# Instructions

### 1 Data Source:

The dataset comes from Data.gov (<https://catalog.data.gov/dataset/popular-baby-names>) and includes ranking data for baby names in the United States from 2011 to 2018. The dataset includes variables with column names such as "Year of Birth, Gender, Ethnicity, Child's First Name, Count, Rank" that can be categorized by gender and ethnicity. The data is collected through civil birth registration, and each record represents the ranking of baby names arranged by frequency.

### 2 Dataset Overview:

The dataset includes ranking data for baby names in the United States from 2011 to 2018, including the number of occurrences and rankings for each name. The dataset can be used to understand the trends and cultural backgrounds of baby names in the United States. The variables in the dataset include gender, ethnicity, name, count, ranking, as well as the number of occurrences and rankings for each name.

### 3 Analysis Directions:

1. Explore the popular baby name trends for different genders and ethnicities: By analyzing variables such as year, gender, ethnicity, name, and count in the dataset, the most popular baby names can be identified each year, as well as their usage among different genders and ethnicities. Comparing trends in different years and between different genders/ethnicities can reveal cultural and societal trends in baby names.
2. Baby name diversity: In the dataset, each name has a ranking. The diversity of baby names can be understood by calculating the number of occurrences for each name and the number of different names. Comparing the diversity of names among different genders and ethnicities can reveal cultural differences and societal trends in baby names.
3. Cultural background of baby names: By analyzing variables such as ethnicity and name in the dataset, the differences in baby name choices among different ethnicities can be explored. For example, the cultural and historical backgrounds of baby names among Latinx, Asian, and African American populations can be analyzed.
4. Length and syllable structure of baby names: By analyzing features such as length and syllable structure of names, the language characteristics of baby names in different genders and ethnicities can be understood. Comparing differences between different genders and ethnicities can explore the influence of different languages and cultural backgrounds on baby names.

### 4 Possible Outcomes:

By analyzing the "Popular Baby Names" dataset, the most popular baby names among different genders and ethnicities can be determined. The trends in different years and between different genders/ethnicities can be understood, as well as the cultural and historical backgrounds of baby name choices among different ethnicities. Additionally, the language characteristics of baby names in different genders and ethnicities can be explored, revealing the influence of different languages and cultural backgrounds on baby names. Through these analyses, a deeper understanding of the cultural background and societal trends of baby names in the United States can be gained, providing valuable insights into understanding American culture and social phenomena.

### 5 Analysis Tools:

Various data analysis tools can be used to analyze the "Popular Baby Names" dataset, such as Python, R, and Excel. In Python, libraries such as pandas, numpy, and matplotlib can be used to import the dataset, perform data cleaning and processing, and visualize the results. In R, libraries such as tidyverse and ggplot2 can be used for the same analysis. In Excel, functions such as pivot tables and charts can be used for data analysis.

### 6 Conclusion:

Through the analysis of the "Popular Baby Names" dataset, the trends and cultural backgrounds of baby names in the United States can be understood, revealing differences and cultural influences among different genders and ethnicities. These analysis results have significant reference value for understanding American culture and social phenomena, and can provide useful suggestions for future baby name choices.