Creating a network from a table of cooccurring items

Table of Contents

Gephi workshops
Update
Presentation of the plugin
Installing the plugin
Opening the plugin
Using the plugin
to go further

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.

This plugin is created by Clement Levallois.

Update



This plugin has been replaced by a new and better function available on the web, developed by the same author. It is free, click and point, without registration at this address: https://nocodefunctions.com

Presentation of the plugin

This plugin is created by Clement Levallois.

It converts a spreadsheet or a csv file into a network.

This plugin enables you to:

- Start from a data table in Excel or csv format
- In the data table, each row describes an "occurrence" (of an event, a purchase, a relation, etc.)
- In columns A, B, C, D, we have the entities involved: column A for persons, column B for what they bought, etc.
- Connections will be created between entities, when they appear in the same occurrence (so, when they are on thee same row)
- Occurrences can have dates, multiple instances of an entity can be listed in a given column.

1. The input

_	1	
1	А	В
1	Client	Purchases
2	Α	sofa bed, chairs, table, dressing, toys, diapers
3	В	toys, diapers, powder milk, babyshoes
4	С	beer, snacks, vodka, pizza, sofa bed, chairs, table
5	D	cereals,toys,table,chairs
6	E	cleaning product, diapers, toys
7	F	powder milk, diapers
8	G	beer, snacks
9	н	pizza, beer
10	I.	table,chair
11		

Figure 1. An Excel file

2. The output

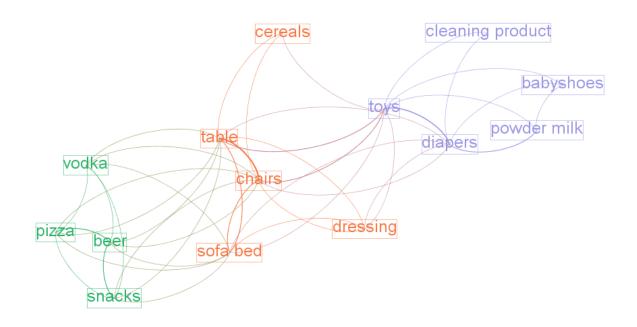


Figure 2. Resulting network

Installing the plugin

Gephi 0.9.1		
File Workspace Tools Window Help		
 Dvarve Plugins 	Y 💭 Preview	U Xá
Options		~ 0
Appearance # Language	Graph = # 🗗 Web Browser = #	ALL Context =
Nodes: Edges: Online docs and support		liodes
		Edges:
		Fil. = Statist
		Reset 📑 🖬 An
	e Tools Window Help	
	Vit Plugins y	
	Options	
	I Language	
	Language	
Layout #	Online docs and support	
Choose a layout		
0		
•		
<no pro<="" td=""><td>roperfact-</td><td></td></no>	roperfact-	
	·	
P Presets Reset] 🕈 🗰 + T \ \ \ \ T A+ A+ Aniston, 32 ■®	a Fra

Figure 3. Choose the menu Tools then Plugins

Check	for Updates		Search:
Update	Name	Category	Algorithms
1	Algorithms	Gephi Core	Agonums
	Statistics API	Gephi Core	Installed version: 0.9.1
1	Appearance API	Gephi Core	Available version: 0.9.1.1
1	Data Laboratory API	Gephi Core	Author: Gephi Consortium
1	Filters Impl	Gephi Core	Date: 12/30/16
× × × × × ×	Visualization Module	Gephi Core	Source: Gephi Update Center
1	Desktop Context	Gephi UI	Homepages http://gephi.org/modules/algorithms-plugin
1	Tools Plugin	Plugin	
1	Statistics Plugin	Plugin	Plugin Description
X	Data Laboratory Plugin	Plugn	Basic graph theory algorithms

Figure 4. Click on the tab Available Plugins

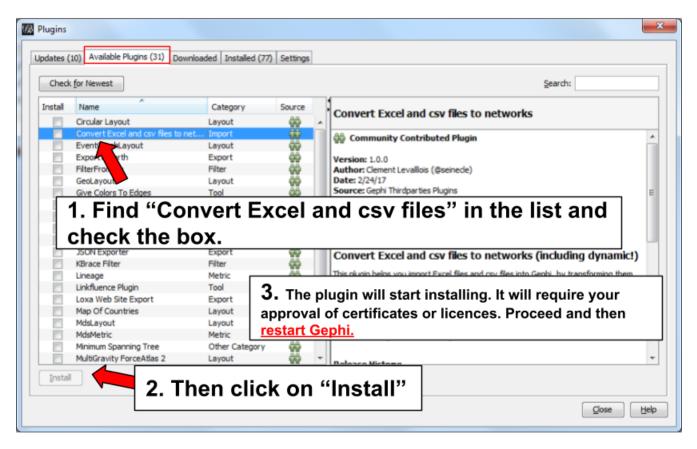


Figure 5. Install the plugin Convert Excel and Csv files and restart Gephi

Opening the plugin

Gephi 0.9.1 - Project 1	And AT STARLEY SAME AND	Come of March 1 and 1 and 1	and all Blance at Blance & Manuel & Manuel	
File Words Window Plugins Help				
Ctrl+Shift+N aboratory Preview Ctrl+O				XXII.
Open mark			a a	Context N
Close Project				Nodes: 0
Properties				Edges: 0
Import spreadsheet				Directed Graph
Import Database	Import Wizard		×	
Import	Import Wizard			Iters N Statistics
Generate /				eset 📑 🖶 Ao
	Steps	Select Wizard (1. from 2)		Library
save As (2)	1. Select Wizard			- Attributes
	2	Category:	Wizard Type:	- Dynamic
Export ,	£			- Edges
Exit		Data importer (co-occurrences)	Convert Excel and csv files to networks	- Topology
				- a Saved queries
13				
15				
		(3)		
= Apply				
Layout =				
1				
Choose a layout				P Queries
Run				- 😡 Drag filter here
P Mar				
		Description:		
		This plugin helps you import Excel or csv Feedback and feature requests are welc	files into Gephi, based on entities which co-occur line by line.	
		contact: @seinede on Twitter.	come!	
		contact: gsenece on I witter.		
<no properties=""></no>				
		1		
		1		
		(4)	
			4)	
<i>p</i>		< Back	Next > Einish Cancel Help	
•		< Back		
<u>A</u>				
A.				
🏴 Presets Reset 🛛 🦞 📾 📲 🏋 💽	A· A· Arial Bold, 32	0	<u>ه</u>	Filter
👩 🗽 🛅 Equila- 🔑 👹 🛞 📉 Macrael 🗰			📄 Spotify 🚺 IBM S 🚺 Slack 🥋 impor 🌠	FR . ** 64
🚯 🐪 🚞 Epilo 🛃 😴 🐼 🌇 Micr 🕨	🛛 💁 Inbox 🚯 🗊 gephi 🧿	аани 🎲 🌺 🕺 🖸 📝	🛛 🛜 Spotify 🕠 IBM S 🚺 Slack 🛞 impor 🕼) Gephi 🔤 🙀 🔺 🖏 🙀 📶 🕩 🍐 23/02/2017

Figure 6. Open the plugin via the menu File - Import

Using the plugin

2nd panel

Import Wizard					
Steps 1. Select Wizard 2. Select Excel or csv file 3. Select agents 4. Subfields in agents 5. Dynamic network? 6. Options	Select Excel or csv file (2. from 7) select file Image: select file				
7. Ready to import					
	<pre></pre>	-			

Figure 7. Select a file

	А	В
1	Α	sofa bed, chairs, table, dressing, toys, diapers
2	В	toys, diapers, powder milk, babyshoes
3	С	beer, snacks, vodka, pizza, sofa bed, chairs, table
4	D	cereals,toys,table,chairs
5	E	cleaning product, diapers, toys
6	F	powder milk,diapers
7	G	beer, snacks
8	н	pizza, beer
9	I.	table, chairs
10		

Figure 8. A file without headers

	А	В
1	Client	Purchases
2	Α	sofa bed, chairs, table, dressing, toys, diapers
3	В	toys, diapers, powder milk, babyshoes
4	С	beer, snacks, vodka, pizza, sofa bed, chairs, table
5	D	cereals,toys,table,chairs
6	E	cleaning product, diapers, toys
7	F	powder milk, diapers
8	G	beer, snacks
9	н	pizza, beer
10	I	table,chair
11		

Figure 9. A file with headers

To describe the next screens of the plugin, we will take the example of **the Excel file just shown**, with headers.

