**[](file:///\\pv-psych-align\deansl$\Mina%20dokument\Mina%20bilder\melbourne%202006\2007-01-20-1504-14\MOV00009.3gp)**

**Exercise**

**Simulating random graphs in MPNet**

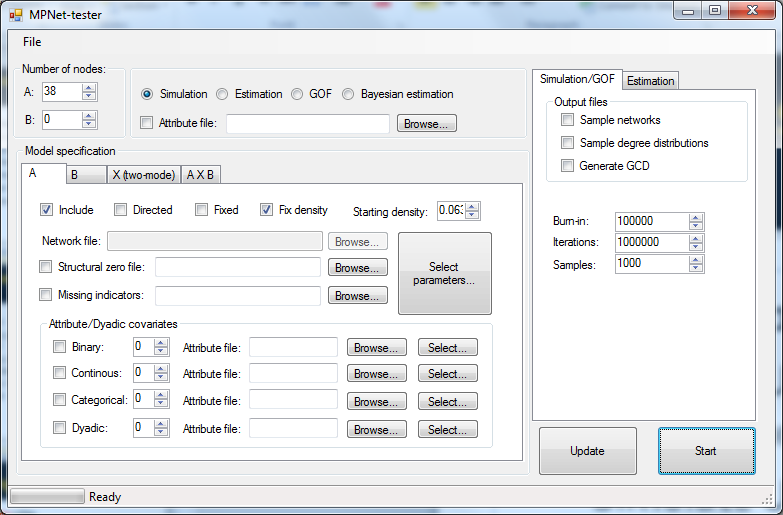
Swinburne University of Technology, Australia ©

* 1. Generating Bernoulli random graphs in MPNet

10.1.1 Simulations

Open MPNet by clicking on the **MPNet.exe** file. We are using the simulation function for this exercise, so make sure you click the ‘**simulation**’ option at the top of the program.

Set up MPNet to simulate a distribution of graphs of 38 nodes, with fixed density of 0.06259 (the number of actors and density of the “**comm\_undirected.txt**” network of The Corporation – the example dataset from Lusher, Koskinen & Robins, 2013).



Choose session folder

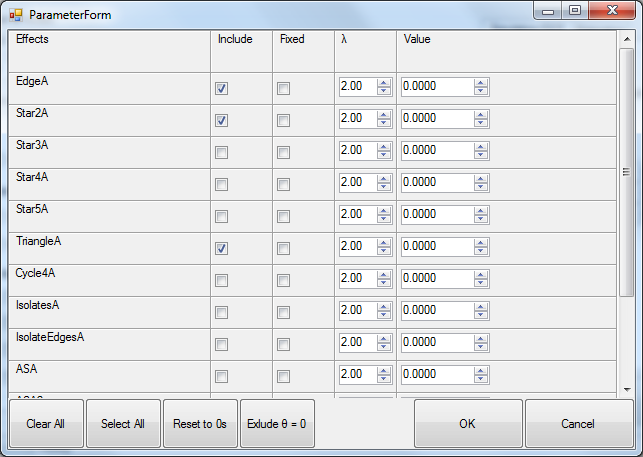
Number of nodes

Fix Density

Density = 0.06259

Click **Select parameters**

Clicking the **Select Parameter**  button opens a new window that permits you to select parameters for the model and parameter values. In this case, we won’t have any parameters (other than the fixed density of 0.06259), but by selecting parameters and leaving the parameter value at 0, MPNet will count the relevant configurations in the simulated graphs. In this case, select **edge**, **2-star** and **triangle**. Then click **OK**.



Back on the main window, you are now set up to simulate a number of 38 node graphs, each of which will have a density of 0.06259. Notice that there is a “burn-in” of 100,000 (a preliminary simulation before the final simulation starts), you will simulate 1,000,000 graphs (“number of iterations”), but the program will record results for 1,000 graphs (“Number of samples ...” – that is every 1,000th graph in the simulation.) Click on **Start!**

1.2 “*session name*\_sim.txt” files

There are a number of output files in the Session folder. The most important is the “simulation-*session name*.txt” file, which contains the statistics for all the 1000 graphs in your sample from the graph distribution. It lists the number of the graph (‘id’), and the number of edges, 2-stars and triangles in each graph. Notice that each graph has the same number of edges, as it must do for a fixed density simulation.

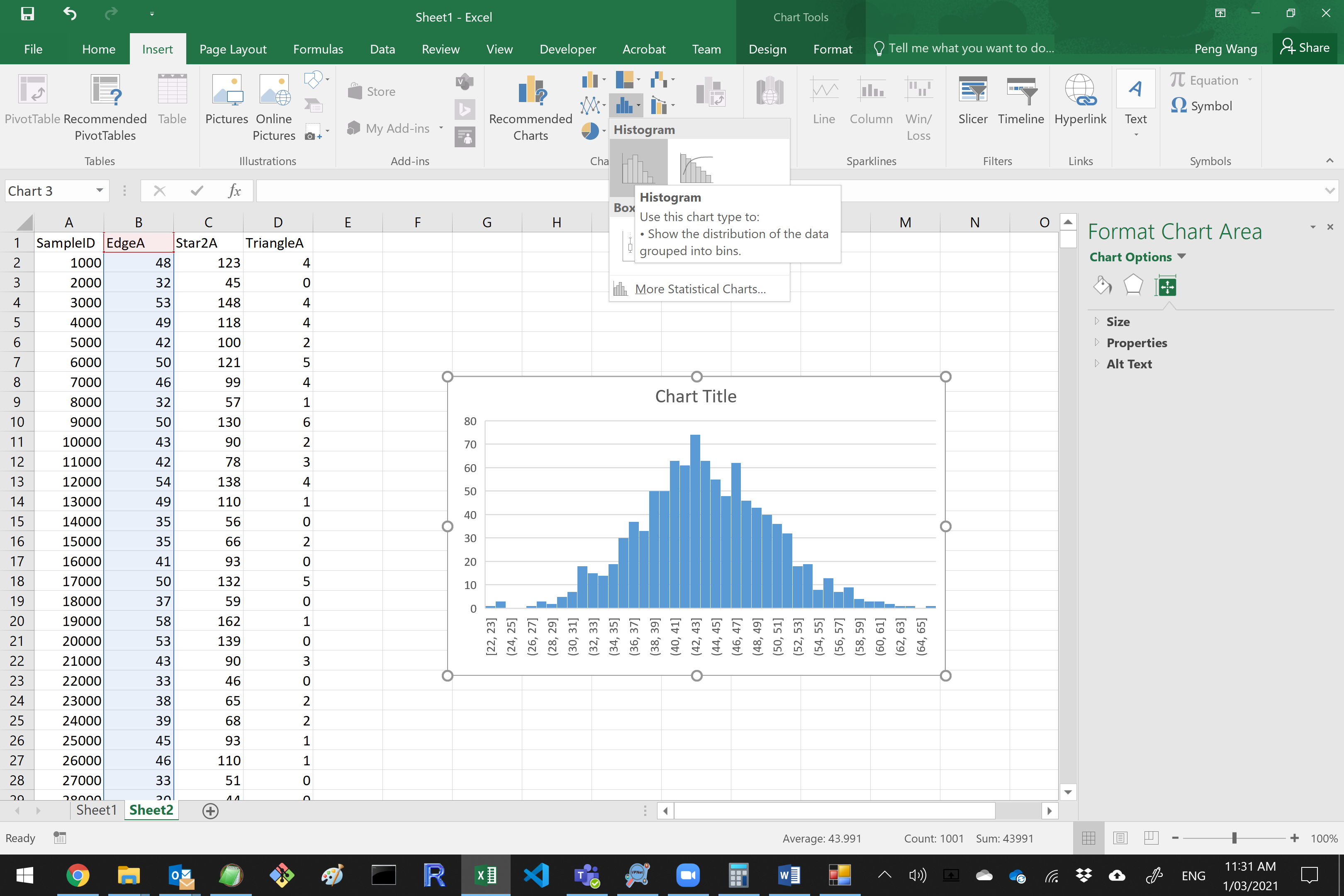
1.3 “*session name*\_Network\_A\_1001000.txt” files

The output file “*session name*\_Network\_M\_1001000.txt”, or “*session name*\_Network\_A\_1001000.txt” has details for the very last graph in the simulation. It is in the form to be read into Pajek (another useful SNA visualization program) to produce a network visualization.

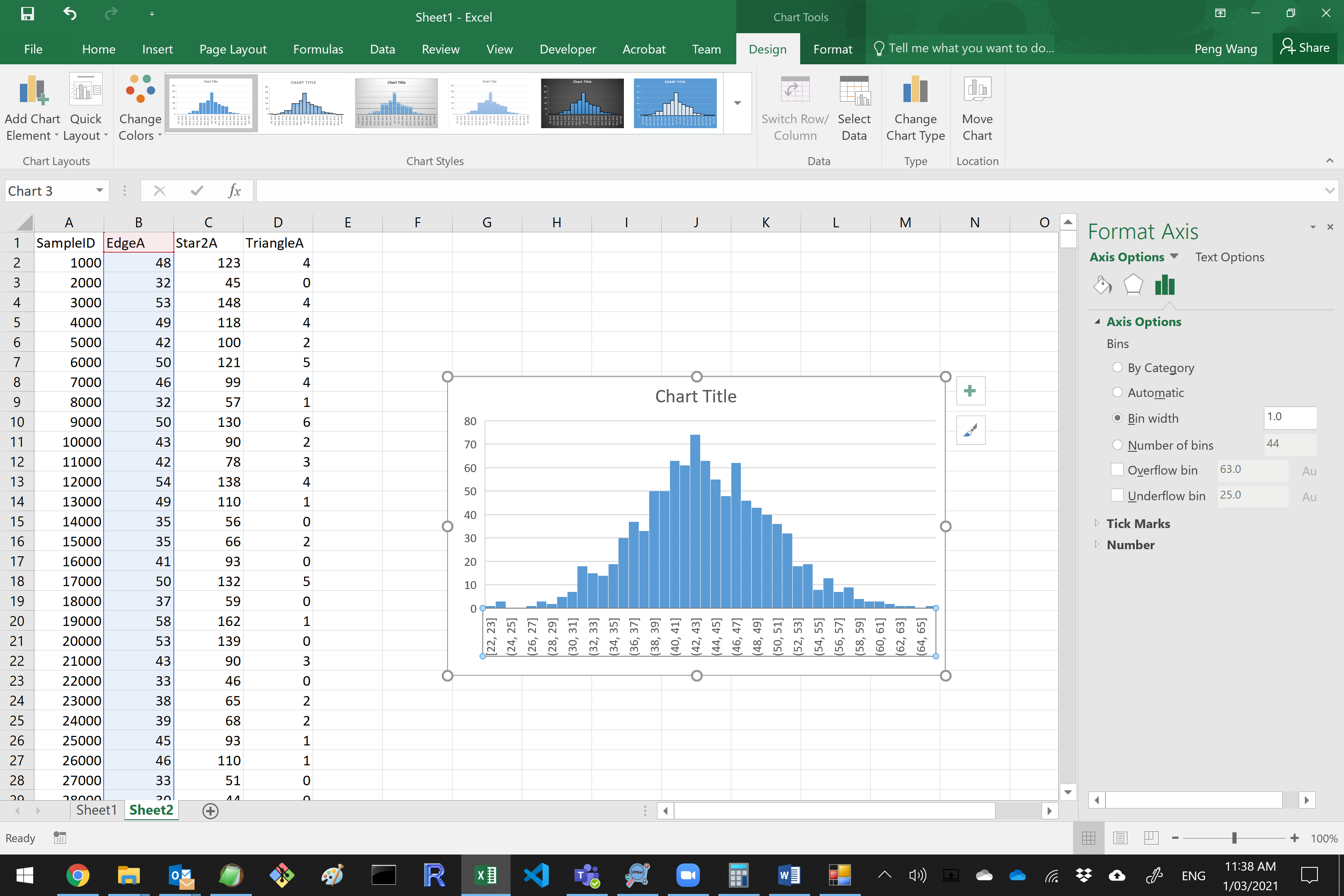
What is the average number of triangles in your graphs? Import or copy/paste the “simulation-*session name*.txt” file into Excel or a statistical package and calculate means and standard deviations. Or, if you have SPSS on your computer, open the “sim\_*session name*.sps” file, which is an SPSS syntax file that will read in this data and produce histograms, with means and standard deviations, and scatterplots of the statistics across the simulation.

**Histograms in Excel**

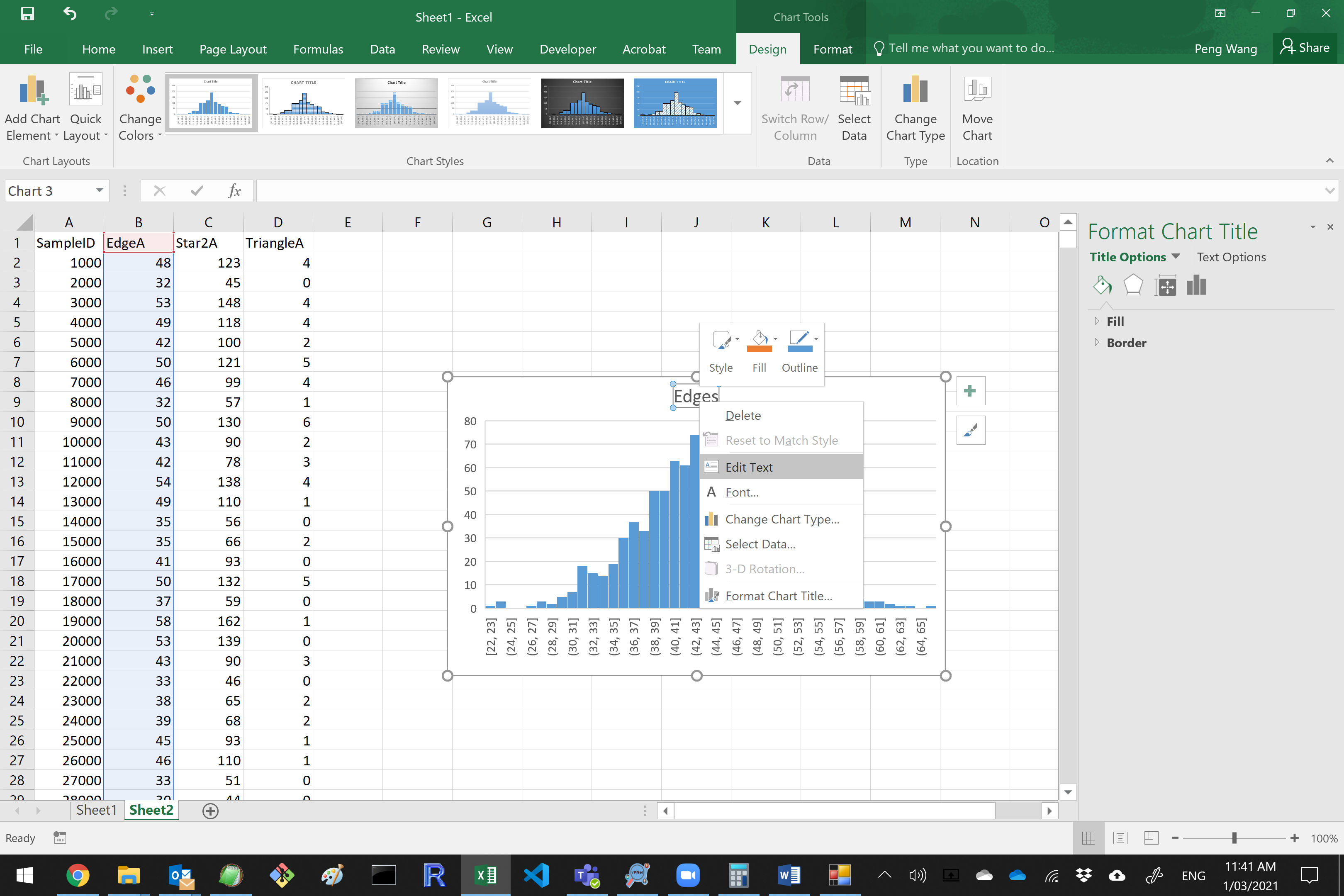
1. Copy and paste the content of *session name*\_sim.txt to an empty Excel spread sheet.
2. Select the Column of data we want to plot in a Histograms (e.g. Column EdgeA which contains the simulated numbers of edges).
3. Under the “Insert” tab, click on the Histogram button.



1. Double Click on the horizontal axis of the Histogram, the Format Axis options will show up on the right hand side.
2. Adjust Bin width to 1.0.



1. Right-click on the Title of the chart, and select “Edit Text”. Then type in the name of the statistic (e.g. Edges) as the Title of the chart.



**Something to consider:**

There are 12 triangles in the original The Corporation’s “**comm\_undirected.txt**” dataset. In the 1,000 graphs of your simulation, how many have 12 or more triangles?

Is the fixed density distribution a good model for the data (or in other words, is it likely that the The Corporation’s “comm\_undirected” data come from a distribution of graphs with fixed density of 0.06259)?

10.2 Simulating Bernoulli random graph distributions using MPNet

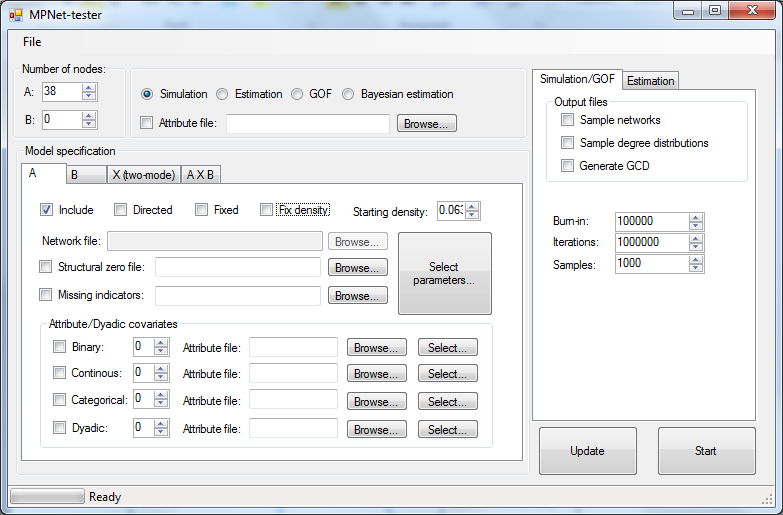
Maybe the fixed density limits the number of triangles. Instead of fixing the density at exactly 0.06259, we could have a model where the chances of a tie occurring between two nodes was 0.06259. This is the Bernoulli graph distribution. The only non-zero parameter in a Bernoulli graph model is the edge parameter.

**The only parameter in a Bernoulli graph distribution (a simple random graph distribution) is an edge parameter (L)**. Selecting a value for the edge parameter fixes the probability *p* of an edge between any pair of nodes. The relationship between *p* and the L parameter is:

*L parameter =* ln*(p/(1-p))*

It follows from the calculation that an L parameter value of -2.7065 (approx) corresponds to a probability of 0.06259. A distribution of graphs with that probability should have an AVERAGE density of 0.06259.

Start back with MPNet. Give this session a new name (or delete or remove the previous output files from the Session folder). Change the Fix density instruction. The starting graph density is now not important (it will be removed through the burn-in), so you can just leave it.

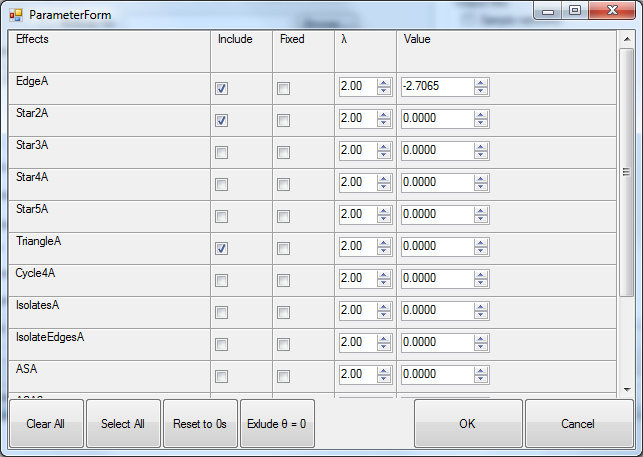


This doesn’t matter now

**DO NOT** Fix Density

Click **Select parameters**

When the structural parameter window appears, enter -2.7065 as the edge parameter value. Then run the simulation as before.



Then run the simulation as before. Notice now when the graph density is *not* fixed that the number of edges varies from graph to graph.

Enter the data from “simulation-*session name*.txt” data into SPSS using the syntax file, or into Excel or another statistical package.

**Given the number of triangles in the data, is a Bernoulli distribution a good model?**

Draw a scatterplot of the number of triangles against the number of edges. The Corporation data has 44 edges and 12 triangles. How many graphs from your sample have BOTH 44 edges and 12 triangles?

**[](file:///\\pv-psych-align\deansl$\Mina%20dokument\Mina%20bilder\melbourne%202006\2007-01-20-1504-14\MOV00009.3gp)**

**Exercise**

**Density, reciprocity and clustering:**

**VPNet**

Swinburne University of Technology, Australia ©

3.1 The data: net1234

Go to **Data > net1234**

This set of *directed* networks represents a communication network within a government instrumentality, with employees asked who they mainly communicated with to get their work done.

* Network 4 (net4) is a slightly altered directed network from a real dataset
* net1, net2 and net3 are all simulated networks. All networks have 85 nodes.

**Names**: Node\_file\_net1234.txt

net1.txt

net2.txt

net3.txt

net4.txt

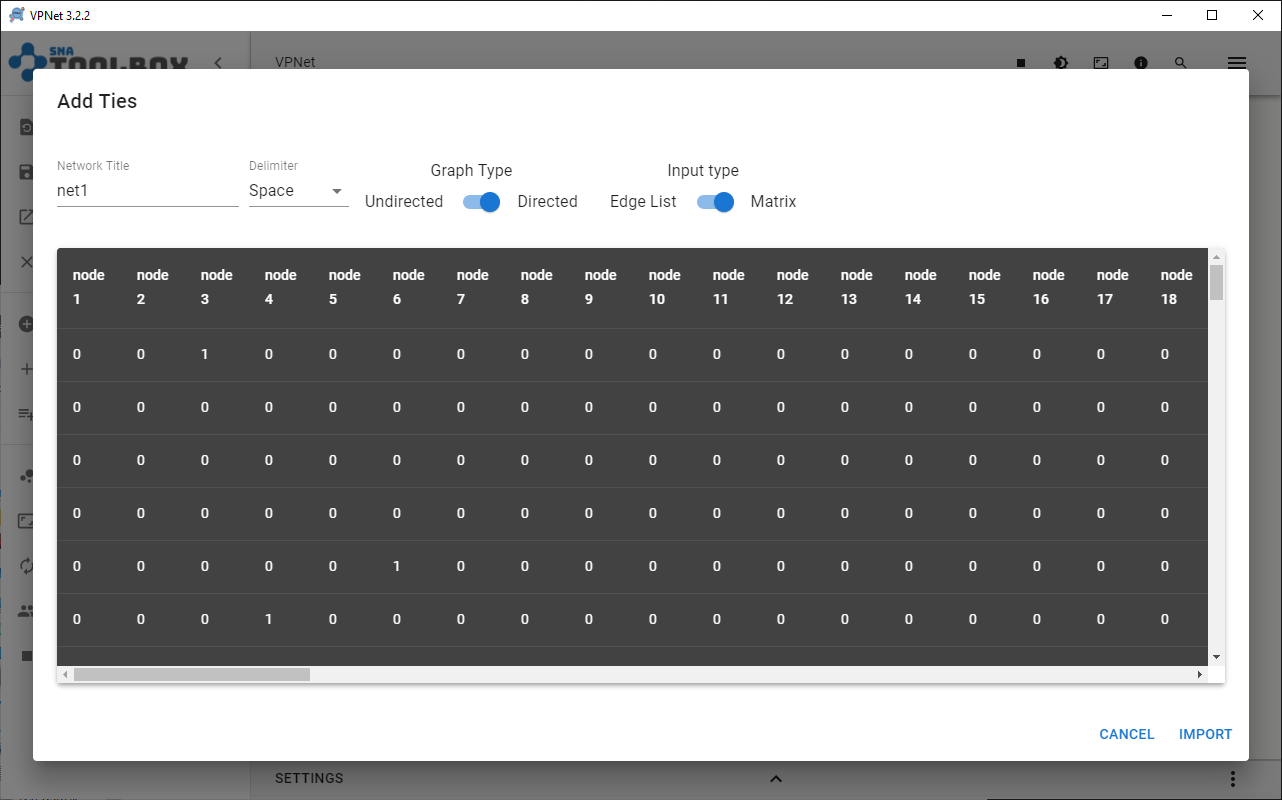
**Nodes**: 85

**Type**: Communication

In this exercise we use VPNet to calculate network indices relating to density, reciprocity and clustering.

