E-Commerce User Behavior & Time-Sensitive Recommendation System

Said Tabharit | UMKC Hack-a-Roo 2025 | AI/Data Science Track

Team X

Introduction

The rapid growth of e-commerce platforms has resulted in a deluge of user behavior data. However, despite rich user interactions, cart abandonment rates remain extremely high (over 93%), and most systems still rely on static recommendations. Our goal is to build a **data-driven, time-sensitive recommendation engine** that adapts to users' browsing patterns and purchase timing, increasing engagement and conversion.

This project integrates:

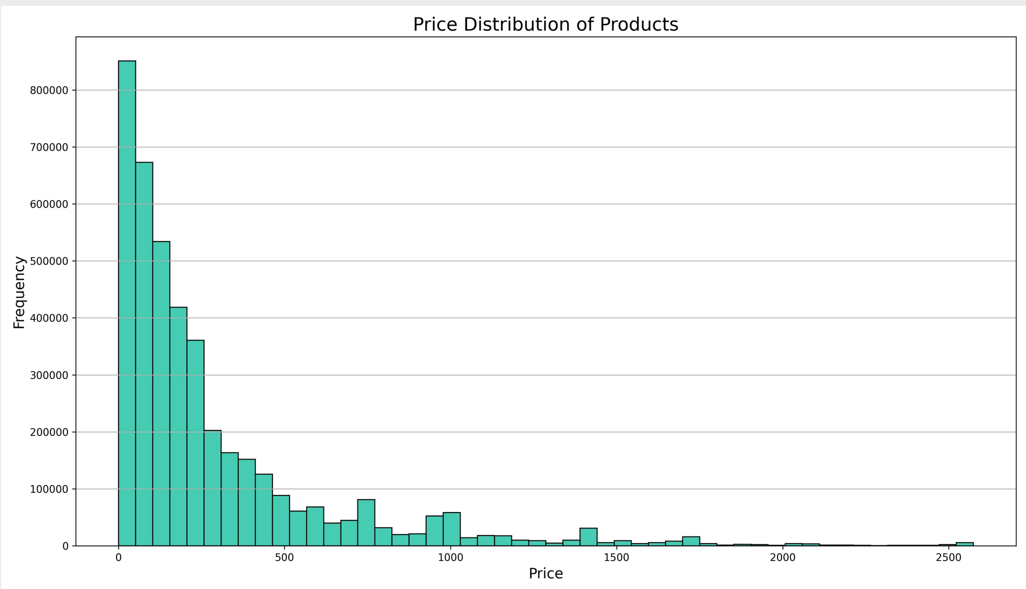
- **4.2 million behavioral events** across 1.39 million users from October 2019.
- **Comprehensive data cleaning, EDA, segmentation, and modeling** pipeline.
- A focus on **personalization, user retention**, and **conversion optimization**.

Data Cleaning

The dataset spans 2019-10-01 to 2019-10-31 and originally contained 4,244,876 records across 9 attributes (event type, product, category, brand, price, etc.). After removing 583 duplicate entries and correcting inconsistent values, we retained 4,244,293 high-quality records for analysis.

Key Stats:

