Reports on Mini Project

Name : Manoj Kumar S

Batch Code : DA_TN_FNB07

Contact no : 7397463066

Email ID : Manoj161326@gmail.com

Project Title : Sales & Profit Analysis Dashboard (2023-2025)

Project Domin : Sales & E-commerce

Submission Date : 10/12/2025

Mentor name : Kumaran M

Raw dataset link : https://github.com/Manoj-1613/Mini_Project/blob/main/Snitch_Fashion_Sales_Uncleaned.csv

clean dataset link : https://github.com/Manoj-1613/Mini_Project/blob/main/Snitch_Fashion_Sales_Cleaned_Dataset.xlsx

Purpose of the Analysis :

The purpose of this Sales & Profit Analysis is to understand overall business performance across different products, customer segments, and cities. The dataset includes order details such as unit price, quantity sold, discounts, sales amounts, and profit. By analyzing these variables, the goal is to identify high-performing products, profitable customer segments, and key markets that contribute most to total sales. This analysis also helps uncover low-profit or loss-making areas, enabling better decision-making for pricing, marketing, and inventory planning.

Objective of the Report :

The main objective of this report is to convert raw sales data into meaningful insights through data cleaning, data modeling, and interactive visualizations. This includes calculating total sales, total profit, discount impact, order counts, and city-wise performance. The report aims to highlight top-selling products, profit contribution by customer segments (B2B & B2C), and monthly sales trends. The objective is to support managers and business teams with a clear understanding of sales growth patterns and profitability to make strategic business decisions.

Step - 1

Data Cleaning into Data Set:

This is my Raw data set, first Step Changed file to table format **Ctrl + T**

Order_ID	Customer_Name	Product_Category	Product_Name	Units_Sold	Unit_Price	Discount_%	Sales_Amount	Order_Date	City	Segment	Profit
1000	Brian Thompson	Jeans	Slim Fit Jeans		842	0.6	0	27-02-2025	Delhi	B2C	2137.45
1001	Shaun Ross	Jeans	Slim Fit Jeans	1			0	15-07-2025	Ahmedaba	B2B	1588.15
1002	Sarah Snyder	Jackets	Puffer Coat	1	637.82		0	02-01-2025	Mumbai	B2B	-158.03
1003	Jay Briggs	Shoes	Loafers	2	2962.27		0	18-06-2025	Bangalore	B2B	2296.5
1004	Maria Blake	Accessories	Belts	1	2881.07	0.27	2103.18	28-09-2024	Hyderabac	B2B	63.66
1005	Samuel Miller	T-Shirts	Crop Top				0	05-12-2023	Mumbai	B2B	1477.73
1006	John Mclean	Jeans	Boyfriend Jeans	1	2060.85		0	02-04-2024	Ahmedaba	B2B	198.36
1007	Dr. Kathryn Bass PhD	Dresses	Casual Midi		3669.56		0	28-09-2024	Bangalore	B2C	2490.01
1008	Marie Fisher	Accessories	Sunglasses				0	28-09-2024	Ahmedaba	B2C	-886.8
1009	Dale Perry	T-Shirts	Oversized T-shirt	5			0	20-07-2025	Bangalore	B2B	-972.73
1010	Oscar Turner	T-Shirts	Crop Top	0	4007.02		0	28-09-2024	Pune	B2B	-950.38
1011	Jacob Sanchez	Accessories	Belts	0			0	27-05-2025	Delhi	B2C	1370.67
1012	Valerie Hampton	Jeans	Straight Cut	4		0.21	0	22-08-2024	Ahmedaba	B2B	843.61
1013	James Delgado	T-Shirts	Crop Top			0.62	0	28-09-2024	Pune	B2C	1279.58
1014	Richard Hood	Shoes	Loafers	4			0	21-09-2024	Hyderabac	B2B	-889.25
1015	Julie Fox	Shoes	Sneakers	6			0	28-09-2024	Ahmedaba	B2B	793.8
1016	Valerie Garner	T-Shirts	Crop Top				0	26-10-2024	Hyderabac	B2B	1881.89
1017	Alejandro Johnson	T-Shirts	Oversized T-shirt	3	468.53		0	28-09-2024	Mumbai	B2B	-195.65
1018	Andrew Washington	Dresses	Maxi Dress	4			0	13-04-2024	Bangalore	B2C	2102.56
1019	Mark Allison	Dresses	Casual Midi		4064.4	0.7	0	22-10-2024	Ahmedaba	B2C	877.21
1020	Russell Eaton	Jeans	Boyfriend Jeans	1			0	19-03-2025	Pune	B2B	669.75
1021	Amber Bowman	Dresses	Wrap Dress				0	05-05-2024	Bangalore	B2B	-705.72
1031	Lauren Jensen	Dresses	Wrap Dress		1872.05	0.15	0	27-05-2025	Pune	B2C	-863.9
1023	David Arnold	Shoes	Slip-ons				0	28-09-2024	Mumbai	B2B	-875.51
1024	Thomas Long	Jackets	Puffer Coat	-2		1.19	0	19-02-2025	Delhi	B2B	491.1
1025	Christine Reed	Jeans	Straight Cut	4	4595.08		0	18-07-2025	Bangalore	B2B	330.82

Step – 2 City

Problem Identifying :

Data set look like inconsistencies format such as missing values, incorrect date formats & duplicate data repeated.

Action Taken :

I performed comprehensive data cleaning by correcting date formats, removing negative or incorrect values, and filling or recalculating missing fields.

