

EDA Using Power BI

Project Title: Retail Sales Performance Analysis

Problem Statement

A nationwide retail company operates multiple stores across four regions (North, South, East, West). The business sells electronics and office supplies to different customer segments.

Senior management has noticed that overall revenue is growing, but profit margins are inconsistent across regions and products. Some regions report high sales but low profits, while others perform steadily.

The leadership team wants a single Power BI dashboard that:

- Cleans and validates raw sales data
- Identifies profitable and non-profitable products
- Highlights regional performance gaps
- Supports data-driven decisions for promotions and inventory planning

You are assigned as a Data Analyst to explore the data, build a robust data model, apply DAX calculations, and present actionable insights.

Questions:

- Identify missing values and duplicates in the Sales dataset. What cleaning steps are required?
- Perform univariate analysis on Sales and Profit. What distribution patterns do you observe?
- Which product categories contribute the highest revenue and profit?
- How does sales performance vary across regions?
- Create a star schema using Sales, Customer, and Product tables. Why is this model effective?
- Write a DAX measure to calculate Total Sales and Total Profit.
- Calculate Profit Margin (%) using DAX and identify low-margin products.
- Design a dashboard to show regional and category-wise performance.
- Which products should be discontinued or promoted based on analysis?

Identify missing or Duplicate Sales data.

The screenshot shows the Power Query Editor interface with the following details:

- File** tab selected.
- Home**, **Transform**, **Add Column**, **View**, **Tools**, **Help** tabs visible.
- Queries [3]** pane on the left showing **Customers**, **Products**, and **Sales**.
- Properties** ribbon bar with **Advanced Editor** button highlighted.
- Transform** ribbon bar with **Data Type: Any**, **Use First Row as Headers**, **Merge Queries**, **Append Queries**, **Combine Files**, and **Combine** buttons.
- Table** view showing the data from the **Changed Type** step:

	OrderID	OrderDate	CustomerID	ProductID	Region
1	1001	01-02-2024	C004	P006	East
2	1002	28-01-2024	C003	P001	South
3	1003	15-01-2024	C002	P006	North
4	1004	04-02-2024	C027	P001	South
5	1005	10-02-2024	C026	P002	North
6	1006	01-03-2024	C030	P004	West
7	1007	24-02-2024	C012	P004	South
8	1008	20-01-2024	C005	P003	West
9	1009	17-02-2024	C020	P003	North
10	1010	04-01-2024	C017	P004	East
11	1011	07-01-2024	C006	P004	North
12	1012	01-01-2024	C020	P005	East
13	1013	14-02-2024	C030	P003	West
14	1014	13-01-2024	C017	P004	North
15	1015	21-02-2024	C014	P003	South
16	1016	18-02-2024	C015	P001	East
17	1017	16-01-2024	C010	P001	East
18	1018	13-01-2024	C009	P006	West
19	1019	18-01-2024	C001	P006	North
20	1020	22-02-2024	C015	P003	West
21	1021	21-01-2024	C029	P005	North
22	1022	28-01-2024	C026	P005	West
23	1023	20-01-2024	C003	P007	West
- Query Settings** pane on the right showing **PROPERTIES** (Name: Sales) and **APPLIED STEPS** (Removed Duplicates).

The screenshot shows the Power Query Editor interface with the 'Replace Values' dialog box open. The dialog box has the following fields:

- Value To Find:** null
- Replace With:** q

At the bottom right of the dialog box are the **OK** and **Cancel** buttons. In the background, the main Power Query Editor window displays a table with columns: Region, Quantity, UnitPrice, Discount, and Profit. The table has 23 rows of data. The 'Applied Steps' pane on the right shows the steps taken: Source, Navigation, Promoted Headers, Changed Type, and Removed Duplicates.

Screenshot of Power BI Query Editor showing a table named "Sales".

Query Settings:

- Name: Sales
- Applied Steps: Removed Duplicates

Properties:

- Name: Sales

Applied Steps:

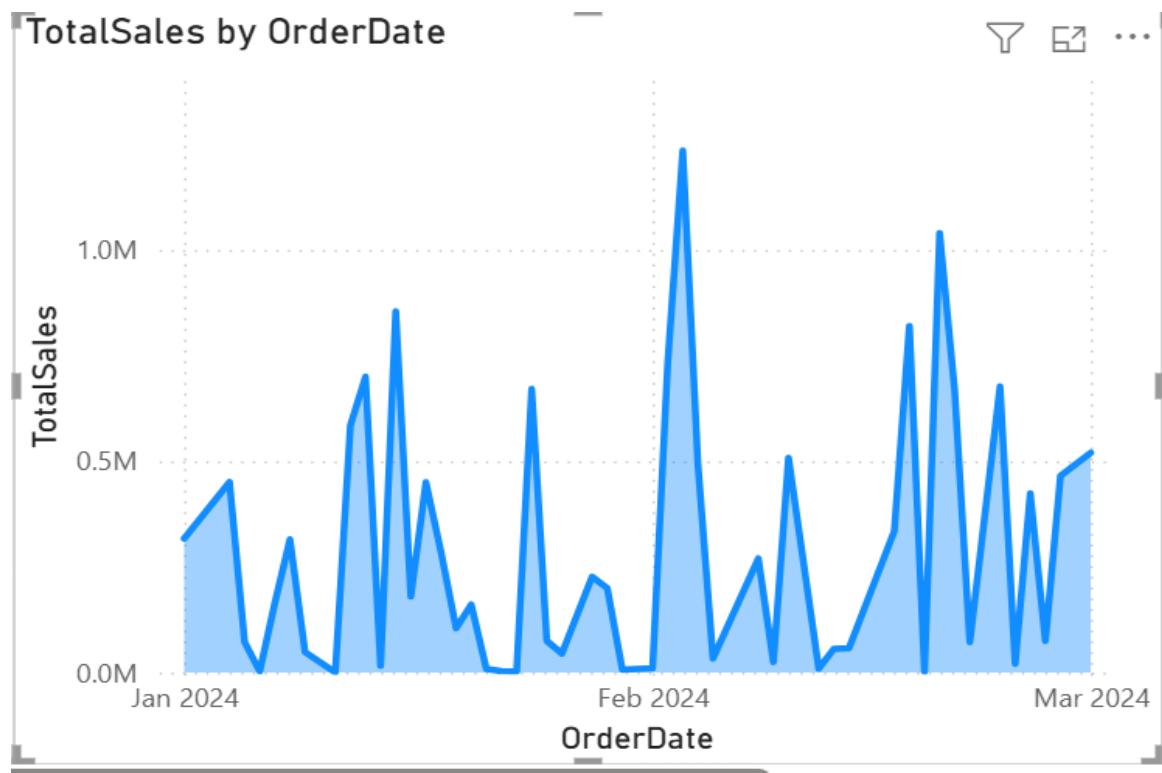
- Removed Duplicates

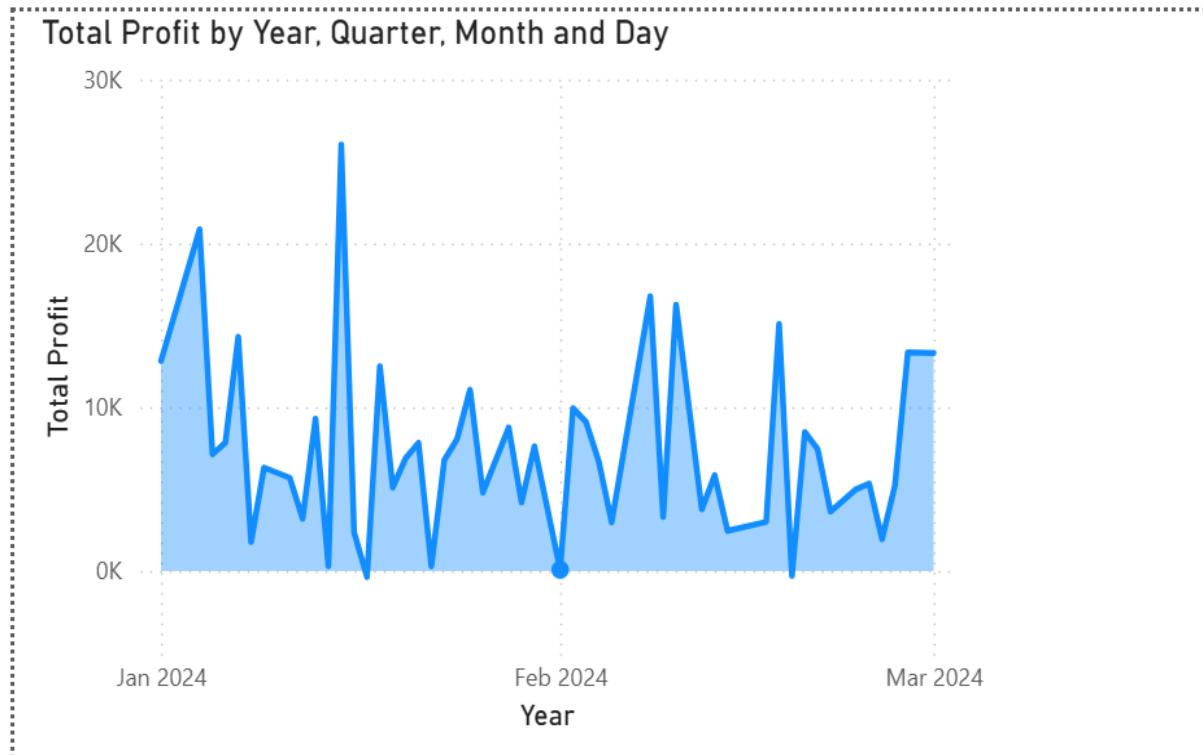
Data Preview:

OrderID	OrderDate	CustomerID	ProductID	Region
1001	01-02-2024	C004	P006	East
1002	28-01-2024	C003	P001	South
1003	15-01-2024	C002	P006	North
1004	04-02-2024	C027	P001	South
1005	10-02-2024	C026	P002	North
1006	01-03-2024	C030	P004	West
1007	24-02-2024	C012	P004	South
1008	20-01-2024	C005	P003	West
1009	17-02-2024	C020	P003	North
1010	04-01-2024	C017	P004	East
1011	07-01-2024	C006	P004	North
1012	01-01-2024	C020	P005	East
1013	14-02-2024	C030	P003	West
1014	13-01-2024	C017	P004	North
1015	21-02-2024	C014	P003	South
1016	18-02-2024	C015	P001	East
1017	16-01-2024	C010	P001	East
1018	13-01-2024	C009	P006	West
1019	18-01-2024	C001	P006	North
1020	22-02-2024	C015	P003	West

9 COLUMNS, 100 ROWS Column profiling based on top 1000 rows PREVIEW DOWNLOADED AT 16:30

Univariate Analysis on Sales and Profit





product categories contribute the highest revenue and profit

File Home Help Table tools Measure tools

Name: Total Profit Format: General Data category: Uncategorized

Home table: Sales \$ % Auto New Quick measure measure

Structure Formatting Properties Calculations

1 Total Profit = SUM(Sales[Profit])

OrderID	OrderDate	CustomerID	ProductID	Region	Quantity	UnitPrice	Discount	Profit	Sales
1001	01-02-2024	C004	P006	East	2	5000	0	76	10000
1002	28-01-2024	C003	P001	South	5	45000	0.19	6146	225000
1003	15-01-2024	C002	P006	North	7	5000	0.06	3449	35000
1004	04-02-2024	C027	P001	South	8	45000	0.24	-720	360000
1005	10-02-2024	C026	P002	North	1	1200	0.11	1578	1200
1006	01-03-2024	C030	P004	West	2	65000	0.1	5511	130000
1007	24-02-2024	C012	P004	South	4	65000	0.08	524	260000
1008	20-01-2024	C005	P003	West	7	8000	0.09	-816	56000
1009	17-02-2024	C020	P003	North	9	8000	0.04	1539	72000
1010	04-01-2024	C017	P004	East	6	65000	0.18	7561	390000
1011	07-01-2024	C006	P004	North	2	65000	0.1	7849	130000
1012	01-01-2024	C020	P005	East	2	300	0.13	4921	600
1013	14-02-2024	C030	P003	West	7	8000	0.11	-247	56000
1014	13-01-2024	C017	P004	North	10	65000	0.21	1680	650000
1015	21-02-2024	C014	P003	South	2	8000	0.1	3389	16000
1016	18-02-2024	C015	P001	East	9	45000	0.12	7505	405000
1017	16-01-2024	C010	P001	East	2	45000	0.24	-283	90000
1018	13-01-2024	C009	P006	West	10	5000	0.01	7629	50000
1019	18-01-2024	C001	P006	North	5	5000	0.09	4596	25000
1020	22-02-2024	C015	P003	West	9	8000	0.22	3634	72000
1021	21-01-2024	C029	P005	North	1	300	0.24	7254	300
1022	28-01-2024	C026	P005	West	5	300	0.08	2617	1500
1023	30-01-2024	C003	P002	West	6	1200	0.1	7604	7200

