Leafs

Kaisa Roggeveen, Scott Graham

March 22nd 2018

Dr. Steven M. Vamosi Associate Dean, Diversity, Equity and Inclusion Professor, Population Biology 2500 University Drive NW Department of Biological Sciences University of Calgary Calgary AB T2N 1N4 Canada

Introduction

The intent of this paper is to develop a method for classifying leaves as either Cherry or Pear, based on their measured length and width. This method was developed for Dr. Steven Vamosi, a botanist from the Universtiy of Calgary.

The classification method used was Linear discriminant analysis, developed by R.A Fischer. To .... To take training samples from a sampled population take measurements, from these measurements classification rules are created. This will then be tested against the classifying sample to see if there are any miss classifications.

Cherry and Pear leaves are both leaves from fruit trees. Cherry trees belong to the genus Prunus and Pear trees belong to the genus Pyrus [2],[3]. A common feature amongst the leaves is that they both have a midrib, which is the centreal vein of the leaf which extends along the leaf's centerline.

Data

Measurement Process

The first step taken in the measurement of the leaves was to give each leaf an identification number based on the species. The method used to meausre the dimensions was to create a box with the minimum length and width in which the entire leaf would be encompassed in the box.

To begin creating the sides of the box, a ruler was aligned parallel to the midrib, which is the central vein in the leaf and moved towards the left and the right of the picture until only one point on the leaf remained [1]. From the single point on the side of the leaf, a line was drawn parallel to the midrib of the leaf.

Next, the base and point of the leaf were measured, a ruler was placed perpendicular to the midrib and the ruler was moved towards to tip of the leaf until a single point remained, a line was draw perpendicular to the midrib at this point. At the base of the leaves the length of the leaf was set as the point where the leaf ends and the stem begins, at this point a line was drawn perpendicular to the midrib.

After all the boxes were created, the width (lines parallel to midrib) and the length (lines perpendicular to midrib) were measured and the results were recorded in a spread sheet.

Data Creation

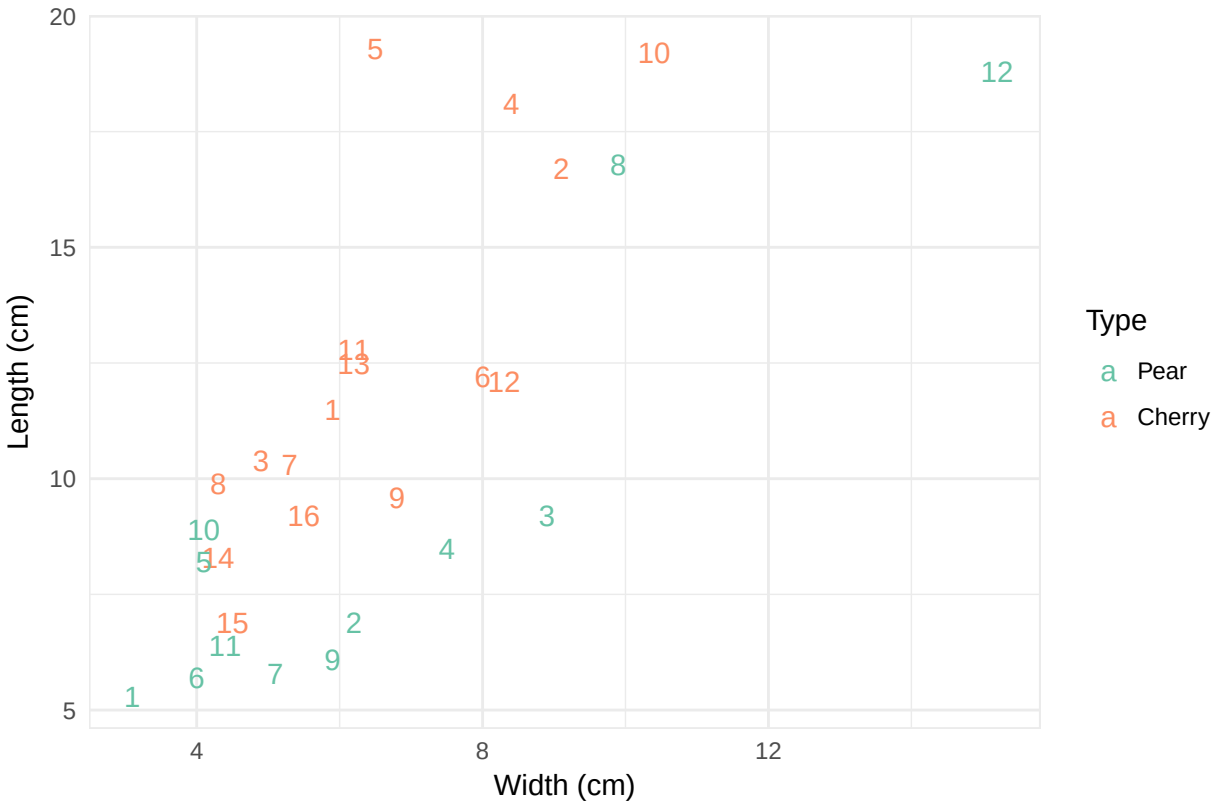
Table 1: Data Summary

Type	Length	Width
Pear :12	Min. : 5.300	Min. : 3.100
Cherry:16	1st Qu.: 7.875	1st Qu.: 4.475

Type	Length	Width
NA	Median : 9.750	Median : 6.050
NA	Mean :10.914	Mean : 6.536
NA	3rd Qu.:12.575	3rd Qu.: 8.075
NA	Max. :19.300	Max. :15.200

In this orginal dataset there are a few issues that need to be ackowledged. The first issues that occured during the data measurements was the result of the leaves that were distributed as the training sample were images, in which the images were not to scale. This resulted in a few outliers, which much larger lengths and widths compared to the other leaves in the set. These outliers included Pear#12, Cherry#10 and Cherry#5. However, based on the nature of this project in just observing the ratio between the length and width, this should not be affected by the size of the image, unless the image was streched in either direction.

