

CS 503 – Data Visualization

Assignment 1: Visualization Design

1. Introduction

In this assignment, you will design a visualization for a small data set and provide a rigorous rationale for your design choices.

After the World War II, antibiotics were considered as "wonder drugs", since they were easy remedy for what had been intractable ailments. To learn which drug worked most effectively for which bacterial infection, performance of the three most popular antibiotics on 16 bacteria were gathered. The values Table 1 represent the minimum inhibitory concentration (MIC), a measure of the effectiveness of the antibiotic, which represents the concentration of antibiotic required to prevent growth in vitro. The reaction of the bacteria to Gram staining is described by the covariate “gram staining”. Bacteria that are stained dark blue or violet are Gram-positive. Otherwise, they are Gram-negative.

Bacteria	Penicilin	Streptomycin	Neomycin	Gram Staining
Aerobacter aerogenes	870	1	1.6	negative
Brucella abortus	1	2	0.02	negative
Brucella anthracis	0.001	0.01	0.007	positive
Diplococcus pneumoniae	0.005	11	10	positive
Escherichia coli	100	0.4	0.1	negative
Klebsiella pneumoniae	850	1.2	1	negative
Mycobacterium tuberculosis	800	5	2	negative
Proteus vulgaris	3	0.1	0.1	negative
Pseudomonas aeruginosa	850	2	0.4	negative
Salmonella (Eberthella) typhosa	1	0.4	0.008	negative
Salmonella schottmuelleri	10	0.8	0.09	negative
Staphylococcus albus	0.007	0.1	0.001	positive
Staphylococcus aureus	0.03	0.03	0.001	positive
Streptococcus fecalis	1	1	0.1	positive
Streptococcus hemolyticus	0.001	14	10	positive
Streptococcus viridans	0.005	10	40	positive

2. Tasks

Your task is to design a visualization technique that you believe effectively communicates the data and provide a short write-up (no more than 4 paragraphs) describing your design. This design is must be implemented in Python. The Python code provided on [Python data Visualization: demos](#) will give you an idea how to use different visualization techniques in Python. In addition, the lecture [Visualization Design](#) contains several hints and explanations about this assignment. While you must use the data set given, note that you are free to transform the data as you see fit. Your chart image should be interpretable without recourse to your short write-up. Do not forget to include title, axis labels or legends as needed!

As different visualizations can emphasize different aspects of a data set, you should document what aspects of the data you are attempting to communicate effectively. In short, what story (or stories) are you trying to tell? Just as important, also note which aspects of the data might be obscured or down-played due to your visualization design. In your write-up, you should provide a rigorous rationale for your design decisions. Document the visual encodings you used and why they are appropriate for the data. These decisions include the choice of visualization type, size, color, scale, and other visual elements, as well as the use of sorting or other data transformations.

3. Submission

You must submit the *pdf* file of your report. In fact, it must contain the illustration of your visualization technique in term of images (three images that illustrates three visualization techniques) and a rigorous explanation. Indeed, your explanation must justify which one of the three visualization techniques will help doctors to make a relevant decision. The comparison of the three visualization techniques is an asset. In addition, you have to submit your *.py* or *.ipynb* if you prefer using *Jupyter Notebook*. **Please, do not submit your assignment in .zip or .rar files.**