

Gephi for Network Analysis

Siti Lei

Data and Visualization Librarian, Duke Kunshan University

Last updated: December 8, 2025

This tutorial introduces the basics of network analysis and how to use Gephi to import and visualize data, apply different layouts, install and use plugins, and export the graphs. For demonstration, we will use a subset of character relationships from Sherlock Holmes stories.

Before we start	2
Introduction	2
Glossary	2
About the tool	3
Setup	3
Datasets	3
Node Table	3
Edge Table	4
Install Plugins	4
Import Network Data	5
Visualize Data	8
Left Toolbar	8
Appearance	15
Bottom Toolbar	26
Layout	28
Filters & Statistics	39
Filters	40
Statistics	43
Export the Graph	45
Capture the current view	45
Save the graph	46
Export as a website	50
Publish to GitHub	52
Learning Resources	52

Before we start

Introduction

In this tutorial, you will learn the basic concepts of network analysis and how to use Gephi to create network graphs for qualitative research, from importing and visualizing data to installing plugins, applying different layouts, and exporting the final graphs. For demonstration purposes, we will use a subset of character relationships from the Sherlock Holmes stories.

To give you a sense of its wide applicability, network analysis can be used in many contexts, such as social relationships, communication patterns, transportation systems, biological interactions, or character connections in a story. Some project examples include [DKU's Natural and Applied Sciences Research](#), [Map of the Interstate Highway System in the United States](#), and [Map of Scientific Paradigms](#).

For any questions about network analysis and Gephi, contact the Data and Visualization Librarian, Siti Lei (siti.lei@dukekunshan.edu.cn) for support.

Glossary

Network analysis is a method used to study how different items (**nodes**) and their connections (**edges**) form patterns and relationships. A network graph visually represents these connections, helping audiences understand how things are linked, how information or influence flows, and which nodes play important roles within a system.

Nodes (aka. actors, vertices, points, or entities) are individual items that participate in relationships or are connected to others in a network. They can represent many things, such as people, groups, organizations, locations, events, or concepts. Each node can also have additional information attached to it, such as a category, a location, or a size value.

Edges (aka. links, connections, ties, relations) refer to the relationships or interactions between nodes in a network. They show how two nodes are connected and can represent many types of relationships, such as communication or influence. Each edge can also include additional information, such as the type of relationship. Multiple edges between a pair of items is allowed.

Link direction in a network can be either **directed** or **undirected**. A **directed edge** shows a one-way relationship, such as an email sent or a citation from one paper to another. An **undirected edge** refers to a two-way or mutual relationship, such as friendship or co-authorship.

Degree refers to the number of edges connected to a node which shows how many direct relationships that node has within the network. A node with a high degree is more connected, while a node with a low degree has fewer connections.

Weight refers to the strength, frequency, or importance of an edge between two nodes. A higher weight means the two nodes interact more often or have a stronger relationship.

About the tool

We will use Gephi, an open-source and free software compatible with Windows, Mac, and Linux devices, in this workshop. Visit [Gephi](#) for more information about the tool.

Setup

Download [Gephi](#) and complete the installation on your local computer.

Datasets

Download the sample datasets: [Workshop_Files.zip](#)

Prepare Data

In network analysis, it is recommended to have a node table and an edge table that complement each other. A node table lists all the individual elements in your network. An edge table describes how those nodes are connected.

Node Table

Each row in the node table describes a single node and should include a node label. You can also add a unique identifier (such as the **ID** column shown in the example sheet below), which helps simplify the creation of the edge table. Additional relevant attributes can be added as needed.

Example:

1	ID	Name	Role	Role Type	Citizenship
2	1	Sherlock Holmes	Consulting detective	Protagonist	British
3	2	John Watson	Doctor & Holmes's	Protagonist	British
4	3	Mary Morstan	Watson's wife; clos	Supporting character	British
5	4	James Moriarty	Criminal masterm	Antagonist	British
6	5	Eurus Holmes	Holmes's sister (BB	Antagonist	British
7	6	Irene Adler	Adler: "The Woman	Supporting character	American
8	7	Mycroft Holmes	Government official	Supporting character	British
9	8	Martha Louise Hudson	Landlady of 221B	Supporting character	British
10	9	Inspector Lestrade	Detective inspector	Ally	British
11	10	Colonel Sebastian Moran	Moriarty's chief her	Antagonist	British
12	11	Molly Hooper	Pathologist assistir	Supporting character	British
13	12	Inspector Gregson	Detective inspector	Ally	British
14	13	Inspector Bradstreet	Detective inspector	Ally	British
15	14	Inspector Hopkins	Detective inspector	Ally	British
16	15	Mr. Holmes	Father of Sherlock	Minor character	British

Edge Table

Each row describes a single edge between a **Source** and a **Target** node. If you assigned unique identifiers in the node table, you can use those identifiers as the **Source** and **Target** (see the example sheet below). If not, you can use the node labels instead.

Example:

1	Source	Target	Label	Weight	Type
2		1	5 Family Member		4 Undirected
3		1	7 Family Member		4 Undirected
4		1	15 Family Member		2 Undirected
5		1	16 Family Member		2 Undirected
6		1	8 Landlady & Tenant		3 Undirected
7		2	3 Family Member		4 Undirected
8		2	17 Family Member		4 Undirected
9		3	17 Family Member		4 Undirected
10		1	2 Partner & Colleague		5 Undirected
11		1	11 Partner & Colleague		3 Undirected
12		1	12 Partner & Colleague		3 Undirected

Install Plugins

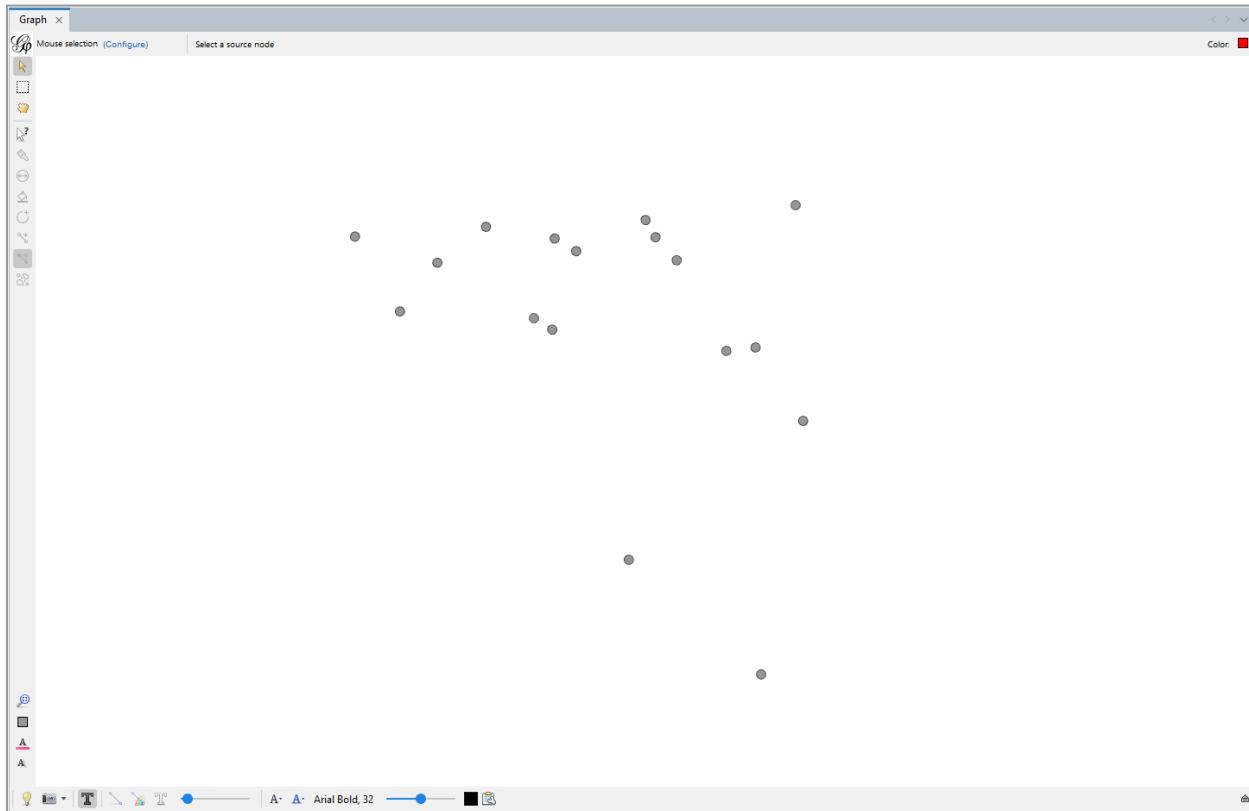
Gephi provides a series of plugins that allows better experience and function in working with network analysis. Let's try some of them in this tutorial.

1. Open **Gephi** on your computer. In the welcome popup, click on **New Project**.
2. In the top menu, click on **Tools** then select **Plugins**. Note that you can also change the language in the **Tools** menu.
3. Under Available Plugins, select **Publish your network to the web**, **SigmaExporter**, and **Circular Layout**. Then click **Install**.
4. Click **Next**.
5. Check **I accept the terms in all of the license agreements**, then click **Install**.
6. When finished, select **Restart Now**, then click **Finish**.
7. In the new pop-up, click on **Do not save for now**.
8. Gephi will then relaunch on your computer. In the welcome window, click **New Project**.
9. Next, go to **Tools** in the top menu and select **Plugins**. Under the **Installed** tab, you will find **Publish your network to the web**, **SigmaExporter**, and **Circular Layout** listed.

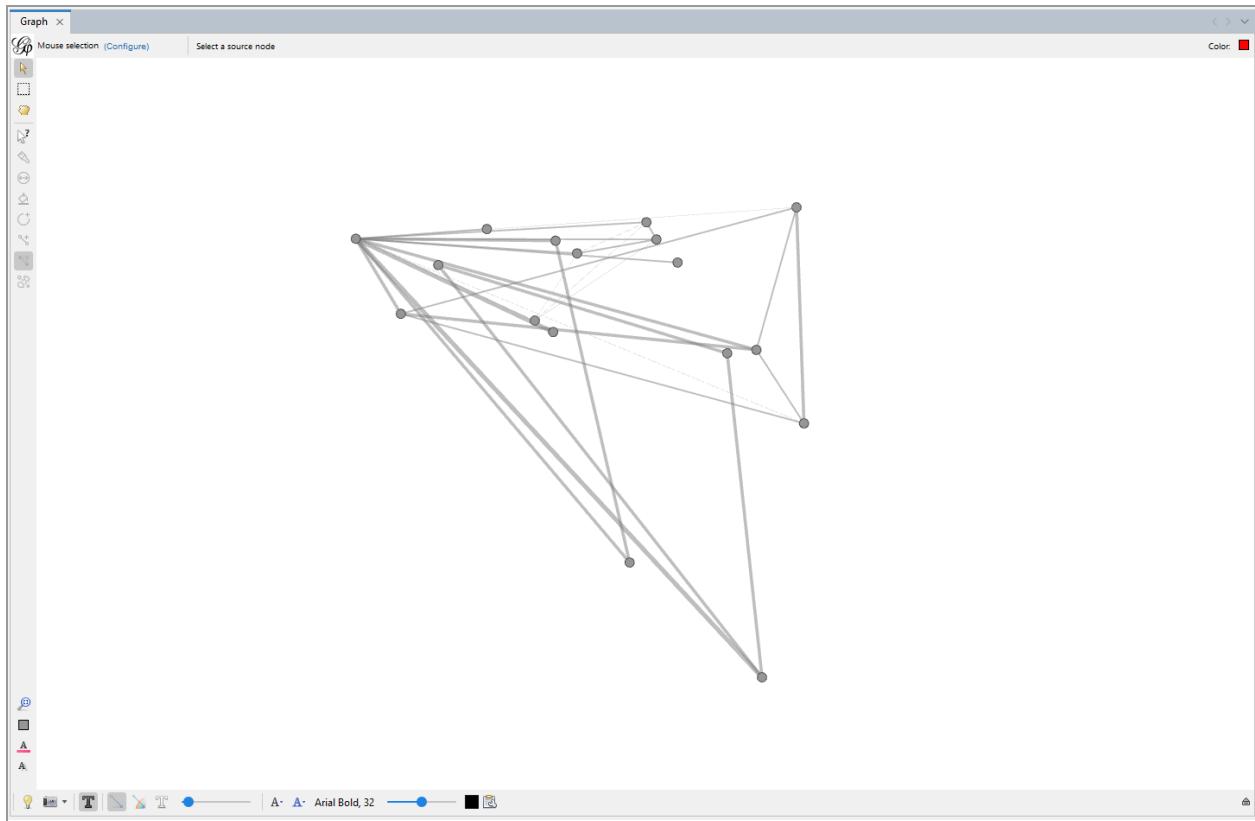
Import Network Data

Gephi allows you to import network data from various file types, including CSV files, Excel sheets, and graph files such as GEXF, GML, and GraphML. In this tutorial, we will demonstrate how to import the nodes and edges tables in CSV file format.

1. Go to the **File** tab, then select **Import spreadsheet** from the menu.
2. Select the **nodes.csv** from the sample datasets downloaded in your computer, then click **Open**.
3. Click **Next >**.
4. Review the data type for each column, then click **Finish**.
5. Check **Append to existing workspace**, then click **OK**.
6. A few dots will now appear on the canvas, with each dot representing a node. Note that the exact layout may differ, as the positioning of data is randomly assigned.



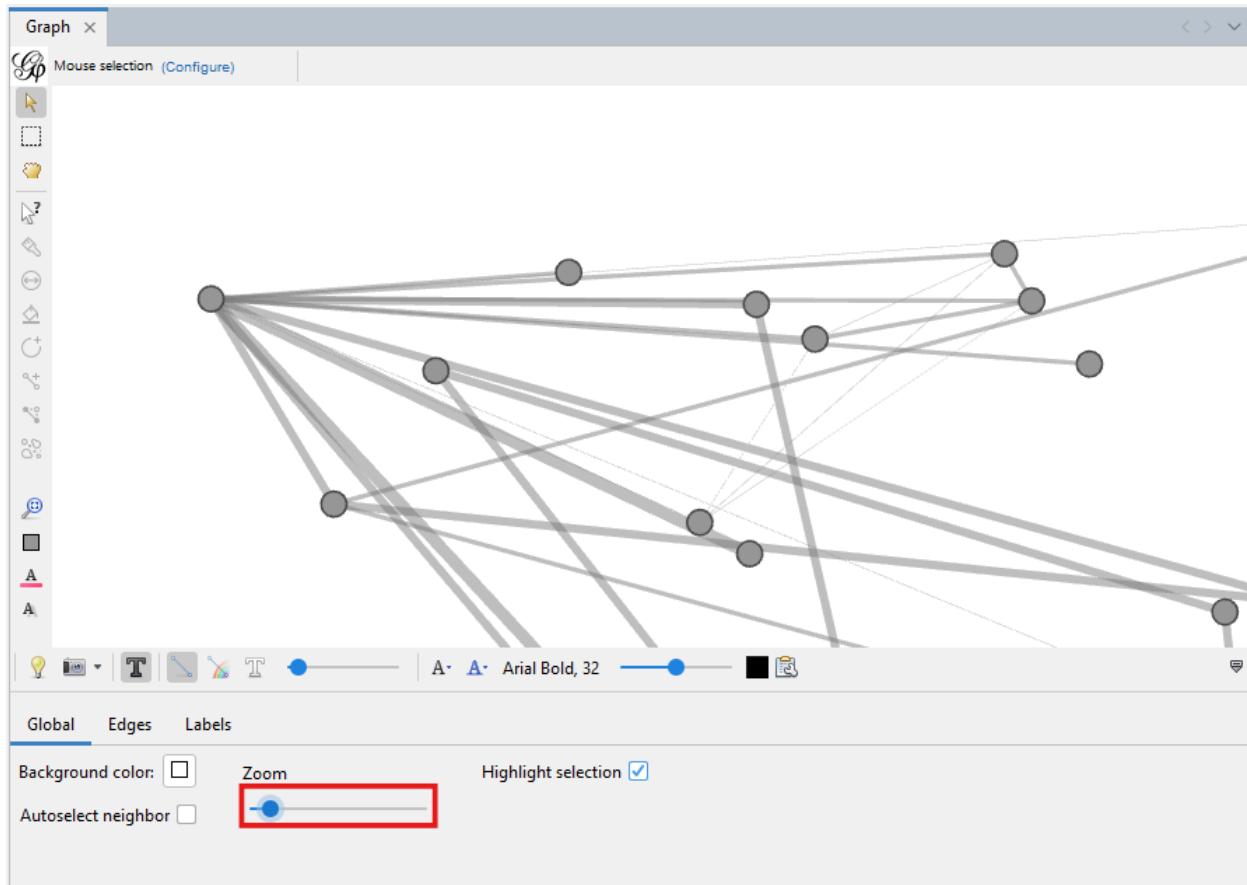
7. Next, go to the **File** tab again, select **Import spreadsheet** from the menu.
8. Select the **edges.csv** from the sample datasets downloaded on your computer, then click **Open**.
9. Click **Next >**.
10. Review the data type for each column, then click **Finish**.
11. Check **Append to existing workspace**, then click **OK**.
12. Now, few connection lines will show up connecting each dot in the canvas. Each line represents one edge. Note that the exact layout may differ, as the positioning of data is randomly assigned.



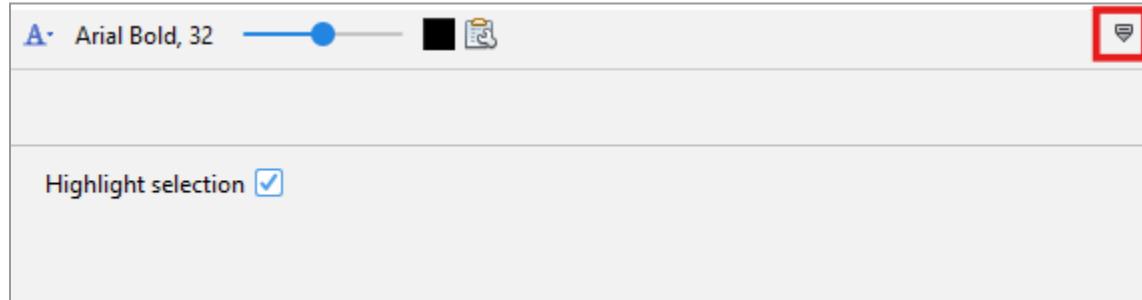
13. Click the arrow on the right side of the bottom panel.



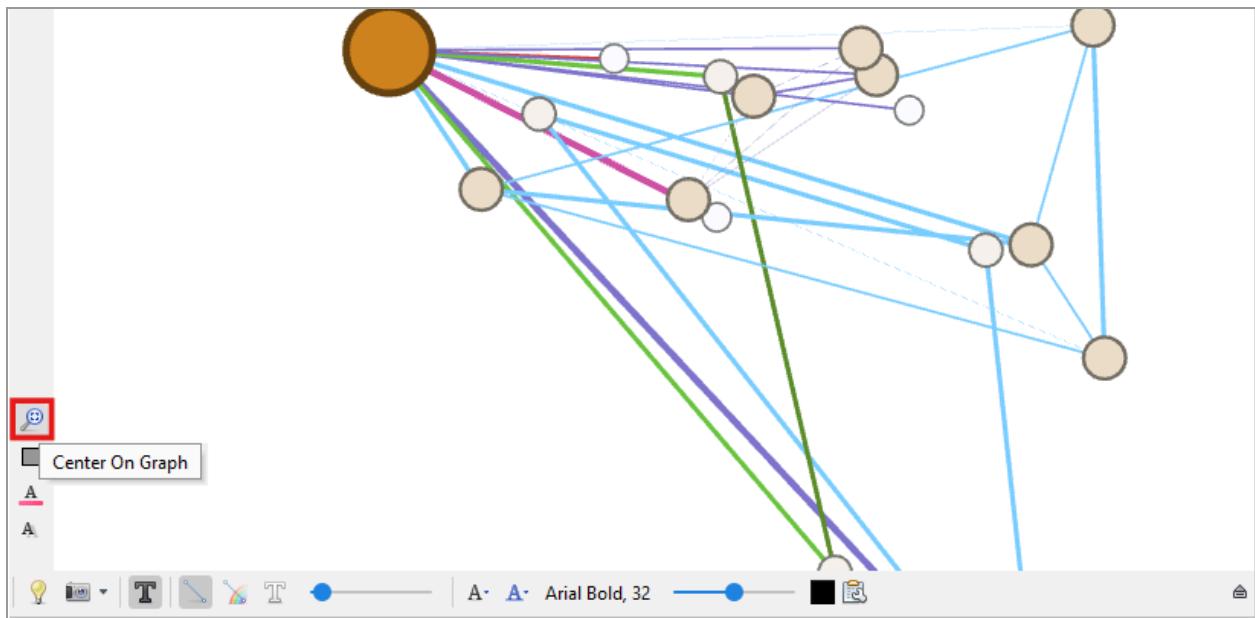
14. Drag the scale bar under **Zoom** to zoom in and view more detail. You can also use your mouse wheel to zoom. Note that you can also change the background color here.



15. Click the arrow on the right side of the bottom panel again to hide it.



16. To reset the view, click on the **Center on graph** icon on the left-bottom toolbar.

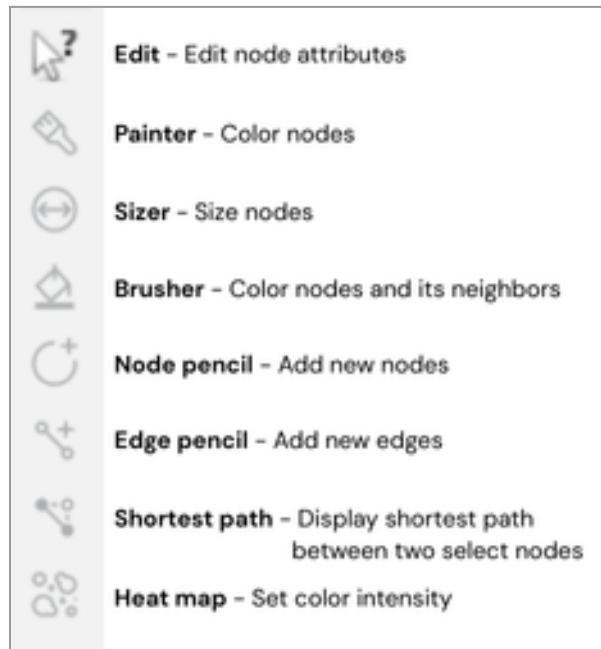


Visualize Data

In Gephi, you can customize the visual appearance of your network by adjusting the nodes, edges, background, and layout positions using a variety of built-in and plugin tools.

Left Toolbar

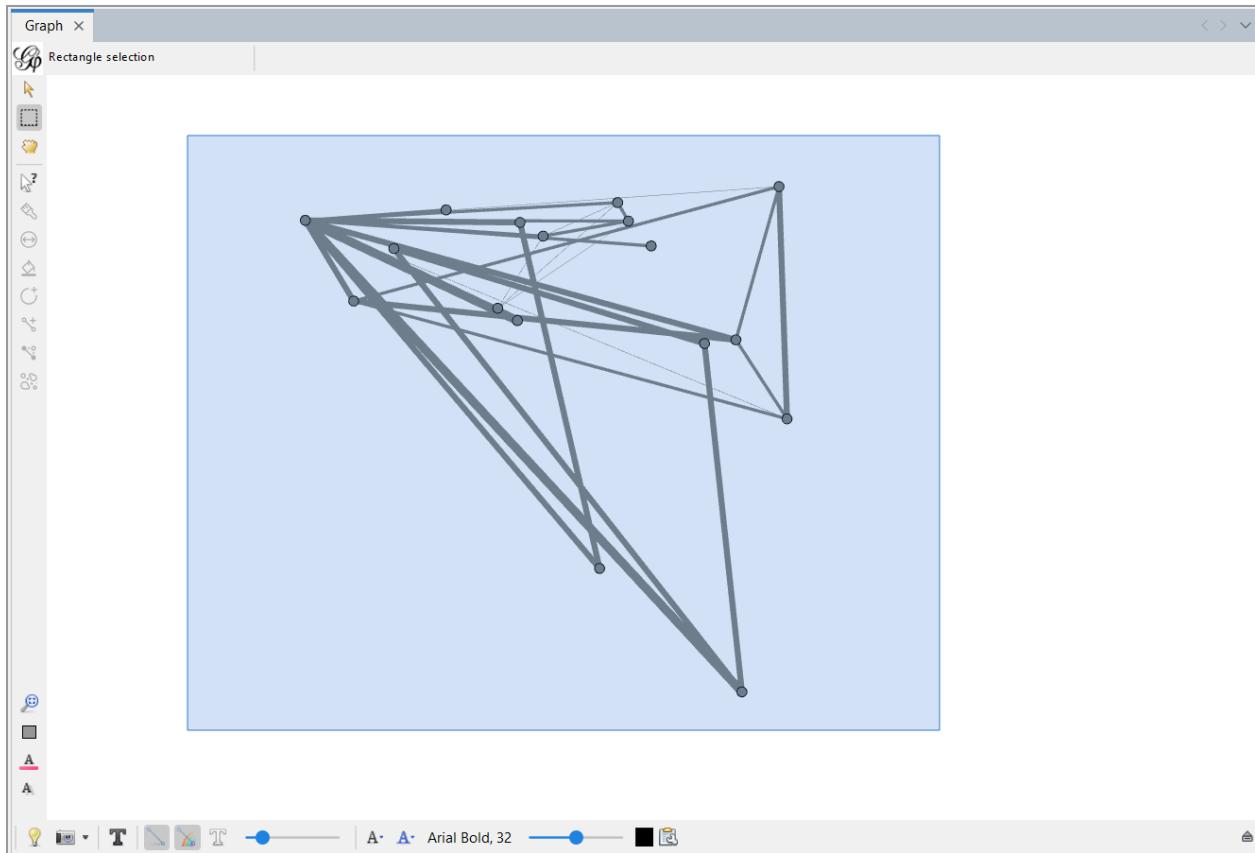
1. The left toolbar provides editing tools for modifying the visuals of the nodes and edges in the graph. We will only cover some of the tools here, but it is encouraging to explore and try them on your own after the tutorial.



2. Click on the **Rectangle selection** icon on the left toolbar.



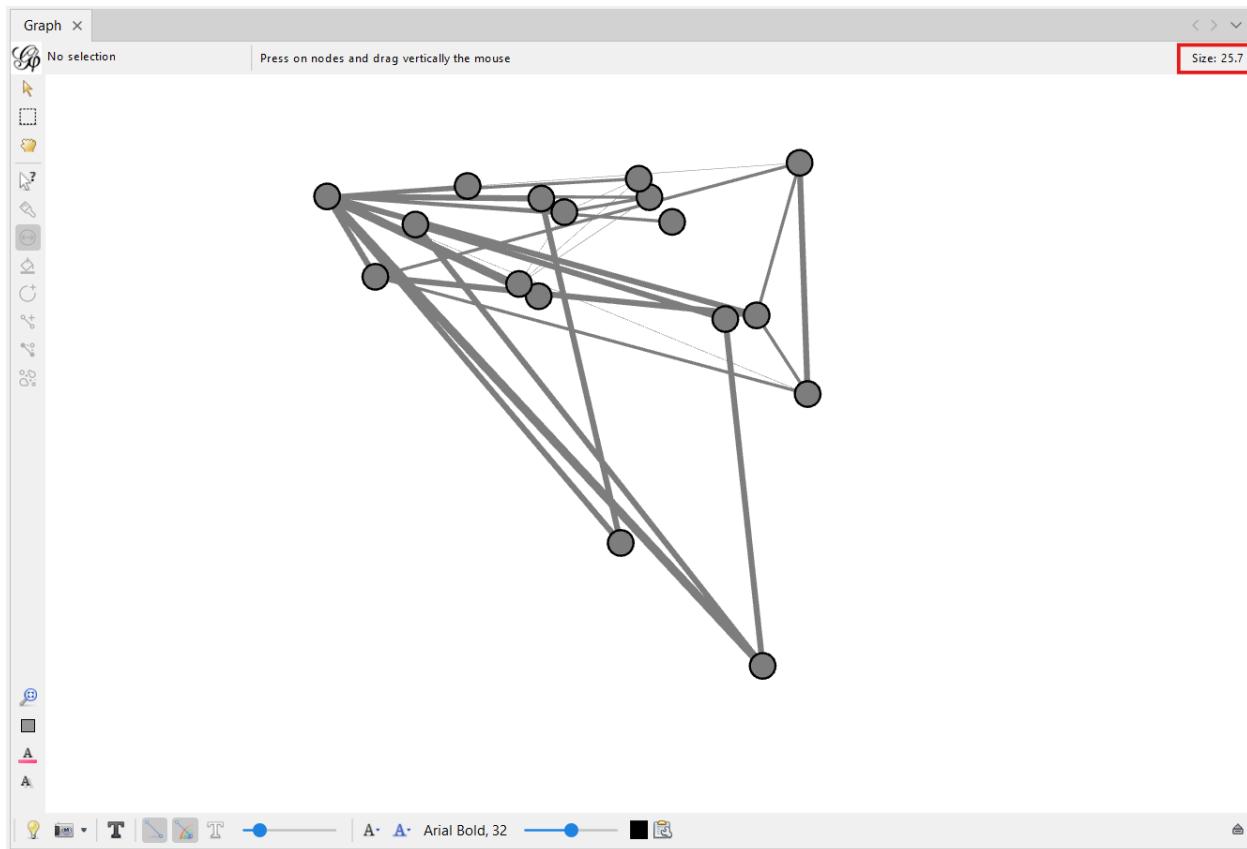
3. Click and drag to draw a square that selects all the imported items.



4. Select the Size icon in the toolbar.



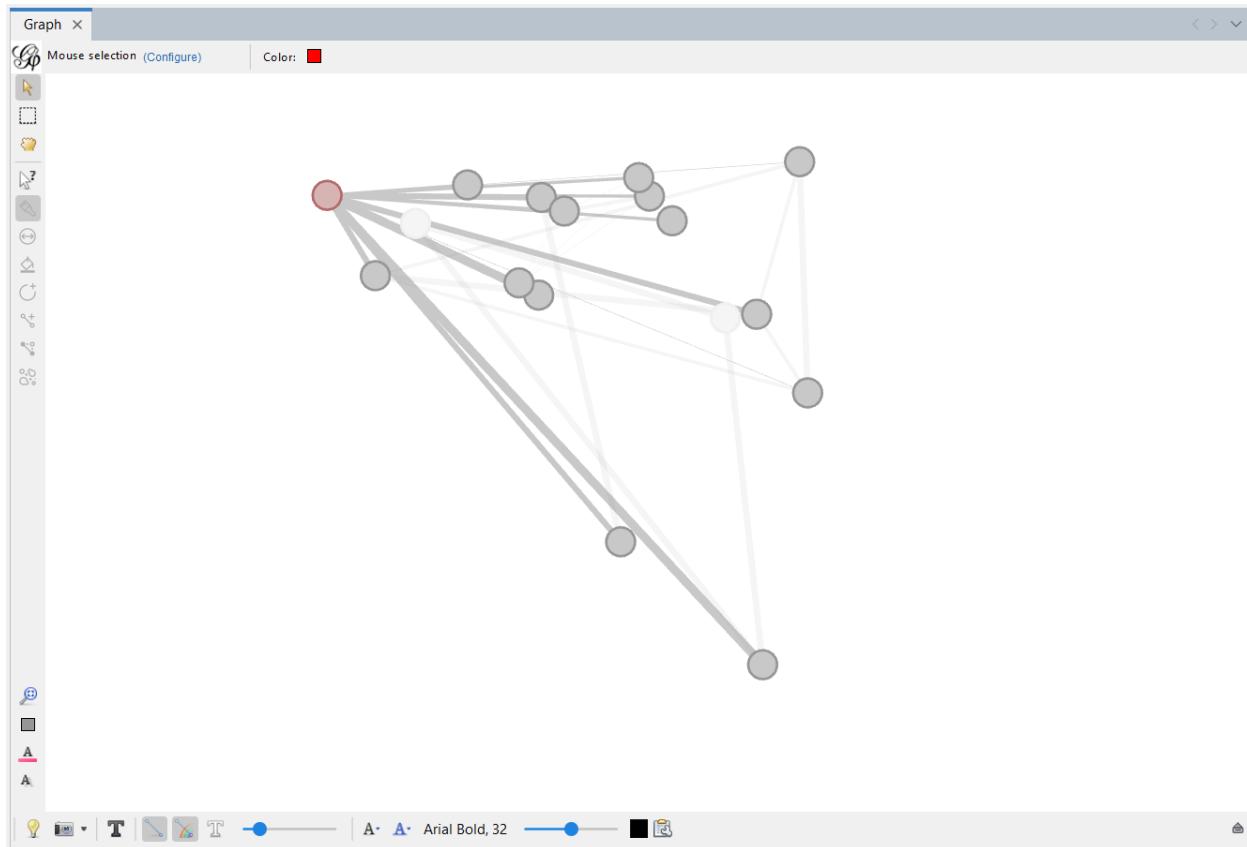
5. Click anywhere on the canvas, then drag the mouse vertically to adjust the node size. Click again to confirm the change. The updated size value will appear in the top-right corner.



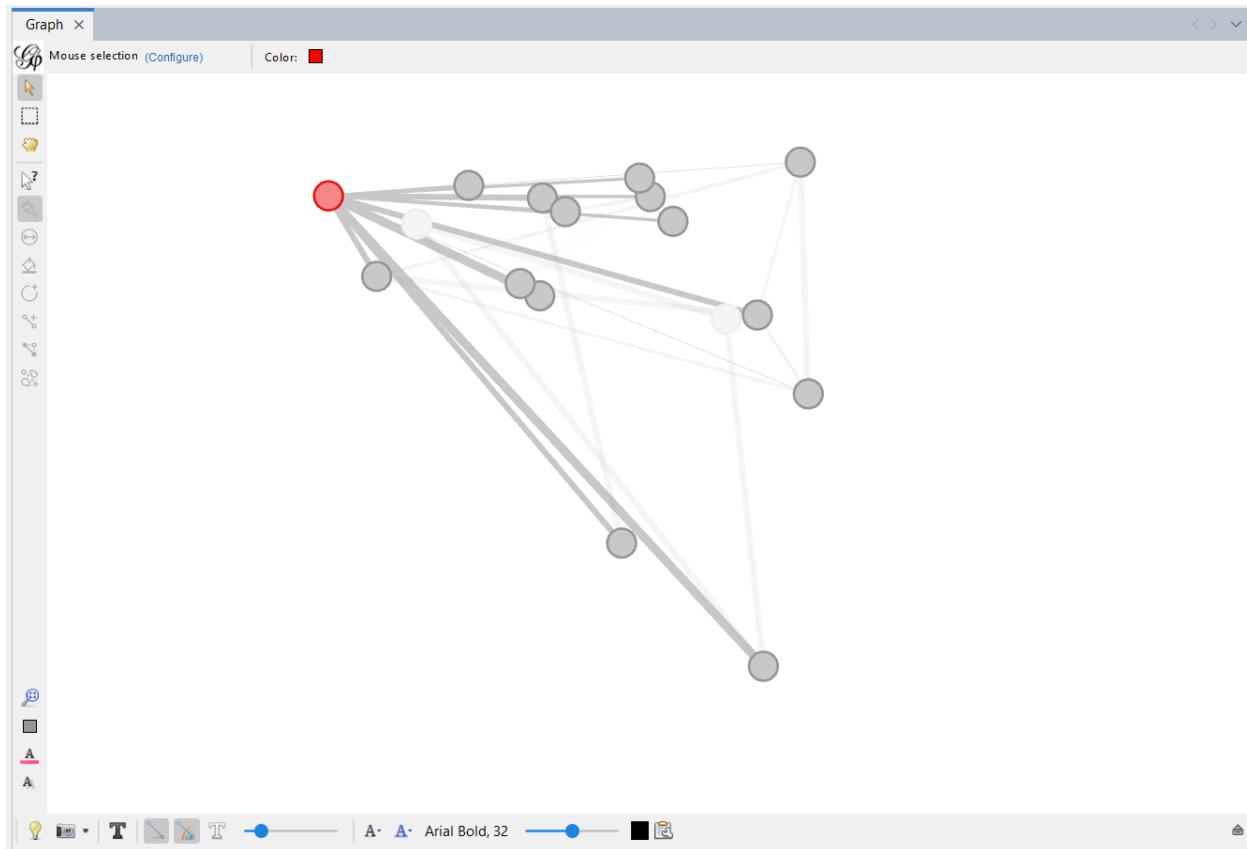
6. Now, select the **Direct selection** icon and **Painter** icon in the toolbar.



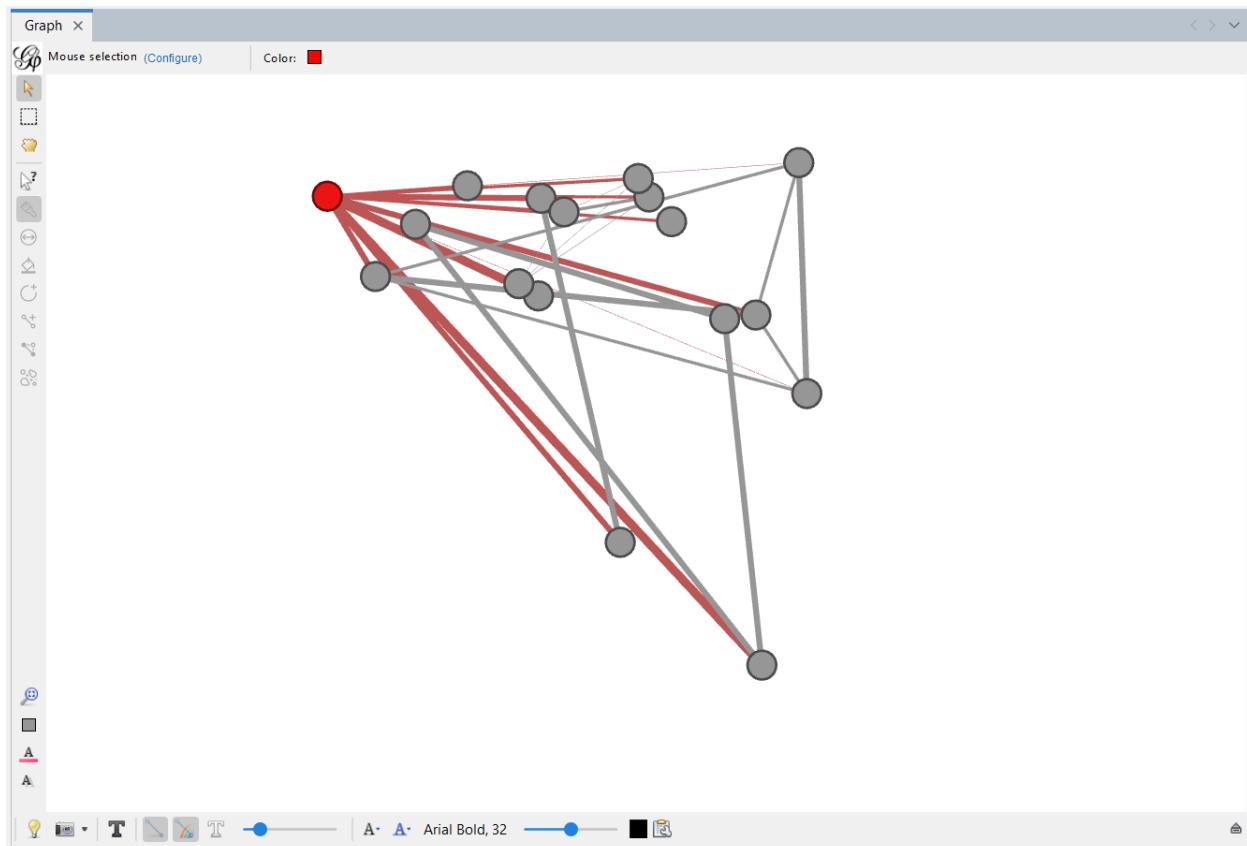
7. Click on any node to apply a color to it.



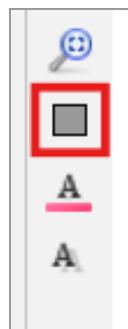
8. Double-click the same node to intensify the color. Each additional click makes the color more vibrant.



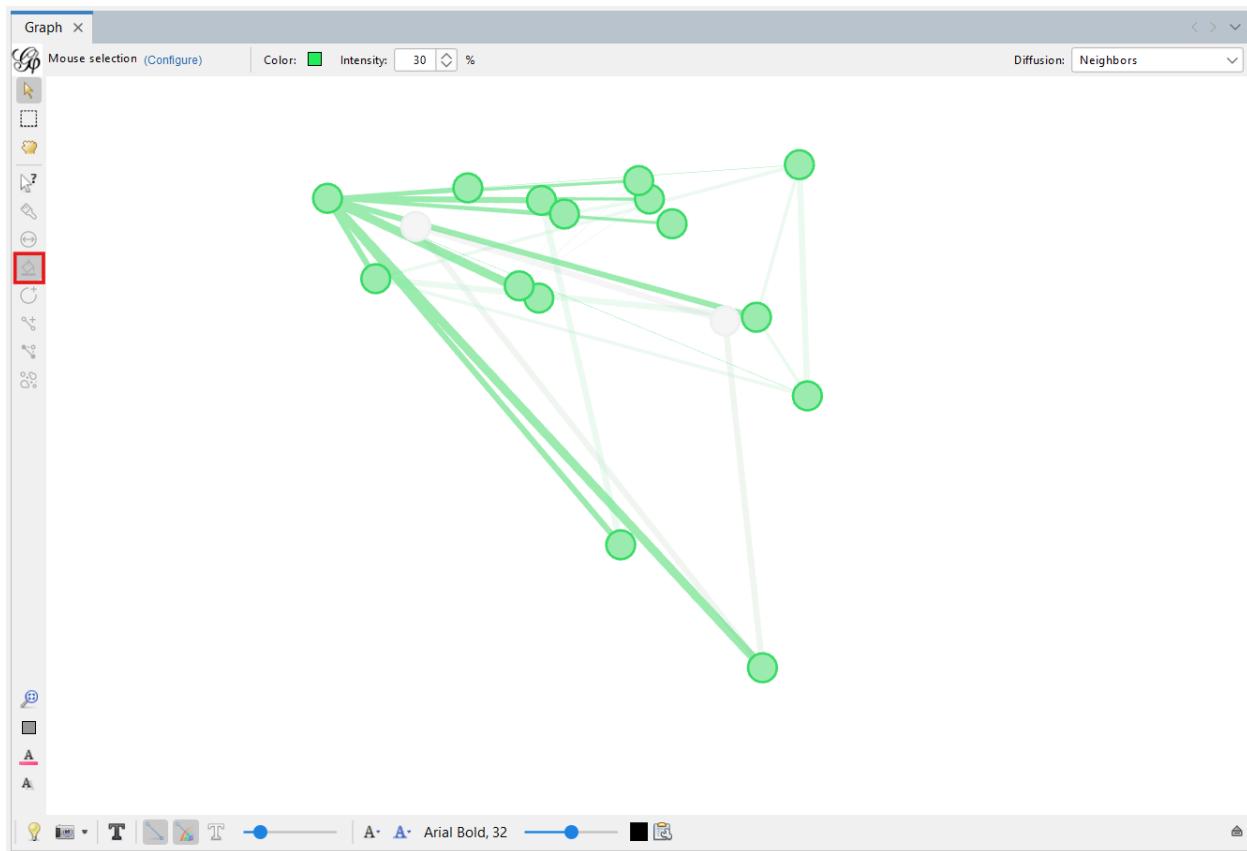
9. Move the cursor away. The selected node and all edges directed toward it will now be highlighted in color.



