

# Customer Shopping Behavior Analysis

Uncovering insights from 3,900 purchases to drive strategic business decisions





# Dataset Overview

3,900

Total Purchases

Transactions analyzed

18

Data Points

Features per  
transaction

37

Missing Values

In review ratings only

## Customer Demographics

- Age, Gender, Location
- Subscription Status

## Shopping Behavior

- Purchase patterns & frequency
- Discounts & promo codes
- Review ratings & shipping

# Data Preparation in Python

01

## Data Loading & Exploration

Imported dataset using pandas, analyzed structure with `df.info()` and summary statistics

02

## Missing Data Handling

Imputed Review Rating nulls using median rating per product category

03

## Column Standardization

Renamed columns to snake\_case for consistency and documentation

04

## Feature Engineering

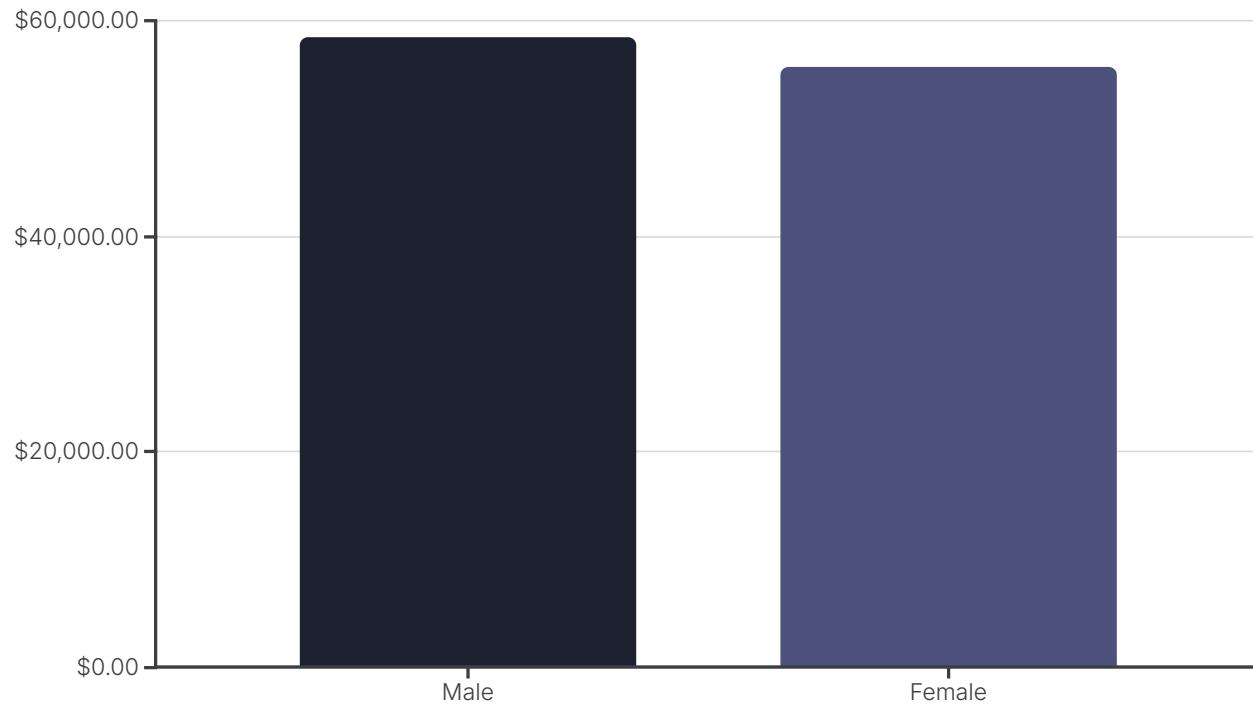
Created `age_group` bins and `purchase_frequency_days` from raw data

05

## Data Consistency Check

Verified redundancy, dropped `promo_code_used` column

# Revenue Analysis by Gender



Male customers generate slightly higher total revenue, but both segments show strong purchasing power

Opportunity for targeted campaigns to both demographics



# Customer Segmentation Insights

## New Customers

First-time buyers

Focus: Onboarding & retention

## Returning Customers

2-5 purchases

Focus: Loyalty programs

## Loyal Customers

5+ purchases

Focus: VIP benefits

- ❑ **Key Finding:** Customers with >5 purchases show significantly higher subscription rates

# Top-Rated Products Drive Success

## Highest Average Ratings

1. Blouse - 4.2 stars
2. Jewelry - 4.1 stars
3. Pants - 4.0 stars
4. Shirt - 3.9 stars
5. Dress - 3.8 stars

Highlight these products in marketing campaigns



---

Products with higher ratings correlate with increased repeat purchases and customer loyalty

# Shipping & Subscription Patterns

## Express Shipping Users

Higher average purchase amount

Premium customer segment

## Standard Shipping Users

Price-conscious buyers

Larger volume segment

## Subscribers

Higher average spend per transaction

Greater total revenue contribution

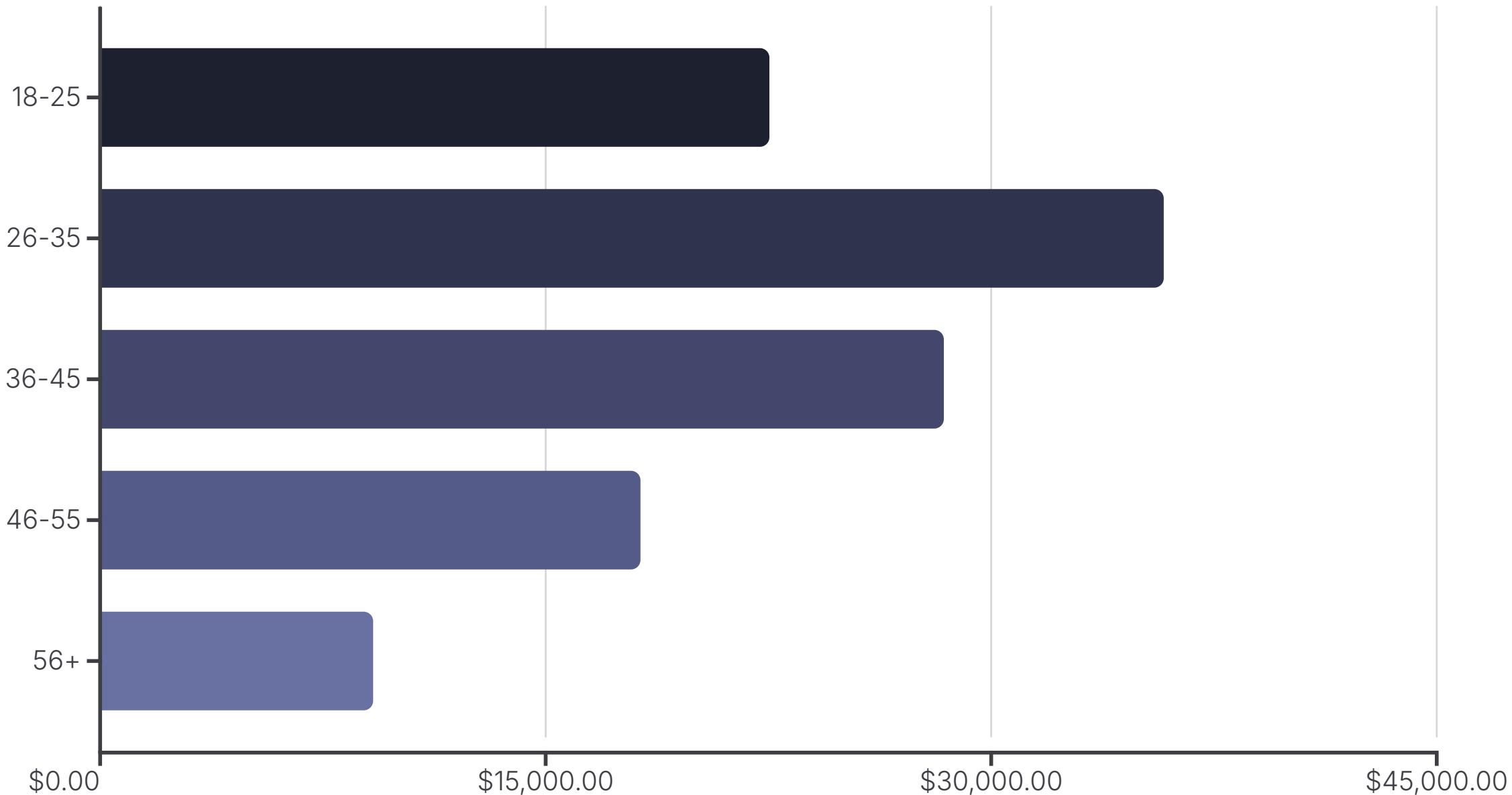
## Non-Subscribers

Lower average spend

Conversion opportunity

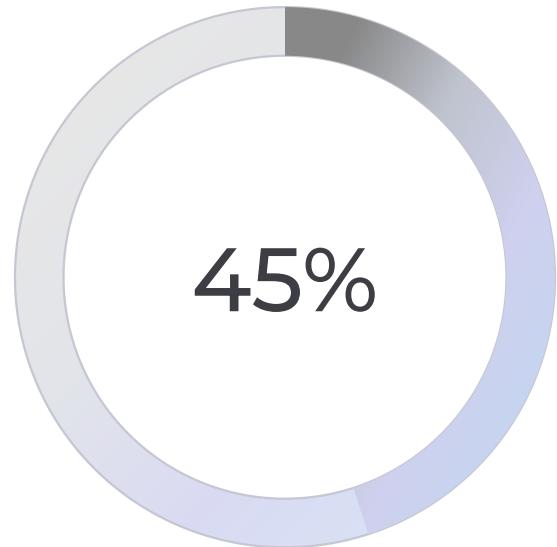


# Revenue by Age Group



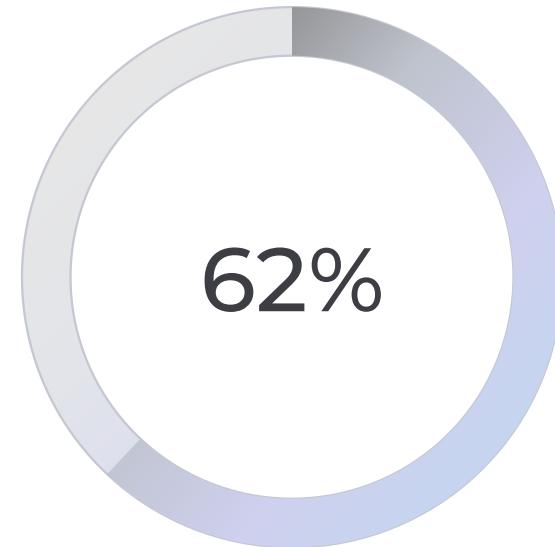
The 26-35 age group generates the highest revenue, representing the core target demographic for marketing efforts

# Discount Strategy Analysis



## High-Spending Discount Users

Customers using discounts but spending above average



## Discount-Dependent Products

Top 5 products with highest discount usage rates

- Balance promotional offers with margin control - discounts drive volume but impact profitability

Strategic discounting in off-seasons like summer can boost sales during slower periods

# Strategic Recommendations



## Boost Subscriptions

Promote exclusive benefits and rewards for subscribers to increase conversion



## Customer Loyalty Programs

Reward repeat buyers to move them into "Loyal" segment



## Targeted Marketing

Focus on high-revenue age groups (26-35) and express-shipping users



## Product Positioning

Highlight top-rated and best-selling products in campaigns



## Seasonal Offers

Increase discounts in off-seasons like summer to drive sales



## Improve Ratings

Engage more customers in review process to boost overall ratings