UCL Depthmap 7: Data Analysis

Alasdair Turner

Version 7.12.00c

Outline

Introduction

Getting started
Preparing VGA and axial analyses
Importing a MIF/MID data file

Entering observation data
Adding a column
Using the table window
Entering data

Summary statistics

Scatter plots
Simple scatter plots
Pushing values to layers
Cross map comparisons

Exporting data

Conclusion

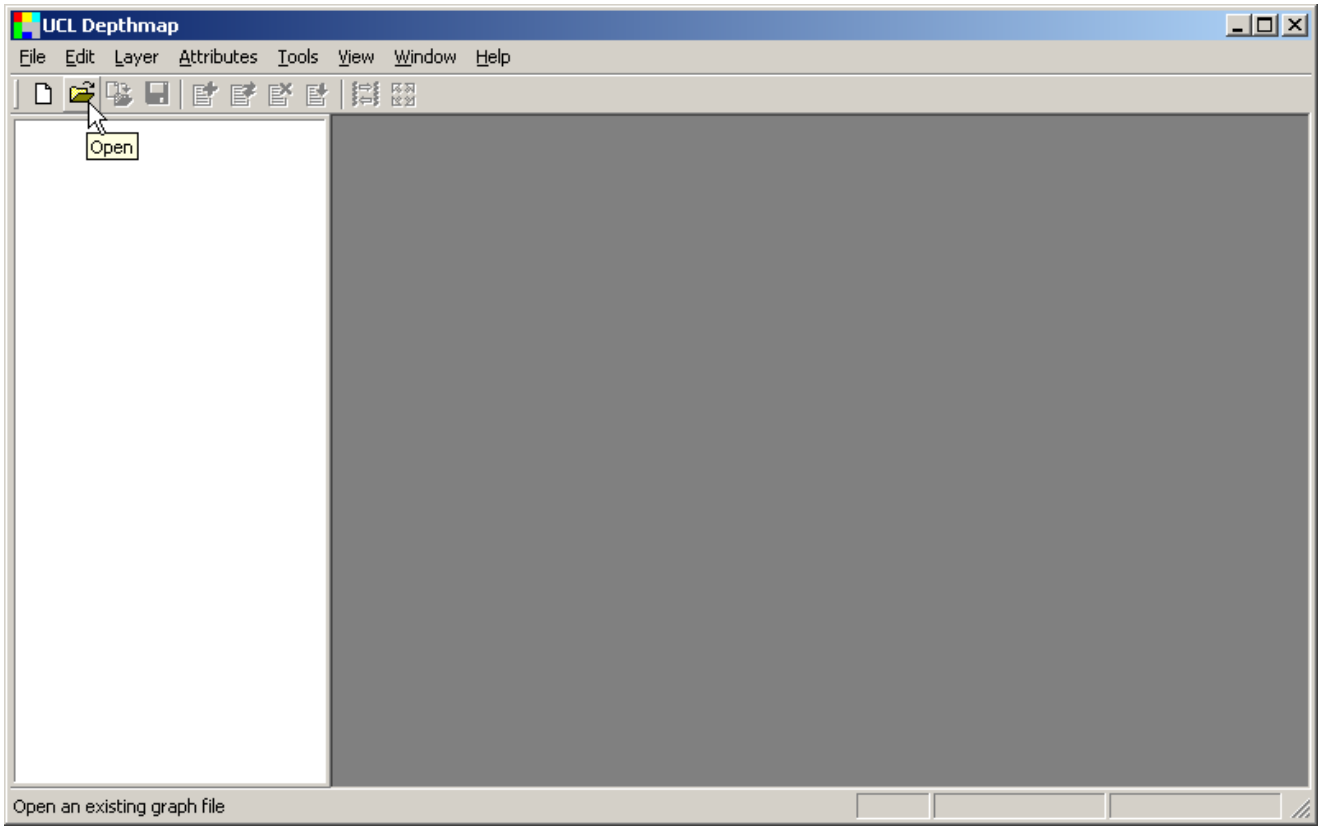
Data analysis in Depthmap

Although Depthmap is primarily a graph analysis tool, it does allow you to investigate data that you produce.
This tutorial will show you how to look at summary statistics about attributes, compare attributes with each other through scatter plots, and compare attributes to observation data that you have collated.
If you want to do more detailed data analysis than are described here, you may also want to export data from Depthmap, and so a basic text file export is described in this tutorial.
The tutorial assumes that you can already produce visibility graph analyses and axial map analyses.

Getting started

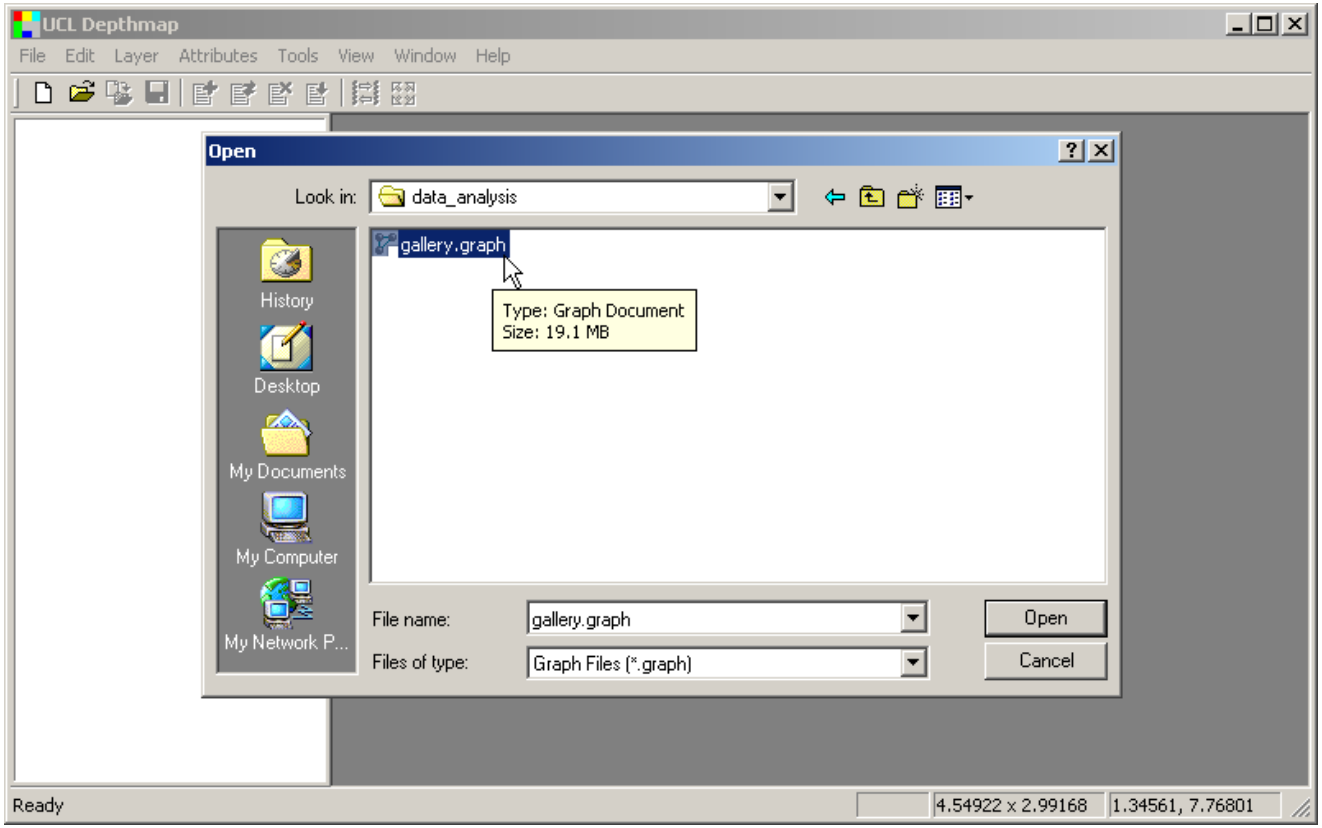
In this section, we load some data into Depthmap ready for analysis. Pointers on how to display Depthmap layers and how to import MIF/MID files from MapInfo are given.

Open file



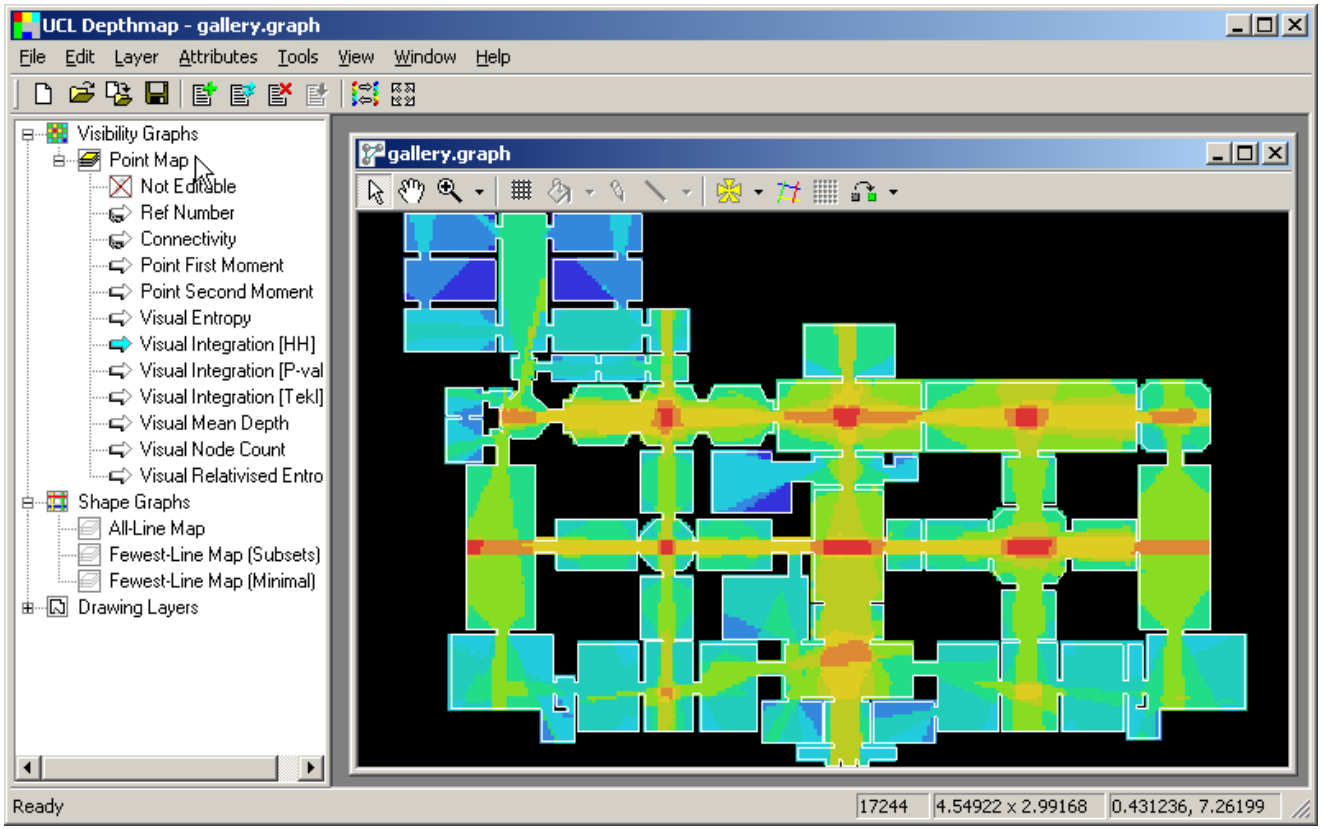
We will start by opening a graph previously prepared which includes both a VGA map and an axial map. To open a saved file, select ‘Open’ from the ‘File’ menu or click the ‘Open’ icon on the main tool bar.

Open file



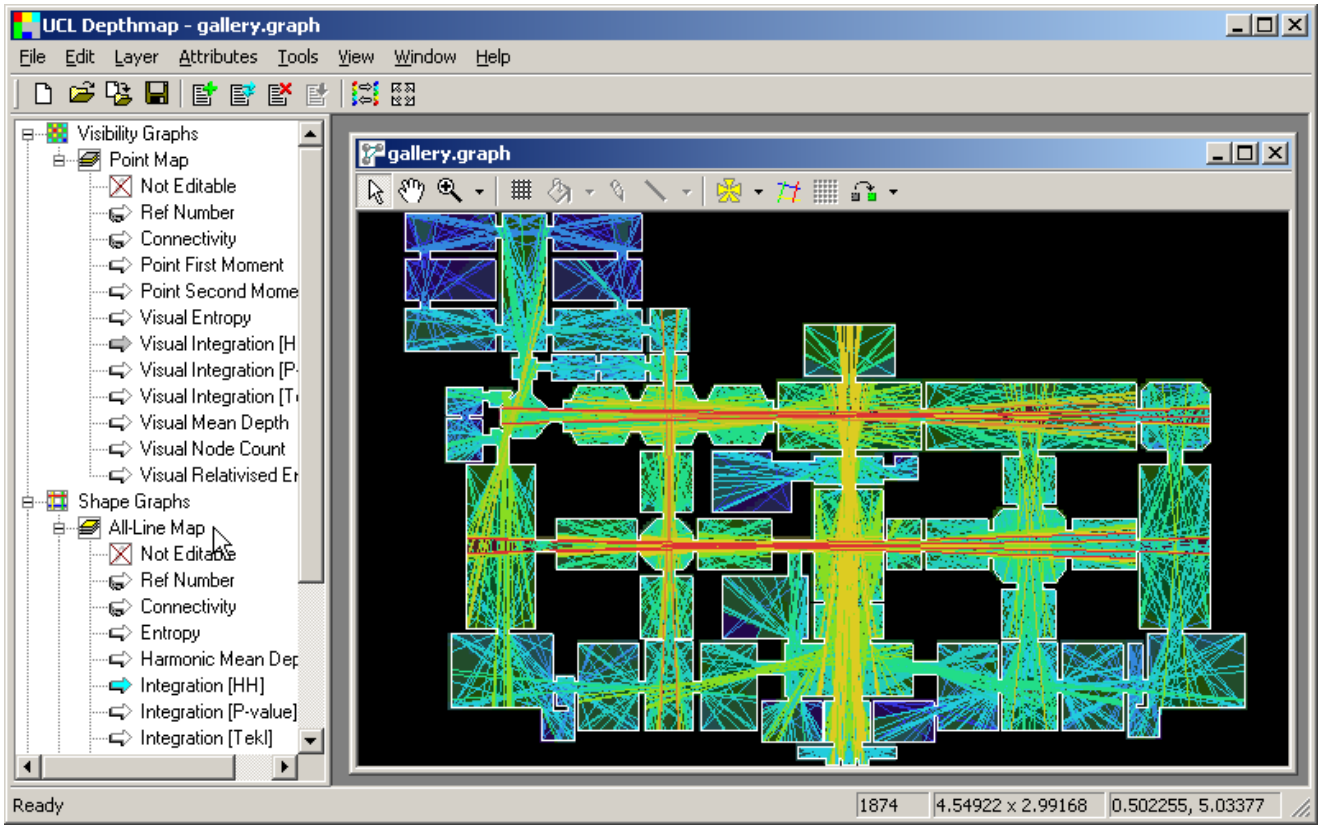
We will open the `gallery.graph` graph file, which contains both VGA and axial analysis of a gallery layout.
This file is available from the tutorials folder on the Depthmap website as one of the files included in `gallery.zip`.

Open layers



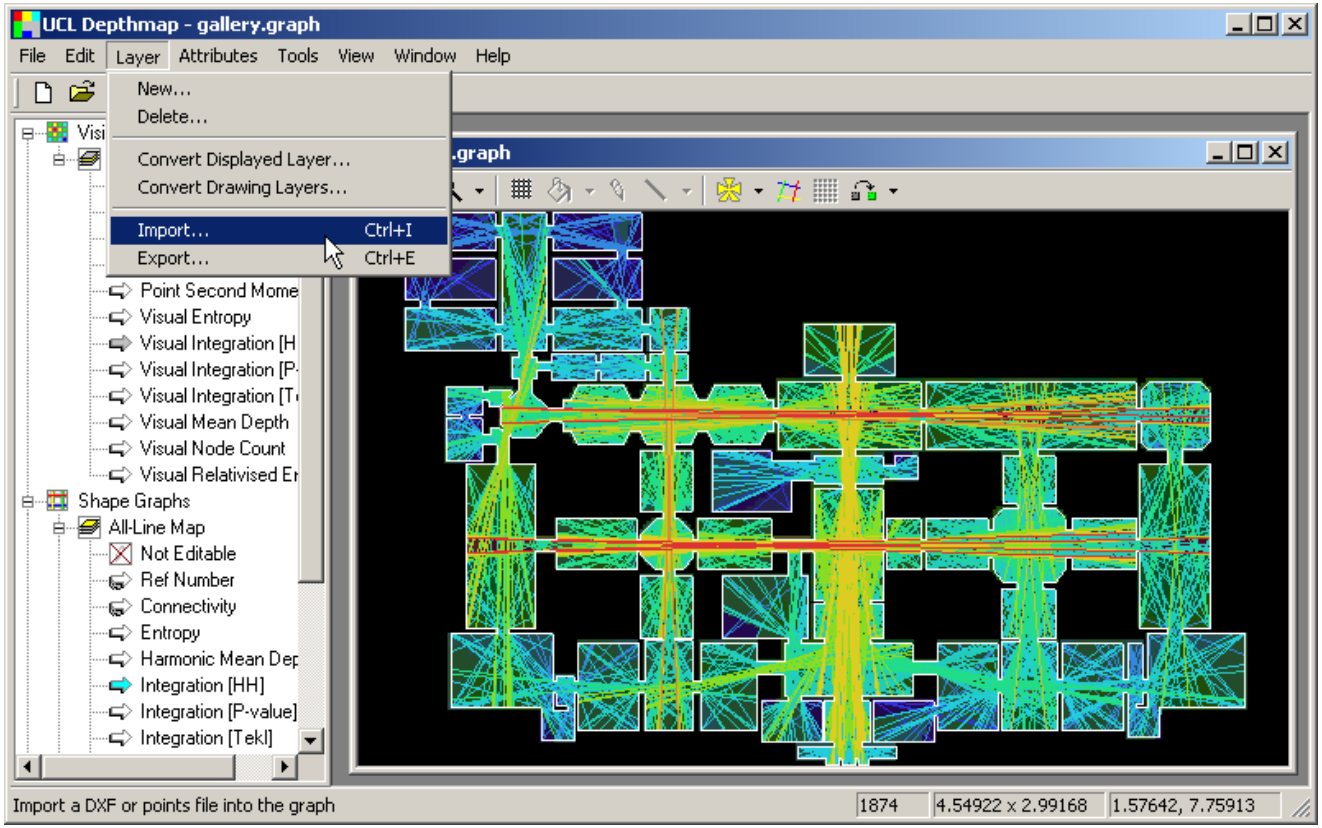
The graph file displays the visibility graph layer by default.

Open layers



Clicking on a second layer of a different graph type will promote the new layer to the topmost layer, and place the original layer behind it. You can hide the topmost layer by clicking on it again.

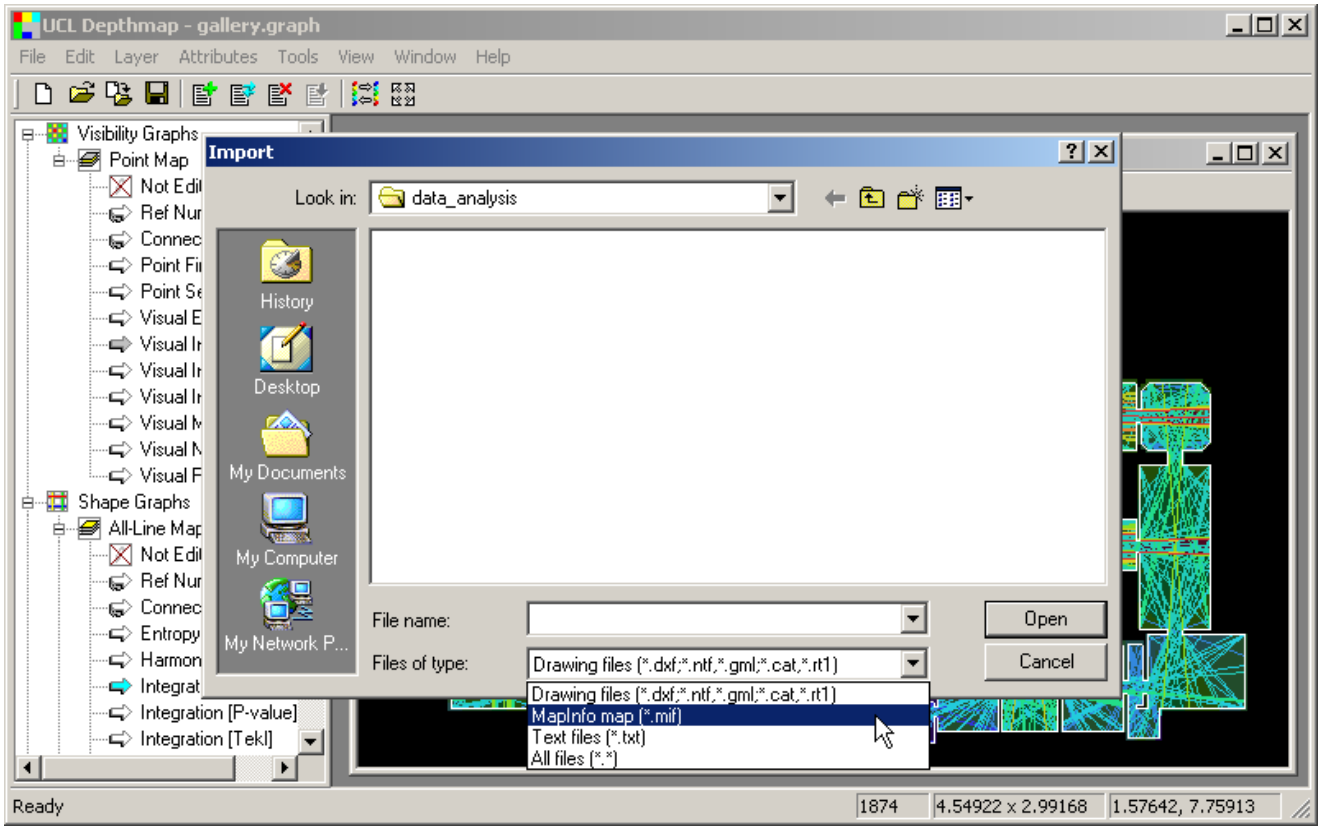
Import data layer



In addition to the graph information, we will import MapInfo data from a pair of MIF and MID files.

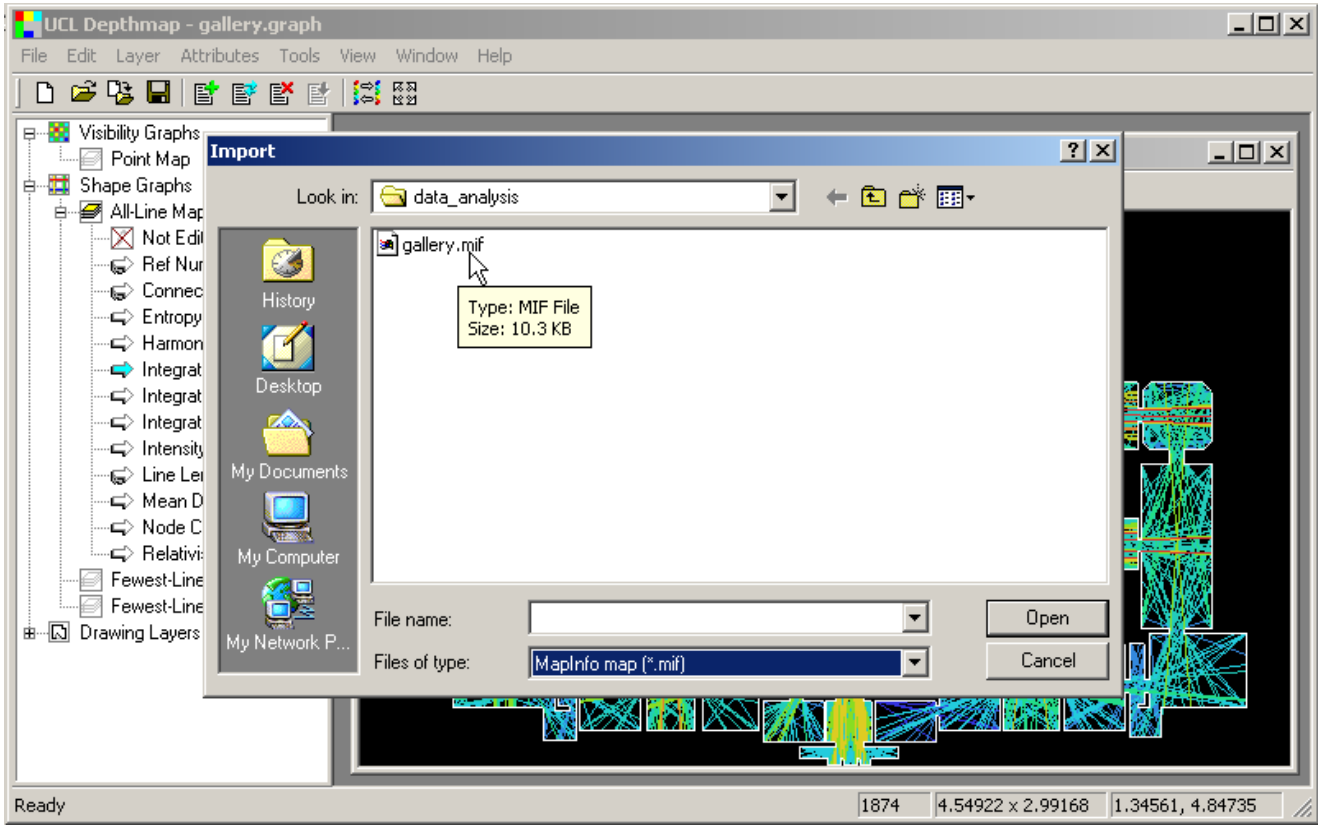
Use ‘Import’ from the ‘Layer’ menu, or click the ‘Import’ tool from the main tool bar, to import the files.

Import data layer



Change the import file type to ‘mif’ using the drop down menu.

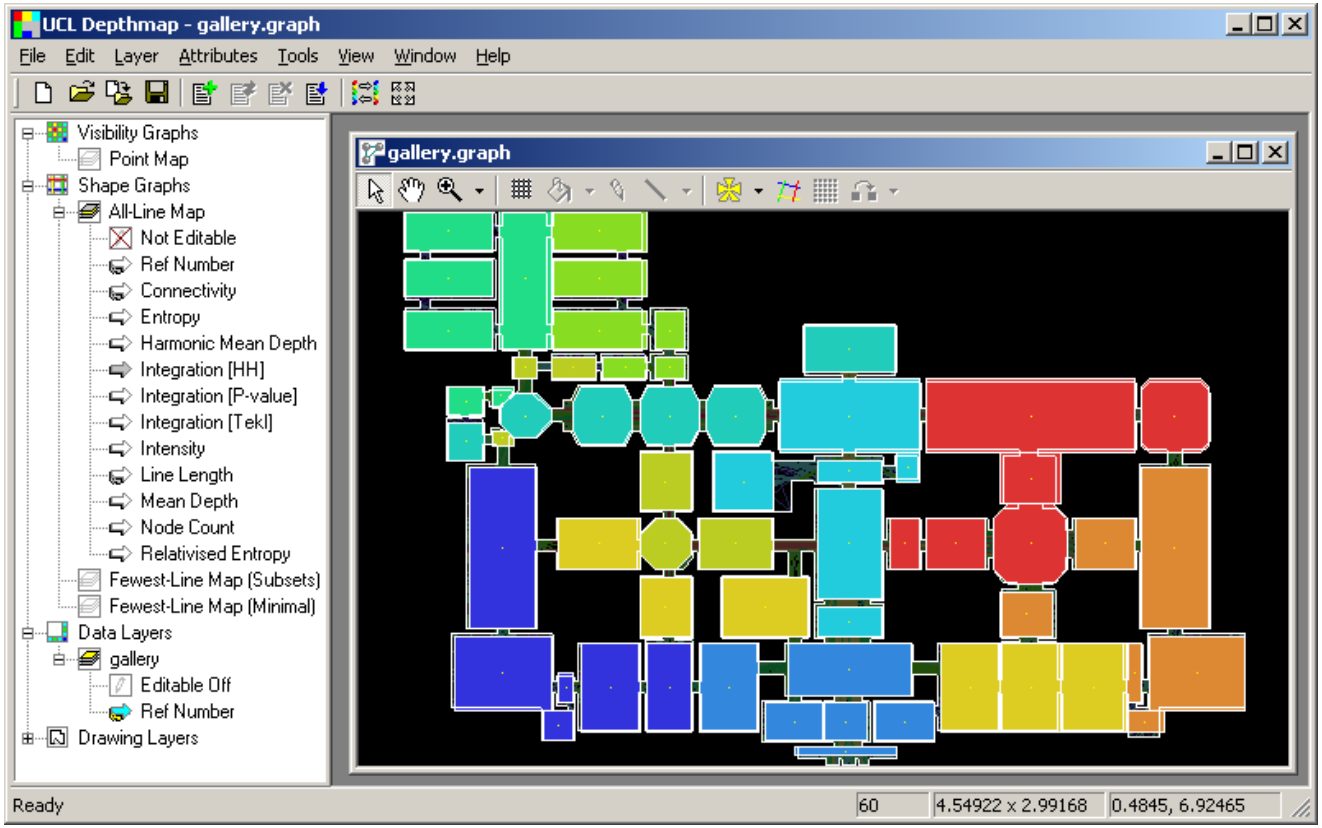
Import data layer



We will import the pair of files `gallery.mif` and `gallery.mid`. These files are available from the tutorials folder on the Depthmap website as two of the files included in `gallery.zip`.

Note that *both* files need to be in the same folder, even though we open them by selecting `gallery.mif`.

Import data layer



We will use the polygons in these files for a basis to enter some dummy observation data.

However, note that you do not need to use MapInfo to create them. The polygons in `gallery.mif` and `gallery.mid` were drawn in Depthmap. See the Convex Space Analysis tutorial for information about how to draw your own polygon data.

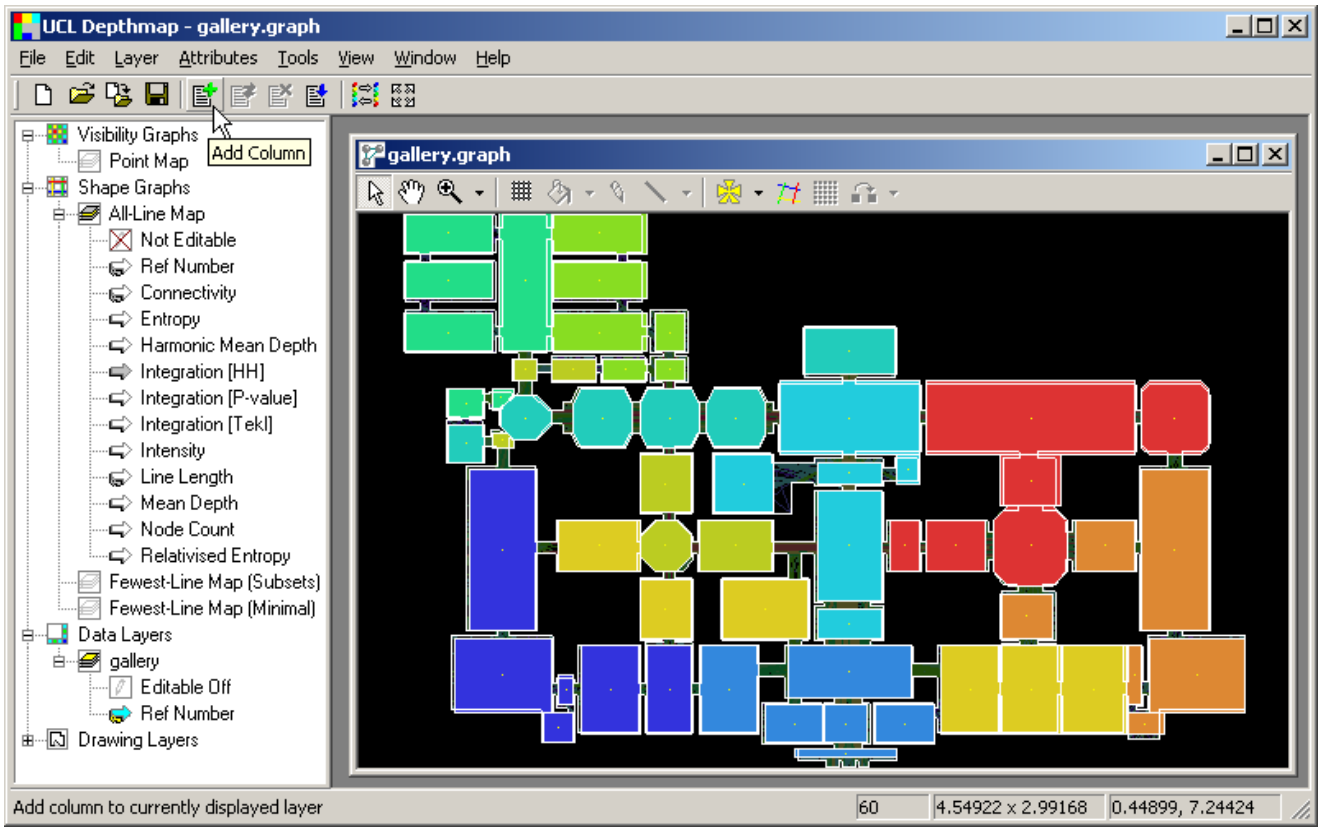
Entering observation data

You might import observation data from other software — for example, from MapInfo — but it is also fairly easy to enter observation data into Depthmap directly.

For the purposes of this section, I will assume that we have *room through movement* data. That is, the number of people passing into and out of a room during a period of time. We will mark the through movement on each room.

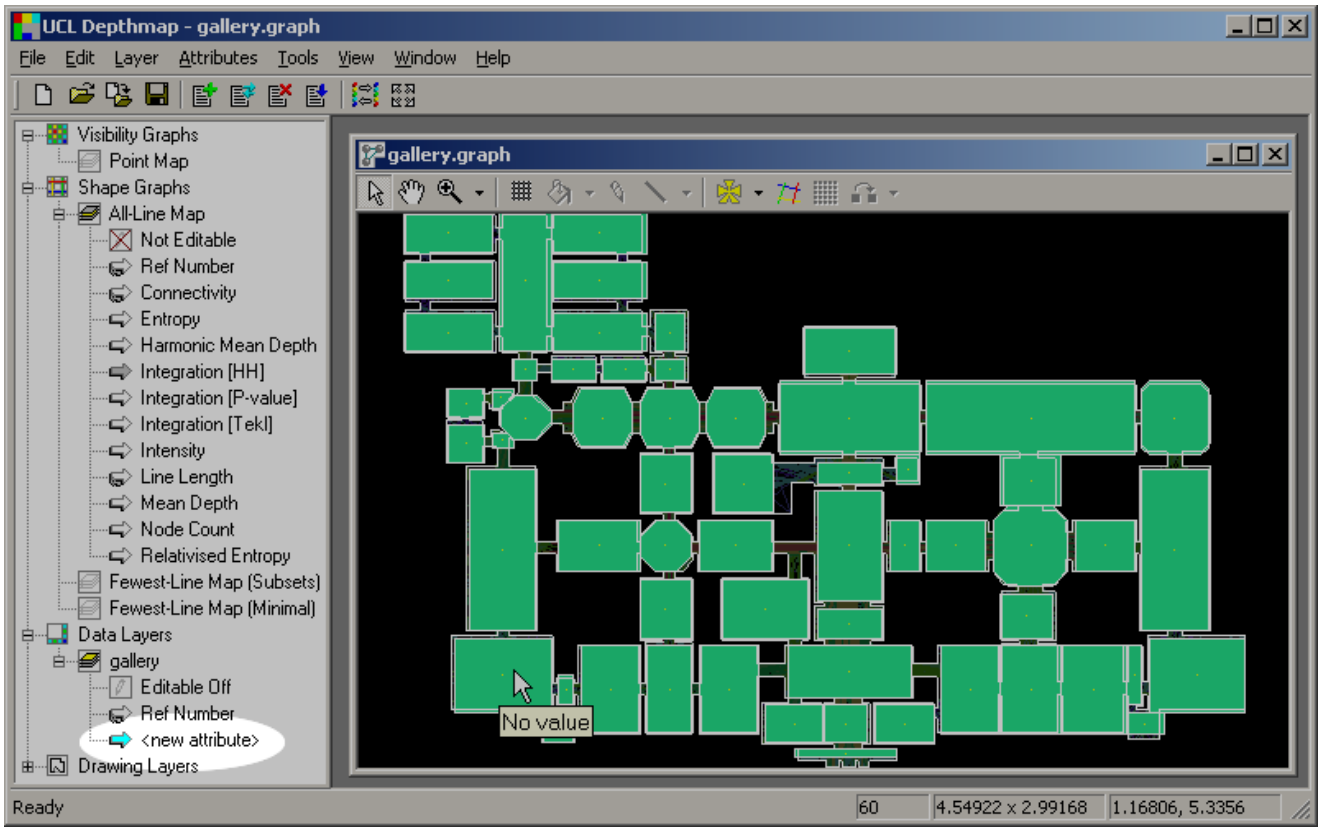
Note that typically, in space syntax analysis, researchers will be entering gate counts rather than through movement. That is, the number of pedestrians crossing a notional line across a pavement, door threshold or sidewalk.

Add a new column



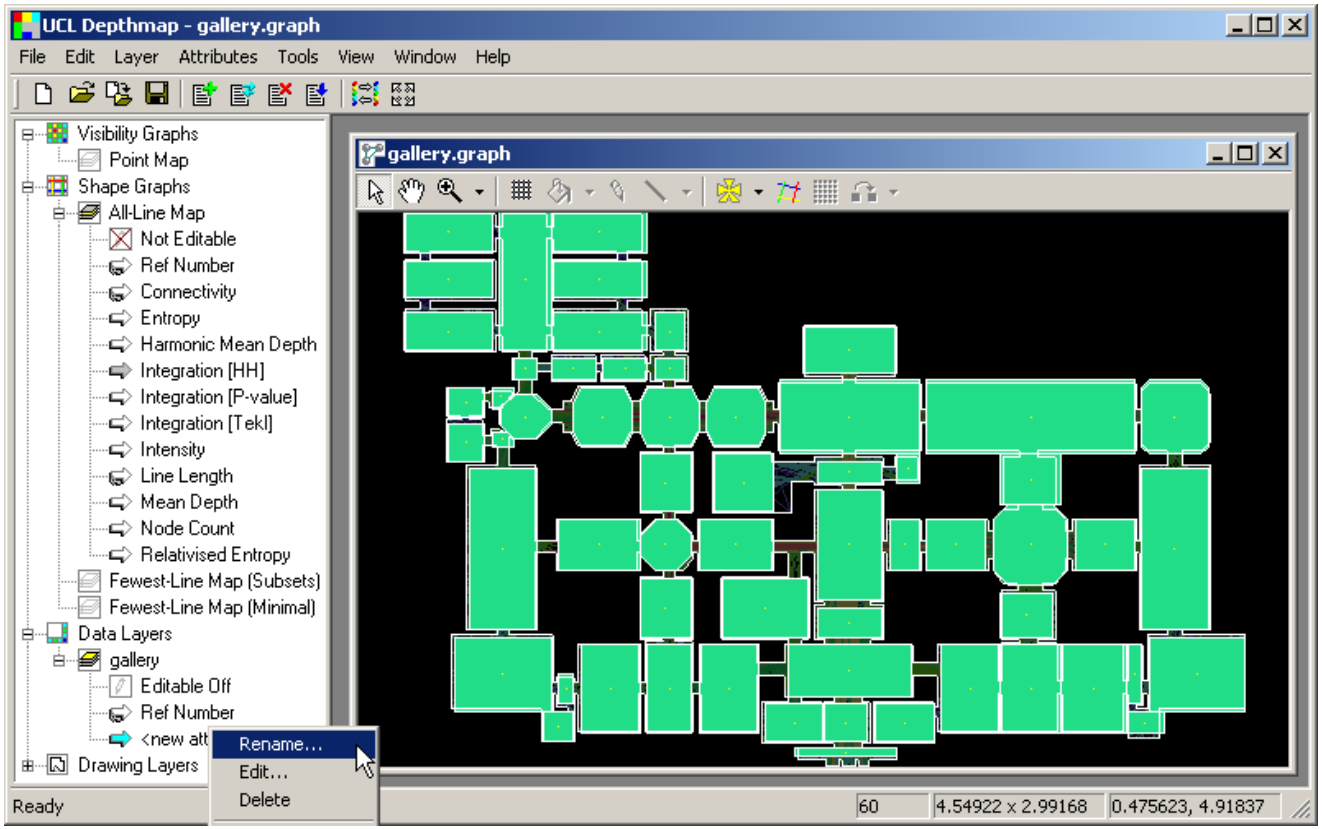
To begin we need a column into which to enter the data. With the data layer uppermost, click on ‘Add Column’ from the main tool bar, or choose ‘Add Column’ from the ‘Attributes’ menu.

Add a new column



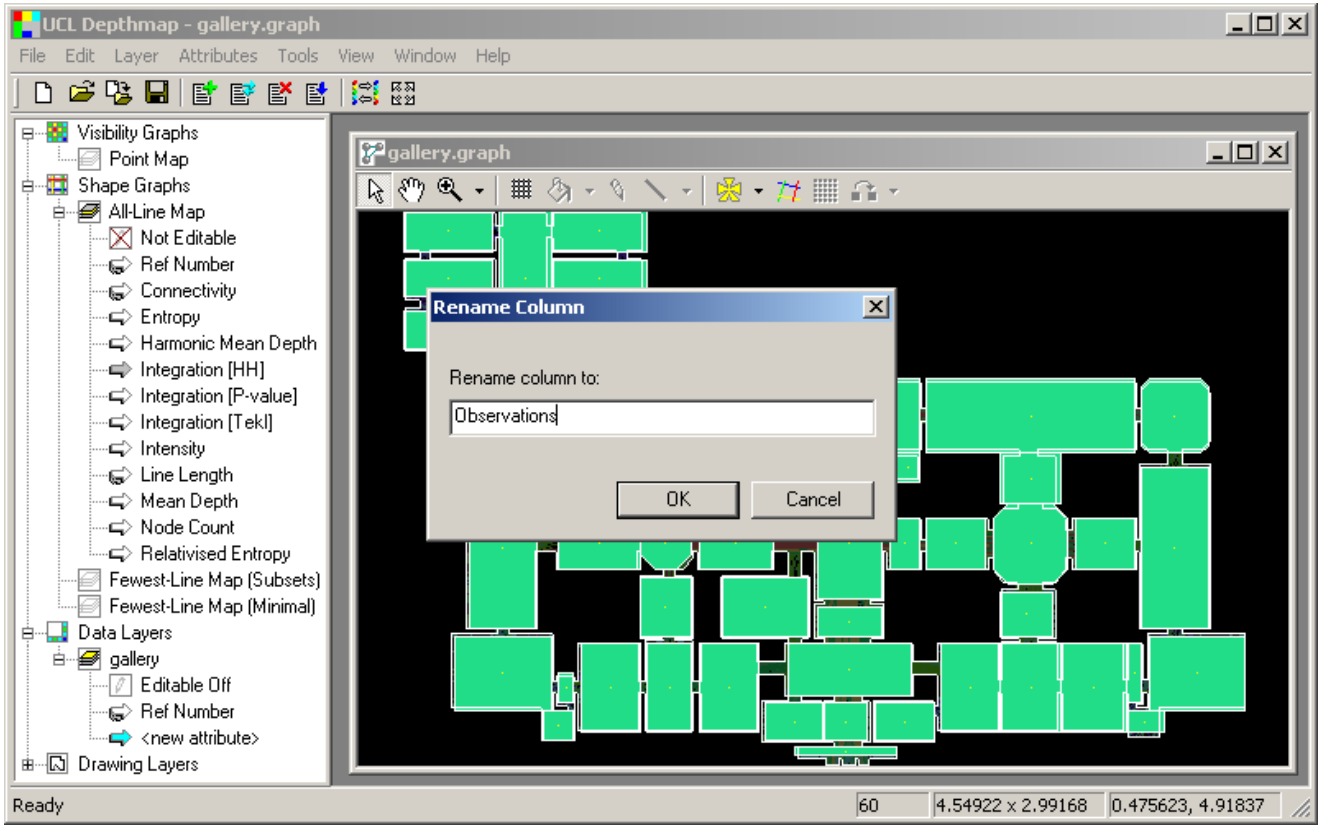
A new column, called `<new attribute>` is added, with no values inserted.

Add a new column



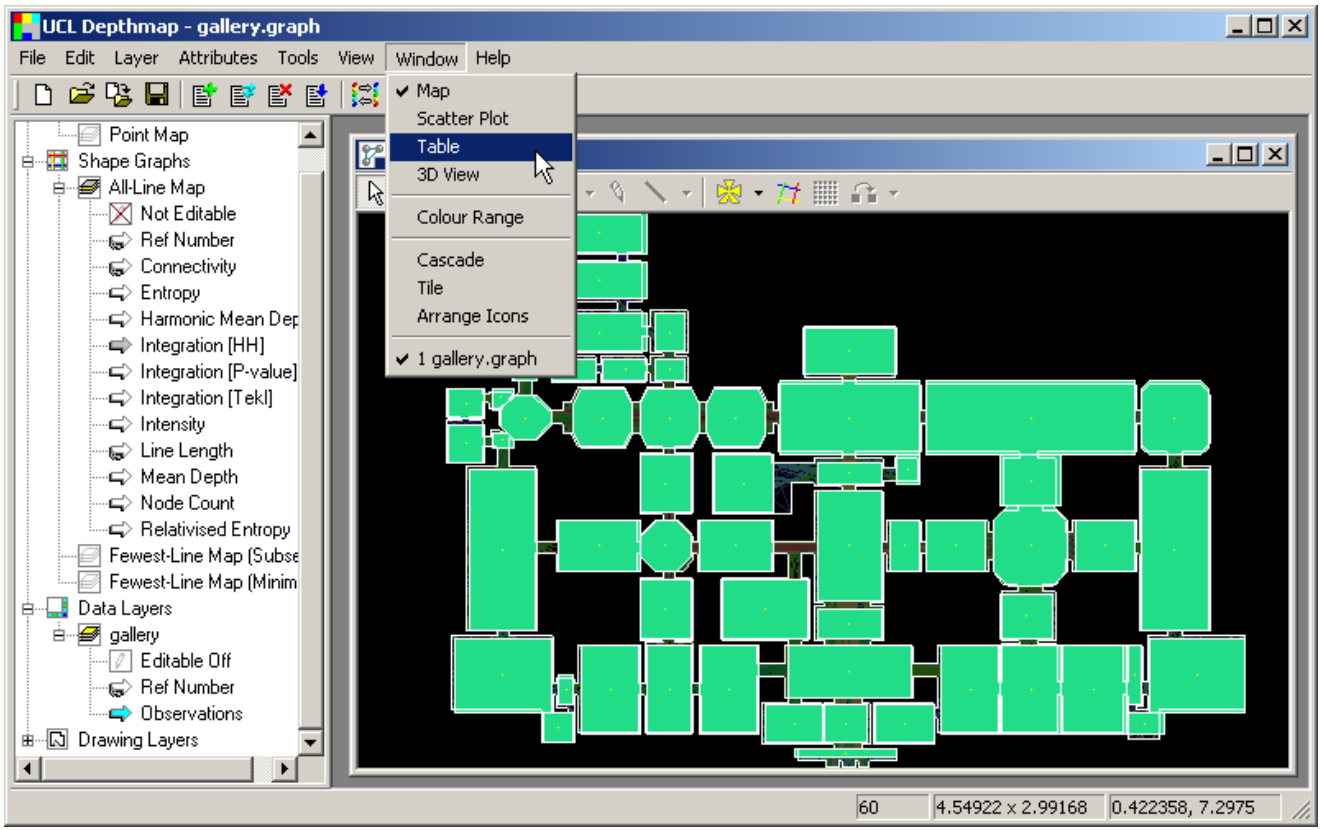
First, let us change the name of the column: either right-click on the column name in the sidebar, or choose ‘Rename Column’ from the ‘Attributes’ menu.

Add a new column



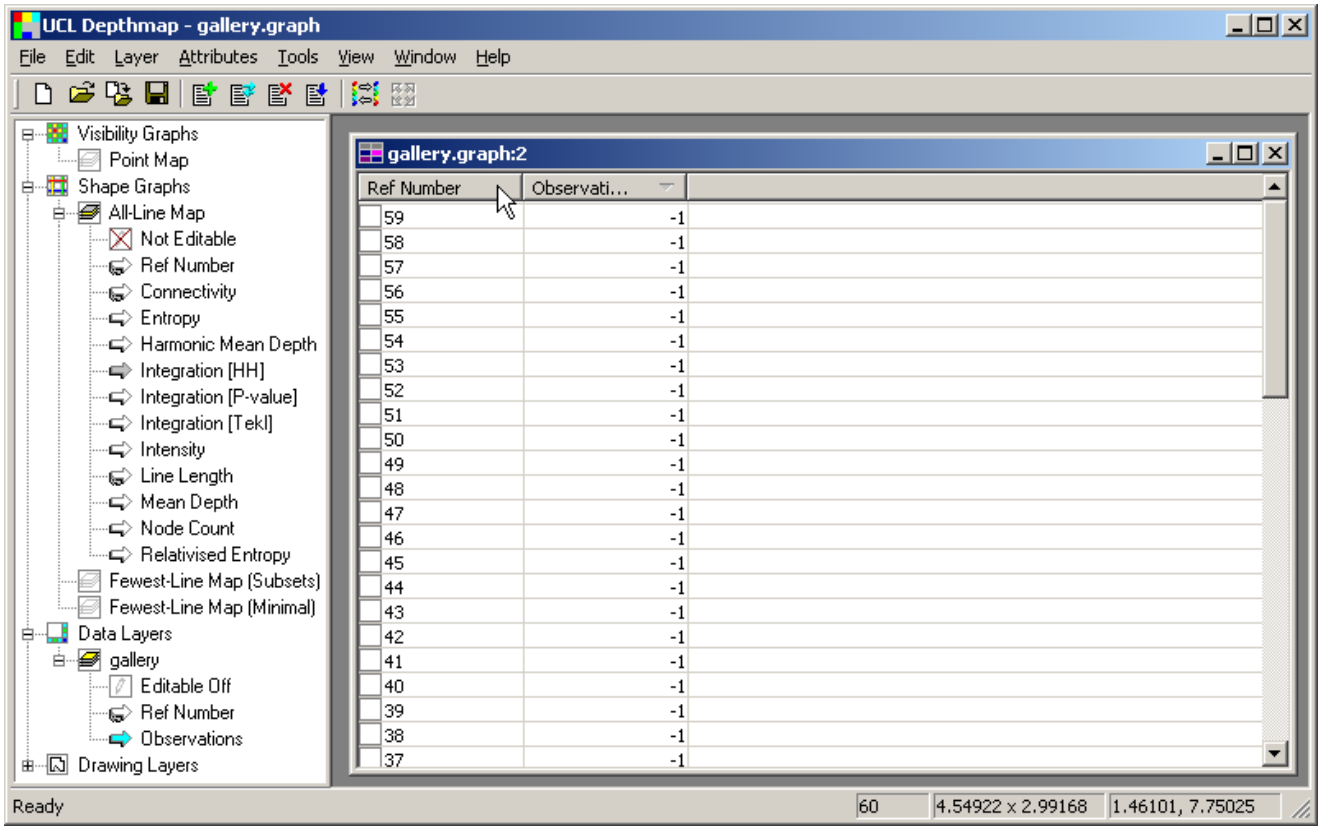
Type in the name for your new column. Note that your version of `gallery.mif` and `gallery.mid` already contain the dummy observation data that we will enter here in a column called `Dummy_Observations`.

Table view



Now that we have a new column, let us open a table or spreadsheet view of the map in order to enter the observation data. From the 'Window' menu, select 'Table'.

Table window



Note that the table shows that every row in the 'Observation' column is set to -1, which represents 'No Value' in Depthmap. The ordering is also by 'Observation', and as all the values are the same, no particular order is shown. Let us order by 'Ref Number' instead. Click on the 'Ref Number' column header.

Table window

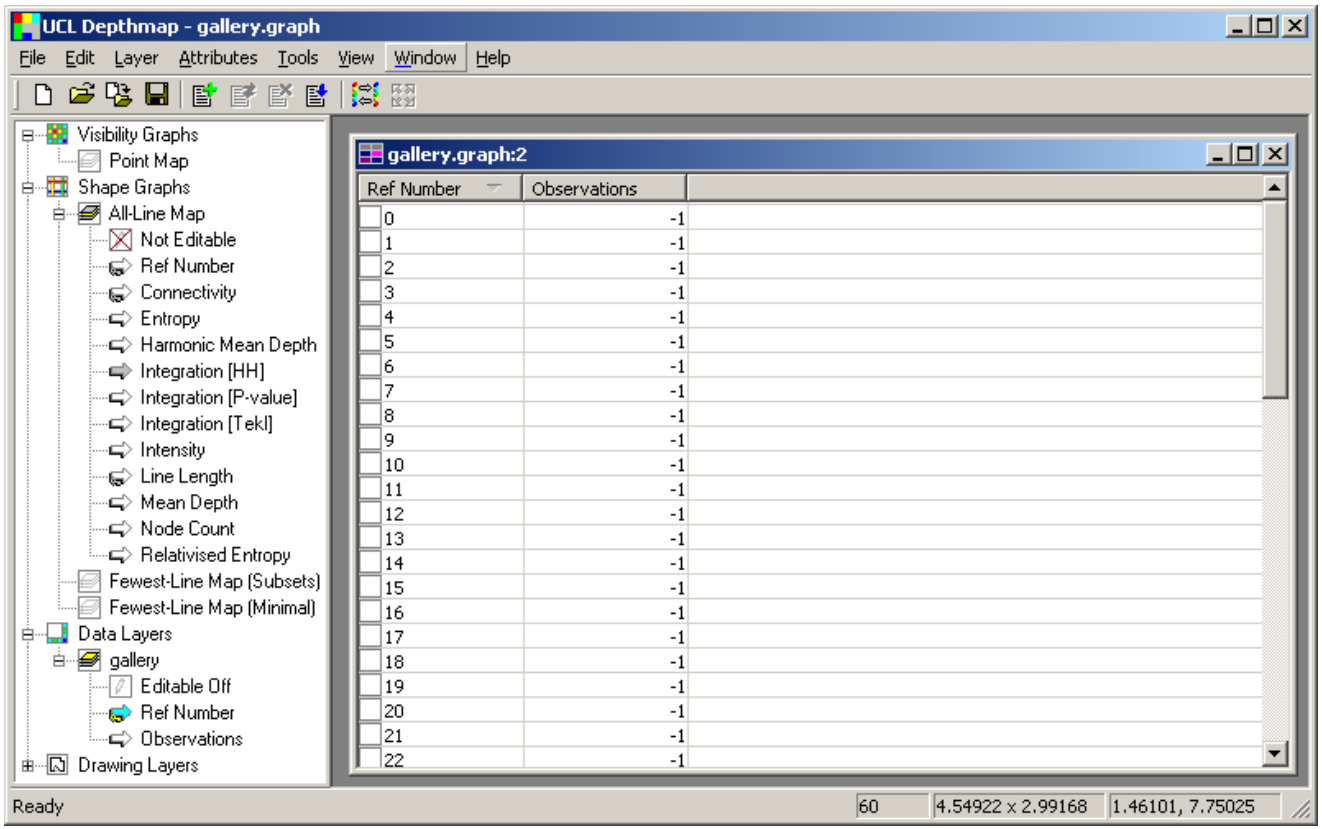
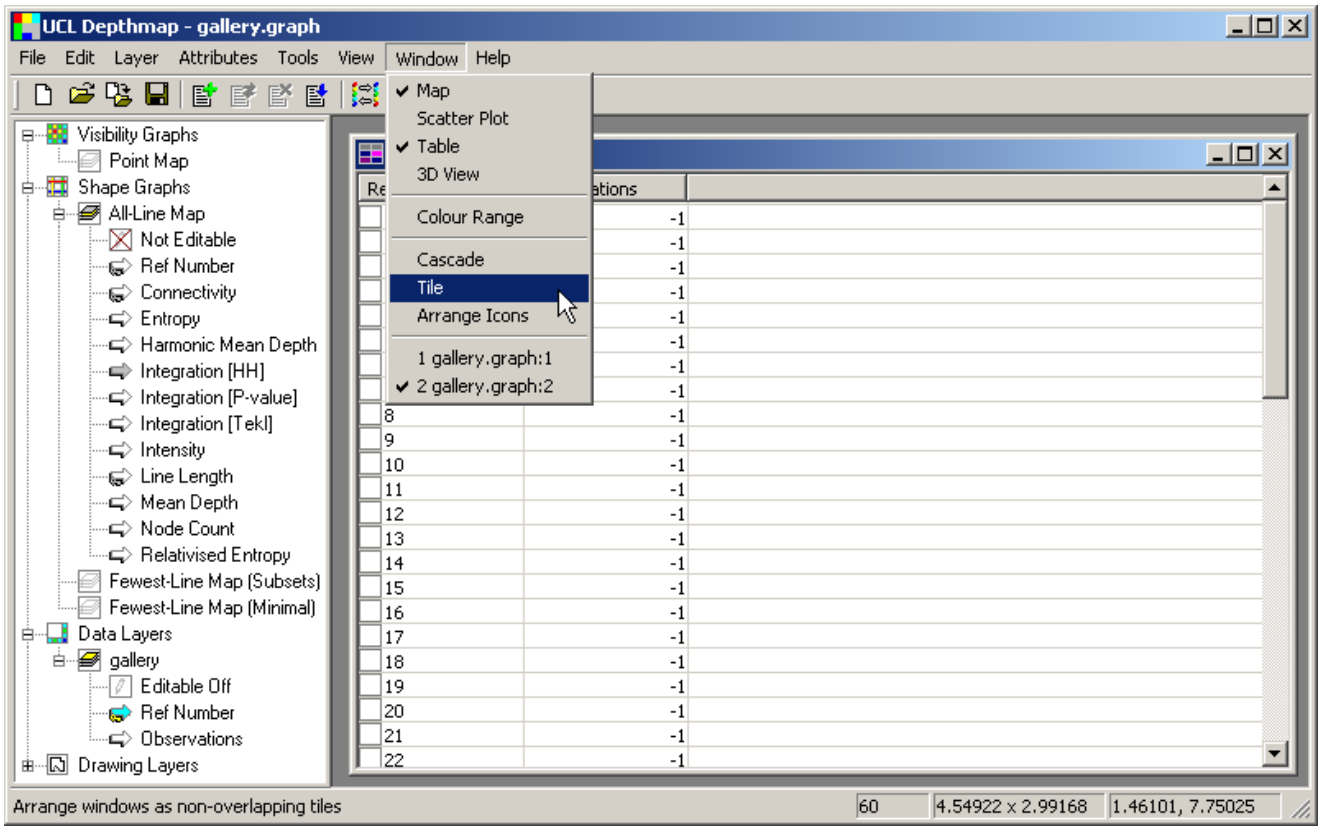
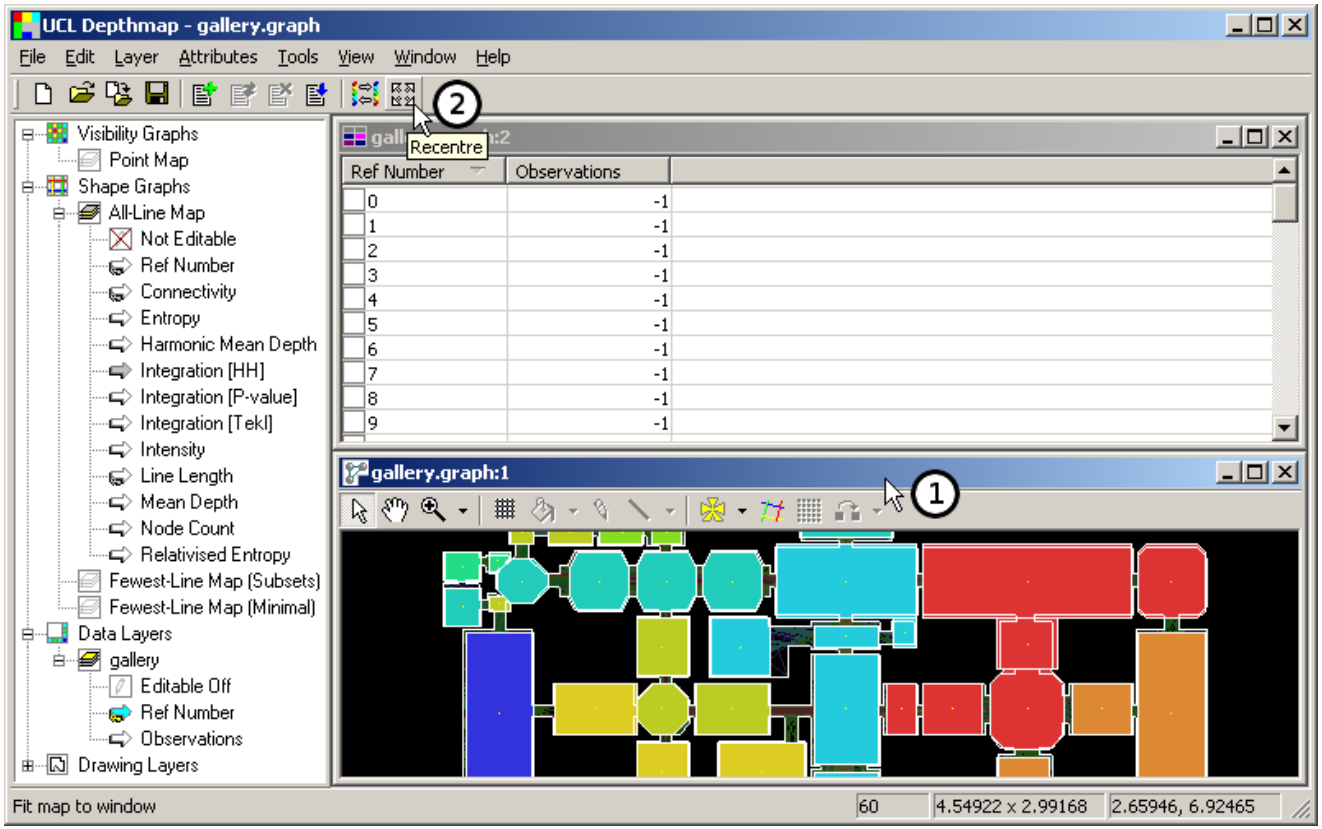


Table window



We can view both the map window and the table window at the same time by selecting ‘Tile’ from the ‘Window’ menu.

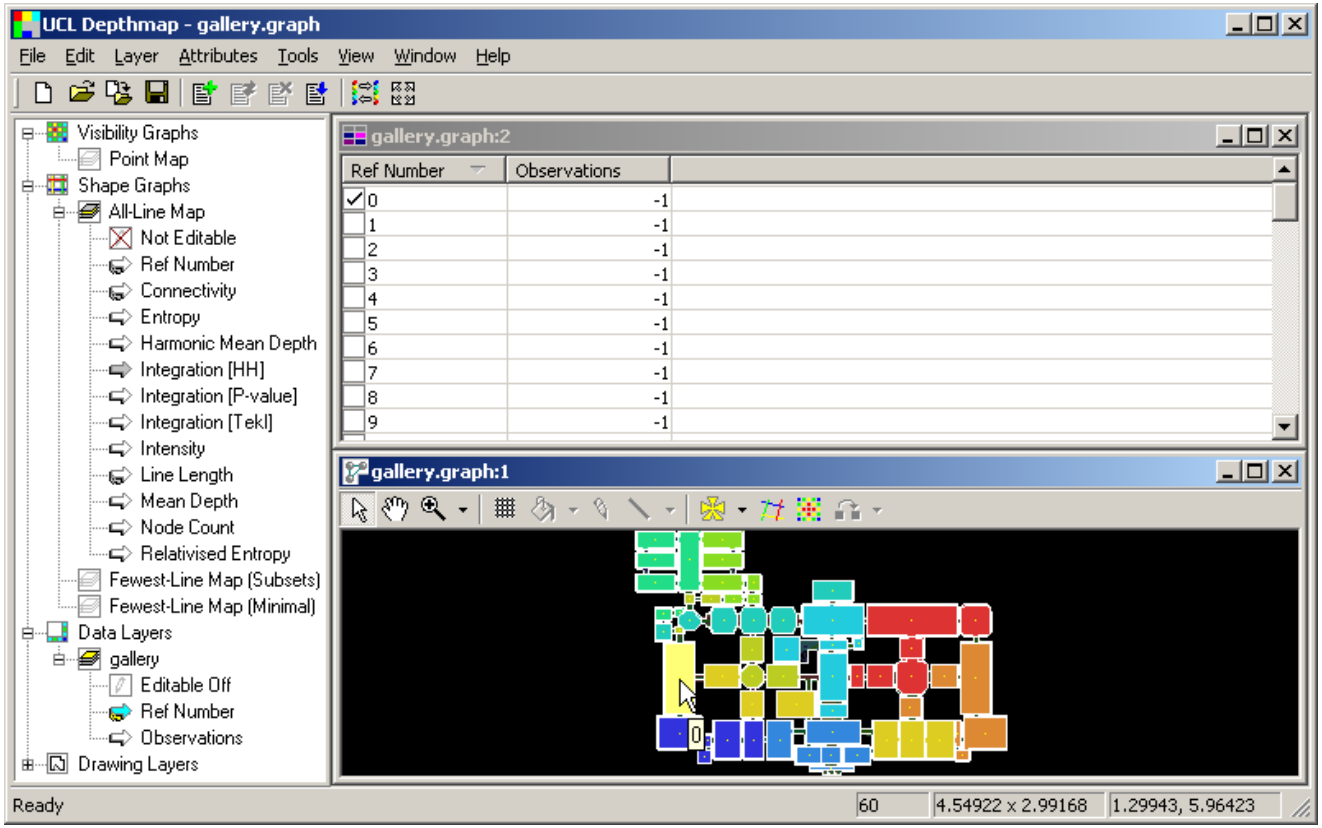
Table window



You may want to display Depthmap full screen in order to have a little more room.

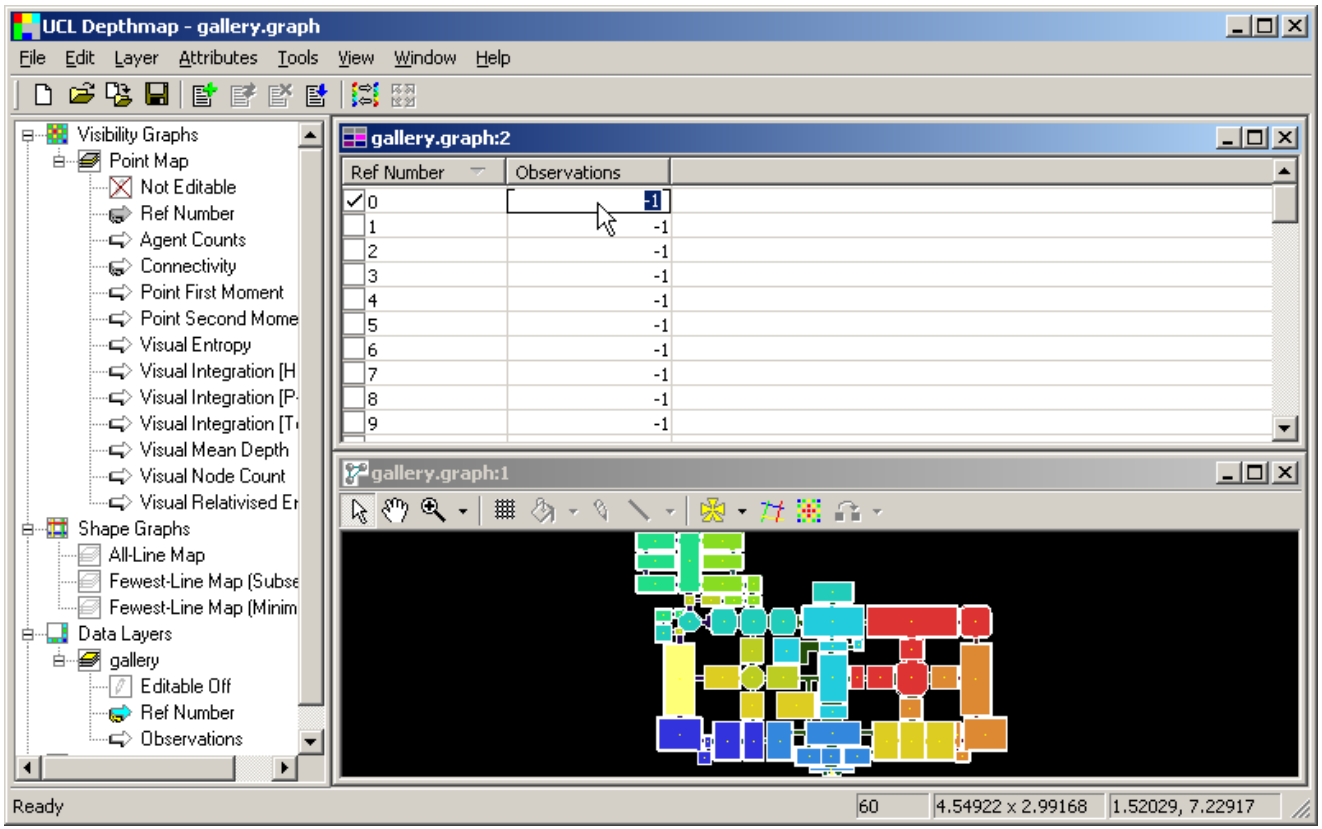
I will recentre the map window view by first clicking on the map window title bar to select it, and then clicking on the recentre button.

Table window



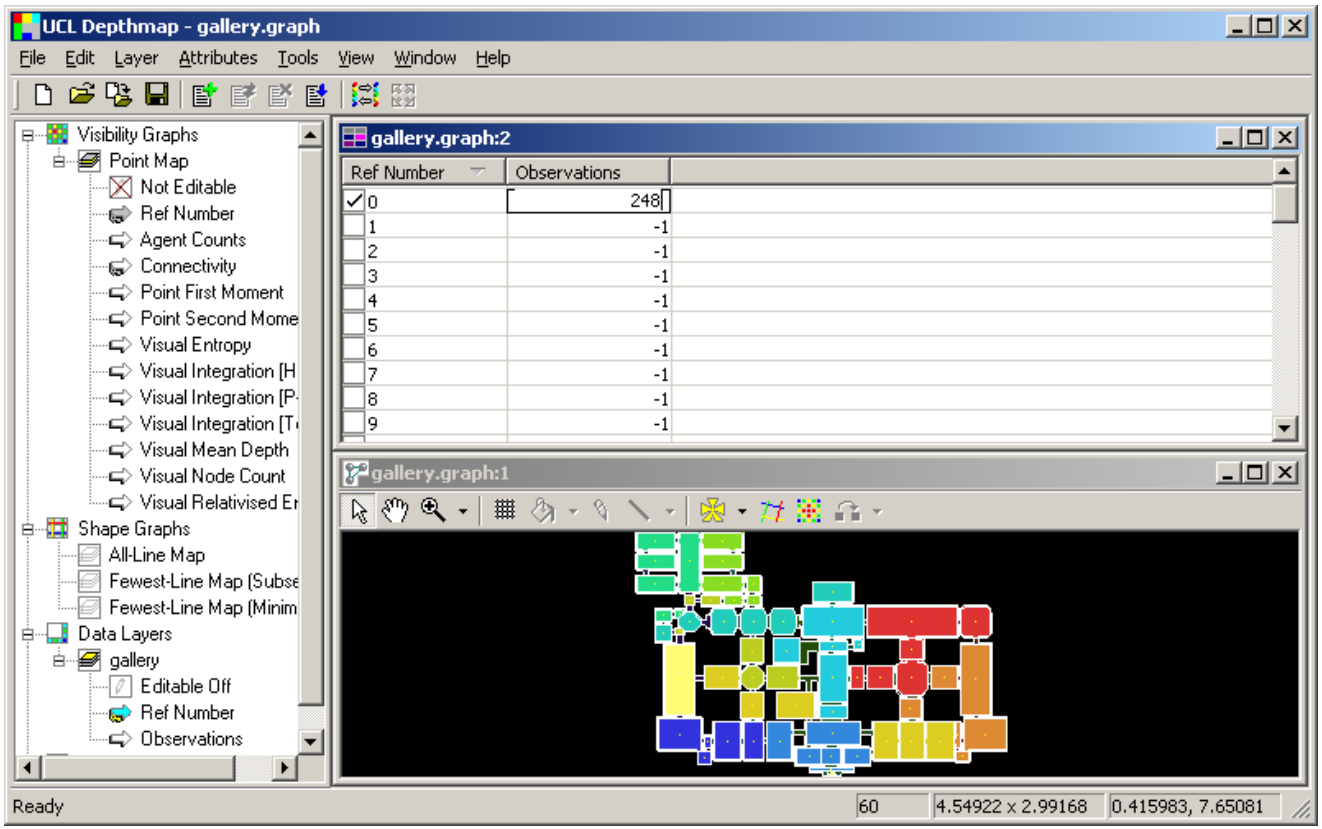
Note that if you select a shape by clicking on it in the map window, the equivalent row in the table window is shown with a tick next to it. Equally, if you select a row in the table window by clicking on the tick box for the row, it is highlighted in the map window.

Entering data



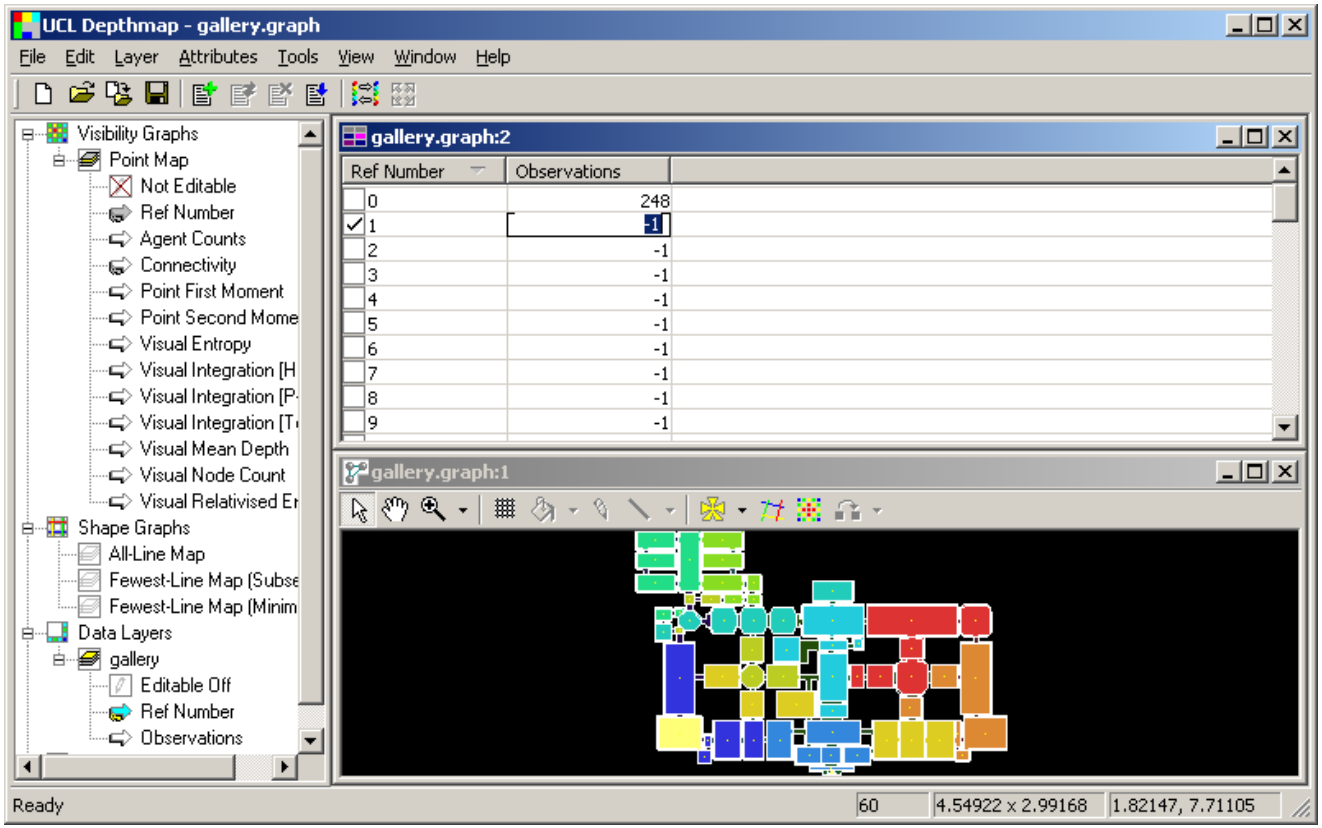
In order to enter a value for an observation, click on the value you want to change.

