

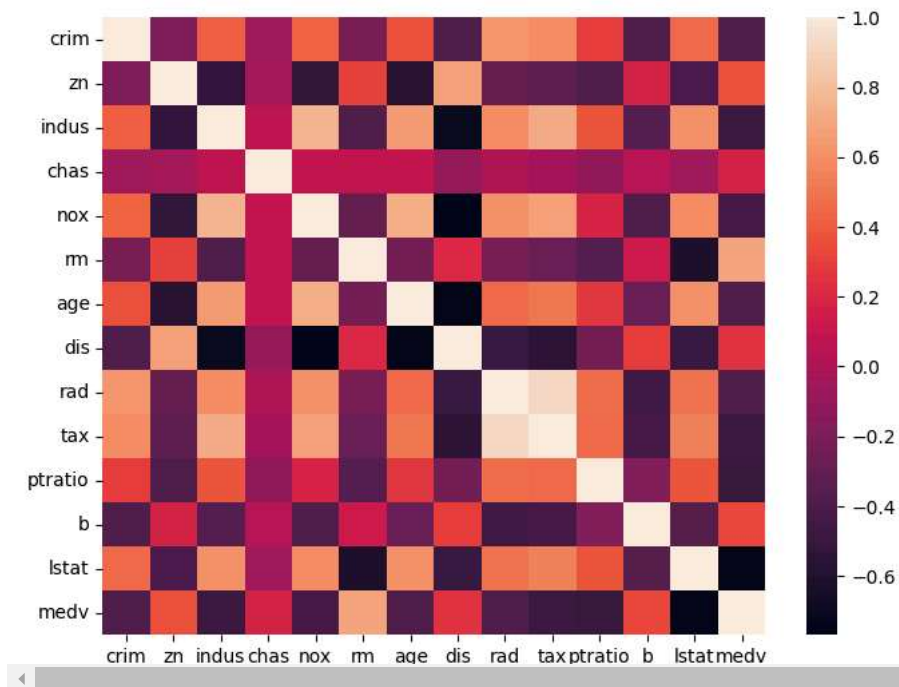
```
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
```

```
df = pd.read_csv("BostonHousing.csv")
df.isnull().sum()
```

```
crim      0
zn        0
indus     0
chas      0
nox       0
rm        0
age       0
dis       0
rad       0
tax       0
ptratio   0
b         0
lstat     0
medv      0
dtype: int64
```

```
plt.figure(figsize=(8,6))
sns.heatmap(df.corr())
plt.show
```

```
<function matplotlib.pyplot.show(close=None, block=None)>
```



```
y = df["medv"]
x = df[["rm", "b", "zn"]]
```

```
X_train, X_test, y_train, y_test = train_test_split(x, y, test_size=0.2, random_state=42)
model = LinearRegression()
model.fit(X_train, y_train)
y_pred = model.predict(X_test)
comparison = pd.DataFrame({'Actual': y_test, 'Predicted': y_pred})
print(comparison)
```

```
Actual Predicted
173    23.6  24.274655
274    32.4  28.626080
491    13.6  20.381880
72     22.8  21.113451
452    16.1  22.977111
```

```
...
412  17.9  -0.584340
436   9.6  15.279719
411  17.2  17.172385
86   22.5  20.809365
75   21.4  22.834276
```

```
[102 rows x 2 columns]
```

Start coding or [generate](#) with AI.