

System for Massive Datasets CSC 502

Customer Segmentation using PySpark

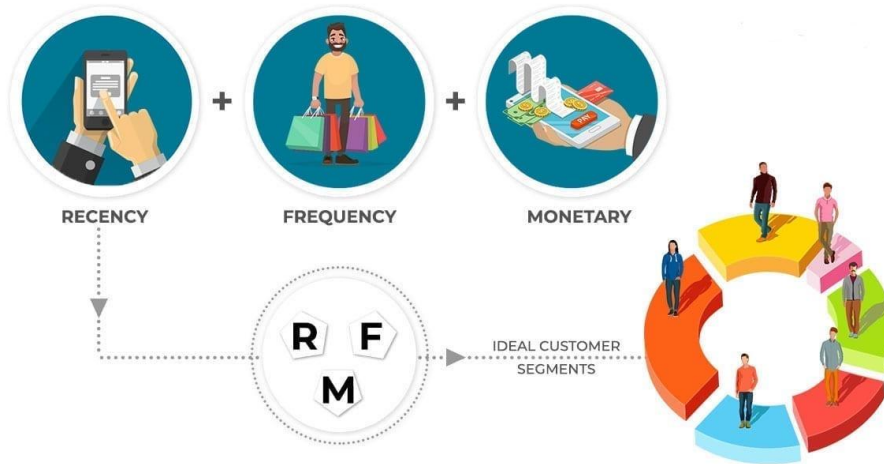
Jharna Kumari - V01046956






RFM SEGMENTATION

There are at least two main goals of customer segmentation:

1. Continue to provide the best service to your best customers.
2. Focus on prospective customers who resemble your best customers.



RFM Metrics		
 RECENCY The freshness of the customer activity, be it purchases or visits E.g. Time since last order or last engaged with the product	 FREQUENCY The frequency of the customer transactions or visits E.g. Total number of transactions or average time between transactions/engaged visits	 MONETARY The intention of customer to spend or purchasing power of customer E.g. Total or average transactions value

Methodology

Data Preparation

Load Data: PySpark for CSV handling with schema inference.

Clean Data:

- Remove null Customer IDs.
- Filter out canceled transactions.
- Correct data types for Price and InvoiceDate.

RFM Calculation

Recency: Days since last purchase, calculated from the max transaction date.

Frequency: Count of unique transactions per customer.

Monetary Value: Sum of all transaction values per customer.

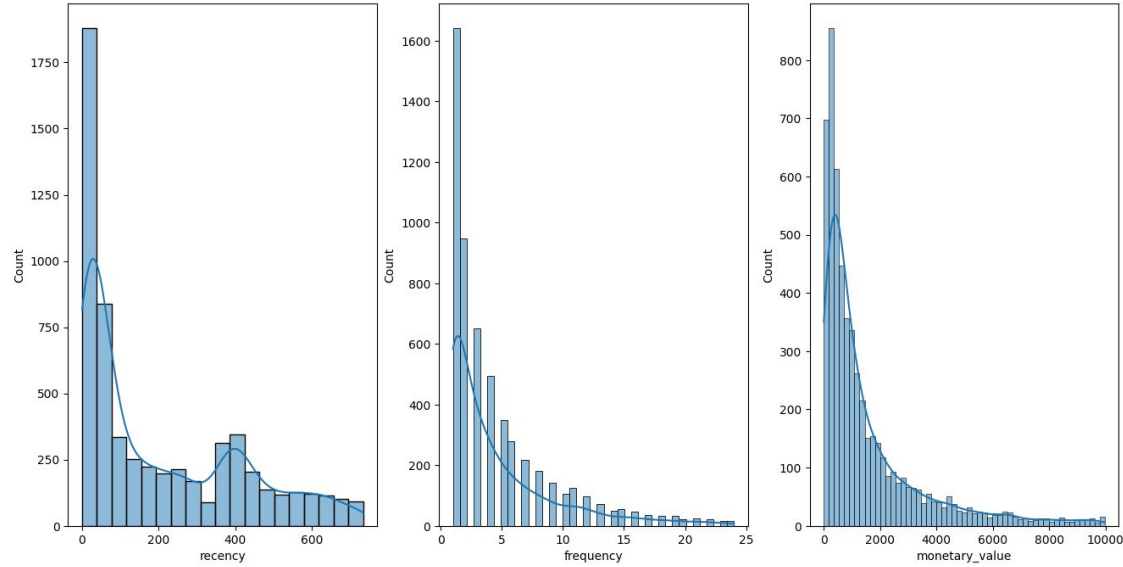


Methodology

RFM Segmentation

Score Assignment:

- Quantiles for R, F, M metrics.
- Lower scores for better recency and higher monetary value.
- Frequency scoring inverted.

