In [1]:
```
import pandas as pd
import numpy as np
df1 = {
'Name':['George','Andrea','micheal','maggie','Ravi','Xien','Jalpa',np.nan],
'State':['Arizona','Georgia','Newyork','Indiana','Florida','California',np.nan,np.na
'Gender':["M","F","M","F","M","M",np.nan,np.nan],
'Score':[63,48,56,75,np.nan,77,np.nan,np.nan]
}
df1 = pd.DataFrame(df1,columns=['Name','State','Gender','Score'])
print(df1)
```

```
      Name       State Gender  Score
0   George     Arizona      M   63.0
1   Andrea     Georgia      F   48.0
2  micheal     Newyork      M   56.0
3   maggie     Indiana      F   75.0
4     Ravi     Florida      M    NaN
5     Xien  California      M   77.0
6    Jalpa         NaN    NaN    NaN
7      NaN         NaN    NaN    NaN
```

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

In [3]:
```
# Assign data
data = {'Name': ['Jai', 'Princi', 'Gaurav',
                 'Anuj', 'Ravi', 'Natasha', 'Riya'],
        'Age': [17, 17, 18, 17, 18, 17, 17],
        'Gender': ['M', 'F', 'M', 'M', 'M', 'F', 'F'],
        'Marks': [90, 76, 'NaN', 74, 65, 'NaN', 71]}
```

In [4]:
```
# Import pandas package
import pandas as pd

# Assign data
data = {'Name': ['Jai', 'Princi', 'Gaurav',
                 'Anuj', 'Ravi', 'Natasha', 'Riya'],
        'Age': [17, 17, 18, 17, 18, 17, 17],
        'Gender': ['M', 'F', 'M', 'M', 'M', 'F', 'F'],
        'Marks': [90, 76, 'NaN', 74, 65, 'NaN', 71]}

# Convert into DataFrame
df = pd.DataFrame(data)

# Display data
df
```

Out[4]:

	Name	Age	Gender	Marks
0	Jai	17	M	90
1	Princi	17	F	76
2	Gaurav	18	M	NaN
3	Anuj	17	M	74
4	Ravi	18	M	65
5	Natasha	17	F	NaN

	Name	Age	Gender	Marks
6	Riya	17	F	71

In [5]:
```
# Compute average
c = avg = 0
for ele in df['Marks']:
# Compute average
    c = avg = 0
    for ele in df['Marks']:
        if str(ele).isnumeric():
            c += 1
            avg += ele
avg /= c

# Replace missing values
df = df.replace(to_replace="NaN",value=avg)

# Display data
df
```

Out[5]:

	Name	Age	Gender	Marks
0	Jai	17	M	90.0
1	Princi	17	F	76.0
2	Gaurav	18	M	75.2
3	Anuj	17	M	74.0
4	Ravi	18	M	65.0
5	Natasha	17	F	75.2
6	Riya	17	F	71.0

In [6]:
```
# Categorize gender
df['Gender'] = df['Gender'].map({'M': 0,
                                 'F': 1, }).astype(float)
 # Display data
df
```

Out[6]:

	Name	Age	Gender	Marks
0	Jai	17	0.0	90.0
1	Princi	17	1.0	76.0
2	Gaurav	18	0.0	75.2
3	Anuj	17	0.0	74.0
4	Ravi	18	0.0	65.0
5	Natasha	17	1.0	75.2
6	Riya	17	1.0	71.0

In [7]:
```
# Filter top scoring students
df = df[df['Marks'] >= 75]
# Remove age row
```

```
df = df.drop(['Age'], axis=1)
 # Display data
df
```

Out[7]:

	Name	Gender	Marks
0	Jai	0.0	90.0
1	Princi	1.0	76.0
2	Gaurav	0.0	75.2
5	Natasha	1.0	75.2

In [8]:

```
# import module
import pandas as pd
 # creating DataFrame for Student Details
details = pd.DataFrame({
    'ID': [101, 102, 103, 104, 105, 106,
           107, 108, 109, 110],
    'NAME': ['Jagroop', 'Praveen', 'Harjot',
             'Pooja', 'Rahul', 'Nikita',
             'Saurabh', 'Ayush', 'Dolly', "Mohit"],
    'BRANCH': ['CSE', 'CSE', 'CSE', 'CSE', 'CSE',
               'CSE', 'CSE', 'CSE', 'CSE', 'CSE']})

# printing details
print(details)
```

```
    ID     NAME BRANCH
0  101  Jagroop    CSE
1  102  Praveen    CSE
2  103   Harjot    CSE
3  104    Pooja    CSE
4  105    Rahul    CSE
5  106   Nikita    CSE
6  107  Saurabh    CSE
7  108    Ayush    CSE
8  109    Dolly    CSE
9  110    Mohit    CSE
```