- Price Range: \$0–\$2574 (mean \$290.6, std \$358.5)
- Median Price: \$163.0

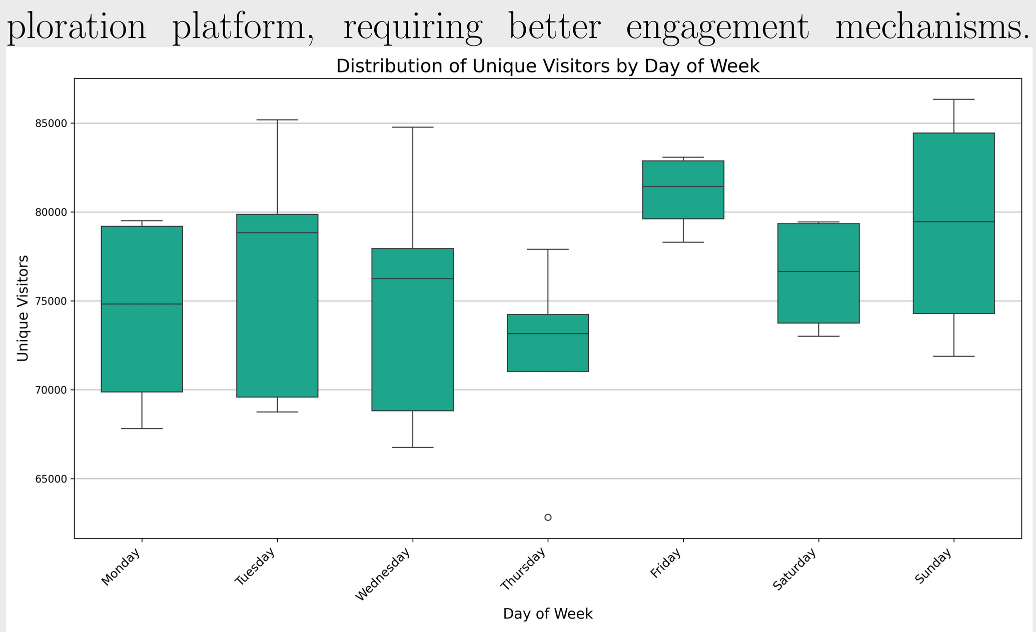


Exploratory Insights

Exploratory analysis reveals:

- **Dominant Event:** Views (4.08M) — users mainly browse rather than purchase.
- **Top Brands:** Samsung, Apple, Xiaomi — electronics dominate.
- **Top Categories:** 43% in *electronics.smartphone*, followed by appliances and computers.
- **Peak Traffic:** Weekends (Sunday 86K visits) and afternoon sessions.

Implication: The site functions primarily as a product exploration platform, requiring better engagement mechanisms.



Methodology

Our pipeline integrates multiple analytical layers to drive personalization:

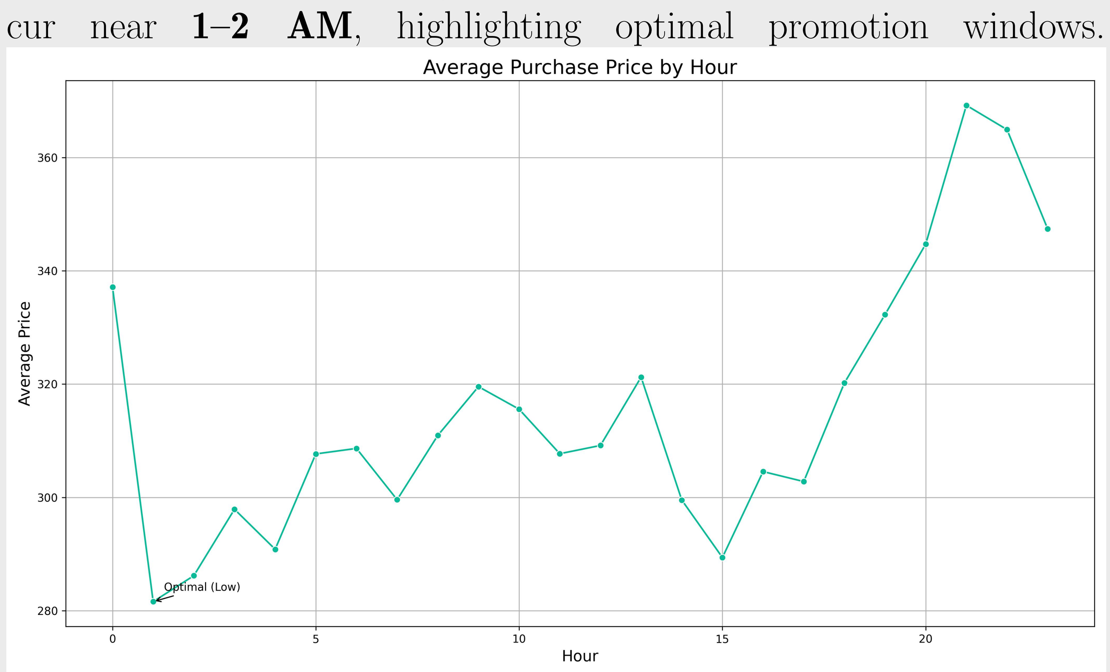
1. Exploratory Data Analysis (EDA) — Identified browsing dominance, repeat customers (537K), and low repeat purchase rate (0.05).

