Zadanie 2.1

```
# 2.1

library("klaR")
library("MASS")

w <- read.table("http://www.ipipan.eu/~teisseyrep/TEACHING/DM/DANE/wine.data",
    sep = ",")
head(w, 2)
```

```
##   V1    V2   V3   V4   V5  V6   V7   V8   V9  V10  V11  V12  V13  V14
## 1  1 14.23 1.71 2.43 15.6 127 2.80 3.06 0.28 2.29 5.64 1.04 3.92 1065
## 2  1 13.20 1.78 2.14 11.2 100 2.65 2.76 0.26 1.28 4.38 1.05 3.40 1050
```

```
# a)

m1 <- lda(V1 ~ V2 + V8, data = w)
z1 <- predict(m1, newdata = w)$posterior
head(z1, 2)  # pstwo aposteriori
```

```
##        1         2         3
## 1 0.9996 0.0003965 2.206e-05
## 2 0.8124 0.1866492 9.655e-04
```

```
# b)

klasy <- predict(m1, newdata = w)$class
t <- table(w$V1, klasy)
t
```

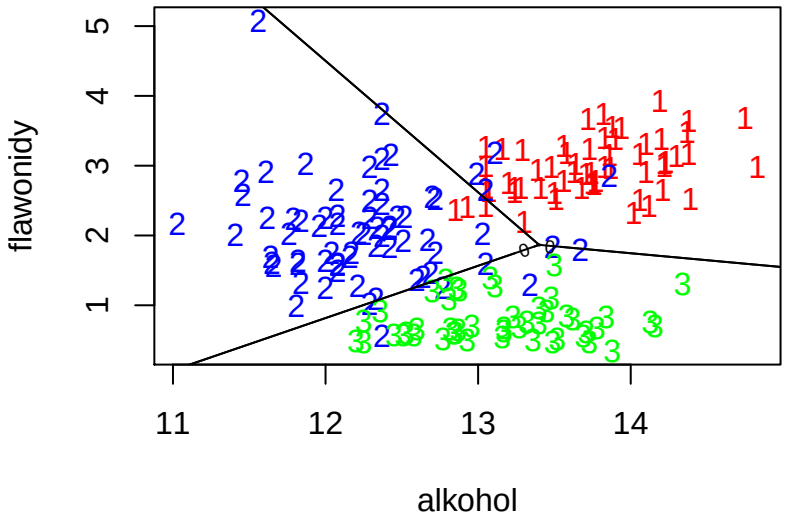
```
##    klasy
##      1  2  3
##   1 56  3  0
##   2  4 60  7
##   3  0  0 48
```

```
100 * sum(diag(t))/nrow(w)  # procent poprawnych klasyfikacji
```

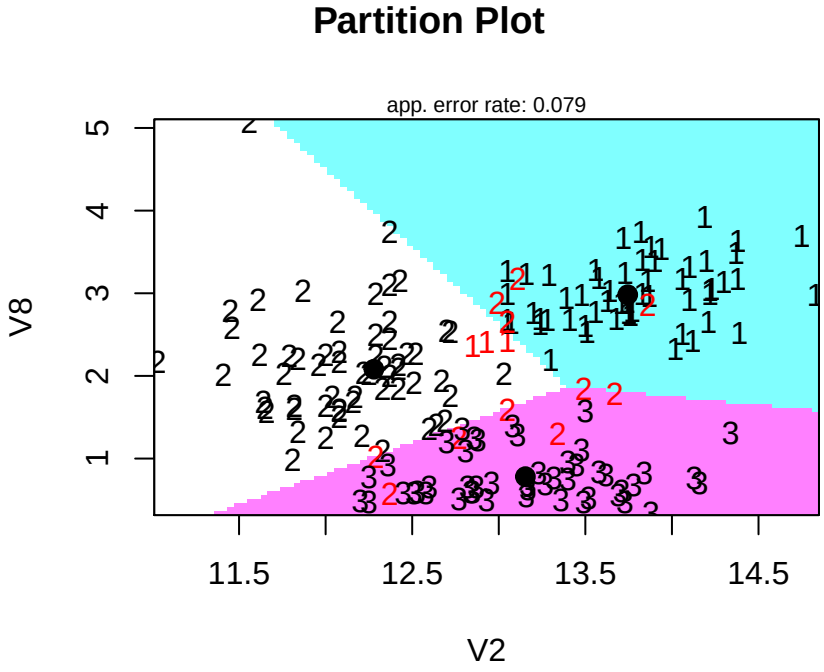
```
## [1] 92.13
```

```
# c)

n1 <- length(which(w$V1 == 1))
n2 <- length(which(w$V1 == 2))
n3 <- length(which(w$V1 == 3))

plot(w$V2, w$V8, xlab = "alkohol", ylab = "flawonidy", pch = c(rep("1", n1),
    rep("2", n2), rep("3", n3)), col = c(rep("red", n1), rep("blue", n2), rep("green",
    n3)))

x <- seq(10.5, 15.5, length.out = 200)
y <- seq(0, 5.5, length.out = 200)
siatka <- expand.grid(V2 = x, V8 = y)
pred <- predict(m1, newdata = siatka)$posterior

z1 <- pred[, 1] - pmax(pred[, 2], pred[, 3])
```

```
z2 <- pred[, 2] - pmax(pred[, 1], pred[, 3])
z3 <- pred[, 3] - pmax(pred[, 2], pred[, 1])

contour(x, y, matrix(z1, 200), level = 0, add = T)
contour(x, y, matrix(z2, 200), level = 0, add = T)
contour(x, y, matrix(z3, 200), level = 0, add = T)
```

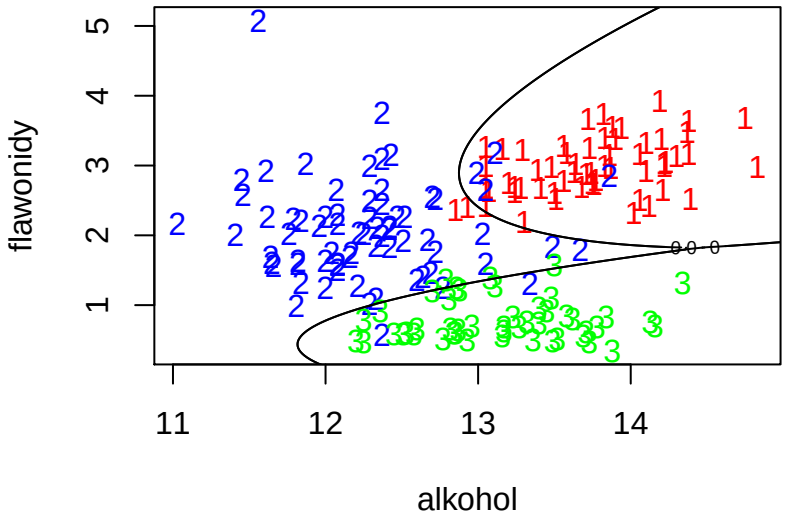



```
partimat(as.factor(V1) ~ V8 + V2, data = w)  # albo inaczej
```




```
# d)

m2 <- qda(V1 ~ V2 + V8, data = w)
z1 <- predict(m2, newdata = w)$posterior
head(z1, 2)

##         1         2         3
```

```
## 1 0.9996 0.0004125 4.547e-14
## 2 0.8735 0.1265091 1.904e-10

klasy <- predict(m2, newdata = w)$class
t <- table(w$V1, klasy)
t

##    klasy
##      1  2  3
##   1 57  2  0
##   2  4 65  2
##   3  0  3 45

100 * sum(diag(t))/nrow(w)  # wiekszy niz w lda

## [1] 93.82

n1 <- length(which(w$V1 == 1))
n2 <- length(which(w$V1 == 2))
n3 <- length(which(w$V1 == 3))

plot(w$V2, w$V8, xlab = "alkohol", ylab = "flawonidy", pch = c(rep("1", n1),
    rep("2", n2), rep("3", n3)), col = c(rep("red", n1), rep("blue", n2), rep("green",
    n3)))

x <- seq(10.5, 15.5, length.out = 200)
y <- seq(0, 5.5, length.out = 200)
siatka <- expand.grid(V2 = x, V8 = y)
pred <- predict(m2, newdata = siatka)$posterior

z1 <- pred[, 1] - pmax(pred[, 2], pred[, 3])
z2 <- pred[, 2] - pmax(pred[, 1], pred[, 3])
z3 <- pred[, 3] - pmax(pred[, 2], pred[, 1])

contour(x, y, matrix(z1, 200), level = 0, add = T)
contour(x, y, matrix(z2, 200), level = 0, add = T)
contour(x, y, matrix(z3, 200), level = 0, add = T)
```



Zadanie 2.2

```
# 2.2

m <- read.table("http://www.ipipan.eu/~teisseyrep/TEACHING/DM/DANE/kredit.asc",
    header = TRUE)
head(m, 2)
```

```
##   kredit laufkont laufzeit moral verw hoehe sparkont beszeit rate famges
## 1      1        1       18     4    2  1049        1       2    4      2
## 2      1        1        9     4    0  2799        1       3    2      3
##   buerge wohnzeit verm alter weitkred wohn bishkred beruf pers telef
## 1      1        4    2    21        3    1        1     3    1     1
## 2      1        2    1    36        3    1        2     3    2     1
##   gastarb
## 1       1
## 2       1
```

```
# a)

m_lda <- lda(kredit ~ ., data = m)
m_qda <- qda(kredit ~ ., data = m)

t_lda <- table(m$kredit, predict(m_lda, newdata = m)$class)
t_lda
```

```
##
##       0   1
##   0 149 151
##   1  79 621
```

```
100 * sum(diag(t_lda))/nrow(m)
```

```
## [1] 77
```

```
t_qda <- table(m$kredit, predict(m_qda, newdata = m)$class)
t_qda
```

```
##
##       0   1
##   0 200 100
##   1 121 579
```

```
100 * sum(diag(t_qda))/nrow(m)
```

```
## [1] 77.9
```

```
# b)

s_lda <- stepclass(m[, 2:ncol(m)], m[, 1], method = "lda", direction = "backward")
```

```
s_lda$formula
```

```
## m[, 1] ~ laufkont + laufzeit + moral + verw + sparkont + beszeit +
##     rate + famges + buerge + wohnzeit + verm + alter + weitkred +
##     wohn + bishkred + telef + gastarb
## <environment: 0x000000000da72d30>
```

```
m_lda2 <- lda(s_lda$formula, data = m)
t_lda2 <- table(m$kredit, predict(m_lda2, newdata = m)$class)
t_lda2
```

```
##
##       0   1
##   0 150 150
##   1  74 626
```

```
100 * sum(diag(t_lda2))/nrow(m)
```

```
## [1] 77.6
```

```
s_qda <- stepclass(m[, 2:ncol(m)], m[, 1], method = "qda", direction = "backward")
```

```
s_qda$formula
```

```
## m[, 1] ~ laufkont + laufzeit + moral + verw + hoehe + sparkont +
##     famges + buerge + verm + alter + weitkred + wohn + bishkred +
##     beruf
## <environment: 0x000000000d3697e0>
```

```
m_qda2 <- qda(s_qda$formula, data = m)
t_qda2 <- table(m$kredit, predict(m_qda2, newdata = m)$class)
t_qda2
```

```
##
##       0   1
##   0 170 130
##   1  83 617
```

```
100 * sum(diag(t_qda2))/nrow(m)
```

```
## [1] 78.7
```

Zadanie 2.3

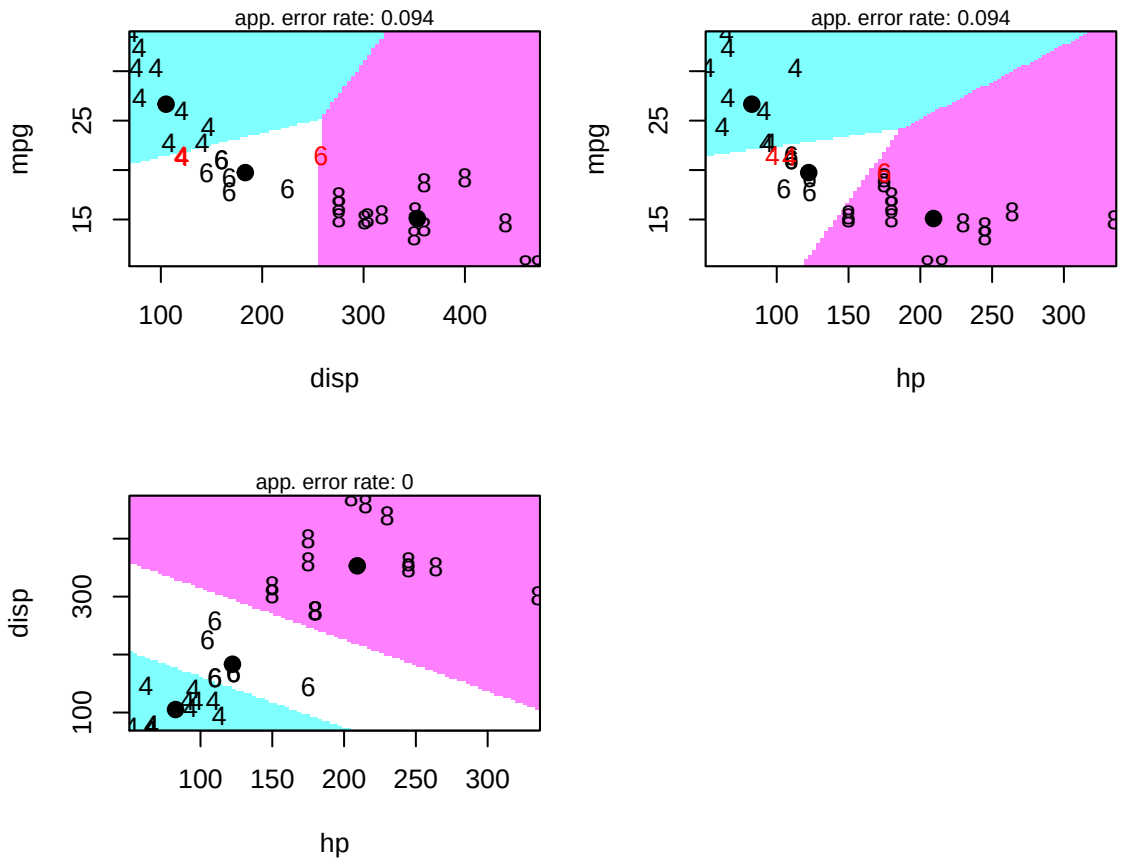
```
# 2.3

data(mtcars)
head(mtcars, 3)
```

```
##                mpg cyl disp  hp drat    wt  qsec vs am gear carb
## Mazda RX4     21.0   6  160 110 3.90 2.620 16.46  0  1    4    4
## Mazda RX4 Wag 21.0   6  160 110 3.90 2.875 17.02  0  1    4    4
## Datsun 710    22.8   4  108  93 3.85 2.320 18.61  1  1    4    1
```

```
c <- mtcars[, c(2, 1, 3, 4)]
head(c, 3)
```

```
##               cyl  mpg disp  hp
## Mazda RX4       6 21.0  160 110
## Mazda RX4 Wag   6 21.0  160 110
## Datsun 710      4 22.8  108  93
```

```
partimat(as.factor(c$cyl) ~ ., data = c, method = "lda", nplots.vert = 2)
```




```
partimat(as.factor(c$cyl) ~ ., data = c, method = "qda", nplots.vert = 2)
```