## Congratulations! You passed!

Grade received 100% To pass 70% or higher

Go to next item

## Practice quiz: Multiple linear regression

Latest Submission Grade 100%

1. In the training set below, what is  $x_4^{(3)}$ ? Please type in the number below (this is an integer such as 123, no decimal points).

1/1 point

| Size in feet <sup>2</sup> | Number of<br>bedrooms | Number of floors | Age of home in years | Price (\$) in<br>\$1000's |
|---------------------------|-----------------------|------------------|----------------------|---------------------------|
| X <sub>1</sub>            | X2                    | Хз               | X4                   |                           |
| 2104                      | 5                     | 1                | 45                   | 460                       |
| 1416                      | 3                     | 2                | 40                   | 232                       |
| 1534                      | 3                     | 2                | 30                   | 315                       |
| 852                       | 2                     | 1                | 36                   | 178                       |
|                           |                       |                  |                      |                           |

| 30 |   |
|----|---|
|    |   |
|    | Correct   |
|    | Yes! $x_4^{(3)}$ is the 4th feature (4th column in the table) of the 3rd training example (3rd row in the table). |

1/1 point  $Which of the following are potential benefits of vectorization? \ Please choose the best option. \\$ O It makes your code run faster O It can make your code shorter O It allows your code to run more easily on parallel compute hardware All of the above

**⊘** Correct Correct! All of these are benefits of vectorization!

 $\textbf{3.} \quad \text{True/False? To make gradient descent converge about twice as fast, a technique that almost always works is to} \\$ double the learning rate alpha.

1/1 point

O True

False

**⊘** Correct

Doubling the learning rate may result in a learning rate that is too large, and cause gradient descent to fail to find the optimal values for the parameters  $\boldsymbol{w}$  and  $\boldsymbol{b}$ .