

Customer Shopping Behavior Analysis

Uncovering insights into spending patterns, product preferences, and subscription behavior to guide strategic business decisions.



PROJECT INITIATION

Project Overview

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

Dataset Summary: Key Features

Data Volume

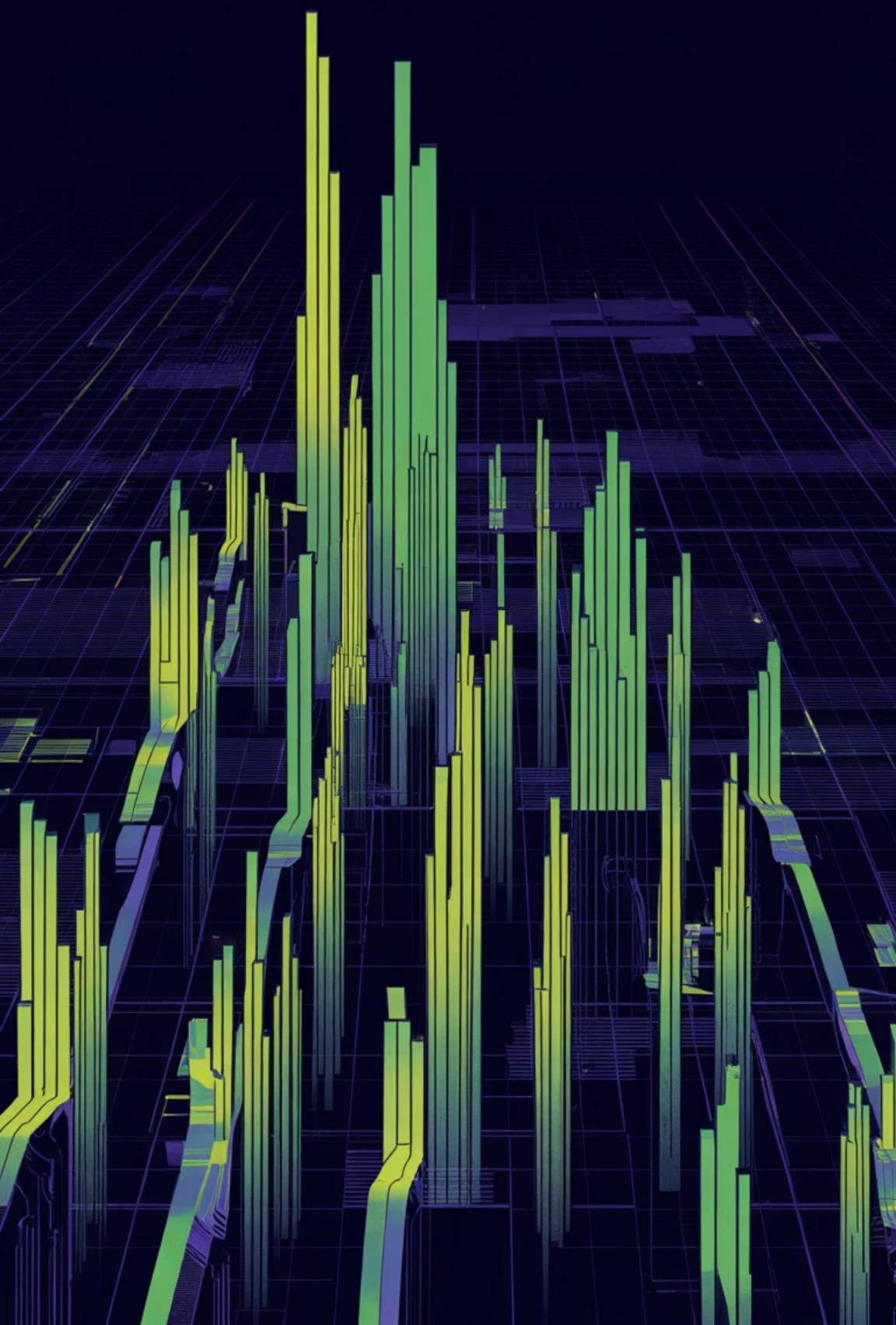
3,900 Rows and **18** Columns of transactional data.

Key Features

Includes Customer demographics, Purchase details (Category, Items), and Shopping behavior (Discount, Frequency).

Missing Data

Only **37** missing values, all in the `review_rating` column.



DATA PREPARATION

Exploratory Data Analysis & Cleaning

1 Data Loading & Exploration

Imported dataset using pandas; used `df.info()` and `.describe()` for initial structure and summary statistics.

2 Missing Data Handling

Imputed missing values in Review Rating using the median rating of each product category.

3 Standardization & Consistency

Renamed columns to snake case. Dropped redundant `promo_code_used` column.

	Customer ID	Age	Gender	Item Purchased	Category	Purchase Amount (USD)	Location	Size	Color	Season	Review Rating	Subscription Status	Shipping Type	Discount Applied	Promo Code Used	Previous Purchases	Payment Method
count	3900.000000	3900.000000	3900	3900	3900	3900.000000	3900	3900	3900	3900	3863.000000	3900	3900	3900	3900	3900.000000	3900
unique	NaN	NaN	2	25	4	NaN	50	4	25	4	NaN	2	6	2	2	NaN	6
top	NaN	NaN	Male	Blouse	Clothing	NaN	Montana	M	Olive	Spring	NaN	No	Free Shipping	No	No	NaN	PayPal
freq	NaN	NaN	2652	171	1737	NaN	96	1755	177	999	NaN	2847	675	2223	2223	NaN	677
mean	1950.500000	44.068462	NaN	NaN	NaN	59.764359	NaN	NaN	NaN	NaN	3.750065	NaN	NaN	NaN	NaN	25.351538	NaN
std	1125.977353	15.207589	NaN	NaN	NaN	23.685392	NaN	NaN	NaN	NaN	0.716983	NaN	NaN	NaN	NaN	14.447125	NaN
min	1.000000	18.000000	NaN	NaN	NaN	20.000000	NaN	NaN	NaN	NaN	2.500000	NaN	NaN	NaN	NaN	1.000000	NaN
25%	975.750000	31.000000	NaN	NaN	NaN	39.000000	NaN	NaN	NaN	NaN	3.100000	NaN	NaN	NaN	NaN	13.000000	NaN
50%	1950.500000	44.000000	NaN	NaN	NaN	60.000000	NaN	NaN	NaN	NaN	3.800000	NaN	NaN	NaN	NaN	25.000000	NaN
75%	2925.250000	57.000000	NaN	NaN	NaN	81.000000	NaN	NaN	NaN	NaN	4.400000	NaN	NaN	NaN	NaN	38.000000	NaN
max	3900.000000	70.000000	NaN	NaN	NaN	100.000000	NaN	NaN	NaN	NaN	5.000000	NaN	NaN	NaN	NaN	50.000000	NaN

A vertical decorative column on the left side of the slide features abstract, overlapping geometric shapes in shades of yellow, green, and black. The shapes include various polygons and curved forms, creating a complex, layered pattern.

Feature Engineering for Deeper Insights

Age Group Creation

Created the `age_group` column by binning customer ages for demographic analysis.

Purchase Frequency

Calculated `purchase_frequency_days` from purchase data to analyze buying habits.

Discount Verification

Verified consistency between `discount_applied` and `promo_code_used`.



CORE ANALYSIS

Key Questions Answered with SQL

A series of 10 targeted questions were answered using SQL queries to extract specific business intelligence from the dataset.

Revenue & Spend

Total revenue by gender, average spend comparison (shipping/subscription), and age group contribution.

Product & Discounts

Top 5 products by rating, top 3 products per category, and products with the highest discount percentage.

Customer Segmentation

Segmenting customers (New, Returning, Loyal) and analyzing repeat buyer subscription likelihood.

Subscriber Value & Loyalty



Subscriber Spend

Subscribed customers spend significantly more on average and contribute higher total revenue.



Top Products

Identified the top 5 products with the highest average review ratings for promotion.



Shipping Preference

Compared average purchase amounts between Standard and Express Shipping users.

The analysis confirmed that repeat buyers (more than 5 previous purchases) are also highly likely to subscribe.

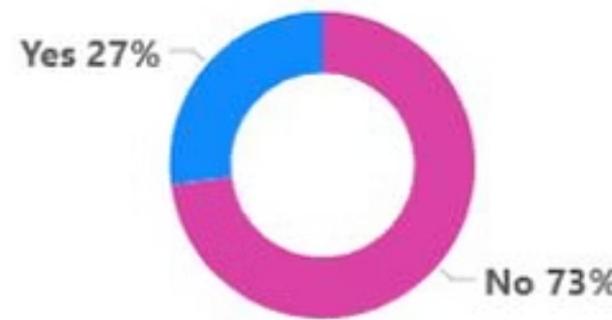
Gender



Category



Subscription Status



Revenue by Category



Sales by Category



VISUAL INSIGHTS

Power BI Dashboard Preview

The final dashboard provides an interactive summary of the key findings, allowing stakeholders to drill down into specific segments and metrics.

Strategic Business Recommendations

1

Boost Subscriptions

Promote exclusive benefits to increase subscriber base and leverage higher average spend.

2

Customer Loyalty

Implement programs to reward repeat buyers and move them into the "Loyal" customer segment.

3

Review Discount Policy

Carefully balance sales boosts from discounts with necessary margin control.



ACTIONABLE STEPS

Targeted Marketing & Product Focus

→ Product Positioning

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

→ Targeted Marketing

Focus efforts on high-revenue age groups and customers who prefer express shipping.

