

# **RETAIL ANALYSIS PROJECT**

**- Prepared By : Pulkit Mehrotra**

# RETAIL DATA ANALYSIS

# TABLE OF CONTENTS

S.No.	Contents	Slide No.
1.	Business Context & Problem Statement	4
2.	Entity Relationship Diagram & it's Observations	5-6
3.	Tables Overview	7
4.	Data Inconsistencies	8-13
5.	Data Cleaning	14-19
6.	Exploratory Data Analysis [High Level Metrics]	20-23
7.	Customer Level Analysis	24-40
8.	Category Level Analysis	41-50
9.	Cross Selling Analysis	51-54
9.	Sales Trend Analysis	55-63
10.	Store Level Analysis	64-69
11.	Final Observations & Recommendations	70-71

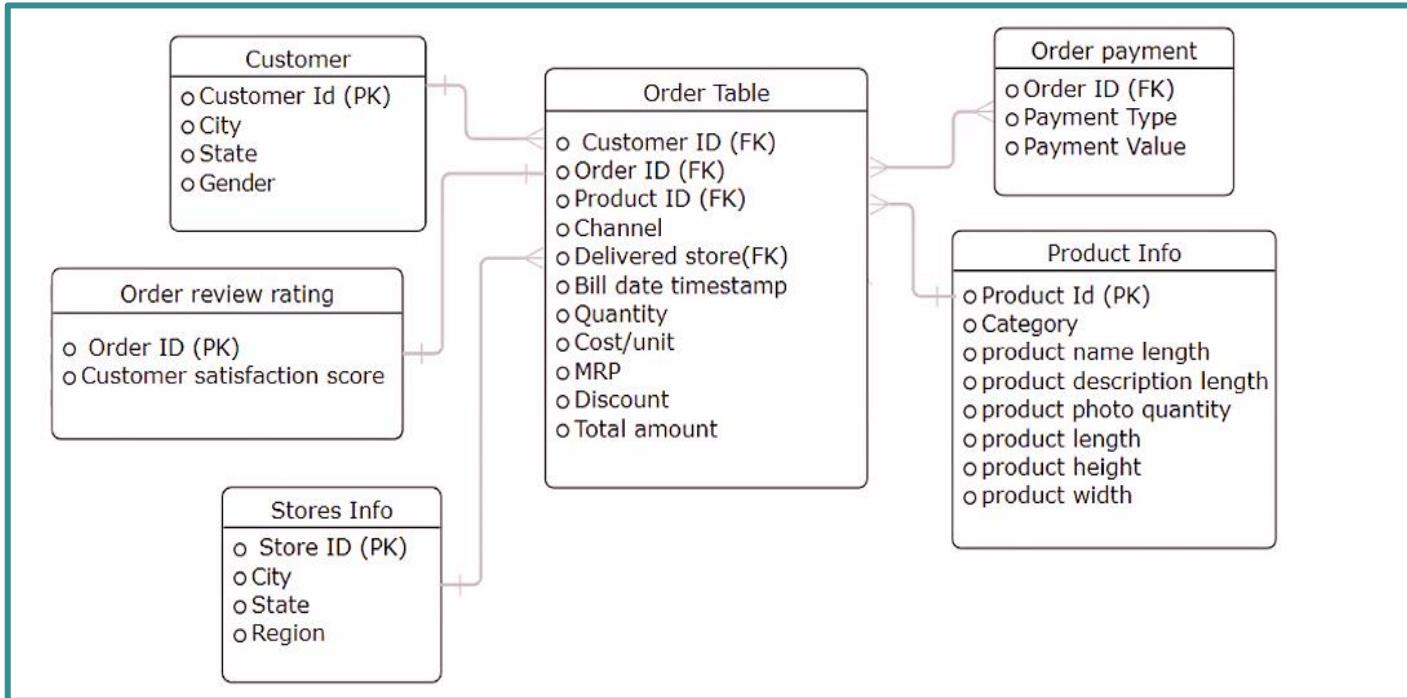
# Business Context

- **About the Client**: The client is one of the leading retail chains in India and would like to partner with Analytixlabs.
- **Client's Expectations**: To provide data-driven insights from the point of sales data to define CRM/marketing/Campaign/Sales strategies going forward. The client also wants help in measuring, managing, and analyzing the performance of the business.

# Problem Statement

- Data-driven insights about business and understand customer behaviors, product behavior, store behavior, channel behavior, etc.
- Key metrics and different types of analysis, including product analysis, product category analysis, customer category analysis, customer segment analysis,
- Sales patterns and trends, seasonality impact on sales, cross selling, customer satisfaction, cohort analysis and store-level analysis to understand category performance.
- Define a strategy to increase sales for the upcoming year.

# Entity-Relationship Diagram



Legend:  
→ Many-to-one relationship  
+— One-to-Many relationship  
↔ Many-to-Many relationship  
— One-to-One relationship

PK - Primary Key  
FK - Foreign Key

# OBSERVATIONS FROM → E-R DIAGRAM

- One order id have multiple payment methods
- One customer have multiple orders
- One order have multiple products
- In 'Orders' table, there are some records for which bill\_date\_timestamp is different from the given time period (i.e. before Sept, 2021 and after Oct, 2023)

# TABLES OVERVIEW

Data has been provided from Sep 2021 to Oct 2023 for specific categories of products for randomly selected customers.

TABLES	CUSTOMERS	ORDERS	ORDERPAYMENTS	ORDERREVIEWS_RATING	PRODUCTSINFO	Stores Info
No. of Rows	99441	112650	103886	100000	32951	536
No. of Columns	4	11	3	2	9	4
No. of Duplicate Records	0	0	615	350	0	2
No. of Unique Records	99441	112650	103271	99650	32951	534
Type of Data	Categorical	Categorical, Numerical	Categorical, Numerical	Ordinal	Categorical, Numerical	Categorical
Primary Key	CustID	order_id + product_id	order_id	order_id	product_id	StoreID
Foreign Key	No	Customer_id, product_id, Delivered_StoreID	order_id	order_id	No	No

# DATA INCONSISTENCIES

# Data Inconsistencies in Customers Table

- There is discrepancy in one of the customer\_city due to the presence of '@' symbol.

226	7154230469	Palasa @ Kasibugga	Andhra Pradesh	F
227	9762563896	Palwancha	Andhra Pradesh	M

- At many places, Delhi Municipal Corporation (U) is shown as a city in Delhi state, but this is a mistake because it can't be shown as a city.

54...	9205630219	Naharlagun	Arunachal Pra...	F
54...	7664990607	Delhi Municipal Corporation (U)	Delhi	F
54...	7184055837	Badepalle	Andhra Pradesh	F

## Solutions :

- ✓ Inform the Sales Department about the presence of '@' symbol in one of customer\_city and gather the correct information that needs to be updated there.
- ✓ Update all the customer\_city of customer\_state : Delhi where it's Delhi Municipal Corporation after consultation with Sales Department.

# Data Inconsistencies in Orders Table (1/2)

- Values in Discount column are not clear, whether they are in monetary terms or in percentage.

	Customer_id	Channel	Cost_Per_Unit	Discount
1	5046435989	Instore	144	0
2	5197899150	Instore	155	0
3	2012608304	Instore	179.899993896484	14
4	2012608304	Instore	179.899993896484	0
5	4971740883	Instore	259	0
6	1402471011	Online	29.9899997711182	0
7	7963455455	Instore	29.9899997711182	0
8	4418191368	Instore	69.9000015258789	0
9	2797327664	Instore	46.2999992370605	8
10	9514471988	Instore	1240	13
11	8304011962	Instore	98.8000030517578	11

- Some orders with same order id, same Customer id and purchased on same date & time but having different MRP & different Discounts.

Customer_id	order_id	product_id	Channel	Delivery_Status	Bill_date_timestamp	Quantity	Cost_Per_Unit	MRP	Discount	Total_Amount
5046435989	000aed2e25dbad2f9ddb70584c5a2d4fa33915031a8cde03dd0d3e8fb27f01		Instore	ST101	5/16/2023 20:57	1	144	152.77	0	152.77
5197899150	002d040018d12a3853c059f7f23ab5b97c58626cc3e15d514d2095f69650220		Instore	ST101	11/27/2022 3:47	1	155	169.84	0	169.84
2012608304	00571ded73b3c061925584feab0db48695c431b31927fefef5343e675f279e7		Instore	ST101	5/25/2022 21:10	1	179.9	208.91	14	194.91
2012608304	00571ded73b3c061925584feab0db48695c431b31927fefef5343e675f279e7		Instore	ST101	5/25/2022 21:10	2	179.9	194.91	0	389.81

# Data Inconsistencies in Orders Table (2/2)

- Entries for a Customer\_id = 7816049273 can be seen duplicated as it has same product id, same order id and same time along with a discrepancy where MRP. & Discount is different from others.

62293	7816049273	3b12b05b1ef9fd1f18face4e10c50c2e	c7dc39bcc7f9bf8a9b253b8727a0cde5	Instore	ST301	2022-09-14 01:10:00.0000000	1	40	55.0999984741211	0
62294	7816049273	3b12b05b1ef9fd1f18face4e10c50c2e	c7dc39bcc7f9bf8a9b253b8727a0cde5	Instore	ST301	2022-09-14 01:10:00.0000000	2	40	55.0999984741211	0
62295	7816049273	3b12b05b1ef9fd1f18face4e10c50c2e	c7dc39bcc7f9bf8a9b253b8727a0cde5	Instore	ST301	2022-09-14 01:10:00.0000000	3	40	55.0999984741211	0
62296	7816049273	3b12b05b1ef9fd1f18face4e10c50c2e	c7dc39bcc7f9bf8a9b253b8727a0cde5	Instore	ST301	2022-09-14 01:10:00.0000000	4	40	55.0999984741211	0
62297	7816049273	3b12b05b1ef9fd1f18face4e10c50c2e	c7dc39bcc7f9bf8a9b253b8727a0cde5	Instore	ST301	2022-09-14 01:10:00.0000000	5	40	63.0999984741211	8

- One more similar case with Customer\_id = 5052834682.

62326	5052834682	49f879aa078787#9ca5e831515407b9	98712b7f46e0fe4f4da2f7df1b9231a7	Instore	ST301	2022-08-31 02:15:00.0000000	1	50	66.7900009155273	0	66.7900009155273
62327	5052834682	49f879aa078787#9ca5e831515407b9	98712b7f46e0fe4f4da2f7df1b9231a7	Instore	ST301	2022-08-31 02:15:00.0000000	2	50	81.7900009155273	15	133.580001831055

## Solutions :

- The discrepancy regarding same products with same time of purchase but different MRP. And Discount needs to be communicated to Sales Department and correct information should be updated.
- It should also need to be confirmed whether Discount column has values in monetary terms or in percentage.

