社會網絡分析工具
Social Network Analysis SNA

臺大圖書館學科服務組
范蔚敏
wmfan@ntu.edu.tw

學員普查

學習動機 & 先備知識

初學的痛苦/今天的磨難

1. 專有名詞
   (完全沒聽過!!)
2. 各類演算法
   (數學!!理解後可以解讀、畫出最適切的網絡圖與驗證)
3. 各種應用案例
   (好玩的地方!!)

課程目標

- 學習社會網絡分析軟體應用
- 軟體：Gephi
- 軟體操作
- 資料處理
- 數據解讀

→網絡分析的視覺化高手!
　不會寫程式也可以!

課程大綱

01 社會網絡分析實際應用
02 社會網絡分析基本概念
03 Gephi軟體入門─基本功能
04 Gephi軟體入門─Overview
05 Gephi軟體入門─Data Lab & Preview
06 實務操作─資料匯入
07 實務操作─Layout
08 Gephi Statistic
09 Gephi 篩選Filter

01社會網絡特色與應用
The features & practical application

社會(科學)網絡分析源起

- 緣起：從圖形理論(Graph Theory)發展的量化技巧
  - 量化演算法
  - 圖形視覺化

- 其他學科怎麼看
  - 數學領域 vertex arc
  - 電腦科學 node link
  - 社會科學 actor relation

社會網絡分析特性

和傳統研究方法比較

比較	研究問題	資料蒐集	資料分析	分析項目
傳統量化分析	體重、心血管疾病	公斤、血壓	量化統計檢定	T檢定或相關係數
社會網絡分析	學術研究共著	作者、文章	社會網絡分析	重要的作者、領域分析

差異

① 可以將隱性知識視覺化，EX：關係、合作經驗
② 不受量化統計母體樣本的限制
③ 利用社會網絡分析演算法探究研究面向

A network of science: 150 years of Nature papers

社會網絡分析面向

1) 整體社會網絡的全貌 EX小世界網絡
2) 節點與節點彼此的連結，來定義節點屬性 EX權威性
3) 依據節點屬性分群其子群代表的意義 EX學術領域分析
4) 子群之間彼此的關係 EX交流密切與否
5) 子群內節點的情形分析 EX子群凝聚力，各子群內使用者特性
6) 子群結構的對應 EX使用者相關推薦

社會網絡分析應用
金曲獎最佳男女歌手提名人創作角色分析

圖 2　金曲獎最佳國語男女歌手獎之提名音樂工作者社群樣貌

范蔚敏、唐牧群（2019）

金曲獎最佳男女歌手提名人創作角色分析

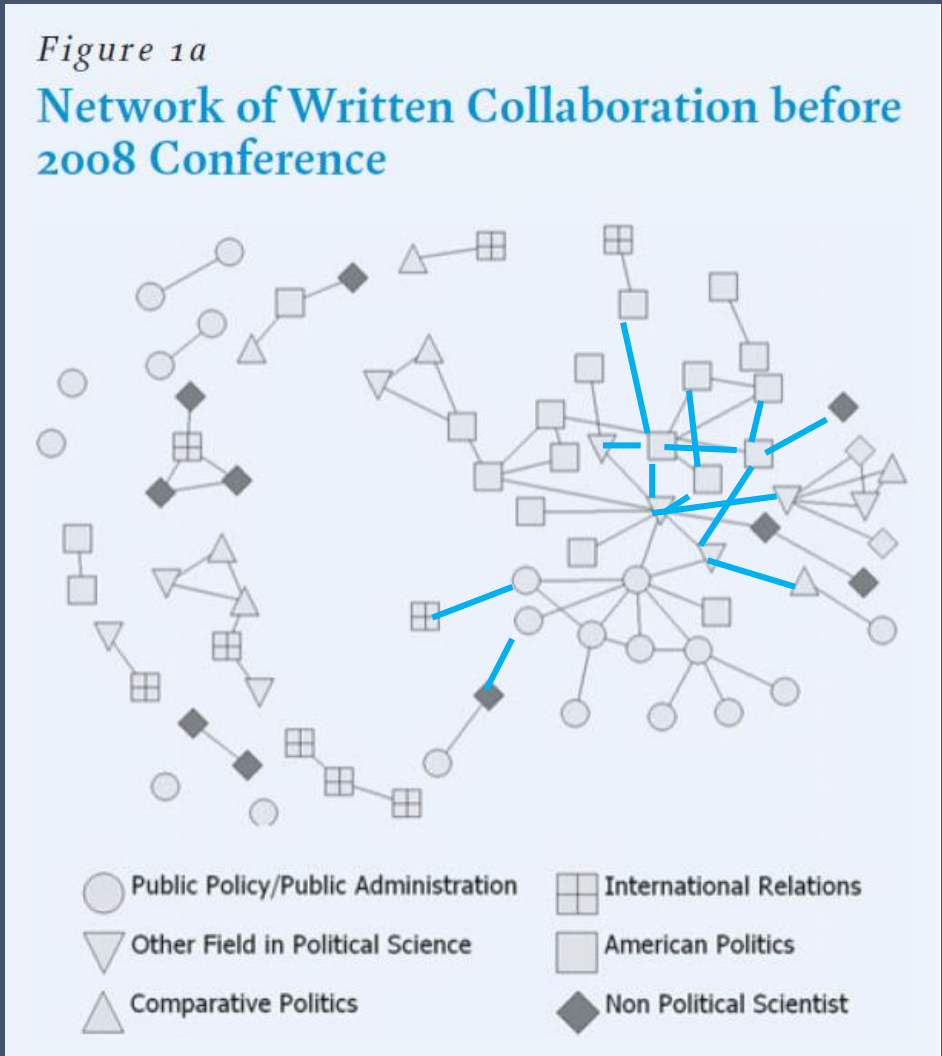
關係形式：歌一創作人

1.社群網絡具有相當規模，且社群成員連結程度高，但是有別於小世界網絡現象，整體網絡平均群聚係數數值相對較低，顯示音樂工作者彼此之間合作對象多元，且並未偏好與特定音樂創作人合作之情形。

2.每位音樂工作者合作程度進行分群，顯示整體網絡有明顯分群，且各分群社群網絡凝聚程度較高。

范蔚敏、唐牧群（2019）。第18屆到第29屆臺灣金曲獎最佳國語男女歌手提名人創作角色分析。數位典藏與數位人文，(3)，43-62。doi:10.6853/DADH.201904_3.0003

社會網絡分析應用
學術研究者合作情形



政治科學領域學者參與研討會對於新社群成員協作的模式影響

1. 資料形式：期刊文獻、問卷調查
2. 研究問題：研討會

→研究者共同協作論文

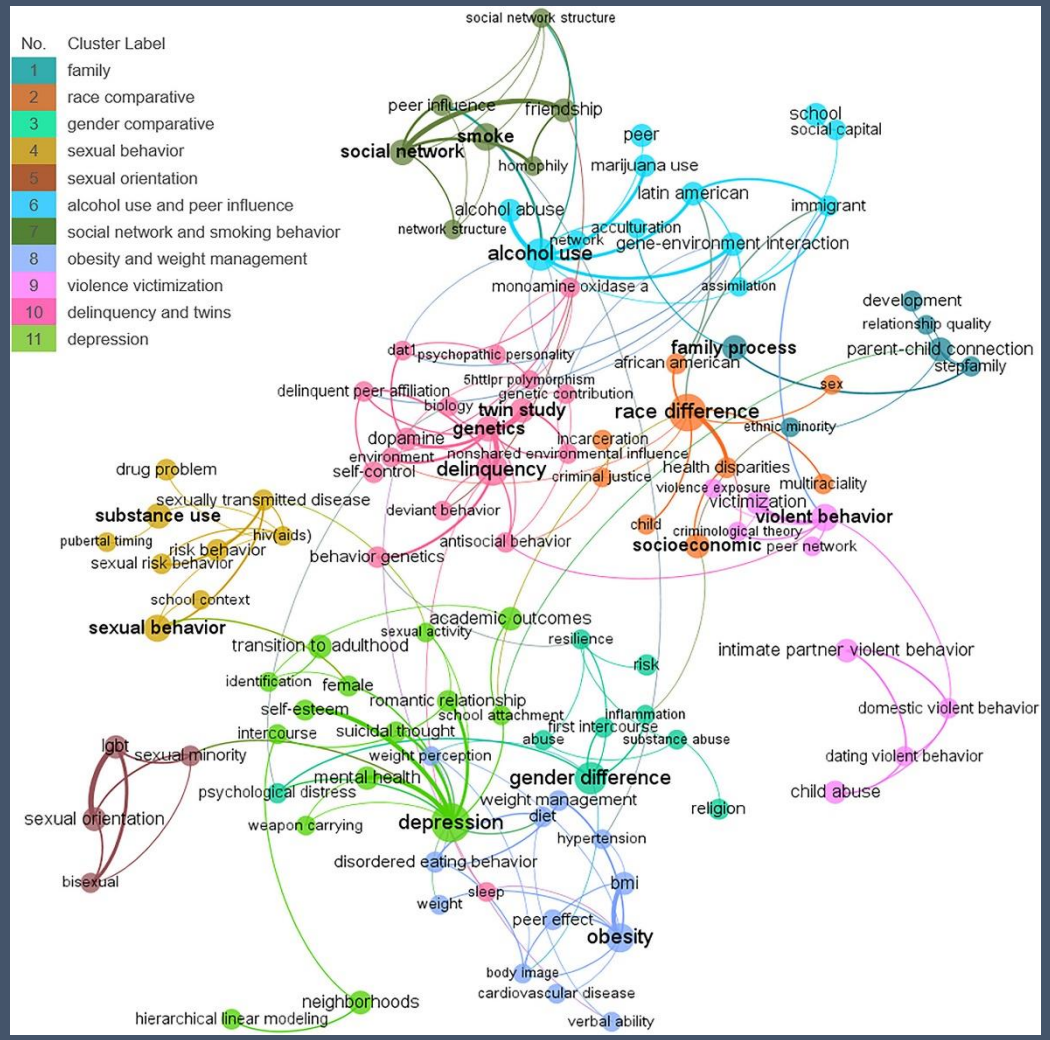
→促進研究者彼此交流專業意見?

→有更多跨領域學者進入政治學領域?

Berardo, R. (2011). Networking Networkers: An Initial Exploration of the Patterns of Collaboration among the Members of a New Community in Political Science. *PS: Political Science & Politics, 44(1), 69-75. doi:10.1017/S1049096510001885*

社會網絡分析應用
領域分析主題共現字

Fan, W.-M., Jeng, W., & Tang, M.-C. (2023). Using data citation to define a knowledge domain: A case study of the Add-Health dataset. *Journal of the Association for Information Science and Technology, 74(1), 81-98. doi:https://doi.org/10.1002/asi.24688*



以引用美國國家青少年成人長期研究資料(the National Longitudinal Study of Adolescent to Adult Health, Add-Health)的期刊論文為研究範圍，利用作者所下的關鍵字分析引用資料的期刊文獻主題的發展

社會網絡分析應用
環境與兒童肥胖關係

Knapp, E. A., Bilal, U., Burke, B. T., Dougherty, G. B., & Glass, T. A. (2018). A network approach to understanding obesogenic environments for children in Pennsylvania. *Connections, 38(1), 1-11. doi:doi:10.21307/connections-2018-001*

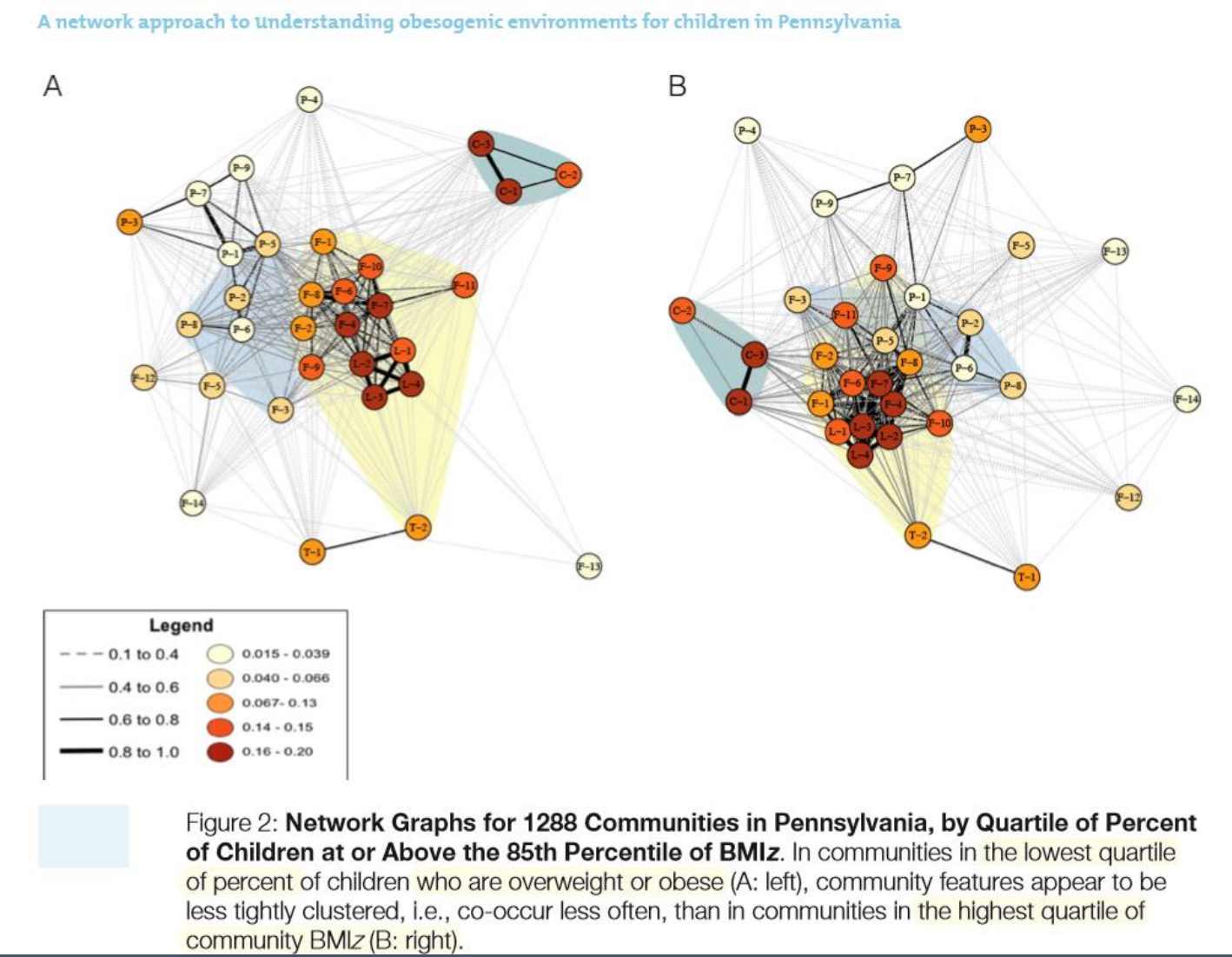
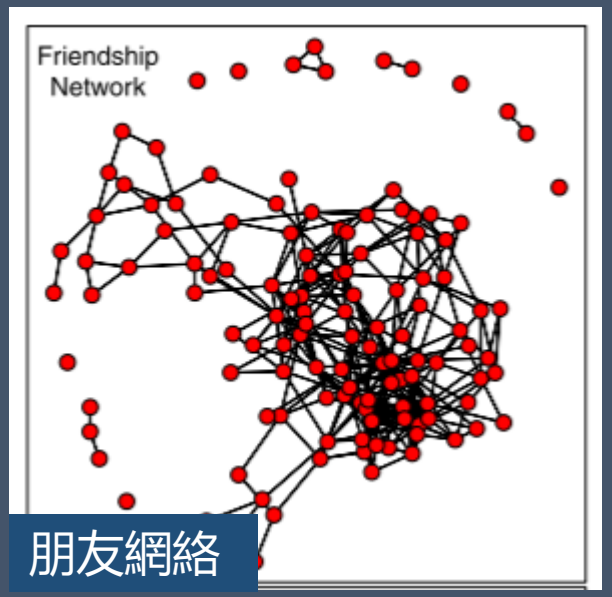
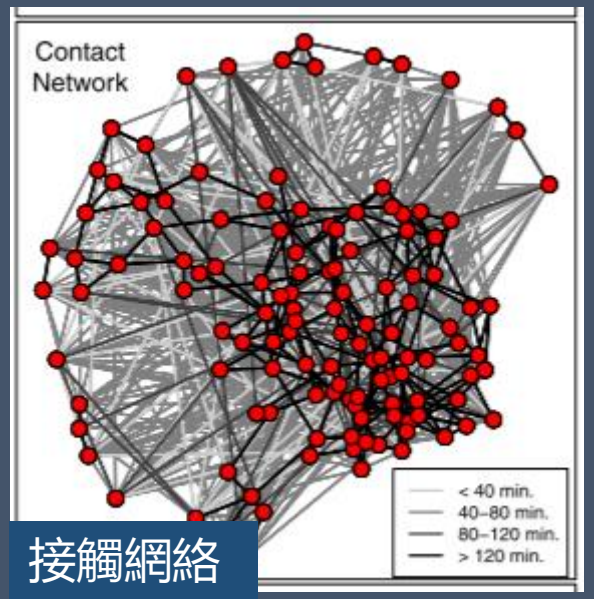


Figure 2: **Network Graphs for 1288 Communities in Pennsylvania, by Quartile of Percent of Children at or Above the 85th Percentile of BMIz**. In communities in the lowest quartile of percent of children who are overweight or obese (A: left), community features appear to be less tightly clustered, i.e., co-occur less often, than in communities in the highest quartile of community BMIz (B: right).

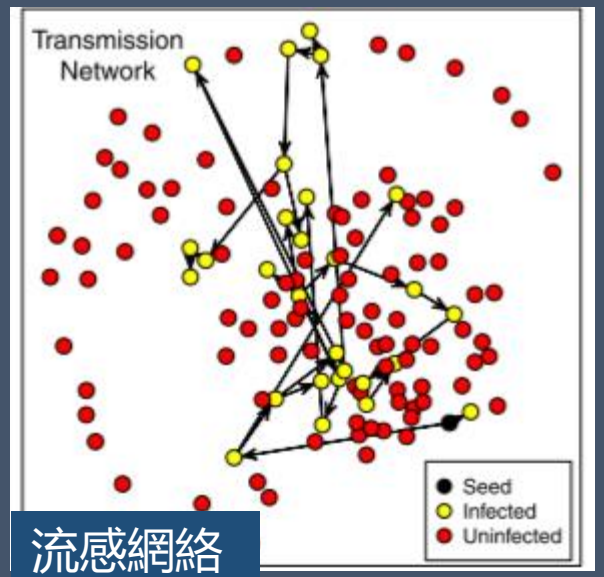
社會網絡分析應用
校園內流感傳染途徑



朋友網絡



接觸網絡



流感網絡

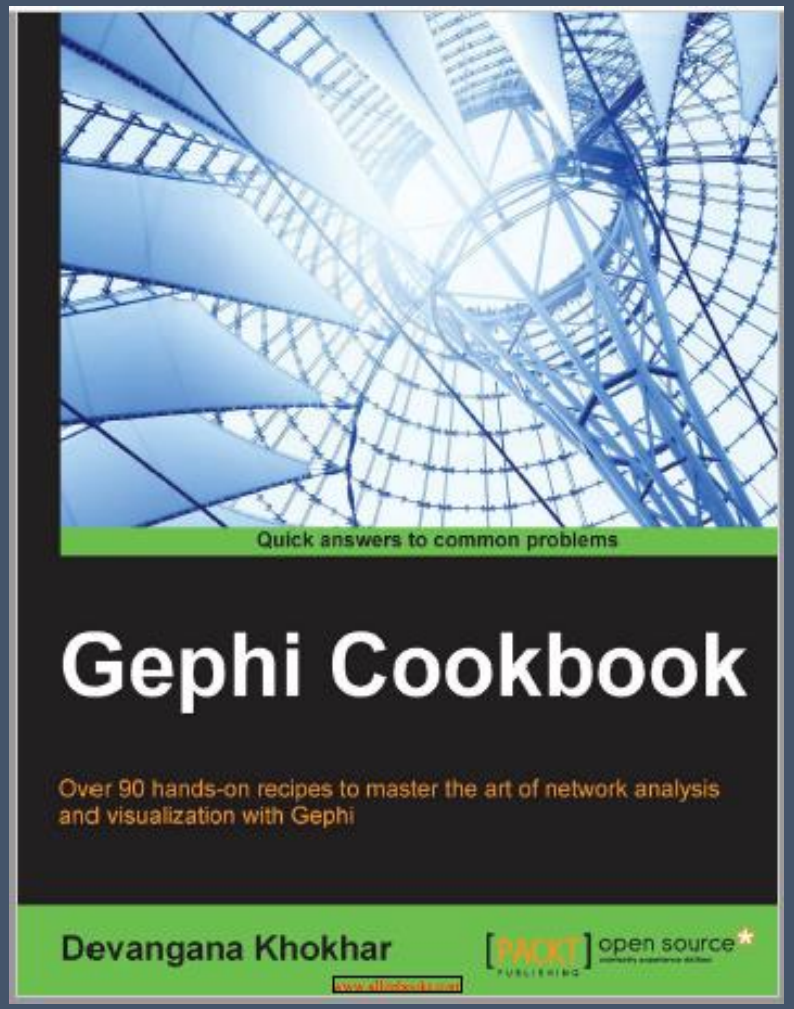
Potter, G. E., Handcock, M. S., Longini, I. M., & Halloran, M. E. (2012). Estimating within-school contact networks to understand influenza transmission. *The Annals of Applied Statistics, 6(1), 1-26, 26. Retrieved from https://doi.org/10.1214/11-AOAS505*

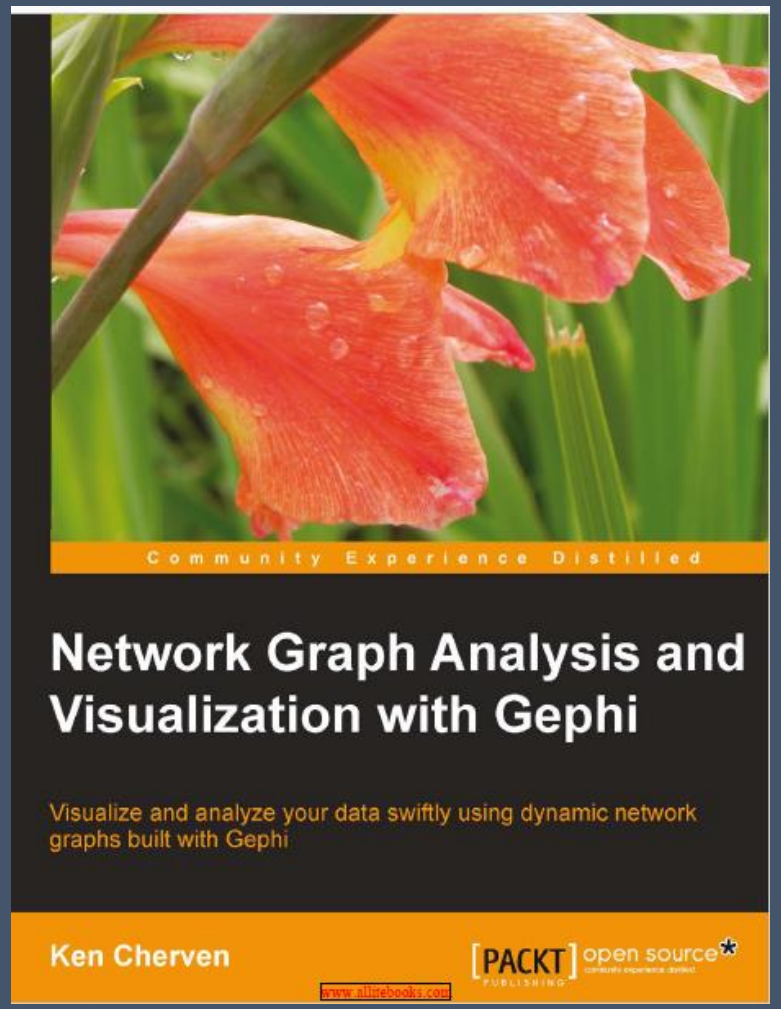
鄭君, 馬曉君, & 華強. (2014). 社會網絡分析工具Ucinet和Gephi的比較研究. *情報理論與實踐(TIA), 37*(8), 133-138.

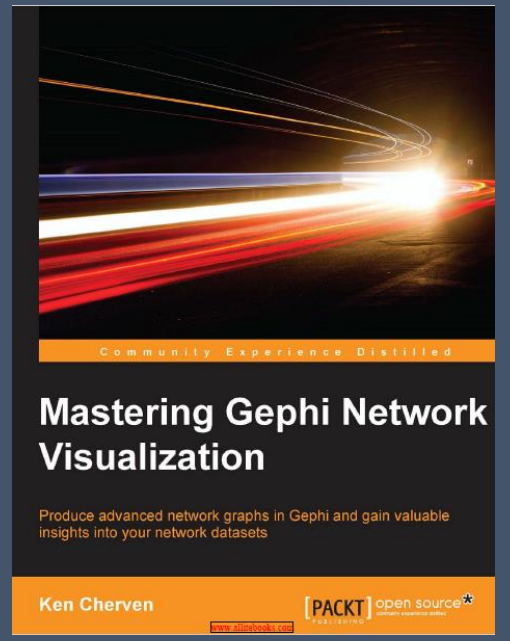
社會網絡分析工具介紹

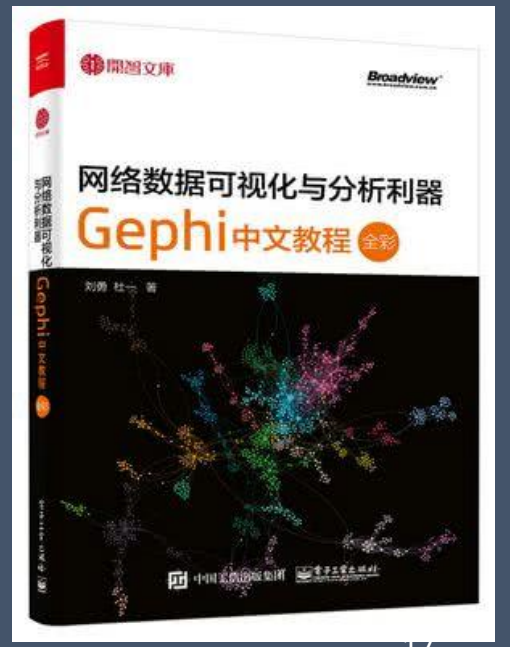
項目	Ucinet	Gephi
作業系統	Windows	任何支持Java. 6 和OpenGL 的系統
支援演算法擴展	否	是
支持矩陣	是	是
網路規模	中型	大型
原始程式碼開放情況	不開放	開放
付費類型	商業60天試用	免費
基本功能	社會網路資料和其他1-mode及2-mode數據分析，可做矩陣代數和多元統計分析等	圖像資料探索性分析、關聯式資料分析、社交網路分析、生物網路分析等
比較	Network File=>Gephi edge Attribute File=>Gephi node 不支援中文資料 統計功能較多元	Node table Edge table 可處理中文資料 使用介面友善，視覺化功能佳

社會網絡工具書 Gephi









02 社會網絡基本觀念
Basic concept of SNA

基本觀念

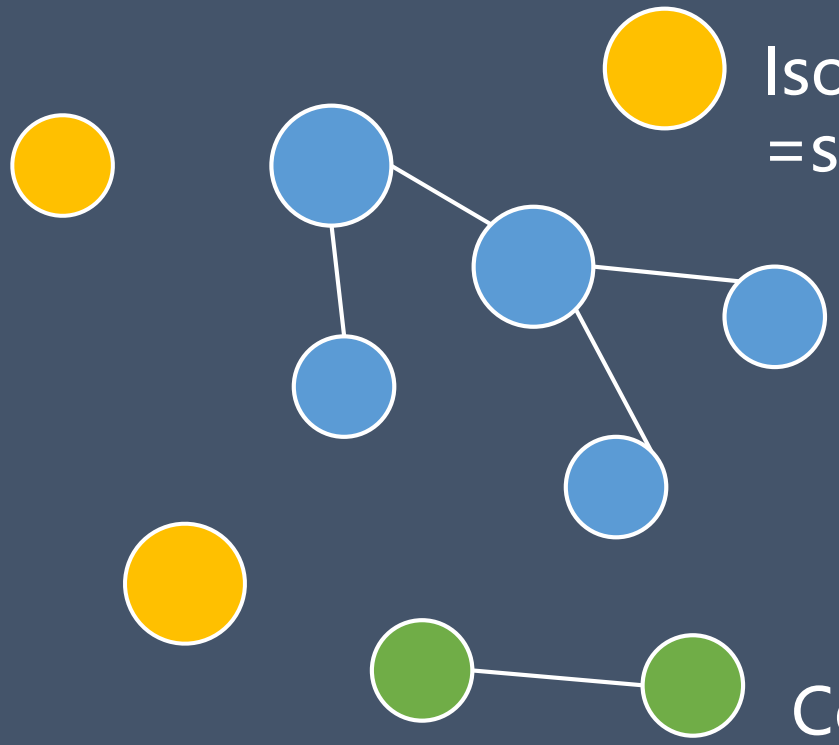
網絡

=節點+連結= node + edge

Actor Vertex Node

個人、產品、
事件、國家

概念解釋



Isolate
=single node

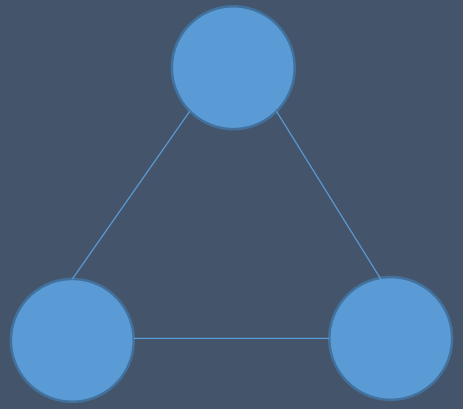
Main Component
= the largest pair of nodes

Component
= a pair of nodes

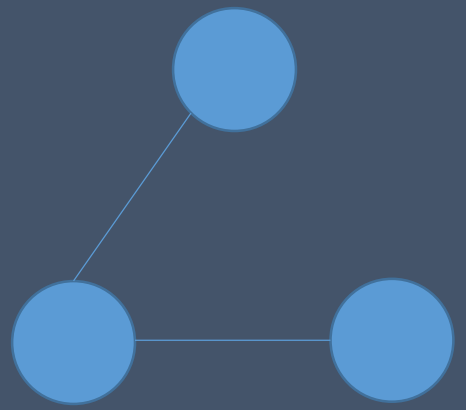
名詞解釋

Dyad/ Pair

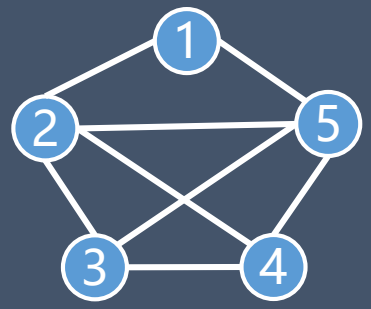
Closure Triad



Triplet



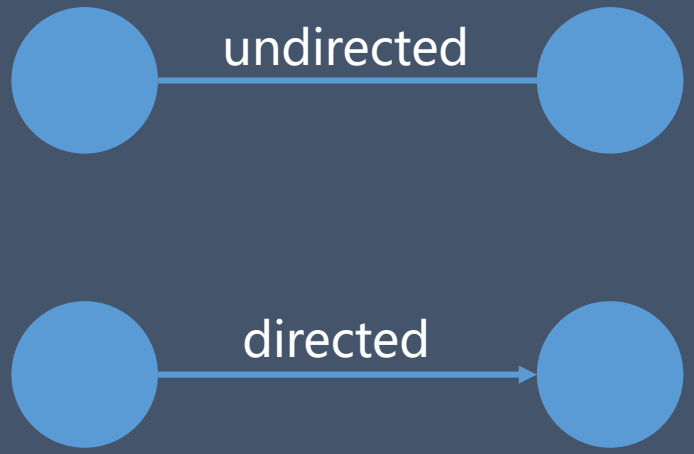
名詞解釋



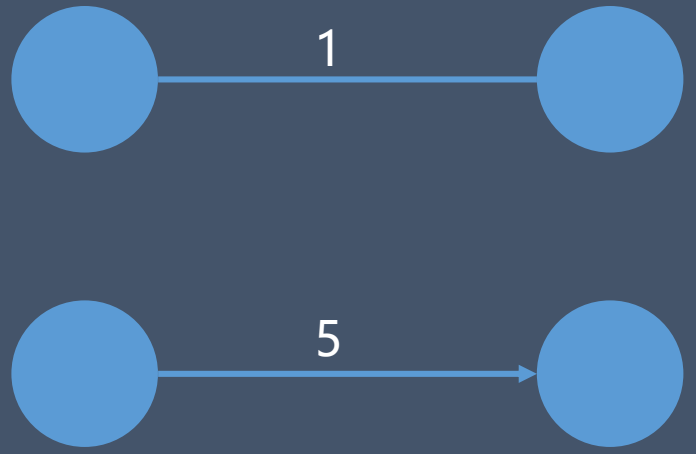
名詞	定義	例子	EX
walk	節點與連結皆可無限計算	金錢流向	(1-2-5-3-5)
path	節點與連結僅能計算一次	病毒感染	(1-2-3-4)
trial	每一條連結進能計算一次 (節點可重複計算)	八卦訊息	(1-2-3-5-2)
geodesics distance	任兩節點之間最短的path	包裹快遞	(1-2-4)

名詞解釋

- Direction



- Value/weight



概念解釋—圖像基本定義

Graph G { a finite set of nodes, N (each node is unique)
a finite set of edges, E (each edge is unique)

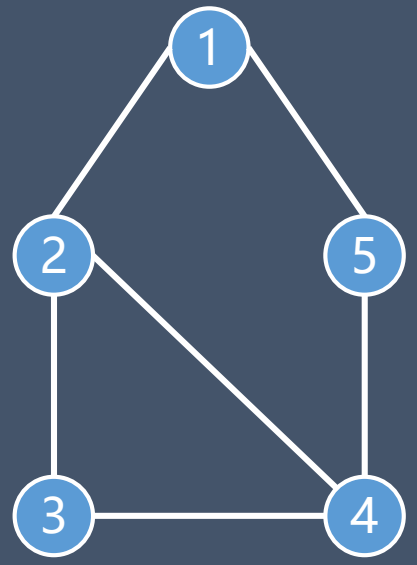
- Each edge is a pair (n1,n2) where n1, n2 ∈ N

概念解釋—無向圖(undirected)

N = { 1, 2 ,3, 4, 5}

E = { (1,2), (2, 3), (2, 4), (3, 4), (4, 5), (5, 1)
(2,1), (3, 2), (4, 2), (4, 3), (5, 4), (1, 5) }

All edges are two-way. Edges are unordered pairs.



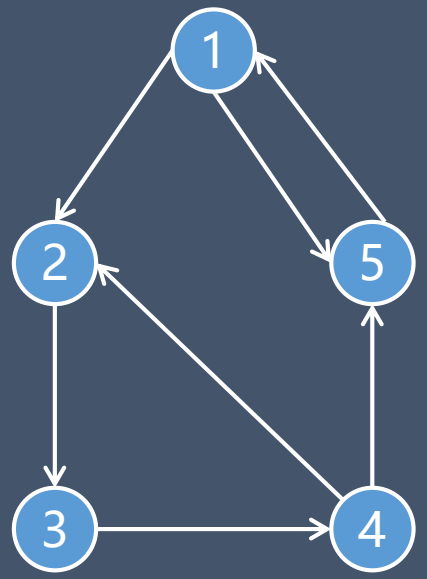
Target / Source	1	2	3	4	5
1	0	1	0	0	1
2	1	0	1	1	0
3	0	1	0	1	0
4	0	1	1	0	1
5	1	0	0	1	0

概念解釋—有向圖(directed)

N = { 1, 2, 3, 4, 5}

E = { (1, 2), (1, 5), (2, 3), (3, 4), (4, 2), (4, 5), (5, 1) }

All edges are one-way as indicated by the arrows.
Edges are ordered pairs.



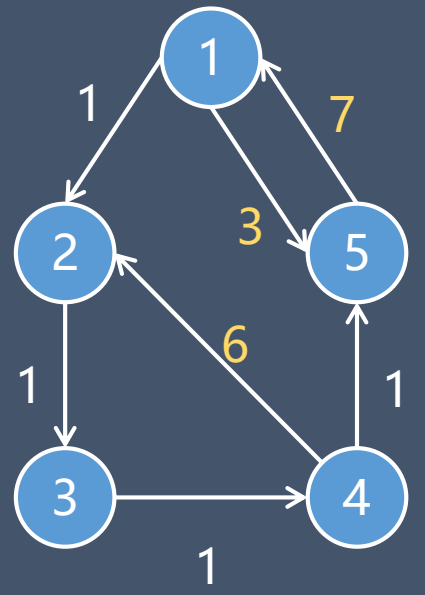
Target / Source	1	2	3	4	5
1	0	1	0	0	1
2	0	0	1	0	0
3	0	0	0	1	0
4	0	1	0	0	1
5	1	0	0	0	0

概念解釋—有向圖+權重

N = { 1, 2, 3, 4, 5}

E = { (1, 2), (1, 5), (2, 3), (3, 4), (4, 2), (4, 5), (5, 1) }

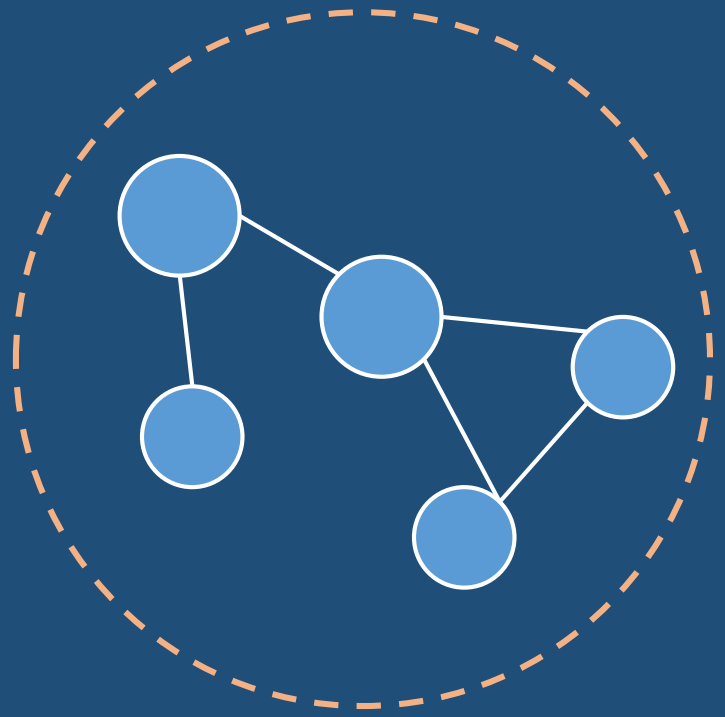
All edges are one-way as indicated by the arrows.
Edges are ordered pairs.



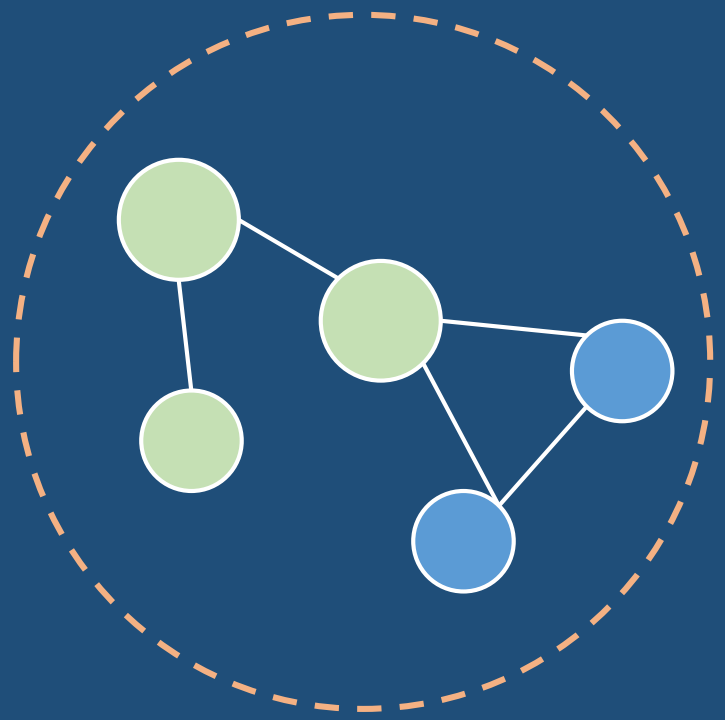
Target / Source	1	2	3	4	5
1	0	1	0	0	3
2	0	0	1	0	0
3	0	0	0	1	0
4	0	6	0	0	1
5	7	0	0	0	0

概念解釋

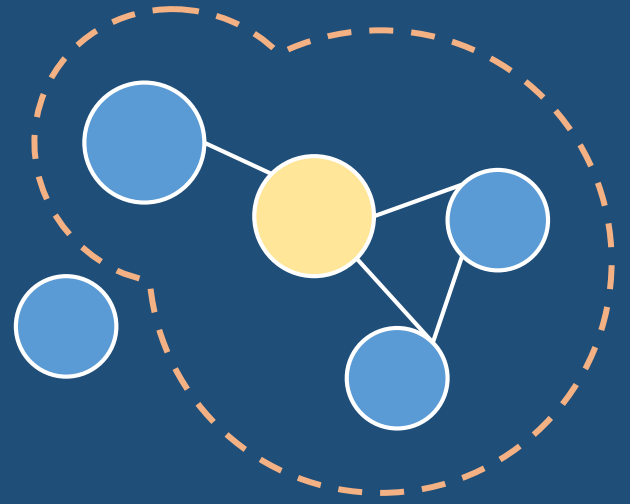
Social Network
One Mode



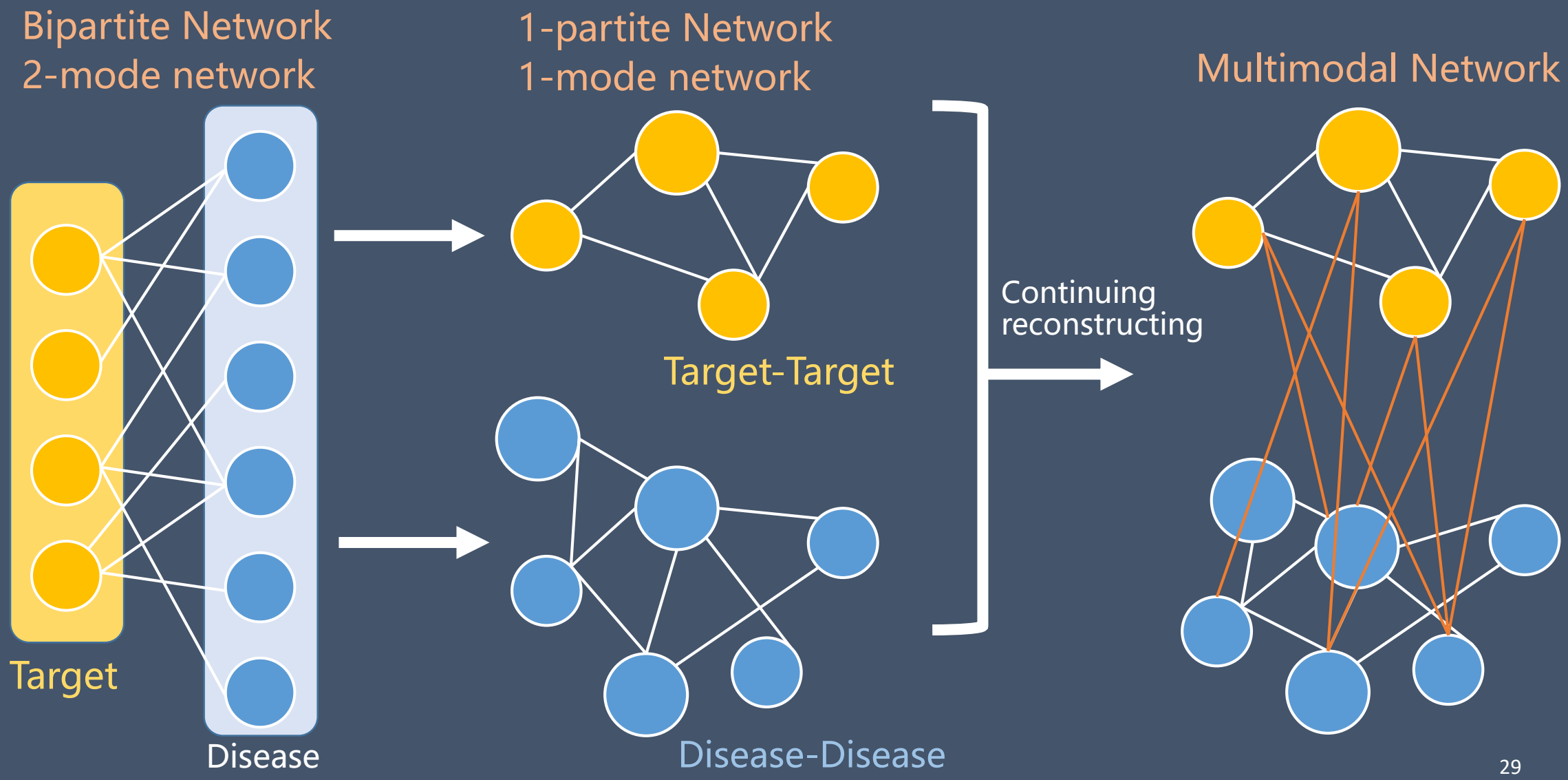
Social Network
Two Mode



Ego-Centered Network
EgoNet



概念解釋


Bipartite Network
2-mode network
1-partite Network
1-mode network
Multimodal Network
Target
Disease
Target-Target
Disease-Disease
Continuing
reconstructing

概念解釋—計算圖像可能的連結數

Q: For a set N with n elements,
how many possible edges there?

A:

For undirected graphs, the number of pairs in N:

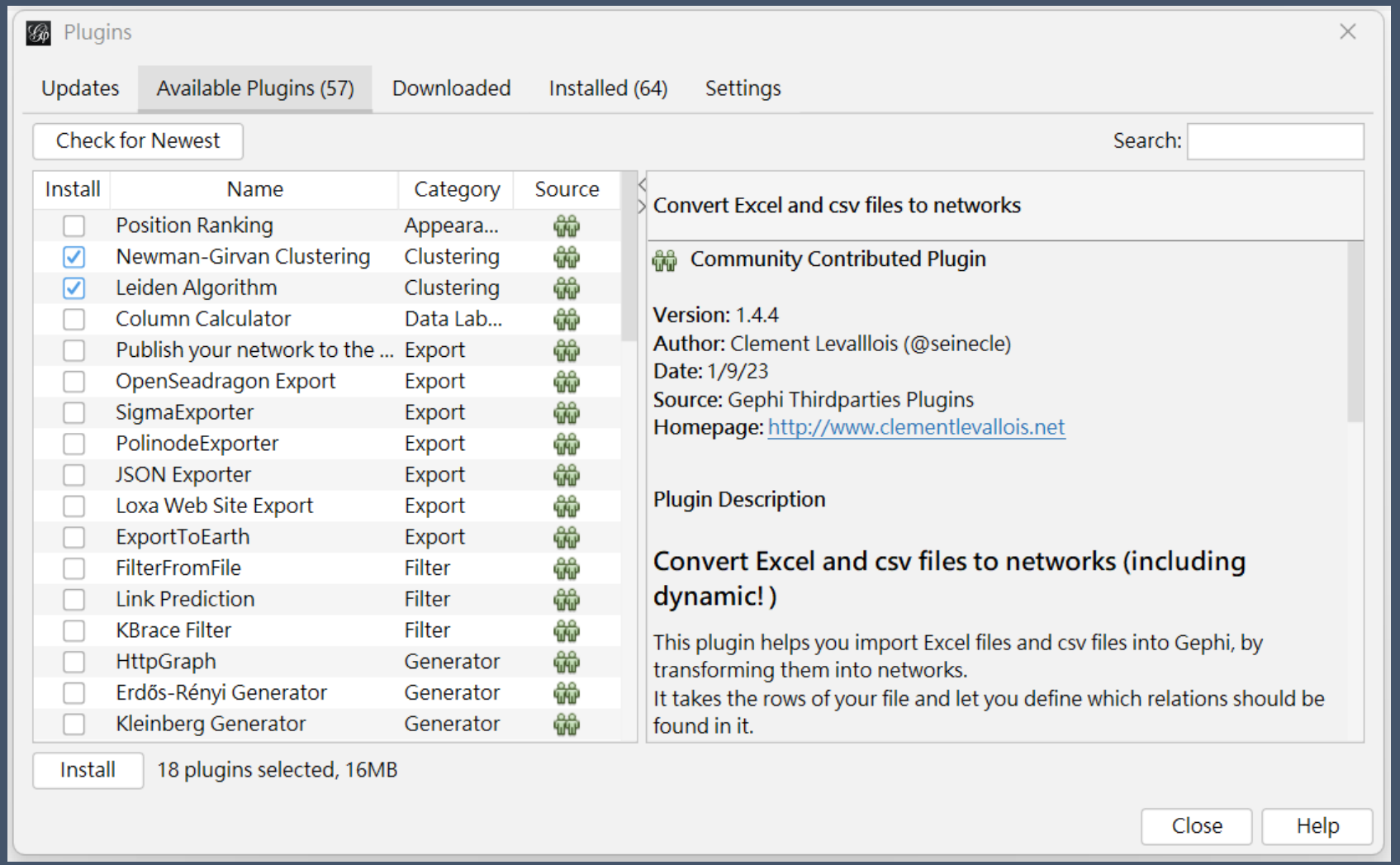
$= n \cdot (n - 1) / 2$

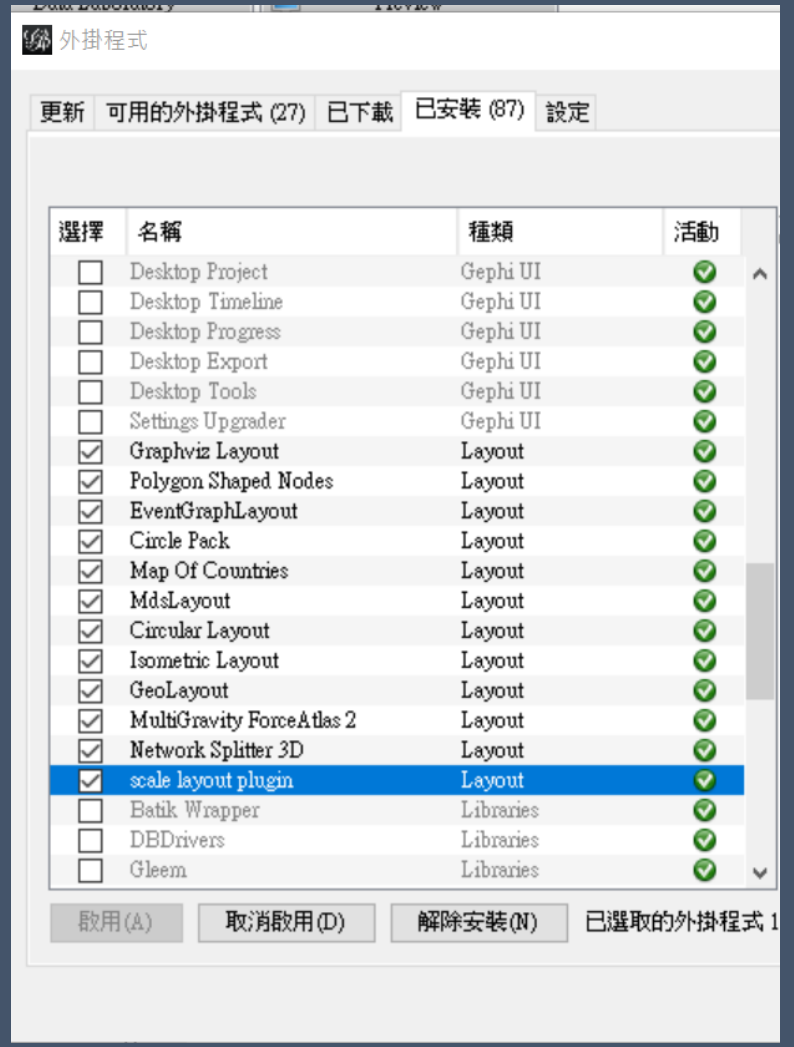
For directed graphs, the number of ordered pairs in N:

$= n^2 - n = n \cdot (n - 1)$

Gephi軟體安裝(連結)

安裝外掛程式 種類：Layout、Import、Clustering、Metrics



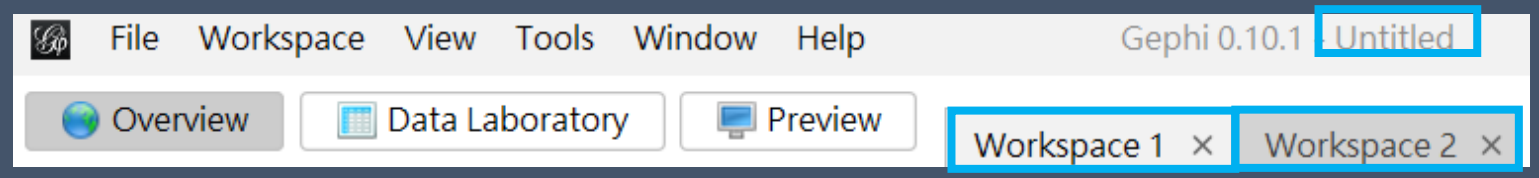


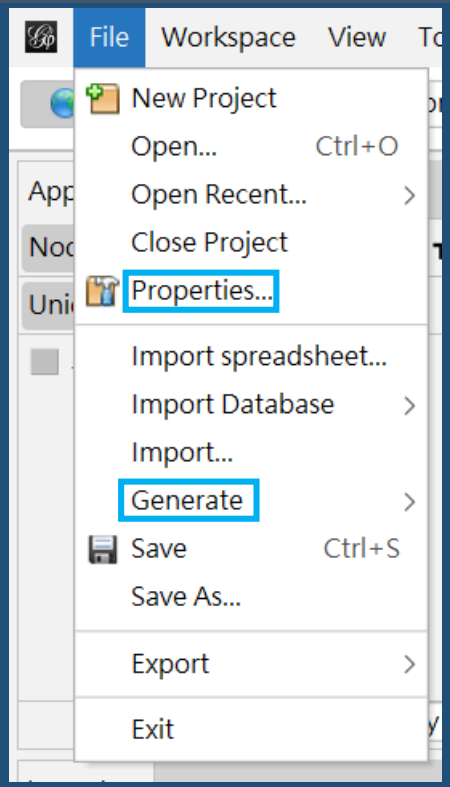
03 Gephi軟體入門

基本功能列

開啟新專案

- Project 專案 (每次只能開一個)
  >Workspace 空白頁(可同時開啟多個)

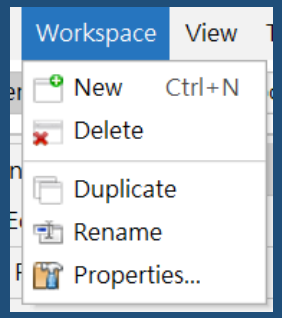




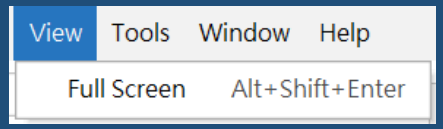
資料匯入
資料匯出
另存新檔

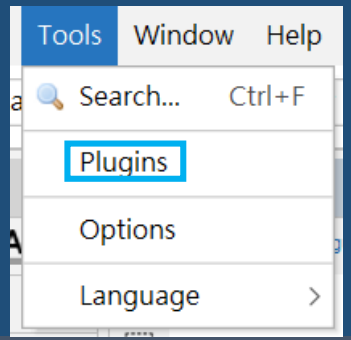
專案資料紀錄
專案名、作者名
關鍵字、描述

可以產生範例檔案
隨機靜態、多途靜態
動態資料

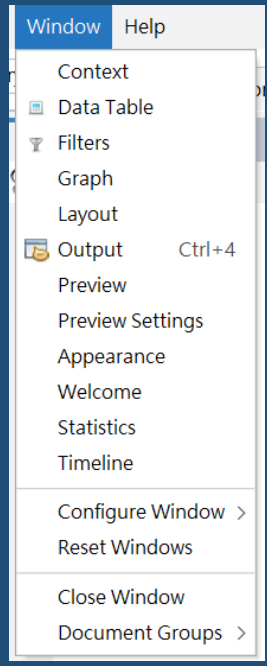


新增空白頁
刪除空白頁
空白頁重新命名

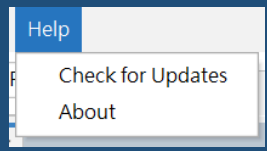




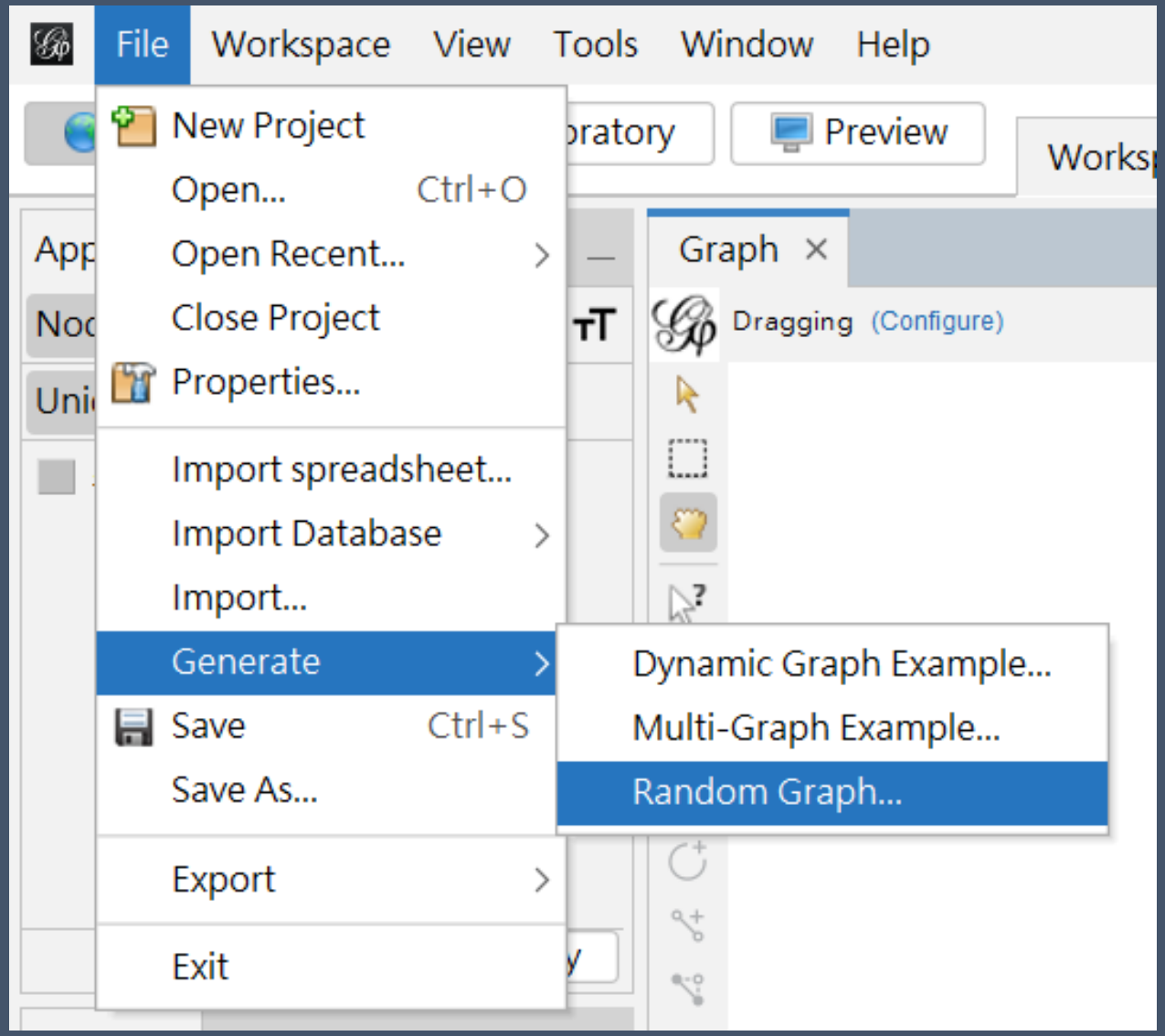
搜尋
外掛程式
進階設定
語言

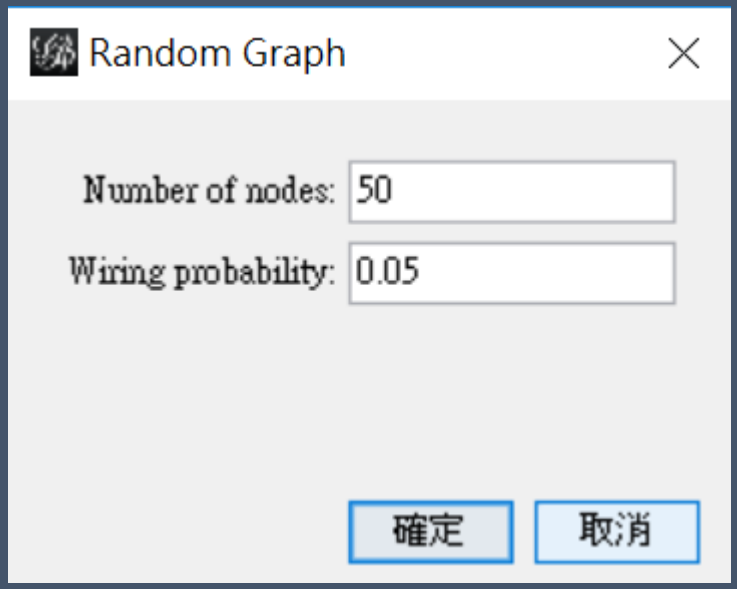


選用需要的視窗



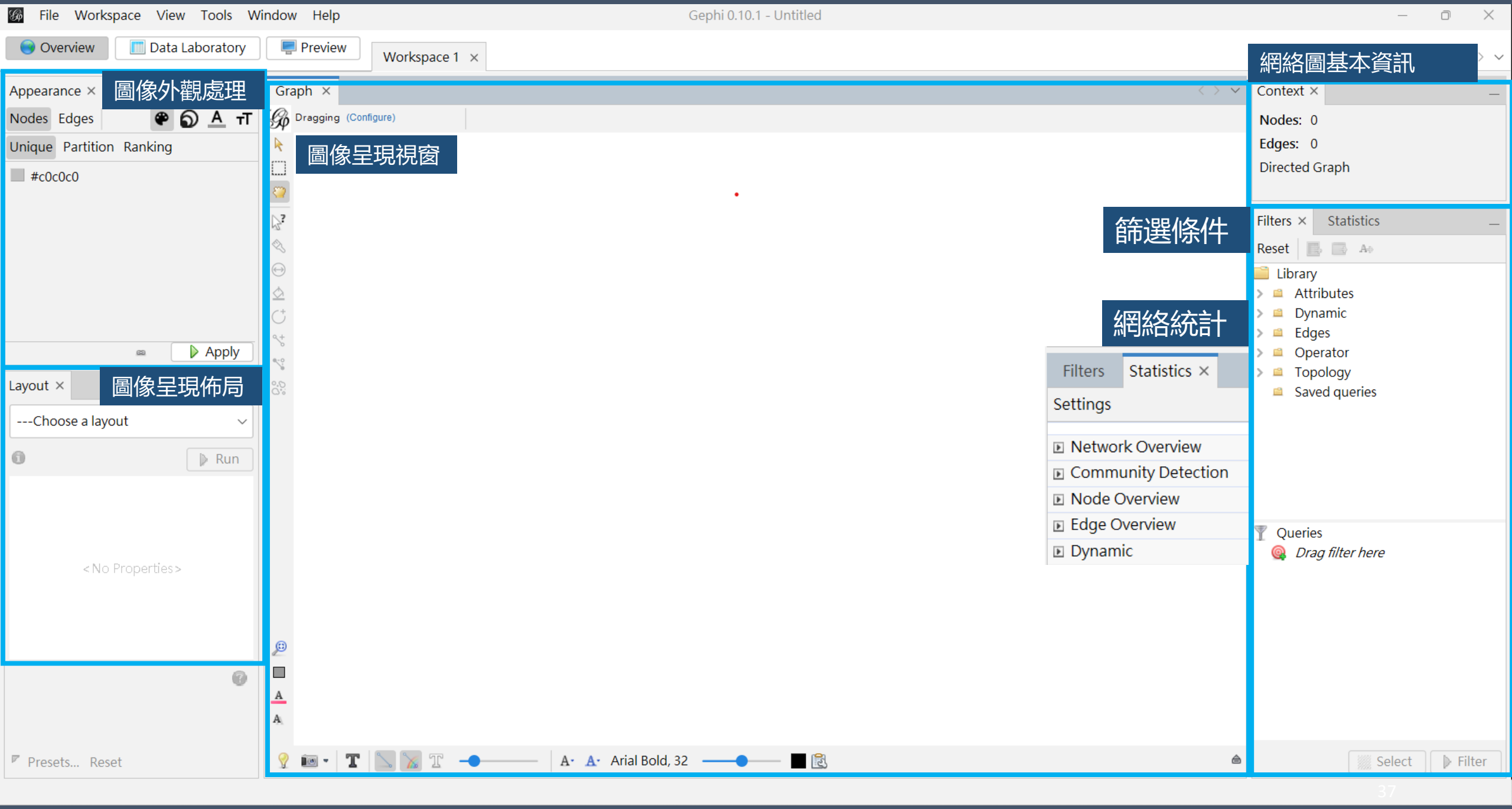
軟體更新
軟體基本訊息


File
Workspace
View
Tools
Window
Help
New Project
Open... Ctrl+O
Open Recent...
Close Project
Properties...
Import spreadsheet...
Import Database
Import...
Generate
Save Ctrl+S
Save As...
Export
Exit
Preview
Graph
Dragging (Configure)
Dynamic Graph Example...
Multi-Graph Example...
Random Graph...

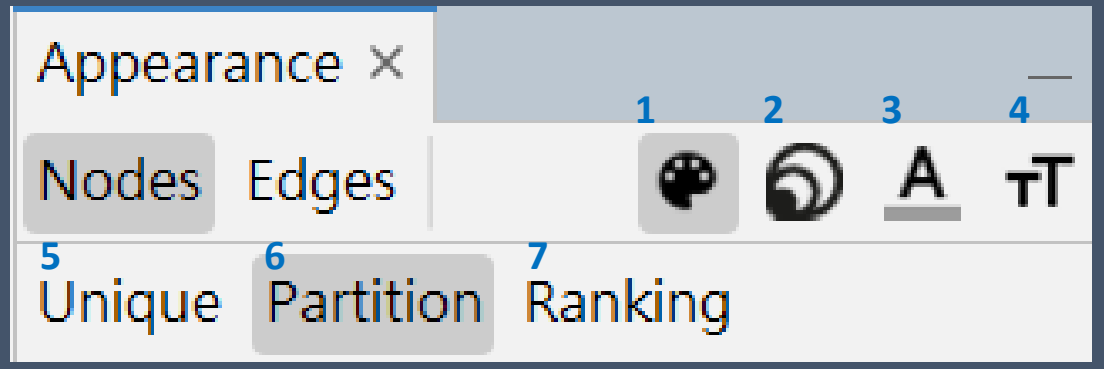

Random Graph
Number of nodes: 50
Wiring probability: 0.05
確定
取消

Gephi軟體入門
04 Overview

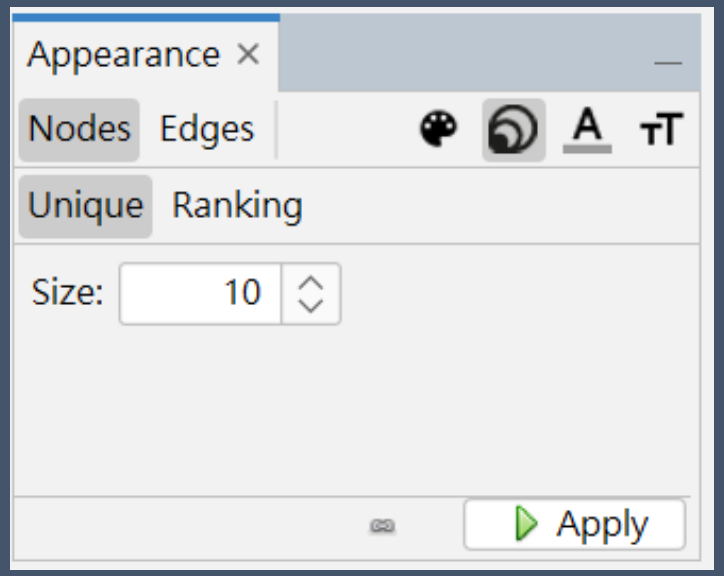
4.1 Appearance 圖像外觀處理

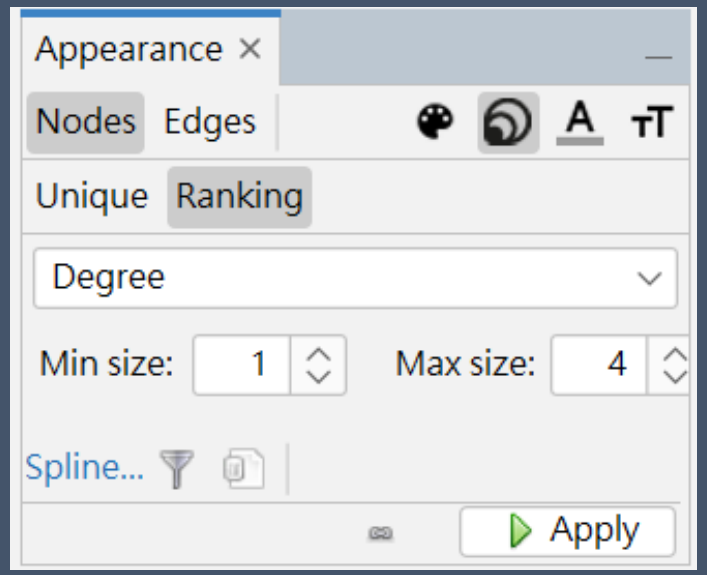

File Workspace View Tools Window Help
Gephi 0.10.1 - Untitled
Overview
Data Laboratory
Preview
Workspace 1
網絡圖基本資訊
Appearance
圖像外觀處理
Nodes Edges
Unique Partition Ranking
#c0c0c0
Apply
Graph
Dragging (Configure)
圖像呈現視窗
Context
Nodes: 0
Edges: 0
Directed Graph
篩選條件
Filters
Statistics
Reset
Library
Attributes
Dynamic
Edges
Operator
Topology
Saved queries
網絡統計
Filters
Statistics
Settings
Network Overview
Community Detection
Node Overview
Edge Overview
Dynamic
Layout
圖像呈現佈局
---Choose a layout
Run
<No Properties>
Queries
Drag filter here
Presets... Reset
Arial Bold, 32
Select
Filter

圖像外觀處理(節點/連結)



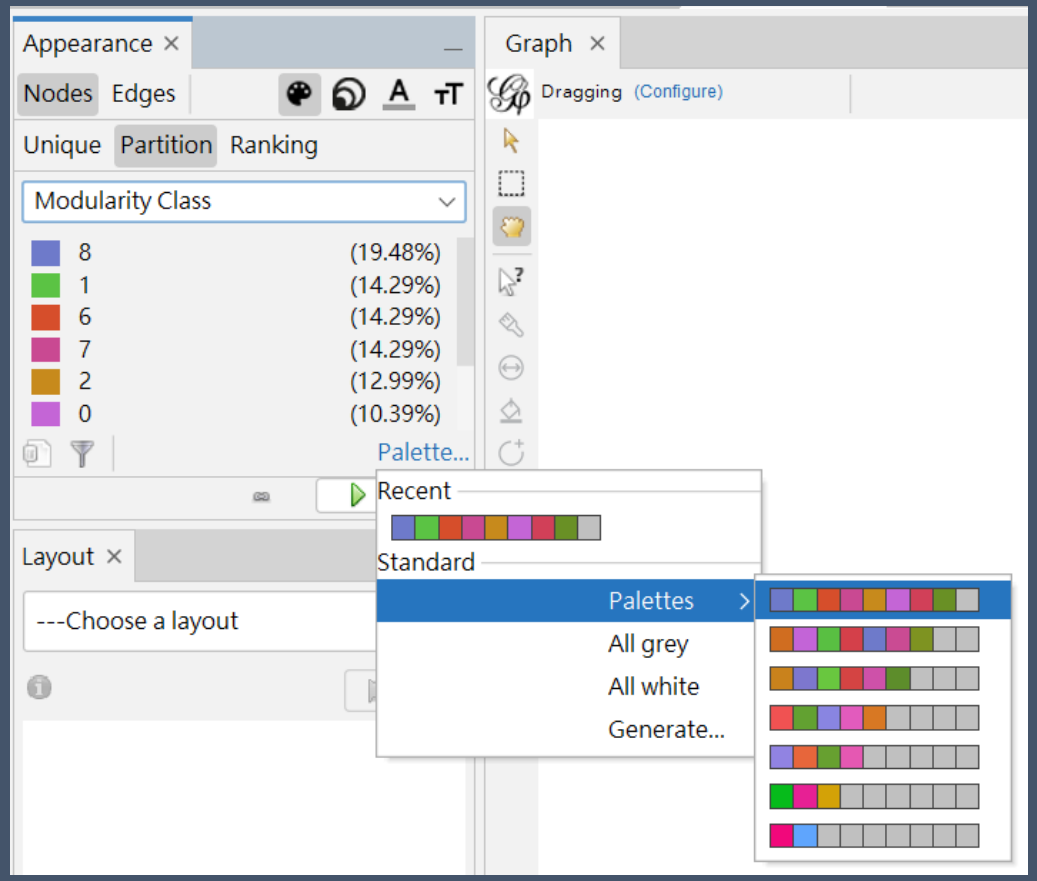
1. 節點顏色
2. 節點大小
3. 節點字體顏色
4. 節點字體大小
5. Unique 每一節點的進行相同動作處理
6. Partition 依據屬性類別進行標示
7. Ranking 依據屬性級距進行標示



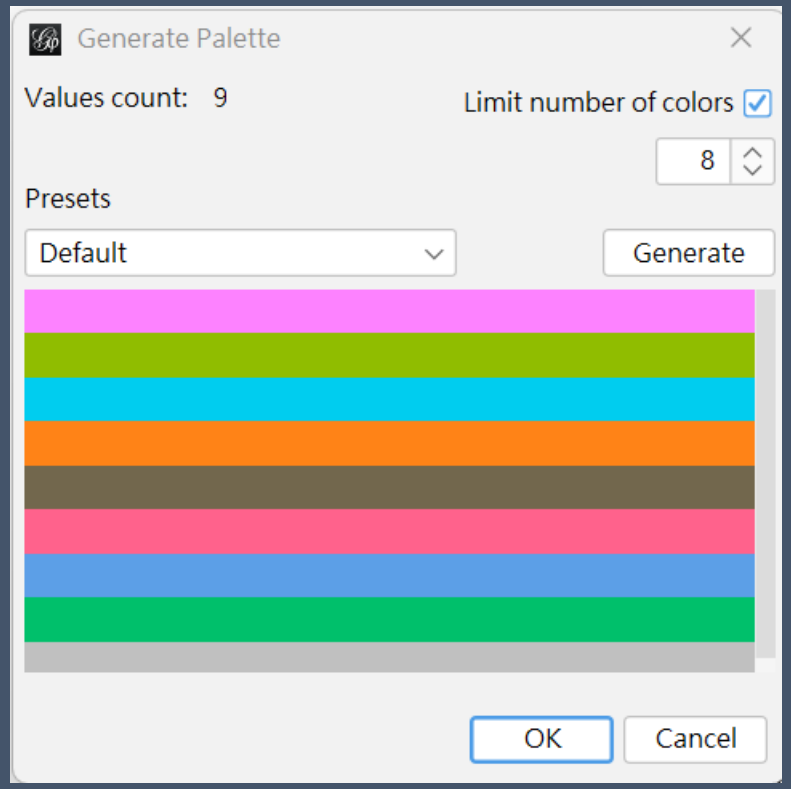


圖像外觀處理(節點/連結)

想要換節點顏色
Palette... > Palettes

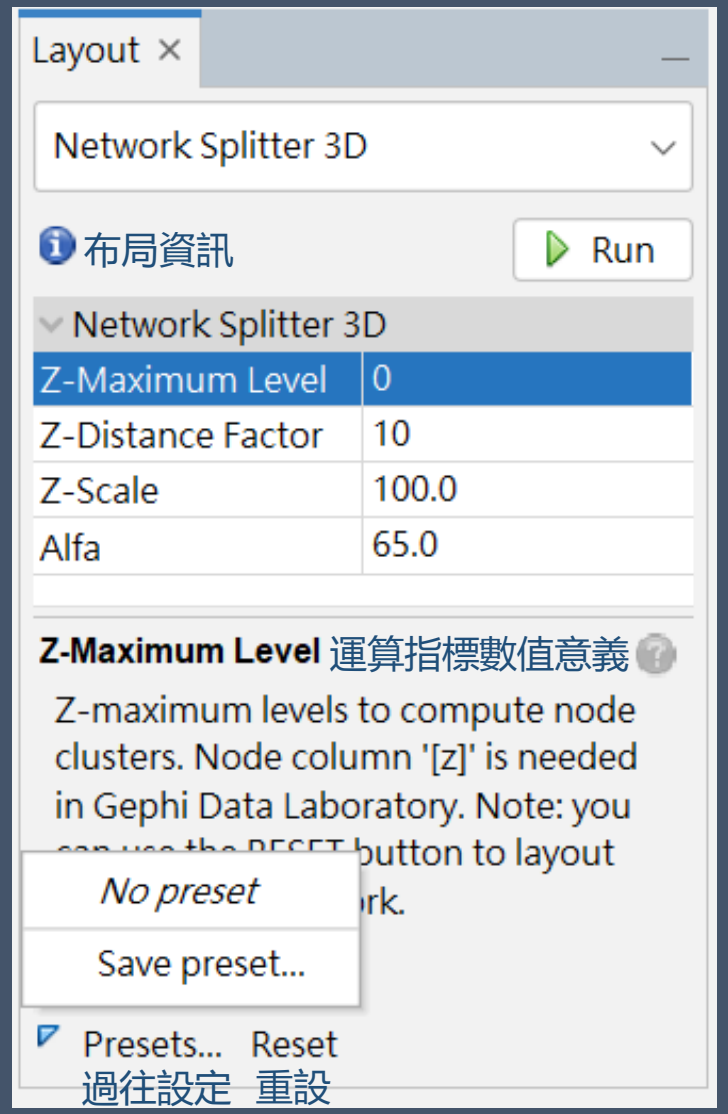


節點顏色重新產生，上限20色
Palette... > Generate...



Gephi軟體入門

4.2 Overview Layout 圖像呈現佈局



基本圖像佈局 Basic Layout

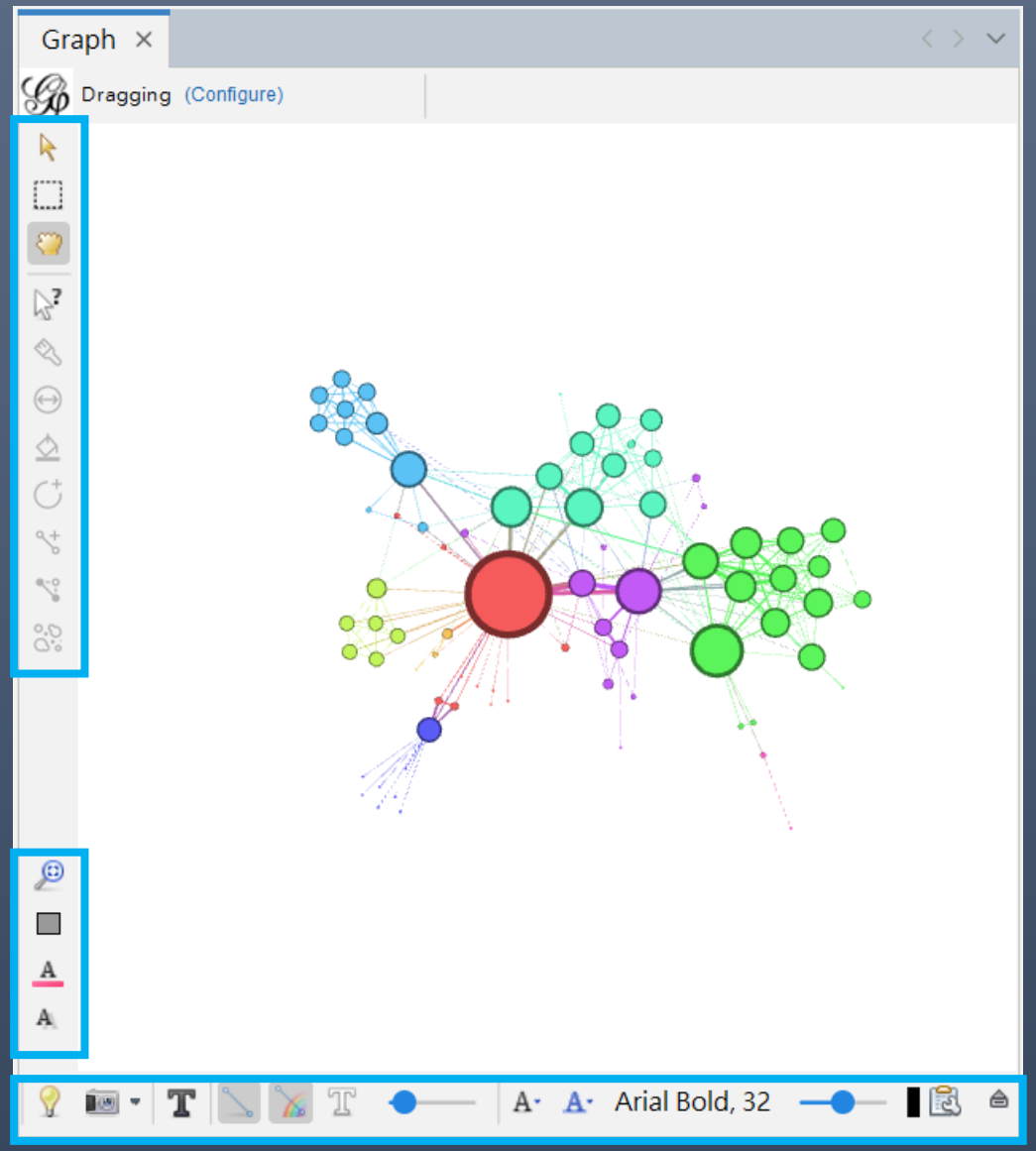
Layout 名稱	作用
Contraction	縮小
Expansion	放大
Rotate	90度選轉
Random Layout	隨機分布
Noverlap	節點不重疊

Noverlap	
speed	3.0
ratio	10.0
margin	50.0

速度
節點不重疊比率，數值↑圖像↑
邊緣

Gephi軟體入門

4.3 Overview Graph

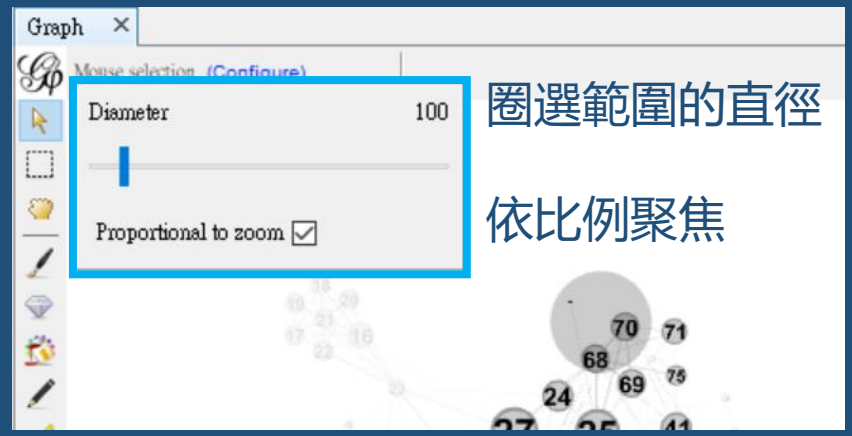


選擇節點以及與其相鄰的點

正方形

圓形

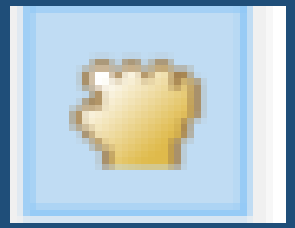
長按左鍵



圈選範圍的直徑

依比例聚焦

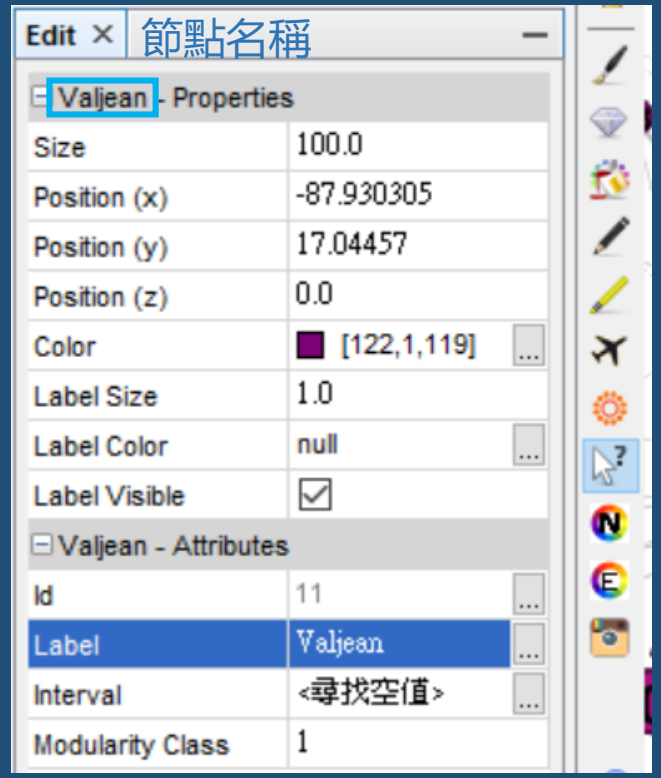
移動節點



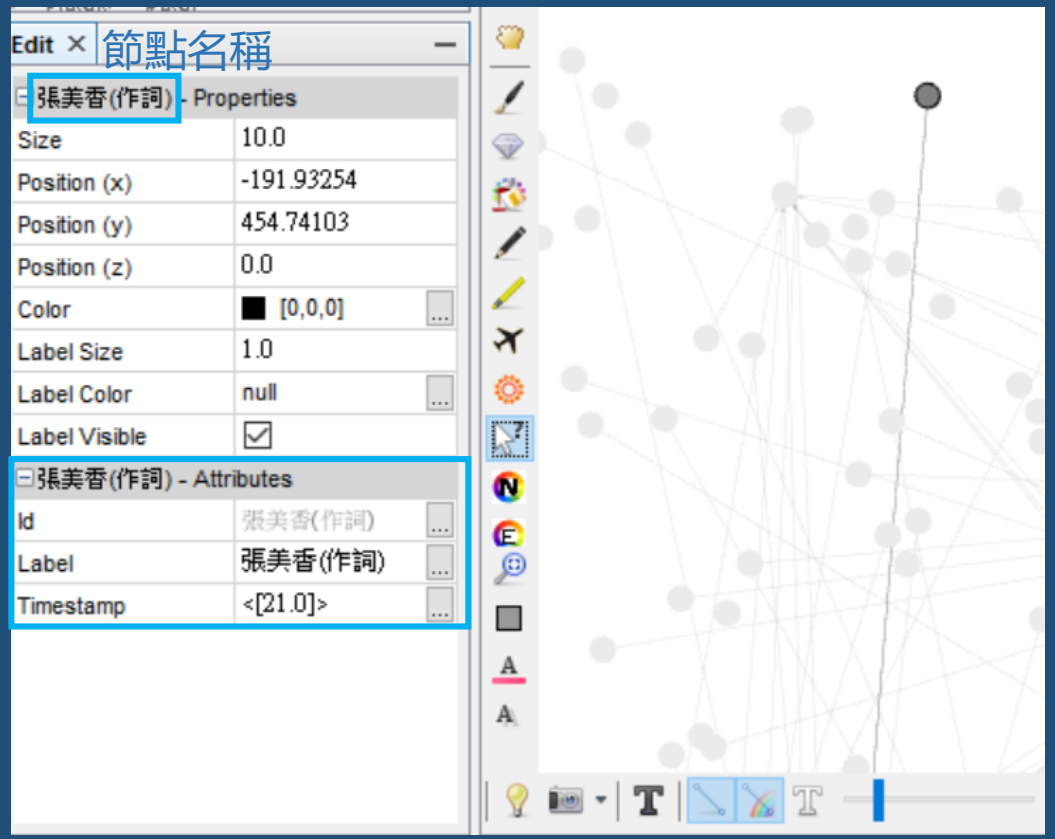
用滑鼠左鍵拖曳要移動的節點

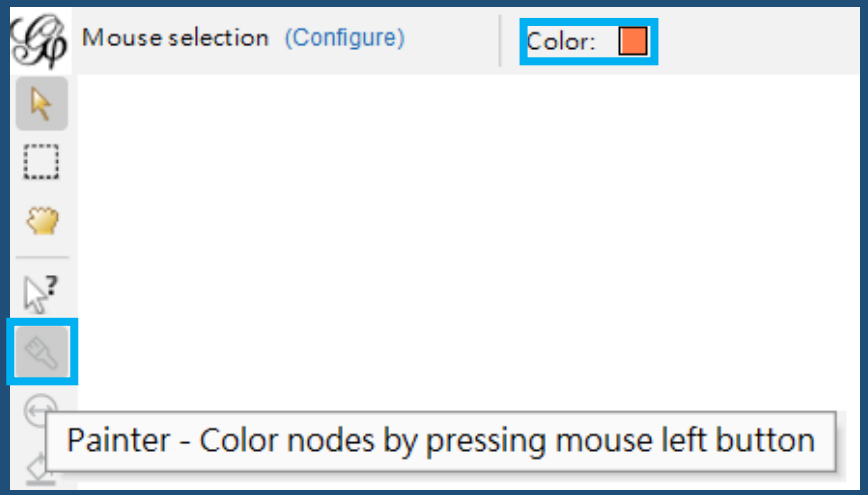
節點訊息
Properties 在Overview的節點資訊
Attributes 在Data Laboratory的節點資訊

靜態資料



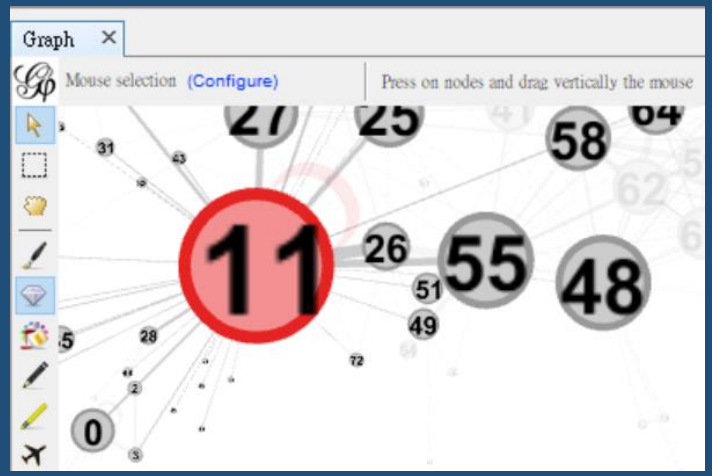
動態資料



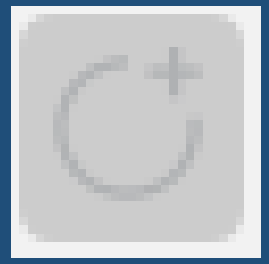


選擇右上角顏色框，可以按左鍵將結點填滿顏色

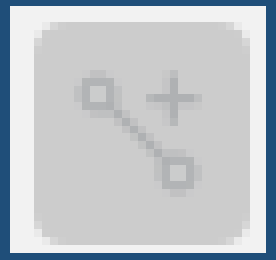
特定節點放大或縮小
用滑鼠左鍵按住節點
滑鼠向上，節點放大，
滑鼠向下，節點縮小



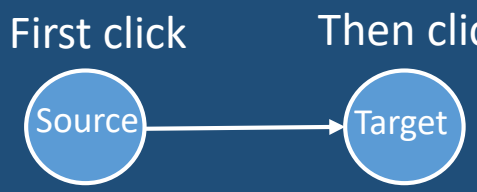
針對某一節點及其鄰近節點進行著色



節點鉛筆
可增加節點
one click one node



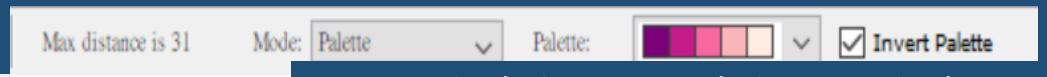
連結鉛筆
可增加連結
If edge is directed, click from Source to Target



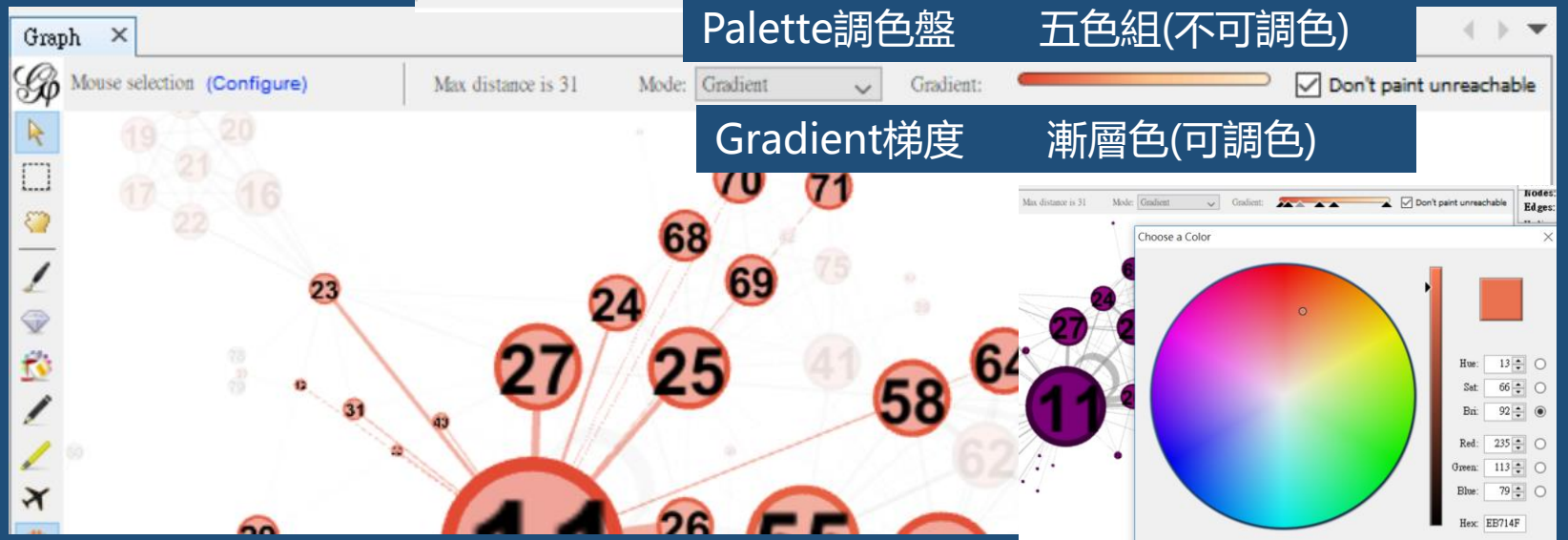
顯示兩節點之間最短距離

以單一節點為中心，依照距離塗上不同色

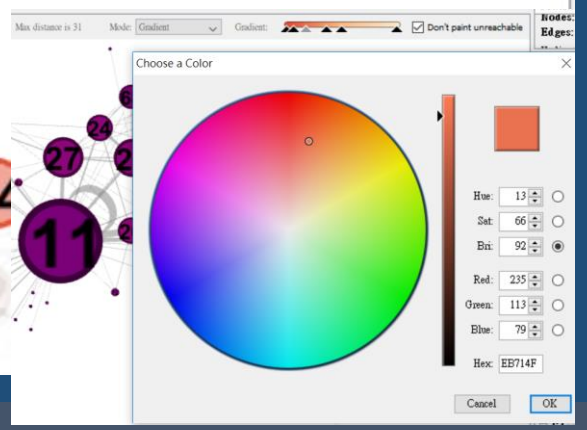
色組顏色可以反置：近/深色→遠/淺色 、近/淺色→遠/深色



Palette調色盤　五色組(不可調色)

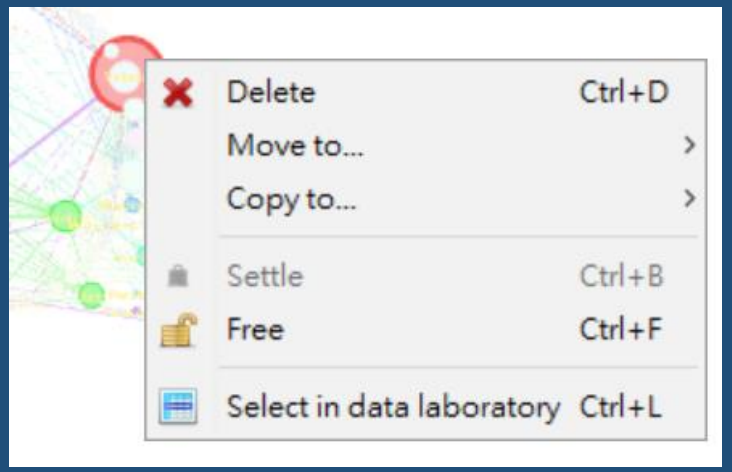


Gradient梯度　漸層色(可調色)



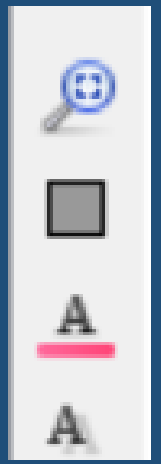
中心節點無法觸及到的節點不加以著色

針對節點按右鍵還可以...



