2.5 Regression Lines (Part 2): Interpreting a Regression Line

* $y = mx + b$ *

In the regression line $\hat{y} = a + bx$, a is the **y-intercept** and b is the **slope**. The **y-intercept** *a* is the predicted value of y when x = 0. The **slope** *b* of a regression line described the predicted change in the y variable for each 1-unit increase in the x variable.

In the Ford F-150 example, the equation of the regression line is $\hat{y} = 38{,}257 - 0.1629x$.
The slope is the coefficient of x, b = -0.1629. This means that the *predicted* price of a Ford F-150 goes down by 0.1629 dollars for each additional mile that the truck is driven.
The y-intercept is a = 38,257. This means that the *predicted* price of a truck that has been driven 0 miles is $38,257.

It is very important to include the word "predicted" (or its equivalent) in the interpretation of the slope and y-intercept. Otherwise, it may seem that our predictions will be exactly correct.

Example: In the example yesterday about tapping on cans, the equation of the regression line is $\hat{y} = \boxed{248.6} + \boxed{2.63}x$, where x is the tapping time (in seconds) and y is the amount of soda remaining (in milliliters).

$\hat{y} = a + bx$

a.) Interpret the $\overset{b}{\boxed{\text{slope}}}$ of the regression line.

$b = 2.63$

The predicted amount of soda remaining in the can goes up by 2.63mL for every second tapped.

b.) Does the value of the $\underset{a}{\boxed{\text{y-intercept}}}$ have meaning in this context? If so, interpret the y-intercept. If not, explain why.

$a = 248.6$

When the tapping time is 0 seconds (x=0), the amount of soda remaining is 248.6mL.

In some contexts, the y-intercept doesn't have meaning because a value of x = 0 doesn't make sense. For example, in a scatterplot relating x as height and y as weight for a sample of students, it wouldn't make sense to predict the weight for a student with a height of 0.