Tyler Paulsen
cs420-01

1) Estimation of how long the homework will take.

The homework will take about 3 hours. start -10:30am

2) 1D Clustering using Otsu's method.

a) Yes, a machine is very black and white. There is no way to validate if the reason the person is speeding is legitimate or not.

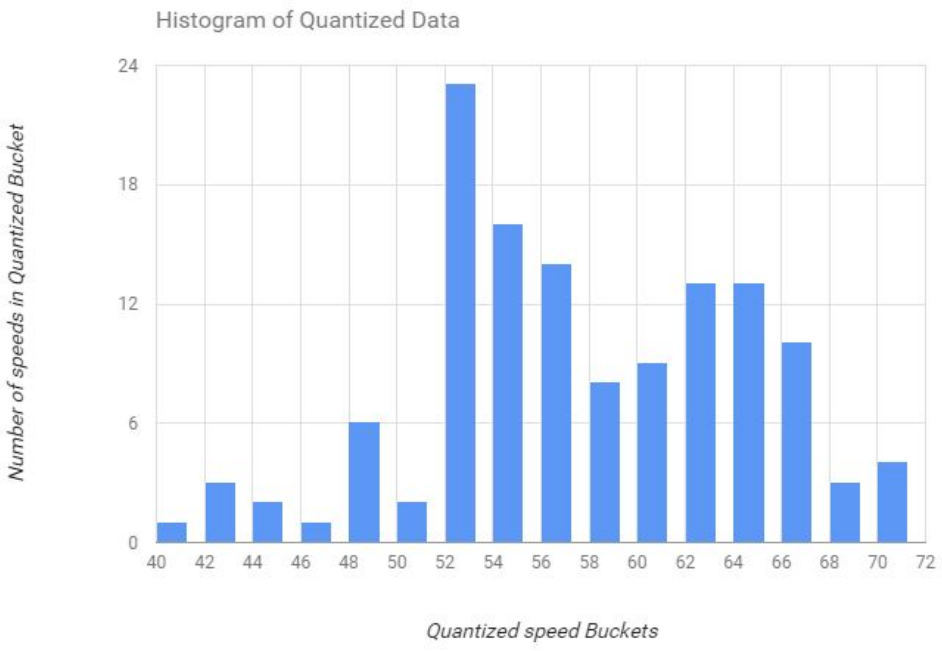
b) No they do not, I would still not be ethically for it.

e) With Otsu's method, the bin that had the lowers mixed variance was the 52-54 bin.

f) The minimum mixed variance was 26.05

g) The program would take the first mixed variance that was the lowest. No, the data had no tie, so this situation did not have to be dealt with.

h)



3) Exploratory Data Analysis

a) The mode is 7, the median is 15, and the average is 15.775

Tyler Paulsen
cs420-01

b) With the last element removed the mode median and average is as follows: 7, 14, and 15.769. The value for the median and the average went down.

c) With Otsu's method, the data was placed into buckets the size of 2 with the 16 not omitted. When the clustering routine was run, the threshold that best split the data was the 18-20 bin with the mixed variance of 1.98.

4) reality versus perception.

a) I finished the rough draft at 2:00pm, so it took me 3.5 hours. Proofreading took 1 hour when coming back to it the next day, so I totaled 4.5 hours on the homework.
b) The fraction would be 3/4.5 = .67
c) Creating an idea on how to program something is easy, finding the right libraries, and how to use them is the tough part. With this assignment, the most time consuming part was figuring out the syntax for Python again.

The first google search i did was: "why is estimating software time hard". The search gave me a stackexchange post that relates to what i said. http://programmers.stackexchange.com/questions/60994/why-is-software-schedule-estimation-so-hard. A summary of this post in one sentence would be: it is hard to estimate something we have not done before.

5) bonus