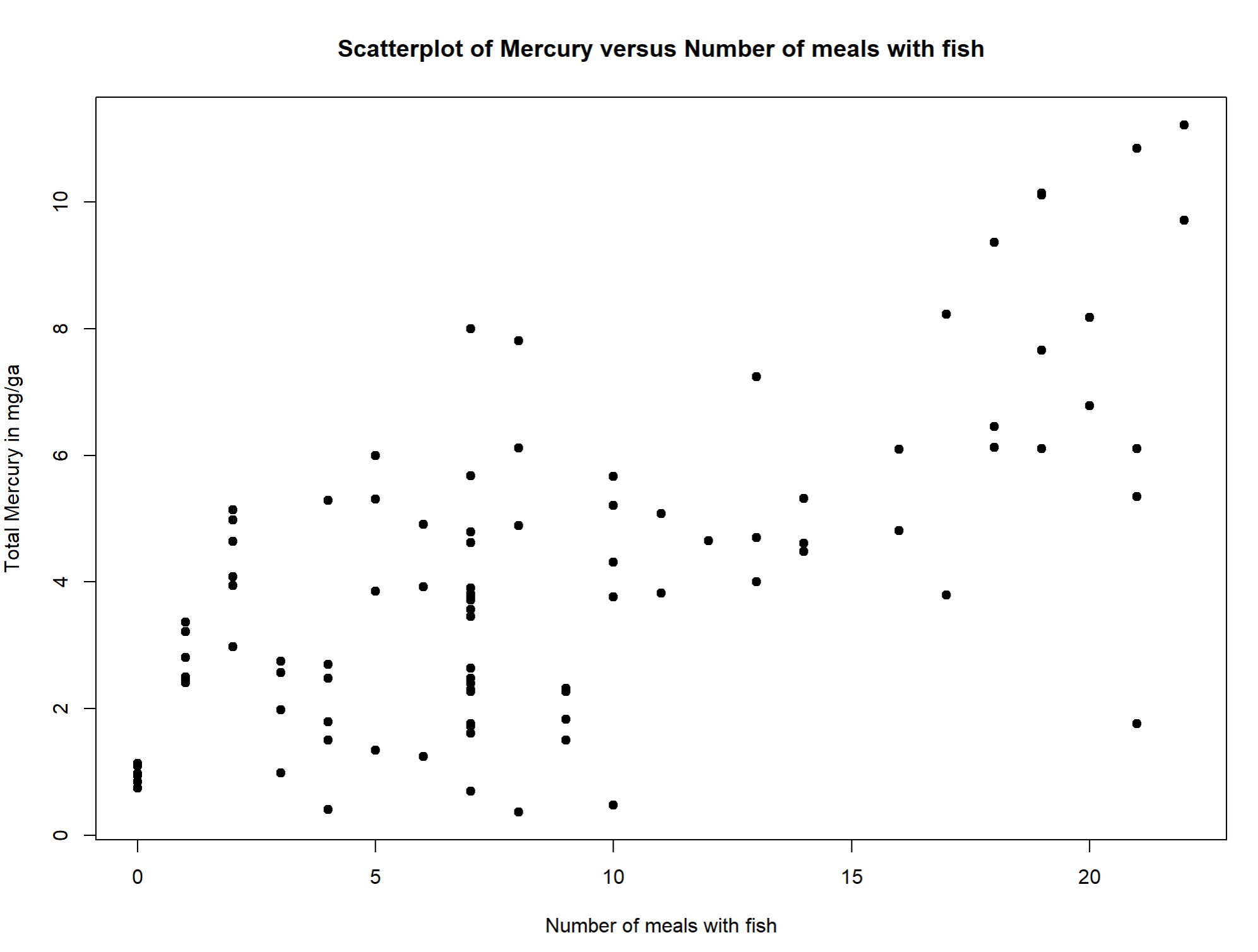
Yiduo Feng

CS 555

Homework 3

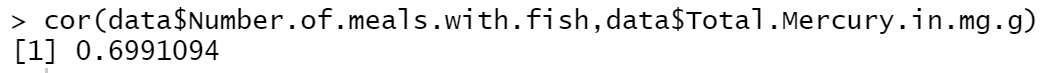
07/26/2022

**(1) To get a sense of the data, generate a scatterplot (using an appropriate window, label the axes, and title the graph). Consciously decide which variable should be on the x-axis and which should be on the y-axis. Using the scatterplot, describe the form, direction, and strength of the association between the variables.**



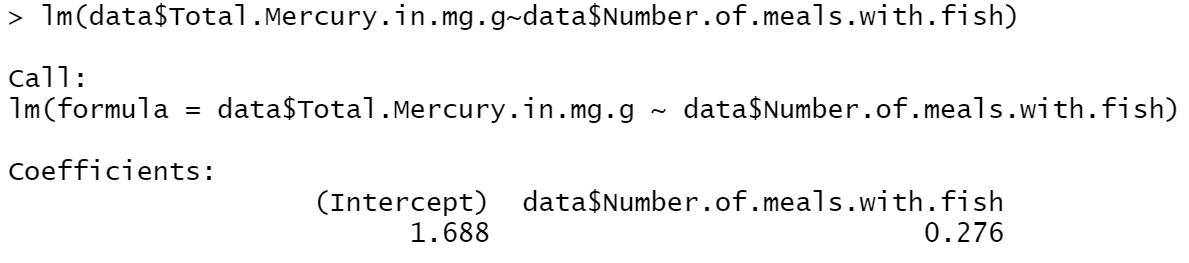
We can see a upward trend in the graph. The number of meals increases the mercury, so the direction is positive. The form is linear, and it is strong relationship but not very strong because we can see there are some outliers.

**(2) Calculate the correlation coefficient. What does the correlation tell us?**

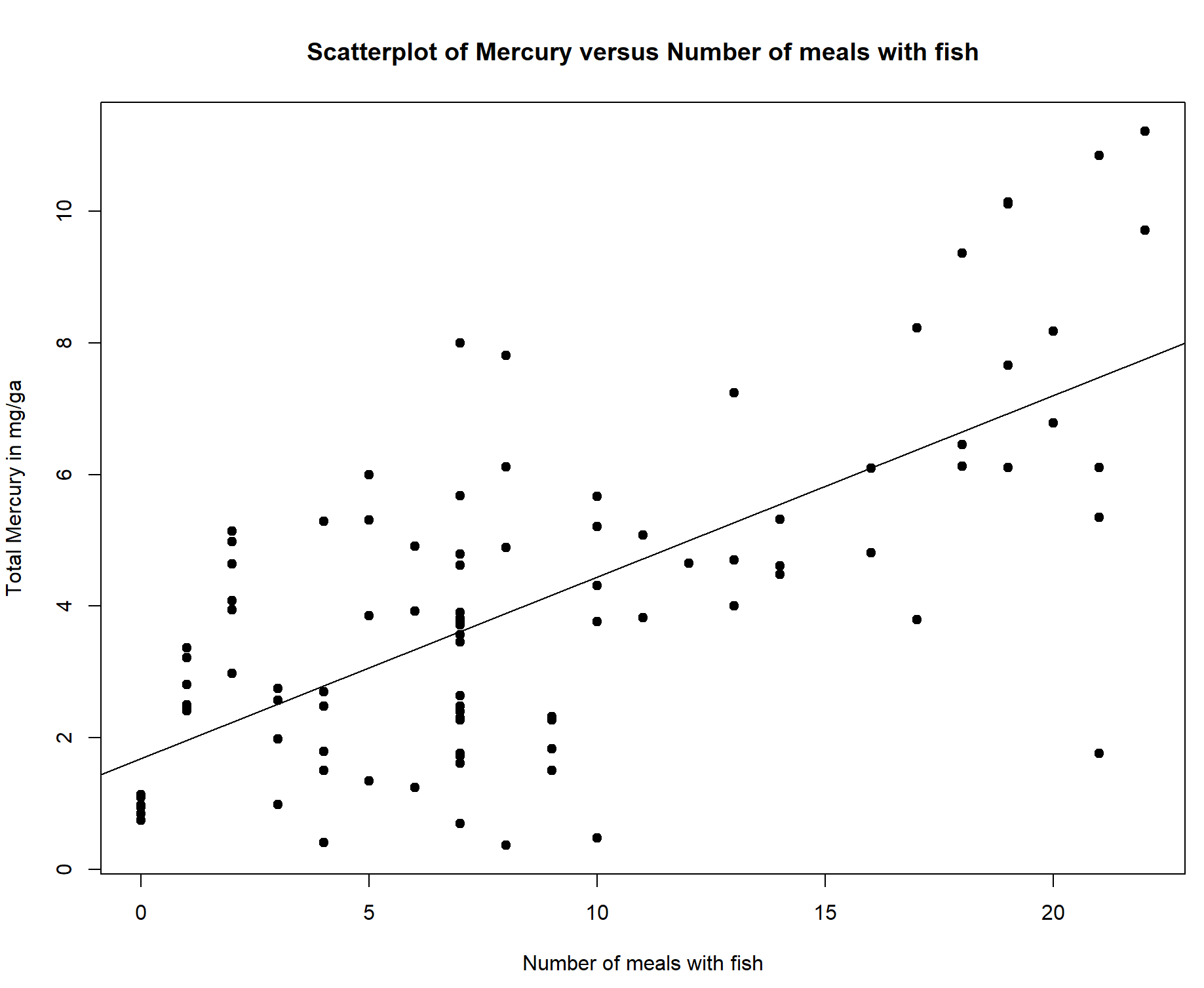


By using R, the correlation coefficient is 0.6991094. The greater the absolute value of the correlation coefficient, the stronger the correlation, the closer the correlation coefficient is to 1 or -1, the stronger the correlation, the closer the correlation coefficient is to 0, the weaker the correlation.0.6991094 is bigger than 0.5 and kind of close to 1, so this correlation is strong.

**(3) Find the equation of the least squares regression equation and write out the equation. Add the regression line to the scatterplot you generated above.**



The equation is: y = 1.688+0.276x

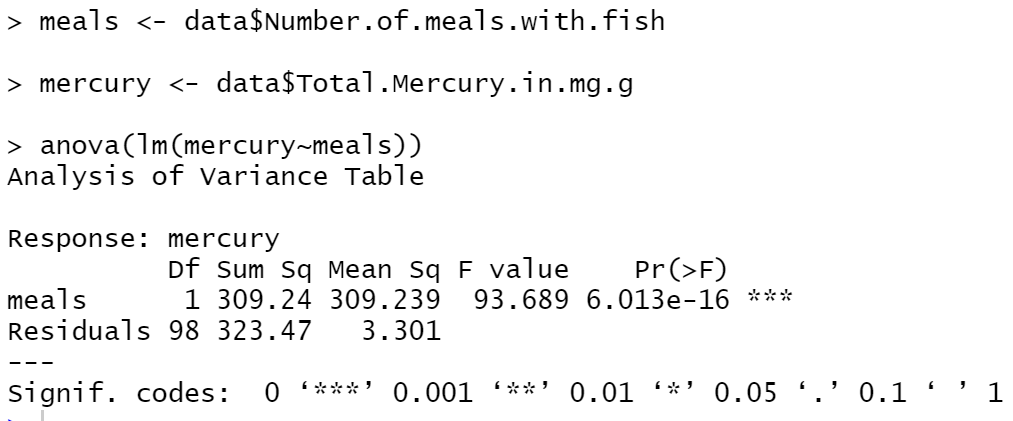


**(4) What is the estimate for ? How can we interpret this value? What is the estimate for ? What is the interpretation of this value? For the interpretations, you should be interpreting them in the context of this specific data set.**

Beta1 is 0.276 which is the slope of the line. Beta0 is 1.688 which is the intercept of the line. Beta 1 represents the degree to which the number of meals affects Mercury. When the number of meal increases, Mercury increases by about 0.276. Beta 0 represents the control value of the experiment. When the number of meals is 0, Mercury is about 1.688.

**(5) Calculate the ANOVA table AND the table which gives the standard error of . Formally test the hypothesis that = 0 using either the F-test or the t-test at the level. Either way, present your results using the 5-step procedure, as described in the course notes.**

**Within your conclusion, calculate the R-squared value and interpret this. Also, calculate (using R) and interpret the 90% confidence interval for .**



According to the code above, we get the ANOVA table.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | SS | df | MS | F-statixtic | p-value |
| Regression | 309.24 | 1 | 309.239 | 93.689 | 6.01E-16 |
| Residual | 323.47 | 98 | 3.301 |  |  |
| Total | 632.71 |  |  |  |  |

Standard error of estimate = sqrt(SSE/(n-2)) = sqrt(323.47/98)= 1.816787

1. Set up the hypotheses and select the alpha level

H0 : 𝛽1= 0 (there is no linear association between meals and mercury)

H1 : 𝛽1 ≠ 0 (there is a linear association between these factors)

𝛼 = 0.05

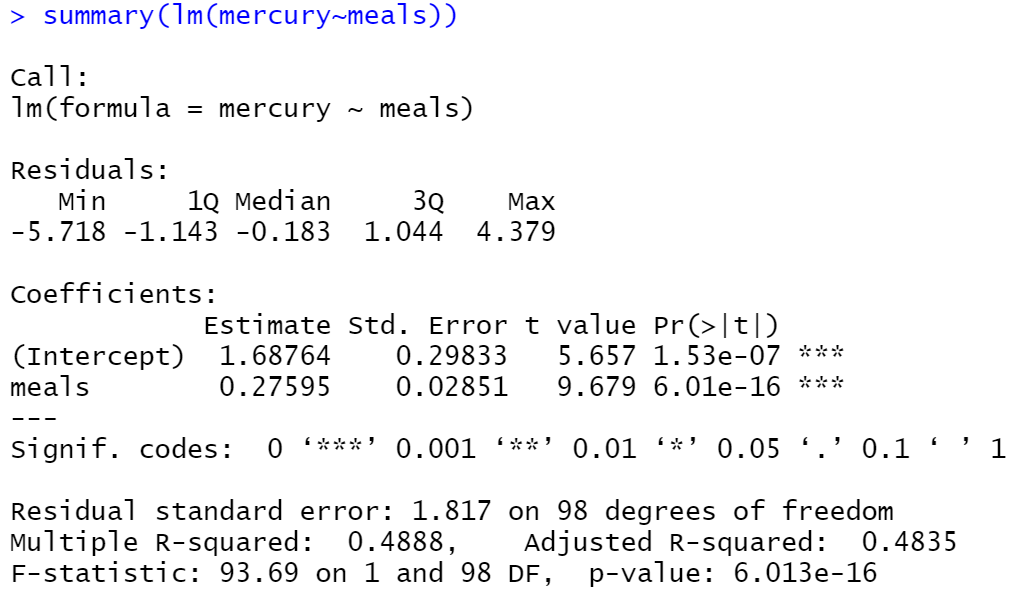
1. Select the appropriate test-statistic

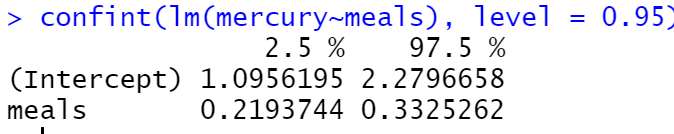
𝐹 = 𝑀𝑆 𝑅𝑒𝑔 𝑀𝑆 𝑅𝑒𝑠 with 1 and n-2 degrees of freedom

1. State the decision rule

Decision Rule: Reject H0 if 𝑝 ≤ 𝛼. Otherwise, do not reject H0

1. Compute the test statistic





According to the summary function, it gives F-statistic: 93.69 on 1 and 98 DF.

p-value: 6.013e-16.

1. Conclusion

Reject H0 since 𝑝 ≤ 𝛼. We have significant evidence at the 𝛼 = 0.05 level that 𝛽1≠ 0. That is, there is evidence of a significant linear association between meals with fish and mercury. The beta coefficient for the regression is 0.276 indicating an increase of approximately 0.276 percentage mercury for each additional number of meal with fish. The 95% confidence interval for the beta coefficient is approximately 0.219 to 0.3325 indicating that we are 95% confident that the underlying increase in mercury for each number of meal with fish is between 0.219 and 0.3325.