



You are working at Rivian, who are rolling out their new battery powered R1S SUV. You are charged with developing a strategy for capturing some of the resale market - once the first vehicles will be sold as used cars. You will have to figure out how to price these used vehicles. As there are no used vehicles available, you are interested in the relationship between what people are willing to pay for used vehicles with different odometer readings. You know that higher odometer readings will result in lower prices, but you want to know how much lower.

1. You have hired a research firm to provide you with data on sales of the closest competitor vehicle, the Tesla Model X with a comparable battery pack size. You have data for 5,000 transactions for Tesla Model X standard range model year 2019, which contain the price paid in US\$ and the odometer reading (in miles). Open the dataset “tesla_x.xlsx”.
2. Now let's get fancy. Red the data into R. Regress price on odometer readings. Generate the regression output. To make sure that we are getting the same numbers, the intercept estimate should be 75063.2882.

a. Please let me know what values you are getting for the following statistics:

Number of Observations:

Estimate of the Slope on Odometer Readings:

- b. Interpret the intercept in words, making sure you use the correct units.
 - c. Interpret the slope in words, making sure you use the correct units.
 - d. Sketch the scatterplot with regression line. Please label the units on both axes.
 - e. Is this relationship linear? Does the model tend to underpredict or overpredict prices for cars of particular odometer readings?
 - f. Generate and plot the residuals. What do you notice?
3. Estimate three alternative models. You will need to first create additional variables with $\text{LN}(\text{Price})$ and $\text{LN}(\text{Odometer})$. Be sure to use the natural log rather than base10 log.

Model #1: $\text{Log Price} = b_0 + b_1 \text{ Odometer}$

Model #2: $\text{Price} = b_0 + b_1 \text{Log Odometer}$

Model #3: $\text{Log Price} = b_0 + b_1 \text{Log Odometer}$

- a. Examine the residual plots for all three models. Which model best fits the data (circle one)? Explain briefly.

Model #1

Model #2

Model #3

- b. Interpret the slope from your preferred model in words with all appropriate units.
- c. What does your preferred model predict for the price (in dollars) of a Tesla Model X that has an odometer reading of 70,000 total miles?