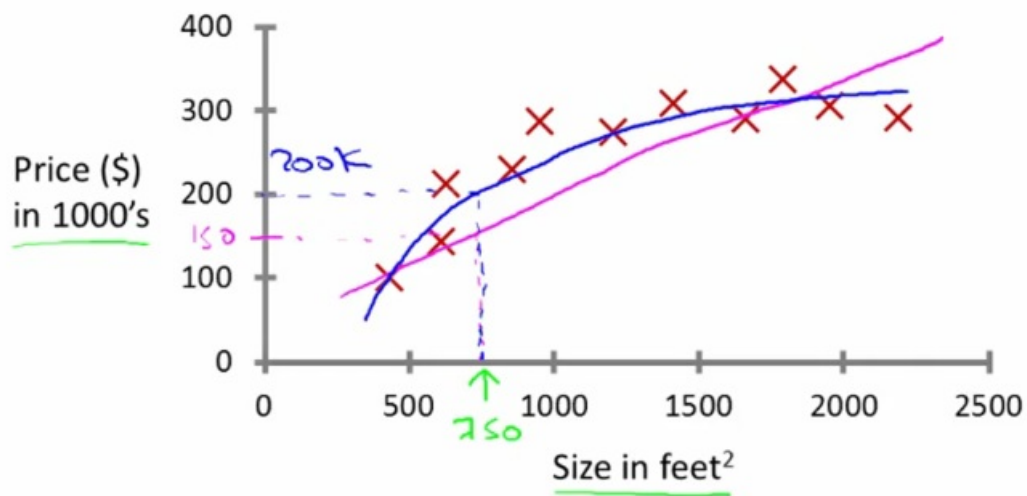


1. 监督学习 (Supervised Learning) : 对于数据集中的每个数据都有正确的答案, 算法就是基于这些作出预测。

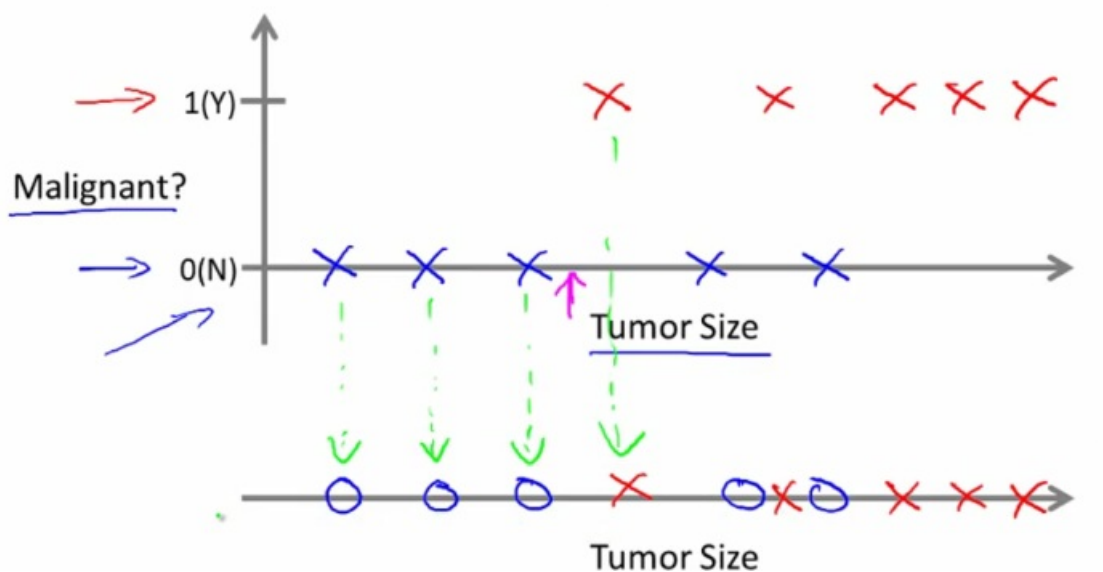
回归 (Regression) 问题: 预测一个连续值的输出。例如: 根据以往数据预测未来房价、根据公司以前销售量预测未来销售量。

## Housing price prediction.



分类 (Classification) 问题: 预测一个离散值的输出。例如: 给定一个数据集, 根据肿瘤大小预测良性还是恶性、判断某人是否被黑过。

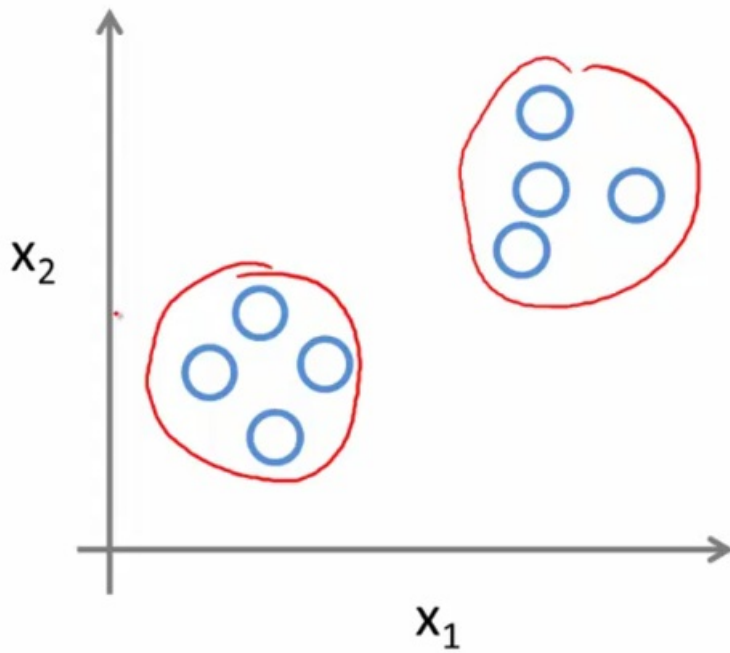
## Breast cancer (malignant, benign)



2. 无监督学习: 数据集中没有属性或标签, 算法自动从数据中寻找规律。

聚类 (Cluster) 算法: 把不同个体归入不同的类。例如: 新闻聚类、市场分割。

# Unsupervised Learning



鸡尾酒会算法：两个不同位置的人对着不同位置的麦克风说话，算法可以将两个人的声音分离出来，或者一个人在一个有音乐的背景下讲话，算法将人声和音乐声分离开来。

## Cocktail party problem

