Project Background and work plan

You are working as a *data analyst* for **Northwind**: A wholesale distributor that supplies a wide range of food and beverage products to retail businesses, restaurants, and other companies around the world. The company operates through a digital platform where customers can browse products, place orders, and manage their accounts. Northwind's mission is to provide quality products efficiently and to build strong relationships with its clients, ensuring they receive what they need when they need it.

Your role is to analyze data from Northwind's digital store to generate insights that could **help the company improve its profitability**. Imagine you're preparing this analysis for Northwind's board, who will use your findings to make strategic decisions.

Contents

Defining Your 3 Research Questions

Q1: Product Portfolio & Pricing Strategy

- **The Goal:** Identify which products are "star performers" and which are "profit drains."
- **The Question:** Which product categories have the highest profit margins versus those with the highest discount rates, and how is discounting affecting the bottom line?
- **Actionable Insight:** Adjust pricing for low-margin items or reduce aggressive discounting on high-demand products.

Q2: Customer Segmentation (RFM Lite)

- **The Goal:** Identify the most valuable customers and those at risk of churning.
- **The Question:** Who are the top 20% of customers by total revenue, and are there significant geographic regions where sales are declining?
- **Actionable Insight:** Targeted loyalty programs for "Whale" customers and re-engagement campaigns for declining regions.

Q3: Logistics & Employee Performance

- **The Goal:** Find bottlenecks in fulfillment and identify top sales talent.
- **The Question:** What is the average lead time (order to ship date) per shipper, and which employees are generating the highest average order value?
- **Actionable Insight:** Re-negotiate contracts with slow shippers or replicate the sales techniques of top-performing employees across the team.

Data Exploration & SQL Querying

- **Task 1:** Inspect the Schema. Map how Orders connects to Order Details (via OrderID) and Products (via ProductID).
- **Task 2: SQL Extraction.** Write queries to calculate:
- **Gross Profit:** (UnitPrice * Quantity * (1 - Discount)) - (SupplierCost * Quantity). *Note: If Supplier Cost isn't explicitly in a column, use the Products table UnitPrice as a proxy for cost and Order Details UnitPrice for the sale.*
- **Shipping Delays:** ShippedDate - OrderDate.
- **Task 3:** Export results to CSV/Excel for Tableau.

Dashboard Development (Tableau)

- **Task 4:** Create a "General Metrics" landing page (Total Revenue, Total Orders, Average Discount).
- **Task 5:** Build 3 specific worksheets for your Research Questions (e.g., a Heatmap for regions, a Scatter plot for Price vs. Quantity, and a Bar chart for Employee sales).

- **Task 6:** Assemble the Dashboard using "Actions" so clicking a region filters the product list.

Presentation & Insights

- **Task 7:** Draft the slides. Don't just show charts; explain **why** the data looks that way.
- **Task 8:** Formulate the "So What?" (e.g., "Our shipping delay in Brazil is 15% higher than average; we should switch to Shipper B for that region").

Final Polish & Appendix

- **Task 9:** Organize the Appendix. Clean up your SQL code with comments.
- **Task 10:** Record a practice run of your 7-10 minute presentation to ensure timing.

Task 1 – Schema of the database.

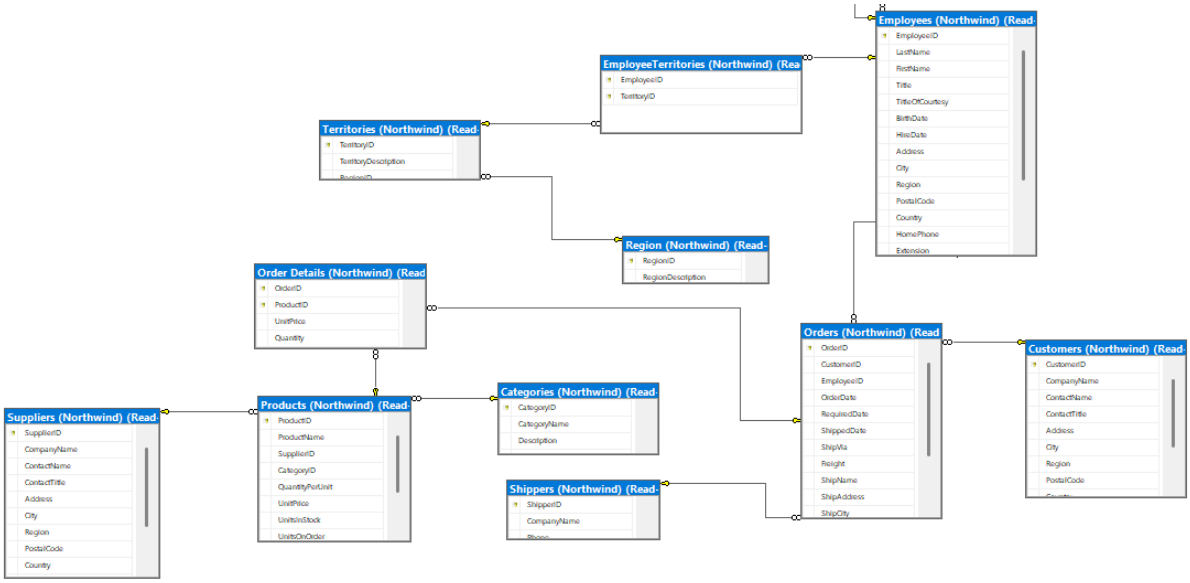

Figure 1 - Schema of whole Northwind dataset

Task 2 – Q1: The "Profit vs. Discount" Breakdown.

