Superwise learning lab -2

NAME:- Ayub Alam

Rollno:-TNU2021053100031L

Assigment Date:-

Submission date:-

Gradient desent

```python
In [2]: import pandas as pd
        import numpy as np
        import matplotlib.pyplot as plt
```

```python
In [3]: data=pd.read_csv(r'linear_regression.csv')
        data
```

Out[3]:

	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
5	2.9	56642.0
6	3.0	60150.0
7	3.2	54445.0
8	3.2	64445.0
9	3.7	57189.0
10	3.9	63218.0
11	4.0	55794.0
12	4.0	56957.0
13	4.1	57081.0

	YearsExperience	Salary
14	4.5	61111.0
15	4.9	67938.0
16	5.1	66029.0
17	5.3	83088.0
18	5.9	81363.0
19	6.0	93940.0
20	6.8	91738.0
21	7.1	98273.0
22	7.9	101302.0
23	8.2	113812.0
24	8.7	109431.0
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

In [4]:
```python
x=data.iloc[:,0]
y=data.iloc[:,1]
```

In [5]:
```python
plt.scatter(x, y)
plt.plot(x,y,color='red')
plt.xlabel('x')
plt.ylabel('y')
plt.title('Regression plot')
plt.show()
```

Regression plot
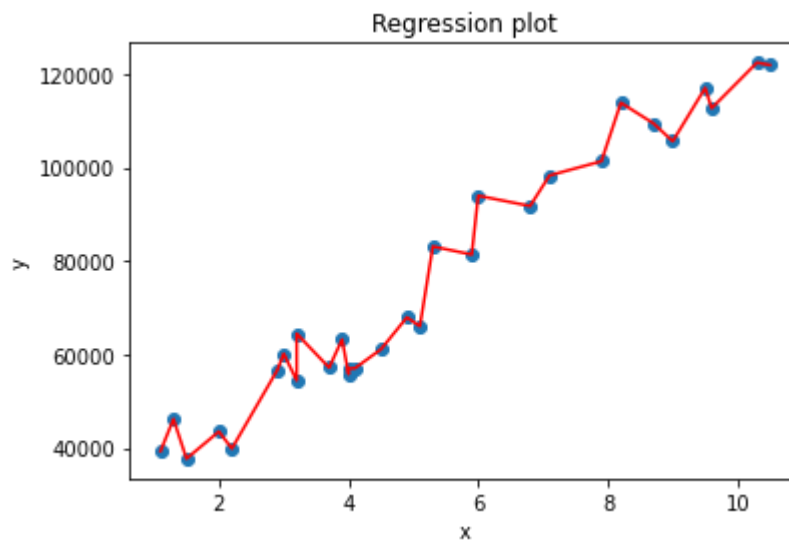
```
m=0
c=0
L=0.0001
e=1000
n=float(len(x))
for i in range (e):
    y_pred=m*x+c
    d_m=(-2/n)*sum(x*(y-y_pred))
    d_c=(-2/n)*sum(y-y_pred)
    m=m - L*d_m
    c=c - L*d_c
print(m, c)
```

12836.600965885045 2915.2044856014018

```
y_=m*x+c
y_
```

```
0       17035.465548
1       19602.785741
2       22170.105934
3       28588.406417
4       31155.726611
5       40141.347287
6       41425.007383
7       43992.327576
8       43992.327576
9       50410.628059
10      52977.948253
11      54261.608349
12      54261.608349
13      55545.268446
14      60679.908832
15      65814.549218
16      68381.869412
17      70949.189605
18      78651.150184
19      79934.810281
20      90204.091054
21      94055.071343
22     104324.352116
23     108175.332406
```

```
24      114593.632889
25      118444.613179
26      124862.913662
27      126146.573758
28      135132.194434
29      137699.514627
Name: YearsExperience, dtype: float64
```

