Case - Segmenting Retail Customers using the RFM Model of Customer Behavior

Situation - Recency, frequency and monetary (RFM) value of purchase are classic marketing variables that are used often in what is called 'RFM Customer Segmentation.' RFM segmentation allows for a simple and intuitive way for marketers to target specific groups of customers with marketing interventions that match the characteristics of the *typical* user in the group.

Complication - Is there a data driven way to segment customers using the RFM variables that we have, as opposed to using managerial judgement/intuition to decide the levels of the RFM cutoffs to make high-value, medium-value and low-value segments? How would these work with more than two dimensions. In our case we have recency, frequency, monetary value of purchases, as well as customer tenure with the company. In other words we have 4 dimensions. How do we take all 4 dimensions into account to group similar customers in one segment/cluster and make sure that the segments themselves are dissimilar from each other.

Key question - How many segments should we have? How do we know which customer belongs to which segment?

Data: We have 2000 rows of following type of data in CA-customerData.csv

	Monetary	Recency	Frequency	tenure	ID#
	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
1	138	28	3	40	2
2	240	14	1	14	30
3	97	6	2	10	59
4	348	2	7	38	89
5	239	20	2	28	96
6	253	10	4	20	120

Solution approach - Deploy k-means clustering algorithm to create three cluster which should correspond to some interesting (perhaps high, medium low value) segments of customers.

Code - customerSegmentation-classVersion.Rmd