

Customer Segmentation in E-Commerce Using RFM Analysis and K-Means Clustering

Priyanka Dilip Maisgawali¹, Dr. Ujwala Narkhede²

¹MBA – Business Analytics (Appearing), Wainganga College of Engineering and Management, Nagpur, Maharashtra, India

²Professor, Wainganga College of Engineering and Management, Nagpur, Maharashtra, India

Corresponding Author Details:

Priyanka Dilip Maisgawali, MBA – Business Analytics (Appearing), Wainganga College of Engineering and Management, Nagpur, Maharashtra, India.

Email: priyankamaisgawali@gmail.com

Abstract

E-commerce businesses collect a large amount of customer data such as purchase history, spending, frequency of purchases and product preferences. Customer segmentation helps businesses divide customers into different groups based on their behaviour. This research explains how data analytics can be used for customer segmentation in e-commerce. The study focuses on RFM analysis (Recency, Frequency and Monetary Value) and K-Means clustering. Customer segmentation helps businesses provide personalised offers, improve customer satisfaction, retain customers and make better marketing decisions.

Keywords: E-Commerce, Customer Segmentation, Data Analytics, RFM Analysis, K-Means Clustering, Customer Behaviour, Digital Marketing, Personalization

1. Introduction

E-commerce has changed the way people buy products. Customers purchase products through websites and mobile applications, creating large amounts of data.

This data can help companies understand:

- What customers buy
- How often they buy
- How much they spend
- When they last purchased
- Which products they prefer

Customer segmentation means dividing customers into groups with similar behaviour. Data analytics helps businesses identify these groups and create suitable marketing strategies.

Background of the Study

Traditional marketing often treats customers as one large group. However, customers have different needs, preferences and purchasing habits.

For example:

- Some customers purchase frequently.
- Some customers purchase only during discounts.
- Some customers spend a large amount.
- Some customers have not purchased for a long time.
- Some customers are new to the business.

Treating all these customers in the same way may not be effective.

Data analytics allows businesses to identify these differences. A commonly used approach is RFM analysis, where:

- Recency (R): How recently the customer made a purchase.
- Frequency (F): How often the customer makes purchases.
- Monetary (M): How much money the customer spends.

RFM can then be combined with clustering methods to create customer groups. Recent research continues to use and extend RFM with additional behavioural variables such as basket size, campaign participation and shopping regularity.

Problem Statement

E-commerce businesses serve customers with different purchasing patterns and expectations. Treating all customers as a single group can result in ineffective marketing campaigns, inappropriate promotional offers, inefficient use of advertising budgets, and difficulty in retaining valuable customers.

The major problems include:

1. Difficulty in understanding different customer behaviours.
2. Lack of personalised marketing strategies.
3. Difficulty in identifying valuable customers.
4. Difficulty in retaining inactive customers.
5. Inefficient use of customer data.
6. Marketing campaigns being sent to customers without considering their individual behaviour.

Therefore, there is a need for data-driven customer segmentation.

Objectives of the Study

The main objectives of this research are:

1. To understand the concept of customer segmentation in e-commerce.
 2. To study the role of data analytics in customer segmentation.
 3. To understand the use of RFM analysis.
 4. To examine the use of K-Means clustering for customer segmentation.
 5. To identify different customer groups based on purchasing behaviour.
 6. To understand how segmentation can improve marketing decisions.
 7. To study the benefits and challenges of data-driven segmentation.
-

2.Literature Review Customer Segmentation

Customer segmentation means dividing customers into groups based on similar characteristics. Traditional methods use factors like age, gender, income, location, and occupation. In e-commerce, customer behaviour and purchase history are also used to understand customers better.

Data Analytics in E-Commerce

Data analytics helps e-commerce businesses convert customer data into useful information. It can be used to understand buying habits, customer preferences, product demand, and customer value.

The main types of analytics are:

- Descriptive Analytics: What happened?
- Diagnostic Analytics: Why did it happen?
- Predictive Analytics: What may happen?
- Prescriptive Analytics: What action should be taken?

Customer segmentation mainly uses descriptive and predictive analytics to understand customers and plan marketing activities.

RFM Analysis

RFM stands for Recency, Frequency, and Monetary Value. It helps businesses identify valuable customers based on how recently and frequently they purchase and how much they spend. Customers who buy recently, purchase often, and spend more are usually considered valuable. Customers who buy rarely or have not purchased for a long time may need special offers or re-engagement.

Clustering Techniques

Clustering is a machine-learning method used to group similar customers. K-Means is a commonly used clustering technique. It divides customers into different groups based on their similarities. This is useful when businesses want to analyse several customer characteristics at the same time. This version is suitable for a college research paper and is easier to understand while keeping all four topics and key points.

3. Research Methodology

Research Design

This study uses a descriptive and analytical research design. It explains customer segmentation and how data analytics can help e-commerce businesses understand their customers.

Data Sources segmentation and Sampling

The study can use a sample of 100 e-commerce customers. The respondents can be selected using convenience sampling. The sample includes customers who regularly use online shopping platforms.

- Primary Data: Data collected through customer surveys or questionnaires.
- Secondary Data: Data collected from research papers, journals, e-commerce reports, books, websites, and public datasets.

Variables Used

The main variables are Recency, Frequency, Monetary Value, Product Preference, and Discount Behaviour.

RFM Analysis

- Recency: How recently a customer purchased.
- Frequency: How often a customer purchases.
- Monetary: How much a customer spends.

RFM helps identify valuable customers.

K-Means Customer Segmentation

K-Means groups customers with similar buying behaviour.

Data → RFM Analysis → K-Means → Customer Groups → Marketing Strategy

Main groups can include high-value, regular, new, at-risk, and low-value customers.

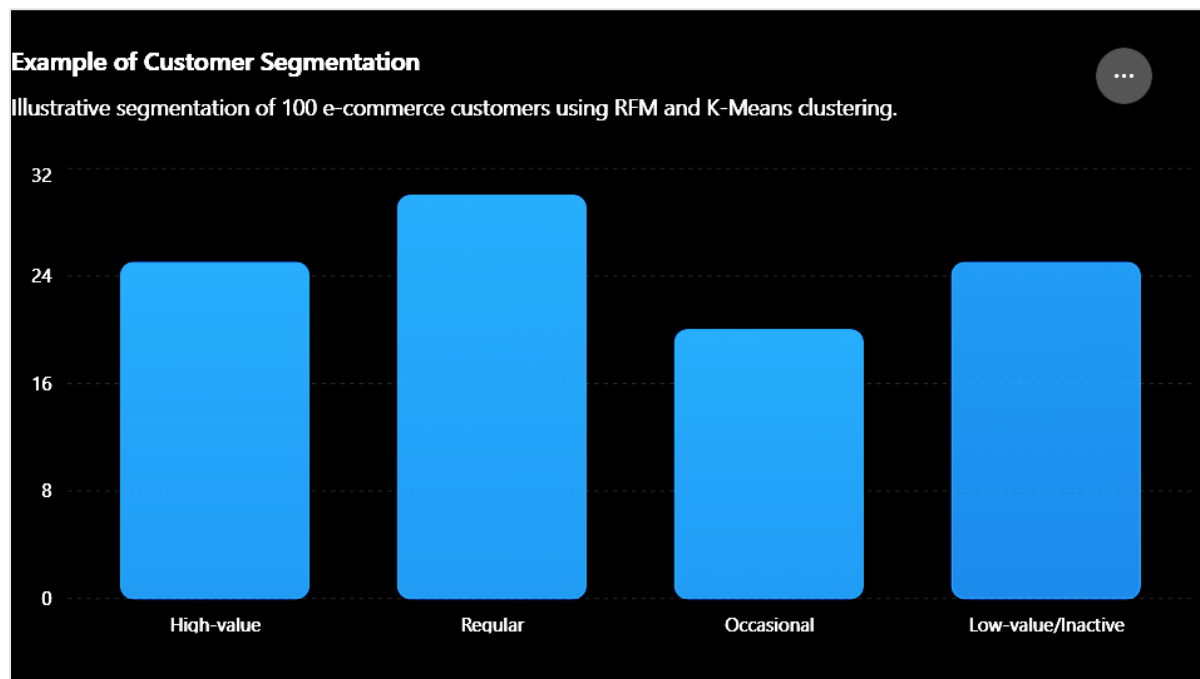
4.Data Analysis and Interpretation

For this study, data of 100 customers can be analysed using Excel, Python, R or Power BI. After removing duplicate records and handling missing values, Recency, Frequency and Monetary (RFM) values can be calculated. K-Means clustering can then be used to divide customers into groups.

Customer Segment	No. of Customers	Percentage	Customer Type
Segment 1	25	25%	High-value customers
Segment 2	30	30%	Regular customers
Segment 3	20	20%	Occasional customers
Segment 4	25	25%	Low-value/inactive customers
Total	100	100%	

Example of Customer Segmentation

Illustrative segmentation of 100 e-commerce customers using RFM and K-Means clustering.



Interpretation

The example shows that 30% are regular customers, 25% are high-value customers, 20% are occasional customers and 25% are low-value/inactive customers.

Based on these groups, businesses can use different marketing strategies:

- High-value: Loyalty rewards and premium offers.

- Regular: Discounts and personalised recommendations.
- Occasional: Special offers to increase purchases.
- Low-value/Inactive: Re-engagement messages and attractive offers.

The Elbow Method can be used to select the suitable number of clusters, while the Silhouette Score can help check how well the customers are separated into groups.

5. Findings and Discussion

The study finds that:

- Customers have different buying habits.
- Data analytics helps identify customer groups.
- RFM measures customer value.
- K-Means groups similar customers.
- Segmentation helps in personalised marketing.
- Valuable and inactive customers can be identified.

Benefits

- Personalised offers
- Better customer retention
- Better customer experience
- Effective marketing
- Improved customer loyalty

Challenges

- Poor or missing data
- Large amount of data
- Changing customer behaviour
- Difficulty in understanding groups, Privacy and security issues

Managerial Implications

Managers can use customer segments for targeted marketing, loyalty programmes, better recommendations and customer retention.

Conclusion

Customer segmentation is an important use of data analytics in e-commerce. It helps businesses understand different types of customers and their buying behaviour.

RFM analysis helps measure customers based on recency, frequency, and spending. K-Means clustering helps group customers with similar behaviour. These methods can support personalised marketing, customer retention, better customer experience, and decision-making.

However, businesses should maintain data quality, customer privacy, and responsible use of data.

Overall, data analytics helps e-commerce companies understand customers and create better marketing strategies.

References

1. Gomes, M. A., & Meisen, T. (2023). A review on customer segmentation methods for personalized customer targeting in e-commerce use cases. *Information Systems and e-Business Management*, 21, 527–570.
2. Özcan, T. (2026). Customer Segmentation Using an Extended RFM Model and Clustering Algorithms in E-Commerce. *Journal of Theoretical and Applied Electronic Commerce Research*, 21(5), 142.
3. Wong, C. G., Tong, G. K., & Haw, S. C. Exploring Customer Segmentation in E-Commerce using RFM Analysis with Clustering Techniques. *Journal of Telecommunications and the Digital Economy*.
4. Salminen, J., Mustak, M., Sufyan, M., et al. (2023). How can algorithms help in segmenting users and customers? *Journal of Marketing Analytics*, 11, 677–692.
5. Das, A., Mishra, S., & Mayaluri, Z. L. (2026). Business intelligence, AI-driven personalization, and ethical customer segmentation in e-commerce. *SN Business & Economics*.
6. Akter, S., & Wamba, S. F. (2016). Big data analytics in E-commerce: A systematic review and agenda for future research. *Electronic Markets*, 26, 173–194.