where the supplier company is located
- **Region:** Region where the supplier company is located
- **Postal Code:** Postal code of the supplier company
- **Country:** Country where the supplier company is located
- **Phone:** Phone number of the supplier company
- **Fax:** Fax number of the supplier company
- **Home Page:** Website of the supplier company

Shippers Table

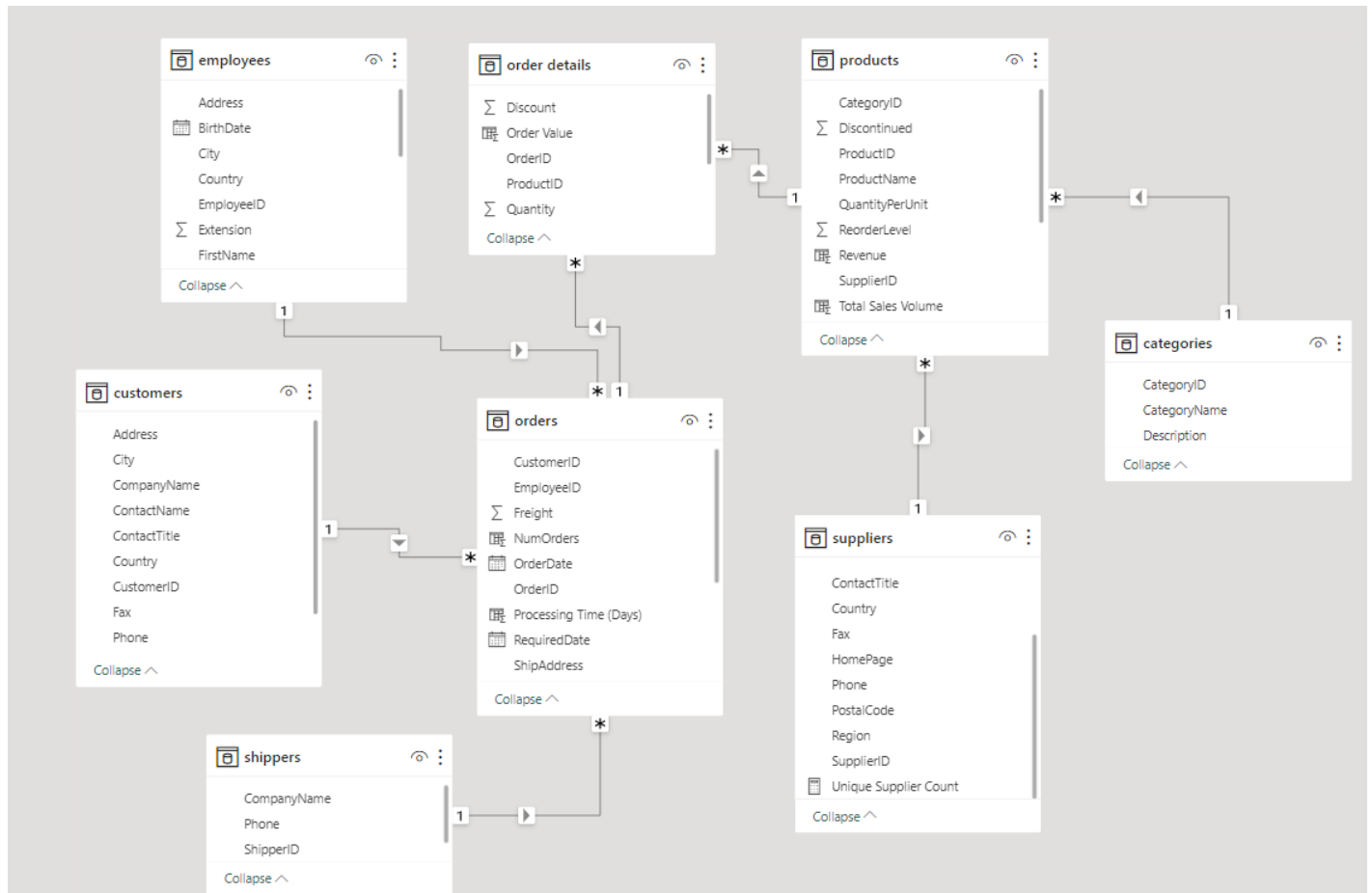
- **Shipper ID:** Unique identifier for each shipping company
- **Company Name:** Name of the shipping company
- **Phone:** Phone number of the shipping company
- **Order Count:** Total order count of company

Categories Table

- **Category ID:** Unique identifier for each product category
- **Category Name:** Name of the product category
- **Description:** Description of the product category

E-R Diagram

An Entity-Relationship (ER) diagram is a visual representation of the major entities within a system, along with their properties (attributes) and the relationships between these entities. Here's how you might describe the ER diagram for the Northwind Traders database



Steps to Connect to the Data

For EDA

1. **Retrieve Data from GitHub:** Navigate to the GitHub repository and download the required data files. These files are located in the CSV and SQL folders.
2. **Import SQL Queries to SQL Workbench:** Open SQL Workbench and import or copy-paste the SQL queries provided in the GitHub repository. These queries will help you create the necessary database and tables.

-
3. **Create Database and Tables:** Execute the imported SQL queries in SQL Workbench to create the database and tables.
 4. **Generate Required Tables Using Queries:** Create the necessary tables by writing and executing SQL queries in SQL Workbench. These tables should be structured according to your requirements for the EDA.
 5. **Export Tables to Excel:** Copy the tables created in SQL Workbench and paste them into an Excel spreadsheet.
 6. **Visualize Data in Excel:** Use Excel's data visualization tools to analyse the data and create visual representations.

For PowerBi

1. **Open Power BI Desktop:** Start by launching Power BI Desktop.
2. **Load CSV Data:**
 - Click on "Home" in the top menu, then select "Get Data".
 - In the dropdown menu, select "Text/CSV".
 - Navigate to the location of your CSV file, select it, and click "Open".
 - In the preview window, check if the data has been correctly identified. You can adjust the settings if necessary. Click "Load" to import the data.
3. **Data Transformation:**
 - Once the data is loaded, click on "Edit Queries" in the Home tab to open the Power Query Editor.
 - Here, you can perform various transformations like removing columns, changing data types, splitting columns, etc. Use the options in the "Transform" tab for these operations.
 - After making the necessary transformations, click "Close & Apply" in the Home tab to apply these changes.
4. **Create a Dashboard:**
 - In the "Fields" pane, you'll see your tables and fields. You can drag and drop these onto the report canvas to create visuals.
 - Use the "Visualizations" pane to choose the type of visual (bar chart, pie chart, table, etc.). The visual will update automatically as you add fields.
 - You can add multiple visuals to create a dashboard. Use the "Format" pane to customize the appearance of your visuals.
 - Save your report by clicking "File" > "Save".
5. **Answer Questions with Dashboard:**
 - Use the created dashboard to answer the given questions. You can filter and slice the data in different ways to get the required insights.

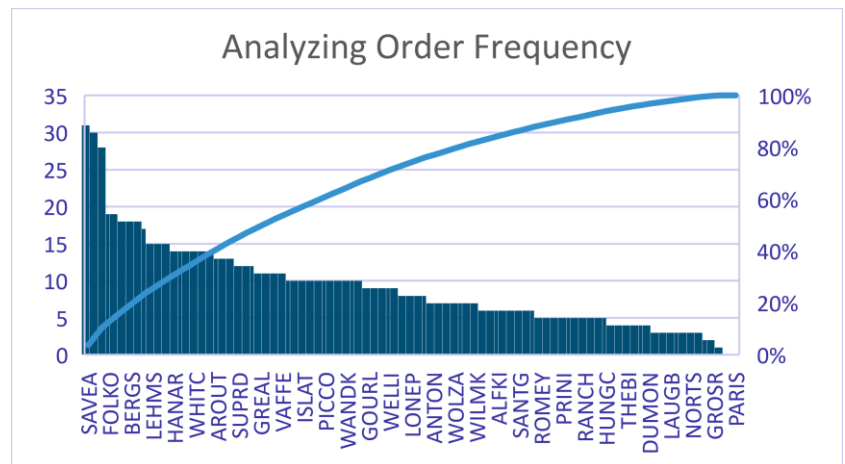
Problem Statement with Solution and Insights

EDA Questions

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

➡ To identify key factors influencing customer retention or loyalty based on the Northwind database, we use SQL queries to extract relevant insights from the dataset.

CustomerID	CompanyName	OrderCount
SAVEA	Save-a-lot Markets	31
ERNSH	Ernst Handel	30
QUICK	QUICK-Stop	28
FOLKO	Folk och fä HB	19
HUNGO	Hungry Owl All-Night Grocers	19
HILAA	HILARIÄN-Abastos	18
BERGS	Berglunds snabbköp	18
RATTC	Rattlesnake Canyon Grocery	18
BONAP	Bon app'	17
LEHMS	Lehmanns Marktstand	15
FRANK	Frankenversand	15
WARTH	Wartian Herkku	15
HANAR	Hanari Carnes	14
BOTTM	Bottom-Dollar Markets	14
LILAS	LILA-Supermercado	14
WHITC	White Clover Markets	14
KOENE	Königlig Essen	14
LAMAI	La maison d'Asie	14
AROUT	Around the Horn	13
QUEEN	Queen Cozinha	13
MEREP	Mère Paillarde	13
SUPRD	Suprêmes d'Olives	12
REGGC	Reggiani Caseifici	12
LINOD	LINO-Delicatesses	12
GREAL	Great Lakes Food Market	11
BLONP	Blondel père et fils	11



Row Labels	Sum of ProductCount
Beverages	404
Condiments	216
Confections	334
Dairy Products	366
Grains/Cereals	196
Meat/Poultry	173
Produce	136
Seafood	330
(blank)	0



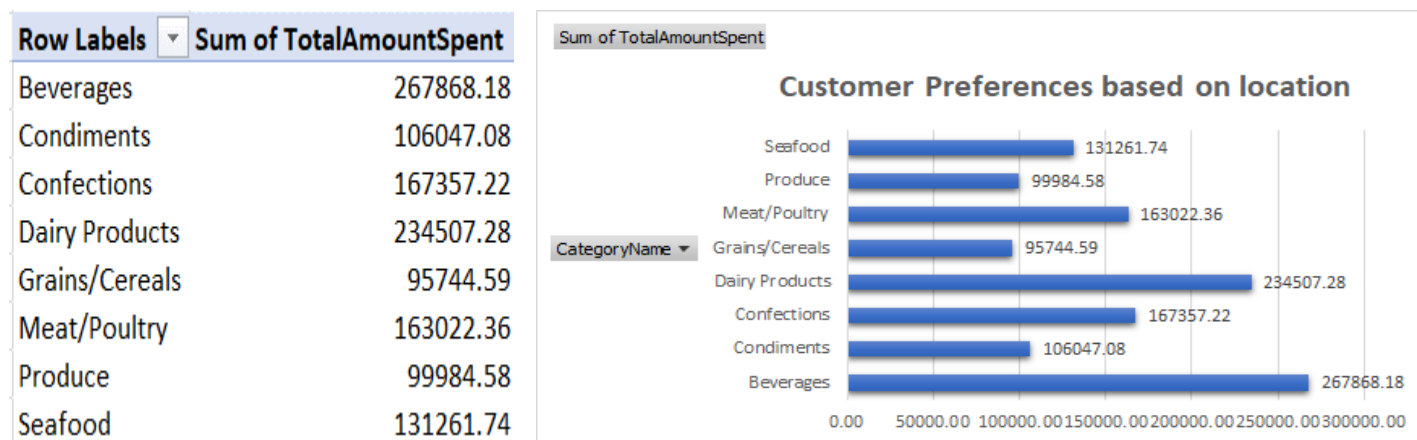
This is how we can get Key factors

Insight

1. **Order Frequency:** Customers who place orders more frequently are likely to be more loyal.
2. **Product Categories:** Customer loyalty may be influenced by specific product categories that customers purchase from most frequently.
3. **Discounts:** Customers who receive larger discounts tend to place more orders, suggesting that discounts may impact customer loyalty.
4. **Employee Influence:** Interactions with specific employees could affect customer loyalty, with some employees having more interactions with customers.
5. **Product Availability:** The availability of products, including stock levels and whether products are discontinued, could influence customer loyalty.
6. **Shipping Experience:** The performance of shipping companies, particularly those with on-time deliveries, could influence customer loyalty.

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

➡ To explore how customer preferences vary based on their location or demographics and create interactive visualizations, you can use SQL to extract relevant data.



Insight

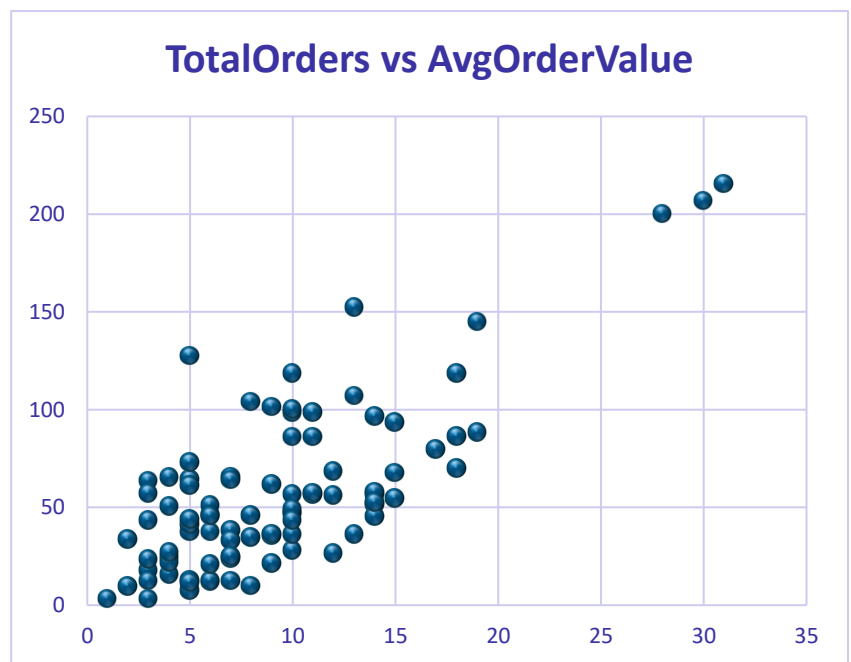
We can extract customer preferences vary based on their location or demographic by using SQL queries from the Northwind database. we need data from the "Customers," "Orders," "Order Details," and "Products" tables .by utilize SQL to calculate customer preferences based on relevant factors such as product categories, order history, discounts, and location. For instance, you can calculate the total amount spent by customers in each product category.

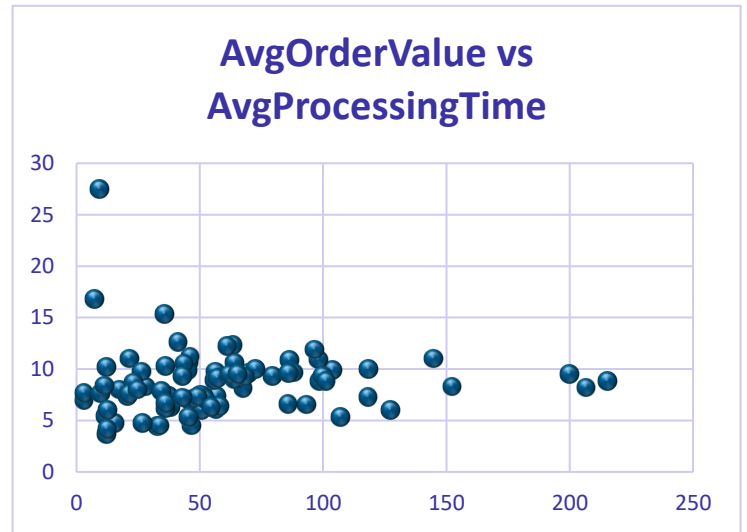
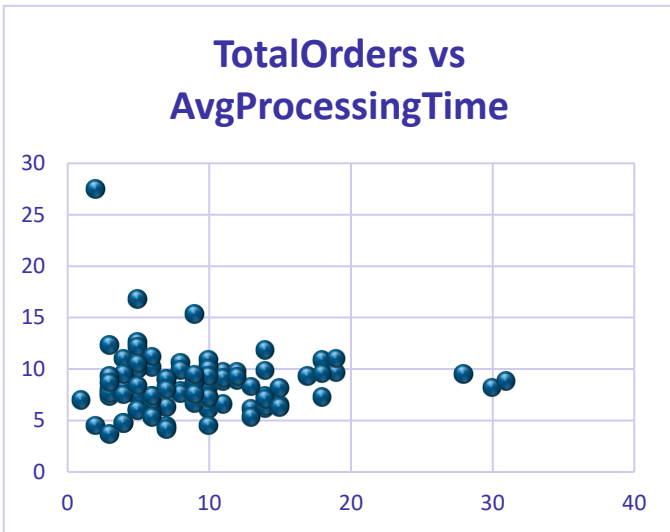
To answering the second question we can definately explore this through interactive visualisations. By using pivot table and slicer as well as chart we can get desired visualisation.

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

➡ To identify interesting patterns or clusters in customer behavior and visualize potential market segments using the Northwind dataset, we use SQL to analyze the data and Excel for visualization. Here's how to do it

CustomerID	TotalOrders	AvgOrderValue	AvgProcessingTime
ALFKI	6	37.59666667	7.3333
ANATR	4	24.355	8.25
ANTON	7	38.36	6.2857
AROUT	13	36.30384615	6.1538
BERGS	18	86.64	10.8333
BLAUS	7	24.03714286	8.1667
BLONP	11	56.69636364	9.7273
BOLID	3	63.72333333	12.3333
BONAP	17	79.87470588	9.3125
BOTTM	14	56.71071429	6.1538
BSBEV	10	28.131	8.3
CACTU	6	12.12666667	5.4
CENTC	1	3.25	7
CHOPS	8	45.905	10.625
COMMI	5	37.564	7.4
CONSH	3	17.87333333	8
DRACD	6	51.00666667	6
DUMON	4	15.925	4.75
EASTC	8	104.0425	9.875
ERNSH	30	206.8463333	8.1786
FAMIA	7	33.25	4.4286
FOLIG	5	127.588	6
FOLKO	19	88.32	9.6316
FRANK	15	93.56266667	6.5333
FRANR	3	57.14	7.3333





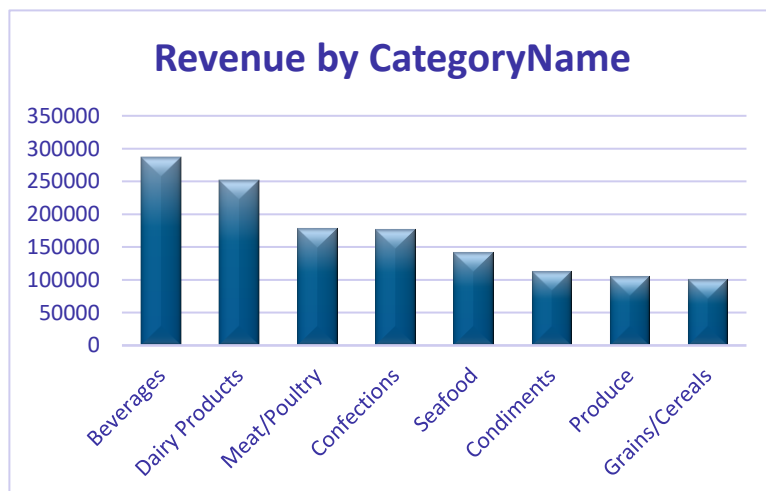
Insight

From these plots, we can see some patterns and clusters which could represent different market segments. For example, customers with high total orders and high average order value could be a potential market segment. Similarly, customers with low total orders but high average processing time could be another market segment.

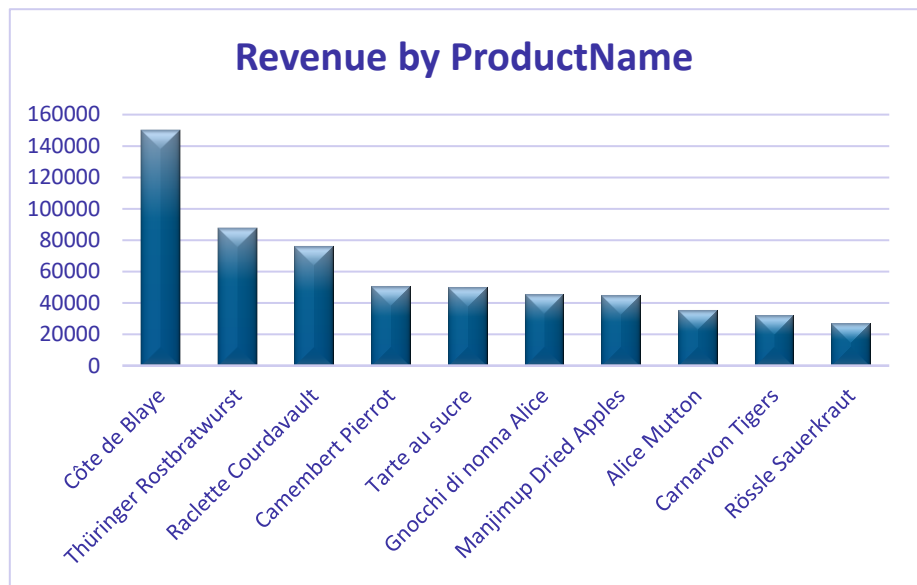
4.Are there any specific product categories or SKUs that contribute significantly to order revenue? Can we identify them through visualizations?

➡ To identify specific product categories or SKUs that contribute significantly to order revenue and visualize them using SQL and Excel, we can follow these steps:

CategoryName	Revenue
Beverages	286526.95
Dairy Products	251330.5
Meat/Poultry	178188.8
Confections	177099.1
Seafood	141623.09
Condiments	113694.75
Produce	105268.6
Grains/Cereals	100726.8



ProductNam	Revenue
Côte de Blaye	149984.2
Thüringer Rostbratwurst	87736.4
Raclette Courdavault	76296
Camembert Pierrot	50286
Tarte au sucre	49827.9
Gnocchi di nonna Alice	45121.2
Manjimup Dried Apples	44742.6
Alice Mutton	35482.2
Carnarvon Tigers	31987.5
Rössle Sauerkraut	26865.6



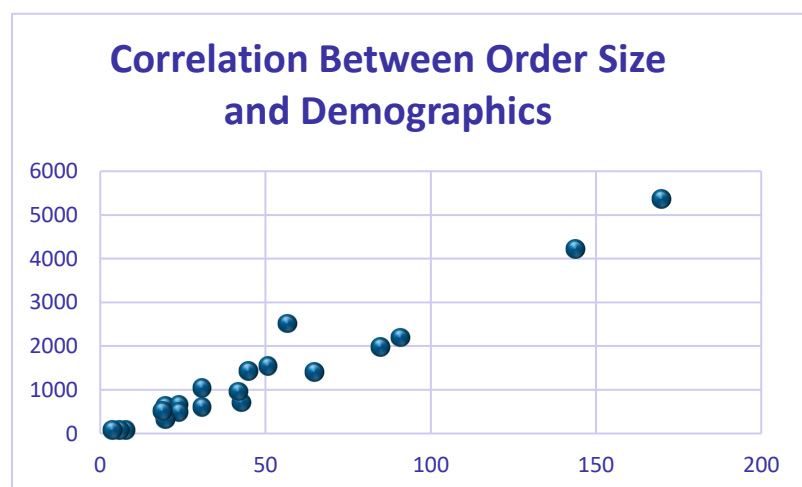
Insight

These visualizations provide a clear picture of which product categories and SKUs contribute significantly to the total order revenue.

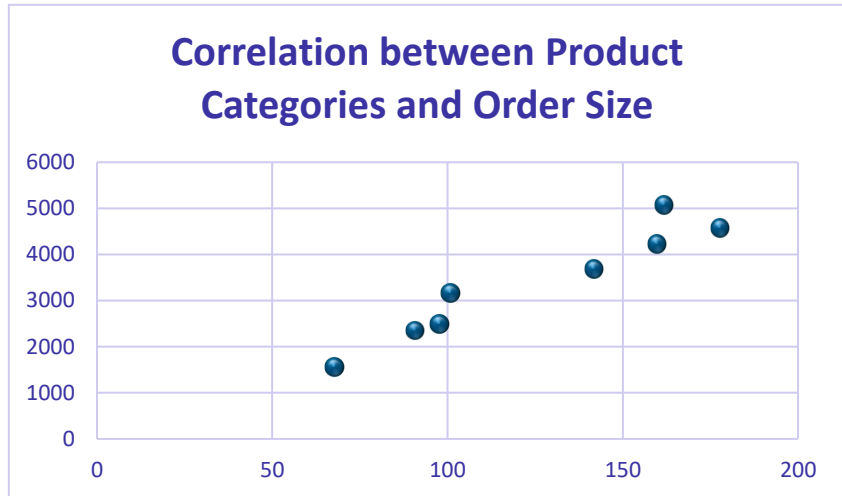
5.Are there any correlations between order size and customer demographics or product categories? Can we explore this visually using scatter plots or heatmaps?

➡To explore correlations between order size (order quantity) and customer demographics or product categories and visualize this using scatter plots or heatmaps, we use SQL to analyze the data and Excel for visualization. Here are the steps to do this:

Count of CustomerCountry	Sum of OrderSize
8	71
57	2510
20	617
91	2194
45	1420
24	656
31	599
85	1972
170	5359
31	1033
20	380
43	706
6	69
4	76
24	477
20	309
42	957
19	510
65	1401
144	4220
51	1532



Count of ProductCategory	Sum of OrderSize
178	4571
101	3154
160	4216
162	5066
98	2484
91	2350
68	1551
142	3676



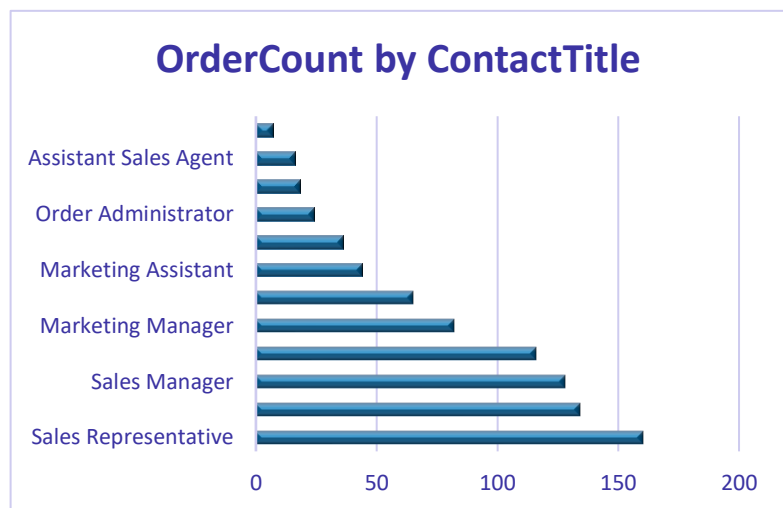
Insight

Yes there is a strong positive correlation between order size and customer demographics as well as product categories , we can visually see then via scatter plot.

