

## R output for Exercise 4, Question 2

```

> data(iris)
> irstype = as.vector(iris[, 5])
> IrisVe = as.matrix(iris[irstype == "versicolor", -5])
> IrisVi = as.matrix(iris[irstype == "virginica", -5])
> mVe = apply(IrisVe, 2, mean)
> mVi = apply(IrisVi, 2, mean)
> mVe

Sepal.Length Sepal.Width Petal.Length Petal.Width
           5.936           2.770           4.260           1.326

> mVi

Sepal.Length Sepal.Width Petal.Length Petal.Width
           6.588           2.974           5.552           2.026

> vVe = var(IrisVe)
> vVi = var(IrisVi)
> sp = (vVe + vVi)/2
> d = mVe - mVi
> d

Sepal.Length Sepal.Width Petal.Length Petal.Width
          -0.652          -0.204          -1.292          -0.700

> a = solve(sp) %*% d
> a

              [,1]
Sepal.Length  3.556303
Sepal.Width   5.578621
Petal.Length -6.970128
Petal.Width -12.386041

> nVe = nrow(IrisVe)
> nVi = nrow(IrisVi)
> nVe

[1] 50

> nVi

[1] 50

```

```
> D = ((nVi * nVe)/(nVi + nVe)) * t(a) %>% d
> D
```

```
      [,1]
[1,] 355.4721
```

```
> Da = 95 * D/(98 * 4)
> Da
```

```
      [,1]
[1,] 86.14759
```

```
> 1 - pf(Da, 4, 95)
```

```
      [,1]
[1,] 0
```

```
> rule = as.vector(t(a) %>% (mVe + mVi)/2)
> rule
```

```
[1] -16.66309
```

```
> t(a) %>% mVe
```

```
      [,1]
[1,] -9.553643
```

```
> t(a) %>% mVi
```

```
      [,1]
[1,] -23.77253
```

```
> vepred = as.vector(IrisVe %>% a)
```

```
> vipred = as.vector(IrisVi %>% a)
```

```
> table(vepred > rule)
```

```
FALSE TRUE
      2   48
```

```
> table(vipred <= rule)
```

```
FALSE TRUE
      1   49
```

```
> x = c(7, 4.5, 5.1, 2)
```

```
> x %>% a
```

```
      [,1]
[1,] -10.32182
```

## R output for Week 4, Question 3

```

> library(class)
> library(MASS)
> data(crabs)
> dim(crabs)

[1] 200  8

> colnames(crabs)

[1] "sp"    "sex"   "index" "FL"    "RW"    "CL"    "CW"    "BD"

> crabs.knn = knn(train = crabs[, 4:7], test = crabs[, 4:7], cl = crabs[, "sex"])
> table(crabs.knn, crabs[, "sex"])

crabs.knn  F  M
          F 100  0
          M  0 100

> index = sample(1:200, 50)
> crabs.knn1=knn(train=crabs[-index,4:7],test=crabs[index,4:7],cl=crabs[-index, "sex"])
> table(crabs.knn1, crabs[index, "sex"])

crabs.knn1  F  M
           F 24  1
           M  2 23

> newx = c(19.3, 14.5, 40, 30.9)
> knn(train = crabs[-index, 4:7], test = newx, cl = crabs[-index, "sex"])

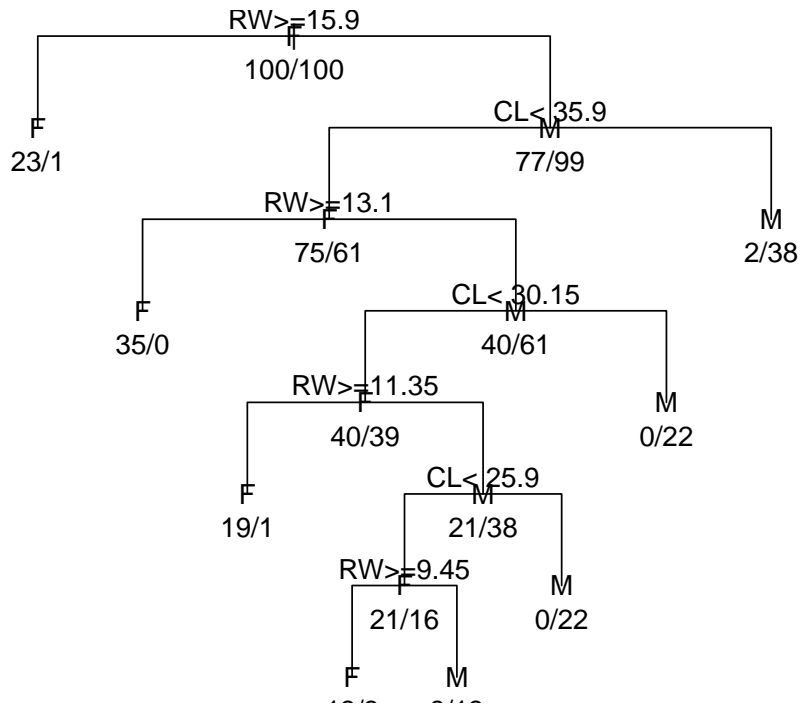
[1] M
Levels: F M

> knn(train = crabs[, 4:7], test = newx, cl = crabs[, "sex"])

[1] M
Levels: F M

> library(rpart)
> crabs.tree = rpart(sex ~ FL + RW + CL + CW, crabs)
> plot(crabs.tree, uniform = TRUE)
> text(crabs.tree, all = TRUE, use.n = TRUE)

```



```

> set.seed(123)
> cvKNN = c()
> cvTree = c()
> neworder = sample(1:200, 200)
> for (i in 0:4) {
+   index = c((i * 40 + 1):((i + 1) * 40))
+   TS = crabs[neworder[index], ]
+   LS = crabs[-neworder[index], ]
+   cv.knn = knn(train = LS[, 4:7], test = TS[, 4:7], cl = LS["sex"])
+   cv.tree = rpart(sex ~ FL + RW + CL + CW, LS)
+   err.knn = table(cv.knn, TS[, "sex"])
+   pred.tree = predict(cv.tree, TS[, 4:7], type = "class")
+   err.tree = table(pred.tree, TS[, "sex"])
+   cvKNN = c(cvKNN, 40 - sum(diag(err.knn)))
+   cvTree = c(cvTree, 40 - sum(diag(err.tree)))
+ }
> cvKNN

[1] 2 3 5 4 4

> cvTree

[1] 8 10 7 10 11

```