Assignment 4 Report

Professor Nelson

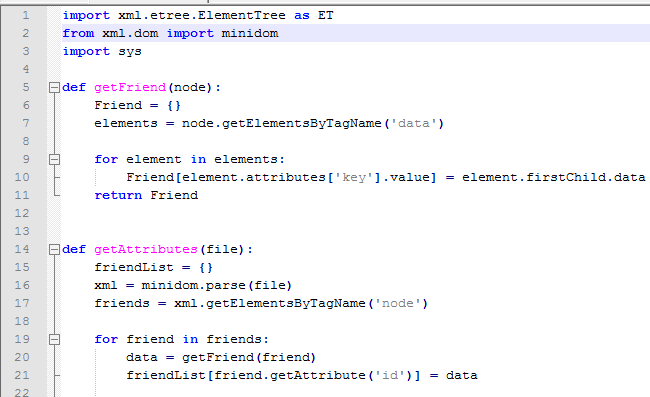
By Tyler Medina

3/2/17

**Part 1**

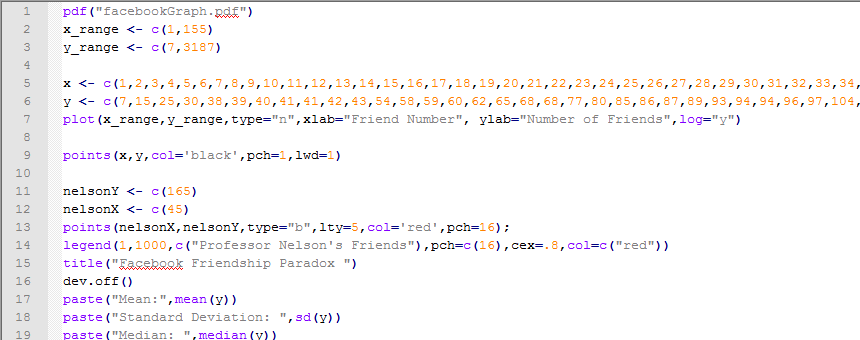
**1.1 Extracting data from GRAPHML file**

To get the desired data from the GRAPHML file that was provided, I had to create a python program that could extract the data. Using the ElementTree and minidom libraries I extracted the needed elements with the getElementsByTagName() function.



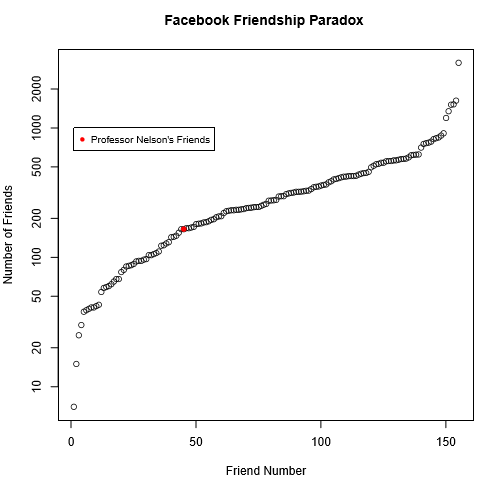
**1.2 Generating R program**

Once the key elements that were needed to make the graph were extracted, I created an R program to plot the data. To get the mean, median and standard deviations, I used the built-in R functions.



|  |  |
| --- | --- |
| Mean | 357.735483870968 |
| Standard Deviation | 370.704490379149 |
| Median | 259 |

This space was intentionally left blank



**1.3 Interpreting the data**

Because Professor Nelson is positioned below the median, 45 and 259 respectively, it means that at least more than half of his friends contain more friends than he does. The friendship paradox holds true for Professor Nelson.

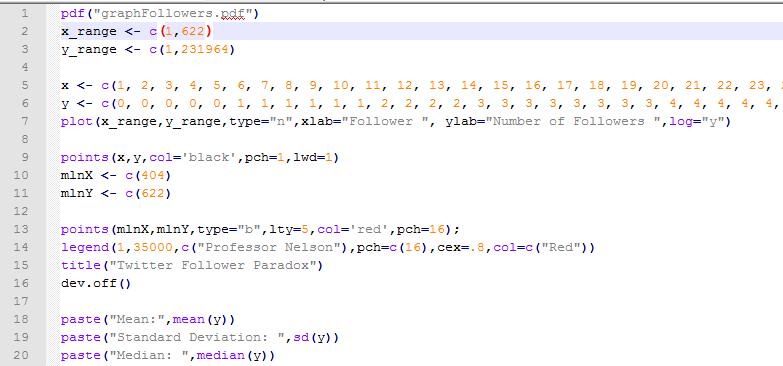
**Part2**

**2.1 Getting number of followers**

To get the number of Professor Nelson’s followers, I used a program very similar to the one in assignment two. This streams the how many followers the followers of Professor Nelson have, but the data is out of order. To put the data in order, I used a small sort function.

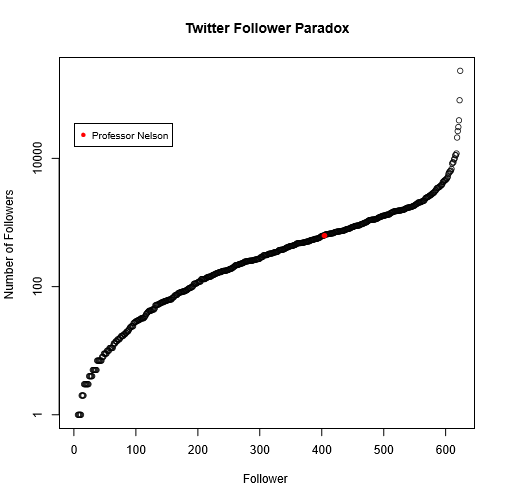
**2.2 Generating R program**

Using the same code from part 1, I generated a program that graphs the follower count of Professor Nelson’s followers. Just like part 1, I use the built-in R mean, standard deviation, and median functions to calculate those values.



|  |  |
| --- | --- |
| Mean | 1514.8202247191 |
| Standard Deviation | 10167.031694119 |
| Median | 312 |

This space was intentionally left blank



**2.3 Interpreting the Data**

Professor Nelson is above the median, so he has more followers than most of his followers. This indicates that the friendship paradox does not hold true for his Twitter profile. There could be a couple of reasons for this. One, because you don’t have to follow someone back that follows you, like in Facebook, there can be a growing disparity between popular and un poplar tweeters. Another reason, is that people generally follow a person whose content or views they like and agree with. This is a different relation from Facebook, where people have personal connections.

References

The Document Object Model API. Python Software Foundation. Version 2.0. Available at:

https://docs.python.org/2/library/xml.dom.html