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

➡To analyze how order frequency varies across different customer segments and visualize this using bar charts or treemaps, you can use SQL to extract and summarize the data and Excel for visualization.

ContactTitle	OrderCount
Sales Representative	160
Owner	134
Sales Manager	128
Accounting Manager	116
Marketing Manager	82
Sales Associate	65
Marketing Assistant	44
Sales Agent	36
Order Administrator	24
Assistant Sales Representative	18
Assistant Sales Agent	16
Owner/Marketing Assistant	7



Insight

From the chart, we can see that the 'Sales Representative' segment has the highest order frequency, followed by 'Owner', 'Sales Manager', 'Accounting Manager', and 'Marketing Manager'.

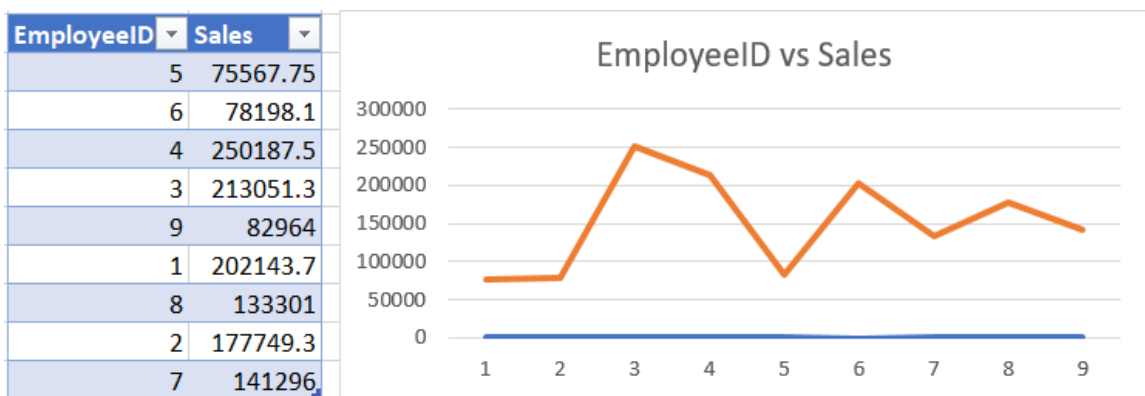
7.Are there any correlations between employee satisfaction levels and key performance indicators?

➡ Analyzing the correlation between employee satisfaction levels and key performance indicators (KPIs) in the Northwind dataset can provide valuable insights. To explore this visually using scatter plots or line charts, we can use SQL to extract and summarize the data and Excel for visualization.

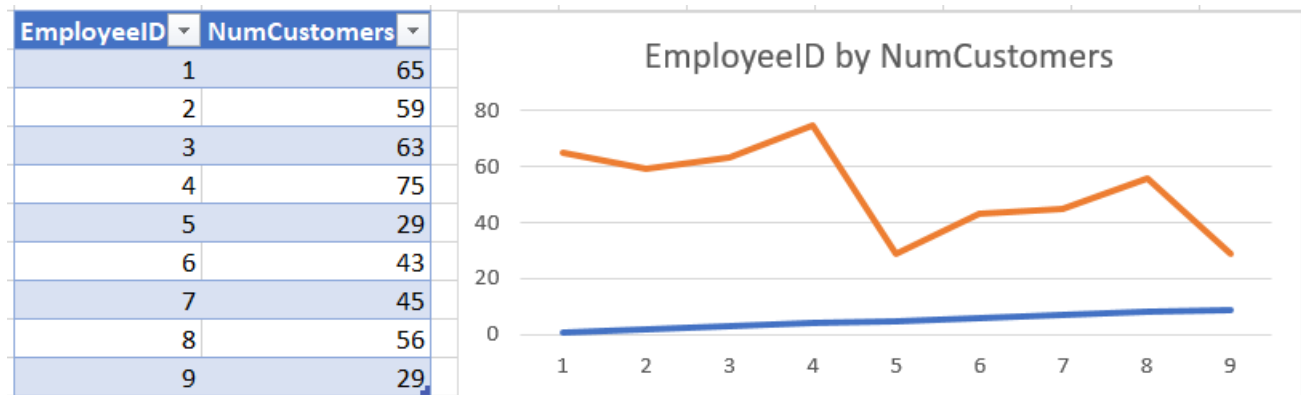
Number of order handled by each employee



Total sales generated by each employee



Number of unique customers served by each employee



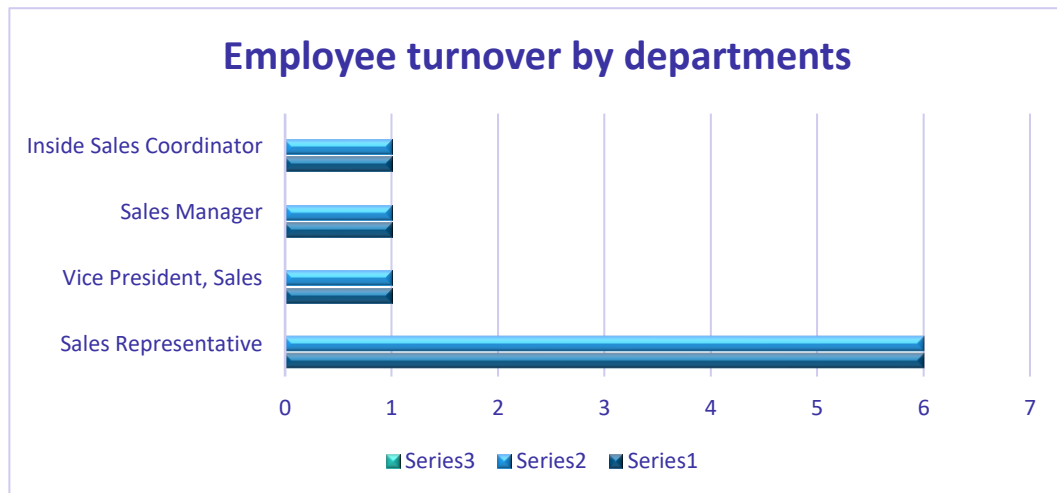
Insight

This analysis provides valuable insights into the performance of each employee. However, please note that these are just raw performance indicators. To get a more accurate picture of each employee's performance, we might need to consider other factors such as the time period over which these sales were made, the difficulty level of each sale, the employee's role and responsibilities, etc.

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

➡To analyze how employee turnover varies across different departments or job roles in the Northwind dataset and visualize this using bar charts, you can use SQL to extract and summarize the data, and then use Excel for visualization.

JobTitle	TotalEmployees	EmployeesJoinedBefore2022	EmployeesJoinedIn2022	
Sales Representative	6	6	0	66.67%
Vice President, Sales	1	1	0	11.11%
Sales Manager	1	1	0	11.11%
Inside Sales Coordinator	1	1	0	11.11%



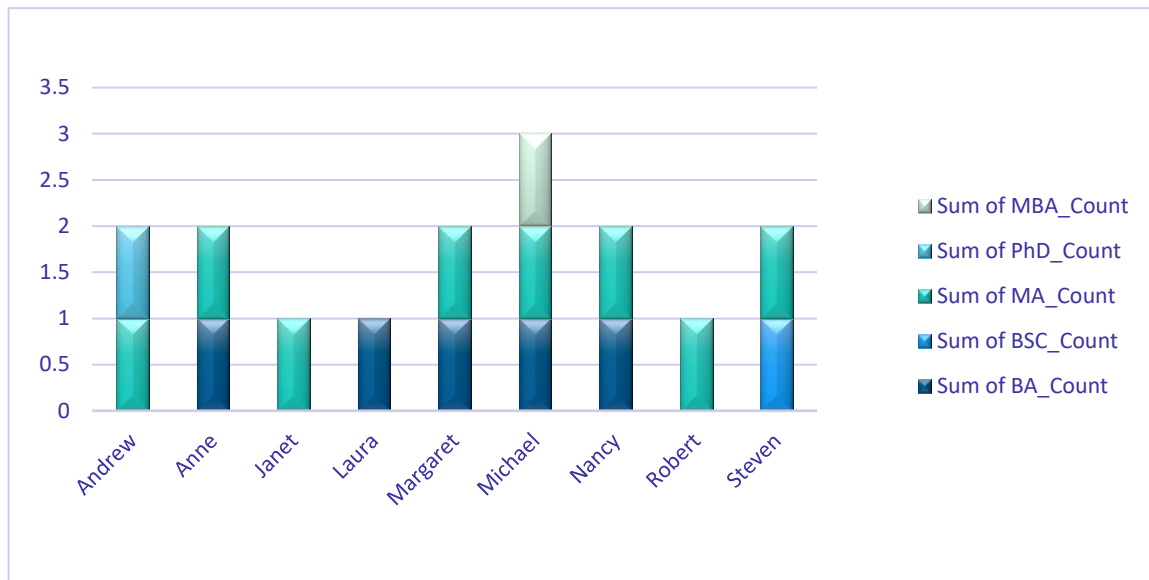
Insight

We can see that 'Sales Representative' turnover rate is more than any other role and we also can visualize this using clustered bar chart. This will help you gain insights into the turnover patterns within the organization.

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

➡ To identify patterns or clusters in employee skill sets or qualifications through visualizations and use this information for talent management, we can follow these steps using SQL and Excel.

EmployeeID	FirstName	LastName	BA_Count	BSC_Count	MA_Count	PhD_Count	MBA_Count
1	Nancy	Davolio	1	0	1	0	0
2	Andrew	Fuller	0	0	1	1	0
3	Janet	Leverling	0	0	1	0	0
4	Margaret	Peacock	1	0	1	0	0
5	Steven	Buchanan	0	1	1	0	0
6	Michael	Suyama	1	0	1	0	1
7	Robert	King	0	0	1	0	0
8	Laura	Callahan	1	0	0	0	0
9	Anne	Dodsworth	1	0	1	0	0



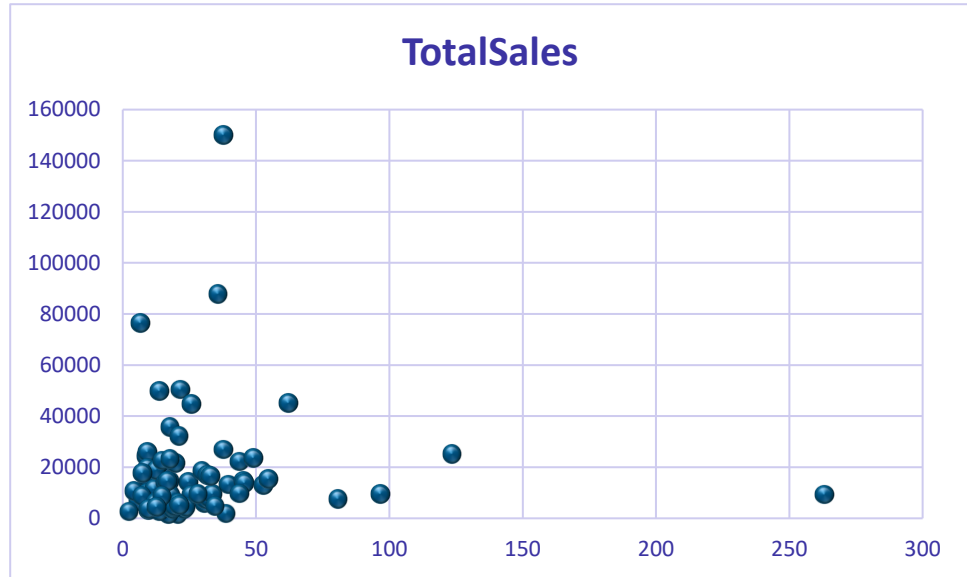
Insight

- 1.Examine the visualizations to identify the most common qualifications and skills among employees. These are your areas of strength, and you can use this information to.
- 2.Identify less common qualifications or skills. These are potential areas of improvement or gaps, and you can.
- 3.Identify employees who have unique qualifications or skills that are valuable to your organization. Consider these individuals for promotions or leadership roles, as their expertise can be a valuable asset

10. Are there any correlations between product attributes (e.g., size, color, features) and sales performance? Can we explore this visually using scatter plots or heatmaps?

➡ To Identify this we need to check data, Due to the limitation of data we don't have specific attributes like size, color, or features for the products.We can still analyze the correlation between available attributes and sales performance. For instance, we can check if there's a correlation between 'UnitPrice' and the quantity sold, or if the 'UnitsInStock' affects sales performance.

UnitPrice	TotalSales
18	35482.2
19	3080
10	19048.3
22	50286
21.35	31987.5
25	14277.6
30	18559.2
40	13150.8
97	9424.8
31	5801.15
21	1542.75
38	149984.2
6	6664.75
23.25	3383.8
15.5	20876.5
17.45	1713.5
39	1813.5
62.5	45121.2
9.2	16172.5
81	7345
10	3047.2
21	4782.6
9	24307.2
4.5	10524.2
14	21534.9



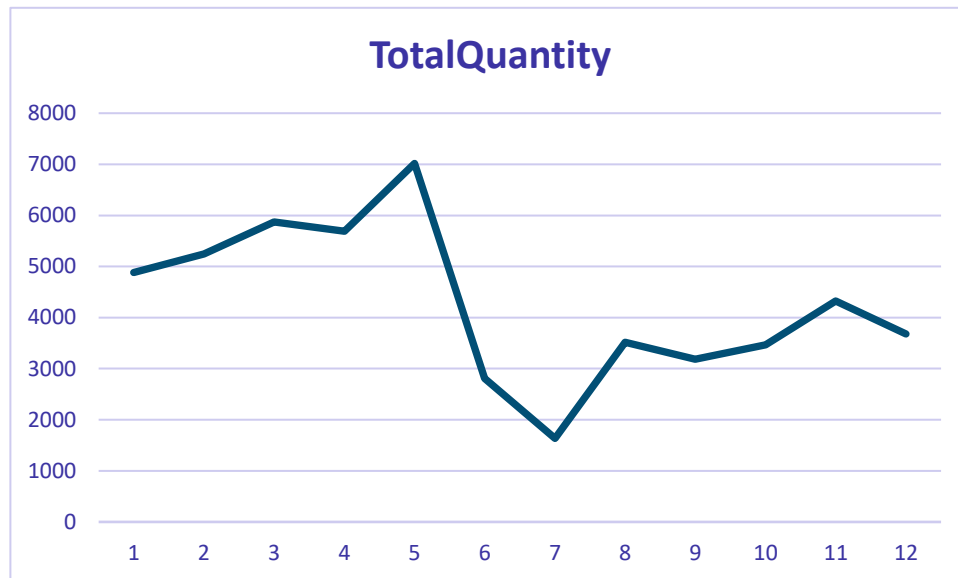
Insight

The table displays the sales performance of different products, which could be used to identify which products are performing well and which are not. The highlighted text suggests that the user is interested in visually exploring correlations between product attributes and sales performance using scatter plots or heatmaps. This could be a valuable approach for identifying key factors that drive sales.

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

➡ Definitely we can visualize a drastic downfall from the month of may to july in the demand of products through the line chart.

Months	TotalQuantity
1	4882
2	5244
3	5870
4	5687
5	7017
6	2808
7	1635
8	3516
9	3183
10	3467
11	4326
12	3682



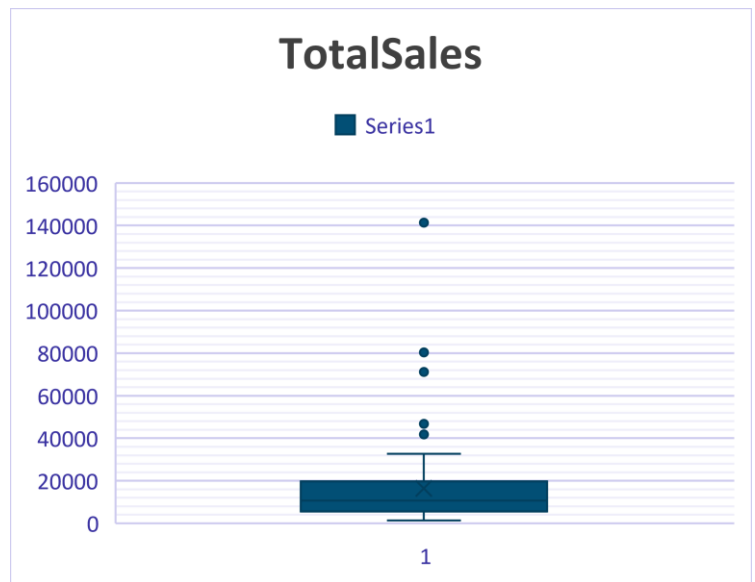
Insight

- 1.Sales Seasonality: By grouping the total quantity of orders by month, you can identify any seasonal trends in the company's sales. For example, some months might have higher sales due to seasonal demand for certain products.
- 2.Inventory Management: This information can also help with inventory management. If certain months have higher sales, the company can ensure they have sufficient stock during these months to meet the increased demand.
- 3.Sales Performance: Comparing the total quantity of orders month by month can give an indication of the company's sales performance over time. Any significant increases or decreases in sales can be further investigated to identify the causes.

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

➡ We can identify the outliers by visualize the box plot and wishker chart.
we can discontinue the products that are outliers.

ProductID	ProductName	TotalSales
1	Chai	12788.09998
2	Chang	16355.95997
3	Aniseed Syrup	3043.999999
4	Chef Anton's Cajun Seasoning	8567.899988
5	Chef Anton's Gumbo Mix	5347.199996
6	Grandma's Boysenberry Spread	7136.999999
7	Uncle Bob's Organic Dried Pears	22044.29999
8	Northwoods Cranberry Sauce	12771.99999
9	Mishi Kobe Niku	7226.499988
10	Ikura	20867.33997
11	Queso Cabrales	12901.76999
12	Queso Manchego La Pastora	12257.65998
13	Konbu	4960.439991
14	Tofu	7991.489994
15	Genen Shouyu	1784.825
16	Pavlova	17215.77547
17	Alice Mutton	32698.37998
18	Carnarvon Tigers	29171.87496
19	Teatime Chocolate Biscuits	5862.619996
20	Sir Rodney's Marmalade	22563.35999
21	Sir Rodney's Scones	9103.999991
22	Gustaf's Knäckebröd	7122.36



Insight

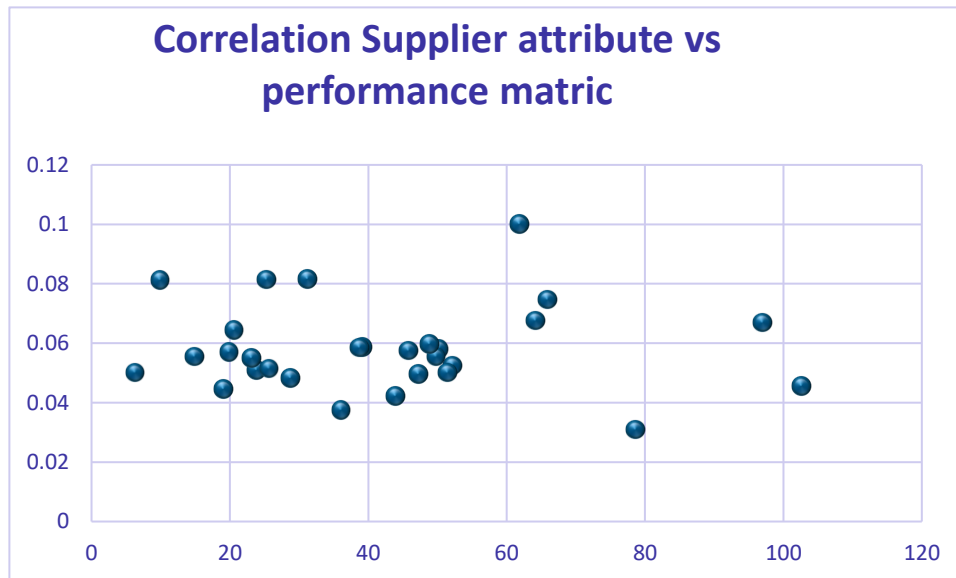
We calculate the total sales for each product in the 'Products' table. It does this by summing the product of unit price, quantity, and discount from the 'Order Details' table for each product. This query can provide insights into which products generate the most revenue, helping to identify best-selling products and inform business strategies.

13. Are there any correlations between supplier attributes (e.g., location, size, industry) and performance metrics (e.g., on-time delivery, product quality)?

➡ we can see that there is a negative correlation between them. We use the `correl()` function in Excel to check the correlation between them.

Correlation = -0.0336405496139262

SupplierID	CompanyName	Country	AvgUnitsInStock	AvgDiscount
1	Exotic Liquids	UK	25.383	0.08138298
2	New Orleans Cajun Delights	USA	50.3429	0.057857144
3	Grandma Kelly's Homestead	USA	36.1667	0.037407408
4	Tokyo Traders	Japan	23.9216	0.050980393
5	Cooperativa de Quesos 'Las Cabras'	Spain	39.2308	0.058653847
6	Mayumi's	Japan	28.8824	0.048235295
7	Pavlova, Ltd.	Australia	20.6994	0.064294479
8	Specialty Biscuits, Ltd.	UK	14.9683	0.055476191
9	PB Knäckebröd AB	Sweden	78.7059	0.030882353
10	Refrescos Americanas LTDA	Brazil	20	0.056862746
11	Heli Süßwaren GmbH & Co. KG	Germany	38.7966	0.058474577
12	Plutzer Lebensmittelgroßmärkte AG	Germany	47.3966	0.049608939
13	Nord-Ost-Fisch Handelsgesellschaft mbH	Germany	10	0.08125
14	Formaggi Fortini s.r.l.	Italy	6.4135	0.050000001
15	Norske Meierier	Norway	52.2095	0.052380953
16	Bigfoot Breweries	USA	51.5231	0.050000001
17	Svensk Sjöföda AB	Sweden	97.098	0.066862746
18	Aux joyeux ecclésiastiques	France	45.8889	0.057407408
19	New England Seafood Cannery	USA	102.7045	0.045454546
20	Leka Trading	Singapore	23.2195	0.05487805



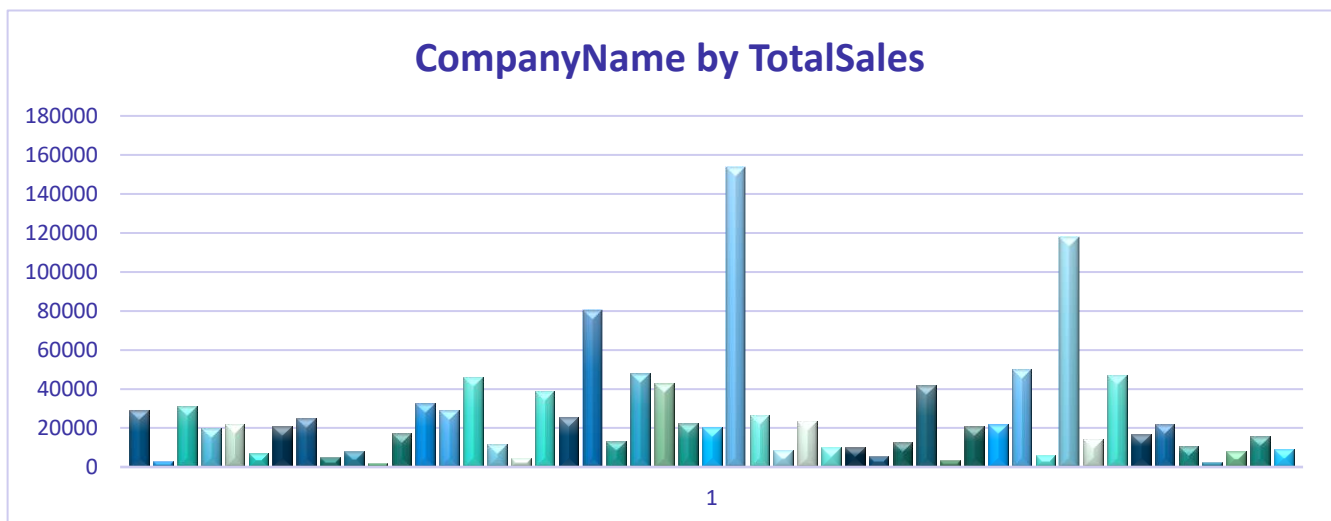
Insight

calculates the average units in stock for each supplier in the 'Suppliers' table. It does this by joining the 'Suppliers' table with the 'Products' table. This can provide insights into the inventory levels maintained by each supplier, which can be useful for managing stock and ensuring a steady supply of products. It also includes the country of each supplier, which could provide insights into geographical trends in inventory levels.

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

➡ To analyze supplier performance across different product categories, we can calculate the total sales for each supplier and category. We can definitely visualize this using a stacked bar chart.

SupplierID	CompanyName	CategoryID	CategoryName	TotalSales
1	Exotic Liquids	1	Beverages	29144.05995
1	Exotic Liquids	2	Condiments	3043.999999
2	New Orleans Cajun Delights	2	Condiments	31167.98996
3	Grandma Kelly's Homestead	2	Condiments	19908.99999
3	Grandma Kelly's Homestead	7	Produce	22044.29999
4	Tokyo Traders	6	Meat/Poultry	7226.499988
4	Tokyo Traders	8	Seafood	20867.33997
5	Cooperativa de Quesos 'Las Cabras'	4	Dairy Products	25159.42997
6	Mayumi's	8	Seafood	4960.439991
6	Mayumi's	7	Produce	7991.489994
6	Mayumi's	2	Condiments	1784.825
7	Pavlova, Ltd.	3	Confections	17215.77547
7	Pavlova, Ltd.	6	Meat/Poultry	32698.37998
7	Pavlova, Ltd.	8	Seafood	29171.87496
8	Specialty Biscuits, Ltd.	3	Confections	46243.97997
9	PB Knäckebröd AB	5	Grains/Cereals	11724.05999
10	Refrescos Americanas LTDA	1	Beverages	4504.364997



Insight

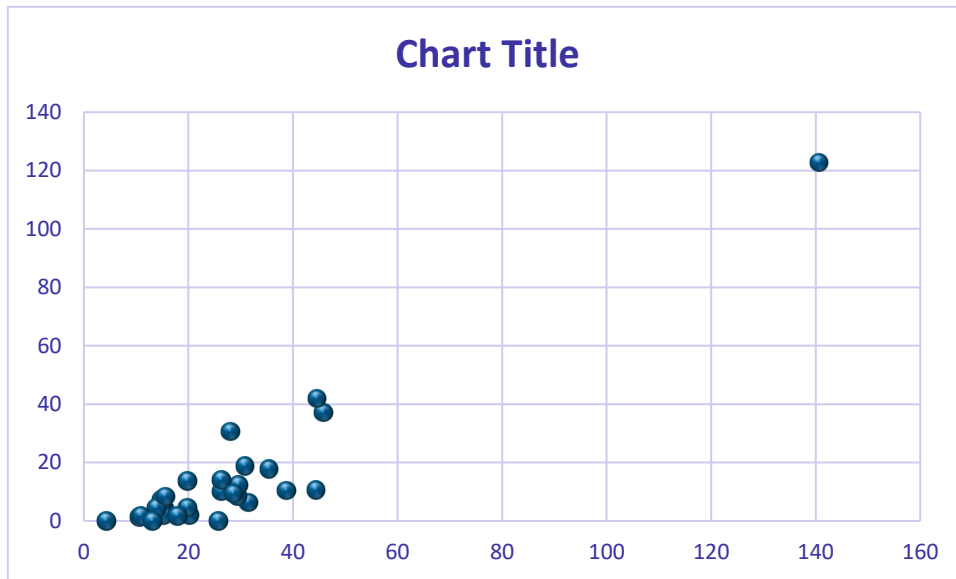
It shows the total sales for different companies. Company 1 and Company 2 have the highest sales, indicating they might be the market leaders or most preferred by customers. On the other hand, Company 3 and Company 4 have the lowest sales,

suggesting they might be facing challenges in the market or are less preferred by customers. This graph can help in understanding market dynamics and customer preferences.

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

➡ We can conclude from the visualization that suppliers with lower average prices and lower variability might be more reliable in terms of cost management. On the other hand, suppliers with higher variability might offer opportunities for cost savings when their prices are at the lower end of their range.

SupplierID	CompanyName	AvgUnitPrice	StdDevUnitPrice
1	Exotic Liquids	15.66666667	4.027681991
2	New Orleans Cajun Delights	20.35	1.964370128
3	Grandma Kelly's Homestead	31.66666667	6.236095645
4	Tokyo Traders	46	37.06750599
5	Cooperativa de Quesos 'Las Cabras'	29.5	8.5
6	Mayumi's	14.91666667	7.054352478
7	Pavlova, Ltd.	35.57	17.65127757
8	Specialty Biscuits, Ltd.	28.175	30.52280909
9	PB Knäckebröd AB	15	6
10	Refrescos Americanas LTDA	4.5	0
11	Heli Süßwaren GmbH & Co. KG	29.71	12.2538511
12	Plutzer Lebensmittelgroßmärkte AG	44.678	41.85052133
13	Nord-Ost-Fisch Handelsgesellschaft mbH	25.89	0
14	Formaggi Fortini s.r.l.	26.43333333	9.918445219
15	Norske Meierier	20	13.71738556
16	Bigfoot Breweries	15.33333333	1.885618083
17	Svensk Sjöföda AB	20	4.546060566
18	Aux joyeux ecclésiastiques	140.75	122.75
19	New England Seafood Cannery	14.025	4.375
20	Leka Trading	26.48333333	13.97857488



Insight

The data points do not seem to follow a linear trend, suggesting that the two variables on the x and y axes may not be correlated, or their relationship is non-linear. There's an outlier data point at (140, 120), which could be an exception or a data entry error. Without more context or access to the actual data, these are just preliminary observations. user Give me Meaning full insight of it.

PowerBI Questions

Customer Analysis

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

➡ By using either a bar chart or a map visualization, you can effectively display how customer distribution varies across different regions or customer segments in Power BI. The choice between a bar chart and a map depends on your data and the visual representation that best conveys the information to your audience.

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

➡ The line chart shows the number of new customers acquired each year. We can see that the number of new customers has been fluctuating over the years.

3. Can we visualize the distribution of customer demographics such as age, gender, or income using histograms or pie charts?

➡ The provided datasets do not contain demographic information such as age, gender, or income for the customers. The available customer-related information is limited to company name, contact name, contact title, address, city, region, postal code, country, and phone.

While we don't have specific demographic data like age, gender, or income, we can still explore and visualize other aspects of the customer data that we do have. For example, we can look at the distribution of customers by country or city.

NorthWind_Sales_Data, Customer Analysis

Q1.Customer distribution Across Region



Region

All

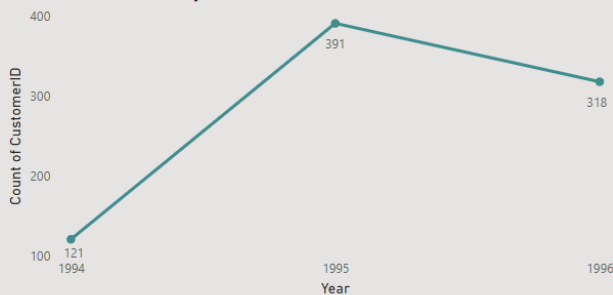
Country

All

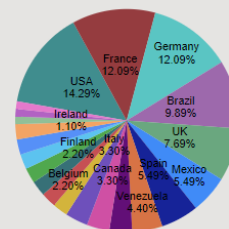
City

All

Q2. Customer Acquisition Over Time



Q3.Distribution of customers by country



Order Analysis

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

➡ It appears there was an error due to the 'Period' data type of the 'YearMonth' column. To resolve this, we can convert the 'Period' data type to 'datetime' before plotting. The line chart shows the number of orders placed each month. We can see that the order volume has been fluctuating over time.

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

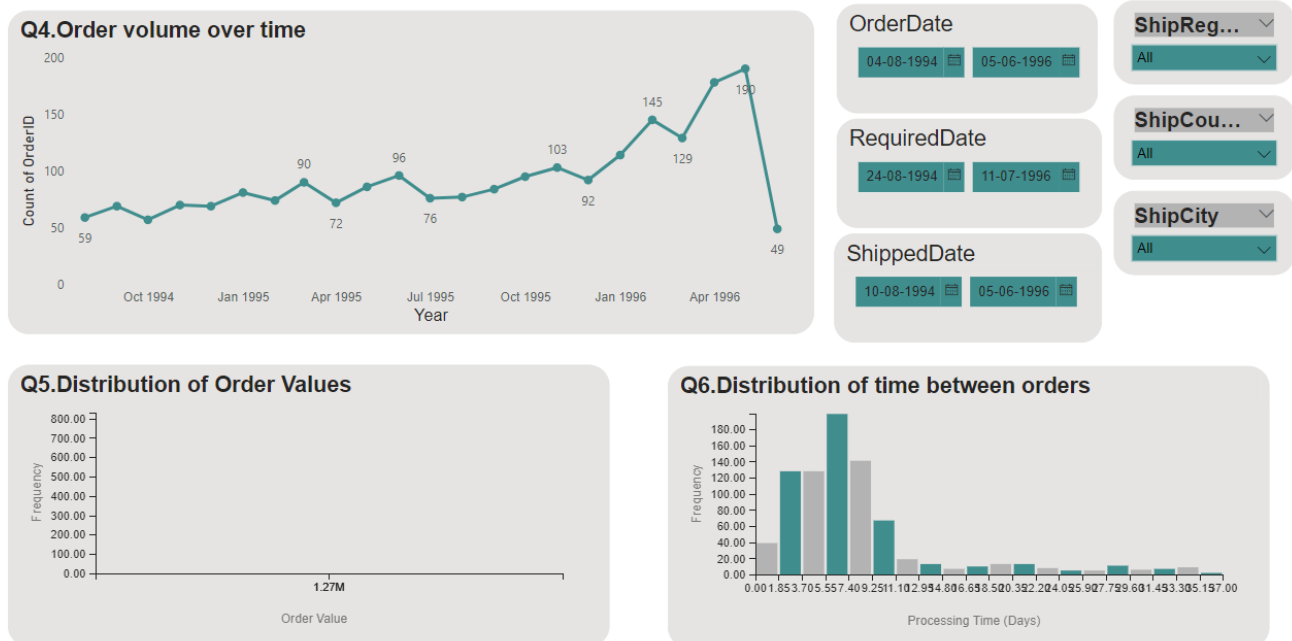
➡ The histogram shows the frequency of different order values. We can see that most of the order values are concentrated on the lower end, indicating that smaller orders are more common.

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

➡ To visualize the average order processing time or shipping duration, we first need to calculate these metrics. The order processing time can be calculated as the difference between the order date and the shipped date. However, the provided datasets do not contain a shipped date, so we cannot calculate this directly.

As an alternative, if we assume that the order processing time is constant, we can analyze the trend of orders over time to infer the shipping duration. This would involve calculating the time difference between consecutive orders for each customer.

NorthWind_Sales_Data, Orders Analysis



Employee Analysis

7. How does employee productivity vary across different departments or job roles?
Can we create a stacked bar chart or grouped column chart to visualize it?

➡ To analyze employee productivity across different departments or job roles, we need data that includes information about employees' departments or job roles and their productivity measures. However, the provided datasets do not contain this information.

The 'employees.csv' file contains information about employees, but it does not specify their departments or job roles. The 'orders.csv' and 'order details.csv' files contain information about orders, but they do not specify which employee processed each order. We need to merge the orders and order details dataframes to get the total value of each order. Then, we can merge this with the employees dataframe to get the employee details for each order. After that, we can calculate productivity measures such as the number of orders processed by each employee and the total value of orders processed by each employee. Finally, we can visualize these measures using a bar chart or grouped column chart.

8. What is the distribution of employee tenure? Can we create a histogram or box plot to display it?

➡ The tenure of each employee in years, calculated as the difference between the current date (2023-11-08) and the hire date of each employee.

9. Can we visualize employee performance ratings or KPIs using a radar chart or bullet graph?

➡ we can visualize employee performance ratings or KPIs using a radar chart or bullet graph in PowerBI. However, PowerBI does not natively support radar charts or bullet graphs. Instead, we can use a combination of bar charts, line charts, and card visuals to create a similar effect.

Before we proceed, we need to define the performance ratings or KPIs for the employees. These could be based on various factors such as the number of orders processed, the total value of orders processed, the tenure of the employee, etc.

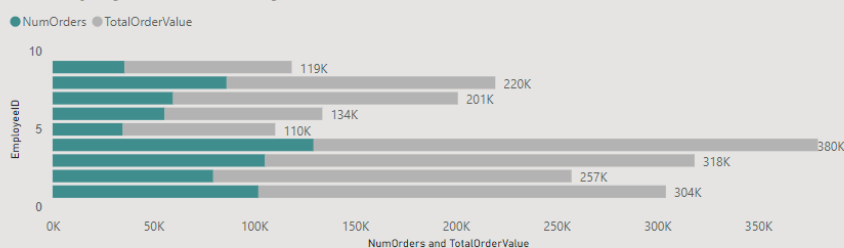
NorthWind_Sales_Data, Employee Analysis

Region
All

Country
All

City
All

Q7.Employee Productivity



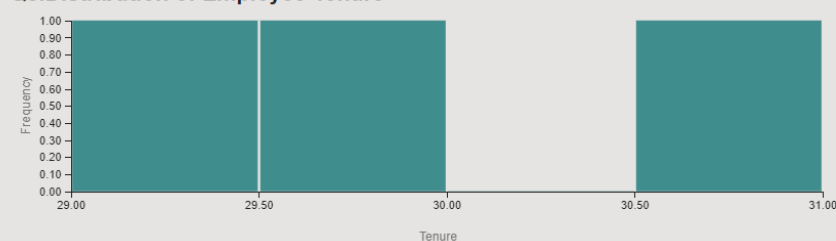
689K

NumOrders

1.35M

TotalOrderValue

Q8.Distribution of Employee Tenure



689K

Sum of NumOrders

1.35M

Sum of TotalValue

Product Analysis

10.What is the distribution of product ratings or reviews? Can we create a histogram or stacked bar chart to visualize it?

➡Unfortunately, there doesn't seem to be a column that directly represents product ratings or reviews in this dataframe either. We might need to infer this information from other dataframes or columns. However, without any explicit rating or review data, it's not possible to create a distribution of product ratings or reviews. we can create a distribution of other product-related metrics with the available data. For instance, we can use the 'Quantity' column in the 'order_details' dataframe to create a distribution of product quantities sold. This can give us an idea of which products are most popular.

11.How does the sales volume vary across different product categories? Can we create a bar chart or treemap to display it?

➡Merge the 'products' dataframe with the 'categories' dataframe to get the category for each product.Merge the resulting dataframe with the 'order_details'

dataframe to get the sales volume for each product. Group the resulting dataframe by 'CategoryID' and sum the 'Quantity' for each category.

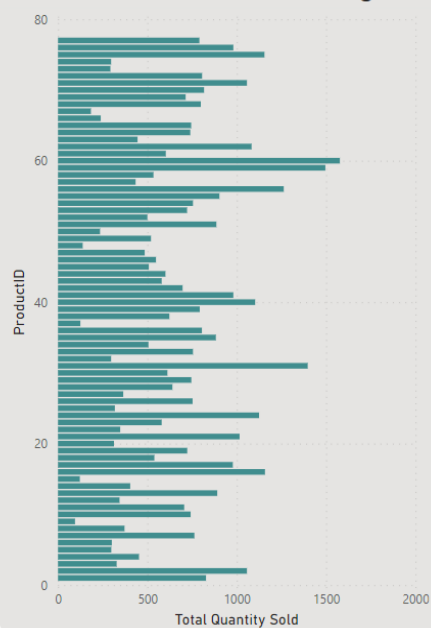
12. Can we visualize the pricing distribution of products using a box plot or violin plot?

➡ We can visualize the pricing distribution of products using a box plot or violin plot.

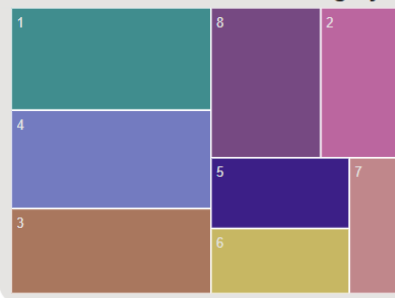
However, Power BI does not natively support violin plots. Instead, we can use a histogram to visualize the distribution of prices. the histogram shows the frequency of different price ranges.

NorthWind_Sales_Data, Product Analysis

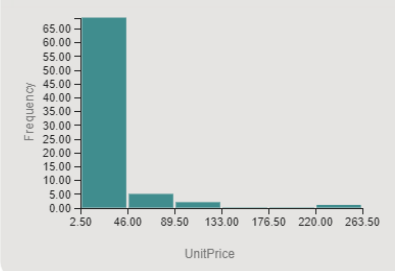
Q10. Distribution of Product Rating



Q11. Sales Volume Across Category



Q12. Price Distribution of Product



ProductName

- ☐ Alice Mutton
- ☐ Aniseed Syrup
- ☐ Boston Crab Meat
- ☐ Camembert Pierrot
- ☐ Carnarvon Tigers
- ☐ Chai
- ☐ Chang
- ☐ Chartreuse verte
- ☐ Chef Anton's Cajun Seasoning
- ☐ Chef Anton's Gumbo Mix
- ☐ Chocolate
- ☐ Côte de Blaye
- ☐ Escargots de Bourgogne
- ☐ Filo Mix
- ☐ Fløtemysost
- ☐ Geitost
- ☐ Genen Shouyu
- ☐ Gnocchi di nonna Alice
- ☐ Gorgonzola Telino
- ☐ Grandma's Boysenberry Spread

Supplier Analysis

13. What is the distribution of supplier ratings or performance metrics? Can we create a bar chart or radar chart to visualize it?

➡ Unfortunately, this dataset does not contain any explicit supplier ratings or performance metrics. We might need to derive some metrics based on other available data. For instance, we could consider the number of products supplied, the total quantity of products sold, or the total revenue generated from each supplier as a proxy for their performance. We calculate the total quantity of products sold for each supplier. For this, we'll need to join the 'products' and 'order details' dataframes on 'ProductID', and then group by 'SupplierID'. We have calculated the number of products supplied by each supplier, the total quantity of products sold for each supplier, and the total revenue generated from each supplier's products.

14. How does the cost or pricing structure vary across different suppliers? Can we create a box plot or stacked bar chart to display it?

➡ To analyze the cost or pricing structure across different suppliers, we can use the 'UnitPrice' column in the 'products' table. To analyze the cost structure, we can calculate the average unit price for each supplier. In DAX, we can do this with the AVERAGEX function, which calculates the average result of an expression evaluated for each row in a table. The measure might look something like this:

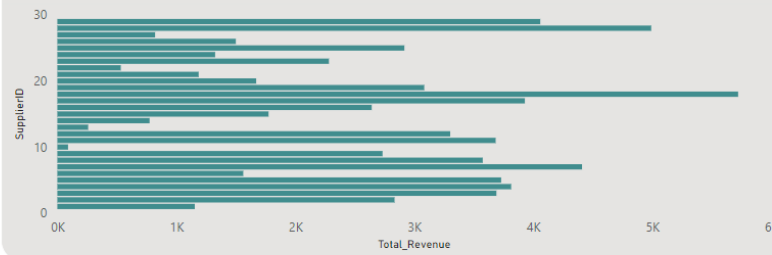
Average Unit Price = AVERAGEX('products', 'products'[UnitPrice])

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

➡ To visualize the geographical distribution of suppliers using a map or bubble chart in Power BI, you would need to have geographical data such as the latitude and longitude, or the city and country of the suppliers. However, based on the files provided, it seems we only have the 'Country' and 'Region' information for the suppliers.

NorthWind_Sales_Data, Suppliers Analysis

Q13. Distribution Of Total Revenue across Suppliers



CompanyName

- ☐ Aux joyeux ecclésiastiques
- ☐ Bigfoot Breweries
- ☐ Cooperativa de Que...
- ☐ Escargots Nouveaux
- ☐ Exotic Liquids
- ☐ Forêts d'érables
- ☐ Formaggi Fortini s.r.l.
- ☐ Gai pâturage
- ☐ G'day, Mate
- ☐ Grandma Kelly's Ma...

Country

All

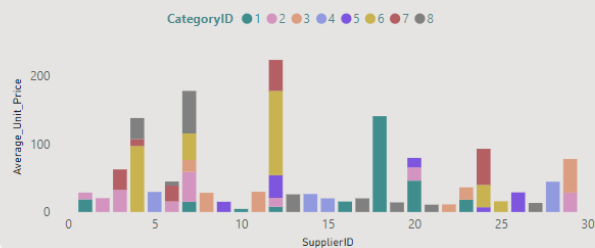
Region

All

City

All

Q14. Pricing across Suppliers



Q15. Geographical Distribution of Suppliers



Conclusion

The Northwind Sales data provides a wealth of information for analysis and visualization. Effective data visualization in Power BI enables stakeholders to gain valuable insights and make data-driven decisions in various aspects of their business, including customer behavior, sales trends, employee performance, and product optimization. The insights derived from this analysis can help Northwind Traders stay competitive and excel in the wholesale market landscape.