Varun_MSiA400_Assignment2

```
library(tidyverse)
```

```
## Loading tidyverse: ggplot2
## Loading tidyverse: tibble
## Loading tidyverse: tidyr
## Loading tidyverse: readr
## Loading tidyverse: purrr
## Loading tidyverse: dplyr
```

```
## Conflicts with tidy packages ----------------------------------------------
```

```
## filter(): dplyr, stats
## lag():    dplyr, stats
```

```
redwine <- read.csv("redwine.txt", sep = "\t", header = T)
```

Problem 1

Mean of RS disregarding missing values

```
mean(redwine$RS, na.rm = T)
```

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

Similarly, mean of SD disregarding missing values

```
mean(redwine$SD, na.rm = T)
```

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

Problem 2

```
SD.obs <- redwine$SD[!is.na(redwine$SD)]
FS.obs <- redwine$FS[!is.na(redwine$SD)]
ABC <- lm(SD.obs ~ FS.obs)
coefficients(ABC)
```

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

The model is therefore $\hat{SD.obs} = 13.185505 + 2.086077$ FS.obs

```
summary(ABC)
```

```
##
## Call:
## lm(formula = SD.obs ~ FS.obs)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -54.489 -13.530  -7.155   7.252 197.587
##
## Coefficients:
```

```
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 13.18551    1.11502   11.82   <2e-16 ***
## FS.obs       2.08608    0.05867   35.56   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 24.39 on 1580 degrees of freedom
## Multiple R-squared:  0.4445, Adjusted R-squared:  0.4441
## F-statistic:  1264 on 1 and 1580 DF,  p-value: < 2.2e-16
```

A quick look at the summary of this model shows us that the coefficients are highly significant, indicating a definite linear relationship.

Problem 3

```
SD.predicted <- predict(ABC, newdata = data.frame(FS.obs = redwine$FS[is.na(redwine$SD)]))
redwine$SD[is.na(redwine$SD)] <- SD.predicted
mean(redwine$SD)
```

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

The mean of SD after imputation is 46.30182

Problem 4

```
avg.imp <- function(a,  avg){
missing <- is.na(a)
n.missing <- sum(missing)
a.obs <- a[!missing]
imputed <- a
imputed[missing]<- avg
return(imputed)
}
rsavg <- mean(redwine$RS[!is.na(redwine$RS)])
redwine$RS <- avg.imp(redwine$RS, rsavg)
mean(redwine$RS)
```

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

Obviously the mean remains unchanged, at 2.53792

Problem 5

```
fullmodel <- lm(QA~., data = redwine)
coefficients(fullmodel)
```

```
##   (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
```

The coefficients of the full linear model are seen above.

Problem 6

```
summary(fullmodel)
```

```
##
## Call:
## lm(formula = QA ~ ., data = redwine)
##
## 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
```

Based on the p-values, the attribute PH is least likely to be related to QA given its large 0.414 p-value.

Problem 7

```
library(DAAG)
```

```
## Warning: package 'DAAG' was built under R version 3.4.2
```

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

```
validation <- CVlm(m = 5, form.lm = formula(QA ~ .), data = redwine)
```

```
## Warning in predict.lm(subs.lm, newdata = data[rows.out, ]): prediction from
## a rank-deficient fit may be misleading
```

```
## Warning in predict.lm(subs.lm, newdata = data[rows.out, ]): prediction from
## a rank-deficient fit may be misleading
```

