Project Overview

The project used information related to sales, customers, products, suppliers, and operations from Northwind's historical data with the aim of revealing insights that could be turned into actions for the benefit of the company. By utilizing PostgreSQL queries, I examined revenues across several lenses, including trends, customer behaviors, product and category performance, employee performance, supplier impact, shipping performance, and geographic revenue distributions. The analysis showed the growth phase of 1996-1998 was remarkable, as monthly revenue nearly tripled and year-end months attracted the most sales. Generally, the best customers, most productive employees, and most impactful suppliers receive a small part of the total reached revenue, while only a few categories, in particular Beverages, Dairy, and Confections, consistently report the highest revenues and profits.

In addition, the findings identified operational and strategic opportunities. Issues were highlighted regarding items of high demand but low stock, thus an opportunity for proactive inventory management, while differences in shipping performance, and so on, were apparent. The geographic revenue split showed how revenue and profit are largely a function of selected countries and localities, while market-conscious solutions, potentially aimed at increasing sales and revenues, were awaiting responses.

Dataset

This comprehensive dataset offers a detailed overview of a business operation, documenting everything from customer demographics to purchase details, order placement dates, and the employee responsible for processing the sale and order. The dataset includes a table of customers, employees, orders, and order_details, to see how the flow of sales and interactions occurs. Product information is organized in the products and categories tables, including prices, stock, and where it is supplied from, while the suppliers and shippers tables assist in tracking the supply chain of inventory sales. The dataset also connects people and sales to locations through region, territories, and us_states, to consider how important geography is for performance. Other tables, such as customer_demographics are empty but facilitate future use once the information is there. Overall, this dataset is an authentic and rich dataset that enables exploration and understanding of the complete picture of a business.

Queries

Monthly Revenue Analysis (1996–1998)

```sql
--1: What are the monthly total sales (revenue) for each year?--
CREATE VIEW V_Monthly_Revenue AS
SELECT
    DATE_PART('year', o.order_date) AS ORDER_YEAR,
    DATE_PART('month', o.order_date) AS ORDER_MONTH,
    SUM(od.unit_price * od.quantity * (1 - od.discount)) AS TOTAL_REVENUE
FROM orders AS o
INNER JOIN order_details AS od
    ON od.order_id = o.order_id
GROUP BY
    ORDER_YEAR, ORDER_MONTH
ORDER BY
     ORDER_YEAR, ORDER_MONTH;
```

Output

	order_year double precision	order_month double precision	total_revenue double precision
1	1996	7	27861.89512966156
2	1996	8	25485.275070743264
3	1996	9	26381.400132587554
4	1996	10	37515.72494547888
5	1996	11	45600.04521113701
6	1996	12	45239.630493214434
7	1997	1	61258.0701679784
8	1997	2	38483.6349503243
9	1997	3	38547.22010972678
10	1997	4	53032.95238894149
11	1997	5	53781.28982514166
12	1997	6	36362.80233480245
13	1997	7	51020.85751860481

Total rows: 23 | Query complete 00:00:00.063

This query summarizes total monthly revenue over three years (1996–1998) to help identify sales patterns and movement over time. This type of analysis helps highlight growth patterns, seasonality, and any additional unusual movements in revenue. The observations included consistent revenue growth, although with some anomalies from month to month.

- Revenue is clearly on an upward trajectory, almost tripling between 1996 and the beginning of 1998.
- 1996 revenue generally varied between $25k–$45k while 1998 begins at a higher number of over $94k.
- All months with the highest revenue are clustered in the second half of 1997 and the entire year of 1998, indicating massive ramping of business activity.
- The annual revenue totals indicate strong growth in revenue: 1996 approximately equal to $208k, 1997 approximately equal to $617k, and for 1998 (Jan–May) approximately equal to $441k.
- May of 1998 is a major data outlier as the only $18k month surrounded by months $94k–$123k, indicating either missing data or seasonal fluctuation in revenue.
- By May of 1998, the company achieved 71% of the annual total revenue of 1997, suggesting significant year-over-year revenue growth.

Top 10 Customers by Total Spending

```sql
--2: Who are the 10 customers with the highest total spending?--
CREATE VIEW V_Top_Customers_by_Spending AS
SELECT
    c.customer_id,
    c.company_name,
    COUNT(DISTINCT o.order_id) AS ORDER_COUNT,
    SUM(od.unit_price * od.quantity * (1 - od.discount)) AS TOTAL_SPENDING
FROM customers AS c
INNER JOIN orders AS o
    ON c.customer_id = o.customer_id
INNER JOIN order_details od
    ON o.order_id = od.order_id
GROUP BY c.customer_id, c.company_name
ORDER BY total_spending DESC
LIMIT 10;
```

Output

	customer_id [PK] character varying (5)	company_name character varying (40)	order_count bigint	total_spending double precision
1	QUICK	QUICK-Stop	28	110277.30503039382
2	ERNSH	Ernst Handel	30	104874.97814367746
3	SAVEA	Save-a-lot Markets	31	104361.94954039395
4	RATTC	Rattlesnake Canyon Grocery	18	51097.80082826822
5	HUNGO	Hungry Owl All-Night Grocers	19	49979.90508149548
6	HANAR	Hanari Carnes	14	32841.369948457475
7	KOENE	Königlich Essen	14	30908.383872538416
8	FOLKO	Folk och fä HB	19	29567.562490026656
9	MEREP	Mère Paillarde	13	28872.19015611842
10	WHITC	White Clover Markets	14	27363.604900210797

This query reveals the 10 customers who account for the most total revenue based on aggregated spend across all orders, indicating which customers are most important to fleet revenue, as well as how spending is distributed among high, mid, and low-tiered customers. This information may guide an analysis of customer revenue concentration and potential customer vulnerability.

- The top 3 customers are “QUICK,” “ERNSH,” and “SAVEA,” each exceeded $100k in total spend and form a definite high-tiered value bucket well above the remaining customers.
- There was a 50%+ decrease in revenue activity from the 3rd to the 4th highest spender, confirming established heavy reliance on limited top-tier customers.
- Mid-tiered customers spent in the $27k–$51k range, which is more evenly distributed than the value generated by high-tier customers.
- Order count does not equate to spend, as several customers placed more orders but at a lesser value than other customers.
- As segmentation is developed based on revenue generation, the following value tiers are apparent: high-value (>100k), mid-value ($30k–$60k), and standard value ($27k–$33k).
- Current revenue activity and total sales are highly concentrated, as losing even one of the top 3 customers will likely significantly impact revenue.

Revenue by Product Category

```
--3: Which product categories are bringing in the most revenue?--
CREATE VIEW V_Top_Categories_by_Revenue AS
SELECT
    c.category_name,
    SUM(od.unit_price * od.quantity * (1 - od.discount)) AS TOTAL_REVENUE
FROM products AS p
INNER JOIN categories AS c
    ON p.category_id = c.category_id
INNER JOIN order_details AS od
    ON p.product_id = od.product_id
GROUP BY c.category_name
ORDER BY TOTAL_REVENUE DESC;
```

Output

	category_name character varying (15)	total_revenue double precision
1	Beverages	267868.17978624784
2	Dairy Products	234507.2845305777
3	Confections	167357.22547338586
4	Meat/Poultry	163022.3602687388
5	Seafood	131261.7365558489
6	Condiments	106047.08460955074
7	Produce	99984.58007357619
8	Grains/Cereals	95744.58735543818

This query calculates the total sales for each product category to understand which areas contribute the most to mixed sales. It identifies the top, medium, and low sales categories to inform inventory needs, marketing efforts, and strategic focus. The findings show a strong two-category dominance.

- Beverages and Dairy Products are the top categories contributing to a combined $502k of sales, nearly 34% of total category sales.
- There is a strong medium tier amongst Confection, Meat/Poultry, and Seafood, contributing sales between $131k and $167k.
- Below those categories are the low categories of Condiments, Produce, and Grains/Cereals, generating sales below $110k, albeit not too much of a gap between them.
- The largest dollar gap is between Dairy Products ($234k) and Confections ($167k) for a total of a $67k gap, marking the total shift from high dollar categories to medium dollar categories.

- The smallest computer gap is between Produce ($100k) and Grains/Cereals ($95k), which highlights similar sales performance.
- Beverages and Dairy Products would be strategic areas to focus on, with expansion opportunities being a potential for the medium categories, and possibly looking at pricing or SKUs for the lower categories.

Units Sold by Product

```
--4: Which product has overall the highest number of units sold?--
CREATE VIEW V_Top_Products_by_Units_Sold AS
SELECT
    p.product_name,
    p.product_id,
    SUM(od.quantity) AS TOTAL_UNITS_SOLD
FROM products AS p
INNER JOIN order_details AS od
        ON od.product_id = p.product_id
GROUP BY p.product_name, p.product_id
ORDER BY TOTAL_UNITS_SOLD DESC;
```

Output

	product_name character varying (40)	product_id [PK] smallint	total_units_sold bigint
1	Camembert Pierrot	60	1577
2	Raclette Courdavault	59	1496
3	Gorgonzola Telino	31	1397
4	Gnocchi di nonna Alice	56	1263
5	Pavlova	16	1158
6	Rhönbräu Klosterbier	75	1155
7	Guaraná Fantástica	24	1125
8	Boston Crab Meat	40	1103
9	Tarte au sucre	62	1083
10	Flotemysost	71	1057
11	Chang	2	1057
12	Sir Rodney's Scones	21	1016
13	Jack's New England Clam Chow...	41	981

Total rows: 77 | Query complete 00:00:00.108

This query ranks products by total number of units sold, ultimately highlighting which products sold in high volume and which sold in low volume. This analysis communicates demand concentration, increasing trends in terms of category performance, and volume tiers for every SKU available through the platform. The results of this analysis provide support for future inventory plans and product strategy.

- Camembert Pierrot is the top-selling product at 1,577 units and is followed closely by other dairy items, such as Raclette Courdavault and Gorgonzola Telino.
- There are only 10 SKUs showing sales in excess of 1,000 units, creating a very strong top tier of unit sales.
- Dairy again dominates the bestselling list of SKUs, which aligns with a previous finding that dairy is one of the highest revenue categories.
- There are clear unit sales tiers by SKU: Tier 1 (> 1000 units), Tier 2 (700 to 1000), Tier 3 (400 to 700), and Tier 4 (< 400) units sold, and the unit sales sharply fall off after approximately the top 15 best-selling products.
- There are products that are sold in low volume, such as Mishi Kobe Niku (95 units) or Chocolade (138 units), which do show some demand in their low sales, and may warrant a review of inventory and pricing.
- High volume categories (Dairy, Beverage, Confectionery) are responsible for the most unit sales, while high price categories (Seafood, Meat) make a higher contribution to revenue than they do for volume total.

Top Employees by Total Sales

```sql
--5: Who are the employees with the most sales?--
CREATE VIEW V_Top_Employees_by_Sales AS
SELECT
    e.employee_id,
    e.first_name || ' ' || e.last_name AS EMPLOYEE_NAME,
    SUM(od.unit_price * od.quantity * (1 - od.discount)) AS TOP_SALES_BY_EMPLOYEE
FROM employees AS e
INNER JOIN orders AS o
    ON e.employee_id = o.employee_id
INNER JOIN order_details AS od
    ON o.order_id = od.order_id
GROUP BY
    e.employee_id,
    EMPLOYEE_NAME
ORDER BY TOP_SALES_BY_EMPLOYEE DESC;
```

Output

	employee_id [PK] smallint	employee_name text	top_sales_by_employee double precision
1	4	Margaret Peacock	232890.84598203757
2	3	Janet Leverling	202812.84238787484
3	1	Nancy Davolio	192107.60465689204
4	2	Andrew Fuller	166537.75528742478
5	8	Laura Callahan	126862.2774778762
6	7	Robert King	124568.23485103557
7	9	Anne Dodsworth	77308.06623830303
8	6	Michael Suyama	73913.12937797053
9	5	Steven Buchanan	68792.28239394959

The report ranks employees by total sales revenue from orders. It identifies the highest-performing sales employees, emphasizes revenue concentration among those, and points out the gaps between high-performing employees and lower-performing employees. This is useful in assessing productivity and productivity contributions from members of the sales team.

- Top-performing employee Margaret Peacock shows considerable sales revenue ($232,890.85) ahead of her teammates.
- The top three performers (Margaret, Janet, Nancy) generate a base of ~ $627k and are establishing themselves as a revenue-generating peer group.
- A noteworthy gap between high and low performance is demonstrated, with nearly 3.4× revenue between the highest performer (Margaret) and the lowest performer (Steven Buchanan at ~ $68k).
- Andrew Fuller and Laura Callahan are examples of lower mid-tier performers who are contributing to productivity, but are still a considerable distance away from the revenue generated from the top three.
- All three bottom performers are under $80k in revenue; it is worth considering if any of them could benefit from training and/or assessing territory discussions.
- Overall distribution of sales is top-heavy, meaning a greater share of revenue is contributed by fewer employees, an important point in workforce planning and incentive strategies.

Annual Average Order Value (AOV) by Year

```sql
--6: What is the annual average order value (AOV) for each year?--
CREATE VIEW V_Annual_Average_Order_Value AS
SELECT
    DATE_PART('year', o.order_date) AS YEAR,
    SUM(od.unit_price * od.quantity * (1 - od.discount)) / COUNT(od.order_id) AS AVG_ORDER_VALUE
FROM order_details AS od
INNER JOIN orders AS o
    ON od.order_id = o.order_id
GROUP BY YEAR
ORDER BY YEAR;
```

Output

	year double precision	avg_order_value double precision
1	1996	513.787582673636
2	1997	582.7055735530691
3	1998	637.6611653803774

This query evaluates the Average Order Value for every year in order to gain a better understanding of how customer spending per order changes over time. It can help measure the quality of revenue growth and indicate whether customers place larger or more valuable orders every year.

- AOV increases every year from 1996, moving to 1997, and then in 1998 shows clear growth upward.
- Total growth from 1996 to 1998 = +123.87, equating to a ~24.1% increase in 2 years.
- Most importantly, if we notice, customers spend more per order over time, indicating improved order value, robust product mix, or restricted discounting.
- Growth rate is increasing, but the pace is slowing (13.4%→ 9.4%).
- The pacing slowing down may indicate we have stabilized as a market, or we have reached a natural limit of customer spending peak.
- Overall, the trend denotes robust improvement in revenue even without asking for additional orders — all while providing a durable customer experience, indicating a good service-oriented business model.

