

# RETAIL STORE SALES ANALYSIS

-SUBMITTED BY UVESH SHAIKH

## DATA INFO

Column Name	Data Type	Notes
Transaction ID	Text(ID)	Unique Identifier
Customer ID	Text(ID)	Customer Unique ID
Category	Text	Product category
Item	Text	Item name with code
Price per Item	Float	Static price of the one item
Quantity	Integer	Quantity of the purchased item
Total Spent	Float	Total amount spent on the items
Payment Method	Text	Mode of payment
Location	Text	Location where transaction takes place
Transaction Date	DateTime	Date of transaction
Discount	Text	Categorical data

# DATA CLEANING STEPS

## 1. Removed Invalid and Duplicate Entries

- ❖ Removed duplicate rows to avoid data redundancy.
- ❖ Deleted any rows that were completely empty or contained nonsensical values.

## 2. Column Renaming for Clarity

- ❖ Renamed Discount Applied → Discount for simplicity.
- ❖ Renamed Price Per Unit → Price Per Item to better reflect product-level pricing.

## 3. Standardized Discount Values

- ❖ Replaced empty cells in the Discount column with "No Discount".
- ❖ Converted Boolean values:
- ❖ TRUE → "Given Discount"
- ❖ FALSE → "No Discount"

## 4. Normalized Payment Method Values

- ❖ Standardized inconsistent values in the Payment Method column:
- ❖ Digital Wallet → Digital Payment
- ❖ Cash → Cash Payment

## 5. Handled Missing Item Values

- ❖ Used category-based logic to impute missing values in the Item column based on the corresponding Category.

## 6. Imputed Missing Price Values

- ❖ Calculated and filled missing values in the Price Per Item column using the formula:  
$$\text{Price Per Item} = \text{Total Spent} / \text{Quantity}$$

## 7. Created Pivot Table for Analysis

- ❖ Prepared a pivot table summarizing:
- ❖ Sum of Total Sales by Category and location.
- ❖ Discount impact on Category.
- ❖ Monthly Sales across 4 years.
- ❖ Total Sales and frequency of payment method.
- ❖ Quantity sold by category.



## Final Output




The dataset is now cleaned, complete, and well-structured for further reporting, visualization, and dashboarding in Excel, Power BI, or Python.

# PIVOT TABLE

Sum of Total Sales by Category and location			
Sum of Total Spent	Location		
Category	In-store	Online	Grand Total
Beverages	98019.5	98978.5	196998
Butchers	101777	106341	208118
Computers and electric accessories	87323.5	103369	190692.5
Electric household essentials	97778.5	106035	203813.5
Food	95926	98886	194812
Furniture	99611	95699	195310
Milk Products	88813	91299	180112
Patisserie	91415	90550.5	181965.5
<b>Grand Total</b>	<b>760663.5</b>	<b>791158</b>	<b>1551821.5</b>

Discount impact on Category	
Discount	Sum of Total Spent
<b>Given discount</b>	<b>524470.5</b>
Beverages	67723
Butchers	72490
Computers and electric accessories	62741
Electric household essentials	68706.5
Food	60874.5
Furniture	64439.5
Milk Products	60364.5
Patisserie	67131.5
<b>No discount</b>	<b>1027351</b>
Beverages	129275
Butchers	135628
Computers and electric accessories	127951.5
Electric household essentials	135107
Food	133937.5
Furniture	130870.5
Milk Products	119747.5
Patisserie	114834
<b>Grand Total</b>	<b>1551821.5</b>

Monthly Sales across 4 years	
Years (Transaction Date)	(All) 
Months 	Sum of Total Spent
Jan	174387.5
Feb	119685
Mar	122419
Apr	125618.5
May	124551.5
Jun	129771
Jul	131509
Aug	123287.5
Sep	129344
Oct	119413.5
Nov	122146.5
Dec	129688.5
<b>Grand Total</b>	<b>1551821.5</b>

Quantity sold by category						
Sum of Quantity	Year 					
Category 	2022	2023	2024	2025	Grand Total	
Beverages	2722	2429	3062	143	8356	
Butchers	3079	2387	2598	142	8206	
Computers and electric accessories	2519	2760	2821	172	8272	
Electric household essentials	2921	2594	2666	128	8309	
Food	2530	2830	2847	180	8387	
Furniture	2648	2770	2934	110	8462	
Milk Products	2761	2843	2607	128	8339	
Patisserie	2717	2422	2707	87	7933	
<b>Grand Total</b>	<b>21897</b>	<b>21035</b>	<b>22242</b>	<b>1090</b>	<b>66264</b>	
Total sales and frequency of payment method						
Payment Method 	Sum of Total Spent	Count of Transaction ID				
Cash Payment	537703.5	4102				
Credit Card	506839	3925				
Digital Payment	507279	3941				
<b>Grand Total</b>	<b>1551821.5</b>	<b>11968</b>				