

QUESTION 1: Baby Names

Talyah Greyling^a

^a*Stellenbosch University, Cape Town, South Africa*

Abstract

This report is a concise analysis of baby naming trends in the US through the years utilising graphs, tables and bullet summaries to explain the findings to the client, a New York based kids' toy design agency.

1. Introduction

This analysis examines US baby naming trends to identify which trends endure versus fade. It then explores how a cultural influence like popular movies can shape naming choices. For brevity and ease of use all findings are presented in bullet point summaries.

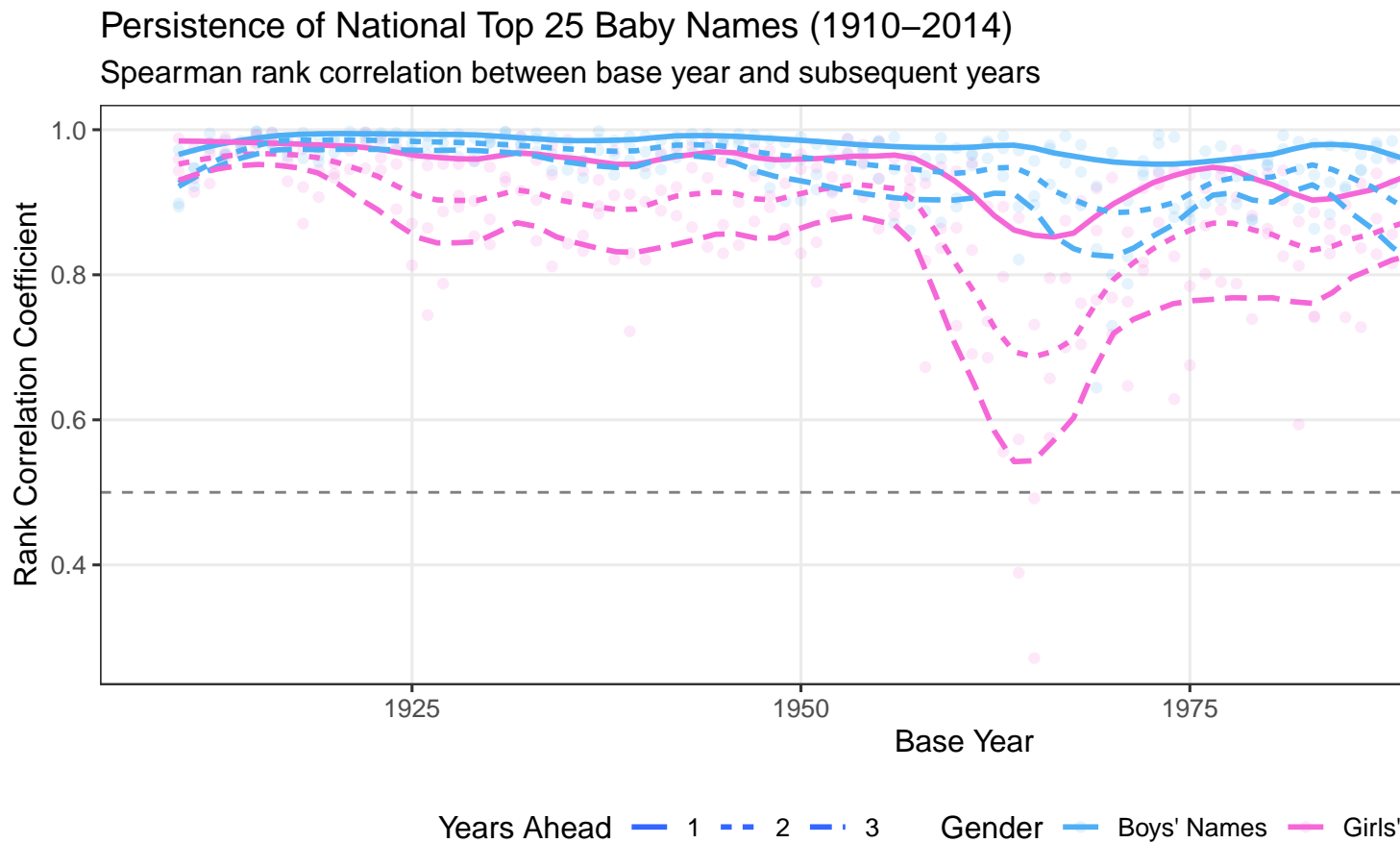
Corresponding author: Talyah Greyling

Email address: 23761067@sun.ac.za (Talyah Greyling)

Contributions:

The authors would like to thank the DNS server syncing for taking its time so that she could finish her morning cup of coffee before starting.

2. Persistence of baby names



Note: Correlations computed between each year's top 25 names and names in subsequent years. Only includes names that were in the top 25 in both years.

Figure 2.1: Persistence of National Top 25 Baby Names (1910-2014) by Spearman rank-correlation.

