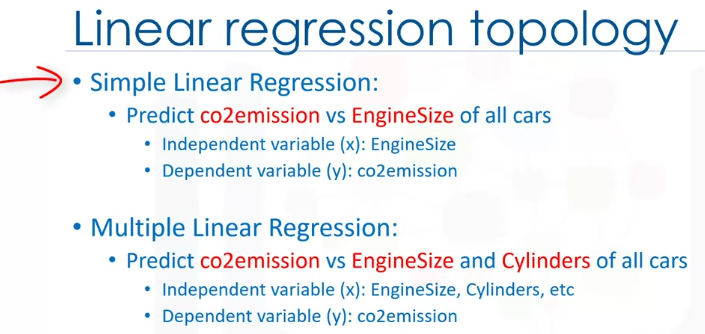


we can use linear regression to predict a continuous value such as co2-emission.

In simple linear regression there are two variables, a dependent variable and an independent variable.



There are two types of linear regression models.  
They are simple regression and multiple regression.

simple linear regression: when one independent variable is used to estimate a dependent variable.

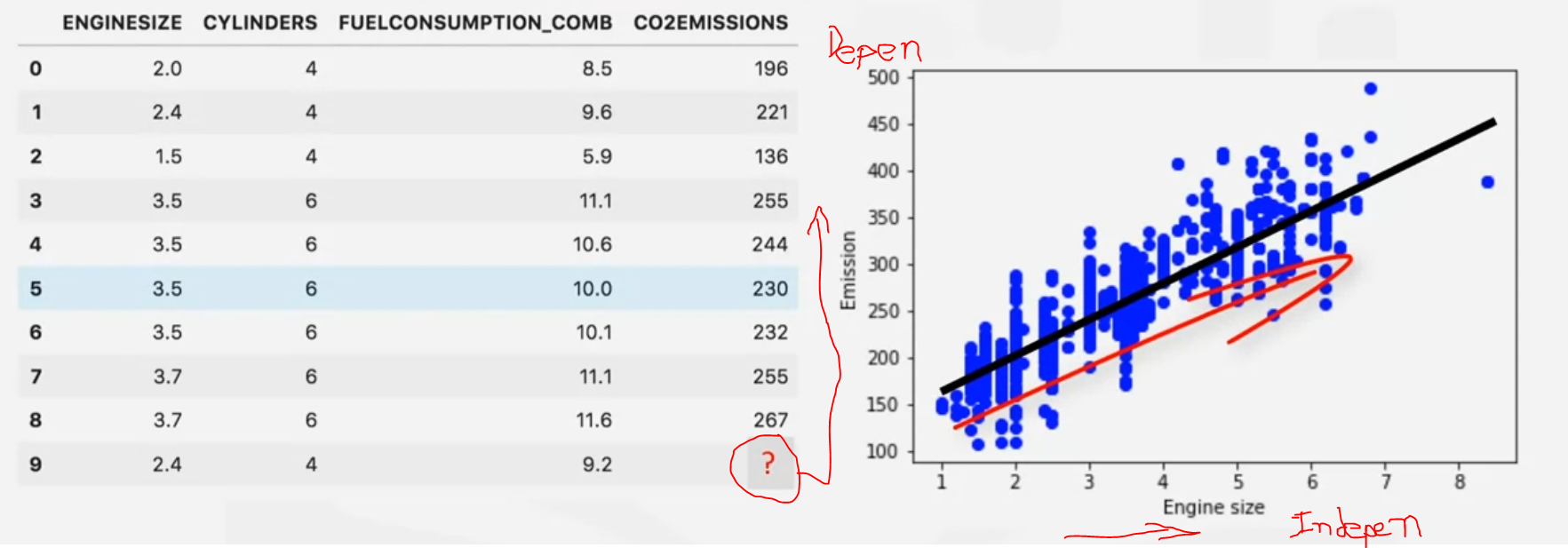
For example: predicting co2 emission using the engine size variable.

When more than one independent variable is present, the

process is called multiple linear regression.

For example: predicting co2-emission using engine size and cylinders of cars.

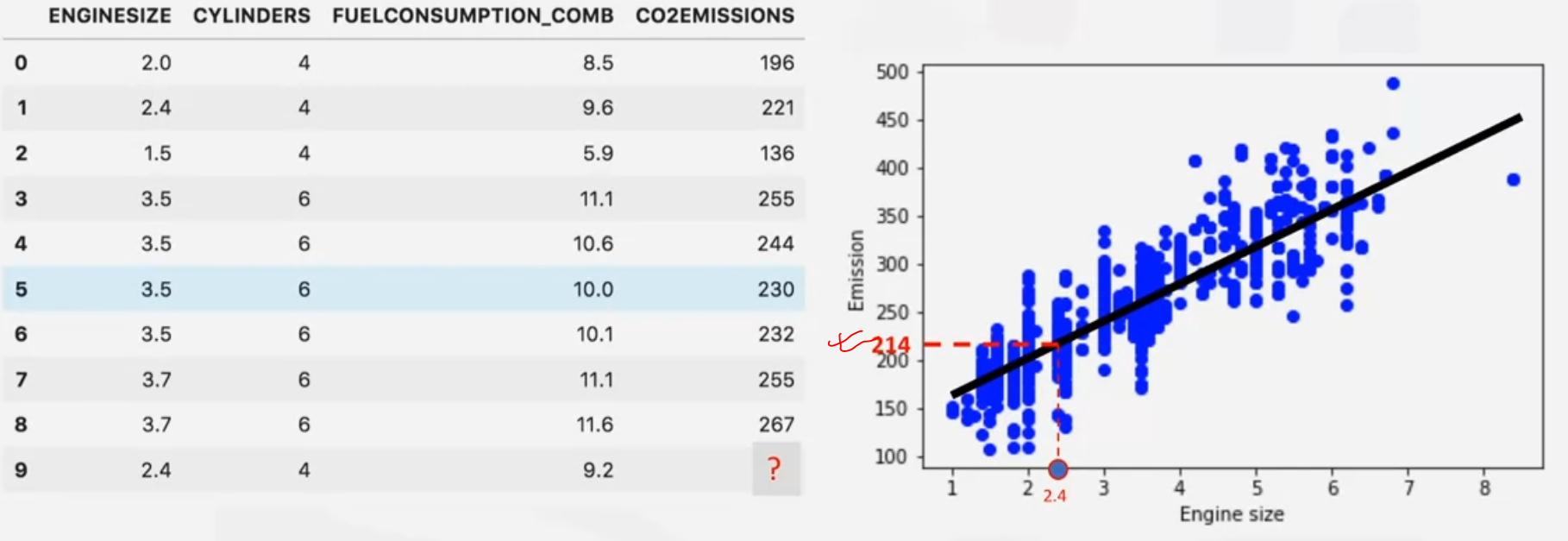
NOW, using simple linear regression, find the value,



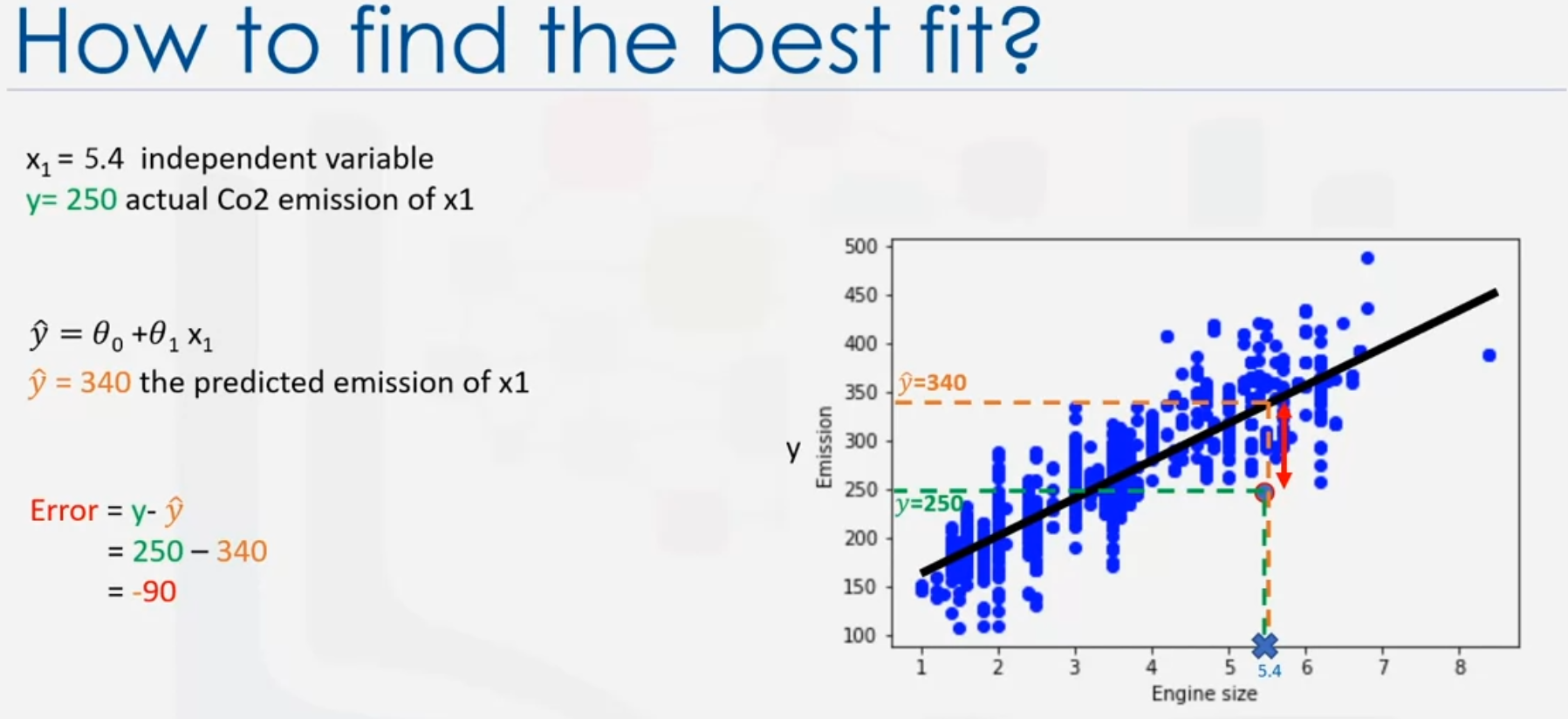
let us assume for a moment that the line is a good fit of the data we can use it to predict the emission of an unknown car.

