Summary

This assignment walked me through how to analyze drug testing data on mice. I first cleaned the dataset by identifying the duplicate mouse and removing its ID from the data frame. The cleaned data frame was then used to calculate summary statistics and create graphs. For all the graphs, I grouped the data into a usable data frame and then generated the graph. For the boxplot, I struggled to create the visual because of the way it was organized in the list. I think it would be helpful to know how to do it in a data frame that organizes the results cleanly.

The results show that Capomulin and Ramicane are the most frequently used drugs in the study. Male mice are used slightly more than female, but it is almost and equal split. Furthermore, there is a pretty strong correlation (84%) between weight and tumor size. This indicates that weight might be a cause of tumor size, but further analysis would need to be conducted to determine causation.