

## ▼ Wordclouds with stylecloud

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from wordcloud import WordCloud, STOPWORDS, ImageColorGenerator
!pip install stylecloud
import stylecloud
```

```
names = pd.read_csv("https://raw.githubusercontent.com/fivethirtyeight/data/master/
names.head()
```



	name	total	male_share	female_share	gap
1	Casey	176544.328149	0.584287	0.415713	0.168573
2	Riley	154860.665173	0.507639	0.492361	0.015278
3	Jessie	136381.830656	0.477834	0.522166	0.044331
4	Jackie	132928.788740	0.421133	0.578867	0.157735
5	Avery	121797.419516	0.335213	0.664787	0.329574

```
names.tail()
```

	name	total	male_share	female_share	gap
915	Eaden	102.319579	0.572761	0.427239	0.145521
916	Inioluwa	101.384201	0.352729	0.647271	0.294542
917	Gwin	101.243364	0.562137	0.437863	0.124273
918	Yacine	100.230400	0.544599	0.455401	0.089198
919	Aeon	100.211040	0.464835	0.535165	0.070331

```
names.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 919 entries, 1 to 919
Data columns (total 5 columns):
#   Column          Non-Null Count  Dtype
---  -
0   name            919 non-null    object
1   total           919 non-null    float64
2   male_share      919 non-null    float64
3   female_share    919 non-null    float64
4   gap             919 non-null    float64
dtypes: float64(4), object(1)
memory usage: 43.1+ KB
```

```
female = names.query("female_share > male_share")
female
```

	name	total	male_share	female_share	gap
3	Jessie	136381.830656	0.477834	0.522166	0.044331
4	Jackie	132928.788740	0.421133	0.578867	0.157735
5	Avery	121797.419516	0.335213	0.664787	0.329574
7	Peyton	94896.395216	0.433719	0.566281	0.132561
8	Kerry	88963.926250	0.483949	0.516051	0.032102
...	...	...	...	...	...
910	Eudell	103.921273	0.419673	0.580327	0.160655
911	Mikele	103.518292	0.394806	0.605194	0.210388
914	Olayinka	102.625879	0.418250	0.581750	0.163500
916	Inioluwa	101.384201	0.352729	0.647271	0.294542
919	Aeon	100.211040	0.464835	0.535165	0.070331

428 rows × 5 columns

```
male = names.query("male_share > female_share")
male
```

	name	total	male_share	female_share	gap
1	Casey	176544.328149	0.584287	0.415713	0.168573
2	Riley	154860.665173	0.507639	0.492361	0.015278
6	Jaime	109870.187290	0.561793	0.438207	0.123586
12	Skyler	53486.390419	0.646053	0.353947	0.292106
13	Frankie	51288.068109	0.623671	0.376329	0.247343
...	...	...	...	...	...
912	Alija	103.229976	0.576972	0.423028	0.153945
913	Carel	102.758657	0.564701	0.435299	0.129403
915	Eaden	102.319579	0.572761	0.427239	0.145521
917	Gwin	101.243364	0.562137	0.437863	0.124273
918	Yacine	100.230400	0.544599	0.455401	0.089198

491 rows × 5 columns

```
text = " ".join(name for name in female.name)
print("There are {} names that are used more for females than for males".format(len(text)))
text_male = " ".join(name for name in male.name)
print("There are {} names that are used more for males than for females".format(len(text_male)))
```

There are 2876 names that are used more for females than for males  
There are 3206 names that are used more for males than for females

[illegible]

<https://colab.research.google.com/drive/1q5LavX4U8pvzswbTxT0Na3pguHoxOLmP#scrollTo=A9R2HWB3WJTV&printMode=true>

