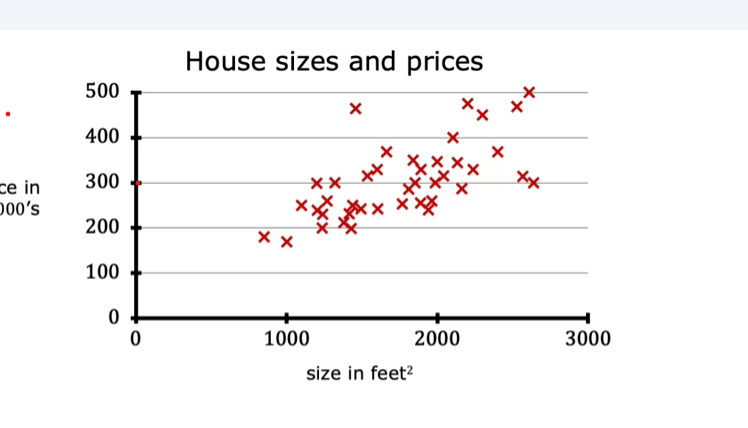
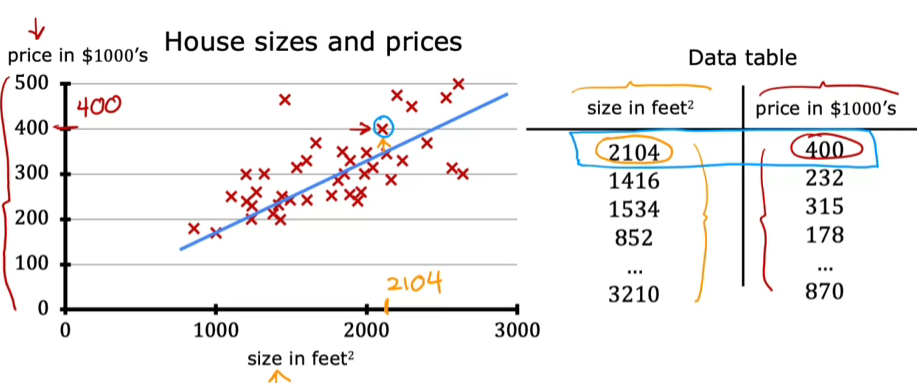
**LINEAR REGRESSION MODEL**

**Example: Predict the price of the house**

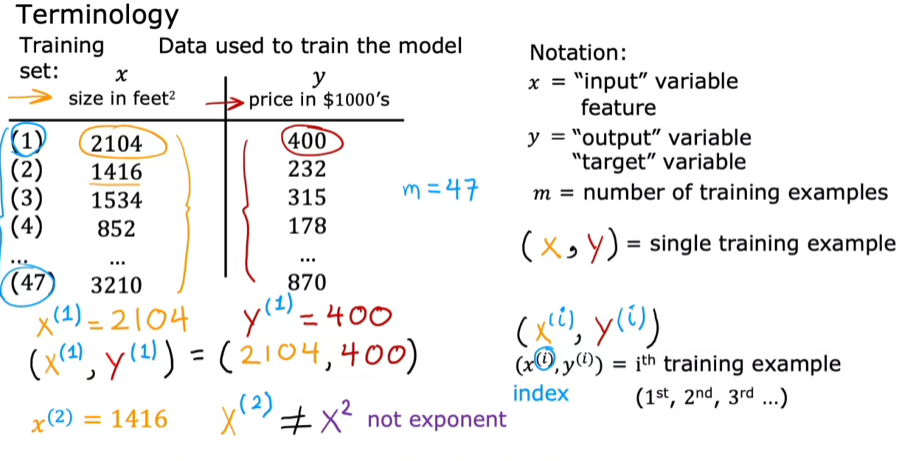


**Using the linear regression model. We can plot a straight line like this.**



**Here are some terminology and definitions:**

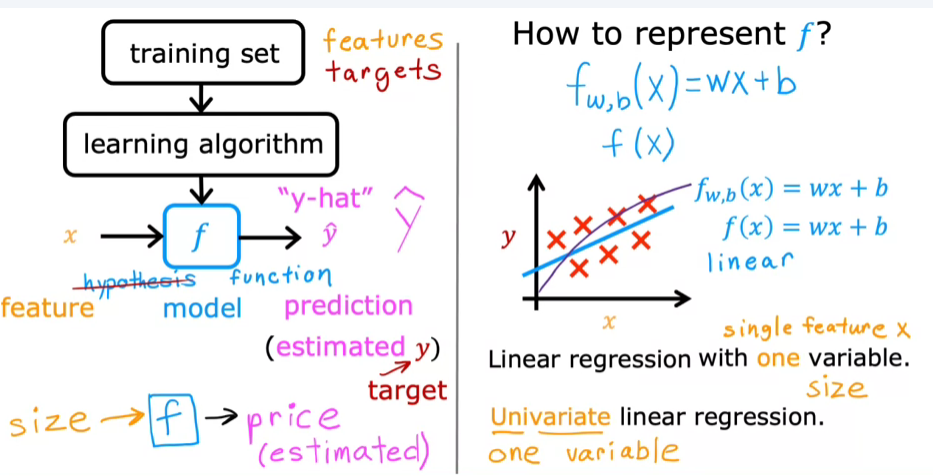
1. **Supervised learning requires a training set, which consists of input-output pairs where the correct outputs (prices) are known for the inputs (house sizes).**
2. **The input variable is denoted as lowercase x (e.g., size of the house), while the output variable is denoted as lowercase y (e.g., price of the house).**



**The learning algorithm produces a function, denoted as f, that takes new inputs and outputs predictions (y-hat), estimating the true values (y). Where y is the true value for that training example, referred to as the output variable, or “target”.**

**f(x) = w. x + b where w, b are parameters that influence the predictions based on input features on x.**

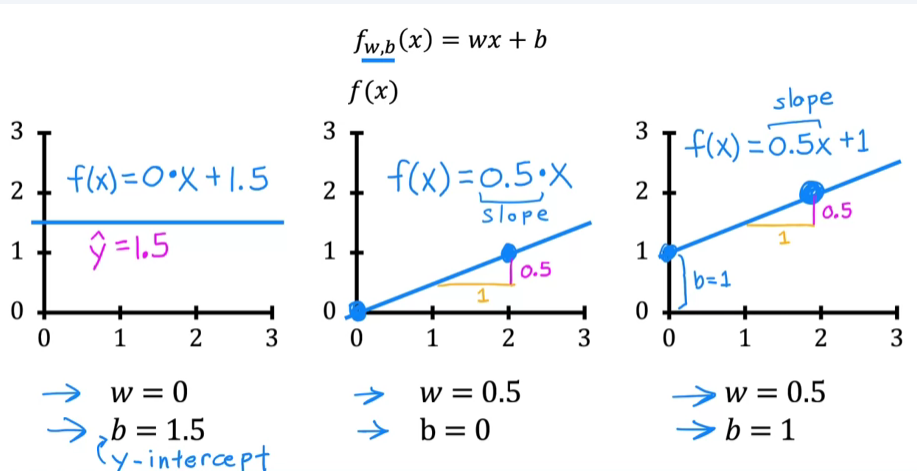
**The model generates a best-fit line on a graph, making predictions for the output target y using this linear function.**



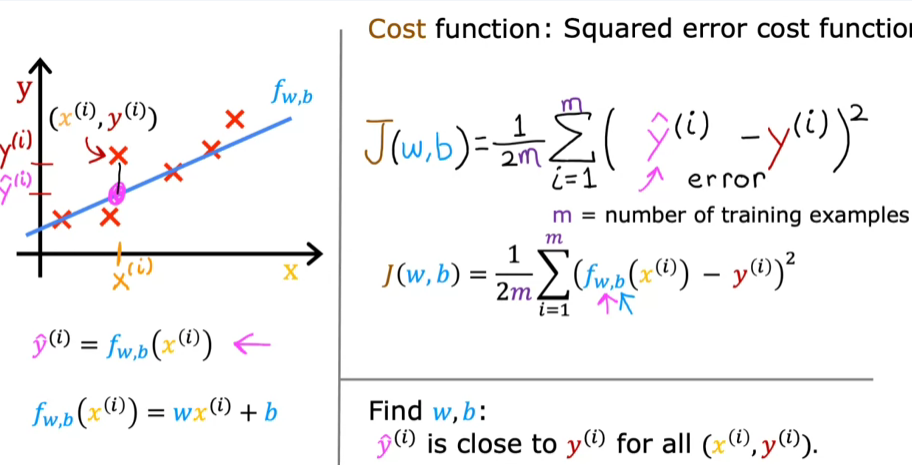
**COST FUNCTION FORMULA**

**f(x) = wx + b, here w, b are called parameters, w is called weight and b is called bias.**

* **The cost function measures how well the linear regression model predicts the output targets (y) based on input features (x). It helps in adjusting the model parameters (w and b) to improve predictions.**



* **The cost function is defined as the average of the squared errors between predicted values (y hat) and actual targets (y), allowing for a standardized measure of model performance.**

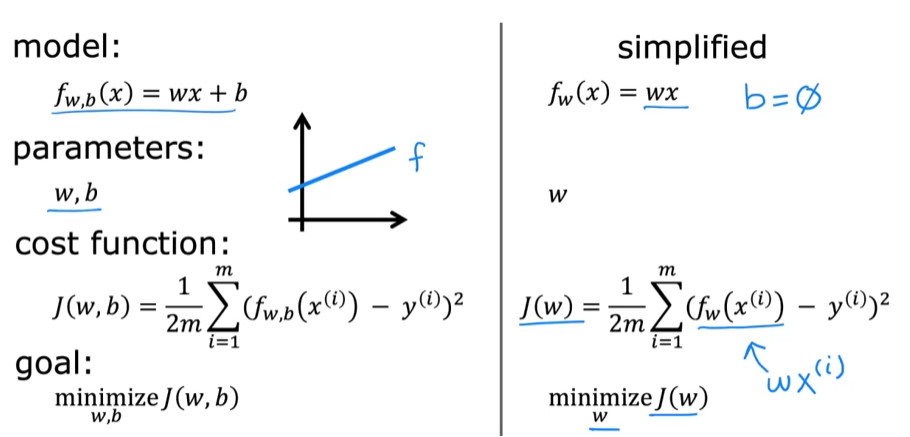


* **The parameters w (weights) and b (bias) determine the slope and intercept of the linear function used in the regression. Different values of w and b result in different lines on a graph, affecting the model's predictions.**
* **The goal is to find optimal values for w and b that minimize the cost function, ensuring the line fits the training data as closely as possible**
* **A well-fitting line visually passes close to the training examples, indicating that the model is making accurate predictions. The cost function quantifies this fit by calculating the total squared error across all training examples.**

**COST FUNCTION INTUITION**

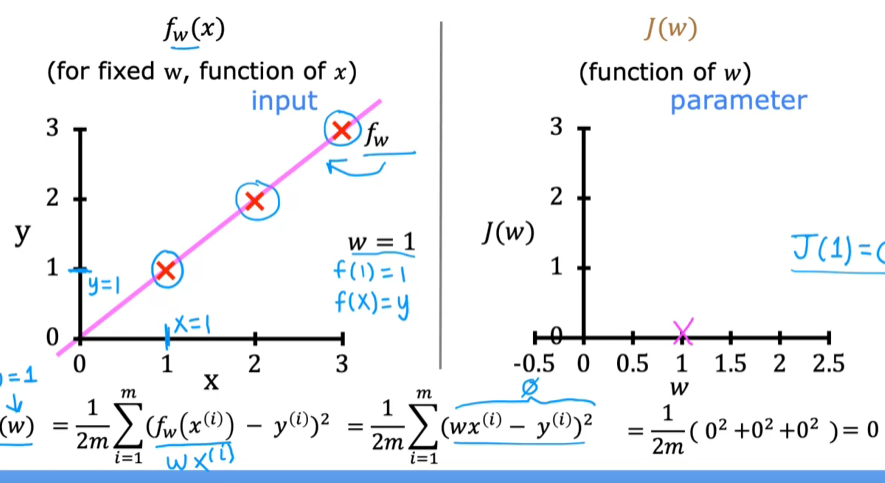
UNDERSTANDING THE COST FUNCTION

* **The cost function, denoted as J, measures the difference between the model's predictions and the actual values for y. The goal is to minimize J to achieve a better fit for the training data.**
* **In a simplified model, where the parameter b is set to 0, the cost function becomes a function of just one parameter, w, allowing for a clearer visualization of how changes in w affect the cost.**

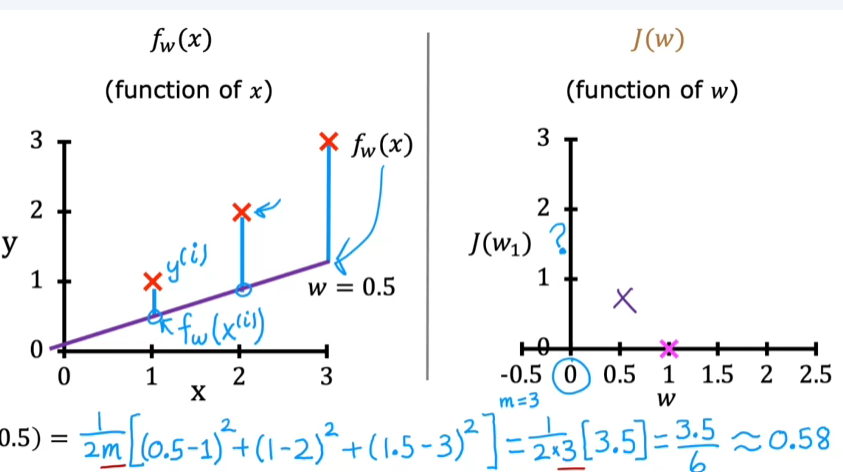


**Visualizing The Relationship Between F And J**

**When w = 1.**

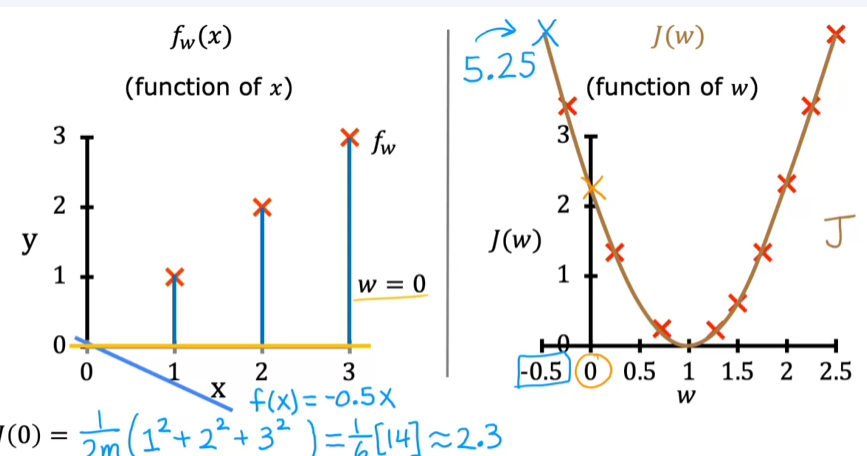


**When w = 0.5**



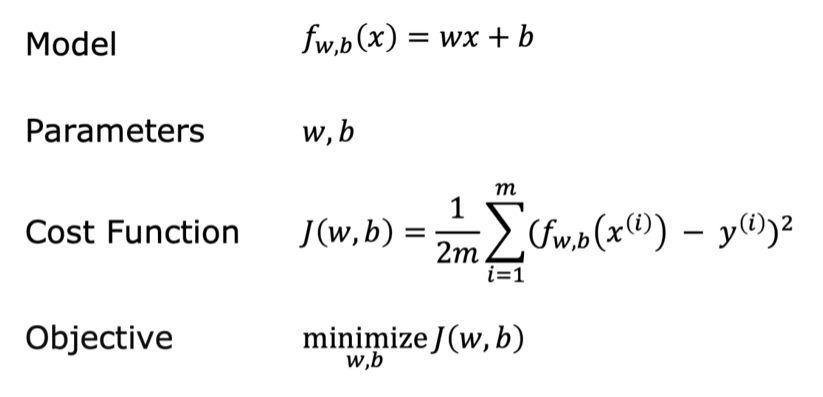
* **As different values of w are chosen, the corresponding straight line fit (f) changes, and each value of w results in a specific cost (J). For example, when w equals 1, the cost J is 0, indicating a perfect fit.**
* **By plotting the cost function J against various values of w, one can observe how the cost varies, helping to identify the optimal value of w that minimizes J.**

