



# INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

## RFM ANALYSIS USING DATA ANALYTICS

<sup>1</sup>Avula Prasanth Kumar Reddy, <sup>2</sup>Gurrula Yatish Durga Rao, <sup>3</sup>Manda Jayanth, <sup>4</sup>V. Kapil

<sup>1</sup>Student, <sup>2</sup>Student, <sup>3</sup>Student, <sup>4</sup>Assistant Professor

<sup>1</sup>CSE - Data Science,

<sup>1</sup>SRK Institute of Technology, Vijayawada, India

**Abstract:** This project delves into the application of RFM (Recency, Frequency, Monetary) analysis for customer segmentation, utilizing Python as the primary analytical tool. Beginning with a theoretical overview, it highlights RFM's significance in modern marketing, emphasizing its role in understanding customer behavior. Through Python's data analysis capabilities, the project demonstrates the extraction of insights from transaction data, ensuring accuracy through meticulous cleaning procedures. By computing RFM scores, it enables the segmentation of customers, facilitating targeted marketing and resource allocation. Visualization techniques are explored to represent segmentation results intuitively. Practical implications are discussed, showcasing how businesses can enhance customer-centric strategies through RFM analysis and Python programming. Ultimately, the project underscores the transformative potential of these tools in reshaping customer relationship management paradigms, empowering businesses to navigate the complexities of the digital age effectively.

**Index Terms** - RFM (Recency, Frequency, Monetary) analysis, Python, Customer behavior, Transaction data, Data analysis

### I. INTRODUCTION

RFM analysis (Recency, Frequency, Monetary) is a strategic marketing tool that uses customer behavior data to forecast future actions. Emerging from direct mail marketing, it became notable in a 1995 article by Tom Wansbeek and Jan Roelf Bult, which highlighted the Pareto Principle, indicating that a small percentage of customers generate a large share of sales.

RFM analysis focuses on three key factors: recency of transaction, frequency of engagement, and monetary expenditure. This methodology allows businesses to categorize customers based on their transaction patterns, enabling personalized marketing and customer communication. The main benefits of RFM analysis include improved revenue generation through targeted customer segmentation, leading to higher response rates, better customer retention, and increased customer lifetime value (CLTV).

It provides insights into customer behavior: recency helps identify likely repeat customers, frequency indicates loyalty, and monetary value reflects potential high-spending customers. By segmenting customers based on RFM scores, businesses can create targeted marketing campaigns. For example, customers with frequent high-value transactions might respond well to premium product promotions, while those with consistent lower spending may appreciate loyalty rewards or referral incentives. Traditional segmentation often relied on demographic data, but with modern analytics and customer data platforms (CDPs), behavioral segmentation has become more precise. RFM analysis uses these granular insights to develop dynamic customer segments and tailor marketing efforts effectively.

Customer relationship management (CRM) systems can track RFM metrics, aiding businesses in segmenting customers and devising tailored marketing strategies. This approach empowers businesses to foster personalized customer experiences, build stronger loyalty, and drive sustainable growth.

## II. RESEARCH METHODOLOGY

RFM analysis (Recency, Frequency, Monetary) is a strategic marketing tool that uses customer behavior data to forecast future actions. Emerging from direct mail marketing, it became notable in a 1995 article by Tom Wansbeek and Jan Roelf Bult, which highlighted the Pareto Principle, indicating that a small percentage of customers generate a large share of sales.

RFM analysis focuses on three key factors: recency of transaction, frequency of engagement, and monetary expenditure. This methodology allows businesses to categorize customers based on their transaction patterns, enabling personalized marketing and customer communication. The main benefits of RFM analysis include improved revenue generation through targeted customer segmentation, leading to higher response rates, better customer retention, and increased customer lifetime value (CLTV).

It provides insights into customer behavior: recency helps identify likely repeat customers, frequency indicates loyalty, and monetary value reflects potential high-spending customers. By segmenting customers based on RFM scores, businesses can create targeted marketing campaigns. For example, customers with frequent high-value transactions might respond well to premium product promotions, while those with consistent lower spending may appreciate loyalty rewards or referral incentives.

Traditional segmentation often relied on demographic data, but with modern analytics and customer data platforms (CDPs), behavioral segmentation has become more precise. RFM analysis uses these granular insights to develop dynamic customer segments and tailor marketing efforts effectively. Customer relationship management (CRM) systems can track RFM metrics, aiding businesses in segmenting customers and devising tailored marketing strategies.

This approach empowers businesses to foster personalized customer experiences, build stronger loyalty, and drive sustainable growth.

This proposed project focuses on leveraging data analytics techniques to enhance customer segmentation and analysis without relying on machine learning algorithms. By utilizing the Pandas library for data manipulation and Plotly sub-frameworks like express and graph objects for visualization, we aim to provide businesses with actionable insights into customer behavior. Our approach emphasizes the importance of understanding and interpreting data trends through dynamic visualizations, enabling stakeholders to make informed decisions.

### Advantages:

- Utilizes data analytics techniques for customer segmentation and analysis
- Relies on Pandas library for efficient data manipulation
- Utilizes Plotly sub-frameworks for dynamic and interactive visualizations
- Provides actionable insights into customer behavior
- Offers customizations tailored to specific business needs
- Enhances decision-making processes for stakeholders
- Improves understanding of customer preferences and trends
- Empowers businesses to optimize marketing campaigns and promotions
- Enhances customer satisfaction and retention strategies

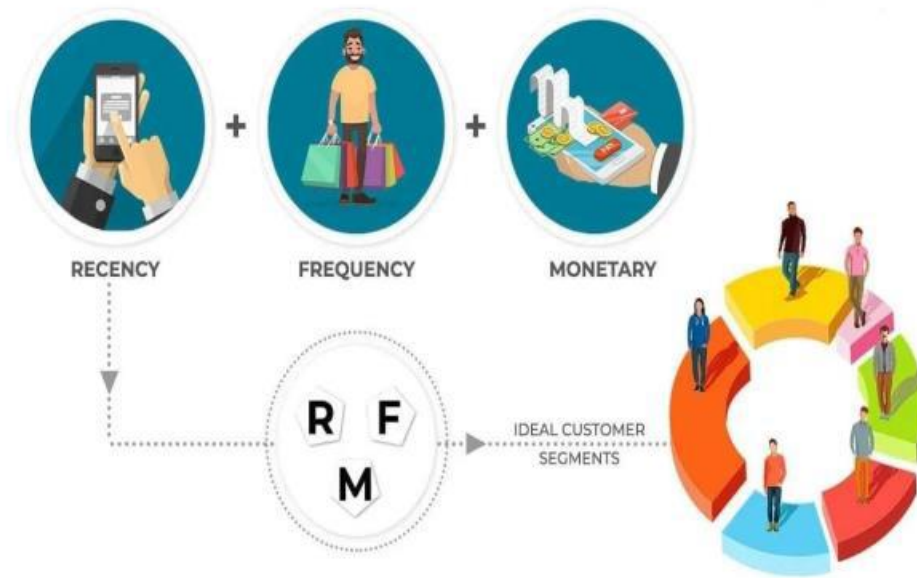


Fig 1: RFM Analysis

### III. RESULTS

CustomerID	PurchaseDate	TransactionAmount	ProductInformation	OrderID
0	8814	2023-04-11	943.31	Product C 890075
1	2188	2023-04-11	463.70	Product A 176819
2	4608	2023-04-11	80.28	Product A 340062
3	2559	2023-04-11	221.29	Product A 239145
4	9482	2023-04-11	739.56	Product A 194545

Location
0 Tokyo
1 London
2 New York
3 London
4 Paris

Fig 2: Original Dataset

CustomerID	PurchaseDate	TransactionAmount	ProductInformation	OrderID	Location	Recency	Frequency	MonetaryValue	Recency_score	Frequency_score	Monetary_score
0	8814	2023-04-11	943.31	Product C 890075	Tokyo	339	1	943.31	1	1	2
1	2188	2023-04-11	463.70	Product A 176819	London	339	1	463.70	1	1	1
2	4608	2023-04-11	80.28	Product A 340062	New York	339	1	80.28	1	1	1
3	2559	2023-04-11	221.29	Product A 239145	London	339	1	221.29	1	1	1
4	9482	2023-04-11	739.56	Product A 194545	Paris	339	1	739.56	1	1	2

Fig 3: Dataset after calculating scores of RFM

CustomerID	PurchaseDate	TransactionAmount	ProductInformation	OrderID	Location	Recency	Frequency	MonetaryValue	Recency_score	Frequency_score	Monetary_score	RFM_score	Value Segment
0	2023-04-11	943.31	Product C	890075	Tokyo	339	1	943.31	1	1	2	4	Low-Value
1	2023-04-11	463.70	Product A	178819	London	339	1	463.70	1	1	1	3	Low-Value
2	2023-04-11	80.28	Product A	340062	New York	339	1	80.28	1	1	1	3	Low-Value
3	2023-04-11	221.29	Product A	238145	London	339	1	221.29	1	1	1	3	Low-Value
4	2023-04-11	739.56	Product A	194545	Paris	339	1	739.56	1	1	2	4	Low-Value

Fig 4: Dataset after segmenting customers based on RFM score

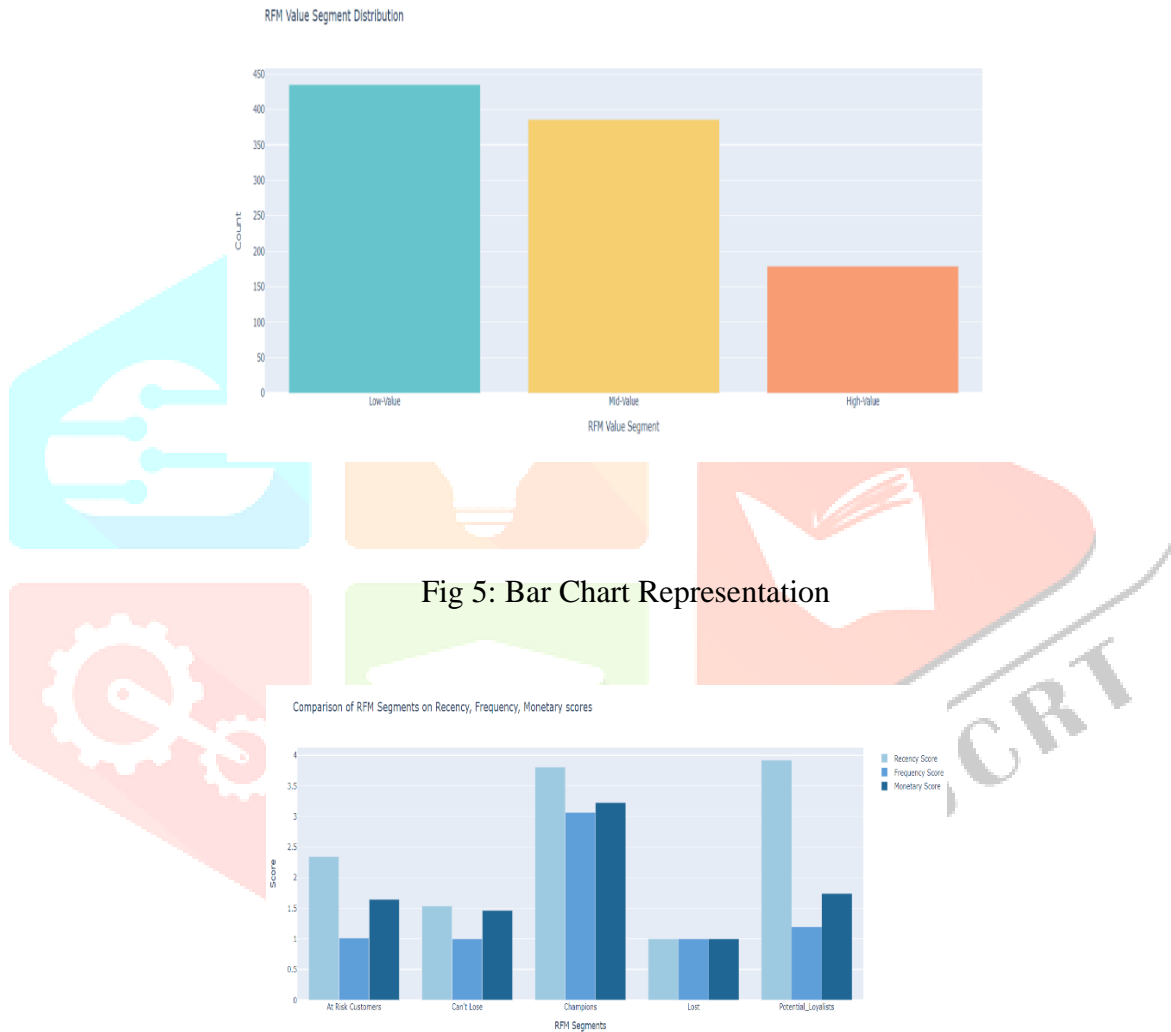


Fig 5: Bar Chart Representation

Fig 6: Comparison of RFM Segments using grouped Bar Chart

#### IV. CONCLUSION

This project demonstrates the effectiveness of RFM Analysis in Python, utilizing advanced data analytics techniques to segment customers based on recency, frequency, and monetary value. Through interactive visualizations with Plotly, businesses can gain actionable insights, optimize marketing strategies, and enhance customer engagement in today's competitive landscape.

## V. REFERENCES

- [1] Aggarwal, A. G., & Yadav, S. (2020). Customer segmentation using Fuzzy-AHP and RFM model. 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO).
- [2] Christy, A., Umamakeswari, A., Priyatharsini, L., & Neyaa, A. (2021). RFM ranking – An effective approach to customer segmentation. *Journal Of King Saud University - Computer And Information Sciences*, 33(10), 1251-1257.
- [3] Dursun, A., & Caber, M. (2016). Using data mining techniques for profiling profitable hotel customers: An application of RFM analysis. *Tourism Management Perspectives*, 18, 153-160.
- [4] Martinez, R. G., Carrasco, R. A., García-Madariaga, J., Gallego, C. P., & Herrera-Viedma, E. (2019). A comparison between fuzzy linguistic RFM model and traditional RFM model applied to campaign management. case study of retail business. *Procedia Computer Science*, 162, 281–289.

