

RETAIL ANALYSIS PROJECT



SCALETRIX.AI

1

BUSINESS CONTEXT

Context

- The client is a leading Retail Chain in India
- The client has provided the Point of Sales data for period Sep 2021 to Oct 2023 for 37 stores out of 534 stores across 7 States

Expected outcome

- To provide data driven analysis of key metrics at product level, customer level, sales level, store level, etc.
- To define CRM/ marketing/ Campaign/ Sales strategies for the upcoming year
- To suggest a strategy to increase sales for the upcoming year



SCALETRIX.AI

Image source: iStock

2

BUSINESS CONTEXT

Problem statement

The retail client is facing challenges in understanding and improving various aspects of their business:

- Identifying sales trends
- Understanding customer buying patterns and preferences
- Understanding product behaviors and store performance
- Analyzing the effectiveness of promotional campaigns
- Identifying customer churn
- Optimizing pricing strategies



SCALETRIX.AI

Image source: SCALETRIX.AI

3

OVERVIEW OF TABLES (1/2)

Total 6 tables have been given in data set, as under:

| Table name | No. of columns | No. of rows | Granularity | Primary key | Foreign key | Parameters/details |
|---------------|----------------|-------------|---------------------|-------------|-------------|---|
| Customer | 4 | 99,441 | Customer level data | Customer id | NA | <ul style="list-style-type: none"> • Customer-id, city, State and Gender. • Data given for total 99,441 unique customers |
| Stores info | 4 | 535 | Store level data | Store id | NA | <ul style="list-style-type: none"> • Store-id, seller city, seller State and Region • Data given for 534 unique stores across 4 region and 19 States/UTs |
| Products info | 9 | 32,951 | Product level data | Product id | NA | <ul style="list-style-type: none"> • Product-id, category, product name length and description length, product photos quantity, product dimensions and weight • Data given for 32,951 unique products over 13 categories while category of some product-id is not given |



SCALETRIX.AI

4

OVERVIEW OF TABLES (2/2)

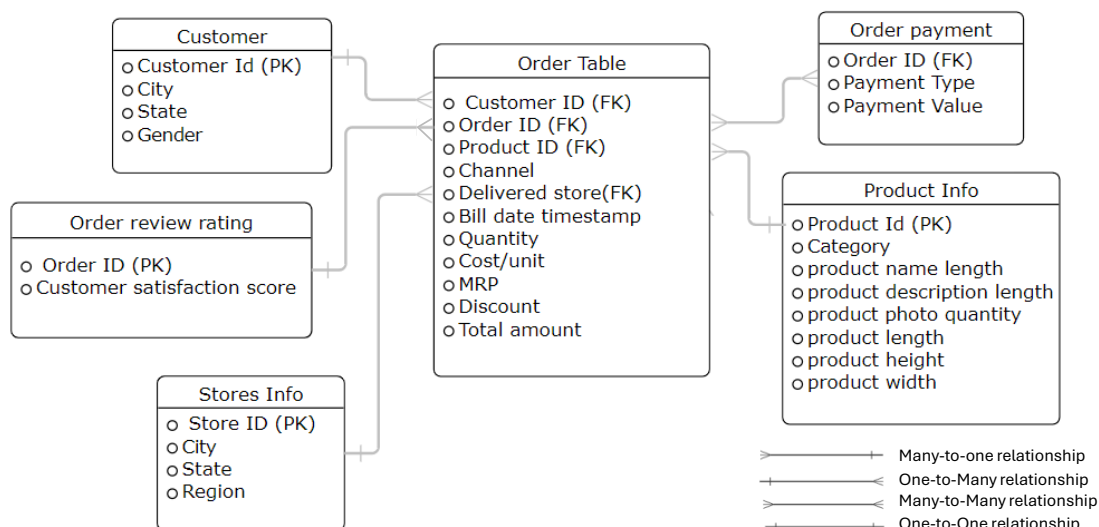
| Table name | No. of columns | No. of rows | Granularity | Primary key | Foreign key | Parameters/details |
|----------------------|----------------|-------------|------------------|-------------|---------------------------------------|---|
| Orders | 11 | 1,12,650 | Order level data | NA | Customer-ID Product-ID Store-ID | <ul style="list-style-type: none"> Customer-id, order-id, product-id, channel, store-id, bill date time stamp, quantity, unit cost, MRP, discount and total amount Data given for 98,666 unique orders for all 32,951 products across 37 stores from 7 States |
| Order payment | 3 | 1,03,886 | Order level data | NA | NA | <ul style="list-style-type: none"> Order-id, payment type and payment value Data has been given for 99,440 unique order regarding their payment details |
| Order review ratings | 2 | 1,00,000 | Order level data | Order ID | NA | <ul style="list-style-type: none"> Order-id, customer satisfaction score Data has been given for 99,441 unique order id including all 98,666 order id given in order table |



SCALETRIX.AI

5

ENTITY RELATIONSHIP DIAGRAM



SCALETRIX.AI

6

OBSERVATIONS

- 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)



SCALETRIX.AI

7

DATA DISCREPANCIES (1/7)

• Inconsistencies in data in 'Orders' table:

- Same product id with same order id has been shown in multiple records and quantity is increasing consecutively. Further, total amount is also in consecutive manner.

| Customer_id | order_id | product_id | Channel | Delivered_StoreID | Bill_date_timestamp | Qty | Cost Per Unit | MRP | Discount | Total Amount |
|-------------|----------------------------------|----------------------------------|---------|-------------------|---------------------|-----|---------------|--------|----------|--------------|
| 1115885858 | 713eda1fb337fff2cccfae60fd0b411e | 781afe929e3016a667f5f439afd55fce | Instore | ST230 | 2/15/2023 16:20 | 1 | 109.9 | 122.17 | 0 | 122.17 |
| 1115885858 | 713eda1fb337fff2cccfae60fd0b411e | 781afe929e3016a667f5f439afd55fce | Instore | ST230 | 2/15/2023 16:20 | 2 | 109.9 | 122.17 | 0 | 244.34 |
| 1115885858 | 713eda1fb337fff2cccfae60fd0b411e | 781afe929e3016a667f5f439afd55fce | Instore | ST230 | 2/15/2023 16:20 | 3 | 109.9 | 122.17 | 0 | 366.51 |
| 1115885858 | 713eda1fb337fff2cccfae60fd0b411e | 781afe929e3016a667f5f439afd55fce | Instore | ST230 | 2/15/2023 16:20 | 4 | 109.9 | 122.17 | 0 | 488.68 |

- One order_id is mentioned against more than one customer id. No. of such orders id: 5.

| Customer_id | order_id | product_id | Channel | Delivered_StoreID | Bill_date_timestamp | Qty | Cost Per Unit | MRP | Discount | Total Amount |
|-------------|----------------------------------|----------------------------------|---------|-------------------|---------------------|-----|---------------|-------|----------|--------------|
| 1687730899 | 001d8f0e34a38c37f7dba2a37d4eba8b | e67307ff0f15ade43fcb6e670be7a74c | Instore | ST103 | 5/18/2022 17:35 | 1 | 18.99 | 26.77 | 0 | 26.77 |
| 7341229049 | 001d8f0e34a38c37f7dba2a37d4eba8b | e67307ff0f15ade43fcb6e670be7a74c | Instore | ST103 | 5/18/2022 17:35 | 2 | 18.99 | 26.77 | 0 | 53.54 |



SCALETRIX.AI

8

DATA DISCREPANCIES (2/7)

- Inconsistencies in data in 'Orders' table (contd.):**

➤ Different store ids have been mentioned for a single order_id. No. of such orders:1,007.

| Customer_id | order_id | product_id | Channel | Delivered_StoreID | Bill_date_timestamp | Qty | Cost Per Unit | MRP | Dis-count | Total Amount |
|-------------|--------------------------------------|--------------------------------------|---------|-------------------|---------------------|-----|---------------|-------|-----------|--------------|
| 6513964772 | 002f98c0f7efd42638e d6100ca699b42 | 880be32f4db1d9f6e2b ec38fb6ac23ab | Instore | ST103 | 10/8/2022 9:30 | 2 | 44.9 | 52.06 | 0 | 104.12 |
| 6513964772 | 002f98c0f7efd42638e d6100ca699b42 | d41dc2f297f52d75d7 8714b378d4068 | Instore | ST301 | 10/8/2022 9:30 | 1 | 8.99 | 41.56 | 0 | 41.56 |

➤ Similarly, for some orders, multiple bill date timestamp is given. No. of such orders: 334.

| Customer_id | order_id | product_id | Channel | Delivered_StoreID | Bill_date_timestamp | Qty | Cost Per Unit | MRP | Dis-count | Total Amount |
|-------------|--------------------------------------|--------------------------------------|---------|-------------------|---------------------|-----|---------------|-------|-----------|--------------|
| 2002461216 | 01cce1175ac3c4a450 e3a0f856d02734 | 415dfa57292b8b7 360d3f4cf2f9bff06 | Instore | ST133 | 7/25/2023 20:43 | 1 | 60 | 94.24 | 9 | 85.24 |
| 2002461216 | 01cce1175ac3c4a450 e3a0f856d02734 | 9d0aa87e8df1bdbe0f7 9353520a2d538 | Instore | ST112 | 7/23/2023 20:43 | 2 | 36.23 | 38.50 | 0 | 77.00 |



SCALETRIX.AI

9

DATA DISCREPANCIES (3/7)

- Inconsistencies in data in 'Orders' table (contd.):**

➤ Same order id with same product id has been reported in two records but MRP is different but there is no difference in the cost and in amount also due to different discounts.

| Customer_id | order_id | product_id | Channel | Delivered_Store_ID | Bill_date_timestamp | Qty | Cost Per Unit | MRP | Discount | Total Amount |
|-------------|--------------------------------------|--------------------------------------|---------|--------------------|---------------------|-----|---------------|-------|----------|--------------|
| 3185511294 | 0008288aa423d2a3f00f cb17cd7d8719 | 368c6c730842d78016ad 823897a372db | Instore | ST218 | 2/21/2023 2:55 | 1 | 49.9 | 71.27 | 8 | 63.27 |
| 3185511294 | 0008288aa423d2a3f00f cb17cd7d8719 | 368c6c730842d78016ad 823897a372db | Instore | ST218 | 2/21/2023 2:55 | 2 | 49.9 | 73.27 | 10 | 126.54 |

- Inconsistencies in data while mapping the tables:**

➤ For some orders 'Total_Amount' in the 'Orders' table is not matching with 'Payment_value' in the 'OrderPayments' table for that order id. No. of such orders: 576.

| order_id | No_of_Products_Purchased | total_quantity | total_cost | total_MRP | total_discount | total_amount | total_profit | no_of_payment_type | total_payment |
|--------------------------------------|--------------------------|----------------|------------|-----------|----------------|--------------|--------------|--------------------|---------------|
| 00789ce015e7e5791c7914f32bb4fa d4 | 1 | 1 | 154 | 168.83 | 0 | 168.83 | 14.83 | 1 | 190.81 |



SCALETRIX.AI

10

DATA DISCREPANCIES (4/7)

- Inconsistencies in data while mapping the tables (contd.):**

- In 'Orders' table, 98,666 unique order id is provided, while in 'OrderPayments' table, payment details is provided for 99,440 unique order and in 'OrderReview_Ratings' table, reviews are provided for 99,441 unique orders. On further reconciliation, the following has been observed:
 - Payment details for order id **bfb0f9bdef84302105ad712db648a6c** (in 'Orders' table) is not provided in the 'OrderPayments' table.
 - Payment details for 775 orders are provided in 'OrderPayments' table, which are not in 'Orders' table.
 - 'Customer_Satisfaction_Score' for 775 orders are provided in 'OrderReview_Ratings' table, which are not in 'Orders' table
- Customers are from 20 different States, while stores are in 19 states. Customers of Goa haven't got any store in their state.



SCALETRIX.AI

11

DATA DISCREPANCIES (5/7)

- In 'OrderPayments' table, for some orders, payments have been done in more than one parts. No. of such orders: 4,446.

| order_id | payment_type | payment_value |
|----------------------------------|--------------|---------------|
| 002f19a65a2ddd70a090297872e6d64e | voucher | 44.11 |
| 002f19a65a2ddd70a090297872e6d64e | voucher | 33.18 |

Similarly, for order id **fa65dad1b0e818e3ccc5cb0e39231352**, payment details is provided 29 times

- In 'OrderReview_Ratings' table, for some orders, customer ratings is provided more than once. No. of such orders: 559.

| order_id | Customer_Satisfaction_Score |
|----------------------------------|-----------------------------|
| 03c939fd7fd3b38f8485a0f95798f1f6 | 3 |
| 03c939fd7fd3b38f8485a0f95798f1f6 | 4 |
| 03c939fd7fd3b38f8485a0f95798f1f6 | 3 |



SCALETRIX.AI

12

DATA DISCREPANCIES (6/7)

- Duplicate data:**

- In 'Store Info' table, one duplicate store id is given (*Store id – ST410*)

- Data type discrepancy:**

- In 'Orders' table, the format for data in bill date timestamp column is not uniform. Some dates are provided in date format while some are in text format.

- Special characters and missing data:**

- In 'ProductsInfo' table, for some product id, #NA is mentioned against the product category. For the product with #NA category, the data for some other fields of these products such as product name length, product description length and photos quantity are missing



SCALETRIX.AI

13

DATA DISCREPANCIES (7/7)

- Total Amount of one order id in 'Orders' table

| Customer_id | order_id | product_id | Channel | Delivered StoreID | Bill_date_timestamp | Qty | Cost Per Unit | MRP | Dis-count | Total Amount | Net qty | Net amount |
|-------------|----------------------------------|----------------------------------|----------------|-------------------|---------------------|-----|---------------|-------|-----------|--------------|---------|------------|
| 6927074336 | 00bcee890eba57a9767c7b5ca12d3a1b | 6c90c0f6c2d89eb816b9e205b9d6a36a | Phone Delivery | ST186 | 7/26/2022 21:05 | 1 | 165.5 | 181.3 | 0 | 181.3 | 1 | 181.3 |
| 6927074336 | 00bcee890eba57a9767c7b5ca12d3a1b | b7d94dc0640c7025dc8e3b46b52d8239 | Phone Delivery | ST233 | 7/26/2022 21:05 | 2 | 175.91 | 246.6 | 18 | 457.2 | 1 | 228.6 |
| 6927074336 | 00bcee890eba57a9767c7b5ca12d3a1b | d143bf43abb18593fa8ed20cc990ae84 | Phone Delivery | ST186 | 7/26/2022 21:05 | 3 | 165.5 | 181.3 | 0 | 543.9 | 1 | 181.3 |
| 6927074336 | 00bcee890eba57a9767c7b5ca12d3a1b | 55939df5d8d2b853fbc532bf8a00dc32 | Phone Delivery | ST186 | 7/26/2022 21:05 | 4 | 165.5 | 181.3 | 0 | 725.2 | 1 | 181.3 |
| Total | | | | | | | | | | 1907.6 | 4 | 772.5 |

- 'Payment_value' of the same order id in 'OrdersPayment' table

| order_id | payment_type | payment_value |
|----------------------------------|--------------|---------------|
| 00bcee890eba57a9767c7b5ca12d3a1b | credit_card | 772.5 |

It can be seen from above tables that the payment value is matching with sum of 'Net amount' and not with sum of 'total_amount'. When we take **net qty=1** and **net_amount = (MRP – discount)* Net_qty**, the values match.



SCALETRIX.AI

14

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.



SCALETRIX.AI

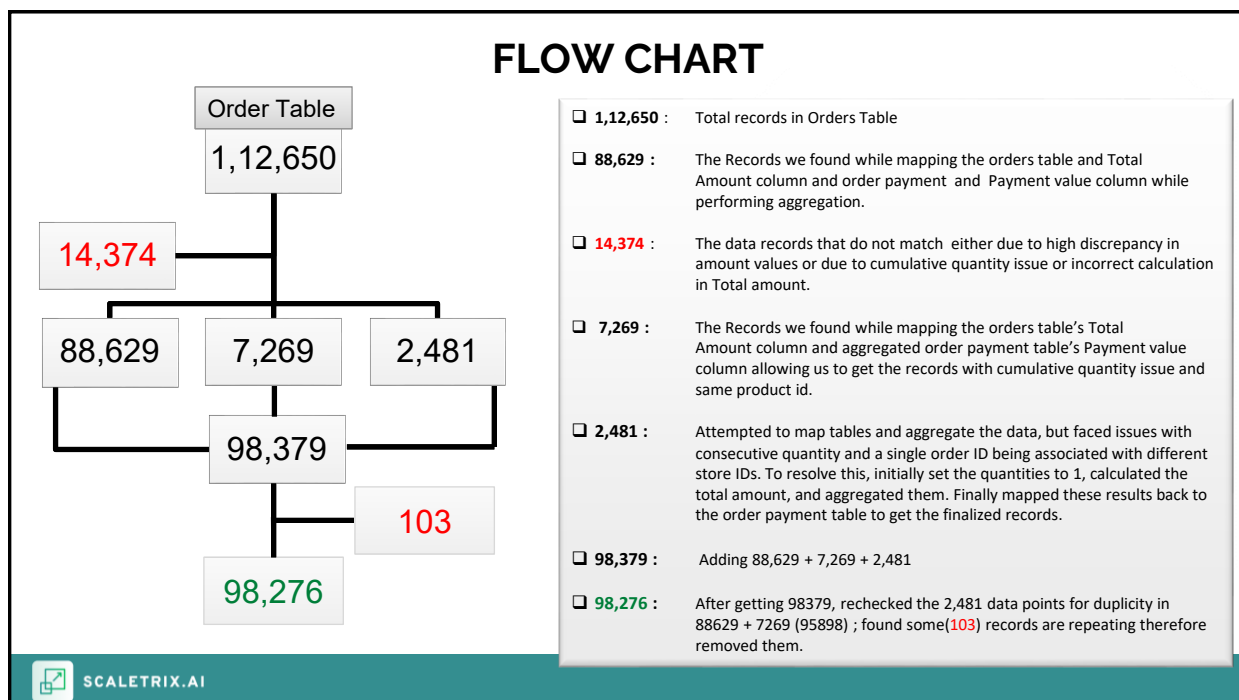
15

DATA CLEANING

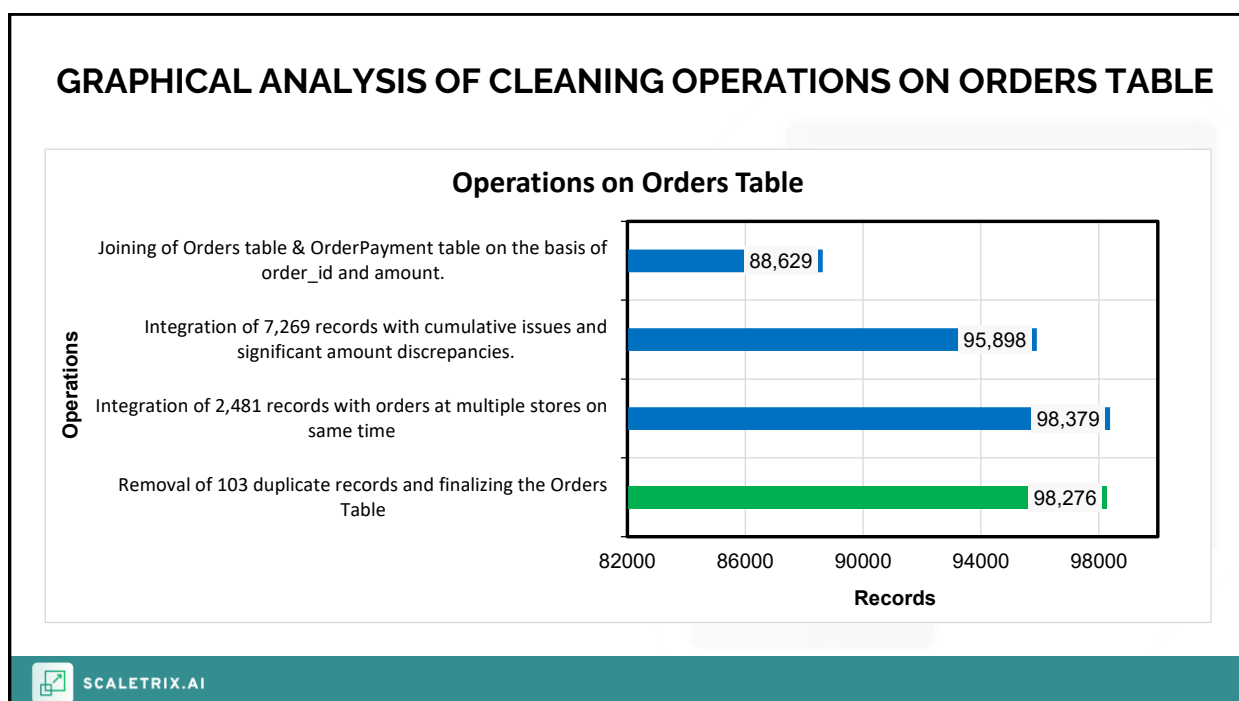


SCALETRIX.AI

16



17



18

EXPLORATORY DATA ANALYSIS



SCALETRIX.AI

19

EXPLORATORY DATA ANALYSIS-1

Sales Information

Total Revenue: ₹ 15.45M

Total Cost: ₹ 13.27M

Total Discount: ₹ 5.38 L

Total Profit: ₹ 2.18M

Percentage of Discount: 3.07%

Percentage of Profit: 16.46%

Order Information

Order Count: 96,896

Total Payment Method: 4

Total channel: 3

Average Discount Per Order: ₹ 5.56

Average Order Value: ₹ 159.49

Average Categories per order: 1.004

Average Items Per Order: 1.12



SCALETRIX.AI

Image source: iStock

20

EXPLORATORY DATA ANALYSIS-1

Customer Information

Total Customer: 96,805

Average profit per customer: ₹ 22.5

Average discount per customer: ₹ 5.56

Transaction per customer: 1.0009

Average sales per customer: ₹159.72

Avg no. of days between two transactions: 4

Product Information

Total product: 31,757

Total category: 14

Seller store information

Total quantity: 1,08,881

Total states: 7

Total store: 37

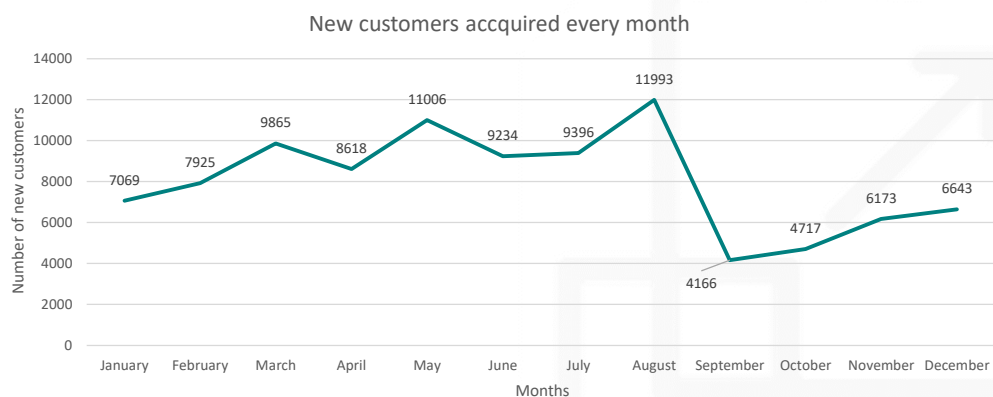
Total region: 4



SCALETRIX.AI

21

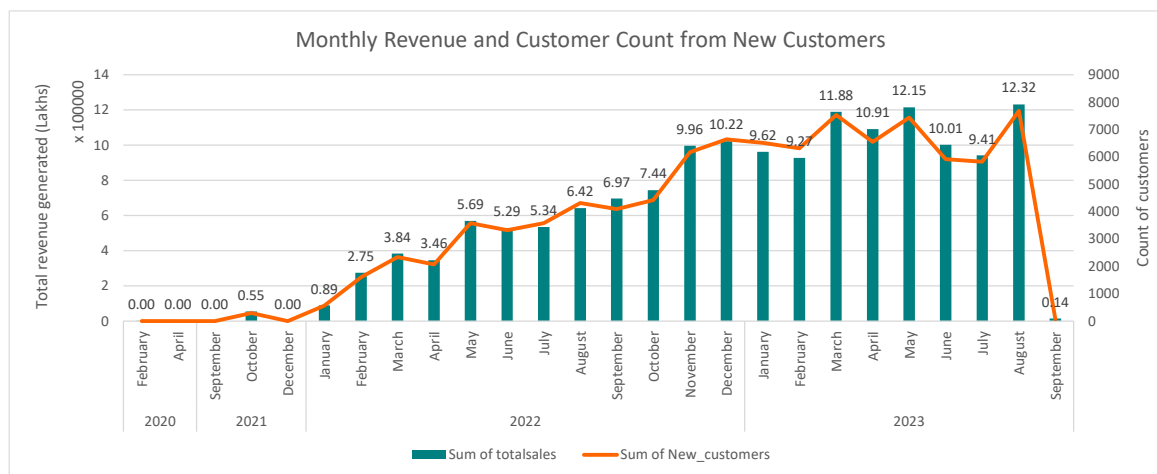
EXPLORATORY DATA ANALYSIS



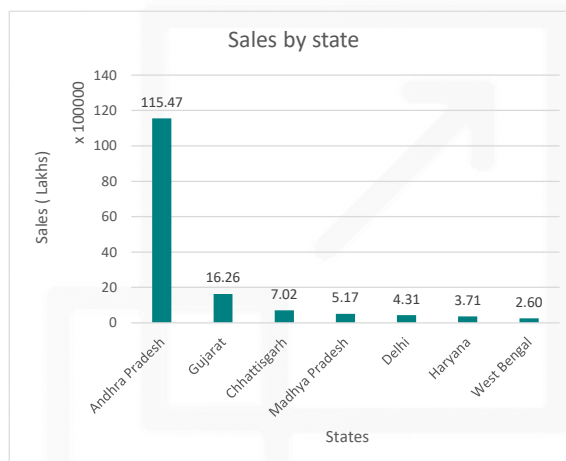
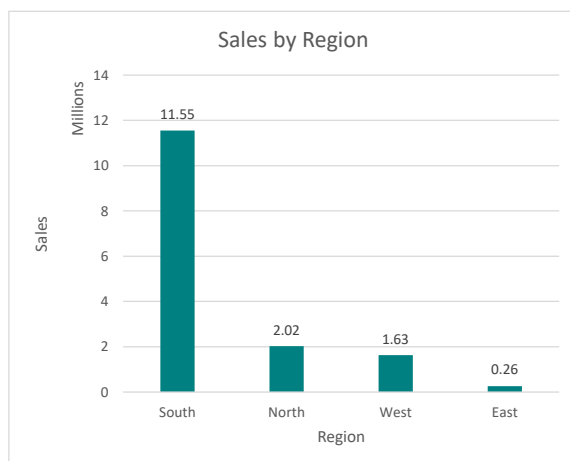
SCALETRIX.AI

22

EXPLORATORY DATA ANALYSIS



EXPLORATORY DATA ANALYSIS



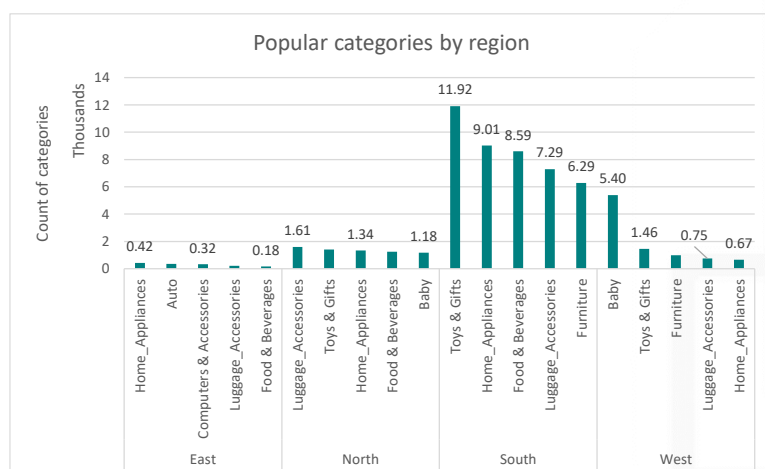
EXPLORATORY DATA ANALYSIS



SCALETRIX.AI

25

EXPLORATORY DATA ANALYSIS



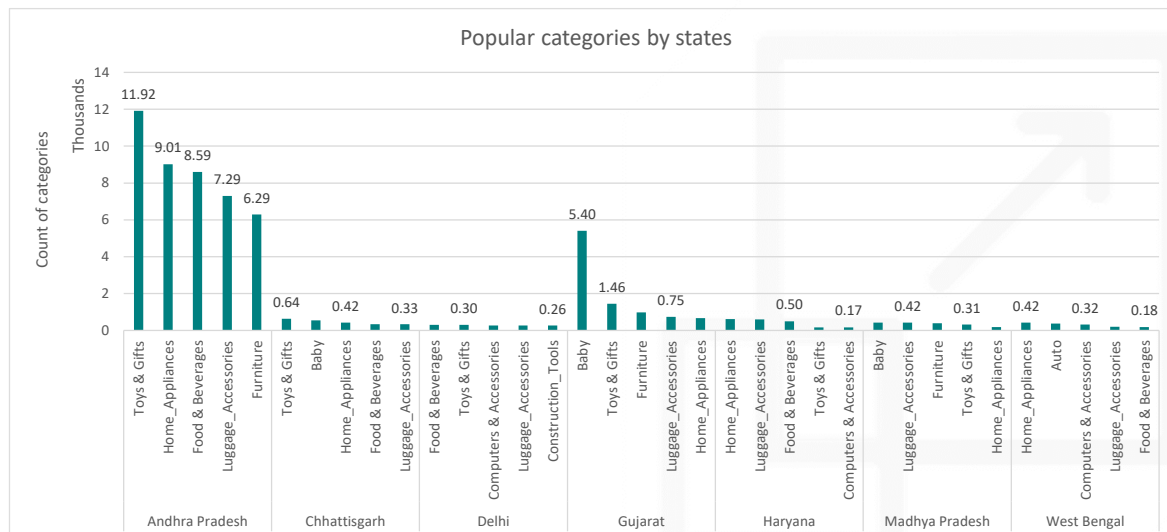
| Region | Category | Orders | Sales |
|--------|---------------------|--------|--------|
| South | Toys & Gifts | 11,919 | 20.38L |
| West | Baby | 5402 | 6.18L |
| North | Luggage Accessories | 1605 | 3.17L |
| East | Home Appliances | 422 | 0.73L |



SCALETRIX.AI

26

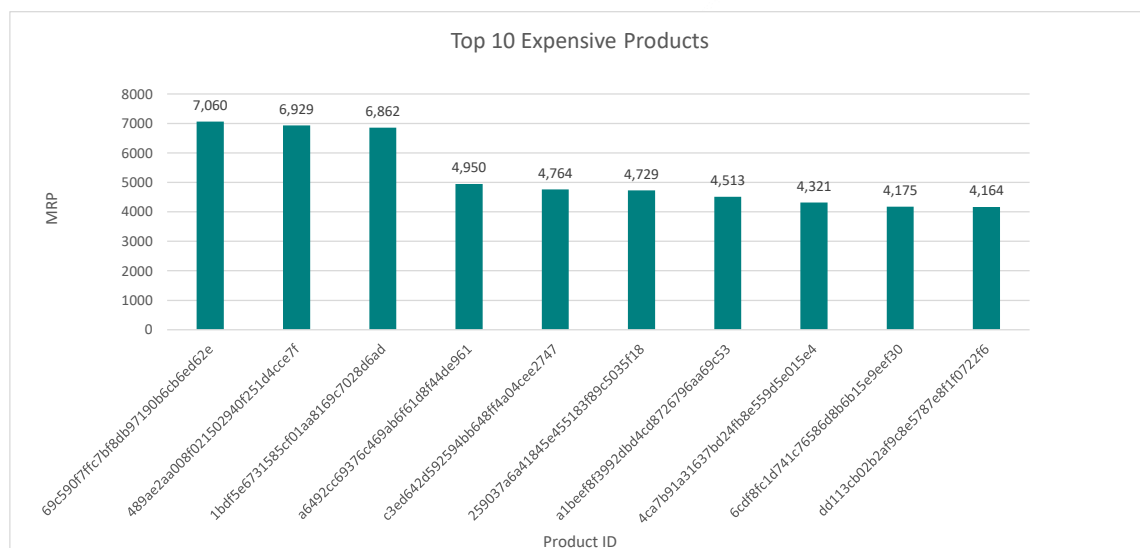
EXPLORATORY DATA ANALYSIS



SCALETRIX.AI

27

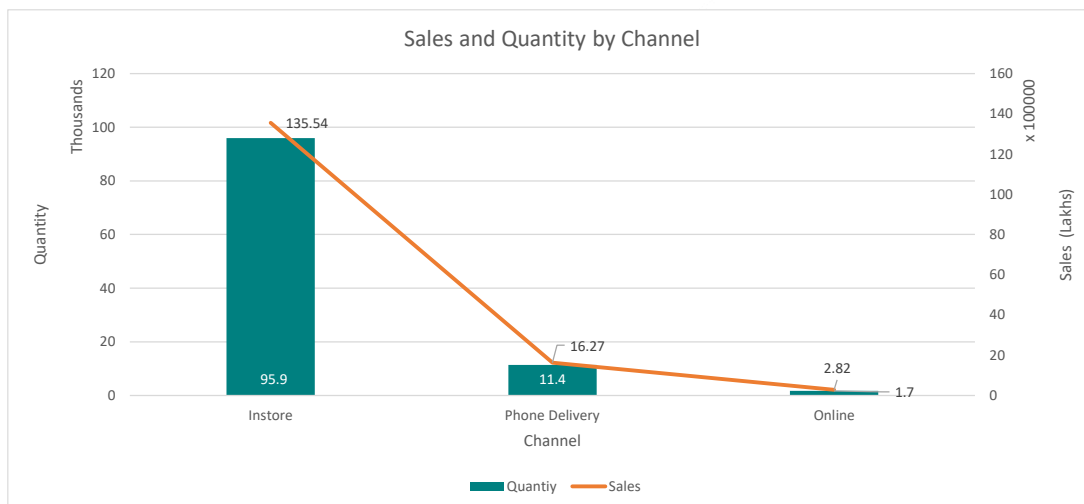
EXPLORATORY DATA ANALYSIS



SCALETRIX.AI

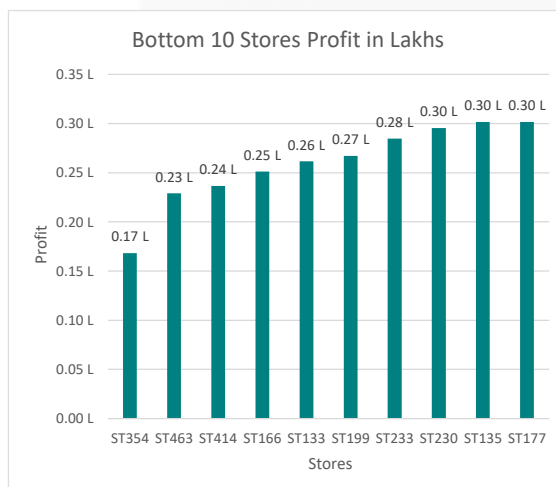
28

EXPLORATORY DATA ANALYSIS



29

EXPLORATORY DATA ANALYSIS



30

EXPLORATORY DATA ANALYSIS

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.



SCALETRIX.AI

31

CUSTOMER BEHAVIOUR

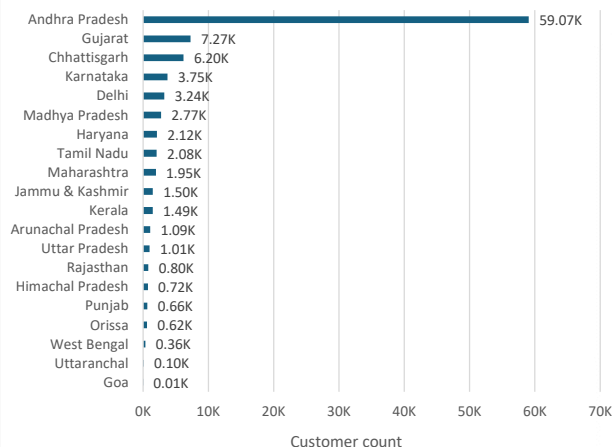


SCALETRIX.AI

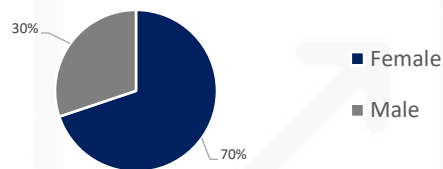
32

CUSTOMER BEHAVIOUR

Number of customers in each state



Gender Count



| Gender | No Of Customers | Total Sales | Sales Percentage |
|--------|-----------------|-------------|------------------|
| Female | 67,610 | 10.81 M | 69.92 % |
| Male | 29,195 | 4.65 M | 30.08 % |

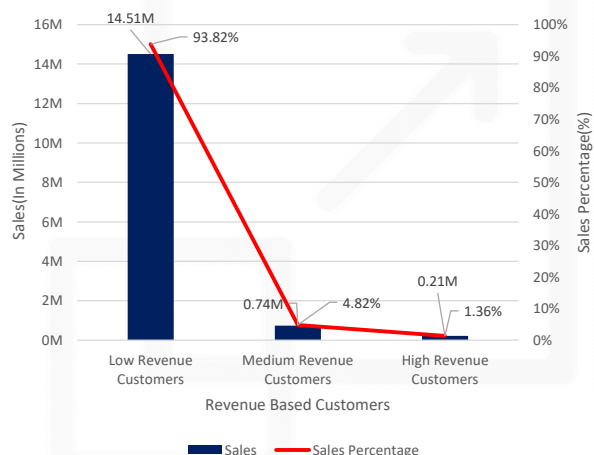
CUSTOMER BEHAVIOUR

Customers Based On The Revenue

| Types Of Customers | No Of Customers | Percentage |
|--------------------------|-----------------|------------|
| High Revenue Customers | 50 | 0.05% |
| Medium Revenue Customers | 380 | 0.39% |
| Low Revenue Customers | 96375 | 99.56% |

- High Revenue Customers Having Revenue (≥ 3 K)
- Medium Revenue Customers Having Revenue (≥ 1.5 K)
- Low revenue customers Having Revenue (< 1.5 K)

Sales Distribution Based On Customer



CUSTOMER BEHAVIOUR

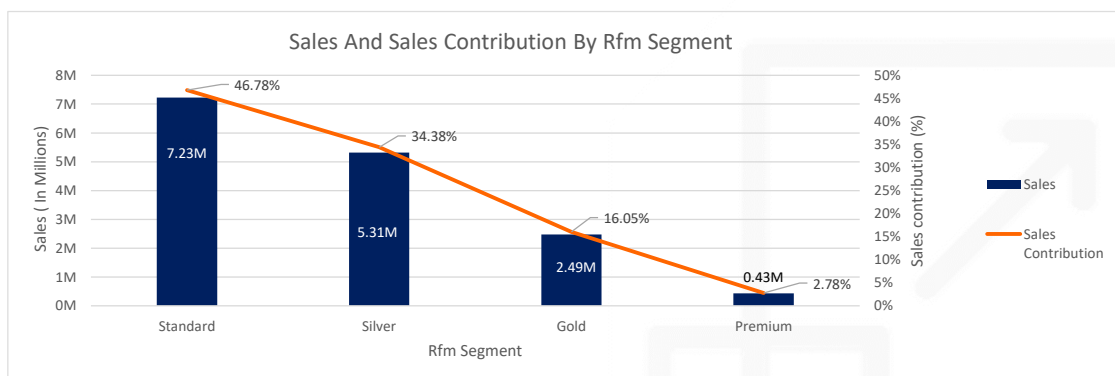
RFM SEGMENTATION

- **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).
- **R_Score:** A score for recency, where the most recent buyers get higher scores.
- **F_Score:** A score for frequency, where customers who have made more purchases get higher scores.
- **M_Score:** A score for monetary value, where customers who have spent more money get higher scores.
- **Total_scoring:**
 - This step combines the R, F, and M scores into a **Total_score** by summing R_Score, F_Score, and M_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).



35

RFM SEGMENTATION



| Customers Segments | No Of Customers | Percentage |
|--------------------|-----------------|------------|
| Standard | 36,263 | 37.46 % |
| Silver | 24,055 | 24.85 % |
| Gold | 24,132 | 24.93 % |
| Premium | 12,355 | 12.76 % |

Based on the **Total score**, customers are classified into different segments:

- **Premium:** Total score between 11 and 12
- **Gold:** Total score between 09 and 10.
- **Silver:** Total score between 08 and 07
- **Standard:** Total score between 03 and 06



36

CUSTOMER BEHAVIOUR

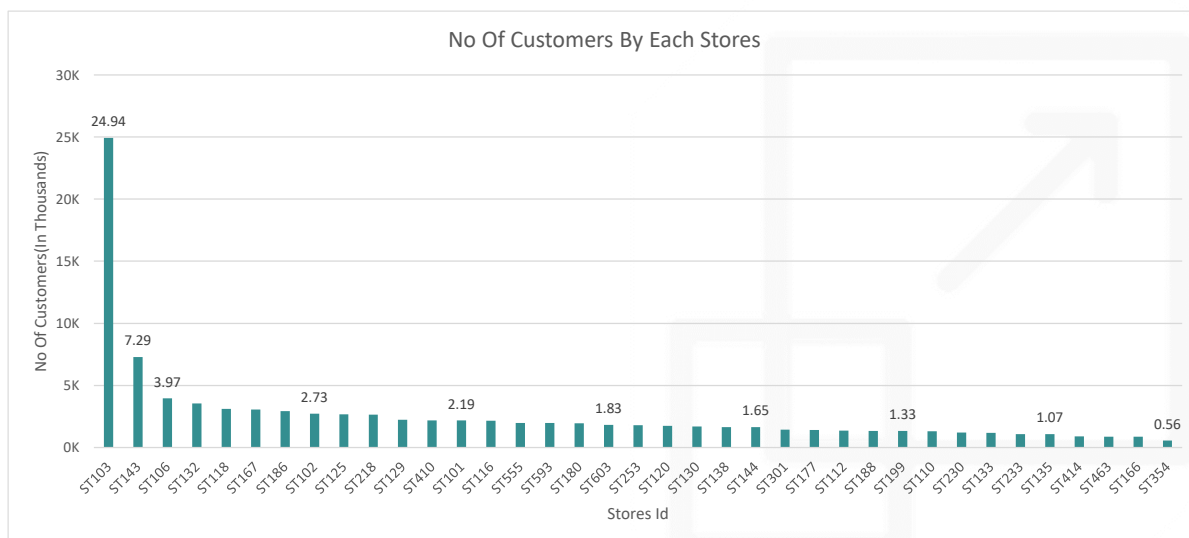
| Types Of Customers | No Of Customers | Percentage |
|--|-----------------|------------|
| One-time Buyers | 96,769 | 99.96 % |
| Repeat Buyers | 36 | 0.04 % |
| Discount Seeker Customers | 39,084 | 40.37% |
| Non-discount Seeker Customers | 57,721 | 59.63% |
| Number Of Customers Purchasing From A Single Category | 96,566 | 99.49% |
| Number Of Customers Purchasing From A Multiple Category | 488 | 0.51% |
| Customers who made purchases across all available channels: Online, Instore, and Phone Delivery | 1 | 0.001 % |



SCALETRIX.AI

37

BEHAVIOUR OF ONE TIME BUYERS

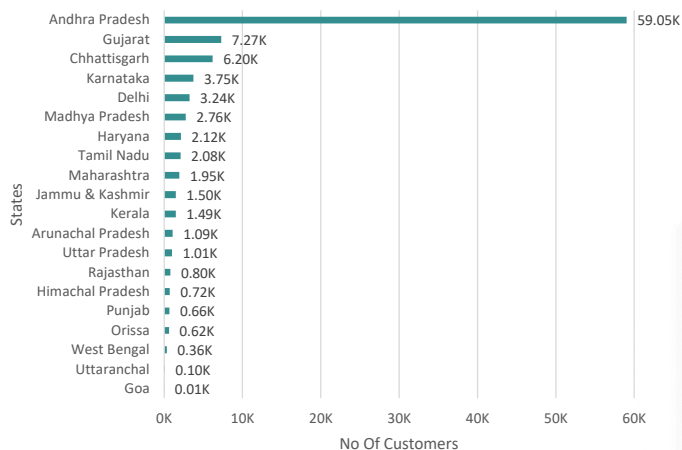


SCALETRIX.AI

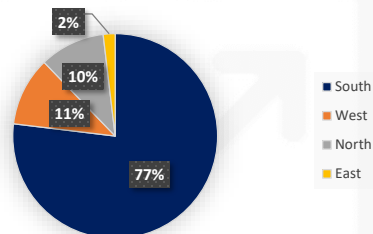
38

BEHAVIOUR OF ONE TIME BUYERS

No Of Customers By States



No Of Customers in Each Region



| Region | No Of Customers |
|--------|-----------------|
| South | 74,757 |
| West | 10,572 |
| North | 10,032 |
| East | 1,828 |



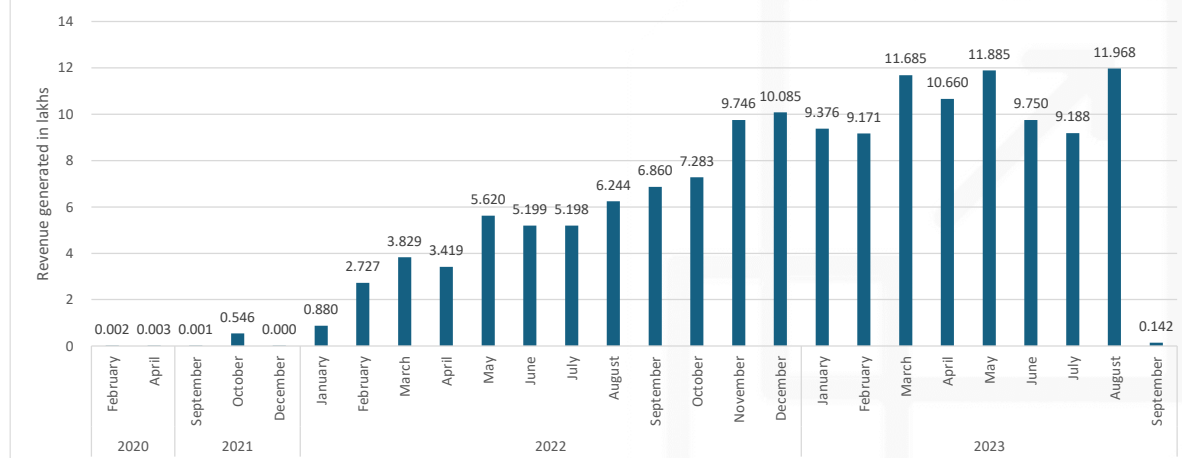
SCALETRIX.AI

Image source: SCALETRIX.AI

39

BEHAVIOUR OF ONE TIME BUYERS

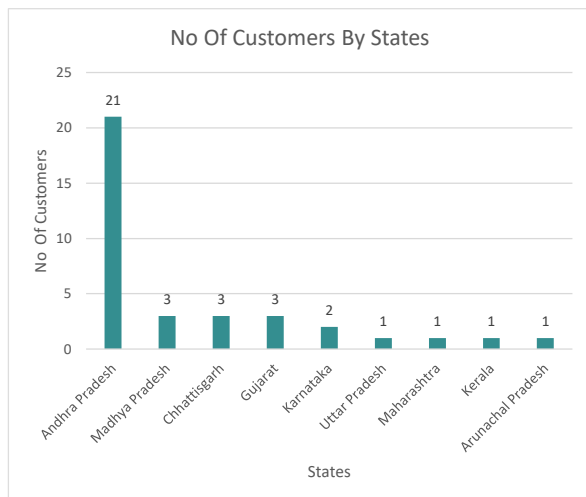
Revenue generated over time by one time buyers



SCALETRIX.AI

40

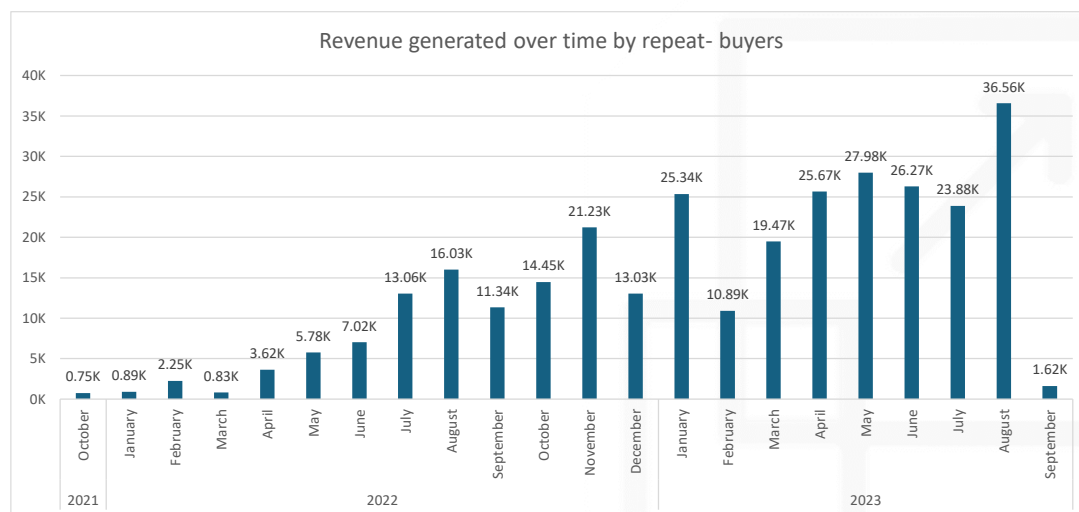
BEHAVIOUR OF REPEAT BUYERS



SCALETRIX.AI

41

BEHAVIOUR OF REPEAT BUYERS

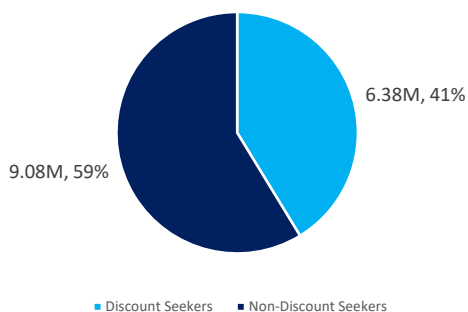


SCALETRIX.AI

42

BEHAVIOUR OF NON-DISCOUNT vs DISCOUNT SEEKER

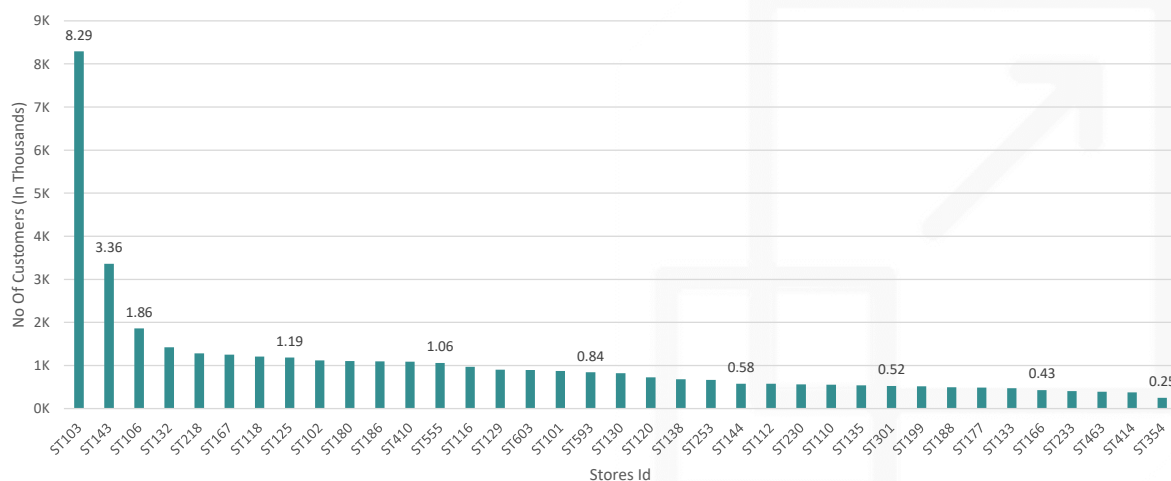
Sales Distribution By Discount Seekers And Non Discount Seekers Customers



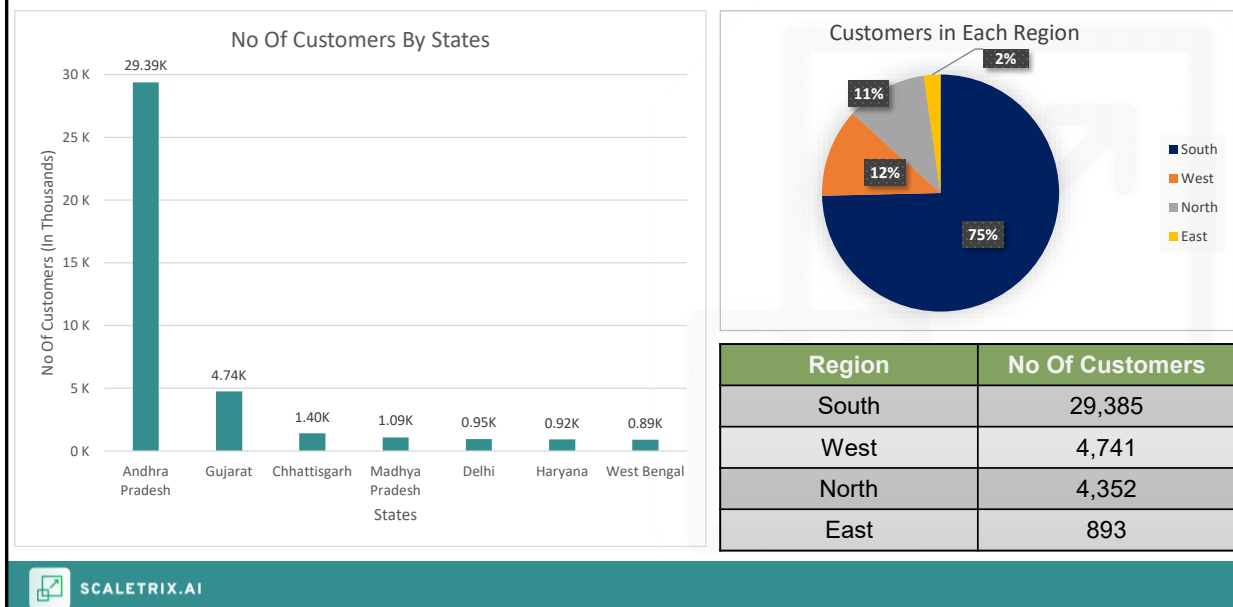
| Customer | Average order value |
|---------------------|---------------------|
| Discount seeker | ₹ 160.80 |
| Non-discount seeker | ₹ 157.28 |

BEHAVIOUR OF DISCOUNT SEEKER

No Of Customers By Stores



BEHAVIOUR OF DISCOUNT SEEKER



45

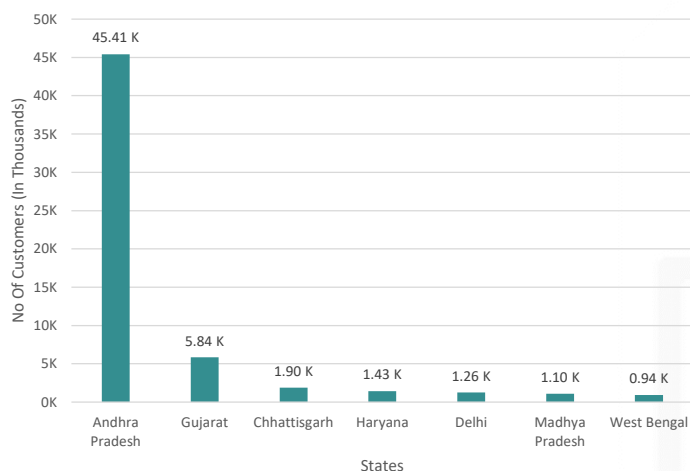
BEHAVIOUR OF NON-DISCOUNT SEEKER



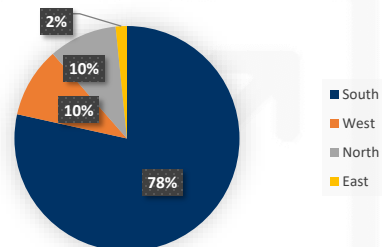
46

BEHAVIOUR OF NON-DISCOUNT SEEKER

No Of Customers By States



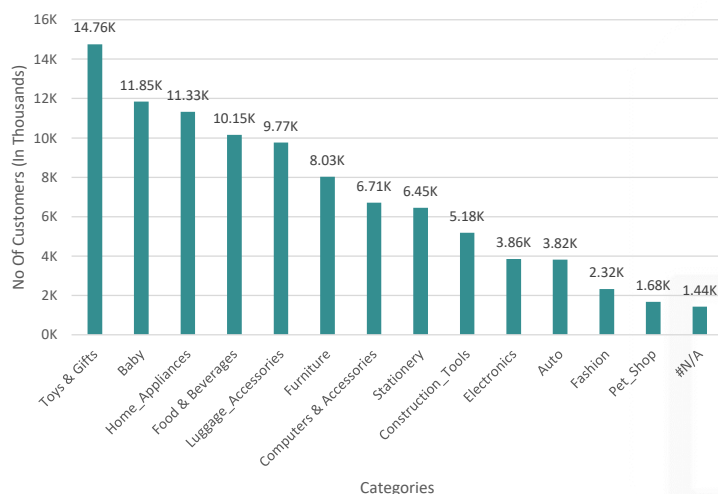
Customers in Each Region



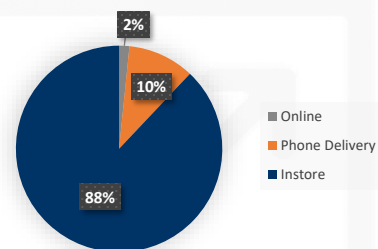
| Region | No Of Customers |
|--------|-----------------|
| South | 45,408 |
| West | 5,841 |
| North | 5,686 |
| East | 936 |

CUSTOMER PREFERENCES ANALYSIS

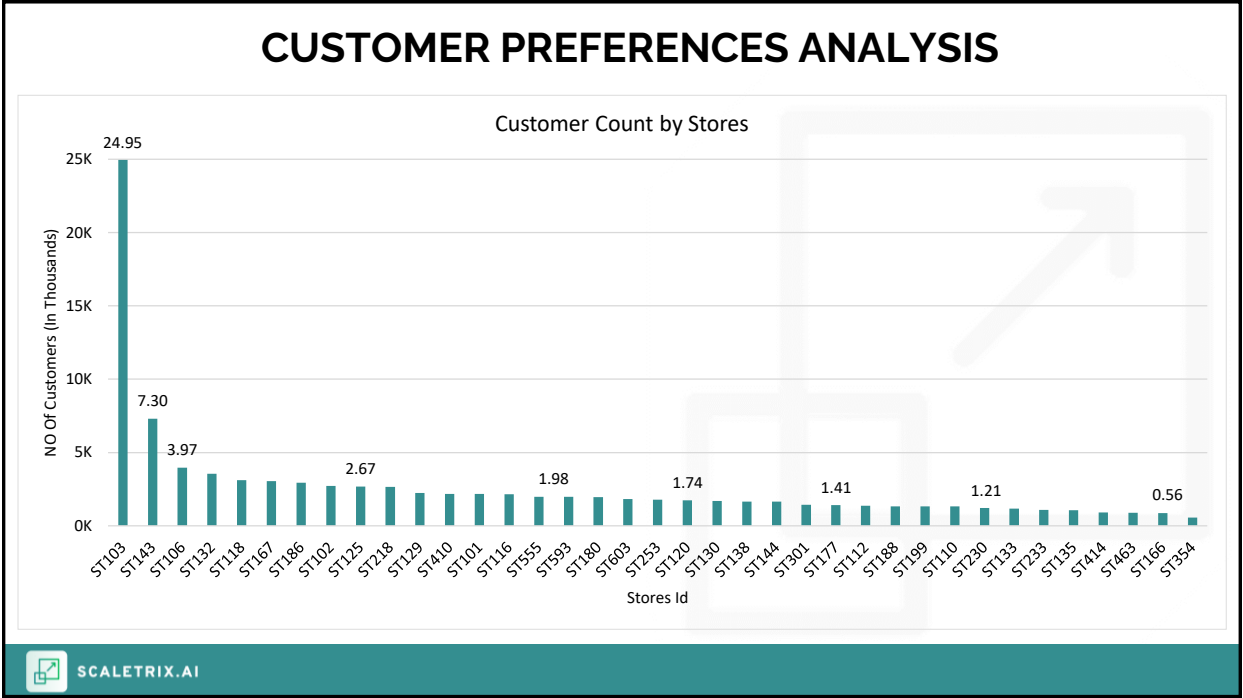
Customer Count by Categories



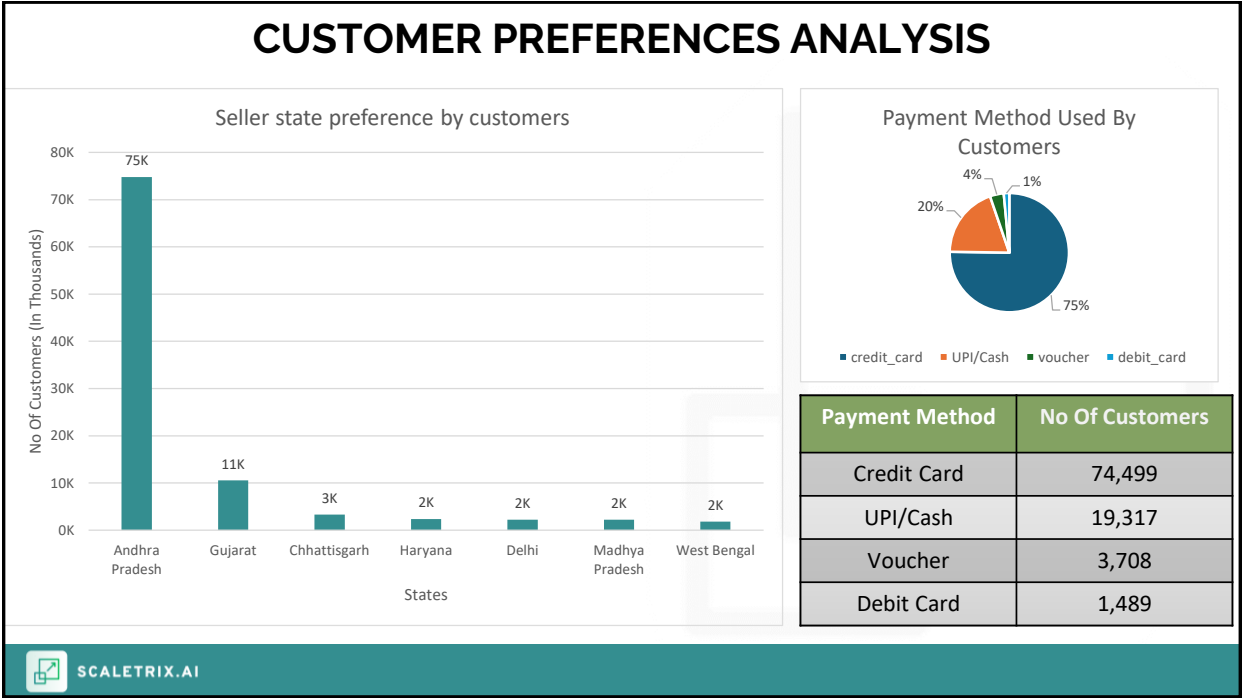
Channels Used By Customers



| Channel | No Of Customers |
|----------------|-----------------|
| Instore | 85,158 |
| Phone Delivery | 10,102 |
| Online | 1,556 |



49



50

CUSTOMER BEHAVIOUR

Gender Distribution

Observation: The customer base has more female customers (67.61 k) than male (29.19K). 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 (96.7K) compared to repeat buyers (36)
Recommendation: Introduce loyalty programs or personalized marketing to encourage repeat purchases.

Customer Segmentation by Revenue

Observation:

The majority of customers fall into the low-revenue segment, generating less than 1.5K revenue.

Medium-revenue customers (1.5K to 3K revenue) are fewer in number compared to low-revenue customers.

High-revenue customers are significantly lower than the above.

Recommendations

Targeted Marketing: Focus marketing efforts on converting low-revenue customers to medium-revenue customers by offering personalized promotions and loyalty programs.

Customer Retention: Implement strategies to retain medium-revenue customers and encourage them to move to the high-revenue segment.



SCALETRIX.AI

51

CUSTOMER BEHAVIOUR

RFM Segmentation

Observations:

- The majority of customers fall into the Standard segment 34.17K (28.47%)
- The distribution indicates that a large portion of customers are in the lower-value Standard segment, with fewer customers in the higher-value Gold and Platinum segments.

Recommendations:

- Develop targeted strategies to move customers from the Standard segment to higher-value segments like Gold and Platinum. This could include personalized offers, exclusive deals, and enhanced customer service.
- Implement retention programs for Gold and Platinum customers to maintain their loyalty and encourage higher spending. Offer exclusive benefits and rewards to these high-value customers.
- Educate customers in the Standard segment about the benefits of moving to higher-value segments. Highlight the advantages of frequent purchases and higher spending, such as access to premium products and services.



SCALETRIX.AI

52

CUSTOMER BEHAVIOUR

Discount Seekers Analysis :

Observation :

There are more non-discount seeker customers (57.72K) compared to discount seekers (39.08K). Discount seeking customer's mostly preferred stores in States like Andhra Pradesh (Store : ST103) and Gujarat are (Store: ST143) indicating these stores have a high engagement with discount seeking customer. The South region was seen to have the highest percentage of unique customers seeking discount.

Recommendations :

- Analyze the strategies used by top-performing stores (ST103 and ST104) and replicate successful tactics in other stores.
- Develop targeted campaigns for discount seekers and non-discount seekers to maximize engagement and sales.
- Tailor discount promotions to regions with higher engagement, such as the South regions.
- Focus on states with high discounted customer counts for targeted marketing campaigns to boost sales and customer loyalty.



SCALETRIX.AI

53

CUSTOMER BEHAVIOUR

Non-Discount Seeker Analysis

Observations:

- The southern region was seen to have the highest number of unique customers(78 %) with no preference on discounts indicating a strong preference for non-discounted purchases in this region. Out of the southern states Andhra Pradesh stands out with 45.41K unique customers with no preference on discounts.
- Store ID ST103 leads with 16.66K unique customers with no preference on discounts. Other top stores include ST143, ST132, ST106, and ST118.

Recommendations:

- Focus marketing efforts on the South region and Andhra Pradesh, as these areas have the highest concentration of non-discount seekers. Tailor campaigns to highlight the value and exclusivity of non-discounted products.
- Analyze the strategies and practices of top-performing stores like ST103 to identify key factors contributing to their success. Implement similar strategies in other stores to boost non-discounted sales.
- Develop loyalty programs and personalized offers for non-discount seekers to enhance their shopping experience and encourage repeat purchases. Highlight the benefits of non-discounted products, such as quality and exclusivity.



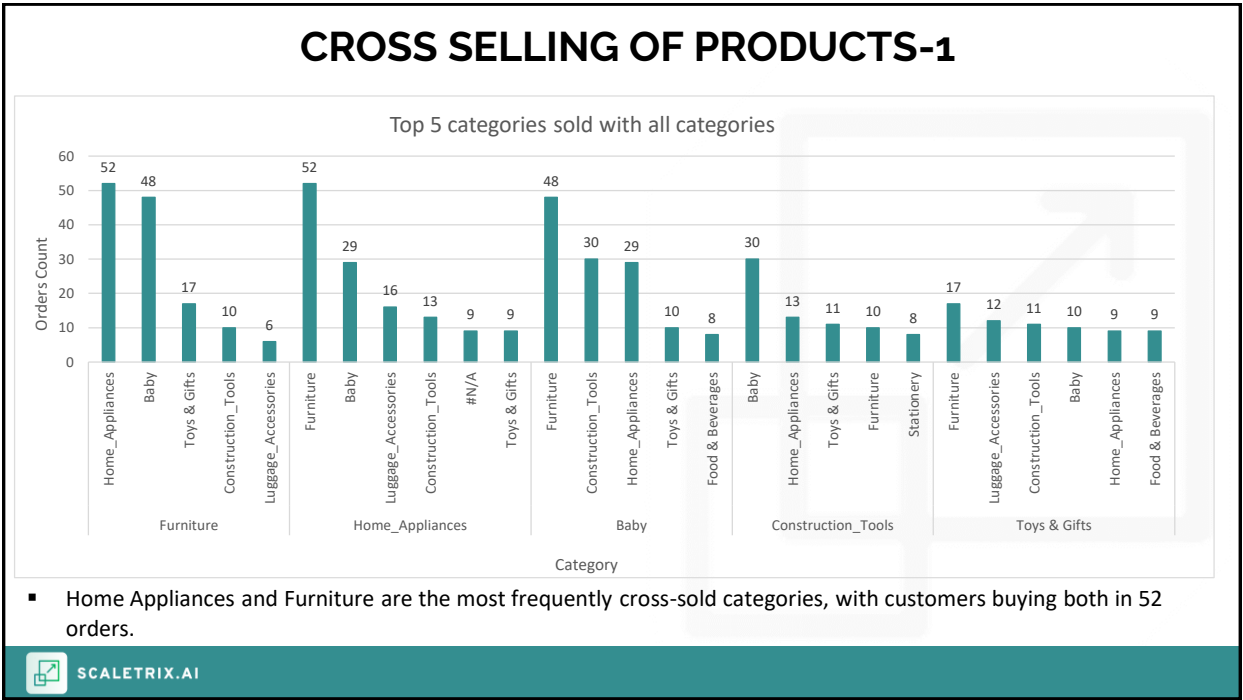
SCALETRIX.AI

54

CROSS – SELLING ANALYSIS

SCALETRIX.AI

55



56

CROSS SELLING OF PRODUCTS-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 three unique product categories occurs only once in order.



SCALETRIX.AI

57

CROSS SELLING OF PRODUCTS-3

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.



SCALETRIX.AI

58

SALES TREND ANALYSIS

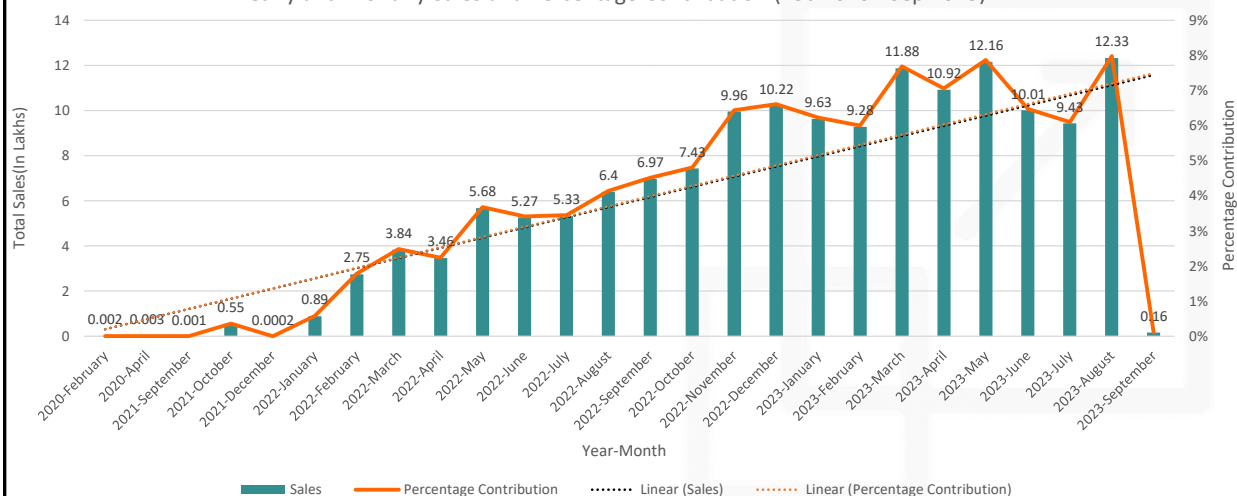


SCALETRIX.AI

59

SALES TRENDS, PATTERNS & SEASONALITY-1

Yearly and Monthly Sales and Percentage Contribution (Feb 2020 - Sep 2023)



SCALETRIX.AI

60

SALES TRENDS, PATTERNS & SEASONALITY-2

Observations

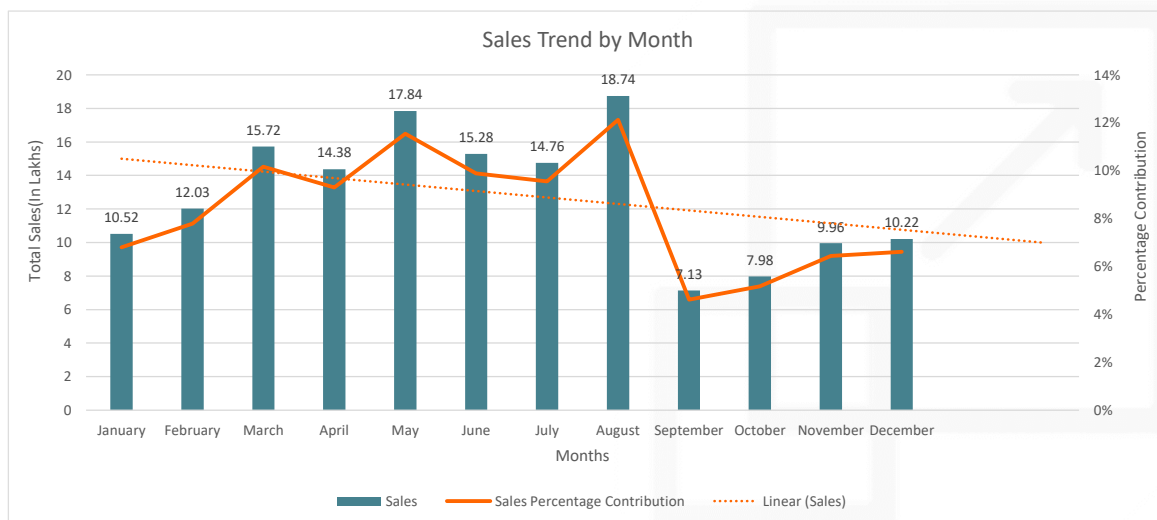
- **Trend:** The upward trend suggests that the business is growing or expanding
- **Seasonality:** The recurring peaks and troughs suggest that there are factors influencing sales that are cyclical or seasonal. These could be related to factors like holidays, weather, or economic conditions
- **Correlation:** The "Percentage contribution" line seems to follow the same seasonal pattern as the "Total Sales" line, indicating that the percentage is influenced by the overall sales volume.



SCALETRIX.AI

61

SALES TRENDS, PATTERNS & SEASONALITY-3



SCALETRIX.AI

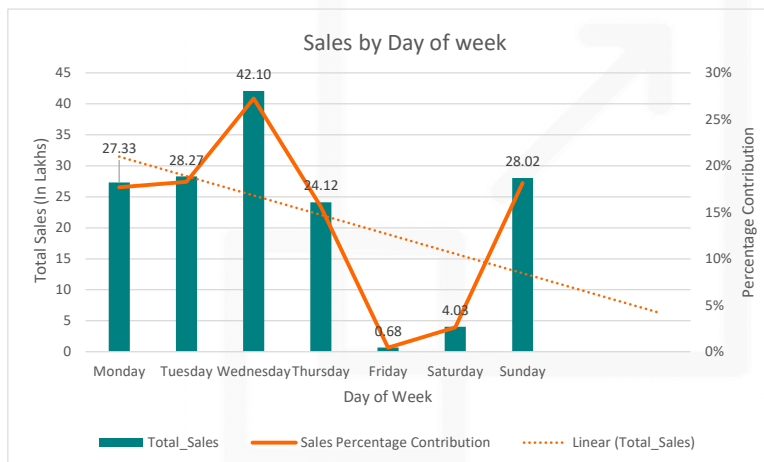
Image source: SCALETRIX.AI

62

SALES TRENDS, PATTERNS & SEASONALITY-4

Sales by Weekday vs. Weekend

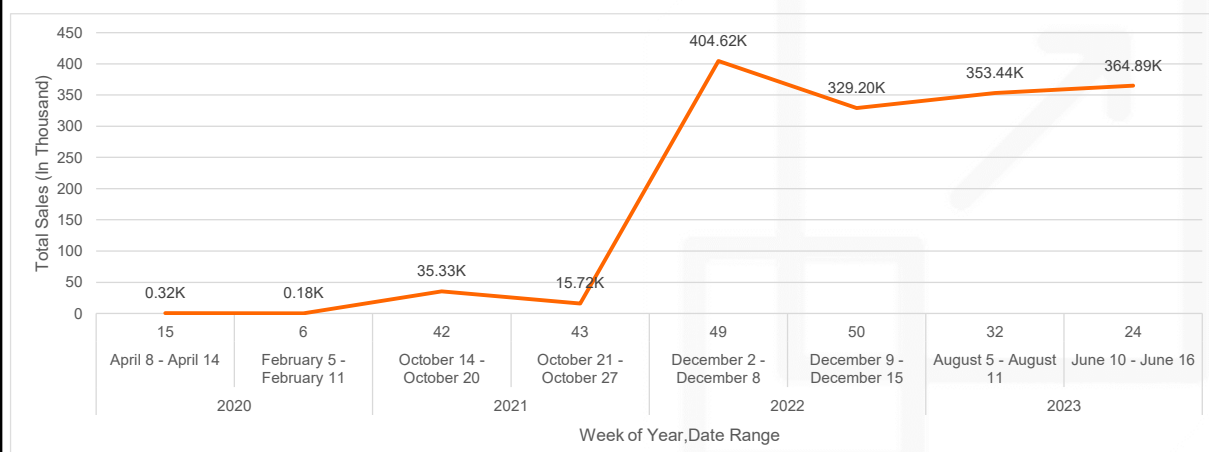
| Day Of Week | Day Type | Total Sales |
|-------------|----------|-------------|
| Monday | Weekday | 27.33 Lakh |
| Tuesday | Weekday | 28.27 Lakh |
| Wednesday | Weekday | 42.10 Lakh |
| Thursday | Weekday | 24.12 Lakh |
| Friday | Weekday | 67.71 K |
| Saturday | Weekend | 4.03 Lakh |
| Sunday | Weekend | 28.02 Lakh |



63

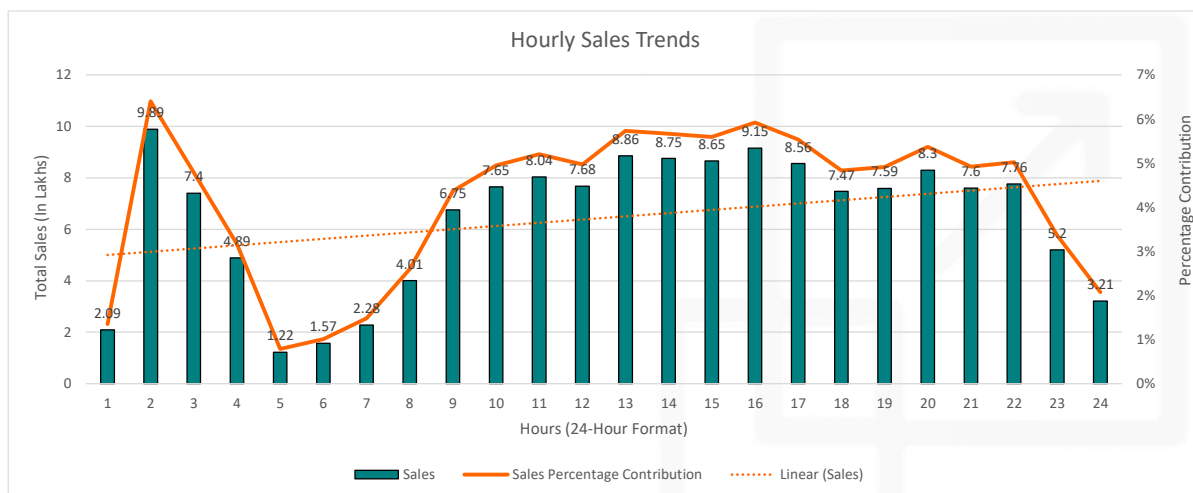
SALES TRENDS, PATTERNS & SEASONALITY-5

Top 2 Weeks with Highest Sales Each Year



64

SALES TRENDS, PATTERNS & SEASONALITY-6



Sales increase during morning (9-11 AM), afternoon (1-5 PM), and evening (8 PM)



SCALETRIX.AI

65

SALES TRENDS, PATTERNS & SEASONALITY-7

Observations :

- Sales peaks initially at 2 AM with the highest sales percentage contribution of 9.89 lakhs (6.4 %), followed by a sudden decrease in customer activity and sales.
- Around 8 AM the sales starts to increase , reaching a peak around 4 PM of 9.15 lakhs (5.92 %).
- Sales are maintained at a level but with some fluctuations from 9 AM to 10 PM , which fits with typical shopping times, when people are more likely to buy in the late morning and early evening.
- However, after 11 PM , the sales decreases rapidly.



SCALETRIX.AI

66

SALES TRENDS, PATTERNS & SEASONALITY-8

Recommendations:

- Analyze the factors contributing to the sharp decline in sales and sudden growth spurt. Consider external factors like economic conditions, industry trends, or competitive pressures. Understanding these factors can help to enhance future growth
- Use historical data and analysis to forecast future sales trends. This can aid in planning and resource allocation
- Analyze customer data to identify different segments. This can help tailor marketing strategies and product offerings
- Evaluate the competitive landscape and identify opportunities to differentiate products or services
- Review operational costs and identify areas for improvement to enhance profitability
- Consider implementing promotional strategies during off-peak seasons to stimulate demand and maintain sales levels
- Monitor economic indicators and industry trends that may impact sales. Be prepared to adjust strategies in response to changing economic conditions



SCALETRIX.AI

67

SALES TRENDS, PATTERNS & SEASONALITY-9

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



SCALETRIX.AI

68

CATEGORY BEHAVIOUR



SCALETRIX.AI

69

CATEGORYWISE – SALES (PARETO ANALYSIS)

| CATEGORY | SALES | CUMULATIVE SALES | CUM% OF SALES |
|-------------------------|-------|------------------|---------------|
| Toys & Gifts | 2.6M | 2.6M | 17% |
| Home Appliances | 1.8M | 4.4M | 28% |
| Baby | 1.6M | 6.0M | 39% |
| Food & Beverages | 1.6M | 7.6M | 49% |
| Luggage_Accessories | 1.6M | 9.2M | 60% |
| Furniture | 1.3M | 10.5M | 68% |
| Computers & Accessories | 1.3M | 11.8M | 76% |
| Construction Tools | 1.1M | 12.9M | 83% |
| Stationery | 0.7M | 13.6M | 88% |
| Auto | 0.7M | 14.2M | 92% |
| Electronics | 0.5M | 14.8M | 96% |
| Pet Shop | 0.2M | 15.0M | 97% |
| Fashion | 0.2M | 15.2M | 99% |
| # NA | 0.2M | 15.5M | 100% |

Pareto Principle:

Pareto Principle, also known as the 80/20 rule. According to the Pareto Principle, approximately 80% of the effects come from 20% of the causes.

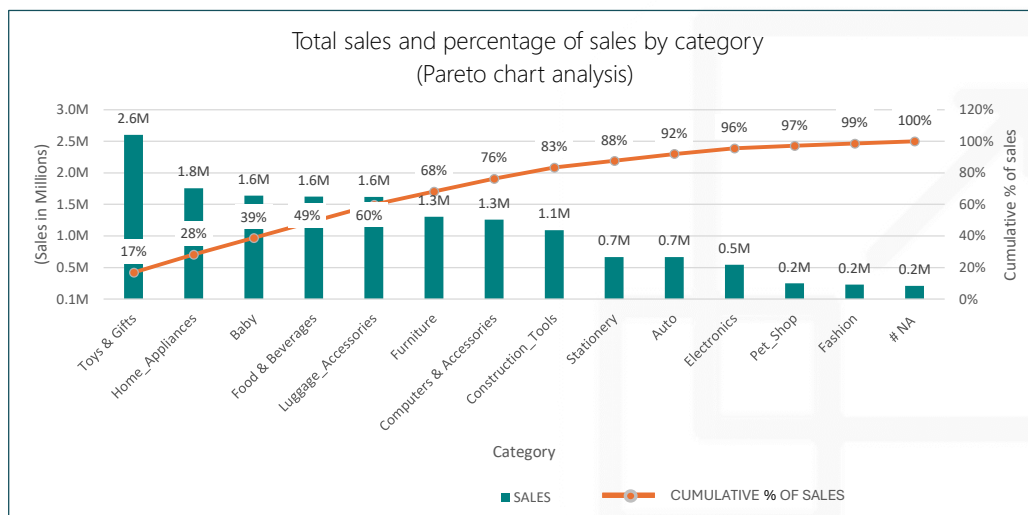
- We observe that around 50% of the categories contribute nearly 80% of total sales.



