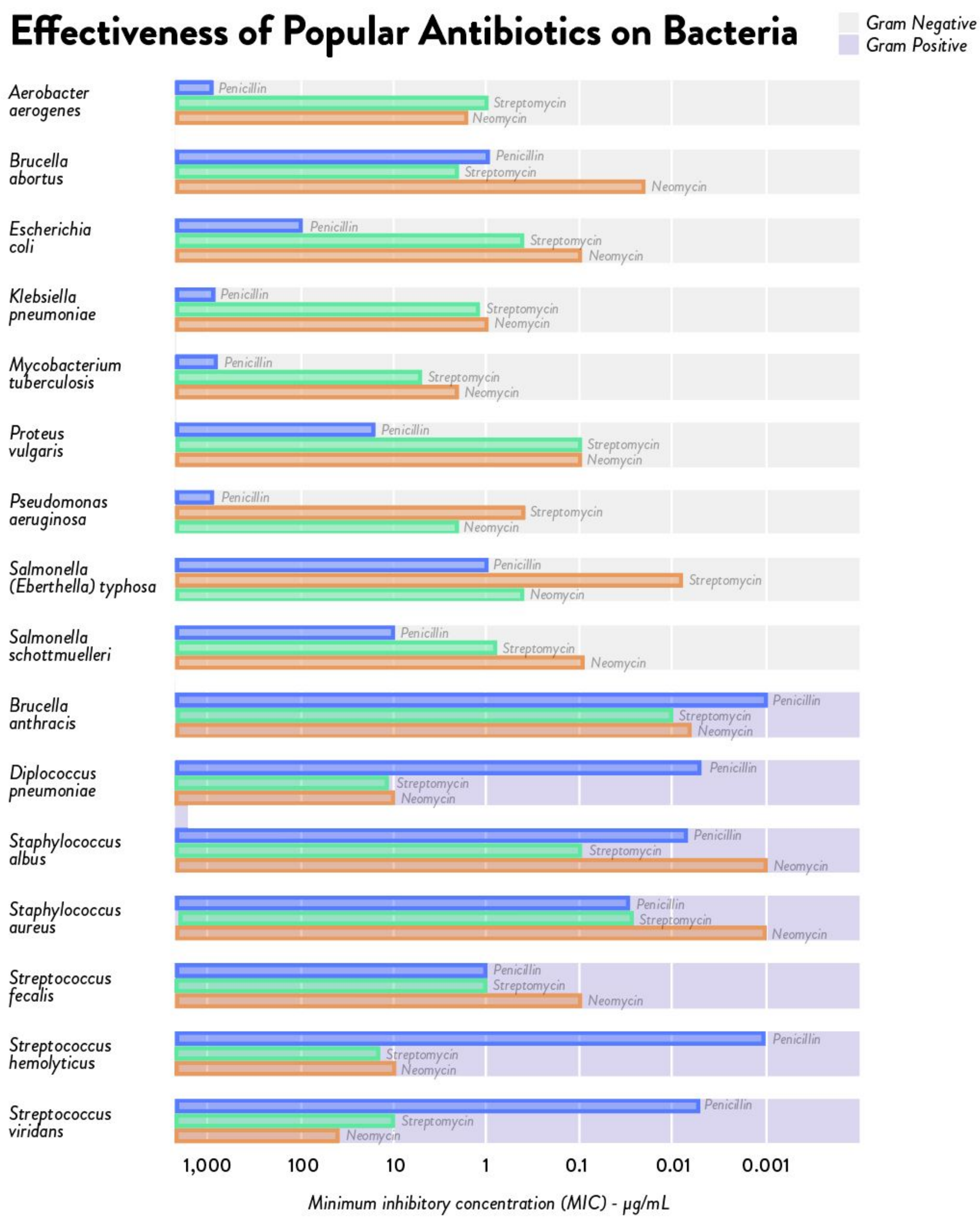
Thomas Li
EECS 396/496: Interactive Information Visualization
Assignment 1: Visualization Design



Design Decisions

Encodings/Mappings

- **Type of bacteria** is a nominal variable, mapped to y-position
- **Type of antibiotic** is a nominal variable, mapped to hue
- **Gram staining** is a nominal variable, mapped to hue
- **Minimum inhibitory concentration** is a quantitative interval variable, mapped to bar length

Summary

These decisions were made with the primary goal of preserving the goal of collecting this data and representing those relationships. The prompt states that this data was gathered “to learn which drug worked most effectively for which bacterial infection” of the tested antibiotics; with this in mind, I found it most important to allow viewers to focus on each bacteria type and easily perceive which antibiotic was most effective (or had the lowest MIC) of the three. I also wanted to point out that the numerical values for MIC vary widely but also have significant variance at low levels (0.001 v. 0.007), acknowledging that in order for those small numbers to look any different on a graph that I would have to be flexible with the scale of my visualization. This tradeoff allows the viewer to see the single best antibiotic for each bacteria most clearly by providing high levels of distinction for high-performing antibiotics.

I wanted to capitalize on our tendency as English speakers to read left-to-right, so that by reading one line the viewer could see one bacteria type and how all three antibiotics could be used for it. I also wanted to use position and length to represent the mapping between type of bacteria and MIC for each antibiotic, since they’re very effective visual encodings for these types of variables. It was important for the MICs of each antibiotic for the same bacteria to be grouped together for the ease of comparison, so I chose to use the next best visual encoding, hue, to distinguish between them (using colors chosen with Adobe Colors to avoid differences in value). I added the text so that the viewer wouldn’t have to refer to a legend in an arbitrary location each time, and reduced the opacity so that it would not distract from the larger data patterns. Gram staining did not seem to be the focus of this visualization, so I wanted it to be simply and minimally represented within the data; thus, I chose to use variation in hue (since it's still a good choice of encoding) but changed the subtle background colors rather than the colors of the data points themselves. I then sorted the bacteria types in alphabetical order based on their gram staining classification to minimize the visual effect of the gram staining color differences while preserving relative ease in finding a desired bacteria.

I also wanted to emphasize that the antibiotic with the lowest MIC was the most effective by thinking of the plotted data as a measure of effectiveness rather than just MIC. I did this by reversing the axes so that it started with the highest MIC value (1000) and ended with the lowest (0.001); this allows us to perceive the bars as "effectiveness" and to see the longest bar as the most effective. To partially soften the distribution of this data (which has differences that range from hundredths to hundreds), I chose a logarithmic scale in base 10 so that all data points and their differences would be visible. These decisions also allow the viewer to quickly analyze the data more easily by scanning the right side of the visualization, which isn't crowded by all of the bacteria type names.

Tools

I plotted the data using Microsoft Excel, and then copied the plot data itself into Adobe Illustrator to add the text labels and the hue.