Encoding

In [20]:
```
import pandas as pd
```

In [21]:
```
data1 = pd.read_csv('encoding.csv' , usecols = ['Status'] )
```

In [22]:
```
data1.head()
```

Out[22]:

	Status
0	Spam
1	Ham
2	Spam
3	Ham
4	Ham

Label Encoding Using Pandas Library Function

In [23]:
```
data1.loc[data1['Status'] == 'Spam', 'Status' ] = 0
```

In [24]:
```
data1.loc[data1['Status'] == 'Ham' , 'Status'] = 1
```

In [25]:
```
data1.head()
```

Out[25]:

	Status
0	0
1	1
2	0
3	1
4	1

In [26]:
```
data1.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10 entries, 0 to 9
Data columns (total 1 columns):
Status    10 non-null int64
dtypes: int64(1)
memory usage: 160.0 bytes
```

In [27]:
```
data2 = pd.read_csv('encoding.csv' , usecols = ['State'])
```

In [28]:
```
data2
```

Out[28]:

	State
0	West Bengal
1	Bihar
2	UP
3	UP
4	Bihar
5	West Bengal
6	West Bengal
7	Bihar
8	Bihar
9	UP

OneHotEncoder using Pandas Library Functions

In [32]:
```
dummies = pd.get_dummies(data2.State)
```

In [33]:
```
dummies.head()
```

Out[33]:

	Bihar	UP	West Bengal
0	0	0	1
1	1	0	0
2	0	1	0
3	0	1	0
4	1	0	0

In [44]:
```
onehotencoder = pd.concat([data2,dummies] , axis = 'columns')
```

In [45]:
```
onehotencoder
```

Out[45]:

	State	Bihar	UP	West Bengal
0	West Bengal	0	0	1
1	Bihar	1	0	0
2	UP	0	1	0
3	UP	0	1	0
4	Bihar	1	0	0
5	West Bengal	0	0	1
6	West Bengal	0	0	1
7	Bihar	1	0	0
8	Bihar	1	0	0
9	UP	0	1	0

Dummy Variable Trap

In [46]:
```
onehotencoder.drop(['West Bengal', 'State'] , axis = 1 , inplace = True)
```

In [47]:
```
onehotencoder
```

Out[47]:

	Bihar	UP
0	0	0
1	1	0
2	0	1
3	0	1
4	1	0
5	0	0
6	0	0
7	1	0
8	1	0
9	0	1

In [ ]:
```

```