

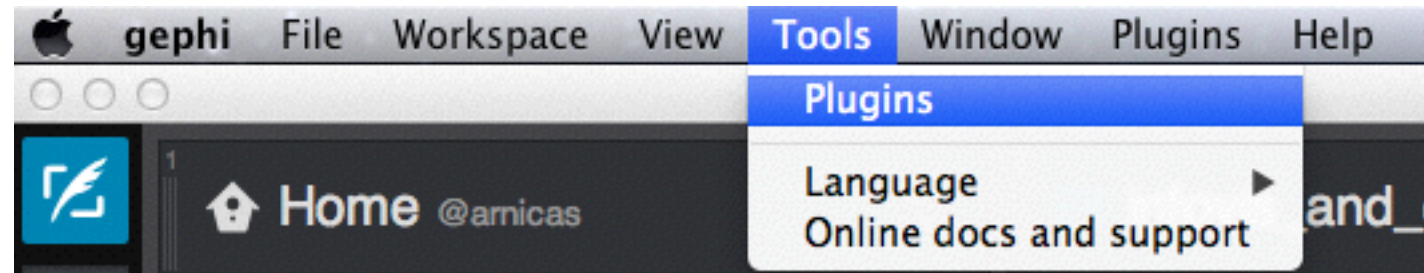
Some Layout and Export Tricks in Gephi (Mixed Chapters Version)

Lynn Cherny (@arnicas)

If you're following along...

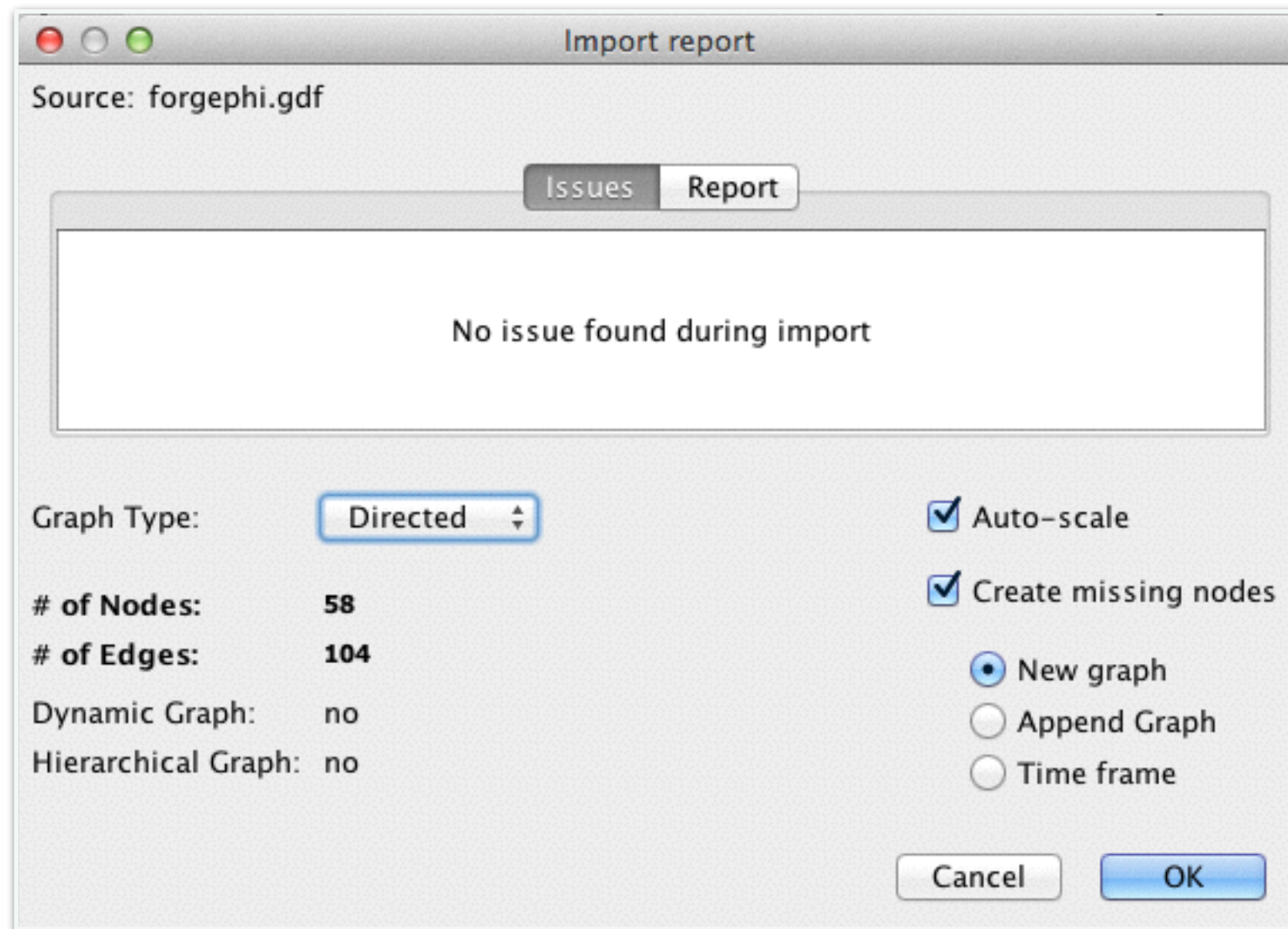
- All the files are in <https://github.com/arnicas/TopicsPythonGephi>
- You should have a gephi appropriate files now, found in files/forgephi.gdf: <https://github.com/arnicas/TopicsPythonGephi/blob/master/files/forgephi.gdf>

Plugins to Add to Gephi



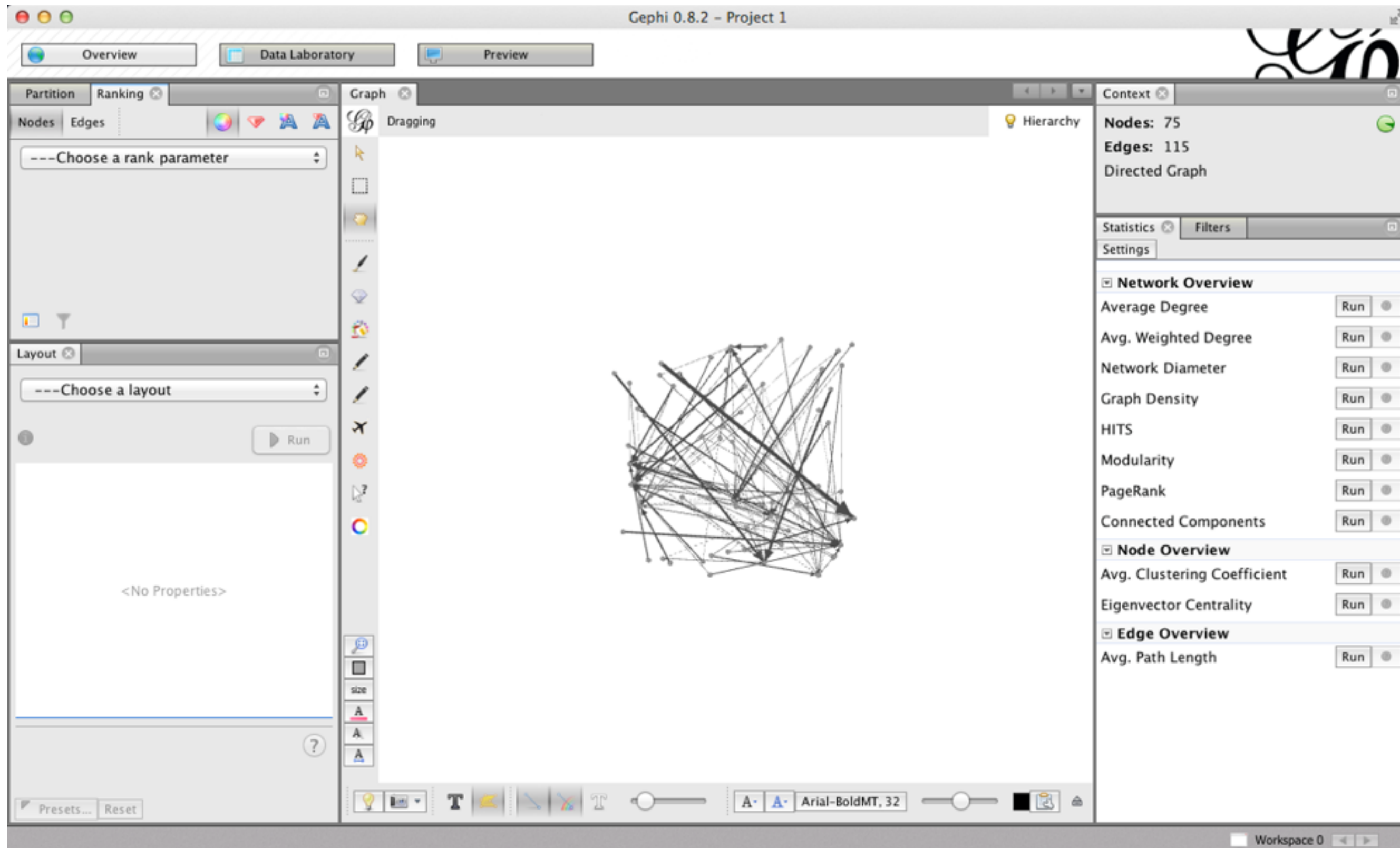
- Circular Layout
- SigmaExporter

Launch Gephi. “Open” on file menu... Use defaults you see here.

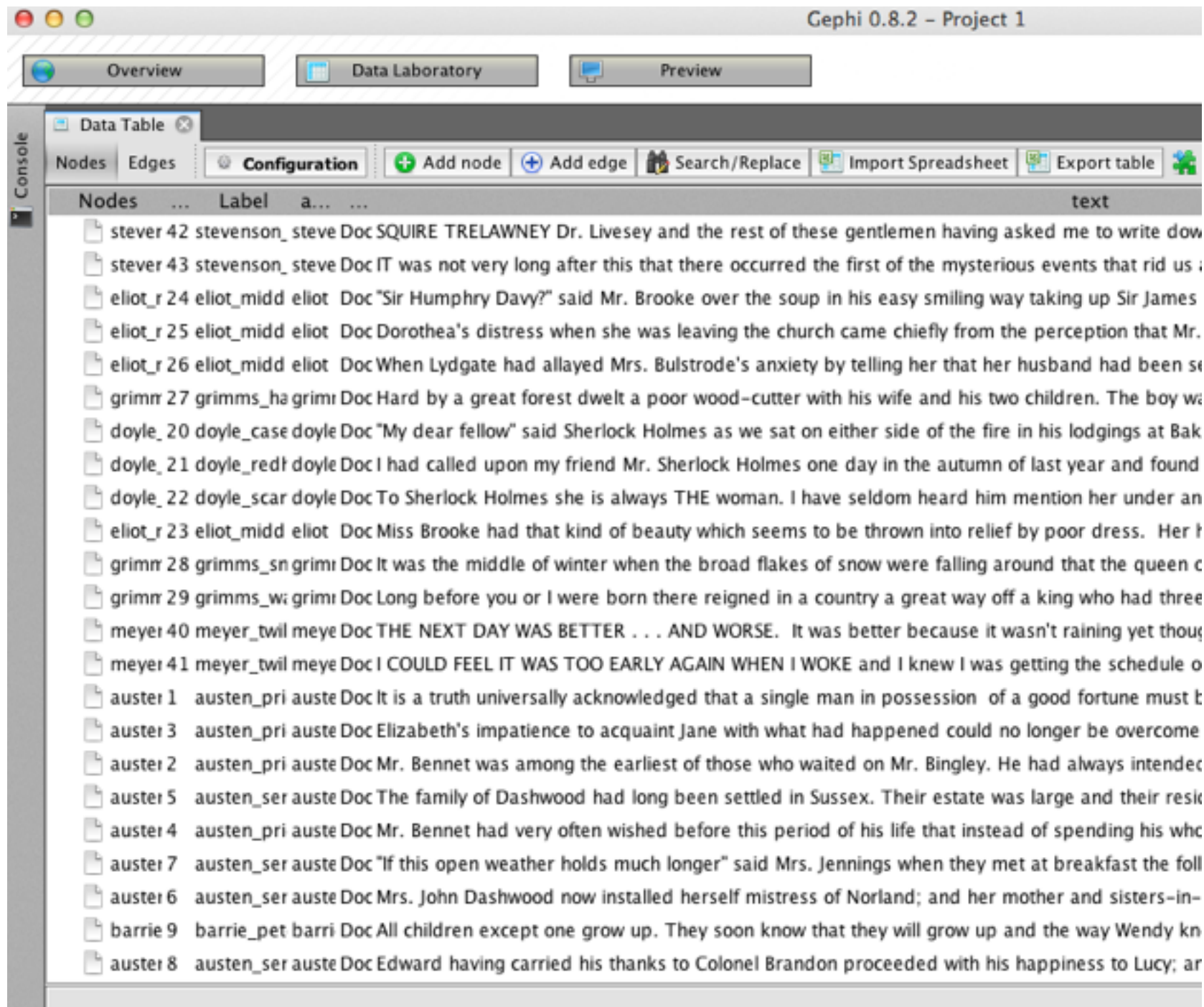


If your number of nodes does NOT = # docs + # topics, you have a problem! (Check your nodes files as csv.)

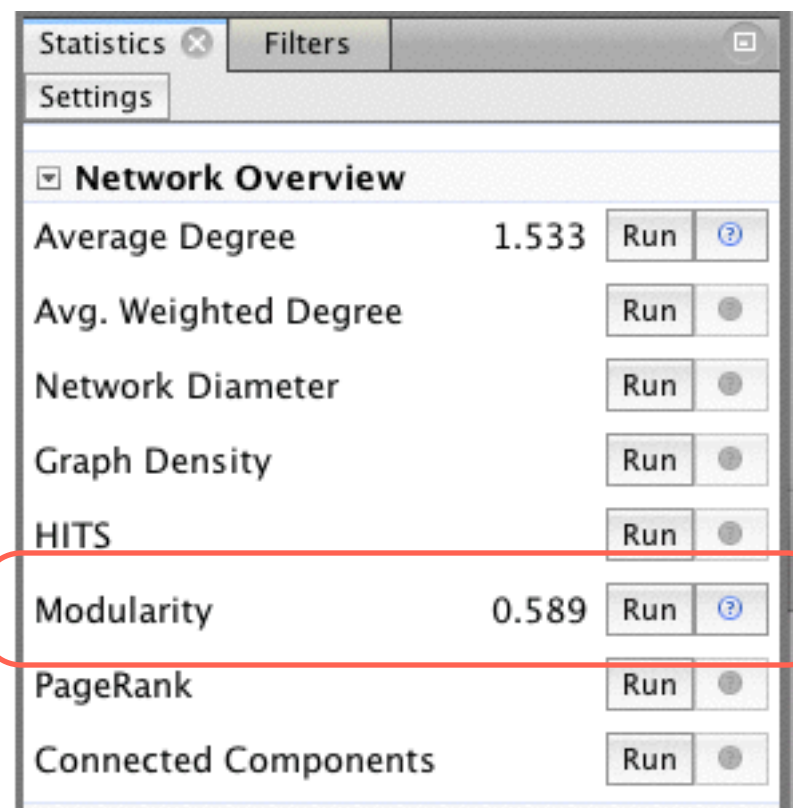
If all goes well, you should see something like this....



On the Data Laboratory Tab, you'll see what we imported. This is another method for debugging your file...

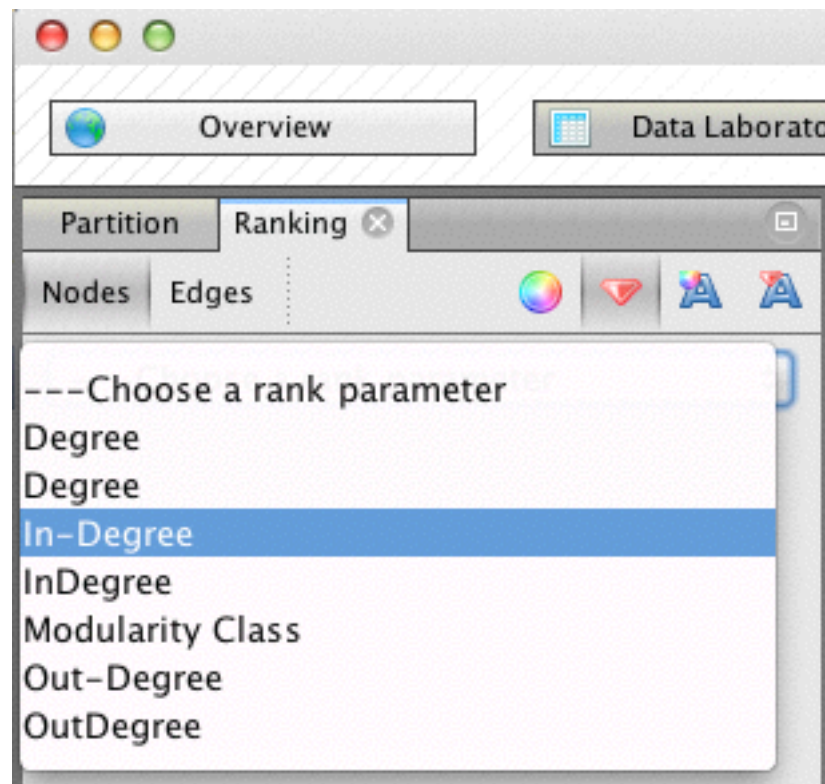


Over on the right side, you should have a “statistics” panel. “Run” a few of them and dismiss the dialogs. This adds stats to your data file you can use in layout and design. Run Modularity at least -- it’s a “community finding” algorithm.

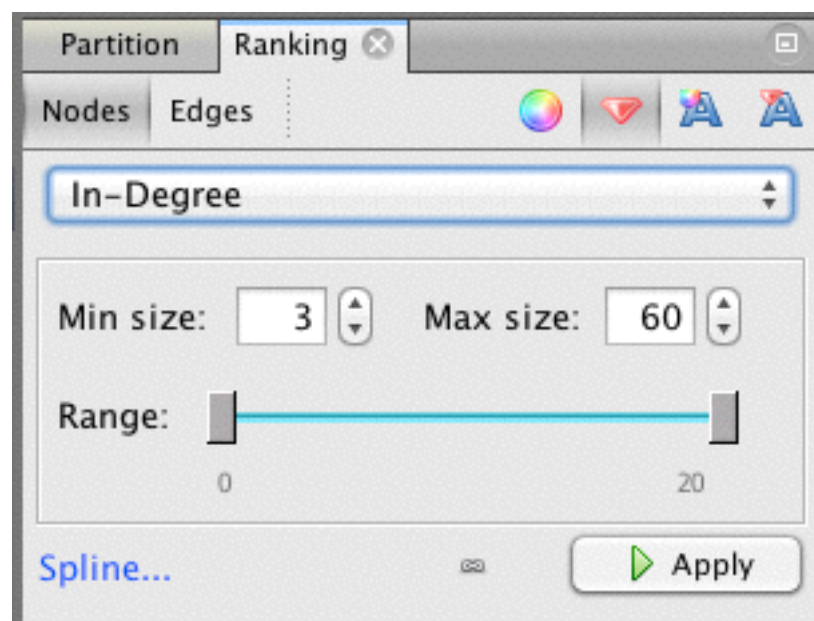


Now, on upper left side, in Ranking, select nodes, and the diamond. Pick “In-Degree” to size by.

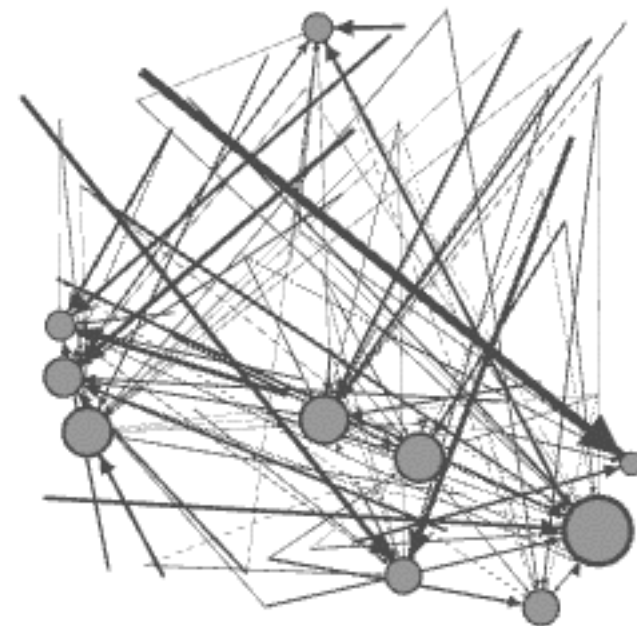
1.



2.



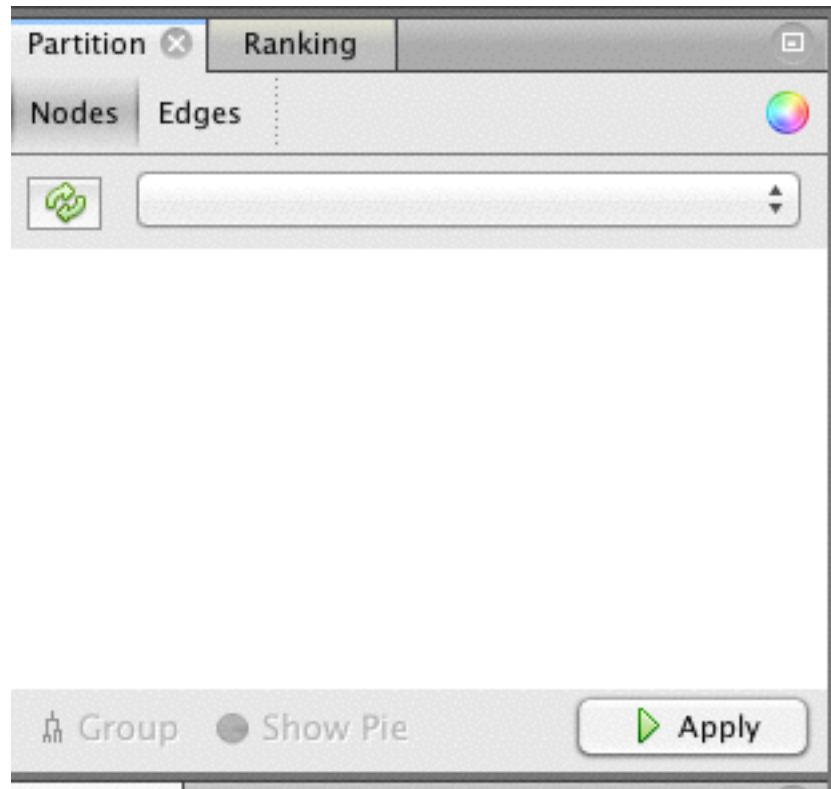
3. after
Apply:



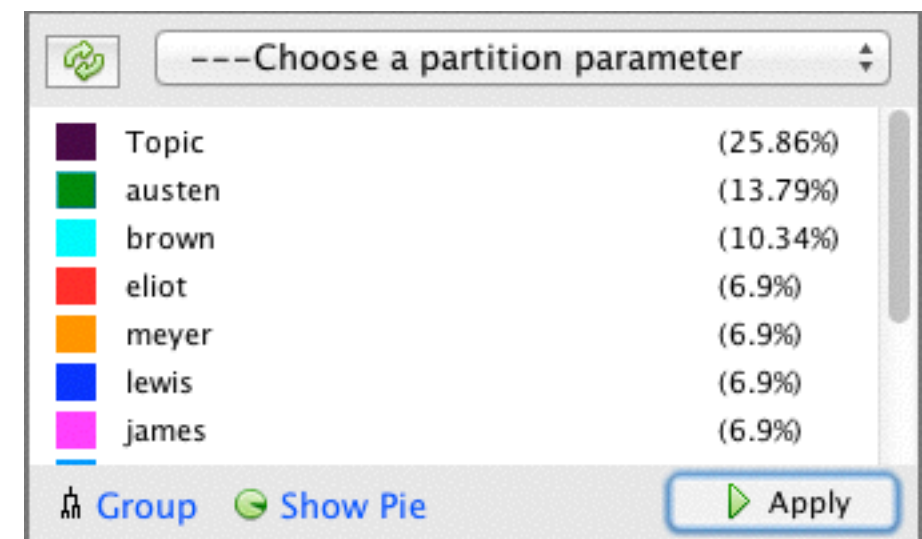
These are the nodes with the most “in arrows,” or highest in-degree. They are the “topics.”

Color by “Author” First

1. Hit green refresh arrow:



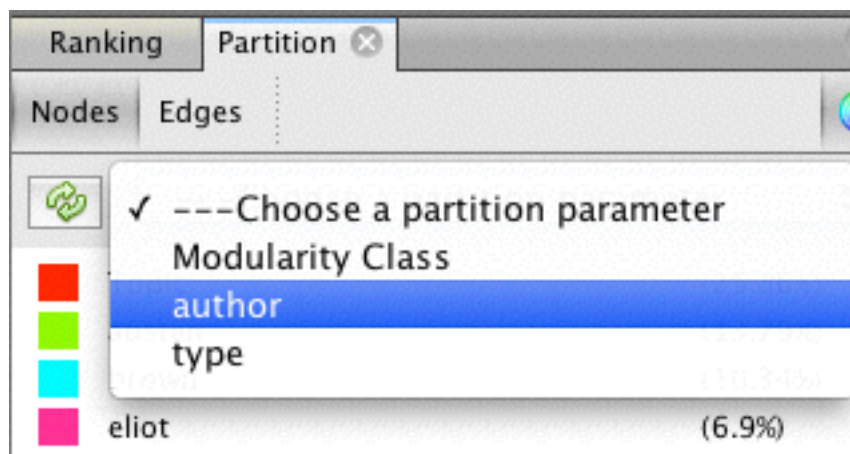
3. You'll see the “groups” and random assigned colors. Click on a color to pick a new one. Make them as different as you can.



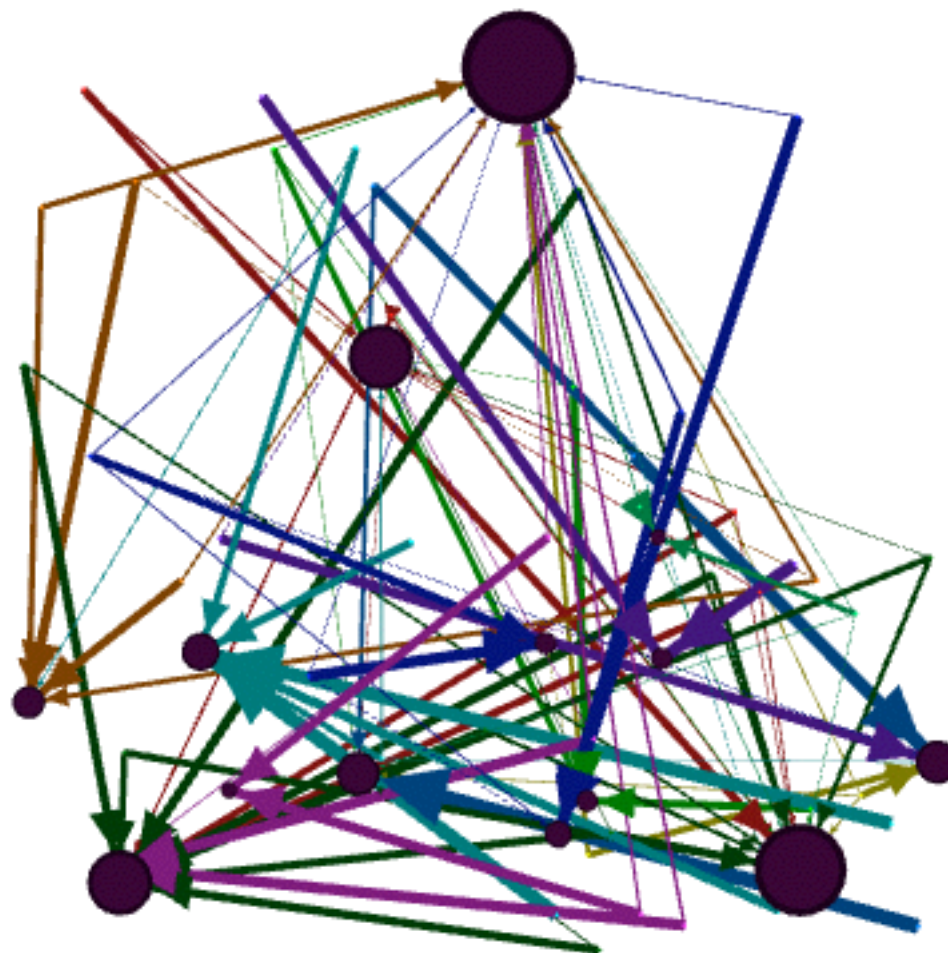
Tip: If you hate all of them, try a right click on one - pick randomize.

4. Then hit “Apply.”

2. Then choose “author” from menu:

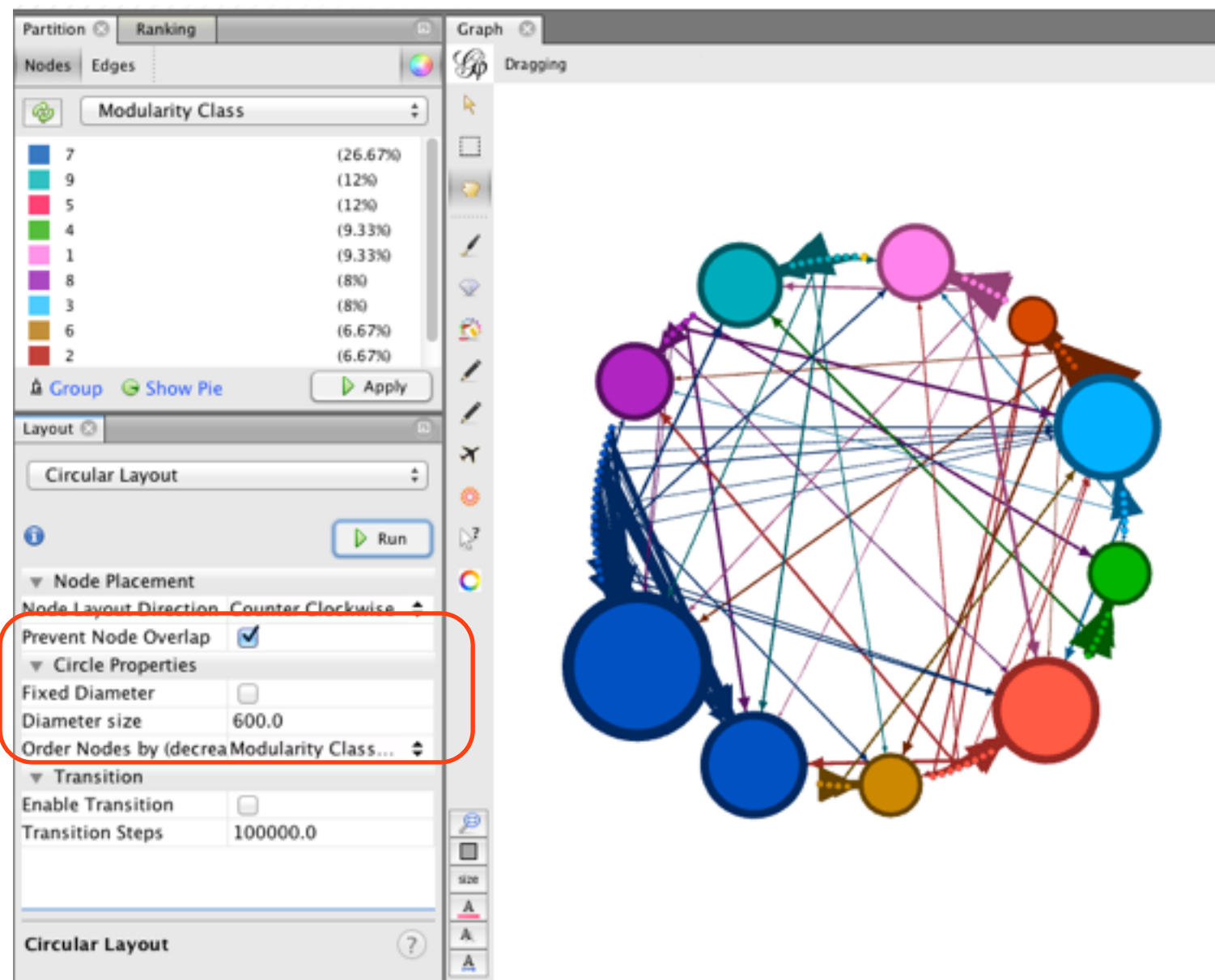


What are the big nodes?



Let's try to lay it out a bit...
Circular is nice for topics.
You'll need to hand-adjust in
any case.

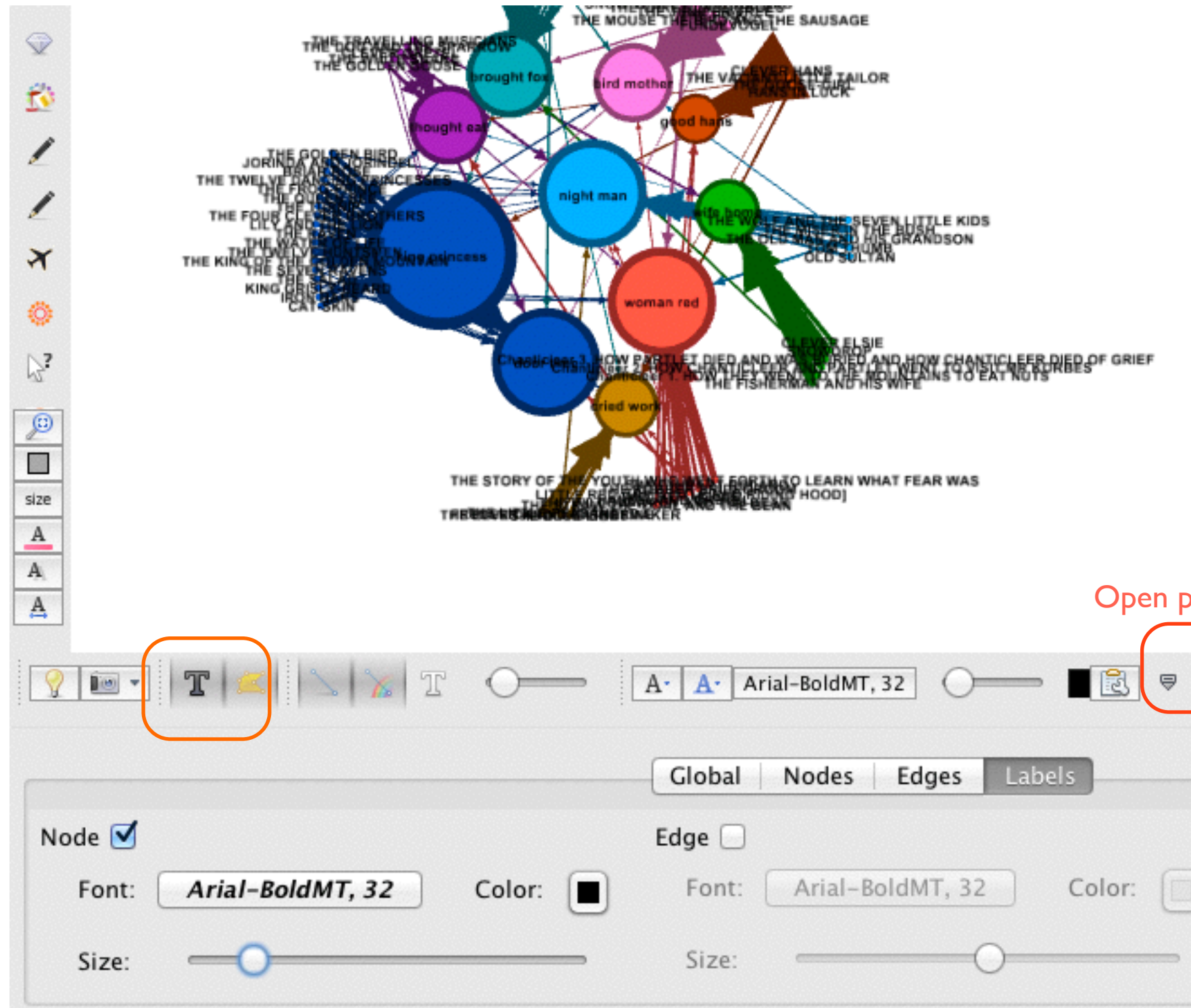
Order nodes by modularity
class!



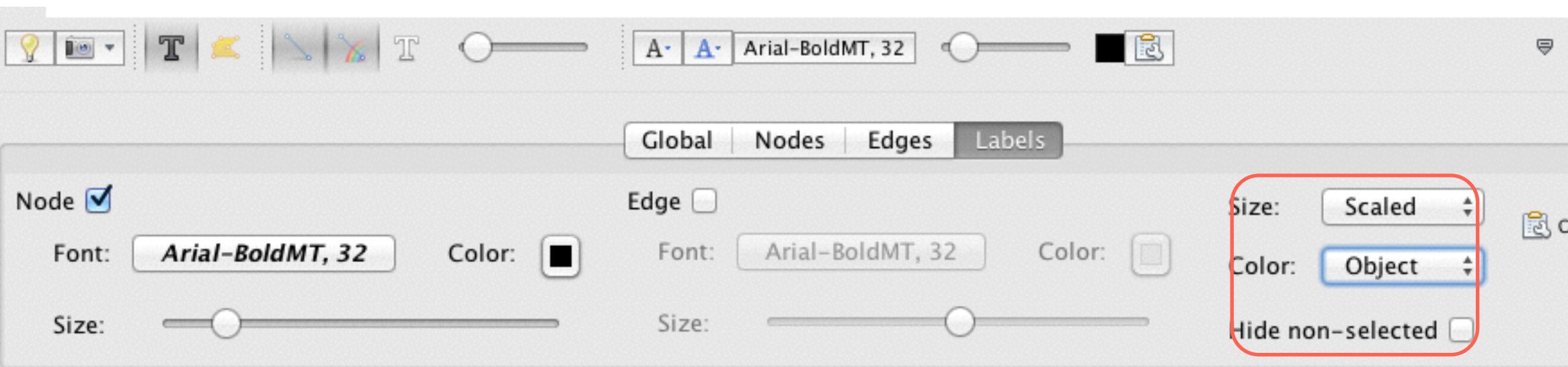
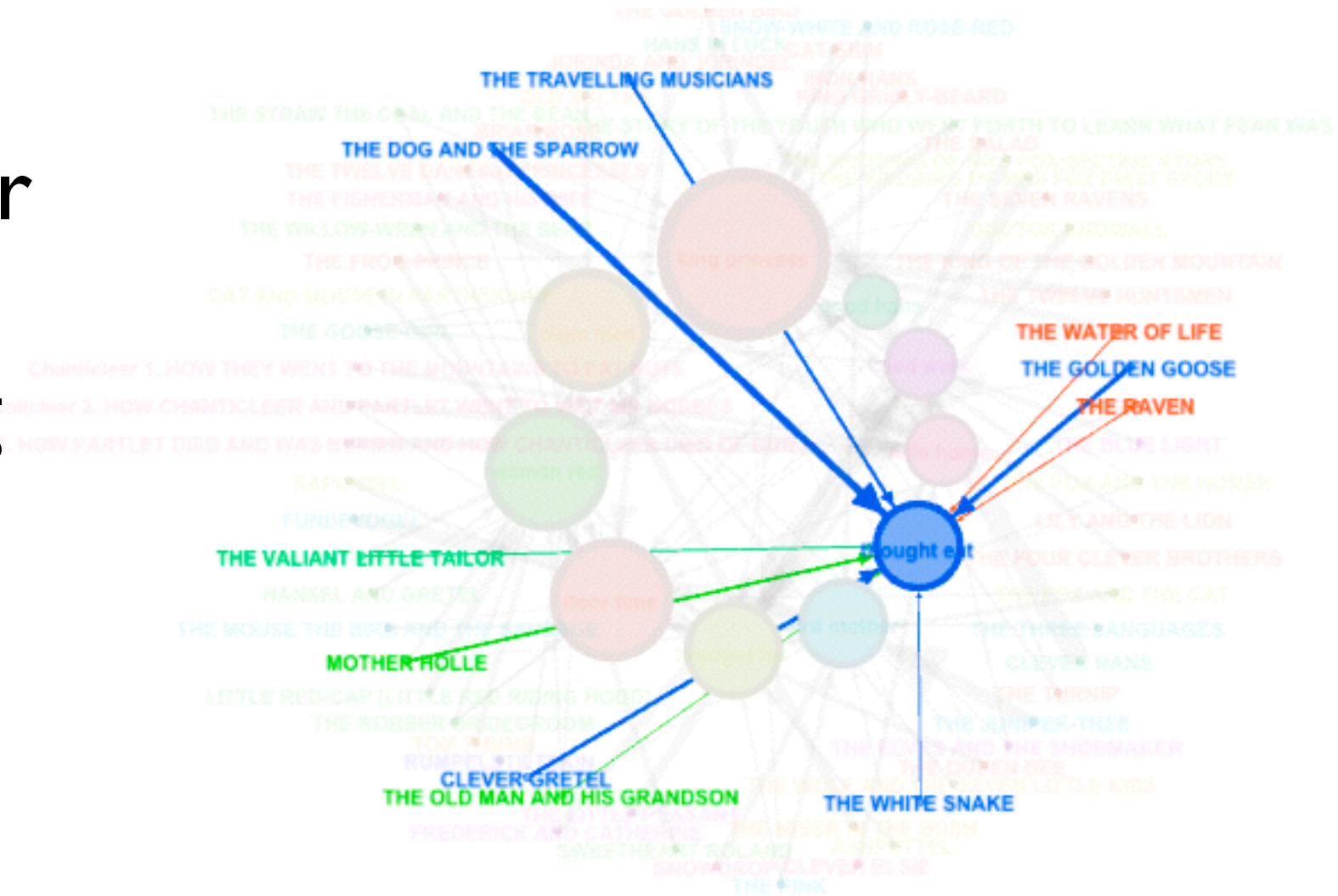
Hand-tweak to
move the topic
nodes inside.



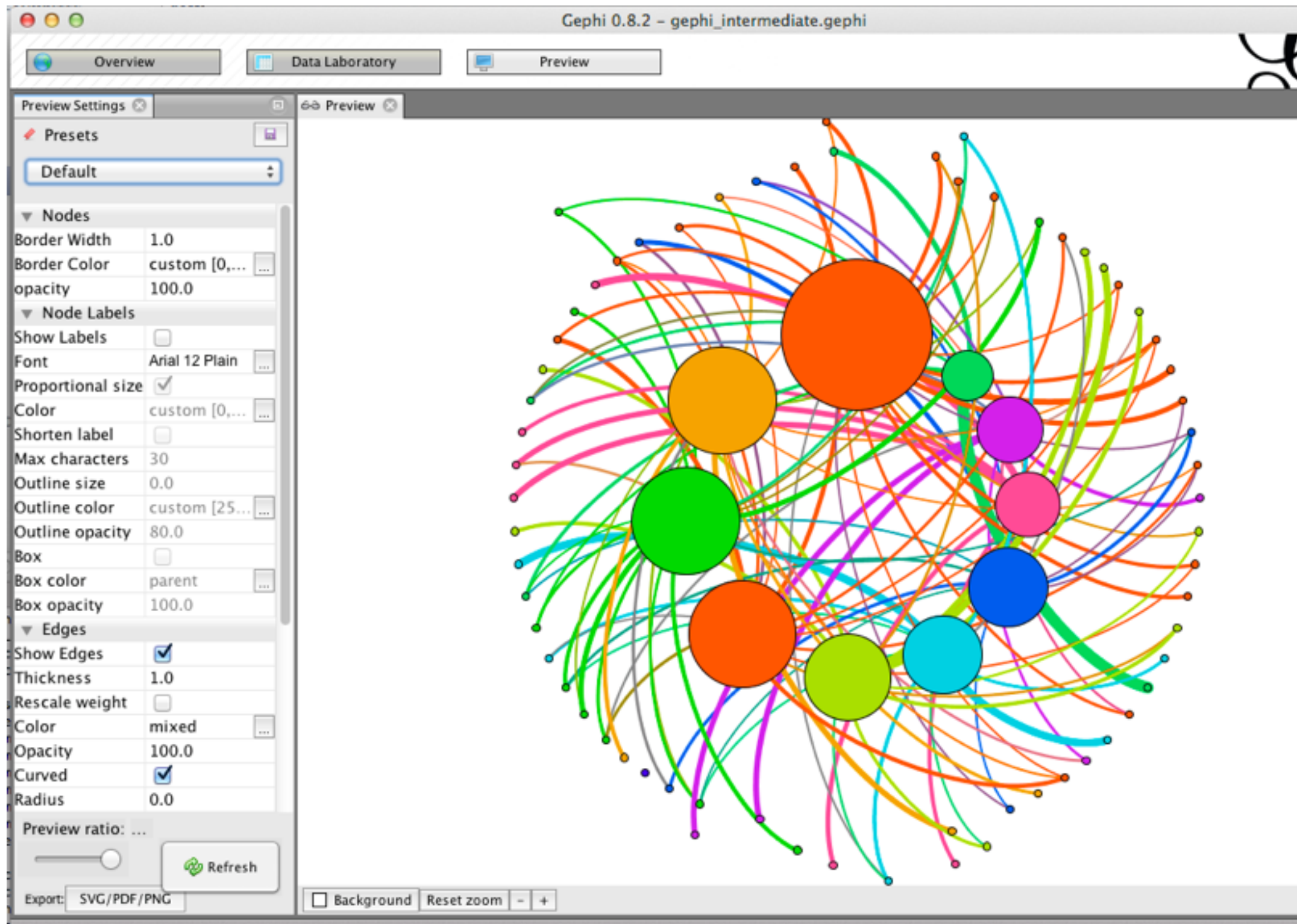
Turn on
labels and
adjust
sizes...



There are
various
options for
coloring
and sizing
at
the
bottom.

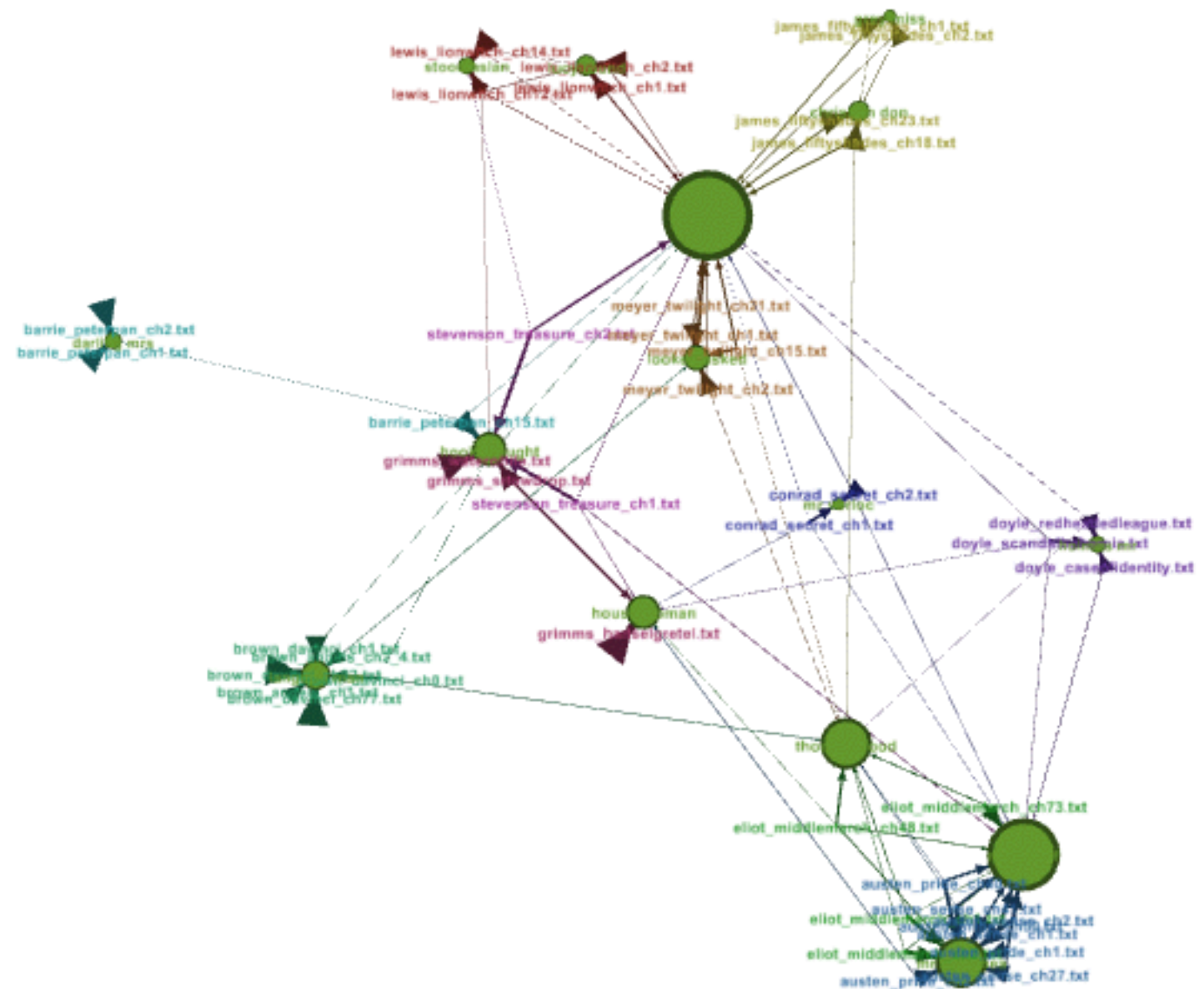


The “pretty” exportable version is on the Preview Tab.
Finish adjusting in here. But for adjustments to locations
you need to go back to Overview, then come here and
REFRESH.



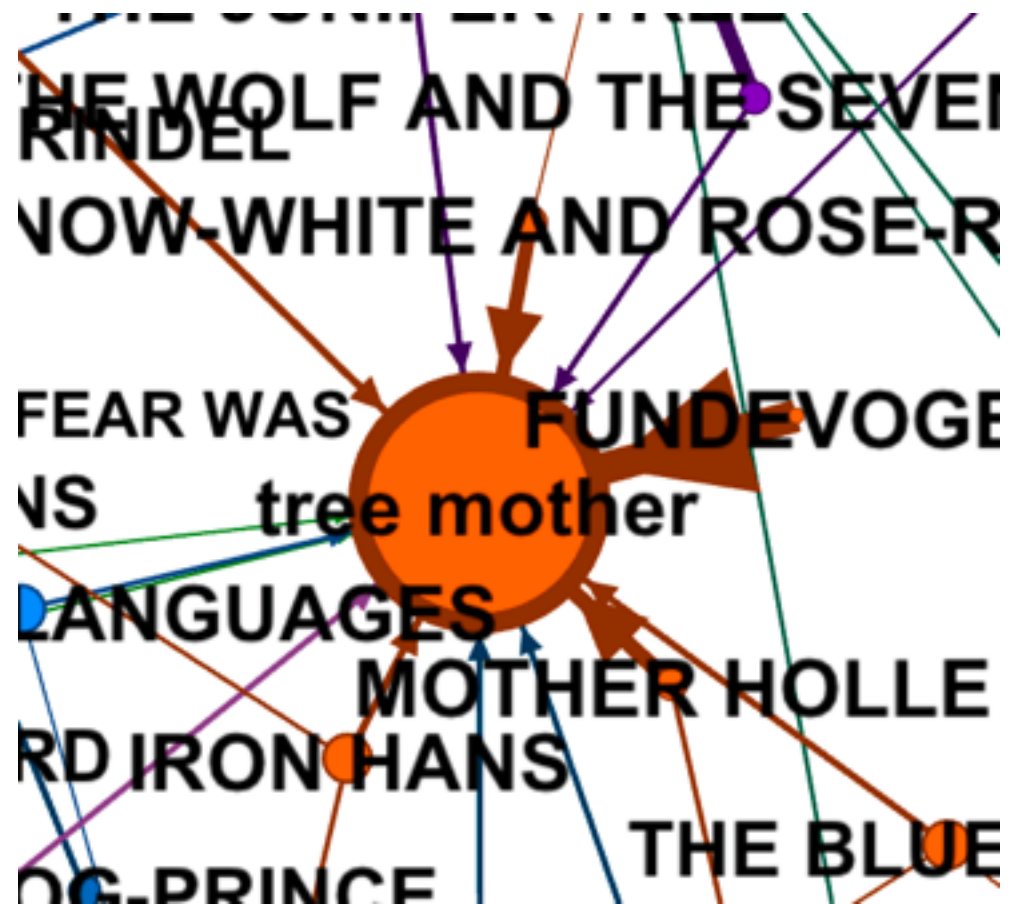
A Force Atlas2

Layout for the Mixed Chapters:

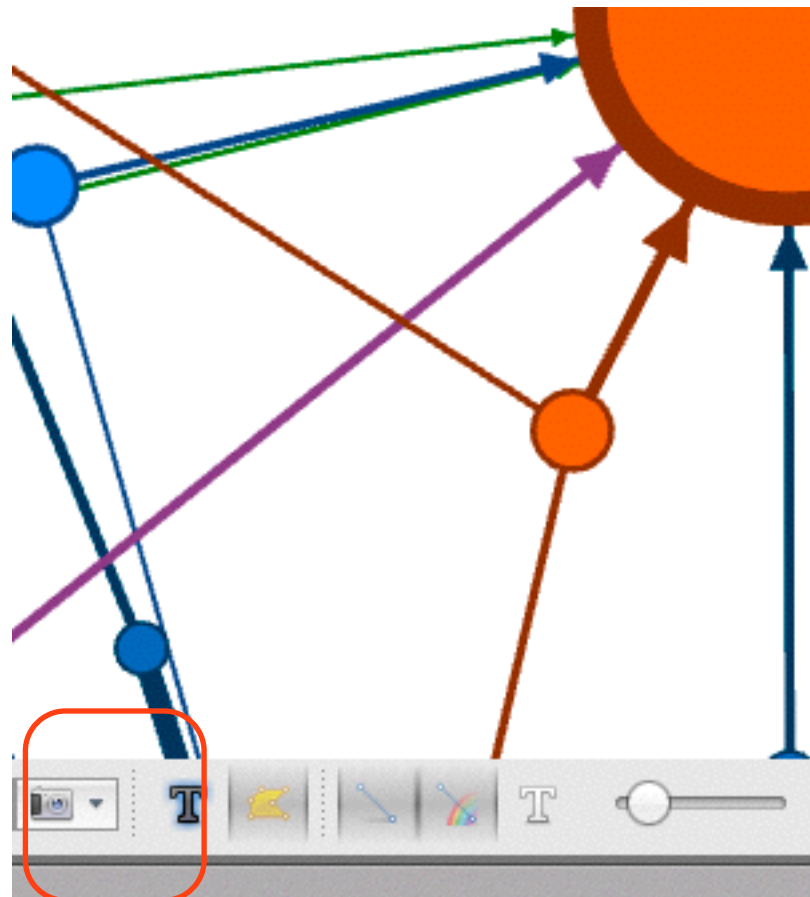


Tip: After running this, stop it before it gets too big.

A nice final adjustment workflow is to zoom in on an area....



Turn off the labels if they're on...

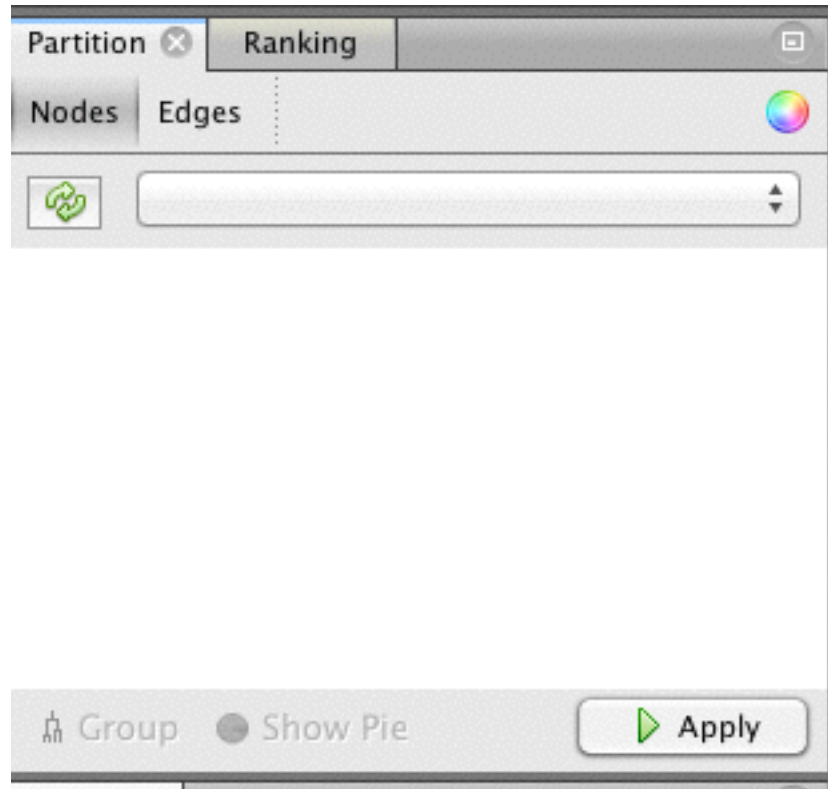


Now drag & move nodes around carefully to space things as you prefer...

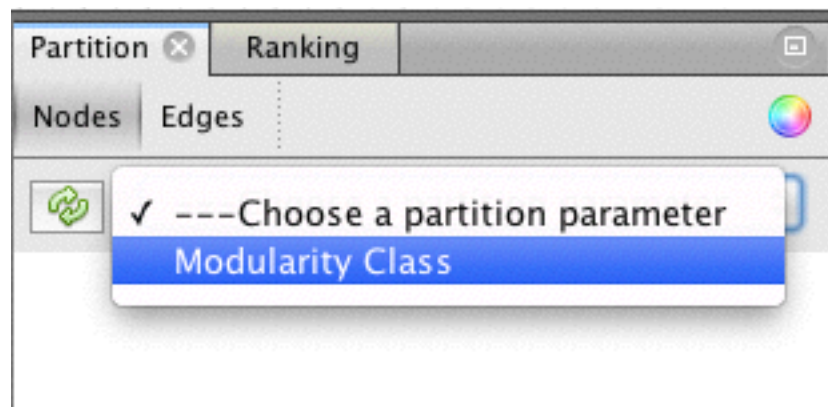
Then switch back to Preview and hit REFRESH to see if you like that layout better.

For an interesting analysis of your topics, try changing to color by “Modularity” and compare to Authors. What clusters do you see?

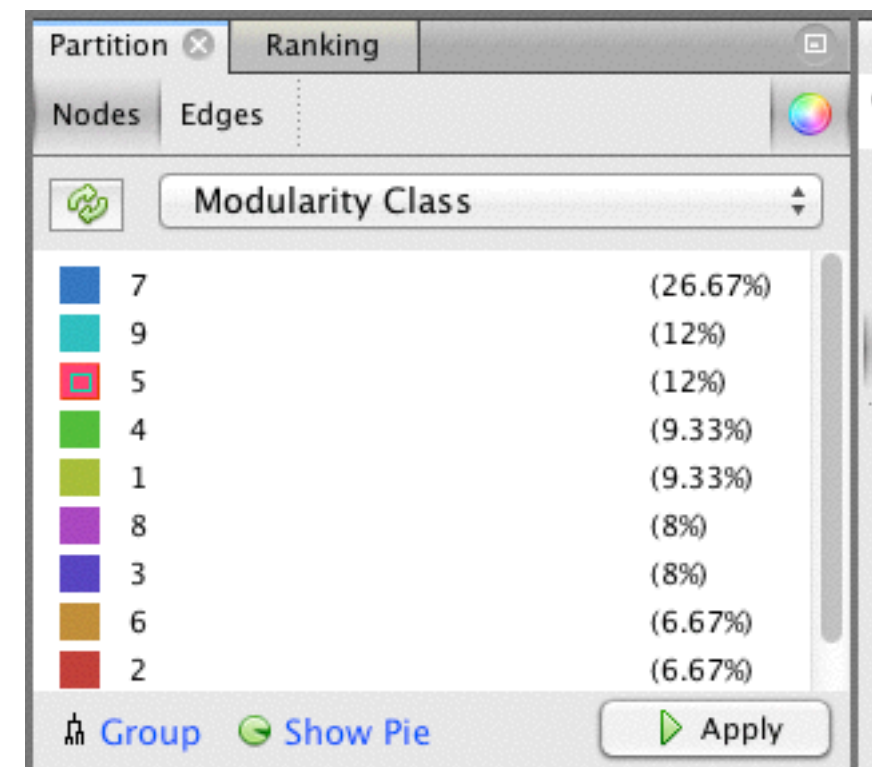
1. Hit green refresh arrow:



2. Then choose “Modularity Class” from menu:



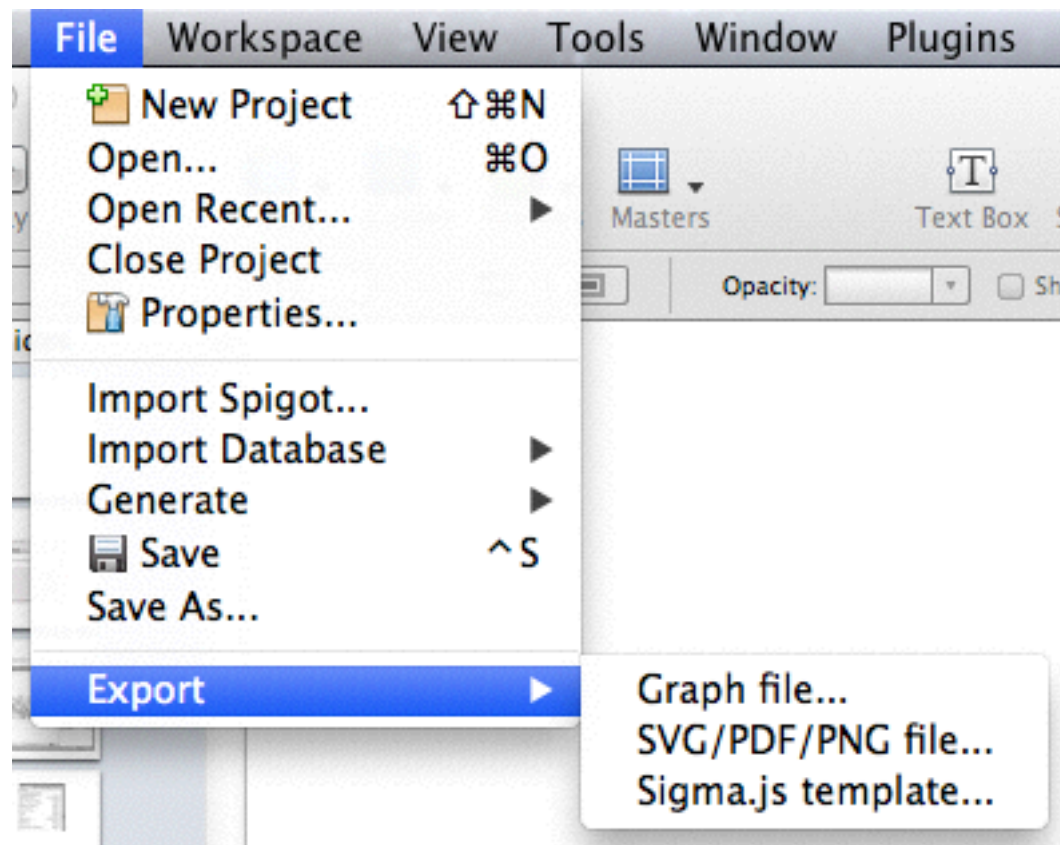
3. You'll see the “groups” and random assigned colors. Click on a color to pick a new one. Make them as different as you can.



4. Then hit “Apply.”

Export to SigmaJS

To export to an
interactive web
project after your
Layout, pick Export...
Sigma.js template.



Put in whatever useful, informative text you can supply. I like “Dim” for Hover behavior. Let’s use Modularity Class for Group.

Sigma.js Export

Export to Sigma.js template

/Users/lynn/Documents/Talks/TopicsPythonGephi/files Browse...

Legend

Node* Topic/Story

Edge* Strength

Color* Community

Branding

Logo (url)

Link

Author* Lynn Cherny

Title* Jell of Grimms

Features

☒ Include search? ☒ Group edges by direction?

Hover behavior Dim

Group Selector? Modularity Class

Image attribute? None (Default)

Attributes

Coming soon

Short Description*

An example of exporting topic model relations from Gephi using the Sigma.js plugin.

Long Description*

An example of a sigma.js export of Grimm’s Fairytales topics – chapters mapped to topics.

☐ Replace node ids with numbers

Cancel OK

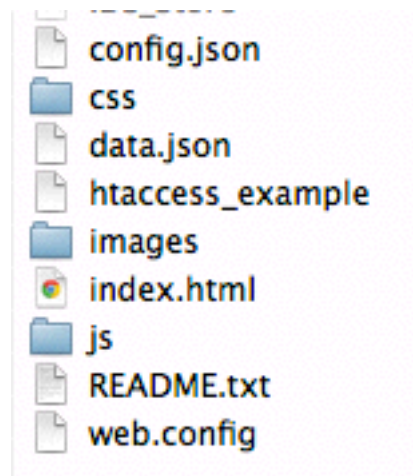
It will output a “network” folder - pick where!

A shortcut to fix and run the network you made:

If you are in the directory with the sigmajs network directory and the script `run_network.py`, you can type:

```
>python run_network.py
```

The sigmajs
“network”
folder has
these:



The run_network.py
script will fix
these values for you and
start the server.

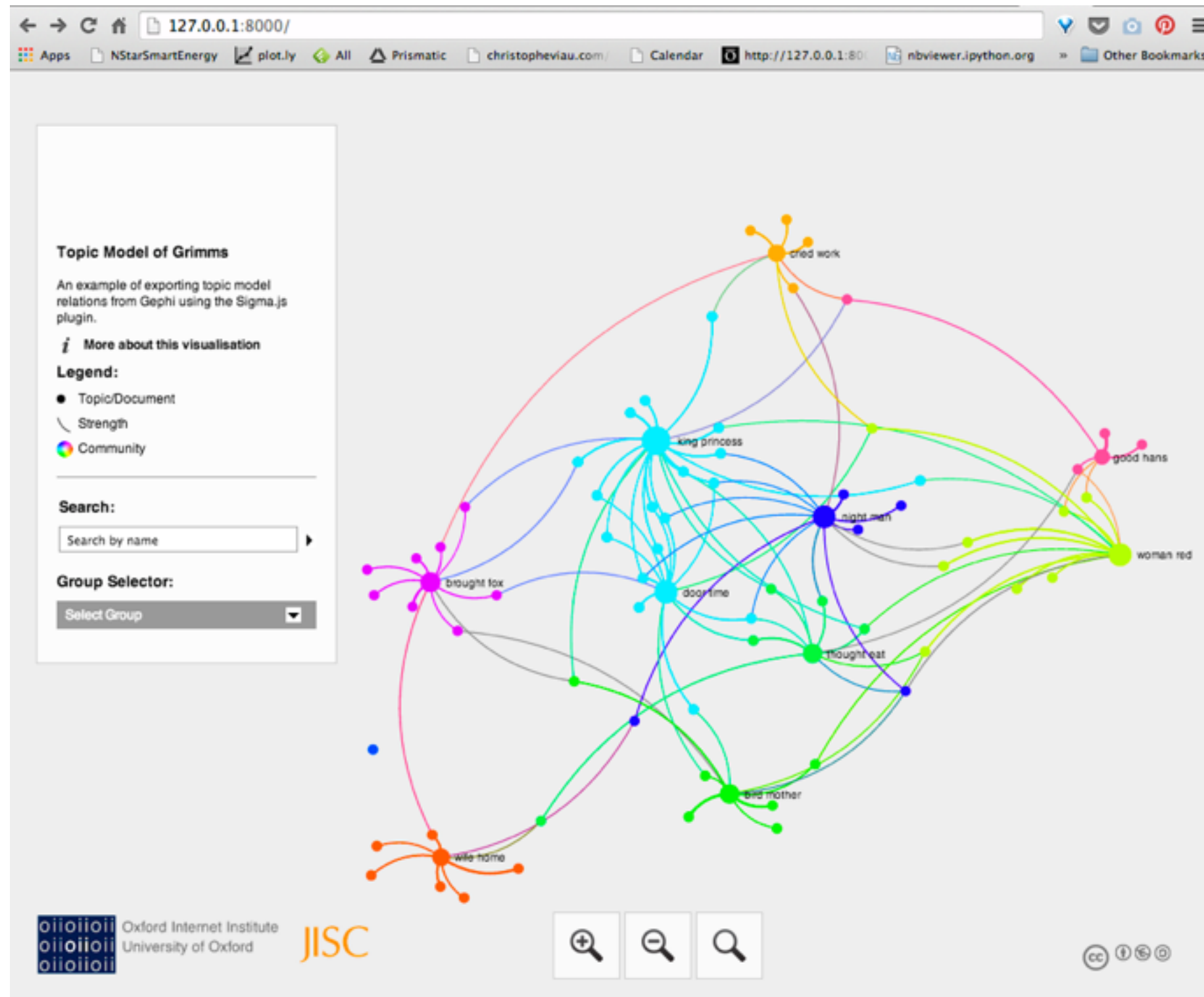
```
config.json
17  "colorLabel": "Community",
18  "nodeLabel": "Topic/Document"
19  },
20  "features": {
21    "search": true,
22    "groupSelectorAttribute": "Modularity Class",
23    "hoverBehavior": "default"
24  },
25  "informationPanel": {
26    "imageAttribute": false,
27    "groupByEdgeDirection": true
28  },
29  "sigma": {
30    "graphProperties": {
31      "minEdgeSize": 1,
32      "maxNodeSize": 20,
33      "maxEdgeSize": 8,
34      "minNodeSize": 7
35    },
36    "drawingProperties": {
37      "labelThreshold": 10,
38      "hoverFontStyle": "bold",
39      "defaultEdgeType": "curve",
40      "defaultLabelColor": "#000",
41      "defaultLabelHoverColor": "#fff",
42      "defaultLabelSize": 14,
43      "activeFontStyle": "bold",
44      "fontStyle": "bold",
45      "defaultHoverLabelBGColor": "#002147",
46      "defaultLabelBGColor": "#ddd"
47    },
48    "mouseProperties": {
49      "minRatio": 0.75,
50      "maxRatio": 20
51    }
52  }
53 }
```

Edit config.json to
increase some
sizes:

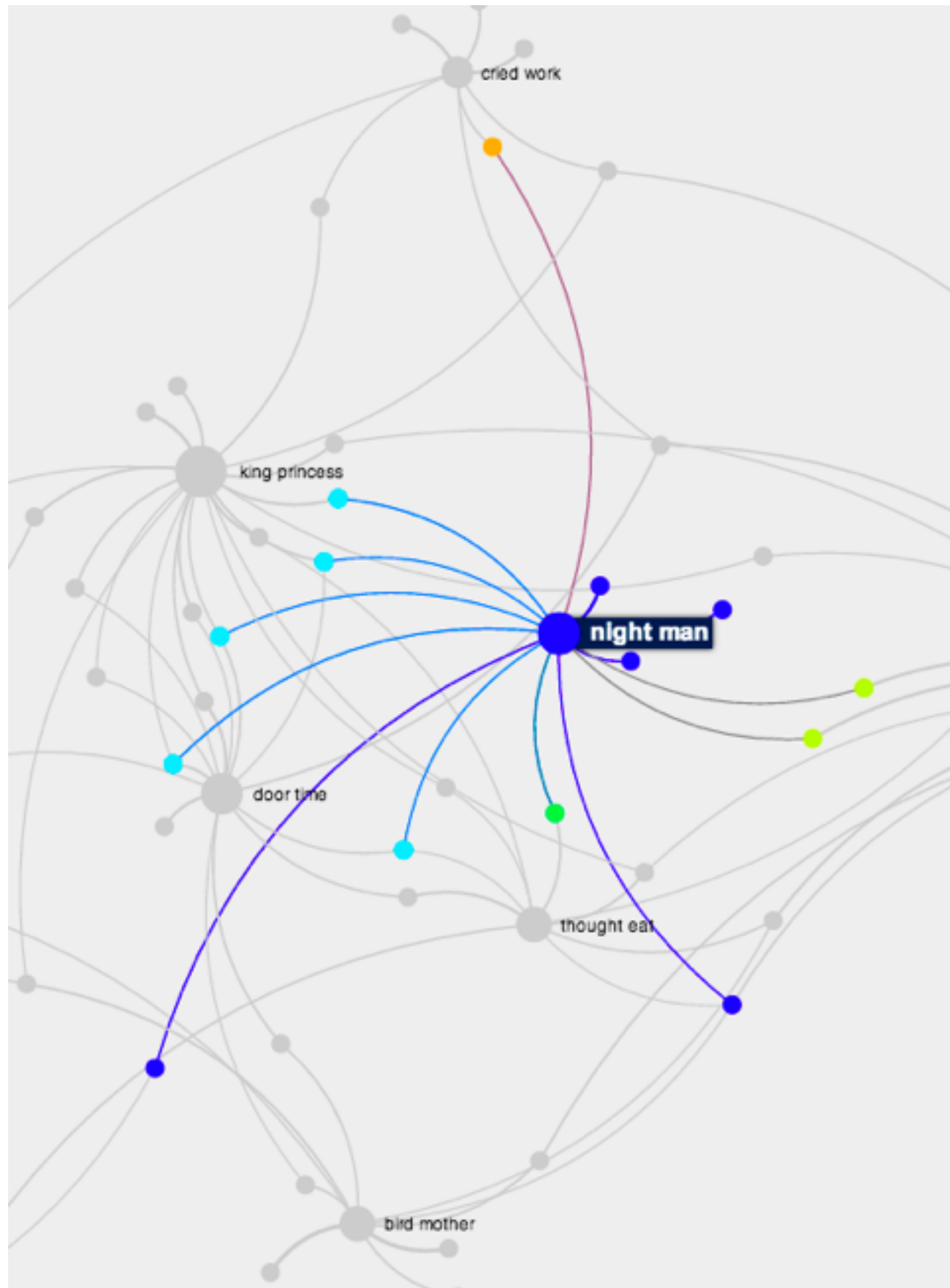
But decrease labelThreshold to maybe 7...

Move your network folder to a web server or start one locally (or use the run_network.py script):

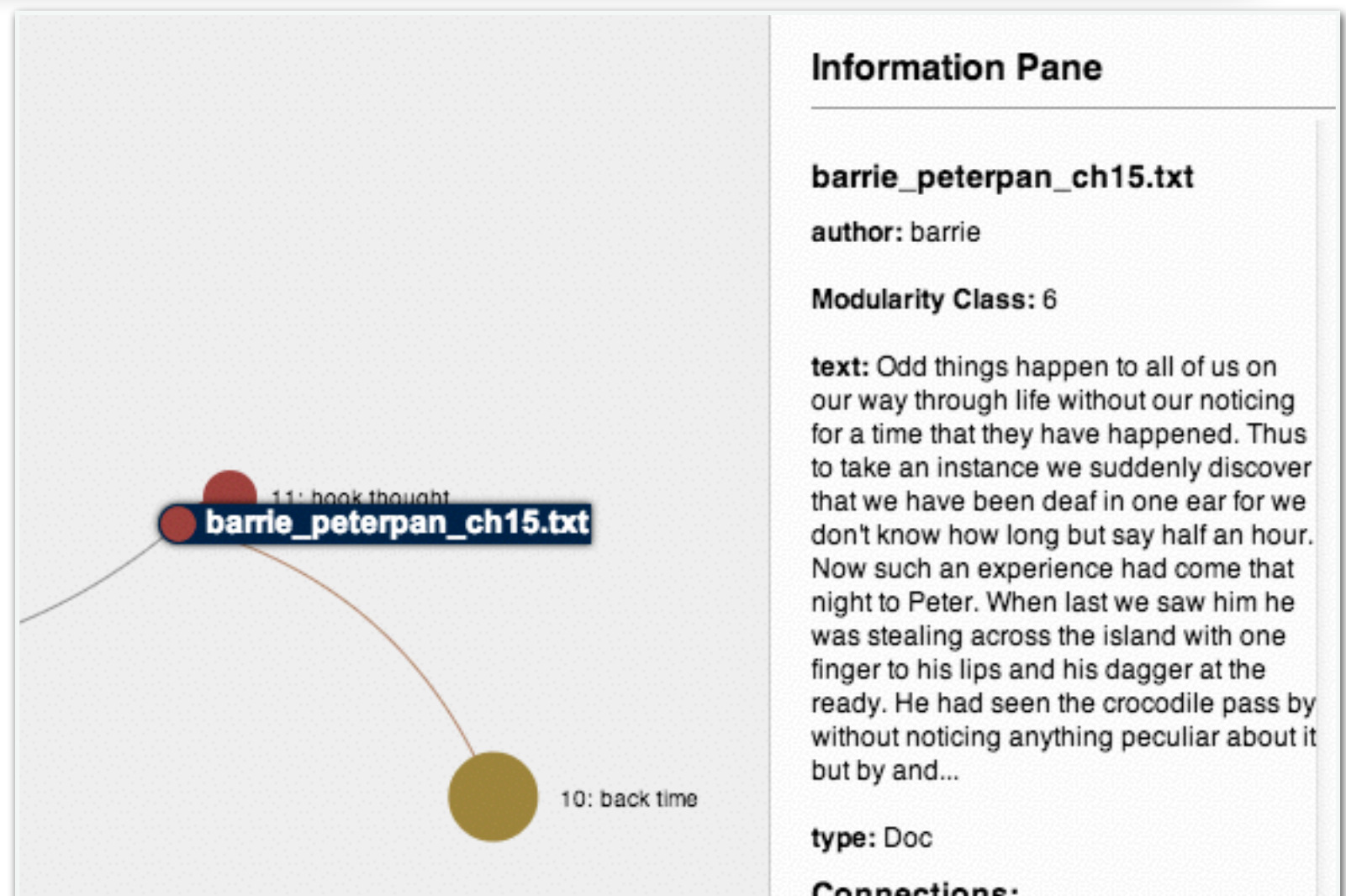
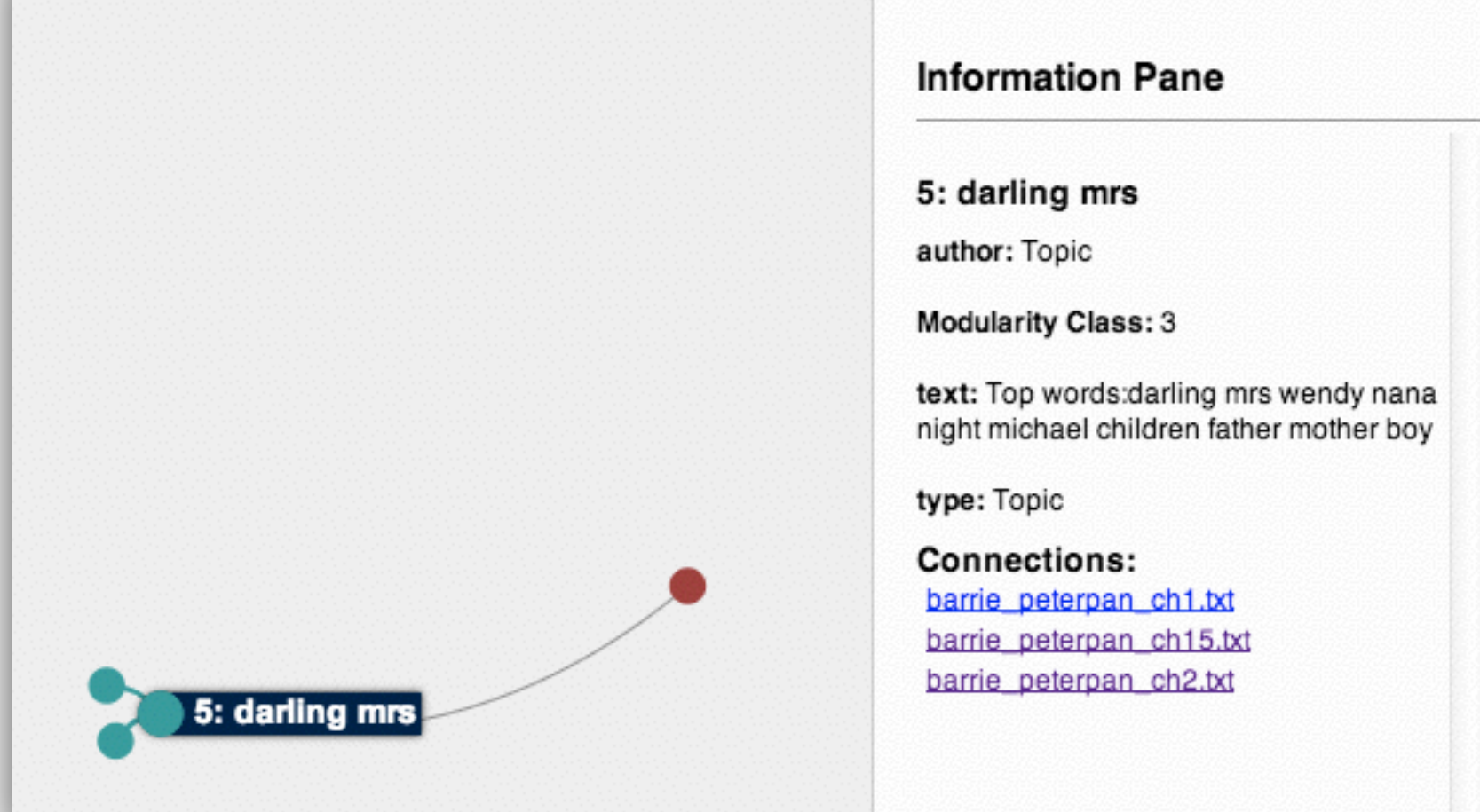
```
new-host-2:files lynn$ cd network
new-host-2:network lynn$ python -m SimpleHTTPServer 8000
Serving HTTP on 0.0.0.0 port 8000 ...
127.0.0.1 - - [30/Mar/2014 15:45:06] "GET / HTTP/1.1" 200 -
127.0.0.1 - - [30/Mar/2014 15:45:07] "GET /js/jquery/jquery.min.js HTTP/1.1" 200 -
```



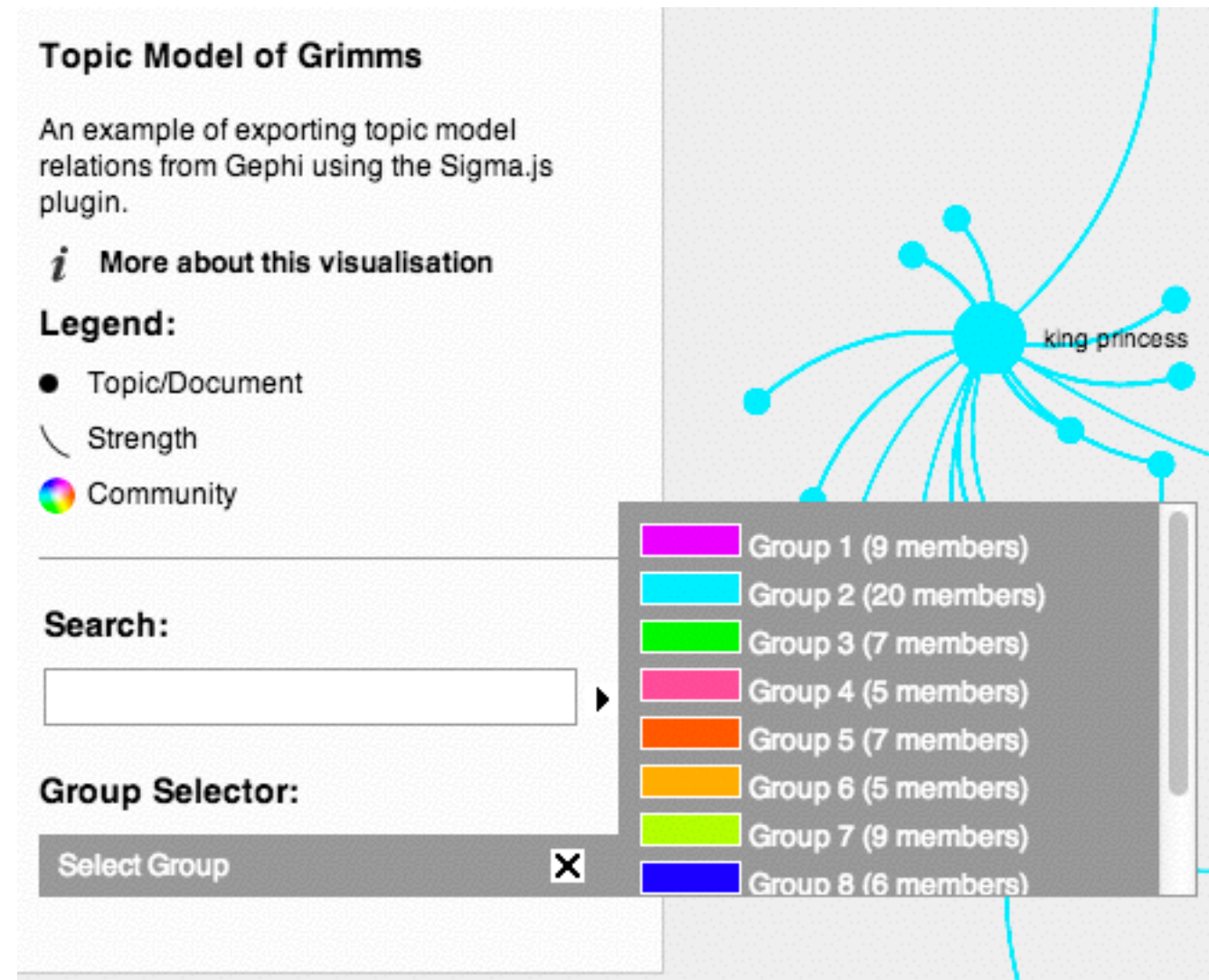
At this level, rollovers dim unconnected nodes, showing only connected nodes. Only the largest node labels are visible (topics).



Clicking
gives you
info on the
right side,
including
the text
we
uploaded
to the
graph file
from
python!



The groups
are the
modularity
classes found.
You can
search for a
node name
too...



As a last cleanup step, consider using the network to rejigger your names and labels.

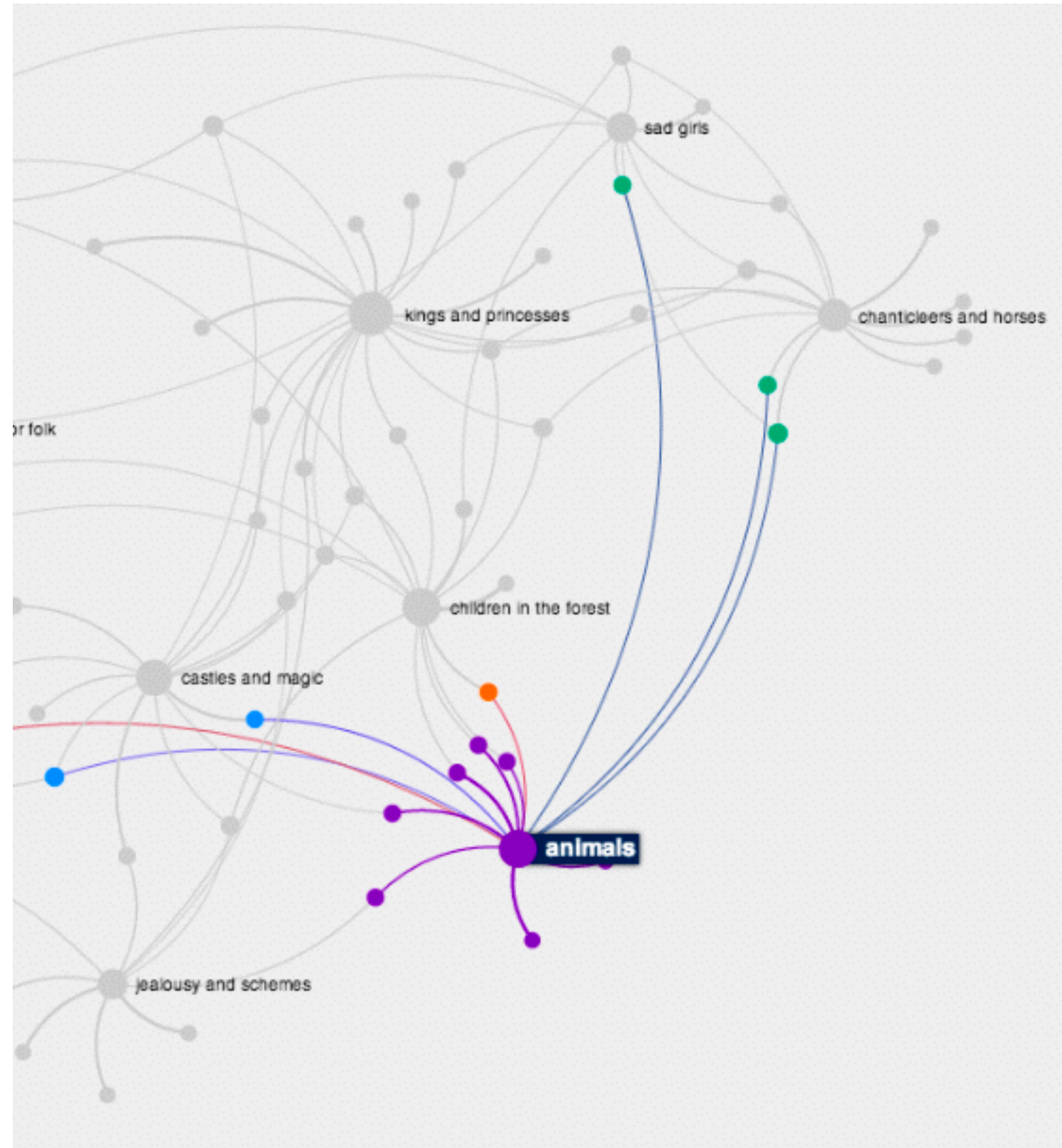
Copy the data.json file to **another name** and edit that, as you browse the network.

```
sons and last of all one daughter. Although the little girl was very
pretty she was so weak and small that they thought she could not live;
but they said she should at once be christened. So the father sent one
of his sons in haste to the spring to get some water but the other six
ran with him. ...\""",
  "title": "THE SEVEN RAVENS"
},
"color": "rgb(15,249,249)",
"size": 3.0
},
{
  "label": "bird mother",
  "x": 8.915628433227539,
  "y": 432.99188232421875,
  "id": "66",
  "attributes": {
    "Modularity Class": "1",
    "text": "\"Top words:bird mother beautiful white tree children gold cook
home father\"",
    "title": "bird mother"
  },
  "color": "rgb(62,249,15)",
  "size": 31.5
},
{
  "label": "CAT AND MOUSE IN PARTNERSHIP",
  "x": -674.8820190429688,
  "y": 9.39062786102295,
  "id": "13",
  "attributes": {
    "Modularity Class": "9",
    "text": "\" Chapter CAT AND MOUSE IN PARTNERSHIP A certain cat had made the
```

relabel the topics with a more human interpretable name based on inspection

For example...
swapping in my
data_renamed.json
file...

But you may find
you want to redo
the topic modeling
with more or
fewer topic.



- The utility of this graph is improved by useful text/content additions in the graph file (the python exercise we went thru first) and some tweaks in the output visuals. For more info on “rollover”, add info to each node you export in python!
- It could be even better by post-processing the modularity groups to have better names than “group 2”... more cleanup work.
- NOTE: The sigma.js exporter from Gephi is out of date, using an old sigma.js format. It also seems to not differentiate weights/sizes of edges

Hope this helped!

lynn@ghostweather.com