Important links:

http://stackoverflow.com/questions/18974437/how-to-build-a-clean-word-cloud-using-pytagcloud-without-a-crowded-image-pytho

<http://www.nltk.org/book/ch01.html#counting-vocabulary>

python d3?

Competitive Analysis:

* [http://rappers.mdaniels.com .s3-website-us-east-1.amazonaws.com/](http://rappers.mdaniels.com.s3-website-us-east-1.amazonaws.com/)
  + This is the project from which I received my initial motivation. This project is nice in that it marries a charming aesthetic with a clearly understandable data representation.
  + I like that the rappers can be split into region (or just the wu tang clan)
  + I like how each rapper’s image is clear and easily differentiable from one another.
  + I also like that when the user mouses over the picture of the rappers the exact number of unique words used by said rapper is displayed.
  + I do not like that there is no ability to search for the rapper.
  + I do not like that nothing happens when the rapper is clicked
  + I do not like that there is no animation in the project to give it a dynamic appearance.
* <http://www.cameronmusco.com/rap_collaboration_graph/>
  + This project is interesting because it does a similar thing to how I plan to compare lyrics between rappers and see which rappers are most derivative of other rappers (meaning that they have the highest number of connections between rappers)
    - NetworkX?
    - Python d3?
  + I like how thorough this graph is, but I do not like how hard to navigate and how aesthetically unappealing it is.
  + I like how on click it shows the rank of a rapper’s influence
  + I like how searchable this chart is
* <http://nickbloom.net/d3/rapnetwork/>
  + I like how dynamic this graph looks with the fluid animation that happens when it is initially loaded.
  + I also like how rappers are searchable, but not how the search is case-sensitive
  + I do not like how the rappers are the user has no idea who they are looking at unless they mouse over them, and how nothing happens on click

Because I am interested in hip hop and rap the direction that I want to go with my term project is something in the same vein as this: [http://rappers.mdaniels.com.s3-website-us-east-1.amazonaws.com](http://rappers.mdaniels.com.s3-website-us-east-1.amazonaws.com/" \t "_blank) where I would use data from [http://rap.genius.com](http://rap.genius.com/" \t "_blank) and the twitter api to analyze certain rappers for their lyrical abilities. I think that the project could benefit from more interactivity and a deeper data analysis. I could, for example, see what specific words are common amongst certain rappers and perhaps even incorporate the comments on the Rap Genius pages of certain songs. I would also like to use the data that I would be collecting to find which rapper influenced others by seeing which rappers are lyrically similar (judging to see with whom a rapper has the most words in common).I think that I would also like to analyze the social impact that certain rappers or rap songs have had on the culture by using the twitter API and using some metric to measure their influence, and also see which rapper is currently at “the top of the game.” Furthermore, I would like to do a complete analysis on rap data by doing a tag cloud on the data to see which words or moods were most prevalent in the entirety of rap from year to year.