Assignment 1

Analysing Company Performance with SQL

Fatemeh Elyasifar

Student ID: 25589351

09/09/2024

94692 - Data Science Practise

Master of Data Science and Innovation

University of Technology of Sydney

Table of Contents

1. Executive Summary

Overview

This project is undertaken to assist various teams within the company, such as the Product, HR, Logistics, and Sales teams, by providing them with crucial data insights derived from the company's Northwind database. The main goal is to create SQL queries to answer specific business questions. The project is crucial because it can boost efficiency and improve pricing strategies by using the company's data.

Problem Statement and Context

Each department faced some challenges. For instance, building a list of KPIs to measure employees' performances, selecting products with offer for a specific price range, analysing global performance, calculating the percentage increase, showing the performance of each category according to their price, and comparing the price of offered products against their categories' average and median price, are required. The project aims to provide accurate insights to assist decision-making in various areas.

Achieved Outcomes and Results

Throughout the project, the requirements of each team were successfully achieved. Products within the $20 to $50 range were identified, with significant price increases highlighted. A list of key performance indicators (KPIs) was developed to measure employee performance. Additionally, total sales, average order values, and discount percentages were identified to enhance pricing strategies. These outcomes enabled the company to gain a better understanding of its performance and make beneficial decisions for future development.

In summary, the project successfully provided valuable data insights across different parts of the company and helped to enhance strategic planning and overall efficiency.

2. Introduction

Objectives

The key objective of this project is to generate useful insights for various departments in the company by analysing data from the Northwind dataset. The goal is to answer business questions that would help each team to improve their strategies. The project aims to provide SQL queries that could assist in optimising pricing strategies and evaluating sales metrics.

Stakeholders and Requirements

Each department had its own challenges. For instance, the Product Team needed to identify products with specific price ranges to review pricing strategies. The HR Team wanted to look at their employees' performance. Moreover, the Sales Team wanted to evaluate their employee performance across each category. The Logistic Team needed to know the current state of their regional suppliers' stocks for each category and see how orders were handled across different regions and time periods.

Project Goal for Requirements

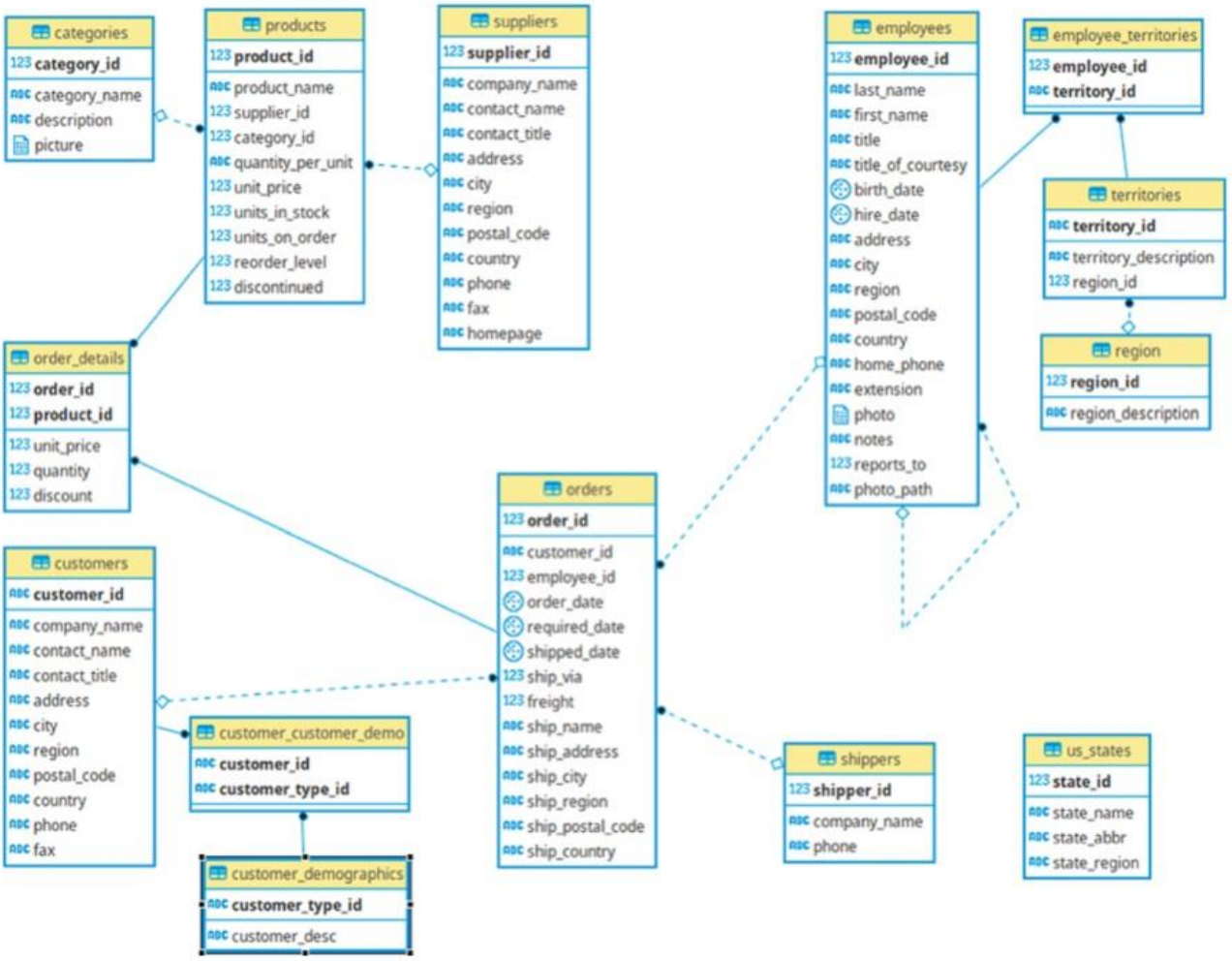
The project aims to meet the mentioned requirements by developing SQL queries related to each department's needs. These queries are designed to extract valuable data to provide useful insights. For this project, the data we gathered and analysed could optimise operations across the company.

In summary, the project is all about appropriate responses to business questions to analyse and make better decisions, work more efficiently, and drive the business forward.

3. Data Understanding

Dataset

For this project, the Northwind database was used that has 14 tables, including sales, products, employees, region, and orders, that simulates a company's operations. The Order Details' table with 2155 rows, is the biggest table of the schema. Tables are connected to each other through foreign keys. The dataset provides a source of information for generating insights across different departments of the company.



Data Sources, Collection methods and Limitations

The Northwind dataset is designed to reflect the real-world operations of a company. The data is collected from different parts of the company during 1996 to 1998. However, based on the collection time, it might not be an appropriate source to make real decisions. Also, the way the data is structured might limit flexibility to dig deeper into certain relationships between variables.

Variables and Significance

The dataset contains various tables, each with several important features. These includes:

- Products Table: Includes variables such as ‘product_id’, ‘product_name’, ‘category_id’, ‘unit_price’, and ‘units_in_stock’. These variables are crucial for understanding product pricing, stock levels, and product availability.

- Orders Table: Contains ‘order_id’, ‘customer_id’, ‘order_date’, and ‘shipped_date’, which are necessary to analysing shipping delays and regional differences.

