

Description HW2

From PCA analyse, for 10 components, we can see that most variance lie in two components (69 in 10 and 16 in 9). So we can see what features have the biggest impact in these components, and it is CASH_ADVANCE_TRX for PC10 and INSTALLMENTS_PURCHASES and ONEOFF_PURCHASES for PC9. We can also group the data by clusters (we've used 4 cluster KMean) and see means values in clusters. All this data can help us to interpreted results.

From the results we can see that cluster 2 is mainly made up of conservative customers, they use cash more often and have generally large balances, although they do not often buy expensive goods or take them on instalments. Savings accounts would be more interesting for these clients. The third cluster of clients also has large balances and clients also often use cash, but in this cluster people buy quite expensive items and take large instalment payments. It can be assumed that this is a fairly loyal class of clients that like using credit cards. Low-cost customer retention programmes can be used for them. Class 1 is the less affluent clients, but who keep their spending in their budget. They don't have very expensive purchases, they don't use cash as often, and they take instalments. They should be targeted with programmes for lower price segment. And the last 0 class is characterised by a very frequent use of the card with large purchases and small balances. It can be seen that these people tend to spend a lot and are prone to impulse purchases. You should be careful with such clients as they can easily have problems with paying their bills, you should keep a close eye on them and set their limit carefully.