RFM ANALYSIS

When it comes to finding out who your best customers are, the old RFM matrix principle is the best. RFM stands for Recency, Frequency, and Monetary. It is a customer segmentation technique that uses past purchase behavior to divide customers into groups.

### RFM Score Calculations

RECENCY (R): Days since last purchase  
FREQUENCY (F): Total number of purchases  
MONETARY VALUE (M): Total money this customer spent

### The Process

Step 1: Calculate the RFM metrics for each customer.  
Step 2: Add segment numbers to RFM table.  
Step 3: Sort according to the RFM scores from the best customers (score 111).

Since RFM is based on user activity data, the first thing we need is data.

## Data

The dataset we will use is the same as when we did [Market Basket Analysis](https://datascienceplus.com/a-gentle-introduction-on-market-basket-analysis%E2%80%8A-%E2%80%8Aassociation-rules/) — Online retail data set that can be downloaded from [UCI Machine Learning Repository](https://archive.ics.uci.edu/ml/datasets/online+retail).

The dataset contains all the transactions occurring between 01/12/2010 and 09/12/2011 for a UK-based and registered online retailer.

Solution:

### Steps: Explore the data — validation and new variables

1. Missing values in important columns;  
2. Customers’ distribution in each country;  
3. Unit price and Quantity should > 0;  
4. Invoice date should < today.

## RFM Customer Segmentation

**Create a RFM table and Calculate RFM metrics for each customer**

### Split the metrics

### Create a segmented RFM table

### Add segment numbers to the newly created segmented RFM table