

```
1  #%%
2  import pandas as pd
3  #%%
4  df = pd.read_csv('melb_data.csv')
5  #%%
6  df.describe()
7  #%%
8  df.columns
9  #%%
10 df.info()
11 #%%
12 # Removendo valores nulos
13 df.dropna(axis=0, inplace=True)
14
15 #%%
16 # Selecionando as colunas que serão usadas para
   treinar o modelo
17 features = ['Rooms', 'Bathroom', 'Landsize', '
   Lattitude', 'Longitude']
18
19 # Definindo as variáveis preditoras e a variável
   alvo
20 y = df['Price']
21 X = df[features]
22 #%%
23 X.describe()
24 #%%
25 X.head()
26 #%%
27 # Importando o train_test_split
28 from sklearn.model_selection import
   train_test_split
29
30 # Separando os dados em treino e validação
31 train_X, val_X, train_y, val_y = train_test_split
   (X, y, random_state = 1)
32
33 # Importando o modelo arvore de decisão
34 from sklearn.tree import DecisionTreeRegressor
35
```

```
36 # Definindo o modelo
37 model = DecisionTreeRegressor(random_state=1)
38
39 # Treinando o modelo
40 model.fit(train_X, train_y)
41
42 # Fazendo as previsões
43 predictions = model.predict(val_X)
44 #%%
45 # Calculando o erro absoluto médio
46 from sklearn.metrics import mean_absolute_error
47
48 print(mean_absolute_error(val_y, predictions))
49
```