

# CLUSTERING

## Goal:

## Cluster crypto-currencies based on their daily behaviour to help traders.

## Data Preparation:

For analysis of behaviour of currencies, since high and low are extremes in the day and they are not representative, I have been left with choice of opening or closing. And I choose closing price.

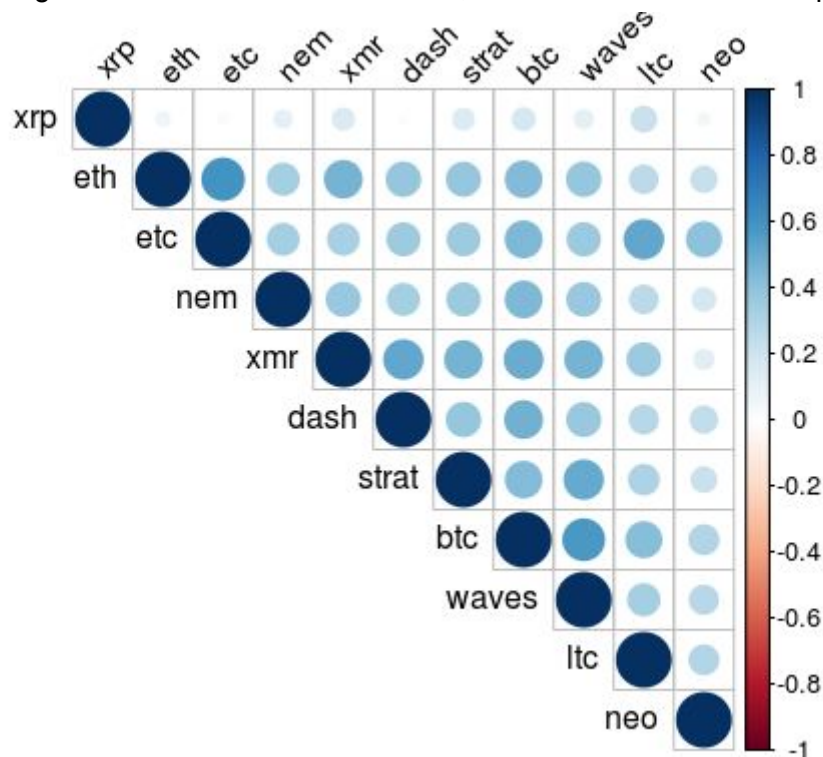
Out of all currencies we have, I have only considered currencies which are present since 1st Jan 2017. This had left me with 11 currencies: btc, eth, xrp, waves, strat, nem, ltc, xmr, etc, neo, dash.

The closing price itself does not convey any insights, what we are actually interested in is increase or decrease for that day. So I take consecutive difference of closing price.

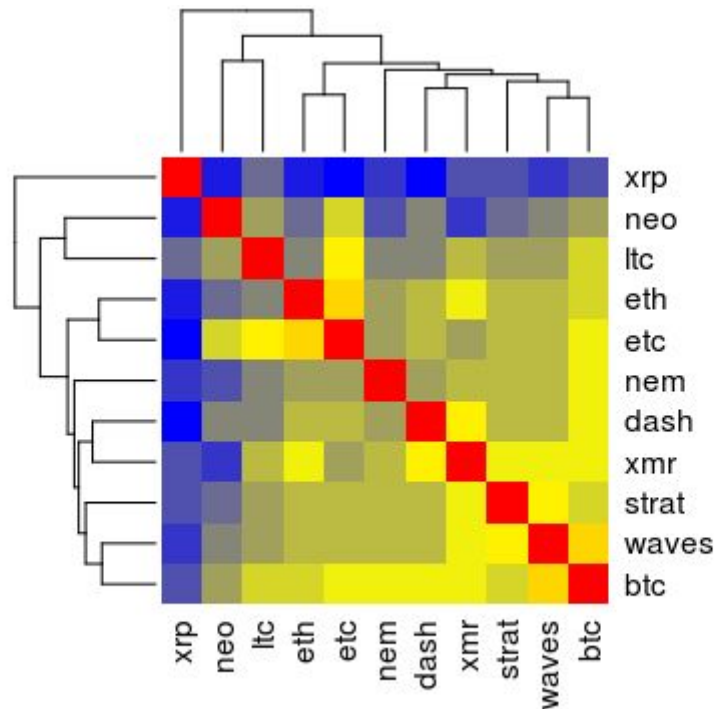
As the prices of currencies are very different, the change in closing price are in different scales. I performed scaling.

## Clustering Approaches:

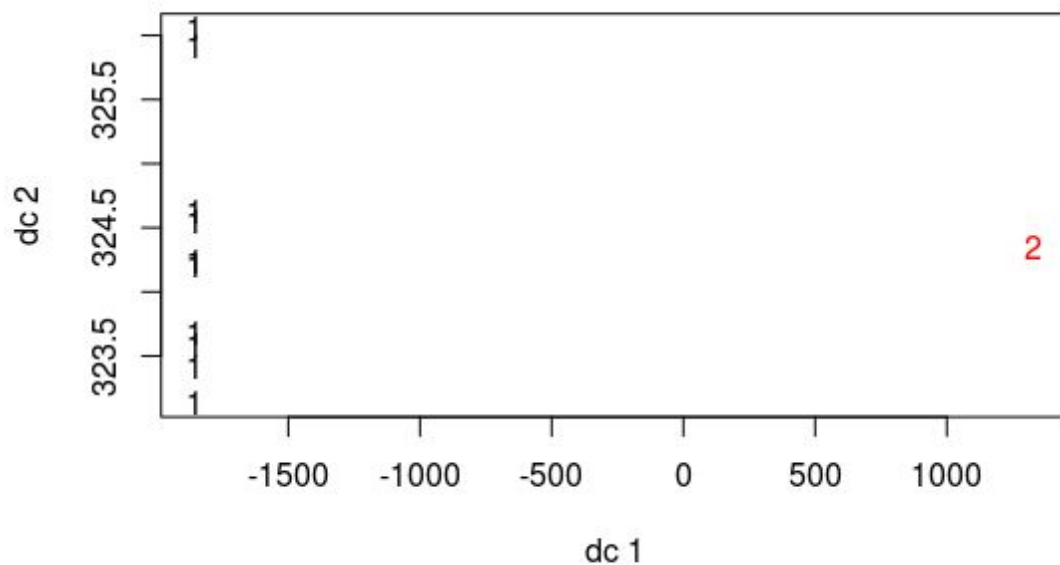
After computing Pearson Correlation coefficient, I have made a correlation plot.