- Order Details Table: This connects orders and products, including ‘order_id’, ‘product_id’, ‘unit_price’, ‘quantity’, and ‘discount’. It’s essential for analysing sales performance.
-
- Categories Table: Contains ‘category_id’ and ‘category_name’, which is crucial for analysing how different categories perform.
-
- Employees Table: Includes ‘employee_id’, ‘first_name’, ‘last_name’, ‘region’, ‘hire_date’, and ‘reports_to’, that provide insights into employee information.

Introductions to create and connect to database

- Download DBeaver from the official website.
- Connect to PostgreSQL:
    - Click on the New Database Connection button on DBeaver
    - Select PostgreSQL from the list
    - Enter the PostgreSQL server details
- Create a new database
- Write SQL queries

4. Business Question 1

a. Description

The Product Team wants to know the following information about each product:

- Product name
- Product unit price

This information should be provided by products that their unit price is between $20 and $50 and not discontinued. The result should be ordered by unit price in descending order, so they can focus on their annual review of the company pricing strategy and the higher-priced items within this range.

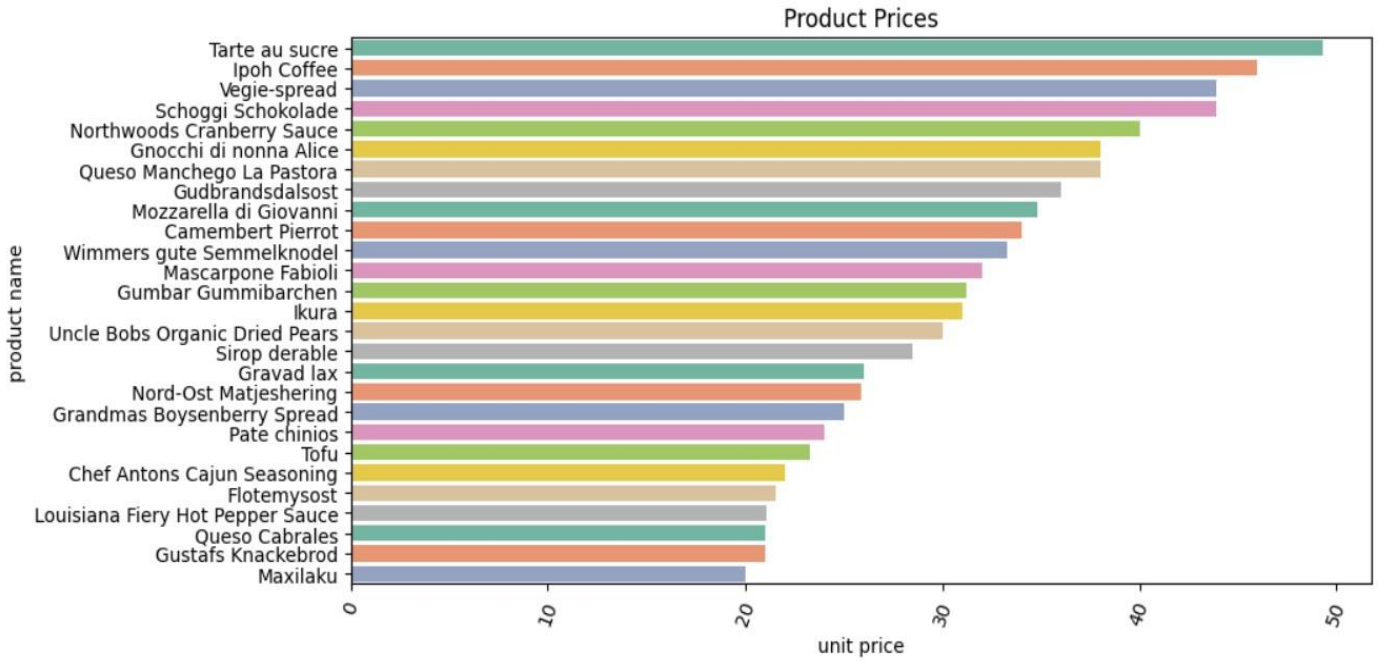
b. Results

	product_name	unit_price
1	Tarte au sucre	49.3
2	Ipoh Coffee	46
3	Vegie-spread	43.9
4	Schoggi Schokolade	43.9
5	Northwoods Cranberry Sauce	40
6	Gnocchi di nonna Alice	38
7	Queso Manchego La Pastora	38
8	Gudbrandsdalsost	36
9	Mozzarella di Giovanni	34.8
10	Camembert Pierrot	34
11	Wimmers gute Semmelknödel	33.25
12	Mascarpone Fabioli	32
13	Gumbär Gummibärchen	31.23
14	Ikura	31
15	Uncle Bob's Organic Dried Pear	30
16	Sirop d'érable	28.5
17	Gravad lax	26
18	Nord-Ost Matjeshering	25.89
19	Grandma's Boysenberry Spread	25
20	Pâté chinois	24
21	Tofu	23.25
22	Chef Anton's Cajun Seasoning	22
23	Flotemysost	21.5
24	Louisiana Fiery Hot Pepper Sauc	21.05
25	Queso Cabrales	21
26	Gustaf's Knäckebröd	21
27	Maxilaku	20

The table shows a list of products that are currently being offered at prices between $20 and $50. Tarte au sucre was the most expensive product while the Maxilaku was the cheapest.

These products highlight the more expensive items first sorted by their unit price in descending order.

c. Key insights and Findings



- There is a varied distribution from the $20 to $50 price range.
- There are some products with the same price such as Gnocchi and Queso.
- Tarte au sucre priced 49.3, offering an expensive product.

This table helps the Product Team assess the current pricing strategy to enhance market performance. The team can use this data to make decisions about pricing adjustments to boost profitability.

5. Business Question 2

a. Description

The Logistics Team wants to know the following information about different countries to assess their performance:

- average days between the order date and the shipping date
- total number of unique orders

This information should be ordered by country name in alphabetic order, so they can focus on countries that may have faced challenges and require improvement.

b. Results

	ship_country	total_orders	avg_days
1	Austria	11	5.89
2	Brazil	28	8.12
3	France	23	9.43
4	Germany	34	5.38
5	Spain	12	7.83
6	Sweden	14	13.29
7	UK	16	6.25
8	USA	39	7.89
9	Venezuela	18	8.73

Based on result, USA had the highest number of orders while Austria had the lowest number of orders. Sweden had the longest shipping times in 1998 while Germany had the shortest shipping time. Generally, all countries' shipping time was more than 5 days which is considerable.

c. Key insights and Findings

- This analysis shows that 9 countries did not perform well in 1998. This helps the Logistics Team to improve delivery times and overall customer satisfaction in those regions.