# Data Inconsistencies in Order Payments Table

- Some orders have a payment value equivalent to 0 which is an error.

	order_id	payment_type	payment_value
1	8bcbe01d44d147f901cd3192671144db	voucher	0
2	fa65dad1b0e818e3ccc5cb0e39231352	voucher	0
3	6ccb433e00daae1283ccc956189c82ae	voucher	0
4	4637ca194b6387e2d538dc89b124b0ee	UPI/Cash	0
5	00b1cb0320190ca0daa2c88b35206009	UPI/Cash	0
6	45ed6e85398a87c253db47c2df9f48216	voucher	0

- There are some orders whose payment value is less than 1 and appears to be inconsistent or wrong.

	order_id	payment_type	payment_value
1	4884bd0f5624b0b791920965686cf6ff	credit_card	0.0299999993294477
2	bb1d699eccd1fe0e0b22aea230eab2d0	credit_card	0.0299999993294477
3	4884bd0f5624b0b791920965686cf6ff	credit_card	0.0299999993294477
4	7db5f2eb8f5f54db9f9e71ba4296bcbf	voucher	0.00999999977648258
5	fb4de3600d359f84927517e78ff9ba54	voucher	0.00999999977648258
6	636f0241ddc83a3b9e37a8088167bd45	voucher	0.00999999977648258
7	25b5b0ea53b7d5a2d5712a0d9d0b3649	credit_card	0.00999999977648258
8	0218c7a4fb8d5b1bd22c82b783b8359c	credit_card	0.00999999977648258
9	7db5f2eb8f5f54db9f9e71ba4296bcbf	voucher	0.00999999977648258
10	fb4de3600d359f84927517e78ff9ba54	voucher	0.00999999977648258
11	ca4b9f3ce6fc19e8533501cf8c6b832e	credit_card	0.00999999977648258

# Data Inconsistencies in Products Info Table

- Some products have **zero** product weight which is not possible, hence needs to be corrected.

	product_id	product_weight_g	product_length_cm	product_height_cm	product_width_cm
1	81781c0fed9fe1ad6e8c81fca1e1cb08	0	30	25	30
2	8038040ee2a71048d4bdbbdc985b69ab	0	30	25	30
3	36ba42dd187055e1fbe943b2d11430ca	0	30	25	30
4	e673e90efa65a5409ff4196c038bb5af	0	30	25	30

- There are many products for which the category, product's name length, description & photos quantity is **NULL**.

	product_id	Category	product_name_lenght	product_description_lenght	product_photos_qty	product_weight_g	product_length_cm
1	a41e356c76fab66334f36de622ecbd3a	#N/A	NULL	NULL	NULL	650	17
2	d8dee61c2034d6d075997acef1870e9b	#N/A	NULL	NULL	NULL	300	16
3	56139431d72cd51f19eb9f7dae4d1617	#N/A	NULL	NULL	NULL	200	20
4	46b48281eb6d663ced748f324108c733	#N/A	NULL	NULL	NULL	18500	41
5	5fb61f482620cb672f5e586bb132eaef9	#N/A	NULL	NULL	NULL	300	35
6	e10758160da97891c2fdcbc35f0f031d	#N/A	NULL	NULL	NULL	2200	16

## Solutions :

- ✓ Reconcile with Products Department regarding the products where product weight is 0 and update it with correct values.
- ✓ Inform the Products Department regarding the products whose category, name length, description length and photos quantity is NULL and get the correct information about it.

# DATA CLEANING

# RECOMMENDATIONS FOR DATA CLEANING

- For instances where order ID is same with different product ID but having consecutive quantity ,the analysis for transaction to be done on net quantity and net amount.
- For multiple customers for a single order in 'Orders' table, either we replace the multiple customer id with one best suited customer id or we can exclude those order\_ids for the purpose of data analysis.
- Similar approach may be adopted for multiple stores for a single order in 'Orders' table.
- Similar approach can be adopted for multiple bill\_date\_timestamp for a single order in 'Orders' table.
- For the order id in which there is mismatch between total of 'Net amount' in 'Orders' table and 'Payment\_value' in 'OrderPayments' table, we may exclude them for the purpose of data analysis.
- For multiple customer ratings for a single order id, average of customer ratings can be taken.
- As 'Stores Info' table has unique store level data, the duplicate store id can be removed.
- To rectify inconsistent bill date timestamps in the 'Orders' table, the datetime format should be standardized.
- In 'Orders' table, there is an ambiguity in discount column that is it in % or value. For practical assumption, we take it as value because some numbers are greater than 100
- Currencies are not mentioned in tables. For practical assumption, we take it as INR.

# DATA CLEANING

## **CUSTOMERS TABLE :**

- Removing Special Characters from Customer city and replacing with space.

## **ORDERS TABLE [ FACT TABLE]:**

- Keeping the Entries for only those where Quantities were maximum for the same customer id, product id and order id where Quantities were cumulative.
- Eliminating the customer ids with same order id, just keeping the one record in Orders Table.
- Removing Duplicate Records from Orders Table where order ids, customer id, product id, Channel, Delivered StoreID, Bill\_date\_timestamp and MRP are same.
- Correcting the data format IN ORDERS TABLE Where in BILLSTAMP THE DATE IS INAPPROPRIATE FORMAT. Standardizing all date formats to 'YYYY-MM-DD HH:MI:SS'.
- Replacing all Quantity with 1 for cumulative quantities where Customer id and Order id are same for different records but Product id is different.
- Recalculating the Total Amount by using formula : Total Amount = (Quantity\*MRP)-Discount and updating the Total Amount column in Orders Table for the same.

# DATA CLEANING

- Adding a New Column = "Profit" to Orders table and Calculating the Profit by using formula : Profit = [Total Amount] - ([Cost Per Unit]\*Quantity) and updating the same in Orders table.
- Removing all Records from Orders Table of Year 2020 because it's not part of our analysis.
- Updating the Orders Table with Keeping one customer id and remove other where one order id is assigned to multiple customer id.

## **ORDER PAYMENTS TABLE:**

- Updating all payment values in Order Payments Table with Total Amount Values in Orders Table having same order ids.
- Removing the records in Order Payments Table for which there is no similar order ids in Orders Table.
- Removing all Duplicate records in Order Payments Table with same order id, payment type and payment value. Keeping one instance of each such records.

# DATA CLEANING

## ORDER REVIEW RATINGS TABLE :

- Removing the records in Order Review Ratings Table for which there is no similar order ids in Orders Table.
- Removing all Duplicate records in Order Review Ratings Table with same order id and Customer Satisfaction Score

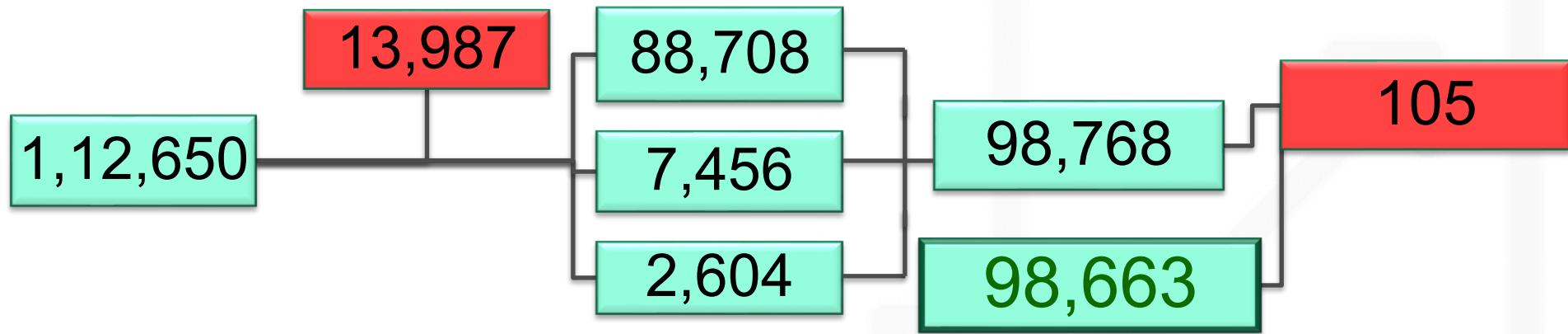
## ORDER PAYMENTS TABLE :

- Removing all records from Product Info Table where product length, product description length and product photos qty is NULL and Category is #N/A.
- Removing all those records from Product Info Table which does not have corresponding records in Orders Table on the basis of product id.

## STORES INFO TABLE :

- No Data Cleaning Required.

# Orders Table → Flow Chart



- **1,12,650** : Total records in Orders Table.
- **88,708** : Records found while mapping aggregated Total Amount column of Orders Table & Payment value column of Order Payments Table.
- **13,987** : Records Dropped due to no match found either because of high discrepancy in amount values, cumulative quantity issues or total amount calculation error.
- **7,456** : Records found while mapping Order's table Total Amount column & aggregated Order Payment's Payment Value column, to get records with cumulative quantity issue with same product id.
- **2,604** : Mapped results again with Order Payments table after resolving issues of consecutive quantity and single order id associated with multiple different Store IDs; by setting quantities to 1, total amount calculation & aggregation.
- **98,768** : Consolidating →  $88,708 + 7,456 + 2,604$
- **98,663** : Final records found after removing more **105** records due to duplication.

# EXPLORATORY DATA ANALYSIS (EDA)

# HIGH LEVEL METRICS (1/3)

## Customer Information

**Total Customers :** 98,572

**Average Discount per Customer :** ₹5.16

**Average Revenue per Customer :** ₹160.50

**Average Profit per Customer :** ₹22.95

**Average Customer Satisfaction Score :** 4

**Customer Retention Rate :** 0.329%

## Sales Information

**Total Revenue :** ₹15,820,372

**Total Profit :** ₹2,262,459

**Total Cost :** ₹13,557,912

**Total Discount :** ₹508,878

**Percentage of Discount :** 3.75%

**Percentage of Profit :** 16.68%

# HIGH LEVEL METRICS (2/3)

## Orders Information

**Total Orders :** 98,633  
**Average Discount per Order :** ₹5.15  
**Average Revenue per Order :** ₹160.40  
**Average Profit per Order :** ₹22.94  
**Total Payment Methods :** 4  
**Average Items per Order :** 1.13  
**Total Channels :** 3

## Products Information

**Total No. Products :** ₹32,950  
**Total No. of Categories :** 14  
**Most Frequent Purchased Product :**  
99a4788cb24856965c36a24e339b6058 (467 times)  
**Product Category with Maximum Revenue :** Toys & Gifts  
**Product Category with Maximum Profit :** Home Appliances

# HIGH LEVEL METRICS (3/3)

## Seller Store Information

**Total Quantity :** 111,880

**Total Stores :** 37

**Total Regions :** 4

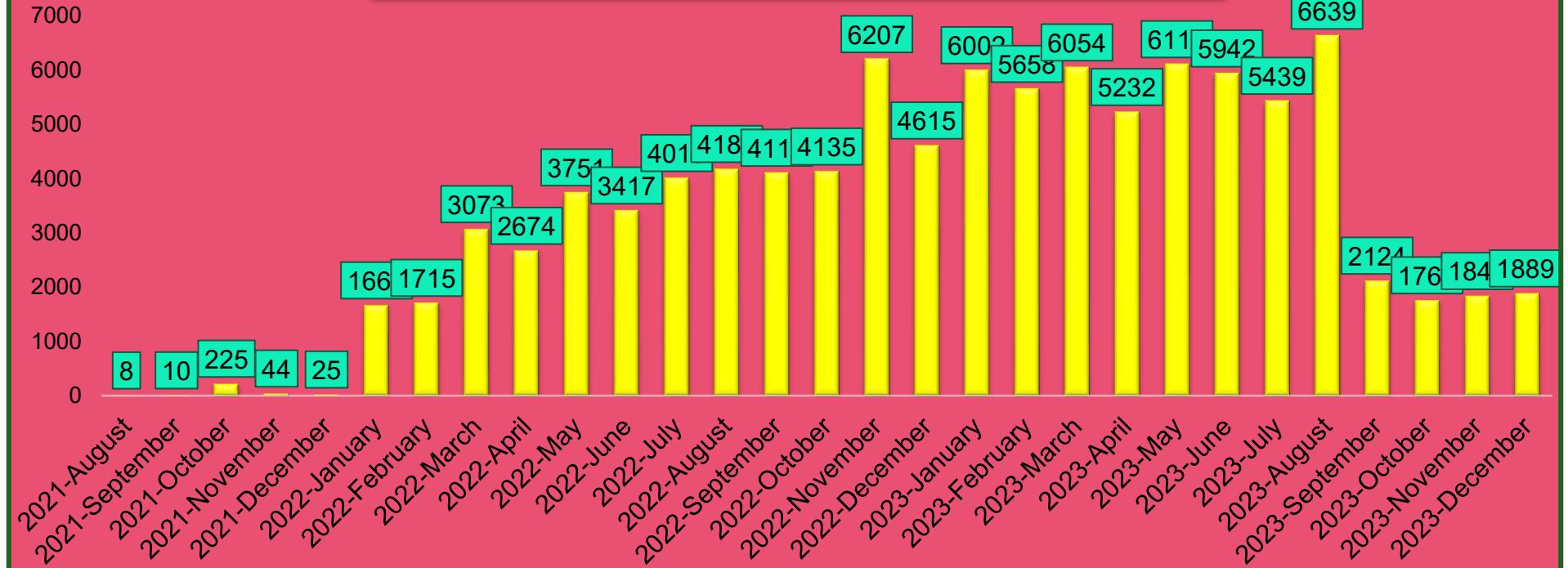
**Most Profitable Store :** ST103 (Profit : ₹61,0331)

**Region with Maximum Revenue :** South

# CUSTOMER LEVEL ANALYSIS

# Customer Level Analysis (1/11)

## New Customers Acquired Every Month



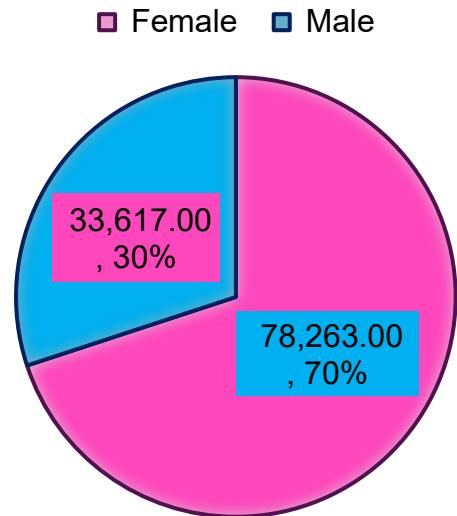
# Customer Level Analysis (2/11)

Revenues From New Customers On Monthly Basis



# Customer Level Analysis (3/11)

## GENDER – QUANTITY RELATIONSHIP



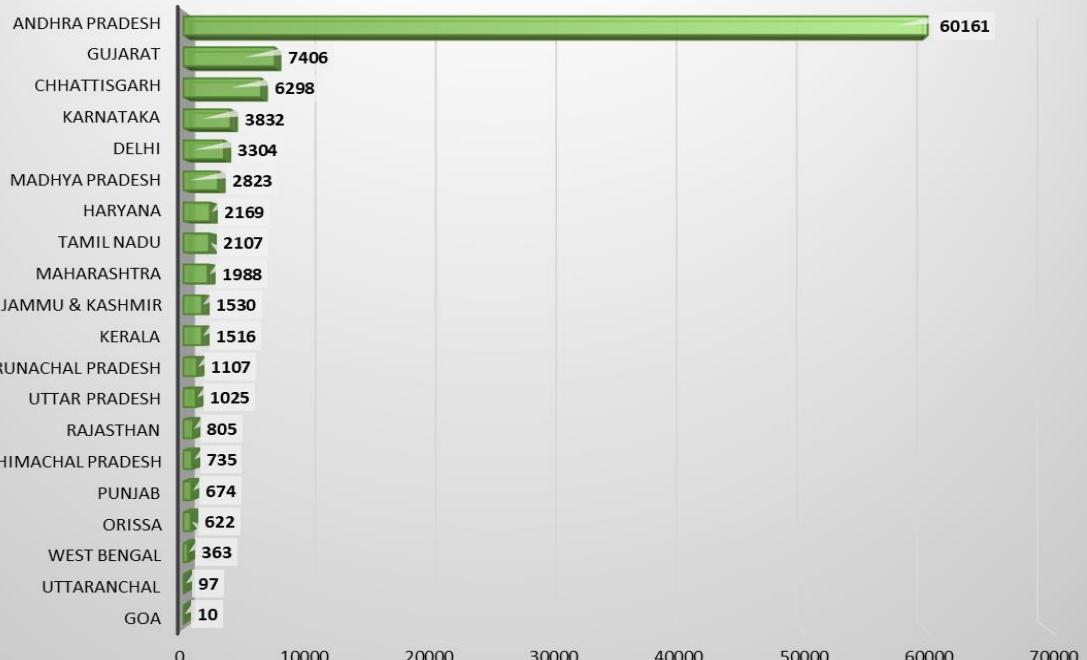
## Gender Wise Contribution :

Gender	Total Revenue	Total Quantity
Female	₹11,065,239.88	78,263
Male	₹4,755,132.36	33,617

# Customer Level Analysis (4/11)

Count of Custid

## STATE - WISE CUSTOMERS DISTRIBUTION



States	No. of Customers
Andhra Pradesh	60161
Gujarat	7406
Chhattisgarh	6298
Karnataka	3832
Delhi	3304
Madhya Pradesh	2823
Haryana	2169
Tamil Nadu	2107
Maharashtra	1988
Jammu & Kashmir	1530
Kerala	1516
Arunachal Pradesh	1107
Uttar Pradesh	1025
Rajasthan	805
Himachal Pradesh	735
Punjab	674
Orissa	622
West Bengal	363
Uttaranchal	97
Goa	10
<b>Grand Total</b>	<b>98572</b>

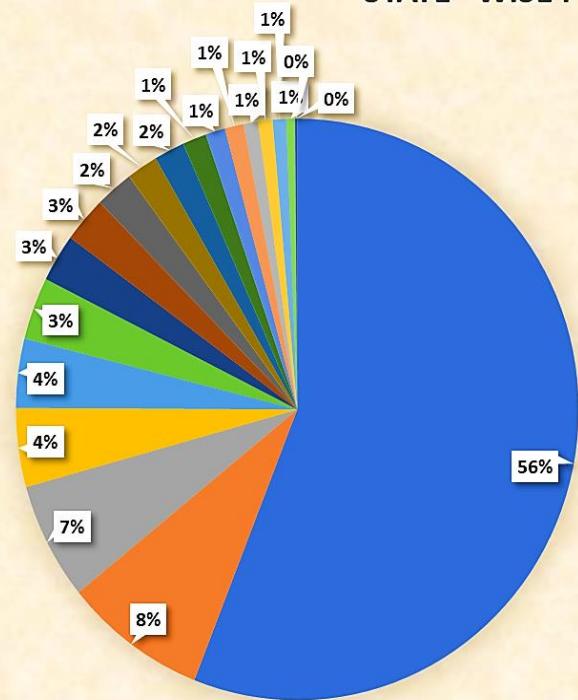
# Customer Level Analysis (5/11)



Cities	No. of Customers
Akkarampalle	15385
Badepalle	6825
Anantapur	2750
Eddumailaram	2113
Anakapalle	1510
Alwal	1429
Eluru	1371
Pedana	1237
Bestavaripetta	1177
Adoni	925
Jaggayyapet	845
Bheemavaram	791
Nagari	741
Kyathampalle	706
Bobbili	685
Bheemunipatnam	684
Dowleswaram	649
Bade Bacheli	623
Moragudi	611
Bhadrachalam	565

# Customer Level Analysis (6/11)

Sum of Profit



customer_state	
■ Andhra Pradesh	■ Gujarat
■ Chhattisgarh	■ Karnataka
■ Delhi	■ Madhya Pradesh
■ Tamil Nadu	■ Maharashtra
■ Haryana	■ Kerala
■ Jammu & Kashmir	■ Uttar Pradesh
■ Arunachal Pradesh	■ Rajasthan
■ Punjab	■ Himachal Pradesh
■ Odissa	■ West Bengal
■ Uttarakhand	■ Goa

States	Total Profit
Andhra Pradesh	1560034.35
Gujarat	228867.1
Chhattisgarh	182109.23
Karnataka	122153.73
Delhi	107337.89
Madhya Pradesh	96657.93
Tamil Nadu	74384.58
Maharashtra	71066.18
Haryana	60926.39
Kerala	49965
Jammu & Kashmir	48954.81
Uttar Pradesh	38205.36
Arunachal Pradesh	31587.75
Rajasthan	30062.24
Punjab	23413.94
Himachal Pradesh	22956.8
Orissa	21563.73
West Bengal	14022.9
Uttaranchal	3142.8
Goa	305.93
<b>Grand Total</b>	<b>2787718.64</b>

# Customer Level Analysis (7/11)



Cities	Total Profit
Akkarampalle	309851.94
Badepalle	200937
Anantapur	76373.84
Eddumailaram	62397
Pedana	44482.99
Eluru	41788.4
Anakapalle	40888.4
Alwal	30122.62
Dowleswaram	26326.35
Moragudi	24893.33
Jaggayyapet	24877.96
Bestavaripeta	23425.92
Bheemunipatnam	22689.74
Kuppam	19695.06
Adoni	17554.24
Bhadrachalam	17335.97
Kyathampalle	15839.65
Chilakaluripet	15769.89
Bheemavaram	15356.7
Nagari	14434.46

# RFM SEGMENTATION

- **Recency (R):** How recently a customer made a purchase.
  - **Frequency (F):** How often a customer makes a purchase over a certain period.
  - **Monetary (M):** How much money a customer spends on purchases over a certain period.
- 
- Each customer is assigned a score for Recency, Frequency, and Monetary using the NTILE(4) function, which divides the customers into quartiles (groups of 4).
  - Recency\_Score:** A score for recency, where the most recent buyers get higher scores.
  - Frequency\_Score:** A score for frequency, where customers who have made more purchases get higher scores.
  - Monetary\_Score:** A score for monetary value, where customers who have spent more money get higher scores.

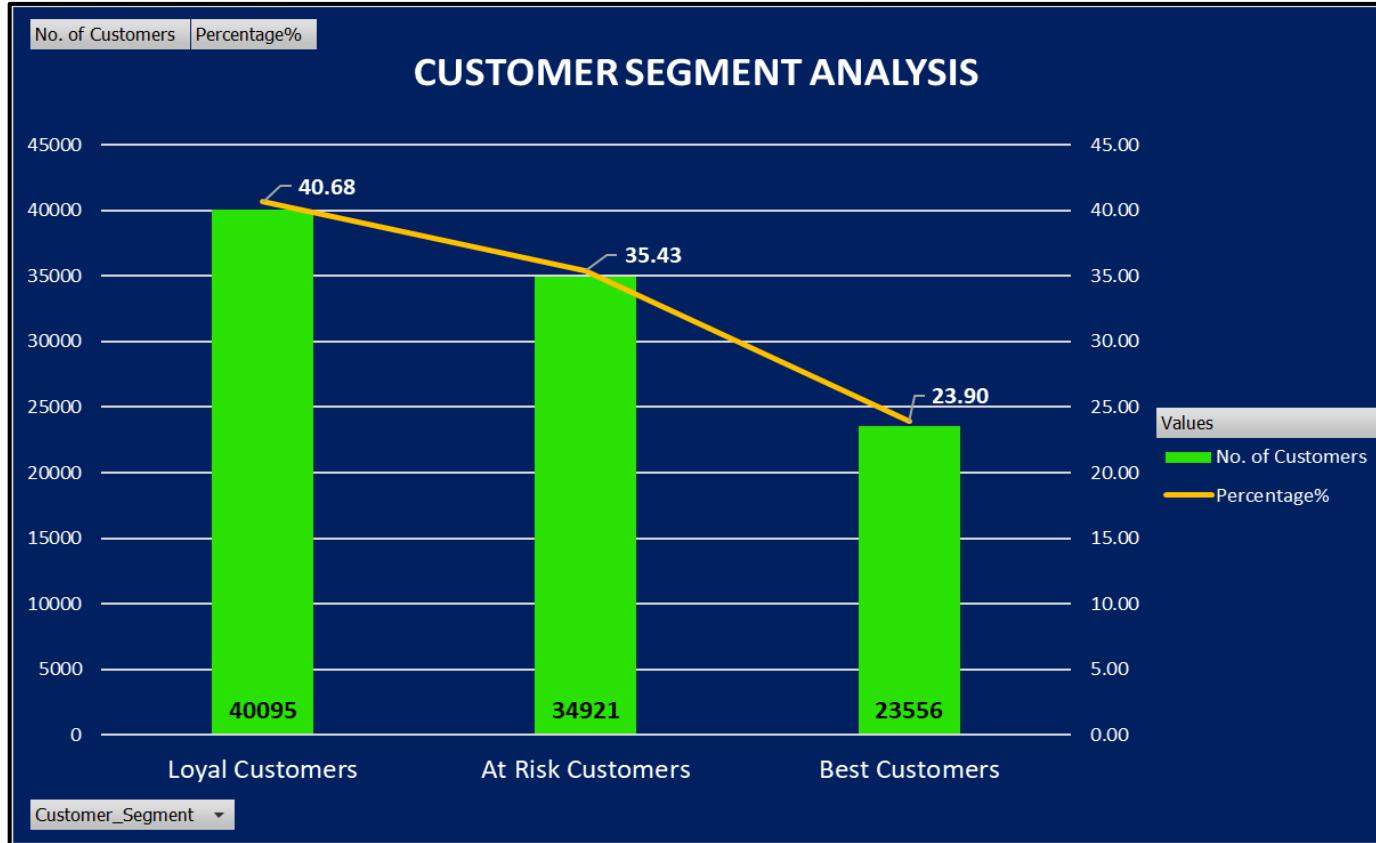
# RFM SEGMENTATION

## Total\_scoring:

This step combines the R, F, and M scores into a Total\_score by summing Recency\_Score, Frequency\_Score, and Monetary\_Score.

