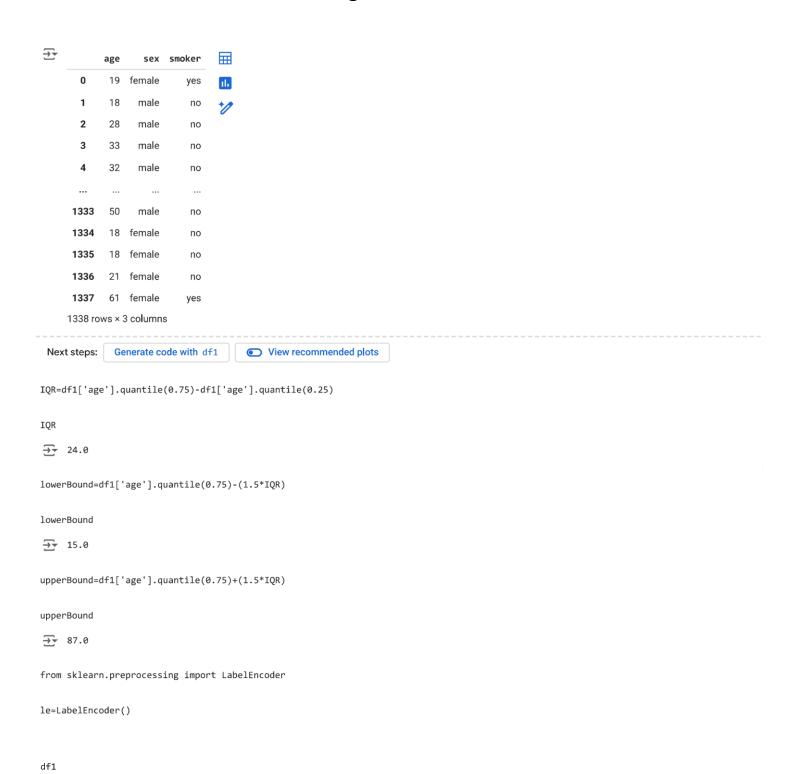
df1

## Assignment-2

```
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
data=pd.read_csv("/content/insurance.csv")
data.head()
sex
                         bmi children smoker
                                                   region
                                                               charges
                                                                          \blacksquare
          19
             female 27.900
                                                southwest 16884.92400
      1
          18
                male 33.770
                                                 southeast
                                                            1725.55230
                                            no
      2
          28
                male 33.000
                                     3
                                                 southeast
                                                            4449.46200
                                     0
                                                northwest 21984.47061
      3
          33
                male 22.705
                                            no
          32
                male 28.880
                                     0
                                                northwest
                                                            3866.85520
              Generate code with data
                                          View recommended plots
 Next steps:
data.tail()
\overrightarrow{\Rightarrow}
                                                                           \overline{\Pi}
                           bmi children smoker
                                                     region
                                                                charges
            age
                    sex
      1333
             50
                   male 30.97
                                       3
                                                  northwest 10600.5483
                                                                           ılı
      1334
             18 female 31.92
                                                   northeast
                                                               2205.9808
                                               no
      1335
                                       0
                                                   southeast
                                                               1629.8335
             18
                 female 36.85
                                               no
      1336
             21
                 female 25.80
                                       0
                                                  southwest
                                                              2007.9450
                                               no
      1337
             61 female 29.07
                                       0
                                              yes northwest 29141.3603
data.shape
→ (1338, 7)
data.isnull().sum()
     age
                  0
                  0
     sex
     bmi
                  0
     children
     smoker
                  0
     region
     charges
                 0
     dtype: int64
df1=data.drop(columns=['bmi','children','charges','region'],axis=1)
```

# Assignment-2



### Assignment-2

```
model=LogisticRegression()
model.fit(x_train,y_train)
▼ LogisticRegression
LogisticRegression()
pred=model.predict(x_test)
pred
0, 0, 0, 0])
```

### Assignment-2

```
y_test
```

#### df1.tail()

<b>→</b>		age	sex	smoker
	1333	50	1	0
	1334	18	0	0
	1335	18	0	0
	1336	21	0	0
	1337	61	0	1

model.predict([[50,1]])

→ array([0])

from sklearn.metrics import accuracy\_score

accuracy\_score(y\_test,pred)

0.8134328358208955

#

Could not connect to the reCAPTCHA service. Please check your internet connection and reload to get a reCAPTCHA challenge.