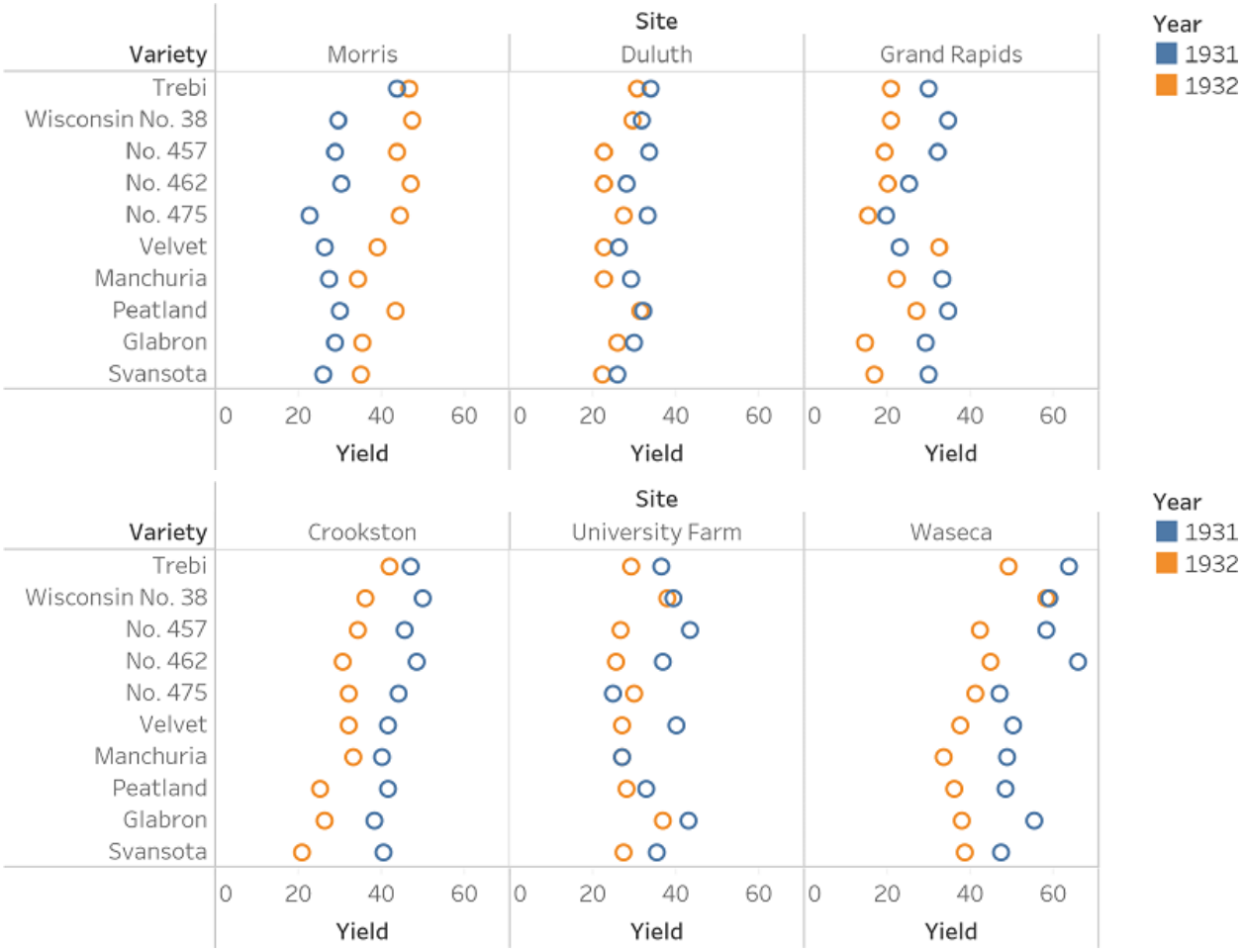
Anika Benons

CS448B Assignment #1

Minnesota Barley Yield (bushels/acres) in 1931 and 1932



Tool Used:

To create the above data visualization, I utilized Tableau. Tableau offered an easy manipulation of data, design, coloring, and other formatting functions that made it very convenient to use.

Design Decisions:

For the data visualization of the Minnesota barley yield data, I ended up utilizing all four given variables: site, variety, yield, and year. I believed all four fields contributed to the story the data was trying to tell which I believe is a comparison of the barley yield of each variety within each given Minnesota site in both 1931 and 1932. Therefore, emphasis was placed on the years 1931 and 1932 by giving a color code to each year and emphasis was placed on each site by giving each site their own section for barley yield data; this allows for an easy visual comparison for each variety based on the year within each site.

In addition, I had originally planned to place the data in a way that emphasized comparison between the sites rather than comparison within the sites. I decided against this as I concluded that such a comparison would be more meaningful if position (latitude/longitude) of each site was also considered. And because I did focus on comparison within the sites, the visualization design downplays comparison between the barley yield data of each site.

Moreover, I chose to use a circle plot since I believe it gave the most clear, precise design while still serving the purpose and story of the data. Because only two years were given, the circle plot does not create too much clutter, and with the varieties placed in a vertical fashion, along with the yield data for both years, visual comparisons are easy to make not only

between the years but also between the varieties within each site. Furthermore, the hollow circles used to plot the data allowed for overlaps to be more easily seen between the data for each year.

The yield data's units are included within the title as I thought it would create less clutter and two tables were used to get a more holistic view of the data. The yield data is set up to be plotted from 0 to about 70 bushels/acres to encompass the minimum and maximum yield and to make it visually easy to determine that the further left a circle is, the higher the yield.