

## Abstract

This work is a course research project that collaborates with Fifth Third Bank. The motivation of the project is that through collaboration with reputable external organizations, it has become apparent that these organizations, while eager to foster partnerships with educational institutions, encounter substantial limitations in sharing comprehensive datasets, particularly concerning customer personal information. The purpose of this project is to use the user transaction dataset which contains the daily transaction activities to address the development of an AI-based recommendation system. The issue here is. In this unique setting, the AI model is built on only essential purchase data. The study focuses on leveraging minimal transaction data to enhance the recommendation system's performance.

## Datasets

The dataset we used is a transaction dataset from Fifth Third bank. The dataset captures a snapshot of user transactions, offering a nuanced view of consumer behaviors and spending habits. It provides valuable insights into the diverse array of transactions, encompassing various merchant categories and spanning different locations. Each transaction is characterized by details such as the transaction amount, merchant information, and geographic coordinates, shedding light on the patterns of financial activity in different cities and countries. The dataset's temporal information, including transaction and posting dates, enables a temporal analysis of consumer spending trends. This dataset serves as a rich resource for researchers and analysts interested in delving into the dynamics of user-initiated payments, with potential applications in financial analytics, consumer behavior studies, and the development of intelligent recommendation systems.

## Project Description

- **Data Preprocessing**
  - Within this stage, our primary emphasis will be directed towards the meticulous preparation of the dataset. This intricate process encompasses not only the thorough organization and structuring of the data but also, significantly, the implementation of robust anonymization and minimization techniques applied to the bank transaction dataset.
- **Model Design**
  - We would like to implement collaborative filtering [1] (Figure 1) and extrapolative collaborative filtering models for recommendation generation. Extrapolative Collaborative Filtering [2] is an advanced extension of traditional collaborative filtering methods in recommendation systems. While conventional collaborative filtering primarily relies on user-item interactions and similarities, extrapolative collaborative filtering goes a step further by incorporating additional contextual information to enhance the accuracy and relevance of recommendations.
- **Evaluation and Optimization**
  - Assessing the performance of the recommendation system using relevant metrics. Several metrics are commonly used to evaluate recommendation system performance.
  - Fine-tuning models and algorithms for improved accuracy and efficiency. Fine-tuning machine learning involves optimizing the parameters and configurations to enhance the system's performance.

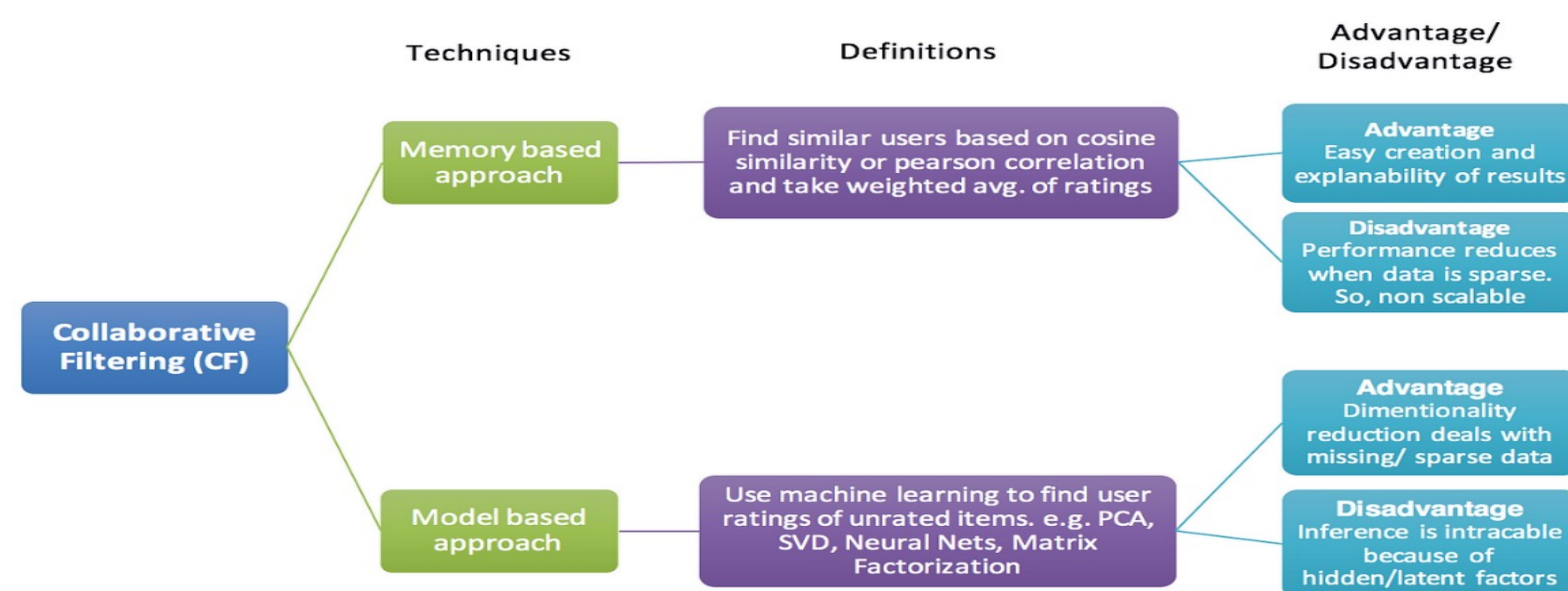


Figure 1. Collaborative Filtering.

## Results

So far, we have conducted a comprehensive preliminary analysis of the dataset, unearthing valuable insights into user transactions. The initial findings are particularly encouraging, revealing promising trends and patterns that have been visually elucidated in Figure 2 and Figure 3, enhancing our understanding of the intricate dynamics within the dataset.

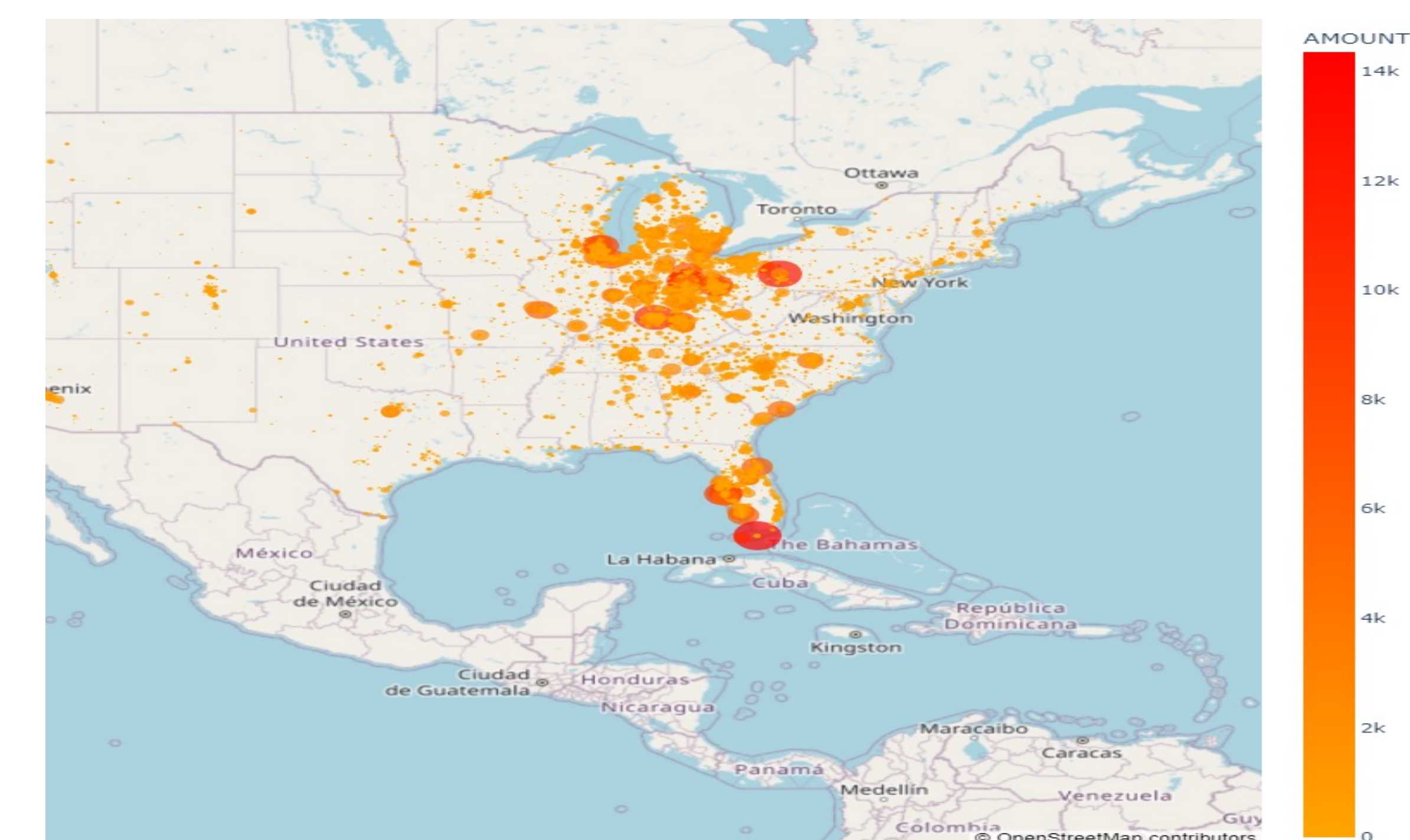


Figure 2. Visualization of the dataset.

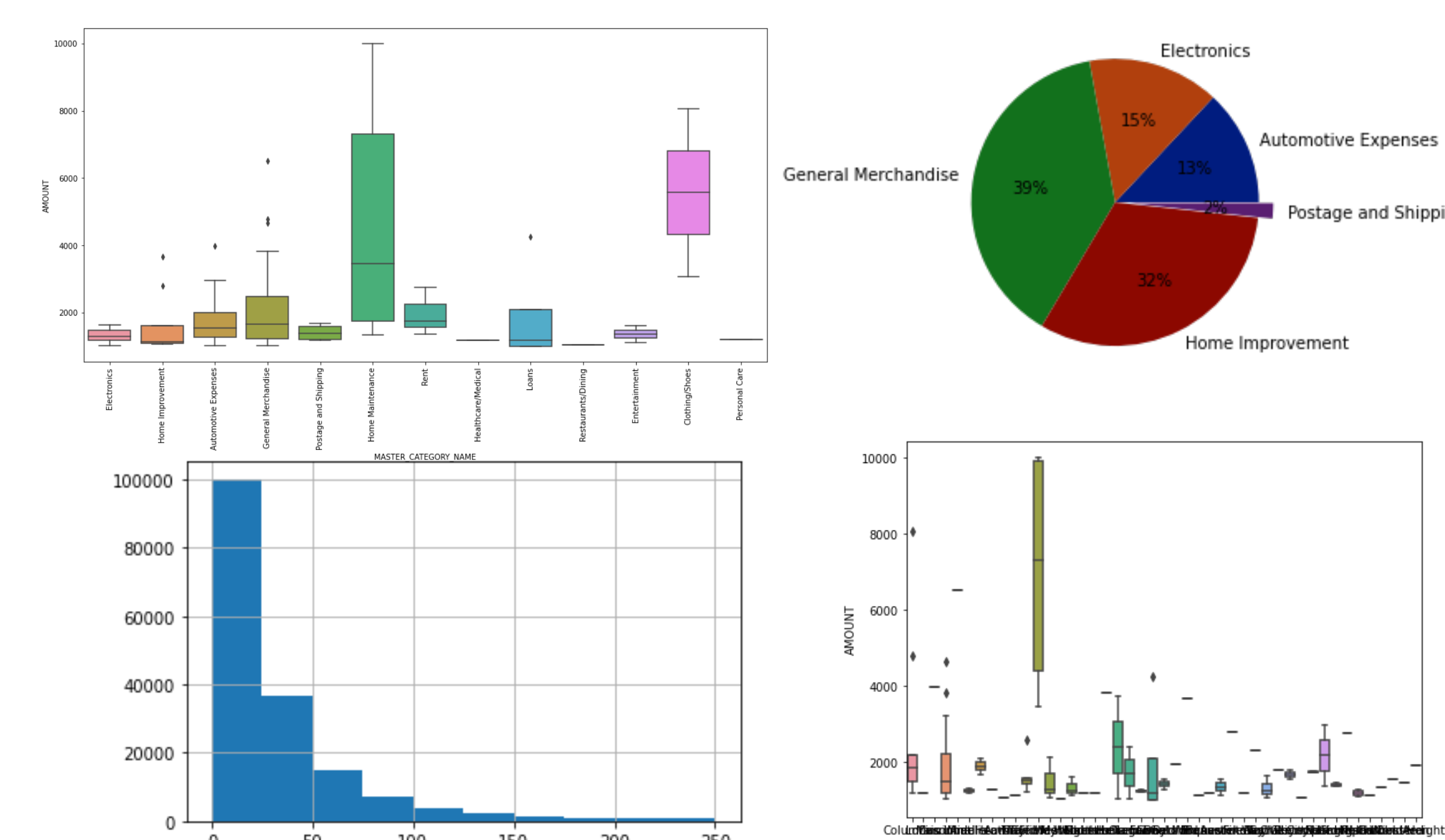


Figure 3. Visualization of the dataset.

## Conclusion & Future Directions

So far, we have understood the dataset through various data exploration approaches. The next step is using a data transformation strategy such as PCA to hide the sensitive information and developing suitable models for the recommendation system which includes topics like user segmentation, product recommendation, fraud detection, visualization, (visualize the dataset with provided geographic info) and predict user spending amount.

## References

- [1] Su, X., & Khoshgoftaar, T. M. (2009). A survey of collaborative filtering techniques. *Advances in artificial intelligence*, 2009.
- [2] Koren, Y. (2009, June). Collaborative filtering with temporal dynamics. In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 447-456).