

Name:Hamza Patel Roll:2305215 simple lr

```
In [1]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt

from sklearn.datasets import fetch_california_housing
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error, r2_score
```

Load the California Housing Dataset

```
In [2]: housing = fetch_california_housing(as_frame=True)

df = housing.frame
print(df.head())
```

	MedInc	HouseAge	AveRooms	AveBedrms	Population	AveOccup	Latitude	\
0	8.3252	41.0	6.984127	1.023810	322.0	2.555556	37.88	
1	8.3014	21.0	6.238137	0.971880	2401.0	2.109842	37.86	
2	7.2574	52.0	8.288136	1.073446	496.0	2.802260	37.85	
3	5.6431	52.0	5.817352	1.073059	558.0	2.547945	37.85	
4	3.8462	52.0	6.281853	1.081081	565.0	2.181467	37.85	

	Longitude	MedHouseVal
0	-122.23	4.526
1	-122.22	3.585
2	-122.24	3.521
3	-122.25	3.413
4	-122.25	3.422

Select Feature and Target

```
In [3]: X = df[['MedInc']]
y = df[['MedHouseVal']]
```

Split the Dataset

```
In [5]: X_train, X_test, y_train, y_test = train_test_split(X,y,test_size=0.2, random_state
'''print(X_train)
print(X_test)
print(y_train)
print(y_test)''')
```

```
Out[5]: 'print(X_train)\nprint(X_test)\nprint(y_train)\nprint(y_test)'
```

Create and Train the Linear Regression Model

```
In [6]: model = LinearRegression()
model.fit(X_train, y_train)
y_pred = model.predict(X_test)
```

```
print("Mean sq. error:", mean_squared_error(y_test, y_pred))  
print("R2 score:", r2_score(y_test, y_pred))
```

Mean sq. error: 0.7091157771765548
R2 score: 0.45885918903846656

```
In [11]: plt.figure()  
plt.scatter(X_test, y_test, color='lightblue')  
plt.plot(X_test, y_pred, color='purple')  
plt.xlabel("Median Income")  
plt.ylabel("Median House Value")  
plt.title("Simple Linear regression: Housing Price Prediction")
```

Out[11]: Text(0.5, 1.0, 'Simple Linear regression: Housing Price Prediction')



In []: