Deggendorf Institute of Technology
Faculty of Computer Science

Polio in the United States: Socioeconomic Patterns and the Effects of the Salk and Sabin Vaccines (1928-1970)

A mini research project for Data Visualization under Life Science Informatics

Submitted by

Wilbroda Onyango

Beverly Odhiambo

Faculty of Computer Science
Deggendorf Institute of Technology
2025

1. Introduction

Poliomyelitis, commonly known as polio is a highly contagious illness that primarily affects young children. It is an enteric (intestinal) disease contracted through contact with faecal matter typically via contaminated food or water, shared objects, or poor hand hygiene. The disease targets the neurological system and can cause respiratory and spinal paralysis, as well as death in certain cases. Since ancient times, polio has been recognized as a debilitating and sometimes lethal disease. The 1950s saw a particularly severe impact on the United States, as enormous hospital wards were crowded with people on respirators. With 20,000 instances of paralytic poliomyelitis in 1952, the disease's incidence peaked in the US. Polio continues to be a leading cause of disability in underdeveloped nations (Forum et al., 1996).

For the most part, local residents followed the common public health guidance of the time: to treat filth as the enemy and strive for cleanliness as the primary defence (Oshinsky, 2005). However, this strategy was highly unsuccessful in combating polio in the US, in contrast to many infectious diseases where better sanitation dramatically lowers transmission. Polio outbreaks continued despite extensive hygiene efforts and were actually more common in high income states with superior sanitation infrastructure. This contradiction demonstrated the poliovirus's peculiarity, since it defied the usual pattern of illness linked to substandard living conditions. While overcrowding and inadequate healthcare were problems in low-income states, the wealthier, cleaner locations frequently recorded the highest number of instances.

In recent years, studies have suggested that improved public sanitation may have created the polio epidemic. Before improvements to public sanitation, almost everyone became infected with poliovirus during early childhood. Infection most often occurred between 6 months and 4 years of age. However, symptoms were usually mild, resembling the common cold, because babies acquired passive immunity from their mothers while in utero. Thereby developing lifelong immunity to polio (Johnson, 2017).

Then came better sanitation including improved sewage disposal and clean water supplies in the United States. Without contact with contaminated food, water, and dirt, children missed the opportunity to contract the illness at a young age and develop lifelong immunity to polio. Prior to cleaning up the public water supply, almost all children were exposed to the virus through faeces-contaminated water. However, in affluent states where sanitation improvements were implemented earlier children were no longer exposed during infancy.

This delayed exposure meant that more individuals encountered the virus later in childhood or adolescence, when the risk of infection was higher and immunity less likely. As a result, polio incidence was paradoxically higher in these wealthier, cleaner regions, despite their better living conditions (Johnson, 2017).

1.1. Salk Vaccine Development

The first successful polio vaccine was created by US physician Jonas Salk. It is commonly known as the Salk Vaccine, an inactivated virus vaccine that is administered intravenously. It was developed in 1954 in the state of Pennsylvania. Development of this vaccine significantly reduced polio incidence but did not fully eradicate the virus.

1.1.1. The Cutter Incident

In 1955, the Cutter incident occurred when a batch of Salk vaccines produced by Cutter Laboratories contained live poliovirus instead of the inactivated form. The faulty vaccine led to hundreds of polio cases and several deaths, primarily affecting children in a few U.S. states, including California, Idaho, Illinois, Washington, and Colorado. As a result, the vaccination program was temporarily halted.

1.2. Sabin Vaccine Development

In 1961, Albert Sabin developed a live attenuated vaccine that could be given orally, as drops or on a sugar cube. The ease of administering the oral vaccine made it the ideal candidate for mass vaccination campaigns which led to a dramatic and sustained decline in polio cases across the United States (*History of Polio Vaccination*, n.d.).

However, in the late 1990s, it was discovered that the oral vaccine could, in rare cases, cause vaccine-associated paralytic poliomyelitis (VAPP) in some patients. As a result, the United States reverted to using the inactivated Salk vaccine, which eliminated the risk of VAPP and contributed to the complete eradication of polio and related illnesses in the country (Rai et al., 2022).

2. Aim

This study aims to analyze and visualize the relationship between income levels and polio incidence across U.S. states. Through data cleaning, transformation, and visual exploration, the project seeks to identify patterns and trends that reveal how socioeconomic status may

have influenced polio case distribution. The goal is to provide clear, data-driven insights using charts, graphs, and maps to support historical understanding of the disease's spread.

3. Methods

Our research and analysis were conducted on secondary data from Project Tycho on polio incidence in the US. Our findings were further informed and supported by income statistics in the US from the Federal Reserve Bank of St. Louis. The data, their sources, data analysis conducted, and the tools and packages used for this project are further elaborated below.

3.1. Data sources

3.1.1. Polio incidence in the United States (1928 – 1968)

Project Tycho is a public health data initiative developed by the University of Pittsburgh that provides open-access data on infectious diseases in the United States and other countries. The database features data on disease cases and deaths sources from public health surveillance. As of 2018, Project Tycho featured over 3.6 million counts of reported incidences and deaths in the United States, with data that dates as far back as 1888 (van Panhuis et al., 2018).

This research project uses level 1 data from Project Tycho on the incidence and deaths from various infectious diseases in the United States between 1916 and 2011 (Van Panhuis et al., 2018). The original dataset contains counts for eight diseases, including smallpox, polio, hepatitis A and rubella, although our research project focuses only on polio data. It is organised into polio case counts and incidence rates per 100,000 people, spanning all 50 U.S states and Washington D.C from 1910 to 1968. The polio data used, however, starts in 1928 and spans until 1968.

3.1.2. Per Capita Personal Income in the United States (1929 – 1970)

As we have focused our research on the relationship between economic status and polio incidence in the late 1920s to late 1960s, a significant part of research entails comparing economic trends in the United States to polio incidence rates. All economic data was downloaded from the Federal Reserve Economic Data – FRED in short – database. FRED is an online database compiling time series economic data from several sources in the United States, managed by the Federal Reserve Bank of St. Louis (*What Is FRED?*, n.d.).

Per capita personal income (PCPI) is a widely recognised and standardised measure of economic well-being at the individual level, making it an appropriate metric for analysing

state-level economic trends over time in the US. PCPI captures the average income received by residents of a state from all sources, adjusted for population size (*Income & Saving | U.S. Bureau of Economic Analysis (BEA)*, n.d.). It provides insight into relative living standards and income dynamics. As such, PCPI was our chosen economic metric for comparisons across states over time.

The economic data analysed herein is time series PCPI data from FRED. The vast amount of data on the database was filtered for PCPI data from 1929 to 1970. 1929 is the first year with available PCPI data on FRED, while 1970 is two years after the end of our available polio Tycho dataset. The downloaded dataset contains annual PCPI in US dollars (*2024, Release Tables: Per Capita Personal Income by State, Annual | FRED | St. Louis Fed*, n.d.) for each of the twelve aforementioned states of interest (Alabama, Arkansas, Mississippi, South Carolina, Louisiana, New York, California, Illinois, District of Columbia, New Jersey, Pennsylvania and Washington.

3.2. Variables

The aim of the data analysis was to find trends and potentially identify a relationship between income and polio incidence in the US. The main variables of interest in the Tycho dataset were year, state, total polio cases, and incidence per 100,000. The data for all other diseases were filtered out. For the FRED data, the main variables of interest were the observation date (year) and the states' PCPI.

3.3. Data Cleaning and Preparation

Dates in the Tycho dataset are presented in the form of epidemiological weeks (epi_week). These were transformed to actual dates in the DD-MM-YYYY format. The dates were also later transformed to years. Minimal cleaning was required for the FRED PCPI data, as the download presents it clean and ready to use.

3.4. Analytical Approach

Descriptive statistics and time-series visualizations were generated. Animated choropleth maps were created using the usmap and gganimate packages in R to visualize spatiotemporal trends in polio incidence. Additionally, comparisons were made between pre-vaccine (before 1955) and post-vaccine periods to visualise the impact of the Salk and Sabin vaccines.

3.5. Tools Used

All analyses were conducted in R (version 4.5.0). Several packages were used for different analytical and visualisation purposes:

Data Manipulation and Import

Package	Version	Purpose
dplyr	1.1.4	Core package for data manipulation: filtering, selecting, summarising, mutating
readr	2.1.5	Importing csv files into R
tidyr	1.3.1	Reshape and tidy the data
tidyverse	2.0.0	Loads the core packages of the tidyverse in one go: ggplot2, dplyr, tidyr, readr, etc
ggplot2	3.5.2	Used for creating static data visualisations

Maps and Geo-spatial Data

Package	Version	Purpose
usmap	0.8.0	Contains tools for plotting U.S. maps by state or country
sf	1.0-21	Handles spatial vector data and used for plotting U.S. maps, specifically highlighting states
usdata	0.3.1	Includes datasets and tools for working with U.S. state codes, abbreviations and regions

Animation and Interactive Graphics

Package	Version	Purpose

gganimate	1.0.10	Extends ggplot2 to make animated plots using time or transitions
plotly	4.11.0	Turns ggplot2 plots or data frames into interactive HTML plots

Visual Elements and Output

Package	**Version**	**Purpose**
shadowtext	0.1.5	Adds readable text labels with shadows to ggplot2 plots
gifski	1.32.0-2	Renders animated gganimate plots into high-quality GIFs
png	0.1-8	Reads and writes PNG images
magick	2.8.7	Advanced image processing, including annotating and animating

4. Results

To explore economic trends across U.S. states from 1929 to 1970, we visualized changes in Per Capita Personal Income (PCPI) over time. The graph below displays annual PCPI values for each state, providing a comparative overview of income growth patterns during this period.

