Customer Shopping Behaviour Analysis

Agenda

- Problem Statement

- Approach

- Data Overview

- Project Workflow

- Exploratory Data Analysis

- Key Insights

- Dashboard Preview

- Recommendation

Problem Statement

Problem Statement

Despite having a high volume of daily sales, the business struggles to understand which products drive profit, how customer demographics affect sales, and what factors lead to sales fluctuations. This project aims to analyze shopping sales data to uncover actionable insights that can improve marketing strategies, inventory planning, and overall business performance.

Objectives

- To analyze overall sales performance.

- To identify top-performing and low-performing products.

- To study customer purchasing behavior.

- To understand revenue drivers.

Approach

Approach – Two phase solution

Exploratory Data Analysis

Interactive Dashboard in Power BI

1. Deep dive into sales data.

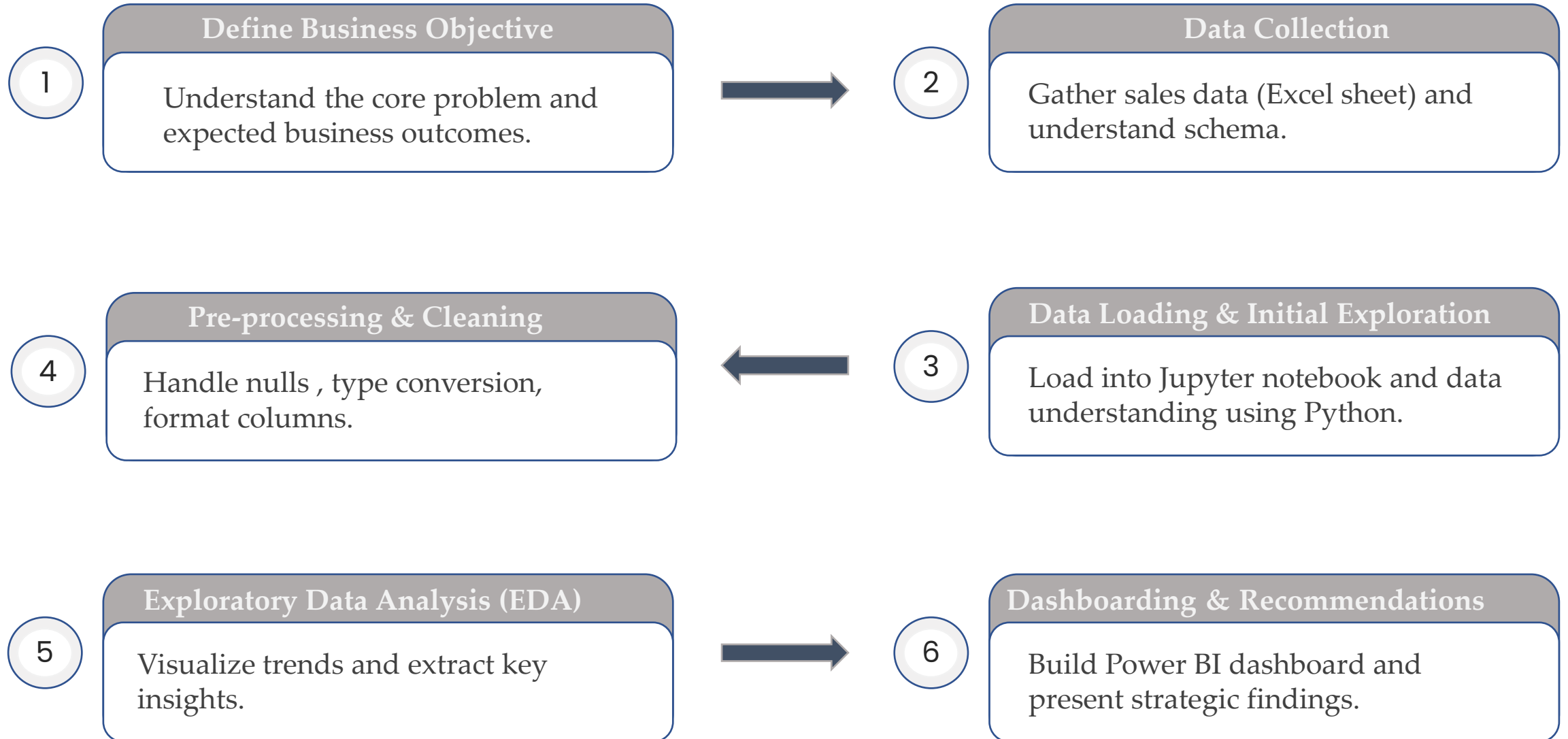2. Understanding of data, trends and patterns.

1. Build a live view for business Users to self serve insights.

2. Enable slicing by session for dynamic view of the dashboard.

Project Workflow

Project Workflow

1 — **Define Business Objective**
Understand the core problem and expected business outcomes.

2 — **Data Collection**
Gather sales data (Excel sheet) and understand schema.

4 — **Pre-processing & Cleaning**
Handle nulls , type conversion, format columns.

3 — **Data Loading & Initial Exploration**
Load into Jupyter notebook and data understanding using Python.

5 — **Exploratory Data Analysis (EDA)**
Visualize trends and extract key insights.

6 — **Dashboarding & Recommendations**
Build Power BI dashboard and present strategic findings.

Exploratory Data Analysis

Raw Dataset Structure – Before Processing

	Customer ID	Age	Gender	Item Purchased	Category	Purchase Amount (USD)	Location	Size	Color	Season	Review Rating	Subscription Status	Shipping Type	Discount Applied	Promo Code Used	Payment Method
0	1	55	Male	Blouse	Clothing	53	Kentucky	L	Gray	Winter	3.1	Yes	Express	Yes	Yes	Venmo
1	2	19	Male	Sweater	Clothing	64	Maine	L	Maroon	Winter	3.1	Yes	Express	Yes	Yes	Cash
2	3	50	Male	Jeans	Clothing	73	Massachusetts	S	Maroon	Spring	3.1	Yes	Free Shipping	Yes	Yes	Credit Card
3	4	21	Male	Sandals	Footwear	90	Rhode Island	M	Maroon	Spring	3.5	Yes	Next Day Air	Yes	Yes	PayPal
4	5	45	Male	Blouse	Clothing	49	Oregon	M	Turquoise	Spring	2.7	Yes	Free Shipping	Yes	Yes	PayPal

Age	Gender	Item Purchased	Category	Purchase Amount (USD)	Location	Size	Color	Season	Review Rating	Subscription Status	Shipping Type	Discount Applied	Promo Code Used	Payment Method	Frequency of Purchases	Custom
40	Female	Hoodie	Clothing	28	Virginia	L	Turquoise	Summer	4.2	No	2-Day Shipping	No	No	Venmo	Weekly	Adult
52	Female	Backpack	Accessories	49	Iowa	L	White	Spring	4.5	No	Store Pickup	No	No	Bank Transfer	Bi-Weekly	Senior
46	Female	Belt	Accessories	33	New Jersey	L	Green	Spring	2.9	No	Standard	No	No	Venmo	Quarterly	Senior
44	Female	Shoes	Footwear	77	Minnesota	S	Brown	Summer	3.8	No	Express	No	No	Venmo	Weekly	Senior
52	Female	Handbag	Accessories	81	California	M	Beige	Spring	3.1	No	Store Pickup	No	No	Venmo	Quarterly	Senior

Exploratory Data Analysis (EDA)

Uncovering patterns, trends, and business insights from historical data !

Understanding the "What, Where & Why" behind the sales numbers

Exploring data through visuals, aggregations, and comparisons

Laying the groundwork for informed recommendations

Let's Dive In →

EDA: Step-by-Step Process

- Import Libraries

- Load Data

- Initial Exploration

- Pre-processing & Cleaning

- Feature Engineering

- EDA & Visualization

- Key Insights

- Recommendations

Pre-processing & Feature Engineering

The necessary steps applied to prepare this dataset for analysis.

- Load dataset and perform Data profiling.

- Replace missing value of review rating with help of backward filling technique.

- Performing data type conversion on review rating.

- Create Custom column with help of Age column of dataset.

- Drop unwanted columns to make dataset understandable.

- Checking and Removing of duplicated data from dataset.

