**Network analysis**

**Data**

Choose a dataset we have not used in class for network analysis from the course datasets page, https://kaggle.com, or somewhere else on the web. Choose one you find is interesting and from which you can extract reasonable nodes and edges (by reasonable, I mean they should make sense and not be forced; nodes should be entities or events of the same type and edges should represent some sort of relationship between the nodes). Before you process it, think about what you expect to see from a network diagram of the data.

**Report details**

In a word processor of your choice, describe the data you decided to analyze and why you chose that data (what about it interested you?). Include links to the data. Then describe the network: what are edges, what are nodes? Make a table of all the network metrics we discussed (e.g., density, diameter, etc.).

Explore different visualizations of your network. What interesting patterns emerge? How does the real network match or differ from what you expected? Describe these and include images of your network to aide your analysis.

Finally, end with some questions that you formed during your analysis that may be answered with further analysis or additional data collection and/or mining.

Use Jupyter Notebooks or Excel to perform any data extraction or cleaning you need to do. Use Gephi to extract network statistics and make visualizations.

**What to upload**

* your report (with an appropriate header, including your name, the date, the class, and a title)
* a nicely curated Jupyter notebook or Excel spreadsheet (only if you did any pre-processing to extract your edge file)
* your Gephi project (should have a .gephi extension)