**DATA MINING PROJECT**

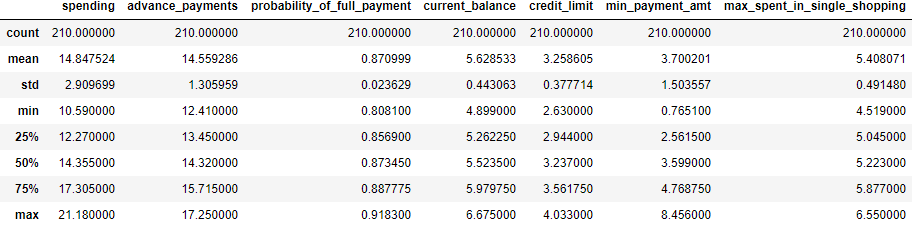
**December 2019 Batch**

**Bank Marketing**

**Q1. Read the data and do exploratory data analysis. Describe the data briefly.**

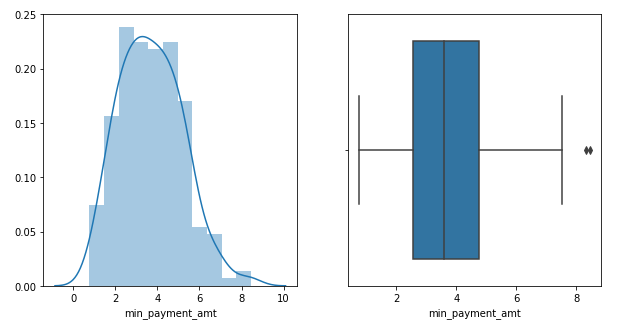
**Observation:**

**Data consist of total 7 variables-** spending, advance payments, probability of full payment, current balance, credit limit, min payment amount, and max spent in single shopping



No null values present in the data and data is equally distributed across each column.

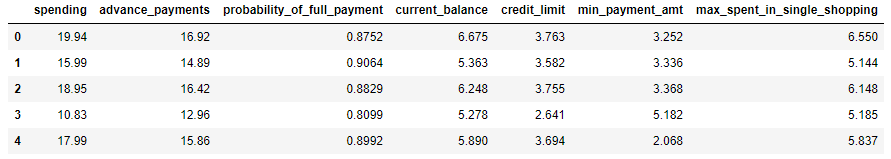
There is some outlier in the field min payment amount as we can see below



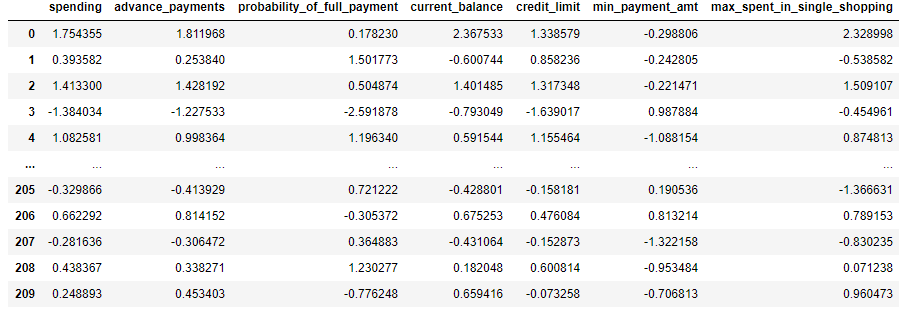
**Q2. Do you think scaling is necessary for clustering in this case? Justify.**

**Observation:**

**Scaling is must. There are some outliers present in some fields and also high values may cause our model bias towards high values of the variables.**



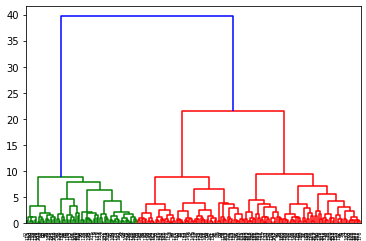
**As we can see in above 5 rows the value is higher for variable spending and advance payments as compared with other variables, this can cause our model bias towards the higher values.**



**Above figure shows the data after applying scaling. This can help model to measure distance without any bias.**

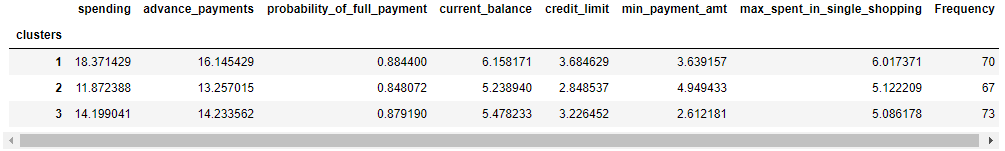
**Q3 Apply hierarchical clustering to scaled data. Identify the number of optimum clusters using Dendrogram and briefly describe them**

**Observation:**



As we can see above figure dendrogram have many of the clusters. But the chosen number of cluster is three because we have a small dataset with only 210 rows.

**Apply hierarchical clustering to scaled data**



As per results using hierarchical clustering we have form total 3 clusters for our values. And the observations are:

1st cluster- High spending and advance payments and the probability of full payment is high. Current balance and credit limit is also high. Max spent in single shopping is also high for this customers.

2nd cluster- less spending and advance payments and the probability of full payment is also less. But minimum payment amount is higher for this type of customer.

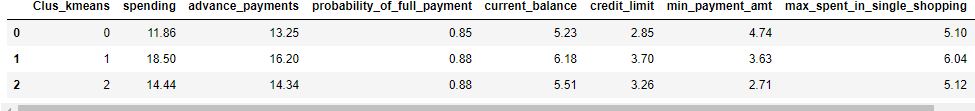
3rd cluster- low spending and advance payment, but the probability of full payment is good. Min payment amount is very less for this type of customers.

**Q 4 Apply K-Means clustering on scaled data and determine optimum clusters. Apply elbow curve and determine the number of clusters.**

**Observation:**

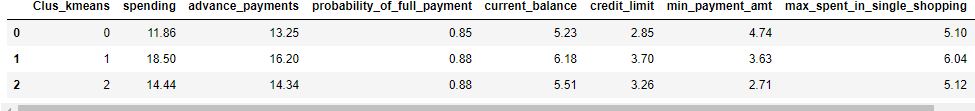


As we can see in above elbow plot inertia changed after cluster 3 i.e. variance changed sharply after cluster 3. So the number of cluster chosen for K means is 3.



Above fig shows the clusters formed after applying K-means to the scaled dataset and given number of cluster is 3.

**Q5 Describe cluster profiles for the clusters defined. Recommend different promotional strategies for different clusters.**



1st cluster (0) – spending, advance payment and current balance is low for this type of customer. But the minimum payment amount is high and maximum spent in single shopping is less.

2nd cluster (1) - spending, advance payment and current balance is very high for this type of customer. But the minimum payment amount is high and maximum spent in single shopping is very high.

3nd cluster (2) - spending, advance payment and current balance is high for this type of customer. But the minimum payment amount is low and maximum spent in single shopping is high.

**Recommend different promotional strategies for different clusters.**

**For customers in cluster 1 shows that they prefer for less spending and their credit limit is also less. So the promotional strategies applied to this type of customers is increasing their credit card limit.**

**For customers in cluster 2 shows that they prefer for very high spending and their credit limit is also high. We don’t let go this type of customers. So the promotional strategies applied to this type of customers is increase their card limits giving them emi offers and other special promotional strategy should be applied on this type of customers.**

**For customers in cluster 3 shows that they prefer for high spending and their credit limit is also not low and also no high. So the promotional strategies applied to this type of customers is giving cash back offers and also EMI offers.**