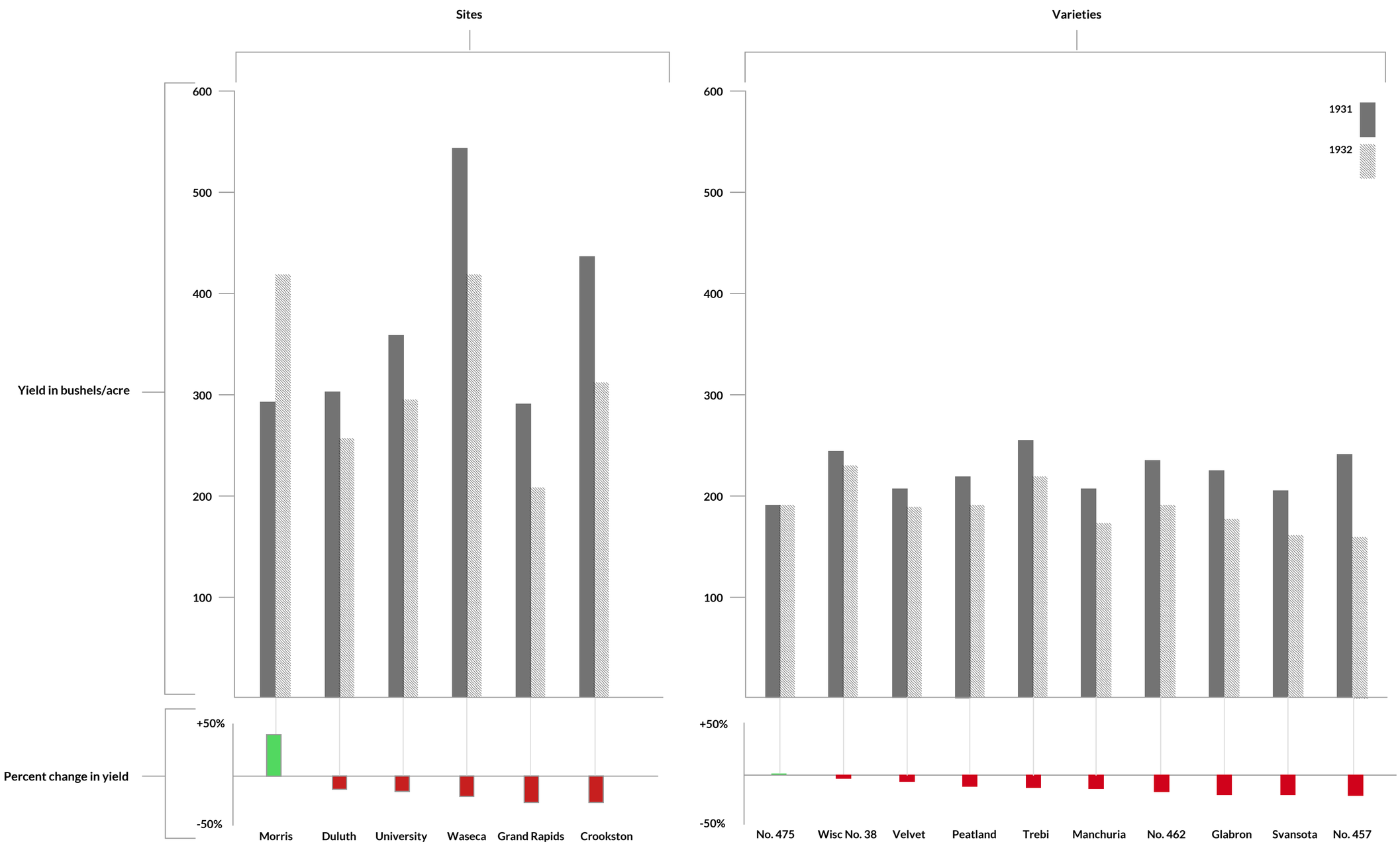
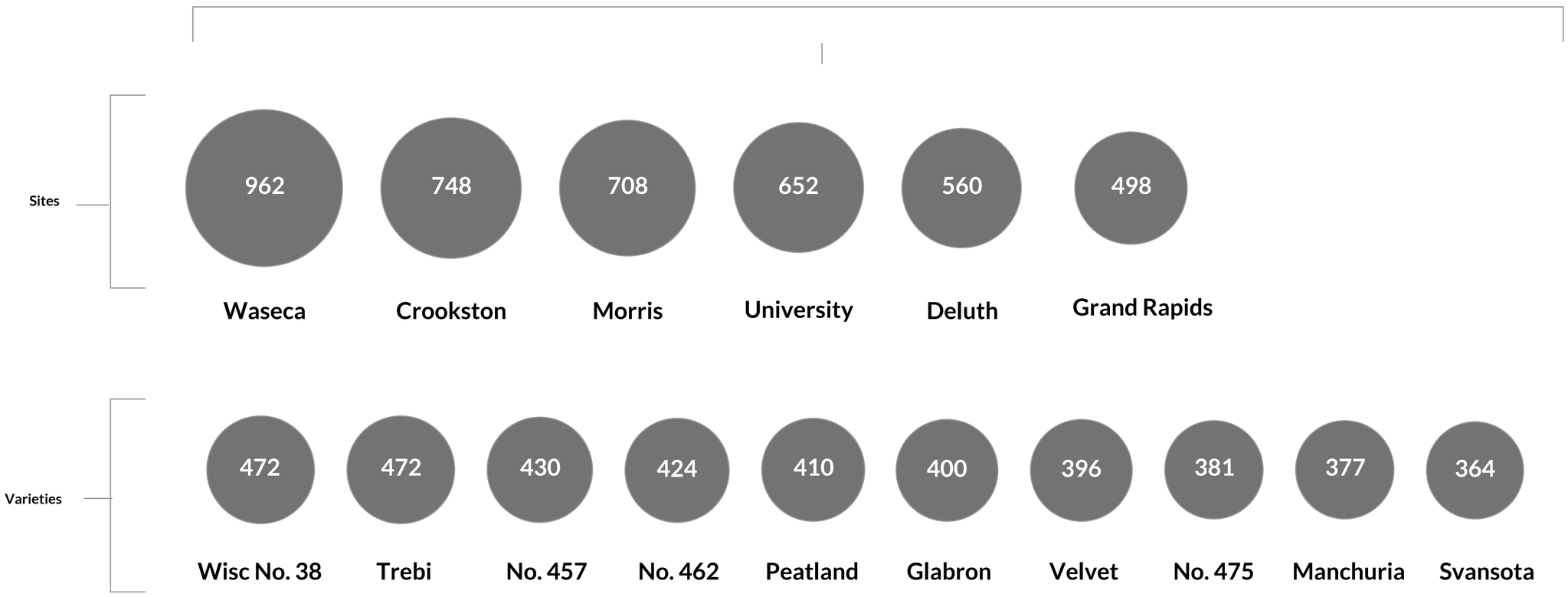
Bronson Duran
CS 448B: Data Visualization
A1: Visualization Design

Barley Yield Data Across Sites and Varieties From 1931 and 1932



Total yield in bushels/acre (1931 + 1932)



Bronson Duran
CS 448B
A1: Visualization Design

To begin, I asked myself what were the most important variables were in the data, as well as what were the most pertinent questions I should be asking about the data. It seemed that yield was the most important variable, thus I felt that the most pertinent questions were regarding how yield differs between farm sites and barley variations, as well as how the yield changes over time.

I chose to differentiate sites and varieties along the X axis because they are independent variables. I chose to display the yield and percent change in yield along the Y axis because they are dependent variables. I chose to display the yield for 1931 and 1932 next to one another in order to show change over time. I chose to color the positive percent change green because green has a positive growing connotation, and I chose to color the negative percent change red because red has an urgent or negative connotation. I chose not to color the sites or varieties as I felt that would be distracting for comparing quantities. I chose to order the the sites and varieties by greatest positive percent change in yield, as I felt that this was one of the most important factors in the data, which you can see displays the anomalous 42% increase in yield by the Morris site. I also felt that total yield was an important variable to display, so I used circles to display relative quantities. My goal was to allow the reader to quickly pick out the sites and varieties with the greatest total yields as well as with the greatest positive change in yield over time.

To create this visualization, I drew by hand the concept as well as the justifications. I then wrote a short Python script to calculate the numbers. Lastly, I used Sketch to create the visualization. There are improvements that could be made. For example, I do not show a breakdown of variations within a specific site. A more thorough visualization would show the percent change in yield for a specific variety within a specific site from 1931 – 1932, but I didn't have time to do this.