**Linear Regression**

In this section, you'll use linear regression to predict life expectancy from [body mass index (BMI)](https://en.wikipedia.org/wiki/Body_mass_index). 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](http://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LinearRegression.html) class. This class provides the function [fit()](http://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LinearRegression.html#sklearn.linear_model.LinearRegression.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()](http://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LinearRegression.html#sklearn.linear_model.LinearRegression.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 second 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 BMI for males across the world. The data comes from [Gapminder](https://www.gapminder.org/).

The data file can be found under the "bmi\_and\_life\_expectancy.csv" tab in the quiz below. It includes three columns, containing the following data:

* **Country** – The country the person was born in.
* **Life expectancy** – The average life expectancy at birth for a person in that country.
* **BMI** – The mean BMI of males in that country.

**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](http://pandas.pydata.org/pandas-docs/stable/generated/pandas.read_csv.html) to load the data into a dataframe (don't forget to import pandas!)
* Assign the dataframe to the variable bmi\_life\_data.

**2. Build a linear regression model**

* Create a regression model using scikit-learn's [LinearRegression](http://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LinearRegression.html) 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.