

Recency, Frequency and Monetary

Introduction

RFM (Recency, Frequency, Monetary) is a customer segmentation technique used in marketing and data analytics. It evaluates customer behavior based on three key metrics:

1. **Recency (R):** How recently a customer made a purchase. Customers who bought more recently are more likely to return.
2. **Frequency (F):** How often a customer makes a purchase within a specific period. High-frequency customers are considered more loyal.
3. **Monetary (M):** The total amount a customer has spent. Higher spending customers are often more valuable.

How RFM is Used

- Segmentation: Grouping customers based on their RFM scores (e.g., VIP customers, dormant customers, churned customers).
- Marketing Campaigns: Sending targeted promotions based on customer behavior.
- Churn Prediction: Identifying customers at risk of leaving and taking action.
- Customer Lifetime Value (CLV): Estimating long-term profitability of customers.

Formula of RFM

RFM itself is not a single formula but rather a scoring model based on three key customer behavior metrics: **Recency (R)**, **Frequency (F)**, and **Monetary (M)**. Each metric is calculated separately and then combined for segmentation.

1. Recency (R) Calculation

Recency measures how recently a customer made a purchase.

$$R = \text{Current Date} - \text{Last Purchase Date}$$

- The smaller the value, the more recent the purchase (better for customer retention).
- Recency can be grouped into quantiles (e.g., scoring from 1 to 5).

2. Frequency (F) Calculation

Frequency counts how many purchases a customer has made in a given time period.

$$F = \text{Total Number of Purchases in the Given Period}$$

- Higher frequency indicates more engagement.
- Customers are ranked into quantiles (e.g., scoring from 1 to 5).

2. Monetary (M) Calculation

Monetary measures the total amount a customer has spent.

$$M = \sum(\text{Purchase Amount})$$

- The more a customer spends, the higher their value.
- Like Recency and Frequency, customers are grouped into quantiles (e.g., 1 to 5).

Final RFM Score

Each metric (R, F, M) is assigned a score (typically 1 to 5), and the final RFM score is a combination of these values:

$$RFM_Score = R_Score \times 100 + F_Score \times 10 + M_Score$$

- Example: A customer with **R = 5, F = 4, M = 3** would have an **RFM score of 543**.
- Higher scores generally indicate more valuable customers.

Assigning the Score RFM

The **RFM score** is assigned based on **quantiles (percentiles), ranking, or custom thresholds**. Typically, customers are **ranked** into **5 groups** for each metric (**Recency, Frequency, Monetary**) using **quantiles (1 to 5)**, with **5 being the best and 1 being the lowest**.

1. Assigning Scores for Recency (R)

- Customers who **spend more** get a **higher score (5)**.
- Lower spenders receive **lower scores (1)**.

Recency (Days Since Last Purchase)	M Score
0 - 30 days	5
31 - 60 days	4
61 - 90 days	3
91 - 120 days	2
121+ days	1

2. Assigning Scores for Frequency (F)

- Customers who purchase **more frequently** get a **higher score (5)**.
- Those with **fewer transactions** receive **lower scores (1)**.

Number of Purchases	F Score
10+	5
7-9	4
4-6	3
2-3	2
1	1

3. Assigning Scores for Monetary (M)

- Customers who **spend more** get a **higher score (5)**.
- Lower spenders receive **lower scores (1)**.

Total Spend (\$)	M Score
\$1000+	5
\$750 - \$999	4
\$500 - \$749	3
\$250 - \$499	2
< \$250	1

4. Calculating the Final RFM Score

Each customer gets a 3-digit **RFM Score** based on their **R, F, and M** values.

Customer	Recency (R)	Frequency (F)	Monetary (M)	RFM Score
A	5	2	5	525
B	2	3	5	235
C	4	4	3	443
D	3	5	2	352
E	5	2	1	521

5. Interpreting the RFM Score

Different RFM scores help in customer segmentation:

RFM Score	Customer Segment	Interpretation
555, 554, 545, etc.	Best Customers	Recent, frequent, high spenders
455, 454, 445, etc.	Loyal Customers	Buy often, but not the highest spenders
155, 144, 133, etc.	Churned Customers	Haven't purchased recently
511, 411, 311, etc.	New Customers	First-time buyers, high potential
111, 112, 121, etc.	Lost Customers	Not engaged, low frequency & spend

Implementing Code

1. SQL SERVER CODE

Using AdventureWorks databases

```

1. WITH RFM AS (
2.     -- Calculate Recency, Frequency, and Monetary per Customer
3.     SELECT
4.         soh.CustomerID,
5.         DATEDIFF(DAY, MAX(soh.OrderDate), GETDATE()) AS Recency,
6.         COUNT(soh.SalesOrderID) AS Frequency,
7.         SUM(sod.LineTotal) AS Monetary
8.     FROM Sales.SalesOrderHeader soh
9.     JOIN Sales.SalesOrderDetail sod ON soh.SalesOrderID = sod.SalesOrderID
10.    WHERE soh.CustomerID IS NOT NULL
11.    GROUP BY soh.CustomerID
12. ),
13. RFM_Scored AS (
14.     -- Assign scores using NTILE(5) for segmentation (1 = low, 5 = high)
15.     SELECT
16.         CustomerID,
17.         Recency,
18.         NTILE(5) OVER (ORDER BY Recency ASC) AS R_Score, -- Lower recency = higher score
19.         Frequency,
20.         NTILE(5) OVER (ORDER BY Frequency DESC) AS F_Score, -- Higher frequency = higher score
21.         Monetary,
22.         NTILE(5) OVER (ORDER BY Monetary DESC) AS M_Score -- Higher monetary = higher score
23.     FROM RFM
24. )
25. SELECT
26.     CustomerID,
27.     R_Score, F_Score, M_Score,
28.     CAST(R_Score AS VARCHAR) + CAST(F_Score AS VARCHAR) + CAST(M_Score AS VARCHAR) AS RFM_Score
29. FROM RFM_Scored
30. ORDER BY RFM_Score DESC;
31.

```

2. PYTHON CODE

Using AdventureWorks Database

```

import pandas as pd
import pyodbc

conn = pyodbc.connect(
    'DRIVER={SQL Server};'
    'SERVER=YourServerName;'
    'DATABASE=AdventureWorks;'
    'Trusted_Connection=yes;'
)

query = """
WITH RFM AS (
    SELECT
        soh.CustomerID,
        DATEDIFF(DAY, MAX(soh.OrderDate), GETDATE()) AS Recency,
        COUNT(soh.SalesOrderID) AS Frequency,
        SUM(sod.LineTotal) AS Monetary
    FROM Sales.SalesOrderHeader soh

```

```

        JOIN Sales.SalesOrderDetail sod ON soh.SalesOrderID = sod.SalesOrderID
        WHERE soh.CustomerID IS NOT NULL
        GROUP BY soh.CustomerID
    ),
    RFM_Scored AS (
        SELECT
            CustomerID,
            Recency,
            NTILE(5) OVER (ORDER BY Recency ASC) AS R_Score,
            Frequency,
            NTILE(5) OVER (ORDER BY Frequency DESC) AS F_Score,
            Monetary,
            NTILE(5) OVER (ORDER BY Monetary DESC) AS M_Score
        FROM RFM
    )
    SELECT
        CustomerID,
        R_Score, F_Score, M_Score,
        CAST(R_Score AS VARCHAR) + CAST(F_Score AS VARCHAR) + CAST(M_Score AS VARCHAR) AS RFM_Score
    FROM RFM_Scored
    ORDER BY RFM_Score DESC;
"""

rfm_df = pd.read_sql(query, conn)
print(rfm_df.head())

```