Leonardo Gomez

March 6, 2018

CS 596

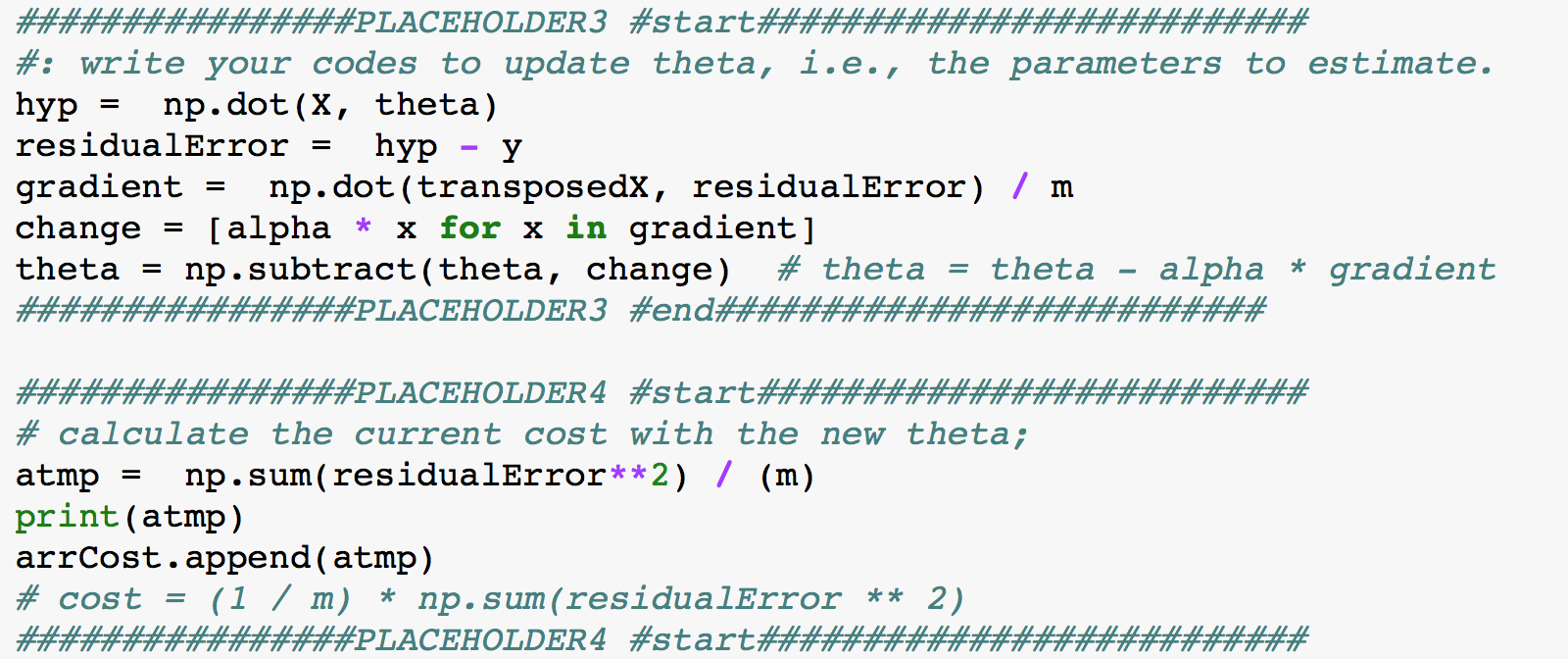
Homework 3

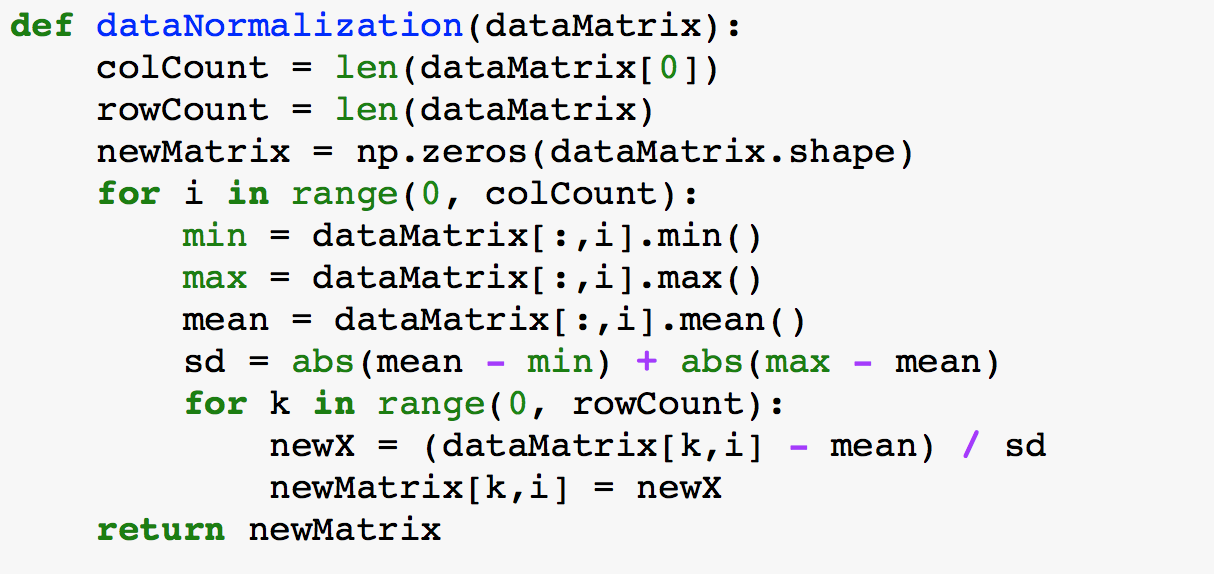
**Algorithm:**

For this problem, I used a Multivariable Linear Regression algorithm (MLR). Statistically speaking, multivariate analysis refers to statistical models that have 2 or more dependent or outcome variables. This type of statistical model can be used to attempt to assess the relationship between several variables; one can assess independent relationships while adjusting for potential confounders.

**Adjustments to the code**:

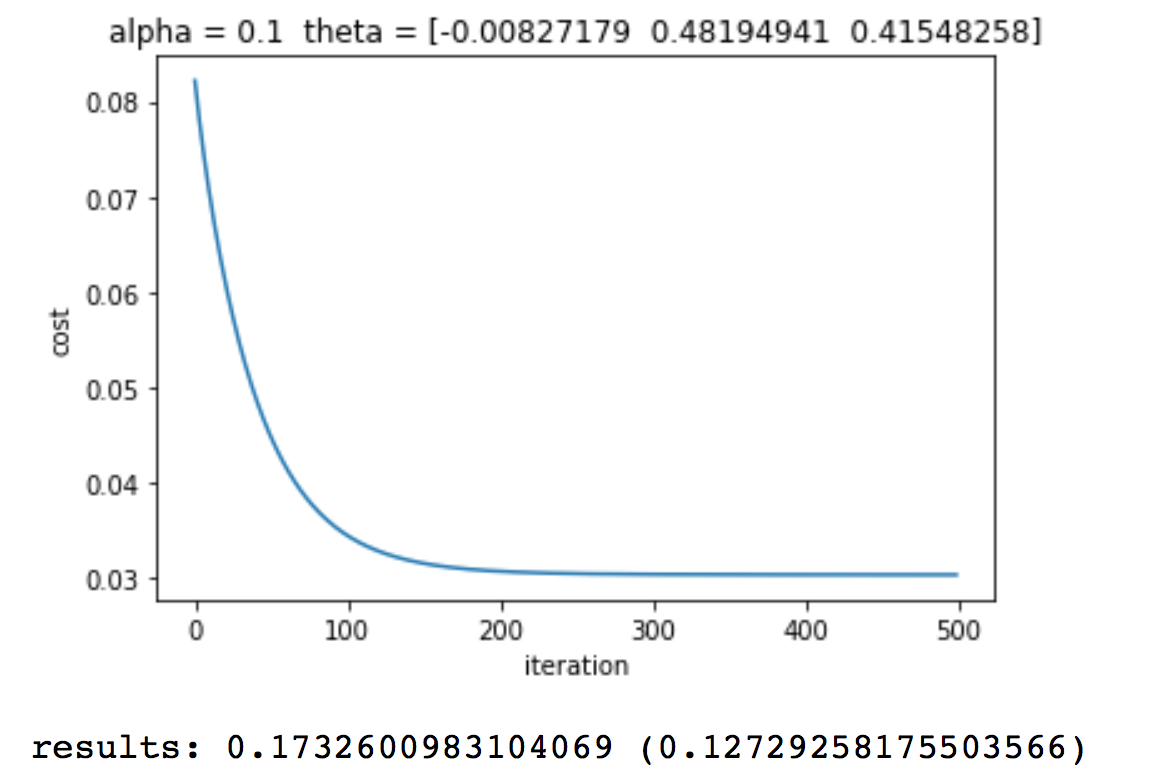
Besides the require changes made in PLACEHOLDER 1,2,3, AND 4. I added a few changes to the gradient and decent function. I created a section for hypothesis in which I do the dot product of the X and the θ passed down to the function. Once we have the hypothesis we subtract the Y to get the residual error. Then the gradient becomes the dot product between the transposed X and the residual error divided by the number of training samples.



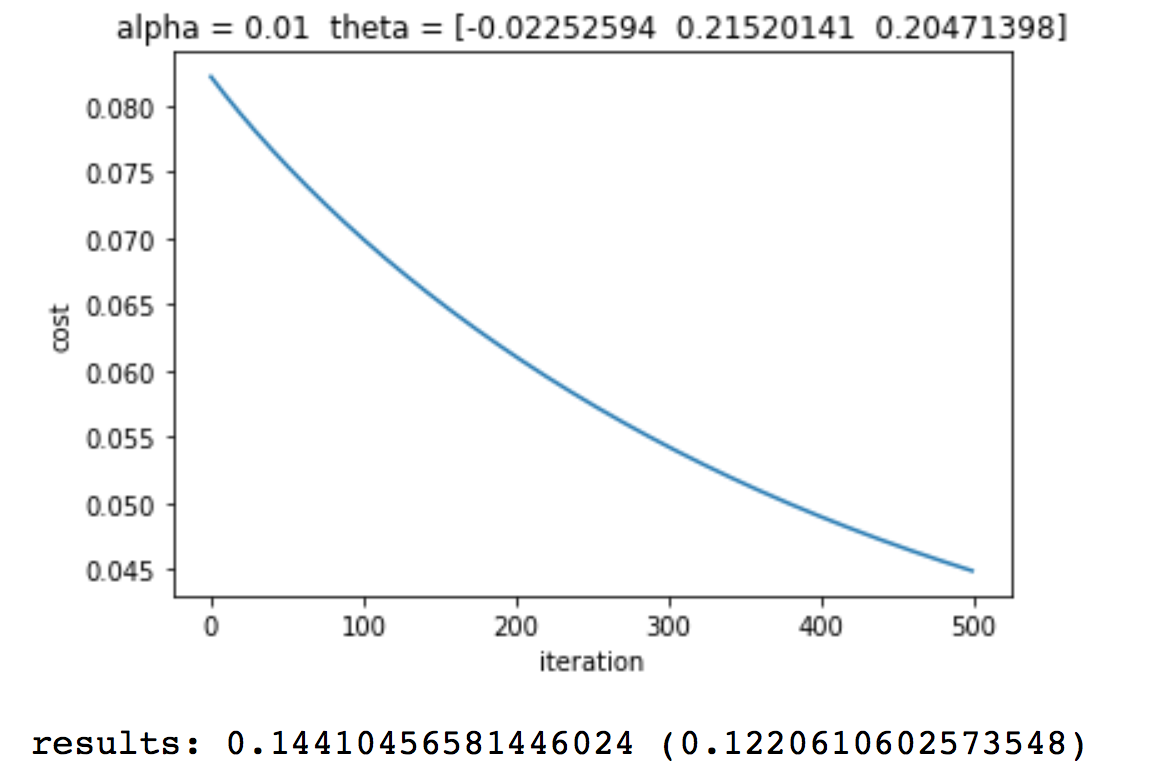
 To normalize the data, I used Mean Normalization. This involved subtracting the average value for an input variable from the values for that input variable resulting in a new average value divided by the Standard Deviation. This was noted by the formula: where is the average of all values for feature (i) and is the Standard Deviation.

**Results:**

ALPHA = 0.1

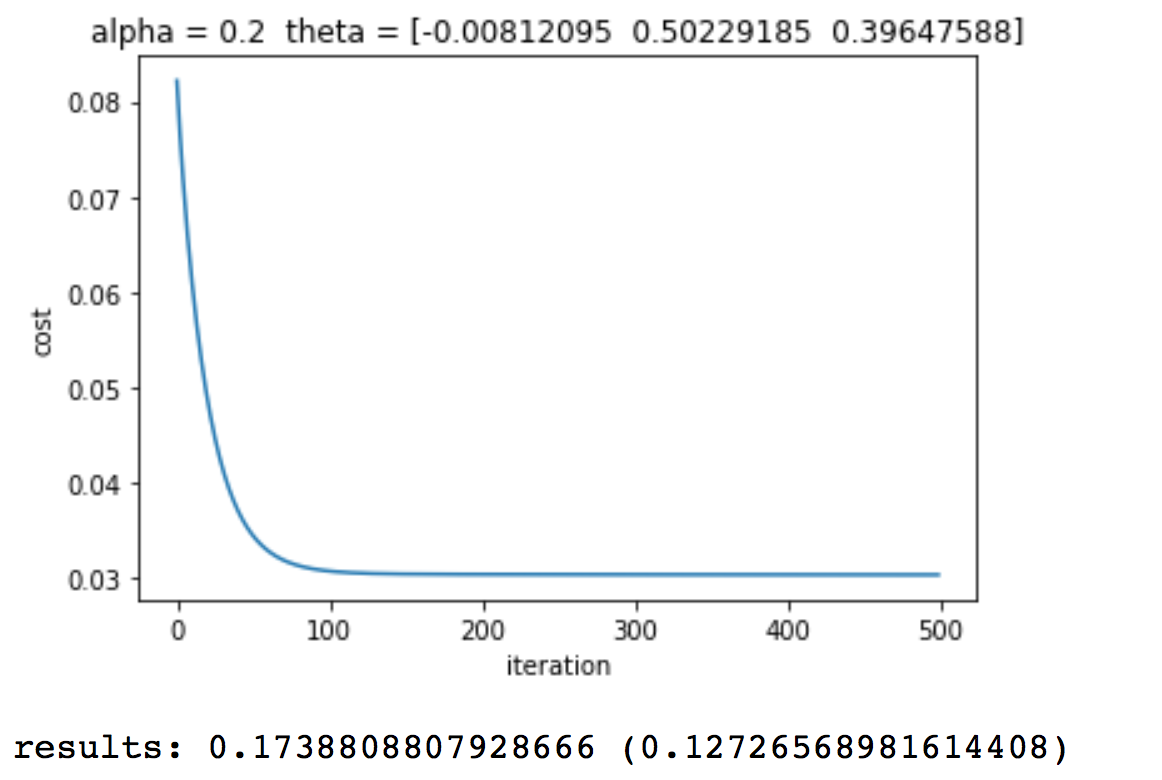
MAX\_ITER = 500

ALPHA = 0.01

MAX\_ITER = 500

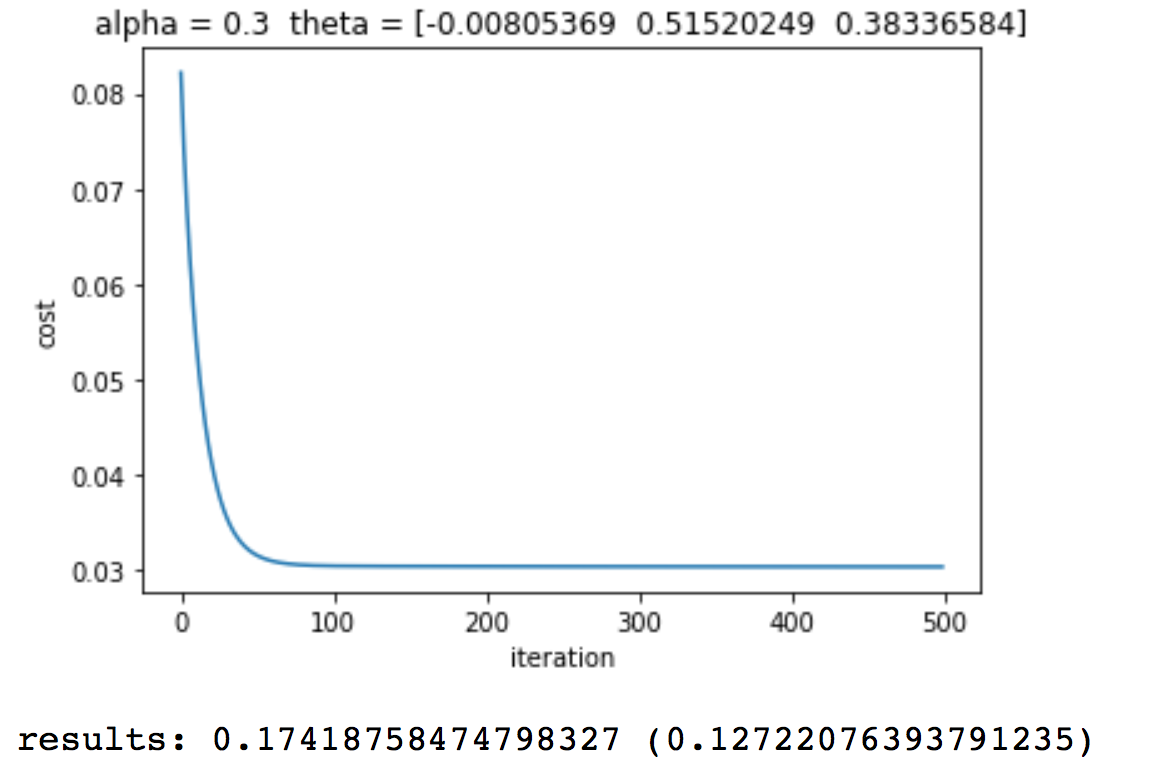
ALPHA = 0.2

MAX\_ITER = 500



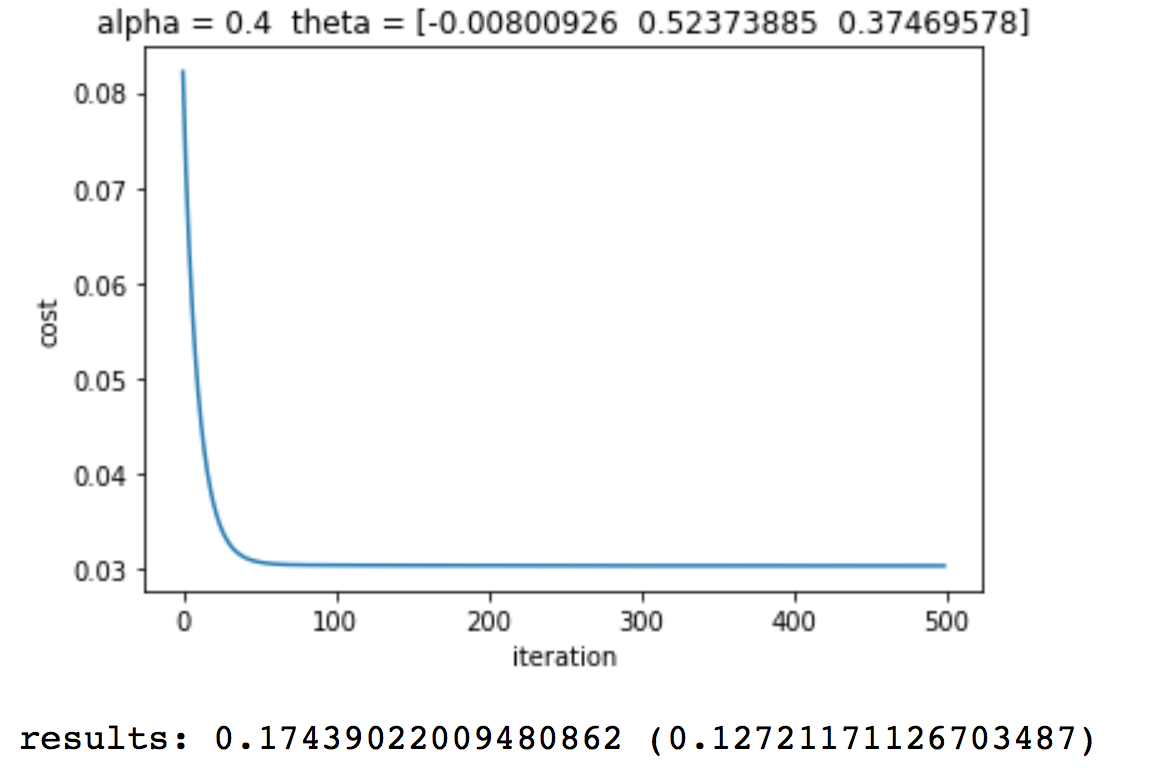
ALPHA = 0.3

MAX\_ITER = 500



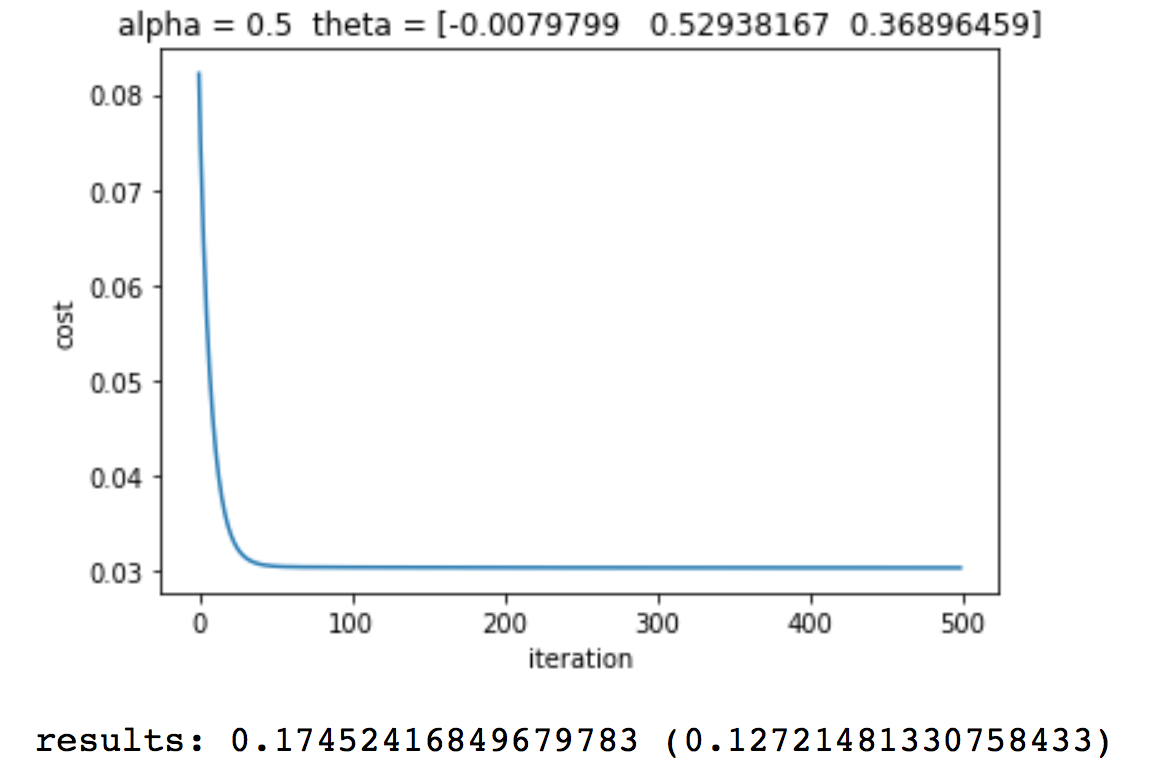
ALPHA = 0.4

MAX\_ITER = 500



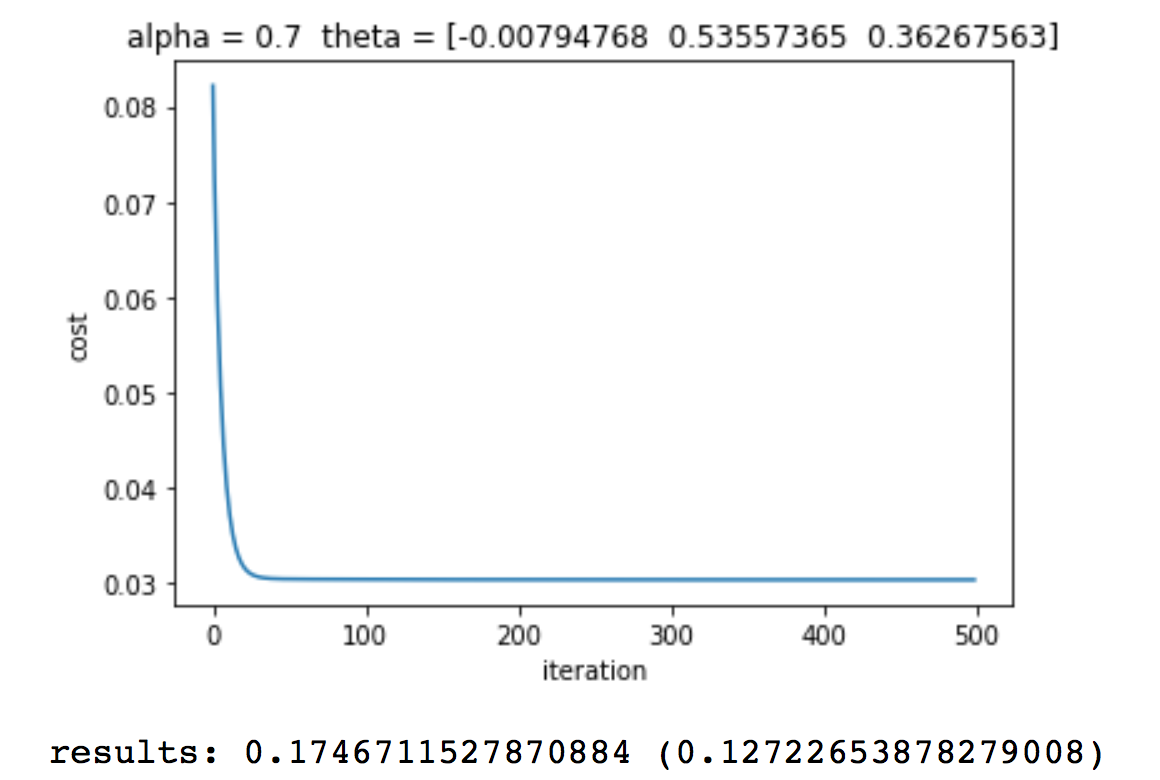
ALPHA = 0.5

MAX\_ITER = 500



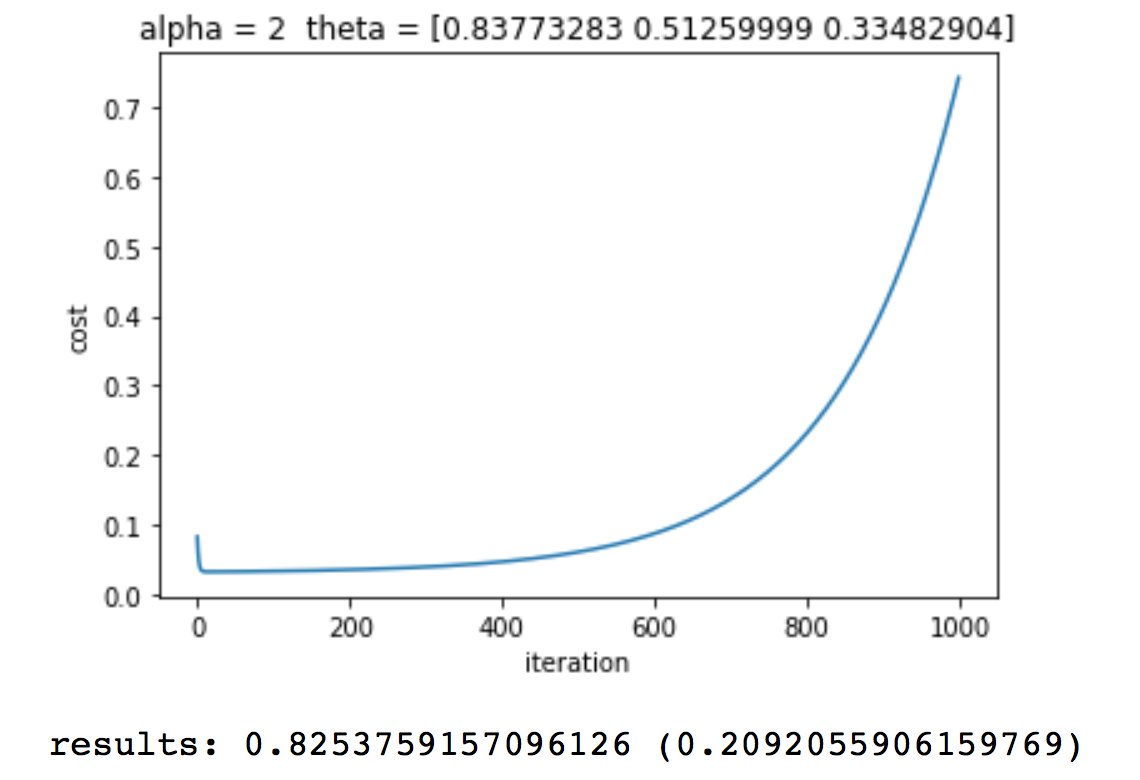
ALPHA = 0.7

MAX\_ITER = 500



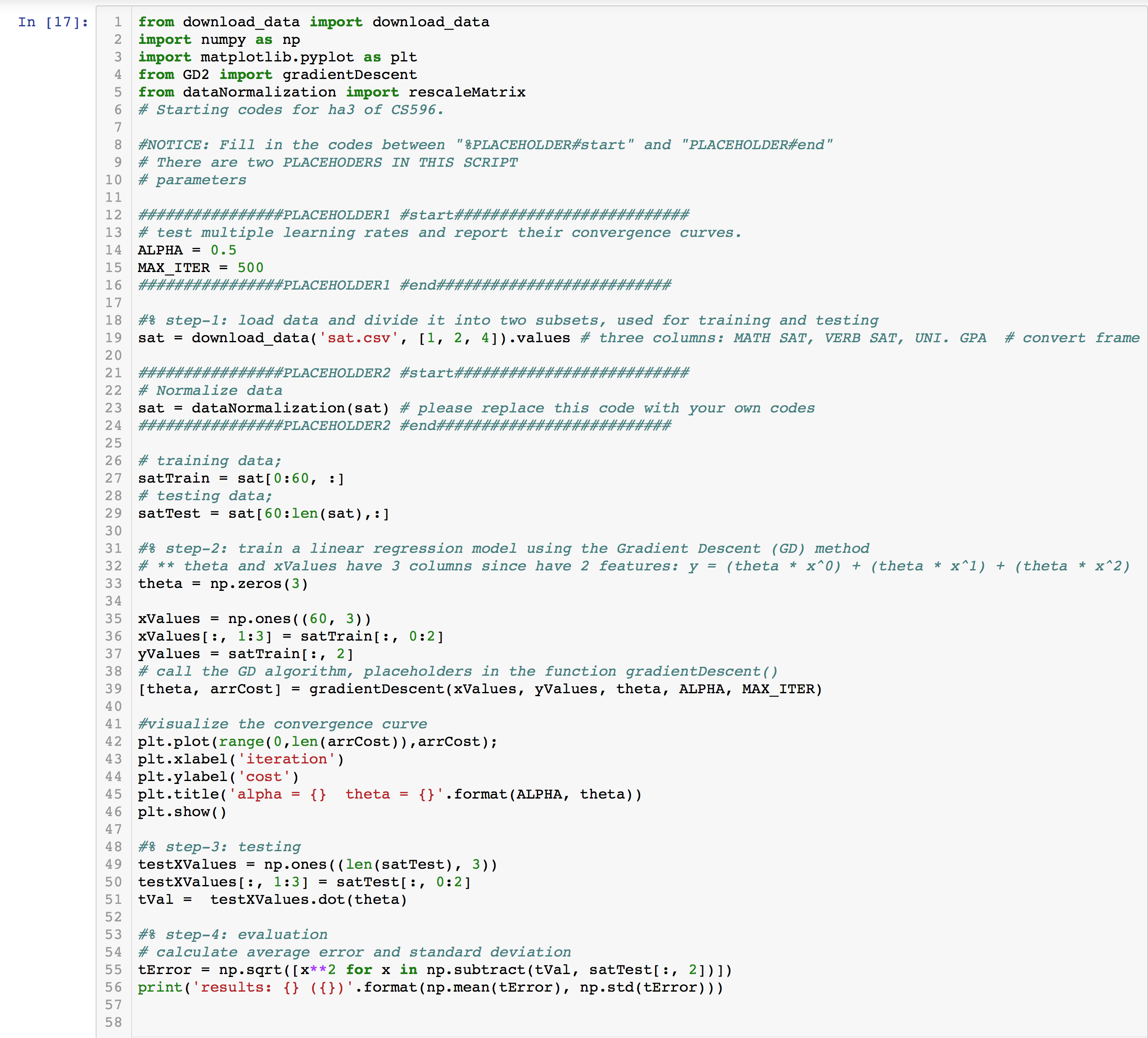
ALPHA = 2

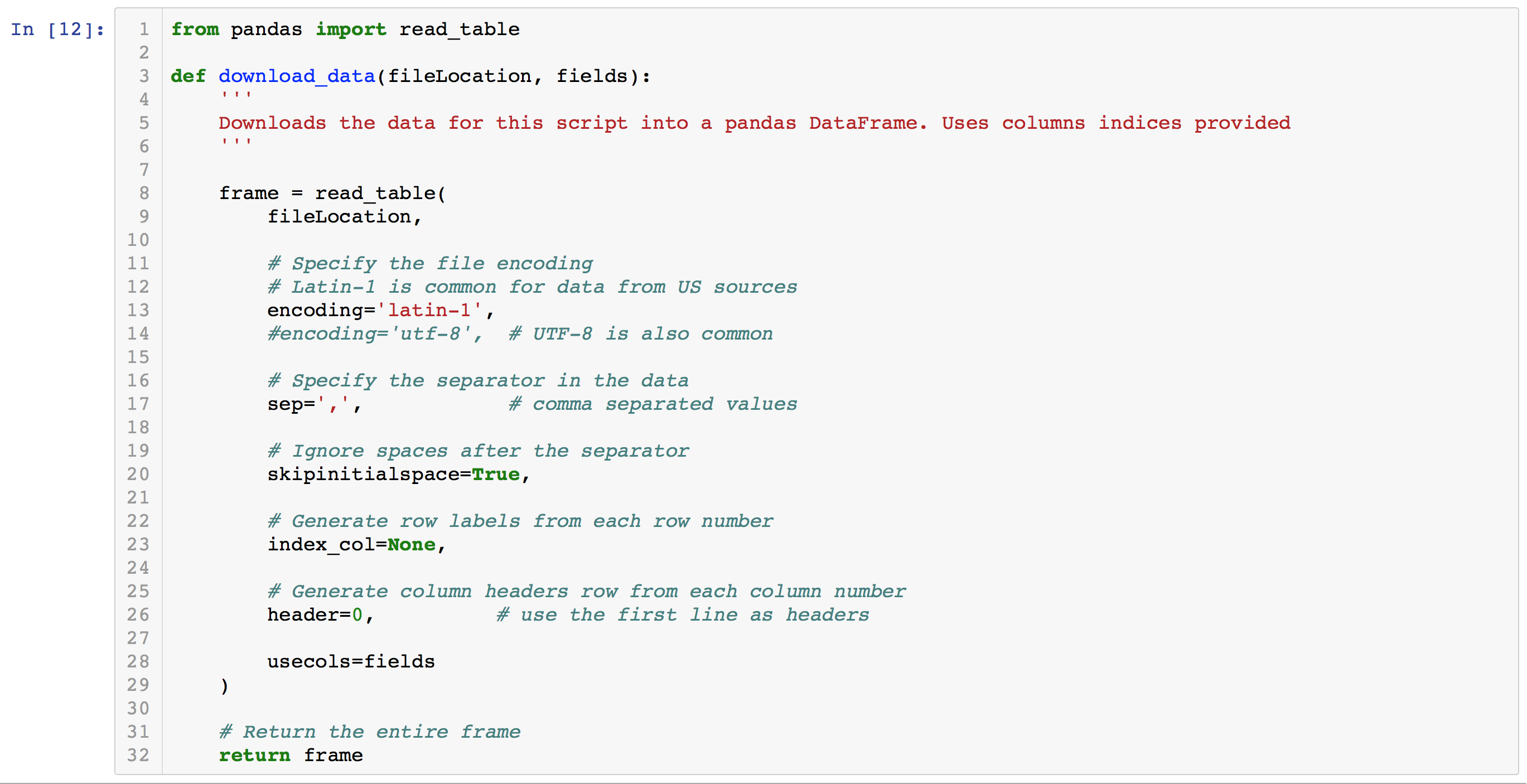
MAX\_ITER = 1000



**Results Analysis:**

In the previous results, we noticed that if we kept increasing the learning rate the cost of every iteration becomes lower. However, when we increase it too much the cost starts to go up exponentially. This is because the function skips too much and it overshoots the local minima. Also, something to note is that Alpha must remain below 2 or the function will not overshoot its target. Therefore, Alpha=2 is the cutoff point. Furthermore, all the learning rates that we tested below 2 start to flatten out around 100 iterations, this means that there is no need to have more than 100-150 iterations. Even with fewer iterations the difference in error is so insignificant that it does not matter.

**Code:**

****

