MSIA400_LAB2_ZIWEN_WANG

Ziwen Wang

11/4/2017

Loading Data

```
dt<- read.table("redwine.txt",header = T)
head(dt)
```

```
##   QA   FA   VA   CA  RS    CH FS SD     DE   PH   SU  AL
## 1  5  7.4 0.70 0.00 1.9 0.076 11 34 0.9978 3.51 0.56 9.4
## 2  5  7.8 0.88 0.00 2.6 0.098 25 67 0.9968 3.20 0.68 9.8
## 3  5  7.8 0.76 0.04 2.3 0.092 15 54 0.9970 3.26 0.65 9.8
## 4  6 11.2 0.28 0.56 1.9 0.075 17 60 0.9980 3.16 0.58 9.8
## 5  5  7.4 0.70 0.00 1.9 0.076 11 34 0.9978 3.51 0.56 9.4
## 6  5  7.4 0.66 0.00 1.8 0.075 13 40 0.9978 3.51 0.56 9.4
```

```
#Q1
m_RS<-mean(dt$RS,na.rm=T)
m_RS
```

```
## [1] 2.537952
```

```
M_SD<-mean(dt$SD,na.rm=T)
M_SD
```

```
## [1] 46.29836
```

```
#Q2
dt2<-dt[!is.na(dt$SD),c('SD','FS')]
md1<-lm(SD~FS,data = dt2)
md1$coefficients
```

```
## (Intercept)          FS
##   13.185505    2.086077
```

```
#Q3
dt3<-dt[is.na(dt$SD),c('FS')]
predict.list <- predict(md1, newdata = data.frame(FS = dt3), se.fit = FALSE)
new_sd<-predict.list
new_fs<-dt3
dt4<-dt2
whole_sd<-c(dt4$SD,new_sd)
whole_fs<-c(dt4$FS,new_fs)
avg_sd<-mean(whole_sd)
avg_sd
```

```
## [1] 46.30182
```

```
#Q4
avg.imp<- function (a, avg)
  { missing<- is.na(a)
  n.missing <-sum(missing)
  a.obs <-a[!missing]
```

```
  imputed <- a
  imputed[missing]<-avg
  return (imputed)
}

AVG_RS<-mean(dt$RS,na.rm=T)
head(AVG_RS)
```

```
## [1] 2.537952
```

```
mean(dt[!is.na(dt$RS),'RS'])
```

```
## [1] 2.537952
```

```
RS_IMP<-avg.imp(dt$RS,AVG_RS)
head(RS_IMP)
```

```
## [1] 1.9 2.6 2.3 1.9 1.9 1.8
```

```
new_AVG_RS<-mean(RS_IMP)
new_AVG_RS
```

```
## [1] 2.537952
```

```
#Q5
#We have imputed all missing values in the data set.
#Build multiple linear regression model for the new data set and save it as winemodel.
#Print out the coefficients of the regression model.

imp_dt<-dt
imp_dt$RS<- RS_IMP
imp_dt$SD[is.na(imp_dt$SD) == T] <- predict.list
winemodel<-lm(QA~.,data =imp_dt )
summary(winemodel)
```

```
##
## Call:
## lm(formula = QA ~ ., data = imp_dt)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.78010 -0.36249 -0.06331  0.44595  1.98828
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  4.720e+01  1.782e+01   2.649 0.008151 **
## FA           6.841e-02  1.872e-02   3.654 0.000267 ***
## VA          -1.098e+00  1.213e-01  -9.053  < 2e-16 ***
## CA          -1.789e-01  1.474e-01  -1.214 0.224954
## RS           2.593e-02  1.419e-02   1.827 0.067944 .
## CH          -1.631e+00  4.097e-01  -3.982 7.14e-05 ***
## FS           3.530e-03  2.159e-03   1.635 0.102262
## SD          -2.855e-03  7.248e-04  -3.939 8.54e-05 ***
## DE          -4.482e+01  1.789e+01  -2.505 0.012329 *
## PH           3.600e-02  4.409e-02   0.816 0.414413
## SU           9.449e-01  1.136e-01   8.321  < 2e-16 ***
## AL           2.470e-01  2.265e-02  10.906  < 2e-16 ***
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.6491 on 1587 degrees of freedom
## Multiple R-squared:  0.3584, Adjusted R-squared:  0.354
## F-statistic:  80.6 on 11 and 1587 DF,  p-value: < 2.2e-16
```

```
winemodel$coefficients
```

```
##   (Intercept)            FA            VA            CA            RS
##  47.202815335   0.068406796  -1.097686420  -0.178949797   0.025926958
##            CH            FS            SD            DE            PH
##  -1.631290466   0.003530106  -0.002854970 -44.816652166   0.035996993
##            SU            AL
##   0.944871182   0.247046550
```

```
#Q6
#Print out the summary of the model.
#Pick one attribute that is least likely to be related to QA based on p-values.
summary(winemodel)
```

```
##
## Call:
## lm(formula = QA ~ ., data = imp_dt)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.78010 -0.36249 -0.06331  0.44595  1.98828
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  4.720e+01  1.782e+01   2.649 0.008151 **
## FA           6.841e-02  1.872e-02   3.654 0.000267 ***
## VA          -1.098e+00  1.213e-01  -9.053  < 2e-16 ***
## CA          -1.789e-01  1.474e-01  -1.214 0.224954
## RS           2.593e-02  1.419e-02   1.827 0.067944 .
## CH          -1.631e+00  4.097e-01  -3.982 7.14e-05 ***
## FS           3.530e-03  2.159e-03   1.635 0.102262
## SD          -2.855e-03  7.248e-04  -3.939 8.54e-05 ***
## DE          -4.482e+01  1.789e+01  -2.505 0.012329 *
## PH           3.600e-02  4.409e-02   0.816 0.414413
## SU           9.449e-01  1.136e-01   8.321  < 2e-16 ***
## AL           2.470e-01  2.265e-02  10.906  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.6491 on 1587 degrees of freedom
## Multiple R-squared:  0.3584, Adjusted R-squared:  0.354
## F-statistic:  80.6 on 11 and 1587 DF,  p-value: < 2.2e-16
```

```
# We choose PH since its p value is 0.414413 >0.05, we have confidence to believe PH is not significant
```

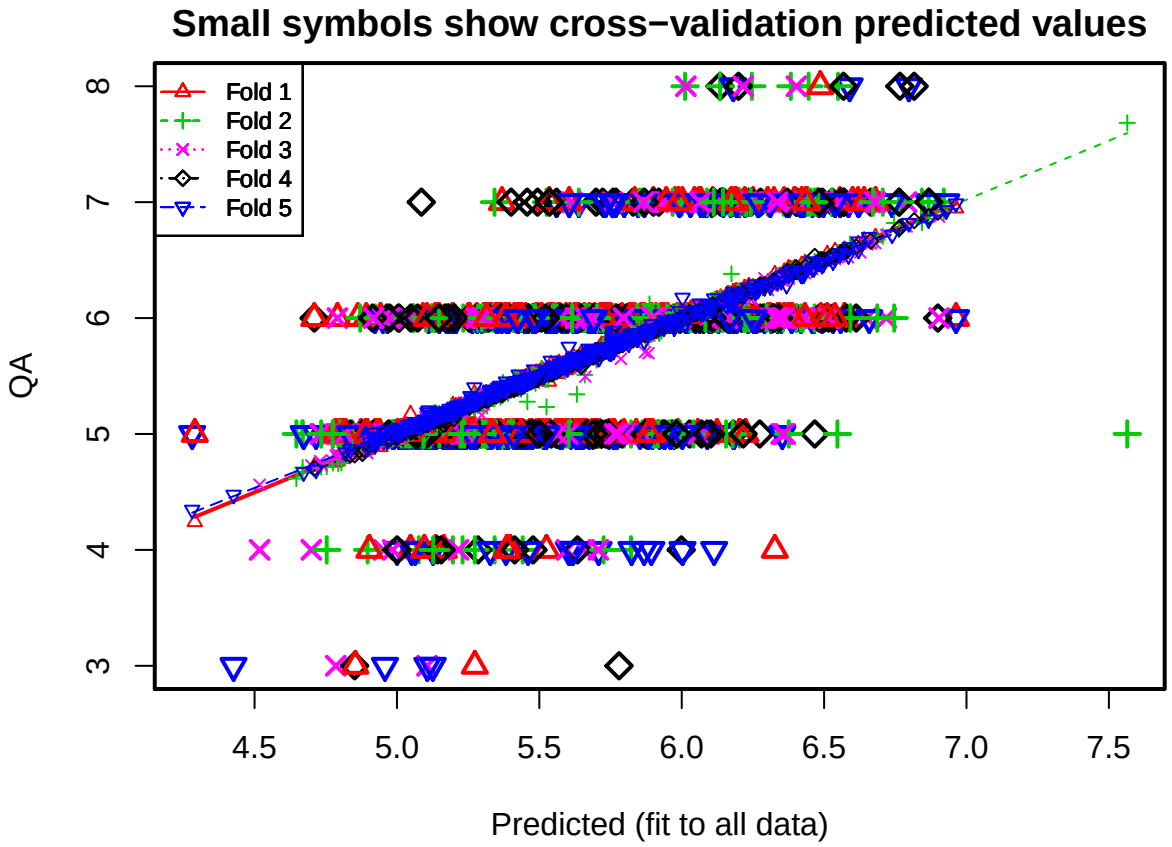
```
#Q7
#Perform 5-fold cross validation for the model you just built. Print out the average error rate.
library(DAAG)
```

```
## Loading required package: lattice
```

```r
validation = CVlm(data=imp_dt, m=5,  form.lm=formula(winemodel))
```

```
## Analysis of Variance Table
##
## Response: QA
##              Df Sum Sq Mean Sq F value  Pr(>F)
## FA            1     16    16.0   38.07 8.7e-10 ***
## VA            1    144   143.6  340.77 < 2e-16 ***
## CA            1      0     0.0    0.06   0.810
## RS            1      0     0.1    0.18   0.673
## CH            1     13    13.1   30.99 3.0e-08 ***
## FS            1      3     2.9    6.84   0.009 **
## SD            1     30    29.5   70.02 < 2e-16 ***
## DE            1     60    60.1  142.62 < 2e-16 ***
## PH            1      2     1.9    4.55   0.033 *
## SU            1     56    56.3  133.57 < 2e-16 ***
## AL            1     50    50.1  118.95 < 2e-16 ***
## Residuals  1587    669     0.4
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
## Warning in CVlm(data = imp_dt, m = 5, form.lm = formula(winemodel)):
##
##  As there is >1 explanatory variable, cross-validation
##  predicted values for a fold are not a linear function
##  of corresponding overall predicted values.  Lines that
##  are shown for the different folds are approximate
```



```
##
## fold 1
## Observations in test set: 319
##                   1       2      3     4      7    8     12      15     23
## Predicted    5.0673  5.0805  5.174 5.733  5.108 5.37  5.666  5.1113  5.688
## cvpred       5.0644  5.0605  5.158 5.767  5.114 5.39  5.676  5.0716  5.698
## QA           5.0000  5.0000  5.000 6.000  5.000 7.00  5.000  5.0000  5.000
## CV residual -0.0644 -0.0605 -0.158 0.233 -0.114 1.61 -0.676 -0.0716 -0.698
##                  25     31     40     45    55     58     61    65    66
## Predicted   5.536  5.148  5.889  5.197 5.552  5.243  5.415  5.39  5.39
## cvpred      5.561  5.132  5.924  5.221 5.562  5.252  5.454  5.40  5.40
## QA          6.000  5.000  5.000  5.000 6.000  5.000  5.000  5.00  5.00
## CV residual 0.439 -0.132 -0.924 -0.221 0.438 -0.252 -0.454 -0.40 -0.40
##                  69    70     72     75     81     84    95   101    113
## Predicted    6.06 5.835 4.9234  5.711  5.290 5.0590  5.05 5.428  5.255
## cvpred       6.09 5.832 4.9197  5.738  5.329 4.9765  5.17 5.435  5.227
## QA           5.00 6.000 5.0000  5.000  5.000 5.0000  4.00 6.000  5.000
## CV residual -1.09 0.168 0.0803 -0.738 -0.329 0.0235 -1.17 0.565 -0.227
##                  115   118  120   122     139    143     148  149   151
## Predicted    5.558 5.197 4.89 5.675  5.0939  6.963  5.0710 5.57 5.719
## cvpred       5.563 5.212 4.89 5.645  5.0889  6.947  5.0191 5.59 5.767
## QA           5.000 6.000 6.00 6.000  5.0000  6.000  5.0000 6.00 6.000
## CV residual -0.563 0.788 1.11 0.355 -0.0889 -0.947 -0.0191 0.41 0.233
##                  158    161     165   168   169    170     179    186    187
## Predicted    5.594 4.9302  5.0336  5.14 5.394  5.534  5.0679  5.431  5.193
## cvpred       5.608 4.9272  5.0707  5.17 5.394  5.452  5.0749  5.461  5.222
## QA           5.000 5.0000  5.0000  4.00 6.000  5.000  5.0000  5.000  5.000
## CV residual -0.608 0.0728 -0.0707 -1.17 0.606 -0.452 -0.0749 -0.461 -0.222
##                  189   192    198  199    203    214    216   224   238  240
## Predicted   5.0034 5.488  6.221 6.04  5.534  5.351  5.582 5.325 5.091 4.71
## cvpred      4.9784 5.508  6.258 5.99  5.556  5.378  5.599 5.319 5.096 4.72
## QA          5.0000 6.000  6.000 7.00  5.000  5.000  5.000 6.000 6.000 6.00
## CV residual 0.0216 0.492 -0.258 1.01 -0.556 -0.378 -0.599 0.681 0.904 1.28
##                249   254    257    264   269  289    290    296    298   301
## Predicted   5.33 4.873  5.569  5.556 5.531 5.90  5.616  5.437  5.290 5.658
## cvpred      5.34 4.872  5.612  5.595 5.541 5.89  5.625  5.446  5.284 5.655
## QA          6.00 5.000  5.000  5.000 6.000 7.00  5.000  5.000  5.000 6.000
## CV residual 0.66 0.128 -0.612 -0.595 0.459 1.11 -0.625 -0.446 -0.284 0.345
##                305   310  315   320  326    334  335   350    368    373
## Predicted   4.797 5.509  5.9 5.366 5.54  5.656 5.98 5.304  5.260  6.382
## cvpred      4.802 5.553  5.9 5.343 5.60  5.636 5.95 5.287  5.295  6.421
## QA          5.000 6.000  5.0 6.000 6.00  5.000 7.00 6.000  5.000  6.000
## CV residual 0.198 0.447 -0.9 0.657 0.40 -0.636 1.05 0.713 -0.295 -0.421
##                374     375    377   386   387   398     401    412    420
## Predicted    5.132  6.1005  6.235 5.369 5.313 5.751  5.0174  5.355  5.243
## cvpred       5.135  6.0922  6.129 5.362 5.332 5.747  5.0275  5.366  5.242
## QA           5.000  6.0000  6.000 6.000 6.000 6.000  5.0000  5.000  5.000
## CV residual -0.135 -0.0922 -0.129 0.638 0.668 0.253 -0.0275 -0.366 -0.242
##                421    429    436   445   449    451    457    458   469
## Predicted   6.060  5.324  5.473 6.324 5.502 5.9093  5.414  5.206 5.489
## cvpred      6.044  5.345  5.502 6.305 5.503 5.9256  5.411  5.205 5.527
## QA          7.000  5.000  5.000 7.000 6.000 6.0000  5.000  5.000 6.000
## CV residual 0.956 -0.345 -0.502 0.695 0.497 0.0744 -0.411 -0.205 0.473
##                474    478    484   489   494    495   500   501   508   515
```

```
## Predicted    6.12  6.571  5.887 6.251 5.637  6.180 5.637 5.195 5.357 6.067
## cvpred       6.16  6.587  5.909 6.262 5.578  6.221 5.578 5.214 5.374 6.065
## QA           5.00  6.000  5.000 7.000 6.000  6.000 6.000 6.000 6.000 7.000
## CV residual -1.16 -0.587 -0.909 0.738 0.422 -0.221 0.422 0.786 0.626 0.935
##                522    532   539    558  567   569     572  575   577   579
## Predicted    5.278  5.745 6.398  5.987 4.79 5.594  6.0209 5.74  5.53  5.25
## cvpred       5.317  5.742 6.389  5.956 4.76 5.551  6.0293 5.75  5.54  5.25
## QA           5.000  5.000 7.000  5.000 6.00 6.000  6.0000 6.00  4.00  5.00
## CV residual -0.317 -0.742 0.611 -0.956 1.24 0.449 -0.0293 0.25 -1.54 -0.25
##                586   588     595   604   609    612   633    636   649
## Predicted   5.347 4.807  5.0475 5.412 5.280  5.596 5.732  5.220 6.002
## cvpred      5.355 4.787  5.0567 5.443 5.237  5.625 5.746  5.239 6.016
## QA          6.000 5.000  5.0000 6.000 6.000  5.000 6.000  5.000 7.000
## CV residual 0.645 0.213 -0.0567 0.557 0.763 -0.625 0.254 -0.239 0.984
##                656   660   661    662    671   673   677    678    680
## Predicted    5.262  5.40 5.549  5.273  5.603 4.290 5.622  5.069  5.797
## cvpred       5.248  5.38 5.549  5.297  5.638 4.241 5.657  5.062  5.861
## QA           5.000  4.00 6.000  5.000  5.000 5.000 6.000  5.000  5.000
## CV residual -0.248 -1.38 0.451 -0.297 -0.638 0.759 0.343 -0.062 -0.861
##                689    696   705    708    714   727    729   730    733
## Predicted    5.207  6.374  5.16  5.573  5.217 5.566  5.363 5.720 4.9392
## cvpred       5.237  6.382  5.15  5.588  5.242 5.518  5.374 5.687 4.9412
## QA           5.000  6.000  4.00  5.000  5.000 6.000  5.000 6.000 5.0000
## CV residual -0.237 -0.382 -1.15 -0.588 -0.242 0.482 -0.374 0.313 0.0588
##                736    745     758  777    779    786   790   793    809
## Predicted   4.786  5.259  5.0167 4.94  5.701  5.410 4.890 5.146  5.385
## cvpred      4.784  5.254  5.0079 4.89  5.711  5.429 4.862 5.126  5.381
## QA          5.000  5.000  5.0000 6.00  5.000  5.000 5.000 6.000  5.000
## CV residual 0.216 -0.254 -0.0079 1.11 -0.711 -0.429 0.138 0.874 -0.381
##                819     820    821    823    825    846    847    851
## Predicted    5.0028  5.0561  5.230  5.383  5.537  5.273  5.273  5.467
## cvpred       5.0148  5.0296  5.231  5.395  5.555  5.286  5.286  5.509
## QA           5.0000  5.0000  5.000  5.000  5.000  5.000  5.000  5.000
## CV residual -0.0148 -0.0296 -0.231 -0.395 -0.555 -0.286 -0.286 -0.509
##                854    872    886   890   899    907   908    910    913
## Predicted    6.12  5.631  5.415 4.783 6.600  5.659 5.826  6.160  6.313
## cvpred       6.15  5.646  5.394 4.754 6.609  5.639 5.824  6.199  6.322
## QA           6.00  5.000  5.000 5.000 7.000  5.000 6.000  6.000  6.000
## CV residual -0.15 -0.646 -0.394 0.246 0.391 -0.639 0.176 -0.199 -0.322
##                914    915   929    934    935    937   945   962   967
## Predicted   6.359  6.160  6.20  5.363  5.572  6.336 6.396  5.30 6.176
## cvpred      6.377  6.199  6.23  5.359  5.557  6.346 6.426  5.33 6.175
## QA          7.000  6.000  5.00  5.000  5.000  6.000 7.000  5.00 7.000
## CV residual 0.623 -0.199 -1.23 -0.359 -0.557 -0.346 0.574 -0.33 0.825
##                968    969   970   975    976  979   981   986   987    988
## Predicted   4.821  6.360  5.49 6.421  5.284 5.95 5.815 5.857 6.221  5.314
## cvpred      4.801  6.386  5.50 6.457  5.318 5.99 5.825 5.852 6.252  5.347
## QA          5.000  6.000  5.00 7.000  5.000 7.00 6.000 6.000 7.000  5.000
## CV residual 0.199 -0.386 -0.50 0.543 -0.318 1.01 0.175 0.148 0.748 -0.347
##                990    991   993   997  1002  1015  1017   1022   1024 1025
## Predicted    6.050  5.441 5.251 6.109 6.187 5.790 6.555  6.211  6.303 5.60
## cvpred       6.084  5.478 5.182 6.082 6.221 5.788 6.555  6.228  6.307 5.61
## QA           6.000  5.000 6.000 7.000 7.000 6.000 7.000  6.000  6.000 7.00
## CV residual -0.084 -0.478 0.818 0.918 0.779 0.212 0.445 -0.228 -0.307 1.39
```

```
##             1031    1043    1046   1048   1049   1050 1053  1054   1058
## Predicted   5.84  6.0692  6.0375  5.696 5.9837 5.9374  6.2 6.681  5.135
## cvpred      5.83  6.0662  6.0444  5.709 5.9873 5.9437  6.2 6.696  5.138
## QA          7.00  6.0000  6.0000  5.000 6.0000 6.0000  5.0 7.000  5.000
## CV residual 1.17 -0.0662 -0.0444 -0.709 0.0127 0.0563 -1.2 0.304 -0.138
##             1062    1063  1073  1074   1075   1077  1083   1088   1098
## Predicted   6.49  6.0423 5.422 5.507 4.8960  6.361 5.414  6.375  5.263
## cvpred      6.50  6.0547 5.449 5.509 4.9231  6.398 5.445  6.427  5.291
## QA          8.00  6.0000 6.000 6.000 5.0000  6.000 6.000  6.000  5.000
## CV residual 1.50 -0.0547 0.551 0.491 0.0769 -0.398 0.555 -0.427 -0.291
##              1099   1100  1103   1105   1113  1114 1122  1123   1124
## Predicted   6.441  5.263 5.825  6.361  6.137 5.717  6.2  6.21  6.147
## cvpred      6.419  5.291 5.837  6.377  6.174 5.759  6.2  6.23  6.159
## QA          7.000  5.000 6.000  6.000  6.000 6.000  6.0  6.00  6.000
## CV residual 0.581 -0.291 0.163 -0.377 -0.174 0.241 -0.2 -0.23 -0.159
##               1139  1140   1143  1144   1145  1147    1152   1154   1156
## Predicted    5.188 5.256  6.166  6.03  5.758 5.607  6.0925  6.041  5.246
## cvpred       5.194 5.252  6.197  6.08  5.796 5.598  6.0832  6.055  5.251
## QA           5.000 6.000  6.000  6.00  5.000 6.000  6.0000  6.000  5.000
## CV residual -0.194 0.748 -0.197 -0.08 -0.796 0.402 -0.0832 -0.055 -0.251
##               1164   1166  1178   1180   1186  1193  1197  1198   1208
## Predicted    5.349  5.421 6.302  6.127 5.9498 6.619 5.148 5.452  5.523
## cvpred       5.357  5.384 6.293  6.138 5.9545 6.635 5.149 5.489  5.532
## QA           5.000  5.000 7.000  6.000 6.0000 7.000 6.000 6.000  5.000
## CV residual -0.357 -0.384 0.707 -0.138 0.0455 0.365 0.851 0.511 -0.532
##             1210   1215   1216   1219   1220   1221   1224   1230   1232
## Predicted   6.19 5.8696  6.164 5.8943  6.236  6.259  6.489  5.393  5.428
## cvpred      6.23 5.9156  6.202 5.9114  6.275  6.254  6.499  5.414  5.417
## QA          7.00 6.0000  6.000 6.0000  6.000  6.000  6.000  5.000  5.000
## CV residual 0.77 0.0844 -0.202 0.0886 -0.275 -0.254 -0.499 -0.414 -0.417
##              1237   1238   1242   1243   1246  1259   1263   1273   1274
## Predicted   5.419 5.9969  5.644  6.345  5.640 5.695  5.564  5.840  5.196
## cvpred      5.434 5.9897  5.677  6.363  5.645 5.674  5.598  5.852  5.247
## QA          6.000 6.0000  5.000  6.000  5.000 6.000  5.000  5.000  5.000
## CV residual 0.566 0.0103 -0.677 -0.363 -0.645 0.326 -0.598 -0.852 -0.247
##              1277   1287  1288 1294  1313  1316  1319  1320   1322    1334
## Predicted    6.33  6.538  6.22  5.1 4.815 5.135 5.135 5.105 5.9277  5.0276
## cvpred       6.38  6.577  6.22  5.1 4.804 5.129 5.129 5.052 5.9228  5.0252
## QA           4.00  6.000  5.00  4.0 5.000 6.000 6.000 6.000 6.0000  5.0000
## CV residual -2.38 -0.577 -1.22 -1.1 0.196 0.871 0.871 0.948 0.0772 -0.0252
##              1340  1342   1348   1350  1352   1356   1357   1364    1367
## Predicted   5.632 5.632  5.085  5.628 5.863  5.512  5.501  4.903  5.0566
## cvpred      5.636 5.636  5.104  5.636 5.854  5.541  5.531  4.898  5.0487
## QA          6.000 6.000  5.000  5.000 6.000  5.000  5.000  4.000  5.0000
## CV residual 0.364 0.364 -0.104 -0.636 0.146 -0.541 -0.531 -0.898 -0.0487
##              1370  1381   1387   1394   1413   1414   1417  1418    1420
## Predicted    5.40 5.724  5.237  5.407  6.433  5.684  5.763 6.435  5.0233
## cvpred       5.42 5.724  5.243  5.419  6.471  5.666  5.795 6.447  5.0147
## QA           4.00 6.000  5.000  5.000  6.000  5.000  5.000 7.000  5.0000
## CV residual -1.42 0.276 -0.243 -0.419 -0.471 -0.666 -0.795 0.553 -0.0147
##              1425   1427  1440 1442  1451   1454  1460  1461   1466 1467
## Predicted   5.744  6.389 5.623 4.86 6.174  5.116 6.640 5.714  5.436 5.54
## cvpred      5.775  6.379 5.597 4.86 6.198  5.145 6.667 5.719  5.465 5.56
## QA          6.000  6.000 6.000 6.00 7.000  5.000 7.000 6.000  5.000 7.00
```

```
## CV residual 0.225 -0.379 0.403 1.14 0.802 -0.145 0.333 0.281 -0.465 1.44
##              1470   1477   1482  1485  1491   1494  1499   1501  1506
## Predicted    4.85  5.272  5.873  5.39  6.51  5.074 5.315  5.207  5.27
## cvpred       4.84  5.349  5.899  5.38  6.55  5.047 5.308  5.227  5.28
## QA           3.00  5.000  5.000  4.00  6.00  5.000 6.000  5.000  3.00
## CV residual -1.84 -0.349 -0.899 -1.38 -0.55 -0.047 0.692 -0.227 -2.28
##              1511  1525  1528 1530 1535  1536   1539  1543  1546  1559
## Predicted   5.727 5.718 5.820 5.44 5.99 5.384  5.898 5.381 5.438 4.890
## cvpred      5.768 5.748 5.827 5.46 6.00 5.409  5.883 5.385 5.472 4.902
## QA          6.000 6.000 6.000 6.00 7.00 6.000  5.000 6.000 6.000 5.000
## CV residual 0.232 0.252 0.173 0.54 1.00 0.591 -0.883 0.615 0.528 0.098
##               1563   1565   1582   1584
## Predicted    5.333  5.333  5.884  5.336
## cvpred       5.349  5.349  5.906  5.352
## QA           5.000  5.000  5.000  5.000
## CV residual -0.349 -0.349 -0.906 -0.352
##
## Sum of squares = 122    Mean square = 0.38    n = 319
##
## fold 2
## Observations in test set: 320
##                9     10     16    19    22      29    34     35    36
## Predicted   5.34  5.666  5.190  5.08  5.44  5.0531 5.204  5.235 5.257
## cvpred      5.33  5.632  5.237  5.04  5.44  5.0317 5.149  5.201 5.218
## QA          7.00  5.000  5.000  4.00  5.00  5.0000 6.000  5.000 6.000
## CV residual 1.67 -0.632 -0.237 -1.04 -0.44 -0.0317 0.851 -0.201 0.782
##                37   43    47     49    52     59   63      64    71    76
## Predicted   5.600 5.54 4.647  5.371 5.427  5.345 5.53  5.0947 5.233  5.65
## cvpred      5.579 5.53 4.615  5.334 5.392  5.339 5.23  5.0775 5.227  5.62
## QA          6.000 6.00 5.000  5.000 6.000  5.000 7.00  5.0000 6.000  5.00
## CV residual 0.421 0.47 0.385 -0.334 0.608 -0.339 1.77 -0.0775 0.773 -0.62
##                78    80      83     88       99   102     108    123   124
## Predicted   5.288  5.44  5.0928  5.471  5.05256 5.611  5.1415 5.0180  5.11
## cvpred      5.261  5.45  5.0678  5.447  5.00682 5.598  5.0852 4.9799  5.10
## QA          6.000  4.00  5.0000  5.000  5.00000 6.000  5.0000 5.0000  5.00
## CV residual 0.739 -1.45 -0.0678 -0.447 -0.00682 0.402 -0.0852 0.0201 -0.10
##                126   128   131  135    140    153    155    156     164
## Predicted    5.133 4.805 4.804 4.92  5.089  5.283  5.606  5.594  5.0311
## cvpred       5.186 4.751 4.801 4.90  5.109  5.316  5.588  5.576  5.0137
## QA           5.000 5.000 5.000 6.00  5.000  5.000  5.000  5.000  5.0000
## CV residual -0.186 0.249 0.199 1.10 -0.109 -0.316 -0.588 -0.576 -0.0137
##                167    175    180    183    191   209   213   215    219
## Predicted    5.151  5.363  5.320 4.9894  5.129  5.31 5.873 5.254  5.384
## cvpred       5.161  5.357  5.329 4.9678  5.149  5.32 5.885 5.239  5.375
## QA           5.000  5.000  5.000 5.0000  5.000  5.00 6.000 6.000  5.000
## CV residual -0.161 -0.357 -0.329 0.0322 -0.149 -0.32 0.115 0.761 -0.375
##                225   229    230  232   237   239  250   252      259    261
## Predicted    5.34 5.703  5.563 5.55 5.058 5.058 5.59 5.456  5.04290  5.575
## cvpred       5.34 5.718  5.543 5.58 5.045 5.045 5.54 5.441  5.00276  5.584
## QA           4.00 6.000  5.000 6.00 6.000 6.000 6.00 6.000  5.00000  5.000
## CV residual -1.34 0.282 -0.543 0.42 0.955 0.955 0.46 0.559 -0.00276 -0.584
##                262    263   266    272    273   277    283   294    297
## Predicted    4.753  5.542 6.197  6.314  5.861 5.531  5.211 5.642  5.089
## cvpred       4.719  5.554 6.234  6.324  5.879 5.468  5.198 5.612  5.121
```

```
## QA           4.000  5.000 7.000  6.000  5.000 6.000  5.000 6.000  5.000
## CV residual -0.719 -0.554 0.766 -0.324 -0.879 0.532 -0.198 0.388 -0.121
##                300    302    303   308   312   317    323  327    330
## Predicted    5.214  6.109  5.342 5.336 5.119  5.42  5.261 6.11  5.610
## cvpred       5.164  6.152  5.317 5.356 5.137  5.47  5.242 6.12  5.629
## QA           5.000  6.000  5.000 6.000 6.000  5.00  5.000 7.00  5.000
## CV residual -0.164 -0.152 -0.317 0.644 0.863 -0.47 -0.242 0.88 -0.629
##                331    348   349  359     361    363   364    369   370
## Predicted    6.175  6.592 5.761 5.97  5.0656  5.454  6.04  5.327 6.867
## cvpred       6.381  6.639 5.767 5.98  5.0659  5.479  6.07  5.364 6.857
## QA           6.000  6.000 6.000 7.00  5.0000  5.000  5.00  5.000 7.000
## CV residual -0.381 -0.639 0.233 1.02 -0.0659 -0.479 -1.07 -0.364 0.143
##                372   376   388  391    409   414    415    416     423
## Predicted   5.748 6.223 5.190 6.01  6.110 6.358  5.142  5.208  5.1113
## cvpred      5.718 6.225 5.193 5.98  6.121 6.328  5.174  5.183  5.0953
## QA          6.000 7.000 6.000 8.00  6.000 7.000  5.000  5.000  5.0000
## CV residual 0.282 0.775 0.807 2.02 -0.121 0.672 -0.174 -0.183 -0.0953
##                431    434   439    442  446   447   454  456   464   465
## Predicted   6.436  5.473 5.850  6.135 5.27  5.97 6.105 6.38 4.668 5.722
## cvpred      6.463  5.512 5.862  6.179 5.29  6.05 6.112 6.41 4.666 5.755
## QA          7.000  5.000 6.000  6.000 6.00  5.00 7.000 8.00 5.000 6.000
## CV residual 0.537 -0.512 0.138 -0.179 0.71 -1.05 0.888 1.59 0.334 0.245
##                470  477  482   490    493    498  499   529    531    534
## Predicted    5.284  5.8 6.55 5.853 6.9204  5.956 6.13 5.317  6.146  6.746
## cvpred       5.317  5.8 6.54 5.847 6.9191  5.991 6.16 5.298  6.148  6.821
## QA           5.000  5.0 8.00 6.000 7.0000  5.000 8.00 6.000  6.000  6.000
## CV residual -0.317 -0.8 1.46 0.153 0.0809 -0.991 1.84 0.702 -0.148 -0.821
##                536    541    543    548   556    561   564   573    582
## Predicted    6.109  5.358  5.363  6.232  5.89  5.917 5.658  5.99  5.374
## cvpred       6.116  5.364  5.362  6.243  6.12  5.966 5.657  6.02  5.417
## QA           6.000  5.000  5.000  6.000  5.00  5.000 6.000  5.00  5.000
## CV residual -0.116 -0.364 -0.362 -0.243 -1.12 -0.966 0.343 -1.02 -0.417
##                584   598  606   608    611    624    626     627   634
## Predicted   6.000 5.626 5.17 5.561  5.464  6.356  5.284  5.0355  5.20
## cvpred      6.061 5.648 5.16 5.574  5.473  6.342  5.247  5.0175  5.21
## QA          7.000 6.000 6.00 6.000  5.000  6.000  5.000  5.0000  4.00
## CV residual 0.939 0.352 0.84 0.426 -0.473 -0.342 -0.247 -0.0175 -1.21
##                637    643   653  658   659    664   667   670     674   675
## Predicted   4.669  5.458  7.56 5.86 5.549  6.125 5.130 5.658  5.1090 5.622
## cvpred      4.711  5.279  7.68 5.93 5.525  6.145 5.125 5.699  5.0897 5.646
## QA          5.000  5.000  5.00 7.00 6.000  6.000 6.000 6.000  5.0000 6.000
## CV residual 0.289 -0.279 -2.68 1.07 0.475 -0.145 0.875 0.301 -0.0897 0.354
##                681   683   694    699   703   706     713    719   725
## Predicted    5.373  5.43  5.18  5.115 5.285 4.794  5.0257  5.240  5.23
## cvpred       5.444  5.45  5.21  5.134 5.248 4.743  5.0328  5.261  5.16
## QA           5.000  5.00  5.00  5.000 6.000 5.000  5.0000  5.000  4.00
## CV residual -0.444 -0.45 -0.21 -0.134 0.752 0.257 -0.0328 -0.261 -1.16
##                728   741    744   746    754   766   769    780    783
## Predicted    5.363 5.816  5.415 5.632  5.172 5.123 5.103  5.279  5.016
## cvpred       5.318 5.771  5.489 5.342  5.173 5.136 5.113  5.288  5.051
## QA           5.000 6.000  5.000 6.000  5.000 6.000 6.000  5.000  5.000
## CV residual -0.318 0.229 -0.489 0.658 -0.173 0.864 0.887 -0.288 -0.051
##                784   799   806   810   816   822    828  829   830   834
## Predicted    5.390 5.668 6.763 5.571  5.97 6.844  5.634 6.45 5.921  5.82
```

```
## cvpred       5.349 5.667 6.762 5.549  6.02 6.823  5.634 6.47 5.869  5.87
## QA           5.000 6.000 7.000 6.000  5.00 7.000  5.000 8.00 6.000  4.00
## CV residual -0.349 0.333 0.238 0.451 -1.02 0.177 -0.634 1.53 0.131 -1.87
##                837   848    855    857   858    860    867   870    871
## Predicted   6.167 5.304  6.117  6.117 6.302 6.0236  6.146 5.714 5.9213
## cvpred      6.213 5.272  6.154  6.154 6.309 5.9898  6.111 5.723 5.9407
## QA          7.000 6.000  6.000  6.000 7.000 6.0000  6.000 6.000 6.0000
## CV residual 0.787 0.728 -0.154 -0.154 0.691 0.0102 -0.111 0.277 0.0593
##                  880   882    883    891   901  903   918   925   930   932
## Predicted    5.0740 5.836  6.333  5.713  6.38 5.64 5.701  6.20 6.526  5.36
## cvpred       5.0731 5.841  6.354  5.721  6.39 5.62 5.637  6.21 6.525  5.35
## QA           5.0000 6.000  6.000  5.000  5.00 7.00 6.000  5.00 7.000  5.00
## CV residual -0.0731 0.159 -0.354 -0.721 -1.39 1.38 0.363 -1.21 0.475 -0.35
##                933    936    940   953    955    956    958    972   973
## Predicted   5.410  6.336  5.929 6.227  6.144  5.994 5.9240  6.419 6.203
## cvpred      5.416  6.366  5.888 6.216  6.123  5.987 5.9471  6.428 6.221
## QA          6.000  6.000  5.000 7.000  6.000  5.000 6.0000  6.000 7.000
## CV residual 0.584 -0.366 -0.888 0.784 -0.123 -0.987 0.0529 -0.428 0.779
##                978   1000 1001  1006   1010   1016   1020   1027  1034
## Predicted   4.844  6.198 6.19 6.453  5.822  6.401  5.616  6.467 5.582
## cvpred      4.883  6.183 6.16 6.443  5.857  6.434  5.592  6.489 5.578
## QA          5.000  6.000 7.00 7.000  5.000  6.000  5.000  6.000 6.000
## CV residual 0.117 -0.183 0.84 0.557 -0.857 -0.434 -0.592 -0.489 0.422
##               1037  1038   1051    1061   1064 1065  1066   1081 1082
## Predicted   6.478 4.775  5.696  6.0443  6.383 5.88 5.473  6.389 5.85
## cvpred      6.506 4.749  5.699  6.0765  6.399 5.88 5.452  6.425 5.92
## QA          7.000 5.000  5.000  6.0000  6.000 6.00 6.000  6.000 7.00
## CV residual 0.494 0.251 -0.699 -0.0765 -0.399 0.12 0.548 -0.425 1.08
##                1093   1104  1106  1109  1112 1116  1119  1120   1127  1131
## Predicted   5.9210  6.178  6.16 4.978 6.089 5.68  6.59  6.18  6.687 5.676
## cvpred      5.9098  6.171  6.14 4.945 6.056 5.66  6.62  6.14  6.677 5.698
## QA          6.0000  6.000  5.00 5.000 7.000 6.00  6.00  5.00  6.000 6.000
## CV residual 0.0902 -0.171 -1.14 0.055 0.944 0.34 -0.62 -1.14 -0.677 0.302
##                1132  1135  1151  1159   1160   1165   1173   1179 1188
## Predicted    5.508 6.467 6.706  6.22  5.905  5.349  6.425  5.521 5.95
## cvpred       5.565 6.476 6.706  6.16  5.938  5.333  6.437  5.468 5.95
## QA           5.000 7.000 7.000  6.00  5.000  5.000  6.000  5.000 6.00
## CV residual -0.565 0.524 0.294 -0.16 -0.938 -0.333 -0.437 -0.468 0.05
##                1190   1192 1195   1199 1203  1211  1213    1214  1225
## Predicted    4.897  5.114 4.96  6.167 6.25 5.548 5.548  6.0334  6.10
## cvpred       4.884  5.068 4.95  6.198 6.24 5.521 5.521  6.0401  6.15
## QA           4.000  5.000 6.00  6.000 8.00 6.000 6.000  6.0000  6.00
## CV residual -0.884 -0.068 1.05 -0.198 1.76 0.479 0.479 -0.0401 -0.15
##                1233  1234   1248    1261  1262  1267  1292   1296   1301
## Predicted    5.393  5.73  5.610  5.1100  5.27 5.554 5.767  5.150 5.9818
## cvpred       5.388  5.77  5.595  5.0796  5.22 5.548 5.785  5.171 5.9384
## QA           5.000  4.00  5.000  5.0000  4.00 6.000 6.000  5.000 6.0000
## CV residual -0.388 -1.77 -0.595 -0.0796 -1.22 0.452 0.215 -0.171 0.0616
##              1302   1303 1305   1307   1310  1314   1321  1324  1326  1327
## Predicted   5.66  6.298  4.8  5.268  5.149 5.563  5.297 6.243 5.722 5.722
## cvpred      5.51  6.299  4.8  5.294  5.151 5.569  5.315 6.269 5.707 5.707
## QA          6.00  6.000  5.0  5.000  5.000 6.000  5.000 7.000 6.000 6.000
## CV residual 0.49 -0.299  0.2 -0.294 -0.151 0.431 -0.315 0.731 0.293 0.293
##               1330  1332   1335   1338   1339    1349    1362   1366
```

```
## Predicted   5.147 4.734 4.9854  5.310  5.310  5.0850  5.0507  5.317
## cvpred      5.149 4.714 4.9557  5.298  5.298  5.0702  5.0407  5.326
## QA          6.000 5.000 5.0000  5.000  5.000  5.0000  5.0000  5.000
## CV residual 0.851 0.286 0.0443 -0.298 -0.298 -0.0702 -0.0407 -0.326
##                1377  1380   1384  1386   1388   1390   1393    1395   1398
## Predicted    5.0311 5.724  5.149 4.774  5.237  5.273  5.332  5.0413  5.117
## cvpred       5.0225 5.723  5.178 4.779  5.221  5.323  5.323  5.0245  5.127
## QA           5.0000 6.000  5.000 5.000  5.000  5.000  5.000  5.0000  5.000
## CV residual -0.0225 0.277 -0.178 0.221 -0.221 -0.323 -0.323 -0.0245 -0.127
##                1401  1406   1419   1421   1422   1428  1430  1435   1438
## Predicted    4.9844 6.444  5.402  5.402  5.407  5.916  6.55 5.298  5.316
## cvpred       5.0262 6.445  5.416  5.416  5.422  5.906  6.57 5.265  5.241
## QA           5.0000 7.000  5.000  5.000  5.000  5.000  5.00 6.000  5.000
## CV residual -0.0262 0.555 -0.416 -0.416 -0.422 -0.906 -1.57 0.735 -0.241
##               1441 1446  1453   1455  1456   1458  1463  1476   1486  1490
## Predicted   6.174 4.87 6.112 5.9488 5.436  5.116 5.484 6.671  5.222 5.883
## cvpred      6.167 4.85 6.124 5.9747 5.387  5.115 5.468 6.657  5.187 5.831
## QA          7.000 6.00 7.000 6.0000 6.000  5.000 6.000 7.000  5.000 6.000
## CV residual 0.833 1.15 0.876 0.0253 0.613 -0.115 0.532 0.343 -0.187 0.169
##                1493  1504  1507  1509   1512  1521  1522   1524  1529
## Predicted    5.899 5.873 5.486 5.895  5.348 5.844  5.13  5.603 5.702
## cvpred       5.867 5.811 5.437 5.861  5.322 5.828  5.10  5.604 5.704
## QA           5.000 6.000 6.000 6.000  5.000 6.000  4.00  5.000 6.000
## CV residual -0.867 0.189 0.563 0.139 -0.322 0.172 -1.10 -0.604 0.296
##                1534  1537  1541  1542   1544    1551    1552   1561   1562
## Predicted    5.302 5.615 5.919 6.145 5.8685  5.0981  5.1074  5.093  5.093
## cvpred       5.319 5.588 5.878 6.148 5.9037  5.0794  5.0891  5.134  5.134
## QA           5.000 6.000 6.000 7.000 6.0000  5.0000  5.0000  5.000  5.000
## CV residual -0.319 0.412 0.122 0.852 0.0963 -0.0794 -0.0891 -0.134 -0.134
##                1567  1569    1577   1581  1593
## Predicted    6.234  5.23  6.0829  6.223 5.898
## cvpred       6.216  5.22  6.0626  6.218 5.883
## QA           6.000  5.00  6.0000  6.000 6.000
## CV residual -0.216 -0.22 -0.0626 -0.218 0.117
##
## Sum of squares = 132    Mean square = 0.41    n = 320
##
## fold 3
## Observations in test set: 320
##                  11     14    20    21     27   32    33     39     50
## Predicted    5.0342  5.959 5.416 5.608  5.505 5.41  5.37  4.518  5.193
## cvpred       5.0211  5.973 5.418 5.579  5.497 5.42  5.37  4.561  5.164
## QA           5.0000  5.000 6.000 6.000  5.000 6.00  5.00  4.000  5.000
## CV residual -0.0211 -0.973 0.582 0.421 -0.497 0.58 -0.37 -0.561 -0.164
##                 51    53     54    60     67     77     79     82     89
## Predicted    5.304 5.457  5.306 5.369  5.332  5.649  5.022  5.262  5.870
## cvpred       5.313 5.442  5.255 5.345  5.336  5.672  5.023  5.292  5.839
## QA           5.000 6.000  5.000 6.000  5.000  5.000  5.000  5.000  5.000
## CV residual -0.313 0.558 -0.255 0.655 -0.336 -0.672 -0.023 -0.292 -0.839
##                 90     92     98    107   109   110    112   134    142
## Predicted    5.0006  6.458  5.308  5.264 5.584 4.736  5.234 5.520  5.265
## cvpred       5.0202  6.425  5.303  5.291 5.566 4.731  5.204 5.511  5.269
## QA           5.0000  6.000  5.000  5.000 6.000 5.000  5.000 6.000  5.000
## CV residual -0.0202 -0.425 -0.303 -0.291 0.434 0.269 -0.204 0.489 -0.269
```

```
##                146     159   163   171    177   178    196    202    217
## Predicted    4.838  5.0771 5.503  4.70  5.363 5.515 4.9792  5.608  5.599
## cvpred       4.831  5.0853 5.502  4.73  5.344 5.532 4.9602  5.569  5.603
## QA           5.000  5.0000 6.000  4.00  5.000 6.000 5.0000  5.000  5.000
## CV residual 0.169 -0.0853 0.498 -0.73 -0.344 0.468 0.0398 -0.569 -0.603
##                223   226  231   233    234    241    251    253   260
## Predicted    5.359 5.703 5.98 5.592  5.563  5.253 5.9751  5.838 6.016
## cvpred       5.358 5.682 5.95 5.571  5.576  5.264 5.9727  5.828 6.027
## QA           5.000 6.000 7.00 6.000  5.000  5.000 6.0000  5.000 7.000
## CV residual -0.358 0.318 1.05 0.429 -0.576 -0.264 0.0273 -0.828 0.973
##                265    274    275    278  280   285    299   309   313   318
## Predicted    6.07  5.170  5.228  6.314 5.89  5.44  5.313 5.421 5.294 5.366
## cvpred       6.09  5.176  5.175  6.282 5.85  5.44  5.348 5.445 5.269 5.377
## QA           5.00  5.000  5.000  6.000 7.00  5.00  5.000 6.000 6.000 6.000
## CV residual -1.09 -0.176 -0.175 -0.282 1.15 -0.44 -0.348 0.555 0.731 0.623
##                321    322   324    329   333   340    346  347   354    355
## Predicted    5.80  5.197 5.380 5.9246 5.040 6.554  5.449 6.07  6.35  6.303
## cvpred       5.82  5.197 5.379 5.9143 5.007 6.542  5.471 6.09  6.37  6.242
## QA           7.00  5.000 6.000 6.0000 6.000 7.000  5.000 7.00  5.00  6.000
## CV residual 1.18 -0.197 0.621 0.0857 0.993 0.458 -0.471 0.91 -1.37 -0.242
##                379   381   383   394    397     400    402   407  408
## Predicted    6.719 5.787 5.787 4.772 5.0174  5.0886 5.9429 5.887 5.94
## cvpred       6.717 5.774 5.774 4.772 4.9674  5.0837 5.9172 5.895 5.95
## QA           6.000 6.000 6.000 5.000 5.0000  5.0000 6.0000 6.000 7.00
## CV residual -0.717 0.226 0.226 0.228 0.0326 -0.0837 0.0828 0.105 1.05
##                424   430    433  441   452   455   460    461   463    466
## Predicted    6.436 5.353  6.649 6.01 5.244  6.10  5.11 5.9546  6.36  5.911
## cvpred       6.427 5.393  6.642 6.00 5.272  6.08  5.13 5.9734  6.35  5.919
## QA           7.000 6.000  6.000 8.00 6.000  5.00  3.00 6.0000  5.00  5.000
## CV residual 0.573 0.607 -0.642 2.00 0.728 -1.08 -2.13 0.0266 -1.35 -0.919
##                 473    475    476    479    481    483    487   492   502
## Predicted    6.0427 5.9452  5.360  5.360  5.871  5.876  5.349 6.790 6.377
## cvpred       6.0306 5.9362  5.389  5.389  5.707  5.878  5.367 6.791 6.342
## QA           6.0000 6.0000  5.000  5.000  5.000  5.000  5.000 7.000 7.000
## CV residual -0.0306 0.0638 -0.389 -0.389 -0.707 -0.878 -0.367 0.209 0.658
##                504    523    525    526     542   545    546       549
## Predicted    6.507  5.688  5.303  5.409  6.0481 5.683  5.148  6.016528
## cvpred       6.499  5.668  5.273  5.426  6.0172 5.679  5.128  6.000511
## QA           7.000  5.000  5.000  5.000  6.0000 6.000  5.000  6.000000
## CV residual 0.501 -0.668 -0.273 -0.426 -0.0172 0.321 -0.128 -0.000511
##                552    565  568    571    578    583  587   590    593
## Predicted    5.722  6.291 4.79  6.285  5.215  5.251 6.31 6.140  5.324
## cvpred       5.708  6.344 4.83  6.289  5.191  5.267 6.30 6.126  5.323
## QA           6.000  6.000 6.00  6.000  5.000  5.000 7.00 7.000  5.000
## CV residual 0.292 -0.344 1.17 -0.289 -0.191 -0.267 0.70 0.874 -0.323
##                594   597   599    601   605    614   629    630   631
## Predicted    5.140 5.613 5.290  4.953 5.102  5.756 5.275 4.9731 5.275
## cvpred       5.147 5.609 5.288  4.979 5.093  5.747 5.271 4.9621 5.271
## QA           5.000 6.000 6.000  4.000 6.000  5.000 6.000 5.0000 6.000
## CV residual -0.147 0.391 0.712 -0.979 0.907 -0.747 0.729 0.0379 0.729
##                635    640   641   642   644   645   650    651    652
## Predicted    5.480  6.558  5.35  5.24  5.24  5.35 5.786  5.408  5.100
## cvpred       5.444  6.536  5.36  5.25  5.25  5.36 5.649  5.407  5.144
## QA           5.000  6.000  5.00  5.00  5.00  5.00 6.000  5.000  5.000
```

```
## CV residual -0.444 -0.536 -0.36 -0.25 -0.25 -0.36 0.351 -0.407 -0.144
##                654    657    665    666   672    686     687   688   692
## Predicted    6.146  5.408  5.735  5.346  5.11  5.565  5.0636  5.21 4.731
## cvpred       6.155  5.407  5.744  5.352  5.13  5.598  5.0741  5.23 4.765
## QA           6.000  5.000  5.000  5.000  5.00  5.000  5.0000  5.00 5.000
## CV residual -0.155 -0.407 -0.744 -0.352 -0.13 -0.598 -0.0741 -0.23 0.235
##                693   698    700    709     712    715    723    726   737
## Predicted    5.120 5.282  6.008 5.9344  5.0307  5.106  5.567  5.495 4.786
## cvpred       5.142 5.294  6.015 5.9346  5.0284  5.106  5.545  5.504 4.814
## QA           5.000 6.000  6.000 6.0000  5.0000  5.000  5.000  5.000 5.000
## CV residual -0.142 0.706 -0.015 0.0654 -0.0284 -0.106 -0.545 -0.504 0.186
##                738   750    751    761   765    767    770   771   772
## Predicted   5.143 5.464  5.172  5.156 5.113  5.132  5.146 5.103 4.891
## cvpred      5.164 5.473  5.174  5.131 5.122  5.109  5.153 5.094 4.866
## QA          6.000 6.000  5.000  5.000 6.000  5.000  5.000 6.000 5.000
## CV residual 0.836 0.527 -0.174 -0.131 0.878 -0.109 -0.153 0.906 0.134
##                782   788   789    794   800   805    813    824   827
## Predicted    5.390 5.421 5.421  5.700 5.668 5.491  5.975  5.383 6.176
## cvpred       5.395 5.433 5.433  5.702 5.675 5.482  5.973  5.396 6.154
## QA           5.000 6.000 6.000  5.000 6.000 6.000  5.000  5.000 7.000
## CV residual -0.395 0.567 0.567 -0.702 0.325 0.518 -0.973 -0.396 0.846
##                831    835   839    840   843    844  856   859    863
## Predicted    5.60  5.101 6.360  5.317 5.901 4.9267 5.93 6.226  5.563
## cvpred       5.61  5.118 6.352  5.322 5.889 4.9216 5.95 6.201  5.553
## QA           4.00  5.000 7.000  5.000 6.000 5.0000 7.00 7.000  5.000
## CV residual -1.61 -0.118 0.648 -0.322 0.111 0.0784 1.05 0.799 -0.553
##                 864     865     866     869   873    878     884   885
## Predicted    5.0825  5.0394  5.0678  6.0236  5.66 5.9213  5.0740 5.298
## cvpred       5.0761  5.0272  5.0607  6.0244  5.64 5.9119  5.0619 5.277
## QA           5.0000  5.0000  5.0000  6.0000  4.00 6.0000  5.0000 6.000
## CV residual -0.0761 -0.0272 -0.0607 -0.0244 -1.64 0.0881 -0.0619 0.723
##                888   889     894   895   896   897  905   909    911    927
## Predicted   6.420 5.751  5.0694 5.128 5.640 6.600 5.76 5.740  6.487  6.233
## cvpred      6.402 5.733  5.0668 5.125 5.634 6.577 5.71 5.721  6.481  6.208
## QA          7.000 6.000  5.0000 6.000 6.000 7.000 7.00 6.000  6.000  6.000
## CV residual 0.598 0.267 -0.0668 0.875 0.366 0.423 1.29 0.279 -0.481 -0.208
##                928  939   942  943   946   949   951    961    964    977
## Predicted    4.985 6.54 6.584 5.61 6.226 6.664 6.664  6.188  6.240  5.284
## cvpred       4.979 6.51 6.524 5.60 6.211 6.645 6.645  6.174  6.222  5.272
## QA           4.000 7.00 7.000 7.00 7.000 7.000 7.000  6.000  6.000  5.000
## CV residual -0.979 0.49 0.476 1.40 0.789 0.355 0.355 -0.174 -0.222 -0.272
##                980   983   984    985    989    995   996  1004  1008
## Predicted    5.963  6.52 5.815  5.963  5.419  5.218 5.455 6.453 6.555
## cvpred       5.959  6.52 5.831  5.959  5.402  5.202 5.462 6.443 6.534
## QA           5.000  6.00 6.000  5.000  5.000  5.000 6.000 7.000 7.000
## CV residual -0.959 -0.52 0.169 -0.959 -0.402 -0.202 0.538 0.557 0.466
##               1014  1029    1033  1039    1040   1041  1044  1047 1056
## Predicted   5.304 5.673  5.0280 6.627  6.0692  5.103 6.245 5.480 4.95
## cvpred      5.321 5.642  5.0385 6.567  6.0719  5.147 6.159 5.501 4.95
## QA          6.000 6.000  5.0000 7.000  6.0000  5.000 7.000 6.000 6.00
## CV residual 0.679 0.358 -0.0385 0.433 -0.0719 -0.147 0.841 0.499 1.05
##               1059  1060 1067  1068  1071  1072 1080  1084  1085  1087 1091
## Predicted   6.076 6.142 6.35 6.429 6.410 4.896 5.88  6.31 5.414 6.268 6.40
## cvpred      6.088 6.133 6.34 6.399 6.388 4.834 5.70  6.31 5.402 6.247 6.36
```

```
## QA          7.000 7.000 7.00 7.000 7.000 5.000 7.00  6.00 6.000 7.000 8.00
## CV residual 0.912 0.867 0.66 0.601 0.612 0.166 1.30 -0.31 0.598 0.753 1.64
##              1110  1111 1126   1129  1141   1149   1150   1153  1161  1162
## Predicted   5.795 5.650 6.32  5.600 5.420  6.124  6.302  5.246 6.177 5.857
## cvpred      5.803 5.646 6.28  5.579 5.409  6.124  6.299  5.241 6.169 5.852
## QA          6.000 6.000 7.00  5.000 6.000  6.000  6.000  5.000 7.000 6.000
## CV residual 0.197 0.354 0.72 -0.579 0.591 -0.124 -0.299 -0.241 0.831 0.148
##              1168   1171   1176   1181   1183   1187   1189  1196    1204
## Predicted   6.650 5.9747 5.9024  6.127  6.156  5.793  5.406 5.338  5.0955
## cvpred      6.636 5.9728 5.9085  6.107  6.141  5.759  5.391 5.331  5.0548
## QA          7.000 6.0000 6.0000  6.000  6.000  5.000  5.000 6.000  5.0000
## CV residual 0.364 0.0272 0.0915 -0.107 -0.141 -0.759 -0.391 0.669 -0.0548
##               1212  1217   1228  1231    1235   1247    1249   1252
## Predicted    5.380 5.240  5.444  6.52 5.99692  5.147  6.0510  5.355
## cvpred       5.356 5.229  5.456  6.47 5.99303  5.145  6.0405  5.345
## QA           5.000 6.000  5.000  6.00 6.00000  5.000  6.0000  5.000
## CV residual -0.356 0.771 -0.456 -0.47 0.00697 -0.145 -0.0405 -0.345
##                1257  1260   1265  1266   1268  1269   1271 1279   1286
## Predicted    5.1114 5.695  6.350 5.554  6.406 5.344  6.904 4.97  5.865
## cvpred       5.0913 5.697  6.329 5.544  6.395 5.338  6.875 4.97  5.853
## QA           5.0000 6.000  6.000 6.000  6.000 6.000  6.000 6.00  5.000
## CV residual -0.0913 0.303 -0.329 0.456 -0.395 0.662 -0.875 1.03 -0.853
##               1291   1306  1325  1333   1337   1347  1351   1354  1358
## Predicted    5.577  5.161 5.722 5.488  5.310  5.806  5.42  5.228 5.799
## cvpred       5.592  5.151 5.716 5.464  5.294  5.812  5.41  5.237 5.792
## QA           5.000  5.000 6.000 6.000  5.000  5.000  5.00  5.000 6.000
## CV residual -0.592 -0.151 0.284 0.536 -0.294 -0.812 -0.41 -0.237 0.208
##               1359  1363  1371   1372  1374  1375   1376  1396   1397
## Predicted   4.9849 5.752 4.910  6.160 4.816  4.79  5.134 5.279  5.307
## cvpred      4.9431 5.744 4.937  6.164 4.793  4.81  5.114 5.301  5.327
## QA          5.0000 6.000 5.000  6.000 5.000  3.00  5.000 6.000  5.000
## CV residual 0.0569 0.256 0.063 -0.164 0.207 -1.81 -0.114 0.699 -0.327
##              1400   1403   1407  1424  1432   1433  1436    1437  1452
## Predicted   5.799  6.426  6.433  5.71 5.417  6.329 5.298  5.0553 6.064
## cvpred      5.806  6.432  6.386  5.70 5.389  6.328 5.164  5.0472 6.071
## QA          6.000  6.000  6.000  4.00 6.000  6.000 6.000  5.0000 7.000
## CV residual 0.194 -0.432 -0.386 -1.70 0.611 -0.328 0.836 -0.0472 0.929
##              1462  1472  1478 1495  1503  1505  1510  1514 1516   1519
## Predicted    5.22  6.06 6.680 5.87  5.24 5.884  6.36 5.573 4.92  5.600
## cvpred       5.24  6.04 6.646 5.84  5.23 5.899  6.32 5.539 4.92  5.614
## QA           4.00  5.00 7.000 7.00  5.00 6.000  5.00 6.000 6.00  5.000
## CV residual -1.24 -1.04 0.354 1.16 -0.23 0.101 -1.32 0.461 1.08 -0.614
##                1531 1533   1547   1549 1550   1570   1571    1572  1573
## Predicted    6.0891 5.63  5.564  5.765 6.22 5.9148  6.342  6.0564 4.956
## cvpred       6.0746 5.62  5.561  5.758 6.20 5.9069  6.343  6.0438 4.946
## QA           6.0000 6.00  5.000  5.000 8.00 6.0000  6.000  6.0000 5.000
## CV residual -0.0746 0.38 -0.561 -0.758 1.80 0.0931 -0.343 -0.0438 0.054
##              1575  1578  1579   1583  1585   1587   1597   1599
## Predicted   5.662 5.796 5.686  5.777 6.336  6.358 5.8981 5.9652
## cvpred      5.495 5.791 5.712  5.777 6.324  6.331 5.9005 5.9402
## QA          6.000 6.000 6.000  5.000 7.000  6.000 6.0000 6.0000
## CV residual 0.505 0.209 0.288 -0.777 0.676 -0.331 0.0995 0.0598
##
## Sum of squares = 121    Mean square = 0.38    n = 320
```

```
##
## fold 4
## Observations in test set: 320
##                13     18    30     41    44     57     73     91    93
## Predicted    5.20  5.378 5.374  5.750  6.03  5.687 4.9179  5.130  6.47
## cvpred       5.19  5.361 5.356  5.756  6.05  5.679 4.9108  5.117  6.53
## QA           5.00  5.000 6.000  5.000  5.00  5.000 5.0000  5.000  5.00
## CV residual -0.19 -0.361 0.644 -0.756 -1.05 -0.679 0.0892 -0.117 -1.53
##                 96   103     104    111   114   116  129   133    136
## Predicted   5.9382 5.200  5.0764  5.558 5.736 5.766 5.92  6.18  5.321
## cvpred      5.9592 5.181  5.0647  5.559 5.718 5.756 5.93  6.20  5.321
## QA          6.0000 6.000  5.0000  5.000 6.000 6.000 7.00  5.00  5.000
## CV residual 0.0408 0.819 -0.0647 -0.559 0.282 0.244 1.07 -1.20 -0.321
##                137    138    141     147    154   157   160     166   172
## Predicted    5.265  5.324  5.321  5.0836  5.283  5.61 5.008  5.0210 5.438
## cvpred       5.264  5.314  5.321  5.0792  5.273  5.61 5.002  5.0256 5.419
## QA           5.000  5.000  5.000  5.0000  5.000  5.00 6.000  5.0000 6.000
## CV residual -0.264 -0.314 -0.321 -0.0792 -0.273 -0.61 0.998 -0.0256 0.581
##                173 185     190    194    195   200   207   210   212
## Predicted   5.438   5  5.0211  5.346  5.346  5.14 6.202 6.277 5.151
## cvpred      5.419   5  5.0239  5.332  5.332  5.14 6.199 6.285 5.146
## QA          6.000   6  5.0000  5.000  5.000  4.00 7.000 7.000 6.000
## CV residual 0.581   1 -0.0239 -0.332 -0.332 -1.14 0.801 0.715 0.854
##                218     220   221  235   236     242   243   244   245
## Predicted   5.00308  5.0715 5.479 4.71 5.058  6.0554 5.135 6.179 6.179
## cvpred      4.99337  5.0762 5.463 4.70 5.035  6.0459 5.116 6.169 6.169
## QA          5.00000  5.0000 6.000 6.00 6.000  6.0000 6.000 7.000 7.000
## CV residual 0.00663 -0.0762 0.537 1.30 0.965 -0.0459 0.884 0.831 0.831
##                255     256   271   276  282   286   287   288  291   295
## Predicted   5.456  5.0498 5.647 5.755 5.70  5.44 5.884 5.793 5.90 5.827
## cvpred      5.441  5.0496 5.582 5.737 5.67  5.41 5.871 5.771 5.89 5.811
## QA          6.000  5.0000 6.000 6.000 7.00  5.00 6.000 6.000 7.00 6.000
## CV residual 0.559 -0.0496 0.418 0.263 1.33 -0.41 0.129 0.229 1.11 0.189
##                306    307   311   328    332   336    337    341   343
## Predicted   5.311  5.049 5.311  6.27  6.286 6.144  6.486  6.188 5.784
## cvpred      5.294  5.042 5.294  6.27  6.278 6.142  6.495  6.197 5.774
## QA          6.000  5.000 6.000  5.00  6.000 7.000  6.000  6.000 6.000
## CV residual 0.706 -0.042 0.706 -1.27 -0.278 0.858 -0.495 -0.197 0.226
##                344    353    356   360    366  367    371   382   384  390
## Predicted   5.784  5.294 5.9911 5.688  6.403 5.56  5.339 5.894 5.787 5.72
## cvpred      5.774  5.277 5.9815 5.675  6.406 5.55  5.327 5.883 5.778 5.71
## QA          6.000  5.000 6.0000 6.000  6.000 7.00  5.000 6.000 6.000 7.00
## CV residual 0.226 -0.277 0.0185 0.325 -0.406 1.45 -0.327 0.117 0.222 1.29
##                392    395   403    406   411    419  422   428     432
## Predicted   5.894  5.494 5.692 5.9528 5.362 5.9358 5.76 5.313  5.1109
## cvpred      5.883  5.484 5.677 5.9528 5.347 5.9221 5.75 5.298  5.0957
## QA          6.000  5.000 6.000 6.0000 6.000 6.0000 7.00 6.000  5.0000
## CV residual 0.117 -0.484 0.323 0.0472 0.653 0.0779 1.25 0.702 -0.0957
##                435   437  443   450   453     467    468    471    472
## Predicted   5.850 5.410 5.77 5.909 5.435  6.0503  6.900  5.965 5.9817
## cvpred      5.844 5.394 5.76 5.892 5.414  6.0424  6.913  5.949 5.9743
## QA          6.000 6.000 7.00 6.000 6.000  6.0000  6.000  5.000 6.0000
## CV residual 0.156 0.606 1.24 0.108 0.586 -0.0424 -0.913 -0.949 0.0257
##                480    486   488   491  496   497   505   506   507   512
```

```
## Predicted   5.478  5.349 5.199 5.436 6.13 5.195 6.529 6.434 6.494 5.449
## cvpred      5.466  5.336 5.131 5.424 6.13 5.178 6.538 6.442 6.503 5.425
## QA          6.000  5.000 6.000 6.000 8.00 6.000 7.000 7.000 7.000 6.000
## CV residual 0.534 -0.336 0.869 0.576 1.87 0.822 0.462 0.558 0.497 0.575
##                517   518    521    530   538   547  550   553    555    562
## Predicted    6.13  4.85  6.104  5.281 5.507 5.417 5.18 5.703  5.980  5.148
## cvpred       6.12  4.82  6.102  5.268 5.507 5.408 5.16 5.689  5.958  5.144
## QA           6.00  3.00  6.000  5.000 6.000 6.000 6.00 6.000  5.000  5.000
## CV residual -0.12 -1.82 -0.102 -0.268 0.493 0.592 0.84 0.311 -0.958 -0.144
##                566     570   574   576    581   585  589   591   592
## Predicted    5.917  6.0209  5.28  6.18  5.374 6.113 6.77  5.32 5.683
## cvpred       5.905  6.0087  5.27  6.17  5.346 6.084 6.78  5.31 5.701
## QA           5.000  6.0000  4.00  6.00  5.000 7.000 8.00  5.00 6.000
## CV residual -0.905 -0.0087 -1.27 -0.17 -0.346 0.916 1.22 -0.31 0.299
##                 596   600   602   607    610   613    619    620   621
## Predicted    5.0905 5.496 5.412 6.475  6.244 5.428  5.675  5.790  5.14
## cvpred       5.0711 5.479 5.391 6.475  6.229 5.406  5.654  5.779  5.14
## QA           5.0000 6.000 6.000 7.000  6.000 6.000  5.000  5.000  5.00
## CV residual -0.0711 0.521 0.609 0.525 -0.229 0.594 -0.654 -0.779 -0.14
##                622    623     625    632  639  646   648  663   668    676
## Predicted    5.138  5.349  5.1633  5.669 5.09 5.49  5.48 5.39 5.658  5.732
## cvpred       5.136  5.332  5.0819  5.651 5.07 5.47  5.47 5.38 5.638  5.718
## QA           5.000  5.000  5.0000  5.000 7.00 7.00  4.00 6.00 6.000  5.000
## CV residual -0.136 -0.332 -0.0819 -0.651 1.93 1.53 -1.47 0.62 0.362 -0.718
##                682    684    690  697     707   710    718    720   731
## Predicted   5.544  5.565  5.738 5.28  5.1098 5.805  5.490  5.129 4.879
## cvpred      5.518  5.558  5.737 5.27  5.0905 5.793  5.471  5.104 4.832
## QA          6.000  5.000  5.000 6.00  5.0000 6.000  5.000  5.000 5.000
## CV residual 0.482 -0.558 -0.737 0.73 -0.0905 0.207 -0.471 -0.104 0.168
##                740    742    748   749    752    753     759     764   774
## Predicted    5.155  5.129  5.327 5.476  5.172  5.210 5.01673  5.0727 5.560
## cvpred       5.126  5.122  5.308 5.465  5.145  5.187 4.99775  5.0527 5.557
## QA           5.000  5.000  5.000 6.000  5.000  5.000 5.00000  5.0000 6.000
## CV residual -0.126 -0.122 -0.308 0.535 -0.145 -0.187 0.00225 -0.0527 0.443
##                776   778   781    785    787   791    792    795    796
## Predicted   4.886 5.169 5.049  5.210  5.410 5.462 4.9902  6.501  5.638
## cvpred      4.867 5.135 5.015  5.196  5.393 5.458 4.9793  6.498  5.659
## QA          5.000 6.000 6.000  5.000  5.000 6.000 5.0000  6.000  5.000
## CV residual 0.133 0.865 0.985 -0.196 -0.393 0.542 0.0207 -0.498 -0.659
##                797   798     801    802   803   804   808    811   814
## Predicted    5.48 6.234  5.0302  5.410 6.402 5.323 6.763  5.614  6.00
## cvpred       5.46 6.242  5.0144  5.382 6.434 5.302 6.773  5.594  5.99
## QA           5.00 7.000  5.0000  5.000 7.000 6.000 7.000  5.000  4.00
## CV residual -0.46 0.758 -0.0144 -0.382 0.566 0.698 0.227 -0.594 -1.99
##                815   817    818    832   833     836   838    842    845
## Predicted    6.209 5.686  6.578 5.9100  5.78  5.0838 6.200  5.419  6.323
## cvpred       6.198 5.666  6.574 5.9126  5.76  5.0624 6.201  5.394  6.321
## QA           6.000 6.000  6.000 6.0000  3.00  5.0000 7.000  5.000  6.000
## CV residual -0.198 0.334 -0.574 0.0874 -2.76 -0.0624 0.799 -0.394 -0.321
##                849    852   853    868   874 876    881   893   898
## Predicted    5.273  5.467  5.71  6.110 6.244 6.4  5.470 5.570 5.640
## cvpred       5.262  5.452  5.72  6.116 6.233 6.4  5.452 5.552 5.619
## QA           5.000  5.000  5.00  6.000 7.000 7.0  5.000 6.000 6.000
## CV residual -0.262 -0.452 -0.72 -0.116 0.767 0.6 -0.452 0.448 0.381
```

```
##                912    916    917    919    920    921    922   924   938
## Predicted    6.0698  6.388  5.249 5.9986  6.291  5.743 5.9986 5.701  5.63
## cvpred       6.0414  6.385  5.249 5.9894  6.288  5.722 5.9894 5.673  5.63
## QA           6.0000  6.000  5.000 6.0000  6.000  5.000 6.0000 6.000  4.00
## CV residual -0.0414 -0.385 -0.249 0.0106 -0.288 -0.722 0.0106 0.327 -1.63
##               944   950     957  959    965    982    994   998  1003  1007
## Predicted   5.40 6.664  6.0663 5.83  6.188  5.234  5.319 6.109 6.569 6.569
## cvpred      5.38 6.664  6.0587 5.81  6.193  5.226  5.291 6.109 6.574 6.574
## QA          7.00 7.000  6.0000 7.00  6.000  5.000  5.000 7.000 7.000 7.000
## CV residual 1.62 0.336 -0.0587 1.19 -0.193 -0.226 -0.291 0.891 0.426 0.426
##               1013   1019   1021   1023  1026   1028  1035   1045  1057
## Predicted   4.836  6.531  6.211  5.642 5.254  5.614 5.073  6.276 6.142
## cvpred      4.821  6.533  6.201  5.626 5.232  5.604 5.053  6.286 6.136
## QA          5.000  6.000  6.000  5.000 6.000  5.000 6.000  6.000 7.000
## CV residual 0.179 -0.533 -0.201 -0.626 0.768 -0.604 0.947 -0.286 0.864
##               1076   1078 1090    1092  1094  1095  1097   1101   1107 1118
## Predicted   6.137  5.940 5.87  6.0400 6.559 5.367 5.367  6.613  6.484 5.68
## cvpred      6.122  5.927 5.86  6.0344 6.563 5.339 5.339  6.614  6.477 5.66
## QA          7.000  5.000 7.00  6.0000 7.000 6.000 6.000  6.000  6.000 6.00
## CV residual 0.878 -0.927 1.14 -0.0344 0.437 0.661 0.661 -0.614 -0.477 0.34
##               1121   1136   1138  1157  1158   1170  1174  1182   1185
## Predicted   6.57  6.322  6.103 6.482 6.371 5.9719 5.461  6.07  5.406
## cvpred      6.58  6.318  6.101 6.478 6.404 5.9663 5.445  6.06  5.423
## QA          8.00  6.000  6.000 7.000 7.000 6.0000 6.000  5.00  5.000
## CV residual 1.42 -0.318 -0.101 0.522 0.596 0.0337 0.555 -1.06 -0.423
##                1191    1194  1205  1206  1207   1218  1223   1227   1244
## Predicted    6.321  5.1076 6.146 6.146 6.146  6.516 5.179  5.205  5.148
## cvpred       6.315  5.0838 6.149 6.149 6.149  6.516 5.156  5.171  5.143
## QA           6.000  5.0000 7.000 7.000 7.000  6.000 6.000  5.000  5.000
## CV residual -0.315 -0.0838 0.851 0.851 0.851 -0.516 0.844 -0.171 -0.143
##               1251   1253   1255   1256   1264 1270    1272  1275 1276
## Predicted   5.731  5.169  5.666  5.563  5.001 6.82  6.0723 5.740 4.97
## cvpred      5.713  5.148  5.654  5.546  4.971 6.84  6.0718 5.724 4.97
## QA          6.000  5.000  5.000  5.000  4.000 8.00  6.0000 6.000 6.00
## CV residual 0.287 -0.148 -0.654 -0.546 -0.971 1.16 -0.0718 0.276 1.03
##               1278  1281  1282  1283   1289   1290   1297  1298    1299
## Predicted   5.200 5.692 5.692 5.597  5.698  5.698  5.150  6.18  6.0683
## cvpred      5.179 5.677 5.677 5.587  5.718  5.718  5.107  6.18  6.0669
## QA          6.000 6.000 6.000 6.000  5.000  5.000  5.000  6.00  6.0000
## CV residual 0.821 0.323 0.323 0.413 -0.718 -0.718 -0.107 -0.18 -0.0669
##              1304   1312  1323  1328   1329  1331   1345  1346  1368 1369
## Predicted    6.1  6.341  6.22 5.722  5.216 5.147  5.935 5.644 5.806 4.97
## cvpred       6.1  6.339  6.21 5.711  5.199 5.143  5.916 5.629 5.837 4.96
## QA           5.0  6.000  5.00 6.000  5.000 6.000  5.000 6.000 6.000 6.00
## CV residual -1.1 -0.339 -1.21 0.289 -0.199 0.857 -0.916 0.371 0.163 1.04
##               1378  1379  1383   1389 1399  1409    1410  1411   1416
## Predicted   5.884 5.459  5.15  5.421 5.46 6.868  6.0475 5.437  5.482
## cvpred      5.881 5.438  5.14  5.417 5.44 6.876  6.0418 5.432  5.479
## QA          6.000 6.000  5.00  5.000 7.00 7.000  6.0000 6.000  5.000
## CV residual 0.119 0.562 -0.14 -0.417 1.56 0.124 -0.0418 0.568 -0.479
##                 1423   1429   1431   1447   1448 1450 1457  1465  1468 1469
## Predicted    6.1077  5.690  5.741  5.444  5.353  6.2 5.58  5.44  5.16 5.54
## cvpred       6.0938  5.672  5.732  5.428  5.336  6.2 5.57  5.42  5.15 5.52
## QA           6.0000  5.000  5.000  5.000  5.000  8.0 6.00  5.00  4.00 7.00
```

```
## CV residual -0.0938 -0.672 -0.732 -0.428 -0.336  1.8 0.43 -0.42 -1.15 1.48
##               1471    1473   1480  1481   1488   1489   1492    1497  1500
## Predicted    5.285  6.0829  5.873  5.41  5.718  5.872  5.872  5.0735 5.772
## cvpred       5.264  6.0827  5.859  5.41  5.719  5.871  5.871  5.0764 5.769
## QA           5.000  6.0000  5.000  4.00  5.000  5.000  5.000  5.0000 6.000
## CV residual -0.264 -0.0827 -0.859 -1.41 -0.719 -0.871 -0.871 -0.0764 0.231
##              1515  1523   1526  1538   1540   1548    1554    1557    1560
## Predicted   4.91  6.09  5.542 5.522  5.718  5.982  5.1061  5.0643  5.0934
## cvpred      4.91  6.09  5.526 5.511  5.707  5.981  5.0814  5.0515  5.0866
## QA          6.00  5.00  5.000 6.000  5.000  5.000  5.0000  5.0000  5.0000
## CV residual 1.09 -1.09 -0.526 0.489 -0.707 -0.981 -0.0814 -0.0515 -0.0866
##                1574    1576   1589  1590   1598
## Predicted    6.0631  6.0742  6.157 4.971  5.499
## cvpred       6.0663  6.0736  6.171 4.948  5.487
## QA           6.0000  6.0000  6.000 5.000  5.000
## CV residual -0.0663 -0.0736 -0.171 0.052 -0.487
##
## Sum of squares = 151    Mean square = 0.47    n = 320
##
## fold 5
## Observations in test set: 320
##                   5      6   17     24     26     28   38    42    46
## Predicted    5.0673  5.100 5.85  5.140  5.388  5.688 5.73  5.13  6.11
## cvpred       5.0878  5.118 5.82  5.188  5.404  5.701 5.72  5.13  6.17
## QA           5.0000  5.000 7.00  5.000  5.000  5.000 7.00  4.00  4.00
## CV residual -0.0878 -0.118 1.18 -0.188 -0.404 -0.701 1.28 -1.13 -2.17
##                 48     56     62     68    74    85    86     87     94
## Predicted    5.515  5.213 4.9051  5.433  5.00 5.815  5.46  6.458  5.471
## cvpred       5.501  5.241 4.9263  5.461  5.01 5.826  5.48  6.488  5.472
## QA           5.000  5.000 5.0000  5.000  4.00 6.000  5.00  6.000  5.000
## CV residual -0.501 -0.241 0.0737 -0.461 -1.01 0.174 -0.48 -0.488 -0.472
##                 97   100    105     106   117   119   121    125  127
## Predicted    5.328 5.200  5.231  5.0764 5.427 5.675 4.281  5.133 4.82
## cvpred       5.381 5.201  5.248  5.0846 5.427 5.692 4.345  5.138 4.84
## QA           5.000 6.000  5.000  5.0000 6.000 6.000 5.000  5.000 5.00
## CV residual -0.381 0.799 -0.248 -0.0846 0.573 0.308 0.655 -0.138 0.16
##                 130   132    144    145   150   152   162   174    176
## Predicted    5.441  6.18  5.461  6.963 5.682  5.60  5.61 5.784  5.522
## cvpred       5.466  6.23  5.487  6.988 5.691  5.75  5.62 5.789  5.538
## QA           5.000  5.00  5.000  6.000 6.000  4.00  4.00 6.000  5.000
## CV residual -0.466 -1.23 -0.487 -0.988 0.309 -1.75 -1.62 0.211 -0.538
##                 181    182     184    188     193    197   201    204  205
## Predicted    5.320  5.115  5.0232  5.098  4.9524  5.378 6.088  5.272 5.26
## cvpred       5.321  5.167  5.0491  5.123  5.0222  5.395 6.047  5.276 5.27
## QA           5.000  5.000  5.0000  5.000  5.0000  5.000 7.000  5.000 6.00
## CV residual -0.321 -0.167 -0.0491 -0.123 -0.0222 -0.395 0.953 -0.276 0.73
##                206      208    211    222   227    228   246     247
## Predicted   6.202  5.00668  6.493  5.169 5.763  5.175 5.590  5.0518
## cvpred      6.138  5.00797  6.457  5.189 5.827  5.185 5.622  5.0717
## QA          7.000  5.00000  6.000  5.000 6.000  5.000 6.000  5.0000
## CV residual 0.862 -0.00797 -0.457 -0.189 0.173 -0.185 0.378 -0.0717
##                 248    258   267  268    270  279    281  284    292   293
## Predicted    5.0344  5.064  5.05 6.59  6.314 6.80 5.9937 6.03  5.773 5.593
## cvpred       5.0464  5.101  5.08 6.61  6.298 6.82 5.9607 6.03  5.805 5.578
```

```
## QA           5.0000  5.000  4.00 8.00  6.000 8.00 6.0000 7.00  5.000 6.000
## CV residual -0.0464 -0.101 -1.08 1.39 -0.298 1.18 0.0393 0.97 -0.805 0.422
##                 304     314    316  319   325    338    339    342  345
## Predicted    4.9897  5.0115 5.8884 5.80 5.540  5.654  6.183  6.349 5.81
## cvpred       5.0356  5.0529 5.9084 5.79 5.631  5.657  6.144  6.319 5.78
## QA           5.0000  5.0000 6.0000 7.00 6.000  5.000  6.000  6.000 6.00
## CV residual -0.0356 -0.0529 0.0916 1.21 0.369 -0.657 -0.144 -0.319 0.22
##                351   352   357   358   362  365   378   380    385   389
## Predicted   5.674 5.265  6.18 6.222 5.791 5.56 6.867 5.793  5.327 5.385
## cvpred      5.663 5.282  6.14 6.201 5.791 5.55 6.851 5.795  5.332 5.385
## QA          6.000 6.000  5.00 7.000 6.000 7.00 7.000 6.000  5.000 6.000
## CV residual 0.337 0.718 -1.14 0.799 0.209 1.45 0.149 0.205 -0.332 0.615
##                393   396   399   404    405   410    413     417    418
## Predicted    5.435 6.459 5.749 5.599  5.128  5.71 4.9280  6.0758  5.242
## cvpred       5.412 6.428 5.711 5.568  5.147  5.65 4.9613  6.0477  5.261
## QA           5.000 7.000 6.000 6.000  5.000  4.00 5.0000  6.0000  5.000
## CV residual -0.412 0.572 0.289 0.432 -0.147 -1.65 0.0387 -0.0477 -0.261
##                425  426   427    438    440   444    448   459   462
## Predicted    5.111 5.76 5.619 6.0219  5.167 6.172  5.973 6.105  5.20
## cvpred       5.118 5.77 5.637 5.9854  5.183 6.145  5.977 6.085  5.22
## QA           5.000 7.00 6.000 6.0000  5.000 7.000  5.000 7.000  5.00
## CV residual -0.118 1.23 0.363 0.0146 -0.183 0.855 -0.977 0.915 -0.22
##                485   503   509   510    511   513   514    516   519
## Predicted    6.558 6.377 5.449 6.368  5.761 5.717 6.067  5.258  6.46
## cvpred       6.558 6.376 5.427 6.285  5.716 5.693 6.044  5.326  6.43
## QA           6.000 7.000 6.000 7.000  5.000 6.000 7.000  5.000  6.00
## CV residual -0.558 0.624 0.573 0.715 -0.716 0.307 0.956 -0.326 -0.43
##                520    524    527    528    533  535   537    540   544
## Predicted    5.881  5.207  5.881 5.9768  5.745 5.48  5.28  5.897 5.756
## cvpred       5.885  5.188  5.885 5.9893  5.706 5.47  5.27  5.894 5.733
## QA           5.000  5.000  5.000 6.0000  5.000 6.00  5.00  5.000 6.000
## CV residual -0.885 -0.188 -0.885 0.0107 -0.706 0.53 -0.27 -0.894 0.267
##                551    554    557    559    560    563   580    603   615
## Predicted   5.368  5.200 5.8970 5.8970  6.336  5.115 5.710 4.9325 5.759
## cvpred      5.393  5.228 5.9042 5.9042  6.314  5.102 5.659 4.9619 5.801
## QA          6.000  5.000 6.0000 6.0000  6.000  5.000 6.000 5.0000 6.000
## CV residual 0.607 -0.228 0.0958 0.0958 -0.314 -0.102 0.341 0.0381 0.199
##                616    617     618     628  638    647    655   669    679
## Predicted    5.196  5.196  6.1285  5.0355 4.67  5.362  5.244  5.64  5.098
## cvpred       5.168  5.168  6.0747  5.0391 4.67  5.415  5.241  5.61  5.093
## QA           5.000  5.000  6.0000  5.0000 5.00  5.000  5.000  5.00  5.000
## CV residual -0.168 -0.168 -0.0747 -0.0391 0.33 -0.415 -0.241 -0.61 -0.093
##                685   691    695  701   702   704   711   716    717    721
## Predicted   4.72  4.96  5.181 4.92 5.282  5.46 4.897 5.208  5.217  5.240
## cvpred      4.69  5.02  5.186 4.90 5.307  5.47 4.882 5.226  5.222  5.246
## QA          5.00  3.00  5.000 6.00 6.000  4.00 5.000 6.000  5.000  5.000
## CV residual 0.31 -2.02 -0.186 1.10 0.693 -1.47 0.118 0.774 -0.222 -0.246
##                722   724    732    734    735    739    743   747   755
## Predicted    5.132  6.35  5.414  5.229  5.179  5.170  5.216 5.205 5.111
## cvpred       5.132  6.37  5.445  5.256  5.194  5.174  5.244 5.208 5.202
## QA           5.000  5.00  5.000  5.000  5.000  5.000  5.000 6.000 6.000
## CV residual -0.132 -1.37 -0.445 -0.256 -0.194 -0.174 -0.244 0.792 0.798
##                756   757    760     762   763    768     773   775   807
## Predicted   5.195 5.185  5.228  5.0727 5.350 4.9362  5.0153 5.551 6.739
```

```
## cvpred      5.193 5.216  5.221  5.0703 5.346 4.9786  5.0145 5.589 6.719
## QA          6.000 6.000  5.000  5.0000 6.000 5.0000  5.0000 6.000 7.000
## CV residual 0.807 0.784 -0.221 -0.0703 0.654 0.0214 -0.0145 0.411 0.281
##                812    826   841    850     861   862   875   877   879
## Predicted   5.9739  5.634 6.417  5.270  5.0394 5.174 6.268  5.87 5.298
## cvpred      5.9167  5.658 6.386  5.293  5.0757 5.195 6.242  5.89 5.313
## QA          6.0000  5.000 7.000  5.000  5.0000 6.000 7.000  4.00 6.000
## CV residual 0.0833 -0.658 0.614 -0.293 -0.0757 0.805 0.758 -1.89 0.687
##                887    892   900  902  904    906    923   926    931   941
## Predicted   5.354  5.069  5.13 5.64 5.76  5.124  6.291 6.423  5.572 6.621
## cvpred      5.364  5.103  5.16 5.66 5.82  5.141  6.267 6.372  5.607 6.589
## QA          6.000  5.000  3.00 7.00 7.00  5.000  6.000 7.000  5.000 7.000
## CV residual 0.636 -0.103 -2.16 1.34 1.18 -0.141 -0.267 0.628 -0.607 0.411
##                947   948   952   954    960    963    966    971    974
## Predicted   6.556 6.482 6.482 6.586  5.421  5.326  6.263  6.419  5.960
## cvpred      6.552 6.476 6.476 6.542  5.435  5.371  6.247  6.375  5.943
## QA          7.000 7.000 7.000 7.000  5.000  5.000  6.000  6.000  5.000
## CV residual 0.448 0.524 0.524 0.458 -0.435 -0.371 -0.247 -0.375 -0.943
##                992  999   1005  1009  1011  1012   1018 1030 1032 1036
## Predicted    5.319 4.97  5.649 6.371 6.497 5.880  6.531 5.60 5.77 5.99
## cvpred       5.318 4.94  5.629 6.382 6.468 5.864  6.488 5.62 5.79 5.97
## QA           5.000 6.00  5.000 7.000 7.000 6.000  6.000 7.00 7.00 7.00
## CV residual -0.318 1.06 -0.629 0.618 0.532 0.136 -0.488 1.38 1.21 1.03
##               1042   1052 1055  1069   1070   1079   1086 1089   1096
## Predicted   5.524  5.428 4.95 6.429  5.580  5.940  5.353 5.87  5.675
## cvpred      5.534  5.511 4.98 6.378  5.584  5.981  5.349 5.80  5.664
## QA          6.000  5.000 6.00 7.000  5.000  5.000  5.000 7.00  5.000
## CV residual 0.466 -0.511 1.02 0.622 -0.584 -0.981 -0.349 1.20 -0.664
##               1102  1108   1115  1117  1125  1128  1130   1133  1134
## Predicted    6.178 6.539  6.658 5.676  5.82 5.697 5.842 6.9290 5.993
## cvpred       6.181 6.516  6.698 5.706  5.87 5.729 5.824 6.9287 6.058
## QA           6.000 7.000  6.000 6.000  4.00 6.000 6.000 7.0000 7.000
## CV residual -0.181 0.484 -0.698 0.294 -1.87 0.271 0.176 0.0713 0.942
##                1137    1142   1146  1148  1155  1163   1167   1169  1172
## Predicted    6.1033  6.0959  6.164 6.193 5.884 6.419  5.922  6.063 5.768
## cvpred       6.0738  6.0972  6.143 6.211 5.891 6.406  5.918  6.125 5.781
## QA           6.0000  6.0000  6.000 7.000 6.000 7.000  5.000  6.000 6.000
## CV residual -0.0738 -0.0972 -0.143 0.789 0.109 0.594 -0.918 -0.125 0.219
##               1175  1177   1184  1200  1201  1202  1209   1222   1226 1229
## Predicted   5.461  5.38  5.099 5.148 5.452 6.077 6.146  6.259  5.185 6.62
## cvpred      5.454  5.45  5.118 5.159 5.451 6.068 6.142  6.224  5.189 6.66
## QA          6.000  4.00  5.000 6.000 6.000 7.000 7.000  6.000  5.000 7.00
## CV residual 0.546 -1.45 -0.118 0.841 0.549 0.932 0.858 -0.224 -0.189 0.34
##               1236  1239  1240    1241 1245  1250   1254  1258  1280 1284
## Predicted    6.00  5.06  5.89  5.0836 5.76 5.731  5.277 5.672 6.311 5.37
## cvpred       6.17  5.08  5.95  5.0687 5.89 5.745  5.286 5.735 6.274 5.38
## QA           4.00  4.00  4.00  5.0000 6.00 6.000  5.000 6.000 7.000 6.00
## CV residual -2.17 -1.08 -1.95 -0.0687 0.11 0.255 -0.286 0.265 0.726 0.62
##               1285   1293  1295  1300  1308   1309   1311  1315   1317
## Predicted    5.853  6.386 5.767  4.43  5.62  5.269  5.161 5.559 5.9551
## cvpred       5.863  6.412 5.768  4.47  5.66  5.262  5.167 5.574 5.9749
## QA           5.000  6.000 6.000  3.00  4.00  5.000  5.000 6.000 6.0000
## CV residual -0.863 -0.412 0.232 -1.47 -1.66 -0.262 -0.167 0.426 0.0251
##               1318  1336  1341 1343  1344   1353   1355 1360  1361  1365
```

```
## Predicted    6.193  6.23 5.632 5.54 5.632  5.228  5.391 5.75  5.65 5.815
## cvpred       6.175  6.26 5.643 5.55 5.643  5.249  5.417 5.70  5.65 5.842
## QA           6.000  6.00 6.000 6.00 6.000  5.000  5.000 6.00  5.00 6.000
## CV residual -0.175 -0.26 0.357 0.45 0.357 -0.249 -0.417 0.30 -0.65 0.158
##                 1373   1382  1385   1391   1392   1402 1404  1405    1408
## Predicted    4.91014 4.9202 4.836  6.258  5.611 4.9844 6.18 5.579  6.0475
## cvpred       5.00816 4.9475 4.866  6.289  5.631 4.9951 6.19 5.593  6.0585
## QA           5.00000 5.0000 5.000  6.000  5.000 5.0000 8.00 6.000  6.0000
## CV residual -0.00816 0.0525 0.134 -0.289 -0.631 0.0049 1.81 0.407 -0.0585
##                1412   1415  1426 1434   1439   1443   1444  1445   1449
## Predicted   5.759  5.763 5.744 5.76  5.455  5.444  5.881 5.632  5.476
## cvpred      5.783  5.737 5.731 5.78  5.471  5.465  5.889 5.649  5.479
## QA          6.000  5.000 6.000 7.00  5.000  5.000  5.000 6.000  5.000
## CV residual 0.217 -0.737 0.269 1.22 -0.471 -0.465 -0.889 0.351 -0.479
##                1459  1464   1474   1475  1479  1483   1484   1487    1496
## Predicted    5.938 5.545  5.629  5.272  5.11  5.33  6.002  5.410  6.0135
## cvpred       5.919 5.573  5.656  5.401  5.19  5.35  5.977  5.431  6.0301
## QA           5.000 6.000  5.000  5.000  3.00  4.00  5.000  5.000  6.0000
## CV residual -0.919 0.427 -0.656 -0.401 -2.19 -1.35 -0.977 -0.431 -0.0301
##                1498   1502  1508  1513  1517  1518   1520  1527   1532  1545
## Predicted   5.772 4.9325 5.884 5.393  6.09 5.844  5.264 5.534  5.357 6.271
## cvpred      5.798 4.9666 5.898 5.422  6.12 5.853  5.306 5.548  5.401 6.258
## QA          6.000 5.0000 6.000 6.000  5.00 6.000  5.000 6.000  5.000 7.000
## CV residual 0.202 0.0334 0.102 0.578 -1.12 0.147 -0.306 0.452 -0.401 0.742
##                1553  1555 1556  1558   1564  1566   1568  1580   1586  1588
## Predicted   5.683 5.423 5.73 5.423  5.333 5.846  5.333  5.88  6.253 5.692
## cvpred      5.732 5.453 5.74 5.453  5.348 5.857  5.348  5.89  6.257 5.716
## QA          6.000 6.000 7.00 6.000  5.000 6.000  5.000  5.00  6.000 6.000
## CV residual 0.268 0.547 1.26 0.547 -0.348 0.143 -0.348 -0.89 -0.257 0.284
##                1591  1592  1594   1595   1596
## Predicted    6.167 5.482 5.486  5.525 5.9465
## cvpred       6.179 5.556 5.493  5.542 5.9581
## QA           6.000 6.000 6.000  5.000 6.0000
## CV residual -0.179 0.444 0.507 -0.542 0.0419
##
## Sum of squares = 155    Mean square = 0.48    n = 320
##
## Overall (Sum over all 320 folds)
##    ms
## 0.426
```

```
attr(validation,'ms')
```

```
## [1] 0.426
```

```
#Q8
#Mr. Klabjan is informed that the attribute picked in Problem 6 actually contains outliers.
#Calculate the average and standard deviation of the selected attribute.
#Create a new data set after removing observa- tions that is outside of the
#range and name the data set as redwine2.
#Print out the dimension of redwine2 to know how many observations are removed.

head(imp_dt$PH)
```

```
## [1] 3.51 3.20 3.26 3.16 3.51 3.51
```

```r
PH_avg<-mean(imp_dt$PH)
head(PH_avg)
```

```
## [1] 3.31
```

```r
PH_std<-sd(imp_dt$PH)
head(PH_std)
```

```
## [1] 0.392
```

```r
up<-PH_avg+3*PH_std
up
```

```
## [1] 4.48
```

```r
down<-PH_avg-3*PH_std
down
```

```
## [1] 2.13
```

```r
redwine2<-subset(imp_dt,imp_dt$PH<up & imp_dt$PH>down)
remove<-length(imp_dt$PH)-length(redwine2$PH)
remove
```

```
## [1] 19
```

```r
#Q9
#Build regression model winemodel2 using the new data set from Problem 8
#and print out the summary. Compare this model with the model obtained in
#Problem 6 and decide which one is better. Pick 5 attributes that
#is most likely to be related to QA based on p-values.

new_md<-lm(QA~.,data=redwine2)
summary(new_md)
```

```
##
## Call:
## lm(formula = QA ~ ., data = redwine2)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.6893 -0.3634 -0.0437  0.4522  2.0127
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.90e+01   2.12e+01    0.90    0.370
## FA           2.46e-02   2.60e-02    0.95    0.344
## VA          -1.07e+00   1.22e-01   -8.79  < 2e-16 ***
## CA          -1.78e-01   1.48e-01   -1.20    0.230
## RS           1.30e-02   1.50e-02    0.87    0.387
## CH          -1.90e+00   4.21e-01   -4.52  6.6e-06 ***
## FS           4.42e-03   2.18e-03    2.03    0.043 *
## SD          -3.14e-03   7.38e-04   -4.26  2.2e-05 ***
## DE          -1.50e+01   2.17e+01   -0.69    0.489
## PH          -4.25e-01   1.93e-01   -2.20    0.028 *
## SU           9.13e-01   1.15e-01    7.95  3.5e-15 ***
## AL           2.83e-01   2.66e-02   10.65  < 2e-16 ***
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.648 on 1568 degrees of freedom
## Multiple R-squared:  0.363,  Adjusted R-squared:  0.358
## F-statistic: 81.2 on 11 and 1568 DF,  p-value: <2e-16
```

```
#Since the R^2 of the new model is 0.363 and the adjusted R^2 is 0.358, the old model's value is
#0.358 and 0.354, both R^2 and R^2 adjusted of new model is higher than old model. The new
#model is better.

#VA,SU,AL,SD,CH the vlause most related with the model.
```