



Aplicação de Medidas de Dispersão em Data Science

# Aula	11
<input checked="" type="checkbox"/> Preparada	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/> Revisada	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/> Lecionada	<input checked="" type="checkbox"/>

▼ Considere os erros de um modelo de regressão...

Vamos defini-los como sendo:

$$e = y - \hat{y}$$

Podemos repetir o todo racional que fizemos dos desvios da médias.

Então temos:

▼ Aplicação Prática em modelos de regressão:

▼ **MAE** (*Mean Absolute Error* [Erro Absoluto Médio])

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

▼ **MSE** (*Mean Squared Error* [Erro Quadrático Médio])

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

▼ **RMSE** (*Root Mean Squared Error* [Raiz do Erro Quadrático Médio])

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

▼ Então... Quais destes indicadores vamos usar?

O **MSE** é expresso na unidade ao quadrado.

Já o **MAE** e o **RMSE** estão na mesma unidade que o dado original, então usamos os dois!

▼ Faz sentido?

A comparação entre o **MAE** e o **RMSE** vai nos mostrar quantos valores extremos temos nos erros do modelo em que estamos medindo estes indicadores.