

# Practical Social Network Analysis With Gephi

Dr. Derek Greene  
Insight @ UCD

# Gephi - Introduction

- Gephi: Open source interactive network exploration and visualisation tool for Windows, Linux and Mac OS X.

The screenshot shows the Gephi website homepage. At the top, there's a navigation bar with links for Download, Blog, Store, Wiki, Forum, Support, and Bug tracker. Below the navigation is a main menu with Home, Features, Learn, Develop, Plugins, Services, and Consortium. The central part of the page features a large image of the Gephi software interface, which displays a complex network graph with many nodes and edges, color-coded by community. To the left of the interface, there's a section titled "The Open Graph Viz Platform" with text about Gephi being an interactive visualization and exploration platform for networks and complex systems. It also mentions that Gephi runs on Windows, Linux, and Mac OS X, is open-source, and is free. A "Download FREE Gephi 0.8.2-beta" button is prominently displayed, along with links for Release Notes and System Requirements. Below the download button are links for Features, Quick start, Screenshots, and Videos.

[Gephi: an open source software for exploring and manipulating networks.](#)

[M Bastian, S Heymann, M Jacomy - ICWSM, 2009 - aaai.org](#)

Abstract Gephi is an **open source software** for graph and **network** analysis. It uses a 3D render engine to display large **networks** in real-time and to speed up the exploration. A flexible and multi-task architecture brings new possibilities to work with complex data sets ...

Cited by 708 Related articles All 15 versions Import into BibTeX Saved More

<http://gephi.org>

# Input Data - Graph Formats

- First task - loading your network data into Gephi.
- Gephi supports loading and saving graphs in a number of formats.
- Simplest approach is to use **comma-separated (CSV)** data exported from tools such as Excel or R.

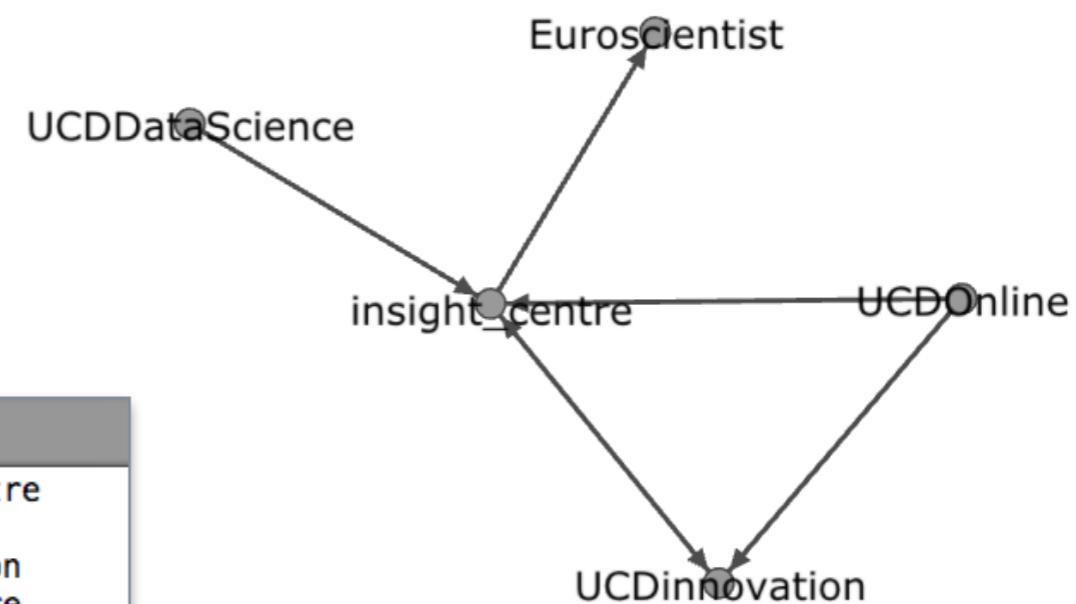
Simple binary graph

No weights or attributes

|   | A              | B              | C |
|---|----------------|----------------|---|
| 1 | UCDDataScience | insight_centre |   |
| 2 | UCDOnline      | insight_centre |   |
| 3 | insight_centre | UCDInnovation  |   |
| 4 | UCDInnovation  | insight_centre |   |
| 5 | insight_centre | Euroscientist  |   |
| 6 | UCDOnline      | UCDInnovation  |   |
| 7 |                |                |   |

eg\_ucd.csv

```
UCDDataScience,insight_centre
UCDOnline,insight_centre
insight_centre,UCDInnovation
UCDInnovation,insight_centre
insight_centre,Euroscientist
UCDOnline,UCDInnovation
```



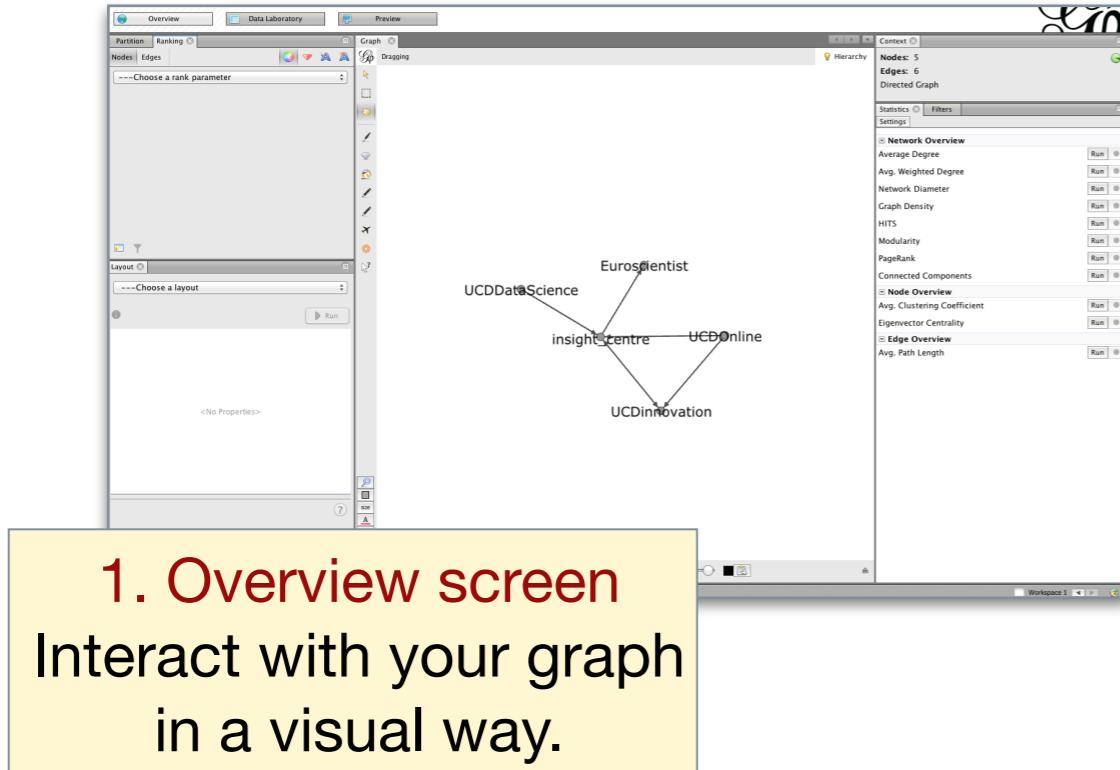
On loading, Gephi will ask whether graph is directed or undirected.

# Input Data - Graph Formats

- Gephi supports more comprehensive file formats which can store node and edge attributes, together with layout and presentation information (e.g. position, size, colour etc).
- Native format is a **.gephi** file, which can contain multiple “workspaces” (i.e. several different graphs).
- Gephi also supports several open formats which can be used to exchange data with other tools. Common examples:
  - **GEXF**: XML Graph Exchange Format for complex networks.  
<http://gexf.net>
  - **GraphML**: XML Graph Markup language  
<http://graphml.graphdrawing.org>  
<http://cs.brown.edu/~rt/gdhandbook/chapters/graphml.pdf>
  - **GML**: Plain text Graph Modeling Language for describing graphs  
[http://en.wikipedia.org/wiki/Graph\\_Modelling\\_Language](http://en.wikipedia.org/wiki/Graph_Modelling_Language)

# Gephi Screens

- Three key screens in Gephi, accessible from top buttons...



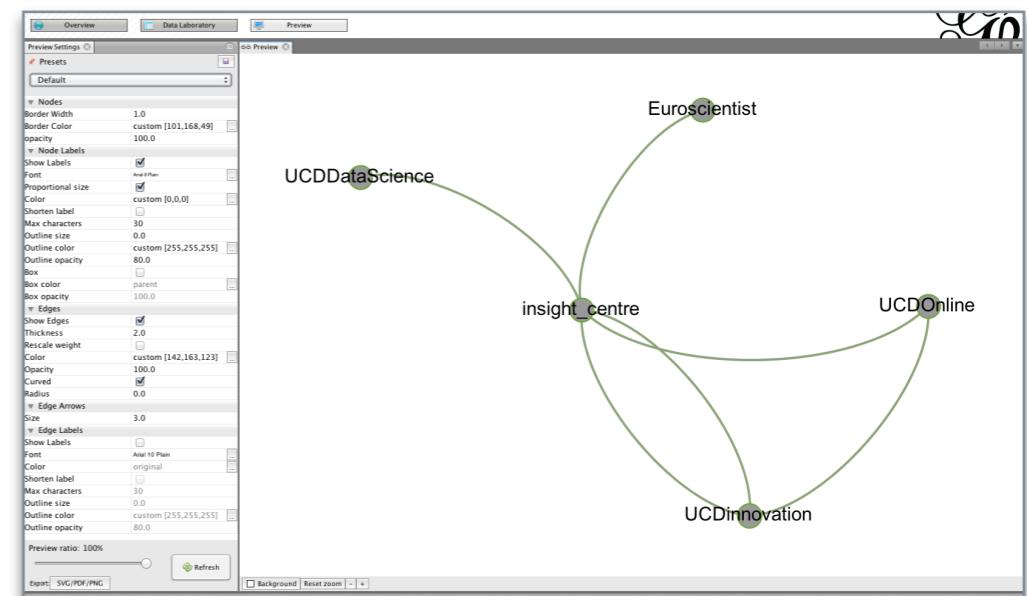
**1. Overview screen**  
Interact with your graph  
in a visual way.

| Nodes          | id             | Label          |
|----------------|----------------|----------------|
| UCDDataScience | UCDDataScience | UCDDataScience |
| insight_centre | insight_centre | insight_centre |
| UCDOnline      | UCDOnline      | UCDOnline      |
| UCDInnovation  | UCDInnovation  | UCDInnovation  |
| Euroscientist  | Euroscientist  | Euroscientist  |

