

Assignment 2: Exploring the Baby Names Dataset

For this assignment I chose to explore the [baby names dataset](#) which contains information statistics relating to baby names. This dataset contains baby name counts on a country and state-by-state level. The notable features of this dataset include baby names, the number of babies named each name (split by country and state-level), organized by years from 1880-2014.

I conducted a two part investigation into this dataset.

I began this project by exploring some specific questions about my name and the common names given to babies around the year of my birth:

1. In which states are the most Taylors born?
2. What were the most common names given to babies born in California in the year 1993?
3. How has the popularity of the name Taylor changed over time in the United States from 1950-2014?

After discovering information about my name I then decided to perform some additional analysis to determine:

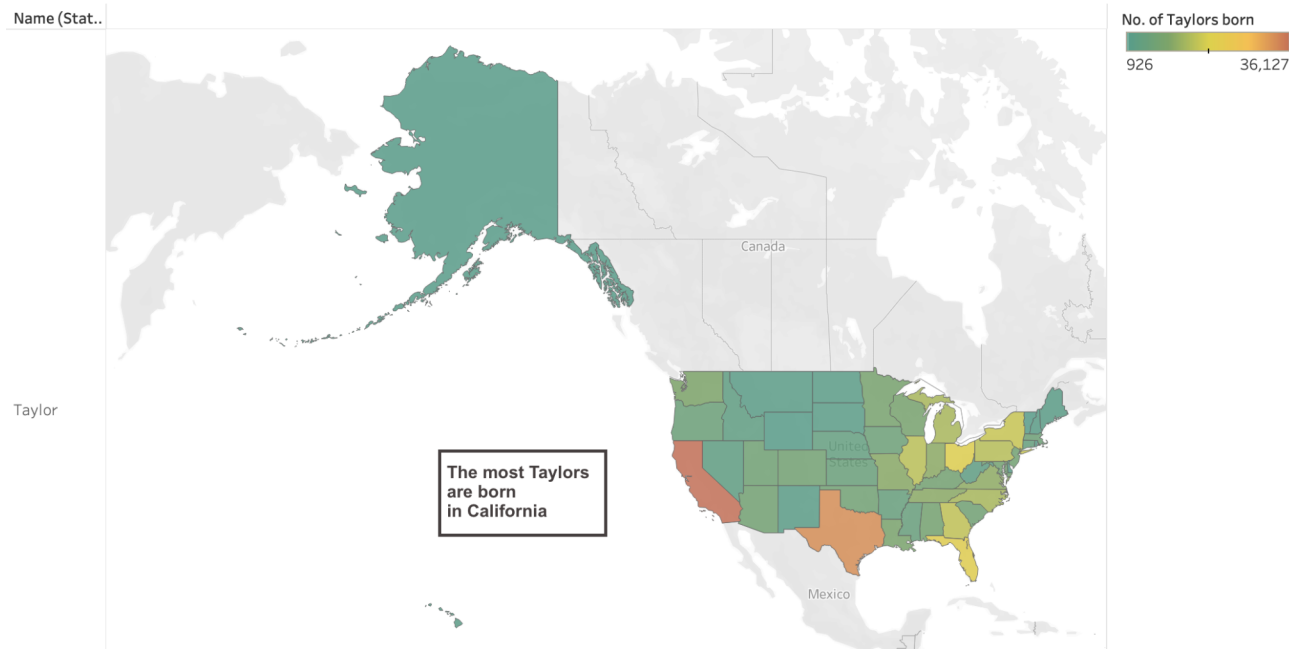
4. In what year were the greatest number of unique names given to children born in the United States? Which gender has the most unique names and how this has changed over time?
5. How have popular movie and book series that I read or watched during my life impacted the number of babies named after these movie/book characters?

Before addressing my research questions I first looked through the data to see if any cleaning needed to be done. I mostly found the dataset to be complete minus a few null values. I removed the incomplete data from the dataset for analysis.

RQ1: In which states are the most Taylors born?

As a child, my parents always told me that they decided to name me Taylor because it wasn't a very popular name (especially for girls) at the time of my birth. So I chose to investigate the statistics to see how much truth there was behind their explanation. I began by plotting the number of Taylors born in each state from 1950-2014. From the resulting figure, the most Taylors are born in California, followed by Texas, with Ohio having the third highest concentration of babies being named Taylor.

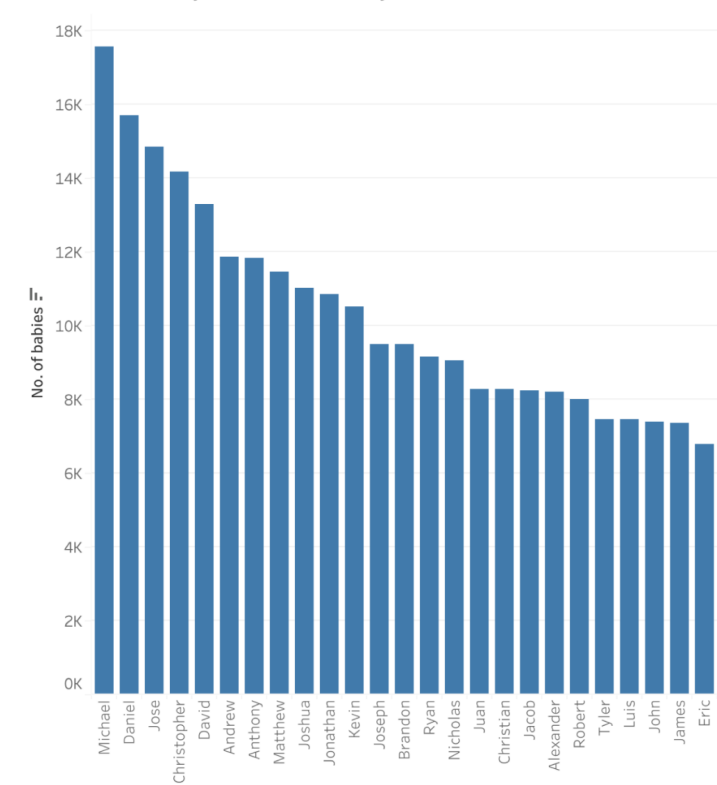
State by State Popularity of the name Taylor



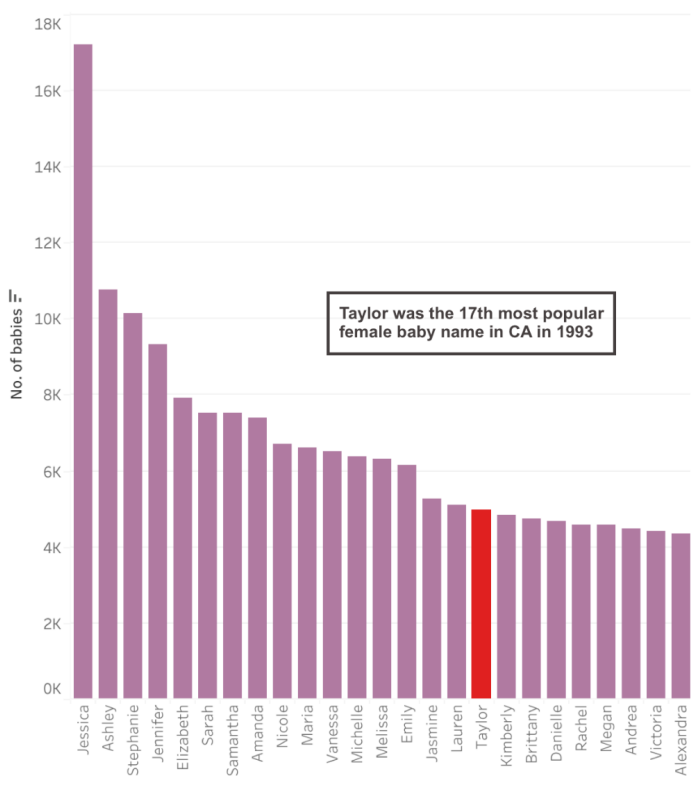
RQ2: What were the most common names given to babies born in California in the year 1993?

I chose to perform a general search to see what the most popular baby names were in the state and year in which I was born. I began by filtering the data for data specific to the state of California and only for the year of 1993. Low and behold, it seems that Taylor was the 17th most popular name for baby girls.

Most commonly used male baby names in CA 1993



Most commonly used female baby names in CA 1993

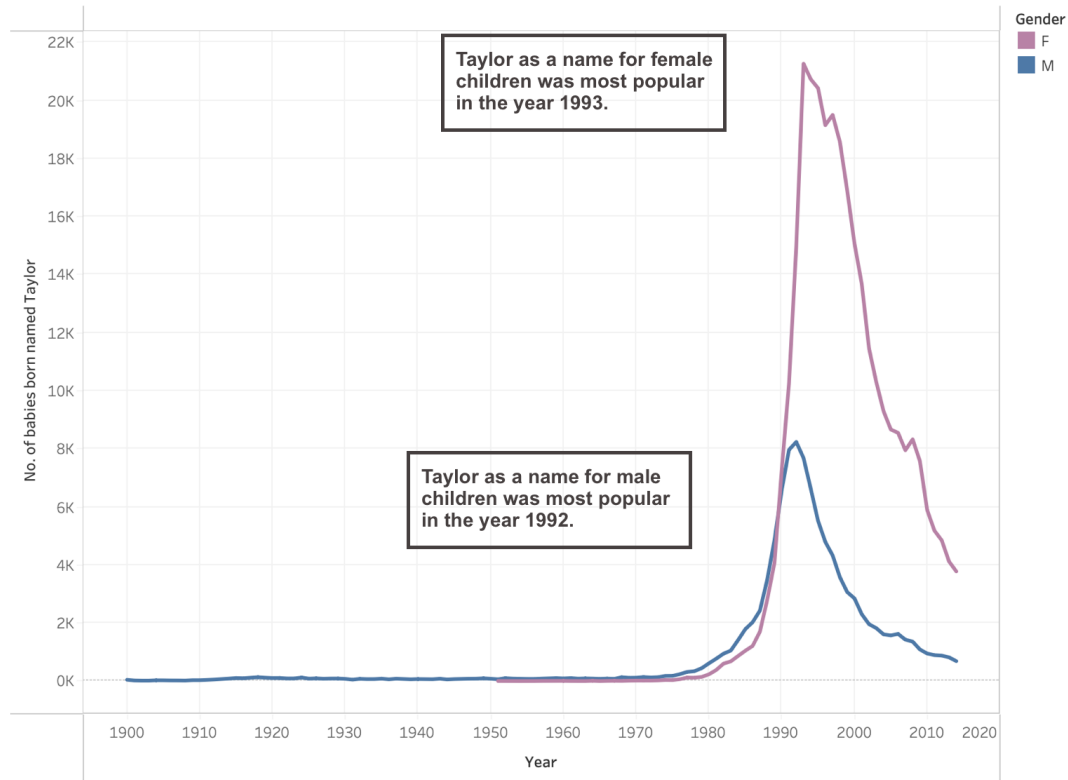


Taylor was the 17th most popular female baby name in CA in 1993

RQ3: How has the popularity of the name Taylor changed over time in the United States from 1900-2014?

I then explored how the trend of naming your child Taylor changed over time across genders. From this analysis we can see that both male and female popularity was similar until around 1990 when the number of baby girls being named Taylor surpassed the number of baby boys with the name. Looking specifically at my year of birth (1993); the most Taylors were born then with two times as many girls being named Taylor than boys.

Number of babies named Taylor by gender

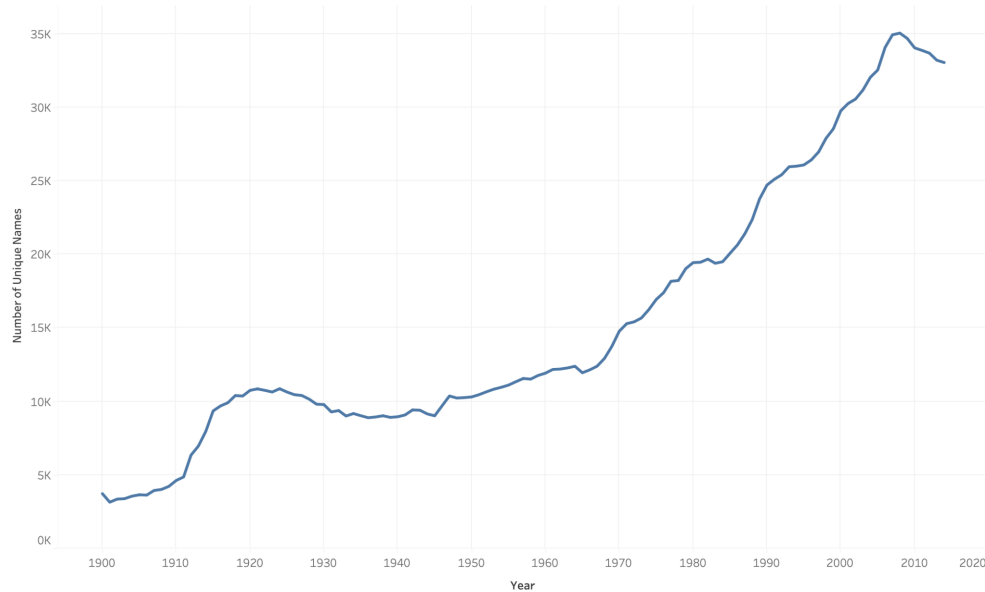


Conclusion: After performing this analysis on my name, I have come to the conclusion that my parents are just uncreative.

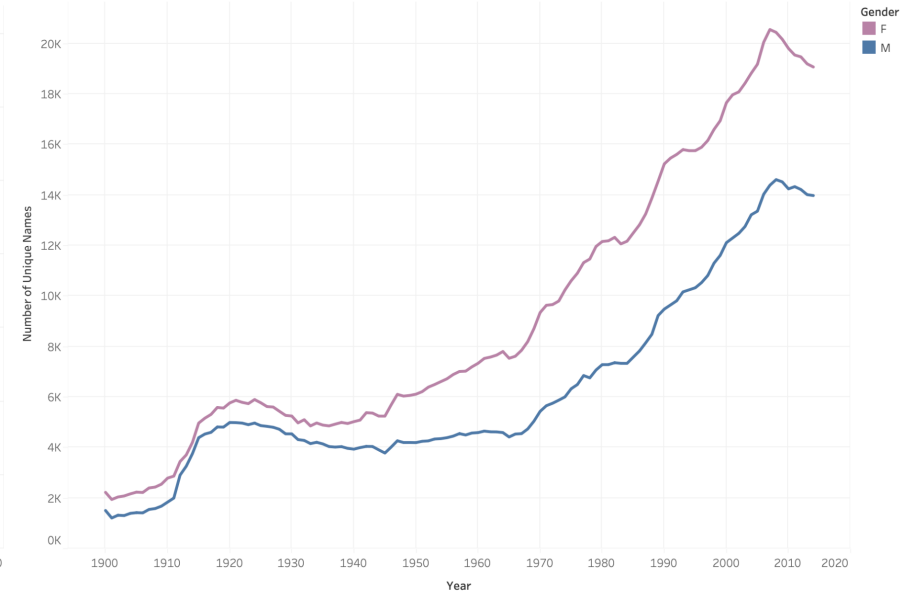
RQ4: In what year were the greatest number of unique names given to children born in the United States?

I then chose to explore how the variety in baby names has changed since the 1900s. For this research question I had to modify my data to reflect the number of unique names for each year using distinct counts of unique instances of names and mapped the spread across time from the 1900s to 2014. From the left chart we see that 2008 had the most unique names given to babies born that year in the US. I further investigated the difference in use of unique baby names across genders. From the right chart below we can see that girls have consistently had a larger number of unique names.

No. of Unique Baby Names in the US from 1900-2014



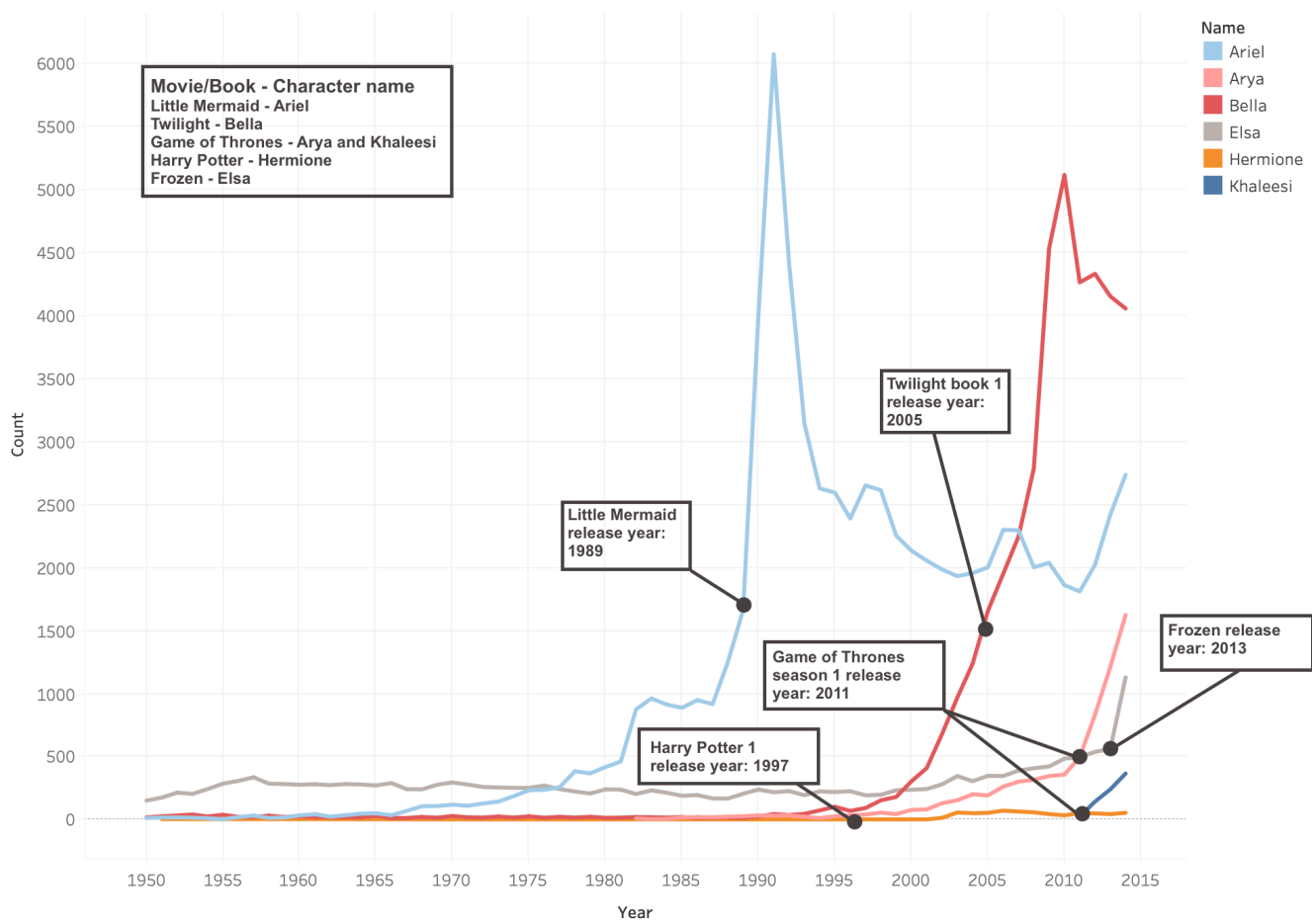
No. of Unique Names by Gender from 1900-2014



RQ5: How have popular movie and book series that I read or watched during my life impacted the number of babies named after these movie/book characters?

For this analysis I first chose what books/movie series I would select names from. I chose The Little Mermaid, Harry Potter, Twilight, and Game of Thrones. I then selected a name or two from each franchise and plotted their popularity over time. I also added annotations with each movie/book's release year so it would be easier to understand the trends that evolved around its release time and how this impacted baby naming in the future. It was really interesting to see how much popular franchises impacted the increasing popularity of baby girl names but almost the opposite effect for baby boy names. For most of the male names, the frequency of naming baby boys after popular movie/book characters decreased after the movie/book's release.

Popularity of Girls Names From Popular Movies and Books from the 2000's-2010's



Popularity of Boys Names From Popular Movies and Books from the 2000's-2010's

