PSOSM Assignment 1
Divay Prakash - 2014039

Ans 1) a, b, c
Parts a, b and c are present in the form of graphs in 2 different folder titled 'Panama Papers graphs' and 'Fake News graphs'. All were generated using the 'twitter.py' script and changing the 'searchQuery' parameter to either '#PanamaPapers' or '#FakeNews'. Folders also contain files titled 'tweets.txt' containing the tweets obtained.

Ans 1) d
Part d is present in the form of graphs in the aforementioned folders. I analysed the media, images, urls, hashtags, mentions etc. Also I analysed the account creation and timezone split of Twitter users. Other conclusions are in the logs below:

#PanamaPapers logs
Downloaded 475 tweets, saved to tweets.txt
322 tweets contained URLs
67 tweets contained images
0 tweets contained media other than images
163 tweets mentioned other users
448 tweets contained hashtags
68 tweets were quote tweets
Out of 67 tweets containing images, 51 were retweeted by other users
Only 76.119403%% of tweets containing images were retweeted!
Out of 475 tweets, 419 were retweeted by other users
However 90.196078%% of all tweets except those containing images were retweeted!

#FakeNews logs
Downloaded 10025 tweets, saved to tweets.txt
5373 tweets contained URLs
472 tweets contained images
0 tweets contained media other than images
6850 tweets mentioned other users
8934 tweets contained hashtags
2027 tweets were quote tweets
Out of 472 tweets containing images, 150 were retweeted by other users
Only 31.779661%% of tweets containing images were retweeted!
Out of 10025 tweets, 1652 were retweeted by other users
However 15.722810%% of all tweets except those containing images were retweeted!

Ans 2a) i, ii, iii
Parts i, ii and iii and present in the form of graphs in the folder title 'Facebook graphs'. All were generated using the 'facebook_process_data.py' script.
Data from Facebook was obtained using the 'facebook_get_data.py' script.

Ans 2c)
No, this is not possible. Given an access token for User A, irrespective of the relationship of A and B, it is necessary to have B's explicit permission to access his/her data. B will need to allow the application to access the data and only then we will be able to read it.

Ans 3)

Parameter	Facebook	Twitter
Rate limit	200 API calls / 60 minutes	180 API calls / 15 mins
Data format	JSON	JSON
Data Access	Very limited	Much more as compared to FB
Data Scraping	- Requires user-specific OAuth token - Requires user to grant permission to app, depends on privacy settings	- Can be done using application OAuth token only - Can be done without user permission

Bonus question)
The solution is in the file 'bonus.py'.

****