SQL query 1:

```
SELECT  c.CategoryName,

    -- Total value if everything was sold at full price

    ROUND(SUM(od.UnitPrice * od.Quantity), 2) AS Potential_Revenue,

    -- The actual money lost to discounts

    ROUND(SUM(od.UnitPrice * od.Quantity * od.Discount), 2) AS
Total_Discount_Value,

    -- The actual money that hit the bank

    ROUND(SUM(od.UnitPrice * od.Quantity * (1 - od.Discount)), 2) AS
Actual_Revenue,

    -- Estimated Profit (Actual Revenue minus Wholesale Cost)

    ROUND(SUM((od.UnitPrice * od.Quantity * (1 - od.Discount)) - (p.UnitPrice *
od.Quantity)), 2) AS Estimated_Gross_Profit,

    -- Average Discount Rate per Category

    ROUND(AVG(od.Discount) * 100, 2) AS Avg_Discount_Percent

FROM [Northwind].[Categories] c

JOIN [Northwind].[Products] p ON c.CategoryID = p.CategoryID

JOIN [Northwind].[Order Details] od ON p.ProductID = od.ProductID

GROUP BY c.CategoryName

ORDER BY Estimated_Gross_Profit DESC;
```

Why are we doing this? (The Rationale)

In a wholesale business like Northwind, volume does not always equal value.

- **The Problem:** Sales teams often use heavy discounts to close deals. If they discount a product that already has a low margin (the difference between the Products price and Order Details price), we might be losing money on every sale.
- **The Goal:** We are looking for the "Discount Sweet Spot"—the point where discounts drive enough volume to increase profit without eroding the bottom line.

Results

Table 1 - Profit vs. Discount

Category Name	Potential Revenue	Total Discount Value	Actual Revenue	Estimated Gross Profit	Avg Discount Percent
Beverages	$ 286,526.95	$ 18,658.77	$ 267,868.18	$ 113,077.06	6.19%
Dairy Products	$ 251,330.50	$ 16,823.22	$ 234,507.29	$ 99,943.14	5.34%
Confections	$ 177,099.10	$ 9,741.88	$ 167,357.22	$ 72,192.95	5.69%
Meat/Poultry	$ 178,188.80	$ 15,166.44	$ 163,022.36	$ 67,681.01	**6.45%**
Seafood	$ 141,623.09	$ 10,361.35	$ 131,261.74	$ 56,731.97	6.02%
Condiments	$ 113,694.75	$ 7,647.67	$ 106,047.08	$ 44,875.58	5.26%
Produce	$ 105,268.60	$ 5,284.02	$ 99,984.58	$ 44,287.08	4.54%
Grains/Cereals	$ 100,726.80	$ 4,982.21	$ 95,744.59	$ 42,320.59	4.53%

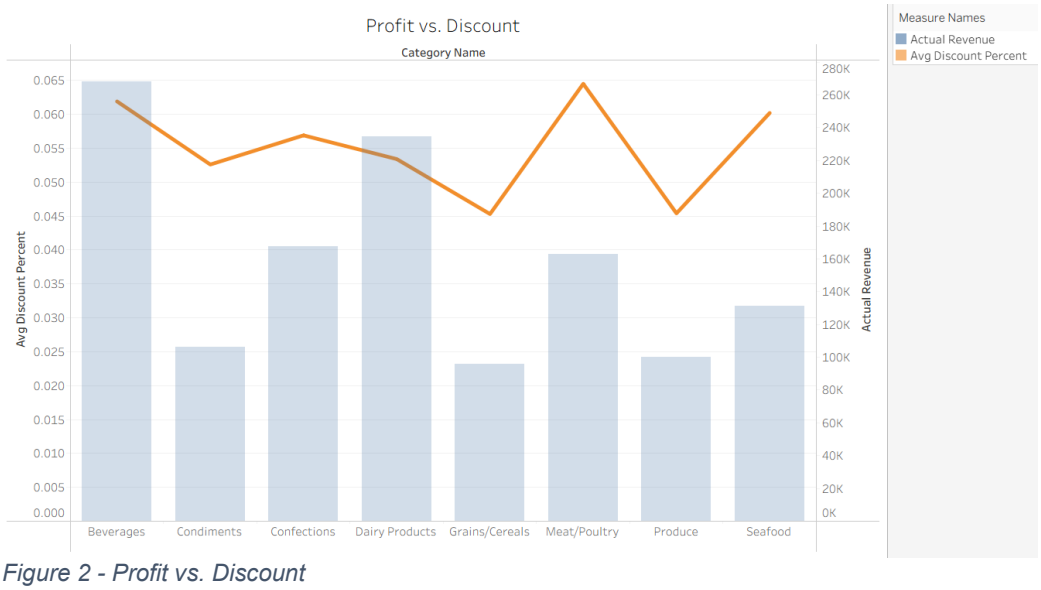

Figure 2 - Profit vs. Discount

Insights -
We are looking for **Efficiency**.

- **Beverages** has a high discount (6.19%), but it also has the highest profit. This suggests that the discounts are successfully driving high-volume sales that more than make up for the price cut.
- **Meat/Poultry**, however, has the **highest discount rate (6.45%)** but significantly lower profit than Dairy or Beverages. This suggests we are discounting expensive items too heavily without seeing a massive return in total profit.

