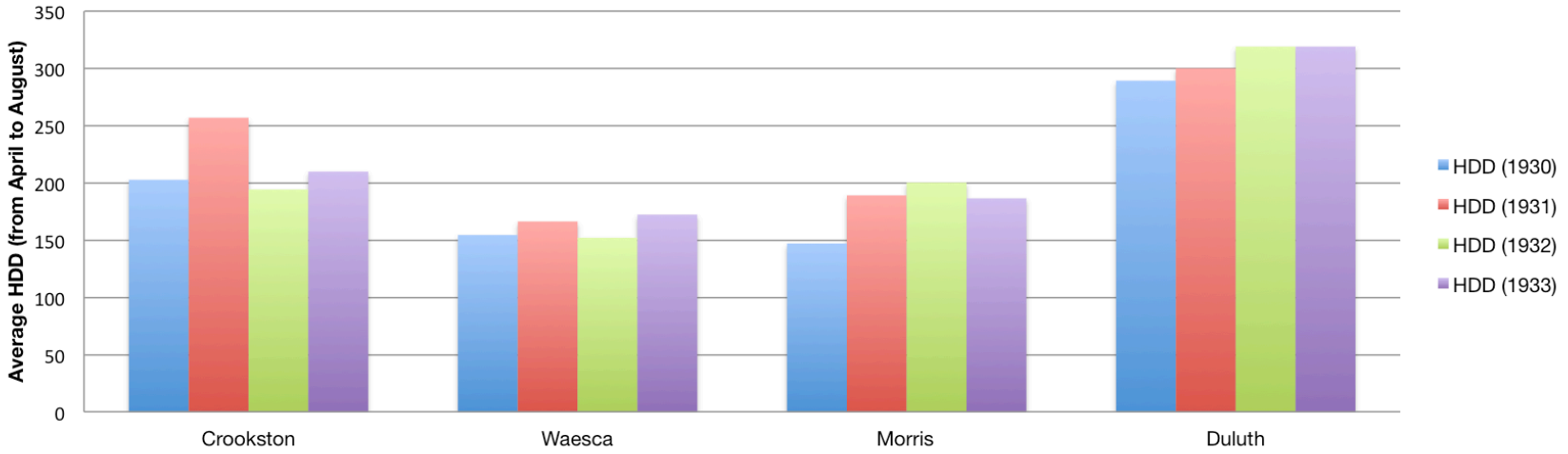
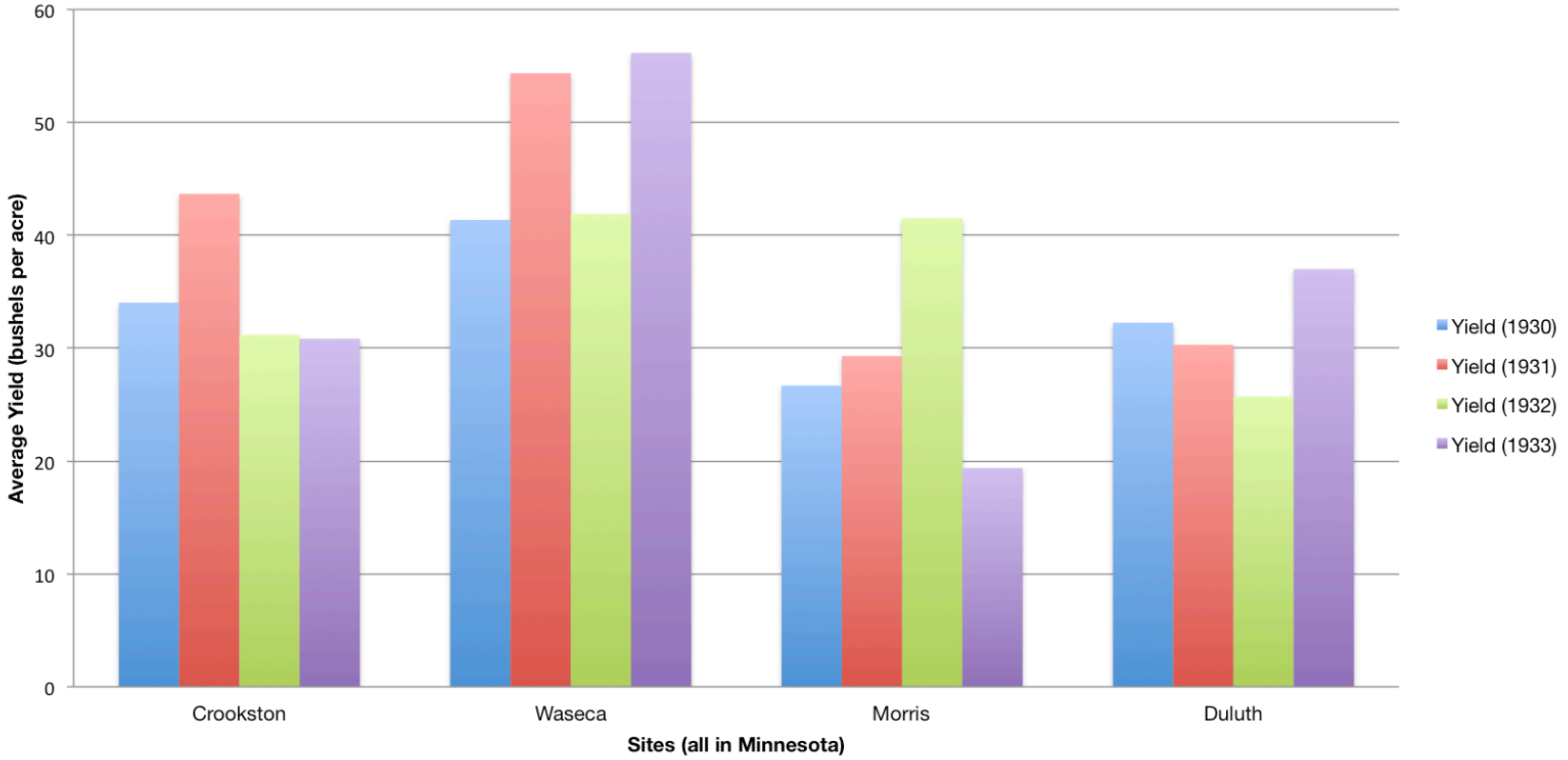
Wright, Kevin (2013). Revisiting Immer's Barley Data. The American Statistician, 67(3), 129–133.

of Heating Degree Days from 1930 to 1933 by Sites



Minnesota Barley Yields from 1930 to 1933 by Sites



HDD: number of degrees that a day's average temperature is below 65 Fahrenheit; degree of coldness *(higher HDD, lower overall temperature)*

In this visualization, I tried to express a possible correlation between climate factors and barley yield. In order to display the **relationship between low temperature and barley yield**, I augmented the original dataset with the barley yield data of 1930 and 1933 and climate report from 1930 to 1933. Hence, I aimed to show that higher barley yields are generally associated with cooler overall temperature. I used **Microsoft Excel** to view the data and Excel's "Chart" functionality to generate visualization.

Among several climate factors, I chose "heating degree days (or HDD)" as a measurement for coolness. In addition, since barley is grown from April until August, I took average of HDD of those five months per year. To display the relationships between HDD and yields in time series, I decided to mimic what Charles Minard did with his famous Napoleon's March graphic; I juxtaposed a bar graph that has average HDD for y-axis and site for x-axis on top of a bar graph that has barley yield for y-axis and site for x-axis. The third parameter, "year," was expressed with different colors. For example, 1930 data was presented with blue and 1931 data was presented with red in both graphs.

More precisely, I tried to use **x, y position cue**, one of the most intuitive cues, for denoting a pair of value (either HDD or yield) and site. By quickly glancing at x and y coordinate, any viewer can easily get HDD or barley yield of given site. Then, when expressing years, the important third variable, I used a visual cue that people tend to perceive easily, **color**. Applying consistent color for each year throughout the graphs, I emphasized the correlation between HDD and barley yields for each site. If one skims two graphs and compares them, s/he will notice that up-downs of two graphs throughout the years have very similar tendency for each site. For instance, Waesca shows a strong positive correlation between HDD and yield. When it comes to Waesca, the higher the HDD is, the higher the barley yield as well for every year. One can also notice that the year with the highest HDD reported the highest barley yield for every site. **The chart implies that there might exist a positive correlation between average HDD from April to August and average barley yield.** This even aligns with some research result (the research of Gunderson, Carr, and Martin in 2007).

In a nutshell, I used 1) **position**, 2) **color**, and 3) **parallel juxtaposition** of bar graphs to present a potential **positive relationship between yearly coldness and barley yields**. However, at the same time, the graphic does not show a very strong correlation and suggests that there might be other factors besides climate factors affecting barley yields.