In [4]:
```
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
```

In [5]:
```
data =pd.read_csv('insurance_pre.csv')
```

In [6]:
```
data
```

Out[6]:

	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 × 6 columns

In [7]:
```
data=pd.get_dummies(data,drop_first=True)
```

In [8]:
```
data
```

Out[8]:

	age	bmi	children	charges	sex_male	smoker_yes
0	19	27.900	0	16884.92400	0	1
1	18	33.770	1	1725.55230	1	0
2	28	33.000	3	4449.46200	1	0
3	33	22.705	0	21984.47061	1	0
4	32	28.880	0	3866.85520	1	0
...	...	...	...	...	...	...
1333	50	30.970	3	10600.54830	1	0
1334	18	31.920	0	2205.98080	0	0
1335	18	36.850	0	1629.83350	0	0
1336	21	25.800	0	2007.94500	0	0
1337	61	29.070	0	29141.36030	0	1

1338 rows × 6 columns

In [9]:
```
independent=data[["age","bmi","children","sex_male","smoker_yes"]]
dependent=data[["charges"]]
```

In [11]:
```
from sklearn.model_selection import train_test_split
```

In [12]:
```
X_train,X_test,Y_train,Y_test=train_test_split(independent,dependent,test_size=1/3,random_state=0)
```

In [13]:
```
import numpy as np
from sklearn import linear_model
regressor=linear_model.RidgeCV(alphas=np.logspace(-6,6,13))
regressor.fit(X_train,Y_train)
```

Out[13]:
```
RidgeCV(alphas=array([1.e-06, 1.e-05, 1.e-04, 1.e-03, 1.e-02, 1.e-01, 1.e+00, 1.e+01,
       1.e+02, 1.e+03, 1.e+04, 1.e+05, 1.e+06]))
```

In [14]:
```
weight=regressor.coef_
print("weight of the model={}",format(weight))
```

```
weight of the model={} [[  260.12162836   315.19307298   545.76990774   -70.78078423
  23235.68568187]]
```

In [15]:
```
bias=regressor.intercept_
print("bias of the model={}",format(bias))
```

```
bias of the model={} [-12009.23874689]
```

In [16]:
```
y_pred=regressor.predict(X_test)
```

In [17]:
```
from sklearn.metrics import r2_score
r_score=r2_score(Y_test,y_pred)
print("r_score",r_score)
```

```
r_score 0.7864532389752167
```

In [18]:
```
age=int(input("enter the prediction input value:"))
bmi=float(input("enter the prediction input value:"))
children=int(input("enter the  prediction input value:"))
sex_male=int(input("enter the prediction input value:"))
smoker_yes=int(input("enter the prediction value:"))
future_prediction=regressor.predict([[age,bmi,children,sex_male,smoker_yes]])
print("future_prediction={}",format(future_prediction))
```

```
enter the prediction input value:20
enter the prediction input value:25.800
enter the  prediction input value:0
enter the prediction input value:0
enter the prediction value:0
future_prediction={} [[1325.17510298]]
C:\ProgramData\Anaconda3\lib\site-packages\sklearn\base.py:450: UserWarning: X does not have valid feature names, but RidgeCV was fitted with feature names
  warnings.warn(
```

In [ ]: