End-to-End Customer Purchase & Behaviour Analytics (E-commerce)

By Aayush Kumbhar, Data Analyst

Date: 24/12/2025 – 31/12/2025

Executive Summary: From Raw Data to Actionable Insights

Workflow

Transformed raw customer shopping data into meaningful business insights.

Dataset

~3,900 records, 19 columns: demographics, purchasing, transactions.

Tools

Python (Pandas) for cleaning, MySQL for analysis, Power BI for visualization.

Business Problem: Hidden Insights

Raw customer shopping data makes it challenging to understand customer behavior, spending, and influencing factors.

Customer Segmentation

Difficulty in understanding which customer types drive sales.

Behavioral Differences

Hard to see how purchasing behavior varies across transactions.

Trend Identification

Inability to identify basic trends for decision-making.

Objectives of the Analysis

To study customer shopping behavior and turn raw data into useful insights for business decisions.

01

Understand Purchasing Behavior

Overall customer purchasing patterns.

02

Identify Key Metrics

Important business performance indicators.

03

Analyze Segments

Customer behavior across categories and segments.

04

Create Interactive Dashboard

Easy-to-understand visualization for tracking performance.

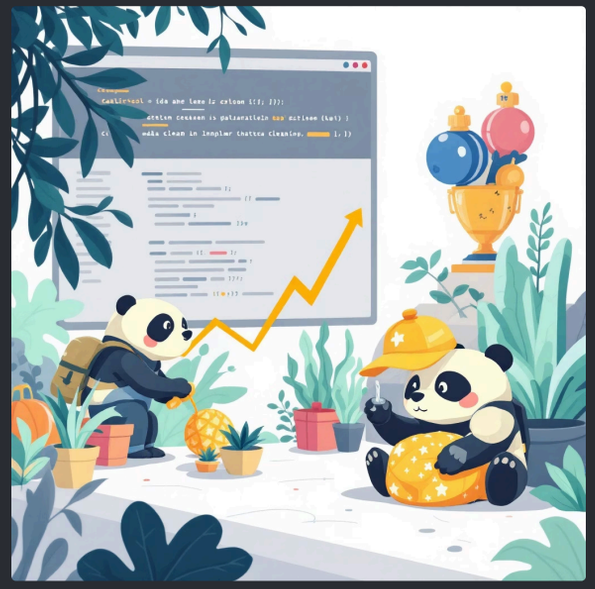
Data Overview: Customer Shopping Dataset

Excel-based dataset with ~3,900 rows and 19 columns after preprocessing.

- Customer demographics & order purchase behavior.
- Subscription status & order locations (US).
- Seasonal patterns & time gap between orders.
- Number of previous purchases (new vs. returning customers).
- Product review ratings for satisfaction insights.

Data Cleaning & Feature Engineering (Python – Pandas)

