```
# implementation of multiple regression on given dataset
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error, mean_absolute_error
from sklearn import preprocessing
df = pd.read_csv('D:\Real-estate1.csv')
df.drop('No', inplace = True,axis=1)


print(df.head())
print(df.columns)
sns.scatterplot(x='X4 number of convenience stores',
                y='Y house price of unit area', data=df)
X = df.drop('Y house price of unit area',axis= 1)
y = df['Y house price of unit area']
print(X)
print(y)
X_train, X_test, y_train, y_test = train_test_split(
    X, y, test_size=0.3, random_state=101)
model = LinearRegression()
model.fit(X_train,y_train)
predictions = model.predict(X_test)
```

```
print(  'mean_squared_error : ', mean_squared_error(y_test, predictions))

print(  'mean_absolute_error : ', mean_absolute_error(y_test, predictions))
```

Output

```
  X1 transaction date  X2 house age  ...  X6 longitude  Y house price of unit area

0           2012.917        32.0  ...     121.54024                  37.9

1           2012.917        19.5  ...     121.53951                  42.2

2           2013.583        13.3  ...     121.54391                  47.3

3           2013.500        13.3  ...     121.54391                  54.8

4           2012.833         5.0  ...     121.54245                  43.1


[5 rows x 7 columns]

Index(['X1 transaction date', 'X2 house age',

       'X3 distance to the nearest MRT station',

       'X4 number of convenience stores', 'X5 latitude', 'X6 longitude',

       'Y house price of unit area'],

      dtype='object')

     X1 transaction date  X2 house age  ...  X5 latitude  X6 longitude

0            2012.917        32.0  ...     24.98298     121.54024

1            2012.917        19.5  ...     24.98034     121.53951

2            2013.583        13.3  ...     24.98746     121.54391

3            2013.500        13.3  ...     24.98746     121.54391

4            2012.833         5.0  ...     24.97937     121.54245

..                ...         ...  ...          ...          ...

409          2013.000        13.7  ...     24.94155     121.50381
```

410 2012.667 5.6 ... 24.97433 121.54310

411 2013.250 18.8 ... 24.97923 121.53986

412 2013.000 8.1 ... 24.96674 121.54067

413 2013.500 6.5 ... 24.97433 121.54310

[414 rows x 6 columns]

0 37.9

1 42.2

2 47.3

3 54.8

4 43.1

409 15.4

410 50.0

411 40.6

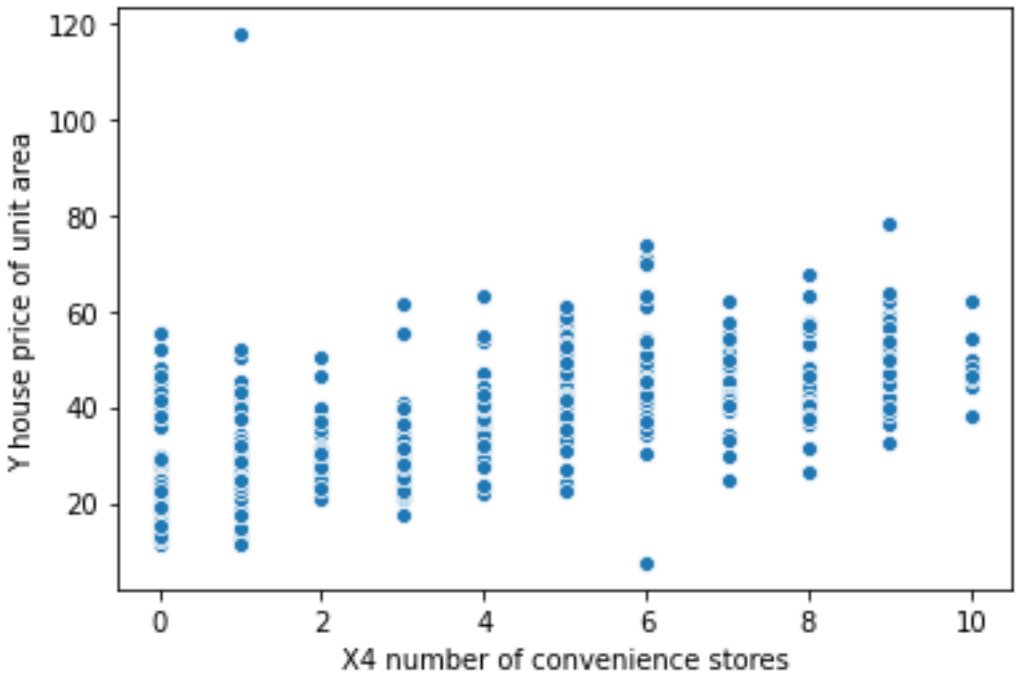
412 52.5

413 63.9

Name: Y house price of unit area, Length: 414, dtype: float64

mean_squared_error : 46.21179783493681

mean_absolute_error : 5.392293684756571


Y house price of unit area
120
100
80
60
40
20
0
2
4
6
8
10
X4 number of convenience stores