

474 Assignment 1 Write Up:
Gabe Bizar

First, I utilized grouped bar charts to utilize the x-axis (position) to create categorized buckets for each of the strains of bacteria and how each of the antibiotics behave in accordance with that bucket (bacterial strain). The behavior of the strain in relation with the antibiotics is listed on the y-axis where the larger the value, the less effective the antibiotic is within that strain. Due to the sheer size of some of the effectiveness values (e.g. 800+), but frequent occurrence, I was unable to remove these points as outliers, but instead opted to logarithmically scale the y-axis to contain all the y-values, while maintaining the visual readability of having three-digit values and values in the thousandths place on the same visualization. Y-axis as positioning was ideal, as the effectiveness value was a quantitative variable and falls within Bertin's visualization principles, wherein quantitative variables such as effectiveness is easily represented by positioning.

Next, I used colors within each of the buckets (accompanied by the legend) to differentiate the different antibiotics because the different types of antibiotics are nominal variables that can be easily distinguished if given color. Color is ideal because antibiotics do not have a discernable rank/hierarchy and neither does color. This also falls in line with Bertin's design principles of allocating color to nominal variables, due again to the lack of discernable order.

Also important to note, the mixture of both position in color in the binary encoding of the positive/negative Gram Staining was used in creating a subplot directly to the right with opposing colors (but still sharing the y-axis). Such a dramatic x-position divide in addition to opposing colors is an appropriate representation of a binary encoded variable (positive/negative), as it falls within Bertin's design principles and offers a very dramatic way of showcasing a nominal variable without breaking the continuity/flow of the visualization. That is to say, when broken up into two subplots as opposed to one large conglomerate plot we are actually able to more easily discern variations at a glance than without the separation, but still maintain all the data.