Using the predictors we found to be significant from the previous hw, our hypotheses tests proceeds as follows:

, there is no relationship between a student’s high school GPA / SAT verbal scores and their first-year college GPA.

, there is a relationship between a student’s high school GPA / SAT verbal scores and their first-year college GPA.

Note, is used to index our predictors, where represents HSGPA, and represents SATV.

The test-statistics, degrees of freedom, and corresponding p-values for each of the predictors are:

, ,

, ,

Thus, using a significance level , we find that there’s sufficient evidence to reject the null (for both cases) in favor of the alternative. We conclude that there is a relationship between a student’s first year college GPA and their high school GPA / SAT verbal scores.

solution

library(readr)

## Warning: package 'readr' was built under R version 4.1.2

data <- read\_csv("data.csv")

## Rows: 219 Columns: 10  
## ── Column specification ────────────────────────────────────────────────────────  
## Delimiter: ","  
## dbl (10): GPA, HSGPA, SATV, SATM, Male, HU, SS, FirstGen, White, CollegeBound  
##   
## ℹ Use `spec()` to retrieve the full column specification for this data.  
## ℹ Specify the column types or set `show\_col\_types = FALSE` to quiet this message.

attach(data)  
head(data)

## # A tibble: 6 × 10  
## GPA HSGPA SATV SATM Male HU SS FirstGen White CollegeBound  
## <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>  
## 1 3.06 3.83 680 770 1 3 9 1 1 1  
## 2 4.15 4 740 720 0 9 3 0 1 1  
## 3 3.41 3.7 640 570 0 16 13 0 0 1  
## 4 3.21 3.51 740 700 0 22 0 0 1 1  
## 5 3.48 3.83 610 610 0 30.5 1.5 0 1 1  
## 6 2.95 3.25 600 570 0 18 3 0 1 1

model = lm(GPA ~ HSGPA + SATV)  
summary(model)

##   
## Call:  
## lm(formula = GPA ~ HSGPA + SATV)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.97894 -0.27639 0.02867 0.30133 0.87956   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.6351217 0.2955033 2.149 0.03272 \*   
## HSGPA 0.4975320 0.0750569 6.629 2.66e-10 \*\*\*  
## SATV 0.0012283 0.0003373 3.641 0.00034 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.4061 on 216 degrees of freedom  
## Multiple R-squared: 0.246, Adjusted R-squared: 0.239   
## F-statistic: 35.23 on 2 and 216 DF, p-value: 5.711e-14

anova(model)

## Analysis of Variance Table  
##   
## Response: GPA  
## Df Sum Sq Mean Sq F value Pr(>F)   
## HSGPA 1 9.433 9.4329 57.210 1.111e-12 \*\*\*  
## SATV 1 2.186 2.1861 13.258 0.0003397 \*\*\*  
## Residuals 216 35.615 0.1649   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1