

Cours DH UniNe

# Analyse de réseau avec Gephi

Simon Gabay

Genève, Lundi 20 avril 2020

# Installation

# Java

Pour faire fonctionner Gephi, il vous faut installer Java 8 (*Java SE Runtime Environment 8* de son petit nom). Pour ce faire:

- Allez à cette adresse:

<https://www.oracle.com/technetwork/java/javase/downloads/jre8-downloads-2133155.html>

- Créez si besoin un compte *Oracle*

## Java SE Runtime Environment 8 Downloads

Do you want to run Java™ programs, or do you want to develop Java programs? If you want to run Java programs, but not develop them, download the Java Runtime Environment, or JRE™.

### Important Oracle JDK License Update

**The Oracle JDK License has changed for releases starting April 16, 2019.**

The new [Oracle Technology Network License Agreement for Oracle Java SE](#) is substantially different from prior Oracle JDK licenses. The new license permits certain uses, such as personal use and development use, at no cost -- but other uses authorized under prior Oracle JDK licenses may no longer be available. Please review the terms carefully before downloading and using this product. An FAQ is available [here](#).

Commercial license and support is available with a low cost [Java SE Subscription](#).

Oracle also provides the latest OpenJDK release under the open source [GPL License](#) at [jdk.java.net](#).

If you want to develop applications for Java, download the Java Development Kit, or JDK™. The JDK includes the JRE, so you do not have to download both separately.

[JRE 8u251 Checksum](#)

### Java SE Runtime Environment 8u251

This software is licensed under the [Oracle Technology Network License Agreement for Oracle Java SE](#)

Product / File Description	File Size	Download
Linux x86 RPM Package	68.41 MB	 <a href="#">jre-8u251-linux-i586.rpm</a>
Linux x86 Compressed Archive	84.22 MB	 <a href="#">jre-8u251-linux-i586.tar.gz</a>
Linux x64 RPM Package	67.6 MB	 <a href="#">jre-8u251-linux-x64.rpm</a>
Linux x64 Compressed Archive	83.49 MB	 <a href="#">jre-8u251-linux-x64.tar.gz</a>
macOS x64 Installer	80.66 MB	 <a href="#">jre-8u251-macosx-x64.dmg</a>
macOS x64 Compressed Archive	72.24 MB	 <a href="#">jre-8u251-macosx-x64.tar.gz</a>
Solaris SPARC 64-bit	46.28 MB	 <a href="#">jre-8u251-solaris-sparcv9.tar.gz</a>
Solaris x64 Compressed Archive	50.06 MB	 <a href="#">jre-8u251-solaris-x64.tar.gz</a>
Windows x86 Online	1.97 MB	 <a href="#">jre-8u251-windows-i586-iftw.exe</a>
Windows x86 Offline	65.52 MB	 <a href="#">jre-8u251-windows-i586.exe</a>
Windows x86	67.99 MB	 <a href="#">jre-8u251-windows-i586.tar.gz</a>
Windows x64	73.73 MB	 <a href="#">jre-8u251-windows-x64.exe</a>
Windows x64	73.12 MB	 <a href="#">jre-8u251-windows-x64.tar.gz</a>

# Instructions supplémentaires pour les utilisateurs Windows

1. Attention: Gephi s'installe tout seul en **32-bit** et il nous faut du **64-bit**

- ouvrir `C:\Program Files (x86)\Gephi-0.9.1\bin\`.
- Il faut executer `gephi64.exe` et pas `gephi.exe` si vous avez une machine récente (*Windows 10 64bit*).

2. pour trouver `gephi64` : aller dans le menu Démarrage

- Chercher `gephi`;
- Clic droit (propriétés): donne l'adresse du dossier du fichier où est `gephi.exe`, qui est aussi le dossier où est `gephi64.exe`
- Lancer `properties`. Contrôler que le chemin de fichier est celui de ``gephi64.exe`?.

3. Penser (très sérieusement) à arrêter d'utiliser *Windows*.

# Préparation des données

# Types de réseau

- réseau normal
- réseau bipartite

# Réseau normal

On a qu'un seul type de sommet, et chaque individu est relié à un autre par une valeur commune. Dans ce cas les auteurs, imprimeurs et éditeurs sont reliés par l'identifiant de l'imprimé qui les a réunis.

Clef	Valeur
Imprimé_1	Auteur_1
Imprimé_1	Imprimeur_1
Imprimé_1	Imprimeur_2
Imprimé_1	Editeur_1
Imprimé_2	Auteur_1
Imprimé_2	Editeur_2

# Réseau bipartite

On a deux types de sommets, chaque individu étant lié à celui à qui il est associé dans le tableau.

Clef	Valeur
Auteur_1	Imprimeur_1
Auteur_1	Imprimeur_1
Auteur_1	Imprimeur_2
Auteur_1	Editeur_1
Auteur_1	Editeur_2
Imprimeur_1	Imprimeur_1
Imprimeur_1	Editeur_1
Imprimeur_2	Editeur_1
Auteur_1	Editeur_2

## Créer le réseau pour Gephi

À partir de ce type de tableau, vous pouvez créer les fichiers nécessaires pour *Gephi* en utilisant l'application en ligne

<https://medialab.github.io/table2net>.

# Premier test

# Ouverture

Nous allons devoir importer nos données. Elles peuvent être de deux types:

- des fichiers de tableau classiques ( `.tsv` , `.csv` )
- des fichiers `.gexf` (*Graph Exchange XML Format*) – l'équivalent d'un objet *igraph* dans *R*.

# Mes fichiers

Il faut faire attention à utiliser des en-têtes pour les colonnes

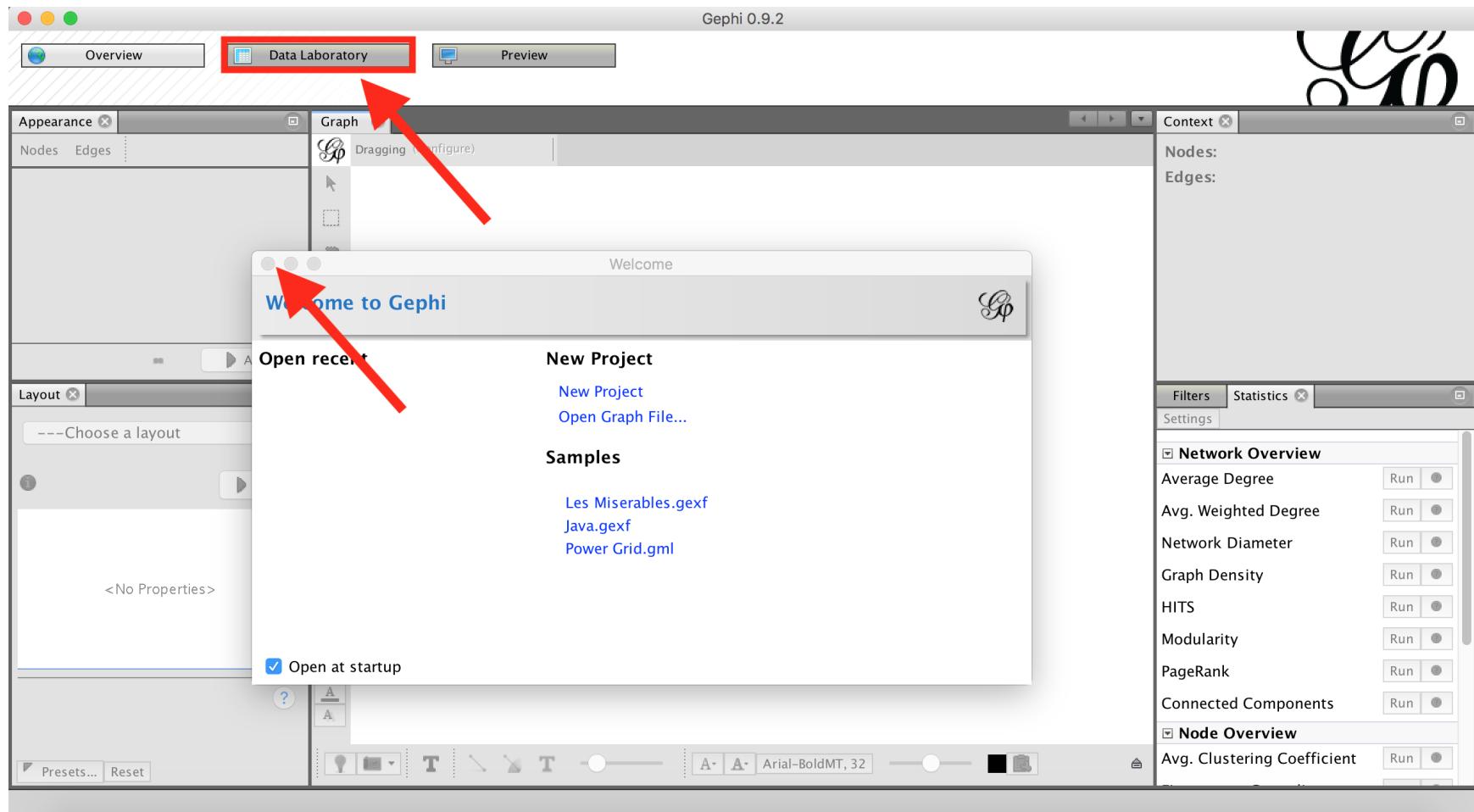
1	Source	Target
2	1	2
3	1	3
4	1	4
5	1	5
6	1	6
7	1	7
8	2	3
9	2	4
10	2	5
11	2	6
12	2	7
13	3	4
14	3	5
15	3	6
16	3	7

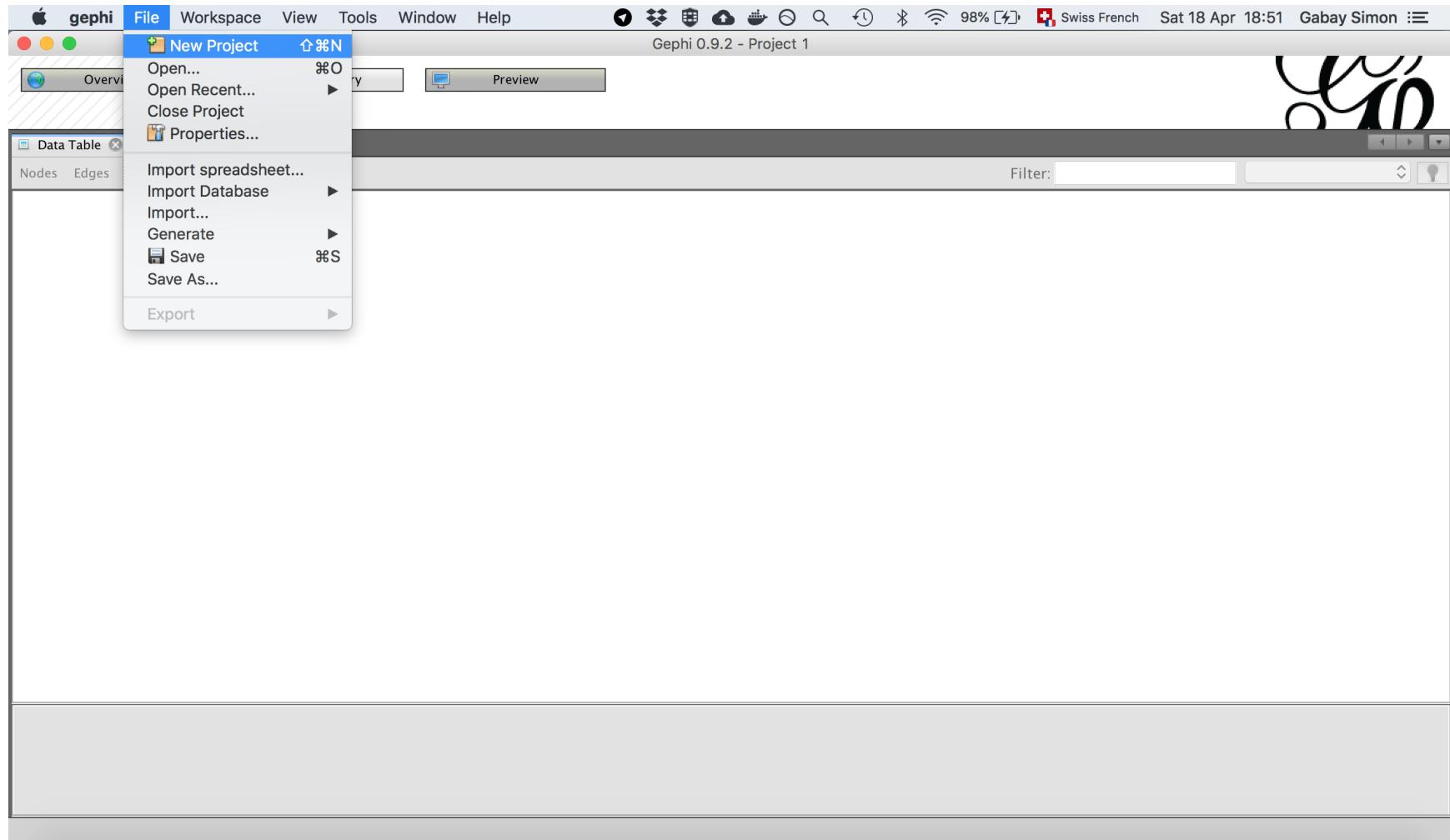
1	Id	Label	Type
2	1	Molière	Auteur
3	2	Guillaume de Luyne	Libraire
4	3	Claude Barbin	Libraire
5	4	Charles de Sercy	Libraire
6	5	Jean Hénault	Imprimeur
7	6	François Noël	Imprimeur
8	7	Christophe Journel	Imprimeur
9	8	Sieur de Neuf-Villenaine	Libraire
10	9	Jean Ribou	Libraire
11	10	Jean Guignard	Libraire
12	11	Gabriel Quinet	Libraire

# Ouverture

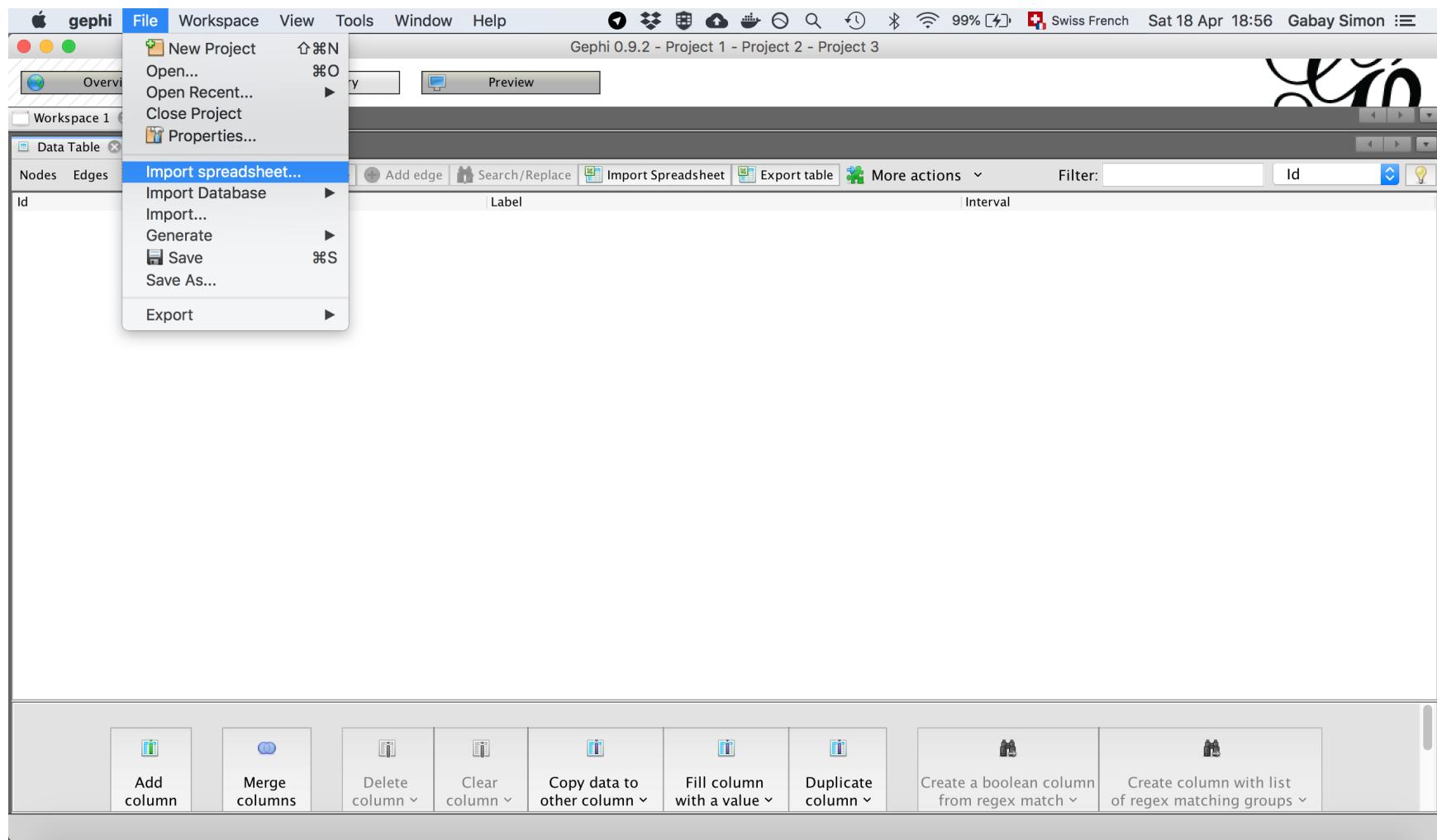
Fermer la fenêtre d'accueil et aller dans le Data Laboratory



