Stat 470/670 Homework 1

Not due, just for practice.

1. Download the CyTOF data from
   http://jfukuyama.github.io/teaching/stat670/notes/cytof_one_experiment.csv.

   The csv has 50000 rows; if the computation is too slow because of the number of observations, feel free to subset down to a smaller number (see the `subset` function).

2. Choose one of the columns of the dataset and make at least two plots showing its distrubition (choose from ECDF, quantile, histogram, density estimate). What does the plot tell you about the distribution of the values in that column?

3. For that same column you used above, make a Q-normal plot. What does the plot tell you about the distribution of the values in that column?

4. Choose any two of the columns of the dataset and make a Q-Q plot comparing the distributions. What does the Q-Q plot tell you about similarities or differences between the distributions of the values in the two columns?