Supplier Revenue Summary

```
--7: Which supplier's products account for the highest share of total revenue?--
CREATE VIEW V_Supplier_Revenue_Summary AS
SELECT
    s.company_name AS SUPPLIER_COMPANY_NAME,
    SUM(od.unit_price * od.quantity * (1 - od.discount)) AS REVENUE
FROM products AS p
INNER JOIN suppliers AS s
    ON s.supplier_id = p.supplier_id
INNER JOIN order_details AS od
    ON od.product_id = p.product_id
GROUP BY SUPPLIER_COMPANY_NAME
ORDER BY REVENUE DESC;
```

Output

	supplier_company_name character varying (40)	revenue double precision
1	Aux joyeux ecclésiastiques	153691.2755027201
2	Plutzer Lebensmittelgroßmärkte AG	145372.39966947702
3	Gai pâturage	117981.18020485654
4	Pavlova, Ltd.	106459.77607838642
5	G'day, Mate	65626.77016849845
6	Forêts d'érables	61587.5698526758
7	Specialty Biscuits, Ltd.	59032.0801634851
8	Pasta Buttini s.r.l.	50254.609883744415
9	Formaggi Fortini s.r.l.	48225.164374038046
10	Norske Meierier	43141.5101294113
11	Leka Trading	42017.64502122432
12	Grandma Kelly's Homestead	41953.299975756556
13	Heli Süßwaren GmbH & Co. KG	38653.419428995025

Total rows: 29 | Query complete 00:00:00.147

This query shows revenue recognized from product sales at each supplier. It shows which suppliers are the main contributors to total revenue, and the impact of supplier revenue concentration across the supply base.

- Aux joyeux ecclésiastiques comes in as the top supplier with $153,691.28, leading all other suppliers by a wide margin.

- The top four suppliers generate greater than $523k collectively, which indicates strong supplier revenue concentration.
- There is approximately a $35k revenue gap between Rank 1 and Rank 3, indicating a strong lead from the top suppliers.
- After the top 4 suppliers, revenue decreases quickly (from $106k to $65k), suggesting somewhat of a reliance on a small number of top suppliers.
- Several suppliers at the bottom of the list earn very little, including Zaanse Snoepfabriek ($5.3k) and Refrescos Americanas LTDA ($4.5k).
- High supplier concentration could entail a level of supply chain risk if anything were to happen to one of the top suppliers.

Average Shipping Time by Carrier

```sql
--8: What is the average shipping time, per carrier?--
CREATE VIEW V_Average_Shipping_Time_by_Shipper AS
SELECT
    sh.company_name AS SHIPPER_NAME,
    AVG(o.shipped_date - o.order_date) AS AVG_SHIPPING_DAYS
FROM orders AS o
INNER JOIN shippers AS sh
    ON o.ship_via = sh.shipper_id
WHERE o.shipped_date IS NOT NULL
GROUP BY sh.company_name
ORDER BY AVG_SHIPPING_DAYS;
```

Output

	shipper_name character varying (40)	avg_shipping_days numeric
1	Federal Shipping	7.4738955823293173
2	Speedy Express	8.5714285714285714
3	United Package	9.2349206349206349

This query estimates the average number of days each shipping carrier takes to deliver orders. This makes comparing shipper performance and determining which carrier has the fastest shipping speeds easy or the slowest possible.

- Federal Shipping is the fastest carrier with an average of 7.47 days.
- United Package is the slowest at 9.23 days with a noticeable 1.76-day gap behind the fastest.
- Speedy Express is in between, with a mid-range performance of 8.57 days, which is slower than Federal Shipping but faster than United Package.
- Faster shipping is likely to result in better customer satisfaction and repeat purchases.
- If reducing delivery time is desired, Federal Shipping should always be used for key and/or urgent orders.
- These shipping differences are helpful for assessing the overall cost-speed tradeoff when negotiating contracts and possible premium shipping.

High-Demand and Low-Stock Products

```sql
--9: Which items are low on stock yet are in high demand?--
CREATE VIEW V_High_Demand_Low_Stock_Products AS
SELECT
    p.product_name,
    p.units_in_stock,
    SUM(od.quantity) AS DEMAND
FROM products AS p
INNER JOIN order_details AS od
    ON p.product_id = od.product_id
GROUP BY
    p.product_name,
    p.units_in_stock
HAVING
    p.units_in_stock < 20
    AND SUM(od.quantity) > 30
ORDER BY
    DEMAND DESC;
```

Output

	product_name character varying (40)	units_in_stock smallint	demand bigint
1	Camembert Pierrot	19	1577
2	Gorgonzola Telino	0	1397
3	Tarte au sucre	17	1083
4	Chang	17	1057
5	Sir Rodney's Scones	3	1016
6	Alice Mutton	0	978
7	Outback Lager	15	817
8	Mozzarella di Giovanni	14	806
9	Scottish Longbreads	6	799
10	Uncle Bob's Organic Dried Pears	15	763
11	Gumbär Gummibärchen	15	753
12	Thüringer Rostbratwurst	0	746
13	Perth Pasties	0	722

Total rows: 26 | Query complete 00:00:00.083

This query shows products with strong customer demand and low inventory (stock < 20 units and total demand > 30). It provides answers to address low-stock products that are vulnerable to stockouts and helps with rebuilding and replenishing decisions.

