Customer Behavior Analysis

Using SQL, Python & Power BI

Project Overview

Analyzing customer shopping behavior with transactional retail data.

Python

Data cleaning, preparation, EDA.

MySQL Workbench

Business-driven analytical queries.

Power BI

Interactive dashboards, insights.

Business Problem Statement

Understanding changing customer shopping behaviour.

Revenue Drivers

Which customer groups drive revenue?

Discount Influence

How do discounts affect spending?

Product Preference

Which products are most preferred?

Loyalty & Subscriptions

Converting returning customers, subscription impact.

Dataset Summary

Transactional records for individual customer purchases.

Dataset Structure

- Rows: 3,900
- Columns: 18

Key Columns

- **Demographics:** Gender, Age_Group, Subscription_Status
- **Purchases:** Purchase_Amount, Discount_Applied, Previous_Purchases
- **Products:** Category, Item_Purchased
- **Shipping:** Shipping_Type
- **Review:** Review_Rating

Python Data Preparation & EDA

Performed in Jupyter Notebook using pandas.

01

Data Loading

Loaded CSV, inspected structure.

02

Missing Value Handling

Filled missing Review_Rating with median of respective category

03

Data Cleaning

Standardized names, converted types, cleaned columns.

04

Feature Engineering

Created age_group, purchase_frequency_days, customer_segment.

SQL Analysis: Revenue & Discounts

Extracting patterns in customer spending and discount behavior.

Q1: Gender Revenue

Male: €157,890 (~68%)

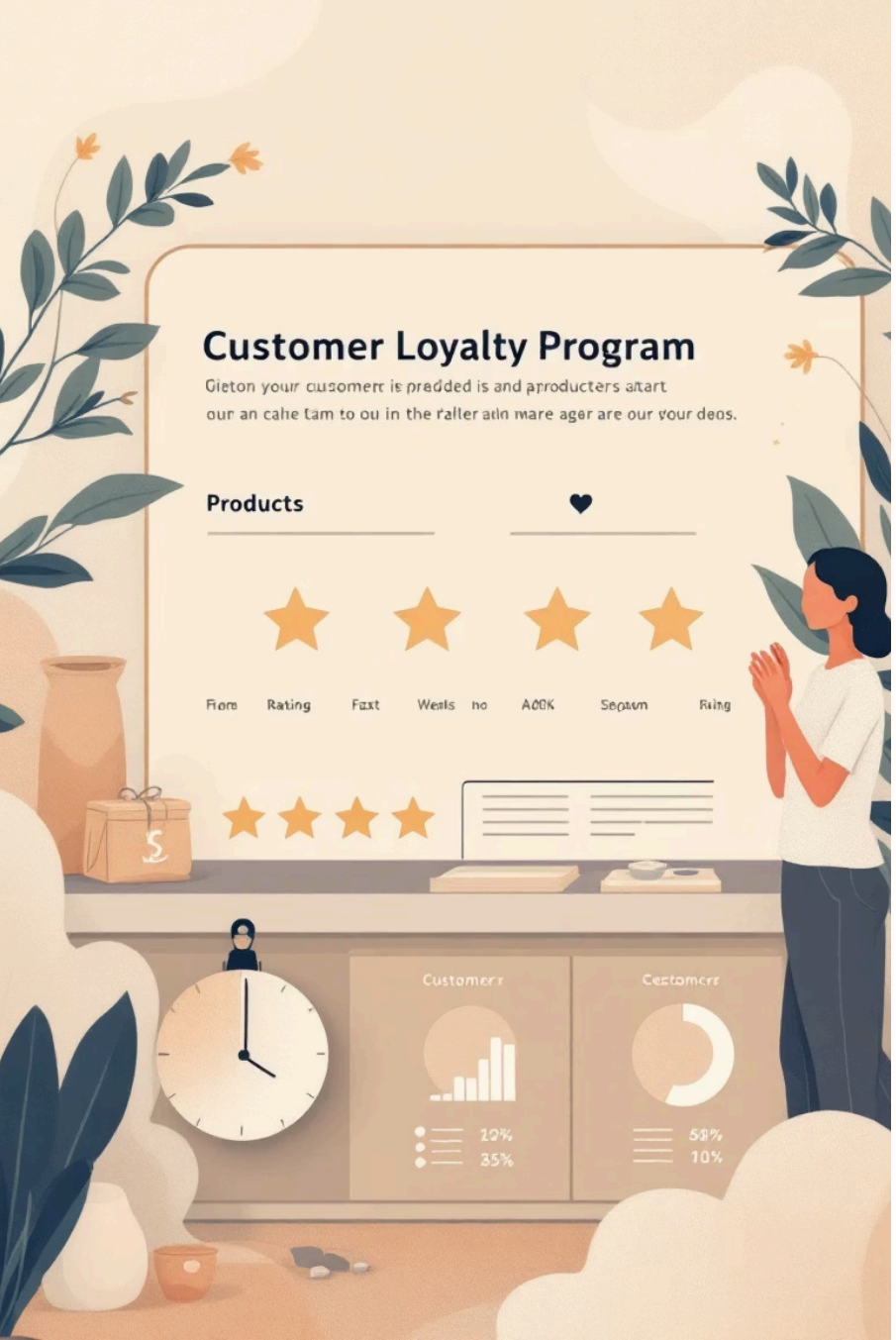
Female: € 75,191 (~32%)

Insight: Male shoppers are the dominant revenue driver.

Q2: Discount & Spend

Customers using discounts still spending above average.

Insight: High-value, promotion-responsive buyers. Target for personalized offers.



SQL Analysis: Products & Loyalty

Identifying product preferences and customer lifecycle stages.

Q3: Top 5 Rated Products

- Gloves: 3.9
- Sandals, Boots, Hat, Skirt: 3.8

Insight: High satisfaction, leverage for marketing.

Q7: Customer Segmentation

- Loyal: 3,116
- Returning: 701
- New: 83

Insight: Excellent retention, but struggling with new customer acquisition.

SQL Analysis: Shipping & Subscriptions

Evaluating impact of shipping types and subscription programs.

Q4: Avg Purchase by Shipping

- Express: €60.48
- Standard: €58.46

Insight: Express orders have ~3.5% higher basket value.

Q5: Subscriber Spend

- Subscribers: 1,053 customers, €62,645 revenue
- Non-Subscribers: 2,847 customers, €170,436 revenue

Insight: Average spend is similar, but subscriber base is smaller. Opportunity to grow subscribers.

Key Findings

Summary of insights from customer behavior analysis.

Demographics

Male, Young Adult, Middle-aged customers drive most revenue.

Subscriptions

27% subscribers, but large untapped opportunity to convert loyal non-subscribers.

Customer Base

80% Loyal, 2% New. Strong retention, weak acquisition.

Products

Clothing is top category. Hat, Sneakers, Coat highly discount-dependent.

Business Recommendations

Actionable strategies for growth.

Strengthen Subscriptions

Target loyal buyers with exclusive offers to increase revenue per customer.

Boost New Acquisition

Implement social media, referral programs, first-purchase discounts.

Optimize Discounts

Reduce reliance on discounts for price-sensitive products; bundle items.

Promote High-Rated Products

Leverage top-rated items in campaigns to drive conversions.

Personalized Marketing

Tailor ads for Young Adults and Middle-aged customers.