Mathieu Rolfo CS448B Assignment One



Minnesota Barley Yields by Location, 1931 vs. 1932



These visualizations were created using the online software tool Infogram.

For this assignment, I chose to depict the 1931-1932 barley yields from various Minnesotan plots in two grouped bar charts. The main goal of my visualization was to depict the decline of barley yields in 1932, while showing the outlier of Morris to not obscure the facts. The use of "vs." in the title was intended to signal this focus to viewers. I chose the grouped bar chart as my format because I felt it was able to depict the decline of barley yield between 1931 and 1932 in a clear format. Area graphs and line graphs were considered, but in my opinion they disguise the temporal element of change over time between values from 1931 and 1932. Also, the continuous line in each of those formats suggested that the graph was a time-series, which would have been misleading.

I chose to create two separate grouped bar charts rather than one after testing both formats. When on a single chart, it was difficult to understand the difference between variety and location. Also the sheer number of lines obscured Morris' yield increase in 1932, the main point of displaying location, and a separate graph made this easier.

I chose two easily distinguishable colors (tested to appeal to color blind and regular-sighted individuals) for 1931 and 1932 values. This makes the difference between years, for a single variety/location or across the whole graph, easy to glean. Horizontal lines were added to make this even easier. For additional clarity, I provided averages within the plots. I chose to display the x-axis in alphabetical order to avoid incorrect interpretations when people scan left-to-right across the graphic, and indicated this in the axis label for clarity. Additionally, the y-axis runs from 0 to the maximum value to avoid creating a distorted image for viewers in terms of height differences between bars.

I chose to use average values, rather than plotting individual data points, which led to some limitations of this visualization. I chose to do this because I felt the strong trends in the data, where every nearly every specimen for a given plot or variety increased or decreased, meant that using averages would illustrate the same point while being less complex than displaying all points in a dot/scatter plot. However, as a result this graphic doesn't explicitly tell viewers the strength of these trends; for example, these averages could have been skewed by a couple outlier points.