

Setup

Eliminate Duplicate Records

1

Remove Duplicate Records from the Data Set

Jupyter pymaceuticals_starter Last Checkpoint: 23 minutes ago

File Edit View Run Kernel Settings Help Trusted

JupyterLab Python 3 (ipykernel)

```
Cleaned DataFrame:
[22]:
```

	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
...
1888	z969	Naftisol	Male	9	30	25	63.145652	2
1889	z969	Naftisol	Male	9	30	30	65.841013	3
1890	z969	Naftisol	Male	9	30	35	69.176246	4
1891	z969	Naftisol	Male	9	30	40	70.314904	4
1892	z969	Naftisol	Male	9	30	45	73.867845	4

1880 rows x 8 columns

[]:

Number of mice in the clean DataFrame: 248

Summary Statistics

Jupyter pymaceuticals_starter Last Checkpoint: 5 minutes ago

File Edit View Run Kernel Settings Help Trusted

JupyterLab Python 3 (ipykernel)

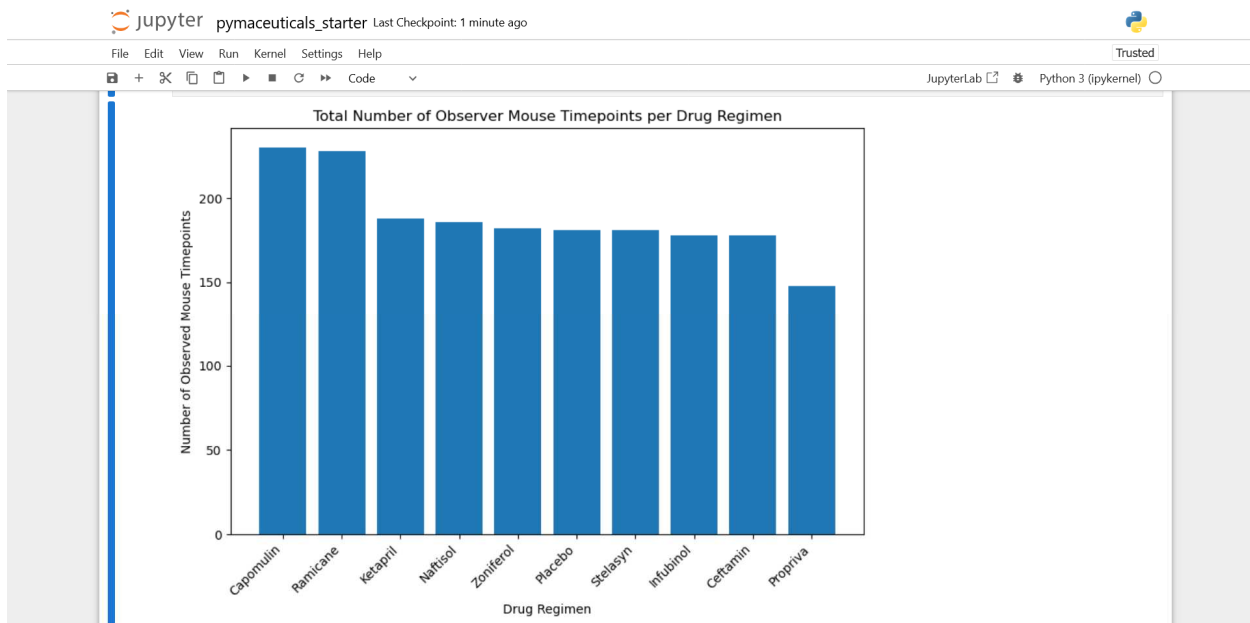
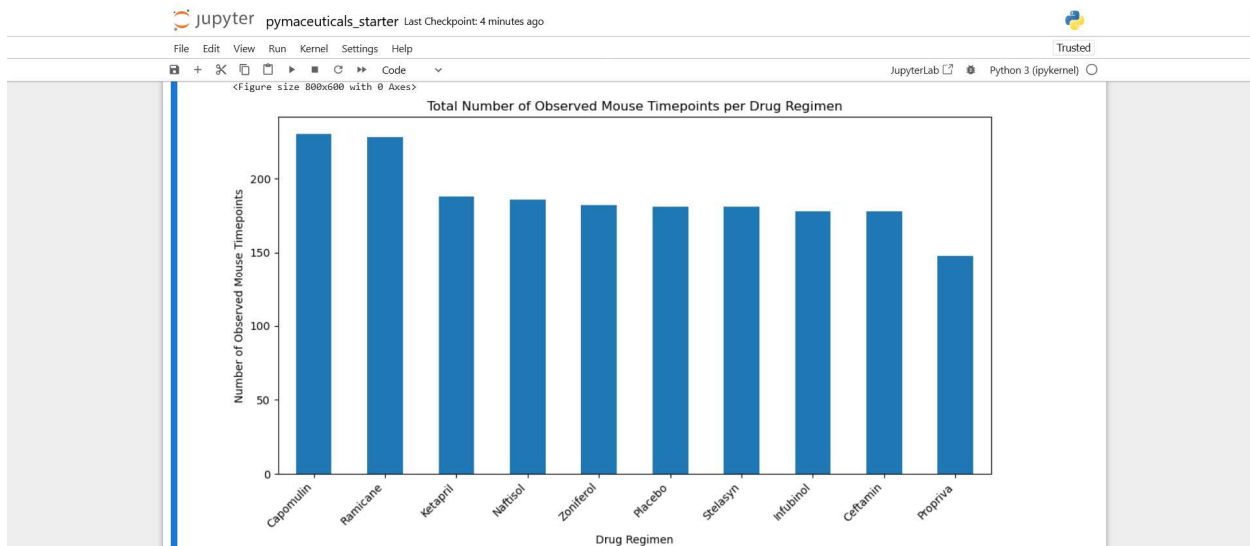
```
Summary Statistics of Tumor Volume by Drug Regimen:
[34]:
```