A customer with the lowest scores for recency, frequency, and monetary will have a total score of 3 ( $1 + 1 + 1$ ), while a customer with the highest scores in all three categories will have a total score of 12 ( $4 + 4 + 4$ ).

# Customer Level Analysis (8/11)



# Customer Level Analysis (9/11)

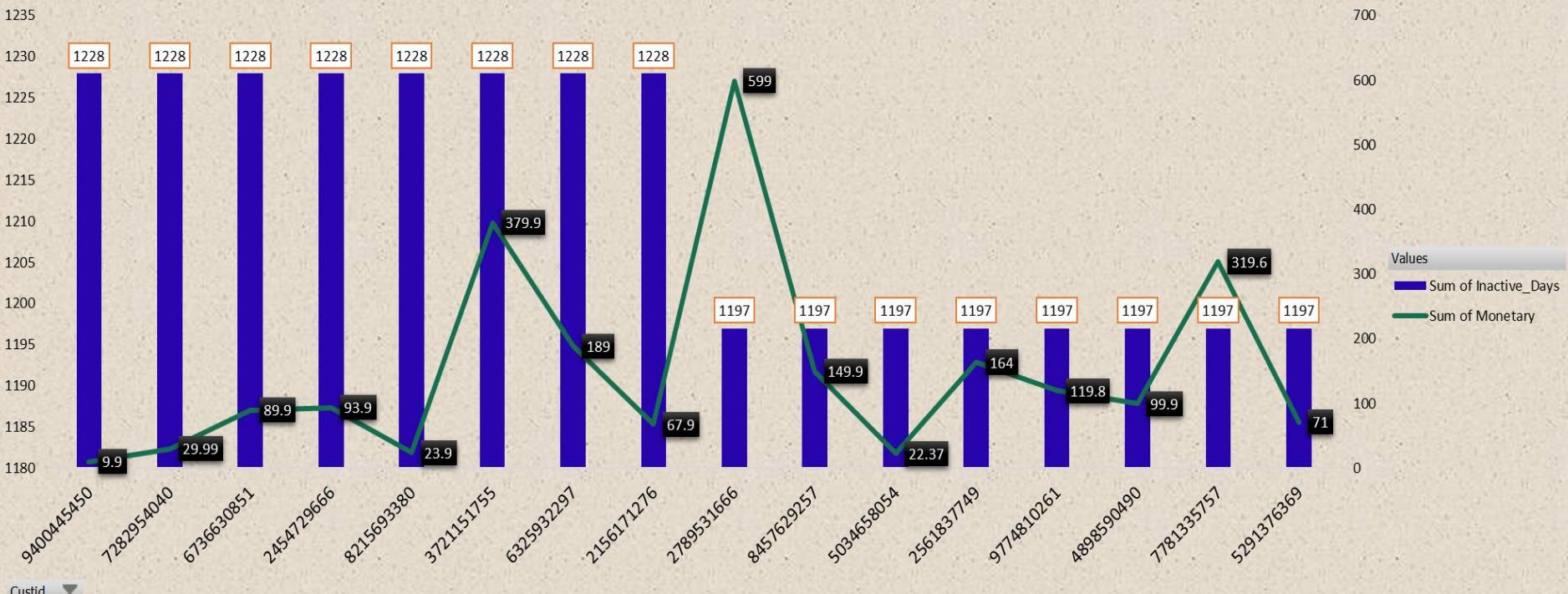
Customer Segment	No. of Customers	Percentage%
Loyal Customers	40095	40.68
At Risk Customers	34921	35.43
Best Customers	23556	23.90
Grand Total	98572	100

Customer Segments	Total Score
Best	Between 10 & 12
Loyal	Between 7 & 9
At Risk	Between 3 & 6

# Customer Level Analysis (10/11)

Sum of Inactive\_Days Sum of Monetary

## TOP 10 CUSTOMERS WITH MAXIMUM NO. OF INACTIVE DAYS & THEIR CONTRINUTION TO REVENUE



# Customer Level Analysis (11/11)

Customer Segment	No. of Customers	Percentage %
One-time Buyers	98534	99.961%
Repeat Buyers	38	0.039%
Discount Seekers	38578	39.137%
Non-discount Seekers	58388	59.234%
Mixed Discount Seekers	1606	1.629%
Single Category Buyers	97821	99.238%
Multiple Category Buyers	751	0.762%
Customers using All Channels	1	0.001%

# OBSERVATIONS & RECOMMENDATIONS

## Gender Distribution

- **Observation:**

The customer base has more female customers (68.87 k) than male (29.70K). Female customers contributed to a higher sales percentage than males.

## One-Time Buyers vs Repeat Buyers

- **Observation:**

The customer base has significantly high number of one-time buyers (98.5K) compared to repeat buyers (38)

- **Recommendation:**

Introduce loyalty programs or personalized marketing to encourage repeat purchases.

# OBSERVATIONS & RECOMMENDATIONS

## RFM Segmentation

- **Observations:**

The majority of customers fall into the Loyal Customer segment 40.01K (40.68%) and At-Risk Customer segment 34.92K (35.43%).

The distribution indicates that a large portion of customers are in the middle-value Loyal Customer segment as it has a total score between 7 & 9 and low-value At Risk Customer segment of total score between 3 & 6, with fewer customers in the higher-value Best Customer Segment with a total score of between 10 & 12. This is a situation where quick actions are required.

# OBSERVATIONS & RECOMMENDATIONS

## RFM Segmentation

- **Recommendations:**

Develop targeted strategies to move customers from the At Risk Customer segment to higher-value segments like Loyal and Best Customer segments. This could include personalized offers, exclusive deals, and enhanced customer service.

Implement retention programs for Loyal & Best customers to maintain their loyalty and encourage higher spending. Offer exclusive benefits and rewards to these high-value customers.

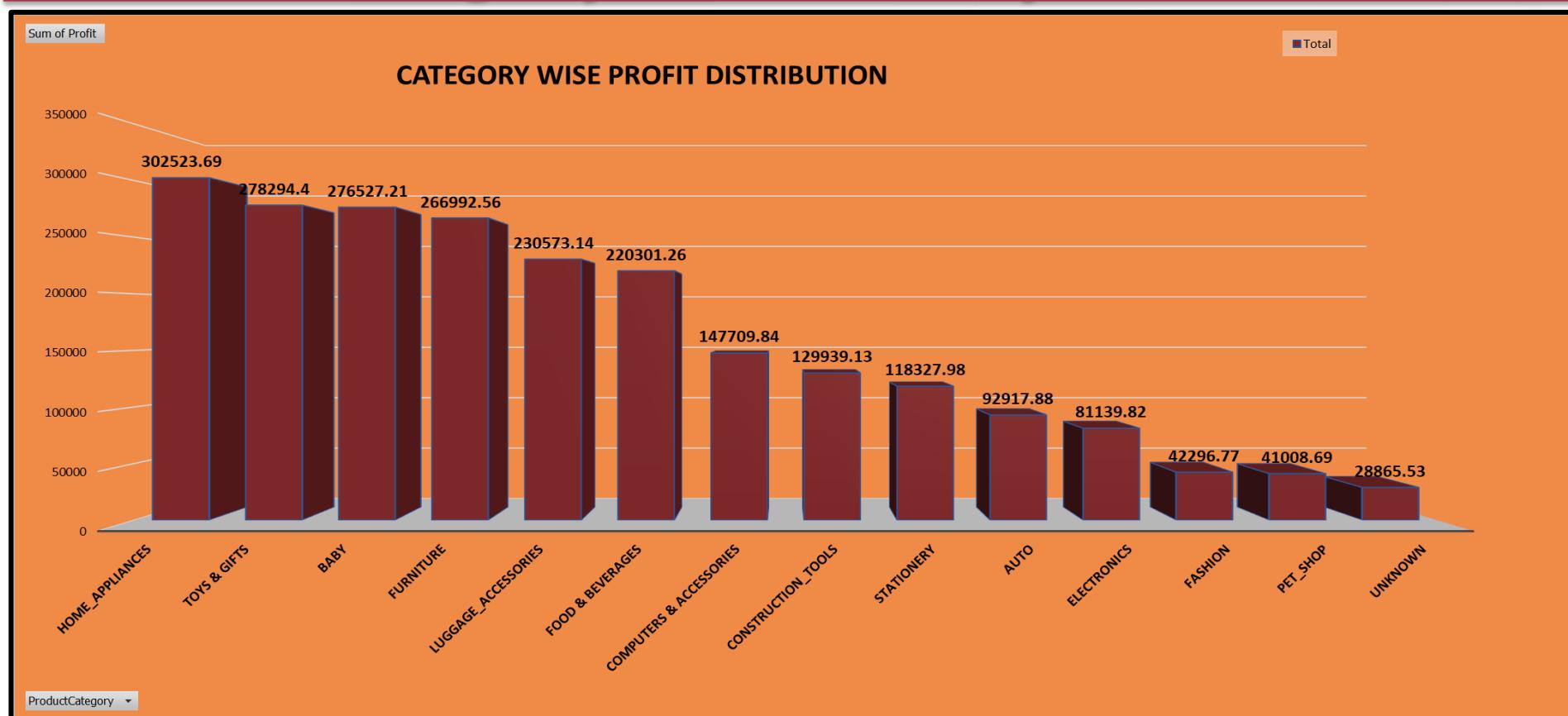
Educate customers in the At Risk segment about the benefits of moving to higher-value segments. Highlight the advantages of frequent purchases and higher spending.

# CATEGORY LEVEL ANALYSIS

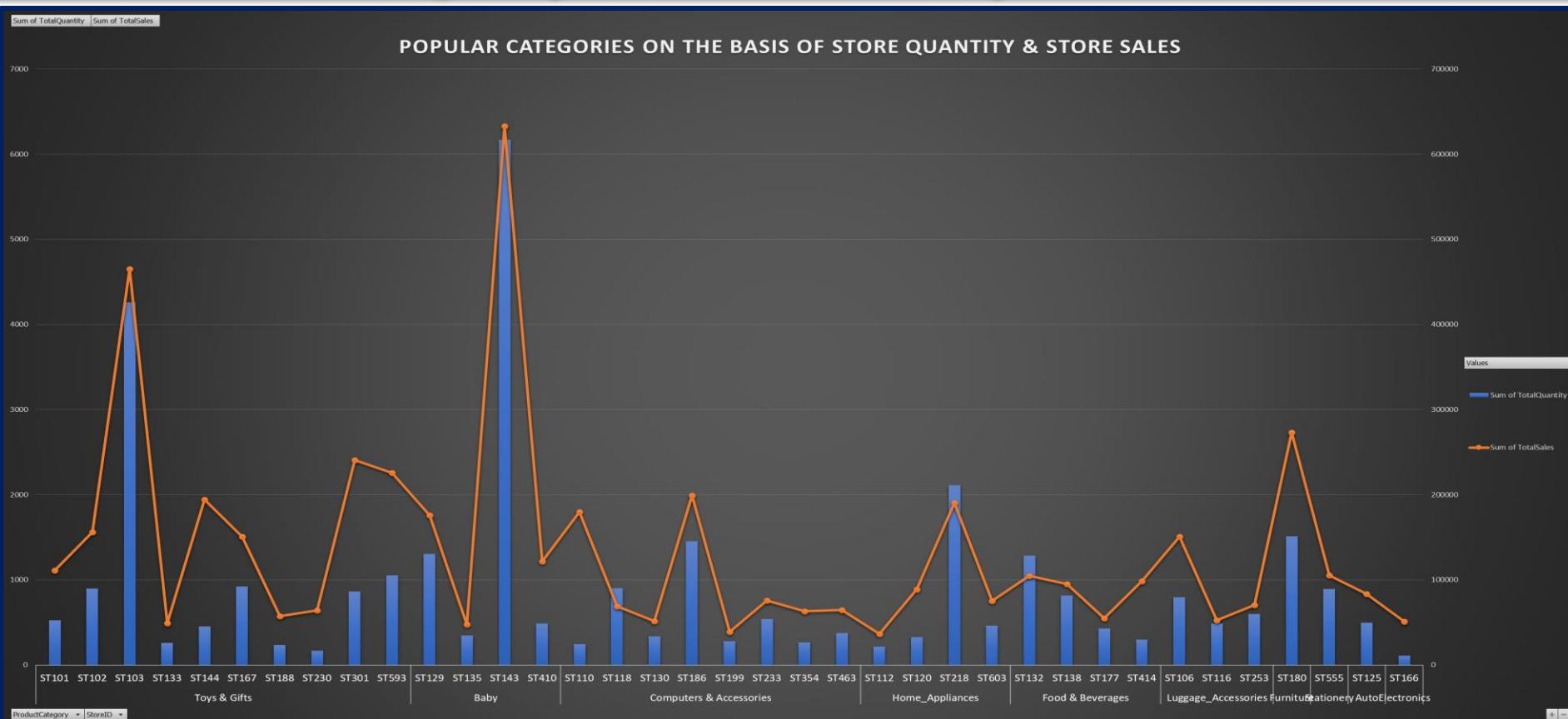
# Category Level Analysis (1/9)

Product Category	Total Sales	Total Quantity	Total Profit
Toys & Gifts	2,640,831.35	16,054.00	276,474.97
Home Appliances	1,790,766.86	13,502.00	303,588.42
Baby	1,713,268.84	14,047.00	276,376.93
Food & Beverages	1,652,594.65	11,355.00	220,177.09
Luggage_Accessories	1,645,974.23	10,958.00	233,175.47
Furniture	1,355,219.27	10,506.00	268,363.69
Computers & Accessories	1,287,577.30	7,963.00	150,174.60
Construction_Tools	1,104,901.04	5,736.00	129,125.95
Auto	685,751.90	4,225.00	92,840.86
Stationery	682,152.87	7,134.00	118,443.28
Electronics	550,497.34	4,210.00	81,505.49
Pet_Shop	253,855.23	1,937.00	41,008.69
Fashion	243,830.83	2,635.00	42,250.37
Unknown	213,150.53	1,618.00	28,954.16

# Category Level Analysis (2/9)



# Category Level Analysis (3/9)



# Category Level Analysis (4/9)

Row Labels	Sum of TotalQuantity	Sum of TotalSales
<b>Toys &amp; Gifts</b>	<b>9636</b>	<b>1713992.91</b>
ST101	523	111055.04
ST102	898	156112.28
ST103	4261	464964.64
ST133	257	48971.68
ST144	455	194235.72
ST167	924	150960.88
ST188	234	57404.95
ST230	167	63964.61
ST301	865	240461.94
ST593	1052	225861.17
<b>Baby</b>	<b>8307</b>	<b>978246.93</b>
ST129	1303	175780.11
ST135	346	47853.88
ST143	6170	632892.52
ST410	488	121720.42
<b>Computers &amp; Accessories</b>	<b>4395</b>	<b>741746.76</b>
ST110	245	179982.46
ST118	901	68904.28
ST130	338	51373.92
ST186	1456	199066.81
ST199	278	38846.55
ST233	539	75837.61
ST354	263	63192.38
ST463	375	64542.75

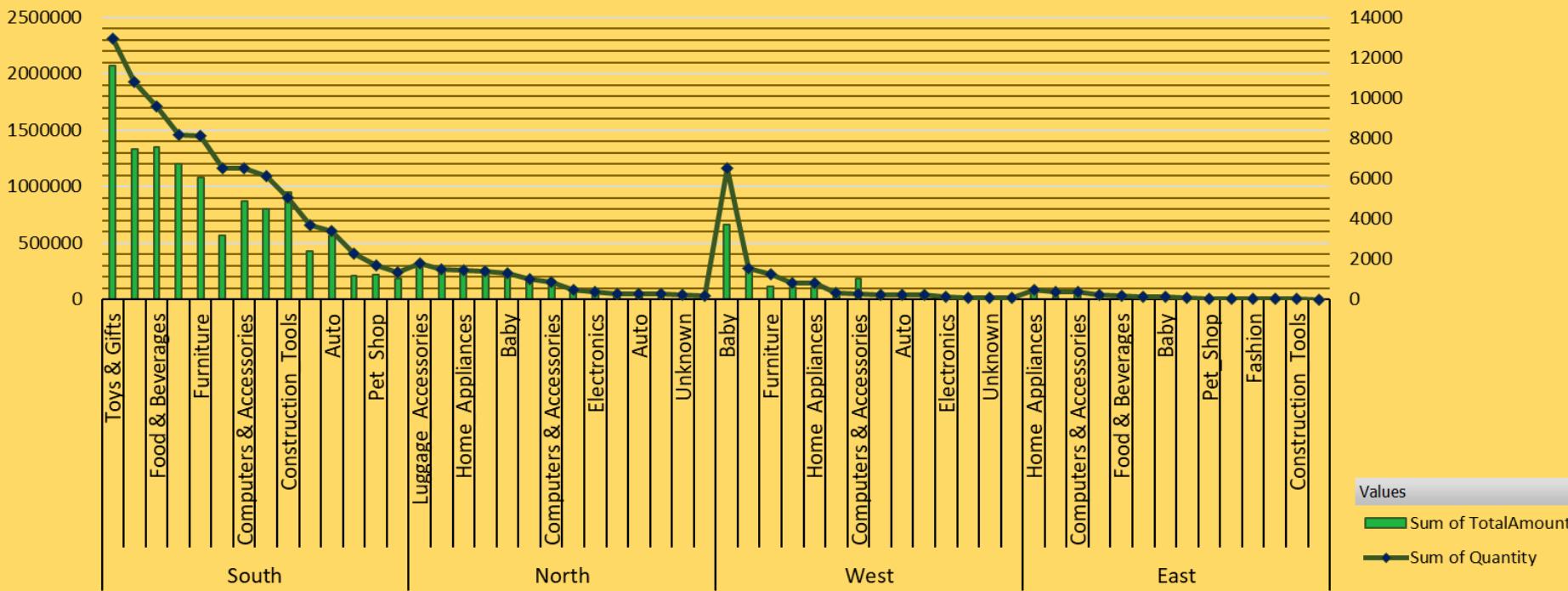
# Category Level Analysis (5/9)

<b>Home_Appliances</b>	<b>3118</b>	<b>391575.06</b>
ST112	216	36495.15
ST120	327	89032.29
ST218	2113	190548.9
ST603	462	75498.72
<b>Food &amp; Beverages</b>	<b>2827</b>	<b>353364.37</b>
ST132	1287	104901.07
ST138	814	95239.72
ST177	430	54706.85
ST414	296	98516.73
<b>Luggage_Accessories</b>	<b>1877</b>	<b>273310.06</b>
ST106	795	150528.33
ST116	486	52410.08
ST253	596	70371.65
<b>Furniture</b>	<b>1513</b>	<b>273068.6</b>
ST180	1513	273068.6
<b>Stationery</b>	<b>894</b>	<b>105123.3</b>
ST555	894	105123.3
<b>Auto</b>	<b>494</b>	<b>83569.46</b>
ST125	494	83569.46
<b>Electronics</b>	<b>107</b>	<b>50962.72</b>
ST166	107	50962.72

# Category Level Analysis (6/9)

Sum of TotalAmount Sum of Quantity

## Popular Categories By Region



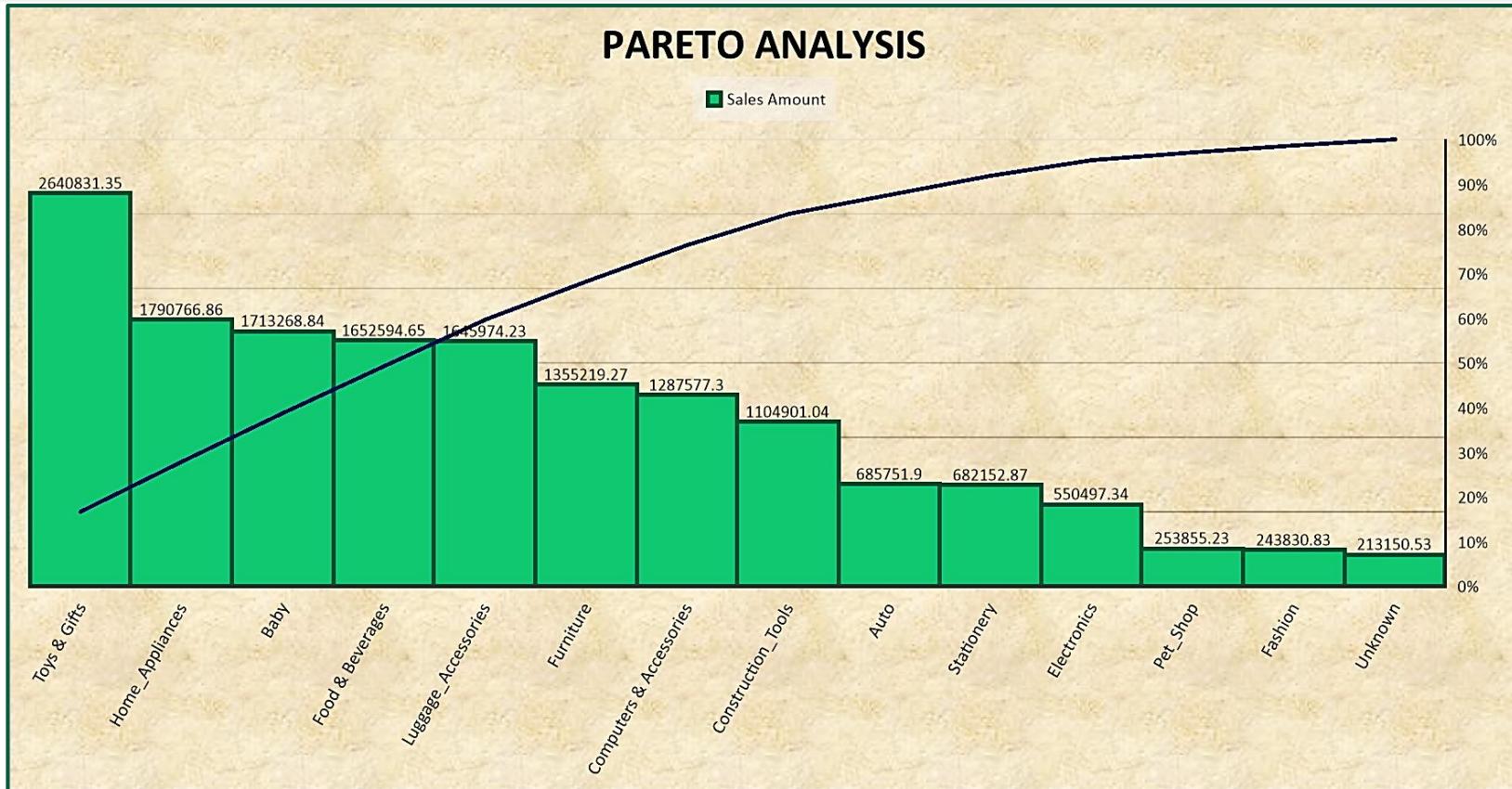
Region ▾ ProductCategory ▾

+ -

# Category Level Analysis (7/9)

Region	Most Popular Category	Quantity Sold	Sales Amount
SOUTH	TOYS & GIFTS	12,954	₹ 20.72 Lakh
WEST	BABY	6,497	₹ 6.65 Lakh
NORTH	LUGGAGE ACCESSORIES	1,779	₹ 3.21 Lakh
EAST	HOME APPLIANCES	462	₹ 0.75 Lakh

# Category Level Analysis (8/9)



# Category Level Analysis (9/9)

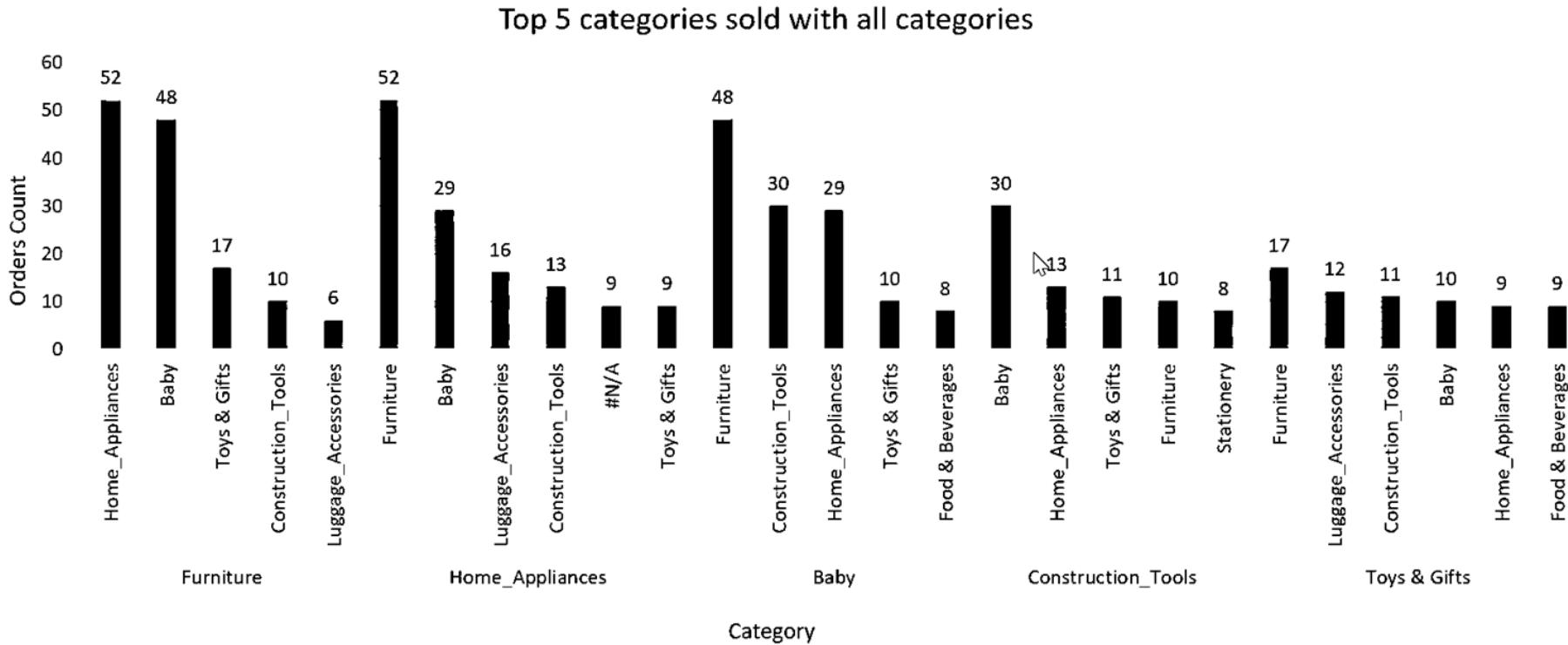
Product Category	Sales Amount	Cumulative Sales	Cumulative % Sales
Toys & Gifts	2640831.35	2640831.35	16.69%
Home_Appliances	1790766.86	4431598.21	28.01%
Baby	1713268.84	6144867.05	38.84%
Food & Beverages	1652594.65	7797461.7	49.29%
Luggage_Accessories	1645974.23	9443435.93	59.69%
Furniture	1355219.27	10798655.2	68.26%
Computers & Accessories	1287577.3	12086232.5	76.40%
Construction_Tools	1104901.04	13191133.54	83.38%
Auto	685751.9	13876885.44	87.72%
Stationery	682152.87	14559038.31	92.03%
Electronics	550497.34	15109535.65	95.51%
Pet_Shop	253855.23	15363390.88	97.11%
Fashion	243830.83	15607221.71	98.65%
Unknown	213150.53	15820372.24	100.00%

## Pareto Principle (80/20 Rule) :

The Pareto Principle, also known as the **80/20 Rule**, states that **roughly 80% of outcomes result from 20% of causes**. From our analysis, it can be observed that around **50% of categories contribute to nearly 80% of total sales**.

# CROSS SELLING ANALYSIS

# Cross Selling of Products (1/2)



Home Appliances and Furniture are the most frequently cross sold product categories, with customers buying both in 52 orders.

# Cross Selling of Products (2/2)

Top 10 combinations of 3 Product Category are selling together in each transaction

Category 1	Category 2	Category 3
Unknown	Pet_Shop	Construction_Tools
Auto	Baby	Construction_Tools
Baby	Unknown	Luggage_Accessories
Baby	Construction_Tools	Home_Appliances
Baby	Furniture	Home_Appliances
Baby	Toys & Gifts	Luggage_Accessories
Computers & Accessories	Construction_Tools	Home_Appliances
Construction_Tools	Baby	Home Appliances
Construction_Tools	Toys & Gifts	Stationery
Food & Beverages	Luggage_Accessories	Computers & Accessories

Each combination of 3 unique product categories occurs only once in order.

# RECOMMENDATIONS

- **Marketing Strategies:**

**Bundle Offers:** Create bundle offers for the top cross-selling categories, such as Home Appliances and Furniture, to encourage customers to purchase these items together.

**Targeted Promotions:** Use targeted promotions for combinations like Baby products and Home Appliances to attract specific customer segments.

- **Customer Insights:**

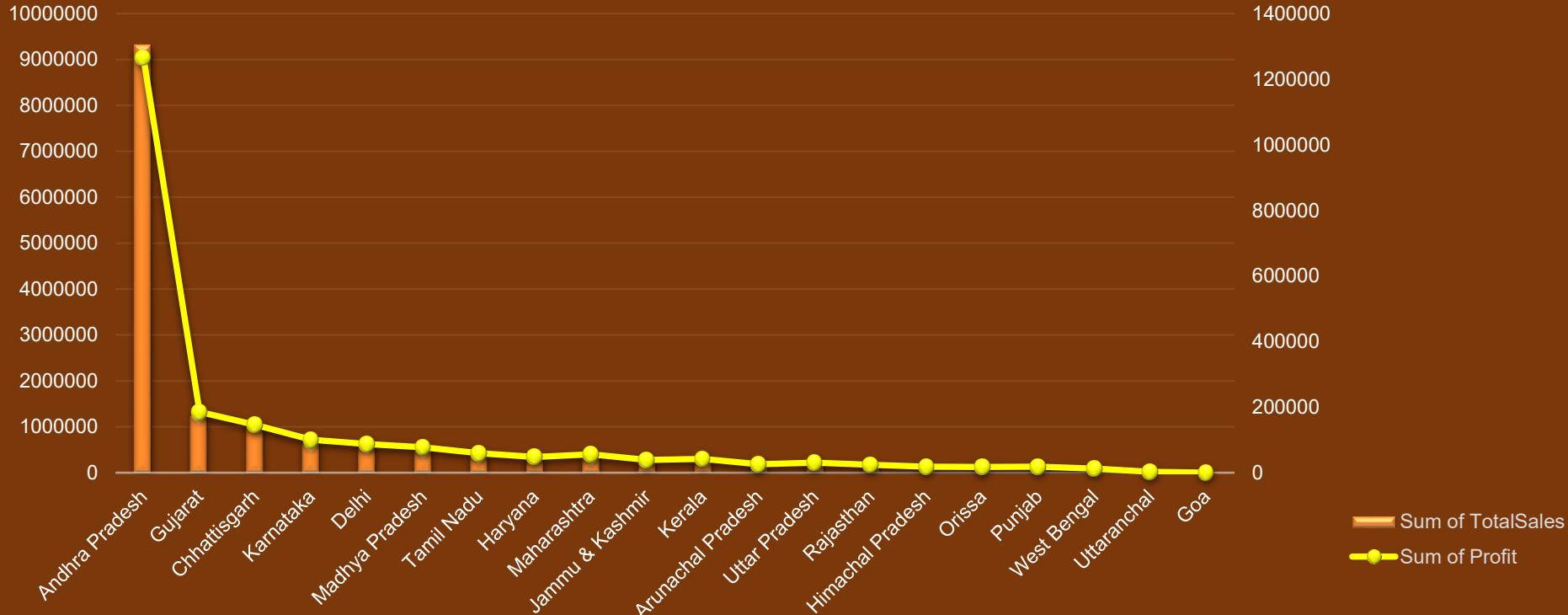
**Personalized Recommendations:** Use the data to provide personalized product recommendations to customers based on their purchase history.

**Customer Segmentation:** Segment customers based on their purchasing patterns and tailor marketing campaigns to each segment.

# SALES TREND ANALYSIS

# Sales Trend Analysis (1/8)

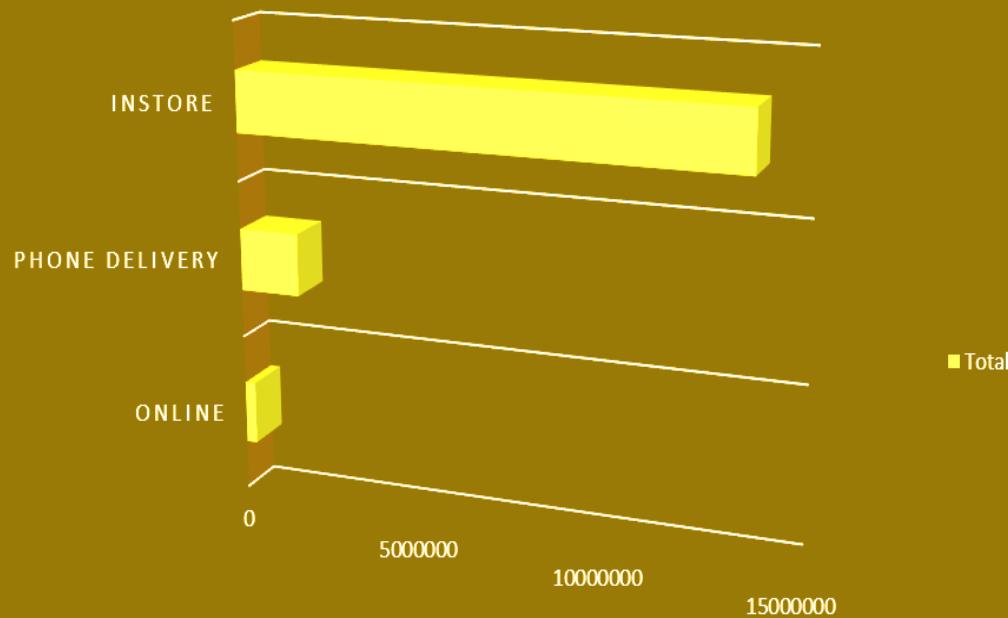
## STATE WISE SALES TREND ANALYSIS



# Sales Trend Analysis (2/8)

Sum of TotalAmount

## CHANNEL WISE SALES TREND

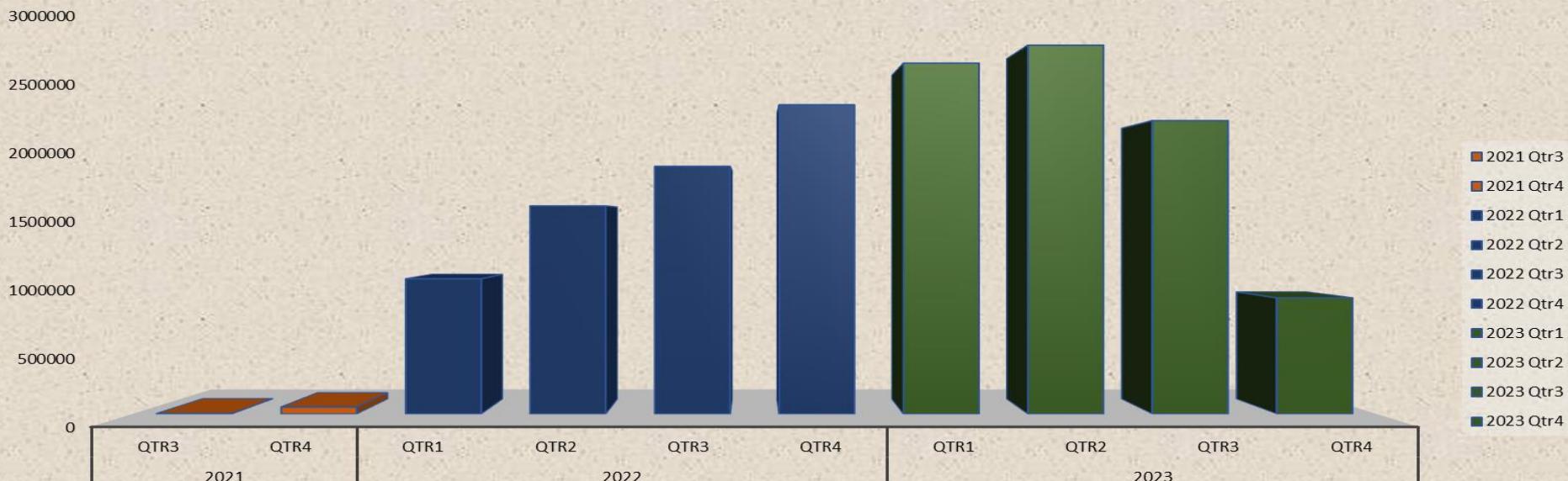


Channels	Sales Amount
Instore	13873917.23
Phone Delivery	1660948.47
Online	285506.54
Grand Total	15820372.24

# Sales Trend Analysis (3/8)

Sum of TotalAmount

## YEARLY - QUARTERLY SALES TREND ANALYSIS



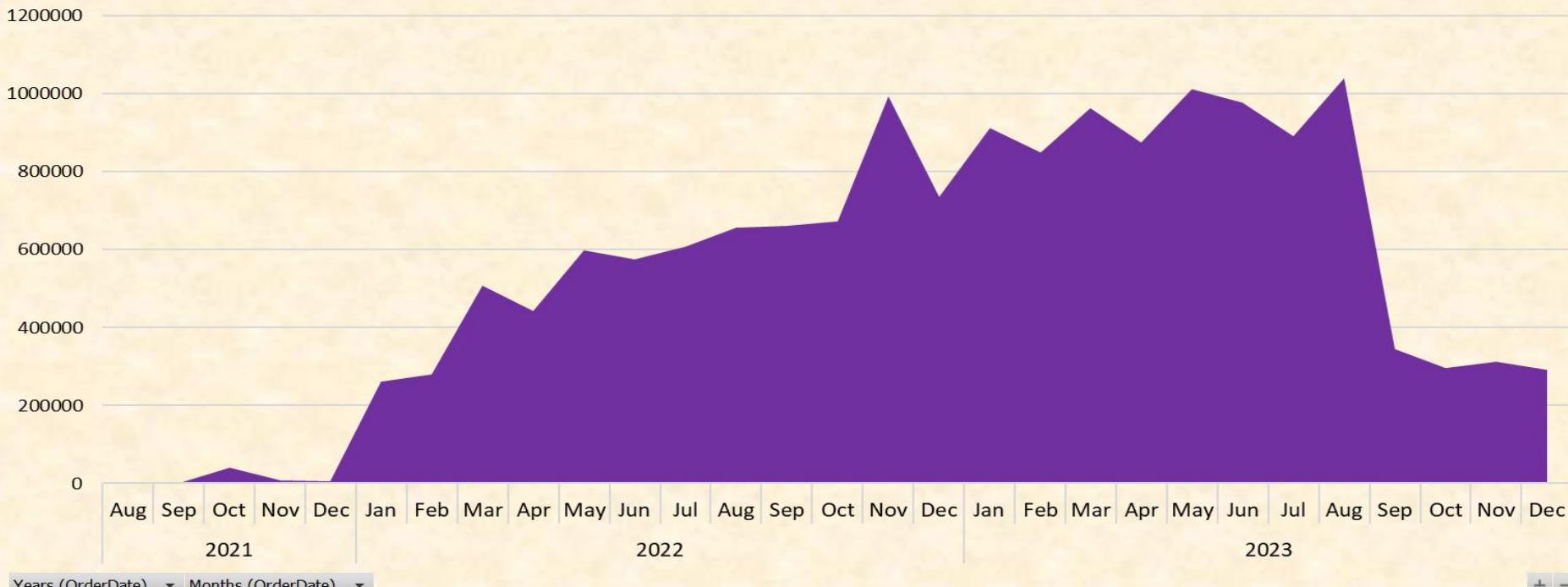
Years (OrderDate) ▾ Quarters (OrderDate) ▾

+ -

# Sales Trend Analysis (4/8)

Sum of TotalAmount

## YEAR - MONTHLY SALES TREND ANALYSIS



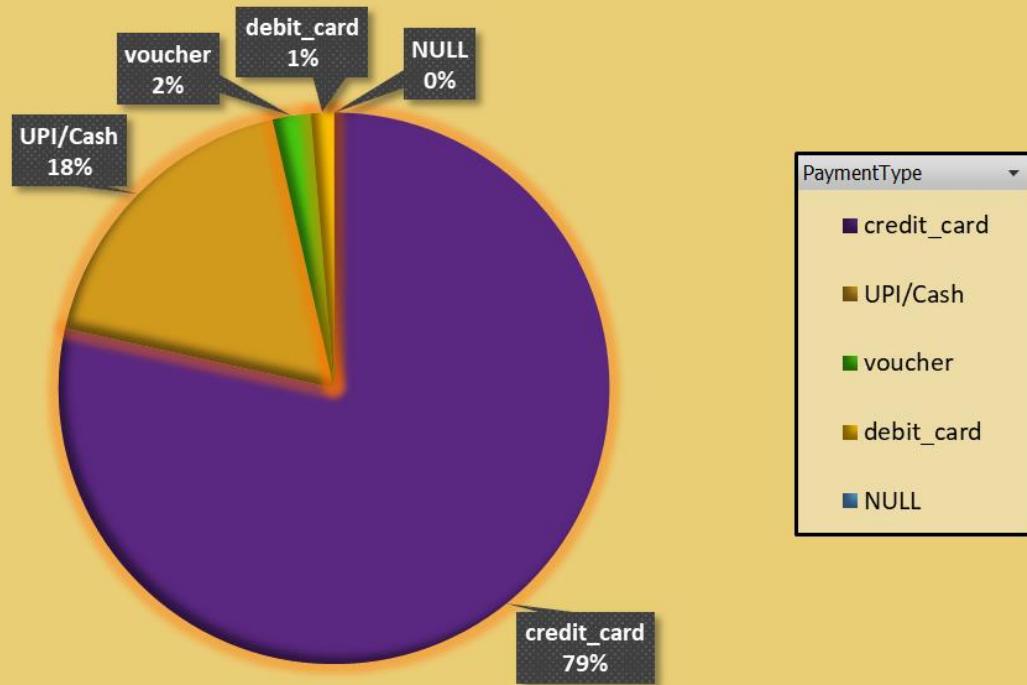
Years (OrderDate) ▾ Months (OrderDate) ▾

+

# Sales Trend Analysis (5/8)

Sales Amount

## PAYMENT METHODS & SALES TREND ANALYSIS

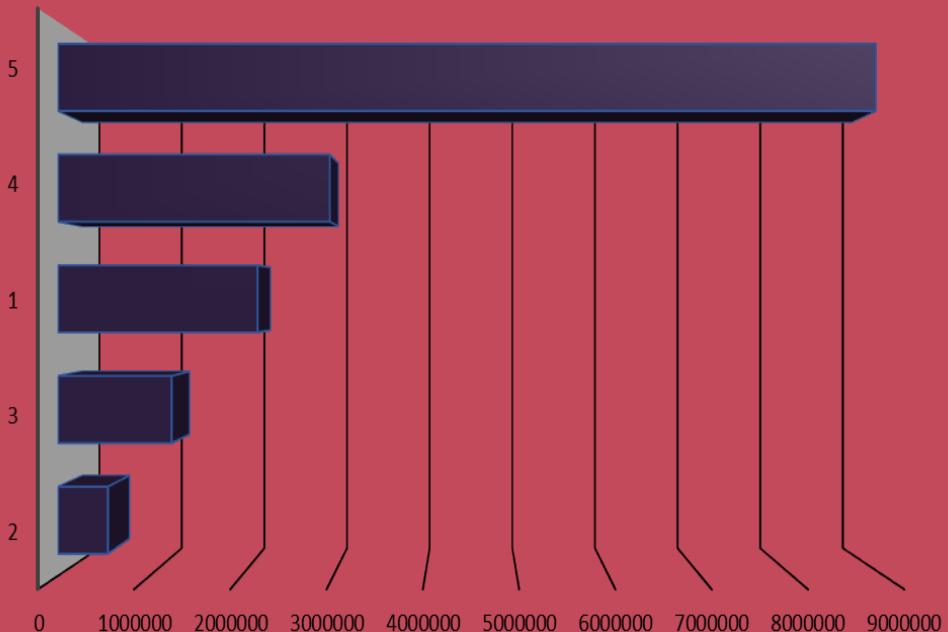


Payment Type	Sales Amount
credit_card	12415882.77
UPI/Cash	2837616.58
voucher	352152.97
debit_card	214576.46
Grand Total	15820372.24

# Sales Trend Analysis (6/8)

Sales Amount

## Customer Satisfaction Score & Sales Trend Analysis

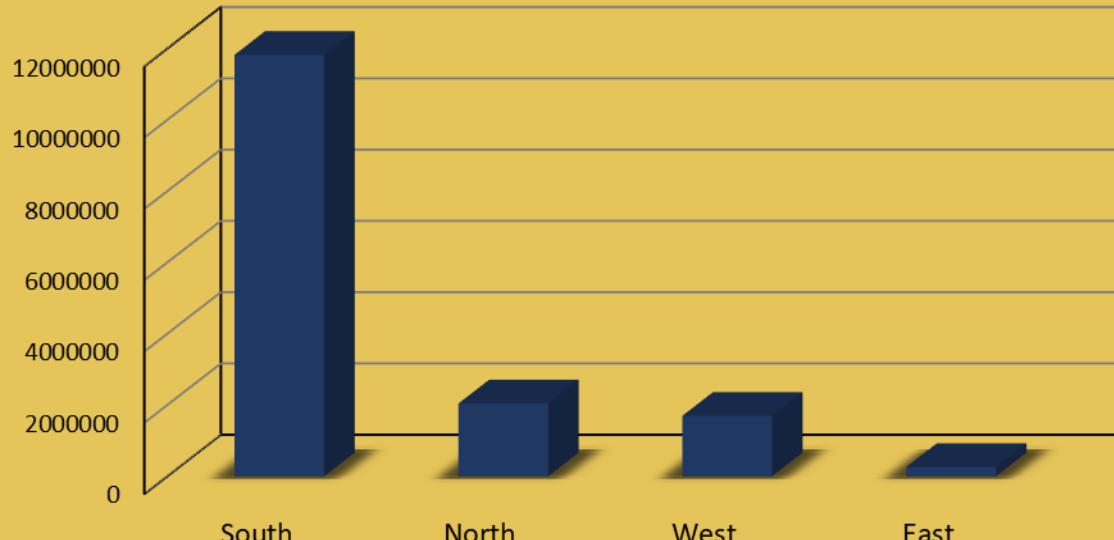


Satisfaction Score	Sales Amount
2	540464.64
3	1235550.14
4	2172595.93
5	2960316.13
Grand Total	8911445.4
	15820372.24

# Sales Trend Analysis (7/8)

Sales Amount

## Region Wise - Sales Trend Analysis



Region ▾

Region	Sales Amount
South	11823568.26
North	2041266.83
West	1695609.58
East	259927.57
Grand Total	15820372.24

# Sales Trend Analysis (8/8)

## Top 10 Most Expensive Products Sorted By Price

ProductID	Product Category	Maximum Price	Total Sales	Contribution To Sales Percentage
489ae2aa008f021502940f251d4cce7f	Home_Appliances	6735	6929.31	4.38%
69c590f7ffc7bf8db97190b6cb6ed62e	Computers & Accessories	6729	6922.21	4.38%
1bdf5e6731585cf01aa8169c7028d6ad	Toys & Gifts	6499	6726.66	4.25%
a6492cc69376c469ab6f61d8f44de961	Luggage_Accessories	4799	4950.34	3.13%
c3ed642d592594bb648ff4a04cee2747	Luggage_Accessories	4690	4764.34	3.01%
259037a6a41845e455183f89c5035f18	Computers & Accessories	4590	4681.78	2.96%
a1beef8f3992dbd4cd8726796aa69c53	Electronics	4399.87	4513.32	2.85%
6cdf8fc1d741c76586d8b6b15e9eef30	Toys & Gifts	4099.99	4175.26	2.64%
dd113cb02b2af9c8e5787e8f1f0722f6	Luggage_Accessories	4059	7990.31	5.05%
6902c1962dd19d540807d0ab8fade5c6	Toys & Gifts	3999.9	4016.91	2.54%

# STORES LEVEL ANALYSIS

# Stores Level Analysis (1/5)

Sum of Total\_Sales

## TOP 10 STORES WITH MAXIMUM SALES



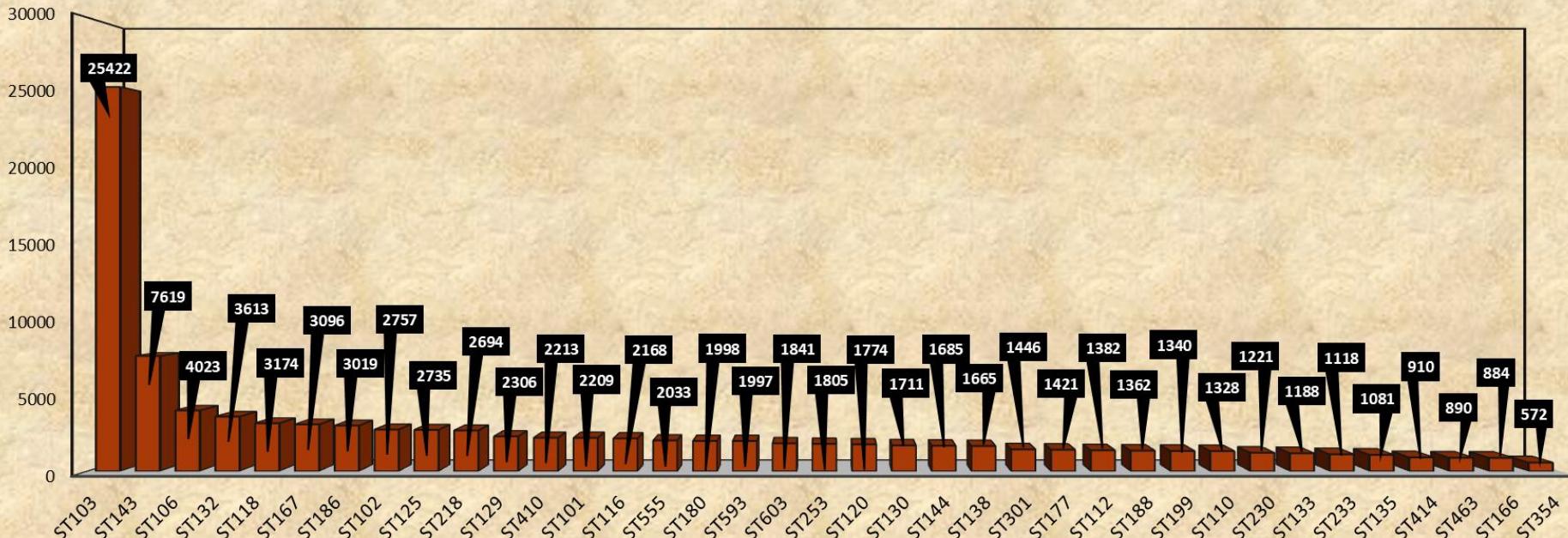
StoreID

Store ID <input type="button" value="▼"/>	Total Sales
ST103	3387650.12
ST143	934233.15
ST106	722928.37
ST102	585010.84
ST125	542533.18
ST410	516239.34
ST167	469434.69
ST180	448258.72
ST132	408140.88
ST118	397590.6
Grand Total	8412019.89

# Stores Level Analysis (2/5)

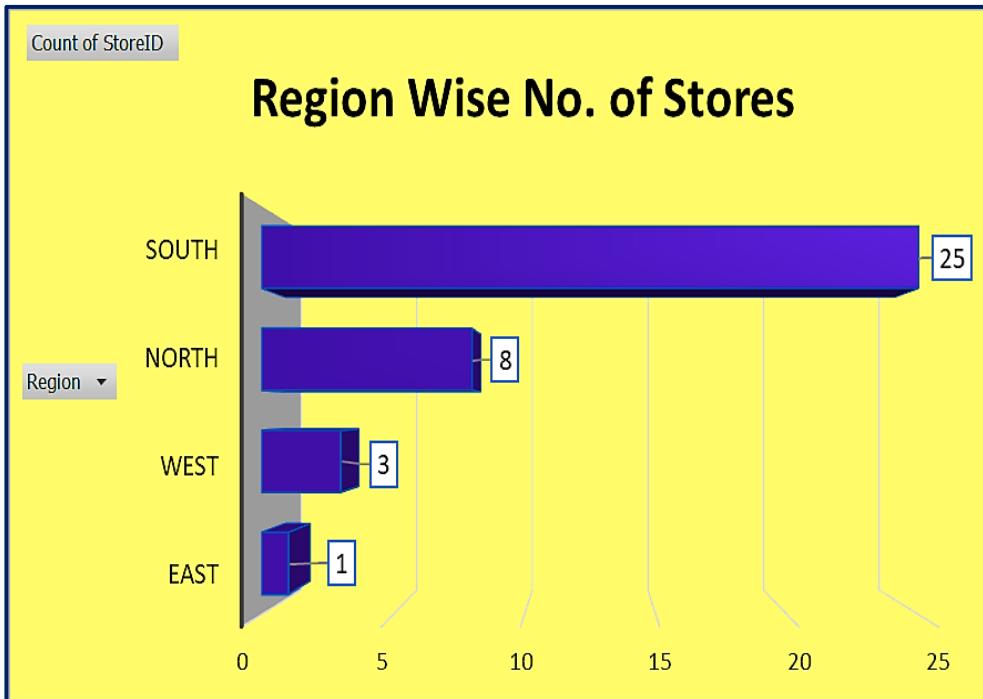
Sum of Distinct\_Customers\_Visited

## Store Wise Distinct Customers Visited



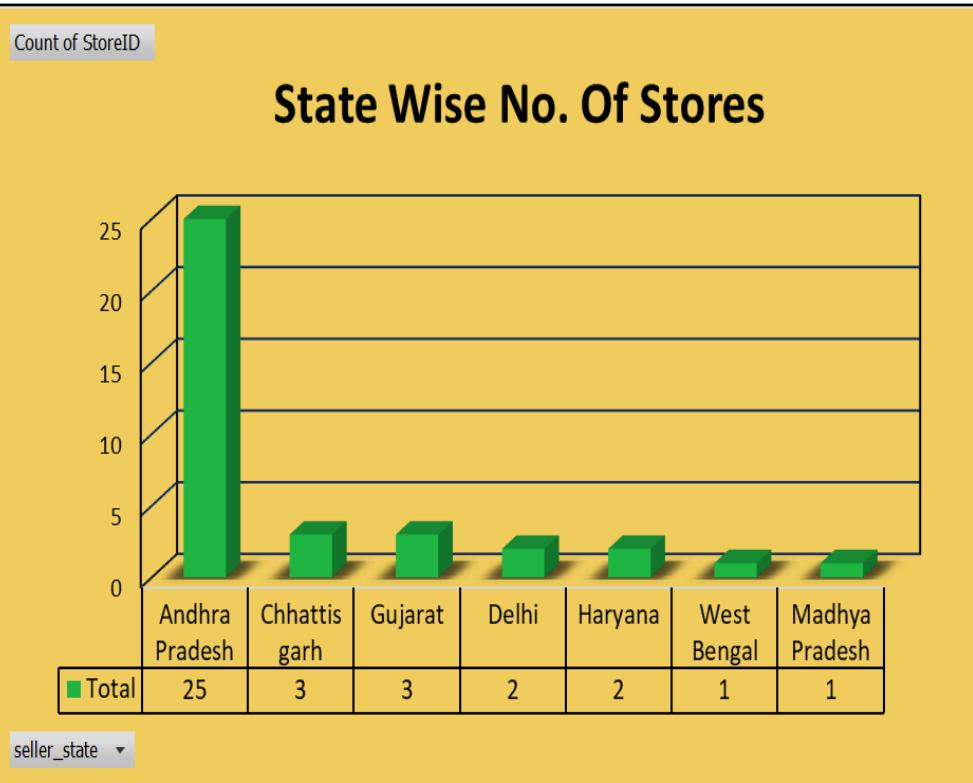
StoreID ▾

# Stores Level Analysis (3/5)



Regions	No. of Stores
South	25
North	8
West	3
East	1
<b>Grand Total</b>	<b>37</b>

# Stores Level Analysis (4/5)

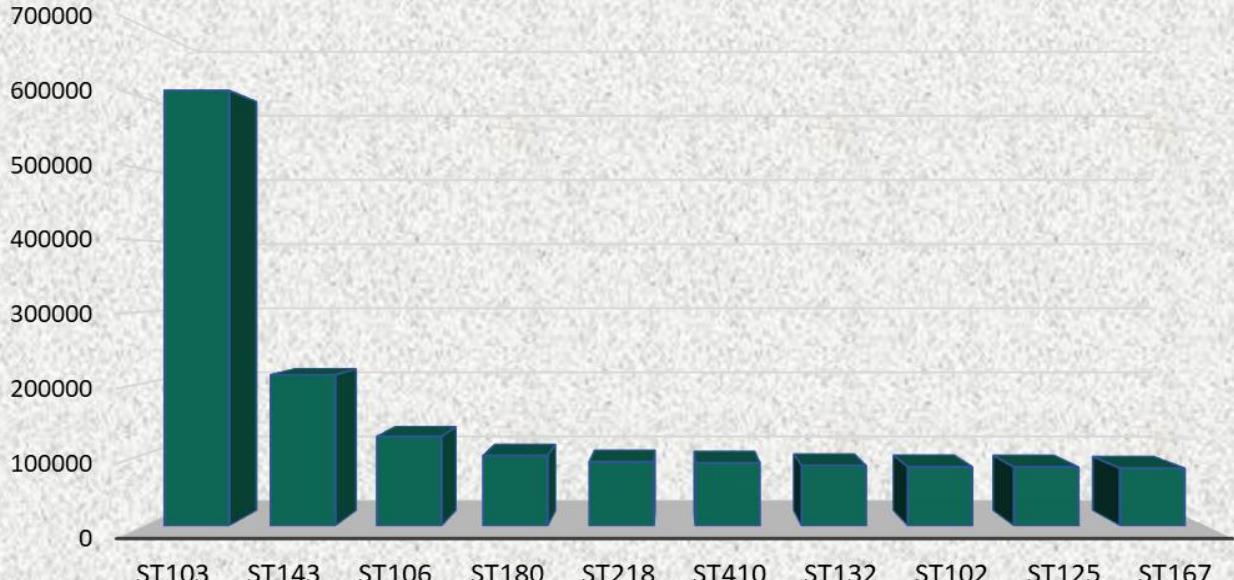


States	No. of Stores
Andhra Pradesh	25
Chhattisgarh	3
Delhi	2
Gujarat	3
Haryana	2
Madhya Pradesh	1
West Bengal	1
<b>Grand Total</b>	<b>37</b>

# Stores Level Analysis (5/5)

Sum of Total\_Profit

## TOP 10 STORES WITH MAXIMUM PROFIT



StoreID

Store ID	Total Profit
ST103	610331.54
ST143	211528.62
ST106	124994.81
ST180	98491.46
ST218	89681.78
ST410	88079.32
ST132	84122.73
ST102	82511.97
ST125	82261.78
ST167	80541.69
Grand Total	1552545.7

# FINAL OBSERVATIONS

- The number of new customers were significantly low in 2020 and 2021, after which the business started to acquire a greater number of new customers which peaked in months May and August.
- Revenue generation increased after December 2021 and kept on escalating till August, 2023 followed by a sudden decline in September 2023.
- By region the highest sales were seen in South and North, while the least in East.
- By state the highest sales were seen Andhra Pradesh and Gujarat, while the least in West Bengal.
- Toys & Gifts, Home appliances and Baby category product contribute to the highest sales amount .
- The stores in southern region especially in Andhra Pradesh have the most bought products which has Toys & Gifts, Home appliances and Food & Beverages categories that are most sought after.

# FINAL RECOMMENDATIONS

- Explore the impact of factors such as industry-specific events, regulatory changes, or competitive dynamics on sales trends.
- Gather customer feedback to understand their needs and preferences. This can help identify product improvements or new opportunities.
- Analyze long-term trends to identify any underlying shifts in customer preferences or market dynamics.
- Align sales strategies with the company's long-term goals and objectives.

# Thank you!