**Web Analytics Lab 6**

**Gephi – Network Visualization & Network Statistics**

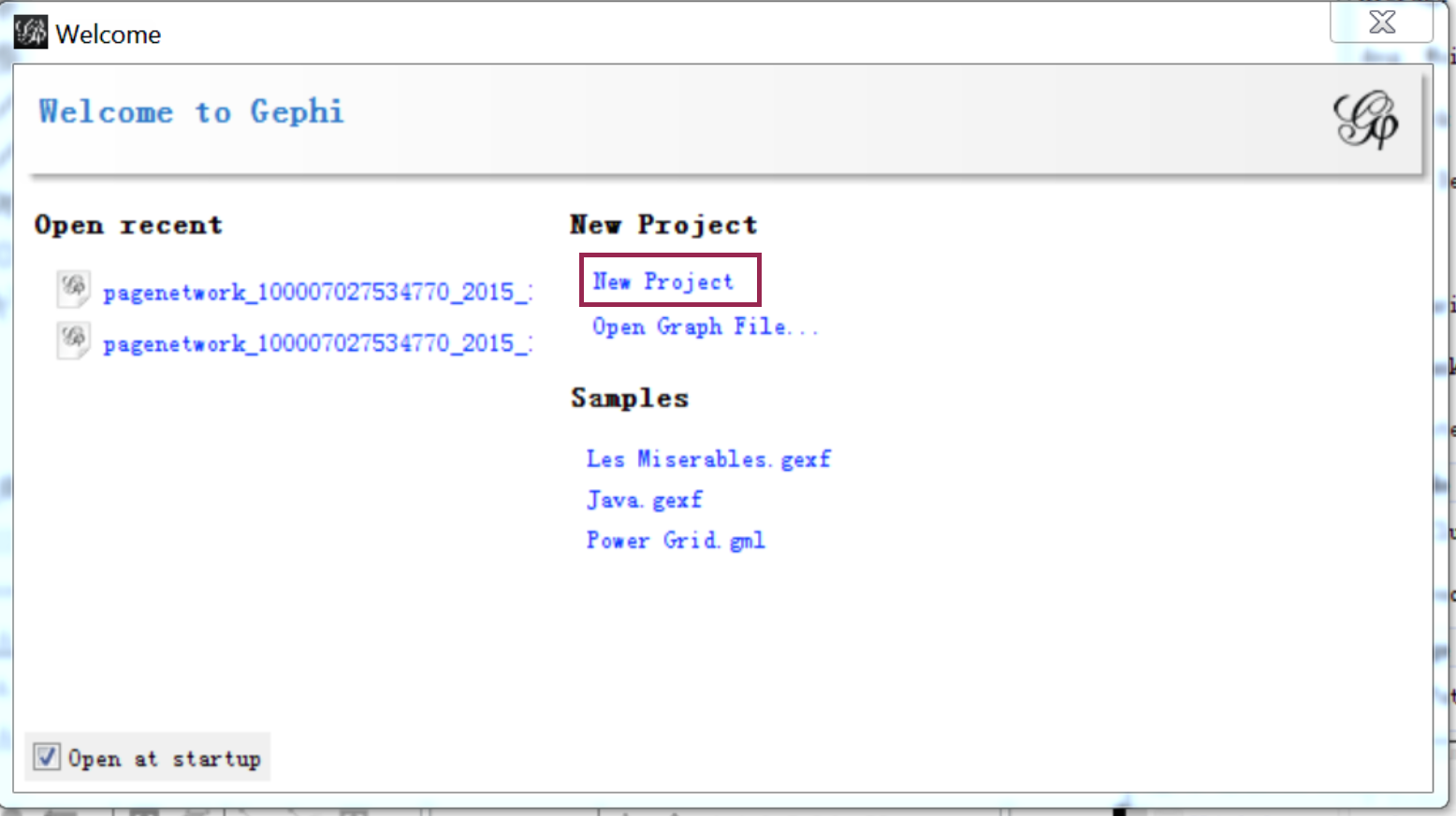
1. **Installing Gephi**

* Download and install Gephi

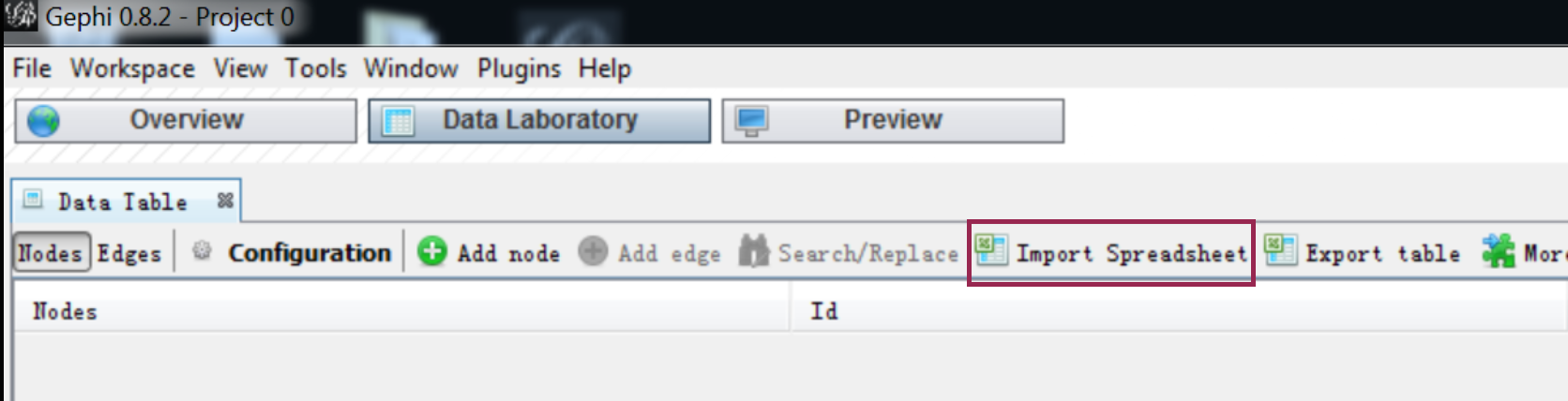
<https://gephi.org/users/download/>

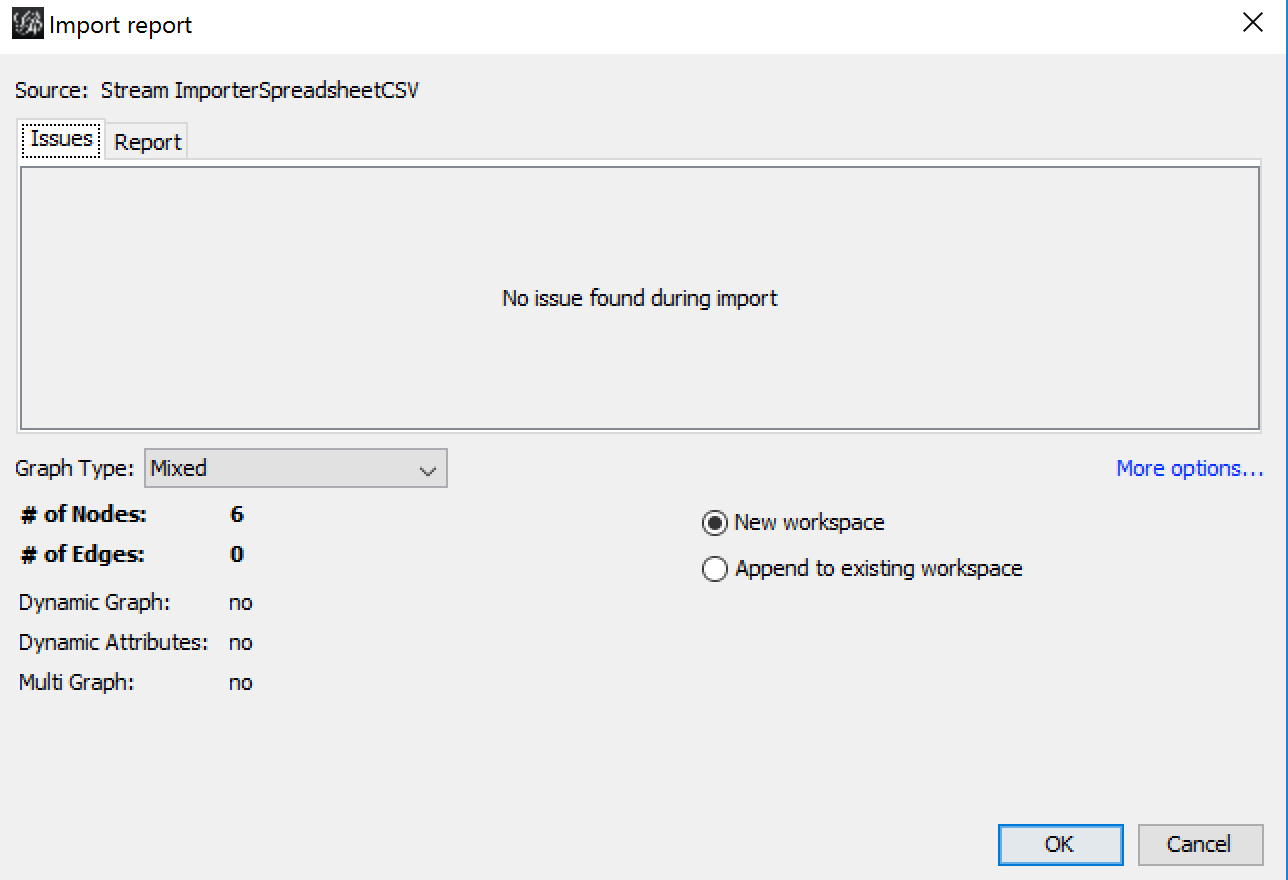
1. **Download three datasets from Blackboard**
   1. EdgeTable.csv
   2. NodeTable.csv
2. **Network visualization using Gephi**

* Open Gephi, click New Project in the window that pops out.

