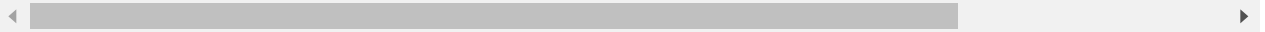


Read the dataset from the below link

https://raw.githubusercontent.com/guipsamora/pandas_exercises/master/06_Stats/US_Baby_Names/
[\(https://raw.githubusercontent.com/guipsamora/pandas_exercises/master/06_Stats/US_Baby_Names](https://raw.githubusercontent.com/guipsamora/pandas_exercises/master/06_Stats/US_Baby_Names/)



```
In [18]: import numpy as np
import pandas as pd
```

```
In [33]: url='https://raw.githubusercontent.com/guipsamora/pandas_exercises/master/06_Stats/US_Baby_Names/'
df = pd.read_csv(url)
```

1. Delete unnamed columns

```
In [ ]: df.head()
```

```
In [36]: mydf=df.drop('Unnamed: 0',axis=1)
```

```
In [37]: mydf.head()
```

Out[37]:

	Id	Name	Year	Gender	State	Count
0	11350	Emma	2004	F	AK	62
1	11351	Madison	2004	F	AK	48
2	11352	Hannah	2004	F	AK	46
3	11353	Grace	2004	F	AK	44
4	11354	Emily	2004	F	AK	41

2. Show the distribution of male and female

```
In [44]: len(mydf)
```

Out[44]: 1016395

```
In [54]: len(mydf[mydf['Gender']=='M'])
```

Out[54]: 457549

```
In [53]: len(mydf[mydf['Gender']=='F'])
```

Out[53]: 558846

```
In [59]: mydf['Gender'].value_counts()
```

```
Out[59]: F      558846  
        M      457549  
        Name: Gender, dtype: int64
```

3. Show the top 5 most preferred names

```
In [49]: mydf['Name'].value_counts().head()
```

```
Out[49]: Riley      1112  
        Avery      1080  
        Jordan     1073  
        Peyton     1064  
        Hayden     1049  
        Name: Name, dtype: int64
```

4. What is the median name occurrence in the dataset

```
In [ ]: mydf.head()
```

5. Distribution of male and female born count by states

```
In [104]: mydf.groupby(['State', 'Gender']).size().head(10)
```

```
Out[104]: State  Gender  
        AK      F      2404  
           M      2587  
        AL      F      9878  
           M      8419  
        AR      F      7171  
           M      6475  
        AZ      F     14518  
           M     10820  
        CA      F     45144  
           M     31637  
        dtype: int64
```

```
In [ ]:
```