project-2-622

February 21, 2024

```
[ ]: ## importing pandas
     ## load the medical

     import pandas as pd
     data = pd.read_csv("/content/medical_insurance.csv")
     data
```

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

     [1338 rows x 7 columns]
```

```
[ ]: ##finding the size of the dataframe

     data.shape
```

```
[ ]: (1338, 7)
```

```
[ ]: ##checking if the dataframe values are containing any null values or not

     data.isna().sum()
```

```
[ ]: age         0
     sex         0
     bmi         0
     children    0
     smoker      0
```

```
region      0
charges     0
dtype: int64
```

[ ]:
```
# converting sex with numerical values

data['sex'].replace({'male': 0, 'female': 1},inplace = True)
```

[ ]:
```
##replacing smoker with integer values

data['smoker'].replace({'no': 0, 'yes': 1},inplace = True)
```

[ ]:
```
data
```

[ ]:
```
      age  sex     bmi  children  smoker     region      charges
0      19    1  27.900         0       1  southwest  16884.92400
1      18    0  33.770         1       0  southeast   1725.55230
2      28    0  33.000         3       0  southeast   4449.46200
3      33    0  22.705         0       0  northwest  21984.47061
4      32    0  28.880         0       0  northwest   3866.85520
…     …    …    …       …      …       …          …
1333   50    0  30.970         3       0  northwest  10600.54830
1334   18    1  31.920         0       0  northeast   2205.98080
1335   18    1  36.850         0       0  southeast   1629.83350
1336   21    1  25.800         0       0  southwest   2007.94500
1337   61    1  29.070         0       1  northwest  29141.36030

[1338 rows x 7 columns]
```

[ ]:
```
# printing the columns of dataframe

data.columns
```

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

[ ]:
```
# assigning valuesv to the x and y varibales:

x = data[['age', 'sex', 'bmi', 'children', 'smoker']]
y = data[['charges']]
```

[ ]:
```
response =  data[['age', 'sex', 'bmi', 'children', 'smoker']]
response
```

[ ]:
```
      age  sex     bmi  children  smoker
0      19    1  27.900         0       1
1      18    0  33.770         1       0
```

```
2       28    0  33.000         3       0
3       33    0  22.705         0       0
4       32    0  28.880         0       0
…    …  …     …       …      …
1333    50    0  30.970         3       0
1334    18    1  31.920         0       0
1335    18    1  36.850         0       0
1336    21    1  25.800         0       0
1337    61    1  29.070         0       1

[1338 rows x 5 columns]
```

[ ]:
```
# importing the modules and training ,testing

from sklearn.model_selection import train_test_split
x_train,x_test,y_train,y_test = train_test_split(x,y,test_size = 0.2,␣
 ↪random_state=56)
from sklearn.linear_model import LinearRegression

model = LinearRegression()
model
```

[ ]:
```
LinearRegression()
```

[ ]:
```
# fitting the model

model.fit(x_train,y_train)
```

[ ]:
```
LinearRegression()
```

[ ]:
```
# predicting the model value

