In [ ]:

Practical#2(Linear Regression Multiple Variables)

In [ ]:
```python
import numpy as np
import pandas as pd
from sklearn.linear_model import LinearRegression
```

In [ ]:
```python
data =pd.read_csv('home.csv')
data.head()
```

Out[ ]:

	area	bedroom	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

In [ ]:
```python
import math
```

In [ ]:
```python
data.bedroom.median()
```

Out[ ]: 4.0

In [ ]:
```python
median_bedroom=math.floor(data.bedroom.median())
```

In [ ]:
```python
median_bedroom
```

```
Out[ ]: 4
```

```
In [ ]: data.bedroom = data.bedroom.fillna(median_bedroom)
```

```
In [ ]: clf = LinearRegression()
```

```
In [ ]: clf.fit(data[['area','bedroom','age']],data.price)
```

```
Out[ ]: LinearRegression(copy_X=True, fit_intercept=True, n_jobs=None, normaliz
        e=False)
```

```
In [ ]: clf.coef_
```

```
Out[ ]: array([  112.06244194, 23388.88007794, -3231.71790863])
```

```
In [ ]: clf.intercept_
```

```
Out[ ]: 221323.00186540396
```

```
In [ ]: a=clf.predict([[3000,3,40]])
        print(a)

        [498408.25158031]
```

```
In [ ]: b=clf.predict([[2500,4,5]])
        print(b)

        [578876.03748933]
```

```python
In [3]: import numpy as np
```

```python
In [4]: import pandas as pd
```

```python
In [7]: from sklearn.linear_model import LinearRegression
```

```python
In [8]: data =  pd.read_csv('hiring.csv')
```

```python
In [ ]: data.head()
```

```python
In [13]: data["experience"].fillna("zero",inplace = True)
```

```python
In [14]: data.head()
```

Out[14]:

	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

```python
In [19]: # from skleaarn.preprocessing import LabelEncoder
         # a=LabelEncoder()
         from sklearn import preprocessing
         labelencoder=preprocessing.LabelEncoder()
         # data['experience']=labelencoder.fit_transform(data['experience'])
         # data['experience'].unique()
```

```python
In [17]: data
```

Out[17]:

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

In [ ]:

In [30]:
```python
import math
import math
data['test_score(out of 10)'].median()
t1 = math.floor(data['test_score(out of 10)'].median())
# median_test_score(out of 10)
data['test_score(out of 10)']= data['test_score(out of 10)'].fillna(t1)
```

In [ ]:

In [31]:
```python
clf.fit(data[['experience','test_score(out of 10)','interview_score(out of 10)']],data[['salary($)']])
```

Out[31]: LinearRegression(copy_X=True, fit_intercept=True, n_jobs=None, normalize=False)

In [32]:
```python
print(clf.predict([[2,9,6]]))
```

[[72016.43913945]]

```
In [33]:  print(clf.predict([[12,10,10]]))

[[28973.8879756]]
```