

**Name: Akond Rahman**  
**Unity ID: aarahman**  
**Title: HW#4**  
**Date: April 13, 2016**

Answer to Question 1:

For code see the following:

Naïve Bayes: nb.R  
kNN: knn3.R  
Decision Tree with entropy: dtEntropy.R  
Decision tree with Gini: dtGini.R  
MLP: mlp.R

Results:

kNN classifier with k=3:

	Class =3	Class = 8
Class =3	<b>163</b>	6
Class = 8	3	<b>160</b>

Individual class accuracy for class = 3: 0.981

Individual class accuracy for class = 8: 0.963

Overall class accuracy = 0.973

kNN classifier with k=5:

	Class =3	Class = 8
Class =3	<b>162</b>	7
Class = 8	4	<b>159</b>

Individual class accuracy for class = 3: 0.976

Individual class accuracy for class = 8: 0.957

Overall class accuracy = 0.967

kNN classifier with k=50:

	Class =3	Class = 8
Class =3	<b>161</b>	18
Class = 8	5	<b>148</b>

Individual class accuracy for class = 3: 0.968

Individual class accuracy for class = 8: 0.891

Overall class accuracy = 0.931

Naïve Bayes with laplace=0.0:

	Class =3	Class = 8
Class =3	<b>166</b>	166
Class = 8	0	<b>0</b>

Individual class accuracy for class = 3: 1.0

Individual class accuracy for class = 8: 0.0

Overall class accuracy = 0.50

Naïve Bayes with laplace=0.25:

	Class =3	Class = 8
Class =3	<b>166</b>	166
Class = 8	0	<b>0</b>

Individual class accuracy for class = 3: 1.00

Individual class accuracy for class = 8: 0.0

Overall class accuracy = 0.50

Naïve Bayes with laplace=0.50:

	Class =3	Class = 8
Class =3	<b>166</b>	166
Class = 8	0	<b>0</b>

Individual class accuracy for class = 3: 1.00

Individual class accuracy for class = 8: 0.00

Overall class accuracy = 0.50

Linear SVM:

	Class =3	Class = 8
Class =3	<b>159</b>	4
Class = 8	7	<b>162</b>

Individual class accuracy for class = 3: 0.957

Individual class accuracy for class = 8: 0.975

Overall class accuracy = 0.9668

Polynomial SVM:

	Class =3	Class = 8
Class =3	<b>159</b>	10
Class = 8	7	<b>156</b>

Individual class accuracy for class = 3: 0.957

Individual class accuracy for class = 8: 0.939

Overall class accuracy = 0.948

Sigmoid SVM:

	Class =3	Class = 8
Class =3	<b>159</b>	7
Class = 8	7	<b>159</b>

Individual class accuracy for class = 3: 0.957

Individual class accuracy for class = 8: 0.957

Overall class accuracy = 0.957

Linear SVM with cost=10:

	Class =3	Class = 8
Class =3	<b>159</b>	4
Class = 8	7	<b>162</b>

Individual class accuracy for class = 3: 0.957

Individual class accuracy for class = 8: 0.975

Overall class accuracy = 0.9668

Linear SVM with cost=100:

	Class =3	Class = 8
Class =3	<b>159</b>	4
Class = 8	7	<b>162</b>

Individual class accuracy for class = 3: 0.957

Individual class accuracy for class = 8: 0.975

Overall class accuracy = 0.9668

Multi-layer perceptron with 10 trials:

	Class =3	Class = 8
Class =3	<b>0</b>	0
Class = 8	166	<b>166</b>

Individual class accuracy for class = 3: 0  
 Individual class accuracy for class = 8: 1.0  
 Overall class accuracy = 0.50

Multi-layer perceptron with 100 trials:

	Class =3	Class = 8
Class =3	<b>0</b>	0
Class = 8	166	<b>166</b>

Individual class accuracy for class = 3: 0.0  
 Individual class accuracy for class = 8: 1.00  
 Overall class accuracy = 0.50

Multi-layer perceptron with 100 trials, and size=50:

	Class =3	Class = 8
Class =3	<b>0</b>	0
Class = 8	166	<b>166</b>

Individual class accuracy for class = 3: 0.0  
 Individual class accuracy for class = 8: 1.00  
 Overall class accuracy = 0.50

Multi-layer perceptron with 100 trials, and size=100:

	Class =3	Class = 8
Class =3	<b>0</b>	0
Class = 8	166	<b>166</b>

Individual class accuracy for class = 3: 0.0  
 Individual class accuracy for class = 8: 1.00  
 Overall class accuracy = 0.50

Decision tree with C5.0 that uses entropy:

	Class =3	Class = 8
Class =3	<b>156</b>	16
Class = 8	10	<b>150</b>

Individual class accuracy for class = 3: 0.939

Individual class accuracy for class = 8: 0.925  
Overall class accuracy = 0.921

Decision tree with C5.0 that uses entropy, with 10 trials:

	Class =3	Class = 8
Class =3	<b>159</b>	6
Class = 8	7	<b>160</b>

Individual class accuracy for class = 3: 0.957  
Individual class accuracy for class = 8: 0.963  
Overall class accuracy = 0.961

Decision tree with C5.0 that uses entropy with 50 trials:

	Class =3	Class = 8
Class =3	<b>162</b>	3
Class = 8	4	<b>163</b>

Individual class accuracy for class = 3: 0.976  
Individual class accuracy for class = 8: 0.982  
Overall class accuracy = 0.978

Decision tree with C5.0 that uses entropy with 100 trials:

	Class =3	Class = 8
Class =3	<b>162</b>	3
Class = 8	4	<b>163</b>

Individual class accuracy for class = 3: 0.976  
Individual class accuracy for class = 8: 0.982  
Overall class accuracy = 0.78

Decision tree with RPART that uses gini:

	Class =3	Class = 8
Class =3	<b>152</b>	11
Class = 8	14	<b>155</b>

Individual class accuracy for class = 3: 0.915  
Individual class accuracy for class = 8: 0.9337  
Overall class accuracy = 0.924

Decision tree with RPART that uses gini, cp=0.01:

	Class =3	Class = 8
Class =3	<b>152</b>	11
Class = 8	14	<b>155</b>

Individual class accuracy for class = 3: 0.915  
Individual class accuracy for class = 8: 0.9337  
Overall class accuracy = 0.9246

Decision tree with RPART that uses gini, cp=0.10:

	Class =3	Class = 8
Class =3	<b>145</b>	18
Class = 8	21	<b>148</b>

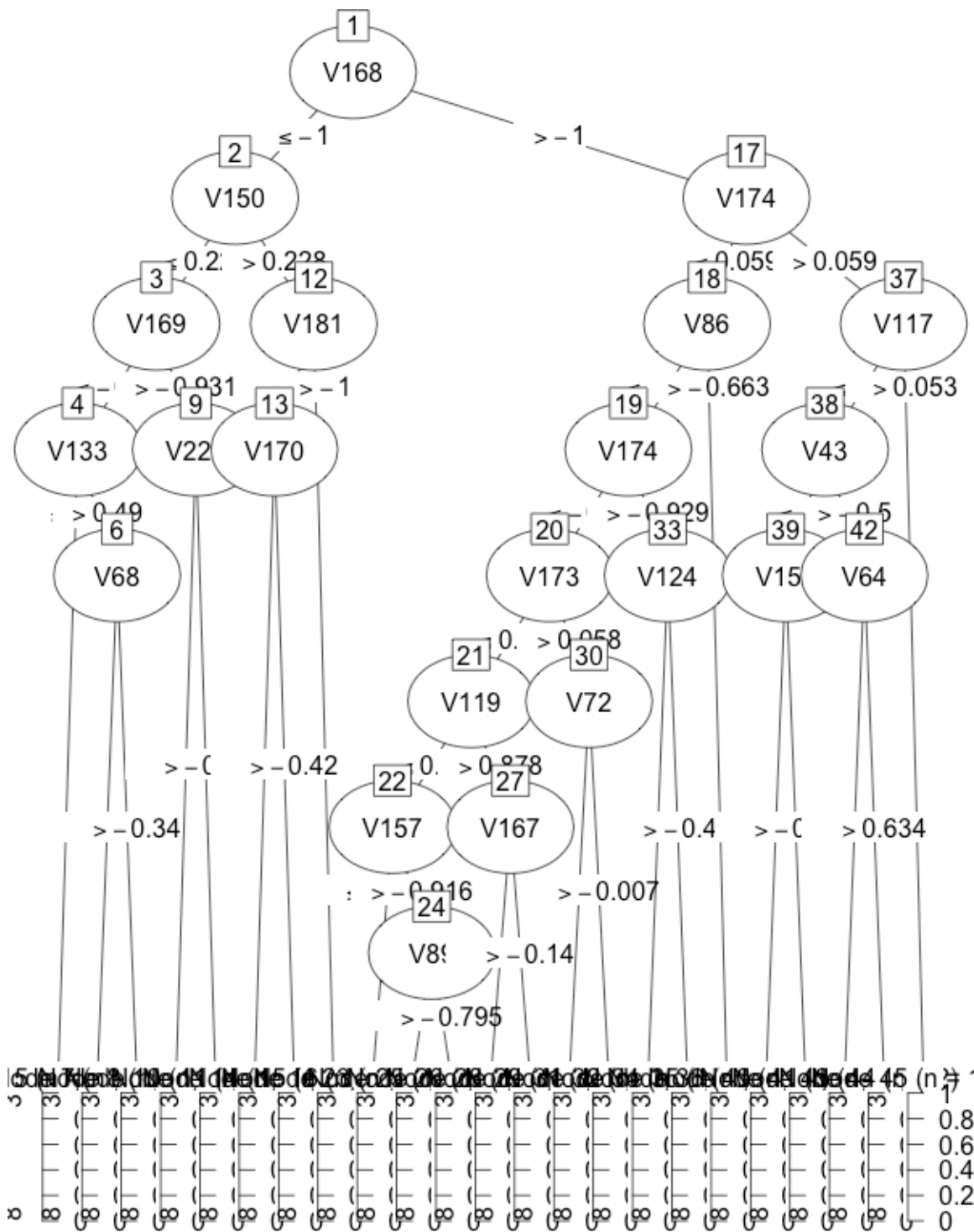
Individual class accuracy for class = 3: 0.873  
Individual class accuracy for class = 8: 0.891  
Overall class accuracy = 0.8825

Decision tree with RPART that uses gini, cp=0.25:

	Class =3	Class = 8
Class =3	<b>145</b>	18
Class = 8	21	<b>148</b>

Individual class accuracy for class = 3: 0.873  
Individual class accuracy for class = 8: 0.891  
Overall class accuracy = 0.882

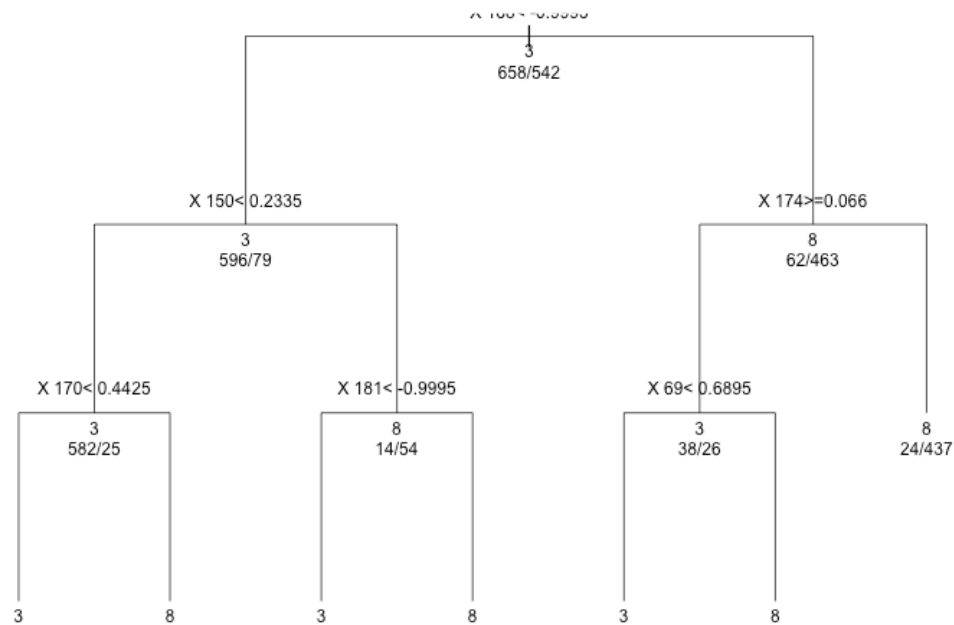
Plots for decision tree using entropy:



Plots for decision tree using Gini:



### Classification Tree Using Gini



Decision tree with error:

No implementation of decision tree with classification error was found

Explanation:

Decision tree with classification errors can generate a gain of zero using one split incorrectly generating rules for classification, whereas, using entropy and Gini, more branches of the tree will be generated. For this reason, implementation of decision tree algorithms with classification error has not been developed.

