Assignment 1  
Analysing Company Performance with SQL

short line

**Bui The Hai**

Student ID: 24683423

September 8, 2023

94692 - Data Science Practise

Master of Data Science and Innovation

University of Technology of Sydney

Table of Contents

[1. Executive Summary 3](#_Toc143719974)

[2. Introduction 3](#_Toc143719975)

[3. Data Understanding 4](#_Toc143719976)

[4. Business Question 1 5](#_Toc143719977)

[a. Description: 5](#_Toc143719978)

[b. Results 5](#_Toc143719979)

[c. Key insights and Findings: 6](#_Toc143719980)

[5. Business Question 2 6](#_Toc143719981)

[a. Description 6](#_Toc143719982)

[b. Results 7](#_Toc143719983)

[c. Key insights and Findings 8](#_Toc143719984)

[6. Business Question 3 9](#_Toc143719986)

[a. Description 9](#_Toc143719987)

[b. Results 9](#_Toc143719988)

[c. Key insights and Findings 10](#_Toc143719989)

[7. Business Question 4 11](#_Toc143719990)

[a. Description 11](#_Toc143719991)

[b. Results 11](#_Toc143719992)

[c. Key insights and Findings 12](#_Toc143719993)

[8. Business Question 5 14](#_Toc143719997)

[a. Description 14](#_Toc143719998)

[b. Results 14](#_Toc143719999)

[c. Key insights and Findings 15](#_Toc143720000)

[9. Business Question 6 16](#_Toc143720001)

[a. Description 16](#_Toc143720002)

[b. Results 16](#_Toc143720003)

[c. Key insights and Findings 17](#_Toc143720004)

[10. Business Question 7 20](#_Toc143720005)

[a. Description 20](#_Toc143720006)

[b. Results 20](#_Toc143720007)

[c. Key insights and Findings 22](#_Toc143720008)

[11. Business Question 8 24](#_Toc143720009)

[a. Description 24](#_Toc143720010)

[b. Results 25](#_Toc143720011)

[c. Key insights and Findings 28](#_Toc143720012)

[12. Business Question 9 31](#_Toc143720013)

[a. Description 31](#_Toc143720014)

[b. Results 32](#_Toc143720015)

[c. Key insights and Findings 33](#_Toc143720016)

[13. Business Question 10 37](#_Toc143720017)

[a. Description 37](#_Toc143720018)

[b. Results 37](#_Toc143720019)

[c. Key insights and Findings 39](#_Toc143720020)

[14. Conclusion 42](#_Toc143720021)

[15. References 43](#_Toc143720022)

# Executive Summary

In the report, we present the results of our project focused on analyzing various aspects of the company’s operation to gain valuable insights and inform decision-making. Our primary aims are to assess performance, identify trends and uncover opportunities for improvement across several dimensions of the business by answering a series of targeted questions.

The outcome of this project is a set of comprehensive analyses of various business aspects, such as sales trends, employee performance, product pricing and supplier stocks. Each section of the report is designed to answer specific questions and extract insights from the data. These insights will reveal the company’s status, thereby highlighting potential areas for improvement.

Overall, this project serves as a tool for the company’s leadership, assisting them to make informed decisions based on data-driven insights.

# Introduction

The key objectives of the project are to provide insights that empower decision making process of various facets of the company, identify opportunities for enhancing efficiency or allocating resources more proficiently, introduce a quantitative approach to evaluate employee performance, supplier relationships and product strategies, and uncover growth potential by several analysis on sale trends, customer preferences and market dynamics.

Many stakeholders play a vital role in shaping the objectives and the outcomes of this project. The executive team, who are seeking strategic insights to guide the organization’s direction, will be able to improve profitability by analyzing the project’s results on various criteria of company’s performance. The pricing team can gain more market insights for optimizing the pricing strategies across various products. The sales team will obtain a clearer understanding of employee performance, sales patterns, and customer behavior to optimize their sales strategies. The logistics team can utilize this report to improve inventory management solutions, thereby ensuring timely product availability while minimizing excessive stock.

To effectively address the above requirement, we utilize the power of SQL to extract valuable insights from the database. The table results will then be visualized to illustrate the key findings. SQL is a helpful tool, which allows us to perform complex data manipulations, aggregations, and comparisons, enabling us to provide meaningful insights from the database.

# Data Understanding

The data used for this project is the Northwind database, which contains sales data for a fictitious company named Northwind Traders, which operates in importing and exporting specialty goods around the world. The dataset offers valuable insights into the company’s operations, customer interactions and products. The data consists of 13 tables, which provide a comprehensive view of the company’s performance. It covers various facets of the company’s activities, from product categories and sales to customer interactions, employee performance or management, and more.

Here are the following steps to set up the database on PostgreSQL:

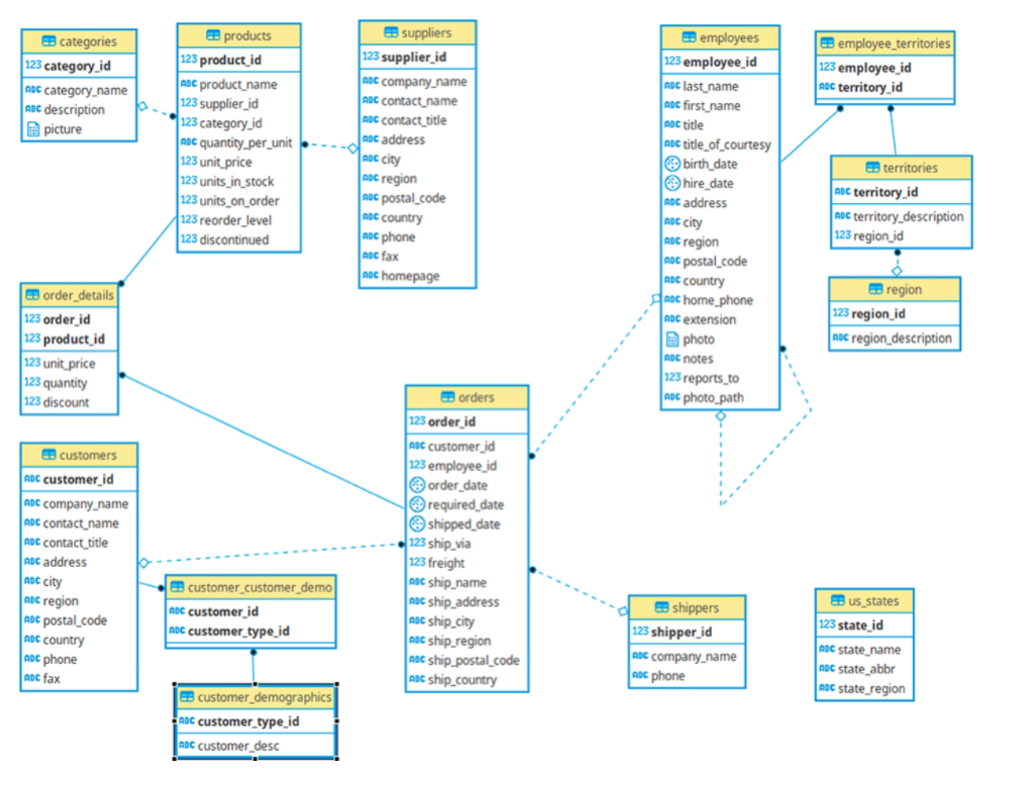
a. Download the PostgreSQL Database Management System from the official website: https://www.postgresql.org/download/

b. Follow the installation to set up password for superuser (usually ‘postgres’)

c. Create a database on postgresql named “northwind” and connect to ‘northtwind’ database

e. Download the Northwind Dataset and run the creation script to create table on ‘nothwind’ database.

The figure below is the entity relationship diagram (ERD) for this dataset after successfully created.



**Figure 1. ERD for the dataset**

# Business Question 1

## Description:

Filtered on the following conditions:

1. their unit price is between 20 and 50 (greater or equal to 20 but less or equal than 50)

2. they are not discontinued

### Results

To answer this question, we filtered the products based on the criteria mentioned above. Here are figures displaying the products that meet these conditions.

**Figure 2. Unit prices of various products**

**Figure 3. Count percentage of various products**

## Key insights and Findings:

The unit prices are distributed across the range of 20 to 49.3. The highest unit price is 49.3 for “Tarte au sucre” and lowest unit price is 20 for “Maxilaku”. The wide range of unit prices and the diversity of products likely reflect the company's pricing strategy aimed at attracting customers from different segments. Products like "Tarte au sucre" and "Ipoh Coffee" have higher unit prices compared to others. These high-value items could potentially be premium products or specialty items. Since the query filters out discontinued products, all the items in the result set are currently active and available for purchase. This shows the company’s commitment to providing those products to customers.

# Business Question 2

## Description

The Logistics Team wants to do a retrospection of their performances for the year 1998,

in order to identify for which countries, they didn’t perform well.

They asked you to provide them a list of countries with the following information:

1. Their average days between the order date and the shipping date (formatted to have only 2 decimals)

2. Their total number of unique orders (based on the order id)

Filtered on the following conditions:

1. the year of order date is 1998

2. their average days between the order date and the shipping date is greater or equal 5 days

3. their total number of orders is greater than 10 orders

Finally order the results by country name in an ascending order (following alphabetical order).

### Results

The analysis produced the following results, showcasing countries’ names, the average days between the order date and the shipping date, and the total number of unique orders for each country based on the above conditions:

**Figure 4. Average days between order and shipping for countries**

**Figure 5. Total unique orders for countries**

## Key insights and Findings

The average days between the order date and the shipping date indicates the average time it takes for orders to be shipped between countries. Among countries, Germany, Austria, and the UK have relatively lower average days, which indicate faster orders processing and shipping. By contrast, Sweden, France, Spain, and the USA have higher average days, suggesting potential insufficient in their shipping processes. Faster shipping plays an important role in increasing repeat business, which can lead to higher customer satisfaction. However, the company should balance it with logistics costs.

Germany with higher total unique orders tends to have relatively lower average days between order and shipping. Countries with lower total unique orders, like Sweden and Spain, exhibit longer average times, highlighting room for optimization in their logistics process.

# Business Question 3

## Description

The HR Team wants to know for each employee what was their age on the date they joined the company and who they currently report to. Provide them with a list of every employees with

+ their full name (first name and last name combined in a single field)

+ their job title

+ their age at the time they were hired

+ their manager full name (first name and last name combined in a single field)

+ their manager job title

Finally order the results by employee age and employee full name in an ascending order (lowest first).

### Results

The table provided below presents the detailed information about each employee based on above criteria:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| employee\_full\_name | employee\_job\_title | age\_at\_hire | manager\_full\_name | manager\_job\_title |
| Anne Dodsworth | Sales Representative | 28 | Steven Buchanan | Sales Manager |
| Janet Leverling | Sales Representative | 28 | Andrew Fuller | Vice President, Sales |
| Michael Suyama | Sales Representative | 30 | Steven Buchanan | Sales Manager |
| Robert King | Sales Representative | 33 | Steven Buchanan | Sales Manager |
| Laura Callahan | Inside Sales Coordinator | 36 | Andrew Fuller | Vice President, Sales |
| Steven Buchanan | Sales Manager | 38 | Andrew Fuller | Vice President, Sales |
| Andrew Fuller | Vice President, Sales | 40 |  |  |
| Nancy Davolio | Sales Representative | 43 | Andrew Fuller | Vice President, Sales |
| Margaret Peacock | Sales Representative | 55 | Andrew Fuller | Vice President, Sales |

**Table 1. Detailed information of employees**

## Key insights and Findings

**Figure 6. Average age of employees at hire**

**Figure 7. Average age of employees working under supervision of**

The graphs above illustrate the average age of employees by job title and the average age of employees working under their managers. The average age of Sales Managers (38) and the average age of Vice Presidents of Sales (40) are higher compared to the average age of Inside Sales Coordinators (36) and Sales Representatives (36.1). Higher-level roles like Sales Manager and Vice President require more experience and a deep understanding of the sale process and strategic decision making.

A noticeable connection emerges between an employee's age when they are hired and the managerial positions they fall under. Employees who work under the "Vice president, sales" have a higher average age (40) at the time employed than those under the sales manager. The positions reporting to the Vice president could demand more excellent professional experience and specialized skills due to higher level strategic responsibilities associated with a vice president role. By contrast, employees reporting to Sales Manager might be in more junior or entry levels (average age at 30.3 when arrived).

# Business Question 4

## Description

The Logistics Team wants to do a retrospection of their global performances over 1997-1998,

in order to identify for which month they perform well. They asked you to provide them a list with:

+ their year/month as single field in a date format (e.g. “1990-01-01” January 1990)

+ their total number of orders

+ their total freight (formatted to have no decimals)

Filtered on the following conditions:

+ the order date is between 1997 and 1998

+their total number of orders is greater than 35 orders

Finally order the results by total freight (descending order).

### Results

The table below presents the comprehensive overview, detailing the performance of the logistics team across various months during the year 1997-1998. It highlights the total number of orders and total freight where the total number of orders is greater than 35 orders. Finally, the results are ordered in a descending order based on the total freight value.

|  |  |  |
| --- | --- | --- |
| year\_month | total\_orders | total\_freight |
| 1998-04-01 | 74 | 6393 |
| 1998-01-01 | 55 | 5463 |
| 1998-03-01 | 73 | 5379 |
| 1998-02-01 | 54 | 4272 |
| 1997-10-01 | 38 | 3945 |
| 1997-12-01 | 48 | 3757 |
| 1997-09-01 | 37 | 3237 |

**Table 2. Detailed performances of logistics team**

## Key insights and Findings

## **Figure 8. Pie chart for total orders**

**Figure 9. Bar chart for total orders**

## **Figure 10. Pie chart for total freights by months**

**Figure 11. Bar chart for total freights by months**

The results illustrate the year and month, the total number of orders, and the total freight for each period. Generally, there is a positive correlation between total orders and freight. However, the fluctuations in total freight costs appear to be more pronounced than the fluctuations in total orders. This may suggest that other factors beyond volume, such as shipping method, fuel costs or other logistics matters, contribute significantly to the freight costs.

