CUSTOMER SEGMENTATION & PURCHASE ANALYSIS

U.S. SHOPPING DATASET

DATA ANALYSIS & INSIGHTS

NAME: - VIKRAM BHAGNANI

DATE: - 23/11/2025

OBJECTIVE:
ANALYZE CUSTOMER PURCHASE BEHAVIOR TO:

➢ IDENTIFY WHICH PRODUCT CATEGORIES NEED DISCOUNTS

➢ UNDERSTAND CARD SPENDING BY AGE

➢ EXPLORE SEASONAL AND LOCATION-BASED IMPACT

➢ SEGMENT CUSTOMERS INTO 3+ BEHAVIOR-BASED GROUPS

➢ PROVIDE DATA-DRIVEN MARKETING RECOMMENDATIONS

DATASET SUMMARY

➢Total Records: 25
➢Columns: Customer ID, Age, Gender, Item Purchased, Category, Amount,   Location, Season, Rating, Payment Method, etc.
➢Format: SQL → Excel → Power BI → Python
➢Goal: Generate insights + build segments + create strategy

Customer _ID	Age	Gender	Item_Purc hased	Category	Purchase_ Amount_ USD	Location	Size	Color	Season	Review_R ating	Subscripti on_Status	Payment_ Method	Shipping_ Type	Discount_ Applied	Promo_Co de_Used	Previous_ Purchases	Preferred_ Payment_ Method	Frequency _of_Purch ases
1	55	Male	Blouse	Clothing	53	Kentucky	L	Gray	Winter	3.1	Yes	Credit Card	Express	Yes	Yes	14	Venmo	Fortnightly
2	19	Male	Sweater	Clothing	64	Maine	L	Maroon	Winter	3.1	Yes	Bank Transfer	Express	Yes	Yes	2	Cash	Fortnightly
3	50	Male	Jeans	Clothing	73	Massachus etts	S	Maroon	Spring	3.1	Yes	Cash	Free Shipping	Yes	Yes	23	Credit Card	Weekly
4	21	Male	Sandals	Footwear	90	Rhode Island	M	Maroon	Spring	3.5	Yes	PayPal	Next Day Air	Yes	Yes	49	PayPal	Weekly
5	45	Male	Blouse	Clothing	49	Oregon	M	Turquoise	Spring	2.7	Yes	Cash	Free Shipping	Yes	Yes	31	PayPal	Annually

DATA PIPELINE

SQL → Excel → Python Clustering → Power BI Dashboard → Insights → Recommendation

KEY METRICS (OVERVIEW)