For example:

for a sample car with engine size 2.4, you can find the emission is 214

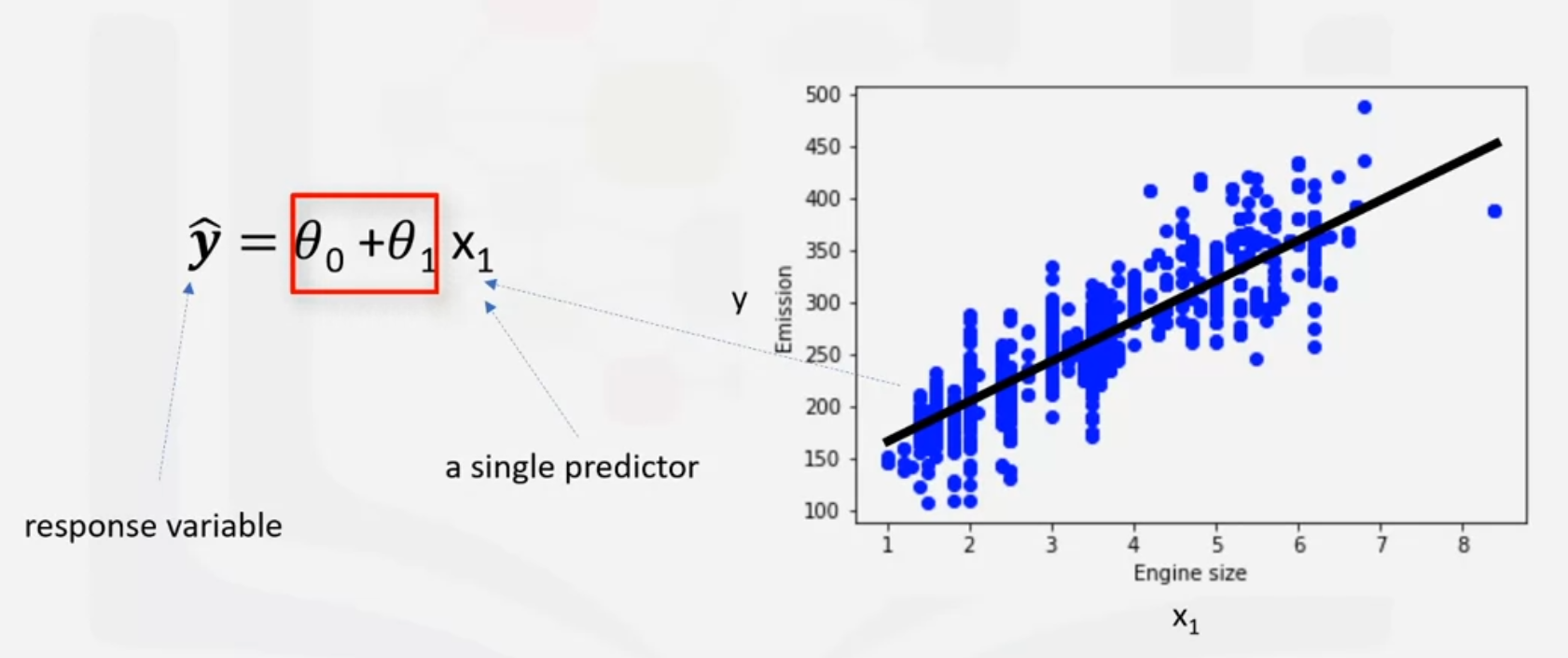


But here is the problem:



So the fit line is not accurate.

We need to find the best Fit line.  
the formula is:



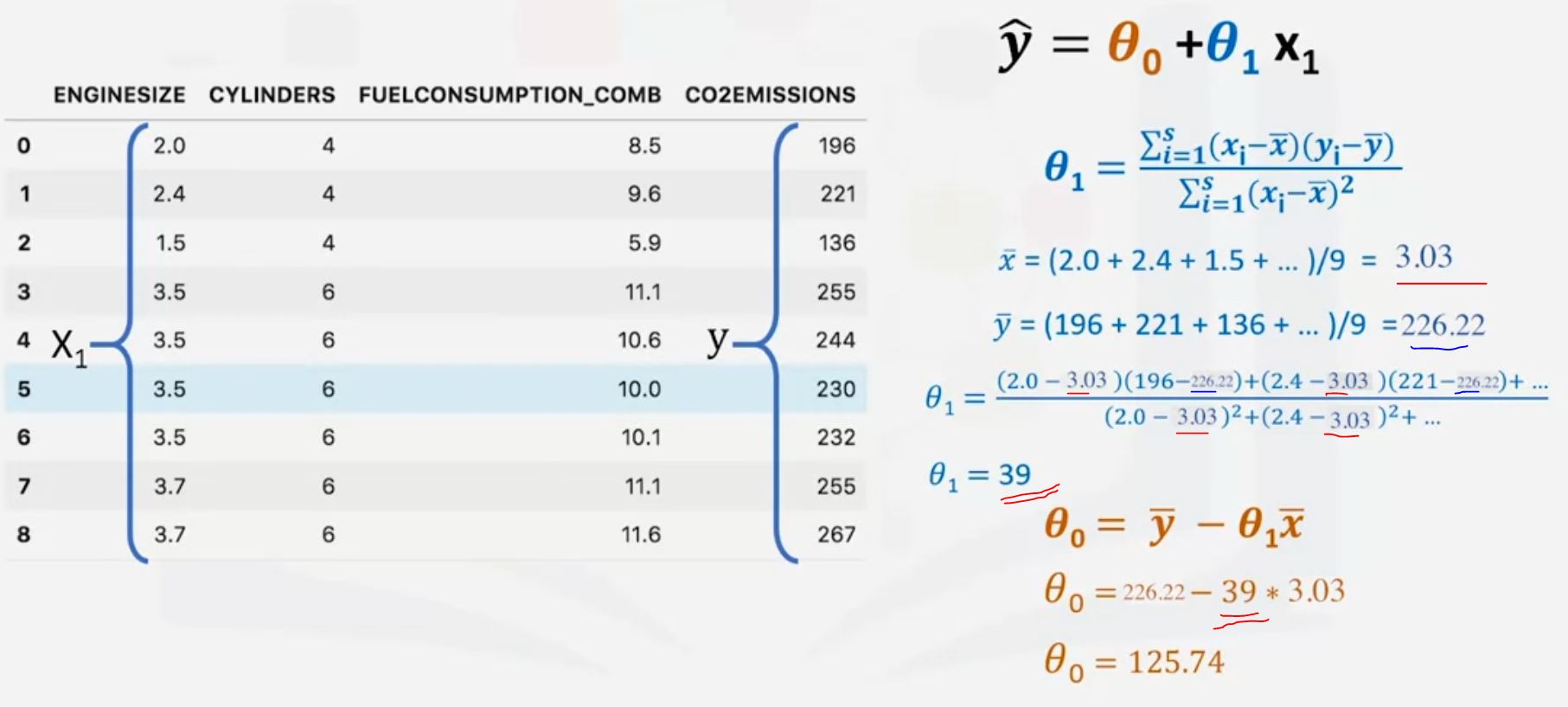
theta 0 and theta 1 are the parameters of the line that we must adjust theta 1 is known as the slope or gradient of the fitting line and theta 0 is known as the intercept theta 0 and theta 1 are also called the coefficients of the linear equation you can interpret this equation.

theta 0 and theta 1 are the intercept and slope of the line.

We can estimate them directly from our data it requires that we

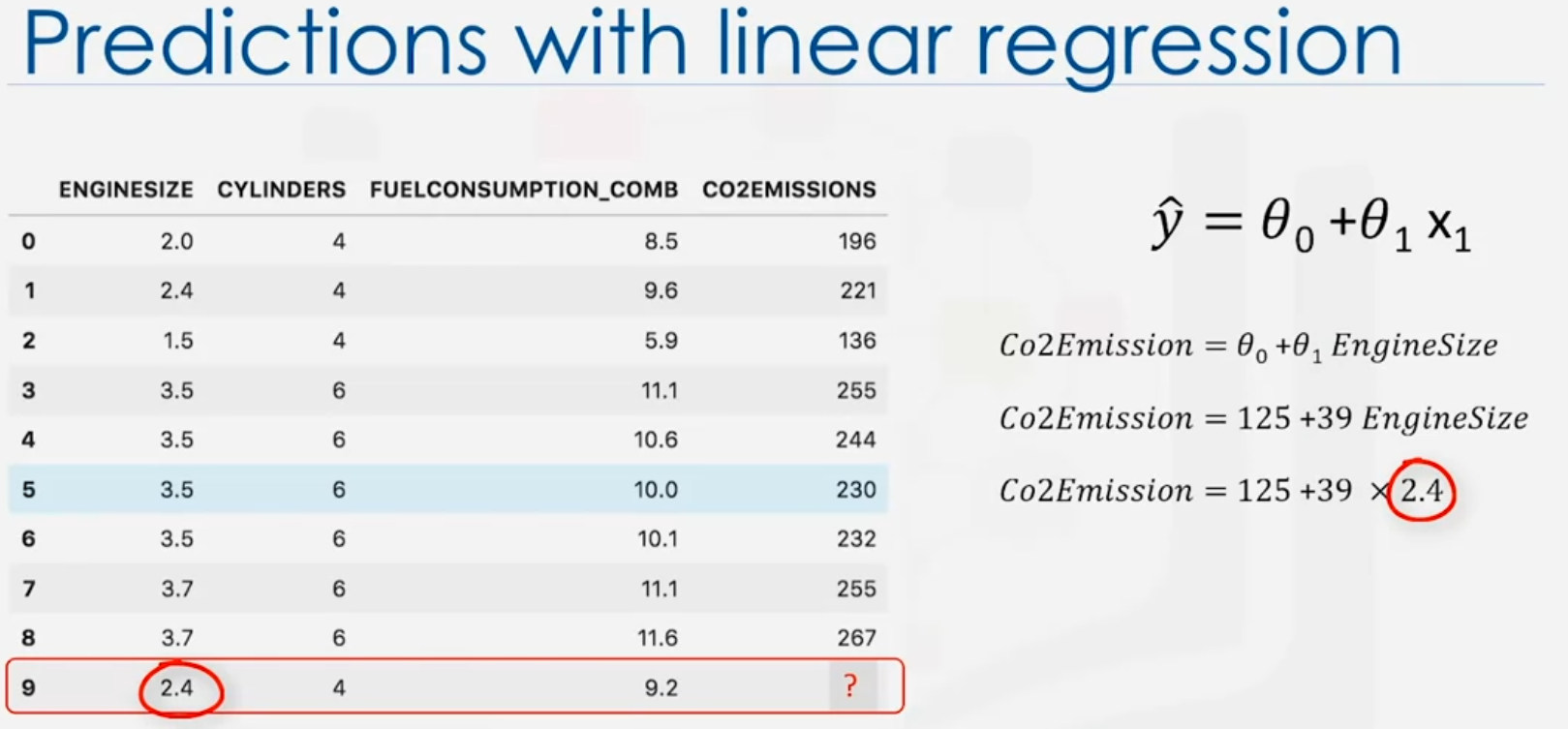
calculate the mean of the independent and dependent or target columns from the data set.

Notice that all of the data must be available to traverse and calculate the parameters.





NOW,



ANS: 218.6  
  
Now, why we use linear regression:

1. Very fast
2. No parameter tuning
3. Easy to understand, and highly interpretable