Customer segmentation using RFM technique

(Recency, Frequency, and Monetary)

* UserId - Unique identifier of a user.
* TransactionId - Unique identifier of a transaction. If the same TransactionId is present in multiple rows, then all those products are bought together in the same transaction.
* TransactionTime - Time at which the transaction is performed
* ItemCode - Unique identifier of the product purchased
* ItemDescription - Simple description of the product purchased
* NumberOfItemsPurchased - Quantity of the product bought in the transaction
* CostPerItem - Price per each unit of the product
* Country - Country from which the purchase is made.

Using RFM technique I have separated customers. Objective is to find customers with recent purchase records (low recency), frequent purchase record (high frequency) and have spent good money (monetary is high). These customers likely to be more important than customers with low frequency and low monetary. There are also those customers in the middle, with moderately high frequency and monetary, but also very high recency. Perhaps these customers bought something in the past which is no longer offered or available.

Step 1 – Exploratory Data analysis. In this step the data set is explored and all us useful changes for analysis are done. For example, checking for Na values, changing the format of TransactionTime column to Date format. Removing negative and suspicious values from the dataset.

Step 2- Creating the a rfm data set which has columns Recency, Frequency and Momentary. Recency: how recent the customer has bought a product. Frequency: how often the customer is buying and Momentary: sum of all the purchases i.e. how much the customer is spending.

Step 3- after creating the rfm dataset we segregate the customers into those who have brought once and more than once. Then we create levels from 1 to 5 for the RFM columns. after that we give scores to the customers according to that.

Step 4- we segregate customers into 'Inactive', 'Temporary idle’,’ Frequent' and 'Highly Active' based on their recency values.

Step 5- we segregate customers into potential. Returning, regular and stellar customers based on their frequency scores

Step 6 – we segregate customers into Best, top75%, top25% and worst based on their momentary scores

The customers are segmented based on their transitional data for making better business decisions. By this analysis we know what type of customers to be approached and other financial decisions can be made.