y_pred = model.predict(x_test)
y_pred
```

[ ]:
```
array([[10497.55082509],
       [ 4445.3141368 ],
       [ 2954.03192707],
       [ 7498.99579688],
       [31164.09019604],
       [ 1578.82547813],
       [ 7531.05082207],
       [11186.68833031],
       [ 6272.24780508],
       [36203.23930768],
       [ 9319.83062278],
       [11048.3299926 ],
       [10764.41995348],
       [ 6697.57516913],
       [ 6753.21985264],
       [14317.90766598],
       [14418.83201408],
       [ 2878.96487508],
       [-2694.91900168],
       [ -807.72404282],
       [  705.94687688],
       [ 7955.61220972],
       [34829.71545793],
       [ 8730.90295494],
       [11878.76608438],
       [ 9438.95551076],
       [12140.46980554],
       [14250.37021776],
       [12016.23026466],
       [26826.15100105],
       [ 2812.71449421],
       [ 7149.28919514],
       [11823.36628633],
       [10264.411789  ],
       [11849.22923542],
       [ 7006.68469493],
       [10347.28986265],
       [10215.6578853 ],
       [ 9714.12804541],
       [ 4575.4925934 ],
       [ 7606.26053953],
       [12032.03398861],
       [ 4098.50630987],
       [ 7774.66071614],
       [ 5760.55260701],
       [ 3563.77545779],
       [ 9014.70918895],
       [ 5782.27610638],
       [ 2314.40473593],
       [15028.87480361],
       [ 5224.80291213],
       [10165.43667955],
       [10440.24659573],
       [14461.3190369 ],
       [15696.00593738],
       [ 8789.16266744],
       [ 7873.76675193],
       [ 9093.85233101],
       [10865.64356041],
       [ 9323.73534105],
       [35604.47191889],
       [16802.50913862],
       [ 8442.4806936 ],
       [32880.10132281],
       [ 9316.4433764 ],
       [12234.2979367 ],
       [ 7704.83982314],
       [ 9925.48225642],
       [ 4177.46618617],
       [40331.16906568],
       [10377.4411679 ],
       [ 4240.46777909],
       [ 5487.23333198],
       [ 7487.83014823],
       [ 1157.77689596],
       [33978.66881703],
       [ 7907.96210761],
       [ 5213.15473889],
       [ 9917.35312963],
       [37315.90558469],
       [ 7285.00350353],
       [11103.19883968],
       [ 3640.535263  ],
       [10621.27522202],
       [ 5832.66447689],
       [ 5638.33973148],
       [12508.19962838],
       [30742.05089767],
       [ 8420.50636179],
       [15329.95509469],
       [39858.87319687],
       [31561.68787231],
       [ 3436.76994753],
       [13192.3832811 ],
       [30692.86854876],
       [13193.48316965],
       [38738.76039579],
       [ 7385.36790034],
       [15038.09091963],
       [29117.43859015],
       [ 8114.67433027],
       [ -908.27561842],
       [ 5362.84285071],
       [14423.33495601],
       [12194.11653003],
       [ 1118.34886036],
       [12442.69870554],
       [10723.22623899],
       [13834.62941428],
       [26037.62342848],
       [24467.23069607],
       [ 7862.30575749],
       [33848.78367227],
       [ 7611.20743969],
       [30652.73615646],
       [ 3588.05990678],
       [11154.04697543],
       [12227.13065741],
       [ 1700.94208064],
       [23968.91107777],
       [ 5103.50665939],
       [33833.25948709],
       [11968.29279769],
       [ 5539.69650023],
       [ 3798.12675645],
       [34075.24448863],
       [ 1566.10974172],
       [10350.48398324],
       [35543.81509646],
       [32230.56322498],
       [ 5329.96356271],
       [ 2215.95101142],
       [ 3511.89861923],
       [ 8179.66996918],
       [ 6775.02976501],
       [ 1895.22825969],
       [ 5677.06078844],
       [ 4116.73614407],
       [ 5597.17442933],
       [12303.55263744],
       [ 6527.75279179],
       [11189.03904779],
       [ 5768.44700235],
       [ 5321.78600954],
       [ 2319.82459465],
       [11854.97111231],
       [26377.05229802],
       [ 5344.63839779],
       [ 4933.41372956],
       [ 4103.84308889],
       [18081.77134699],
       [15500.14995053],
       [11404.19575707],
       [11904.49723339],
       [ 6541.09371825],
       [10375.26980314],
       [29583.44481872],
       [ 8603.45581439],
       [34698.94239672],
       [ 2989.58786377],
       [ 1840.43322027],
       [38269.70083298],
       [23691.08465377],
       [13358.60223293],
       [32939.95751551],
       [ 9385.09681417],
       [ 8339.52011412],
       [12343.39563093],
       [14392.67545915],
       [15598.87422753],
       [16604.23719554],
       [ 3832.67516458],
       [ 5505.51491761],
       [ 5997.46152199],
       [10490.75661229],
       [30750.54351667],
       [ 1716.12219966],
       [ 5966.46419995],
       [ 1352.51298022],
       [ 3028.49842747],
       [ 7460.42900535],
       [10182.23212508],
       [ 6373.91898928],
       [31435.97175734],
       [12684.01167576],
       [15829.1853888 ],
       [ 3915.6016646 ],
       [ 9984.40414022],
       [10364.88593789],
       [ 4221.43369612],
       [ 4139.48210528],
       [  168.21141094],
       [12103.47615511],
       [ 3945.02947284],
       [ 4852.43110094],
       [10294.80102737],
       [ 9702.27384706],
       [  738.93229792],
       [28352.38076979],
       [ 5490.35988588],
       [ -384.60428195],
       [11656.63580958],
       [15969.73060428],
       [30462.94102537],
       [ 6612.61301746],
       [13968.8612016 ],
       [ 1786.24235149],
       [ 4807.69790859],
       [ 6274.09091068],
       [ 9774.54926265],
       [11970.42559609],
       [38308.15395938],
       [ 2282.52935745],
       [11166.09429956],
       [ 8301.3040471 ],
       [14024.35697867],
       [26108.9920813 ],
       [13580.23417614],
       [ 1296.75651068],
       [ 2215.95101142],
       [ 4304.24753638],
       [13464.75098603],
       [30364.24632844],
       [ 2802.48156935],
       [12077.39107991],
       [26666.11164336],
       [ 6285.227853  ],
       [ 6990.95448464],
       [14853.96053272],
       [ 9227.34060728],
       [ 3038.70497394],
       [10887.33747971],
       [18383.69112814],
       [36663.15364622],
       [11378.15316127],
       [30257.67275742],
       [  983.7831609 ],
       [ 2665.33316627],
       [35080.65226397],
       [ 3549.19385642],
       [29961.31753439],
       [37334.65026886],
       [11916.55674549],
       [ 7044.07387721],
       [34425.24643615],
       [ 9625.37920247],
       [ 4453.08267905],
       [30835.66052177],
       [ 2578.6723144 ],
       [26420.24397154],
       [-1252.98559426],
       [36345.16815753],
       [33216.27474605],
       [14375.47258781],
       [ 9100.41455767],
       [11125.2595845 ],
       [ 6050.40750883],
       [34569.17295103],
       [11548.47923744],
       [16392.76728546],
       [15907.04508256],
       [12526.00101737],
       [ -431.95903827],
       [10131.82794756],
       [11566.84810981],
       [30086.74425055],
       [-1184.92472716],
       [ 2086.83779005]])
```

[ ]:

```python
# predicting the model value

inputdata = [[19,1,27,0,1]]
prediction = model.predict(inputdata)
prediction
```

```
/usr/local/lib/python3.10/dist-packages/sklearn/base.py:439: UserWarning: X does
not have valid feature names, but LinearRegression was fitted with feature names
  warnings.warn(
```

[ ]:

```
array([[25687.46097455]])
```

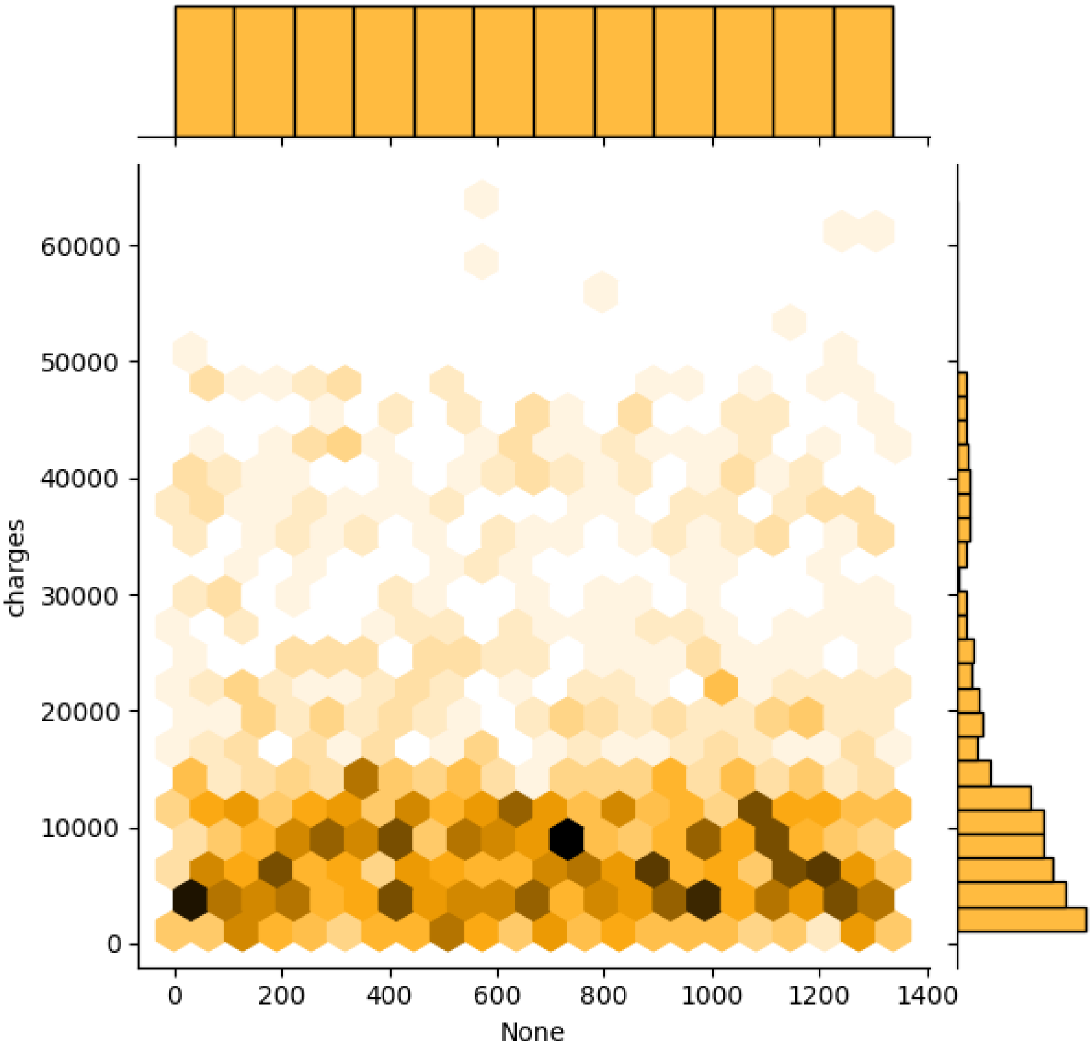
[139]:

