Formative Report

Abstract

This report offers an analytical exploration of customer behavior within a commercial transaction dataset. By excluding personal transfer data, the focus narrows to commercial transactions to discern spending patterns and preferences, utilizing the RFM model for customer classification.

Introduction

The study investigates commercial transactions, segregating personal transfers to scrutinize customer and merchant behavior more effectively. The purpose is to identify spending trends and group customers to facilitate targeted marketing strategies.

Literature Review

The concept of customer segmentation has been extensively explored, with significant emphasis on the RFM model, which considers recency, frequency, and monetary value as the key metrics for segmenting customers. Shirole et al. conducted a study that employed the RFM model alongside the K-means clustering algorithm, delineating customers into meaningful clusters to enhance marketing strategies and customer relationship management [1]. This sentiment is echoed by Aliyev et al., who asserted the importance of behavioral segmentation in banking, employing unsupervised machine learning algorithms to refine customer segmentation [2]. Furthermore, the work by Qiasi et al. outlines the role of RFM technique and clustering algorithms in measuring customer loyalty and value, demonstrating their critical role in customer relationship management strategies [3].

Methodology

The dataset was pruned of personal transfer data, emphasizing commercial transactions. Two primary dimensions, customer and merchant, were scrutinized. The RFM model was employed for customer classification, and spending was categorized into 19 behavioral types.

Data Description/Preparation

Data mining techniques revealed spending patterns across different merchant categories, with particular attention to monthly and yearly expenditures. Manual labeling was employed in the RFM model, albeit with a risk of subjectivity-induced errors.

Results and Discussions

December showed a universal spending increase, attributed to holiday sales events. A disproportionate spending concentration in bars and supermarkets was observed, suggesting potential growth areas. Customer ranking by bar spending yielded targeted customer lists for potential marketing strategies.

Further Work and Improvement

The need for objective classification prompts a move away from manual labeling towards machine learning methods like K-means and DBSCAN. Future work involves refining these techniques and incorporating mentor insights for enhanced data interpretation.

Conclusion

Preliminary findings suggest significant seasonal influences on spending and concentration of expenditures in specific categories. These insights guide targeted marketing initiatives, although a shift towards automated classification models is essential for reliability.

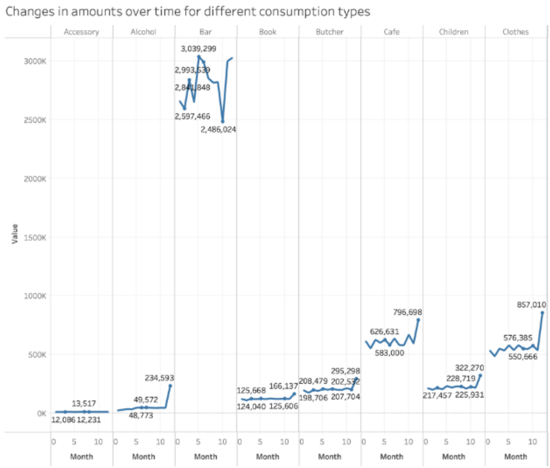
References

[1] R. Shirole, L. Salokhe, and S. Jadhav, "Customer Segmentation using RFM Model and K-Means Clustering," International Journal of Scientific Research in Science and Technology, vol. 8, no. 3, pp. 591-597, May-June 2021.

[2] M. Aliyev, E. Ahmadov, H. Gadirli, A. Mammadova, and E. Alasgarov, "Segmenting Bank Customers via RFM Model and Unsupervised Machine Learning," ADA University School of Information Technologies and Engineering, 2021.

[3] R. Qiasi, M. Baqeri-Dehnavi, B. Minaei-Bidgoli, "Developing a model for measuring customer loyalty and value with RFM technique and clustering algorithms," The Journal of Mathematics and Computer Science, vol. 4, no. 2, pp. 172-181, 2012.

Appendices



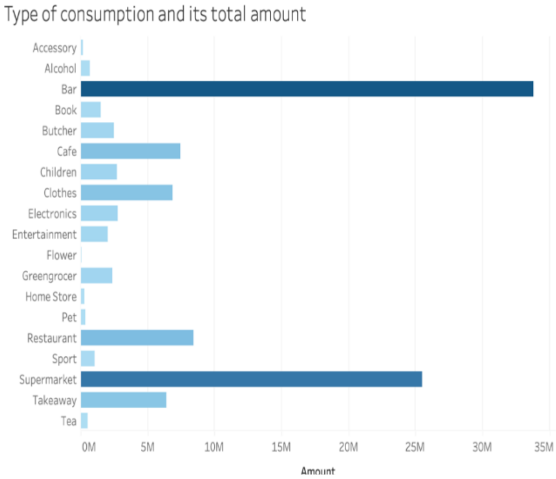


Fig. 1. 12 Months of Different Types of Amount Curves Fig. 2. Consumption Types and Total Amount

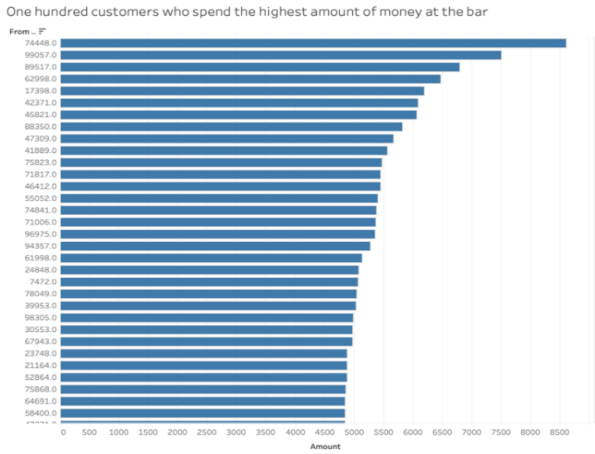


Fig. 3. Ranking Bar Customers in Descending Order of Consumption Amount

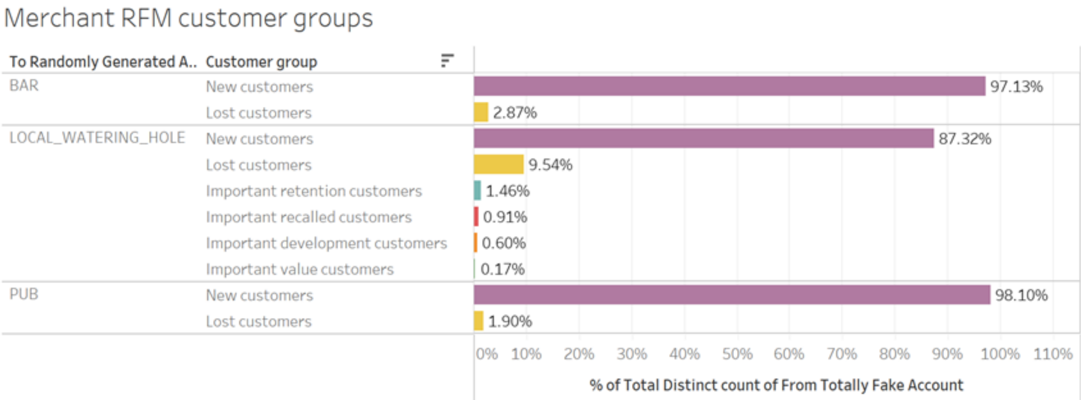


Fig. 4. RFM Analysis for The Top Three Stores in The Bar Category