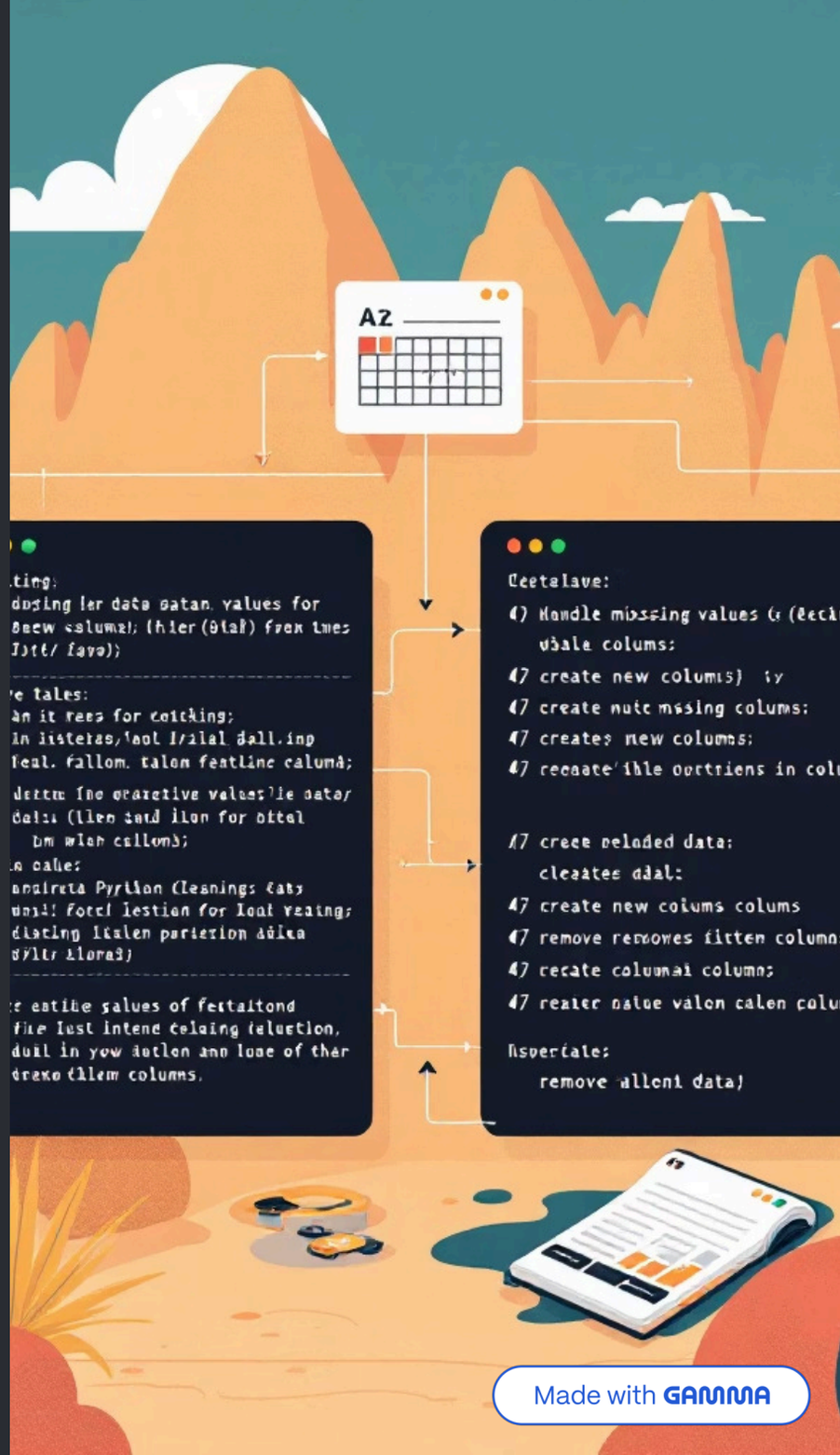
Initial inspection found 37 missing values in 'review_rating', handled by replacing with median per product category.

