Chris Krokus

Homework #5 – Matplotlib – Pymaceuticals

My jupyter notebook did not like me moving folders around while I was attempting to write this up and run the whisker and box plots for the other 3 drugs….

In the ‘Total Data Points by Drug’ bar graph, we see that Capomulin and Ramicane have the largest number of data points, which either means that these were the most common drug regimens for all the mice involved in the study or more likely, these were the drugs that reduced the tumor size, fought the cancer and kept the mice alive the longest. ‘The Percentage of Males vs Females Tested’ pie graph does provide any valuable information other than the fact that there were approximately and even number (51% to 49%) of male and female mice involved in the study.

For the four most promising treatment regimens I developed box and whisker plots to display each regimen’s data for tumor volume in cubic millimeters. 23 mice were on the Capomulin drug regimen. The median tumor volume was 38.1 mm3 with the interquartile range being 7.8 mm3. The lower quartile aka the median of the lower region is 32.4 mm3 and the upper quartile aka the median of the upper region is 40.16 mm3. No mice on the Capomulin regimen were outliers as there was no tumor size less than 20.7 mm3 or greater than 51.8 mm3.

The correlation coefficient for the Average Mouse Weight vs. Tumor Volume is an r-squared of .709 meaning that 70.9% of the variation in tumor volumes (dependent variable) can be explained by the relationship to the mouse’s weight (independent variable).