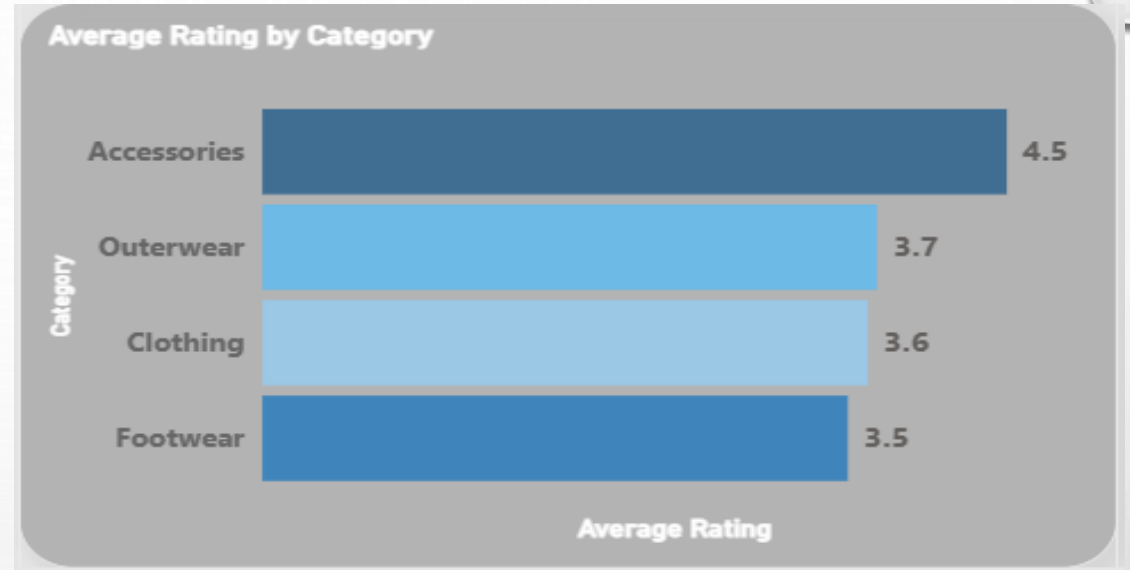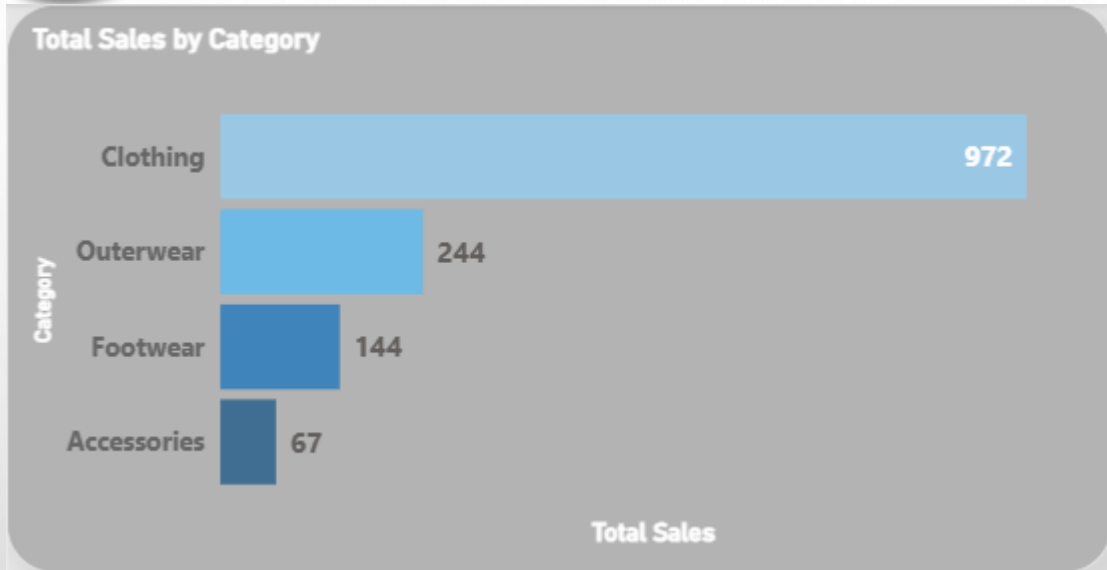| 1427	25	3.68
Total Sales	Total Customers	Average Rating

Short bullets:
➤ Clothing is the top-selling category
➤ Winter accounts for 46% of total revenue
➤ Age 31 is the highest-spending age group

PRODUCT INSIGHTS

Total Sales by Category

Category	Total Sales
Clothing	972
Outerwear	244
Footwear	144
Accessories	67

Average Rating by Category

| Category	Average Rating
Accessories	4.5
Outerwear	3.7
Clothing	3.6
Footwear	3.5

Insights:

➢Clothing drives most revenue but has lower satisfaction → **needs discount**

➢Footwear has moderate sales → **small discount**

➢Accessories have highest rating → **no discount needed**

➢Outerwear performs well mainly in Winter → **seasonal discount**

DISCOUNTS RECOMMENDATION

Discount_Applied by Category

Category	Discount_Applied
Clothing	16
Outerwear	4
Footwear	3
Accessories	2
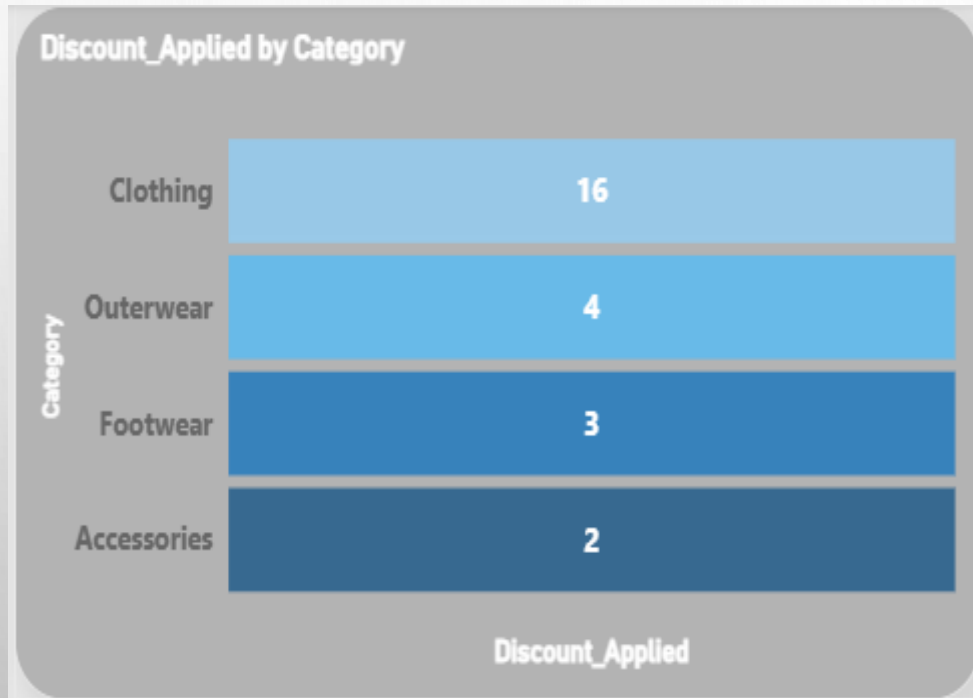
➢Apply **discounts to Clothing** (high sales, lower rating)

➢Limited, small discounts for Footwear

➢**No discount** on Accessories → strong satisfaction

➢Apply **seasonal discounts** for Outerwear

DEMOGRAPHICS INSIGHTS

Total Sales by Age

Total Sales by Season

Total Sales by Location
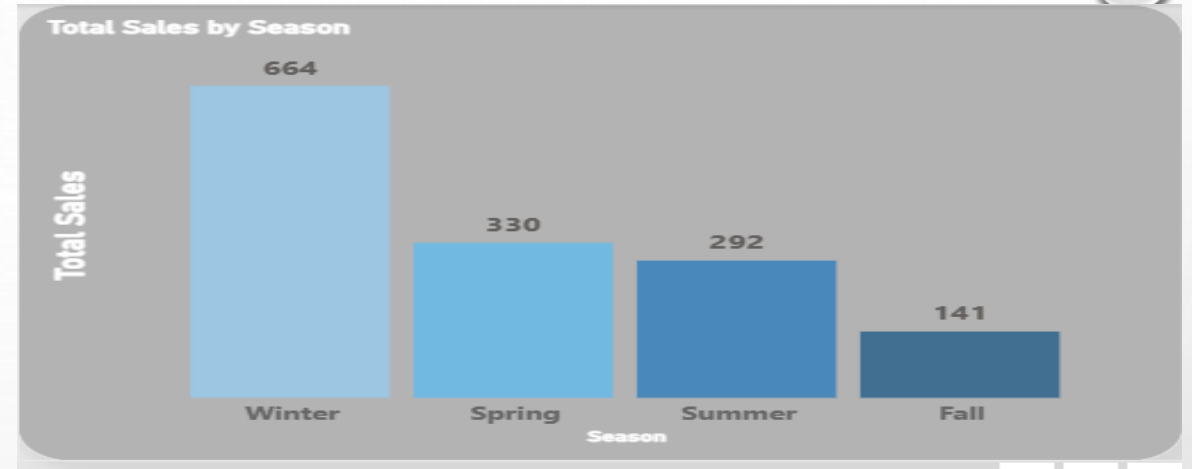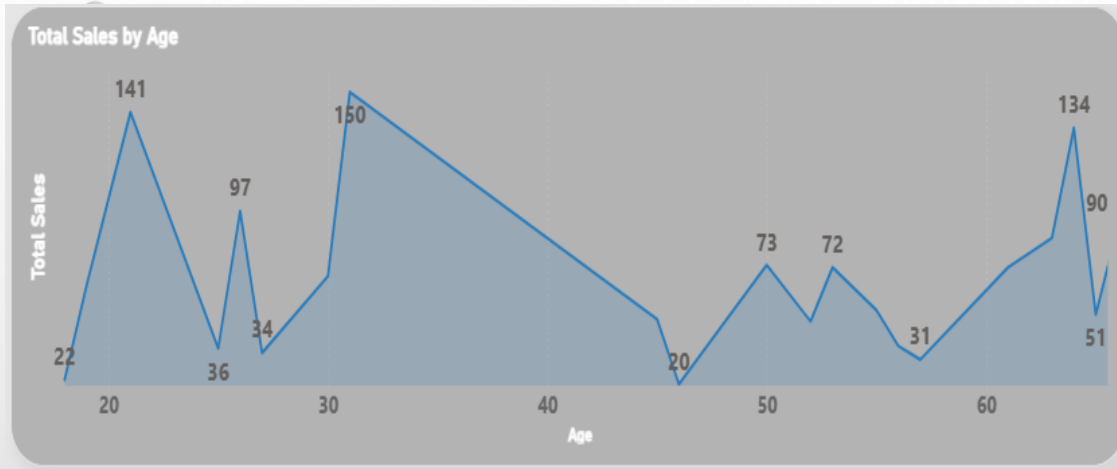
Insights:
- Highest spenders: **Age 31 & 64**
- Winter = **Peak season (46%)**
- Top spending states: Rhode Island, Montana
- Lowest spend: Wyoming, Florida → target campaigns

CLUSTERING APPROACH (PYTHON)

METHODS USED:

➢ ENCODING CATEGORICAL FEATURES

➢ STANDARDSCALER

➢ K-MEANS CLUSTERING (K = 3)

➢ FEATURES USED:

- Age
- PURCHASE AMOUNT
- RATING
- SEASON
- CATEGORY

CLUSTER PROFILES

Customer_ID	Age	Sales	Cluster
1	55	53	1
2	19	64	1
3	50	73	0
4	21	90	1
5	45	49	1
6	46	20	1
7	63	85	1
8	27	34	2
9	26	97	1
10	57	31	1
11	53	34	2
12	30	68	2
13	61	72	1
14	65	51	1
15	64	53	2
16	64	81	2
17	25	36	2
18	53	38	1
19	52	48	1
20	66	90	2
21	21	51	2
22	31	62	2
23	56	37	2
24	31	88	2
25	18	22	2
Total	1099	1427	36

Average Spends by Cluster

Cluster	Average Spends
2	55
1	58
0	73
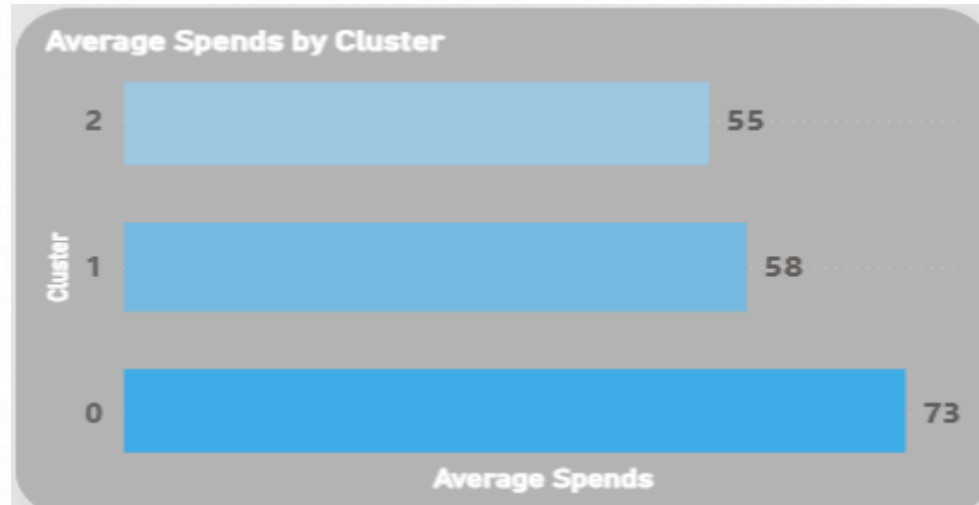
Cluster	Characteristics	Interpretation
0	Low spend, basic purchases, low variance	Budget Shoppers
1	High spenders, loyal payment methods, fast shipping	Premium Loyal Customers
2	Medium spend, variety items, convenience-based	Convenience Shoppers

RECOMMENDATIONS (MARKETING STRATEGY)

1. PRODUCT STRATEGY

➤ DISCOUNT CLOTHING

➤ PROMOTE ACCESSORIES WITHOUT DISCOUNT

➤ SEASONALLY MARKET OUTERWEAR

2. CUSTOMER STRATEGY

➤ TARGET AGES **31–35** (TOP SPENDERS)

➤ CREATE SENIOR DISCOUNTS (FOR 60+ GROUP)

3. SEASONAL STRATEGY

➤ RUN **WINTER PREMIUM CAMPAIGNS**

➤ BOOST FALL DISCOUNTS

4. REGIONAL STRATEGY

➤ HIGH-SPEND STATES → RETENTION PLANS

➤ LOW-SPEND → TARGETED PROMOTIONS

CONCLUSION

Summarize

➢Product, demographic, and seasonal patterns clarified

➢Clustering identified 3 customer groups

➢Discounts strategy optimized

➢Final recommendations support targeted marketing

➢The analysis can help increase conversion & retention

THANK YOU