Popularity of Baby Names

Tianyu Duan

2025-11-15

Introduction

This document contains: (1) Code for the Popular Baby Names Project, (2) Visualization for the Popular Baby Names Project, and (3) Narrative Text for the Popular Baby Names Project.

Code for the Popular Baby Names Project

```r
library(tidyverse)
library(ggplot2)

if (!requireNamespace("babynames", quietly = TRUE)) {
  install.packages("babynames")
}
library(babynames)

# Harmonize columns to match BabyNames schema
BabyNames <- babynames::babynames %>%
  select(name, sex, year, n)

names_to_plot <- c("Michael", "Jennifer", "Emma", "Liam")

baby_sel <- BabyNames %>%
  filter(name %in% names_to_plot) %>%
  mutate(sex = dplyr::recode(sex, F = "Female", M = "Male"))

baby_year <- baby_sel %>%
  group_by(name, year) %>%
  summarise(n = sum(n), .groups = "drop")

p_names_clean <- ggplot(
  data = baby_year,
  mapping = aes(x = year, y = n, color = name)
) +
  geom_line(linewidth = 1) +
```

```r
  scale_y_continuous(labels = scales::label_number(big.mark = ",")) +
  scale_color_viridis_d(option = "D", end = 0.9) +
  labs(
    title    = "Popularity of Selected U.S. Baby Names Over Time",
    subtitle = "Counts by year; male and female combined",
    x        = "Year",
    y        = "Number of Babies",
    color    = "Name"
  ) +
  theme_minimal(base_size = 12) +
  theme(plot.title.position = "plot", legend.position = "bottom")

p_facets <- ggplot(
  data = baby_year,
  mapping = aes(x = year, y = n)
) +
  geom_line(linewidth = 0.9) +
  facet_wrap(~ name, ncol = 2, scales = "free_y") +
  scale_y_continuous(labels = scales::label_number(big.mark = ",")) +
  labs(
    title    = "Popularity of Selected U.S. Baby Names (Small Multiples)",
    subtitle = "Counts by year; male and female combined",
    x        = "Year",
    y        = "Number of Babies"
  ) +
  theme_minimal(base_size = 12) +
  theme(
    plot.title.position = "plot",
    strip.text = element_text(face = "bold"),
    panel.grid.minor = element_blank()
  )

p_facets
```
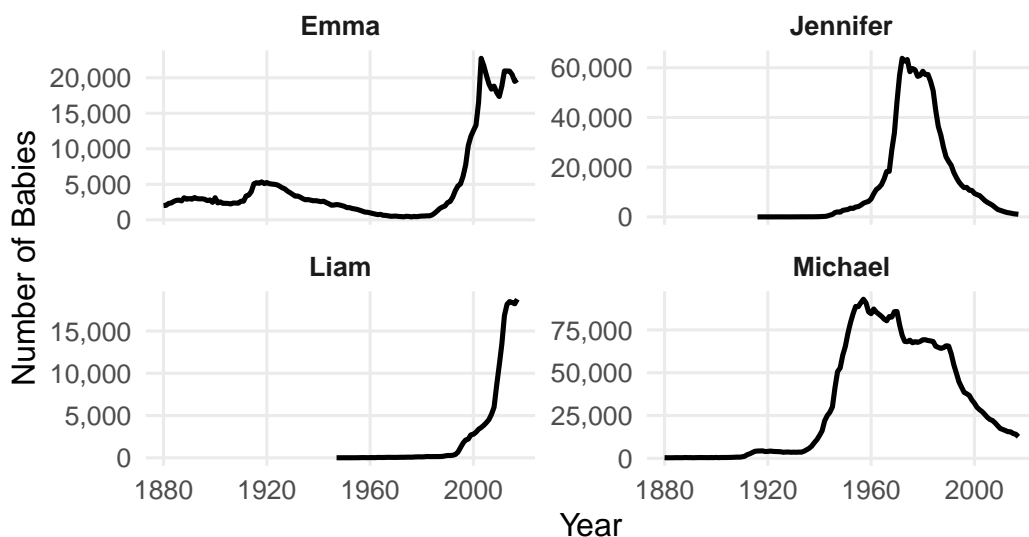
Popularity of Selected U.S. Baby Names (Small Multiples)
Counts by year; male and female combined

Visualization for Popular Baby Names Project

```r
if (!exists("baby_year")) {
  if (!requireNamespace("babynames", quietly = TRUE)) install.packages("babynames")
  library(babynames); library(tidyverse)
  BabyNames <- babynames::babynames %>% select(name, sex, year, n)
  names_to_plot <- c("Michael", "Jennifer", "Emma", "Liam")
  baby_year <- BabyNames %>%
    filter(name %in% names_to_plot) %>%
    group_by(name, year) %>%
    summarise(n = sum(n), .groups = "drop")
}

#| fig-cap: "Figure 1. U.S. Baby Names Popularity Over Time (Combined Sexes)"
#| fig-alt: "Time-series lines for four U.S. baby names-Michael, Jennifer, Emma, and Liam-show
#| fig-width: 8
#| fig-height: 4.5
#| fig-dpi: 300
library(ggplot2)

p_names <- ggplot(
  data    = baby_year,
  mapping = aes(x = year, y = n, color = name, linetype = name)
) +
  geom_line(linewidth = 1) +
  scale_y_continuous(labels = scales::label_number(big.mark = ",")) +
```
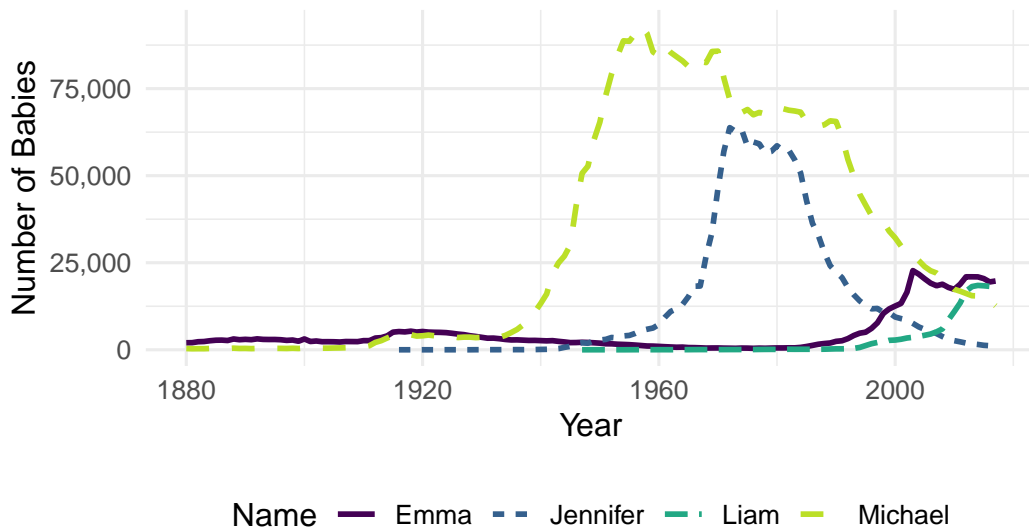
```
  scale_color_viridis_d(option = "D", end = 0.9) +      #
  labs(
    title    = "Popularity of Selected U.S. Baby Names Over Time",
    subtitle = "Counts by year; male and female combined",
    x        = "Year",
    y        = "Number of Babies",
    color    = "Name",
    linetype = "Name"
  ) +
  theme_minimal(base_size = 12) +
  theme(plot.title.position = "plot", legend.position = "bottom")

p_names
```

Popularity of Selected U.S. Baby Names Over Time
Counts by year; male and female combined

Narrative Text for the Popular Baby Names Project

The time-series graph demonstrates how the popularity of four U.S. baby names has varied over time. Michael and Jennifer were very popular in the middle to late 1900s, but they lost popularity after that. Emma and Liam, on the other hand, were much more frequent around 2000. I picked these names to highlight two opposite trends: older names that have become less popular and younger names that have become more popular. This means that the names people give their babies change throughout time, not because they have always liked them. The picture is clear, works for people who are colorblind, and employs both line type and color to make the data easy to understand.