**Data for grade3**

lm(formula = G3 ~ ., data = finalnumericdata)

Residuals:

Min 1Q Median 3Q Max

-8.8372 -0.5133 -0.0169 0.6064 5.6715

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) -0.0006693 0.9259720 -0.001 0.999423

age 0.0222177 0.0474924 0.468 0.640082

Pstatus -0.0866044 0.1601915 -0.541 0.588957

Medu -0.0373899 0.0611778 -0.611 0.541313

Fedu 0.0304015 0.0598311 0.508 0.611549

traveltime 0.1241927 0.0737015 1.685 0.092479 .

studytime 0.0509324 0.0653211 0.780 0.435852

failures -0.2447822 0.0982942 -2.490 0.013025 \*

famrel -0.0236495 0.0541330 -0.437 0.662353

freetime -0.0418228 0.0521594 -0.802 0.422961

goout -0.0241505 0.0495568 -0.487 0.626197

Dalc -0.0699276 0.0707945 -0.988 0.323659

Walc -0.0064709 0.0546809 -0.118 0.905836

health -0.0454001 0.0355448 -1.277 0.201988

absences 0.0171964 0.0113831 1.511 0.131377

G1 0.1326353 0.0371462 3.571 0.000384 \*\*\*

G2 0.8747815 0.0347310 25.187 < 2e-16 \*\*\*

internetyes 0.1174885 0.1256736 0.935 0.350220

higheryes 0.2003229 0.1814631 1.104 0.270052

nurseryyes -0.0732449 0.1267589 -0.578 0.563590

activitiesyes 0.0282219 0.1035868 0.272 0.785369

paidyes -0.1831286 0.2147892 -0.853 0.394213

familysupport 0.1239321 0.1062758 1.166 0.244009

schoolsupport -0.1501707 0.1686312 -0.891 0.373529

guardianother 0.2403286 0.2250423 1.068 0.285971

reasonrepo -0.1620956 0.1367574 -1.185 0.236363

reasonother -0.4119268 0.1694597 -2.431 0.015348 \*

resonhome -0.0902566 0.1327717 -0.680 0.496893

famlsizedum 0.0348521 0.1140587 0.306 0.760040

address 0.1826722 0.1181534 1.546 0.122602

sexdum -0.1004888 0.1154785 -0.870 0.384532

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.25 on 618 degrees of freedom

Multiple R-squared: 0.8573, Adjusted R-squared: 0.8504

F-statistic: 123.8 on 30 and 618 DF, p-value: < 2.2e-16

**Data for g1**

lm(formula = G1 ~ ., data = dataforg1)

Residuals:

Min 1Q Median 3Q Max

-11.0106 -1.4372 -0.1118 1.5210 6.5992

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 10.23687 1.66000 6.167 1.26e-09 \*\*\*

age -0.08647 0.08708 -0.993 0.321153

Pstatus 0.29655 0.29625 1.001 0.317216

Medu 0.22609 0.11286 2.003 0.045580 \*

Fedu 0.16735 0.11048 1.515 0.130340

traveltime -0.09063 0.13638 -0.665 0.506600

studytime 0.43227 0.11966 3.612 0.000328 \*\*\*

failures -0.99602 0.17602 -5.658 2.34e-08 \*\*\*

famrel 0.04297 0.09973 0.431 0.666676

freetime -0.08150 0.09648 -0.845 0.398564

goout -0.01733 0.09170 -0.189 0.850182

Dalc -0.20593 0.13079 -1.574 0.115885

Walc -0.05526 0.10118 -0.546 0.585148

health -0.02522 0.06558 -0.385 0.700634

absences -0.04879 0.02098 -2.325 0.020370 \*

internetyes 0.27715 0.23236 1.193 0.233407

higheryes 1.57908 0.32949 4.792 2.06e-06 \*\*\*

nurseryyes -0.22036 0.23449 -0.940 0.347709

activitiesyes 0.23855 0.19151 1.246 0.213387

paidyes -0.61172 0.39649 -1.543 0.123374

familysupport -0.13841 0.19666 -0.704 0.481817

schoolsupport -1.06299 0.30917 -3.438 0.000625 \*\*\*

guardianother 0.16987 0.41654 0.408 0.683542

reasonrepo 0.56512 0.25206 2.242 0.025311 \*

reasonother -0.14877 0.31347 -0.475 0.635253

resonhome 0.34646 0.24532 1.412 0.158358

famlsizedum 0.25399 0.21090 1.204 0.228923

address 0.45127 0.21788 2.071 0.038754 \*

sexdum -0.35336 0.21330 -1.657 0.098106 .

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.313 on 620 degrees of freedom

Multiple R-squared: 0.3207, Adjusted R-squared: 0.29

F-statistic: 10.45 on 28 and 620 DF, p-value: < 2.2e-16

**DATA for g2**

lm(formula = G2 ~ ., data = dataforg2)

Residuals:

Min 1Q Median 3Q Max

-13.4327 -1.3762 -0.0357 1.5310 6.8427

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 7.06365 1.77543 3.979 7.75e-05 \*\*\*

age 0.10486 0.09314 1.126 0.26069

Pstatus 0.34900 0.31685 1.101 0.27113

Medu 0.23577 0.12071 1.953 0.05124 .

Fedu 0.20864 0.11816 1.766 0.07794 .

traveltime -0.09549 0.14587 -0.655 0.51296

studytime 0.35551 0.12798 2.778 0.00564 \*\*

failures -1.22158 0.18826 -6.489 1.77e-10 \*\*\*

famrel 0.18821 0.10666 1.765 0.07814 .

freetime -0.10517 0.10319 -1.019 0.30851

goout -0.05816 0.09807 -0.593 0.55339

Dalc -0.19137 0.13989 -1.368 0.17180

Walc -0.08373 0.10821 -0.774 0.43936

health -0.10469 0.07014 -1.493 0.13603

absences -0.04025 0.02244 -1.794 0.07335 .

internetyes 0.29509 0.24851 1.187 0.23552

higheryes 1.62223 0.35240 4.603 5.05e-06 \*\*\*

nurseryyes -0.17777 0.25079 -0.709 0.47870

activitiesyes 0.18123 0.20483 0.885 0.37661

paidyes -0.26265 0.42406 -0.619 0.53590

familysupport -0.11873 0.21034 -0.564 0.57263

schoolsupport -0.82228 0.33067 -2.487 0.01315 \*

guardianother 0.18676 0.44550 0.419 0.67521

reasonrepo 0.59612 0.26958 2.211 0.02738 \*

reasonother -0.29224 0.33527 -0.872 0.38373

resonhome 0.20446 0.26237 0.779 0.43613

famlsizedum 0.22010 0.22557 0.976 0.32955

address 0.49298 0.23303 2.116 0.03478 \*

sexdum -0.32175 0.22813 -1.410 0.15894

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.474 on 620 degrees of freedom

Multiple R-squared: 0.3101, Adjusted R-squared: 0.279

F-statistic: 9.954 on 28 and 620 DF, p-value: < 2.2e-16

#########

Trainfor g1

lm(formula = G1 ~ ., data = trainforg1)

Residuals:

Min 1Q Median 3Q Max

-10.959 -1.345 -0.128 1.476 6.823

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 11.83818 2.18644 5.414 1.07e-07 \*\*\*

age -0.12001 0.11302 -1.062 0.288974

Pstatus 0.12117 0.38203 0.317 0.751277

Medu 0.24631 0.14407 1.710 0.088129 .

Fedu 0.13336 0.14091 0.946 0.344513

traveltime -0.14737 0.17995 -0.819 0.413304

studytime 0.36210 0.15296 2.367 0.018406 \*

failures -1.12306 0.21298 -5.273 2.22e-07 \*\*\*

famrel 0.07195 0.12575 0.572 0.567500

freetime -0.15641 0.12521 -1.249 0.212355

goout -0.06875 0.11558 -0.595 0.552329

Dalc -0.25729 0.17205 -1.495 0.135610

Walc -0.06107 0.12520 -0.488 0.625959

health -0.01623 0.08323 -0.195 0.845533

absences -0.04063 0.02574 -1.579 0.115185

internetyes 0.31727 0.29350 1.081 0.280368

higheryes 1.37767 0.39529 3.485 0.000547 \*\*\*

nurseryyes -0.19324 0.30876 -0.626 0.531781

activitiesyes 0.39352 0.24128 1.631 0.103696

paidyes -0.53317 0.50669 -1.052 0.293328

familysupport -0.19421 0.25715 -0.755 0.450549

schoolsupport -0.99132 0.40765 -2.432 0.015471 \*

guardianother 0.13070 0.52389 0.249 0.803122

reasonrepo 0.74915 0.31977 2.343 0.019638 \*

reasonother -0.39056 0.39558 -0.987 0.324100

resonhome 0.32092 0.30989 1.036 0.301030

famlsizedum 0.20898 0.26246 0.796 0.426374

address 0.12601 0.27672 0.455 0.649094

sexdum -0.38073 0.26908 -1.415 0.157891

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.331 on 393 degrees of freedom

Multiple R-squared: 0.3451, Adjusted R-squared: 0.2984

F-statistic: 7.396 on 28 and 393 DF, p-value: < 2.2e-16

Step Model:

**For G1:**

Call:

lm(formula = G1 ~ Medu + studytime + failures + Dalc + internetyes +

higheryes + address + sexdum, data = trainforg1)

Residuals:

Min 1Q Median 3Q Max

-7.429 -1.505 -0.114 1.615 6.494

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 8.7226 0.5312 16.420 < 2e-16 \*\*\*

Medu 0.2841 0.1073 2.648 0.008400 \*\*

studytime 0.5224 0.1376 3.796 0.000169 \*\*\*

failures -1.2791 0.1956 -6.540 1.82e-10 \*\*\*

Dalc -0.3244 0.1209 -2.684 0.007557 \*\*

internetyes 0.4094 0.2692 1.521 0.129141

higheryes 1.4408 0.3787 3.805 0.000163 \*\*\*

address 0.4068 0.2411 1.687 0.092334 .

sexdum -0.3580 0.2423 -1.477 0.140399

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.256 on 413 degrees of freedom

Multiple R-squared: 0.3129, Adjusted R-squared: 0.2996

F-statistic: 23.51 on 8 and 413 DF, p-value: < 2.2e-16

**For G2:**

Call:

lm(formula = G2 ~ Pstatus + Medu + studytime + failures + Walc +

higheryes + nurseryyes + schoolsupport + reasonrepo + reasonother +

address, data = trainforg2)

Residuals:

Min 1Q Median 3Q Max

-12.6238 -1.4377 -0.0936 1.4567 6.1591

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 8.37864 0.66745 12.553 < 2e-16 \*\*\*

Pstatus 0.58212 0.34316 1.696 0.09058 .

Medu 0.24749 0.10714 2.310 0.02139 \*

studytime 0.23832 0.14501 1.643 0.10106

failures -1.23640 0.21325 -5.798 1.34e-08 \*\*\*

Walc -0.26620 0.09163 -2.905 0.00387 \*\*

higheryes 1.81843 0.39332 4.623 5.07e-06 \*\*\*

nurseryyes 0.50576 0.29181 1.733 0.08382 .

schoolsupport -0.80829 0.39443 -2.049 0.04107 \*

reasonrepo 0.72063 0.29604 2.434 0.01535 \*

reasonother -0.53166 0.35459 -1.499 0.13455

address 0.64119 0.25709 2.494 0.01302 \*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.354 on 410 degrees of freedom

Multiple R-squared: 0.3012, Adjusted R-squared: 0.2825

F-statistic: 16.07 on 11 and 410 DF, p-value: < 2.2e-16

**Predicted and Residuals for G1:**

pridict Actaul residual

602 12.268196 8 4.26819605

426 9.366401 7 2.36640123

211 11.956028 14 -2.04397179

191 12.268196 8 4.26819605

379 11.023067 5 6.02306678

24 11.585799 13 -1.41420054

58 12.069069 14 -1.93093090

27 11.387839 15 -3.61216057

338 13.358778 9 4.35877836

19 7.510181 7 0.51018135

254 10.989090 11 -0.01091027

297 11.910233 9 2.91023298

440 11.623049 11 0.62304897

614 12.142938 13 -0.85706222

328 12.212688 11 1.21268817

83 12.552290 11 1.55229043

248 12.153988 14 -1.84601177

160 11.984102 9 2.98410166

174 7.896665 12 -4.10333463

386 11.858843 14 -2.14115661

**Predicted and Residuals for G2:**

pridict Actual residual

602 12.125784 8 4.12578450

426 9.736491 10 -0.26350873

211 12.621990 13 -0.37801019

191 12.047797 12 0.04779695

379 11.790758 18 -6.20924175

24 12.553554 10 2.55355423

58 13.847141 15 -1.15285852

27 12.127016 12 0.12701636

338 13.819258 18 -4.18074229

19 8.132889 8 0.13288915

254 10.813019 10 0.81301909

297 12.365337 16 -3.63466282

440 11.493765 16 -4.50623518

614 12.879029 10 2.87902851

328 12.940137 17 -4.05986262

83 13.599655 11 2.59965475

248 13.048528 13 0.04852769

160 11.878298 12 -0.12170223

174 8.102355 8 0.10235539

386 12.631542 9 3.63154178

accuracy(pridictg1,testforg1$G1)

ME RMSE MAE MPE MAPE

Test set 0.0599252 3.171156 2.50191 -Inf Inf

accuracy(pridictg2,testforg2$G2)

ME RMSE MAE MPE MAPE

Test set -0.03640495 2.758307 2.017264 -Inf Inf

Predictions:

For G1:

#Predict grade 1 when a male student wants to do higher studies,studies for 8 hours, has school support, has never failed, has internet at home, lives in urban area and consumes very less alcohol daily, mother has secondary education..

predict\_df <-data.frame(higheryes=1,studytime=3,schoolsupport=1,failures=0,Medu=3, Dalc=1, address=1, internetyes=1, sexdum=1)

> #predict

> predit\_out <-predict(stepg1, predict\_df, interval="prediction", level=0.95)

> #print output

> print(predit\_out)

fit lwr upr

1 12.71672 8.248602 17.18484

#Predict grade 1 when a female student wants to do higher studies,studies for 12 hours, has school support, has never failed, has internet at home, lives in urban area and consumes very less alcohol daily,mother has higher education.

predict\_df <-data.frame(higheryes=1,studytime=4,schoolsupport=1,failures=0,Medu=4, Dalc=1, address=1, internetyes=1, sexdum=0)

#predict

predit\_out <-predict(stepg1, predict\_df, interval="prediction", level=0.95)

#print output

print(predit\_out)

fit lwr upr

13.88117 9.399138 18.36321