Open the Attribute file first (using ‘**Load Nodes’**) and then open each these 4 text files of networks (using ‘**Add Network Ties**’).

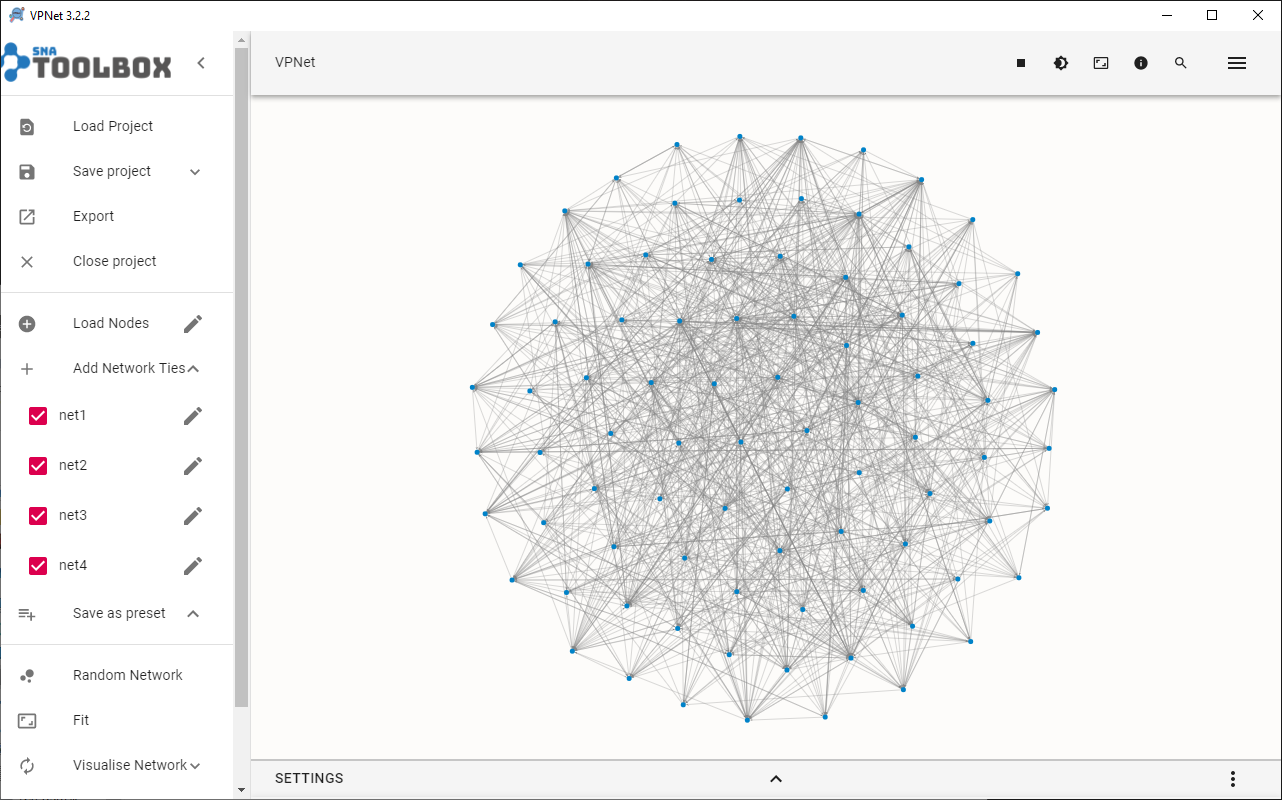
Note that the networks are directed and matrices, and VPNet should detect this and select these options in the import window for you.



3.2 Drawing the networks

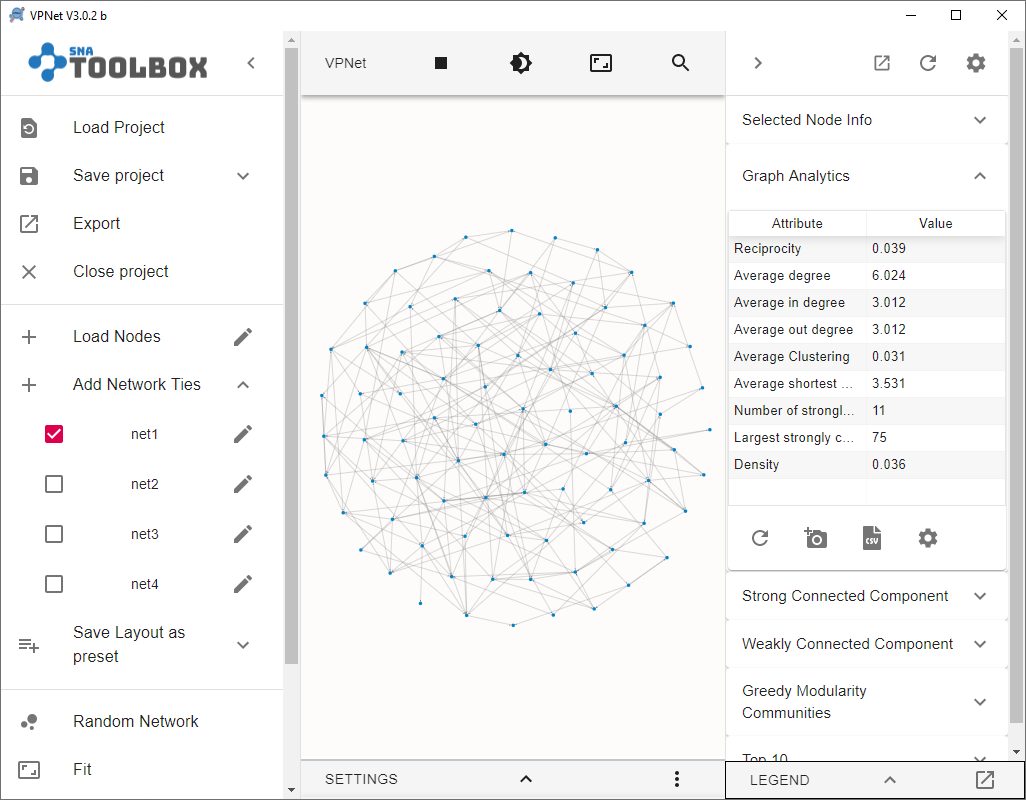
Of course, we learned before how to visualize a network in VPNet. VPNet draws your network automatically for you once you enter the data.

Importantly, VPNet can visualise multiple networks. You can visualise all 4 networks together, or you can select the network you wish to view by clicking the box next to the name of the network.



3.2.1 Density

With the network open, go to the Graph Analytics tab on the right-hand side panel, and then click on the Refresh Analytics button: 



Calculate density for all four networks.

*Question:* List the density of each of the four networks.

Net 1 \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Net 2 \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Net 3 \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Net 4 \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

*Question:* Given that each network has 85 nodes, calculate the number of arcs in each?

(Recall that for a directed network, **density = (number of arcs)/(*n*\*(*n*-1))**, where *n* is the number of nodes.

What was a feature that we imposed when we simulated networks *net1*, *net2* and *net3*?

3.2.2 Reciprocity

You will note that there are reciprocity measures.

Calculate **reciprocity** for all four networks.

*Question:* List the reciprocity of each of the four networks.

Net 1 \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Net 2 \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Net 3 \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Net 4 \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

3.2.3 Clustering

VPNet can calculate the ‘**average clustering**’ coefficient for each network. The clustering coefficient for a node is the density of ties among nodes that are tied to it (i.e. the personal or egocentric network of the node). The clustering coefficient of a network is the average across all nodes.

Under the ‘**Graph Analytics**’ tab, find the ‘**average clustering**’ coefficient which you will need to do network by network.

Calculate the **average** **clustering** for all four networks.

*Question:* What are the **average clustering coefficients** for each of the four networks?

Net 1 \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Net 2 \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Net 3 \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Net 4 \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

In summary:

*Question:* Reviewing your results, what are some of the differences among *net1*, *net2* and *net3*?

Which network or networks does the real network (*net4*)most resemble?

Exercise: Degrees and Centrality

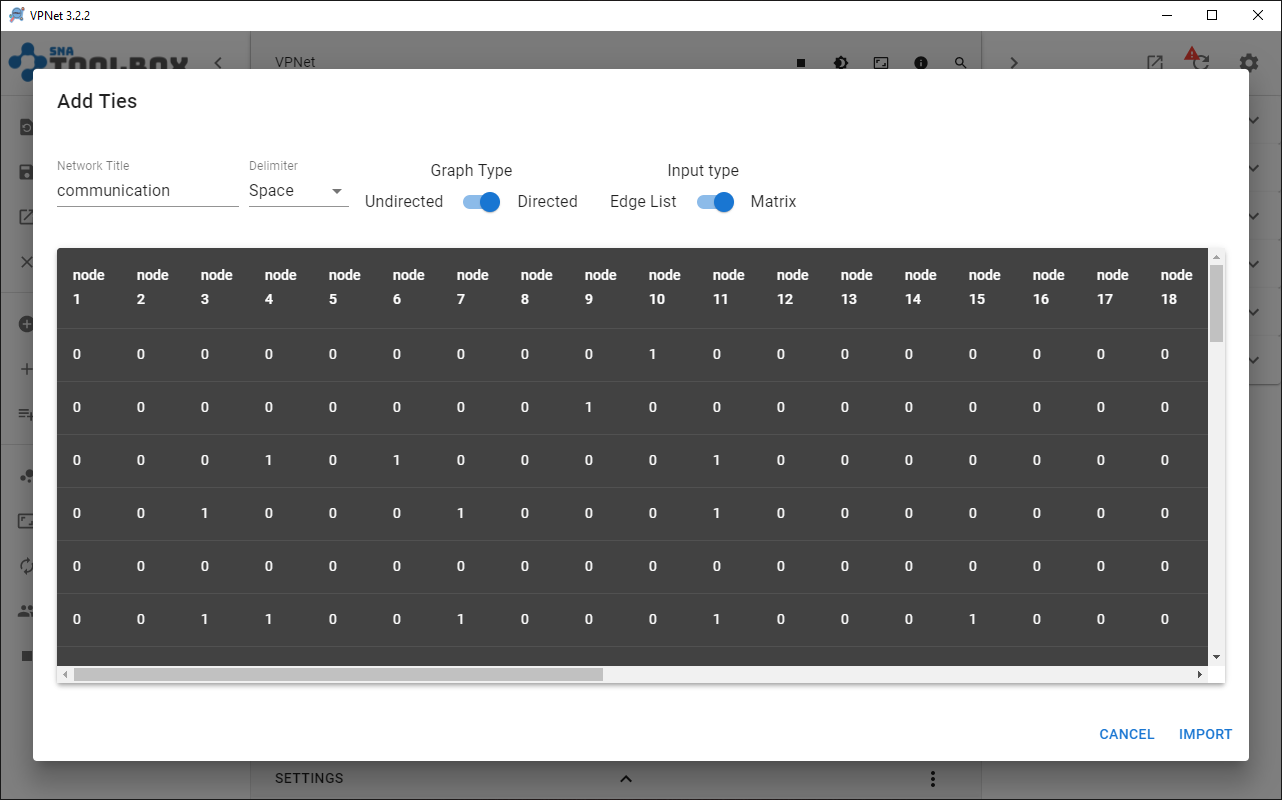
4.1 The data – Communication in The Corporation

In this exercise we will examine the network of collaboration ties from **The Corporation** data. There are 38 actors in this network of managers.

Rather than create data, this time we will import data that we already have which is in textfile format. As we found out in Exercise 1, for VPNet we need two files:

* **An attribute file** (which has the nodes and attribute data)
  + the\_corporation\_attributes.txt
* **A network file** (which has the ties between nodes)
  + communication.txt

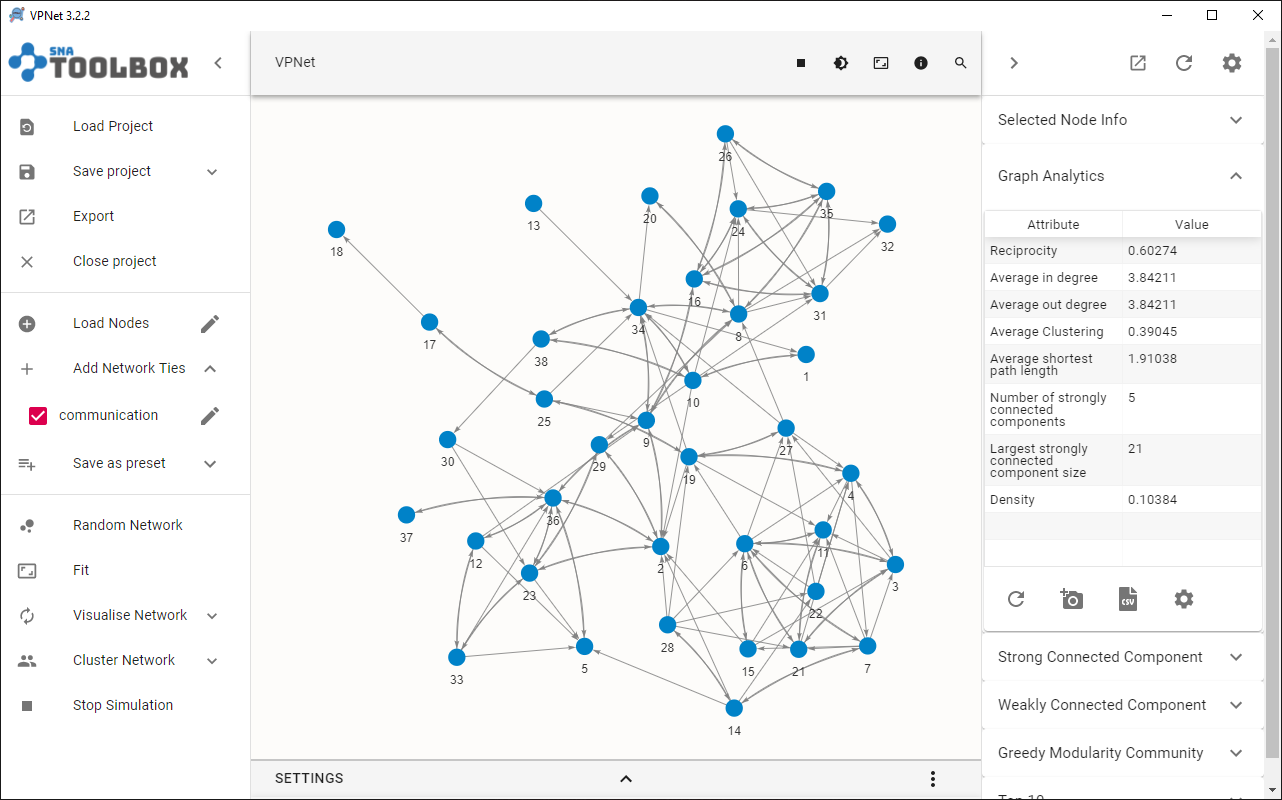
NB: This is a directed network so make sure that you choose ‘Graph Type’ = ‘Directed’.



4.2 Degrees

Here we are interested in the number of ties a node receives (indegree centrality) and the number of ties a node sends (outdegree centrality).

To calculate the **Average Indegree**, we can select **Graph Analytics** tab on the right hand side panel, and hit the *Refresh Analytics* icon



As we can see, for the communication network, the average indegree = 3.842

(Why the average in degree and average out degree are the same?)

Exercise: Bipartite & Multiplex Visualisation

In this part of the exercise we want to show you how to visualize bipartite networks – that is, networks with more than one mode.

