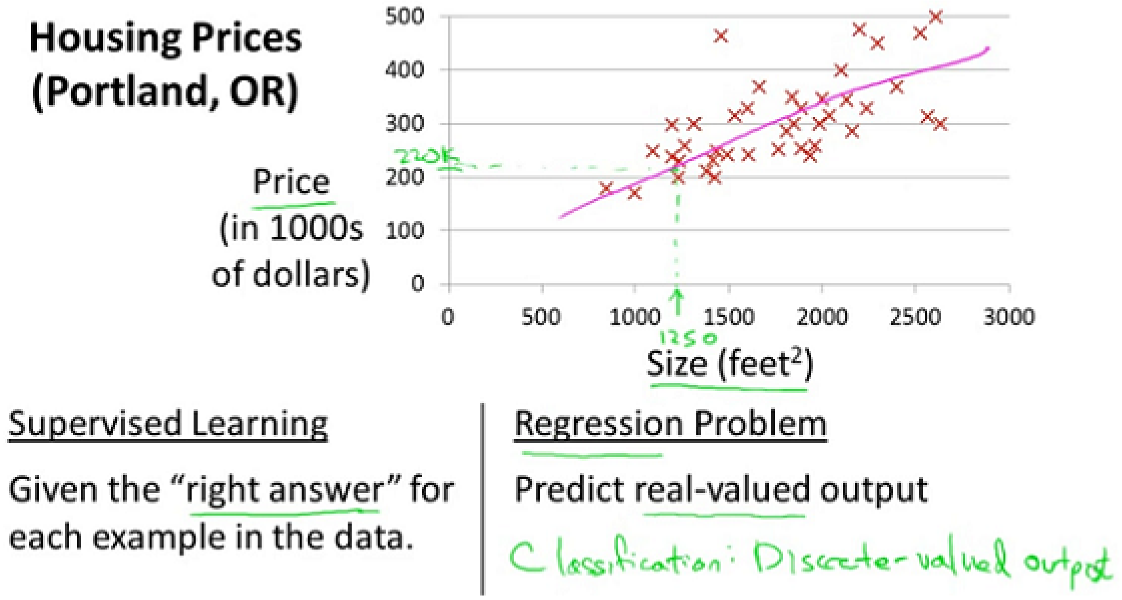
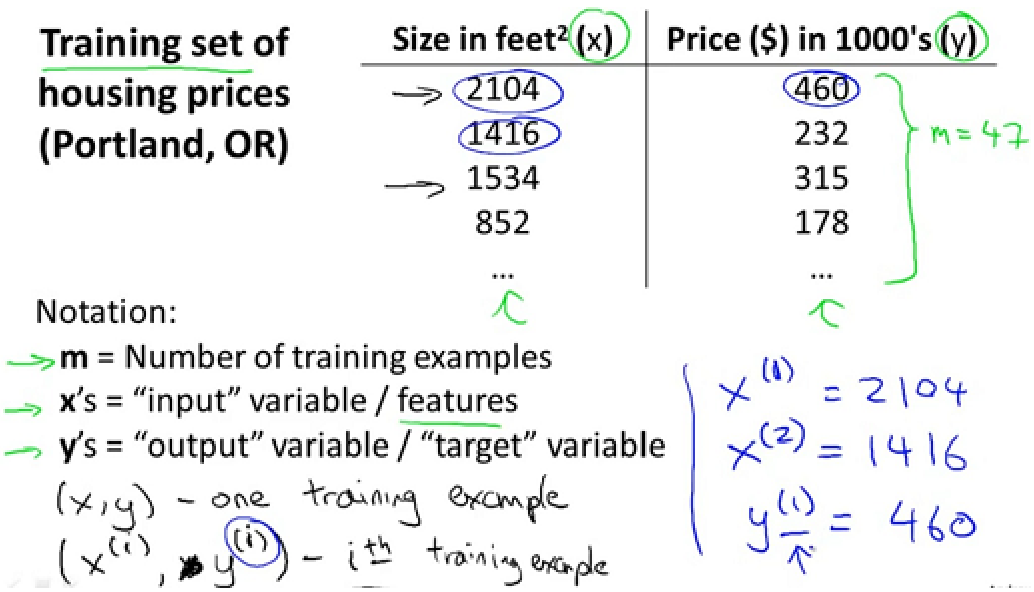
Model Representation

We're going to use a data set of housing prices from the city of Portland, Oregon. Let's say that given this data set, we have a friend that's trying to sell a house and let's see if friend's house is size of 1250 square feet and we want to tell them how much they might be able to sell the house for. One thing we could do is fit a model and we fit a straight line to this data.

And we could tell our friend that he can sell the house for around $ 220,000. This is an example of supervised learning algorithm and moreover this is an example of a regression problem where the term regression refers to the fact that we are predicting a real-valued output namely the price.



In supervised learning, we have a data set and this data set is called a training set. For housing prices example, we have a training set of different housing prices and our job is to learn from this data how to predict prices of the houses. In the data set we have 47 rows in the table. Then we have 47 training examples and m equals 47.



- (x, y) - denote a single training example
- $(x^{(i)}, y^{(i)})$ - denote a specific training example

The superscript i in parentheses that's just an index into our training set and refers to the i^{th} row in the table. For example, $x^{(1)}$ refers to the input value for the first training example so that's 2104, $x^{(2)}$ will be equal to 1416, $y^{(1)}$ will be equal to 460 and so on.

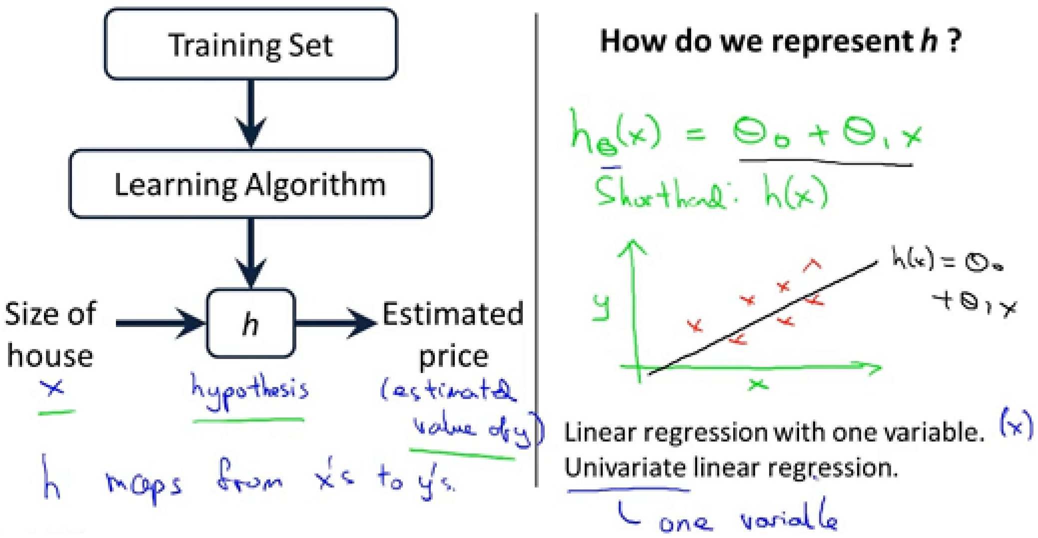
Video Question: Consider the training set shown below. $(x^{(i)}, y^{(i)})$ is the i^{th} training example. What is $y^{(3)}$?

Size in feet^2	Price ($) in 1000's (y)
2104	460
1416	232
1534	315
852	178
...	...

- 1416
- 1534
- 315
- 0

We saw that with the training set like our training set of housing prices and we feed that to our learning algorithm. It's the job of a learning algorithm to then output a function which by convention is usually denoted lowercase h and h stands for hypothesis.

And what the job of the hypothesis is?, it's a function that takes as input the size of a house like the size of the new house our friend's trying to sell so it takes in the value of "x" and it tries to output the estimated value of "y" for the corresponding house.



In supervised learning we represent hypotesis with h_θ or h. We are going to predict that "y" is a linear function of x, and what this function is doing, is predicting that "y" is some straight line function of "x". This model is called linear regression or this model, for example, is actually linear regression with one variable, with the variable being x. And another name for this model is univariate linear regression.

Summary

We'll use $x^{(i)}$ to denote the "input" variables (living area in this example), also called input features, and $y^{(i)}$ to denote the "output" or target variable that we are trying to predict (price).

A pair $(x^{(i)}, y^{(i)})$ is called a training example, and the dataset that we'll be using to learn—a list of m training examples $(x^{(i)}, y^{(i)}); i = 1, \ldots, m$ —is called a training set. We will also use X to denote the space of input values, and Y to denote the space of output values. In this example, $X = Y = \mathbb{R}$.

To describe the supervised learning problem slightly more formally, our goal is, given a training set, to learn a function $h : X \rightarrow Y$ so that $h(x)$ is a “good” predictor for the corresponding value of y. For historical reasons, this function h is called a hypothesis.

When the target variable that we’re trying to predict is continuous, such as in our housing example, we call the learning problem a regression problem. When y can take on only a small number of discrete values (such as if, given the living area, we wanted to predict if a dwelling is a house or an apartment, say), we call it a classification problem.