

```
In [1]: import pandas as pd

Baby_Names = pd.read_csv('https://raw.githubusercontent.com/guipsamora/pandas_exercises/master/01_Pandas_Basics/01_Popular_Baby_Names.csv')
Baby_Names.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1016395 entries, 0 to 1016394
Data columns (total 7 columns):
Unnamed: 0    1016395 non-null int64
Id            1016395 non-null int64
Name          1016395 non-null object
Year          1016395 non-null int64
Gender        1016395 non-null object
State         1016395 non-null object
Count         1016395 non-null int64
dtypes: int64(4), object(3)
memory usage: 54.3+ MB
```

```
In [2]: # Delete Unnamed columns.

del Baby_Names['Unnamed: 0']
del Baby_Names['Id']

Baby_Names.head()
```

Out[2]:

|   | Name    | Year | Gender | State | Count |
|---|---------|------|--------|-------|-------|
| 0 | Emma    | 2004 | F      | AK    | 62    |
| 1 | Madison | 2004 | F      | AK    | 48    |
| 2 | Hannah  | 2004 | F      | AK    | 46    |
| 3 | Grace   | 2004 | F      | AK    | 44    |
| 4 | Emily   | 2004 | F      | AK    | 41    |

```
In [27]: # Show the distribution of male and female

Baby_Names['Gender'].value_counts()
```

```
Out[27]: F    558846
         M    457549
         Name: Gender, dtype: int64
```

In [12]: *# What is the median name occurrence in the dataset*

```
names[names.Count == names.Count.median()]
```

Out[12]:

|                  | Year  | Count |
|------------------|-------|-------|
| Name             |       |       |
| <b>Aishani</b>   | 14078 | 49    |
| <b>Alara</b>     | 16079 | 49    |
| <b>Alysse</b>    | 16057 | 49    |
| <b>Ameir</b>     | 16086 | 49    |
| <b>Anely</b>     | 16071 | 49    |
| <b>Antonina</b>  | 18081 | 49    |
| <b>Aveline</b>   | 12065 | 49    |
| <b>Aziah</b>     | 16073 | 49    |
| <b>Baily</b>     | 16064 | 49    |
| <b>Caleah</b>    | 18106 | 49    |
| <b>Carlota</b>   | 14077 | 49    |
| <b>Cristine</b>  | 14042 | 49    |
| <b>Dahlila</b>   | 14063 | 49    |
| <b>Darvin</b>    | 16078 | 49    |
| <b>Deante</b>    | 18064 | 49    |
| <b>Deserae</b>   | 18061 | 49    |
| <b>Devean</b>    | 8019  | 49    |
| <b>Elizah</b>    | 16063 | 49    |
| <b>Emmaly</b>    | 16075 | 49    |
| <b>Emmanuela</b> | 18074 | 49    |
| <b>Envy</b>      | 16070 | 49    |
| <b>Esli</b>      | 16059 | 49    |
| <b>Fay</b>       | 16072 | 49    |
| <b>Gurshaan</b>  | 14070 | 49    |
| <b>Hareem</b>    | 14090 | 49    |
| <b>Iven</b>      | 14062 | 49    |
| <b>Jaice</b>     | 16098 | 49    |
| <b>Jaiyana</b>   | 14068 | 49    |
| <b>Jamiracle</b> | 18091 | 49    |
| <b>Jelissa</b>   | 16070 | 49    |
| ...              | ...   | ...   |
| <b>Kyndle</b>    | 16082 | 49    |

|            | Year  | Count |
|------------|-------|-------|
| Name       |       |       |
| Kynsley    | 14084 | 49    |
| Leylanie   | 16070 | 49    |
| Maisha     | 14047 | 49    |
| Malillany  | 14087 | 49    |
| Mariann    | 16060 | 49    |
| Marquell   | 16053 | 49    |
| Maurilio   | 16058 | 49    |
| Mckynzie   | 14068 | 49    |
| Mehdi      | 16070 | 49    |
| Nabeel     | 18070 | 49    |
| Nalleli    | 16058 | 49    |
| Nassir     | 16058 | 49    |
| Nazier     | 16061 | 49    |
| Nishant    | 16050 | 49    |
| Rebecka    | 16061 | 49    |
| Reghan     | 18073 | 49    |
| Ridwan     | 16082 | 49    |
| Riot       | 16104 | 49    |
| Rubin      | 16055 | 49    |
| Ryatt      | 16103 | 49    |
| Sameera    | 18083 | 49    |
| Sanjuanita | 10035 | 49    |
| Shalyn     | 18061 | 49    |
| Skylie     | 16086 | 49    |
| Sriram     | 14054 | 49    |
| Trinton    | 16069 | 49    |
| Vita       | 14075 | 49    |
| Yoni       | 16060 | 49    |
| Zuleima    | 14050 | 49    |

66 rows × 2 columns

In [23]: *# 5. Distribution of male and female born count by states*

```
counts = Baby_Names.groupby(['Gender', 'State', 'Year']).count()
counts
```

Out[23]:

|        |       |      | Name | Count |
|--------|-------|------|------|-------|
| Gender | State | Year |      |       |
| F      | AK    | 2004 | 204  | 204   |
|        |       | 2005 | 214  | 214   |
|        |       | 2006 | 221  | 221   |
|        |       | 2007 | 199  | 199   |
|        |       | 2008 | 225  | 225   |
|        |       | 2009 | 212  | 212   |
|        |       | 2010 | 222  | 222   |
|        |       | 2011 | 235  | 235   |
|        |       | 2012 | 216  | 216   |
|        |       | 2013 | 225  | 225   |
|        |       | 2014 | 231  | 231   |
|        | AL    | 2004 | 834  | 834   |
|        |       | 2005 | 847  | 847   |
|        |       | 2006 | 899  | 899   |
|        |       | 2007 | 942  | 942   |
|        |       | 2008 | 937  | 937   |
|        |       | 2009 | 919  | 919   |
|        |       | 2010 | 918  | 918   |
|        |       | 2011 | 908  | 908   |
|        |       | 2012 | 869  | 869   |
|        |       | 2013 | 881  | 881   |
|        |       | 2014 | 924  | 924   |
|        | AR    | 2004 | 607  | 607   |
|        |       | 2005 | 634  | 634   |
|        |       | 2006 | 653  | 653   |
|        |       | 2007 | 657  | 657   |
|        |       | 2008 | 683  | 683   |
|        |       | 2009 | 659  | 659   |
| 2010   |       | 659  | 659  |       |
|        |       | 2011 | 645  | 645   |

|        |       |      | Name | Count |
|--------|-------|------|------|-------|
| Gender | State | Year |      |       |
| ...    | ...   | ...  | ...  | ...   |
| M      | WI    | 2007 | 816  | 816   |
|        |       | 2008 | 827  | 827   |
|        |       | 2009 | 823  | 823   |
|        |       | 2010 | 847  | 847   |
|        |       | 2011 | 836  | 836   |
|        |       | 2012 | 830  | 830   |
|        |       | 2013 | 852  | 852   |
|        |       | 2014 | 840  | 840   |
| WV     | WV    | 2004 | 311  | 311   |
|        |       | 2005 | 336  | 336   |
|        |       | 2006 | 321  | 321   |
|        |       | 2007 | 329  | 329   |
|        |       | 2008 | 333  | 333   |
|        |       | 2009 | 352  | 352   |
|        |       | 2010 | 330  | 330   |
|        |       | 2011 | 342  | 342   |
|        |       | 2012 | 352  | 352   |
|        |       | 2013 | 362  | 362   |
|        |       | 2014 | 365  | 365   |
| WY     | WY    | 2004 | 153  | 153   |
|        |       | 2005 | 171  | 171   |
|        |       | 2006 | 181  | 181   |
|        |       | 2007 | 177  | 177   |
|        |       | 2008 | 177  | 177   |
|        |       | 2009 | 183  | 183   |
|        |       | 2010 | 174  | 174   |
|        |       | 2011 | 166  | 166   |
|        |       | 2012 | 163  | 163   |
|        |       | 2013 | 178  | 178   |
|        |       | 2014 | 181  | 181   |

1122 rows × 2 columns

In [ ]:

