

Learning Module - Attribute Editor and Histogram

As we covered in the learning module in Exercise 1, Attributes are the key pieces of information that provide context and information about the different elements on our graph, so they aren't just an expanse of pointless unknown interconnected nodes. To learn how to use attributes to answer questions we have of the data, this module will cover off on two views in Constellation:

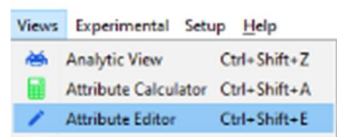
- Attribute Editor, used to display attributes of selected graph elements
- Histogram, used to count, and filter on attributes

We'll be using the Greg and Sarah Network again, so go ahead and open it if you haven't already. You might also find it handy to have the Constellation PDF user reference guide open to explain any terms that come up that you can't remember.



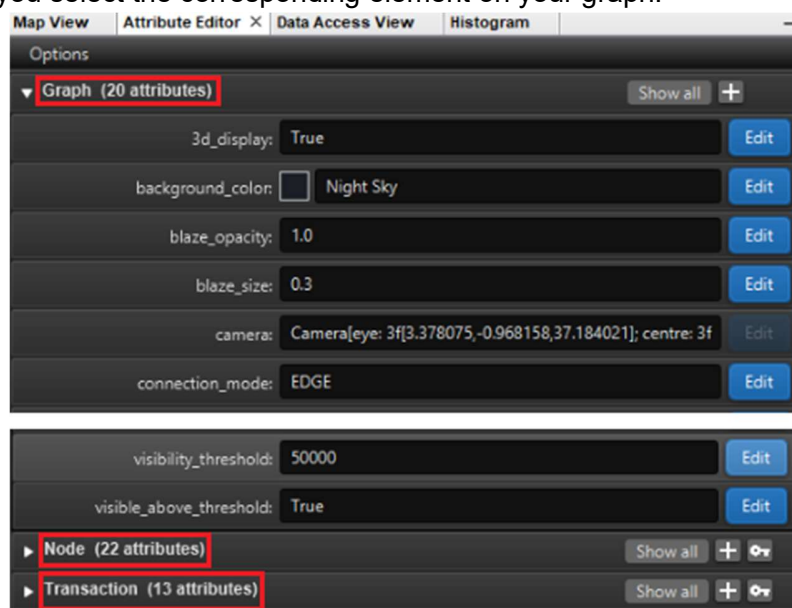
Attribute Editor

To open the Attribute Editor, click Views > Attribute Editor.

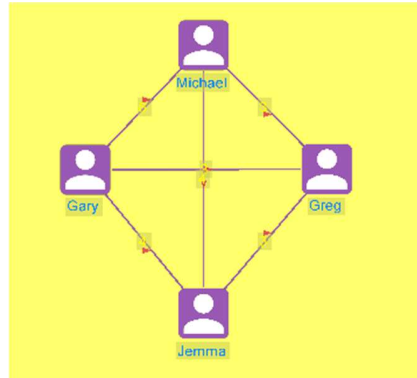


Graph Attributes

The default display when you open the Attribute Editor with nothing selected is Graph Attributes. If you look at the bottom of the view, you'll see Node and Transaction attributes collapsed; these will open when you select the corresponding element on your graph.



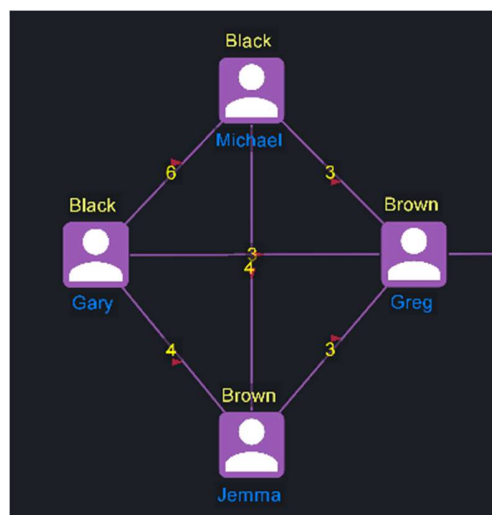
Graph Attributes describe the basic visual properties of the graph. You can see attributes like 'background_color' and 'highlight_color', which do what you think they would do. You can edit these by pressing the Edit button, and change them to your preferred colour. Let's change the background colour to 'banana', and see what that does to the graph.



As you can see, the background is now 'banana' coloured. This is awful, but thankfully like most things in Constellation you can undo your decision with Ctrl+Z. Please do so.

The main Graph Attributes you might want to change in future are 'node_labels_top' and 'node_labels_bottom'. These determine which attributes from nodes get visually displayed on the graph, and where. The default setting is for 'node_labels_bottom' to be set to 'Label', which means that whatever nodes have in their 'Label' attribute will be displayed on the graph. Our graph has the person's name in the Label attribute of each node, which is why that appears under their respective node.

Try changing the attribute 'node_labels_top' from blank to 'Hair Colour', and change the colour from blue to banana. The graph will update to add the new labels on top of your nodes, based on what each node has in that attribute field.



Depending on what kind of analysis you are doing, this can be a very useful tool to visualise certain node attributes that you don't want to look up individually each time.

Node Attributes

If you select a node on your graph, the Attribute Editor will automatically change focus to look at the attributes of the selected node. This is a great way to see all of the available attributes for a specific node on the graph.

A screenshot of the 'Node (21 attributes)' editor window. It displays a list of attributes for a selected node. The attributes are: Gender: Male, Hair colour: Brown, Height: 180, Identifier: Node 1, Label: Greg, Raw: Node 1<Person>, Source: Manually Created, and Type: Person. Each attribute has an 'Edit' button to its right. The 'Identifier' and 'Type' rows are highlighted in red.

Attribute	Value	Action
Gender	Male	Edit
Hair colour	Brown	Edit
Height	180	Edit
Identifier	Node 1	Edit
Label	Greg	Edit
Raw	Node 1<Person>	Edit
Source	Manually Created	Edit
Type	Person	Edit

If you select more than one node, you'll see all of the attributes for all of the nodes selected. If the values for the same attributes are the same across the different nodes, they will merge on top of each other. If there are different values for the same attribute across the different nodes, the attribute will collapse with a ▶ symbol appearing, and <Multiple Values> will be shown in the value field for the attribute. If you click that symbol, you will see the multiple values stored underneath it. Unfortunately, there is no way to select one of those attributes to identify the corresponding node on the graph in the Attribute Editor view, this functionality lies in the Histogram View.

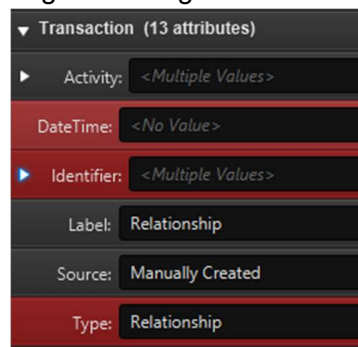
The image shows a graph on the left with three nodes: Michael (top), Gary (bottom left), and Greg (bottom right). They are connected by edges with weights 6, 3, and 4 respectively. To the right is the 'Node (21 attributes)' editor for the selected nodes. The 'Hair colour' attribute is collapsed, showing '<Multiple Values>' and a list of 'Black' and 'Brown'. The 'Height' and 'Identifier' attributes are also collapsed, showing '<Multiple Values>'. The 'Identifier' row is highlighted in red.

Attribute	Value	Action
Gender	Male	Edit
Hair colour	<Multiple Values> Black Brown	Edit
Height	<Multiple Values>	Edit
Identifier	<Multiple Values>	Edit

Transaction Attributes

Transaction attributes are viewed in the attribute editor in the same way as node attributes. As transactions by their nature are often stacked on top of each other, you are much more likely to see collapsed values. A common use for looking at transaction attributes in the attribute editor is to quickly look at attributes like 'Type', 'Activity', and 'DateTime', all of which give you a good overview of what the nature of the relationship between the nodes on the graph is. While helpful

as an overview, it isn't useful for analysis, because you can't make selections from the attribute editor; we'll learn how to do this using the Histogram later in this module.



Transaction (13 attributes)

Activity: <Multiple Values>

DateTime: <No Value>

Identifier: <Multiple Values>


Label: Relationship

Source: Manually Created

Type: Relationship

Creating new Attributes

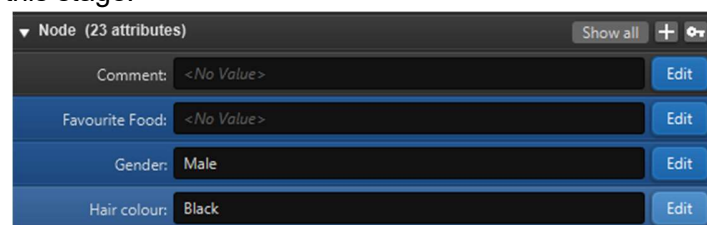
Occasionally you will want to add your own attributes to the data, if they didn't exist already. Most commonly, adding an additional label or comment for manually filled values to indicate something

about your analysis on the graph. To create a new attribute for nodes, click the  button at the top right corner of the Attribute Editor in the Node Tab (you can do the same for Graph and Transaction Attributes). A menu will open to prompt you to choose from a range of existing attributes that Constellation knows about, or to give you the option to make your own Custom attribute.

In your graph, create two new attributes:

- Under the 'Analysis' category, choose 'Comment'
- Make a Custom attribute, name it 'Favourite Food', and with the Type set to 'string' indicating that constellation is expecting text for this attribute.

The attribute editor will update to show the new attributes on your nodes, although they won't have any values at this stage.



Node (23 attributes) Show all +

Comment: <No Value> Edit

Favourite Food: <No Value> Edit

Gender: Male Edit

Hair colour: Black Edit

We can then use the Edit button to add values. On the Greg node, edit 'Comment' to have the value 'Head of Synister Lunches inc.' Edit the 'Favourite Food' attribute to have the value 'banana'. We can now see these values if you select Greg and look at the attribute editor.

If we want to visually see the values you have input on the graph, we need to Edit the Graph properties. Edit the Graph attribute of 'node_labels_top', and add 'Comment' and 'Favourite Food'

in whichever colour you prefer. We can now see those values above the nodes on the graph, where they exist.



This could be useful for saving conclusions from analysis into attributes on a node for future reference. For example, if you found that Greg only eats at vegetarian restaurants you may want to flag him as a Vegetarian and then conduct future analysis on just vegetarians.


Looking at the attributes of a single or small group of nodes or transactions at a time using the attribute editor is not an efficient method for analysis, particularly as you start working on larger graphs. To do better, we need to use the Histogram.

Histogram

The Histogram counts attributes, be they node attributes, transaction attributes, or graph properties (if you can't remember graph properties, read the PDF guide or go back to the learning module in Exercise 1). We use it to count, and then select on, particular attributes that we are interested in. Often, you'll need to use the Filter function to drill-down on a subset of attributes to combine multiple questions together.

Histogramming Attributes

On the Greg and Sarah Network graph, there are a number of attributes on the nodes that can be used to group them together or distinguish them apart, including Name (the Label attribute), Gender, Hair Colour, and Height.

Open the Histogram with Views > Histogram, or by clicking the  button in the toolbar. You'll see the Histogram tools down the bottom of the view; with the red boxed tools below you'll likely use the most:

Graph Element: Node Transaction Edge Link

Category: Attribute

Property:

Sort: Label ▲

Selection Mode: Free Selection Select

Filter: Filter On Selection Clear Filter Actions

- The 'Graph Element' lets us choose whether to look at the attributes of Nodes, or of Transactions.
- The 'Category' lets us choose whether we are looking at Attributes, or Graph Properties
- The 'Property' lets us choose which specific attribute to count
- Ignore 'Sort' and 'Selection Mode'
- 'Filter' lets us... filter the Histogram on our current selection, meaning it will now only count attributes that are included in that filter.

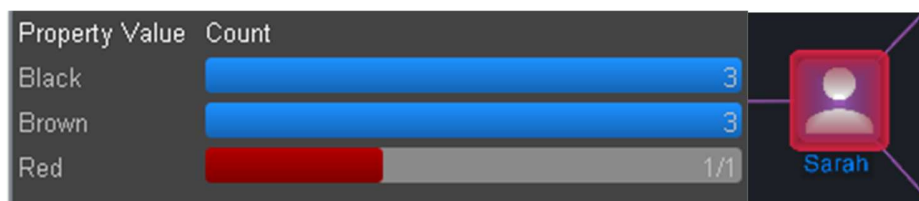
Let's count the number of different hair colours on the graph. Select Graph Element: Node (hair colour is an attribute of nodes, not transactions), Category: Attribute (it is an attribute, not a graph property), Property: Hair Colour.

Graph Element: Node Transaction

Category: Attribute

Property: Hair colour

The Histogram will update to show the attribute being counted in the left column, with the number of nodes with that attribute in the right column.



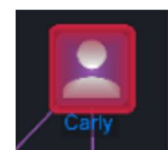
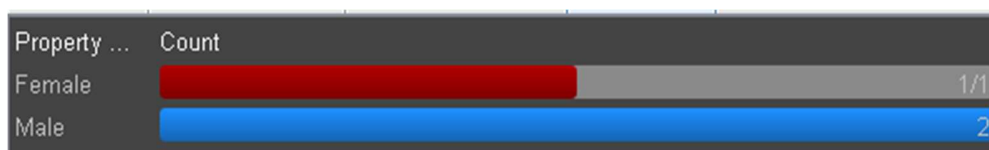
As we can see, there are three people with black hair, three with brown hair, and one person with red hair. If we click the blue bar next to 'red', Constellation will select the node that has that attribute on the graph.

What if we wanted to combine two questions together? For example: find a female with black hair. To do this, we need to use Filtering.

Starting with the attribute of Hair Colour, click 'black' to select all people on the graph with black hair. Click the Filter on Selection button to... filter on to this selection. The Histogram will now only show the attributes of people with black hair. As the Histogram is currently looking at the Hair Colour attribute, and the only attributes we are counting are of those people with black hair, the Histogram should only show 'black', with a count of three.

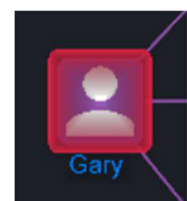


Change the Property to Gender. The Histogram will update to show the attribute counts for the different Gender values for people with black hair. We can now click on the bar next to Female, to select the one female with black hair: Carly.



Let's use this technique to answer another question: which male is the tallest? The first thing you'll need to do is click the Clear Filter button, otherwise you'll only be counting the attributes of people with black hair again.

Set the Property to Gender, and select Male. Filter on this selection so we are only counting the attributes of males. Now change the Property to Height. The Histogram will now show the values for the Height attribute for every male on the graph. You can then select the highest value (187) to select the node with that Height value on the graph: Gary.



Using the Histogram to select, filter, select again, and filter again, is one of the core techniques you will use to conduct analysis in Constellation, so try to master it. If you are analysing a much larger graph, it can be useful to do your first selection and then use Ctrl+U to move that selection to a new graph. This makes follow-on filtering and selection much easier.

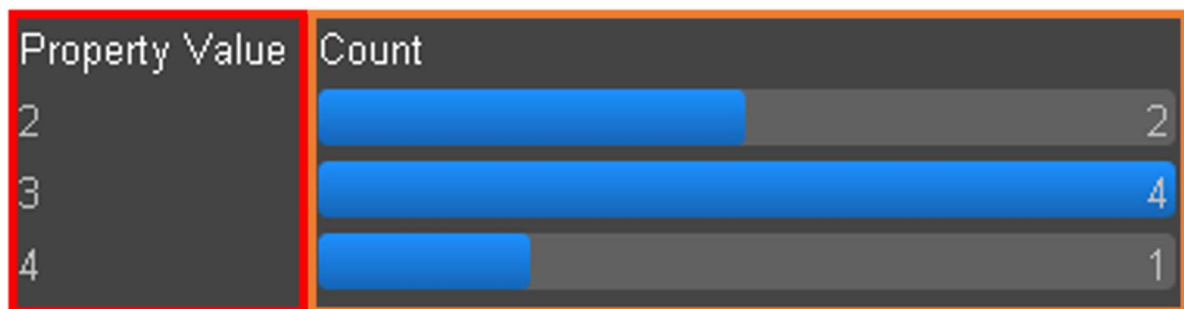
Histogramming Graph Properties

There are also common questions that arise when doing graph analysis that relate to the Graph Properties, rather than the Attributes. Re-read the material in Exercise 1, or the PDF document, if you want more background. The common questions are:

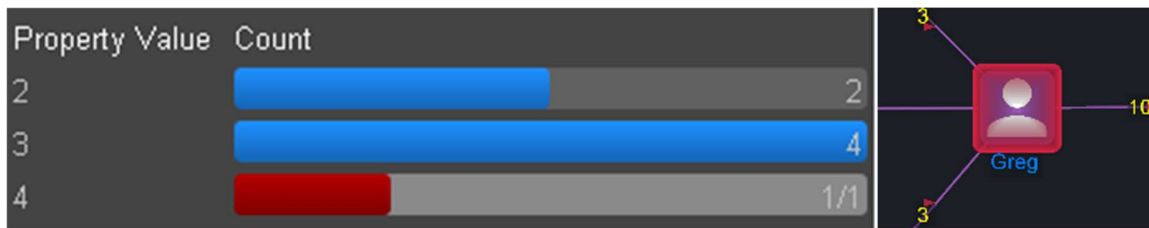
- Which node is connected to the highest amount of other nodes (in graph terminology, which node has the most 'neighbours')?
- Which node has the highest number of transactions?

These are both graph property questions because they refer to the underlying structure of the graph, rather than the attributes. You could strip all of the attributes off the nodes and transactions, and you could still do the analysis.

To answer the first question, open the Histogram. Select the options Graph Element: Node, Category: Graph Property (not Attribute, as we have done previously), Property: Neighbour Count. The Neighbour Count and Transaction Count properties are the two most prevalent options you'll choose in this menu. Selecting Neighbour Count will make the Histogram count how many neighbours each node on the graph has.

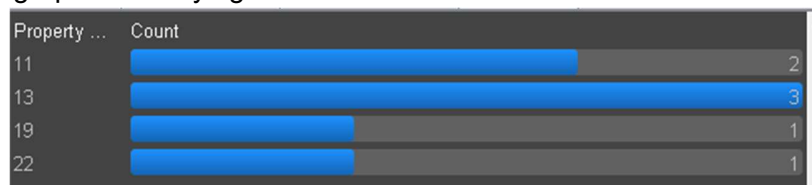


When counting things with numerical values the Histogram can sometimes appear confusing. Just remember the column headings - the property being counted is on the left (red box); the frequency of that property is on the right (orange box). We can see in the Histogram table now that there are two nodes with two neighbours, four nodes with 3 neighbours, and one node with four neighbours. We can then select the node with the count of 'four' on the histogram to find it on the graph: in this case, Greg.

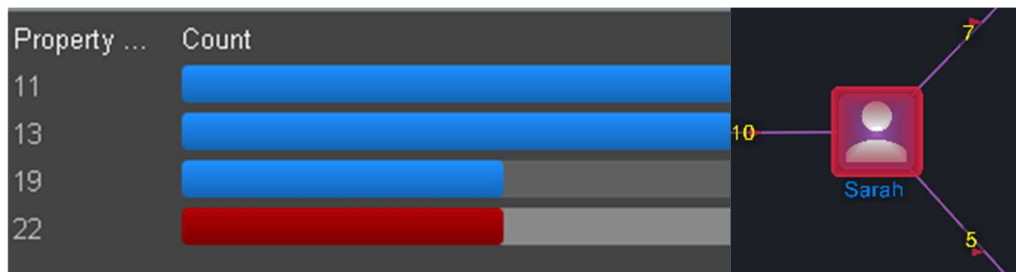


On our small graph you can also do this easily with your eyes, but on graphs with thousands of nodes that is much more difficult.

Answering the question about the node with the highest number of transactions (who had the most lunches) is already difficult to do with your eyes even on our small graph, so you can easily imagine the impossibility of doing this visually on a larger graph. To answer this question, change the Histogram Property from Neighbour Count to Transaction Count. This will show the number of nodes on the graph with varying levels of transactions connected to them.



We can then select the node with the highest transaction count, 22, to find that node on the graph. In this case, the person going to the most lunches was Sarah, even though she had fewer people (neighbours) she was going to lunch with than Greg.



Hopefully after this module you will have the requisite knowledge and skills to answer many of the questions in the coming Exercises.

Common 'gotcha' issues when using the Histogram:

- Not clearing the filter when you are finishing using it.
- Not selecting the correct graph element: is the attribute you want to count an attribute of a Node, or a Transaction?
- Histogramming on Attributes, rather than Graph Properties, or vice versa.
- Looking at the wrong side in the Histogram table to get counts, especially when counting numerical values.