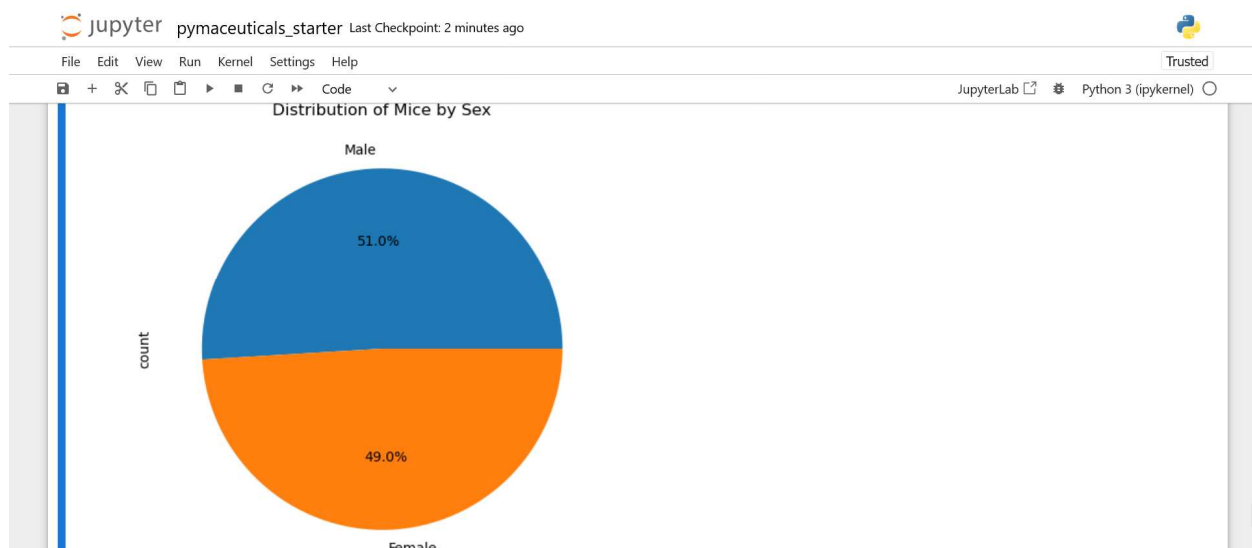
	mean_tumor_volume	median_tumor_volume	variance_tumor_volume	std_tumor_volume	sem_tumor_volume
Capomulin	40.675741	41.557809	24.947764	4.994774	0.329346
Ceftamin	52.591172	51.776157	39.290177	6.268188	0.469821
Infubinol	52.884795	51.820584	43.128684	6.567243	0.492236
Ketapril	55.235638	53.698743	68.553577	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
Propriova	52.320930	50.446266	43.852013	6.622085	0.544332
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