Insight A: The "Discount Trap" in Meat/Poultry

Meat/Poultry has the highest average discount rate (6.45%) among all categories. Despite this, its profit is nearly **$45,000 less** than Beverages. We are potentially "giving away" profit on high-value items where customers might be willing to pay closer to the list price.

Insight B: The "Untapped Potential" in Produce & Grains

Produce and Grains have the **lowest discount rates** (approx. 4.5%). These categories are very healthy in terms of margin-to-revenue ratio. This indicates these products are "necessities" that customers buy even without big price cuts.

Actionable Recommendations for the Board

- **Recommendation 1 (Protect the Margin):** Implement a "Discount Cap" on the **Meat/Poultry** category. Reducing the average discount from 6.45% to 5.5% could recover thousands of dollars in bottom-line profit without significantly impacting volume.
- **Recommendation 2 (Drive Volume):** Since **Produce** has a low discount rate and solid margins, Northwind should trial a "Bundle Discount" (e.g., buy Grains, get 10% off Produce). This uses high-margin items to increase the "Total Basket Size" of the customer.
- **Recommendation 3 (Star Performer):** Continue the aggressive strategy for **Beverages**. The data shows this category handles the 6.19% discount well, maintaining its spot as the #1 profit generator.

A more drill down to Meat/Poultry –
The Theory: Within the "Meat/Poultry" category, there might be one specific product that is responsible for most of the "Discount Trap." Conversely, in "Produce," there might be one "Super-Star" product that deserves its own marketing budget.
The Objective: Identify the specific Top 5 and Bottom 5 products by **Total Profit Contribution**.

SQL query 2

```
SELECT TOP 10
    p.ProductName, c.CategoryName,
    SUM(od.Quantity) AS Total_Units_Sold,
    ROUND(SUM((od.UnitPrice * od.Quantity * (1 - od.Discount)) - (p.UnitPrice * od.Quantity)), 2) AS Product_Gross_Profit,
    ROUND(AVG(od.Discount) * 100, 2) AS Avg_Discount
FROM [Northwind].[Products] p
JOIN [Northwind].[Categories] c ON p.CategoryID = c.CategoryID
JOIN [Northwind].[Order Details] od ON p.ProductID = od.ProductID
WHERE c.CategoryName = 'Meat/Poultry' -- Use single quotes for text
GROUP BY p.ProductName, c.CategoryName
ORDER BY Product_Gross_Profit ASC;
```

Why are we doing this? (The Rationale)

The Theory: Within the "Meat/Poultry" category, there might be one specific product that is responsible for most of the "Discount Trap." Conversely, in "Produce," there might be one "Super-Star" product that deserves its own marketing budget.

The Objective: Identify the specific Top 5 and Bottom 5 products by **Total Profit Contribution**.

Results

Table 2 - Drill down on Meat-Poultry

Product Name	Category Name	Total Units Sold	Product Gross Profit	Avg Discount
Tourtiᴜre	Meat/Poultry	755	$ 1,915.86	7.22
Mishi Kobe Niku	Meat/Poultry	95	$ 2,619.00	10
Pɑt ᑫchinois	Meat/Poultry	903	$ 6,590.40	7.73
Perth Pasties	Meat/Poultry	722	$ 8,733.37	3.83
Alice Mutton	Meat/Poultry	978	$ 13,627.38	6.22
Th□ringer Rostbratwurst	Meat/Poultry	746	$ 34,195.00	6.41

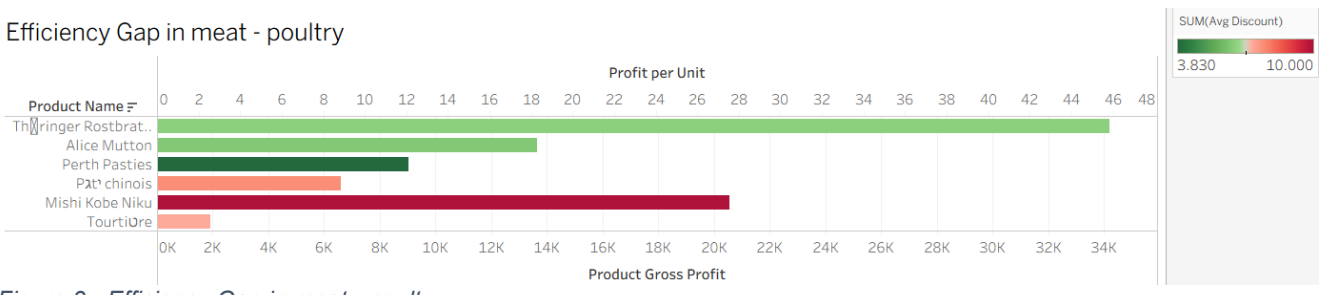

Figure 3 - Efficiency Gap in meat - poultry

Insights –

1. The "Star Performer": Thüringer Rostbratwurst

- **Performance**: This is your absolute powerhouse, generating over $34K in profit with a healthy Profit per Unit of nearly $46.
- **Discount Resilience**: Even with a moderate discount, it remains highly profitable. It is the "Anchor" of the meat department.

2. The "Luxury Leak": Mishi Kobe Niku

- **The Red Flag**: This product is highlighted in deep red, indicating the highest average discount at 10%.
- **The Contradiction**: Despite being a high-value item, it produces significantly less total profit ($20K) than the Rostbratwurst. We are effectively "paying" people to take our most expensive inventory.

3. The "Efficiency Voids": Tourtière & Pâté chinois

- **The Struggle**: Both products show very short bars (low gross profit) and are colored in light red/orange (high discounts).
- **Operational Strain**: These items require the same warehouse space and shipping effort as the Rostbratwurst but return only a fraction of the profit (Tourtière is particularly low at roughly $2 per unit).

Actionable Recommendations for the Board

Protect the Anchor

- **Strategy**: Maintain current pricing for **Thüringer Rostbratwurst**. It is currently optimized. Use it as a "lead" item in marketing to bring customers in without needing to increase its discount.

Stop the "Luxury" Discount

- **Strategy**: Immediate "Discount Freeze" on **Mishi Kobe Niku**.
- **Expected Result**: By reducing the discount from 10% to 5% or 0%, the company can significantly recover profit margin on every single unit sold, as this premium product likely has lower price sensitivity.

Inventory Re-allocation

- **Strategy**: Review the viability of **Tourtière**.
- **Action**: If profit remains at $2 per unit, consider replacing its warehouse footprint with higher-margin items from the **Beverages** or **Dairy** categories identified in our Task 1 macro-analysis.

Task 2 – Q2: The " Customer Segmentation (RFM Lite)" Breakdown

SQL query 1

```
SELECT
    c.CustomerID, c.CompanyName, c.Country, c.City,
    -- Count of unique orders to see loyalty
    COUNT(DISTINCT o.OrderID) AS Order_Frequency,
    -- Total Net Revenue (LTV)
    ROUND(SUM(od.UnitPrice * od.Quantity * (1 - od.Discount)), 2) AS
Lifetime_Value,
    -- Average Discount given to this customer
    ROUND(AVG(od.Discount) * 100, 2) AS Avg_Customer_Discount
FROM [Northwind].[Customers] c
JOIN [Northwind].[Orders] o ON c.CustomerID = o.CustomerID
JOIN [Northwind].[Order Details] od ON o.OrderID = od.OrderID
GROUP BY c.CustomerID, c.CompanyName, c.Country, c.City
ORDER BY Lifetime_Value DESC;
```

SQL query 2

```
SELECT
    Country,
    -- Creates a string like '1997-Q1'
    CONCAT(CAST(YEAR(OrderDate) AS VARCHAR), '-Q', CAST(DATEPART(qq,
OrderDate) AS VARCHAR)) AS Year_Quarter,
    SUM(od.UnitPrice * od.Quantity * (1 - od.Discount)) AS Quarterly_Revenue
FROM [Northwind].[Orders] o
JOIN [Northwind].[Order Details] od ON o.OrderID = od.OrderID
JOIN [Northwind].[Customers] c ON o.CustomerID = c.CustomerID
GROUP BY Country, YEAR(OrderDate), DATEPART(qq, OrderDate)
ORDER BY Year_Quarter ASC;
```

Why are we doing this? (The Rationale)

- **The Pareto Principle:** We want to confirm if a small group of customers (the "Whales") is responsible for the majority of Northwind's profits.
- **Geographic Strategic Planning:** Understanding revenue distribution by country allows the board to decide where to invest in localized distribution hubs or specialized sales teams.
- **Churn Prevention:** By identifying top customers, we can monitor their ordering patterns to prevent them from leaving (churning) for a competitor.

Results

Table 3 – Top 10 Customer Value & Geographic Strongholds

Customer ID	Company Name	Country	City	Order Frequency	Lifetime Value	Avg Customer Discount
QUICK	QUICK-Stop	Germany	Cunewalde	28	$ 110,277.30	6.92
ERNSH	Ernst Handel	Austria	Graz	30	$ 104,874.98	6.81
SAVEA	Save-a-lot Markets	USA	Boise	31	$ 104,361.95	8.28
RATTC	Rattlesnake Canyon Grocery	USA	Albuquerque	18	$ 51,097.80	3.51
HUNGO	Hungry Owl All-Night Grocers	Ireland	Cork	19	$ 49,979.91	11.36
HANAR	Hanari Carnes	Brazil	Rio de Janeiro	14	$ 32,841.37	6.72
KOENE	Kxniglich Essen	Germany	Brandenburg	14	$ 30,908.38	3.97
MEREP	Mere Paillarde	Canada	Montrıal	13	$ 28,872.19	5.78
WHITC	White Clover Markets	USA	Seattle	14	$ 27,363.61	6.88
FRANK	Frankenversand	Germany	Minchen	15	$ 26,656.56	6.56
QUEEN	Queen Cozinha	Brazil	Sao Paulo	13	$ 25,717.50	9

Table 4 – Top 2 countries by alphabetical order - Declining regions

Country	Year_Quarter	Quarterly_Revenue
Argentina	1997-Q1	$ 762.60
Argentina	1997-Q2	$ 335.50
Argentina	1997-Q4	$ 718.50
Argentina	1998-Q1	$ 5,921.50
Argentina	1998-Q2	$ 381.00
Austria	1996-Q3	$ 3,488.68
Austria	1996-Q4	$ 22,112.66
Austria	1997-Q1	$ 11,349.92
Austria	1997-Q2	$ 13,354.04
Austria	1997-Q3	$ 13,894.92
Austria	1997-Q4	$ 18,802.97
Austria	1998-Q1	$ 21,007.90
Austria	1998-Q2	$ 23,992.75

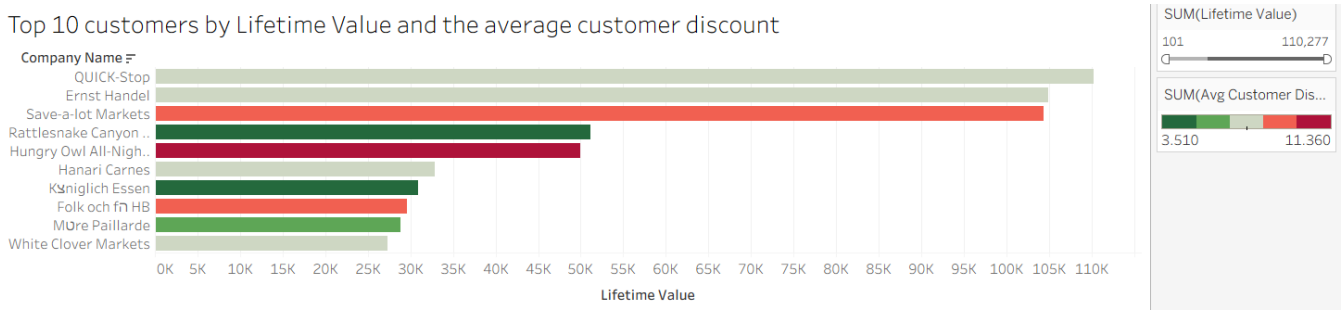

Figure 4 - Top 10 customers by life value and theirs average discount

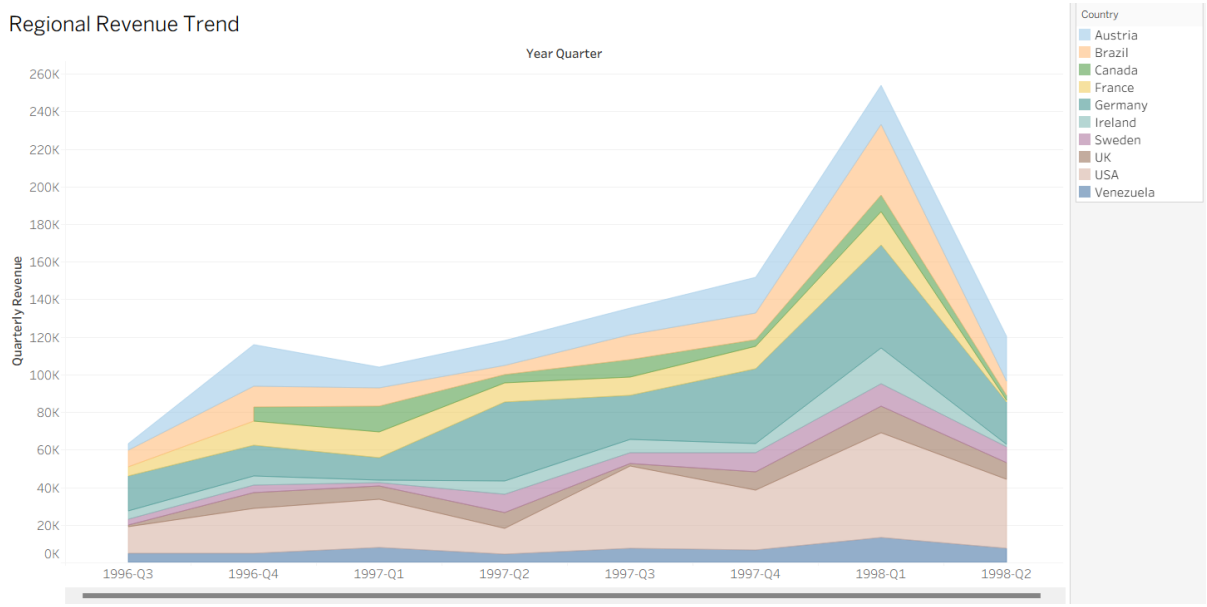

Figure 5 - Regional Revenue Trend

Figure 6 - Geographic map of the most valuable city and discount

Insights –

Insight A: **Insight**: **Save-a-lot Markets** (USA) generates nearly the same revenue as the top German customer but requires a much higher discount (8.28% vs 6.92%) to do so.

Insight B: **The "Discount Trap"**: **Hungry Owl All-Night Grocers** (Ireland) has an extremely high average discount of **11.36%**. We should investigate if this high discount is necessary to keep their business or if it's unnecessarily eroding profit.

Insight C: Based on the data, Northwind's revenue is anchored in three primary regions:

- **Western Europe**: Germany and Austria host two of the top three "Whales". Germany, in particular, shows multiple high-value customers (QUICK, KOENE, FRANK).
- **North America**: The USA (Boise, Albuquerque, Seattle) and Canada (Montreal) represent a massive, high-frequency cluster.
- **South America**: Brazil (Rio de Janeiro and Sao Paulo) is a emerging stronghold with high-volume customers like Hanari Carnes.\

Insight D: **The "1998-Q1" Growth Spike**

- **The Observation:** There is a massive, company-wide revenue peak in the first quarter of 1998 across almost every major region.
- **The Insight:** This indicates either a highly successful global marketing campaign or a major expansion in the product catalog during that period.
- **Actionable Recommendation :** You should investigate what happened in **1998-Q1** and try to replicate that "success formula" for the upcoming quarters.

Insight E: **The Q2 Downturn Warning**

- **The Observation:** After the 1998-Q1 peak, there is a sharp downward trend heading into **1998-Q2**.
- **The Insight:** This is your **"Churn Alert."** Because revenue is dropping across multiple regions simultaneously, this isn't just a local issue—it's a systemic slowdown.
- **Actionable Recommendation:** This is the perfect time to launch the **"VIP Outreach Program"** we discussed for your Top 10 Whales to stabilize the decline before the next quarter ends.

Task 2 – Q3: Fulfillment & Sales Talent

SQL query 1:

```
SELECT
    e.FirstName + ' ' + e.LastName AS Employee_Name, s.CompanyName AS
Shipper_Name,
    COUNT(o.OrderID) AS Total_Orders,
    -- Fulfillment Lead Time
    AVG(DATEDIFF(day, o.OrderDate, o.ShippedDate)) AS Avg_Days_to_Ship,
    -- Sales Effectiveness (Average Order Value)
    ROUND(SUM(od.UnitPrice * od.Quantity * (1 - od.Discount)) / COUNT(DISTINCT
o.OrderID), 2) AS Avg_Order_Value,
    -- Total Revenue Contribution
    ROUND(SUM(od.UnitPrice * od.Quantity * (1 - od.Discount)), 2) AS
Total_Revenue
FROM [Northwind].[Orders] o
JOIN [Northwind].[Employees] e ON o.EmployeeID = e.EmployeeID
JOIN [Northwind].[Shippers] s ON o.ShipVia = s.ShipperID
JOIN [Northwind].[Order Details] od ON o.OrderID = od.OrderID
WHERE o.ShippedDate IS NOT NULL
GROUP BY e.FirstName, e.LastName, s.CompanyName
ORDER BY Avg_Days_to_Ship DESC;
```

Why are we doing this? (The Rationale)

The Rationale: Why Lead Time & AOV Matter?

- **The Fulfillment Bottleneck:** In wholesale, speed is a competitive advantage. If "United Package" is consistently slower than "Speedy Express," it affects our "Whales" in the USA and Germany.
- **Sales Talent (AOV):** Total Revenue can be misleading if an employee just happens to handle more orders. Average Order Value (AOV) reveals who is better at up-selling and cross-selling high-margin items (like the Beverages we found in Task 1).

Results

Table 5 - Top 10 Fulfillment and Sales Talent.

Employee Name	Shipper Name	Total Orders	Avg Days to Ship	Avg Order Value	Total Revenue
Anne Dodsworth	United Package	45	13	$ 1,941.86	$ 36,895.31
Robert King	United Package	62	10	$ 2,255.07	$ 51,866.69
Janet Leverling	United Package	114	9	$ 1,955.55	$ 87,999.65
Janet Leverling	Speedy Express	97	9	$ 1,401.55	$ 50,455.87
Anne Dodsworth	Speedy Express	27	9	$ 1,769.46	$ 17,694.56
Laura Callahan	Speedy Express	65	9	$ 1,338.74	$ 33,468.58
Andrew Fuller	Speedy Express	83	9	$ 1,813.82	$ 61,669.89
Margaret Peacock	United Package	174	9	$ 1,495.90	$ 100,225.38
Michael Suyama	United Package	63	9	$ 1,053.55	$ 25,285.12
Laura Callahan	Federal Shipping	71	9	$ 1,225.15	$ 35,529.22

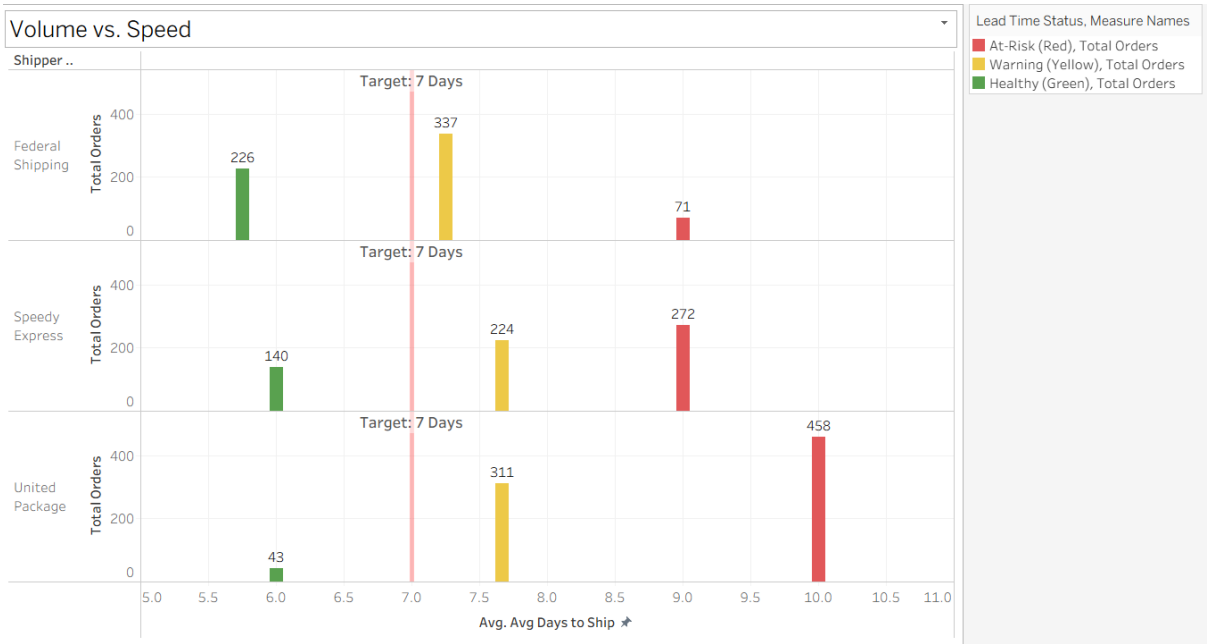

Figure 7 - Volume vs. Speed

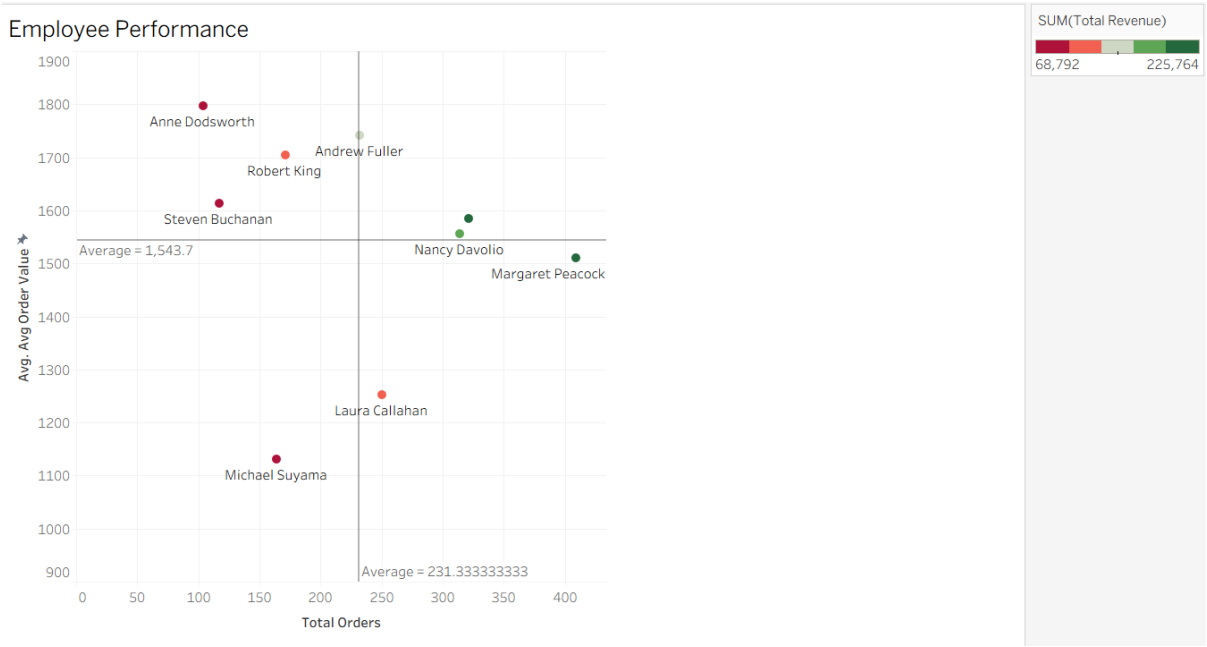

Figure 8 - Employee Performance

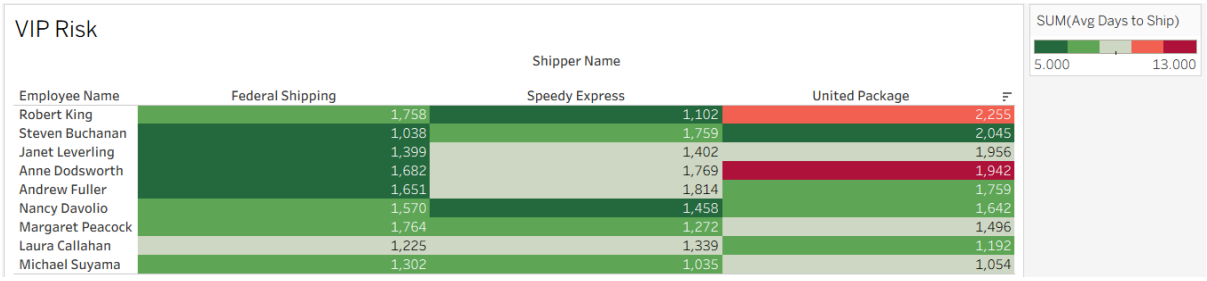

Figure 9 - VIP Risk

Insights – from figure 7:
Insight A: The "United Package" Operational Crisis: United Package has a massive pillar of 458 orders sitting at a 10-day lead time.

Insight B: The "Federal Shipping" Efficiency Model: Federal Shipping successfully delivers its largest volume (337 orders) in roughly 7.25 days, with another 226 orders being delivered in a "Healthy" 5.75 days.

From figure 8:
Insight C: The "Revenue Stars" (Top-Right) - High Volume + High Average Order Value (AOV): **Janet Leverling** and **Nancy Davolio** are your most efficient performers. They don't just process a high volume of orders; they ensure each order is significant in size.

Insight D: The "Premium Closers" (Top-Left) - Low Volume + High Average Order Value (AOV)**: Anne Dodsworth** and **Robert King** are your specialized "deal makers." While they handle fewer total orders, the orders they do close are among the most valuable in the company (exceeding **$1,800+** AOV).

Insight E: The "Operational Engines" (Bottom-Right) - *High Volume + Lower Average Order Value (AOV)***: Laura Callahan** is a "Workhorse." She processes a high number of orders (over 250), but her average order value is significantly lower than

the "Stars." This suggests she may be focusing on "quick wins" or smaller, commodity-style items.

Insight F: The "Developing Talent" (Bottom-Left) - *Low Volume + Lower Average Order Value (AOV)*: **Michael Suyama** is currently underperforming relative to the team average in both quantity and quality.

From figure 9:
Insight G: The "Premium" Bottleneck - Robert King and Anne Dodsworth represent the company's highest-ticket sellers, with average order values (AOV) of $2,255 and $1,942 respectively when using United Package. However, these premium transactions are being met with the company's worst delays (10–13 days), indicated by the deep red cells.

Insight H: Fulfillment Reliability Gap - Federal Shipping is consistently the most reliable partner across almost all employees, maintaining a "Healthy" green status with ship times as low as 5–6 days.

Insight I: Revenue Jeopardy - High-value orders are currently being funneled through the slowest shipping lane. Specifically, the $2,255 AOV deals closed by Robert King are taking 10 days to ship via United Package, whereas a similar high-value deal by Andrew Fuller ($1,651 AOV) via Federal Shipping takes only 6 days.

Actionable Recommendations for the Board

1. This is your single biggest operational bottleneck. While other shippers have "At-Risk" orders, United Package's largest volume of work is also its slowest. This suggests their core process is failing, not just a few outlier orders.

2. This shipper is your most reliable partner. They have proven they can handle high volumes while staying close to or under the 7-day target.

3. Benchmark & Replicate - Use Margaret's sales scripts and workflow management as the "Northwind Gold Standard." She is successfully balancing workload without sacrificing order quality.

4. Luxury Lead Assignment. Assign these employees to your highest-margin "Star" products (like *Thüringer Rostbratwurst*) and your "Whale" customers. Their talent lies in high-value negotiation rather than high-speed processing.

5. Upselling Training. Laura has the stamina for high volume; she now needs training on Cross-Selling. Transition her from just taking orders to bundling high-margin items (e.g., "Would you like to add Beverages to this Meat order?").

6. Mentorship Pipeline. Pair Michael with a "Star" like Margaret Peacock for two weeks. The goal is to determine if the issue is a lack of leads (volume) or a lack of product knowledge (AOV).

7. Implement a VIP Routing Rule: Automatically mandate that any order with an AOV exceeding $1,800 must be shipped via Federal Shipping. This protects our "Whale" customers from the 10+ day delays currently experienced with United Package.

8. Logistics Contract Renegotiation: Present this "Red Flag" data to United Package. Use the documented 13-day delay for Anne Dodsworth’s high-value accounts as leverage to demand improved service levels or discounted rates for non-priority shipments.

9. Sales-Logistics Alignment: Align top-performing "Closers" (those in the top-left of our performance matrix) with our most efficient shippers. We are currently pairing our best sales talent with our weakest logistics link, which risks long-term customer churn.

Task 4 – Tableau dashboard

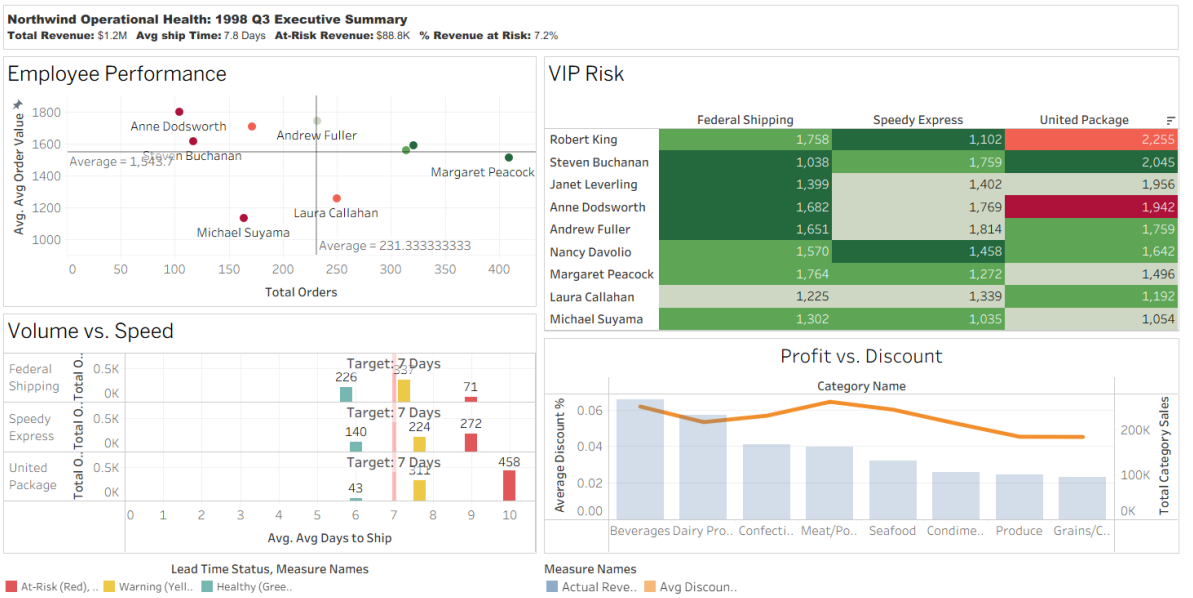

Figure 10 - Tableau dashboard

Northwind Operational Health Summary. Note the $88.8K Revenue at Risk (7.2%) linked primarily to United Package shipping delays exceeding 9 days.