Clayton Nagle

UCSD Extension Data Science Bootcamp

13 March 2021

Pymaceuticals Writeup: 3 Observations

1. There are effectively no outliers in the dataset. It’s surprising to me, and perhaps the data has already been filtered to some degree beforehand, but there is only 1 outlier point across the 4 best performing drugs for tumor volume. I would expect there to be at least a handful of mice with oddly sized tumors, but the data is very neat in that respect.
2. Capomulin treatment appears to reduce tumor size over time. Running the data with multiple different random mice shows that the Capomulin treatment is effective in reducing tumor size, at least slightly. Although I don’t have enough information in front of me to draw a strong conclusion, nor did we perform a robust enough analysis for me to feel confident, the eyeball test suggests to me that this is the case.
3. Interquartile range of tumor size varies wildly across the top 4 treatments. The IQR of tumor size for Capomulin is 7.78 mm3, while the IQR of tumor size for Ceftamin is 15.58. This makes me concerned about how the mice were assigned to treatment groups for this experiment, and raises other questions: are the IQRs different because certain drugs are more or less effective in some cases? Are mice being placed in groups based on their initial tumor size? It isn’t clear given the information provided.