Entering data



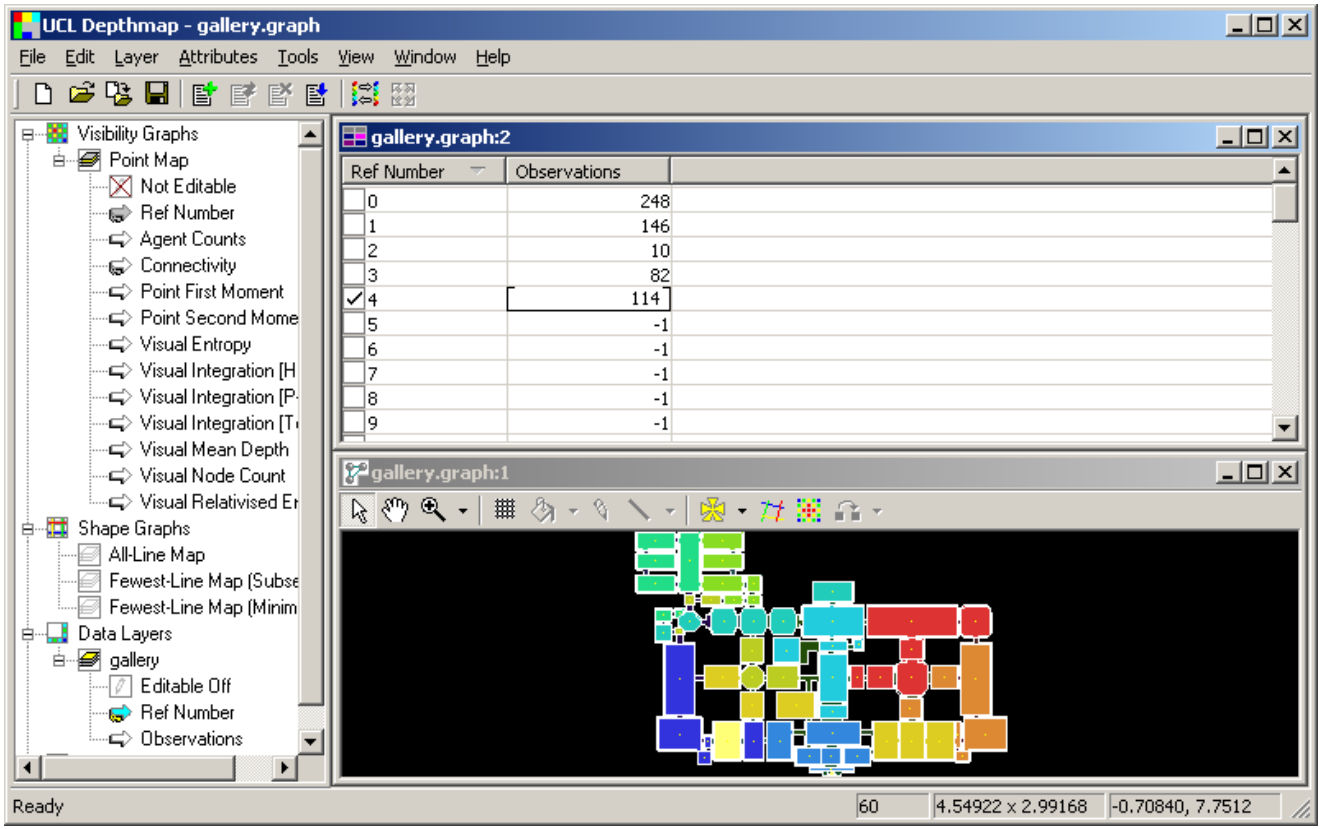
Then type in the value for the row. For example, if 248 people moved through this room in an hour, you might type in ‘248’.

Entering data



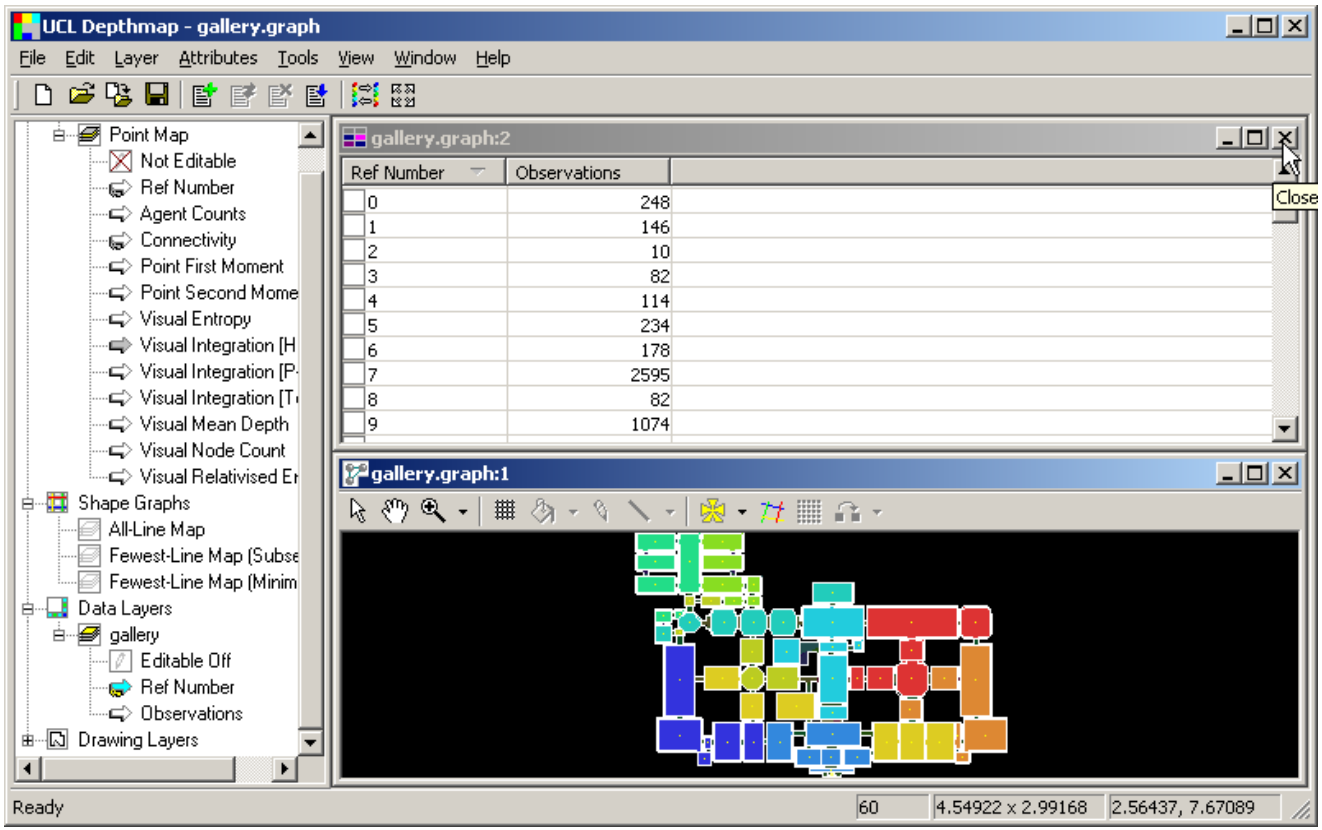
To finish entering the value, press the `Return` key. Notice that as soon as you press `Return`, the cursor moves to the next row down, and the next room is highlighted in the map view.

Entering data



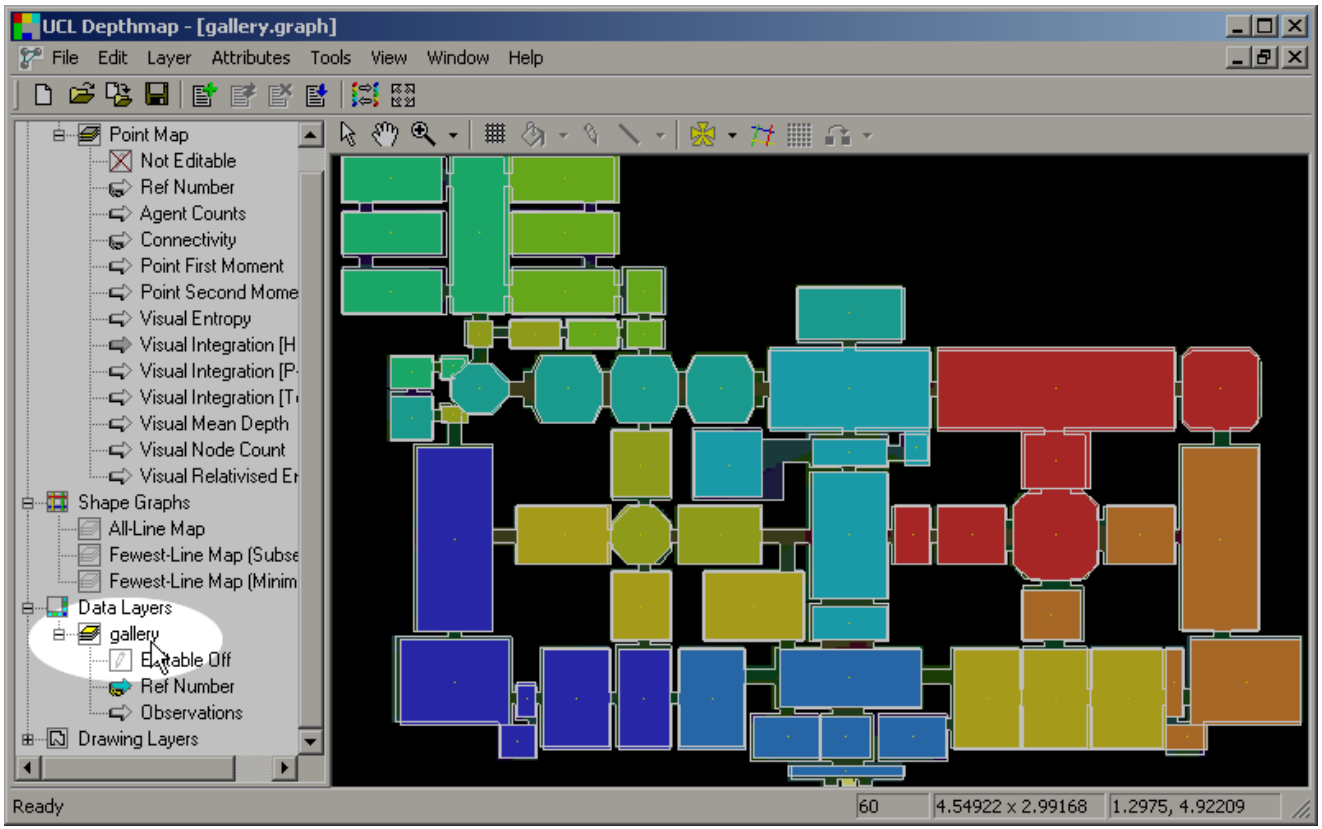
Continue to enter the observation values for each room in the plan.

Entering data



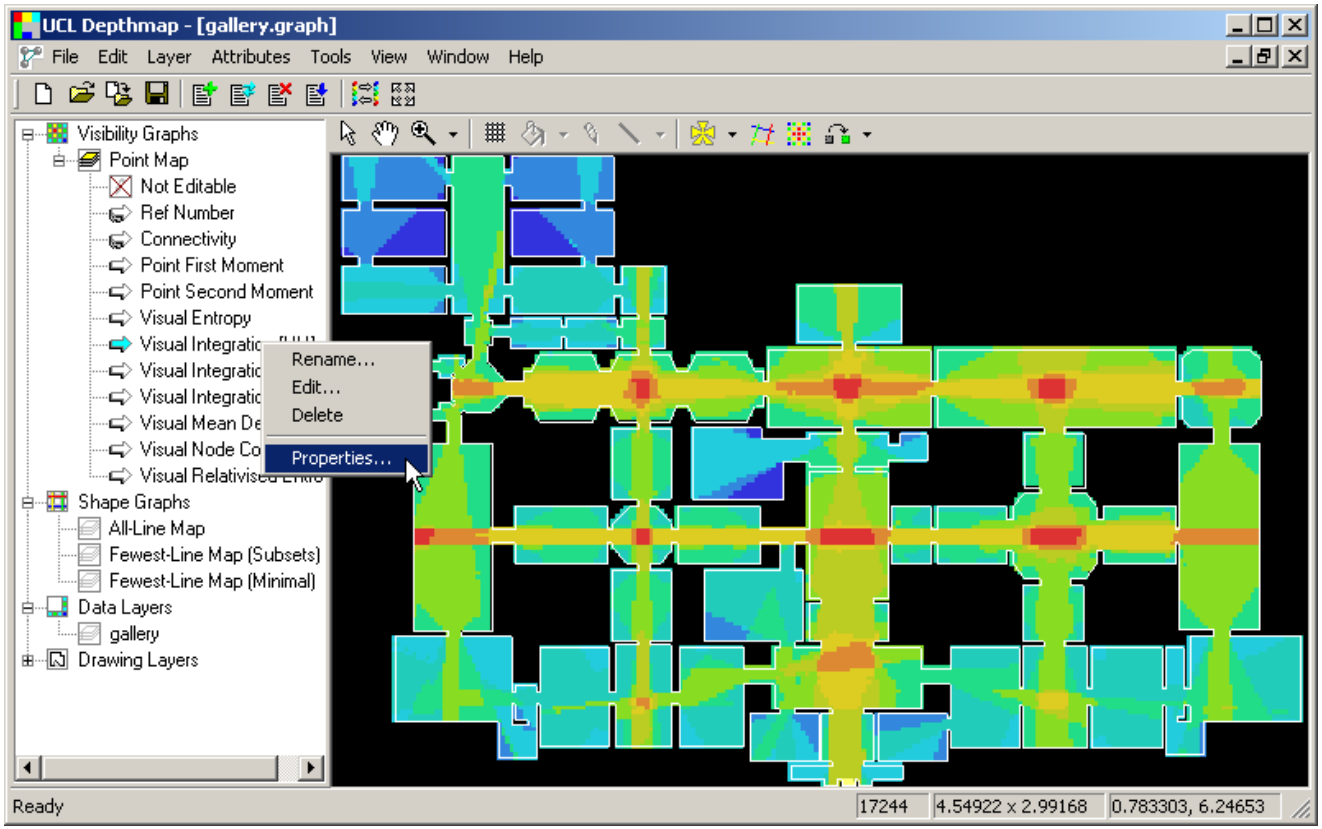
Once the data are entered, the table window can be closed.

Entering data



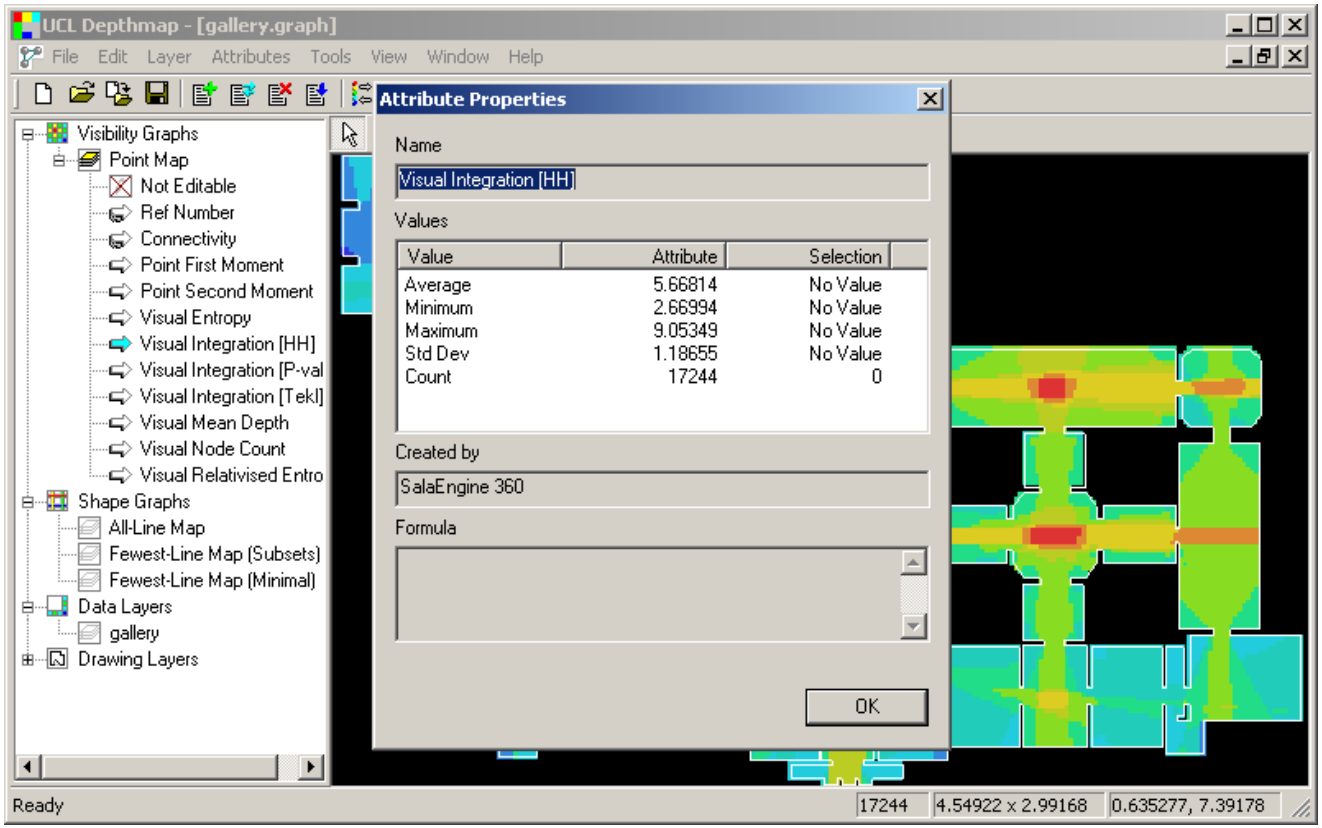
For now we will also hide the observation layer, by clicking on the 'gallery' layer title in the side bar.

Summary statistics



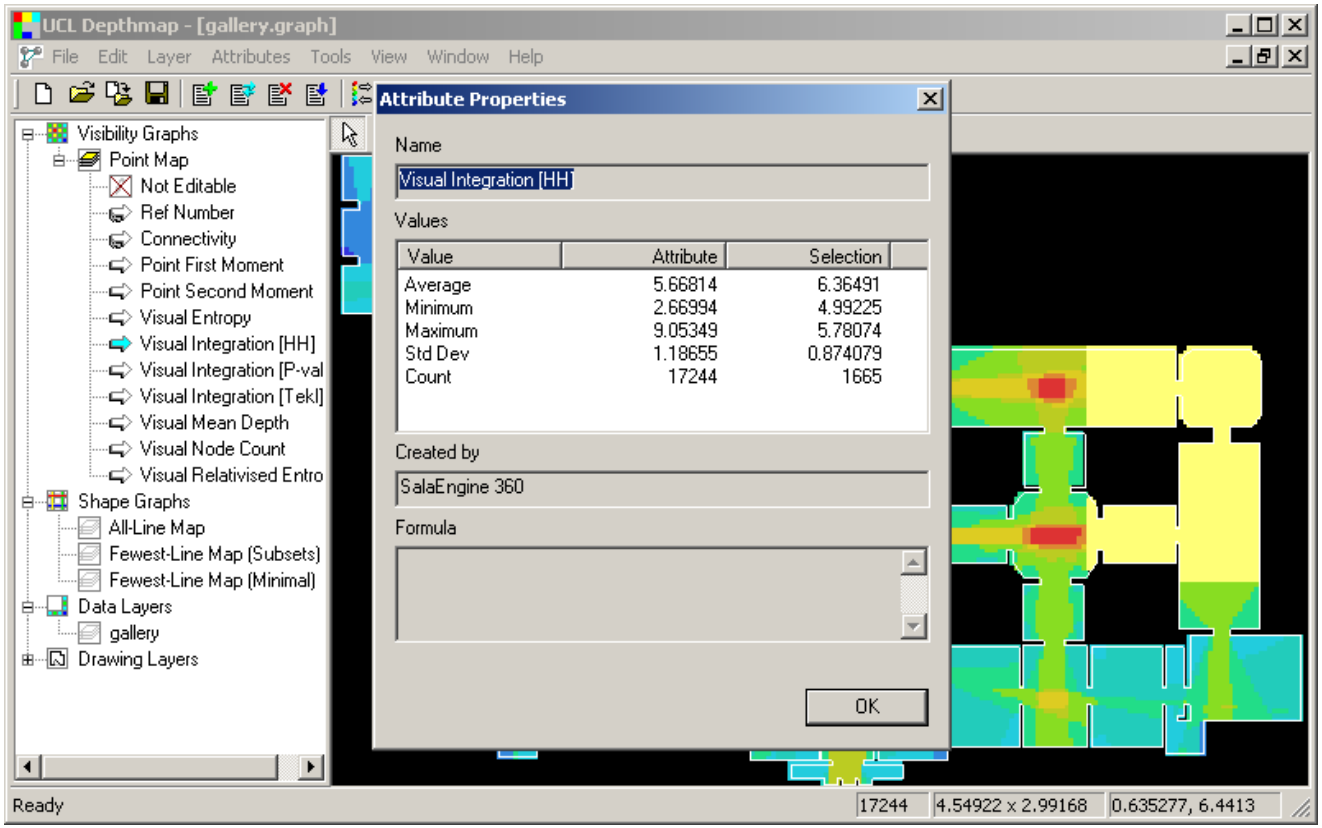
You can obtain summary statistics about a column by right clicking on the column name in the side bar, and then selecting ‘Properties’ from the menu. Alternatively, you can select ‘Column Properties’ from the ‘Attributes’ menu.

Summary statistics



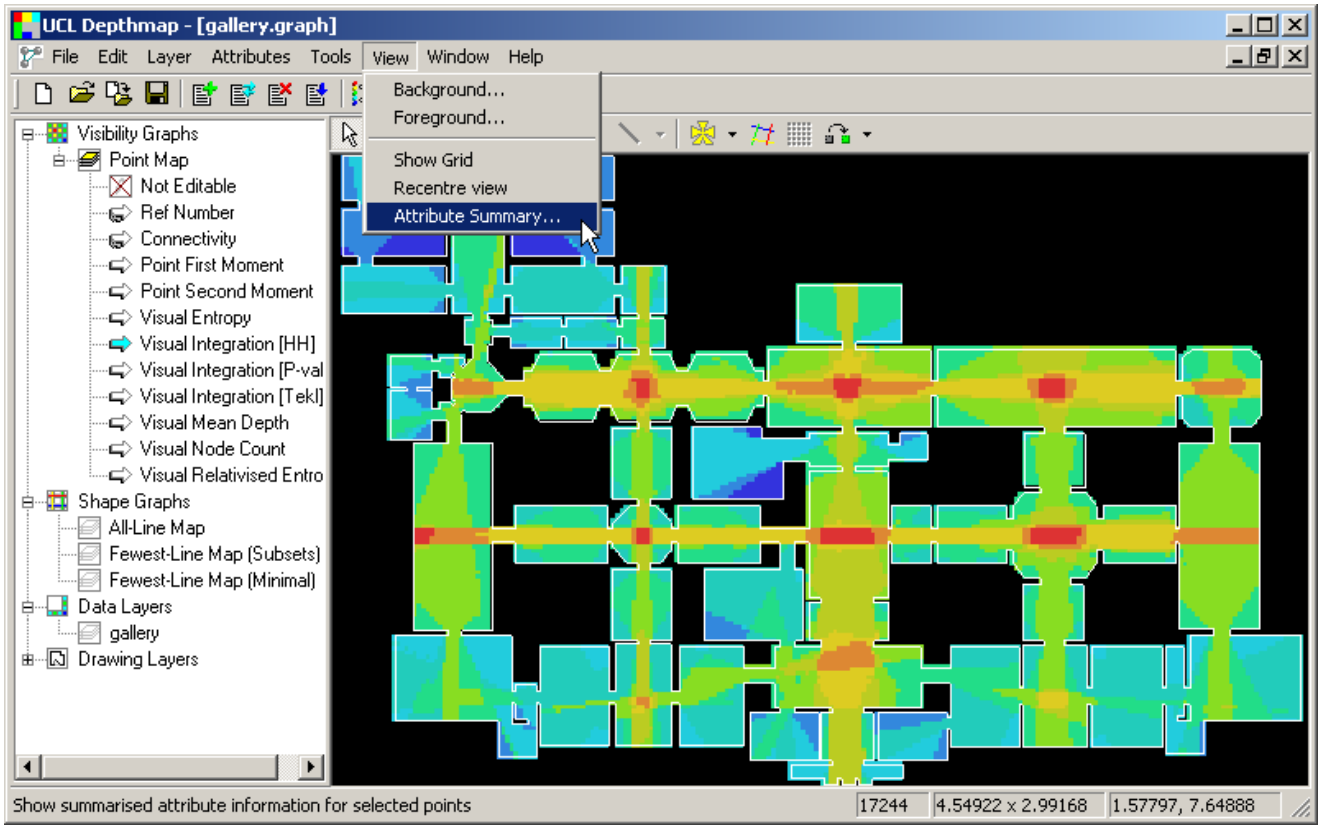
The summary statistics include maximum, minimum, average (mean) and standard deviation of the values, as well as a count of the number of points, lines or shapes in the system.

Summary statistics



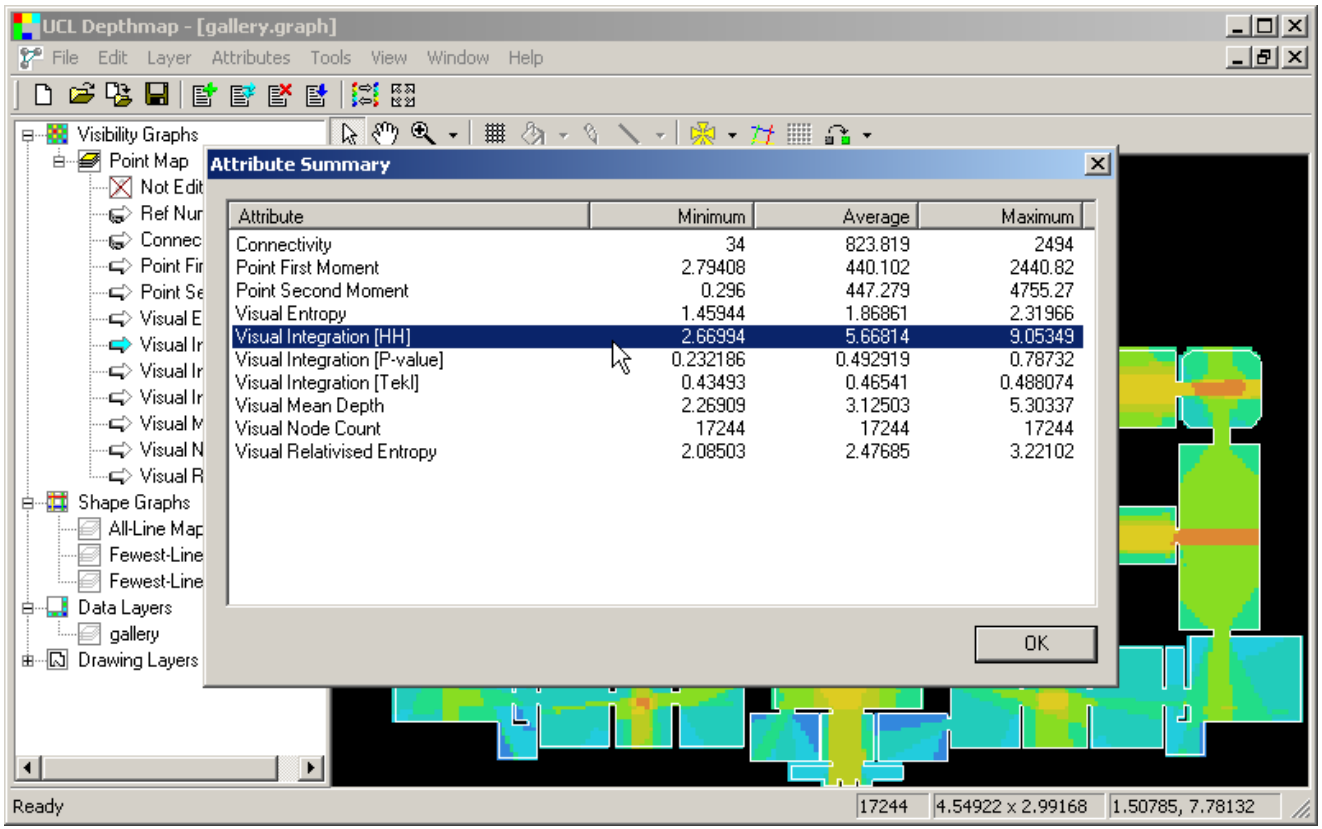
You can compare summary statistics for a range of values against all the values in the system by first selecting the range (or single point) and then choosing ‘Properties’.

Summary statistics



In addition to the per column summary statistics, you can get a briefer set of statistics for all the attributes that have been calculated by choosing ‘Attribute Summary’ from the ‘View’ menu.

Summary statistics



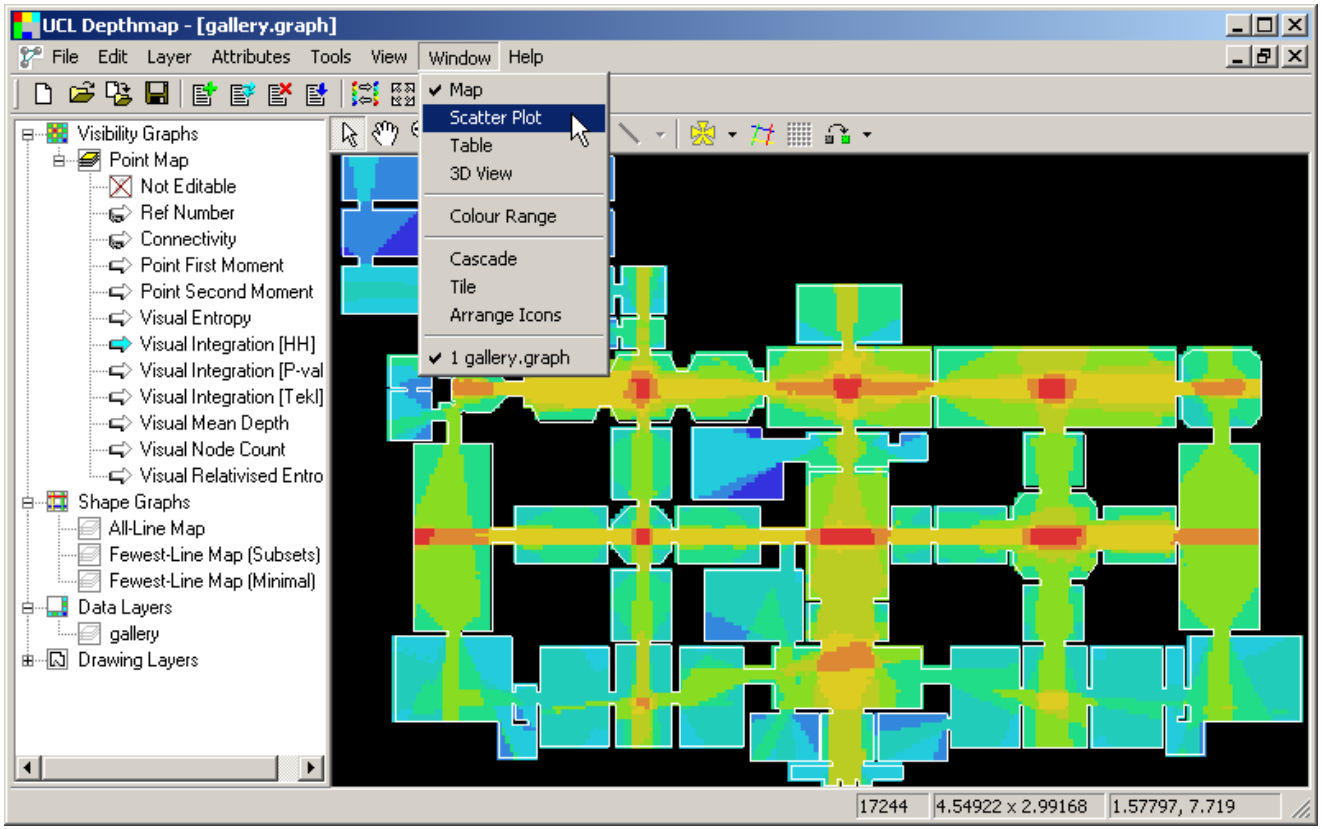
If you double click on any attribute in the summary table, then it will take you to the column properties dialog box for that column.

Scatter plots

In this section, we will cover the investigation of data through scatter plots. We will compare attributes with each other, and across maps.

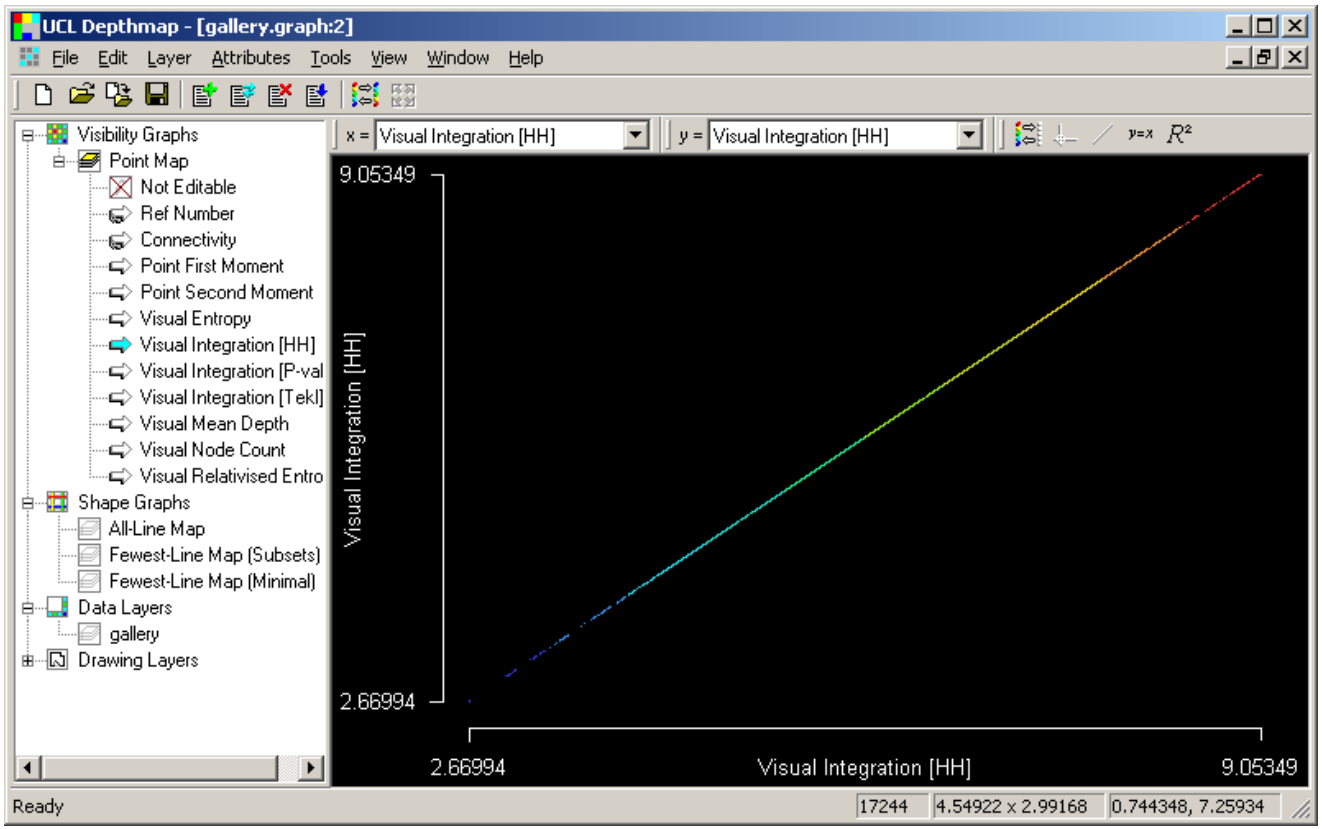
In order to compare across maps, we will need to introduce the 'Push Values to Layer' feature, which copies attributes from one layer to another.

Simple scatter plots



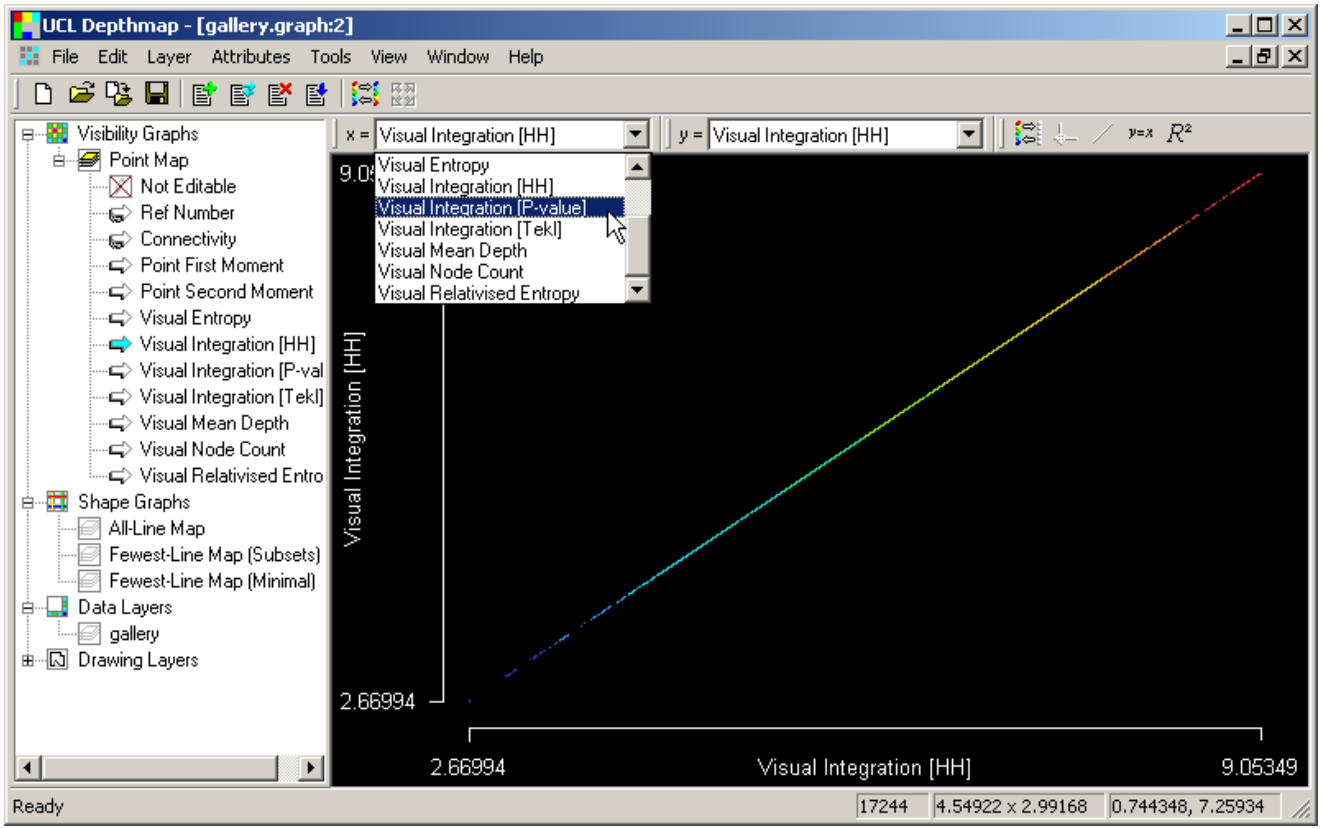
You can open a scatter plot window by selecting ‘Scatter Plot’ from the ‘Window’ menu.

Simple scatter plots



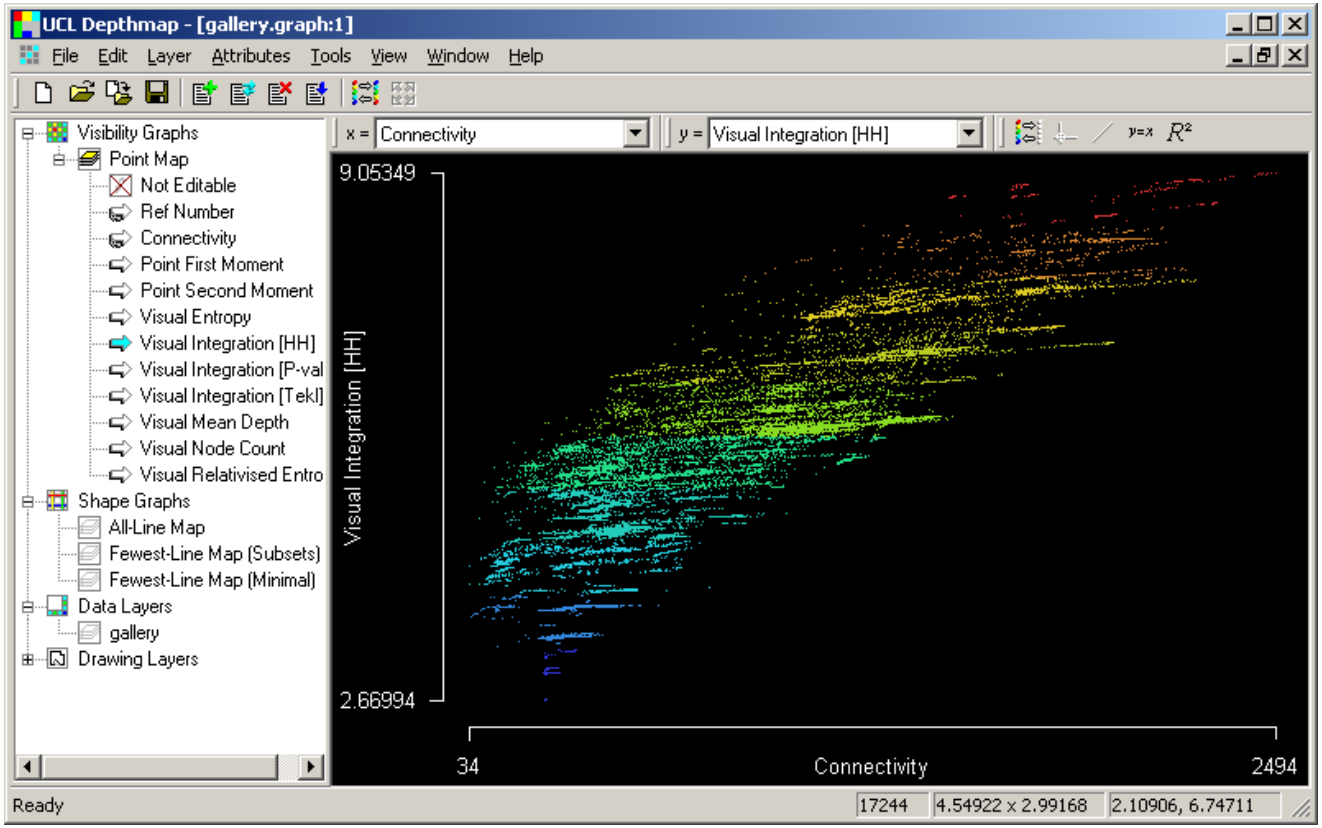
The default is to show the currently displayed attribute on both the x and y axes.

Simple scatter plots



We can change the axes by using the drop down menus at the top of the screen.

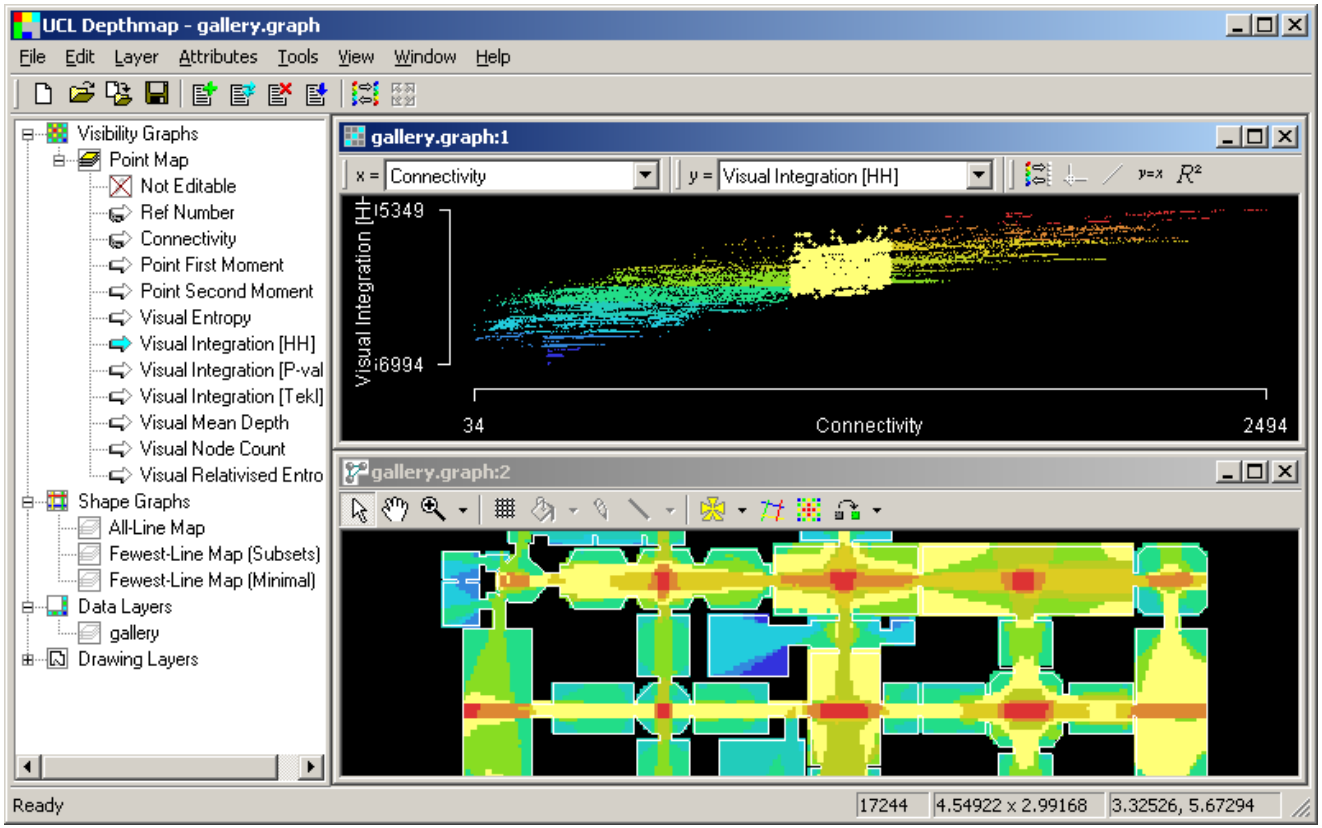
Simple scatter plots



For example, this scatter plot shows the Visual Integration as a function of the Connectivity.

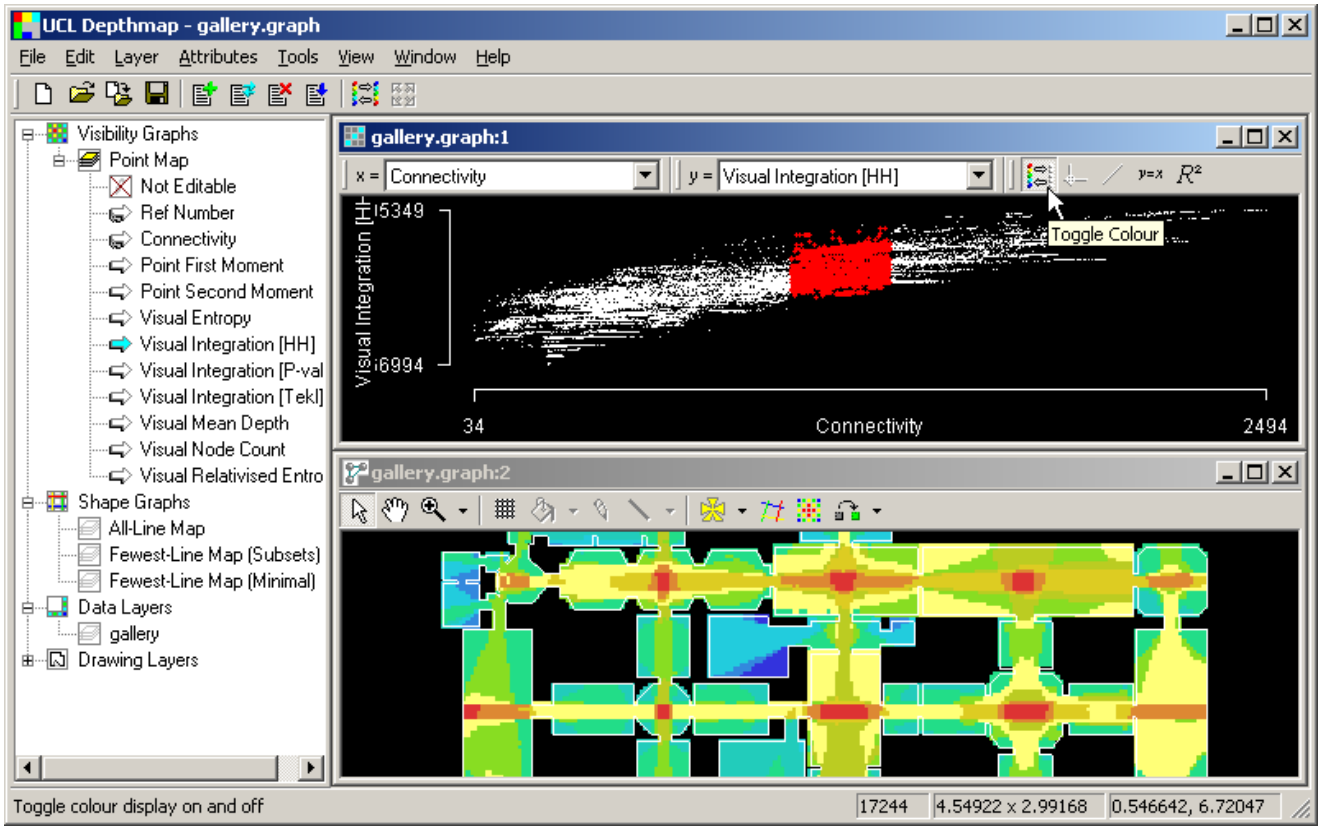
Tip: you can copy the screen using `Ctrl-C` or by selecting 'Copy Screen' from the 'Edit' menu. The screen shot, which is in vector graphics (which are smooth when printed), can then be pasted into your favourite word-processing or presentation software.

