Northwind Traders

Sales Analysis

Northwind Traders, a made-up gourmet food supplier, has provided sales data from 1994 to 1996 for conducting Exploratory Data Analysis (EDA).

OBJECTIVES

The project's major goal was to design and build a dynamic KPI dashboard using Power BI that would allow the executive team to get insights into critical parts of Northwind Traders' operations. The dashboard emphasized the following critical areas:

Sales Trends: An in-depth examination of sales performance over time, including revenue trends, sales growth, and comparisons to previous periods. These insights would allow management to assess the company's financial health, recognize developing patterns, and make strategic decisions based on them.

Product Performance: Provides an in-depth look into the performance of several gourmet food products supplied by Northwind Traders. The dashboard emphasized top-selling products, product category trends, and profitability analyses, making it easier to manage product portfolios and discover chances for improvement.

Key Customer: Identifying and profiling Northwind Traders' most valued customers, allowing the executive team to better understand consumer preferences, evaluate purchase trends, and develop strategies to improve customer happiness and retention. Customer revenue contribution, order frequency, and customer segmentation were all supplied by the dashboard.

Shipping Costs: Assessing the influence of shipping costs on the company's profitability. The dashboard exhibited shipping cost trends, comparisons of different delivery methods, and high-cost shipping regions. These insights aided in the optimization of logistics and the reduction of operational costs.

Other objectives of this analysis are as follows:

- Estimate the total number of orders and the revenue generated.
- Determine sales trends over time.
- Identify the best and worst selling products.
- Determine the most frequently sold product categories.
- Identify the most commonly used shipper by customers.
- Perform geographic analysis of sales.

APPROACH

To achieve the project goals, I used Power BI's powerful capabilities to transform raw data into interactive representations. The following actions were taken during the development process:

Dashboard Design and Visualization: I created a visually stunning and user-friendly dashboard layout using Power BI's intuitive interface. I chose relevant charts, graphs, and visual components to properly convey critical performance metrics. The dashboard was created to support changing data while also allowing for easy exploration and filtering for deeper insights.

Data Modelling and Analysis: In Power BI, I constructed data models by building linkages between various datasets. This allowed for cross-referencing and detailed analysis across various dimensions. To extract significant insights, advanced analytical approaches such as time series analysis, segmentation, and statistical computations were used.

Insights and Recommendations: In addition to creating the dashboard, I conducted extensive research to produce useful insights and recommendations. I developed practical recommendations for the executive team by integrating data-driven insights with subject knowledge. These ideas intended to improve sales strategies, product offerings, customer connections, and logistics operations.

DATA SOURCE

The data used for this analysis was obtained from Maven Analytics, a reputable source known for its comprehensive datasets.

Methodology:

1. Data collection: The data was collected from Maven Analytics and consisted of several CSV files containing information on customers, products, orders, shippers, and employees.

2. Data preprocessing: The data from different tables were merged using full joins in Power Query, utilizing their primary keys. This integration resulted in a dataset consisting of 2179 rows and 8 columns.

orders.csv

File Origin: 65001: Unicode (UTF-8) | Delimiter: Comma | Data Type Detection: Based on first 200 rows

OrderID	CustomerID	EmployeeID	OrderDate	RequiredDate	ShippedDate	ShipVia	Freight	ShipName
10248	VINET	5	04-08-1994	01-09-1994	16-08-1994	3	32.38	Vins et alcools Chevalier
10249	TOMSP	6	05-08-1994	16-09-1994	10-08-1994	1	11.61	Toms Spezialitäten
10250	HANAR	4	08-08-1994	05-09-1994	12-08-1994	2	65.83	Hanari Carnes
10251	VICTE	3	08-08-1994	05-09-1994	15-08-1994	1	41.34	Victuailles en stock
10252	SUPRD	4	09-08-1994	06-09-1994	11-08-1994	2	51.3	Suprêmes délices
10253	HANAR	3	10-08-1994	24-08-1994	16-08-1994	2	58.17	Hanari Carnes
10254	CHOPS	5	11-08-1994	08-09-1994	23-08-1994	2	22.98	Chop-suey Chinese
10255	RICSU	9	12-08-1994	09-09-1994	15-08-1994	3	148.33	Richter Supermarkt
10256	WELLI	3	15-08-1994	12-09-1994	17-08-1994	2	13.97	Wellington Importadora
10257	HILAA	4	16-08-1994	13-09-1994	22-08-1994	3	81.91	HILARIÓN-Abastos
10258	ERNSH	1	17-08-1994	14-09-1994	23-08-1994	1	140.51	Ernst Handel
10259	CENTC	4	18-08-1994	15-09-1994	25-08-1994	3	3.25	Centro comercial Moctezur
10260	OTTIK	4	19-08-1994	16-09-1994	29-08-1994	1	55.09	Ottilies Käseladen
10261	QUEDE	4	19-08-1994	16-09-1994	30-08-1994	2	3.05	Que Delícia
10262	RATTC	8	22-08-1994	19-09-1994	25-08-1994	3	48.29	Rattlesnake Canyon Grocer
10263	ERNSH	9	23-08-1994	20-09-1994	31-08-1994	3	146.06	Ernst Handel
10264	FOLKO	6	24-08-1994	21-09-1994	23-09-1994	3	3.67	Folk och fä HB
10265	BLONP	2	25-08-1994	22-09-1994	12-09-1994	1	55.28	Blondel père et fils
10266	WARTH	3	26-08-1994	07-10-1994	31-08-1994	3	25.73	Wartian Herkku

Extract Table Using Examples | Load | Transform Data | Cancel

To ensure data integrity, comprehensive data cleaning was done, which included resolving missing data, standardizing data formats, and fixing spelling problems. A final dataset of 2155 rows and 8 columns was produced in **order_details** table after irrecoverably missing data points were eliminated since the number of rows with null values were less than 1% so removed those columns.

Entity-Relationship Diagram (ER diagram) is setup with all the tables given. Here, I have used ***2 Fact tables and 7 Fact tables.***

Then I cleaned up the customers table by removing columns that aren't important to my investigation. The categories table also required cleaning, so I removed columns that were unnecessary for my study. The workers table also needed cleaning, so I cleaned it to remove columns that were unnecessary for my study.

customer_table: Provides information on the company's customers, such as company names, contact names, and locations

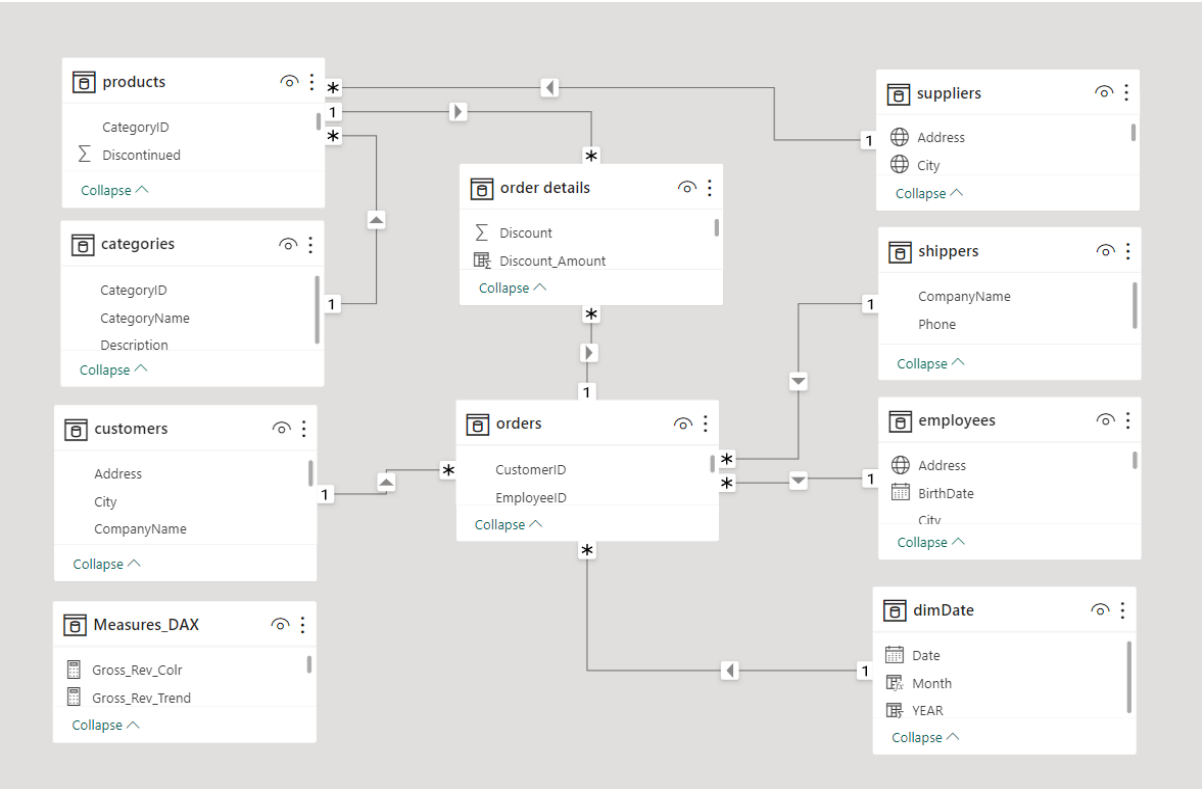
order_table: Contains information about individual orders, including order dates, customer IDs, employee IDs, and shipping details

employees_table: Contains employee records, including names, titles, and contact information

products_table: Offers information about the products Northwind Traders sells, including product names, unit prices, and categories

shippers_table: Contains information about shipping companies used by Northwind Traders.

Categories_table: Provides details about product categories, such as category names and descriptions.



	OrderID	ProductID	UnitPrice	Quantity	Discount
	Valid 100%	Valid 100%	Valid 100%	Valid 100%	Valid 100%
	Error 0%	Error 0%	Error 0%	Error 0%	Error 0%
	Empty 0%	Empty 0%	Empty 0%	Empty 0%	Empty 0%
	378 distinct, 51 unique	77 distinct, 0 unique	114 distinct, 5 unique	46 distinct, 6 unique	6 distinct, 0 unique
1	10248	11	14	12	0
2	10248	42	9.8	10	0
3	10248	72	34.8	5	0
4	10249	14	18.6	9	0
5	10249	51	42.4	40	0
6	10250	41	7.7	10	0
7	10250	51	42.4	35	0.15
8	10250	65	16.8	15	0.15
9	10251	22	16.8	6	0.05
10	10251	57	15.6	15	0.05
11	10251	65	16.8	20	0
12	10252	20	64.8	40	0.05
13	10252	33	2	25	0.05
14	10252	60	27.2	40	0
15	10253	31	10	20	0
16	10253	39	14.4	42	0

After performing all the cleaning and transformation operations, I checked the column quality and column distribution which is shown above in diagram.

Based on the given data fields, I calculated ***Gross_Revenue, Discount_Amount*** and ***Net_Revenue*** using DAX formula shown below:

```
Gross_Revenue = 'order details'[UnitPrice]*'order details'[Quantity]
```

OrderID	ProductID	UnitPrice	Quantity	Discount	Gross_Revenue	Discount_Amount	Net_Revenue
10251	65	$16.80	20	0	$336.00	$0.00	$336.00
10253	31	$10.00	20	0	$200.00	$0.00	$200.00
10255	2	$15.20	20	0	$304.00	$0.00	$304.00
10261	21	$8.00	20	0	$160.00	$0.00	$160.00
10261	35	$14.40	20	0	$288.00	$0.00	$288.00
10265	70	$12.00	20	0	$240.00	$0.00	$240.00
10273	33	$2.00	20	0	$40.00	$0.00	$40.00
10274	71	$17.20	20	0	$344.00	$0.00	$344.00
10277	28	$36.40	20	0	$728.00	$0.00	$728.00
10280	55	$19.20	20	0	$384.00	$0.00	$384.00
10283	15	$12.40	20	0	$248.00	$0.00	$248.00
10287	34	$11.20	20	0	$224.00	$0.00	$224.00
10290	5	$17.00	20	0	$340.00	$0.00	$340.00
10292	20	$64.80	20	0	$1,296.00	$0.00	$1,296.00
10297	72	$27.80	20	0	$556.00	$0.00	$556.00
10299	70	$12.00	20	0	$240.00	$0.00	$240.00
10300	68	$10.00	20	0	$200.00	$0.00	$200.00
10301	56	$30.40	20	0	$608.00	$0.00	$608.00
10309	4	$17.60	20	0	$352.00	$0.00	$352.00
10309	43	$36.80	20	0	$736.00	$0.00	$736.00
10312	53	$26.20	20	0	$524.00	$0.00	$524.00

3.Exploratory data analysis: To better comprehend the data and discover connections between variables, the clean dataset was thoroughly examined. Insights from the data were extracted using pivot tables. All these were performed in Excel to answer EDA questions.

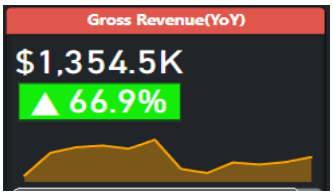
4.Visualization: Charts and maps were built to effectively show critical insights, key performance indices (KPIs), and trends. These visuals make the analysis more accessible and understandable to the intended audience.

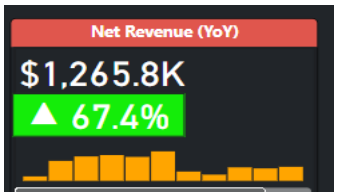
Tool used: Excel, Power BI

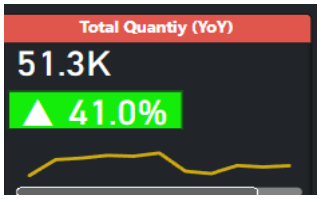
ANALYSIS AND OBSERVATIONS

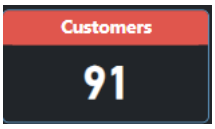
1. Sales Analysis:

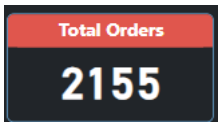
Between June 1994 and June 1996, a total of 2155 orders were recorded, generating $1265.8K in sales as Net revenue and $1354.5K as gross revenue.

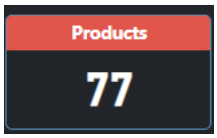


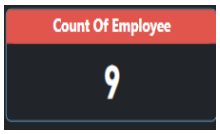






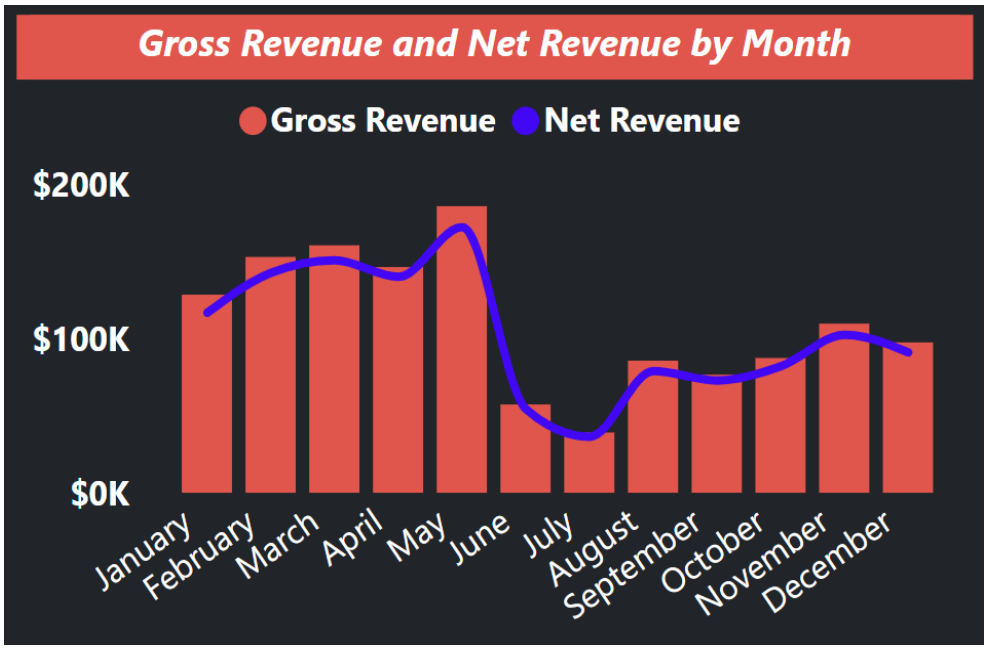






2. Sales Trend Overtime:

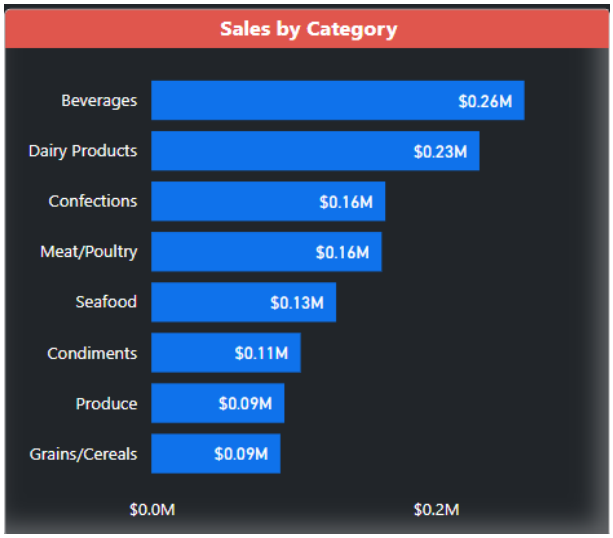
Sales recorded maximum in May and experienced steady decline throughout June and July, and then started increasing slowly in coming months and ended up as approximately average of $100K net revenue monthly.

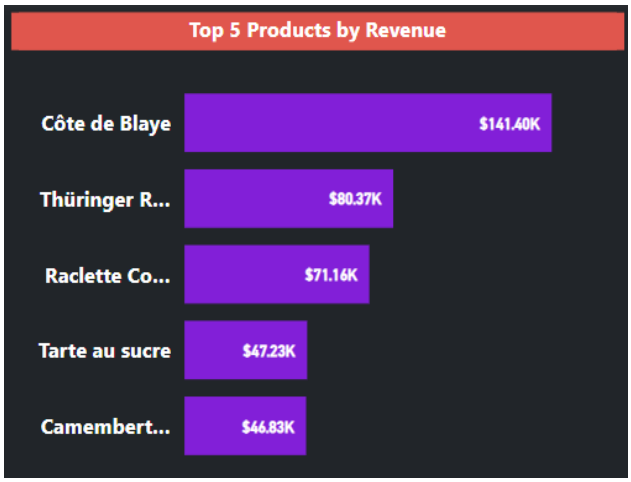


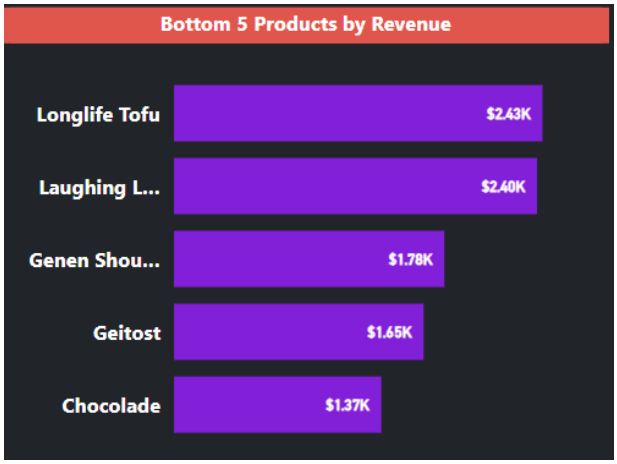
3. Product Analysis:

- The top three best-selling product categories were Beverages, Dairy Products and Confections, generating revenue of $0.26 M, $0.23 M and $0.16 M of sales respectively.
- Grains/Cereals and Produce were the least revenue generating categories, contributing $ 0.09 M for both.
- The top 5 and bottom 5 selling products were identified based on the Revenue generated.

ProductName	Quantity Sold	Net_Revenue
Côte de Blaye	623	$1,41,396.74
Thüringer Rostbratwurst	746	$80,368.67
Raclette Courdavault	1496	$71,155.70
Tarte au sucre	1083	$47,234.97
Camembert Pierrot	1577	$46,825.48
Gnocchi di nonna Alice	1263	$42,593.06
Manjimup Dried Apples	886	$41,819.65
Alice Mutton	978	$32,698.38
Carnarvon Tigers	539	$29,171.88
Rössle Sauerkraut	640	$25,696.64
Mozzarella di Giovanni	806	$24,900.13
Ipoh Coffee	580	$23,526.70

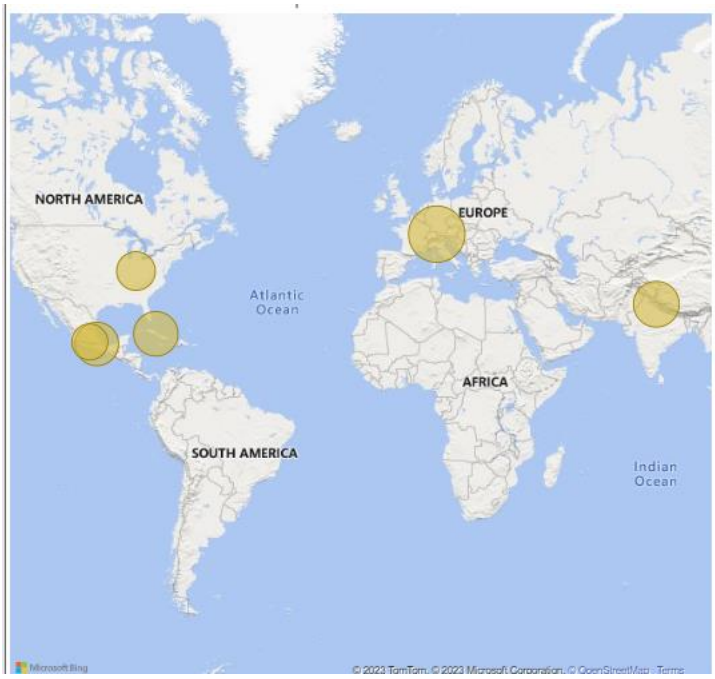






4. Geographic Analysis:

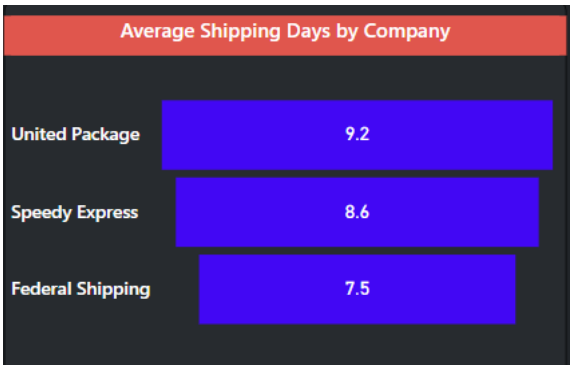
- The majority of orders originated from the United States, Germany, and Austria, accounting for 19.65%, 18.37%, and 9.53% of total net revenue respectively.
- The lowest number of orders came from Norway and Poland, each accounting less than 1% of total revenue.



ShipCountry	Sum of Net_Revenue	Sum of Gross_Revenue
USA	$2,43,618.89	$2,61,482.38
Germany	$2,27,796.66	$2,41,909.13
Austria	$1,18,104.94	$1,29,375.13
Brazil	$1,03,060.70	$1,10,745.68
France	$80,529.57	$84,396.76
UK	$58,971.31	$60,616.51
Sweden	$54,495.14	$59,523.70
Venezuela	$53,046.71	$56,962.33
Ireland	$49,979.91	$57,317.39
Canada	$48,886.79	$54,024.60
Belgium	$33,824.86	$35,134.98
Denmark	$32,428.94	$34,537.95
Switzerland	$31,194.56	$32,333.50
	$25,937.43	$27,443.76
Mexico	$23,282.08	$23,773.45
Finland	$18,810.05	$19,778.45
Spain	$17,983.20	$19,431.89
Italy	$15,363.76	$16,197.15
Portugal	$11,472.36	$12,468.65
Argentina	$7,738.10	$7,738.10
Norway	$5,735.15	$5,735.15
Poland	$3,531.95	$3,531.95

5. Shipper Analysis:

The United Package Shipping Company was the most commonly used shipper, handling approximately 40% of all orders. The Average shipping days for all the shipper companies are shown below and average shipping days was found to be 8.5 days.



CONCLUSION AND RECOMMENDATIONS:

Based on the analysis conducted, the following conclusions and recommendations can be made:

- Northwind Traders saw tremendous sales increase from 1994 to 1996, but then encountered a precipitous drop in sales by June 1996, necessitating further study
- For inventory management and marketing tactics, the best-selling product categories should be emphasized
- Additional investigation into client behaviour and preferences may yield useful insights for focused marketing strategies

- Geographic inquiry reveals that nations with smaller order quantities may have potential market expansion prospects
- Collaboration with the United Package shipment Company has been useful; seeking collaborations with other dependable shippers could improve shipment efficiency

These observations and recommendations aim to support Northwind Traders in optimizing their sales performance and strategic decision-making processes.