Exploring Networks using Gephi

Installations:

Download:

https://gephi.org/users/install/

Install on linux:

update Oracle Java (JDK and JRE):

```
sudo add-apt-repository ppa:webupd8team/java

sudo apt-get update

sudo apt-get install oracle-java8-installer
```

Install gephi:

Extract the downloaded file in a directory.

Run it by executing ./bin/gephi script file

Install on Windows:

Make sure the latest version of java jre is installed - if not, download here: https://www.oracle.com/technetwork/java/javase/downloads/jre8-downloads-2133155.html

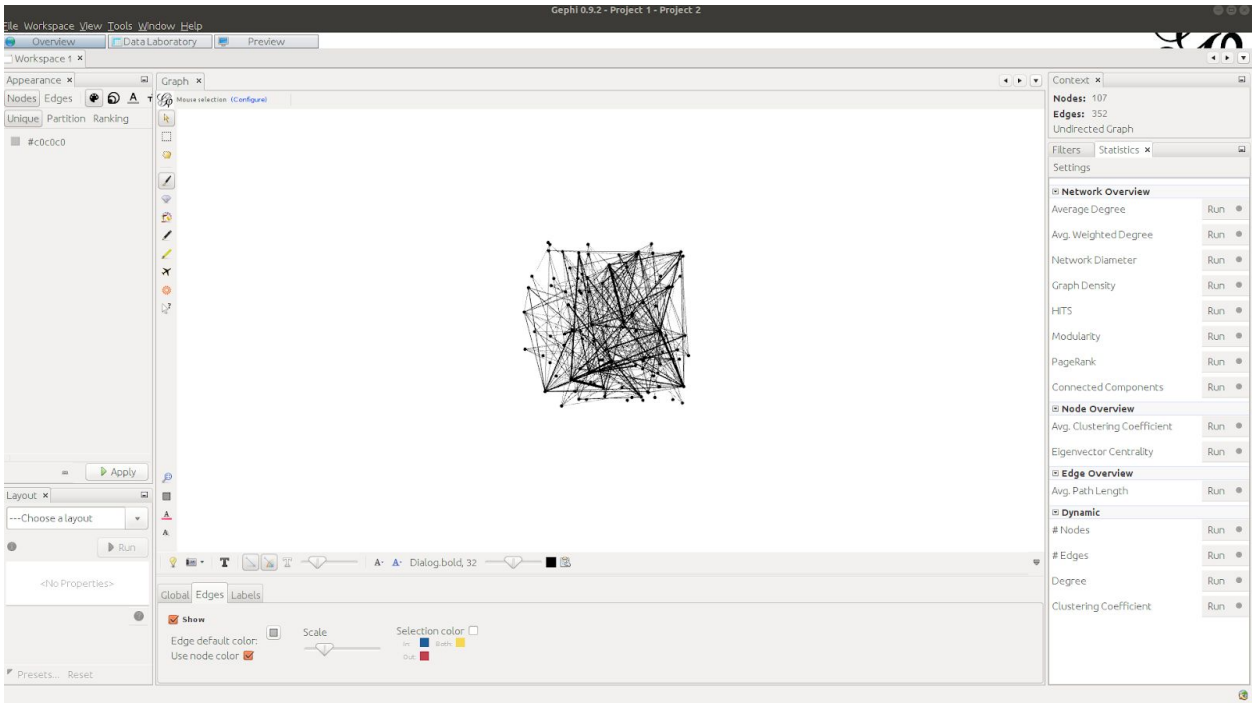
Install gephi:

Run the downloaded .exe file

1. Upload network data

- In the *file* dropdown menu select *open* and navigate to the data folder
- Select the stormofswords.csv file and click *OK*
- Make sure the *Separator* dropdown is pointing to “Comma” and the *Import as* dropdown is pointing to “Edges table”, make sure the Weight column is imported and click *Next*
- In the *Weight* dropdown menu select “Integer” and click *Finish*
- In the import report window select “Undirected” in the *Graph Type* dropdown menu and click *OK*

If your import was successful, you should see a graphical representation of your network on your screen



2. Calculating network properties with Gephi

- Go to the *Statistics* window (if it does not appear on the screen - go to the Window dropdown menu and select Statistics)

Degrees and degree distribution

- In the Network Overview section run Average Degree

  Gephi will generate an html report containing the average degree in the network and a png image of the network's degree distribution.

  2.1 what is the average degree of this network?

  2.2 what is the maximum and minimum degrees in the network?

  2.3 how many nodes have a degree equals to the minimum degree?

  2.4 how many nodes have a degree equals to the maximum degree?

  2.5 what can we say about the degree distribution of this network?

Distance metrics

- In the Network Overview section run Network Diameter
- In the Graph Distance Setting window, select Undirected and check Normalize Centralities, click OK

  Gephi will generate an html Graph Distance Report containing the diameter, radius and average path length of the network. It will also generate the Betweenness Centrality, Closeness Centrality and Eccentricity distribution images for the network

  2.6 what is the diameter of the network?

  2.7 what is the radius of the network?

  2.8 what is the average path length of the network?

Clustering Coefficient

- In the Node Overview section run Clustering Coefficient

  Gephi will generate an html Clustering Coefficient Metric Report containing the average clustering coefficient of the network (global) and the total triangles. It will also generate a clustering coefficient distribution image

  2.9 what is the average cluster coefficient of the network?

3. Exploring nodes and edges

- Click on the Data Laboratory tab
- Click on the Nodes button if it is not already clicked
- Sort the table by the Degree column

    3.1 list the five nodes with the highest degree

- Do the same with the Eccentricity, Closeness Centrality and Betweenness Centrality columns

    3.2 What is the minimum eccentricity of the network? which node has the lowest eccentricity?

    3.3 What is the closest centrality value of the node with the lowest eccentricity?

    3.4 What is the closest centrality rank of the node with the lowest eccentricity?

    3.5 What is the betweenness centrality value of the node with the lowest eccentricity?

    3.6 What is the betweenness centrality rank of the node with the lowest eccentricity?

    3.7 What is the PageRank value of the node with the lowest eccentricity?

    3.8 What is the PageRank rank of the node with the lowest eccentricity?

    3.9 which are the nodes with the highest clustering coefficient? What are the highest and lowest degrees of these nodes and how can this be interpreted?

    3.10 explore the authority and hub scores - what do these score mean and how can we interpret them for this network?

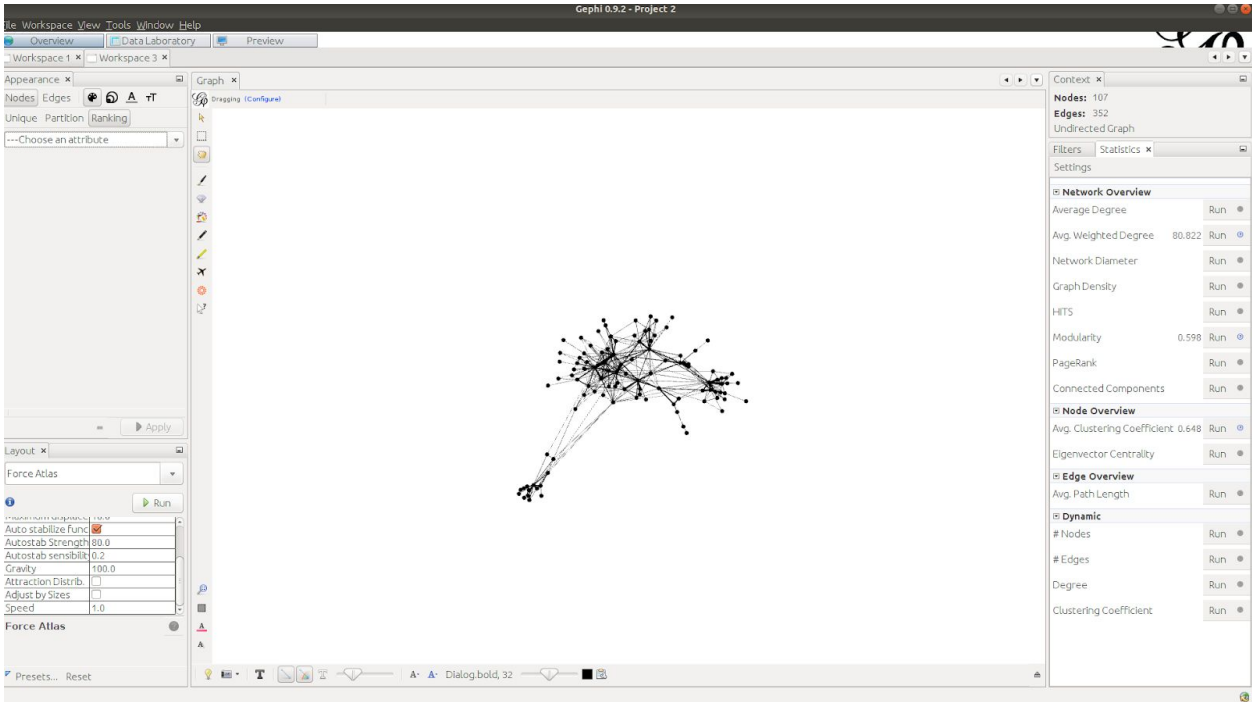
- Click on the Edge button
- Sort the table by Weight

    3.11 which edge in the network has the highest weight and what is that weight?

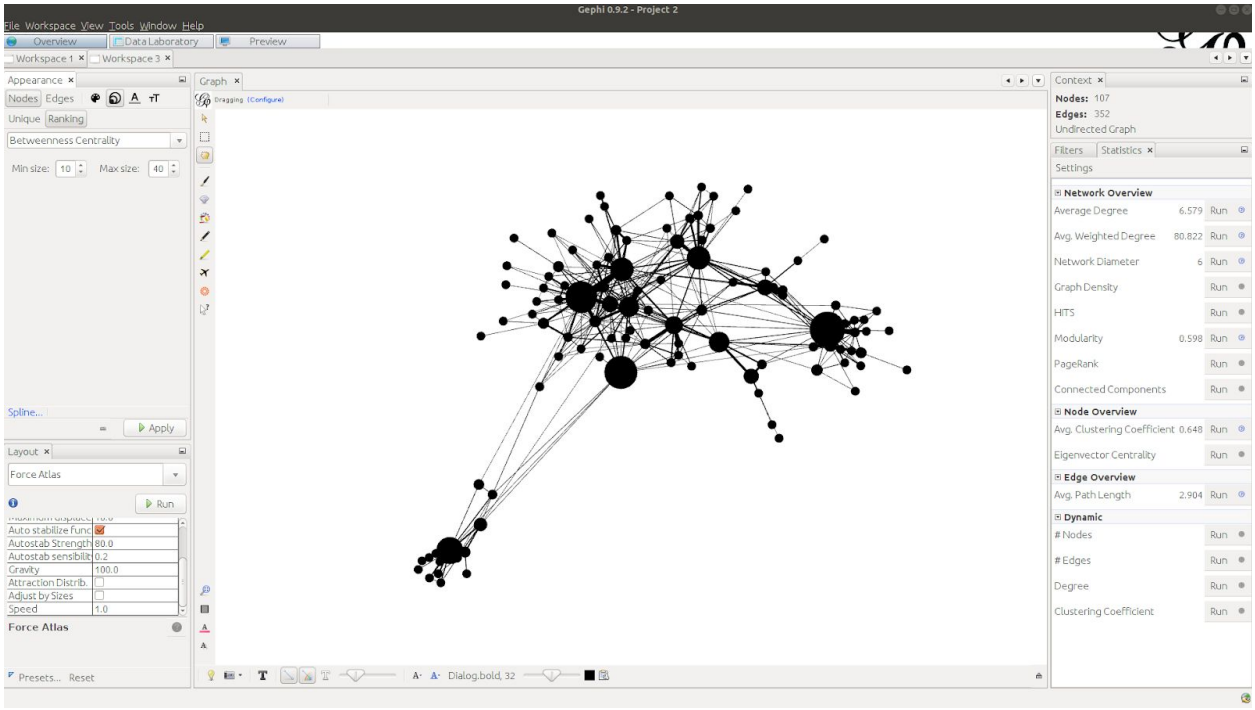
4. Visualization

- In the layout pane, select “Force Atlas” from the dropdown nemu

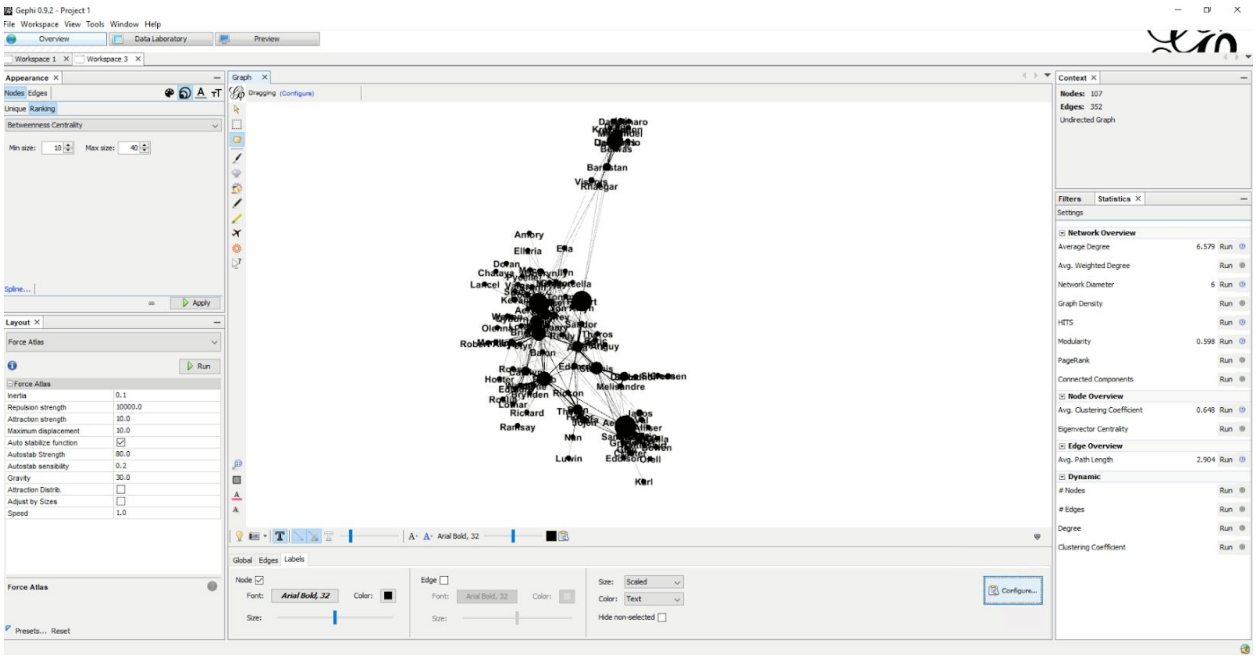
  Force Atlas layout is used for visualizing real-world networks and small-world networks and is efficient in spatially emphasizing groups and communities in a network
- Set the repulsion strength to 10000 - so that the nodes are drawn apart from one another
- You can play around with other parameters to see if that improves the network layout
- Click run and then stop



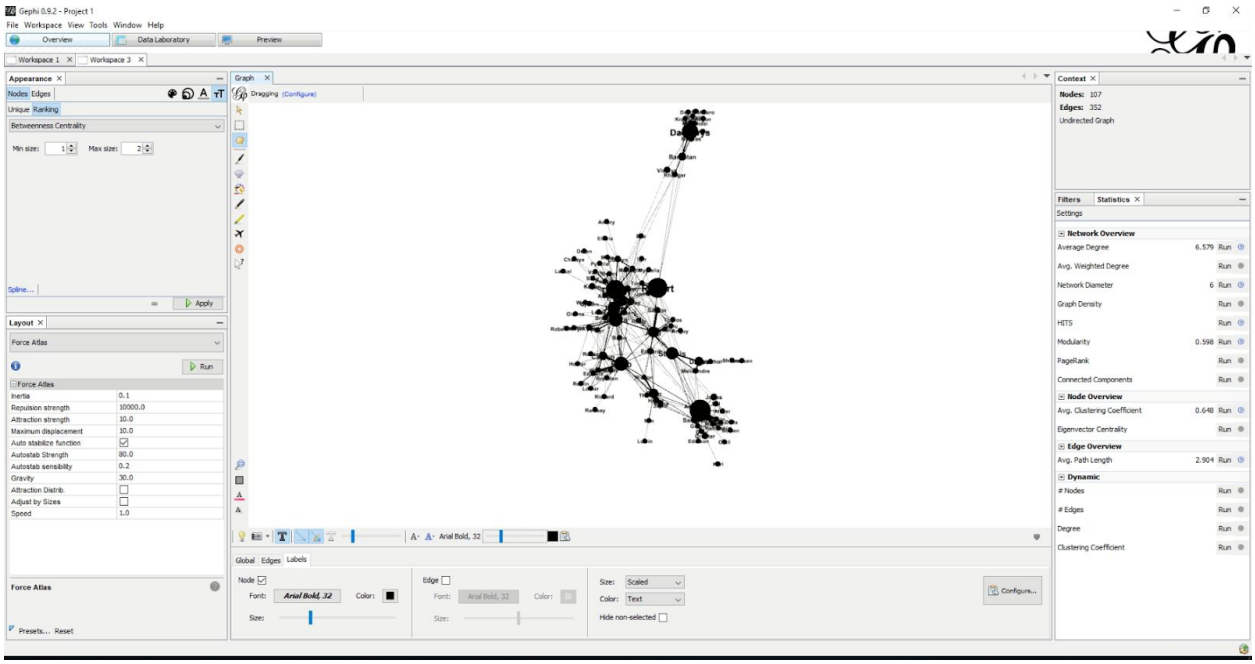
- In order to rank our nodes - lets create a rank based on a certain parameter. The parameters Gephi makes available for us are the ones that we calculated in the statistics pane. If you are starting a session from scratch - calculate the desired metrics again as you previously did in section 2 of this document
- Choose a matric and click on the ranking button in the Appearance pane
- Select that metric in the dropdown menu
- In order to rank the node size click on the size button and play around with the min and max sizes until the result pleases you



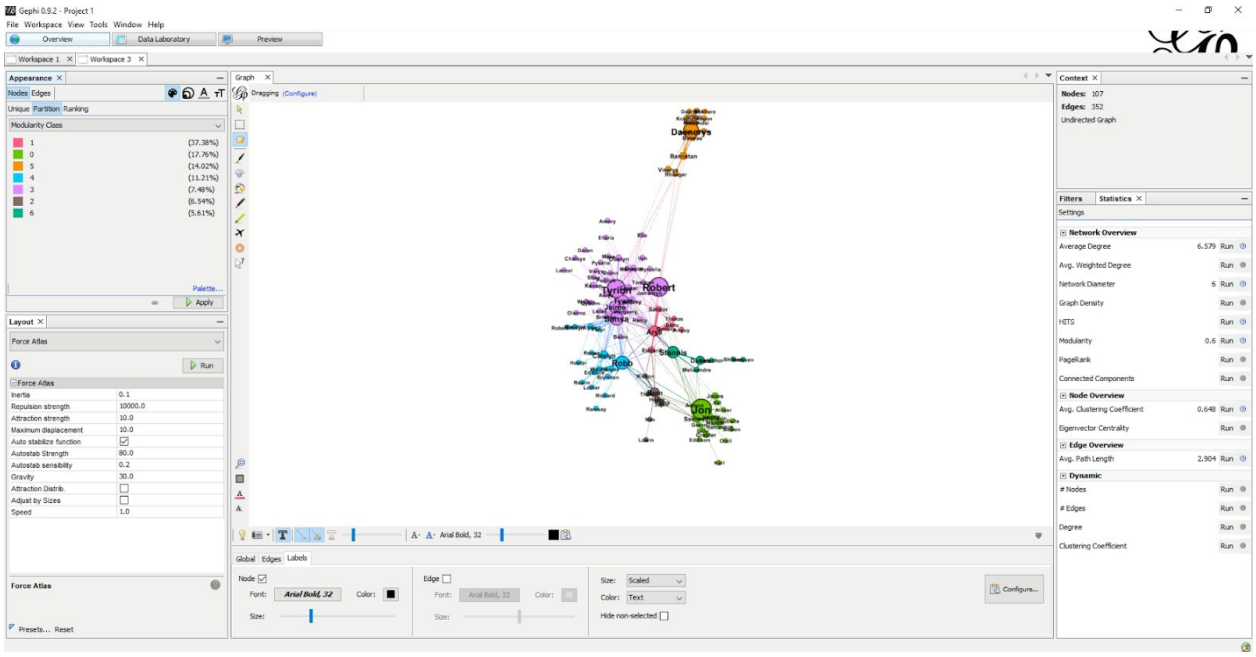
- To show the nodes' labels click on the T icon at the bottom of the graph window, if nothing shows, click on the small arrow the right and click configure. Then click on Id as the field to label the nodes by



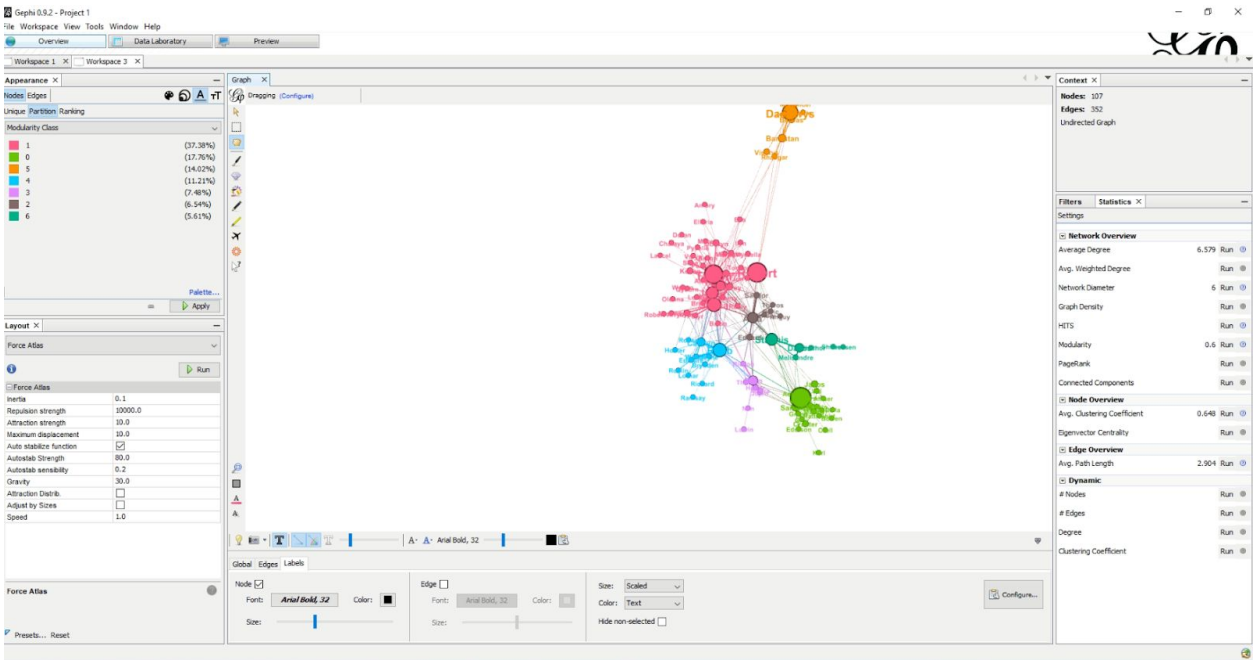
- Scale the labels by a matric: go to the Appearance pane and click on the nodes and then on the ranking button. Click on the the label size button - select your matric from the dropdown menu and scale



- Now let's color the nodes by modularity class: make sure to run the Modularity analysis by clicking on the Modularity button in the statistics pane first
- In the appearance pane click on partition, click the color button and select modularity class from the dropdown menu, click apply



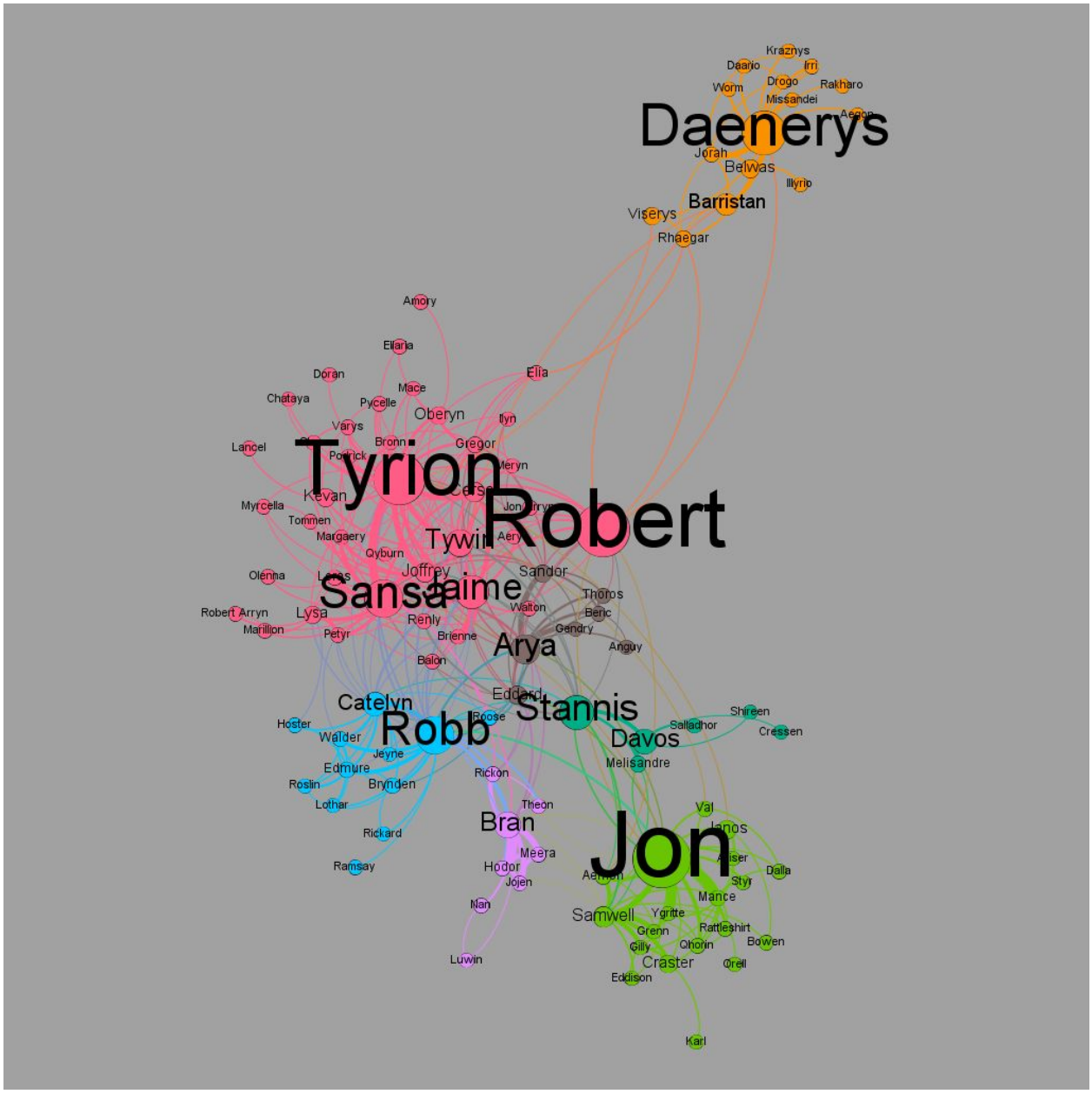
- You can also do the same with the labels, clicking on the Label Color Button



- To further adjust the label positions check the adjust by sizes box in the layout pane
- Some labels are still overlapping. In order to fix this, select label adjust from the layout pane dropdown menu and click run.

Creating print-ready images

- Click on the preview tab and select Default Curved from the dropdown menu, click refresh
- In the node labels menu check show labels
- The network appears without labels. This is because we have to copy our Id column into the label column. To do that, go to the data laboratory tab, click on copy data to other column, select Id and select to copy the data to the label column
- Return to the preview tab and refresh
- Adjust the font size, outline and any other styling
- Change the edges thickness and rescale
- Change the background color from the button on the bottom of the screen
- Finally, click Export at the bottom of the screen and save the image



- Save your project