Simple scatter plots



If you tile the windows, using ‘Tile’ from the ‘Window’ menu, and then select on the scatter plot window, the range will be highlighted on the map window as well.

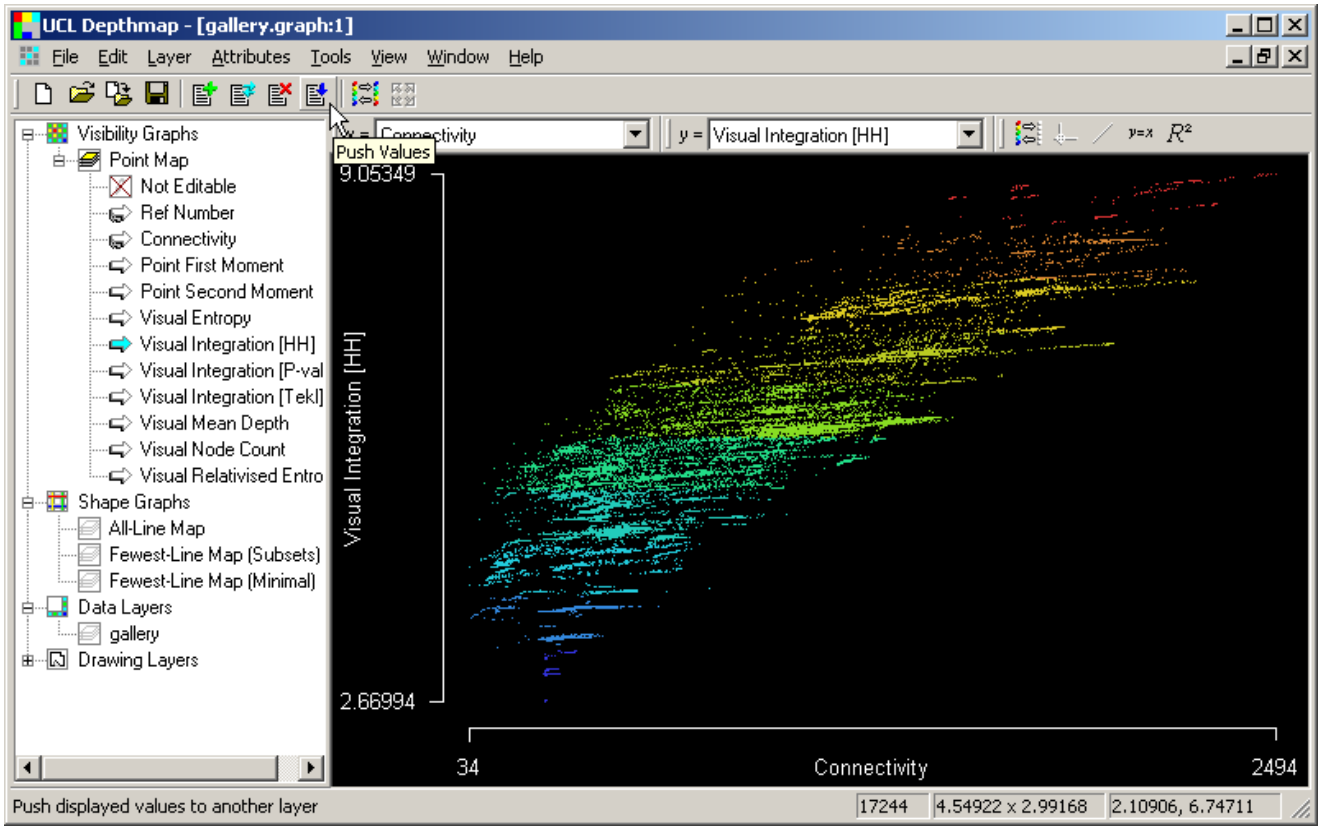
Simple scatter plots



The 'Toggle Colour' button on the scatter-plot window tool bar turns off the colour scale, which you may find useful for preparing printed material. Note that the highlight colour for the selection changes to red, so that it is easily visible against white data points or a white background.

The remaining buttons will be described in the section on comparing data from different maps.

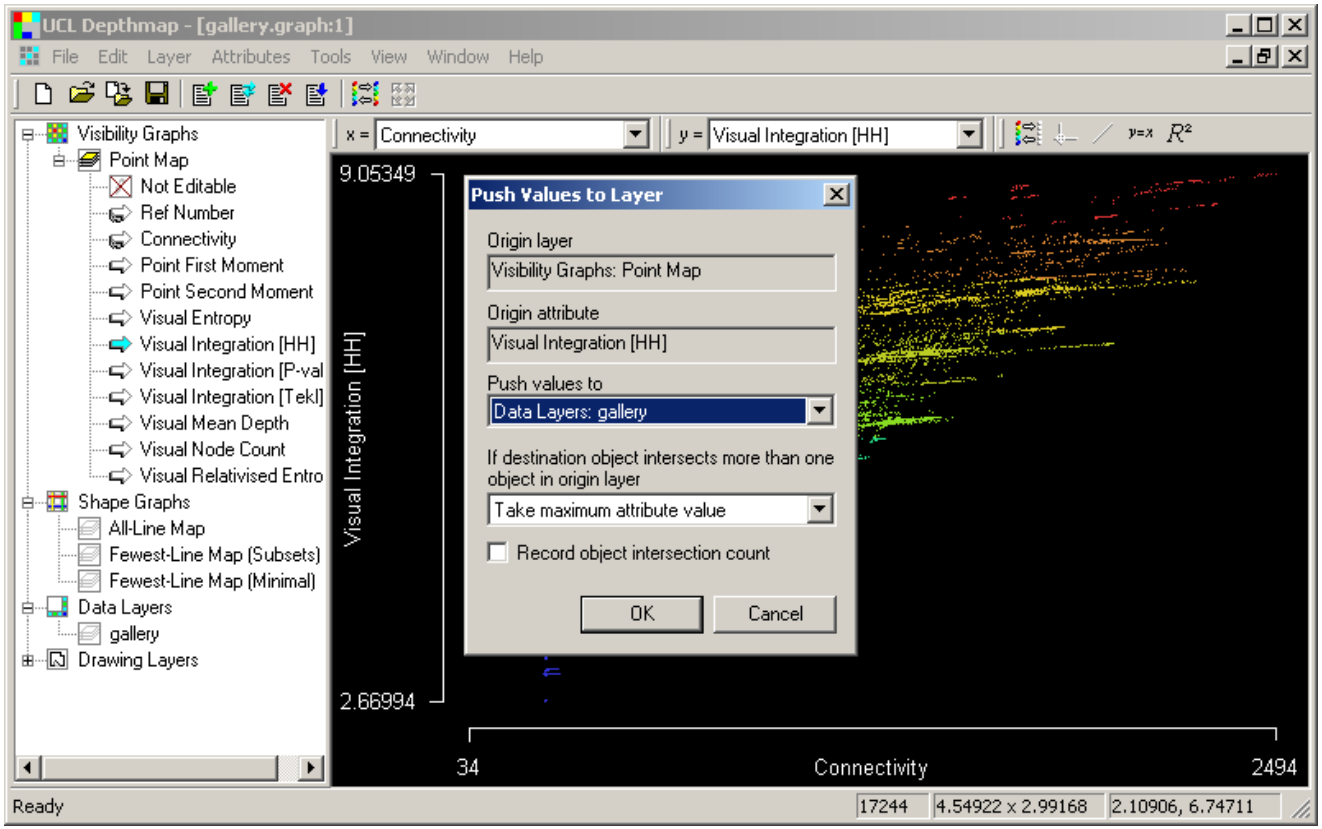
Pushing values to another layer



Simple scatter plots are all very well for theoretical comparisons of measures, but they cannot easily be used to compare VGA maps to axial maps, or VGA or axial values to observation data.

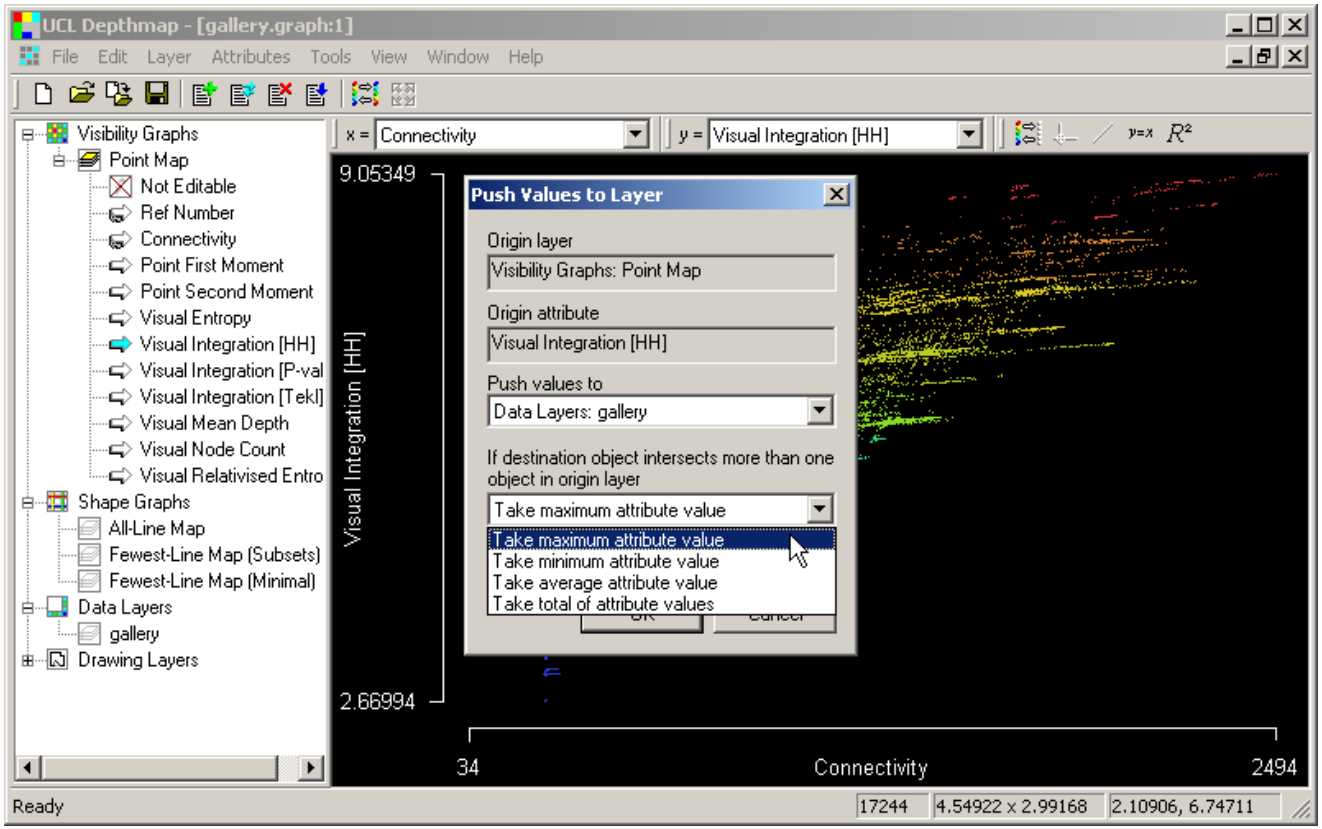
In order to make comparable data, we can *push values* from one layer to another layer, either by clicking the 'Push Values' button on the main tool bar, or by selecting 'Push Values to Layer' from the 'Attributes' menu.

Push values options



For example, to compare VGA data to observation data, we can collate VGA data by room.
Selecting ‘Push Values’ when we are viewing the VGA layer automatically selects the ‘gallery’ layer as the default to which to push the data.

Push values options

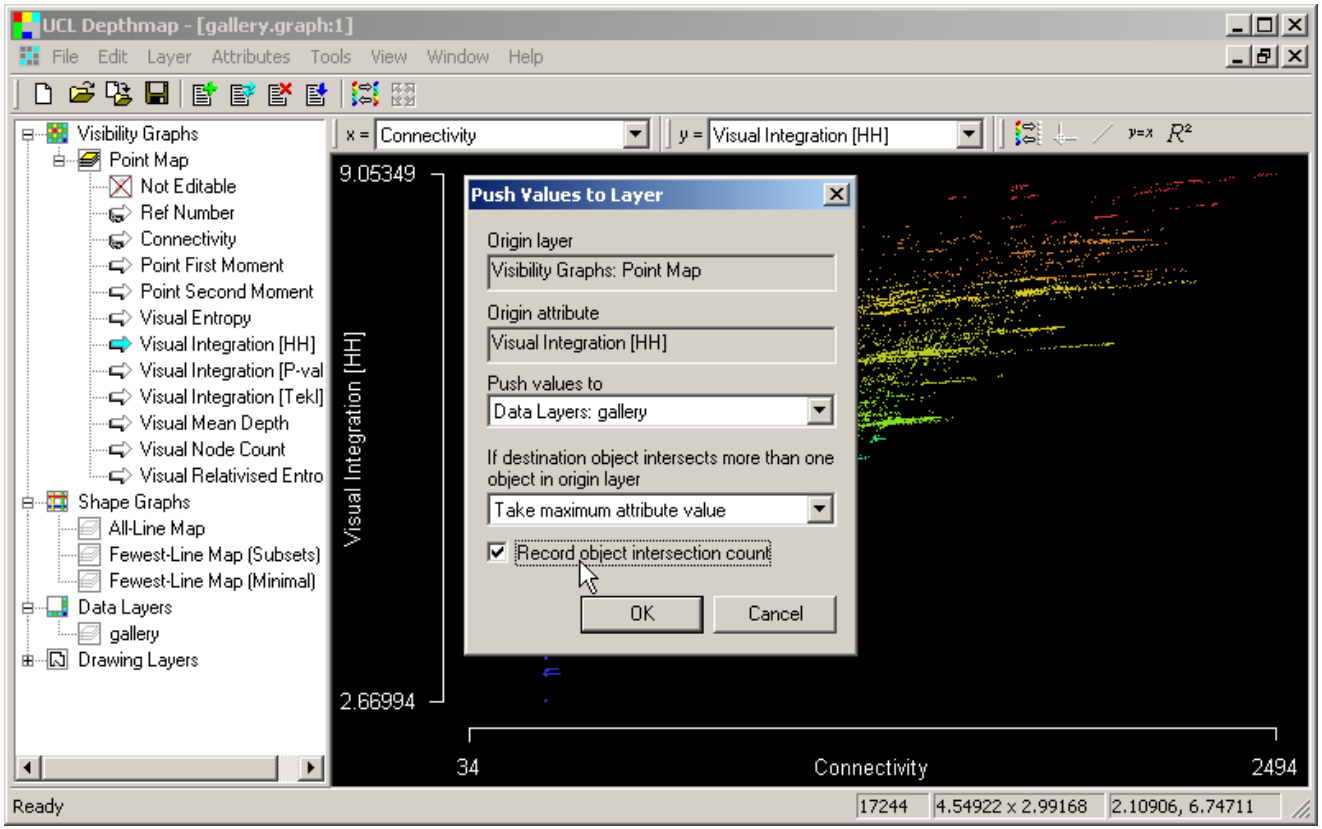


A drop down menu gives us various methods for how the data will be pushed from the VGA layer to the data layer.

Remember that the VGA layer contains points, and the data layer contains polygons representing rooms.

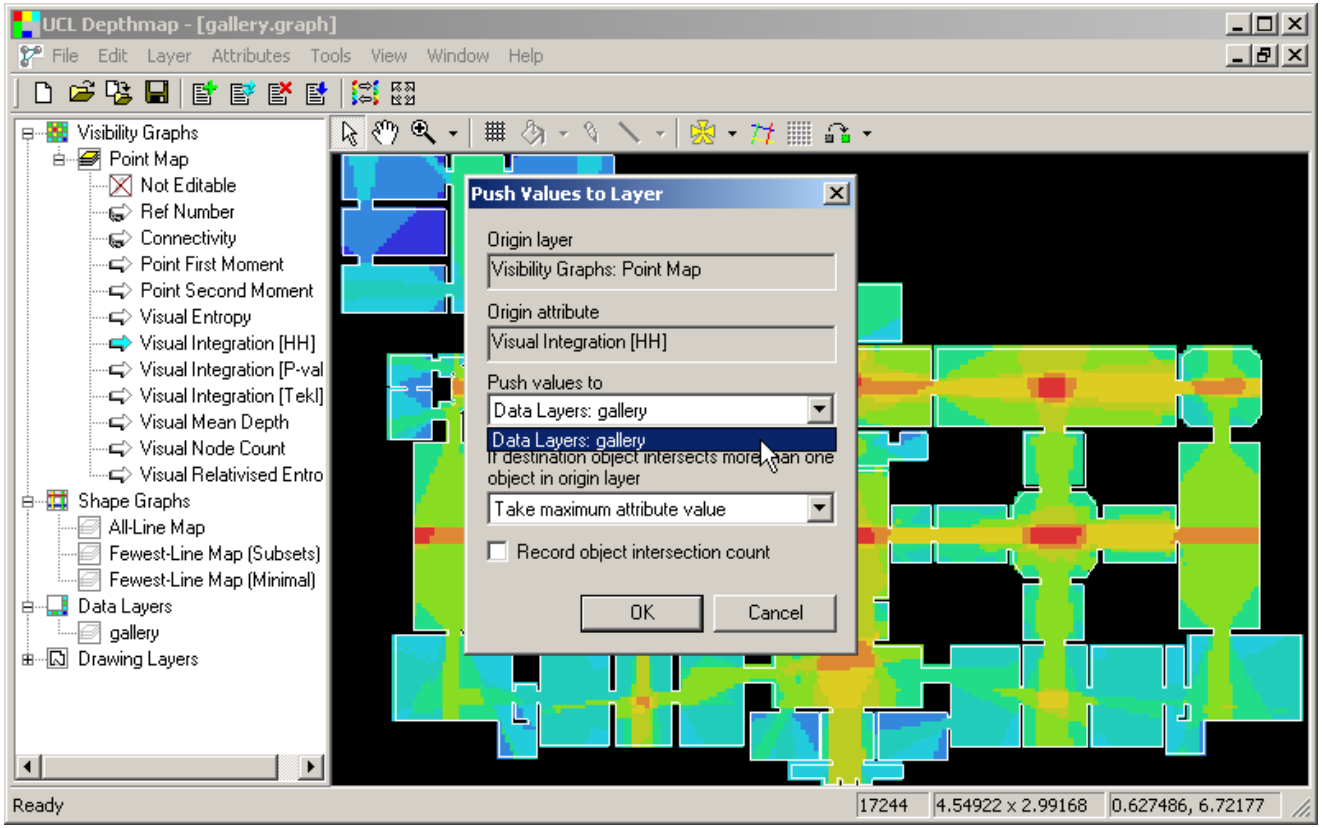
When transferring the data from the VGA layer, many points will intersect each room. We might want to take the maximum point value, the minimum point value, or the average for the whole room.

Push values options



We may also want to know exactly how many points were used to obtain the value transferred to the gallery rooms layer. The 'Record object intersection count' options will create an extra column in the gallery rooms layer to say how many points were used to form the value transferred.

Push values restrictions

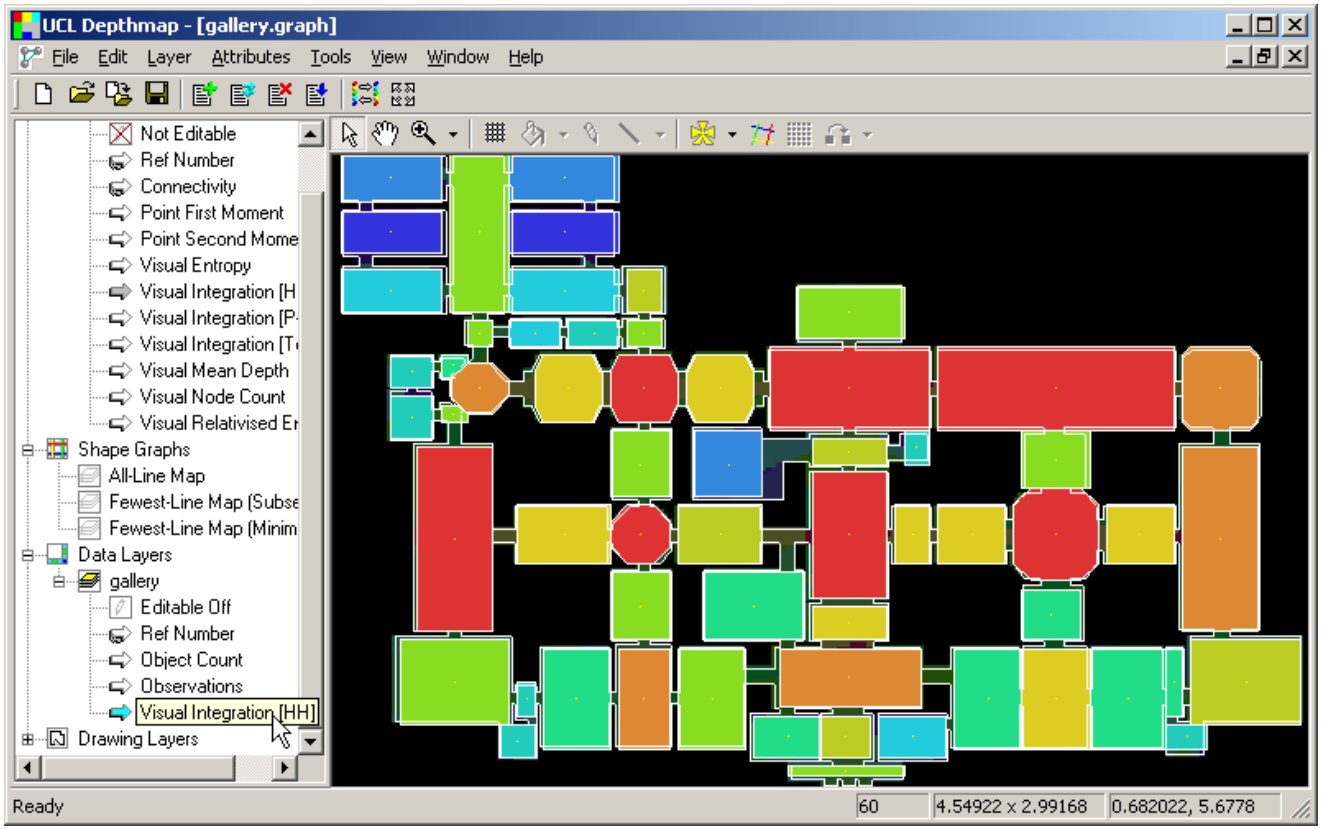


Note that the drop down menu for where to push data does *not* allow us to push data to the axial line layer.

This is because VGA data are recorded for points, and axial line data are recorded for lines. Unless lines are drawn exactly over the VGA grid there will be no overlap of the shapes in the VGA layer and the axial layer.

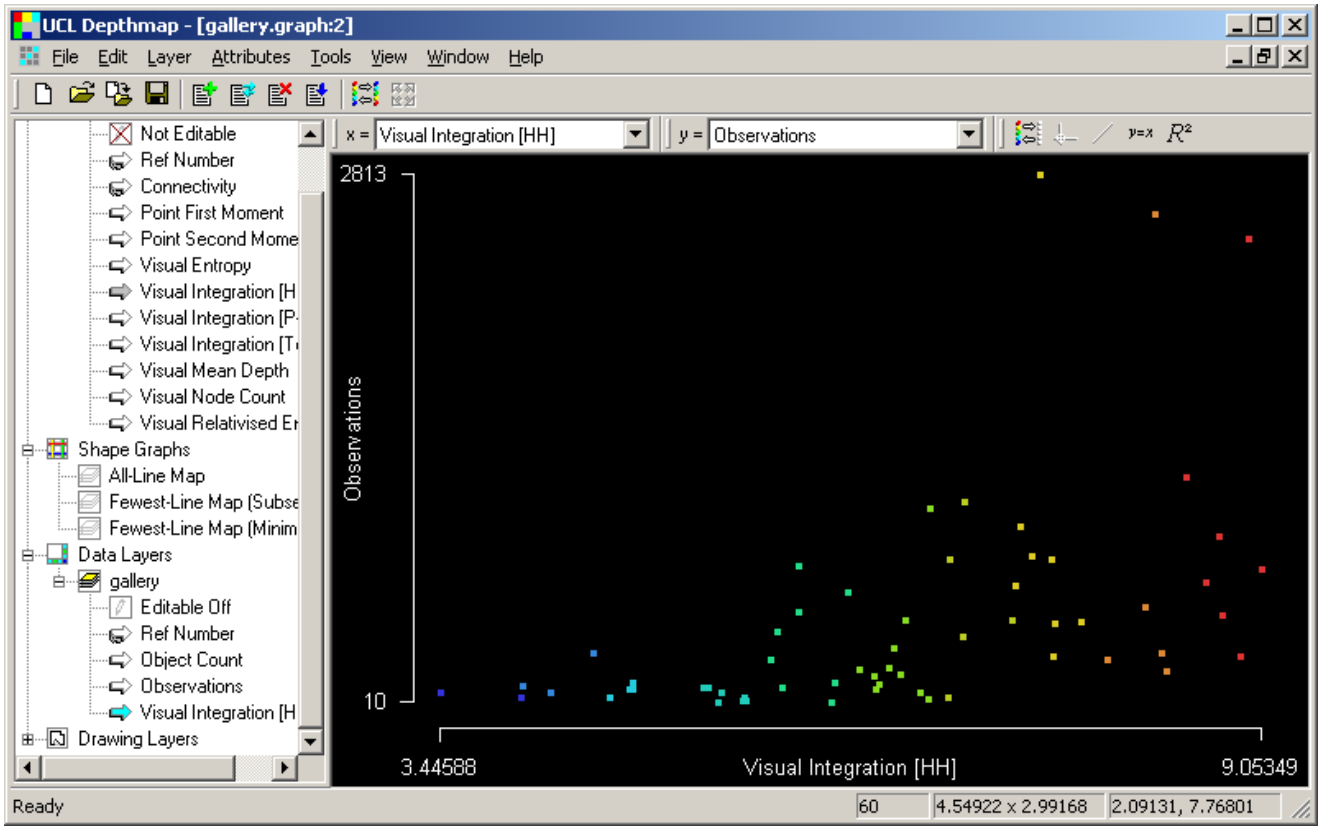
Therefore, in order to compare the two, a proxy layer containing polygons, such as the gallery rooms layer, is required.

