> Data<-read.csv('wineQuality.csv',header=TRUE,sep=";")

> apply(Data,2,var) #Computing variances

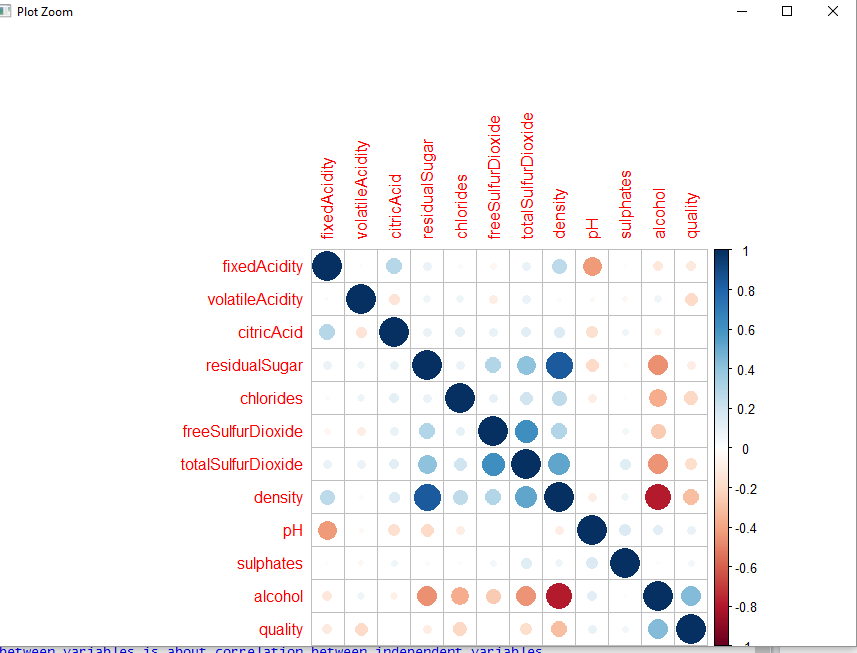
> a<-apply(Data,2,var) #Computing variances

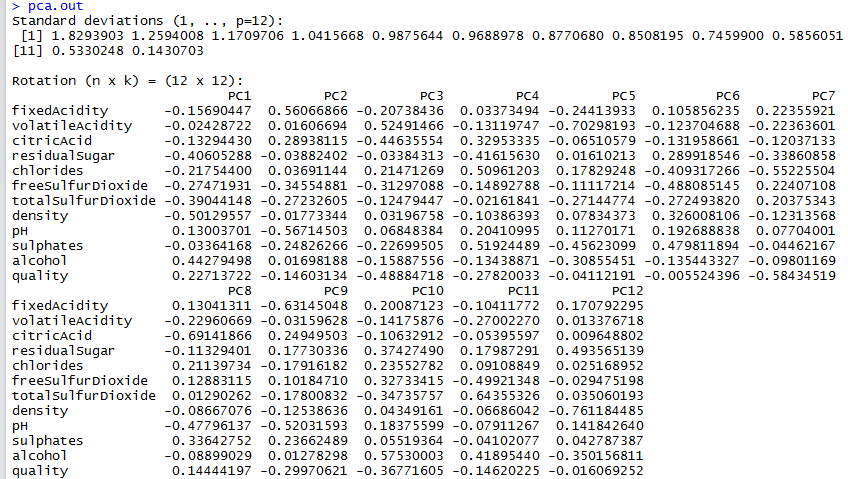
> max(a)

[1] 1806.085

> #totalSulfurDioxide has the largest variance, so Data should be standardized

> pca.out=prcomp(Data, scale=TRUE)





Interpretation of the principal components is based on finding which variables are most strongly correlated with each component, I mean, which of these numbers(PC values) are large in magnitude, the farthest from zero in either positive or negative direction.

I determined the correlation value above 0.5 as important

I observed that the first principal component -PC1 correlates most strongly with the Density, based on correlation of 0.50129557

PC2- correlates most strongly with FixedAcidity(0.5606) and pH(0.5671)

PC3- correlates most strongly with VolatileAcidity(0.5249)

PC4- correlates most strongly with chlorides(0.5096) and sulphates(0.5192)

PC5- correlates most strongly with VolatileAcidity(0.7029)

PC6- correlates most strongly with none of variables

PC7- correlates most strongly with Chlorides(0.5522) and quality(0.5843)

PC8- correlates most strongly with citricAcid(0.6914)

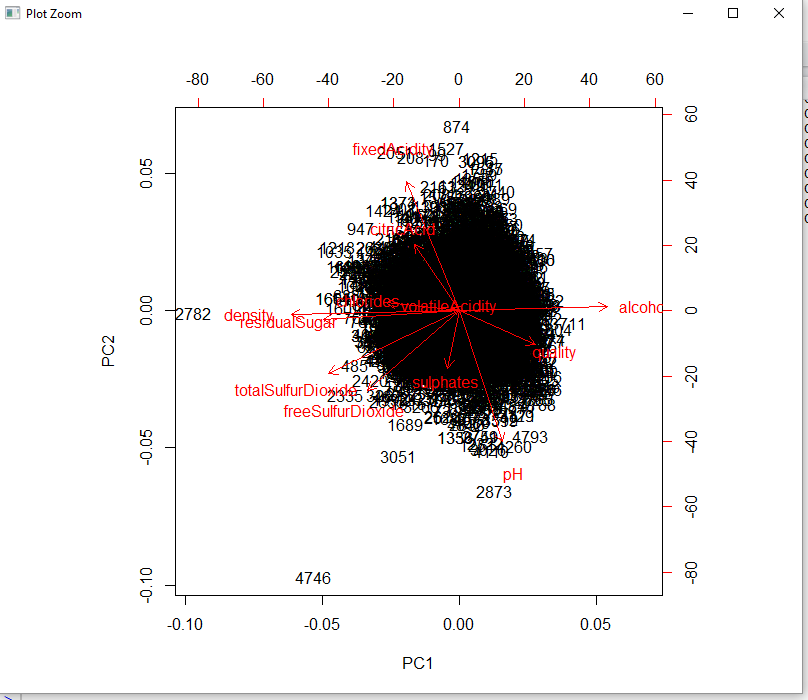
PC9- correlates most strongly with fixedAcidity(0.6314)

PC10- correlates most strongly with alcohol(0.5753)

PC11- correlates most strongly with totalSulfurDioxide(0.6435)

PC12- correlates most strongly with density(0.7611)

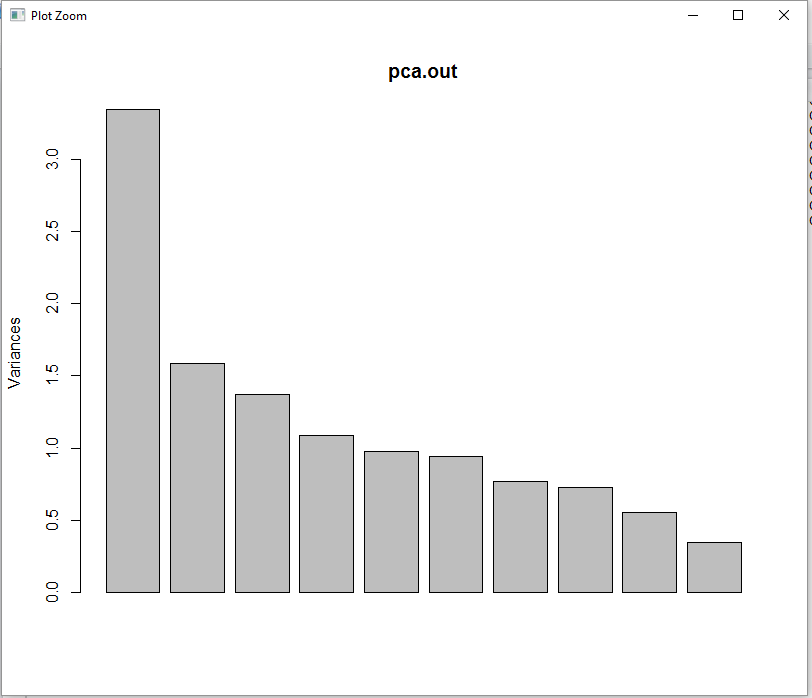
Biplot



The left and bottom axes are showing normalized principal component scores; the top and right axes are showing the loadings

So I was looking at the red dot out by itself to the right(3711) I conclude that this particular dot has a very high value for the first principal component and this kind of wine is going to have high values for the Density. Whereas if I look at red dot at the left(2782), I expect to have low values for Density.The top red dot(874) has a high value for the second principal component. So I expect that this kind of wine would have low values for FixedAcidity and pH. And conversely when I look at the red dot on the bottom, the corresponding wine would have high values for FixedAcidity and pH.

Using eigenvalues > 1 is only one indication of how many factors(components) to retain.Since first 4 eigenvalues are greater than 1,PC1,PC2,PC3,PC4 are retained.Scree Plot will be drawn to define Principal Components with eigen values greater than 1



The percentage of variances explained is magnitude of eigenvalues divided by sum of eigenvalues by scaling data we get mean=0 and standard deviation=1.

I get below variance(eigen vectors) and proportion of variance vector.PC1-0.27\*100=27%; PC2- 0.13\*100=13%; PC3-0.11\*100=11%;PC4-0.09\*100=9%

