Regression for dummies is a web app; a school project of Programming in Python at NYU. Below is the landing page.

What is Linear Regression?

Linear regression is an approach to representing a relationship between an outcome variable (dependent variable) to one or more predictor variables (independent variable). In the most simplest of forms ordinary least squares, also known as simple linear regression, fits a straight line through a set of points. The line is placed in the middle of the set of points and does so by reducing the vertical distances between the points and the line.

Step 1. Upload Data | **Step 2. Choose Variable** | **Step 3. Explanation**

Select a file to upload:

Choose File | no file selected

Upload

Don't have a dataset? No problem! Click next to see how this works using our dataset!

Next

Note: The file must be a .csv file, with headers and no missing values.

Choose the Variable!

File Name: sample3.csv. Wrong file?

	Outcome	Predictor
Name:	bmi	age
		Submit
Type:	Continuous	Continuous
Suggestion:	OLS Regression model	

Go back

Descriptive Statistics

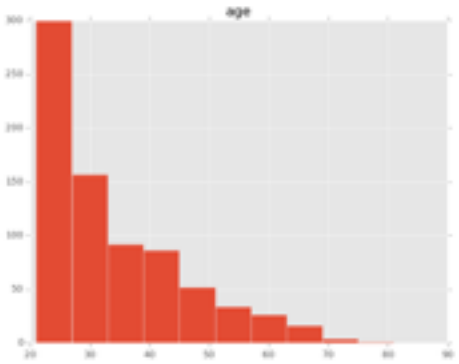
Welcome to the first step in understanding the relationship between these two variables in your dataset!

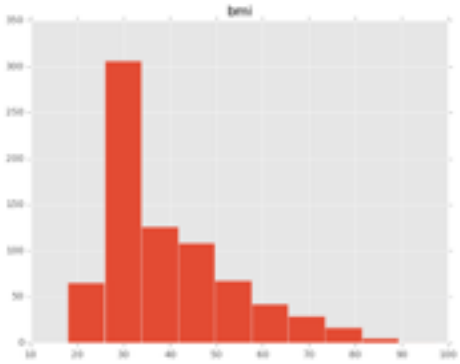
You have selected "**age**" as your independent variable and "**bmi**" as your dependent variable. What does this mean? This means that you are interested in seeing the relationship that "**age**" plays on "**bmi**". If this is an error, please click **here** to reset your dependent and independent variable.

There are a couple of descriptive coefficients that will summarize your dataset. Depending on the source of your data, these coefficients can represent an entire population or just your sample. These coefficients will explain the measure of central tendency (three different ways the center of the variable can be defined) and how stretched or squeezed the data may be.

sample3.csv	age	bmi
Average:	33.241	39.535
Standard Deviation:	11.76	13.915
Range of Values:	21.0 to 81.0	18.0 to 97.2
Median:	29.0	34.8
Mode:	22.0	26.4

For a detailed explanation for Average, Standard Deviation, Range, Median or Mode click the titles!





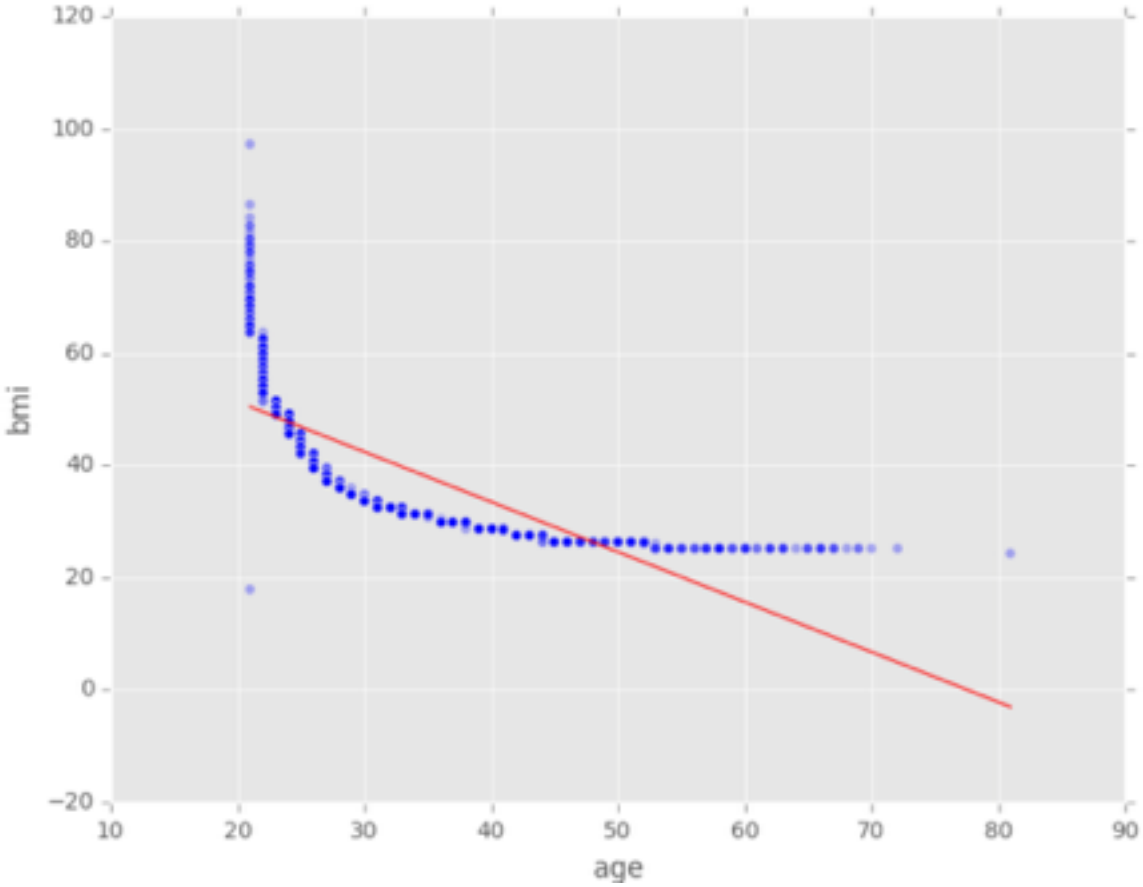
Regression Result

Regression Results

We have found that there is a very strong negative relationship between these two variables!

There are a total of 768 observations in the model.
When the value of "**age**" is zero, you can expect a value of 69.155 for "**bmi**". Additionally for every additional one-unit increase in "**age**" you can expect a 0.891 decrease in "**bmi**". The regression's "goodness of fit" of the regression model is another important value to consider in this regression. In your dataset, "**age**" explains 56.72% of the change in "**bmi**".

When looking at a bivariate analysis it is important to know the strengths and limitation of this analysis. The analysis will determine whether there is a relationship between them, seeing their covariates will show how they each change each other. It is a way to test the association between the dependent and independent variables. Remember that there are relationships between two variables that may be affected by a third variable. For example, if you are interested in the association between years of completed schooling and income, age might be effecting both of these values.


120
100
80
60
40
20
0
−20
bmi
10
20
30
40
50
60
70
80
90
age