Final Dataset Structure – Ready for Analysis

Age	Gender	Item Purchased	Category	Location	Size	Color	Season	Review Rating	Subscription Status	Shipping Type	Discount Applied	Promo Code Used	Payment Method	Frequency of Purchases	Custom	Amount
55	Male	Blouse	Clothing	Kentucky	L	Gray	Winter	3	Yes	Express	Yes	Yes	Venmo	Fortnightly	Senior	53
19	Male	Sweater	Clothing	Maine	L	Maroon	Winter	3	Yes	Express	Yes	Yes	Cash	Fortnightly	Young Adult	64
50	Male	Jeans	Clothing	Massachusetts	S	Maroon	Spring	3	Yes	Free Shipping	Yes	Yes	Credit Card	Weekly	Senior	73
21	Male	Sandals	Footwear	Rhode Island	M	Maroon	Spring	3	Yes	Next Day Air	Yes	Yes	PayPal	Weekly	Young Adult	90
45	Male	Blouse	Clothing	Oregon	M	Turquoise	Spring	2	Yes	Free Shipping	Yes	Yes	PayPal	Annually	Senior	49

```
df.shape

(3900, 18)

df.columns

Index(['Customer ID', 'Age', 'Gender', 'Item Purchased', 'Category',
       'Purchase Amount (USD)', 'Location', 'Size', 'Color', 'Season',
       'Review Rating', 'Subscription Status', 'Shipping Type',
       'Discount Applied', 'Promo Code Used', 'Payment Method',
       'Frequency of Purchases', 'Custom'],
      dtype='object')
```
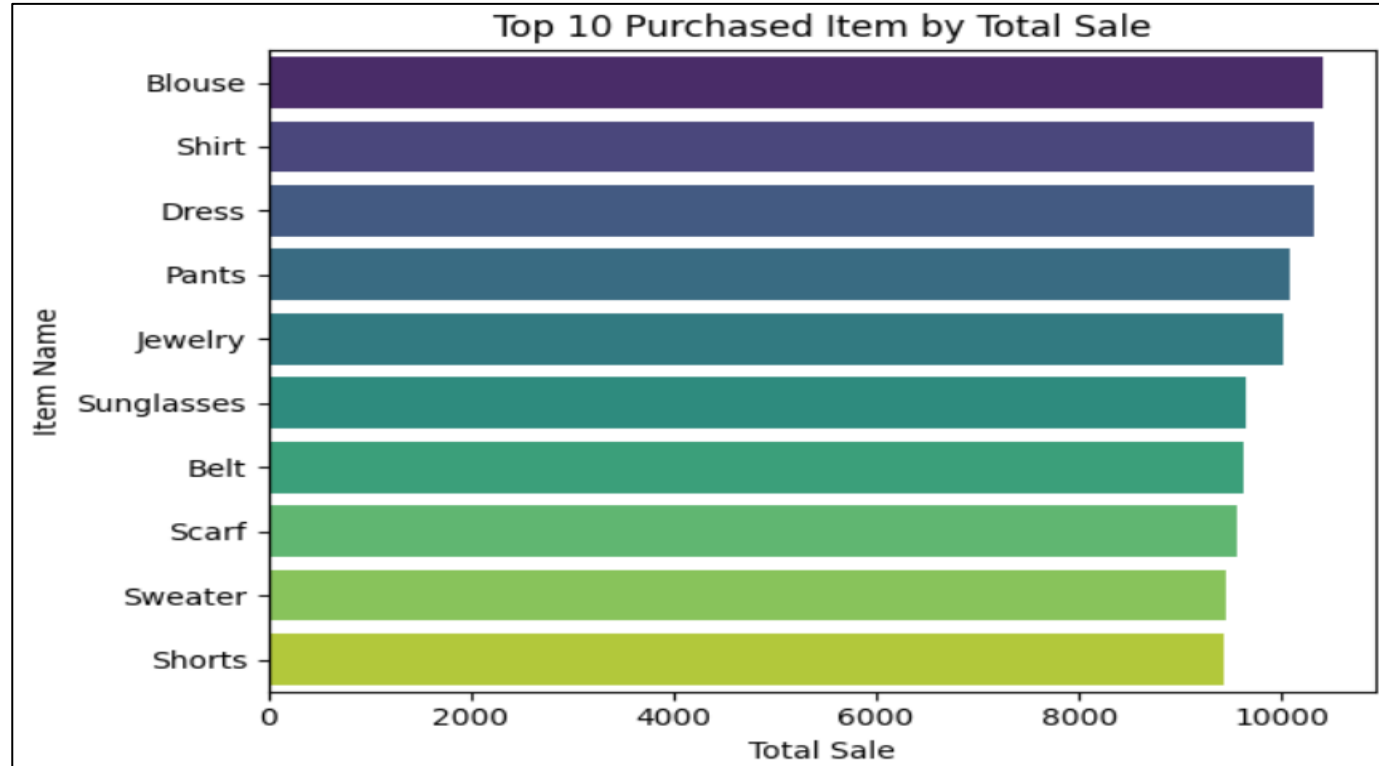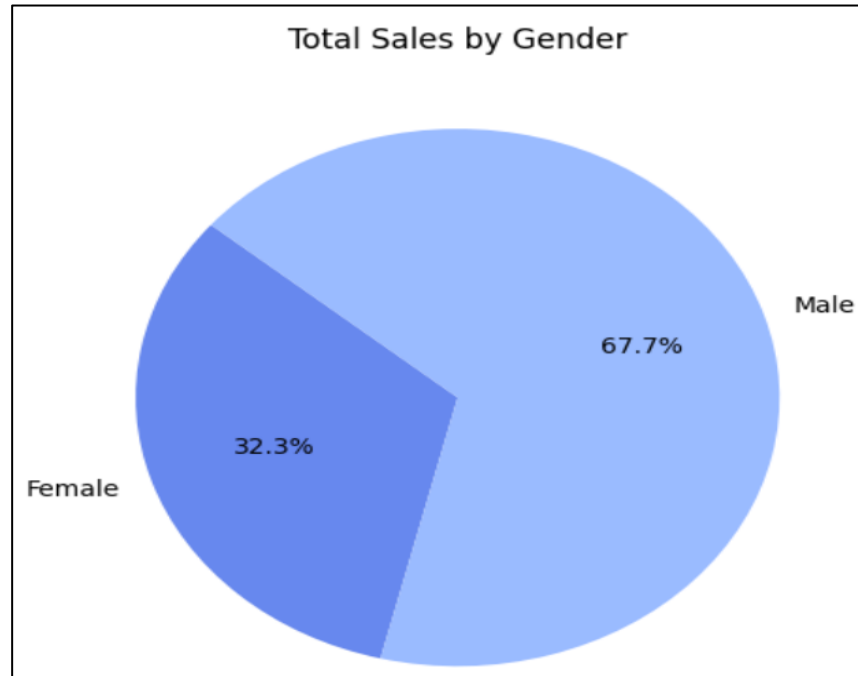
Charts & Insights

