EDA Homework 5

Due: Friday, April 2, 5pm.

The files wine-1.csv, wine-2.csv, and wine-3.csv contain information about the chemical properties of some wines and the soils the grapes were grown in. In particular, the first contains the levels of a collection of fermentative volatile compounds in each of the wines, the second contains the levels of a collection of varietal compounds in these same wines, and the third contains information about the type of soil the grape was grown in and its water-holding capacity.

The data are taken from a study of grenache wines from the Rhone Valley [1].

- Using the `recode` function (or any other way you would like), create a variable that describes the water holding capacity of the soil as either $\leq$ 80mm, 90-140mm, or $\geq$ 200mm.
- Using as predictors the fermentative volatile compounds given in `wine-1.csv`, perform LDA so as to separate the three different water-holding capacity groups you just made. Make a biplot of the result and describe what it tells us about the relationship between the water holding capacities of the soil and the fermentative volatile compounds in the resulting wine.
- Using as predictors the varietal compounds in `wine-2.csv`, do the same thing as in the previous part. Make a biplot of the result and describe what it tells us about the relationships among the variables. Which set of predictors is better at separating the different water capacities?

References

[1] Isabelle Sabon, Gilles de Revel, Yorgos Kotseridis, and Alain Bertrand. "Determination of Volatile Compounds in Grenache Wines in Relation with Different Terroirs in the Rhone Valley," Journal of Agricultural and Food Chemistry 2002 50 (22), 6341-6345. DOI: 10.1021/jf025611k