Import Library | pandas | numpy

```
In [1]: import numpy as np
        import pandas as pd
```

Import Datasets

```
In [2]: df = pd.read_csv('customer.csv')
```

```
In [5]: df.sample(5)
```

Out[5]:

	age	gender	review	education	purchased
40	39	Male	Good	School	No
14	15	Male	Poor	PG	Yes
41	23	Male	Good	PG	Yes
34	86	Male	Average	School	No
39	76	Male	Poor	PG	No

Eliminate Unnecessary Columns

```
In [6]: df = df.iloc[:,2:]
```

```
In [7]: df.sample(5)
```

Out[7]:

	review	education	purchased
45	Poor	PG	Yes
29	Average	UG	Yes
21	Average	PG	No
46	Poor	PG	No
27	Poor	PG	No

Train_Test_Split

```
In [10]: from sklearn.model_selection import train_test_split
         x_train,x_test,y_train,y_test = train_test_split(df.iloc[:,0:2],df.iloc[:,-1:],test_s
```

Import_OrdinalEncoder

```
In [11]: from sklearn.preprocessing import OrdinalEncoder
```

```
In [12]: x_train
```

Out[12]:

	review	education
42	Good	PG
7	Poor	School
36	Good	UG
23	Good	School
0	Average	School
43	Poor	PG
3	Good	PG
49	Good	UG
13	Average	School
12	Poor	School
24	Average	PG
5	Average	School
28	Poor	School
21	Average	PG
4	Average	UG
10	Good	UG
33	Good	PG
31	Poor	School
18	Good	School
19	Poor	PG
44	Average	UG
1	Poor	UG
2	Good	PG
17	Poor	UG
41	Good	PG
30	Average	UG
47	Good	PG
26	Poor	PG
20	Average	School
22	Poor	PG
40	Good	School
14	Poor	PG
8	Average	UG
38	Good	School
6	Good	School
48	Good	UG
11	Good	UG
25	Good	School
29	Average	UG
27	Poor	PG

Define Categories

```
In [13]: oe = OrdinalEncoder(categories=[['Poor','Average','Good'],['School','UG','PG']])
```

```
In [14]: oe.fit(x_train)
```

```
Out[14]: OrdinalEncoder(categories=[['Poor', 'Average', 'Good'], ['School', 'UG', 'PG']])
```

In a Jupyter environment, please rerun this cell to show the HTML representation or trust the notebook.

On GitHub, the HTML representation is unable to render, please try loading this page with nbviewer.org.

Transform Ordinal Encoding

```
In [15]: x_train = oe.transform(x_train)
         x_test = oe.transform(x_test)
```

```
In [16]: x_train
```

```
Out[16]: array([[2., 2.],
                [0., 0.],
                [2., 1.],
                [2., 0.],
                [1., 0.],
                [0., 2.],
                [2., 2.],
                [2., 1.],
                [1., 0.],
                [0., 0.],
                [1., 2.],
                [1., 0.],
                [0., 0.],
                [1., 2.],
                [1., 1.],
                [2., 1.],
                [2., 2.],
                [0., 0.],
                [2., 0.],
                [0., 2.],
                [1., 1.],
                [0., 1.],
                [2., 2.],
                [0., 1.],
                [2., 2.],
                [1., 1.],
                [2., 2.],
                [0., 2.],
                [1., 0.],
                [0., 2.],
                [2., 0.],
                [0., 2.],
                [1., 1.],
                [2., 0.],
                [2., 0.],
                [2., 1.],
                [2., 1.],
                [2., 0.],
                [1., 1.],
                [0., 2.]])
```

```
In [17]: oe.categories_
```

```
Out[17]: [array(['Poor', 'Average', 'Good'], dtype=object),
           array(['School', 'UG', 'PG'], dtype=object)]
```

Label Encoding

```
In [18]: from sklearn.preprocessing import LabelEncoder
```

```
In [19]: le = LabelEncoder()
```

```
In [20]: le.fit(y_train)
```

```
C:\Users\ASUS\anaconda3\Lib\site-packages\sklearn\preprocessing\_label.py:97: DataCo
nversionWarning: A column-vector y was passed when a 1d array was expected. Please c
hange the shape of y to (n_samples, ), for example using ravel().
  y = column_or_1d(y, warn=True)
```

```
Out[20]: LabelEncoder()
```

In a Jupyter environment, please rerun this cell to show the HTML representation or trust the notebook.
On GitHub, the HTML representation is unable to render, please try loading this page with nbviewer.org.

```
In [21]: le.classes_
```

```
Out[21]: array(['No', 'Yes'], dtype=object)
```

```
In [22]: y_train = le.transform(y_train)
         y_test = le.transform(y_test)
```

```
C:\Users\ASUS\anaconda3\Lib\site-packages\sklearn\preprocessing\_label.py:132: DataC
onversionWarning: A column-vector y was passed when a 1d array was expected. Please
change the shape of y to (n_samples, ), for example using ravel().
  y = column_or_1d(y, dtype=self.classes_.dtype, warn=True)
C:\Users\ASUS\anaconda3\Lib\site-packages\sklearn\preprocessing\_label.py:132: DataC
onversionWarning: A column-vector y was passed when a 1d array was expected. Please
change the shape of y to (n_samples, ), for example using ravel().
  y = column_or_1d(y, dtype=self.classes_.dtype, warn=True)
```

```
In [23]: y_train
```

```
Out[23]: array([1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 1, 1, 0, 0, 0, 1, 1, 1, 0, 1, 0, 0,
                0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 0, 0, 0, 1, 1, 0, 1, 0])
```

```
In [ ]:
```