

Exploring Social Networks

Sofia Huang
due 11/03/2022

1. Friendship Paradox on Facebook

Determine if the friendship paradox holds for a user's Facebook account. (This used to be more interesting when you could more easily download your friend's friends list from Facebook. Facebook now requires each friend to approve this operation, effectively making it impossible.)

acnwala_friends_friends_count.csv (provided on Piazza) contains a user's friends' names and number of friends they each have.

Create a graph of the number of friends (y-axis) and the friends themselves (x-axis), sorted by number of friends. The friends don't need to be labeled on the x-axis: 1, 2, 3,..., n should be sufficient. Include the user in the graph in the appropriate sorted position (count the number of their friends) and label as U.

I used the .csv file given by Prof. Nwala and imported it as a DataFrame. Then, I sorted the values based on friend count and added Prof. Nwala and his own friendcount into the DataFrame at the correct position. I used Matplotlib to plot the bar graph to show the data and to prove the friendship paradox holds.

```
1 def plot_friends(df, filename):
2     # sort df values and reset index
3     df = df.sort_values('FRIENDCOUNT')
4     df = df.reset_index(drop=True)
5     index = np.arange(1,99)
6     df['INDEX'] = index
7     # add Prof. Nwala's data to the df
8     df2 = {'USER': 'Alexander Nwala', 'FRIENDCOUNT': 98, 'INDEX': 'U'}
9     line = pd.DataFrame(df2, index=[10])
10    df2 = pd.concat([df.iloc[:10], line, df.iloc[10:]]).reset_index(
drop=True)
11    # plot the bar graph and fix details
12    plt.figure(figsize=(36, 8))
13
14    ax = plt.gca()
15    ax.margins(x=0)
16
17    plt.bar(df2.index, df2['FRIENDCOUNT'], width=0.6)
18    plt.xticks(np.arange(len(df2['INDEX']), df2['INDEX'], fontsize='18'
))
19    plt.xlabel('Users', fontsize=24)
```

```

20 plt.ylabel('Friend Count', fontsize=24)
21 plt.title('Friendship Paradox', fontsize=30)
22 # save plot to file
23
24 plt.tight_layout()
25 plt.savefig(filename)
26 #plt.show()

```

Listing 1: plotting Prof. Nwala's Facebook friends' friend count

Q: What is the mean, standard deviation, and median of the number of friends that the user's friends have?

```

1 The mean of the number of friends that the user's friends have:
  542.6734693877551
2 The median of the number of friends that the user's friends have: 396.0
3 The standard deviation of the number of friends that the user's friends
  have: 536.6744685696292

```

Listing 2: summary statistics for Facebook friends

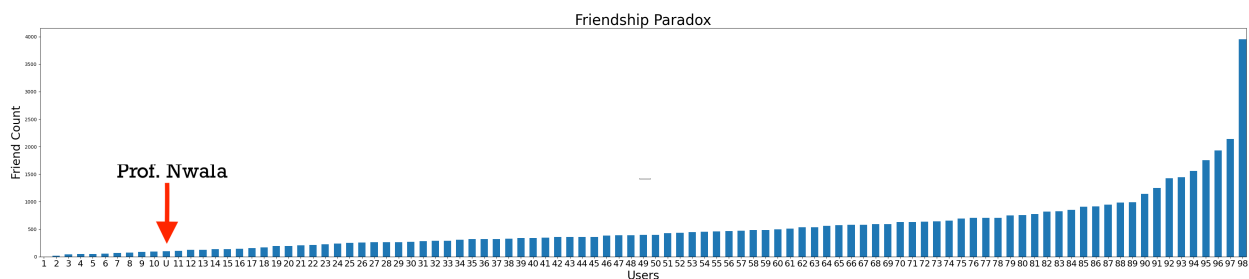


Figure 1: Friend count of Prof. Nwala and his friends on Facebook, sorted by friend count.

Q: Does the friendship paradox hold for this user and their friends on Facebook?

The friendship paradox does hold for this user and their friends on Facebook because the average of the user's friends' friends is 876, while the user only has 98 friends on Facebook. So, on average, the user's friends, have more friends than the user themselves.

2. Friendship Paradox on Twitter

Determine if the friendship paradox holds for your Twitter account. Since Twitter is a directed graph, use followers as the value you measure (i.e., "do your followers have more followers than you?").

I used twarc to get a list of my own followers as a .jsonl file. Using this file, I was able to obtain their follower counts by parsing the data structure returned from Twitter. I used a dictionary

to store the follower's username and their follower count. Then, I converted the dictionary to a .csv file to save it to my computer.

```
1 twarc2 followers _sofiahuang myfollowers.jsonl
```

Listing 3: obtaining my own followers using twarc

```
1 def get_followers_followers(jsonl_file):
2     f = open(jsonl_file)
3     # read in all the lines
4     lines = f.readlines()
5     followers_stats = {}
6     # add a header entry to dictionary for the csv later
7     followers_stats.update({'USER': 'FRIENDCOUNT'})
8     # each line is a json
9     for line in lines:
10        follower_data = json.loads(line)
11        # get my follower count
12        num_my_followers = follower_data['meta']['result_count']
13        print(num_my_followers)
14        # loop through my followers and get their username + follower
count
15        # to add to a dictionary
16        for i in range(num_my_followers):
17            follower_username = follower_data['data'][i]['username']
18            follower_followers = follower_data['data'][i]['
public_metrics']['followers_count']
19            print('username: {} \nfollower_count: {}'.format(
follower_username, follower_followers))
20            followers_stats.update({follower_username:
follower_followers})
21        # add myself to the dictionary
22        followers_stats.update({'_sofiahuang': num_my_followers})
23
24        # open file for writing dictionary to csv file
25        w = csv.writer(open("/Users/sofiahuang/Documents/WM/FALL2022/
DATA440/followers_followers.csv", "w"))
26        # loop over dictionary keys and values
27        for key, val in followers_stats.items():
28            # write every key and value to file
29            w.writerow([key, val])
```

Listing 4: obtaining my followers' follower count

Q: What is the mean, standard deviation, and median of the number of followers that your followers have?

```
1 The mean of the number of friends that my friends have on Twitter:
876.3305084745763
```

```
2 The median of the number of friends that my friends have on Twitter:
  175.0
3 The standard deviation of the number of friends that my friends have on
  Twitter: 5092.913779325692
```

Listing 5: summary statistics for Twitter followers

Q: Does the friendship paradox hold for you and your followers on Twitter?

The friendship paradox holds for me and my friends on Facebook. The average of my followers' number of followers is 876. This number is skewed due to an outlier that has over 30k followers, but even the median of my followers' follower count is 175, compared to my own follower count of 117. So, my followers have more followers than me, on average.

References

- Twarc - Getting followers for a list of usernames <https://twittercommunity.com/t/twarc2-getting-followers-for-a-list-of-username/176959/2>