刪除
移動到>新分頁或其他分頁
複製到>新分頁或其他分頁
鎖住(在轉換layout時節點不會移動)
解除
在資料實驗室中選取

Id	Label
11	Valjean
48	Gavroche
55	Marius
27	Javert
25	Thenardier
23	Fantine
58	Enjolras
62	Courfeyrac
64	Bossuet
63	Bahorel
65	Joly
24	MmeThenardier
26	Cosette
41	Eponine



圖像復位

全部節點顏色復位(顏色可選)

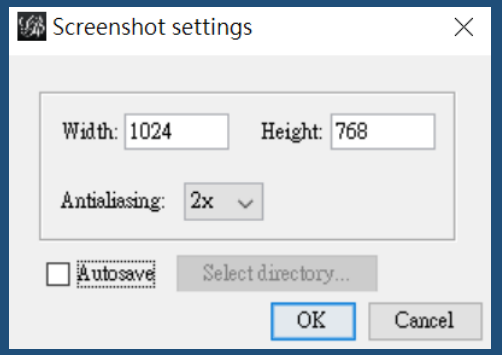
節點標籤顏色復位

節點標籤大小復位


Arial Bold, 32

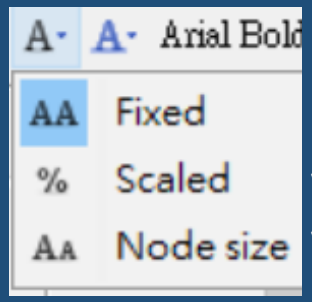
按右鍵可選取顏色


圖像背景顏色
圖像快照


Screenshot settings
Width: 1024
Height: 768
Antialiasing: 2x
Autosave
Select directory...
OK
Cancel

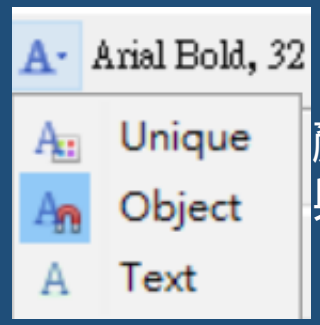

顯示節點標籤
顯示連結
連結與來源節點
相同顏色
顯示連結標籤
調整連結粗細

節點標籤
字體大小

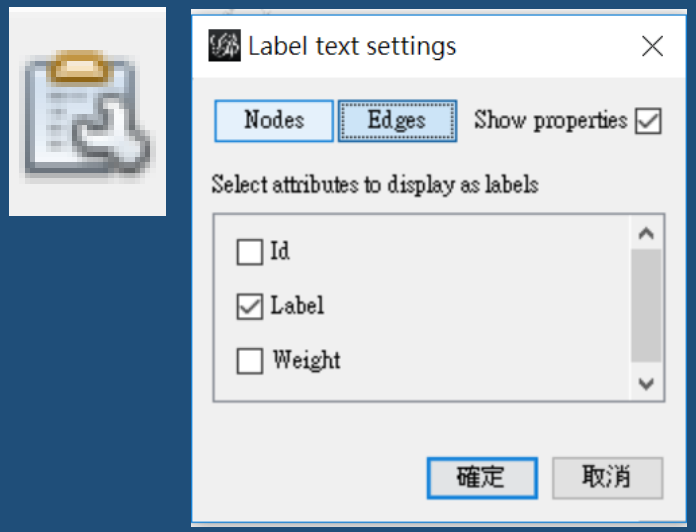


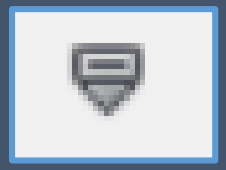
固定大小
依據比例放大
依據節點尺寸縮放

節點標籤
字體顏色

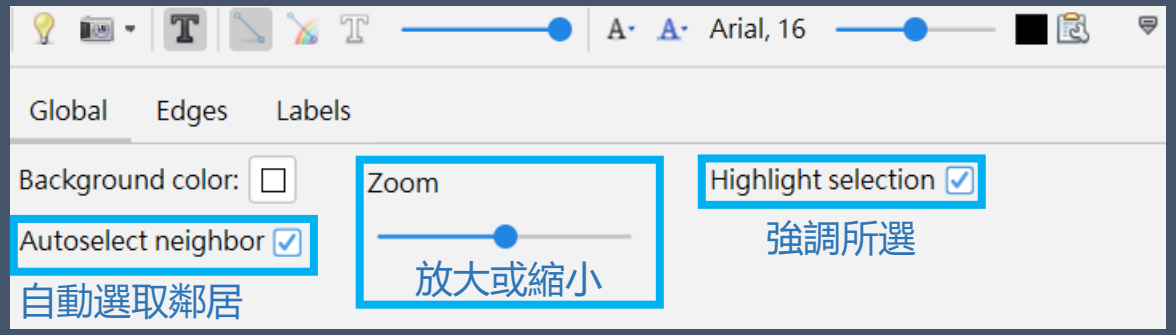


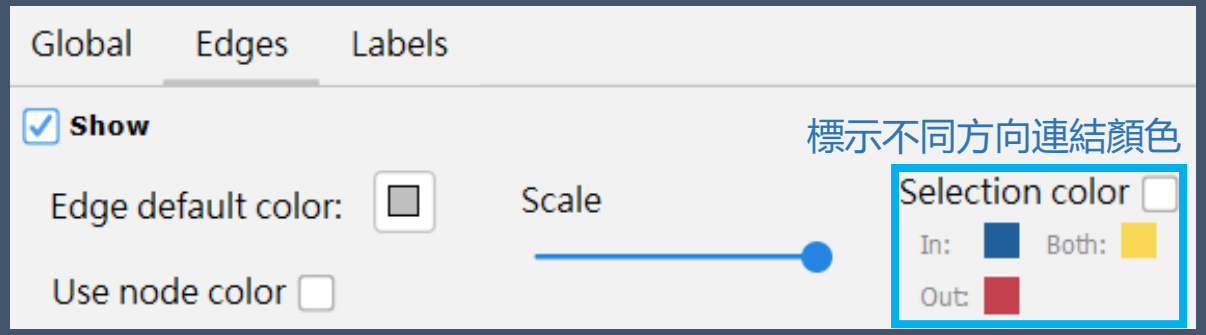
顏色統一顏色
與來源節點相同色

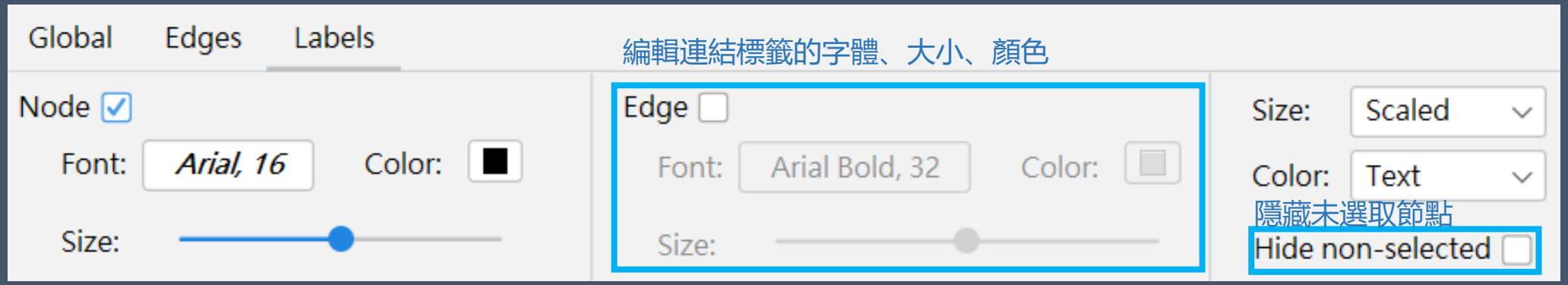




更多選項

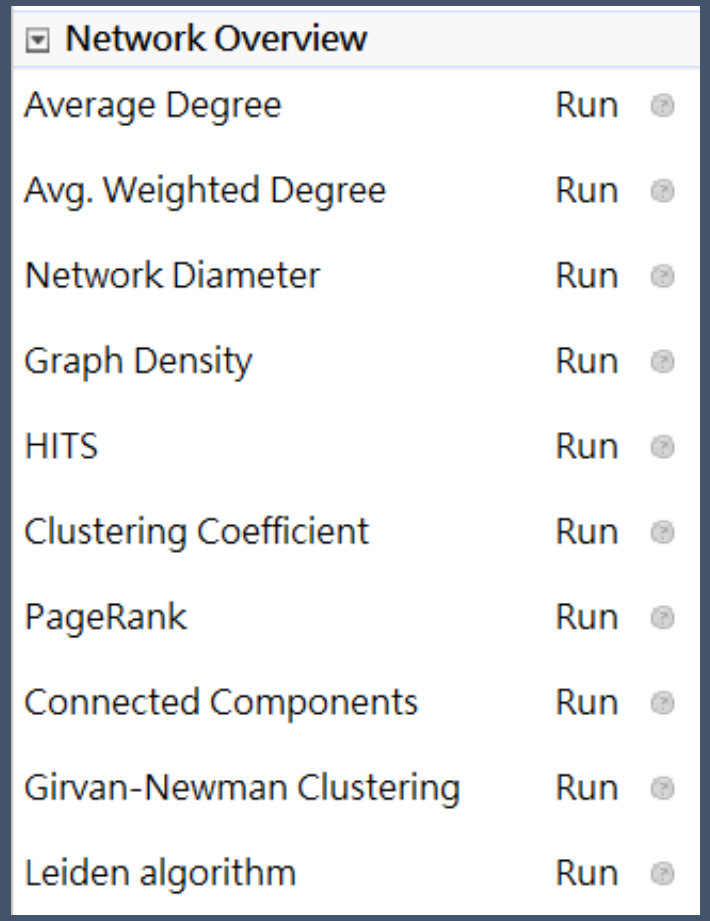

Arial, 16
Global
Edges
Labels
Background color:
Zoom
Highlight selection
Autoselect neighbor
強調所選
放大或縮小
自動選取鄰居


Global
Edges
Labels
Show
標示不同方向連結顏色
Edge default color:
Scale
Selection color
In:
Both:
Out:
Use node color

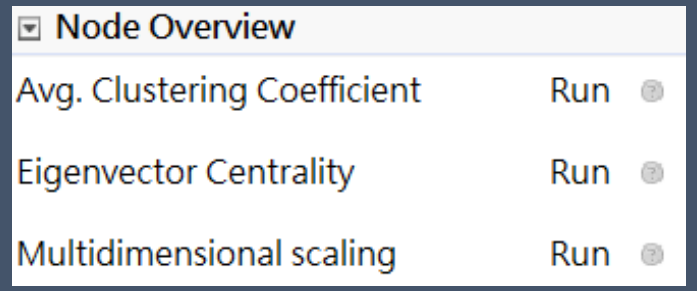

Global
Edges
Labels
編輯連結標籤的字體、大小、顏色
Node
Font:
Arial, 16
Color:
Size:
Edge
Font:
Arial Bold, 32
Color:
Size:
Size:
Scaled
Color:
Text
隱藏未選取節點
Hide non-selected

Gephi軟體入門
Overview 4.4 Statistic

網絡概覽

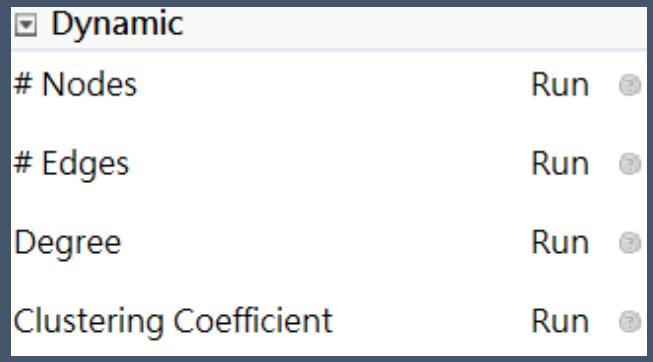


節點概覽



連結概覽

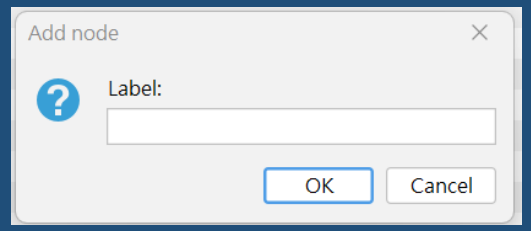
動態資料



Gephi軟體入門
05 Data laboratory & Preview

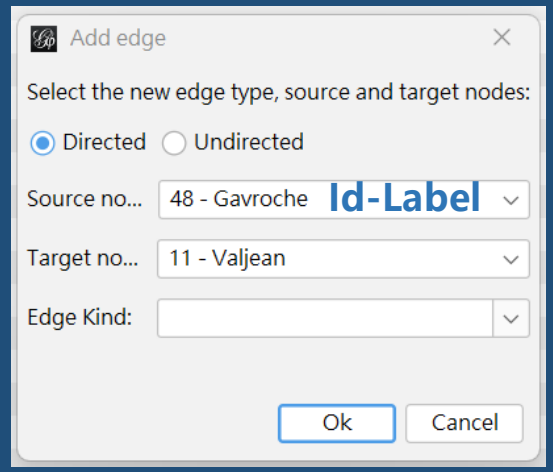
5.1 Data laboratory

Add node 新增節點

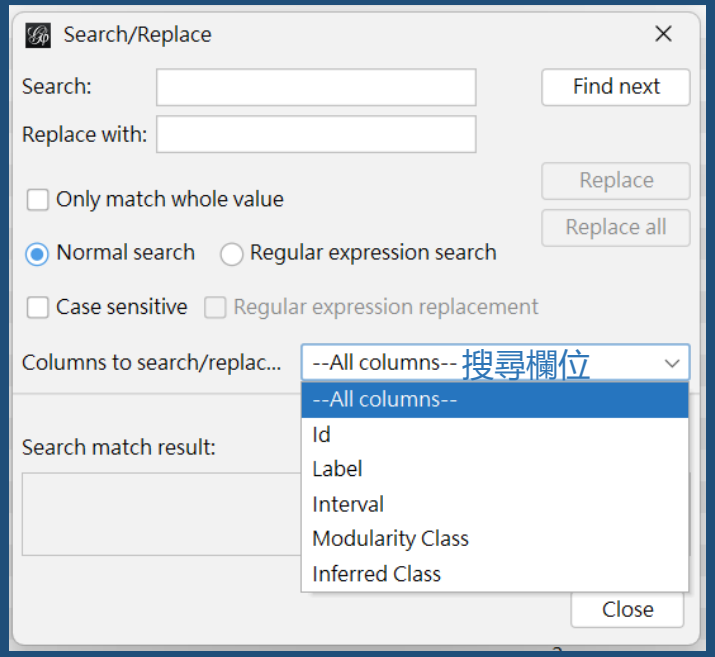


輸入標籤就會新增節點
之後自動產生流水號 Id

Add edge 新增連結



Search/Replace 搜尋與取代



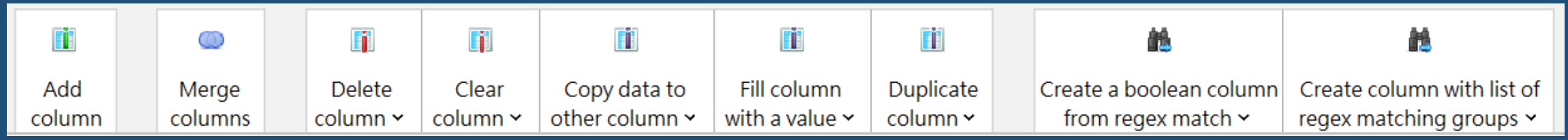
Import Spreadsheet　Export table

匯入Excel分頁
格式：xlsx、csv

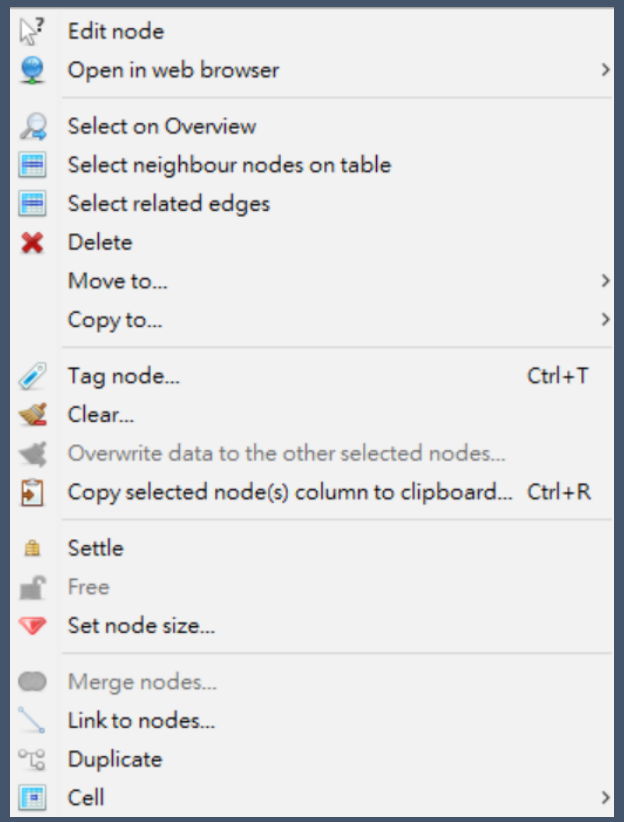
匯出表格資料
格式：csv

篩選特定節點　　選擇欄位

合併欄位　　清除欄位　　填滿欄位

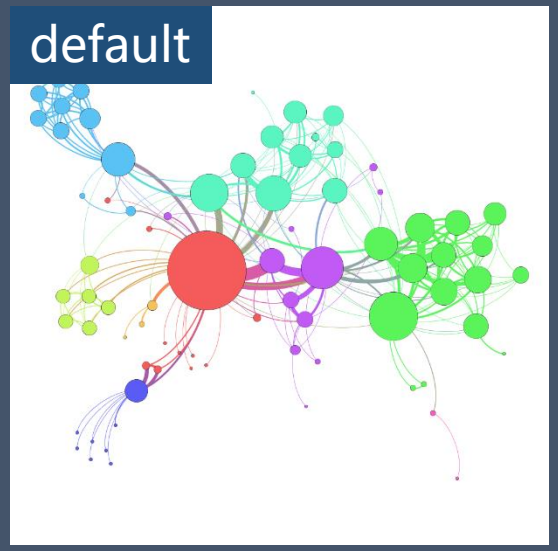


新增欄位　　刪除欄位　　複製資料 Id→Label　　複製欄位

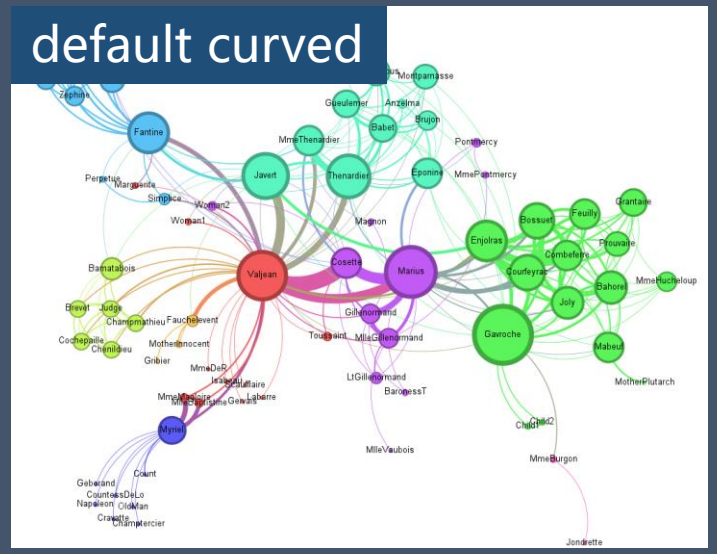


5.2 Preview

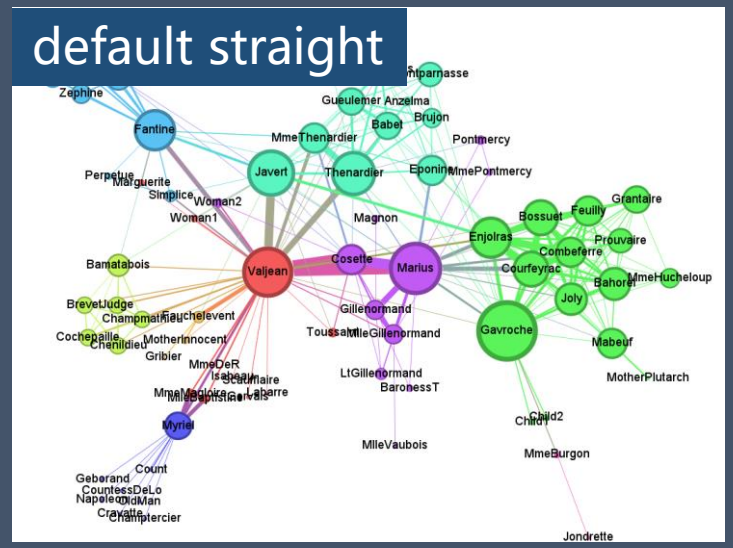
default



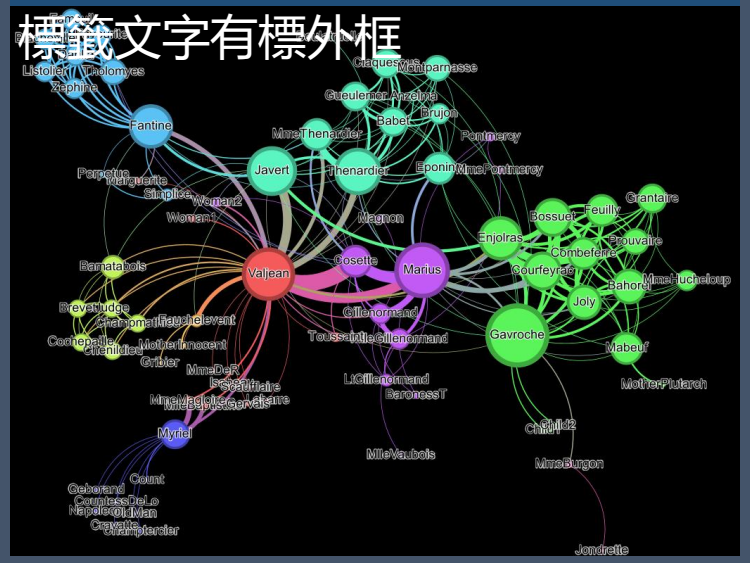
default curved



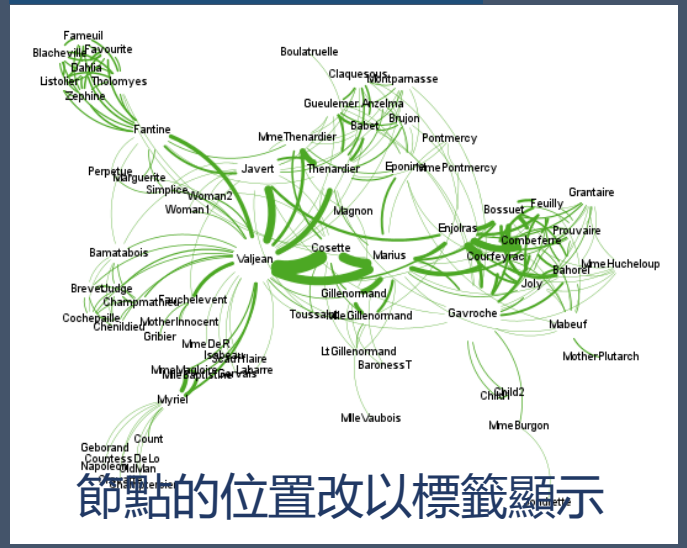
default straight



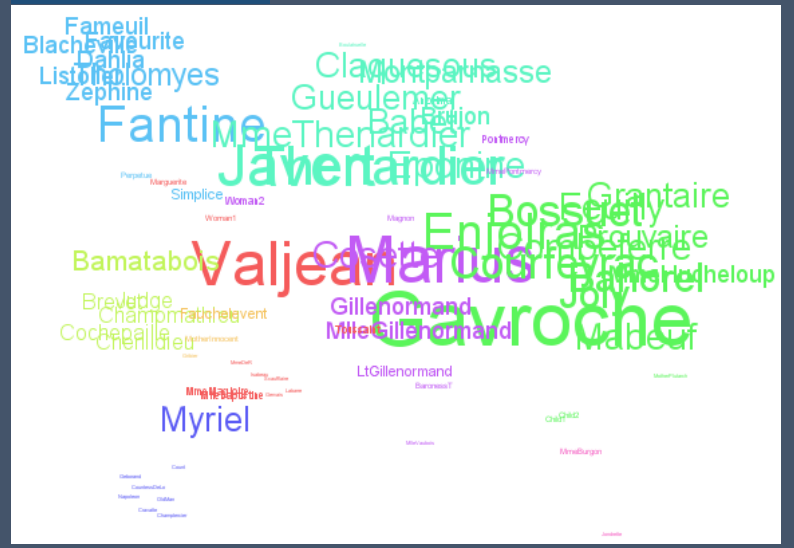
text outline + black background

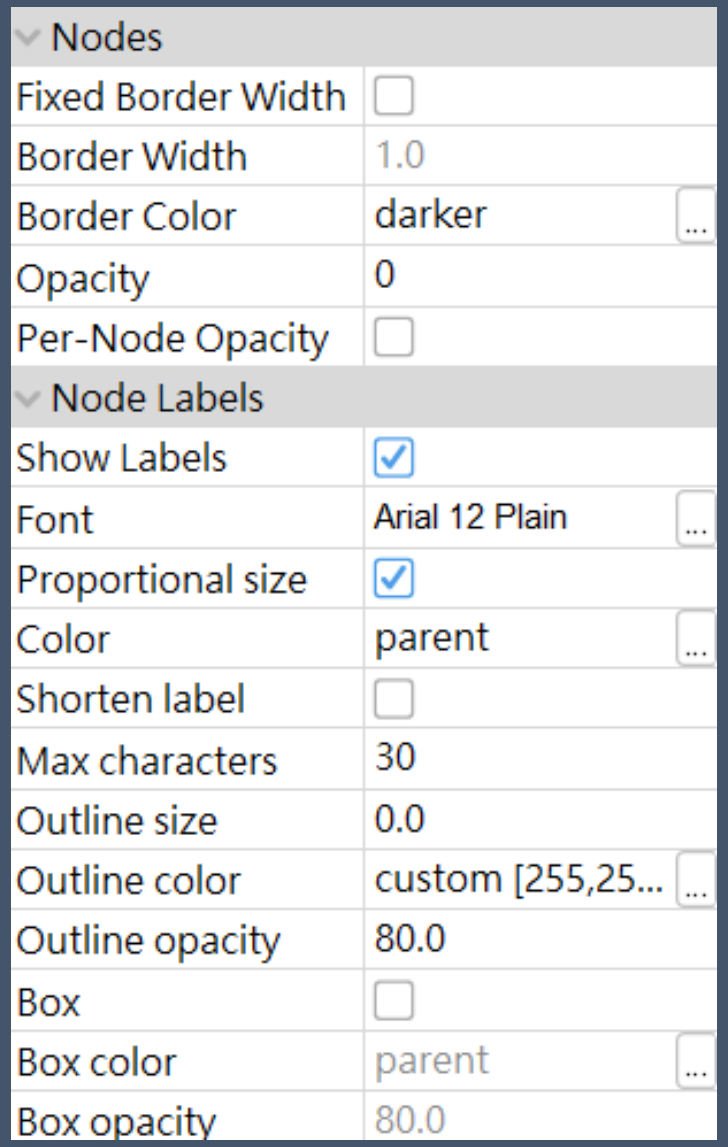


edge custom color



tag cloud



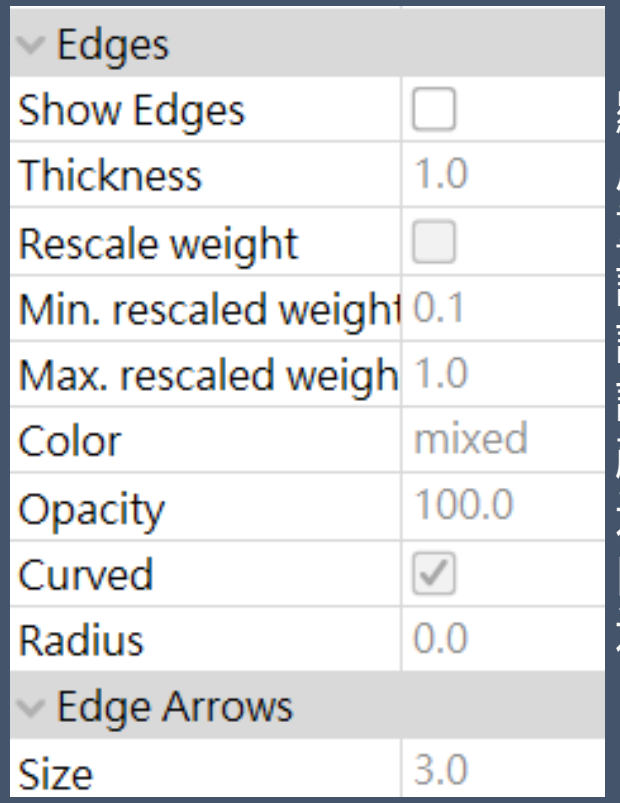


節點

固定邊寬
邊框框度
邊框顏色
透明度
個別節點透明度

節點標籤

顯現標籤
字體(字形、大小)
比例大小
顏色
縮短標籤
縮短字元範圍
輪廓大小
輪廓顏色
輪廓透明度
標籤框
標籤框顏色
標籤框透明度

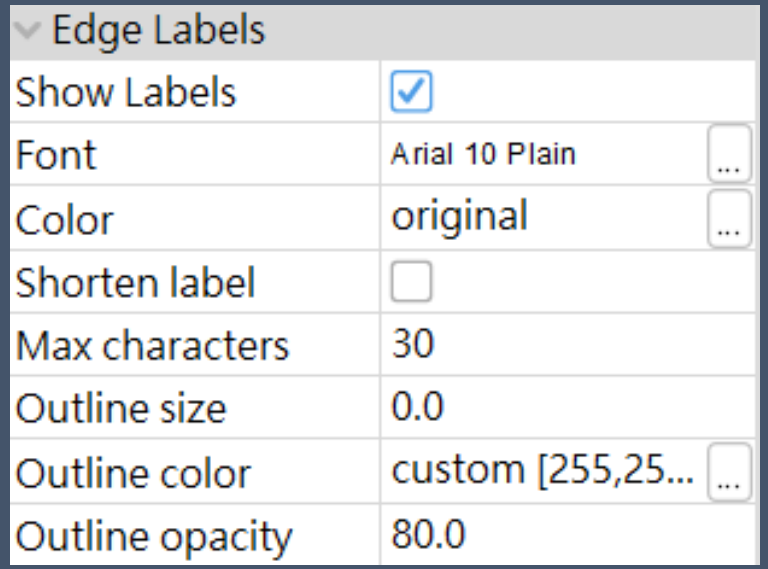


連結

顯現連結
厚度
重新調整權重
調整權重
調整權重最小值
調整權重最大值
顏色
透明度
曲線
連結節點

節點箭頭

大小

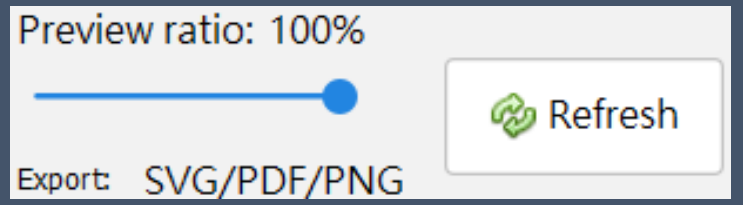


連結標籤

顯現標籤
字體(字形、大小)
顏色
縮短標籤
最大字體
輪廓大小
輪廓顏色
輪廓透明度

輸出圖像

顯示多少比例的節點



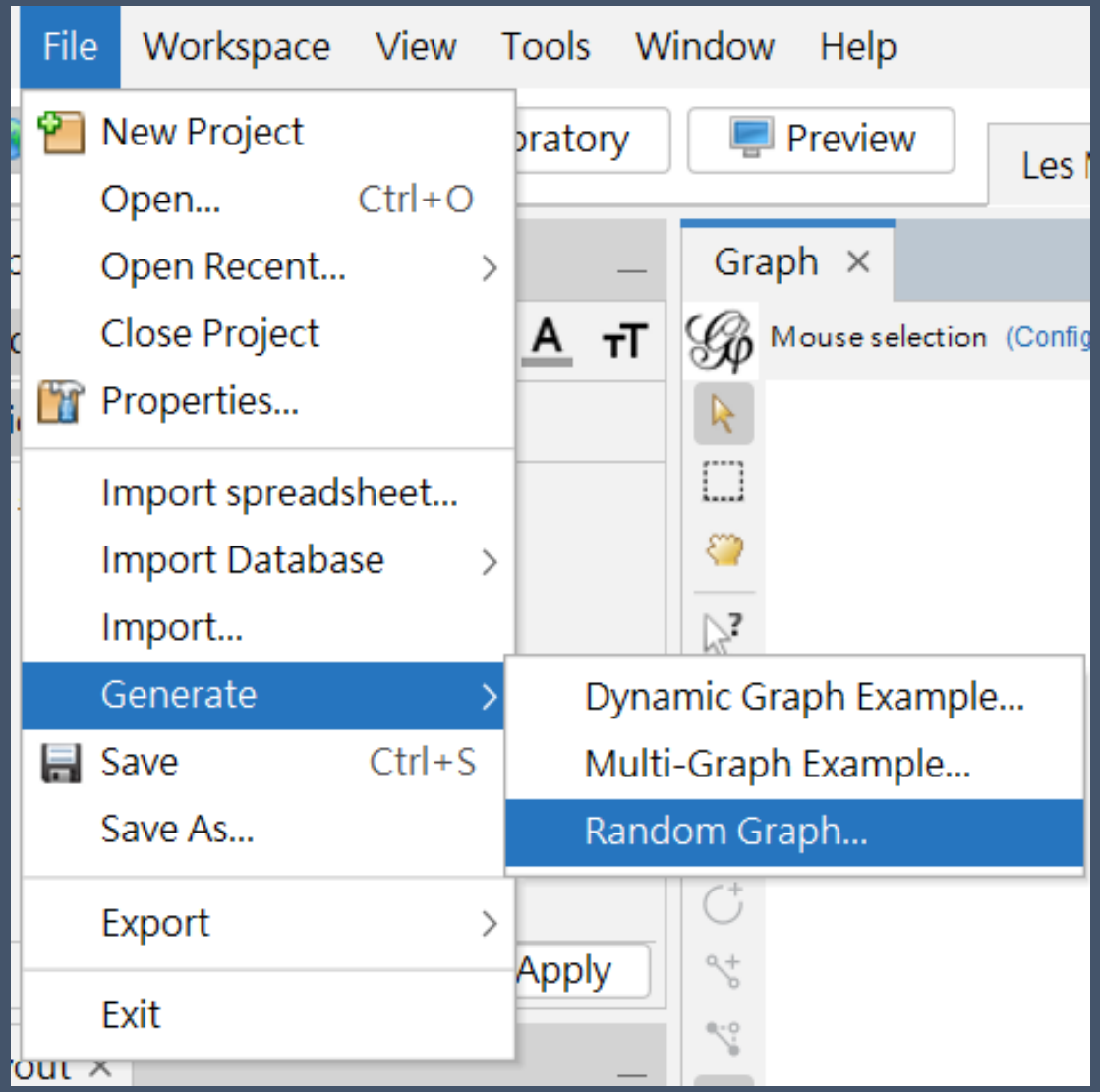
刷新

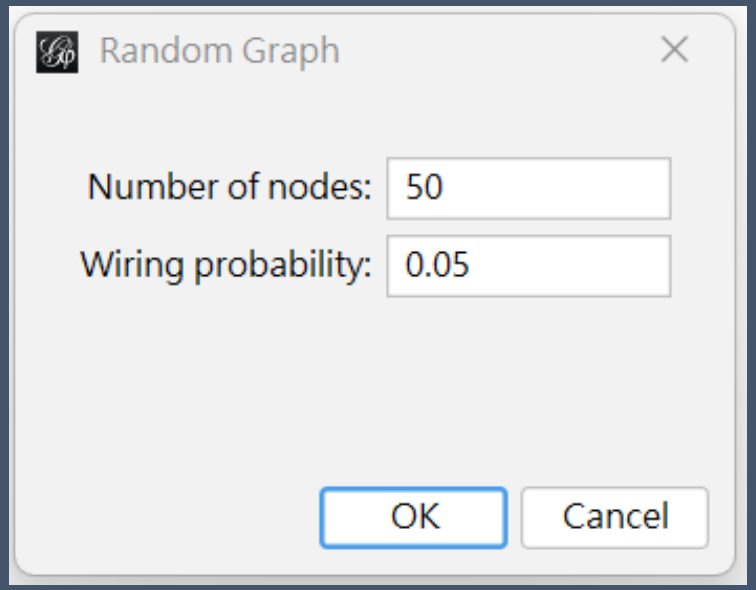
儲存的檔案格式
SVG 可縮小放大的向量檔

Gephi實務操作
06資料匯入

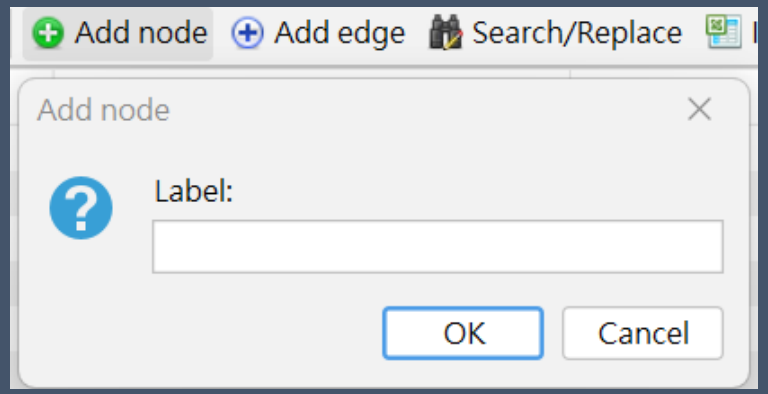
6.1 隨機產生
6.2 單筆建置
6.3 Import Spreadsheet
6.4 Export and Re-import
6.5 Import (Wizard)

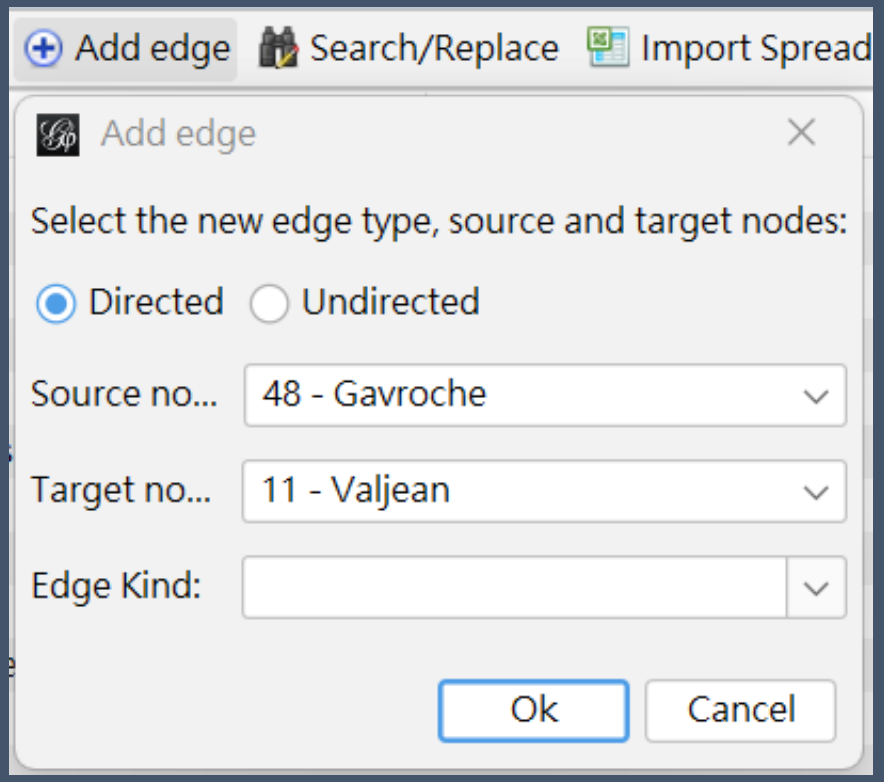
6.1 隨機產生有向圖


File
Workspace
View
Tools
Window
Help
New Project
Open... Ctrl+O
Open Recent...
Close Project
Properties...
Import spreadsheet...
Import Database
Import...
Generate
Save Ctrl+S
Save As...
Export
Exit
Preview
Graph
Mouse selection
Apply
Dynamic Graph Example...
Multi-Graph Example...
Random Graph...


