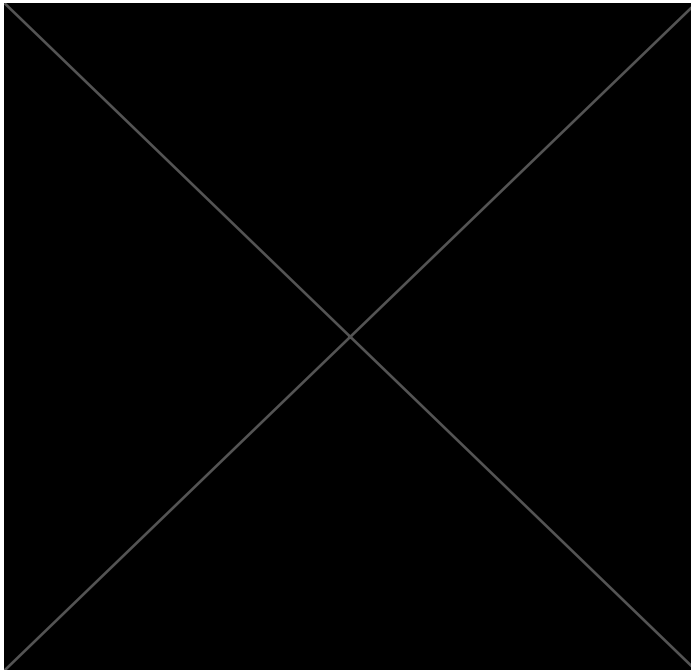


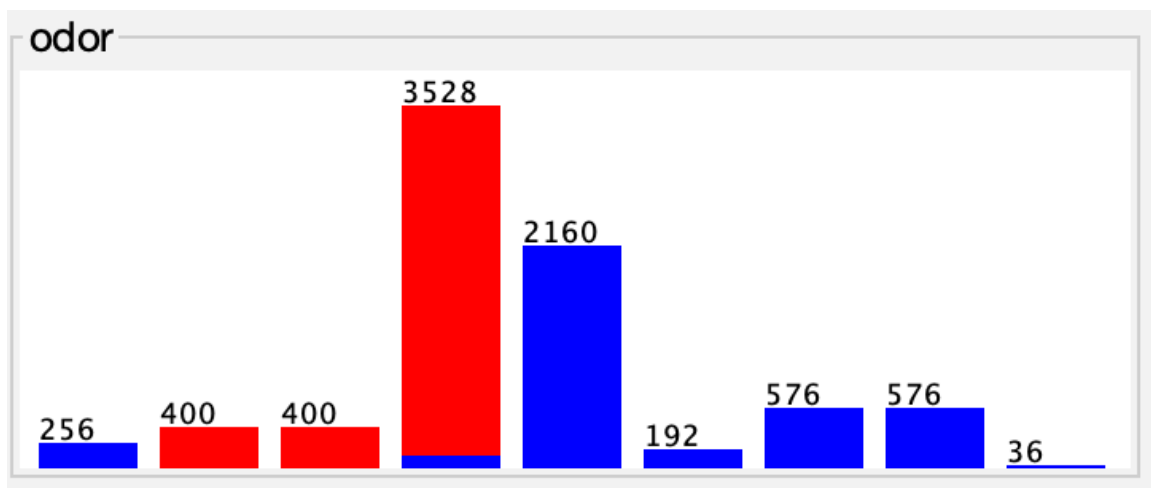
Oblig 6

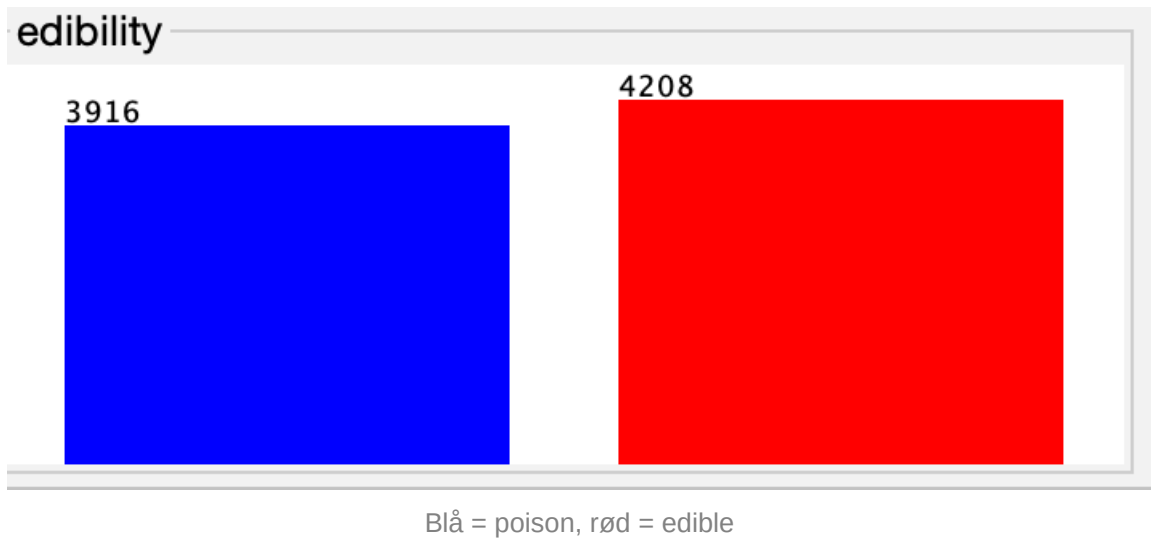


Exercise 6

For this exercise you can use either Python with sklearn or Weka.