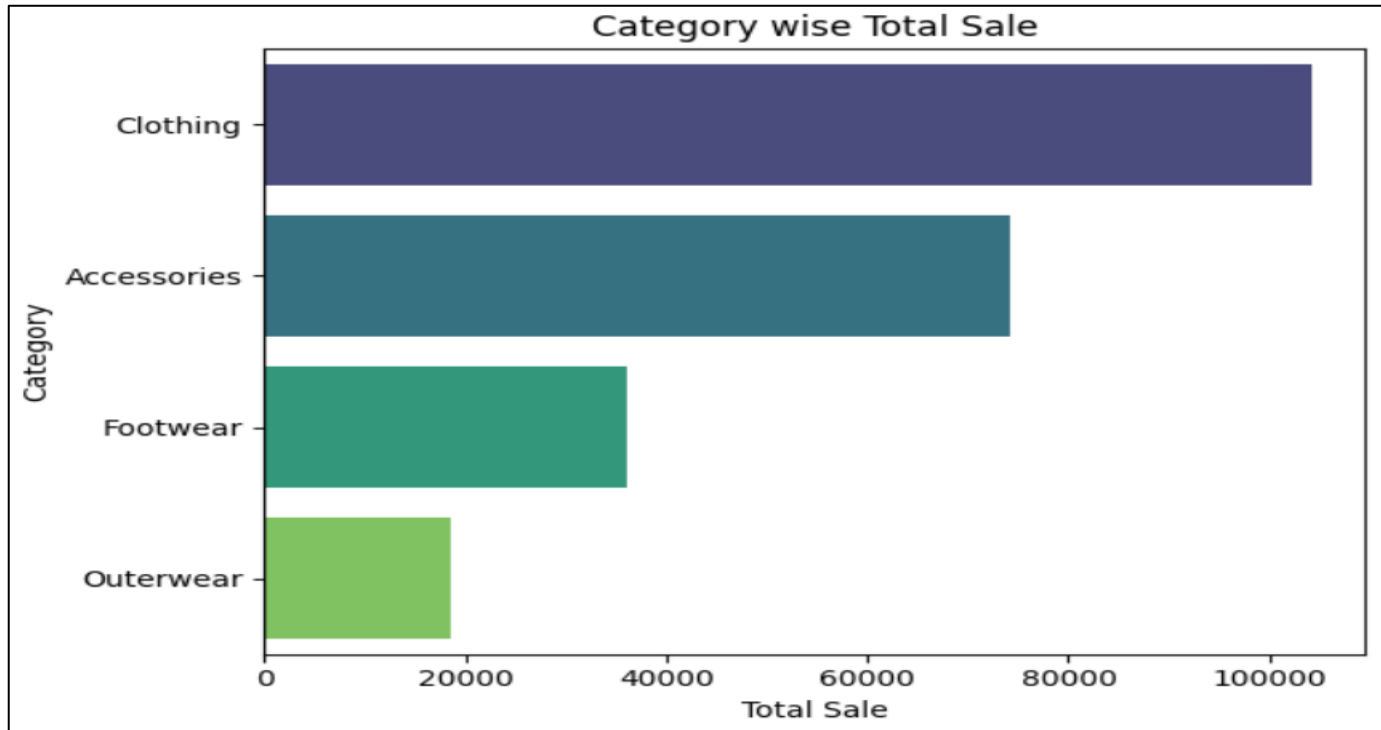# Top 10 Purchased Items.

💡 **Insights**

- Most sale Items : **Blouse, Shirt & Dress**.
- From Item 6 to 10 nearly give same sale.
- **Pants & Jewellery** are Mid-Range sale Items.
- **Most** of the above Items are belongs to **clothing** Category.

Gender wise Total Sale.