**2. Data Laboratory screen**  
Explore graph node & edge  
attribute data in tabular form.

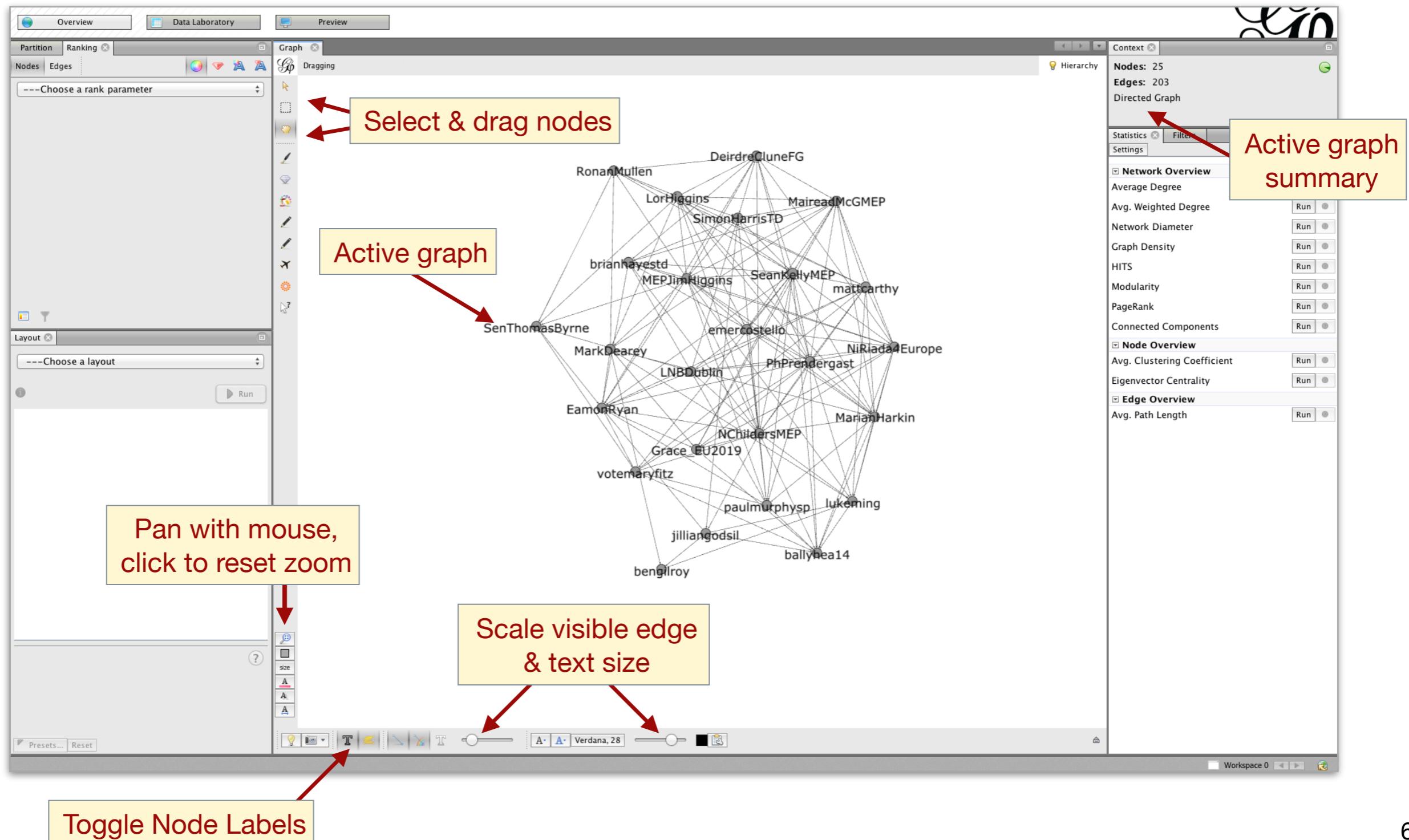
Each screen  
contains multiple  
tabs with specific  
functionality.

**3. Preview screen**  
Fine-tune and render  
your graph for export.



# Graph Overview Screen

- Gephi defaults to the **Overview** screen - the “draft” interactive view.



# Graph Layouts

- Common first step is to apply a **layout algorithm** to re-position nodes in the graph so as to improve its readability and aesthetics.

Choose algorithm in “Layout” tab

Force Atlas

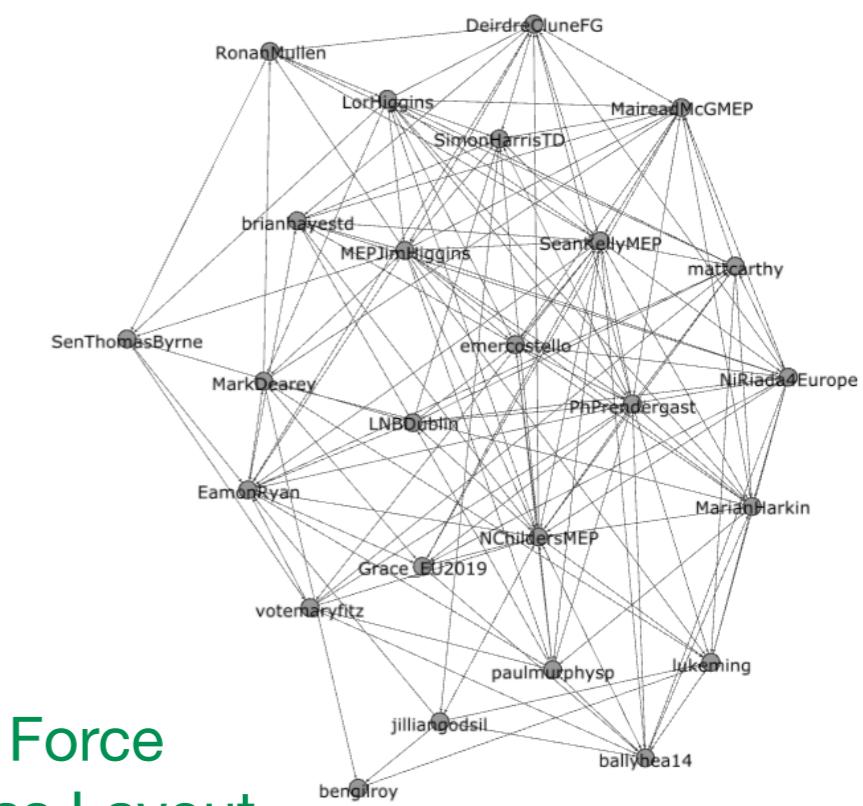
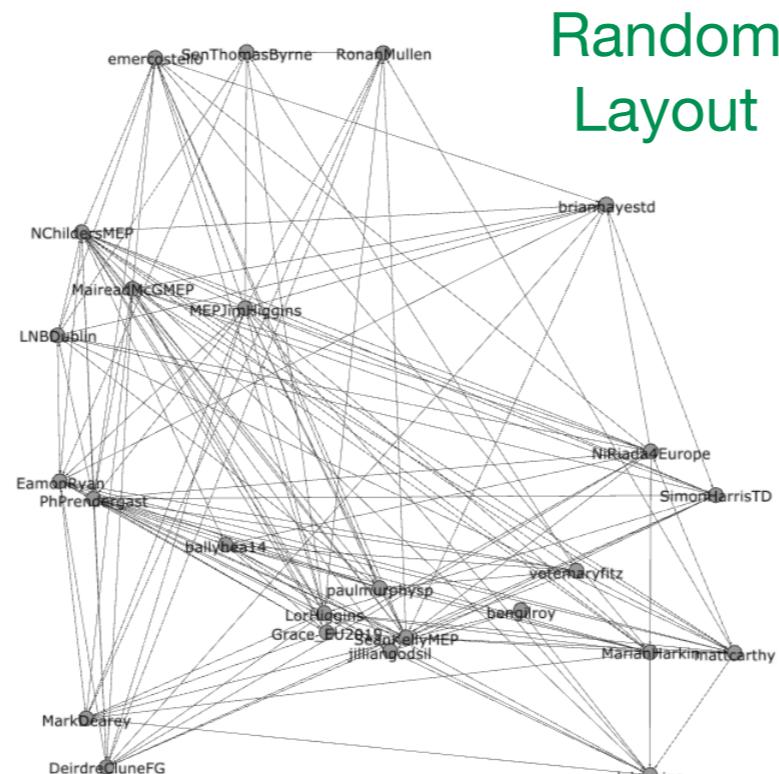
Run

Force Atlas

|                         |                                     |
|-------------------------|-------------------------------------|
| Inertia                 | 0.1                                 |
| Repulsion strength      | 1000.0                              |
| Attraction strength     | 1.0                                 |
| Maximum displacement    | 10.0                                |
| Auto stabilize function | <input checked="" type="checkbox"/> |
| Autostab Strength       | 80.0                                |
| Autostab sensibility    | 0.2                                 |
| Gravity                 | 30.0                                |
| Attraction Distrib.     | <input type="checkbox"/>            |
| Adjust by Sizes         | <input checked="" type="checkbox"/> |
| Speed                   | 1.0                                 |

Properties to tweak algorithm

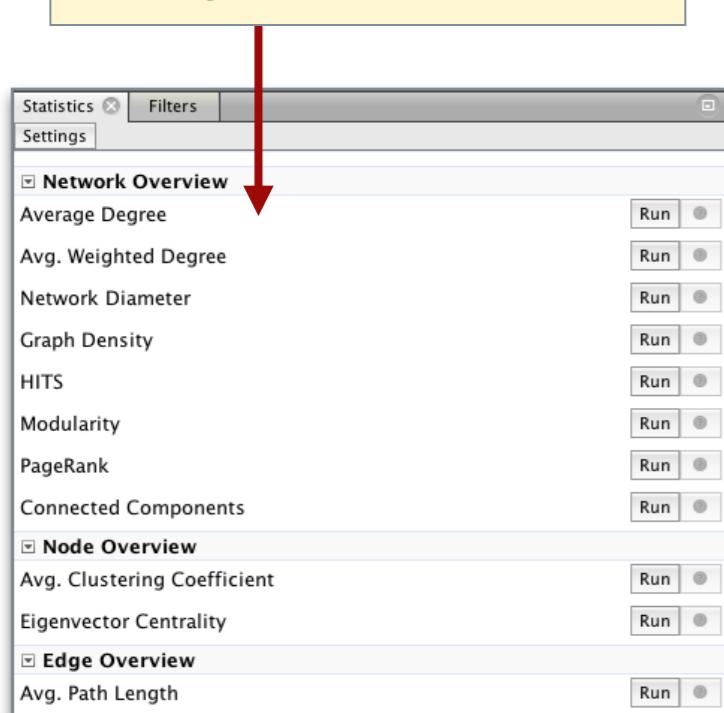
→ Frequent choice is a “**force atlas**” algorithm, which tries to ensure as few edges cross as possible.



# Graph Statistics

- Gephi provides a range of metrics for calculating statistics that characterise a graph and its nodes.

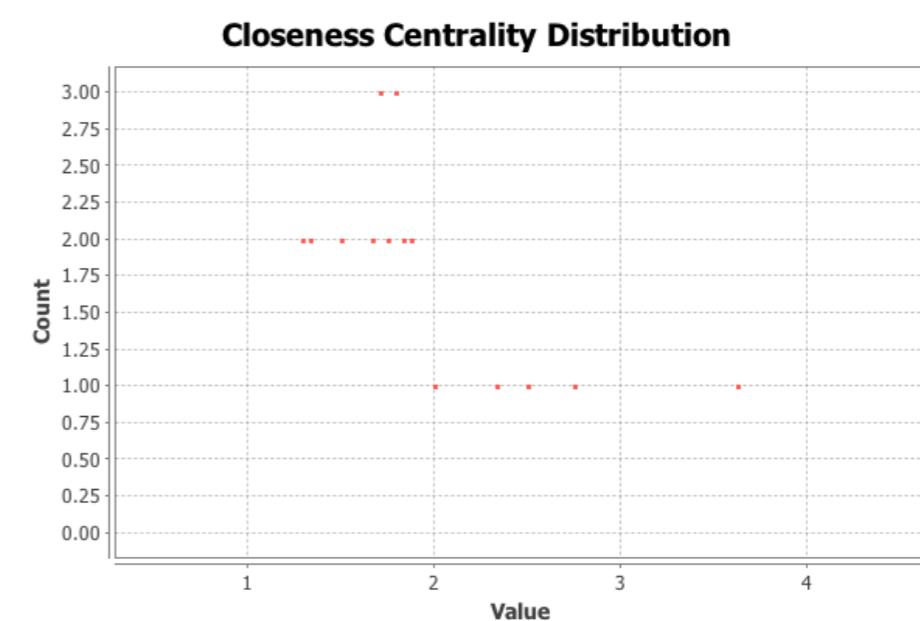
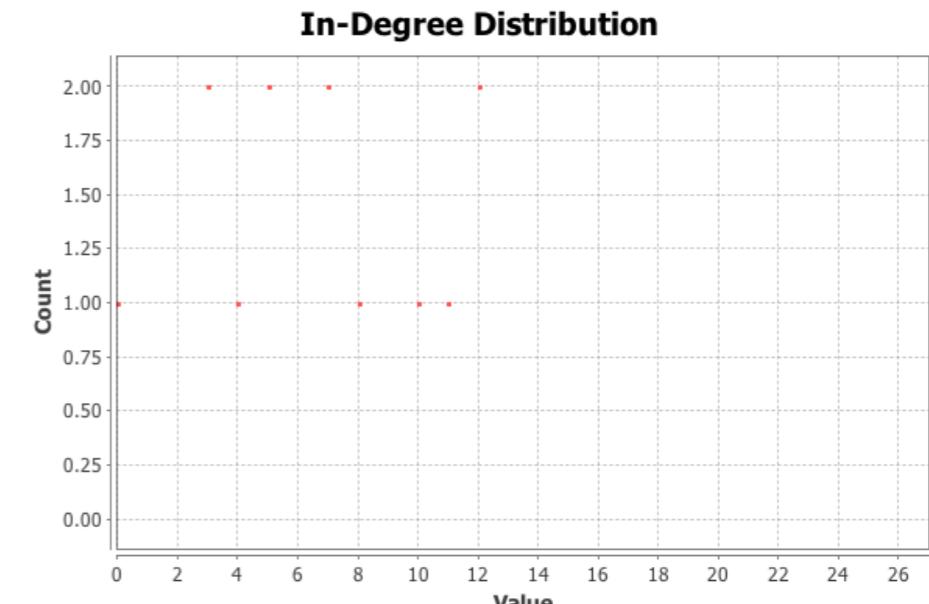
Run a metric from  
“Statistics” tab



- Run “Average Degree”:  
Produces distribution plot  
of node in-degrees  
(user “popularity”)
- Run “Avg Path Length”:  
Produces a report of  
centrality scores for all  
nodes.

## Results:

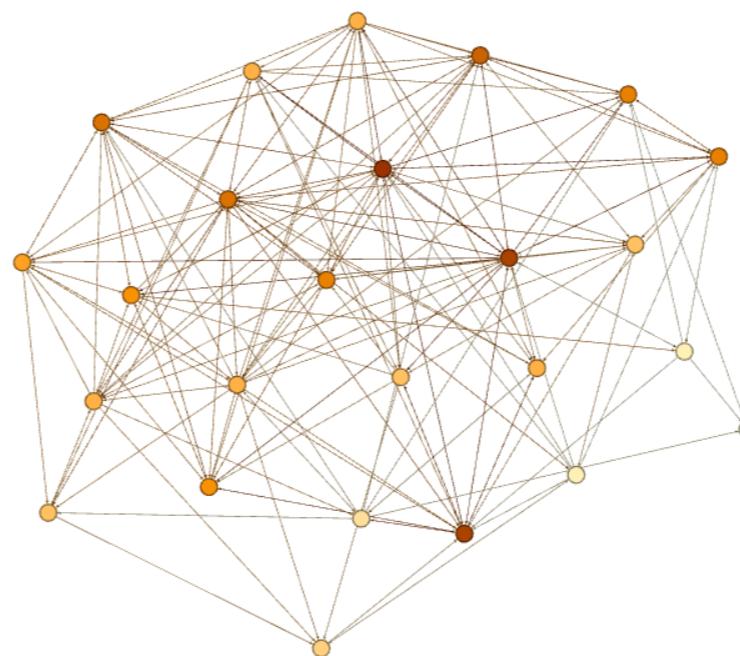
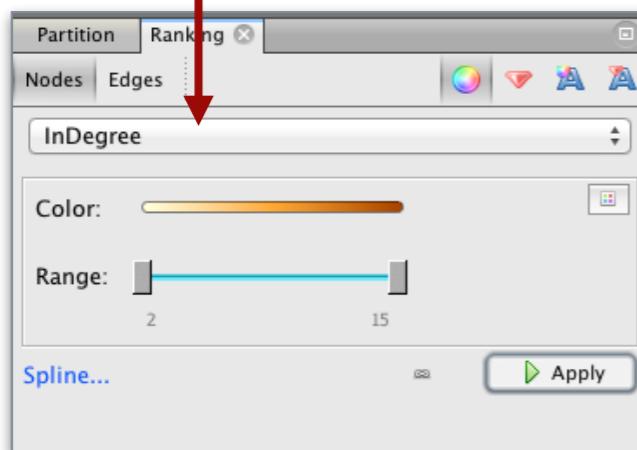
Diameter: 5  
Radius: 2  
Average Path length: 1.848333333333334  
Number of shortest paths: 600



# Ranking Nodes

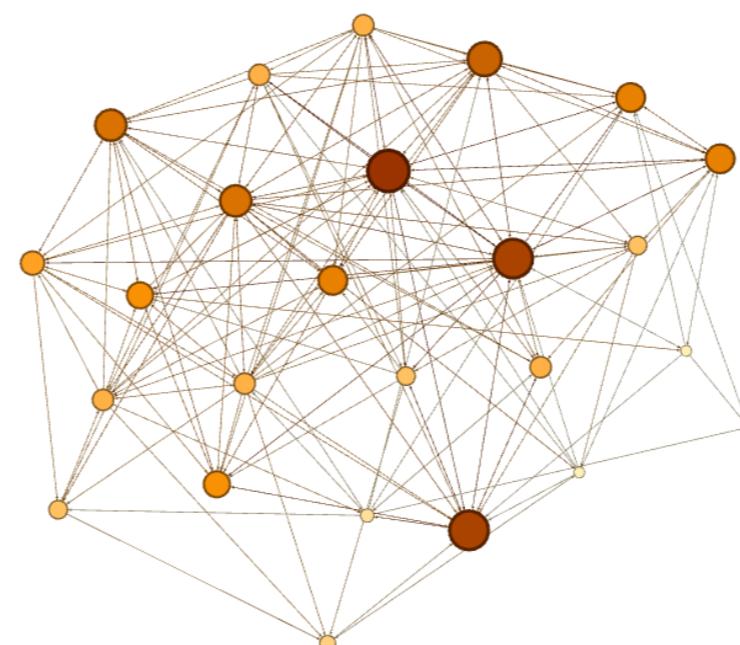
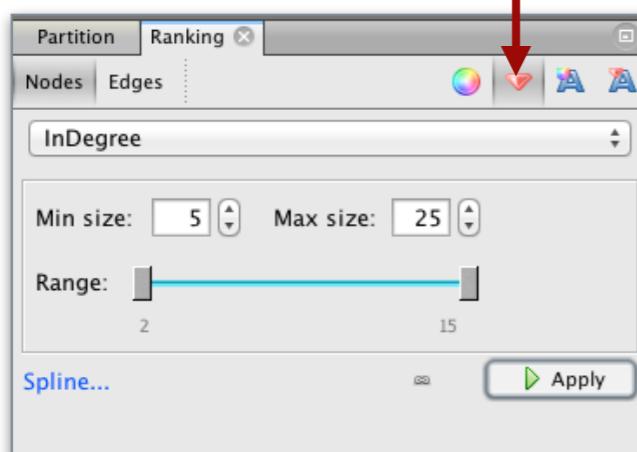
- Nodes can be re-sized and/or coloured based on their statistics, using the “Ranking” tab.