- Using the UCI mushroom dataset from the last exercise, perform a feature selection using a classifier evaluator. Which features are most discriminative?

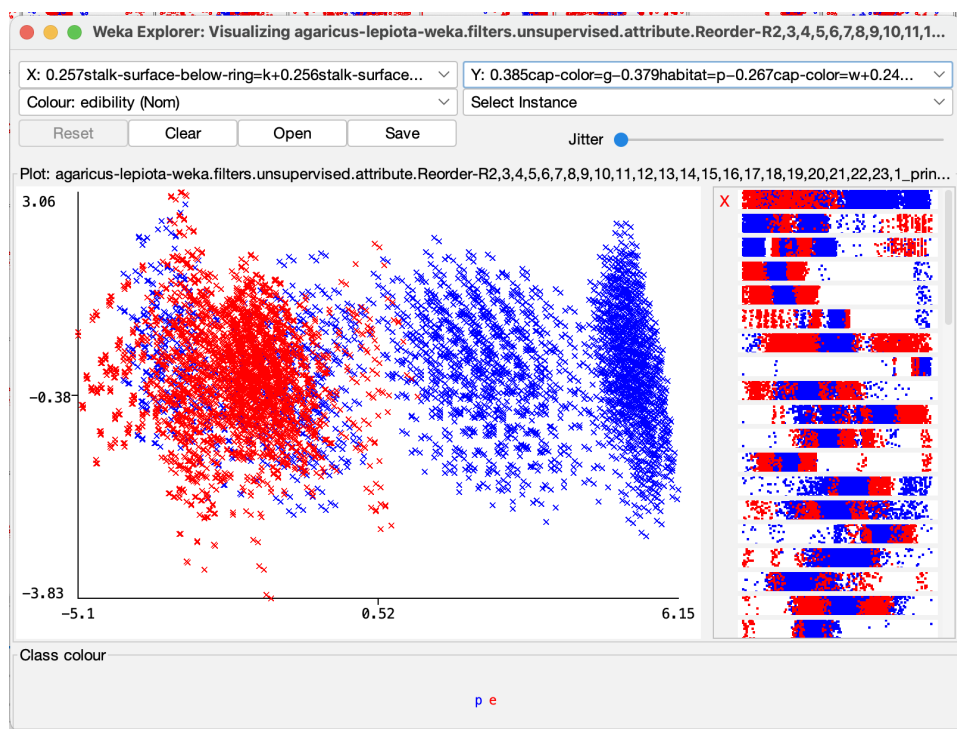
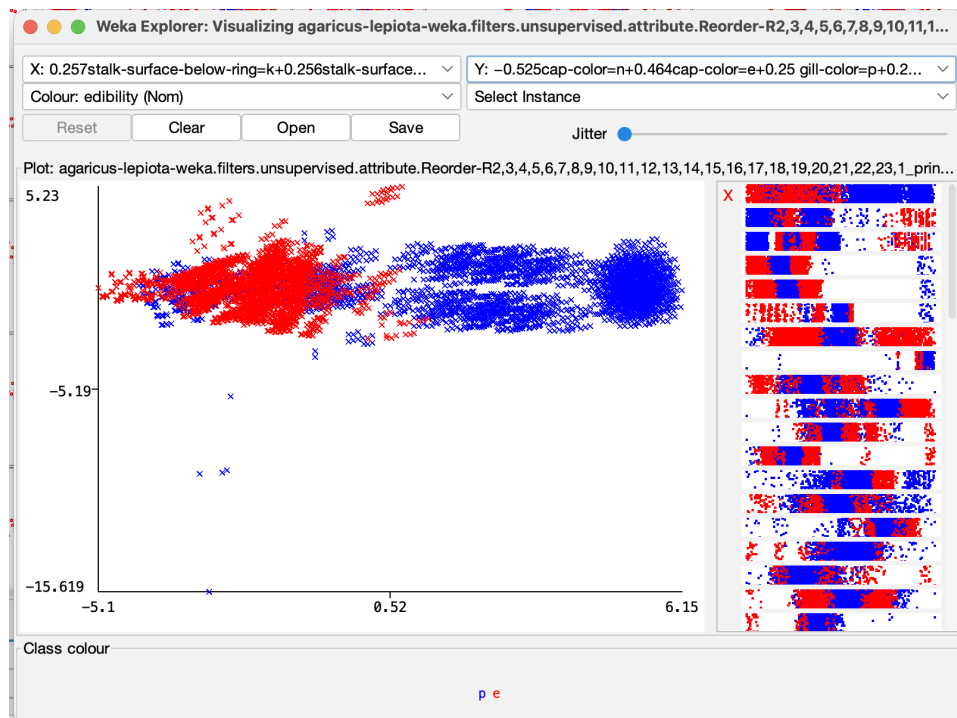


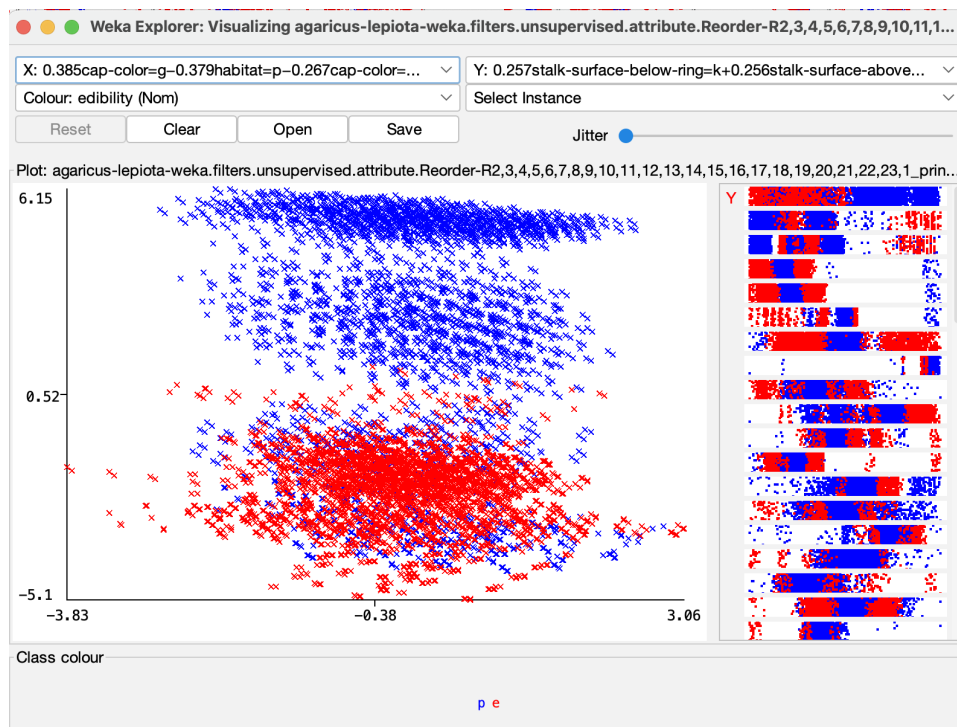


Selected attributes: 5,7,12,17 : 4
odor
gill-spacing
stalk-surface-above-ring
veil-color

Stemmer overens med det vi fant i visualization, at odor var en discriminatory feature.

- Use principal components analysis to construct a reduced space. Which combination of features explain the most variance in the dataset?





- Do you see any overlap between the PCA features and those obtained from feature selection?

Finner et overlapp med variabelene stalk-surface-above-ring = {k} og odor = {f, c, s}.

