

MAI AHMED BAKR  
DECEMBER 2024

# Strategic Analysis of Northwind Sales Data.

---

**Objective:** To provide a detailed, data-driven analysis of customer behavior, sales performance, and operational trends, offering actionable insights to optimize decision-making and enhance business strategy.

## AGENDA

# Analyzing Northwind sales data for actionable business insights.

1. Customer Segmentation Analysis
2. Product Performance Analysis
3. Order Analysis
4. Employee Performance Evaluation
5. Insights & Business Implications

# 1. Customer Segmentation: Grouping Based on Key Metrics

This section segments customers based on RFM (Recency, Frequency, Monetary) and Order Value to identify behavior patterns and inform targeted strategies for engagement and retention.

# Segment customers into categories based RFM Analysis: Recency

Assumption: The dataset's orderdate is historical. Query 1 uses the current date (DATE('now')) as the reference, while Query 2 uses the most recent order date in the dataset (2023-10-28) for a more realistic timeframe.

## Query 1: Recency Based on Current Date

```

6 SELECT
7 customerid,
8 julianday(DATE('now')) - julianday(MAX(orderdate)) AS Recency_Days
9 FROM Orders
10 GROUP BY customerid
11 ORDER BY Recency_Days

```

Result: Recency days range from 400 to 600 days, reflecting outdated data relative to today.

## Query 2: Recency Based on Dataset's Last Order Date

```

15 SELECT
16 customerid,
17 julianday((SELECT date(max(orderdate)) FROM orders)) - julianday(DATE(MAX(orderdate))) AS Recency_Days
18 FROM Orders
19 GROUP BY customerid
20 ORDER BY Recency_Days

```

Result: Recency days range from 0 to 170 days, better representing customer activity within the dataset's timeframe.

# Segment customers into categories based RFM Analysis: Frequency

Assumption:

The total number of orders indicates the customer's purchase frequency, and higher frequencies reflect more active and engaged customers.

Query : Frequency: Total number of orders (volume).

```
23 SELECT
24 customerid,
25 COUNT(orderid) AS Frequency
26 FROM Orders
27 GROUP BY 1
28 ORDER BY 2
```

Result:

The results show that the total number of orders per customer falls between 150 and 210, indicating a consistent level of engagement across customers.

# Segment customers into categories based RFM Analysis: Monetary

Assumption:

Total revenue per customer reflects their overall spending, where higher values indicate high-value customers with significant contributions to overall sales.

Query : Monetary Value: Total amount spent (revenue).

```

31 SELECT
32 customerid,
33 round(Sum((unitprice * quantity) * (1 - discount))) AS RevenuePerCustomer
34 FROM Orders
35 INNER JOIN "Order Details"
36 ON orders.OrderID = "Order Details"."OrderID"
37 GROUP BY 1
38 ORDER BY 2

```

Result:

The total revenue per customer ranges from 3,965,465 to 6,154,115, indicating a varied spending pattern across customers, with some contributing significantly more to total sales than others.

# Customer Segmentation:

## Identifying Key Customer Groups

### Assumption:

Using the RFM analysis view, we defined customer segments based on Recency, Frequency, and Monetary behavior with the following criteria:

- Champion: Customers who made their last purchase within 30 days, placed 170+ orders, and spent over 5.5 million.
- Potential Loyalist: Customers who made 170+ orders or spent over 5.5 million, but didn't meet the Champion criteria.
- At Risk: Customers who did not qualify as either Champion or Potential Loyalist.

### Query 1: Create RFM View

```

42 CREATE VIEW RFM AS
43 SELECT
44 o.customerid,
45 julianday((SELECT (date(max(orderdate))) FROM orders)) - julianday(DATE(MAX(orderdate))) AS Recency_Days,
46 COUNT(DISTINCT o.orderid) AS Frequency,
47 ROUND(SUM((od.unitprice * od.quantity) * (1 - od.discount))) AS RevenuePerCustomer
48 FROM Orders o
49 INNER JOIN "Order Details" od
50 ON o.OrderID = od."OrderID"
51 GROUP BY o.customerid

```

### Query 2: Create Segments

```

54 SELECT
55 CASE
56 WHEN Recency_Days <= 31 AND Frequency >= 170 AND RevenuePerCustomer >= 5500000 THEN 'Champion'
57 WHEN (Frequency >= 170) OR (RevenuePerCustomer >= 5500000) THEN 'Potential Loyalist'
58 ELSE 'At Risk'
59 END AS Segment,
60 COUNT(customerid) AS CustomerCount
61 FROM RFM
62 GROUP BY Segment
63 ORDER BY CustomerCount DESC

```

### — RESULTS

The customer segmentation results are as follows:

- Champion: 6 customers
- Potential Loyalist: 49 customers
- At Risk: 38 customers

The majority of customers fall into the Potential Loyalist category, indicating a strong group with potential for long-term loyalty. However, there is also a notable portion of customers in the At Risk segment, highlighting the need for targeted retention efforts.

# Segment customers into categories based on Order Values

Assumption:

After creating the CustomerOrderValue view, we observed that the average order values ranged between 710 and 760. To categorize customers meaningfully, we used percentiles to define the segments:

- Low Value: 710 to 730
- Medium Value: 730 to 750
- High Value: 750 to 760

This segmentation ensures a balanced classification of customers based on their average order value.

Query 1: Create CustomerOrderValue view.

```

68 CREATE VIEW CustomerOrderValue AS
69 SELECT
70 o.customerid,
71 AVG((od.UnitPrice * od.Quantity) * (1 - od.Discount)) AS AvgOrderValue
72 FROM "Order Details" od
73 INNER JOIN Orders o
74 ON od.orderid = o.orderid
75 GROUP BY 1
76 ORDER BY 2

```

Result: The average order value ranged between 710 and 760.

Query 2: Select customer order value segments.

```

79 SELECT
80 CASE
81 WHEN AvgOrderValue BETWEEN 710 AND 730 THEN 'Low Value'
82 WHEN AvgOrderValue BETWEEN 730 AND 750 THEN 'Medium Value'
83 ELSE 'High Value'
84 END AS OrderValueSegment,
85 COUNT(customerid) AS CustomerCount
86 FROM CustomerOrderValue
87 GROUP BY 1
88 ORDER BY 1

```

— RESULTS

The customer segmentation based on order value resulted in the following distribution:

- Low Value: 24 customers
- Medium Value: 63 customers
- High Value: 6 customers

This analysis highlights that the majority of customers fall into the medium value category, with fewer customers at the extremes of low and high value.

## 2. Product Performance Analysis: Revenue, Volume, and Slow Movers

This analysis highlights top-performing products by revenue and sales volume, while identifying slow-moving items to inform strategic decisions and optimize inventory.

# Product Performance Analysis: Top Revenue Generators

Assumption:

The goal is to identify the top 10 products contributing the most revenue. Revenue is calculated as the total sales value of each product after applying discounts.

Query : Top 10 revenue generator products

```
92 SELECT
93 productname,
94 ROUND(SUM((od.unitprice * od.quantity) * (1 - od.discount))) AS TotalRevenuePerProduct
95 FROM "Order Details" od
96 INNER JOIN Products p
97 ON od.productid = p.ProductID
98 GROUP BY 1
99 ORDER BY 2 DESC
100 LIMIT 10
```

Result:

The product “Côte de Blaye” generated the highest revenue, with a total of 53 million, significantly outperforming the other products in terms of revenue. This highlights its importance as a key revenue driver.

# Product Performance Analysis: Top Sales Volume Products

Assumption:

We are investigating whether the products with the highest order frequency also generate the highest revenue. To do this, we compare the most frequently ordered products with the top revenue-generating products.

Query : Top 10 most frequently ordered products.

```
103 SELECT
104 productname,
105 COUNT(od.orderid) AS QuantityPerProduct
106 FROM "Order Details" od
107 INNER JOIN Products p
108 ON od.productid = p.ProductID
109 GROUP BY 1
110 ORDER BY 2 DESC
111 LIMIT 10
```

Result:

Only two products, Sir Rodney's Marmalade and Raclette Courdavault, appear in both the top revenue and top frequency lists. The other products in the frequency list differ, suggesting that high order frequency does not always correlate with high revenue.

# Product Performance Analysis: Slow Movers

Assumption:

In this analysis, we are focusing on identifying products with low sales volume, defined by their relatively low total revenue. We are considering the top 5 products with the lowest revenue generation, based on sales data from the "Order Details" and "Products" tables.

Query : Products with low sales volume. (5 products)

```
114 SELECT
115 productname,
116 ROUND(SUM((od.unitprice * od.quantity) * (1 - od.discount))) AS TotalRevenuePerProduct
117 FROM "Order Details" od
118 INNER JOIN Products p
119 ON od.productid = p.ProductID
120 GROUP BY 1
121 ORDER BY 2
122 LIMIT 5
```

Result:

The product Geitost has the lowest total revenue, generating 507,121, which is nearly half of the next lowest product in the list. This significant difference highlights that Geitost stands out as the weakest performer among the slow-moving products.

### 3. Order Analysis: Identifying Trends and Patterns

This section analyzes order trends, including seasonality, popular order days, and order size distribution to uncover actionable insights for optimizing sales and inventory management.

# Order Analysis: Seasonality By Month

We assumed that analyzing seasonal fluctuations by month will help identify any patterns or variations in order volume. By examining these trends, we can better understand peak periods and adjust strategies for inventory management, marketing, and sales forecasts.

## Query 1: Seasonality by Month

```
128 SELECT
129  STRFTIME('%m', OrderDate) AS "Month",
130  COUNT(orderid) AS "Order volume"
131 FROM orders
132 GROUP BY "Month"
133 ORDER BY "Month"
```

Result:

Order volume shows consistency throughout the year, with peaks in March (1446) and August (1443), and the lowest in February (1221). This indicates stable demand, with slight seasonal fluctuations.

# Order Analysis: Seasonality By Day of Weak

We assumed that identifying the most popular order days will help determine patterns in customer behavior throughout the week. This can provide insights into the days with the highest order volume, indicating when demand is highest.

## Query 1: Seasonality by Day of Weak

```
136 SELECT  
137 strftime('%w', OrderDate) AS "DayOfWeek",  
138 COUNT(orderid) AS "OrderCount"  
139 FROM Orders  
140 GROUP BY "DayOfWeek"  
141 ORDER BY "DayOfWeek"
```

### Result:

Based on the results, you can observe the following:

- Monday has the highest number of orders, with 2448 orders.
- Wednesday and Thursday show relatively lower order volumes compared to the other days.
- Sunday has the lowest number of orders, with 2309, but it is not too far from the other weekdays.

## SEASONALITY BY MONTH

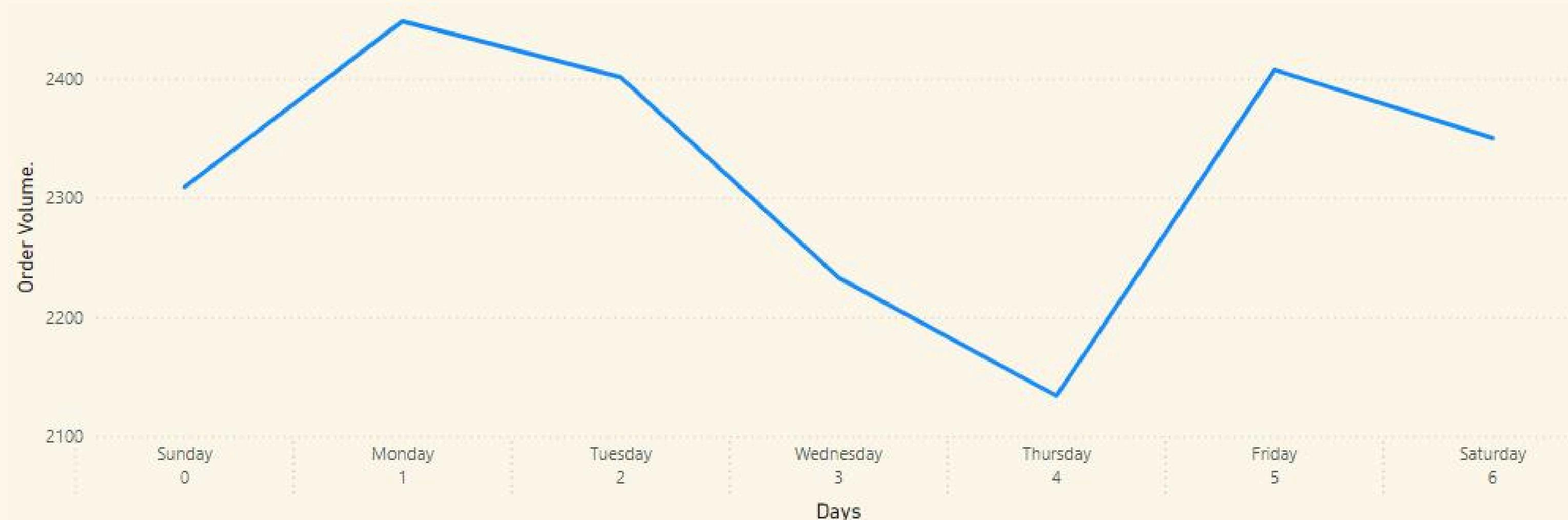


JANUARY

DECEMBER

SUNDAY

SATURDAY



## SEASONALITY BY DAY OF WEEK

# Order Analysis: Order Size

Assumption:

A view was created to calculate the total quantity per order by grouping by orderid, as each order might contain multiple products with repeated entries for the same order. This view consolidated the total quantity for each order.

To simplify the analysis, total order quantities were categorized into three segments:

- Small Order: 500 or fewer.
- Medium Order: 501 to 1000.
- Large Order: Greater than 1000.

## Query 1: Create View for Total Quantity per Order

```

145 CREATE VIEW TotalOrderQuantity AS
146 SELECT orderid,
147 SUM(quantity) AS TotalOrderQuantity
148 FROM "Order Details"
149 GROUP BY 1

```

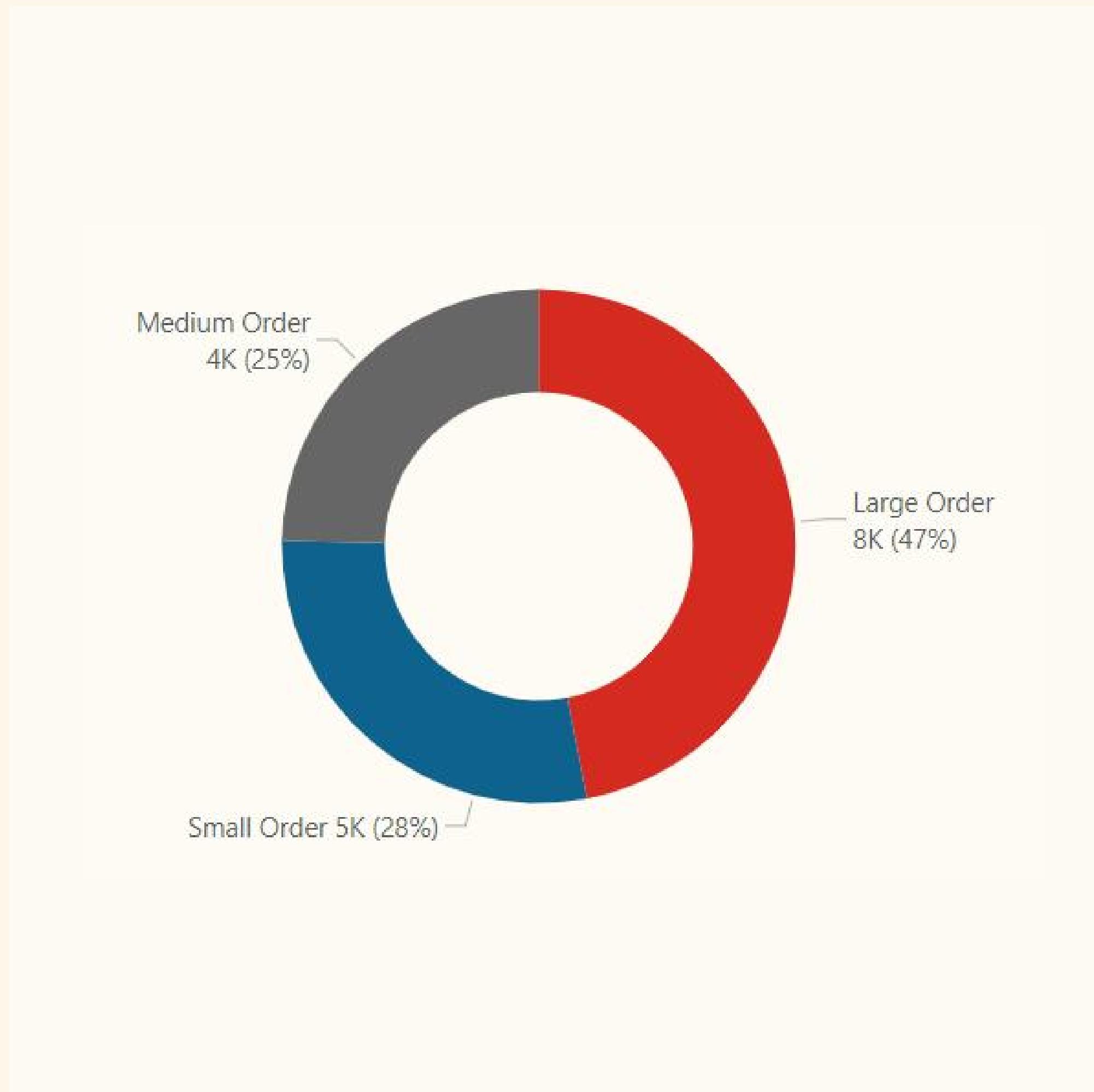
## Query 2: Categorize Orders into Segments

```

152 SELECT
153 CASE
154 WHEN TotalOrderQuantity <= 500 THEN 'Small Order'
155 WHEN TotalOrderQuantity BETWEEN 500 AND 1000 THEN 'Medium Order'
156 ELSE 'Large Order'
157 END AS OrderSizeCategory,
158 COUNT(*) AS OrderCount
159 FROM TotalOrderQuantity
160 GROUP BY 1

```

## DISTRIBUTION OF ORDER SIZES



## ORDER SIZE SEGMENTATION RESULTS

The distribution of orders based on size is as follows:

- Large Orders: 7,651 orders
- Medium Orders: 4,013 orders
- Small Orders: 4,618 orders

This segmentation shows that the majority of orders fall into the small and medium categories, with a significant proportion in the large order category.

## 4. Employee Performance Analysis

This section evaluates employee performance using key metrics: total revenue generated, orders processed, and average order value, providing insights into individual contributions.

# Employee Performance: Total Revenue Generated

Assumption:

To assess the revenue contribution of each employee, we calculated the total revenue generated by orders they processed.

Query : Total Revenue Generated Per Employee

```
165 SELECT
166 o.employeeid,
167 ROUND(SUM((od.unitprice * od.quantity) * (1 - od.discount))) AS TotalRevenuePerEmployee
168 FROM "order details" od
169 INNER JOIN Orders o
170 ON od.orderid = o.OrderID
171 GROUP BY 1
172 ORDER BY 1
```

Result:

Total revenue per employee ranges from 48,314,101 to 51,488,395. Employee ID 4 generated the highest revenue, while employee ID 2 generated the lowest. The average revenue is 49,820,737, with 6 employees performing below this average, indicating generally consistent performance with a few exceptions.

# Employee Performance: Total Sales Volume

Assumption:

To evaluate employee performance based on the number of orders processed, I calculated the total number of orders processed by each employee by grouping orders by employee ID.

Query : Total Sales Volume Per Employee.

```
175 SELECT
176 employeeid,
177 COUNT(orderid) AS TotalOrdersPerEmployee
178 FROM Orders
179 GROUP BY 1
180 ORDER BY 1
```

Result:

The number of orders processed by employees ranges from 1,754 to 1,908. Employee ID 4 processed the most orders, while employee ID 6 processed the least. This distribution is consistent with the revenue analysis, reflecting stable performance with slight variations across employees.

# Employee Performance: Average Order Value

Assumption:

The Average Order Value (AOV) is calculated by determining the average revenue generated per order processed by each employee. This provides insight into the performance of each employee in terms of handling higher-value orders.

Query : Average Order Value Per Employee

```
183 SELECT
184 o.employeeid,
185 ROUND(avg((od.unitprice * od.quantity) * (1 - od.discount))) AS TotalRevenuePerEmployee
186 FROM "order details" od
187 INNER JOIN Orders o
188 ON od.orderid = o.OrderID
189 GROUP BY 1
190 ORDER BY 1
```

Result:

The average order value per employee ranges from 728 to 742, with minimal variation, indicating consistent performance across the team.

## 5. Insights & Business Implications

This section provides actionable insights and recommendations to optimize revenue, customer retention, and employee performance.

# Customer Segmentation Insights and Strategic Recommendations

- The majority of customers fall into the Potential Loyalist segment, making them a key group for long-term engagement.
- The At Risk segment, though slightly smaller than the Potential Loyalist group, still represents a significant portion of the customer base, indicating the need for focused retention efforts.
- To maximize customer loyalty and revenue, strategies should prioritize converting Potential Loyalists into Champions and implement targeted campaigns to address the At Risk segment, reducing churn and enhancing customer stability.

**6%**

Champion Customers

**53%**

Potential Loyalists

**41%**

At Risk customers

# Product Performance Insights and Strategic Recommendations

- Maximize High-Performing Products: Côte de Blaye leads in revenue. Focus on promoting it to sustain its performance.
- Align Frequency with Revenue: Sir Rodney's Marmalade and Raclette Courdavault excel in both frequency and revenue. For others, create strategies to boost revenue alongside frequency.
- Boost Slow Movers: Geitost generates the lowest revenue among slow movers. Consider marketing strategies like bundling or discounts to increase sales or phase it out.

# Order Analysis: Insights and Strategic Recommendations

1. Seasonality by Month: Stable demand year-round with peaks in March and August. Focus on seasonal trends and maintain consistent supply.
2. Seasonality by Day: Highest volume on Monday, with lower orders on Wednesday and Thursday. Optimize staffing and marketing for high-volume days like Monday and Friday.
3. Order Size Distribution: Most orders are small and medium, with a significant number of large orders. Streamline processes for smaller orders and encourage larger ones.

## Employee Performance Insights and Recommendations:

- Six employees fall below the average of 49,820,737, with employees 2, 6, and 9 being the lowest performers. To address this, implement targeted training, motivation strategies, and performance-based incentives to boost their order volume and revenue.

DECEMBER 2024

Thank you