

Assignment 5

Web Science CS595
Name: Amara Naas

Answer to question 1

The Python program (q1.py)¹ will read the file mln.graphml¹ and parse it by using BeautifulSoup and XML element tree. the program will extract only the friend count data for each friend and save it in a file called resultdata.txt¹. By using the data in the resultdata file I created a graph of the number of friends (y-axis) and the friends sorted by number of friends (x-axis) as shown in figure 1. We can see from figure 1 that friendship paradox holds for mln Facebook account where total number of his friends was 154 with friend count data represented (The friends who do not have friend count are eliminated by using if statement in q1.py program). Also from the resultdata file I computed the mean, standard deviation, and median of the number of friends that mls' friends have as following:

```
mean = 357.66
standard deviation = 370.74
median = 259
```

Answer to question 2

The Python program (q2.py)¹ is a modified program from assignment 1. The only change is in the url value. The program will read all the followers of phonedude_mln twitter account by extracting only the followers count data for each follower and save it in a file called resultdata1.txt¹. By using the data in the resultdata1 file I created a graph of the number of followers (y-axis) and the followers sorted by number of followers (x-axis) as shown in figure 2. We can see from figure 2 that friendship paradox holds for phonedude_mln twitter account where total number of his followers was 196. Also from the resultdata1 file I computed the mean, standard deviation, and median of the number of followers that phonedude_mln's followers have as following:

```
mean = 524.3
standard deviation = 1281
median = 207
```

¹File uploaded to github

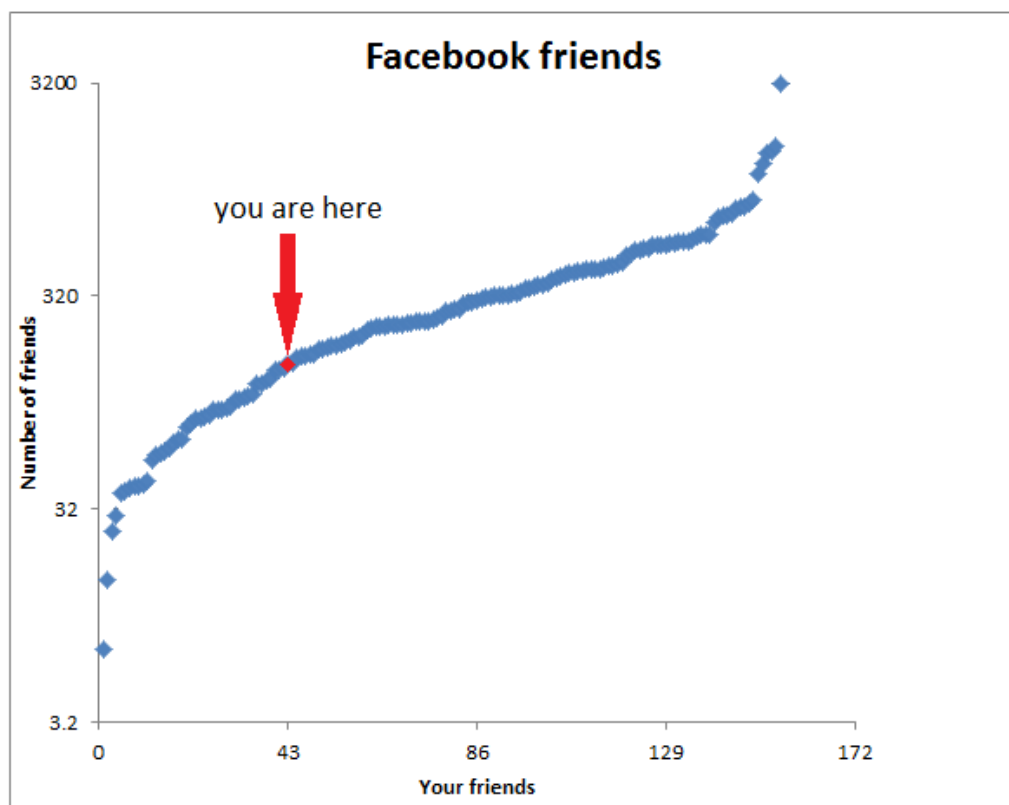


Figure 1: Friends count

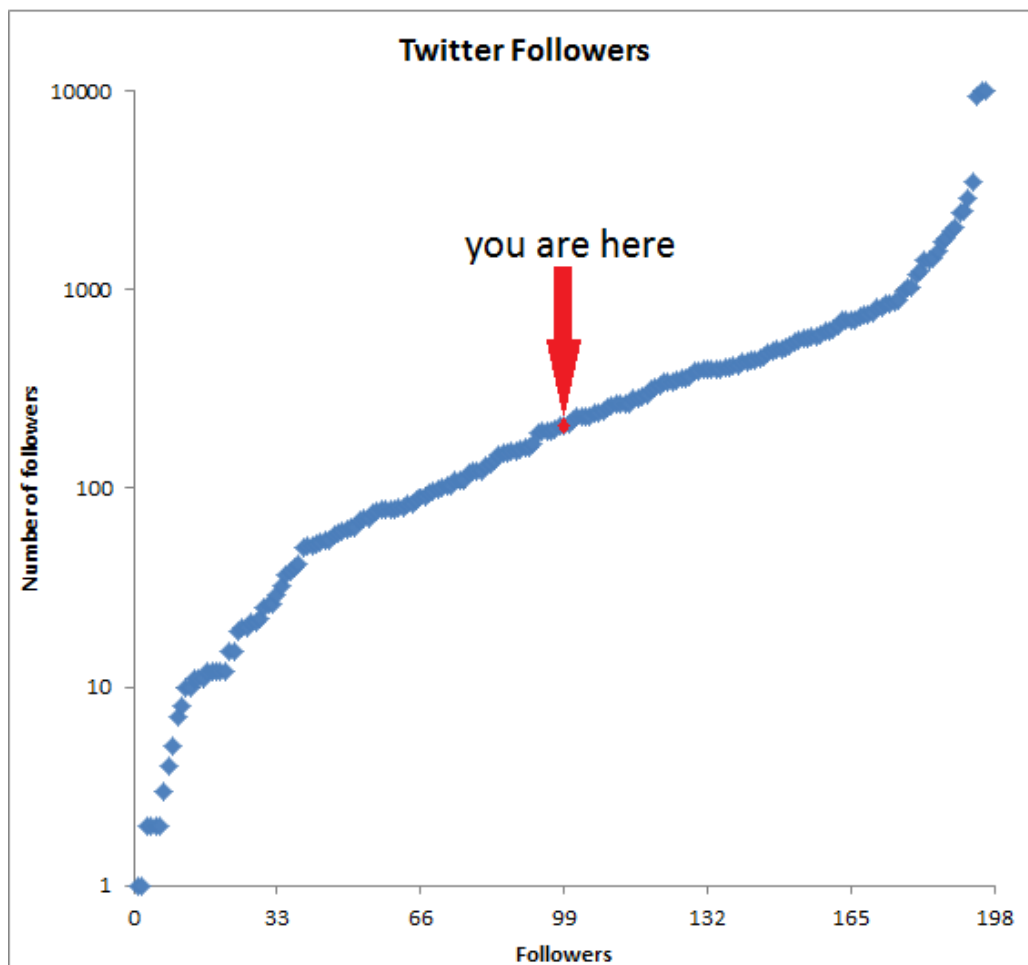


Figure 2: Followers count

```

% python q3.py
{'status': '200', 'content-length': '375', 'content-location': u'http://api.linkedin.com/v1/people/~?oauth_body_hash=2jmj715rSw0yVb%2Fv1WAYkK%2FYBwk%3D&oauth_nonce=88238545&oauth_timestamp=1382420086&oauth_consumer_key=yjhoa5dnxfk8&oauth_signature_method=HMAC-SHA1&oauth_version=1.0&oauth_token=5d7702ea-86bc-42fc-b70e-67476a84f52e&oauth_signature=CfwB%2Fe3Q97RJITQ9K8oSDk8sqkd8%3D', 'x-li-pop': 'PROD-ELA4', 'transfer-encoding': 'chunked', 'vary': '*', 'x-li-uuid': 'iChEir5XLxPwjdzTisAAA==', 'server': 'Apache-Coyote/1.1', 'x-li-fabric': 'PROD-ELA4', 'connection': 'keep-alive', '-content-encoding': 'gzip', 'date': 'Tue, 22 Oct 2013 05:34:45 GMT', 'x-li-request-id': '41NSAZKEW1', 'access-control-allow-origin': 'http://www.linkedin.com', 'x-li-format': 'xml', 'content-type': 'text/xml; charset=UTF-8'}
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<person>
  <first-name>Ama</first-name>
  <last-name>Naa</last-name>
  <headline>Student at odu</headline>
  <site-standard-profile-request>
    <url>http://www.linkedin.com/profile/view?id=298876082&authType=name&authToken=brqB&trk=api*a3380471*s3453881*</url>
  </site-standard-profile-request>
</person>

```

Figure 3: Followers count

Answer to question 3

In the Python program (q3.py)² will access linkedIn API and retrieve the data. First thing to do is to create an application, and generate your own access token. Secondly, register the new application, get the consumer key and consumer secret, and save them at the q3.py². Because I do not have friends in linkedIn I only get the my data by generating the Oauth_Tokens and run q3 python file, and set Oauth.Token and Oauth.Token.Secret as shown in figure 3.

Answer to question 4

The Python program (q2.py)² is a modified python program from question 2. The only change is in the url value where followers/list is replaced by friends/list. The program will reads all the followings of phonedude_mln twitter account by extracting only the friends count data for each following and save it in a file called resultdata2.txt². By using the data in the resultdata2 file I created a graph of the number of followings (y-axis) and the followings sorted by number of followings (x-axis) as shown in figure 4. We can see from figure 3 that friendship paradox holds for phonedude_mln

²File uploaded to github

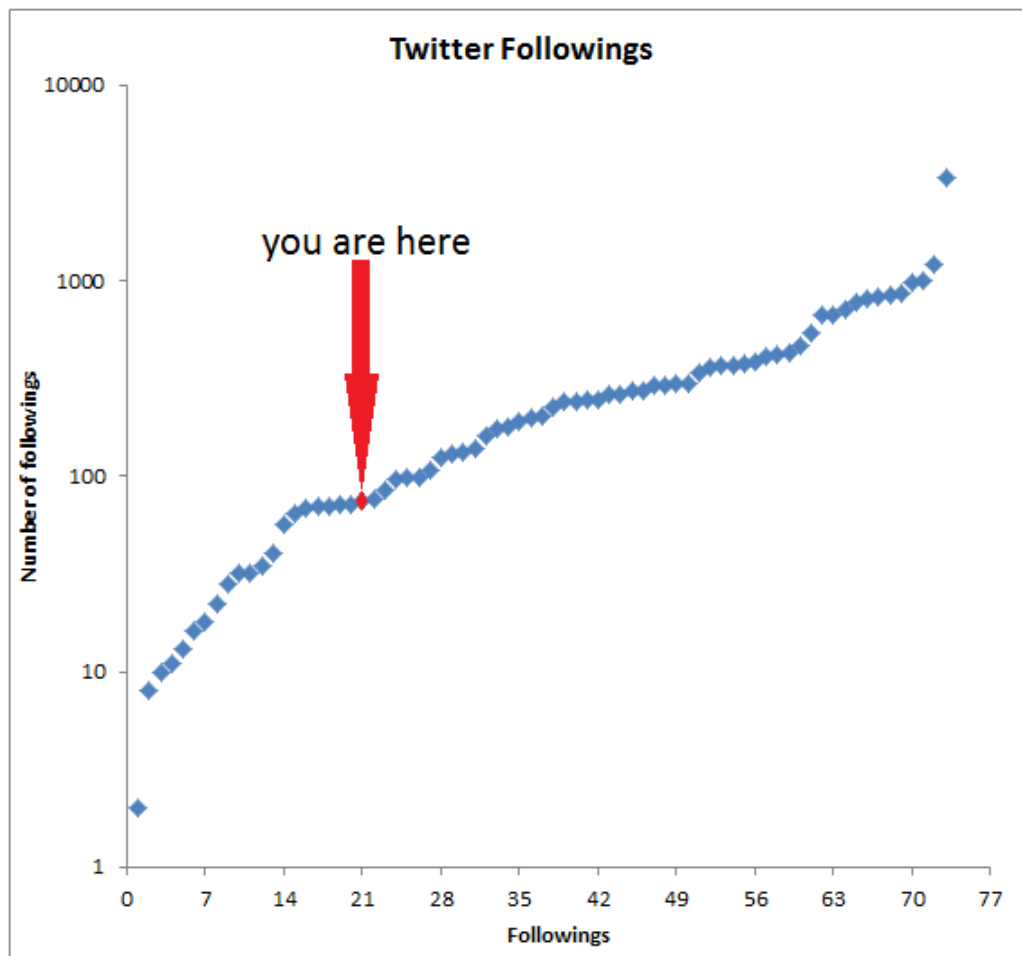


Figure 4: Followers count

twitter account where total number of his followings was 73. Also from the resultdata2 file I computed the mean, standard deviation, and median of the number of followers that phonedude_mln's followings have as following:

mean = 323.52
standard deviation = 455.76
median = 202