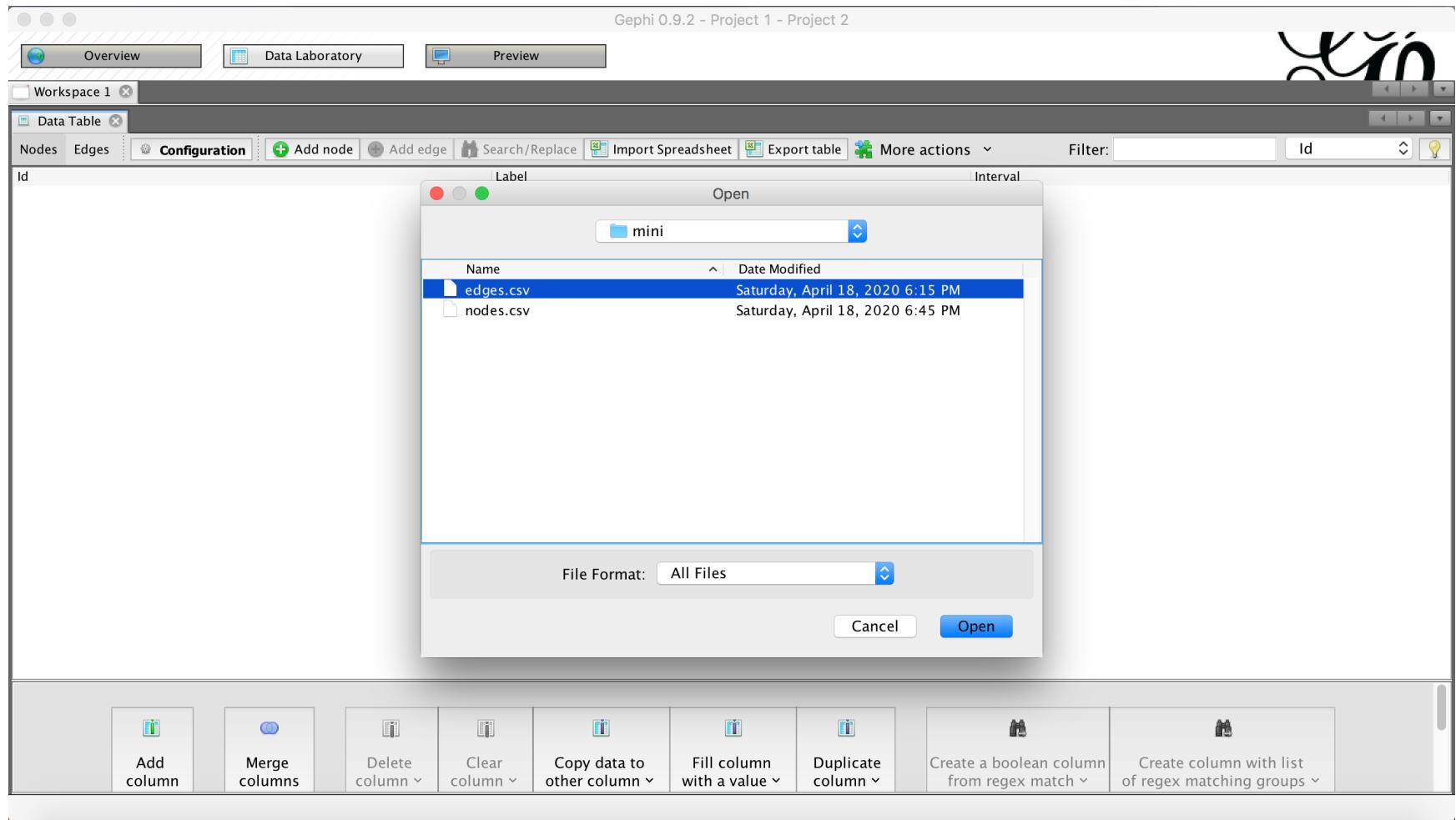
# Nouveau projet



# Importer une feuille de tableur

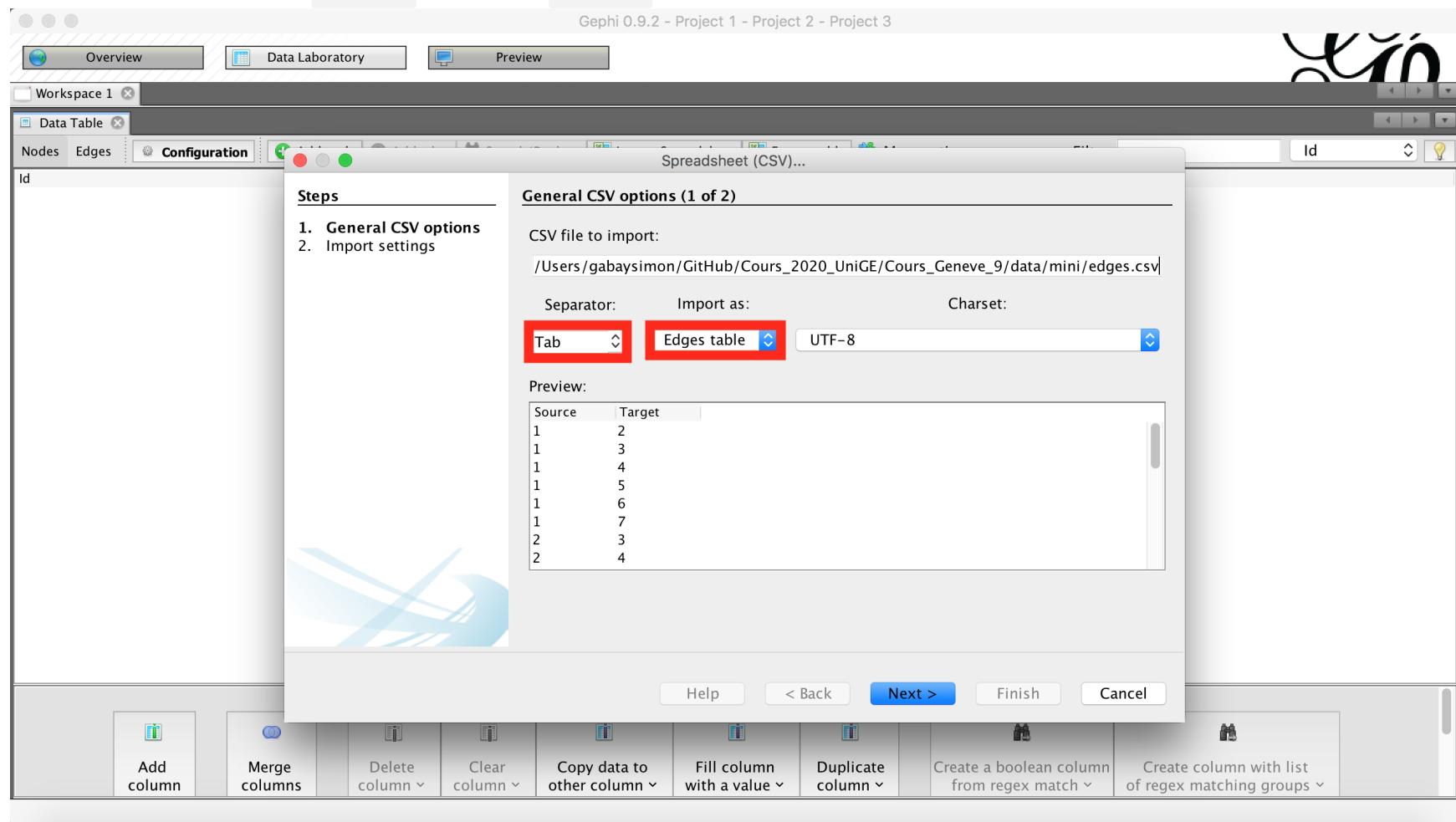


# Choisir la feuille des arêtes



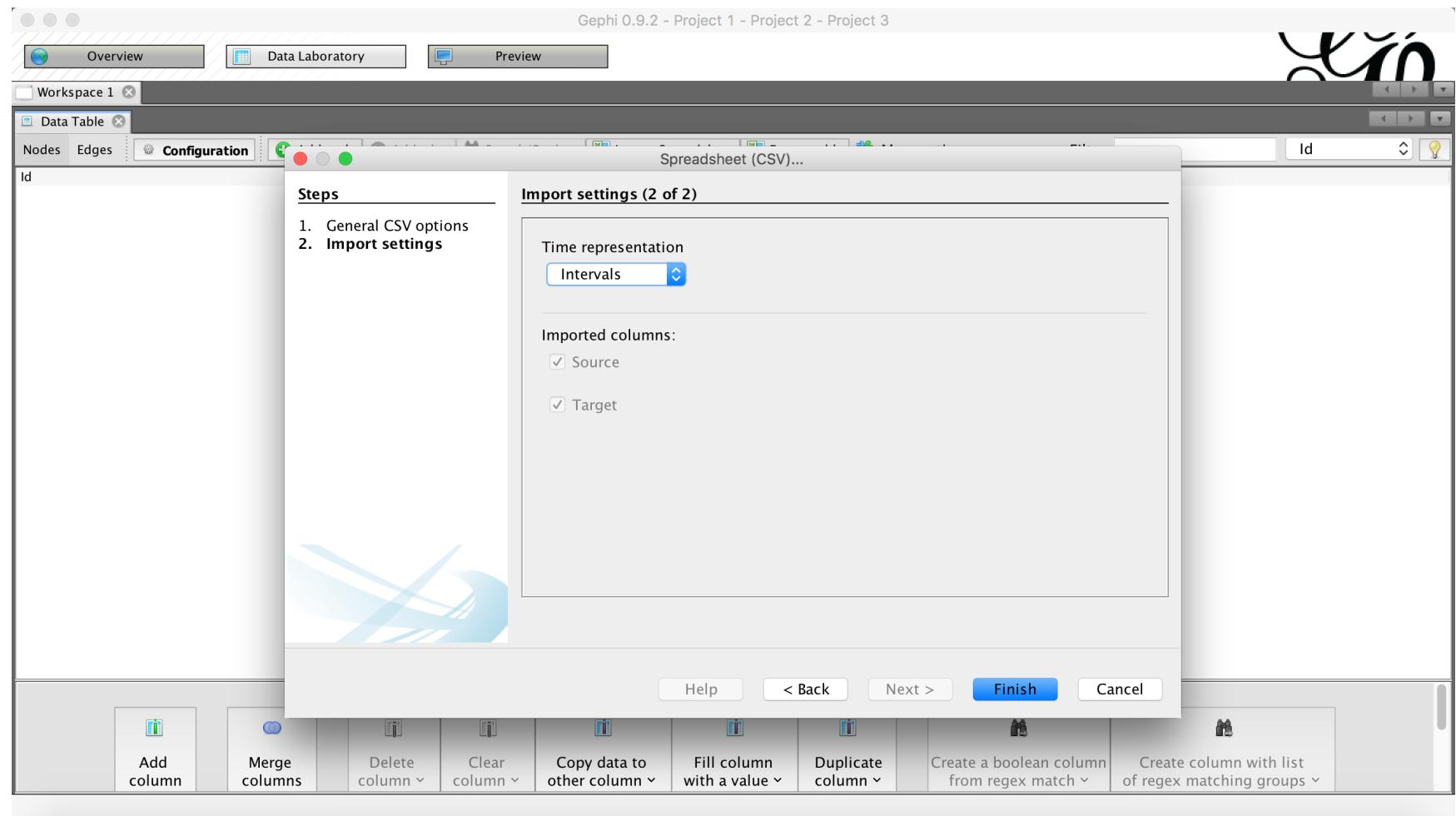
# Définir les paramètres d'import

Normalement les paramètres sont pré-remplis, mais on peut choisir si nous avons un `tsv` ou un `csv`, les nœuds ou les arêtes...



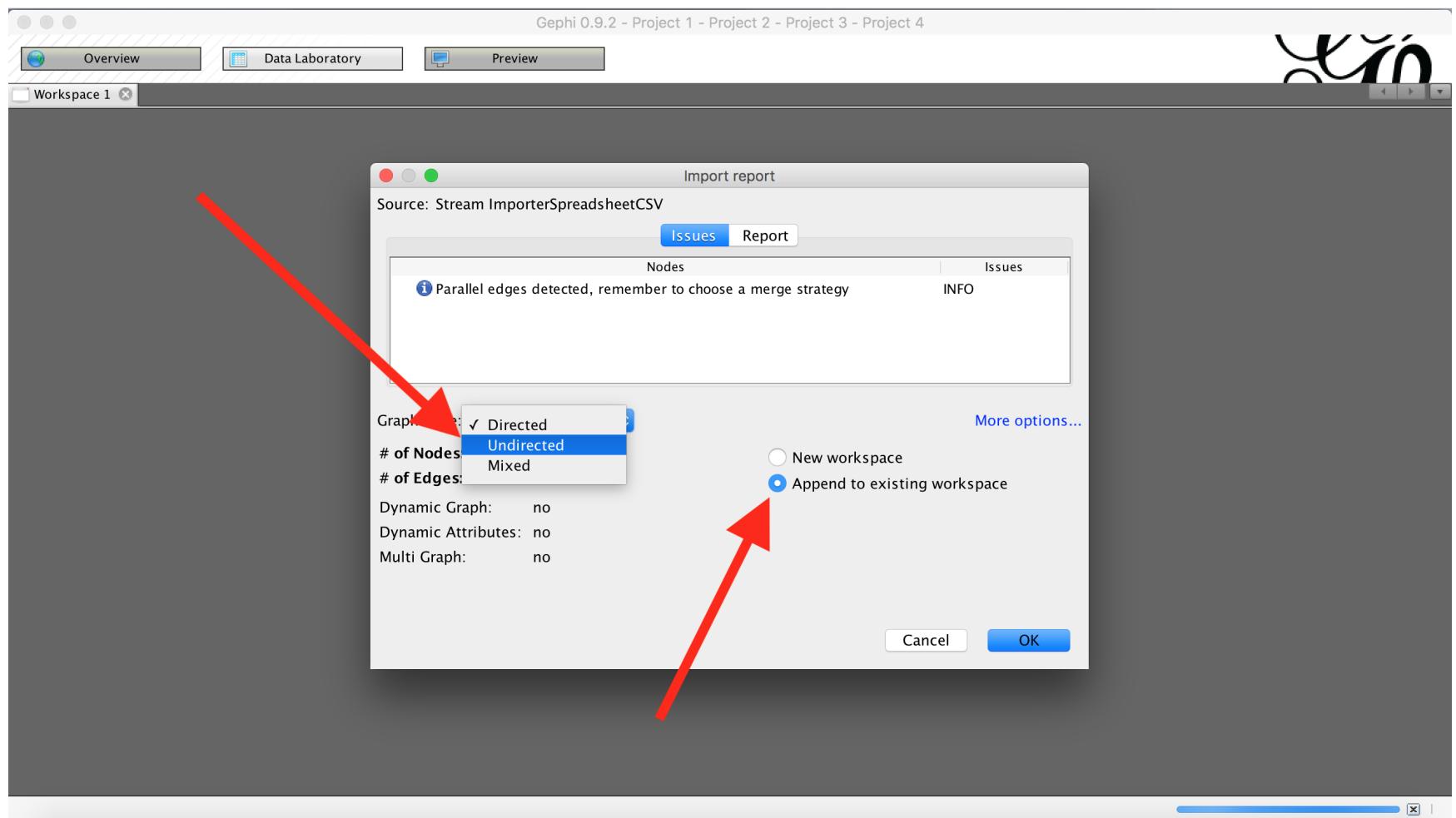
# Définir les paramètres d'import II

Normalement les paramètres sont une nouvelle fois pré-remplis (on se rappelle ici que nous avions utilisé des en-tête...)



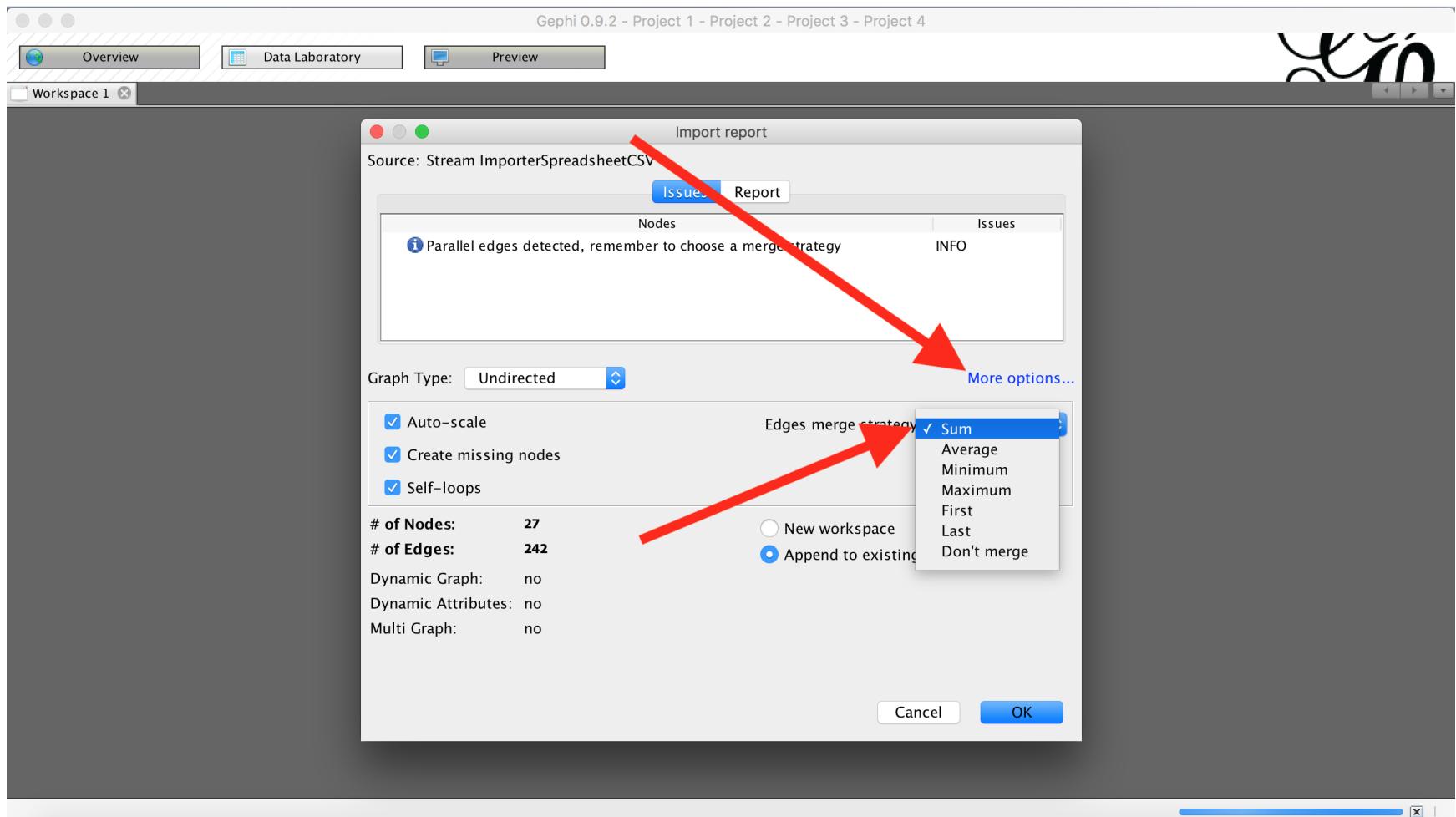
# Finalisation de l'import

On n'oublie pas de cocher la case `Append to existing workspace`, car on veut ouvrir plusieurs feuilles dans le même espace de travail.  
Précisez que le graphe n'est pas dirigé (dans notre cas).



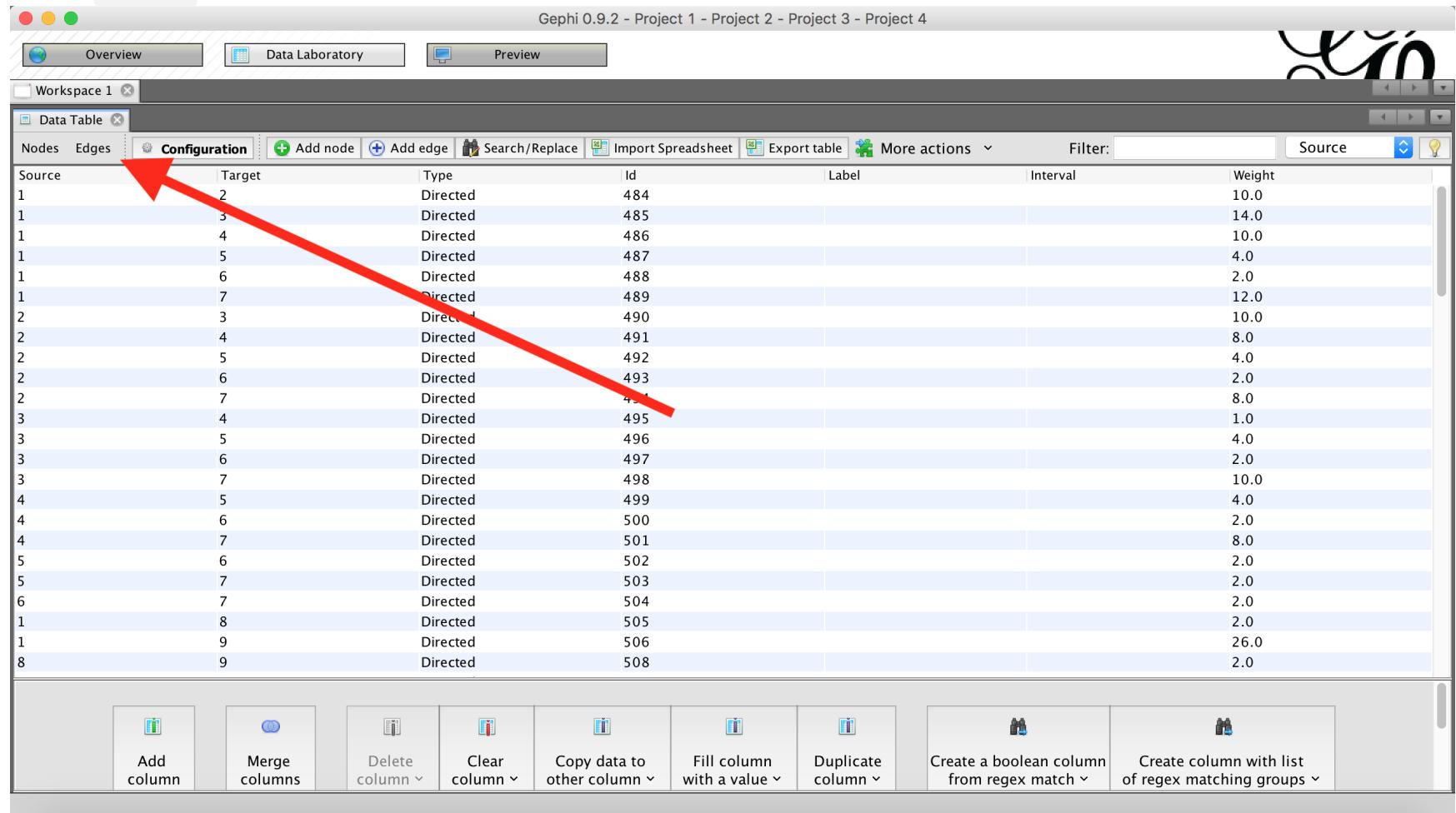
# Finalisation de l'import II

Il existe un bouton **More options...**. Vous pouvez notamment y gérer la question des liens multiples (transformer plusieurs fois le même lien en un poids en faisant la somme, par exemple).



# Observation du résultat

Et voilà! Normalement il n'y a rien si on clique sur `nodes` sauf la liste des `IDs` utilisés pour les arêtes. Importons donc les noeuds!



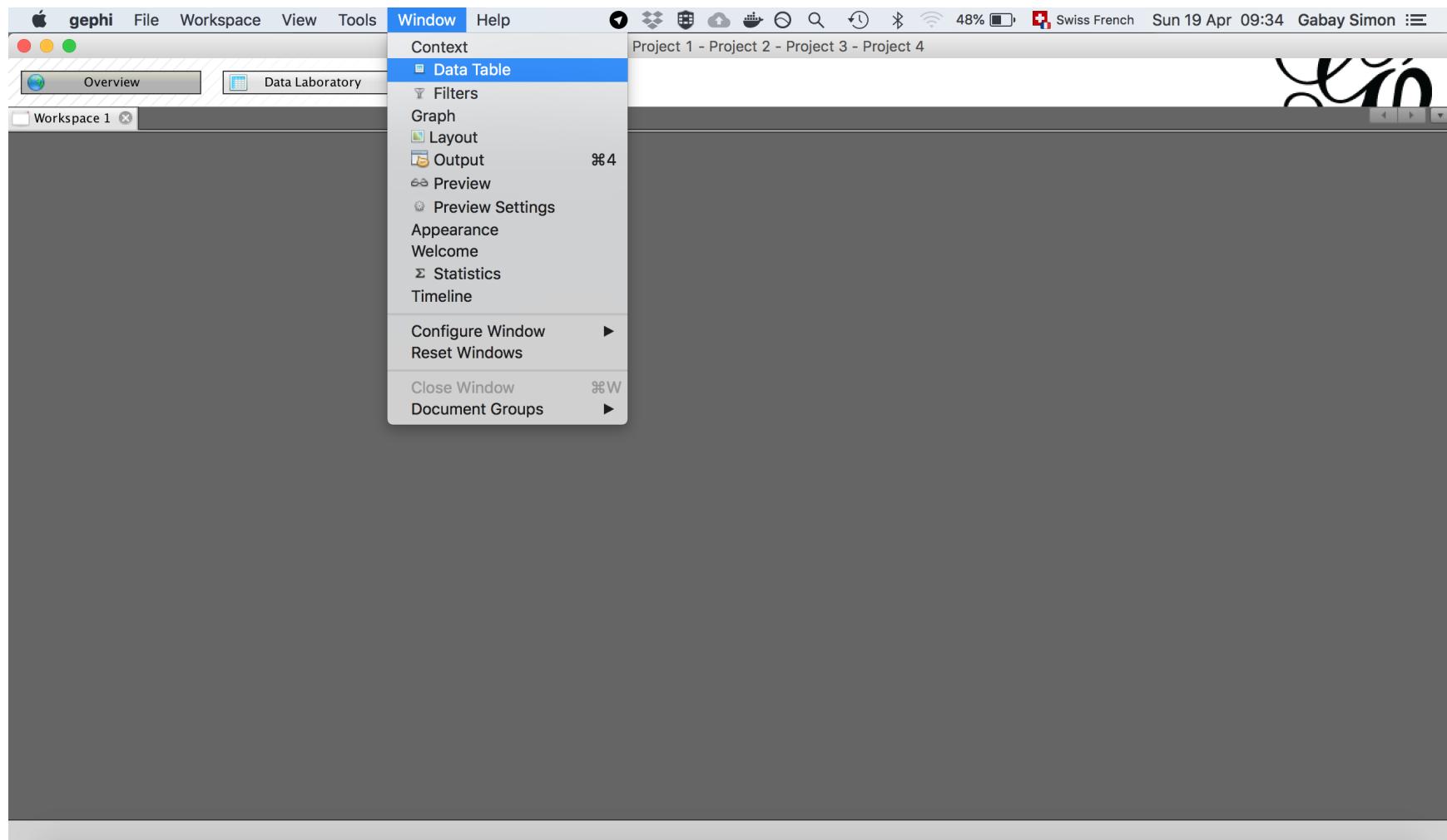
The screenshot shows the Gephi 0.9.2 Data Laboratory interface. At the top, there are tabs for Overview, Data Laboratory, and Preview. Below that is a workspace titled 'Workspace 1'. A 'Data Table' window is open, showing a list of edges. The table has columns: Source, Target, Type, Id, Label, Interval, and Weight. A red arrow points from the text above to the 'Configuration' tab in the table header. The table contains the following data:

Source	Target	Type	Id	Label	Interval	Weight
1	2	Directed	484			10.0
1	3	Directed	485			14.0
1	4	Directed	486			10.0
1	5	Directed	487			4.0
1	6	Directed	488			2.0
1	7	Directed	489			12.0
2	3	Directed	490			10.0
2	4	Directed	491			8.0
2	5	Directed	492			4.0
2	6	Directed	493			2.0
2	7	Directed	494			8.0
3	4	Directed	495			1.0
3	5	Directed	496			4.0
3	6	Directed	497			2.0
3	7	Directed	498			10.0
4	5	Directed	499			4.0
4	6	Directed	500			2.0
4	7	Directed	501			8.0
5	6	Directed	502			2.0
5	7	Directed	503			2.0
6	7	Directed	504			2.0
1	8	Directed	505			2.0
1	9	Directed	506			26.0
8	9	Directed	508			2.0

At the bottom of the table window, there is a toolbar with various icons for column operations: 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, and Create column with list of regex matching groups.

# Observation du résultat

Si vous ne voyez pas apparaître automatiquement les données,  
affichez-les



## Les nœuds

On recommence la même procédure avec les nœuds. À vous de jouer!

# Observation du résultat

Gephi 0.9.2 - Project 1 - Project 2 - Project 3 - Project 4

Overview Data Laboratory Preview

Workspace 1

Data Table

Nodes Edges Configuration	Add node	Add edge	Search/Replace	Import Spreadsheet	Export table	More actions	Filter:	Id	
Id	Label	Interval					type		
1	Molière						Auteur		
2	Guillaume de Luyne						Libraire		
3	Claude Barbin						Libraire		
4	Charles de Sercy						Libraire		
5	Jean Héault						Imprimeur		
6	François Noël						Imprimeur		
7	Christophe Journel						Imprimeur		
8	Sieur de Neuf-Villeneuve						Libraire		
9	Jean Ribou						Libraire		
10	Jean Guignard						Libraire		
12	Gabriel Quinet						Libraire		
13	François II Noël						Imprimeur		
14	Thomas Jolly						Libraire		
15	Louis Billaine						Libraire		
16	Estienne Loysen						Libraire		
18	Pierre Trabouillet						Libraire		
19	Nicolas Le Gras						Libraire		
20	Théodore Girard						Libraire		
21	Etienne Maucroy						Imprimeur		
17	Claude Blageart						Imprimeur		
22	Claude II Calleville						Imprimeur		
23	Claude Audinet						Imprimeur		
24	Pierre Le Monnier						Libraire		
25	Pierre Corneille						Auteur		

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

# Naviguer dans Gephi

- Overview permet de créer le graphe
- Data Laboratory permet de gérer les données
- Preview permet de finaliser l'image du graphe produite

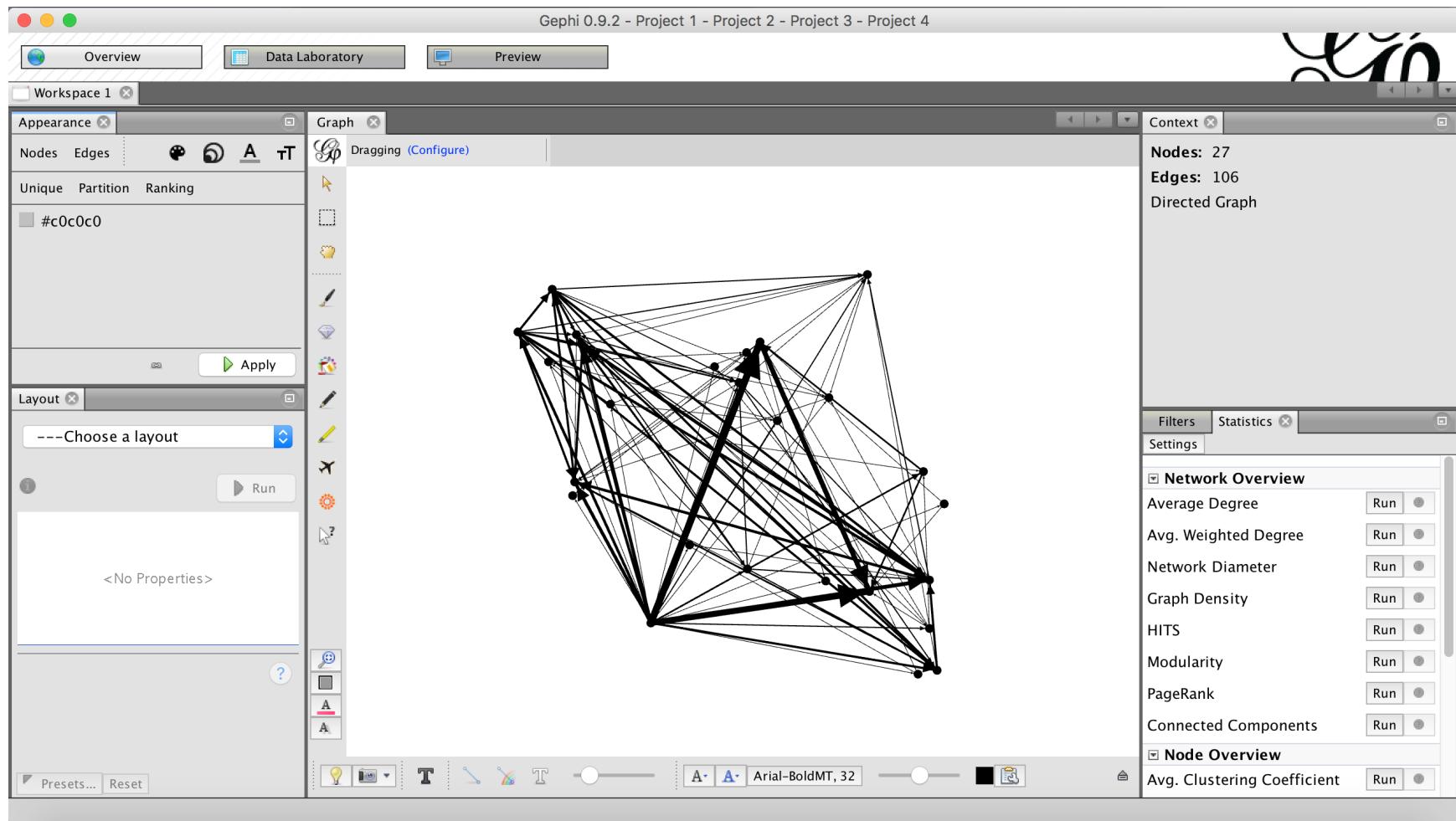
The screenshot shows the Gephi 0.9.2 interface with the 'Data Laboratory' tab selected. The window title is 'Gephi 0.9.2 - Project 1 - Project 2 - Project 3 - Project 4'. The main area displays a 'Data Table' for 'Nodes' with 25 entries. The columns are 'Id', 'Label', 'Interval', and 'type'. The 'type' column contains values like 'Auteur', 'Libraire', and 'Imprimeur'. Below the table is a toolbar with various data manipulation icons.

Id	Label	Interval	type
1	Molière		Auteur
2	Guillaume de Luyne		Libraire
3	Claude Barbin		Libraire
4	Charles de Sercy		Libraire
5	Jean Hénault		Imprimeur
6	François Noël		Imprimeur
7	Christophe Journeel		Imprimeur
8	Sieur de Neuf-Villenaine		Libraire
9	Jean Ribou		Libraire
10	Jean Guignard		Libraire
12	Gabriel Quinet		Libraire
13	François II Noël		Imprimeur
14	Thomas Jolly		Libraire
15	Louis Billaine		Libraire
16	Estienne Loyson		Libraire
18	Pierre Trabouillet		Libraire
19	Nicolas Le Gras		Libraire
20	Théodore Girard		Libraire
21	Etienne Maucroy		Imprimeur
17	Claude Blageart		Imprimeur
22	Claude II Calleville		Imprimeur
23	Claude Audinet		Imprimeur
24	Pierre Le Monnier		Libraire
25	Pierre Corneille		Auteur

Toolbar icons (from left to right): 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.

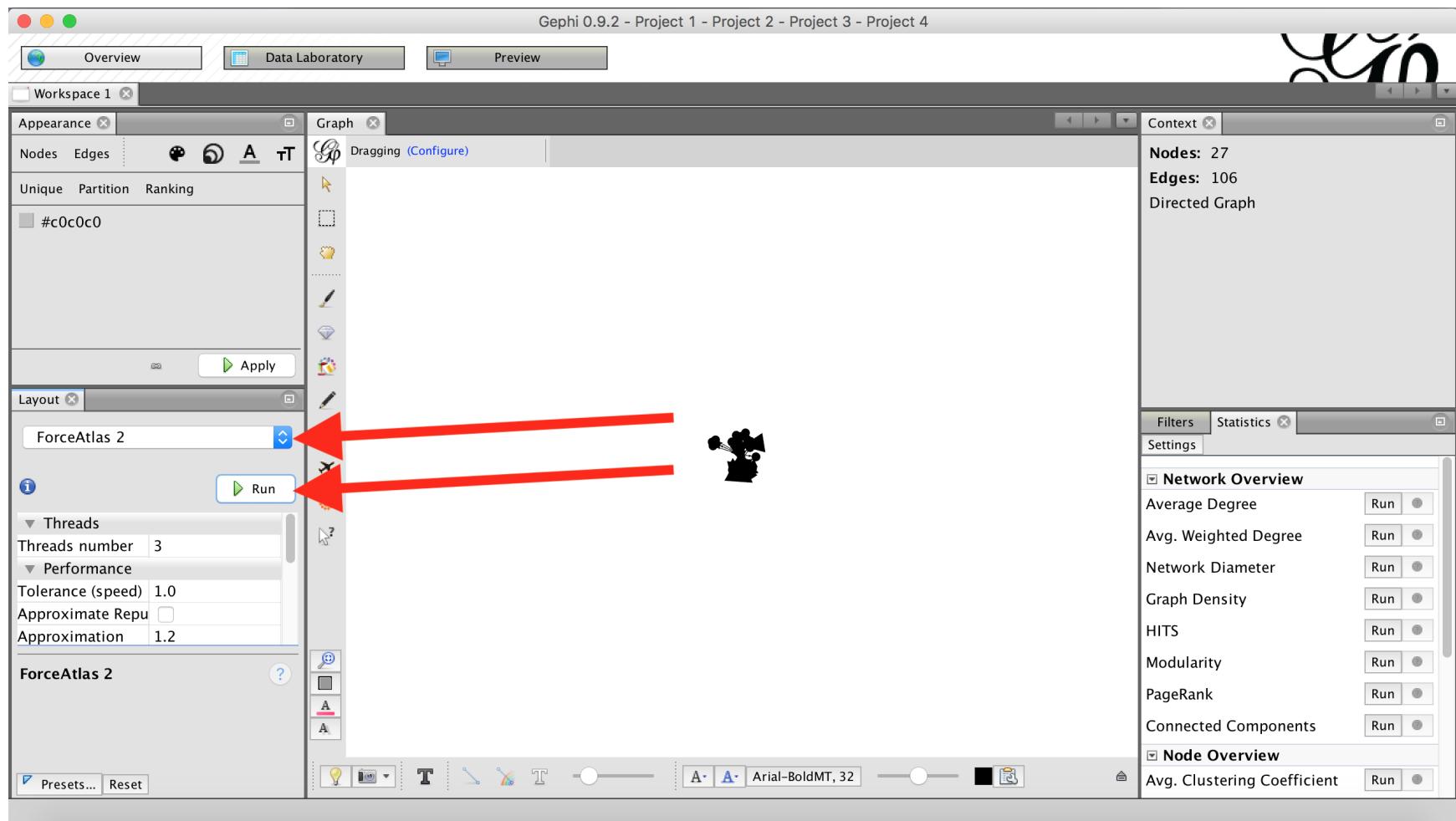
# Overview

Ici nous allons pouvoir manipuler le graphe: choisir le layout, faire des calculs de centralité, etc.



# Layout

Choisissons un layout au hasard, pour voir ce que ça donne. Par exemple **Force Atlas 2**. On clique sur **Run**, on attend que le graphe se stabilise, puis sur **Stop**.

