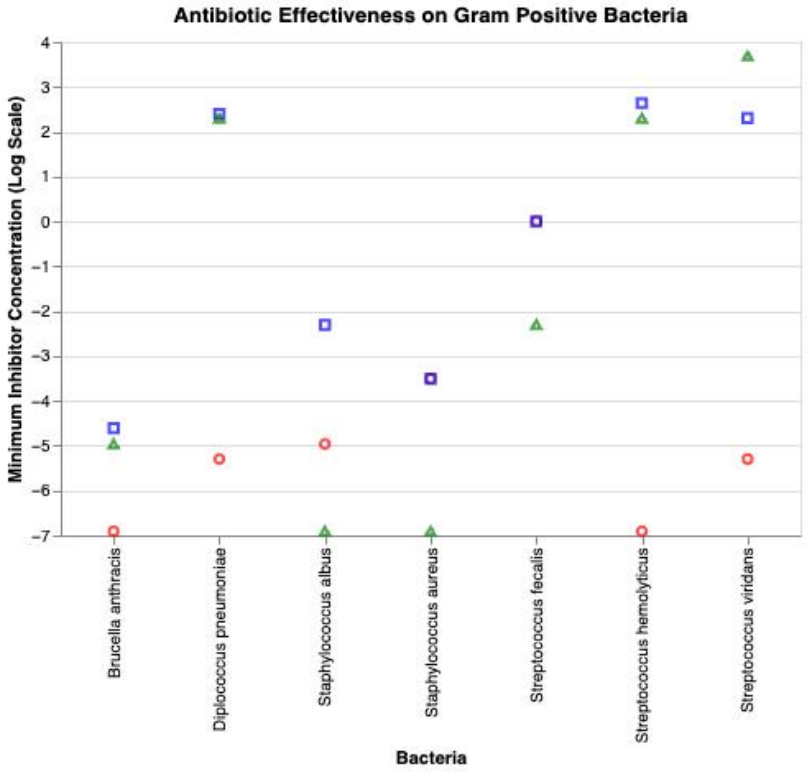
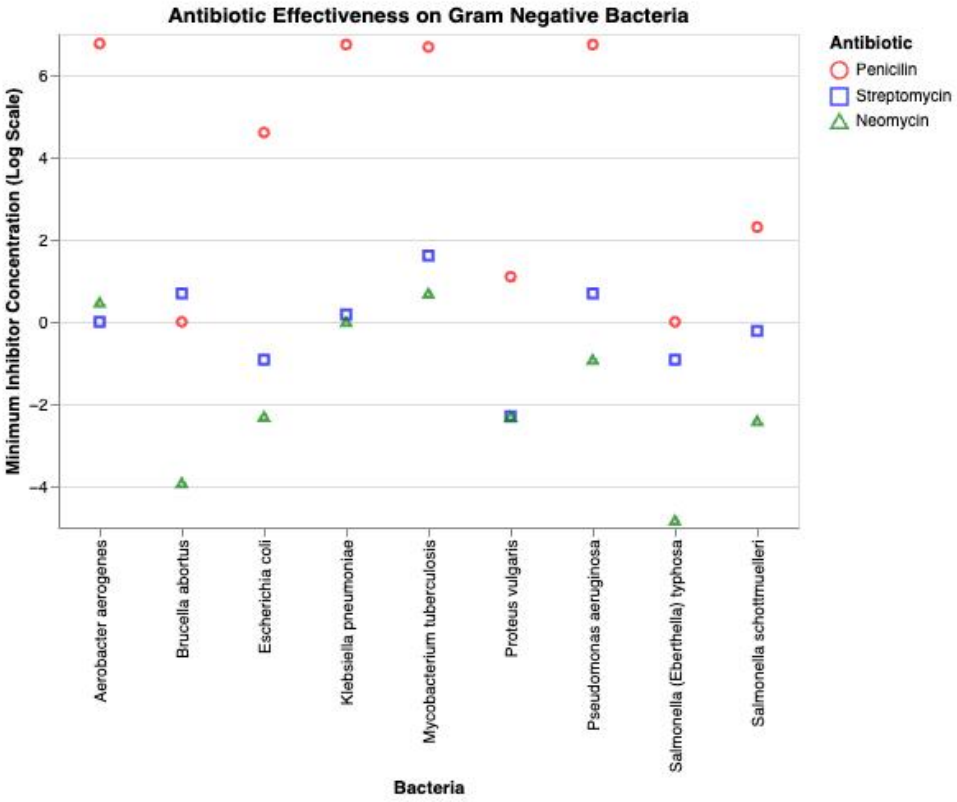
Yuyan Wang

Assignment 1: Visualization design





The tools used to create this graph are Python libraries; including pandas (for creating and managing the dataframe), numpy (for converting data into a log scale), and altair (for plotting the graph).

The first decision to make when starting this assignment is to establish the main goal of this visualization. It is evident that we need to convey a message or information, but what exactly? In this case, I have focused on providing a means of comparing the effectiveness of three antibiotics (Penicillin, Streptomycin, and Neomycin) for which we have numerical records.

Since the purpose is to provide a visual comparison, I considered that the actual numerical values (minimum inhibitor concentration - MIC) are not crucial; what matters more is the relative effectiveness among these three antibiotics. In other words, it is not essential to convey that Penicillin's MIC for "Aerobacter aerogenes" is 870; instead, it is more important to highlight that the MIC for the other two antibiotics is much lower than 870. Moreover, upon analyzing the original dataset, it becomes apparent that there is a wide range of values, making it challenging to discern differences among small values. Therefore, considering these observations, I decided to perform a Log Scale transformation on the numerical values in the dataset to emphasize relative differences among smaller values. Additionally, all MIC values related to the same bacteria are grouped in the same column to facilitate comparison.

Since all numerical data points are positioned in the same space, there is another problem to address: how to differentiate them. My solution was to apply color encoding to the antibiotic type to make it visually appealing and use shape encoding to support those who are color blind. A legend is included on the right side of the image to provide clarity.

Up to this point, most design patterns have been explained, except for the rationale behind using two separate graphs (one for gram-positive bacteria and one for gram-negative bacteria). The answer lies in allowing researchers to investigate if the effectiveness of different antibiotics is influenced by whether a bacterium is gram-positive or gram-negative. It is important to acknowledge that this dataset is not large enough (only 16 samples) to conclusively demonstrate whether antibiotic effectiveness is affected by gram staining. However, from the final plot, we can intuitively assume that Penicillin may not be an effective remedy for gram-negative bacteria (all red dots are in the upper layer of the left graph, indicating a high MIC). On the other hand, it could be a viable alternative for gram-positive bacteria since the MIC is lower for them. This observation can lead to further studies, such as exploring the reasons behind this behavior, possible antimicrobial resistance, and why Penicillin exhibits a high MIC for gram-negative bacteria, among other questions.