Bubble Plots

Interact with a Bubble Plot

Using longitude and latitude for the x and y axes on a bubble plot

Have you ever wanted to color points on a scatter plot by some other variable? I do it often, by using the Row Legend command. This way, not only are the points colored, but a legend is added too.

As an example, let's create a scatter plot for which the points correspond to zip codes in the continental United States. The x and y coordinates are the longitude and latitude of the zip code, so the scatter plot forms a rough approximation of the country. The data is not current, which is a partial explanation of the sparse data in the western half of the country. Another reason for the sparsity might be that the land areas associated with those zip codes are larger.

Say I want to color the points by another variable: the proportion of the county that is covered with water. I right click on the scatter plot and select Row Legend. The Mark by Column dialog appears. Choose the column which you want to color by. In this case, I want to color by PropWater.
The default JMP color scheme for a continuous variable is called Blue to Gray to Red, and is shown on the picture above. Notice that not only can you color the points according to a column, but you can change the marker as well. The default marker is the regular black dots shown on the scatter plot. For this example, we’ll only color the points.

This color scheme goes from blue (low proportion of water) to gray (middle proportion of water) to red (high proportion of water). The color scheme has various shades of the colors along the spectrum, to account for the fact that Prop Water is a continuous factor. As expected, most of the country is blue. Those counties with more water (the grays and reds) are there, but are hard to isolate and view on the scatter plot.

If your purpose is to highlight and focus on those counties with a higher proportion of water, then one of the other JMP color schemes might work better. The other color scheme options can be accessed from the Colors menu on the Mark by Column dialog.

Let’s try the White to Blue color scheme. Counties with little water will be closer to the white end of the spectrum, and counties with lots of water will be colored closer to the blue end of the spectrum. That way, the counties with lots of water will stand out, and will be easier to see on the scatter plot.

As you can see, by using a different color scheme, we are able to focus the attention on the counties with lots of water. Notice the blue spot all by its lonesome in the Utah area. That spot didn’t stand out as quickly using the default color scheme. My guess is that the high proportion of water is due to the Great Salt Lake and/or Utah Lake.
The color red is often associated with negative/bad values. If a high proportion of water was somehow a bad thing, then you may want to use the **White to Red** scheme, as shown here.

As we can see from this example, JMP’s color schemes are a powerful asset when visualizing different aspects of your data!