



# Projet de cartographie de l'information

La cartographie de l'information est une technique qui permet de représenter graphiquement les données et les relations entre elles. Cette technique est utilisée dans de nombreux domaines, notamment la science des données, la gestion de projet, la sociologie et l'analyse de réseaux.

Un outil couramment utilisé pour la cartographie de l'information est Gephi, un logiciel open source qui permet de visualiser et d'analyser des données de réseau. Gephi offre une interface graphique intuitive qui permet aux utilisateurs de charger des données, de les préparer et de les visualiser sous différentes formes.

La cartographie de l'information avec Gephi permet de mieux comprendre les relations entre les différentes données et de découvrir des patterns et des tendances qui pourraient être difficiles à repérer autrement. Elle peut également être utilisée pour visualiser des réseaux sociaux, des données de communication, des données de trafic et de nombreux autres types de données.

Ce projet explore le réseau de personnages du roman de Victor Hugo, Les Misérables (1862). Le roman d'Hugo suit la vie et les interactions de nombreux personnages en France avant la rébellion de 1832 à Paris. J'ai lu « Les Mis » il y a un certain temps, et je ne peux me rappeler que le scénario général et quelques-unes des intrigues de personnages. Comme je ne connais plus les nombreux personnages du roman et leurs relations, je suis intéressé à comprendre le réseau social global des personnages et leur connectivité. À l'aide d'un ensemble de données libre produit en 1993, le but de ce projet est de distinguer clairement les groupes de caractères en fonction de leur connectivité, les caractères les plus centraux du roman et le degré de co-apparitions.

Anas ELKACEMI

Elève ingénieur des données et connaissance,

ESI, Rabat, Maroc

anas.elkacemi@esi.ac.ma

---

# Output

Lien de la version web de la visualisation :

[https://nocodefunctions.com/html/vosviewer/index.html?json=data/public/vosviewer\\_4ffd1df49dc2ede926f3.json](https://nocodefunctions.com/html/vosviewer/index.html?json=data/public/vosviewer_4ffd1df49dc2ede926f3.json)

+

Joint à ce document : un fichier *output.pdf* qui contient la visualisation avec une bonne qualité.

## Table de matières

Output.....	2
Table de matières .....	2
Table des figures .....	2
Environnement technique .....	3
Dataset.....	3
Data Cleaning .....	3
Network visualisation.....	4
<b>Amélioration de la lisibilité du réseau</b> .....	6
Réglage de l'épaisseur des maillons .....	8
Partage et diffusion .....	9
Résultats et discussion.....	11

## Table des figures

Figure 1 onglet laboratoire de données .....	3
Figure 2 dataset sur gephi .....	3
Figure 3 Apparence du fichier ouvert .....	4
Figure 4 Affichage des étiquettes de nœuds\$ .....	4
Figure 5 modification de quelques paramètres et lancement de la mise en page .....	5
Figure 6 resultat de Force Atlas 2 .....	5
Figure 7 Nous colorerons les nœuds en fonction de leur sexe. Pour ce faire, nous sélectionnons Genre dans le panneau Apparence :.....	6
Figure 8 Après coloration des caractères selon leur sexe .....	6
Figure 9 Noverlap ou Label Adjust .....	7
Figure 10 Ajustement des label .....	7
Figure 11 Ajustement de l'épaisseur .....	8
Figure 12 Classement des tailles de noeuds par centralité .....	8
Figure 13 Résultat .....	9
Figure 14 visualisation sur onglet "prévisualisation" .....	9
Figure 15 interface du site nocodefunctions .....	10

# Environnement technique

Pour ce projet , j'ai acquis l'ensemble de données « Les Misérables » provenant de la base de données open source Github. Le logiciel utilisé pour créer des visualisations à partir de cet ensemble de données était le programme de cartographie de réseau, Gephi. En créant ma visualisation.

## Dataset

Lien de la Dataset : <https://visdatasets.github.io>

Les misérables dataset est un ensemble de données original compilé par Donald Knuth pour la Stanford GraphBase ; extrait du graphique D3 de Mike Bostock.

## Data Cleaning

On a essayé de faire quelques manipulations de nettoyage sur GEPHI à travers l'onglet « laboratoire de données ».

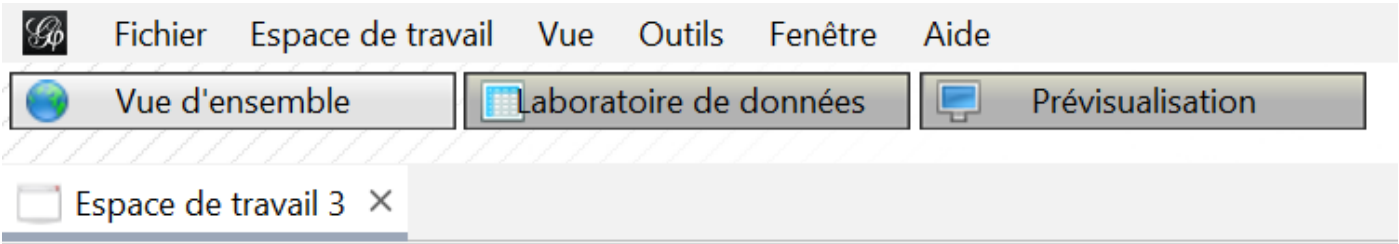


Figure 1 onglet laboratoire de données

- On a essayé de codifier le champ de GENDER pour pouvoir visualiser ce champ sur Gephi.