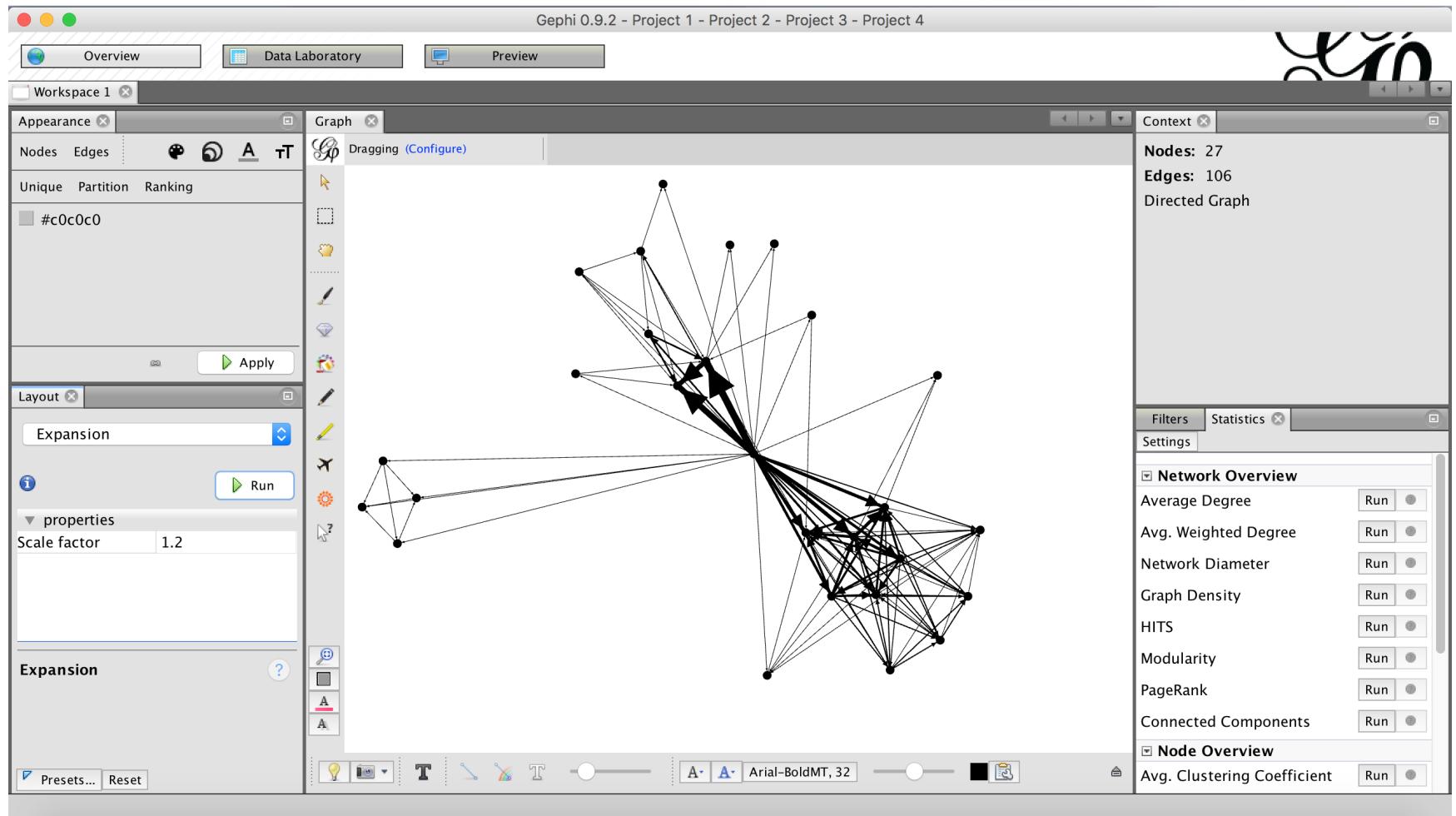


## Layout II

Rappel: on ne choisit pas les *layouts* au hasard: on le fait en fonction du type de données, et des options proposées. *Force Atlas* offre certaines fonctionnalités utile, comme *no overlap* ("Empêcher le recouvrement") qui permet d'éviter que deux nœuds se chevauchent, ce qui peut être un problème pour la production du rendu final – pensez à le cocher dans les options!

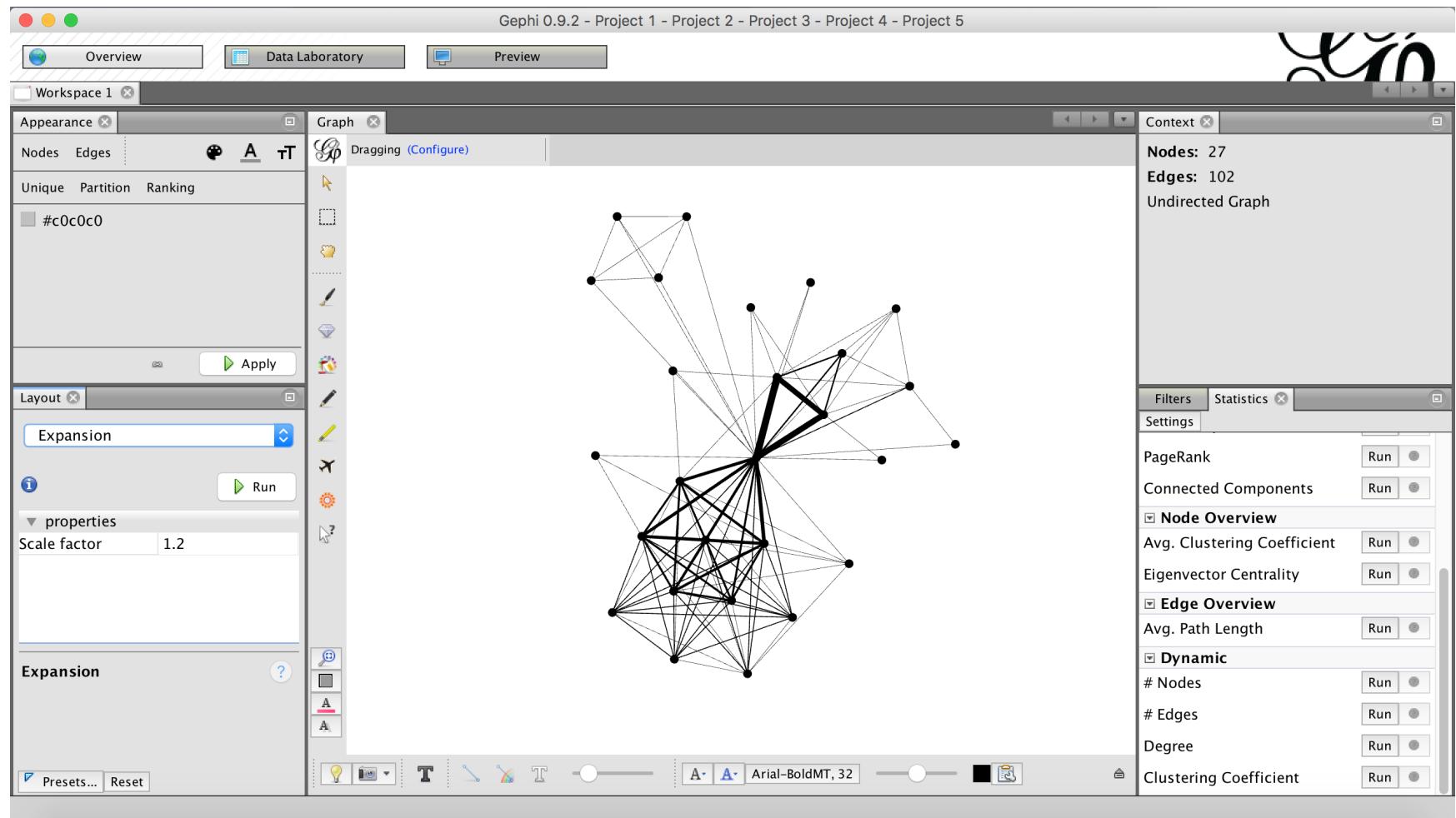
# Layout III

Si le résultat est illisible, on peut "l'agrandir" avec le layout **expansion**. Répéter l'opération autant de fois que nécessaire.



# Layout IV

Le même graphe sans avec l'option *no overlap*

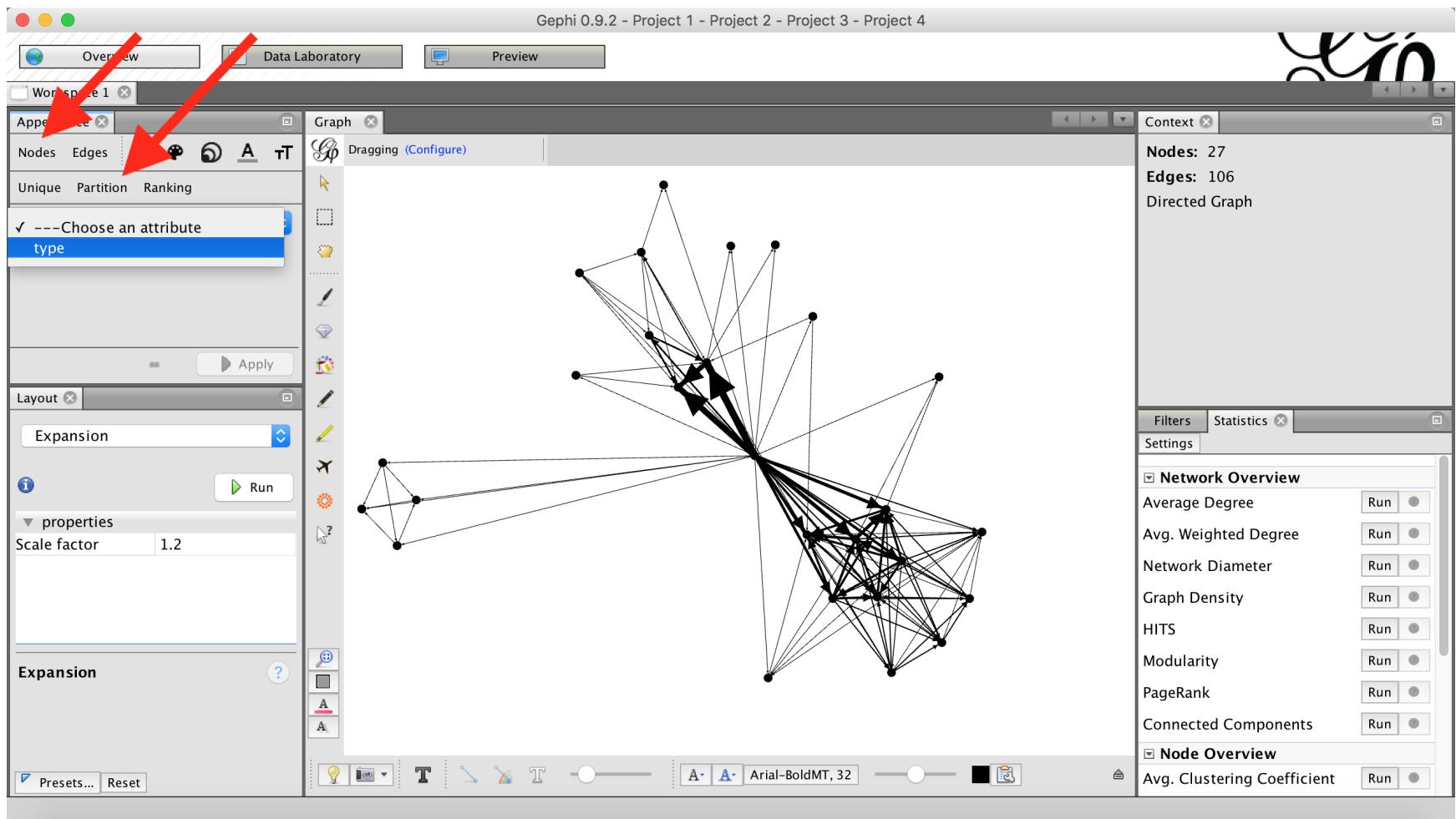


## **Layout IV**

Explorez les différents layouts, les différentes options

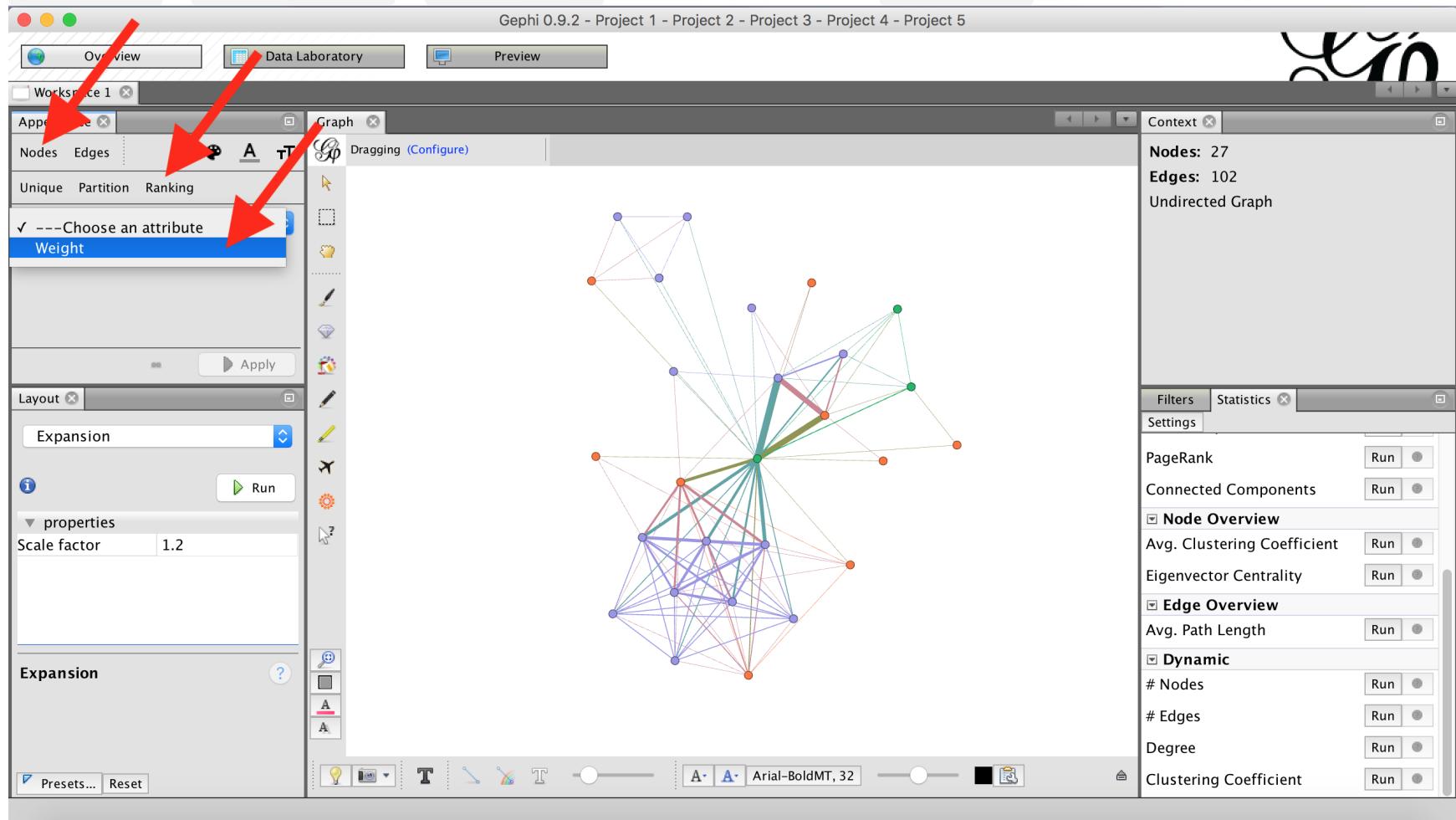
# Coloriage

Il est possible de customiser son graphe simplement. On se rappelle que, dans nos données, nous avons un **Type** pour les nœuds: colorions ces derniers en fonction du type dans **Appearance**.



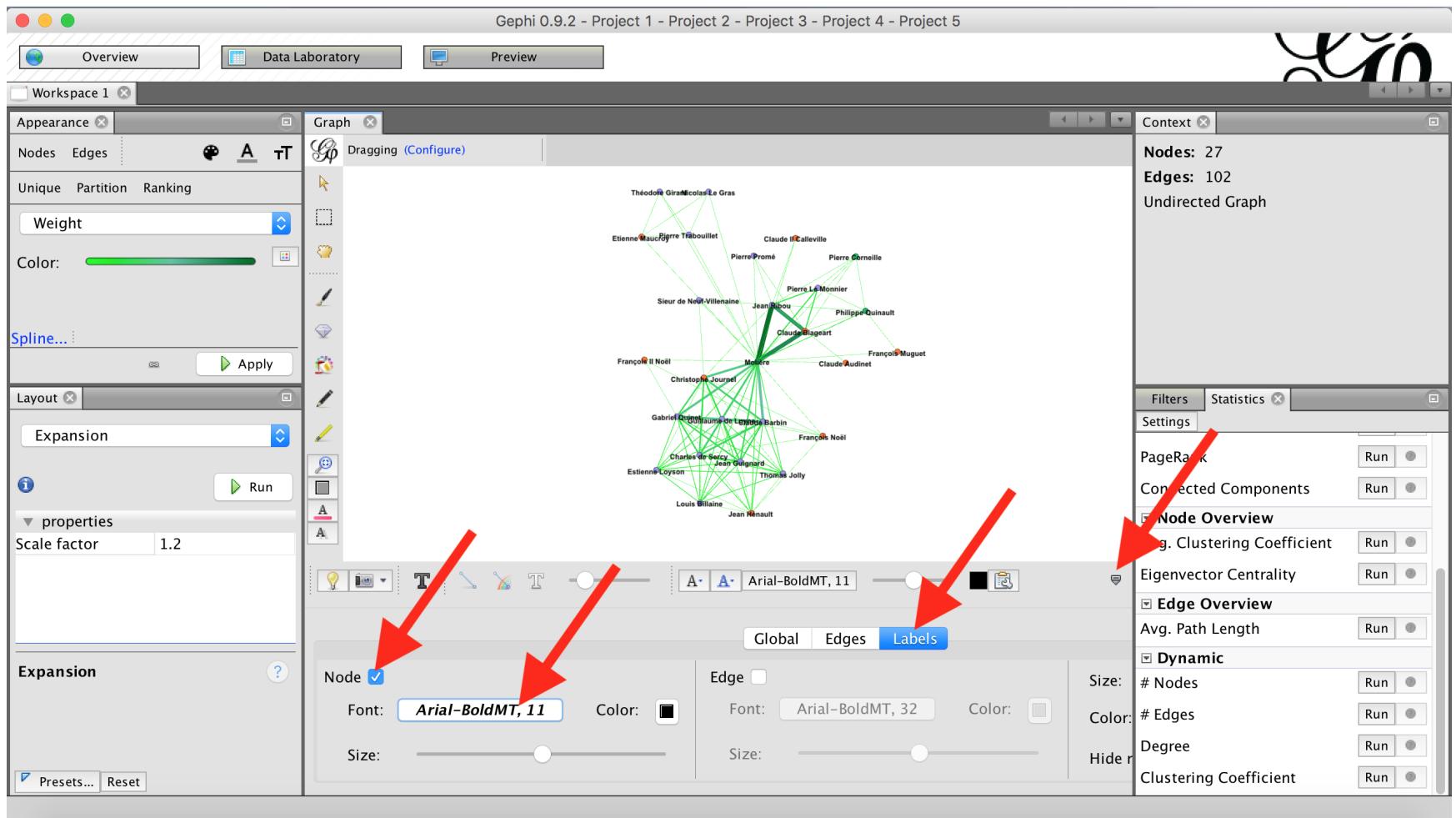
# Coloriage II

Comme les arêtes ont un poids, on peut ajuster la couleurs des liens en fonction du poids. Retournez dans **Appearance**, choisissez **Edges > Ranking > Weight** et cliquez sur **Apply**.



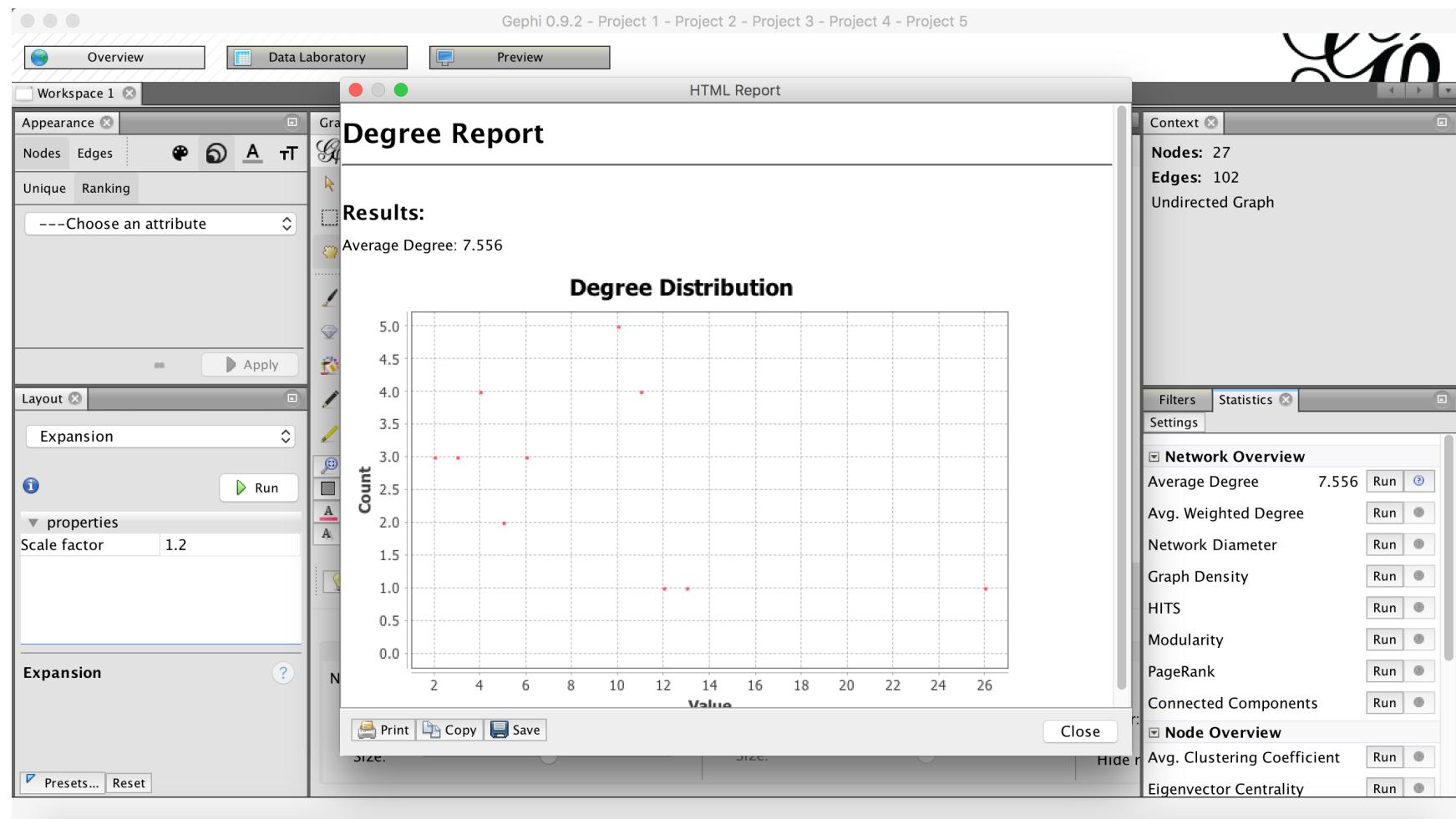
# Coloriage III

On se rappelle aussi que les nœuds représentent des personnes, qui ont un nom. Affichons-les! Cliquez sur l'icone en bas à droite du graphe, choisissez Labels , cliquez sur Nodes , choisissez la police et la taille



# Statistics

Les différentes indicateurs (densité, centralité...) sont produits simplement dans le panneau **Statistics**.



# Statistics II

Le résultat est ajouté dans le Data laboratory . Il en ira de même pour chaque calcul effectué.

The screenshot shows the Gephi 0.9.2 interface with the 'Data Laboratory' tab selected. A table titled 'Data Table' displays 25 rows of node data. The columns are labeled 'Id', 'Label', 'Interval', 'type', and 'Degree'. The 'Degree' column is highlighted with a red box. The data shows various historical figures and their degrees of connectivity. The 'Configuration' tab is also visible at the top of the table window.

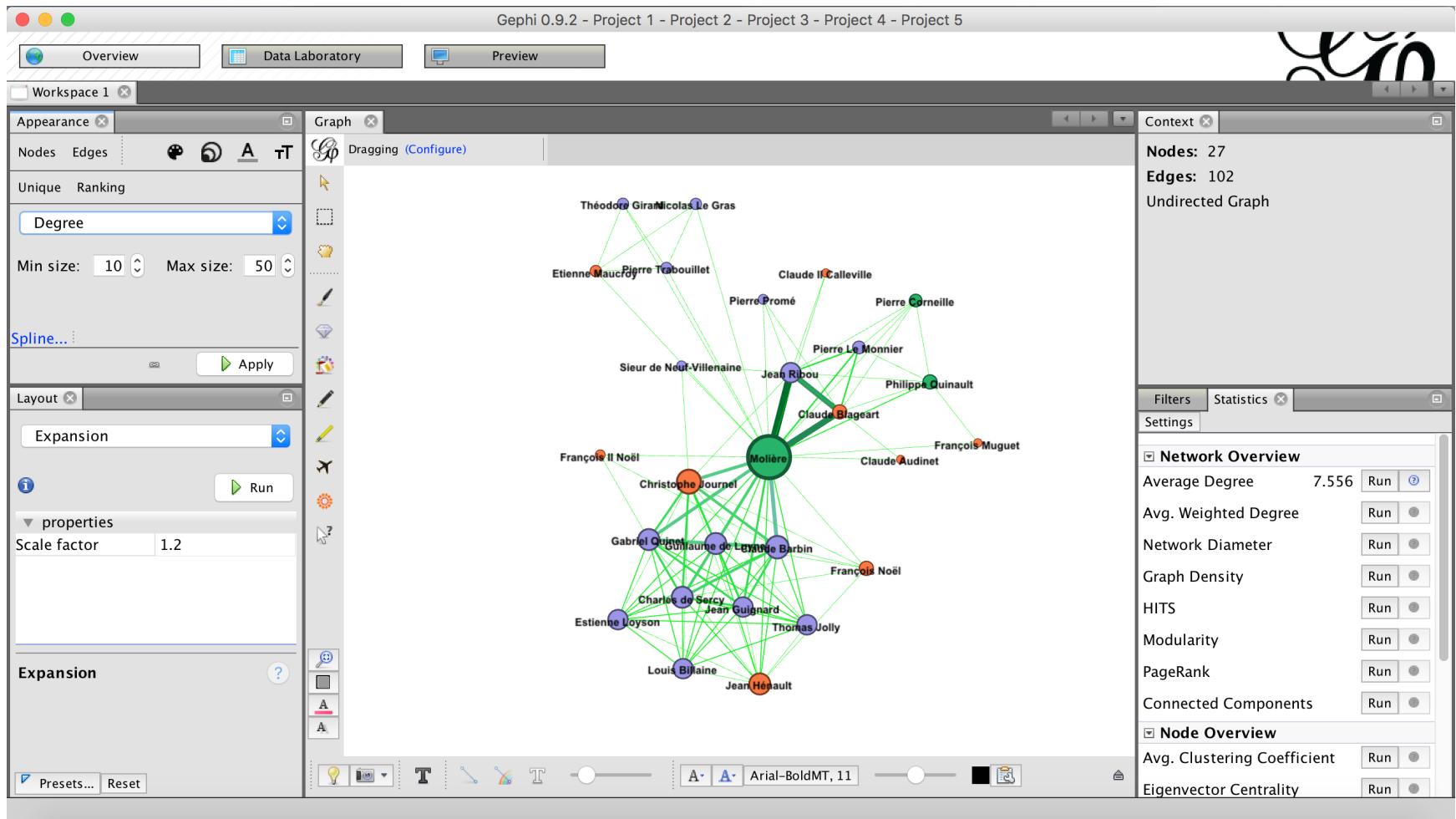
Id	Label	Interval	type	Degree
1	Molière		Auteur	26
2	Guillaume de Luyne		Libraire	11
3	Claude Barbin		Libraire	12
4	Charles de Sercy		Libraire	11
5	Jean Hénault		Imprimeur	11
6	François Noël		Imprimeur	6
7	Christophe Journel		Imprimeur	13
8	Sieur de Neuf-Villenaïne		Libraire	3
9	Jean Ribou		Libraire	10
10	Jean Guignard		Libraire	10
12	Gabriel Quinet		Libraire	11
13	François II Noël		Imprimeur	3
14	Thomas Jolly		Libraire	10
15	Louis Billaine		Libraire	10
16	Estienne Loyson		Libraire	10
18	Pierre Trabouillet		Libraire	4
19	Nicolas Le Gras		Libraire	4
20	Théodore Girard		Libraire	4
21	Etienne Maucroy		Imprimeur	4
17	Claude Blageart		Imprimeur	6
22	Claude II Calleville		Imprimeur	2
23	Claude Audinet		Imprimeur	2
24	Pierre Le Monnier		Libraire	5
25	Pierre Corneille		Auteur	5

Below the table are several toolbar buttons for data manipulation:

- 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 ▾

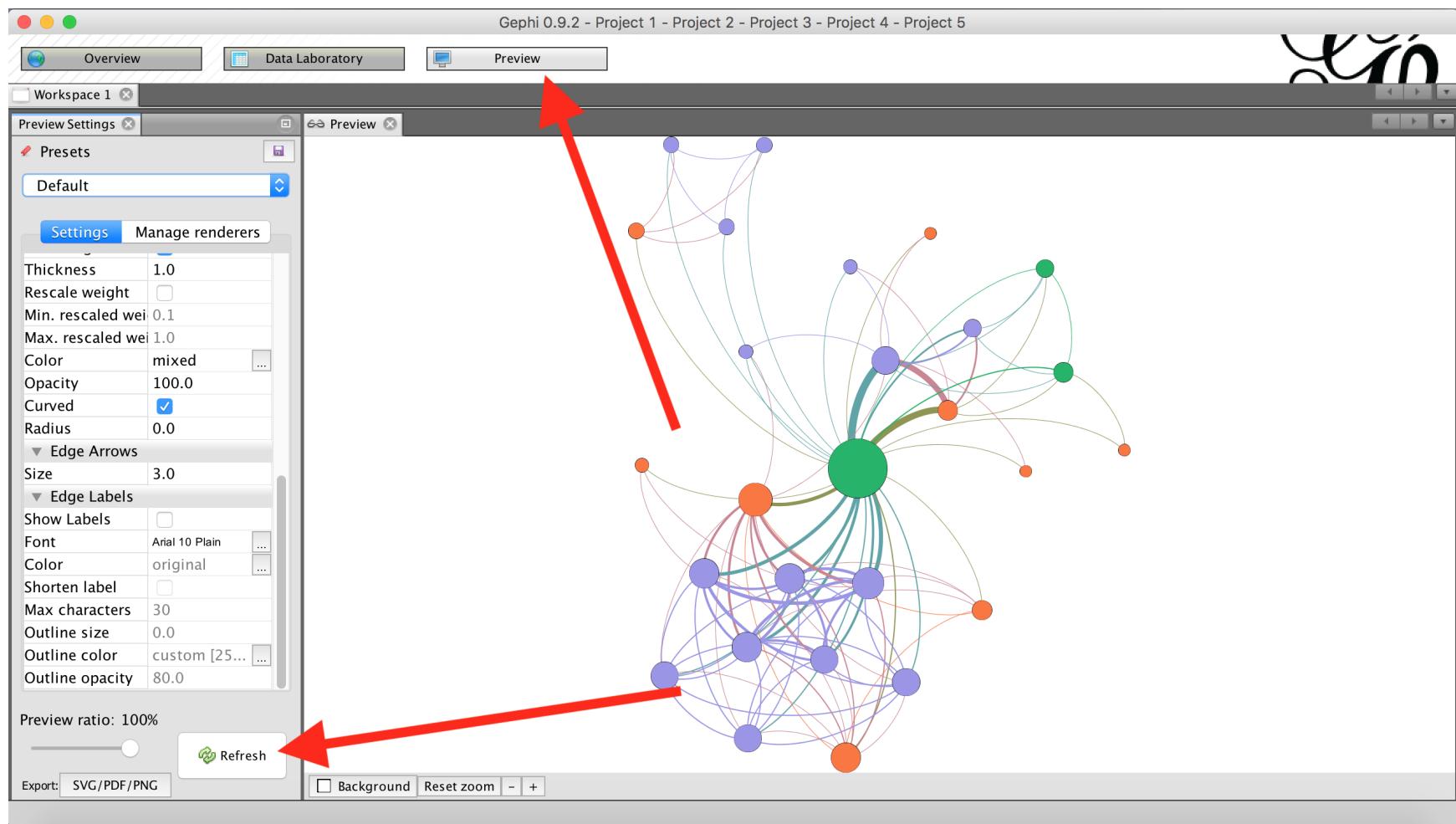
# Coloriage IV

On peut récupérer les valeurs produites pour les afficher. Retournez dans **Appearance**, cliquez sur **Edges** puis l'icone **Size**, et choisissez **Ranking > Degree** et cliquez sur **Apply**.



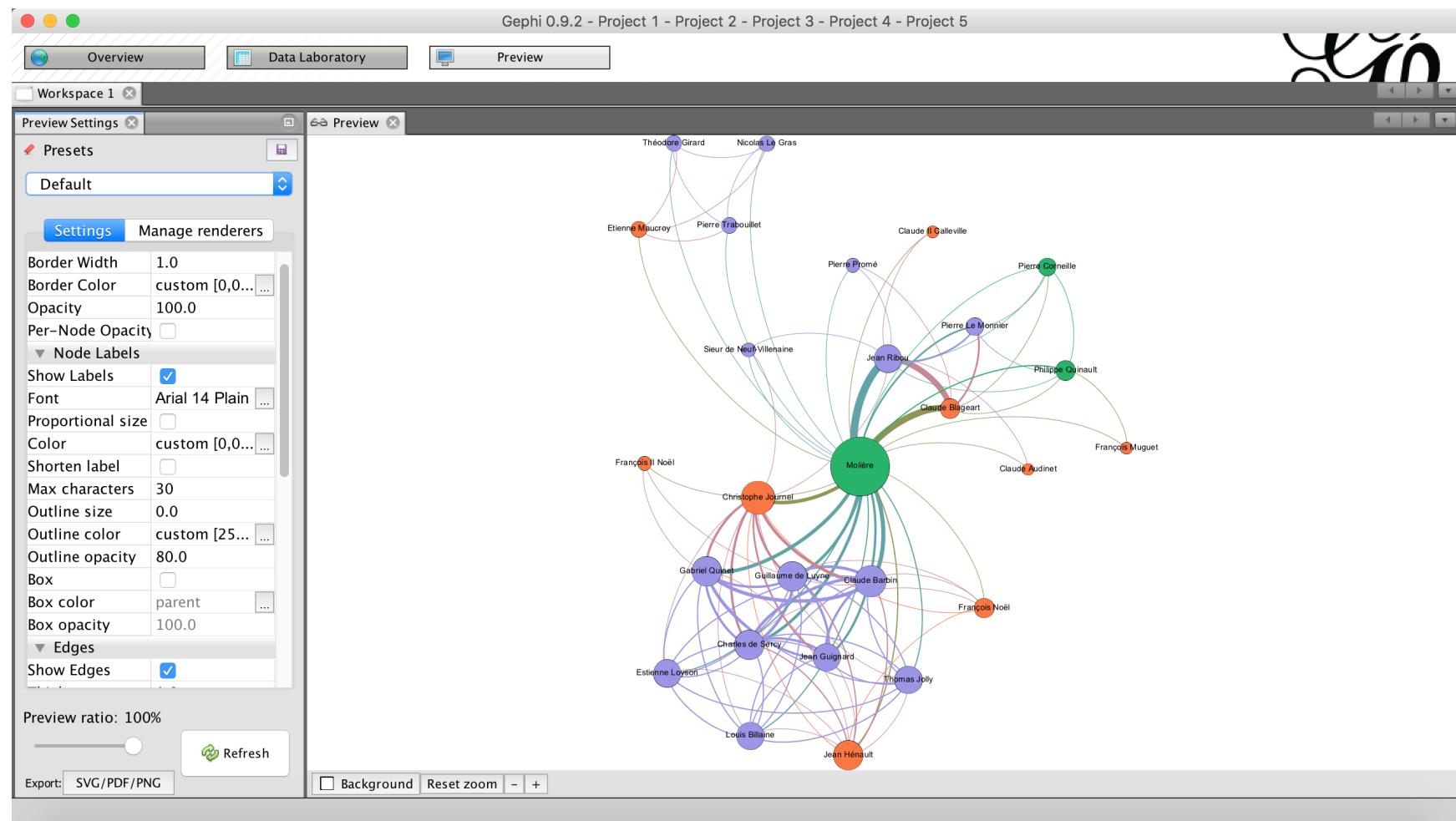
# Preview

Allez dans l'onglet Preview , et cliquez sur Refresh (il faudra cliquer sur ce bouton après chaque manipulation).



# Preview II

Finalisez votre publication en ajoutant toutes les informations nécessaires: nom des labels, etc.



# Export

Vous pouvez exporter/sauvegarder votre document en différents formats:

- L'image en JPG, PDF ou SVG (recommandé pour l'ouvrir dans *Inkscape* ou *Illustrator*)
- Les données en CSV, TSV, GEFX, GML...

**À vous de jouer**

Recommencez l'exercice, mais avec un graphe de très grande taille, surtout pour voir l'impact des différents algorithmes de tracé de graphe sur la spatialisation du résultat. Nous vous invitons à importer et visualiser les données dans le dossier Data/Maxi (que nous reprenons à Martin Grandjean, dont [le site](#) et les travaux vous seront d'une grande utilité si vous vous intéressez aux réseaux).