6. Business Question 3

a. Description

The HR team wants to know the following information about each employee:

- Full name and job title
- Age at hiring date
- Reporting manager and job title

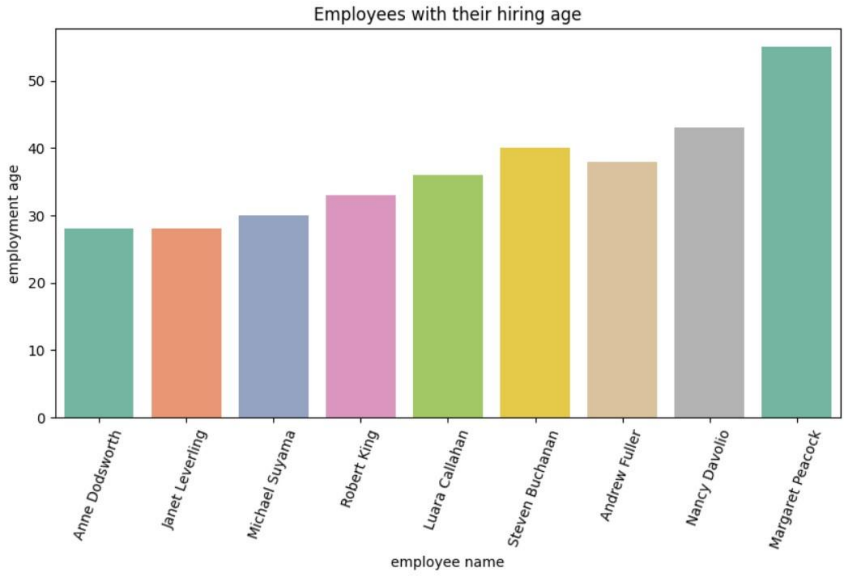
This information should be provided by ascending age and surname in alphabetic order. So, this can provide a better understanding of employees.

b. Results

	employee_full_name	employee_job_title	employee_age	manager_full_name	manager_job_title
1	Anne Dodsworth	Sales Representative	28	Steven Buchanan	Sales Manager
2	Janet Leverling	Sales Representative	28	Andrew Fuller	Vice President, Sales
3	Michael Suyama	Sales Representative	30	Steven Buchanan	Sales Manager
4	Robert King	Sales Representative	33	Steven Buchanan	Sales Manager
5	Laura Callahan	Inside Sales Coordinator	36	Andrew Fuller	Vice President, Sales
6	Steven Buchanan	Sales Manager	38	Andrew Fuller	Vice President, Sales
7	Andrew Fuller	Vice President, Sales	40	[NULL]	[NULL]
8	Nancy Davolio	Sales Representative	43	Andrew Fuller	Vice President, Sales
9	Margaret Peacock	Sales Representative	55	Andrew Fuller	Vice President, Sales

The analysis provides a detailed list of 8 employees with their full name, job title, their age at the time of hiring, their manager's full name, and manager's job title. There were 6 people as sales representatives, including Anne Dodsworth at the top of the list, as the youngest at the time of hiring. Andrew as a manager does not have any manager.

c. Key insights and Findings



- Employees were older than 27 at the time of hiring.
- Margarent was the oldest candidate for sales representative.
- Only 2 managers were at the company.

The table shows how responsibilities are distributed across different levels of the organization.
So, the HR Team can use this analysis to gain insights into hiring practices, career development, and the overall organizational structure.

7. Business Question 4

a. Description

The Logistics Team wants to know the following information about their global performance over 1997-1998:

- year/month as single field in a date format
- total number of orders which is greater than 35 orders
- total freight

This information should be ordered by total freight in descending order. The goal is to identify patterns that could inform future logistics strategies.

b. Results

	date	total_orders	total_freight
1	1998-04-01	74	6,394
2	1998-01-01	55	5,463
3	1998-03-01	73	5,379
4	1998-02-01	54	4,273
5	1997-10-01	38	3,946
6	1997-12-01	48	3,758
7	1997-09-01	37	3,237

The analysis reveals the performance for each month for the years 1997 and 1998 based on the total number of orders and the sum of freight.

This result shows that all dates with total orders greater than 35 belong to the last four months of 1997 and the first four months of 1998. During these 8 months their total freight was more than 3,000 and less than 7,000.

c. Key insights and Findings

- The team had a strong performance in April and March in terms of total number of orders.
- The team also experienced peak performance in April and January 1998 in terms of total freight, which means through these months they shipped a higher volume of goods.
- Based on the plot, there is a weak positive correlation between total orders and freight. It means the higher the total number of orders the higher total freight. But this is not always the case.

So, these insights can help the logistics Team in optimising shipping strategies to balance costs with efficiency.

8. Business Question 5

a. Description

The Pricing Team wants to know the following information about each product:

- product name
- current unit price
- initial unit price
- percentage increase which is not between 20% and 30%

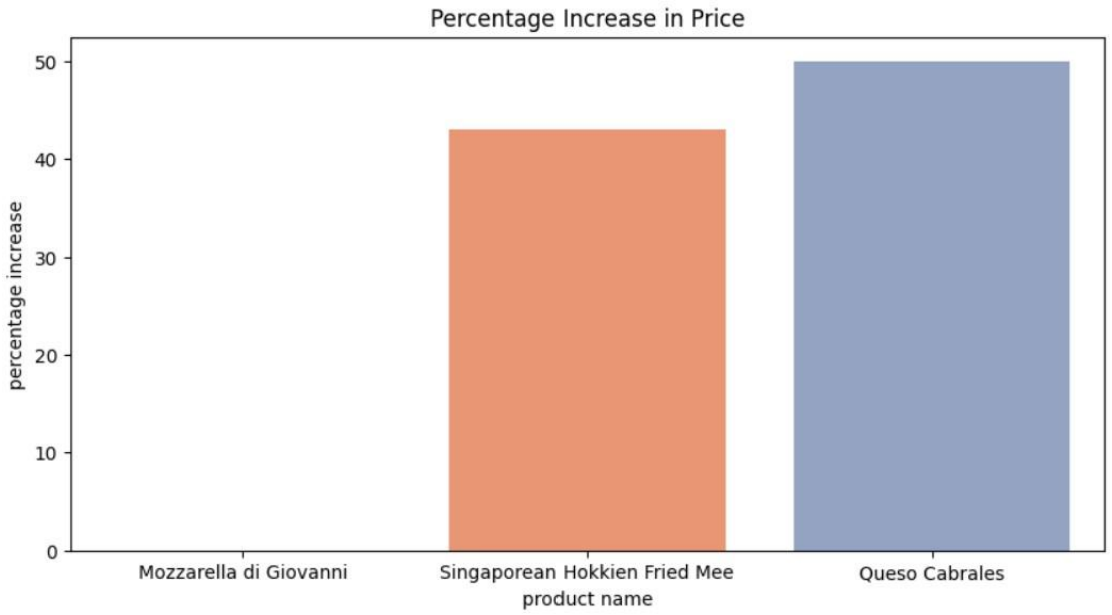
This information should be ordered by percentage increase in ascending order. This analysis will help the team understand pricing trends and identify any products that may need further adjustment.

b. Results

	product_name	current_price	initial_price	percentage_increase
1	Mozzarella di Giovanni	34.8	34.8	0
2	Singaporean Hokkien Fried Mee	14	9.8	43
3	Queso Cabrales	21	14	50

These results indicate that certain products have experienced significant price increases above 30%. For example, Queso Cabrales saw a 50% increase, which is well above the 30% threshold.

c. Key insights and Findings



- As there are just 2 products with more than a 30% price increase, the company can decrease these prices to make them affordable for customers.
- No price change occurred for Mozzarella di Giovanni.

