Color: Class

Color: Cluster

Color: Cluster
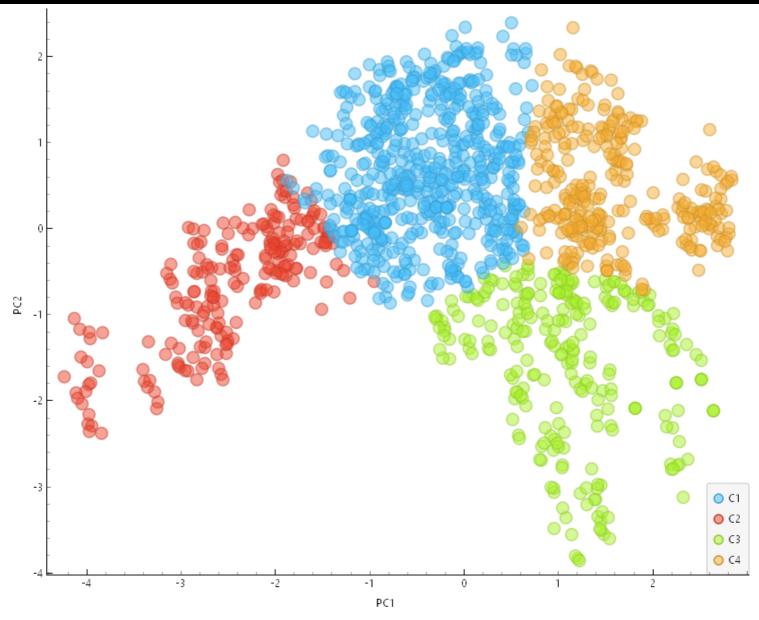
2. The clusters with k=4 have less variance between clusters but that is what you would expect with more clusters. I think that the k=4 is the better clustering solution because it has more options where it can get a better result.

3. Both models predicted that there was most likely a left right split when in reality there is more of a top down split. This doesn't surprise me as I myself would have made a similar assumption.

4. We probably use PCA to reduce dimensions to 2 dimensions because anything more is hard to visualise.