**ISLR 2.4 Prob#7**

Table

Description automatically generated

* Euclidean Distance, d(p,q) =

Obs.1 Euclidean Distance = = 3

Obs.2 Euclidean Distance = = 2

Obs.3 Euclidean Distance = = 3.162

Obs.4 Euclidean Distance = = 2.236

Obs.5 Euclidean Distance = = 1.414

Obs.6 Euclidean Distance = = 1.732



* For k =1 , it will be green because from Observation#5, closest to k=1 is 1.414 which is green. So the prediction will be gree.



* For k=3, the prediction will be Red because when k=3 the closest Euclidean Distance is 3.162 which is Red.

Text

Description automatically generated with medium confidence

* For the case of highly non linear decision boundary, smaller k will be best value for covering the closest points as much as possible. If the decision boundary is linear then larger k will cover most of the decision points.

**ISLR 4.7 Prob#5**

Graphical user interface, text, Word

Description automatically generated with medium confidence

* QDA will perform better in training set because of its better flexibility and it will fit the training dataset with better accuracy compare to LDA.

On the test set, LDA should perform better than QDA because QDA will overfit the test set.

A screenshot of a computer

Description automatically generated with low confidence

* For the case of non-linear decision boundary, QDA will perform better than LDA in training and test set because of greater flexibility of QDA.

Text

Description automatically generated

* With the increment of n, a more flexible method will yield a better fit. As a result, I expect that test prediction accuracy of QDA relative to LDA will improve.

Text

Description automatically generated

* False
* With fewer sample point, QDA as a flexible model will yield to overfit the test set and performance will be worser than LDA.