

# Gephi – network analysis and visualization

2019 March 6  
Eric Johnson  
Mark Dahlquist

# Goals for today

- Learn about network analysis
- Practice using Gephi for network visualization

# Schedule

- Overview of network analysis
- Tools for network analysis
- About Gephi
- Installing Gephi
- How to use Gephi
- Practice with Gephi

# Networks



Image: [https://cdn.pixabay.com/photo/2015/05/24/14/02/spaghetti-781795\\_960\\_720.jpg](https://cdn.pixabay.com/photo/2015/05/24/14/02/spaghetti-781795_960_720.jpg)

# Networks



Image:

[https://assets.epicurious.com/photos/55f72d733c346243461d496e/2:1/w\\_1260,h\\_630/09112015\\_15minute\\_pastasauce\\_tomato.jpg](https://assets.epicurious.com/photos/55f72d733c346243461d496e/2:1/w_1260,h_630/09112015_15minute_pastasauce_tomato.jpg)

# Nodes and Edges

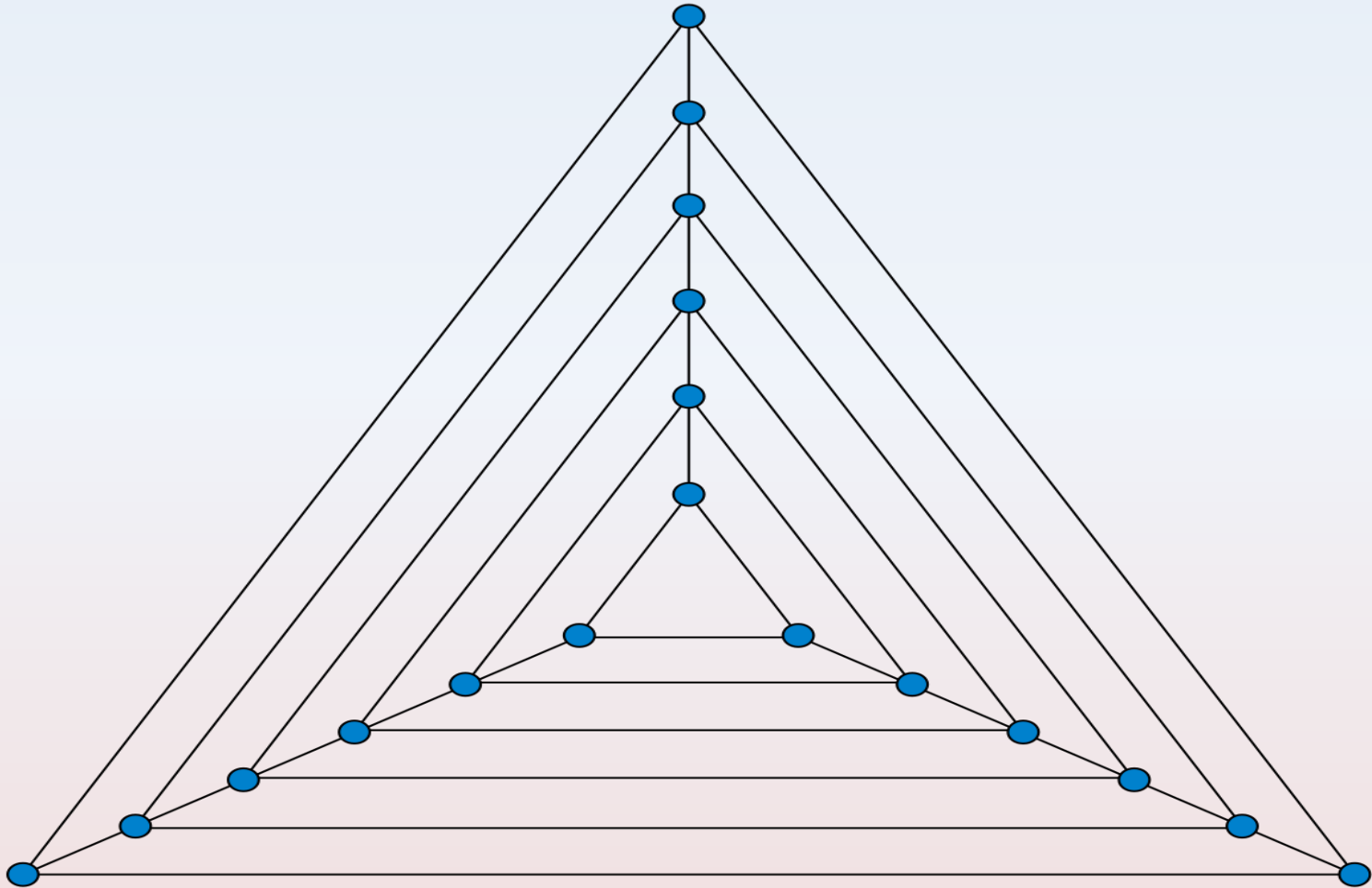
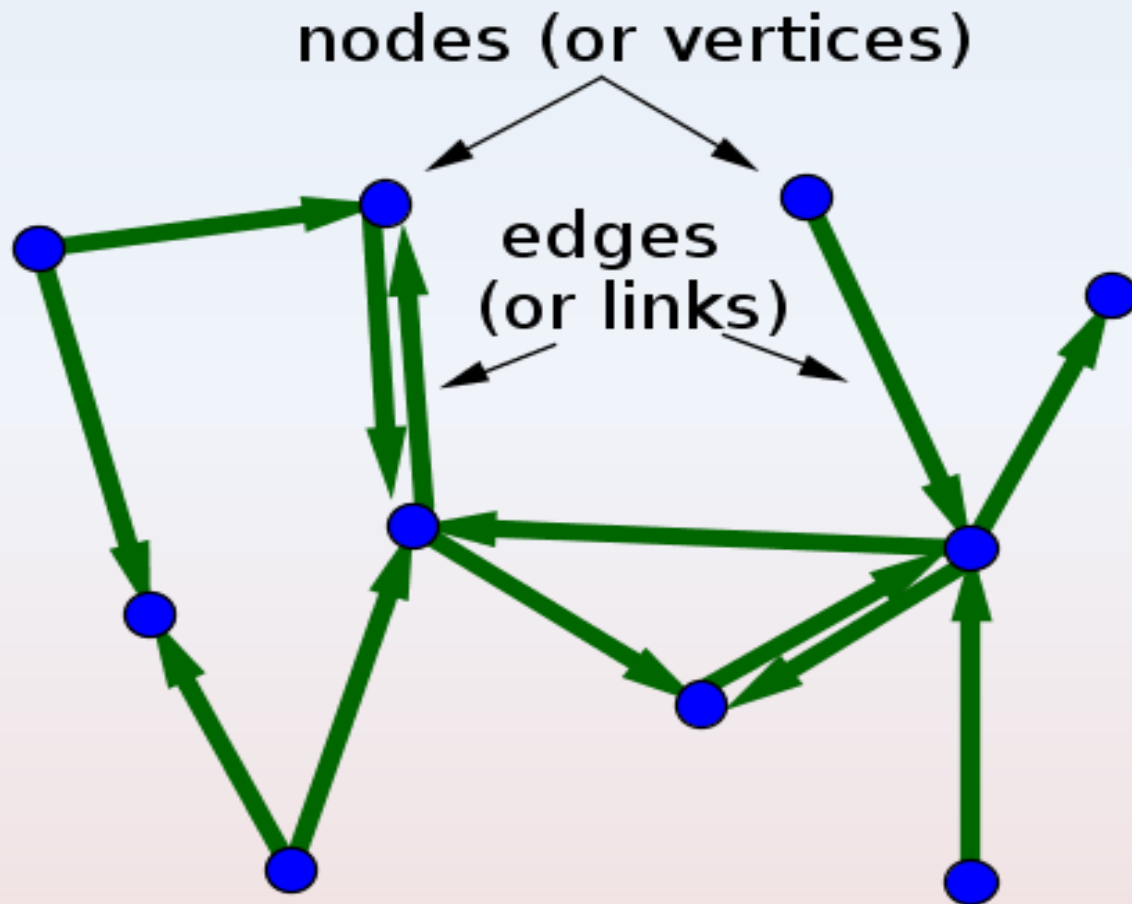
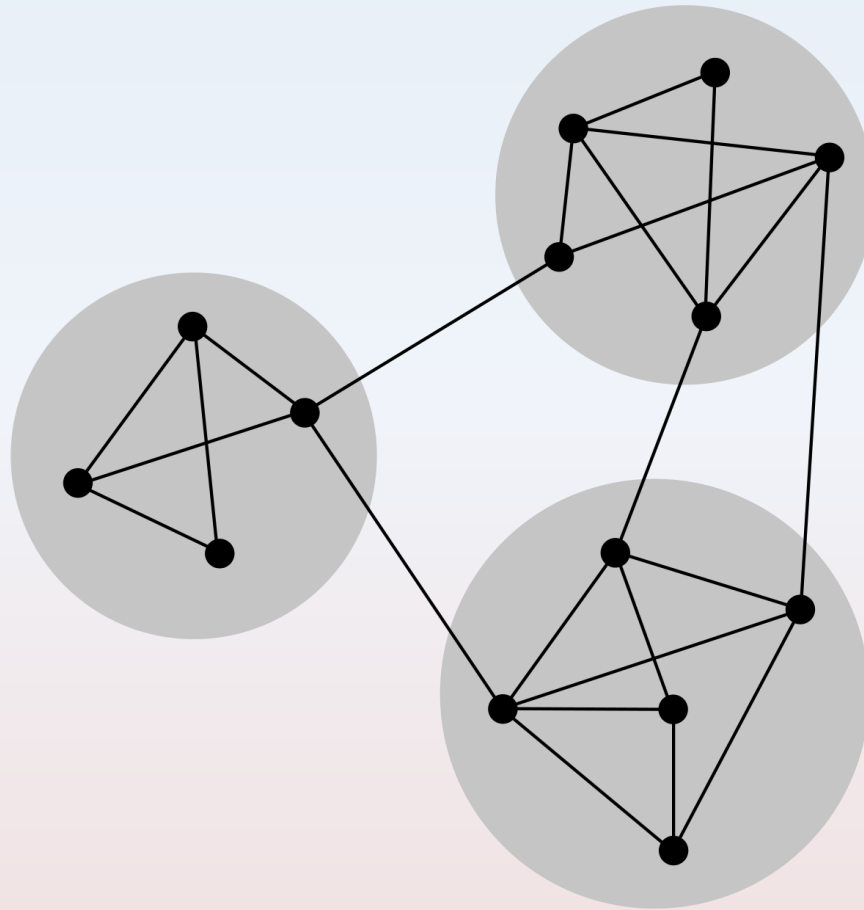


Image: [https://upload.wikimedia.org/wikipedia/commons/thumb/2/28/Nested\\_triangle\\_graph\\_18.svg/1200px-Nested\\_triangle\\_graph\\_18.svg.png](https://upload.wikimedia.org/wikipedia/commons/thumb/2/28/Nested_triangle_graph_18.svg/1200px-Nested_triangle_graph_18.svg.png)

# Directed network

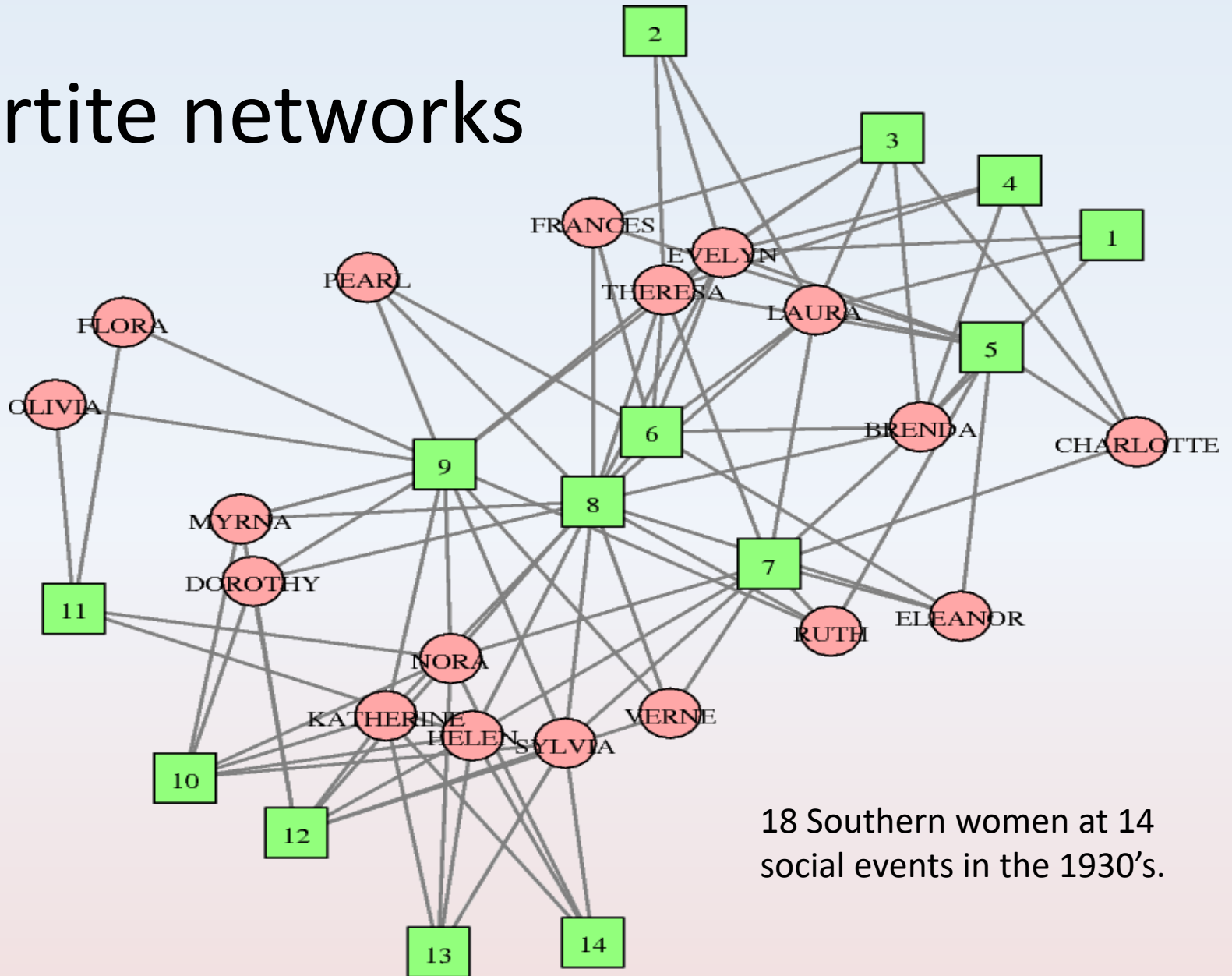


# Clusters



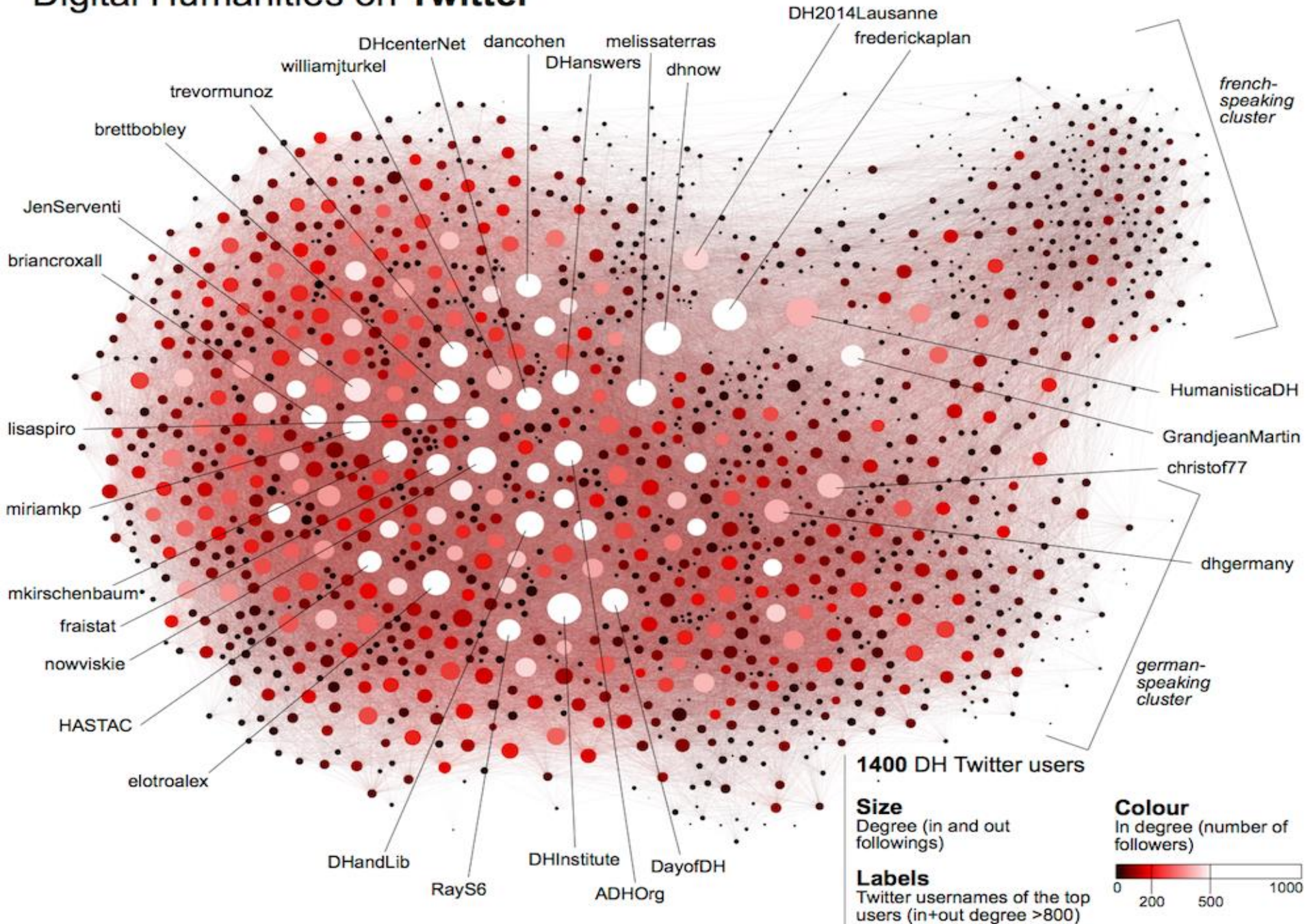


# Bipartite networks



18 Southern women at 14 social events in the 1930's.

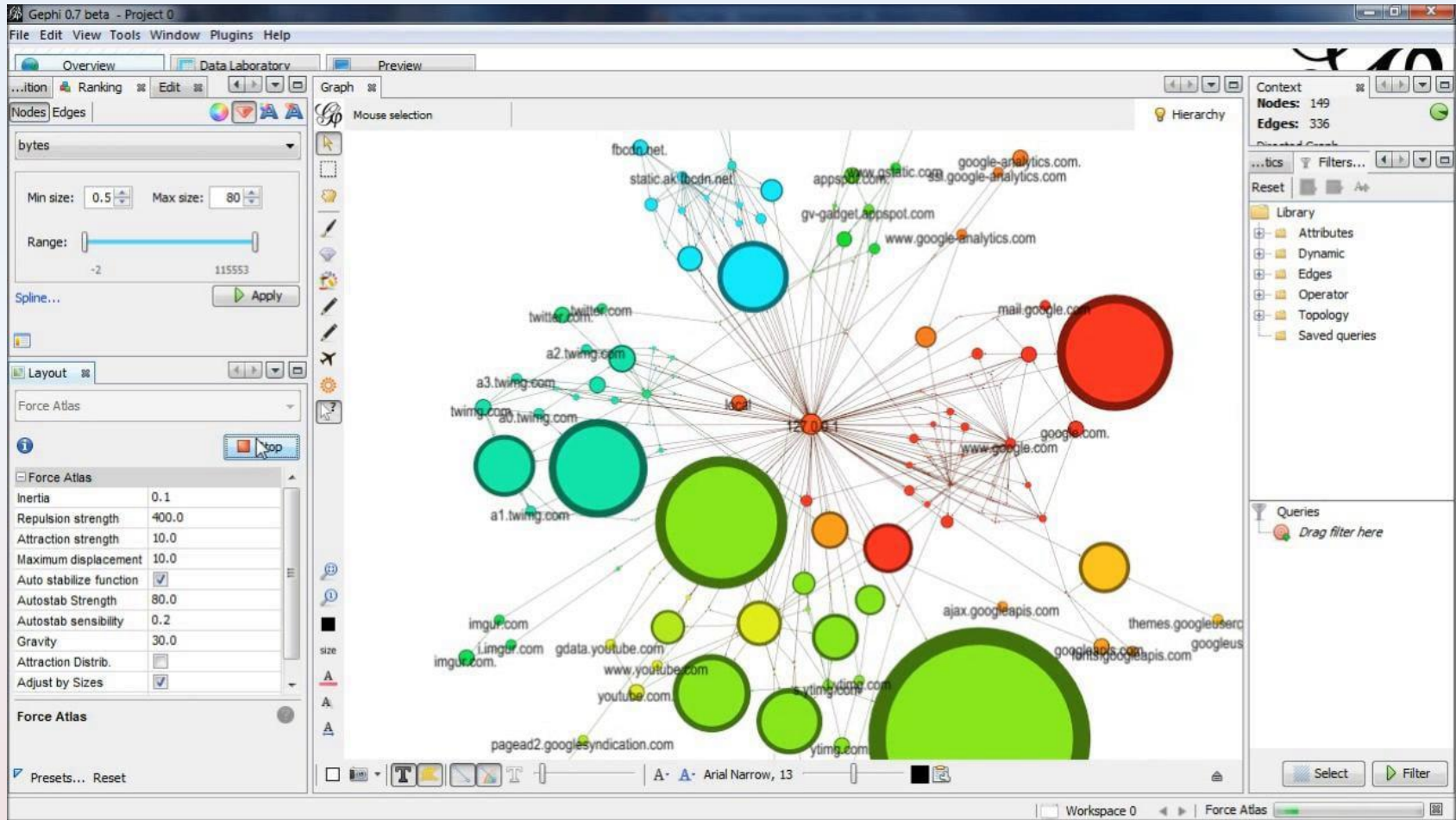
# Digital Humanities on Twitter



# Free Tools for Network Analysis

- Commetrix - <http://www.commetrix.de/>
- Cuttlefish - <https://github.com/dev-cuttlefish/cuttlefish>
- Cytoscape - <https://cytoscape.org/>
- EgoNet - <https://sourceforge.net/projects/egonet/>
- Gephi - <https://gephi.org/>
- Graph-tool - <https://graph-tool.skewed.de/>
- GraphChi - <https://github.com/GraphChi/graphchi-cpp>
- Graphviz - <http://graphviz.org/>
- JUNG - <http://jung.sourceforge.net/>
- Netlytic - <https://netlytic.org/home/>
- Network Workbench - <http://nwb.cns.iu.edu/>
- NetworKit - <https://networkit.github.io/>
- NetworkX - <http://networkx.github.io/>
- NodeXL - <https://archive.codeplex.com/?p=nodexl>
- Pajek - <http://mrvar.fdv.uni-lj.si/pajek/>
- R + visualization libraries - <https://www.r-project.org/>
- SocNetV - <https://socnetv.org/>
- Socioviz - <http://socioviz.net/SNA/eu/sna/login.jsp>
- Statnet - <http://statnetproject.org/>
- Tulip - <http://tulip.labri.fr/TulipDrupal/>
- Visone - <http://visone.info/html/demo.html>

# Gephi



# Gephi

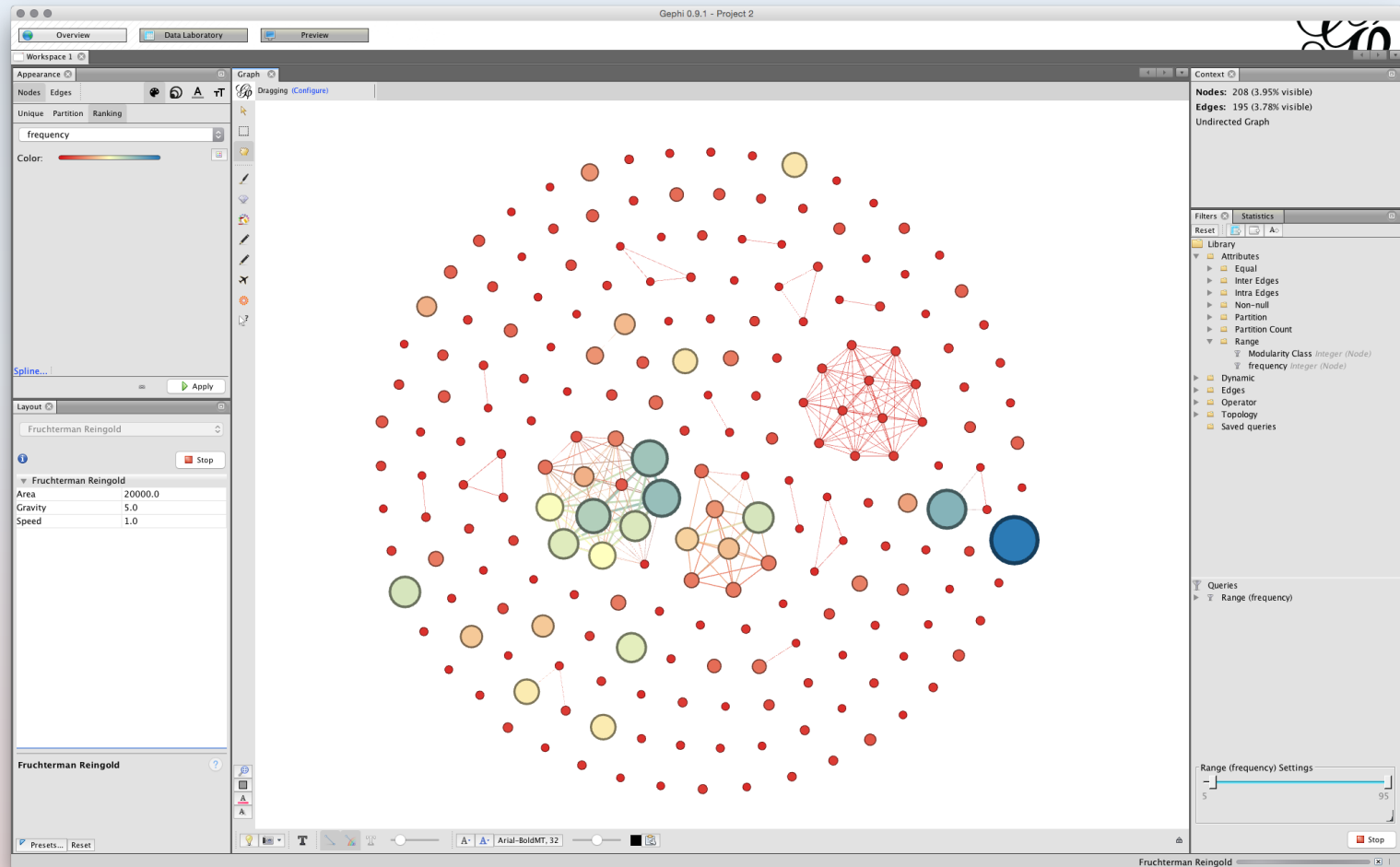


Image: [https://www.pauloldham.net/gephi\\_patent\\_network/](https://www.pauloldham.net/gephi_patent_network/)



# Gephi

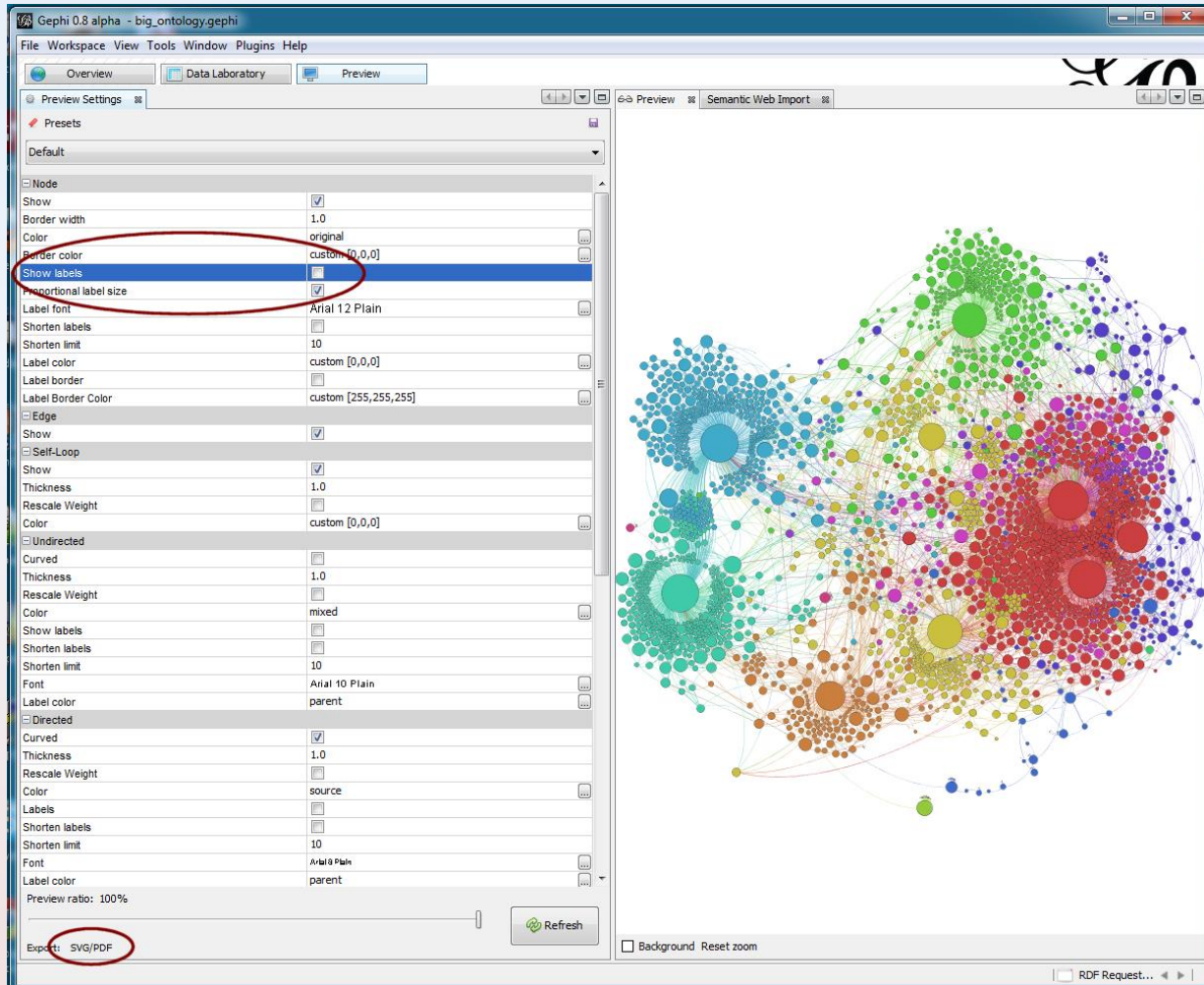


Image: [http://www.mkbergman.com/wp-content/themes/ai3v2/images/2011Posts/110808\\_gephi\\_preview.png](http://www.mkbergman.com/wp-content/themes/ai3v2/images/2011Posts/110808_gephi_preview.png)

Questions?



# Installing Gephi

- Requires Java JRE version 7 or 8.
  - PC: <https://www.java.com/en/download/>
  - Mac: Java is preinstalled
  - Linux: Update your distribution with JRE

# Installing Gephi

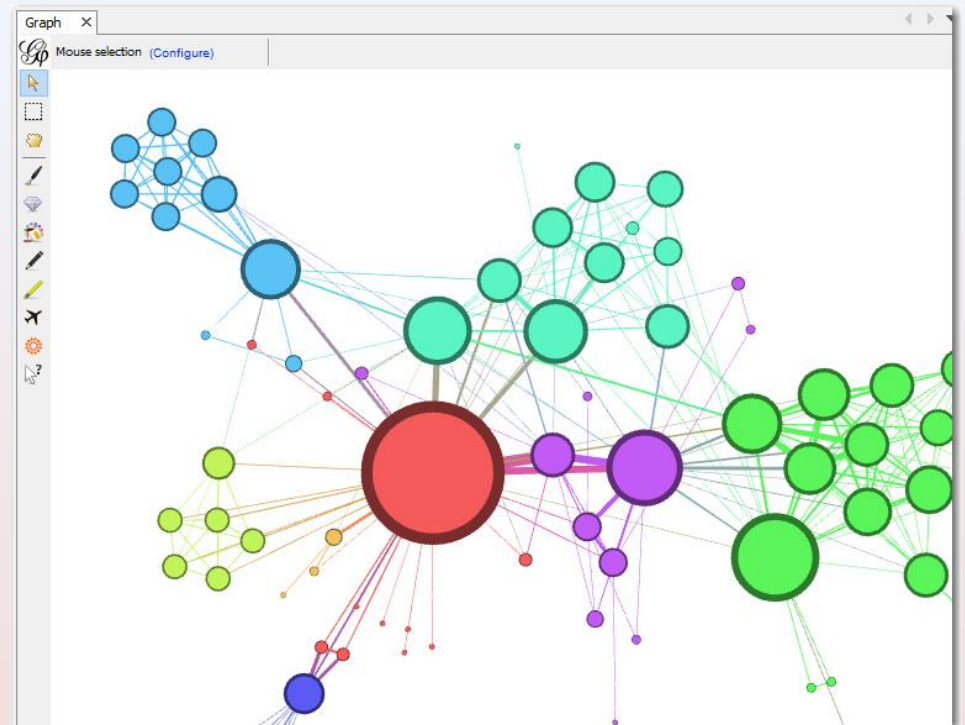
- Gephi installer
  - <https://gephi.org/users/download/>
  - Windows:
    - run installer and follow prompts
  - Mac:
    - click the downloaded .dmg file
    - Drag the gephi app into the application folder
  - Linux:
    - Untar/unzip then execute `./bin/gephi` script file

# Install Gephi now

Mark and I will walk around to help if needed.

# Run Gephi

- Welcome window
  - Select Les Miserables.gexf sample
    - (Window > Welcome will show that window)
  - This is our goal.



# Gephi datasets

<https://github.com/gephi/gephi/wiki/Datasets>

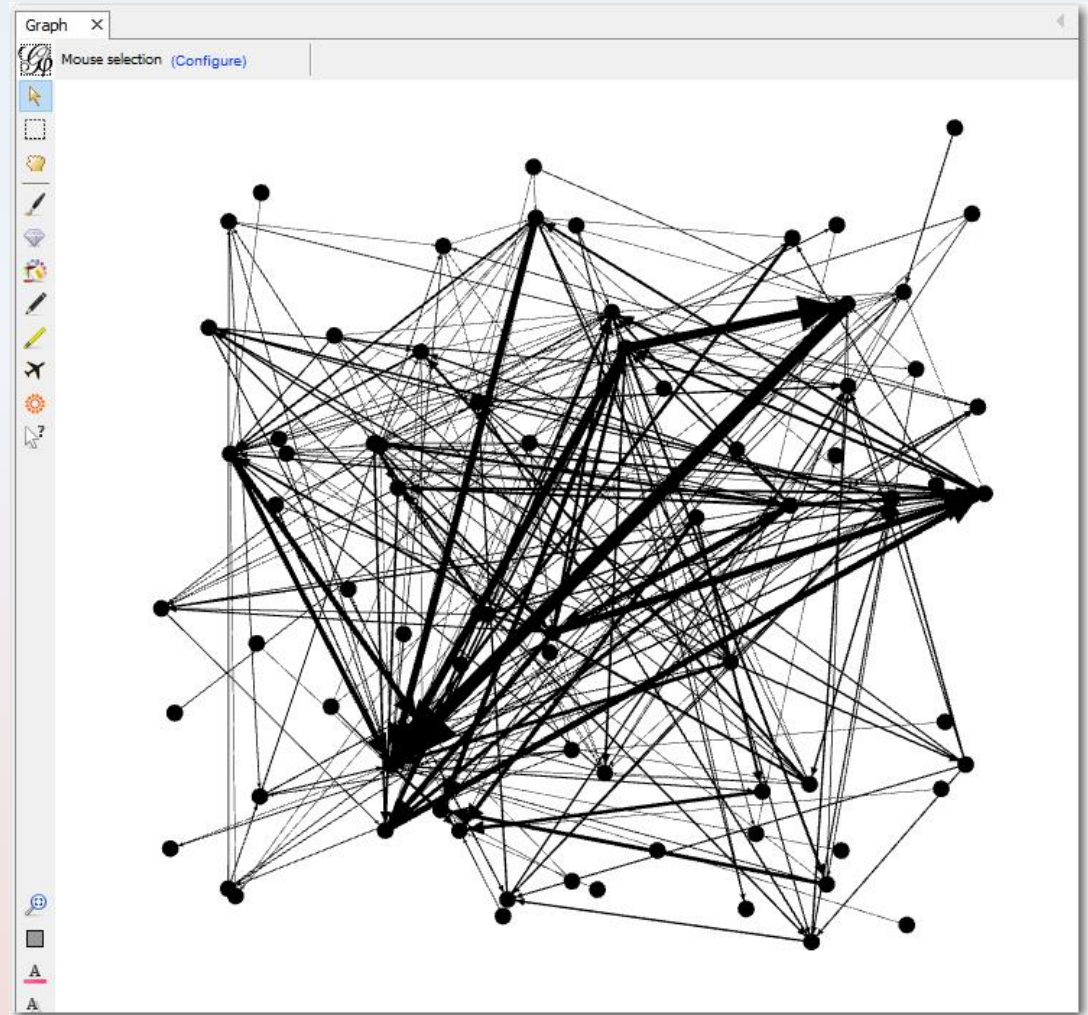
- Download the [“GML file. Les Miserables: ...”](#)
- Unzip and save the file

```
Creator "Mark Newman on Fri Jul 21 12:44:53 2006"  
Graph  
[  
  node [ id 0 label "Myriel" ]  
  node [ id 1 label "Napoleon" ]  
  ....  
  edge [ source 1 target 0 value 1 ]  
  edge [ source 2 target 0 value 8 ]  
  ....  
]
```