Id	Label	Interval	Gender	Ec
0	Myriel		M	4.0
1	Napoleon		M	5.0
6	Champfercier		M	5.0
7	Cravatte		M	5.0
8	Count		M	5.0
9	OldMan		M	5.0
10	Labarre		M	4.0
11	Valjean		M	3.0
14	Isabeau		M	4.0
15	Gervais		M	4.0
16	Tholomyes		M	4.0
17	Listolier		M	4.0
18	Fameuil		M	4.0
19	Blacheville		M	4.0
27	Javert		M	3.0
28	Fauchelevant		M	4.0
29	Bamatabois		M	4.0
32	Scaufflaire		M	4.0
34	Judge		M	4.0
35	Champmathieu		M	4.0
36	Brevet		M	4.0
37	Chouffe		M	4.0

Ajouter une colonne

Fusionner les colonnes

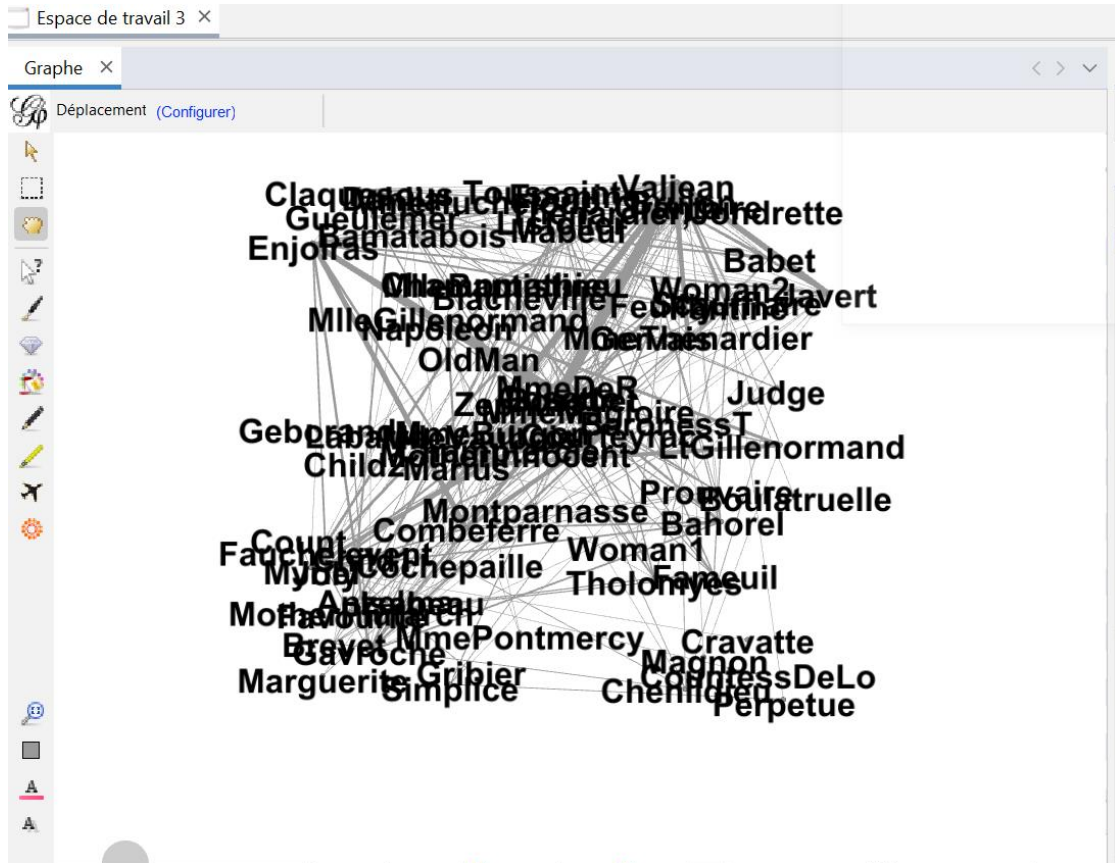
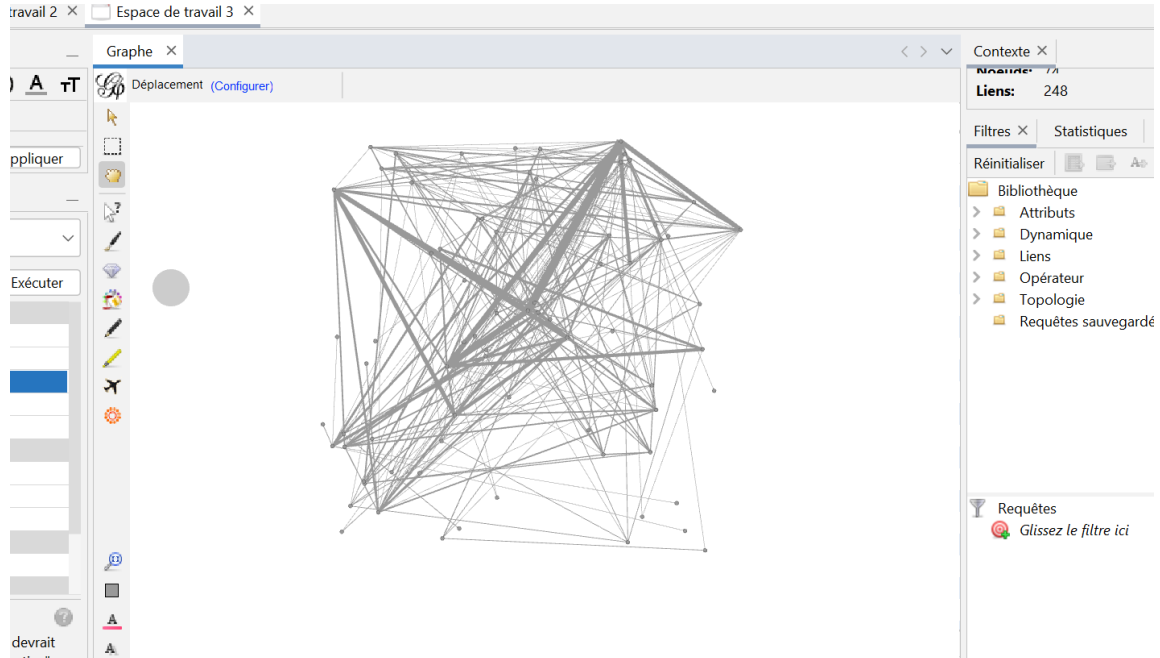
Supprimer la colonne

Effacer la colonne

Copier les données vers une colonne

Remplir avec u

Figure 2 dataset sur gephi



Après on va appliquer la spatialisation :

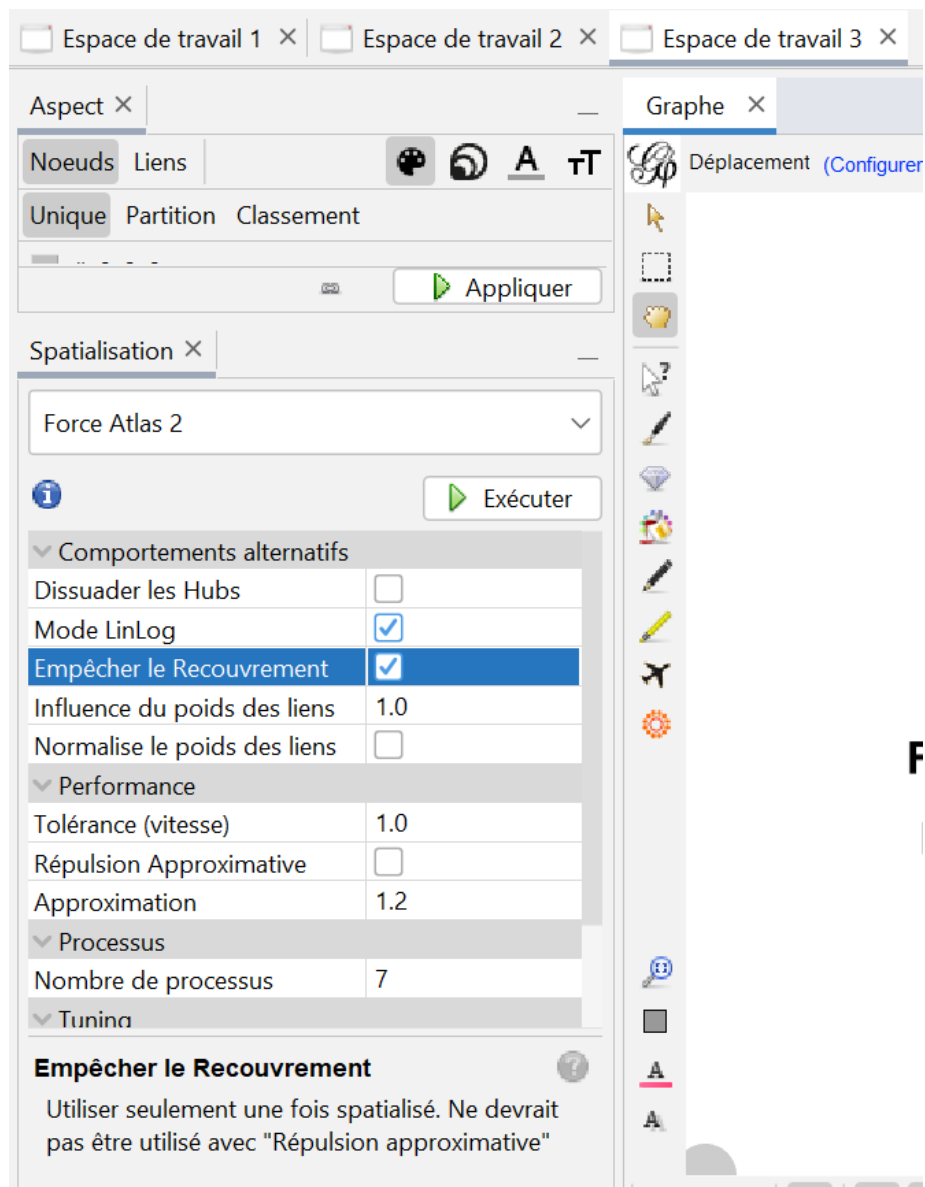


Figure 5 modification de quelques paramètres et lancement de la mise en page

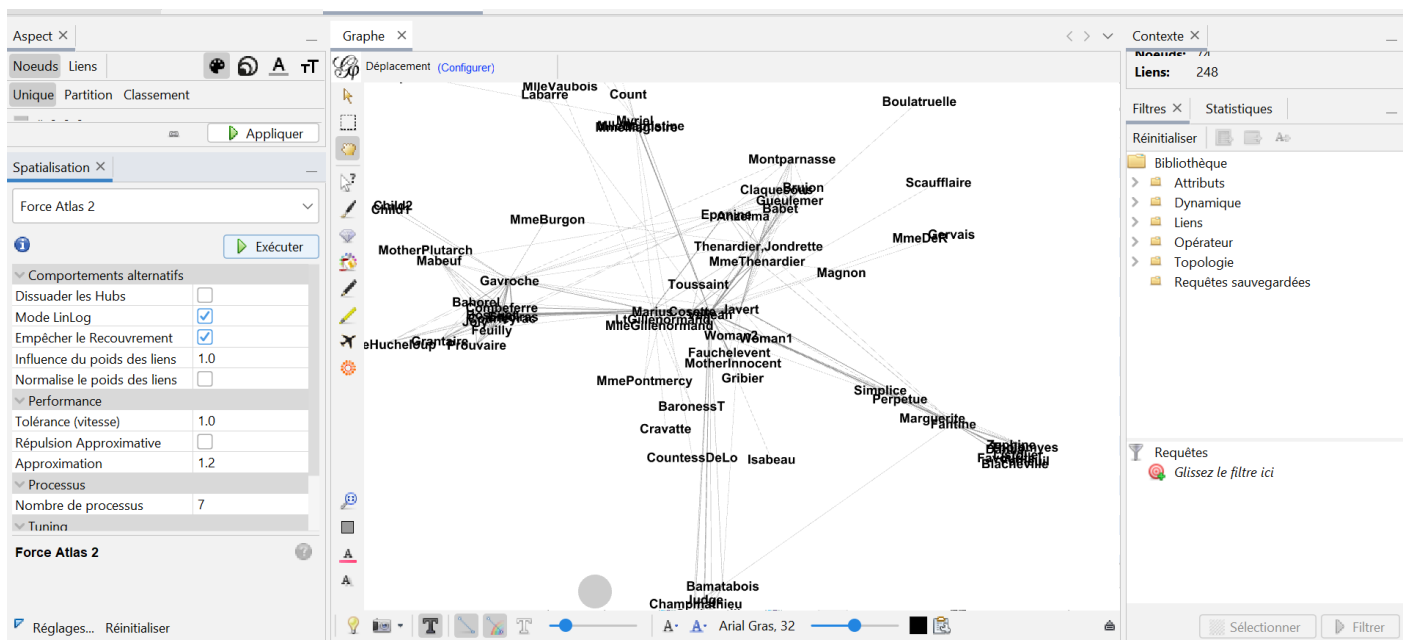


Figure 6 resultat de Force Atlas 2

La force atlas 2 permet de détecter facilement les communautés.

We will color the nodes based on their gender. To do that, we select **Gender** in the **Appearance** panel:

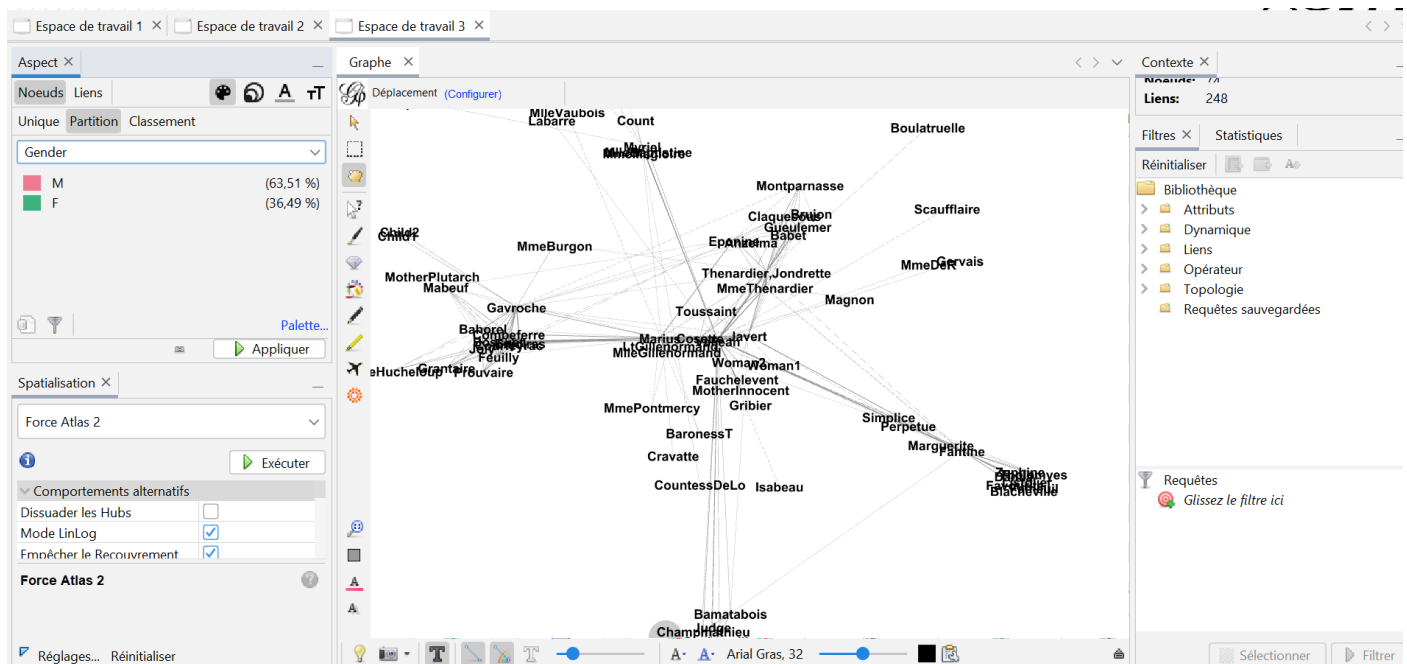


Figure 7 Nous colorerons les nœuds en fonction de leur sexe. Pour ce faire, nous sélectionnons Genre dans le panneau Apparence :

Le résultat est :

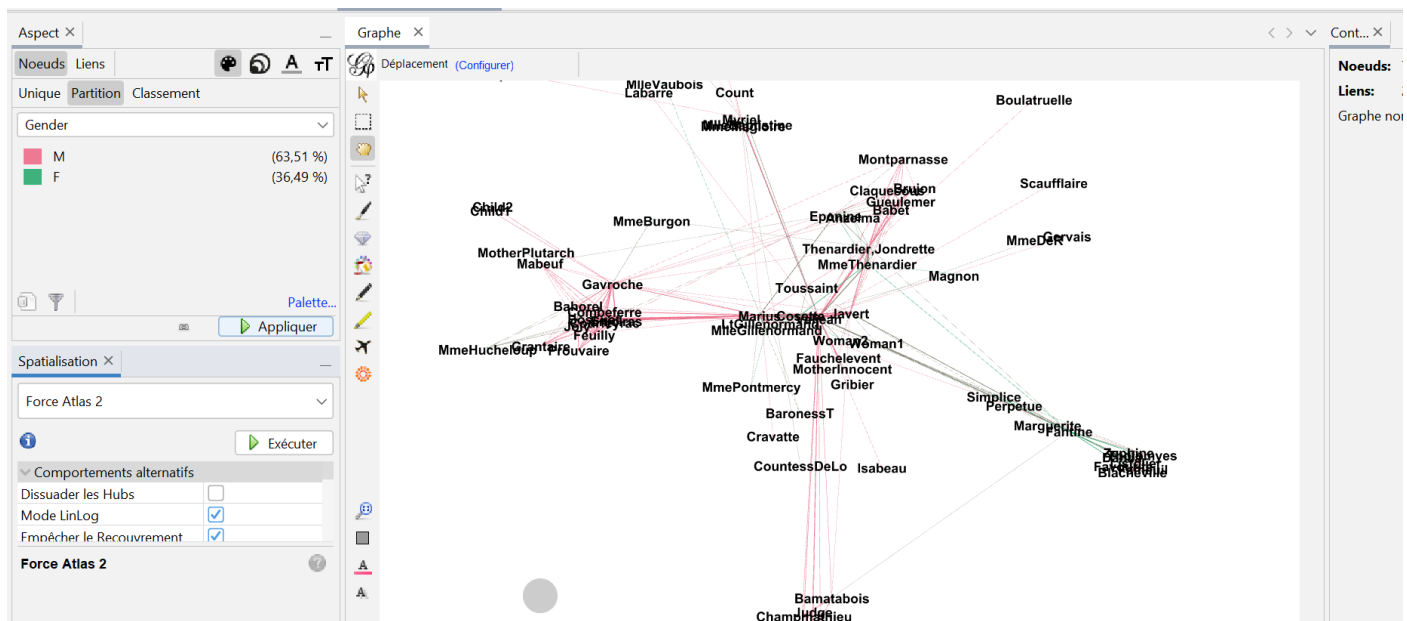


Figure 8 Après coloration des caractères selon leur sexe

## Amélioration de la lisibilité du réseau

Le résultat que nous obtenons soulève un certain nombre de questions :

1. Le réseau est trop grand ou trop petit, il est difficile à lire
2. Les étiquettes des caractères se chevauchent.
3. La taille des étiquettes peut être trop grande ou trop petite.
4. Les liens sont parfois trop grands.

Dans le panneau de mise en page, choisissez "Label Adjust" ou "Noverlap" : ces mises en page déplaceront les nœuds juste pour que les étiquettes cessent de se chevaucher :

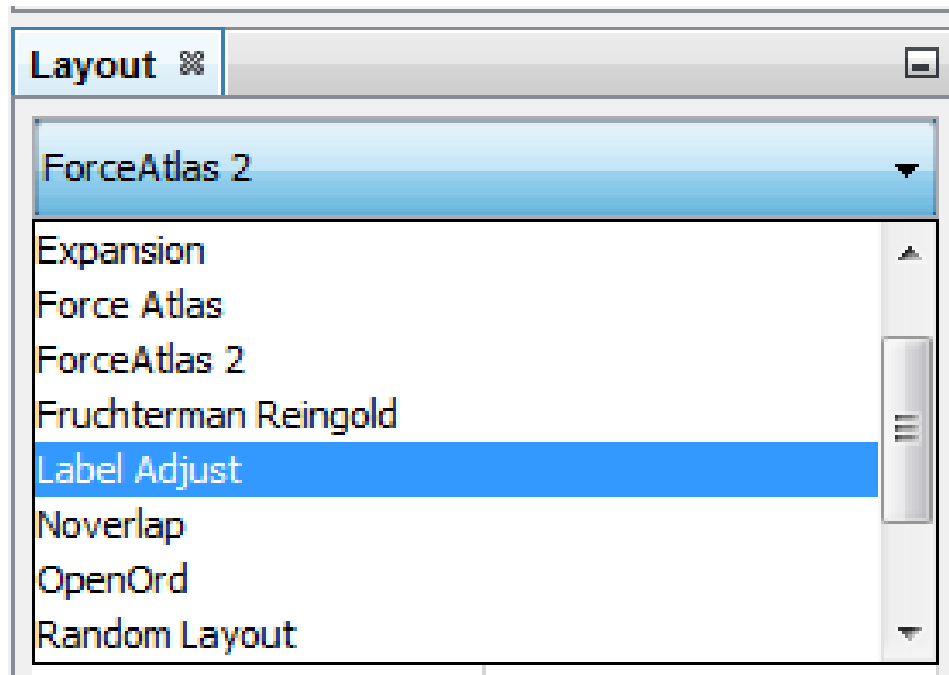


Figure 9 Noverlap ou Label Adjust

Dans le panneau inférieur de Gephi en cliquant sur la tête de flèche minuscule. Puis en sélectionnant "nœuds" et en déplaçant le curseur :

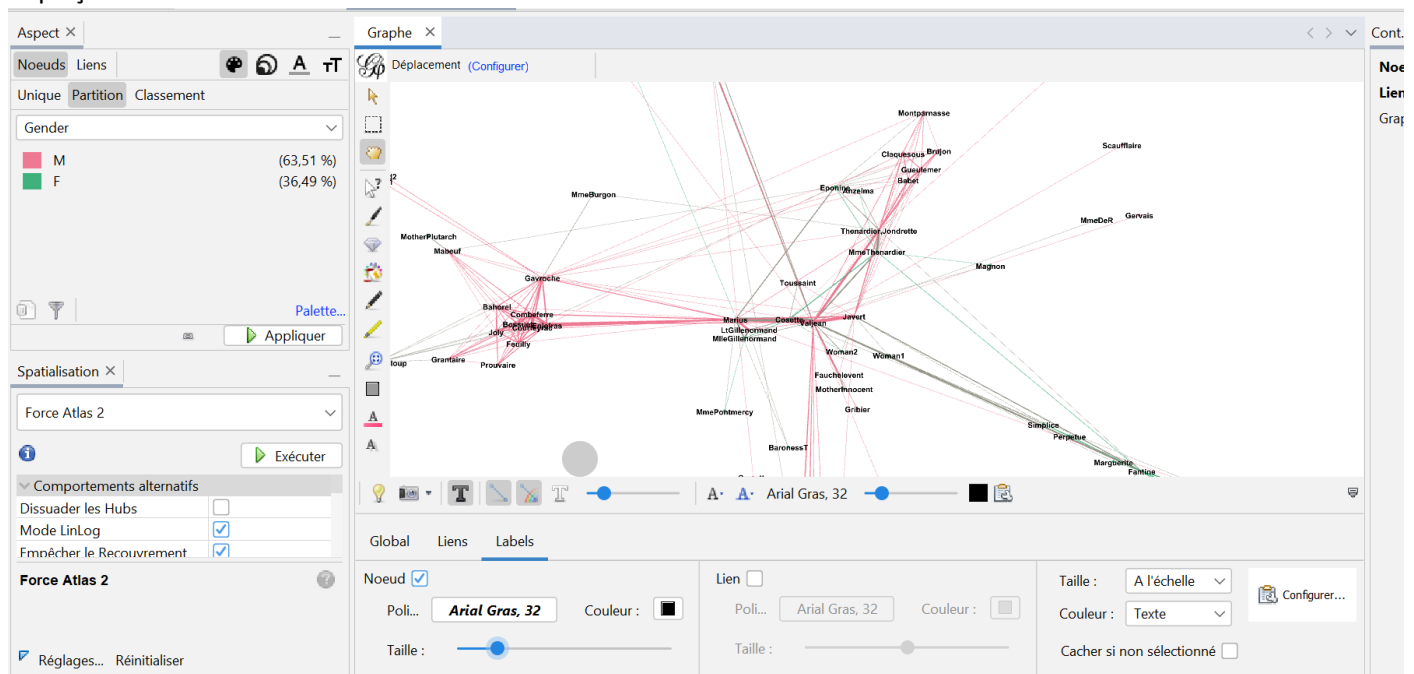


Figure 10 Ajustement des label

## Réglage de l'épaisseur des maillons

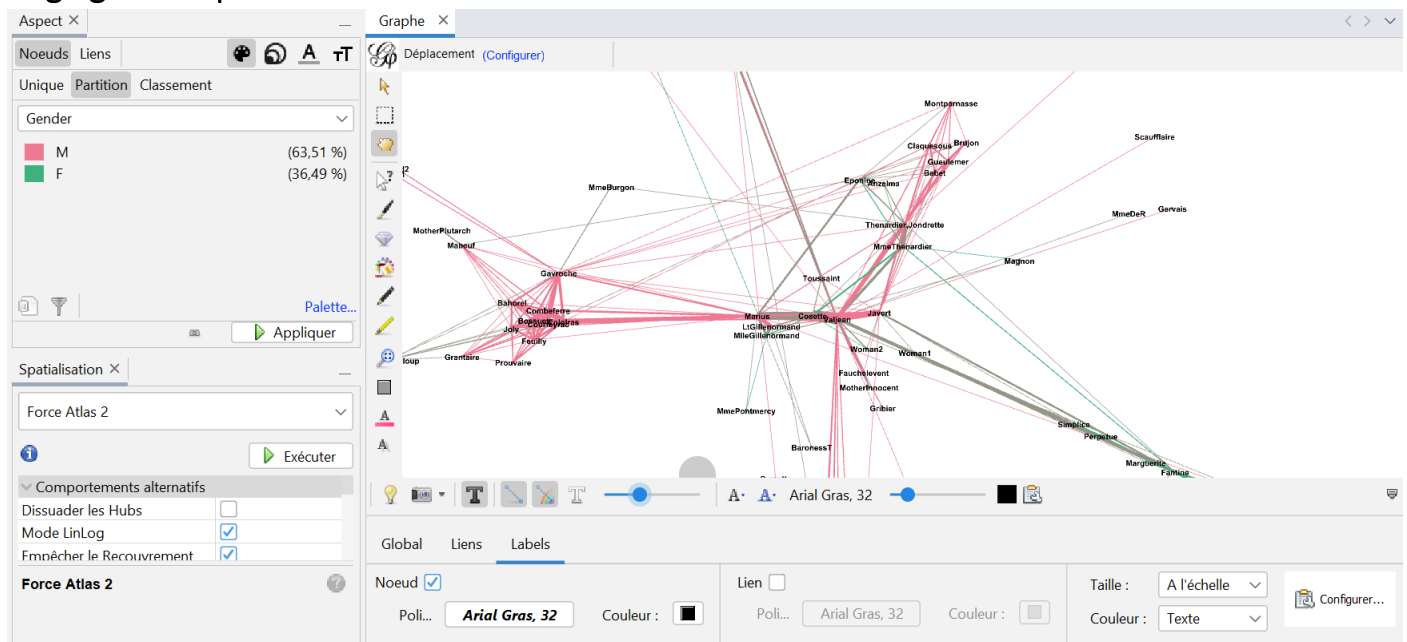


Figure 11 Ajustement de l'épaisseur

Pour redimensionner les nœuds en fonction de la valeur de leur centralité intermédiaire, nous utilisons le panneau Apparence :

