Supply Chain Optimization through Data Analytics

Principles of Business Analytics Project

Team members:
Xiaokang Li
Anass Saissi Hassani
Ettore Bertolini

20 October 2025

Abstract

This project explores how data organisation and SQL-based analytics can improve operational decision-making for Northwind Traders, a fictional gourmet food supplier. Our team analysed a public Kaggle dataset to identify opportunities for supply chain optimization, revenue growth, and customer segmentation. Using PostgreSQL, we created a relational schema with seven core tables, enabling the development and execution of ten analytical SQL queries aligned with Northwind's strategic aims. Our findings show that over 20% of total revenue comes from the top three products, the primary shipping partner is responsible for most delays, and several categories are experiencing declining sales. These insights offer actionable strategies for employee training, improved shipment reliability, and better inventory management. This project demonstrates how structured databases and targeted analytics can drive measurable improvements in efficiency and profitability for organisations dependent on supply chain operations.

Contents

1. Business Scenario and Client Context

1.1. Client Background

Northwind Traders is a wholesale distribution company that supplies food and beverage products to global retailers and restaurants. The company processes hundreds of orders each month involving employees, customers, suppliers, and shippers. Although Northwind has a strong data collection system, its information is stored in separate files, which limits the ability to perform comprehensive performance analysis.

1.2. Business Challenge

The client faces three significant issues:

- Lack of integrated visibility across customer orders, suppliers, and shippers results in poor supply chain coordination.
- Limited insight into customer profitability and product performance reduces pricing and marketing efficiency.
- Manual reporting processes that delay data-driven decision-making.

1.3. Objectives

The goal of this analysis is to assist Northwind Traders by organizing data and enabling SQL-based analytics, allowing the company to achieve the following objectives and monitor trends, identify inefficiencies, and make informed decisions that reduce costs and enhance service quality.

- Enhance operational efficiency by analyzing delivery times and carrier performance.
- Improve profitability by identifying top-performing customers, employees, and products.
- Build a structured database architecture that supports ongoing analytics and reporting.

2. Data Source and Relational Schema

2.1. Data Source

We used the *Northwind Traders* dataset from Kaggle (https://www.kaggle.com), a publicly available sample business database. It includes seven interrelated tables containing transactional, customer, and product data:

- **Categories**: Product category information.
- **Customers**: Company and contact details for clients.
- **Employees**: Staff responsible for managing orders.
- **OrderDetails**: Line-item information linking orders and products.
- **Orders**: Transactional order information.
- **Products**: Product-level information, including price and category.
- **Shippers**: Details about logistics partners.

2.2. Relational Schema

We implemented the data in PostgreSQL, ensuring referential integrity through foreign key relationships. The relational schema is designed as shown in **Figure 1**.
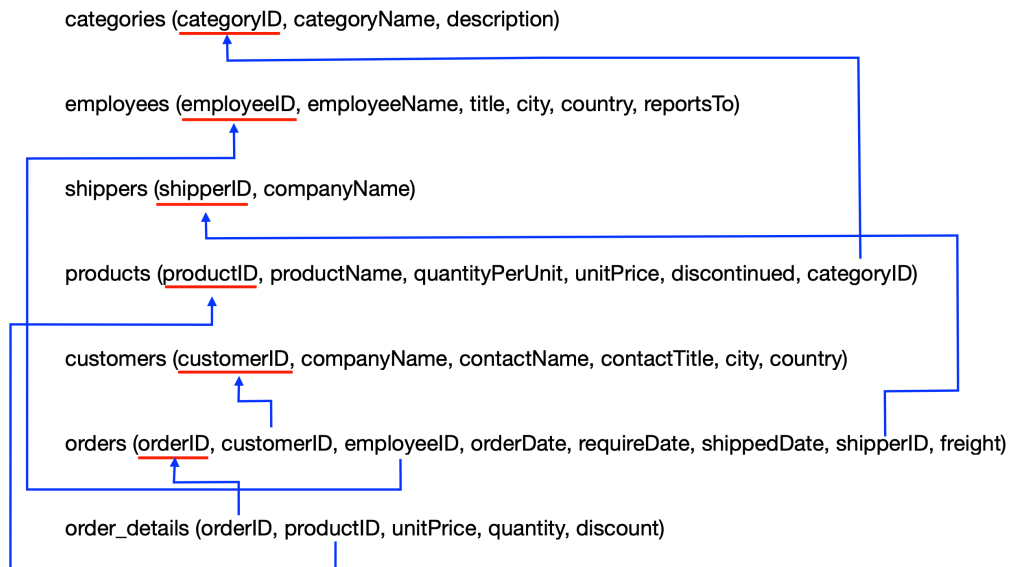

Figure 1. The relational schema of the *Northwind Traders* database. The primary keys (PKs) in each table are marked in red. The blue arrows illustrate the relationships between the foreign keys (FKs) in each table and the corresponding PKs in others.