# Open lesmiserables.gml file in Gephi

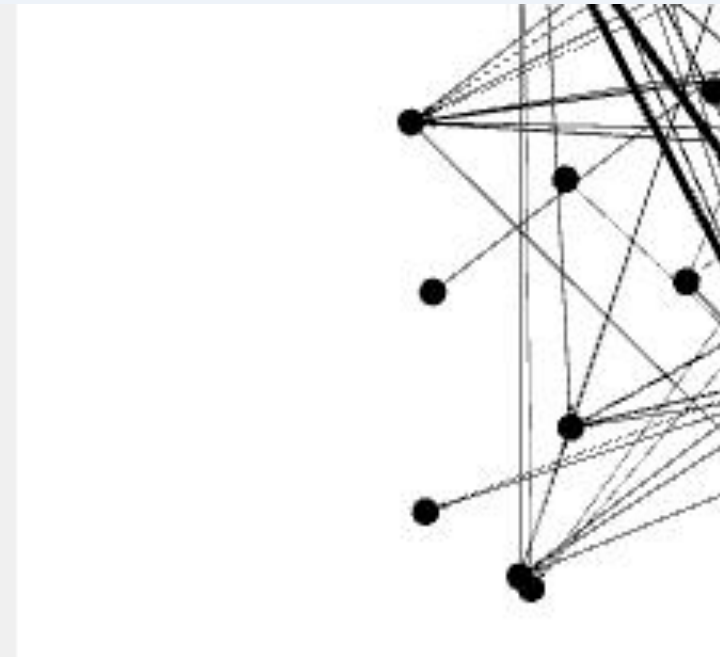
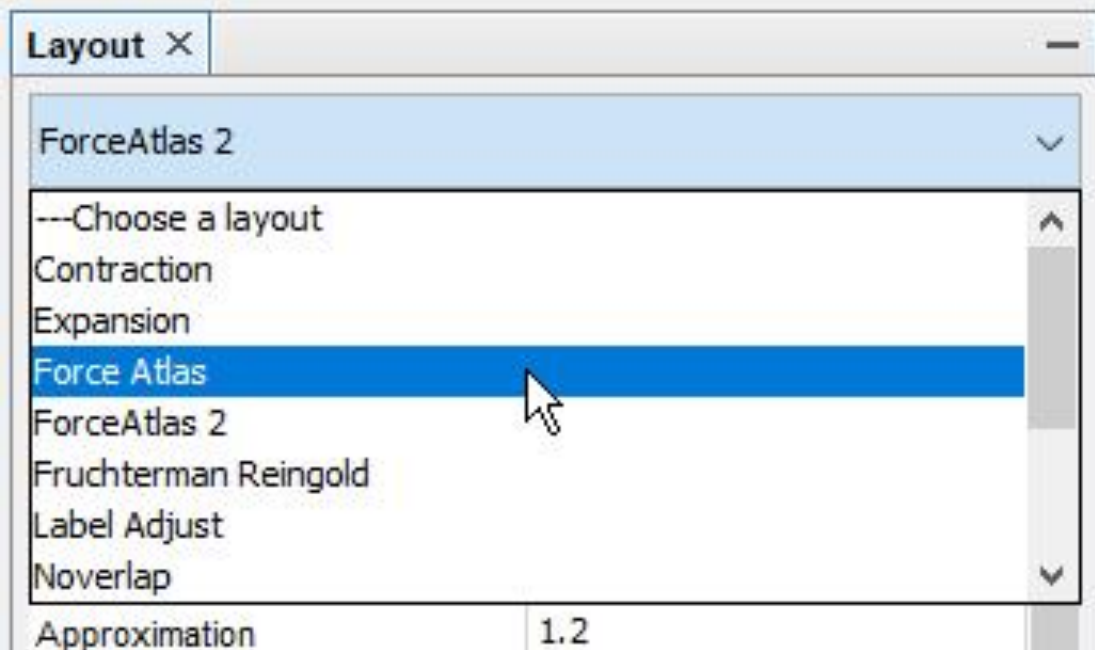
Zoom with  
mouse wheel  
or two-finger  
up-down  
stroke on the  
touch pad.

Right mouse  
button will  
move the  
visualization.



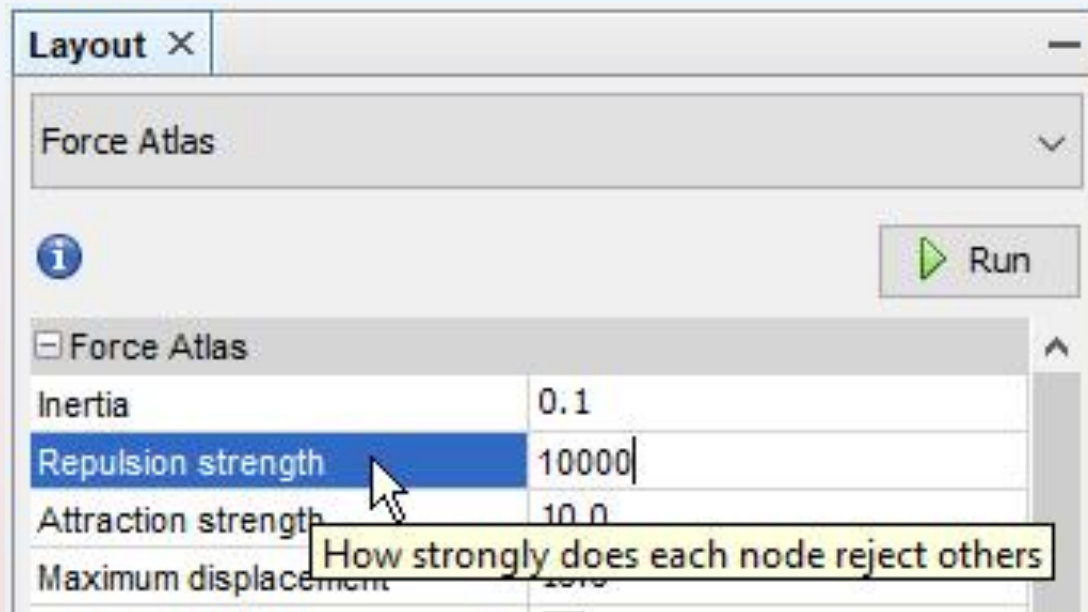
# Layout

Gephi has many different options for how to distribute the nodes



# Stretch out the graph

Change the “Repulsion strength”  
to 10,000 and click “Run”



Layout ×

Force Atlas

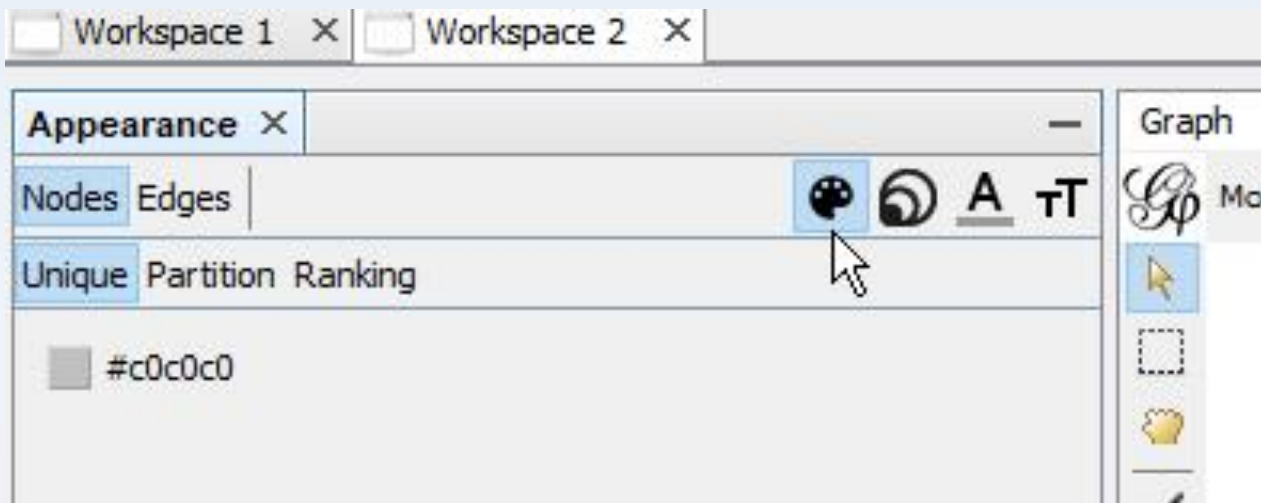
Run

Force Atlas	
Inertia	0.1
Repulsion strength	10000
Attraction strength	10.0
Maximum displacement	20.0

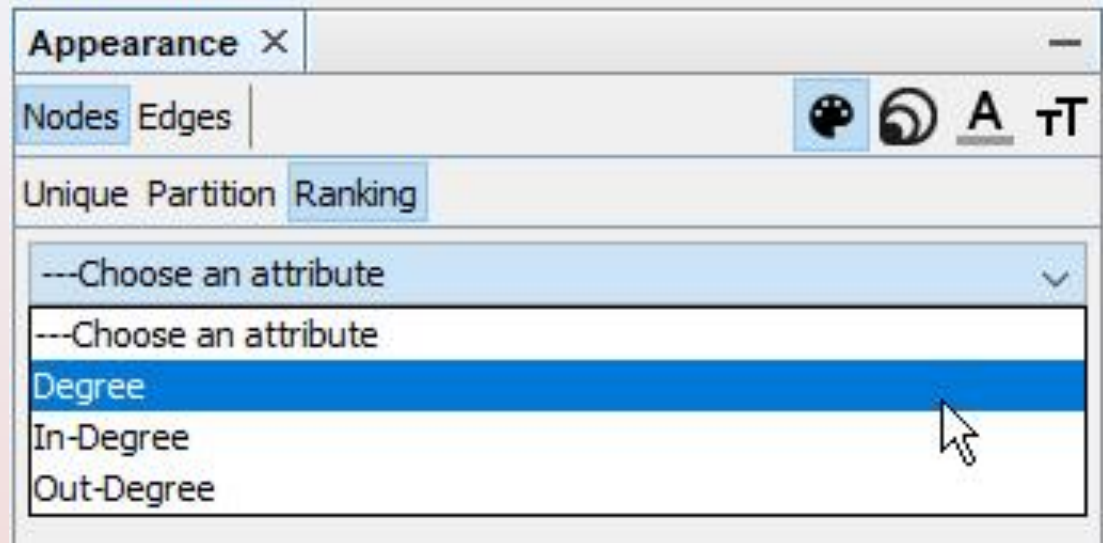
How strongly does each node reject others



# Appearance > Nodes > Color palette icon > Ranking > Degree

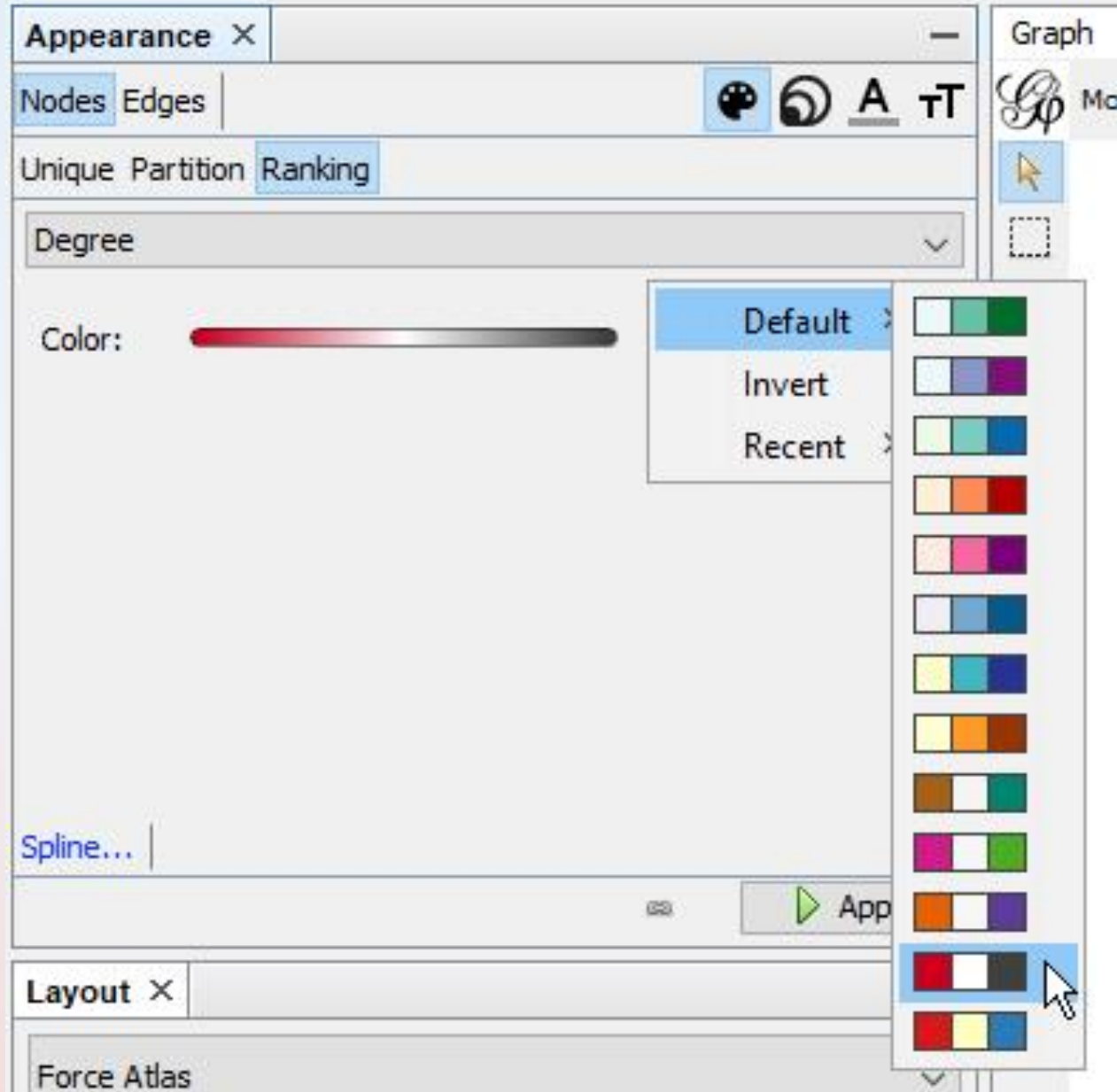


Highly  
connected  
nodes  
become  
darker



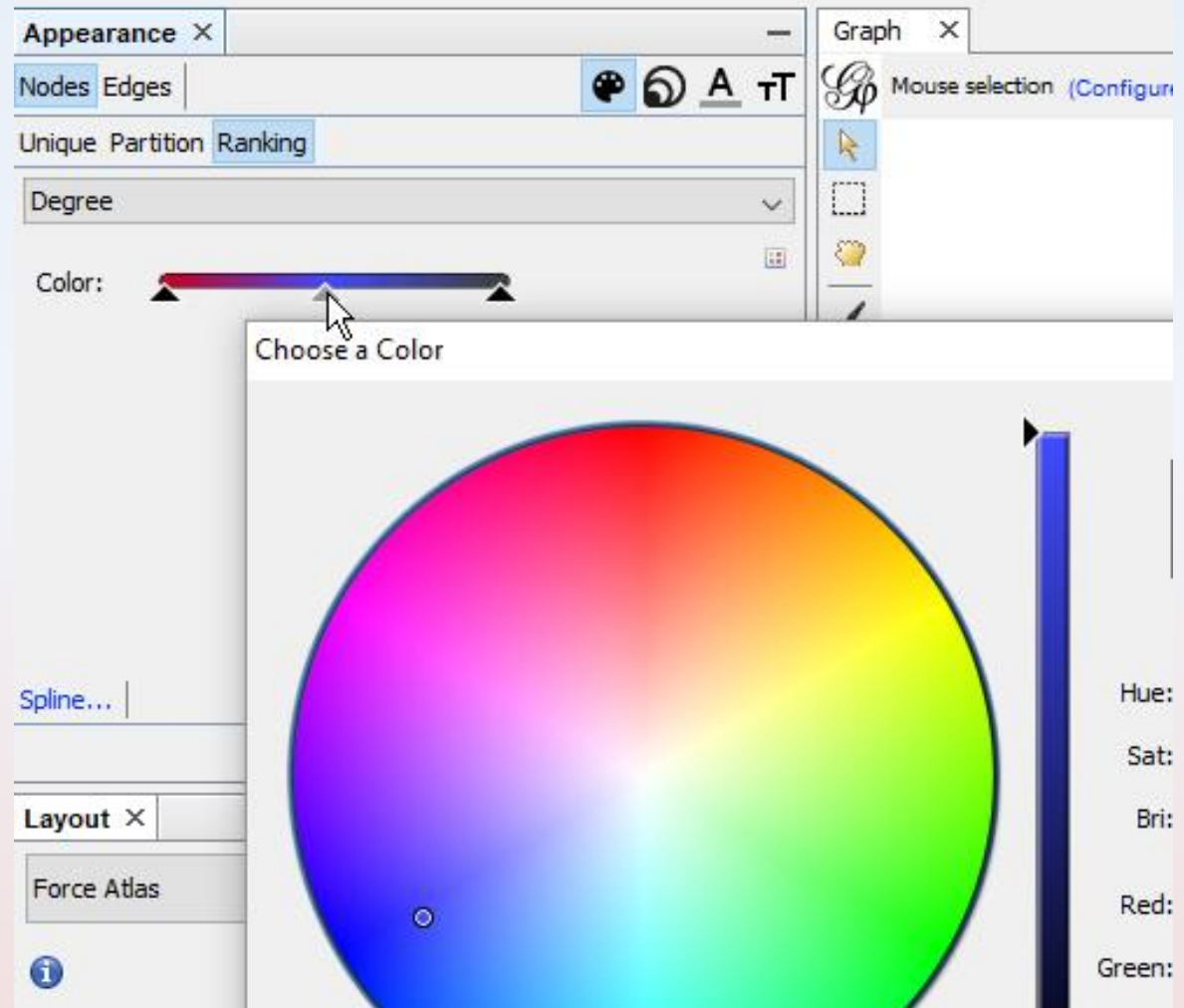
# Set colors

To the right of the color bar, click and select your desired colors then click Apply



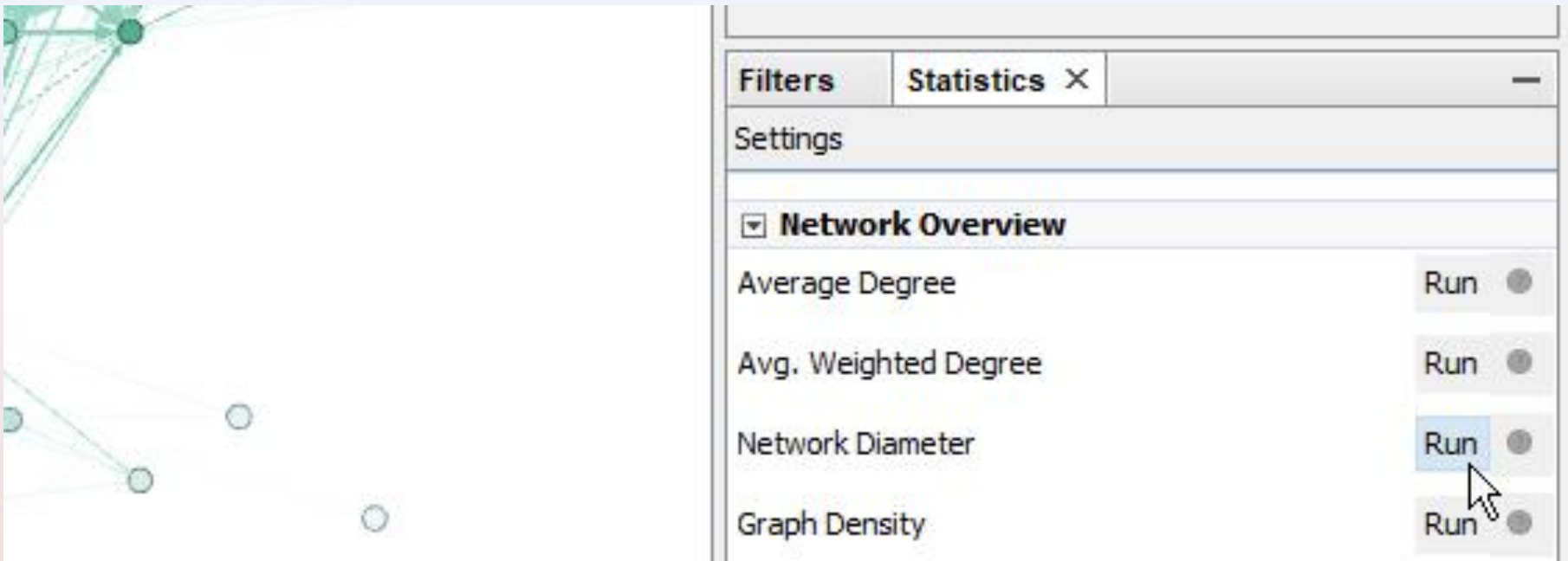
# Set color gradient

Click on a triangle in the color bar to chose a color



# Statistics

- On the right, click
  - Statistics > Network Diameter > Run
  - Select Undirected

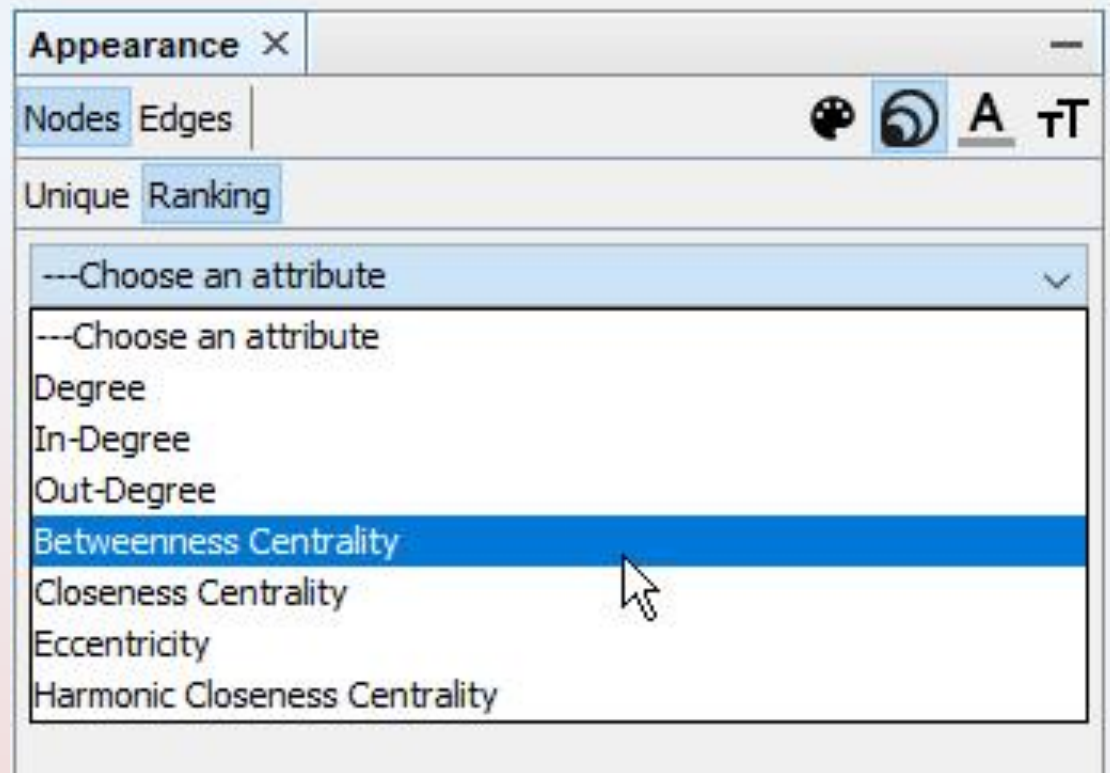


The image shows a network visualization software interface. On the left, a network graph is displayed with nodes and edges. On the right, a statistics panel is visible, showing various network metrics and their corresponding 'Run' buttons. The 'Network Diameter' button is highlighted, and a mouse cursor is pointing at it.

Filters	Statistics X	
Settings		
Network Overview		
Average Degree	Run	●
Avg. Weighted Degree	Run	●
Network Diameter	Run	●
Graph Density	Run	●

# Size

- More options are now available
- Size > Ranking > Betweenness Centrality



# Ranking > Size

Set size to range from 20 to 119

The screenshot displays a software interface for graph visualization, divided into two main panels: 'Appearance' and 'Graph'.

**Appearance Panel:**

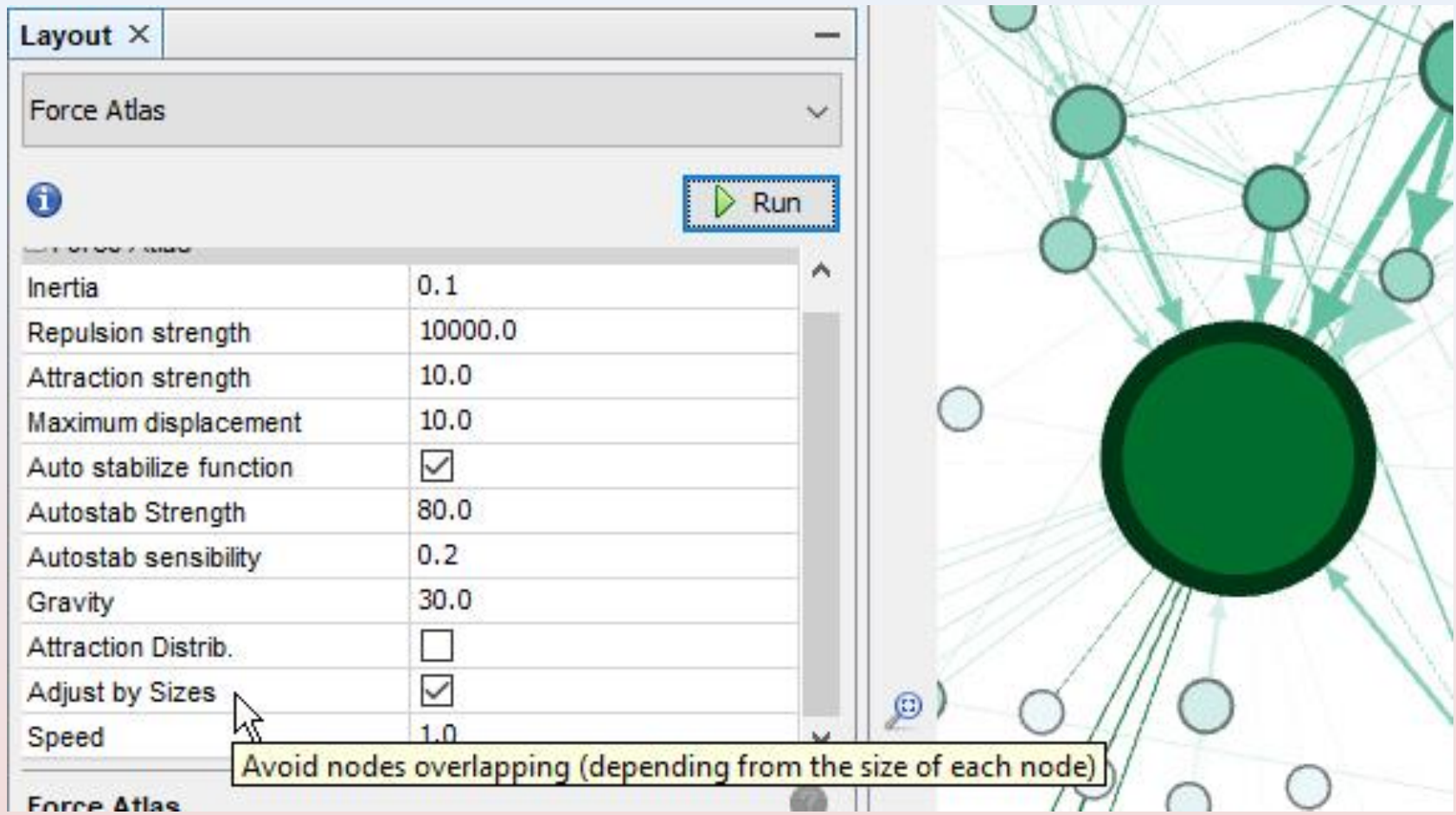
- Buttons: Nodes, Edges, Unique, Ranking (selected), Betweenness Centrality (selected).
- Size controls: Min size: 20, Max size: 119.
- Icons: A palette icon, a circular arrow icon (highlighted with a mouse cursor), an 'A' icon, and a 'T' icon. A tooltip labeled 'Size' is visible over the circular arrow icon.

**Graph Panel:**

- Header: Graph, Mouse selection (Configure).
- Tools: A vertical toolbar with various icons for graph manipulation (e.g., selection, pan, zoom, delete, add node, add edge).
- Visualization: A network graph with nodes of varying sizes and colors (green and light blue) connected by edges. The largest node is a dark green circle, indicating it has the highest betweenness centrality.

# Layout by size

In Layout, click Adjust by Sizes



The screenshot shows a software interface for network layout. On the left is a 'Layout' panel with a dropdown menu set to 'Force Atlas'. Below the menu is an information icon and a 'Run' button. A table lists various parameters for the layout algorithm:

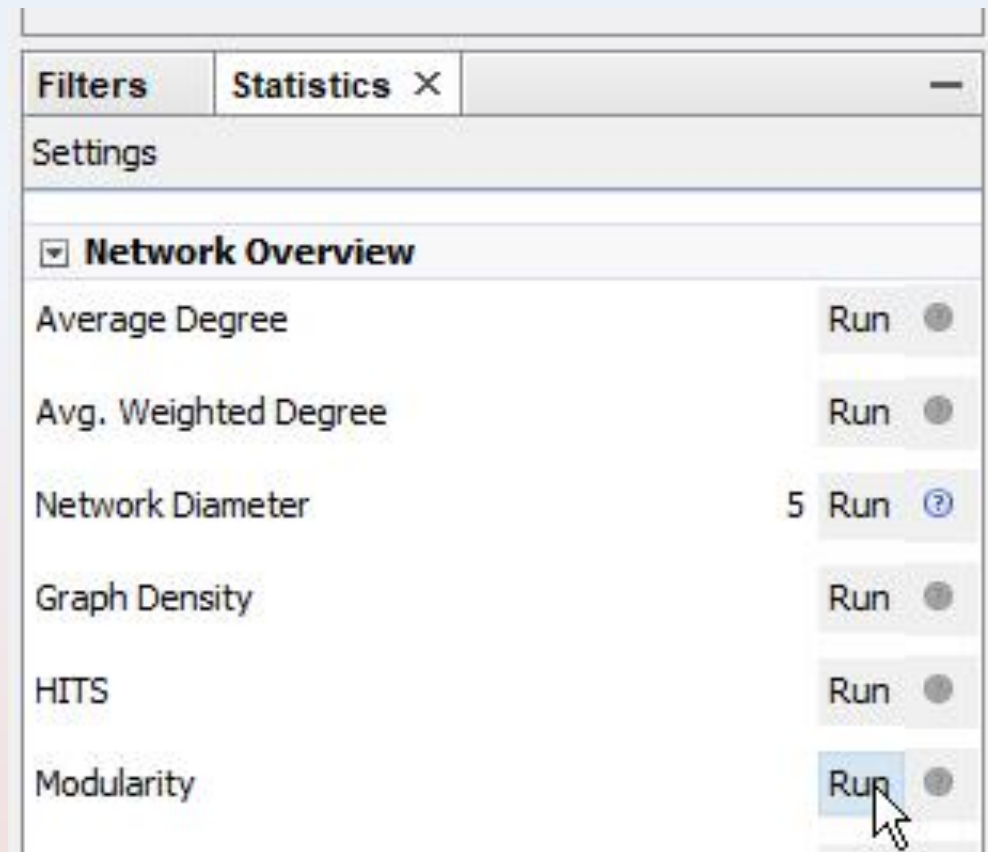
Inertia	0.1
Repulsion strength	10000.0
Attraction strength	10.0
Maximum displacement	10.0
Auto stabilize function	<input checked="" type="checkbox"/>
Autostab Strength	80.0
Autostab sensibility	0.2
Gravity	30.0
Attraction Distrib.	<input type="checkbox"/>
Adjust by Sizes	<input checked="" type="checkbox"/>
Speed	1.0

A mouse cursor is pointing at the 'Adjust by Sizes' checkbox. A callout box at the bottom of the panel contains the text: "Avoid nodes overlapping (depending from the size of each node)". To the right of the panel is a network diagram with several nodes of varying sizes and colors (green and blue) connected by edges. A large green node is prominent in the center.

# Community Detection

What clusters of nodes belong together?

-In Statistics,  
run Modularity



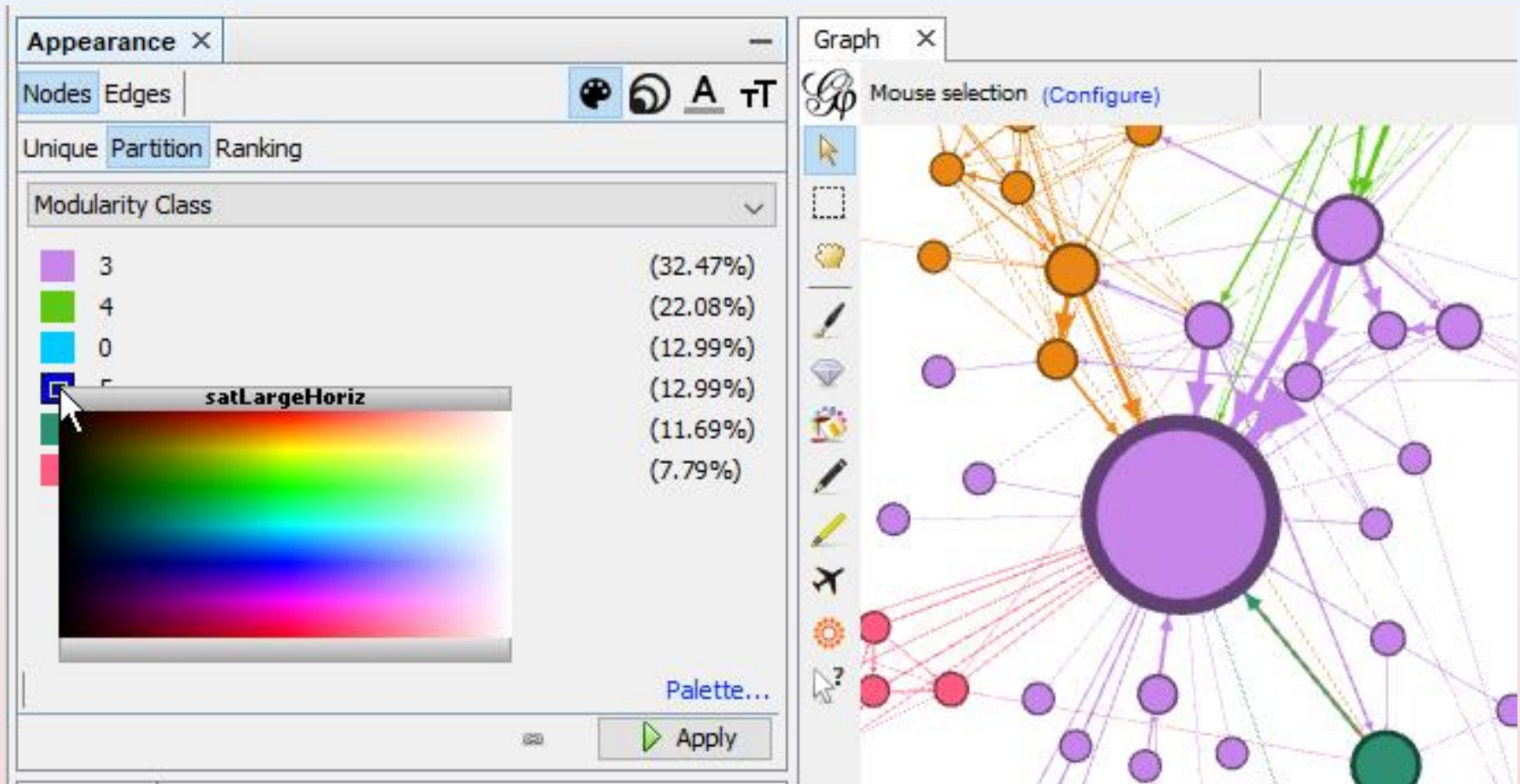
The screenshot shows a software interface with a 'Statistics' tab. Under the 'Network Overview' section, several metrics are listed, each with a 'Run' button and a status indicator (a grey circle). The 'Modularity' button is highlighted in blue, and a mouse cursor is pointing at it.

Metric	Value	Run Button	Status
Average Degree		Run	●
Avg. Weighted Degree		Run	●
Network Diameter	5	Run	?
Graph Density		Run	●
HITS		Run	●
Modularity		Run	●



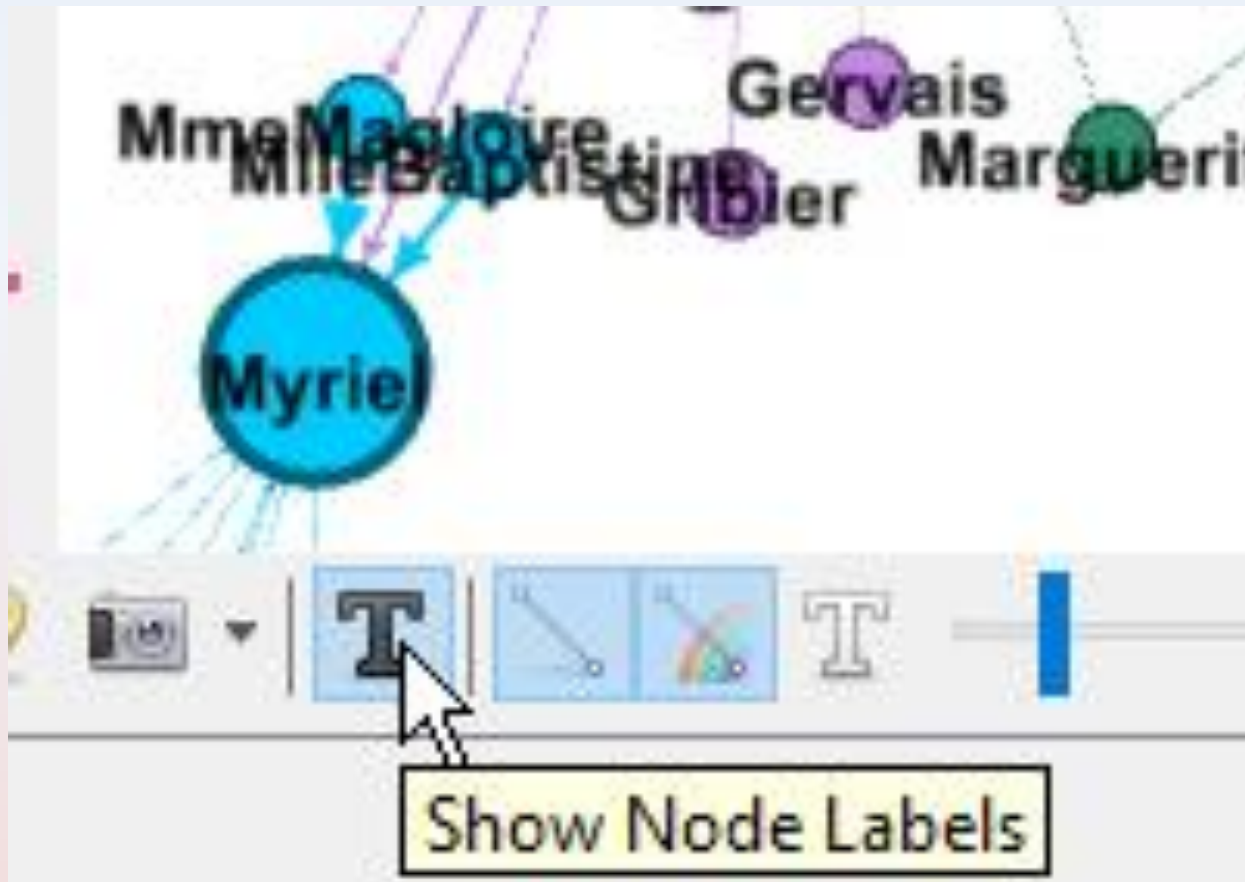
# Display modularity

Nodes > Colors > Partition > Modularity Class



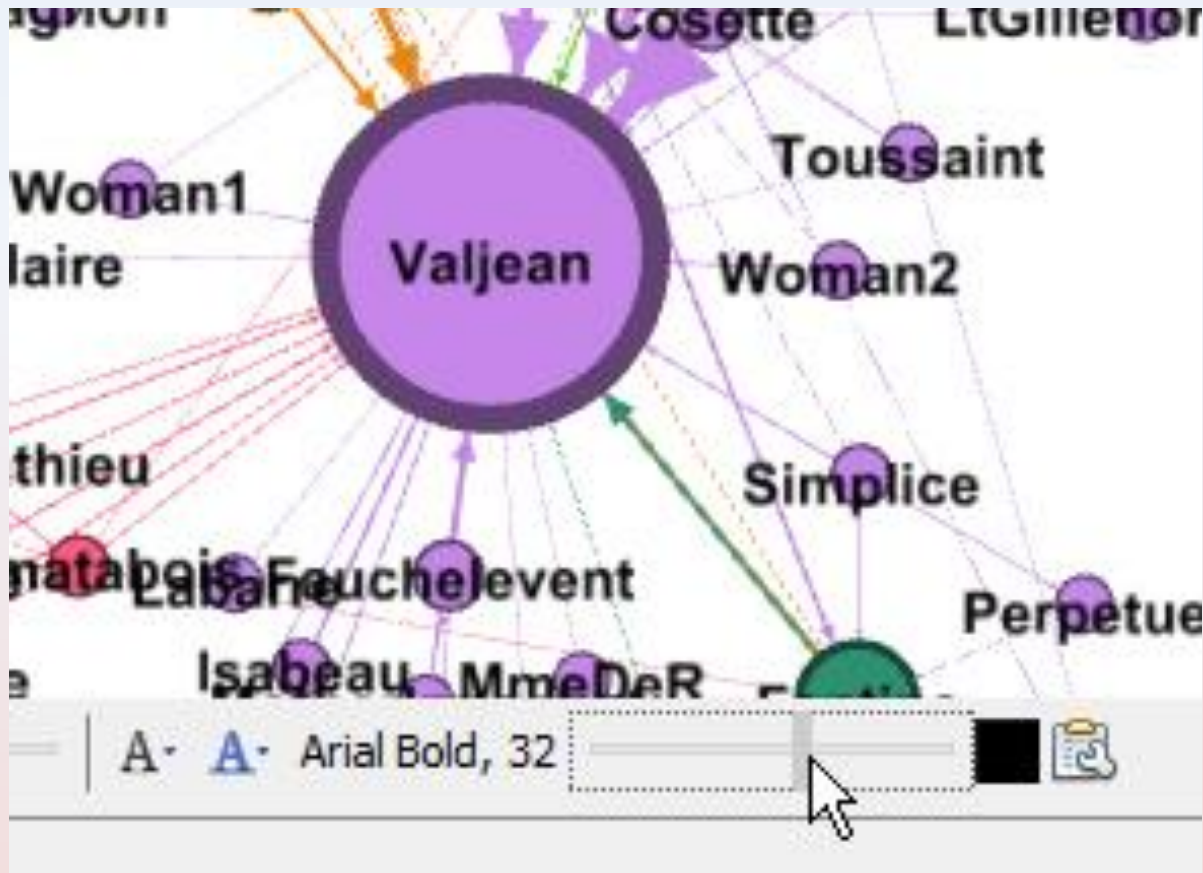
# Node labels

Click the solid “T” to show node labels



# Node Label Size

The slider can change the size of node labels

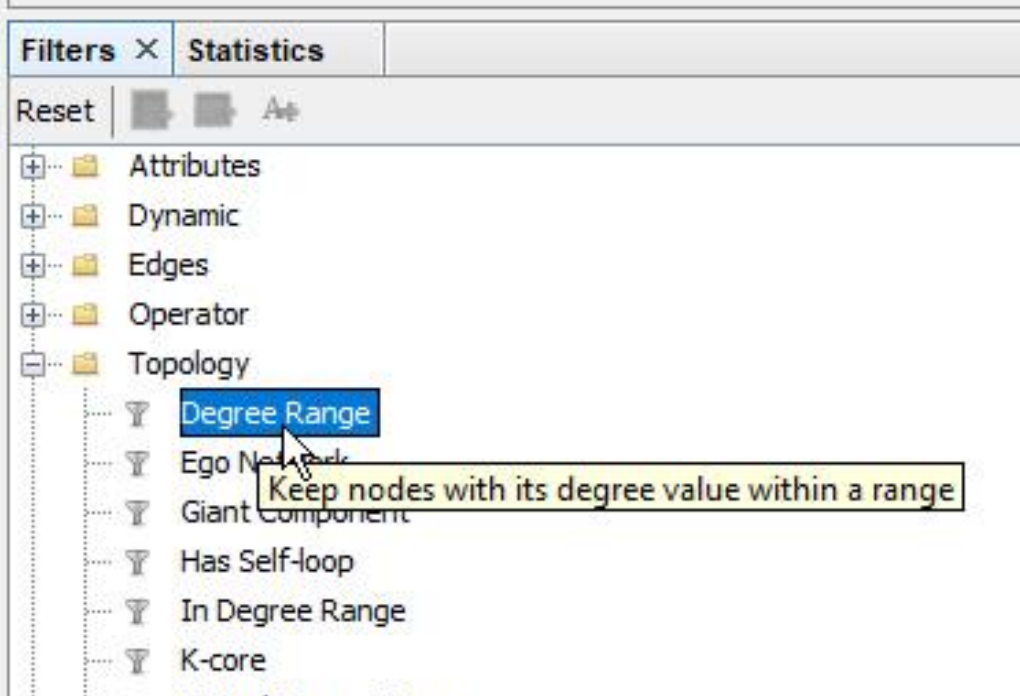


# Filters

Beside Statistics is Filters

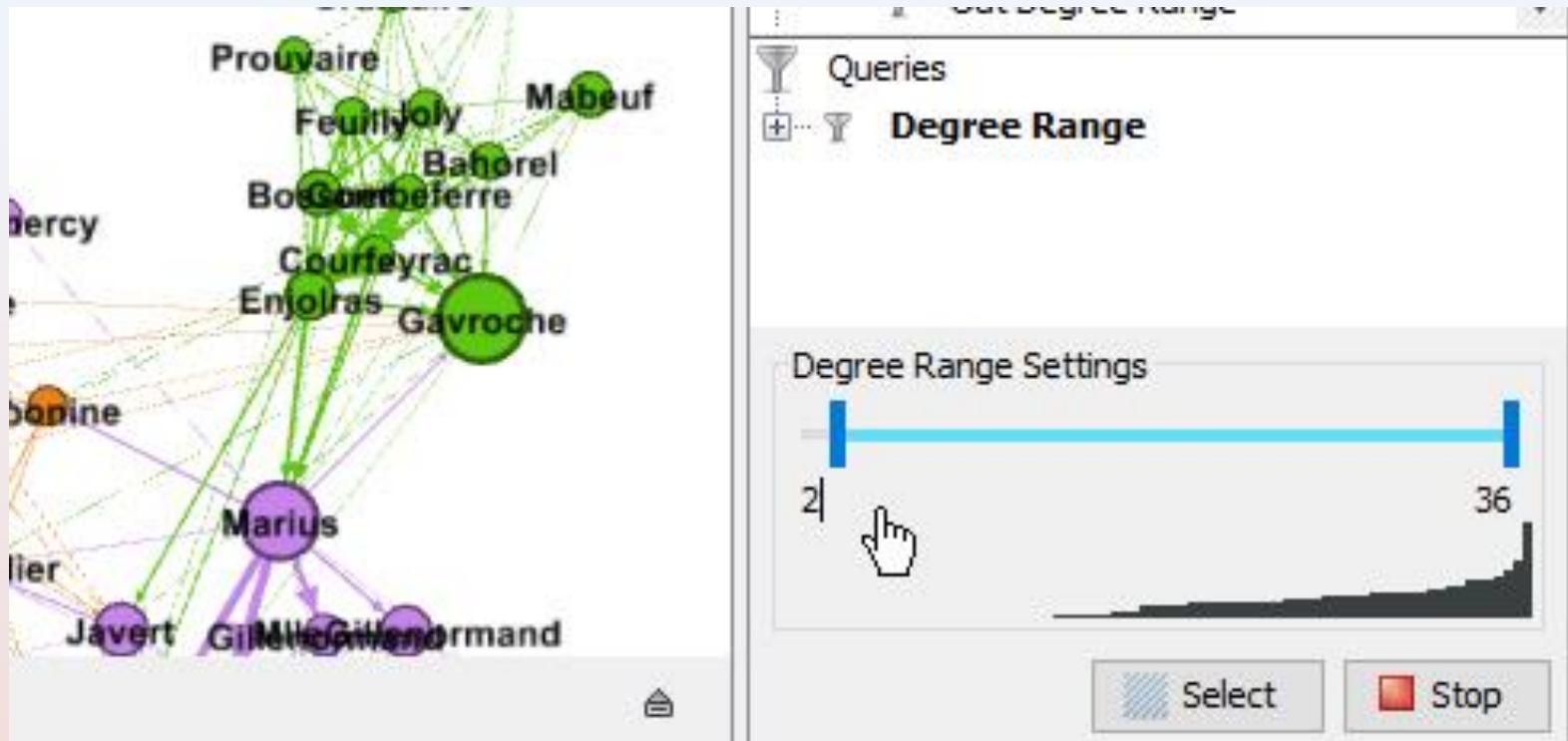
Select Filters > Topology > Degree Range

And drag it down to the queries section



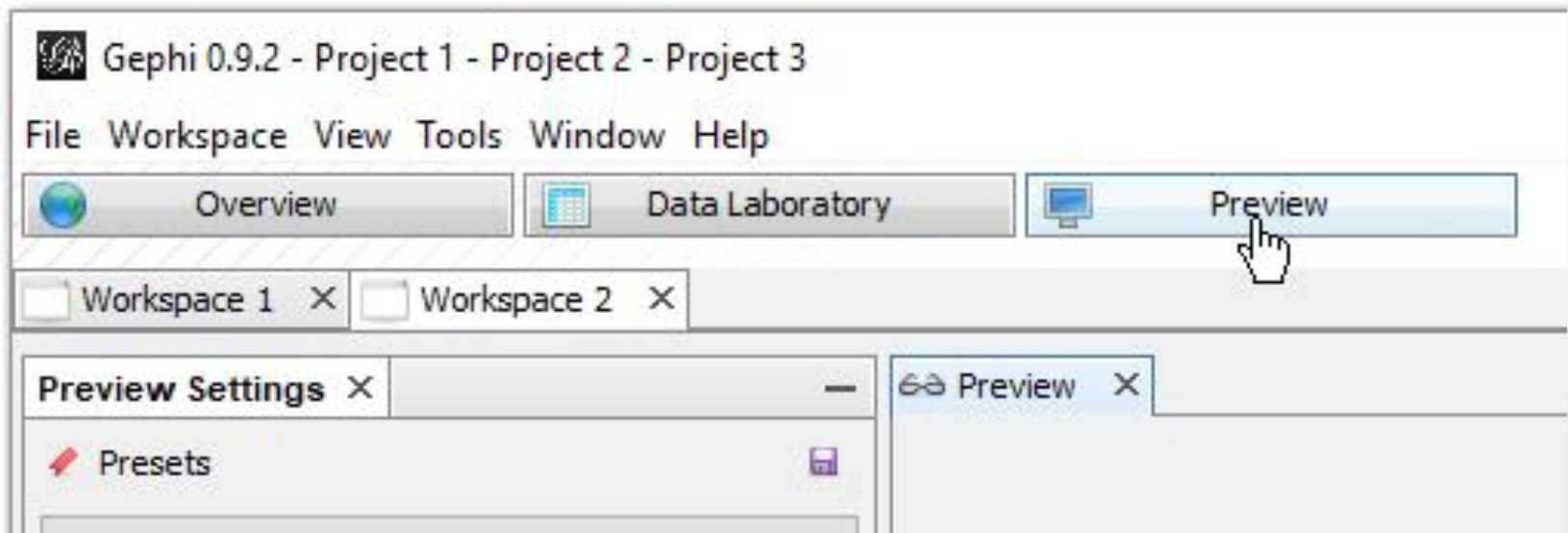
# Filters

Adjust the degree range slider to show significant items.



# Exporting the graph

Click the preview tab at the very top



# Preview settings

Click Show Labels and any other changes,  
then click Refresh

The image shows a software interface with two windows. The left window is titled 'Preview Settings' and contains a 'Node Labels' section with the following settings:

Setting	Value
Border Width	1.0
Border Color	custom [0,0,0]
Opacity	100.0
Per-Node Opacity	<input type="checkbox"/>
<b>Node Labels</b>	
Show Labels	<input checked="" type="checkbox"/>
Font	Arial 12 Plain
Proportional size	<input checked="" type="checkbox"/>
Color	custom [0,0,0]
Shorten label	<input type="checkbox"/>
Max characters	30
Outline size	0.0
Outline color	custom [255,255,255]
Outline opacity	80.0
Box	<input type="checkbox"/>
Box color	parent

The right window is titled 'Preview' and displays a network graph. The graph consists of nodes connected by lines. The nodes are labeled with names: 'Montparnasse', 'Brujon', 'Claquesous', 'Eponine', 'Gueulemer', 'Babet', 'Anzelm', 'Theardier', 'Jav', 'Mme Theardier', 'Valje', and 'Gammathieu'. The 'Valje' node is the largest and most prominent, colored purple. The 'Theardier' node is also large and colored orange. The 'Gammathieu' node is small and colored red. The 'Anzelm' node is small and colored blue. The 'Babet' node is small and colored orange. The 'Gueulemer' node is small and colored orange. The 'Claquesous' node is small and colored orange. The 'Brujon' node is small and colored orange. The 'Montparnasse' node is small and colored orange. The 'Mme Theardier' node is small and colored orange. The 'Jav' node is small and colored orange. The 'Eponine' node is small and colored orange.

# Export

Select PNG or other format at the lower left.

The screenshot displays the 'Edges' settings panel on the left and a network graph on the right. The 'Edges' panel includes a table with the following settings:

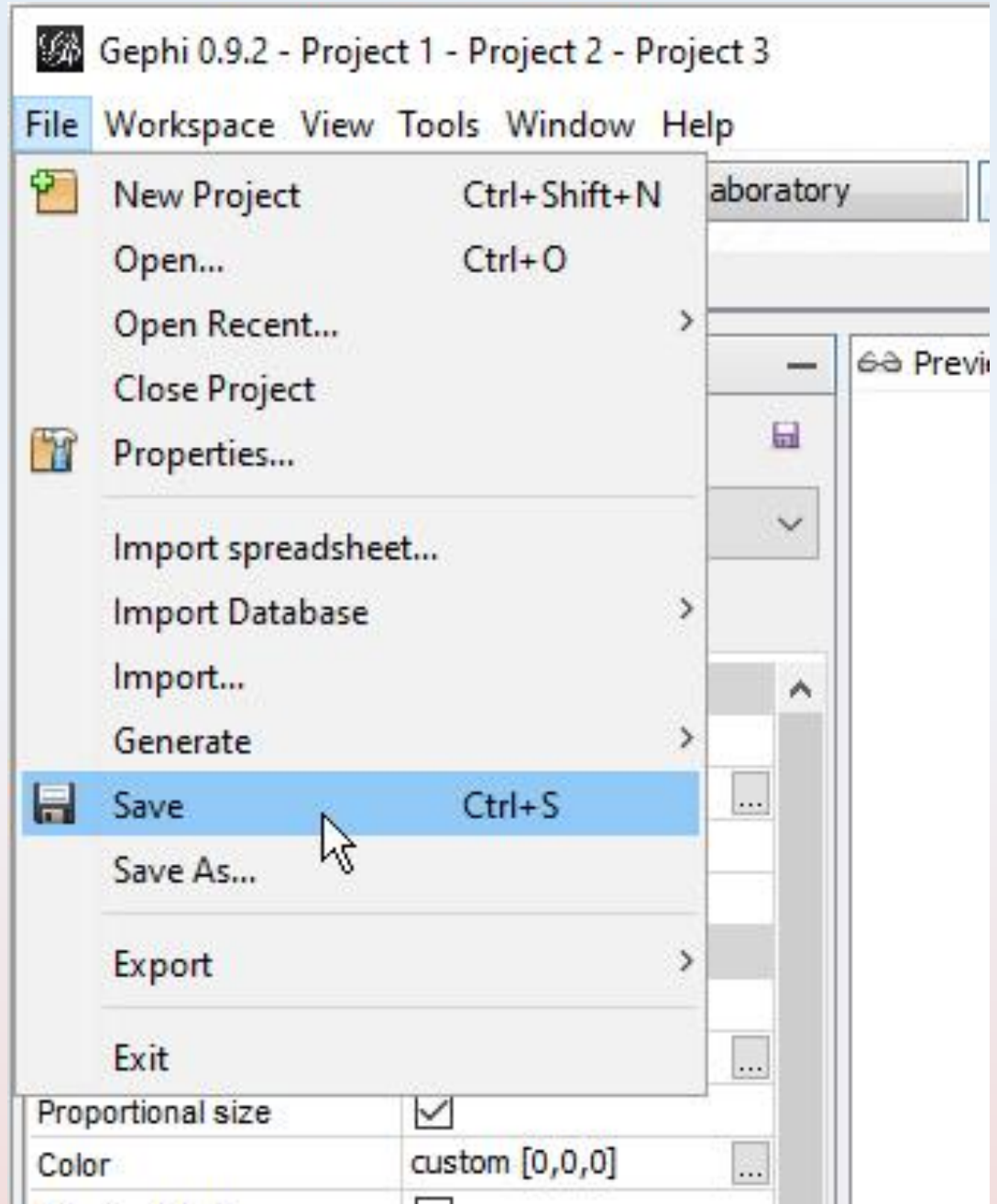
Show Edges	<input checked="" type="checkbox"/>
Thickness	1.0
Rescale weight	<input type="checkbox"/>
Min. rescaled weight	0.1
Max. rescaled weight	1.0

Below the table, the 'Preview ratio' is set to 100%. A 'Refresh' button is located to the right of the preview slider. At the bottom left of the panel, the 'Export:' dropdown menu is open, showing 'SVG/PDF/PNG' as the selected option. A tooltip below the dropdown reads 'Export as SVG of PDF format'. The network graph on the right shows several nodes (red circles) connected by edges. The nodes are labeled with names: 'Brevet', 'Champmathieu', 'Chenildieu', 'Cocheville', 'Jouy', 'Reaumont', and 'Tabou'. A 'Background' checkbox is visible at the bottom right of the graph area.



# Save Project from the File menu

This saves  
your project  
settings



Questions?

# Practice Time

# More Information

- <http://www.martingrandjean.ch/gephi-introduction/>
- <https://www.kdnuggets.com/2015/06/top-30-social-network-analysis-visualization-tools.html/3>
- <https://video.search.yahoo.com/search/video?fr=crmas&p=gephi+tutorial#id=2&vid=597f4e60a1ba9270030a1c6b3be6cc9e&action=click>
- <https://github.com/gephi/gephi/wiki/Datasets>
- <https://youtu.be/371n3Ye9vVo>

# Image sources

- [https://cdn.pixabay.com/photo/2015/05/24/14/02/spaghetti-781795\\_960\\_720.jpg](https://cdn.pixabay.com/photo/2015/05/24/14/02/spaghetti-781795_960_720.jpg)
- [https://assets.epicurious.com/photos/55f72d733c346243461d496e/2:1/w\\_1260,h\\_630/09112015\\_15minute\\_pastasauce\\_tomato.jpg](https://assets.epicurious.com/photos/55f72d733c346243461d496e/2:1/w_1260,h_630/09112015_15minute_pastasauce_tomato.jpg)