💡 **Insights**

- In total sale Male & Female contribute **67.7%** and **32.3%** respectively.
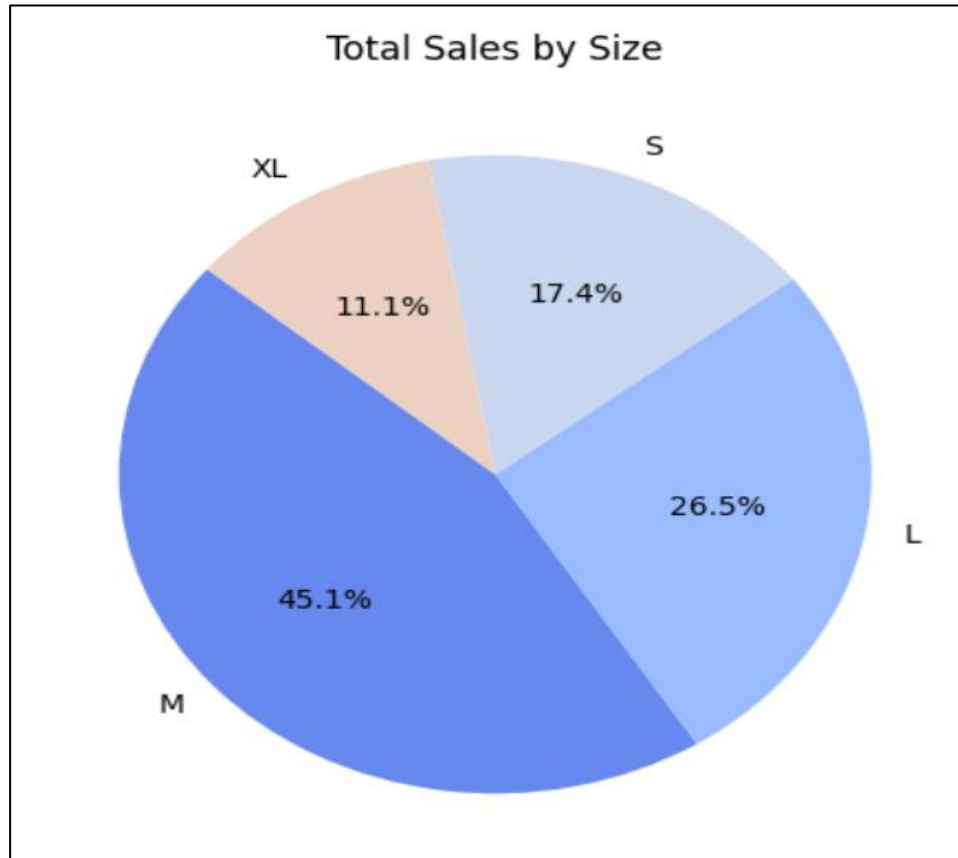- Total sale by Male is **Twice** to Female.

Category wise Total Sale

💡 **Insights**

- **Top Sale Category**: Clothing with **~$104K**
- Accessories & Footwear are Mid – range Category.
- **Bottom Sale Category**: Outerwear with **~19K**.
- **~85K** difference between Top and Bottom Sale Category.
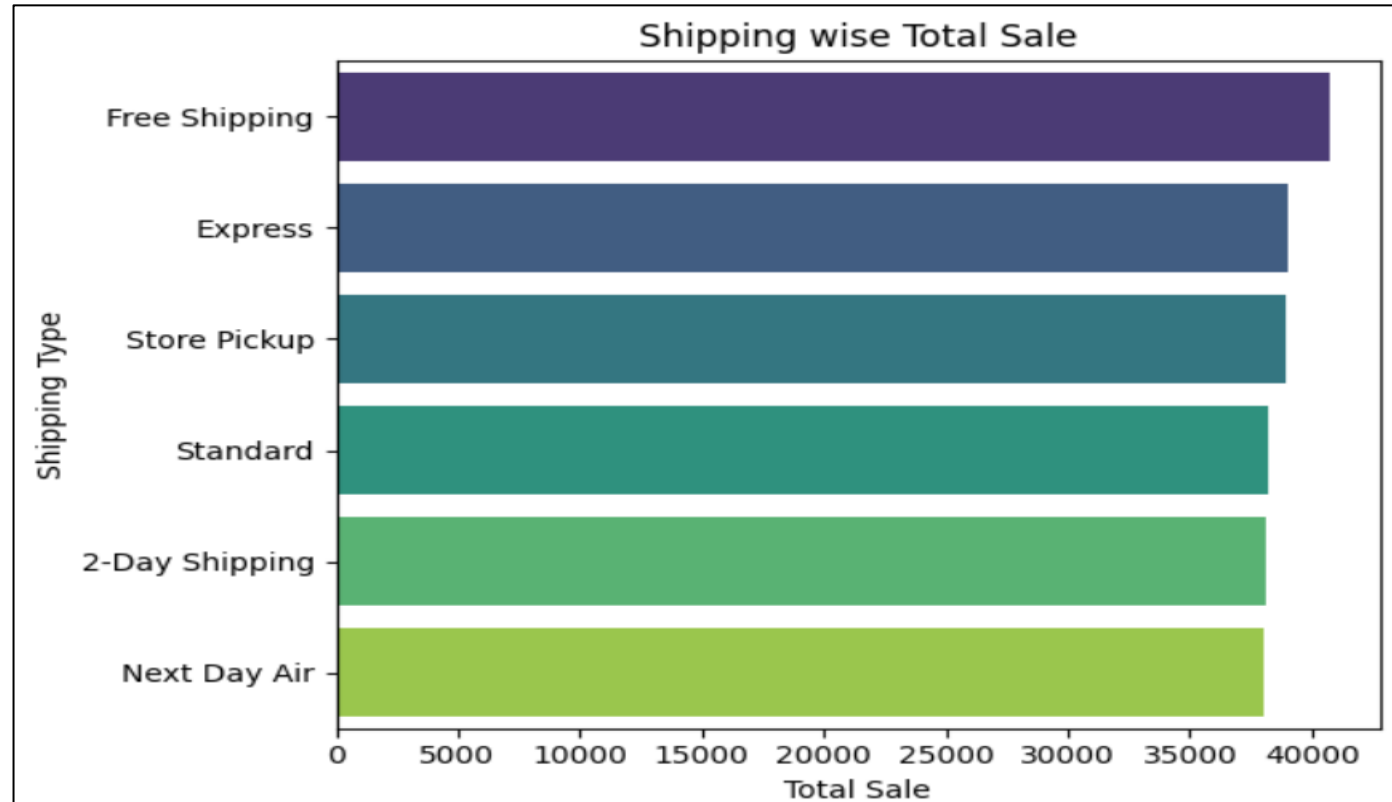
Size Wise Sale

Total Sales by Size

💡 **Insights**

- **M size** dominates: Generates the majority of total sales at **45.1%**.

- **L size** is significant: Contributes a substantial **26.5%** to total sales.

- Size S & XL Contribute **17.4%** and **11.1%** Sale respectively.
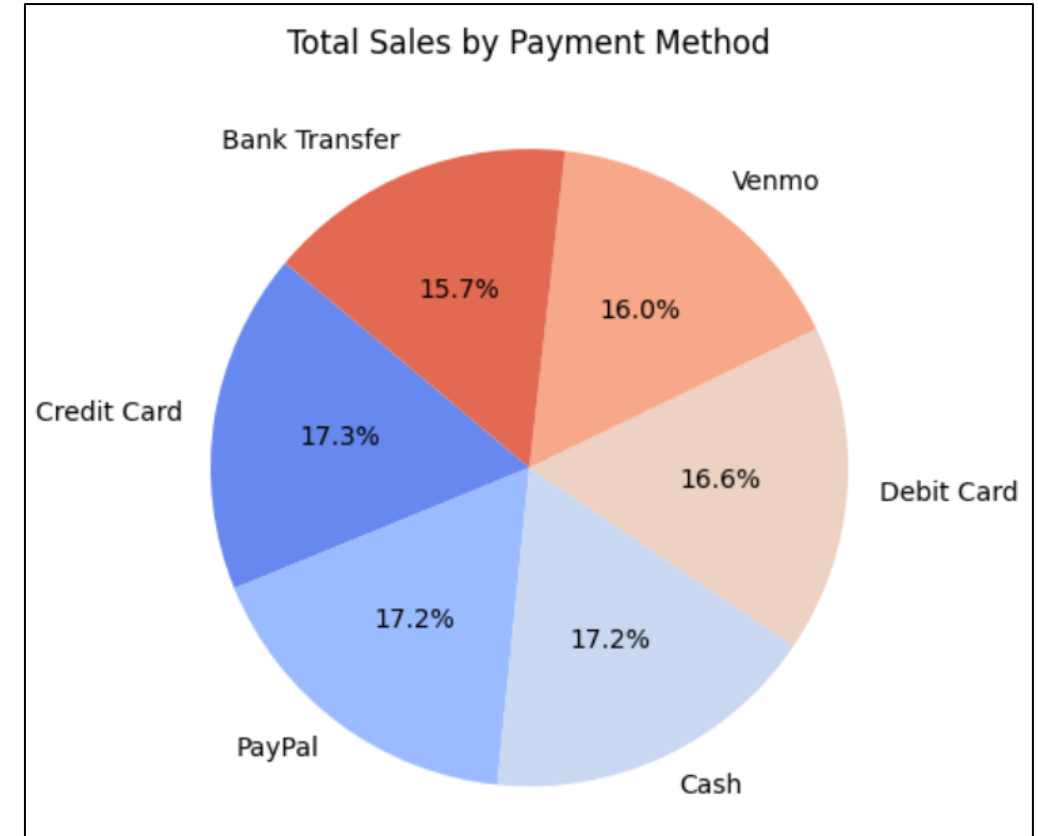
Shipping wise Sale

💡 **Insights**

- **Free Shipping** is most Preferred Shipping type give ~41K Sale.
- Standard, 2-Day Shipping, Next Day Air shipping type contributed equally with ~38K.

Payment Mode wise Sale

💡 **Insights**

- **PayPal, Credit Card & Cash** equally use for Payment.
- Venmo & Debit Card contributed **16% & 16.6%** in total sale.

Total Sales by Payment Method

Key Insights

Key Insights 💡📊

- Total Sale - **~233K**, Avg Rating – **3.7**.

- **Clothing** - Most Sale Category.

- Male are more Active in shopping as compare to Female.

- **Outwear** is less Sale Category as compare to others.

- **M** is most Sale and **S & XL** are Less sale Size respectively.

- **Free Shipping** Mostly Preferred by Customers.

- All Payment mode on an avg contribute **16.9%** in Total Sale.

Dashboard Preview
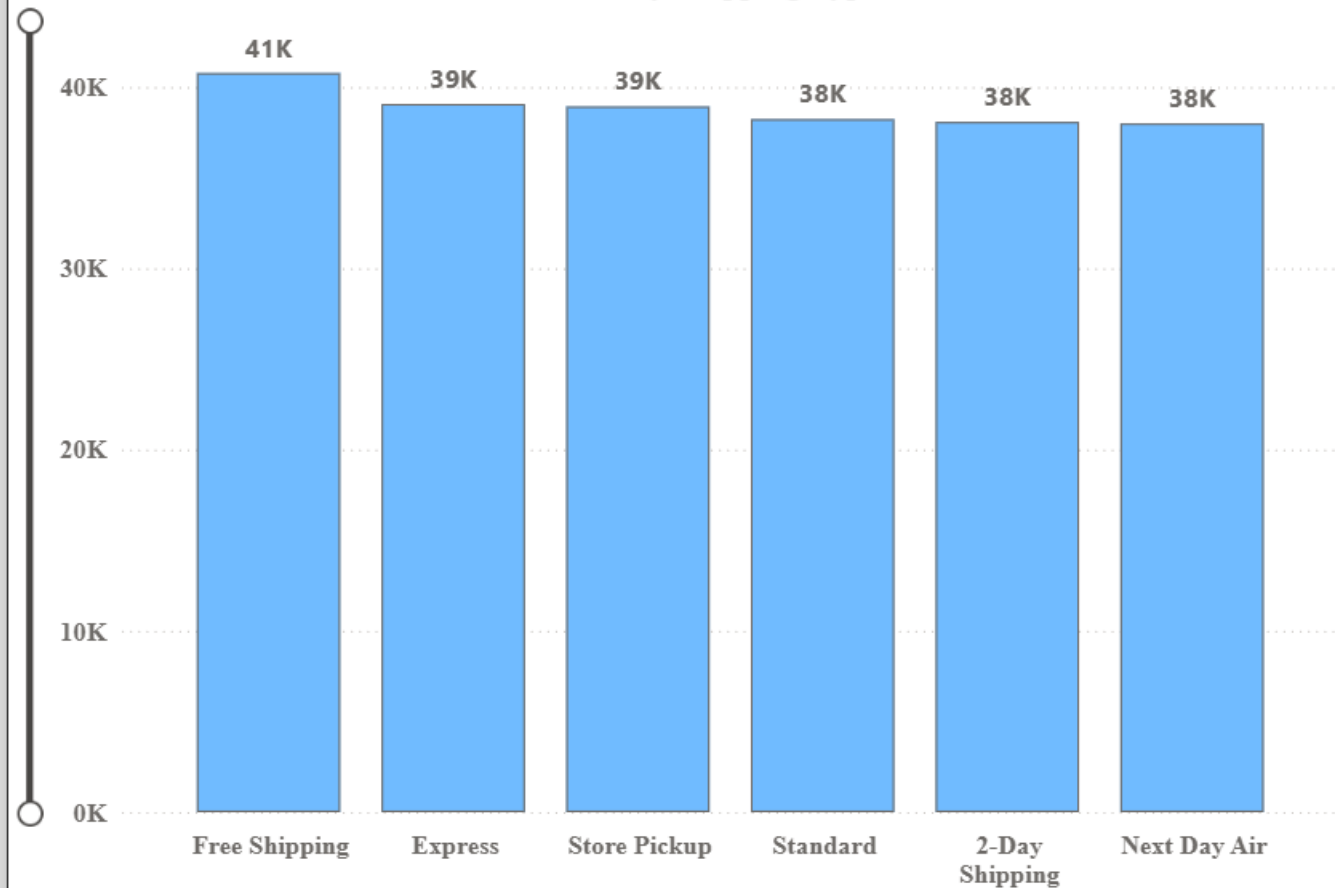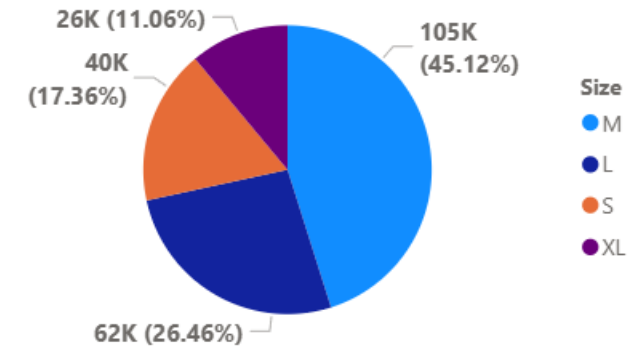
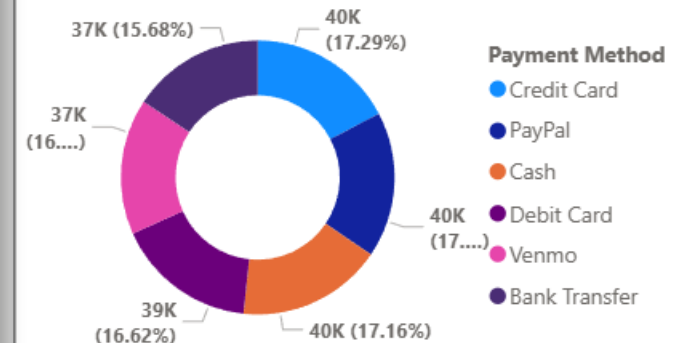# Page 1

