1. *I came up with ideas based on the suggestions and what I saw in the first few rows of train. Then I created some features. I tried out each feature on its own to see how well it predicted. If it didn't do very well, I decided not to use that feature. I plotted my feature grouped by its spam or ham category to see if my feature was making any meaningful divisions. Another thing I did was look at model weights (for appropriately scaled features like words). I removed some of the words with zero weights.*
2. *I tried counting the number of uppercase words in the subject, but then I decided that the percentage of uppercase words might be more appropriate. The percentages worked better. I also tried a small number of words that I thought of myself, which wasn't very good and then I found about 460 or so spam words from the internet, which worked quite well.*
3. *When I looked at the weights for my word model alone, I was surprised that some of the words that I thought would be very indicative of spam had low or zero weights. I realized later that this could be due in part to other words or phrases that tend to co-occur in spam emails.*