Bryan Carey
CS 432
Assignment 4

1. I extracted the amount of friends each of Dr. Nelson's friends had, & placed them into a .txt file. . I ran the text file through a simple Python program I made use of by utilizing the statistics library (installed through command prompt). I found about the statistics library after I began looking into the math library where I initially thought the three required functionalities belonged to. I used code from https://pythonprogramming.net/statistics-python-3-module-mean-standard-deviation/ and https://developers.google.com/edu/python/regular-expressions as the basis of my program to complete the computations. The snap below is from my initial framework gathered from the first link.

```
import statistics

example_list = [5,2,5,6,1,2,6,7,2,6,3,5,5]

x = statistics.mean(example_list)
print(x)
```

The resulting code I created is shown below-

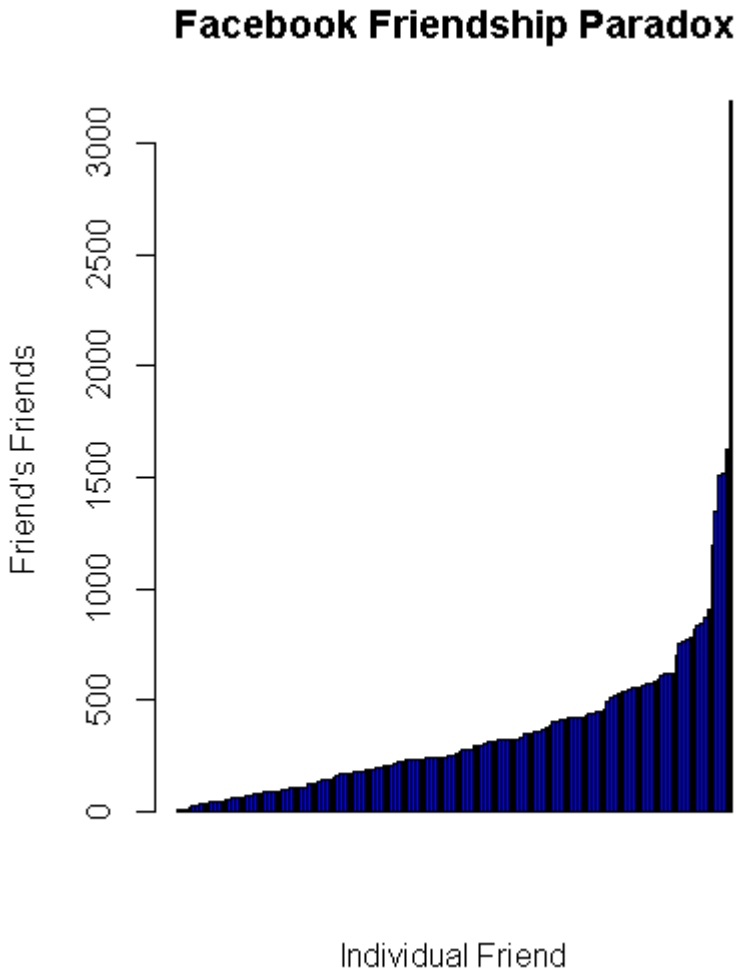
```
1   import statistics
2   import math
3   import re
4
5   if __name__ == "__main__":
6
7       with open('friend_num.txt') as input:
8           followers = input.read()
9       #https://docs.python.org/2/library/re.html
10      followers = map(int, re.findall('\d+', followers))
11
12      # https://pythonprogramming.net/statistics-python-3-module-mean-standard-deviation/
13      print ' The mean number of friends is ' , statistics.mean(followers)
14
15      print ' The median number of friends is ' , statistics.median(followers)
16
17      print ' The standard deviation of the number of friends is ' , statistics.stdev(followers)
18
19
```

I ran my code and stored the generated output in a .txt file, friends_stat.txt.

When graphing this information, initially I tried to follow the link here http://stackoverflow.com/questions/8774002/setting-the-color-for-an-individual-data-point , however the point marked for Dr. Nelson did not change color. Another student sent me the following link, http://stackoverflow.com/questions/22781685/different-colors-for-each-bar-in-stacked-bar-graph-base-graphics and was able to fix the problem. I was told that instead of using a color on top of a color, I had to do so as a range. The revision to my code was as follows-

Bryan Carey
CS 432
Assignment 4
barplot(x,main="Facebook Friendship Paradox",xlab="Individual Friend",ylab="Friend's Friends", **col=col**)

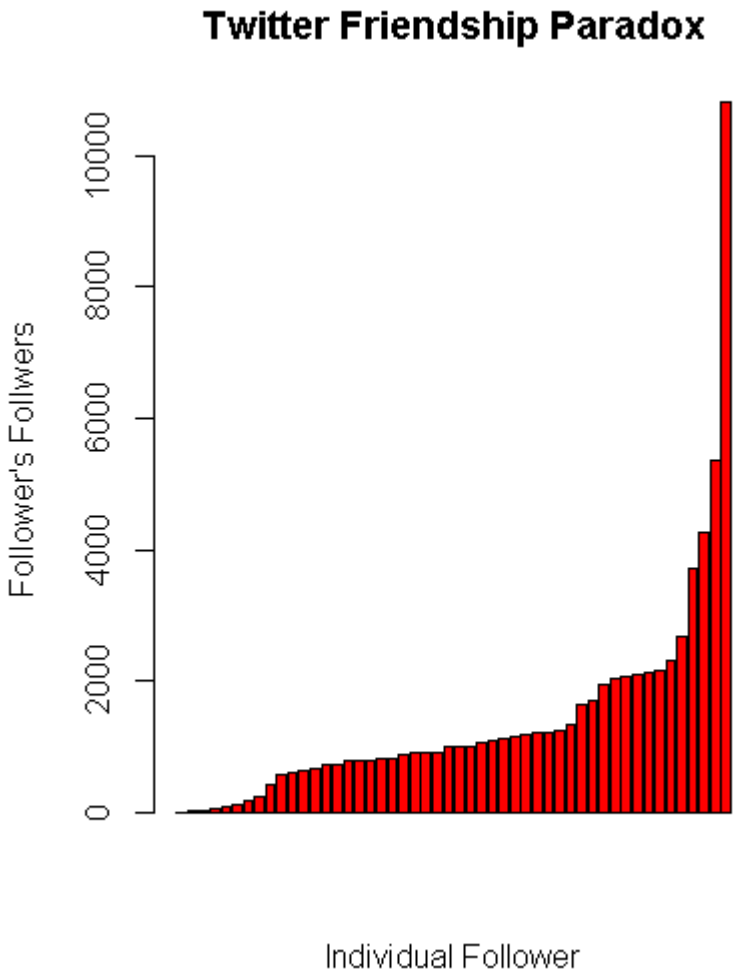
col <- c(rep(c("blue"),42),rep(c("white"),1),rep(c("blue"),111)). The first 42 elements in the list are blue, followed by Dr. Nelson(white bar), and he is followed by the rest of the blue bars(111). You will have to zoom in to see the differentiation.



The friendship paradox holds true.

Bryan Carey
CS 432
Assignment 4

2. For question 2, I acquired exactly 50 followers on Twitter and created a .txt file listing my follower's followers given the relatively small size of the group. I took that document, followers.txt and ran a program identical to the one above to yield my mean, median, and standard deviation values. These are stored in a file, followers_stat.txt.



The friendship paradox holds true for my twitter account as well. All .R files used to generate the graphs are located in my repository.