

Behavioral customer insights: RFM clustering based on transaction data

B2B wholesale company wants to understand the behavior of its customer base. Recency, Frequency, Monetary value framework leverages existing customer transaction data to arrive at segmentation based on actual behavior.

Challenge

The client is a B2B wholesale company, which has **recently integrated its customer database** to arrive at a harmonized transaction data overview across different systems.¹ Now the client wants to utilize the data to **understand basic customer behavior** and how customers can be clustered based on their transaction history. This should serve to get a **better prioritization of the customer base** for sales steering and marketing campaigns.



Solution

1 Implement RFM based customer segmentation

RFM stands for Recency, Frequency, Monetary value and is a practitioners' concept that has been used in a basic form over the last two decades:

- **Recency:** When did the customer buy the last time?
- **Frequency:** How often did a customer buy over a specific period?
- **Monetary value:** For what value/margin did the customer buy?

Simple RFM scoring of the customer base can be done by partitioning the data. We employ a continuous transformation of the data combined with k-means clustering:

01. Calculate RFM metrics on a customer level: days since last purchase, number of purchases over last year, total monetary value of purchases.
02. Transform the data to facilitate segmentation by reducing outlier values.
03. Employ k-means clustering algorithm building on business logic and experience.

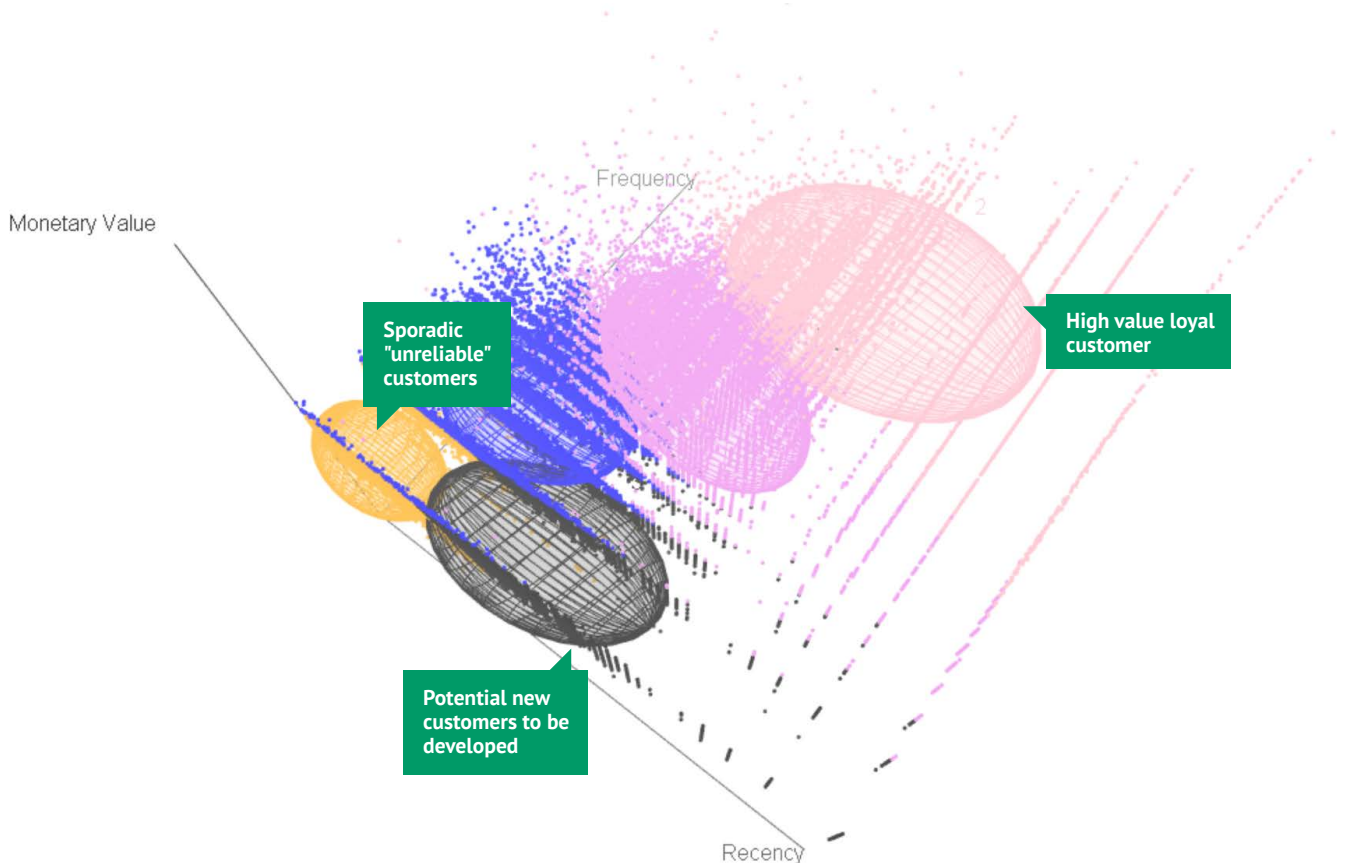
2 Visualize RFM clustering to use for daily sales steering

The visualization of the clustering result helps to communicate the logic and to draw attention to this concept. It is key to explain the segmentation to both marketing and sales departments in order to facilitate implementation into daily operation and decision making.

Insights:

Number of clusters **need to make sense from a business perspective**. For example, three clusters do not allow targeted steering, 20 clusters will likely result in confusion and lose the value of practicability. With our client, **we arrived at five clusters as an ideal trade-off**. In the discussion it helps to **identify typical well-known customers from each segment** and to discuss why they have been assigned and how similar they are to other customers.

¹ See case study: "Where CRM is missing: towards an integrated view of customer data"



Results

With the RFM segmentation, the client was able to benefit on multiple dimensions:

- **Facilitate improved sales steering and marketing campaigns:**
Which customers can we develop? Who has changed behavior?
- **Learn about the customer base:**
Who is rarely buying? Which percentage of customers are only buying arbitrarily?
- **Cross-check with existing customer prioritization:**
If these are our 250 prioritized customers with individual service so far, which customers are similar in behavior and should also belong to this group?

Example:

One cluster was identified as "development customers", who buy with a medium frequency but only at low values. It was shown for these customers, our client is not the preferred B2B wholesaler. These customers prefer competitors' offerings. Using a targeted steering of the sales force, these customers can now be addressed and some of the potential seized.