CS732/DS732: Data Visualization -- Datathon 3
Jaya Sreevalsan Nair, IIIT Bangalore

Announcement: September 25, 2020, 11:59 pm IST
Submission: October 05, 2020, 11:59 pm IST on LMS

Dataset:
Source and description:
Use one of two networks for a warm-up exercise to test out any 2 of the 3 graph visualization tools/APIs – D3.js, Gephi, NetworkX (py).

- Animal network of bottlenose dolphins with 62 nodes, and 159 edges, http://konect.cc/networks/dolphins/
- Zachary karate club network, with 34 nodes and 78 edges, of the members of a university karate club collected by Wayne Zachary in 1977. http://konect.cc/networks/ucidata-zachary/

Use tabular datasets published by the World Health Organization to create new networks (similarity networks, correlation networks, etc.) and visualize network communities. https://www.kaggle.com/sudalairajkumar/novel-corona-virus-2019-dataset?select=time_series_covid_19_recovered.csv or https://bit.ly/2FWr56j

Data processing will involve determining an appropriate threshold to filter the completely connected networks.

Use the row- or column-major ordering in the tabular datasets in the WHO COVID-19 dataset collection to compute the networks – explain your rationale for your network generation.

You can use 2-3 csv files in the dataset collection to make your storyline. Each csv file must be processed as a network.

The data in a csv file must be maximally used.

Indicative Tasks:
Explain your inference from these networks and the communities you discover.

A few hints:

- ***be as creative and as curious as you can.***

Report:
Your report must contain images and text to say the following;
1. which parts of the dataset were you able to use, and how have you been able to use – the more you use the merrier.
2. which visualizations did you choose, why, what technologies did you use for the visualizations,
3. your inferences.