



Exploratory Data Analysis: OmniMart Retailers

Uncovering Actionable Insights for Strategic Business Growth

Team: SIGMA BOYS • Date: 2025-09-14

- JOY TALUKDAR — Lead Analyst
- DEBMALYA DAS — Visualization Specialist
- DEVKMAR PAL — Data Engineer
- SRINJAN BHAKAT — Report Writer

The EDA Mission Briefing

Objective

Analyze transactional data to identify opportunities for enhancing Customer Experience (CX), optimizing product portfolio, and improving operations.

Data Source

Internal retail_data.csv
transaction logs.

Key Areas

- Customer Behavior & Segmentation
- Product Performance
- Operational Efficiency

Data Overview & Preprocessing

Our initial dataset contained 302,010 records across 30 variables, spanning customer demographics, purchase specifics, product information, feedback, and logistics details.

01

Initial Assessment

Identified missing values primarily in **Amount/Total_Amount/Ratings** and inconsistent date formats.

02

Cleaning Highlights

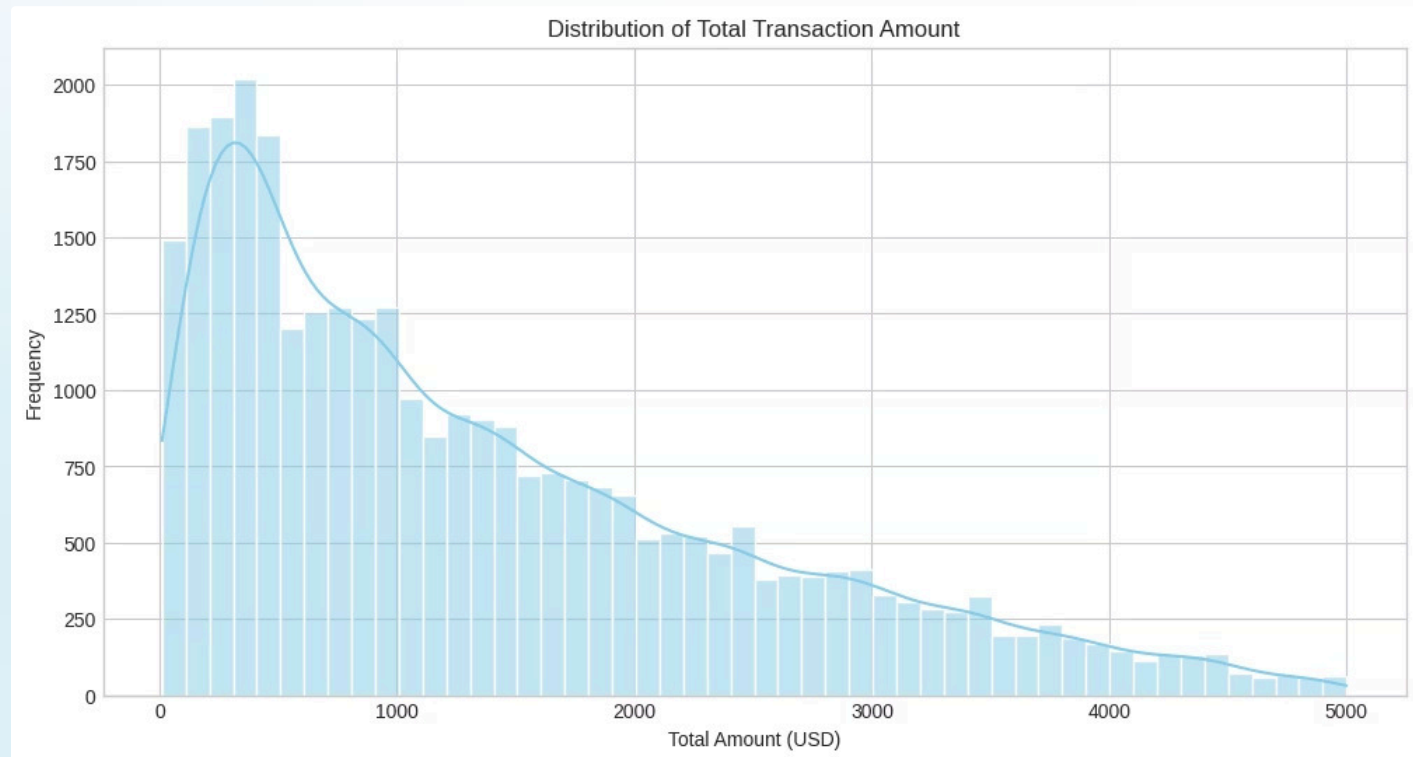
Rows with critical missing data were dropped. Date fields were standardized, and monetary/rating fields were converted to appropriate numeric types.

