Use wine data set from sklearn library and try to form clusters of wine behaviour using ***Malic Acid*** and ***Proline*** features. Drop the other features for simplicity.

-Create a scatter plot of the above mentioned features of the wine data set.

-Figure out if any pre-processing such as scaling would here.

-Draw elbow plot and from that figure out an optimal value of k.