SCALETRIX.AI

70

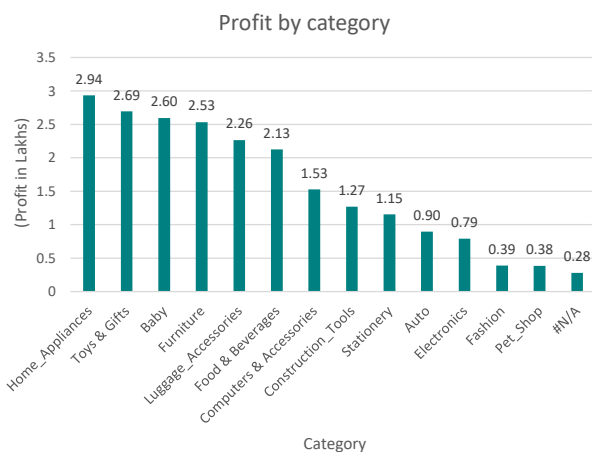
CATEGORYWISE SALES (PARETO ANALYSIS)



SCALETRIX.AI

71

PROFITABLE CATEGORY



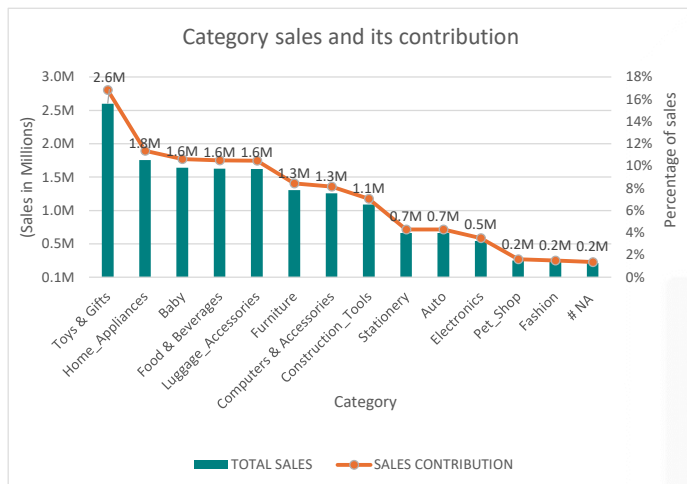
| CATEGORY | TOTAL PROFIT | PERCENTAGE OF PROFIT |
|-------------------------|--------------|----------------------|
| Home Appliances | 2.94 L | 13.44% |
| Toys & Gifts | 2.69 L | 12.33% |
| Baby | 2.60 L | 11.88% |
| Furniture | 2.53 L | 11.60% |
| Luggage_Accessories | 2.26 L | 10.36% |
| Food & Beverages | 2.13 L | 9.73% |
| Computers & Accessories | 1.53 L | 7.00% |
| Construction Tools | 1.27 L | 5.81% |
| Stationery | 1.15 L | 5.29% |
| Auto | 0.90 L | 4.11% |
| Electronics | 0.79 L | 3.61% |
| Fashion | 0.39 L | 1.79% |
| Pet Shop | 0.38 L | 1.76% |
| #NA | 0.28 L | 1.28% |



SCALETRIX.AI

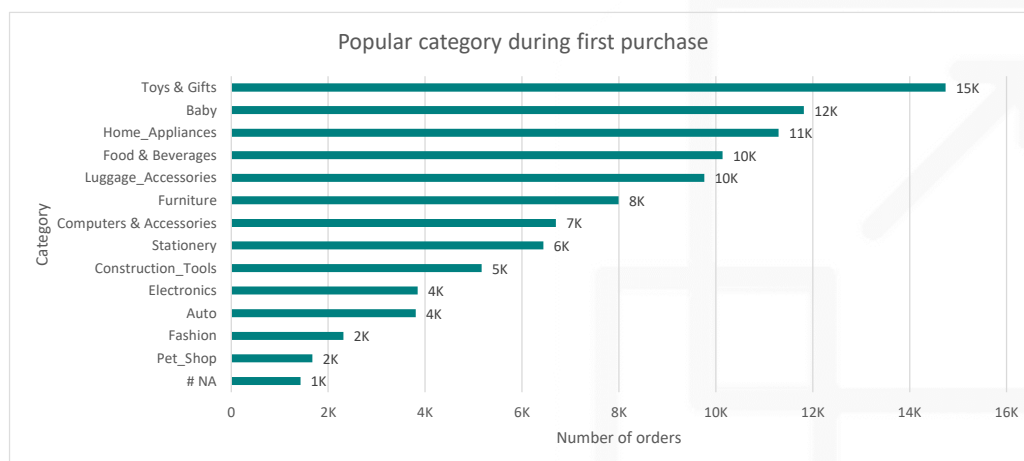
72

SALES CONTRIBUTION OF EACH CATEGORY



| CATEGORY | TOTAL SALES | SALES CONTRIBUTION |
|-------------------------|-------------|--------------------|
| Toys & Gifts | 2.60M | 17% |
| Home_Appliances | 1.76M | 11% |
| Baby | 1.64M | 11% |
| Food & Beverages | 1.62M | 11% |
| Luggage_Accessories | 1.62M | 10% |
| Furniture | 1.30M | 8% |
| Computers & Accessories | 1.26M | 8% |
| Construction_Tools | 1.09M | 7% |
| Stationery | 0.66M | 4% |
| Auto | 0.66M | 4% |
| Electronics | 0.54M | 4% |
| Pet_Shop | 0.25M | 2% |
| Fashion | 0.23M | 1% |
| # NA | 0.21M | 1% |

POPULAR CATEGORY DURING FIRST PURCHASE



CATEGORY PENETRATION ANALYSIS

$$\text{Category Penetration} = \frac{\text{Number of Orders Containing the Category}}{\text{Total Number of Orders}} \times 100$$

Category Penetration Analysis by Month-on-Month focuses on tracking how often a particular product category is included in customer orders over time. The goal is to understand the popularity or reach of a category within the overall orders for each month.



SCALETRIX.AI

75

CATEGORY PENETRATION ANALYSIS

YEAR

2022

2020

2023

MONTH

JANUARY

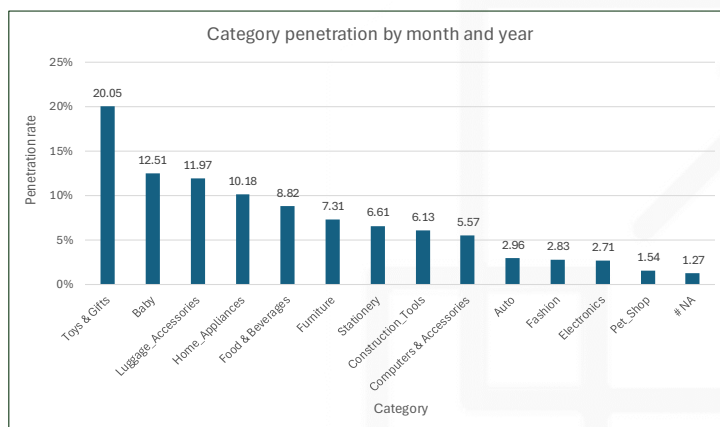
MARCH

APRIL

MAY

JUNE

JULY



SCALETRIX.AI

76

CATEGORY PENETRATION ANALYSIS

Observations

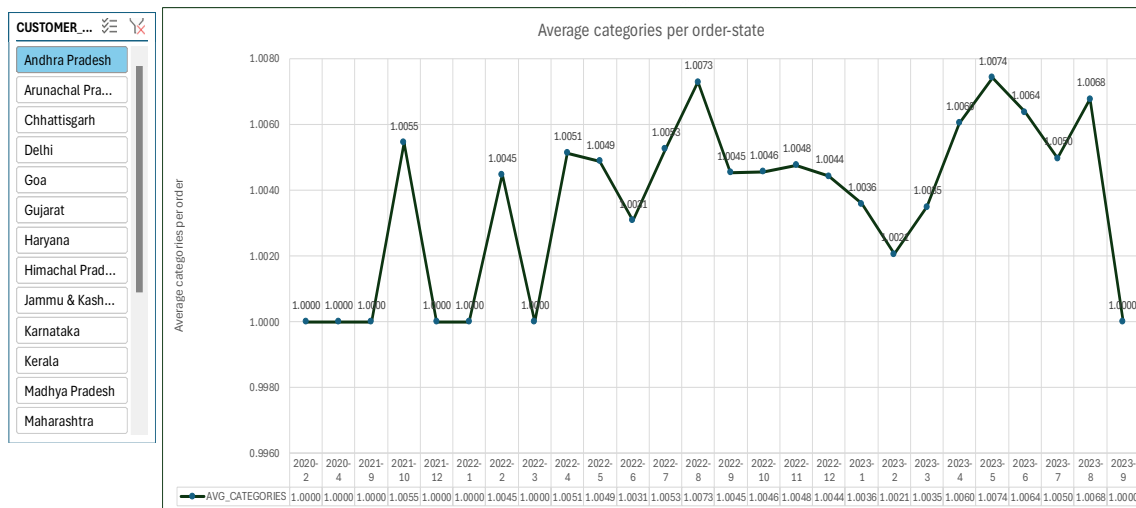
- In October 2021, **Toys and Gifts** category has the highest penetration when compared to other categories
- In the year 2022
 - From January to March, **Furniture** category has the highest penetration and **Pet Shop** category has the least penetration.
 - From April to June, **Toys and Gifts** category has highest penetration and **Pet Shop** category has the least penetration.
 - In July month, **Baby** category has highest penetration and **# NA** category has the least penetration.
 - From August to December, **Toys and Gifts** category has the highest penetration where as **# NA** and **Pet Shop** categories occupy the last position.
- In the year 2023
 - From January to August, **Toys and Gifts** category has the highest penetration and **# NA** category has the least penetration.
 - In September, **Food and Beverages** category has the highest penetration and **Pet Shop** category has the least penetration.



SCALETRIX.AI

77



AVERAGE CATEGORIES PER ORDER-CUSTOMER STATE



SCALETRIX.AI

78

AVERAGE CATEGORIES PER ORDER-REGIONWISE



YEAR  

2022

2023

2020

2021

MONTH  

JANUARY

MARCH

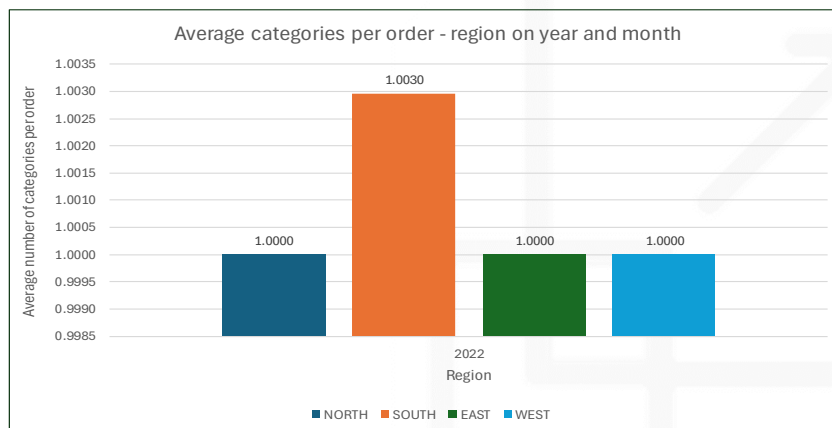
APRIL

MAY

JUNE

JULY

AUGUST



SCALETRIX.AI

79

AVERAGE CATEGORIES PER ORDER - PER REGION

Observations :

- In **East** region, average number of categories per order is exactly one for all months. This means that all of the orders placed in east region have items from just one category in every month.
- In **North** region, average number of categories per order is higher in November 2022 , January 2023 and from May 2023 to July 2023 compared to other months.
- In **South** region , average number of categories per order is at par or more than other regions across all months
- In **West** region, average number of categories per order is exactly one for all the months except from May 2023 to July 2023.



SCALETRIX.AI

80

CUSTOMER SATISFACTION

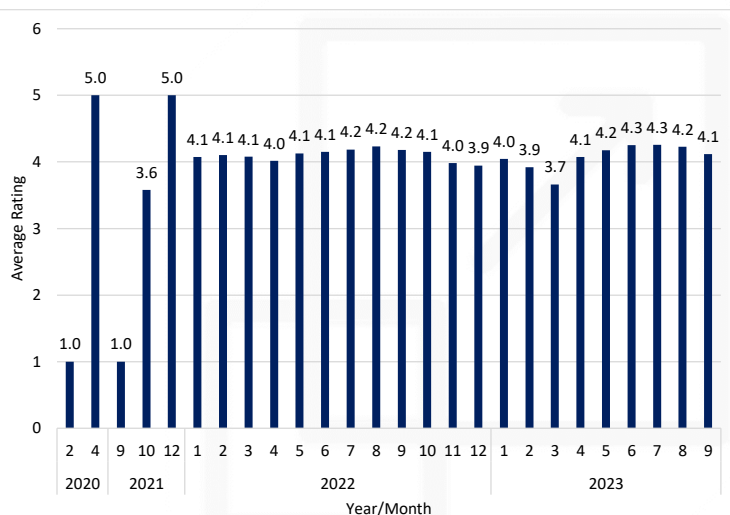


SCALETRIX.AI

81

AVERAGE RATING BY MONTH

| Year | Month | Average_Rating |
|------|-------|----------------|
| 2020 | 2 | 1.00 |
| 2020 | 4 | 5.00 |
| 2021 | 9 | 1.00 |
| 2021 | 10 | 3.58 |
| 2021 | 12 | 5.00 |
| 2022 | 1 | 4.07 |
| 2022 | 2 | 4.10 |
| 2022 | 3 | 4.08 |
| 2022 | 4 | 4.02 |
| 2022 | 5 | 4.13 |
| 2022 | 6 | 4.15 |
| 2022 | 7 | 4.19 |
| 2022 | 8 | 4.23 |
| 2022 | 9 | 4.18 |
| 2022 | 10 | 4.15 |
| 2022 | 11 | 3.98 |
| 2022 | 12 | 3.94 |
| 2023 | 1 | 4.04 |
| 2023 | 2 | 3.92 |
| 2023 | 3 | 3.66 |
| 2023 | 4 | 4.07 |
| 2023 | 5 | 4.18 |
| 2023 | 6 | 4.25 |
| 2023 | 7 | 4.25 |
| 2023 | 8 | 4.23 |
| 2023 | 9 | 4.12 |



SCALETRIX.AI

82

AVERAGE RATING BY MONTH

Observations :

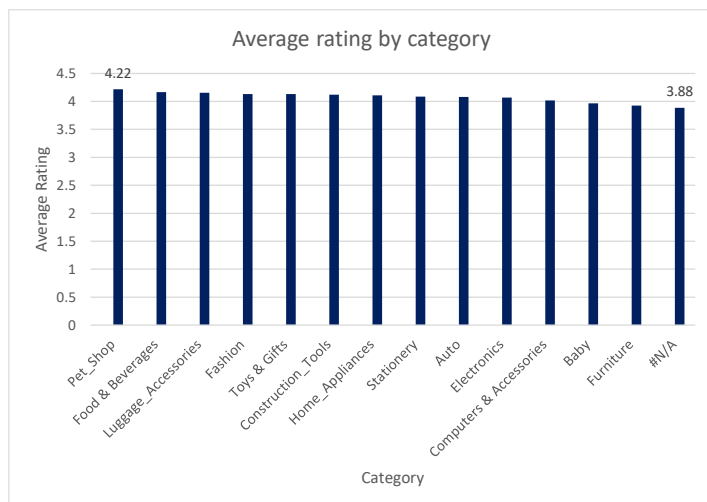
- The seasonal trends in ratings show that After the initial peaks, the values stabilize and remain relatively uniform for the rest of the intervals, with minor fluctuations.
- There is a slight dip in the middle and a small rise towards the end, suggesting some variation in the measured parameter.
- The relatively stable values across most of the remaining intervals suggest consistent performance with minor fluctuations.
- Inferences: Understand the period of stability for possible optimizations. The slight fluctuations toward the end may indicate changes in trends or customer behavior.



SCALETRIX.AI

83

AVERAGE RATING BY CATEGORY



| Category | Average Rating |
|-------------------------|----------------|
| Pet_Shop | 4.22 |
| Food & Beverages | 4.16 |
| Luggage_Accessories | 4.16 |
| Fashion | 4.13 |
| Toys & Gifts | 4.13 |
| Construction_Tools | 4.12 |
| Home_Appliances | 4.11 |
| Stationery | 4.09 |
| Auto | 4.08 |
| Electronics | 4.07 |
| Computers & Accessories | 4.01 |
| Baby | 3.97 |
| Furniture | 3.92 |
| #N/A | 3.88 |



SCALETRIX.AI

84

AVERAGE RATING BY CATEGORY

Observations :

- **Pet Shop** category has the **highest satisfaction score at 4.22**, indicating strong customer satisfaction in this category.
- **# NA** category has the **lowest satisfaction score at 3.93**.
- Most categories have scores above 4.0, indicating that overall customer satisfaction is good, though there is room for improvement in certain segments.

Recommendations :

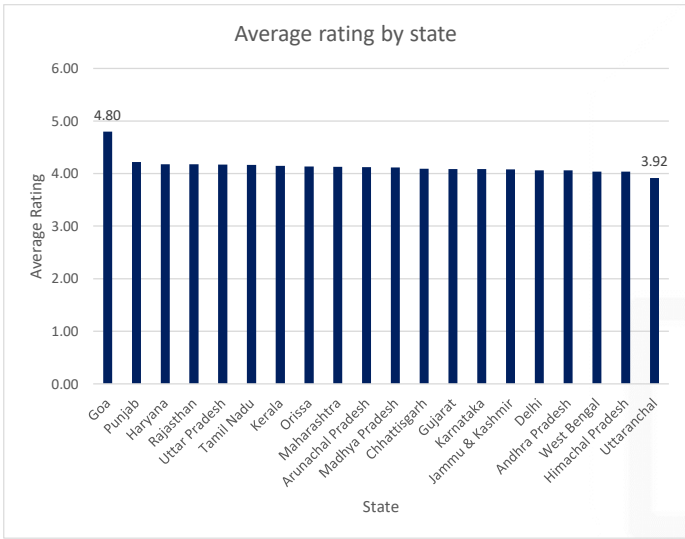
- Consider improving product quality, delivery times, or customer service in the Furniture, Baby, and # NA categories to boost satisfaction.
- Sustain high performance in Pet Shop, Food & Beverages, and Luggage Accessories by rewarding loyal customers with promotions or discounts to strengthen loyalty.
- Launch customer satisfaction improvement programs across all categories with a focus on personalization, better customer service, and optimized shopping experiences.



SCALETRIX.AI

85

AVERAGE RATING BY STATE



| State | Average_Rating |
|-------------------|----------------|
| Goa | 4.80 |
| Punjab | 4.22 |
| Haryana | 4.18 |
| Rajasthan | 4.18 |
| Uttar Pradesh | 4.17 |
| Tamil Nadu | 4.16 |
| Kerala | 4.15 |
| Orissa | 4.14 |
| Maharashtra | 4.13 |
| Arunachal Pradesh | 4.12 |
| Madhya Pradesh | 4.12 |
| Chhattisgarh | 4.10 |
| Gujarat | 4.09 |
| Karnataka | 4.09 |
| Jammu & Kashmir | 4.08 |
| Delhi | 4.06 |
| Andhra Pradesh | 4.06 |
| West Bengal | 4.04 |
| Himachal Pradesh | 4.04 |
| Uttarakhand | 3.92 |



SCALETRIX.AI

86

AVERAGE RATING BY STATE

Observations :

- **Goa** has the highest average rating of **4.8**, indicating high customer satisfaction.
- Additionally, states like Punjab, Haryana, Rajasthan, and Uttar Pradesh also demonstrate strong customer satisfaction with ratings above 4.15.
- States such as Uttaranchal, West Bengal, and Himachal Pradesh have a low average rating of around 3.92.

Recommendations :

- Prioritize improvement in low-performing states like Uttaranchal, West Bengal, and Himachal Pradesh by identifying specific customer dissatisfaction points and implementing targeted efforts to improve customer satisfaction.
- Analyze and replicate the strategies and services contributing to high customer satisfaction in Goa to maintain high satisfaction levels and consider implementing similar approaches in other states.
- Enhance customer engagement, service consistency, and introduce customer feedback loops in mid-performing states like Gujarat, Karnataka, and Chhattisgarh to prevent ratings from falling below the 4.0 threshold.



SCALETRIX.AI

87

AVERAGE RATING BY STORE



SCALETRIX.AI

88

AVERAGE RATING BY STORE

Observations :

- The top-performing stores, including **ST138**, **ST301**, and **ST106**, display **strong customer satisfaction** and service with average ratings above **4.20**.
- However, there are underperforming stores such as ST180, ST143 and ST463 which have **lower ratings** and may need operational improvements to enhance customer satisfaction.
- Additionally, the wide rating distribution ranging from 3.76 to 4.26 indicates clear opportunities for improvement across the board, especially in the lower-rated stores.

Recommendations :

- Investigate Low-Performing Stores: Conduct detailed audits to identify specific issues and initiate training programs to improve ratings.
- Gathering specific customer feedback from the lower-rated stores and using surveys or interaction data to improve specific areas is crucial for enhancing performance.



SCALETRIX.AI

89

COHORT ANALYSIS



SCALETRIX.AI

90

COHORT ANALYSIS – FIXED MONTH RETENTION

- In fixed month retention cohort analysis, we analyze how many customers from a cohort month made purchases in other months.
- We compare cohort customers and repeating customers by comparing their revenue and purchases. We also calculate retention rate.
- **Retention rate** is a critical information that is calculated to know what percentage of customers from that cohort month makes purchases in other months.
- We also calculate average months taken for subsequent purchases for each cohort month.
 - Suppose customers 1 , 2 and 3 are from Jan 2023 cohort. They make their next purchases in March, May and July 2023 respectively.
 - That means customer 1 purchases after 2 month ,customer 2 purchases after 4 months and customer 3 purchases after 6 months.
 - Hence average months to make subsequent purchase is $(2+4+6)/3$ (number of customers) . Therefore the answer is 4 in this scenario
- The analysis results for all cohort months have been presented in tabular form in upcoming slides.



SCALETRIX.AI

Image source: Scaletrix

91

TABLE COLUMN DESCRIPTIONS

| Column Name | Description |
|---------------------------------|---|
| Cohort Month | First purchase month of a customer |
| Cohort Customers | Group of customers who made their first purchase in a particular month |
| Repeat Customers | Customers who made purchases in months other than cohort month |
| Retention Rate | Number of repeat customers *100/Number of cohort customers |
| Avg. Months to Repeat | Sum of months taken by each repeat customer to make subsequent purchases / Number of cohort customers |
| Total Orders - Cohort Customers | Number of orders from cohort customers |
| Total Revenue- Cohort Customers | Sum of revenue from cohort customers |
| Total Orders- Repeat Customers | Number of orders from customers who made purchases in months other than cohort month |
| Total Revenue- Repeat Customers | Sum of revenue from customers who made purchases in months other than cohort month |



SCALETRIX.AI

Image source: Scaletrix

92

COHORT ANALYSIS – FIXED MONTH

| Cohort Month | Cohort customers | Repeat_ Customers | Retention_ Rate | Avg_Months_ to_Repeat | Total_Orders_ Cohort_Customers | Total_Revenue_ Cohort_Customers | Total_Orders_ Repeat_Customers | Total_Revenue_ Repeat_Customers |
|--------------|------------------|-------------------|-----------------|-----------------------|--------------------------------|---------------------------------|--------------------------------|---------------------------------|
| 2020-2 | 2 | 0 | 0 | 0 | 2 | 175 | 0 | 0 |
| 2020-4 | 1 | 0 | 0 | 0 | 1 | 323 | 0 | 0 |
| 2021-9 | 1 | 0 | 0 | 0 | 1 | 75 | 0 | 0 |
| 2021-10 | 299 | 0 | 0 | 0 | 310 | 55,370 | 0 | 0 |
| 2021-12 | 1 | 0 | 0 | 0 | 1 | 20 | 0 | 0 |
| 2022-1 | 560 | 0 | 0 | 0 | 564 | 88,904 | 0 | 0 |
| 2022-2 | 1607 | 1 | 0.062% | 11 | 1622 | 2,74,917 | 2 | 86 |
| 2022-3 | 2342 | 2 | 0.085% | 9 | 2343 | 3,83,391 | 6 | 851 |
| 2022-4 | 2076 | 1 | 0.048% | 6 | 2097 | 3,45,510 | 3 | 256 |
| 2022-5 | 3573 | 2 | 0.056% | 3 | 3603 | 5,67,735 | 13 | 1842 |
| 2022-6 | 3318 | 3 | 0.090% | 5 | 3348 | 5,26,880 | 19 | 2273 |
| 2022-7 | 3574 | 2 | 0.056% | 5 | 3618 | 5,32,736 | 13 | 1624 |
| 2022-8 | 4313 | 3 | 0.070% | 2 | 4383 | 6,40,346 | 16 | 1811 |



SCALETRIX.AI

Image source: Scaletrix

93

COHORT ANALYSIS – FIXED MONTH

| Cohort Month | Cohort Customers | Repeat_ Customers | Retention_ Rate | Avg_Months_ to_Repeat | Total_Orders_ Cohort_Customers | Total_Revenue_ Cohort_Customers | Total_Orders_ Repeat_Customers | Total_Revenue_ Repeat_Customers |
|--------------|------------------|-------------------|-----------------|-----------------------|--------------------------------|---------------------------------|--------------------------------|---------------------------------|
| 2022-9 | 4096 | 0 | 0 | 0 | 4142 | 6,97,013 | 0 | 0 |
| 2022-10 | 4418 | 2 | 0.045% | 1 | 4494 | 7,42,595 | 13 | 1424 |
| 2022-11 | 6173 | 5 | 0.081% | 6 | 6254 | 9,95,503 | 12 | 1670 |
| 2022-12 | 6642 | 2 | 0.030% | 5 | 6716 | 10,20,979 | 4 | 834 |
| 2023-1 | 6509 | 1 | 0.015% | 6 | 6610 | 9,62,325 | 2 | 190 |
| 2023-2 | 6316 | 2 | 0.032% | 5 | 6371 | 9,27,201 | 4 | 379 |
| 2023-3 | 7523 | 3 | 0.040% | 4 | 7604 | 11,87,254 | 7 | 1127 |
| 2023-4 | 6541 | 1 | 0.015% | 4 | 6671 | 10,90,367 | 2 | 165 |
| 2023-5 | 7433 | 0 | 0 | 0 | 7562 | 12,14,468 | 0 | 0 |
| 2023-6 | 5916 | 1 | 0.017% | 1 | 6016 | 10,00,345 | 2 | 378 |
| 2023-7 | 5822 | 1 | 0.017% | 1 | 5909 | 9,39,668 | 2 | 329 |
| 2023-8 | 7680 | 0 | 0 | 0 | 7832 | 12,30,676 | 0 | 0 |
| 2023-9 | 69 | 0 | 0 | 0 | 71 | 14,333 | 0 | 0 |



SCALETRIX.AI

Image source: Scaletrix

94

OBSERVATIONS

- There are 26 cohorts in total
- 32 customers from all cohorts have made purchase in months after cohort month
- Aug 2023 cohort has maximum number of customers (7,680)
- Apr 2020 , Sep 2021 and Dec 2021 have the lowest number of cohort customers (1)
- June 2022 cohort has the highest retention rate (0.09 %)
- Total revenue from retained customers was highest in June 2022 (INR 2,273)
- Avg. months for repeat purchase was highest for cohort Feb 2022 (11)
- There are 10 cohorts with no repeat customers
- **As a general observation, we can say retention is extremely low for all cohort months**



SCALETRIX.AI

Image source: Scaletrix

95

COHORT ANALYSIS – MONTH BY MONTH RETENTION

- **In month-by-month retention cohort analysis, we analyze how many customers from a cohort month made purchases in each of the next 12 months**
 - Suppose Jan 2023 is our cohort month. This is shown as month_0 in the table in upcoming slides.
 - Suppose 5 customers from this month made a purchase in Feb 2023- the month after cohort month. Then those customers will be included in month_1.
 - Similarly, we input number of repeat customers upto month_12 column.
- This form of analysis helps us to compare the behavior of customers from different cohorts
- The analysis result for all cohort months are shown in tabular form in upcoming slides



SCALETRIX.AI

Image source: Scaletrix

96

TABLE COLUMN DESCRIPTIONS

| Column Name | Descriptions |
|--------------|---|
| Cohort Month | First purchase month of a customer |
| Month_0 | Cohort customers |
| Month_1 | Customers in month_0 who make purchases after 1 month |
| Month_12 | Customers in month_0 who make purchases after 12 months |



SCALETRIX.AI

Image source: Scaletrix

97

COHORT ANALYSIS – RETENTION BY MONTH

| Cohort Month | Retention by Month | | | | | | | | | | | | |
|--------------|--------------------|---------|---------|---------|---------|---------|---------|---------|---------|---------|----------|----------|----------|
| | Month_0 | Month_1 | Month_2 | Month_3 | Month_4 | Month_5 | Month_6 | Month_7 | Month_8 | Month_9 | Month_10 | Month_11 | Month_12 |
| 2020-2 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2020-4 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2021-9 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2021-10 | 299 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2021-12 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2022-1 | 560 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2022-2 | 1607 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 2022-3 | 2342 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 |
| 2022-4 | 2076 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2022-5 | 3573 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| 2022-6 | 3318 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 |
| 2022-7 | 3574 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2022-8 | 4313 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| 2022-9 | 4096 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |



SCALETRIX.AI

Image source: Scaletrix

98

COHORT ANALYSIS – RETENTION BY MONTH

| Cohort Month | Retention by Month | | | | | | | | | | | | |
|--------------|--------------------|---------|---------|---------|---------|---------|---------|---------|---------|---------|----------|----------|----------|
| | Month_0 | Month_1 | Month_2 | Month_3 | Month_4 | Month_5 | Month_6 | Month_7 | Month_8 | Month_9 | Month_10 | Month_11 | Month_12 |
| 2022-10 | 4418 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 2022-11 | 6173 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 2 | 1 | 0 | | |
| 2022-12 | 6642 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | | | |
| 2023-1 | 6509 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | | | | |
| 2023-2 | 6316 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | | | | | |
| 2023-3 | 7523 | 0 | 0 | 0 | 2 | 1 | 0 | | | | | | |
| 2023-4 | 6541 | 0 | 0 | 0 | 1 | 0 | | | | | | | |
| 2023-5 | 7433 | 0 | 0 | 0 | 0 | | | | | | | | |
| 2023-6 | 5916 | 1 | 0 | 0 | | | | | | | | | |
| 2023-7 | 5822 | 1 | 0 | | | | | | | | | | |
| 2023-8 | 7680 | 0 | | | | | | | | | | | |
| 2023-9 | 69 | | | | | | | | | | | | |



SCALETRIX.AI

Image source: Scaletrix

99

OBSERVATIONS

- Cohort month Nov 2022 has the most number of customers who made repeat purchases. 5 customers from this cohort made repeat purchases in next 10 months
- Customers from 7 cohorts didn't make any purchase in next 12 months
- First 6 cohort months didn't have a single customer who made purchase in next 12 months
- Customers from 4 cohorts made purchases in the immediate month after cohort month
 - Aug 2022
 - Oct 2022
 - June 2023
 - July 2023
- Cohort Feb 2022 had the longest delay for a repeat purchase (11 months)



SCALETRIX.AI

100

FINAL OBSERVATIONS



101

Based of the entire analysis some major points have been observed :

❑ Customer Growth and Retention:

- New customer acquisition started slow but gradually increased, showing initial growth.
- Despite the increase, retention rates dropped, and many existing customers were lost.
- Revenue generated by new customers surpassed that of some existing ones, highlighting an imbalance in customer loyalty.
- **Probable Reason:** Limited customer engagement or lack of loyalty programs to retain new customers.
- **Further Investigation:** Explore customer feedback to identify reasons for attrition. Analyze loyalty strategies of top-performing stores.

❑ Store and Region Performance:

- Certain stores consistently outperformed others in terms of sales and profit.
- A similar pattern was observed across state, regions and product categories.
- **Probable Reason:** Variations in store management, location demographics, or product availability could influence these differences.
- **Further Investigation:** Compare in-store practices or regional marketing strategies for high- and low-performing areas.



102

Based of the entire analysis some major points have been observed (contd.) :

❑ Discount Seekers vs. Non-Discount Seekers:

- Non-discount seekers outnumbered discount seekers.
- Customers purchasing a single category dominated over those purchasing across multiple categories.
- **Probable Reason:** There may be a stronger product preference among non-discount shoppers.
- **Further Investigation:** Examine why single-category purchases dominate and whether cross-category promotions could drive sales.

❑ Sales Trends Across Time:

- Sales generally increased over time, but with fluctuations, especially towards the end of the time period, where a sudden drop was observed.
- Weekly sales dropped notably on Fridays and Saturdays.
- **Probable Reason:** External factors such as seasonality, market trends, or economic conditions might explain these variations.
- **Further Investigation:** Study market conditions during low-sales periods and compare weekend foot traffic across different timeframes.



SCALETRIX.AI

103

RECOMMENDATIONS



SCALETRIX.AI

104

RECOMMENDATIONS FOR WEAK POINTS

Low customer retention

- Retail chain can try personalized marketing strategies to retain customers. They could collect Email address of the customer if they are willing to provide. Thus, retail chain can remind them of any new products or upcoming offers.
- Loyalty rewards can be given for retained customers like coupons, discounts etc.
- Credit points can be given for each purchase and when customer reaches a certain threshold, they can get selected products for free

Target Less Popular Categories

- For categories with fluctuating penetration rates, run targeted promotions or bundle these with top-selling categories to increase their appeal.
- Offer discounts or perks for purchasing these bundles.
- Stock Optimization: Ensure that frequently cross-sold items are well-stocked to meet customer demand and avoid stockouts.
- Warehouse Layout: Arrange the warehouse layout to place frequently cross-sold items closer together to streamline the picking process.



SCALETRIX.AI

105

RECOMMENDATIONS FOR WEAK POINTS

Focus on area of low sales

- Enhance Weekend Sales: Since weekend sales are lower, consider implementing special weekend promotions or discounts to attract more customers.
- Explore activities or events that can drive more traffic to your business during weekends.
- Boost Friday Sales: Given the low sales on Fridays, consider introducing "Friday Specials" or limited-time offers to encourage more purchases.
- Analyze customer behavior and preferences on Fridays to tailor your offerings accordingly.

Investigate the pain points

- Analyze the factors contributing to the sharp decline in sales and sudden growth spurt. Consider external factors like economic conditions, industry trends, or competitive pressures. Understanding these factors can help to enhance future growth.
- Use historical data and analysis to forecast future sales trends. This can aid in planning and resource allocation.
- Use insights from high-performing days to replicate successful strategies on other days of the week.



SCALETRIX.AI

106

RECOMMENDATIONS FOR STRONG POINTS

Leverage High Sales Times

- Since sales are highest after midnight and pick up after 8:30 AM, consider running time-sensitive promotions during these periods to maximize revenue.
- Use targeted ads or reminders during these peak hours to drive traffic to the platform.

Promote Preferred Channels

- Encourage the use of instore and phone delivery channels, as they are the most popular among customers.
- Highlight the convenience and benefits of these channels in marketing campaigns.