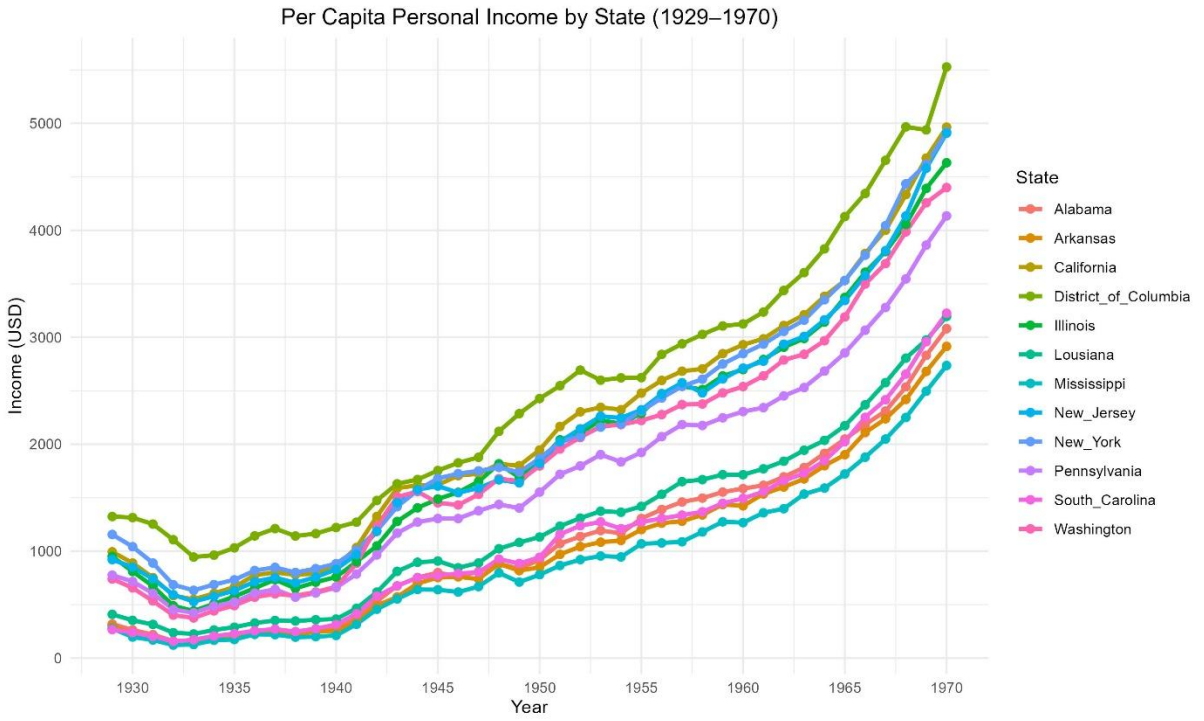

Figure 4.1 Per Capita Personal Income trends by state from 1929 to 1970.

Polio was a major public health concern in the United States throughout the early to mid-20th century. Historical records and literature highlight 1952 as the year with the highest recorded incidence, particularly in New York, marking a critical peak in the national outbreak.

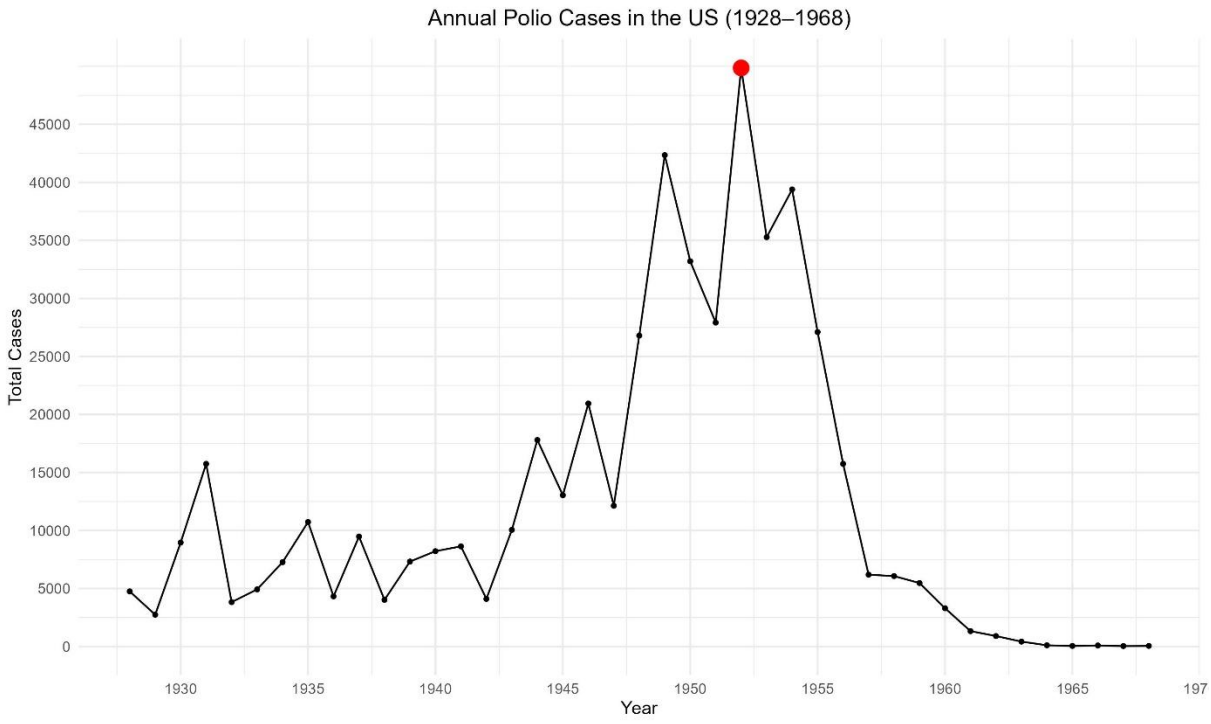

Figure 4.2 Annual Polio Incidence trends in the US from 1928 to 1968.

In 1954, the Salk vaccine was developed in Pennsylvania and rolled out countrywide. By 1955, polio incidence reduced significantly across US states. In 1955, thousands of Salk vaccines were contaminated in the Cutter Laboratory in California; the incident was named the Cutter Incident. Four other states were affected: Illinois, Washington, Colorado and Idaho. While polio incidence dropped across the US, these five states continued to record high numbers.

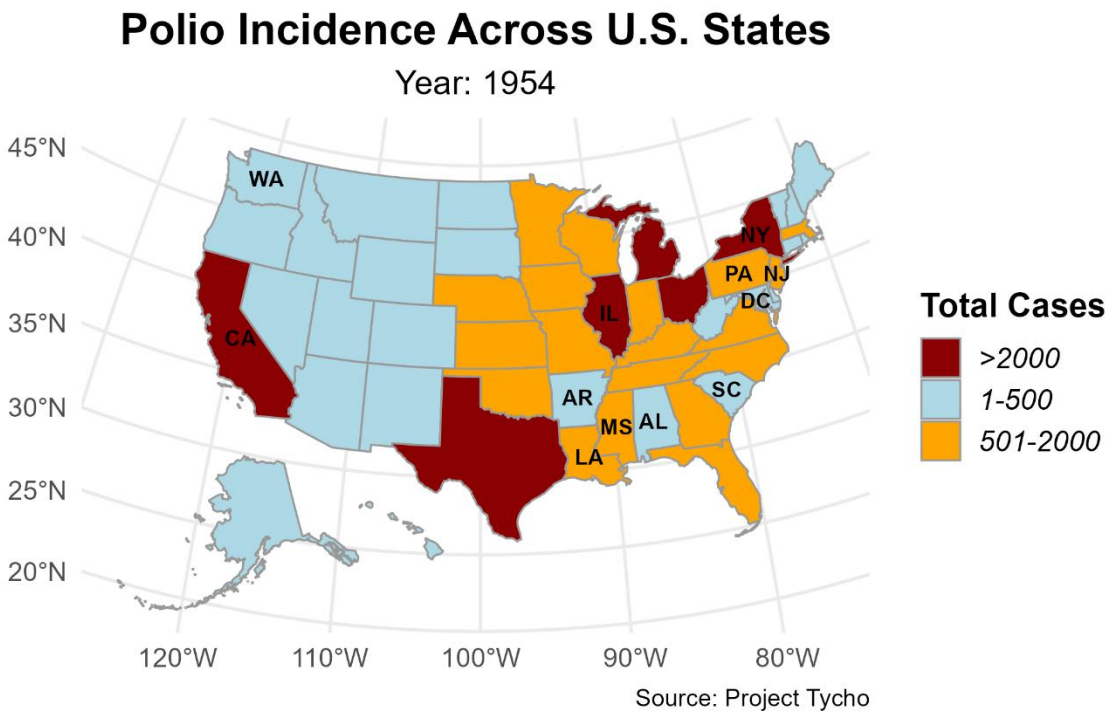

Figure 4.3 Polio incidence in the US in 1954 with 12 highlighted states.