This design ensures that queries can promptly access transactional insights across various tables. Based on the schema shown in the above figure, the Entity-Relationship Diagram can be derived as shown in **Figure 2**.
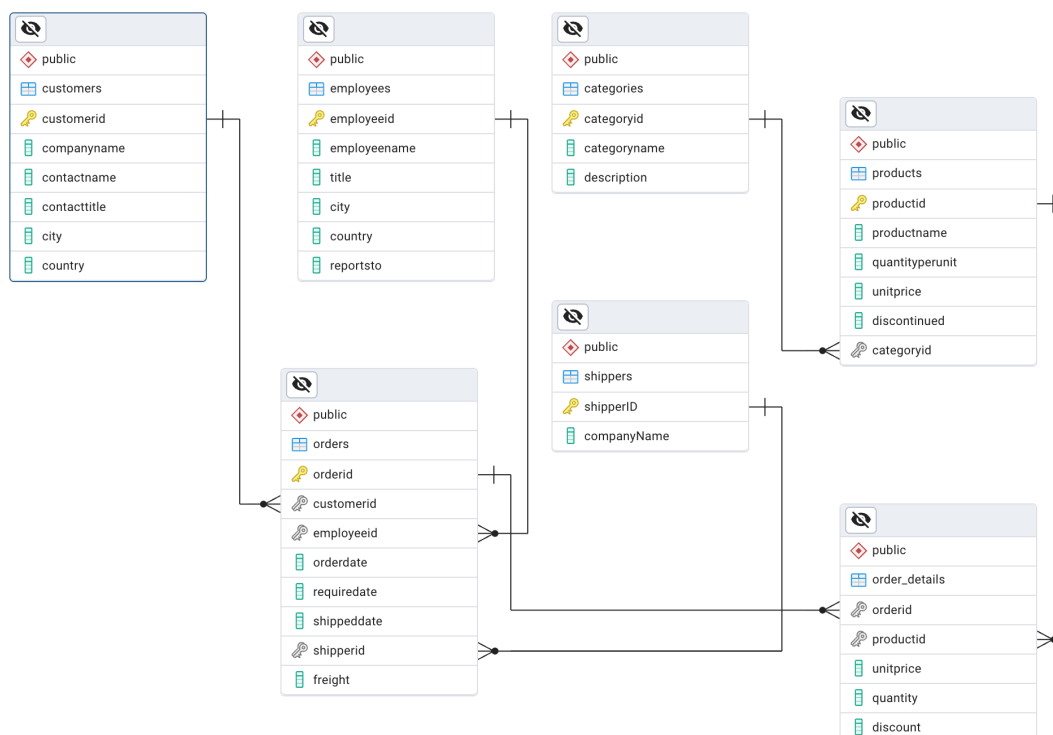

Figure 2. The Entity-Relationship (ER) Diagram of the *Northwind Traders* database.

2

3. SQL Methodology and Analytical Design

To address the client's business objectives, we adopted a question-driven analytical approach. Each query was designed to answer a practical business question, yielding insights that could inform strategic actions.

The analytical process involved three key steps:

- **Schema design and data validation**: Cleaning and normalizing tables to prevent redundancy and maintain referential integrity.
- **SQL query development**: Constructing efficient queries using "JOIN", "GROUP BY", "WHERE", and arithmetic functions to extract insights.
- **Interpretation and communication**: Translating numerical findings into business implications to support management decisions.

The ten SQL queries were categorized into three analytical themes:

- **Revenue performance analysis**
- **Customer and employee performance**
- **Supply chain efficiency**

4. Analytical Results and Insights

4.1. Revenue and Sales Trends

Query 1: *Annual revenue and order volume analysis*

By aggregating total sales and order counts by year (see **Figure 3**), we observe that *Northwind Traders* achieved the highest revenue in 2014, consistent with that year's prominent order volume.
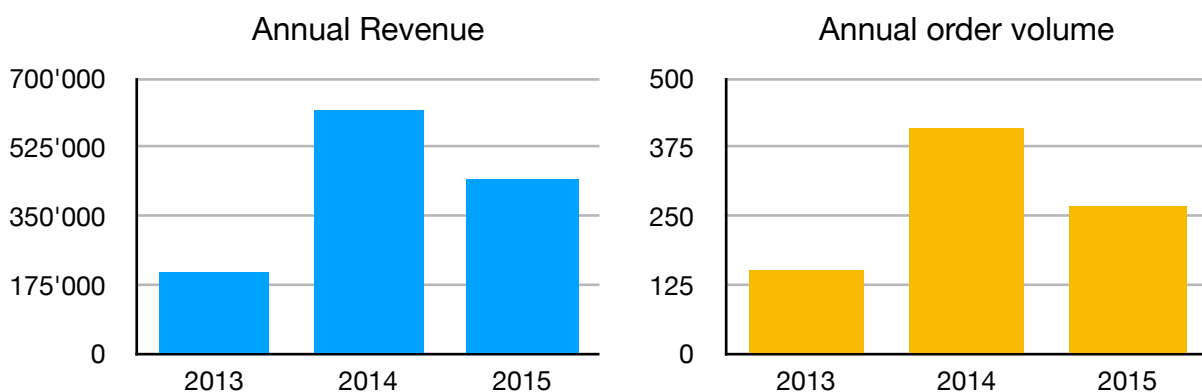
Figure 3. The annual revenue (left) and the order volume (right) of the Northwind Traders database.

Query 2: *Top-revenue products*

We initially calculated the total sales from 2013 to 2015, which is in total $1,265,793.11 (refer to Query 1). By joining the OrderDetails table with the Products table, we identified the top 10 items generating the most revenue, as shown in **Table 1**. It is evident that **over 20%** of the total revenue

Table 1. Top 10 products

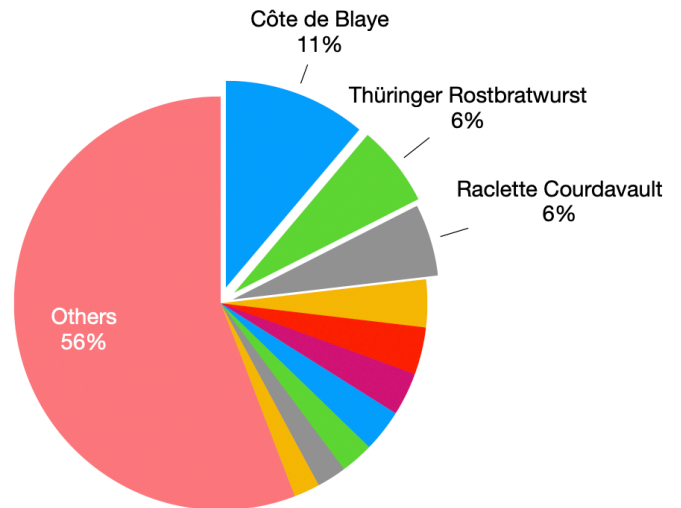| product ranking	product name	revenue
1	Côte de Blaye	141,396.74
2	Thüringer Rostbratwurst	80,368.67
3	Raclette Courdavault	71,155.7
4	Tarte au sucre	47,234.97
5	Camembert Pierrot	46,825.48
6	Gnocchi di nonna Alice	42,593.06
7	Manjimup Dried Apples	41,819.65
8	Alice Mutton	32,698.38
9	Carnarvon Tigers	29,171.88
10	Rössle Sauerkraut	25,696.64

Figure 4. Sales contribution of the top 3 best-selling products.

comes from the **top three** products (see **Figure 4**), indicating a high concentration of revenue that requires careful monitoring.

Query 3: *Top customers by revenue*

By combining *Orders*, *OrderDetails*, and *Customers*, we identified the top 10 customers (**Table 2**). We found that five corporate clients account for over one-third of total revenue (see **Figure 5**), highlighting potential dependency risks and opportunities for loyalty programs.

