



Customer Shopping Behavior Analysis

Uncovering strategic insights from 3,900 customer transactions to drive data-informed business decisions across product categories, customer segments, and purchasing patterns.



Dataset Overview

3,900

Total Transactions

Customer purchases analyzed across multiple categories

18

Data Columns

Comprehensive features tracking behavior and demographics

37

Missing Values

Only in Review Rating column, handled through imputation

The dataset captures rich customer information including demographics (Age, Gender, Location, Subscription Status), detailed purchase information (Item Purchased, Category, Purchase Amount, Season, Size, Color), and behavioral indicators (Discount Applied, Promo Code Used, Previous Purchases, Frequency of Purchases, Review Rating, Shipping Type). This comprehensive data structure enables multi-dimensional analysis of shopping patterns and customer preferences.



Data Preparation & Python Analysis

Data Loading & Exploration

- Imported dataset using pandas
- Used df.info() to check structure
- Applied .describe() for summary statistics
- Verified data types and distributions

Cleaning & Standardization

- Checked for null values across all columns
- Imputed missing Review Rating values using median rating per product category
- Renamed columns to snake_case format
- Verified data consistency

01

Feature Engineering

Created age_group column by binning customer ages into meaningful segments for demographic analysis

02

Purchase Frequency Calculation

Generated purchase_frequency_days column from purchase data to track customer engagement patterns

03

Redundancy Check

Verified if discount_applied and promo_code_used were redundant; dropped promo_code_used to streamline dataset

04

Database Integration

Connected Python script to PostgreSQL and loaded cleaned DataFrame for SQL analysis

Revenue Analysis by Gender

Male Customers

Total revenue generated by male customer segment analyzed to understand spending patterns and purchasing power across product categories

Female Customers

Total revenue generated by female customer segment compared against male counterparts to identify gender-based preferences and opportunities

We performed structured SQL analysis in PostgreSQL to compare total revenue generated by male versus female customers. This gender-based revenue comparison reveals important insights about customer value, product preferences, and potential marketing opportunities. Understanding these patterns enables targeted campaigns and product positioning strategies that resonate with each demographic segment.



Discount & Spending Behavior

High-Spending Discount Users

Our analysis identified customers who used discounts but still spent above the average purchase amount. This segment represents valuable customers who are price-conscious yet willing to invest in quality products. Understanding this behavior helps optimize discount strategies to attract customers without unnecessarily eroding margins on purchases that would have occurred at full price.

Top 5 Products by Rating

We identified products with the highest average review ratings to understand which items deliver exceptional customer satisfaction. These top-rated products represent opportunities for featured positioning in marketing campaigns and can serve as anchor products to build customer trust and drive conversion rates.

Key Insights

- Discount users aren't always low spenders
- Quality products maintain high ratings
- Customer satisfaction drives repeat purchases
- Strategic discounting can attract premium buyers



Shipping & Subscription Insights

1

Shipping Type Comparison

Compared average purchase amounts between Standard and Express shipping options. Express shipping users tend to have different purchasing behaviors and may represent higher-value customer segments willing to pay premium for convenience.

2

Subscribers vs. Non-Subscribers

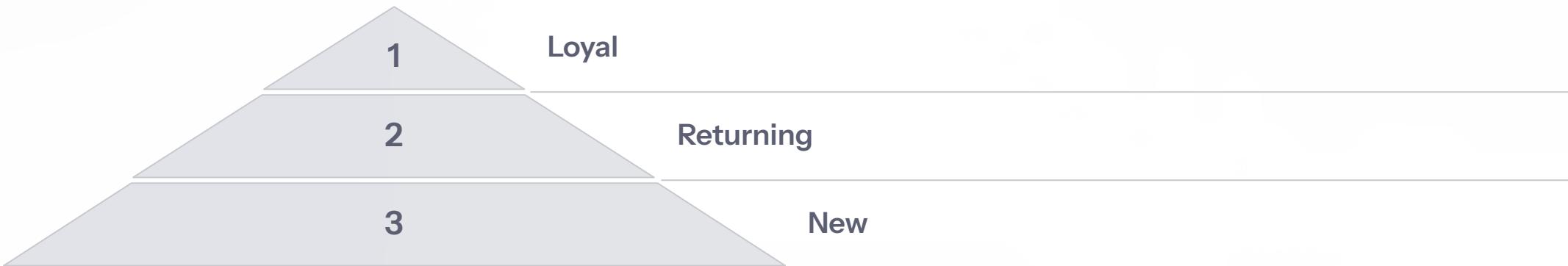
Analyzed average spend and total revenue across subscription status. Subscribers typically demonstrate higher lifetime value, more frequent purchases, and stronger brand loyalty compared to one-time buyers.

3

Discount-Dependent Products

Identified 5 products with the highest percentage of discounted purchases. These items may require pricing strategy review or could be strategically used as loss leaders to drive traffic and cross-selling opportunities.

Customer Segmentation Strategy



We classified customers into three distinct segments based on purchase history: **New customers** (first-time buyers requiring onboarding and conversion strategies), **Returning customers** (repeat buyers showing initial loyalty signals), and **Loyal customers** (high-frequency purchasers representing the most valuable segment). This segmentation enables targeted marketing approaches, personalized communication strategies, and differentiated retention programs for each group.

Repeat Buyers & Subscriptions

Analysis revealed whether customers with more than 5 purchases are more likely to subscribe. This correlation helps identify the optimal point in the customer journey to promote subscription benefits and convert high-frequency buyers into committed subscribers.

Strategic Implications

Understanding the relationship between purchase frequency and subscription adoption enables data-driven decisions about when and how to present subscription offers for maximum conversion rates.

Product Performance Analysis

Top Products by Category

Listed the most purchased products within each category to identify bestsellers and customer preferences across different product lines

Our comprehensive product analysis identified the top 3 products per category, revealing which items drive the most sales volume within each product line. Combined with age group revenue analysis, these insights enable precise inventory management, targeted product development, and age-specific marketing strategies that maximize revenue potential across all customer demographics.

Highest Rated Items

Products with exceptional review ratings represent quality benchmarks and opportunities for featured marketing campaigns

Revenue by Age Group

Calculated total revenue contribution of each age group to understand demographic spending power and targeting opportunities

Interactive Power BI Dashboard



We built an interactive dashboard in Power BI to present insights visually, enabling stakeholders to explore data dynamically and make informed decisions in real-time. The dashboard consolidates all key metrics including revenue by gender, customer segmentation, product performance, subscription analysis, and demographic breakdowns into an intuitive, user-friendly interface.



Visual Analytics

Interactive charts and graphs bring data to life with drill-down capabilities



Dynamic Filtering

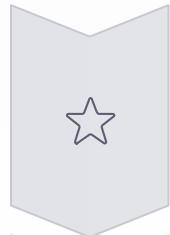
Slice and dice data by category, time period, demographics, and behavior



Real-Time Updates

Connected to live data sources for up-to-date business intelligence

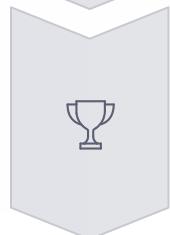
Strategic Business Recommendations



Boost Subscriptions



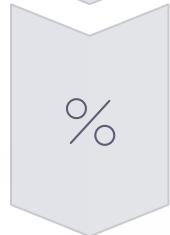
Promote exclusive benefits for subscribers including early access to new products, special pricing, and VIP customer service to convert high-frequency buyers into committed subscribers



Customer Loyalty Programs



Reward repeat buyers with points, tiered benefits, and exclusive perks to move them from the "Returning" segment into the high-value "Loyal" segment



Review Discount Policy



Balance sales boosts with margin control by strategically applying discounts to specific products, customer segments, and time periods rather than blanket promotions



Product Positioning



Highlight top-rated and best-selling products in marketing campaigns to leverage social proof and drive conversion rates across all channels



Targeted Marketing



Focus marketing efforts on high-revenue age groups and express-shipping users who demonstrate higher purchase values and willingness to pay premium for convenience