This rigorous cleaning ensures data integrity for reliable analysis and actionable insights.

Key Finding 1: Customer Demographics

OmniMart serves a diverse customer base. Analysis of total transaction amount distribution reveals varying spending patterns, including both concentrated groups and high-value outliers.

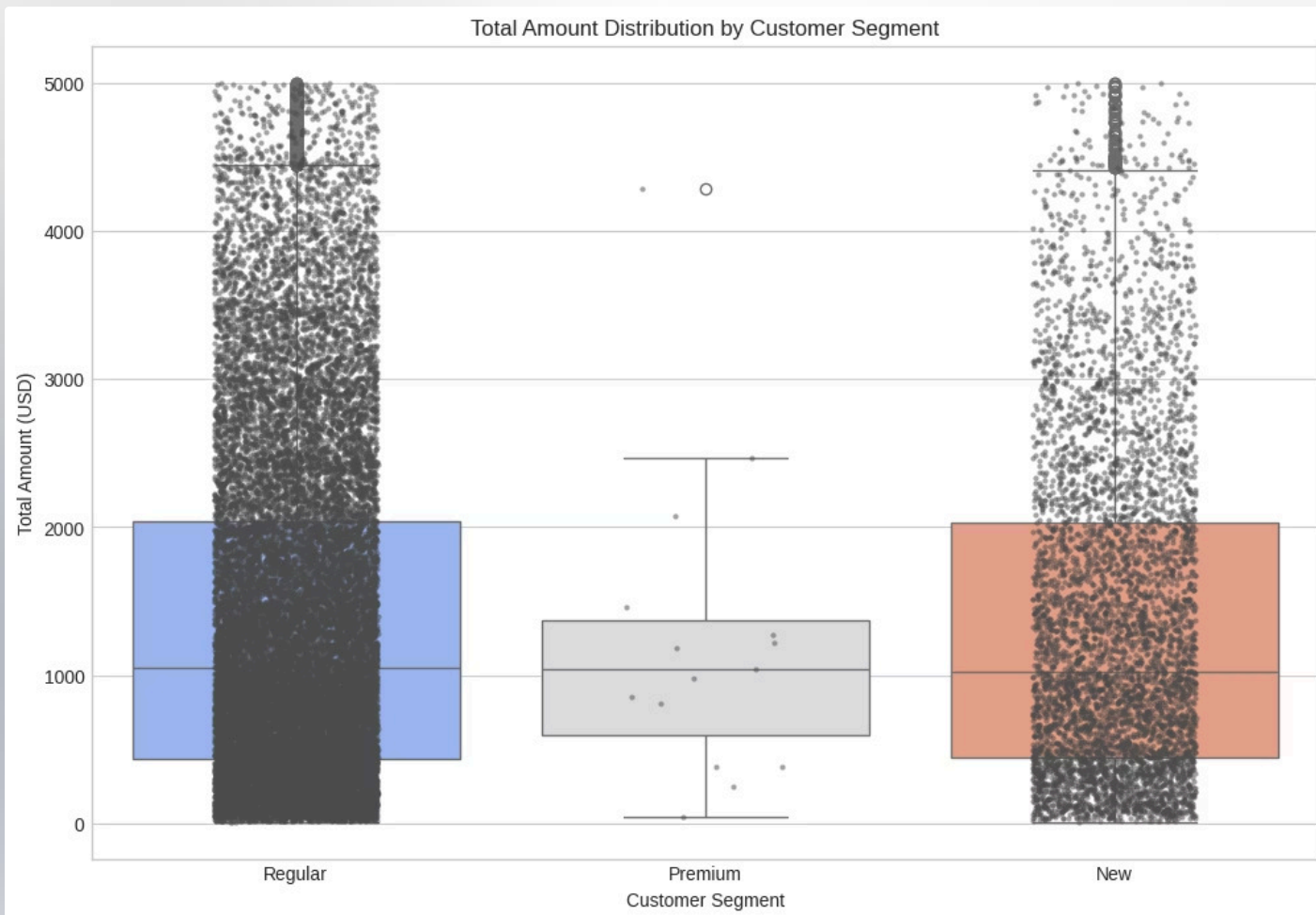
Understanding these demographics is crucial for tailored marketing efforts.



Key Finding 2: Customer Segments & Spending

A deep dive into spending patterns across customer segments highlights that **Premium customers exhibit significantly higher median spending** and contribute more high-value outliers. This segment represents a key opportunity for targeted engagement and loyalty programs.

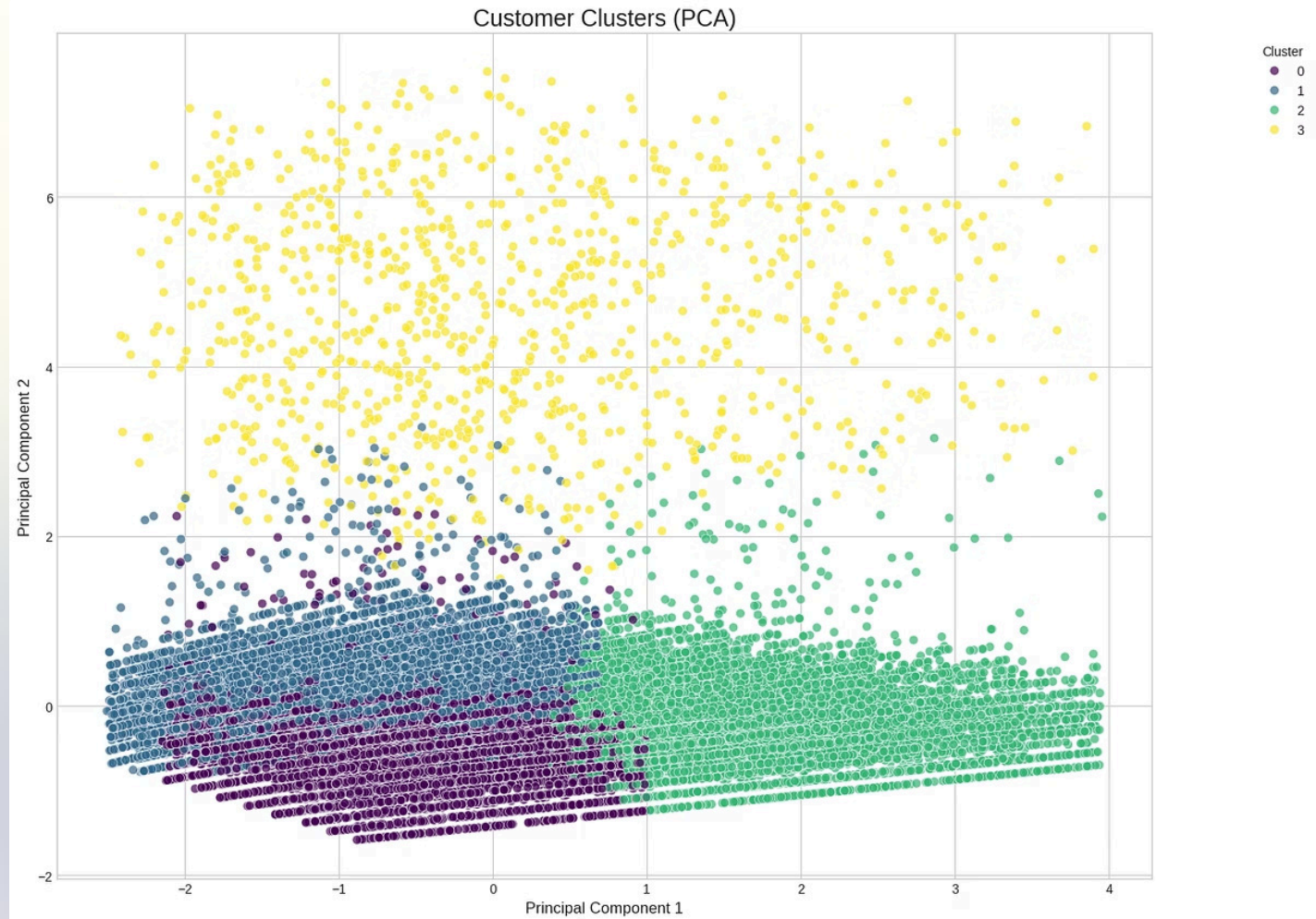
The visualization clearly demonstrates the elevated spending of premium customers.

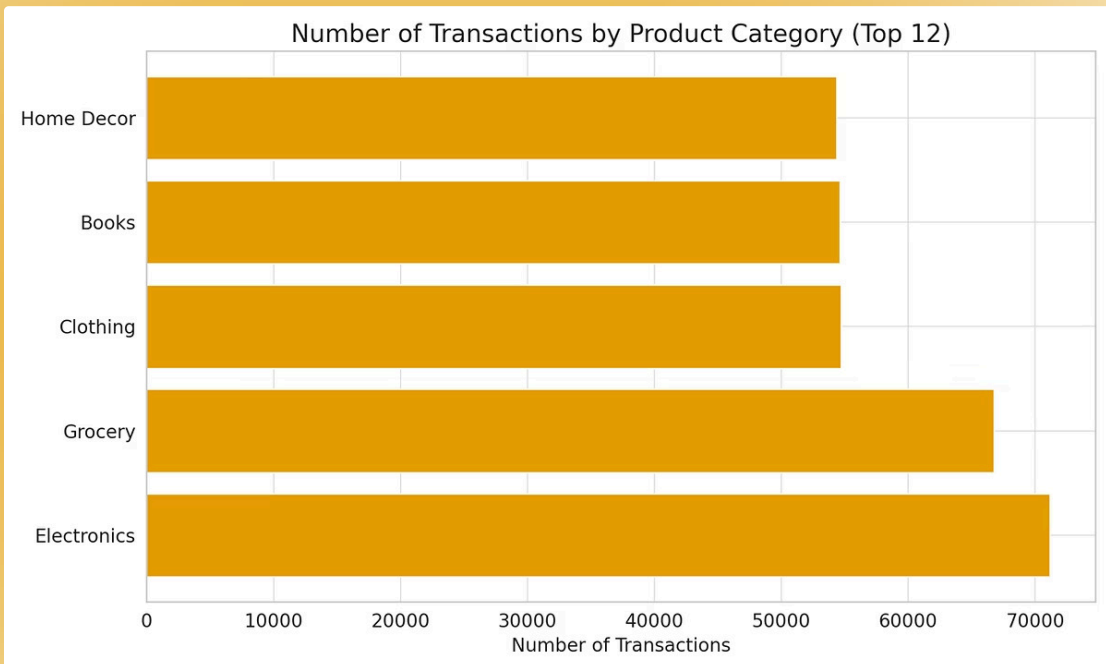


Key Finding 3: Customer Cluster Analysis

Utilizing PCA (Principal Component Analysis) and KMeans clustering, we identified four distinct customer clusters. These clusters are pivotal for developing highly targeted marketing campaigns and personalized customer experiences, moving beyond broad segmentation.

The scatter plot visualizes the separation of these clusters, revealing the underlying patterns in customer behavior.





Key Finding 4: Product Performance

While the **Grocery category** drives significant **transaction volume**, Electronics and Home Decor categories are notable for generating higher revenue per transaction. This suggests different strategic approaches for inventory management and promotional activities across product lines.

The transaction count per product category highlights Grocery's popularity.

Key Findings 5 & 6: Ratings & Operational Efficiency

A significant number of **low ratings (1-2 stars)** indicate potential issues with product quality or customer service.



Additionally, our analysis points to fulfillment bottlenecks, particularly with high demand for expedited shipping, impacting overall operational efficiency.



Actionable Strategic Recommendations

1 Optimize Fulfillment Pipeline

Conduct an audit to reduce Processing/Pending delays, directly improving customer satisfaction and order turnaround times.

2 Refine Customer Segmentation & Marketing

Develop loyalty programs and targeted campaigns for high-potential clusters, such as "Low-Income Premium" customers.

3 Two-Pronged Product Review

Investigate whether low ratings are due to product quality or logistics issues before considering product delisting.

4 Leverage Top Categories

Focus promotional efforts and inventory management on Electronics & Home Decor to capitalize on their higher revenue per transaction.

Next Steps & Thank You

Our exploratory analysis has laid a solid foundation. The next phase will focus on predictive modeling and further segmentation to refine strategies.

→ Next Steps

- Predictive modeling (sales forecasting)
- Churn analysis
- Geographic normalization for regional strategies

→ Deliverables

- analysis.ipynb
- Presentation
- README.md ready for GitHub submission

Thank you! — SIGMA BOYS