Table 2. Top 10 customers

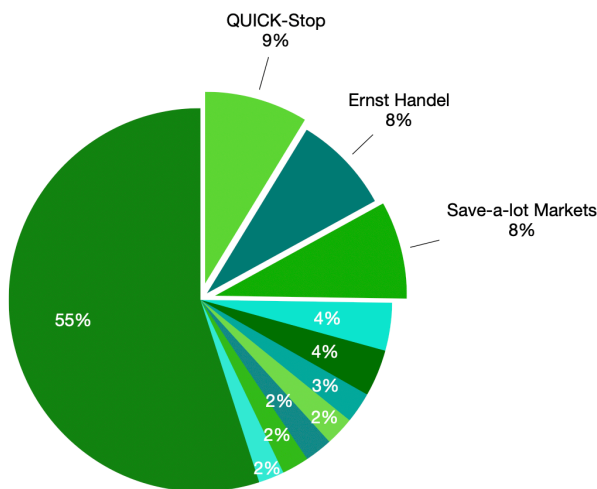| company_name	revenue	orders
QUICK-Stop	110277.32	28
Ernst Handel	104874.99	30
Save-a-lot Markets	104361.96	31
Rattlesnake Canyon Grocery	51097.80	18
Hungry Owl All-Night Grocers	49979.91	19
Hanari Carnes	32841.37	14
Königlich Essen	30908.39	14
Folk och fä HB	29567.57	19
Mère Paillarde	28872.19	13
White Clover Markets	27363.61	14

Figure 5. Sales contribution of the top 5 customers.

4.2. Employee and Customer Performance

Query 4: *Employee sales performance*

Joining Orders, OrderDetails, and Employees revealed significant variation among sales staff (see **Table 3**). The top 3 performers account for almost half of the revenue and orders, more than twice that of the bottom quartile of employees, suggesting room for targeted training and incentive alignment.

Table 3. Top performing employees

employee_name	revenue	revenue ratio%	No. of orders	Order%
Margaret Peacock	232890.87	18.40	156	18.80
Janet Leverling	202812.88	16.02	127	15.30
Nancy Davolio	192107.65	15.18	123	14.82
Andrew Fuller	166537.76	13.16	96	11.57
Laura Callahan	126862.29	10.02	104	12.53
Robert King	124568.24	9.84	72	8.67
Anne Dodsworth	77308.08	6.11	43	5.18
Michael Suyama	73913.15	5.84	67	8.07
Steven Buchanan	68792.30	5.43	42	5.06

Query 5: *Customer order frequency and basket size*

By counting the total number of orders and averaging the order values per customer, we identified distinct buyer segments (see **Table 4**). Some small-to-medium enterprises, such as *Hanari Carnes,*

Table 4. Customer order frequency and basket size

company_name	orders_count	avg_order_value	company_name	orders_count	avg_order_value
Save-a-lot Markets	31	3366.51	Bon app'	17	1291.96
Ernst Handel	30	3495.83	Frankenversand	15	1777.10
QUICK-Stop	28	3938.48	Lehmanns Marktstand	15	1284.09
Hungry Owl All-Night Grocers	19	2630.52	Wartian Herkku	15	1043.25
Folk och fä HB	19	1556.19	Hanari Carnes	14	2345.81
Rattlesnake Canyon Grocery	18	2838.77	Königlich Essen	14	2207.74
Berglunds snabbköp	18	1384.87	White Clover Markets	14	1954.54
HILARION-Abastos	18	1264.93			

Königlich Essen, and White Clover Markets, tend to place fewer but higher-value orders than most other buyers with fewer than 20 orders, implying potential for personalized account management.

4.3. Supply Chain and Shipping Insights

Query 6: *Average delivery time by shipper*

Northwinde Traders works with three shippers, and it is crucial to assess their performance to see if any are delivering services without substantial delays. Using date arithmetic functions on the Orders table, we measured the average delivery time and the number of orders delivered per shipping partner, as shown in **Table 5**. It is surprising that United Package, the partner handling the majority of shipments, is also responsible for most delayed deliveries, highlighting potential opportunities to renegotiate contracts or improve routing strategies.

Table 5. Shipper performance

shipper	avg_days_order_to_ship	On-time delivery %	shipped_orders
United Package	9.23	94.92	315
Speedy Express	8.57	95.10	245
Federal Shipping	7.47	96.39	249

Query 7: *Monthly revenue versus shipping cost trend*

To investigate the monthly revenue growth and shipping costs, we aggregated the monthly sales and shipping expenses, as shown in **Figure 6**. It is evident that *Northwind Traders* experienced steady but insignificant revenue growth from July 2013 to November 2014, with minor fluctuations. There was a surge in sales over the next few months, until April 2015, the reason for which was unclear. However, the freight costs are almost negligible, accounting for around 5%.

To further validate this phenomenon, we investigated the correlation between monthly revenue and freight costs. As shown in **Figure 7**, the two values exhibit a strong positive correlation, indicating that increases in freight costs effectively reflect increases in company income. Nonetheless, the delivery costs are not yet a significant obstacle to the company's growth at this stage.

Query 8: *Regional product preferences*

A significant determinant influencing Northwind's trading activities is the regional preferences for various products. Culinary distinctions, coupled with cultural heritage, significantly influence consumers' daily dietary choices. Consequently, suppliers often need to tailor their product offerings and marketing strategies to align with local consumers' preferences.

Cross-referencing Categories, Customers, and Orders revealed prominent markets and regional differences in demand (**Figure 8**). Northwind Traders' largest market is in the USA, followed by Germany. Nonetheless, 15 of the 21 countries are in Europe, generating more than three times the sales of the USA. Regarding regional demand variations, beverages dominate the entire sales market, especially in the USA, Germany, and Brazil. Dairy products represent the second most popular choice, mostly consumed in Europe. These insights can inform the development of tailored regional marketing strategies.

Figure 6. Monthly growth of revenue and freight costs (above) and the ratio of freight costs.
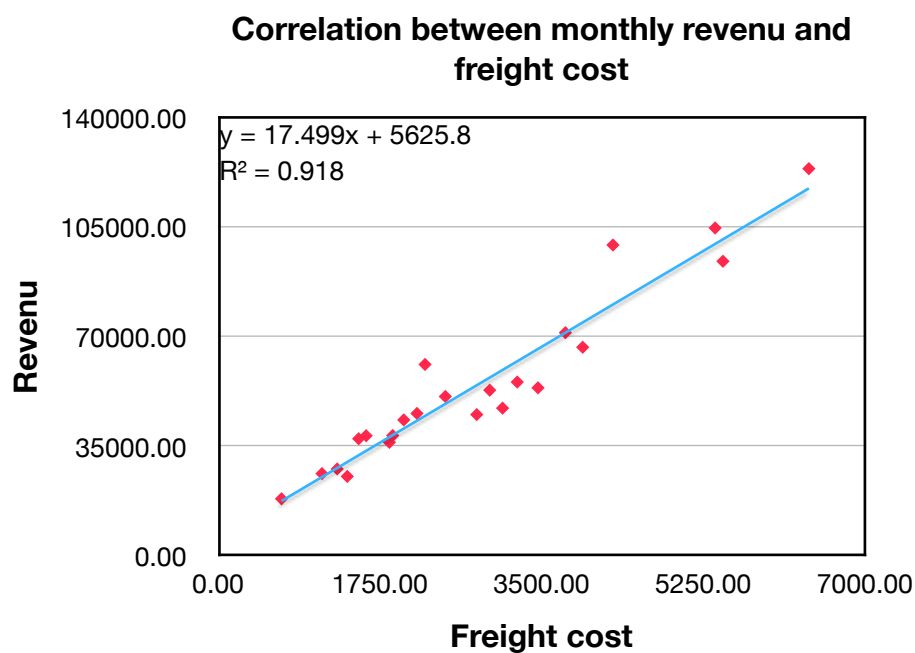