9. Business Question 6

a. Description

The Pricing Team wants to know the following information about each category:

- category name
- price range as:
    - “1. Below $20”
    - “2. $20 - $50”
    - “3. Over $50”
- total amount considering the offered discount
- volume of orders

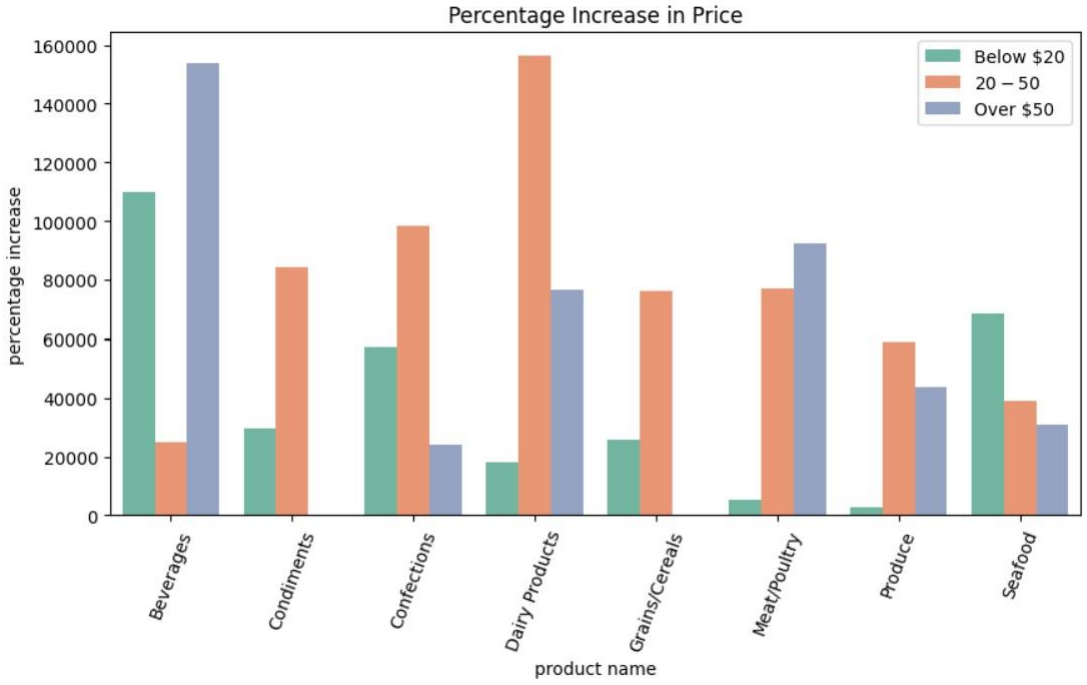
This information should be ordered by category name and then price range in ascending order. This analysis will help the Pricing Team understand which price ranges contribute most to the overall sales within each category and identify any product strategy that may need further adjustment.

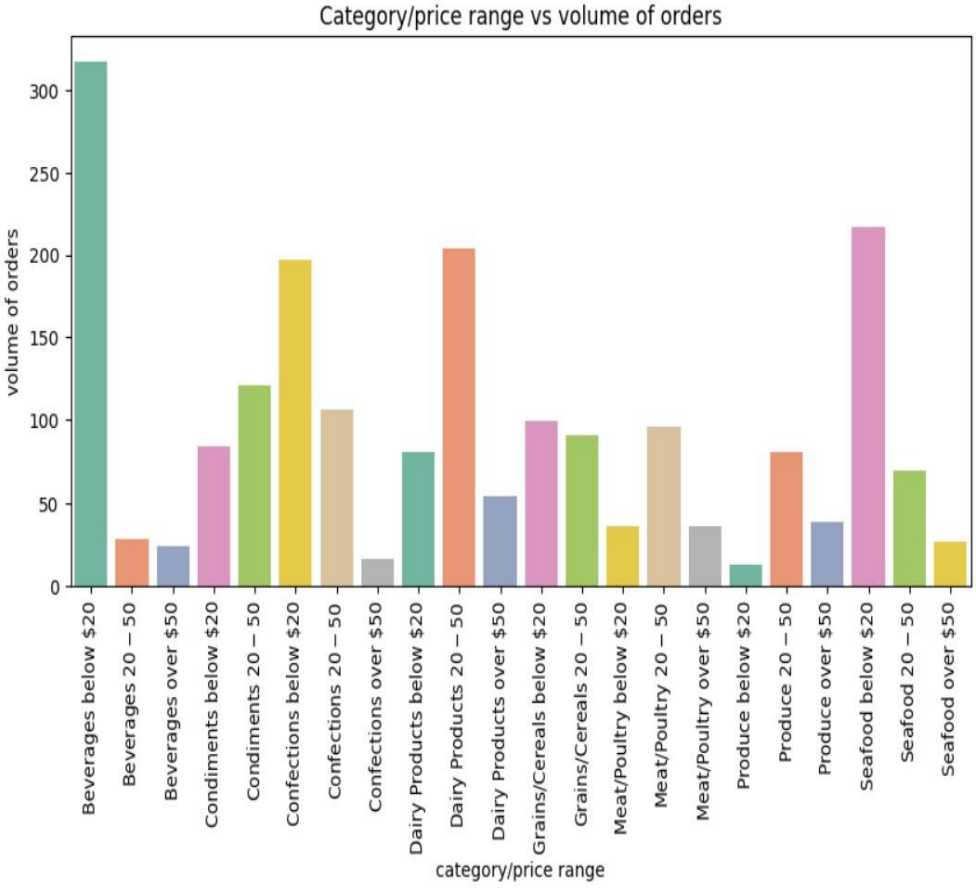
b. Results

	category_name	price_range	total_amount	volume_of_orders
1	Beverages	1. Below $20	102,945	317
2	Beverages	2. $20 - $50	23,527	28
3	Beverages	3. Over $50	141,397	24
4	Condiments	1. Below $20	27,299	85
5	Condiments	2. $20 - $50	78,748	121
6	Confections	1. Below $20	53,365	197
7	Confections	2. $20 - $50	91,429	106
8	Confections	3. Over $50	22,563	16
9	Dairy Products	1. Below $20	16,569	81
10	Dairy Products	2. $20 - $50	146,783	204
11	Dairy Products	3. Over $50	71,156	54
12	Grains/Cereals	1. Below $20	24,071	99
13	Grains/Cereals	2. $20 - $50	71,673	91
14	Meat/Poultry	1. Below $20	4,728	36
15	Meat/Poultry	2. $20 - $50	70,699	96
16	Meat/Poultry	3. Over $50	87,595	36
17	Produce	1. Below $20	2,432	13
18	Produce	2. $20 - $50	55,732	81
19	Produce	3. Over $50	41,820	39
20	Seafood	1. Below $20	65,110	217
21	Seafood	2. $20 - $50	36,980	70
22	Seafood	3. Over $50	29,172	27

There are 8 distinct categories in the table. This result indicates that for categories like Diary products and Beverages, the "$20 - $50" and the "Over $50" price range respectively generates the highest total revenue while the other "Over $50" range generates less revenue and fewer orders. Also, the "Below $20" range for Beverages and Seafood suggests strong demand for lower-priced products.

c. Key insights and Findings





- The analysis shows that products less than $20 are the most significant contributor to revenue. This means a strong market preference for low-priced items.
- Beverages has the highest total amount for the price over $50.
- Costumers prefer to buy decent dairy products, meat or poultry in range of $20 to $50 rather than the cheap ones. This suggests that they like to eat high-quality dairy products, meat or poultry.
- The cheapest Beverages, confections, and seafood are preferred among customers.

So, the result highlights the importance of maintaining various products offering across different price ranges. This helps the Pricing Team to focus on promotional strategies to maximize their sales.

10. Business Question 7

a. Description

The Logistics Team wants to know the following information about regional suppliers' stock:

- supplier region" as:
    - "America"
    - "Europe"
    - "Asia-Pacific"
- category name
- total units in stock
- total units on order
- total reorder level

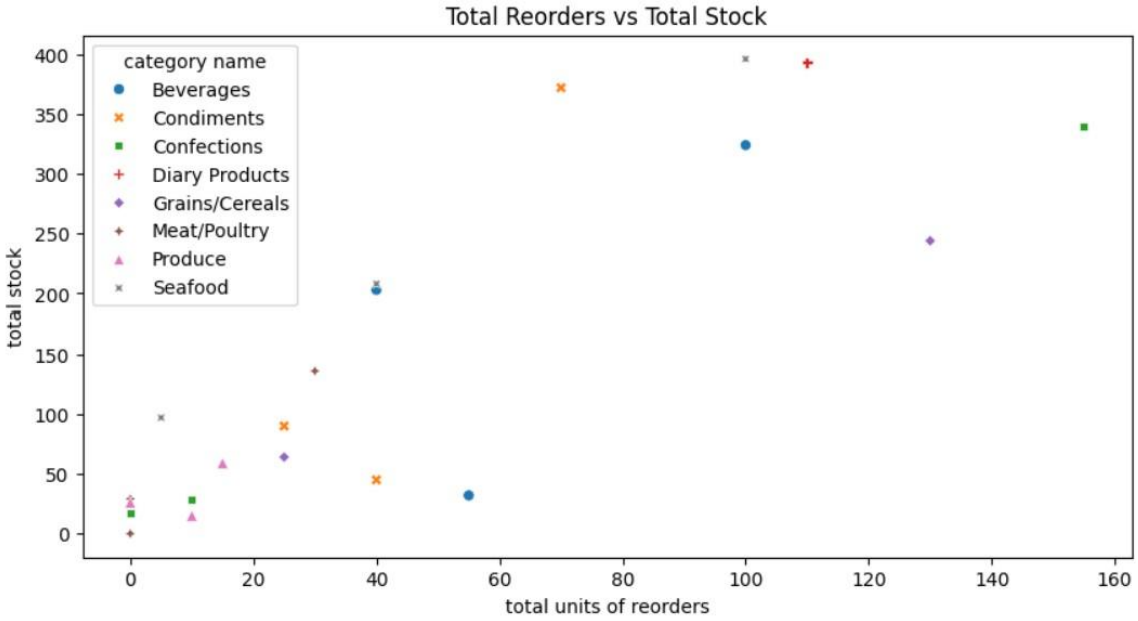
This information should be ordered by category name and then supplier region and reorder level in ascending order. This will help them monitor inventory levels and make informed decisions about restocking and distribution to avoid shortages situations.

b. Results

	category_name	supplier_region	total_stock	total_orders	total_reorders
1	Beverages	America	203	0	40
2	Beverages	Asia-Pacific	32	20	55
3	Beverages	Europe	324	40	100
4	Condiments	America	372	100	70
5	Condiments	Asia-Pacific	90	0	25
6	Condiments	Europe	45	70	40
7	Confections	America	17	0	0
8	Confections	Asia-Pacific	29	0	10
9	Confections	Europe	340	180	155
10	Dairy Products	Europe	393	140	110
11	Grains/Cereals	Asia-Pacific	64	0	25
12	Grains/Cereals	Europe	244	90	130
13	Meat/Poultry	America	136	0	30
14	Meat/Poultry	Asia-Pacific	29	0	0
15	Meat/Poultry	Europe	0	0	0
16	Produce	America	15	0	10
17	Produce	Asia-Pacific	59	20	15
18	Produce	Europe	26	0	0
19	Seafood	America	208	0	40
20	Seafood	Asia-Pacific	97	0	5
21	Seafood	Europe	396	120	100

The result shows the significant distribution of stock across different regions. America has the highest stock levels for Condiments, and Meat/Poultry. Also, the highest level of reorder belongs to Europe when there is no meat or poultry available in this region.

c. Key insights and Findings



- Europe has higher stock levels for Beverages and Seafood, which may indicate higher demand in that region.
- More beverages are needed in Asia-Pacific.
- More condiments are needed in Europe.
- In other places, there is enough stock considering the total orders and reorders across different regions.

So, this result helps the Logistics Team to consider adjusting reorder levels based on regional demand patterns to better align inventory with needs.

11. Business Question 8

a. Description

The Pricing Team wants to know the following information about products:

- their category name
- product name
- unit price
- category average unit price
- category median unit price
- position against the category average unit price as:
    - “Below Average”
    - “Equal Average”
    - “Over Average”
- position against the category median unit price as:
    - “Below Median”
    - “Equal Median”
    - “Over Median”

This information should not be discontinued and be ordered by category name in alphabetic order and then product name in ascending order.

