

PROJECT TITLE : RETAIL SALES PERFORMANCE ANALYSIS

Register Number : 727823TUCS104

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:

1. Identify missing values and duplicates in the Sales dataset.
What cleaning steps are required?

➤ Find missing values :

The screenshot shows the Power BI interface with three queries: Sales, Customers, and Products. The Sales query is selected. A context menu is open over the first column of the Sales table, which contains Customer IDs. The menu includes options like Sort Ascending, Sort Descending, Clear Sort, Clear Filter, Remove Empty, and Number Filters. A search bar and a list of numerical filters are also visible.

The screenshot shows the Power BI interface with the same three queries. The Sales query is selected. A context menu is open over the Quantity column of the Sales table. The menu includes Sort Ascending, Sort Descending, Clear Sort, Clear Filter, Remove Empty, and Number Filters. A search bar and a list of numerical filters are visible, with '(null)' selected.

➤ Find duplicate values :

Queries [3]

10 COLUMNS, 23 ROWS Column profiling based on top 1000 rows

	ProductID	Category	Quantity	Sales	Profit
	distinct, 1 unique	3 distinct, 1 unique	12 distinct, 7 unique	14 distinct, 7 unique	17 distinct, 13 unique
1	1	Electronics	2	40000	6000
2	2	Office Supplies	5	7500	1500
3	3	Electronics	1	30000	-2000
4	4	Office Supplies	10	12000	3000
5	5	Electronics	3	90000	18000
6	6	Office Supplies	8	null	-800
7	1	Electronics	1	20000	2000
8	2	Office Supplies	6	9000	null
9	3	Electronics	2	60000	5000
10	4	Office Supplies	12	14400	3600
11	5	Electronics	1	30000	6000
12	6	Office Supplies	10	8000	-1200
13	2	Office Supplies	4	6000	1200
14	1	Electronics	1	20000	3000
15	4	Office Supplies	9	10800	2700
16	3	Electronics	2	60000	-4000
17					

Queries [3]

10 COLUMNS, 23 ROWS Column profiling based on top 1000 rows

	ProductID	Category	Quantity	Sales	Profit
	distinct, 0 unique	2 distinct, 0 unique	11 distinct, 6 unique	14 distinct, 8 unique	17 distinct, 14 unique
1	1	Electronics	2	40000	6000
2	2	Office Supplies	5	7500	1500
3	3	Electronics	1	30000	-2000
4	4	Office Supplies	10	12000	3000
5	5	Electronics	3	90000	18000
6	6	Office Supplies	8	null	-800
7	1	Electronics	1	20000	2000
8	2	Office Supplies	6	9000	null
9	3	Electronics	2	60000	5000
10	4	Office Supplies	12	14400	3600
11	5	Electronics	1	30000	6000
12	6	Office Supplies	10	8000	-1200
13	2	Office Supplies	4	6000	1200
14	1	Electronics	1	20000	3000
15	4	Office Supplies	9	10800	2700
16	3	Electronics	2	60000	-4000
17					

➤ Cleaning Steps:

1. Remove missing values and replacing them :

Queries [3]

Sales

Customers

Products

= Table.ReplaceValue(#"Changed Type",null,0,Replacer.ReplaceValue,{"Sales"})

	ProductID	Category	Quantity	Sales	Profit
1	P001	Electronics	2	40000	60
2	P002	Office Supplies	5	7500	15
3	P003	Electronics	1	30000	-20
4	P004	Office Supplies	10	12000	30
5	P005	Electronics	3	90000	180
6	P006	Office Supplies	8	0	-8
7	P001	Electronics	1	20000	20
8	P002	Office Supplies	6	9000	n
9	P003	Electronics	2	60000	50
10	P004	Office Supplies	12	14400	36
11	P005	Electronics	1	30000	60
12	P006	Office Supplies	10	8000	-12
13	P002	Office Supplies	4	6000	12
14	P001	Electronics	1	20000	30
15	P004	Office Supplies	9	10800	27
16	P003	Electronics	2	60000	-40
17	P006	Office Supplies	7	5600	7
18	P005	Electronics	2	60000	120
19	P002	Office Supplies	5	7500	15
20	P001	Electronics	3	60000	90
21					

10 COLUMNS, 23 ROWS Column profiling based on top 1000 rows PREVIEW DOWNLOADED AT 11:34

Query Settings

PROPERTIES

Name: Sales

All Properties

APPLIED STEPS

Source: Sales
Promoted Headers:
Changed Type:
Replaced Value:

Queries [3]

Sales

Customers

Products

= Table.ReplaceValue(#"Replaced Value",null,0,Replacer.ReplaceValue,{"Profit"})

	ProductID	Category	Quantity	Sales	Profit
1	1	Electronics	2	40000	6000
2	2	Office Supplies	5	7500	1500
3	3	Electronics	1	30000	-2000
4	4	Office Supplies	10	12000	3000
5	5	Electronics	3	90000	18000
6	6	Office Supplies	8	0	-800
7	1	Electronics	1	20000	2000
8	2	Office Supplies	6	9000	0
9	3	Electronics	2	60000	5000
10	4	Office Supplies	12	14400	3600
11	5	Electronics	1	30000	6000
12	6	Office Supplies	10	8000	-1200
13	2	Office Supplies	4	6000	1200
14	1	Electronics	1	20000	3000
15	4	Office Supplies	9	10800	2700
16	3	Electronics	2	60000	-4000
17	6	Office Supplies	7	5600	700
18	5	Electronics	2	60000	12000
19	2	Office Supplies	5	7500	1500
20	1	Electronics	3	60000	9000
21					

10 COLUMNS, 23 ROWS Column profiling based on top 1000 rows PREVIEW DOWNLOADED AT 11:35

Query Settings

PROPERTIES

Name: Sales

All Properties

APPLIED STEPS

Source: Sales
Promoted Headers:
Changed Type:
Replaced Value:
Replaced Value1:

2. Remove duplicate values :

Queries [3]

Sales
Customers
Products

= Table.Distinct(#"Replaced Value1", {"OrderID"})

1	2	3	OrderID	OrderDate	Avg_Region	Avg_StoreID	Avg_CustomerID
1	2	3	2001	05-01-2024	North	N01	C001
2	2	3	2002	07-01-2024	South	S01	C002
3	2	3	2003	10-01-2024	East	E01	C003
4	2	3	2004	15-01-2024	West	W01	C004
5	2	3	2005	18-01-2024	North	N02	C005
6	2	3	2006	20-01-2024	South	S02	C006
7	2	3	2007	02-02-2024	East	E02	C001
8	2	3	2008	05-02-2024	West	W02	C002
9	2	3	2009	10-02-2024	North	N01	C003
10	2	3	2010	12-02-2024	South	S01	C004
11	2	3	2011	15-02-2024	East	E01	C005
12	2	3	2012	18-02-2024	West	W01	C006
13	2	3	2013	20-02-2024	North	N02	C001
14	2	3	2014	22-02-2024	South	S02	C002
15	2	3	2015	25-02-2024	East	E02	C003
16	2	3	2016	28-02-2024	West	W02	C004
17	2	3	2017	03-03-2024	North	N01	C005
18	2	3	2018	05-03-2024	South	S01	C006
19	2	3	2019	08-03-2024	East	E01	C001
20	2	3	2020	10-03-2024	West	W01	C002
21	2	3					