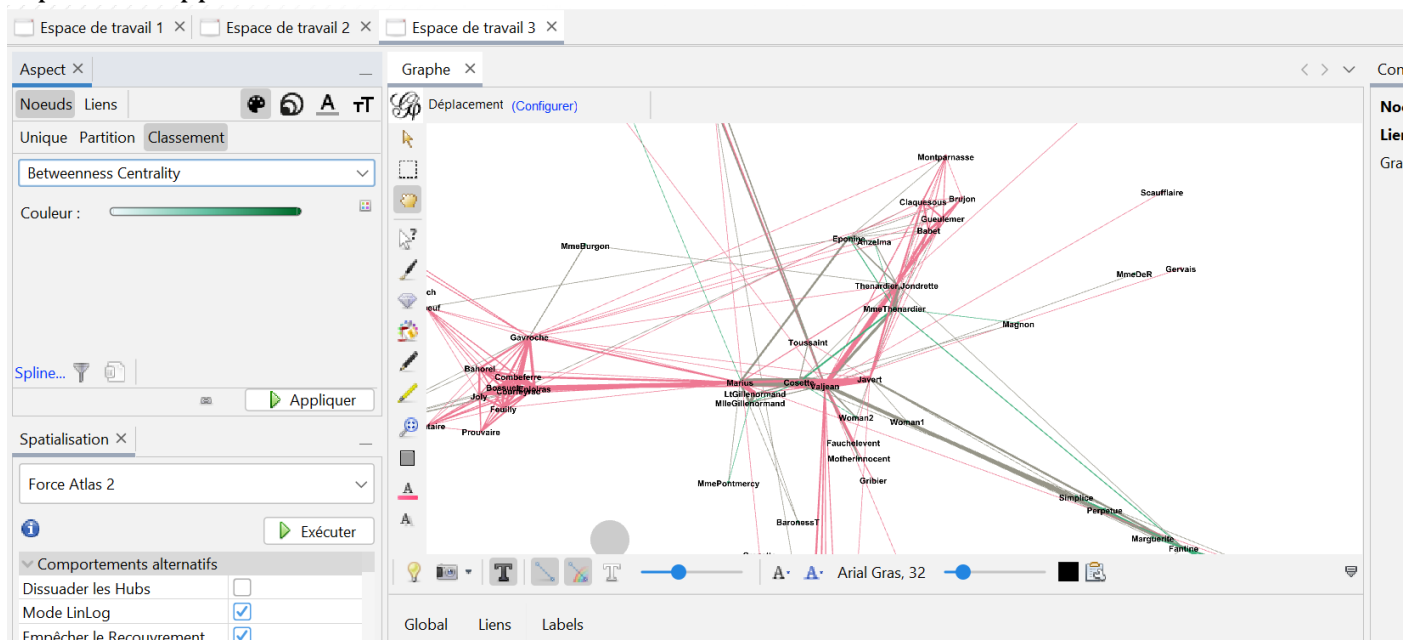


Figure 12 Classement des tailles de noeuds par centralité



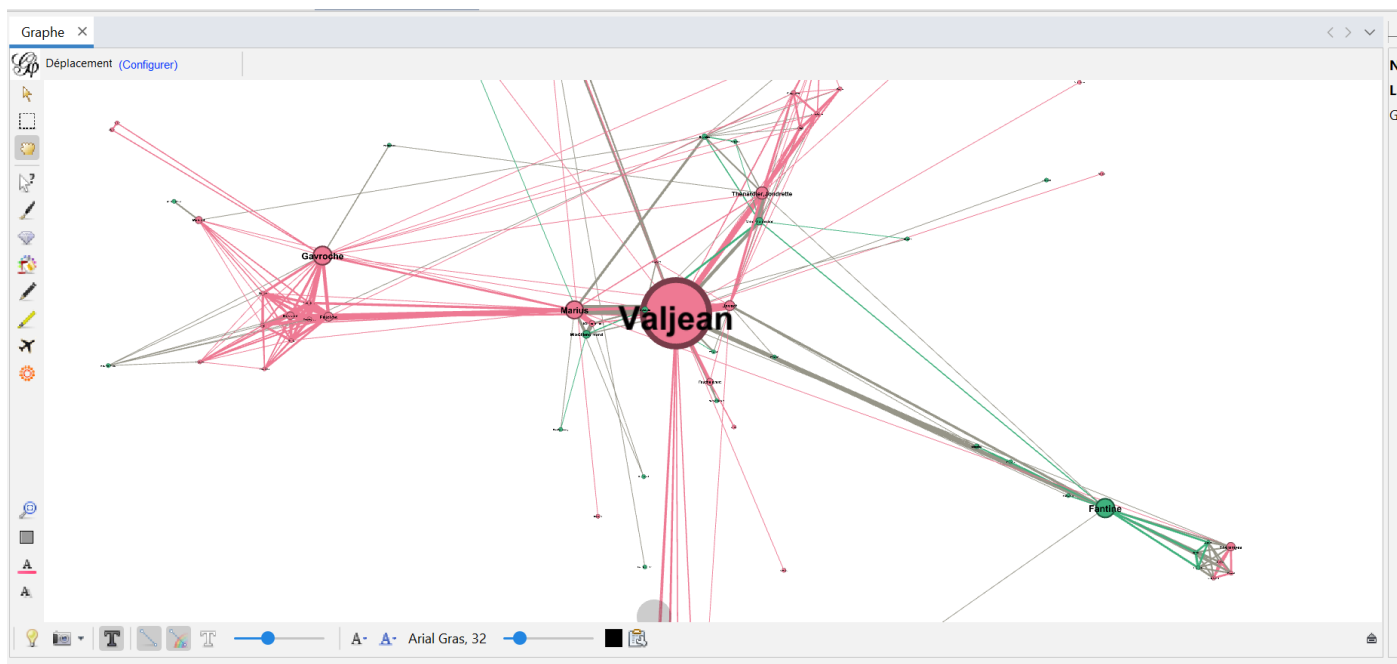


Figure 13 Résultat

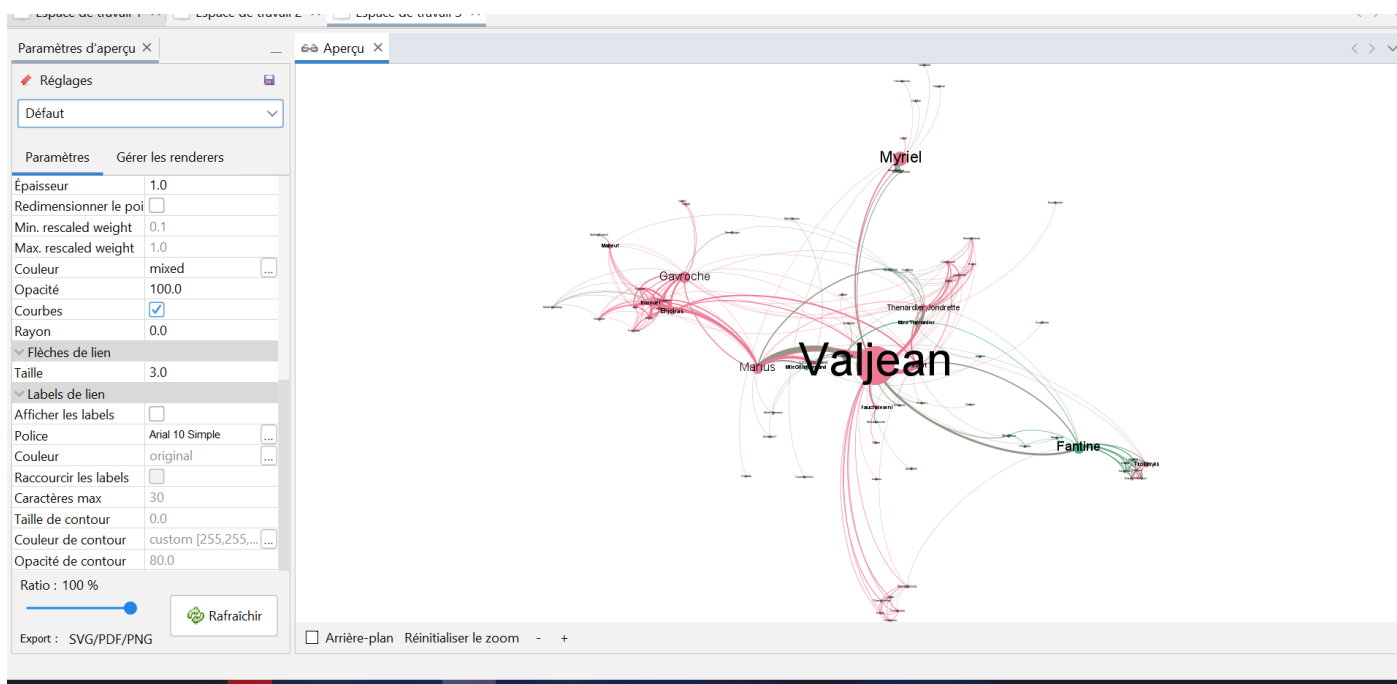


Figure 14 visualisation sur onglet "prévisualisation"

## Partage et diffusion

On va essayer d'exporter ce réseau comme visualisation web

La version web utilise VOSviewer en ligne, qui fonctionne mieux pour les réseaux jusqu'à ~ 500 nœuds

- 1- On va essayer d'exporter le fichier sous format . *gexf*

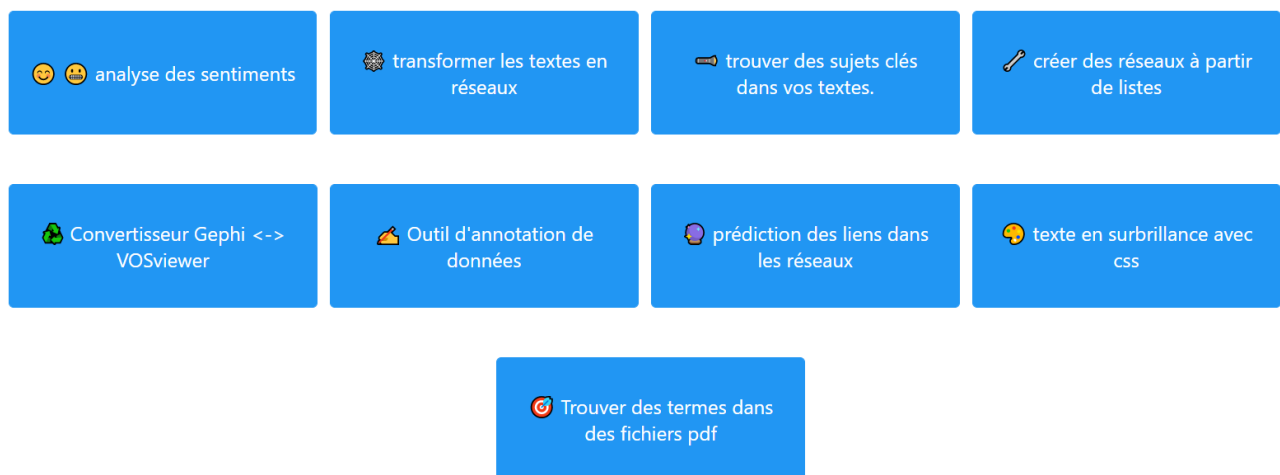


Figure 15 interface du site nocodefunctions

Choisissez un fichier à convertir ou lisez ce qui suit sur Gephi et VOSviewer.

☐ Option 1 : convertir un fichier Gephi gexf en un fichier json de VOSviewer [et éventuellement, partager immédiatement le fichier en tant que visualisation w  
☒ Option 2 : convertir un fichier json de VOSviewer en un fichier Gephi (format .gexf)  
☐

1 Choisir un fichier gexf (Gephi) Choisir un fichier Aucun fichier choisi

⚡ Le fichier doit être un **fichier .gexf**, pas un **fichier .gephi**! [Voir tous les détails ici](#)  
 📢 Le réseau doit être composé de 500 nœuds ou moins - sinon la visualisation en ligne sera impossible à lire. Voir ci-dessous pour plus de détails.  
 📄 Taille maximale du fichier: 100Mb

2 Convertir en fichier json de VOSviewer

☐ Rendre la visualisation de VOSviewer partageable sur le web avec une url publique ?

VOSviewer

The screenshot shows the VOSviewer web interface. At the top, there's a browser address bar with 'nocodefunctions.com/html/vosviewer/index.html'. Below it, a file explorer shows various folders like 'startups', 'Nouveau dossier', 'reporting', etc. The main area displays a network graph with nodes labeled with names like 'Geborand', 'Napoleon', 'Count Myriel', 'Boulatruelle', 'Montparnasse', 'Brulon', 'Eponine', 'Labarre', 'Toussaint', 'Marius', 'Valjean', 'Magnon', 'Scauffaire', 'Gervais', 'Perpetue', 'Tholomyes', 'Bamatabois', 'Brevet', 'CountessDeLo', 'BaronessT', 'MmePontmercy', 'Fauchelevent', 'Gribier', 'MmeDeR', 'Bahorel', 'Courfeyrac', 'MotherPlutarch', 'Child1', and 'MmeBurgon'. The nodes are connected by lines representing edges. On the right side, there are zoom controls (+, -, reset, and info icons). At the bottom left, a status bar shows ': 74 | : 248 | : 813 | Clusters: 11'. At the bottom right, there's a small circular icon with the number 22.

© 2015 Pearson Education, Inc. or its affiliate(s). All rights reserved. Pearson Education, Inc., publishing as Pearson Benjamin Cummings, 101 Philip Drive, Assinippi Park, New York, NY 10964-2133