b. Results

	category_name	product_name	unit_price	avg_unit_price	median_unit_price	price_vs_avg	price_vs_median
1	Beverages	Chai	18	46.03	18	Below Average	Equal median
2	Beverages	Chang	19	46.03	18	Below Average	Over median
3	Beverages	Chartreuse verte	18	46.03	18	Below Average	Equal median
4	Beverages	Côte de Blaye	263.5	46.03	18	Over Average	Over median
5	Beverages	Guaraná Fantástica	4.5	46.03	18	Below Average	Below median
6	Beverages	Ipoh Coffee	46	46.03	18	Below Average	Over median
7	Beverages	Lakkalikööri	18	46.03	18	Below Average	Equal median
8	Beverages	Laughing Lumberjack Lager	14	46.03	18	Below Average	Below median
9	Beverages	Outback Lager	15	46.03	18	Below Average	Below median
10	Beverages	Rhönbräu Klosterbier	7.75	46.03	18	Below Average	Below median
11	Beverages	Sasquatch Ale	14	46.03	18	Below Average	Below median
12	Beverages	Steeleye Stout	18	46.03	18	Below Average	Equal median
13	Condiments	Aniseed Syrup	10	22.99	21.05	Below Average	Below median
14	Condiments	Chef Anton's Cajun Seasonin	22	22.99	21.05	Below Average	Over median
15	Condiments	Chef Anton's Gumbo Mix	21.35	22.99	21.05	Below Average	Over median
16	Condiments	Genen Shouyu	13	22.99	21.05	Below Average	Below median
17	Condiments	Grandma's Boysenberry Spre	25	22.99	21.05	Over Average	Over median
18	Condiments	Gula Malacca	19.45	22.99	21.05	Below Average	Below median
19	Condiments	Louisiana Fiery Hot Pepper S	21.05	22.99	21.05	Below Average	Below median
20	Condiments	Louisiana Hot Spiced Okra	17	22.99	21.05	Below Average	Below median
21	Condiments	Northwoods Cranberry Sauce	40	22.99	21.05	Over Average	Over median
22	Condiments	Original Frankfurter grüne Sc	13	22.99	21.05	Below Average	Below median
23	Condiments	Sirop d'érable	28.5	22.99	21.05	Over Average	Over median
24	Condiments	Vegie-spread	43.9	22.99	21.05	Over Average	Over median
25	Confections	Chocolade	12.75	25.16	16.25	Below Average	Below median
26	Confections	Gumbär Gummibärchen	31.23	25.16	16.25	Over Average	Over median
27	Confections	Maxilaku	20	25.16	16.25	Below Average	Over median
28	Confections	NuNuCa Nuß-Nougat-Creme	14	25.16	16.25	Below Average	Below median
29	Confections	Pavlova	17.45	25.16	16.25	Below Average	Over median
30	Confections	Schoggi Schokolade	43.9	25.16	16.25	Over Average	Over median
31	Confections	Scottish Longbreads	12.5	25.16	16.25	Below Average	Below median
32	Confections	Sir Rodney's Marmalade	81	25.16	16.25	Over Average	Over median
33	Confections	Sir Rodney's Scones	10	25.16	16.25	Below Average	Below median
34	Confections	Tarte au sucre	49.3	25.16	16.25	Over Average	Over median
35	Confections	Teatime Chocolate Biscuits	9.2	25.16	16.25	Below Average	Below median
36	Confections	Valkoinen suklaa	16.25	25.16	16.25	Below Average	Equal median
37	Confections	Zaanse koeken	9.5	25.16	16.25	Below Average	Below median
38	Dairy Products	Camembert Pierrot	34	28.73	33	Over Average	Over median
39	Dairy Products	Flotemysost	21.5	28.73	33	Below Average	Below median
40	Dairy Products	Geitost	2.5	28.73	33	Below Average	Below median
41	Dairy Products	Gorgonzola Telino	12.5	28.73	33	Below Average	Below median
42	Dairy Products	Gudbrandsdalsost	36	28.73	33	Over Average	Over median
43	Dairy Products	Mascarpone Fabioli	32	28.73	33	Over Average	Below median
44	Dairy Products	Mozzarella di Giovanni	34.8	28.73	33	Over Average	Over median
45	Dairy Products	Queso Cabrales	21	28.73	33	Below Average	Below median
46	Dairy Products	Queso Manchego La Pastora	38	28.73	33	Over Average	Over median
47	Dairy Products	Raclette Courdavault	55	28.73	33	Over Average	Over median
48	Grains/Cereals	Filo Mix	7	21.29	20.25	Below Average	Below median
49	Grains/Cereals	Gnocchi di nonna Alice	38	21.29	20.25	Over Average	Over median
50	Grains/Cereals	Gustaf's Knäckebröd	21	21.29	20.25	Below Average	Over median
51	Grains/Cereals	Ravioli Angelo	19.5	21.29	20.25	Below Average	Below median
52	Grains/Cereals	Singaporean Hokkien Fried M	14	21.29	20.25	Below Average	Below median
53	Grains/Cereals	Tunnbröd	9	21.29	20.25	Below Average	Below median
54	Grains/Cereals	Wimmers gute Semmelknöd	33.25	21.29	20.25	Over Average	Over median
55	Meat/Poultry	Alice Mutton	39	15.72	15.72	Over Average	Over median
56	Meat/Poultry	Mishi Kobe Niku	97	15.72	15.72	Over Average	Over median
57	Meat/Poultry	Pâté chinois	24	15.72	15.72	Over Average	Over median
58	Meat/Poultry	Perth Pasties	32.8	15.72	15.72	Over Average	Over median
59	Meat/Poultry	Thüringer Rostbratwurst	123.79	15.72	15.72	Over Average	Over median
60	Meat/Poultry	Tourtière	7.45	15.72	15.72	Below Average	Below median
61	Produce	Longlife Tofu	10	29.06	26.63	Below Average	Below median
62	Produce	Manjimup Dried Apples	53	29.06	26.63	Over Average	Over median
63	Produce	Rössle Sauerkraut	45.6	29.06	26.63	Over Average	Over median
64	Produce	Tofu	23.25	29.06	26.63	Below Average	Below median
65	Produce	Uncle Bob's Organic Dried Pe	30	29.06	26.63	Over Average	Over median
66	Seafood	Boston Crab Meat	18.4	20.68	16.7	Below Average	Over median
67	Seafood	Carnarvon Tigers	62.5	20.68	16.7	Over Average	Over median
68	Seafood	Escargots de Bourgogne	13.25	20.68	16.7	Below Average	Below median
69	Seafood	Gravad lax	26	20.68	16.7	Over Average	Over median
70	Seafood	Ikura	31	20.68	16.7	Over Average	Over median
71	Seafood	Inlagd Sill	19	20.68	16.7	Below Average	Over median
72	Seafood	Jack's New England Clam Ch	9.65	20.68	16.7	Below Average	Below median
73	Seafood	Konbu	6	20.68	16.7	Below Average	Below median
74	Seafood	Nord-Ost Matjeshering	25.89	20.68	16.7	Over Average	Over median
75	Seafood	Röd Kaviar	15	20.68	16.7	Below Average	Below median
76	Seafood	Rogede sild	9.5	20.68	16.7	Below Average	Below median
77	Seafood	Spegesild	12	20.68	16.7	Below Average	Below median

Most of the products in beverages are below the average and median price, indicating affordable products. This applies to other categories such as condiments, confections, grains, and seafood as well.

In the meat and poultry category, most of the products are over average or median, offering expensive products. This suggests changing the price strategies to make it cheaper for customers.

c. Key insights and Findings

- The result shows a diverse range of products within each category.
- Most prices are below their categories' average and median which means these products are more affordable
- The Product Team can decrease the price of meat to make it more affordable.
- The Product Team can increase the price in those positioned below average to increase the profit.
- The Pricing Team may consider some new strategics based on price positions.

12. Business Question 9

a. Description

The Sales Team wants to know the following information about each employee:

- full name
- job title
- total sales amount excluding discount
- total number of unique orders
- total number of orders
- average product amount excluding discount
- average order amount excluding discount
- total discount amount
- total sales amount including discount
- total discount percentage

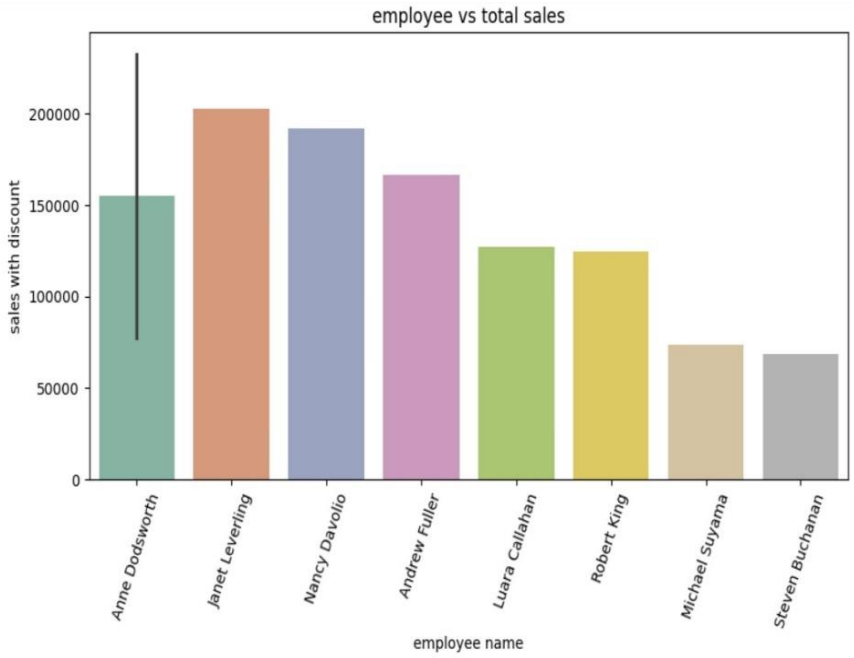
The result should be ordered by total sales amount including discount in descending order. As a result, the team can make decisions regarding performance evaluations, bonuses, and potential areas for improvement.