[]:

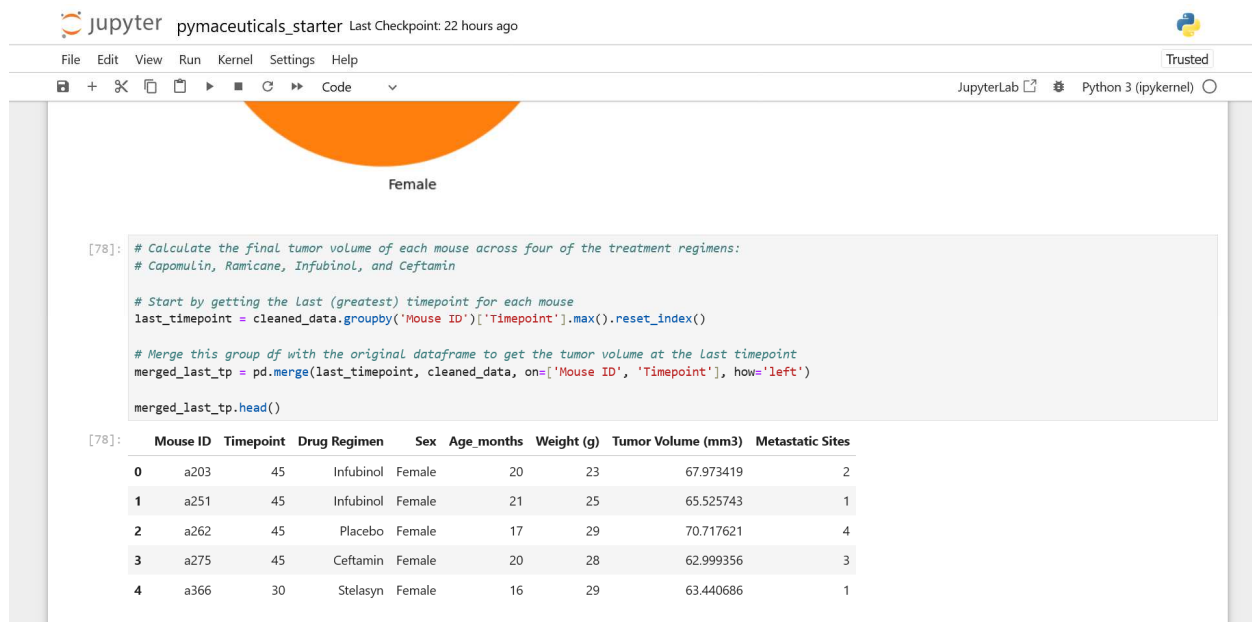
[]:

Bar and Pie Plots





Quartiles, Outliers, and Box Plots



```
jupyter pymaceuticals_starter Last Checkpoint: 11 minutes ago

File Edit View Run Kernel Settings Help Trusted

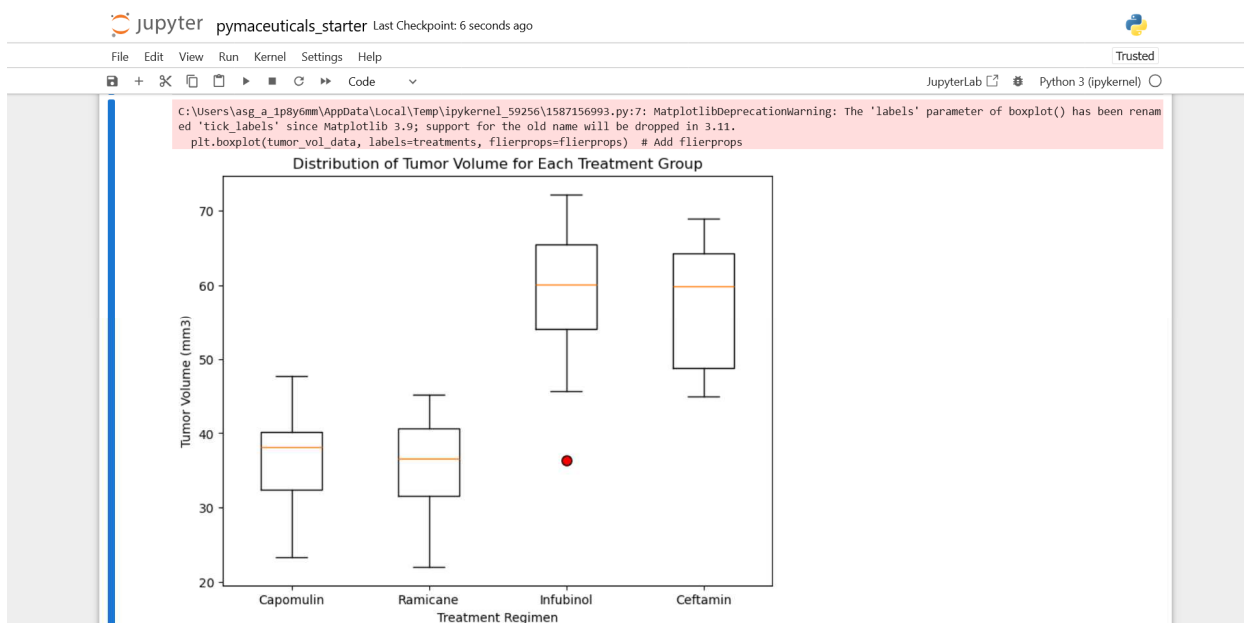
JupyterLab Python 3 (ipykernel)

Capomulin IQR data:
The lower quartile of Capomulin is: 32.37735684
The upper quartile of Capomulin is: 40.1592203
The interquartile range of Capomulin is: 7.7818634600000004
The the median of Capomulin is: 38.1251644
Values below 20.70 could be outliers.
Values above 51.83 could be outliers.
No potential outliers found for Capomulin.

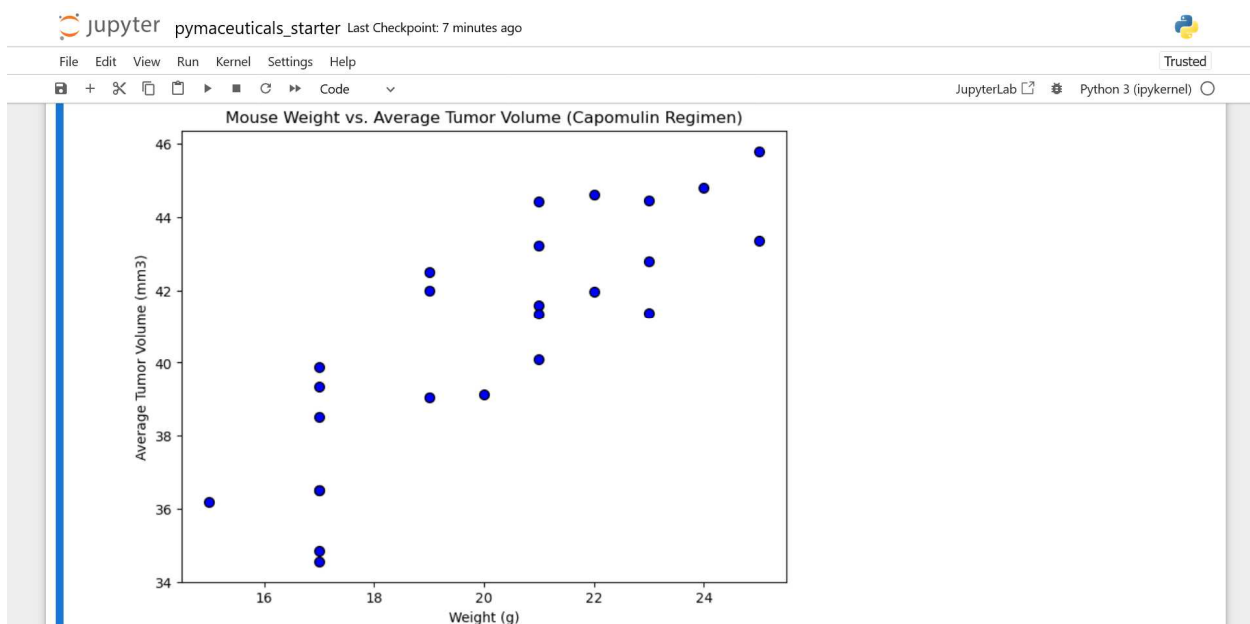
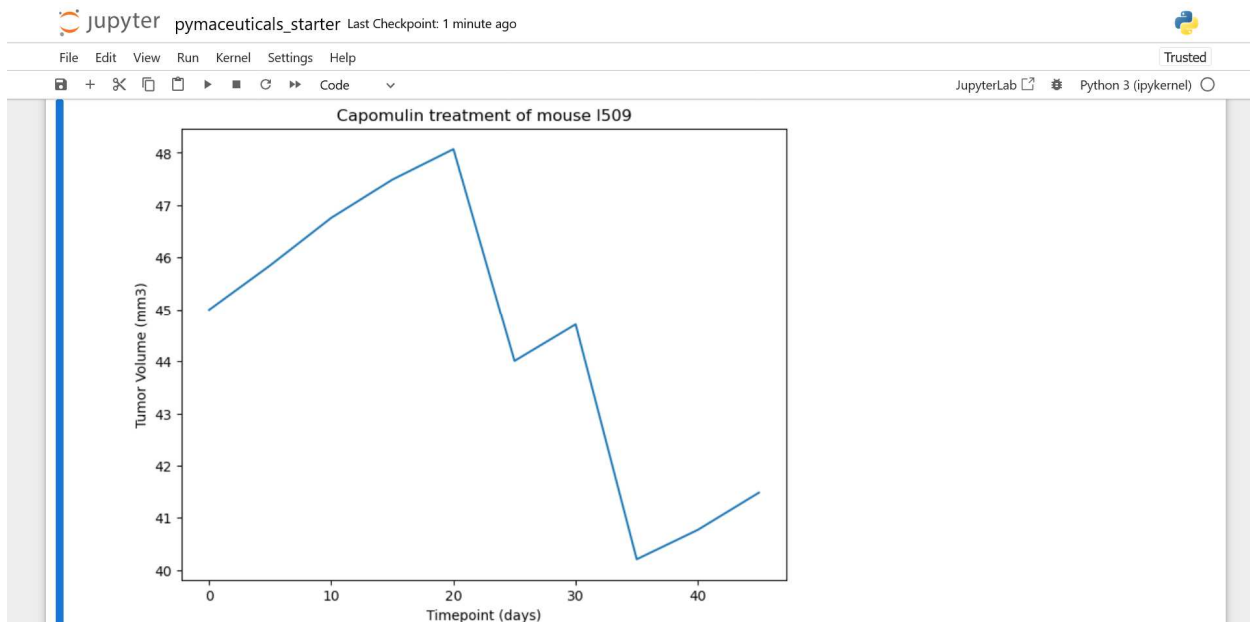
Ramican IQR data:
The lower quartile of Ramican is: 31.56046955
The upper quartile of Ramican is: 40.65908627
The interquartile range of Ramican is: 9.0985367199999998
The the median of Ramican is: 36.56165229
Values below 17.91 could be outliers.
Values above 54.31 could be outliers.
No potential outliers found for Ramican.

Infubinol IQR data:
The lower quartile of Infubinol is: 54.04860769
The upper quartile of Infubinol is: 65.52574285
The interquartile range of Infubinol is: 11.477135160000003
The the median of Infubinol is: 60.16518046
Values below 36.83 could be outliers.
Values above 82.74 could be outliers.
Potential outliers for Infubinol:
Index: 31, Mouse ID: c326, Timepoint: 5, Tumor Volume: 36.3213458

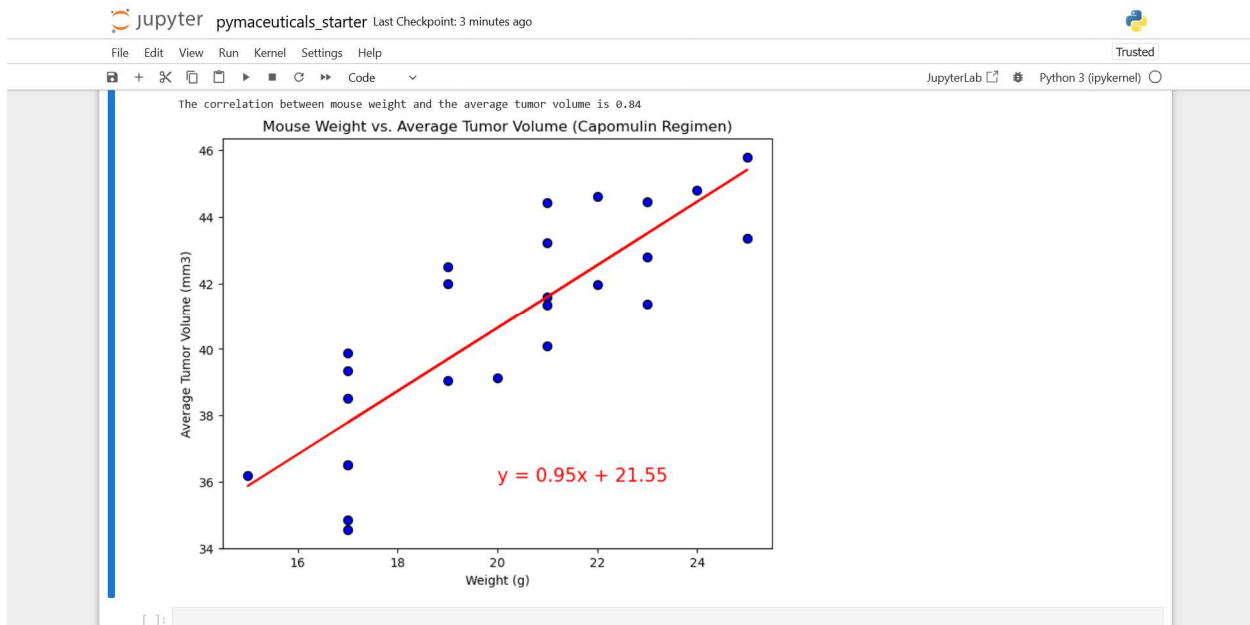
Ceftamin IQR data:
The lower quartile of Ceftamin is: 48.72207785
The upper quartile of Ceftamin is: 64.29983003
The interquartile range of Ceftamin is: 15.577521799999997
The the median of Ceftamin is: 59.85195552
Values below 25.36 could be outliers.
Values above 87.67 could be outliers.
No potential outliers found for Ceftamin.
```



Line and Scatter Plots



Correlation and Regression



Addendum

