




```
import numpy as np
import pandas as pd
```

```
df = pd.read_csv('customer.csv')
```

```
df.sample(5)
```




	age	gender	review	education	purchased
5	31	Female	Average	School	Yes
9	74	Male	Good	UG	Yes
19	97	Male	Poor	PG	Yes
35	74	Male	Poor	School	Yes
27	69	Female	Poor	PG	No





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

```
df.head()
```



	review	education	purchased
0	Average	School	No
1	Poor	UG	No
2	Good	PG	No
3	Good	PG	No
4	Average	UG	No

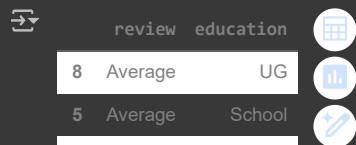


Next steps:

[Generate code with df](#)[View recommended plots](#)[New interactive sheet](#)

```
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_size=0.2)
```

```
X_train
```



	review	education
8	Average	UG
5	Average	School
4	Average	UG
30	Average	UG
40	Good	School
47	Good	PG
0	Average	School
29	Average	UG
13	Average	School
1	Poor	UG
36	Good	UG
28	Poor	School
19	Poor	PG
49	Good	UG
32	Average	UG
26	Poor	PG
33	Good	PG
11	Good	UG
17	Poor	UG
48	Good	UG
15	Poor	UG
43	Poor	PG
41	Good	PG
27	Poor	PG
20	Average	School
37	Average	PG
39	Poor	PG
35	Poor	School
16	Poor	UG
2	Good	PG
38	Good	School
21	Average	PG
18	Good	School
7	Poor	School
3	Good	PG
10	Good	UG
12	Poor	School
42	Good	PG
45	Poor	PG
34	Average	School

Next steps:

[Generate code with X\\_train](#)[View recommended plots](#)[New interactive sheet](#)

```
from sklearn.preprocessing import OrdinalEncoder
```

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

```
oe.fit(X_train)
```



OrdinalEncoder

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

```
X_train = oe.transform(X_train)
```

X\_train



```
array([[1., 1.],
       [1., 0.],
       [1., 1.],
       [1., 1.],
       [2., 0.],
       [2., 2.],
       [1., 0.],
       [1., 1.],
       [1., 0.],
       [0., 1.],
       [2., 1.],
       [0., 0.],
       [0., 2.],
       [2., 1.],
       [1., 1.],
       [0., 2.],
       [2., 2.],
       [2., 1.],
       [0., 1.],
       [2., 1.],
       [0., 1.],
       [0., 2.],
       [2., 2.],
       [0., 2.],
       [1., 0.],
       [1., 2.],
       [0., 2.],
       [0., 0.],
       [0., 1.],
       [2., 2.],
       [2., 0.],
       [1., 2.],
       [2., 0.],
       [0., 0.],
       [2., 2.],
       [2., 1.],
       [0., 0.],
       [2., 2.],
       [0., 2.],
       [1., 2.],
       [2., 1.]])
```

oe.categories\_



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

X\_train




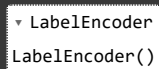
```
array([[1., 1.],
       [1., 0.],
       [1., 1.],
       [1., 1.],
       [2., 0.],
       [2., 2.],
       [1., 0.],
       [1., 1.],
       [1., 0.],
       [0., 1.],
       [2., 1.],
       [0., 0.],
       [0., 2.],
       [2., 1.],
       [1., 1.],
       [0., 2.],
       [2., 2.],
       [2., 1.],
       [0., 1.],
       [2., 1.]])
```

```
[0., 1.],  
[0., 2.],  
[2., 2.],  
[0., 2.],  
[1., 0.],  
[1., 2.],  
[0., 2.],  
[0., 0.],  
[0., 1.],  
[2., 2.],  
[2., 0.],  
[1., 2.],  
[2., 0.],  
[0., 0.],  
[2., 2.],  
[2., 1.],  
[0., 0.],  
[2., 2.],  
[0., 2.],  
[1., 0.]])
```


```
from sklearn.preprocessing import LabelEncoder
```

```
le = LabelEncoder()
```

```
le.fit(y_train)
```

```
le.classes_
```

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