# Pymaceuticals Inc.



**Removed duplicated mice using unique and clean function.**

The total number of mice is 248.

**Generated bar plot showing the total number of timepoints for all mice tested for each drug regimen using pandas and pyplot**.

* The bar graph showed the Drug Regimen Capomulin has the maximum mice number (230), and Propriva has the smaller mice number (182).

Graphical user interface

Description automatically generatedGraphical user interface

Description automatically generated

**Generate a pie plot showing the distribution of female versus male mice using Pandas and Pyplot**

The total count of mice by gender showed that 124 female mice and 125 male mice.Chart

Description automatically generated

**Calculated the IQR and quantitatively determined if there are any potential outliers.**

Chart, box and whisker chart

Description automatically generated

**Generated a line plot of tumor volume vs. time point for a mouse treated with Capomulin**

From the selected treatments Capomulin reduces the size of tumors better.

Graphical user interface, application

Description automatically generated

**Generated scatter plot of average tumor volume vs. mouse weight for the Capomulin regimen**

The correlation between mouse weight, and average tumor volume is 0.84. It is a strong positive correlation, when the mouse weight increases the average tumor volume also increases.

Chart, scatter chart

Description automatically generated

**Calculated the correlation coefficient and linear regression model for mouse weight and average tumor volume for the Capomulin regimen.**

* The regression analysis helped to understand how much the average tumor volume (dependent variable) will change when weight of mice changes (independent variables). The R-squared value is 0.70. Higher R-squared values represent smaller differences between the observed data, and the fitted value. 70% the model explains all the variation in the response variable around its mean.

Chart, scatter chart

Description automatically generated

**Observations**

* There were more trials done for certain drugs than others, so the information is skewed. Capomulin, Ramicane ,Ketapril, and Naftisol had more # of trials than other drugs given.

**Other comments:**

When doing the Summary statistics question, particularly when I used the mean and median functions, I received an error that the function was deprecated. I had to google it and later discovered that it required a numeric\_only = true parameter.

Had difficulty solving the linear and regression model had to google and referred various sources to answer the question.