- Several of the top-demand products—like Camembert Pierrot, Gorgonzola Telino, and Tarte au sucre—are at critically low stock.
- Several items are already out of stock, such as Gorgonzola Telino and Alice Mutton, presenting an immediate revenue loss risk.
- Chang and Mozzarella di Giovanni—those will run out quickly soon, with the demand being quite high.
- Several items with moderately low stock levels (10–19), including Chang and Mozzarella di Giovanni, will likely run out of stock, given the respectable demand levels.
- Identified and quickly out-of-stock products (0–4 units) must have some time-sensitive decisions made by sellers, especially if the demand levels are 700–1400+ units.
- These items will contribute to the overall sales, meaning stockouts will lead to significant lost sales and customer dissatisfaction.

Revenue Share by Country

```sql
--10: What is the proportion of total revenue by country?--
CREATE VIEW V_Country_Revenue_Share AS
WITH COUNTRY_REVENUE AS (
    SELECT
        c.country AS COUNTRY,
        SUM(od.unit_price * od.quantity * (1 - od.discount)) AS TOTAL_REVENUE
    FROM customers  AS c
    INNER JOIN orders AS o
        ON o.customer_id = c.customer_id
    INNER JOIN order_details AS od
        ON o.order_id = od.order_id
    GROUP BY COUNTRY
)
SELECT
    COUNTRY,
    TOTAL_REVENUE,
    ROUND(
        CAST(TOTAL_REVENUE * 100 / SUM(TOTAL_REVENUE) OVER() AS numeric), 2) AS REVENUE_PERCENTAGE
FROM COUNTRY_REVENUE
ORDER BY REVENUE_PERCENTAGE DESC;
```

Output

	country character varying (15)	total_revenue double precision	revenue_percentage numeric
1	USA	245584.61030220677	19.40
2	Germany	230284.63325421108	18.19
3	Austria	128003.8381586512	10.11
4	Brazil	106925.7764042853	8.45
5	France	81358.322529562	6.43
6	UK	58971.309912234916	4.66
7	Venezuela	56810.628758941646	4.49
8	Sweden	54495.13992099263	4.31
9	Canada	50196.29010077998	3.97
10	Ireland	49979.90508149549	3.95
11	Belgium	33824.85528823436	2.67
12	Denmark	32661.02250821233	2.58

Total rows: 21 Query complete 00:00:00.095

This query finds the total revenue collected by country and that country's share of total revenue. It allows the team to identify key markets and revenue concentration, and identify potential regional opportunities for growth.

- Just three countries, the USA, Germany, and Austria account for nearly 48% of total revenue, and the USA is the single largest market, accounting for 19.4% of total revenue.
- European countries as a block account for a large revenue share and have strong regional performance.
- Brazil and Venezuela are emerging mid-tier markets with 13% combined revenue share.
- Several countries (Poland, Norway, Italy, Spain, etc.) are contributing less than 1.5% total revenue, indicating a less developed or niche market for those countries.

Top 5 Most Profitable Orders

```sql
--10: What is the proportion of total revenue by country?--
CREATE VIEW V_Country_Revenue_Share AS
WITH COUNTRY_REVENUE AS (
    SELECT
        c.country AS COUNTRY,
        SUM(od.unit_price * od.quantity * (1 - od.discount)) AS TOTAL_REVENUE
    FROM customers  AS c
    INNER JOIN orders AS o
        ON o.customer_id = c.customer_id
    INNER JOIN order_details AS od
        ON o.order_id = od.order_id
    GROUP BY COUNTRY
)
SELECT
    COUNTRY,
    TOTAL_REVENUE,
    ROUND(
        CAST(TOTAL_REVENUE * 100 / SUM(TOTAL_REVENUE) OVER() AS numeric), 2) AS REVENUE_PERCENTAGE
FROM COUNTRY_REVENUE
ORDER BY REVENUE_PERCENTAGE DESC;
```

Output

	order_id [PK] smallint	total_profit double precision
1	10865	16387.4999871477 5
2	10981	15810
3	11030	12615.05006790161 1
4	10889	11380
5	10417	11188.40013980865 5

This query identifies the orders that yielded the highest net profit after applying discounts, which is helpful to shed light on orders generating the most margin for the business, as well as behavioral patterns relating to higher-value sale activity.

- The ranked order ID 10865 was a profit of 16,387.50, which is significantly higher than rank #5, indicating profit concentration.
- The 5,200-profit differential between rank #1 and rank #5 demonstrates that only a limited number of orders accumulated significant profit.
- Some of these orders will likely include high-value products, larger quantities, or very little discounting.
- High profit orders are important for maintaining good margins, which often indicate lucrative customer segments and/or product mixes.
- By analyzing each of these high-profit orders, we may be able to determine patterns associated with those orders (recurring customers, territory, or sales rep), which can give us insight into determining who and where high-profit orders come from on a consistent basis.

- Focusing on each of these orders, along with exerting marketing and loyalty communication towards the specific customers, can provide additional margin sustainability over the long run.

Year-over-Year (YOY) Customer Spending Changes

```sql
--12: Which customers have the biggest increase or decrease in spending from year to year?--
CREATE VIEW V_Customer_YOY_Spending AS
WITH CUSTOMER_YEARLY_SPENDING AS (
    SELECT
        c.customer_id,
        DATE_PART('year', o.order_date) AS YEAR,
        c.company_name,
        SUM(od.unit_price * od.quantity * (1 - od.discount)) AS TOTAL_SPENDING
        FROM customers AS c
        INNER JOIN orders AS o
            ON c.customer_id = o.customer_id
        INNER JOIN order_details AS od
            ON o.order_id = od.order_id
        GROUP BY
            c.customer_id,
            c.company_name,
            YEAR
),
CUSTOMER_YEAR_OVER_YEAR AS (
    SELECT
        customer_id,
        company_name,
        YEAR,
        TOTAL_SPENDING,
        LAG(TOTAL_SPENDING) OVER (PARTITION BY customer_id ORDER BY YEAR) AS PREVIOUS_YEAR_SPENDING
    FROM CUSTOMER_YEARLY_SPENDING
)
SELECT
    customer_id,
    company_name,
    year,
    TOTAL_SPENDING,
    PREVIOUS_YEAR_SPENDING,
    TOTAL_SPENDING - COALESCE(PREVIOUS_YEAR_SPENDING, 0) AS YOY_CHANGE
FROM CUSTOMER_YEAR_OVER_YEAR
WHERE PREVIOUS_YEAR_SPENDING IS NOT NULL
ORDER BY ABS(TOTAL_SPENDING - PREVIOUS_YEAR_SPENDING) DESC;
```

Output

	customer_id [PK] character varying (5)	company_name character varying (40)	year double precision	total_spending double precision	previous_year_spending double precision	yoy_change double precision
1	QUICK	QUICK-Stop	1997	61109.91012364827	11950.080003829004	49159.830119819264
2	SAVEA	Save-a-lot Markets	1997	57713.57465326578	10338.264898133875	47375.30975513191
3	ERNSH	Ernst Handel	1997	48096.26317733167	15568.065068242922	32528.198109088746
4	QUICK	QUICK-Stop	1998	37217.31490291655	61109.91012364827	-23892.595220731717
5	SAVEA	Save-a-lot Markets	1998	36310.1099889943	57713.57465326578	-21403.46466427148
6	HANAR	Hanari Carnes	1998	23821.199969008565	6022.7699720579385	17798.429996950625
7	MEREP	Mère Paillarde	1997	23332.310070187003	5539.88008593142	17792.429984255585
8	SIMOB	Simons bistro	1998	232.08500996507703	16232.412604600191	-16000.327594635113
9	SIMOB	Simons bistro	1997	16232.412604600191	352.59999763965607	15879.812606960535
10	WARTH	Wartian Herkku	1998	269.99999955296516	12262.942562392802	-11992.942562839837
11	HUNGO	Hungry Owl All-Night Grocers	1997	20454.4048591581	9123.380134060679	11331.024725097423
12	FOLKO	Folk och fä HB	1997	13314.669975699782	2608.8249821439385	10705.844993555844
13	HILAA	HILARION-Abastos	1997	13482.74399596101	3242.8200022834535	10239.923993677556

Total rows: 145 Query complete 00:00:00.165

This query assesses how the total spending of each customer fluctuates over the course of one year compared to the previous year. It shows which customers made the largest increases in their spending and which the biggest decrease. Understanding both spending increases and decreases is important to find opportunities for growth or the risk of a customer decreasing their spending.

- Several customers made considerable increases in their spending on a year-over-year (YOY) basis. For example, QUICK, SAVEA, and ERNSH each increased spending by 30K–49K within the span of a single year.
- Other example customers like HANAR and MEREP grew as well, suggesting either growing demand or a new buying behavior.
- These increases could be due to either one-off bulk purchases, new contracts, or an increased adoption of their product lines.
- Some customers also had large drops in YOY spending, like QUICK (-23k), SAVEA (-21k), and SIMOB (-16k), suggesting sudden drops in purchasing from last year.
- Drops could also suggest the contract period ended, competitor growth, issues with service level, or other typical market fluctuations.
- High volatility, meaning both big increases and then a big decrease, in a customer's spending, has risk for revenue but also may have a high opportunity for growth if managed appropriately.

Top Revenue-Generating Territories

```sql
--13: Which regions are the ones with the highest revenue?--
CREATE VIEW V_Territory_Revenue_Summary AS
SELECT
    t.territory_id,
    t.territory_description,
    SUM(od.unit_price * od.quantity * (1 - od.discount)) AS TOTAL_REVENUE
FROM territories AS t
INNER JOIN employee_territories AS et
    ON t.territory_id = et.territory_id
INNER JOIN employees AS e
    ON et.employee_id = e.employee_id
INNER JOIN orders AS o
    ON e.employee_id = o.employee_id
INNER JOIN order_details AS od
    ON o.order_id = od.order_id
GROUP BY
    t.territory_id,
    t.territory_description
ORDER BY TOTAL_REVENUE DESC;
```

Output

	territory_id [PK] character varying (20)	territory_description character varying (60)	total_revenue double precision
1	20852	Rockville	232890.84598203757
2	27511	Cary	232890.84598203757
3	27403	Greensboro	232890.84598203757
4	32859	Orlando	202812.84238787484
5	30346	Atlanta	202812.84238787484
6	33607	Tampa	202812.84238787484
7	31406	Savannah	202812.84238787484
8	19713	Neward	192107.60465689204
9	06897	Wilton	192107.60465689204
10	01730	Bedford	166537.75528742478

Total rows: 49 | Query complete 00:00:00.153

This query seeks to determine which sales territories generate the most total revenue. This is valuable as it elevates the highest performing sales locations and aids in determining if there are geographic similarities in sales performance and customer interest.

- The highest revenue-generating territories are Rockville, Cary, and Greensboro, each bringing in roughly $232,891 in revenue.
- Following a significant gap, the next tier would either be Orlando, Atlanta, Tampa, and Savannah, with one indicated $202,813 in revenue, indicating another cluster of high-performing sales territories.
- Revenue is distributed across many of the largest U.S. urban markets, indicating that the company has diversified and geographically dispersed sales strength.
- In high revenue-generating territories, it is likely that these markets are home to high performers, high revenue customers, or heavily populated commercial areas.
- Areas generating a consistent revenue range between several cities suggest that the company has stable market penetration, rather than being reliant on one geographic area.
- These observations can shape or inform the company's regional sales strategy, including resource allocation or expansion planning information.

Frequently Ordered Product Pairs (Market Basket Analysis)

```sql
--14: What are the products which are often ordered together? (This is Market basket analysis.)--
CREATE VIEW V_Frequently_Ordered_Products AS
SELECT
    p1.product_name AS product_1,
    p2.product_name AS product_2,
    COUNT(*) AS TIMES_ORDERED_TOGETHER
FROM order_details od1
INNER JOIN order_details od2
    ON od1.order_id = od2.order_id
    AND od1.product_id < od2.product_id
INNER JOIN products p1
    ON od1.product_id = p1.product_id
INNER JOIN products p2
    ON od2.product_id = p2.product_id
GROUP BY
    product_1,
    product_2
ORDER BY TIMES_ORDERED_TOGETHER DESC
LIMIT 40;
```

Output

	product_1 character varying (40)	product_2 character varying (40)	times_ordered_together bigint
1	Sir Rodney's Scones	Sirop d'érable	8
2	Pavlova	Gorgonzola Telino	7
3	Pavlova	Camembert Pierrot	6
4	Pavlova	Tarte au sucre	6
5	Gorgonzola Telino	Mozzarella di Giovanni	6
6	Camembert Pierrot	Flotemysost	6
7	Nord-Ost Matjeshering	Tourtière	6
8	Gorgonzola Telino	Manjimup Dried Apples	5
9	Pâté chinois	Tarte au sucre	5
10	Chang	Raclette Courdavault	5

Total rows: 40 Query complete 00:00:00.124

This query uses a self-join on the order_details table to find the product combinations ordered together most frequently (co-occurred). This is useful for identifying customer purchasing behaviors to promote and cross-sell, increase promotional activity, and more accurately manage inventory levels.

- Importantly, the top product pair is "Sir Rodney's Scones" & "Sirop d'érable", which were ordered together 8 times.
- Pavlova appears frequently in multiple top pairs, one with Gorgonzola Telino (7 times), and another with Camembert Pierrot (6 instances), identifying Pavlova as a key product bundled with other products.
- Some pairs obviously reflect combinations of dressings (e.g., Chang & Raclette Courdavault) or logical options (e.g., Louisiana Fiery Hot Pepper Sauce & Flotemysost).
- As an example, the analysis can create cross-selling opportunities in two ways: one is to recommend one of the products if one of the frequently co-occurred products is in a customer's cart.
- Also, the analysis is useful for promotional bundling. For example, two of the products sold most frequently together could be included in a promotional effort, suggesting that customers could save money if they purchased both, to increase average order value.
- Lastly, this analysis informs inventory management simultaneously to ensure that complementary products are defective together to avoid missed sales..

Top Revenue Generating Product Per Category

```sql
--15: For each category, which product is the one with the highest revenue?--
CREATE VIEW V_Top_Product_by_Category AS
WITH PRODUCT_REVENUE AS (
    SELECT
        p.product_id,
        p.product_name,
        c.category_name,
        SUM(od.unit_price * od.quantity * (1 - od.discount)) AS total_revenue
    FROM products p
    INNER JOIN categories AS c
        ON p.category_id = c.category_id
    INNER JOIN order_details AS od
        ON p.product_id = od.product_id
    GROUP BY p.product_id, p.product_name, c.category_name),
RANKED_PRODUCTS AS (
    SELECT
        product_id,
        product_name,
        category_name,
        total_revenue,
        RANK() OVER (PARTITION BY category_name ORDER BY total_revenue DESC) AS REVENUE_RANK
    FROM PRODUCT_REVENUE)
SELECT
    product_name,
    category_name,
    total_revenue
FROM RANKED_PRODUCTS
WHERE REVENUE_RANK = 1
ORDER BY category_name;
```

Output

	product_name character varying (40)	category_name character varying (15)	total_revenue double precision
1	Côte de Blaye	Beverages	141396.7356273254
2	Vegie-spread	Condiments	16701.095047264098
3	Tarte au sucre	Confections	47234.969978504174
4	Raclette Courdavault	Dairy Products	71155.69990943
5	Gnocchi di nonna Alice	Grains/Cereals	42593.0598222503
6	Thüringer Rostbratwurst	Meat/Poultry	80368.6724385033
7	Manjimup Dried Apples	Produce	41819.65024582073
8	Carnarvon Tigers	Seafood	29171.874963399023

Total rows: 8 | Query complete 00:00:00.088

This query identifies the one product in each category that produces the greatest total revenue. It allows us to concentrate marketing, inventory, and supply planning on the products that generate the most revenue in each category.

- Within the Beverages category, Côte de Blaye generates $141,396.74 in revenue, more than any other product within the category.
- For Dairy Products, the category leader is Raclette Courdavault with $71,155.70 in revenue, which reinforces Dairy as one of the most lucrative revenue categories.
- Within Meat/Poultry, Thüringer Rostbratwurst creates value at $80,368.67, illustrating that high ticket items move more revenue even though they do not necessarily sell more units.
- One product can contribute extensively to the revenue for the related category; future promotions and stock levels should be assigned to these products.
- Full focus of campaigns and inventory on the related categories is important for marketing and inventory accumulated sales.

Total Profit by Country

```
--Total Profit by Country--
CREATE VIEW V_Country_Profit AS
SELECT
    c.country AS country,
    SUM(od.unit_price * od.quantity * (1 - od.discount)) - SUM(o.freight) AS TOTAL_PROFIT_COUNTRY
FROM customers AS c
INNER JOIN orders AS o
    ON c.customer_id = o.customer_id
INNER JOIN order_details AS od
    ON o.order_id = od.order_id
GROUP BY c.country
ORDER BY TOTAL_PROFIT_COUNTRY DESC;
```

Output

	country character varying (15)	total_profit_country double precision
1	USA	199418.75873970677
2	Germany	192113.86762921108
3	Austria	100272.3908930262
4	Brazil	92342.7998417853
5	France	68731.750263937
6	UK	50480.781591922416
7	Venezuela	48938.188329254146
8	Sweden	44349.02371005513
9	Canada	43874.39117499873

Total rows: 21 Query complete 00:00:00.088

This query computes the total profit generated per country by deducting freight expenses from overall sales revenue. This analysis will provide insight into which countries are providing the biggest contribution to overall profit and will help drive a focused strategy towards high-margin markets.

- The USA ($199,418.76) leads the way, followed by Germany ($192,113.87), making it the two most lucrative markets.
- A number of countries made a considerable contribution to overall profits, but are far behind the two leaders: Austria ($100,272.39), Brazil ($92,342.80), and France ($68,731.75).
- Poland and Norway ($3,070.42) and Argentina have generated little profit and therefore probably have low sales volumes or high shipping costs.
- Profits are being contributed by a small number of countries, and the risk of dependency of those countries demands changes or a desire to expand outside those two countries.
- The USA and Germany should remain the focus of marketing, determining inventory allocation, and the potential for expansion efforts.
- Countries that are low profit should somehow optimize costs (e.g., shipping or price constraints) or consider exiting the market, or focusing on specific customers or products.

Total Profit by Product

```sql
--Total Profit by product--
CREATE VIEW V_Product_Profit AS
SELECT
    p.product_name AS product_name,
    SUM(od.unit_price * od.quantity * (1 - od.discount)) - SUM(o.freight) AS TOTAL_PROFIT_PRODUCT
FROM products AS p
INNER JOIN order_details AS od
    ON p.product_id = od.product_id
INNER JOIN orders AS o
    ON od.order_id = o.order_id
GROUP BY p.product_name
ORDER BY TOTAL_PROFIT_PRODUCT DESC;
```

Output

	product_name character varying (40)	total_profit_product double precision
1	Côte de Blaye	133936.1565257629
2	Thüringer Rostbratwurst	73882.25202834705
3	Raclette Courdavault	63147.05147193
4	Tarte au sucre	41248.480232410424
5	Camembert Pierrot	39179.86115480155
6	Manjimup Dried Apples	37435.95053878948
7	Gnocchi di nonna Alice	36593.91968553155
8	Alice Mutton	29395.750333560703
9	Carnarvon Tigers	26897.034875508398

Total rows: 77 Query complete 00:00:00.113

This query computes total product profit by deducting freight costs from total sales after discounting sales. It identifies which products are most profitable to direct stock, marketing, and price development efforts toward the most profitable products.

- Côte de Blaye tops the list of products with a total profit of $133,936.16, well ahead of the next product.
- The top products in terms of profit include dairy, meat/poultry, and some specialty foods.
- Some of the closest low price or niche products (e.g., Geitost) are minimal profit, approximately $45 per year.
- Most of the profit is generated by a relatively small number of products, following the same patterns as revenue volume.
- To aid in stock management, promotions, and supplier deals, identify which products have the highest profit.

- As a potential optimization of portfolio profit, low-profit products may be worth evaluating and analyzing for price and marketing, or discontinuing.