10 COLUMNS, 21 ROWS Column profiling based on top 1000 rows

Query Settings

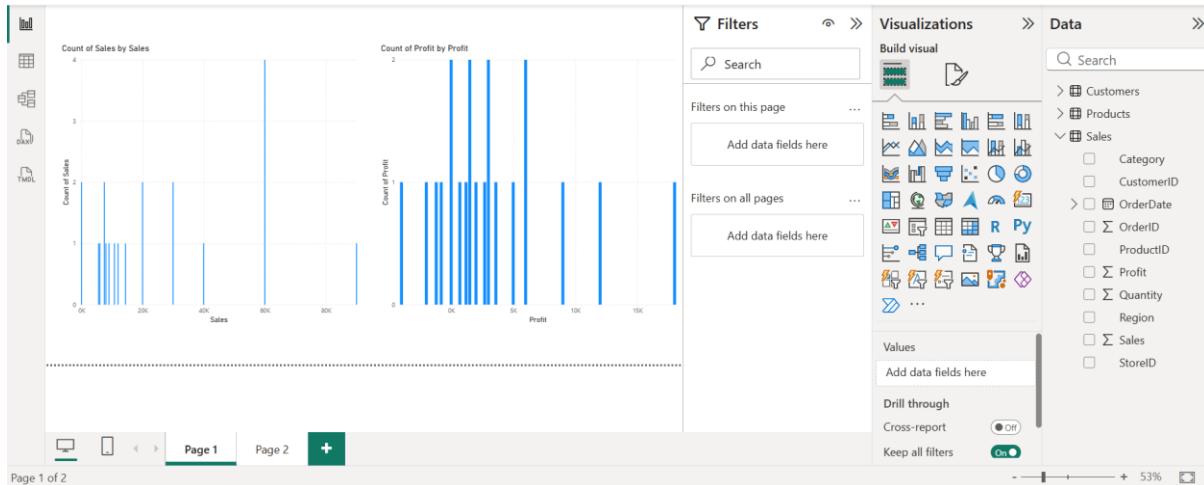
Properties
Name: Sales
All Properties

Applied Steps
Source
Promoted Headers
Changed Type
Replaced Value
Replaced Value1
X Removed Duplicates

PREVIEW DOWNLOADED AT 11:36

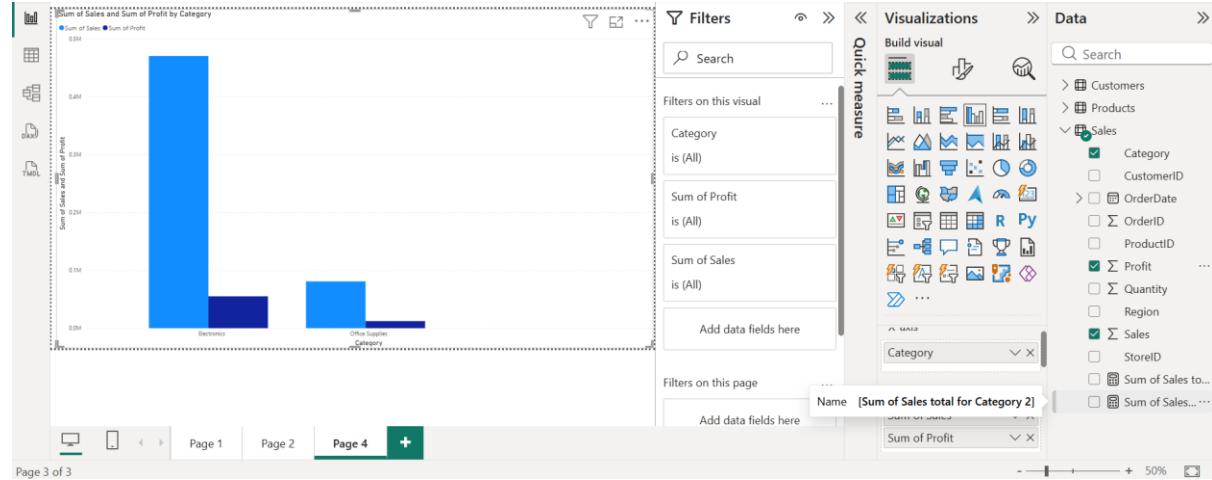
Final Output : Missing values were found in Sales and Profit columns and replaced with zero.
Duplicate records based on OrderID were removed.

2. Perform univariate analysis on Sales and Profit. What distribution patterns do you observe?



Final Output : Sales data is right-skewed, indicating that a few high-value transactions contribute most of the revenue.
Profit distribution shows variability with both positive and negative values, indicating inconsistent profitability across transactions.

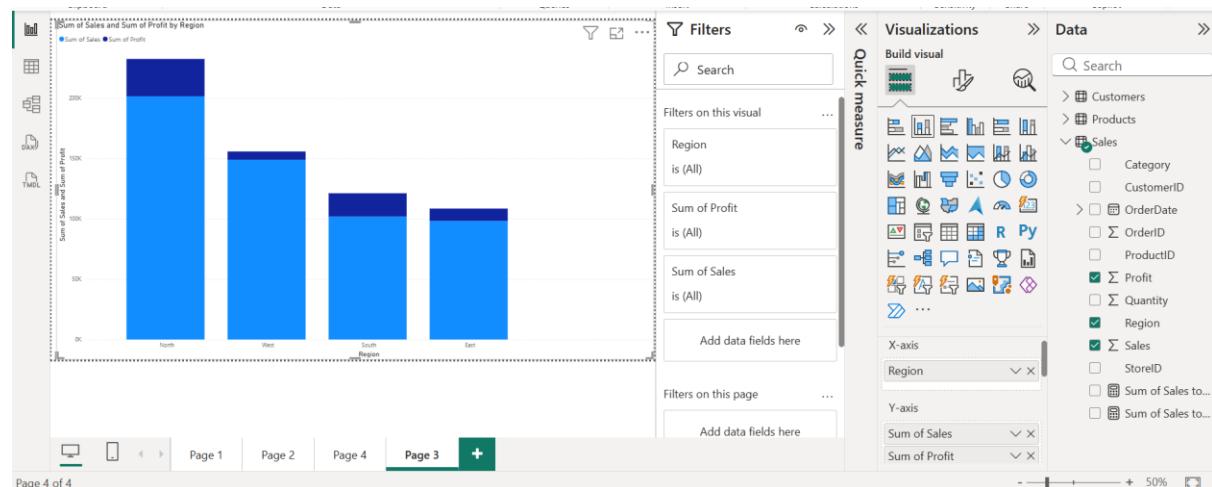
3. Which product categories contribute the highest revenue and profit?



Final Output : Electronics contributes the highest revenue and profit.

Office Supplies generates lower revenue and has comparatively lower profit margins.

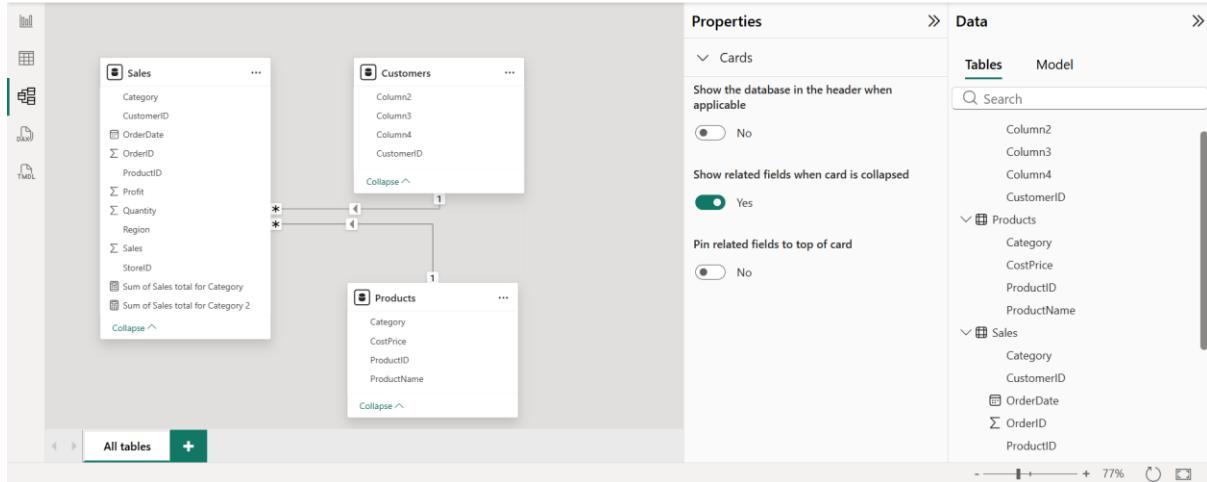
4. How does sales performance vary across regions?



Final Output : The North region records the highest sales and profit.

South shows stable performance, while East and West experience fluctuating and lower profitability.

5. Create a star schema using Sales, Customer, and Product tables. Why is this model effective?



Final Output : Star schema improves query performance, simplifies data analysis, and allows easy slicing of sales data by customer and product attributes.

6. Write a DAX measure to calculate Total Sales and Total Profit.

The screenshot shows the Power BI Data View. A table is displayed with columns: CustomerID, Column2, Column3, Column4, Total Sales, and Total Profit. The data rows are: C001 (John Retail, Consumer, Delhi, 550800, 67200), C002 (ABC Corp, Corporate, Mumbai, 550800, 67200), C003 (XYZ Ltd, Corporate, Bangalore, 550800, 67200), C004 (OfficeMart, Small Business, Chennai, 550800, 67200), C005 (QuickBuy, Consumer, Hyderabad, 550800, 67200), and C006 (TechZone, Small Business, Pune, 550800, 67200). The Total Sales column is defined by the DAX measure `Total Sales = SUM(Sales[Sales])`. The Data pane on the right shows the structure of the Sales table, including columns for Category, CustomerID, OrderDate, ProductID, Profit, Quantity, and Region. The Total Sales measure is highlighted in the pane.

1 Total Profit = SUM(Sales[Profit])

CustomerID	Column2	Column3	Column4	Total Sales	Total Profit
C001	John Retail	Consumer	Delhi	550800	67200
C002	ABC Corp	Corporate	Mumbai	550800	67200
C003	XYZ Ltd	Corporate	Bangalore	550800	67200
C004	OfficeMart	Small Business	Chennai	550800	67200
C005	QuickBuy	Consumer	Hyderabad	550800	67200
C006	TechZone	Small Business	Pune	550800	67200

Table: Customers (7 rows) Column: Total Profit (1 distinct values)

Final Output : DAX measures were created using SUM() to calculate total sales and total profit from the Sales table.

7. Calculate Profit Margin (%) using DAX and identify low-margin products.

1 Profit Margin % =
2 DIVIDE([Total Profit], [Total Sales], 0) * 100

CustomerID	Column2	Column3	Column4	Total Sales	Total Profit	Profit Margin %
C001	John Retail	Consumer	Delhi	550800	67200	12.2004357298475
C002	ABC Corp	Corporate	Mumbai	550800	67200	12.2004357298475
C003	XYZ Ltd	Corporate	Bangalore	550800	67200	12.2004357298475
C004	OfficeMart	Small Business	Chennai	550800	67200	12.2004357298475
C005	QuickBuy	Consumer	Hyderabad	550800	67200	12.2004357298475
C006	TechZone	Small Business	Pune	550800	67200	12.2004357298475

Table: Customers (7 rows) Column: Profit Margin % (1 distinct values)

The screenshot shows the Power BI Report view interface. On the left, there is a table visual displaying product information: Laptop, Office Chair, Paper Pack, Printer, Smartphone, and Tablet, along with their Total Sales, Total Profit, and Profit Margin %. In the center, the 'Filters' pane is open, showing a dropdown for 'Profit Margin %' with options '(All)', '(Blank)', and '12.20'. A checkbox for 'Require single selection' is checked. Below the filters, there is a 'Quick measure' section with 'Total Profit'. On the right, the 'Visualizations' and 'Data' panes are visible, showing various chart and table icons under 'Build visual', and a detailed list of data fields under 'Customers', 'Products', and 'Sales'.