2. Customer Segmentation (K-Means) — Divided users into three key segments:

- **Segment 0: Browsers** — 1.16M users, low purchase rate (0.36%).
- **Segment 1: Buyers** — 197K users, high-value purchases (avg \$988).
- **Segment 2: Loyal Customers** — 29.9K users, 64% purchase rate.

3. Hybrid Recommendation System — Combines collaborative filtering (user-item) with temporal filtering based on browsing hour and recency. Achieved **Precision = 0.123**, demonstrating strong relevance.

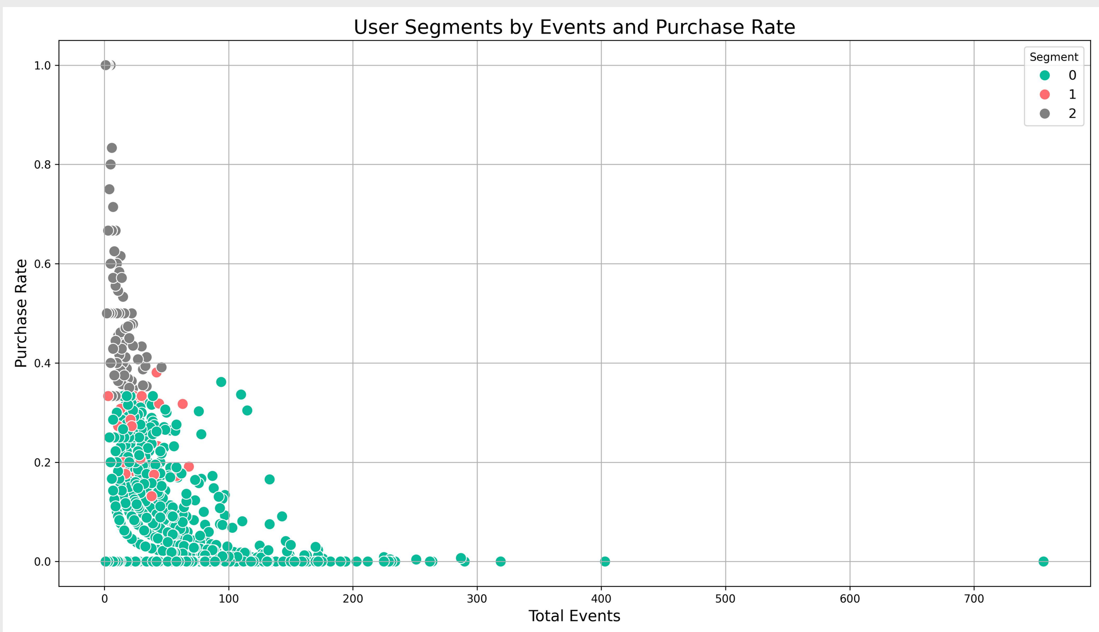
4. Predictive Modeling — Hourly analysis showed price and activity fluctuate significantly; lowest purchase prices occur near **1–2 AM**, highlighting optimal promotion windows.



User Segmentation Visualization

Clusters reveal clear distinctions in engagement and purchase value:

- **Browsers:** frequent page views, low spend.
- **Buyers:** medium events, high-value transactions.
- **Abandoners:** browsing without checkout completion.



Recommendation Engine

Personalized suggestions for individual users are generated based on shared interests and activity recency.

Example Recommendations:

- User 512475445 → [4700419]
- User 512365995 → [1002544, 1002524]