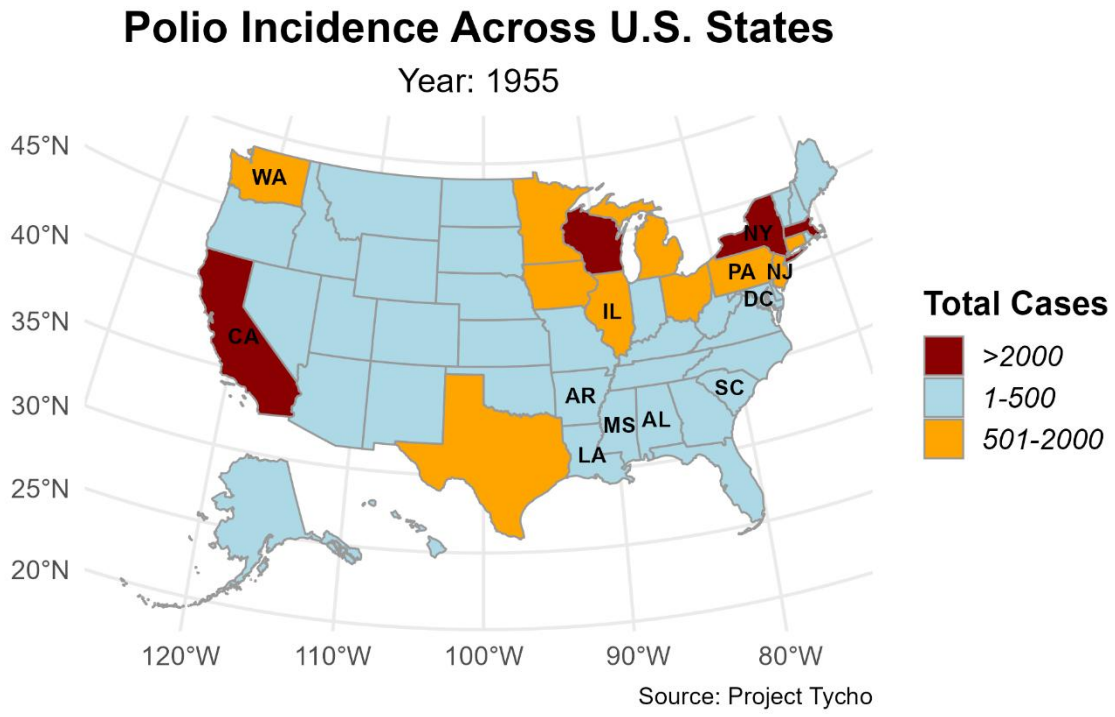

Figure 4.4 Polio incidence in the US in 1955 with 12 highlighted states.

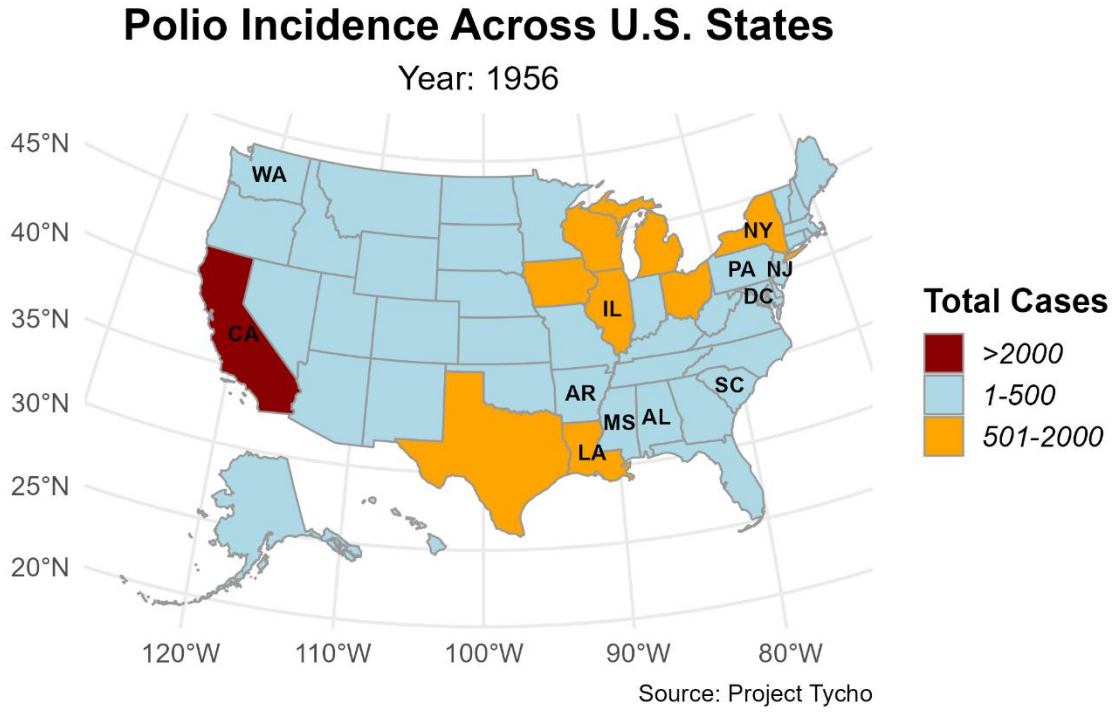

Figure 4.5 Polio incidence in the US in 1956 with 12 highlighted states.

The line graph below illustrates the total number of reported polio cases from 1950 to 1960 across twelve states. This period encompasses both the pre-vaccine era and the years following the introduction of the polio vaccine, allowing for a comparison of disease trends before and after vaccination efforts began. Notably, the graph highlights the 1955 Cutter incident, a critical event involving a batch of improperly inactivated vaccine doses that temporarily impacted polio case numbers. This visualization provides insight into the effectiveness of vaccination programs as well as the consequences of vaccine-related setbacks during this decade.

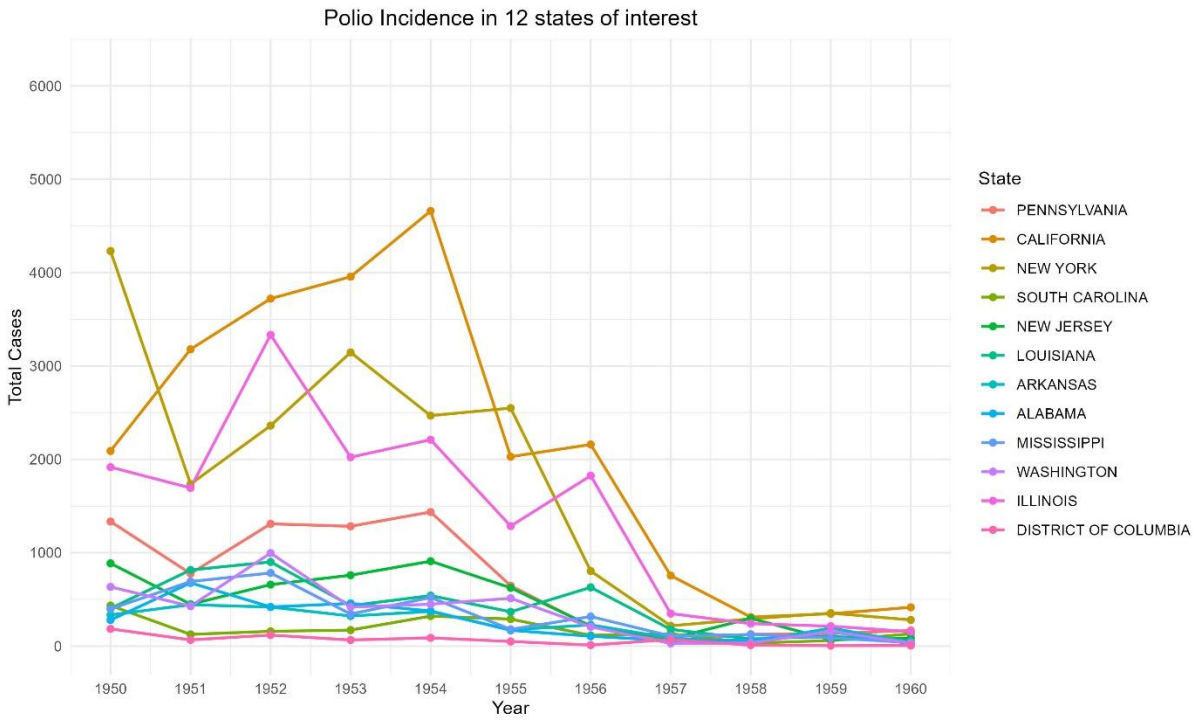

Figure 4.6 Polio cases in the US (1950-1960), including the Cutter incident of 1955. The figure shows a rise in Polio cases after the Cutter Incident of 1955.

5. Discussion

The data visualizations from this project reveal a paradox: polio incidence was higher in high income states, despite their earlier access to improved sanitation and healthcare infrastructure. The graphs show that high income states such as California, New York, Illinois and Pennsylvania experienced some of the highest case rates, while many lower income states such as South Carolina, Mississippi, Arkansas and Alabama reported fewer cases. The idea that wealth equates to better health outcomes is being called into question by this tendency, particularly in light of infectious diseases like polio.

Our findings are supported by historical evidence. Children's exposure to the polio virus was delayed in wealthier states due to the early implementation of sanitation systems. Because of this, they were exposed to the virus later in childhood, when they were more vulnerable to symptoms, as opposed to during infancy, when protection could be provided by natural immunity from maternal antibodies. This delayed exposure made polio more common in cleaner, wealthier regions.

A turning point was reached with the release of the Salk vaccine, but public confidence was momentarily damaged in 1956 when contaminated vaccine batches led to new illnesses in multiple states. Despite this, widespread vaccination efforts were revived and subsequently reinforced by the oral Sabin vaccine, which made mass vaccination possible due to its simplicity of administration. Nationwide, this resulted in a sharp drop in polio cases.

6. Conclusion

This study demonstrates a clear decline in polio cases following the introduction of the vaccine in the early 1950s, underscoring the significant impact of vaccination programs on controlling infectious diseases. The 1955 Cutter incident highlights the critical importance of vaccine safety and quality control, as it temporarily reversed progress by causing a surge in cases. Additionally, our analysis revealed that states with higher income levels experienced higher polio incidence rates, while lower-income states tended to have fewer reported cases. This income-related pattern suggests that socioeconomic factors influenced disease transmission or reporting during this period. Overall, the findings emphasize the tremendous benefits of immunization, the need for rigorous vaccine oversight, and the role of socioeconomic context in shaping public health outcomes. Continued vigilance and targeted strategies remain essential for effective disease control.

References

2024, Release Tables: Per Capita Personal Income by State, Annual | FRED | St. Louis Fed. (n.d.). Retrieved 11 July 2025, from https://fred.stlouisfed.org/release/tables?rid=110&eid=257197#snid=257202

Forum, I. of M. (US) V. S., Howe, C. J., & Johnston, R. B. (1996). History and Current Status. In *Options for Poliomyelitis Vaccination in the United States: Workshop Summary*. National Academies Press (US). https://www.ncbi.nlm.nih.gov/books/NBK231547/

History of polio vaccination. (n.d.). Retrieved 16 September 2025, from https://www.who.int/news-room/spotlight/history-of-vaccination/history-of-polio-vaccination

Income & Saving | U.S. Bureau of Economic Analysis (BEA). (n.d.). Retrieved 11 July 2025, from https://www.bea.gov/resources/learning-center/what-to-know-income-saving

Johnson, H. (2017, June 1). How Improved Public Sanitation Created the Polio Epidemic. *Parenting Patch*. https://parentingpatch.com/improved-public-sanitation-created-polio-epidemic/

Oshinsky, D. M. (2005). *Polio: An American Story*. Oxford University Press, USA.

Rai, A., Uwishema, O., Uweis, L., El Saleh, R., Arab, S., Abbass, M., Wellington, J., Musabirema, F., Adanur, I., & Patrick Onyeaka, C. V. (2022). Polio returns to the USA: An epidemiological alert. *Annals of Medicine and Surgery*, *82*. https://doi.org/10.1016/j.amsu.2022.104563

Van Panhuis, W., Cross, A., & Burke, D. (2018). *Project Tycho Level 1 data: Counts of multiple diseases reported in UNITED STATES OF AMERICA, 1916-2011* [Dataset]. Project Tycho. https://zenodo.org/records/12608992

van Panhuis, W. G., Cross, A., & Burke, D. S. (2018). Project Tycho 2.0: A repository to improve the integration and reuse of data for global population health. *Journal of the*

American Medical Informatics Association, *25*(12), 1608–1617. https://doi.org/10.1093/jamia/ocy123

What is FRED? | Getting To Know FRED. (n.d.). Retrieved 11 July 2025, from https://fredhelp.stlouisfed.org/fred/about/about-fred/what-is-fred/