In [8]:
```python
plt.scatter(x,y,color='Blue')
plt.plot(x,y_,color='red')
plt.xlabel('x')
plt.ylabel('y')
plt.title('Regression plot')
plt.show()
```
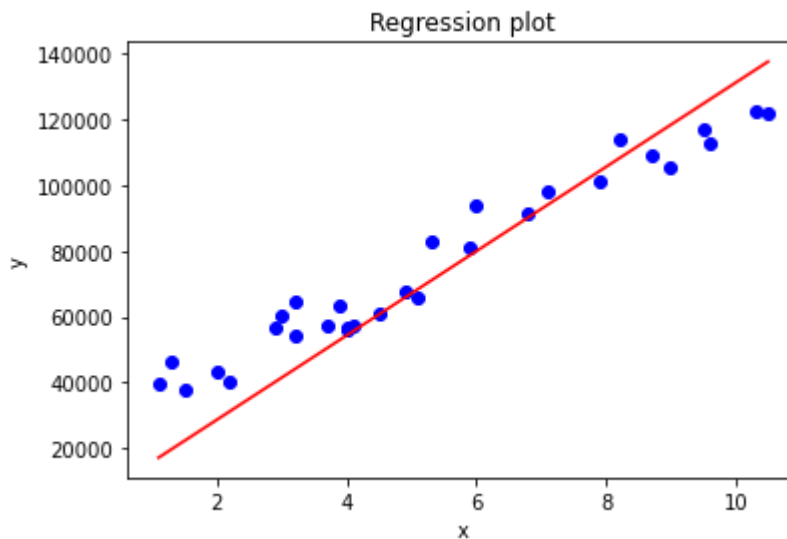

In [9]:
```python
from sklearn.model_selection import train_test_split
```

In [10]:
```python
x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=1/3,random_state=0)
```

In [11]:
```python
x_train
```

Out[11]:
```
5       2.9
16      5.1
8       3.2
14      4.5
23      8.2
20      6.8
1       1.3
29      10.5
6       3.0
4       2.2
18      5.9
19      6.0
9       3.7
7       3.2
25      9.0
3       2.0
0       1.1
21      7.1
```

```
15    4.9
12    4.0
Name: YearsExperience, dtype: float64
```

In [12]:
```
y_train
```

Out[12]:
```
5        56642.0
16       66029.0
8        64445.0
14       61111.0
23      113812.0
20       91738.0
1        46205.0
29      121872.0
6        60150.0
4        39891.0
18       81363.0
19       93940.0
9        57189.0
7        54445.0
25      105582.0
3        43525.0
0        39343.0
21       98273.0
15       67938.0
12       56957.0
Name: Salary, dtype: float64
```

In [13]:
```
l=0.0001
m=0
c=0

epochs=1000
n=float(len(x))
n
```

Out[13]: 30.0

In [14]:
```
for i in range(epochs):
    y_pred=m*x_train+c
    d_m=(-2/n)*sum(x_train*(y_train-y_pred))
    d_c=(-2/n)*sum(y_train-y_pred)
    m=m - l*d_m
    c=c - l*d_c
print(m,c)
```

```
13044.458512709201 2879.8182712986677
```

In [15]:
```
y_pred=m*x_train+c
```

In [16]:
```
y_pred
```

Out[16]:
```
5        40708.747958
16       69406.556686
8        44622.085512
14       61579.881578
23      109844.378076
```

```
20      91582.136158
1       19837.614338
29     139846.632655
6       42013.193809
4       31577.626999
18      79842.123496
19      81146.569348
9       51144.314768
7       44622.085512
25     120279.944886
3       28968.735297
0       17228.722635
21      95495.473712
15      66797.664984
12      55057.652322
Name: YearsExperience, dtype: float64
```

In [21]:
```python
plt.scatter(x,y,color='Blue')
plt.plot(x_train,y_pred,color='red')
plt.xlabel('x')
plt.ylabel('y')
plt.title('Regression plot')
plt.show()
```