QNC Sept 30, Class Notes

1. Correlation and Regression
   1. Similarities: they both deal with associations in a linear fashion
   2. Assuming linear data, scatter plot is the visual tool
   3. Difference: Regression each of the different variables have a special, different role (i.e. independent vs dependent variable), while for a correlation you can switch out the two variables
   4. Correlation: can utilize a histogram of the variables, to isolate them
      1. Can’t claim causality
      2. Go from 2 dimension to one dimension – a line
      3. PCA plot – plot the most variable with the least variable (the most oblong shape)
   5. Correlation coefficient - if you know the value of one, how well do you know the value of another
      1. Extreme 1: correlation coefficient of 1 is a straight line
      2. Extreme 2: correlation coefficient of 0 is a cloud