Figure XX: Length vs Width Scatter Plot



Classification Procedure (LDA)

Training Data

Table 2: LDA Prior Probabilities

Type	Probability
Pear	0.4286
Cherry	0.5714

Table 3: LDA Group Means

	Length	Width
Pear	8.8833	6.5333
Cherry	12.4375	6.5375

Table 4: LDA Coefficients of Linear Discriminants

Dimension	Coefficient
Length	0.4194
Width	-0.5311

Table 5: LDA Misclassification Results

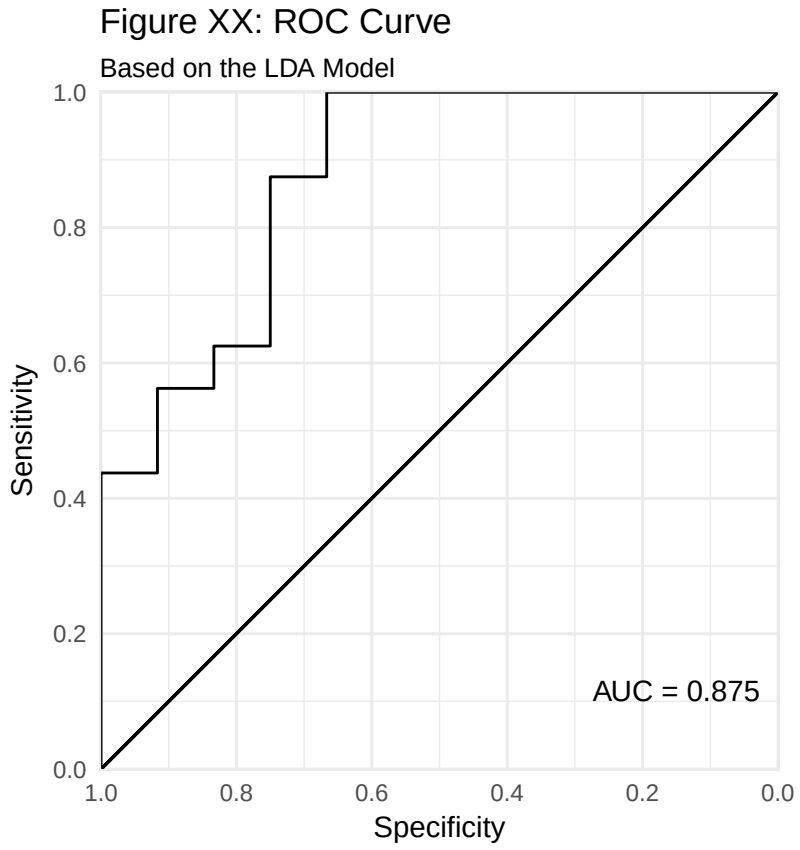
Predicted	Actual	Length	Width	Cherry Probability	Pear Probability	Correct Prediction
Pear	Cherry	9.6	6.8	0.3582	0.6418	FALSE
Pear	Cherry	12.1	8.3	0.4481	0.5519	FALSE
Pear	Cherry	6.9	4.5	0.3892	0.6108	FALSE
Cherry	Pear	8.2	4.1	0.6631	0.3369	FALSE
Cherry	Pear	16.8	9.9	0.8116	0.1884	FALSE
Cherry	Pear	8.9	4.1	0.7528	0.2472	FALSE

Table 6: LDA Confusion Matrix

	Pear	Cherry
Pear	9	3
Cherry	3	13

Table 7: LDA Confusion Matrix Stats

	x
Sensitivity	0.7500000
Specificity	0.8125000
Pos Pred Value	0.7500000
Neg Pred Value	0.8125000
Precision	0.7500000
Recall	0.7500000
F1	0.7500000
Prevalence	0.4285714
Detection Rate	0.3214286
Detection Prevalence	0.4285714
Balanced Accuracy	0.7812500



New Data

Table 8: LDA New Data Predictions

Predicted	Cherry	Pear	Number	Length	Width
Cherry	0.8003229	0.1996771	1	8.2	3.2
Pear	0.2772400	0.7227600	2	5.2	3.8
Cherry	0.5942515	0.4057485	3	7.6	4.0

Observation Space

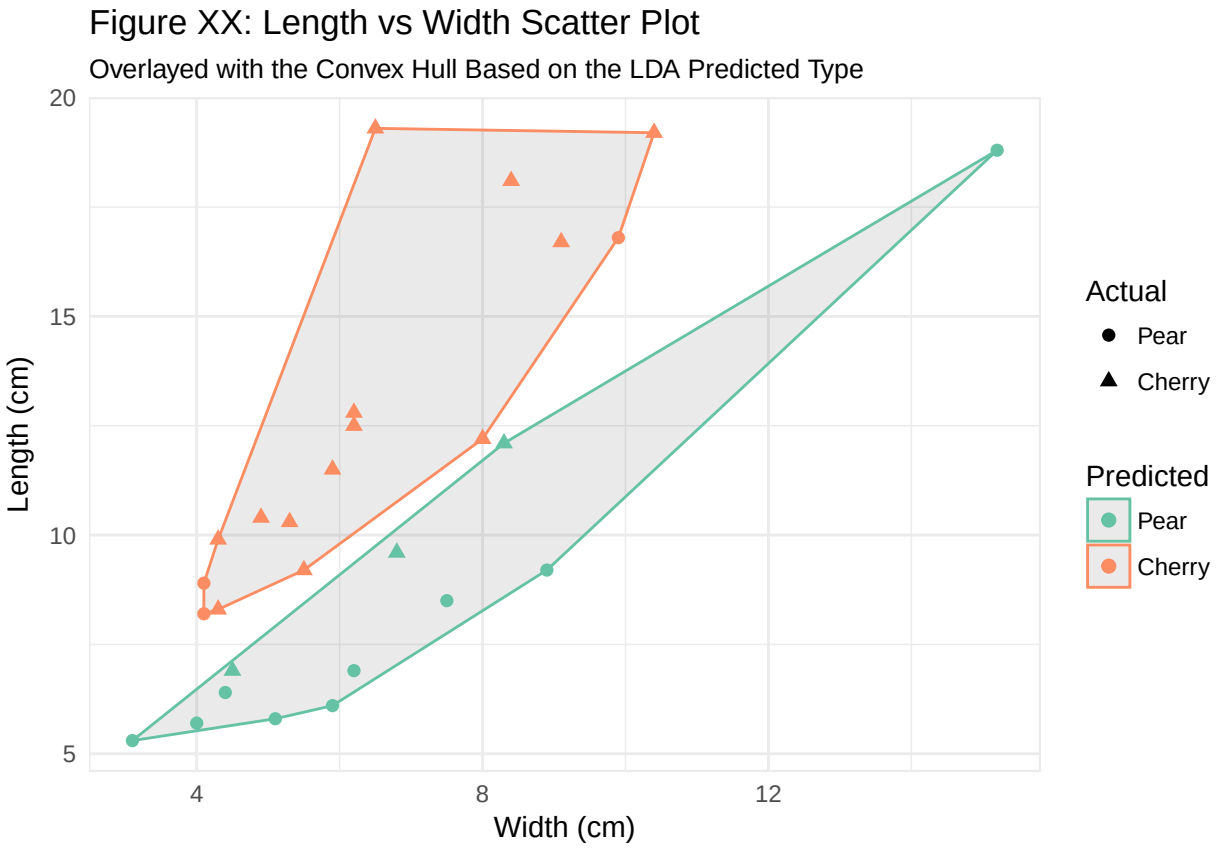
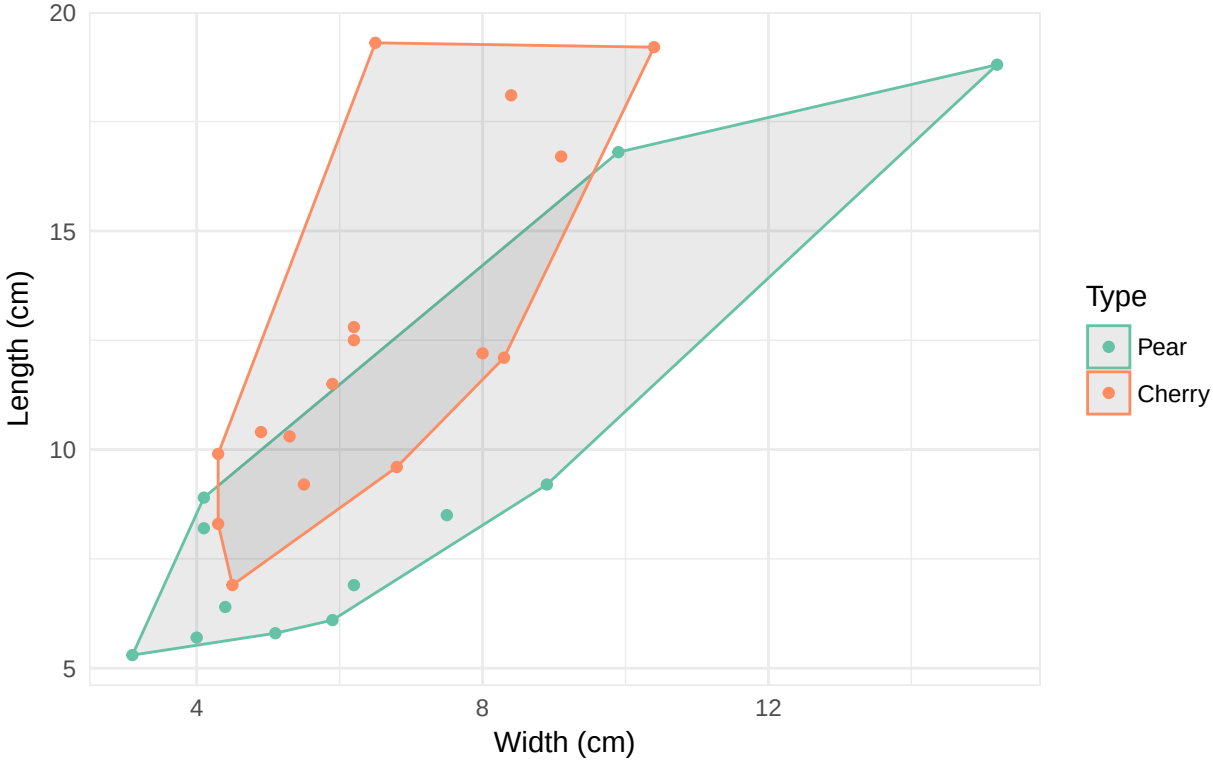


Figure XX: Length vs Width Scatter Plot

Overlayed with the Convex Hull of that Type



Probability Distributions

Contour

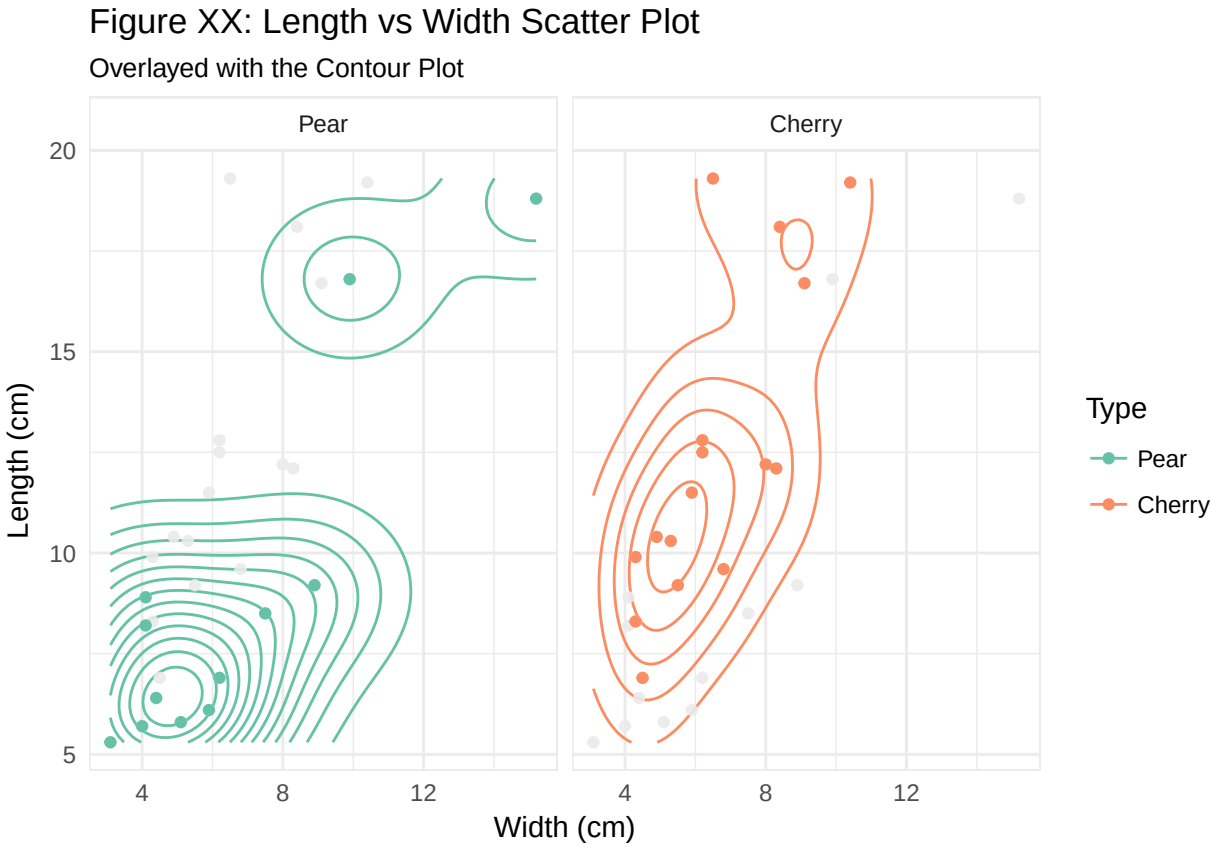


Figure XX: Length vs Width Scatter Plot