| Customer Shopping Behavior Analysis

| Select all | Fall | Spring | Summer | › |

Total Sale
$233.1K

Avg Rating
3.7

Avg Sale
$59.8

Most Sale Category
Clothing

Item Purchased	Sum of Purchase Amount (USD)
Jeans	$7,548
Gloves	$8,477
Sneakers	$8,635
Backpack	$8,636
Hoodie	$8,767
Handbag	$8,857
Boots	$9,018
Sandals	$9,200
Shoes	$9,240

Alabama
$5,261   3.8
Sum of Purchase Am...   Avg_Rating

Alaska
$4,867   3.8
Sum of Purchase Am...   Avg_Rating

Arizona
$4,326   3.7
Sum of Purchase Am...   Avg_Rating

Arkansas
$4,828   3.7
Sum of Purchase Am...   Avg_Rating

California
$5,605   3.8
Sum of Purchase Am...   Avg_Rating

Colorado
$4,222   3.7
Sum of Purchase Am...   Avg_Rating

Connecticut

Gender wise Total Sale

75K
(32.26%)

Gender
- Male
- Female

158K
(67.74%)

Category wise Total Sale

19K
(7.95%)

36K
(15.49...)

104K
(44.73%)

Category
- Clothing
- Accessories
- Footwear
- Outerwear

74K
(31.83%)

Page 2

| Customer Shopping Behavior Analysis

Fall Spring Summer Winter

Total Sale by Shipping Type

- Free Shipping: 41K
- Express: 39K
- Store Pickup: 39K
- Standard: 38K
- 2-Day Shipping: 38K
- Next Day Air: 38K

Total sale by Size

- 105K (45.12%) — M
- 62K (26.46%) — L
- 40K (17.36%) — S
- 26K (11.06%) — XL

Size
- M
- L
- S
- XL

Total sale wrt Payment Method

- 40K (17.29%) — Credit Card
- 40K (17....) — PayPal / Debit Card
- 40K (17.16%)
- 39K (16.62%)
- 37K (16....)
- 37K (15.68%)

Payment Method
- Credit Card
- PayPal
- Cash
- Debit Card
- Venmo
- Bank Transfer

Recommendations

Recommendations 👍

- **Increase stock & promotions for Clothing**, since it is the highest-selling category.

- **Introduce discounts and improve visibility for Outwear** to boost its low sales.

- **Target male customers with personalized offers**, as they shop more than females.

- **Stock more size M products**, and optimize inventory for S & XL due to low demand.

- **Continue free shipping**, as it is the most preferred option and improves conversions.

- **Maintain multiple payment options**, and add small digital-payment rewards to improve usage.

- **Improve product quality and service** to raise the average customer rating above 3.7.

- **Use personalized product recommendations** based on customer shopping patterns.

- **Run strong seasonal and festive campaigns** to increase overall sales.

- **Expand variety within Clothing**, the best-performing category, to attract more buyers.

Thank You…