

The Normalist Name

Our dataset lets anyone interested in names, writers, and anyone with a name see how often a certain name appears, similar names, and popularity trends about a given name.

The Dataset:

Location and Rights:

The data is obtained from the Social Security Administration at:
<https://www.ssa.gov/oact/babynames/limits.html> - Initially downloaded January 17 2024.

This information in the dataset is sourced from The SSA databases as of March 2023. It covers names from 1880 - 2022, in the states and territories of the United States of America. This specific dataset does not come with a readme or specific rights. It represents ~70% of the population as names with less than 5 occurrences in a year are hidden for privacy reasons.

CC0 - Available for public use, as said in:
<https://catalog.data.gov/dataset/baby-names-from-social-security-card-applications-national-data>

Data Layout:

The data is in 142 text files separated by year, with comma separated fields and a newline for each name. It is ordered by number of occurrences, and in the case of equal occurrences, alphabetically.

Name, Sex, Number of Occurrences

Questions it answers:

1. How popular is my name? - This takes the |Name| and |Number of Occurrences| and could use the year.
2. What is a random 90's style name? - year, |Name|, |Number of Occurrences|
3. What name was really popular but fell out of fashion? - again, year, |Name|, |Number of Occurrences|
4. Was at least 5 people ever named "Blert" in one year? - This could be answered with just |Name|
5. What is the average age of someone with this name? - year, |Name|, |Occurrences|
6. How often is this name given to Males and Females? - |Name|, |Sex|, |Occurrences|, and could use year

Audience:

Expecting **Parents** - Questions: 1 (popularity), 4 (Do people have this name), 5 (Age of name)

-> *To find a unique name that is not too uncommon, and does not sound "old."*

Student confirming gender identity- 1 (popularity), 4 (Do people have this name), 6 (Name's sex)

-> *To find a name that is not primarily masculine or feminine, and is common.*

Stumped **Writer** - 2 (name from time), 3 (dated name), 5 (age of name)

-> *To get era specific names and test if planned names are suitable for a given time.*