```python
## plotting the jointplot graph for the indices of x and y

import seaborn as sns
sns.jointplot(x=x.index,y='charges',data = data,kind='hex',color='orange')
```

[139]:

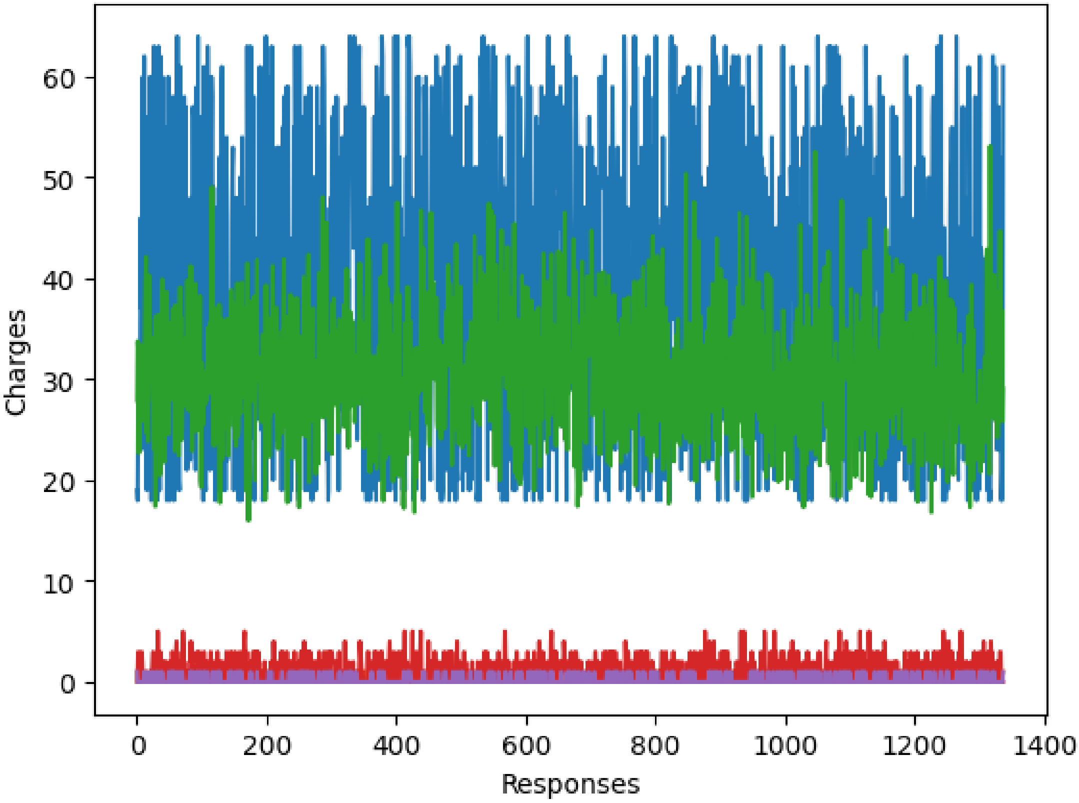
```
<seaborn.axisgrid.JointGrid at 0x7a01d4864460>
```