Table: Sales (100 rows) Columnn: Total Profit (0 distinct values)

File Home Help Table tools Measure tools

Name: Total Sales Format: General Data category: Uncategorized

Home table: Sales \$ % Auto New Quick measure measure Calculations

Structure Formatting Properties Calculations

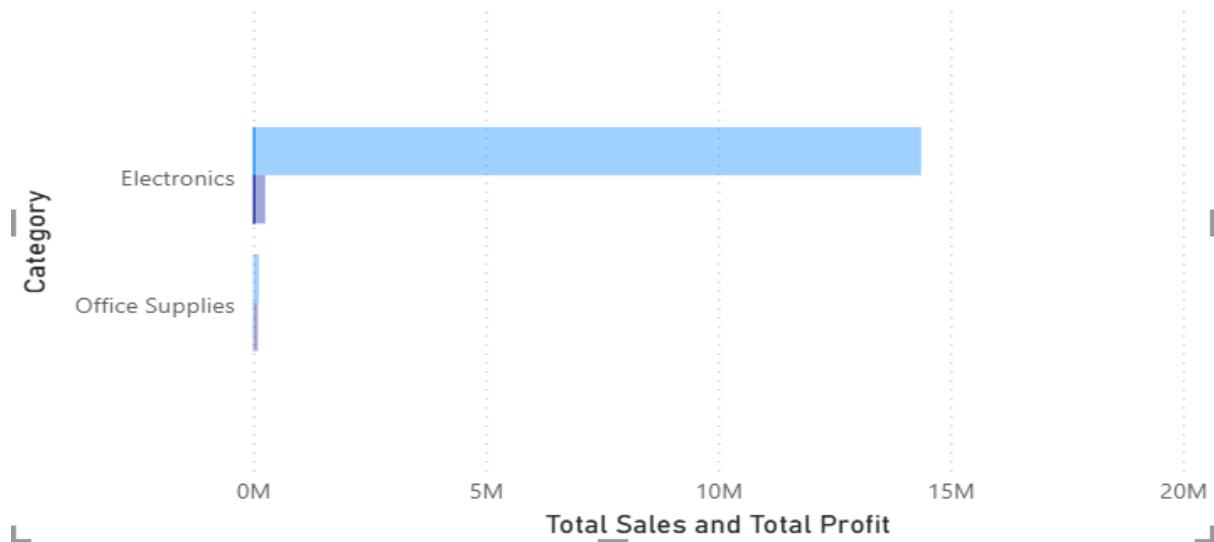
1 Total sales = SUMX(Sales, Sales[Quantity] * Sales[UnitPrice])

OrderID	OrderDate	CustomerID	ProductID	Region	Quantity	UnitPrice	Discount	Profit	Sales
1001	01-02-2024	C004	P006	East	2	5000	0	76	10000
1002	28-01-2024	C003	P001	South	5	45000	0.19	6146	225000
1003	15-01-2024	C002	P006	North	7	5000	0.06	3449	35000
1004	04-02-2024	C027	P001	South	8	45000	0.24	-720	360000
1005	10-02-2024	C026	P002	North	1	1200	0.11	1578	1200
1006	01-03-2024	C030	P004	West	2	65000	0.1	5511	130000
1007	24-02-2024	C012	P004	South	4	65000	0.08	524	260000
1008	20-01-2024	C005	P003	West	7	8000	0.09	-816	56000
1009	17-02-2024	C020	P003	North	9	8000	0.04	1539	72000
1010	04-01-2024	C017	P004	East	6	65000	0.18	7567	390000
1011	07-01-2024	C006	P004	North	2	65000	0.1	7849	130000
1012	01-01-2024	C020	P005	East	2	300	0.13	4921	600
1013	14-02-2024	C030	P003	West	7	8000	0.11	-247	56000
1014	13-01-2024	C017	P004	North	10	65000	0.21	1680	650000
1015	21-02-2024	C014	P003	South	2	8000	0.1	3389	16000
1016	18-02-2024	C015	P001	East	9	45000	0.12	7505	405000
1017	16-01-2024	C010	P001	East	2	45000	0.24	-283	90000
1018	13-01-2024	C009	P006	West	10	5000	0.01	7629	50000
1019	18-01-2024	C001	P006	North	5	5000	0.09	4596	25000
1020	22-02-2024	C015	P003	West	9	8000	0.22	3634	72000
1021	21-01-2024	C029	P005	North	1	300	0.24	7254	300
1022	28-01-2024	C026	P005	West	5	300	0.08	2617	1500
1023	30-01-2024	C003	P002	West	6	1200	0.1	7604	7200

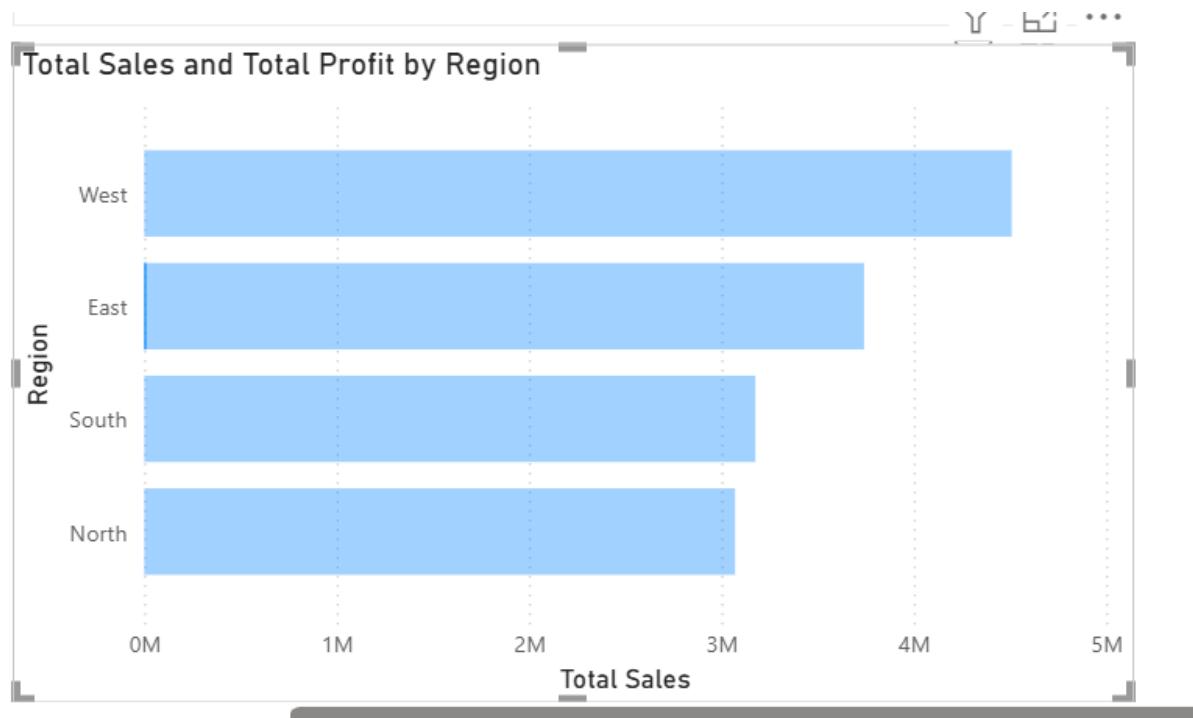
Table: Sales (100 rows) Column: Total Sales (0 distinct values)

Total Sales and Total Profit by Category

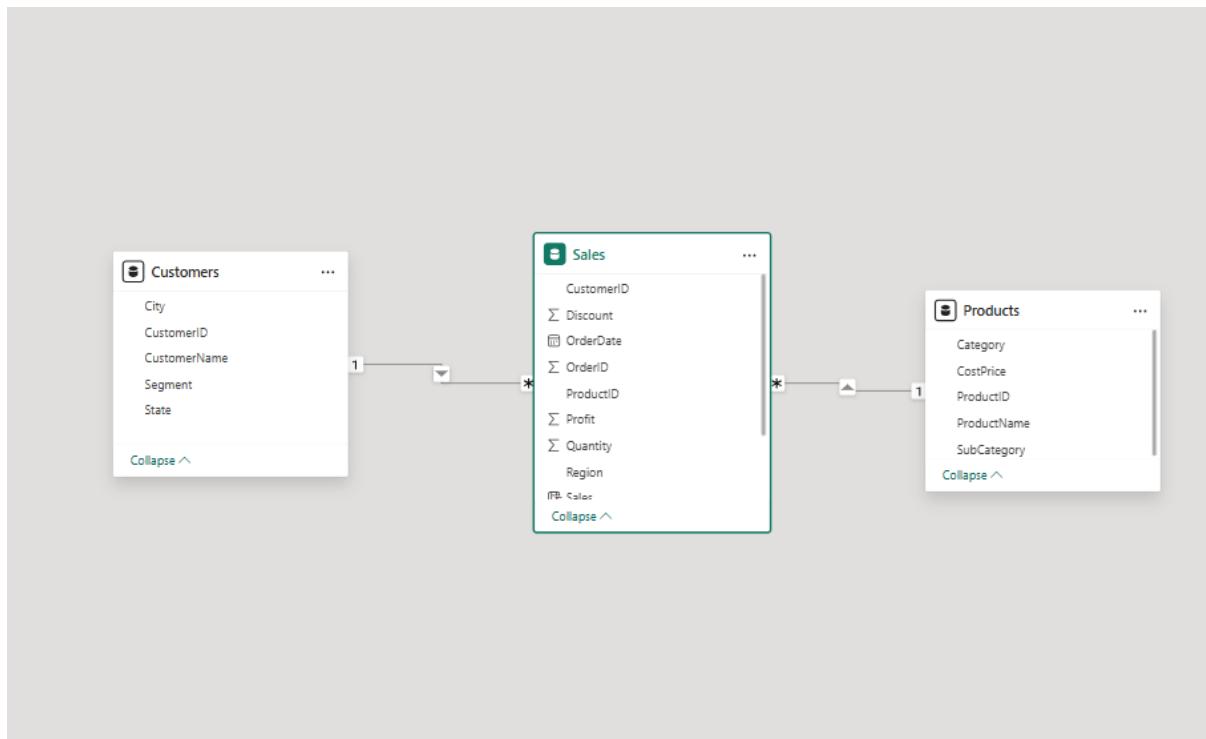
● Total Sales ● Total Profit



Sales performance vary across regions



Star Schema using Sales, Customer, and Product Tables



DAX measure to calculate Total Sales and Total Profit.

Table tools

Name: Total Sales
Home table: Sales

Format: General
Data category: Uncategorized

Structure: Calculations

1 Total Sales = SUMX(Sales, Sales[Quantity] * Sales[UnitPrice])

OrderID	OrderDate	CustomerID	ProductID	Region	Quantity	UnitPrice	Discount	Profit	Sales
1001	01-02-2024	C004	P006	East	2	5000	0	76	10000
1002	28-01-2024	C003	P001	South	5	45000	0.19	6146	225000
1003	15-01-2024	C002	P006	North	7	5000	0.06	3449	35000
1004	04-02-2024	C027	P001	South	8	45000	0.24	-720	360000
1005	10-02-2024	C026	P002	North	1	1200	0.11	1578	1200
1006	01-03-2024	C030	P004	West	2	65000	0.1	5511	130000
1007	24-02-2024	C012	P004	South	4	65000	0.08	524	260000
1008	20-01-2024	C005	P003	West	7	8000	0.09	-816	56000
1009	17-02-2024	C020	P003	North	9	8000	0.04	1539	72000
1010	04-01-2024	C017	P004	East	6	65000	0.18	7561	390000
1011	07-01-2024	C006	P004	North	2	65000	0.1	7849	130000
1012	01-01-2024	C020	P005	East	2	300	0.13	4921	600
1013	14-02-2024	C030	P003	West	7	8000	0.11	-247	56000
1014	13-01-2024	C017	P004	North	10	65000	0.21	1680	650000
1015	21-02-2024	C014	P003	South	2	8000	0.1	3389	16000
1016	18-02-2024	C015	P001	East	9	45000	0.12	7505	405000
1017	16-01-2024	C010	P001	East	2	45000	0.24	-283	90000
1018	13-01-2024	C009	P006	West	10	5000	0.01	7629	50000
1019	18-01-2024	C001	P006	North	5	5000	0.09	4596	25000
1020	22-02-2024	C015	P003	West	9	8000	0.22	3634	72000
1021	21-01-2024	C029	P005	North	1	300	0.24	7254	300
1022	28-01-2024	C026	P005	West	5	300	0.08	2617	1500
1023	30-01-2024	C003	P002	West	6	1200	0.1	7604	7200

Table: Sales (100 rows) Column: Total Sales (0 distinct values)

File Home Help Table tools Measure tools

Name: Total Profit
Home table: Sales

Format: General
Data category: Uncategorized

Structure: Calculations

1 Total Profit = SUM(Sales[Profit])

OrderID	OrderDate	CustomerID	ProductID	Region	Quantity	UnitPrice	Discount	Profit	Sales
1001	01-02-2024	C004	P006	East	2	5000	0	76	10000
1002	28-01-2024	C003	P001	South	5	45000	0.19	6146	225000
1003	15-01-2024	C002	P006	North	7	5000	0.06	3449	35000
1004	04-02-2024	C027	P001	South	8	45000	0.24	-720	360000
1005	10-02-2024	C026	P002	North	1	1200	0.11	1578	1200
1006	01-03-2024	C030	P004	West	2	65000	0.1	5511	130000
1007	24-02-2024	C012	P004	South	4	65000	0.08	524	260000
1008	20-01-2024	C005	P003	West	7	8000	0.09	-816	56000
1009	17-02-2024	C020	P003	North	9	8000	0.04	1539	72000
1010	04-01-2024	C017	P004	East	6	65000	0.18	7561	390000
1011	07-01-2024	C006	P004	North	2	65000	0.1	7849	130000
1012	01-01-2024	C020	P005	East	2	300	0.13	4921	600
1013	14-02-2024	C030	P003	West	7	8000	0.11	-247	56000
1014	13-01-2024	C017	P004	North	10	65000	0.21	1680	650000
1015	21-02-2024	C014	P003	South	2	8000	0.1	3389	16000
1016	18-02-2024	C015	P001	East	9	45000	0.12	7505	405000
1017	16-01-2024	C010	P001	East	2	45000	0.24	-283	90000
1018	13-01-2024	C009	P006	West	10	5000	0.01	7629	50000
1019	18-01-2024	C001	P006	North	5	5000	0.09	4596	25000
1020	22-02-2024	C015	P003	West	9	8000	0.22	3634	72000
1021	21-01-2024	C029	P005	North	1	300	0.24	7254	300
1022	28-01-2024	C026	P005	West	5	300	0.08	2617	1500
1023	30-01-2024	C003	P002	West	6	1200	0.1	7604	7200

Table: Sales (100 rows) Column: Total Profit (0 distinct values)

Profit Margin (%) and low-margin products

Table tools **Measure tools**

Name: Profit Margin % | Format: General | Data category: Uncategorized | New measure | Quick measure

