**KU Edwards – Data Analytics BootCamp**

**MatPlotLib Homework**

Using pymaceuticals\_starter

**Observations and Insights**

Below is an analysis of the Pymaceuticals HomeWork Assignment. This process examined the results of 9 drug regimens: Capomulin, Ceftamin, Infubinol, Ketapril, Naftisol, Propriva, Ramicane, Stlasyn and Zoniferol; along with a control Placebo, and involved 250 mice evenly distributed between male and female.

After examining the results of the study:

-Four Regimens showed the greatest reduction of Tumor Volume (whether controlling for variance

via StdDeviation or SEMs)

**Ramicane** 0.320 **Capomulin** 0.329 **Ceftamin** 0.469 **Infubinol** 0.492

-The Mice in these regimens had the greatest weight retention

-The number of Metastic sites were the lowest (with some showing no sites after the study period)

-The Regimens appeared to be equally effective for Male and Female mice

Below the code, calculations and results can be seen…

**Dependencies and starter code**[**¶**](http://localhost:8888/notebooks/Instructions/Pymaceuticals/pymaceuticals_starter.ipynb#Dependencies-and-starter-code)

In [1]:

*# Dependencies and Setup*

**import** matplotlib.pyplot **as** plt

**import** pandas **as** pd

**import** numpy **as** np

**import** scipy.stats **as** st

**from** matplotlib **import** pyplot **as** plt

**from** scipy.stats **import** linregress

**from** scipy.stats **import** sem

**from** sklearn **import** datasets

​

*# Study data files*

mouse\_metadata **=** "data/Mouse\_metadata.csv"

study\_results **=** "data/Study\_results.csv"

​

*# Read the mouse data and the study results*

mouse\_metadata **=** pd.read\_csv(mouse\_metadata)

study\_results **=** pd.read\_csv(study\_results)

*#*

mouse\_metadata\_pd **=** pd.DataFrame (mouse\_metadata)

study\_results\_pd **=** pd.DataFrame (study\_results)

​

*# Combine the data into a single dataset*

*# -------Merge two dataframes using an inner join --- merge\_table = pd.merge(\_pd, \_pd, on="")*

MergedMouse\_pd **=** pd.merge(mouse\_metadata\_pd, study\_results\_pd, on**=**'Mouse ID')

MergedMouse\_pd.head()

Out[1]:

|  | **Mouse ID** | **Drug Regimen** | **Sex** | **Age\_months** | **Weight (g)** | **Timepoint** | **Tumor Volume (mm3)** | **Metastatic Sites** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **0** | k403 | Ramicane | Male | 21 | 16 | 0 | 45.000000 | 0 |
| **1** | k403 | Ramicane | Male | 21 | 16 | 5 | 38.825898 | 0 |
| **2** | k403 | Ramicane | Male | 21 | 16 | 10 | 35.014271 | 1 |
| **3** | k403 | Ramicane | Male | 21 | 16 | 15 | 34.223992 | 1 |
| **4** | k403 | Ramicane | Male | 21 | 16 | 20 | 32.997729 | 1 |
| **5** | k403 | Ramicane | Male | 21 | 16 | 25 | 33.464577 | 1 |
| **6** | k403 | Ramicane | Male | 21 | 16 | 30 | 31.099498 | 1 |
| **7** | k403 | Ramicane | Male | 21 | 16 | 35 | 26.546993 | 1 |
| **8** | k403 | Ramicane | Male | 21 | 16 | 40 | 24.365505 | 1 |
| **9** | k403 | Ramicane | Male | 21 | 16 | 45 | 22.050126 | 1 |
| **10** | s185 | Capomulin | Female | 3 | 17 | 0 | 45.000000 | 0 |
| **11** | s185 | Capomulin | Female | 3 | 17 | 5 | 43.878496 | 0 |
| **12** | s185 | Capomulin | Female | 3 | 17 | 10 | 37.614948 | 0 |
| **13** | s185 | Capomulin | Female | 3 | 17 | 15 | 38.177232 | 0 |
| **14** | s185 | Capomulin | Female | 3 | 17 | 20 | 36.866876 | 0 |
| **15** | s185 | Capomulin | Female | 3 | 17 | 25 | 33.949940 | 0 |
| **16** | s185 | Capomulin | Female | 3 | 17 | 30 | 32.959671 | 1 |
| **17** | s185 | Capomulin | Female | 3 | 17 | 35 | 28.328531 | 1 |
| **18** | s185 | Capomulin | Female | 3 | 17 | 40 | 25.472143 | 1 |
| **19** | s185 | Capomulin | Female | 3 | 17 | 45 | 23.343598 | 1 |

In [2]:

*## Summary statistics*

In [2]:

*# Generate a summary statistics table of mean, median, variance, standard deviation, and SEM*

*# of the tumor volume for each regimen*

*#TumVol = MergedMouse\_pd['Tumor Volume (mm3)']*

*#TumVol*

*# calculating measures of central tendency*

mean\_TumVol **=** MergedMouse\_pd.groupby("Drug Regimen").mean()["Tumor Volume (mm3)"]

print(f"The mean TumVol is: \n {mean\_TumVol}")

​

median\_TumVol **=** MergedMouse\_pd.groupby("Drug Regimen").median()["Tumor Volume (mm3)"]

print(f"The median TumVol is: \n {median\_TumVol}")

​

*#mode\_TumVol = MergedMouse\_pd.groupby("Drug Regimen").mode()["Tumor Volume (mm3)"]*

*#print(f"The mode TumVol is: \n {mode\_TumVol}")*

The mean TumVol is:

Drug Regimen

Capomulin 40.675741

Ceftamin 52.591172

Infubinol 52.884795

Ketapril 55.235638

Naftisol 54.331565

Placebo 54.033581

Propriva 52.322552

Ramicane 40.216745

Stelasyn 54.233149

Zoniferol 53.236507

Name: Tumor Volume (mm3), dtype: float64

The median TumVol is:

Drug Regimen

Capomulin 41.557809

Ceftamin 51.776157

Infubinol 51.820584

Ketapril 53.698743

Naftisol 52.509285

Placebo 52.288934

Propriva 50.854632

Ramicane 40.673236

Stelasyn 52.431737

Zoniferol 51.818479

Name: Tumor Volume (mm3), dtype: float64

In [3]:

*# calculating the variance and standard deviation using the different modules*

var\_TumVol **=** MergedMouse\_pd.groupby("Drug Regimen").var()["Tumor Volume (mm3)"]

print(f"The population variance is: \n {var\_TumVol}")

*# = np.(TumVol,ddof = 0)*

sd\_TumVol **=** MergedMouse\_pd.groupby("Drug Regimen").std()["Tumor Volume (mm3)"]

print(f"The population standard deviation is: \n {sd\_TumVol}")

​

The population variance is:

Drug Regimen

Capomulin 24.947764

Ceftamin 39.290177

Infubinol 43.128684

Ketapril 68.553577

Naftisol 66.173479

Placebo 61.168083

Propriva 42.351070

Ramicane 23.486704

Stelasyn 59.450562

Zoniferol 48.533355

Name: Tumor Volume (mm3), dtype: float64

The population standard deviation is:

Drug Regimen

Capomulin 4.994774

Ceftamin 6.268188

Infubinol 6.567243

Ketapril 8.279709

Naftisol 8.134708

Placebo 7.821003

Propriva 6.507770

Ramicane 4.846308

Stelasyn 7.710419

Zoniferol 6.966589

Name: Tumor Volume (mm3), dtype: float64

In [4]:

*# Calculate "standard error on means" or SEM*

​

semsTumVol **=** MergedMouse\_pd.groupby("Drug Regimen").sem()["Tumor Volume (mm3)"]

print(f"The population SEM using the scipy.stats module is: \n {semsTumVol}")

The population SEM using the scipy.stats module is:

Drug Regimen

Capomulin 0.329346

Ceftamin 0.469821

Infubinol 0.492236

Ketapril 0.603860

Naftisol 0.596466

Placebo 0.581331

Propriva 0.512884

Ramicane 0.320955

Stelasyn 0.573111

Zoniferol 0.516398

Name: Tumor Volume (mm3), dtype: float64

In [5]:

*# ---------------------------------Merge and Rename the values*

MrgTumRegm\_pd **=** pd.DataFrame() *#=========creates a new dataframe that's empty*

MrgTumRegm\_pd **=** pd.merge(mean\_TumVol, median\_TumVol, on**=**'Drug Regimen')

MrgTumRegm\_pd **=** MrgTumRegm\_pd.rename(columns**=**{"Tumor Volume (mm3)\_x":"Tumor Volume Mean",

"Tumor Volume (mm3)\_y":"Tumor Volume Median"})

MrgTumRegm\_pd **=** pd.merge(MrgTumRegm\_pd, var\_TumVol, on**=**'Drug Regimen')

MrgTumRegm\_pd **=** pd.merge(MrgTumRegm\_pd, sd\_TumVol, on**=**'Drug Regimen')

MrgTumRegm\_pd **=** MrgTumRegm\_pd.rename(columns**=**{"Tumor Volume (mm3)\_x":"Tumor Volume StdDev",

"Tumor Volume (mm3)\_y":"Tumor Volume Variance"})

MrgTumRegm\_pd **=** pd.merge(MrgTumRegm\_pd, semsTumVol, on**=**'Drug Regimen')

MrgTumRegm\_pd **=** MrgTumRegm\_pd.rename(columns**=**{"Tumor Volume (mm3)":"Tumor Volume StdErrMeans"})

MrgTumRegm\_pd

Out[5]:

|  | **Tumor Volume Mean** | **Tumor Volume Median** | **Tumor Volume StdDev** | **Tumor Volume Variance** | **Tumor Volume StdErrMeans** |
| --- | --- | --- | --- | --- | --- |
| **Drug Regimen** |  |  |  |  |  |
| **Capomulin** | 40.675741 | 41.557809 | 24.947764 | 4.994774 | 0.329346 |
| **Ceftamin** | 52.591172 | 51.776157 | 39.290177 | 6.268188 | 0.469821 |
| **Infubinol** | 52.884795 | 51.820584 | 43.128684 | 6.567243 | 0.492236 |
| **Ketapril** | 55.235638 | 53.698743 | 68.553577 | 8.279709 | 0.603860 |
| **Naftisol** | 54.331565 | 52.509285 | 66.173479 | 8.134708 | 0.596466 |
| **Placebo** | 54.033581 | 52.288934 | 61.168083 | 7.821003 | 0.581331 |
| **Propriva** | 52.322552 | 50.854632 | 42.351070 | 6.507770 | 0.512884 |
| **Ramicane** | 40.216745 | 40.673236 | 23.486704 | 4.846308 | 0.320955 |
| **Stelasyn** | 54.233149 | 52.431737 | 59.450562 | 7.710419 | 0.573111 |
| **Zoniferol** | 53.236507 | 51.818479 | 48.533355 | 6.966589 | 0.516398 |

In [6]:

*# ------------------------display in formated style for readability -----*

MrgTumRegm\_pd.style.format({'Tumor Volume Mean': "{:.2f}",

'Tumor Volume Median': "{:.2f}",

'Tumor Volume StdDev': "{:.2f}",

'Tumor Volume Variance': "{:.2f}",

'Tumor Volume StdErrMeans': "{:.2f}"

})

Out[6]:

|  | **Tumor Volume Mean** | **Tumor Volume Median** | **Tumor Volume StdDev** | **Tumor Volume Variance** | **Tumor Volume StdErrMeans** |
| --- | --- | --- | --- | --- | --- |
| **Drug Regimen** |  |  |  |  |  |
| **Capomulin** | 40.68 | 41.56 | 24.95 | 4.99 | 0.33 |
| **Ceftamin** | 52.59 | 51.78 | 39.29 | 6.27 | 0.47 |
| **Infubinol** | 52.88 | 51.82 | 43.13 | 6.57 | 0.49 |
| **Ketapril** | 55.24 | 53.70 | 68.55 | 8.28 | 0.60 |
| **Naftisol** | 54.33 | 52.51 | 66.17 | 8.13 | 0.60 |
| **Placebo** | 54.03 | 52.29 | 61.17 | 7.82 | 0.58 |
| **Propriva** | 52.32 | 50.85 | 42.35 | 6.51 | 0.51 |
| **Ramicane** | 40.22 | 40.67 | 23.49 | 4.85 | 0.32 |
| **Stelasyn** | 54.23 | 52.43 | 59.45 | 7.71 | 0.57 |
| **Zoniferol** | 53.24 | 51.82 | 48.53 | 6.97 | 0.52 |

In [ ]:

​

**Bar plots**

In [7]:

*# calculating counts by regimen*

count\_TumVol **=** MergedMouse\_pd.groupby("Drug Regimen").count()["Mouse ID"]

*#*

SortedCntTumvol **=** count\_TumVol.sort\_values(ascending**=False**)

CntTumvol\_pd **=** pd.DataFrame(SortedCntTumvol)

CntTumvol\_pd **=** CntTumvol\_pd.rename(columns**=**{"Mouse ID":"RegimenCounts"})

​

CntTumvol\_pd

Out[7]:

|  | **RegimenCounts** |
| --- | --- |
| **Drug Regimen** |  |
| **Capomulin** | 230 |
| **Ramicane** | 228 |
| **Ketapril** | 188 |
| **Naftisol** | 186 |
| **Zoniferol** | 182 |
| **Stelasyn** | 181 |
| **Placebo** | 181 |
| **Infubinol** | 178 |
| **Ceftamin** | 178 |
| **Propriva** | 161 |

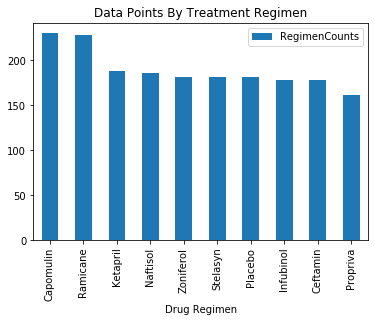
In [10]:

*# Generate a bar plot showing number of data points for each treatment regimen using pandas*

*#*

count\_chart**=**CntTumvol\_pd.plot(kind**=**'bar', title**=**'Data Points By Treatment Regimen')

​



In [9]:

*#*

CntTumvol\_pd **=** CntTumvol\_pd.reset\_index()

In [15]:

*# Generate a bar plot showing number of data points for each treatment regimen using pyplot*

*# set x-axis and tick locations*

x\_axis **=** np.arange(len(CntTumvol\_pd))

tick\_locations **=** [value**+**0.4 **for** value **in** x\_axis]

*# set Figure*

plt.figure(figsize**=**(5,4))

*# build Bar Chart and display it*

plt.bar(x\_axis, CntTumvol\_pd["RegimenCounts"], color**=**'b', alpha**=**0.5, align**=**"edge")

plt.xticks(tick\_locations, CntTumvol\_pd["Drug Regimen"], rotation**=**"vertical")

plt.xlim(**-**0.25, len(x\_axis))

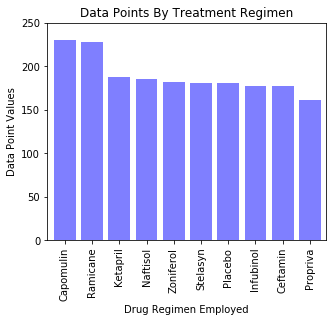
plt.ylim(0, max(CntTumvol\_pd["RegimenCounts"])**+**20)

plt.title("Data Points By Treatment Regimen")

plt.xlabel("Drug Regimen Employed")

plt.ylabel("Data Point Values")

plt.show()



In [ ]:

​

**Pie plots**

In [16]:

*# calculating counts by sex*

count\_Sex **=** MergedMouse\_pd.groupby("Sex").count()["Mouse ID"]

CntSexTot **=** MergedMouse\_pd.count()['Sex']

*#*

SortedCntSex **=** count\_Sex.sort\_values()

CntSex\_pd **=** pd.DataFrame(SortedCntSex)

CntSex\_pd **=** CntSex\_pd.rename(columns**=**{"Mouse ID":"Number Mice By Sex"})

PctSex **=** CntSex\_pd['Number Mice By Sex'] **/** CntSexTot **\*** 100

*#*

PctSex\_pd **=** pd.DataFrame(PctSex)

PctSex\_pd **=** PctSex\_pd.rename(columns**=**{"Number Mice By Sex":"% Mice By Sex"})

PctSex\_pd

​

Out[16]:

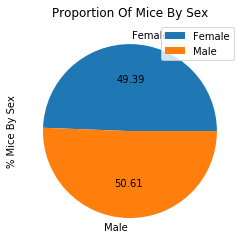
|  | **% Mice By Sex** |
| --- | --- |
| **Sex** |  |
| **Female** | 49.392499 |
| **Male** | 50.607501 |

In [17]:

*# Generate a pie plot showing the distribution of female versus male mice using pandas*

*#*

count\_chart**=**PctSex\_pd.plot(kind**=**'pie', title**=**'Proportion Of Mice By Sex', y**=**"% Mice By Sex", autopct**=**'%.2f')



In [18]:

​

PctSex\_pd **=** PctSex\_pd.reset\_index()

PctSex\_pd

Out[18]:

|  | **Sex** | **% Mice By Sex** |
| --- | --- | --- |
| **0** | Female | 49.392499 |
| **1** | Male | 50.607501 |

In [19]:

*# Generate a pie plot showing the distribution of female versus male mice using pyplot*

*# set x-axis and tick locations*

*#x\_axis = np.arange(len(CntTumvol\_pd))*

*#tick\_locations = [value+0.4 for value in x\_axis]*

*# set Figure*

plt.figure(figsize**=**(8,4))

*# build Pie Chart and display it*

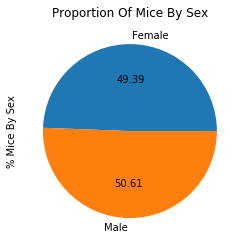
plt.pie(PctSex\_pd["% Mice By Sex"], autopct**=**'%.2f', labels**=**PctSex\_pd["Sex"])

plt.title('Proportion Of Mice By Sex')

plt.ylabel("% Mice By Sex")

*#plt.legend(handles=PctSex\_pd[['Sex']], loc="best")*

plt.show()



In [ ]:

​

**Quartiles, outliers and boxplots**

In [27]:

*# Calculate the final tumor volume of each mouse across*

*# four of the most promising treatment regimens.*

SemsTumvol\_pd **=** pd.DataFrame(semsTumVol)

SemsTumvol\_pd **=** SemsTumvol\_pd.sort\_values(['Tumor Volume (mm3)'],ascending**=True**)

​

SemsTumvol\_pd.head(4)

​

Out[27]:

|  | **Tumor Volume (mm3)** |
| --- | --- |
| **Drug Regimen** |  |
| **Ramicane** | 0.320955 |
| **Capomulin** | 0.329346 |
| **Ceftamin** | 0.469821 |
| **Infubinol** | 0.492236 |

In [39]:

*# Calc the timepoint by mouse MergedMouse\_pd*

MouseTimept\_pd **=** MergedMouse\_pd.groupby(["Mouse ID"])["Timepoint"].max()

MouseTimept\_pd **=** MouseTimept\_pd.reset\_index()

*# Merge TimePoint by mouse id to Mouse dtls*

MergeInfo\_pd **=** MouseTimept\_pd[['Mouse ID', 'Timepoint']].merge(MergedMouse\_pd,

on**=**['Mouse ID', 'Timepoint'], how**=**'left')

MergeInfo\_pd.head()

​

Out[39]:

|  | **Mouse ID** | **Timepoint** | **Drug Regimen** | **Sex** | **Age\_months** | **Weight (g)** | **Tumor Volume (mm3)** | **Metastatic Sites** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **0** | a203 | 45 | Infubinol | Female | 20 | 23 | 67.973419 | 2 |
| **1** | a251 | 45 | Infubinol | Female | 21 | 25 | 65.525743 | 1 |
| **2** | a262 | 45 | Placebo | Female | 17 | 29 | 70.717621 | 4 |
| **3** | a275 | 45 | Ceftamin | Female | 20 | 28 | 62.999356 | 3 |
| **4** | a366 | 30 | Stelasyn | Female | 16 | 29 | 63.440686 | 1 |

In [48]:

Ramicane **=** MergeInfo\_pd.loc[MergeInfo\_pd['Drug Regimen'] **==** 'Ramicane']['Tumor Volume (mm3)']

Ramicane.head()

Out[48]:

6 38.407618

7 43.047543

10 38.810366

12 32.978522

33 38.342008

Name: Tumor Volume (mm3), dtype: float64

In [49]:

Capomulin **=** MergeInfo\_pd.loc[MergeInfo\_pd['Drug Regimen'] **==** 'Capomulin']['Tumor Volume (mm3)']

Capomulin.head()

Out[49]:

19 38.982878

24 38.939633

61 30.485985

64 37.074024

66 40.159220

Name: Tumor Volume (mm3), dtype: float64

In [50]:

Ceftamin **=** MergeInfo\_pd.loc[MergeInfo\_pd['Drug Regimen'] **==** 'Ceftamin']['Tumor Volume (mm3)']

Ceftamin.head()

Out[50]:

3 62.999356

21 45.000000

22 56.057749

25 55.742829

58 48.722078

Name: Tumor Volume (mm3), dtype: float64

In [59]:

Infubinol **=** MergeInfo\_pd.loc[MergeInfo\_pd['Drug Regimen'] **==** 'Infubinol']['Tumor Volume (mm3)']

Infubinol.head()

Out[59]:

0 67.973419

1 65.525743

11 57.031862

13 66.083066

27 72.226731

Name: Tumor Volume (mm3), dtype: float64

In [ ]:

*# Calculate the IQR and quantitatively determine if there are any potential outliers.*

In [56]:

*# calc the Ramicane Quartiles and IQR...*

RamiQuartiles **=** Ramicane.quantile([.25,.5,.75])

RamiLowQ **=** RamiQuartiles[0.25]

RamiUprQ **=** RamiQuartiles[0.75]

RamiIqr **=** RamiUprQ **-** RamiLowQ

*#RamiIqr*

*#RamiQuartiles*

Out[56]:

9.098536719999998

In [55]:

*# calc the Capomulin Quartiles and IQR...*

CapoQuartiles **=** Capomulin.quantile([.25,.5,.75])

CapoLowQ **=** CapoQuartiles[0.25]

CapoUprQ **=** CapoQuartiles[0.75]

CapoIqr **=** CapoUprQ **-** CapoLowQ

*#CapoIqr*

*#CapoQuartiles*

Out[55]:

7.781863460000004

In [57]:

*# calc the Ceftamin Quartiles and IQR...*

CeftQuartiles **=** Ceftamin.quantile([.25,.5,.75])

CeftLowQ **=** CeftQuartiles[0.25]

CeftUprQ **=** CeftQuartiles[0.75]

CeftIqr **=** CeftUprQ **-** CeftLowQ

*#CeftIqr*

CeftQuartiles

Out[57]:

0.25 48.722078

0.50 59.851956

0.75 64.299830

Name: Tumor Volume (mm3), dtype: float64

In [60]:

*# calc the Infubinol Quartiles and IQR...*

InfuQuartiles **=** Infubinol.quantile([.25,.5,.75])

InfuLowQ **=** InfuQuartiles[0.25]

InfuUprQ **=** InfuQuartiles[0.75]

InfuIqr **=** InfuUprQ **-** InfuLowQ

*#InfuIqr*

InfuQuartiles

Out[60]:

0.25 54.048608

0.50 60.165180

0.75 65.525743

Name: Tumor Volume (mm3), dtype: float64

In [61]:

*# Generate a box plot of the final tumor volume of each mouse across four regimens of interest*

Top4Fig,Top4Ax **=** plt.subplots()

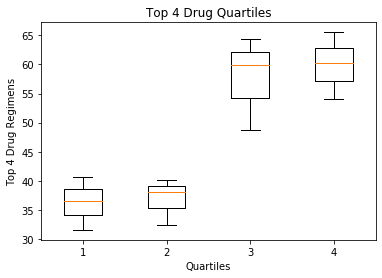
Top4Ax.boxplot([RamiQuartiles, CapoQuartiles, CeftQuartiles, InfuQuartiles])

plt.xlabel('Quartiles')

plt.ylabel('Top 4 Drug Regimens')

plt.title('Top 4 Drug Quartiles')

plt.show()



**## Line and scatter plots**

In [64]:

*# Generate a line plot of time point versus tumor volume for a mouse treated with Capomulin*

CapoMrg **=** MergedMouse\_pd.loc[MergedMouse\_pd['Drug Regimen'] **==** 'Capomulin']

CapoMouse **=** CapoMrg.loc[CapoMrg['Mouse ID'] **==** 's185']

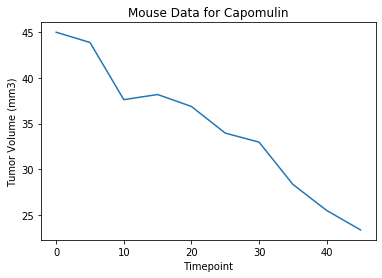
plt.plot(CapoMouse['Timepoint'], CapoMouse['Tumor Volume (mm3)'])

plt.xlabel('Timepoint')

plt.ylabel('Tumor Volume (mm3)')

plt.title('Mouse Data for Capomulin')

plt.show()



In [65]:

*# Generate a scatter plot of mouse weight versus average tumor volume for the Capomulin regimen*

*# CapoMrg calculated above*

CapoAve **=** CapoMrg.groupby(['Mouse ID']).mean()

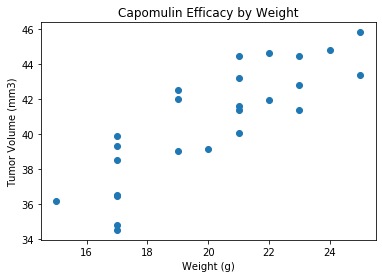
plt.scatter(CapoAve['Weight (g)'], CapoAve['Tumor Volume (mm3)'])

plt.xlabel('Weight (g)')

plt.ylabel('Tumor Volume (mm3)')

plt.title('Capomulin Efficacy by Weight')

plt.show()



In [90]:

*# Calculate the correlation coefficient*

*# and linear regression model for mouse weight*

*# and average tumor volume for the Capomulin regimen*

CapoCorrelation **=** st.pearsonr(CapoAve['Weight (g)'], CapoAve['Tumor Volume (mm3)'])[0],2

print (f"For the Capomulin Regimen, the Calculated Correlation Value is {CapoCorrelation[0]:.2f}")

CapoModel **=** st.linregress(CapoAve['Weight (g)'], CapoAve['Tumor Volume (mm3)'])*#CapoModel*

*#CapoModel*

Yval **=** CapoAve['Weight (g)'] **\*** CapoModel[0] **+** CapoModel[1]

For the Capomulin Regimen, the Calculated Correlation Value is 0.84

In [81]:

plt.scatter(CapoAve['Weight (g)'], CapoAve['Tumor Volume (mm3)'])

plt.plot(CapoAve['Weight (g)'], Yval, color **=** 'red')

plt.xlabel('Capomulin Mouse Weight (g)')

plt.ylabel('Capomulin Ave Tumor Volume (mm3)')

plt.title('Correlated Mouse Weight and Tumor Size using Capomulin')

plt.show()