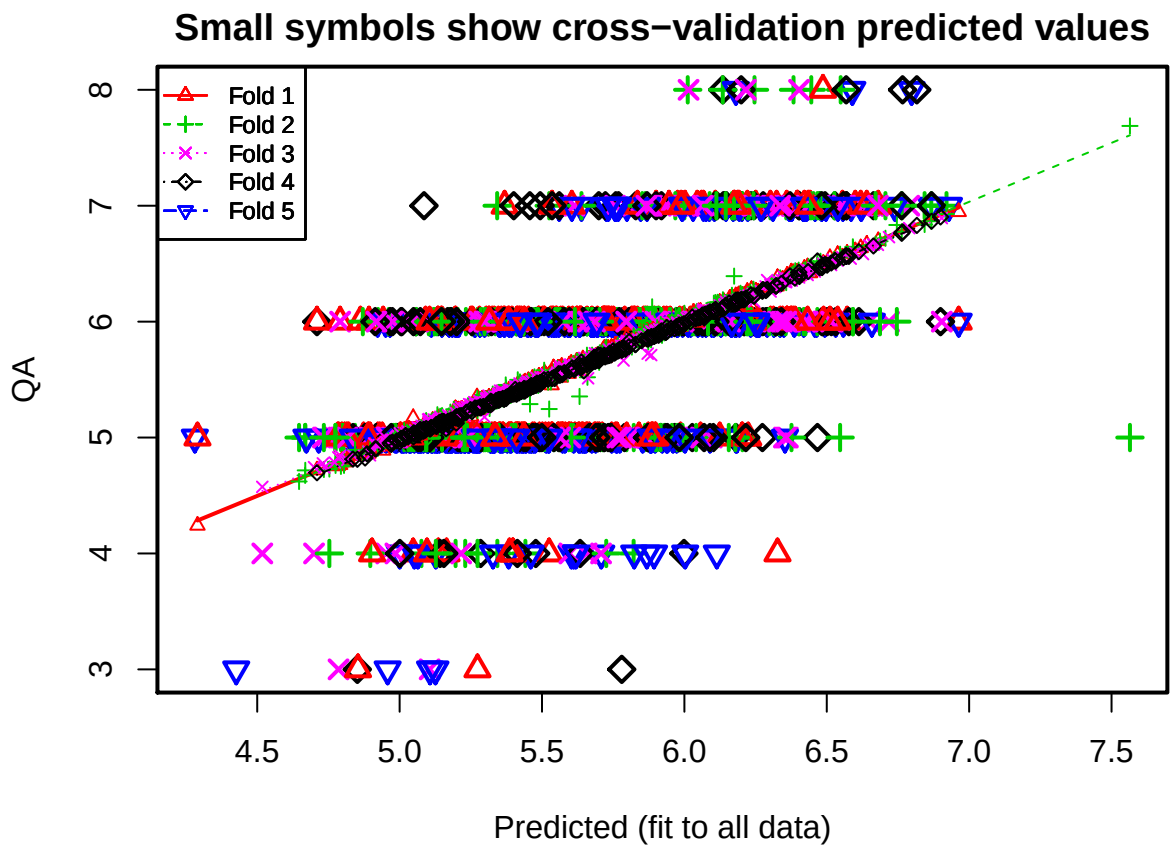
```
## Warning in predict.lm(subs.lm, newdata = data[rows.out, ]): prediction from
## a rank-deficient fit may be misleading

## Warning in predict.lm(subs.lm, newdata = data[rows.out, ]): prediction from
## a rank-deficient fit may be misleading

## Warning in predict.lm(subs.lm, newdata = data[rows.out, ]): prediction from
## a rank-deficient fit may be misleading

## 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(m = 5, form.lm = formula(QA ~ .), data = redwine):
##
##  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    37
## Predicted   5.34  5.67  5.190  5.08  5.445  5.053 5.204  5.235 5.257 5.600
## cvpred      5.35  5.64  5.244  5.05  5.453  5.045 5.159  5.217 5.231 5.592
## QA          7.00  5.00  5.000  4.00  5.000  5.000 6.000  5.000 6.000 6.000
## CV residual 1.65 -0.64 -0.244 -1.05 -0.453 -0.045 0.841 -0.217 0.769 0.408
```

```
##                43    47     49    52     59   63      64   71     76    78
## Predicted   5.543 4.647  5.371 5.427  5.345 5.53  5.0947 5.23  5.649 5.288
## cvpred      5.544 4.621  5.349 5.406  5.351 5.25  5.0897 5.24  5.632 5.275
## QA          6.000 5.000  5.000 6.000  5.000 7.00  5.0000 6.00  5.000 6.000
## CV residual 0.456 0.379 -0.349 0.594 -0.351 1.75 -0.0897 0.76 -0.632 0.725
##                80      83    88       99   102     108     123    124
## Predicted    5.44  5.0928  5.47  5.0526 5.611  5.1415 5.01801  5.112
## cvpred       5.46  5.0791  5.46  5.0184 5.611  5.0957 4.99358  5.113
## QA           4.00  5.0000  5.00  5.0000 6.000  5.0000 5.00000  5.000
## CV residual -1.46 -0.0791 -0.46 -0.0184 0.389 -0.0957 0.00642 -0.113
##                126  128   131  135   140    153    155    156     164
## Predicted    5.133 4.80 4.804 4.92  5.09  5.283  5.606  5.594  5.0311
## cvpred       5.197 4.76 4.807 4.91  5.12  5.328  5.597  5.585  5.0232
## QA           5.000 5.00 5.000 6.00  5.00  5.000  5.000  5.000  5.0000
## CV residual -0.197 0.24 0.193 1.09 -0.12 -0.328 -0.597 -0.585 -0.0232
##                167    175   180    183    191    209   213   215    219
## Predicted    5.151  5.363  5.32 4.9894  5.129  5.309 5.873 5.254  5.384
## cvpred       5.173  5.373  5.34 4.9805  5.158  5.331 5.895 5.249  5.388
## QA           5.000  5.000  5.00 5.0000  5.000  5.000 6.000 6.000  5.000
## CV residual -0.173 -0.373 -0.34 0.0195 -0.158 -0.331 0.105 0.751 -0.388
##                225   229    230   232   237   239  250   252     259    261
## Predicted    5.34 5.703  5.563 5.554 5.058 5.058 5.59 5.456  5.0429  5.575
## cvpred       5.35 5.731  5.557 5.595 5.059 5.059 5.55 5.454  5.0155  5.598
## QA           4.00 6.000  5.000 6.000 6.000 6.000 6.00 6.000  5.0000  5.000
## CV residual -1.35 0.269 -0.557 0.405 0.941 0.941 0.45 0.546 -0.0155 -0.598
##                262    263   266    272    273  277    283   294    297
## Predicted    4.753  5.542 6.197  6.314  5.861 5.53  5.211 5.642  5.089
## cvpred       4.728  5.568 6.244  6.334  5.887 5.48  5.211 5.627  5.129
## QA           4.000  5.000 7.000  6.000  5.000 6.00  5.000 6.000  5.000
## CV residual -0.728 -0.568 0.756 -0.334 -0.887 0.52 -0.211 0.373 -0.129
##                300    302    303   308   312   317    323   327   330
## Predicted    5.214  6.109  5.342 5.336 5.119  5.42  5.261 6.107  5.61
## cvpred       5.176  6.164  5.328 5.369 5.148  5.48  5.255 6.128  5.64
## QA           5.000  6.000  5.000 6.000 6.000  5.00  5.000 7.000  5.00
## CV residual -0.176 -0.164 -0.328 0.631 0.852 -0.48 -0.255 0.872 -0.64
##                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       6.392  6.645 5.776 5.99  5.0752  5.488  6.08  5.375 6.866
## QA           6.000  6.000 6.000 7.00  5.0000  5.000  5.00  5.000 7.000
## CV residual -0.392 -0.645 0.224 1.01 -0.0752 -0.488 -1.08 -0.375 0.134
##               372   376   388  391    409   414    415    416    423   431
## Predicted   5.75 6.223 5.190 6.01  6.110 6.358  5.142  5.208  5.111 6.436
## cvpred      5.73 6.232 5.206 5.99  6.133 6.337  5.183  5.187  5.105 6.473
## QA          6.00 7.000 6.000 8.00  6.000 7.000  5.000  5.000  5.000 7.000
## CV residual 0.27 0.768 0.794 2.01 -0.133 0.663 -0.183 -0.187 -0.105 0.527
##                434  439    442   446   447   454  456   464   465    470
## Predicted    5.473 5.85  6.135 5.273  5.97 6.105 6.38 4.668 5.722  5.284
## cvpred       5.523 5.87  6.189 5.301  6.05 6.123 6.41 4.674 5.766  5.328
## QA           5.000 6.00  6.000 6.000  5.00 7.000 8.00 5.000 6.000  5.000
## CV residual -0.523 0.13 -0.189 0.699 -1.05 0.877 1.59 0.326 0.234 -0.328
##                477  482   490   493   498  499   529    531    534    536
## Predicted    5.797 6.55 5.853 6.920  5.96 6.13 5.317  6.146  6.746  6.109
## cvpred       5.813 6.55 5.858 6.929  6.00 6.17 5.309  6.161  6.834  6.128
## QA           5.000 8.00 6.000 7.000  5.00 8.00 6.000  6.000  6.000  6.000
```

```
## CV residual -0.813 1.45 0.142 0.071 -1.00 1.83 0.691 -0.161 -0.834 -0.128
##                541    543    548   556    561  564   573    582   584
## Predicted    5.358  5.363  6.232  5.89  5.917 5.66  5.99  5.374 6.000
## cvpred       5.375  5.373  6.253  6.13  5.975 5.67  6.03  5.428 6.072
## QA           5.000  5.000  6.000  5.00  5.000 6.00  5.00  5.000 7.000
## CV residual -0.375 -0.373 -0.253 -1.13 -0.975 0.33 -1.03 -0.428 0.928
##                598   606   608    611    624    626     627   634   637
## Predicted   5.626 5.175 5.561  5.464  6.356  5.284  5.0355  5.20 4.669
## cvpred      5.656 5.173 5.586  5.487  6.352  5.259  5.0305  5.22 4.717
## QA          6.000 6.000 6.000  5.000  6.000  5.000  5.0000  4.00 5.000
## CV residual 0.344 0.827 0.414 -0.487 -0.352 -0.259 -0.0305 -1.22 0.283
##                643   653  658   659    664  667  670    674   675    681
## Predicted    5.458  7.56 5.86 5.549  6.125 5.13 5.66  5.109 5.622  5.373
## cvpred       5.289  7.69 5.94 5.538  6.156 5.14 5.71  5.102 5.658  5.454
## QA           5.000  5.00 7.00 6.000  6.000 6.00 6.00  5.000 6.000  5.000
## CV residual -0.289 -2.69 1.06 0.462 -0.156 0.86 0.29 -0.102 0.342 -0.454
##                683    694    699   703   706     713    719   725    728
## Predicted    5.426  5.178  5.115 5.285 4.794  5.0257  5.240  5.23  5.363
## cvpred       5.463  5.219  5.145 5.261 4.751  5.0459  5.276  5.17  5.333
## QA           5.000  5.000  5.000 6.000 5.000  5.0000  5.000  4.00  5.000
## CV residual -0.463 -0.219 -0.145 0.739 0.249 -0.0459 -0.276 -1.17 -0.333
##                741    744   746    754   766   769   780     783    784
## Predicted   5.816  5.415 5.632  5.172 5.123 5.103  5.28  5.0161  5.390
## cvpred      5.781  5.498 5.355  5.185 5.147 5.125  5.30  5.0606  5.364
## QA          6.000  5.000 6.000  5.000 6.000 6.000  5.00  5.0000  5.000
## CV residual 0.219 -0.498 0.645 -0.185 0.853 0.875 -0.30 -0.0606 -0.364
##                799   806   810   816   822    828  829   830   834   837
## Predicted   5.67 6.763 5.571  5.97 6.844  5.634 6.45 5.921  5.82 6.167
## cvpred      5.68 6.775 5.563  6.03 6.838  5.649 6.48 5.882  5.88 6.227
## QA          6.00 7.000 6.000  5.00 7.000  5.000 8.00 6.000  4.00 7.000
## CV residual 0.32 0.225 0.437 -1.03 0.162 -0.649 1.52 0.118 -1.88 0.773
##                848    855    857   858      860    867   870    871     880
## Predicted   5.304  6.117  6.117 6.302  6.02362  6.146 5.714 5.9213  5.0740
## cvpred      5.283  6.166  6.166 6.322  6.00175  6.124 5.736 5.9533  5.0853
## QA          6.000  6.000  6.000 7.000  6.00000  6.000 6.000 6.0000  5.0000
## CV residual 0.717 -0.166 -0.166 0.678 -0.00175 -0.124 0.264 0.0467 -0.0853
##                882    883    891   901  903   918   925   930    932   933
## Predicted   5.836  6.333  5.713  6.38 5.64 5.701  6.20 6.526  5.363 5.410
## cvpred      5.853  6.368  5.736  6.40 5.63 5.649  6.22 6.538  5.363 5.429
## QA          6.000  6.000  5.000  5.00 7.00 6.000  5.00 7.000  5.000 6.000
## CV residual 0.147 -0.368 -0.736 -1.40 1.37 0.351 -1.22 0.462 -0.363 0.571
##                936    940  953    955   956    958    972   973   978
## Predicted    6.336  5.929 6.23  6.144  5.99 5.9240  6.419 6.203 4.844
## cvpred       6.378  5.903 6.23  6.138  6.00 5.9588  6.439 6.233 4.894
## QA           6.000  5.000 7.00  6.000  5.00 6.0000  6.000 7.000 5.000
## CV residual -0.378 -0.903 0.77 -0.138 -1.00 0.0412 -0.439 0.767 0.106
##                1000  1001  1006   1010   1016   1020   1027  1034  1037
## Predicted    6.198 6.188 6.453  5.822  6.401  5.616  6.467 5.582 6.478
## cvpred       6.197 6.173 6.458  5.868  6.444  5.606  6.503 5.592 6.521
## QA           6.000 7.000 7.000  5.000  6.000  5.000  6.000 6.000 7.000
## CV residual -0.197 0.827 0.542 -0.868 -0.444 -0.606 -0.503 0.408 0.479
##               1038   1051    1061  1064  1065  1066   1081 1082   1093
## Predicted   4.77  5.696  6.0443  6.38 5.883 5.473  6.389 5.85 5.9210
## cvpred      4.76  5.713  6.0879  6.41 5.892 5.465  6.439 5.92 5.9244
```

```
## QA          5.00  5.000  6.0000  6.00 6.000 6.000  6.000 7.00 6.0000
## CV residual 0.24 -0.713 -0.0879 -0.41 0.108 0.535 -0.439 1.08 0.0756
##               1104  1106   1109  1112  1116   1119  1120   1127  1131
## Predicted    6.178  6.16 4.9778 6.089 5.676  6.592  6.18  6.687 5.676
## cvpred       6.185  6.15 4.9574 6.071 5.673  6.636  6.16  6.694 5.711
## QA           6.000  5.00 5.0000 7.000 6.000  6.000  5.00  6.000 6.000
## CV residual -0.185 -1.15 0.0426 0.929 0.327 -0.636 -1.16 -0.694 0.289
##               1132  1135  1151   1159   1160   1165   1173   1179   1188
## Predicted    5.508 6.467 6.706  6.219  5.905  5.349  6.425  5.521 5.9498
## cvpred       5.581 6.489 6.719  6.169  5.949  5.343  6.449  5.477 5.9647
## QA           5.000 7.000 7.000  6.000  5.000  5.000  6.000  5.000 6.0000
## CV residual -0.581 0.511 0.281 -0.169 -0.949 -0.343 -0.449 -0.477 0.0353
##               1190    1192 1195  1199 1203  1211  1213    1214   1225
## Predicted    4.897  5.1145 4.96  6.17 6.25 5.548 5.548  6.0334  6.103
## cvpred       4.895  5.0821 4.96  6.21 6.25 5.535 5.535  6.0516  6.159
## QA           4.000  5.0000 6.00  6.00 8.00 6.000 6.000  6.0000  6.000
## CV residual -0.895 -0.0821 1.04 -0.21 1.75 0.465 0.465 -0.0516 -0.159
##              1233  1234   1248    1261  1262  1267  1292   1296   1301
## Predicted    5.39  5.73  5.610  5.1100  5.27 5.554 5.767  5.150 5.9818
## cvpred       5.40  5.78  5.608  5.0885  5.23 5.561 5.796  5.184 5.9533
## QA           5.00  4.00  5.000  5.0000  4.00 6.000 6.000  5.000 6.0000
## CV residual -0.40 -1.78 -0.608 -0.0885 -1.23 0.439 0.204 -0.184 0.0467
##             1302  1303  1305   1307   1310  1314   1321  1324  1326  1327
## Predicted   5.66  6.30 4.802  5.268  5.149 5.563  5.297 6.243 5.722 5.722
## cvpred      5.52  6.31 4.811  5.303  5.163 5.583  5.323 6.279 5.722 5.722
## QA          6.00  6.00 5.000  5.000  5.000 6.000  5.000 7.000 6.000 6.000
## CV residual 0.48 -0.31 0.189 -0.303 -0.163 0.417 -0.323 0.721 0.278 0.278
##             1330  1332   1335   1338   1339    1349    1362  1366    1377
## Predicted   5.15 4.734 4.9854  5.310  5.310  5.0850  5.0507  5.32  5.0311
## cvpred      5.16 4.724 4.9697  5.314  5.314  5.0853  5.0511  5.34  5.0336
## QA          6.00 5.000 5.0000  5.000  5.000  5.0000  5.0000  5.00  5.0000
## CV residual 0.84 0.276 0.0303 -0.314 -0.314 -0.0853 -0.0511 -0.34 -0.0336
##              1380   1384  1386   1388   1390   1393    1395   1398    1401
## Predicted   5.724  5.149 4.774  5.237  5.273  5.332  5.0413  5.117  4.9844
## cvpred      5.736  5.189 4.789  5.235  5.337  5.338  5.0375  5.138  5.0356
## QA          6.000  5.000 5.000  5.000  5.000  5.000  5.0000  5.000  5.0000
## CV residual 0.264 -0.189 0.211 -0.235 -0.337 -0.338 -0.0375 -0.138 -0.0356
##              1406   1419   1421   1422   1428  1430  1435   1438  1441
## Predicted   6.444  5.402  5.402  5.407  5.916  6.55 5.298  5.316 6.174
## cvpred      6.458  5.431  5.431  5.436  5.919  6.58 5.271  5.251 6.181
## QA          7.000  5.000  5.000  5.000  5.000  5.00 6.000  5.000 7.000
## CV residual 0.542 -0.431 -0.431 -0.436 -0.919 -1.58 0.729 -0.251 0.819
##             1446  1453   1455  1456   1458  1463  1476  1486  1490   1493
## Predicted   4.87 6.112 5.9488 5.436  5.116 5.484 6.671  5.22 5.883  5.899
## cvpred      4.86 6.136 5.9839 5.399  5.127 5.481 6.669  5.20 5.844  5.881
## QA          6.00 7.000 6.0000 6.000  5.000 6.000 7.000  5.00 6.000  5.000
## CV residual 1.14 0.864 0.0161 0.601 -0.127 0.519 0.331 -0.20 0.156 -0.881
##              1504  1507  1509   1512  1521  1522   1524  1529   1534  1537
## Predicted   5.873 5.486 5.895  5.348 5.844  5.13  5.603 5.702  5.302 5.615
## cvpred      5.824 5.449 5.874  5.334 5.841  5.12  5.617 5.716  5.332 5.601
## QA          6.000 6.000 6.000  5.000 6.000  4.00  5.000 6.000  5.000 6.000
## CV residual 0.176 0.551 0.126 -0.334 0.159 -1.12 -0.617 0.284 -0.332 0.399
##              1541  1542  1544    1551   1552   1561   1562  1567   1569
## Predicted   5.919 6.145 5.868  5.0981  5.107  5.093  5.093  6.23  5.228
```

```
## cvpred      5.892 6.161 5.915  5.0934  5.103  5.144  5.144  6.23  5.234
## QA          6.000 7.000 6.000  5.0000  5.000  5.000  5.000  6.00  5.000
## CV residual 0.108 0.839 0.085 -0.0934 -0.103 -0.144 -0.144 -0.23 -0.234
##                1577   1581  1593
## Predicted    6.0829  6.223 5.898
## cvpred       6.0745  6.232 5.896
## QA           6.0000  6.000 6.000
## CV residual -0.0745 -0.232 0.104
##
## 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 5.406  5.375  4.518  5.19
## cvpred       5.0372  5.983 5.434 5.597  5.515 5.436  5.384  4.574  5.18
## QA           5.0000  5.000 6.000 6.000  5.000 6.000  5.000  4.000  5.00
## CV residual -0.0372 -0.983 0.566 0.403 -0.515 0.564 -0.384 -0.574 -0.18
##                 51    53     54    60     67     77      79     82    89
## Predicted    5.304 5.457  5.306 5.369  5.332  5.649  5.0217  5.262  5.87
## cvpred       5.327 5.459  5.271 5.362  5.352  5.687  5.0374  5.305  5.85
## QA           5.000 6.000  5.000 6.000  5.000  5.000  5.0000  5.000  5.00
## CV residual -0.327 0.541 -0.271 0.638 -0.352 -0.687 -0.0374 -0.305 -0.85
##                  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       5.0362  6.432  5.319  5.304 5.581 4.741  5.218 5.528  5.282
## QA           5.0000  6.000  5.000  5.000 6.000 5.000  5.000 6.000  5.000
## CV residual -0.0362 -0.432 -0.319 -0.304 0.419 0.259 -0.218 0.472 -0.282
##                146    159   163    171    177   178    196    202    217
## Predicted   4.838  5.077 5.503  4.699  5.363 5.515 4.9792  5.608  5.599
## cvpred      4.842  5.102 5.519  4.745  5.362 5.547 4.9742  5.581  5.618
## QA          5.000  5.000 6.000  4.000  5.000 6.000 5.0000  5.000  5.000
## CV residual 0.158 -0.102 0.481 -0.745 -0.362 0.453 0.0258 -0.581 -0.618
##                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       5.375 5.699 5.97 5.587  5.592  5.276 5.9885  5.846 6.043
## QA           5.000 6.000 7.00 6.000  5.000  5.000 6.0000  5.000 7.000
## CV residual -0.375 0.301 1.03 0.413 -0.592 -0.276 0.0115 -0.846 0.957
##                265   274    275    278  280    285    299   309   313  318
## Predicted    6.07  5.17  5.228  6.314 5.89  5.440  5.313 5.421 5.294 5.37
## cvpred       6.10  5.19  5.188  6.299 5.87  5.454  5.362 5.467 5.284 5.39
## QA           5.00  5.00  5.000  6.000 7.00  5.000  5.000 6.000 6.000 6.00
## CV residual -1.10 -0.19 -0.188 -0.299 1.13 -0.454 -0.362 0.533 0.716 0.61
##                321    322   324   329   333   340    346   347   354   355
## Predicted   5.80  5.197 5.380 5.925 5.040 6.554  5.449 6.065  6.35  6.30
## cvpred      5.83  5.211 5.392 5.931 5.021 6.555  5.484 6.103  6.39  6.26
## QA          7.00  5.000 6.000 6.000 6.000 7.000  5.000 7.000  5.00  6.00
## CV residual 1.17 -0.211 0.608 0.069 0.979 0.445 -0.484 0.897 -1.39 -0.26
##                379   381   383   394    397     400    402    407  408
## Predicted    6.719 5.787 5.787 4.772 5.0174  5.0886 5.9429 5.8874 5.94
## cvpred       6.729 5.791 5.791 4.784 4.9809  5.0975 5.9363 5.9107 5.96
## QA           6.000 6.000 6.000 5.000 5.0000  5.0000 6.0000 6.0000 7.00
## CV residual -0.729 0.209 0.209 0.216 0.0191 -0.0975 0.0637 0.0893 1.04
##                424   430    433  441   452  455   460    461   463    466
```

```
## Predicted   6.436 5.353  6.649 6.01 5.244  6.1  5.11 5.9546  6.36  5.911
## cvpred      6.444 5.405  6.658 6.02 5.288  6.1  5.15 5.9894  6.37  5.935
## QA          7.000 6.000  6.000 8.00 6.000  5.0  3.00 6.0000  5.00  5.000
## CV residual 0.556 0.595 -0.658 1.98 0.712 -1.1 -2.15 0.0106 -1.37 -0.935
##                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       6.0451 5.9518  5.403  5.403  5.724  5.892  5.381 6.807 6.354
## QA           6.0000 6.0000  5.000  5.000  5.000  5.000  5.000 7.000 7.000
## CV residual -0.0451 0.0482 -0.403 -0.403 -0.724 -0.892 -0.381 0.193 0.646
##                504   523    525   526     542   545    546     549   552
## Predicted   6.507  5.69  5.303  5.41  6.0481 5.683  5.148  6.0165 5.722
## cvpred      6.515  5.68  5.288  5.44  6.0323 5.694  5.142  6.0156 5.724
## QA          7.000  5.00  5.000  5.00  6.0000 6.000  5.000  6.0000 6.000
## CV residual 0.485 -0.68 -0.288 -0.44 -0.0323 0.306 -0.142 -0.0156 0.276
##                565  568    571    578    583   587   590    593    594
## Predicted    6.291 4.79  6.285  5.215  5.251 6.314 6.140  5.324  5.140
## cvpred       6.358 4.84  6.304  5.205  5.284 6.314 6.143  5.338  5.162
## QA           6.000 6.00  6.000  5.000  5.000 7.000 7.000  5.000  5.000
## CV residual -0.358 1.16 -0.304 -0.205 -0.284 0.686 0.857 -0.338 -0.162
##                597   599    601   605    614   629    630   631   635
## Predicted   5.613 5.290  4.953 5.102  5.756 5.275 4.9731 5.275  5.48
## cvpred      5.625 5.304  4.992 5.108  5.763 5.287 4.9759 5.287  5.46
## QA          6.000 6.000  4.000 6.000  5.000 6.000 5.0000 6.000  5.00
## CV residual 0.375 0.696 -0.992 0.892 -0.763 0.713 0.0241 0.713 -0.46
##                640    641    642    644    645   650    651    652    654
## Predicted    6.558  5.347  5.236  5.236  5.347 5.786  5.408  5.100  6.146
## cvpred       6.549  5.374  5.264  5.264  5.374 5.665  5.423  5.154  6.172
## QA           6.000  5.000  5.000  5.000  5.000 6.000  5.000  5.000  6.000
## CV residual -0.549 -0.374 -0.264 -0.264 -0.374 0.335 -0.423 -0.154 -0.172
##                657    665    666    672    686     687    688   692    693
## Predicted    5.408  5.735  5.346  5.109  5.565  5.0636  5.209 4.731  5.120
## cvpred       5.423  5.759  5.367  5.145  5.612  5.0889  5.245 4.777  5.155
## QA           5.000  5.000  5.000  5.000  5.000  5.0000  5.000 5.000  5.000
## CV residual -0.423 -0.759 -0.367 -0.145 -0.612 -0.0889 -0.245 0.223 -0.155
##                698     700    709     712   715    723    726   737   738
## Predicted   5.282  6.0085 5.9344  5.0307  5.11  5.567  5.495 4.786 5.143
## cvpred      5.309  6.0312 5.9512  5.0423  5.12  5.562  5.519 4.828 5.179
## QA          6.000  6.0000 6.0000  5.0000  5.00  5.000  5.000 5.000 6.000
## CV residual 0.691 -0.0312 0.0488 -0.0423 -0.12 -0.562 -0.519 0.172 0.821
##                750    751    761   765    767    770  771   772    782
## Predicted   5.464  5.172  5.156 5.113  5.132  5.146 5.10 4.891  5.390
## cvpred      5.488  5.189  5.146 5.137  5.124  5.169 5.11 4.879  5.412
## QA          6.000  5.000  5.000 6.000  5.000  5.000 6.00 5.000  5.000
## CV residual 0.512 -0.189 -0.146 0.863 -0.124 -0.169 0.89 0.121 -0.412
##                788   789    794   800   805    813    824   827   831
## Predicted   5.421 5.421  5.700 5.668 5.491  5.975  5.383 6.176  5.60
## cvpred      5.448 5.448  5.719 5.692 5.499  5.989  5.412 6.173  5.63
## QA          6.000 6.000  5.000 6.000 6.000  5.000  5.000 7.000  4.00
## CV residual 0.552 0.552 -0.719 0.308 0.501 -0.989 -0.412 0.827 -1.63
##                835   839   840    843    844  856   859    863     864
## Predicted    5.101 6.360  5.32 5.9013 4.9267 5.93 6.226  5.563  5.0825
## cvpred       5.133 6.369  5.34 5.9041 4.9348 5.96 6.218  5.571  5.0917
## QA           5.000 7.000  5.00 6.0000 5.0000 7.00 7.000  5.000  5.0000
## CV residual -0.133 0.631 -0.34 0.0959 0.0652 1.04 0.782 -0.571 -0.0917
```

```
##                865     866     869   873    878     884   885   888  889
## Predicted    5.0394  5.0678  6.0236  5.66 5.9213  5.0740 5.298 6.420 5.75
## cvpred       5.0428  5.0761  6.0395  5.65 5.9278  5.0776 5.293 6.417 5.75
## QA           5.0000  5.0000  6.0000  4.00 6.0000  5.0000 6.000 7.000 6.00
## CV residual -0.0428 -0.0761 -0.0395 -1.65 0.0722 -0.0776 0.707 0.583 0.25
##                894  895  896   897  905   909    911    927    928   939
## Predicted    5.0694 5.13 5.64 6.600 5.76 5.740  6.487  6.233  4.985 6.538
## cvpred       5.0819 5.14 5.65 6.594 5.73 5.738  6.499  6.224  4.993 6.526
## QA           5.0000 6.00 6.00 7.000 7.00 6.000  6.000  6.000  4.000 7.000
## CV residual -0.0819 0.86 0.35 0.406 1.27 0.262 -0.499 -0.224 -0.993 0.474
##                942  943   946   949   951    961   964    977    980    983
## Predicted    6.584 5.61 6.226 6.664 6.664  6.188  6.24  5.284  5.963  6.521
## cvpred       6.544 5.62 6.227 6.664 6.664  6.189  6.24  5.288  5.975  6.534
## QA           7.000 7.00 7.000 7.000 7.000  6.000  6.00  5.000  5.000  6.000
## CV residual 0.456 1.38 0.773 0.336 0.336 -0.189 -0.24 -0.288 -0.975 -0.534
##                984    985   989    995   996  1004  1008  1014  1029
## Predicted    5.815  5.963  5.42  5.218 5.455 6.453 6.555 5.304 5.673
## cvpred       5.847  5.975  5.42  5.219 5.478 6.462 6.551 5.336 5.659
## QA           6.000  5.000  5.00  5.000 6.000 7.000 7.000 6.000 6.000
## CV residual 0.153 -0.975 -0.42 -0.219 0.522 0.538 0.449 0.664 0.341
##                1033  1039    1040   1041  1044  1047 1056  1059  1060
## Predicted    5.0280 6.627  6.0692  5.103 6.245 5.480 4.95 6.076 6.142
## cvpred       5.0534 6.583  6.0872  5.159 6.175 5.515 4.96 6.102 6.149
## QA           5.0000 7.000  6.0000  5.000 7.000 6.000 6.00 7.000 7.000
## CV residual -0.0534 0.417 -0.0872 -0.159 0.825 0.485 1.04 0.898 0.851
##                1067  1068  1071  1072 1080   1084  1085  1087 1091  1110
## Predicted    6.351 6.429 6.410 4.896 5.88  6.313 5.414 6.268 6.40 5.795
## cvpred       6.357 6.415 6.403 4.847 5.71  6.325 5.418 6.264 6.38 5.818
## QA           7.000 7.000 7.000 5.000 7.00  6.000 6.000 7.000 8.00 6.000
## CV residual 0.643 0.585 0.597 0.153 1.29 -0.325 0.582 0.736 1.62 0.182
##                1111  1126   1129  1141   1149   1150   1153  1161  1162
## Predicted    5.650 6.319  5.600 5.420  6.124  6.302  5.246 6.177 5.857
## cvpred       5.663 6.299  5.594 5.425  6.142  6.317  5.257 6.185 5.868
## QA           6.000 7.000  5.000 6.000  6.000  6.000  5.000 7.000 6.000
## CV residual 0.337 0.701 -0.594 0.575 -0.142 -0.317 -0.257 0.815 0.132
##                1168   1171  1176   1181   1183   1187   1189  1196    1204
## Predicted    6.650 5.9747 5.902  6.127  6.156  5.793  5.406 5.338  5.0955
## cvpred       6.652 5.9898 5.924  6.123  6.156  5.775  5.404 5.348  5.0685
## QA           7.000 6.0000 6.000  6.000  6.000  5.000  5.000 6.000  5.0000
## CV residual 0.348 0.0102 0.076 -0.123 -0.156 -0.775 -0.404 0.652 -0.0685
##                1212  1217   1228   1231    1235  1247    1249   1252   1257
## Predicted    5.380 5.240  5.444  6.515  5.9969  5.15  6.0510  5.355  5.111
## cvpred       5.372 5.243  5.471  6.488  6.0103  5.16  6.0567  5.361  5.107
## QA           5.000 6.000  5.000  6.000  6.0000  5.00  6.0000  5.000  5.000
## CV residual -0.372 0.757 -0.471 -0.488 -0.0103 -0.16 -0.0567 -0.361 -0.107
##                1260   1265  1266  1268  1269   1271 1279  1286   1291   1306
## Predicted    5.695  6.350 5.554  6.41 5.344  6.904 4.97  5.86  5.577  5.161
## cvpred       5.713  6.346 5.561  6.41 5.354  6.893 4.98  5.87  5.608  5.165
## QA           6.000  6.000 6.000  6.00 6.000  6.000 6.00  5.00  5.000  5.000
## CV residual 0.287 -0.346 0.439 -0.41 0.646 -0.893 1.02 -0.87 -0.608 -0.165
##                1325  1333   1337   1347   1351   1354  1358  1359 1363
## Predicted    5.722 5.488  5.310  5.806  5.420  5.228 5.799 4.985 5.75
## cvpred       5.732 5.483  5.312  5.829  5.424  5.253 5.809 4.958 5.76
## QA           6.000 6.000  5.000  5.000  5.000  5.000 6.000 5.000 6.00
```

```
## CV residual 0.268 0.517 -0.312 -0.829 -0.424 -0.253 0.191 0.042 0.24
##               1371   1372  1374  1375   1376  1396   1397  1400   1403
## Predicted   4.9101  6.160 4.816  4.79  5.134 5.279  5.307 5.799  6.426
## cvpred      4.9482  6.177 4.807  4.83  5.129 5.316  5.343 5.822  6.449
## QA          5.0000  6.000 5.000  3.00  5.000 6.000  5.000 6.000  6.000
## CV residual 0.0518 -0.177 0.193 -1.83 -0.129 0.684 -0.343 0.178 -0.449
##               1407  1424  1432   1433  1436    1437  1452  1462  1472
## Predicted    6.433  5.71 5.417  6.329 5.298  5.0553 6.064  5.22  6.06
## cvpred       6.402  5.72 5.406  6.345 5.177  5.0629 6.087  5.26  6.06
## QA           6.000  4.00 6.000  6.000 6.000  5.0000 7.000  4.00  5.00
## CV residual -0.402 -1.72 0.594 -0.345 0.823 -0.0629 0.913 -1.26 -1.06
##              1478 1495   1503   1505  1510  1514 1516   1519    1531  1533
## Predicted   6.680 5.87  5.240 5.8842  6.36 5.573 4.92  5.600  6.0891 5.627
## cvpred      6.662 5.86  5.246 5.9157  6.34 5.555 4.94  5.629  6.0922 5.636
## QA          7.000 7.00  5.000 6.0000  5.00 6.000 6.00  5.000  6.0000 6.000
## CV residual 0.338 1.14 -0.246 0.0843 -1.34 0.445 1.06 -0.629 -0.0922 0.364
##               1547   1549 1550   1570   1571    1572   1573  1575  1578
## Predicted    5.564  5.765 6.22 5.9148  6.342  6.0564 4.9556 5.662 5.796
## cvpred       5.578  5.774 6.22 5.9241  6.359  6.0615 4.9593 5.511 5.807
## QA           5.000  5.000 8.00 6.0000  6.000  6.0000 5.0000 6.000 6.000
## CV residual -0.578 -0.774 1.78 0.0759 -0.359 -0.0615 0.0407 0.489 0.193
##              1579   1583 1585   1587   1597   1599
## Predicted   5.686  5.777 6.34  6.358 5.8981 5.9652
## cvpred      5.727  5.793 6.34  6.347 5.9161 5.9569
## QA          6.000  5.000 7.00  6.000 6.0000 6.0000
## CV residual 0.273 -0.793 0.66 -0.347 0.0839 0.0431
##
## 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    5.200  5.378 5.374  5.750  6.03  5.687 4.9179  5.130  6.47
## cvpred       5.181  5.352 5.347  5.748  6.04  5.669 4.9028  5.108  6.52
## QA           5.000  5.000 6.000  5.000  5.00  5.000 5.0000  5.000  5.00
## CV residual -0.181 -0.352 0.653 -0.748 -1.04 -0.669 0.0972 -0.108 -1.52
##                 96   103    104    111   114   116  129   133    136
## Predicted   5.9382 5.200  5.076  5.558 5.736 5.766 5.92  6.18  5.321
## cvpred      5.9504 5.172  5.056  5.551 5.708 5.746 5.92  6.19  5.313
## QA          6.0000 6.000  5.000  5.000 6.000 6.000 7.00  5.00  5.000
## CV residual 0.0496 0.828 -0.056 -0.551 0.292 0.254 1.08 -1.19 -0.313
##                137    138    141     147    154    157  160     166  172
## Predicted    5.265  5.324  5.321  5.0836  5.283  5.606 5.01  5.0210 5.44
## cvpred       5.255  5.305  5.313  5.0708  5.264  5.601 4.99  5.0179 5.41
## QA           5.000  5.000  5.000  5.0000  5.000  5.000 6.00  5.0000 6.00
## CV residual -0.255 -0.305 -0.313 -0.0708 -0.264 -0.601 1.01 -0.0179 0.59
##              173  185     190    194    195   200   207   210   212    218
## Predicted   5.44 5.00  5.0211  5.346  5.346  5.14 6.202 6.277 5.151 5.0031
## cvpred      5.41 4.99  5.0154  5.324  5.324  5.13 6.188 6.274 5.138 4.9853
## QA          6.00 6.00  5.0000  5.000  5.000  4.00 7.000 7.000 6.000 5.0000
## CV residual 0.59 1.01 -0.0154 -0.324 -0.324 -1.13 0.812 0.726 0.862 0.0147
##                 220   221  235   236    242   243   244   245   255
## Predicted    5.0715 5.479 4.71 5.058  6.055 5.135 6.179 6.179 5.456
## cvpred       5.0678 5.453 4.70 5.026  6.036 5.108 6.158 6.158 5.432
```

```
## QA           5.0000 6.000 6.00 6.000  6.000 6.000 7.000 7.000 6.000
## CV residual -0.0678 0.547 1.30 0.974 -0.036 0.892 0.842 0.842 0.568
##                 256   271   276  282    286   287   288  291   295   306
## Predicted    5.0498 5.647 5.755 5.70  5.440 5.884 5.793 5.90 5.827 5.311
## cvpred       5.0411 5.575 5.728 5.67  5.401 5.862 5.761 5.88 5.801 5.286
## QA           5.0000 6.000 6.000 7.00  5.000 6.000 6.000 7.00 6.000 6.000
## CV residual -0.0411 0.425 0.272 1.33 -0.401 0.138 0.239 1.12 0.199 0.714
##                 307   311   328    332   336    337    341   343   344
## Predicted    5.0489 5.311  6.27  6.286 6.144  6.486  6.188 5.784 5.784
## cvpred       5.0339 5.286  6.26  6.268 6.133  6.485  6.187 5.764 5.764
## QA           5.0000 6.000  5.00  6.000 7.000  6.000  6.000 6.000 6.000
## CV residual -0.0339 0.714 -1.26 -0.268 0.867 -0.485 -0.187 0.236 0.236
##                 353    356   360    366  367    371   382   384  390   392
## Predicted    5.294 5.9911 5.688  6.403 5.56  5.339 5.894 5.787 5.72 5.894
## cvpred       5.268 5.9728 5.665  6.397 5.54  5.319 5.873 5.768 5.70 5.873
## QA           5.000 6.0000 6.000  6.000 7.00  5.000 6.000 6.000 7.00 6.000
## CV residual -0.268 0.0272 0.335 -0.397 1.46 -0.319 0.127 0.232 1.30 0.127
##                 395   403   406   411    419  422  428     432   435   437
## Predicted    5.494 5.692 5.953 5.362 5.9358 5.76 5.31  5.1109 5.850 5.410
## cvpred       5.475 5.668 5.943 5.338 5.9122 5.74 5.29  5.0875 5.835 5.386
## QA           5.000 6.000 6.000 6.000 6.0000 7.00 6.00  5.0000 6.000 6.000
## CV residual -0.475 0.332 0.057 0.662 0.0878 1.26 0.71 -0.0875 0.165 0.614
##                443   450   453     467    468    471    472   480    486
## Predicted   5.77 5.909 5.435  6.0503  6.900  5.965 5.9817 5.478  5.349
## cvpred      5.75 5.882 5.405  6.0331  6.903  5.939 5.9652 5.457  5.327
## QA          7.00 6.000 6.000  6.0000  6.000  5.000 6.0000 6.000  5.000
## CV residual 1.25 0.118 0.595 -0.0331 -0.903 -0.939 0.0348 0.543 -0.327
##                488   491  496   497   505   506   507   512    517   518
## Predicted   5.199 5.436 6.13 5.195 6.529 6.434 6.494 5.449  6.132  4.85
## cvpred      5.124 5.416 6.12 5.169 6.527 6.432 6.493 5.417  6.111  4.81
## QA          6.000 6.000 8.00 6.000 7.000 7.000 7.000 6.000  6.000  3.00
## CV residual 0.876 0.584 1.88 0.831 0.473 0.568 0.507 0.583 -0.111 -1.81
##                 521    530   538  547   550   553    555    562    566
## Predicted    6.1043  5.281 5.507 5.42 5.181 5.703  5.980  5.148  5.917
## cvpred       6.0921  5.259 5.499 5.40 5.152 5.679  5.948  5.135  5.896
## QA           6.0000  5.000 6.000 6.00 6.000 6.000  5.000  5.000  5.000
## CV residual -0.0921 -0.259 0.501 0.60 0.848 0.321 -0.948 -0.135 -0.896
##                 570   574    576    581   585  589    591   592     596
## Predicted   6.02094  5.28  6.185  5.374 6.113 6.77  5.324 5.683  5.0905
## cvpred      5.99896  5.26  6.159  5.337 6.074 6.77  5.302 5.692  5.0632
## QA          6.00000  4.00  6.000  5.000 7.000 8.00  5.000 6.000  5.0000
## CV residual 0.00104 -1.26 -0.159 -0.337 0.926 1.23 -0.302 0.308 -0.0632
##                600   602   607    610   613    619    620    621    622
## Predicted   5.50 5.412 6.475  6.244 5.428  5.675  5.790  5.143  5.138
## cvpred      5.47 5.382 6.465  6.219 5.397  5.645  5.769  5.132  5.127
## QA          6.00 6.000 7.000  6.000 6.000  5.000  5.000  5.000  5.000
## CV residual 0.53 0.618 0.535 -0.219 0.603 -0.645 -0.769 -0.132 -0.127
##                623     625    632  639  646   648   663   668    676   682
## Predicted    5.349  5.1633  5.669 5.09 5.49  5.48 5.394 5.658  5.732 5.544
## cvpred       5.322  5.0761  5.641 5.07 5.46  5.46 5.371 5.628  5.709 5.509
## QA           5.000  5.0000  5.000 7.00 7.00  4.00 6.000 6.000  5.000 6.000
## CV residual -0.322 -0.0761 -0.641 1.93 1.54 -1.46 0.629 0.372 -0.709 0.491
##                684    690   697     707   710    718     720   731    740
## Predicted    5.57  5.738 5.282  5.1098 5.805  5.490  5.1290 4.879  5.155
```

```
## cvpred       5.55  5.728 5.261  5.0825 5.784  5.462  5.0952 4.824  5.117
## QA           5.00  5.000 6.000  5.0000 6.000  5.000  5.0000 5.000  5.000
## CV residual -0.55 -0.728 0.739 -0.0825 0.216 -0.462 -0.0952 0.176 -0.117
##                742    748   749    752    753    759     764   774  776
## Predicted    5.129  5.327 5.476  5.172  5.210 5.0167  5.0727 5.560 4.89
## cvpred       5.113  5.298 5.456  5.137  5.178 4.9898  5.0441 5.548 4.86
## QA           5.000  5.000 6.000  5.000  5.000 5.0000  5.0000 6.000 5.00
## CV residual -0.113 -0.298 0.544 -0.137 -0.178 0.0102 -0.0441 0.452 0.14
##                778   781    785    787   791   792    795    796    797
## Predicted   5.169 5.049  5.210  5.410 5.462 4.990  6.501  5.638  5.478
## cvpred      5.127 5.006  5.188  5.383 5.449 4.971  6.487  5.651  5.451
## QA          6.000 6.000  5.000  5.000 6.000 5.000  6.000  5.000  5.000
## CV residual 0.873 0.994 -0.188 -0.383 0.551 0.029 -0.487 -0.651 -0.451
##                798      801    802   803   804   808    811   814    815
## Predicted   6.234  5.03018  5.410 6.402 5.323 6.763  5.614  6.00  6.209
## cvpred      6.232  5.00596  5.373 6.425 5.293 6.762  5.585  5.98  6.187
## QA          7.000  5.00000  5.000 7.000 6.000 7.000  5.000  4.00  6.000
## CV residual 0.768 -0.00596 -0.373 0.575 0.707 0.238 -0.585 -1.98 -0.187
##                817    818    832   833    836  838    842    845    849
## Predicted   5.686  6.578 5.9100  5.78  5.084 6.20  5.419  6.323  5.273
## cvpred      5.656  6.563 5.9037  5.75  5.054 6.19  5.386  6.311  5.255
## QA          6.000  6.000 6.0000  3.00  5.000 7.00  5.000  6.000  5.000
## CV residual 0.344 -0.563 0.0963 -2.75 -0.054 0.81 -0.386 -0.311 -0.255
##                852    853    868   874   876    881   893  898     912
## Predicted    5.467  5.705  6.110 6.244 6.404  5.470 5.570 5.64  6.0698
## cvpred       5.442  5.711  6.106 6.222 6.389  5.442 5.542 5.61  6.0316
## QA           5.000  5.000  6.000 7.000 7.000  5.000 6.000 6.00  6.0000
## CV residual -0.442 -0.711 -0.106 0.778 0.611 -0.442 0.458 0.39 -0.0316
##                916    917    919    920    921    922   924   938  944
## Predicted    6.388  5.249 5.9986  6.291  5.743 5.9986 5.701  5.63 5.40
## cvpred       6.375  5.241 5.9797  6.279  5.713 5.9797 5.664  5.62 5.38
## QA           6.000  5.000 6.0000  6.000  5.000 6.0000 6.000  4.00 7.00
## CV residual -0.375 -0.241 0.0203 -0.279 -0.713 0.0203 0.336 -1.62 1.62
##                950     957  959    965    982    994  998  1003  1007  1013
## Predicted   6.664  6.0663 5.83  6.188  5.234  5.319 6.11 6.569 6.569 4.836
## cvpred      6.653  6.0488 5.80  6.183  5.218  5.282 6.10 6.563 6.563 4.814
## QA          7.000  6.0000 7.00  6.000  5.000  5.000 7.00 7.000 7.000 5.000
## CV residual 0.347 -0.0488 1.20 -0.183 -0.218 -0.282 0.90 0.437 0.437 0.186
##                1019  1021   1023  1026   1028  1035   1045  1057  1076
## Predicted    6.531  6.21  5.642 5.254  5.614 5.073  6.276 6.142 6.137
## cvpred       6.522  6.19  5.617 5.224  5.596 5.045  6.276 6.126 6.111
## QA           6.000  6.00  5.000 6.000  5.000 6.000  6.000 7.000 7.000
## CV residual -0.522 -0.19 -0.617 0.776 -0.596 0.955 -0.276 0.874 0.889
##                1078 1090    1092  1094  1095  1097   1101   1107  1118 1121
## Predicted    5.940 5.87  6.0400 6.559 5.367 5.367  6.613  6.484 5.676 6.57
## cvpred       5.918 5.85  6.0246 6.553 5.331 5.331  6.604  6.466 5.651 6.57
## QA           5.000 7.00  6.0000 7.000 6.000 6.000  6.000  6.000 6.000 8.00
## CV residual -0.918 1.15 -0.0246 0.447 0.669 0.669 -0.604 -0.466 0.349 1.43
##                1136    1138  1157  1158  1170  1174  1182   1185   1191
## Predicted    6.322  6.1033 6.482 6.371 5.972 5.461  6.07  5.406  6.321
## cvpred       6.307  6.0917 6.467 6.395 5.957 5.436  6.05  5.415  6.305
## QA           6.000  6.0000 7.000 7.000 6.000 6.000  5.00  5.000  6.000
## CV residual -0.307 -0.0917 0.533 0.605 0.043 0.564 -1.05 -0.415 -0.305
##                1194 1205 1206 1207   1218  1223   1227   1244  1251  1253
```

```
## Predicted    5.1076 6.15 6.15 6.15  6.516 5.179  5.205  5.148 5.731  5.17
## cvpred       5.0761 6.14 6.14 6.14  6.506 5.147  5.162  5.134 5.704  5.14
## QA           5.0000 7.00 7.00 7.00  6.000 6.000  5.000  5.000 6.000  5.00
## CV residual -0.0761 0.86 0.86 0.86 -0.506 0.853 -0.162 -0.134 0.296 -0.14
##                1255   1256   1264 1270    1272  1275 1276  1278  1281  1282
## Predicted    5.666  5.563  5.001 6.82  6.0723 5.740 4.97 5.200 5.692 5.692
## cvpred       5.645  5.537  4.963 6.83  6.0618 5.715 4.96 5.171 5.668 5.668
## QA           5.000  5.000  4.000 8.00  6.0000 6.000 6.00 6.000 6.000 6.000
## CV residual -0.645 -0.537 -0.963 1.17 -0.0618 0.285 1.04 0.829 0.332 0.332
##                1283  1289  1290    1297  1298    1299  1304  1312  1323
## Predicted   5.597  5.70  5.70  5.1501  6.18  6.0683  6.10  6.34  6.22
## cvpred      5.578  5.71  5.71  5.0989  6.17  6.0574  6.09  6.33  6.20
## QA          6.000  5.00  5.00  5.0000  6.00  6.0000  5.00  6.00  5.00
## CV residual 0.422 -0.71 -0.71 -0.0989 -0.17 -0.0574 -1.09 -0.33 -1.20
##                1328  1329  1331   1345 1346  1368 1369  1378  1379   1383
## Predicted   5.722  5.22 5.147  5.935 5.64 5.806 4.97 5.884 5.459  5.149
## cvpred      5.702  5.19 5.135  5.906 5.62 5.829 4.95 5.872 5.429  5.131
## QA          6.000  5.00 6.000  5.000 6.00 6.000 6.00 6.000 6.000  5.000
## CV residual 0.298 -0.19 0.865 -0.906 0.38 0.171 1.05 0.128 0.571 -0.131
##                1389 1399  1409    1410  1411  1416    1423   1429   1431
## Predicted    5.421 5.46 6.868  6.0475 5.437  5.48  6.1077  5.690  5.741
## cvpred       5.409 5.43 6.866  6.0325 5.424  5.47  6.0844  5.663  5.723
## QA           5.000 7.00 7.000  6.0000 6.000  5.00  6.0000  5.000  5.000
## CV residual -0.409 1.57 0.134 -0.0325 0.576 -0.47 -0.0844 -0.663 -0.723
##                1447   1448 1450 1457   1465  1468 1469   1471    1473  1480
## Predicted    5.44  5.353 6.20 5.58  5.436  5.16 5.54  5.285  6.0829  5.87
## cvpred       5.42  5.328 6.19 5.56  5.411  5.14 5.51  5.255  6.0737  5.85
## QA           5.00  5.000 8.00 6.00  5.000  4.00 7.00  5.000  6.0000  5.00
## CV residual -0.42 -0.328 1.81 0.44 -0.411 -1.14 1.49 -0.255 -0.0737 -0.85
##                1481   1488   1489   1492    1497  1500 1515  1523   1526
## Predicted    5.41  5.718  5.872  5.872  5.0735 5.772 4.91  6.09  5.542
## cvpred       5.40  5.711  5.861  5.861  5.0679 5.761 4.90  6.08  5.517
## QA           4.00  5.000  5.000  5.000  5.0000 6.000 6.00  5.00  5.000
## CV residual -1.40 -0.711 -0.861 -0.861 -0.0679 0.239 1.10 -1.08 -0.517
##                1538   1540   1548    1554    1557    1560    1574    1576
## Predicted   5.522  5.718  5.982  5.1061  5.0643  5.0934  6.0631  6.0742
## cvpred      5.502  5.698  5.972  5.0734  5.0441  5.0782  6.0575  6.0645
## QA          6.000  5.000  5.000  5.0000  5.0000  5.0000  6.0000  6.0000
## CV residual 0.498 -0.698 -0.972 -0.0734 -0.0441 -0.0782 -0.0575 -0.0645
##                1589   1590   1598
## Predicted    6.157 4.9706  5.499
## cvpred       6.162 4.9406  5.479
## QA           6.000 5.0000  5.000
## CV residual -0.162 0.0594 -0.479
##
## 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     5.07   5.1   5.85   5.14   5.39   5.69   5.73   5.13   6.11
## cvpred       20.96  21.1  24.19  21.49  22.30  23.53  23.68  21.20  25.30
## QA            5.00   5.0   7.00   5.00   5.00   5.00   7.00   4.00   4.00
## CV residual -15.96 -16.1 -17.19 -16.49 -17.30 -18.53 -16.68 -17.20 -21.30
```

```
##                  48     56     62     68    74     85     86     87     94
## Predicted      5.51   5.21   4.91   5.43   5.0   5.81   5.46   6.46   5.47
## cvpred        22.82  21.56  20.31  22.49  20.7  24.06  22.60  26.69  22.62
## QA             5.00   5.00   5.00   5.00   4.0   6.00   5.00   6.00   5.00
## CV residual -17.82 -16.56 -15.31 -17.49 -16.7 -18.06 -17.60 -20.69 -17.62
##                  97    100    105    106    117    119    121    125    127
## Predicted      5.33   5.2   5.23   5.08   5.43   5.68   4.28   5.13   4.82
## cvpred        22.05  21.5  21.64  21.00  22.45  23.49  17.73  21.24  19.92
## QA             5.00   6.0   5.00   5.00   6.00   6.00   5.00   5.00   5.00
## CV residual -17.05 -15.5 -16.64 -16.00 -16.45 -17.49 -12.73 -16.24 -14.92
##                 130    132    144    145    150    152    162    174    176
## Predicted      5.44   6.18   5.46   6.96   5.68   5.6   5.61   5.78   5.52
## cvpred        22.52  25.59  22.60  28.81  23.50  23.3  23.19  23.93  22.84
## QA             5.00   5.00   5.00   6.00   6.00   4.0   4.00   6.00   5.00
## CV residual -17.52 -20.59 -17.60 -22.81 -17.50 -19.3 -19.19 -17.93 -17.84
##                 181    182    184    188    193    197    201    204    205
## Predicted      5.32   5.12   5.02   5.1   4.95   5.38   6.09   5.27   5.26
## cvpred        22.00  21.19  20.78  21.1  20.50  22.25  25.16  21.82  21.78
## QA             5.00   5.00   5.00   5.0   5.00   5.00   7.00   5.00   6.00
## CV residual -17.00 -16.19 -15.78 -16.1 -15.50 -17.25 -18.16 -16.82 -15.78
##                 206    208    211    222    227    228    246    247    248
## Predicted      6.2   5.01   6.49   5.17   5.76   5.18   5.59   5.05   5.03
## cvpred        25.6  20.72  26.83  21.40  23.85  21.39  23.11  20.90  20.83
## QA             7.0   5.00   6.00   5.00   6.00   5.00   6.00   5.00   5.00
## CV residual -18.6 -15.72 -20.83 -16.40 -17.85 -16.39 -17.11 -15.90 -15.83
##                 258    267    268    270    279    281    284    292    293
## Predicted      5.06   5.05   6.59   6.31   6.8   5.99   6.03   5.77   5.59
## cvpred        20.94  20.88  27.27  26.08  28.1  24.75  24.65  23.91  23.13
## QA             5.00   4.00   8.00   6.00   8.0   6.00   7.00   5.00   6.00
## CV residual -15.94 -16.88 -19.27 -20.08 -20.1 -18.75 -17.65 -18.91 -17.13
##                 304    314    316    319    325    338    339    342    345
## Predicted      4.99   5.01   5.89   5.8   5.54   5.65   6.18   6.35   5.81
## cvpred        20.66  20.96  24.37  24.0  22.87  23.39  25.55  26.23  24.04
## QA             5.00   5.00   6.00   7.0   6.00   5.00   6.00   6.00   6.00
## CV residual -15.66 -15.96 -18.37 -17.0 -16.87 -18.39 -19.55 -20.23 -18.04
##                 351    352    357    358    362    365    378    380    385
## Predicted      5.67   5.27   6.18   6.22   5.79   5.56   6.87   5.79   5.33
## cvpred        23.44  21.76  25.55  25.73  23.94  22.96  28.39  23.97  22.03
## QA             6.00   6.00   5.00   7.00   6.00   7.00   7.00   6.00   5.00
## CV residual -17.44 -15.76 -20.55 -18.73 -17.94 -15.96 -21.39 -17.97 -17.03
##                 389    393    396    399    404    405    410    413    417
## Predicted      5.38   5.43   6.46   5.75   5.6   5.13   5.71   4.93   6.08
## cvpred        22.27  22.49  26.71  23.77  23.1  21.21  23.56  20.40  25.12
## QA             6.00   5.00   7.00   6.00   6.0   5.00   4.00   5.00   6.00
## CV residual -16.27 -17.49 -19.71 -17.77 -17.1 -16.21 -19.56 -15.40 -19.12
##                 418    425    426    427    438    440    444    448    459
## Predicted      5.24   5.11   5.76   5.62   6.02   5.17   6.17   5.97   6.11
## cvpred        21.70  21.14  23.83  23.25  24.89  21.39  25.51  24.71  25.25
## QA             5.00   5.00   7.00   6.00   6.00   5.00   7.00   5.00   7.00
## CV residual -16.70 -16.14 -16.83 -17.25 -18.89 -16.39 -18.51 -19.71 -18.25
##                 462    485    503    509    510    511    513    514    516
## Predicted      5.2   6.56   6.38   5.45   6.37   5.76   5.72   6.07   5.26
## cvpred        21.5  27.14  26.35  22.54  26.31  23.78  23.64  25.10  21.73
## QA             5.0   6.00   7.00   6.00   7.00   5.00   6.00   7.00   5.00
```

```
## CV residual -16.5 -21.14 -19.35 -16.54 -19.31 -18.78 -17.64 -18.10 -16.73
##                519    520    524    527    528    533    535    537   540
## Predicted     6.46   5.88   5.21   5.88   5.98   5.74   5.48   5.28   5.9
## cvpred       26.73  24.34  21.54  24.34  24.74  23.75  22.66  21.83  24.4
## QA            6.00   5.00   5.00   5.00   6.00   5.00   6.00   5.00   5.0
## CV residual -20.73 -19.34 -16.54 -19.34 -18.74 -18.75 -16.66 -16.83 -19.4
##                544    551   554   557   559    560    563    580    603
## Predicted     5.76   5.37   5.2   5.9   5.9   6.34   5.11   5.71   4.93
## cvpred       23.80  22.21  21.5  24.4  24.4  26.20  21.16  23.61  20.40
## QA            6.00   6.00   5.0   6.0   6.0   6.00   5.00   6.00   5.00
## CV residual -17.80 -16.21 -16.5 -18.4 -18.4 -20.20 -16.16 -17.61 -15.40
##                615   616   617    618    628    638    647    655    669
## Predicted     5.76   5.2   5.2   6.13   5.04   4.67   5.36   5.24   5.64
## cvpred       23.82  21.5  21.5  25.32  20.83  19.32  22.18  21.71  23.32
## QA            6.00   5.0   5.0   6.00   5.00   5.00   5.00   5.00   5.00
## CV residual -17.82 -16.5 -16.5 -19.32 -15.83 -14.32 -17.18 -16.71 -18.32
##                679    685    691    695    701    702    704   711    716
## Predicted     5.1   4.72   4.96   5.18   4.92   5.28   5.46   4.9   5.21
## cvpred       21.1  19.49  20.48  21.43  20.35  21.84  22.59  20.2  21.55
## QA            5.0   5.00   3.00   5.00   6.00   6.00   4.00   5.0   6.00
## CV residual -16.1 -14.49 -17.48 -16.43 -14.35 -15.84 -18.59 -15.2 -15.55
##                717    721    722    724    732    734    735    739    743
## Predicted     5.22   5.24   5.13   6.35   5.41   5.23   5.18   5.17   5.22
## cvpred       21.59  21.69  21.24  26.27  22.40  21.63  21.42  21.40  21.58
## QA            5.00   5.00   5.00   5.00   5.00   5.00   5.00   5.00   5.00
## CV residual -16.59 -16.69 -16.24 -21.27 -17.40 -16.63 -16.42 -16.40 -16.58
##                747    755   756    757    760    762    763    768    773
## Predicted     5.2   5.11   5.2   5.19   5.23   5.07   5.35   4.94   5.02
## cvpred       21.5  21.20  21.5  21.45  21.64  20.99  22.12  20.44  20.75
## QA            6.0   6.00   6.0   6.00   5.00   5.00   6.00   5.00   5.00
## CV residual -15.5 -15.20 -15.5 -15.45 -16.64 -15.99 -16.12 -15.44 -15.75
##                775    807    812    826    841    850    861    862    875
## Predicted     5.55   6.74   5.97   5.63   6.42   5.27   5.04   5.17   6.27
## cvpred       22.97  27.87  24.68  23.32  26.52  21.80  20.85  21.39  25.93
## QA            6.00   7.00   6.00   5.00   7.00   5.00   5.00   6.00   7.00
## CV residual -16.97 -20.87 -18.68 -18.32 -19.52 -16.80 -15.85 -15.39 -18.93
##                877   879    887    892    900    902    904    906    923
## Predicted     5.87   5.3   5.35   5.07   5.13   5.64   5.76   5.12   6.29
## cvpred       24.28  21.9  22.14  20.97  21.20  23.32  23.79  21.19  26.00
## QA            4.00   6.0   6.00   5.00   3.00   7.00   7.00   5.00   6.00
## CV residual -20.28 -15.9 -16.14 -15.97 -18.20 -16.32 -16.79 -16.19 -20.00
##                926    931    941    947    948    952    954    960    963
## Predicted     6.42   5.57   6.62   6.56   6.48   6.48   6.59   5.42   5.33
## cvpred       26.54  23.05  27.37  27.09  26.82  26.82  27.22  22.43  22.04
## QA            7.00   5.00   7.00   7.00   7.00   7.00   7.00   5.00   5.00
## CV residual -19.54 -18.05 -20.37 -20.09 -19.82 -19.82 -20.22 -17.43 -17.04
##                966    971    974    992    999   1005   1009  1011   1012
## Predicted     6.26   6.42   5.96   5.32   4.97   5.65   6.37   6.5   5.88
## cvpred       25.89  26.52  24.64  22.01  20.53  23.37  26.35  26.9  24.33
## QA            6.00   6.00   5.00   5.00   6.00   5.00   7.00   7.0   6.00
## CV residual -19.89 -20.52 -19.64 -17.01 -14.53 -18.37 -19.35 -19.9 -18.33
##                1018  1030   1032   1036   1042   1052   1055   1069   1070
## Predicted     6.53   5.6   5.77   5.99   5.52   5.43   4.95   6.43   5.58
## cvpred       27.05  23.2  23.87  24.77  22.85  22.50  20.50  26.56  23.09
```

```
## QA            6.00   7.0   7.00   7.00   6.00   5.00   6.00   7.00   5.00
## CV residual -21.05 -16.2 -16.87 -17.77 -16.85 -17.50 -14.50 -19.56 -18.09
##               1079   1086   1089   1096   1102   1108   1115   1117   1125
## Predicted     5.94   5.35   5.87   5.67   6.18   6.54   6.66   5.68   5.82
## cvpred       24.56  22.16  24.25  23.49  25.55  27.03  27.54  23.49  24.10
## QA            5.00   5.00   7.00   5.00   6.00   7.00   6.00   6.00   4.00
## CV residual -19.56 -17.16 -17.25 -18.49 -19.55 -20.03 -21.54 -17.49 -20.10
##               1128   1130   1133   1134  1137  1142   1146   1148   1155
## Predicted     5.7   5.84   6.93   5.99   6.1   6.1   6.16   6.19   5.88
## cvpred       23.6  24.15  28.66  24.78  25.2  25.2  25.47  25.60  24.34
## QA            6.0   6.00   7.00   7.00   6.0   6.0   6.00   7.00   6.00
## CV residual -17.6 -18.15 -21.66 -17.78 -19.2 -19.2 -19.47 -18.60 -18.34
##               1163   1167   1169   1172   1175   1177  1184   1200   1201
## Predicted     6.42   5.92   6.06   5.77   5.46   5.38   5.1   5.15   5.45
## cvpred       26.55  24.49  25.30  23.85  22.59  22.25  21.1  21.31  22.56
## QA            7.00   5.00   6.00   6.00   6.00   4.00   5.0   6.00   6.00
## CV residual -19.55 -19.49 -19.30 -17.85 -16.59 -18.25 -16.1 -15.31 -16.56
##               1202   1209   1222   1226   1229  1236   1239   1240   1241
## Predicted     6.08   6.15   6.26   5.18   6.62   6.0   5.06   5.89   5.08
## cvpred       25.14  25.42  25.88  21.46  27.40  24.8  20.95  24.37  21.04
## QA            7.00   7.00   6.00   5.00   7.00   4.0   4.00   4.00   5.00
## CV residual -18.14 -18.42 -19.88 -16.46 -20.40 -20.8 -16.95 -20.37 -16.04
##               1245   1250   1254   1258   1280   1284   1285   1293   1295
## Predicted     5.76   5.73   5.28   5.67   6.31   5.37   5.85   6.39   5.77
## cvpred       23.80  23.70  21.82  23.47  26.10  22.21  24.22  26.41  23.85
## QA            6.00   6.00   5.00   6.00   7.00   6.00   5.00   6.00   6.00
## CV residual -17.80 -17.70 -16.82 -17.47 -19.10 -16.21 -19.22 -20.41 -17.85
##               1300   1308   1309   1311   1315   1317   1318   1336   1341
## Predicted     4.43   5.62   5.27   5.16   5.56   5.96   6.19   6.23   5.63
## cvpred       18.30  23.24  21.78  21.36  23.00  24.61  25.62  25.79  23.30
## QA            3.00   4.00   5.00   5.00   6.00   6.00   6.00   6.00   6.00
## CV residual -15.30 -19.24 -16.78 -16.36 -17.00 -18.61 -19.62 -19.79 -17.30
##               1343   1344   1353   1355   1360   1361   1365   1373   1382
## Predicted     5.54   5.63   5.23   5.39   5.75   5.65   5.82   4.91   4.92
## cvpred       22.94  23.30  21.64  22.31  23.77  23.37  24.07  20.36  20.36
## QA            6.00   6.00   5.00   5.00   6.00   5.00   6.00   5.00   5.00
## CV residual -16.94 -17.30 -16.64 -17.31 -17.77 -18.37 -18.07 -15.36 -15.36
##               1385   1391   1392   1402   1404   1405   1408   1412   1415
## Predicted     4.84   6.26   5.61   4.98   6.18   5.58   6.05   5.76   5.76
## cvpred       20.03  25.89  23.21  20.63  25.56  23.09  25.00  23.82  23.83
## QA            5.00   6.00   5.00   5.00   8.00   6.00   6.00   6.00   5.00
## CV residual -15.03 -19.89 -18.21 -15.63 -17.56 -17.09 -19.00 -17.82 -18.83
##               1426   1434   1439   1443   1444   1445   1449   1459   1464
## Predicted     5.74   5.76   5.46   5.44   5.88   5.63   5.48   5.94   5.54
## cvpred       23.77  23.82  22.57  22.52  24.33  23.30  22.65  24.56  22.94
## QA            6.00   7.00   5.00   5.00   5.00   6.00   5.00   5.00   6.00
## CV residual -17.77 -16.82 -17.57 -17.52 -19.33 -17.30 -17.65 -19.56 -16.94
##               1474   1475   1479   1483  1484   1487   1496   1498   1502
## Predicted     5.63   5.27   5.11   5.33   6.0   5.41   6.01   5.77   4.93
## cvpred       23.30  21.76  21.12  22.06  24.8  22.38  24.87  23.87  20.40
## QA            5.00   5.00   3.00   4.00   5.0   5.00   6.00   6.00   5.00
## CV residual -18.30 -16.76 -18.12 -18.06 -19.8 -17.38 -18.87 -17.87 -15.40
##               1508   1513   1517   1518   1520   1527   1532   1545   1553
## Predicted     5.88   5.39   6.09   5.84   5.26   5.53   5.36   6.27   5.68
```

```
## cvpred       24.36  22.30  25.19  24.16  21.78  22.89  22.15  25.93  23.51
## QA            6.00   6.00   5.00   6.00   5.00   6.00   5.00   7.00   6.00
## CV residual -18.36 -16.30 -20.19 -18.16 -16.78 -16.89 -17.15 -18.93 -17.51
##               1555   1556   1558   1564   1566   1568   1580   1586   1588
## Predicted     5.42   5.73   5.42   5.33   5.85   5.33   5.88   6.25   5.69
## cvpred       22.43  23.71  22.43  22.06  24.17  22.06  24.34  25.86  23.55
## QA            6.00   7.00   6.00   5.00   6.00   5.00   5.00   6.00   6.00
## CV residual -16.43 -16.71 -16.43 -17.06 -18.17 -17.06 -19.34 -19.86 -17.55
##               1591   1592   1594   1595   1596
## Predicted     6.17   5.48   5.49   5.52   5.95
## cvpred       25.51  22.92  22.68  22.86  24.59
## QA            6.00   6.00   6.00   5.00   6.00
## CV residual -19.51 -16.92 -16.68 -17.86 -18.59
##
## Sum of squares = 101042    Mean square = 316    n = 320
##
## Overall (Sum over all 320 folds)
##   ms
## 63.5
```

The average error rate for 5-fold validation was seen to be a mean-square error of 63.5.

Problem 8

```
phmean <- mean(redwine$PH[!is.na(redwine$PH)])
phsd <- sd(redwine$PH[!is.na(redwine$PH)])
ph.ub <- phmean + (3*phsd)
ph.lb <- phmean - (3*phsd)
```

```
redwine2 <- subset(redwine, PH < ph.ub & PH > ph.lb)
nrow(redwine2)
```

```
## [1] 1580
```

redwine2 has 1580 rows, which means 19 observations were removed due to outliers in the attribute PH.

Problem 9

```
winemodel2 <- lm(QA ~., data = redwine2)
summary(winemodel2)
```

```
##
## 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
```

The R squared in the new model has gone up from .358 to .363, suggesting an improvement. Therefore the new model has slightly better performance.

Based on the p-values, the 5 attributes that are most likely to be related to QA now are: VA, CH, SD, SU and AL, since they have the lowest p-values.