Linear Regression

In Siraj's video, he predicted the body weight of an animal from the weight of it's brain using **linear regression**. In this section, you'll use linear regression to make prediction on life expectancy from **body mass index (BMI)** from birth. Before you do that, let's go over the tools required to build this model.

For your linear regression model, you'll be using scikit-learn's `LinearRegression`class. This class provides the function `fit()` to fit the model to your data.

```
>>> from sklearn.linear_model import LinearRegression
>>> model = LinearRegression()
>>> model.fit(x_values, y_values)
```

In the example above, the `model` variable is a linear regression model that has been fitted to the data `x_values` and `y_values`. Fitting the model means finding the best line that fits the training data. Let's make two predictions using the model's `predict()`function.

```
>>> print(model.predict([ [127], [248] ]))
[[ 438.94308857, 127.14839521]]
```

The model returned an array of predictions, one prediction for each input array. The first input, `[127]`, got a prediction of `438.94308857`. The seconds input, `[248]`, got a prediction of `127.14839521`. The reason for predicting on an array like `[127]` and not just `127`, is because you can have a model that makes a prediction using multiple features. We'll go over using multiple variables in linear regression later in this lesson. For now, let's stick to a single value.

Linear Regression Quiz

In this quiz, you'll be working with data on the average life expectancy at birth and the average the BMI for males across the world. The data comes from **Gapminder**.

The data file can be found under the "bmi_and_life_expectancy.csv" tab in the quiz below. The data includes the country the person was born in. This data is under the "Country" column. The Life expectancy for a person in that country in the "Life expectancy" column. The media BMI of a child born in that country as "BMI". You'll predict the life expectancy using this BMI.

You'll need to complete each of the following steps:

1. Load the data

- The data is in the file called "bmi_and_life_expectancy.csv".
- Use pandas `read_csv` to load the data into a dataframe.
- Assign the dataframe to the variable `bmi_life_data`.

2. Build a linear regression model

- Create a regression model using scikit-learn's `LinearRegression`and assign it to `bmi_life_model`.
- Fit the model to the data.

3. Predict using the model

- Predict using a BMI of 21.07931 and assign it to the variable `laos_life_exp`.