



1. Draw in a best fit line on the scatterplot above.
2. Label the residual for the point (5, 70) and the point (13, 160). Estimate the value of each.
3. Draw a triangle at the estimated mean number of calories per serving when there are 10 grams of sugar per serving in the cereal.
4. Draw a rectangle at the fitted value for calories when sugar=7.
5. Draw a star at  $\mu(\widehat{cal} | \widehat{sugar} = 3)$ .
6. Write out your estimated regression equation. Replace  $x$  and  $y$  with  $cal$  and  $sugar$ .

$$\widehat{\mu\{cal | sugar\}} = 90 + \frac{30}{15} sugar$$

Now let's have R estimate the regression line for us.

```
cereal.fit <- lm(calories~sugar, data = cereal)
summary(cereal.fit)

##
## Call:
## lm(formula = calories ~ sugar, data = cereal)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -39.65  -9.47   0.47  10.47  38.05
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    89.65      3.45    26.00 < 2e-16 ***
## sugar           2.48      0.42     5.92 9.2e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 16.2 on 75 degrees of freedom
## Multiple R-squared:  0.318, Adjusted R-squared:  0.309
## F-statistic: 35 on 1 and 75 DF, p-value: 9.17e-08
```

7. Write out R's estimated regression equation.

$$\widehat{\mu\{cal|sugar\}} = 89.65 + 2.48sugar$$

8. Interpret the slope estimate.

A one g increase in the g of sugar per serving is estimated to be associated with a 2.48 kcal increase in the mean amount of calories in cereal.

9. Interpret the y-intercept estimate.

For cereals w/ 0 g of sugar per serving, the mean amount of calories is estimated to be 89.65 kcal.

10. What is the estimated mean number of calories per serving when there are 10 grams of sugar per serving in the cereal, according to R's model?

$$89.65 + 2.48(10) = 114.48$$

11. What is  $\widehat{\mu(y|x=3)}$  according to R's model?

$$89.65 + 2.48(3) = 97.09$$