APSTA-GE: 2003 Intermediate Quantitative Methods
Lab Section 003, Week 12

29 November 2020

Introduction

In this lab, we will examine various factors that may be associated with student's math performance. The attached dataset, `student-mat.csv`, is a random sample of student's math scores in secondary education of two Portuguese schools. It was collected by using school reports and quesitonnaires.

Note: the dataset is in `csv` format with semi-colon as separators.

```
dat <- read.csv("student-mat.csv", sep = ";", header = TRUE)
```

Dimensions

```
dim(dat)
```

```
## [1] 395  33
```

```
table(is.na(dat))
```

```
##
## FALSE
## 13035
```

The dataset contains 395 records with 33 columns. No data point is missing.

Variables

Here is a list of column names.

```
as.data.frame(colnames(dat))
```

```
##      colnames(dat)
## 1          school
## 2             sex
## 3             age
## 4         address
## 5         famsize
## 6         Pstatus
## 7            Medu
## 8            Fedu
## 9            Mjob
## 10           Fjob
## 11         reason
## 12       guardian
## 13     traveltime
## 14      studytime
```

```
## 15      failures
## 16     schoolsup
## 17       famsup
## 18         paid
## 19    activities
## 20       nursery
## 21        higher
## 22      internet
## 23      romantic
## 24        famrel
## 25      freetime
## 26         goout
## 27          Dalc
## 28          Walc
## 29        health
## 30      absences
## 31            G1
## 32            G2
## 33            G3
```

Distributions

Plot distribution of numeric variables.
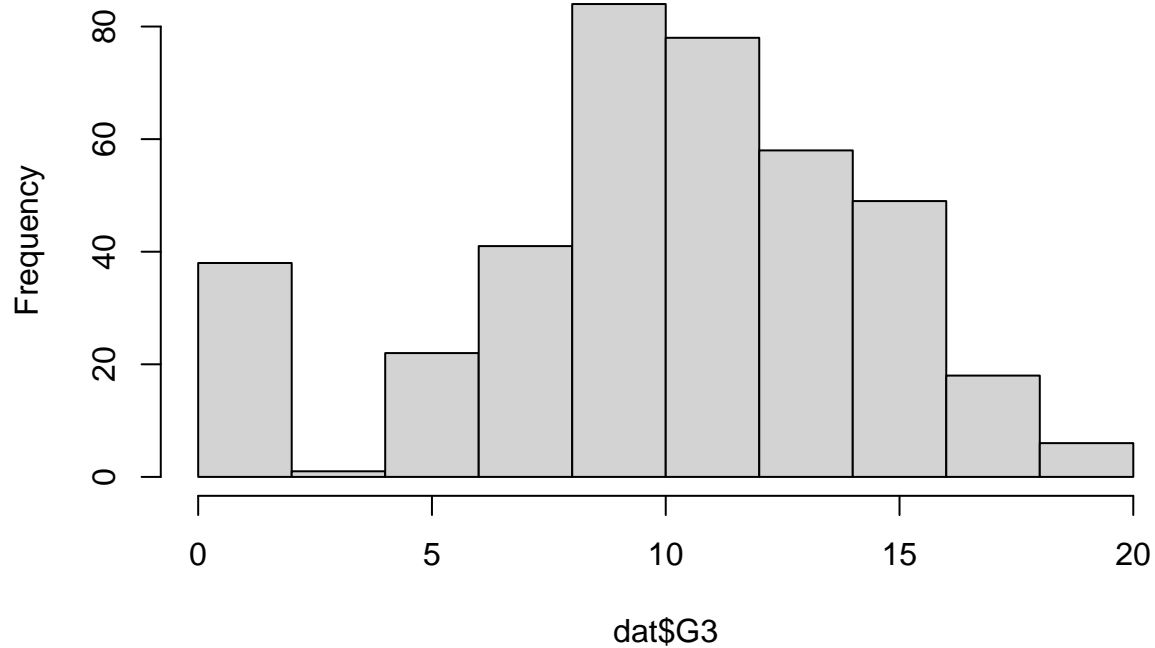
Math Score (Label)   Range: 0 - 20

```
summary(dat$G3)
```

```
##     Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##     0.00    8.00   11.00   10.42   14.00   20.00
```

```
hist(dat$G3)
```

Histogram of dat$G3

Regression Models

Baseline models

```
mod_BL <- lm(G3 ~ 1, data = dat)
summary(mod_BL)
```

```
##
## Call:
## lm(formula = G3 ~ 1, data = dat)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -10.4152  -2.4152   0.5848   3.5848   9.5848
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  10.4152     0.2305   45.18   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.581 on 394 degrees of freedom
```

Y ~ Age

```
mod_age <- lm(G3 ~ age, data = dat)
summary(mod_age)
```

```
##
```

```
## Call:
## lm(formula = G3 ~ age, data = dat)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -11.3992  -1.6588   0.3412   3.1809   9.5014
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  20.1011     2.9928   6.717 6.54e-11 ***
## age          -0.5801     0.1787  -3.246  0.00127 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.527 on 393 degrees of freedom
## Multiple R-squared:  0.02611,    Adjusted R-squared:  0.02363
## F-statistic: 10.54 on 1 and 393 DF,  p-value: 0.001271
```

Y ~ age + sex

```
mod_age_sex <- lm(G3 ~ age + factor(sex), data = dat)
summary(mod_age_sex)
```

```
##
## Call:
## lm(formula = G3 ~ age + factor(sex), data = dat)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -11.8593  -1.8593   0.2806   3.1407   9.7571
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  19.5024     2.9966   6.508 2.33e-10 ***
## age          -0.5700     0.1781  -3.200  0.00149 **
## factor(sex)M  0.9065     0.4547   1.994  0.04688 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.51 on 392 degrees of freedom
## Multiple R-squared:  0.03588,    Adjusted R-squared:  0.03097
## F-statistic: 7.295 on 2 and 392 DF,  p-value: 0.0007751
```

Model with All Predictors

```
for (col in 1:ncol(dat)) {
  if (class(dat[, col]) == "character") {
    dat[, col] <- as.factor(dat[, col])
  }
}
mod_full <- lm(G3 ~ ., data = dat)
summary(mod_full)
```

```
##
## Call:
## lm(formula = G3 ~ ., data = dat)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -7.9339 -0.5532  0.2680  0.9689  4.6461
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)     -1.115488   2.116958  -0.527 0.598573
## schoolMS         0.480742   0.366512   1.312 0.190485
## sexM             0.174396   0.233588   0.747 0.455805
## age             -0.173302   0.100780  -1.720 0.086380 .
## addressU         0.104455   0.270791   0.386 0.699922
## famsizeLE3       0.036512   0.226680   0.161 0.872128
## PstatusT        -0.127673   0.335626  -0.380 0.703875
## Medu             0.129685   0.149999   0.865 0.387859
## Fedu            -0.133940   0.128768  -1.040 0.298974
## Mjobhealth      -0.146426   0.518491  -0.282 0.777796
## Mjobother        0.074088   0.332044   0.223 0.823565
## Mjobservices     0.046956   0.369587   0.127 0.898973
## Mjobteacher     -0.026276   0.481632  -0.055 0.956522
## Fjobhealth       0.330948   0.666601   0.496 0.619871
## Fjobother       -0.083582   0.476796  -0.175 0.860945
## Fjobservices    -0.322142   0.493265  -0.653 0.514130
## Fjobteacher     -0.112364   0.601448  -0.187 0.851907
## reasonhome      -0.209183   0.256392  -0.816 0.415123
## reasonother      0.307554   0.380214   0.809 0.419120
## reasonreputation 0.129106   0.267254   0.483 0.629335
## guardianmother   0.195741   0.252672   0.775 0.439046
## guardianother    0.006565   0.463650   0.014 0.988710
## traveltime       0.096994   0.157800   0.615 0.539170
## studytime       -0.104754   0.134814  -0.777 0.437667
## failures        -0.160539   0.161006  -0.997 0.319399
## schoolsupyes     0.456448   0.319538   1.428 0.154043
## famsupyes        0.176870   0.224204   0.789 0.430710
## paidyes          0.075764   0.222100   0.341 0.733211
## activitiesyes   -0.346047   0.205938  -1.680 0.093774 .
## nurseryyes      -0.222716   0.254184  -0.876 0.381518
## higheryes        0.225921   0.500398   0.451 0.651919
## internetyes     -0.144462   0.287528  -0.502 0.615679
## romanticyes     -0.272008   0.219732  -1.238 0.216572
## famrel           0.356876   0.114124   3.127 0.001912 **
## freetime         0.047002   0.110209   0.426 0.670021
## goout            0.012007   0.105230   0.114 0.909224
## Dalc            -0.185019   0.153124  -1.208 0.227741
## Walc             0.176772   0.114943   1.538 0.124966
## health           0.062995   0.074800   0.842 0.400259
## absences         0.045879   0.013412   3.421 0.000698 ***
## G1               0.188847   0.062373   3.028 0.002645 **
## G2               0.957330   0.053460  17.907  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Residual standard error: 1.901 on 353 degrees of freedom
## Multiple R-squared:  0.8458, Adjusted R-squared:  0.8279
## F-statistic: 47.21 on 41 and 353 DF,  p-value: < 2.2e-16
```

Model Selection: Step-wise AIC

Additive Terms Only

Start from the full model. Run stepAIC by going "backward". You can also use the baseline model and go "forward" or use any model and go with "both" directions.

```
stepAIC(mod_full, scope = . ~ 1, direction = "backward")
```

```
## Start:  AIC=547.02
## G3 ~ school + sex + age + address + famsize + Pstatus + Medu +
##      Fedu + Mjob + Fjob + reason + guardian + traveltime + studytime +
##      failures + schoolsup + famsup + paid + activities + nursery +
##      higher + internet + romantic + famrel + freetime + goout +
##      Dalc + Walc + health + absences + G1 + G2
##
##               Df Sum of Sq     RSS    AIC
## - Mjob         4      1.18  1276.7 539.39
## - Fjob         4      7.60  1283.1 541.37
## - reason       3      8.85  1284.3 543.75
## - guardian     2      2.57  1278.1 543.82
## - goout        1      0.05  1275.5 545.04
## - famsize      1      0.09  1275.6 545.05
## - paid         1      0.42  1275.9 545.15
## - Pstatus      1      0.52  1276.0 545.19
## - address      1      0.54  1276.0 545.19
## - freetime     1      0.66  1276.2 545.23
## - higher       1      0.74  1276.2 545.25
## - internet     1      0.91  1276.4 545.31
## - traveltime   1      1.37  1276.9 545.45
## - sex          1      2.01  1277.5 545.65
## - studytime    1      2.18  1277.7 545.70
## - famsup       1      2.25  1277.8 545.72
## - health       1      2.56  1278.1 545.82
## - Medu         1      2.70  1278.2 545.86
## - nursery      1      2.77  1278.3 545.88
## - failures     1      3.59  1279.1 546.14
## - Fedu         1      3.91  1279.4 546.23
## - Dalc         1      5.28  1280.8 546.65
## - romantic     1      5.54  1281.0 546.74
## - school       1      6.22  1281.7 546.94
## <none>                     1275.5 547.02
## - schoolsup    1      7.37  1282.9 547.30
## - Walc         1      8.55  1284.0 547.66
## - activities   1     10.20  1285.7 548.17
## - age          1     10.68  1286.2 548.32
## - G1           1     33.12  1308.6 555.15
## - famrel       1     35.33  1310.8 555.82
## - absences     1     42.28  1317.8 557.91
## - G2           1   1158.72  2434.2 800.31
```

```
##
## Step:  AIC=539.39
## G3 ~ school + sex + age + address + famsize + Pstatus + Medu +
##     Fedu + Fjob + reason + guardian + traveltime + studytime +
##     failures + schoolsup + famsup + paid + activities + nursery +
##     higher + internet + romantic + famrel + freetime + goout +
##     Dalc + Walc + health + absences + G1 + G2
##
##               Df Sum of Sq    RSS    AIC
## - Fjob         4     7.59 1284.3 533.73
## - reason       3     8.73 1285.4 536.08
## - guardian     2     2.61 1279.3 536.20
## - goout        1     0.05 1276.7 537.40
## - famsize      1     0.06 1276.8 537.41
## - paid         1     0.36 1277.0 537.50
## - address      1     0.56 1277.2 537.56
## - freetime     1     0.65 1277.3 537.59
## - Pstatus      1     0.67 1277.4 537.60
## - higher       1     0.76 1277.5 537.63
## - internet     1     0.99 1277.7 537.70
## - traveltime   1     1.41 1278.1 537.83
## - studytime    1     1.96 1278.7 538.00
## - famsup       1     2.01 1278.7 538.01
## - sex          1     2.20 1278.9 538.07
## - health       1     2.54 1279.2 538.18
## - Medu         1     2.76 1279.5 538.24
## - nursery      1     2.91 1279.6 538.29
## - failures     1     3.72 1280.4 538.54
## - Fedu         1     3.82 1280.5 538.57
## - Dalc         1     4.64 1281.3 538.82
## - romantic     1     5.53 1282.2 539.10
## - school       1     6.33 1283.0 539.34
## <none>                    1276.7 539.39
## - schoolsup    1     7.83 1284.5 539.81
## - Walc         1     7.90 1284.6 539.83
## - activities   1    10.48 1287.2 540.62
## - age          1    10.70 1287.4 540.69
## - G1           1    32.86 1309.5 547.43
## - famrel       1    36.93 1313.6 548.65
## - absences     1    44.45 1321.1 550.91
## - G2           1  1177.44 2454.1 795.52
##
## Step:  AIC=533.73
## G3 ~ school + sex + age + address + famsize + Pstatus + Medu +
##     Fedu + reason + guardian + traveltime + studytime + failures +
##     schoolsup + famsup + paid + activities + nursery + higher +
##     internet + romantic + famrel + freetime + goout + Dalc +
##     Walc + health + absences + G1 + G2
##
##               Df Sum of Sq    RSS    AIC
## - reason       3     9.66 1294.0 530.69
## - guardian     2     3.16 1287.4 530.70
## - goout        1     0.12 1284.4 531.77
## - famsize      1     0.16 1284.4 531.78
```

```
## - paid         1       0.46 1284.8 531.88
## - address      1       0.51 1284.8 531.89
## - Pstatus      1       0.98 1285.3 532.03
## - freetime     1       0.98 1285.3 532.03
## - higher       1       1.09 1285.4 532.07
## - traveltime   1       1.51 1285.8 532.20
## - internet     1       1.60 1285.9 532.22
## - studytime    1       1.84 1286.1 532.30
## - sex          1       2.44 1286.7 532.48
## - famsup       1       2.47 1286.8 532.49
## - Medu         1       2.81 1287.1 532.60
## - nursery      1       3.21 1287.5 532.72
## - health       1       3.47 1287.8 532.80
## - Fedu         1       4.01 1288.3 532.96
## - failures     1       4.12 1288.4 533.00
## - school       1       5.23 1289.5 533.34
## - Dalc         1       5.45 1289.7 533.40
## - romantic     1       5.54 1289.8 533.43
## <none>                 1284.3 533.73
## - Walc         1       7.11 1291.4 533.91
## - schoolsup    1       9.10 1293.4 534.52
## - age          1      10.65 1294.9 534.99
## - activities   1      12.66 1297.0 535.61
## - famrel       1      34.67 1319.0 542.26
## - G1           1      36.47 1320.8 542.79
## - absences     1      44.08 1328.4 545.06
## - G2           1    1190.12 2474.4 790.77
##
## Step:  AIC=530.69
## G3 ~ school + sex + age + address + famsize + Pstatus + Medu +
##     Fedu + guardian + traveltime + studytime + failures + schoolsup +
##     famsup + paid + activities + nursery + higher + internet +
##     romantic + famrel + freetime + goout + Dalc + Walc + health +
##     absences + G1 + G2
##
##               Df Sum of Sq    RSS    AIC
## - guardian     2      3.26 1297.2 527.69
## - goout        1      0.06 1294.0 528.71
## - address      1      0.08 1294.0 528.72
## - famsize      1      0.16 1294.1 528.74
## - paid         1      0.49 1294.4 528.84
## - higher       1      0.50 1294.4 528.85
## - Pstatus      1      0.88 1294.8 528.96
## - freetime     1      1.46 1295.4 529.14
## - sex          1      1.47 1295.4 529.14
## - traveltime   1      1.50 1295.4 529.15
## - internet     1      1.75 1295.7 529.23
## - studytime    1      1.77 1295.7 529.23
## - famsup       1      2.46 1296.4 529.44
## - health       1      2.94 1296.9 529.59
## - nursery      1      3.23 1297.2 529.68
## - Medu         1      3.66 1297.6 529.81
## - Fedu         1      4.35 1298.3 530.02
## - failures     1      4.75 1298.7 530.14
```

```
## - Dalc        1        5.22 1299.2 530.28
## - school      1        5.28 1299.2 530.30
## - romantic    1        5.41 1299.3 530.34
## <none>                       1294.0 530.69
## - Walc        1        7.59 1301.5 531.00
## - schoolsup   1        9.16 1303.1 531.48
## - age         1       10.84 1304.8 531.99
## - activities  1       11.28 1305.2 532.12
## - famrel      1       35.01 1329.0 539.24
## - G1          1       36.35 1330.3 539.64
## - absences    1       43.30 1337.2 541.70
## - G2          1     1210.19 2504.1 789.49
##
## Step:  AIC=527.69
## G3 ~ school + sex + age + address + famsize + Pstatus + Medu +
##     Fedu + traveltime + studytime + failures + schoolsup + famsup +
##     paid + activities + nursery + higher + internet + romantic +
##     famrel + freetime + goout + Dalc + Walc + health + absences +
##     G1 + G2
##
##               Df Sum of Sq    RSS    AIC
## - address     1        0.00 1297.2 525.69
## - famsize     1        0.21 1297.4 525.75
## - goout       1        0.22 1297.4 525.75
## - higher      1        0.33 1297.5 525.79
## - paid        1        0.63 1297.8 525.88
## - Pstatus     1        1.07 1298.3 526.01
## - traveltime  1        1.09 1298.3 526.02
## - freetime    1        1.28 1298.5 526.07
## - sex         1        1.32 1298.5 526.09
## - internet    1        1.87 1299.1 526.26
## - studytime   1        2.19 1299.4 526.35
## - famsup      1        2.27 1299.5 526.38
## - nursery     1        2.76 1300.0 526.53
## - health      1        2.89 1300.1 526.57
## - Medu        1        5.03 1302.2 527.22
## - school      1        5.25 1302.5 527.28
## - romantic    1        5.66 1302.9 527.40
## - Dalc        1        5.70 1302.9 527.42
## - Fedu        1        6.15 1303.3 527.55
## - failures    1        6.42 1303.6 527.64
## <none>                       1297.2 527.69
## - Walc        1        7.85 1305.0 528.07
## - schoolsup   1        8.88 1306.1 528.38
## - activities  1       11.11 1308.3 529.06
## - age         1       14.39 1311.6 530.04
## - famrel      1       34.83 1332.0 536.15
## - G1          1       36.97 1334.2 536.79
## - absences    1       44.05 1341.2 538.88
## - G2          1     1207.82 2505.0 785.63
##
## Step:  AIC=525.69
## G3 ~ school + sex + age + famsize + Pstatus + Medu + Fedu + traveltime +
##     studytime + failures + schoolsup + famsup + paid + activities +
```

9

```
##     nursery + higher + internet + romantic + famrel + freetime +
##     goout + Dalc + Walc + health + absences + G1 + G2
##
##              Df Sum of Sq    RSS    AIC
## - famsize     1      0.21 1297.4 523.75
## - goout       1      0.23 1297.4 523.76
## - higher      1      0.33 1297.5 523.79
## - paid        1      0.63 1297.8 523.88
## - Pstatus     1      1.07 1298.3 524.01
## - traveltime  1      1.16 1298.4 524.04
## - freetime    1      1.28 1298.5 524.07
## - sex         1      1.33 1298.5 524.09
## - internet    1      1.92 1299.1 524.27
## - studytime   1      2.22 1299.4 524.36
## - famsup      1      2.27 1299.5 524.38
## - nursery     1      2.76 1300.0 524.53
## - health      1      2.89 1300.1 524.57
## - Medu        1      5.04 1302.2 525.22
## - school      1      5.49 1302.7 525.35
## - romantic    1      5.66 1302.9 525.41
## - Dalc        1      5.70 1302.9 525.42
## - Fedu        1      6.16 1303.4 525.56
## - failures    1      6.42 1303.6 525.64
## <none>                     1297.2 525.69
## - Walc        1      7.91 1305.1 526.09
## - schoolsup   1      8.88 1306.1 526.38
## - activities  1     11.22 1308.4 527.09
## - age         1     14.40 1311.6 528.05
## - famrel      1     34.84 1332.0 534.16
## - G1          1     37.04 1334.2 534.81
## - absences    1     44.30 1341.5 536.95
## - G2          1   1214.11 2511.3 784.62
##
## Step:  AIC=523.75
## G3 ~ school + sex + age + Pstatus + Medu + Fedu + traveltime +
##     studytime + failures + schoolsup + famsup + paid + activities +
##     nursery + higher + internet + romantic + famrel + freetime +
##     goout + Dalc + Walc + health + absences + G1 + G2
##
##              Df Sum of Sq    RSS    AIC
## - goout       1      0.23 1297.6 521.82
## - higher      1      0.34 1297.8 521.85
## - paid        1      0.63 1298.0 521.94
## - traveltime  1      1.20 1298.6 522.12
## - Pstatus     1      1.25 1298.7 522.13
## - freetime    1      1.27 1298.7 522.14
## - sex         1      1.37 1298.8 522.17
## - internet    1      1.89 1299.3 522.33
## - famsup      1      2.20 1299.6 522.42
## - studytime   1      2.29 1299.7 522.45
## - nursery     1      2.62 1300.0 522.55
## - health      1      2.85 1300.3 522.62
## - Medu        1      4.96 1302.4 523.26
## - school      1      5.58 1303.0 523.45
```

```
## - romantic     1        5.59 1303.0 523.45
## - Dalc         1        5.63 1303.0 523.46
## - Fedu         1        6.30 1303.7 523.66
## - failures     1        6.50 1303.9 523.73
## <none>                       1297.4 523.75
## - Walc         1        8.03 1305.5 524.19
## - schoolsup    1        8.90 1306.3 524.45
## - activities   1       11.13 1308.5 525.13
## - age          1       14.36 1311.8 526.10
## - famrel       1       34.81 1332.2 532.21
## - G1           1       37.12 1334.5 532.89
## - absences     1       44.34 1341.8 535.03
## - G2           1     1216.66 2514.1 783.06
##
## Step:  AIC=521.82
## G3 ~ school + sex + age + Pstatus + Medu + Fedu + traveltime +
##     studytime + failures + schoolsup + famsup + paid + activities +
##     nursery + higher + internet + romantic + famrel + freetime +
##     Dalc + Walc + health + absences + G1 + G2
##
##                Df Sum of Sq    RSS    AIC
## - higher       1        0.35 1298.0 519.93
## - paid         1        0.61 1298.3 520.01
## - traveltime   1        1.20 1298.8 520.19
## - Pstatus      1        1.26 1298.9 520.20
## - sex          1        1.29 1298.9 520.21
## - freetime     1        1.65 1299.3 520.32
## - internet     1        1.81 1299.5 520.37
## - famsup       1        2.17 1299.8 520.48
## - studytime    1        2.22 1299.9 520.50
## - nursery      1        2.55 1300.2 520.59
## - health       1        2.75 1300.4 520.66
## - Medu         1        5.17 1302.8 521.39
## - school       1        5.46 1303.1 521.48
## - romantic     1        5.65 1303.3 521.54
## - Dalc         1        5.73 1303.4 521.56
## - Fedu         1        6.24 1303.9 521.72
## - failures     1        6.42 1304.1 521.77
## <none>                       1297.6 521.82
## - schoolsup    1        8.88 1306.5 522.51
## - Walc         1       10.43 1308.1 522.98
## - activities   1       11.02 1308.7 523.16
## - age          1       14.14 1311.8 524.10
## - famrel       1       35.48 1333.1 530.47
## - G1           1       37.06 1334.7 530.94
## - absences     1       44.14 1341.8 533.03
## - G2           1     1221.43 2519.1 781.84
##
## Step:  AIC=519.93
## G3 ~ school + sex + age + Pstatus + Medu + Fedu + traveltime +
##     studytime + failures + schoolsup + famsup + paid + activities +
##     nursery + internet + romantic + famrel + freetime + Dalc +
##     Walc + health + absences + G1 + G2
##
```

```
##                 Df Sum of Sq    RSS    AIC
## - paid          1      0.73 1298.7 518.15
## - sex           1      1.14 1299.1 518.27
## - traveltime    1      1.15 1299.1 518.28
## - Pstatus       1      1.33 1299.3 518.33
## - freetime      1      1.63 1299.6 518.42
## - internet      1      1.85 1299.8 518.49
## - studytime     1      2.11 1300.1 518.57
## - famsup        1      2.19 1300.2 518.59
## - nursery       1      2.57 1300.6 518.71
## - health        1      2.81 1300.8 518.78
## - Medu          1      5.28 1303.3 519.53
## - Dalc          1      5.65 1303.6 519.64
## - school        1      5.73 1303.7 519.67
## - romantic      1      5.97 1304.0 519.74
## - Fedu          1      6.13 1304.1 519.79
## <none>                      1298.0 519.93
## - failures      1      7.05 1305.0 520.07
## - schoolsup     1      8.93 1306.9 520.64
## - Walc          1     10.40 1308.4 521.08
## - activities    1     10.76 1308.8 521.19
## - age           1     15.08 1313.1 522.49
## - famrel        1     35.70 1333.7 528.64
## - G1            1     37.71 1335.7 529.24
## - absences      1     44.06 1342.0 531.11
## - G2            1   1221.11 2519.1 779.85
##
## Step:  AIC=518.15
## G3 ~ school + sex + age + Pstatus + Medu + Fedu + traveltime +
##     studytime + failures + schoolsup + famsup + activities +
##     nursery + internet + romantic + famrel + freetime + Dalc +
##     Walc + health + absences + G1 + G2
##
##                 Df Sum of Sq    RSS    AIC
## - sex           1      0.97 1299.7 516.45
## - traveltime    1      1.06 1299.8 516.47
## - Pstatus       1      1.22 1300.0 516.52
## - freetime      1      1.50 1300.2 516.60
## - internet      1      1.66 1300.4 516.65
## - studytime     1      1.92 1300.6 516.73
## - nursery       1      2.35 1301.1 516.86
## - health        1      2.66 1301.4 516.96
## - famsup        1      3.14 1301.9 517.10
## - Dalc          1      5.44 1304.2 517.80
## - Medu          1      5.63 1304.4 517.86
## - school        1      5.95 1304.7 517.95
## - romantic      1      5.95 1304.7 517.96
## - Fedu          1      6.40 1305.1 518.09
## <none>                      1298.7 518.15
## - failures      1      7.86 1306.6 518.53
## - schoolsup     1      8.68 1307.4 518.78
## - activities    1     10.87 1309.6 519.44
## - Walc          1     11.14 1309.9 519.53
## - age           1     14.99 1313.7 520.68
```

```
## - famrel       1        36.22 1334.9 527.01
## - G1           1        37.08 1335.8 527.27
## - absences     1        43.75 1342.5 529.24
## - G2           1      1239.93 2538.7 780.90
##
## Step:  AIC=516.45
## G3 ~ school + age + Pstatus + Medu + Fedu + traveltime + studytime +
##     failures + schoolsup + famsup + activities + nursery + internet +
##     romantic + famrel + freetime + Dalc + Walc + health + absences +
##     G1 + G2
##
##                 Df Sum of Sq    RSS    AIC
## - traveltime  1       1.17 1300.9 514.80
## - Pstatus     1       1.20 1300.9 514.81
## - internet    1       1.54 1301.2 514.91
## - freetime    1       1.85 1301.5 515.01
## - nursery     1       2.25 1302.0 515.13
## - famsup      1       2.77 1302.5 515.29
## - studytime   1       2.86 1302.5 515.31
## - health      1       3.08 1302.8 515.38
## - Dalc        1       4.98 1304.7 515.96
## - school      1       5.67 1305.4 516.17
## - Medu        1       6.13 1305.8 516.31
## - romantic    1       6.43 1306.1 516.39
## - Fedu        1       6.59 1306.3 516.44
## <none>                     1299.7 516.45
## - failures    1       7.67 1307.4 516.77
## - schoolsup   1       8.21 1307.9 516.93
## - activities  1      10.27 1310.0 517.56
## - Walc        1      11.92 1311.6 518.05
## - age         1      15.11 1314.8 519.01
## - famrel      1      36.95 1336.7 525.52
## - G1          1      37.72 1337.4 525.75
## - absences    1      42.86 1342.6 527.26
## - G2          1    1242.31 2542.0 779.42
##
## Step:  AIC=514.8
## G3 ~ school + age + Pstatus + Medu + Fedu + studytime + failures +
##     schoolsup + famsup + activities + nursery + internet + romantic +
##     famrel + freetime + Dalc + Walc + health + absences + G1 +
##     G2
##
##                 Df Sum of Sq    RSS    AIC
## - Pstatus     1       1.20 1302.1 513.17
## - freetime    1       1.65 1302.5 513.30
## - internet    1       1.69 1302.6 513.32
## - nursery     1       2.15 1303.0 513.45
## - health      1       2.99 1303.9 513.71
## - studytime   1       3.11 1304.0 513.74
## - famsup      1       3.12 1304.0 513.75
## - Dalc        1       4.67 1305.5 514.22
## - Medu        1       5.86 1306.7 514.58
## - romantic    1       6.38 1307.2 514.73
## <none>                     1300.9 514.80
```

```
## - Fedu        1       7.24 1308.1 515.00
## - school      1       7.32 1308.2 515.02
## - failures    1       7.59 1308.5 515.10
## - schoolsup   1       8.30 1309.2 515.31
## - activities  1       9.89 1310.8 515.79
## - Walc        1      12.47 1313.3 516.57
## - age         1      15.89 1316.8 517.60
## - famrel      1      37.31 1338.2 523.97
## - G1          1      39.82 1340.7 524.71
## - absences    1      42.93 1343.8 525.63
## - G2          1    1257.83 2558.7 780.01
##
## Step:  AIC=513.17
## G3 ~ school + age + Medu + Fedu + studytime + failures + schoolsup +
##      famsup + activities + nursery + internet + romantic + famrel +
##      freetime + Dalc + Walc + health + absences + G1 + G2
##
##               Df Sum of Sq    RSS    AIC
## - freetime    1       1.60 1303.7 511.65
## - nursery     1       1.97 1304.0 511.76
## - internet    1       2.06 1304.1 511.79
## - health      1       2.92 1305.0 512.05
## - famsup      1       2.93 1305.0 512.05
## - studytime   1       3.14 1305.2 512.12
## - Dalc        1       4.49 1306.5 512.52
## - romantic    1       6.15 1308.2 513.03
## - Medu        1       6.34 1308.4 513.08
## <none>                    1302.1 513.17
## - Fedu        1       7.06 1309.1 513.30
## - school      1       7.14 1309.2 513.33
## - failures    1       7.39 1309.5 513.40
## - schoolsup   1       8.48 1310.5 513.73
## - activities  1      10.80 1312.9 514.43
## - Walc        1      12.21 1314.3 514.85
## - age         1      16.23 1318.3 516.06
## - famrel      1      37.32 1339.4 522.33
## - G1          1      39.35 1341.4 522.93
## - absences    1      45.40 1347.5 524.70
## - G2          1    1265.76 2567.8 779.41
##
## Step:  AIC=511.65
## G3 ~ school + age + Medu + Fedu + studytime + failures + schoolsup +
##      famsup + activities + nursery + internet + romantic + famrel +
##      Dalc + Walc + health + absences + G1 + G2
##
##               Df Sum of Sq    RSS    AIC
## - internet    1       1.89 1305.5 510.22
## - nursery     1       1.94 1305.6 510.24
## - health      1       3.06 1306.7 510.58
## - famsup      1       3.18 1306.8 510.61
## - studytime   1       3.79 1307.5 510.80
## - Dalc        1       3.82 1307.5 510.81
## - romantic    1       6.19 1309.8 511.52
## <none>                    1303.7 511.65
```

```
## - Medu          1      6.75 1310.4 511.69
## - failures      1      6.89 1310.6 511.73
## - school        1      7.35 1311.0 511.87
## - Fedu          1      7.53 1311.2 511.92
## - schoolsup     1      8.37 1312.0 512.18
## - activities    1     10.01 1313.7 512.67
## - Walc          1     12.44 1316.1 513.40
## - age           1     16.39 1320.1 514.59
## - G1            1     40.81 1344.5 521.83
## - famrel        1     41.09 1344.8 521.91
## - absences      1     44.23 1347.9 522.83
## - G2            1   1264.31 2568.0 777.44
##
## Step:  AIC=510.22
## G3 ~ school + age + Medu + Fedu + studytime + failures + schoolsup +
##     famsup + activities + nursery + romantic + famrel + Dalc +
##     Walc + health + absences + G1 + G2
##
##                Df Sum of Sq    RSS    AIC
## - nursery      1      1.78 1307.3 508.76
## - famsup       1      2.92 1308.5 509.10
## - health       1      3.54 1309.1 509.29
## - studytime    1      4.01 1309.6 509.43
## - Dalc         1      4.06 1309.6 509.45
## - Medu         1      6.11 1311.7 510.06
## <none>                     1305.5 510.22
## - failures     1      6.99 1312.5 510.33
## - romantic     1      7.01 1312.6 510.34
## - Fedu         1      7.63 1313.2 510.52
## - school       1      7.92 1313.5 510.61
## - schoolsup    1      8.75 1314.3 510.86
## - activities   1     10.13 1315.7 511.28
## - Walc         1     12.36 1317.9 511.94
## - age          1     15.57 1321.1 512.91
## - famrel       1     40.11 1345.7 520.17
## - G1           1     42.19 1347.8 520.78
## - absences     1     43.07 1348.6 521.04
## - G2           1   1267.44 2573.0 776.21
##
## Step:  AIC=508.76
## G3 ~ school + age + Medu + Fedu + studytime + failures + schoolsup +
##     famsup + activities + romantic + famrel + Dalc + Walc + health +
##     absences + G1 + G2
##
##                Df Sum of Sq    RSS    AIC
## - famsup       1      2.94 1310.3 507.65
## - health       1      3.54 1310.9 507.83
## - Dalc         1      3.89 1311.2 507.93
## - studytime    1      4.25 1311.6 508.04
## - Medu         1      5.49 1312.8 508.41
## <none>                     1307.3 508.76
## - failures     1      6.81 1314.1 508.81
## - romantic     1      7.22 1314.6 508.94
## - Fedu         1      7.97 1315.3 509.16
```

```
## - school      1      8.25 1315.6 509.25
## - schoolsup   1      8.48 1315.8 509.31
## - activities  1      9.82 1317.2 509.72
## - Walc        1     12.69 1320.0 510.57
## - age         1     15.40 1322.7 511.39
## - famrel      1     40.31 1347.7 518.76
## - G1          1     42.09 1349.4 519.28
## - absences    1     42.86 1350.2 519.50
## - G2          1   1267.41 2574.8 774.48
##
## Step:  AIC=507.65
## G3 ~ school + age + Medu + Fedu + studytime + failures + schoolsup +
##     activities + romantic + famrel + Dalc + Walc + health + absences +
##     G1 + G2
##
##             Df Sum of Sq    RSS    AIC
## - studytime  1      3.34 1313.6 506.65
## - Dalc       1      3.64 1313.9 506.74
## - health     1      3.77 1314.0 506.78
## - Medu       1      6.13 1316.4 507.49
## <none>                   1310.3 507.65
## - failures   1      6.88 1317.2 507.72
## - Fedu       1      7.01 1317.3 507.75
## - romantic   1      7.16 1317.4 507.80
## - school     1      7.36 1317.6 507.86
## - schoolsup  1      8.77 1319.0 508.28
## - activities 1     10.44 1320.7 508.78
## - Walc       1     11.93 1322.2 509.23
## - age        1     16.22 1326.5 510.51
## - famrel     1     39.68 1350.0 517.43
## - G1         1     40.55 1350.8 517.69
## - absences   1     43.39 1353.7 518.51
## - G2         1   1267.34 2577.6 772.91
##
## Step:  AIC=506.65
## G3 ~ school + age + Medu + Fedu + failures + schoolsup + activities +
##     romantic + famrel + Dalc + Walc + health + absences + G1 +
##     G2
##
##             Df Sum of Sq    RSS    AIC
## - Dalc       1      3.40 1317.0 505.67
## - health     1      4.03 1317.6 505.86
## - Medu       1      5.61 1319.2 506.33
## - failures   1      5.84 1319.5 506.40
## - Fedu       1      6.22 1319.8 506.52
## <none>                   1313.6 506.65
## - romantic   1      7.92 1321.5 507.03
## - schoolsup  1      8.09 1321.7 507.08
## - school     1      8.42 1322.0 507.18
## - activities 1     11.31 1324.9 508.04
## - Walc       1     14.02 1327.6 508.84
## - age        1     17.96 1331.6 510.02
## - G1         1     39.40 1353.0 516.33
## - famrel     1     39.69 1353.3 516.41
```

```
## - absences    1      45.29 1358.9 518.04
## - G2          1    1267.13 2580.8 771.39
##
## Step:  AIC=505.67
## G3 ~ school + age + Medu + Fedu + failures + schoolsup + activities +
##     romantic + famrel + Walc + health + absences + G1 + G2
##
##              Df Sum of Sq    RSS    AIC
## - health      1      3.74 1320.8 504.80
## - Medu        1      4.83 1321.8 505.12
## - Fedu        1      5.97 1323.0 505.46
## - failures    1      6.35 1323.4 505.57
## <none>                    1317.0 505.67
## - schoolsup   1      7.34 1324.3 505.87
## - school      1      7.59 1324.6 505.94
## - romantic    1      8.02 1325.0 506.07
## - activities  1     10.73 1327.7 506.88
## - Walc        1     11.05 1328.1 506.97
## - age         1     18.76 1335.8 509.26
## - G1          1     39.42 1356.4 515.32
## - famrel      1     39.75 1356.8 515.42
## - absences    1     44.98 1362.0 516.94
## - G2          1   1266.03 2583.0 769.75
##
## Step:  AIC=504.8
## G3 ~ school + age + Medu + Fedu + failures + schoolsup + activities +
##     romantic + famrel + Walc + absences + G1 + G2
##
##              Df Sum of Sq    RSS    AIC
## - Medu        1      4.32 1325.1 504.08
## - Fedu        1      5.46 1326.2 504.42
## - failures    1      5.88 1326.6 504.55
## <none>                    1320.8 504.80
## - schoolsup   1      6.72 1327.5 504.80
## - school      1      7.35 1328.1 504.99
## - romantic    1      7.62 1328.4 505.07
## - activities  1     10.49 1331.2 505.92
## - Walc        1     12.51 1333.3 506.52
## - age         1     20.67 1341.4 508.93
## - G1          1     39.86 1360.6 514.54
## - famrel      1     43.03 1363.8 515.46
## - absences    1     44.39 1365.2 515.85
## - G2          1   1262.31 2583.1 767.75
##
## Step:  AIC=504.08
## G3 ~ school + age + Fedu + failures + schoolsup + activities +
##     romantic + famrel + Walc + absences + G1 + G2
##
##              Df Sum of Sq    RSS    AIC
## - Fedu        1      1.89 1327.0 502.65
## - schoolsup   1      5.75 1330.8 503.79
## - failures    1      6.51 1331.6 504.02
## - school      1      6.63 1331.7 504.06
## <none>                    1325.1 504.08
```

```
## - romantic     1        7.08 1332.2 504.19
## - activities   1       10.05 1335.1 505.07
## - Walc         1       11.84 1336.9 505.60
## - age          1       21.91 1347.0 508.56
## - G1           1       38.75 1363.8 513.47
## - famrel       1       43.04 1368.1 514.71
## - absences     1       48.54 1373.6 516.30
## - G2           1     1288.29 2613.4 770.36
##
## Step:  AIC=502.65
## G3 ~ school + age + failures + schoolsup + activities + romantic +
##     famrel + Walc + absences + G1 + G2
##
##              Df Sum of Sq    RSS    AIC
## - schoolsup   1        5.49 1332.5 502.28
## - failures    1        5.50 1332.5 502.28
## - school      1        6.71 1333.7 502.64
## <none>                      1327.0 502.65
## - romantic    1        7.46 1334.4 502.86
## - activities  1       10.85 1337.8 503.86
## - Walc        1       11.42 1338.4 504.03
## - age         1       20.86 1347.8 506.81
## - G1          1       37.46 1364.4 511.64
## - famrel      1       43.08 1370.0 513.27
## - absences    1       47.77 1374.7 514.62
## - G2          1     1293.22 2620.2 769.38
##
## Step:  AIC=502.28
## G3 ~ school + age + failures + activities + romantic + famrel +
##     Walc + absences + G1 + G2
##
##              Df Sum of Sq    RSS    AIC
## - failures    1        5.54 1338.0 501.92
## - school      1        6.32 1338.8 502.14
## <none>                      1332.5 502.28
## - romantic    1        8.16 1340.6 502.69
## - Walc        1        9.94 1342.4 503.21
## - activities  1       10.38 1342.8 503.34
## - age         1       26.72 1359.2 508.12
## - G1          1       33.28 1365.7 510.02
## - famrel      1       43.34 1375.8 512.92
## - absences    1       50.61 1383.1 515.00
## - G2          1     1315.78 2648.2 771.59
##
## Step:  AIC=501.92
## G3 ~ school + age + activities + romantic + famrel + Walc + absences +
##     G1 + G2
##
##              Df Sum of Sq    RSS    AIC
## <none>                      1338.0 501.92
## - school      1        6.90 1344.9 501.95
## - romantic    1        8.72 1346.7 502.48
## - Walc        1        8.82 1346.8 502.51
## - activities  1        9.90 1347.9 502.83
```

```
## - age         1      33.30 1371.3 509.63
## - G1          1      37.24 1375.2 510.76
## - famrel      1      44.94 1382.9 512.96
## - absences    1      50.88 1388.9 514.66
## - G2          1    1336.04 2674.0 773.42
##
## Call:
## lm(formula = G3 ~ school + age + activities + romantic + famrel +
##      Walc + absences + G1 + G2, data = dat)
##
## Coefficients:
##   (Intercept)        schoolMS              age   activitiesyes       romanticyes
##       0.65012         0.45495         -0.26017        -0.32049          -0.32668
##        famrel            Walc         absences              G1                G2
##       0.38361         0.11979          0.04714         0.18058           0.96153
```

This is the selected model. It has the lowest AIC score (best prediction performance).

```
mod_lowest_AIC <- lm(G3 ~ school + age + activities + romantic + famrel + Walc +
                    absences + G1 + G2, data = dat)
summary(mod_lowest_AIC)
```

```
##
## Call:
## lm(formula = G3 ~ school + age + activities + romantic + famrel +
##      Walc + absences + G1 + G2, data = dat)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -8.8416 -0.4534  0.2645  1.0247  4.0315
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.65012    1.45806   0.446 0.655933
## schoolMS       0.45495    0.32293   1.409 0.159693
## age           -0.26017    0.08405  -3.095 0.002109 **
## activitiesyes -0.32049    0.18989  -1.688 0.092278 .
## romanticyes   -0.32668    0.20622  -1.584 0.113989
## famrel         0.38361    0.10668   3.596 0.000365 ***
## Walc           0.11979    0.07521   1.593 0.112050
## absences       0.04714    0.01232   3.826 0.000152 ***
## G1             0.18058    0.05517   3.273 0.001159 **
## G2             0.96153    0.04904  19.607  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.864 on 385 degrees of freedom
## Multiple R-squared:  0.8382, Adjusted R-squared:  0.8344
## F-statistic: 221.6 on 9 and 385 DF,  p-value: < 2.2e-16
```

with Interactive Terms

```
# stepAIC(mod_full, scope = . ~ .^2, direction = "both")
```