In [9]:

```
# Import module
import pandas as pd
# Creating Dataframe for Fees_Status
fees_status = pd.DataFrame(
    {'ID': [101, 102, 103, 104, 105,
            106, 107, 108, 109, 110],
     'PENDING': ['5000', '250', 'NIL',
                 '9000', '15000', 'NIL',
                 '4500', '1800', '250', 'NIL']})
# Printing fees_status
print(fees_status)
```

```
    ID PENDING
0  101    5000
1  102     250
2  103     NIL
3  104    9000
4  105   15000
5  106     NIL
6  107    4500
7  108    1800
```

```
8  109     250
9  110     NIL
```

In [10]:

```
# Import module
import pandas as pd
 # Creating Dataframe
details = pd.DataFrame({
    'ID': [101, 102, 103, 104, 105,
           106, 107, 108, 109, 110],
    'NAME': ['Jagroop', 'Praveen', 'Harjot',
             'Pooja', 'Rahul', 'Nikita',
             'Saurabh', 'Ayush', 'Dolly', "Mohit"],
    'BRANCH': ['CSE', 'CSE', 'CSE', 'CSE', 'CSE',
               'CSE', 'CSE', 'CSE', 'CSE', 'CSE']})
 # Creating Dataframe
fees_status = pd.DataFrame(
    {'ID': [101, 102, 103, 104, 105,
            106, 107, 108, 109, 110],
     'PENDING': ['5000', '250', 'NIL',
                 '9000', '15000', 'NIL',
                 '4500', '1800', '250', 'NIL']})
# Merging Dataframe
print(pd.merge(details, fees_status, on='ID'))
```

```
    ID     NAME BRANCH PENDING
0  101  Jagroop    CSE    5000
1  102  Praveen    CSE     250
2  103   Harjot    CSE     NIL
3  104    Pooja    CSE    9000
4  105    Rahul    CSE   15000
5  106   Nikita    CSE     NIL
6  107  Saurabh    CSE    4500
7  108    Ayush    CSE    1800
8  109    Dolly    CSE     250
9  110    Mohit    CSE     NIL
```

In [11]:

```
# Import module
import pandas as pd
# Creating Data
car_selling_data = {'Brand': ['Maruti', 'Maruti', 'Maruti',
                              'Maruti', 'Hyundai', 'Hyundai',
                              'Toyota', 'Mahindra', 'Mahindra',
                              'Ford', 'Toyota', 'Ford'],
                    'Year':  [2010, 2011, 2009, 2013,
                              2010, 2011, 2011, 2010,
                              2013, 2010, 2010, 2011],
                    'Sold': [6, 7, 9, 8, 3, 5,
                             2, 8, 7, 2, 4, 2]}
 # Creating Dataframe of car_selling_data
df = pd.DataFrame(car_selling_data)
 # printing Dataframe
print(df)
```

```
       Brand  Year  Sold
0     Maruti  2010     6
1     Maruti  2011     7
2     Maruti  2009     9
3     Maruti  2013     8
4    Hyundai  2010     3
5    Hyundai  2011     5
6     Toyota  2011     2
7   Mahindra  2010     8
8   Mahindra  2013     7
```

```
9       Ford  2010     2
10    Toyota  2010     4
11      Ford  2011     2
```

In [12]:

```
# Import module
import pandas as pd
# Creating Data
car_selling_data = {'Brand': ['Maruti', 'Maruti', 'Maruti',
                              'Maruti', 'Hyundai', 'Hyundai',
                              'Toyota', 'Mahindra', 'Mahindra',
                              'Ford', 'Toyota', 'Ford'],
                    'Year':  [2010, 2011, 2009, 2013,
                              2010, 2011, 2011, 2010,
                              2013, 2010, 2010, 2011],
                    'Sold': [6, 7, 9, 8, 3, 5,
                             2, 8, 7, 2, 4, 2]}
 # Creating Dataframe for Provided Data
df = pd.DataFrame(car_selling_data)
 # Group the data when year = 2010
grouped = df.groupby('Year')
print(grouped.get_group(2010))
```

```
       Brand  Year  Sold
0     Maruti  2010     6
4    Hyundai  2010     3
7   Mahindra  2010     8
9       Ford  2010     2
10    Toyota  2010     4
```

DETAILS STUDENTS DATA WHO WANT TO PARTICIPATE IN THE EVENT:

In [13]:

```
# Import module
import pandas as pd
# Initializing Data
student_data = {'Name': ['Amit', 'Praveen', 'Jagroop',
                         'Rahul', 'Vishal', 'Suraj',
                         'Rishab', 'Satyapal', 'Amit',
                         'Rahul', 'Praveen', 'Amit'],
                'Roll_no': [23, 54, 29, 36, 59, 38,
                            12, 45, 34, 36, 54, 23],
                'Email': ['xxxx@gmail.com', 'xxxxxx@gmail.com',
                          'xxxxxx@gmail.com', 'xx@gmail.com',
                          'xxxx@gmail.com', 'xxxxx@gmail.com',
                          'xxxxx@gmail.com', 'xxxxx@gmail.com',
                          'xxxxx@gmail.com', 'xxxxxx@gmail.com',
                          'xxxxxxxxxx@gmail.com', 'xxxxxxxxxx@gmail.com']}
 # Creating Dataframe of Data
df = pd.DataFrame(student_data)
 # Printing Dataframe
print(df)
```

```
        Name  Roll_no                Email
0       Amit       23        xxxx@gmail.com
1    Praveen       54      xxxxxx@gmail.com
2    Jagroop       29      xxxxxx@gmail.com
3      Rahul       36          xx@gmail.com
4     Vishal       59        xxxx@gmail.com
5      Suraj       38       xxxxx@gmail.com
6     Rishab       12       xxxxx@gmail.com
```

```
7   Satyapal       45       xxxxx@gmail.com
8       Amit       34       xxxxx@gmail.com
9      Rahul       36      xxxxxx@gmail.com
10   Praveen       54  xxxxxxxxxx@gmail.com
11      Amit       23  xxxxxxxxxx@gmail.com
```

In [14]:

```
# Import module
import pandas as pd
# Initializing Data
student_data = {'Name': ['Amit', 'Praveen', 'Jagroop',
                         'Rahul', 'Vishal', 'Suraj',
                         'Rishab', 'Satyapal', 'Amit',
                         'Rahul', 'Praveen', 'Amit'],
                'Roll_no': [23, 54, 29, 36, 59, 38,
                            12, 45, 34, 36, 54, 23],
                'Email': ['xxxx@gmail.com', 'xxxxxx@gmail.com',
                          'xxxxxx@gmail.com', 'xx@gmail.com',
                          'xxxx@gmail.com', 'xxxxx@gmail.com',
                          'xxxxx@gmail.com', 'xxxxx@gmail.com',
                          'xxxxx@gmail.com', 'xxxxxx@gmail.com',
                          'xxxxxxxxxx@gmail.com', 'xxxxxxxxxx@gmail.com']}
 # Creating Dataframe of Data
df = pd.DataFrame(student_data)
 # Printing Dataframe
print(df)
```

```
        Name  Roll_no                Email
0       Amit       23        xxxx@gmail.com
1    Praveen       54      xxxxxx@gmail.com
2    Jagroop       29      xxxxxx@gmail.com
3      Rahul       36          xx@gmail.com
4     Vishal       59        xxxx@gmail.com
5      Suraj       38       xxxxx@gmail.com
6     Rishab       12       xxxxx@gmail.com
7   Satyapal       45       xxxxx@gmail.com
8       Amit       34       xxxxx@gmail.com
9      Rahul       36      xxxxxx@gmail.com
10   Praveen       54  xxxxxxxxxx@gmail.com
11      Amit       23  xxxxxxxxxx@gmail.com
```

In [15]:

```
# import module
import pandas as pd
# initializing Data
student_data = {'Name': ['Amit', 'Praveen', 'Jagroop',
                         'Rahul', 'Vishal', 'Suraj',
                         'Rishab', 'Satyapal', 'Amit',
                         'Rahul', 'Praveen', 'Amit'],
                'Roll_no': [23, 54, 29, 36, 59, 38,
                            12, 45, 34, 36, 54, 23],
                'Email': ['xxxx@gmail.com', 'xxxxxx@gmail.com',
                          'xxxxxx@gmail.com', 'xx@gmail.com',
                          'xxxx@gmail.com', 'xxxxx@gmail.com',
                          'xxxxx@gmail.com', 'xxxxx@gmail.com',
                          'xxxxx@gmail.com', 'xxxxxx@gmail.com',
                          'xxxxxxxxxx@gmail.com', 'xxxxxxxxxx@gmail.com']}
 # creating dataframe
df = pd.DataFrame(student_data)
 # Here df.duplicated() list duplicate  Entries in ROllno.
# So that ~(NOT) is placed in order to get non duplicate values.
non_duplicate = df[~df.duplicated('Roll_no')]
 # printing non-duplicate values
print(non_duplicate)
```

	Name	Roll_no	Email
0	Amit	23	xxxx@gmail.com
1	Praveen	54	xxxxxx@gmail.com
2	Jagroop	29	xxxxxx@gmail.com
3	Rahul	36	xx@gmail.com
4	Vishal	59	xxxx@gmail.com
5	Suraj	38	xxxxx@gmail.com
6	Rishab	12	xxxxx@gmail.com
7	Satyapal	45	xxxxx@gmail.com
8	Amit	34	xxxxx@gmail.com

In [ ]: