Our wine model had an R-squared value of 0.83, which tells us how accurate our model is on the data we used to construct the model.

So we know our model does a good job predicting the data it's seen.

But we also want a model that does well on new data or data it's never seen before so that we can use the model to make predictions for later years.

Bordeaux wine buyers profit from being able to predict the quality of a wine years before it matures.

They know the values of the independent variables, age and weather, but they don't know the price the wine.

So it's important to build a model that does well at predicting data it's never seen before.

The data that we use to build a model is often called the training data, and the new data is often called the test data.

The accuracy of the model on the test data is often referred to as out-of-sample accuracy.

Let's see how well our model performs on some test data in R.

We have two data points that we did not use to build our model in the file "wine_test.csv".

Let's load this new data file into R. We'll call it wineTest, and we'll use the read.csv function to read in the data file "wine_test.csv".

If we take a look at the structure of wineTest, we can see that we have two observations of the same variables we had before.

These data points are for the years 1979 and 1980.

To make predictions for these two test points, we'll use the function predict.

We'll call our predictions predictTest, and we'll use the predict function.

The first argument to this function is the name of our model.

Here the name of our model is model4.

Then, we say newdata equals name of the data set that we want to make predictions for, in this case wineTest.

If we look at the values in predictTest, we can see that for the first data point we predict 6.768925, and for the

second data point we predict 6.684910.

If we look at our structure output, we can see that the actual Price for the first data point is 6.95, and the actual Price for the second data point is 6.5.

So it looks like our predictions are pretty good.

Let's verify this by computing the R-squared value for our test set.

Recall that the formula for R-squared is: R-squared equals 1 minus the Sum of Squared Errors divided by the Total Sum of Squares.

So let's start by computing the Sum of Squared Errors on our test set.

The Sum of Squared Errors equals the sum of the actual values wineTest dollar sign Price minus our predictions predictTest squared, and then summed.

The Total Sum of Squares equals, the sum again of the actual values wineTest$price, and difference between the mean of the prices on the training set which is our baseline model.

We square these values and add them up.

To compute the R-squared now, we type 1 minus Sum of Squared Errors divided by the Total Sum of Squares.

And we see that the out-of-sample R-squared on this test set is .7944278.

This is a pretty good out-of-sample R-squared.

But while we do well on these two test points, keep in mind that our test set is really small.

We should increase the size of our test set to be more confident about the out-of-sample accuracy of our model.

We can compute the test set R-squared for several different models.

This shows the model R-squared and the test set R-squared for our model as we add more variables.

We saw that the model R-squared will always increases or stays the same as we add more variables.

However, this is not true for the test set.

We want to look for a model with a good model R-squared but also with a good test set R-squared.

In this case we would need more data to be conclusive since two data points in the test set are not really enough to reach any conclusions.

However, it looks like our model that uses Average Growing Season Temperature, Harvest Rain, Age, and Winter Rain does very well in sample on the training set as well as out-of-sample on the test set.

Note here that the test set R-squared can actually be negative.

The model R-squared is never negative since our model can't do worse on the training data than the baseline model.

However, our model can do worse on the test data then the baseline model does.

This leads to a negative R-squared value.

But it looks like our model Average Growing Season Temperature, Harvest Rain, Age, and Winter Rain beats the baseline model.

We'll see in the next video how well Ashenfelter did using this model to make predictions.