

Boston Housing Linear Regression Analysis

Boston Housing Dataset Linear Regression Script

Import libraries

```
import pandas as pd
```

```
import numpy as np
```

```
import matplotlib.pyplot as plt
```

```
import seaborn as sns
```

```
from sklearn.datasets import load_boston
```

```
from sklearn.model_selection import train_test_split
```

```
from sklearn.linear_model import LinearRegression
```

```
from sklearn.metrics import mean_absolute_error, mean_squared_error, r2_score
```

Load dataset

```
boston = load_boston()
```

```
df = pd.DataFrame(boston.data, columns=boston.feature_names)
```

```
df['PRICE'] = boston.target
```

Display first 5 rows

```
print(df.head())
```

Split into features and target

Boston Housing Linear Regression Analysis

```
X = df.drop('PRICE', axis=1)
```

```
y = df['PRICE']
```

Train-test split

```
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
```

Create and train model

```
lr = LinearRegression()
```

```
lr.fit(X_train, y_train)
```

Predictions

```
y_pred = lr.predict(X_test)
```

Evaluate model

```
mae = mean_absolute_error(y_test, y_pred)
```

```
mse = mean_squared_error(y_test, y_pred)
```

```
r2 = r2_score(y_test, y_pred)
```

```
print(f'MAE: {mae:.2f}')
```

```
print(f'MSE: {mse:.2f}')
```

Boston Housing Linear Regression Analysis

```
print(f'R2 Score: {r2:.2f}')
```

Coefficients interpretation

```
coeff_df = pd.DataFrame(lr.coef_, X.columns, columns=['Coefficient'])
```

```
print(coeff_df)
```

Check multicollinearity using heatmap

```
plt.figure(figsize=(10,8))
```

```
sns.heatmap(df.corr(), annot=True, cmap='coolwarm')
```

```
plt.title('Correlation Heatmap')
```

```
plt.show()
```