Steps Followed :

Find and replacement use change to inconsistence values (Ex: hyd to Hyderabad)

Before action

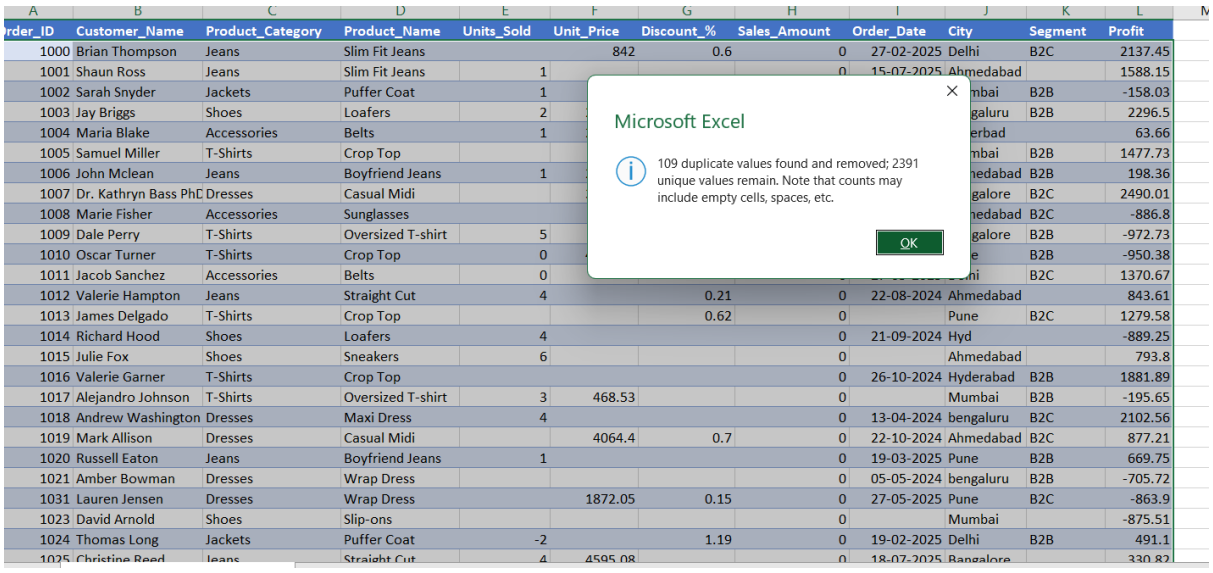
City
bengaluru
Bangalore
Bangalore
Hyd
Hyderabad
bengaluru
bengaluru
Bangalore
Hyd
Bangalore
Hyd
bengaluru
Hyd
bengaluru
Hyd
Hyd
Bangalore
bengaluru
bengaluru
bengaluru
bengaluru
Hyd
Bangalore
Hyderabad
Bangalore

Action Taken

City
Bangalore
Hyderabad
Bangalore
Bangalore
Hyderabad
Hyderabad
Bangalore
Bangalore
Bangalore
Hyderabad
Hyderabad
Hyderabad
Hyderabad
Bangalore
Hyderabad
Bangalore
Hyderabad
Hyderabad
Bangalore
Bangalore
Bangalore
Bangalore
Bangalore
Hyderabad
Bangalore

Step - 3 Remove Duplicate value

- 109 duplicate value remove has per Oder ID (Unique value)
- Navigate -> Data Ribbon -> Remove duplicate



rder_ID	Customer_Name	Product_Category	Product_Name	Units_Sold	Unit_Price	Discount_%	Sales_Amount	Order_Date	City	Segment	Profit
1000	Brian Thompson	Jeans	Slim Fit Jeans		842	0.6	0	27-02-2025	Delhi	B2C	2137.45
1001	Shaun Ross	Jeans	Slim Fit Jeans	1			0	15-07-2025	Ahmedabad		1588.15
1002	Sarah Snyder	Jackets	Puffer Coat	1					mbai	B2B	-158.03
1003	Jay Briggs	Shoes	Loafers	2					galuru	B2B	2296.5
1004	Maria Blake	Accessories	Belts	1					erbad		63.66
1005	Samuel Miller	T-Shirts	Crop Top						mbai	B2B	1477.73
1006	John Mclean	Jeans	Boyfriend Jeans	1					nedabad	B2B	198.36
1007	Dr. Kathryn Bass PhD	Dresses	Casual Midi						galore	B2C	2490.01
1008	Marie Fisher	Accessories	Sunglasses						nedabad	B2C	-886.8
1009	Dale Perry	T-Shirts	Oversized T-shirt	5					galore	B2B	-972.73
1010	Oscar Turner	T-Shirts	Crop Top	0					e	B2B	-950.38
1011	Jacob Sanchez	Accessories	Belts	0					hi	B2C	1370.67
1012	Valerie Hampton	Jeans	Straight Cut	4		0.21	0	22-08-2024	Ahmedabad		843.61
1013	James Delgado	T-Shirts	Crop Top			0.62	0		Pune	B2C	1279.58
1014	Richard Hood	Shoes	Loafers	4			0	21-09-2024	Hyd		-889.25
1015	Julie Fox	Shoes	Sneakers	6			0		Ahmedabad		793.8
1016	Valerie Garner	T-Shirts	Crop Top				0	26-10-2024	Hyderabad	B2B	1881.89
1017	Alejandro Johnson	T-Shirts	Oversized T-shirt	3	468.53		0		Mumbai	B2B	-195.65
1018	Andrew Washington	Dresses	Maxi Dress	4			0	13-04-2024	bengaluru	B2C	2102.56
1019	Mark Allison	Dresses	Casual Midi		4064.4	0.7	0	22-10-2024	Ahmedabad	B2C	877.21
1020	Russell Eaton	Jeans	Boyfriend Jeans	1			0	19-03-2025	Pune	B2B	669.75
1021	Amber Bowman	Dresses	Wrap Dress				0	05-05-2024	bengaluru	B2B	-705.72
1031	Lauren Jensen	Dresses	Wrap Dress		1872.05	0.15	0	27-05-2025	Pune	B2C	-863.9
1023	David Arnold	Shoes	Slip-ons				0		Mumbai		-875.51
1024	Thomas Long	Jackets	Puffer Coat	-2		1.19	0	19-02-2025	Delhi	B2B	491.1
1025	Christine Reed	Jeans	Straight Cut	4	4595.08		0	18-07-2025	Bangalore		330.82

Step – 4 Customer name

- Customer name Split to 3 type Title, First name & Last name. I used tool text to column.
- Navigate -> Data Ribbon -> Text to column

Original Data

Customer_Name
Brian Thompson
Shaun Ross
Sarah Snyder
Jay Briggs
Maria Blake
Samuel Miller
John Mclean
Dr. Kathryn Bass PhD
Marie Fisher
Dale Perry
Oscar Turner
Jacob Sanchez

Standardizing Data

Title	First_Name	Last_Name
Dr	Kathryn	Bass PhD
Dr	Derrick	Scott
Mr	Gary	Wilson
Mr	Lawrence	Parks
Mrs	Denise	Solis DDS
Mr	Nathaniel	Cox
Ms	Brianna	Jones
Mr	Daniel	Knox
Dr	Meredith	Murphy
Mr	John	Carroll Jr
Mr	Robert	Washington
Mr	David	Green MD
Mr	Tyrone	Kaiser

Step – 5 Segment

Problem Identifying :

In segment column, most of the value cells are missing. So we need to fill missing values.

Action Taken :

I should convert entire table to the Power Query and check with Group By, Count rows, In Segment Base, which one is most higher values. I updated this Missing Values.

Original Data

Segment
B2C
B2B
B2B
B2B
B2B
B2C
B2C
B2B
B2B
B2C
B2C
B2B
B2B

Standardizing Data

Segment
B2C
B2B
B2B
B2B
B2B
B2B
B2B
B2C
B2C
B2B
B2B
B2C
B2B
B2C
B2B
B2B
B2B
B2B

Step – 6 Unit Sold

Problem Identifying :

Unit sold column most of value missing, we need to fill missing values.

Action Taken :

For records where recalculation was not possible, I replaced the missing Unit_Sold with “0”

Formula : **=Abs(E3)**

Original Data	Standardizing Data
Units_Sold	Unit_Sold
	0
1	1
1	1
2	2
1	1
	0
1	1
	0
	0
5	5
0	0
0	0
4	4
	0
4	4
6	6
	0
3	3
4	4
	0
1	1
	0

Step – 7 Unit Price

Problem Identifying :

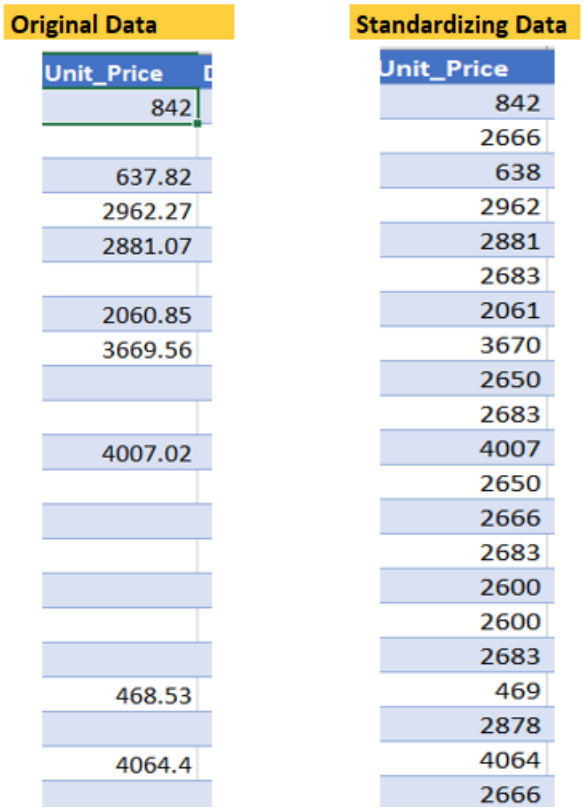
Unit Price column most of value missing, we need to fill missing values.

Action Taken :

Unit price missing, so we calculate by taking an average. product weighs each and every product.

Formula :

=IF(ISBLANK(F2),AVERAGEIF($C$2:$C$2392,C2,$F$2:$F$2392),F2)



Step – 8 Total sales amount

Problem Identifying:

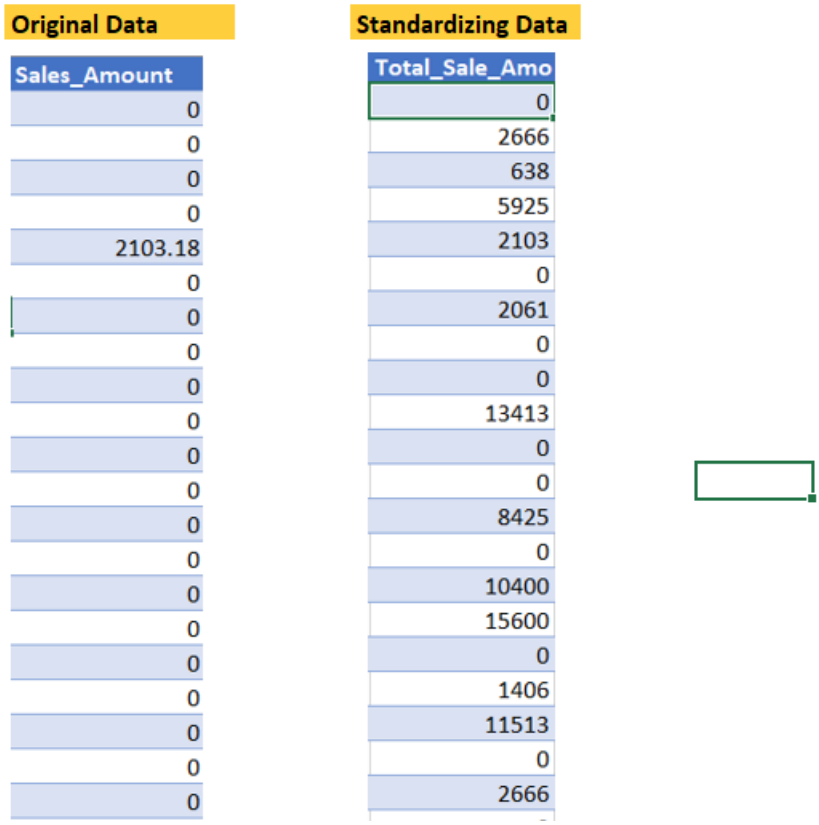
Sale Price column most of value missing, we need to fill missing values.

Action Taken:

Checked for missing or incorrect Sales Amount values using Excel filters and data-quality checks. This helped identify records where Sales Amount was blank, or not matching Unit Sold × Unit Price x (1- Discount).

Formula :

Total sale amount = Unit_Sold × Unit_Price × (1 − Discount)



Original Data
Sales_Amount
0
0
0
0
2103.18
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0

Standardizing Data
Total_Sale_Amo
0
2666
638
5925
2103
0
2061
0
0
13413
0
0
8425
0
10400
15600
0
1406
11513
0
2666

Step – 9 Order Date

Problem Identifying:

Order date column most of value missing, we need to fill missing values.

Action Taken:

Select entrie table change to power query used group by count rows high repat value update on missing cells.

Original Data

Order_Date	
27-02-2025	
15-07-2025	
02-01-2025	
18-06-2025	
05-12-2023	
02-04-2024	
20-07-2025	
27-05-2025	
22-08-2024	
21-09-2024	
26-10-2024	
13-04-2024	
22-10-2024	

Standardizing Data

Order_Date
27-02-2025
15-07-2025
02-01-2025
18-06-2025
28-09-2024
05-12-2023
02-04-2024
28-09-2024
28-09-2024
20-07-2025
28-09-2024
27-05-2025
22-08-2024
28-09-2024
21-09-2024
28-09-2024
26-10-2024
28-09-2024
13-04-2024
22-10-2024

Step – 10 New column Product sale and Return

Problem Identifying:

Unit sold column some cells values reflected as negative.

Action taken:

We consider negative value mean product should be return to store.so we calculate the how many products sale and return. We add new column and measure.

New Column Added Formula used : R2 mean Unit sold value **= IF(R2<0,"Return","Sale")**

Product Return/ Sale
-
Sale
Sale
Return
Sale
Sale
-
-
-
Return
Return
Sale
Sale
-
-
-
-
-
-
-
Return
Sale
Sale

Step – 11 Profit calculation

Problem Identifying:

- Profit column most of amount missing and incorrect values.

Action taken:

- We decided to estimate amount of profit because we don't have cost price column.
- So sale price * 20% margin we assuming. Final amount we get profit.

Formula used : **= sale amount * 0.20 (Assuming 20% profit)**

Original Data

Profit
0
533
128
-1073
2146
642
0
0
0
-32
-1089
-106
1015
0
0
0
0
0
0
0
-39

Standardizing Data

Profit
0
533
128
1185
421
0
412
0
0
2683
0
0
1685
0
2080
3120
0
281
2303
0
533

Step – 12 Add Month and Year

- We prepare for sale dash board so we must require Month and year.
- So we split (MM & YY) wise it’s better view for sale and profit

Formula used : =Month (value) and =Year (value)

Original Data

Order_Date
27-02-2025
15-07-2025
02-01-2025
18-06-2025
05-12-2023
02-04-2024
20-07-2025
27-05-2025
22-08-2024
21-09-2024
26-10-2024
13-04-2024
22-10-2024

Standardizing Data

Month	Year
02	2025
07	2025
01	2025
06	2025
09	2024
12	2023
04	2024
09	2024
09	2024
07	2025
09	2024
05	2025
08	2024
09	2024
09	2024
09	2024
10	2024
09	2024
04	2024
10	2024

Data Visualization in Power BI

After cleaning the dataset, the next step was to convert the raw numbers into meaningful visual insights. Power BI allowed me to represent sales trends, profit performance, and product distribution using easy-to-understand charts and graphs.

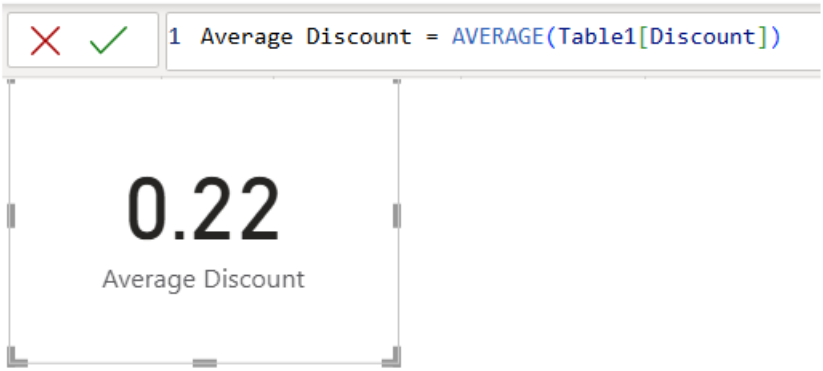
Dax Measure

Total sale Measure

It calculates the overall revenue The Total Sales measure sums up the sales amount for all transactions in the dataset.

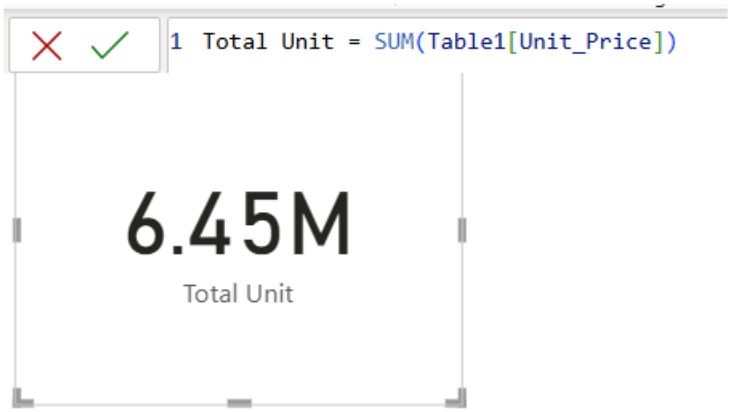
Average Discount Measure

This measure shows how much discount (on average) are giving to customers.



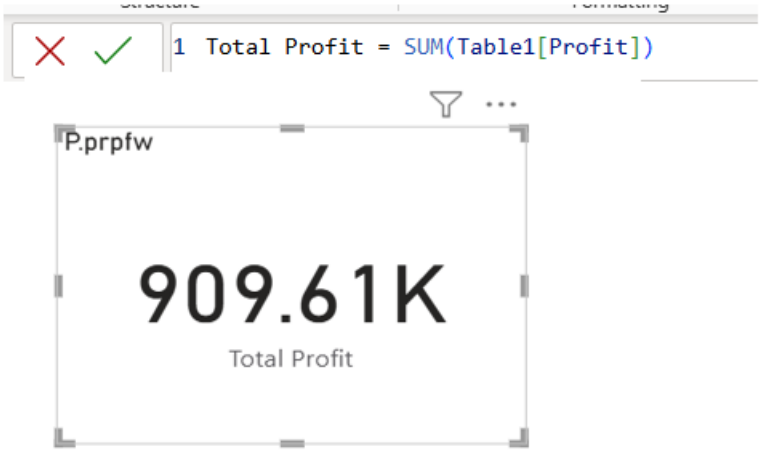
Total Units Measure

It calculates the total number of units sold across your entire dataset. Helps understand product demand.



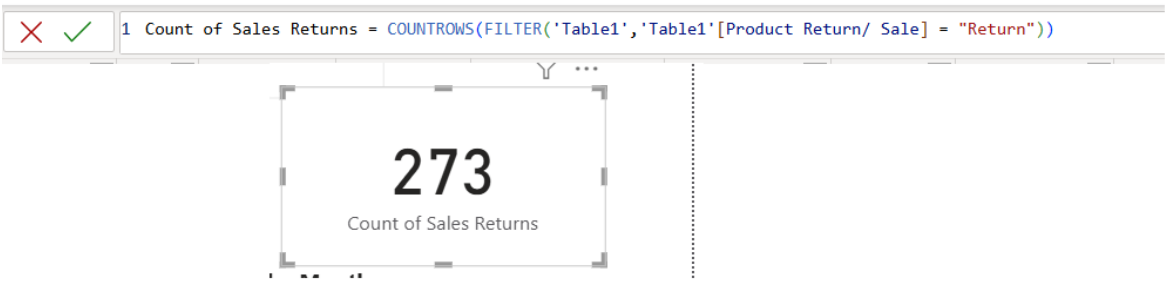
Total Profit Measure

It shows how much money the business actually earns after deducting product cost.



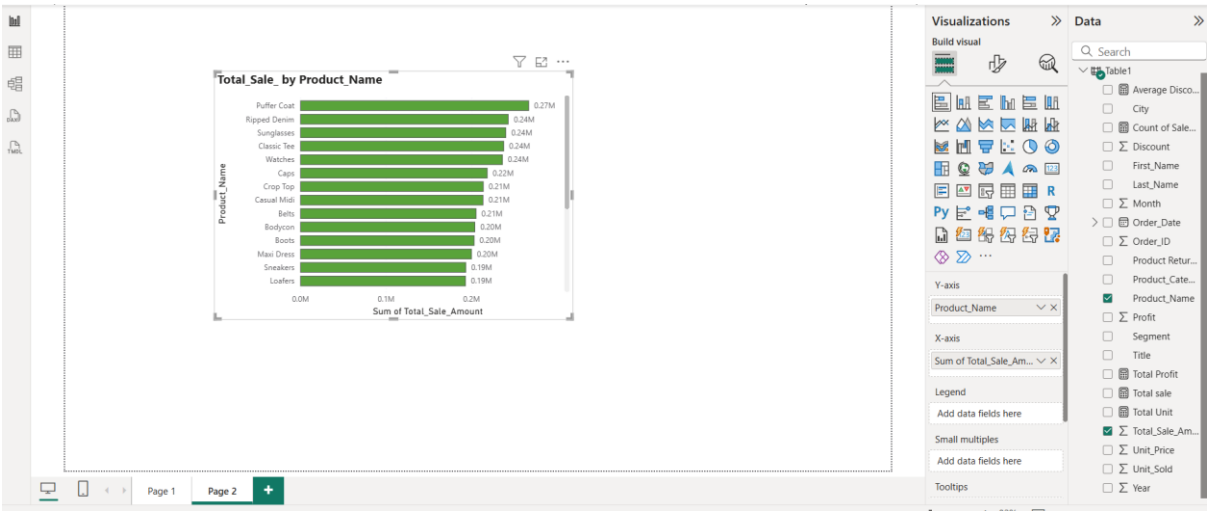
Count of Sales Return – DAX Measure

The Count of Sales Return measure helps identify how many orders were returned by customers. It supports understanding product quality issues and customer dissatisfaction.



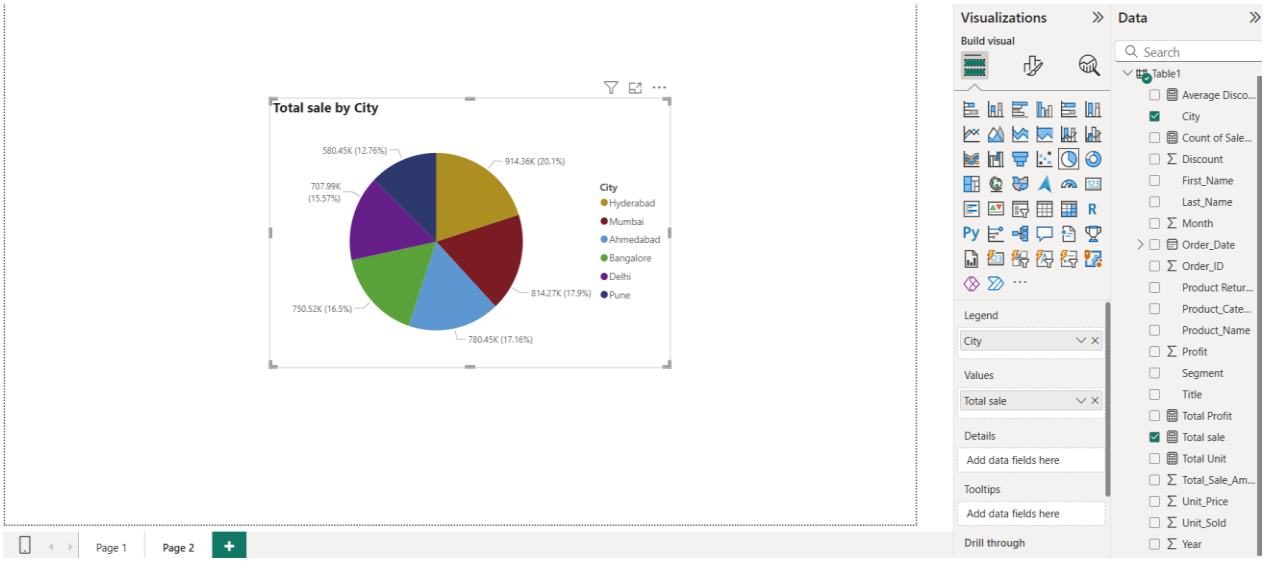
Bar Chart:

- Product Name vs Sales Amount in compare sales performance across different products. To visually identify which products, generate the highest and lowest sales.



Pie Chart:

- Used City and sale amount coulumn to understand which city contributes the most to total sales.
- To compare sales distribution across different geographic locations.
- To identify high-performing and low-performing cities quickly.

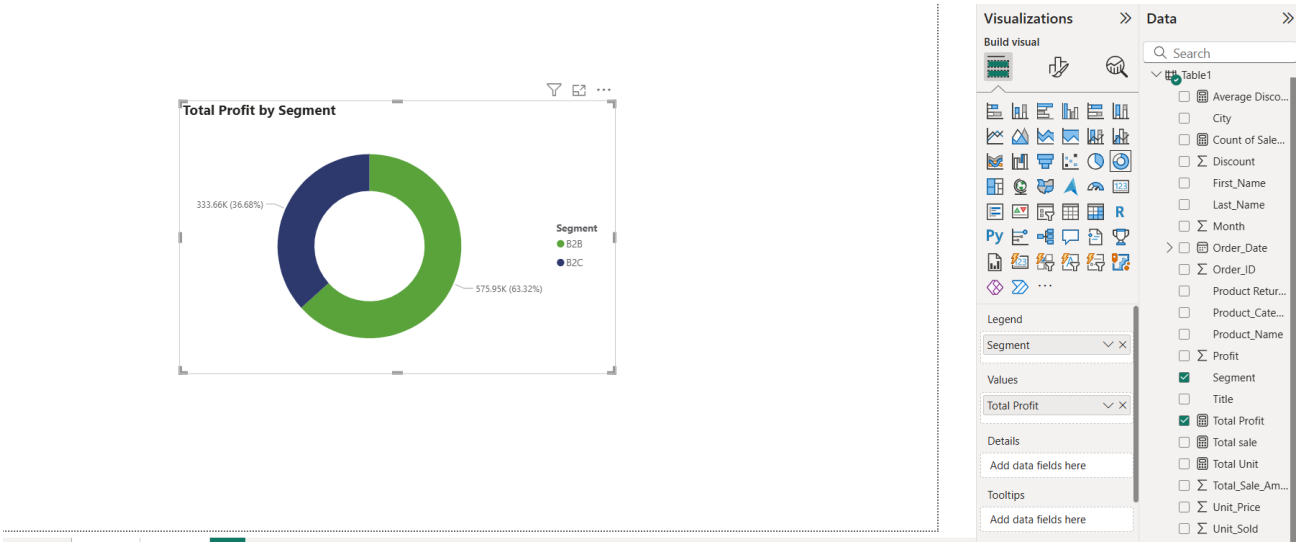


Line Chart :

To track sales and profit performance over time. Shows the time-based performance clearly. To identify trends, seasonal patterns, and growth/decline across months. To compare the movement of Sales vs Profit together in a single timeline.

