

E-Commerce User Behavior Analysis & Recommendation System

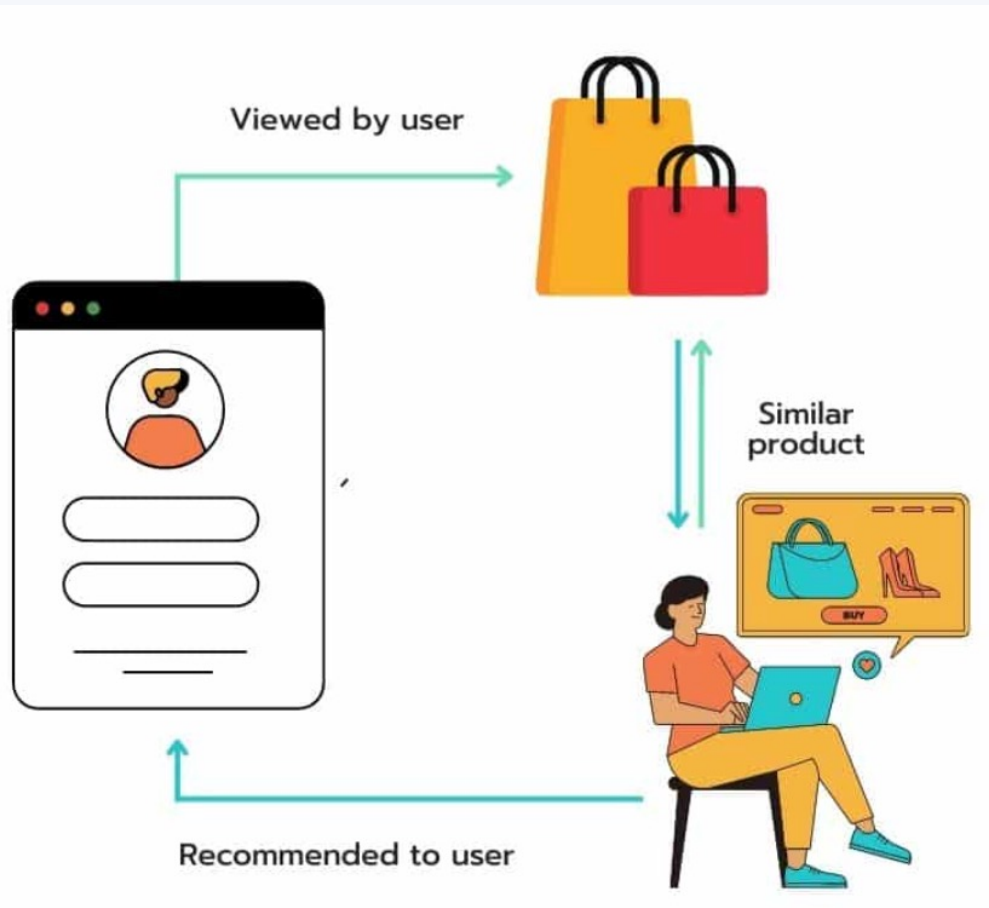
Amy Ngo, Priya Viswanadharao, Tom Steinman, Vy Ngo

ABSTRACT

Understanding consumer behavior is crucial in e-commerce, particularly in reducing cart abandonment and optimizing the user journey. This study analyzes cart abandonment rates, purchase timing patterns, and recommendation system effectiveness using collaborative filtering. Findings reveal that **30.08% of users abandon their carts**, with peak purchasing activity occurring **between 9 AM and 3 PM on weekends**. Additionally, a **user journey analysis highlights critical interaction points** where personalized recommendations enhance engagement and conversion. These insights help optimize the shopping experience, minimize drop-off rates, and refine product recommendations to improve overall sales.

OBJECTIVES

- Analyze cart abandonment behavior** to identify pain points in the checkout process.
- Visualize purchase timing trends** to determine peak shopping hours and their implications.
- Map the user journey** from site entry to purchase, examining how recommendations influence decisions.



MATERIALS

Dataset Overview

•Source: Kaggle
<https://www.kaggle.com/datasets/mkechinov/ecommerce-behavior-data-from-multi-category-store>

File structure	
Property	Description
event_time	Time when event happened at (in UTC).
event_type	Only one kind of event: purchase.
product_id	ID of a product
category_id	Product's category ID
category_code	Product's category taxonomy (code name) if it was possible to make it. Usually present for meaningful categories and skipped for different kinds of accessories.
brand	Downcased string of brand name. Can be missed.
price	Float price of a product. Present.
user_id	Permanent user ID.
** user_session**	Temporary user's session ID. Same for each user's session. Is changed every time user come back to online store from a long pause.

METHODOLOGY

1.Data Cleaning & Preprocessing

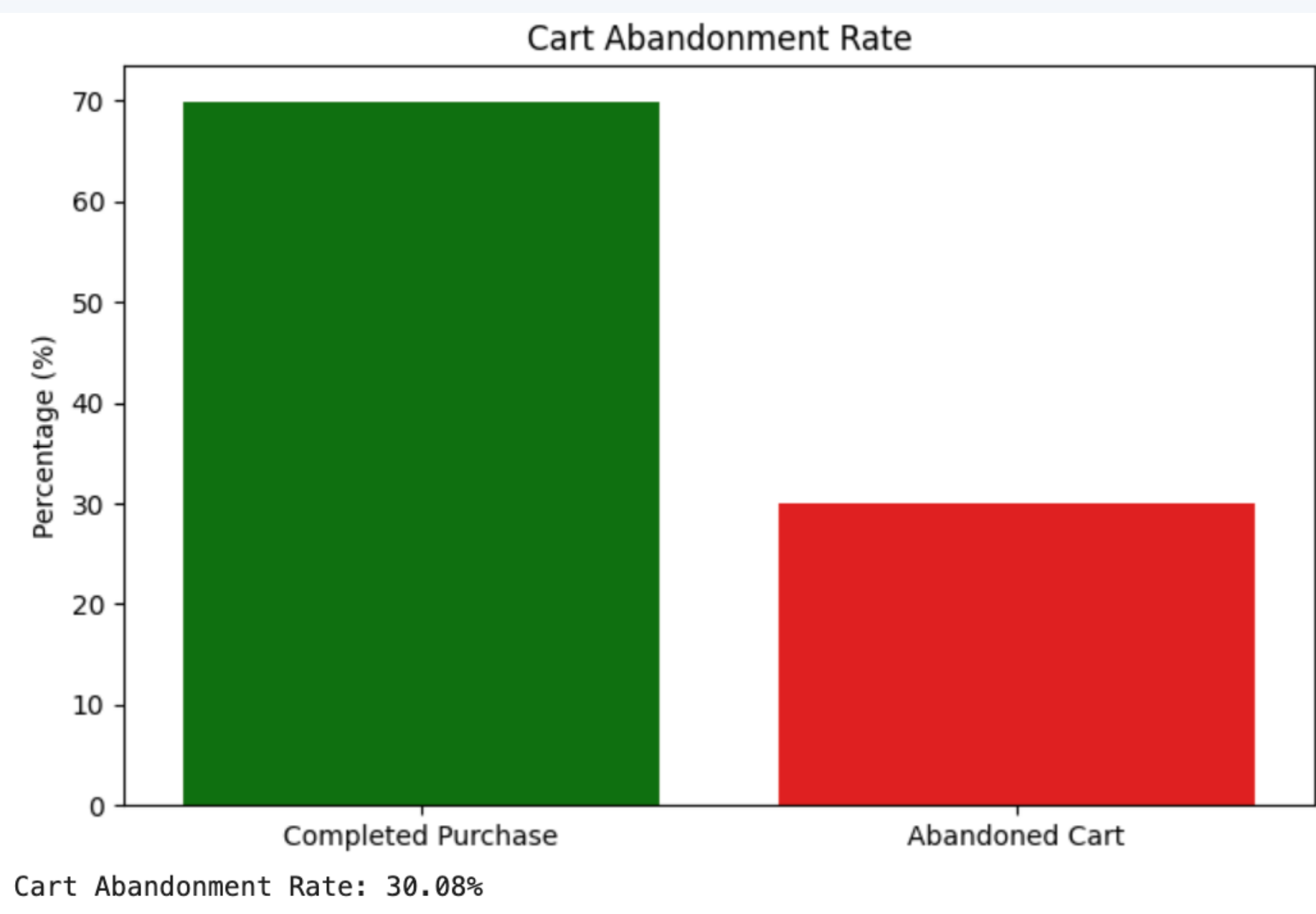
- Handling missing values (remove/fill in), formatting timestamps, combine two datasets.
- Transactional data from October and November (df_filled_oct and df_filled_nov) was combined and sampled for computational efficiency.

