Experiment 4:

Write a Python program to implement Simple Linear Regression.

```
//Implementation of Linear Regression
import pandas as pd
import numpy as np
from sklearn import linear_model
df=pd.read_csv("C:/Users/mitha/Desktop/ML/hyderabad.csv")
print("First 5 rows od dataset:\n",df.head(),"\n")
print ("Available columns:",df.columns.tolist(),"\n")
reg=linear_model.LinearRegression()
reg.fit(df[['Area']],df['Price'])
print("coefficient(slope):",reg.coef_[0])
print("Intercept:",reg.intercept_)
print("Prediction for 3000 sq ft:",reg.predict([[3000]])[0])
print("Prediction for 10000 sq ft:",reg.predict(np.array([[10000]]))[0])
print("Final Coefficient (slope):",reg.coef_[0])
print("Final Intercpet:",reg.intercept_)
```

Output:

```
First 5 rows od dataset:
       Price  Area     Location  No. of Bedrooms  Resale  MaintenanceStaff  \
0   6968000  1340     Nizampet                2       0                 0
1  29000000  3498  Hitech City                4       0                 0
2   6590000  1318    Manikonda                2       0                 0
3   5739000  1295        Alwal                3       1                 0
4   5679000  1145   Kukatpally                2       0                 0

   Gymnasium  SwimmingPool  LandscapedGardens  JoggingTrack  ...  \
0          1             1                  1             1  ...
1          1             1                  1             1  ...
2          1             0                  0             0  ...
3          0             0                  0             0  ...
4          0             0                  1             0  ...

   LiftAvailable  BED  VaastuCompliant  Microwave  GolfCourse  TV  \
0              1    0                1          0           0   0
1              1    0                1          0           0   0
2              0    0                0          0           0   0
3              1    0                0          0           0   0
4              1    0                0          0           0   0

   DiningTable  Sofa  Wardrobe  Stadium
0            0     0         0        0
1            0     0         0        0
2            0     0         0        0
3            0     0         0        0
4            0     0         0        0
```

```
[5 rows x 40 columns]

Available columns: ['Price', 'Area', 'Location', 'No. of Bedrooms', 'Resale', 'MaintenanceStaff', 'Gymnasium', 'Swimm
ingPool', 'LandscapedGardens', 'JoggingTrack', 'RainWaterHarvesting', 'IndoorGames', 'ShoppingMall', 'Intercom', 'Spo
rtsFacility', 'ATM', 'ClubHouse', 'School', '24X7Security', 'PowerBackup', 'CarParking', 'StaffQuarter', 'Cafeteria',
'MultipurposeRoom', 'Hospital', 'WashingMachine', 'Gasconnection', 'AC', 'Wifi', "Children'splayarea", 'LiftAvailabl
e', 'BED', 'VaastuCompliant', 'Microwave', 'GolfCourse', 'TV', 'DiningTable', 'Sofa', 'Wardrobe', 'Stadium']

coefficient(slope): 9753.940608184039
Intercept: -6222669.083283698
Prediction for 3000 sq ft: 23039152.74126842
Prediction for 10000 sq ft: 91316736.9985567
Final Coefficient (slope): 9753.940608184039
Final Intercpet: -6222669.083283698
```

//Implementation of linear regresssion

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
dataset=pd.read_csv('C:/Users/mitha/Desktop/ML/Salary_Data.csv')
dataset.head()
```

Output:

	YearsExperience	Salary
0	1.1	39343.0
1	1.3	46205.0
2	1.5	37731.0
3	2.0	43525.0
4	2.2	39891.0

```
dataset.tail()
```

	YearsExperience	Salary
25	9.0	105582.0
26	9.5	116969.0
27	9.6	112635.0
28	10.3	122391.0
29	10.5	121872.0

X=dataset.iloc[:,:-1].values
y=dataset.iloc[:,-1].values
print(X)
print(y)

```
[[ 1.1]
 [ 1.3]
 [ 1.5]
 [ 2. ]
 [ 2.2]
 [ 2.9]
 [ 3. ]
 [ 3.2]
 [ 3.2]
 [ 3.7]
 [ 3.9]
 [ 4. ]
 [ 4. ]
 [ 4.1]
 [ 4.5]
 [ 4.9]
 [ 5.1]
 [ 5.3]
 [ 5.9]
 [ 6. ]
 [ 6.8]
 [ 7.1]
 [ 7.9]
 [ 8.2]
 [ 8.7]
 [ 9. ]
 [ 9.5]
 [ 9.6]
```

```
 [10.3]
 [10.5]]
[ 39343.  46205.  37731.  43525.  39891.  56642.  60150.  54445.  64445.
  57189.  63218.  55794.  56957.  57081.  61111.  67938.  66029.  83088.
  81363.  93940.  91738.  98273. 101302. 113812. 109431. 105582. 116969.
 112635. 122391. 121872.]
```

```
from sklearn.model_selection import train_test_split
X_train,X_test,y_train,y_test=train_test_split(X,y,test_size=1/3,random_state=0)
X_train
```

```
array([[ 2.9],
       [ 5.1],
       [ 3.2],
       [ 4.5],
       [ 8.2],
       [ 6.8],
       [ 1.3],
       [10.5],
       [ 3. ],
       [ 2.2],
       [ 5.9],
       [ 6. ],
       [ 3.7],
       [ 3.2],
       [ 9. ],
       [ 2. ],
       [ 1.1],
       [ 7.1],
       [ 4.9],
       [ 4. ]])
```

```
from sklearn.linear_model import LinearRegression
regressor=LinearRegression()
regressor.fit(X_train,y_train)
y_pred=regressor.predict(X_test)
y_pred
```

```
array([ 40835.10590871, 123079.39940819,  65134.55626083,  63265.36777221,
       115602.64545369, 108125.8914992 , 116537.23969801,  64199.96201652,
        76349.68719258, 100649.1375447 ])
```

```
plt.scatter(X_train,y_train,color='hotpink')
plt.plot(X_train,regressor.predict(X_train),color='mediumpurple')
plt.title("Salary vs Experience(Training Set)")
plt.xlabel("Years of Experience")
plt.ylabel("Salary")
plt.show()
```

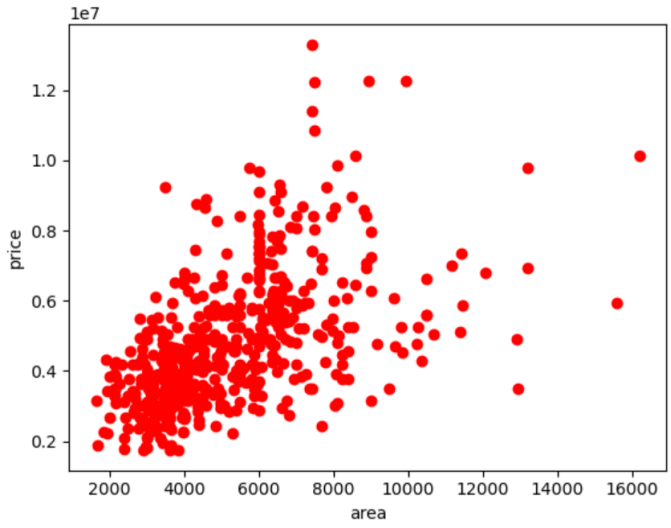

```
plt.scatter(X_test,y_test,color='hotpink')
plt.plot(X_train,regressor.predict(X_train),color='mediumpurple')
plt.title("Salary vs Experience(Testing Set)")
plt.xlabel("Years of Experience")
plt.ylabel("Salary")
plt.show()
```


```
from sklearn import metrics
print('Mean Absolute Error:',metrics.mean_absolute_error(y_test,y_pred))
print('Mean Squared Error:',metrics.mean_squared_error(y_test,y_pred))
print('Root Mean Squared
Error:',np.sqrt(metrics.mean_squared_error(y_test,y_pred)))
```

```
Mean Absolute Error: 3426.4269374307123
Mean Squared Error: 21026037.329511296
Root Mean Squared Error: 4585.4157204675885
```

2.

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn import linear_model
df = pd.read_csv('C:/Users/mitha/Desktop/ML/Housing.csv')
plt.xlabel("area")
plt.ylabel("price")
plt.scatter(df.area,df.price,color='r',marker='o')
```




```
reg=linear_model.LinearRegression()
reg.fit(df[['area']],df.price)
reg.coef_
```

```
array([461.97489427])
```

```
reg.intercept_
```

```
2387308.4823964303
```

```
p=reg.predict(np.array([[5000]]))
p
```

```
array([4697182.95376035])
```

```
d=pd.read_csv("C:/Users/mitha/Desktop/ML/House Price India.csv")
```

```
df.head(2)
```

	price	area	bedrooms	bathrooms	stories	mainroad	guestroom	basement	hotwaterheating	airconditioning	parking
0	13300000	7420	4	2	3	yes	no	no	no	yes	2
1	12250000	8960	4	4	4	yes	no	no	no	yes	3