b. Results

	employee_full_name	job_title	total_sales_without_discount	total_unique_orders	total_orders	avg_product_amount	avg_order_amount	total_discount	total_sales_with_discount	total_discount_percentage
1	Margaret Peacock	Sales Representative	250,187.45	156	420	595.68	1,603.77	17,296.6	232,890.85	6.91
2	Janet Leverling	Sales Representative	213,051.3	127	321	663.71	1,677.57	10,238.46	202,812.84	4.81
3	Nancy Davolio	Sales Representative	202,143.71	123	345	585.92	1,643.44	10,036.11	192,107.6	4.96
4	Andrew Fuller	Vice President, Sales	177,749.26	96	241	737.55	1,851.55	11,211.51	166,537.76	6.31
5	Laura Callahan	Inside Sales Coordinator	133,301.03	104	260	512.7	1,281.74	6,438.75	126,862.28	4.83
6	Robert King	Sales Representative	141,295.99	72	176	802.82	1,962.44	16,727.76	124,568.23	11.84
7	Anne Dodsworth	Sales Representative	82,964	43	107	775.36	1,929.4	5,655.93	77,308.07	6.82
8	Michael Suyama	Sales Representative	78,198.1	67	168	465.46	1,167.14	4,284.97	73,913.13	5.48
9	Steven Buchanan	Sales Manager	75,567.75	42	117	645.88	1,799.23	6,775.47	68,792.28	8.97

The result highlights the sales performance of nine employees. The highest total sales including discounts belongs to Margaret while the lowest to Steven. The total number of orders varies from 168 to 420. Also, the average order amount varies from 1,167 to 1,962.

c. Key insights and Findings



- Margaret, as a sales representative, indicated a strong performance in driving sales volume, having the highest total sales, both excluding and including discounts.
- Luara and Robert have almost the same amount of sales including discount.
- Steven had the lowest amount of sales.

So, the analysis helps the Sales Team with a comprehensive view of each employee's performance.

13. Business Question 10

a. Description

The Sales Team wants to know the following information about each employee:

- their categories name
- full name
- total sales amount including discount
- percentage of total sales amount including discount against his/her total sales amount across all categories
- percentage of total sales amount including discount against the total sales amount across all employees

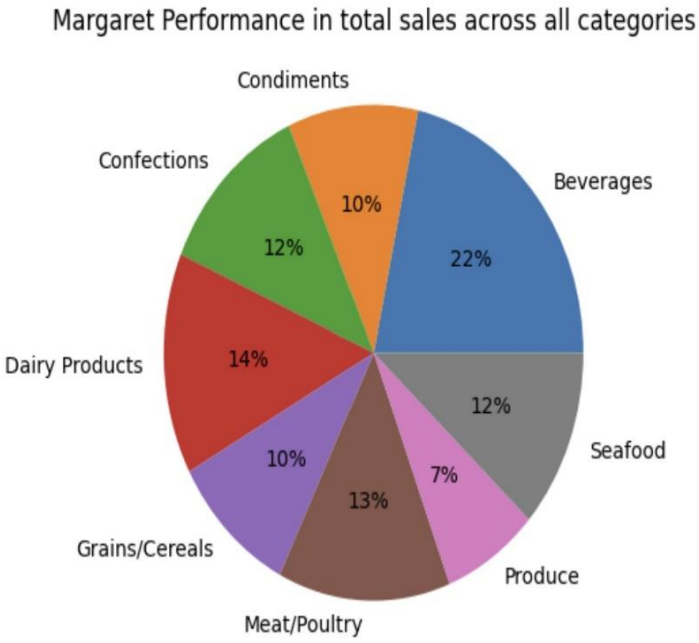
The result should be ordered by category name in alphabetic order and then total sales amount in descending order. As a result, the team can identify overall contribution to the company's sales goals.

b. Results

	category_name	employee_full_name	total_sales_with_discount	percentage_of_employee_sales	percentage_of_category_sales
1	Beverages	Margaret Peacock	50,308.21	21.60163	18.78096
2	Beverages	Nancy Davolio	46,599.35	24.2569	17.39637
3	Beverages	Janet Leverling	44,757.4	22.06833	16.70874
4	Beverages	Andrew Fuller	40,248.25	24.16764	15.02539
5	Beverages	Robert King	27,963.83	22.44861	10.4394
6	Beverages	Anne Dodsworth	19,642.55	25.40816	7.33292
7	Beverages	Laura Callahan	17,897.85	14.10809	6.68159
8	Beverages	Steven Buchanan	11,000.53	15.99093	4.1067
9	Beverages	Michael Suyama	9,450.2	12.78555	3.52793
10	Condiments	Margaret Peacock	23,314.87	10.01107	21.9854
11	Condiments	Andrew Fuller	14,850.67	8.9173	14.00385
12	Condiments	Laura Callahan	14,637.66	11.53823	13.80298
13	Condiments	Nancy Davolio	13,561.56	7.05936	12.78824
14	Condiments	Janet Leverling	13,381.64	6.59802	12.61858
15	Condiments	Anne Dodsworth	10,125.54	13.09765	9.54816
16	Condiments	Robert King	8,851.37	7.10564	8.34664
17	Condiments	Michael Suyama	4,648.47	6.2891	4.3834
18	Condiments	Steven Buchanan	2,675.3	3.88895	2.52275
19	Confections	Janet Leverling	33,622.4	16.57804	20.0902
20	Confections	Nancy Davolio	28,568.92	14.87131	17.07062
21	Confections	Margaret Peacock	27,768.73	11.9235	16.59249
22	Confections	Laura Callahan	21,699.91	17.10509	12.96622
23	Confections	Andrew Fuller	21,455.69	12.88338	12.82029
24	Confections	Robert King	14,518.99	11.65545	8.67545
25	Confections	Anne Dodsworth	8,053.16	10.41698	4.81196
26	Confections	Michael Suyama	6,859.63	9.28067	4.0988
27	Confections	Steven Buchanan	4,809.8	6.99177	2.87397
28	Dairy Products	Nancy Davolio	36,022.98	18.75146	15.36114
29	Dairy Products	Margaret Peacock	33,549.8	14.4058	14.30651
30	Dairy Products	Janet Leverling	32,320.83	15.93628	13.78244
31	Dairy Products	Robert King	27,621.86	22.17408	11.77868
32	Dairy Products	Andrew Fuller	23,812.55	14.29859	10.15429
33	Dairy Products	Steven Buchanan	21,937.63	31.88966	9.35478
34	Dairy Products	Laura Callahan	21,101.47	16.63337	8.99822
35	Dairy Products	Anne Dodsworth	21,101.12	27.29486	8.99807
36	Dairy Products	Michael Suyama	17,039.04	23.05279	7.26589
37	Grains/Cereals	Margaret Peacock	22,579.61	9.69536	23.58317
38	Grains/Cereals	Janet Leverling	21,235.01	10.47025	22.17881
39	Grains/Cereals	Andrew Fuller	11,172.95	6.70896	11.66954
40	Grains/Cereals	Laura Callahan	11,072.05	8.72761	11.56415
41	Grains/Cereals	Michael Suyama	9,410.7	12.73211	9.82896
42	Grains/Cereals	Nancy Davolio	8,465.9	4.40685	8.84217
43	Grains/Cereals	Robert King	6,535.5	5.24652	6.82597
44	Grains/Cereals	Steven Buchanan	4,027.56	5.85467	4.20657
45	Grains/Cereals	Anne Dodsworth	1,245.3	1.61083	1.30065
46	Meat/Poultry	Margaret Peacock	30,867.14	13.25391	18.9343
47	Meat/Poultry	Andrew Fuller	29,873.6	17.93803	18.32485
48	Meat/Poultry	Robert King	21,176.72	17.0001	12.99007
49	Meat/Poultry	Janet Leverling	20,502.62	10.10913	12.57657
50	Meat/Poultry	Laura Callahan	16,395.28	12.92368	10.05707
51	Meat/Poultry	Nancy Davolio	15,038.47	7.82815	9.22479
52	Meat/Poultry	Steven Buchanan	11,488.2	16.69984	7.04701
53	Meat/Poultry	Michael Suyama	9,003.69	12.18145	5.52298
54	Meat/Poultry	Anne Dodsworth	8,676.66	11.22349	5.32237
55	Produce	Nancy Davolio	19,706.25	10.25792	19.70929
56	Produce	Margaret Peacock	17,186.56	7.37966	17.18921
57	Produce	Laura Callahan	12,016.52	9.4721	12.01837
58	Produce	Janet Leverling	11,960.85	5.89748	11.9627
59	Produce	Michael Suyama	11,560.7	15.64093	11.56248
60	Produce	Robert King	10,753.38	8.63252	10.75504
61	Produce	Andrew Fuller	9,376.48	5.63024	9.37793
62	Produce	Steven Buchanan	7,109.02	10.33404	7.11012
63	Produce	Anne Dodsworth	314.81	0.40722	0.31486
64	Seafood	Margaret Peacock	27,315.93	11.72907	20.81028
65	Seafood	Janet Leverling	25,032.09	12.34246	19.07037
66	Seafood	Nancy Davolio	24,144.15	12.56804	18.3939
67	Seafood	Andrew Fuller	15,747.57	9.45586	11.99708
68	Seafood	Laura Callahan	12,041.54	9.49182	9.17369
69	Seafood	Anne Dodsworth	8,148.9	10.54082	6.20813
70	Seafood	Robert King	7,146.58	5.73708	5.44453
71	Seafood	Michael Suyama	5,940.7	8.03741	4.52584
72	Seafood	Steven Buchanan	5,744.25	8.35014	4.37618

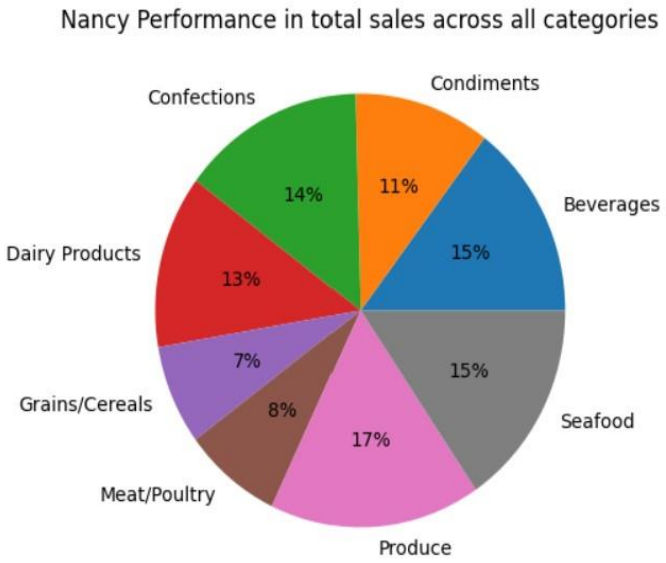
The table shows the distribution of sales across different categories and indicates which employees are most effective within specific product categories. In Beverages, Margaret had the highest total sales including discount but her performance among other employees was moderate.

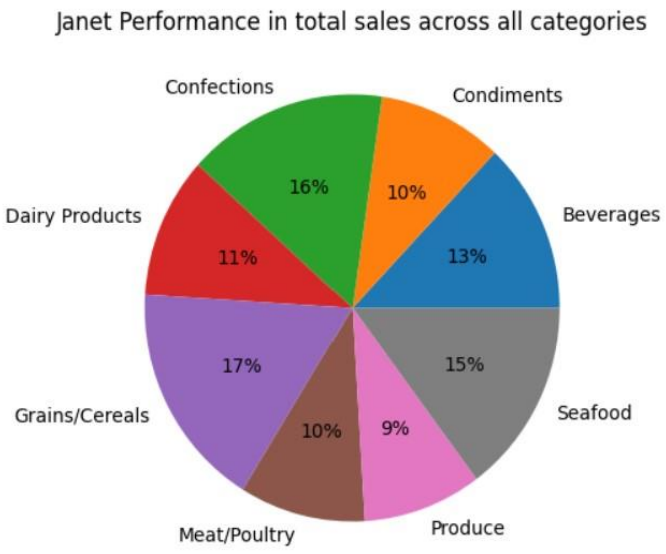
c. Key insights and Findings



- In Beverages, Condiments, Meat/Poultry, Grains/Cereals, and Seafood, Margaret was the best employee among others. She had the highest total sales including discounts across mentioned categories.
- Nancy and Janet performed well after Margaret respectively.
- According to the pie chart, Margaret showed a balanced performance and had her highest sales in beverages and dairy products respectively.
- Janet contributes 22% of his total sales of Condiments, but this only accounts for 10% of the total sales across all employees in that category. This indicates a potential area for growth.

So, the team can utilize this result to measure employees' performance across all categories.





14. Conclusion

Key Insights and Outcomes

- **Employee Performance:** Margaret, who was 55 years old at the time of hiring, demonstrated the best performance as a sales representative by driving sales volume with the highest total sales. On the other hand, Steven had the lowest sales which indicates potential for growth.
- **Pricing:** products in the "$20 - $50" price range are more popular among customers. Also, they prefer to purchase the cheapest options in beverages, confections, and seafood. Since most product prices are above average, the pricing team should consider offering more discounts and changing strategies for prices over $50 to increase their profit.
- **Logistics:** The logistics Team should adjust reorder levels based on regional demand patterns. Although most countries have sufficient stock, the total stock is very close to the number of orders in Europe and Asia-Pacific. Notably, there is a greater number of orders and reorders for beverages and condiments than the current stock level in Asia-Pacific and Europe respectively. So, the team should provide more beverages and condiments in these regions.
- **Global Performance:** The company had a strong performance in April and March 1998 with more than 70 orders. They also experienced peak performance in terms of total freight in April and January 1998. Additionally, Sweden had the longest shipping times in 1998 which means they the need to reduce shipping days.

Project Success

The project was successful in achieving its goals and meeting the specific requirements of the stakeholders. The Northwind database played a significant role to improve decision-making, and the SQL queries were useful in providing the necessary insights. Each department received valuable information for further development.

Recommendations

- More up-to-date data is needed to develop new strategies.
- For better analysis, the company can include additional data sources such as customer feedback.
- The company can also use machine learning algorithms to develop accurate models based on available data to predict future sales and optimise their trends.

In conclusion, this successful project has provided useful insights into decision-making, highlighting opportunities for future improvements. The outcomes not only met the immediate needs of the stakeholders but also paved the way for potential future enhancements.

15. References

Microsoft Corporation. (n.d.). *Northwind Traders sample database.* https://docs.microsoft.com/en-us/sql/samples/northwind-sample-database?view=sql-server-ver15

The PostgreSQL Tutorial Introduction, The PostgreSQL Global Development Group. https://www.postgresql.org/docs/

PostgreSQL Global Development Group. (n.d.). *PostgreSQL documentation.* https://www.postgresql.org/docs/

J. D. Hunter, "Matplotlib: A 2D Graphics Environment", Computing in Science & Engineering, vol. 9, no. 3, pp. 90-95, 2007