ST 518 Homework 2

Eric Warren

September 8, 2023

Contents

1 Problem 1

Read in the data to use for the problem

```
library(tidyverse)
houses <- read.csv("~/ST-518/houses.csv", header = F)
names <- unlist(strsplit(houses[1, 1], " "), recursive = F)
names <- names[2:length(names)]
houses <- houses[2:nrow(houses), ]
colnames(houses) <- names
houses <- as_tibble(lapply(houses, as.numeric))
houses
```

```
## # A tibble: 239 x 4
##    price_k bedrooms bathrooms sqftLiving
##      <dbl>    <dbl>     <dbl>      <dbl>
##  1   2000         3      2.75       3050
##  2    937         3      1.75       2450
##  3    920         5      2.25       2730
##  4    650         3      2.25       2150
##  5    335         2      1.75       1030
##  6    660.        4      2.25       2010
##  7    675         4      3.5        2140
##  8    890         4      1          2550
##  9    673         4      2.25       2590
## 10    269         4      1.75       1490
## # i 229 more rows
```

1.1 Part A

Fit a simple linear regression of `price` on `bedrooms`.

```
priceForBedrooms <- lm(price_k ~ bedrooms, houses)
summary(priceForBedrooms)
```

```
##
## Call:
## lm(formula = price_k ~ bedrooms, data = houses)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -337.59 -162.13  -49.41  116.59 1542.87
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   223.95      63.02   3.554 0.000458 ***
## bedrooms       77.73      18.32   4.244 3.16e-05 ***
## ---
## Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1
##
## Residual standard error: 239.7 on 237 degrees of freedom
## Multiple R-squared:  0.07062,    Adjusted R-squared:  0.0667
## F-statistic: 18.01 on 1 and 237 DF,  p-value: 3.157e-05
```

1.1.1 Part i

As we can see from our summary chart, we can see that the least squares estimate for our slope (under the `estimates` column and `bedrooms` row) is 77.7267421.

1.1.2 Part ii

As we can see from our summary chart, we can see that the standard error for our slope (or $SE(\hat{\beta_1})$) is 18.32. This is found by lining up the `Std. Error` column with the `bedrooms` row.

1.1.3 Part iii

The slope in the case of this problem is saying that for every one bedroom that is added to the house, we expect (or predict) the house value to go up by 77.7267421 thousands of dollars or 77726.74 dollars.

1.1.4 Part iv

We can use the regression model to estimate difference in average home price between 2- and 4-bedrooms homes by using the slope of the regression model as the intercept is a constant and will not change for a 2 or 4 bedroom house. Moreover, the equation is $\hat{y} = \hat{\beta_0} + \hat{\beta_1} x_{i1}$.

- For a 2 bedroom house, we are plugging in the $x_{i1} = 2$ so our equation is $\hat{y} = \hat{\beta_0} + \hat{\beta_1} * 2$
- For a 4 bedroom house, we are plugging in the $x_{i1} = 4$ so our equation is $\hat{y} = \hat{\beta_0} + \hat{\beta_1} * 4$

Thus we can say the difference between them is $(4-2)\hat{\beta_1} = 2 * \hat{\beta_1}$. We know that $\hat{\beta_1} = 7.7726742 \times 10^4$ (our slope was given in thousands of dollars), so our predicted difference in pricing between a 2 and 4 bedroom house is 2 * 77.7267421 = 155453.5 dollars more for the 4 bedroom house than a 2 bedroom house.

Now to find the standard error in the difference between a 2 and 4 bedroom house, we can say that there are two population means to consider:

- $$\mu(x=4) = \beta_0 + \beta_1(4)$$
- $$\mu(x=2) = \beta_0 + \beta_1(2)$$

As we can see, the difference between $\mu(4)$ and $\mu(2)$ is $2 * \beta_1$. The standard error of this difference then is $2 * SE(\hat{\beta_1})$. We saw from part ii that the standard error of the slope was 18.32, which is the same as saying the $SE(\hat{\beta_1})$. Thus, we should do 2 * 18.32 to get $2 * SE(\hat{\beta_1})$ which is 36.64. Therefore, the standard error for the estimated difference is 36.64.

1.2 Part B

First we are going to fit a multiple linear regression of `price` on `bedrooms` and `sqftLiving`.

```
priceForBedroomsAndLiving <- lm(price_k ~ bedrooms + sqftLiving,
    houses)
summary(priceForBedroomsAndLiving)
```

```
##
## Call:
## lm(formula = price_k ~ bedrooms + sqftLiving, data = houses)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -463.42 -125.51  -22.73   88.60 1275.70
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 230.03120   48.26269   4.766 3.28e-06 ***
## bedrooms    -55.46896   17.38723  -3.190  0.00161 **
## sqftLiving    0.21662    0.01671  12.965  < 2e-16 ***
## ---
## Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1
##
## Residual standard error: 183.5 on 236 degrees of freedom
## Multiple R-squared:  0.4572, Adjusted R-squared:  0.4526
## F-statistic: 99.39 on 2 and 236 DF,  p-value: < 2.2e-16
```

1.2.1 Part i

As we can see from our summary table, the least squares estimate of the partial slope for `bedrooms` is -55.4689588.

1.2.2 Part ii

We can see that the estimated partial slope for `bedrooms` is different from the estimated slope in **part (a)**. When adding another explanatory variable our partial slopes can be different from that when we do simple linear regression with our slope being the only variable in question. Before our interpretation for our simple linear regression on `bedrooms` was that for each additional bedroom that was added to the house we predicted that the house's price would go up by 77.7267421 thousands of dollars. Now with our multiple linear regression model, our partial slope for bedrooms is saying that if we keep our square footage constant (that is comparing houses with the same square footage) then we expect our house price to go down by 55.4689588 thousands of dollars for each one addtional bedroom added. In this new case, we have to make sure that we keep our square footage of a house at a constant so we can see just the affect bedrooms have. As in our simple linear regression case, we didn't have to worry about square footage as we were just looking at the number of bedrooms in the house and having that explain (and predict) the house value.

1.2.3 Part iii

Here we are going to formally compare this multiple linear regression with the simple linear regression from **part (a)** using either a T-test or an F-test. Then we are going to write out the models being compared, the hypothesis being tested and report the test statistic (T or F) and p-value.

The models we are comparing is the simple linear regression model which is $\hat{y} = \hat{\beta_0} + \hat{\beta_1} x_{i1}$ to our multiple linear regression model which is $\hat{y} = \hat{\beta_0} + \hat{\beta_1} x_{i1} + \hat{\beta_2} x_{i2}$. In this case we are going to use an F-test to see if we should keep our multiple regression model or go back to our simple linear regression model. The null hypothesis is $H_0 : \hat{\beta_2} = 0$ and our alternative hypothesis is $H_A : \hat{\beta_2} \neq 0$. We are going to look at our models by obtaining an F-statistic and comparing it to the critical value of $F_{\alpha/2,df_1,df_2}$, with $\alpha = .05$ for our statistical test.

The first is using ANOVA to find our F-value.

```
anova(priceForBedrooms, priceForBedroomsAndLiving)  #ANOVA table
```

```
## # A tibble: 2 x 6
##   Res.Df       RSS    Df ‘Sum of Sq‘     F  ‘Pr(>F)‘
##    <dbl>     <dbl> <dbl>       <dbl> <dbl>     <dbl>
## 1    237 13613626.    NA          NA    NA  NA
## 2    236  7950930.     1    5662696.  168.  2.20e-29
```

```
# Getting our F-critical value
qf(p = 0.05/2, df1 = length(priceForBedroomsAndLiving$coefficients) -
    length(priceForBedrooms$coefficients), df2 = nrow(houses) -
    length(priceForBedroomsAndLiving$coefficients), lower.tail = FALSE)
```

```
## [1] 5.088606
```

We can see here that our F-statistic is 168.0805171. If we use our *alpha level* to be .05 for our hypothesis test using the F-distribution, we would get an F-value correspondence that is 5.0886062. Therefore, since our F-statistic is greater than our F-critical value we have statistically significant evidence to reject our null hypothesis ($\hat{\beta}_2 = 0$). Moreover, we have significantly significant evidence that there is association between `price` and `sqftLiving` after accounting for dependence on `bedrooms`. Therefore, we should proceed with our multiple linear regression model.

1.2.4 Part iv

Here we are going to give an interpretation for the partial slope estimate for `sqftLiving`.

```
summary(priceForBedroomsAndLiving)
```

```
##
## Call:
## lm(formula = price_k ~ bedrooms + sqftLiving, data = houses)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -463.42 -125.51  -22.73   88.60 1275.70
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 230.03120   48.26269   4.766 3.28e-06 ***
## bedrooms    -55.46896   17.38723  -3.190  0.00161 **
## sqftLiving    0.21662    0.01671  12.965  < 2e-16 ***
## ---
## Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1
##
## Residual standard error: 183.5 on 236 degrees of freedom
## Multiple R-squared:  0.4572, Adjusted R-squared:  0.4526
## F-statistic: 99.39 on 2 and 236 DF,  p-value: < 2.2e-16
```

As we can see, our partial slope value for `sqftLiving` is 0.2166151. Our partial slope for `sqftLiving` is saying that if we keep the number of bedrooms in every house constant (that is comparing houses with the same number of bedrooms) then we expect (or predict) our house price to go up by 0.2166151 thousands of dollars (or to go up by a predicted 216.62 total dollars) for each additional one square foot added to the house.

1.3 Part C

Here using a separate random sample, multiple regression was fit with output below. For each pair of regression models for home price given below, report the F-ratio for a model comparison. Obtain the appropriate critical value and draw a conclusion from the model comparison. If the comparison is not possible from the provided output, write "NP". Note, for purposes of notation I used the term $MS[E]_f$ to denote the highest model for that situation which is why this varies per part of this problem. For example, in part (i) $MS[E]_f = MS[E]_{M2}$.

1.3.1 Part i

This is looking at Model 1, which is looking at a simple linear regression model with `price` fitted on `sqftLiving` (we can write this model as $\hat{y} = \beta_0 + \beta_1 * x_{i1}$) and comparing it to Model 2, which is a multiple linear regression model with `price` fitted on `sqftLiving` and `bedrooms` (we can write this model as $\hat{y} = \beta_0 + \beta_1 * x_{i1} + \beta_2 * x_{i2}$). To compare the models, our null hypothesis is $H_0 : \beta_2 = 0$ and our alternative hypothesis is $H_A : \beta_2 \neq 0$. In this case, we are going to evaluate our $R(\beta_2|\beta_0, \beta_1)$ which equals the Type I Sum of Squares for `bedrooms`, since this is in sequential order. As we can see where `bedrooms` and the `Type I SS` line up, we can see this value is 10961801. So, $R(\beta_2|\beta_0, \beta_1) = 10961801$. Using this, we are going to find its corresponding F-statistic. To find this, we are going to use the following formula: $F = \frac{R(\beta_2|\beta_0,\beta_1)}{MS[E]_f}$. From our output we can see that $MS[E]_f = \frac{SS[T]-R(\beta_1,\beta_2|\beta_0)}{df_f} = \frac{663612197-320863213-10961801}{4428-2-1} = \frac{331787183}{4425} = 74980.15$. We know that $df_f = n - p - 1 = 4428 - 2 - 1 = 4425$ since there are 4428 observations (n) and 2 predictors (p). We can get the sum of squares total from the output chart in the `Corrected Total` row with the `Sum of Squares` column. To get $R(\beta_1, \beta_2|\beta_0)$, we have to look at the Type I Sum of Squares for `sqftLiving` and `bedrooms` (since β_1 and β_2 are both sequential to β_0). Using that information, we can see that $F = \frac{R(\beta_2|\beta_0,\beta_1)}{MS[E]_f} = \frac{10961801}{74980.15} \approx 146.196$. Using our F-statistic being approximately equal to 146.196, we will compare this to our F-critical value of $F_{\alpha/2,df_1,df_2}$, with $\alpha = .05$ for our statistical test. Note, $df_1 = 1$ since that is how many predictors that are different between the two models (just β_2) and $df_2 = n-p-1 = 4428-2-1 = 4425$. Using these values, we can see our F-critical value of $F_{\alpha/2,df_1,df_2} = F_{.05/2,1,4425} = 5.0273$. As we can see our F-statistic is greater than our F-critical value so we should reject our null hypothesis and conclude that there is statistically significant evidence of association between our `price` and `bedrooms` after accounting for dependence on `sqftLiving`.

1.3.2 Part ii

"NP". This is because we cannot compare models that do not have a predictor variable in common. When looking at the simple linear regression model 1, we have the predictor `sqftLiving`, but the multiple linear regression model 3, we have the predictors `bedrooms` and `bathrooms`. With no predictors in common we cannot do any kind of statistical testing (F or T test).

1.3.3 Part iii

This is looking at Model 1, which is looking at a simple linear regression model with `price` fitted on `sqftLiving` (we can write this model as $\hat{y} = \beta_0 + \beta_1 * x_{i1}$) and comparing it to Model 2, which is a multiple linear regression model with `price` fitted on `sqftLiving` and `bathrooms` (we can write this model as $\hat{y} = \beta_0 + \beta_1 * x_{i1} + \beta_3 * x_{i3}$). To compare the models, our null hypothesis is $H_0 : \beta_3 = 0$ and our alternative hypothesis is $H_A : \beta_3 \neq 0$. In this case, we are going to evaluate our $R(\beta_3|\beta_0, \beta_1)$ which equals the Type I Sum of Squares or Type II Sum of Squares value for `bathrooms`, since this is the last value in our output and thus will equal each other. As we can see where `bathrooms` and the `Type I SS` line up, we can see this value is 109.62936. So, $R(\beta_3|\beta_0, \beta_1) = 109.62936$. Using this, we are going to find its corresponding F-statistic. To find this, we are going to use the following formula: $F = \frac{R(\beta_3|\beta_0,\beta_1)}{MS[E]_f}$. From our output we can see that $MS[E]_f = \frac{SS[T]-R(\beta_1,\beta_3|\beta_0)}{df_f} = \frac{663612197-320863213-109.62936}{4428-2-1} = \frac{342748874}{4425} = 77457.37$. We know that

$df_f = n - p - 1 = 4428 - 2 - 1 = 4425$ since there are 4428 observations (n) and 2 predictors (p). We can get the sum of squares total from the output chart in the `Corrected Total` row with the `Sum of Squares` column. To get $R(\beta_1, \beta_3|\beta_0)$, we have to look at the Type I Sum of Squares for `sqftLiving` and `bathrooms` (since β_1 is sequential to β_0 and β_3 is the last one in the output table). Using that information, we can see that $F = \frac{R(\beta_3|\beta_0,\beta_1)}{MS[E]_f} = \frac{109.62936}{77457.37} \approx 0.0014$. Using our F-statistic being approximately equal to 0.0014, we will compare this to our F-critical value of $F_{\alpha/2,df_1,df_2}$, with $\alpha = .05$ for our statistical test. Note, $df_1 = 1$ since that is how many predictors that are different between the two models (just β_3) and $df_2 = n - p - 1 = 4428 - 2 - 1 = 4425$. Using these values, we can see our F-critical value of $F_{\alpha/2,df_1,df_2} = F_{.05/2,1,4425} = 5.0273$. As we can see our F-statistic is less than our F-critical value so we should fail to reject our null hypothesis and conclude that there is not statistically significant evidence of association between our `price` and `bathrooms` after accounting for dependence on `sqftLiving`. Note, when our F-statistic is less than 1, we can automatically fail to reject our null hypothesis since F-critical values are at least 1.

1.3.4 Part iv

This is looking at Model 2, which is looking at a multiple linear regression model with `price` fitted on `sqftLiving` and `bedrooms`, (we can write this model as $\hat{y} = \beta_0 + \beta_1 * x_{i1} + \beta_2 * x_{i2}$) and comparing it to Model 5, which is a multiple linear regression model with `price` fitted on `sqftLiving`, `bedrooms` and `bathrooms` (we can write this model as $\hat{y} = \beta_0 + \beta_1 * x_{i1} + \beta_2 * x_{i2} + \beta_3 * x_{i3}$). To compare the models, our null hypothesis is $H_0 : \beta_3 = 0$ and our alternative hypothesis is $H_A : \beta_3 \neq 0$. In this case, we are going to evaluate our $R(\beta_3|\beta_0, \beta_1, \beta_2)$ which equals the Type I Sum of Squares or Type II Sum of Squares, since this is the last value in our output and thus will equal each other. As we can see where `bathrooms` and the `Type I SS` line up, we can see this value is 109.62936. So, $R(\beta_3|\beta_0, \beta_1, \beta_2) = 109.62936$. Using this, we are going to find its corresponding F-statistic. To find this, we are going to use the following formula: $F = \frac{R(\beta_3|\beta_0,\beta_1,\beta_2)/\Delta_{df}}{MS[E]_f}$. From our output we can see that $MS[E]_f = 74997$. We know that $\Delta_{df} = 1$ since there is one additional predictor we are evaluating. Using that information, we can see that $F = \frac{R(\beta_3|\beta_0,\beta_1,\beta_2)}{MS[E]_f} = \frac{109.62936/1}{74997} \approx 0.0015$. Using our F-statistic being approximately equal to 0.0015, we will compare this to our F-critical value of $F_{\alpha/2,df_1,df_2}$, with $\alpha = .05$ for our statistical test. Note, $df_1 = 1$ since that is how many predictors that are different between the two models (just β_3) and $df_2 = n - p - 1 = 4428 - 3 - 1 = 4424$. Using these values, we can see our F-critical value of $F_{\alpha/2,df_1,df_2} = F_{.05/2,1,4424} = 5.0273$. As we can see our F-statistic is less than our F-critical value so we should fail to reject our null hypothesis and conclude that there is not statistically significant evidence of association between our `price` and `bathrooms` after accounting for dependence on `sqftLiving` and `bedrooms`. Note, when our F-statistic is less than 1, we can automatically fail to reject our null hypothesis since F-critical values are at least 1.

1.3.5 Part v

This is looking at Model 4, which is looking at a multiple linear regression model with `price` fitted on `sqftLiving` and `bathrooms`, (we can write this model as $\hat{y} = \beta_0 + \beta_1 * x_{i1} + \beta_3 * x_{i3}$) and comparing it to Model 5, which is a multiple linear regression model with `price` fitted on `sqftLiving`, `bedrooms` and `bathrooms` (we can write this model as $\hat{y} = \beta_0 + \beta_1 * x_{i1} + \beta_2 * x_{i2} + \beta_3 * x_{i3}$). To compare the models, our null hypothesis is $H_0 : \beta_2 = 0$ and our alternative hypothesis is $H_A : \beta_2 \neq 0$. In this case, we are going to evaluate our $R(\beta_2|\beta_0, \beta_1, \beta_3)$ which equals the Type II Sum of Squares, this is **not** in sequential order and thus we are looking at just the partial value for β_2. As we can see where `bedrooms` and the `Type II SS` line up, we can see this value is 10662674. So, $R(\beta_2|\beta_0, \beta_1, \beta_3) = 10662674$. Using this, we are going to find its corresponding F-statistic. To find this, we are going to use the following formula: $F = \frac{R(\beta_2|\beta_0,\beta_1,\beta_3)/\Delta_{df}}{MS[E]_f}$. From our output we can see that $MS[E]_f = 74997$. We know that $\Delta_{df} = 1$ since there is one additional predictor we are evaluating. Using that information, we can see that $F = \frac{R(\beta_2|\beta_0,\beta_1,\beta_3)}{MS[E]_f} = \frac{10662674/1}{74997} \approx 142.1747$. Using our F-statistic being approximately equal to 142.1747, we will compare this to our F-critical value of $F_{\alpha/2,df_1,df_2}$, with $\alpha = .05$ for our statistical test. Note, $df_1 = 1$ since that is how many predictors that are different between the two models (just β_3) and $df_2 = n - p - 1 = 4428 - 3 - 1 = 4424$. Using these values, we can see our F-critical value of $F_{\alpha/2,df_1,df_2} = F_{.05/2,1,4424} = 5.0273$. As we can see our F-statistic is greater than our F-critical value so we should reject our null hypothesis and conclude that there is statistically significant evidence of association between our `price` and `bedrooms` after accounting for dependence on `sqftLiving` and `bathrooms`.

2 Problem 2

Consider a matrix formulation for Model 5 in problem 1(c). Use the output to answer questions below. Note: Model 5 was the multiple linear regression model with `price` fitted on `sqftLiving`, `bedrooms` and `bathrooms` (we can write this model as $\hat{y} = \beta_0 + \beta_1 * x_{i1} + \beta_2 * x_{i2} + \beta_3 * x_{i3}$).

2.1 Part A

Write out the model in matrix form, giving the dimension of every component of the model.

We know that a linear regression model in matrix form can be written as $Y = X\beta + E$, where

- Y denotes a response vector that has a dimension of (n x 1)
- X denotes a design matrix that has a dimension of (n x (p + 1))
- β denotes a vector of regression parameters that has a dimension of ((p + 1) x 1)
- E denotes an error vector that has a dimension of (n x 1), assumed $MVN(0, \sigma^2 I_n)$

Knowing this, we can write our multiple linear regression model using matrix form. Note that we have three predictors so $p = 3$ and there are 4428 observations (found by using $n = df_{Total} + 1 = 4427 + 1 = 4428$) so $n = 4428$. Thus our multiple linear regression model in matrix form can be written as $Y = X\beta + E$, where

- Y denotes a response vector that has a dimension of (4428 x 1)
- X denotes a design matrix that has a dimension of (4428 x (3 + 1)) or simplified as a dimension of (4428 x 4)
- β denotes a vector of regression parameters that has a dimension of ((3 + 1) x 1) or simplified as a dimension of (4 x 1)
- E denotes an error vector that has a dimension of (4428 x 1), assumed $MVN(0, \sigma^2 I_n)$

2.2 Part B

Using the same data as was used in the output, give the observed value of matrix product $(X'X)^{-1}X'y$.

We know from our lectures that $\hat{\beta} = (X'X)^{-1}X'y$. Knowing this, we can look at our output from *Problem 1c* and look at the `Parameter Estimate` section to get the 4 values we should get for our (4 x 1) dimension vector of regression parameters. By looking at our output we can determine that $(X'X)^{-1}X'y = \hat{\beta} = \begin{pmatrix} \hat{\beta_0} \\ \hat{\beta_1} \\ \hat{\beta_2} \\ \hat{\beta_3} \end{pmatrix} = \begin{pmatrix} 87.73953 \\ 0.32773 \\ -66.83803 \\ 0.32081 \end{pmatrix}$.

2.3 Part C

Suppose the columns of the X matrix are the same as that given in the output. Let $(X'X)^{-1}_{ij}$ denote the element in the i^{th} row and j^{th} column of the matrix $(X'X)^{-1}$.

2.3.1 Part i

What is $\sqrt{MS(E)(X'X)^{-1}_{11}}$?

We know that $\sqrt{MS(E)(X'X)^{-1}_{11}}$ is the first row and first column in the matrix. From our lecture we know $MS(E)(X'X)^{-1}_{11} = \hat{Var}(\hat{\beta}_0)$. Furthermore, $\sqrt{MS(E)(X'X)^{-1}_{11}} = \sqrt{\hat{Var}(\hat{\beta}_0)}$. We also know that the Standard Error of $\hat{\beta}_0$ can be found under the `Standard Error` column and the `Intercept` row for our output we were given. In our output we can see that $SE(\beta_0) = 16.29818$. Therefore, $\sqrt{MS(E)(X'X)^{-1}_{11}} = 16.29818$.

2.3.2 Part ii

What is $\sqrt{MS(E)(X'X)^{-1}_{22}}$?

We know that $\sqrt{MS(E)(X'X)^{-1}_{22}}$ is the second row and second column in the matrix. From our lecture we know $MS(E)(X'X)^{-1}_{22} = \hat{Var}(\hat{\beta}_1)$. Furthermore, $\sqrt{MS(E)(X'X)^{-1}_{22}} = \sqrt{\hat{Var}(\hat{\beta}_1)}$. We also know that the Standard Error of $\hat{\beta}_1$ can be found under the `Standard Error` column and the `sqftLiving` row for our output we were given. In our output we can see that this is $SE(\beta_1) = 0.00725$. Therefore, $\sqrt{MS(E)(X'X)^{-1}_{22}} = 0.00725$.

2.3.3 Part iii

If $\hat{\beta}$ denotes the observed least squares estimate of the vector of regression coefficients, then what is $(y - X\hat{\beta})'(y - X\hat{\beta})$?

We know that $\hat{y} = X\hat{\beta}$. We also know that $e = y - \hat{y}$ and thus, $e' = (y - \hat{y})'$. Furthermore, we know that $SS(E) = e'e$. Thus, $(y - X\hat{\beta})'(y - X\hat{\beta}) = (y - \hat{y})'(y - \hat{y}) = e'e = SS(E)$. We know from our output under the `Sum of Squares` column and the `Error` row that the $SS(E) = 331787073$. Therefore, $(y - X\hat{\beta})'(y - X\hat{\beta}) = 331787073$.