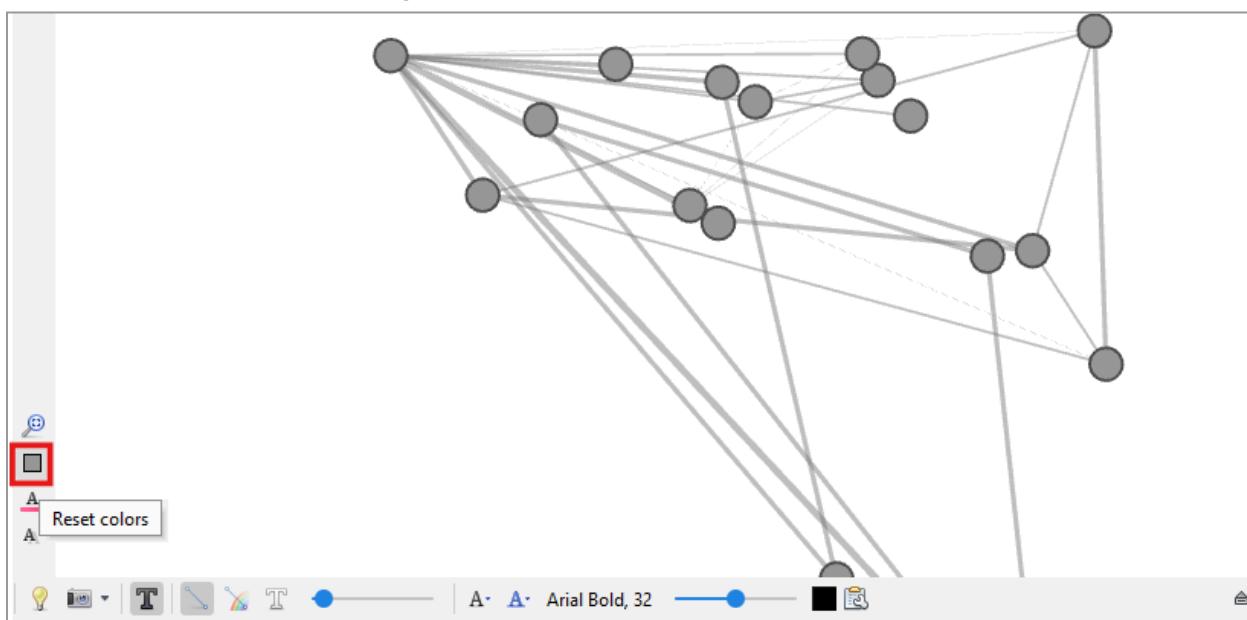
10. Click on the Reset color icon at the bottom of the toolbar. This will reset the color setting.



11. Next, select the **Brush** icon and switch the color to green. Then click a node to apply the color. This will color the selected node, all connected nodes, and the edges between them.

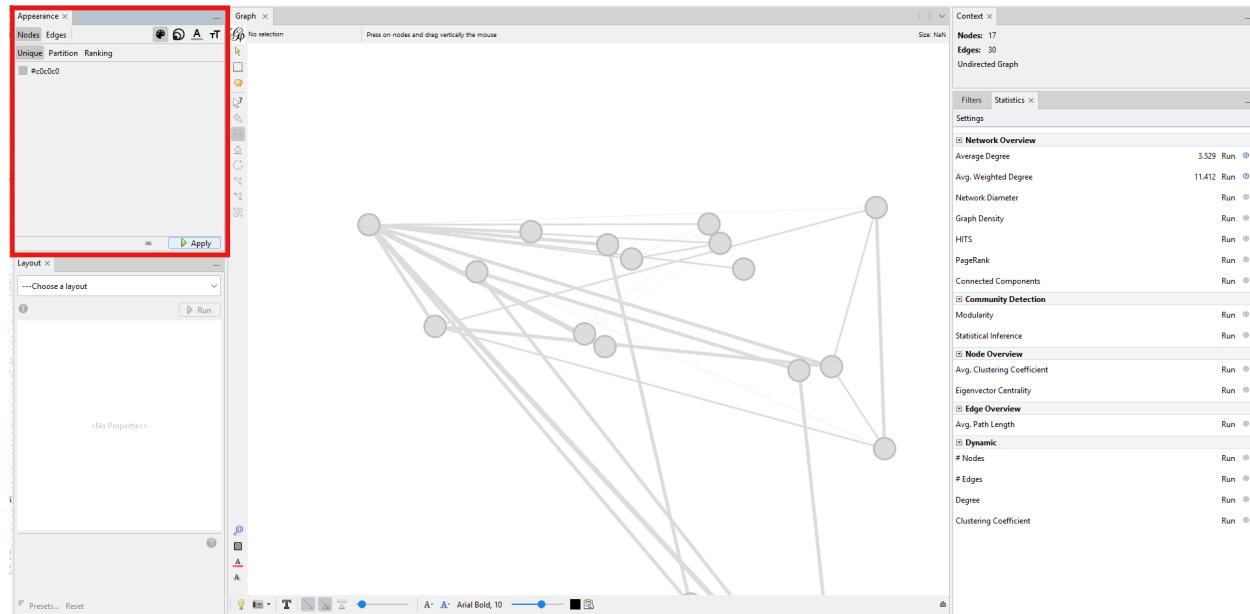


12. To reset the color setting, click on the **Reset colors** icon in the left-bottom toolbar.

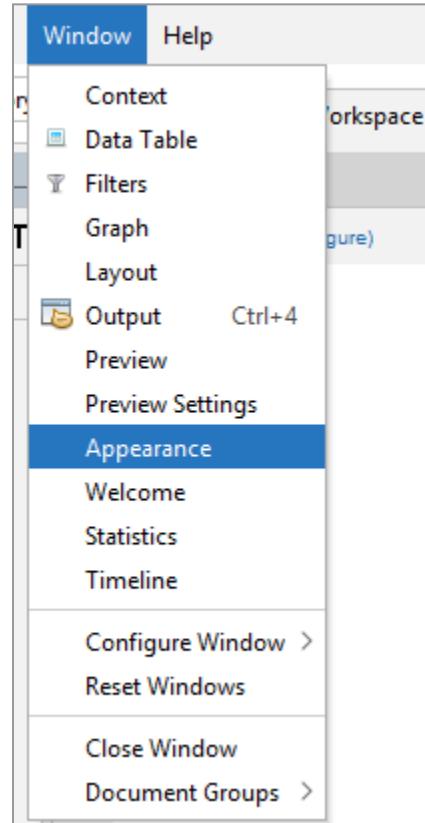


Appearance

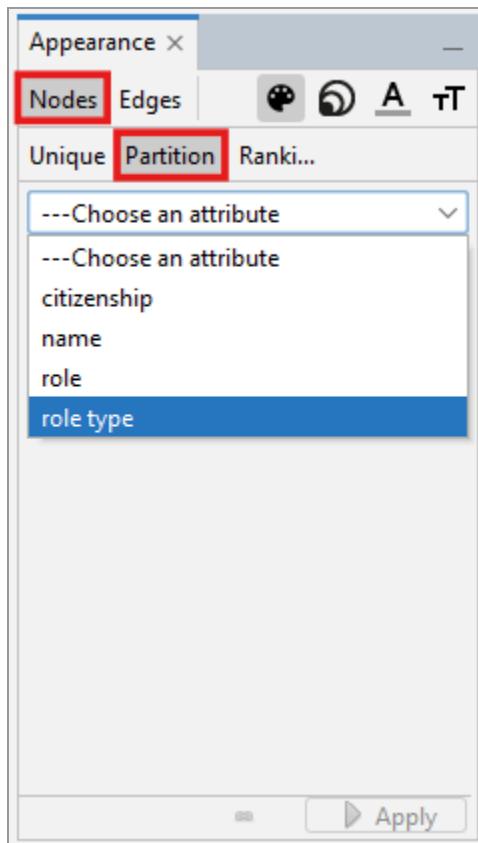
- As default, the Appearance window appears in the top left of the workplace.



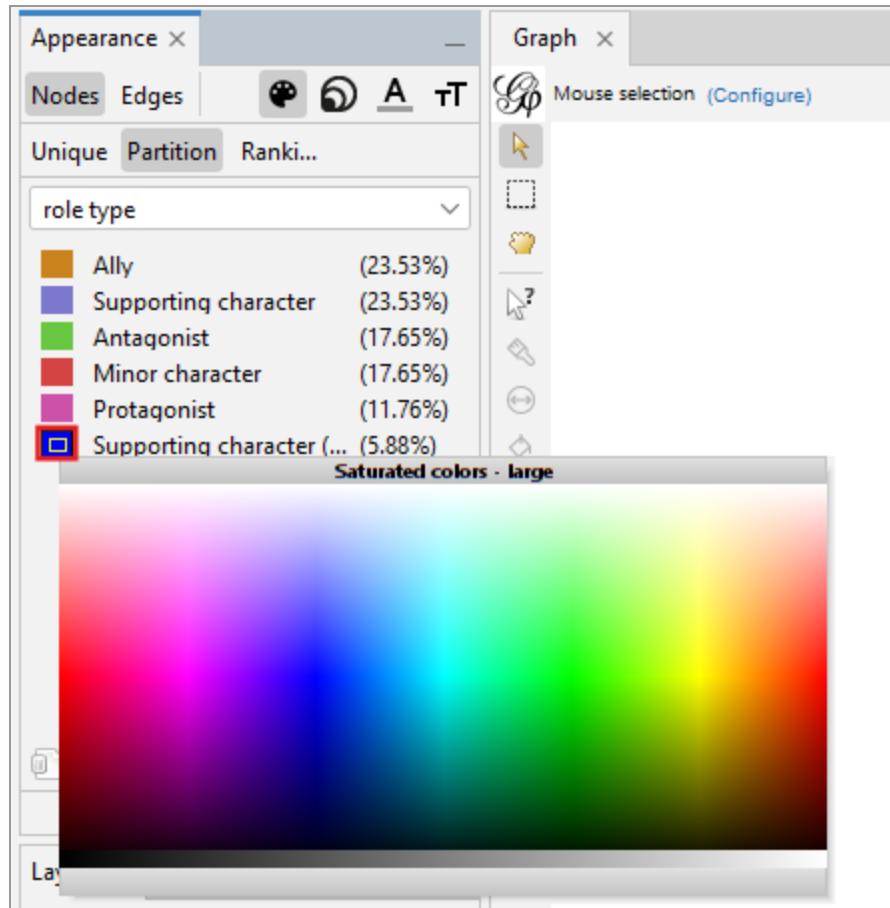
- If you don't have that open, go to the Window tab, then select Appearance. The Appearance window should appear in the workplace again. You can also select Reset Windows to go back to the default view.



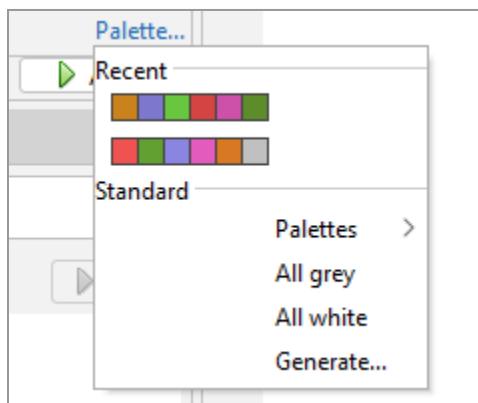
3. In the Appearance window, select **Nodes**, then **Partition**, choose the role types as the attribute.



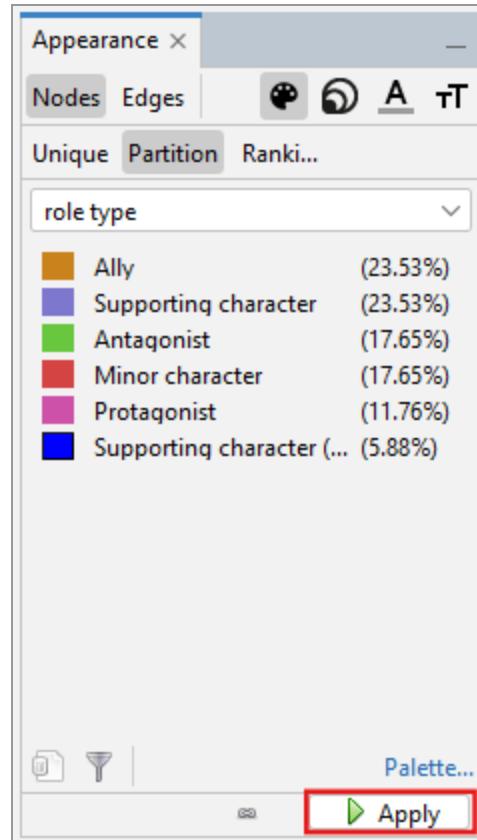
4. The system will automatically assign colors to different values. To adjust a single color, click the color block and choose a new color from the pop-up.



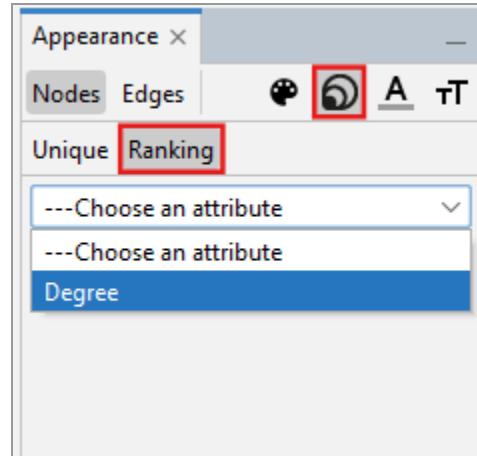
5. Alternatively, you can click **Palette...** in the bottom-right corner and select a new color scale from the list.



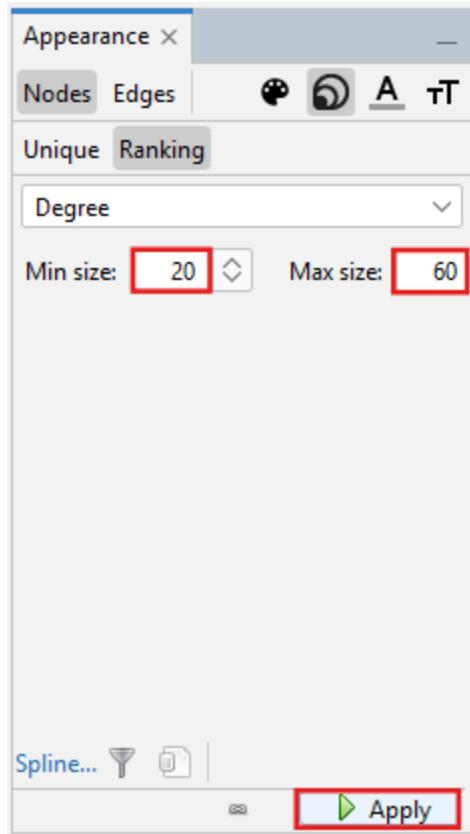
6. Once the color selection is complete, click **Apply**. The nodes on the canvas should now be colored.



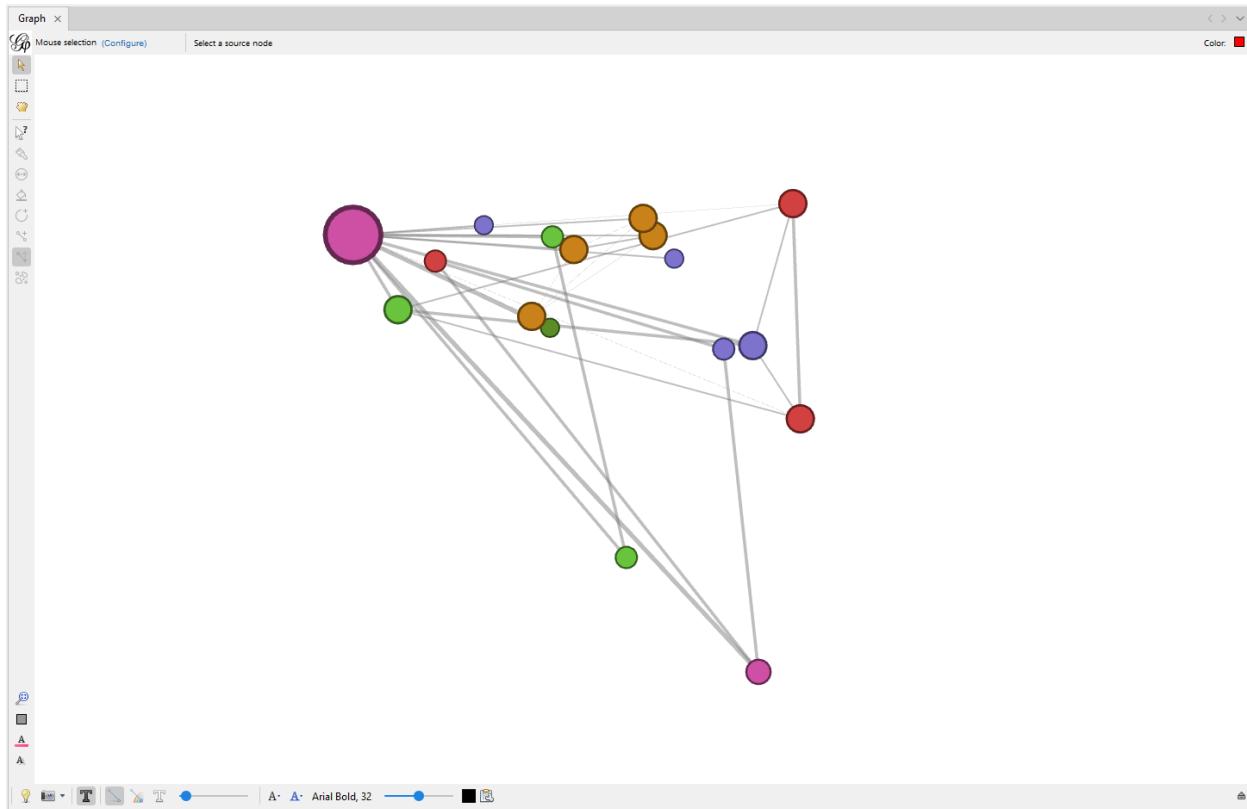
7. Now, let's switch to the **Size** tab, then **Ranking**, select **Degree** as the attribute.



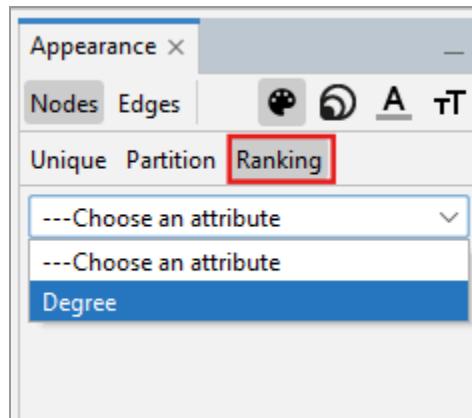
8. Set the minimal size to **20**, and maximal size to **60**. Then, click **Apply**.



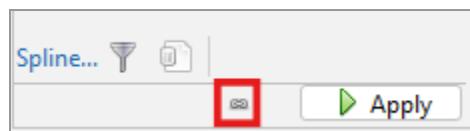
9. The graph should now appear similar to the example below. Note that the exact layout may differ, as the positioning of nodes and edges is randomly assigned.



10. Now, switch to the **Ranking** tab, and select **Degree** as the attribute.



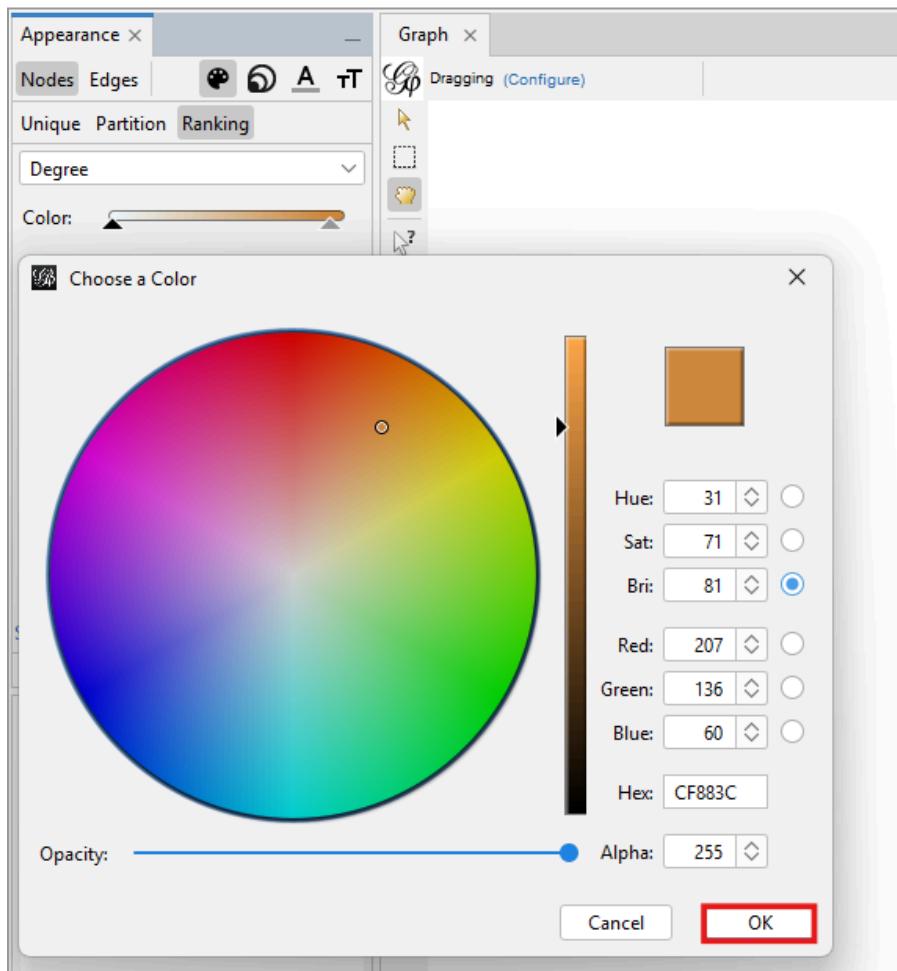
11. Click on the link icon next to the **Apply** button - this will enable the auto apply feature.



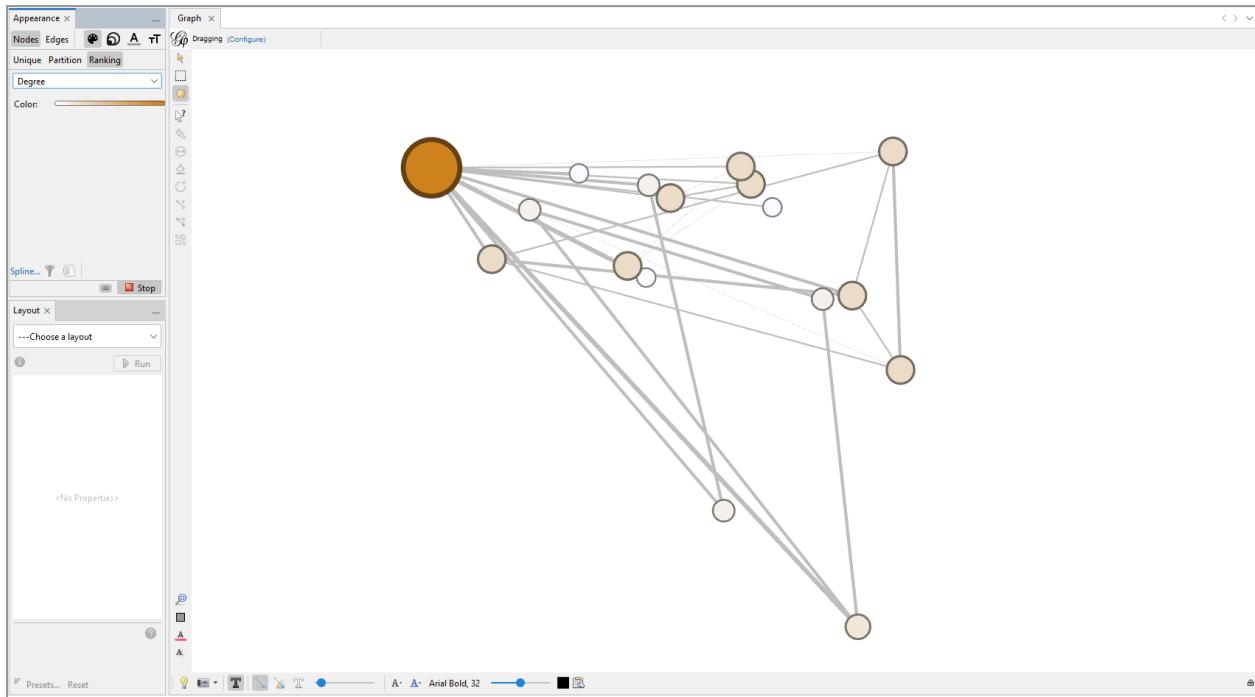
12. Click on **Auto Apply**.



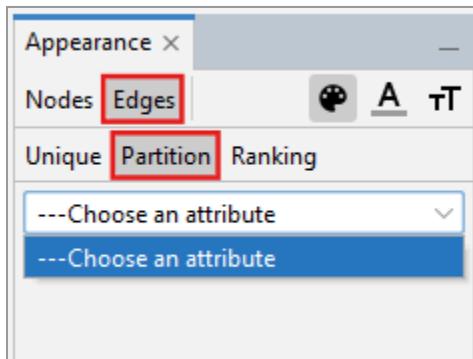
13. Double-click the arrows to adjust the colors, then click **OK** to confirm the changes.



14. The nodes should update their colors automatically. You can delete an arrow by selecting it and dragging it downward, or add a new arrow by clicking anywhere along the bar.



15. Next, we will configure and categorize the edges based on the relationship type, which is the **Label** field in our dataset. In the **Appearance** window, switch to **Edges**, then select **Partition**, and click the arrow to expand the attribute list.

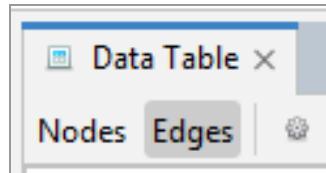


16. You will notice that **Label** is not in the list. This is because we did not set its data type during import (see Step 11 in **Import Data**). To correct this, we need to assign the appropriate data type to the column. Switch to the **Data Laboratory** view, where all imported datasets are displayed.

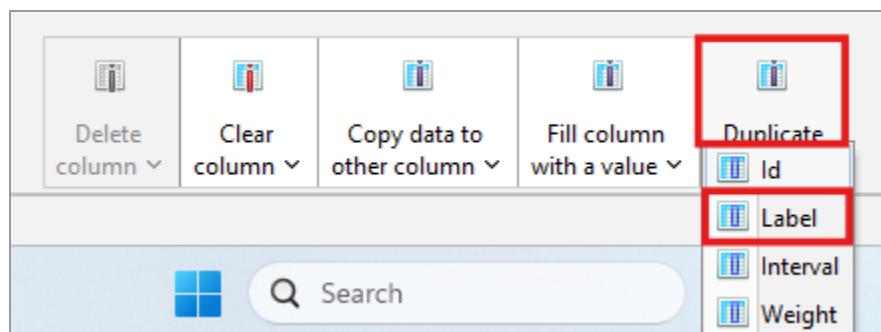
The screenshot shows the Data Laboratory interface with a data table titled 'Data Table X'. The table has columns: Id, Label, Interval, name, role, role type, and citizenship. The data includes entries for Sherlock Holmes, John Watson, Mary Morstan, James Moriarty, Eurus Holmes, Irene Adler, Mycroft Holmes, Martha Louise Hudson, Inspector Lestrade, Colonel Sebastian Moran, Molly Hooper, Inspector Gregson, Inspector Bradstreet, Inspector Hopkins, Mr. Holmes, Mrs. Holmes, and Rosamund Mary Watson. The citizenship column shows British for most characters except Irene Adler (American). The bottom toolbar contains buttons for Add column, Merge columns, Delete column, Clear column, Copy data to other column, Fill column with a value, Duplicate column, Create a boolean column from regex match, Create column with list of regex matching groups, Negate boolean values, and Convert column to dynamic.

Id	Label	Interval	name	role	role type	citizenship
1			Sherlock Holmes	Consulting detective	Protagonist	British
2			John Watson	Doctor & Holmes's partner	Protagonist	British
3			Mary Morstan	Watson's wife; close associate...	Supporting character	British
4			James Moriarty	Criminal mastermind; Holme...	Antagonist	British
5			Eurus Holmes	Holmes's sister (BBC)	Antagonist	British
6			Irene Adler	Adler: "The Woman"; key adv...	Supporting character (not vill...	American
7			Mycroft Holmes	Government official & Holme...	Supporting character	British
8			Martha Louise Hudson	Landlady of 221B Baker Street	Supporting character	British
9			Inspector Lestrade	Detective inspector	Ally	British
10			Colonel Sebastian Moran	Moriarty's chief henchman, a...	Antagonist	British
11			Molly Hooper	Pathologist assisting Holmes ...	Supporting character	British
12			Inspector Gregson	Detective inspector	Ally	British
13			Inspector Bradstreet	Detective inspector	Ally	British
14			Inspector Hopkins	Detective inspector	Ally	British
15			Mr. Holmes	Father of Sherlock Holmes	Minor character	British
16			Mrs. Holmes	Mother of Sherlock Holmes	Minor character	British
17			Rosamund Mary Watson	Watson's daughter	Minor character	British

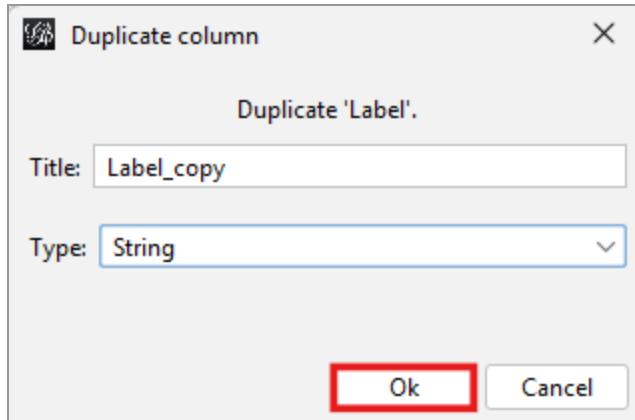
17. In the top-left corner, click **Edges** to switch to the edges data.



18. In the button toolbar, click on the **Duplicate column**, then select **Label**.



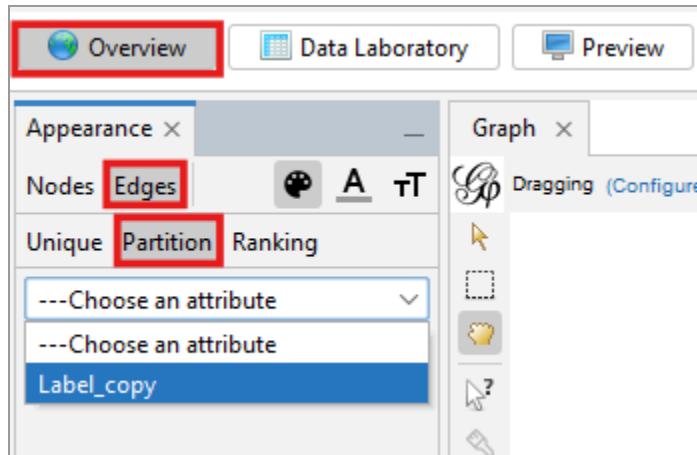
19. In the pop-up, enter a new name for the column and set the type to **String**, then click **Ok**.



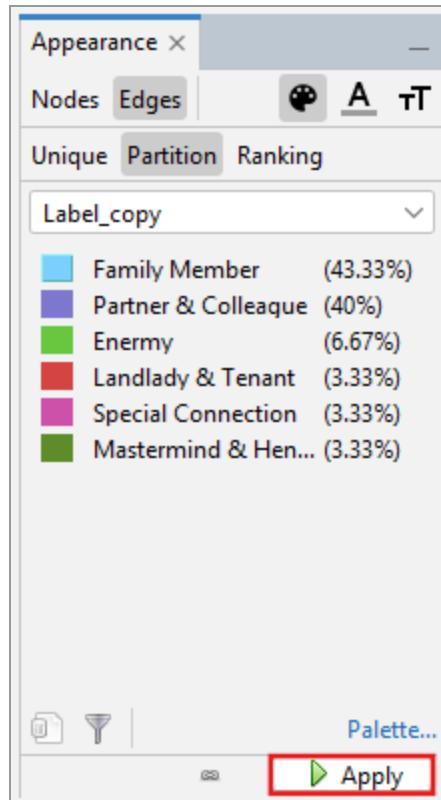
20. A new column **Label_copy** now appears. You may delete the original **Label** column, but we will leave it in this tutorial.

Label	Interval	Weight	Label_copy
Family Member		4.0	Family Member
Family Member		4.0	Family Member
Family Member		2.0	Family Member
Family Member		2.0	Family Member
Landlady & Tenant		3.0	Landlady & Tenant
Family Member		4.0	Family Member
Family Member		4.0	Family Member
Family Member		4.0	Family Member
Partner & Colleague		5.0	Partner & Colleague
Partner & Colleague		3.0	Partner & Colleague
Partner & Colleague		3.0	Partner & Colleague

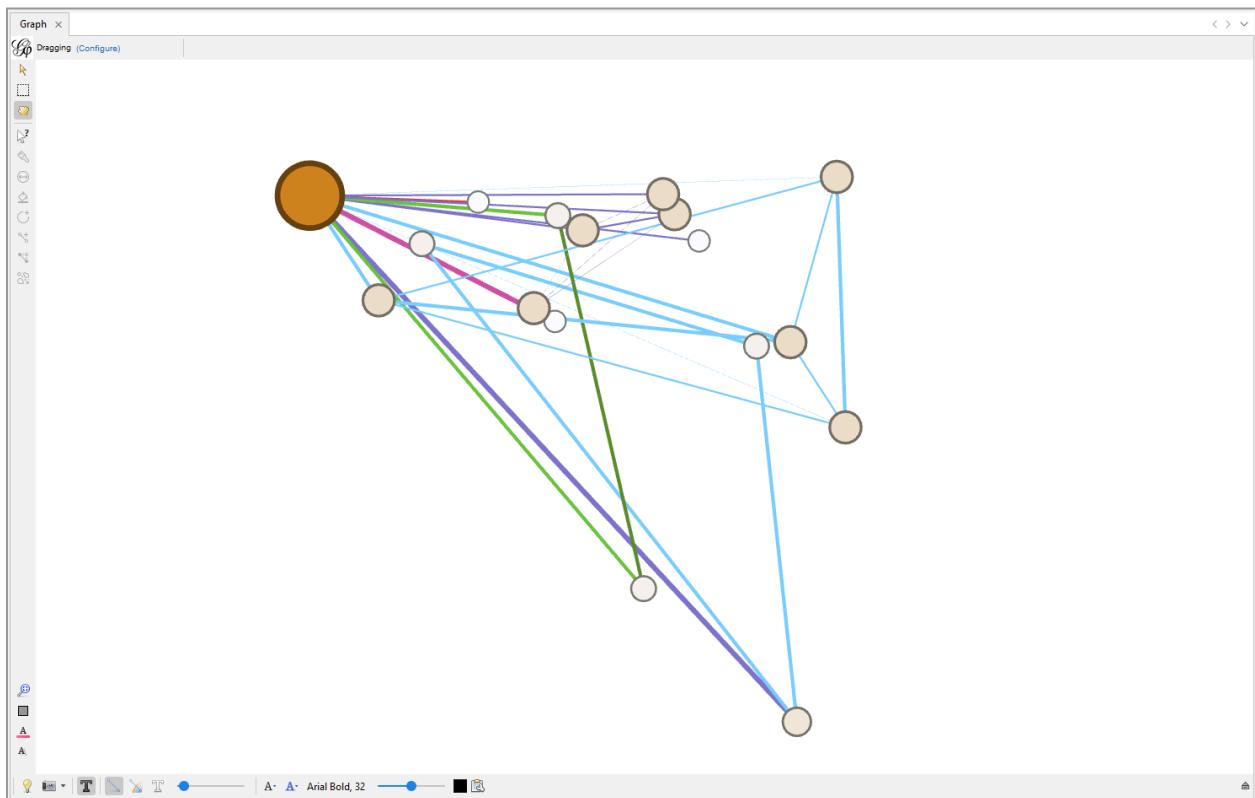
21. Go back to the **Overview** tab. In the **Appearance** window, select **Edges > Partition**, then choose **Label_copy** as the attribute from the dropdown menu.



22. Confirm the color setting, then click **Apply**.

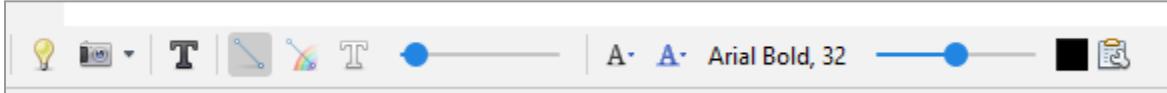


23. The graph now displays edges in different colors representing different relationship categories. Thicker lines indicate higher weights.

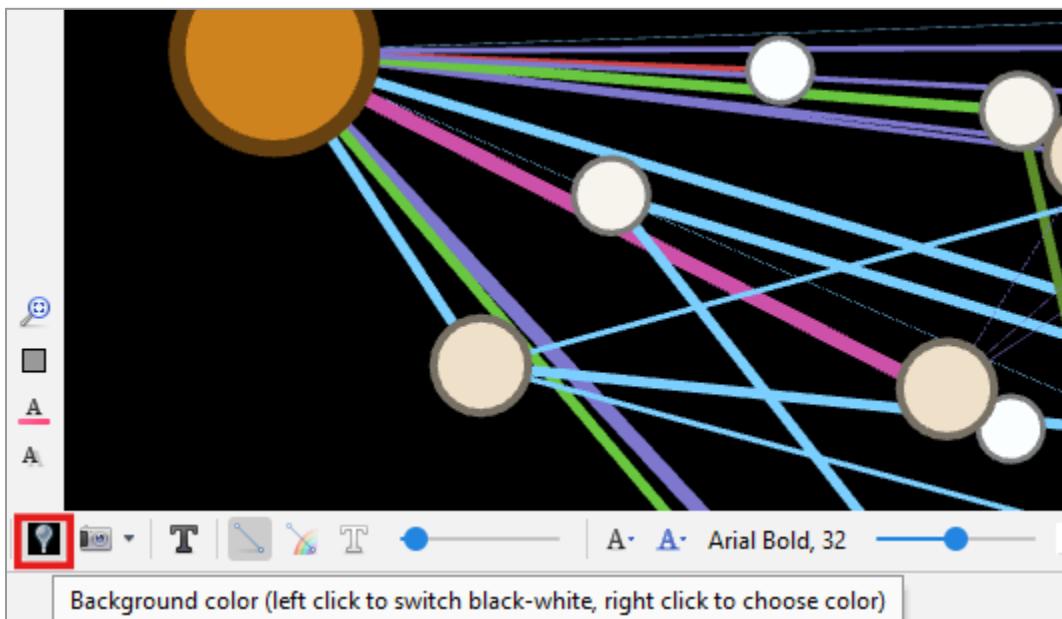


Bottom Toolbar

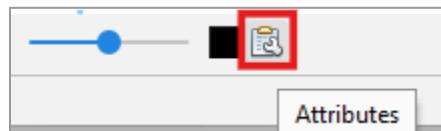
1. The bottom toolbar provides annotating tools for adding or deleting elements, such as the label and edges to the graph. We will only cover some of the tools here, but it is encouraging to explore and try them on your own after the tutorial.



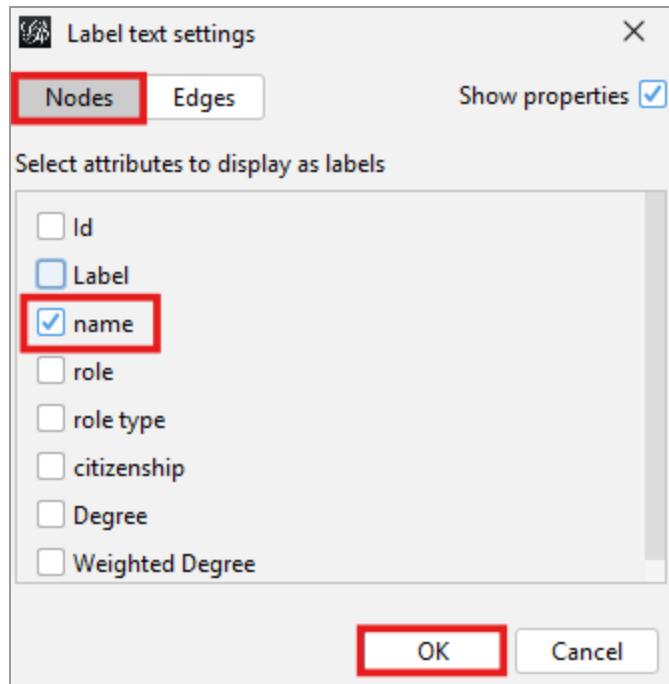
2. First, the **Background color** icon (the lightbulb symbol) allows you to switch the background from bright white to dark black.



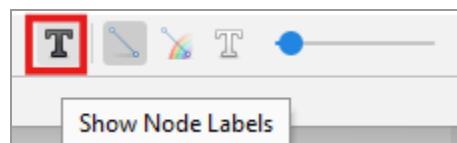
3. Click on the **Attributes** icon in the toolbar.



