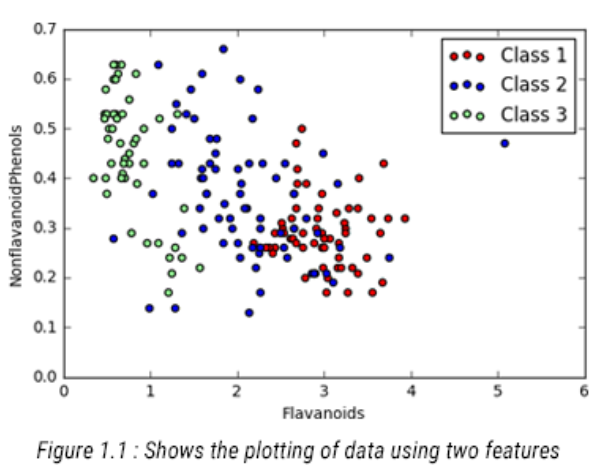
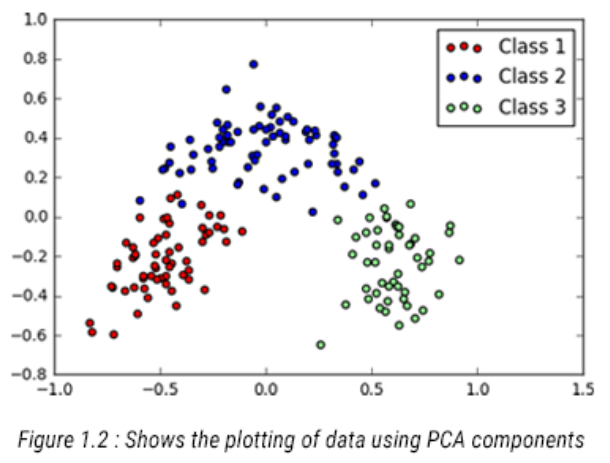
**PCA( Principle Component Analysis )**

Principal Component Analysis (PCA) is an **unsupervised learning algorithm** as it ignores the class labels (the so-called principal components) that maximize the variance in a dataset, to find the directions.

For example, To obtain quality/type of wine, we can use different features of wine such as its pH value, alcohol content, colour of the wine, acidity content and so on, however, many of these features will be redundant or dummy feature variables (can be derived from other features), therefore causing to train model on unnecessary features.

In short, we can get the type of wine with less feature variables and this is what actually PCA does inside the box.





The graph above shows the plotting of data points using two features vs. Plotting of data points using principal components (which is basically summarization of all the features in our dataset) i.e. PCA1 and PCA2. This helps us to identify one more use of PCA i.e. proper visualization of our dataset along with their class tags.  
If you look at the graph, you can visualize the whole dataset properly with only two feature variables as compared to having thirteen feature variables. Though they do not differentiate well among the classes but helps in reducing the feature set with minimal loss of information.