NORTHWIND TRADERS

IEC PORTFOLIO 1

GHULAM NABI HUSSAIN
DANISH YOUNUS KHAN
SIDRA NASIR

Cohort 12

Data Dictionary

Table Name	Field Name	Data type	Description	Example
Categories	Category ID	int	Unique identifier for each category.	1
Categories	Category Name	text	Name of the category	Beverages
Categories	Description	text	Description of the category	Soft drinks, coffees, teas, beers, and ales
Categories	Picture	BLOB	Picture representing the category	...
Customers	Customer ID	text	Unique identifier for each customer	ALFKI
Customers	Company Name	text	Name of the customer company	Alfreds Futterkiste
Customers	Contact Name	text	Name of the contact person	Maria Anders
Customers	Contact Title	text	Title of the contact person	Sales Representative
Customers	Address	text	Address of the customer	Obere Str. 57
Customers	City	text	City of the customer	Berlin
Customers	Region	text	Region of the customer	BC
Customers	Postal Code	text	Postal code of the customer	12209, WA1 1DP
Customers	Country	text	Postal code of the customer	Germany
Customers	Phone	text	Phone number of the customer	030-0074321

Customers	Fax	text	Fax number of the customer	030-0076545
Employees	Employee id	int	Unique identifier for each employee	1
Employees	Last name	text	Last name of the employee	Davolio
Employees	First name	text	First name of the employee	Nancy
Employees	title	text	Job title of the employee	Sales Representative
Employees	Title of courtesy	text	Title of courtesy (e.g., Mr., Mrs.)	Ms.
Employees	birthdate	text	Birth date of the employee	1948-08-12
Employees	Hire date	text	Hire date of the employee	1992-01-05
Employees	address	text	Address of the employee	507 - 20th Ave. E.Apt. 2A
Employees	city	text	City of the employee	Seattle
Employees	region	text	Region of the employee	WA
Employees	Postal code	text	Postal code of the employee	98122, SW1 8JR
Employees	country	text	Country of the employee	USA
Employees	Home phone	text	Home phone number of the employee	(206) 555-9857
Employees	extension	int	Extension number	5467

Employees	photo	BLOB	Photo of the employee	…
Employees	notes	text	Notes about the employee	Education includes a BA. in psychology from Colorado State University in 1970. She also completed The Art of the Cold Call."Nancy is a member of Toastmasters International."
Employees	Report to	text	ID of the manager the employee reports to	2, he was transferred
Employees	photograph	text	Photograph of the employee	http://accweb/emmployees /davolio.bmp
Employees	salary	double	Salary of the employee	2954.55
Employeeterritories	Employee ID	int	ID of the employee	2
Employeeterritories	Territory ID	int	ID of the territory	1581
Order details	Order ID	int	ID of the order	10248
Order details	Product ID	int	ID of the product	11
Order details	Unit Price	double	Unit price of the product	14, 9.8
Order details	Quantity	int	Quantity ordered	12
Order details	Discount	int	Discount applied	0
Orders	Order ID	int	Unique identifier for each order	10248
Orders	Customer ID	text	ID of the customer who placed the order	VINET
Orders	Employee ID	int	ID of the employee who processed the order	5

Orders	Order Date	text	Date when the order was placed	1996-07-04
Orders	Required Date	text	Date by which the order is required	1996-08-01
Orders	Shipped Date	text	Date when the order was shipped	1996-07-16
Orders	Ship Via	int	ID of the shipper used	3
Orders	Freight	double	Freight cost	32.38
Orders	Ship Name	text	Name of the shipper	Vins et alcools Chevalier
Orders	Ship Address	text	Address of the shipper	59 rue de l-Abbaye
Orders	Ship City	text	City of the shipper	Reims
Orders	Ship Region	text	Region of the shipper	RJ
Orders	Ship Postal Code	text	Postal code of the shipper	51100, B-6000
Orders	Ship Country	text	Country of the shipper	France
Products	Product ID	int	Unique identifier for each product	1
Products	ProductName	text	Name of the product	Chai
Products	Supplier ID	int	ID of the supplier providing the product	1
Products	Category ID	int	ID of the category the product belongs to	1
Products	Quantity Per Unit	text	Quantity of product per unit	10 boxes x 20 bags
Products	Unit Price	double	Price per unit of the product	18, 21.35
Products	Units In Stock	int	Number of units in stock	39