Choose a statistic to rank nodes



More saturated (red) colour indicates higher in-degree

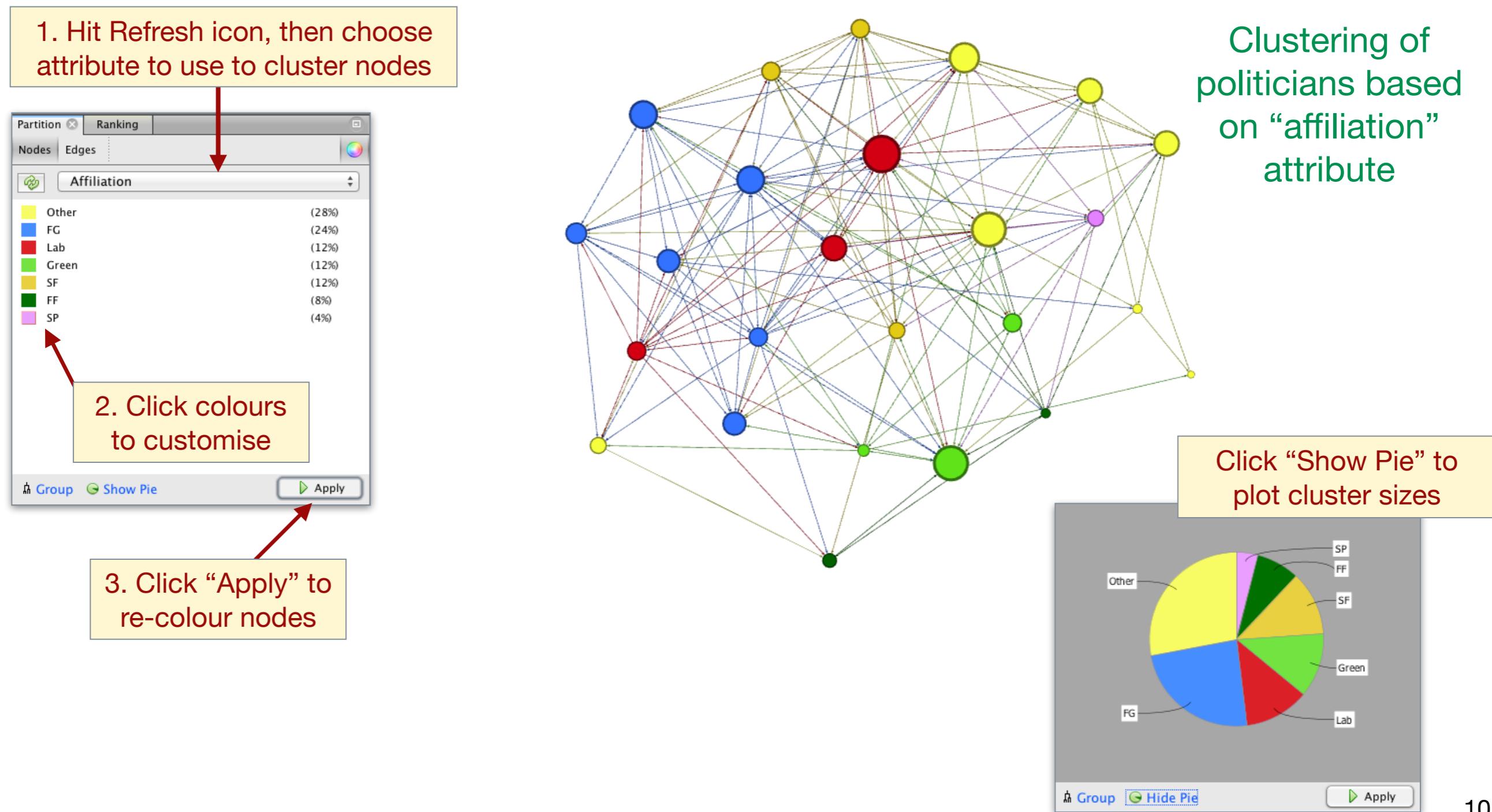
Switch to resize mode



Large node size indicates higher in-degree

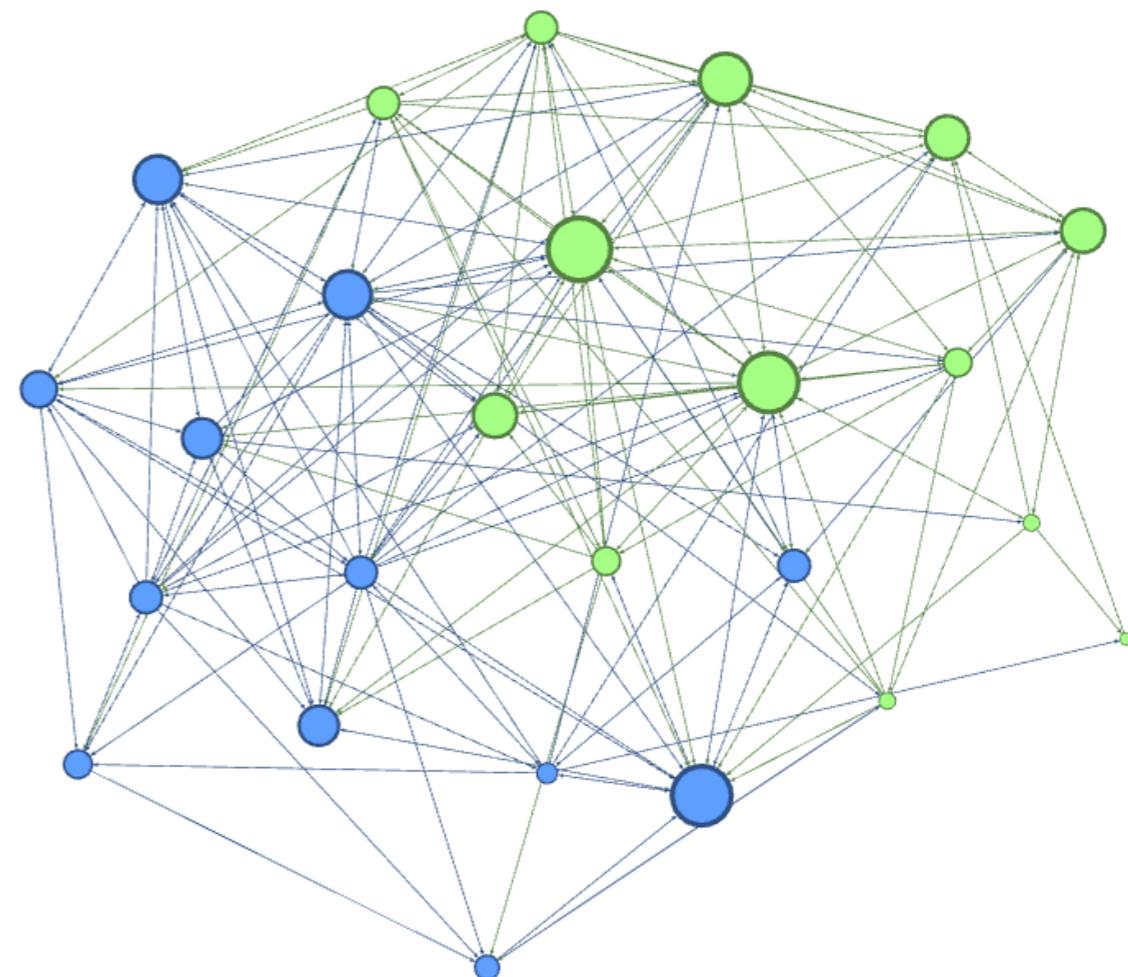
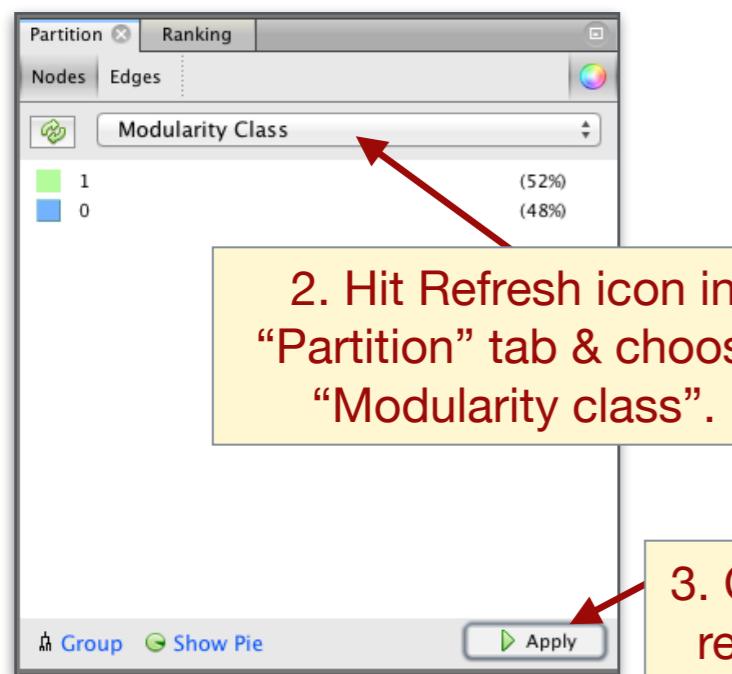
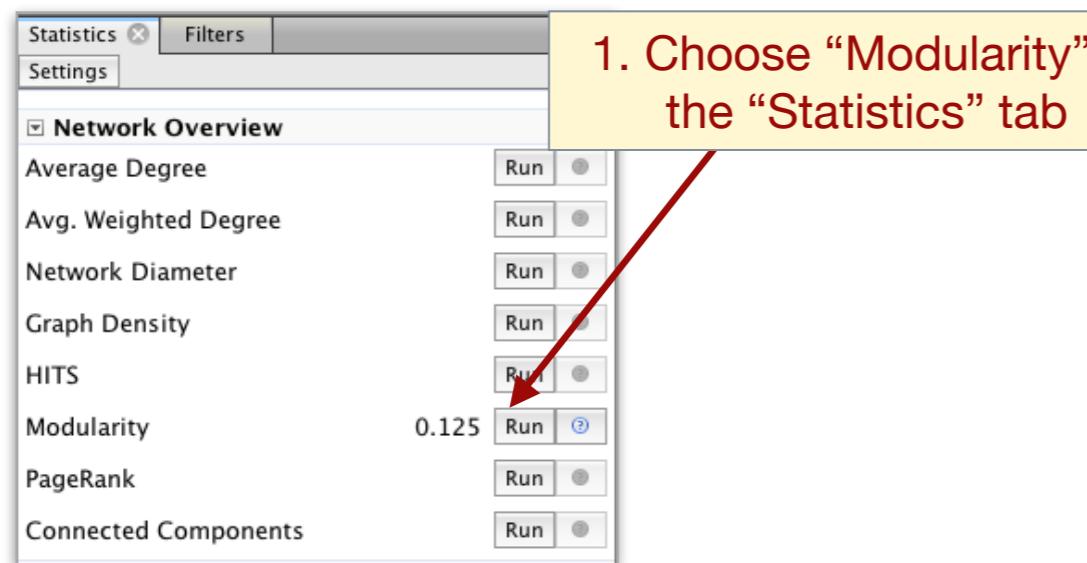
# Graph Clustering

- Nodes can be “clustered” and coloured based on their attribute values (e.g location, affiliation etc) in the “Partition” tab.



# Graph Clustering

- In cases where a grouping of nodes is not known apriori, we can apply cluster analysis methods to automatically detect groups in the data (e.g. communities of similar Twitter users)



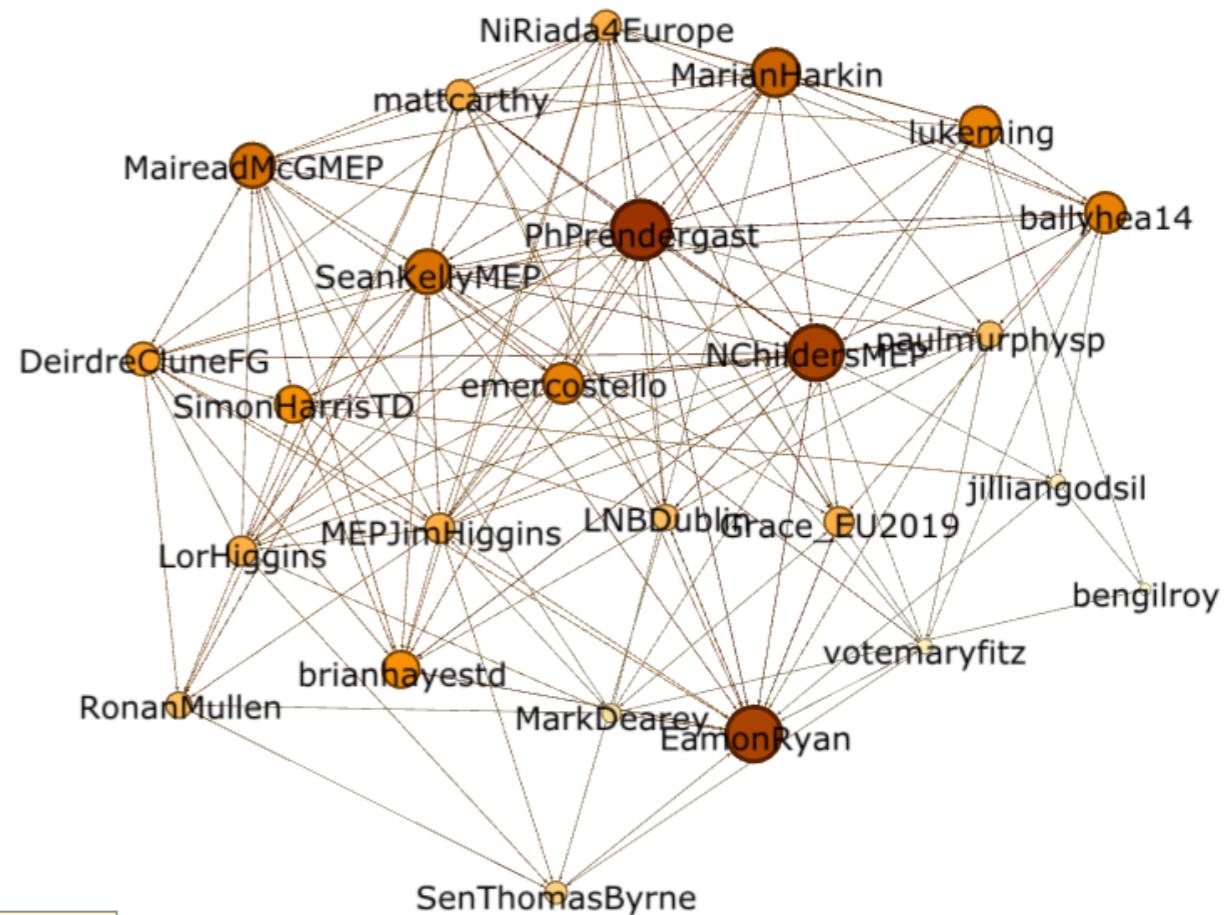
Modularity-based clustering produces 2 communities

# Filtering Nodes

- The “Filters” tab supports complex methods to temporarily highlight or hide subsets of nodes and edges in the graph.
- Nodes can be filtered by attribute value or based on node statistics.