**Data:**  VPNet\_researcher\_node\_data.txt

VPNet\_collaboration\_network.txt

51 researchers (node set = 1)

43 projects (node set = 2)

**Attributes:** Gender =1 (female)

Core = 1 (part of the central research team)

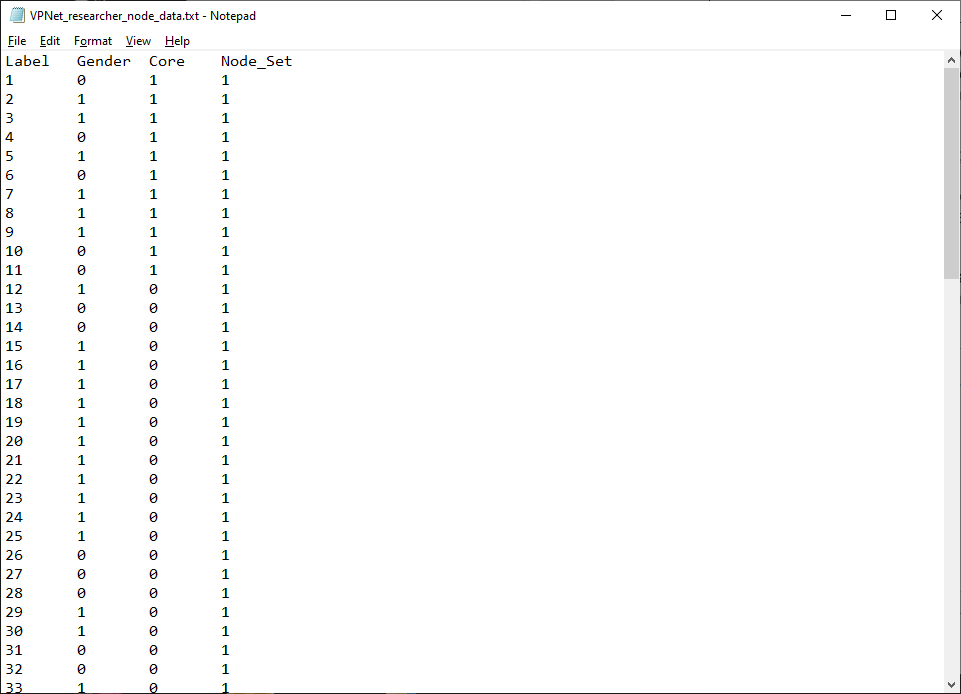
6.1 Bipartite networks

Bipartite networks are also known as two-mode networks. An example of a bipartite network is company boards (comprising one mode or set of actors) and their associated directors (the second mode). Another, which we shall use here, is researchers and the projects they are collaborating on.

6.1.1 Data Structure in VPNet

For the VPNet node/attribute files, the two mandatory fields are **Label** and **Node\_Set**. To create two-mode data, we need to make changes to the **Node\_Set** values in the following way.

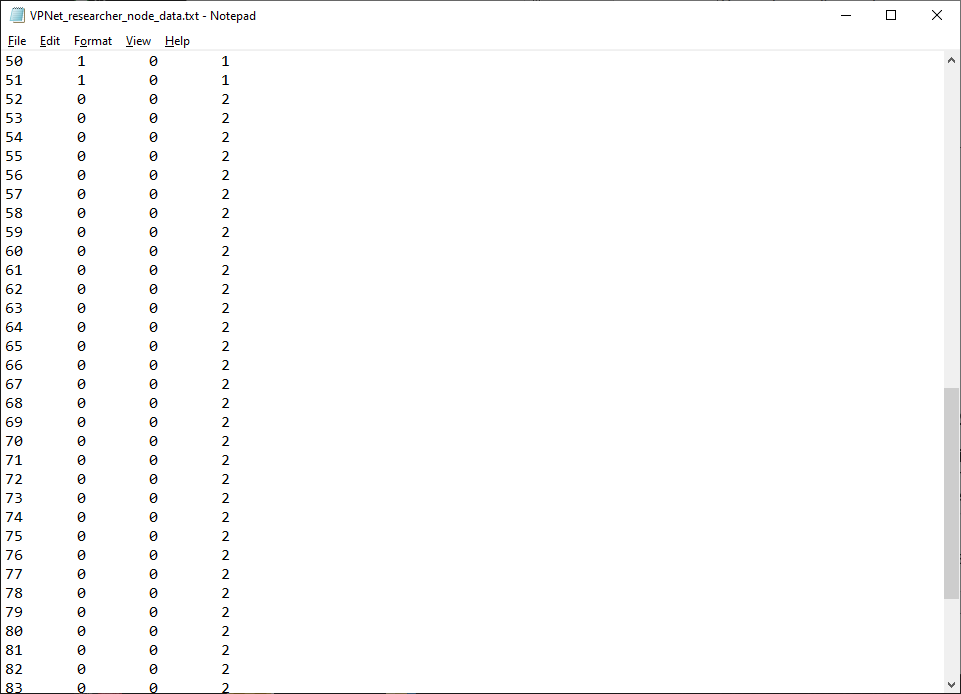
We need to assign the two sets of nodes to different modes or “Node\_Set”. We do that by assigning a node to either **Node\_Set = 1** for *Researchers,*and **Node\_Set = 2** for *Projects*. See below.



…

…

…

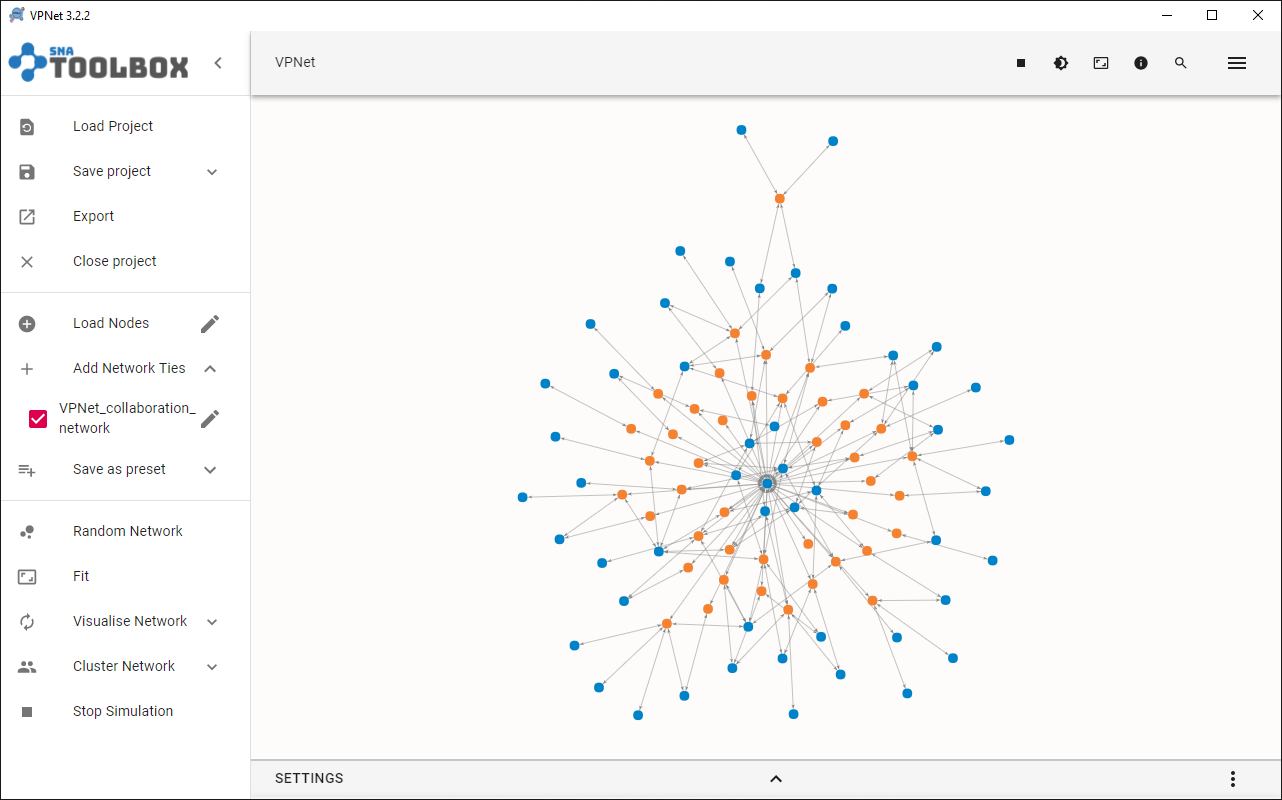


6.1.2 Visualising for different node sets

You need to load up your data in VPNet.

1. **Load nodes** > Researcher Collaborations> VPNet\_researcher\_node\_data.txt
2. **Add Network Ties** > Researcher Collaborations > VPNet\_collaboration\_network.txt

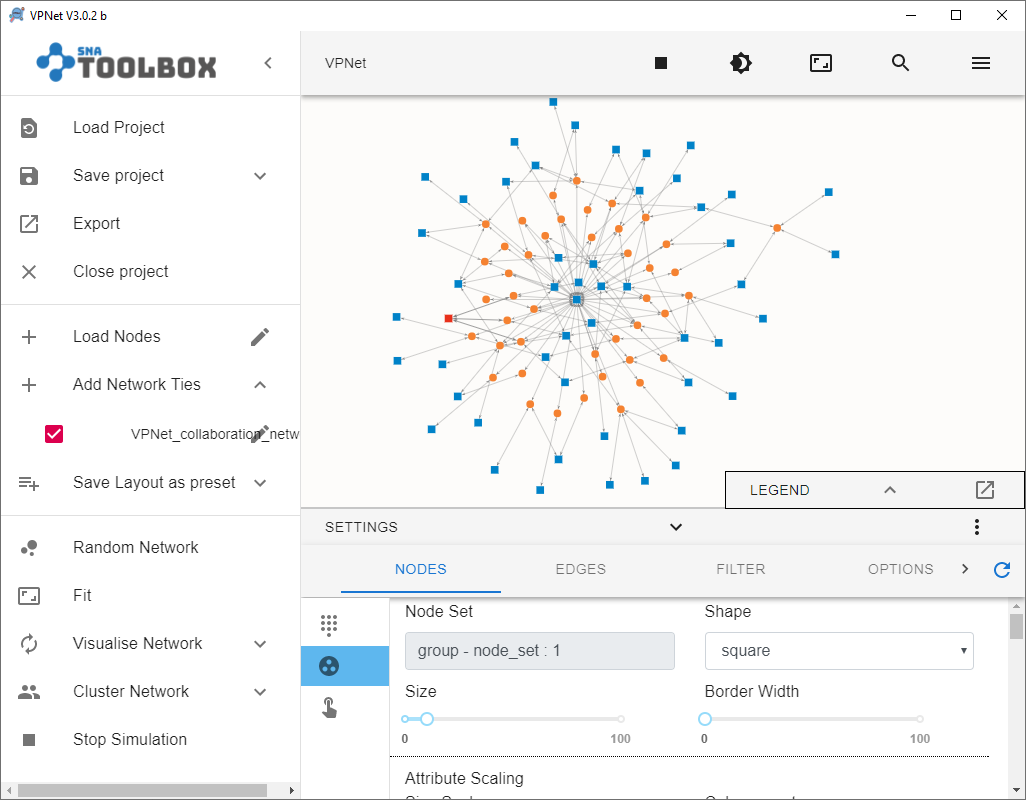
Once you have your data in, it should look like this.



Next, we want to make some changes, but only to one set of nodes.

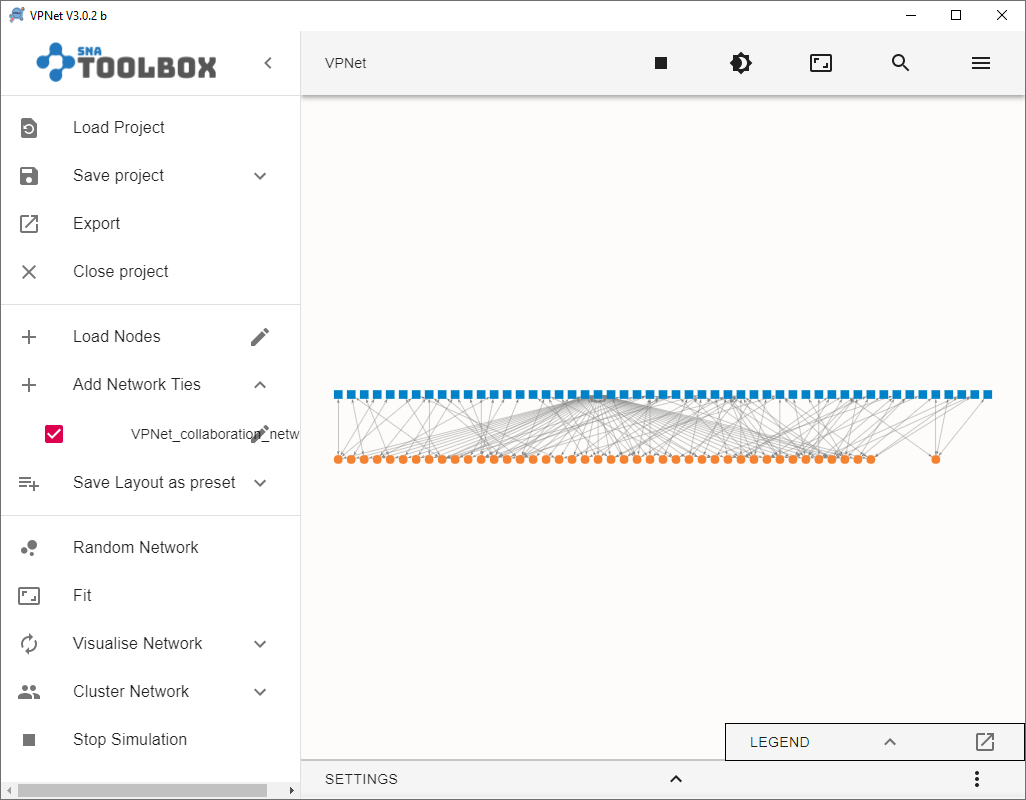
Click on a ‘**Settings > Nodes**’ and then select one of the blue nodes (or one set of colours). On the left, you can select ‘Nodes in the selected set’, you will get options for that specific set.

* Click on ‘**Shape**’ and select ‘**square**’.
* You will now notice that blue nodes are square!

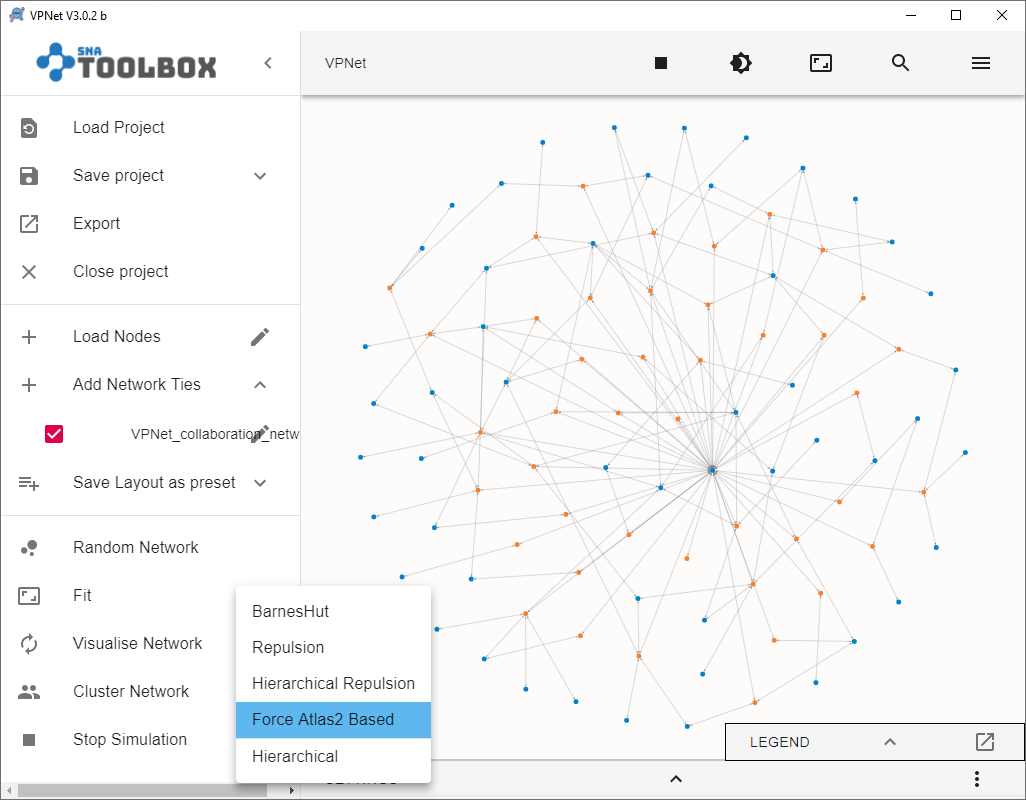


Under ‘**Visualise network**’, choose ‘**Hierarchical**’. It will give you something like this below, but just with two lines of nodes – all of your nodes separated into the two different node sets.

You can drag nodes to new positions to inspect the data, seeing which nodes are more connected.

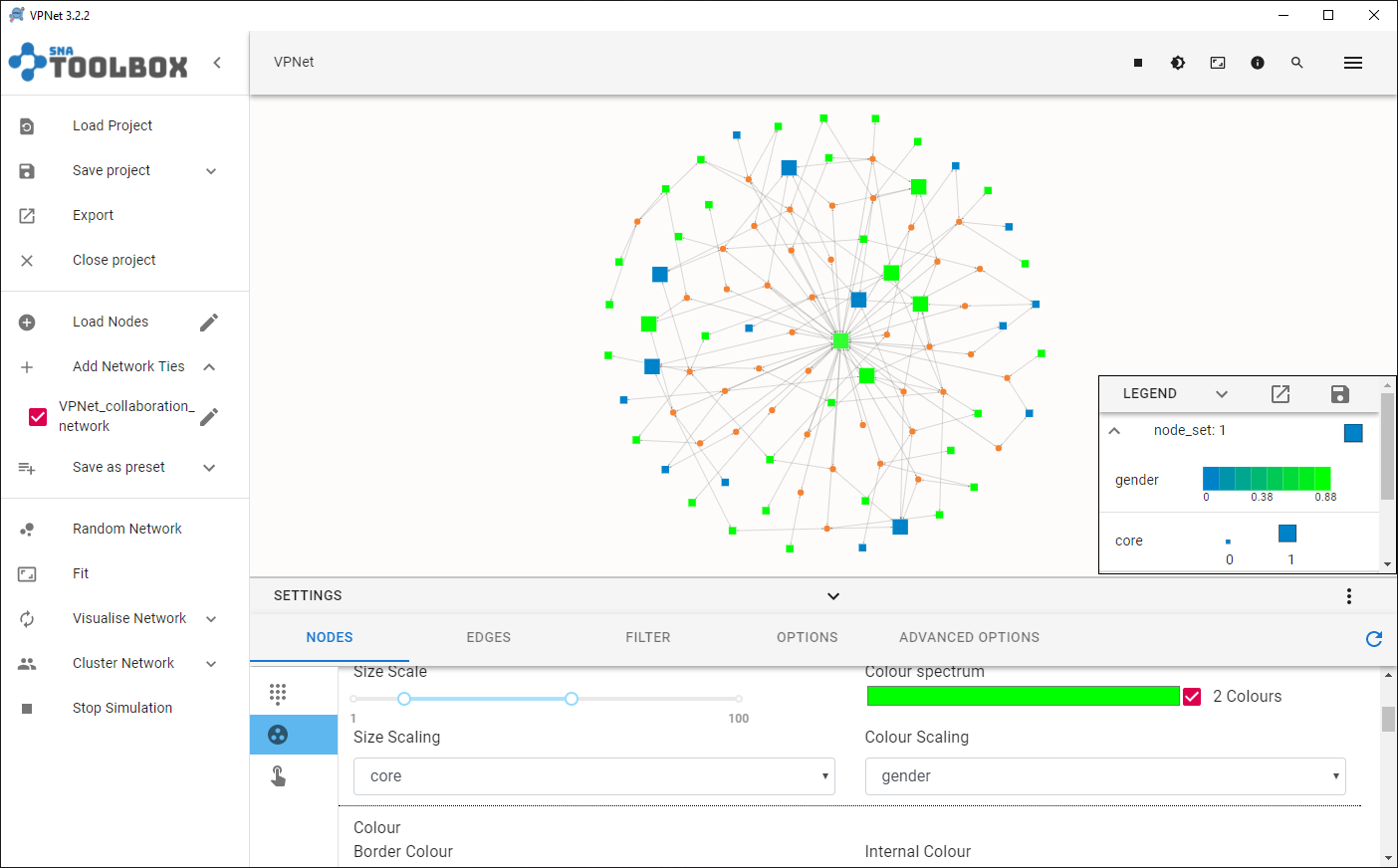


Of course, other layouts will give you other visualisations, and other insights. Using Forced Atlas2 Based we can see that there are few key researchers (blue nodes) that are connected to many projects (orange nodes).



You can also vary the size of nodes for one set of nodes. For example, we have done **Size Scaling** on the researchers (blue squares) by selecting the attribute ‘**Core**’ under **Colour Scaling** in the ‘core central geographic location’ (core), making larger squares for those in the core, and smaller for those outside it.

Further, we have also used **Colour Scaling** for gender, with blue nodes female, green for male. We have done all of this without making any changes to the projects (i.e., the orange nodes).



So what we can see is the ability to make completely different representations for the two node sets.

**NB:** In some instances, you might use the **Node\_Set** to help create different categorical attribute options. You are not restricted to just two sets of nodes, you could have many.

6.2 Multiplex networks

In this part of the exercise we want to show you how to visualize **multiplex** networks – that is, more than network on the same set of nodes. Let’s use data from *The Corporation*.

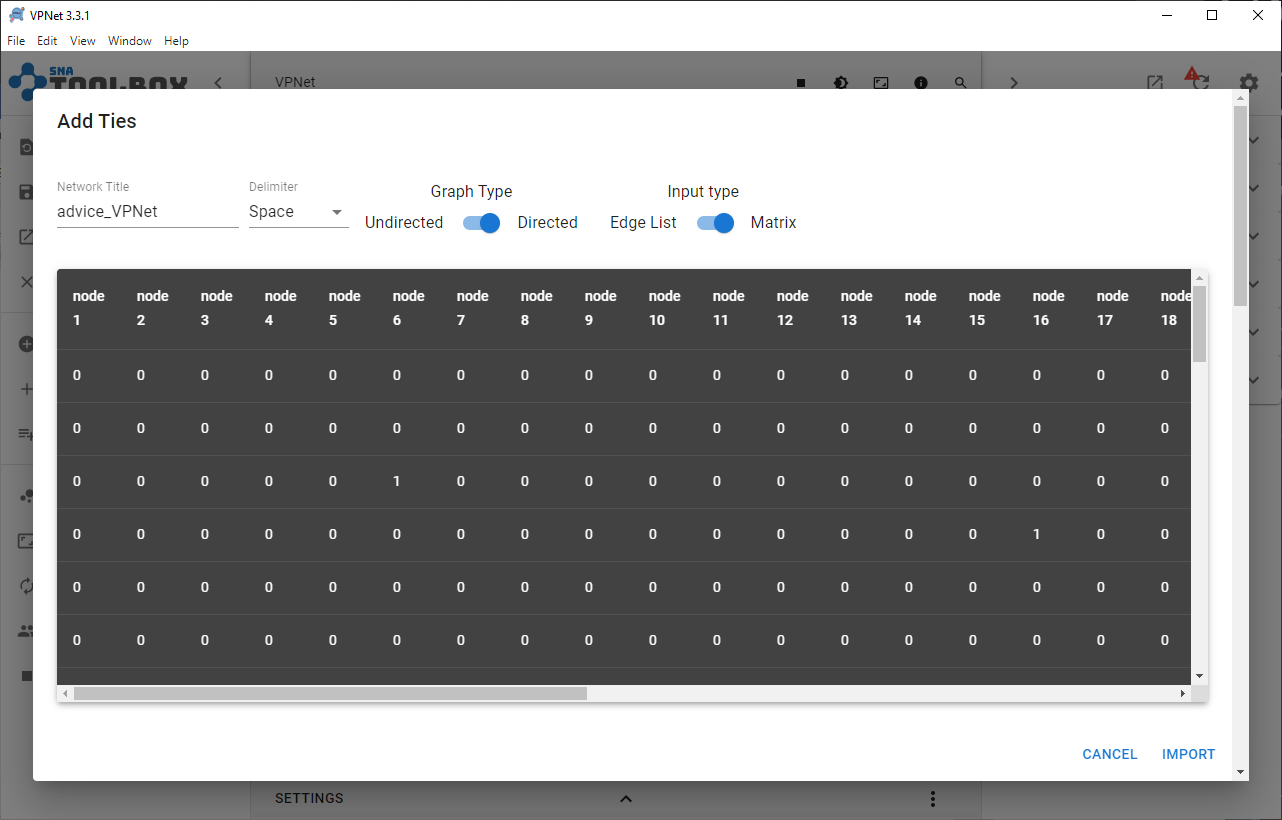
**Data:**  communication.txt

advice\_VPNet.txt

the\_corporation\_attributes.txt

In VPNet, add the nodes first: ‘**Load Nodes > the\_corporation\_attributes.txt’**

Then, select ‘**+Add Network Ties**’ and select the Advice network and import as below:



Then do the same and import the ‘communication.txt’ network by again clicking on ‘**+Add Network Ties**’.

You should now have two networks in your list:

A screenshot of a computer

Description automatically generated



If you click on the ‘tick’ to left of each network name, you can hide or show the network.

If you click on the pencil icon to the right of each network, you can edit properties specifically for that particular network. For the advice network, select red (or any colour you like) and also select ‘smooth enabled’. For the communication network, select a different colour and also smooth enabled. The *smooth enabled* function curves the lines.

A screenshot of a computer

Description automatically generated

You should have something like this below:

