Assignment 1

for

Social Media Analytics
(CS_G519)
in

Master of Engineering (Software Systems)
by

Saoni Deb **(2020H1120300P)**
Debjani Ghosh **(2020H1120263P)**

Under the supervision of
Dr. Poonam Goyal
Assistant Professor
Department of Computer Science and Information Systems

Birla Institute of Technology and Science
Pilani, Pilani Campus, Rajasthan (India)

January 2021

Question 4:
Results for Statistics:

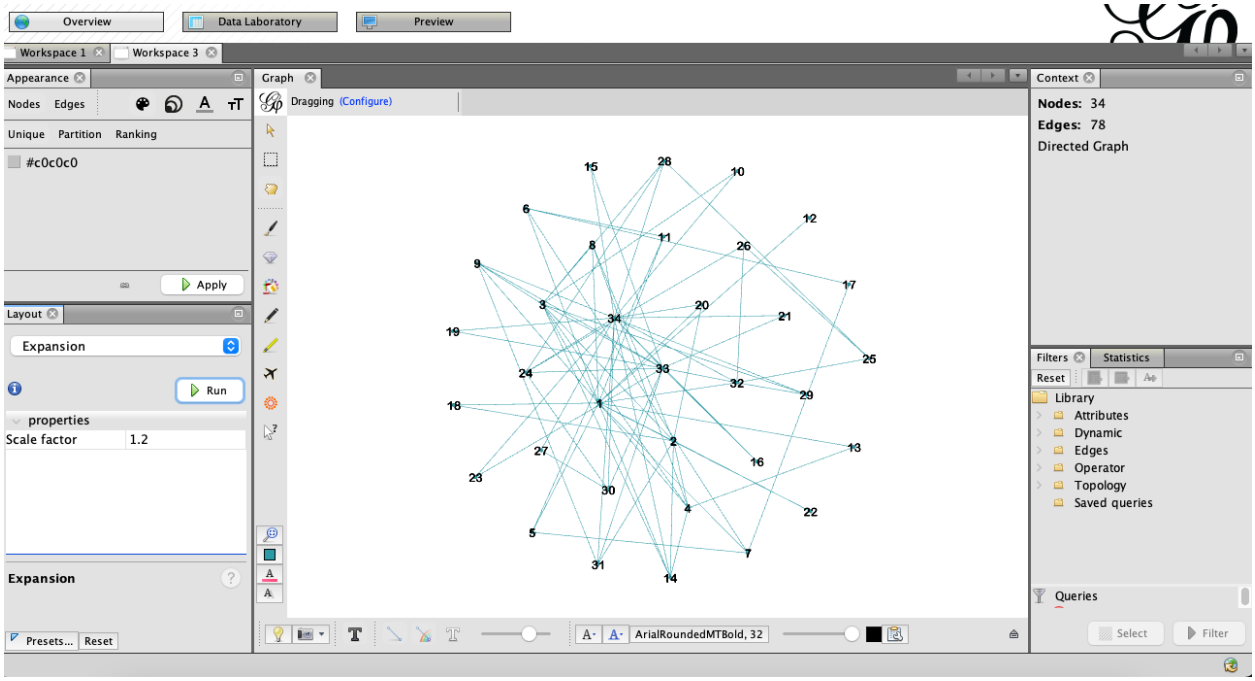
Dataset	No. of nodes	No. of edges	average path length	average clustering coefficient
Karate	34	78	2.4082	0.2557
Dolphins	62	159	3.3569	0.3087
Jazz	198	5484	2.2350	0.5202

Method	Name of the Dataset	Number of clusters found	Modularity score for the clustering	Run time of the algorithm (sec)
Girvan-Newman	Jazz	2	-	8.1014916
	Dolphins	2	-	0.4933638
	Karate	2	-	0.4002154
Modularity Maximization	Jazz	4	0.282	0.2963
	Dolphins	4	0.499	0.0222652
	Karate	3	0.355	0.015399
Spectral Clustering	Jazz	2	-	0.0807788
	Dolphins	2	-	0.0205796
	Karate	2	-	0.004091

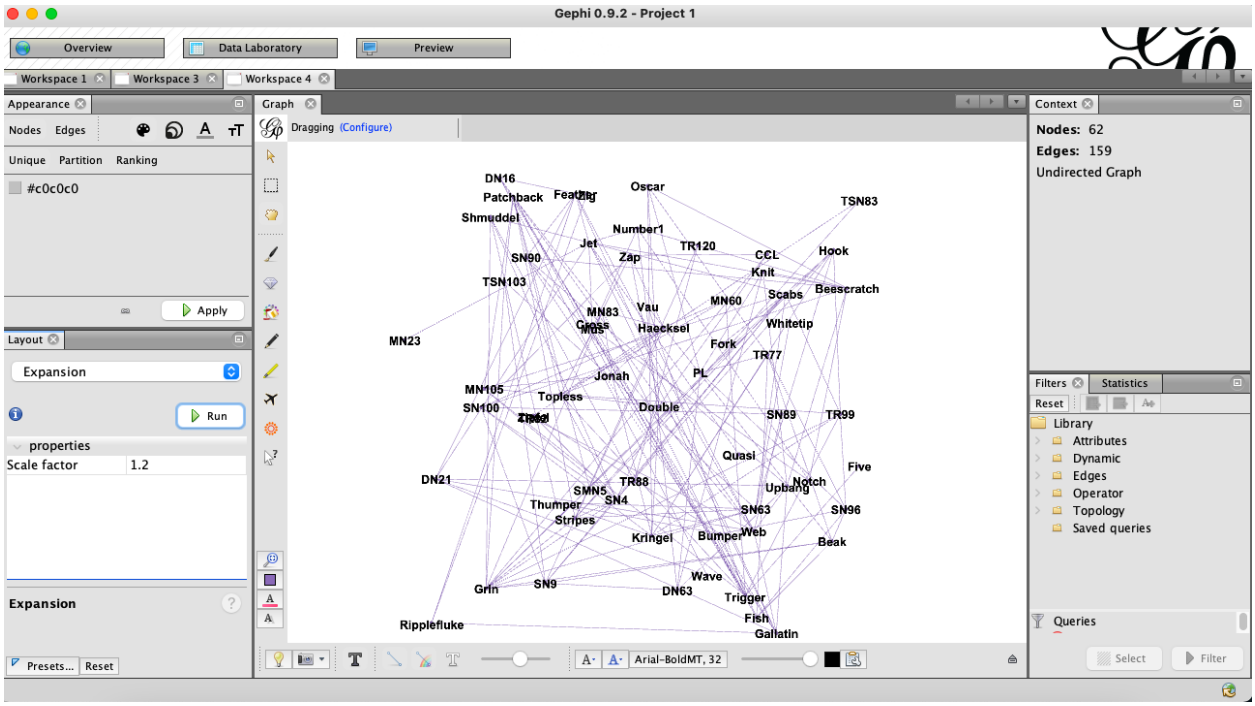
Observation:

As we can observe, spectral clustering is the most efficient as it has the least runtime for all of the above datasets. Also, runtime increases with the increase in the number of nodes and edges.

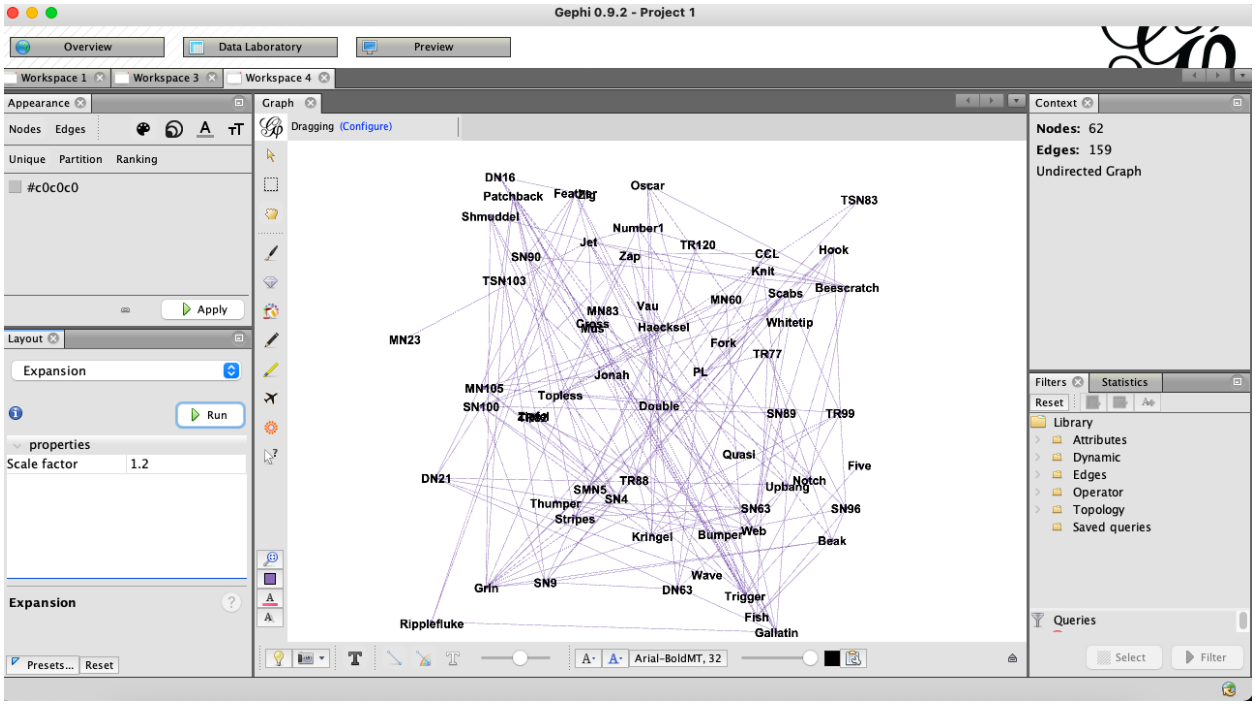
Visualisation for Karate Dataset:



Visualisation for Dolphin Dataset:



Visualisation for Jazz dataset:



Results after visualisation:

The network for Jazz is more dense compared to Dolphins and Karate dataset due to more number of edges and nodes.

Github Repository for the project:

https://github.com/saonideb/Social_Media_Analytics