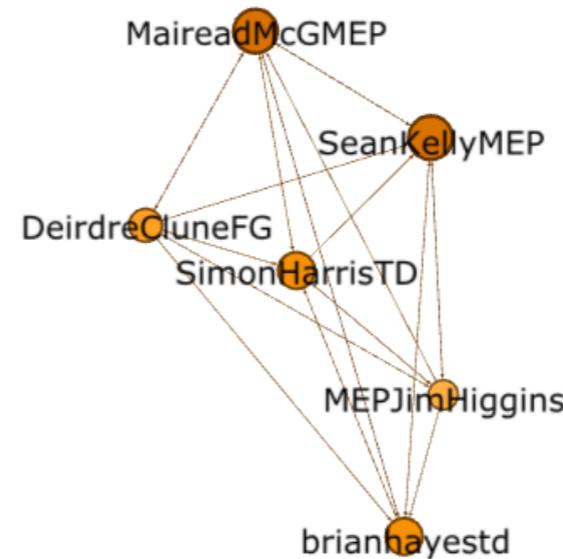
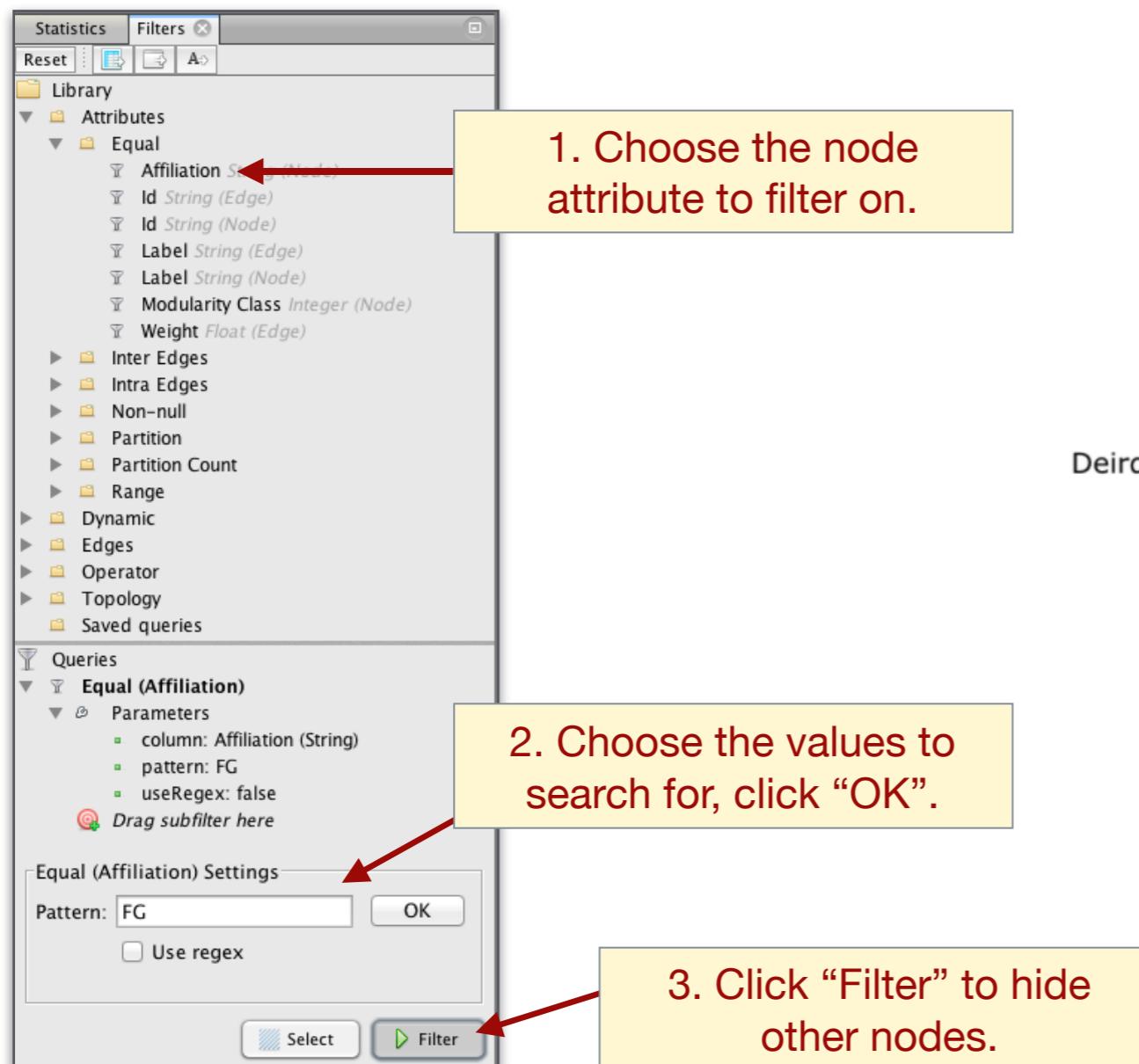
The screenshot shows the NetworkX software interface with the 'Filters' tab selected. On the left, there's a sidebar with 'Statistics' and 'Filters' tabs. Under 'Filters', there's a 'Library' section with various filter types like 'Equal', 'Inter Edges', etc., and a 'Queries' section with a specific 'Equal (Affiliation)' query. The 'Equal (Affiliation)' query has parameters set to 'Pattern: FG' and 'useRegex: false'. At the bottom, there are 'Select' and 'Filter' buttons. Three numbered callouts point to these elements:

1. Choose the node attribute to filter on. (points to the 'Affiliation String (Node)' item in the 'Equal' library)
2. Choose the values to search for, click "OK". (points to the 'Equal (Affiliation)' settings dialog)
3. Click "Filter" to hide other nodes. (points to the 'Filter' button at the bottom)

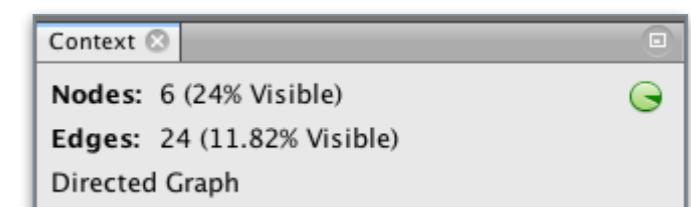


# Filtering Nodes

- The “Filters” tab supports complex methods to temporarily highlight or hide subsets of nodes and edges in the graph.
- Nodes can be filtered by attribute value or based on node statistics.



Filtered set of nodes based on affiliation = “FG”



# Gephi Data Laboratory Screen

- Laboratory: alternative tabular view of the same graph data.

## Graph Nodes

The screenshot shows the Gephi Data Laboratory interface. At the top, there are three tabs: Overview, Data Laboratory (which is selected), and Preview. Below the tabs is a toolbar with buttons for Add node, Add edge, Search/Replace, Import Spreadsheet, Export table, More actions, Filter, and a dropdown for Nodes. A yellow box labeled "Graph operations" covers the first few items in this toolbar. The main area is a data table titled "Data Table" with a "Nodes" tab selected. The table has columns for Nodes, Label, Id, Affiliation, Mod..., In-D..., Out..., Degree, Weig..., Weig..., Weig..., Ecce..., Closene..., and Betweenn... . The data consists of 30 rows of node information. At the bottom of the table are three sections: Nodes, Attributes, and Statistics, each with its own set of icons. A yellow box labeled "Column operations" covers the bottom toolbar, which includes buttons for Add column, Merge columns, Delete column, Clear column, Copy data to other column, Fill column with a value, Duplicate column, Create a boolean column from regex match, and Create column with list of regex matching groups.

| Nodes          | Label          | Id         | Affiliation | Mod... | In-D... | Out... | Degree | Weig... | Weig... | Weig... | Ecce... | Closene... | Between... |
|----------------|----------------|------------|-------------|--------|---------|--------|--------|---------|---------|---------|---------|------------|------------|
| NiRiada4Europe | NiRiada4Europe | 2239747586 | SF          | 1      | 7       | 13     | 20     | 20      | 7       | 13      | 3       | 1.5        | 6.561      |
| ballyhea14     | ballyhea14     | 259885449  | Other       | 1      | 10      | 7      | 17     | 17      | 10      | 7       | 3       | 1.792      | 56.359     |
| PhPrendergast  | PhPrendergast  | 46123918   | Lab         | 1      | 15      | 17     | 32     | 32      | 15      | 17      | 3       | 1.333      | 58.526     |
| MarianHarkin   | MarianHarkin   | 26228880   | Other       | 1      | 12      | 9      | 21     | 21      | 12      | 9       | 3       | 1.667      | 17.841     |
| mattcarthy     | mattcarthy     | 26586771   | SF          | 1      | 7       | 8      | 15     | 15      | 7       | 8       | 3       | 1.75       | 10.261     |
| MEPJimHiggins  | MEPJimHiggins  | 2233065878 | FG          | 0      | 7       | 17     | 24     | 24      | 7       | 17      | 2       | 1.292      | 23.331     |
| SeanKellyMEP   | SeanKellyMEP   | 21440665   | FG          | 0      | 11      | 17     | 28     | 28      | 11      | 17      | 3       | 1.333      | 45.199     |
| LNB Dublin     | LNB Dublin     | 263243802  | SF          | 1      | 6       | 9      | 15     | 15      | 6       | 9       | 3       | 1.667      | 11.879     |
| votemaryfitz   | votemaryfitz   | 213229979  | FF          | 1      | 3       | 6      | 9      | 9       | 3       | 6       | 3       | 1.792      | 12.972     |
| RonanMullen    | RonanMullen    | 6306972    | Other       | 0      | 6       | 2      | 8      | 8       | 6       | 2       | 3       | 2.333      | 3.584      |
| lukeming       | lukeming       | 54172831   | Other       | 1      | 10      | 1      | 11     | 11      | 10      | 1       | 4       | 2.75       | 23.928     |
| NChildersMEP   | NChildersMEP   | 239862568  | Other       | 1      | 14      | 17     | 31     | 31      | 14      | 17      | 2       | 1.292      | 70.499     |
| DeirdreCluneFG | DeirdreCluneFG | 627200944  | FG          | 0      | 8       | 8      | 16     | 16      | 8       | 8       | 3       | 1.708      | 5.371      |
| SenThomasByrne | SenThomasByrne | 32922034   | FF          | 0      | 5       | 2      | 7      | 7       | 5       | 2       | 3       | 2.5        | 5.968      |
| bengilroy      | bengilroy      | 42974146   | Other       | 1      | 2       | 1      | 3      | 3       | 2       | 1       | 5       | 3.625      | 0          |
| Grace_EU2019   | Grace_EU2019   | 2281217988 | Green       | 0      | 7       | 8      | 15     | 15      | 7       | 8       | 3       | 1.708      | 14.606     |
| MarkDearey     | MarkDearey     | 21499082   | Green       | 0      | 4       | 8      | 12     | 12      | 4       | 8       | 3       | 1.792      | 19.703     |
| jilliangodsil  | jilliangodsil  | 21287636   | Other       | 1      | 3       | 6      | 9      | 9       | 3       | 6       | 3       | 1.875      | 13.179     |
| emercostello   | emercostello   | 11587032   | Lab         | 1      | 10      | 5      | 15     | 15      | 10      | 5       | 3       | 1.875      | 2.871      |
| brianhayestd   | brianhayestd   | 228673882  | FG          | 0      | 9       | 3      | 12     | 12      | 9       | 3       | 3       | 2          | 3.11       |
| MaireadMcGMEP  | MaireadMcGMEP  | 166102184  | FG          | 0      | 11      | 7      | 18     | 18      | 11      | 7       | 3       | 1.833      | 7.958      |
| paulmurphyp    | paulmurphyp    | 270850017  | SP          | 1      | 6       | 8      | 14     | 14      | 6       | 8       | 3       | 1.75       | 10.301     |
| SimonHarrisTD  | SimonHarrisTD  | 21117425   | FG          | 0      | 9       | 7      | 16     | 16      | 9       | 7       | 2       | 1.708      | 18.892     |
| EamonRyan      | EamonRyan      | 22005625   | Green       | 0      | 14      | 5      | 19     | 19      | 14      | 5       | 3       | 1.833      | 33.088     |
| LorHiggins     | LorHiggins     | 39019643   | Lab         | 0      | 7       | 12     | 19     | 19      | 7       | 12      | 2       | 1.5        | 33.014     |

**Graph operations**

**Column operations**

# Gephi Data Laboratory Screen

- Laboratory: alternative tabular view of the same graph data.

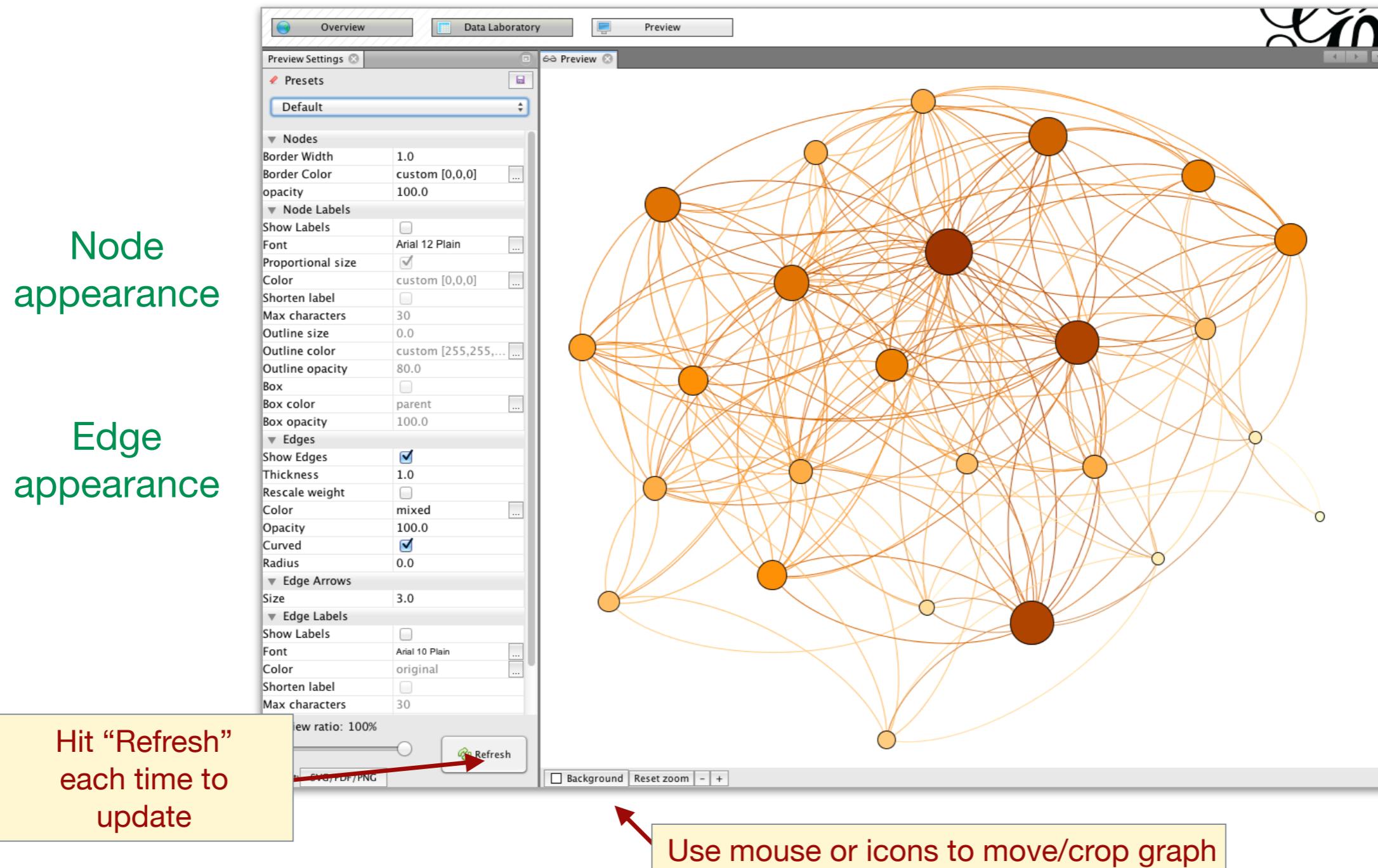
Graph  
Edges

The screenshot shows the Gephi Data Laboratory interface. At the top, there are three tabs: Overview, Data Laboratory (which is selected), and Preview. Below the tabs is a toolbar with various icons for managing data. A yellow box labeled "Edge Pairs (Source, Target)" has a red arrow pointing to the "Source" and "Target" columns in the table. Another yellow box labeled "Edge Weights" has a red arrow pointing to the "Weight" column. The main area is a data table with the following columns: Source, Target, Type, Id, Weight, and Label. The table contains numerous rows of edge data. At the bottom of the interface is a toolbar with icons for adding columns, merging columns, deleting columns, clearing columns, copying data, filling columns, duplicating columns, creating boolean columns from regex matches, and creating columns with lists of regex matching groups.

| Source     | Target     | Type     | Id | Weight | Label |
|------------|------------|----------|----|--------|-------|
| 2239747586 | 259885449  | Directed | 3  | 1      |       |
| 2239747586 | 46123918   | Directed | 4  | 1      |       |
| 2239747586 | 26228880   | Directed | 5  | 1      |       |
| 2239747586 | 26586771   | Directed | 7  | 1      |       |
| 2239747586 | 21440665   | Directed | 9  | 1      |       |
| 2239747586 | 263243802  | Directed | 10 | 1      |       |
| 2239747586 | 54172831   | Directed | 12 | 1      |       |
| 2239747586 | 239862568  | Directed | 2  | 1      |       |
| 2239747586 | 627200944  | Directed | 0  | 1      |       |
| 2239747586 | 2281217988 | Directed | 1  | 1      |       |
| 2239747586 | 11587032   | Directed | 8  | 1      |       |
| 2239747586 | 228673882  | Directed | 11 | 1      |       |
| 2239747586 | 166102184  | Directed | 6  | 1      |       |
| 259885449  | 46123918   | Directed | 16 | 1      |       |
| 259885449  | 26228880   | Directed | 17 | 1      |       |
| 259885449  | 54172831   | Directed | 19 | 1      |       |
| 259885449  | 239862568  | Directed | 15 | 1      |       |
| 259885449  | 2281217988 | Directed | 14 | 1      |       |
| 259885449  | 21287636   | Directed | 18 | 1      |       |
| 259885449  | 270850017  | Directed | 13 | 1      |       |
| 46123918   | 2239747586 | Directed | 21 | 1      |       |
| 46123918   | 259885449  | Directed | 24 | 1      |       |
| 46123918   | 26228880   | Directed | 25 | 1      |       |
| 46123918   | 26586771   | Directed | 29 | 1      |       |
| 46123918   | 2233065878 | Directed | 30 | 1      |       |
| 46123918   | 21440665   | Directed | 33 | 1      |       |
| 46123918   | 263243802  | Directed | 34 | 1      |       |
| 46123918   | 54172831   | Directed | 36 | 1      |       |
| 46123918   | 239862568  | Directed | 26 | 1      |       |
| 46123918   | 627200944  | Directed | 27 | 1      |       |
| 46123918   | 2281217988 | Directed | 22 | 1      |       |
| 46123918   | 11587032   | Directed | 32 | 1      |       |
| 46123918   | 166102184  | Directed | 23 | 1      |       |
| 46123918   | 270850017  | Directed | 20 | 1      |       |

# Gephi Preview Screen

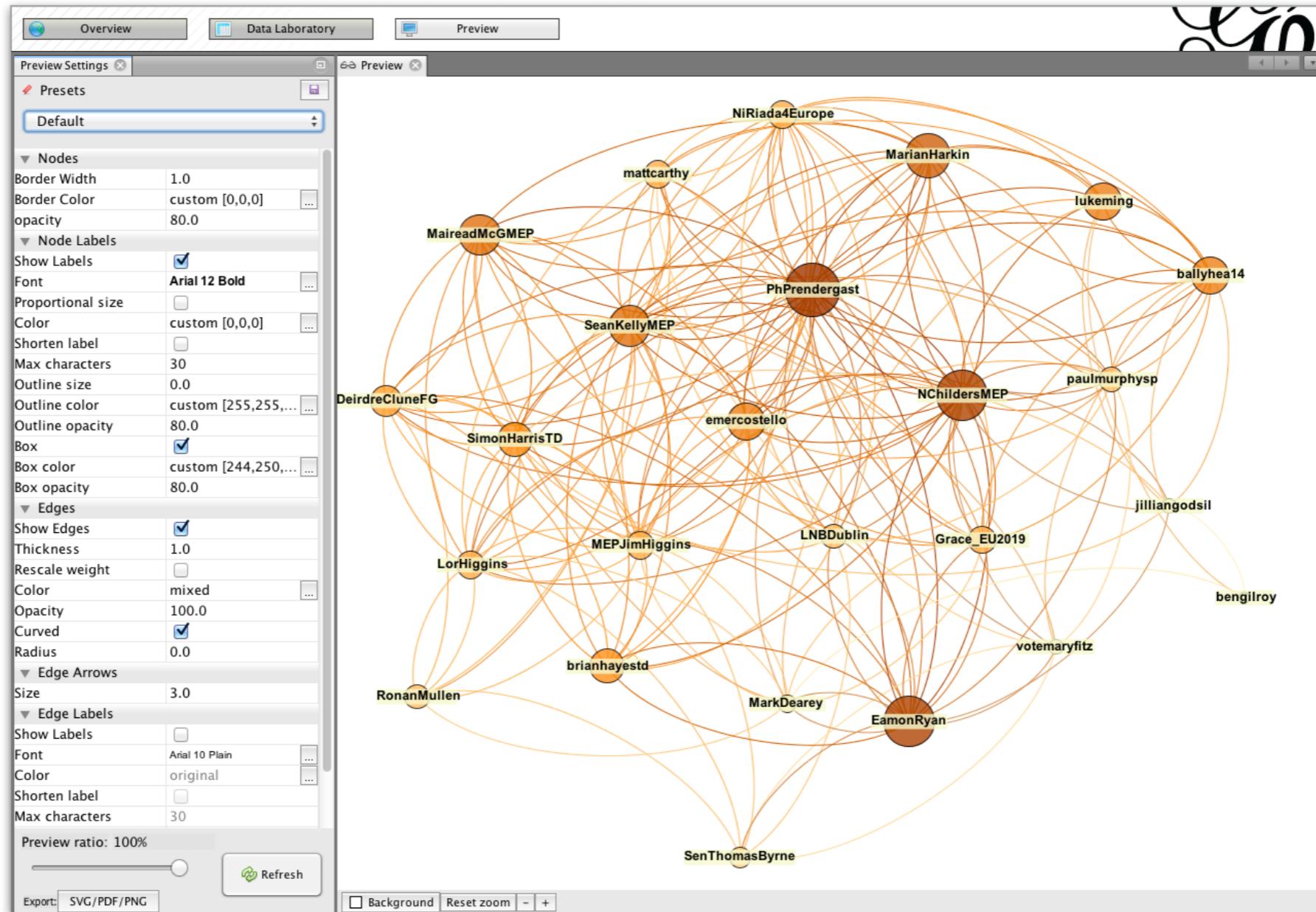
- **Preview:** Tweak the appearance of your graph before exporting a publication quality image.



# Gephi Preview Screen

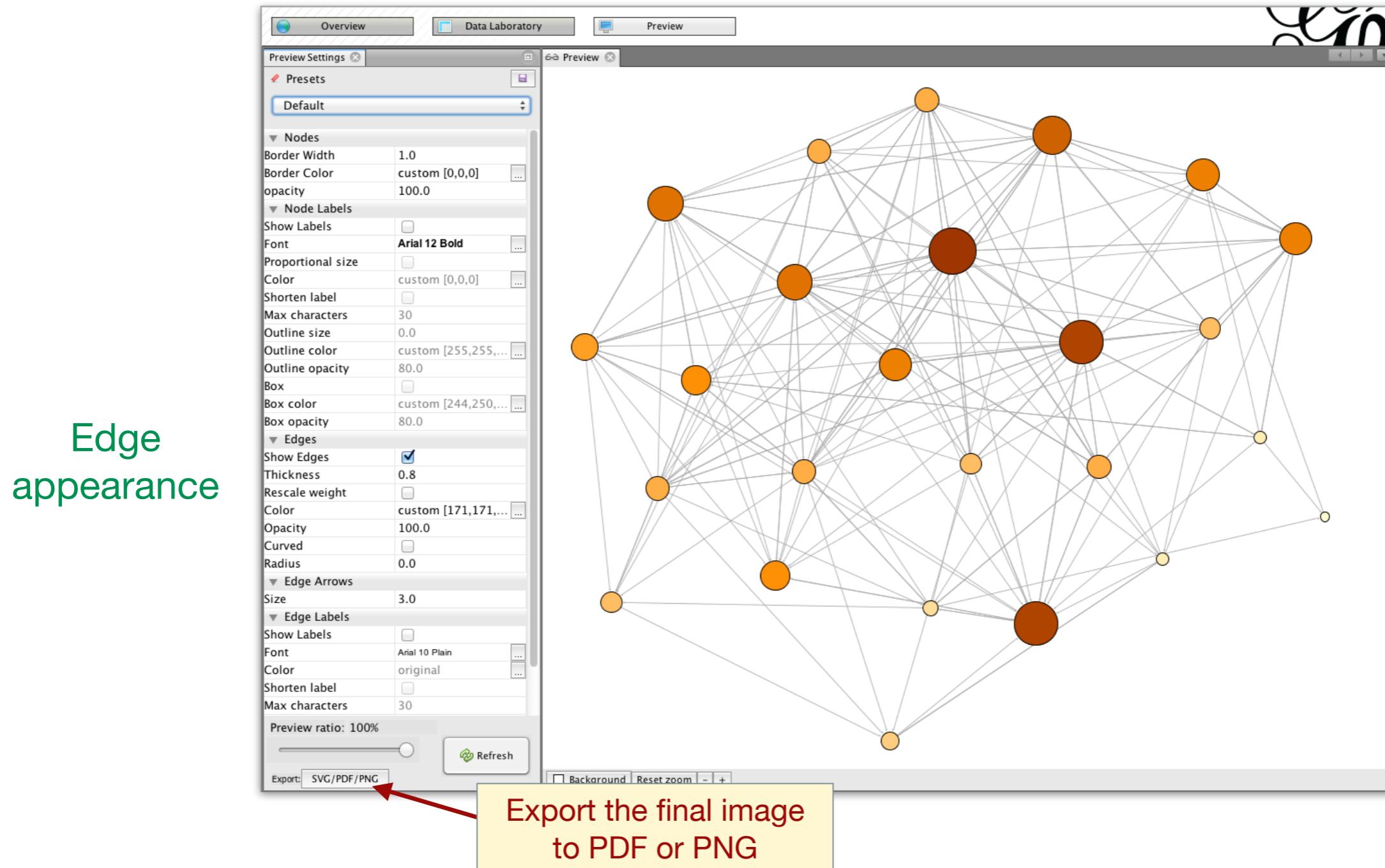
- **Preview:** Tweak the appearance of your graph before exporting a publication quality image.

Node  
appearance

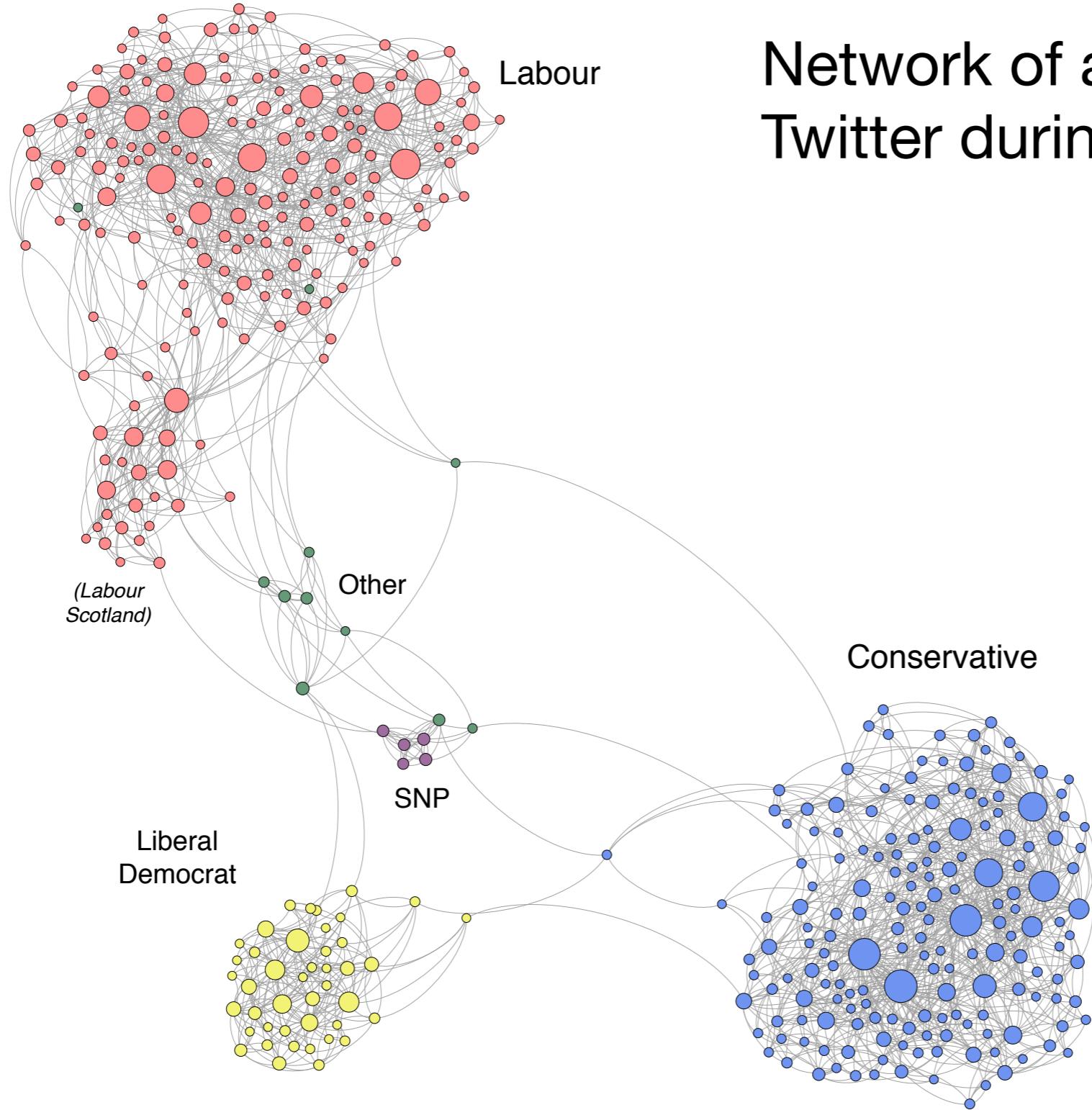


# Gephi Preview Screen

- **Preview:** Tweak the appearance of your graph before exporting a publication quality image.

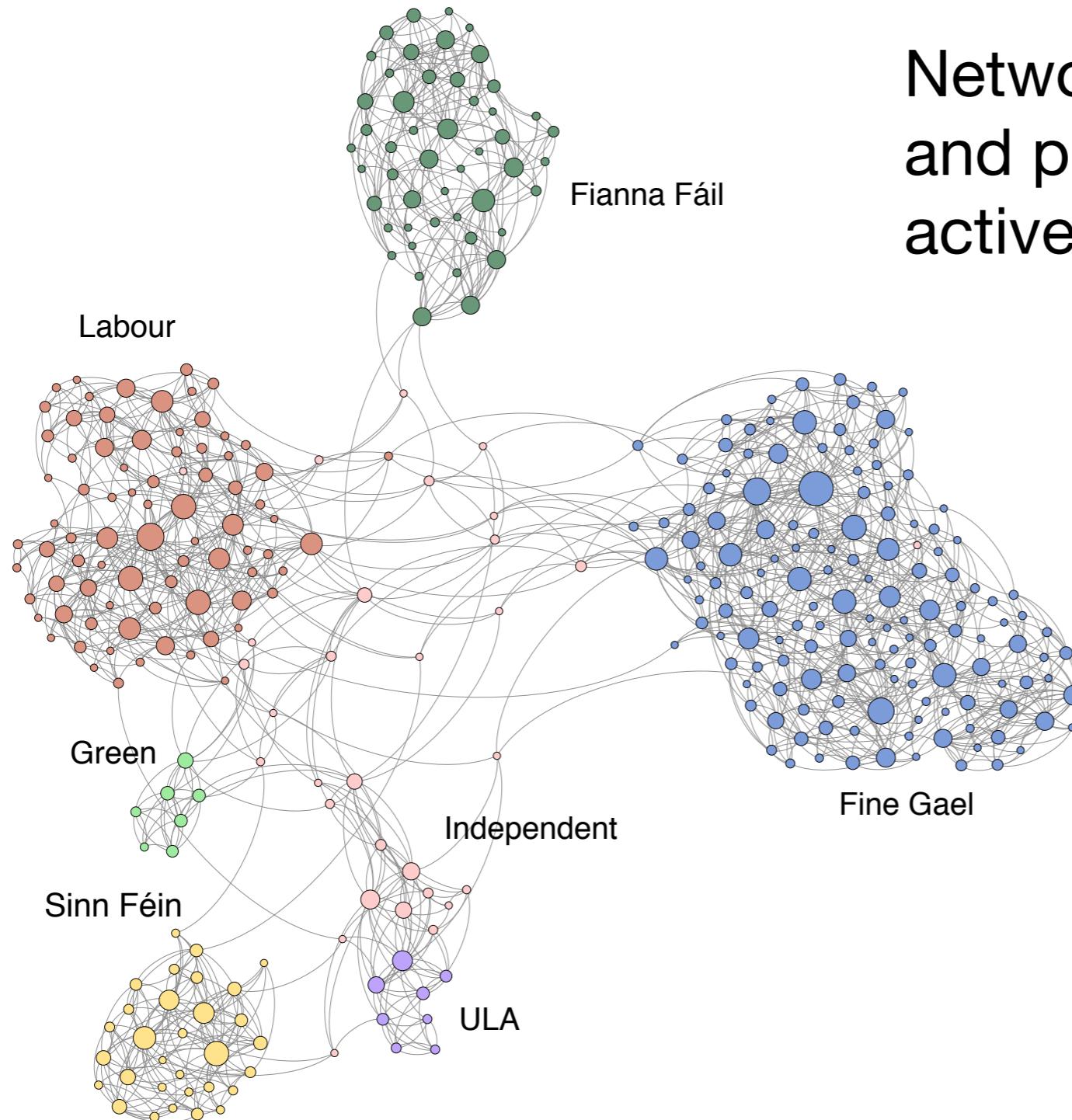


# Example: UK Politics



Network of all UK MPs active on Twitter during 2012.

# Example: Irish Politics



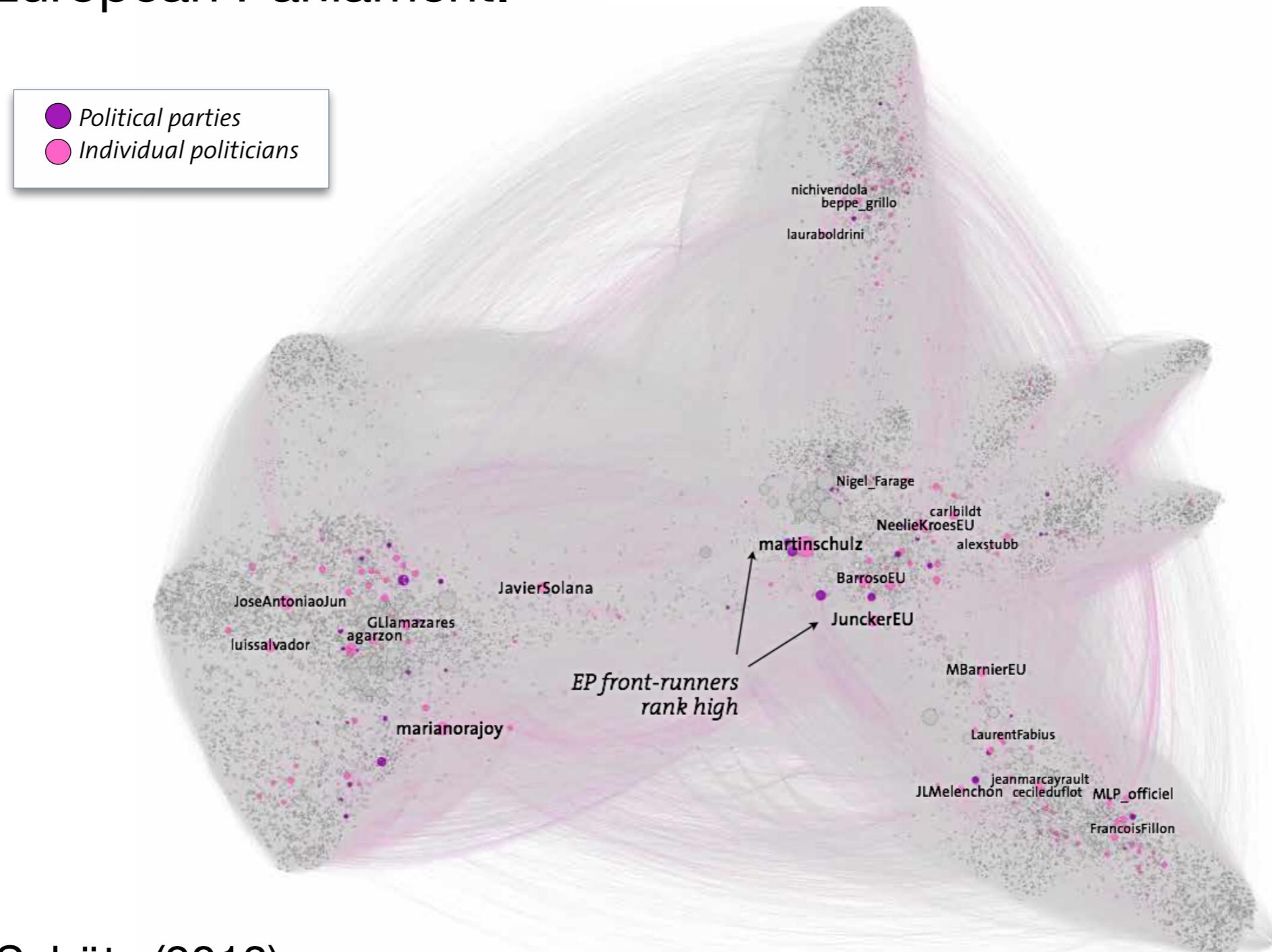
Network of 348 Irish politicians and political organisations active on Twitter in 2011-2012.

<http://mlg.ucd.ie/networks/politics-ie.html>

Greene &  
Cunningham (2013)

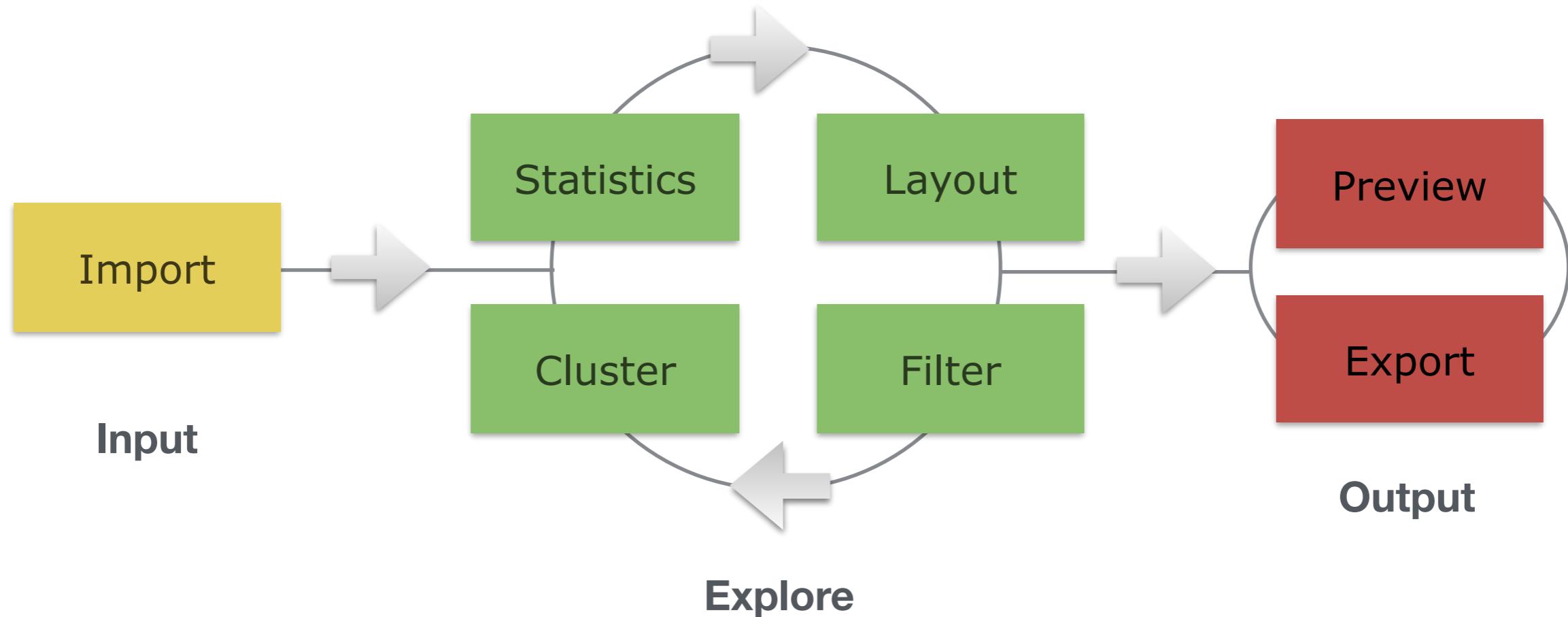
# Example: European Political Twittersphere

Ranking of Twitter accounts for prominent members and groups of European Parliament.



# Conclusion

- Gephi provides a powerful workflow for exploring and visualising graphs...



[Gephi: an open source software for exploring and manipulating networks.](#)  
M Bastian, S Heymann, M Jacomy - ICWSM, 2009 - aaai.org  
Abstract Gephi is an open source software for graph and network analysis. It uses a 3D render engine to display large networks in real-time and to speed up the exploration. A flexible and multi-task architecture brings new possibilities to work with complex data sets ...  
Cited by 708 Related articles All 15 versions Import into BibTeX Saved More

<http://gephi.org>

# References

- M. Bastian, S. Heymann & M. Jacomy (2009). Gephi: an open source software for exploring and manipulating networks. Proc. ICWSM-09.
- A. Maireder & F. Schütz (2014). The European Political Twittersphere: Network of top users discussing the 2014 European Elections. GFK white paper.
- Greene, D. & Cunningham, P. (2013). Producing a Unified Graph Representation from Multiple Social Network Views. Proc. ACM Web Science'13.