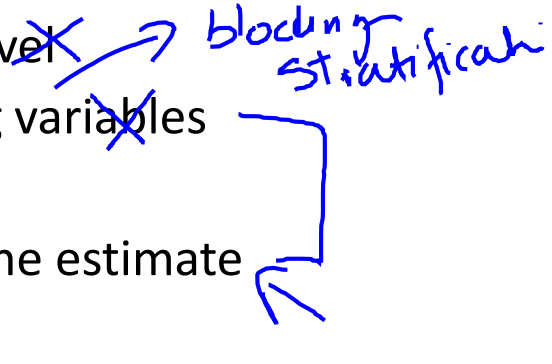
Monday, April 29, 2019

- Warm-up – choose the best answer & justify it
  - A city wants to conduct a poll of taxpayers to determine the level of support for constructing a new city-owned baseball stadium. Which of the following is the main reason for using a large sample size in constructing a confidence interval to estimate the proportion of city taxpayers who would support such a project?
    A. To increase the confidence level
    B. To eliminate any confounding variables
    C. To reduce nonresponse bias
    D. To increase the precision of the estimate
    E. To reduce undercoverage



- Review linear regression

Objectives

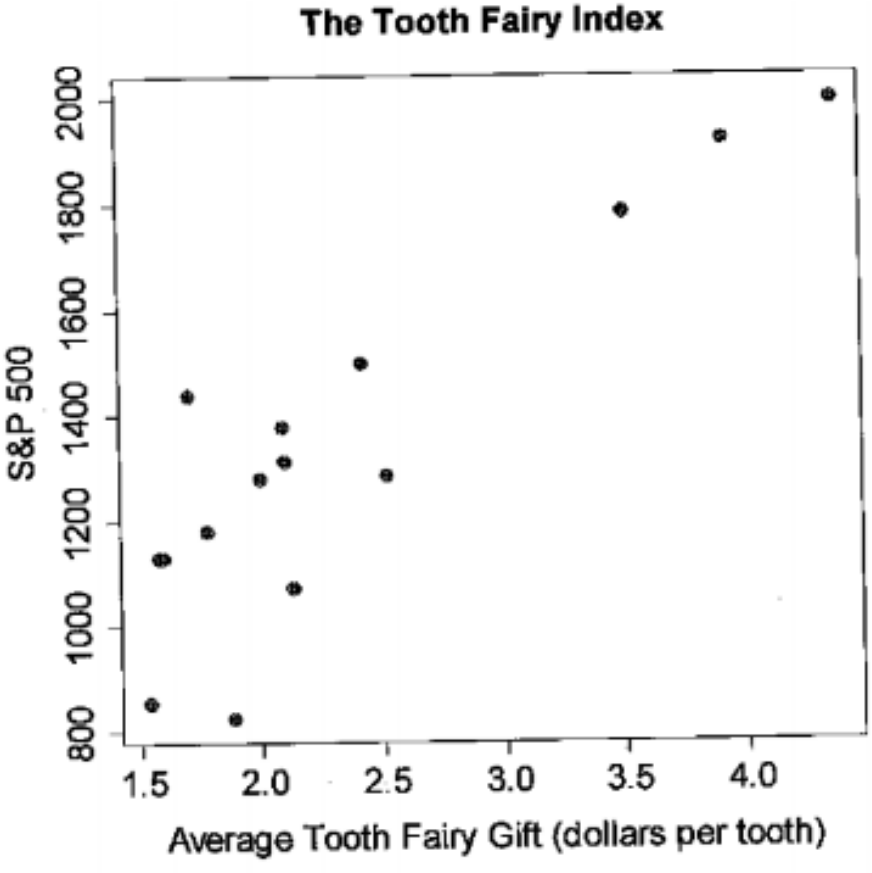
Content: I will review inference with **linear regression** through practice problems.

Social: I will participate in class discussion.

Language: I will look for and define key vocabulary so that I recognize it on assessments.

Day 92

Using annual survey data, Delta Dental estimates that the average tooth fairy gift (measured in dollars per tooth). The S&P 500 is a leading economic indicator based on the stock prices of 500 large companies. Is the Tooth Fairy Index also an economic indicator? The scatter plot below shows the relationship between these variables for the last 15 years.



a) Describe the relationship between the Tooth Fairy Index and the S7P 500.

Day 92

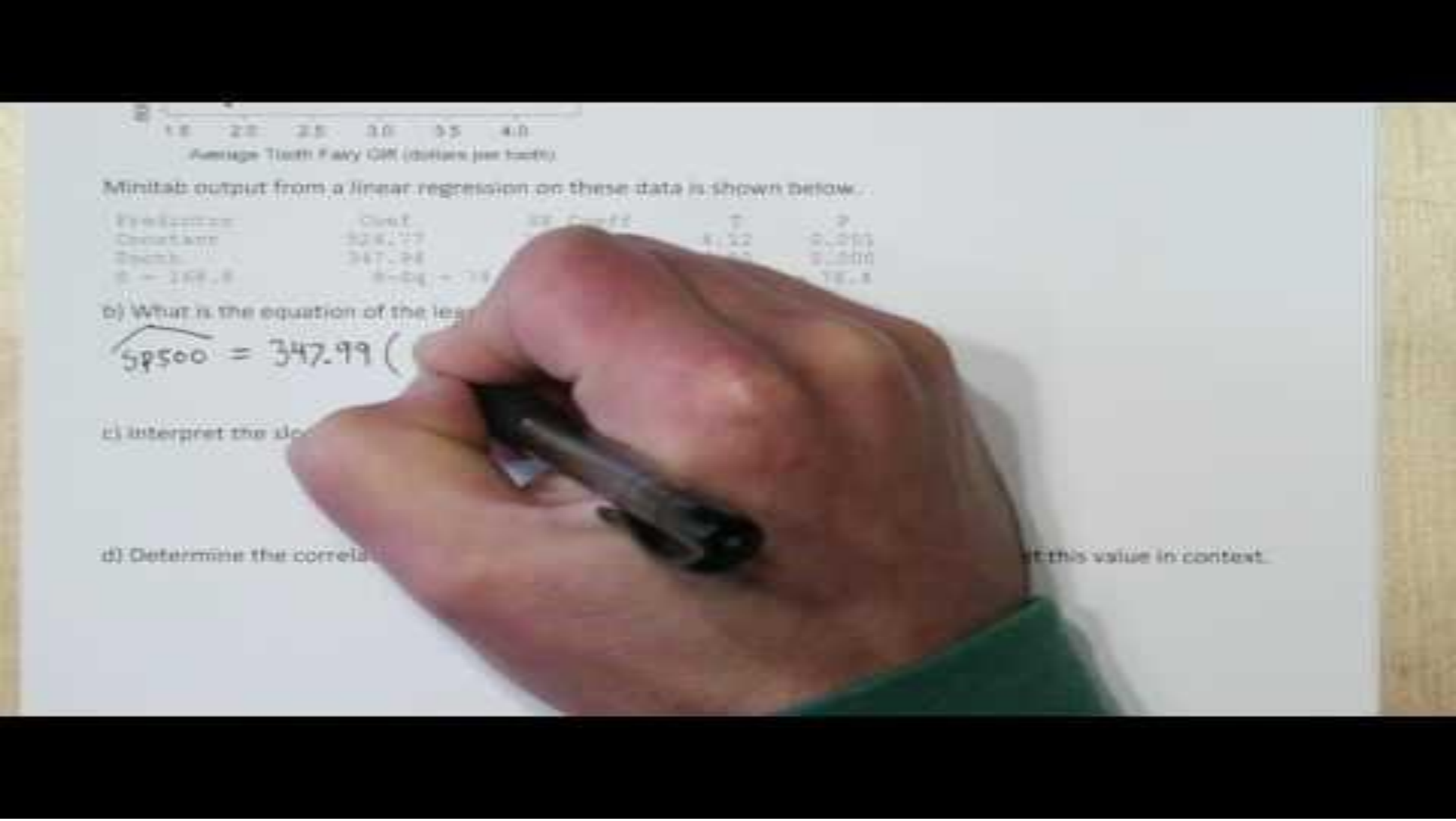
Using annual survey data, Delta Dental estimates that the average tooth fairy gift (measured in dollars per tooth). The S&P 500 is a leading economic indicator based on the stock prices of 500 large companies. Is the Tooth Fairy Index also an economic indicator? The scatter plot below shows the relationship between these variables for the last 15 years. Minitab output from a linear regression on these data is shown below.

