

PRACTICAL - 07

7. Consider a data frame containing data about students i.e. name, gender and passing division:

	Name	Birth_Month	Gender	Pass_Division
0	Mudit Chauhan	December	M	III
1	Seema Chopra	January	F	II
2	Rani Gupta	March	F	I
3	Aditya Narayan	October	M	I
4	Sanjeev Sahni	February	M	II
5	Prakash Kumar	December	M	III
6	Ritu Agarwal	September	F	I
7	Akshay Goel	August	M	I
8	Meeta Kulkarni	July	F	II
9	Preeti Ahuja	November	F	II
10	Sunil Das Gupta	April	M	III
11	Sonali Sapre	January	F	I
12	Rashmi Talwar	June	F	III
13	Ashish Dubey	May	M	II
14	Kiran Sharma	February	F	II
15	Sameer Bansal	October	M	I

a. Perform one hot encoding of the last two columns of categorical data using the `get_dummies()` function.

b. Sort this data frame on the "Birth Month" column (i.e. January to December).
Hint: Convert Month to Categorical.

CODE :

```
[2] import pandas as pd
```

```
[8] df=pd.DataFrame({  
    'Name':['Mohit Chauhan','Seema Chopra','Rani Gupta','Aditya Narayan','Sanjeev Sahni',  
           'Prakash Kumar','Ritu Agarwal','Akshay Goel','Meeta Kulkarni','Preeti Ahuja',  
           'Sunil Das Gupta','Sonali Sapre','Rashmi Talwar','Ashish Dubey','Kiran Sharma',  
           'Sameer Bansal'],  
    'Birth_Month':['December','January','March','October','February','December','September',  
                  'August','July','November','April','January','June','May','February','October'],  
    'Gender':['M','F','F','M','M','M','F','M','F','F','M','F','F','M','F','M'],  
    'Pass_Division':['III','II','I','I','II','III','I','I','II','II','III','I','III','II','II','I']  
})  
df
```

...

	Name	Birth_Month	Gender	Pass_Division
0	Mohit Chauhan	December	M	III
1	Seema Chopra	January	F	II
2	Rani Gupta	March	F	I
3	Aditya Narayan	October	M	I
4	Sanjeev Sahni	February	M	II
5	Prakash Kumar	December	M	III
6	Ritu Agarwal	September	F	I
7	Akshay Goel	August	M	I
8	Meeta Kulkarni	July	F	II
9	Preeti Ahuja	November	F	II
10	Sunil Das Gupta	April	M	III
11	Sonali Sapre	January	F	I
12	Rashmi Talwar	June	F	III
13	Ashish Dubey	May	M	II
14	Kiran Sharma	February	F	II
15	Sameer Bansal	October	M	I

- a. Perform one hot encoding of the last two columns of categorical data using the `get_dummies()` function.

```
df_encoded=pd.get_dummies(df,columns=['Gender','Pass_Division'])
df_encoded
```

[10] ✓ 0.0s

	Name	Birth_Month	Gender_F	Gender_M	Pass_Division_I	Pass_Division_II	Pass_Division_III
0	Mohit Chauhan	December	False	True	False	False	True
1	Seema Chopra	January	True	False	False	True	False
2	Rani Gupta	March	True	False	True	False	False
3	Aditya Narayan	October	False	True	True	False	False
4	Sanjeev Sahni	February	False	True	False	True	False
5	Prakash Kumar	December	False	True	False	False	True
6	Ritu Agarwal	September	True	False	True	False	False
7	Akshay Goel	August	False	True	True	False	False
8	Meeta Kulkarni	July	True	False	False	True	False
9	Preeti Ahuja	November	True	False	False	True	False
10	Sunil Das Gupta	April	False	True	False	False	True
11	Sonali Sapre	January	True	False	True	False	False
12	Rashmi Talwar	June	True	False	False	False	True
13	Ashish Dubey	May	False	True	False	True	False
14	Kiran Sharma	February	True	False	False	True	False
15	Sameer Bansal	October	False	True	True	False	False

- b. Sort this data frame on the “Birth Month” column (i.e. January to December). Hint: Convert Month to Categorical.

```
month_order=['January','February','March','April','May','June',
             'July','August','September','October','November','December']
df_encoded['Birth_Month']=pd.Categorical(df_encoded['Birth_Month'],
                                         categories=month_order,ordered=True)
df_sorted=df_encoded.sort_values(by='Birth_Month')
df_sorted
```

[11] ✓ 0.0s

SUBMITTED TO : DR. PRITI JAGWANI

NAME : GARIMA ROHILLA

ROLL NO : CSC/21/37

UNIVERSITY ROLL NO : 21059570010

...

	Name	Birth_Month	Gender_F	Gender_M	Pass_Division_I	Pass_Division_II	Pass_Division_III
1	Seema Chopra	January	True	False	False	True	False
11	Sonali Sapre	January	True	False	True	False	False
4	Sanjeev Sahni	February	False	True	False	True	False
14	Kiran Sharma	February	True	False	False	True	False
2	Rani Gupta	March	True	False	True	False	False
10	Sunil Das Gupta	April	False	True	False	False	True
13	Ashish Dubey	May	False	True	False	True	False
12	Rashmi Talwar	June	True	False	False	False	True
8	Meeta Kulkarni	July	True	False	False	True	False
7	Akshay Goel	August	False	True	True	False	False
6	Ritu Agarwal	September	True	False	True	False	False
3	Aditya Narayan	October	False	True	True	False	False
15	Sameer Bansal	October	False	True	True	False	False
9	Preeti Ahuja	November	True	False	False	True	False
0	Mohit Chauhan	December	False	True	False	False	True
5	Prakash Kumar	December	False	True	False	False	True