Products	Units On Order	int	Number of units on order	0
Products	Reorder Level	int	Level at which the product needs to be reordered	10
Products	Discontinued	int	Indicates if the product is discontinued	0
Region	Region ID	int	Unique identifier for each region	1
Region	Region Description	text	Description of the region	Eastern
Territories	Territory ID	int	Unique identifier for each territory	1581
Territories	Territory Description	text	Description of the territory	Westboro
Territories	Region ID	int	ID of the region	1
Shippers	Shipper ID	int	Unique identifier for each shipper	1
Shippers	Company Name	text	Name of the shipping company	Speedy Express
Shippers	Phone	text	Phone number of the shipper	(503) 555-9831
Suppliers	Supplier ID	int	Unique identifier for each supplier	1
Suppliers	Company Name	text	Name of the supplier company	Exotic Liquids
Suppliers	Contact Name	text	Name of the contact person	Charlotte Cooper
Suppliers	Contact Title	text	Title of the contact person	Purchasing Manager

Suppliers	Address	text	Address of the supplier	49 Gilbert St.
Suppliers	City	text	City of the supplier	London
Suppliers	Region	text	Region of the supplier	LA
Suppliers	Postal Code	text	Postal code of the supplier	EC1 4SD
Suppliers	Country	text	Country of the supplier	UK
Suppliers	Phone	text	Phone number of the supplier	(171) 555-2222
Suppliers	Fax	text	Fax number of the supplier	(313) 555-3349
Suppliers	Homepage	text	Homepage of the supplier	Mayumi's (on the World Wide Web) #http://www.microsoft.com/accessdev/sampleapps/mayumi.htm#

Sales Analysis

1. **What are the total sales for each month over the past year? (Monthly Sales Trends)**

Objective:

To identify the total sales for each month over the past year. This allows for a detailed understanding of seasonal trends, peak sales periods, and potential areas for improvement in sales strategies.

Query:

```sql
SELECT DATE_FORMAT(OrderDate, '%Y-%m') AS Month, ROUND(SUM(UnitPrice * Quantity),0) AS TotalSales
FROM orders
INNER JOIN orderdetails ON orders.OrderID = orderdetails.OrderID
GROUP BY Month
ORDER BY Month;
```

Result:

Month	TotalSales
1996-07	30192
1996-08	26609
1996-09	27636
1996-10	41204
1996-11	49704
1996-12	50953
1997-01	66693
1997-02	41207
1997-03	39980
1997-04	55699
1997-05	56824
1997-06	39088
1997-07	55465
1997-08	49982
1997-09	59733
1997-10	70329
1997-11	45913

The result identifies total sales according to months. This helps in understanding when sales are highest and planning inventory and promotions accordingly.

Recommendations:

Peak Periods: If certain months show consistently high sales, ensure that inventory levels are sufficient to meet demand during these periods. Plan targeted promotions and marketing campaigns to maximize sales.

Slow Periods: For months with lower sales, consider implementing special promotions, discounts, or marketing efforts to boost sales. Analyze why sales might be lower during these months and address any underlying issues.

Seasonal Trends: If sales show clear seasonal patterns, align your product launches, marketing campaigns, and inventory management with these trends to optimize performance throughout the year.

2. How has the total sales revenue grown year over year? (Yearly Sales Growth)

OBJECTIVE:

      To analyze how the total sales revenue has grown year over year. This helps in understanding the business's growth trajectory, identifying trends, and making informed decisions about future strategies.

Query:

```sql
SELECT YEAR(OrderDate) AS Year, SUM(UnitPrice * Quantity) AS TotalSales
FROM orders
INNER JOIN orderdetails ON orders.OrderID = orderdetails.OrderID
GROUP BY YEAR(OrderDate)
ORDER BY Year;
```

Result:

Year	TotalSales
1996	226298.5000000001
1997	658388.7499999998
1998	469771.34

The result shows total sales growth yearly. This helps us to understand how sales increase or decrease by year or whether there is consistent growth in sales year over year.

Recommendation:

If the results show sustained growth, consider investing more in successful product lines, expanding to new markets, and increasing marketing efforts.

Inventory management:

3. **What is the inventory turnover rate for each product over the past year? (Inventory Turnover Rate)**

Objective:

Calculate the inventory turnover rate for each product over the past year. This metric helps understand how efficiently inventory is being managed and how often stock is being sold and replaced within a year.

Query:

```sql
SELECT products.ProductID, products.ProductName, SUM(orderdetails.Quantity) / AVG(products.UnitsInStock) AS TurnoverRate
FROM orderdetails
INNER JOIN products ON orderdetails.ProductID = products.ProductID
INNER JOIN orders ON orderdetails.OrderID = orders.OrderID
WHERE orderdetails.Quantity IS NOT NULL AND products.UnitsInStock IS NOT NULL
GROUP BY products.ProductID, products.ProductName;
```

Result

ProductID	ProductName	TurnoverRate
72	Mozzarella di Giovanni	57.5714
42	Singaporean Hokkien Fried Mee	26.8077
11	Queso Cabrales	32.0909
51	Manjimup Dried Apples	44.3000
14	Tofu	11.5429
65	Louisiana Fiery Hot Pepper Sauce	9.8026
41	Jack's New England Clam Chowder	11.5412
57	Ravioli Angelo	12.0556
22	Gustaf's Knckebrd	3.3462
60	Camembert Pierrot	83.0000
33	Geitost	6.7411
20	Sir Rodney's Marmalade	7.8250
49	Maxilaku	52.0000

The result indicates how many times the inventory has been sold and replaced over the past year. A higher turnover rate suggests efficient inventory management, while a lower rate may indicate overstocking or slow-moving inventory.

Recommendation:

Improve Inventory Management: For products with low turnover rates, analyze demand patterns and adjust inventory levels accordingly to reduce holding costs and free up capital.

Enhance Sales Strategies: For slow-moving products, consider implementing promotional strategies, discounts, or bundling options to boost sales and improve turnover rates.

Optimize Ordering Practices: For high turnover products, ensure that ordering practices are optimized to maintain adequate stock levels and avoid stockouts. Consider increasing order quantities or frequency for these products.

Product & Customer Analysis

4. **What is the reorder frequency of each product?**

Objective:
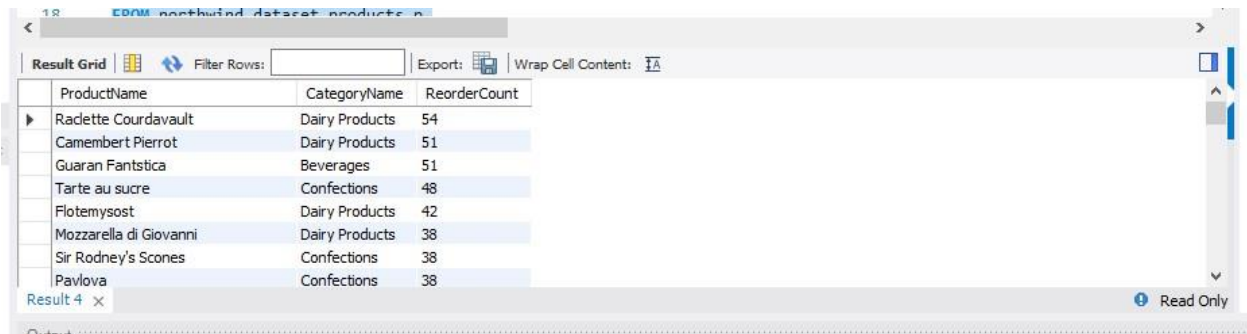
Determine the reorder frequency of each product to ensure adequate inventory management and timely restocking.

Query:

```
78
79 •  SELECT p.ProductName, c.CategoryName, COUNT(*) AS ReorderCount
80     FROM northwind_dataset.products p
81     JOIN northwind_dataset.orderdetails od ON p.ProductID = od.ProductID
82     JOIN northwind_dataset.categories c ON p.CategoryID = c.CategoryID
83     WHERE od.Quantity >= p.ReorderLevel
84     GROUP BY p.ProductName, c.CategoryName
85     ORDER BY ReorderCount DESC;
86
87
```

Result:
The results help in understanding the demand pattern for each product by calculating how frequently they need to be reordered. It is crucial for inventory management to ensure products are available when needed and to identify fast-moving items.

ProductName	CategoryName	ReorderCount
Raclette Courdavault	Dairy Products	54
Camembert Pierrot	Dairy Products	51
Guaran Fantstica	Beverages	51
Tarte au sucre	Confections	48
Flotemysost	Dairy Products	42
Mozzarella di Giovanni	Dairy Products	38
Sir Rodney's Scones	Confections	38
Pavlova	Confections	38

Recommendation:

Use the reorder frequency data to prioritize inventory management, ensuring that products in high demand are adequately stocked to meet customer needs.

5. **Question**
    a) **Total sale of product and which customer bought them?**
    b) **Category-wise breakdown of products and their total sales?**
    c) **Quantity wise breakdown?**

Objectives:
a) To determine the total sales of each product and identify the corresponding customers who purchased them, providing insights into consumer purchasing behavior and facilitating personalized marketing strategies.

b) To analyze the total sales of products within each category, enabling the assessment of category performance and the identification of high-performing product categories for strategic resource allocation and marketing efforts.

c) To examine the breakdown of product quantities available, categorized by product and product category, aiding in inventory management decisions, and ensuring optimal stock levels to meet customer demand efficiently.

Q5a.

```
Q4a.
SELECT o.CustomerID, od.ProductID, p.ProductName, SUM(od.UnitPrice * od.Quantity) AS TotalSales
FROM northwind_dataset.orderdetails od
JOIN northwind_dataset.orders o ON od.OrderID = o.OrderID
JOIN northwind_dataset.products p ON od.ProductID = p.ProductID
GROUP BY o.CustomerID, od.ProductID, p.ProductName;
```
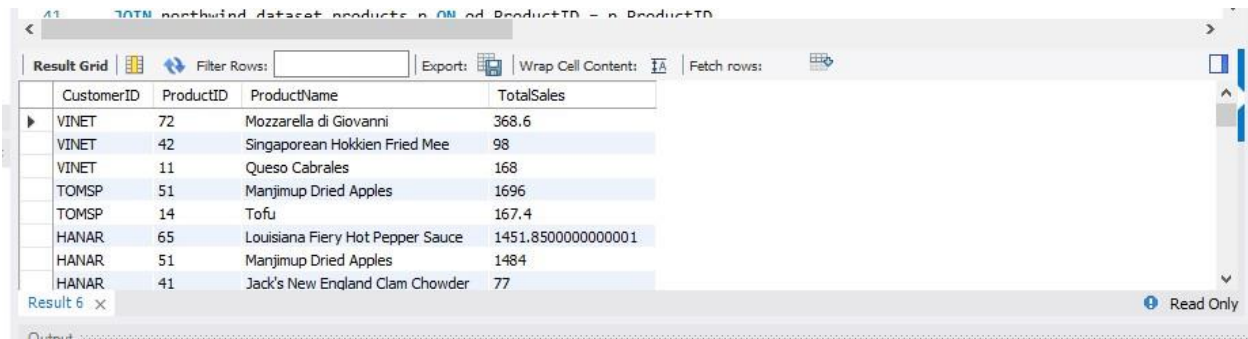
Q5b.

```
Q4b.
SELECT p.CategoryID, c.CategoryName, p.ProductName, SUM(od.UnitPrice * od.Quantity) AS TotalSales
FROM northwind_dataset.orderdetails od
JOIN northwind_dataset.products p ON od.ProductID = p.ProductID
JOIN northwind_dataset.categories c ON p.CategoryID = c.CategoryID
GROUP BY p.CategoryID, c.CategoryName, p.ProductName;
```

Q5c.Quantity wise breakdown

```
5
6 ☒    Q4c.Quantity wise breakdown
7        SELECT p.ProductID, p.ProductName, c.CategoryName, p.UnitsInStock AS Quantity_Present
8        FROM northwind_dataset.products p
9        JOIN northwind_dataset.categories c ON p.CategoryID = c.CategoryID;
0
1        |
```

Result:

Aggregating sales data by product category offers insights into revenue-generating categories, aiding in market trend analysis and resource allocation. Simultaneously, monitoring current stock quantities with category information facilitates effective inventory management, enabling businesses to maintain optimal stock levels and identify products requiring replenishment promptly.
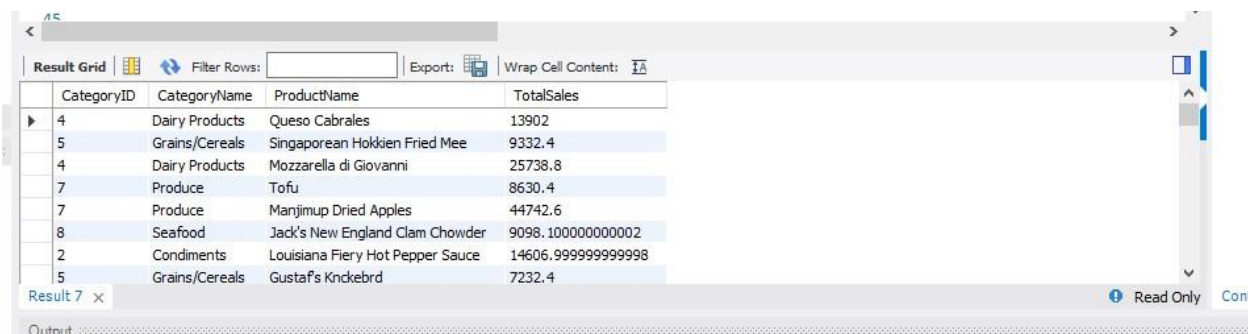

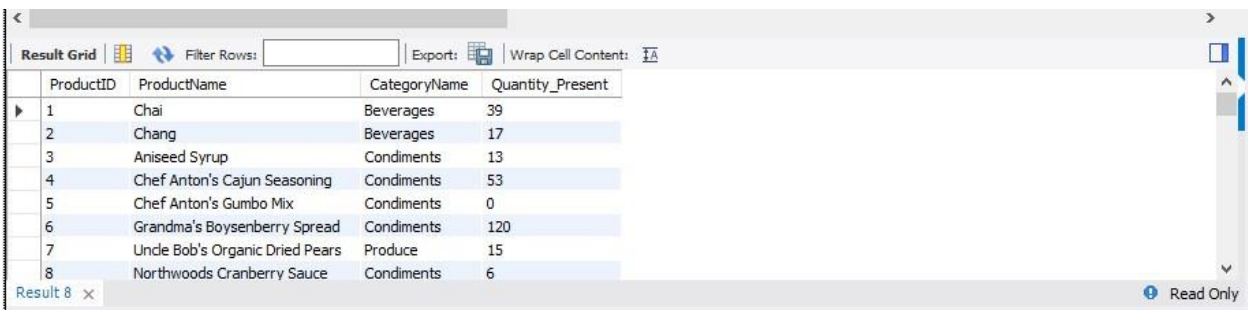| CustomerID	ProductID	ProductName	TotalSales
VINET	72	Mozzarella di Giovanni	368.6
VINET	42	Singaporean Hokkien Fried Mee	98
VINET	11	Queso Cabrales	168
TOMSP	51	Manjimup Dried Apples	1696
TOMSP	14	Tofu	167.4
HANAR	65	Louisiana Fiery Hot Pepper Sauce	1451.8500000000001
HANAR	51	Manjimup Dried Apples	1484
HANAR	41	Jack's New England Clam Chowder	77

Result 6

CategoryID	CategoryName	ProductName	TotalSales
4	Dairy Products	Queso Cabrales	13902
5	Grains/Cereals	Singaporean Hokkien Fried Mee	9332.4
4	Dairy Products	Mozzarella di Giovanni	25738.8
7	Produce	Tofu	8630.4
7	Produce	Manjimup Dried Apples	44742.6
8	Seafood	Jack's New England Clam Chowder	9098.100000000002
2	Condiments	Louisiana Fiery Hot Pepper Sauce	14606.999999999998
5	Grains/Cereals	Gustaf's Knckebrd	7232.4

Result 7

ProductID	ProductName	CategoryName	Quantity_Present
1	Chai	Beverages	39
2	Chang	Beverages	17
3	Aniseed Syrup	Condiments	13
4	Chef Anton's Cajun Seasoning	Condiments	53
5	Chef Anton's Gumbo Mix	Condiments	0
6	Grandma's Boysenberry Spread	Condiments	120
7	Uncle Bob's Organic Dried Pears	Produce	15
8	Northwoods Cranberry Sauce	Condiments	6

Result 8

Recommendation:

Utilize the insights gained from analyzing the total sales of each product and the corresponding customers to tailor personalized marketing strategies. Implement customer segmentation based on purchasing behavior to target specific customer groups effectively. Additionally, consider offering incentives or promotions to encourage repeat purchases from high-value customers. Similarly, leverage the category-wise breakdown of product sales to identify top-performing product categories and capitalize on emerging market trends. Allocate resources towards promoting products within these high-performing categories to maximize profitability. Furthermore, consider diversifying product offerings within successful categories to meet evolving consumer preferences and capture a larger market share.

Freight Analysis:

6. What are the shipping details for each shipped order?

Objective:

Calculate the average freight cost for each shipper over time, allowing for the identification of trends, comparisons between shippers, and potential cost optimization strategies.

Query & Result:

```sql
SELECT
    c.City AS CustomerCity,
    o.ShipCity,
    o.ShippedDate,
    o.RequiredDate,
    o.Freight,
    DATEDIFF(o.ShippedDate, o.OrderDate) AS DaysToShip,
    GREATEST(DATEDIFF(o.ShippedDate, o.RequiredDate), 0) AS DaysLate
FROM orders o
JOIN customers c ON o.CustomerID = c.CustomerID
WHERE o.ShippedDate IS NOT NULL
ORDER BY o.ShippedDate;
```

CustomerCity	ShipCity	ShippedDate	RequiredDate	Freight	DaysToShip	DaysLate
Mnster	Mnster	1996-07-10 00:00:00	1996-08-16 00:00:00	11.61	5	0
Charleroi	Charleroi	1996-07-11 00:00:00	1996-08-06 00:00:00	51.3	2	0
Rio de Janeiro	Rio de Janeiro	1996-07-12 00:00:00	1996-08-05 00:00:00	65.83	4	0
Lyon	Lyon	1996-07-15 00:00:00	1996-08-05 00:00:00	41.34	7	0
Genve	Genve	1996-07-15 00:00:00	1996-08-09 00:00:00	148.33	3	0
Reims	Reims	1996-07-16 00:00:00	1996-08-01 00:00:00	32.38	12	0
Rio de Janeiro	Rio de Janeiro	1996-07-16 00:00:00	1996-07-24 00:00:00	58.17	6	0
Resende	Resende	1996-07-17 00:00:00	1996-08-12 00:00:00	13.97	2	0
San Cristbal	San Cristbal	1996 07 22 00:00:00	1996 08 13 00:00:00	81.91	6	0

Results Interpretation:

The data presents a comprehensive view of shipping activities, encompassing details about the customer's location, the shipment's destination, the associated freight cost, the duration it took to ship the order, and any delays in delivery.

Key observations from the provided data include:

Diverse Customer Base: The orders originate from various cities and countries, indicating a wide geographical spread of customers.

Timely Deliveries: Most orders are delivered within a reasonable timeframe, with DaysToShip ranging from 2 to 12 days. Notably, there are no instances of late deliveries in this dataset, as indicated by DaysLate being 0 for all entries.

Freight Cost Variations: The shipping costs (Freight) exhibit significant variability, ranging from 11.61 to 148.33. This variation is observed even for shipments to the same destination (e.g., Rio de Janeiro), suggesting that factors beyond distance, such as shipper choice or service level, influence the final cost.

Power BI Narrative & Business Insights:

This data is prime for a compelling narrative within a Power BI dashboard. Here is how it could be used:

Shipping Performance Analysis: Visualize the Days to Ship and Days Late data to understand how long it typically takes to fulfill orders and how often they are late. This can reveal bottlenecks in your shipping process and identify areas for improvement.

Freight Cost Analysis: Analyze Freight data across different Customer City and Shipped City pairs to understand how shipping costs vary based on origin and destination. You could create visualizations like:

Operational Efficiency: Identify trends in shipping times and costs over time to see if your operations are becoming more efficient or if there are areas that need attention.

Cost Optimization: Use the insights gained to make data-driven decisions about optimizing shipping routes, negotiating better rates with carriers, or adjusting pricing to account for shipping costs.

By exploring the data visually in Power BI, you can gain valuable insights to optimize your shipping operations, improve customer satisfaction, and ultimately enhance your bottom line.

7. What is the shipping cost summary for each shipping company?

Objective:

      The objective of the query is to analyze the freight costs associated with different shipping companies to assess their overall cost-effectiveness and identify potential cost-saving opportunities.

Query & Result:

```
1  SELECT
2      s.CompanyName,
3      COUNT(*) AS NumShipments,
4      MIN(o.Freight) AS MinFreight,
5      MAX(o.Freight) AS MaxFreight,
6      AVG(o.Freight) AS AvgFreight,
7      STDDEV(o.Freight) AS StdDevFreight
8  FROM orders o
9  JOIN shippers s ON o.ShipVia = s.ShipperID
10 GROUP BY s.CompanyName;
```

CompanyName	NumShipments	MinFreight	MaxFreight	AvgFreight	StdDevFreight
Federal Shipping	255	0.4	1007.64	80.44121568627443	119.12856967427985
Speedy Express	249	0.12	458.78	65.00132530120482	75.24204211914798
United Package	326	0.02	890.78	86.6406441717791	138.00710686960133

Results Interpretation:

The results provide a statistical summary of freight costs for each shipping company:

Federal Shipping: Handled 255 shipments, with freight costs ranging from 0.4 to 1007.64, averaging 80.44 per shipment.

Speedy Express: Handled 249 shipments, with a smaller freight cost range (0.12 to 458.78) and a lower average of 65.00 per shipment.

United Package: Handled the most shipments (326), with the widest range of freight costs (0.02 to 890.78) and the highest average of 86.64 per shipment.

Power BI Narrative & Business Insights:

This data is valuable for crafting a narrative around shipping costs and performance within a Power BI dashboard. Here's how it could be utilized:

- **Shipper Comparison:**

  o A **bar chart** could visualize the average freight cost for each shipper, allowing for easy comparison of their overall cost-effectiveness.

  o A **scatter plot** with NumShipments on one axis and Av Freight on the other could reveal if there's a relationship between shipment volume and average cost.

- **Cost Variability:**

  o Box plots or histograms could display the distribution of freight costs for each shipper, highlighting outliers and showing which shipper has the most consistent pricing.

  o The StdDevFreight value could be included in tooltips or as a separate table to quantify the variability.

8. **What was the average freight cost incurred by each shipping company in each month of the specified time period?**

Objective:

The objective of the query is to analyse the average freight cost per month for each shipping company to identify trends in shipping expenses over time and facilitate comparisons between different carriers.

Query & Result:

```
1  SELECT
2      s.CompanyName,
3      DATE_FORMAT(o.OrderDate, '%Y-%m') AS OrderMonthYear,
4      FORMAT(AVG(o.Freight), 2) AS AvgFreight
5  FROM orders o
6  JOIN shippers s ON o.ShipVia = s.ShipperID
7  GROUP BY s.CompanyName, OrderMonthYear
8  ORDER BY s.CompanyName, OrderMonthYear;
```

CompanyName	OrderMonthYear	AvgFreight
Speedy Express	1998-05	44.98
United Package	1998-05	50.52
Federal Shipping	1998-04	76.03
Speedy Express	1998-04	51.27
United Package	1998-04	115.49
Federal Shipping	1998-03	62.77
Speedy Express	1998-03	50.63
United Package	1998-03	94.88
Federal Shipping	1998-02	46.50

Results Interpretation:

The data shows how the average freight costs fluctuate for each shipping company over time. For example, we see that United Package had the highest average freight cost in April 1998 and the lowest in May 1998, indicating significant volatility.

Power BI Narrative & Business Insights:

This data is prime for a compelling narrative within a Power BI dashboard. Here's how it could be used:

- **Trend Analysis:** A **line chart** with OrderMonthYear on the x-axis and Av Freight on the y-axis, with separate lines for each shipper, would visually showcase the trend of average freight costs over time. This could reveal seasonal patterns, long-term trends, or anomalies in shipping costs.

- **Shipper Comparison:** A **bar chart** comparing the average freight costs of different shippers for each month could easily identify the most and least cost-effective options. This could inform decisions about which shippers to use for future orders.

- **Cost Optimization:** By identifying months with high average freight costs, businesses can investigate the underlying causes and explore strategies to mitigate these costs, such as negotiating better rates with shippers or consolidating shipments.

By visualizing the data in Power BI, the business can gain valuable insights into the shipping costs and make informed decisions to optimize the logistics operations.