```
[ ]: ## plotting the graph between response and charges:

     import matplotlib.pyplot as plt
     plt.plot(response.index,response)
     plt.xlabel('Responses')
     plt.ylabel('Charges')
```

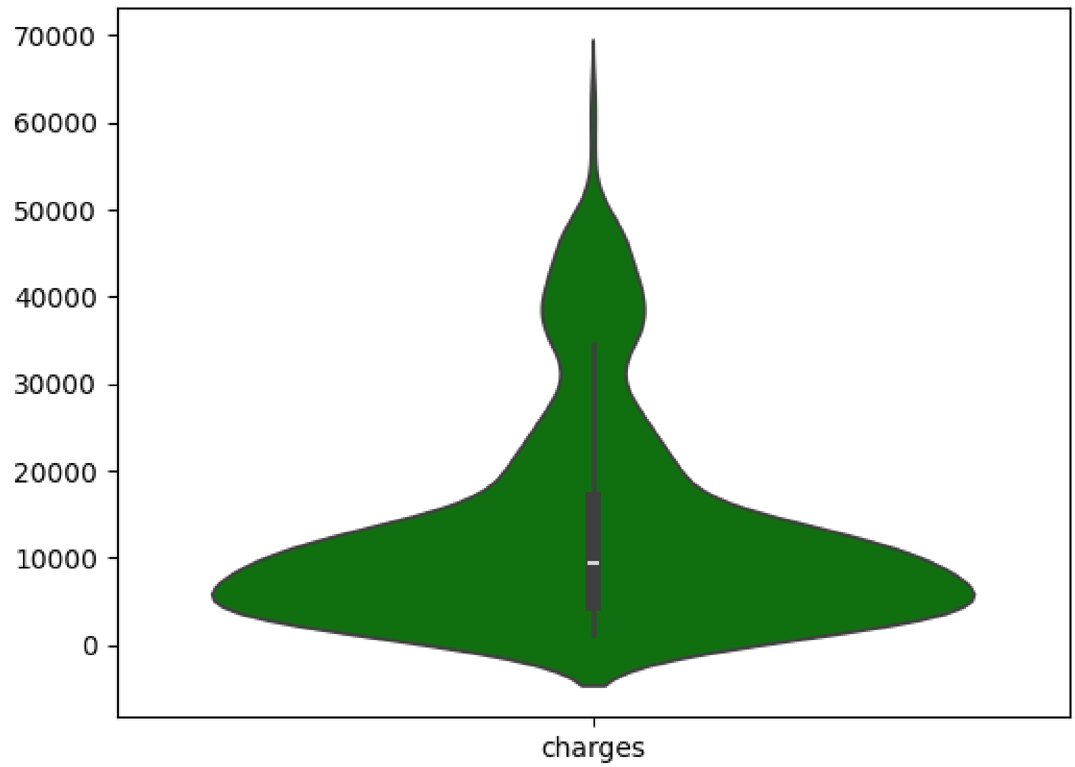
```
[ ]: Text(0, 0.5, 'Charges')
```



[114]:
```
#violinplot for charges

sns.violinplot(y,color='green')
```

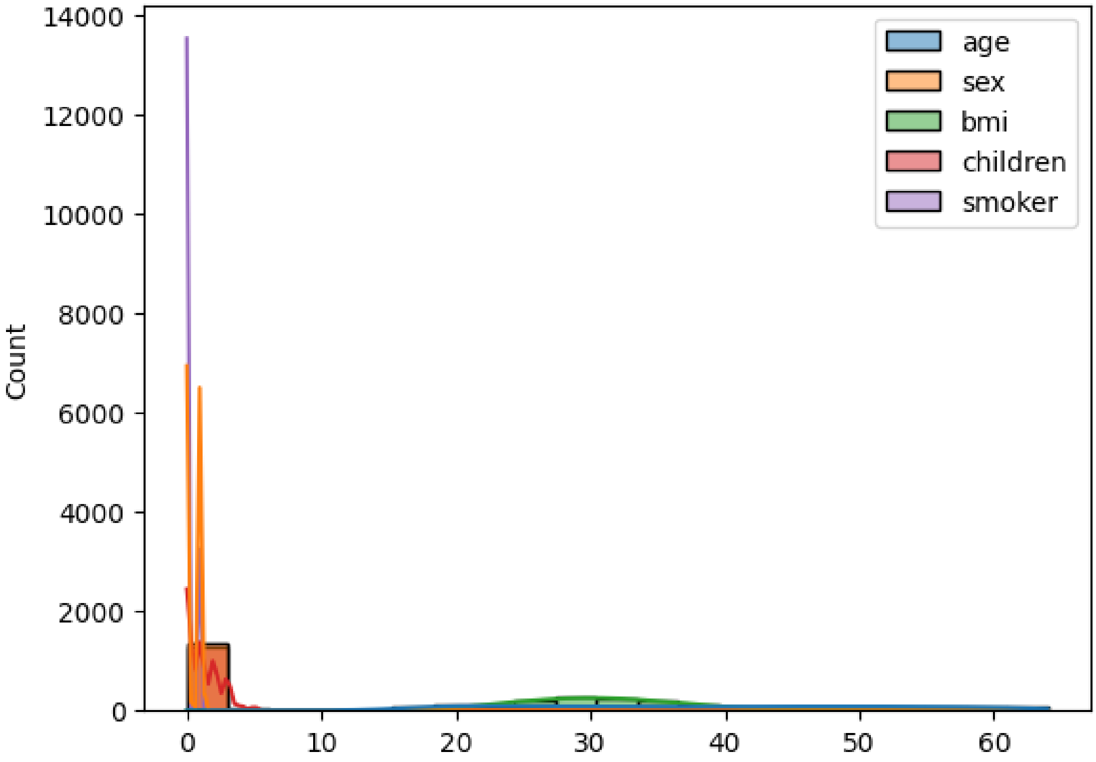
[114]:
```
<Axes: >
```



```
[89]: #hist plot for responses and charges

      import seaborn as sns
      sns.histplot(response,color = 'red',kde= True)
```

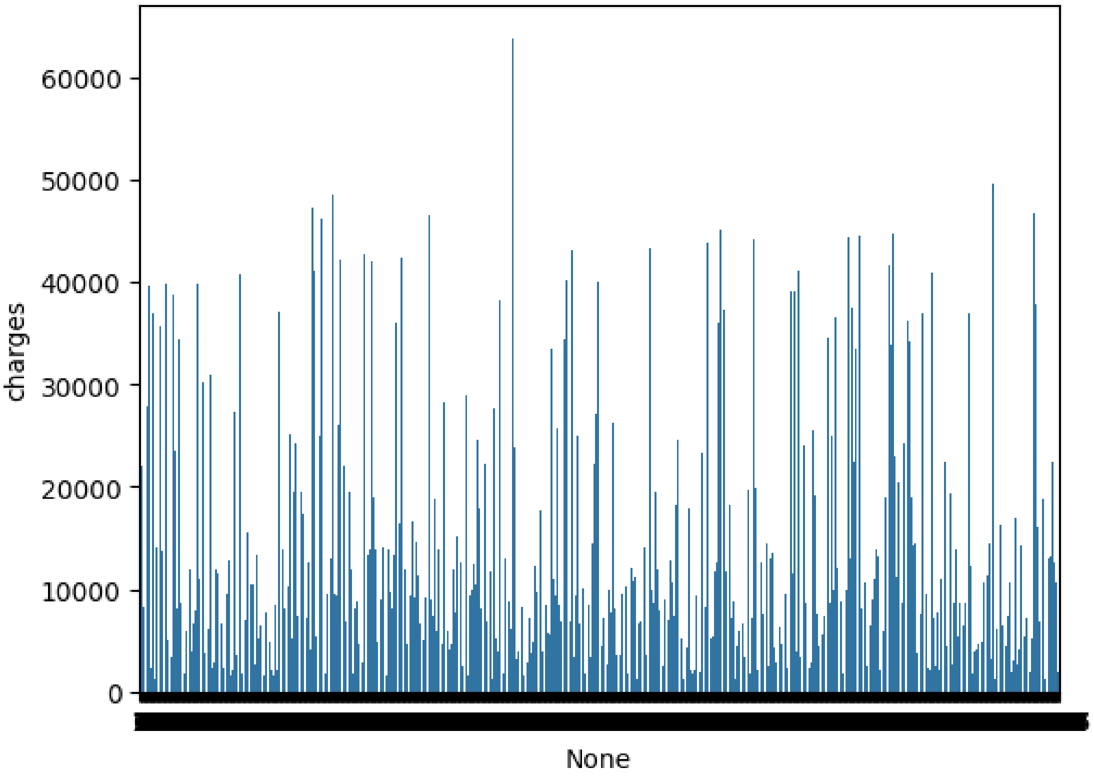
```
[89]: <Axes: ylabel='Count'>
```



```
[ ]: ## barplot between response and charges

     sns.barplot(x = response.index,y=data['charges'])
```

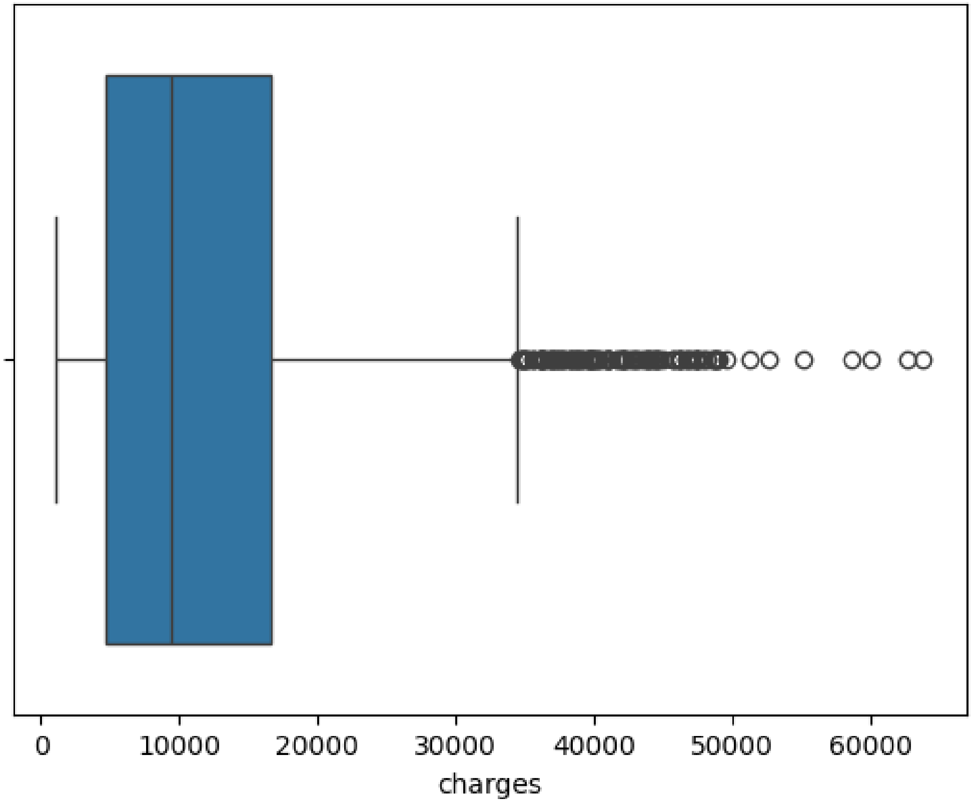
```
[ ]: <Axes: xlabel='None', ylabel='charges'>
```



