

Q5

```
In [ ]: from sklearn.datasets import fetch_california_housing
        from sklearn.linear_model import LinearRegression
        from sklearn.model_selection import train_test_split
        from sklearn.metrics import r2_score, mean_squared_error
        import numpy as np
        import pandas as pd
```

```
In [ ]: obj = fetch_california_housing()
        data = obj["data"]
```

```
In [ ]: for i in range(8):
        for j in range(i+1, 8):
            print(f"Correlations of column {i} with column {j} ")
            print(np.corrcoef(data[:,i], data[:,j])[0,1])
```

Correlations of column 0 with column 1
-0.11903398993785665
Correlations of column 0 with column 2
0.3268954316412956
Correlations of column 0 with column 3
-0.062040133836099076
Correlations of column 0 with column 4
0.004834345627652915
Correlations of column 0 with column 5
0.01876624796696885
Correlations of column 0 with column 6
-0.0798091274597188
Correlations of column 0 with column 7
-0.015175865414173956
Correlations of column 1 with column 2
-0.15327742256198937
Correlations of column 1 with column 3
-0.07774728275376118
Correlations of column 1 with column 4
-0.2962442397735358
Correlations of column 1 with column 5
0.01319135663602974
Correlations of column 1 with column 6
0.011172673530605408
Correlations of column 1 with column 7
-0.10819681311244811
Correlations of column 2 with column 3
0.8476213257130447
Correlations of column 2 with column 4
-0.07221284865893354
Correlations of column 2 with column 5
-0.004852294991781336
Correlations of column 2 with column 6
0.1063889654862552
Correlations of column 2 with column 7
-0.027540053873544787
Correlations of column 3 with column 4
-0.06619740232676065
Correlations of column 3 with column 5
-0.006181201268673116
Correlations of column 3 with column 6
0.0697211298887421
Correlations of column 3 with column 7
0.0133443896399991
Correlations of column 4 with column 5
0.06986273036567671
Correlations of column 4 with column 6
-0.1087847473776677
Correlations of column 4 with column 7
0.09977322287464561
Correlations of column 5 with column 6
0.0023661822637503493
Correlations of column 5 with column 7
0.0024758163767050613
Correlations of column 6 with column 7
-0.9246644339150403

```
In [ ]: X = data
        y = obj.target
```

```
Out[ ]: 20640
```

```
In [ ]: X_train, X_test, y_train, y_test = train_test_split(X , y, random_state=100, test_si
        reg = LinearRegression()
        reg.fit(X_train, y_train)
```

```
Out[ ]: LinearRegression ⓘ ⓘ
        LinearRegression()
```

```
In [ ]: preds = reg.predict(X_test)
        print(f"MSE : {mean_squared_error(y_test, preds)}")
        print(f"R2 score : {r2_score(y_test, preds)}")
```

```
MSE : 0.5088933351158983
R2 score : 0.6223138107295262
```

```
In [ ]: print(reg.coef_, reg.intercept_)

[ 4.33432793e-01  9.22564691e-03 -1.06547768e-01  6.46494007e-01
 -7.07960568e-06 -3.45850134e-03 -4.23282369e-01 -4.37465774e-01] -37.20562128878796
```