

# Discover new regression models

5 minutes

In Unit 2, we looked at fitting a straight line to data points. However, regression can fit many kinds of relationships, including those with multiple factors, and those where the importance of one factor depends on another.

## Experimenting with models

Regression models are often chosen because they work with small data samples, are robust, easy to interpret, and a variety exist.

**Linear regression** is the simplest form of regression, with no limit to the number of features used. Linear regression comes in many forms - often named by the number of features used and the shape of the curve that fits.

**Decision trees** take a step-by-step approach to predicting a variable. If we think of our bicycle example, the decision tree may be first split examples between ones that are during Spring/Summer and Autumn/Winter, make a prediction based on the day of the week. Spring/Summer-Monday may have a bike rental rate of 100 per day, while Autumn/Winter-Monday may have a rental rate of 20 per day.

**Ensemble algorithms** construct not just one decision tree, but a large number of trees - allowing better predictions on more complex data. Ensemble algorithms, such as Random Forest, are widely used in machine learning and science due to their strong prediction abilities.

Data scientists often experiment with using different models. In the following exercise, we'll experiment with different types of models to compare how they perform on the same data.

## Next unit: Exercise - Experiment with more powerful regression models

Continue 