Random Graph
Number of nodes: 50
Wiring probability: 0.05
OK
Cancel

6.2 單筆建置


Add node
Add edge
Search/Replace
Add node
Label:
OK
Cancel


Add edge
Search/Replace
Import Spread
Add edge
Select the new edge type, source and target nodes:
Directed
Undirected
Source no...
48 - Gavroche
Target no...
11 - Valjean
Edge Kind:
Ok
Cancel

Gephi Excel練習匯入檔案

常見匯入教學用Excel CSV檔案匯入

→實際遇到問題

1.中文字碼亂碼
2.爬梳資料隱藏字元、
3.一旦用Excel編修過後檔案編碼會變成ANSI
需用記事本另存UTF-8格式，以致編修不易

→建議作法

直接用Excel建立節點檔和連結檔分別匯入

Gephi Excel匯入資料格式

節點 Node table

欄位	Id	Label	Type
名稱	序號	節點標籤	節點類型
欄數	一組(必要)	一組	可有多組
內容	數字或文字	數字或文字	數字或文字
範例	0	TOM	M
	1	MARRY	F
	2	JENNY	F
	3	TIM	M

Interval 動態資料欄位

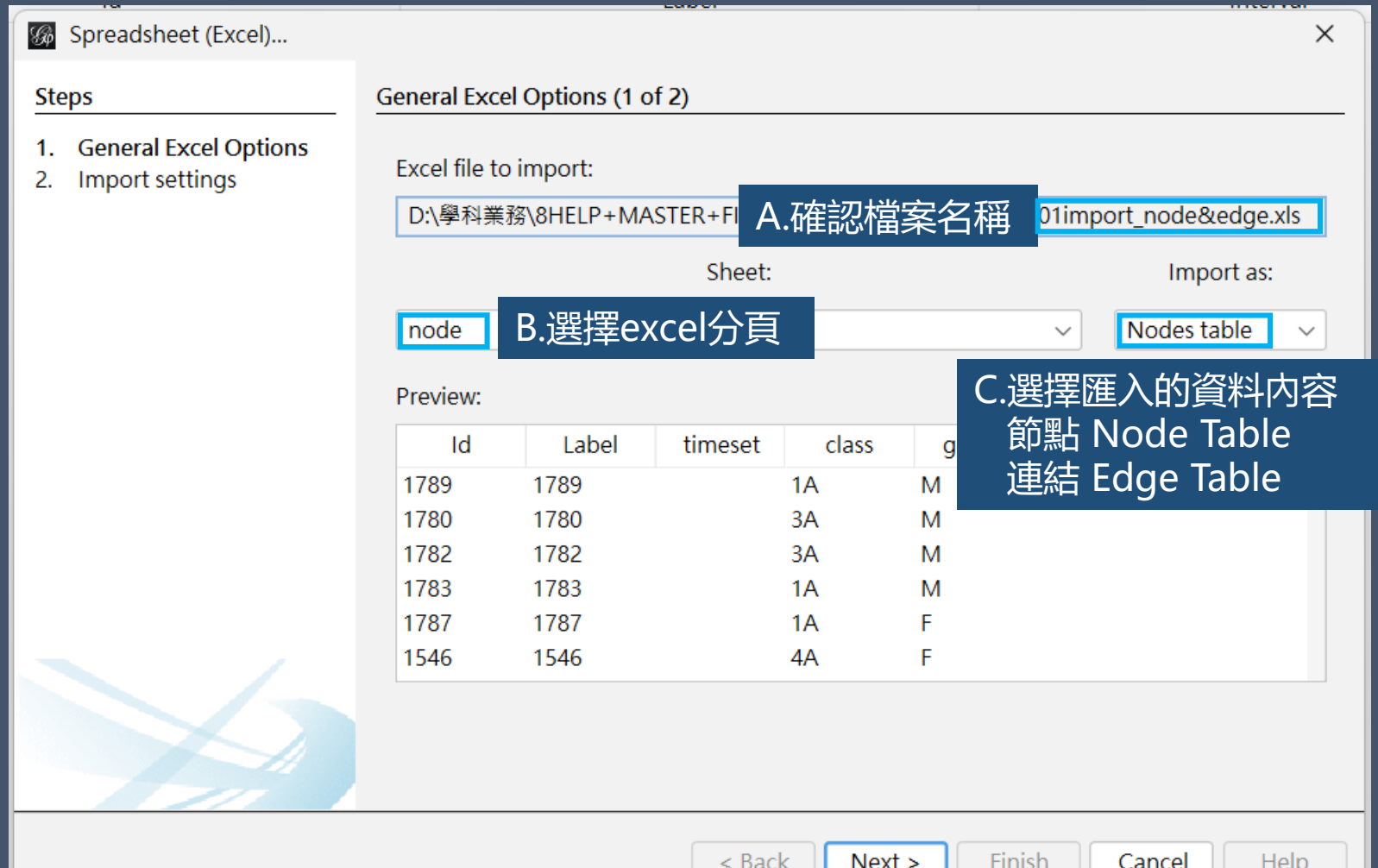
連結 Edge table

欄位	Source	Target	Kind
名稱	來源	目標	種類
欄數	一組(必要)	一組(必要)	一組(非必要)
範例	0	1	Like
	1	2	Comment
	2	3	Share
	3	1	Like
	1	2	Share

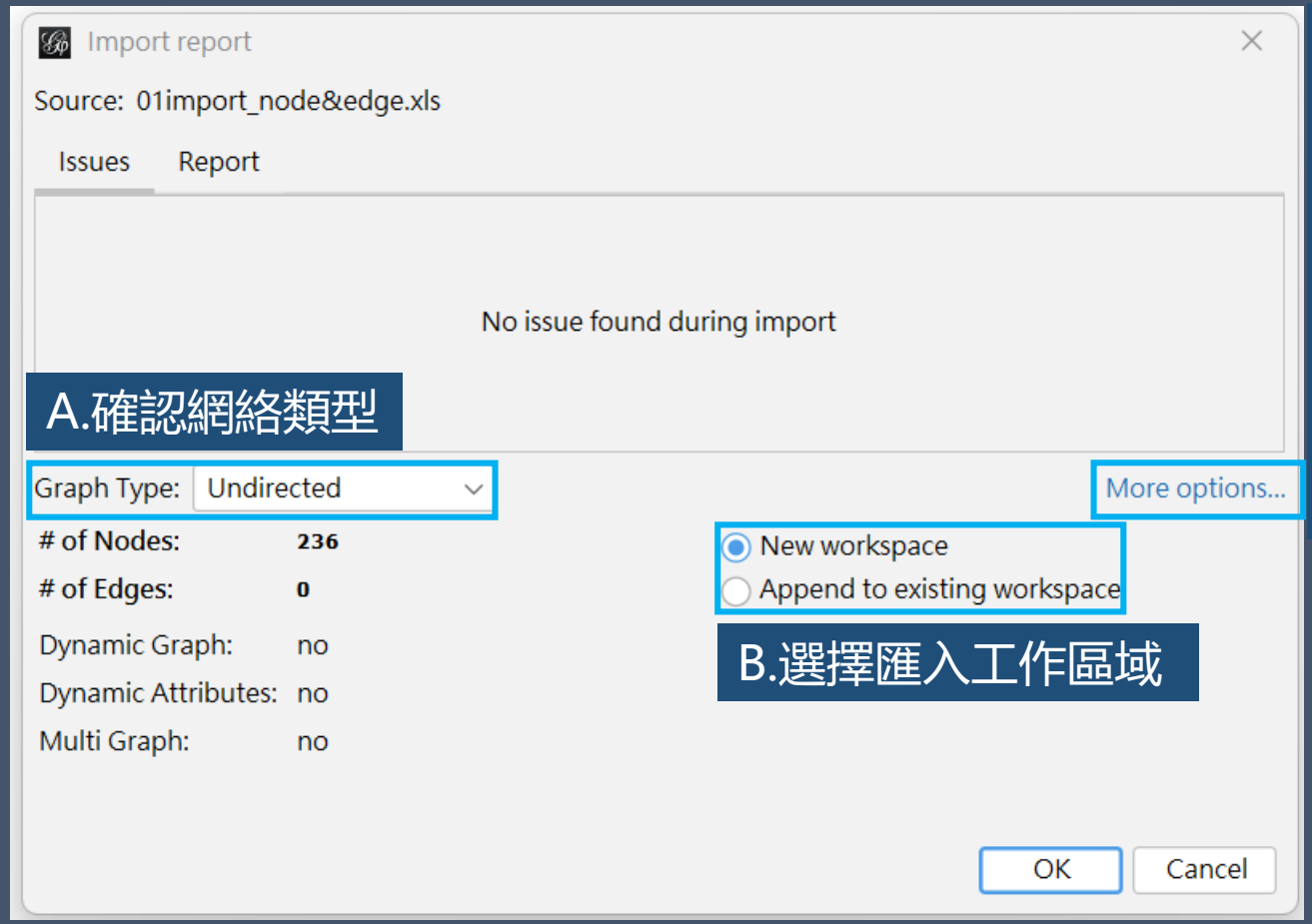
連結若有重複，在匯入時，
Gephi可以自動計算處理，
匯入後數值會顯示在權重欄位

6.3 Import Spreadsheet 匯入資料練習

新專案 >匯入 Node & Edge
確認匯入檔案名稱與分頁



6.3 Import Spreadsheet 匯入資料練習

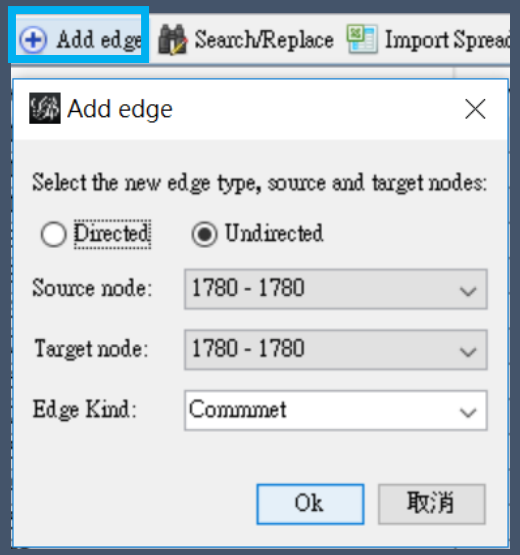


C.連結合併策略
加總 *
平均 *
最大值
最小值
第一個為主
最後一個為主
不要合併 *

*表示常用

Try it!
開新分頁命名school day，將01import_node&edge.xlsx練習資料匯入

6.4 Export and Re-import

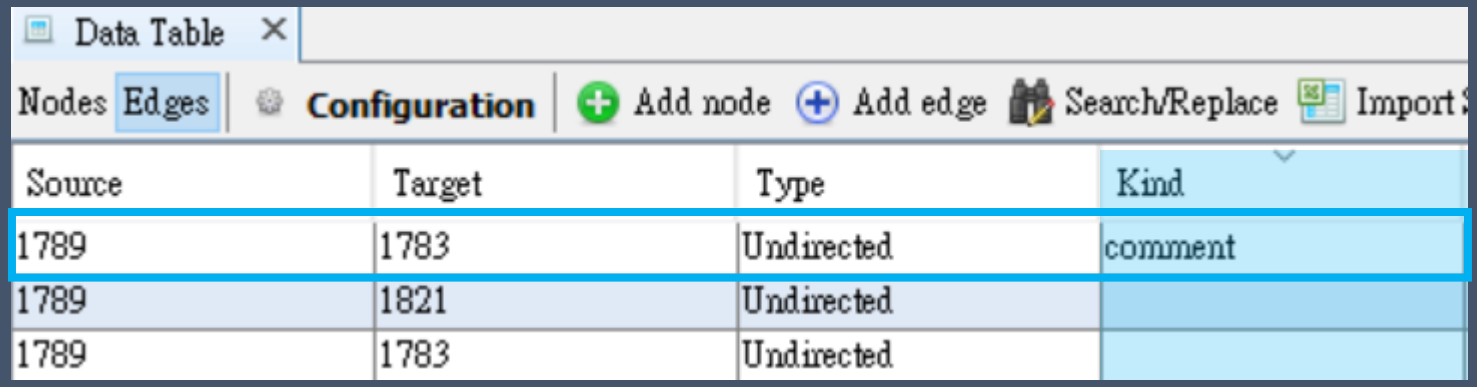


新增單一連結的種類之後，
其餘連結的種類要如何新增?

① [Gephi] Export Table
> School edge csv

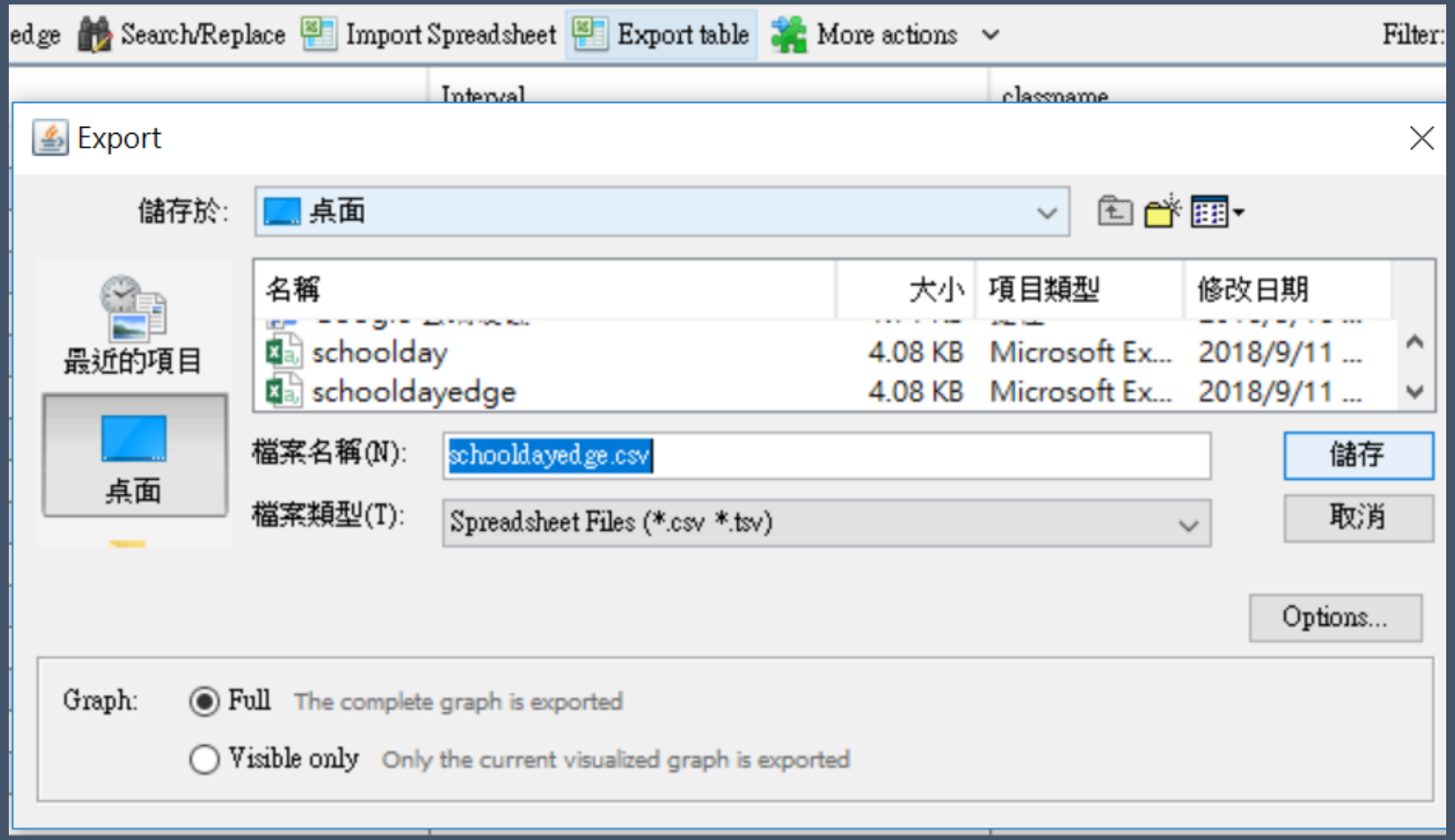
② [Excel] Fulfill Kind Column
>Save as xlsx

③ [Gephi] Import Edge List
> make sure that the number of edges is correct
> append to existing workspace



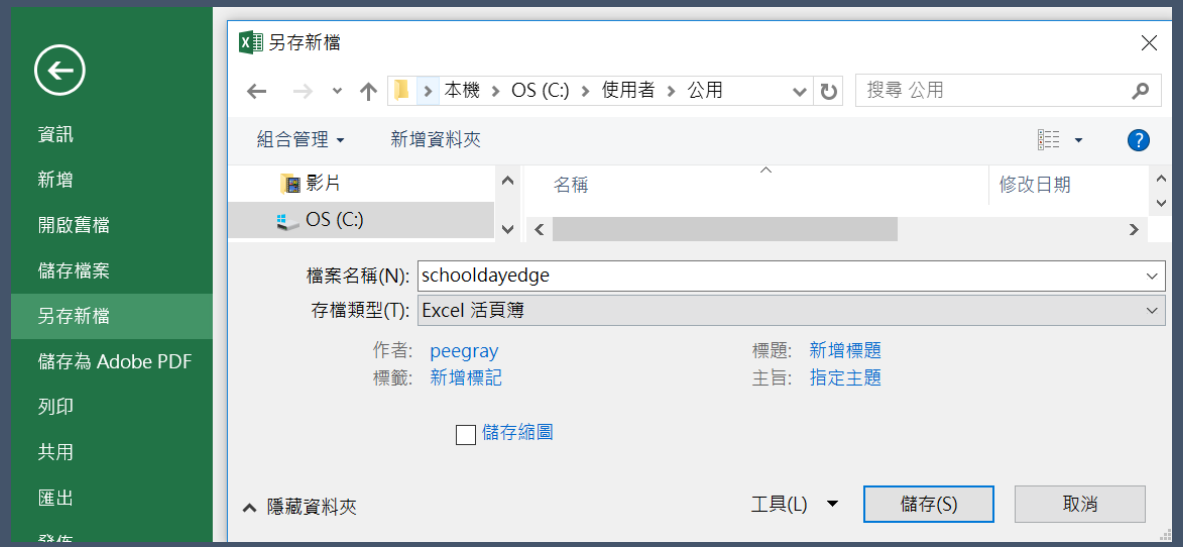
Source	Target	Type	Kind
1789	1783	Undirected	comment
1789	1821	Undirected	
1789	1783	Undirected	

① [Gephi] Export Table
> School edge csv

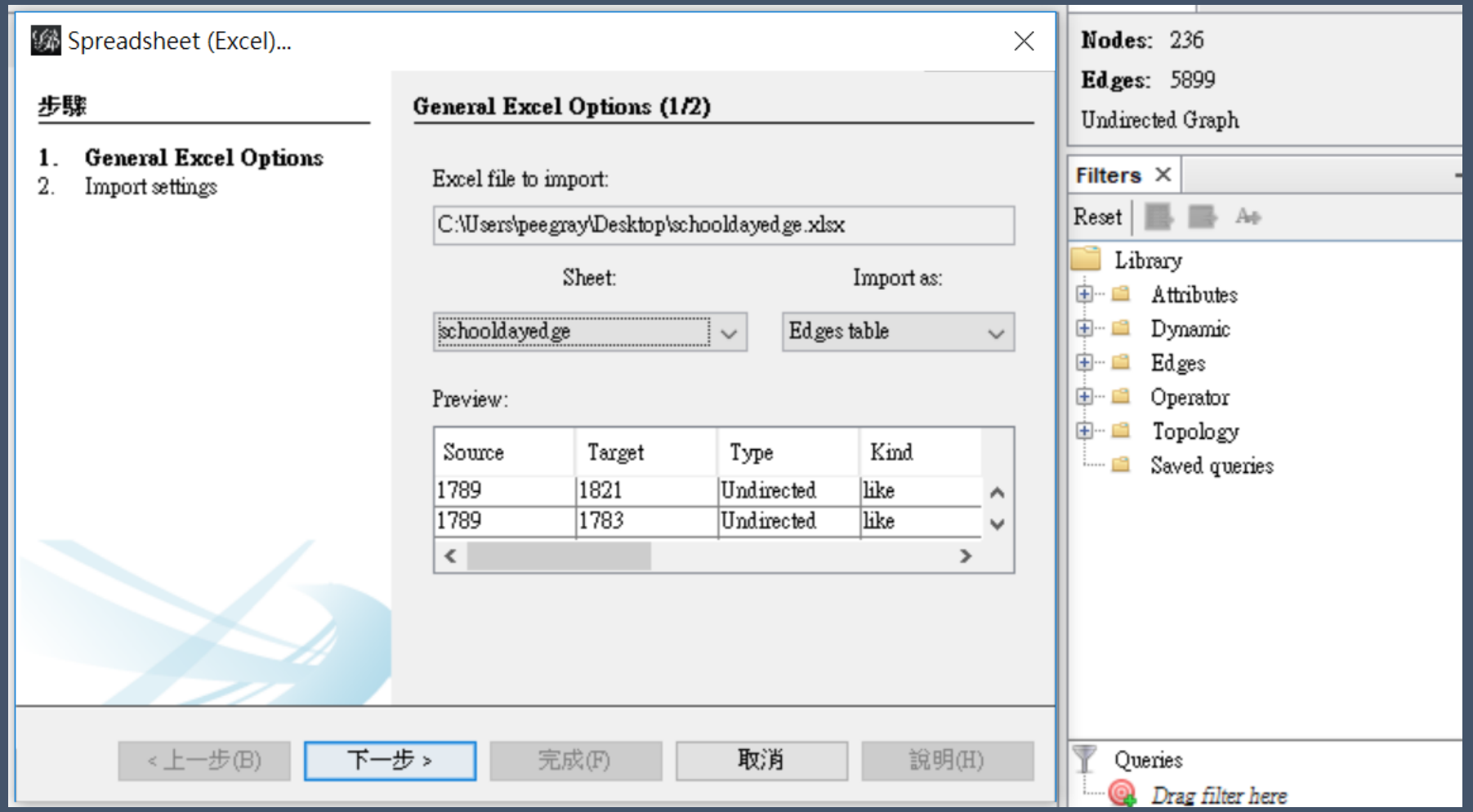


② [Excel] Fulfill Kind Column >Save as xlsx

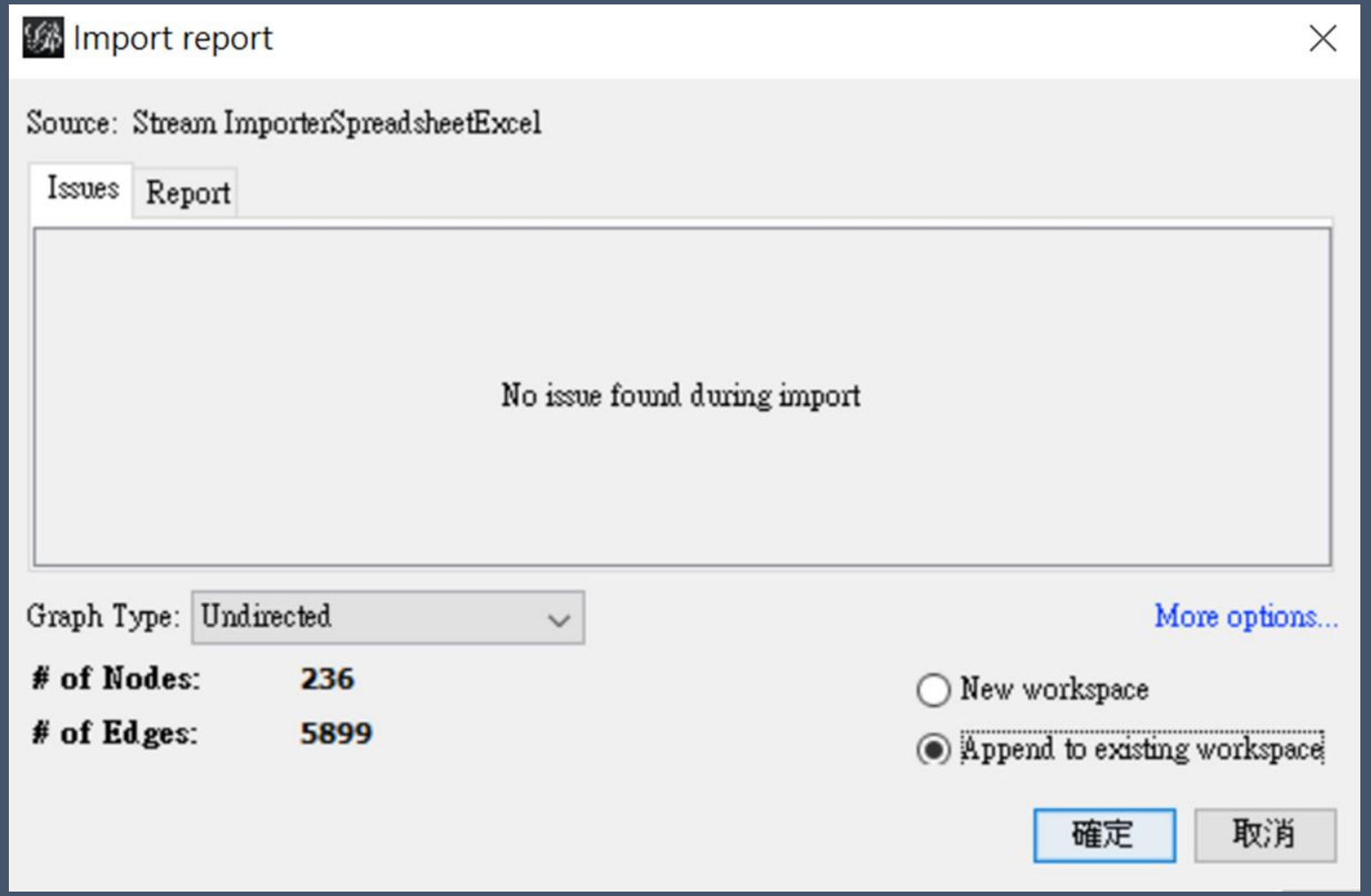
1	Id	Source	Target	Kind		Label	timeset	Weight
2	0	1789	1821	like				1
3	1	1789	1783	like				1
4	2	1789	1778	like				1
5	3	1789	1722	like				1
6	4	1789	1902	like				1
7	5	1789	1892	like				1
8	6	1789	1775	like				1
9	7	1789	1774	like				1
10	8	1789	1711	like				1
11	9	1789	1770	like				1
12	10	1789	1772	like				1
13	11	1789	1732	like				1
14	12	1789	1779	like				1
15	13	1789	1613	like				1
16	14	1789	1500	like				1



③ [Gephi] Import Edge List
> make sure that the number of edges is correct
> append to existing workspace



③ [Gephi] Import Edge List
> make sure that the number of edges is correct
> append to existing workspace



③ [Gephi] Import Edge List
> make sure that the number of edges is correct
> append to existing workspace

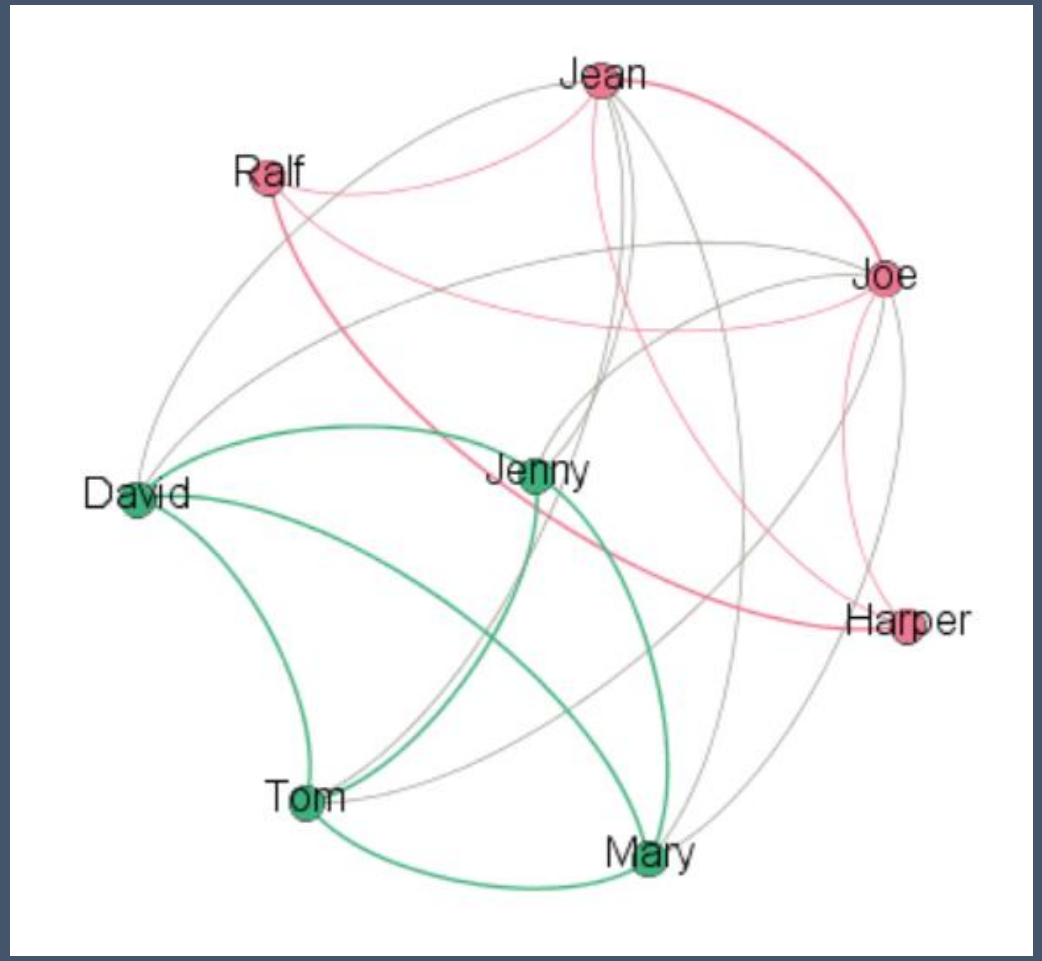
Data Table

Nodes Edges | Configuration | Add node | Add edge | Search/Replace | Import Spreadsheet | Export table | More actio

Source	Target	Type	Kind	Id	Weight	durtion	counts
1789	1778	Undirected	share	5901	1.0	740	23
1789	1732	Undirected	share	5910	1.0	20	1
1789	1821	Undirected	like	5899	1.0	20	1
1789	1783	Undirected	like	5900	1.0	300	13
1789	1722	Undirected	like	5902	1.0	20	1
1789	1902	Undirected	like	5903	1.0	20	1
1789	1775	Undirected	like	5905	1.0	320	13
1789	1774	Undirected	like	5906	1.0	100	
1789	1770	Undirected	like	5908	1.0	20	
1789	1772	Undirected	like	5909	1.0	260	
1789	1779	Undirected	like	5911	1.0	20	
1789	1613	Undirected	like	5912	1.0	60	
1789	1500	Undirected	like	5913	1.0	20	
1789	1759	Undirected	like	5914	1.0	160	
1789	1757	Undirected	like	5915	1.0	1220	32
1789	1653	Undirected	like	5916	1.0	20	1
1789	1753	Undirected	like	5917	1.0	60	3

Try it!
嘗試先匯入Edge檔
看看data laboratory的node檔

6.5 Import-
(1)Data importer (similarities)

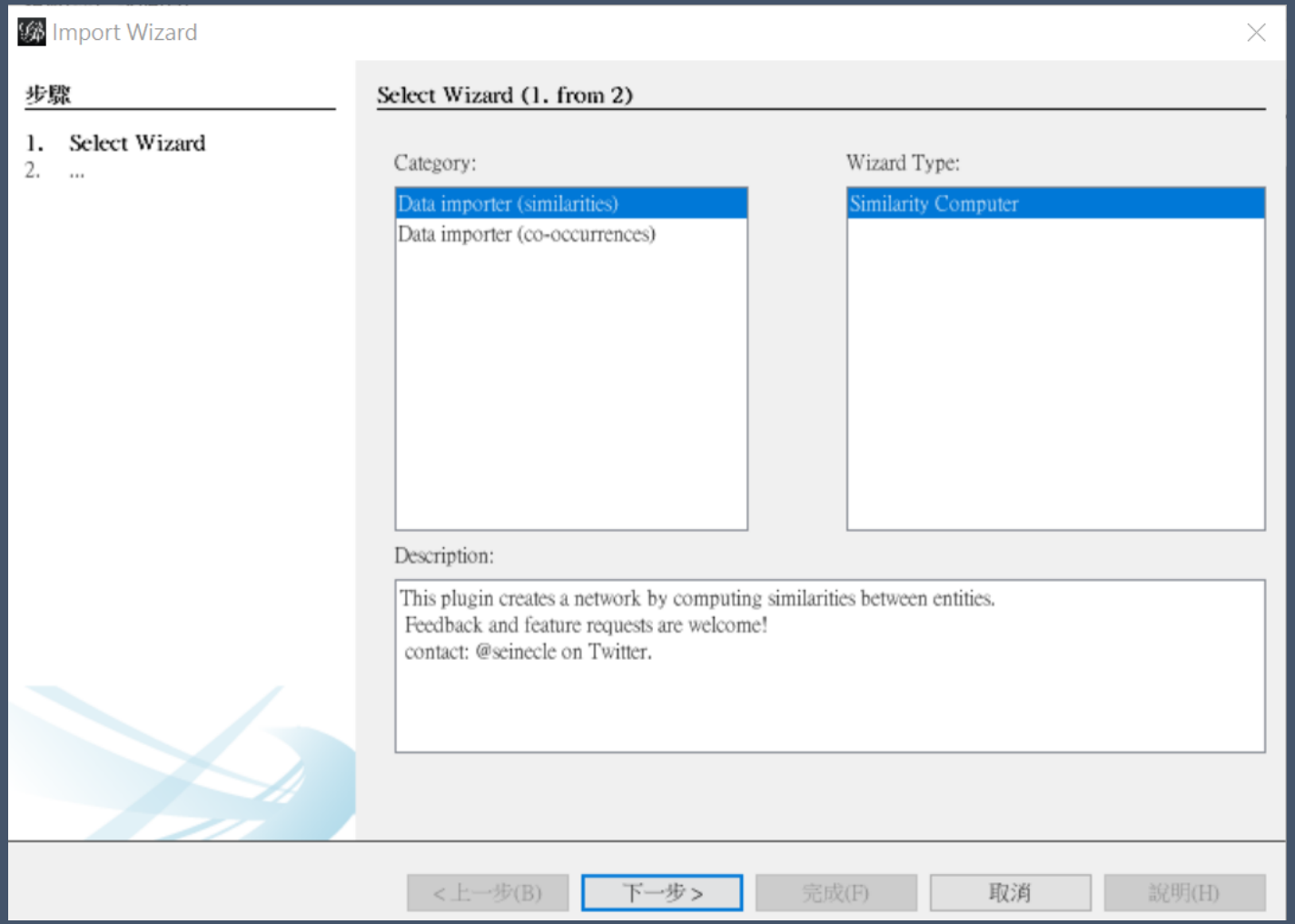


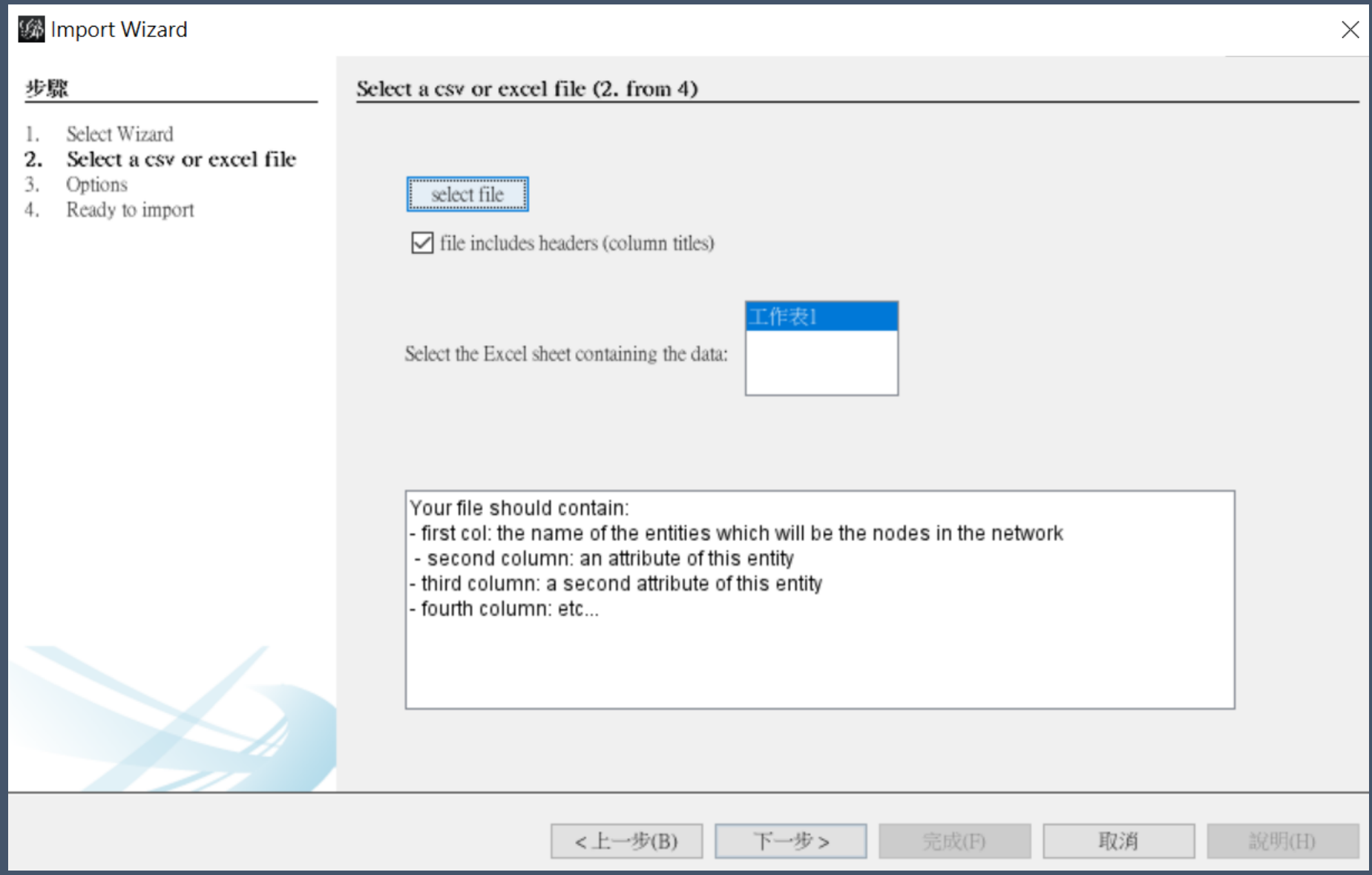
	A	B	C
1	**Author**	**Taste in Ice creams**	**City preference**
2	David	Strawberry	Venice
3	Mary	Strawberry	Venice
4	Jean	Vanilla	Venice
5	Ralf	Vanilla	Paris

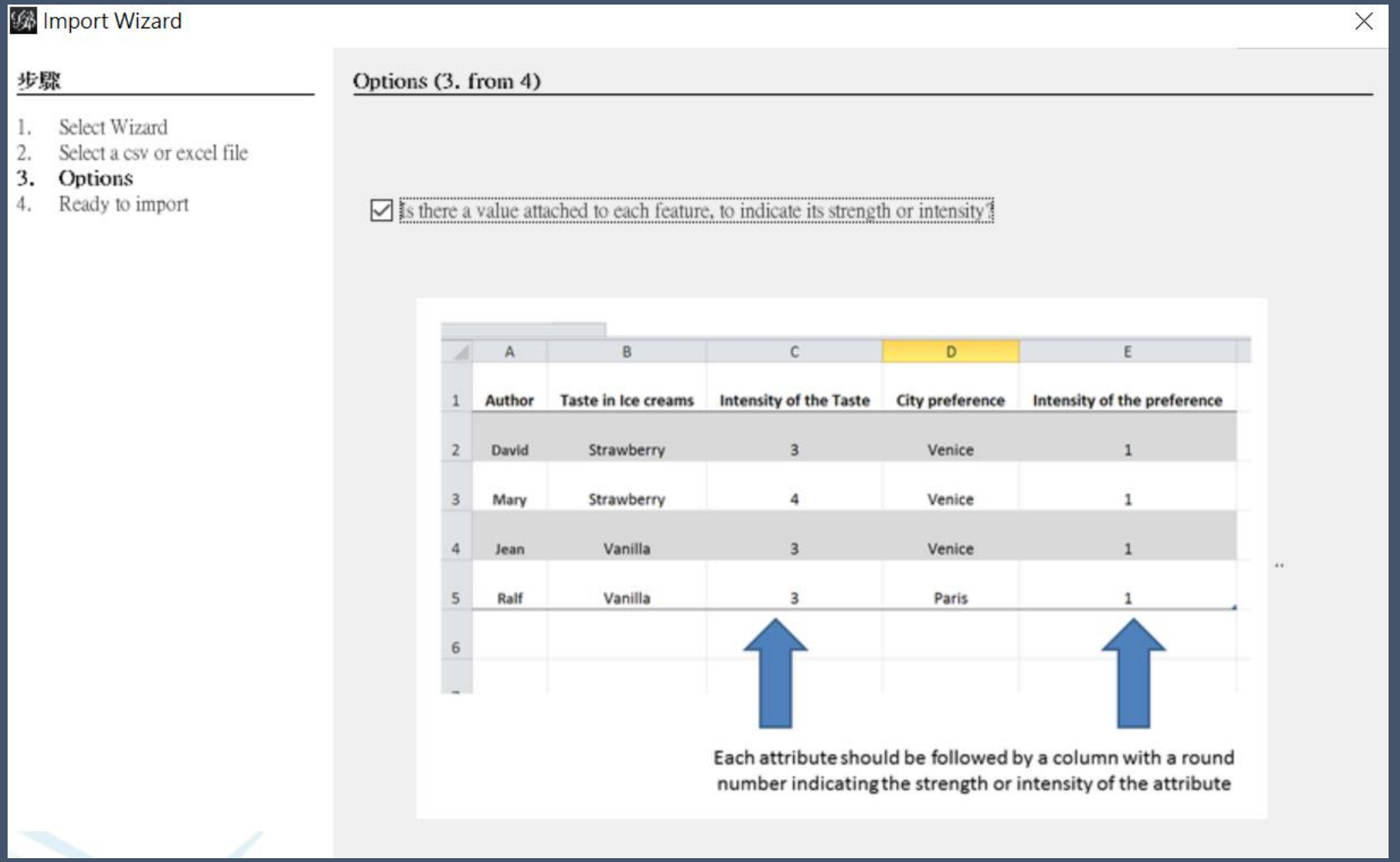
	A	B	C	D	E
1	**Author**	**Taste in Ice creams**	**Intensity of the Taste**	**City preference**	**Intensity of the preference**
2	David	Strawberry	3	Venice	1
3	Mary	Strawberry	4	Venice	1
4	Jean	Vanilla	3	Venice	1
5	Ralf	Vanilla	3	Paris	1
6					

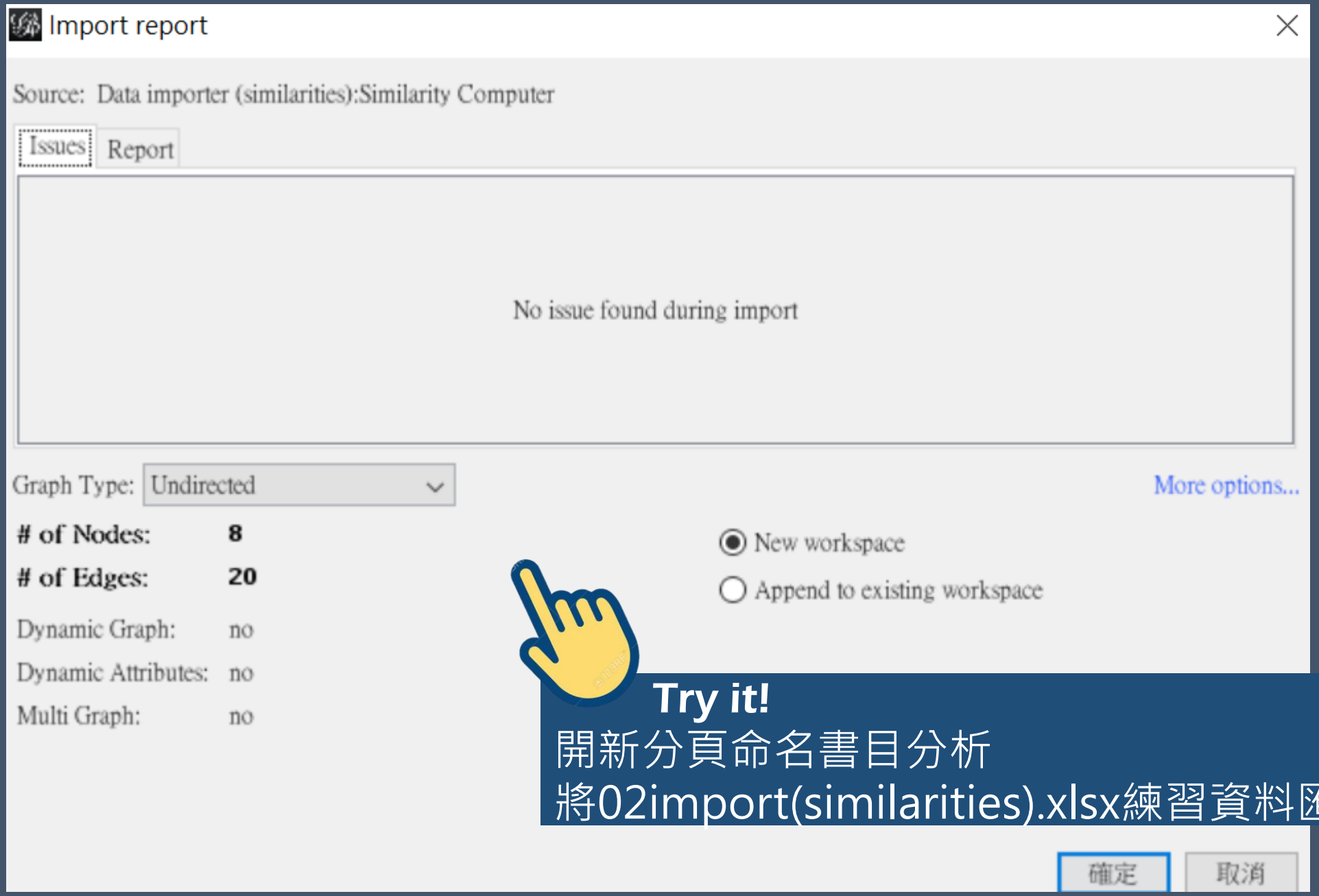
Each attribute should be followed by a column with a round number indicating the strength or intensity of the attribute

檔案>Import>Data importer (similarities)




Import Wizard
步驟
1. Select Wizard
2. Select a csv or excel file
3. Options
4. Ready to import
Select a csv or excel file (2. from 4)
select file
file includes headers (column titles)
Select the Excel sheet containing the data:
工作表1
Your file should contain:
- first col: the name of the entities which will be the nodes in the network
- second column: an attribute of this entity
- third column: a second attribute of this entity
- fourth column: etc...
< 上一步(B)
下一步 >
完成(F)
取消
說明(H)


Import Wizard
步驟
1. Select Wizard
2. Select a csv or excel file
3. Options
4. Ready to import
Options (3. from 4)
Is there a value attached to each feature, to indicate its strength or intensity?
A B C D E
1 Author Taste in Ice creams Intensity of the Taste City preference Intensity of the preference
2 David Strawberry 3 Venice 1
3 Mary Strawberry 4 Venice 1
4 Jean Vanilla 3 Venice 1
5 Ralf Vanilla 3 Paris 1
6
Each attribute should be followed by a column with a round number indicating the strength or intensity of the attribute



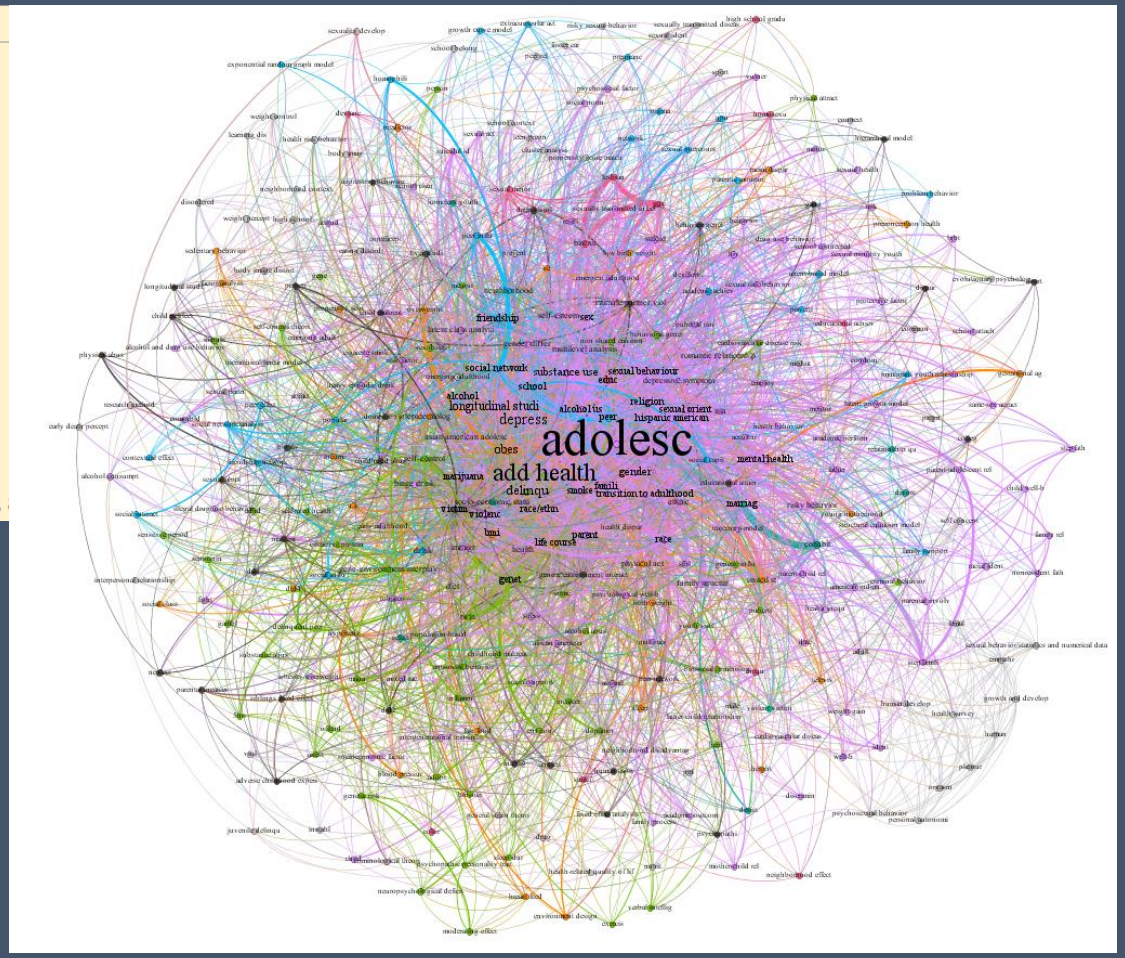
Try it!

開新分頁命名書目分析

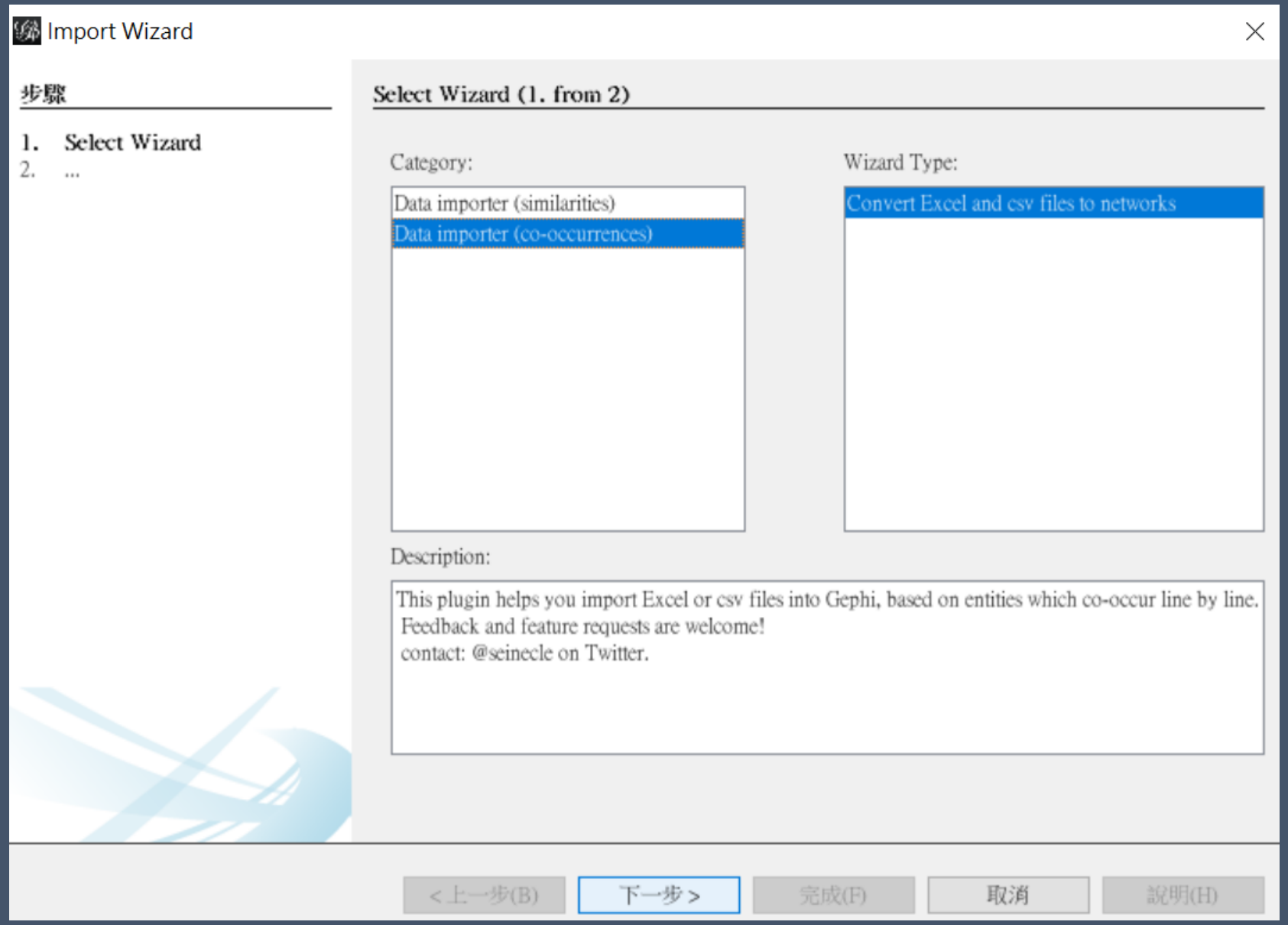
將02import(similarities).xlsx練習資料匯入

6.5 Import-
(2)Data importer (co-occurrences)

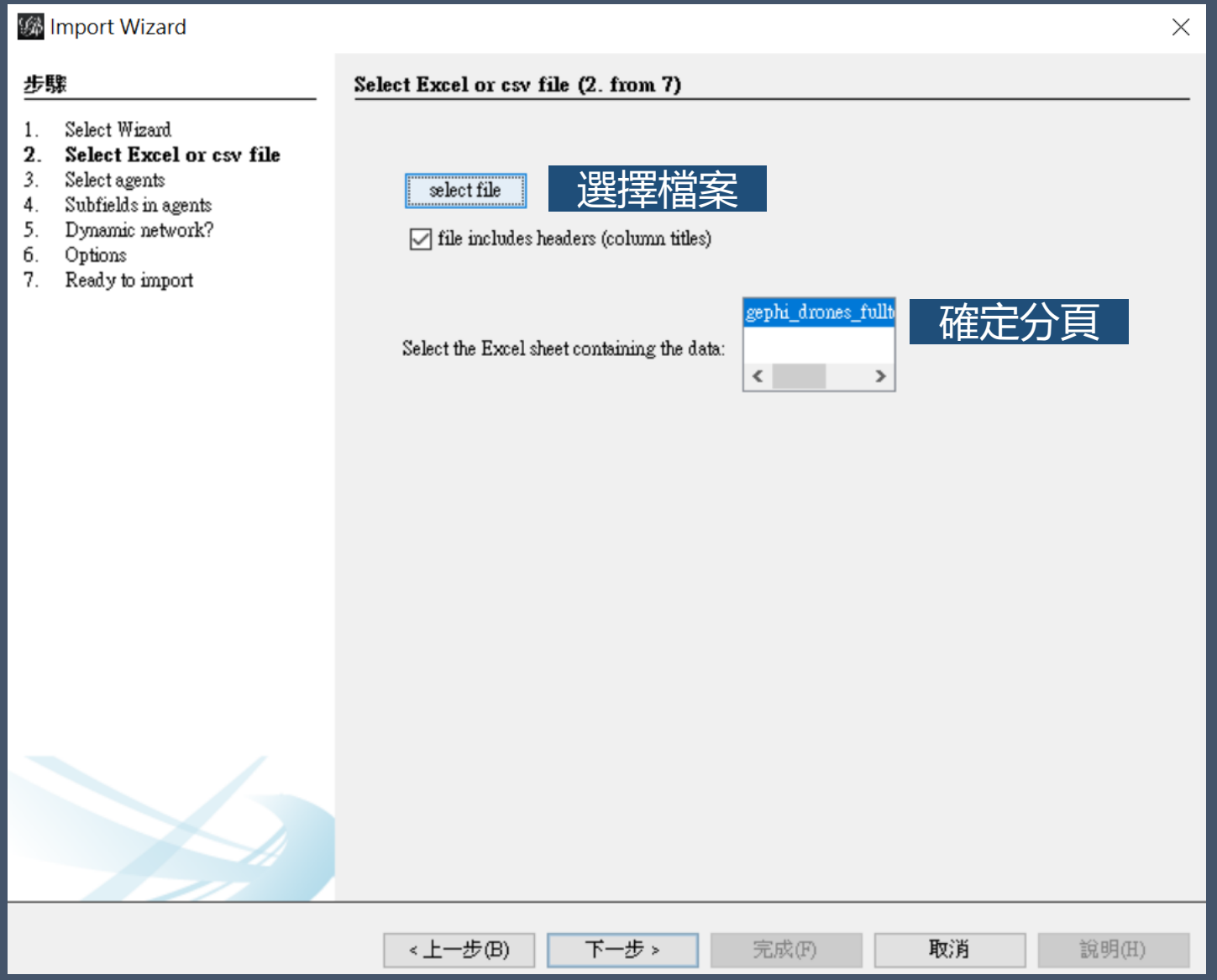
	Authors	Author Keywords (stemming_clean)
1	Yu J.W., Adams S	transition to young adulthood; use of mental health counseling servic
2	Russell S.T., Toor	sexual minor; suicide id; youth
3	Gunn J.F., III, Les	sex differ; sport; suicid
4	Ulloa E., Salazar I	risky sexual behavior; sexual assault; sexually transmitted infect; substance us
5	Manlove J.S., Rya	risk/protective factor; sexual experi; sexual partn; transition to sex
6	Schreck C.J., Bure	religios; routine act; victim
7	Mellor J.M., Freel	religion; substance abus; tobacco; youth
8	Broman C.L.	race and mental health servic; young adult mental health
9	Balistreri K.S., Jo	race and ethnic differ; relationship involv; young adult
10	Lorenzo-Luaces L	public health; race; risk factor; suicid
11	Daigle L.E., Teas	psychopathic trait; psychopathi; replic; revictim; victim
12	Nkansah-Amankra	private religios; public religios; spiritu; suicidal behavior
13	Ehntholt A., Aven	prescription; painkiller substance use; lifecourse; race/ethnicity; multilevel analysi;

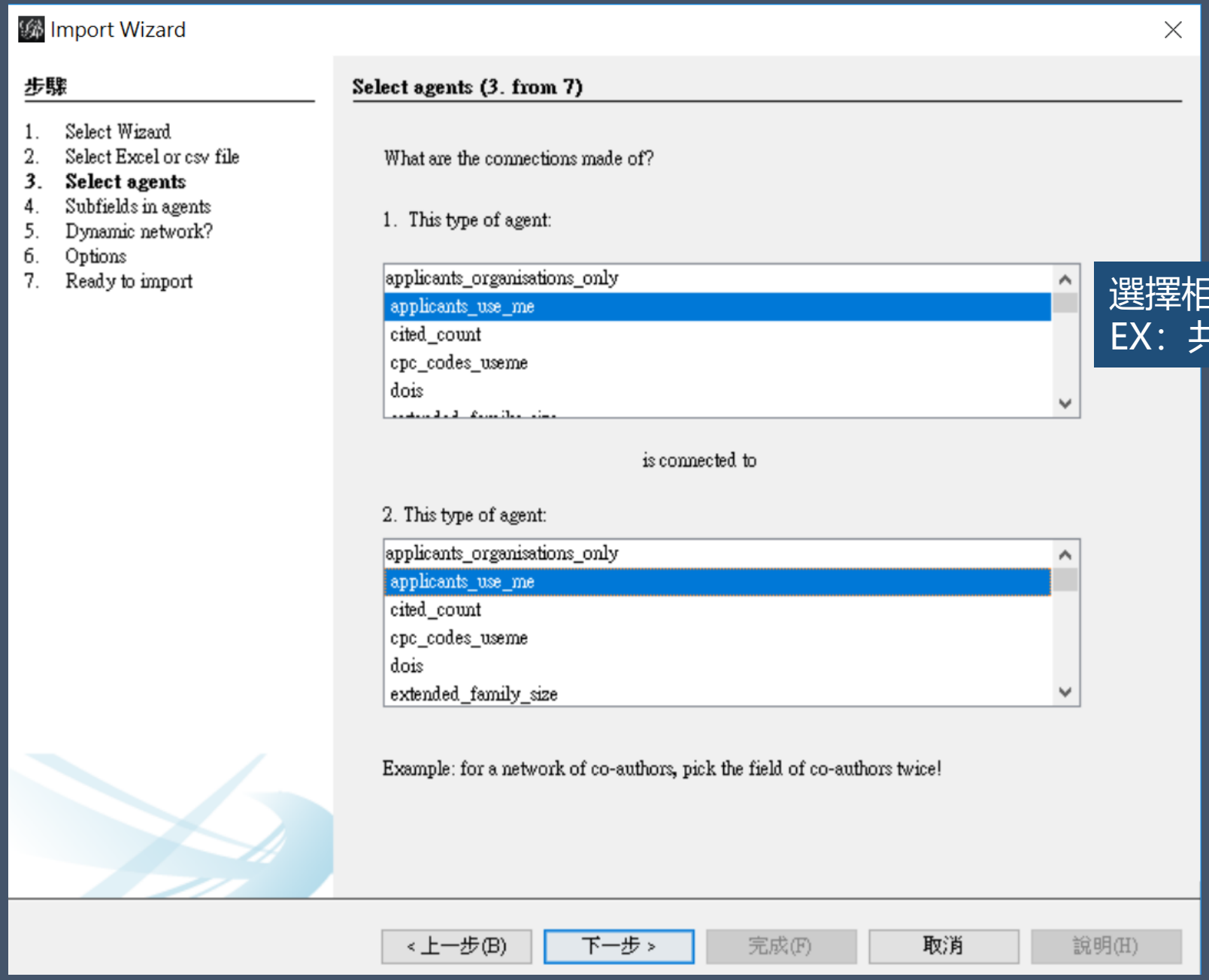


檔案>Import>Data importer (co-occurrences)

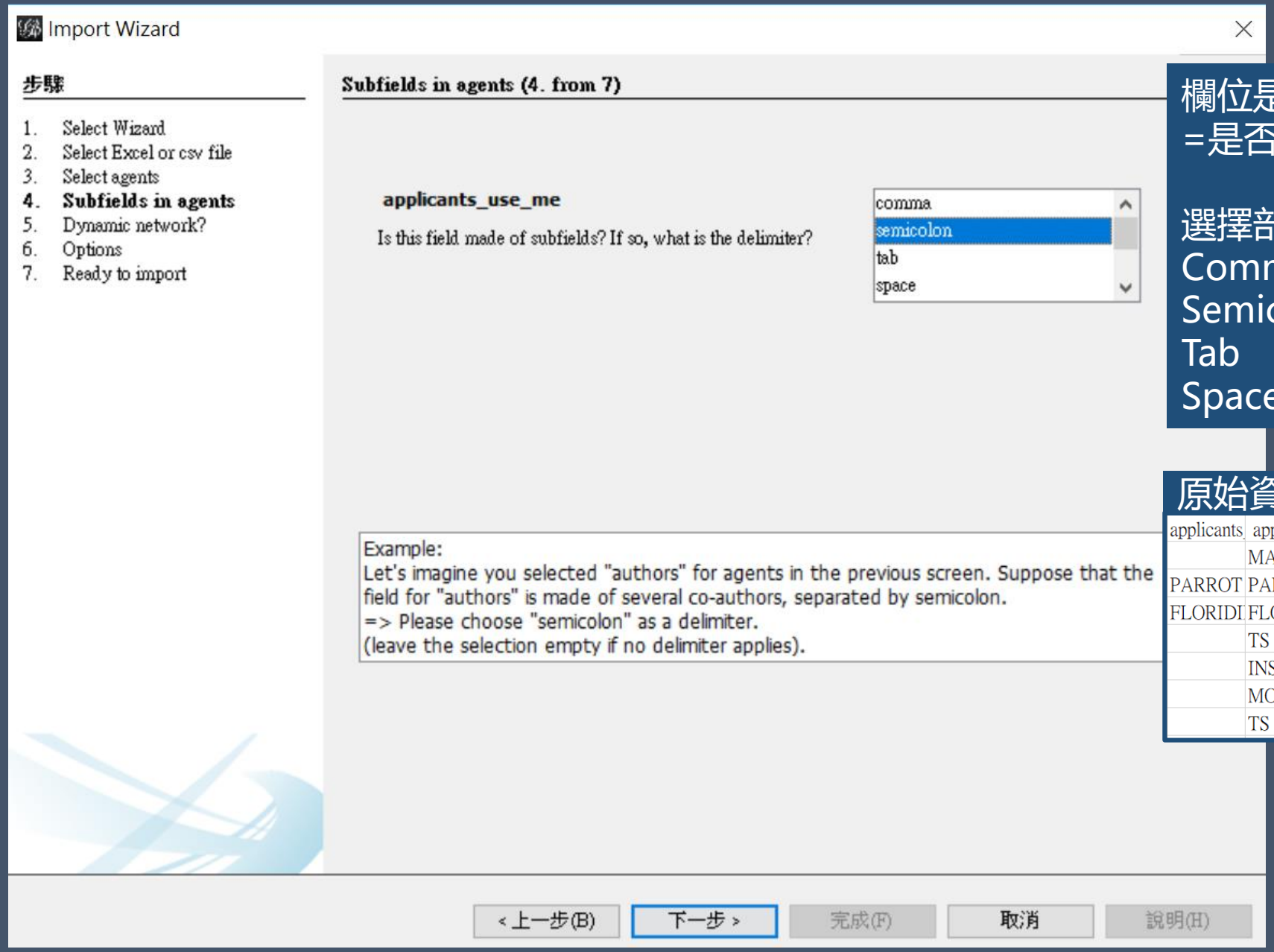


Excel資料記得要有表頭





選擇相連結欄位
EX：共同著作

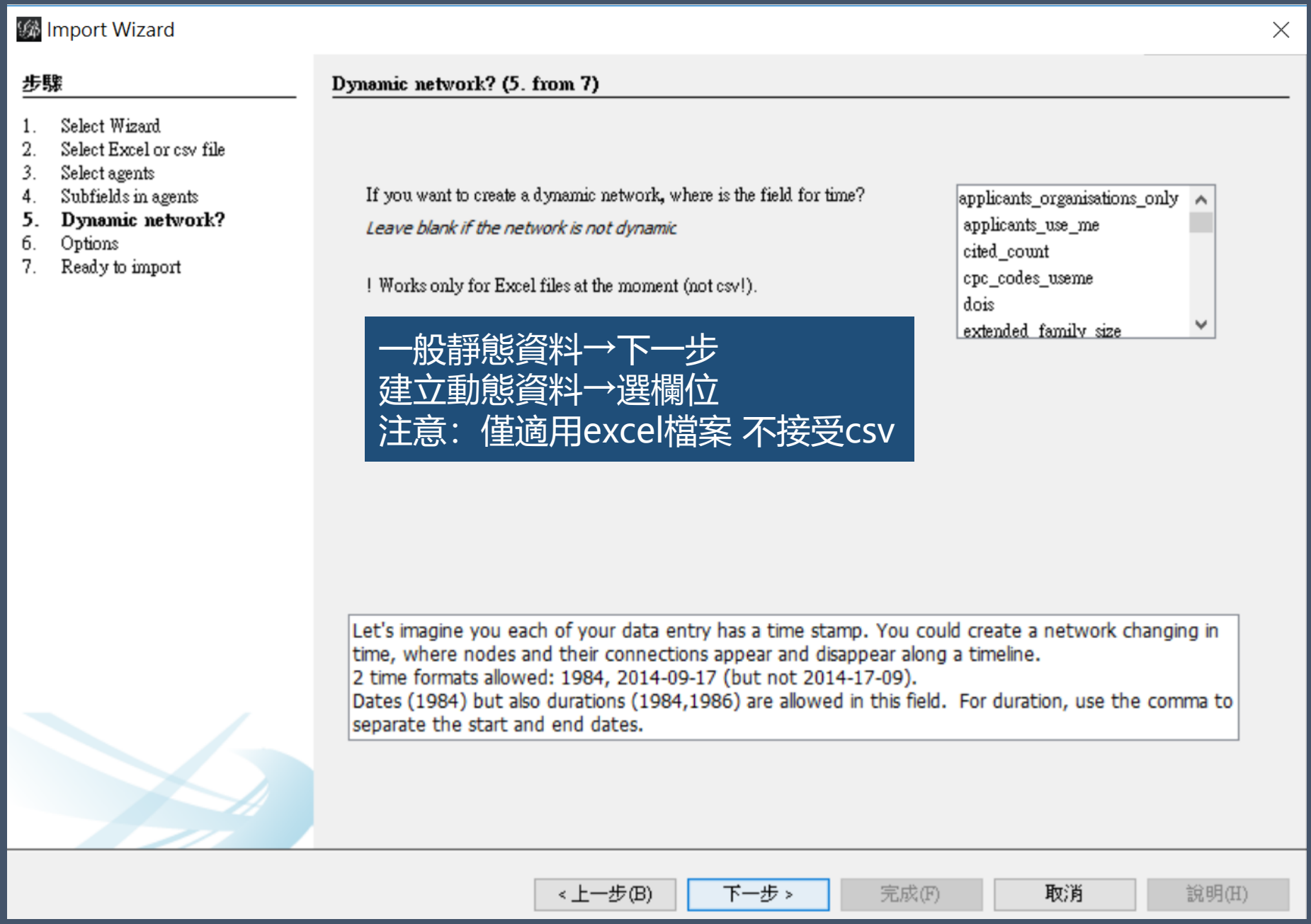


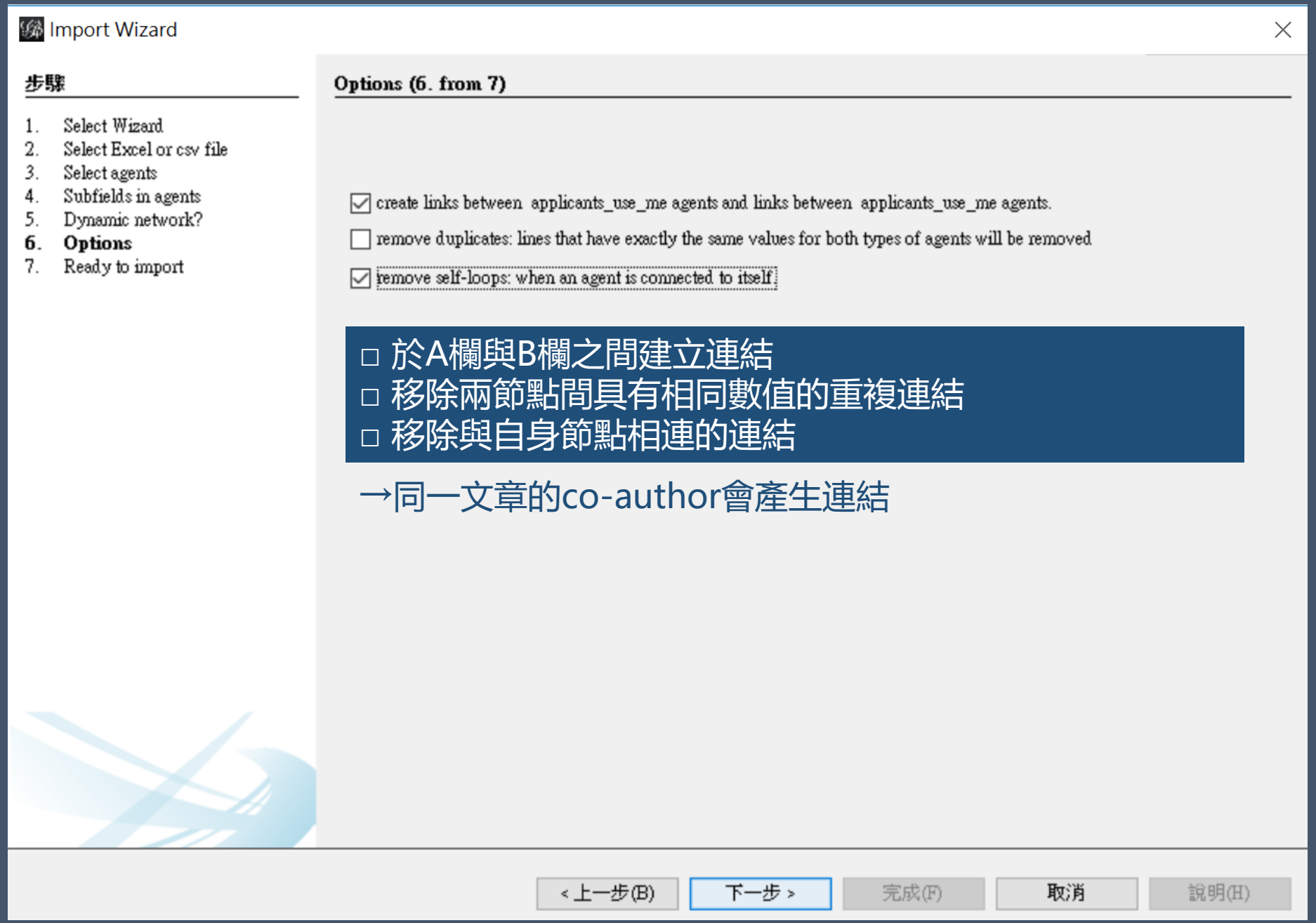
欄位是否還有其他子欄
=是否還需繼續剖析

選擇剖析符號
Comma 逗號
Semicolon 分號
Tab 製表鍵
Space 空白鍵

原始資料，選分號

applicants	applicants_use_me
	MARKOWETZ FRANZ
PARROT	PARROT
FLORIDI	FLORIDIENNE SA; SANDERS PROBEL BIOTECHI
	TS LAB ZA OPTICHEN ZAPIS I OBR
	INST TEKHN KIB I ROBOTIKA PRI
	MONOV
	TS LAB ZA OPTICHEN ZAPIS I OBR

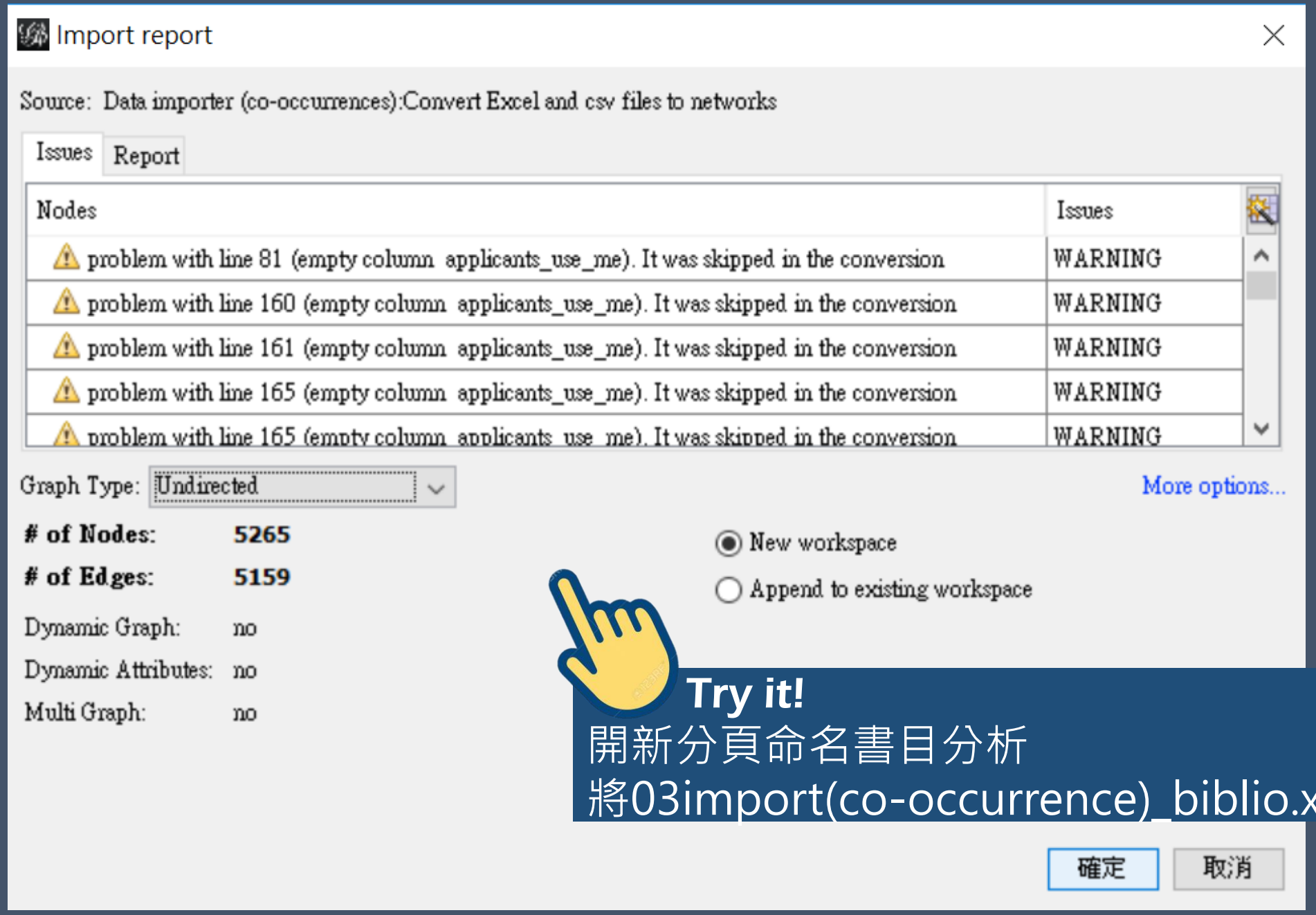

Import Wizard
步驟
1. Select Wizard
2. Select Excel or csv file
3. Select agents
4. Subfields in agents
5. Dynamic network?
6. Options
7. Ready to import
Dynamic network? (5. from 7)
If you want to create a dynamic network, where is the field for time?
Leave blank if the network is not dynamic
! Works only for Excel files at the moment (not csv!).
applicants_organisations_only
applicants_use_me
cited_count
cpc_codes_useme
dois
extended_family_size
一般靜態資料→下一步
建立動態資料→選欄位
注意：僅適用excel檔案 不接受csv
Let's imagine you each of your data entry has a time stamp. You could create a network changing in time, where nodes and their connections appear and disappear along a timeline.
2 time formats allowed: 1984, 2014-09-17 (but not 2014-17-09).
Dates (1984) but also durations (1984,1986) are allowed in this field. For duration, use the comma to separate the start and end dates.
< 上一步(B)
下一步 >
完成(F)
取消
說明(H)



- 於A欄與B欄之間建立連結
- 移除兩節點間具有相同數值的重複連結
- 移除與自身節點相連的連結

→同一文章的co-author會產生連結


Import Wizard
步驟
1. Select Wizard
2. Select Excel or csv file
3. Select agents
4. Subfields in agents
5. Dynamic network?
6. Options
7. Ready to import
Ready to import (7. from 7)
The conversion process will start when you click the "Finish" button.
A network will be created, where:
applicants_use_me are connected to applicants_use_me
Edge weight 連結權重 兩節點之間連結數connection
Frequency 頻率 每一節點出現的次數
- The edge weight will represent the number of connections between any two agents
- Each node (agent) will have an attribute (called "frequency") representing how many times this agent was present in the file.
< 上一步(B)
下一步 >
完成(F)
取消
說明(H)



Try it!

開新分頁命名書目分析

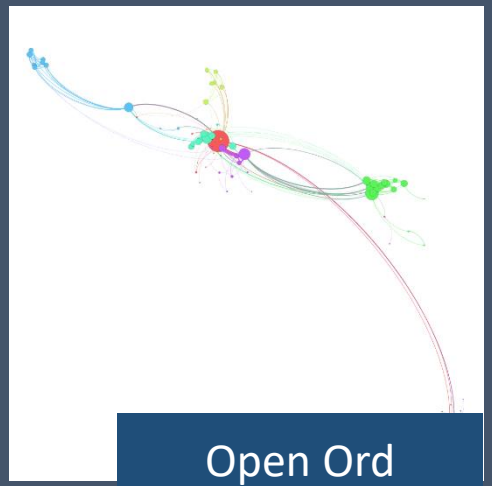
將03import(co-occurrence)_biblio.xlsx練習資料匯入

Gephi實務操作
07 Layout 運算布局

布局的種類

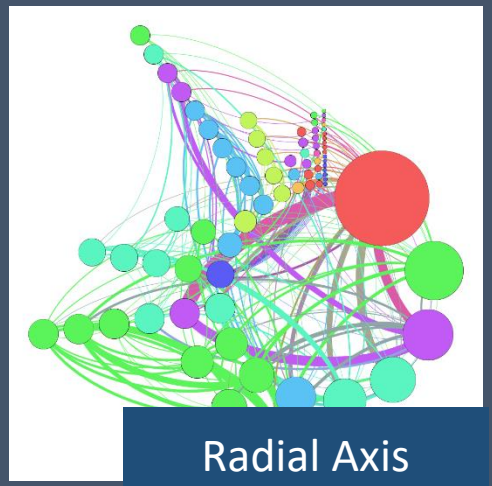
- Force-Direct Layout 力引導布局
- Circle Layout 圓形布局
- Geo Layout 地理布局
- Others

運算的圖像佈局

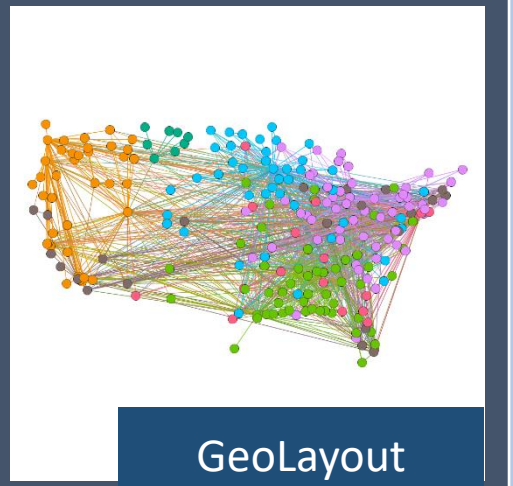


Open Ord

Force Altas



Radial Axis



GeoLayout

分類
Divisions
1) Open Ord

互補
Complementarities
2) ForceAtlas(3)
3) YiFan Hu(2)
4) Fruntchman-Reingold

排序
Ranking
5) Circular
6) Dual Circle
7) Circle Pack
8) Radial Axis

地圖位置
Geographic Repartition
9) GeoLayout

7.1Open Ord

S. Martin, W. M. Brown, R. Klavans, and K. Boyack, "OpenOrd: An Open-Source Toolbox for Large Graph Layout," SPIE Conference on Visualization and Data Analysis (VDA)., 2011

適用圖像：無向+權重

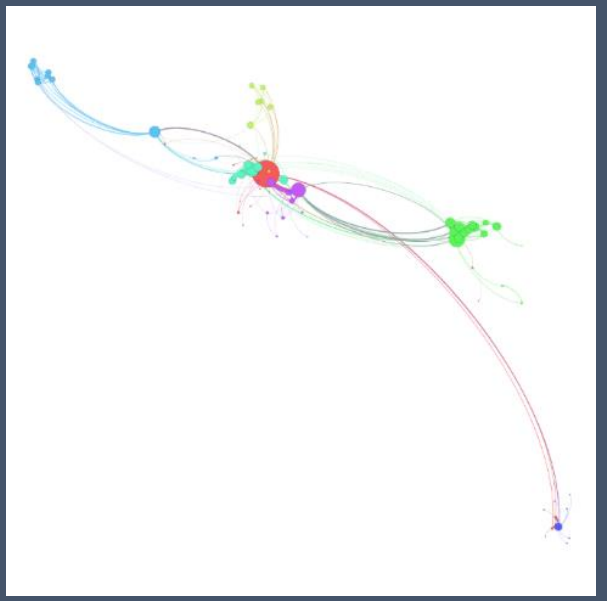
主要用途：用於分群

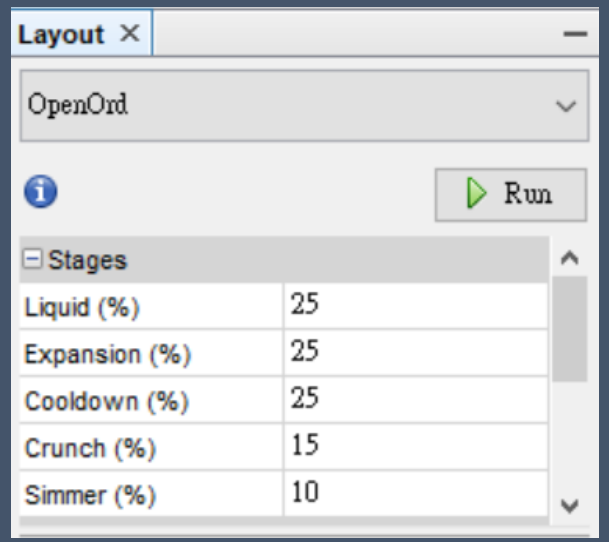
運算方式：力引導+模擬退火法

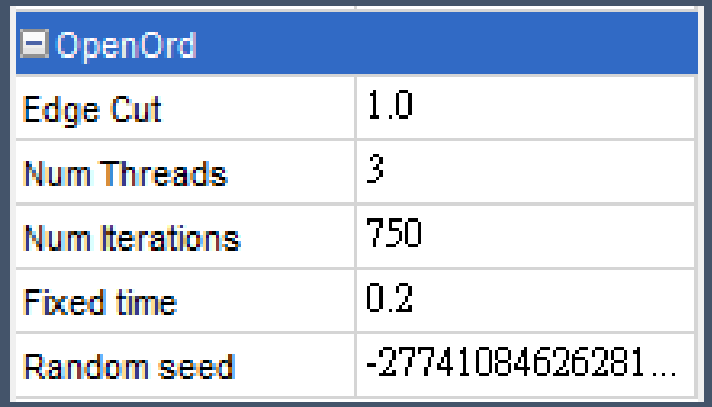
複雜程度：O(N*log(N))

可運算節點範圍：100-1000,000(一百萬)

使用連結權重(edge weight)：是







0(min)-1(max) 連結段開的數量
數值↑分群越多

1(min) core component數量

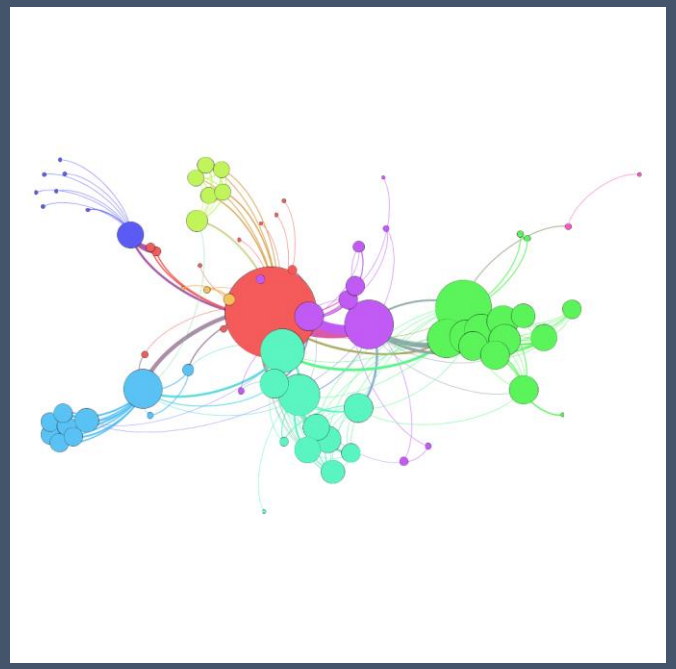
100(min)/750(default)

節點固定的時間 0(不固定)-1(固定)

想像成助鐵匠在選擇敲打的時間
基本都是亂數都可以忽略

7.2.1 Force Atlas

Jacomy, M.Venturini1 T. Heymann S.and Bastian M.(2014)ForceAtlas2, a Continuous Graph Layout Algorithm for Handy Network Visualization Designed for the GephiSoftware.POLS.9(6):p.1-12



適用資料：解讀真實資料，
以最少偏誤地角度明確解讀網絡圖像

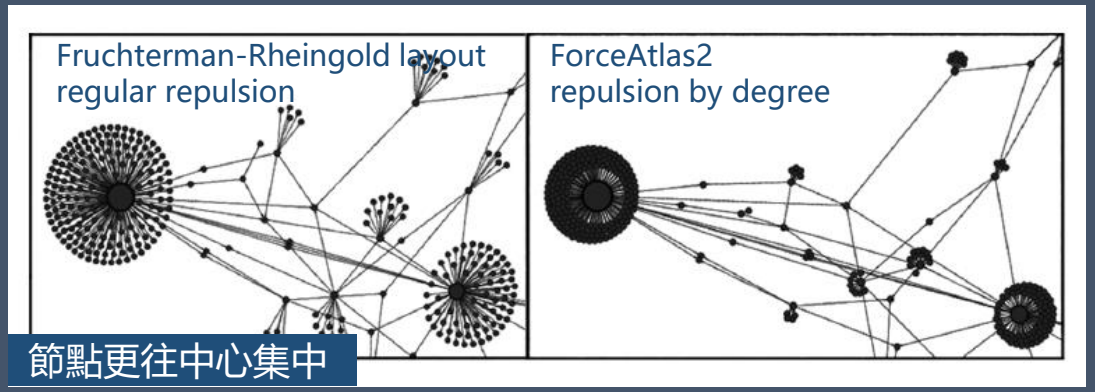
主要用途：小世界網絡圖、無尺度網絡圖
Small-World
Scale-free networks
(小的程度中心性的節點很多，
大的程度中心性的節點很少，冪次法則)

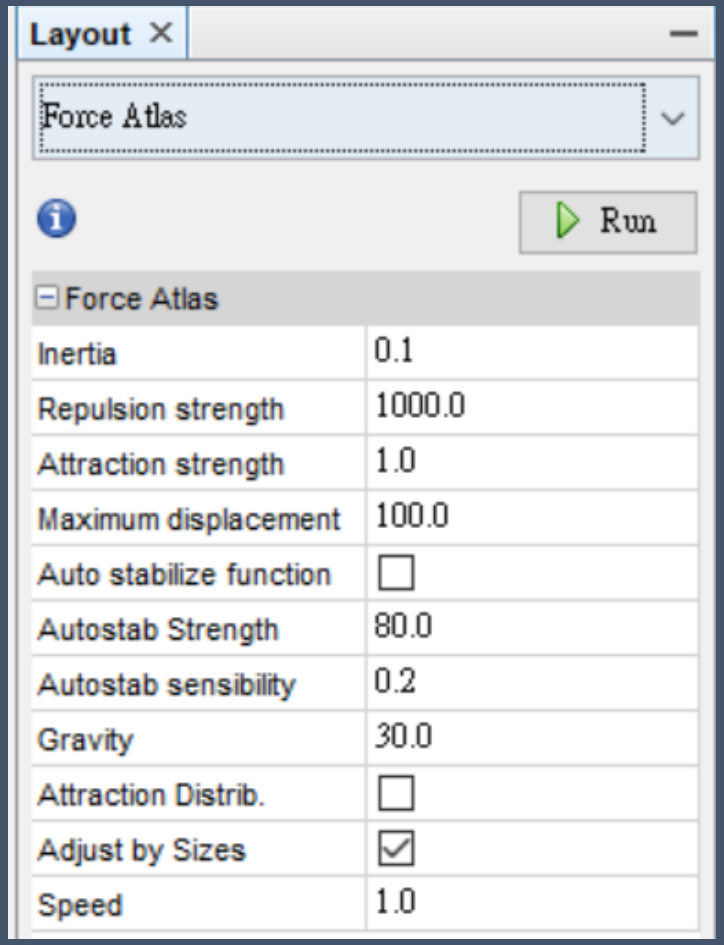
複雜程度：$O(N^2)$

可運算節點範圍：1-10,000(一萬)

使用連結權重(edge weight)：是

7.2.1 Force Atlas





慣性(數值↑ 運算速度↑ 精確率↓)
互斥程度
吸引程度
最大位移量，限制每一節點位移距離，預防部分節點因距離過產生極端互斥
自動穩動功能(活化靜止的不穩定節點，預防節點忽隱忽現，會減少運算效率)
自動穩定功能強度(自動往定功能的強度)
自動穩定功能敏感度
重力(數值↑ 節點越向中心集中)
吸引力分布
依尺寸調整避免節點重疊
速度

7.2.2 Force Altas2 改善第一版演算法 V.1

Layout

Force Atlas

Run

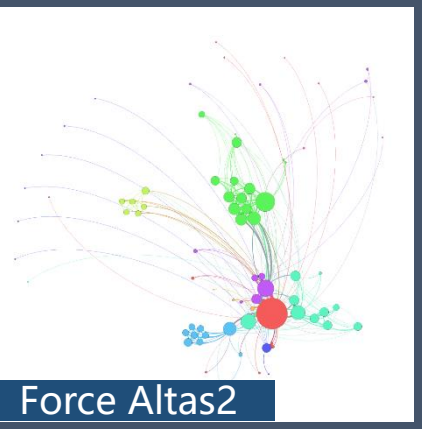
Force Atlas	
Inertia	0.1
Repulsion strength	1000.0
Attraction strength	1.0
Maximum displacement	100.0
Auto stabilize function	☐
Autostab Strength	80.0
Autostab sensibility	0.2
Gravity	30.0
Attraction Distrib.	☐
Adjust by Sizes	☑
Speed	1.0

Layout

ForceAtlas 2

Run

Threads	
Threads number	3
Performance	
Tolerance (speed)	1.0
Approximate Repulsion	☐
Approximation	1.2
Tuning	
Scaling	1000.0
Stronger Gravity	☐
Gravity	100.0
Behavior Alternatives	
Dissuade Hubs	☐
LinLog mode	☐
Prevent Overlap	☑
Edge Weight Influence	1.0



可容許的擺盪程度，數值↓精確率↑，不超過1為佳

節點距離拉大，可形成較大圖像

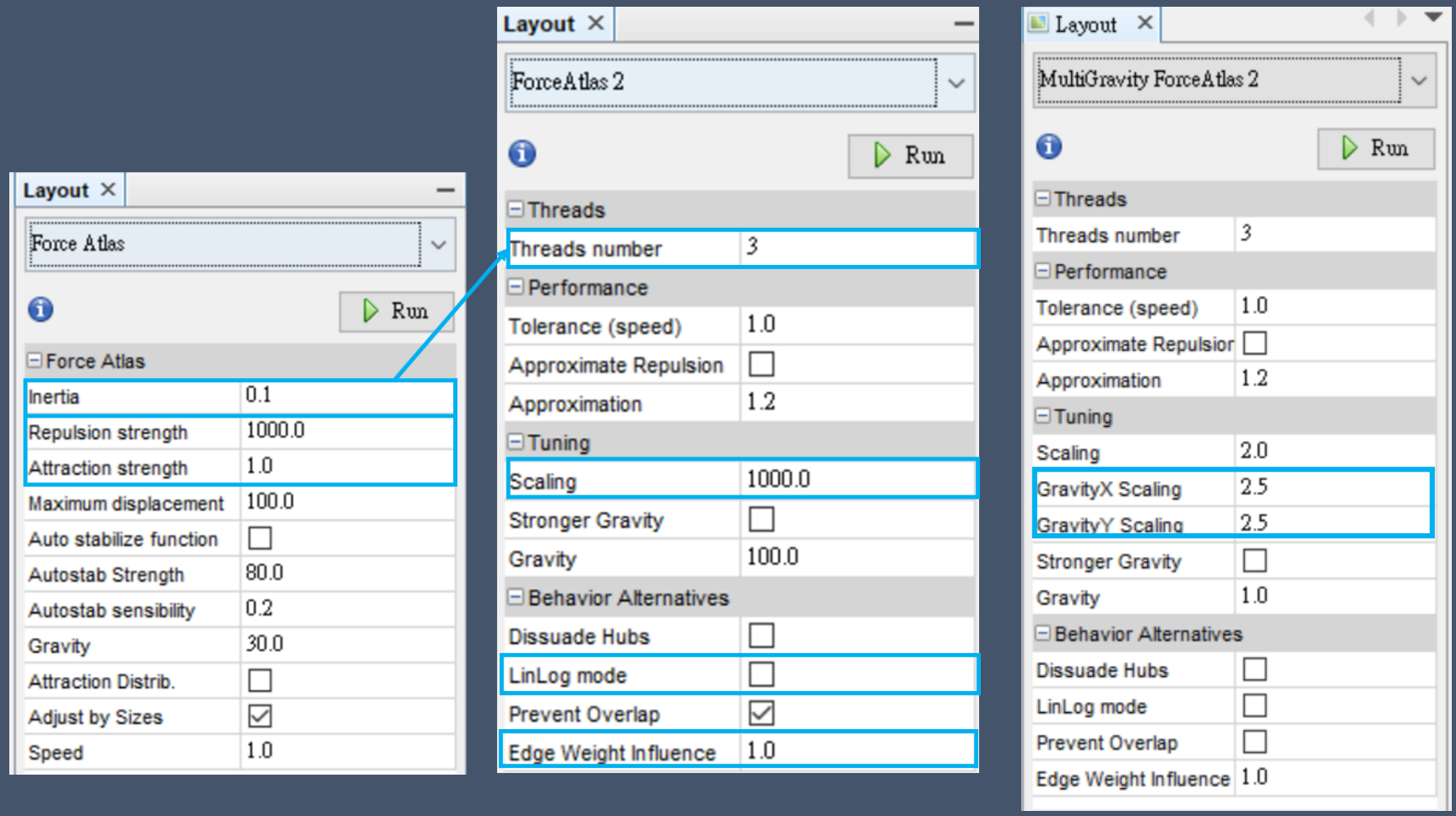
Barnes-Hut演算法最佳化的θ

數值↑ 互斥程度↑ 清晰度↓

模式	吸引力	互斥力	作用
預設	線性	線性	
Linlog	線性	對數	使分群更緊密的結合

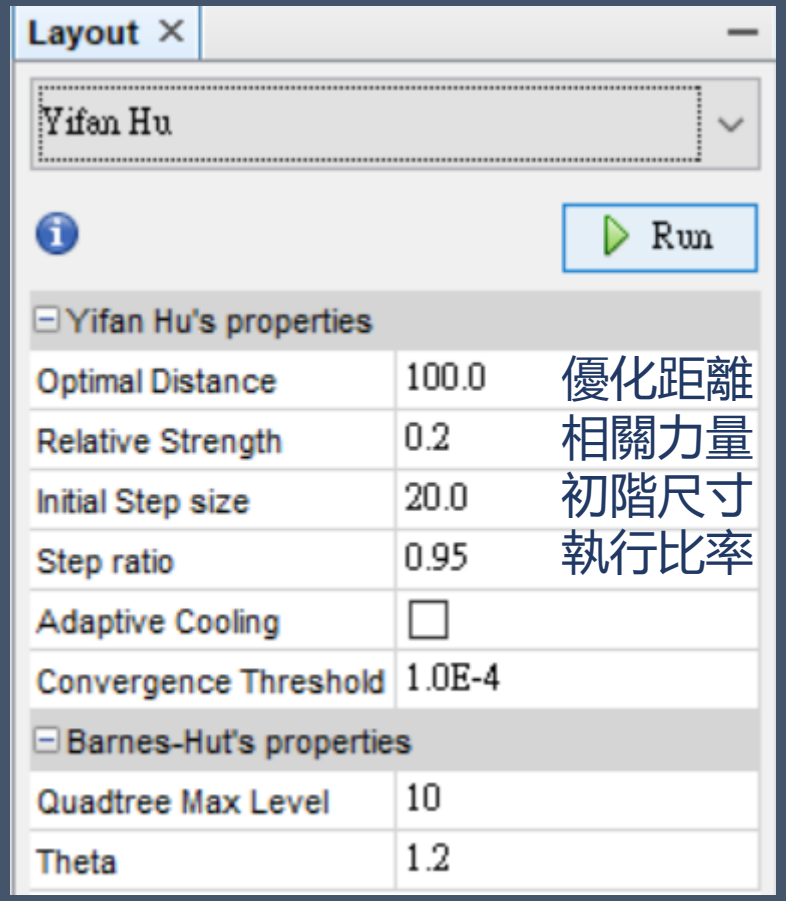
給予每一個連結多少權重，0=沒有

7.2.3 Multigravity ForceAtlas 2



水平集中
垂直集中

7.3.1 YiFan Hu



優化距離：數值↑，圖像節點分散、數值↓，圖像節點越密
圖像互斥(repulsion)與吸引(attraction)兩力之比例
通常預設是10
在執行演算法時運算與更新時，stepsize的比率
用運adoptive cooling
訂演算法停止的能量聚焦程度，門檻值↓精準率↑

四叉樹最大程度：數值↑、正確率↑
Theta θ：數值↓、正確率↑

7.3.2 YiFan Hu Proportional

運算的方式和功能與 YiFan Hu 相似

差異在於

圖形節點的配置以比例置換(proportional displacement)的方式進行

還有，可以處理較大的網絡圖

7.4 Fruchterman Reingold

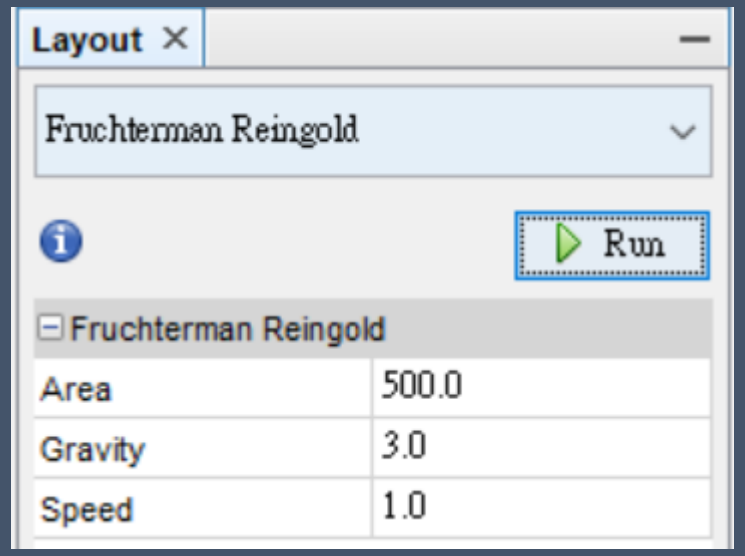
Fruchterman, T. M. J., & Reingold, E. M. (1991). Graph Drawing by Force-Directed Placement. Software: Practice and Experience, 21(11).

運算方式：彈簧模型+靜電力

複雜程度：$O(N^2)$

可運算節點範圍：1-1000

使用連結權重(edge weight)：否!



定義視覺化圖像區域大小
節點與中心距離的程度
演算法聚焦的速度
數值↑、精準度↓

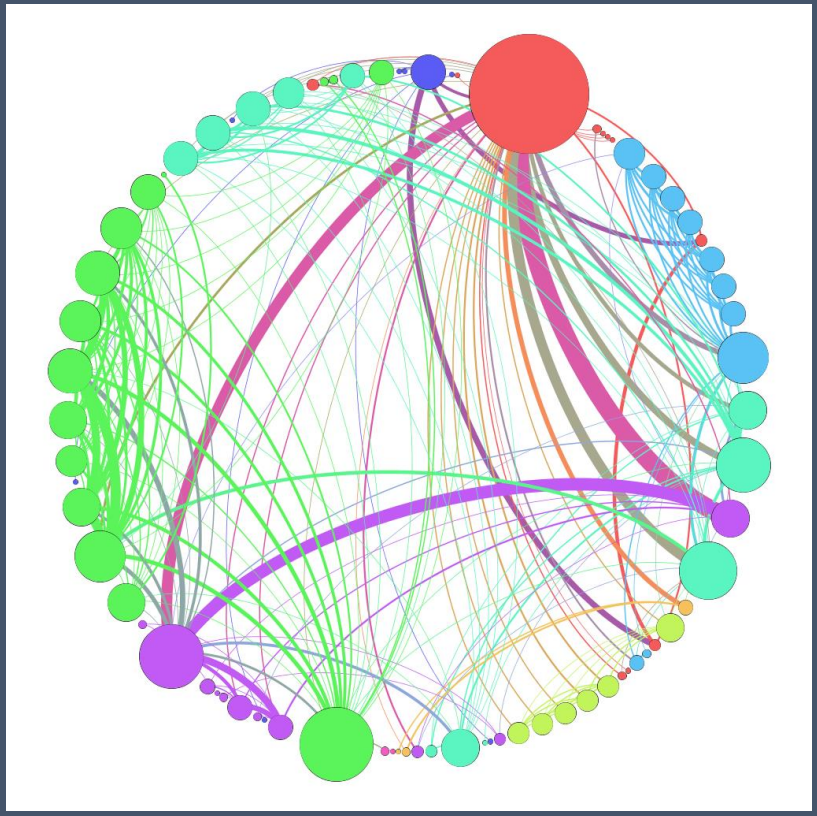
7.5.1Circular Layout

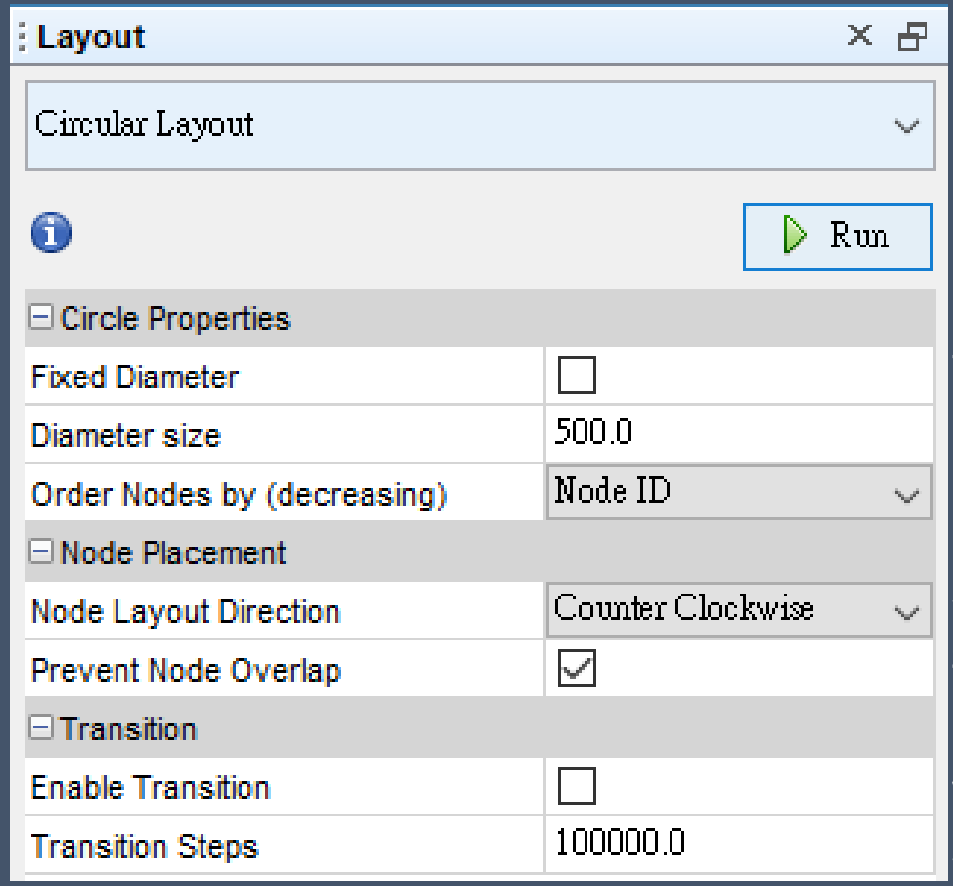
Matt Groeninger. “Circular layout”, Gephi Marketplace, March 2, 2013. https://marketplace.gephi.org/plugin/circular-layout/.

將資料根據設定的屬性(ID、Degree...)排列成圓形，觀察節點與連結之間的分布

複雜程度：$O(N^2)$

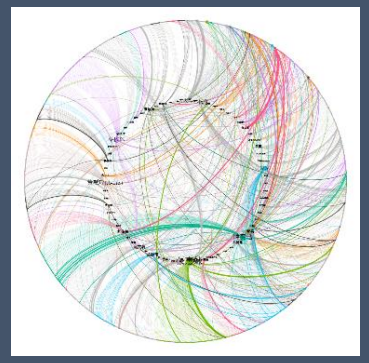
可運算節點範圍：1-1,000,000(一百萬)



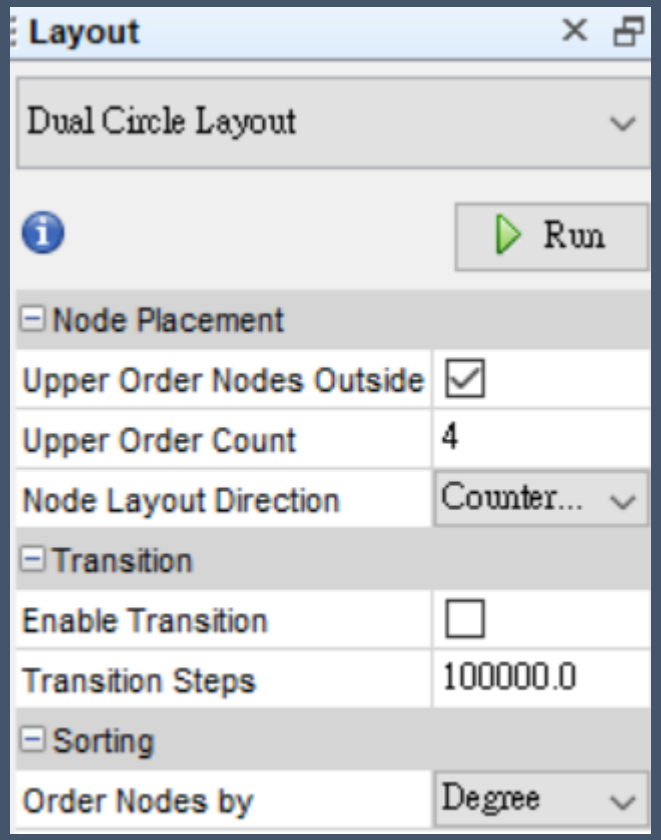


調整圖像直徑
圖像直徑排序的屬性(降冪)

節點排序的方向：順時鐘、逆時鐘
預防節點重疊(如有勾選無法Fixed Diameter同時作用)
做一個平滑的過渡到布局
過度需要的步驟數量

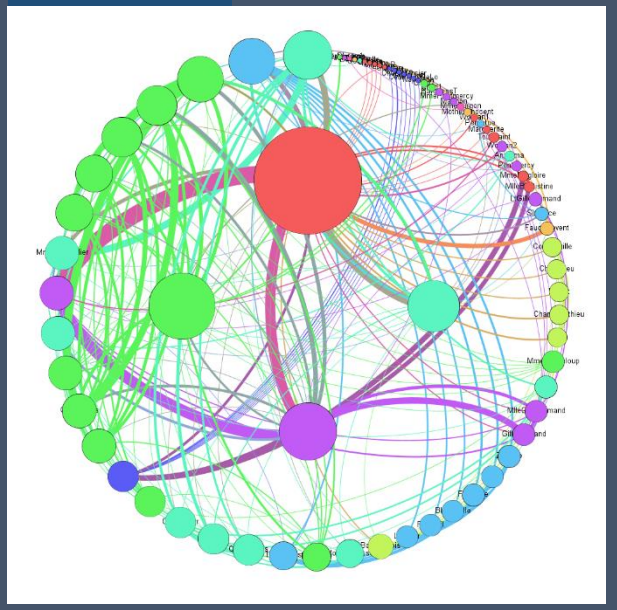


7.5.2 Dual Circle Layout

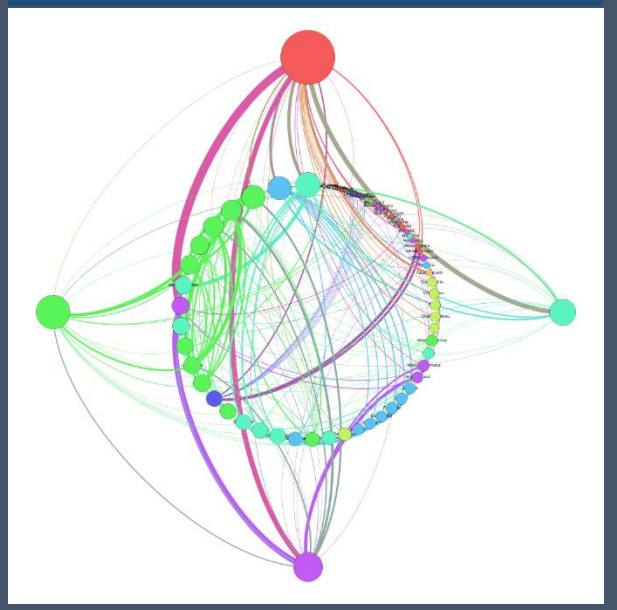


指定節點在外圍
設定外圍節點數
順/逆時鐘

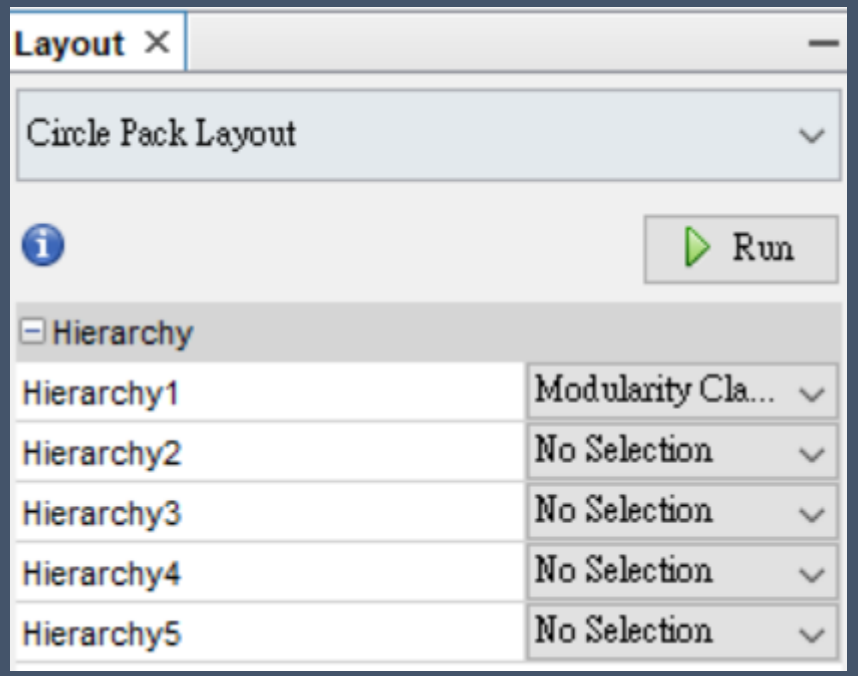
Normal

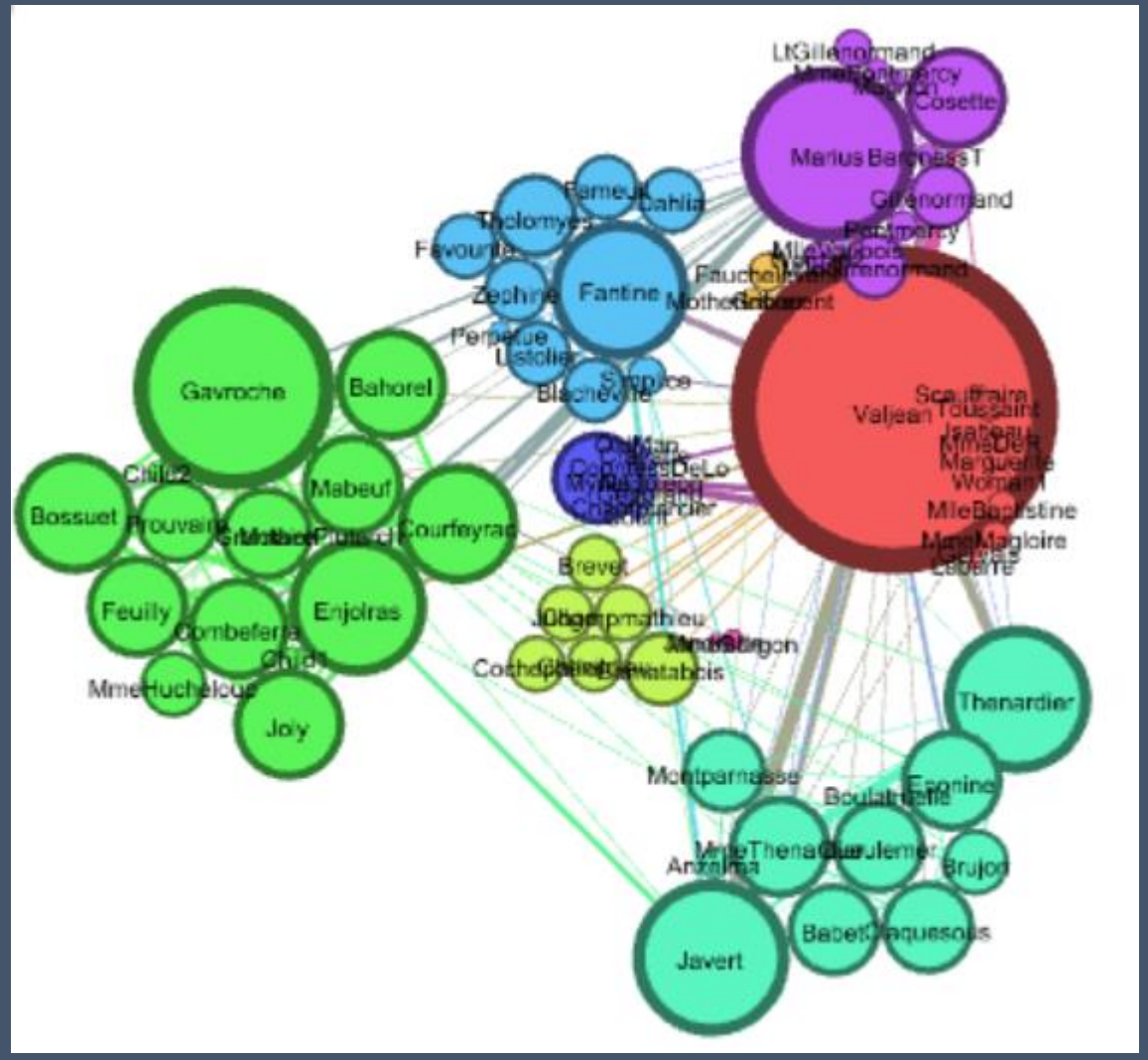


Upper Order Nodes Outside



7.5.3 Circle Pack





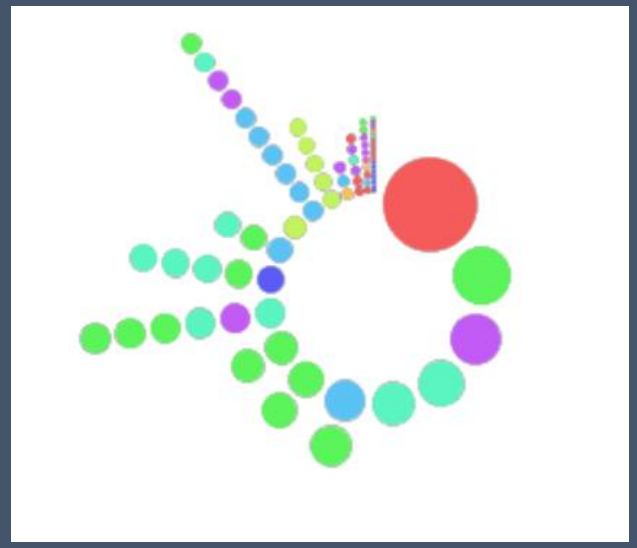
7.5.4 Radial Axis Layout

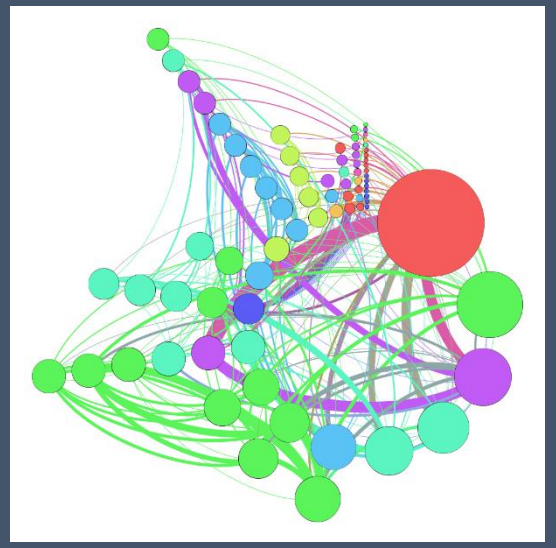
計算方式：將節點依屬性分組，相似節點分為一組並將其繪製成向外放射的軸線，以此類推

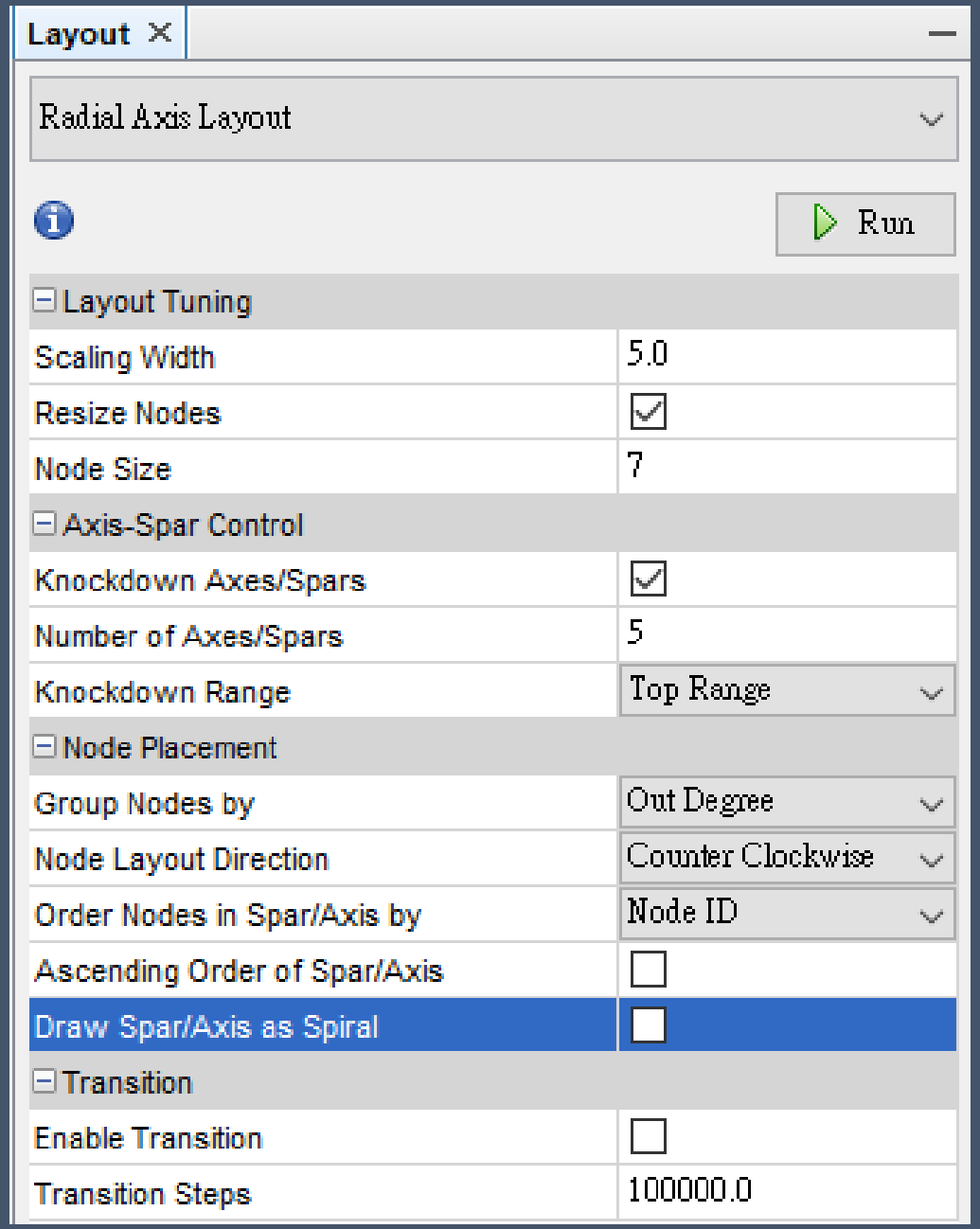
適用目的：用於觀察各組之間節點與連結的分布，進而比較各組之間的同質性(homophily)

複雜程度：$O(N^2)$

可運算節點範圍：1-1,000,000(一百萬)







控制節點之間的空間

重新設定節點的大小

減少軸的數量(=減少分組數)
減成幾之軸(=撿到剩幾組)
減少範圍：top、middle、buttom

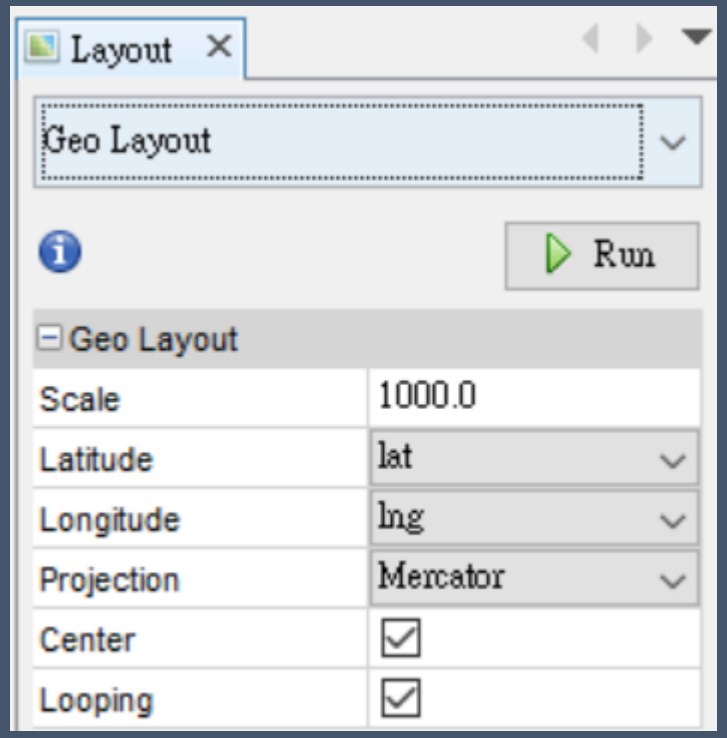
節點分組的屬性
節點佈局的方向
軸內的節點排列順序
軸的升冪排列順序
有勾選→顯現社群內的連結
未勾選→顯現社群之間的連結
做一個平滑的過渡到布局
過度需要的步驟數量

7.6 Geo Layout

使用經緯度來定位節點
複雜程度：O(N²)
可運算節點範圍：1-1,000,000(一百萬)

地理資料匯入格式

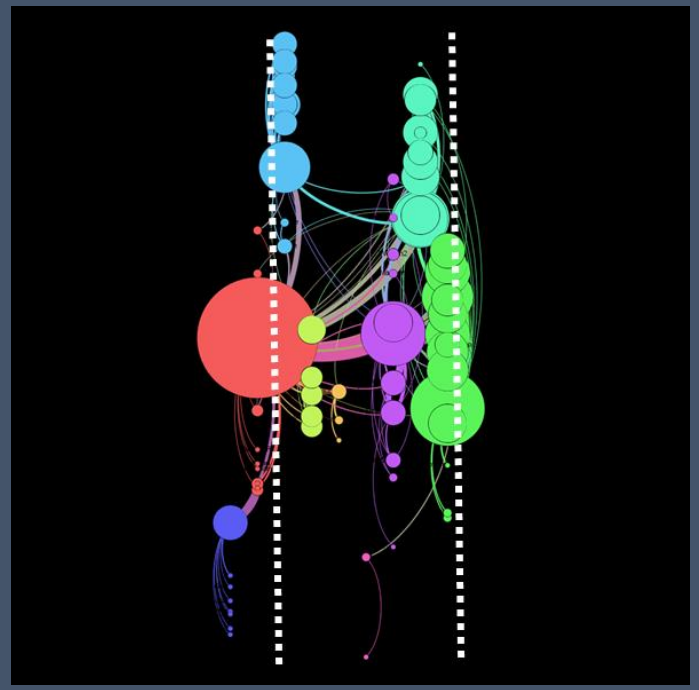
Id	Label	lng	lngx	lat	laty
0	0	-92.2244	可空白	34.72944	可空白
1	1	-81.4422		40.91611	
2	2	-73.8		42.73333	

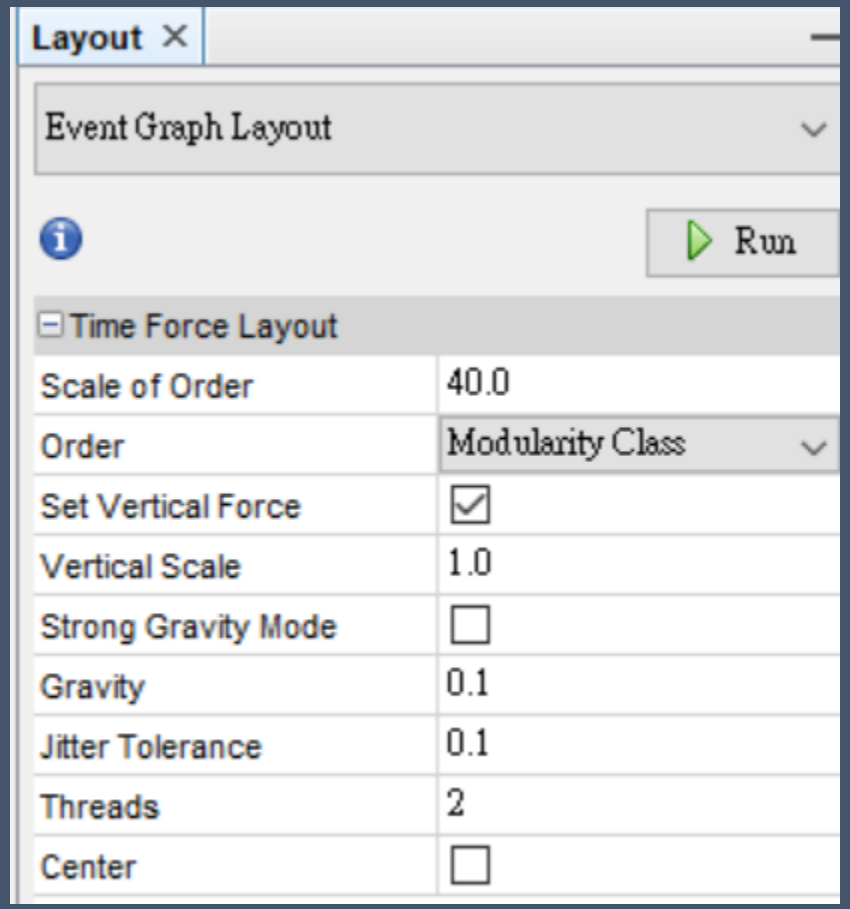


整體圖像的大小
緯度
經度
投射投影法(Mercator麥卡托)
放置於圖像中心
持續進行演算

7.7 Event Graph Layout

剛採用此布局時，節點不會整齊的在同一直線上
→調整Vertical Scale+ Jitter Tolerance，才會慢慢集中





軸線之間的距離
排列的屬性
影響垂直軸上的節點位置(越緊的節點會靠得更近，越鬆散的越遠)
將未連結的節點彼此推開的程度
強重力模式
重力
抖動容差
軸數量
向中心聚集

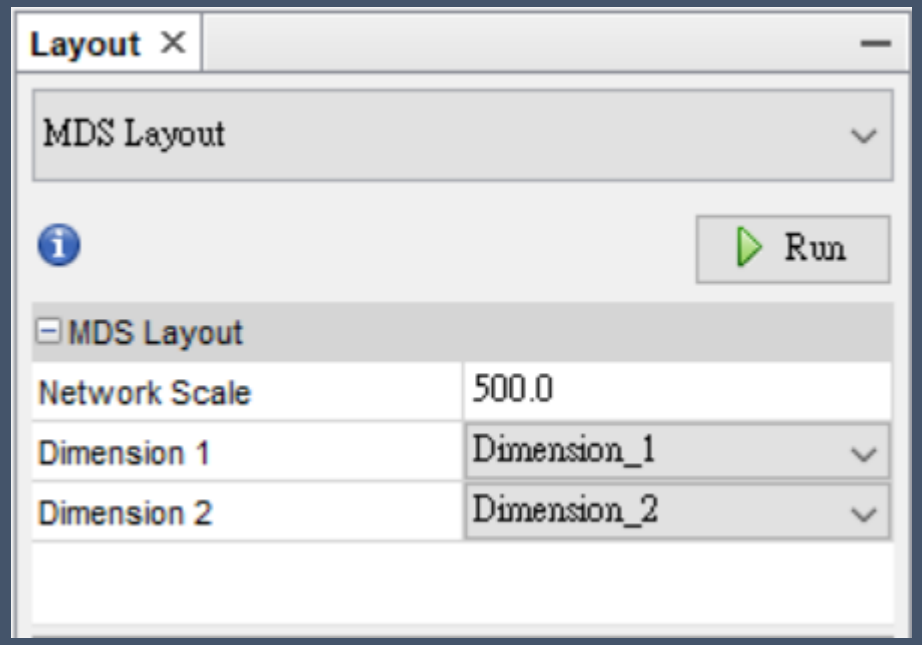
抖動容差:
決定當節點撞到其他節點時允許跳轉的節點數量
可以使水平軸上相同位置的節點更容易相互通過

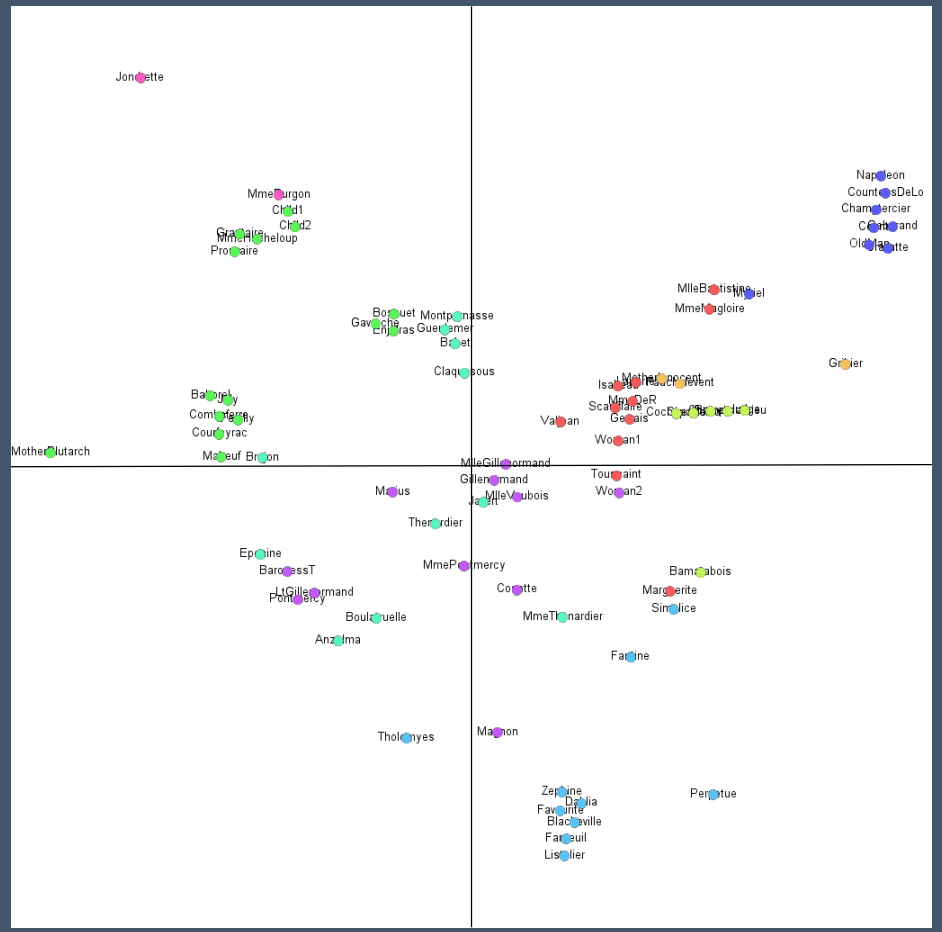
7.8 MDS Layout Multi-Dimensional Scaling 多維尺度

計算圖像節點的路徑距離來表示相似度結構分析

二維向量重疊程度來代表相似程度

Statistic>Multidimensional scaling
Layout> MDS Layout



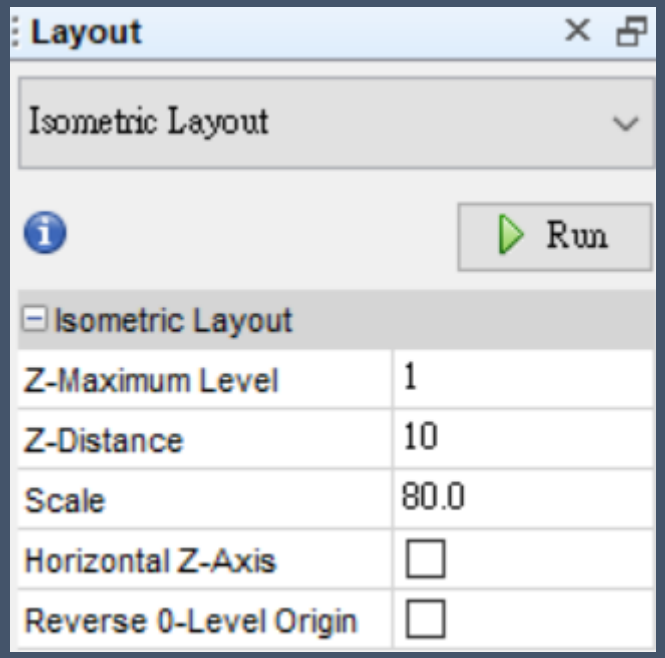


7.9 Isometric Layout

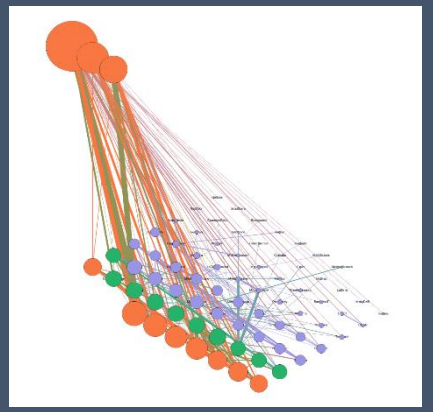
Isometric 等軸測投影

在二維平面呈現三維物體的方法，
三條坐標軸的投影縮放比例相同，
並且任意兩條坐標軸投影之間的角度都是120度。

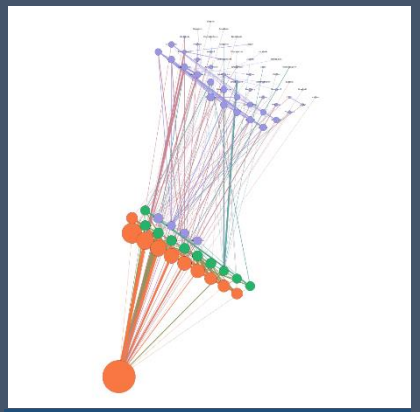
注意 Data laboratory的資料欄位中包含[z] EX：Degree [z]
→執行布局
→Computed Z-Level



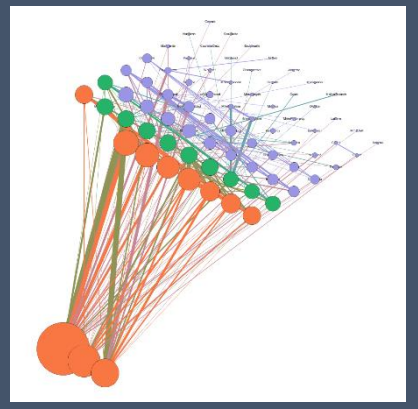
投射的階層數
階層之間的距離
整體網絡大小
水平投射
從Z level=0開始投射



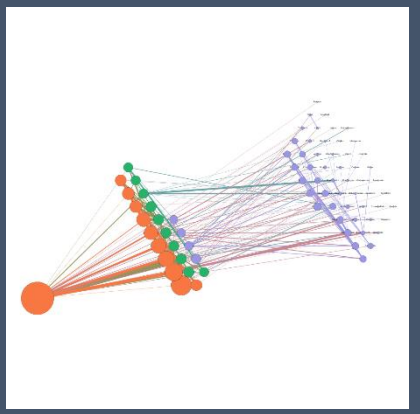
預設 z max level=1



預設 z max level=2



預設 z max level=1
Reverse 0-Level Origin

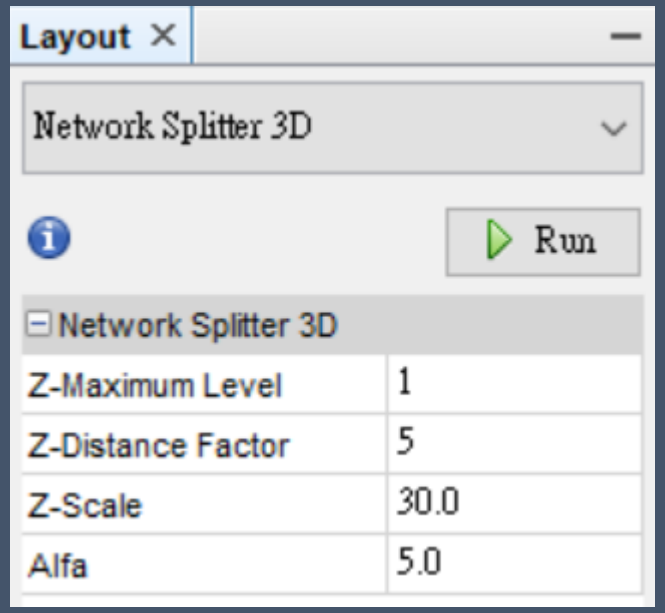


水平投射

Try it!

