Theory:

Multiple linear regression (MLR), also known simply as multiple regression, is a statistical technique that uses several explanatory variables to predict the outcome of a response variable. The goal of multiple linear regression (MLR) is to model the linear relationship between the explanatory (independent) variables and response (dependent) variable.

Import all the necessary machine learning libraries

In [1]:

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import scipy as sp
```

Import and read the csv dataset using panda : "pd.read_csv()"

In [18]:

```
dataset=pd.read_csv('~/Desktop/homeprices-multi.csv')
dataset
```

Out[18]:

| | area | bedrooms | age | price |
|---|---|---|---|---|
| 0 | 2600 | 3.0 | 20 | 550000 |
| 1 | 3000 | 4.0 | 15 | 565000 |
| 2 | 3200 | NaN | 18 | 610000 |
| 3 | 3600 | 3.0 | 30 | 595000 |
| 4 | 4000 | 5.0 | 8 | 760000 |
| 5 | 4100 | 6.0 | 8 | 810000 |

In [19]:

```
X= dataset.iloc[:,:-1].values
Y= dataset.iloc[:,3].values
```

In [20]:

```
from sklearn.impute import SimpleImputer
sim=SimpleImputer(missing_values=np.nan,strategy='mean')
sim.fit(X)
```

Out[20]:

```
SimpleImputer(add_indicator=False, copy=True, fill_value=None,
              missing_values=nan, strategy='mean', verbose=0)
```

In [21]:

```
X = sim.transform(X)
```

In [22]:

```
from sklearn.model_selection import train_test_split
X_train,X_test,Y_train,Y_test=train_test_split(X,Y,test_size=0.35,random_state=0)
```

In [23]:

```
from sklearn.linear_model import LinearRegression
regressor = LinearRegression()
regressor.fit(X_train,Y_train)
```

Out[23]:

```
LinearRegression(copy_X=True, fit_intercept=True, n_jobs=None, normalize=F
alse)
```

In [24]:

```
pred_Y = regressor.predict(X_test)
pred_Y
```

Out[24]:

```
array([770553.01261587, 622483.57465672, 619017.67087196])
```

In [25]:

```
Y_test
```

Out[25]:

```
array([810000, 610000, 565000], dtype=int64)
```

In [26]:

```
X
```

Out[26]:

```
array([[2.6e+03, 3.0e+00, 2.0e+01],
       [3.0e+03, 4.0e+00, 1.5e+01],
       [3.2e+03, 4.2e+00, 1.8e+01],
       [3.6e+03, 3.0e+00, 3.0e+01],
       [4.0e+03, 5.0e+00, 8.0e+00],
       [4.1e+03, 6.0e+00, 8.0e+00]])
```

In [27]:

```
regressor.predict([[5000,4,17]])
```

Out[27]:

```
array([810188.15194815])
```

In [28]:

```
regressor.predict([[4500,5,13]])
```

Out[28]:

```
array([782500.])
```

In [29]:

```
regressor.predict([[3000,5,10]])
```

Out[29]:

```
array([647553.01261587])
```

Conclusion : In this experiment we learnt how to build a multiple linear regression model using machine learning libraries such as sklearn. Also we used the fit and predict functions to fit our data in the model and to predict the output results of our model.We also used the split data function to split our data into training and testing set. Also we used the matplotlib functions to plot the linear regression graph.