The months with the highest total freight costs are April 1998, January 1998, and March 1998. The increase in freight costs could suggest periods of increased business activity and potentially high demand for products. On the other hand, it could indicate seasonal trends or some industry events during this period. January marks the beginning of a new year and shows a surge in total orders and freight.

# Business Question 5

## Description

The Pricing Team wants to know which products had an unit price increase and the percentage increase was not between 20% and 30%. In order to help them they asked you to provide them a list of products with:

+ their product name

+ their current unit price (formatted to have only 2 decimals)

+ their initial unit price (formatted to have only 2 decimals)

+ their percentage increase with the result formatted to an integer (e.g 50 for 50%) using the following calculation:

(Current Unit Price - Initial Unit Price) ÷ Initial Unit Price \* 100

Filtered on the following conditions:

+ their percentage increase is not between 20% and 30% (lower than 20 or greater than 30)

+ Finally order the results by percentage increase (ascending order).

### Results

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| product\_id | product\_name | current\_price | base\_price | percentage\_increase |
| 72 | Mozzarella di Giovanni | 34.80 | 34.80 | 0.0000021923547903 |
| 42 | Singaporean Hokkien Fried Mee | 14.00 | 9.80 | 42.86 |
| 11 | Queso Cabrales | 21.00 | 14.00 | 50.00 |

**Table 3. Detailed information of product price changes**

The analysis filtered out products whose percentage increase fell outside the range of 20% to 30%. The results are then sorted in ascending order based on the percentage increase.

## Key insights and Findings

**Figure 12. Price changes**

Both "Singaporean Hokkien Fried Mee" and "Queso Cabrales" exhibited significant price volatility while "Mozzarella di Giovanni" saw an almost negligible change. This may suggest that "Singaporean Hokkien Fried Mee" and "Queso Cabrales" are likely to be sensitive to external factors that influence their prices.

Different customer segments have varying expectations and sensitivities when it comes to pricing. Each segment might be willing to pay within a specific price range for a particular product. Wild fluctuations outside these ranges could lead to dissatisfaction and potential loss of loyalty.

A stable price range contributes to a consistent perceived value of the product. If customers experience sudden price increases, they might question whether the product's value justifies the higher cost. This can lead to a decline in the product's perceived value. On the other hand, if there's little to no differentiation in quality between your product and competitors' products, price becomes a more critical factor in purchasing decisions.

Brands' reputation is closely tied to pricing strategies. A brand that demonstrates consistency in pricing strategy signals stability and reliability. As a result, in the long term, the sustainability of the price strategy should be the number one priority. Moreover, if the price changes are considerable, effective communication becomes vital. Explaining the reason behind the changes and the potential for more value-added can help mitigate the negative perceptions and maintain customer trust.

# Business Question 6

## Description

The Pricing Team wants to know how each category performs according to their price range.

In order to help them they asked you to provide them a list of categories with:

their category name

their price range as:

“1. Below $20”

“2. $20 - $50”

“3. Over $50”

their total amount (formatted to be integer) taking into account the offered discount

(i.e. subtracting the discounted amount)

their volume of orders (number of orders in which the category was present)

Finally order the results by category name then price range (both ascending order).

### Results

The table below presents the results of our analysis. It provides insights into how different categories of products perform within specific price ranges.

|  |  |  |  |
| --- | --- | --- | --- |
| category\_name | price\_range | total\_amount | order\_volume |
| Beverages | 1. BELOW $20 | 102945 | 317 |
| Beverages | 2. $20 - $50 | 23527 | 28 |
| Beverages | 3. OVER $50 | 141397 | 24 |
| Condiments | 1. BELOW $20 | 47559 | 120 |
| Condiments | 2. $20 - $50 | 58488 | 83 |
| Confections | 1. BELOW $20 | 59837 | 208 |
| Confections | 2. $20 - $50 | 85540 | 96 |
| Confections | 3. OVER $50 | 21980 | 15 |
| Dairy Products | 1. BELOW $20 | 35204 | 124 |
| Dairy Products | 2. $20 - $50 | 163402 | 195 |
| Dairy Products | 3. OVER $50 | 35901 | 28 |
| Grains/Cereals | 1. BELOW $20 | 24684 | 101 |
| Grains/Cereals | 2. $20 - $50 | 71061 | 88 |
| Meat/Poultry | 1. BELOW $20 | 14242 | 50 |
| Meat/Poultry | 2. $20 - $50 | 61185 | 82 |
| Meat/Poultry | 3. OVER $50 | 87595 | 36 |
| Produce | 1. BELOW $20 | 6944 | 28 |
| Produce | 2. $20 - $50 | 65917 | 82 |
| Produce | 3. OVER $50 | 27123 | 24 |
| Seafood | 1. BELOW $20 | 69977 | 228 |
| Seafood | 2. $20 - $50 | 45059 | 70 |
| Seafood | 3. OVER $50 | 16225 | 15 |

**Table 4. Categories’ performances by price ranges**

## Key insights and Findings

|  |
| --- |
| **Figure 13. Total amount by price ranges and categories** |
| **Figure 14. Total order volume by price ranges and categories** |
|  |

**Figure 15. Total amount and total order volume by price ranges**

Regarding category performance, beverages leading total amount and order volume, indicating a strong popularity and revenue generation. Regarding price range and demand, the goods priced below $20 dominate order volume while products priced range from $20 to $50 dominate total amount. The “Over $50” range represents a significant proportion of total amount but low proportion of order volume.

The” Beverage” category is a strong performer in both total amount and order volume. This indicates that beverages are a popular choice among customers, which demonstrates a potential area for further product expansion.

The above $50 price range contributes significantly to the total amount with a relatively low number of order volume. This suggests that customers are more willing to invest more in products in this range. It might be beneficial to focus on high quality or unique products within this price range.

“Below $20” items might have a lower total amount, but they contribute considerably to the total order volume. It indicates that the customer ae more willing to make frequent purchases when the product’s prices are reasonable. Therefore, concentrating on the customer retention and loyalty programs will build a long-term relationship with these customers, which can result in sustained revenue over time. Another strategy for this segment is cross-selling and upselling. Suggesting complementary products can increase the total volume and total revenue overall. Moreover, continue to offer promotions, discounts or other loyalty programs will help to build a long-term relationship with those customers.

On the other hand, “$20 to $50” items dominate in total amount. It means that they contribute significantly to the revenue. Compared to the “Below $20”, while these items hay have a lower order volume, they contribute more to the company in total amount. This means that the “$20 to $50” items have a higher **Average Transaction Value.** Customers who purchase products in this range often spend more per transaction on average. Therefore, due to a strong value proposition or unique features, value recognition is the key to effectively engaging with customers and maximizing sales. Customers need to understand and appreciate the value they receive when purchasing products in this range. Offering bundle deals or packages that combining the products within the “$20 to %50” range will showcase the value customers can get by purchasing multiple items, thereby promoting cross-selling.

Based on the above findings, it is suggested that the business should focus on optimizing its beverage category and maintain a diversified range of products that cover different price points. It will help to satisfy varying consumer preferences.

# Business Question 7

## Description

The Logistics Team wants to know what is the current state of our regional suppliers' stocks

for each category of product. In order to help them they asked you to provide them a list

of categories with:

their supplier region” as:

“America”

“Europe”

“Asia-Pacific”

their category name

their total units in stock

their total units on order

their total reorder level

Finally order the results by category name, then supplier region and reorder level (each in ascending order).

### Results

The graphs below show the current stock status of the company’s regional suppliers across different categories of products, which will help the team understand the availability of products in various regions.

**Figure 16. Total units in stock by categories**

**Figure 17. Total units on order by categories**

**Figure 18. Total reorder levels by categories**

## Key insights and Findings:

**Figure 19. Total units in stock, units on order and reorder level by categories**

Europe is the supplier region with the highest stock levels for total units in stock, total units on order and total reorder level, which indicates its significance as source of products. However, Europe also has units on orders, which suggests there is an ongoing demand for products in this region. The high total reorders level of Europe may suggest that there is a proactive approach to maintaining stocks levels.

Categories such as Produce, Beverages and Dairy products have low or zero total units on order, which may suggest a lack of demand or a specific ordering pattern. The sale team should investigate further to determine the reason behind the lack of demand. This could involve factors like seasonality, changing customer preferences, or even potential inaccuracies in forecasting.

The "Beverages" category exhibits an interesting trend with a high reorder level but relatively low total units on order. This discrepancy could indicate several possibilities. This could be due to the conservative approach to ordering, where products are only reordered when the stock levels approach to a certain point. On the other hand, the lag between the time an order is triggered and the time those arrive in stock might be the reason for this low level. Another possibility is that the reorder level is inaccurately forecast. The sales and logistics team should closely monitor to maintain the balance between reorder level and units on order, thereby enhancing the efficiency of the supply chain.

# Business Question 8

## Description

- The Pricing Team wants to know for each currently offered product how their unit price compares against their categories average and median unit price. In order to help them they asked you to provide them a list of products with:

their category name

their product name

their unit price

their category average unit price (formatted to have only 2 decimals)

their category median unit price (formatted to have only 2 decimals)

their position against the category average unit price as:

“Below Average”

“Equal Average”

“Over Average”

their position against the category median unit price as:

“Below Median”

“Equal Median”

“Over Median”

- Filtered on the following conditions:

+ They are not discontinued

Finally order the results by category name then product name (both ascending).

### Results

The table below displays the price of various products in comparison to the **mean and median** prices of the categories to which they belong, with boxes highlighted in pink color.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Row Labels** | **Below Average** | **Over Average** | **Below Median** | **Equal Median** | **Over Median** |
| Aniseed Syrup |  |  |  |  |  |
| Boston Crab Meat |  |  |  |  |  |
| CÃ´te de Blaye |  |  |  |  |  |
| Camembert Pierrot |  |  |  |  |  |
| Carnarvon Tigers |  |  |  |  |  |
| Chartreuse verte |  |  |  |  |  |
| Chef Anton's Cajun Seasoning |  |  |  |  |  |
| Chocolade |  |  |  |  |  |
| Escargots de Bourgogne |  |  |  |  |  |
| Filo Mix |  |  |  |  |  |
| Flotemysost |  |  |  |  |  |
| Geitost |  |  |  |  |  |
| Genen Shouyu |  |  |  |  |  |
| Gnocchi di nonna Alice |  |  |  |  |  |
| Gorgonzola Telino |  |  |  |  |  |
| Grandma's Boysenberry Spread |  |  |  |  |  |
| Gravad lax |  |  |  |  |  |
| Gudbrandsdalsost |  |  |  |  |  |
| Gula Malacca |  |  |  |  |  |
| GumbÃ¤r GummibÃ¤rchen |  |  |  |  |  |
| Gustaf's KnÃ¤ckebrÃ¶d |  |  |  |  |  |
| Ikura |  |  |  |  |  |
| Inlagd Sill |  |  |  |  |  |
| Ipoh Coffee |  |  |  |  |  |
| Jack's New England Clam Chowder |  |  |  |  |  |
| Konbu |  |  |  |  |  |
| LakkalikÃ¶Ã¶ri |  |  |  |  |  |
| Laughing Lumberjack Lager |  |  |  |  |  |
| Longlife Tofu |  |  |  |  |  |
| Louisiana Fiery Hot Pepper Sauce |  |  |  |  |  |
| Louisiana Hot Spiced Okra |  |  |  |  |  |
| Manjimup Dried Apples |  |  |  |  |  |
| Mascarpone Fabioli |  |  |  |  |  |
| Maxilaku |  |  |  |  |  |
| Mozzarella di Giovanni |  |  |  |  |  |
| Nord-Ost Matjeshering |  |  |  |  |  |
| Northwoods Cranberry Sauce |  |  |  |  |  |
| NuNuCa NuÃŸ-Nougat-Creme |  |  |  |  |  |
| Original Frankfurter grÃ¼ne SoÃŸe |  |  |  |  |  |
| Outback Lager |  |  |  |  |  |
| PÃ¢tÃ© chinois |  |  |  |  |  |
| Pavlova |  |  |  |  |  |
| Queso Cabrales |  |  |  |  |  |
| Queso Manchego La Pastora |  |  |  |  |  |
| RÃ¶d Kaviar |  |  |  |  |  |
| Raclette Courdavault |  |  |  |  |  |
| Ravioli Angelo |  |  |  |  |  |
| RhÃ¶nbrÃ¤u Klosterbier |  |  |  |  |  |
| Rogede sild |  |  |  |  |  |
| Sasquatch Ale |  |  |  |  |  |
| Schoggi Schokolade |  |  |  |  |  |
| Scottish Longbreads |  |  |  |  |  |
| Sir Rodney's Marmalade |  |  |  |  |  |
| Sir Rodney's Scones |  |  |  |  |  |
| Sirop d'Ã©rable |  |  |  |  |  |
| Spegesild |  |  |  |  |  |
| Steeleye Stout |  |  |  |  |  |
| Tarte au sucre |  |  |  |  |  |
| Teatime Chocolate Biscuits |  |  |  |  |  |
| Tofu |  |  |  |  |  |
| TourtiÃ¨re |  |  |  |  |  |
| TunnbrÃ¶d |  |  |  |  |  |
| Uncle Bob's Organic Dried Pears |  |  |  |  |  |
| Valkoinen suklaa |  |  |  |  |  |
| Vegie-spread |  |  |  |  |  |
| Wimmers gute SemmelknÃ¶del |  |  |  |  |  |
| Zaanse koeken |  |  |  |  |  |

**Table 5. Product price vs Mean and Median of Categories**

## Key insights and Findings

**Figure 20. Count of Products Compared to Their Respective Categories' Median Price**

**Figure 21. Count of Products Compared to Their Respective Categories' Average Price**

**Figure 22. Proportions of category**

**Figure 23. Proportions of average price comparison**

**Figure 24. Proportions of median price comparison**

Most products are priced “below average”, which indicates that the company positions itself as a provider of affordable options in the market. Moreover, the balance between “over median” and “below median” products proves that the company caters to different customer segments.

The “confections” category stands out in several comparisons, with higher count of product positioned as “Over median”, “Below median” and “Below average”. This could indicate that the company is targeting a broad range of market segments. This strategy allows them to attract customers from several segments with varying purchasing behavior.

Interestingly, the “Meat/Poultry” and “Grains/Cereals” have a lower count of products, which indicate potentially fewer products offerings or a specific market focus. However, this may suggest that there is limited supplier availability for the company. Therefore, there might be opportunities for the company to explore expanding their offerings within these categories.

# Business Question 9

## Description

The Sales Team wants to build a list of KPIs to measure employees' performances.

In order to help them they asked you to provide them a list of employees with:

+ their full name (first name and last name combined in a single field)

+ their job title

+ their total sales amount excluding discount (formatted to have only 2 decimals)

+ their total number of unique orders

+ their total number of orders

+ their average product amount excluding discount (formatted to have only 2 decimals).

This corresponds to the average amount of product sold (without taking into account any discount applied to it).

their average order amount excluding discount (formatted to have only 2 decimals). This corresponds to the ratio between the total amount of product sold (without taking into account any discount applied to it) against to the total number of unique orders.

their total discount amount (formatted to have only 2 decimals)

their total sales amount including discount (formatted to have only 2 decimals)

Their total discount percentage (formatted to have only 2 decimals)

Finally order the results by total sales amount including discount (descending).

### Results

The following table outline key performance indicators for employees, enabling a comprehensive understanding evaluation of their performance.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| full\_  name | Margaret Peacock | Janet Leverling | Nancy Davolio | Andrew Fuller | Robert King | Laura Callahan | Anne Dodsworth | | Michael Suyama | Steven Buchanan |
| title | Sales Representative | Sales Representative | Sales Representative | Vice President, Sales | Sales Representative | Inside Sales Coordinator | Sales Representative | Sales Representative | | Sales Manager |
| total\_sales\_amount\_excluding\_discount | 232,891 | 202,813 | 192,108 | 166,538 | 124,568 | 126,862 | 77,308 | 73,913 | | 68,792 |
| total\_unique\_orders | 156 | 127 | 123 | 96 | 72 | 104 | 43 | 67 | | 42 |
| total\_orders | 420 | 321 | 345 | 241 | 176 | 260 | 107 | 168 | | 117 |
| average\_product\_amount\_excluding\_discount | 1,759 | 1,772 | 1,750 | 1,926 | 2,100 | 1,381 | 2,008 | 1,245 | | 1,941 |
| average\_order\_amount\_excluding\_discount | 1,599 | 1,674 | 1,640 | 1,848 | 1,957 | 1,279 | 1,924 | 1,164 | | 1,794 |
| total\_discount\_amount | 26 | 16 | 17 | 11 | 13 | 15 | 7 | 9 | | 8 |
| total\_sales\_amount\_including\_discount | 250,187 | 213,051 | 202,144 | 177,749 | 141,296 | 133,301 | 82,964 | 78,198 | | 75,568 |
| total\_discount\_percentage | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | | 0.01 |

**Table 6. Key performances for Employees**

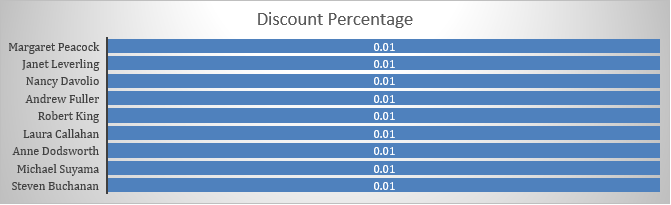
## Key insights and Findings

**Figure 25. Total sales amount (including and excluding discounts) for Employees**

**Figure 26. Total orders and total unique orders for Employees.**

**Figure 27. Average order and average product amount (excluding discount) for Employees.**

**Figure 28. Total discount amount for Employees**



**Figure 29. Discount Percentage for Employees**

Sales representatives, such as Margaret, Janet, Nancy and Robert who have a direct impact on sales volume, show strong performance in terms of total sales amount, total unique orders and total orders. Among them, Robert’s higher average order amount highlights his ability to secure more transactions. On the other hand, the consistent discount policy maintains a standardized approach to evaluate employees’ performance.

Despite handling fewer orders, Andrew Fuller (Vice President) has substantial total sales amount, which is influenced by larger average order and product amounts. This highlights his role in nurturing and closing high-value deals. Similarly, Steven Buchanan, the Sales Manager, may have lower total sales amounts and total orders compared to some other roles, but he stands out in terms of average order and average product amounts. This indicates his ability to generate higher value from fewer transactions. This may suggest that he focuses on cultivating the relationships with clients who make larger purchases, which then contributes to higher average order and product amounts.

By contrast, Laura Callahan (Inside Sales Coordinator) stands out with a high total number of orders, highlighting her effectiveness in managing a large volume of transactions. However, her total sales amount, average product amounts and average order amounts are relatively lower compared to other roles. This may mean that she involves different tasks beyond direct sales as a coordinator. Her role involves supporting the sale team’s activities, maintaining coordination to ensure that orders are processed smoothly. Other criteria including order processing efficiency, coordination feedback or customer satisfaction should be used for assessing her performance.

# Business Question 10

## Description

The Sales Team wants to build another list of KPIs to measure employees' performances across each category. In order to help them they asked you to provide them a list of categories and employees with:

their categories name

their full name (first name and last name combined in a single field)

their total sales amount including discount (formatted to have only 2 decimals)

their percentage of total sales amount including discount against his/her total sales amount across all categories (formatted to have only 5 decimals and maximum value up to 1)

their percentage of total sales amount including discount against the total sales amount across all employees (formatted to have only 5 decimals and maximum value up to 1)

Finally order the results by category name (ascending) then total sales amount (descending).

### Results

The following tables depict employees' contributions to each category and demonstrate how each category contributes to their overall performance.



**Table 7. Employees' performances across categories**

## Key insights and Findings:

**Figure 30. Percentage of total sales per employee**

**Figure 31. Percentage of total sales per category**

**Figure 32. Total sale amount including discount for employees across categories.**

**Figure 33. Percentage of total sales including discount across employees**

**Figure 34. Percentage of total sales including discount across categories**

Margaret Peacock has the largest proportions of total sales across employees (19%). Nancy Davolio and Janet Leverling also maintain a high percentage of total sales, illustrating their importance in generating revenue (totaling 30%).

Beverage and Dairy products contribute the largest proportions to total revenue (totalling36%). Nacy Davolio is the top performer in all these two categories. She also demonstrates her significant impact in many other categories (confections, seafoods, produce, etc.). Her ability and consistent high performance across multiple product categories highlight her versatility as a sales professional. The ability to make a substantial impact across various categories suggests that she possesses a broad skill set and deep understanding of customer needs. Therefore, the company should provide opportunities for her continued growth and development, which can further enhance her performance and overall impact on the company’s success.

On the other hand, Margaret Peacock, the top performer in overall total sales, performs consistently among all categories. She shows an ability to excel in multiple categories including beverages, condiments, confections, seafoods and so on. It highlights her versatility and deep understanding of customer preferences. Given her consistent top performance and adaptability, she may have leadership potential within the sales team. Therefore, considering her for leadership or mentoring position could be a strategic move for her knowledge and expertise.

# Conclusion

* 1. Key findings and insights:

The project unveils diverse insights from different aspects of the company, including sales trends, employee performance, supplier stocks, pricing strategy and so on. First, the company’s pricing strategy is designed to attract a wide range of customers by offering various products at different price points. Second, the age of hired employees is linked to their roles and managerial positions, highlighting the importance of experience for high-level positions. Third, several product categories exhibited strong performance, particularly beer, which generates both revenue and high volume. Fourth, supplier stocks and their manager varied across regions, with Europe emerging as a significant source of products and a potential area. Fifth, employees demonstrated varying levels of performance across categories, with some individuals excelling in specific categories.

* 1. Project success and stakeholder requirements**:**

The project achieved its main goal of providing data-driven insights to inform decision making for stakeholders by addressing specific questions posed by different teams. By utilizing SQL and data visualization, we extracted key information and meaningful insights from the Northwind database.

* 1. Future work and recommendations**:**

Although the project covers many questions on many aspects, it is highly recommended that future works involve more advanced analysis techniques, such as machine learning to forecast trends. Moreover, regularly monitoring KPIs will help the company keep abreast with the changing working environment. Establishing a closer relationship with suppliers is also essential for the company, which may result in more effective inventory management. For customer segmentation, developing tailored strategies for different customer segments will help boost brand recognition and revenue generation.

# References

* 1. *SQL tutorial*. (n.d.). <https://www.w3schools.com/sql/>
  2. *PostgreSQL tutorial – Comprehensive PostGresql tutorial.* (2022, August 28). PostgreSQL Tutorial. https://www.postgresqltutorial.com/