3rd panel

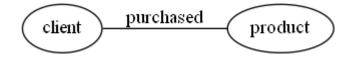
M Import Wizard	
Steps	Select agents (3. from 7)
 Select Wizard Select Excel or csv file Select agents Subfields in agents Dynamic network? Options Ready to import 	What are the connections made of? 1. This type of agent: Client Purchases
	is connected to
	2. This type of agent:
	Client Purchases
	Example: for a network of co-authors, pick the field of co-authors twice!
	< <u>B</u> ack Next > Einish Cancel Help

Figure 10. Which entities should be the nodes?

What does this panel mean?

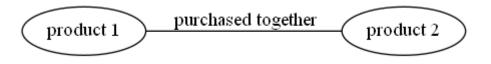
If you look back at the Excel file, you see that we have "Clients" and their "Purchases".

- \rightarrow This means we can build 2 different types of networks, depending on our needs:
- 1. A network showing clients and products, with relations representing purchases from a client to a product.



To create this kind of networks, choose "Client" in the upper window, and "Purchases" in the lower window of the plugin screen.

2. Or a network where 2 products are connected, if one client puchased them together.



To create this kind of networks, choose "Purchases" in the upper <u>and</u> lower windows of the plugin screen.

4th panel

M Import Wizard		x
Steps	Subfields in agents (4. from 7)	
 Select Wizard Select Excel or csv file Select agents Subfields in agents Dynamic network? Options Ready to import 	Purchases Is this field made of subfields? If so, what is the delimiter?	
	Example: Let's imagine you selected "authors" for agents in the previous screen. Suppose that the field for "authors" is made of several co-authors, separated by semicolon. => Please choose "semicolon" as a delimiter. (leave the selection empty if no delimiter applies).	
	< <u>B</u> ack Next > Einish Cancel	Help

Figure 11. Choosing which delimiter is used

This 3rd panel asks: in our Excel file, how are different items separated in a given cell? In our example, we have used commas: the lists of products purchased are comma-separated:

1	А	В
1	Α	sofa bed, shairs, table, dressing, toys, diapers
2	В	toys, diapers, powder milk, babyshoes
3	С	beer, snacks, vodka, pizza, sofa bed, chairs, table
4	D	cereals,toys,table,chairs
5	E	cleaning product, diapers, toys
6	F	powder milk,diapers
7	G	beer, snacks
8	н	pizza, beer
9	I	table, chairs
10		

Figure 12. commas shown in red

5th panel

This panel allows you to specify whether the relations are dynamic in time, or not.

In this case, you need an extra column (column C), where a date is shown. We don't cover this case here.

(read the tutorials on dynamic networks for a starter)

6th panel

M Import Wizard		x
Steps	Options (6. from 7)	
Select Wizard Select Excel or csv file Select agents Subfields in agents Dynamic network? Options Ready to import	Options (6. from 7) Ireate links between Purchases agents and links between Purchases agents. remove duplicates: lines that have exactly the same values for both types of agents will be removed remove self-loops: when an agent is connected to itself.	_
	< Back	

Figure 13. Options panel

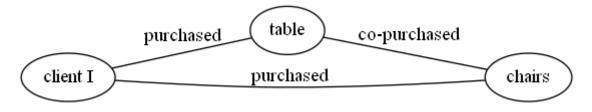
"Create links between Purchases agents and links between Purchase agents"

→ If you chose a Product \leftarrow → Product kind of network in panel 3, then of course you are interested in links between products. **Check the box**.

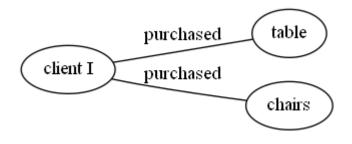
 \rightarrow But if you chose a Client $\leftarrow \rightarrow$ Product kind of network in panel 3, what you need is less obvious.

Let's take the example of client I, who purchased a table and some chairs:

1. Checking the box will create a network where:



1. Not checking the box will create a network where:



"Remove duplicates"

→ Check this option if your Excel or csv file has duplicate rows that you'd like to be removed

```
"Remove self-loops"
```

If a Client has purchased tables twice, so that we have "table, table" in a cell: this would create a link from table to table (a **self loop**).

 \rightarrow Check this option if you'd like self loops to be removed.

7th panel

This panel recaps all the settings. Click on finish to create the network.

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.