Customer Shopping Behavior Analysis

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

Project Overview

1

Data Analysis

Analyzed 3,900 customer purchases.

2

Key Insights

Uncovered spending patterns and product preferences.

3

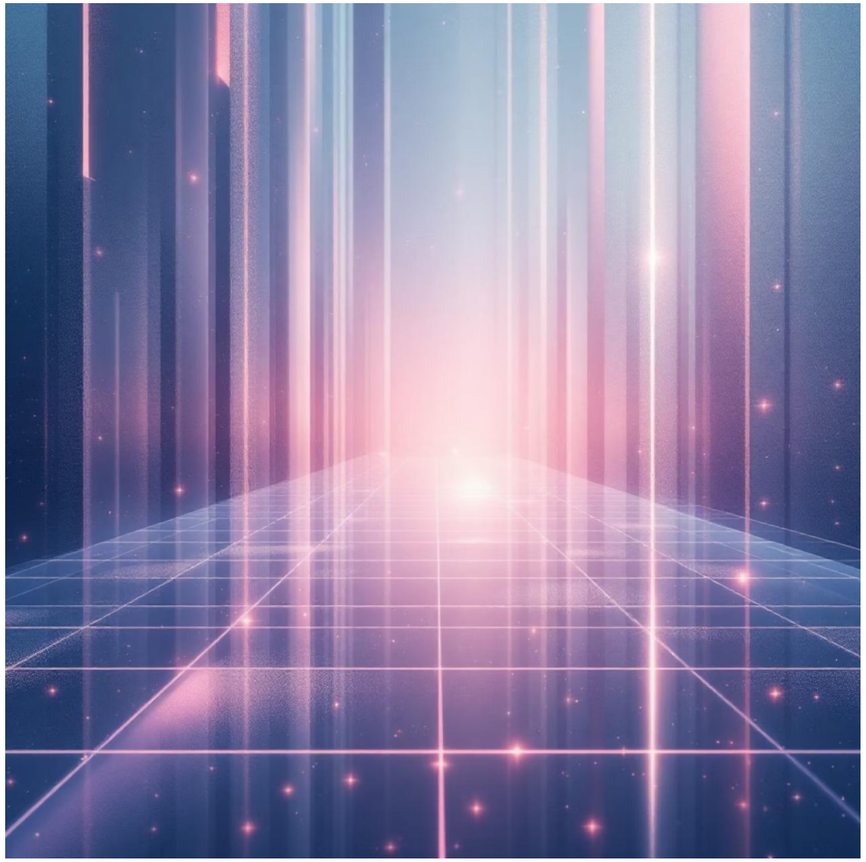
Strategic Decisions

Aimed at guiding business strategies.

Dataset Summary

Our dataset comprises 3,900 rows and 18 columns, detailing customer demographics, purchase specifics, and shopping behaviors.

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



Customer Demographics

Age, Gender, Location, Subscription Status

Purchase Details

Item, Category, Amount, Season, Size, Color

Shopping Behavior

Discount, Promo Code, Previous Purchases, Frequency, Review, Shipping

Exploratory Data Analysis (EDA) in Python

01

Data Loading & Exploration

Imported dataset with pandas, checked structure and summary statistics.

02

Missing Data Handling

Imputed missing 'Review Rating' values using median per product category.

03

Column Standardization

Renamed columns to 'snake_case' for readability.

04

Feature Engineering

Created 'age_group' and 'purchase_frequency_days' columns.

05

Database Integration

Loaded cleaned data into PostgreSQL for SQL analysis.

Data Analysis using SQL: Key Business Insights

Revenue by Gender

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

Top 5 Products by Rating

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

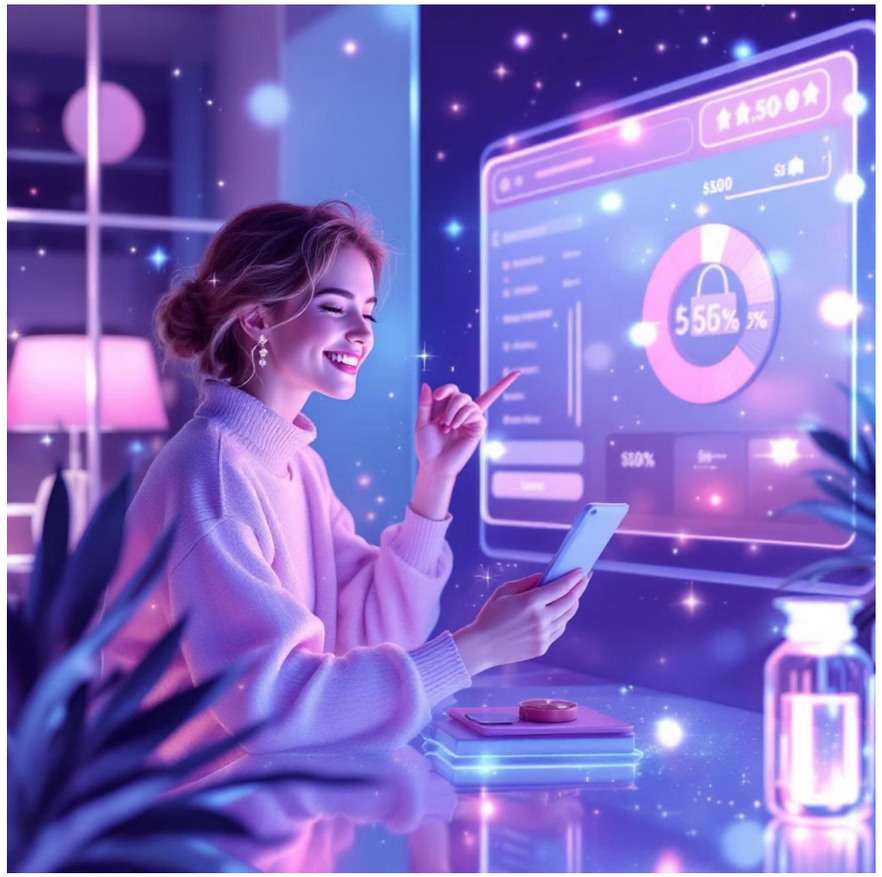
Shipping Type Comparison

Standard: $58.46, Express: $60.48

SQL Insights: Discounts & Subscriptions

High-Spending Discount Users

Identified 839 customers who used discounts but spent above average.



Subscribers vs. Non-Subscribers

Subscribers (1053): Avg Spend $59.49, Total Revenue $62,645

Non-Subscribers (2847): Avg Spend $59.87, Total Revenue $170,436

SQL Insights: Product & Customer Segmentation

1

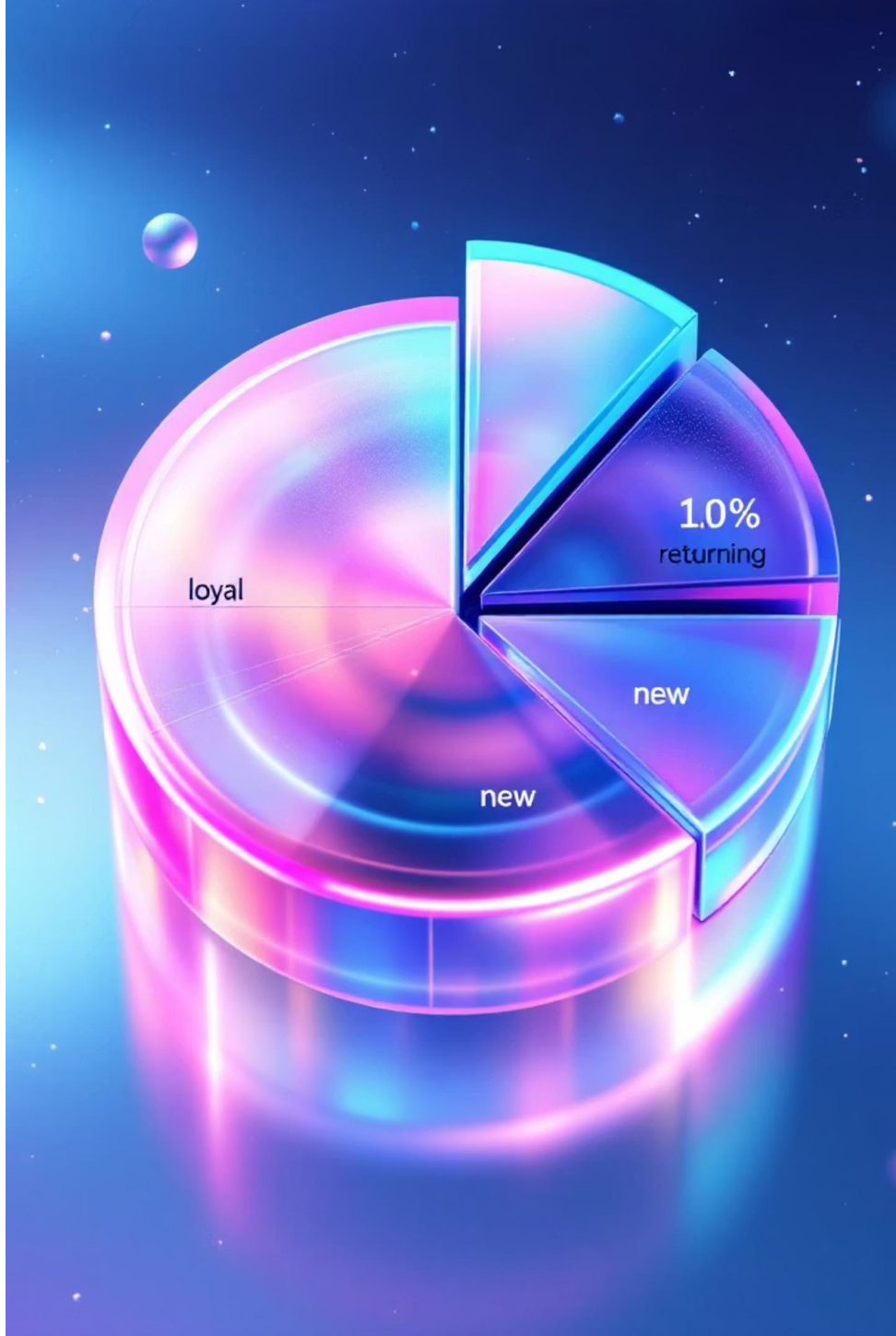
Discount-Dependent Products

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

2

Customer Segmentation

Loyal (3116), Returning (701), New (83)



SQL Insights: Top Products & Age Groups

Top 3 Products per Category

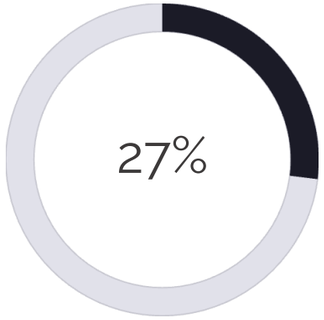
- **Accessories:** Jewelry, Sunglasses, Belt
- **Clothing:** Blouse, Pants, Shirt
- **Footwear:** Sandals, Shoes, Sneakers
- **Outerwear:** Jacket, Coat

Revenue by Age Group

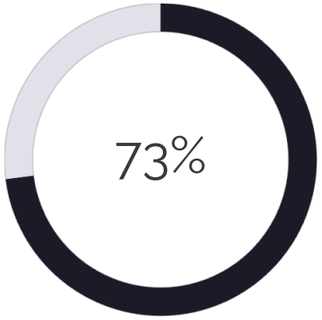
- **Young Adult:** $62,143
- **Middle-aged:** $59,197
- **Adult:** $55,978
- **Senior:** $55,763

Dashboard in Power BI

An interactive dashboard in Power BI visually presents key insights, including customer count (3.9K), average purchase amount ($59.76), and average review rating (3.75).



Subscribers



Non-Subscribers

Business Recommendations

Boost Subscriptions

Promote exclusive benefits to increase subscriber base.

Customer Loyalty Programs

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

Review Discount Policy

Balance sales boosts with margin control for sustainable growth.

Product Positioning

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

Targeted Marketing

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