Overlayed with a Contour Plot of that Type

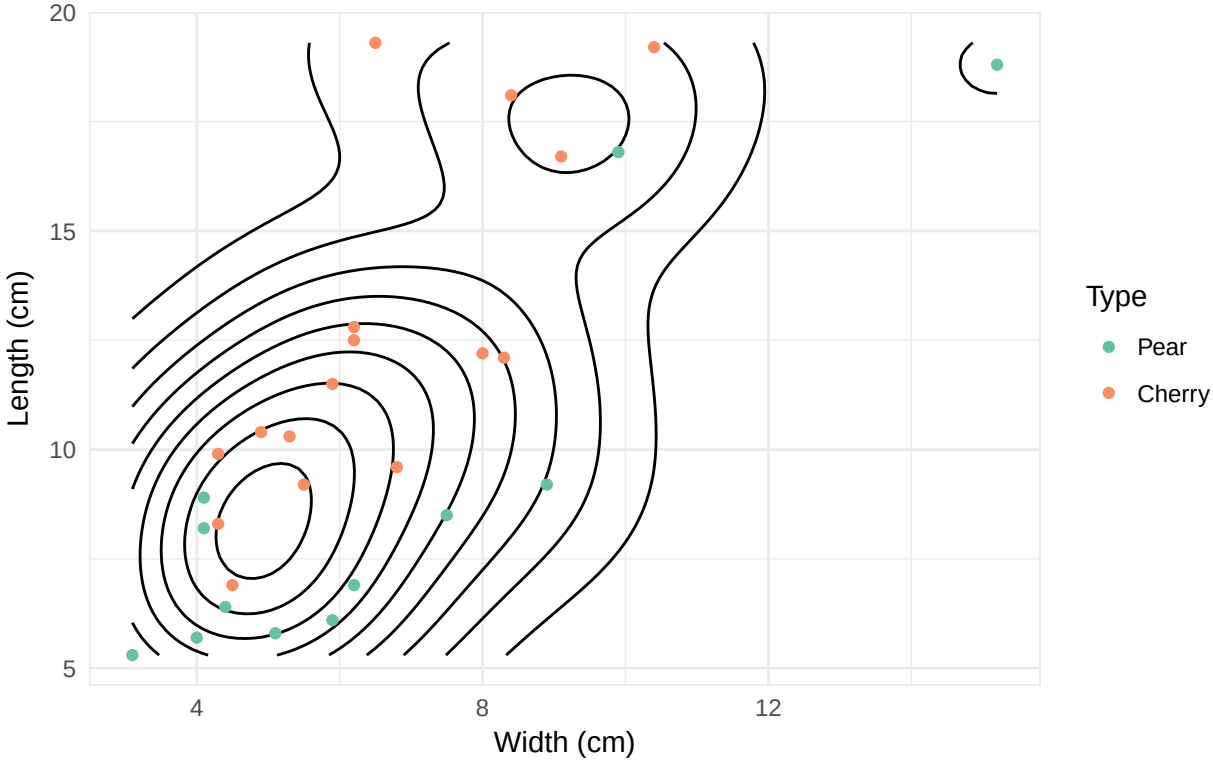


Figure XX: Density Plot by Type

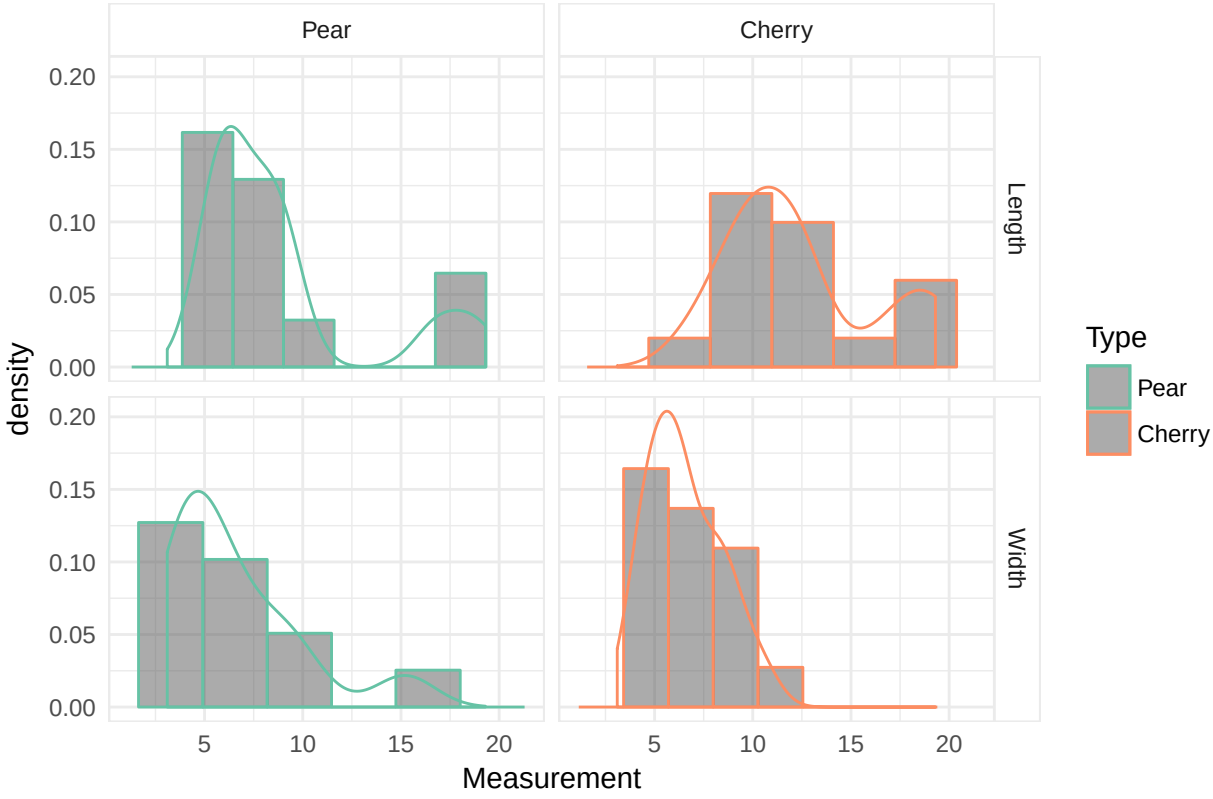
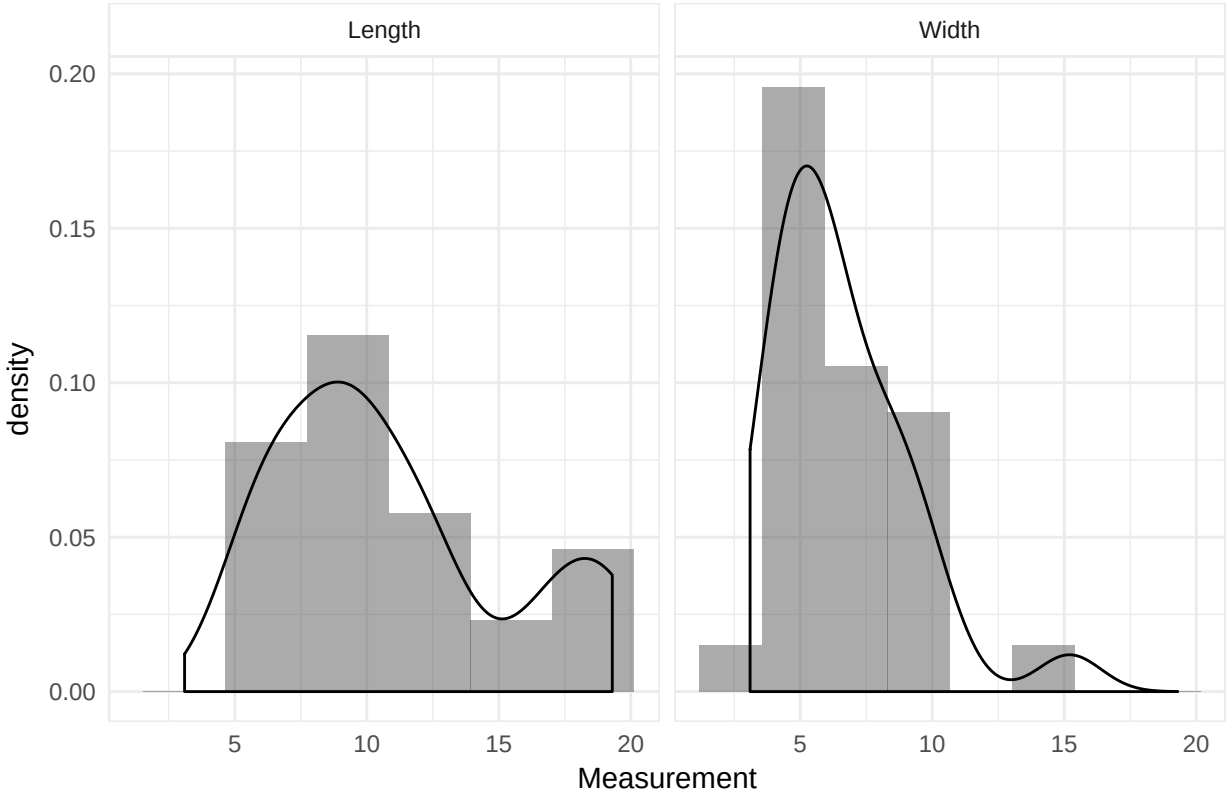


Figure XX: Density Plot



Covariance Matrix

Table 9: Shared Covariance Matrix

	Length	Width
Length	19.422011	8.476508
Width	8.476508	6.685344

Table 10: Cherry Covariance Matrix

	Length	Width
Length	15.047833	5.443833
Width	5.443833	3.357167

Table 11: Pear Covariance Matrix

	Length	Width
Length	19.27788	13.37333
Width	13.37333	11.83152

Classification Procedure (QDA)

Training Data

Table 12: QDA Prior Probabilities

Type	Probability
Pear	0.4286
Cherry	0.5714

Table 13: QDA Group Means

	Length	Width
Pear	8.8833	6.5333
Cherry	12.4375	6.5375

Table 14: QDA Misclassification Results

Predicted	Actual	Length	Width	Cherry Probability	Pear Probability	Correct Prediction
Pear	Cherry	9.6	6.8	0.4686	0.5314	FALSE
Pear	Cherry	12.1	8.3	0.4363	0.5637	FALSE
Pear	Cherry	6.9	4.5	0.4710	0.5290	FALSE
Cherry	Pear	8.2	4.1	0.6431	0.3569	FALSE
Cherry	Pear	16.8	9.9	0.8095	0.1905	FALSE
Cherry	Pear	8.9	4.1	0.7287	0.2713	FALSE

Table 15: QDA Confusion Matrix

	Pear	Cherry
Pear	9	3
Cherry	3	13

Table 16: QDA Confusion Matrix Stats

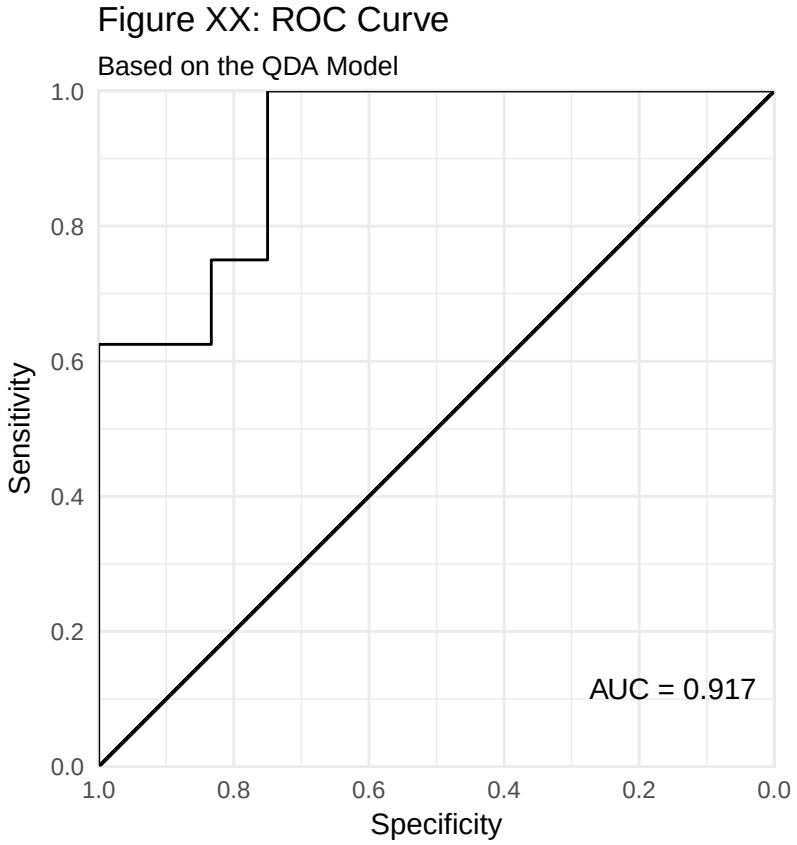
	x
Sensitivity	0.7500000
Specificity	0.8125000
Pos Pred Value	0.7500000
Neg Pred Value	0.8125000
Precision	0.7500000
Recall	0.7500000
F1	0.7500000
Prevalence	0.4285714
Detection Rate	0.3214286
Detection Prevalence	0.4285714
Balanced Accuracy	0.7812500

