## Importing csv data in Gephi

## **Table of Contents**

Gephi workshops	1
Importing an existing network file	1
Importing nodes with labels and their relations: simplest version	2
Importing more than labels: nodes and edges attributes	7
Memo card 1	0
(to be continued)	1
More tutorials on importing data to Gephi	1
to go further	1

last modified: 2023-04-10



## Gephi workshops

I organize online workshops and personalized trainings for Gephi, for beginners and experts. To schedule one or to get more information: analysis@exploreyourdata.com.

## Importing an existing network file

Gephi handles the import of network files in a variety of formats:

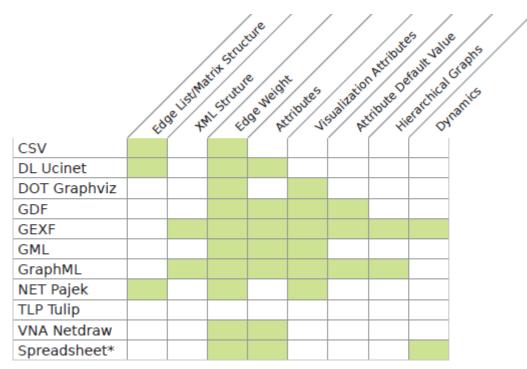


Figure 1. file formats supported by Gephi. Source: gephi.org

To import any of these files, just choose the File → Open menu

## Importing nodes with labels and their relations: simplest version

(looking for the complete, full featured csv import version? skip to here.)

The easiest is probably to just have a text file like this one:

```
Source, Target ①
Jeremy, Jennifer
Valerian, Jeremy
```

① This line is mandatory. Don't put a space between the comma and "Target"

Just write this in a text editor and save it as a file example.txt

To import this file in Gephi, go to File→Import Spreadsheet. Then in the window:

M Import spreadsheet		x						
Steps	General options							
<ol> <li>General options</li> <li>Import settings</li> </ol>								
	cs\src\main\asciidoc\resources\import-tutorial\edges_alone.txt							
	Separator:	As table: Charset:						
	Comma 👻 Edges ta	able 🔹 🗸 UTF-8 🔹						
	Preview:							
	Source	Target						
	Jeremy	Jennifer						
	Valerian	Jeremy						
2. Then click Next	1. Make sure "E	Enish Cancel Help						

Figure 2. Importing nodes with labels and their relations

In this window, make sure the box "Create missing nodes" is checked:

Import spreadsheet	×
Steps	Import settings
<ol> <li>General options</li> <li>Import settings</li> </ol>	New columns are created with the specified type. A generated id is assigned if missing or already existing. Edges need 'Source' and 'Target' columns with the id of the source and target If no 'Type' column is provided, all edges will be directed. If an edge already exists, attributes will be ignored, but their weights will be a Imported columns: Source String Target String Target Target Create missing nodes
	2. Then click Finish

Figure 3. Second screen

Let's look again at the simple text file we used:

Source,Target Jeremy,Jennifer Valerian,Jeremy

There are 2 issues and several missing features, listed below:

### 1. labels containing commas in them will not work.

Imagine our characters have first names and last names separated by a comma:

Dubois, Jeremy

Rodriguez, Valerian

This will look like:

Source, Target

Dubois, Jeremy,Jeremy,Jennifer Rodriguez, Valerian,Dubois, Jeremy

Note that we have 3 commas per line, instead of one! In this mess, Gephi will not detect where the node labels start and end. The import will break.

The proper solution is to put node labels inside **double quotes**:

Source,Target "Dubois, Jeremy","Jeremy,Jennifer" "Rodriguez, Valerian","Dubois, Jeremy"

Be careful to use these straight " " double quotes, not the curvy ones: " " which French keyboards sometimes have.

### 2. labels containing double quotes " " in them will not work.

Example: let's imagine that one of our characters has a middle name: Jeremy "Danger" Dubois. Our text file will look like:

Source,Target "Dubois, Jeremy "Danger"","Jeremy,Jennifer" "Rodriguez, Valerian","Dubois, Jeremy "Danger""

These extra " " will make the Gephi import break:

M Import spreadsheet			×
Steps	General options		
<ol> <li>General options</li> <li>Import settings</li> </ol>	Choose a CSV file to import		
	cs\src\main\asciidoc\resou	rces\import-tutorial\edge	es_alone.txt
	Separator:	As table:	Charset:
	Comma 👻 Edge	s table 🛛 👻 🛛 UT	F-8 👻
	Preview:		
	Source	Target	
	Dubois, Jeremy	Jeremy, Jenni	
	Rodriguez, Valerian	Dubois, Jeren	ny
	< Back Next	> Finish	Cancel Help

Figure 4. The import bugs - the middle name has disappeared

The solution consists in **adding an extra double quote in front of the double quotes**. Our text file will look like:

```
Source,Target
"Dubois, Jeremy ""Danger""","Jeremy,Jennifer"
"Rodriguez, Valerian","Dubois, Jeremy ""Danger"""
```

This time, Gephi imports the network correctly:

M Import spreadsheet					X
Steps	General options				
<ol> <li>General options</li> <li>Import settings</li> </ol>	Choose a CSV file to import: cs\src\main\asciidoc\resources\import-tutorial\edges_alone.txt				
	Separator:	As table		Charset:	
	Comma 👻	Edges table	▼ UTF	-8	•
	Preview:				
	Source		Target		
	Dubois, Jeremy "Dar Rodriguez, Valerian		Jeremy, Jennife Dubois, Jeremy		_
	Rounguez, valenan		Dubois, Jerein	/ Dangei	
	< Back	Next >	Finish	Cancel He	lp
L					

Figure 5. Importing node labels containing double quotes

# Importing more than labels: nodes and edges attributes

To import attributes we will need to proceed differently.

We need 2 text files: one for the list of nodes, one for the list of relations (edges)

An example file with a list of nodes:

```
Id,Label,Date of Birth,Place of Birth,Years of experience,Rating ①
3,"Dubois, Jeremy ""Danger""",17/09/1980,"Paris",8,9.27
1,"Jeremy,Jennifer",25/03/1978,"Tampa",8,4.34
45,"Rodriguez, Valerian",30/04/1985,"Berlin",5,6.66
```

① Nodes **must** have at least an Id and a Label. Don't put spaces after the commas

An example file with a list of edges:

Source,Target,Weight,Type,Where first met (1)

1,45,3,"undirected",London ②

① Edges **must** have at least a Source and Target. Other fields are optional.

② "undirected", the alternative is "directed". Directed edges have arrow heads.

Let's import the list of nodes first. File  $\rightarrow$  Import Spreadsheet

🕵 In	nport spreadsheet						×
Ste	ps	General options					
<b>1.</b> 2.	General options Import settings	Choose a CSV file	to import:				
		C:\Users\Jevallois	(Google Drive \open \gept	hi tutorials\docs\sr	c (main \asciidoc \resour	rces\import-tutorial\node	s.bd
		Separator:	As table:			Charset:	
		Comma 🗸	Nodes table	▼ UTF-8			•
		Preview:	1				
		Id	pel	Date of Birth	Place of Birth	Years of experie	Rating
		3	bois, Jeremy "D	17/09/1980	Paris	8	9.27
		1	eremy, Jennifer	25/03/1978	Татра	8	4.34
		45	Rodriguez, Valerian	30/04/1985	Berlin	5	6.66
		Make	sure "Node	es table"	' is selecte	ed	
					< gack Next >	Enish (	Cancel Help

Figure 6. Importing a list of nodes with attributes

in the next screen, we must be careful with a couple of things:

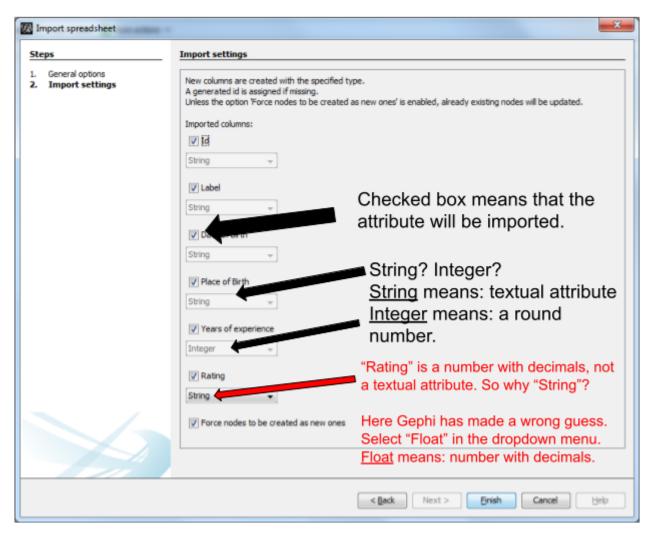


Figure 7. The attributes of the nodes

Then we can import the file with the list of relations. File → Import Spreadsheet

Steps	General options				
General options     Import settings	Choose a CSV file	to import:			
	s\/evallois\/Google	Drive \open \gephi tutor	ials \docs \src \main \a	sciidoc\resources\import	t-tutorial/edges.tx
	Separator:	As table:		Charset:	
	Comma 👻	Edges table	▼ UTF-8		•
	Preview:	1			
	Source	Talet	Weight	Type	Where first met
	1	45	3	undirected	London
	Make	sure "Edges	s table" is s	selected	

Figure 8. Importing a list of edges with attributes

teps	Import settings	
General options Import settings	If no 'Type' column is provided,	sing or already existing. I' columns with the id of the source and target nodes. If any is not provided for a row, it will be ignored.
	Imported columns:	Double means:
	String +	a number with decimals
	⊡ Target String w	(so, a bit like Float but with even more decimals)
	Double -	Gephi has correctly identified the attribute "Where first met"
	String v	as textual (String)
	Where frat net String	In our file, we have just one edge: between nodes with id 1 and id 45. We have already imported
	Create missing nodes	these nodes. If we had not, we could have created them by ticking this box.

Figure 9. The attributes of the edges

### Memo card

#### Nodes

• header must be at least Id, Label

#### Edges

- header must be at least Source, Target
- To have an attribute for the "thickness" or "strength" of a relation, this attribute must be called "Weight".
- Want arrows on your links? Add an attribute "Type", with value "Directed"
- Don't wan't arrows? Add an attribute "Type", with value "Undirected"

Types of attributes: which to choose?

Vhere first met	
String	•
BigDecimal	*
BigInteger	
Boolean	
Byte	
Character	
Double	
Float	
Integer	$\overline{\mathbf{v}}$

Figure 10. Kinds of attributes

- Textual attribute: String. Nodes sharing the same textual value can be colored the same, or filtered together...
- Numerical attribute: Integer, Double or Float. Nodes can be resized according to their value, or colorized in a gradient. Filters can be applied based on the range of values.
- Boolean attribute: just a true or false value. Useful for filtering out some nodes which are true or false on some parameter.
- Other types of attributes: not needed for the moment.

## (to be continued)

### More tutorials on importing data to Gephi

- The Gephi wiki on importing csv
- Video "How to import a CSV into Gephi" by Jen Golbeck

## to go further

Visit the Gephi group on Facebook to get help,

or visit the website for more tutorials

Give a try to nocodefunctions.com, the web application I develop to create networks for Gephi. Click-and-point, free, no registration needed.