Predictor	Coef	SE Coeff	T	P
Constant	524.77	127.41	4.12	0.001
Tooth	347.94	51.13	6.80	0.000
S = 168.8	R-Sq = 78.1		R-Sq(adj) = 76.4	

b) What is the equation of the least-squares regression line?

c) Interpret the slope and y-intercept from the equation.

d) Determine the correlation between the Tooth Fairy Index and the S&P 500. Interpret this in context.

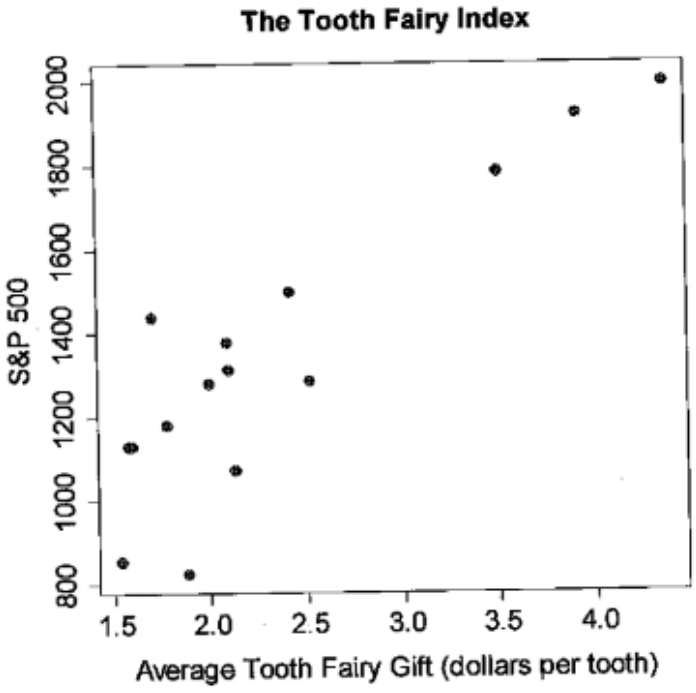

1.5 2.0 2.5 3.0 3.5 4.0
Average Tooth Wavy GR (dollars per tooth)
Minitab output from a linear regression on these data is shown below.
b) What is the equation of the leg
$58500 = 347.99 (
c) Interpret the slo
d) Determine the correla
this value in context.

Day 93

Using annual survey data, Delta Dental estimates that the average tooth fairy gift (measured in dollars per tooth). The S&P 500 is a leading economic indicator based on the stock prices of 500 large companies. Is the Tooth Fairy Index also an economic indicator? The scatter plot below shows the relationship between these variables for the last 15 years. Minitab output from a linear regression on these data is shown below.

Predictor	Coef	SE Coeff	T	P
Constant	524.77	127.41	4.12	0.001
Tooth	347.94	51.13	6.80	0.000

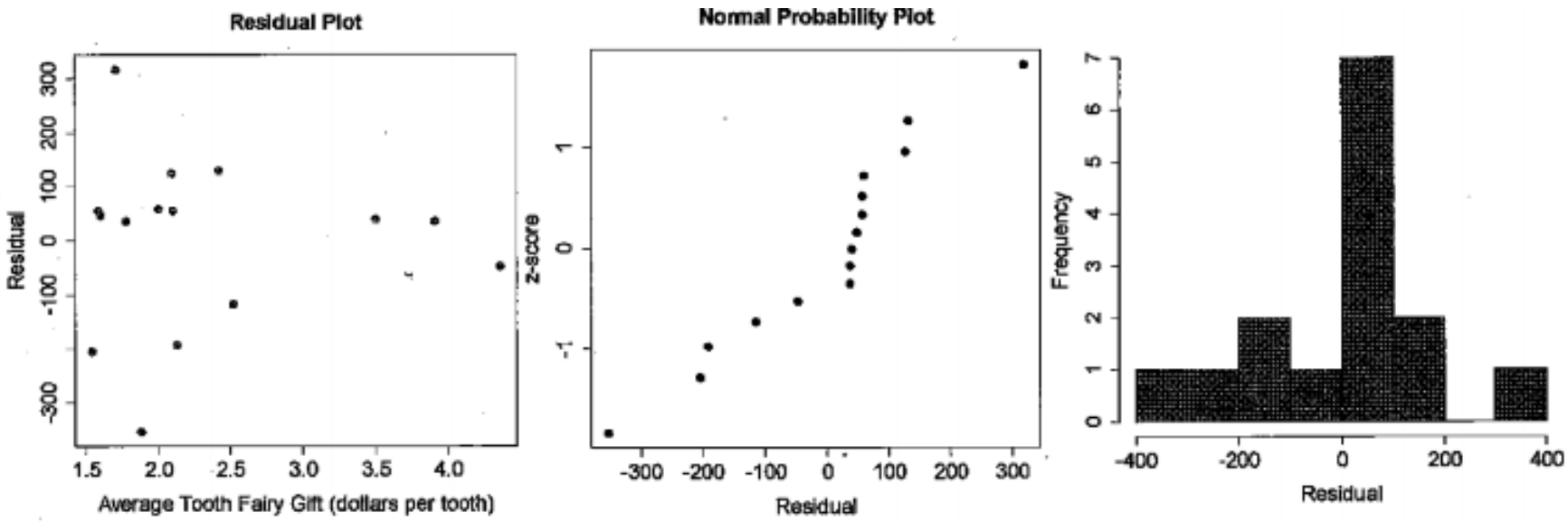
S = 168.8 R-Sq = 78.1 R-Sq(adj) = 76.4



a) Is a linear model appropriate for describing the relationship between the Tooth Fairy Index and the S&P 500?

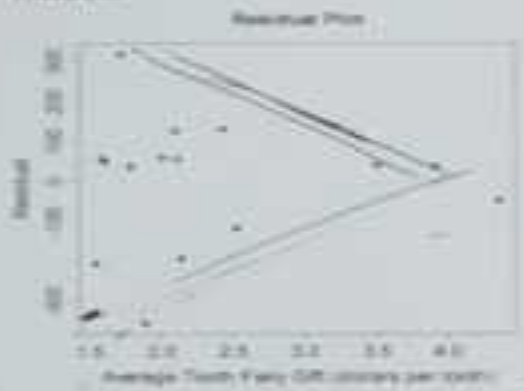
Day 93

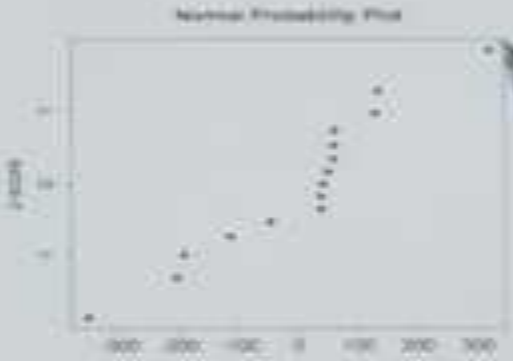
Using annual survey data, Delta Dental estimates that the average tooth fairy gift (measured in dollars per tooth). The S&P 500 is a leading economic indicator based on the stock prices of 500 large companies. Is the Tooth Fairy Index also an economic indicator? The scatter plot below shows the relationship between these variables for the last 15 years.



b) Use the plots and histogram to check the conditions for performing inference about the regression model.

b) Use the plots and histogram below to check the conditions for performing inference about the regression model.





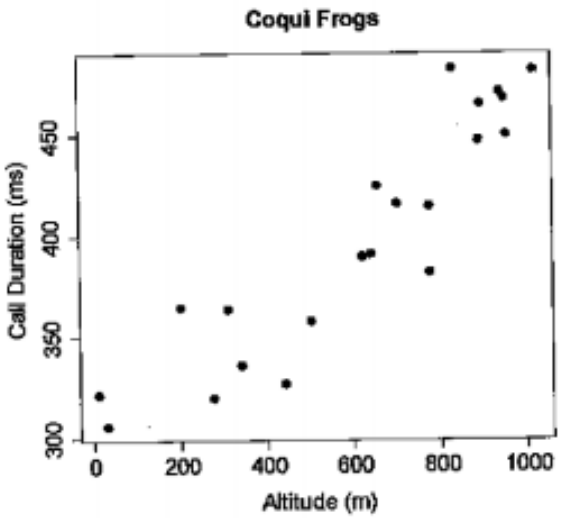
Linear: Our scatterplot appears linear, but the residual plot shows variation in the SP500's value for lower values of average tooth fa

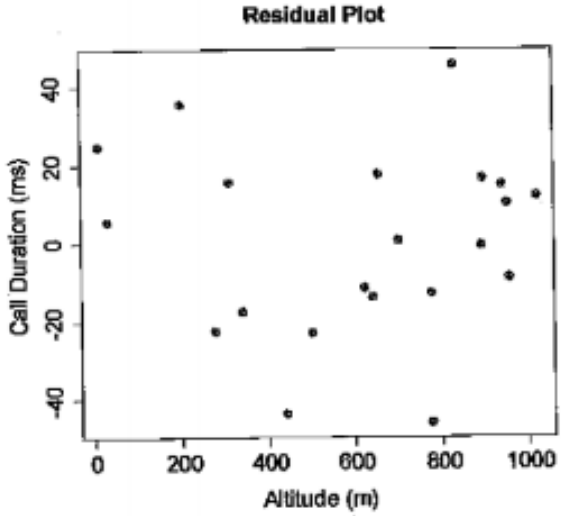
Independent: We suspect the values of our variables are not data comes from consecutive years and the values of average tooth fairy gift are at least partially depe previous years amount. Not Met.

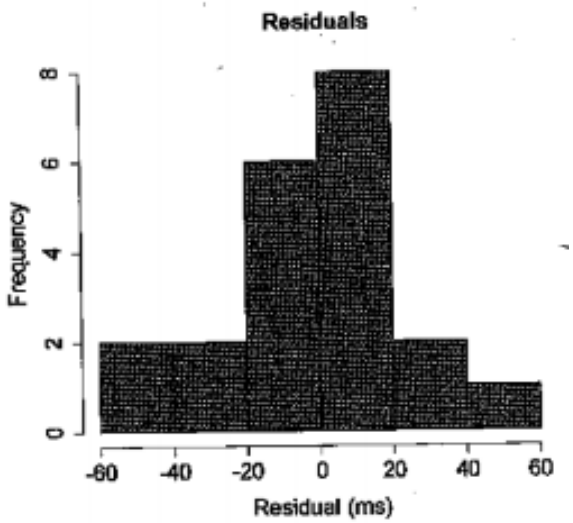
Normal:

Day 94

A coaui frog's call starts with a "CO" followed y a higher pitched "QUI." The duration of this call varies. A linear regression on the relationship between the call duration (measured in milliseconds) and altitude (measured in meters) resulted in the following output from statistical software.



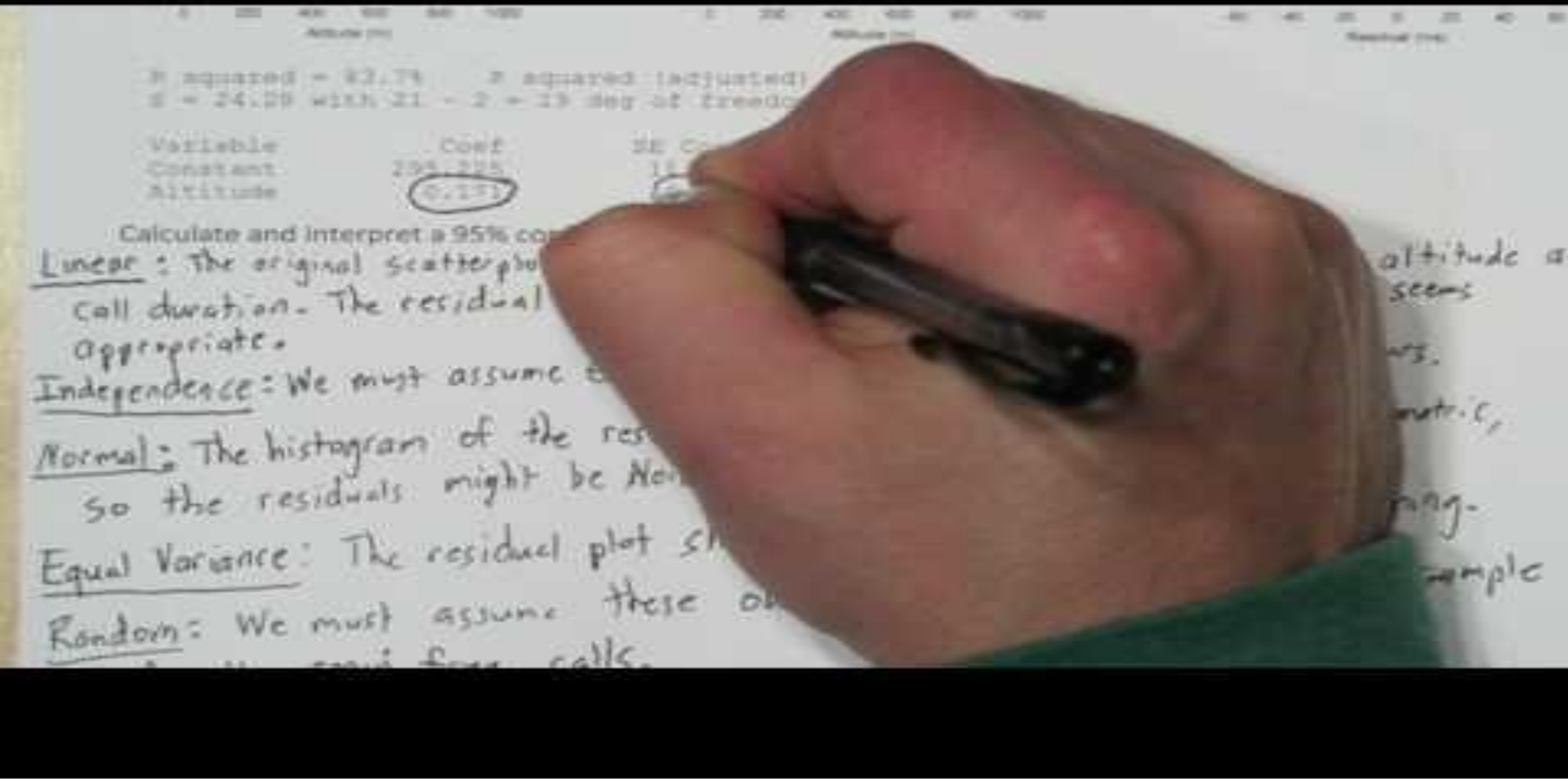




R squared = 83.7% R squared (adjusted) = 82.9%
S = 24.28 with 21 - 2 = 19 deg of freedom

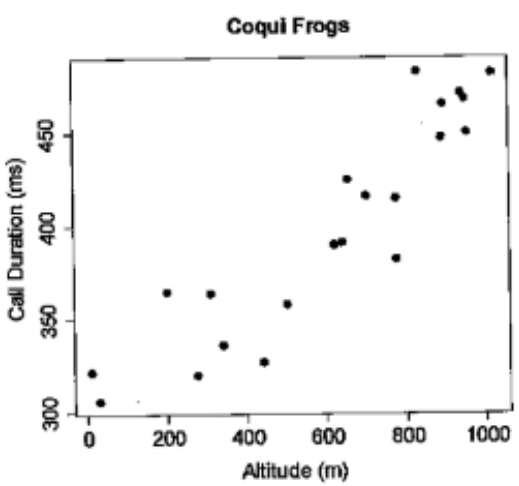
Variable	Coef	SE Coeff	T	P
Constant	295.225	11.7743	25.07	0.000
Altitude	0.171	0.0173	9.89	0.000

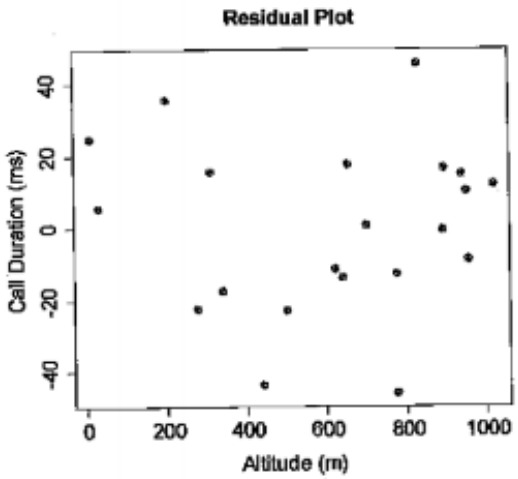
Calculate and interpret a 95% confidence interval for the slope for the slope of the true regression line.

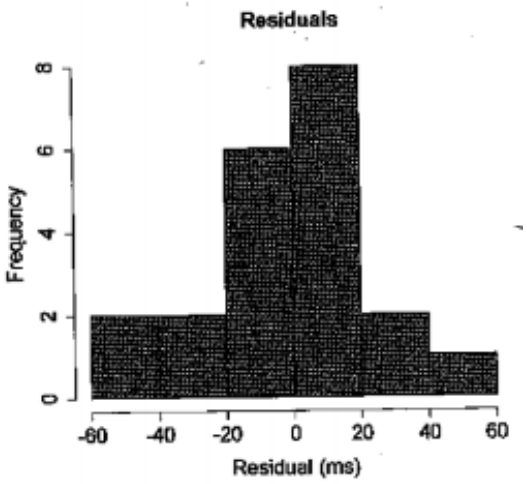

R squared = 82.7%
Variable
Constant
Altitude
Calculate and interpret a 95% co
Linear: The original scatterplo
call duration. The residual
appropriate.
Independence: We must assume
Normal: The histogram of the res
so the residuals might be No
Equal Variance: The residual plot sh
Random: We must assume these o
altitude a
seems
mple

Day 95

A coaui frog's call starts with a "CO" followed y a higher pitched "QUI." The duration of this call varies. A linear regression on the relationship between the call duration (measured in milliseconds) and altitude (measured in meters) resulted in the following output from statistical software.







R squared = 83.7% R squared (adjusted) = 82.9%
S = 24.28 with 21 - 2 = 19 deg of freedom

Variable	Coef	SE Coeff	T	P
Constant	295.225	11.7743	25.07	0.000
Altitude	0.171	0.0173	9.89	0.000

a) Is there statistically convincing evidence of a linear relationship between altitude and call duration?

b) How does the confidence interval constructed on the previous problem relate to your findings in part a?

c) Interpret s, r^2, and the standard error of the slope in context.

Use the context and information from *Day 94* to answer the following.

a) Is there statistically convincing evidence of a linear relationship between altitude and call duration?

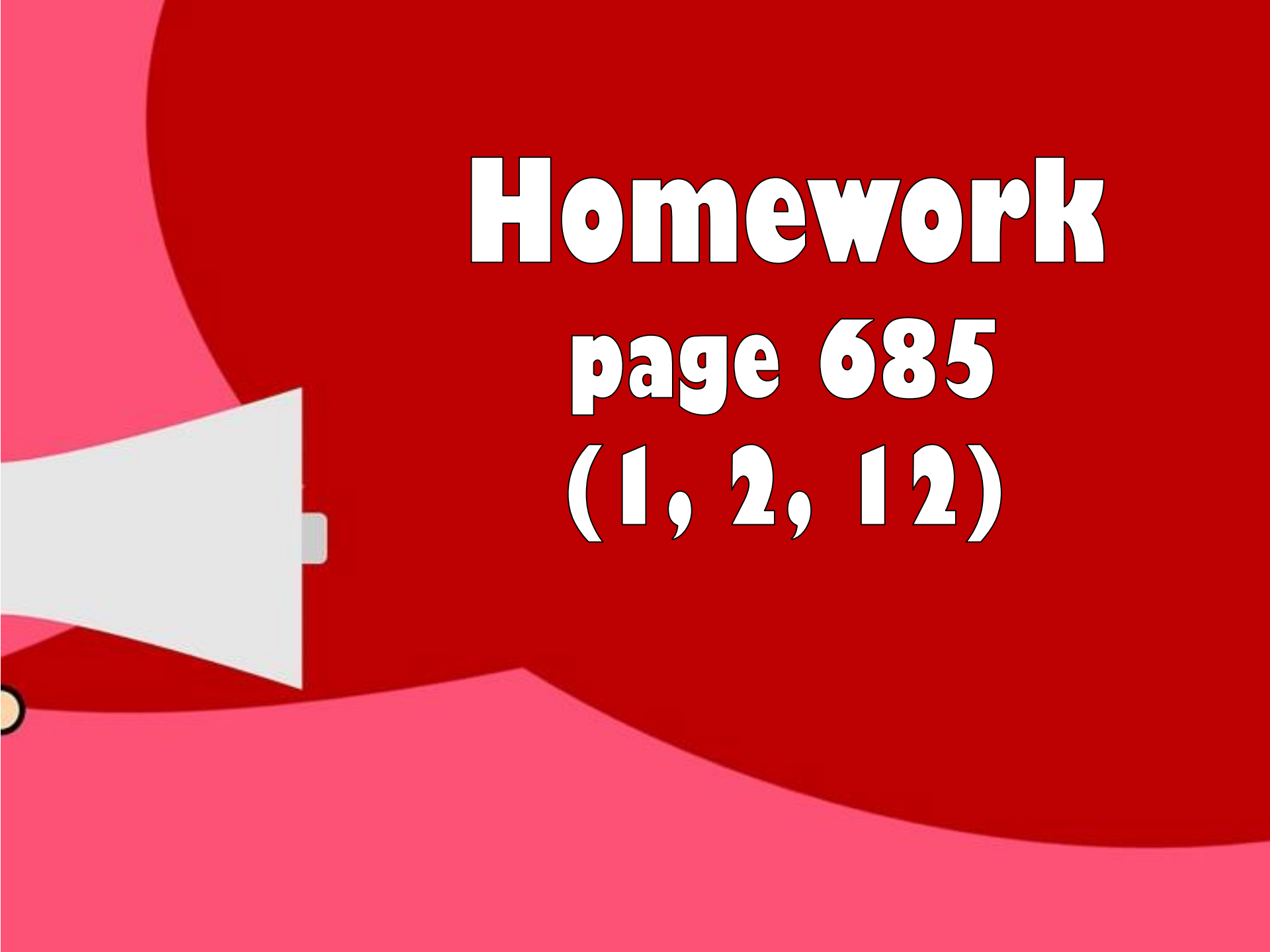
State: We want to test the following hypotheses at the $\alpha = 0.05$ level:

$H_0: \beta = 0$ where β is the true slope of the population regression equation

$H_a: \beta \neq 0$ relating altitude to call duration.

Plan: We checked conditions on day 94, so we will do a t-test for β.

Do: $t = \frac{b - 0}{s_b} = \frac{.171 - 0}{.0173} \approx 9.88$


Homework
page 685
(1, 2, 12)