Homework #5

Data Set Command: mathStat<-read.csv("SuccessInStats.csv", header = TRUE)

1. Commands: mathStat2<-lm(statexam~mathtest+engtest+eng\_gpa+math\_gpa+othr\_gpa, data = mathStat)

summary(mathStat2)

Call:

lm(formula = statexam ~ mathtest + engtest + eng\_gpa + math\_gpa +

othr\_gpa, data = mathStat)

Residuals:

Min 1Q Median 3Q Max

-38.841 -14.726 0.093 13.452 33.126

Coefficients:

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

(Intercept) 6.74484 27.69089 0.244 0.8081

mathtest 0.11580 0.02451 4.726 8.03e-06 \*\*\*

engtest 0.04942 0.02721 1.816 0.0725 .

eng\_gpa -3.36491 7.44624 -0.452 0.6524

math\_gpa 5.47788 6.86508 0.798 0.4269

othr\_gpa -9.70152 8.49977 -1.141 0.2566

---

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

Residual standard error: 17.36 on 94 degrees of freedom

Multiple R-squared: 0.2691, Adjusted R-squared: 0.2302

F-statistic: 6.922 on 5 and 94 DF, p-value: 1.527e-05

a. Based on the p-value which is <.05, meaning the R^2 value is significantly different from 0, the regression model using all the predictor variables predicts test performance in Math 208 relatively well. The R^2 value is .2691(AdjustedR^2= .2302), meaning the predictor variables explain about 26.91% of the variation in test performance in Math 208. However, only one predictor variable was truly significant in the ability to predict math performance, that predictor variable was math test scores with a p-value <.05. Thus, math test scores are significant to the model as a predictor value, while all the other predictor variables are not because their p-values are all >.05. Based off the model, statexam(i)= 6.74 + .11mathtest(i) + .049engtest(i) - 3.36eng\_gpa(i) + 5.48math\_gpa(i) – 9.702other\_gpa(i).

1. Commands: mathStat3<-lm(statexam~mathtest, data = mathStat)

summary(mathStat3)

Call:

lm(formula = statexam ~ mathtest, data = mathStat)

Residuals:

Min 1Q Median 3Q Max

-39.513 -13.474 1.056 13.042 33.634

Coefficients:

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

(Intercept) 3.06131 10.55660 0.290 0.772

mathtest 0.12386 0.02261 5.479 3.31e-07 \*\*\*

---

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

Residual standard error: 17.4 on 98 degrees of freedom

Multiple R-squared: 0.2345, Adjusted R-squared: 0.2267

F-statistic: 30.02 on 1 and 98 DF, p-value: 3.315e-07

a. Since, the original model showed that the only significant predictor variable was math test performance, I only used this variable since it created the best model. The p-value was <.05 so R^2 was significantly different from 0, which makes the model significant in being able to predict stat exam performance. The R^2 = .2345 (AdjustedR^2=.2267), meaning that the predictor variable math test performance can explain 23.45% of the variation in test performance in Math 208. Based off the model, statexam(i) = 3.06 + .12mathtest(i).

b. .12(100) = expected stat exam difference in performance = 12. Parker should score 12 points higher on the stat exams on average than Paige. (Don’t use intercept since they both start off from the same place)

1. Commands: mathStat4<-lm(statexam~mathtest+math\_gpa, data = mathStat)

summary(mathStat4)

Call:

lm(formula = statexam ~ mathtest + math\_gpa, data = mathStat)

Residuals:

Min 1Q Median 3Q Max

-40.643 -13.824 0.086 12.655 33.441

Coefficients:

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

(Intercept) -5.0098 16.9128 -0.296 0.768

mathtest 0.1183 0.0244 4.850 4.71e-06 \*\*\*

math\_gpa 3.8220 6.2443 0.612 0.542

---

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

Residual standard error: 17.46 on 97 degrees of freedom

Multiple R-squared: 0.2374, Adjusted R-squared: 0.2217

F-statistic: 15.1 on 2 and 97 DF, p-value: 1.951e-06

1. I used the best model which included only the mathtest predictor variable, then I added the math\_gpa to the model because Dr. Swain said to. The p-value was <.05 so R^2 was significantly different from 0, which makes the model significant in being able to predict stat exam performance. The R^2 = .2374 (AdjustedR^2=.2217), meaning that the predictor variable math test performance can explain 23.74% of the variation in stat test performance in Math 208. However, the p-value of the predictor variable math\_gpa is >.05, meaning math\_gpa does not significantly explain any variation in stat exam performance. Based off the model, statexam(i) = -5.0098 + .12mathtest(i) + 3.82math\_gpa(i).
2. .12(0) + 3.82(1) = expected stat exam difference in performance = 3.82. Paige should score 3.82 points higher than Parker on the stat exams on average in Math 208.

\*\*\*\*\*SOME VALUES ARE ROUNDED\*\*\*\*\*