Push values to layer



If we click 'OK', and switch to the gallery layer view, by clicking on 'gallery' in the side bar, we see that there are two new columns: 'Visual Integration [HH]' and 'Object Count'.
'Visual Integration [HH]' is, according to the option we chose, the maximum visual integration found within each room polygon.
'Object Count' is the number of points that intersect each room polygon.

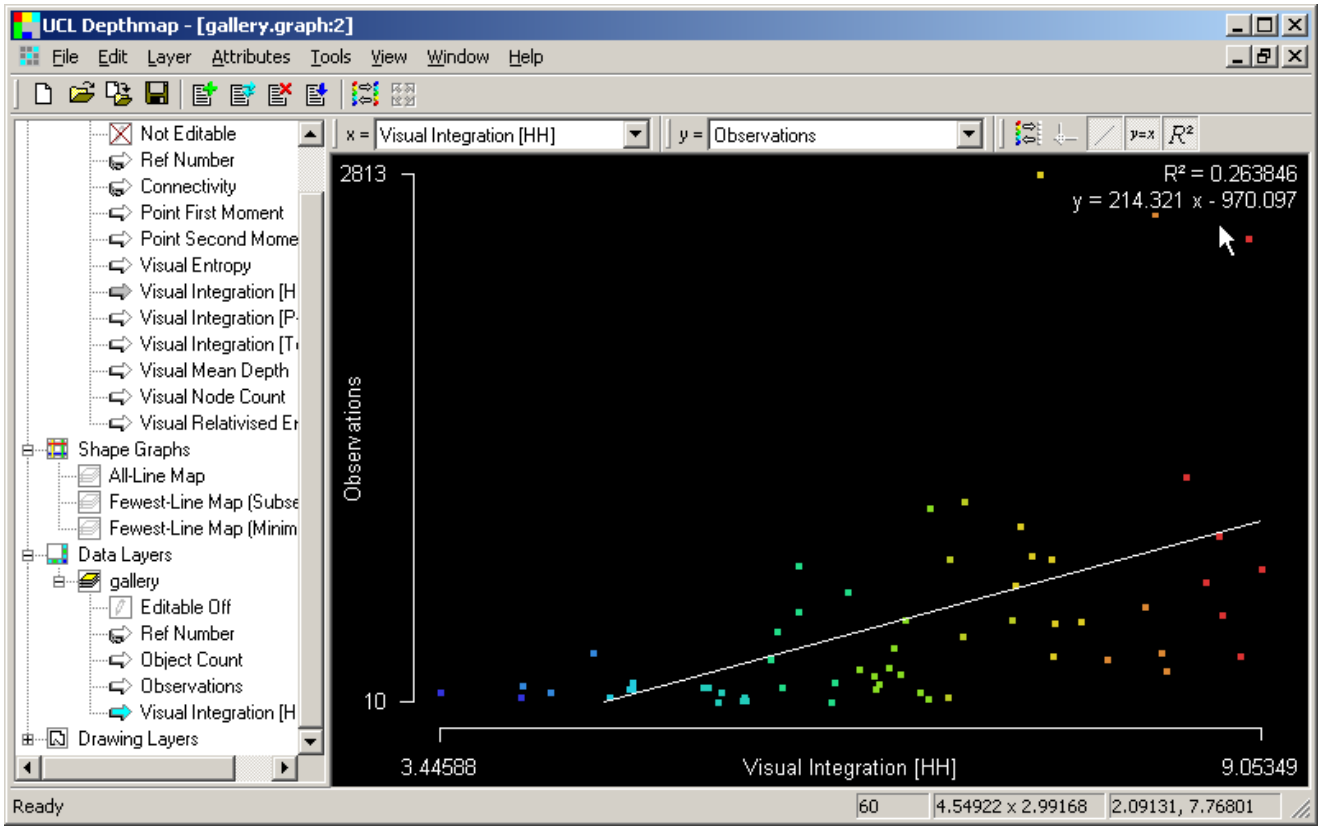
Cross map comparisons



Now that we have data from both observations and VGA in the same layer, we can compare them using the scatter plot window, by selecting ‘Scatter Plot’ from the ‘Window’ menu.

There seems to be a positive trend, if not very convincing, but how do we quantify that trend?

Cross map comparisons

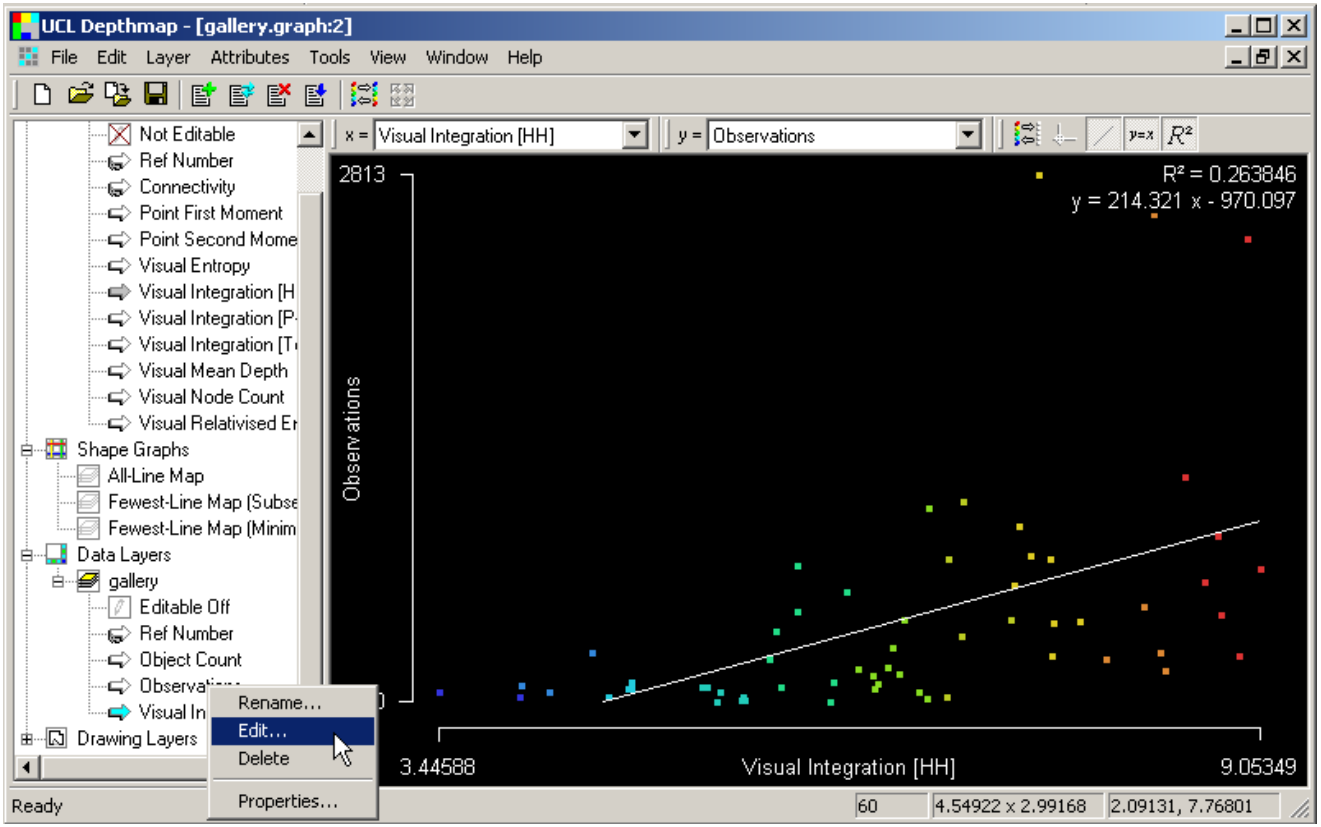


The buttons on the scatter-plot window tool bar give various options for the scatter plot.

- The 'Trend Line' button shows a best fit line through the data
- The $y = x$ button shows the equation of the best fit line in the form $y = mx + c$
- The R^2 button shows the R^2 correlation coefficient for the data.

In this case $R^2 = 0.26$ and $y = 214x - 970$.

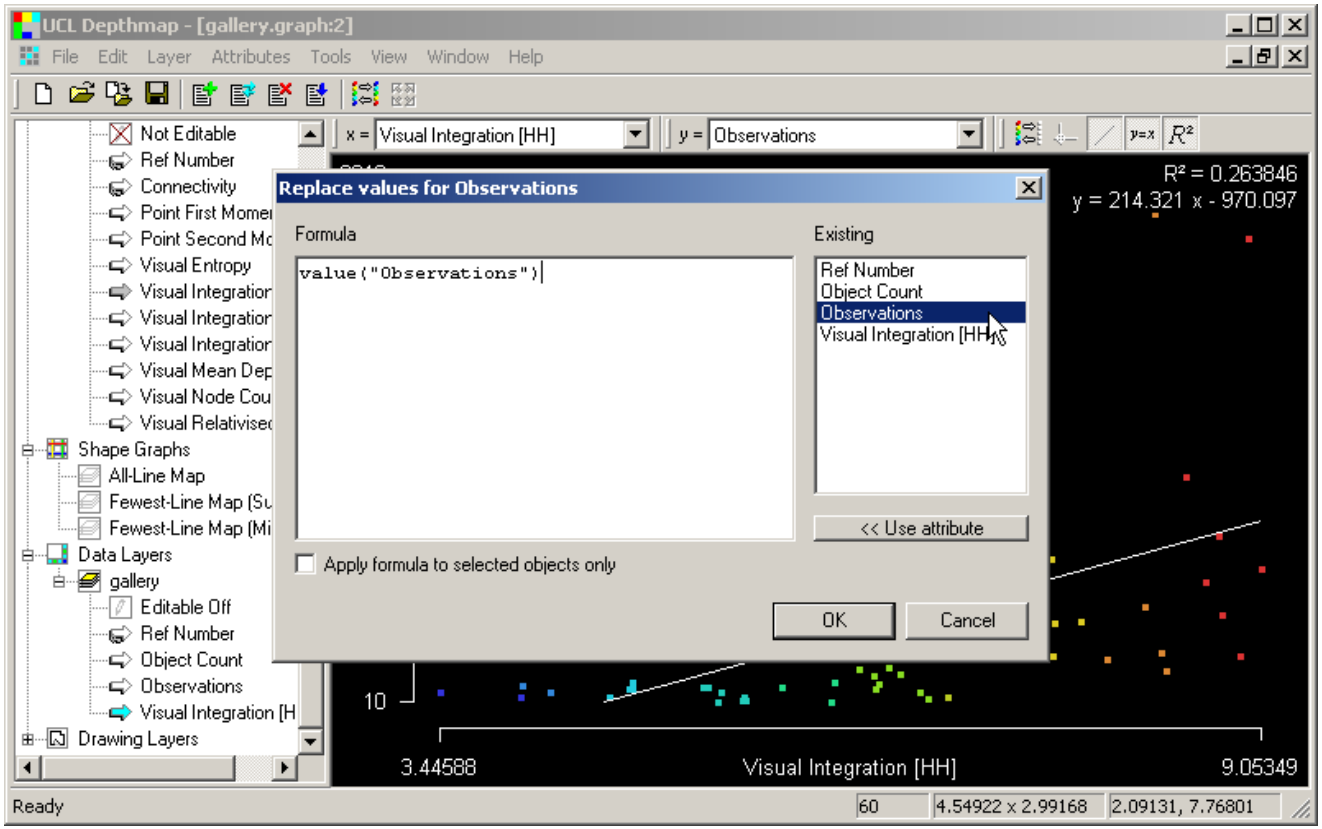
Rescaling an axis



One thing to not about these data is that although the integration data are distributed roughly evenly, the observation data are not. This is a concern for linear correlation, which assumes variables are evenly distributed.

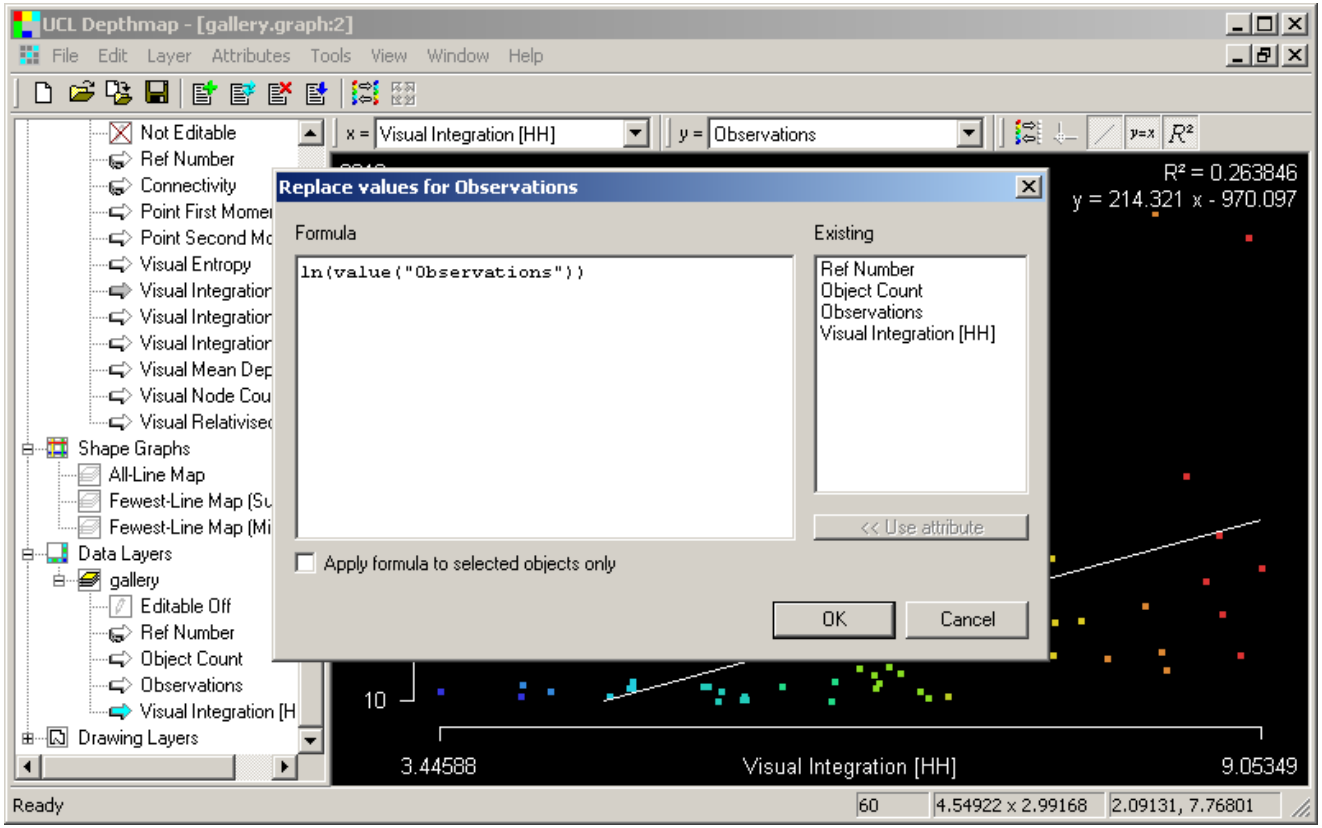
We would be best to convert the observation data to a logarithmic scale. To do this, right click on the ‘Observations’ attribute in the side bar, and select ‘Edit’.

Rescaling an axis



Check that the dialog title bar reads ‘Replace values for Observations’, and not ‘Visual Integration [HH]’.
Now double click on ‘Observations’ in the right hand window. The text `value(''Observations'')` will automatically appear in the entry box.

Rescaling an axis

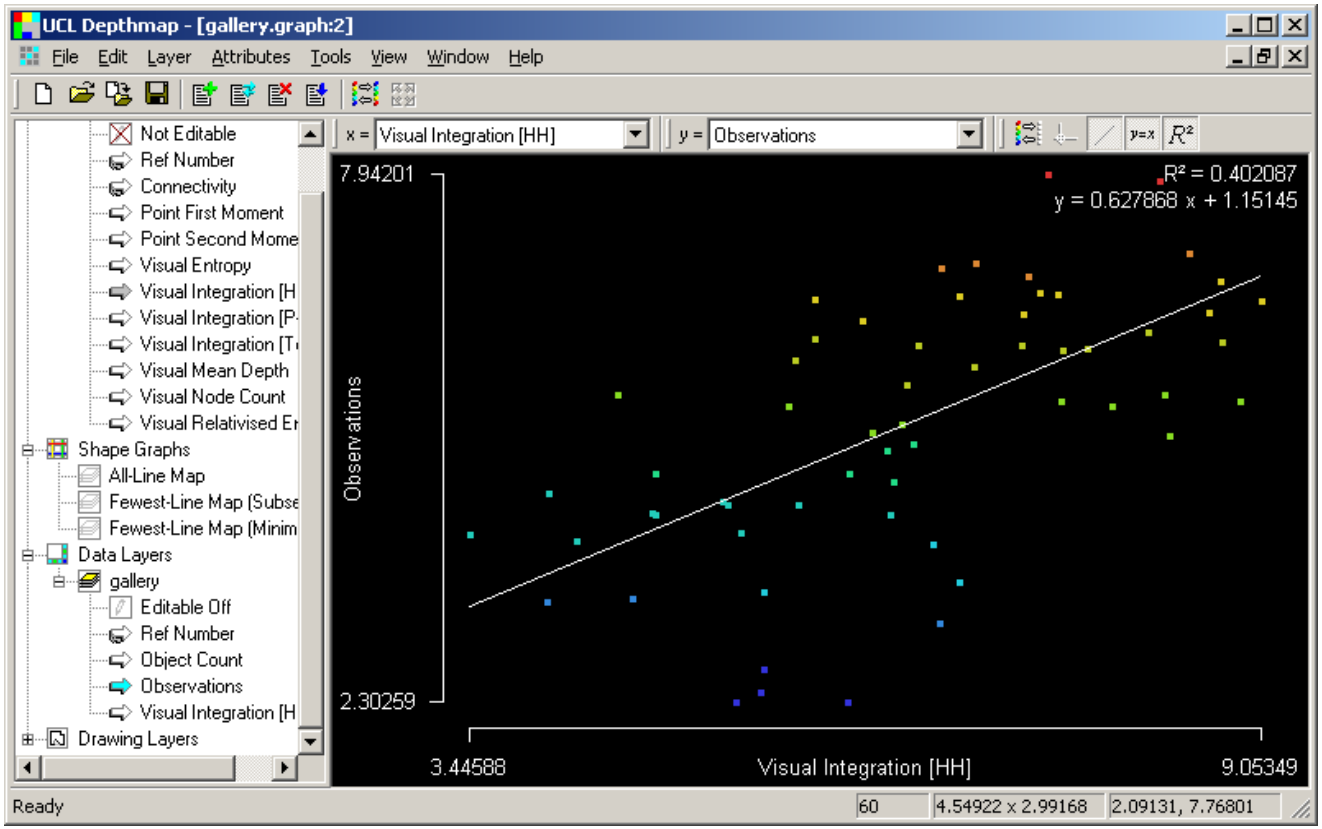


Next alter the entry box text so it reads `ln(value(''Observations''))`, and click 'OK'.

This means: take the natural logarithm of the observation values, and replace them in the observation column.

It might have been best to create a new column to store these values rather than overwriting the originals!

Rescaling an axis

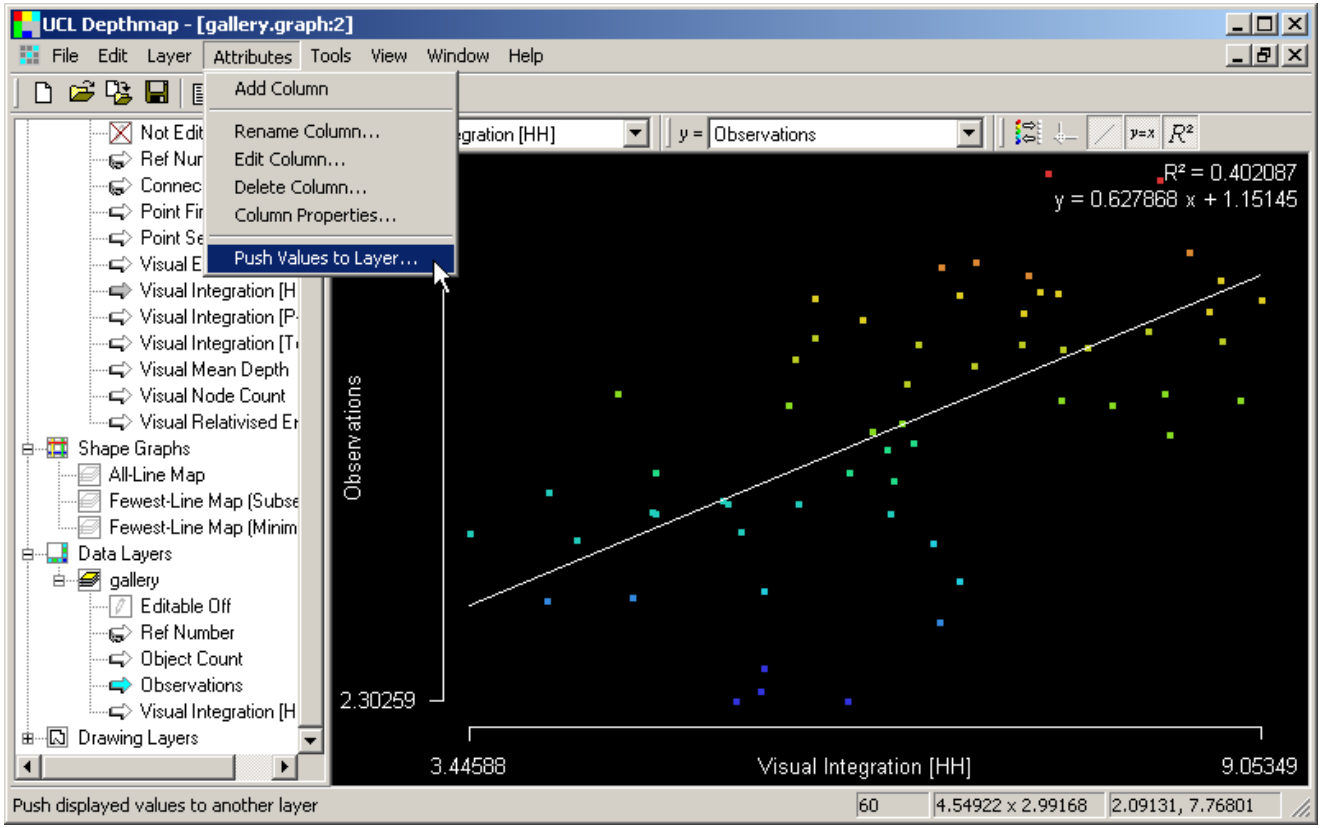


Now both scales are distributed roughly normally, a comparison of the two is possible.

The correlation coefficient R^2 is 0.40.

That is, there is some correspondence between the two, but not striking. This is expected in space syntax theory: a much better measurement of movement is provided by agent-based analysis, not visual integration. For details, see 'To move through space: Lines of vision and movement'.

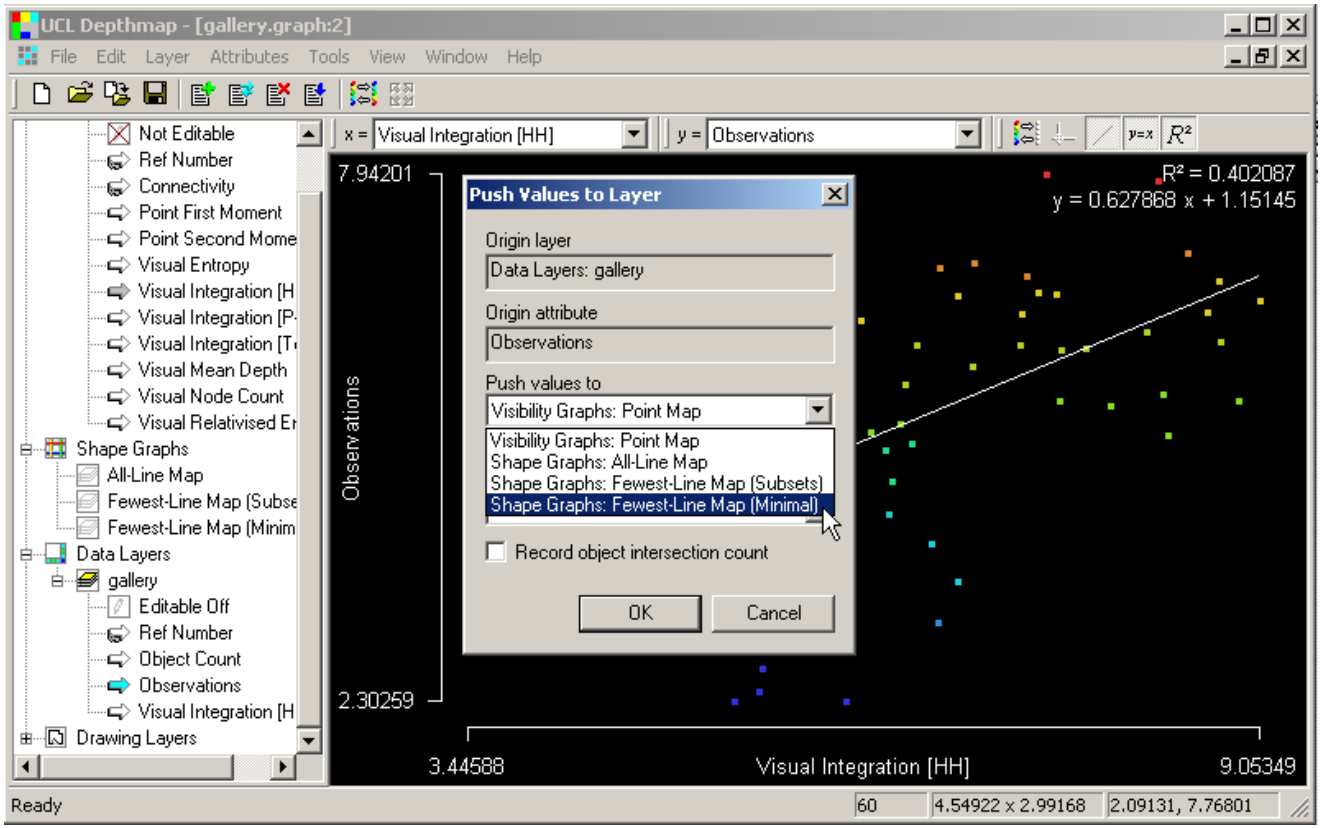
Pushing data the other way



Note that we do not have to just push data from the VGA layer or axial layer to the data layer. We might, for example, push our data layer to the axial layer.

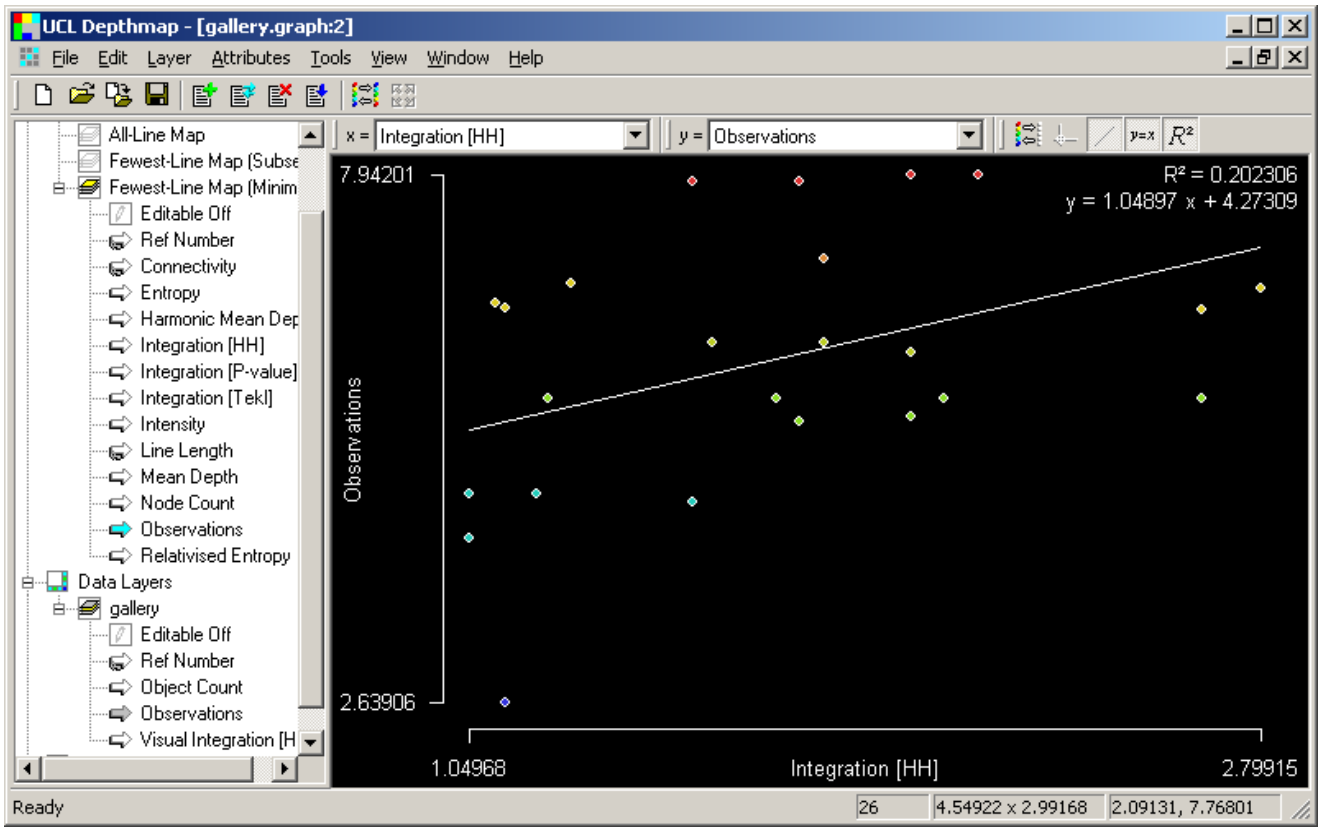
Select ‘Push Values to Layer’, either from the main tool bar or the ‘Attributes’ menu, while the ‘gallery’ layer is uppermost.

Pushing data the other way



I shall push the ‘Observations’ values to the ‘Fewest-Line Map (Minimal)’ axial map.

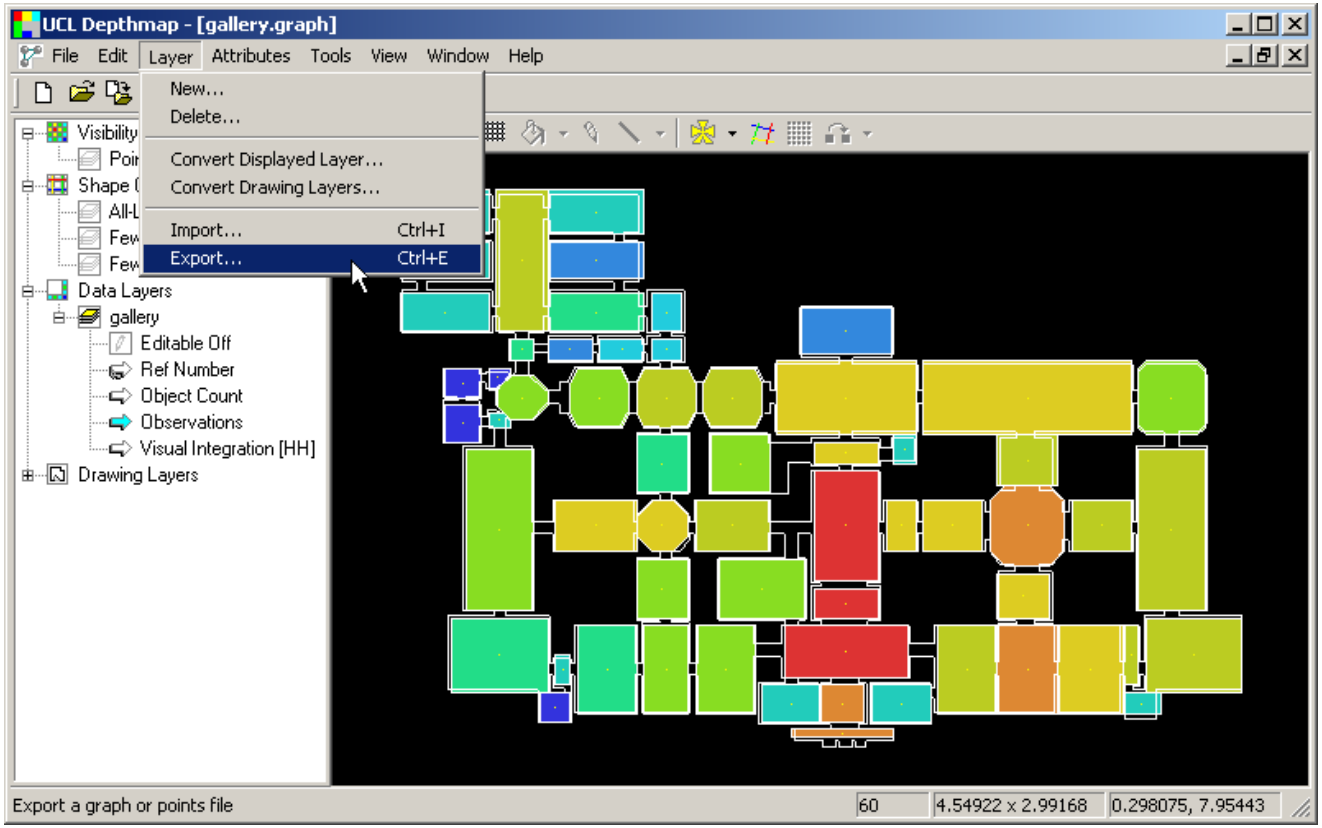
Pushing data the other way



Now I can compare 'Observations' to my axial integration. $R^2 = 0.20$, which goes to show that axial line integration *within a building* is not well related to pedestrian movement[1].

[1] Also note that the data analysis methodology here is flawed: several observation data have been assigned to a single line.

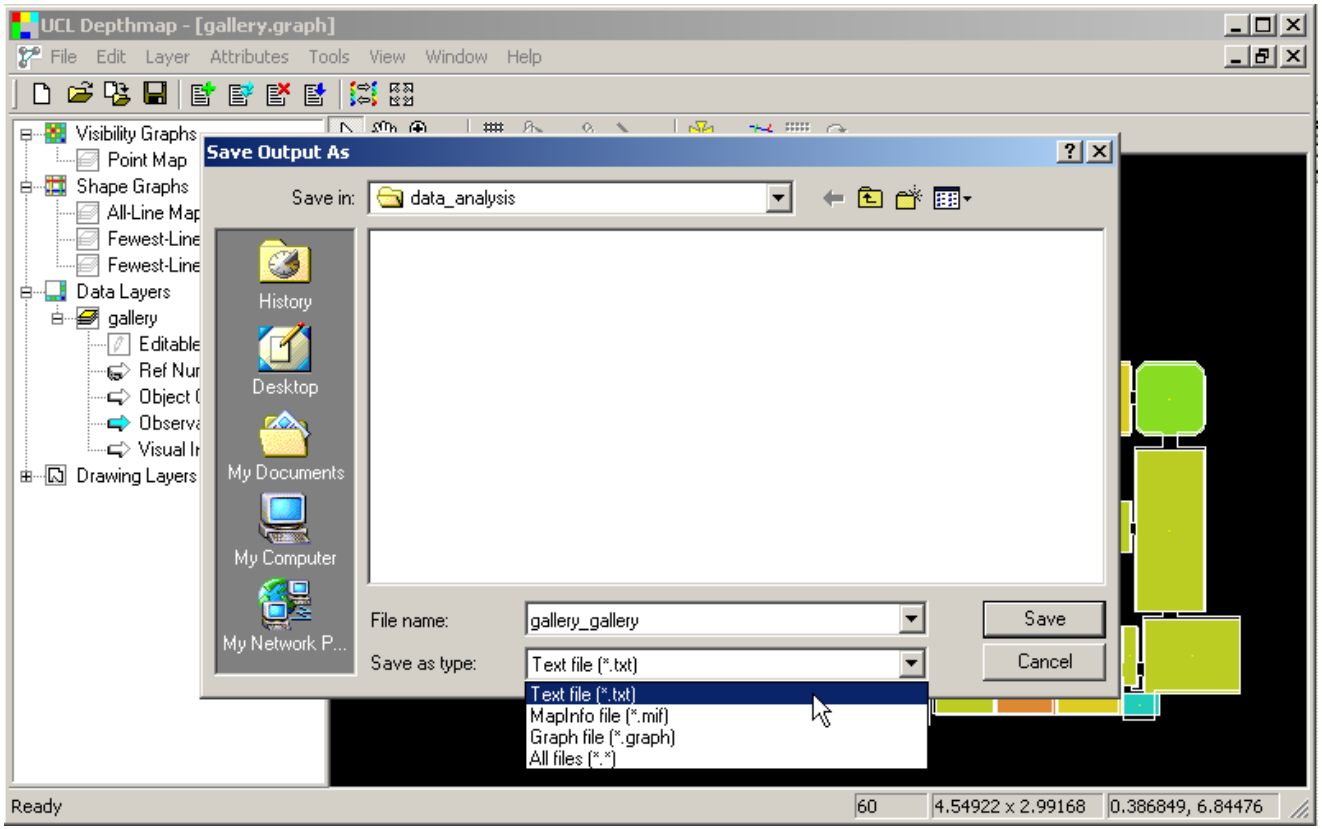
Export data



Of course, there are professional statistical analysis packages which can do far more with the data than Depthmap. To get the data into these programs, Depthmap includes an ‘Export’ option on the ‘Layer’ menu.

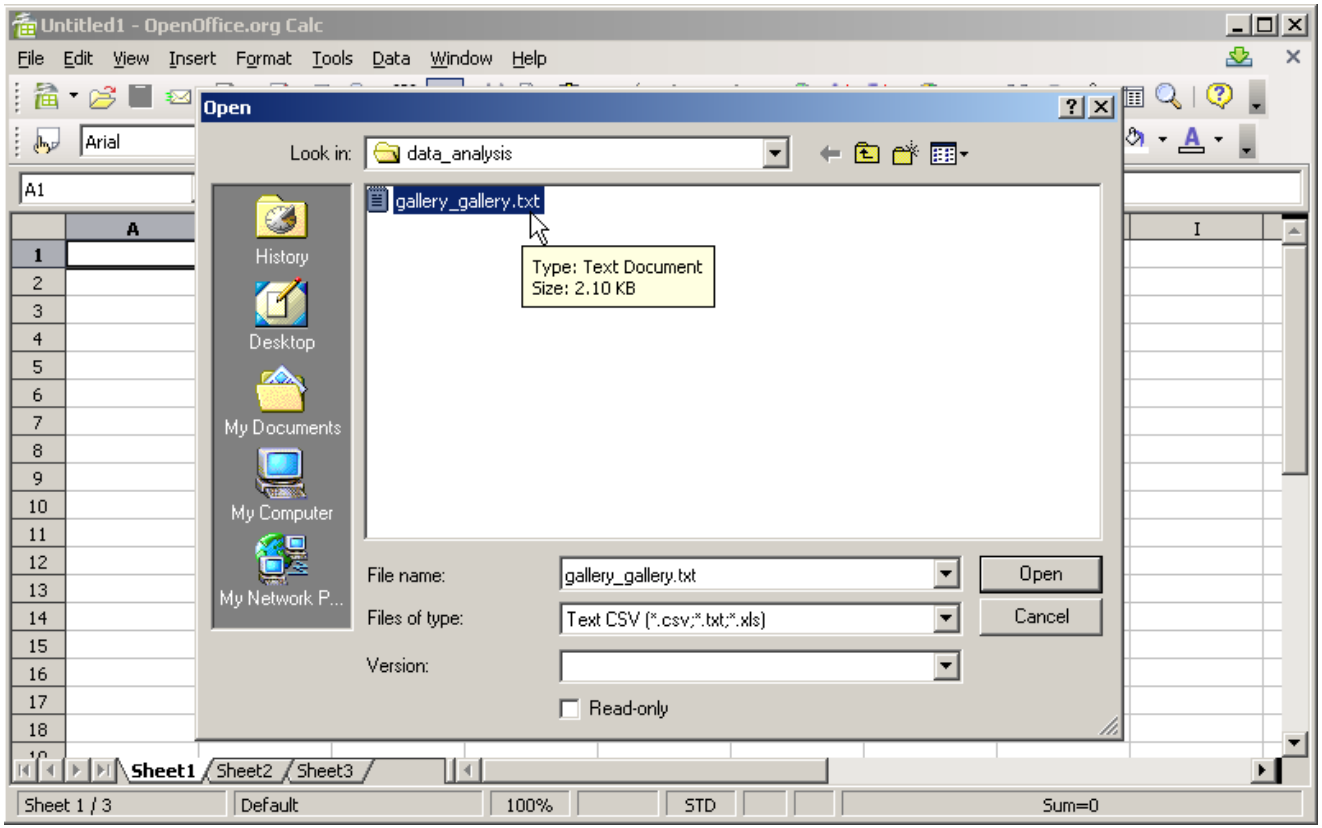
This will export the data from the currently topmost displayed layer: in the case shown, the gallery room layer data.

Export data



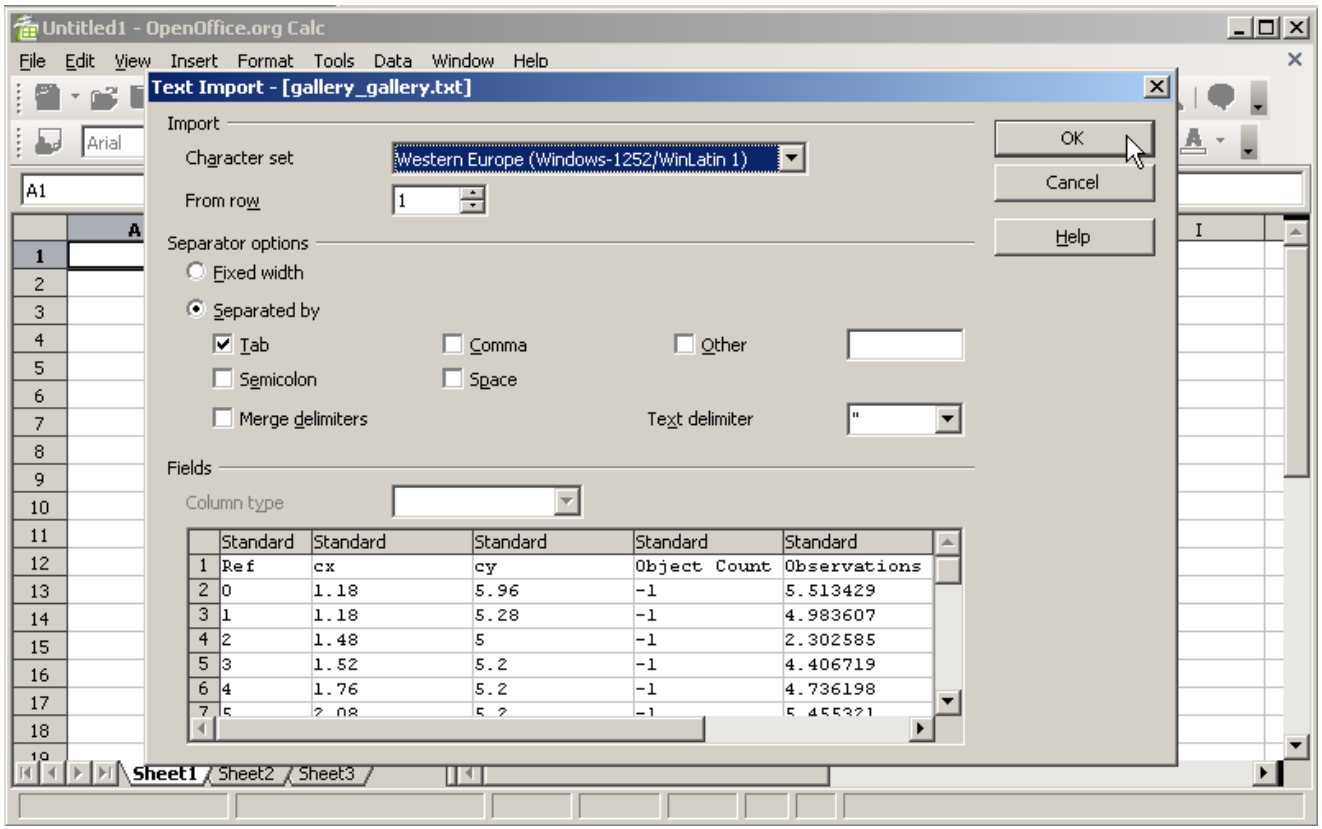
The default save option, a `.txt` file, produces a tab-delimited text file.

Export data



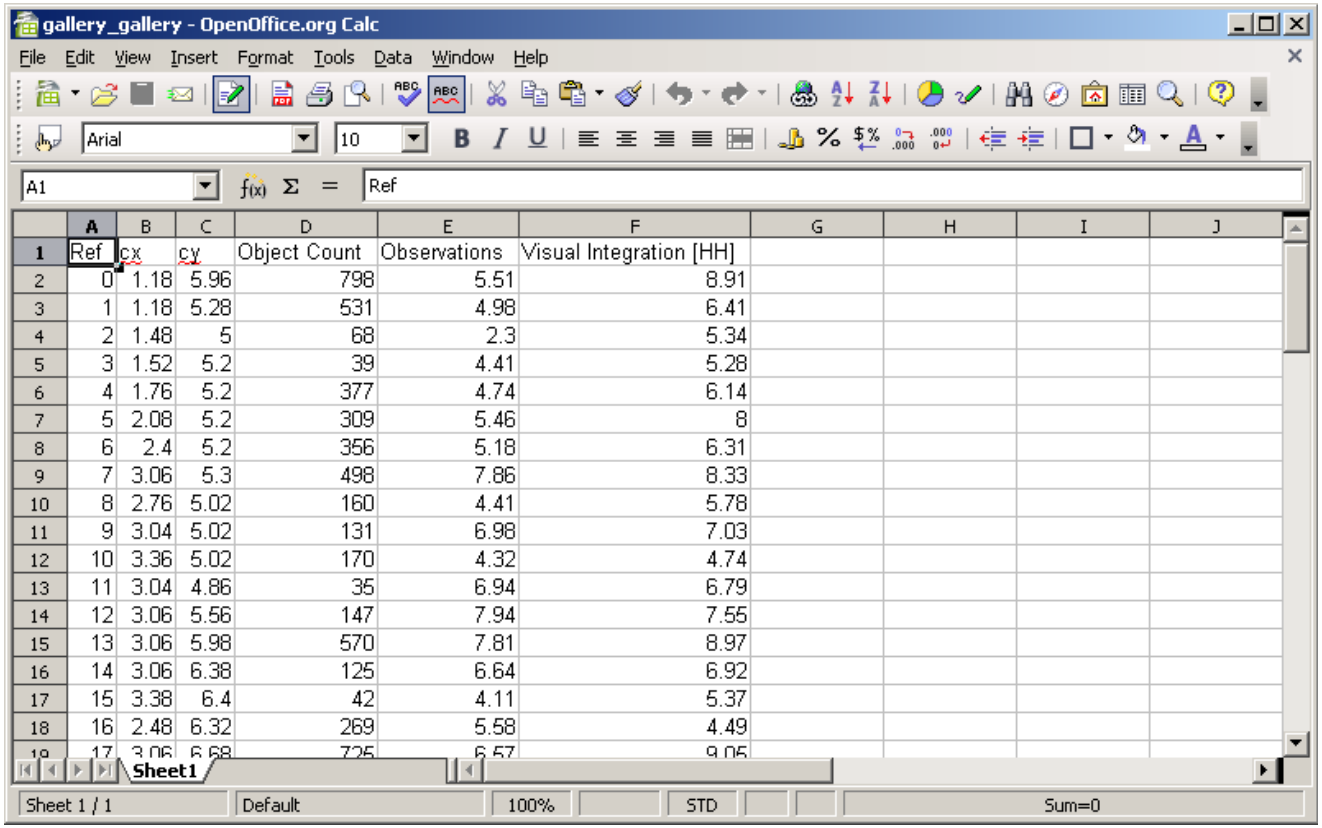
The text file can be imported easily into most data analysis and spreadsheet packages.

Export data



Open Office (shown here) and Excel detect the correct options to import by default.

Export data



	A	B	C	D	E	F	G	H	I	J
1	Ref	cx	cy	Object Count	Observations	Visual Integration [HH]				
2	0	1.18	5.96	798	5.51	8.91				
3	1	1.18	5.28	531	4.98	6.41				
4	2	1.48	5	68	2.3	5.34				
5	3	1.52	5.2	39	4.41	5.28				
6	4	1.76	5.2	377	4.74	6.14				
7	5	2.08	5.2	309	5.46	8				
8	6	2.4	5.2	356	5.18	6.31				
9	7	3.06	5.3	498	7.86	8.33				
10	8	2.76	5.02	160	4.41	5.78				
11	9	3.04	5.02	131	6.98	7.03				
12	10	3.36	5.02	170	4.32	4.74				
13	11	3.04	4.86	35	6.94	6.79				
14	12	3.06	5.56	147	7.94	7.55				
15	13	3.06	5.98	570	7.81	8.97				
16	14	3.06	6.38	125	6.64	6.92				
17	15	3.38	6.4	42	4.11	5.37				
18	16	2.48	6.32	269	5.58	4.49				

For polygonal data, the file contains columns `cx` and `cy` which are the coordinates of the centroid of the polygon. For point data, the point location `x`, `y` is exported, and for line data, the end points `x1`,`y1` and `x2`,`y2` are used.

Conclusion

This tutorial has covered entering observation (or other) data into existing Depthmap geometry, displaying summary statistics for columns, showing scatter plots and correlation coefficients, as well as exporting data to spreadsheets and data analysis packages.
In doing so, it has introduced the key concept within Depthmap of pushing values between layers.
Note that the tutorial has not covered the principles of statistical analysis, and the results shown in the tutorial are for guidance purposes only.