**For all values of w**

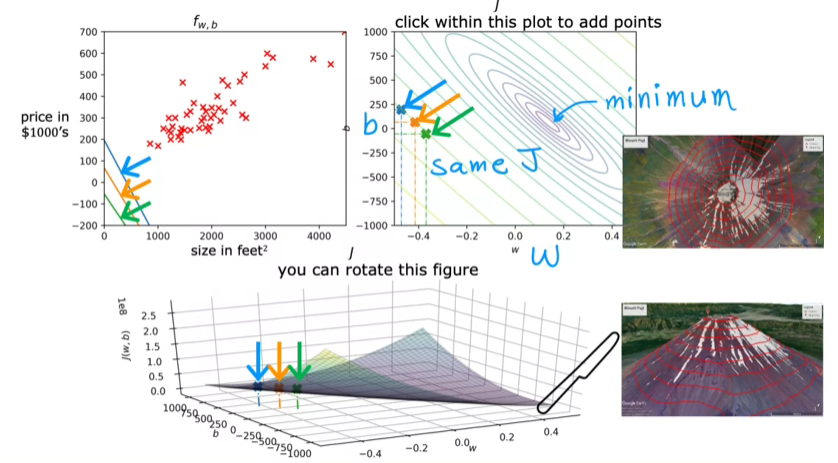


* **The ultimate goal in linear regression is to find the parameters (w and b) that minimize the cost function J. A smaller J indicates a better fit of the model to the training data.**
* **In the example discussed, the best choice for w was found to be 1, which resulted in a straight line that fits the training data very well.**

**VISUALISING THE COST FUNCTION**

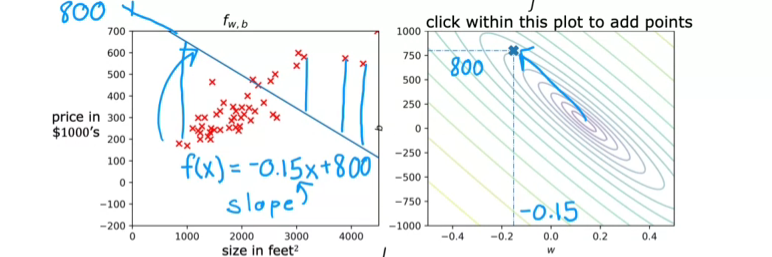
****

* **The 3D surface plot resembles a bowl shape, where each point represents a specific combination of w and b, and the height indicates the cost function's value.**
* **Contour plots provide a 2D view of the same cost function, showing ellipses that represent points with equal cost values, making it easier to visualize the optimization landscape.**

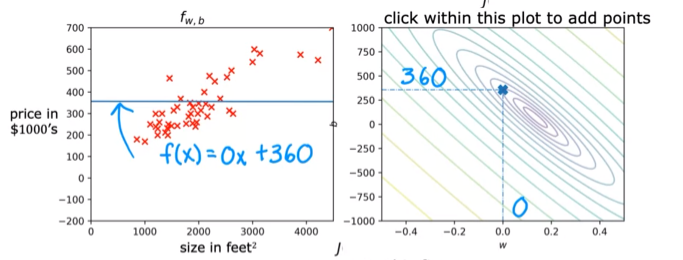


**Let us see for different values of w and b**

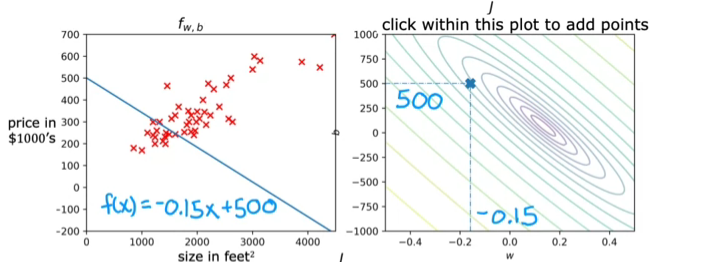
* 1. **w = -0.15 and b = 800**



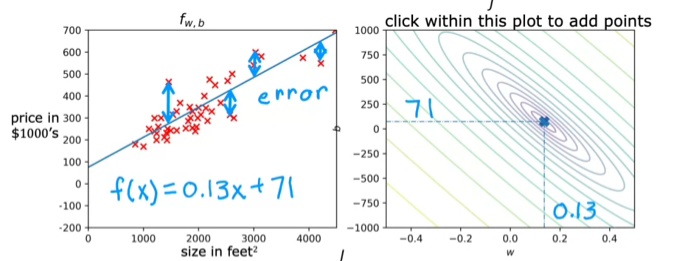
* 1. **w = 0 and b = 360**



* 1. **w = -0.15 and b = 500**



* 1. **w = 0.13 and b = 71**



* **A better fit line is closer to the minimum point on the cost graph, indicating a lower sum of squared errors between predicted and actual values.**

**To find the values of a w and b we are going to learn about an algorithm called the gradient descent,**

* **Gradient descent is introduced as an efficient algorithm for finding optimal values of ( w ) and ( b ) that minimize the cost function ( j ).**
* **This algorithm is crucial not only for linear regression but also for training more complex machine learning models, making it a foundational concept in AI.**