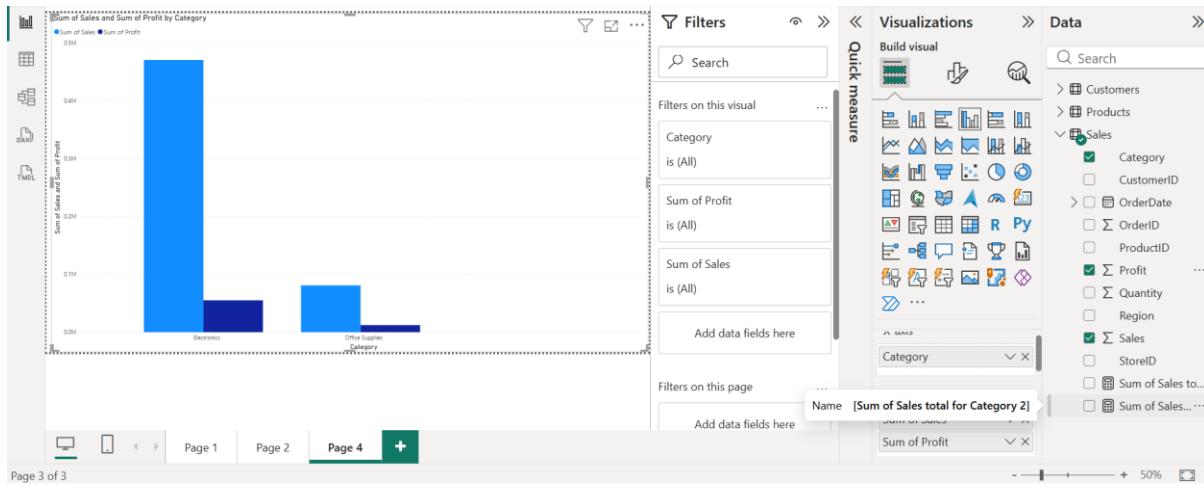
Final Output : Profit Margin (%) was calculated using DAX by dividing total profit by total sales. A table visual was used to identify low-margin products by sorting profit margin in ascending order.

8. Design a dashboard to show regional and category-wise performance.

The screenshot shows a Power BI dashboard with three main components: 1) Three KPI cards at the top displaying '3.86M Sum of Total Sales', '470.40K Sum of Total Profit', and '85.40 Sum of Profit Margin %'. 2) A 'Category-wise clustered column chart' below the first card, comparing Sales and Profit for Electronics and Office Supplies categories. 3) A 'Region-wise stacked column chart' below the second card, comparing Sales and Profit across North, West, South, and East regions. The dashboard also includes filter panes for 'Filters on this page' and 'Filters on all pages', and a 'Values' section. The 'Data' pane on the right lists various data fields and measures.

Final Output : A Power BI dashboard was created using KPI cards and charts to visualize sales and profit across regions and product categories, enabling effective performance comparison.

9. Which products should be discontinued or promoted based on analysis?



Final Output : Based on profit margin analysis, low or loss-making products should be discontinued or repriced, while high-margin electronics should be promoted to maximize profitability.

PROJECT CONCLUSION

- In This Project, Retail Sales Data Was Successfully Cleaned, Validated, And Analyzed Using Power Bi To Understand Revenue And Profitability Patterns Across Regions And Product Categories. Missing Values And Duplicate Records Were Handled To Ensure Data Accuracy, And A Star Schema Data Model Was Created To Improve Analytical Efficiency.
- Dax Measures Were Used To Calculate Total Sales, Total Profit, And Profit Margin, Enabling Clear Identification Of High-Performing And Low-Performing Products. The Analysis Revealed That The Electronics Category Contributes The Highest Revenue And Profit, While Certain Office Supplies Products Show Low Or Negative Profit Margins. Regional Analysis Highlighted Strong Performance In The North Region, With Other Regions Showing Inconsistent Profitability.
- An Interactive Dashboard Was Designed To Present Key Insights Through Kpi Cards, Category-Wise And Region-Wise Visualizations, Supporting Data-Driven Decision-Making. Based On The Findings, High-Margin Products Such As Laptops And Smartphones Should Be Promoted, While Low-Margin Products Like Tablets And Paper Packs Require Review Or Discontinuation.
- Overall, This Dashboard Provides Management With A Clear And Actionable View Of Business Performance, Helping Optimize Promotions, Inventory Planning, And Profitability.