Assignment3

Saumitra Mazumder

June 16, 2019

Question 1

```
## 'data.frame':    4898 obs. of  12 variables:
##  $ fixed.acidity       : num  7 6.3 8.1 7.2 7.2 8.1 6.2 7 6.3 8.1 ...
##  $ volatile.acidity    : num  0.27 0.3 0.28 0.23 0.23 0.28 0.32 0.27 0.3 0.22 ...
##  $ citric.acid         : num  0.36 0.34 0.4 0.32 0.32 0.4 0.16 0.36 0.34 0.43 ...
##  $ residual.sugar      : num  20.7 1.6 6.9 8.5 8.5 6.9 7 20.7 1.6 1.5 ...
##  $ chlorides           : num  0.045 0.049 0.05 0.058 0.058 0.05 0.045 0.045 0.049 0.044 ...
##  $ free.sulfur.dioxide : num  45 14 30 47 47 30 30 45 14 28 ...
##  $ total.sulfur.dioxide: num  170 132 97 186 186 97 136 170 132 129 ...
##  $ density             : num  1.001 0.994 0.995 0.996 0.996 ...
##  $ pH                  : num  3 3.3 3.26 3.19 3.19 3.26 3.18 3 3.3 3.22 ...
##  $ sulphates           : num  0.45 0.49 0.44 0.4 0.4 0.44 0.47 0.45 0.49 0.45 ...
##  $ alcohol             : num  8.8 9.5 10.1 9.9 9.9 10.1 9.6 8.8 9.5 11 ...
##  $ quality             : int  6 6 6 6 6 6 6 6 6 6 ...
```

```
##   fixed.acidity volatile.acidity citric.acid residual.sugar chlorides
## 1           7.0             0.27        0.36           20.7     0.045
## 2           6.3             0.30        0.34            1.6     0.049
## 3           8.1             0.28        0.40            6.9     0.050
## 4           7.2             0.23        0.32            8.5     0.058
## 5           7.2             0.23        0.32            8.5     0.058
## 6           8.1             0.28        0.40            6.9     0.050
##   free.sulfur.dioxide total.sulfur.dioxide density   pH sulphates alcohol
## 1                  45                  170  1.0010 3.00      0.45     8.8
## 2                  14                  132  0.9940 3.30      0.49     9.5
## 3                  30                   97  0.9951 3.26      0.44    10.1
## 4                  47                  186  0.9956 3.19      0.40     9.9
## 5                  47                  186  0.9956 3.19      0.40     9.9
## 6                  30                   97  0.9951 3.26      0.44    10.1
##   quality
## 1       6
## 2       6
## 3       6
## 4       6
## 5       6
## 6       6
```

```
## integer(0)
```

We see that there are no missing data

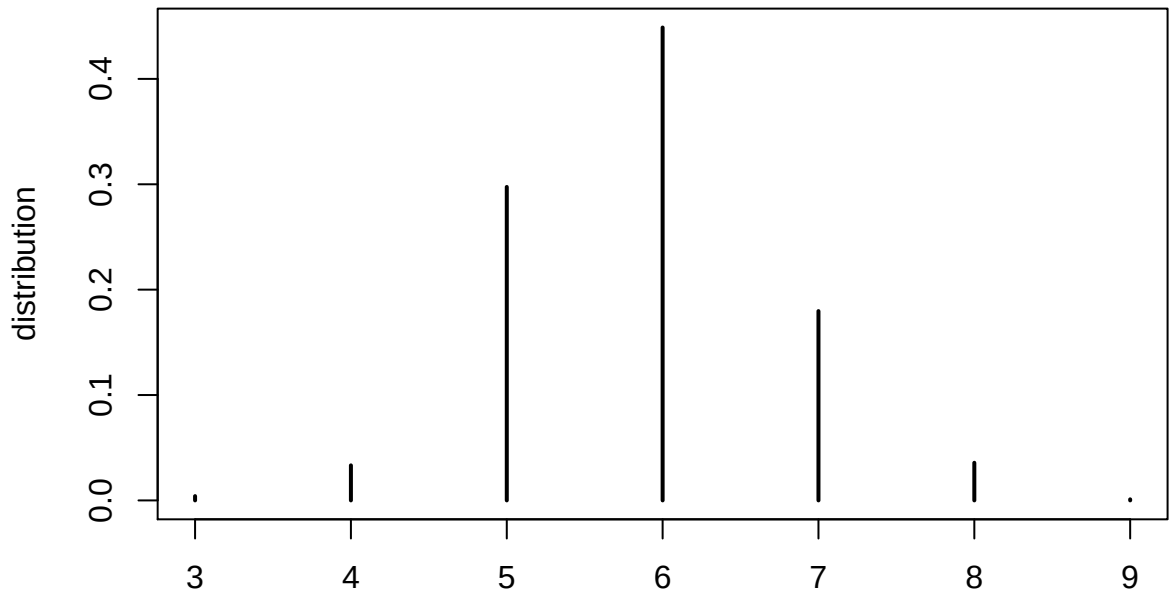
Question 2

```
## 'data.frame':    4898 obs. of  11 variables:
##  $ fixed.acidity       : num  7 6.3 8.1 7.2 7.2 8.1 6.2 7 6.3 8.1 ...
##  $ volatile.acidity    : num  0.27 0.3 0.28 0.23 0.23 0.28 0.32 0.27 0.3 0.22 ...
##  $ citric.acid         : num  0.36 0.34 0.4 0.32 0.32 0.4 0.16 0.36 0.34 0.43 ...
##  $ residual.sugar      : num  20.7 1.6 6.9 8.5 8.5 6.9 7 20.7 1.6 1.5 ...
##  $ chlorides           : num  0.045 0.049 0.05 0.058 0.058 0.05 0.045 0.045 0.049 0.044 ...
```

```
##  $ free.sulfur.dioxide : num  45 14 30 47 47 30 30 45 14 28 ...
##  $ total.sulfur.dioxide: num  170 132 97 186 186 97 136 170 132 129 ...
##  $ density             : num  1.001 0.994 0.995 0.996 0.996 ...
##  $ pH                  : num  3 3.3 3.26 3.19 3.19 3.26 3.18 3 3.3 3.22 ...
##  $ sulphates           : num  0.45 0.49 0.44 0.4 0.4 0.44 0.47 0.45 0.49 0.45 ...
##  $ alcohol             : num  8.8 9.5 10.1 9.9 9.9 10.1 9.6 8.8 9.5 11 ...
```

```
##                       fixed.acidity volatile.acidity citric.acid
## fixed.acidity          1.00000000      -0.02269729  0.28918070
## volatile.acidity      -0.02269729       1.00000000 -0.14947181
## citric.acid            0.28918070      -0.14947181  1.00000000
## residual.sugar         0.08902070       0.06428606  0.09421162
## chlorides              0.02308564       0.07051157  0.11436445
## free.sulfur.dioxide   -0.04939586      -0.09701194  0.09407722
## total.sulfur.dioxide   0.09106976       0.08926050  0.12113080
## density                0.26533101       0.02711385  0.14950257
## pH                    -0.42585829      -0.03191537 -0.16374821
## sulphates             -0.01714299      -0.03572815  0.06233094
## alcohol               -0.12088112       0.06771794 -0.07572873
##                       residual.sugar   chlorides free.sulfur.dioxide
## fixed.acidity            0.08902070  0.02308564       -0.0493958591
## volatile.acidity         0.06428606  0.07051157       -0.0970119393
## citric.acid              0.09421162  0.11436445        0.0940772210
## residual.sugar           1.00000000  0.08868454        0.2990983537
## chlorides                0.08868454  1.00000000        0.1013923521
## free.sulfur.dioxide      0.29909835  0.10139235        1.0000000000
## total.sulfur.dioxide     0.40143931  0.19891030        0.6155009650
## density                  0.83896645  0.25721132        0.2942104109
## pH                      -0.19413345 -0.09043946       -0.0006177961
## sulphates               -0.02666437  0.01676288        0.0592172458
## alcohol                 -0.45063122 -0.36018871       -0.2501039415
##                       total.sulfur.dioxide     density            pH
## fixed.acidity                 0.091069756  0.26533101 -0.4258582910
## volatile.acidity              0.089260504  0.02711385 -0.0319153683
## citric.acid                   0.121130798  0.14950257 -0.1637482114
## residual.sugar                0.401439311  0.83896645 -0.1941334540
## chlorides                     0.198910300  0.25721132 -0.0904394560
## free.sulfur.dioxide           0.615500965  0.29421041 -0.0006177961
## total.sulfur.dioxide          1.000000000  0.52988132  0.0023209718
## density                       0.529881324  1.00000000 -0.0935914935
## pH                            0.002320972 -0.09359149  1.0000000000
## sulphates                     0.134562367  0.07449315  0.1559514973
## alcohol                      -0.448892102 -0.78013762  0.1214320987
##                         sulphates     alcohol
## fixed.acidity        -0.01714299 -0.12088112
## volatile.acidity     -0.03572815  0.06771794
## citric.acid           0.06233094 -0.07572873
## residual.sugar       -0.02666437 -0.45063122
## chlorides             0.01676288 -0.36018871
## free.sulfur.dioxide   0.05921725 -0.25010394
## total.sulfur.dioxide  0.13456237 -0.44889210
## density               0.07449315 -0.78013762
## pH                    0.15595150  0.12143210
## sulphates             1.00000000 -0.01743277
```

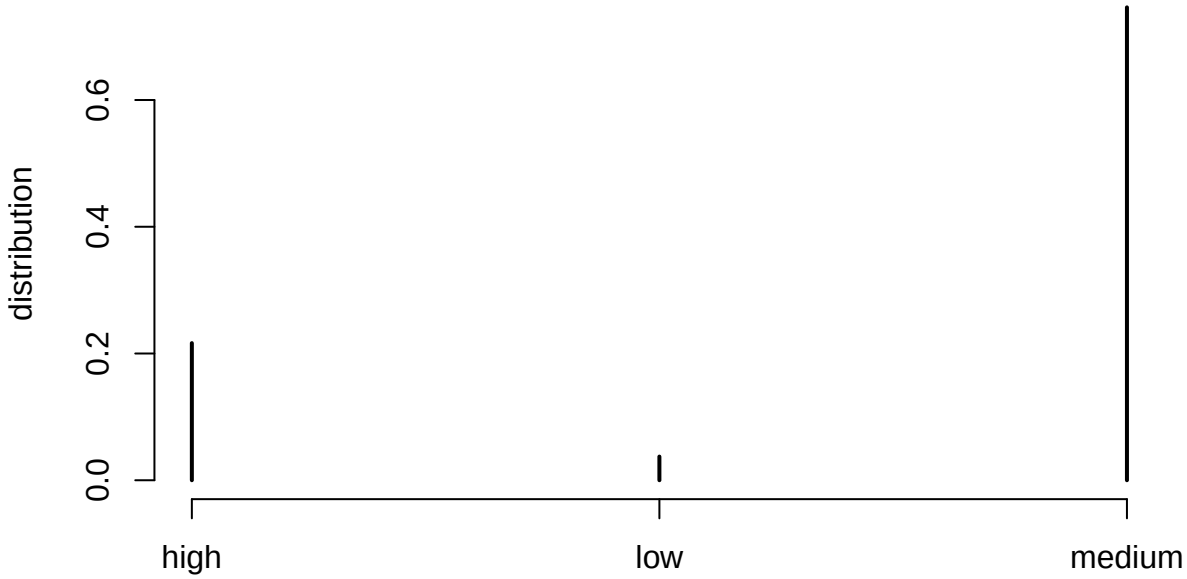
```
## alcohol              -0.01743277  1.00000000
```



Question 3

Question 4

```
##   fixed.acidity volatile.acidity citric.acid residual.sugar chlorides
## 1           7.0             0.27        0.36           20.7     0.045
## 2           6.3             0.30        0.34            1.6     0.049
## 3           8.1             0.28        0.40            6.9     0.050
## 4           7.2             0.23        0.32            8.5     0.058
## 5           7.2             0.23        0.32            8.5     0.058
## 6           8.1             0.28        0.40            6.9     0.050
##   free.sulfur.dioxide total.sulfur.dioxide density   pH sulphates alcohol
## 1                  45                  170  1.0010 3.00      0.45     8.8
## 2                  14                  132  0.9940 3.30      0.49     9.5
## 3                  30                   97  0.9951 3.26      0.44    10.1
## 4                  47                  186  0.9956 3.19      0.40     9.9
## 5                  47                  186  0.9956 3.19      0.40     9.9
## 6                  30                   97  0.9951 3.26      0.44    10.1
##   quality rating
## 1       6 medium
## 2       6 medium
## 3       6 medium
## 4       6 medium
## 5       6 medium
## 6       6 medium
```



Question 5

```
##  fixed.acidity    volatile.acidity  citric.acid     residual.sugar
##  Min.   :0.0000   Min.   :0.0000   Min.   :0.0000   Min.   :0.00000
##  1st Qu.:0.2404   1st Qu.:0.1275   1st Qu.:0.1627   1st Qu.:0.01687
##  Median :0.2885   Median :0.1765   Median :0.1928   Median :0.07055
##  Mean   :0.2937   Mean   :0.1944   Mean   :0.2013   Mean   :0.08883
##  3rd Qu.:0.3365   3rd Qu.:0.2353   3rd Qu.:0.2349   3rd Qu.:0.14264
##  Max.   :1.0000   Max.   :1.0000   Max.   :1.0000   Max.   :1.00000
##    chlorides       free.sulfur.dioxide total.sulfur.dioxide
##  Min.   :0.00000   Min.   :0.00000     Min.   :0.0000
##  1st Qu.:0.08012   1st Qu.:0.07317     1st Qu.:0.2297
##  Median :0.10089   Median :0.11150     Median :0.2900
##  Mean   :0.10912   Mean   :0.11606     Mean   :0.3001
##  3rd Qu.:0.12166   3rd Qu.:0.15331     3rd Qu.:0.3666
##  Max.   :1.00000   Max.   :1.00000     Max.   :1.0000
##     density             pH           sulphates         alcohol
##  Min.   :0.00000   Min.   :0.0000   Min.   :0.0000   Min.   :0.0000
##  1st Qu.:0.08892   1st Qu.:0.3364   1st Qu.:0.2209   1st Qu.:0.2419
##  Median :0.12782   Median :0.4182   Median :0.2907   Median :0.3871
##  Mean   :0.13336   Mean   :0.4257   Mean   :0.3138   Mean   :0.4055
##  3rd Qu.:0.17332   3rd Qu.:0.5091   3rd Qu.:0.3837   3rd Qu.:0.5484
##  Max.   :1.00000   Max.   :1.0000   Max.   :1.0000   Max.   :1.0000
```

Question 6

Question 7

```
## Warning: package 'class' was built under R version 3.5.3
```

Question 8

```
## Warning: package 'gmodels' was built under R version 3.5.3
```

```
##
##
##    Cell Contents
## |-------------------------|
## |                       N |
## |           N / Row Total |
## |           N / Col Total |
## |         N / Table Total |
## |-------------------------|
##
##
## Total Observations in Table:  3419
##
##
##                     | wineTestPredictor
## wineTestingQuality |      high |       low |    medium | Row Total |
## -------------------|-----------|-----------|-----------|-----------|
##               high |       482 |         0 |       274 |       756 |
##                    |     0.638 |     0.000 |     0.362 |     0.221 |
##                    |     0.707 |     0.000 |     0.101 |           |
##                    |     0.141 |     0.000 |     0.080 |           |
## -------------------|-----------|-----------|-----------|-----------|
##                low |         7 |         8 |       129 |       144 |
##                    |     0.049 |     0.056 |     0.896 |     0.042 |
##                    |     0.010 |     0.667 |     0.047 |           |
##                    |     0.002 |     0.002 |     0.038 |           |
## -------------------|-----------|-----------|-----------|-----------|
##             medium |       193 |         4 |      2322 |      2519 |
##                    |     0.077 |     0.002 |     0.922 |     0.737 |
##                    |     0.283 |     0.333 |     0.852 |           |
##                    |     0.056 |     0.001 |     0.679 |           |
## -------------------|-----------|-----------|-----------|-----------|
##       Column Total |       682 |        12 |      2725 |      3419 |
##                    |     0.199 |     0.004 |     0.797 |           |
## -------------------|-----------|-----------|-----------|-----------|
##
##
```

Test data consists of 3419 observations 483 cases predicted accurately as high quality 10 cases predicted accurately as low quality 2315 cases predicted accurately as medium quality total accuracy is 483 +10 + 2315/3419 = 0.82