In [6]:

```python
import pandas as pd
import numpy as np

# Reading the dataset from the given link
dataset=pd.read_csv("https://raw.githubusercontent.com/guipsamora/pandas_exercises/master/06_Stats/US_Baby_Names/US_Baby_Names_right.csv")
```

In [7]:

```python
#Printing the head of the dataset before deleting the unnamed columns
dataset.head()
```

Out[7]:

	Unnamed: 0	Id	Name	Year	Gender	State	Count
0	11349	11350	Emma	2004	F	AK	62
1	11350	11351	Madison	2004	F	AK	48
2	11351	11352	Hannah	2004	F	AK	46
3	11352	11353	Grace	2004	F	AK	44
4	11353	11354	Emily	2004	F	AK	41

In [8]:

```python
# Deleting the unnamed columns.
dataset.drop(labels = "Unnamed: 0", axis=1, inplace=True)
```

In [9]:

```python
#Printing the head of the dataset after deleting the unnamed columns
dataset.head()
```

Out[9]:

	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

In [10]:

```
#Showing the distribution of male and female.
dataset.Gender.value_counts()
```

Out[10]:

```
F    558846
M    457549
Name: Gender, dtype: int64
```

In [11]:

```
#Top 5 most preferred names.
dataset.Name.value_counts()[0:5]
```

Out[11]:

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

In [31]:

```
#Median name occurance in the dataset
medianNameId = dataset['Id'].median()
medianName = dataset[dataset['Id'] == medianNameId]['Name']
print("Median name at index {} and is {}".format(medianNameId, medianName))
```

```
Median name at index 2811921.0 and is 508197    Kasey
Name: Name, dtype: object
```

In [33]:

```python
#Distribution of male and female born count by states
ds_by_states = dataset.groupby('State')
print("Distribution of male and female born count by states")
ds_by_states.Gender.value_counts()
```

Distribution of male and female born count by states

Out[33]:

```
State  Gender
AK     M         2587
       F         2404
AL     F         9878
       M         8419
AR     F         7171
       M         6475
AZ     F        14518
       M        10820
CA     F        45144
       M        31637
CO     F        11424
       M         9183
CT     F         6575
       M         5733
DC     F         3053
       M         3000
DE     F         2549
       M         2440
FL     F        25781
       M        20070
GA     F        19385
       M        15454
HI     M         3546
       F         3255
IA     F         7131
       M         6307
ID     F         4918
       M         4833
IL     F        21268
       M        16828
                ...
OK     F         9519
       M         8138
OR     F         8604
       M         7333
PA     F        17480
       M        14171
RI     F         2558
       M         2468
SC     F         9465
       M         8195
SD     M         2908
       F         2838
TN     F        13063
       M        10588
TX     F        39760
       M        27791
UT     F         9515
       M         8233
VA     F        14759
       M        11997
VT     M         1618
       F         1398
WA     F        13329
       M        11049
WI     F        10549
       M         8940
WV     F         4305
```

```
        M            3733
WY      M            1904
        F            1456
Name: Gender, Length: 102, dtype: int64
```

In [ ]: