ASSIGNMENT REPORT

ON

Community Detection

Course: Social Media Analytics

Instructor: Dr. Poonam Goyal

Submitted by:

Yash Chandak - 2020H1030141P

Swapnil Akash - 2020H1030133P

BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE, Pilani, Pilani Campus

STATISTICS

Dataset	Number of Nodes (n)	Number of Edges (m)	Average Path Length (d)	Average Clustering Coefficient ©
Karate Club Network	34	78	2.40819	0.57063
Dolphin Social Network	62	159	3.35695	0.25895
Jazz musicians network	199	2743	2.24800	0.61383

COMPARISON

- Karate Dataset:

Statistics	Number of Clusters	Modularity Score	Runtime
Girvan-Newman EBC	2	0.3599	0.19
Modularity Maximization	3	0.38	0.00
Spectral Clustering using Graph Laplacian	4	0.41	0.42

- Dolphins Dataset:

Statistics	Number of Clusters	Modularity Score	Runtime
Girvan-Newman EBC	2	0.3787	0.42
Modularity Maximization	4	0.49	0.02
Spectral Clustering using Graph Laplacian	4	0.41	0.44

- Jazz Dataset:

Statistics	Number of Clusters	Modularity Score	Runtime
Girvan-Newman EBC	2	0.0036	15.83
Modularity Maximization	4	0.43	0.35
Spectral Clustering using Graph Laplacian	4	0.48	0.43

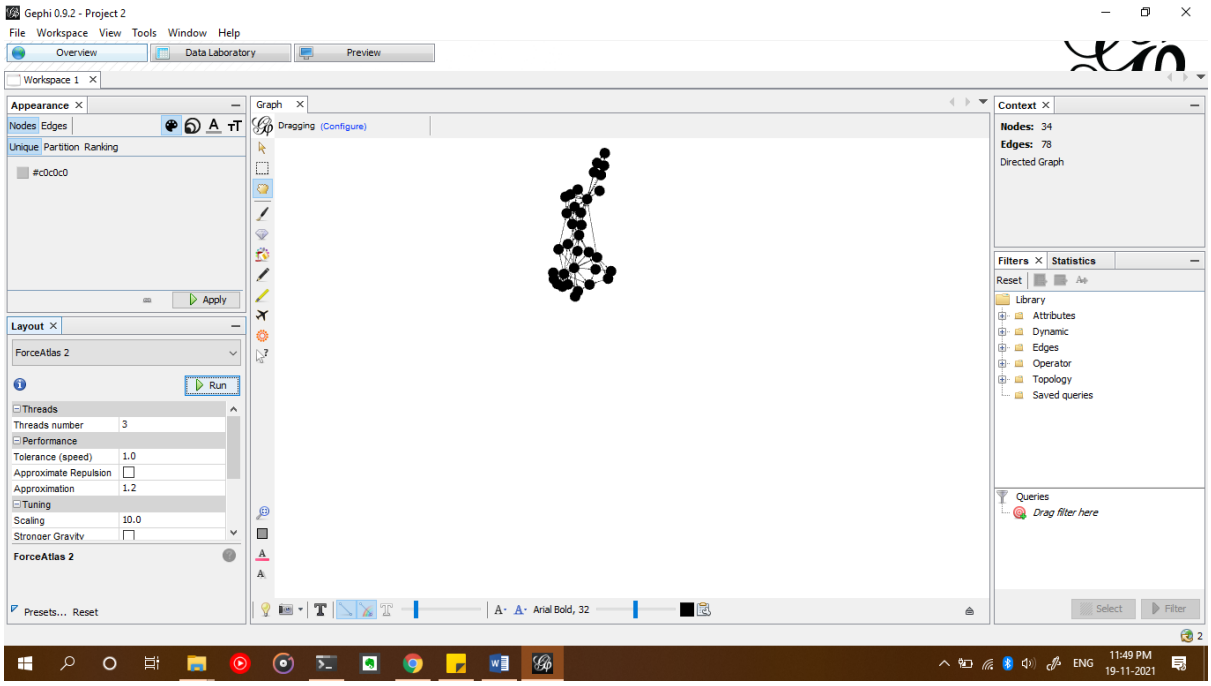
OBSERVATIONS

Average Modularity Score of Girvan-Newman EBC = 0.2474

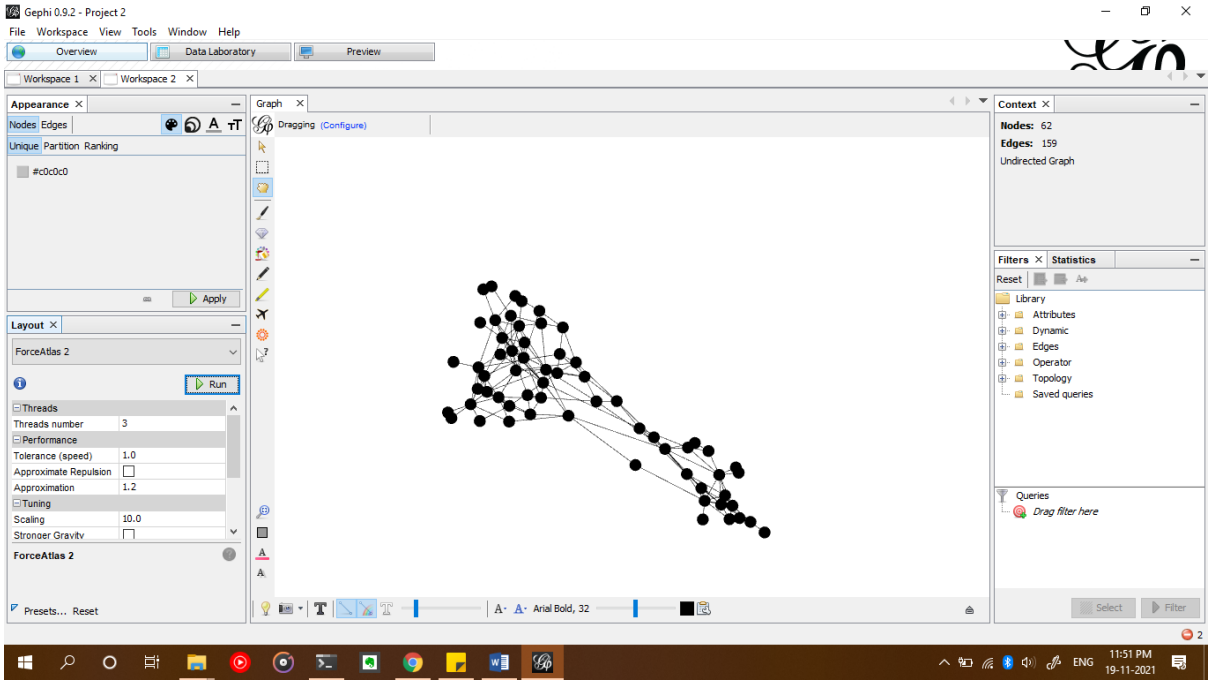
Average Modularity Score of Modularity Maximization = 0.43

Average Modularity Score of Spectral Clustering = 0.43

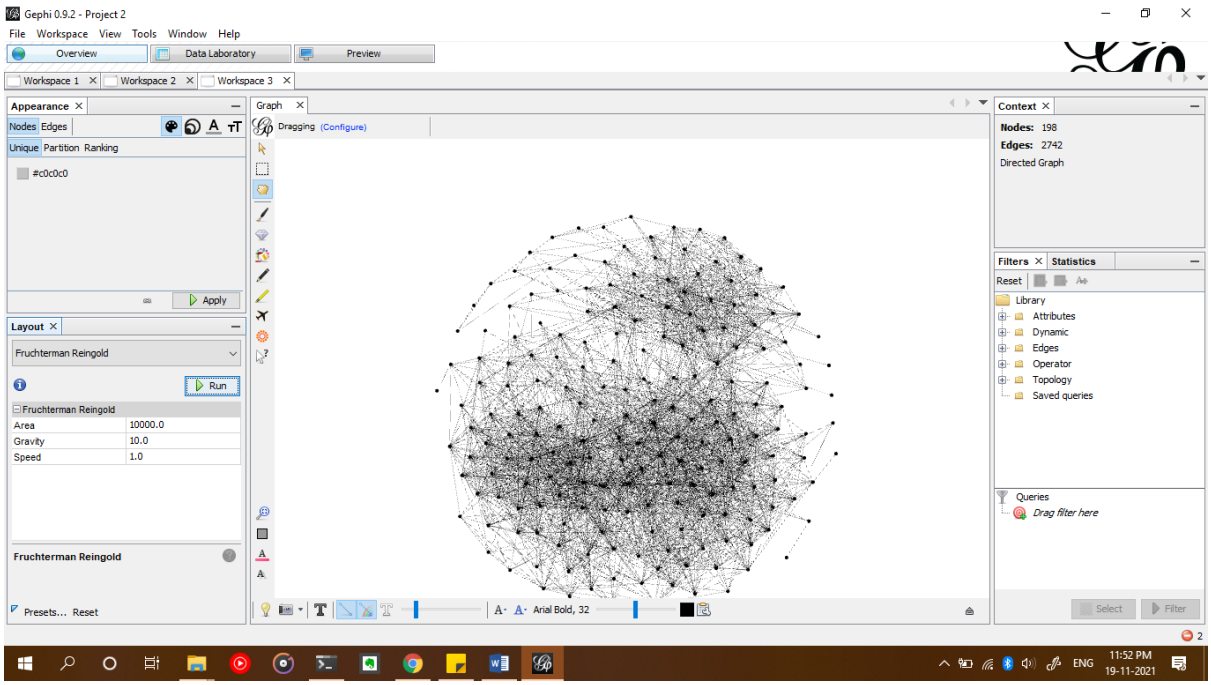
Therefore, Modularity Maximization performs better in terms of avg. modularity score and runtime.



Visualization of Karate Club Dataset. We can see the formation of two clusters as detected by the algorithms.



Visualization of Dolphins Dataset. We can see the formation of 3-4 clusters as detected by the algorithms.



Visualization of Jazz Dataset. We can see the formation of 3-4 clusters as detected by the algorithms.