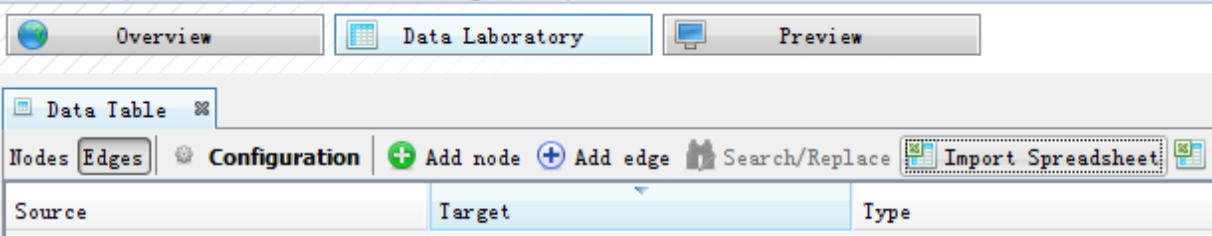


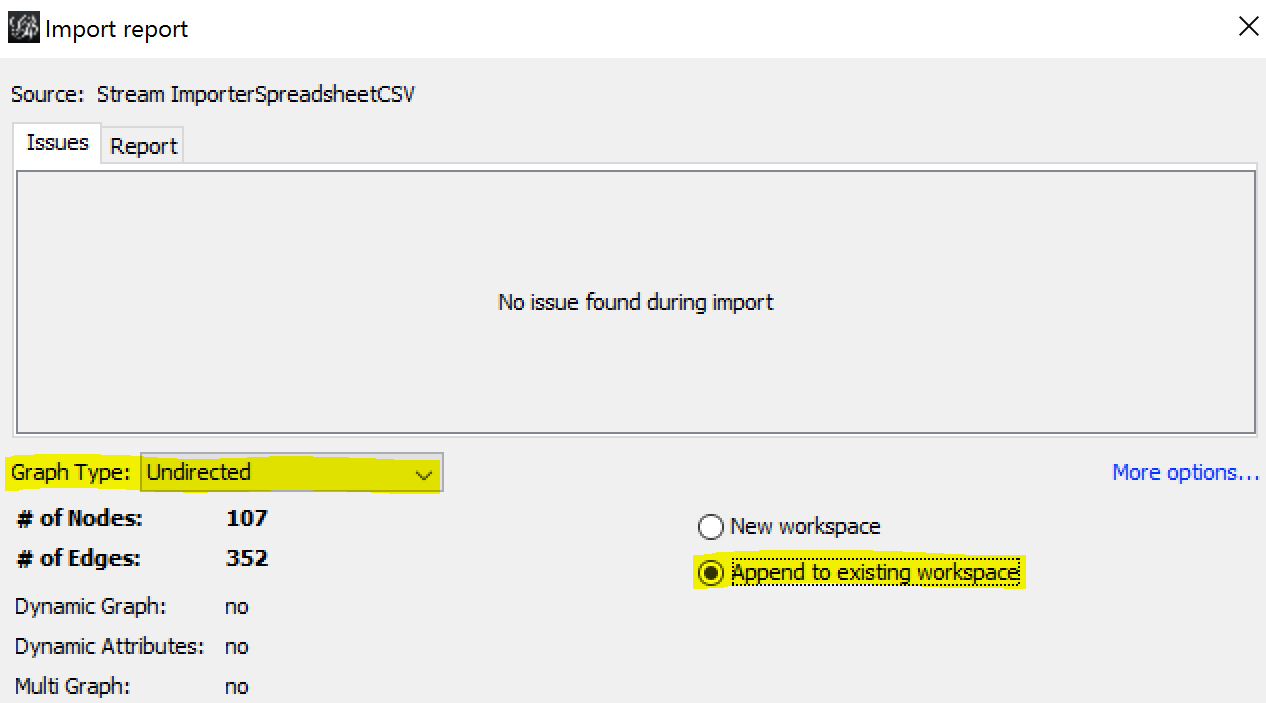
* Next, there are three tabs: **Overview**, **Data Laboratory** and **Preview**.
* First, we will import **nodes file**: Go to **Data Laboratory**, press **Nodes** button, and then click **Import spreadsheet** to import the NodesTable.csv file. Click OK to finish





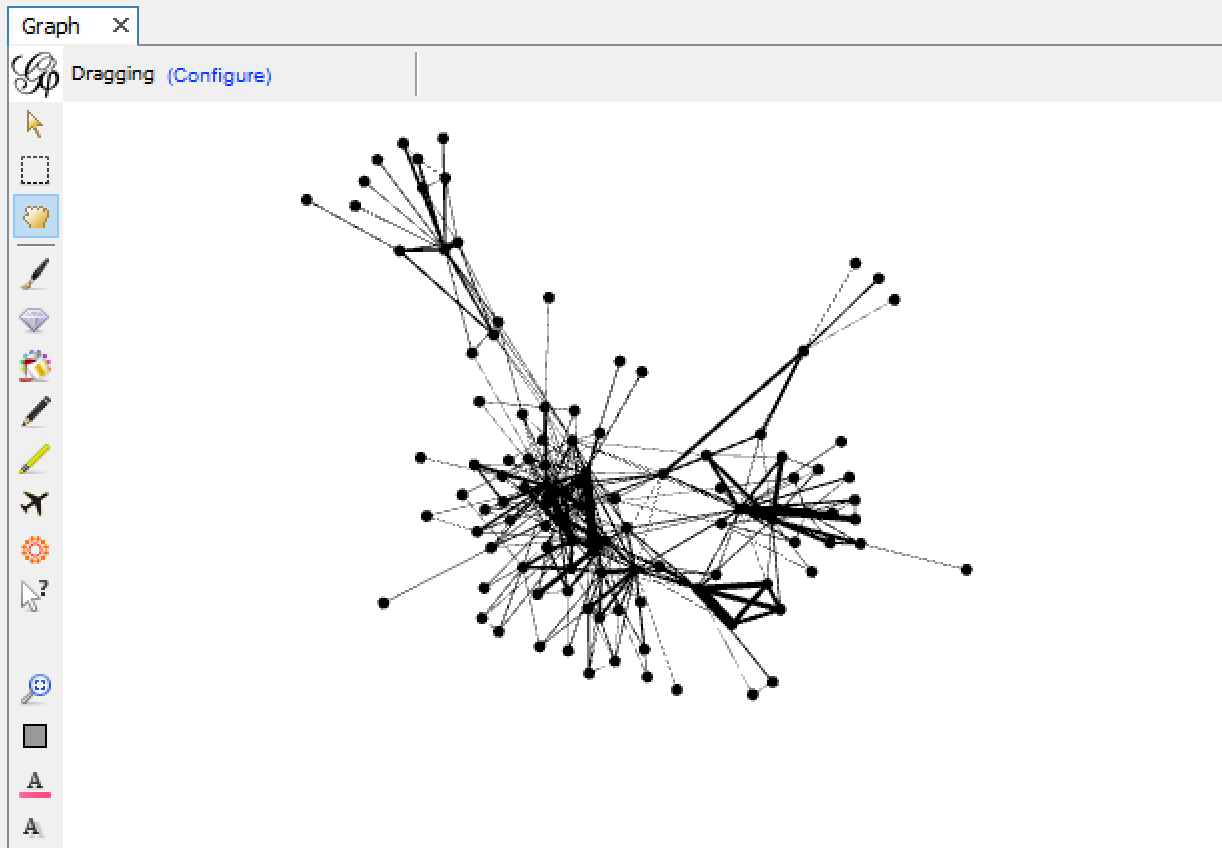
* Second, we will import **edges file**: Go to **Data Laboratory**, press **Edges** button, and then click **Import spreadsheet** to import the EdgesTable.csv file, click **Next**，, then click **Finish.** Remember to choose Append to existing workspace to add edge to existing nodes. (Based on your dataset, you could choose directed or undirected graph type)



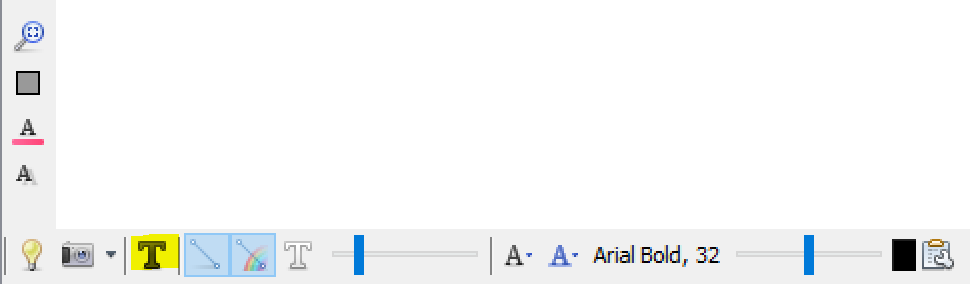


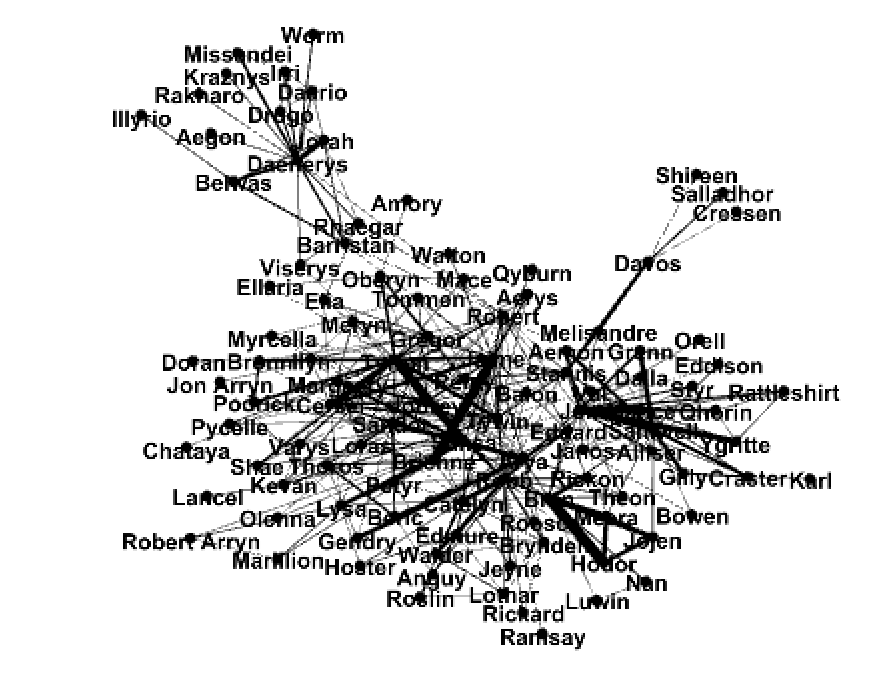
* The data is now imported. You can take a look at the data in the **Nodes**/ **Edges** in the **data laboratory**.
* Then go to the **Overview** Panel.
* In the Overview panel > Layout panel, choose **YIFAN HU proportional** as the layout and then click the **green Run button** to change the layout of the network. When you think the structure of the network is good enough, click the red stop button to **stop** the layout changing.

(layout explain: https://gephi.org/tutorials/gephi-tutorial-layouts.pdf)

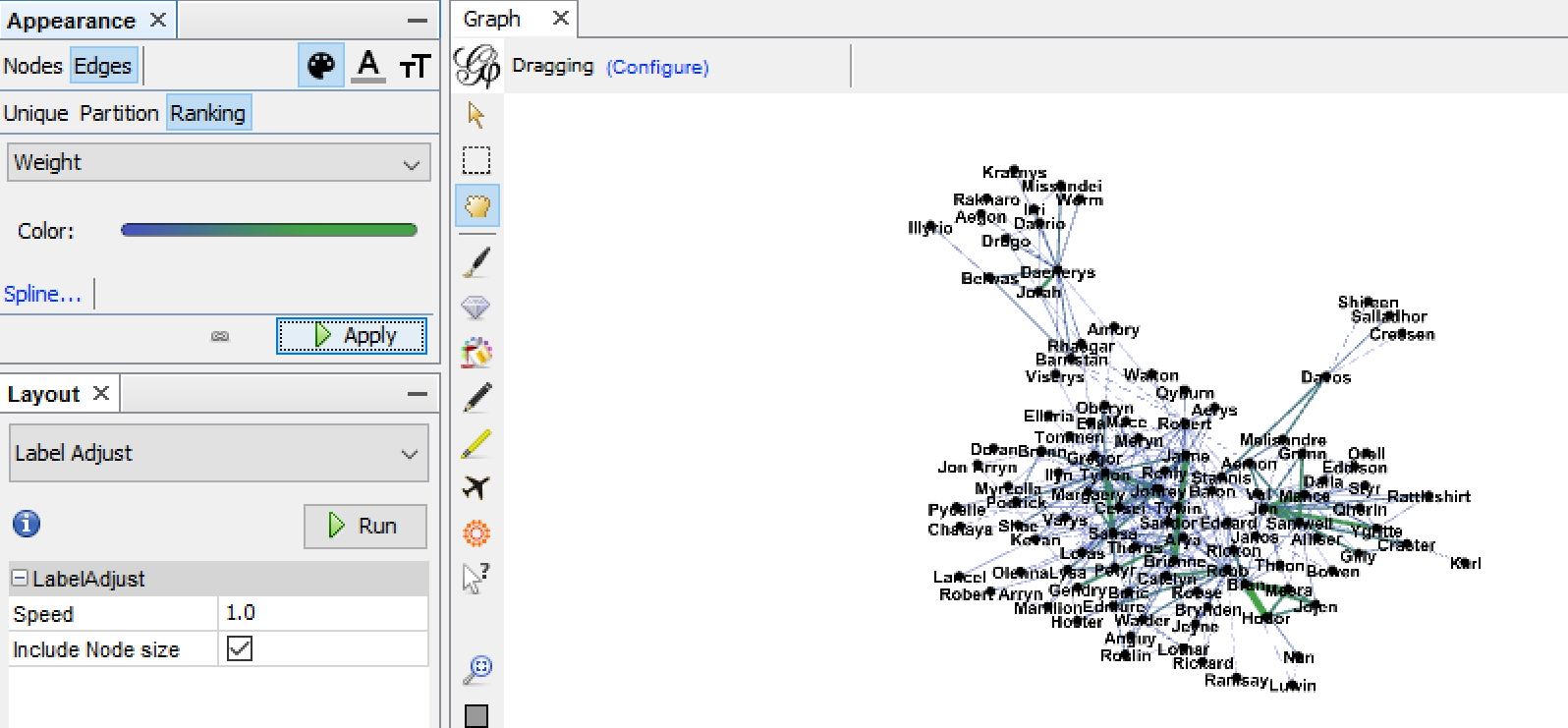


* Now, you can check the ‘**T’** button to show the label for each node. Then you should get a Network similar to this.



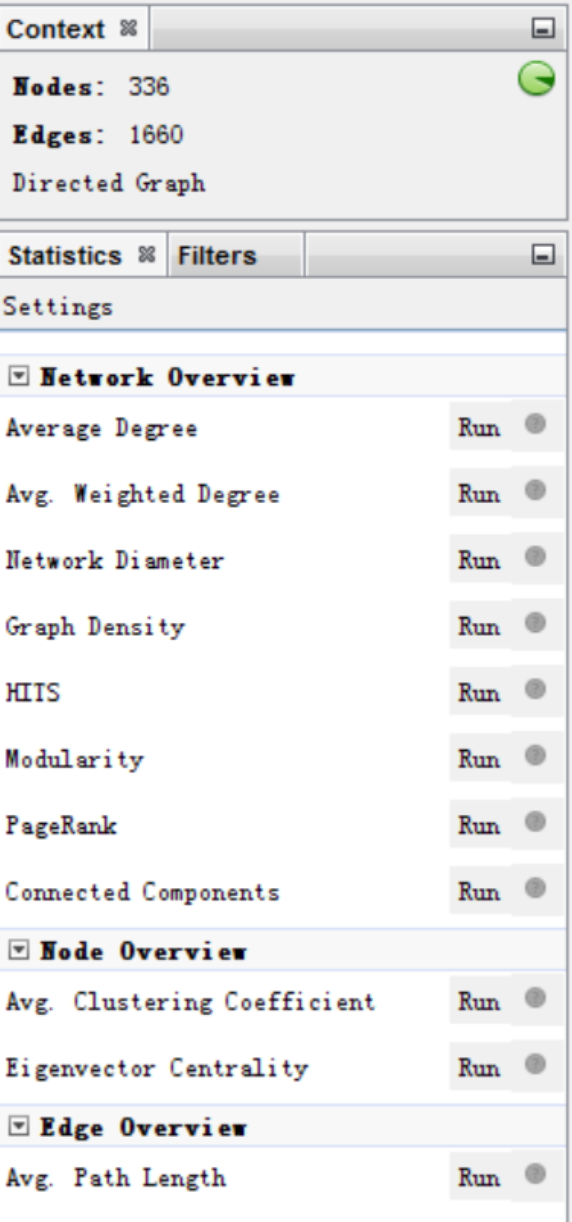


* You can use the graph pad to color different clusters with different colors, edit the display settings of edges and nodes, or zoom in and out to see the nodes more clearly.



1. **Statistics calculation using Gephi**

* There are a bunch of metrics/statistics for network, nodes and edges listed on the right side of the graph pad. You can click the Run button for the metrics that you want to calculate for your network.



* Let’s click the Run button for **PageRank, Average Degree, Avg. Weighted Degree,** and **Network Diameter**. Use the default setting when the settings window pops out and click OK. For each metric, you will get a report window, click close to continue.
* Then go to Data Laboratory tab, you will find that all the metrics you just ran are appended in columns in the data laboratory. You can also click the head of a column to sort the table by this column.