Figure 7. Correlation between the monthly revenue and freight costs. The linear regression line is highlighted in blue with its function and R^2 value shown in the upper left corner.
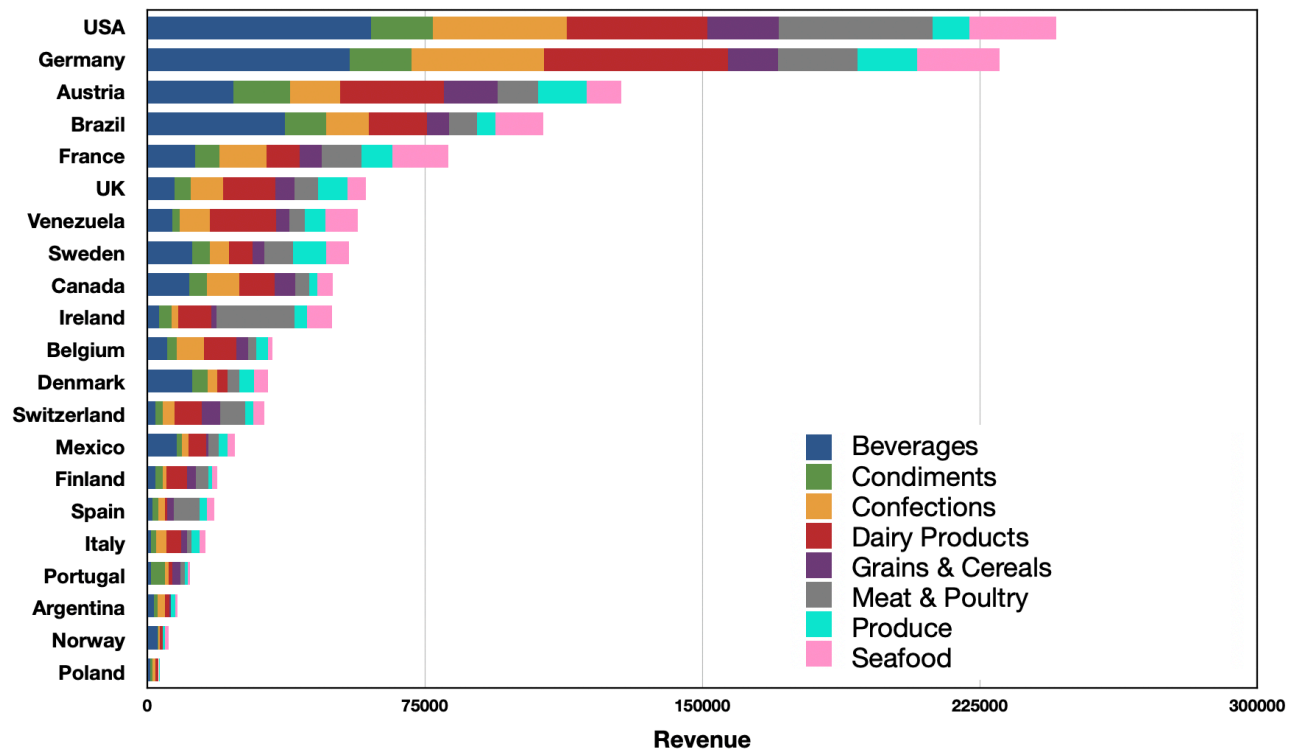
Figure 8. Revenue distribution of each product category across 22 countries.

4.4.Inventory Insights

Query 9: *Products with declining sales trends*

Using CTEs and simple slope calculations, we compared sales trends across consecutive years and identified products with negative sales slopes, representing declining sales trends (**Figure 9**). Products showing a significant decline, such as *Northwoods Cranberry Sauce*, may be overstocked and could benefit from discounts or discontinuation.

Query 10: *Products with increasing sales trends*

Likewise, products with a positive sales slope can be identified as having a growing sales trend (**Figure 9**). For instance, products such as *Raclette Courdavault* and *Thüringer Rostbratwurst* should be stocked in sufficient quantities to meet rising demand.
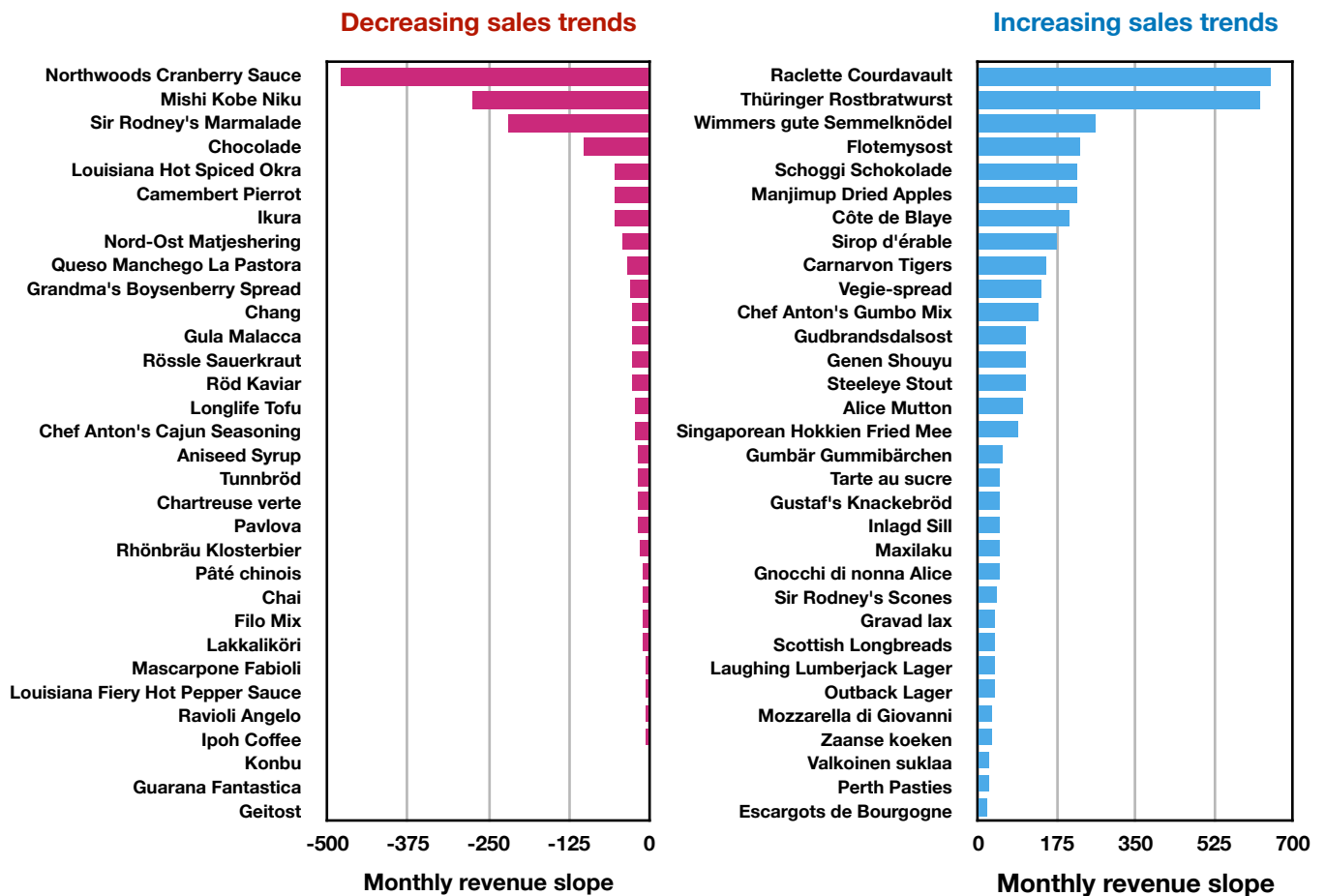
Figure 9. Monthly sales comparison with a declining trend (left) and an increasing trend (right).

5. Business implications

5.1. Short-term (1 ~ 2 months)

Improve shipper reliability

The analysis reveals that the main shipping partner handles most of the company's deliveries but also contributes significantly to delays. This indicates a need to evaluate and enhance their performance. Northwind Traders should renegotiate service-level agreements (SLAs), incorporate penalties or incentives for on-time delivery, and test alternative shipping providers for some orders to compare results. As Mentzer and Moon demonstrated (Mentzer and Moon 2004), optimizing logistics partners can greatly improve customer satisfaction and lower operational costs.

Reduce dependency on a few products and customers

Since a small group of products and customers generates most of the revenue, it's crucial to reduce this concentration risk. Diversify suppliers for top products, expand cross-selling and upselling, and develop loyalty or pricing plans for key customers. The aim is to make revenue more balanced and sustainable (Kotler and Keller 2016).

Review low-performing products

Products with a noticeable drop in sales should be reviewed immediately. The company can either cut back on future purchase orders or run targeted promotions to move slow-selling inventory. Meanwhile, warehouse space can be reallocated to faster-selling items to boost efficiency.

5.2.   Mid-term (2 ~ 6 months)

Introduce automated reporting and dashboards
Instead of running one-time SQL queries, it would be more effective to develop automated and scheduled reports using a business dashboard. This approach saves time, minimizes manual errors, and provides managers with real-time insights. As Davenport notes (Davenport 2018), data visualisation and automation are essential for turning analytics into daily decision-making tools.

Implement sales team performance
The data indicate that a small group of employees generates most of the revenue. This imbalance can be fixed by discovering what top performers do differently and sharing those strategies with the team. Coaching, clearer performance objectives, and better-aligned incentives can help improve overall sales results.

Optimize freight-cost strategy
Even if freight costs make up a relatively small part of total revenue (about 5%), they still strongly influence profitability. The company could analyze these costs by order type and shipment class to find savings opportunities.

5.3. Long-term (6 ~ 12+ months)

Adopt demand forecasting & inventory optimization
Develop a forecasting module, using either time-series analysis or machine learning, that leverages historical monthly sales data and seasonality patterns to determine reorder points and safety stock levels. Incorporating forecast results into procurement processes can help minimise stockout and overstock situations (Choi et al. 2018).

Supplier rationalisation and strategic partnerships
Evaluating each supplier's performance, reliability, and lead time is especially important for key products. Implementing dual sourcing for critical products and establishing long-term agreements with top-performing suppliers can help mitigate risks and strengthen supply chain resilience.

Customer segmentation & margin management
Use metrics like total revenue, order frequency, and average basket size to categorise customers into segments (e.g., frequent high-value, occasional high-value, and low-volume buyers). Customise retention offers, discounts, and communication strategies for each group. Monitoring customer lifetime value can help determine which segments warrant more investment and attention (Gupta and Lehmann 2006).

6.   Conclusion

The SQL-based redesign and analysis conducted for Northwind Traders demonstrate how a well-structured database and targeted queries can transform large quantities of raw data into valuable business insights. By migrating the dataset into a clean, normalized PostgreSQL architecture and employing more advanced SQL techniques, the project uncovered several significant findings. One of the primary outcomes is a pronounced concentration of revenue among a limited number of

products and customers, which could render the company's income more susceptible to fluctuations in demand or customer loyalty. The analysis further reveals notable disparities in sales performance among employees, indicating the necessity for equitable targets and more tailored training or incentive programs.

In examining shipping operations, the data indicate discernible differences in both reliability and cost efficiency between shippers. This presents an opportunity to renegotiate existing contracts or collaborate with new partners to enhance delivery timeliness and manage freight expenses more effectively. Concurrently, the review of sales trends by product highlights steadily growing products and those experiencing a decline in demand, thereby informing inventory management, procurement strategies, and marketing initiatives.

In summary, these findings affirm the project's primary objective: integrating data into a consistent, well-designed database structure facilitates easier access, analysis, and informed decision-making. This methodology establishes a robust foundation for ongoing enhancements, supports improved business performance, and contributes to the organization's resilience and competitiveness across its supply chain and overall operations.

7. *References*

Chopra, S., & Meindl, P. *Supply Chain Management: Strategy, Planning, and Operation*. Pearson Education, 2019.

Choi, T. M., Wallace, S. W., and Wang, Y. *Big Data Analytics in Operations Management*. Springer, 2018.

Davenport, T. H. *Analytics at Work: Smarter Decisions, Better Results*. Harvard Business Press, 2018.

Gupta, S., and Lehmann, D. R. "*Customer Lifetime Value: The Path to Profitability,*" Journal of Service Research 9, no. 2 (2006): 91–112.

Kotler, P., and Keller, K. L. *Marketing Management* (15th ed.). Pearson, 2016.

Mentzer, J. T., and Moon, M. A. S*ales Forecasting Management: A Demand Management Approach*. SAGE Publications, 2004

Appendix

Database: NorthwindTradesProject_dump file (refer to the guidelines in the .sql file for database)

SQL File: northwind_analysis.sql (includes all 10 validated queries)

Team Roles:
- *Xiaokang Li:* schema design, data integration, query validation, data visualisation, and report drafting.
- *Anass Saissi Hassani*: SQL query construction and performance validation.
- *Ettore Bertolini***:** query validation and business insights interpretation.