```
[ ]: ## boxplot for charges

     sns.boxplot(x=data['charges'])
```

```
[ ]: <Axes: xlabel='charges'>
```

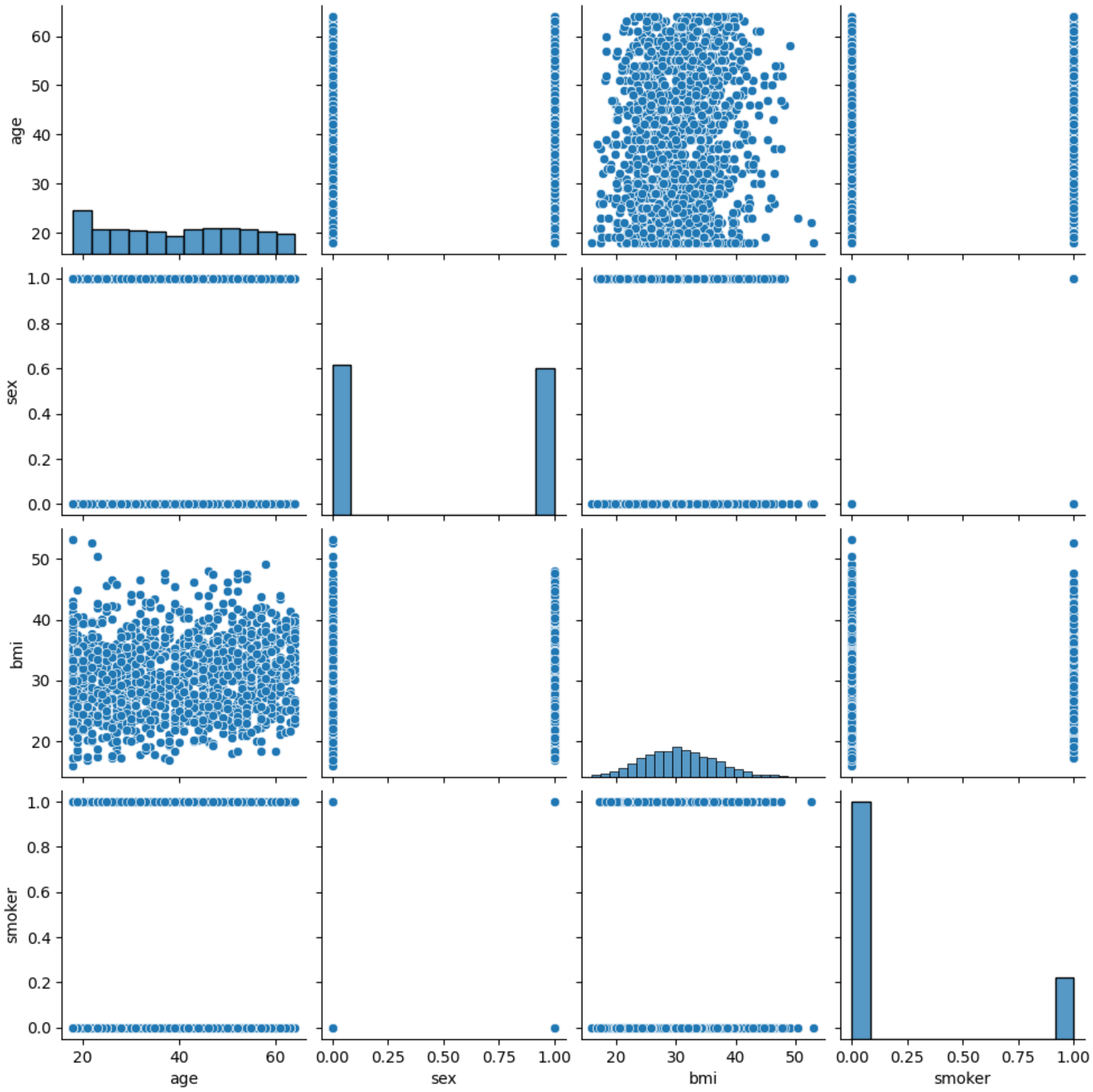


[113]:
```
## printing pairplt for the dataframe

import seaborn as sns

sns.pairplot(data[['age', 'sex', 'bmi', 'smoker']])
```

[113]:
```
<seaborn.axisgrid.PairGrid at 0x7a01d3b2d420>
```


age
60
50
40
30
20
sex
1.0
0.8
0.6
0.4
0.2
0.0
bmi
50
40
30
20
smoker
1.0
0.8
0.6
0.4
0.2
0.0
20
40
60
age
0.00
0.25
0.50
0.75
1.00
sex
20
30
40
50
bmi
0.00
0.25
0.50
0.75
1.00
smoker

[ ]: