Lesson 20: Residuals

Daily Data Collection

Select two topics you think are correlated, make a hypothesis,

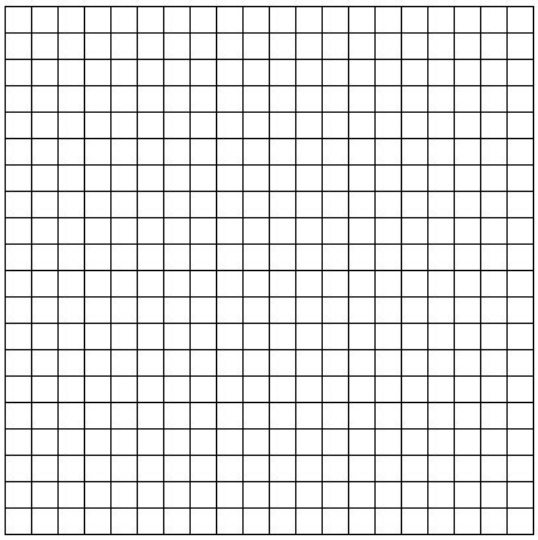
and run a test to see if your assumptions were true.

Class Data:

Explanatory Variable:

Response variable:

Create a scatterplot.



Describe the Direction, Form, and Strength

Write an equation for the regression line

Describe the slope in the context of the situation

Find your residual value

Conclusion/Analysis

Statistical Formulas

Formula for finding the slope and y-intercept in a linear regression line:

Slope: $\boxed{b = r\frac{s_y}{s_x}}$ r = correlation S_y = standard deviation of y S_x = standard deviation of x

The slope of the regression line is important in the sense that it gives us the rate of change of $\hat{y}$ with respect to x. In other words, it gives us the amount of change in $\hat{y}$ when x increases by 1.

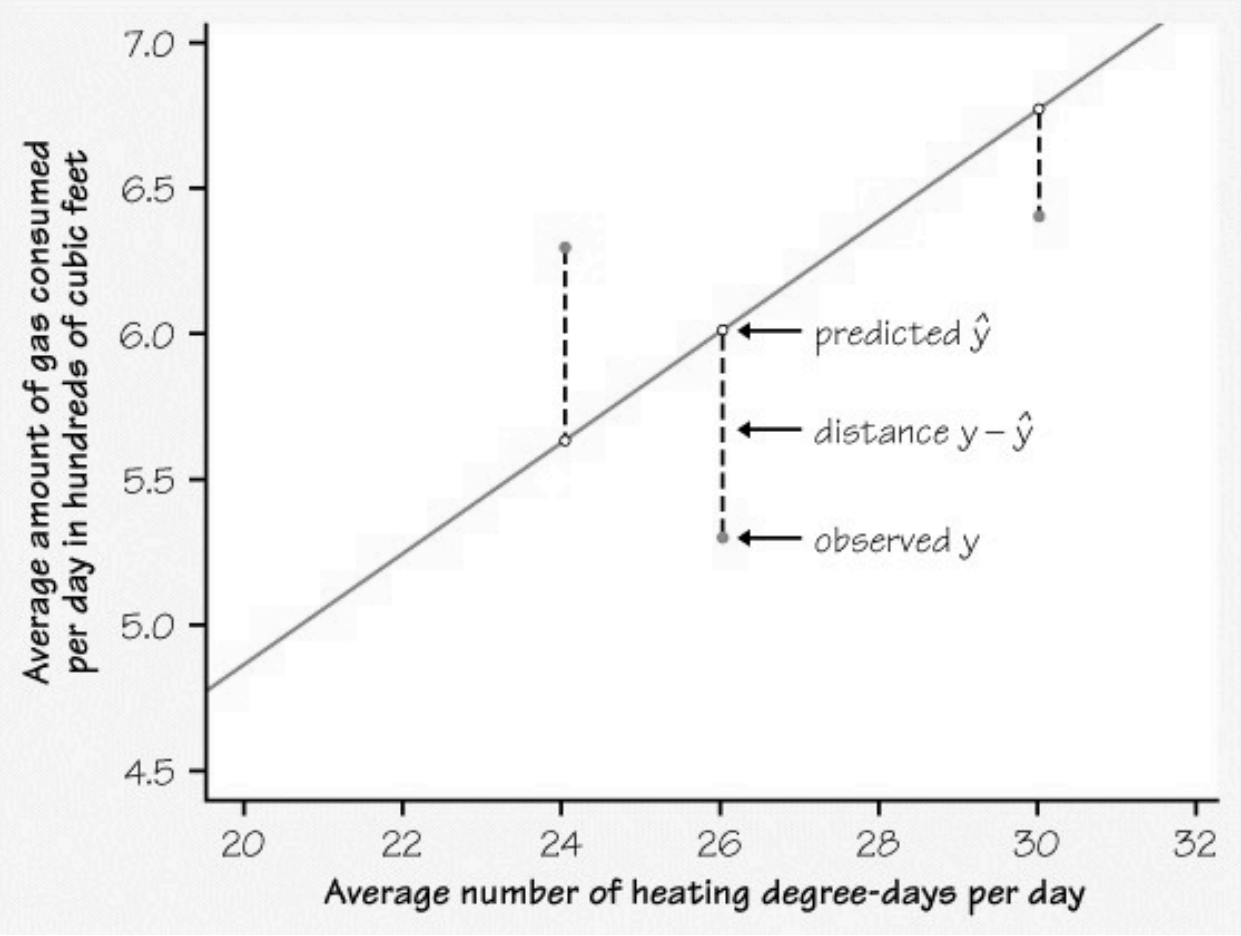
Intercept: $\boxed{a = \bar{y} - b\bar{x}}$

The intercept is statistically meaningful only when x can actually take values close to zero.

Residuals:

Since the LSRL minimized the vertical distance between the data values and a trend line we have a special name for these vertical distances. They are called **residuals**. A residual is simply the difference between the observed y and the predicted y.

Residual = y - $\hat{y}$



Does Fidgeting Keep You Slim?

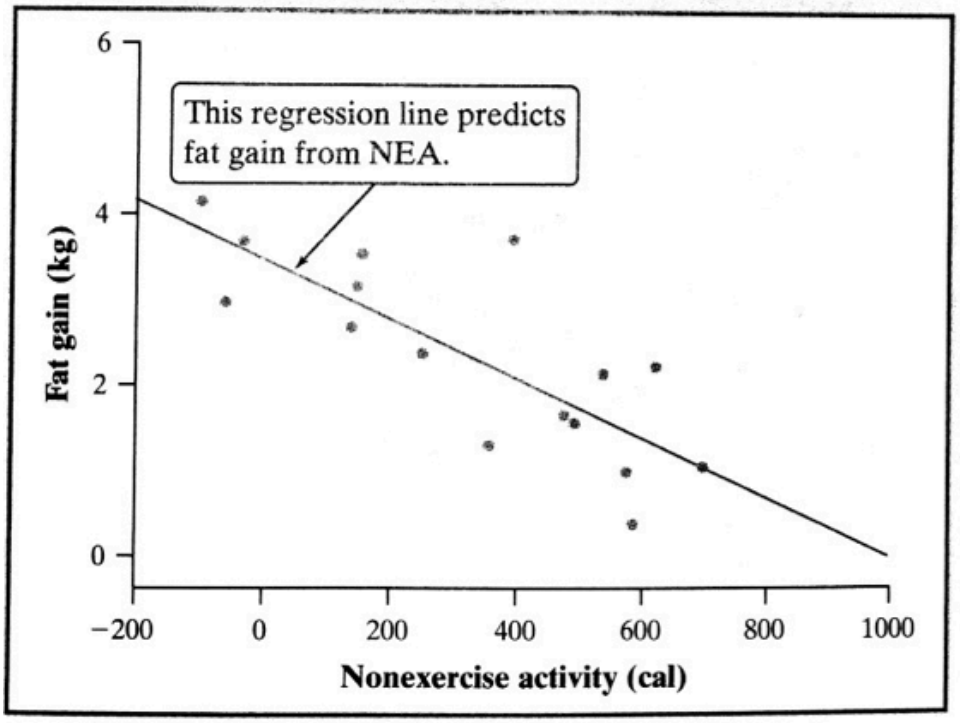
Regression lines as models

Some people don't gain weight even when they overeat. Perhaps fidgeting and other "nonexercise activity" (NEA) explains why—some people may spontaneously increase nonexercise activity when fed more. Researchers deliberately overfed 16 healthy young adults for 8 weeks. They measured fat gain (in kilograms) as the response variable and change in energy use (in calories) from activity other than deliberate exercise—fidgeting, daily living, and the like—as the explanatory variable. Here are the data:[13]

NEA change (cal):	−94	−57	−29	135	143	151	245	355
Fat gain (kg):	4.2	3.0	3.7	2.7	3.2	3.6	2.4	1.3
NEA change (cal):	392	473	486	535	571	580	620	690
Fat gain (kg):	3.8	1.7	1.6	2.2	1.0	0.4	2.3	1.1

Do people with larger increases in NEA tend to gain less fat?

Figure 3.7 is a scatterplot of these data. The plot shows a moderately strong, negative linear association between NEA change and fat gain with no outliers. The correlation is $r = -0.7786$. The line on the plot is a regression line for predicting fat gain from change in NEA.



$$\widehat{\text{fat gain}} = 3.505 - 0.00344(\text{NEA change})$$

Find the residual for the subject who increased NEA by 620 calories and Interpret the value.

For which subject did the regression line over predict fat gain by the most? Explain.

HW 20 - Section 3-2: 43,45,47,53