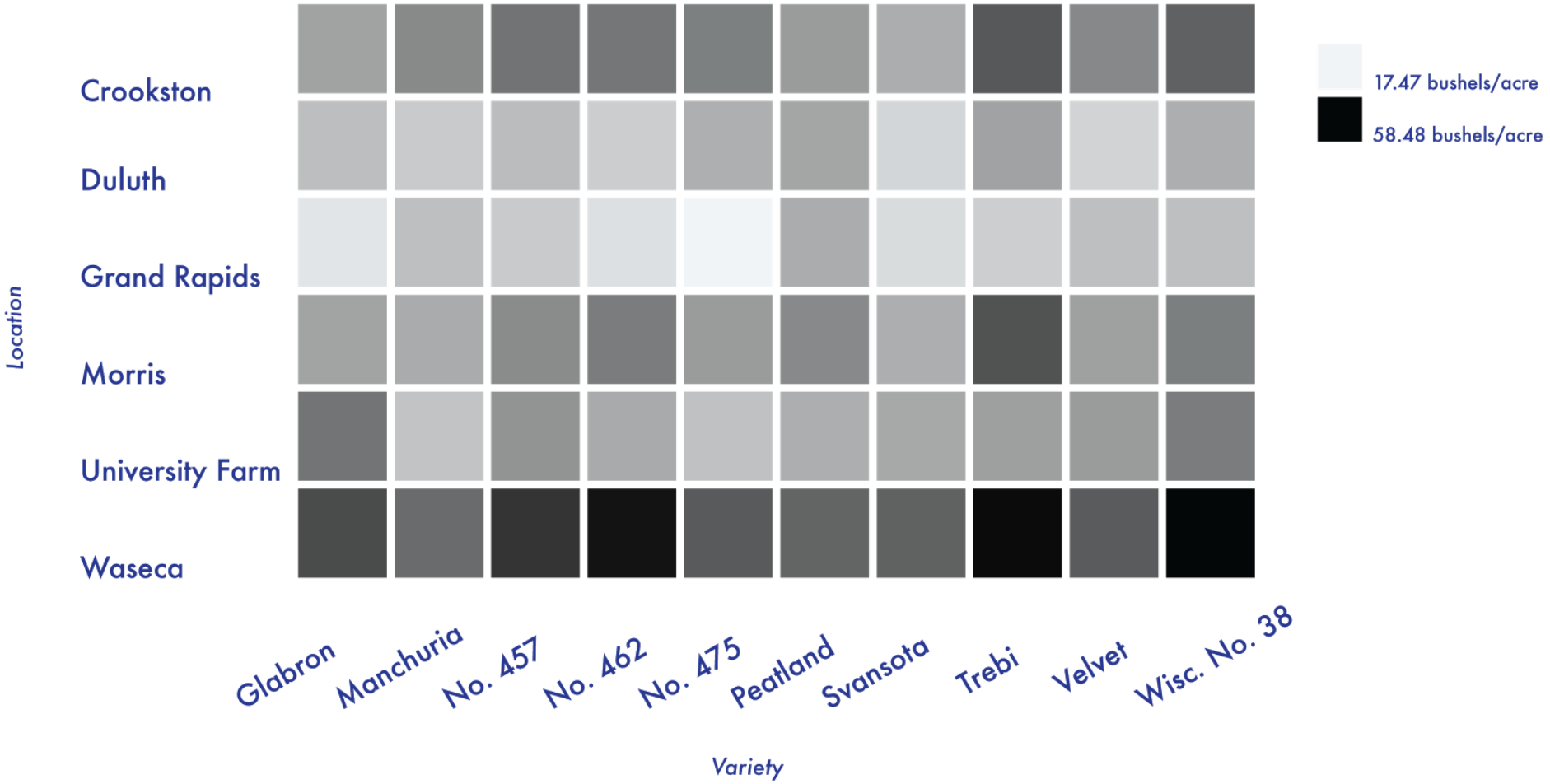
Relative Barley Yield

averaged between 1931 and 1932



Design Documentation
Karen Wang
CS 448B

I used Google Sheets to compute preliminary data filters and transformations, and Adobe Illustrator to produce the final visualization.

For this dataset I noticed that the differences in yield between locations was more pronounced than the differences between varieties of barley. Although the temporal component showed variation as well (1931 had a higher average yield than 1932), I felt that there were not enough data points (only two years' data worth) to make an interesting visualization. Therefore, I chose to focus on showing the variation in yield across the two-dimensional dataspace of location versus variety.

I sought to highlight this aspect of the data by visualizing the relative yields across all 60 location/variety combinations using a "heat grid." The rationale for this is that humans find changes in color intensity more intuitive to understand than symbolic representations. I also chose to use a grayscale instead of color scheme because humans can detect more variation in grayscale compared to color.

This involved some sacrifices in the resolution of the data. For one, I had to take the average of two years instead of the exact values. For another, I needed to transform the absolute yield values into an appropriate color representation. Using the HSB color scale, I kept the hue and saturation values constant while varying the brightness value to correspond to the yield values. In addition, I decided to sacrifice some precision for the sake of visual variation by taking the absolute yield values and re-computing them as brightness values using the formula below:

$$x_{bright} = 1 - [(x - yield_{min})/(yield_{max} - yield_{min})]$$

This takes the absolute yield value, represented by x, and scales it linearly between the global maximum and minimum. The scaled value is subtracted from 1 to match our intuitive interpretation of darker colors as higher values. This makes the cell corresponding the maximum value black, and the cell corresponding to the minimum value white (I ended up making it slightly grey to prevent it from looking like an empty space). In the final visualization, I also chose to make the text blue to not interfere with the grey/black value association, and I added a legend to clarify what the lightest and darkest cells represent.