Interpretation of results:

With respect to overall accuracy decision tree with C5.0 was the best classifier, followed by the k-NN classifier with three neighbors. The C5.0 classifier performed the best, as it was able to create the classification rules with maximum gain based on the features. The relative advantages and disadvantages of the classifiers are presented below:

k-NN:

Advantages:

- simplicity
- fast for small number of neighbors

Disadvantages:

- depends on a distance metric

- lazy learner
- might suffer from the curse of dimensionality

Linear SVM:

Advantages:

- delivers a unique solution as it optimizes the decision boundary
- less susceptible to noise

Disadvantages:

- lack of transparency
- can be come computationally expensive

Decision Tree:

Advantages:

- ease of interpretation
- can explain non-linear models
- construction and prediction can be speedy .

Disadvantages:

- susceptible to noise
- computationally expensive

MLP:

Advantages:

- Efficient computationally
- Can be used as universal computing machines

Disadvantages:

- Convergence can be slow
- Can suffer form local minima

Answer to Question 2(a, b, c, d):

For code see:

Naïve Bayes: naive\_bayes.R

kNN: mlc.R

Results for MLC:

	1	2	3	4	5
1	<b>133</b>	0	40	0	1

2	0	<b>74</b>	21	0	0
3	19	0	<b>360</b>	0	0
4	0	0	0	<b>54</b>	20
5	2	0	4	6	<b>57</b>

Individual class accuracy for class-1: 0.985

Individual class accuracy for class-2: 0.9135

Individual class accuracy for class-3: 0.9756

Individual class accuracy for class-4: 0.857

Individual class accuracy for class-5: 0.7916

Overall accuracy: 0.9416

Results for Naïve Bayes:

	1	2	3	4	5
1	<b>114</b>	0	40	0	1
2	0	<b>81</b>	21	0	0
3	19	0	<b>304</b>	0	0
4	0	0	0	<b>57</b>	20
5	2	0	4	6	<b>51</b>

Individual class accuracy for class-1: 0.844

Individual class accuracy for class-2: 0.100

Individual class accuracy for class-3: 0.8238

Individual class accuracy for class-4: 0.9047

Individual class accuracy for class-5: 0.7083

Overall accuracy: 0.843

Answer to 2 (e):

MLE:

Maximum likelihood classifier (MLC) does not assume that the features are independent and as a result it performs better than the Naïve Bayes classifier.

The disadvantages of MLC are:

- computationally intensive
- the result is dependent on the model of evaluation

The Naïve Bayes classifier assumes the features are independent and uses the Bayesian theorem to classify labels or classes. The Naïve Bayes classifier provides options to add a-priori information. The advantages and disadvantages of Naïve Bayes theorem is:

Naïve Bayes Classifier:

Advantages:

- Fast to train and classify

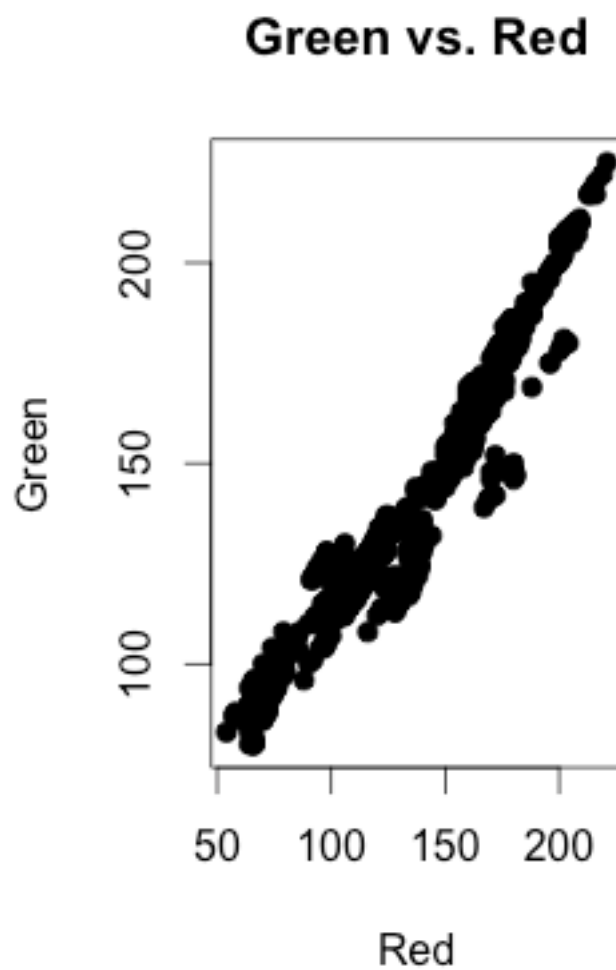
– Conceptually easy to understand

Disadvantages:

- Assumes independence of features
- Does not allow for complex hypotheses

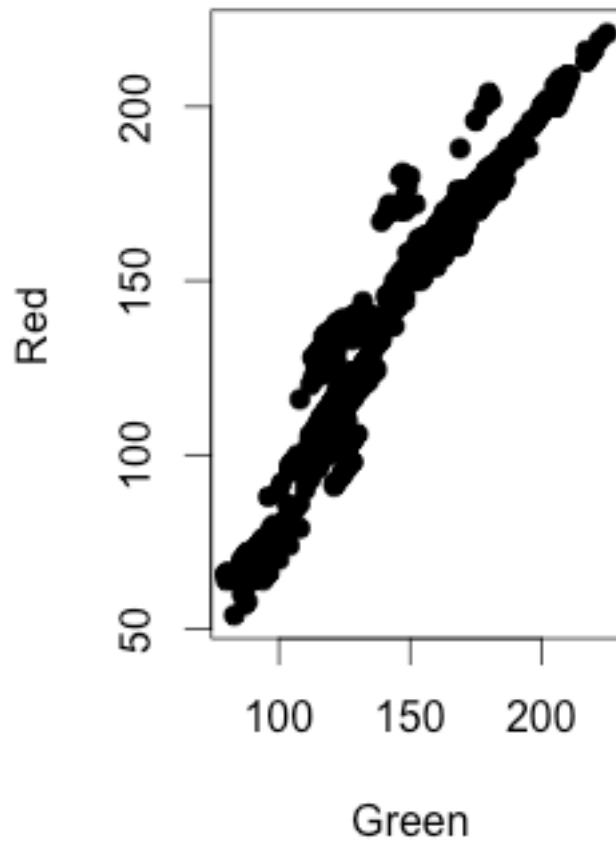
Answer to 2 (f)

Plot of Green vs. Red



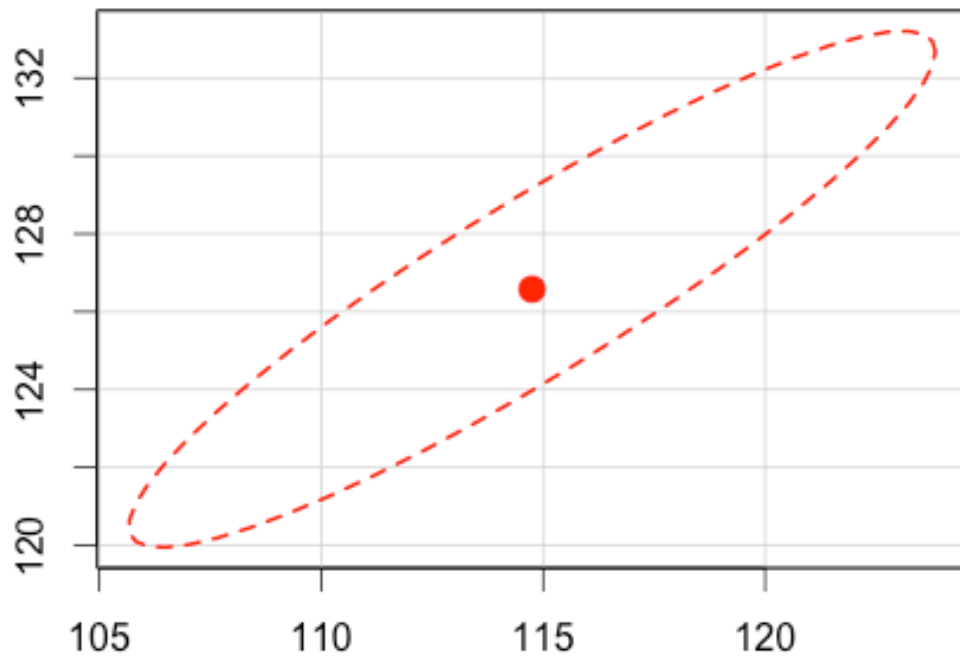
Plot of Red vs Green

Red vs. Green

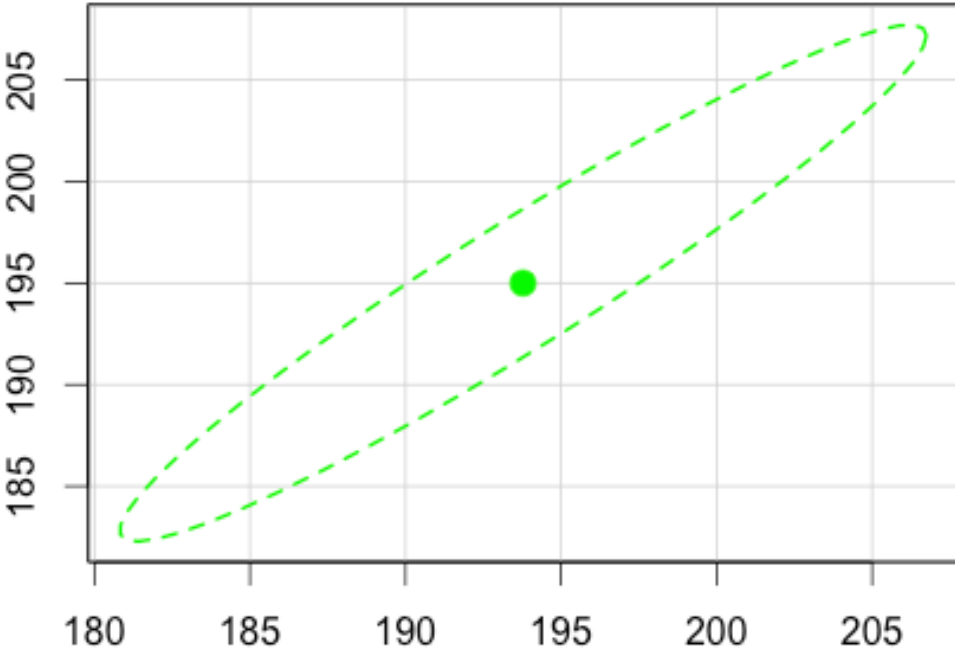


The co-variance matrix plots:

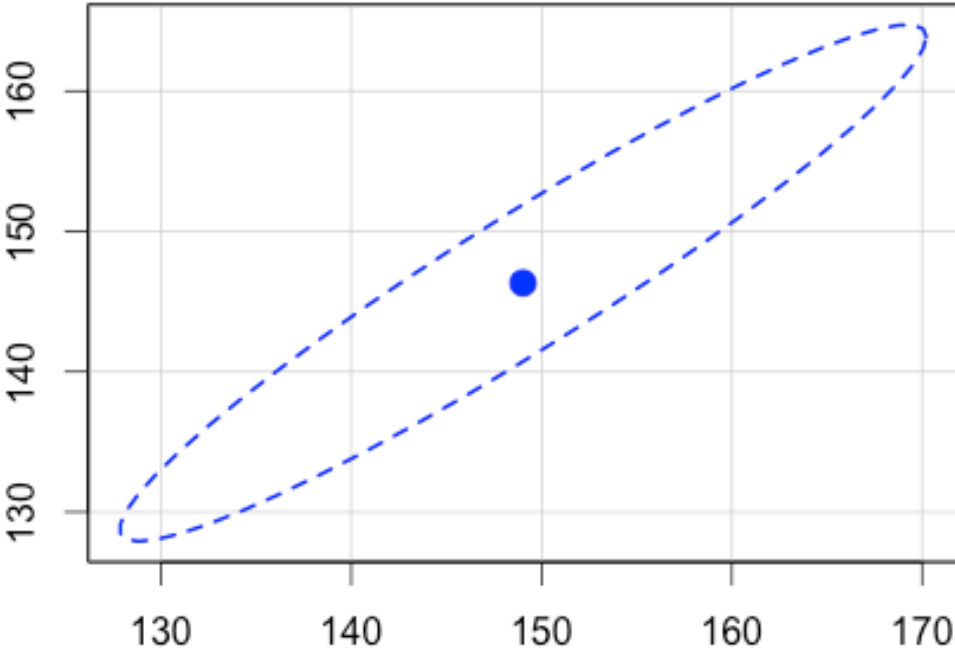
Class-1



Class-2

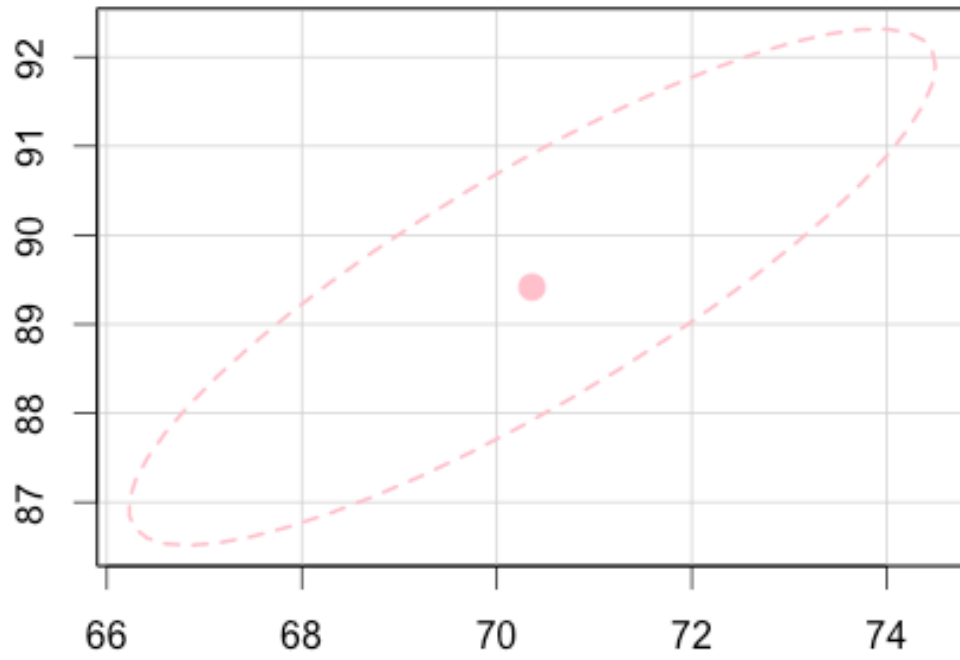


Class-3

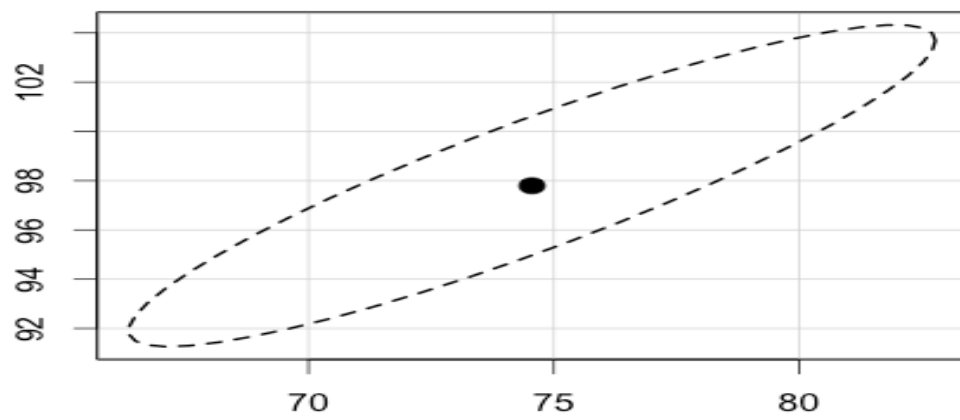




Class-4



Class-5



Comments:

Class 3 had the most variability as the values along the X-axis, Y-axis, as well as the points for the covariance ellipsoid was wide. The next Class with the most variability is Class 2. The least variable class is Class 2.