Customer Shopping Behaviour Analysis

Analyzing customer shopping behaviour and patterns using transactional data from over 40,000 purchases across various products, categories, and age groups. This project focuses on cleaning, organizing, and managing data to derive actionable insights that support strategic business decisions.

DATASET OVERVIEW

Understanding Our Data Foundation

40K

Total Records

Comprehensive purchase transactions analyzed

20

Data Columns

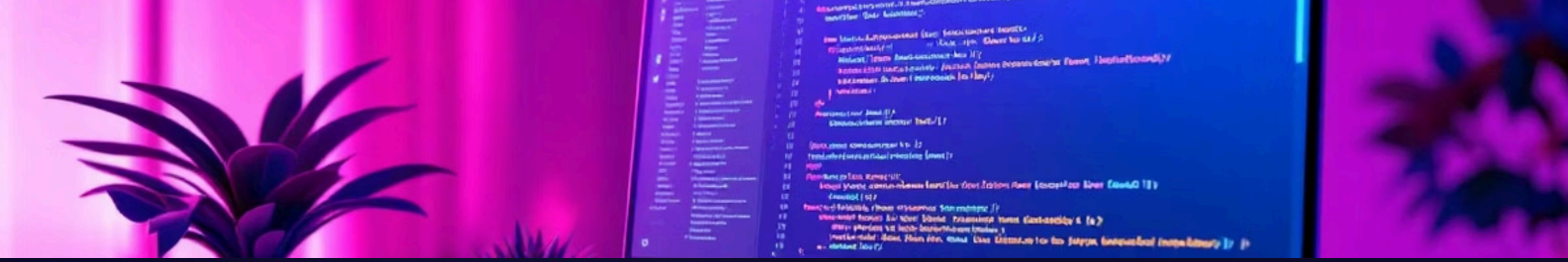
Rich feature set for deep analysis

37

Missing Values

Only in Review Rating column

The dataset encompasses customer demographics (Age, Gender, Location, Subscription Status), purchase details (Item Purchased, Category, Purchase Amount, Season, Size, Colour), and shopping behaviour metrics (Discount Applied, Promo Code Used, Previous Purchases, Frequency of Purchases, Review Rating, Shipping Type).



PYTHON ANALYSIS

Data Cleaning & Preparation Process

01

Data Import

Loaded dataset into pandas using Jupyter Notebook and reviewed sample data with df.head() to understand structure

02

Missing Data Handling

Identified 37 null values in Review Rating column using df.isnull().sum() and replaced with median value

03

Column Standardization

Renamed columns to follow consistent naming conventions for better readability and analysis

04

Feature Engineering

Created new "age_group" column based on numeric age values for segmented reporting and analysis

Advanced Feature Engineering

Purchase Frequency Tracking

Created additional column to maintain days of purchase frequency, enabling time-based customer behaviour analysis. This metric helps identify shopping patterns and customer engagement levels over time.

Data Consistency Optimization

Verified if discount_applied and promo_code_used were redundant fields. After analysis, dropped promo_code_used to streamline the dataset and eliminate duplicate information.

Database Integration

Connected Python script to PostgreSQL and loaded the cleaned DataFrame into the database, enabling powerful SQL-based analysis and queries.

SQL ANALYSIS

Key Business Insights from Data

Revenue by Gender

Calculated total revenue generated by male and female customers to understand gender-based purchasing power

High-Spending Discount Users

Identified customers who used discounts but still spent more than average purchase amount

Top 5 Rated Products

Discovered products with highest review ratings to guide inventory and marketing decisions

Shipping & Subscription Analysis

Shipping Type Comparison

Compared average purchase amount between standard and express shipping methods. Express shipping users showed different spending patterns, indicating premium customer segments.

Subscription Impact

Analyzed average spend and total revenue between subscribed and non-subscribed customers. Subscribers demonstrated higher lifetime value and purchase frequency.

Top Discount Recipients

Identified the top 5 customers who received the highest discounts, revealing opportunities for targeted retention strategies.

Customer Segmentation Insights

Three-Tier Classification

Customers were classified into three distinct segments based on purchase history:

- **New Customers:** First-time buyers exploring the brand
- **Returning Customers:** 2-5 previous purchases showing interest
- **Loyal Customers:** More than 5 purchases demonstrating brand commitment

Repeat Buyer Correlation

Analysis revealed that customers with more than 5 previous purchases are significantly more likely to subscribe, indicating strong brand loyalty and engagement.

Product & Age Group Performance

Top 3 Per Category

Identified the top 3 most purchased products within each category, revealing customer preferences and bestsellers

Age Group Revenue

Analyzed revenue contribution by age group to understand demographic purchasing power and target marketing efforts

These insights enable targeted product positioning and age-specific marketing campaigns that resonate with each demographic segment's unique preferences and spending patterns.

VISUALIZATION

Power BI Dashboard

Built comprehensive dashboard using Power BI to present insights visually. The interactive dashboard enables stakeholders to explore customer behaviour patterns, revenue trends, and product performance metrics in real-time. Visual storytelling transforms complex data into actionable intelligence for strategic decision-making.

Strategic Business Recommendations

Boost Subscriptions

Promote exclusive benefits and perks for subscribers to increase conversion rates and customer lifetime value

Loyalty Programs

Reward repeat buyers to move them into the "Loyal" segment and strengthen brand commitment

Review Discount Policy

Balance sales boosts with margin control to optimize profitability while maintaining competitiveness

Product Positioning

Highlight top-rated and best-selling products in marketing campaigns to drive conversion

Targeted Marketing

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