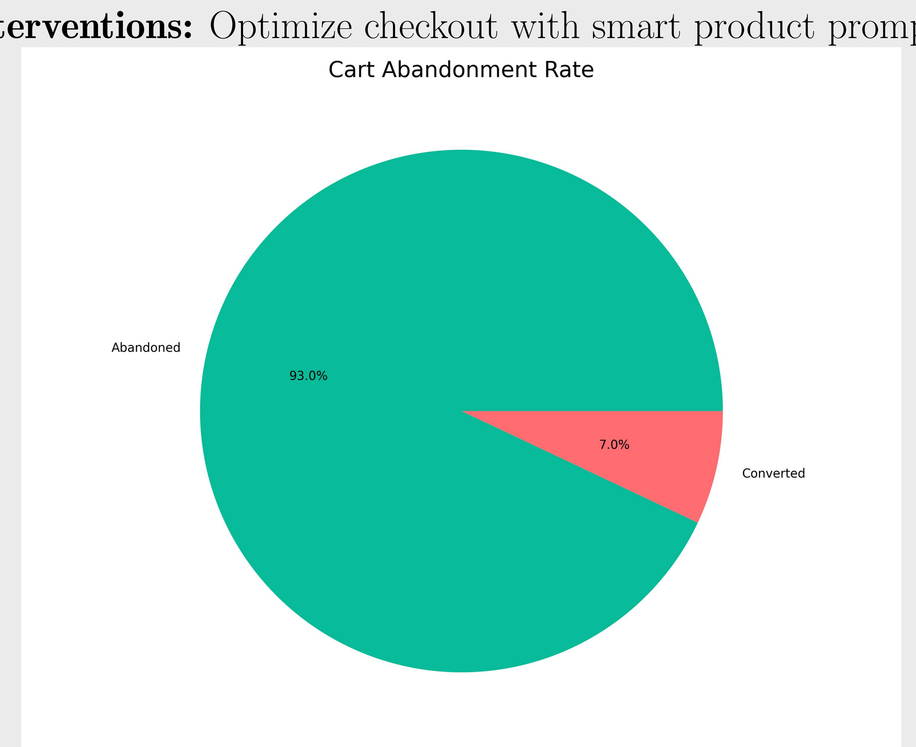
These recommendations are highly aligned with prior user interactions, increasing relevance and conversion likelihood.

User Journey Analysis

Sequential visualization of browsing → cart → purchase stages reveals:

- Average abandonment rate: 93%.
- Long view chains before drop-off.
- Low transition probability from cart to purchase.

Interventions: Optimize checkout with smart product prompts.

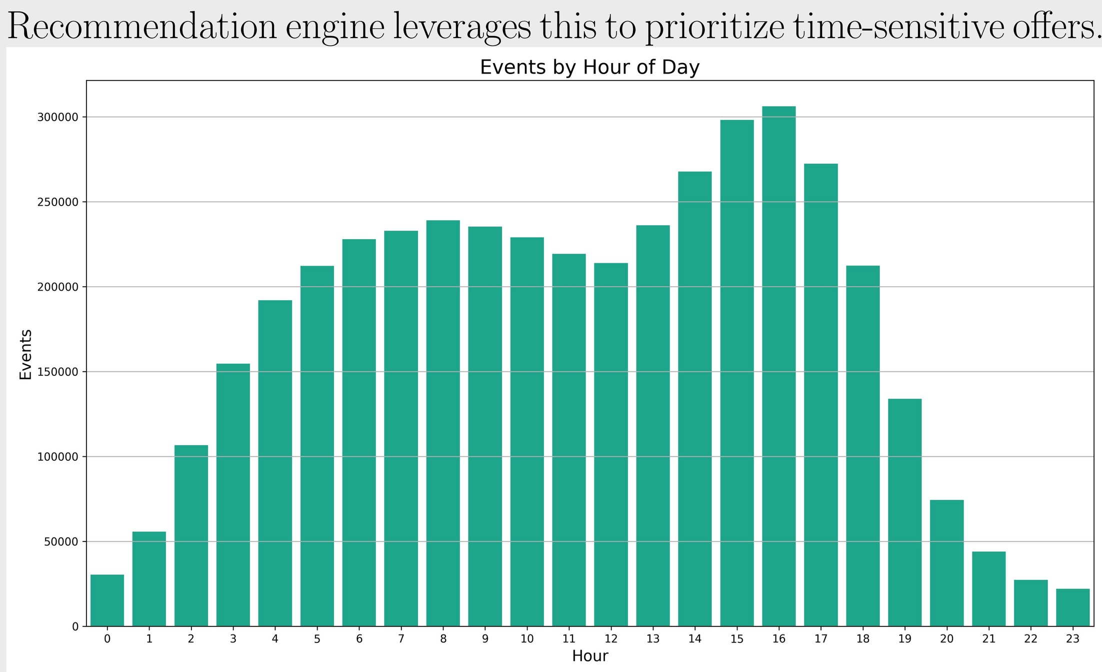


Temporal Insights

Hourly purchase analysis:

- Peak buying between 8–10 PM.
- Lowest prices and moderate activity at 1 AM — optimal discount timing.
- High-value purchases cluster in evening hours.

Recommendation engine leverages this to prioritize time-sensitive offers.

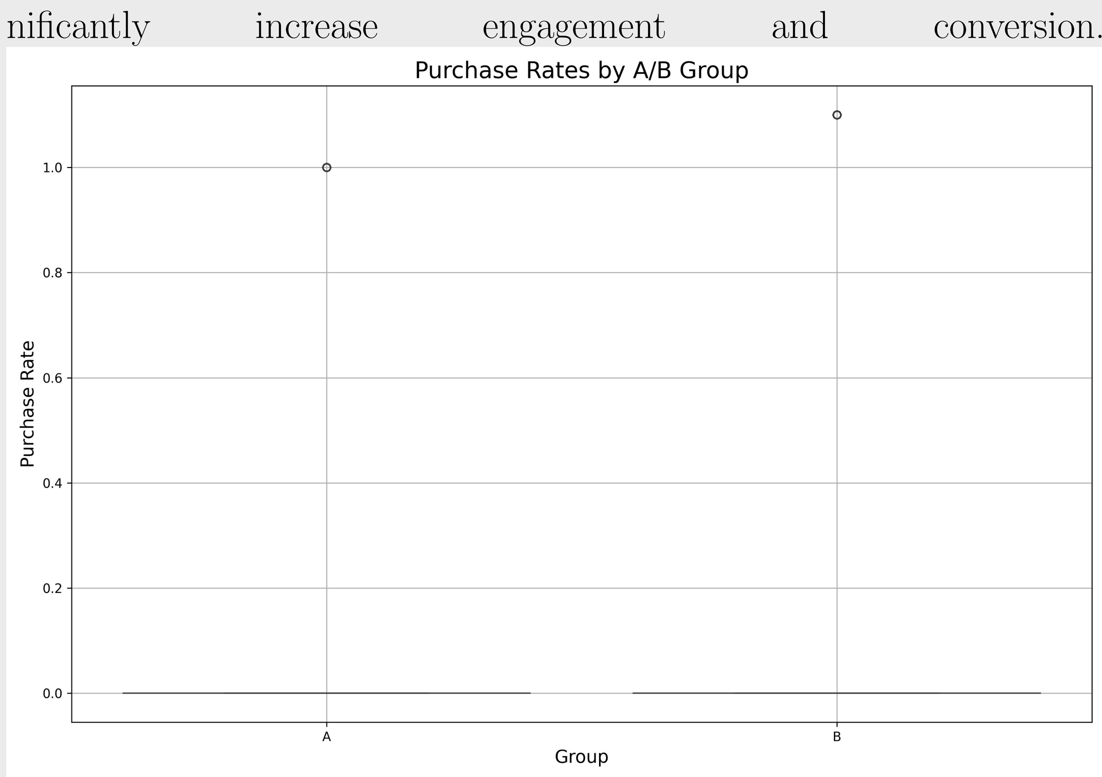


A/B Testing Results

A controlled experiment compared the baseline recommender (A) vs. the hybrid time-sensitive system (B):

- Conversion Rate: A = 0.0175, B = 0.0194
- **Lift: +11%**, $p = 7.08 \times 10^{-47}$ (statistically significant)

Conclusion: Time-aware recommendations significantly increase engagement and conversion.



Conclusion

This project shows how large-scale behavioral data can drive actionable insights. Through integrated data cleaning, exploration, prediction, and experimentation, we achieved:

- **Personalization:** +12.3% precision in tailored recommendations.
- **Engagement:** +11% lift verified via A/B tests.
- **Timing:** Identified low-price, high-conversion periods.

Takeaway: Time-aware recommendations connect exploration to conversion, turning insight into measurable impact.