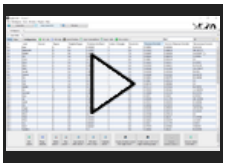
Praxis: Gephi

Re-submit Assignment

Due Feb 10 by 11:59pm **Points** 20 **Submitting** a file upload

For this assignment we will be introducing you to **Gephi** (https://gephi.org/), a graph visualization software that you can use to calculate statistics of interest about a graph as well as make it visually presentable. Go ahead and download and install Gephi and I will walk you through some of its basic functionalities:

gephi tutorial.mp4



Depending upon what version of Gephi you use the location of some buttons and tabs may vary, but the general functionality is roughly the same. If you want to know more about Gephi or what file formats it can read/write you can visit their **documentation page** (https://gephi.org/users/). You can also check out some other tutorials **online** (https://gephi.org/users/).

Now that you have a basic understanding of some of the things Gephi can do we will put that knowledge to use. We first need a graph. You can download the **Les Miserables graph** (http://www-personal.umich.edu/~mejn/netdata/lesmis.zip). **This site** (http://www-personal.umich.edu/~mejn/netdata/) has other graphs which you can download and use with Gephi. They are already in a format that Gephi will recognize (.net, .gexf, .gml, etc).

Once you have the Les Miserables graph do the following:

1. Load the graph into Gephi as an undirected graph.
2. Calculate and record the number of nodes, number of edges, average degree, average clustering coefficient, and average path length of the graph.
3. Save a plot of the degree distribution of the nodes.
4. Then, select a graph layout that you think best visualizes the graph and save it.
5. Create a new Jupyter notebook and use the **random graph generator** (https://networkx.readthedocs.io/en/stable/reference/generated/networkx.generators.random_graphs.erdos_renyi_graph.html#networkx.generators.random_graphs.erdos_renyi_graph) tool introduced in the last assignment to create a random graph with the same number of nodes and about the same average degree and number of edges as the real graph.
6. Save this random graph to a file format that Gephi can read (**see this page for Networkx read/write functions** (https://networkx.readthedocs.io/en/stable/reference/readwrite.html)) and perform steps 1-4 with this new graph. You can save it as pajek (.net), gml, or gexf.

You should then answer the following questions:

1. How do the degree distributions of the real graph and random graph compare? Explain any differences or similarities you see and consider why they might exist.
2. Is the real graph more clustered or less than the random one? What might this tell you about the organizing principles of the real graph?
3. Does the real graph exhibit the small-world property?

When complete, compile your results, figures, and analysis into a **PDF** and upload that, along with the Jupyter notebook you used to generate the random graph.