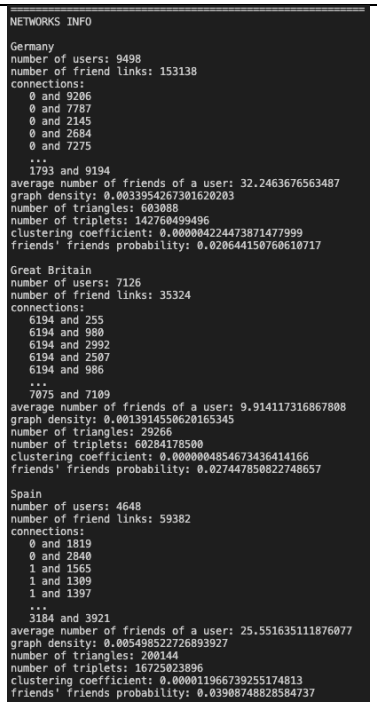
Twitch Network Analysis

Project Summary

In this project I managed to write a code that extracts information about the network of Twitch users from the data files, generates different representations of this network (list of connections in the graph and adjacency matrix), implemented a number of data analysis algorithms including calculations of number of edges and vertices, average number of friends of users in the network, number of triangles in the graph, number of triples in the graph, graph density, clustering coefficient, and average probability that user's friend's friend is also user's friend. I also implemented functions for matrices such as matrix multiplication, power of a matrix, as well as calculation of number of triangles and number of triplets using properties of matrices. However, I did not use these algorithms due to their low efficiency with large matrices.

Results Presentation

Below are the screenshots of the code output



```
==========================================================
NETWORKS INFO

Germany
number of users: 9498
number of friend links: 153138
connections:
   0 and 9206
   0 and 7787
   0 and 2145
   0 and 2684
   0 and 7275
   ...
   1793 and 9194
average number of friends of a user: 32.2463676563487
graph density: 0.0033954267301620203
number of triangles: 603088
number of triplets: 142760499496
clustering coefficient: 0.0000042244738714779 99
friends' friends probability: 0.020644150760610717

Great Britain
number of users: 7126
number of friend links: 35324
connections:
   6194 and 255
   6194 and 980
   6194 and 2992
   6194 and 2507
   6194 and 986
   ...
   7075 and 7109
average number of friends of a user: 9.914117316867808
graph density: 0.0013914550620165345
number of triangles: 29266
number of triplets: 60284178500
clustering coefficient: 0.0000004854673436414166
friends' friends probability: 0.027447850822748657

Spain
number of users: 4648
number of friend links: 59382
connections:
   0 and 1819
   0 and 2840
   1 and 1565
   1 and 1309
   1 and 1397
   ...
   3184 and 3921
average number of friends of a user: 25.551635111876077
graph density: 0.005498522726893927
number of triangles: 200144
number of triplets: 16725023896
clustering coefficient: 0.0000119667392551748 13
friends' friends probability: 0.03908748828584737
```

```
France
number of users: 6549
number of friend links: 112666
connections:
   0 and 6420
   0 and 2941
   0 and 3051
   0 and 5511
   0 and 1020
   ...
   6527 and 6541
average number of friends of a user: 34.40708505115285
graph density: 0.005254594540493715
number of triangles: 422694
number of triplets: 46792338674
clustering coefficient: 0.0000090334018768518 71
friends' friends probability: 0.029339308304626187

Portugal and Brazil
number of users: 1912
number of friend links: 31299
connections:
   0 and 92
   0 and 428
   1 and 689
   1 and 1147
   1 and 1666
   ...
   1805 and 1414
average number of friends of a user: 32.73953974895397
graph density: 0.017132150575067492
number of triangles: 173510
number of triplets: 1163136520
clustering coefficient: 0.0001491742345086026 6
friends' friends probability: 0.06534737696853168

Russia
number of users: 4385
number of friend links: 37304
connections:
   0 and 646
   0 and 2510
   0 and 2425
   0 and 2596
   0 and 85
   ...
   4126 and 4346
average number of friends of a user: 17.01436716077537
graph density: 0.0038810144071111 7
number of triangles: 71445
number of triplets: 14043015120
clustering coefficient: 0.0000050875826444314 19
friends' friends probability: 0.023902203186900645
==========================================================
```

```
==========================================================
AVERAGE NUMBER OF FRIENDS IN DESCENDING ORDER

France: 34.40708505115285 friends
Portugal and Brazil: 32.73953974895397 friends
Germany: 32.2463676563487 friends
Spain: 25.551635111876077 friends
Russia: 17.01436716077537 friends
Great Britain: 9.914117316867808 friends
==========================================================



==========================================================
GRAPH DENSITY IN DESCENDING ORDER

Portugal and Brazil: 0.017132150575067492
Spain: 0.005498522726893927
France: 0.005254594540493715
Russia: 0.0038810144071111 7
Germany: 0.0033954267301620203
Great Britain: 0.0013914550620165345
==========================================================



==========================================================
CLUSTERING COEFFICIENT IN DESCENDING ORDER

Portugal and Brazil: 0.0001491742345086026 6
Spain: 0.0000119667392551748 13
France: 0.0000090334018768518 71
Russia: 0.0000050875826444314 19
Germany: 0.0000042244738714779 99
Great Britain: 0.0000004854673436414166
==========================================================



==========================================================
AVERAGE PROBABILITY THAT A FRIEND OF A FRIEND
OF A USER IS ALSO USER'S FRIEND IN DESCENDING ORDER

Portugal and Brazil: 0.06534737696853168
Spain: 0.03908748828584737
France: 0.029339308304626187
Great Britain: 0.027447850822748657
Russia: 0.023902203186900645
Germany: 0.020644150760610717
==========================================================
```

Results Analysis

I then compared the countries in terms of the calculated metrics and got the following.

	avg. friend number	density	clustering coeff.	friends prob.
1	France	Portugal and Brazil	Portugal and Brazil	Portugal and Brazil
2	Portugal and Brazil	Spain	Spain	Spain
3	Germany	France	France	France
4	Spain	Russia	Russia	Great Britain
5	Russia	Germany	Germany	Russia
6	Great Britain	Great Britain	Great Britain	Germany

From this table we can see how all these metrics are closely related and representative in showing how closely connected users in each community are. For example, Portugal and Brazil region keeps its first place for 3 and gets a second in the fourth. Similar, but opposite performance shows Great Britain. Due to such consistency in results, there is enough evidence to claim that out of these 6 regions, Portugal and Brazil region has the most closely connected user networks, whereas Great Britain has the least closely connected one. It can also be noticed that even though France has the biggest average number of friends per user, according to all the other metrics France gets the third place. That yields that friend connection in France are uniformly distributed among the graph, whereas connections in Portugal and Brazil form community clusters. Other noticeable fact is that with a relatively small average number of friends in Spain network, it has a second place in all other metrics, meaning that these connections form community clusters inside the network.

To conclude, such a complacent table gives an insight about qualitative differences between twitch user networks in different countries.