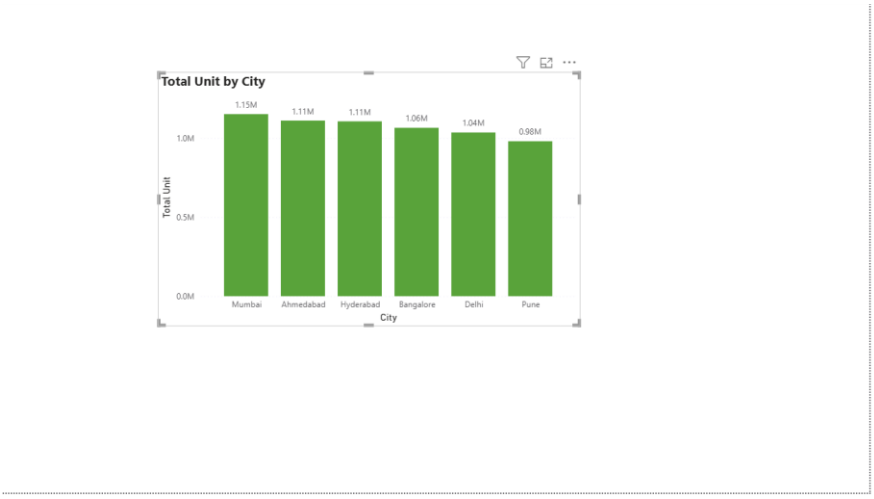
Donut Chart:

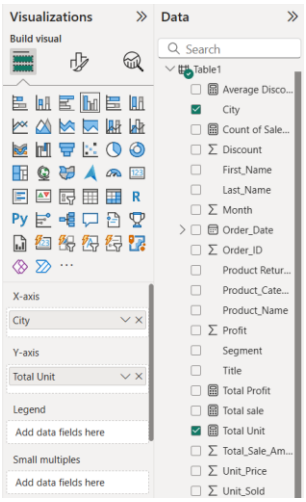
- The Donut Chart displays the total profit contribution of each segment. Each portion of the donut represents one segment, and its size indicates. how much profit that segment generated compared to others.



Column Chart:

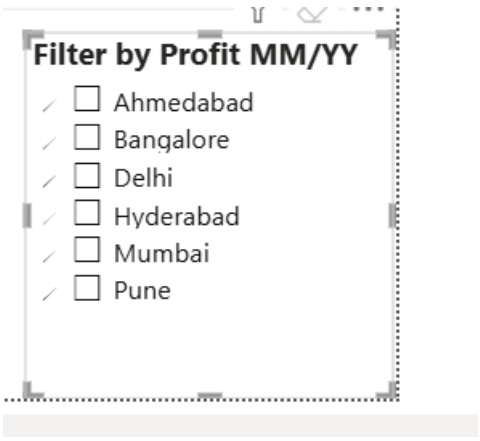
compare unit sales across categories in different cities. The column chart visualizes how many units were sold for each category in different cities. Each column represents the total units sold, grouped by category and separated by city. This helps compare the performance of each category across multiple cities at a glance.



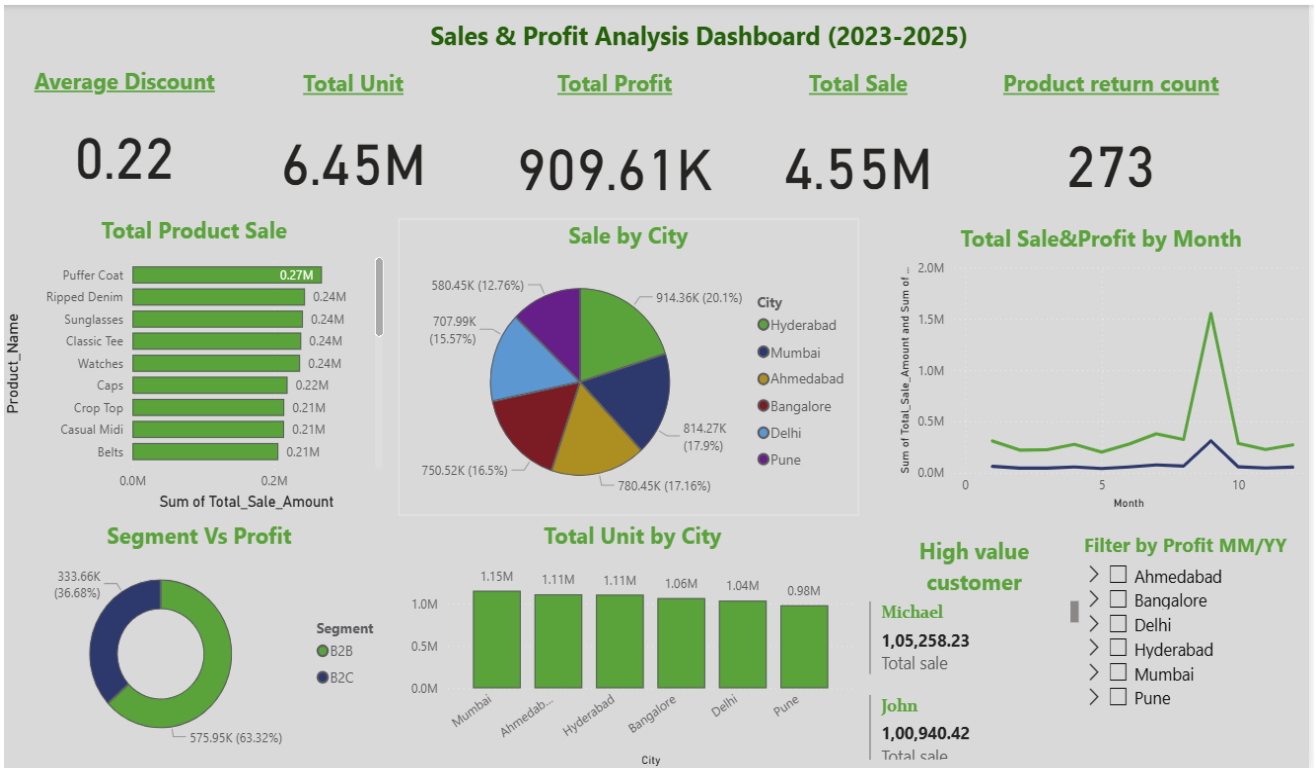


List Slicer:

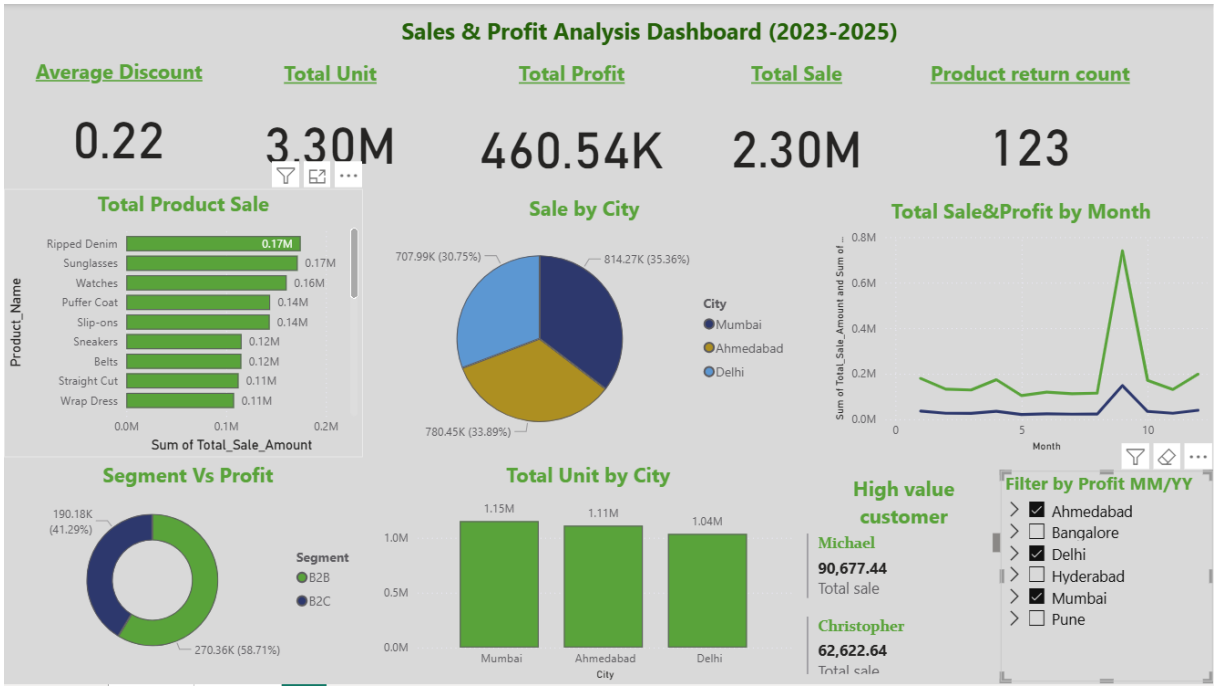
List slicers are added for Month, Year, and City to make the dashboard fully interactive.



Sale Dash Board



After applied list slicer



Key Insights:

- **Product-Level Revenue Concentration:** Sales are heavily concentrated in a few product categories such as Puffer Coats, Ripped Denim, Sunglasses, and Classic Tees, which consistently show the highest total sales amounts. This indicates that these items are your top-performing products and major revenue drivers.

- **Order Volumes & Sales Returns:** Total units sold are over 6.45M, but you also have 273 product sales returns. Even though the return count is small compared to the total orders, it highlights a need to monitor product quality or customer experience for categories that show negative values (returns).

- **Monthly Sales Fluctuation:** The line chart shows visible monthly fluctuations, with peak sales happening around specific months (ex: Month 10 or Month 5 in your chart). This indicates seasonal buying behaviour, which can guide marketing campaigns and inventory planning.

Conclusion :

The Sales & Profit Analysis Dashboard provides a clear and comprehensive overview of business performance across product categories, customer segments. The analysis highlights strong product performance in apparel items such as Puffer Coats and Denim. while cities like Hyderabad and Mumbai continue to drive the highest sales volume. the B2C segment emerges as the most profitable customer group. focus on retail consumers. Although sales returns are limited. Overall, the dashboard delivers actionable insights that support data-driven decision-making to improve sales, profitability, and customer engagement.