1. Explain your data collection process.
I used a clear plastic ruler and measured the leaves from stem to tip. I kept track of leaves I had measured and measured a random sampling from
each plant.

2. What instrument did you use to collect data with?
A clear plastic ruler with 1/16" marks

3. Argue the accuracy and precision of your instrument.
The ruler is a tool built to measure linear distance. It measures inches and fractions of inches which are US customary units and matches my other linear
measuring tools (other rulers/tape measures). The precision is 1/16th of a inch. The abnormal shape of the leaves and the admittedly inconsistent measuring
technique I used make the 1/16" precision more than precise enough

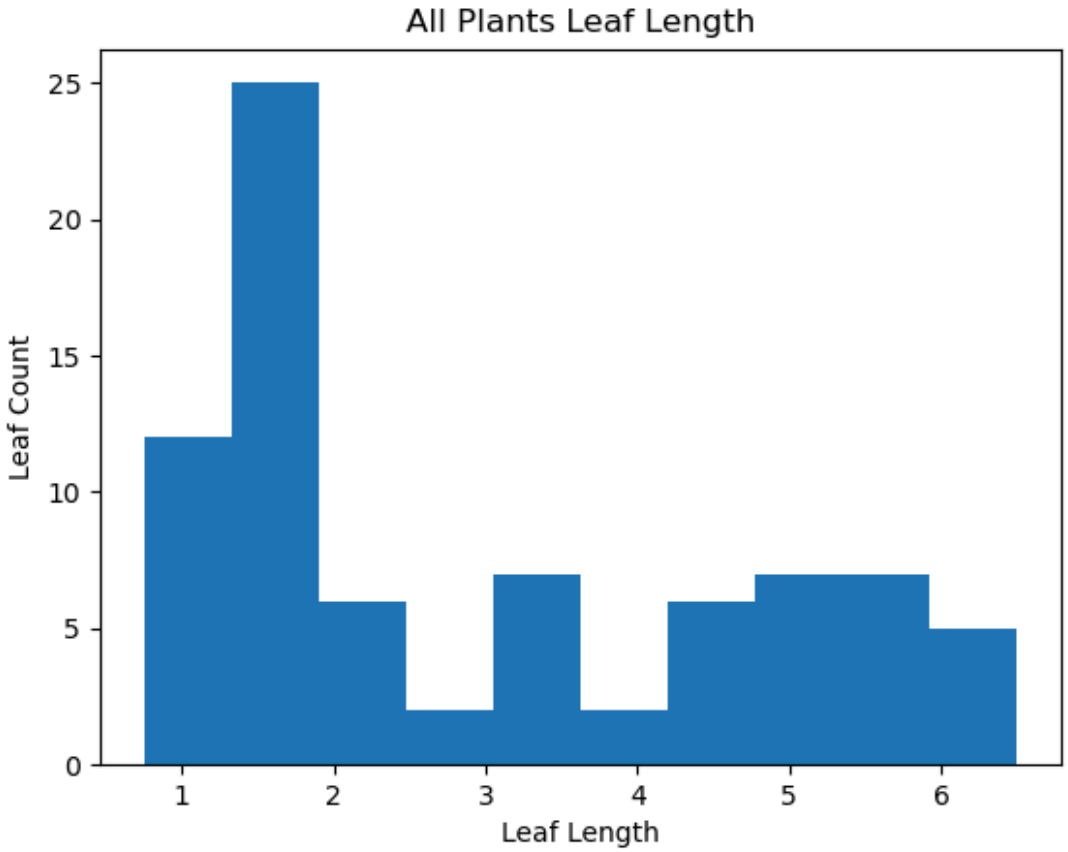
4. How many data points did you collect? Why?
I was able to get 79 total data points. I had 4 plants available (ok I had more, but they don't have leaves) and I just measured around the same amount of leaves
from each plant.

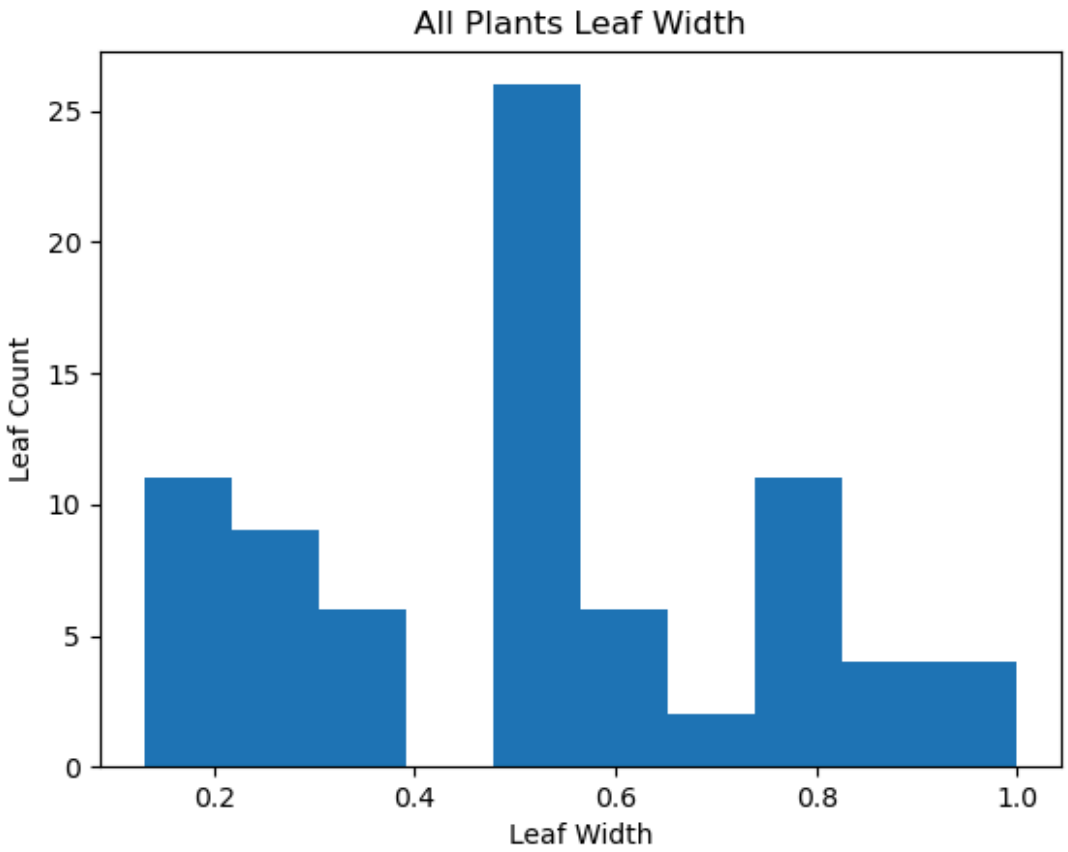
5. Define the size of your data in terms of both N (full data set size) and n (each subset size).
N: 79
Aloe Vera: 18
Corn Plant: 17
Dwarf Umbrella: 24
Parlor Palm: 20

6. Explain any problems that you ran into during the data collection process.
The first problem was figuring out where leaves began on the Aloe Vera and the Corn Plant.

The second problem was figuring out what angle the ruler should be held at on the leaves. Some of the leaves weren't symmetrical, so determining the widest
part was not precise.

The third problem was some of the leaves on the Corn Plant had "dead ends" and I was not sure if those should count towards the length.
I decided they did not.
****
****

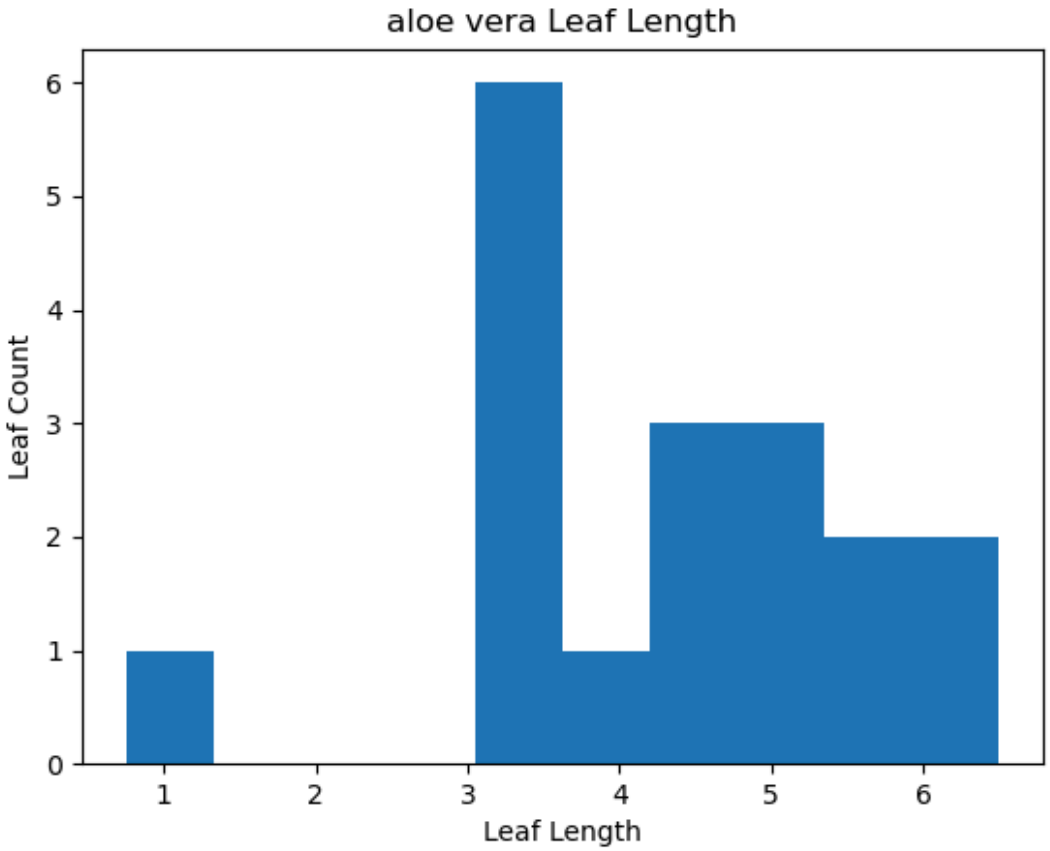

All Plants Leaf Length
Leaf Count
Leaf Length
0
5
10
15
20
25
1
2
3
4
5
6

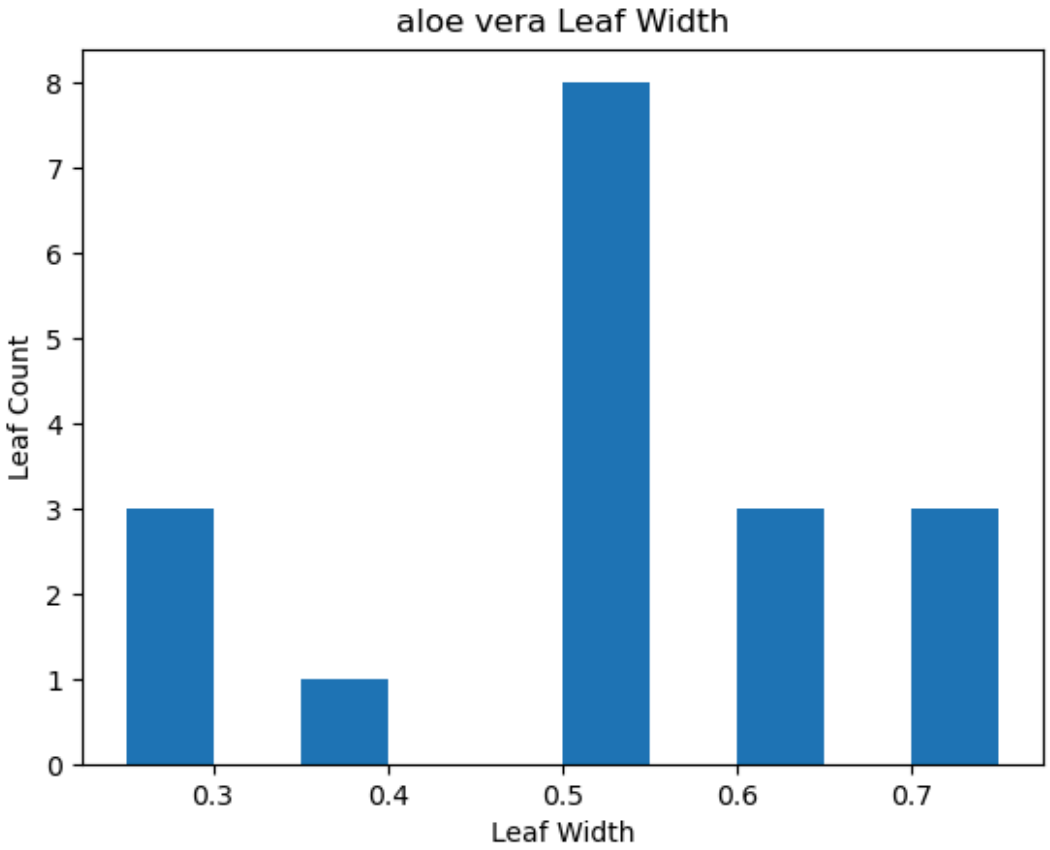


all	length	width
mean	2.98	0.51
median	2.08	0.50
variance	3.10	0.05
standard dev	1.76	0.23

The variance tells me that the data has decent variability. The standard deviation is pretty high considering the range of the data. The mean and median being decently far apart implies this is not a normalized data set.

The widths are more normal, with the mean and the median being fairly close. The “hills” and “valleys” of the histogram do imply that we have multiple populations present, which is true. The variance is pretty small, which makes sense given width is a smaller measurement for the plants I used and I was rounding my measurements.

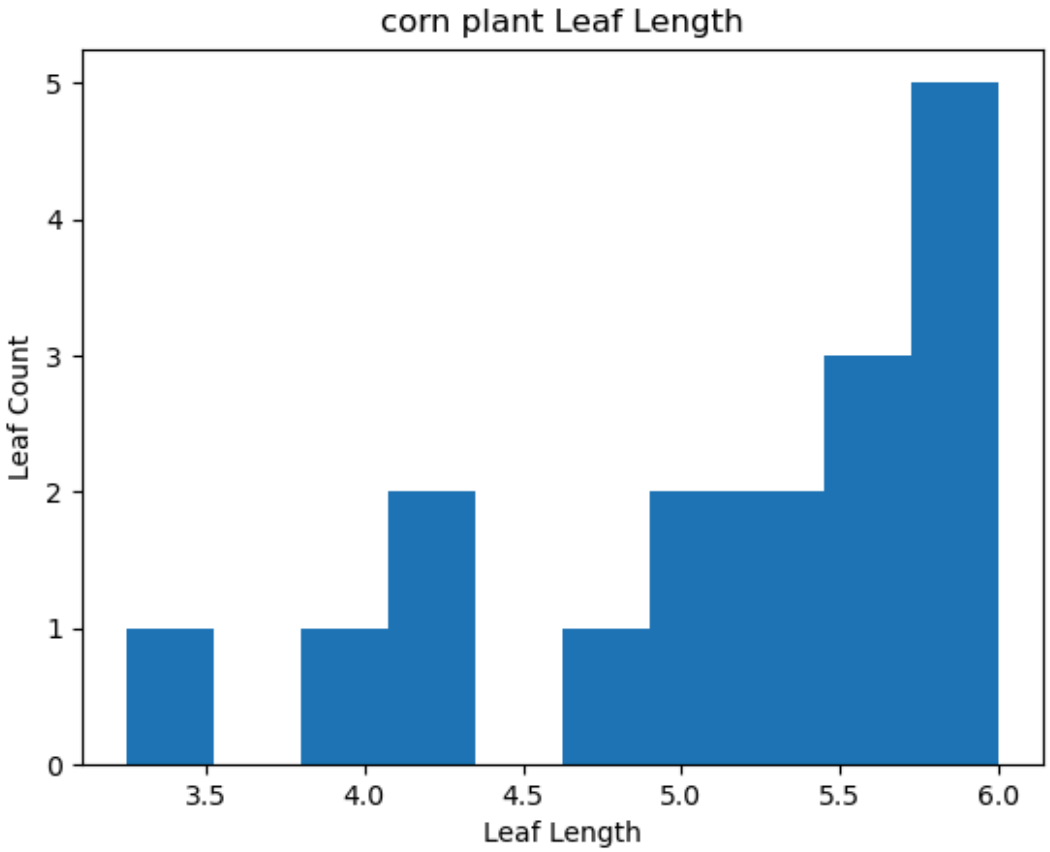

aloe vera Leaf Length
Leaf Count
0
1
2
3
4
5
6
1
2
3
4
5
6
Leaf Length

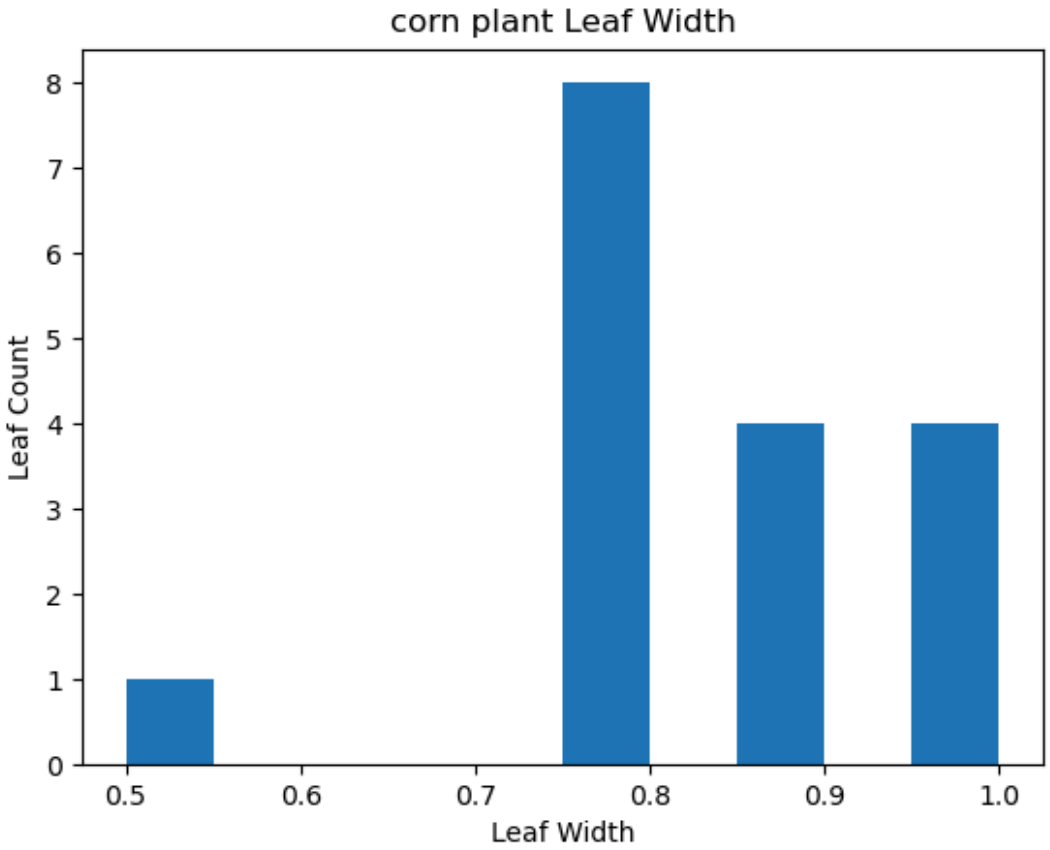


aloe	length	width
mean	4.33	0.51
median	4.50	0.50
variance	1.79	0.03
standard dev	1.34	0.16

Most of the leaves were between 3 and 4 inches long, with one outlier around 1 inch. The variance and standard deviation are still high considering the range of the data. The mean and median are closer than when all plants are considered, which makes sense.

The width of these leaves are very normal compared to the length. The mean and median are almost right on top of each other. The variance and standard deviation reflect how close the measurements of the leaves were to each other.

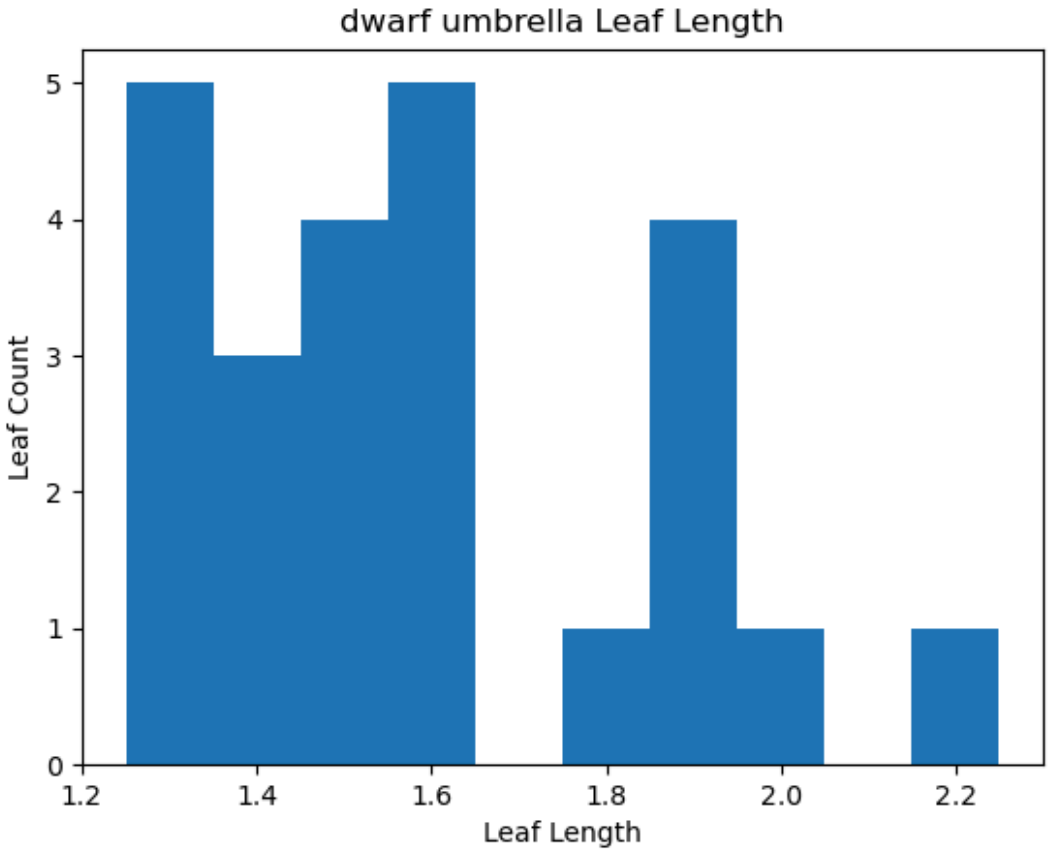

corn plant Leaf Length
Leaf Count
Leaf Length
0
1
2
3
4
5
3.5
4.0
4.5
5.0
5.5
6.0

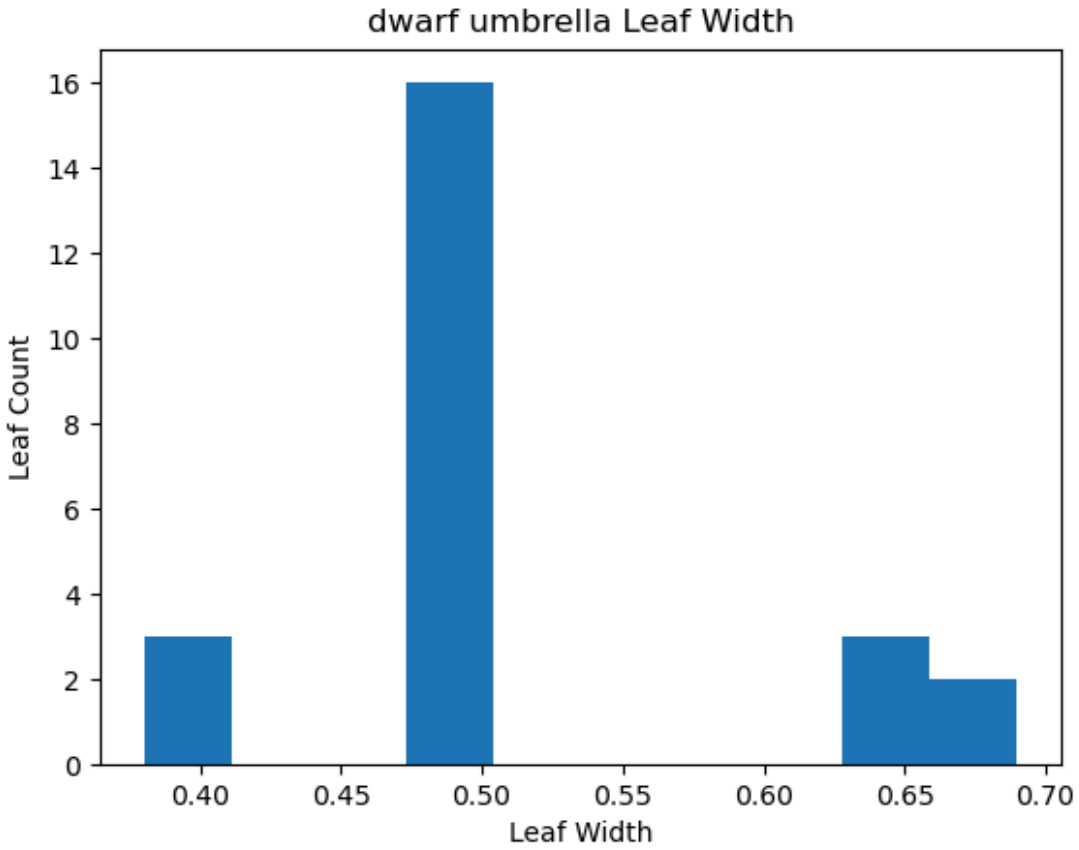


corn plant	length	width
mean	5.13	0.82
median	5.25	0.75
variance	0.64	0.02
standard dev	0.80	0.13

The length, despite the 4 leaves under 4.5 inches long, is very close together with not a lot of variance. The length data is slightly skewed to the right, which is supported by the difference between the mean and median. I do not have much to say here other than this is a goofy looking plant.

The width measurements are much closer together, but follow the same conclusions as the lengths.

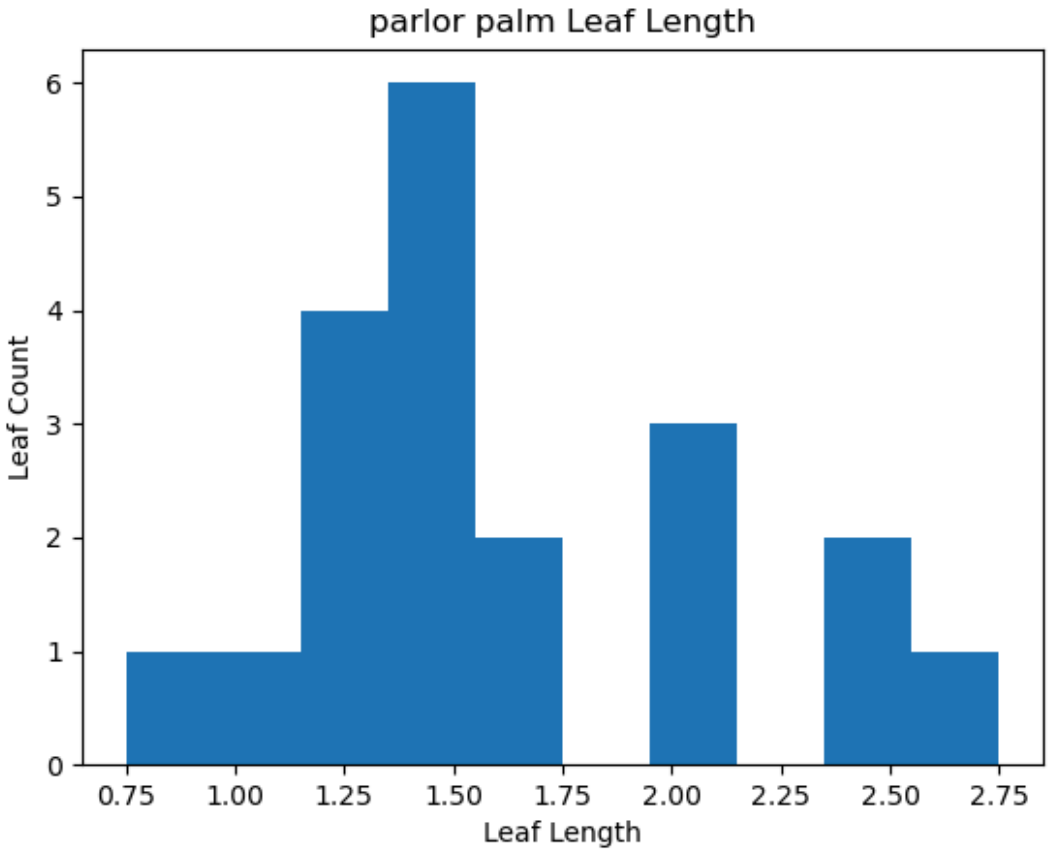

dwarf umbrella Leaf Length
Leaf Count
Leaf Length
0
1
2
3
4
5
1.2
1.4
1.6
1.8
2.0
2.2

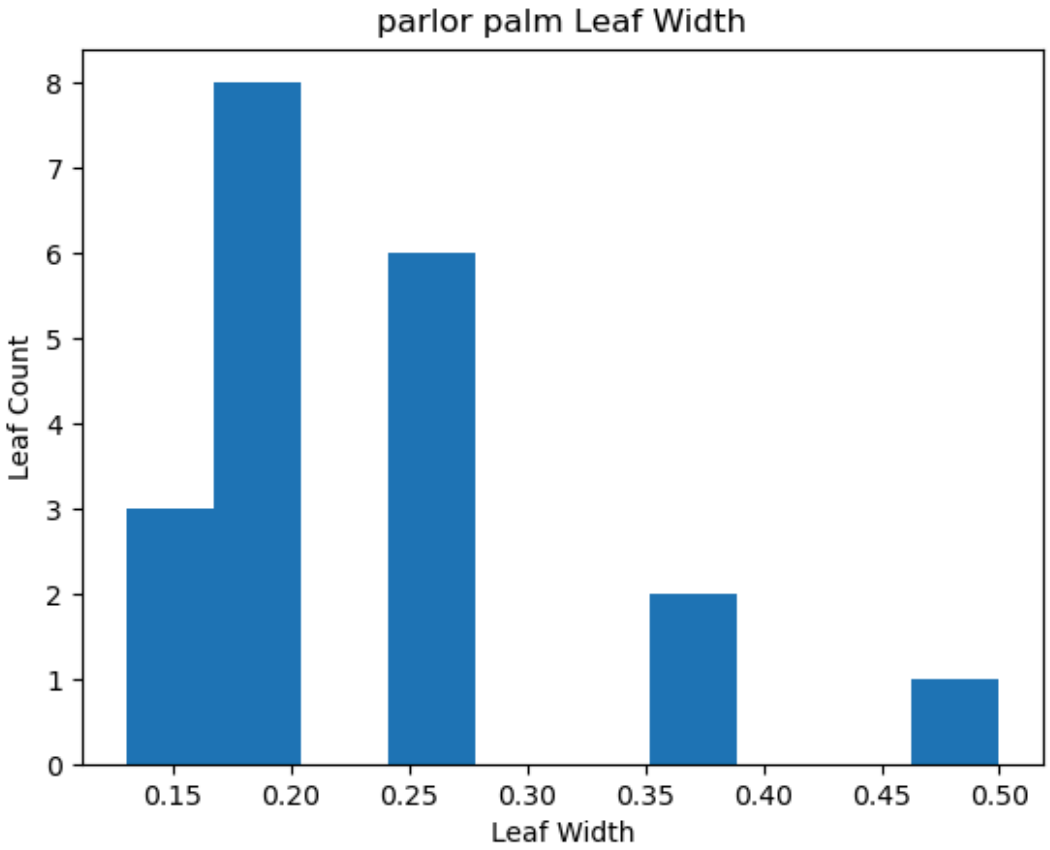


dwarf Umbrella		
mean	1.58	0.52
median	1.56	0.50
variance	0.07	0.01
standard dev	0.27	0.08

The lengths skewed ever so slightly to the left. Seeing the lengths graphed like this does give a pretty good representation of how the leaves look in proportion to each other on the plant. There are somewhat distinct “groups” of lengths the leaves fit into. The data appears to be somewhat normal.

The widths are very close to each other with the mean and median almost the same and very low variance. I bet the data would “spread out” more if I had a more precise measuring device/method.

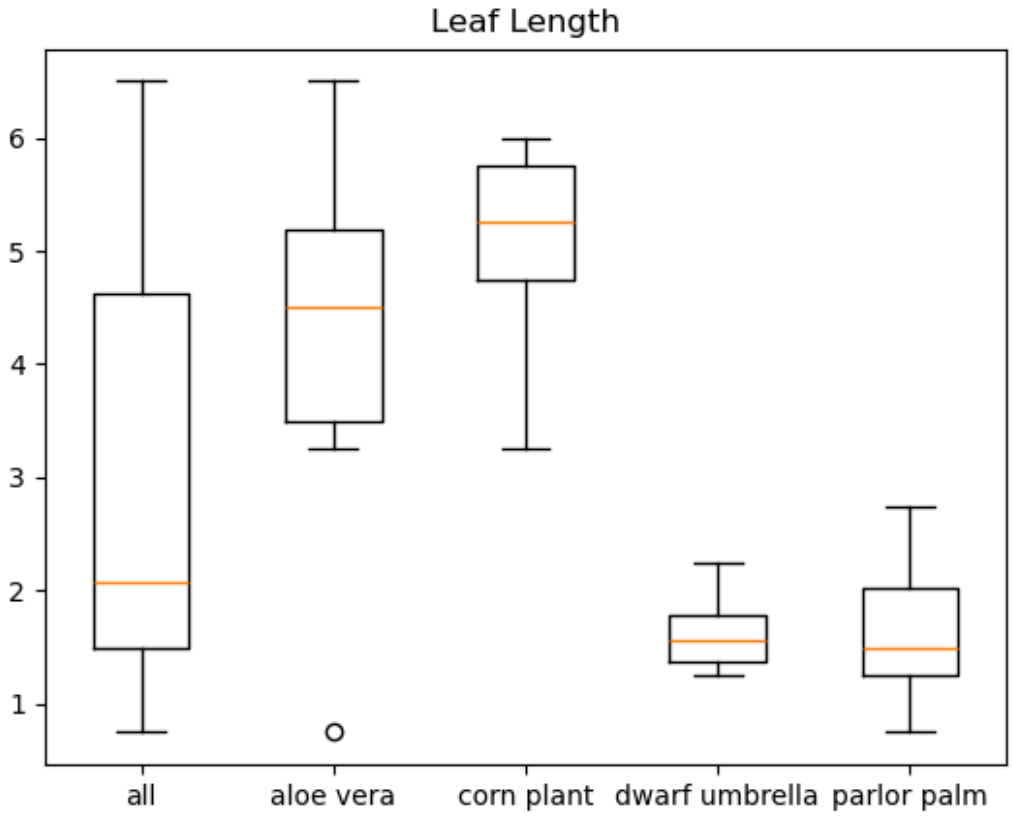

parlor palm Leaf Length
Leaf Count
0
1
2
3
4
5
6
0.75
1.00
1.25
1.50
1.75
2.00
2.25
2.50
2.75
Leaf Length

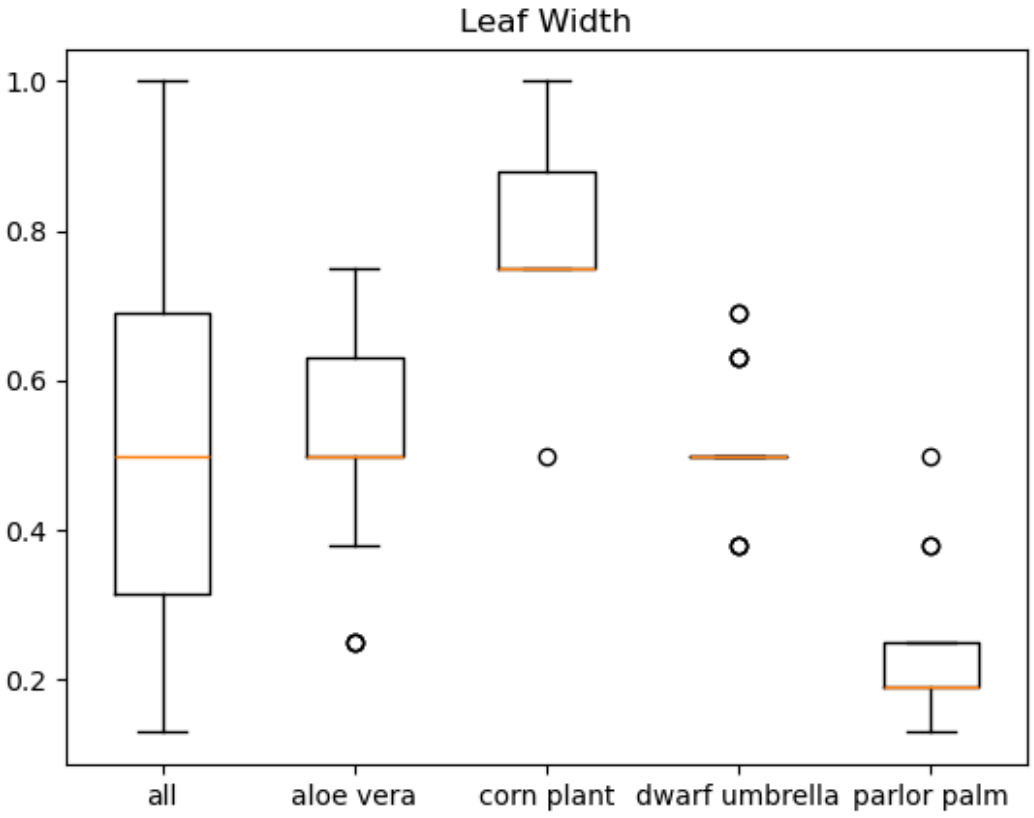


parlor palm	length	width
mean	1.63	0.23
median	1.50	0.19
variance	0.27	0.01
standard dev	0.52	0.09

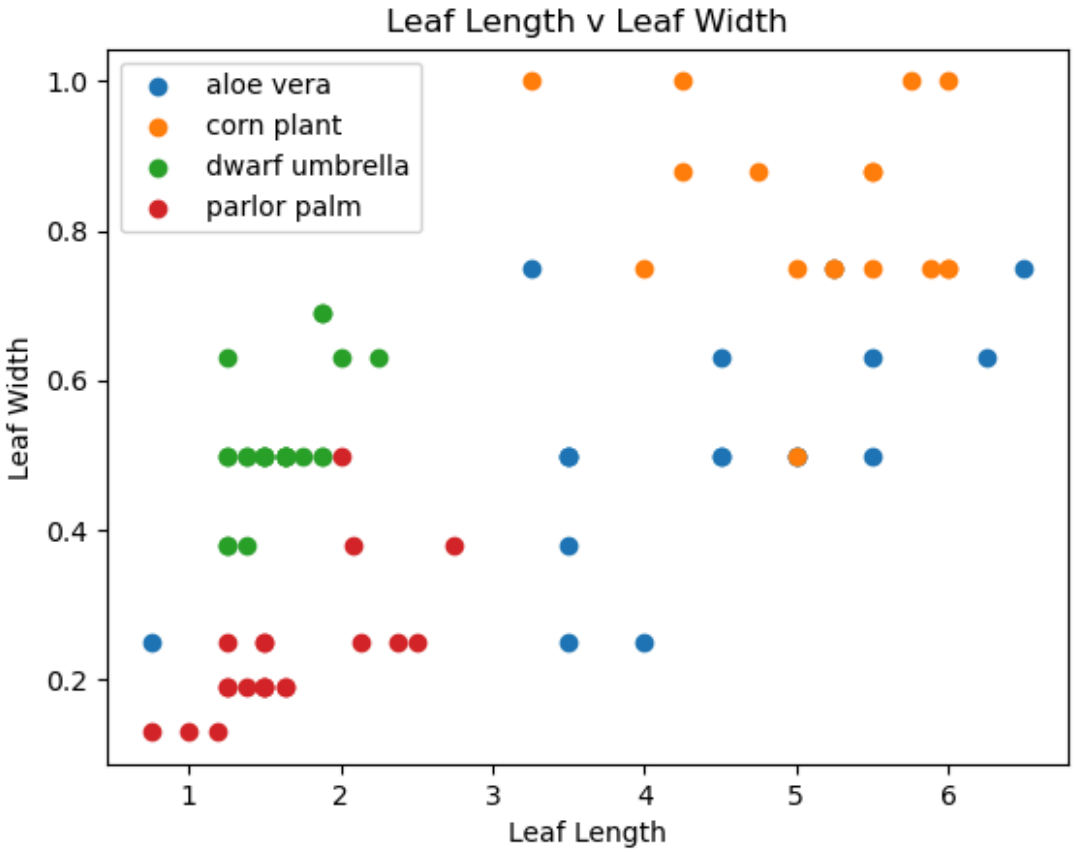
The lengths of the leaves on this plant are very normalized, the graph has a nice bell curve. There are no outliers and the measurements are all decently close to each other.

The widths have 3 “outliers”. Like the length, the width measurements are very close to each other. I think the low variance is due to the lack of precision of my ruler.


Leaf Length
6
5
4
3
2
1
all
aloe vera
corn plant
dwarf umbrella
parlor palm



What immediately sticks out to me is how obvious these graphs make the difference in variance of leaf size. The dwarf umbrella is very consistent, same with the parlor palm. What isn't obvious about this graph is the number of data points for each plant, which could affect conclusions drawn from these graphs. I would argue this graph does demonstrate the variance of each graph fairly well.



The scatter plot shows how the width, length, and plant type cluster together. It also demonstrates the range of the data set as a whole and how each plant type fits into it. Although it doesn't explicitly state how many data points there are, it is shown how many unique measurements were obtained. The spread of the dots for the plant type show the variance for that plant.