

Assignment 6

Theodore Koby-Hercsky

5/3/2021

html_document: <https://rpubs.com/theoKoby/767033>

Set the working directory to the root of your DSC 520 directory

```
setwd("~/Documents/Bellevue University Classes/DSC520/assignments/assignment06")
```

The heights.csv file

I am importing readr from the library so I can use the read_csv function to create my student survey data frame.

```
library(readr)
```

Creating the student survey data frame by using the read_csv function to pull my student survey data.

```
heights_df <- read_csv("data/r4ds/heights.csv")
```

Load the ggplot2 library

```
library(ggplot2)
```

Fit a linear model using the `age` variable as the predictor and `earn` as the outcome

```
age_lm <- lm(earn ~ age, data = heights_df)
```

When we fit our linear model by using our age variable as the predictor and earn as the outcome with the heights_df as our data we see coefficients for intercept and age which is the slope for the age predictor.

```
age_lm
```

```
##
```

```
## Call:
```

```
## lm(formula = earn ~ age, data = heights_df)
```

```
##
```

```
## Coefficients:
```

```
## (Intercept)          age
```

```
##    19041.53         99.41
```

I will view the full report by using summary of my age_lm model

As seen below we see that estimated salary to be \$19,041.53 which is in our intercept.

While we see three stars against regression coefficients that implies the independent variable age is highly correlated with dependent variable earn with a linear relationship.

```
summary(age_lm)
```

```
##
## Call:
## lm(formula = earn ~ age, data = heights_df)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -25098 -12622  -3667   6883 177579
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 19041.53    1571.26  12.119 < 2e-16 ***
## age          99.41       35.46   2.804  0.00514 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 19420 on 1190 degrees of freedom
## Multiple R-squared:  0.006561, Adjusted R-squared:  0.005727
## F-statistic: 7.86 on 1 and 1190 DF, p-value: 0.005137

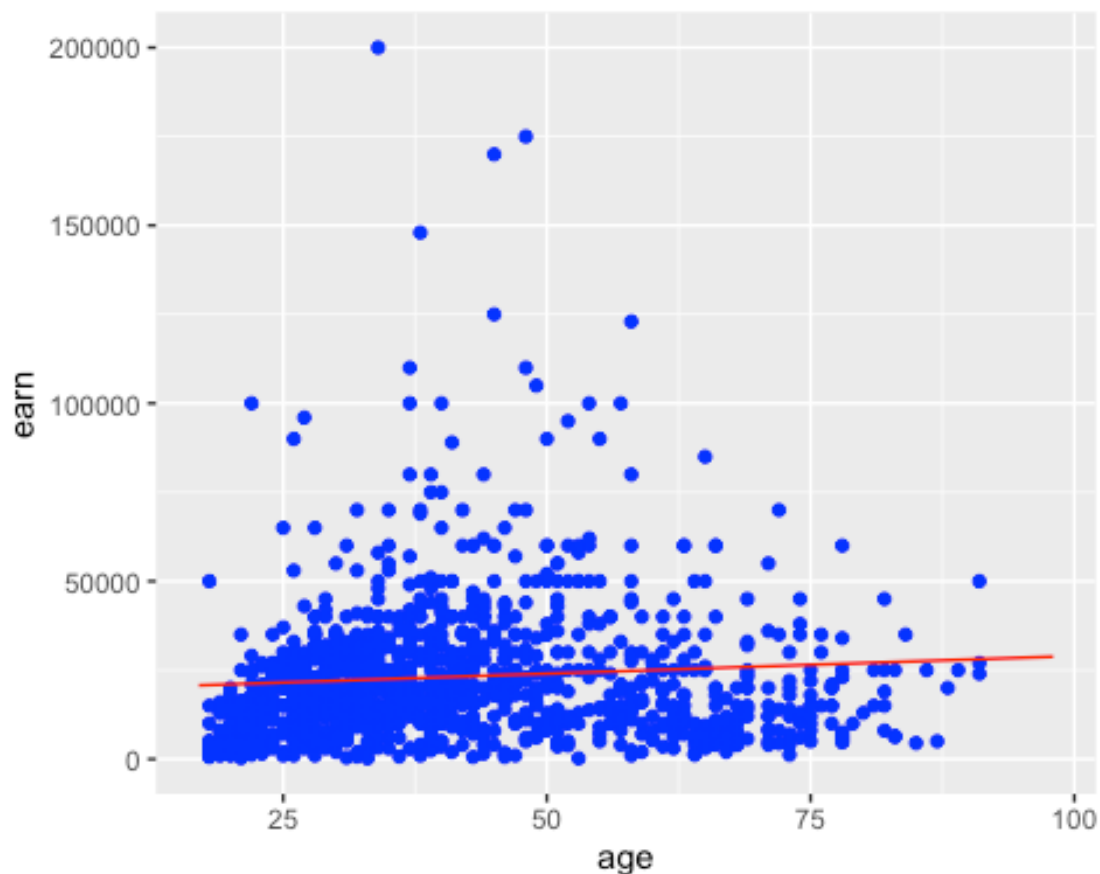
## Created a new age test data frame to incorporate different ages
newage_test_df <- data.frame(age = c(17,19,25,32,59,61,68,75,84,98))
## Created an age list that is a sample of the 1192 from our heights data frame.
age_list <- as.data.frame(sample(18:92, 1192, replace = TRUE))
## Creating predictions using `predict()`
## in this prediction I use the newage_test_df set to equal age and earn set
to predict from our age_lm and the new data that was created in the data frame
newage_test_df.
age_predict_df <- data.frame(age = newage_test_df, earn=predict(age_lm, newdata =
newage_test_df))
## The age_predict_df shows 10 ages with a predicted earnings which we see as
age goes up the predicted earnings goes up.
age_predict_df

##      age      earn
## 1    17 20731.42
## 2    19 20930.24
## 3    25 21526.67
## 4    32 22222.51
## 5    59 24906.47
## 6    61 25105.28
## 7    68 25801.12
## 8    75 26496.96
## 9    84 27391.61
## 10   98 28783.29

## Load the ggplot2 library
library(ggplot2)
## Plot the predictions against the original data
## Created a ggplot that shows our predictions against our original data as w
```

e see an upward movement in earnings as age goes up in our prediction

```
ggplot(data = heights_df, aes(y = earn, x = age)) + geom_point(color='blue')  
+ geom_line(color='red', data = age_predict_df, aes(y = earn, x = age))
```



I am finding the mean by using the mean function to create the mean earn

```
mean_earn <- mean(heights_df$earn)
```

The mean for earn in our heights data set is \$23,154.77

```
mean_earn
```

```
## [1] 23154.77
```

Corrected Sum of Squares Total

```
sst <- sum((mean_earn - heights_df$earn)^2)
```

As seen below we receive a value of 451591883937 which is the sum of squares total for the mean of earn from our heights data set. Which is known as the total amount of differences presented in a basic model that represents the how good the mean is as a model of the observed data.

```
sst
```

```
## [1] 451591883937
```

Corrected Sum of Squares for Model

As seen below we calculate the model sum of squares by taking the mean_earn that we calculated and minus it by our predicted earnings to the second. Wh

ich measures the deviation of data points away from the mean value. As our result indicates a large degree of variability within the data set.

```
ssm <- sum((mean_earn - age_predict_df$earn)^2)
```

The model sum of squares is seen as 89018275

```
ssm
```

```
## [1] 89018275
```

Residuals

We can see our residuals for each and every earnings in the age predict data frame.

```
residuals <- heights_df$earn - age_predict_df$earn
```

```
residuals
```

```
##      [1] 29268.57576 39069.76422 8473.32959 27777.48919 26093.53337
##      [6] -16105.27818 3198.88142 5503.04103 -25391.61092 -1783.29171
##     [11] -14201.42424 9069.76422 -9526.67041 -10222.51081 -2906.46663
##     [16] -8105.27818 14198.88142 17503.04103 -20391.61092 24216.70829
##     [21] -15731.42424 -6930.23578 -16026.67041 17777.48919 9093.53337
##     [26] -15105.27818 1198.88142 23503.04103 13608.38908 -13783.29171
##     [31] 4268.57576 54069.76422 5473.32959 -10222.51081 -17406.46663
##     [36] 4894.72182 -4801.11858 503.04103 -24391.61092 -3783.29171
##     [41] 3268.57576 11069.76422 -11526.67041 -11222.51081 -6206.46663
##     [46] -5105.27818 -22301.11858 -13496.95897 -2391.61092 -7783.29171
##     [51] 13268.57576 -14930.23578 -4526.67041 12777.48919 -20906.46663
##     [56] -11105.27818 -15801.11858 -1496.95897 -11391.61092 -12783.29171
##     [61] -4231.42424 -16930.23578 -17686.67041 -222.51081 -24706.46663
##     [66] 894.72182 -23301.11858 -9496.95897 -19391.61092 -16783.29171
##     [71] -10731.42424 -10930.23578 -6526.67041 -19822.51081 5093.53337
##     [76] 4894.72182 -15801.11858 -21496.95897 -15391.61092 -8783.29171
##     [81] -731.42424 -930.23578 -20326.67041 -21522.51081 -4906.46663
##     [86] -15105.27818 4198.88142 13503.04103 -2391.61092 -18783.29171
##     [91] 39268.57576 -2930.23578 -5486.67041 -7222.51081 -14906.46663
##     [96] 7894.72182 -7801.11858 -11496.95897 -6391.61092 -7783.29171
##    [101] 16268.57576 17069.76422 -4526.67041 9777.48919 2593.53337
##    [106] -8605.27818 -801.11858 503.04103 -22391.61092 41216.70829
##    [111] -15731.42424 -15930.23578 -1526.67041 -18222.51081 35093.53337
##    [116] -20105.27818 4198.88142 43503.04103 22608.38908 15216.70829
##    [121] 9268.57576 -10930.23578 1473.32959 22777.48919 -9906.46663
##    [126] -21105.27818 -8801.11858 3503.04103 108.38908 -23095.29171
##    [131] -2731.42424 22069.76422 10473.32959 -12222.51081 35093.53337
##    [136] -4105.27818 -23401.11858 -25496.95897 -391.61092 -22183.29171
##    [141] -4731.42424 69069.76422 -13526.67041 -2222.51081 -9906.46663
##    [146] -13105.27818 -1801.11858 -6496.95897 -8391.61092 -18783.29171
##    [151] 19268.57576 4069.76422 3473.32959 2777.48919 -5906.46663
##    [156] 18894.72182 -10801.11858 -9496.95897 -3391.61092 -5783.29171
##    [161] -7731.42424 44069.76422 -14526.67041 17777.48919 -9906.46663
##    [166] -5105.27818 -5801.11858 -6496.95897 -2391.61092 20216.70829
##    [171] 4268.57576 -9930.23578 -5526.67041 12777.48919 100093.53337
##    [176] -2105.27818 -8801.11858 503.04103 42608.38908 6216.70829
```

##	[181]	-10731.42424	14069.76422	-6526.67041	-10222.51081	-16906.46663
##	[186]	-17105.27818	9198.88142	18503.04103	-12391.61092	-13783.29171
##	[191]	3268.57576	4069.76422	3473.32959	-2222.51081	-906.46663
##	[196]	18894.72182	43198.88142	35503.04103	4608.38908	-8783.29171
##	[201]	11268.57576	4069.76422	148473.32959	12777.48919	15093.53337
##	[206]	7894.72182	-7801.11858	3503.04103	-1391.61092	-23783.29171
##	[211]	-731.42424	-3930.23578	10473.32959	-7222.51081	25093.53337
##	[216]	-17105.27818	14198.88142	13503.04103	5358.38908	-8783.29171
##	[221]	15268.57576	-14930.23578	-9526.67041	37777.48919	15093.53337
##	[226]	17894.72182	19198.88142	-20496.95897	-19391.61092	-8783.29171
##	[231]	-3731.42424	-18930.23578	43473.32959	27777.48919	-13906.46663
##	[236]	9894.72182	1198.88142	-22996.95897	14608.38908	-8783.29171
##	[241]	-5731.42424	-10930.23578	-15526.67041	-5222.51081	10093.53337
##	[246]	2894.72182	-10801.11858	-6496.95897	-7391.61092	-18283.29171
##	[251]	-7731.42424	-10930.23578	-18526.67041	1777.48919	-7906.46663
##	[256]	-14105.27818	6198.88142	-9496.95897	-24391.61092	-26683.29171
##	[261]	-17539.42424	-3930.23578	8473.32959	-7222.51081	-906.46663
##	[266]	-15105.27818	24198.88142	23503.04103	-7391.61092	1216.70829
##	[271]	1268.57576	6069.76422	-17526.67041	-4722.51081	-8406.46663
##	[276]	2894.72182	26198.88142	-11496.95897	-8391.61092	-1783.29171
##	[281]	-5731.42424	-5930.23578	-7026.67041	1777.48919	-6906.46663
##	[286]	-21105.27818	-21801.11858	-25796.95897	-3391.61092	-1783.29171
##	[291]	-8731.42424	1069.76422	8473.32959	12777.48919	-4906.46663
##	[296]	6894.72182	-19801.11858	-14496.95897	-17391.61092	-27783.29171
##	[301]	-8731.42424	-4930.23578	3473.32959	2777.48919	7093.53337
##	[306]	9894.72182	-19801.11858	53503.04103	-7391.61092	-27783.29171
##	[311]	7268.57576	5069.76422	5473.32959	-4222.51081	-22906.46663
##	[316]	-5105.27818	-19801.11858	-1496.95897	-15391.61092	-16783.29171
##	[321]	9268.57576	-16930.23578	4473.32959	777.48919	-3906.46663
##	[326]	-105.27818	-7801.11858	-22496.95897	-3491.61092	6216.70829
##	[331]	5268.57576	4069.76422	1973.32959	-10222.51081	-9906.46663
##	[336]	-22105.27818	-19801.11858	-12496.95897	-12391.61092	146216.70829
##	[341]	-19731.42424	-10930.23578	23473.32959	-7222.51081	-4906.46663
##	[346]	9894.72182	14198.88142	8503.04103	22608.38908	71216.70829
##	[351]	14268.57576	3069.76422	13473.32959	15777.48919	5093.53337
##	[356]	-20105.27818	122198.88142	3503.04103	-20891.61092	-25783.29171
##	[361]	2268.57576	19069.76422	8473.32959	-8222.51081	-9906.46663
##	[366]	894.72182	-17801.11858	-2496.95897	-22267.61092	-18783.29171
##	[371]	-8731.42424	29069.76422	1473.32959	17777.48919	-18906.46663
##	[376]	-20105.27818	84198.88142	14503.04103	-4391.61092	-7783.29171
##	[381]	-16731.42424	4069.76422	8473.32959	-8222.51081	-18906.46663
##	[386]	-11105.27818	17198.88142	-1496.95897	12608.38908	36216.70829
##	[391]	-4731.42424	-12930.23578	-1526.67041	-8222.51081	20093.53337
##	[396]	-17105.27818	-13801.11858	-18496.95897	-11391.61092	-22783.29171
##	[401]	-1731.42424	69.76422	21473.32959	12777.48919	-16906.46663
##	[406]	-4105.27818	-20001.11858	-9496.95897	-3391.61092	-23783.29171
##	[411]	-9731.42424	-10930.23578	18473.32959	17777.48919	-906.46663
##	[416]	-1105.27818	-5801.11858	33503.04103	-16391.61092	-23783.29171
##	[421]	6268.57576	68069.76422	18473.32959	-7222.51081	-23906.46663
##	[426]	-3105.27818	-13801.11858	-4496.95897	12608.38908	11216.70829

##	[431]	14268.57576	-13930.23578	-7526.67041	-7222.51081	-18906.46663
##	[436]	-14105.27818	-15801.11858	-6496.95897	-20391.61092	9216.70829
##	[441]	8268.57576	-930.23578	-13526.67041	13777.48919	-23406.46663
##	[446]	16894.72182	-801.11858	-13496.95897	2608.38908	-21783.29171
##	[451]	10268.57576	-5930.23578	28473.32959	-17222.51081	-24306.46663
##	[456]	-11105.27818	14198.88142	-2496.95897	-5391.61092	-8783.29171
##	[461]	-619.42424	-8930.23578	19473.32959	-9222.51081	35093.53337
##	[466]	-1105.27818	12198.88142	-23496.95897	-14391.61092	-23783.29171
##	[471]	-731.42424	1069.76422	6473.32959	-222.51081	5093.53337
##	[476]	4894.72182	4198.88142	5503.04103	-1391.61092	-6783.29171
##	[481]	4268.57576	-9930.23578	-8526.67041	-12222.51081	-12906.46663
##	[486]	-18105.27818	31198.88142	9503.04103	1608.38908	-20783.29171
##	[491]	-17731.42424	-930.23578	11473.32959	-7222.51081	-8906.46663
##	[496]	-20105.27818	-17801.11858	5503.04103	-12891.61092	-15783.29171
##	[501]	-13731.42424	29069.76422	-9526.67041	-3222.51081	-15906.46663
##	[506]	-5105.27818	29198.88142	-16496.95897	-14391.61092	-8783.29171
##	[511]	9268.57576	-8930.23578	4473.32959	-9722.51081	-16906.46663
##	[516]	-10105.27818	39198.88142	-1496.95897	-19391.61092	-13783.29171
##	[521]	-8731.42424	14069.76422	18473.32959	6777.48919	-9906.46663
##	[526]	-5105.27818	-22801.11858	-25496.95897	-19391.61092	-23783.29171
##	[531]	4268.57576	-930.23578	8473.32959	-10222.51081	-14906.46663
##	[536]	-9105.27818	19198.88142	13503.04103	-2391.61092	-9783.29171
##	[541]	-3131.42424	-13930.23578	8473.32959	-4222.51081	93.53337
##	[546]	-19105.27818	44198.88142	-11496.95897	-2391.61092	6216.70829
##	[551]	-2731.42424	7069.76422	-6526.67041	-222.51081	10093.53337
##	[556]	-7105.27818	9198.88142	-10096.95897	-5391.61092	1216.70829
##	[561]	-3731.42424	4069.76422	-11526.67041	17777.48919	-12406.46663
##	[566]	9894.72182	79198.88142	73503.04103	-6391.61092	-18783.29171
##	[571]	12268.57576	5069.76422	38473.32959	-4222.51081	-9906.46663
##	[576]	-20105.27818	34198.88142	-14496.95897	608.38908	13216.70829
##	[581]	-18731.42424	-930.23578	6473.32959	-12222.51081	6093.53337
##	[586]	-8105.27818	-9801.11858	8503.04103	-24391.61092	-12783.29171
##	[591]	14268.57576	6069.76422	3473.32959	17777.48919	20093.53337
##	[596]	-11105.27818	-801.11858	13503.04103	6608.38908	11216.70829
##	[601]	14268.57576	-8930.23578	8473.32959	777.48919	-22906.46663
##	[606]	-5105.27818	-15801.11858	33503.04103	-15391.61092	-3783.29171
##	[611]	1268.57576	-12930.23578	-19526.67041	-3622.51081	-4906.46663
##	[616]	74894.72182	24198.88142	503.04103	-26391.61092	-10783.29171
##	[621]	-3731.42424	-10930.23578	-7526.67041	-15022.51081	8093.53337
##	[626]	-2105.27818	-15801.11858	-4496.95897	-14391.61092	4216.70829
##	[631]	15268.57576	-10930.23578	-15526.67041	-10222.51081	-14906.46663
##	[636]	-4105.27818	198.88142	-4496.95897	-4391.61092	-8783.29171
##	[641]	-10731.42424	-5930.23578	-14526.67041	-7222.51081	1093.53337
##	[646]	-15105.27818	-22801.11858	503.04103	608.38908	-13783.29171
##	[651]	34268.57576	-930.23578	-13026.67041	7777.48919	-4906.46663
##	[656]	-3105.27818	-801.11858	-23496.95897	-11391.61092	-18783.29171
##	[661]	-5731.42424	7069.76422	9973.32959	72777.48919	13093.53337
##	[666]	4894.72182	10098.88142	-14496.95897	-17391.61092	16216.70829
##	[671]	19268.57576	569.76422	-7526.67041	15777.48919	-23906.46663
##	[676]	-11105.27818	-10801.11858	-2496.95897	-22891.61092	-10783.29171

##	[681]	-6731.42424	-10930.23578	-12026.67041	-11222.51081	-7906.46663
##	[686]	-19105.27818	-23301.11858	13503.04103	-2391.61092	-783.29171
##	[691]	12268.57576	-16930.23578	-17426.67041	2777.48919	7093.53337
##	[696]	-12105.27818	-17801.11858	-9496.95897	-12391.61092	-10783.29171
##	[701]	-8731.42424	-13930.23578	-20326.67041	-8222.51081	-7906.46663
##	[706]	4894.72182	-801.11858	-4496.95897	-12391.61092	-1783.29171
##	[711]	-731.42424	-11930.23578	-4526.67041	-20422.51081	-6906.46663
##	[716]	-105.27818	-5801.11858	48503.04103	-23891.61092	7216.70829
##	[721]	9268.57576	4069.76422	-20526.67041	37777.48919	10093.53337
##	[726]	-16105.27818	9198.88142	-15496.95897	-15391.61092	1216.70829
##	[731]	-1731.42424	-11930.23578	8473.32959	-13222.51081	-9906.46663
##	[736]	-17105.27818	-15801.11858	6503.04103	-15391.61092	-10783.29171
##	[741]	9268.57576	-8430.23578	-11526.67041	-13722.51081	-3906.46663
##	[746]	-4105.27818	-4801.11858	-21496.95897	-5391.61092	1216.70829
##	[751]	102268.57576	-12930.23578	-19526.67041	-14622.51081	-14906.46663
##	[756]	9894.72182	-10801.11858	-22496.95897	-11391.61092	-28583.29171
##	[761]	2268.57576	-4930.23578	-15526.67041	-21622.51081	3093.53337
##	[766]	-21105.27818	-10801.11858	-11496.95897	-2391.61092	-13783.29171
##	[771]	19268.57576	9069.76422	-16526.67041	12777.48919	-8906.46663
##	[776]	2894.72182	-5801.11858	-5496.95897	-10391.61092	-4783.29171
##	[781]	4268.57576	19069.76422	-13526.67041	-3222.51081	20093.53337
##	[786]	-20105.27818	-1801.11858	-21496.95897	5608.38908	-4783.29171
##	[791]	-14731.42424	69069.76422	35473.32959	-8422.51081	5093.53337
##	[796]	-10105.27818	-10801.11858	4503.04103	5608.38908	1216.70829
##	[801]	-8731.42424	-15930.23578	-526.67041	5777.48919	-12906.46663
##	[806]	-6105.27818	-19801.11858	-21196.95897	-12391.61092	-11783.29171
##	[811]	-16031.42424	-11430.23578	-19826.67041	-16822.51081	-14906.46663
##	[816]	-105.27818	9198.88142	1503.04103	14608.38908	16216.70829
##	[821]	34268.57576	15069.76422	3473.32959	-3222.51081	13093.53337
##	[826]	-10105.27818	-21301.11858	69503.04103	-16391.61092	-19783.29171
##	[831]	-15731.42424	29069.76422	-15526.67041	77777.48919	-8406.46663
##	[836]	6894.72182	4198.88142	-16496.95897	-12391.61092	-22283.29171
##	[841]	5268.57576	-3930.23578	473.32959	-12222.51081	25093.53337
##	[846]	-5105.27818	-25401.11858	-10496.95897	-17391.61092	-14783.29171
##	[851]	-15731.42424	19069.76422	-16526.67041	2777.48919	-12906.46663
##	[856]	32894.72182	-7801.11858	-25496.95897	-6391.61092	-17783.29171
##	[861]	-7731.42424	-930.23578	3473.32959	-21622.51081	93.53337
##	[866]	10894.72182	-10801.11858	-22996.95897	-8386.61092	33216.70829
##	[871]	4268.57576	29069.76422	18473.32959	3777.48919	-2906.46663
##	[876]	-4105.27818	-19693.11858	3503.04103	-21391.61092	-1783.29171
##	[881]	-16731.42424	-16514.23578	-16526.67041	22777.48919	-6906.46663
##	[886]	-19105.27818	-11801.11858	8503.04103	-12391.61092	-16783.29171
##	[891]	-19731.42424	11069.76422	-20026.67041	-19222.51081	10093.53337
##	[896]	44894.72182	-4801.11858	23503.04103	-12391.61092	-19783.29171
##	[901]	-10731.42424	9069.76422	-5526.67041	-2222.51081	75093.53337
##	[906]	-15105.27818	9198.88142	-19496.95897	-19391.61092	-17783.29171
##	[911]	4268.57576	4069.76422	8473.32959	25777.48919	-7906.46663
##	[916]	894.72182	-15801.11858	33503.04103	-5391.61092	-22583.29171
##	[921]	3268.57576	-930.23578	-14526.67041	7777.48919	-1906.46663
##	[926]	-5105.27818	-22801.11858	3503.04103	-15391.61092	11216.70829

##	[931]	-5731.42424	15069.76422	-11526.67041	-9222.51081	-14906.46663
##	[936]	4894.72182	-2801.11858	-6496.95897	-8391.61092	-8783.29171
##	[941]	7268.57576	-7930.23578	-9526.67041	-2222.51081	93.53337
##	[946]	-105.27818	-7801.11858	-20496.95897	-11391.61092	16216.70829
##	[951]	4268.57576	-8930.23578	-19526.67041	-12222.51081	-10906.46663
##	[956]	-10105.27818	-15801.11858	-25296.95897	-12391.61092	-18783.29171
##	[961]	-18731.42424	-16930.23578	-11526.67041	-20722.51081	-6906.46663
##	[966]	-21105.27818	2198.88142	-16496.95897	-24391.61092	-16783.29171
##	[971]	-1731.42424	15069.76422	-14526.67041	12777.48919	15093.53337
##	[976]	-10105.27818	-17801.11858	-1496.95897	608.38908	-6783.29171
##	[981]	768.57576	-17930.23578	1473.32959	-4222.51081	-4906.46663
##	[986]	-9105.27818	-10801.11858	26503.04103	-1391.61092	-24783.29171
##	[991]	-10731.42424	22069.76422	-4526.67041	-18222.51081	25093.53337
##	[996]	29894.72182	2198.88142	-21496.95897	-9391.61092	-7783.29171
##	[1001]	-8731.42424	12069.76422	28473.32959	-16222.51081	-8906.46663
##	[1006]	9894.72182	-801.11858	-6496.95897	-9391.61092	5216.70829
##	[1011]	4268.57576	-930.23578	11473.32959	-18222.51081	-4906.46663
##	[1016]	14894.72182	-11801.11858	-14496.95897	-2391.61092	41216.70829
##	[1021]	17268.57576	-6930.23578	-6526.67041	-5222.51081	35093.53337
##	[1026]	-10105.27818	-801.11858	-8496.95897	42608.38908	1216.70829
##	[1031]	-12731.42424	9069.76422	5473.32959	24777.48919	-12906.46663
##	[1036]	-1105.27818	-12801.11858	53503.04103	-22391.61092	-3783.29171
##	[1041]	7268.57576	-12930.23578	-11526.67041	-13222.51081	-4906.46663
##	[1046]	-13105.27818	-10801.11858	-6496.95897	608.38908	-28518.29171
##	[1051]	14268.57576	5069.76422	5473.32959	-8222.51081	7093.53337
##	[1056]	-15105.27818	-18801.11858	1503.04103	32608.38908	-16783.29171
##	[1061]	-5731.42424	-5930.23578	4473.32959	3777.48919	-6186.46663
##	[1066]	-13105.27818	54198.88142	21503.04103	-9391.61092	56216.70829
##	[1071]	14268.57576	-15930.23578	-12526.67041	7777.48919	55093.53337
##	[1076]	9894.72182	-13801.11858	173503.04103	-23391.61092	-26783.29171
##	[1081]	3268.57576	-18930.23578	-6526.67041	-16222.51081	-19906.46663
##	[1086]	-23605.27818	2198.88142	-17496.95897	32608.38908	-27583.29171
##	[1091]	-19331.42424	-13930.23578	-526.67041	21777.48919	4093.53337
##	[1096]	-22105.27818	-13801.11858	-6496.95897	-26191.61092	8216.70829
##	[1101]	-12731.42424	14069.76422	3473.32959	-4222.51081	5093.53337
##	[1106]	-5105.27818	-13801.11858	-16496.95897	-9391.61092	-8783.29171
##	[1111]	10268.57576	-14930.23578	-9526.67041	-9222.51081	1093.53337
##	[1116]	-10105.27818	4198.88142	-24496.95897	-22391.61092	24216.70829
##	[1121]	-18731.42424	4069.76422	-16526.67041	-1222.51081	-7906.46663
##	[1126]	-13105.27818	16198.88142	-5496.95897	-21391.61092	-3783.29171
##	[1131]	-17731.42424	-17330.23578	-17526.67041	-2222.51081	27093.53337
##	[1136]	34894.72182	-10801.11858	63503.04103	-1391.61092	-26783.29171
##	[1141]	-4731.42424	3069.76422	-6526.67041	19777.48919	-14906.46663
##	[1146]	-8105.27818	4198.88142	5503.04103	8608.38908	-18783.29171
##	[1151]	-731.42424	3069.76422	-4526.67041	777.48919	93.53337
##	[1156]	24894.72182	-5801.11858	3503.04103	-20391.61092	-21783.29171
##	[1161]	9268.57576	-17930.23578	-1526.67041	17777.48919	-14906.46663
##	[1166]	-9105.27818	-14801.11858	-10496.95897	-9391.61092	-16783.29171
##	[1171]	-16731.42424	39069.76422	21473.32959	8777.48919	25093.53337
##	[1176]	1894.72182	4198.88142	-14496.95897	-7391.61092	-13783.29171


```
## [1181] 4268.57576 -18930.23578 -18526.67041 87777.48919 30093.53337
## [1186] 32894.72182 -15801.11858 -7496.95897 -12391.61092 -20783.29171
## [1191] 39268.57576 -14930.23578

## Sum of Squares for Error
## Next we calculate the Sum of Squares for Error which is the deviations predicted from actual empirical values of data. As the error is the difference between the observed value and the predicted value which we want to be as minimal as possible.
sse <- sum(residuals^2)
sse

## [1] 458035419167

## R Squared  $R^2 = SSM/SST$ 
## In R Squared we take the Corrected Sum of Squares for Model (SSM) and divide it by Corrected Sum of Squares Total (sst) Which is a statistical measure of how well the regression predictions approximate the real data points
r_squared <- ssm / sst
r_squared

## [1] 0.0001971211

## Number of observations
## I use the nrow function to determine the number of observations in our heights data set
n <- NROW(heights_df$earn)
n

## [1] 1192

## Number of regression parameters
## As seen below we set the number of regression parameters to two
p <- 2
p

## [1] 2

## Corrected Degrees of Freedom for Model (p-1)
## To get the Corrected Degrees of Freedom for Model we take the Number of regression parameters and minus it by one
dfm <- p - 1
dfm

## [1] 1

## Degrees of Freedom for Error (n-p)
## to calculate the Degrees of Freedom for Error we take the Number of observations and minus it by the Number of regression parameters. As the Corrected Degrees of Freedom Total shows the number of independent values that can vary in an analysis without breaking any constraints
```

```

dfe <- n - p
dfe

## [1] 1190

## Corrected Degrees of Freedom Total: DFT = n - 1
## When calculating the Corrected Degrees of Freedom Total we take the Number
of observations and minus it by one As the Corrected Degrees of Freedom Total
shows the number of independent values that can vary in an analysis without b
reaking any constraints that have been corrected.
dft <- n - 1
dft

## [1] 1191

## Mean of Squares for Model: MSM = SSM / DFM
## When calculating the Mean of Squares for Model we take the Corrected Sum o
f Squares for Model and divide it by the Corrected Degrees of Freedom for Mod
el which gives us 89018275
msm <- ssm / dfm
msm

## [1] 89018275

## Mean of Squares for Error: MSE = SSE / DFE
## When we calculate the Mean of Squares for Error we are taking the Sum of S
quares for Error and dividing it by the Degrees of Freedom for Error which gi
ves us 384903714
mse <- sse / dfe
mse

## [1] 384903714

## Mean of Squares Total: MST = SST / DFT
## When we calculate the Mean of Squares Total we take the Corrected Sum of S
quares Total and divide it by the Corrected Degrees of Freedom Total which gi
ves us 379170348
mst <- sst / dft
mst

## [1] 379170348

## F Statistic F = MSM/MSE
## When we calculate the F Statistic we take the Mean of Squares for Model an
d divide by the Mean of Squares Total to get 0.2312741 This shows that the co
efficients that was used in the model improved the model's fit by a small per
cent seen as 0.2312741.
f_score <- msm / mse
f_score

## [1] 0.2312741

```

```

## Adjusted R Squared  $R^2 = 1 - (1 - R^2)(n - 1) / (n - p)$ 
## When we calculate the Adjusted R Squared we take one and minus it by (one
- R Squared  $R^2$ ) then times it by (Number of observations - one) then divide it
by (Number of observations - Number of regression parameters) which gives us
-0.0006430494. As Adjusted  $R^2$  indicates how well terms fit a curve or line
as we see it is a negative percentage meaning it does not fit well with the curve
adjusted_r_squared <- 1 - (1 - r_squared)*(n-1) / (n-p)
adjusted_r_squared

## [1] -0.0006430494

## Calculate the p-value from the F distribution
## when we calculate the p-value from the F distribution we get 0.630669 Which
means the results have a 63.06% probability of being completely random and
not due to anything in the experiment
p_value <- pf(f_score, dfm, dft, lower.tail=F)
p_value

## [1] 0.630669

```