# Evaluation metrics for simple linear regression

In this reading, we'll provide a more comprehensive overview about evaluation metrics for simple linear regression. In a prior video we covered $R^2$, and mentioned a few other metrics, MAE and MSE. In this reading, we will review the metrics we've previously mentioned, and introduce a few more as well that you may encounter throughout your career as a data professional.

## Review of $R^2$, MSE, and MAE

The main evaluation metric for linear regression is $R^2$, or the coefficient of determination.

## $R^2$: The coefficient of determination

**$R^2$** measures the proportion of variation in the dependent variable, Y, explained by the independent variable(s), X.

- This is calculated by subtracting the sum of squared residuals divided by the total sum of squares from 1.

$R2 = 1 - \text{Total sum of squares}\text{Sum of squared residuals}$

$R^2$ ranges from 0 to 1. So, if a model has an $R^2$ of 0.85, that

means that the X variables explain about 85% of the variation in the Y variable. Although $R^2$ is a highly interpretable and commonly used metric, you may also encounter mean squared error (MSE) and mean absolute error (MAE) when $R^2$ is insufficient in evaluating model performance.

## MSE: Mean squared error

**MSE (mean squared error)** is the average of the squared difference between the predicted and actual values.

- Because of how MSE is calculated, MSE is very sensitive to large errors.

## MAE: Mean absolute error

**MAE (mean absolute error)** is the average of the absolute difference between the predicted and actual values.

- If your data has outliers that you want to ignore, you can use MAE, as it is not sensitive to large errors.

## Other evaluation metrics

Beyond the three metrics listed above, you may also encounter [AIC (Akaike information criterion) and BIC (Bayesian information criterion)](#).

Lastly, there is **adjusted $R^2$**, which will be addressed in more detail in upcoming videos. It is a variation of $R^2$ that accounts for having multiple independent variables present in a linear regression model.

## Key takeaways

- There are many evaluation metrics to choose from with regard to simple linear regression.
- The most common evaluation metric you'll encounter is probably $R^2$. But, there are times when $R^2$ is insufficient or inappropriate to use.
- Based on your experiences and the particulars of a metric, you can use your best judgment to select an appropriate metric to evaluate a regression model.