You are consulting the manager of a high-end restaurant chain. While they employ an excellent sommelier, they wish to keep the quality of their wines in all locations high and at a comparable level. To do this, the restaurants need an easy way to predict high quality wines based on a set of characteristics. Your task is to provide a classifier that can predict top-quality wines based on the observable characteristics in the attached file (winequality.csv) and a short write-up that explains to your client (who knows next to nothing about statistics) what you have done.

In particular, you should

* Familiarize yourselves with the data and prepare it for the analysis
  + Look at some basic summary statistics
  + Plot your data
  + …
* Pick a classification method (naïve Bayes or k Nearest Neighbor) to predict top quality wines
  + Bonus points for picking the classifier based on cross validation
* Assess the performance of your classifier
* Send me your code as Python file together with a **1-page** summary which Explains to a **non-expert** what you were doing which includes a brief description of
  + the data
  + the method you used (including the key assumptions)
  + the results and how the client can apply this when new data comes in