Customer Shopping Behavior Analysis: Unlocking Insights with Python, SQL & Power BI

Sahil Kumar (2300271540105)

Why Analyze Customer Shopping Behavior?

Understand Patterns

Drives smarter marketing and inventory decisions.

Increase Revenue

Target high-value customers and optimize offers.

Powerful Insights

Combines SQL and Python for robust analysis and visualization.

The Dataset: Real Customer Transactions

Our analysis is built upon a rich dataset comprising over **3,900 customer transactions**.

Rich Features

Customer demographics, product details, purchase amounts, and behaviors.

Key Variables

Age, Gender, Product Category, Purchase Frequency, Discounts, and Payment Methods.

Multi-Dimensional

Enables deep analysis of who buys what, when, and how.

Step 1: Data Extraction & Cleaning

with SQL and Python

01

SQL for Data Handling

Used to import, normalize, and efficiently query large datasets.

02

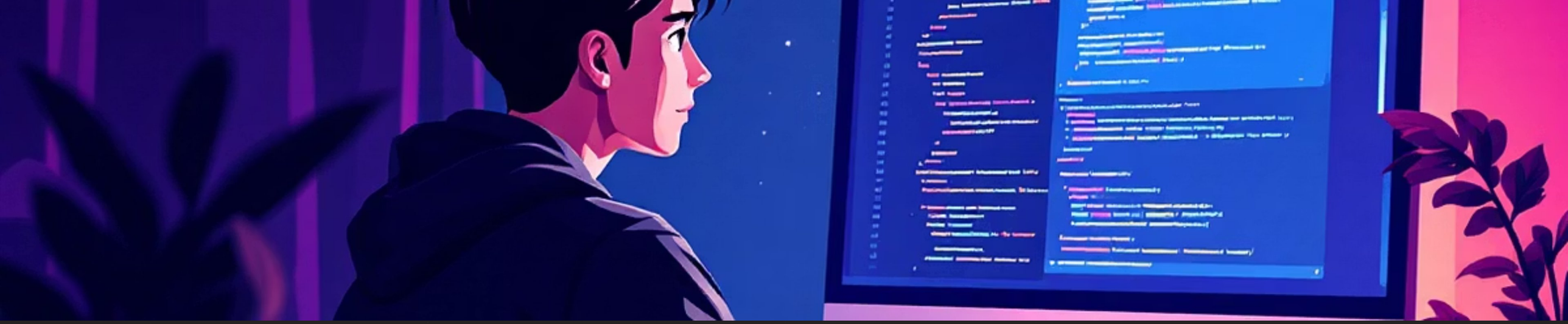
Python (Pandas) for Cleaning

Handles missing values, duplicates, and feature engineering (e.g., age groups).

03

Maintaining Data Quality

Example: Filling missing review ratings with the median per product category.



Step 2: Analyzing Customer Behavior with Python

Calculated KPIs

Total revenue, average purchase value, and purchase frequency trends.

RFM Analysis

Identified top customers by spend and segmented using Recency, Frequency, Monetary.

Customer Grouping

Clustering (K-Means) to group customers by shopping habits for targeted marketing.

Step 3: Visualizing KPIs to Drive Business Decisions

01

Python Libraries

Matplotlib, Seaborn create clear charts for product categories, revenue by gender, discount impact.

02

Power BI Dashboards

Interactive views of purchase trends, customer segments, and product performance.

03

Dynamic Filters

Enable stakeholders to explore data by time period, location, or customer type.

These are some of the Business Question

Q1. What is the total revenue generated by male vs. female customers?

Q2. Which customers used a discount but still spent more than the average purchase amount?

Q3. Which are the top 5 products with the highest average review rating?

Q4. Compare the average Purchase Amounts between Standard and Express Shipping.

Q5. Do subscribed customers spend more? Compare average spend and total revenue between subscribers and non-subscribers.

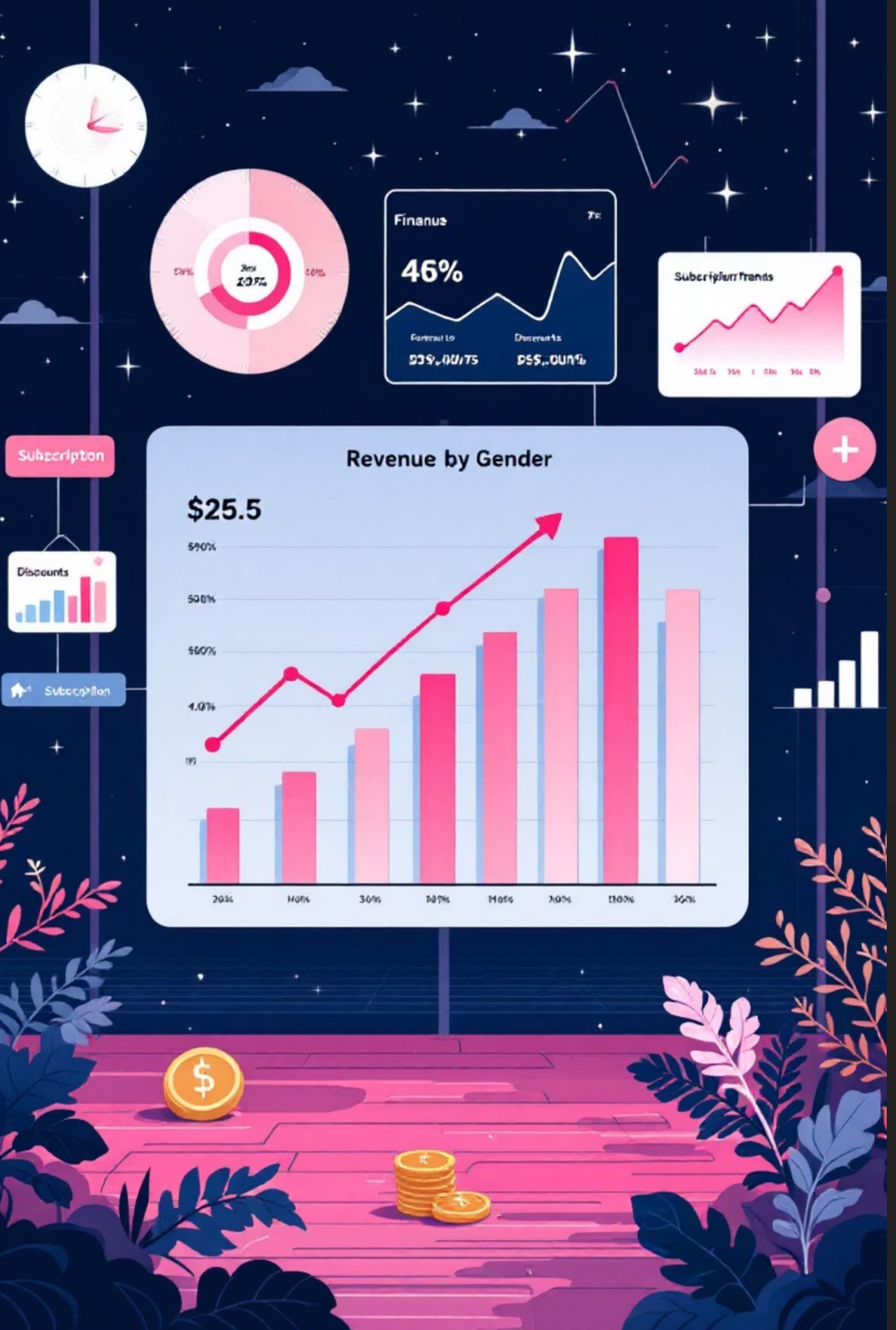
Q6. Which 5 products have the highest percentage of purchases with discounts applied?

Q7. Segment customers into New, Returning, and Loyal based on their total number of previous purchases, and show the count of each segment.

Q8. What are the top 3 most purchased products within each category?

Q9. Are customers who are repeat buyers (more than 5 previous purchases) also likely to subscribe?

Q10. What is the revenue contribution of each age group?



Key Insights Discovered

Gender Revenue

Female customers generated 15% more revenue on average than male customers.

Discount Impact

Discount usage correlated with higher purchase frequency, but not always higher spend.

Loyalty Value

Subscription customers spent 25% more on average, highlighting program value.

Seasonal Trends

Spikes in footwear purchases observed during Fall and Winter seasons.

Visual Storytelling: Sample Dashboard Highlights

These visualizations transform complex data into easily digestible insights, empowering quick decision-making.

CUSTOMER BEHAVIOR DASHBOARD

subscription_status

No | Yes

gender

Female | Male

Category

- Accessories
- Clothing
- Footwear
- Outerwear

Shipping_type

- 2-Day Shipping
- Express
- Free Shipping
- Next Day Air
- Standard
- Store Pickup

3.9K
Number of customer

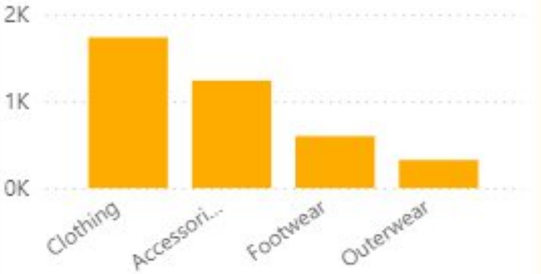
$59.76
Average of purchase amount

3.75
Average Review Rating

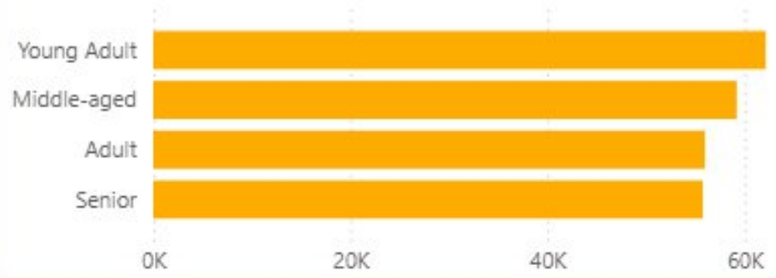
% of Customer by Subscription Status

Revenue By Category

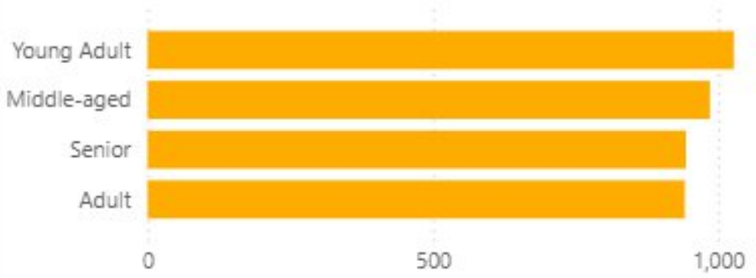
Sales By Category



Revenue By Age Group



Sales By Age Group



Business Impact & Next Steps

Personalized Marketing

Insights enable targeted campaigns and optimized inventory management.

Scalable Workflow

SQL and Python setup supports continuous monitoring and deeper analysis.

Future Enhancements

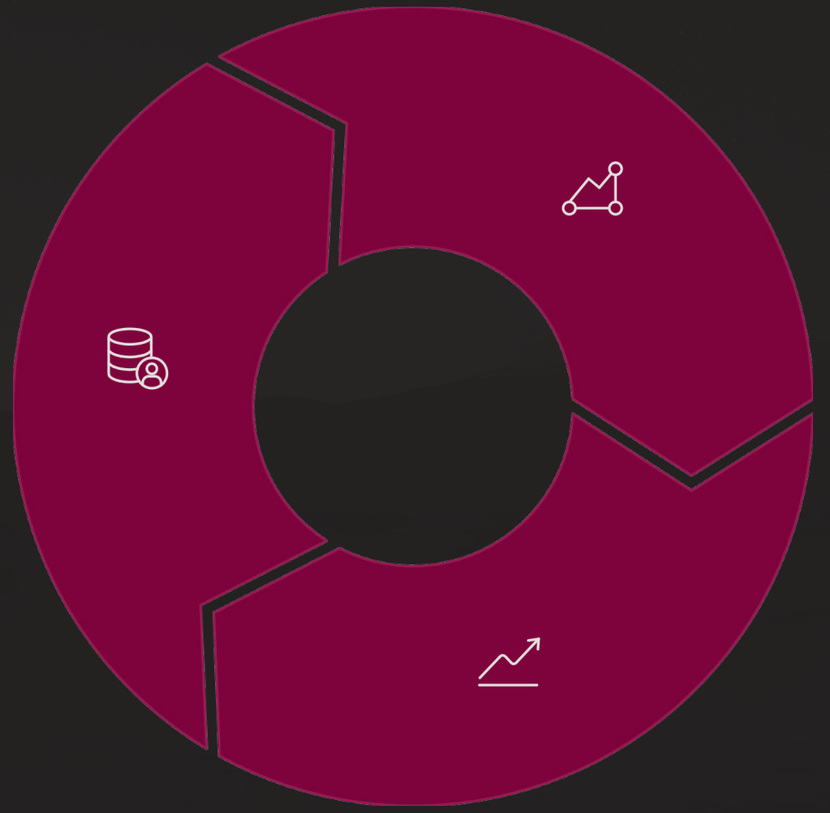
Integrate real-time data and apply predictive models for churn and lifetime value.

Conclusion: Empowering Data-Driven Decisions

This integrated approach drives growth, customer satisfaction, and competitive advantage.

Unlocking Data

Combining SQL, Python, and visualization tools transforms raw customer data.



Actionable KPIs

KPIs turn transactions into strategic business insights.

Strategic Growth

Drives customer satisfaction and competitive advantage.