

```

# Assignment: ASSIGNMENT 7
# Name: Smitshoek, Stephen
# Date: 2022-05-05

## Set the working directory to the root of your DSC 520 directory
setwd("C:\\Users\\sksmi\\PeytoAccess\\Personal\\Bellevue\\DSC520\\dsc520")

## Load the `data/r4ds/heights.csv` to
heights_df <- read.csv("data/r4ds/heights.csv")

# Fit a linear model
earn_lm <- lm(earn ~ height + sex + ed + age + race, data=heights_df)

# View the summary of your model
summary(earn_lm)

predicted_df <- data.frame(
  earn = predict(earn_lm, heights_df),
  ed=heights_df$ed, race=heights_df$race, height=heights_df$height,
  age=heights_df$age, sex=heights_df$sex
)

## Compute deviation (i.e. residuals)
mean_earn <- mean(heights_df$earn)
## Corrected Sum of Squares Total
sst <- sum((mean_earn - heights_df$earn)^2)
## Corrected Sum of Squares for Model
ssm <- sum((mean_earn - predicted_df$earn)^2)
## Residuals
residuals <- heights_df$earn - predicted_df$earn
## Sum of Squares for Error
sse <- sum(residuals^2)
## R Squared
r_squared <- ssm/sst

## Number of observations
n <- NROW(heights_df)
## Number of regression paramaters
p <- 8
## Corrected Degrees of Freedom for Model
dfm <- p - 1
## Degrees of Freedom for Error
dfe <- n - p
## Corrected Degrees of Freedom Total: DFT = n - 1
dft <- n - 1

## Mean of Squares for Model: MSM = SSM / DFM
msm <- ssm / dfm
## Mean of Squares for Error: MSE = SSE / DFE
mse <- sse / dfe
## Mean of Squares Total: MST = SST / DFT
mst <- sst / dft
## F Statistic
f_score <- msm / mse

```

```
## Adjusted R Squared  $R^2 = 1 - (1 - R^2) (n - 1) / (n - p)$   
adjusted_r_squared <- 1 - (1 - r_squared) * (n - 1) / (n - p)
```