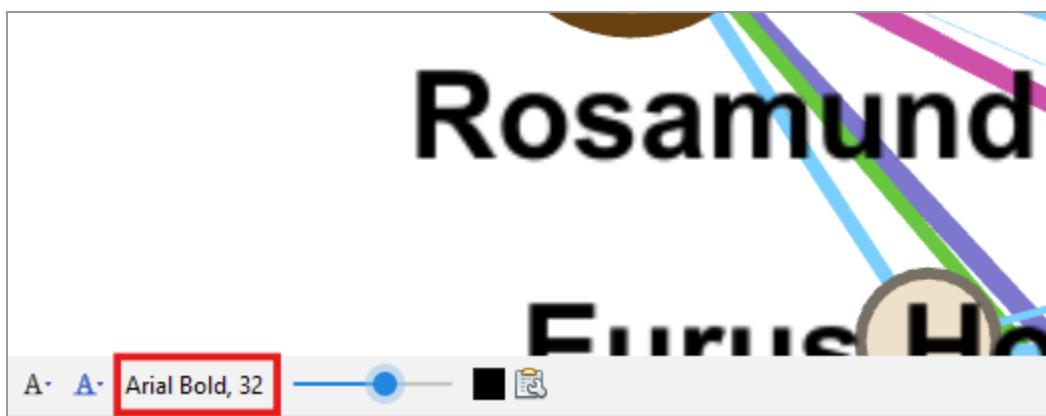
4. For the nodes, select **name** as the attributes, then click **OK**.



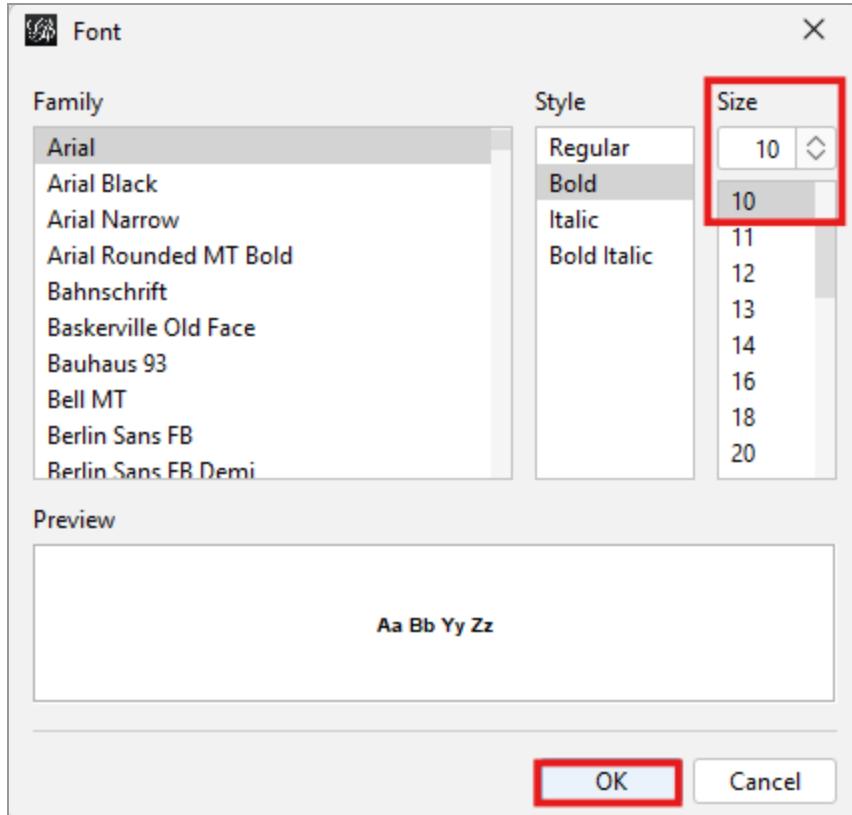
5. Next, click on the **Show Node Labels** icon.



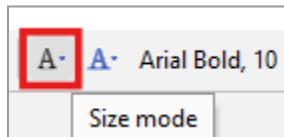
6. You will notice that the labels appear too large in the graph, but we can adjust them. To do so, click on the font settings (**Arial Bold, 32**) or drag the scale bar next to it to change the label size.



7. If you click on the font settings (**Arial Bold, 32**), you can change the font, style, and size of the labels there. Let's set the size to **10**, then click **OK**.

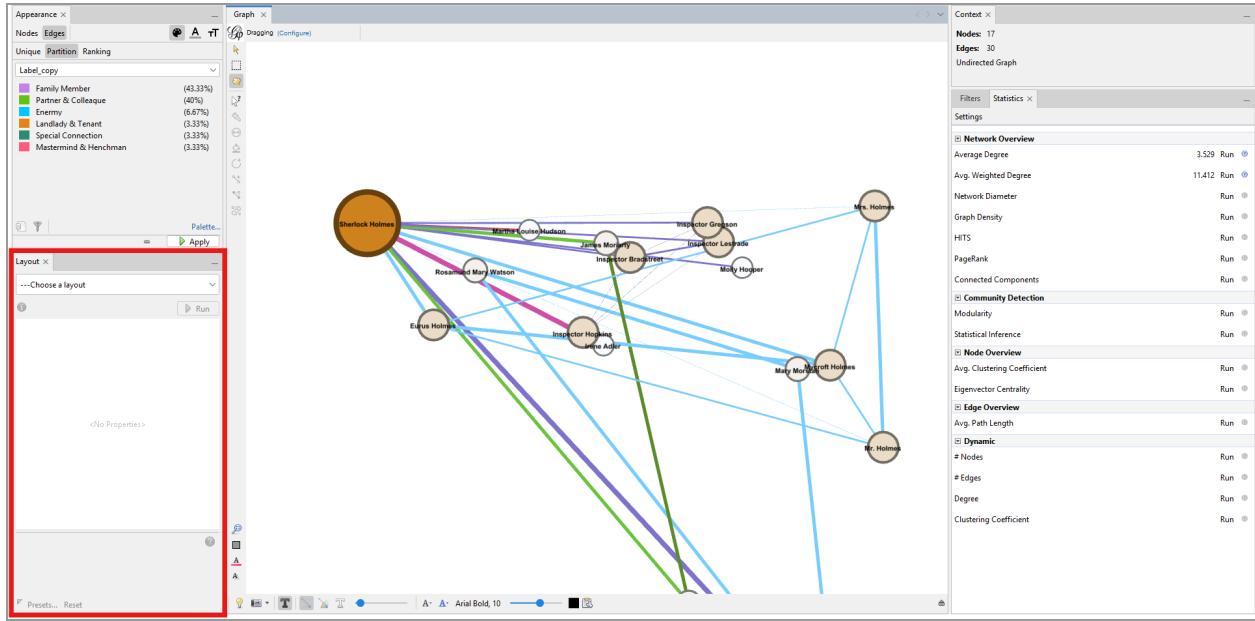


8. You can also use the **Size mode** or **Color mode** icons to control how the label size and color adjust based on the placement and size of the nodes.

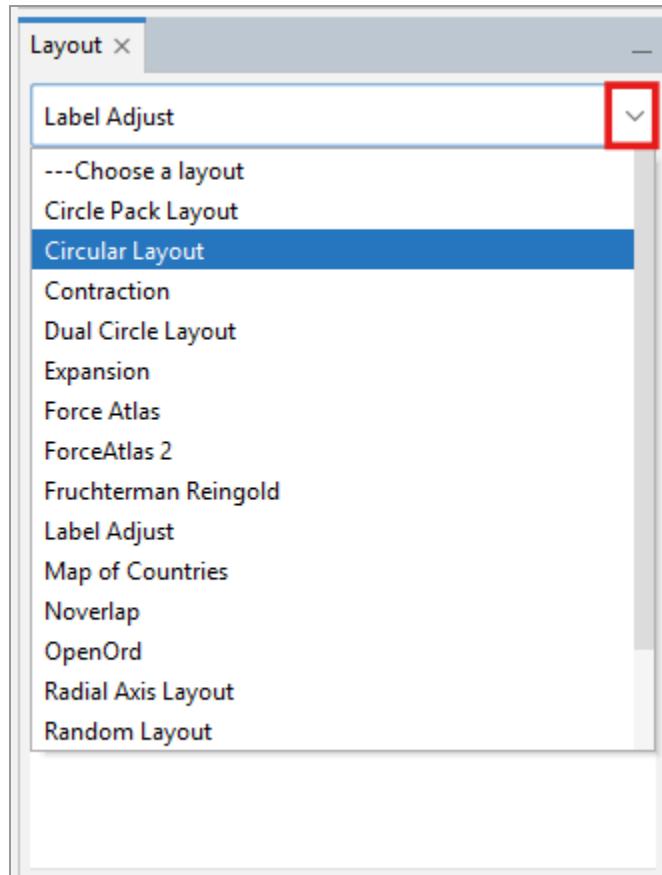


Layout

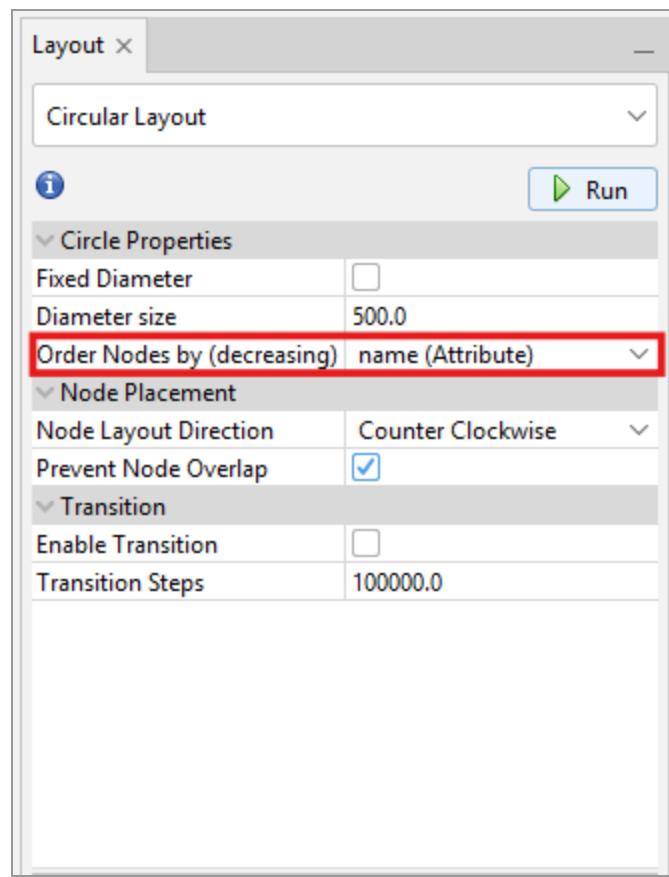
1. As default, the **Layout** window appears in the bottom left of the workplace. We will only cover some of the tools here, but it is encouraging to explore and try them on your own after the tutorial.



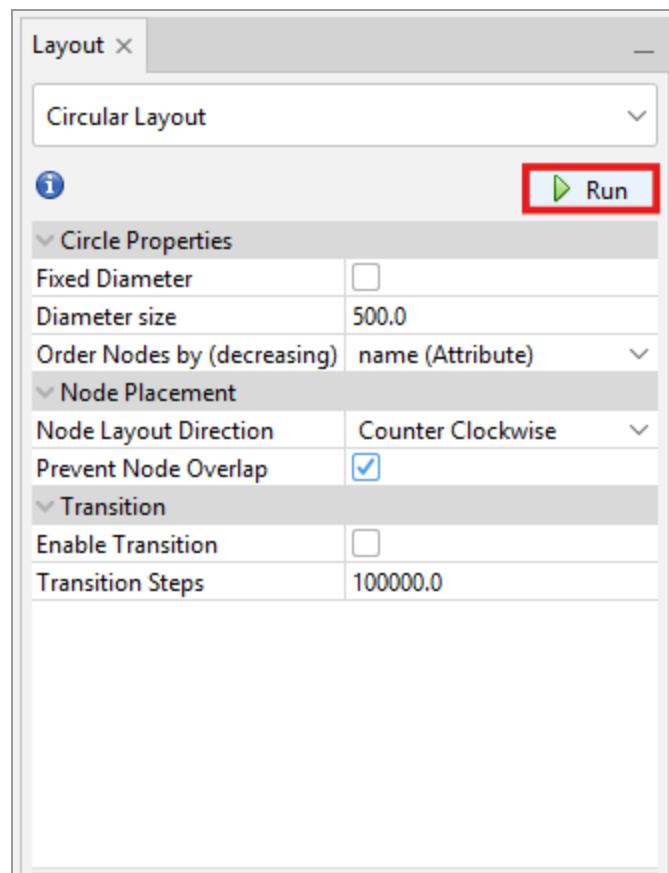
2. Click on the arrow to expand the menu, select **Circular Layout**.



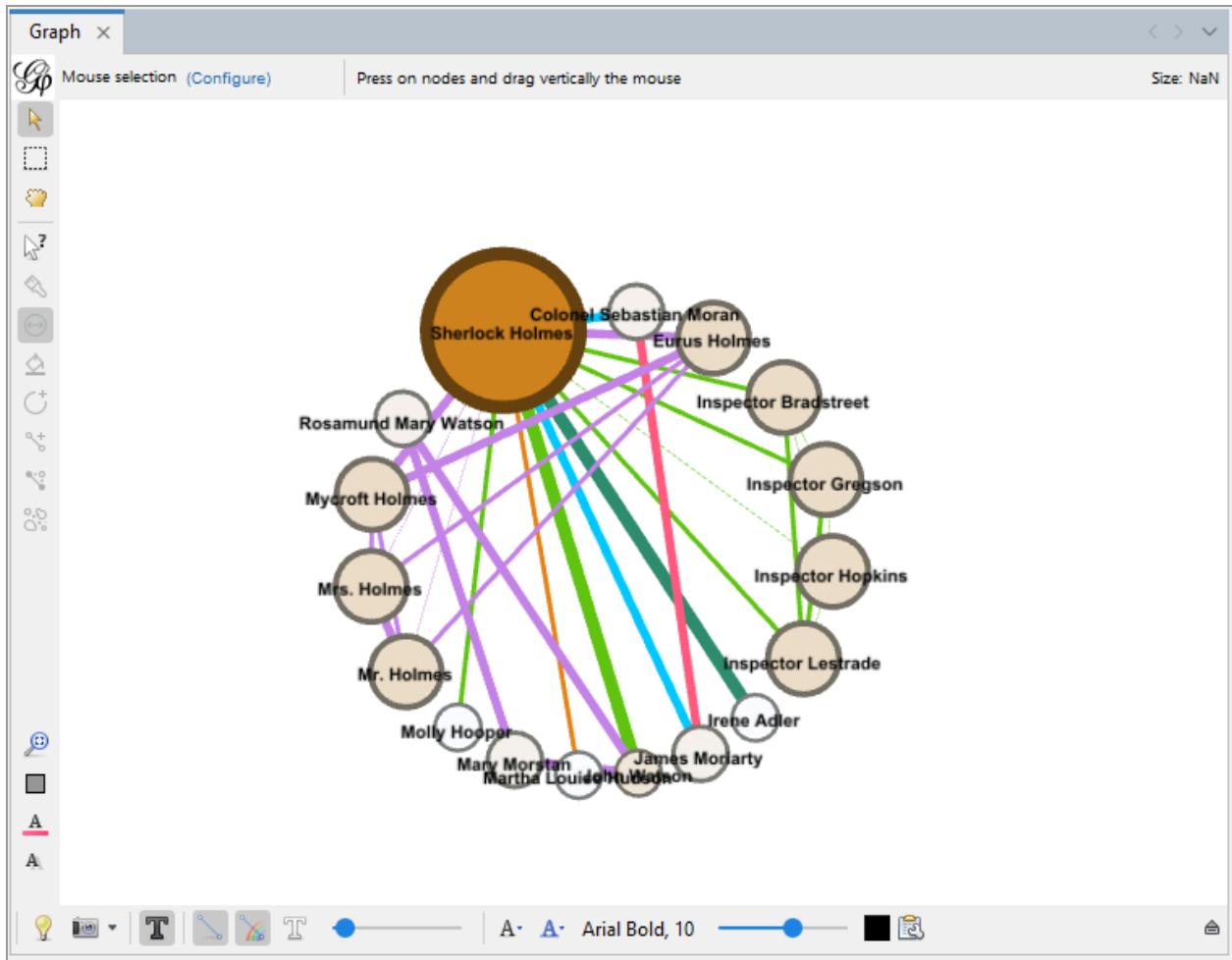
3. Let's say we want the nodes arranged alphabetically by the characters' names, we can set **Order Nodes by (decreasing)** to the **name (Attribute)**.



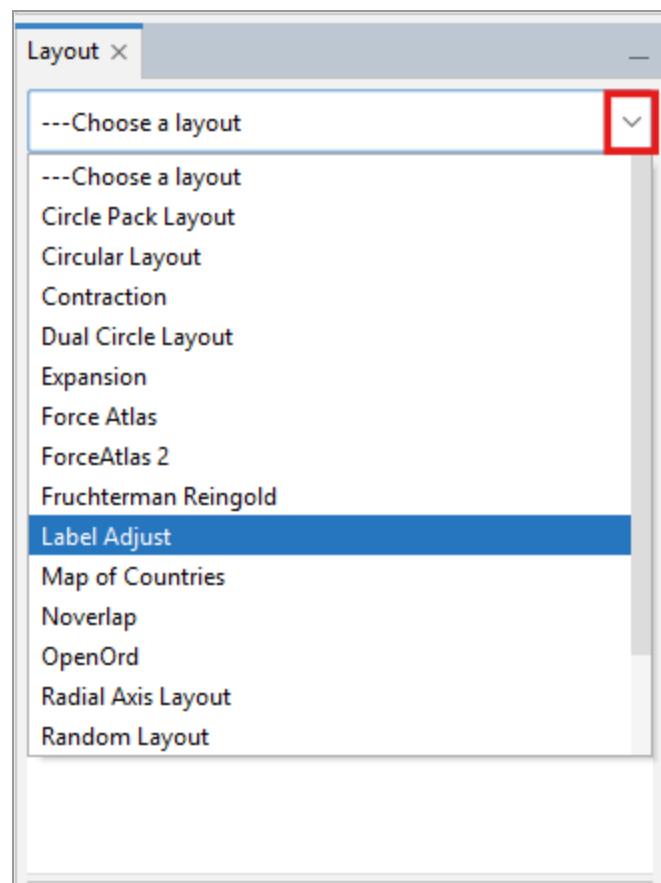
4. Click Run.



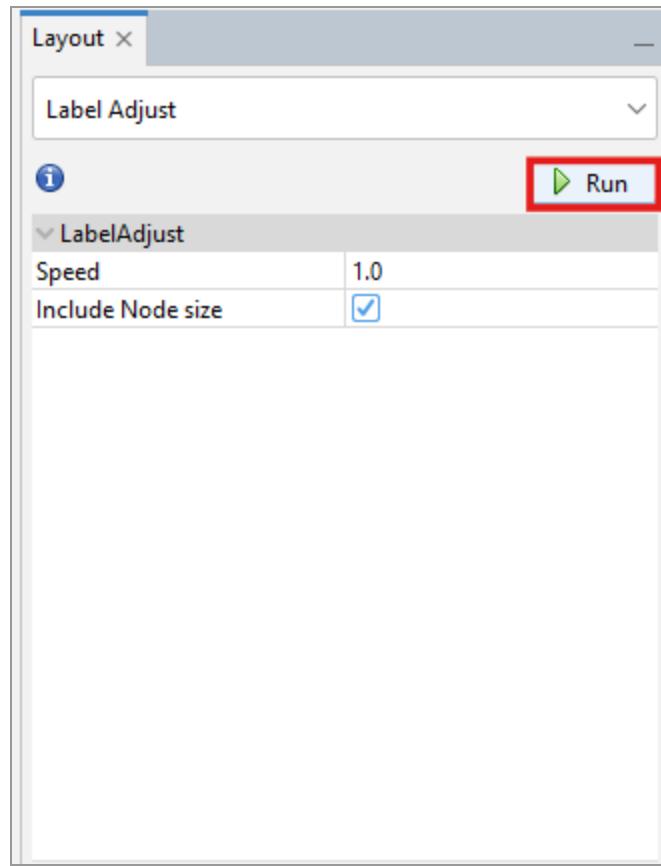
5. Now the nodes should be arranged in a circle, with the edges displayed within the circular arrangement.



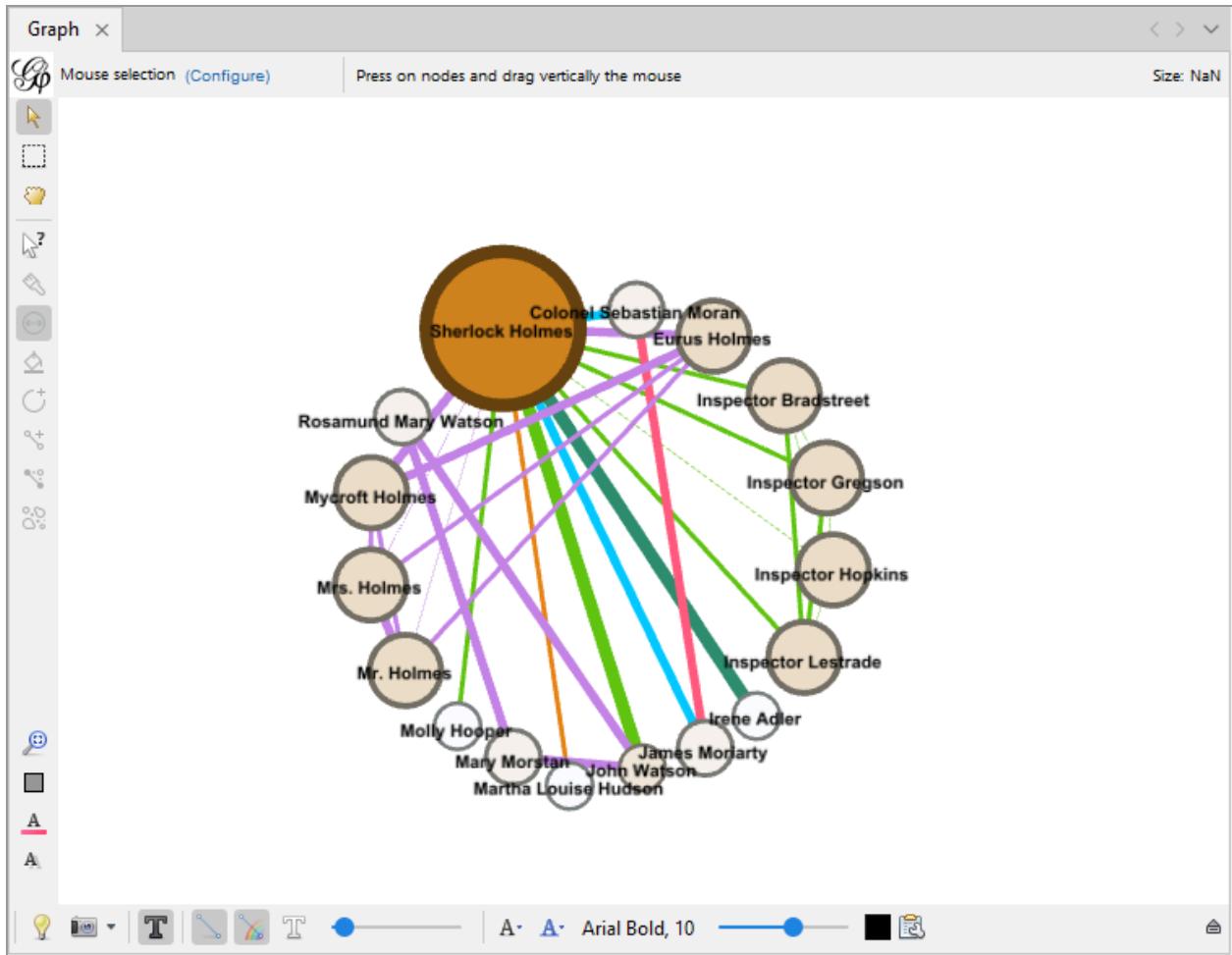
6. However, some labels may now overlap. To refine this, we can try another layout tool. Go to the **Layout** window, click the arrow to expand the menu, and select **Label Adjust**.



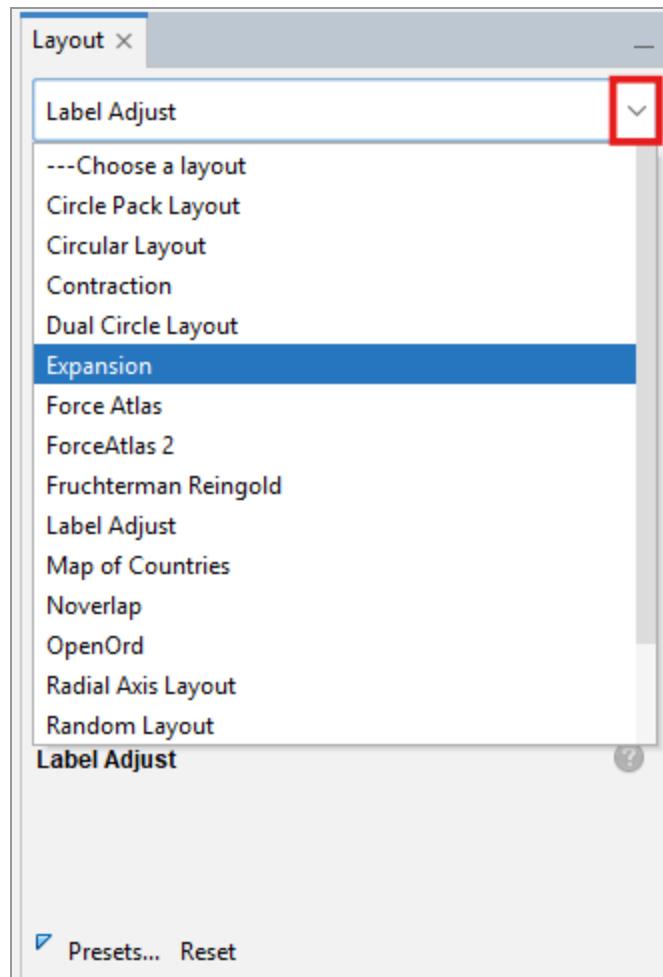
7. Click Run.



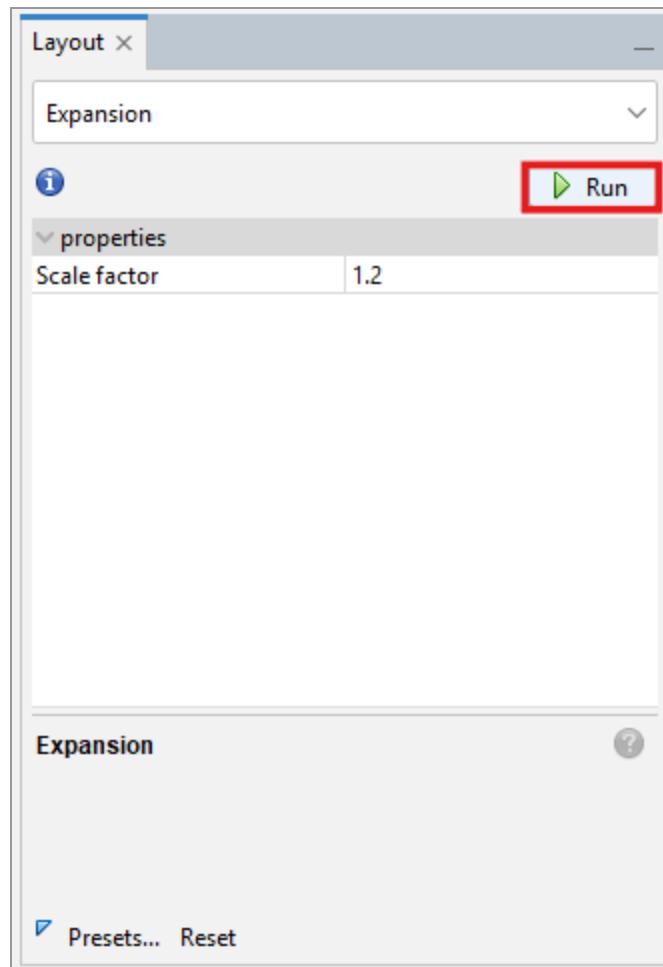
8. The labels will now adjust to clearer positions, creating better spacing and improving readability.



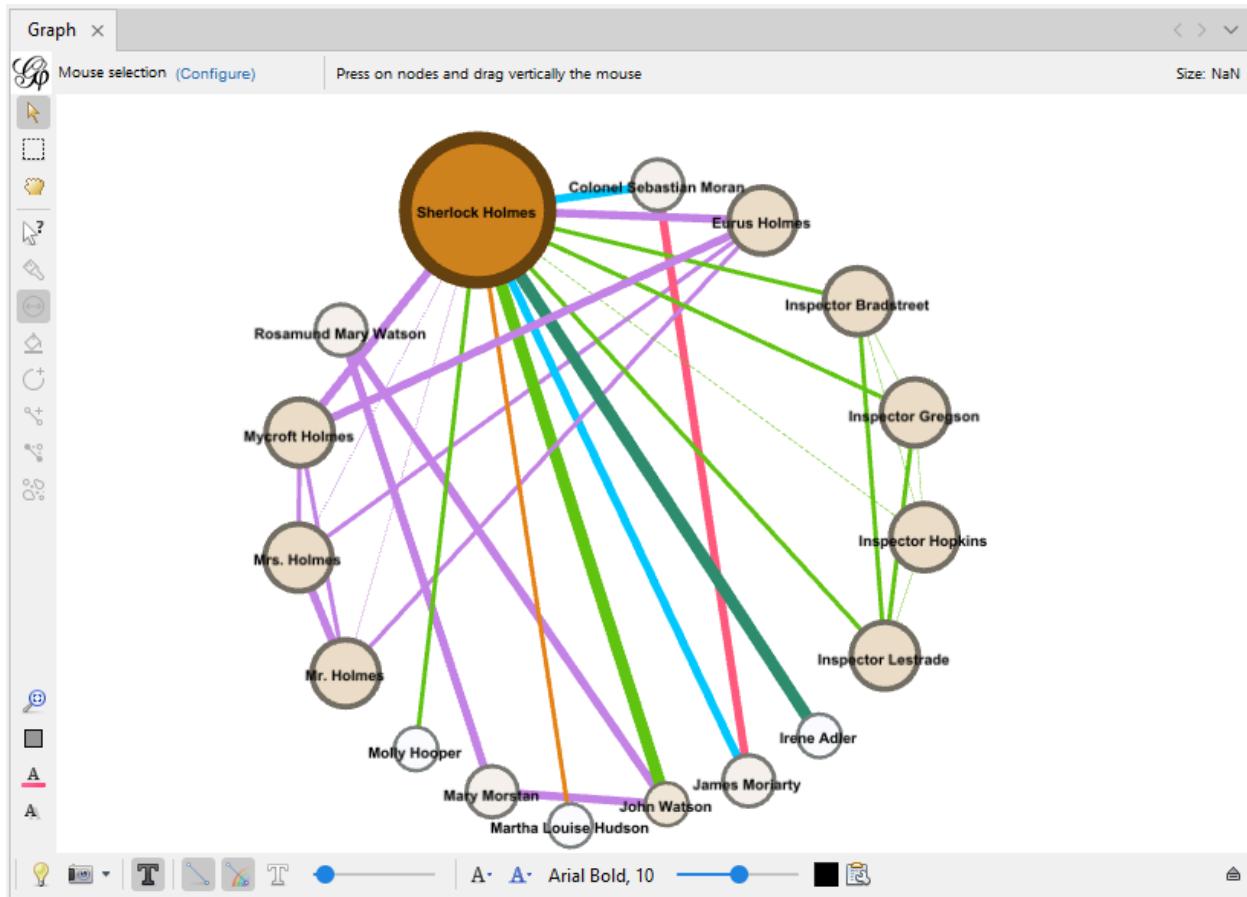
9. Now the graph looks a bit squeezed. Let's say we want to expand it, so go back to the **Layout** window, click the arrow to expand the menu, then select **Expansion**.



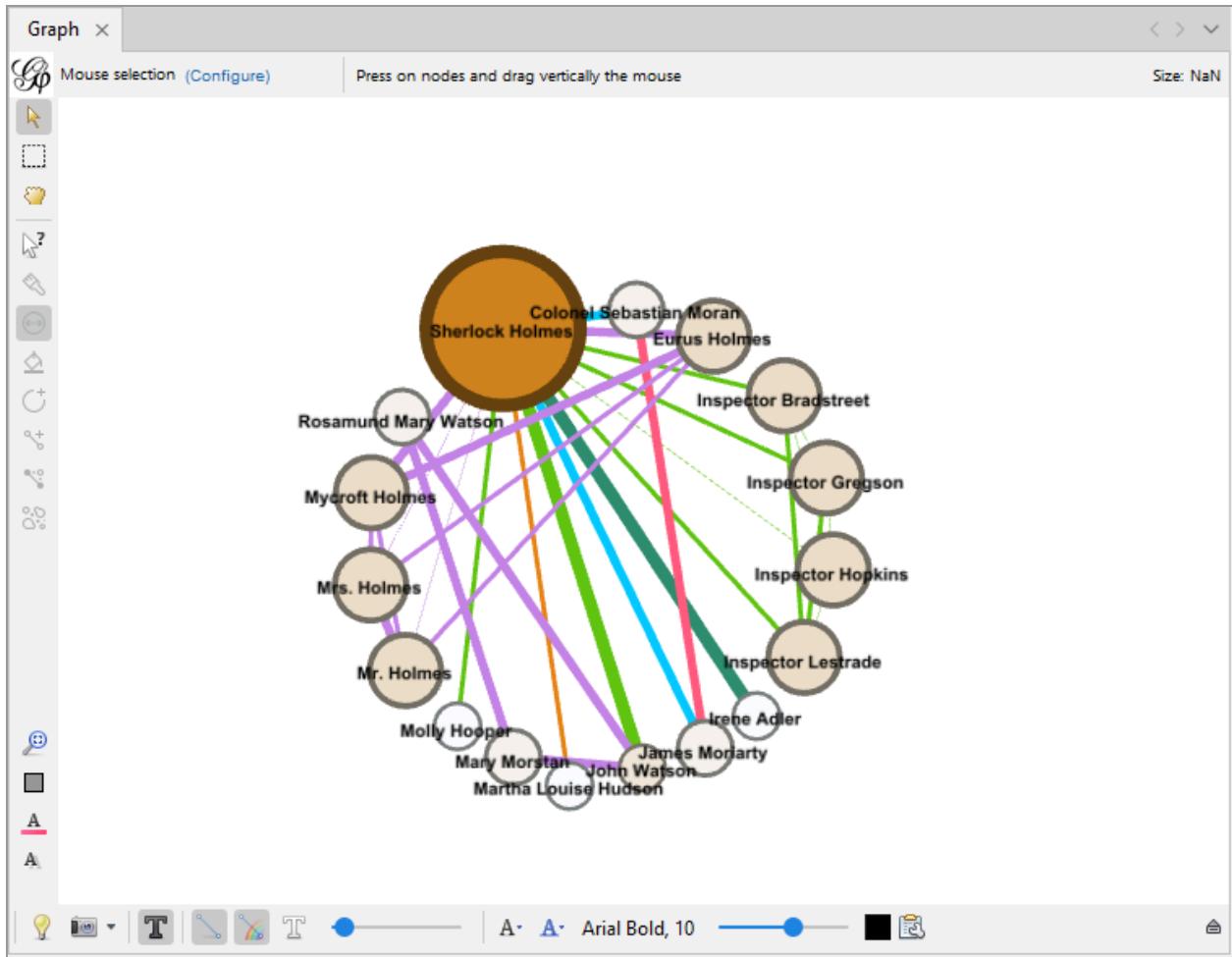
10. Click **Run** once (or multiple times) to expand the layout and reduce the squeezed appearance.



11. The graph now has a better spatial arrangement and is no longer squeezed.

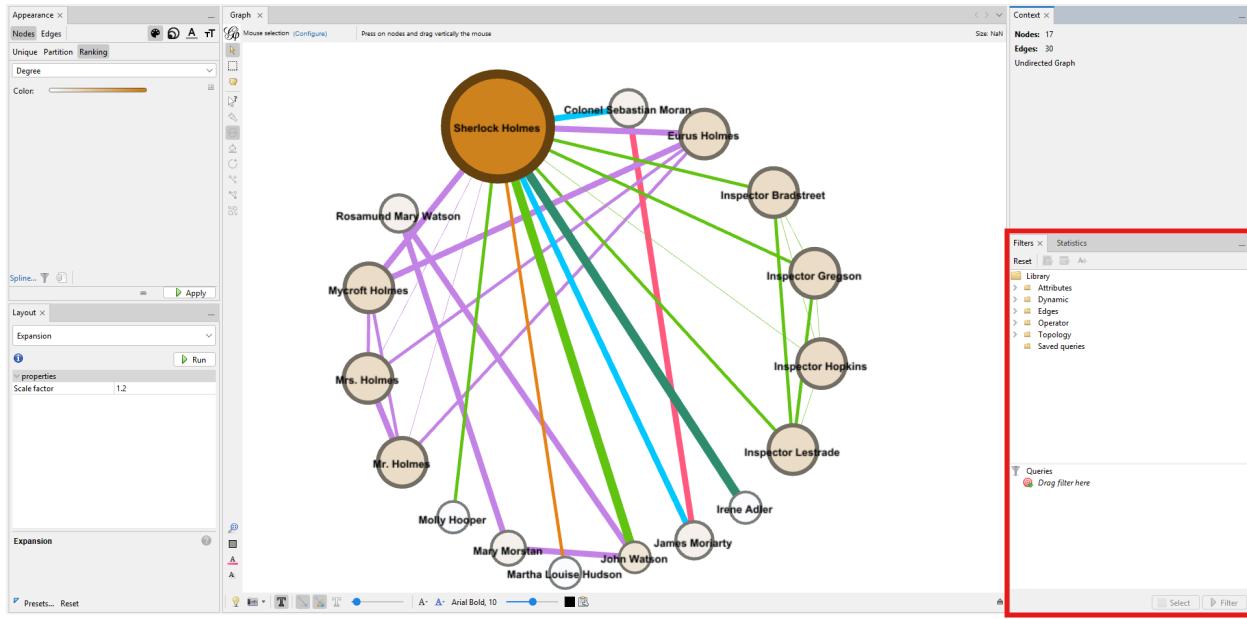


12.



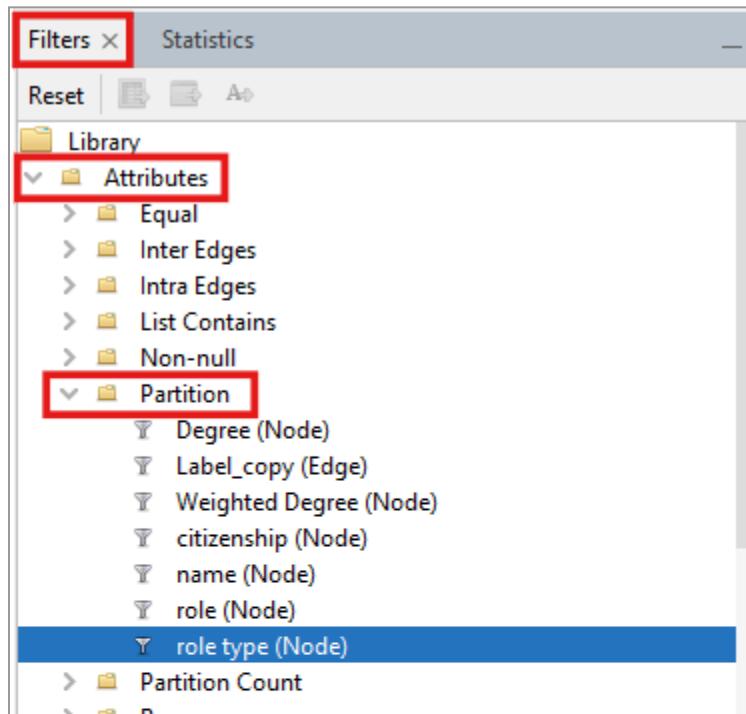
Filters & Statistics

Gephi provides filtering and statistical tools that allow you to narrow down the data for closer examination and perform statistical calculations on the imported dataset. The **Filters & Statistics** window is located at the bottom right of the workspace.

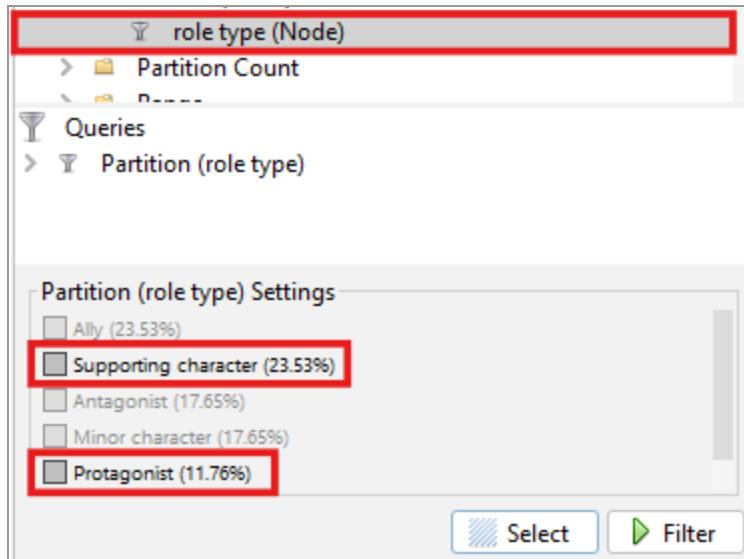


Filters

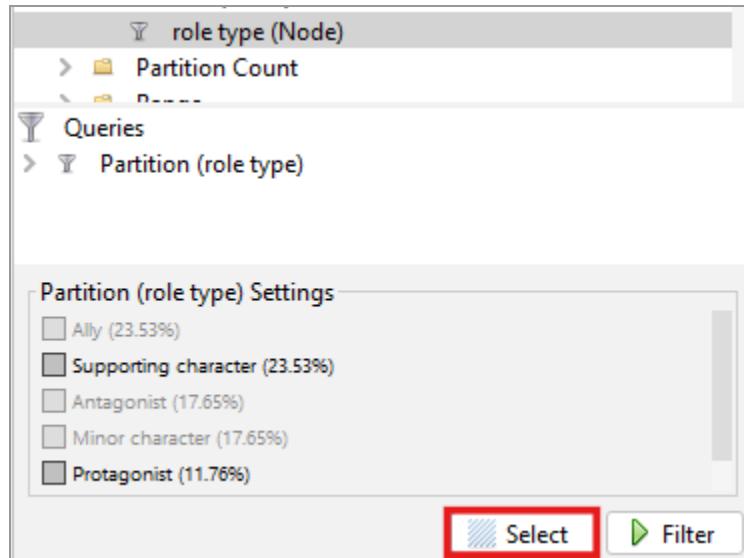
- Let's say we want to look only at the network between protagonists and supporting characters instead of all role types. We can use the filtering tools to build a query to isolate and visualize just these connections. Go to **Filters > Attributes > Partition > role type (Node)**.



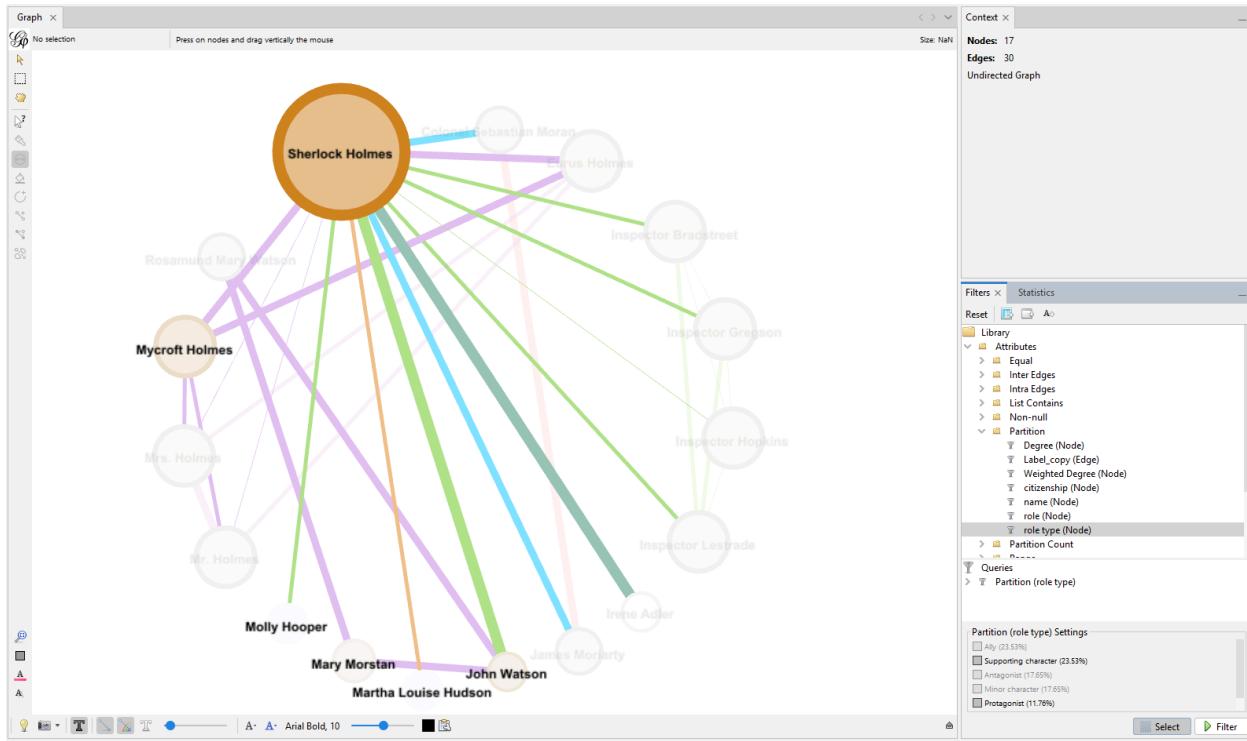
- Double-click on **role type (Node)**, then select **Supporting character (23.53%)** and **Protagonist (11.76%)**.



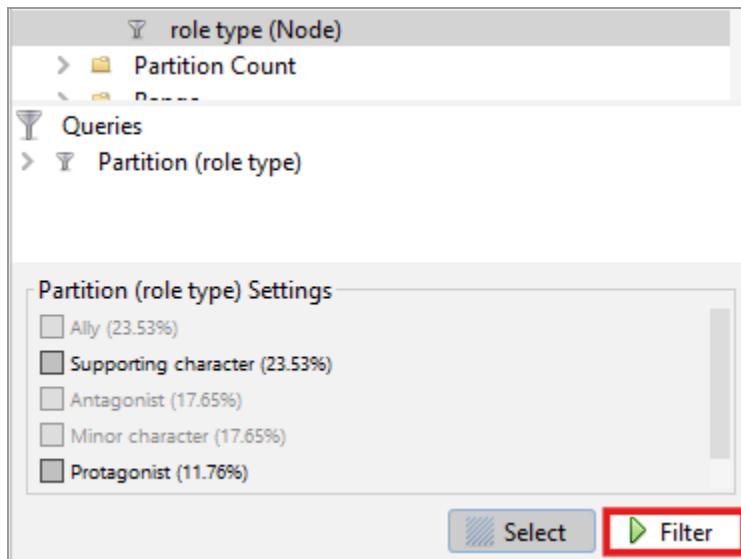
3. Next, click on the **Select** button.



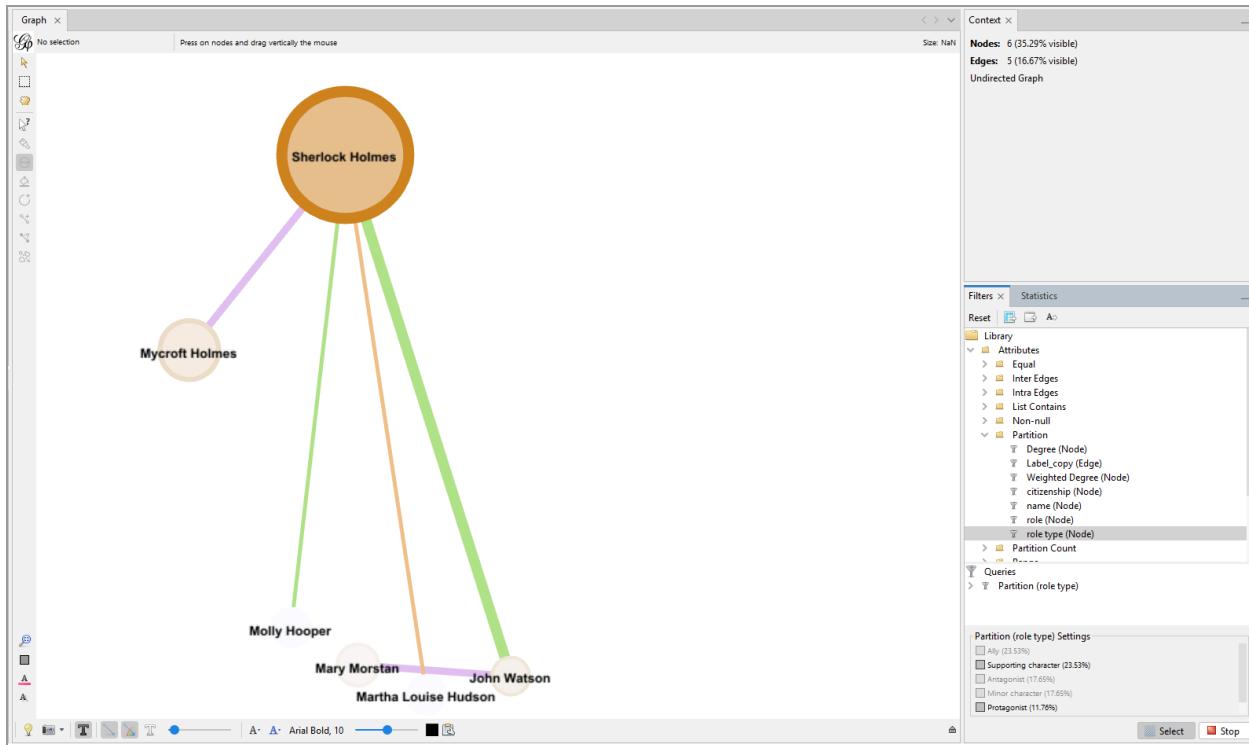
4. The selected roles will be highlighted, while the unselected ones will appear faded in the background.



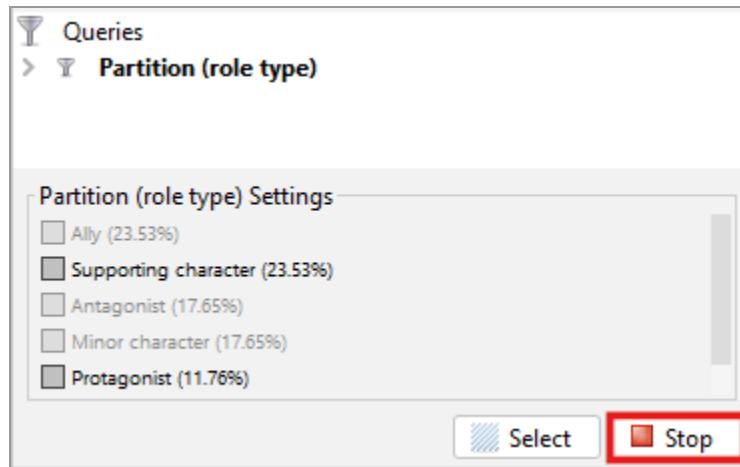
5. Next, let's try out the **Filter** button.



6. Only the selected roles will now remain visible in the workspace.



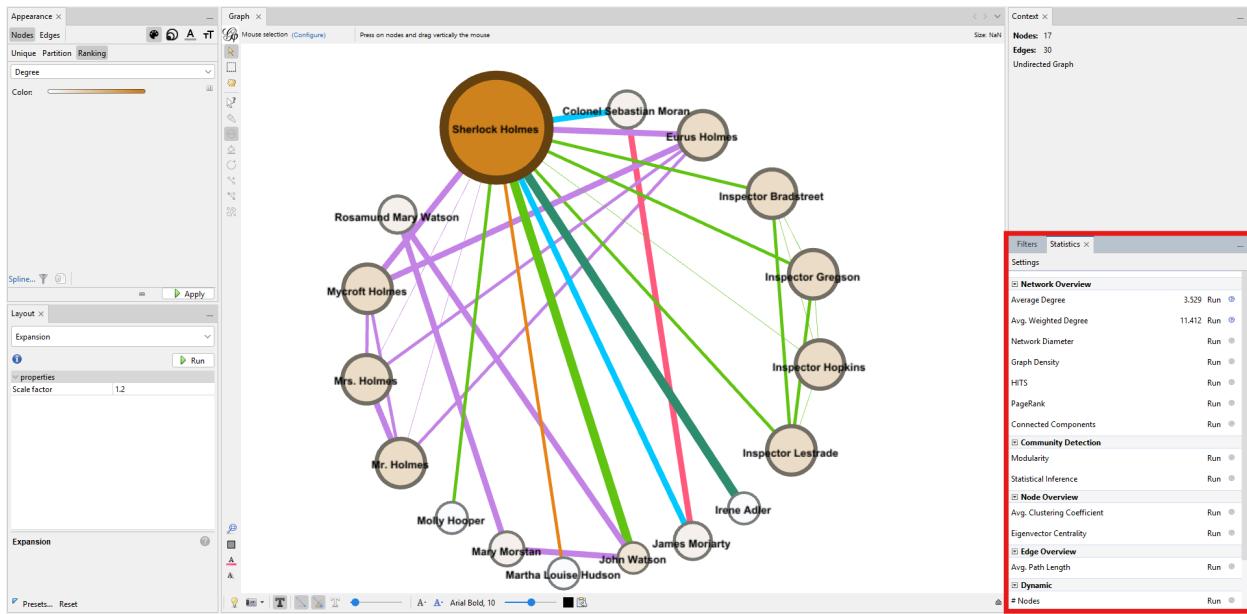
7. Click **Stop** to return to the original, unfiltered graph.



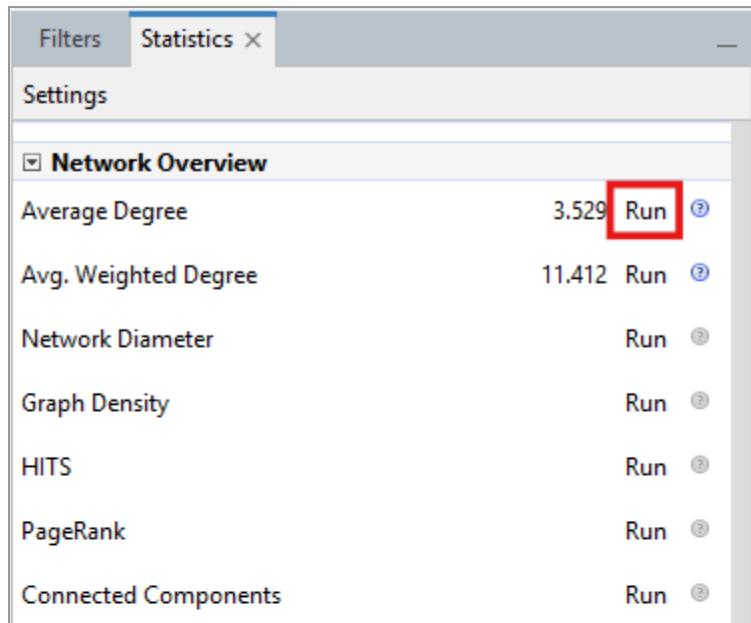
Statistics

Gephi includes a range of built-in statistical tools that help you analyze the structure and characteristics of your network. The results can be visualized directly within the interface or exported for further analysis. We won't cover all tools here, but it is encouraging to explore and try them on your own after the tutorial.

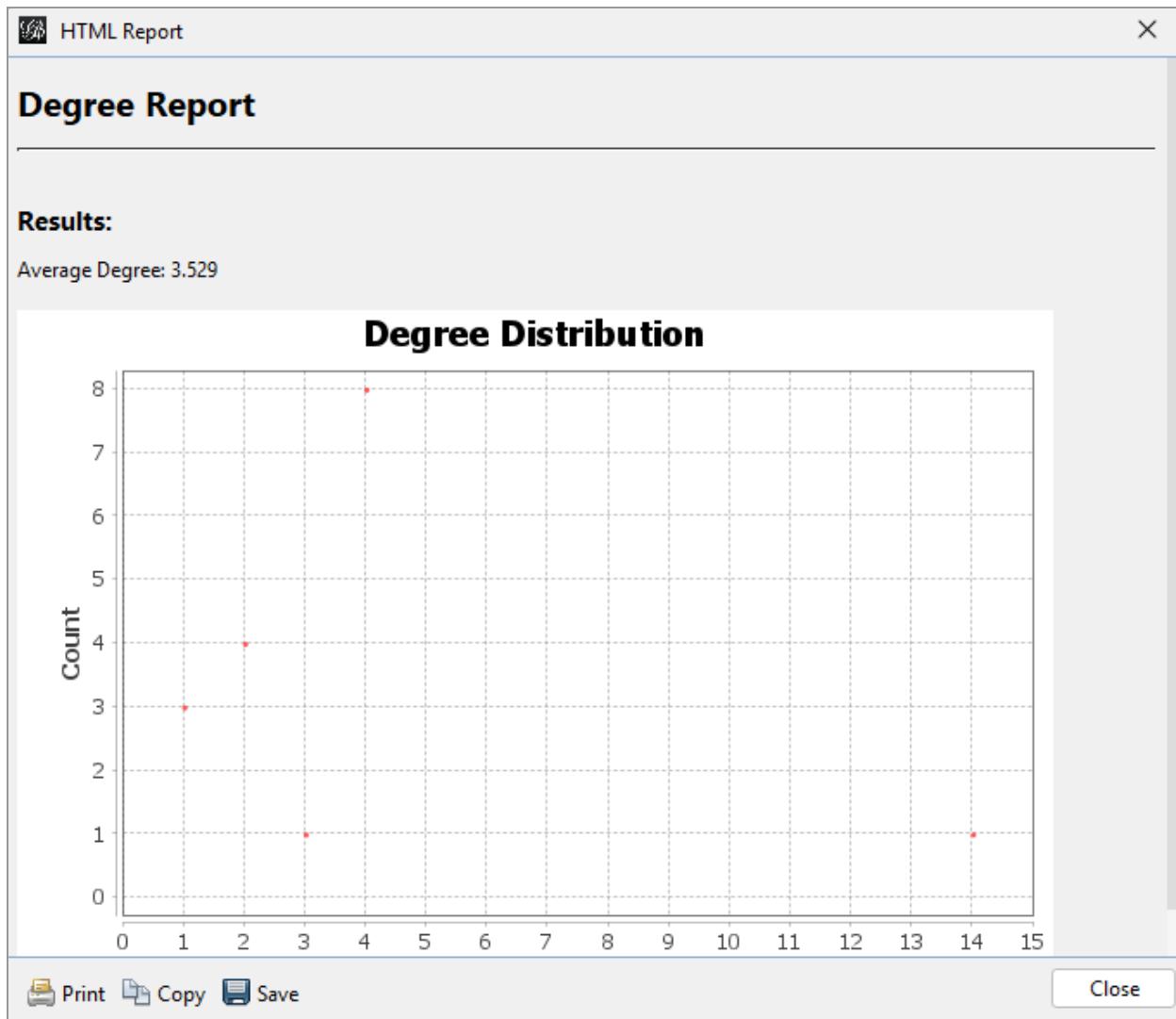
1. Let's say we want to analyze the network's average degree. First, navigate to the **Statistics** tab in the bottom-right **Filters & Statistics** window.



2. Then, click Run next to Average Degree.



3. A pop-up window will appear displaying a line graph of the degree distribution.



4. You can print, copy, or save the results from the pop-up window. Click **Close** when you're finished.



Export the Graph

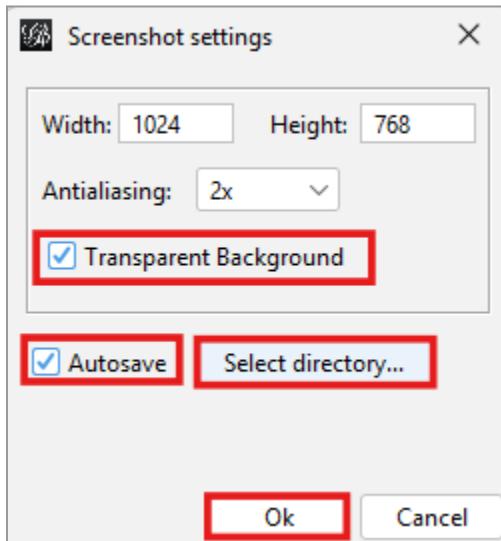
There are multiple ways to export a project from Gephi. You can export the graph as static image files (e.g., PNG, SVG, PDF) for use in reports or presentations, or generate an interactive dashboard that allows users to explore the network online.

Capture the current view

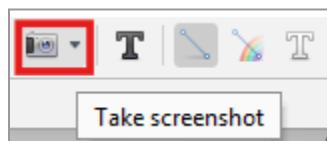
1. Click on the Take screenshot icon (the camera symbol), then go **Configure**.



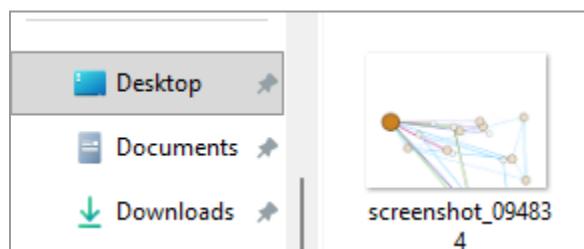
2. Configure the saving settings, such as adjusting the export size, enabling the **Transparent Background** option, and selecting a directory for **Autosave**. Once complete, click **Ok**.



3. Go back to the toolbar and click on the **Take screenshot** icon (the camera symbol) again.

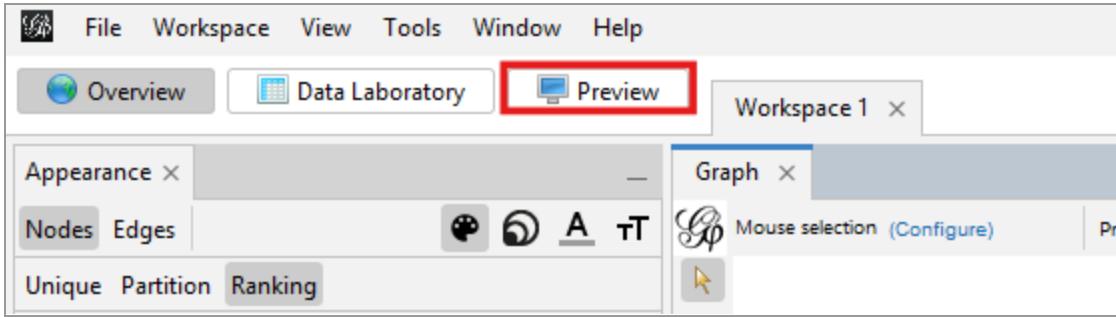


4. Go to the directory where saved as the saving path, you will find the image exported

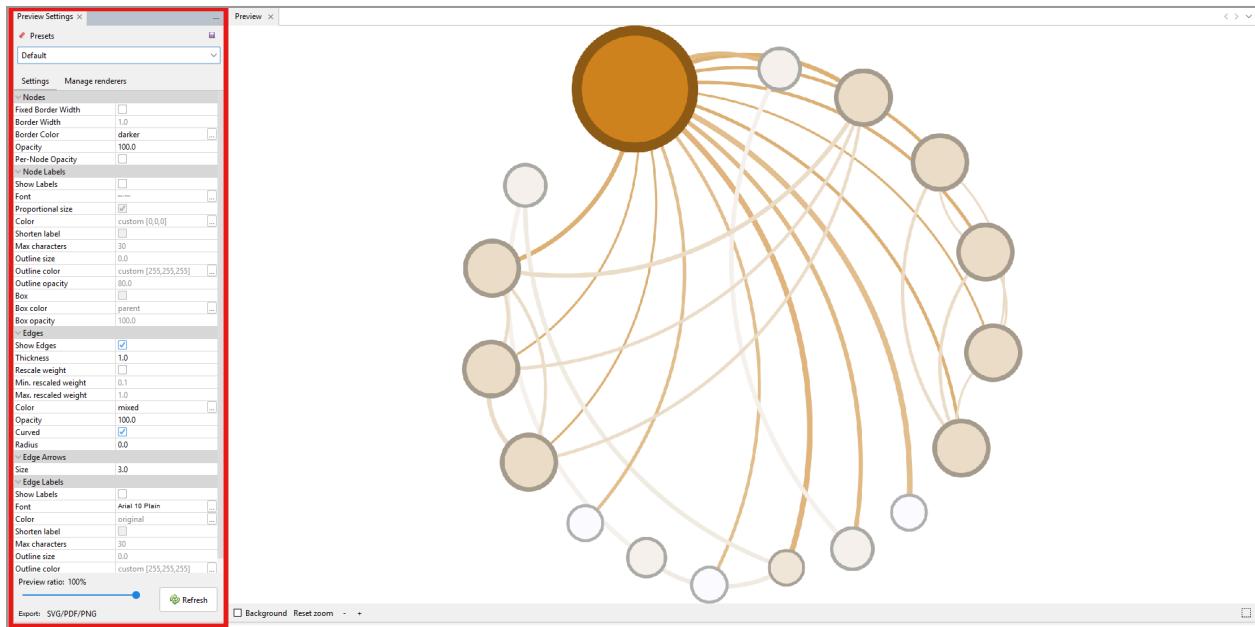


Save the graph

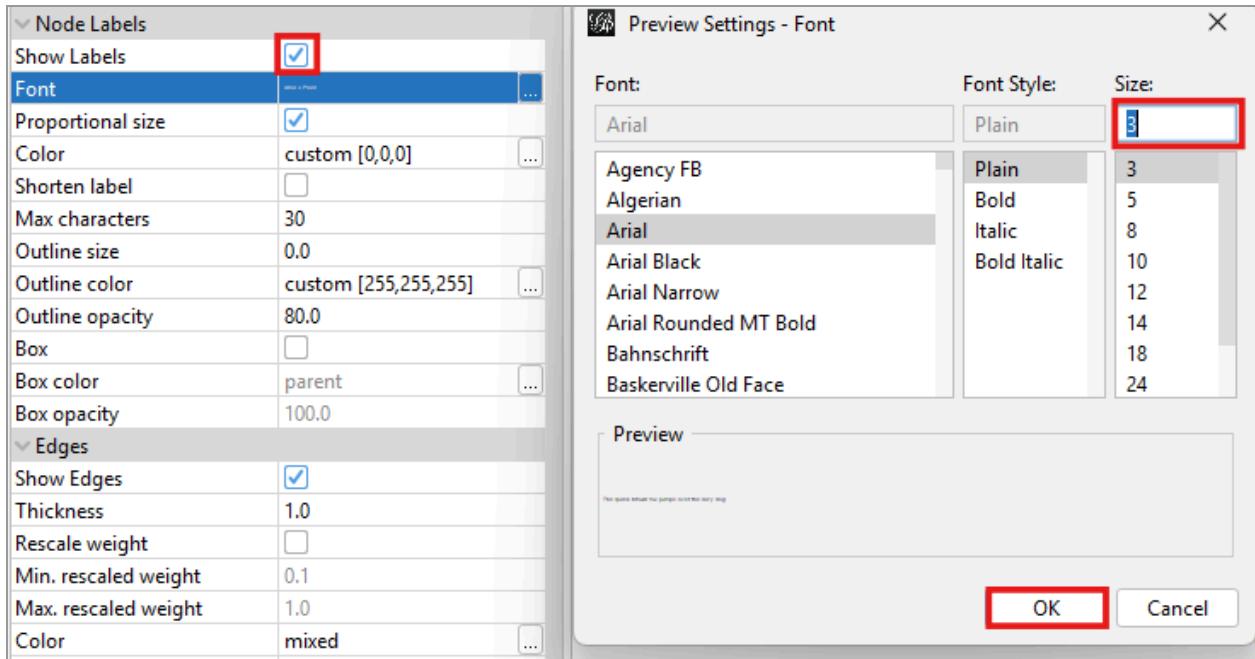
1. Go to **Preview** view first.



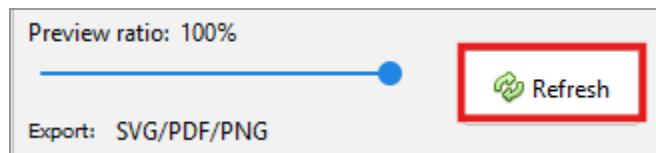
- Now you will see the graph has no label or colored edges we previously set. This is because some settings have not been applied, and we can configure those using the **Preview Settings** pane.



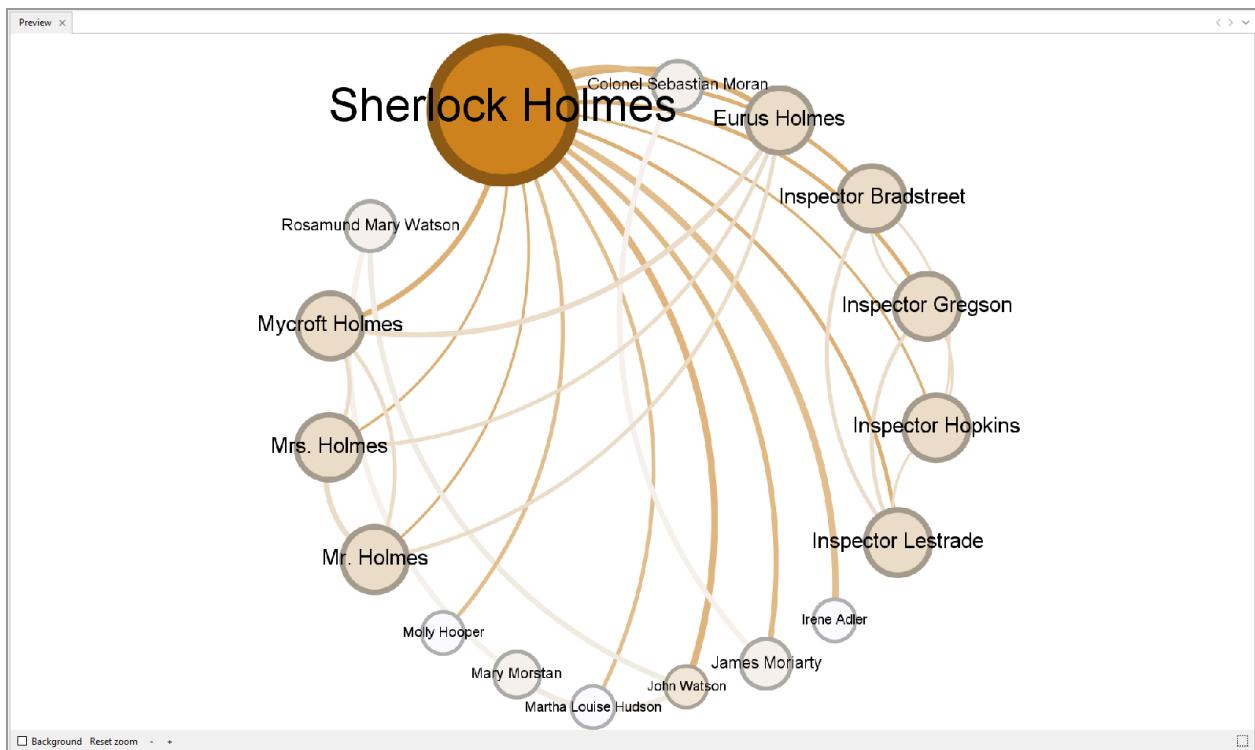
- For example, let's check **Show Labels**, and change the **Font** size to 3 and click **OK**.



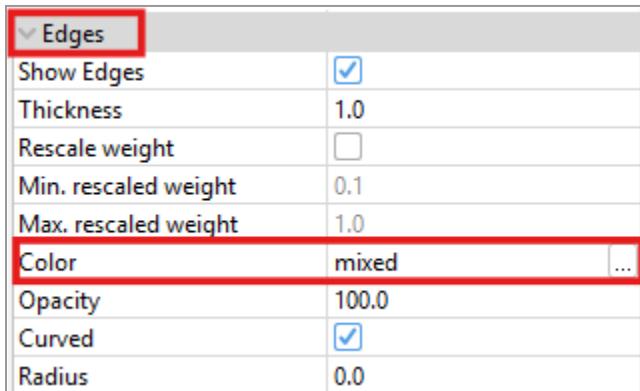
4. Click the Refresh button at the bottom.



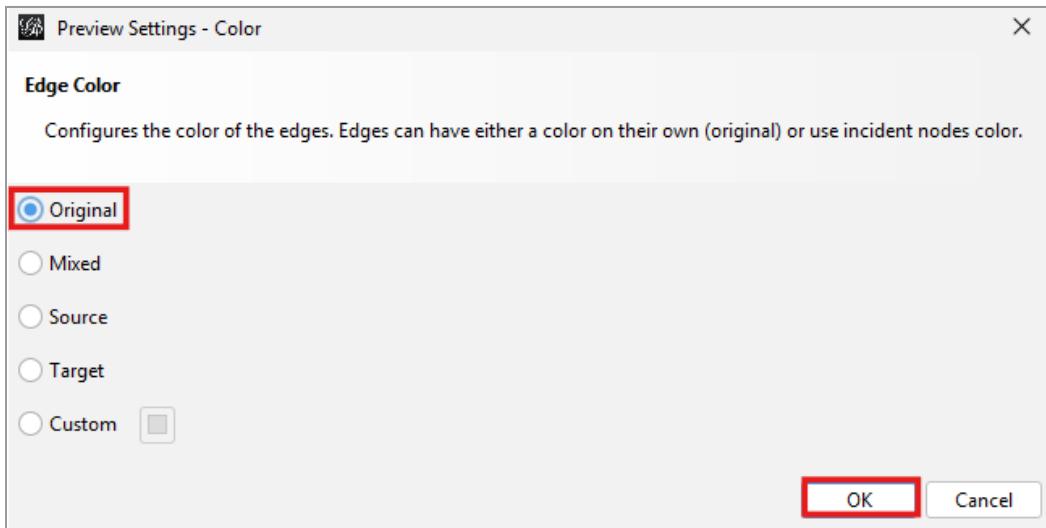
5. The graph is now updated with the characters' names displayed as labels.



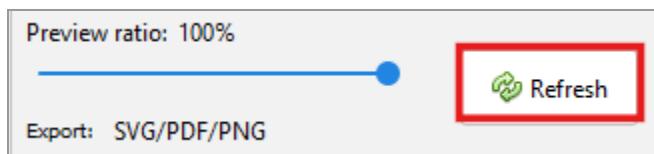
6. Next, we will change the color of the edges. Under **Edges**, in the **Color** field, click the ellipsis (3-dot icon) next to **mixed**.



7. In the pop-up, check **Original**, then click **OK**.



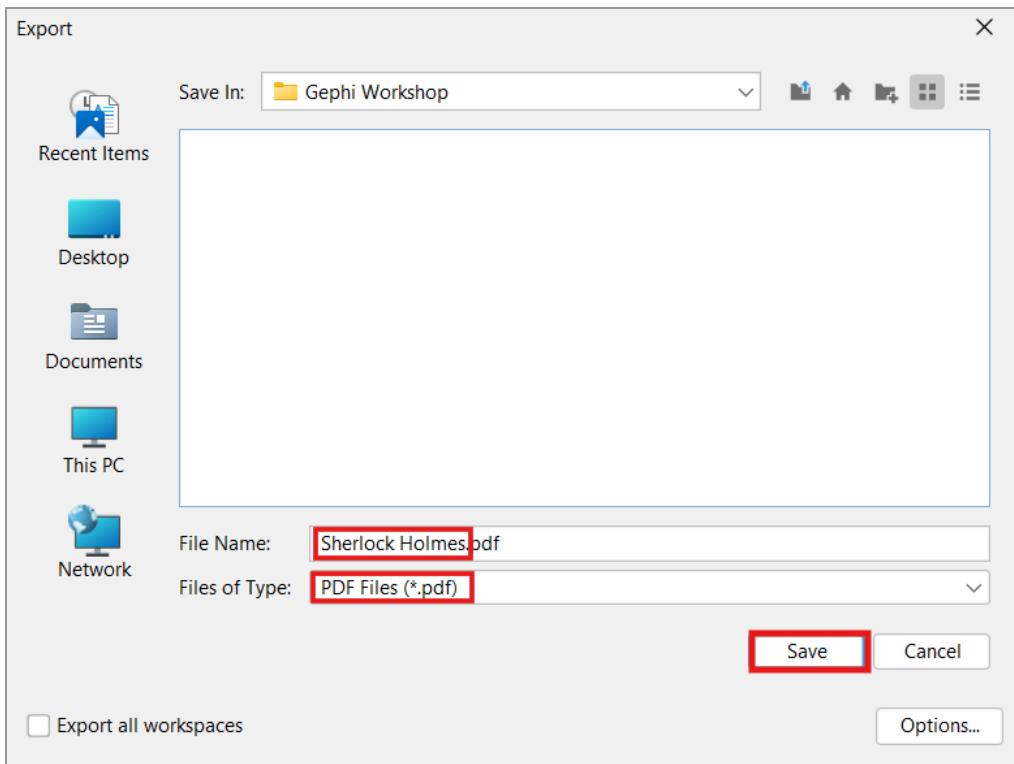
8. Click the **Refresh** button at the bottom again.



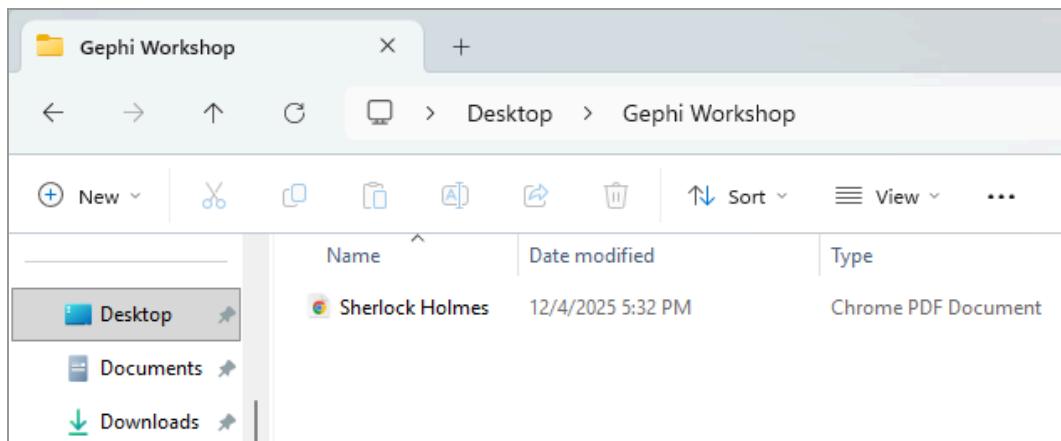
9. Now we have essentially recreated the visualization from the **Overview** tab, even though it may not be an exact match. The preview graph is ready to be exported. To export it, click the **SVG/PDF/PNG** button next to **Export** at the bottom of the **Preview Settings** pane.



10. In the pop-up window, choose the save location, enter a file name, and select the desired format. Then click **Save**.

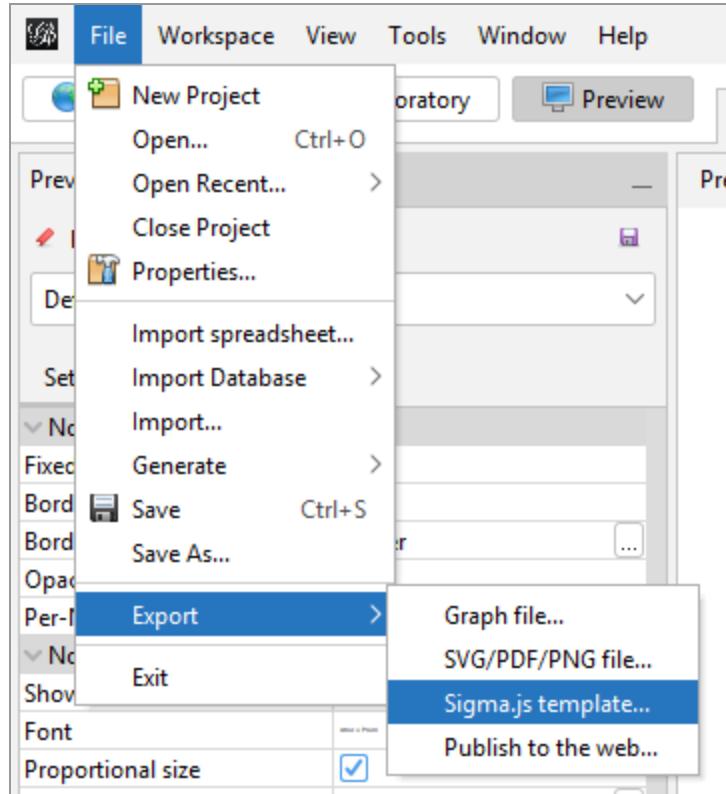


11. The file should now be saved in the designated location.



Export as a website

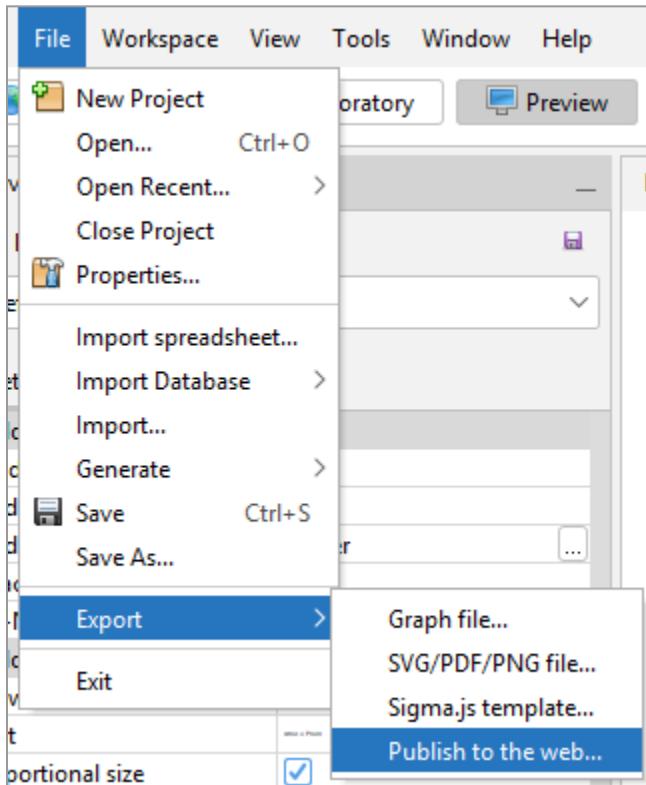
1. In the workspace, go to **File > Export > Sigma.js.template**.



2. In the pop-up window, set the saving path and fill in the details you want to include about the nodes, edges, colors, author, and the graph title. Next, complete the short and long descriptions, and then click **OK**.
3. Now, go to the folder where we just saved the file - this will be a folder called **network**.
4. Open network folder, you should find a HTML file titled **index**. Double-click to open it in your browser.
5. Unfortunately, if you try to open the newly exported index.html file directly, it will not display correctly. This happens because modern web browsers block local HTML files from running JavaScript. There are several ways to work around this, but in this tutorial we will introduce two options. Let's begin with the first one.
6. Install a free software package called [Visual Studio Code](#).
7. Start it up on your computer.
8. In the left toolbar, go to **Extension**, and search for **Live Server** and install it.
9. Next, go to **File > Open Folder** and select the **network** folder that contains the **index.html** file, click **Open** to continue.
10. Once the files are imported, right-click **index.html** in the folder pane and select **Open with Live Server**.
11. This should open up the file in your default browser, and you should be able to view it and interact with it.
12. Note that this link cannot be shared with others because it runs locally on your machine and is not hosted on the internet. To create a shareable version, consider using the **Publish to GitHub** option in the next section.

Publish to GitHub

1. In the workspace, go to **File > Export > Publish to the web.**



2. First, complete the steps under the **Set up (to do just once)** tab.
3. Next, go to the **Publish** tab and click **Publish your network to the web!**
4. Then, under **Link to the network visualization**, click **Open in web browser**. You may also copy and paste the link into any browser.
5. Your project is now hosted on the Retina platform, where you can interact with the graph, adjust visual settings, and download the dataset.
6. Most importantly, you can now share your network using the URL or use the embed code to display the graph as an asset on other websites.

Resources

Network Analysis Theory

- [*Networks, crowds, and markets : reasoning about a highly connected world*](#) by David Easley and Jon Kleinberg
- [*Visual Insights: A Practical Guide to Making Sense of Data*](#) by Katy Börner and David E. Polley
- [*Network data mining and analysis*](#) by Ming Gao, Ee-Peng Lim, David Lo
- [*Fundamentals of big data : network analysis for research and industry*](#) by Hyunjoung Lee and Il Sohn

Gephi

- [Gephi Quickstart](#)
- [Gephi Documentation](#)
- [Workshops](#) by Clément Levallois

Network Data Preparation

- [Preparing Data 1: Making an Edge List](#)
- [Preparing Data 2: Making a Node List from an Edge List](#)

Network Datasets

- [Duke Network Analysis Center](#)
- [Index of Complex Networks](#)
- [UCIrvine Network Data Repository](#)
- [Stanford Large Network Dataset Collection](#)
- [UCINET Datasets](#)
- [Network Datasets](#)