2. Cart Abandonment Analysis: Measured abandonment rates and identified the percentage of users who added products to their cart but did not complete the purchase.

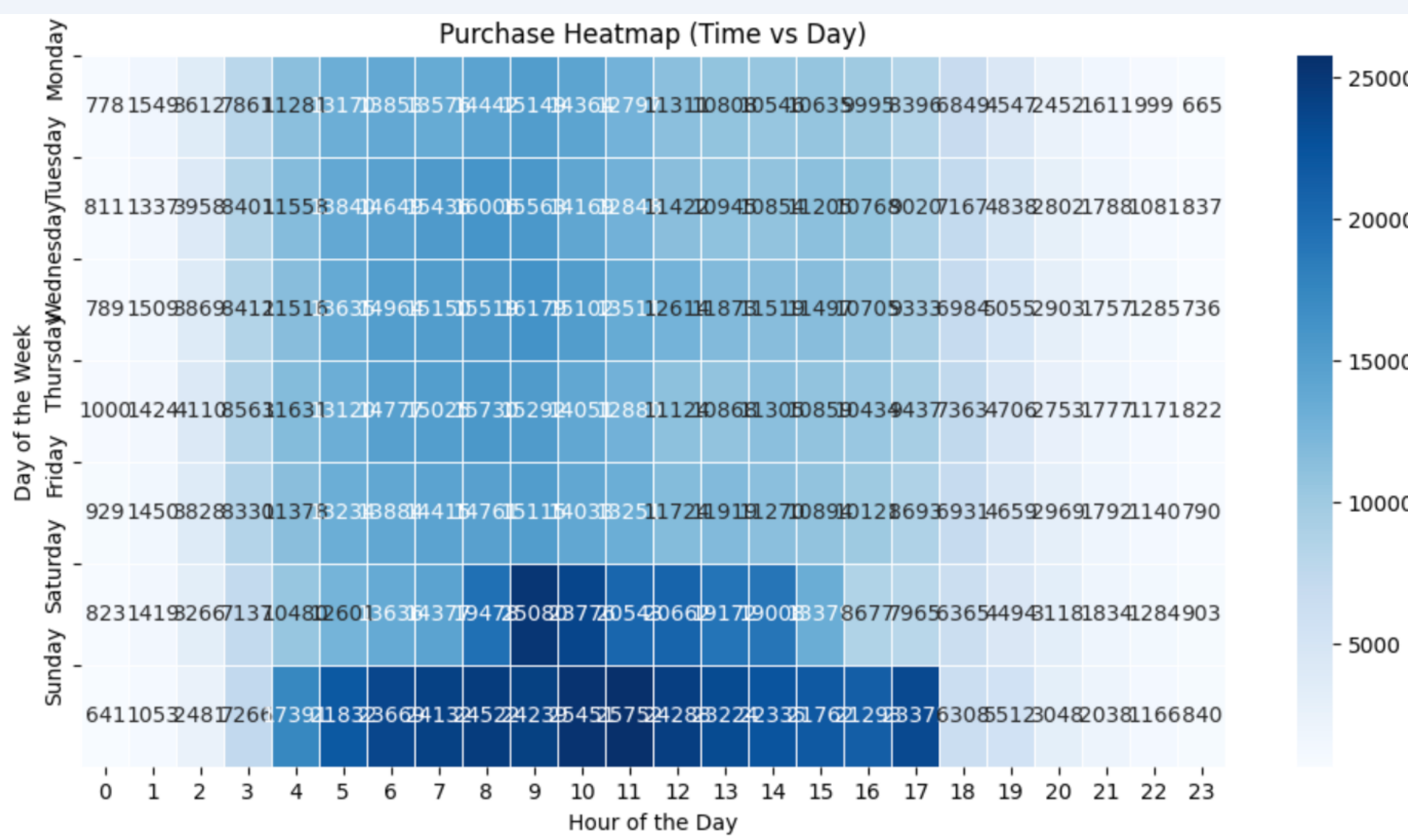
3. Purchase Timing Heatmap: Visualized hourly and daily purchasing patterns to highlight peak activity.

4. User Journey Mapping: Tracked interactions from site visit to checkout, integrating collaborative filtering for product recommendations.

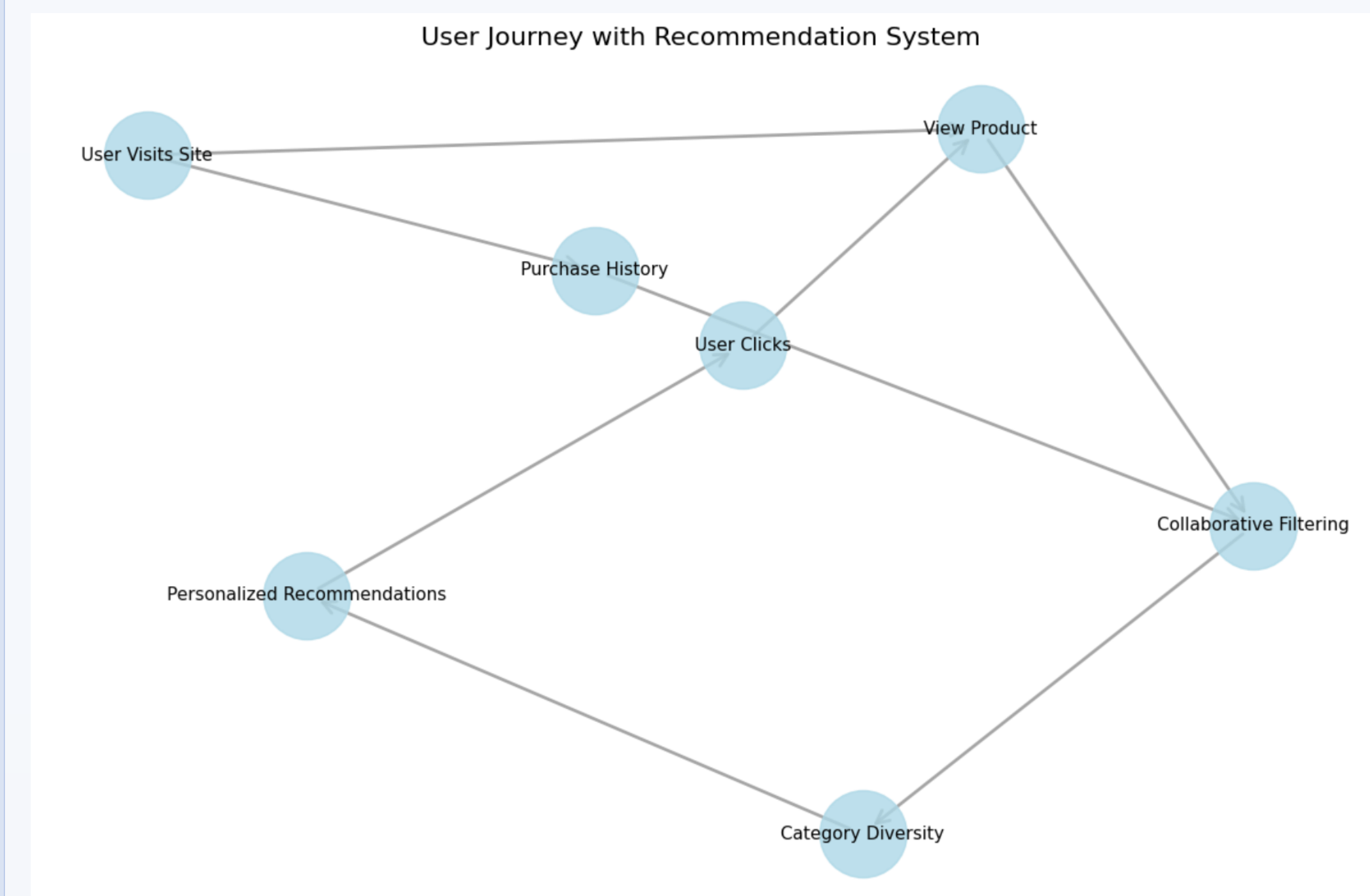
RESULTS



- The chart indicates that **30.08%** of users abandon their carts, while the majority proceed with the purchase.
- Nearly one in three shoppers abandon their carts.
- This indicates potential friction points in the checkout process.



- There's a **clear trend of higher purchasing activity between 9 AM and 3 PM**, particularly on **Friday, Saturday, and Sunday**.
- The highest intensity is seen around **late morning to early afternoon** on the weekends.
- Sunday experiences the highest late-night activity (compared to other days), possibly due to end-of-week shopping habits.



1.User Visits Site (Starting Point)

- The journey begins when a user enters the platform.
- The recommendation system begins collecting **user behavior data** at this stage.

2.User Clicks on a Product

- After browsing, the user engages by **clicking on a product**.
- This click triggers **personalized recommendations** based on past behavior and other users' actions.

3.Purchase History & Collaborative Filtering Influence Recommendations

- Purchase History:** The system suggests products based on what the user (or similar users) previously bought.
- Collaborative Filtering:** Uses data from other shoppers with similar interests to recommend relevant items.

4.View Product & Category Diversity

- Once a user views a product, additional recommendations are shown.
- The system ensures **category diversity**, suggesting different but related products to encourage exploration.

5.Personalized Recommendations & Final Purchase Decision

- The algorithm refines suggestions based on **user preferences, interactions, and session data**.
- If a user **doesn't purchase immediately**, the system may use retargeting (email, notifications, ads) to re-engage them.

Recommendation Model

	user_id	product_id	brand
0	543272936	17301504	creed
1	543272936	17301505	creed
2	543272936	17301506	creed
3	551377651	17301504	creed
4	551377651	17301505	creed
5	551377651	17301506	creed

CONCLUSIONS

- Cart abandonment remains a significant challenge** (30.08%), indicating potential improvements in checkout flow and payment options.
- Peak purchasing occurs between 9 AM and 3 PM, especially on weekends**, suggesting optimal time frames for targeted promotions.
- User journey analysis reveals critical engagement points**, where personalized recommendations can increase conversions.
- Collaborative filtering enhances recommendation accuracy**, leading to a more tailored shopping experience.
- Future improvements include refining recommendation algorithms, improving retargeting strategies, and optimizing the checkout process to reduce abandonment.**

REFERENCES

Source: Kaggle
eCommerce behavior data from multi category store
(<https://www.kaggle.com/datasets/mkechinov/ecommerce-behavior-data-from-multi-category-store>)

CONTACT

Amy Ngo: ango3636@gmail.com
Priya V.: Priya.viswanadharao@gmail.com
Tom Steinman: tasn78@gmail.com
Vy Ngo: vyngo9499@gmail.com