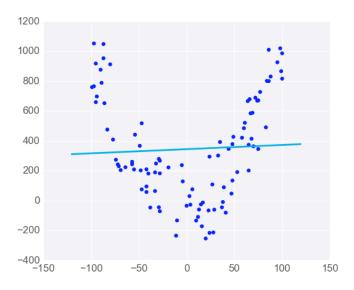
Linear Regression Warnings

Linear regression comes with a set of implicit assumptions and is not the best model for every situation. Here are a couple of issues that you should watch out for.

Linear Regression Works Best When the Data is Linear

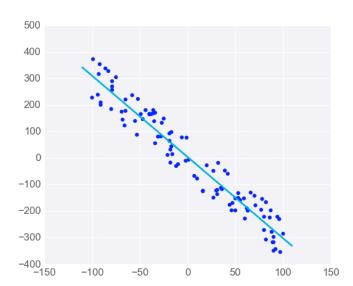
Linear regression produces a straight line model from the training data. If the relationship in the training data is not really linear, you'll need to either make adjustments (transform your training data), add features (we'll come to this next), or use another kind of model.



Linear Regression is Sensitive to Outliers

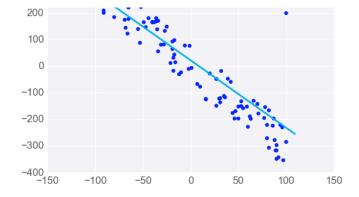
Linear regression tries to find a 'best fit' line among the training data. If your dataset has some outlying extreme values that don't fit a general pattern, they can have a surprisingly large effect.

In this first plot, the model fits the data pretty well.



However, adding a few points that are outliers and don't fit the pattern really changes the way the model predicts.





In most circumstances, you'll want a model that fits most of the data most of the time, so watch out for outliers!

NEXT