



SCHOOL OF
SCIENCE &
TECHNOLOGY

Social Network Analysis

ÁLVARO ROMERO MIRALLES



Tutorial Quick Start

- * Introduction
- * Import file
- * Visualization
- * Layout
- * Ranking (color)
- * Metrics
- * Ranking (size)
- * Layout again
- * Show labels
- * Community-detection
- * Partition
- * Filter
- * Preview
- * Export
- * Save
- * Conclusion

Gephi Tutorial Quick Start

Welcome to this introduction tutorial. It will guide you to the basic steps of network visualization and manipulation in Gephi.

Gephi version 0.7alpha2 was used to do this tutorial.

↓ [Get Gephi](#)

https://gephi.org/tutorials/gephi-tutorial-quick_start.pdf

Download

Gephi is an open-source and multiplatform software distributed under the dual license [CDDL 1.0](#) and [GNU General Public License v3](#).

Official Releases

[Release Notes](#) | [System Requirements](#) | [Installation instructions](#)

Gephi 0.10.1 is the latest stable release.

Download Gephi for Windows

Version 0.10.1

If you have an older Gephi on your computer, you should uninstall it first, [see the installation instructions](#).

All downloads:

[Download Gephi 0.10.1 for Mac OS \(Intel\)](#)
[Download Gephi 0.10.1 for Mac OS \(Silicon\)](#)
[Download Gephi 0.10.1 for Windows](#)
[Download Gephi 0.10.1 for Linux](#)
[Download Older Versions](#)

 **Stay informed**
blog, twitter & more

 **Contribute**
code, doc, test & more

 **Report a Bug**

 **Request a Feature**

 **Share your ideas**

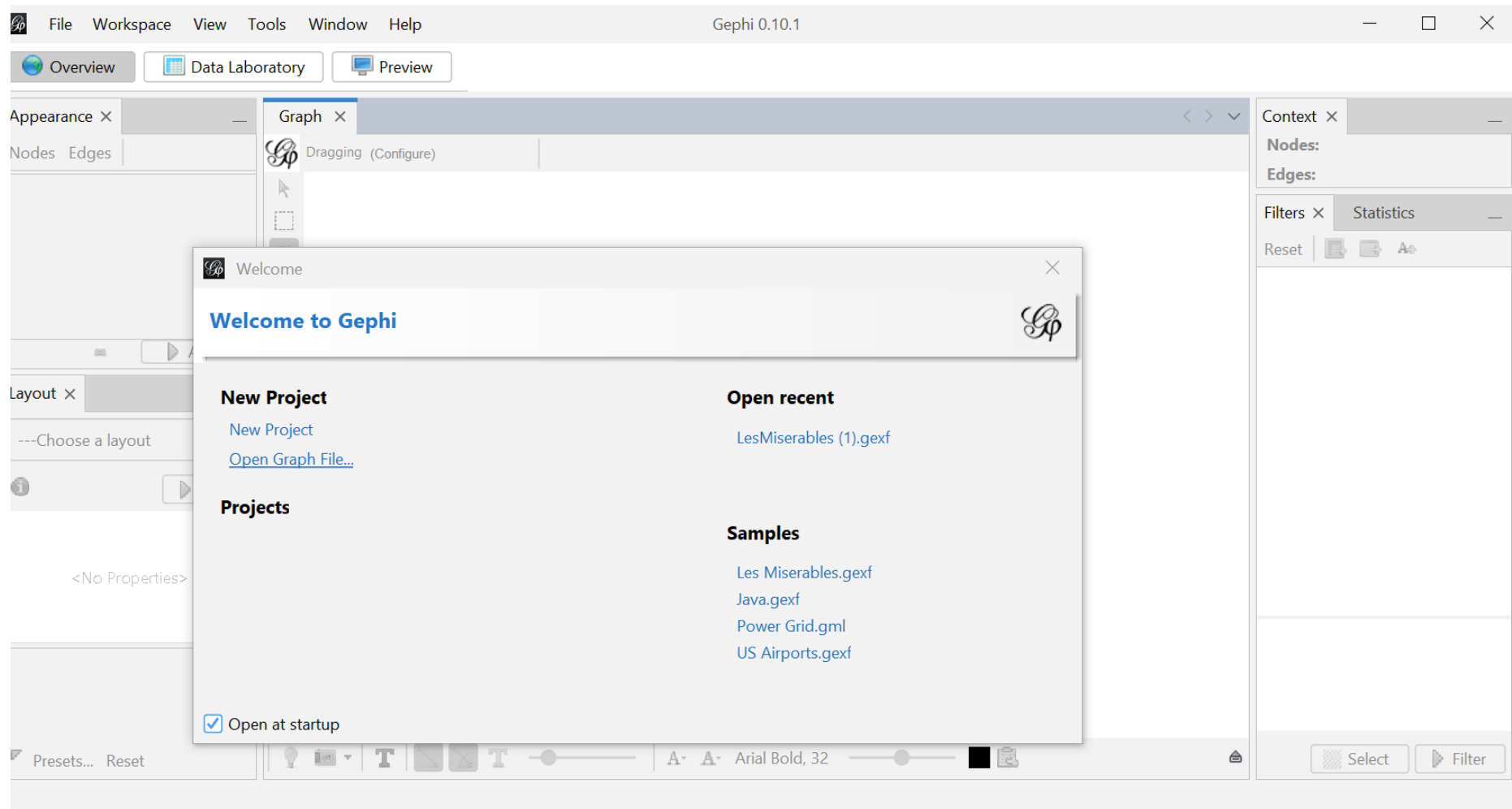
Ways to get help

- 2) Get Personal Help
- 3) Other Gephi Support

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

Open file

- In the menubar, go to File Menu and Open



Import file

- When your file is opened, the report sum up data found and issues.
 - Number of nodes Number of edges
 - Type of graph
- Click on OK to validate and see the graph

Import report

Source: LesMiserables (1).gexf

Issues Report

Nodes	Issues
<i>i</i> GEXF version 1.1 (deprecated)	INFO

Graph Type: Undirected

of Nodes: Undirected

of Edges: Mixed

Dynamic Graph: no

Dynamic Attributes: no

Multi Graph: no

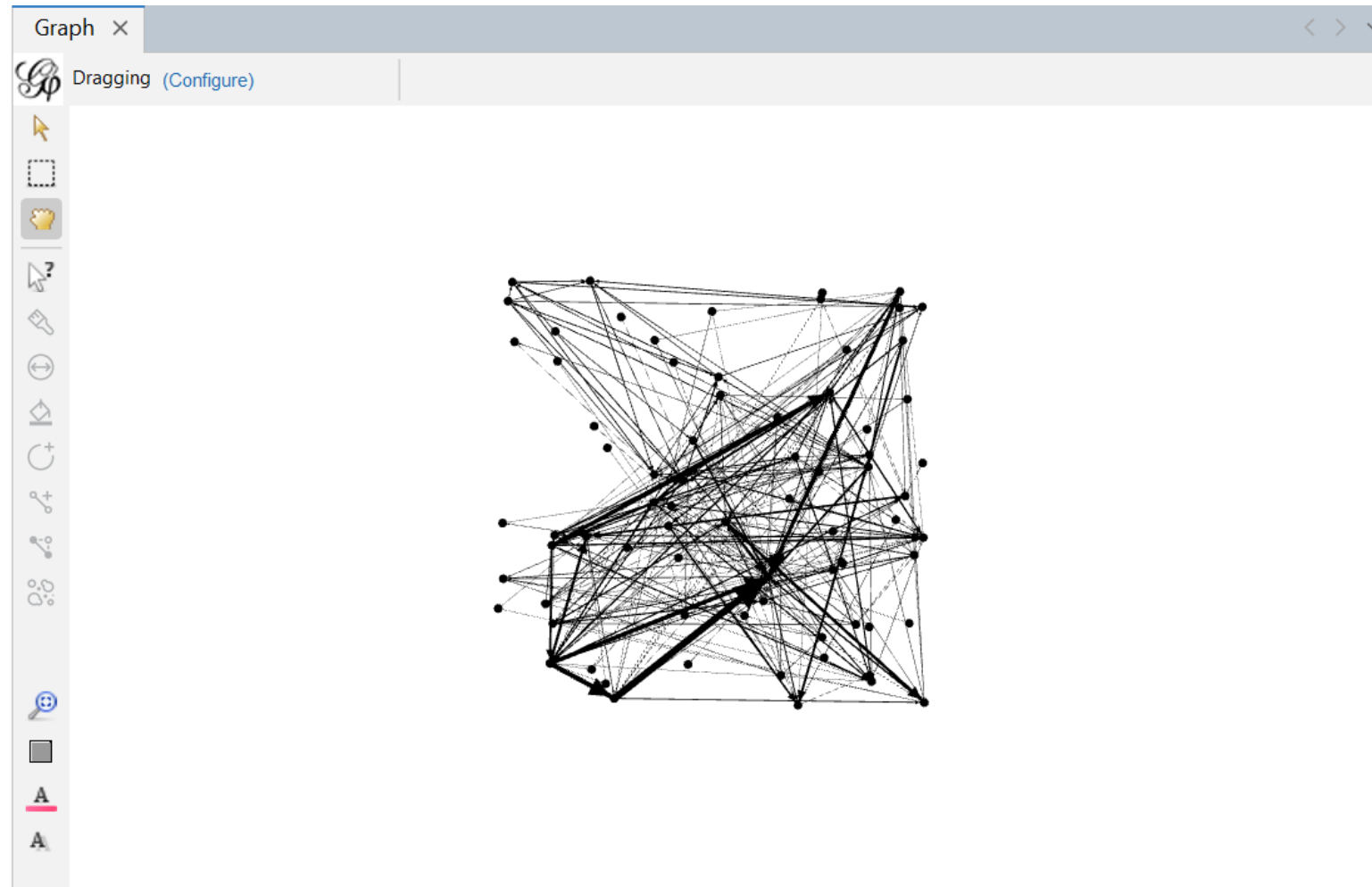
☒ New workspace
☐ Append to existing workspace

More options...

OK Cancel

Visualization

- You should now see this graph
 - ▮ Nodes position is random at first, so you may see a slightly different representation

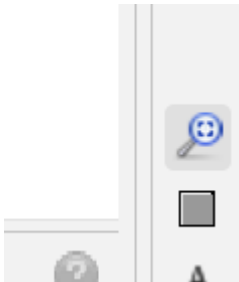


Visualization options

- Locate the “Edge Thickness” slider on the bottom



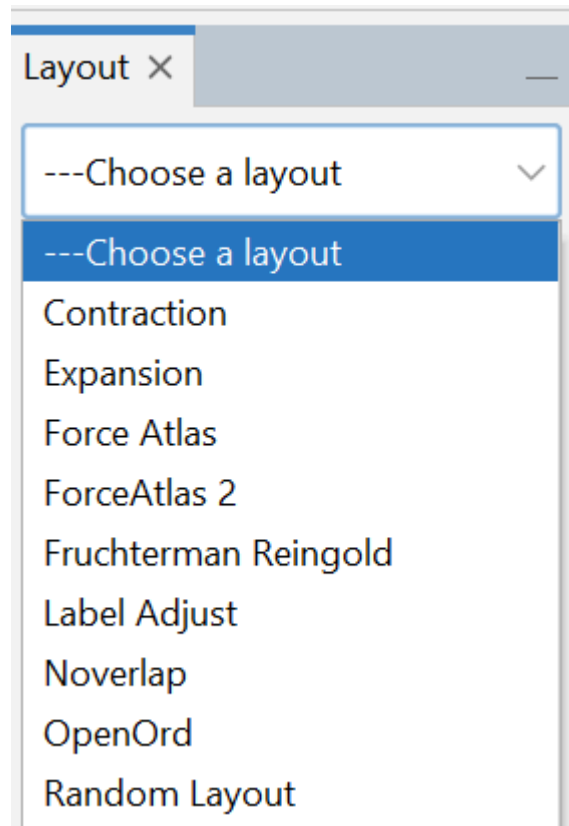
- If you loose your graph, reset the position



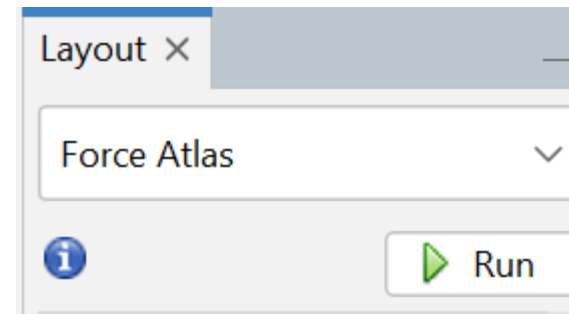
Layout the graph

Layout algorithms sets the graph shape, it is the most essential action.

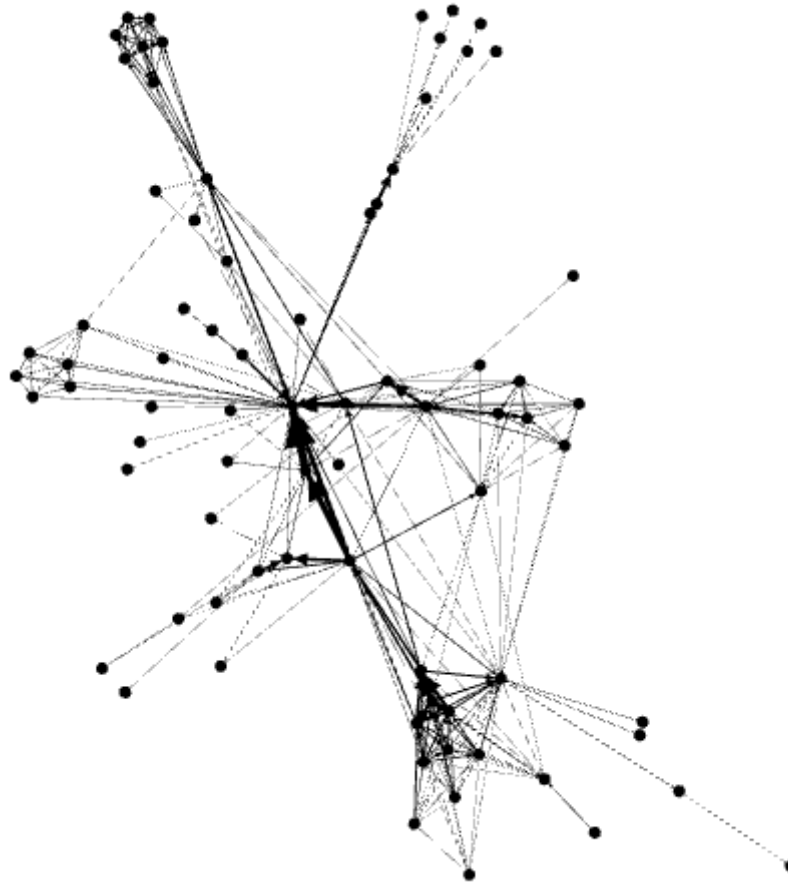
- Locate the Layout module, on the left panel.



- Choose “Force Atlas”
 - Set the “Repulsion strength” at 10 000 to expand the graph.
- Click on run to launch the algorithm

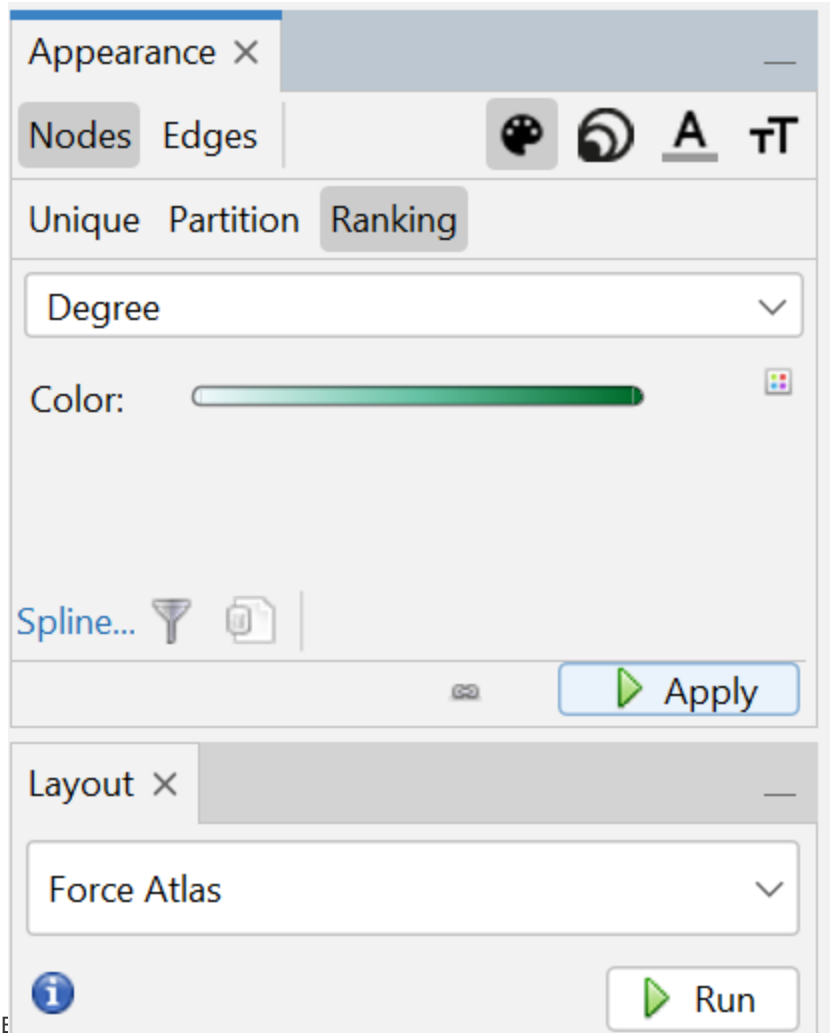


You should now see a layouted graph



Ranking (color)

Ranking module lets you configure node's color and size.



- Locate Ranking module, in the top left.
 - Choose “Degree” as a rank parameter
- Change colors as you prefer
- Click on apply to launch the algorithm

Ranking result table

You can see rank values by enabling the result table. Valjean has 36 links and is the most connected node in the network..

- Locate the Statistics module, on the right panel.
- Run Average Degree
- Who is in the node with the largest degree centrality? You should check the Data Table tab

Filters		
Statistics		
Settings		
Network Overview		
Average Degree	3.299	Run
Avg. Weighted Degree		Run
Network Diameter		Run
Graph Density		Run
HITS		Run
PageRank		Run
Connected Components		Run

Graph					
Data Table					
Nodes					
Edges					
Configuration					
Add node					
Add edge					
Search/Replace					
Import Spreadsheet					
Export table					
More actions					
Id	Label	Interval	In-Degree	Out-Degree	Degree
42.0	Anzelma	0	3	3	3
69.0	Babet	3	7	10	10
63.0	Bahorel	4	8	12	12
29.0	Bamatabois	5	3	8	8
56.0	BaronessT	0	2	2	2
19.0	Blacheville	4	3	7	7
64.0	Bossuet	3	10	13	13
40.0	Boulatruelle	0	1	1	1
36.0	Brevet	2	4	6	6
75.0	Brujon	0	7	7	7
35.0	Champmathieu	3	3	6	6
6.0	Champtercier	0	1	1	1

Metrics

We will calculate the average path length for the network. It computes the path length for all possible pairs of nodes and give information about how nodes are close from each other.

In the report you get some centrality metrics as closeness, betweenness...

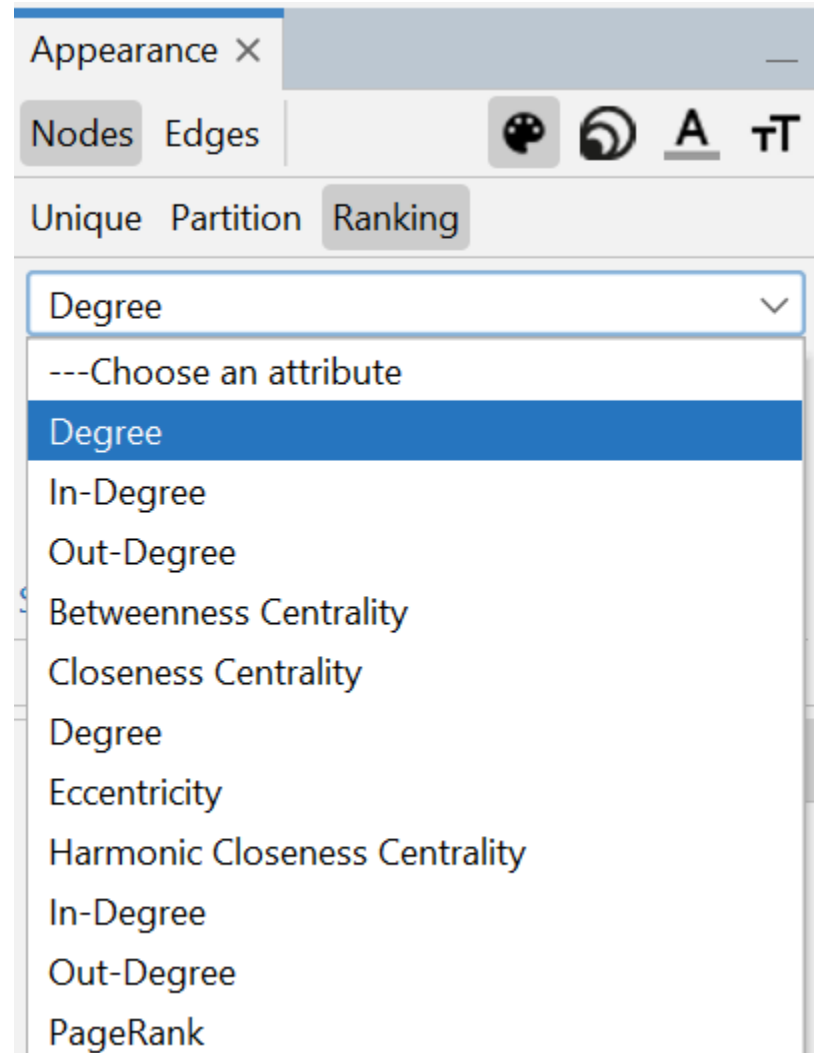
You can check other metrics as centrality ones, density...

Filters	Statistics	×	—
Settings			
Graph Density		Run	
HITS		Run	
PageRank		Run	
Connected Components		Run	
<input type="checkbox"/> Community Detection			
<input checked="" type="checkbox"/> Node Overview			
Avg. Clustering Coefficient		Run	
Eigenvector Centrality		Run	
<input checked="" type="checkbox"/> Edge Overview			
Avg. Path Length		2.4	Run
<input type="checkbox"/> Dynamic			

Ranking (size)

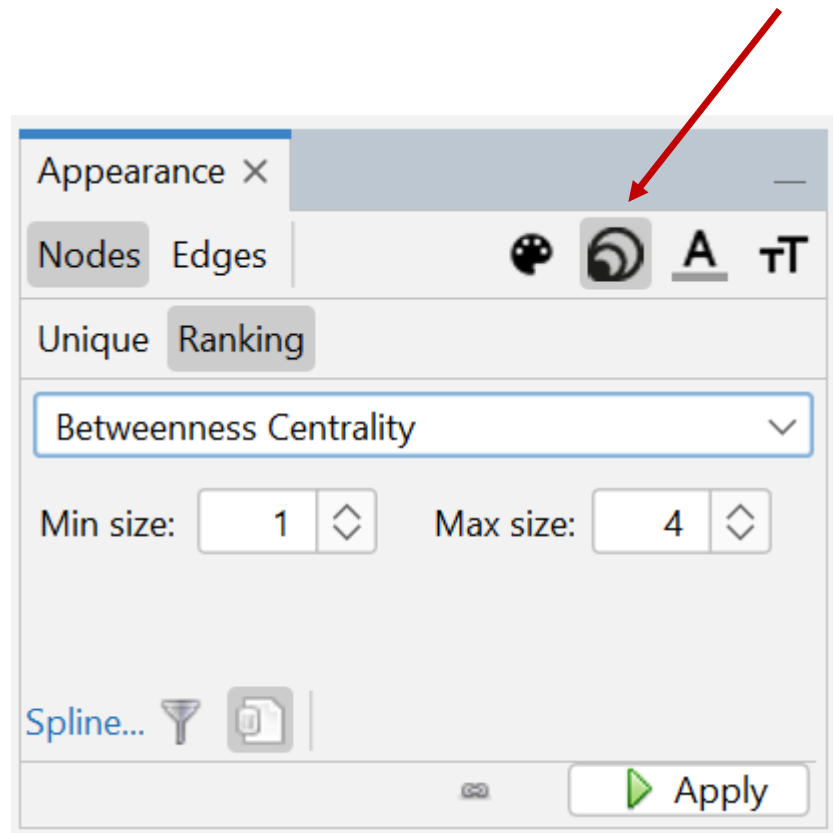
Metrics generates general reports but also results for each node. Thus three new values have been created by the “Average Path Length” algorithm we ran.

- Betweenness Centrality
- Closeness Centrality
- Eccentricity.



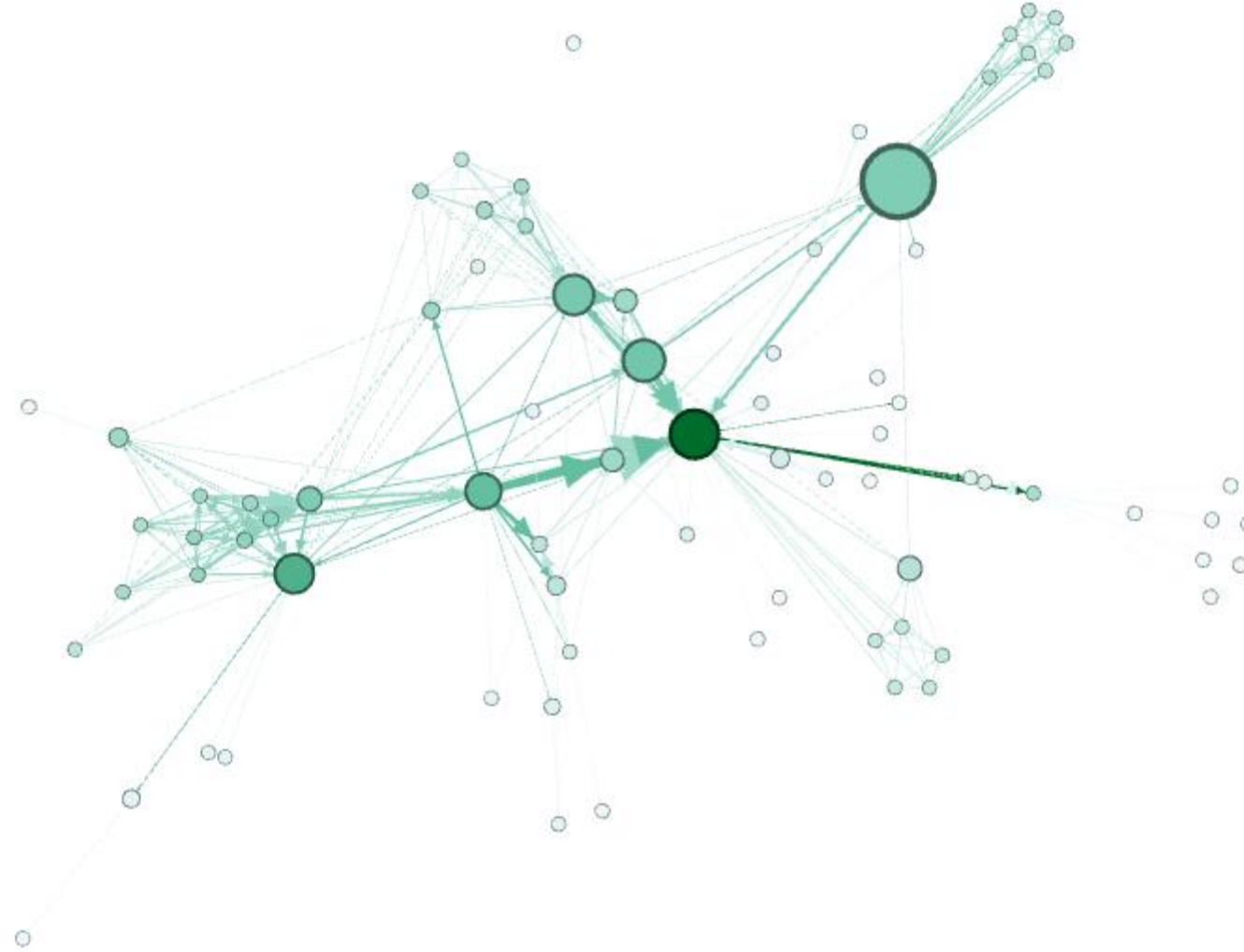
Ranking (size)

The node's size will be set now. Colors remain the "Degree" indicator
Locate the Layout module, on the left panel.



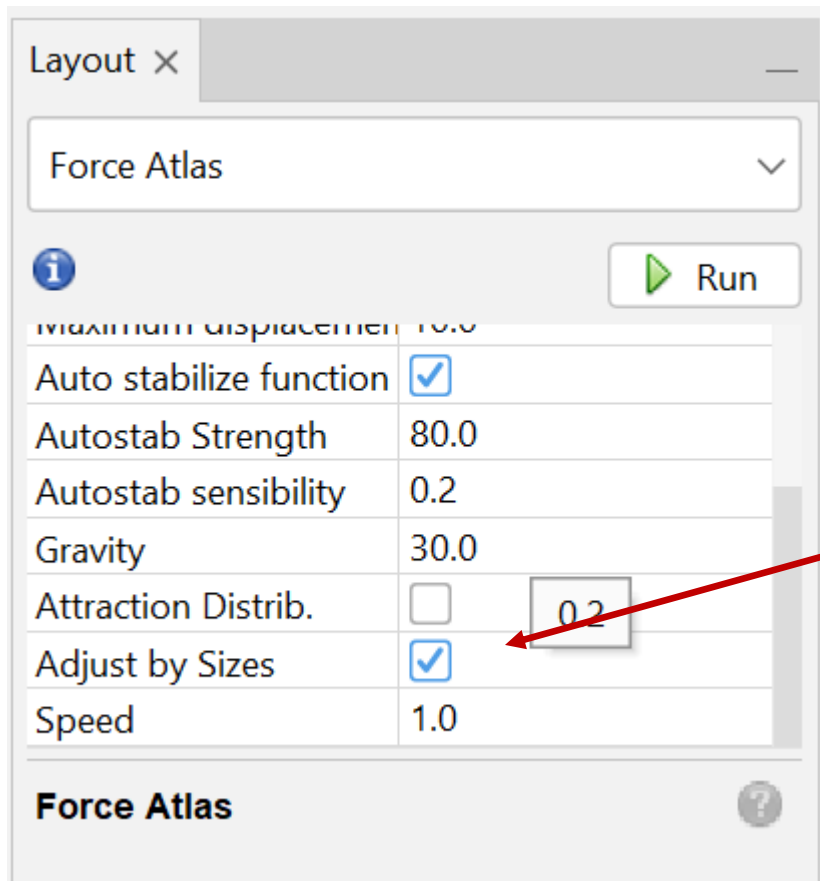
- Select the circles icon in the toolbar for size.
- Change min and max sizes until you find those that work better

You should see a colored and sized graph



Layout again

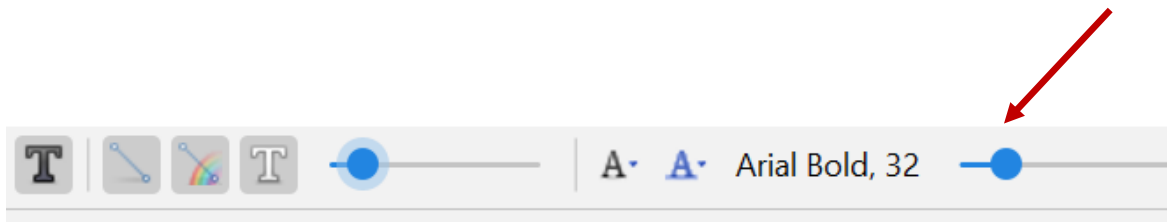
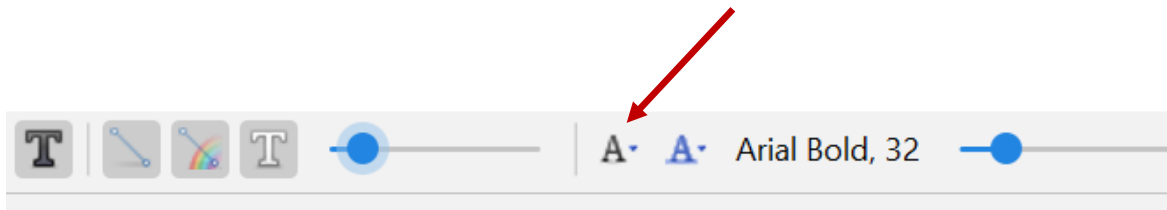
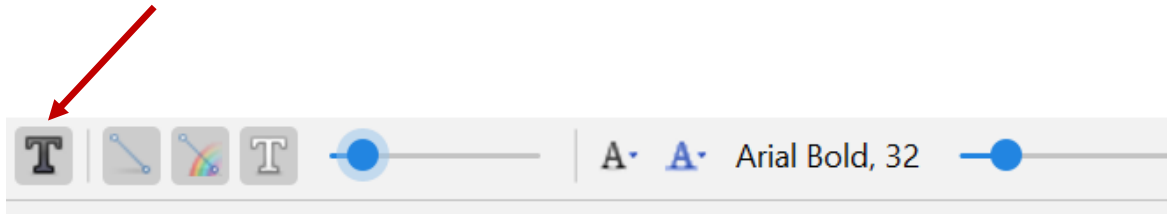
The layout is not completely satisfying, as big nodes can overlap smaller. The “Force Atlas” algorithm has an option to take node size in account when layouting.



- Check the “Adjust by Sizes” option and run again the algorithm for short moment.
- You can see nodes are not overlapping anymore.

Show labels

Let's explore the network more in details now that colors and size indicates central nodes.



- Display node labels
- Set label size proportional to node size
- Set label size with the scale slider

Community detection

The ability to detect and study communities is central in network analysis. We would like to colorize clusters in our example. Gephi implements the Louvain method, available from the Statistics panel.

Modularity settings

Modularity

Community detection algorithm.

☒ Randomize

Produce a better decomposition but increases computation time

☒ Use weights

Use edge weight

Resolution:

1.0

Lower to get more communities (smaller ones) and higher than 1.0 to ge...

Classes start at:

0

OK

Cancel

Filters

Statistics X

Settings

Graph Density

Run

HITS

Run

PageRank

Run

Connected Components

Run

☒ Community Detection

Modularity

Run

Statistical Inference

Run

☒ Node Overview

Avg. Clustering Coefficient

Run

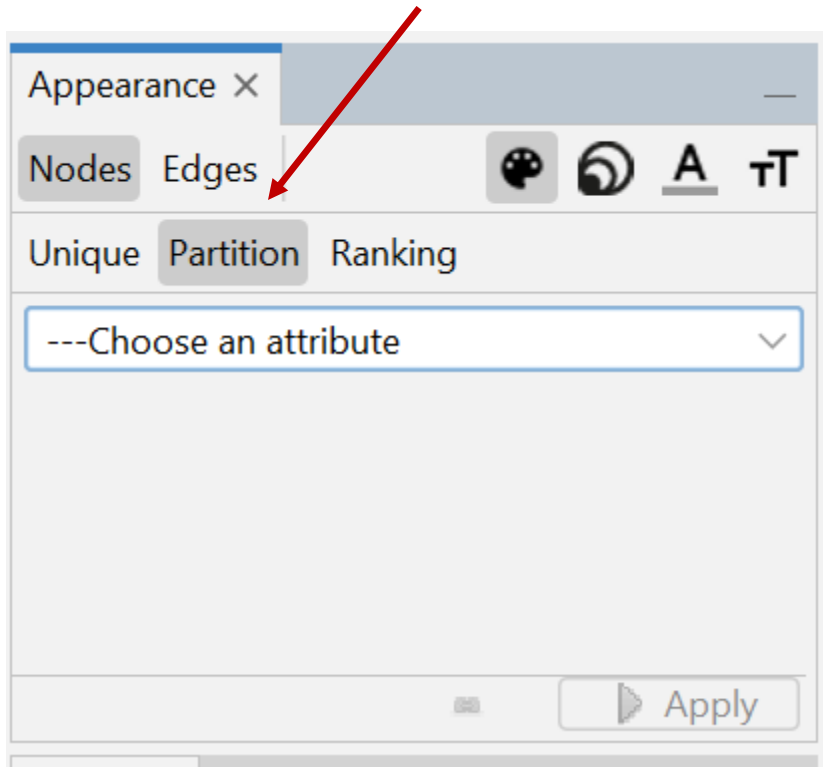
Eigenvector Centrality

Run

Partition

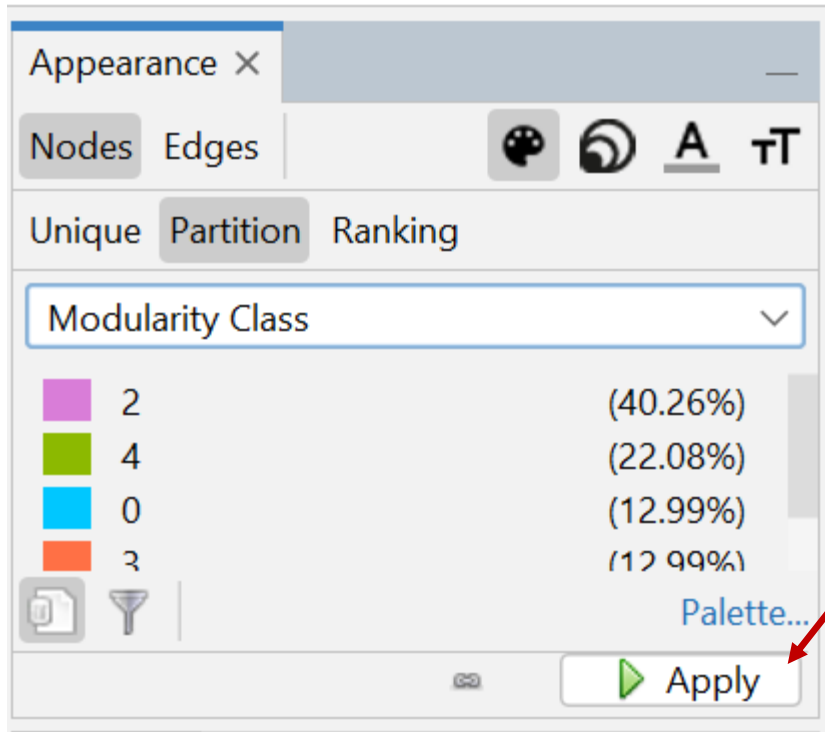
The community detection algorithm created a “Modularity Class” value for each node.

The partition module can use this new data to colorize communities.



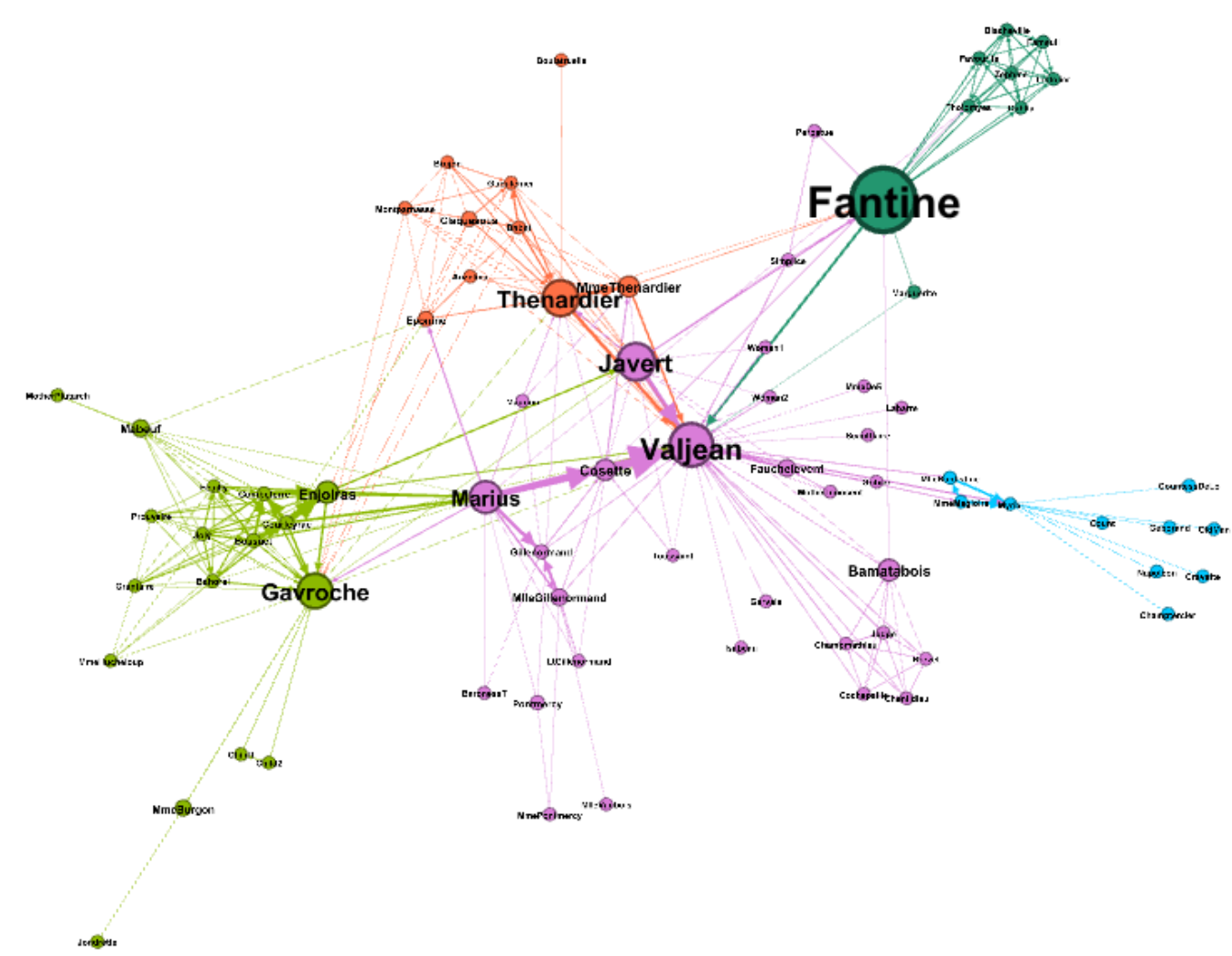
- Select the circles icon in the toolbar for size.
- Change min and max sizes until you find those that work better

Partition



- Select “Modularity Class” in the partition list.
- You can see that 9 communities were found, could be different for you.
- A random color has been set for each community identifier.
- Change it as you prefer

What the network looks like now



Filter

The last manipulation step is filtering. You create filters that can hide nodes and edges on the network. We will create a filter to remove leaves, i.e. nodes with a single edge.

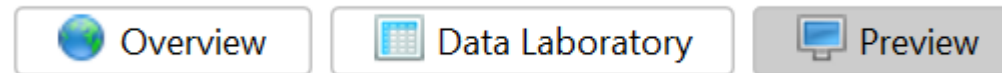
- Select “Degree Range” in the “Attributes” category.
- Click on “Degree Range” to activate the filter
- It shows a range slider and the chart that represents the data, the degree distribution here.
- Move the slider to sets its lower bound to 2.
- Nodes with a degree inferior to 2 are now hidden.

The screenshot displays the software's 'Filters' and 'Statistics' tabs. Under the 'Filters' tab, the 'Library' section is expanded, showing a tree structure of filter categories. The 'Range' category is selected, and the 'Degree Integer (No)' filter is highlighted. Below the library, the 'Queries' section shows the 'Range (Degree)' filter is active. At the bottom, the 'Range (Degree) Settings' panel shows a range slider with a lower bound of 2 and an upper bound of 36. A small chart to the left of the slider shows the degree distribution, with the area under the curve for degrees less than 2 shaded in red, indicating the nodes that will be hidden.

Preview

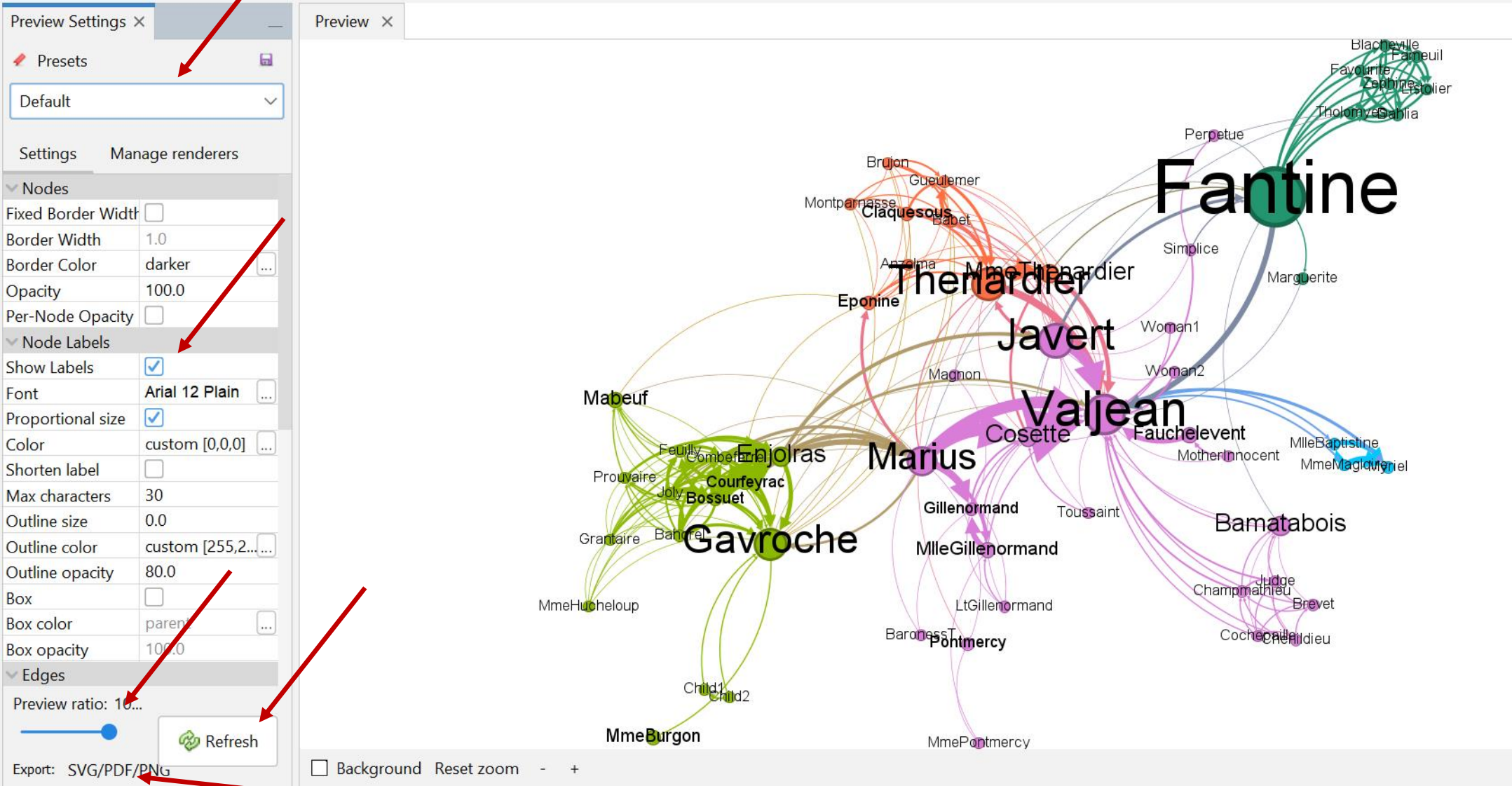
Before exporting your graph as a SVG or PDF file, go to the Preview

- Select the “Preview” tab in the banner



- Click on Refresh to see the preview
- If the graph is big, reduce the “Preview ratio” slider to 50% or 25% to display a partial graph.
- In the Node properties, find “Show Labels” and enable the option.
- Preview Settings supports Presets, click on the presets list and try different configurations.
- Export it as SVG

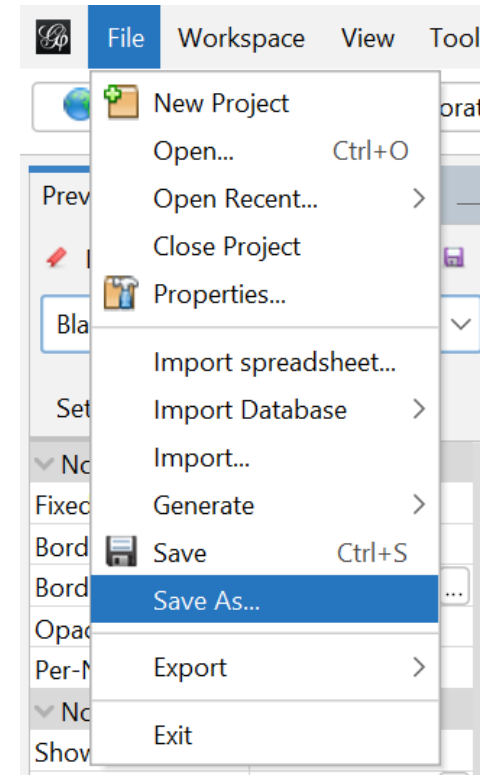
Filter



Export as SVG

Save your project

Saving your project encapsulates all data and results in a single session file.



Conclusion

In this tutorial you learned the basic process to open, visualize, manipulate and render a network file with Gephi.