Enhance Payment Methods

- Since UPI, cash, and credit card payments are the most popular, ensure that these payment methods are prominently featured and offer seamless processing to improve customer satisfaction.



SCALETRIX.AI

107

RECOMMENDATIONS FOR STRONG POINTS

Utilize RFM segmentation

- **Premium Customers:** Offer exclusive deals, early access to new products, and personalized services to maintain their loyalty and encourage frequent purchases.
- **Gold Customers:** Upsell and cross-sell to increase their average spend, potentially moving them into the Premium segment.
- **Standard Customers:** Engage targeted marketing campaigns to convert them into more frequent buyers, potentially moving them into Silver segment.
- **Silver Customers:** Provide incentives (small discounts /loyalty points) to encourage them to make additional purchases and move up to higher segments.

Enhance In-Store and Online Shopping Experience

- Design in-store displays to showcase frequently bought-together products, making it easy for customers to find complementary items.
- Ensure a seamless shopping experience across all channels (online, in-store, mobile). Offer services like click-and-collect, easy returns, and real-time inventory visibility across stores
- Regularly collect customer feedback through surveys, social media, and in-store interactions to understand pain points and areas of improvement.



SCALETRIX.AI

108

APPENDIX



SCALETRIX.AI

109

Recommendations

- Leverage High Sales Days :
 - Capitalize on the high sales on Wednesdays by introducing loyalty programs or special deals to maintain and further increase sales on this day.
 - Use insights from high-performing days to replicate successful strategies on other days of the week.
- Sustain Quarterly Growth :
 - Continue to analyze quarterly sales data to identify successful strategies and replicate them in future quarters.
 - Monitor any external factors that may influence quarterly sales and adjust business strategies accordingly.
- Analyze the factors contributing to the sharp decline in sales and sudden growth spurt. Consider external factors like economic conditions, industry trends, or competitive pressures. Understanding these factors can help to enhance future growth.
- Use historical data and analysis to forecast future sales trends. This can aid in planning and resource allocation.



SCALETRIX.AI

110

Recommendations:

- Analyze customer data to identify different segments. This can help tailor marketing strategies and product offerings.
- Evaluate the competitive landscape and identify opportunities to differentiate products or services.
- Review operational costs and identify areas for improvement to enhance profitability
- Consider implementing promotional strategies during off-peak seasons to stimulate demand and maintain sales levels.
- Monitor economic indicators and industry trends that may impact sales. Be prepared to adjust strategies in response to changing economic conditions.
- Explore the impact of factors such as industry-specific events, regulatory changes, or competitive dynamics on sales trends.



111

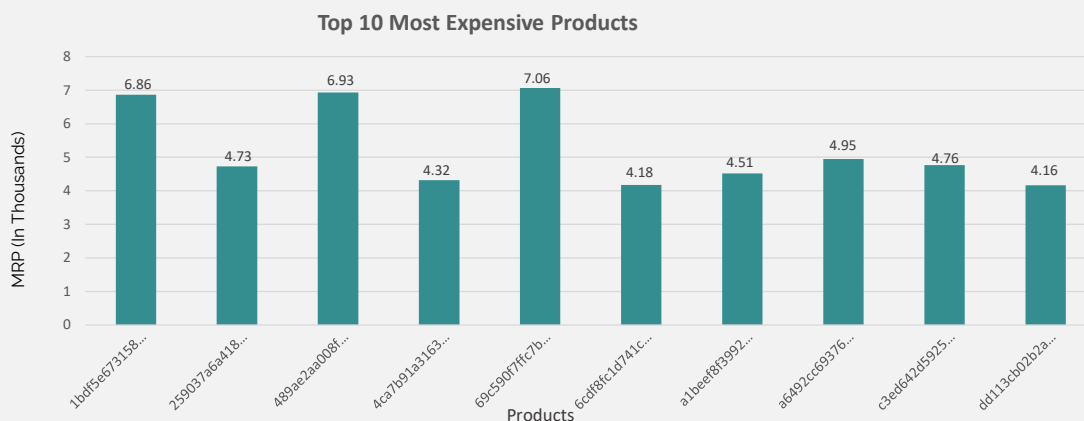
Recommendations

- 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.



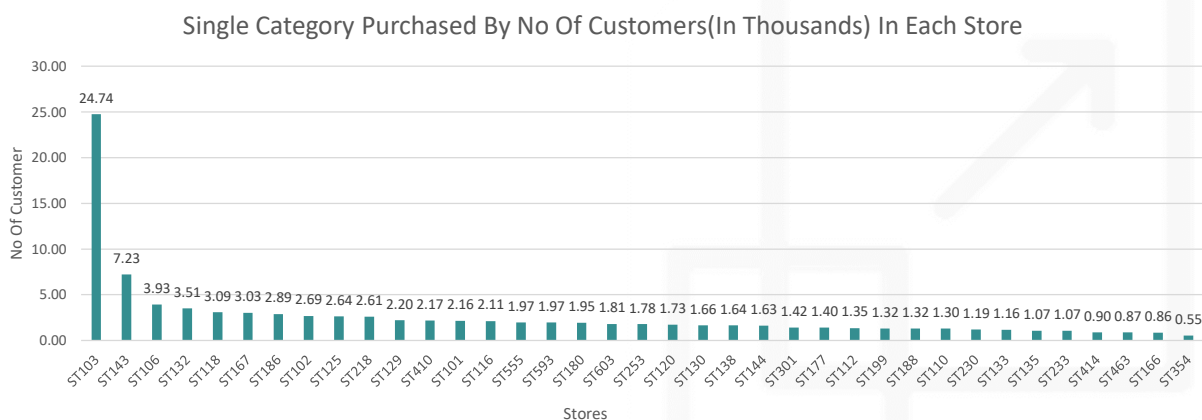
112

Exploratory Data Analysis



TOP 10 most expensive products in terms of maximum retail price(MRP).

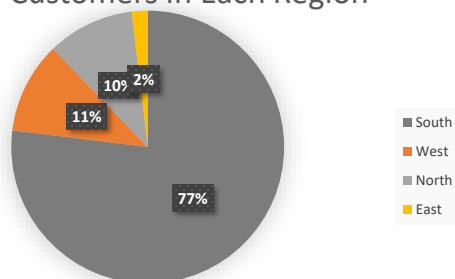
CUSTOMERS WHO PURCHASED ONLY ONE CATEGORY IN A EACH STORE



➤ **ST103** Store has most no of single category purchased by customers i.e. 24.74 K.

CUSTOMERS WHO PURCHASED ONLY ONE CATEGORY IN A EACH REGION

Single Category Purchased By No Of Customers In Each Region

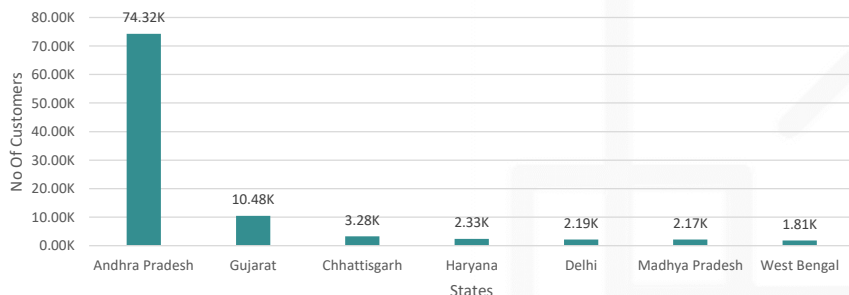


| Region | No Of Customers |
|--------|-----------------|
| South | 74,315 |
| West | 10,482 |
| North | 9960 |
| East | 1809 |

- **South** region has the most no. of single category purchased by customers i.e 74,315

CUSTOMERS WHO PURCHASED ONLY ONE CATEGORY IN STATES

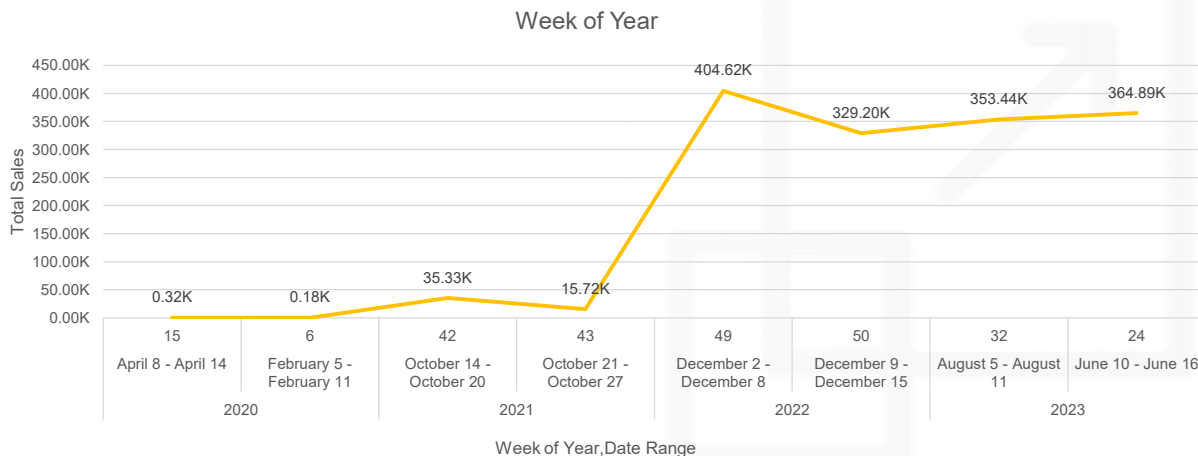
Single Category Purchased By No Of Customers In States



- **Andhra Pradesh** has the most no of single category purchased by 74.32 K customers and **West Bengal** has the lowest with 1.81 K

Sales Trends, Patterns & Seasonality-9

Sales by Top 2 Week of Year



SCALETRIX.AI

117

Sales Trends, Patterns & Seasonality-10

Observations

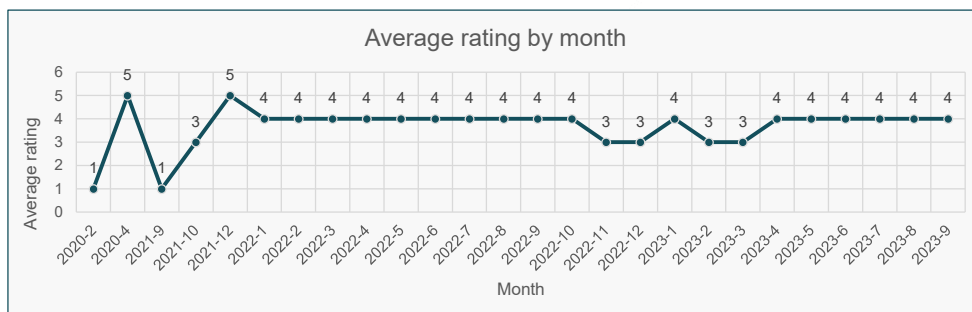
- Sales by Day of the Week:
 - Wednesday has the highest sales at 4.21MM, making up 27.24% of the weekly sales.
 - Friday has the lowest sales at 0.7MM, contributing only 0.43% of the weekly sales.
 - There is a significant drop in sales on Fridays compared to other days of the week.
- Quarterly Sales Overview:
 - The highest quarterly sales were recorded in Q2 2023 at 3.31M.
 - There is a general upward trend in quarterly sales from Q4 2020 to Q2 2023, indicating consistent growth over the years.



SCALETRIX.AI

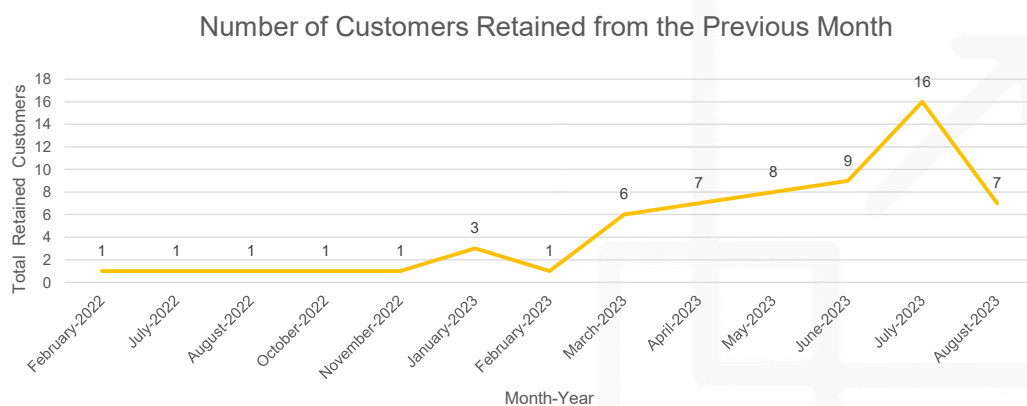
118

AVERAGE RATING BY MONTH



- **Least average ratings** were given on **2020 February and 2021 September** when compared to other months.
- Highest rating of 5 was given on 2020 April and 2021 December
- Most months had average rating of 4

Exploratory Data Analysis



Exploratory Data Analysis

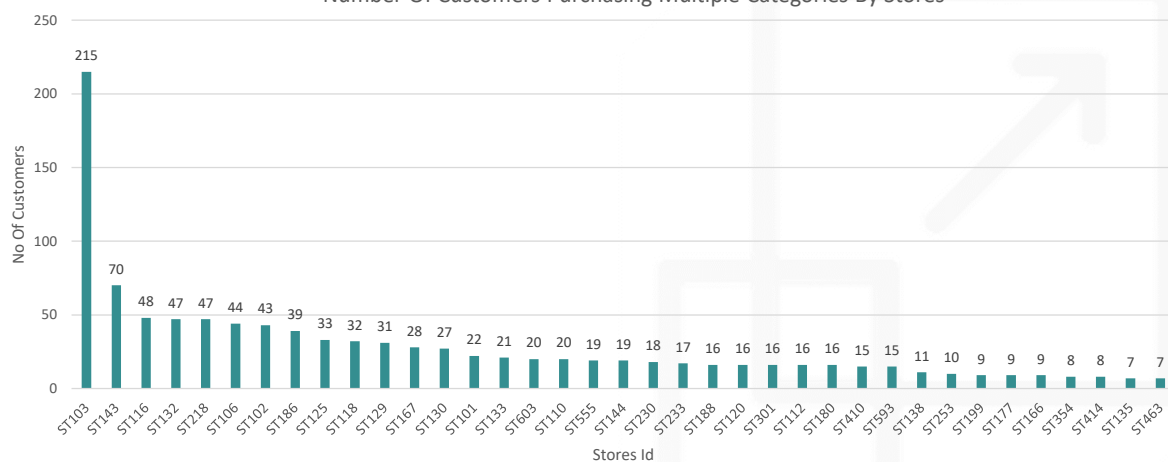
Monthly Revenue and Customer Count from Existing Customers



121

CUSTOMERS WHO PURCHASED IN MULTIPLE CATEGORIES

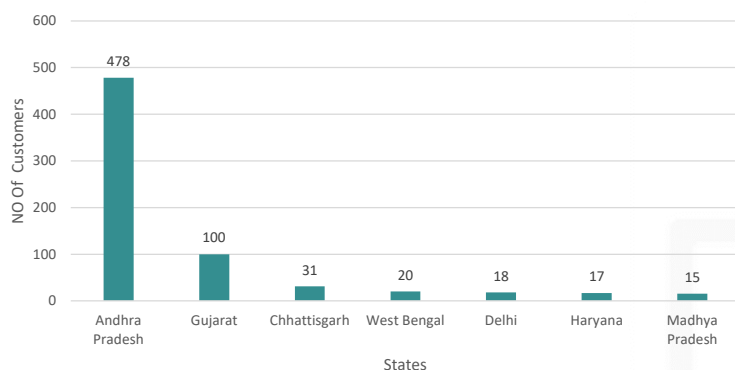
Number Of Customers Purchasing Multiple Categories By Stores



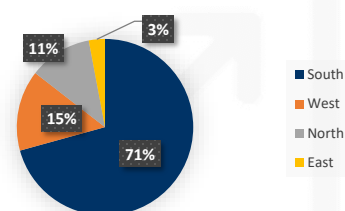
122

CUSTOMERS WHO PURCHASED IN MULTIPLE CATEGORIES

Number of Customers Purchasing Multiple Categories by States



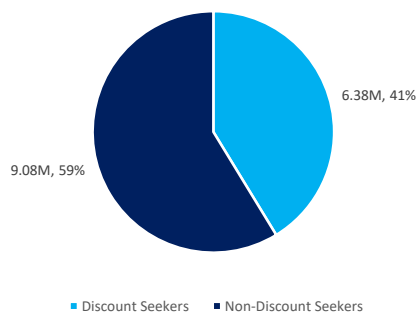
Number of Customers Purchasing Multiple Categories by Region



| Region | No Of Customers |
|--------|-----------------|
| South | 478 |
| West | 100 |
| North | 78 |
| East | 20 |

BEHAVIOUR OF NON-DISCOUNT vs DISCOUNT SEEKER

Sales Distribution By Discount Seekers And Non Discount Seekers Customers



| Customer | Order Count | Percentage |
|---------------------|-------------|------------|
| Discount seeker | 39,080 | 40.10 % |
| Non-discount seeker | 58,304 | 59.87 % |

| Customer | Average order value |
|---------------------|---------------------|
| Discount seeker | ₹ 160.80 |
| Non-discount seeker | ₹ 157.28 |