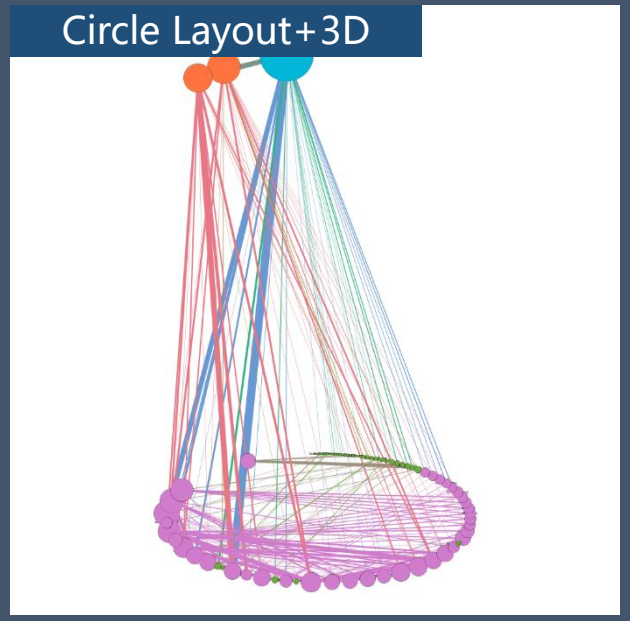
1. 在Data Laboratory練習輸入Z值
2. 應用 Isometric Layout

7.10 Network Splitter 3D

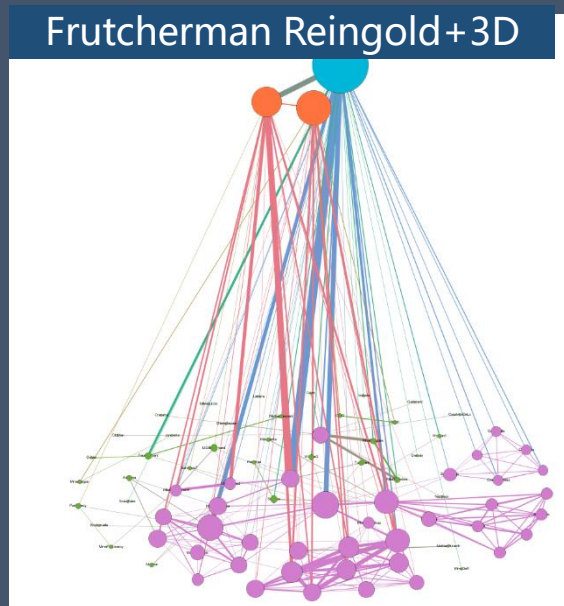


投射的階層數
階層之間的距離
整體網絡大小
旋轉角度0-360度

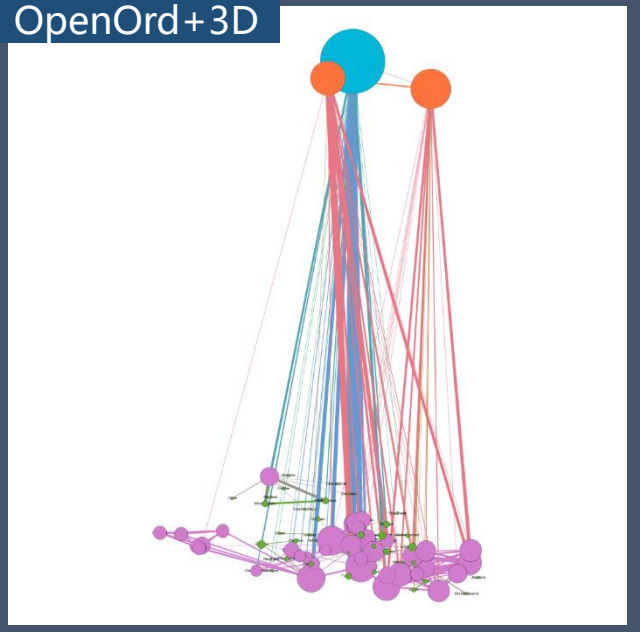
Circle Layout+3D



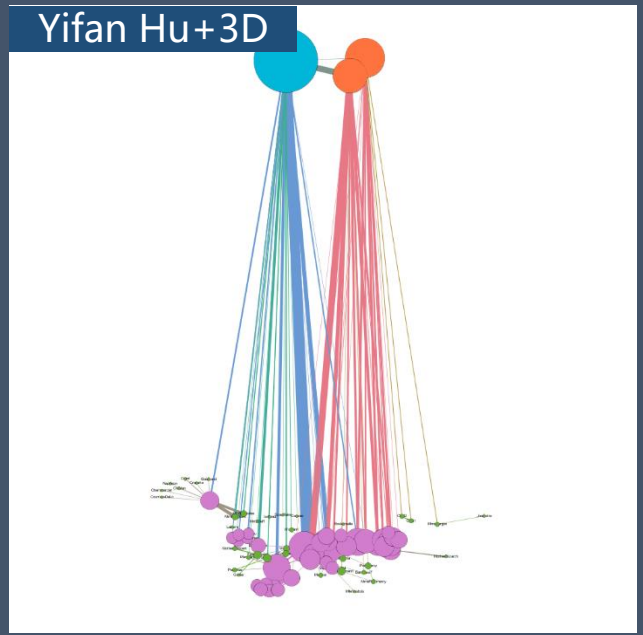
Frutcherman Reingold+3D



OpenOrd+3D

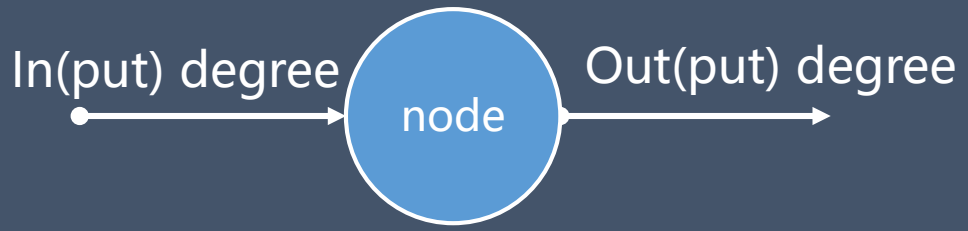


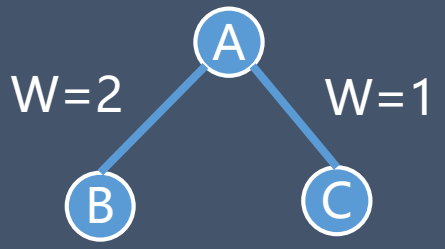
Yifan Hu+3D

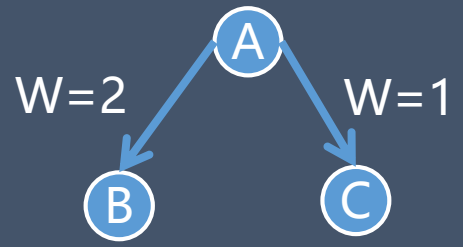


Gephi實務操作
08 Statistic 統計
8.1網絡概觀

Average degree=edges/nodes







Degree 程度=indegree+outdegree

單一節點的連結數

Node A degree =2

Ex：方大同-林俊傑(1個歌手)

Weighted Degree 加權程度

單一節點【其連結數*各自權重】加總

Node A weighted degree =1*2+1*1=3

Ex：方大同-林俊傑(合作5次)

Directed Graph 有向圖

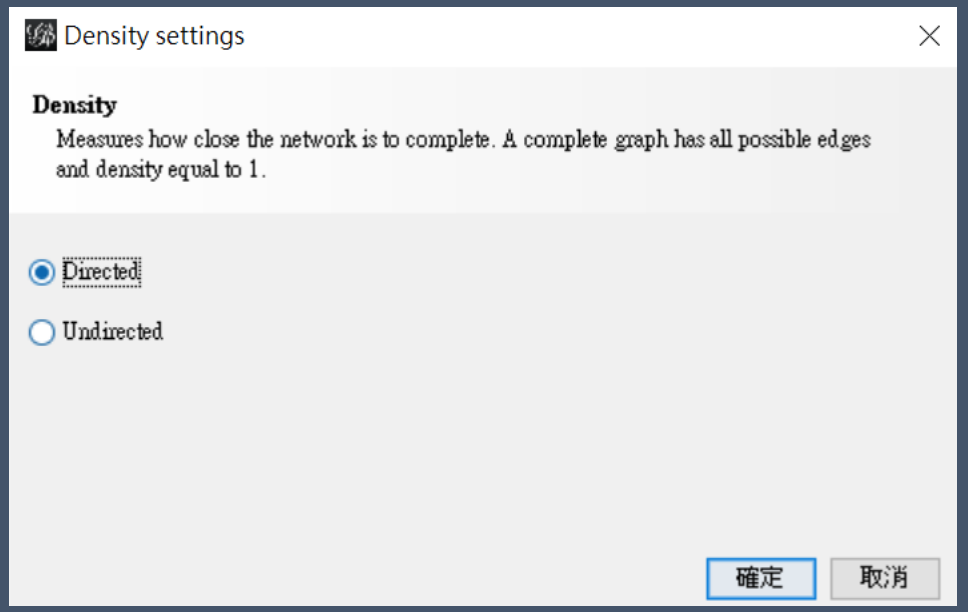
weighted InDegree　Node B =1*2=2

weighted OutDegree Node A =1*2+1*1=3

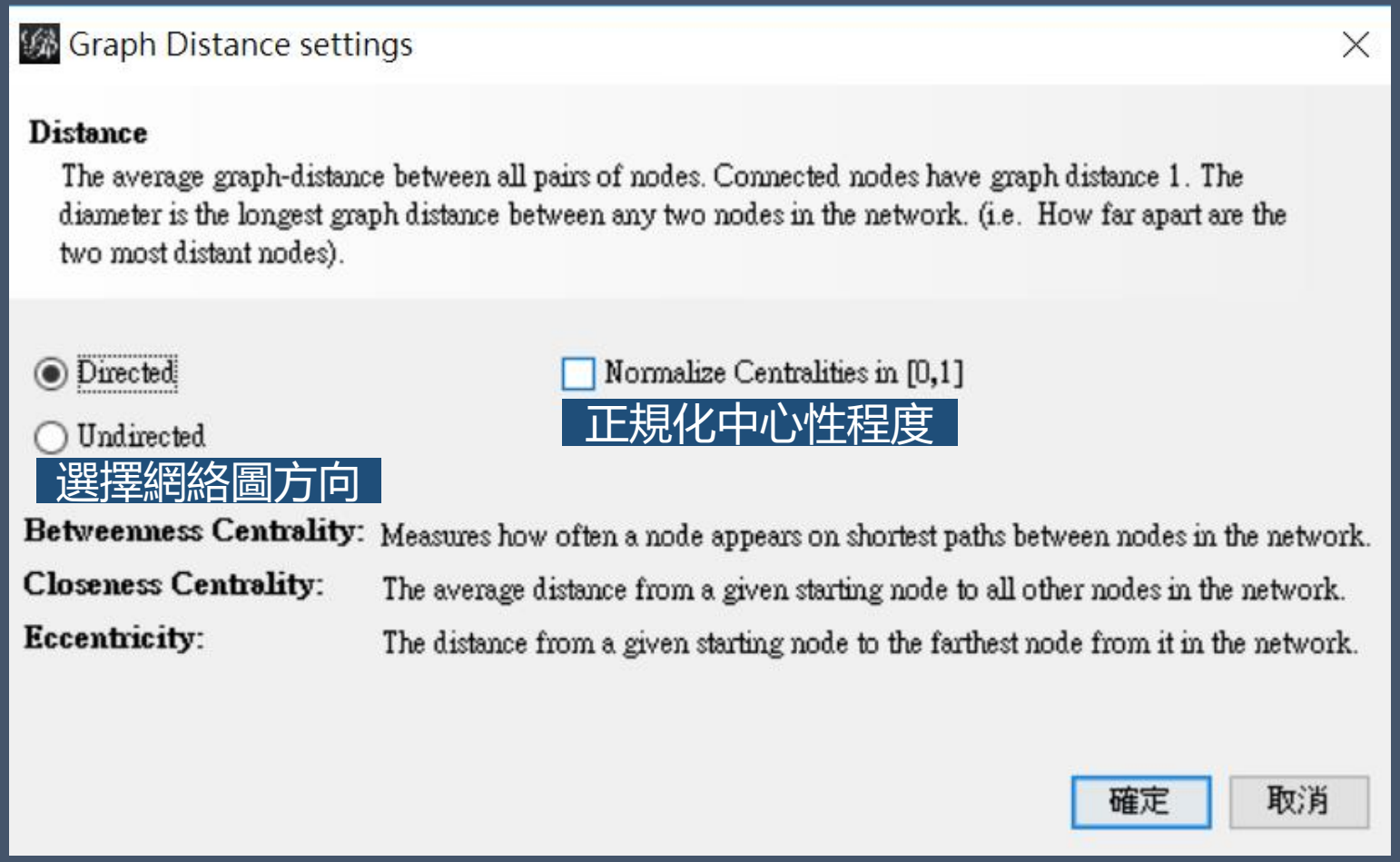
Graph Density

$$\text{Graph Density} = \frac{\text{the total number of edges present in the graph}}{\text{the total number of edges possible in the graph}}$$

- Directed possible edges= n(n-1)
- Undirected possible edges = [n(n-1)]/2

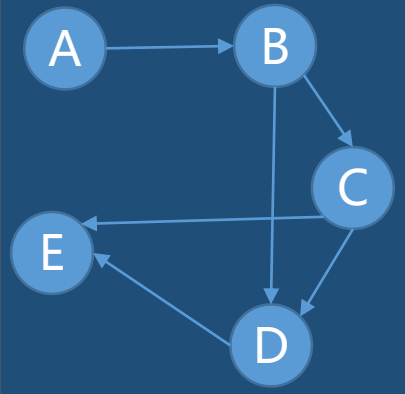


Network Diameter網絡直徑

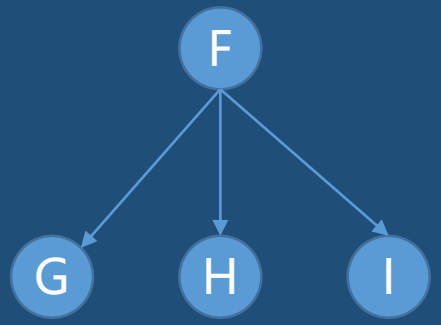


Network Diameter網絡直徑

Distance	the path of pair of vertexes
Eccentricity	the shortest path of a single vertex to the farthest vertex
Radius 半徑	minimum of eccentricity eccentricity = radius → central vertex
Diameter 直徑	maximum of eccentricity eccentricity = diameter → periphery vertex *weighted graph diameter the hops of a vertex to the farthest vertex *If the network is disconnected → the diameter is infinite 無限大



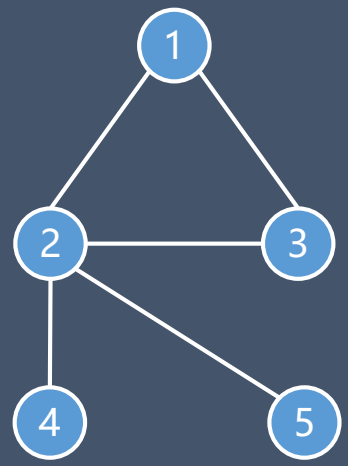
Eccentricity(A)=3
E(B)=2
E(C)=2
E(D)=2
E(E)=3
Diameter=3
Radius=2



Eccentricity(F)=1
E(G)=2
E(H)=2
E(I)=2
Diameter=2
Radius=1

Average path length

the average number of steps along the shortest paths for all possible pairs of nodes.
網絡內所有可能連結最短路徑的平均長度(連結最短路徑長度總和/連結數)
It is a measure of the efficiency of information or mass transport on a network



$$possible\ edges = \frac{5(5-1)}{2} = 10$$

from	to	path	length
1	2	1,2	1
1	3	1,3	1
1	4	1,2,4	2
1	5	1,2,5	2
2	3	2,3	1
2	4	2,4	1
2	5	2,5	1
3	4	3,2,4	2
3	5	3,2,5	2
4	5	4,2,5	2
Sum=15 Average=1.5			

Gephi實務操作
Statistic 統計
8.2網絡與節點中心性_中介中心性

Network Diameter > Betweenness Centrality

中介中心性(間接中心性)

定義：在節點對之間的最短途徑通過單一節點的頻率
→多數節點須通過該節點以觸及其他節點

$$C_B(i) = \sum_{j<k} g_{jk}(i)/g_{jk}$$

g*jk*(i)=the number of geodesics that actor *i* is on
j k 兩節點之間通過 i 的最短捷徑次數

g*jk* = the number of geodesics connecting *jk*
連結 j k 兩節點的最短捷徑個數

Gephi實務操作
Statistic 統計
8.3網絡與節點中心性_接近中心性

Network Diameter > Closeness Centrality

Closeness Centrality接近中心性

(directed graph)

$$C(x) = -\frac{\mathrm{N}}{\sum_y d\,(x,y)}$$ 節點距離加總→取倒數

定義：該節點到所有節點之最短距離總和之倒數

意涵：在最短時間內，某樣東西流經整個網絡

數值越大代表與其他多數節點的距離越短

某一資訊可以透過該接近中心性高的節點迅速傳達給網絡內的多數的其他節點

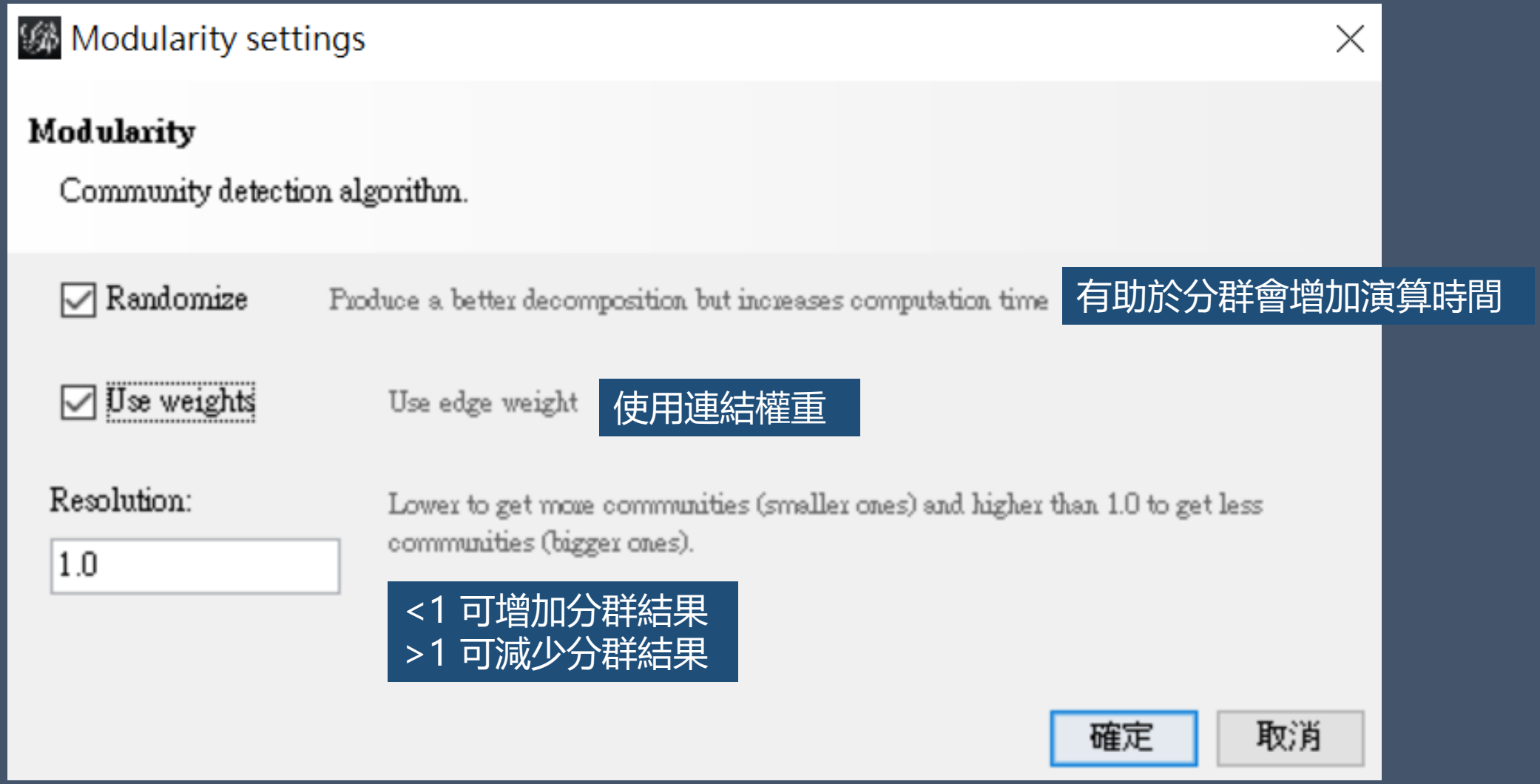
舉例：

(1)人際關係，接近中心性越高，個性越合群

(2)病毒散播，接近中心性越高，越容易感染

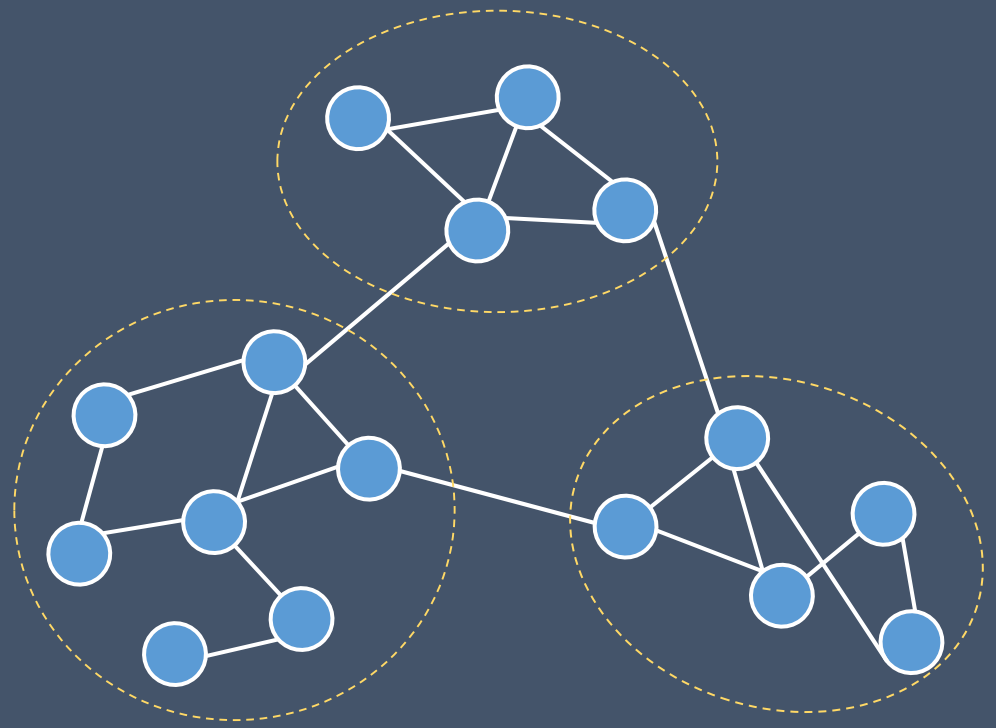
Gephi實務操作
Statistic 統計
8.4 網絡與分群_Modularity

Modularity


Modularity settings
Modularity
Community detection algorithm.
Randomize
Produce a better decomposition but increases computation time
有助於分群會增加演算時間
Use weights
Use edge weight
使用連結權重
Resolution:
1.0
Lower to get more communities (smaller ones) and higher than 1.0 to get less communities (bigger ones).
<1 可增加分群結果
>1 可減少分群結果
確定
取消

Gephi實務操作
Statistic 統計
8.5 網絡與分群Girvan-Newman Clustering

Girvan-Newman Clustering

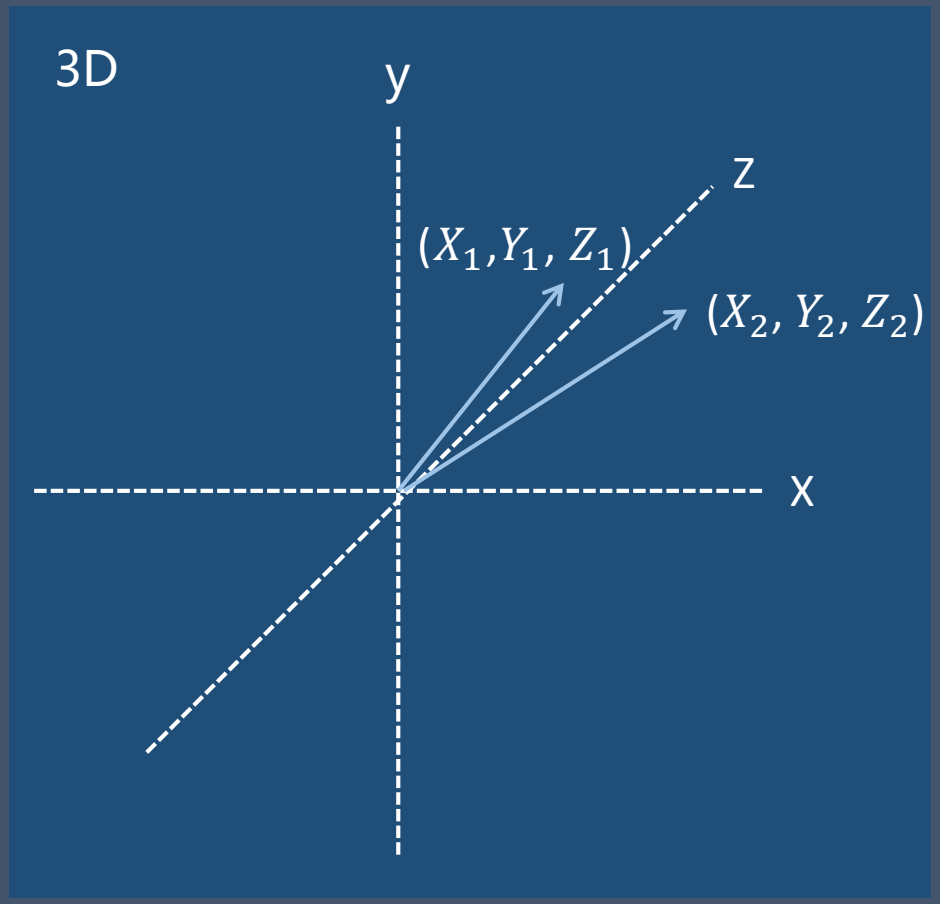


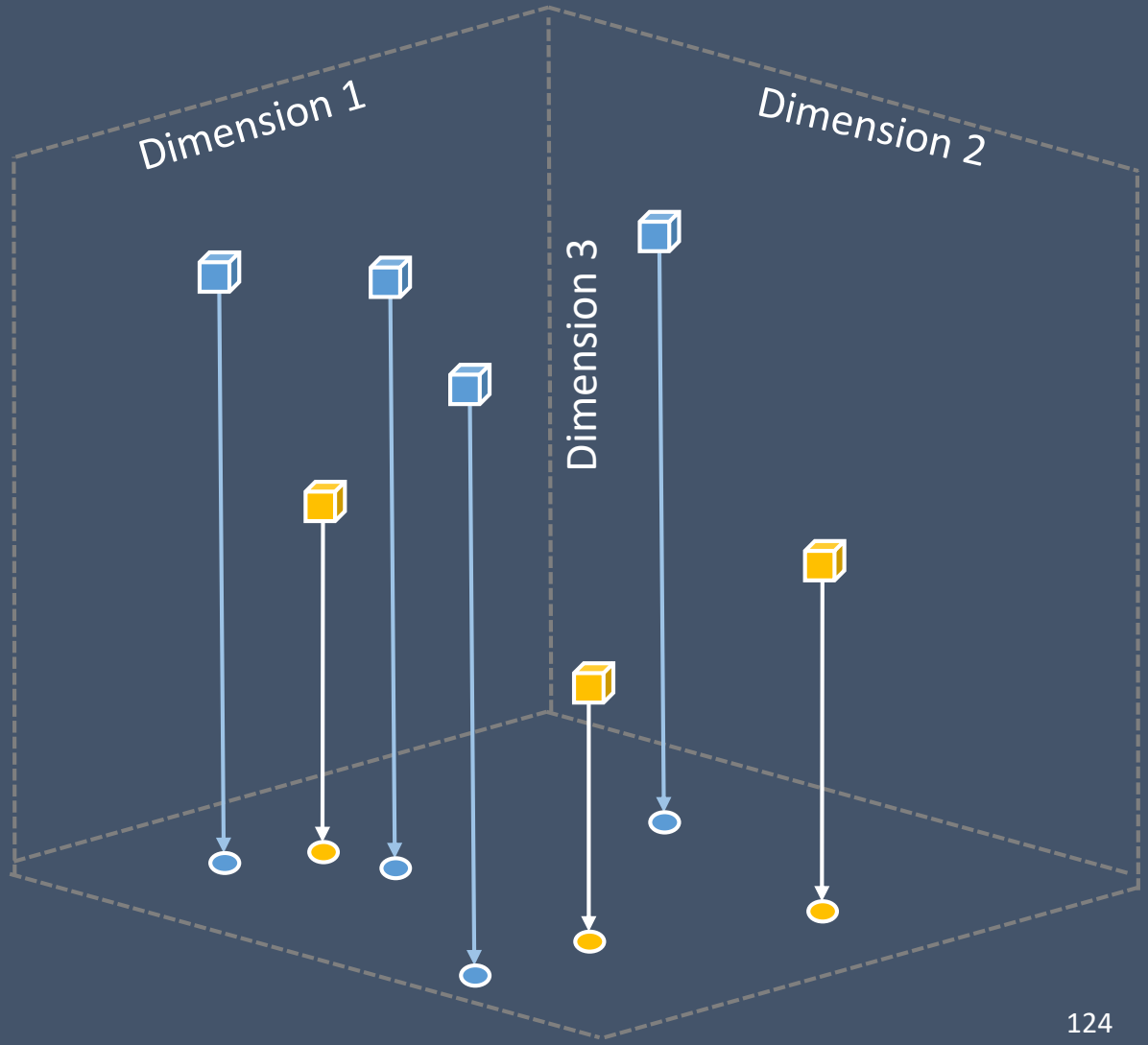
最短路徑中介中心性 計算方式
考量連結方向 → 有向圖
考量平行連結 → 加權計算

分群 計算方式 提供最大的Q值 (網絡分群最佳參數)
考量連結方向 → 有向圖
考量平行連結 → 加權計算

Gephi實務操作
Statistic 統計
8.6 網絡與分群_MDS

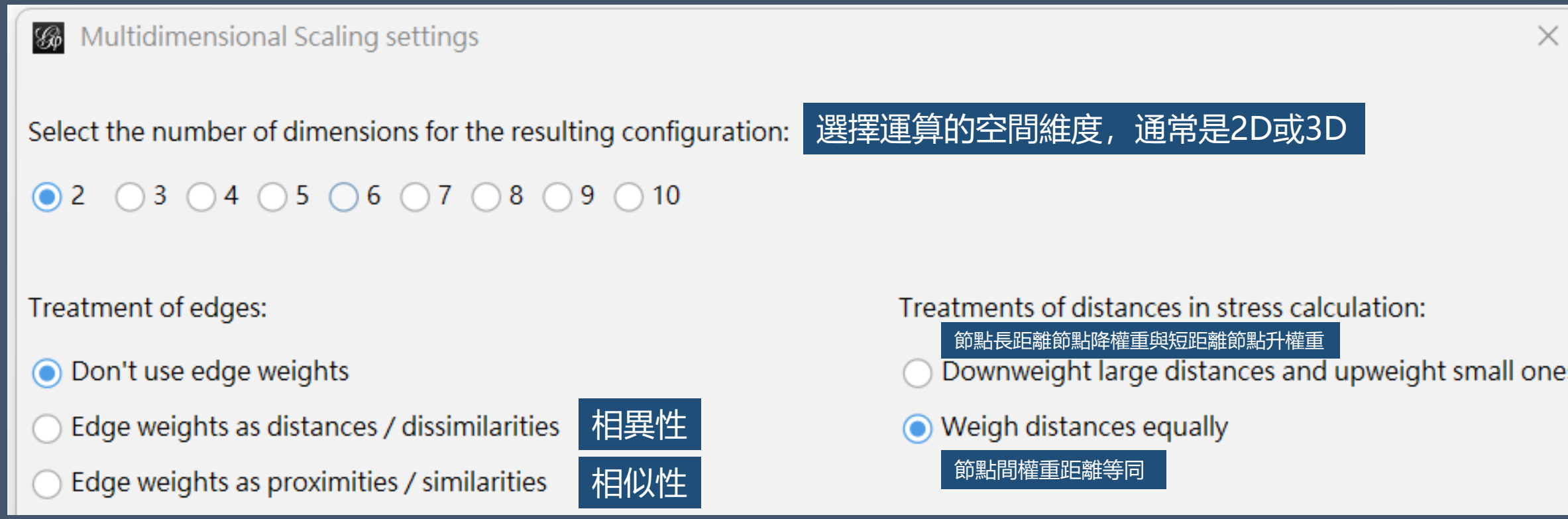
Multi-dimensional Scaling





Multi-dimensional Scaling

以距離遠近在多維空間地圖上呈現節點的相似性或相異性
節點間距離越小代表相似程度越高

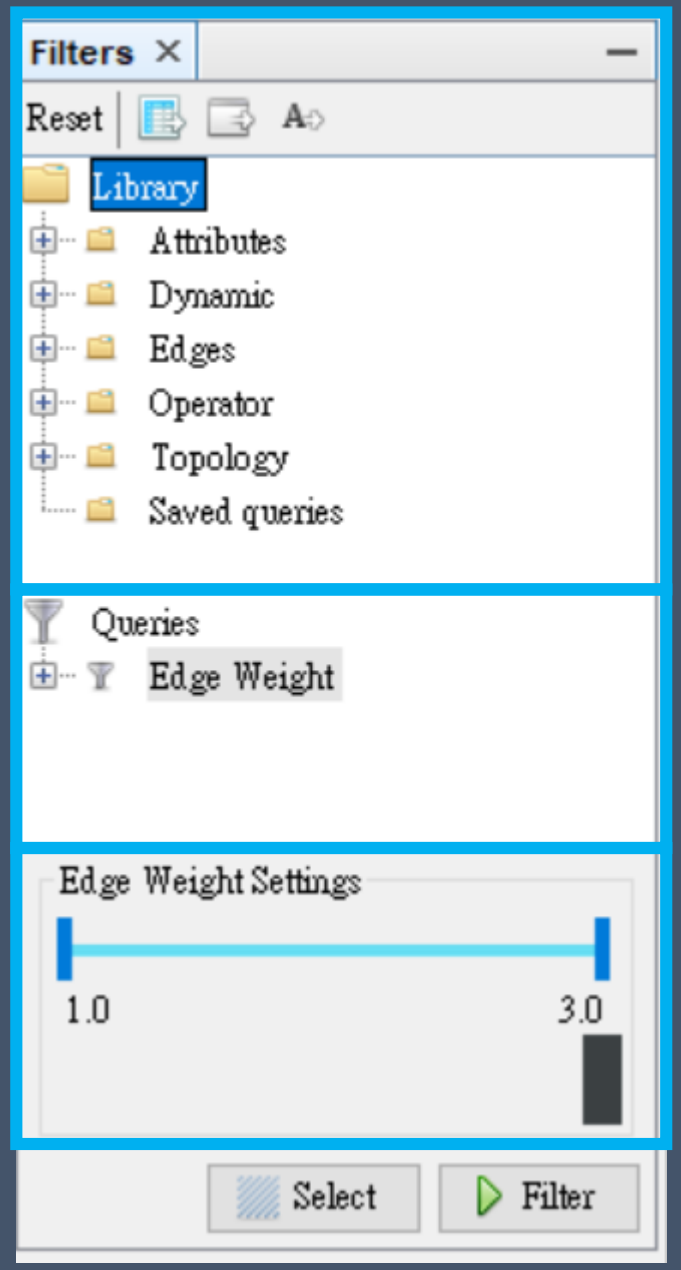


Gephi實務操作
09 Filter 篩選

篩選工具

設定條件

條件門檻值



1. Attributes節點屬性
2. Dynamic動態
3. Edges連結
4. Operator操作
5. Topology拓樸學
6. Save Queries 儲存篩選條件

Filter 將符合條件的節點保留在畫面上

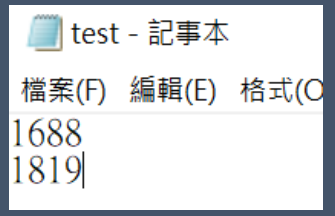
Select 保留整個網絡圖並將符合條件的節點凸顯出來

Gephi實務操作
Filter 篩選 9.1 Attribute

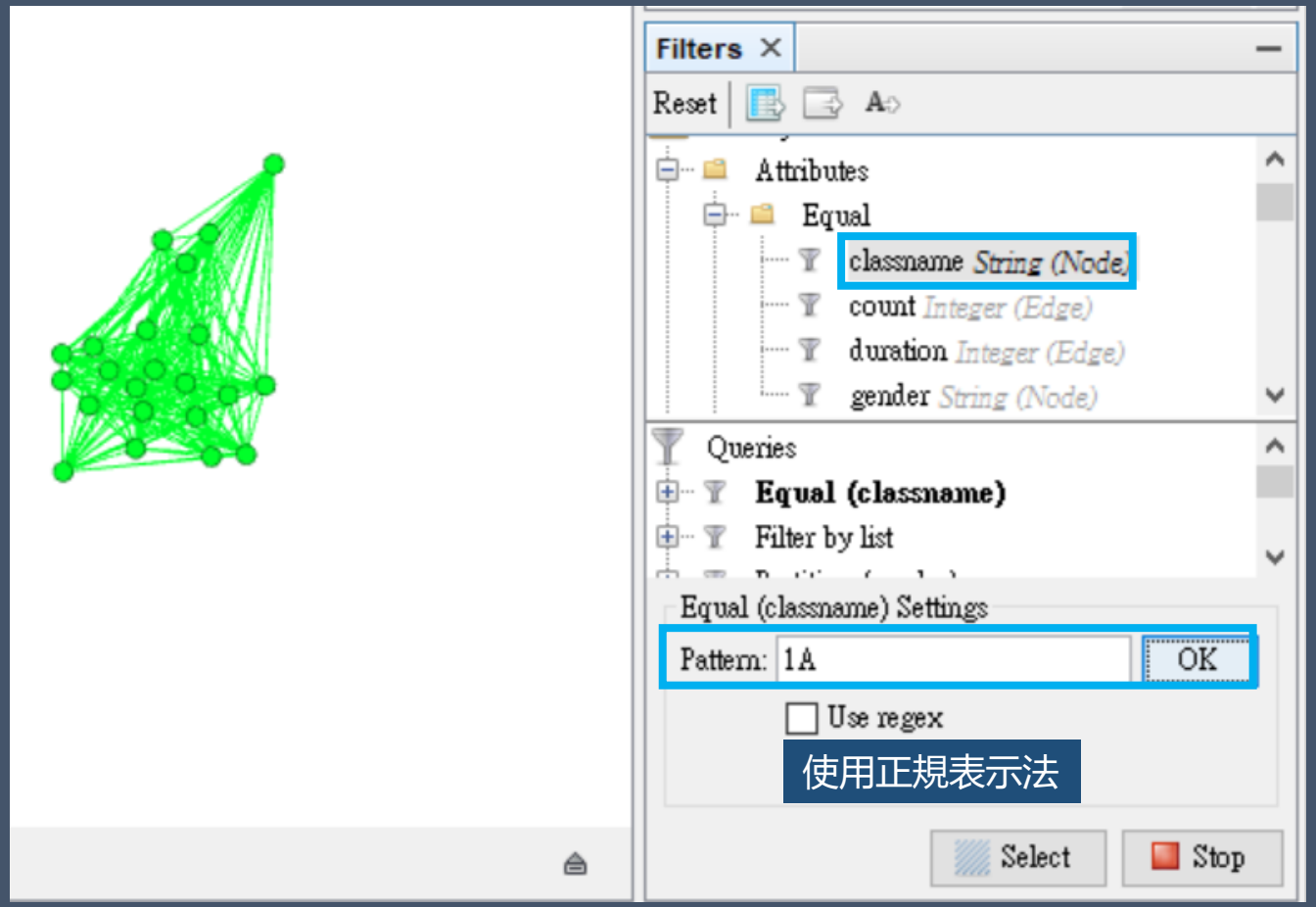
Attributes

Dataset: school day

1) Equal：等於篩選條件
2) Inter Edges：著重在分群後單一子群的內部(within)連結
   [inter edge within a group]
3) Intra Edges：用來找出各分群之間跨群(across-group)的連結路徑
   [intra edges across-group]
4) Non-null: 隱藏圖像的缺值(missing values)的節點與連結
5) Partion: 選擇節點屬性(node)以利觀察網絡內的子群體
6) Partion Count：利用節點次數(node)來篩選要分析的子群體
7) Range：選擇連結(edge)的屬性範圍以利觀察網絡特定節點群
8) Filter by label list: 利用記事本儲存節點標籤，
   鎖定特定節點進行篩選 a label, a row



Attributes 1) Equal

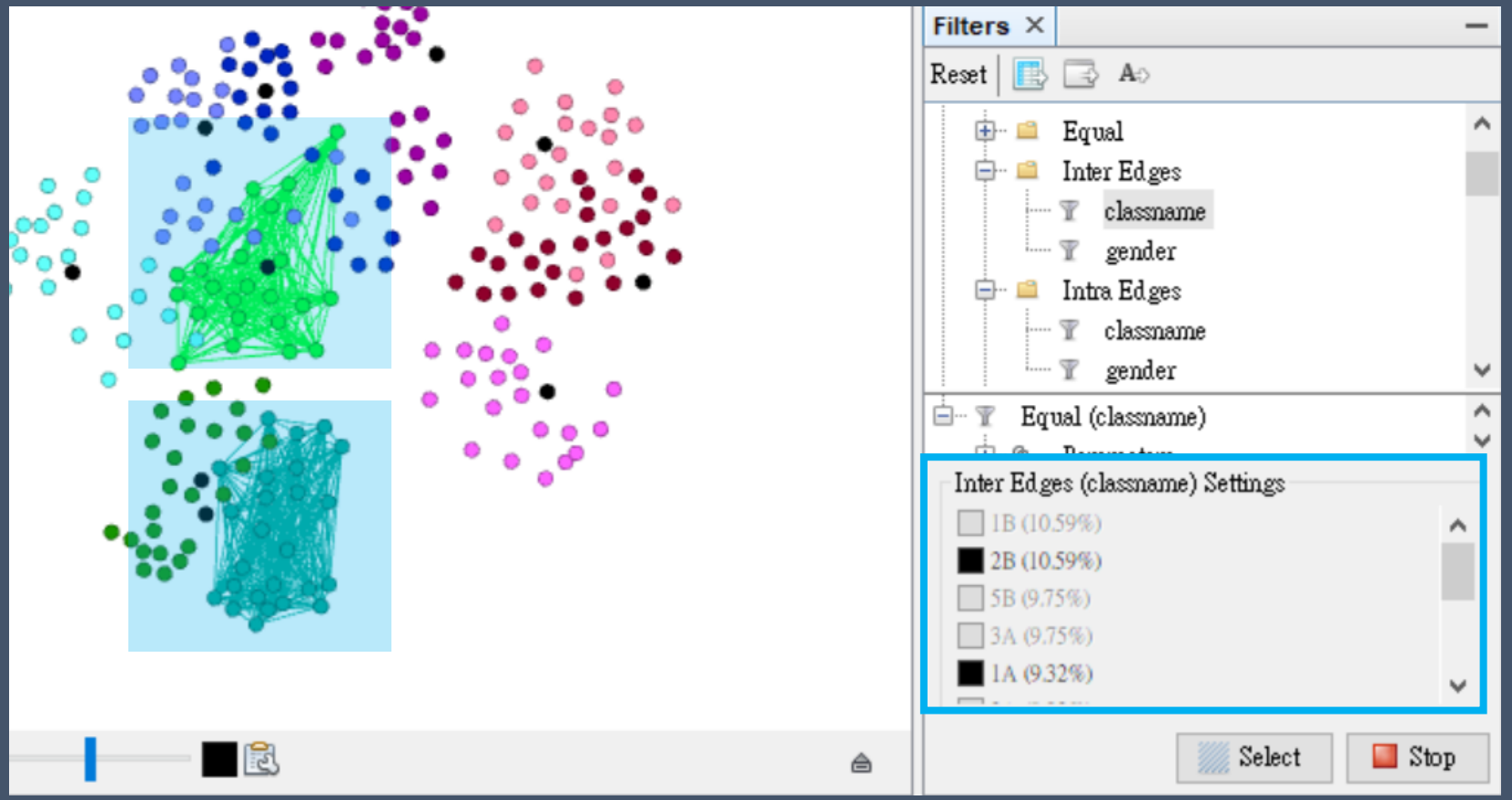


① 選擇篩選的屬性

② 輸入要篩選的對象：1A(班級) 按下OK

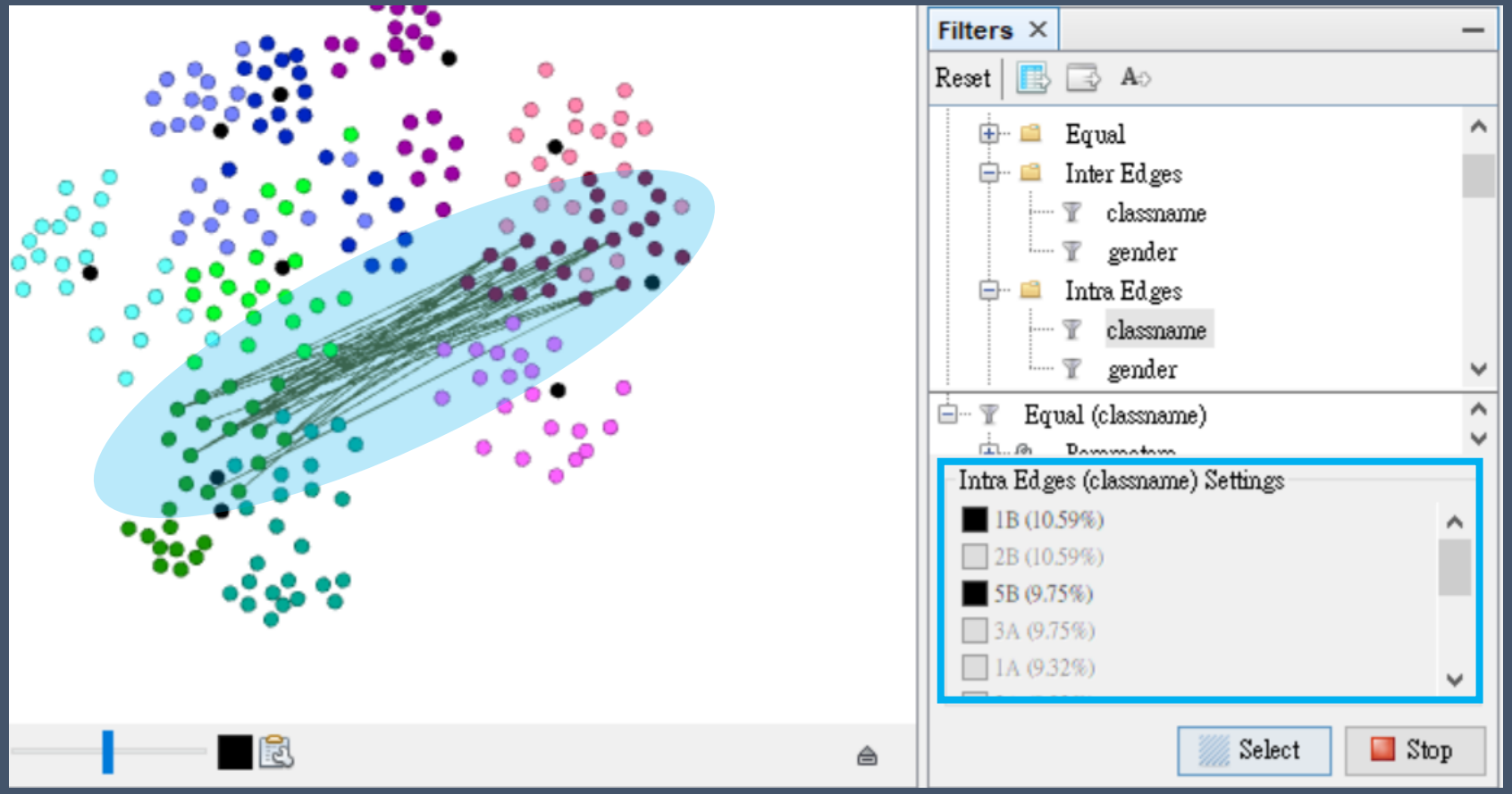
③ 選擇Filter

Attributes 2) Inter Edges



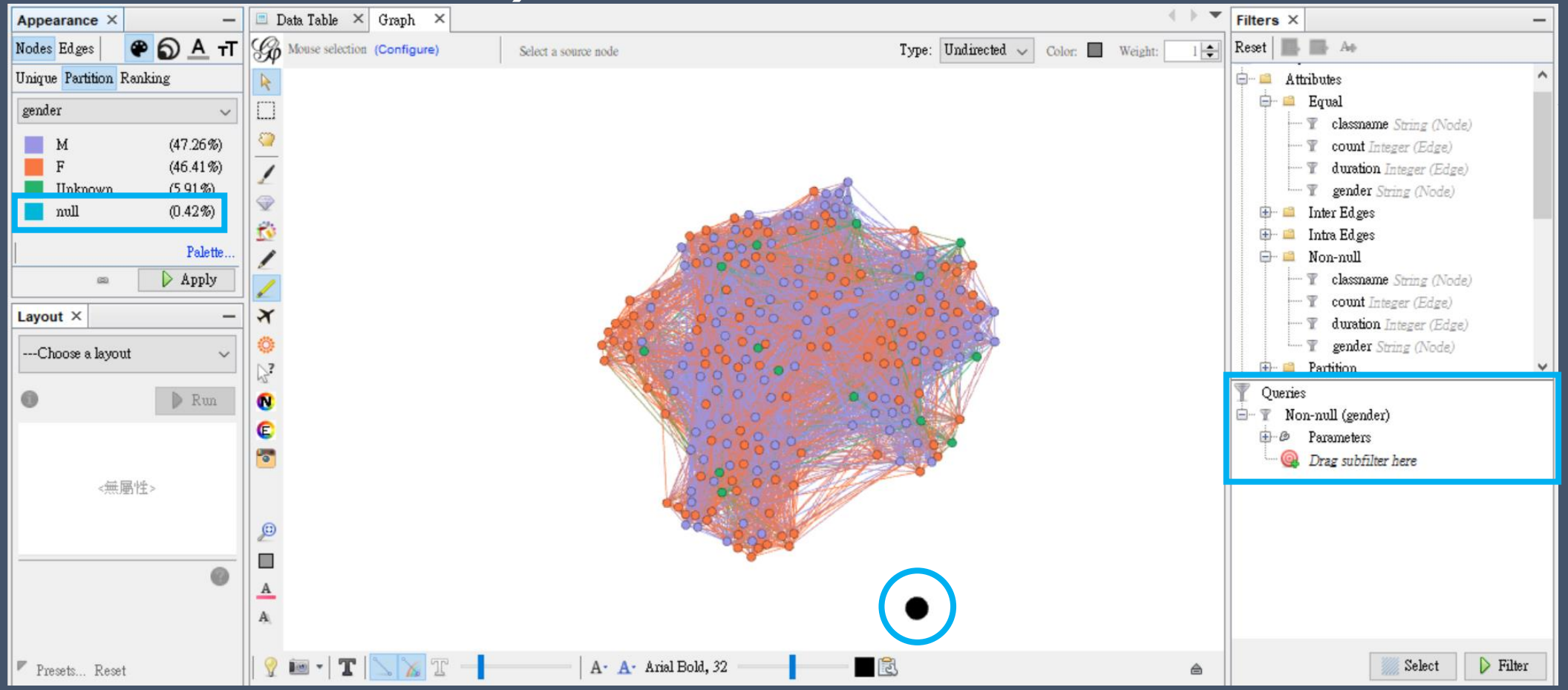
著重在分群後單一子群的內部(within)連結，
用於處理同質性較高的網絡或密度較高的網絡

Attributes 3)Intra Edges



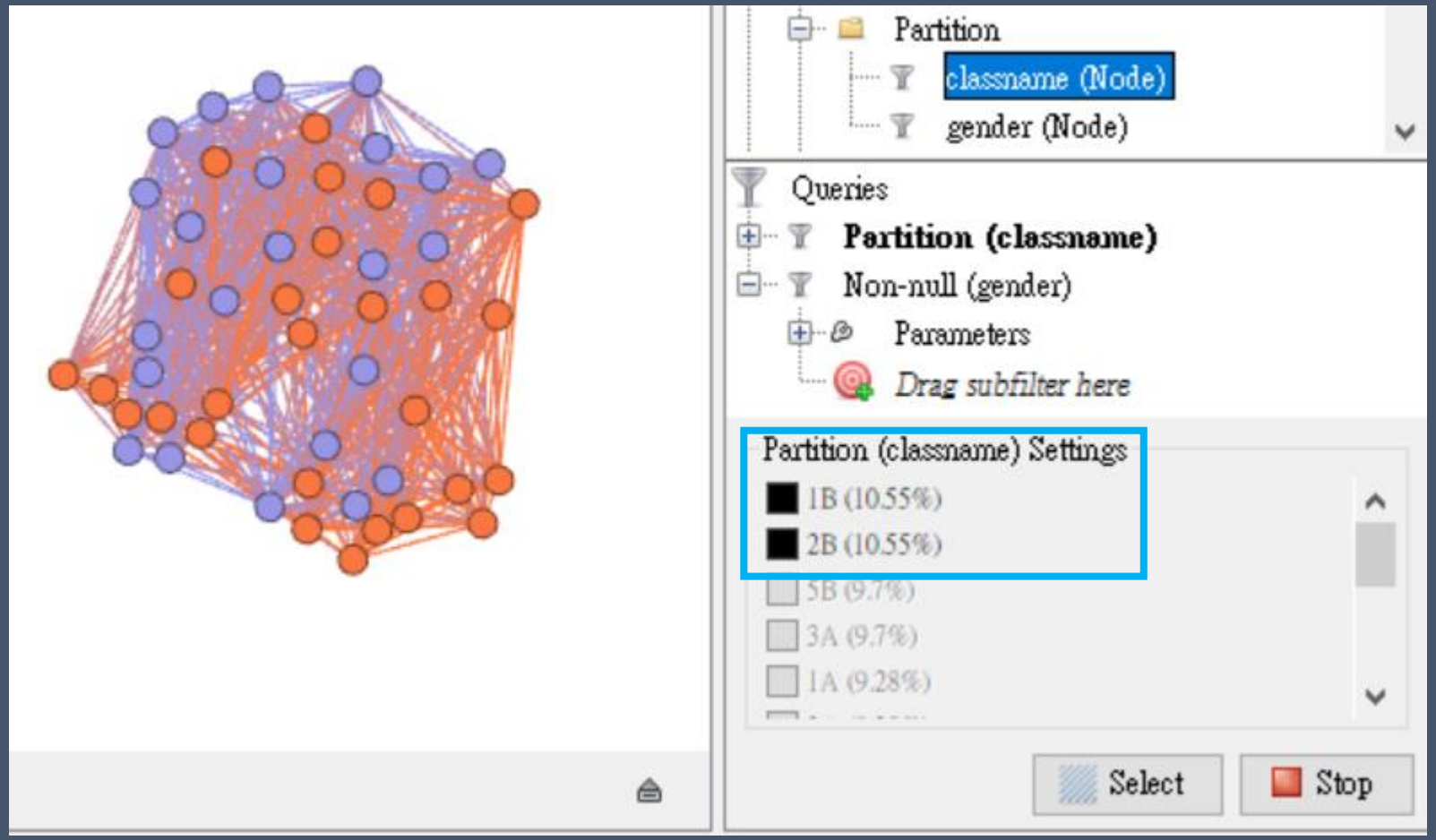
著重分群後跨群(cross-group)之間連結，
用來找出各分群之間的連結路徑

Attributes 4) Non-null



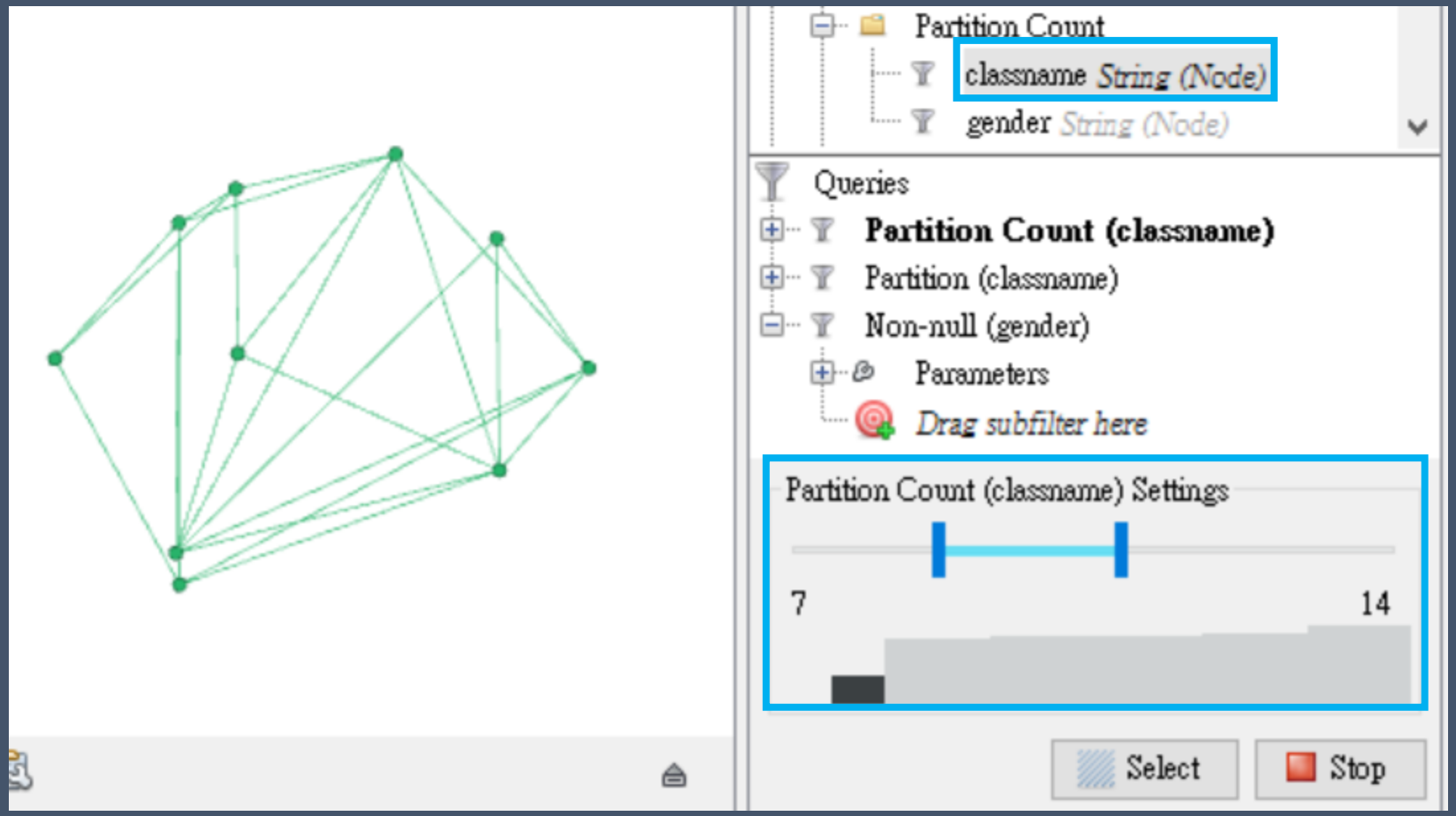
隱藏圖像的缺值(missing values)的節點與連結
注意：範例資料資料集未有空值，操作前請在範例資料集自行新增節點

Attributes 5) Partition



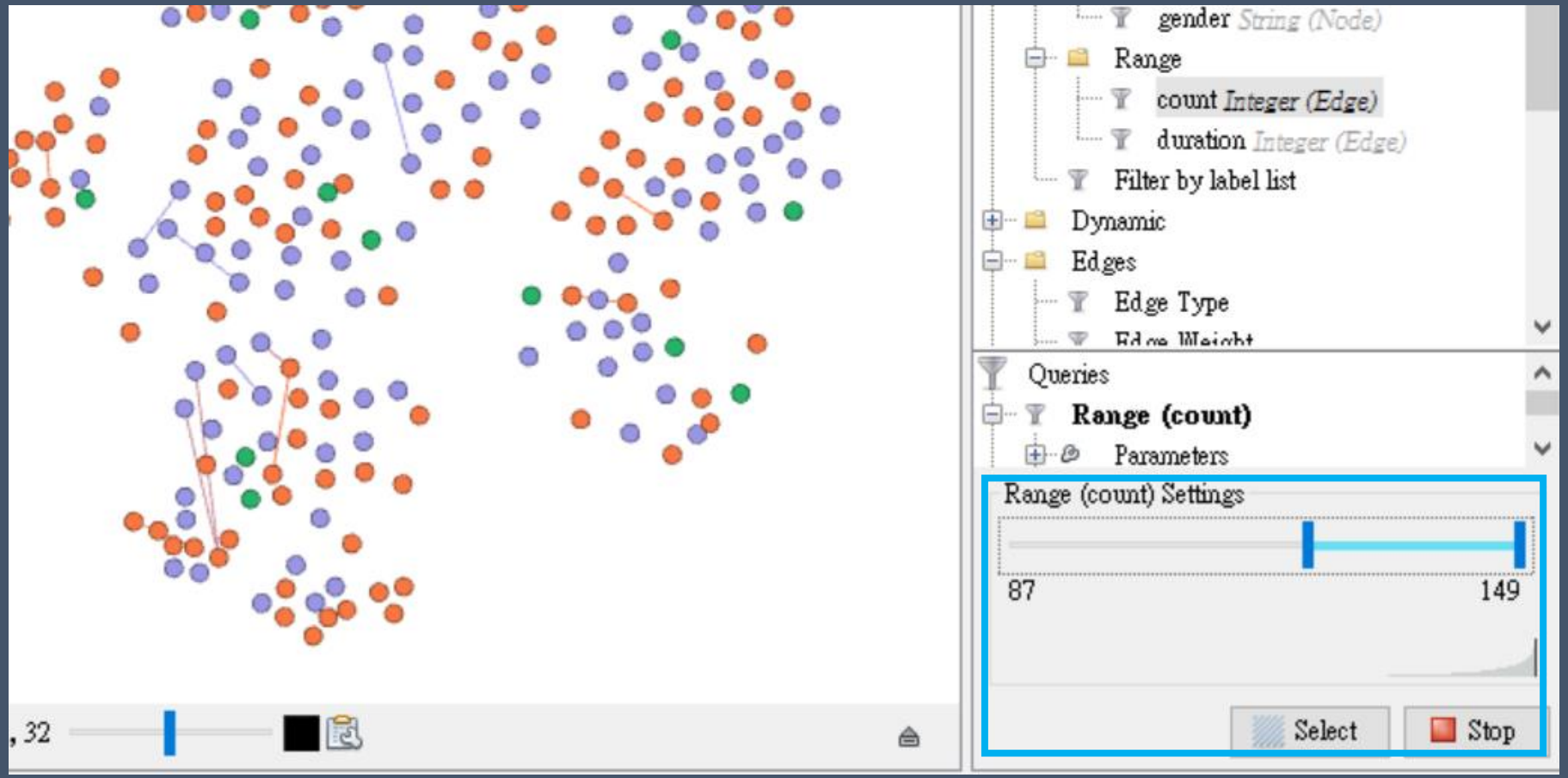
選擇屬性以利觀察網絡內的子集，EX：1B
可選多個子群一同觀察

Attributes 6) Partion Counts



利用節點次數來篩選要分析的子群體，
可鎖定特大或集小的節點群

Attributes 7) Range



選擇連結(edge)的屬性範圍以利觀察網絡特定節點群

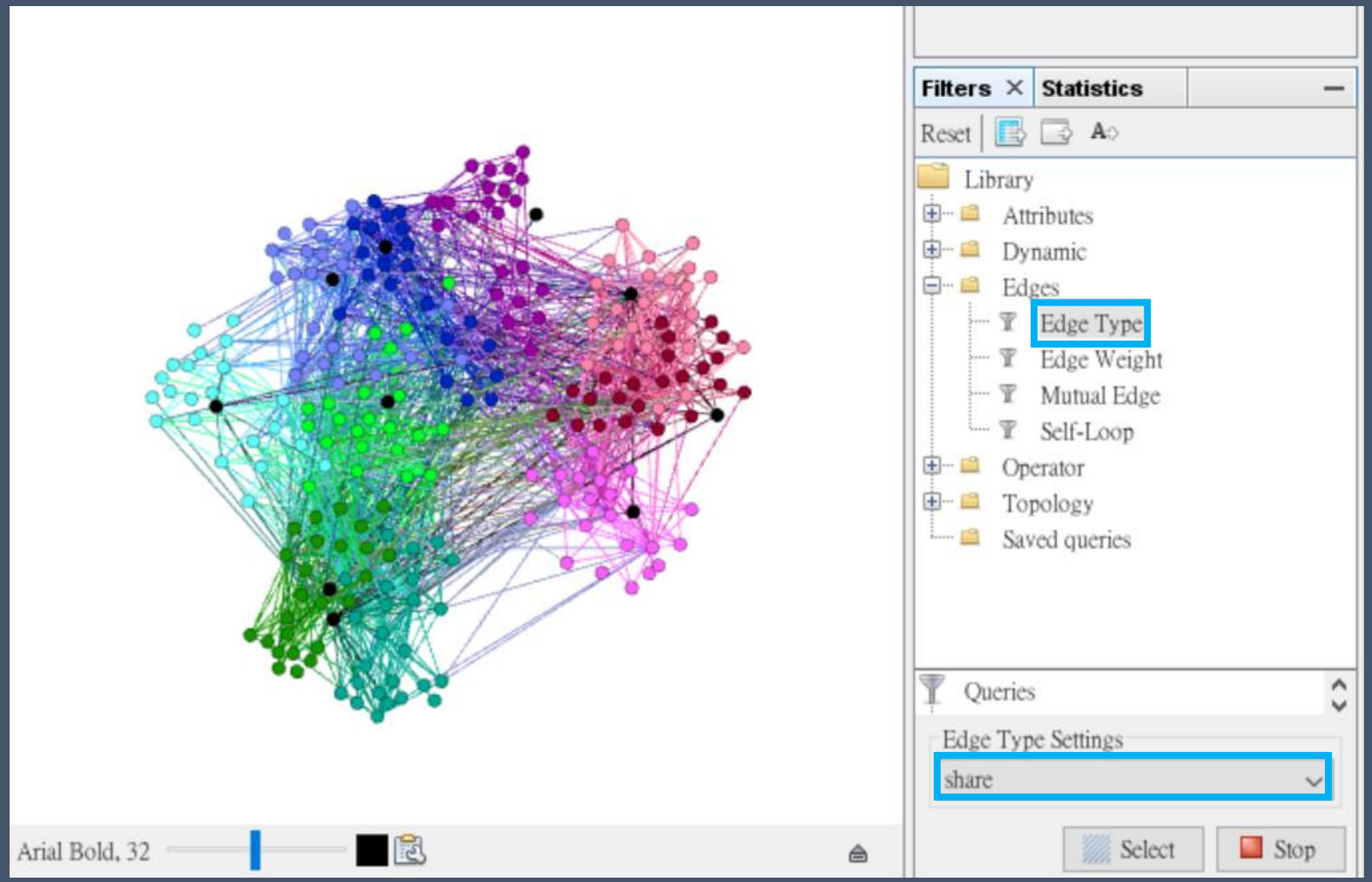
Gephi實務操作
Filter 篩選 9.2 Edges

Edges

Dataset: Multigraph Example

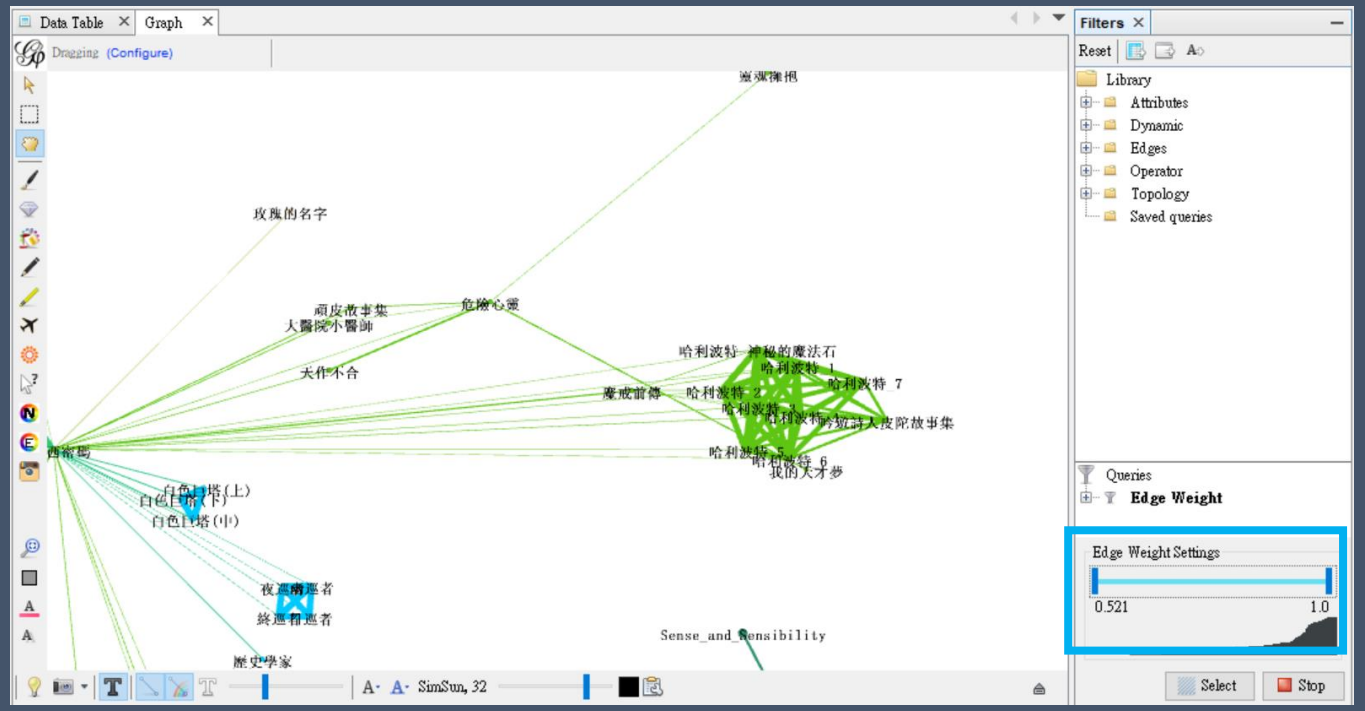
1) Edge type：用來篩選不同種類的連結
2) Edge Weight：利用連結權重值篩選需要的連結
3) Mutual Edge：篩選單向連結，僅保留雙向連結
4) Self-Loop：移除節點的自我連結(自環)

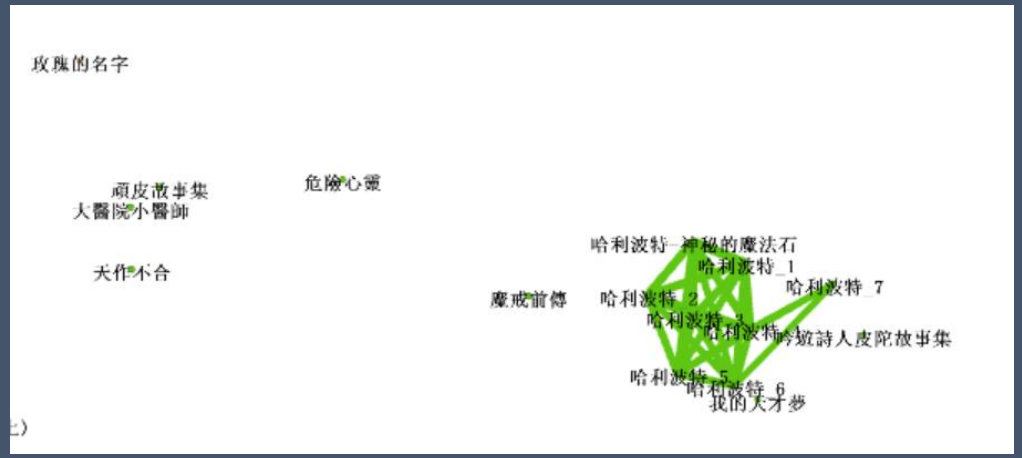
Edges 1) Edge type 用來篩選不同種類的連結



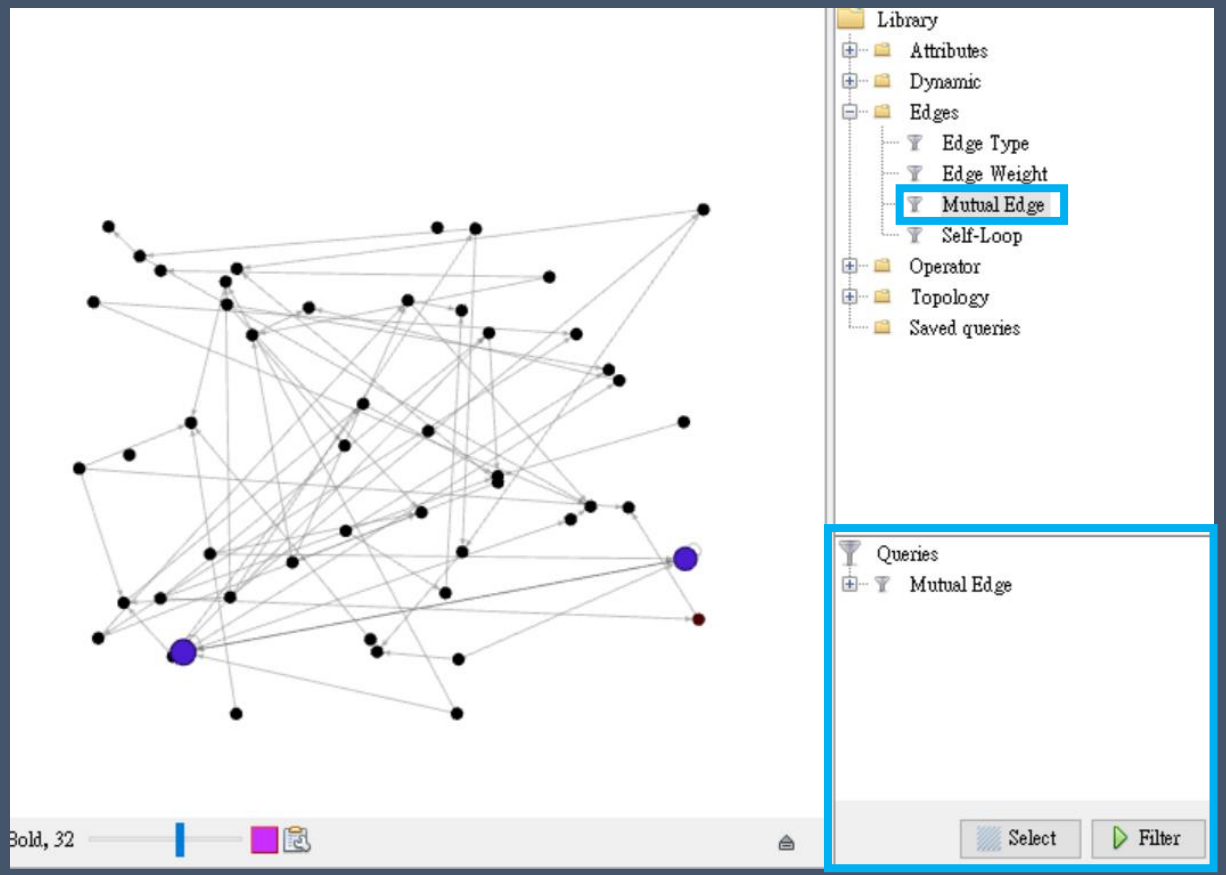
Edges 2) Edge weight

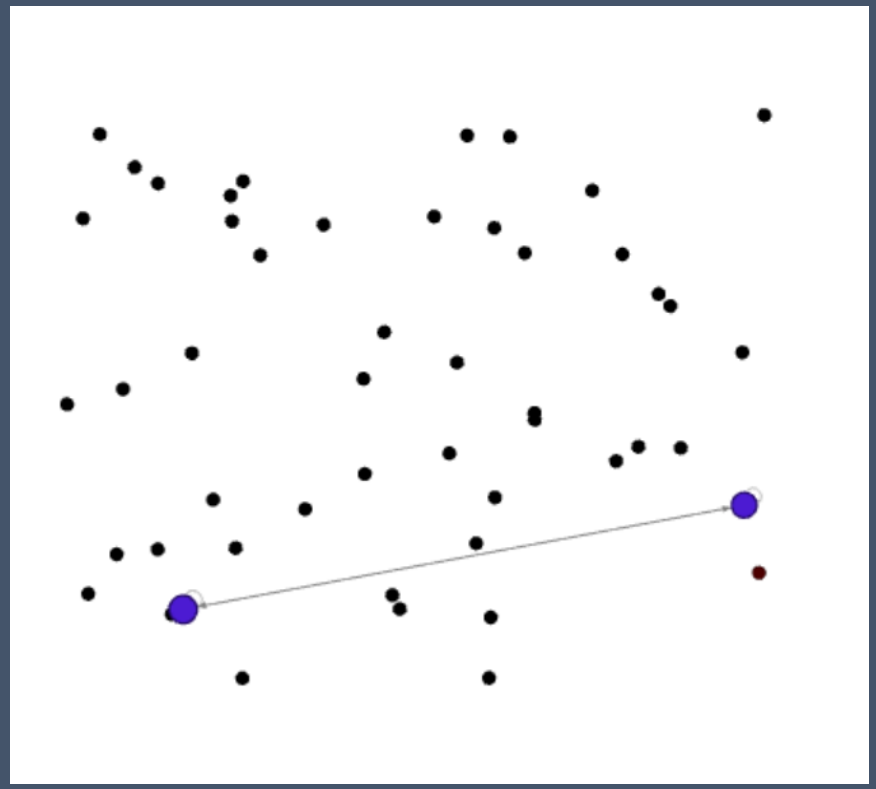
利用連結加權篩選需要的連節





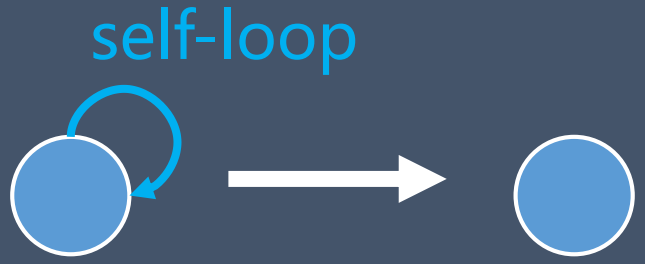
Edges 3) Mutual Edges





篩選單向連結，僅保留雙向連結

Edges 4) Self-Loop

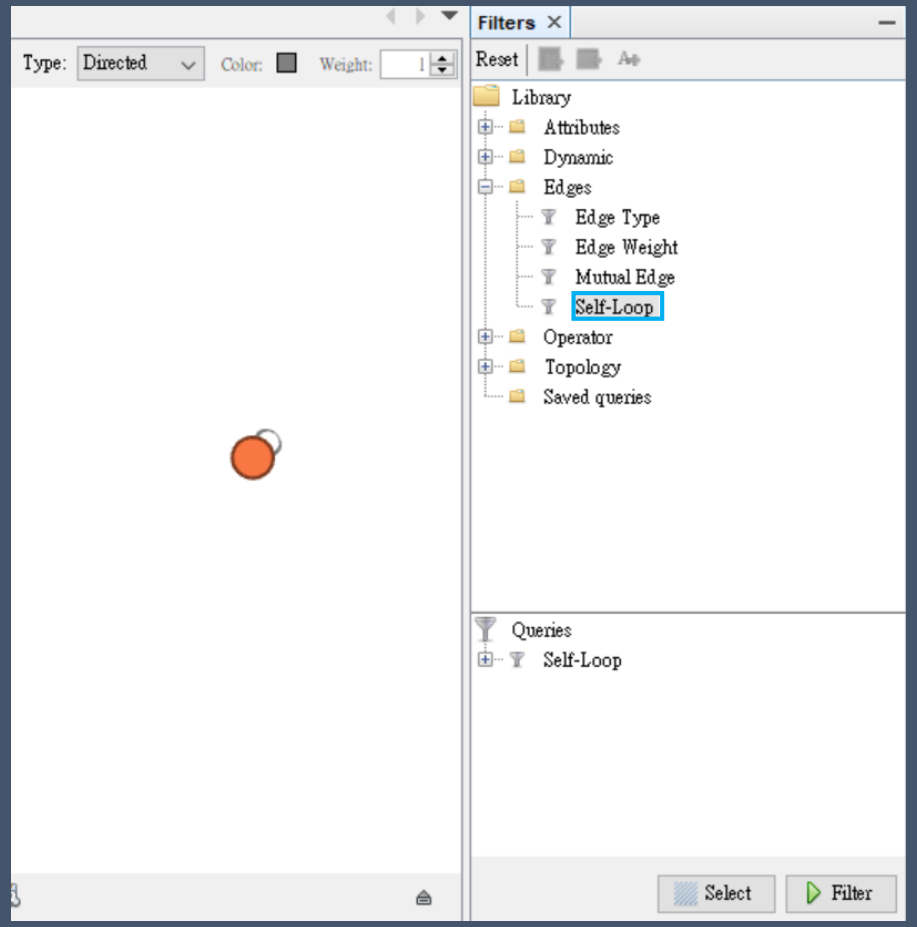


移除節點的自環連結

→移除節點自己連自己的連結

→該節點與其對外連結都不會消失

(自環連結會影響中心性程度的計算)



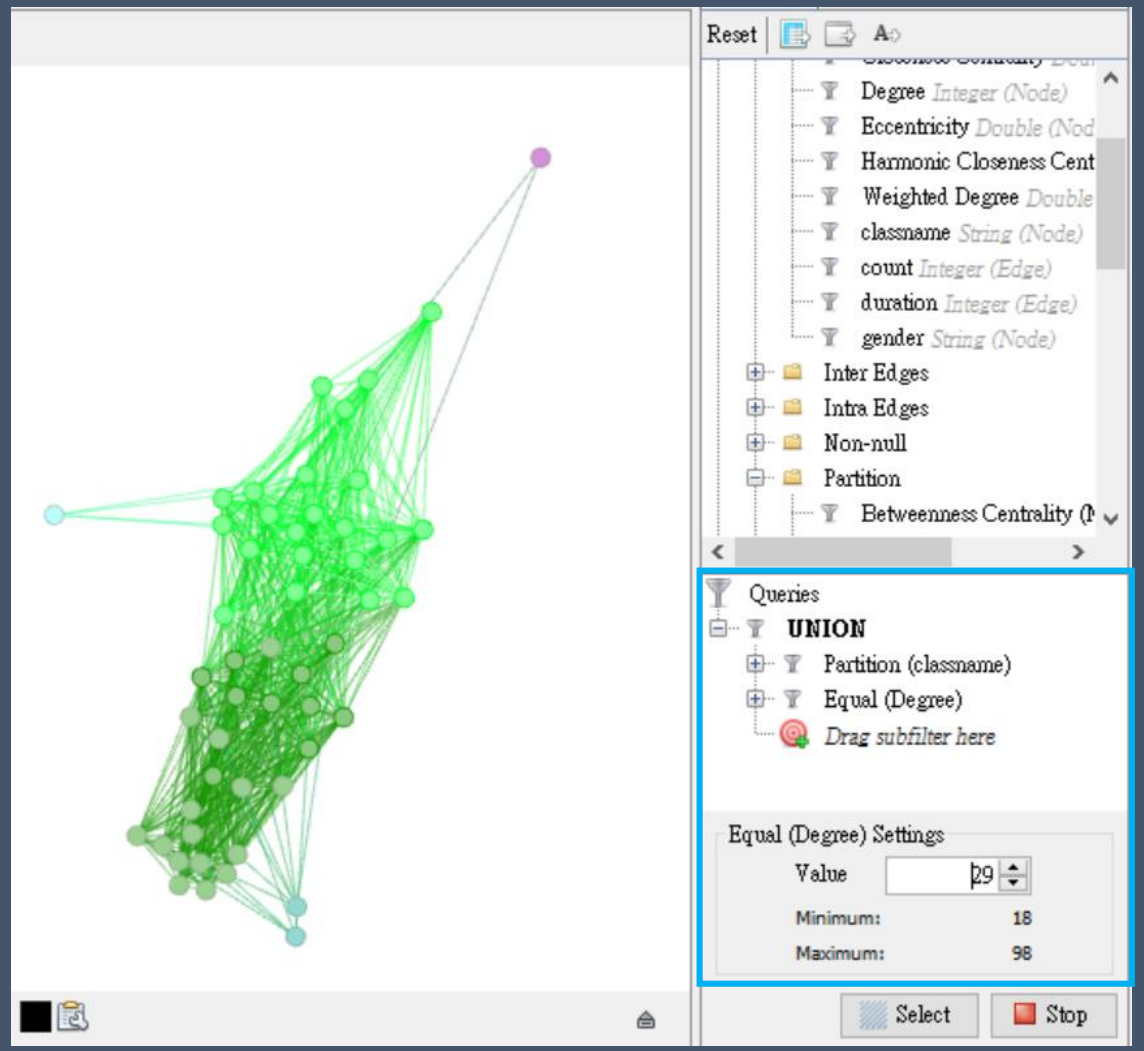
Gephi實務操作
Filter 篩選 9.3 Operator

Operator

Dataset: school day

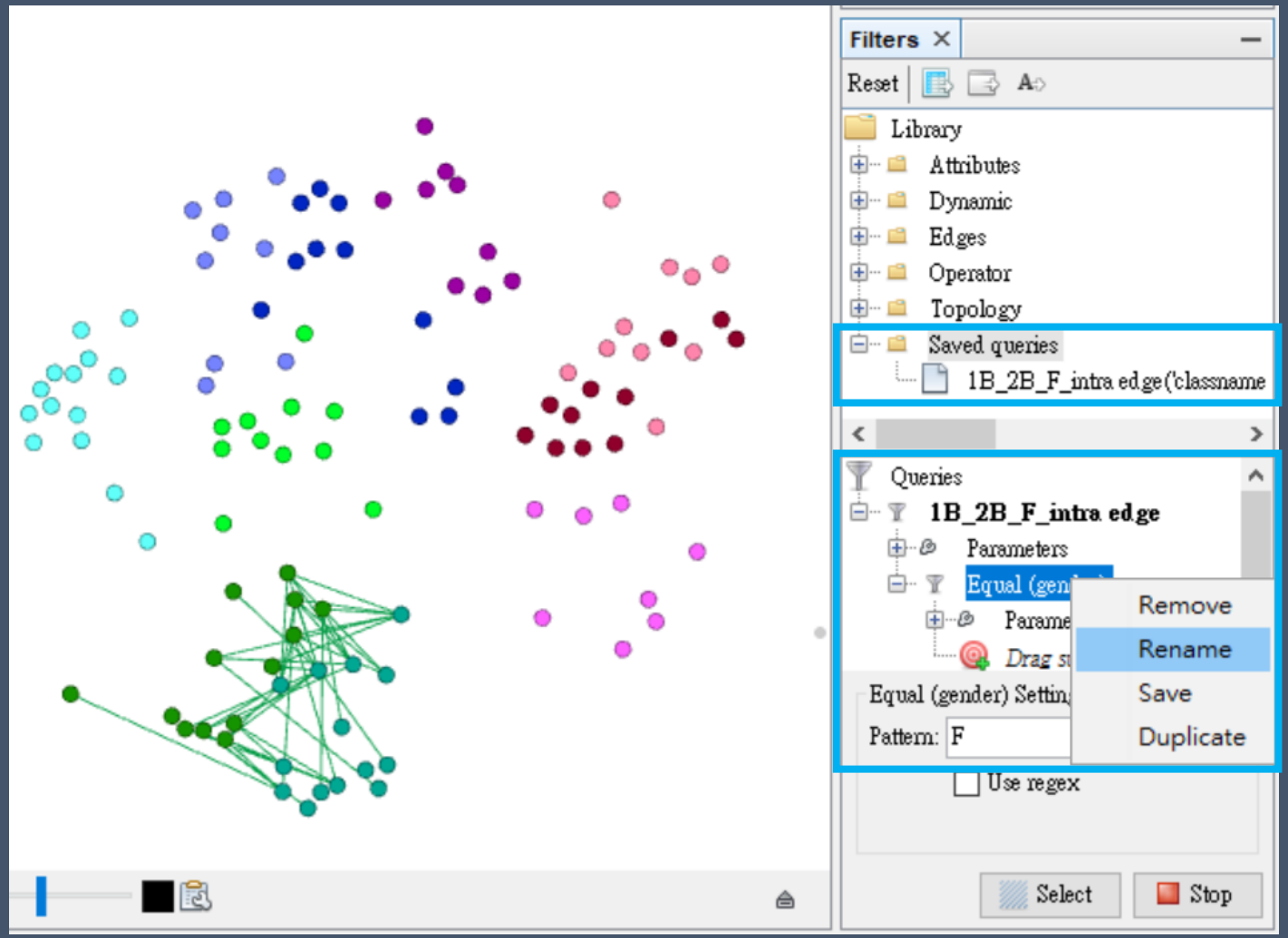
1) Intersection：篩選交叉節點
   +Attribute>PartitionCounts>Degrees
2) Mask(Edges)：篩選特定範圍連結
3) NOT(Edges)：不顯示特定連結　+Attribute>Inter/Intra Edges
4) NOT(Nodes)：不顯示特定節點
5) UNION：聯合設定多個篩選條件 +Attribute>Equal or Partition

5) UNION+Equal/Partition



Gephi實務操作
Filter 篩選 9.4 Saved queries

Saved Queries



完成下列兩步驟後
Saved queries 出現已存的篩選組合

① 選取要設定的篩選組合
Ex：1B與2B兩班女同學來往關係
Partion >1B and 2B
+
Equal(gender) >F [記得按OK]
② 將篩選組合重新命名，再儲存

注意：一定要再Queries視窗先命名，儲存到Saved Queries就無法命名