

Pratik Nanda

# CUSTOMER SHOPPING BEHAVIOR ANALYSIS

Customer Behavior Dashboard



# THE BUSINESS PROBLEM

From raw data to business decisions using  
Python, SQL & Power BI  
(3,900 shopping records analyzed)

Order	Item/Purchaser	Category	Purchase/Location	Size	Color	Season	Review Rating	Subscriptions	Shipping	Discount	Promo/Coupons	Previous
10 Male	Blouse	Clothing	53 Kentucky L	Gray	Winter	3.1	Yes	Express	Yes	Yes	\$4.00	2 Cash
19 Male	Sweater	Clothing	66 Maine L	Marsan	Winter	3.1	Yes	Express	Yes	Yes	\$2.00	2 Cash
50 Male	Jeans	Clothing	73 Massachusetts S	Marsan	Spring	3.1	Yes	Free Ship	Yes	Yes	\$3.00	49 PayP
21 Male	Sandals	Footwear	90 Rhode Island M	Marsan	Spring	3.5	Yes	Next Day	Yes	Yes	\$1.00	49 PayP
45 Male	Blouse	Clothing	49 Oregon M	Turquoise	Spring	2.7	Yes	Free Ship	Yes	Yes	\$1.00	31 PayP
46 Male	Sneakers	Footwear	20 Wyoming M	White	Summer	2.9	Yes	Standard	Yes	Yes	\$4.00	34 Vene
63 Male	Short	Clothing	85 Montana M	Gray	Fall	3.2	Yes	Free Ship	Yes	Yes	\$9.00	49 Cash
27 Male	Shorts	Clothing	34 Louisiana L	Charcoal	Winter	3.2	Yes	Free Ship	Yes	Yes	\$9.00	19 Credit
26 Male	Coat	Outerwear	97 West Virginia L	Silver	Summer	2.6	Yes	Express	Yes	Yes	\$8.00	8 Vene
37 Male	Handbag	Accessories	31 Missouri M	Pink	Spring	4.8	Yes	2-Day Ship	Yes	Yes	\$4.00	4 Cash
53 Male	Shoes	Footwear	34 Arkansas L	Purple	Fall	4.3	Yes	Store Pick	Yes	Yes	\$6.00	26 Bank
30 Male	Shorts	Clothing	68 Hawaii S	Olive	Winter	4.9	Yes	Store Pick	Yes	Yes	\$10.00	10 Bank
61 Male	Coat	Outerwear	72 Delaware M	Gold	Winter	4.5	Yes	Express	Yes	Yes	\$7.00	37 Vene
85 Male	Dress	Clothing	51 New Hampshire M	Violet	Spring	4.7	Yes	Express	Yes	Yes	\$1.00	31 PayP
64 Male	Coat	Outerwear	53 New York L	Teal	Winter	4.7	Yes	Free Ship	Yes	Yes	\$4.00	34 Deli
64 Male	Skirt	Clothing	81 Rhode Island M	Teal	Winter	2.8	Yes	Store Pick	Yes	Yes	\$8.00	8 PayP
25 Male	Sunglasses	Accessories	36 Alabama S	Grey	Spring	4.3	Yes	Next Day	Yes	Yes	\$4.00	44 Deli
53 Male	Dress	Clothing	38 Mississippi M	Lavender	Winter	4.7	Yes	2-Day Ship	Yes	Yes	\$6.00	36 Vene
32 Male	Sweater	Clothing	48 Montana S	Black	Summer	4.6	Yes	Free Ship	Yes	Yes	\$7.00	37 Cash
66 Male	Pants	Clothing	90 Rhode Island M	Green	Summer	3.3	Yes	Standard	Yes	Yes	\$6.00	40 Deli
21 Male	Pants	Clothing	51 Louisiana M	Black	Winter	2.8	Yes	Express	Yes	Yes	\$6.00	2 Cash
31 Male	Pants	Clothing	62 North Carolina M	Charcoal	Winter	4.3	Yes	Store Pick	Yes	Yes	\$2.00	22 Deli
56 Male	Pants	Clothing	37 California M	Peach	Summer	3.2	Yes	Store Pick	Yes	Yes	\$2.00	32 Deli
1 Male	Pants	Clothing	88 Oklahoma M	White	Winter	4.4	Yes	Express	Yes	Yes	\$6.00	40 C
100 Male	Jacket	Outerwear	22 Florida M	Green	Fall	2.9	Yes	Store Pick	Yes	Yes	\$10.00	10 Cash

# DATA PREPARATION

Cleaned and prepared 3,900 records in Python by fixing missing values, creating age groups, and engineering purchase frequency features.

```
In [17]: df.columns  
  
Out[17]: 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', 'previous_purchases',  
       'payment_method', 'frequency_of_purchases'],  
      dtype='object')  
  
In [18]: df = df.rename(columns={"purchase_amount_(USD)": "purchase_amount"})  
  
In [19]: df.columns  
  
Out[19]: Index(['customer_id', 'age', 'gender', 'item_purchased', 'category',  
       'purchase_amount', 'location', 'size', 'color', 'season',  
       'review_rating', 'subscription_status', 'shipping_type',  
       'discount_applied', 'promo_code_used', 'previous_purchases',  
       'payment_method', 'frequency_of_purchases'],  
      dtype='object')  
  
In [20]: labels = ['Young Adult', 'Adult', 'Middle-aged', 'Senior']  
df['age_group'] = pd.qcut(df['age'], q=4, labels=labels)  
  
In [21]: df[['age', 'age_group']].head(10)
```

# BUSINESS ANALYSIS WITH SQL

Connected the dataset to PostgreSQL and answered key business questions on revenue, customer segments, product performance, discounts, and subscriptions.

```
SELECT item_rank, category, item_purchased, total_orders
FROM item_counts
WHERE item_rank <=3;

--Q9. Are customers who are repeat buyers (more than 5 previous purchases) also
SELECT subscription_status,
       COUNT(customer_id) AS repeat_buyers
FROM customer
WHERE previous_purchases > 5
GROUP BY subscription_status;

--Q10. What is the revenue contribution of each age group?
SELECT
    age_group,
    SUM(purchase_amount) AS total_revenue
FROM customer
GROUP BY age_group
ORDER BY total_revenue DESC;
```

# INTERACTIVE DASHBOARD

Designed a Power BI dashboard displaying KPIs, customer segments, product performance, and revenue contribution by age group.



# KEY INSIGHTS

- 1) Male customers generated \$157,890 vs \$75,191 by female customers.
- 2) Young adults contributed the highest revenue (~\$62K).
- 3) Repeat buyers are significantly more likely to be subscribed, and subscribers show higher average spend and total revenue than non-subscribers

**% of Customers by Subscription St...**



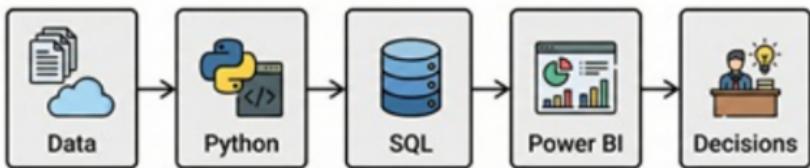
## BUSINESS RECOMMENDATIONS

Reward loyal customers, reduce excessive discounting, prioritize high-value age groups, and promote top-rated products to maximize profitability.

# WHY THIS MATTERS

This end-to-end analytics workflow is repeatable and can help any business convert raw data into data-driven decisions.

Project Data Pipeline



## **FINAL TAKEAWAY**

A complete Data Analytics project demonstrating business impact using Python, SQL, and Power BI.



Pratik Nanda

"DATA/BUSINESS ANALYST | PYTHON  
| SQL | POWER BI"

