

Data Appendix

Dataset: “baby-names-state.csv”

Unit of Observation: Number of babies born with a name in a specific state for a specific year

Data Dictionary:

Column	Description	Potential Response	# of Obs.	Summary Statistics
state_abb	A string referring to the state the babies with that name were born in.	A string of two characters such as AK for Arkansas with 51 possible values for all states + DC	6,215,834	50 states + DC
sex	A binary variable representing the sex of the babies born with a specific name	Either a F representing female babies, or a M representing male babies	6,215,834	F: 3,459,588 M: 2,756,246
year	The year the group of babies born with a specific name were born in	A numeric value ranging from 1910 to 2020	6,215,834	Min: 1910 1st Qu: 1953 Median: 1982 Mean: 1977 3rd Qu: 2004 Max: 2020
name	The name the group of babies born share in a given year	A string of text which will vary in length based on the name chosen, ex. Mary or Margaret	6,215,834	Unique
count	The number of babies born in a specific year and state with the same name	An integer ranging from 1 to 10,025	6,215,834	Min: 5.00 1st Qu: 7.00 Median: 12.00 Mean: 50.98 3rd Qu: 33.0 Max: 10025.00

References

Flowers, A. (2017, September 28). The most common unisex names in America: Is yours one of them?

FiveThirtyEight.

<https://fivethirtyeight.com/features/there-are-922-unisex-names-in-america-is-yours-one-of-them/>

Mekouar, D. (2024, February 23). Gender-Neutral baby names gain popularity, but traditional names still rule. *Voice of America*.

<https://www.voanews.com/a/gender-neutral-baby-names-gain-popularity-but-traditional-names-still-rule/7499777.html#>