

```
## LINEAR REGRESSION USING MULTIPLE VARIABLES
```

```
import pandas as pd
import numpy as np
from sklearn import linear_model
df = pd.read_csv(r'C:\Users\GLAU\Downloads\homeprices.csv')
df
```

	area	bedrooms	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
5	4100	6.0	8	810000

```
df.bedrooms.median()
```

```
4.0
```

```
df.bedrooms = df.bedrooms.fillna(df.bedrooms.median())
df
```

	area	bedrooms	age	price
0	2600	3.0	20	550000
1	3000	4.0	15	565000
2	3200	4.0	18	610000
3	3600	3.0	30	595000
4	4000	5.0	8	760000
5	4100	6.0	8	810000

```
reg = linear_model.LinearRegression()
reg.fit(df.drop('price',axis='columns'),df.price)
```

```
LinearRegression()
```

```
reg.coef_
```

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

```
reg.intercept_
```

```
221323.00186540408
```

```
## Find price of home with 3000 sqr ft area, 3 bedrooms, 40 year old
```

```
reg.predict([[3000, 3, 40]])
```

```
array([498408.25158031])
```

```
112.06244194*3000 + 23388.88007794*3 + -3231.71790863*40 +  
221323.00186540384
```

```
498408.25157402386
```

```
## Find price of home with 2500 sqr ft area, 4 bedrooms, 5 year old
reg.predict([[2500, 4, 5]])
```

```
array([578876.03748933])
```

```
## hiring.csv
```

```
import pandas as pd
import numpy as np
from sklearn import linear_model
from word2number import w2n
url =
"https://github.com/codebasics/py/blob/master/ML/2_linear_reg_multivar
iate/Exercise/hiring.csv" # Make sure the url is the raw version of
the file on GitHub
download = requests.get(url).content
```

```
-----
-----
ModuleNotFoundError                                Traceback (most recent call
last)
<ipython-input-10-53f3a6e7fd47> in <module>
      3 import numpy as np
      4 from sklearn import linear_model
----> 5 from word2number import w2n
      6 url =
"https://github.com/codebasics/py/blob/master/ML/2_linear_reg_multivar
iate/Exercise/hiring.csv" # Make sure the url is the raw version of
the file on GitHub
      7 download = requests.get(url).content
```

```
ModuleNotFoundError: No module named 'word2number'
```

```
## hiring.csv problem
```

```
import pandas as pd
import numpy as np
from sklearn import linear_model
d = pd.read_csv(r"C:\Users\GLAU\Desktop\py-master\ML\
2_linear_reg_multivariate\Exercise\hiring.csv")
d
```

	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			

```
d.experience = d.experience.fillna("zero")
d
```

	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			