Dear ManyEyes Team,

I used ManyEyes to create several visualizations of crime report numbers reported by local police jurisdictions across the USA. I found that setting up and using my dataset within ManyEyes was, for the most part, painless and intuitive. I was thankful for ManyEyes's ability to detect input data types without having to manually specify each column's data type, as this saved a lot of unnecessary clicking and selecting. In addition, the “flip rows and columns” button allowed me to quickly orient the input data in the proper way for ManyEyes to visualize it. The visualizations themselves were also simple to set up, and the documentation included with each visualization put in plain terms how (and more importantly, when) to use each. I wasn't overwhelmed with setting range parameters and limits, which was helpful for a visualization neophyte such as myself.

I also found it useful and fascinating to examine the visualizations created by other users of the site, and create my own visualizations of their datasets. Sometimes I would find a visualization that I thought could be better explored using another method, and I could use the hyperlinks available on each visualization page to create alternative representations of the same dataset. Most importantly, creating these new visualizations was an effortless process and did not deter me from exploring the data in creative ways. I was also impressed by the visualizations' professional look, such as the way data points smoothly flowed to new positions when visualization parameters were changed (e.g., in the histogram and scatterplot), instead of an instant refresh and subsequent loss of context. Highlighted data points also remained highlighted when changing parameters, which was useful for tracking data across several parameters.

I did have some trouble importing my dataset into ManyEyes when pasting my dataset into the text box, which was required for data input. Specifically, when pasting the data from a Linux machine, the data seemed to be imported correctly, but when trying to visualize the data, ManyEyes reported that the dataset was empty. I suspect the problem could be with Unix versus Windows-style newlines, but that's just speculation. One simple way to avoid these platform or browser issues is have an “import from file” option, to allow uploading and processing of a dataset file in the same format as the text box is currently. In addition, I couldn't find any button or field to access certain metadata for each dataset, such as the dataset's description.

One feature I found myself needing over and over was a quick way to find individual data points in my dataset quickly. For example, in the histogram and scatterplot visualizations, I wanted to find crime rates in New York City, but there wasn't an easy way to find them. In the histogram, the labels for each data record were being cut off because the boxes were small, and the labels were being truncated, while the scatterplot had no labels at all. There is a “Search” text field for the histogram visualization, but it didn't seem to work consistently – I ended up having to hunt with the mouse pointer. A simple solution for finding individual data points might be to have a clickable list of labels, one label per data point, where clicking a label or labels in the list highlights the corresponding points on the visualization. The bar graph visualization already has such a list, but it would be useful for other visualizations as well, and should not be too hard to implement. I also found myself wanting to track multiple data points with shared characteristics at the same time. However, there was no quick way to select multiple data points at the same time. One way to enable this functionality could be a click-and-drag box, in the same way you would select multiple files on your computer's desktop. This would save having to Shift + Click for the potentially large number of points you are interested in.

ManyEyes would also benefit from additional, visualization-specific statistical features. For example, for the scatterplot, I wondered what the best-fit line for the data was, as well as other measures such as the correlation and entropy between attributes. Another suggestion I have is to allow users to access certain underlying visualization parameters, such as axis scales (e.g., for the scatterplot) or bucket sizes (histogram). Parameter sliders, such as the small histogram-slider used in the Treemap visualization, would allow the user to find the best parameters by simply playing with the data values until they see something interesting. A potential drawback, however, is that new users could be overwhelmed with widgets, so a happy balance should be found.

In conclusion, I found ManyEyes to be a very promising and useful tool that helped orient me to the world of information visualization, and quickly find interesting parts of my dataset. Keep up the good work!

Sincerely,

Mike Lieberman, Esq.