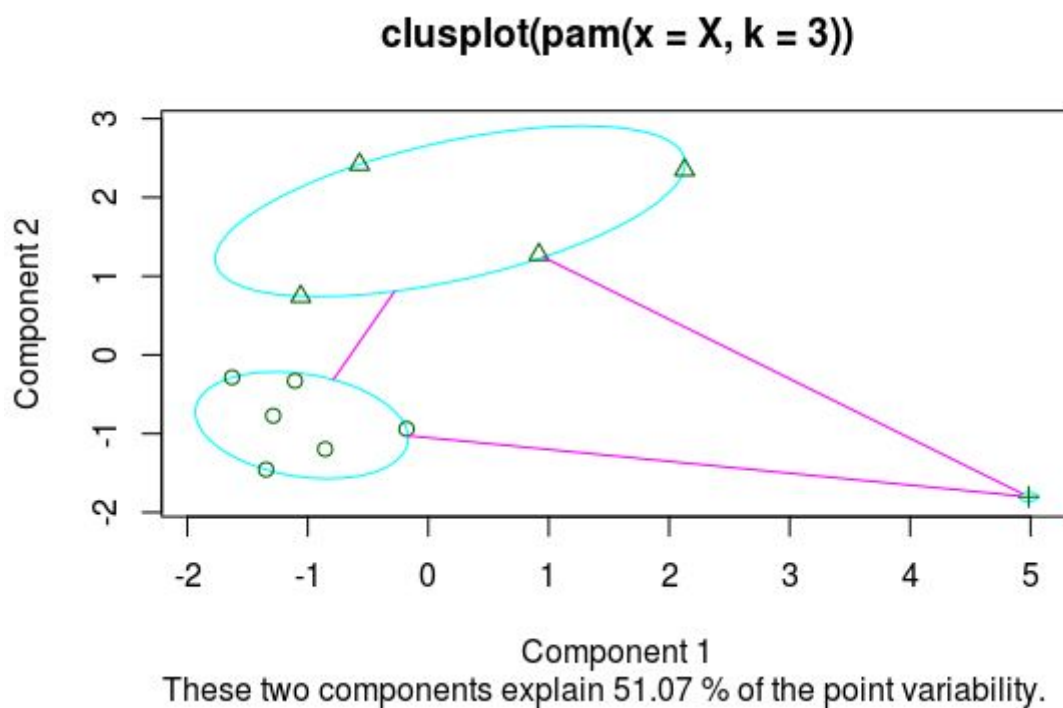
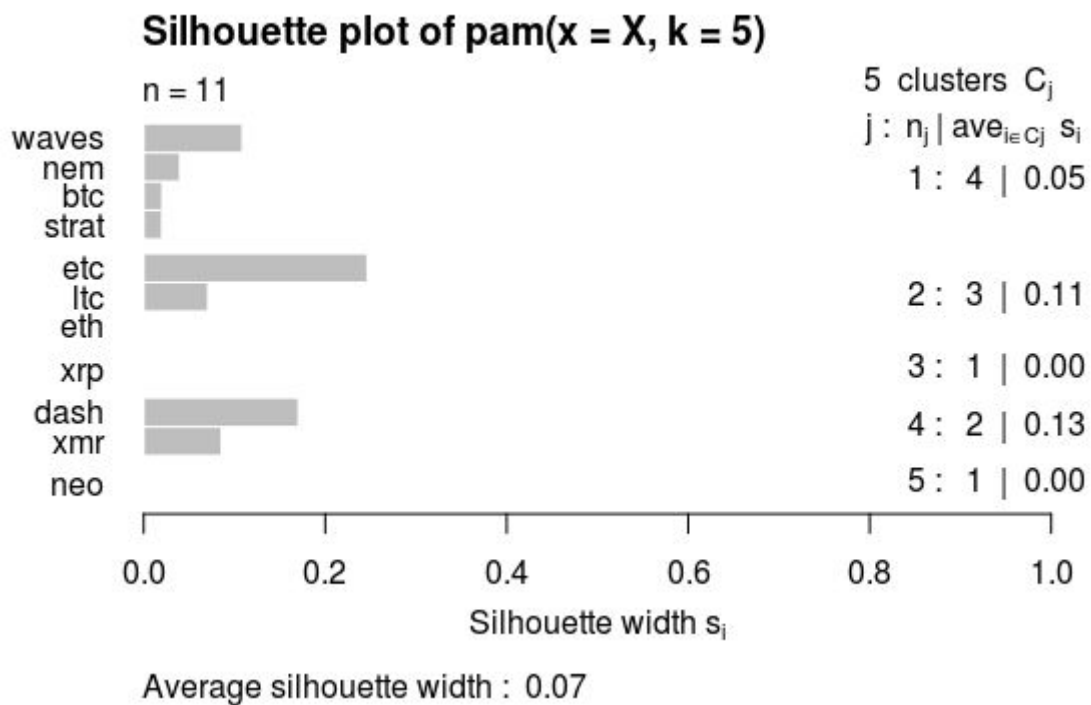
I have also made a heatmap with hierarchical clustering using Pearson correlation coefficient. Color palette: Blue - Yellow- Red

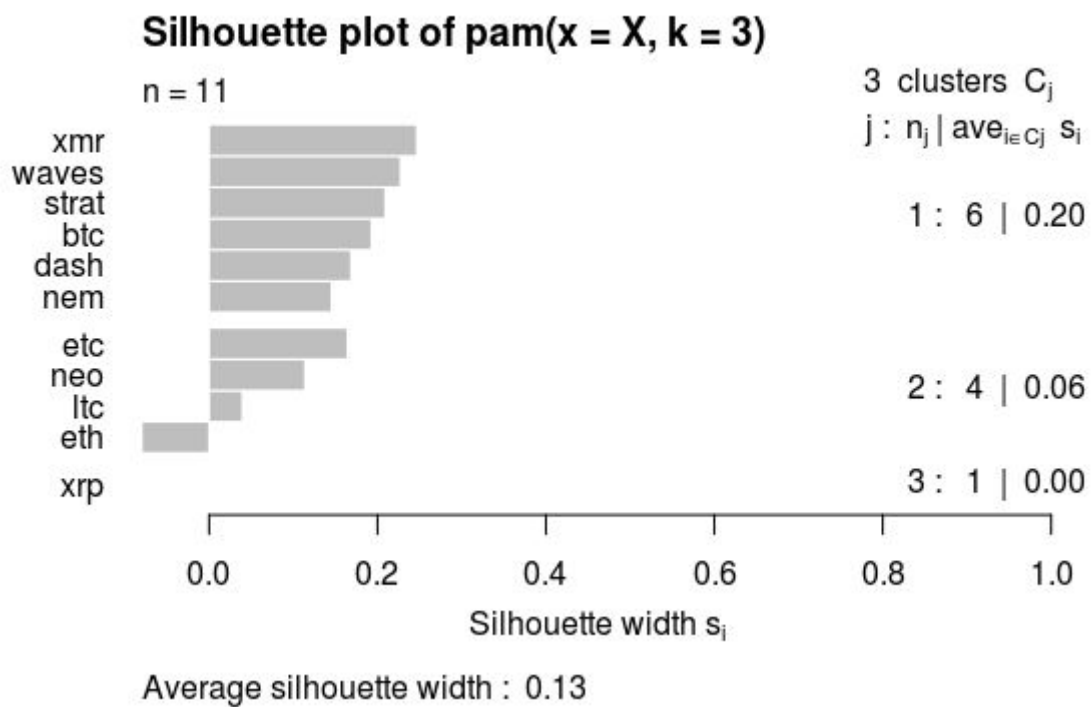


K- Means clustering with k=2 helps in identifying XRP which had no correlation with other currencies.

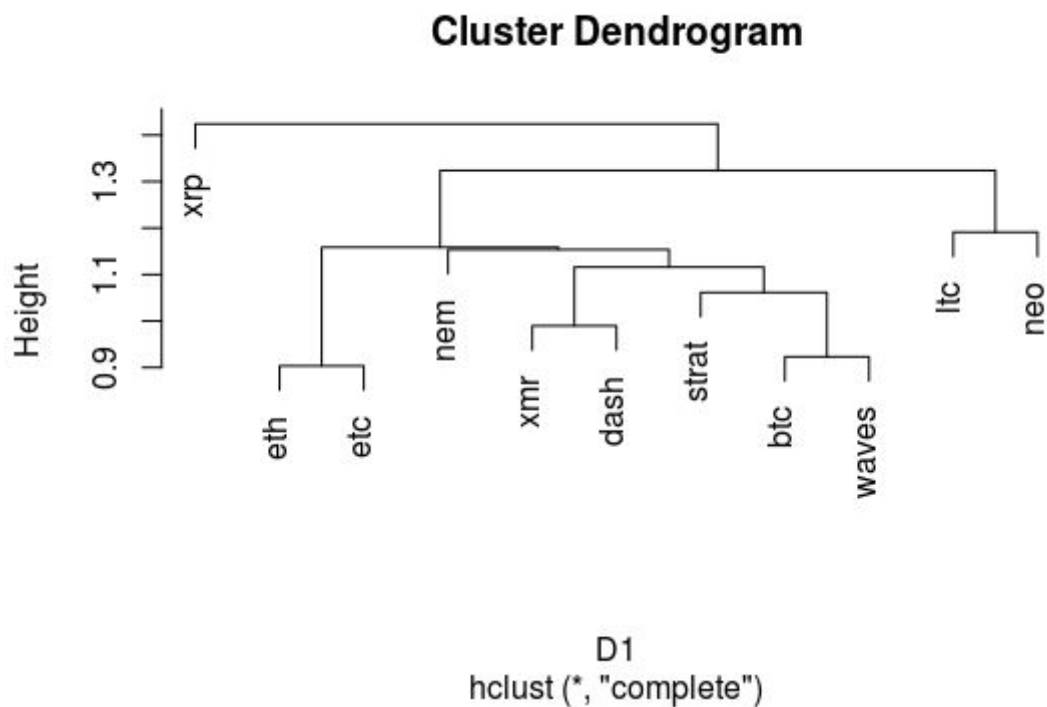


K-Medoids PAM clustering is used to identify the clusters of currencies based on Pearson Correlation coefficient.





Hierarchical clustering with correlation measure of Directional movement.



Clustering by Frechets distance