Home table: Sales | Structure | Formatting | Properties | Calculations

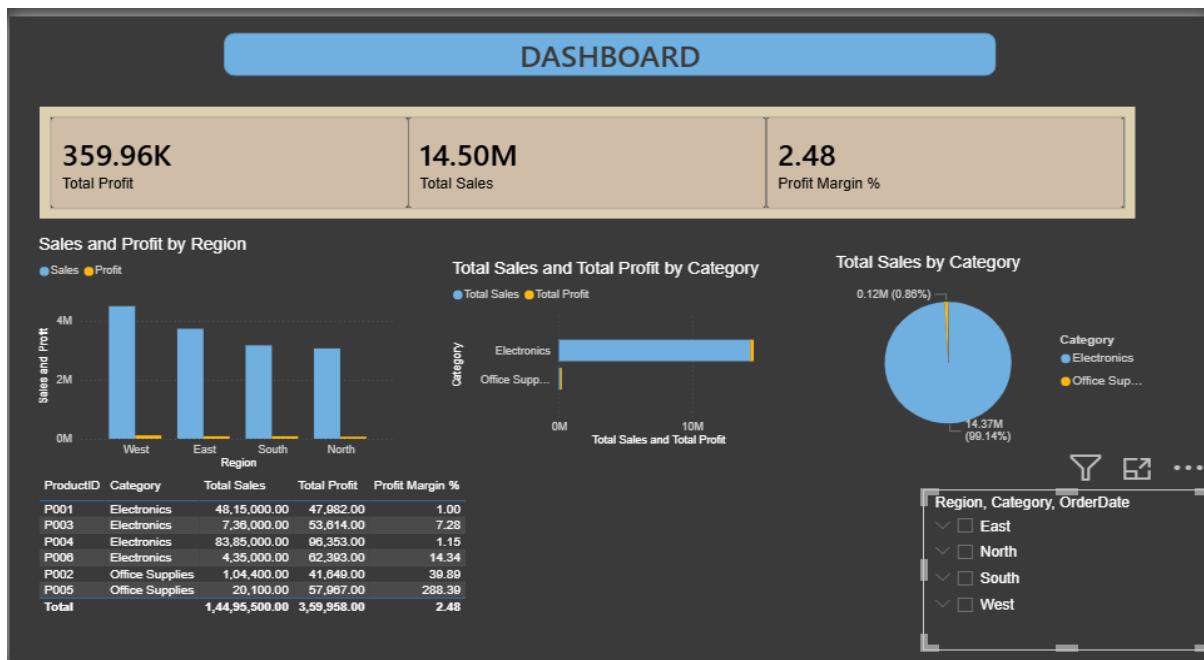
Profit Margin % =
 $\text{DIVIDE}(\text{SUM}(\text{Sales}[Profit]), \text{SUM}(\text{Sales}[Sales]), 0) * 100$

OrderID	OrderDate	CustomerID	ProductID	Region	Quantity	UnitPrice	Discount	Profit	Sales
1001	01-02-2024	C004	P006	East	2	5000	0	76	10000
1002	28-01-2024	C003	P001	South	5	45000	0.19	6146	225000
1003	15-01-2024	C002	P006	North	7	5000	0.06	3449	35000
1004	04-02-2024	C027	P001	South	8	45000	0.24	-720	360000
1005	10-02-2024	C026	P002	North	1	1200	0.11	1578	1200
1006	01-03-2024	C030	P004	West	2	65000	0.1	5511	130000
1007	24-02-2024	C012	P004	South	4	65000	0.08	524	260000
1008	20-01-2024	C005	P003	West	7	8000	0.09	-816	56000
1009	17-02-2024	C020	P003	North	9	8000	0.04	1539	72000
1010	04-01-2024	C017	P004	East	6	65000	0.18	7561	390000
1011	07-01-2024	C006	P004	North	2	65000	0.1	7849	130000
1012	01-01-2024	C020	P005	East	2	300	0.13	4921	600
1013	14-02-2024	C030	P003	West	7	8000	0.11	-247	56000
1014	13-01-2024	C017	P004	North	10	65000	0.21	1680	650000
1015	21-02-2024	C014	P003	South	2	8000	0.1	3389	16000
1016	18-02-2024	C015	P001	East	9	45000	0.12	7505	405000
1017	16-01-2024	C010	P001	East	2	45000	0.24	-283	90000
1018	13-01-2024	C009	P006	West	10	5000	0.01	7629	50000
1019	18-01-2024	C001	P006	North	5	5000	0.09	4596	25000
1020	22-02-2024	C015	P003	West	9	8000	0.22	3634	72000
1021	21-01-2024	C029	P005	North	1	300	0.24	7254	300
1022	28-01-2024	C026	P005	West	5	300	0.08	2617	1500
1023	20-01-2024	C002	P002	West	6	1200	0.1	7604	7200

Table: Sales (100 rows) Column: Profit Margin % (0 distinct values)

ProductID	Profit Margin %
P001	1.00
P004	1.15
P003	7.28
P006	14.34
P002	39.89
P005	288.39
Total	2.48

Dashboard



Discontinued or Promoted based on analysis

Products P001, P004, and P006 have very low profit margins, indicating that they contribute less to profitability. These products should be discontinued or their pricing and cost structure should be revised.

Products P002 and P005 show high profit margins and good sales performance. These products should be promoted through marketing campaigns and increased inventory to maximize business profit.