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

In [2]:
```
df = pd.read_csv('/Users/yash/Downloads/Experience-Salary.csv')
```

In [3]:
```
df
```

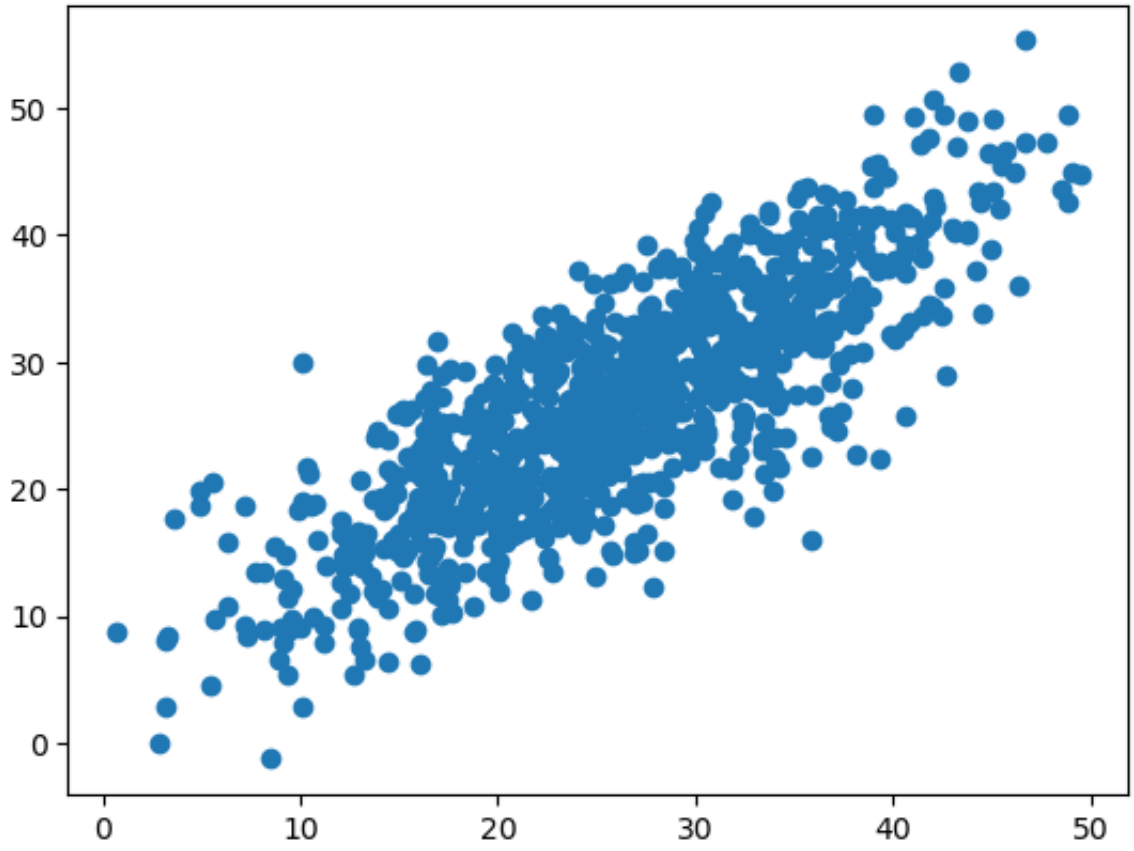
Out[3]:

	exp(in months)	salary(in thousands)
0	18.290293	16.521825
1	17.023407	11.666234
2	26.343613	23.167255
3	19.105834	20.877145
4	27.742516	23.166236
...	...	...
995	22.161741	24.666502
996	32.266497	24.160270
997	17.039030	19.903797
998	25.222124	21.974413
999	16.418028	19.634787

1000 rows × 2 columns

In [4]:
```
plt.scatter(df['exp(in months)'],df['salary(in thousands)'])
```

Out[4]:
```
<matplotlib.collections.PathCollection at 0x13d2c9c50>
```



In [5]:
```
df.shape
```

Out[5]:
```
(1000, 2)
```

In [86]:
```
x = np.array(df['exp(in months)']).reshape((1000,1))
y = np.array(df['salary(in thousands)']).reshape((1000,1))
print("ok")
```

```
ok
```

In [87]:
```
x.shape,y.shape
```

Out[87]:
```
((1000, 1), (1000, 1))
```

In [88]:
```
# adding 1 to each vector xi
one_arr = np.ones((1000,1),order='C')
one_arr.shape
```

Out[88]:
```
(1000, 1)
```

In [91]:
```
x_final =np.concatenate((one_arr,x),axis=1)
```

In [92]:
```
x_final.shape
```

Out[92]:
```
(1000, 2)
```

In [93]:
```
x_final
```

Out[93]:
```
array([[ 1.        , 18.29029332],
       [ 1.        , 17.02340669],
       [ 1.        , 26.34361348],
       ...,
       [ 1.        , 17.03902988],
       [ 1.        , 25.22212418],
       [ 1.        , 16.41802763]])
```

In [94]:
```
x_trans = np.transpose(x_final)
```

In [98]:
```
beta= np.linalg.inv(x_final.T.dot(x_final)).dot(x_final.T).dot(y)
```

In [99]:
```
beta # parameters b0, b1
```

Out[99]:
```
array([[5.19871856],
       [0.82284666]])
```

In [111]:
```
#testing
# x_test = np.array([[1,18.290293],[1,17.023407]],dtype='float')
# x_test
```

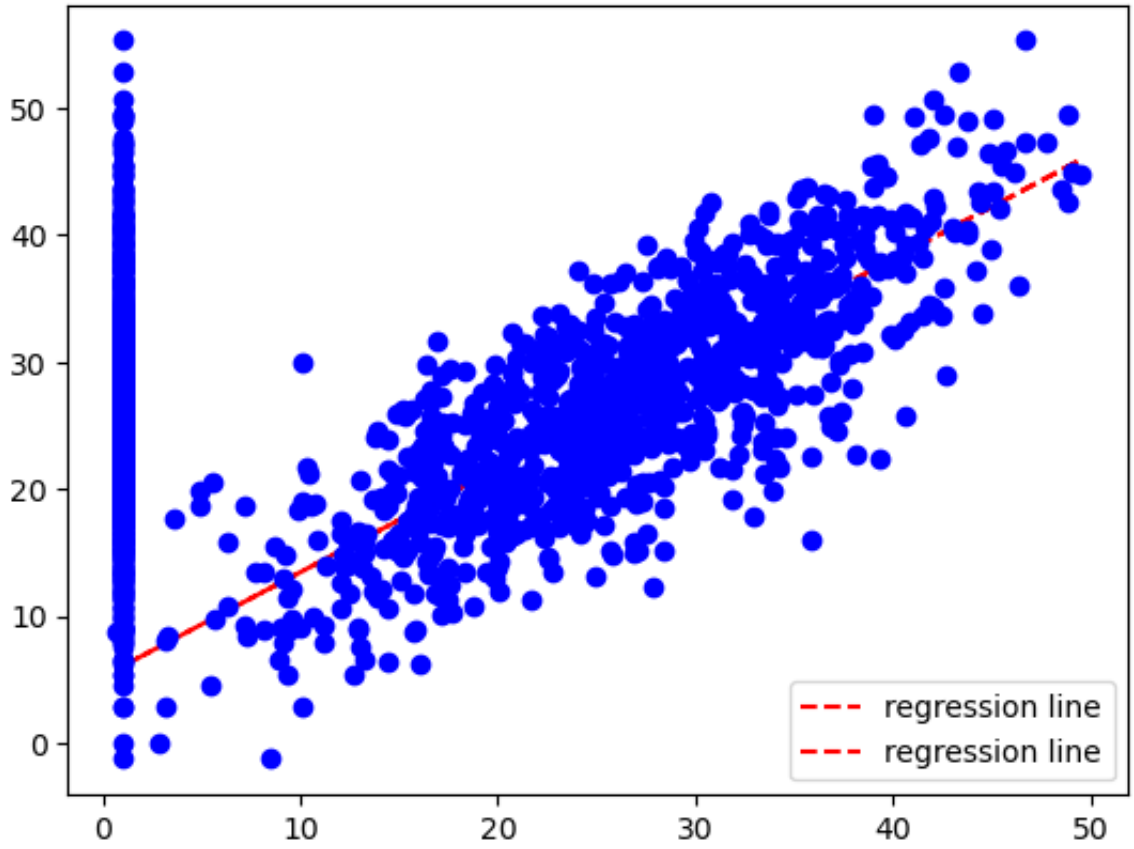
Out[111]:
```
array([[ 1.      , 18.290293],
       [ 1.      , 17.023407]])
```

In [114]:
```
y_pred = x_final.dot(beta)  # testing on the same data,
```

In [115]:
```
y_pred
```

```
       [28.02655563],
       [31.25923615],
       [16.87194932],
       [29.82886015],
       [32.22209514],
       [27.31179205],
       [39.35821682],
       [31.69535009],
       [33.78295173],
       [27.83928797],
       [26.39628385],
       [35.58259329],
       [24.73886029],
       [25.02396459],
       [18.4743638 ],
       [24.33201241],
       [26.28875744],
       [27.90861648],
       [24.05927269],
       [21.69817732],
```

In [119]:
```
plt.plot(x_final,y_pred,'r--',label='regression line') # plot the b
plt.plot(x_final,y,'bo')
plt.legend()
plt.show()
```



In [ ]: