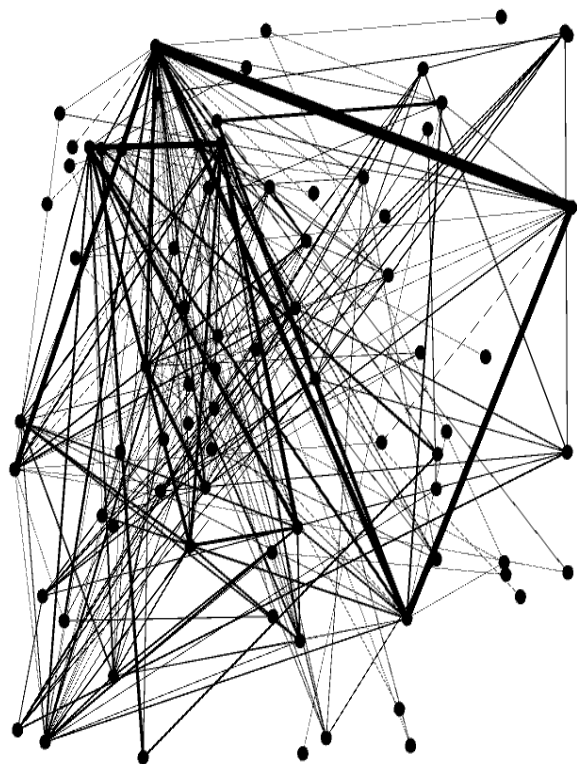


Analysis of the Characters of Les Misérables

Christopher Tran

1) Purpose For Analyzing This Network

- The purpose of analyzing the network of characters within Les Misérables is to inform readers of the impact of having many or few connections.
- Showing the characters with the most links can predict who the most involved characters. The more involved they are can lead to predicting their influence on the plotline

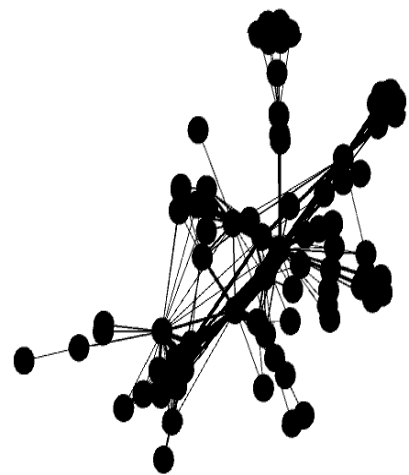


2) Original Layout

- What can you ascertain about the graph you see?
 - From what I can see from initially opening this graph, each node/character is in some way or another connected to at least one other character in the series. Even if there is only one connection, there are no nodes that have zero links.

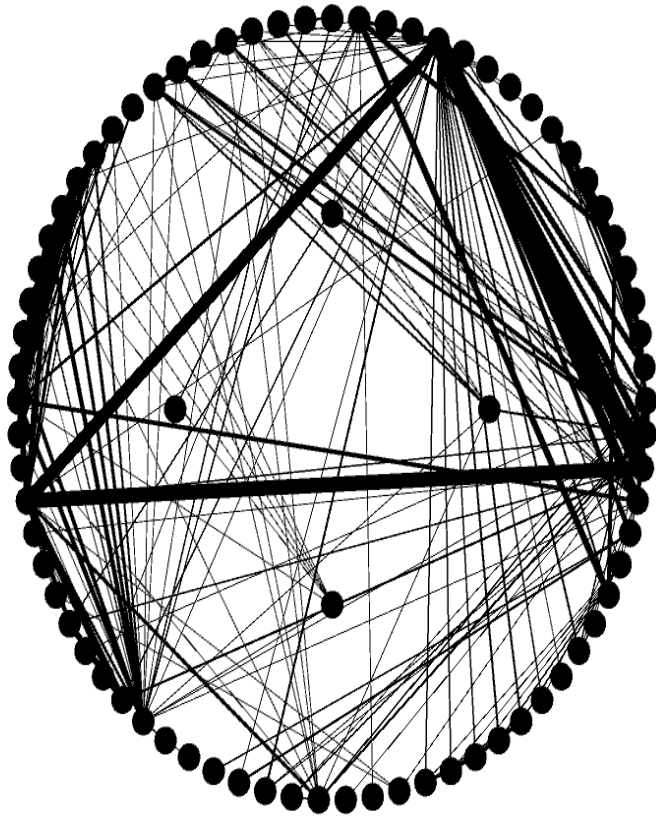
3) Running the algorithms

- Now run three to five Layout algorithms. Make sure to run the Yifan Hu algorithm as one of your choices.
 - Force Atlas 2
 - Dual Circle
 - Yifan Hu
 - Openord
 - Radial Axis



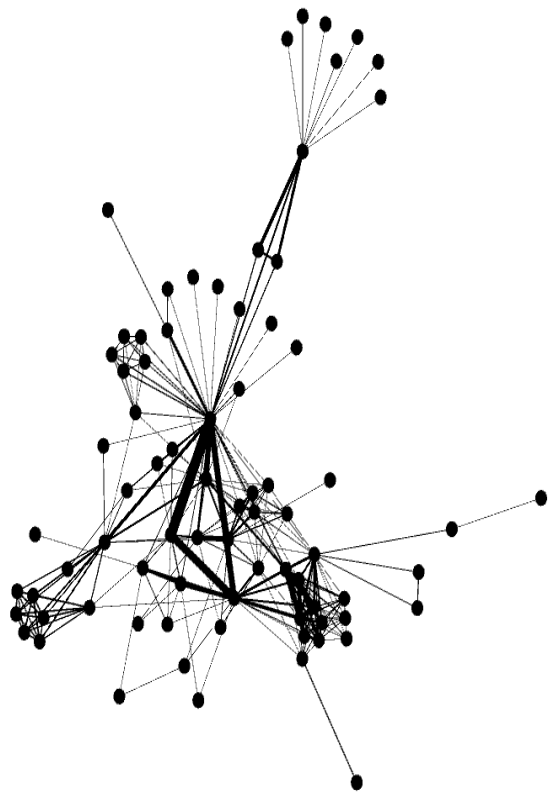
Force Atlas 2 Layout

- Force Atlas 2 layout causes the nodes cluster into their own separate communities



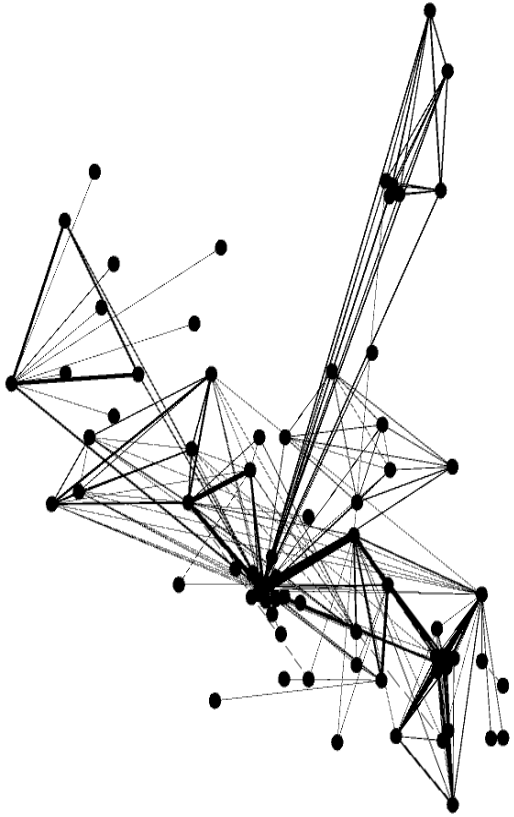
Dual Circle Layout

- The Dual Circle layout forces majority of the nodes into a circle which highlights the degree of each node and weight of each edge.



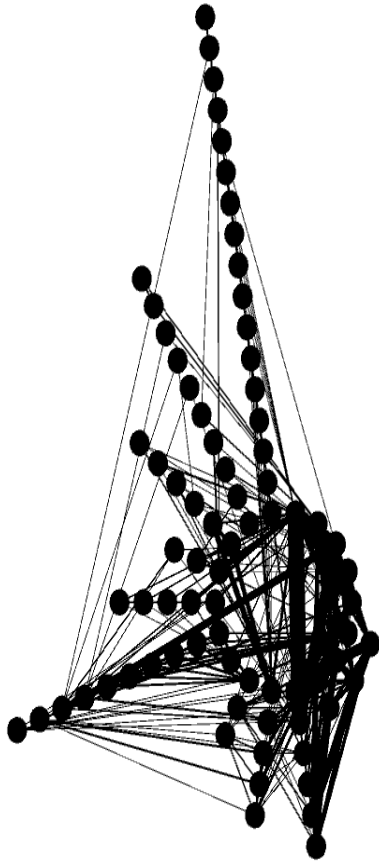
Yifan Hu Layout

- The Yifan Hu Layout spreads out the nodes based on the degree of each node with the higher degree nodes at the core of the network. This is especially good for identifying the lower degree nodes that have only one connection toward the outside.



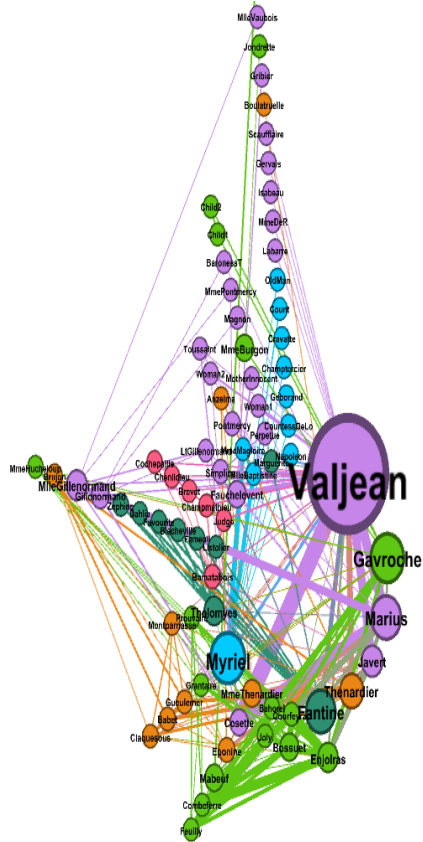
Openord Layout

- Openord Layout seems to group the communities together with the nodes with the highest degrees toward the center of each community.



Radial Axis Layout

- The Radial Axis layout seems like it takes the nodes with the most connections and puts them all in a circle while uniformly lining up the other linked nodes along the outside.



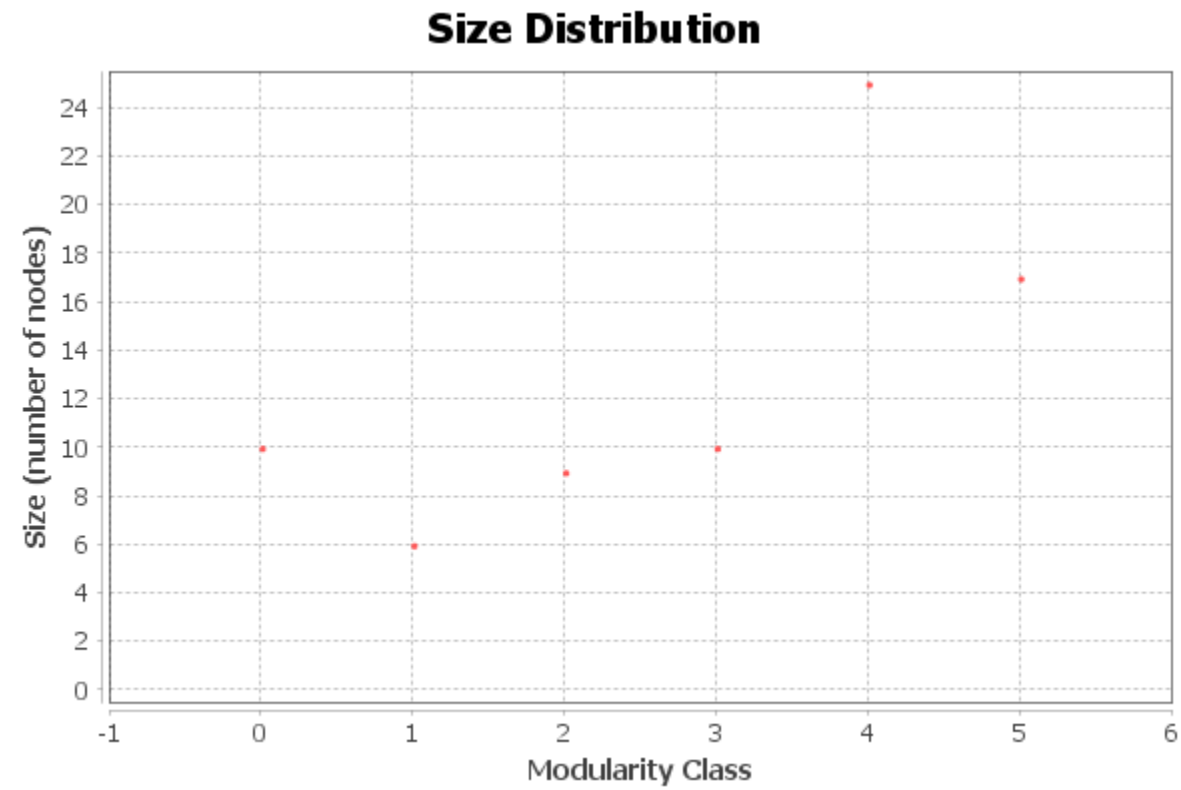
4) Radial Axis Layout W/ Color

- Adding color to the Radial Axis layout seems highlights the what I predicted about nodes with the most connections. The higher degree nodes are highlighted by their size which is also complimented by the stack of connections along the outside.

5) Applying Statistics

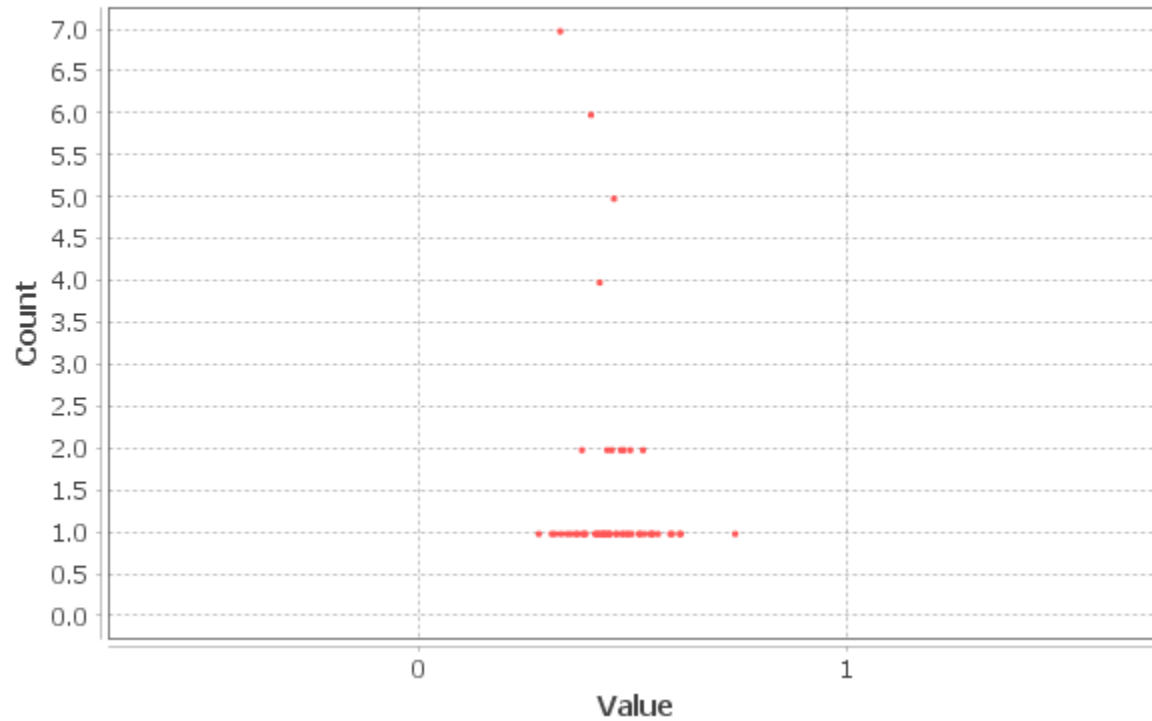
- Now run some statistics about the network such as connectivity, centrality, clustering, etc.
- Explain what statistical results you found. Please show graphs and/or numbers
 - After running various statistics through Gephi, I found that the ones that were most relevant to the Les Mes network were the modularity, network diameter and the average weighted degree.
 - Modularity showed how the character nodes clustered up based on how many connection each node had.
 - Network Diameter was important to show because it should the maximum number of jumps there were at the widest point. The six degrees of separation rule applies here because the maximum jumps in this network was only five.
 - The average weighted degree is important to show how involved a character is based on how many times they appear. This datapoint shows how relevant the character is in that network and throughout the musical.

Modularity

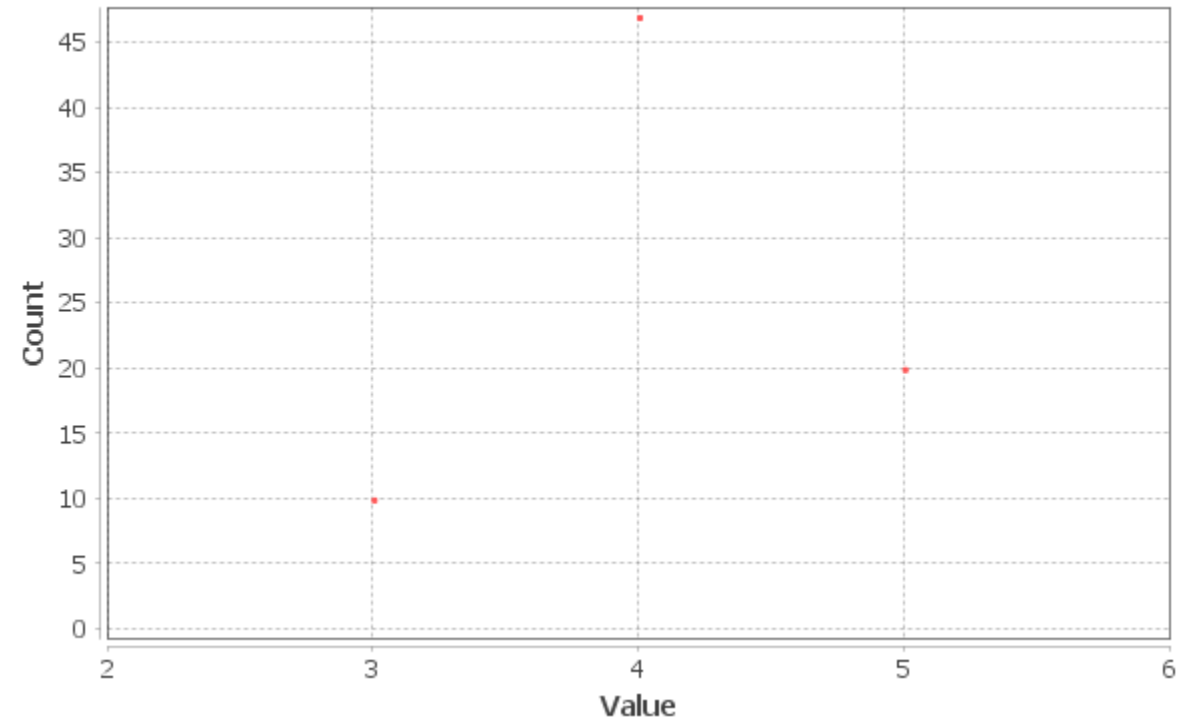


Network Diameter

Harmonic Closeness Centrality Distribution

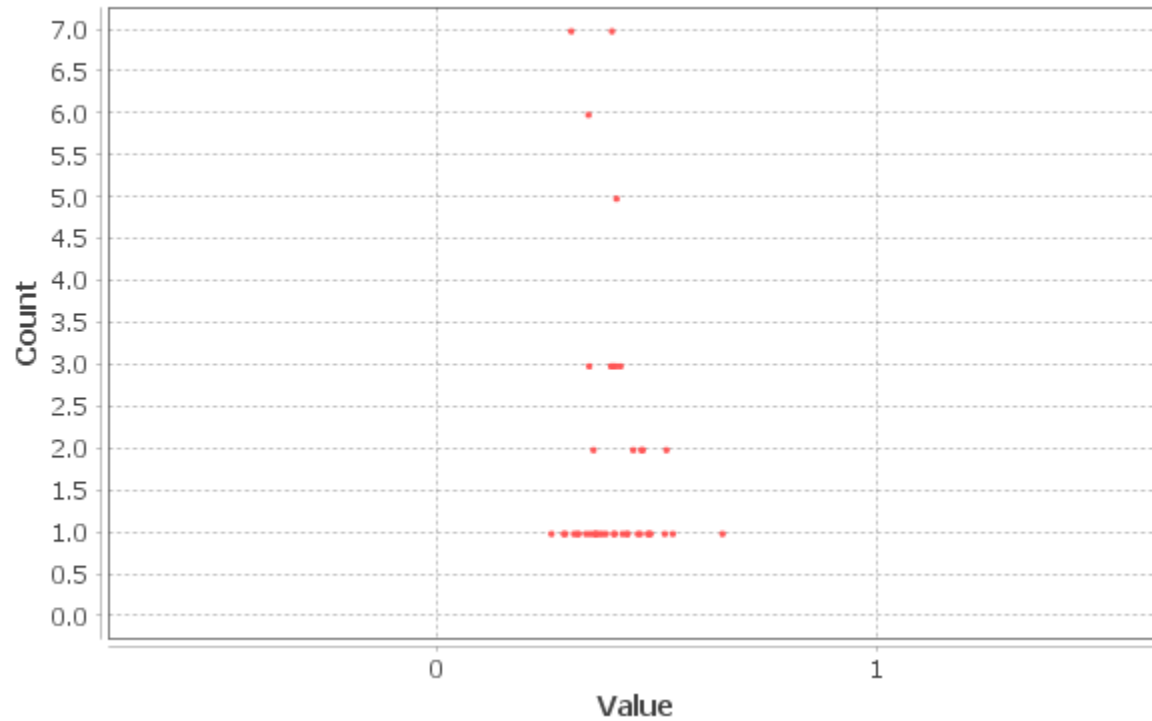


Eccentricity Distribution

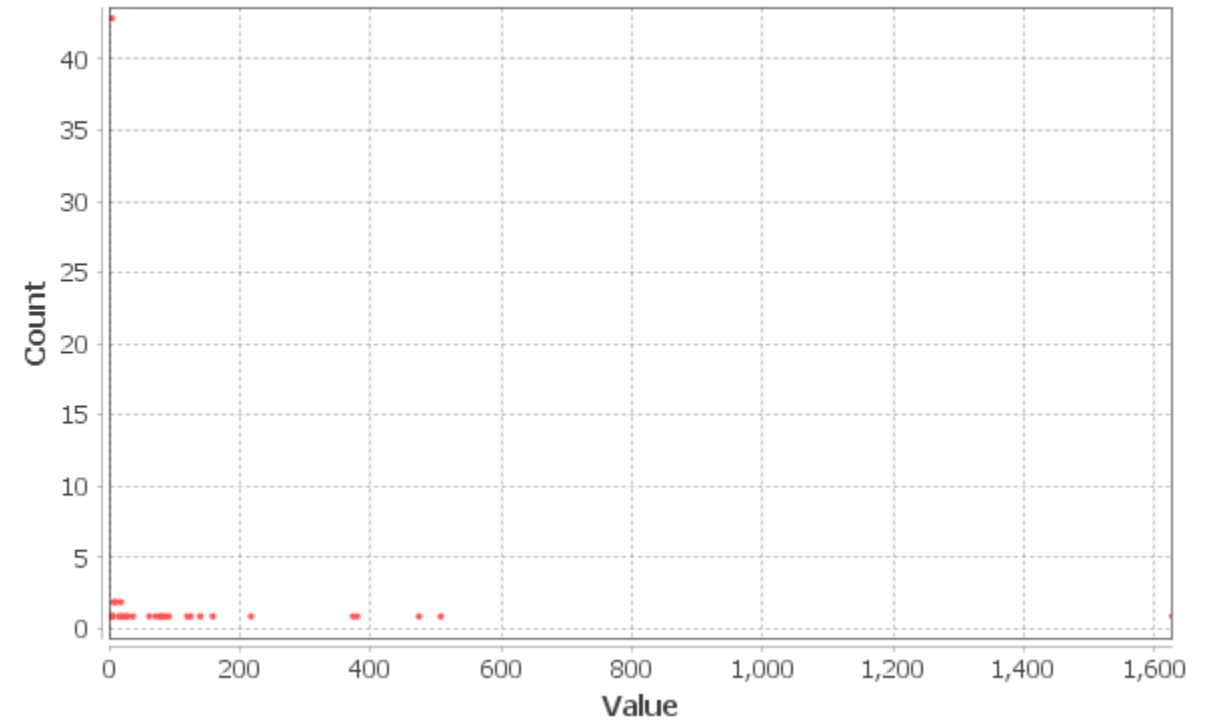


Network Diameter

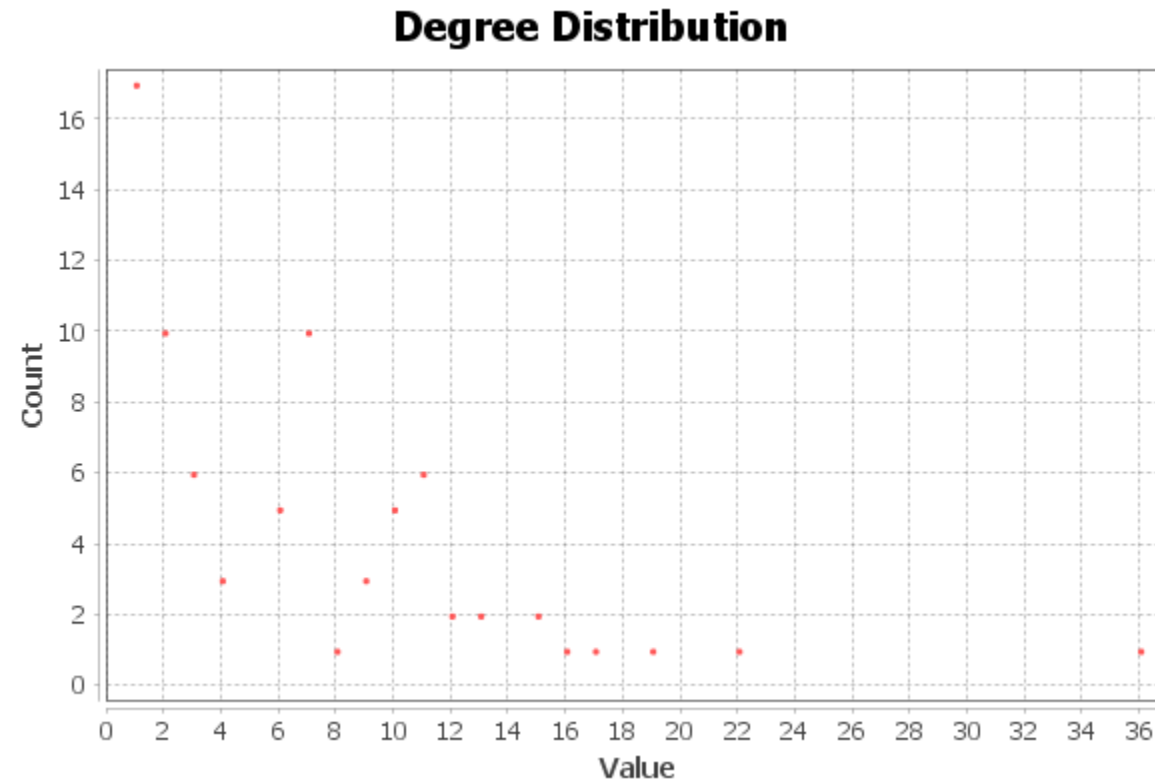
Closeness Centrality Distribution



Betweenness Centrality Distribution



Average Weighted Degree



5) Applying Statistics

- Are there any communities or giant components, please explain?
 - Communities based on the modularity, as modularity splits the nodes into clusters based on how many connections they have. Another thing that should be noted is the weight of each connection. The weight of each connection or the thickness of each line shows how many times the characters interact with each other. The thicker the line the stronger the community.

5) Applying Statistics

- Do you see any homophily or density in your network, please explain?
 - Looking at the network itself, there is no real homophily or density. The value of the graph density only came out to 0.087, being very low.

6) FINAL CONTEMPLATION

- Seeing the nodes grouped together based on their type of relationship or the type of interactions and appearances would have add another level of variety to the dataset. From there, you might be able to analyze how the group of nodes or communities would interact with the other communities.
- The Les Misérables network dataset was an interesting study to say the least. This assignment was a great way of learning how to visualize the different aspects within a network. Upon the initial upload of the dataset, it was hard to draw any conclusions or make any prediction based on the mess of the nodes and links. After running some of the layouts, adding names, colors and some weight to the degrees, it was easier to visualize the relations between the nodes and the communities that formed from them.