

# ScriptRMSE

Cícero

2022-10-01

## Script

```
setwd('C:\\Users\\cicer\\Rstudio')
```

### Setando de onde vou ler os dados

```
dados_de_teste<- read.csv("test.csv")
```

### Lendo os dados do test

```
modeloConjunto1 <- lm(log(SalePrice) ~ LotArea + TotalBsmtSF + Neighborhood + ExterQual,dados )
```

### Criando o modelo que vamos testar sua eficácia

```
summary(modeloConjunto1)
```

### Vendo as informações desse modelo

```
previsões <- predict(modeloConjunto, dados_de_teste)
```

### Pedindo para que o modelo teste os dados

```
summary(previsões)
```

### Vendo as informações da previsao

```
target <- read.csv('Sample_submission.csv')
```

### Lendo os dados alvo para nosso modelo

```
logSalePriceTeste <- log(target$SalePrice)
```

### Fazendo o log da coluna de SalePrice e atribuindo a variavel criada

```
logSalePriceTeste[1:5] Analizando as 5 primeiras colunas
```

```
previsões[1:5]
```

### Olhando as 5 primeiras colunas da nossa previsao para comparar com o gabarito

```
erro <- previsões - logSalePriceTeste
```

### encontrando o erro

```
erro[1:5]
```

### Visualizando o erro para as 5 colunas

```
erro_quadrado <- erro^2
```

### Colocando ao quadrado para deixar tudo positivo

```
erro_quadrado_medio <- mean(erro_quadrado, na.rm = TRUE)
```

### Tirando a media desses erros e removando os valores de NA das colunas

```
RMSE <- sqrt(erro_quadrado_medio)
```

**Encontrando o RMSE ao tirar a raiz do erro quadrado médio**

RMSE

**Analisa o quão o RMSE se aproxima do desvio padrão residual do nosso modelo**