x

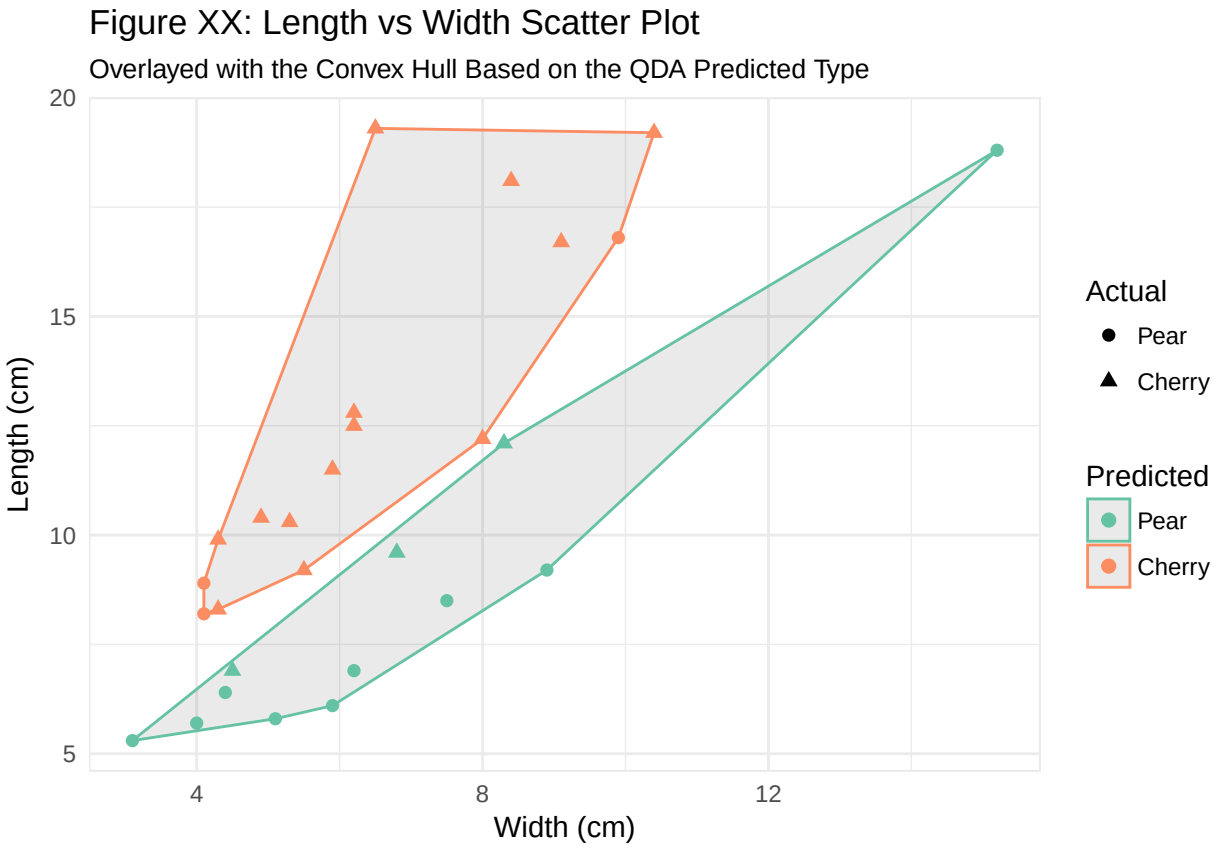
New Data

Table 17: QDA New Data Predictions

Predicted	Cherry	Pear	Number	Length	Width
Cherry	0.6362745	0.3637255	1	8.2	3.2
Pear	0.3382849	0.6617151	2	5.2	3.8
Cherry	0.5712093	0.4287907	3	7.6	4.0



Observation Space



Classification Procedure (GLM)

Training Data

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-1.74	1.519	-1.145	0.2522
Length	0.7764	0.2875	2.7	0.006931
Width	-0.9338	0.3815	-2.448	0.01438

(Dispersion parameter for binomial family taken to be 1)

Table 19: Logit Misclassification Results

Null deviance:	38.24 on 27 degrees of freedom
Residual deviance:	24.26 on 25 degrees of freedom

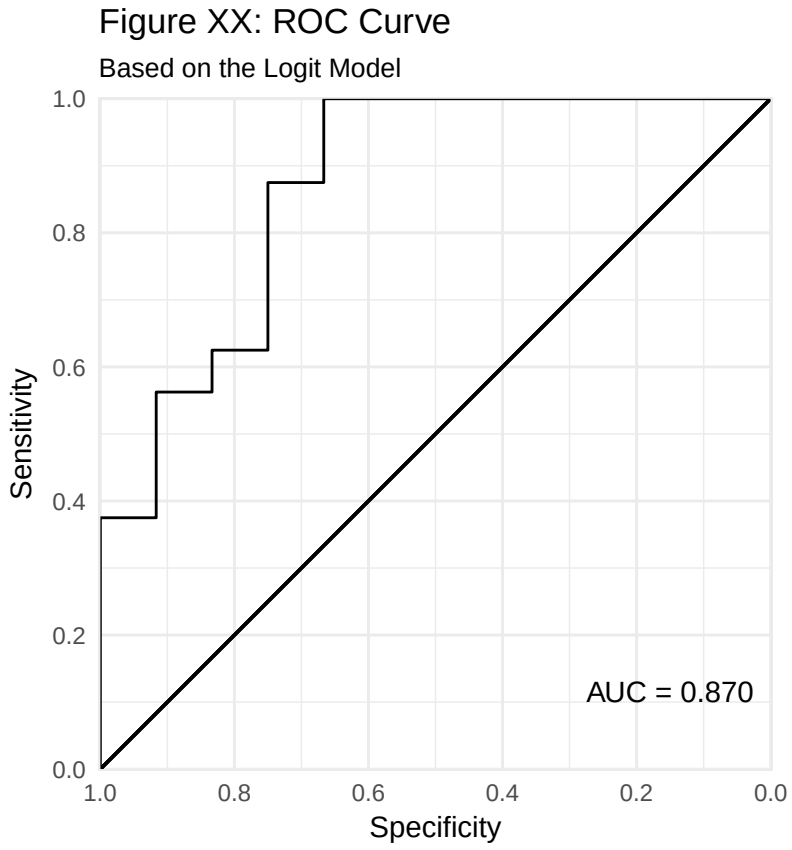
Table 20: Logit Confusion Matrix

Predicted	Actual	Length	Width	Cherry Probability	Pear Probability	Correct Prediction
Pear	Cherry	9.6	6.8	0.3460	0.6540	FALSE
Pear	Cherry	12.1	8.3	0.4759	0.5241	FALSE
Pear	Cherry	6.9	4.5	0.3578	0.6422	FALSE
Cherry	Pear	8.2	4.1	0.6895	0.3105	FALSE
Cherry	Pear	16.8	9.9	0.8868	0.1132	FALSE
Cherry	Pear	8.9	4.1	0.7927	0.2073	FALSE

Table 21: Logit Confusion Matrix Stats

	Pear	Cherry
Pear	9	3
Cherry	3	13

	x
Sensitivity	0.7500000
Specificity	0.8125000
Pos Pred Value	0.7500000
Neg Pred Value	0.8125000
Precision	0.7500000
Recall	0.7500000
F1	0.7500000
Prevalence	0.4285714
Detection Rate	0.3214286
Detection Prevalence	0.4285714
Balanced Accuracy	0.7812500

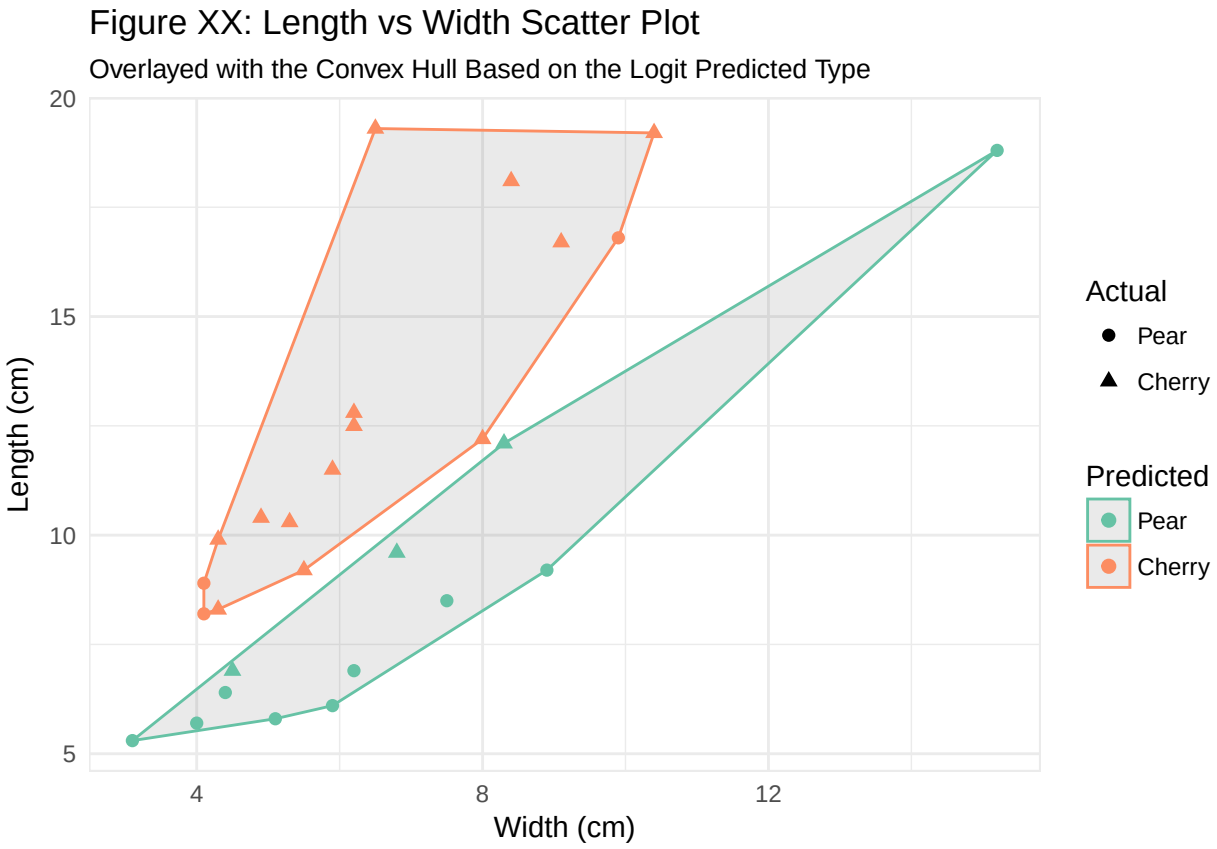


New Data

Table 23: Logit New Data Predictions

Predicted	Cherry Probability	Pear Probability	Number	Length	Width
Cherry	0.8373176	0.1626824	1	8.2	3.2
Pear	0.2225230	0.7774770	2	5.2	3.8
Cherry	0.6047999	0.3952001	3	7.6	4.0

Observation Space



Conclusion

Appendix

Appendix A

Table 24: Data

Number By Type	Type	Length	Width
1	Cherry	11.5	5.9
2	Cherry	16.7	9.1
3	Cherry	10.4	4.9
4	Cherry	18.1	8.4
5	Cherry	19.3	6.5
6	Cherry	12.2	8.0
7	Cherry	10.3	5.3
8	Cherry	9.9	4.3
9	Cherry	9.6	6.8
10	Cherry	19.2	10.4
11	Cherry	12.8	6.2

Number By Type	Type	Length	Width
12	Cherry	12.1	8.3
13	Cherry	12.5	6.2
14	Cherry	8.3	4.3
15	Cherry	6.9	4.5
16	Cherry	9.2	5.5
1	Pear	5.3	3.1
2	Pear	6.9	6.2
3	Pear	9.2	8.9
4	Pear	8.5	7.5
5	Pear	8.2	4.1
6	Pear	5.7	4.0
7	Pear	5.8	5.1
8	Pear	16.8	9.9
9	Pear	6.1	5.9
10	Pear	8.9	4.1
11	Pear	6.4	4.4
12	Pear	18.8	15.2

References

[1] The Parts of a Leaf. (17, October 30). Retrieved March 20, 18, from http://www.robinsonlibrary.com/science/botany/anatomy/leafparts.htm [2] Britannica, T. E. (2016, November 11). Cherry. Retrieved March 20, 2018, from https://www.britannica.com/plant/cherry

[3] Britannica, T. E. (2015, May 13). Pear. Retrieved March 20, 2018, from https://www.britannica.com/plant/pear