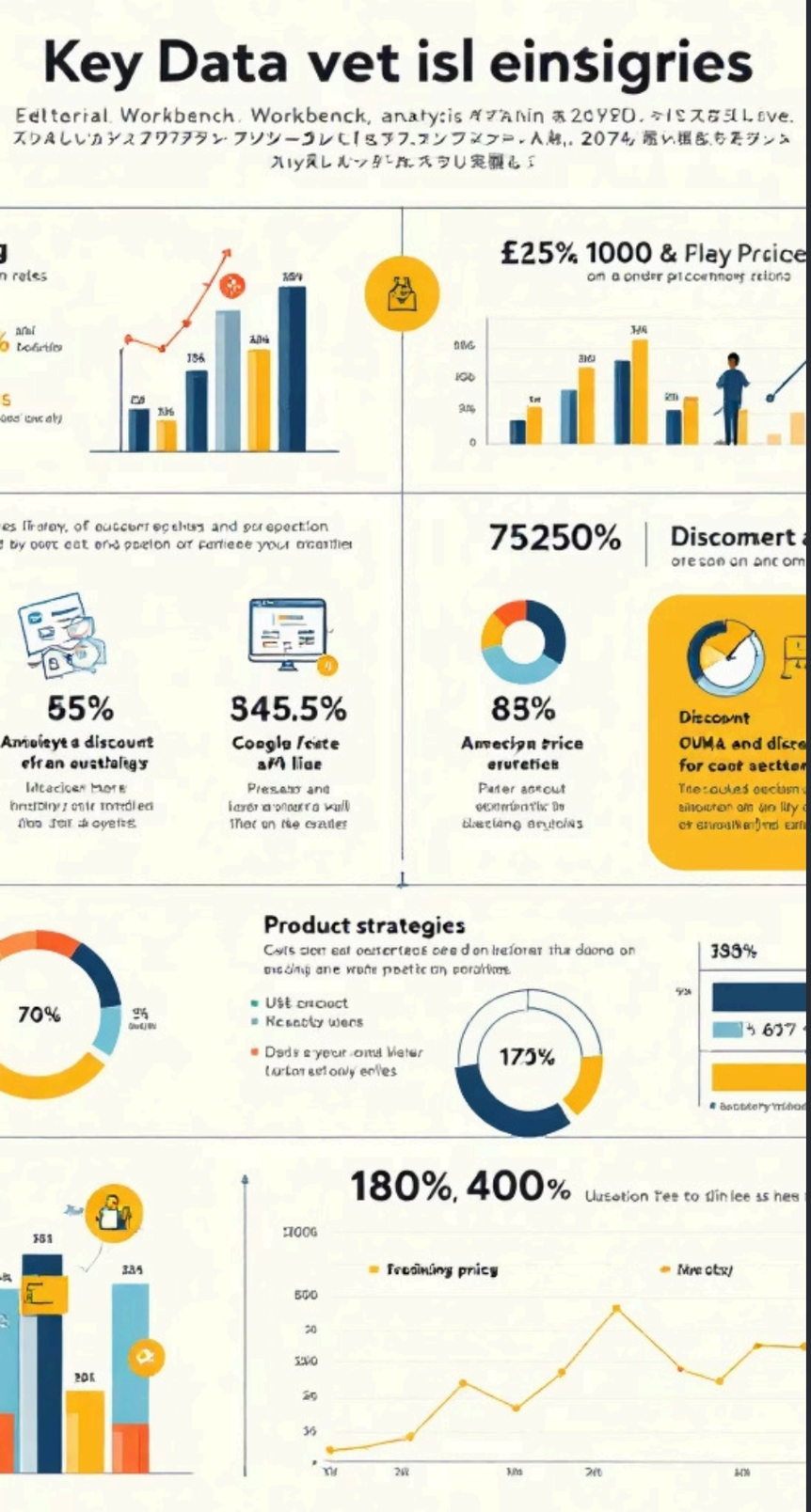


Additional Columns Created:

- Age Group: Young Adult, Adult, Middle Age, Senior.
- Purchase Frequency (Days): Converted values like "weekly" to numerical days.

Removed redundant 'promo_code_used' as it duplicated 'discount_applied'.





SQL Analysis & Aggregations

MySQL Workbench used for structured data analysis, segmentation, and comparative analysis.

Customer Spending

Revenue by gender, AOV for subscribers, age group contribution, high-value customer locations.

Subscription & Retention

Spend comparison, subscriber percentage, gender distribution, repeat buyer likelihood, retention rate.

Discount & Pricing

Effectiveness, revenue comparison, product-specific discounts, seasonal performance.

Product Performance

Top-rated products, most purchased, high frequency/low spend categories, segment preferences.

Key Performance Indicators (KPIs)

Selected KPIs provide a clear view of revenue, customer value, retention, and satisfaction.

3,900

Total Customers

Unique customers analyzed.

$233K

Total Revenue

Generated from all purchases.

$59.76

Average Order Value

Average spent per order.

97.87%

Repeat Customer Rate

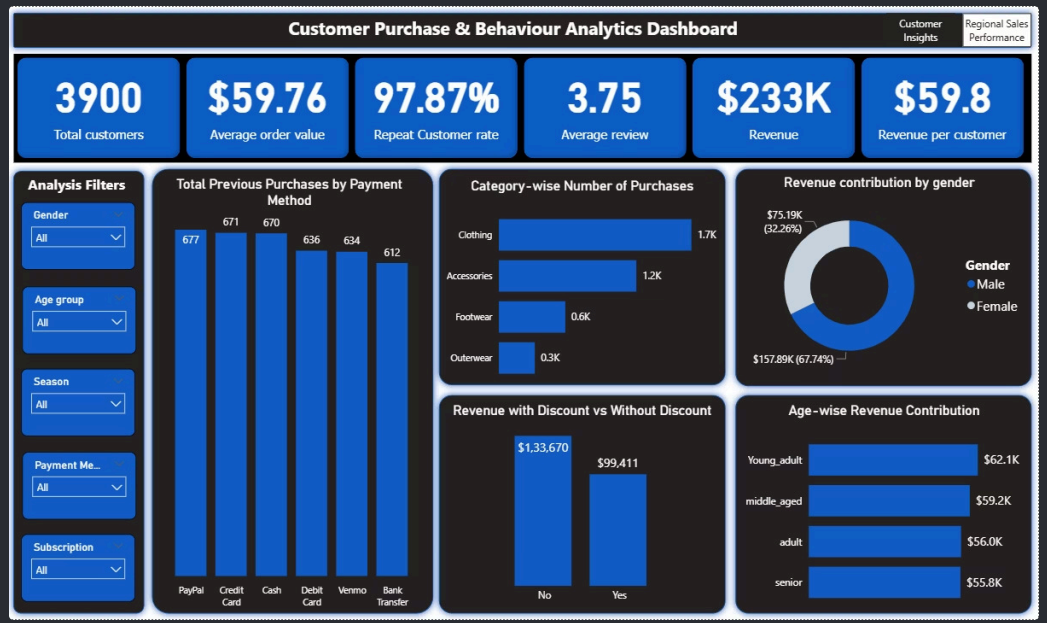
Customers with more than one purchase.

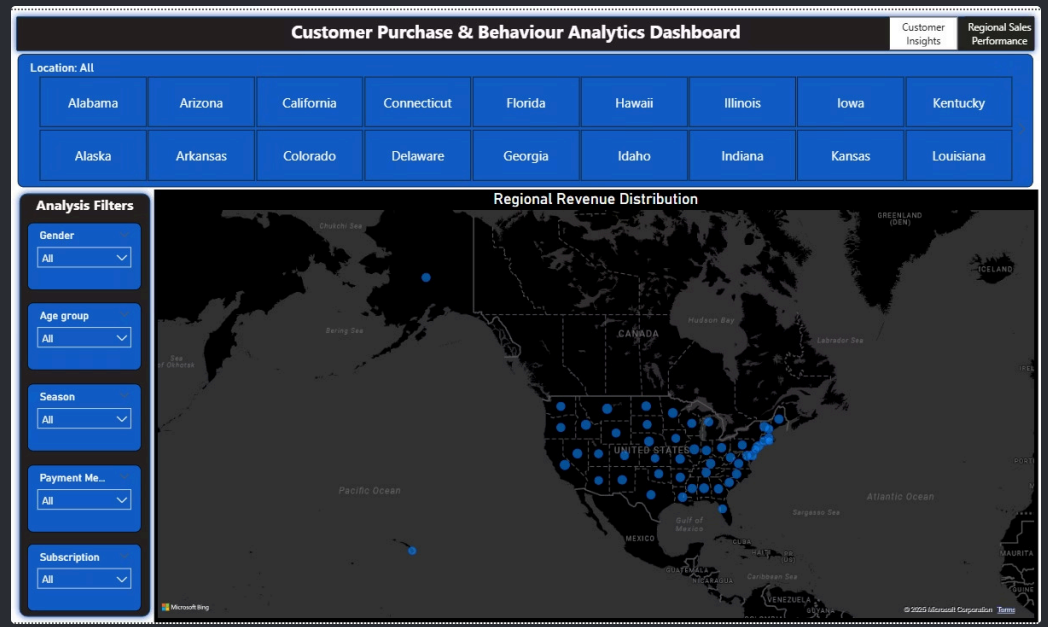
3.75

Average Review Rating

Overall customer satisfaction.

Dashboard Overview: Power BI Visualizations





Interactive dashboard with "Customer Insights" and "Regional Sales Performance" views.

- Key KPI section at the top for quick tracking.
- Visuals for category purchases, revenue by gender/age, discount impact, payment methods.
- Regional view with US map and state-level filters.
- Interactive filters for gender, age, season, payment, subscription, location.

Key Insights & Recommendations

1

Gender Revenue Gap

Male customers generate double the revenue. Targeted campaigns for female customers.

2

Moderate Ratings

Average product rating below 4. Monitor reviews, improve low-rated products.

3

Underperforming Subscription

No significant spend difference for subscribers. Redesign subscription model for value.

4

Seasonal Lows

Purchases lowest in summer. Run targeted promotions during summer.

5

Low-Revenue Locations

Certain states show lower average revenue. Location-based promotions (e.g., free delivery).