Project Overview: Newsfeed Realty

This project looks at the newsfeeds of the friends of the program's user. Specifically the project provides data on who posts most on your friends' walls.

Implimentation

For this project, I used the pattern.web module to mine data and manipulate from Facebook. The program finds the author of each post (for x posts) in each friend's newsfeed and adds one to the frequency that that person has added news to the newsfeed. The frequency tallies are then changed into percentage of total posts. If the author of the post is not the user or one of the user's friends, the author is declared to be of 'other' social groups. To store data, I found it most convenient to use dictionaries because I could define the keys to be people and the elements to be values.
To run the code, open the program and edit the api and the numPostHistory variables in the if __name__ == '__main__' section of the code (bottom of script) to reflect user's Facebook lincense key and have far back in the newsfeed they wish to look.

Results

The code understandable results. Different social circles posted on members walls more than in other social group wall. Ex: Boyfriend/Girlfriend pairs had high values of claimed "Newsfeed Realty" for each other. Oliner had high values for other oliners and no scores for friends of mine from high school.

Reflection

The code does what it is supposed to. The output data is easy to use as input data for other programs. However the output data could be better represented visually as opposed to text. Additionally it would have been interesting to have the code output only certain specified friends instead of everyone on friend list (which can take awhile even for small friend lists). I feel like the scope of the project is not significantly larger or smaller than other homework assignments and is thus appropriately scaled for one person.

One note of interest is that Facebook mining didn't work from spider, so most of my debugging and unit testing was done from the terminal window. The functions were written so that each could be tested for a small set of data (except analyzeAllFriends and newsfeedReality, which by necessity were large).