

```
In [3]: import numpy as np
import matplotlib.pyplot as plt
import pandas as pd

In [4]: dataset=pd.read_csv('insurance_pre.csv')
dataset

Out[4]:
```

	age	sex	bmi	children	smoker	charges
0	19	female	27.900	0	yes	16884.92400
1	18	male	33.770	1	no	1725.55230
2	28	male	33.000	3	no	4449.46200
3	33	male	22.705	0	no	21984.47061
4	32	male	28.880	0	no	3866.85520
...
1333	50	male	30.970	3	no	10600.54830
1334	18	female	31.920	0	no	2205.98080
1335	18	female	36.850	0	no	1629.83350
1336	21	female	25.800	0	no	2007.94500
1337	61	female	29.070	0	yes	29141.36030

1338 rows x 6 columns

```
In [7]: dataset.columns

Out[7]: Index(['age', 'sex', 'bmi', 'children', 'smoker', 'charges'], dtype='object')
```

```
In [13]: dataset=pd.get_dummies(dataset,drop_first=True)
dataset=dataset.astype(int)
dataset

Out[13]:
```

	age	bmi	children	charges	sex_male	smoker_yes
0	19	27	0	16884	0	1
1	18	33	1	1725	1	0
2	28	33	3	4449	1	0
3	33	22	0	21984	1	0
4	32	28	0	3866	1	0
...
1333	50	30	3	10600	1	0
1334	18	31	0	2205	0	0
1335	18	36	0	1629	0	0
1336	21	25	0	2007	0	0
1337	61	29	0	29141	0	1

1338 rows x 6 columns

```
In [15]: dataset.columns

Out[15]: Index(['age', 'bmi', 'children', 'charges', 'sex_male', 'smoker_yes'], dtype='object')
```

```
In [17]: indep=dataset[['age', 'bmi', 'children', 'charges', 'sex_male', 'smoker_yes']]
dep=dataset[['charges']]

In [19]: indep

Out[19]:
```

	age	bmi	children	charges	sex_male	smoker_yes
0	19	27	0	16884	0	1
1	18	33	1	1725	1	0
2	28	33	3	4449	1	0
3	33	22	0	21984	1	0
4	32	28	0	3866	1	0
...
1333	50	30	3	10600	1	0
1334	18	31	0	2205	0	0
1335	18	36	0	1629	0	0
1336	21	25	0	2007	0	0
1337	61	29	0	29141	0	1

1338 rows x 6 columns

```
In [21]: dep

Out[21]:
```

	charges
0	16884
1	1725
2	4449
3	21984
4	3866
...	...
1333	10600
1334	2205
1335	1629
1336	2007
1337	29141

1338 rows x 1 columns

```
In [23]: from sklearn.model_selection import train_test_split
X_train,X_test,y_train,y_test=train_test_split(indep,dep,test_size=1/3,random_state=0)

In [24]: X_train

Out[24]:
```

	age	bmi	children	charges	sex_male	smoker_yes
482	18	31	0	1622	0	0
338	50	32	1	41919	1	1
356	46	43	3	8944	1	0
869	25	24	3	4391	0	0
182	22	19	3	4005	1	0
...
763	27	26	0	3070	1	0
835	42	35	2	7160	1	0
1216	40	25	0	5415	1	0
559	19	35	0	1646	1	0
684	33	18	1	4766	0	0

892 rows x 6 columns

```
In [27]: from sklearn.linear_model import LinearRegression
regressor=LinearRegression()
regressor.fit(X_train,y_train)

Out[27]:
```

LinearRegression

LinearRegression()

```
In [29]: weight=regressor.coef_

In [31]: weight

Out[31]: array([[ -1.08972896e-14, -2.99302300e-13, -1.99250758e-13,
  1.00000000e+00,  4.07119959e-13,  3.98825293e-13]])

In [33]: bais=regressor.intercept_
bais

Out[33]: array([-1.8189894e-12])

In [35]: y_pred=regressor.predict(X_test)

In [37]: from sklearn.metrics import r2_score
r_score=r2_score(y_test,y_pred)

In [39]: r_score

Out[39]: 1.0

In [43]: import pickle
filename="finalized_model_Mul_linear.sav"
pickle.dump(regressor,open(filename,'wb'))

In [59]: loaded_model=pickle.load(open("finalized_model_Mul_linear.sav",'rb'))
result=loaded_model.predict([[16,12,13,15,28,10]])

C:\Users\ADMIN\anaconda3\Lib\site-packages\sklearn\base.py:493: UserWarning: X does not have valid feature names, but LinearRegression was fitted with feature names
warnings.warn(

In [61]: result
```

Out[61]: array([[15.]])

In []: