```
In [2]: import numpy as np
        import pandas as pd
        import matplotlib.pyplot as plt
        import seaborn as sns
```

```
In [3]: pwd
```

Out[3]: '/Users/jatin/Documents/DAV/Practicles'

```
In [4]: ls
```

```
Datasets/          Prac4.ipynb        Prac_4_Sheet.csv    prac7.csv
PDFS/              Prac5.ipynb        Practicle_1.ipynb
Prac2.ipynb        Prac6.ipynb        delete.ipynb
Prac3.ipynb        Prac7.ipynb        delete.pdf
```

```
In [5]: df=pd.read_csv("prac7.csv")
```

```
In [6]: df
```

Out[6]:

	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.

In [27]:
```python
df=pd.get_dummies(df,columns=['Gender','Pass_Division'])
df
```

Out[27]:

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

In [ ]:

In [ ]:

In [ ]:

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

In [12]: `new_order = ['January', 'February', 'March', 'April', 'May', 'June', 'Ju`

In [28]: `df2=df`

In [29]: `df2['Birth_Month']=pd.Categorical(df['Birth_Month'], categories=new_orde`

In [30]: `df2.sort_values(by='Birth_Month')`

Out[30]:

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

In [ ]: