



Customer Shopping Behavior Analysis

Data-Driven Strategic Insights into Consumer Trends,
Category Performance, and Subscription Growth



@Nilesh_Patil_1402



Project Overview

This project analyses customer shopping behaviour using transactional data from 3,900 purchases across various product categories. The goal is to uncover insights into spending patterns, customer segments, product preferences, and subscription behaviour to guide strategic business decisions.

METHODOLOGICAL WORKFLOW



Python EDA

Advanced cleaning using Pandas, handling missing ratings via category-wise median, and feature engineering for age segmentation.



SQL Analysis

Structured querying in PostgreSQL to analyze revenue by demographics, product performance, and subscription value.



Power BI Viz

Transformation of metrics into interactive dashboards for dynamic tracking of KPIs and category performance.

DATA PREPARATION STRATEGY



Missing Value Imputation: Filled 37 missing values in "Review Rating" using `transform(lambda x: x.fillna(x.median()))` to maintain category integrity.



SQL-Friendly Refactoring: Standardized column names to lowercase and replaced spaces with underscores for seamless database integration.



Redundancy Control: Identified 100% correlation between "promo_code_used" and "discount_applied," removing the redundant feature to clean the feature set.

Data Export: Final cleaned dataset exported to a relational structure in PostgreSQL via SQLAlchemy.

KEY FEATURE ENGINEERING



Age Segmentation

Implemented `pd.qcut` to create equitable demographic tiers:

- Young Adult
- Adult
- Middle Aged
- Senior

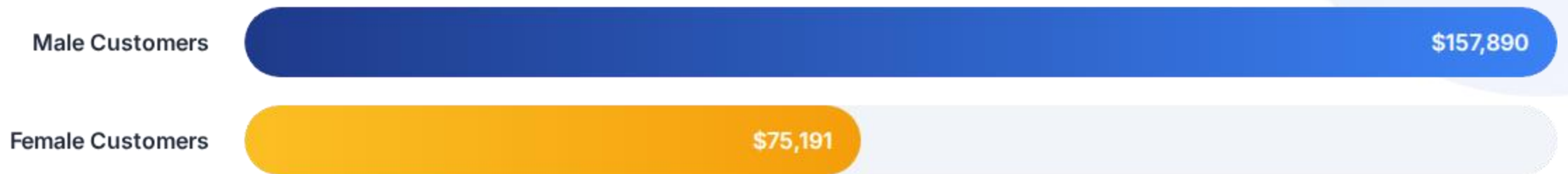


Frequency Mapping

Converted categorical descriptions into numeric "Days" for quantitative trend analysis:

- Weekly → 7 Days
- Fortnightly → 14 Days
- Annually → 365 Days

REVENUE BY GENDER PROFILE



Finding: Male customers contribute approximately 67% of total revenue, representing a primary market segment.

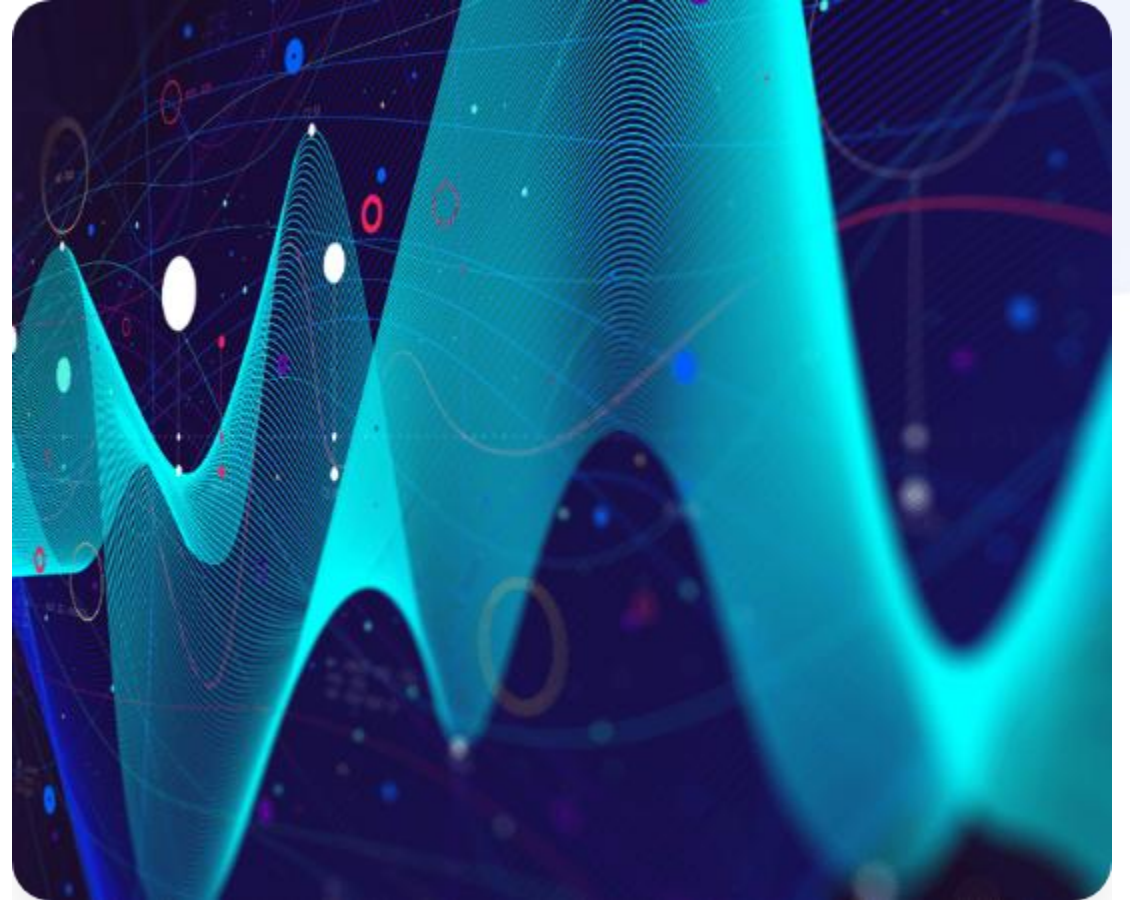
TOP PRODUCTS BY REVIEW RATING

Item Purchased	Product Category	Avg. Review Rating	Performance Tier
Gloves	Accessories	3.86	Top Tier
Sandals	Footwear	3.84	Top Tier
Boots	Footwear	3.82	High Performing
Hat	Accessories	3.80	High Performing
Skirt	Clothing	3.78	Strong

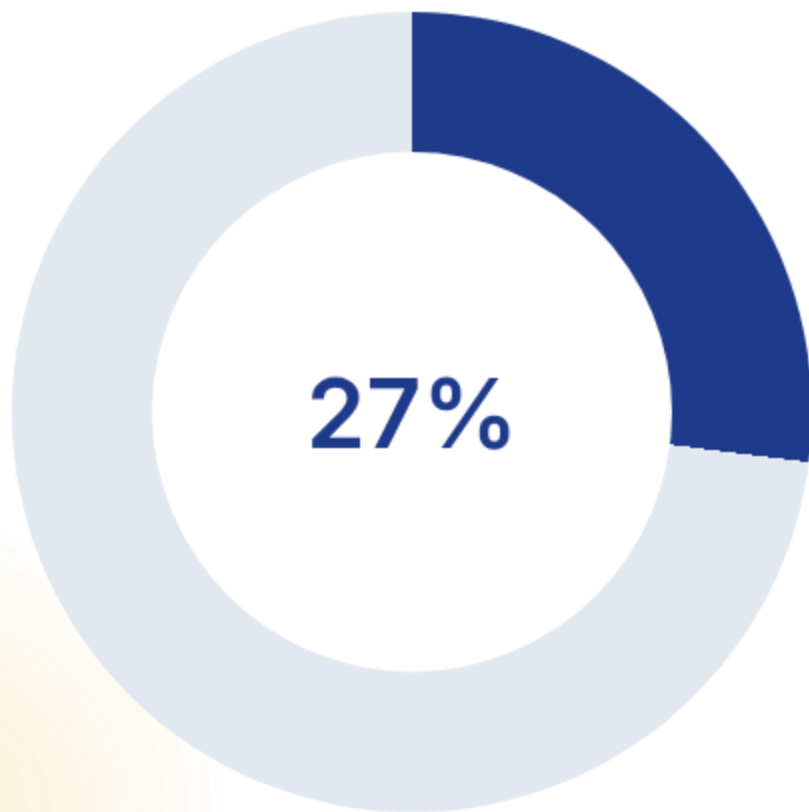
DISCOUNT ADOPTION TRENDS

A significant portion of top-selling items rely heavily on promotional activity:

- 🏷️ **Hats:** 50% purchases with discount.
- 🏷️ **Sneakers:** 49.66% purchases with discount.
- 🏷️ **Coats:** 49.07% purchases with discount.



SUBSCRIPTION PENETRATION



■ Non-Subscribers (73%)

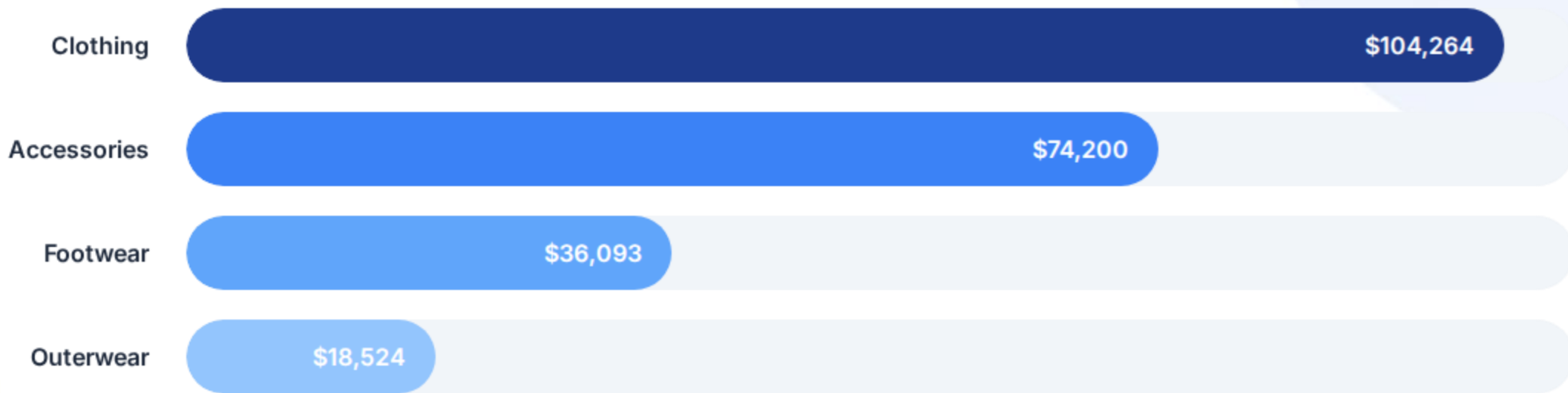
■ Subscribers (27%)

Revenue Comparison:

Subscribers Average Spend: \$59.49

Non-Subscribers Average Spend: \$59.87

REVENUE CONTRIBUTION BY CATEGORY



Clothing remains the dominant revenue driver, contributing nearly 45% of total sales volume.

STRATEGIC GROWTH RECOMMENDATIONS



Subscription Growth

Implement a high-value loyalty program to convert the 73% non-subscriber base into recurring customers.



Target Marketing

Prioritize digital campaigns for the Young Adult and Middle-Aged segments who show highest active volume.



Outerwear Focus

Launch seasonal bundles and targeted promotions to improve Outerwear category performance and ratings.

Final Dashboard in Power BI

Customer Shopping Behavior Dashboard

SubscriptionWise

No

Yes

GenderWise

Female

Male

CategoryWise

Accessories

Clothing

Footwear

Outerwear

ShipingTypeWise

☐ 2-Day Shipping

☐ Express

☐ Free Shipping

☐ Next Day Air

☐ Standard

☐ Store Pickup

Number of customers

3.9K

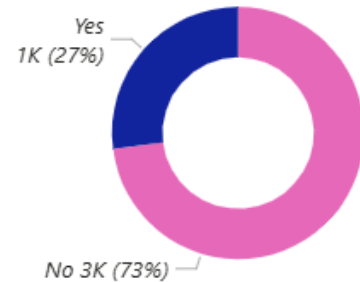
Average purchase smount

\$59.76

Average review rating

3.75

% of customer by subscription status



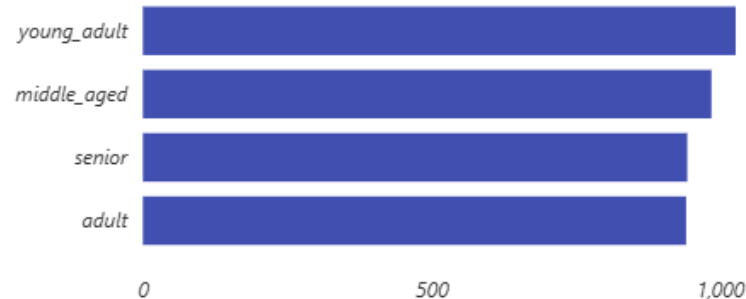
Revenue by category



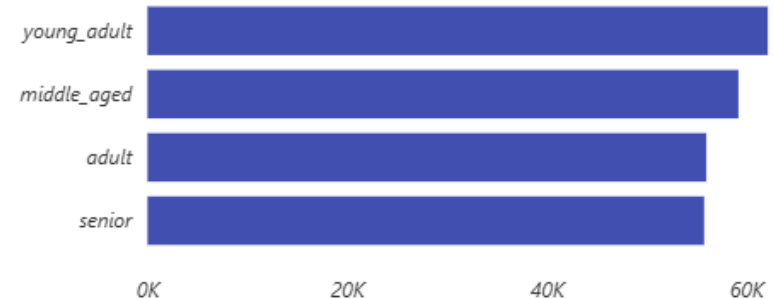
Customer by category



Sales by age-group



Revenue by age-group



Thank you for your time and attention.

Project Create By

Nilesh Patil

Analysis Scope

3.9K Transactions



<https://www.linkedin.com/in/nilesh-patil-347810362/>

Any Questions?

