Customer Shopping Behavior Analysis

Understanding customer behavior is crucial for improving sales, satisfaction, and loyalty. We'll analyze purchasing patterns across demographics, products, and channels to optimize strategies.

The Business Challenge

Changing Patterns

Management observed shifts in purchasing across demographics, categories, and sales channels.

Key Drivers

Identify factors like discounts, reviews, seasons, and payment preferences influencing decisions and repeat purchases.

Our goal: Leverage consumer shopping data to identify trends, improve engagement, and optimize marketing and product strategies.

Project Overview & Deliverables

Data Preparation & Modeling

Clean and transform raw data using Python.

Data Analysis

Extract insights on segments, loyalty, and purchase drivers using SQL.

Visualization & Insights

Build an interactive Power BI dashboard for data-driven decisions.

Report & Presentation
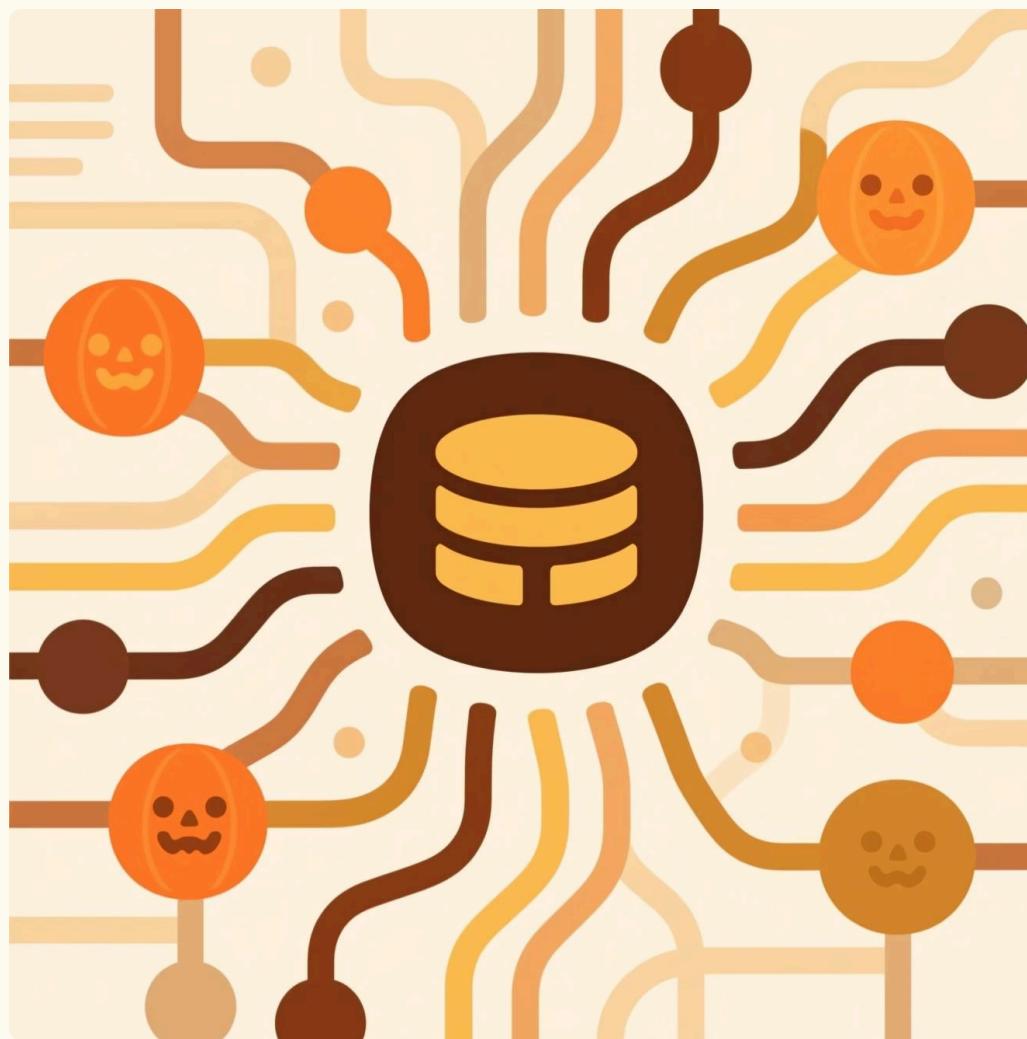
Summarize findings and recommendations for stakeholders.

This project analyzes 3,900 purchases to uncover spending patterns, customer segments, and product preferences.

Dataset Summary

Key Statistics

- **Rows:** 3,900
- **Columns:** 18
- **Missing Data:** 37 values in "Review Rating"

Key Features

- **Customer Demographics:** Age, Gender, Location, Subscription Status
- **Purchase Details:** Item, Category, Amount (USD), Season, Size, Color
- **Shopping Behavior:** Discount, Promo Code, Previous Purchases, Frequency, Review Rating, Shipping Type

Exploratory Data Analysis (Python)

01

Data Loading & Exploration

Imported dataset with pandas; used df.info() and df.describe() for initial checks.

02

Missing Data Handling

Imputed missing "Review Rating" values using the median rating per product category.

03

Column Standardization

Renamed columns to snake_case for improved readability.

04

Feature Engineering

Created age_group and purchase_frequency_days columns.

05

Database Integration

Loaded cleaned data into MySQL for SQL analysis.
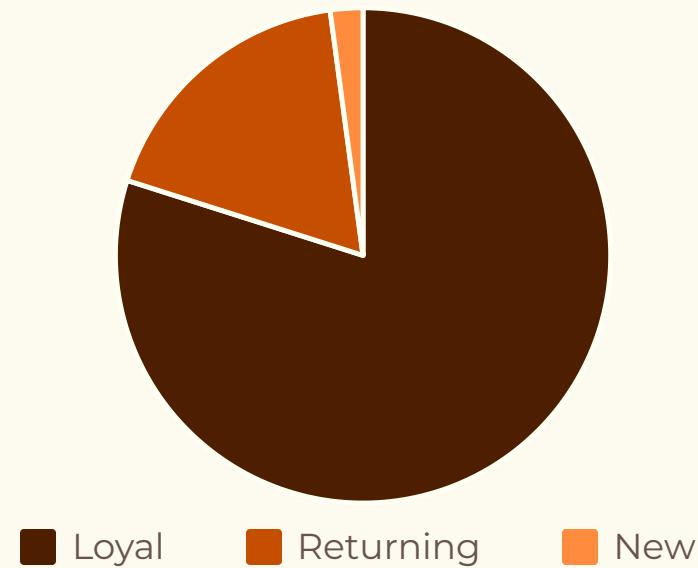
SQL Analysis: Key Business Insights

1

Revenue by Gender

Male: $157,890, Female: $75,191

2

Top 5 Products by Rating

Gloves (3.86), Sandals (3.84), Boots (3.82), Hat (3.80), Skirt (3.78)

3

Shipping Type Comparison

Standard: $58.46, Express: $60.48 (Avg. Purchase)

4

Discount-Dependent Products

Hat (50%), Sneakers (49.66%), Coat (49.07%), Sweater (48.17%), Pants (47.37%)

Customer Segmentation & Loyalty

Customer Segments

- ■ Loyal
- ■ Returning
- ■ New

Subscribers vs. Non-Subscribers

Subscribers: 1,053 customers, Avg. Spend: $59.49, Total Revenue: $62,645

Non-Subscribers: 2,847 customers, Avg. Spend: $59.87, Total Revenue: $170,436

Repeat buyers (>5 purchases) are more likely to subscribe (958 Yes vs. 2518 No).
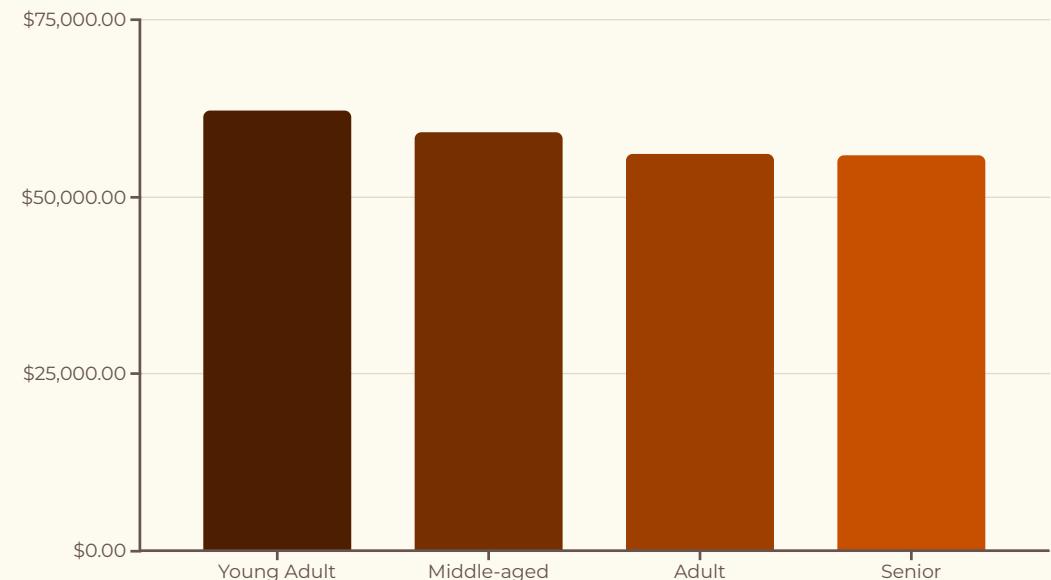
Product Performance & Age Group Revenue

Top Products per Category

- **Accessories:** Jewelry, Sunglasses, Belt
- **Clothing:** Blouse, Pants, Shirt
- **Footwear:** Sandals, Shoes, Sneakers
- **Outerwear:** Jacket, Coat

Revenue by Age Group

Power BI Dashboard: Visualizing Insights

An interactive dashboard provides key metrics and visualizations for data-driven decisions.

3.9K Customers	$59.76 Avg. Purchase	3.75 Avg. Review Rating

Actionable Recommendations

1   Boost Subscriptions

Promote exclusive benefits to increase subscriber base.

2   Customer Loyalty Programs

Reward repeat buyers to foster loyalty and move them into the "Loyal" segment.

3   Review Discount Policy

Balance sales boosts from discounts with margin control.

4   Product Positioning

Highlight top-rated and best-selling products in marketing campaigns.

5   Targeted Marketing

Focus efforts on high-revenue age groups and express-shipping users for maximum impact.