Customer Shopping Behavior Analysis

Uncovering insights into spending patterns, customer segments, and product preferences through data analytics.

Project Overview & Dataset

◎ Project Goal

This project analyzes transactional data to uncover key insights into customer behavior. By examining patterns across 3,900 purchases, we aim to guide strategic business decisions regarding inventory, marketing, and customer retention.

🛢 Dataset Summary

Volume: 3,900 Rows, 18 Columns

Demographics: Age, Gender, Location, Subscription

Metrics: Purchase Amount, Frequency, Ratings

Note: Missing ratings were imputed using category medians.

Analysis Methodology

Python EDA

Data cleaning and preparation. Implemented feature engineering to create 'Age Groups' and 'Purchase Frequency' metrics using pandas.

SQL Analysis

Performed structured queries on PostgreSQL to extract business metrics like revenue per gender and shipping comparisons.

Power BI

Designed an interactive dashboard to visualize trends in subscription status, category sales, and demographic splits.

Revenue Distribution by Gender
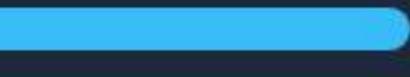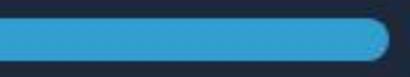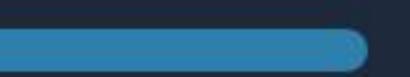
Total Revenue Share

- Male: **$157,890 (67%)**
- Female: **$75,191 (33%)**
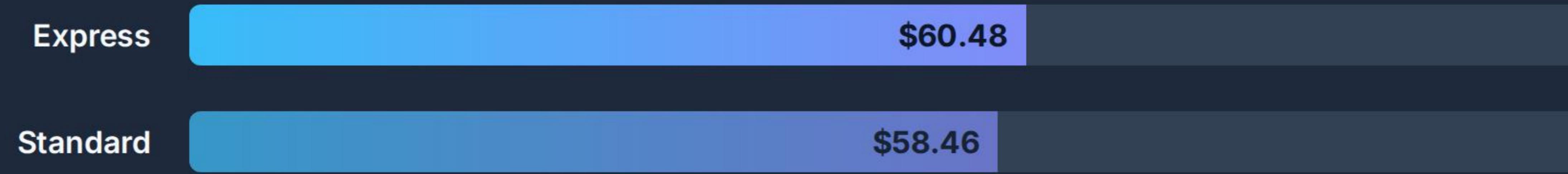
Male customers generate nearly double the revenue of female customers in this dataset, indicating a strong male consumer base.

Top 5 Products by Average Rating

Rank	Item Purchased	Average Rating (Out of 5.0)	Visual Indicator
1	Gloves	3.86	
2	Sandals	3.84	
3	Boots	3.82	
4	Hat	3.80	
5	Skirt	3.78	

Accessories and footwear consistently perform best in customer reviews.

Shipping Type vs. Average Spend

Express	$60.48
Standard	$58.46

🚚 Customers choosing **Express Shipping** tend to spend slightly more per order. Targeting these users with premium offers could yield higher returns.

Subscription Status Impact

73%

Are Non-Subscribers

Revenue Volume vs. Avg Spend

Although the majority of revenue ($170k) comes from non-subscribers due to their sheer volume, the **Average Spend** is remarkably similar between groups:

- **Subscribers:** $59.49 avg spend
- **Non-Subscribers:** $59.87 avg spend

Opportunity: Convert high-volume non-subscribers to recurring revenue models.

Customer Segmentation

We classified customers based on purchase history:

Loyal Customers

3,116

Returning

701

New

83

Products with Highest Discount Rates

Discount Strategy

Certain items are heavily discounted to drive sales. The products listed on the right have the highest percentage of purchases where a discount was applied.

Insight: High discount rates on staple items like Hats and Sneakers suggest price sensitivity in these categories.

Item	Discount Rate
Hat	50.00%
Sneakers	49.66%
Coat	49.07%
Sweater	48.17%
Pants	47.37%

Revenue by Age Group

Even Distribution

Revenue is remarkably consistent across all age demographics. There is no single dominant age group, though "Young Adults" contribute marginally more.

This suggests a broad market appeal for the product lines, allowing for mass-market marketing strategies rather than niche targeting.

Young Adult	$62,143
Middle-aged	$59,197
Adult	$55,978
Senior	$55,763

Business Recommendations

Boost Subscriptions

Promote exclusive benefits. With only 27% of users subscribed, there is significant room to grow recurring revenue.

Customer Loyalty Programs

Reward repeat buyers to transition them into the "Loyal" segment, increasing lifetime value.

Review Discount Policy

Balance sales boosts with margin control, especially for high-discount items like Hats and Sneakers.

Questions?

Thank you for your attention.

Image Sources

https://png.pngtree.com/png-clipart/20250207/original/pngtree-crowd-of-people-various-ages-and-styles-enjoying-a-lively-shopping-png-image_19524132.png

Source: pngtree.com

https://img.freepik.com/premium-photo/blurred-defocused-background-modern-clothing-store_79762-9411.jpg

Source: www.freepik.com