HIRING PROJECT(using linear regression multiple)

In [11]:
```
import pandas as pd
import numpy as np
from sklearn import linear_model
from word2number import w2n
```

In [2]:
```
df = pd.read_csv("G:\\Hiringdata\\project.csv")
df
```

Out[2]:

	experience	test_score(out of 10)	interview_score(out of 10)	salary($)
0	NaN	8.0	9	50000
1	NaN	8.0	6	45000
2	five	6.0	7	60000
3	two	10.0	10	65000
4	seven	9.0	6	70000
5	three	7.0	10	62000
6	ten	NaN	7	72000
7	eleven	7.0	8	80000

In [3]:
```
df.experience = df.experience.fillna("zero")
df
```

Out[3]:

	experience	test_score(out of 10)	interview_score(out of 10)	salary($)
0	zero	8.0	9	50000
1	zero	8.0	6	45000
2	five	6.0	7	60000
3	two	10.0	10	65000
4	seven	9.0	6	70000
5	three	7.0	10	62000
6	ten	NaN	7	72000
7	eleven	7.0	8	80000

In [4]:
```
df.experience = df.experience.apply(w2n.word_to_num)
df
```

Out[4]:

	experience	test_score(out of 10)	interview_score(out of 10)	salary($)
0	0	8.0	9	50000
1	0	8.0	6	45000
2	5	6.0	7	60000
3	2	10.0	10	65000
4	7	9.0	6	70000
5	3	7.0	10	62000
6	10	NaN	7	72000
7	11	7.0	8	80000

In [5]:
```
import math
median_test_score = math.floor(df['test_score(out of 10)'].mean())
median_test_score
```

Out[5]: 7

In [6]:
```
df["test_score(out of 10)"]=df["test_score(out of 10)"].fillna(median_test_score)
df
```

Out[6]:

	experience	test_score(out of 10)	interview_score(out of 10)	salary($)
0	0	8.0	9	50000
1	0	8.0	6	45000
2	5	6.0	7	60000
3	2	10.0	10	65000
4	7	9.0	6	70000
5	3	7.0	10	62000
6	10	7.0	7	72000
7	11	7.0	8	80000

In [8]:
```
reg = linear_model.LinearRegression()
reg.fit(df[["experience","test_score(out of 10)","interview_score(out of 10)"]],df["salary($)"])
```

Out[8]: LinearRegression()

In [9]:
```
reg.predict([[2,9,6]])
```

Out[9]: array([53713.86677124])

In [10]:
```
reg.predict([[12,10,10]])
```

Out[10]: array([93747.79628651])