Data Wrangling and Visualization Summary

Canyi Chen

2025-11-10

Table of contents

1 Data Wrangling, Visualization, and Reflection

This report presents a tidy frequency table for a selected Armed Forces sub-group, a time series for baby names, and a plot of a box-volume function.

1.1 Armed Forces Data Wrangling Redux

This section explores how **sex and rank** interact in one branch of the U.S. Armed Forces. I focus on the **Army** and limit the analysis to **Enlisted ranks (E1–E9)**, because this group has many observations and helps avoid empty categories.

pay_grade	sex	count
E1	Female	1326
E2	Female	4336
E3	Female	10229
E4	Female	15143
E5	Female	10954
E6	Female	7363
E7	Female	4410
E8	Female	1472
E9	Female	394
E1	Male	7429
E2	Male	22338
E3	Male	43775
E4	Male	79234
E5	Male	54803
E6	Male	49502
E7	Male	30264
E8	Male	9482
E9	Male	2865
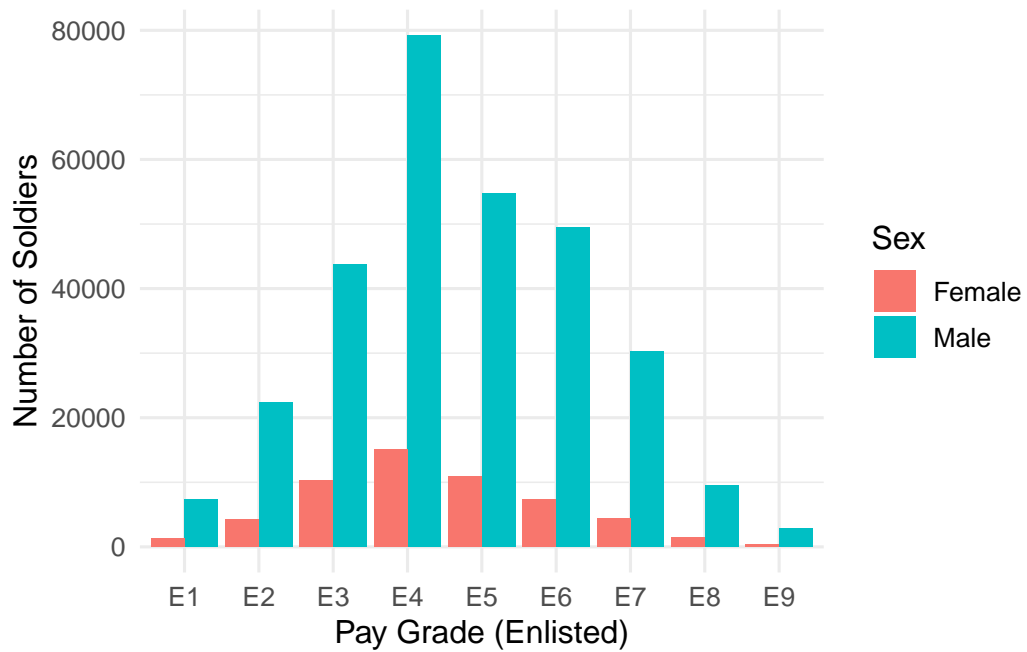
1.2 Popularity of Baby Names

Name of my siblings.

Figure 1. Popularity of Charlie, Cindy, and Poppy Over Time
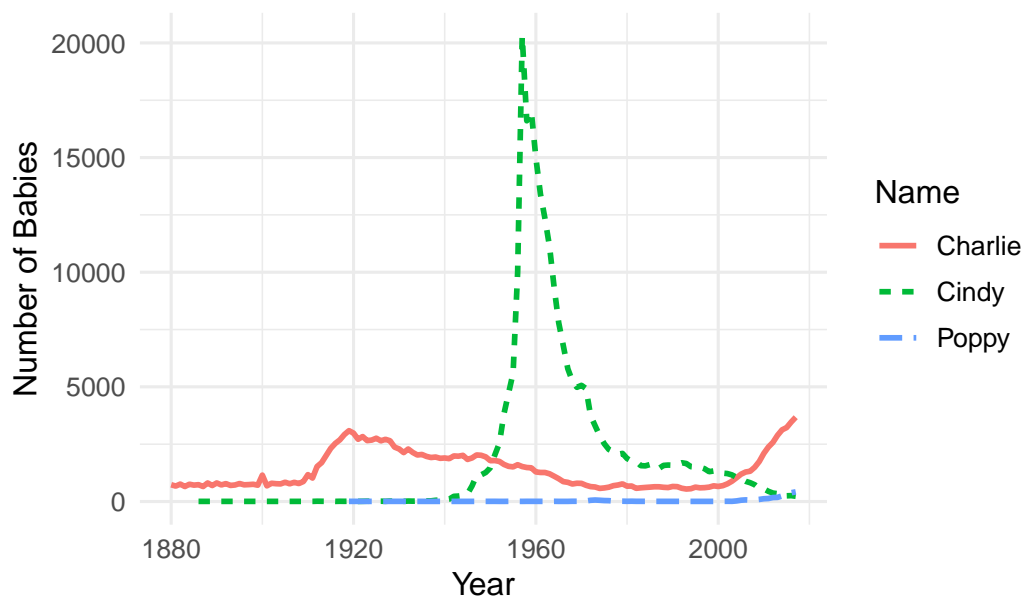
1.3 Box Problem Function and Plot (36 × 48 inches)

The sheet size is 36 by 48 inches. The volume is ($V(x)=x(36-2x)(48-2x)$) for ($0<x<18$). Figure 2 shows the curve over the practical domain.

1.4 What I Have Learned

I have learned how to clean and reshape public datasets, visualize trends clearly, and make reproducible reports using Quarto. This project helped me connect data wrangling, visualization, and mathematical modeling in one workflow.