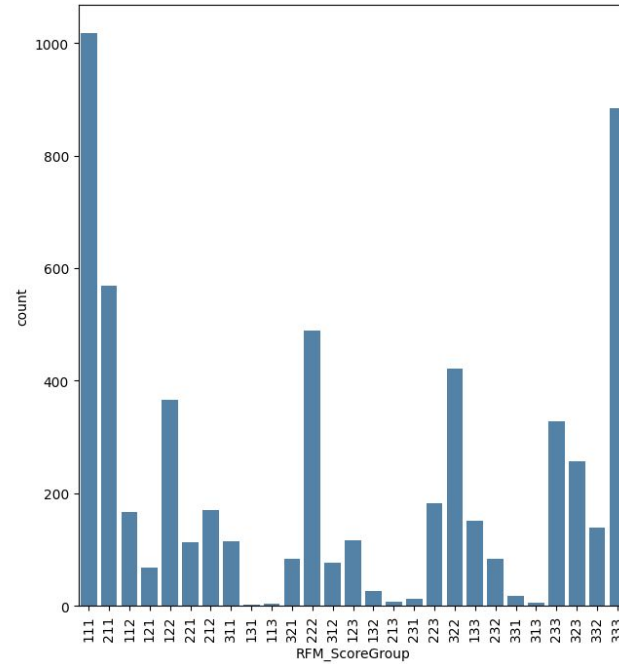
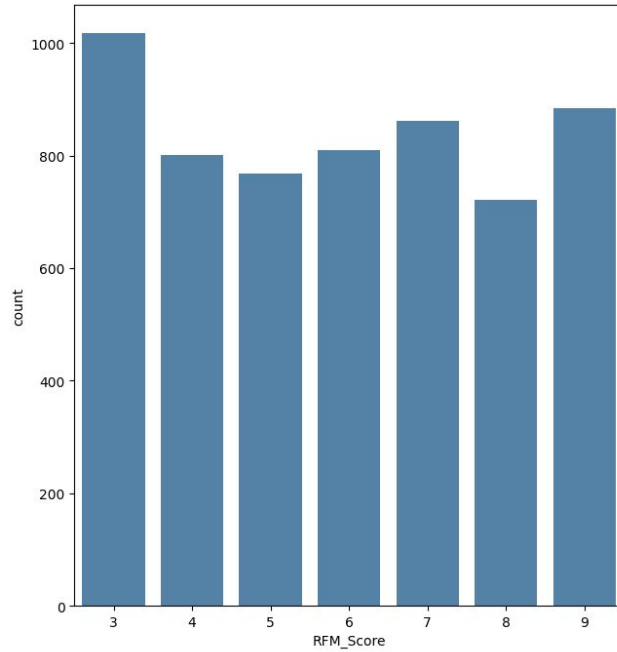


Composite Score: Aggregate R, F, M into a single RFM score.

Loyalty Segments: Assign Silver, Gold, Platinum based on RFM score.

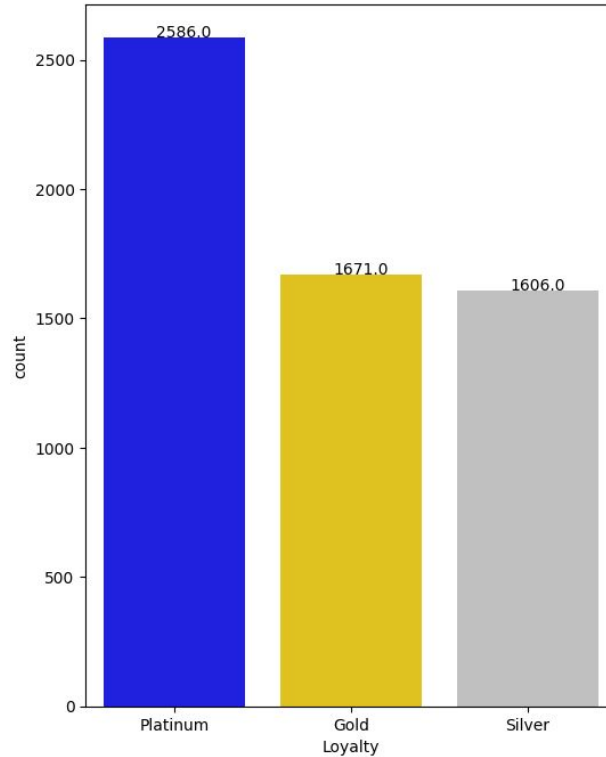
Visualizations

RFM Scores and Segmentation

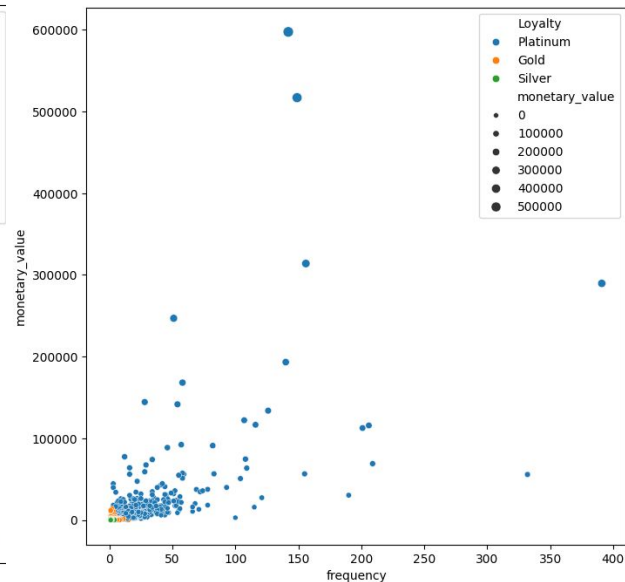
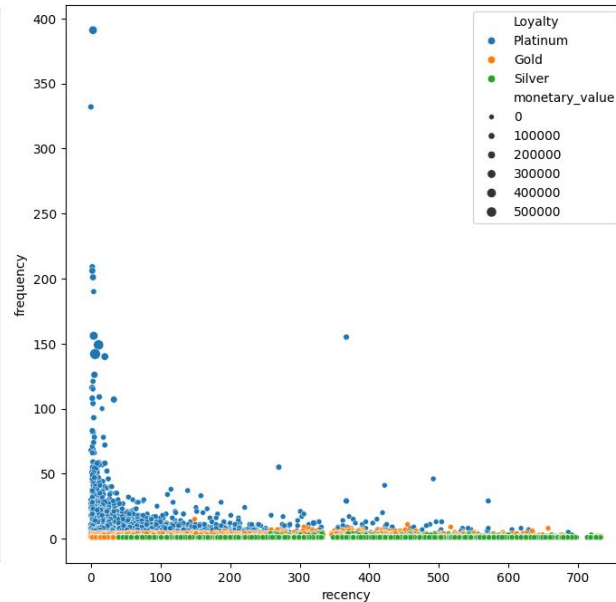
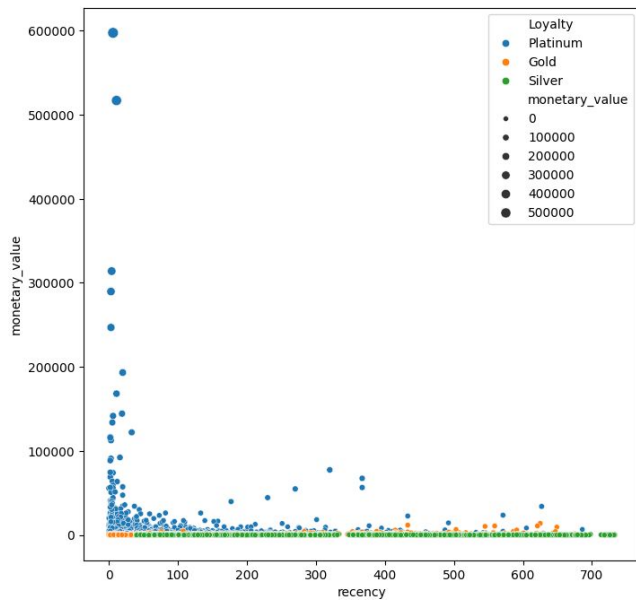


Visualizations

Loyalty level distribution

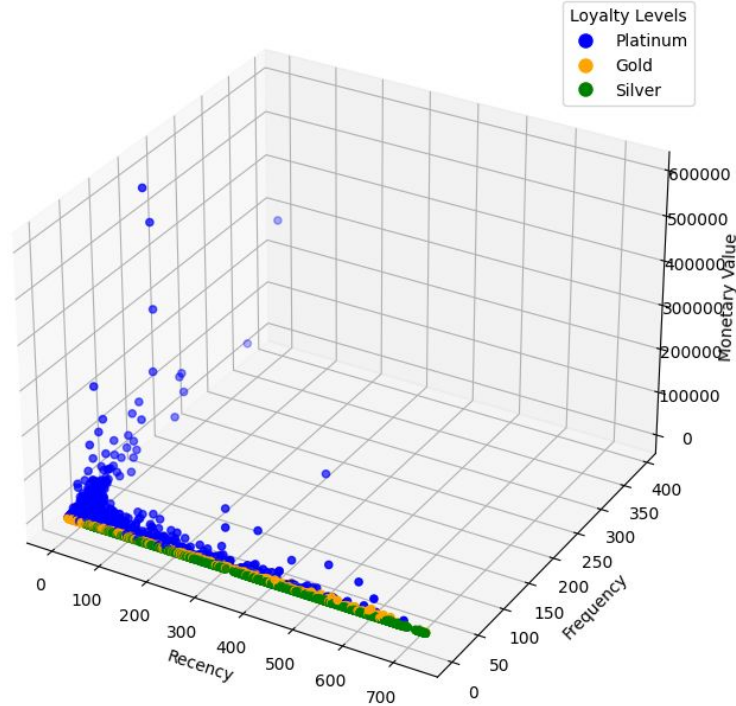


Loyalty level distribution



Loyalty level distribution

3D Scatter Plot by Recency, Frequency, and Monetary Value



Strategic Implications

Segment-Specific Strategies: Personalized Marketing

Platinum Customers

Exclusive Offers:

- Deploy premium strategies to reward loyalty.
- Encourage repeat business and maintain high engagement.

Gold & Silver Customers

Engagement & Upgrading:

- Implement initiatives aimed at increasing transaction frequency and monetary value.
- Offer incentives to move Silver to Gold, and Gold to Platinum