Insights from Figure 2.1: * Name persistence is much stronger among boys than girls. * Generally boy and girl names display downward sloping persistence trends over time. * Persistence becomes weaker for boys and girls the further the future year is from the target year. * The client's suspicion on slower to persistence since the 1990s is: * Correct for boys names. * Incorrect for girls names which display with an unexpected upwards trend since the 1990s.

4. Popular names per decade visualisation

F

M



Insights from Figure 4.1: * Name persistence is much stronger among males. * Popularity is concentrated among less names for males, contributing to their higher persistence levels.

* Most persistently popular name over all decades: * Girls: Mary * Boys: William * High popularity before the 1960s: * Girls: Nancy, Margaret, Helen, Patricia, Barbara, Mary * Boys: Thomas, Charles, Richard, William, John, Robert, James * High popularity after the 1960s: * Girls: Elizabeth, Amanda, Ashley, Emily * Boys: Daniel, Matthew, Joshua, Christopher, James, Michael

5. Movie star names calculations

```
## # A tibble: 29,118 x 7
##   name          Name      Surname character  title release_year tmdb_score
##   <chr>         <chr>    <chr>    <chr>      <chr>      <dbl>      <dbl>
## 1 Judy Garland   Judy      Garland Dorothy Ga~ The ~      1939      7.6
## 2 Ray Bolger     Ray       Bolger   Hunk / Sca~ The ~      1939      7.6
## 3 Jack Haley     Jack      Haley    Hickory / ~ The ~      1939      7.6
## 4 Bert Lahr      Bert      Lahr     Zeke / Cow~ The ~      1939      7.6
## 5 Margaret Hamilton Margaret Hamilton Elmira Gul~ The ~      1939      7.6
## 6 Frank Morgan   Frank     Morgan   Professor ~ The ~      1939      7.6
## 7 Clara Blandick Clara      Blandick Auntie Em   The ~      1939      7.6
## 8 Charley Grapewin Charley    Grapewin Uncle Henry The ~      1939      7.6
## 9 Billie Burke    Billie    Burke     Glinda      The ~      1939      7.6
## 10 Pat Walshe     Pat       Walshe    Nikko        The ~      1939      7.6
## # i 29,108 more rows

## # A tibble: 10 x 11
##   name      Name Surname character title release_year tmdb_score Decade Gender
##   <chr>    <chr> <chr>    <chr>    <chr>      <dbl>      <dbl> <dbl> <chr>
## 1 James Bu~ James Burke Luke      The ~      1941      7.8  1950 M
## 2 James Wh~ James Whitmo~ Gus Mini~ The ~      1950      7.5  1950 M
## 3 James St~ James Stewart Macaulay~ The ~      1940      7.7  1950 M
## 4 James Fl~ James Flavin  Mate Bri~ King~      1933      7.6  1950 M
## 5 James Mc~ James McKech~ Spud Wil~ The ~      1943      7.6  1950 M
## 6 James Kn~ James Knight Club Por~ The ~      1943      7.6  1950 M
## 7 James Re~ James Rennie Frank Mc~ Now,~      1942      7.4  1950 M
## 8 James Ca~ James Cagney Tom Powe~ The ~      1931      7.3  1950 M
## 9 James St~ James Stephe~ Howard J~ The ~      1940      7.4  1950 M
## 10 James Ca~ James Carlis~ <NA>      The ~      1940      7.4  1950 M
## # i 2 more variables: Total_Dec_Count <dbl>, Rank <int>
```

Table 5.1: Top 10 Female Names Coinciding With Actor Names

name	Name	Surname	character	title	release_year	tmdb_score	Decade	Gender	Total_Dec_Count	Rank
Mary Astor	Mary Astor	Brigid O'Shaughnessy		The Maltese Falcon	1941	7.8	1940	F	639986	1

name	Name	Surname	character	title	release_year	imdb_score	Decade	Gender	Total_Reviews	Count
Mary Anderson	Mary	Anderson	Maybelle Merriwether	Gone with the Wind	1939	8.0	1940	F	639986	1
Mary Nash	Mary	Nash	Margaret Lord	The Philadelphia Story	1940	7.7	1940	F	639986	1
Mary Clare	Mary	Clare	Baroness Isabel Nisatona	The Lady Vanishes	1938	7.5	1940	F	639986	1
Mary Wickes	Mary	Wickes	Dora Pickford	Now, Voyager	1942	7.4	1940	F	639986	1
Mary Morris	Mary	Morris	Halima	The Thief of Bagdad	1940	7.2	1940	F	639986	1
Mary Clare	Mary	Clare	Mrs. Corney	Oliver Twist	1948	7.5	1940	F	639986	1
Mary Boland	Mary	Boland	Mrs. Bennet	Pride and Prejudice	1940	6.9	1940	F	639986	1
Mary Jo Ellis	Mary	Jo Ellis	Clementine	The Barkleys of Broadway	1949	7.0	1940	F	639986	1
Mary Astor	Mary	Astor	Brigid O'Shaughnessy	The Maltese Falcon	1941	7.8	1950	F	625498	1
