

Final Project

Kerry Yu

Background: Fertility and infant mortality in the U.S.

Infant mortality is the death of an infant before their first birthday, while the fertility rate refers to the average number of children a woman is expected to have during her reproductive lifetime. Fertility and infant mortality rates are critical indicators of public health and socioeconomic well-being. Fertility and infant mortality have a strong relationship with each other. Changes in fertility rates can influence infant mortality rates, as factors such as maternal age, education, and access to healthcare play significant roles in both outcomes.

For example, higher fertility rates among younger or less-educated mothers may correlate with higher infant mortality rates due to disparities in prenatal care and socioeconomic resources. Similarly, differences in geography and demographic in these rates can call attention to the inequities in healthcare access, economic stability, and public health infrastructure. Together, these metrics offer a window into the broader social, economic, and environmental conditions that shape human health and development.

In 2022, over 20,500 infants passed away in the United States. The top five causes of infant mortality that year were:

- Congenital abnormalities (birth defects).
- Preterm birth and low birth weight.
- Sudden Infant Death Syndrome (SIDS).
- Accidental injuries, such as car accidents.
- Complications arising during pregnancy.

Infant mortality and fertility data

The data is sourced from the Centers for Disease Control and Prevention (CDC), which suppresses counts between 1 and 9 for privacy reasons, resulting in some missing values. Despite this limitation, the dataset provides a robust foundation for analyzing trends and disparities in fertility and infant mortality over time.

Fertility rates, measured as the number of births per 1,000 women aged 15–44, reflect population dynamics and societal trends, while infant mortality rates, measured as the number of infant deaths per 1,000 live births, provide insight into the health and healthcare access of a population. Infant mortality rates in the U.S. have remained close to 5.6 deaths per 1,000 live births, while the US fertility rate is about 55 births per 1,000 women aged 15–44.

The data for **infant mortality** in the US spans from 2003 to 2023 for women aged 15–44. The table below provides a summary of descriptions for the variables measured in the infant mortality dataframe.

Category	Variables
Geographic and Time	state, year, biannual code
Birth Counts by demographic	birth counts by demographic (e.g. non-Hispanic Black, Hispanic, etc.) total birth counts across all demographics
Infant Death Counts	infant death counts by demographic (e.g. non-Hispanic Black, Hispanic, etc.) infant death counts of congenital-related deaths (present a birth) and of non-congenital related deaths infant death counts of neonatal deaths (first 28 days of life) total infant death counts across all categories

While the data for **fertility** in the US spans from 2016 to 2023 for women aged 15–44. The table below provides a summary of descriptions for the variables measured in the fertility dataframe.

Category	Variables	Information
Geographic and Time	state, year, bimonthly code, time, start and end dates of observation	
Birth Counts	age education marital status	grouped into 3 age cohorts (15-24, 25-34, 35-44) without high school education, with high school education, with some college education, or with college degree married or unmarried

Category	Variables	Information
	insurance status	birth covered by Medicaid or not covered by Medicaid
	race/ethnicity	birth to non-Hispanic Black, non-Hispanic White, Hispanic, or other
	total	total births across all categories
Population Statistics	total	total population in the area
	race/ethnicity	population of women of non-Hispanic Black, non-Hispanic White, Hispanic, and other races/ethnicities
	insurance	population of women covered by Medicaid
	education	population of women not covered by Medicaid
		population of women without high school education
		population of women with high school education
		population of women with some college education
	marital status	population of women with a college degree
		population of married women
	age	population of unmarried women
		population of women by age cohorts (15-24, 25-34, 35-44)

Motivation

As stated previously, fertility and infant mortality rates serve as key indicators of public health. They reflect the broader social, economic, and healthcare conditions of a population. Changes in fertility rates—whether due to shifting societal norms, economic conditions, or policy changes—can significantly impact infant mortality. For example, higher fertility rates among certain demographic groups or regions may place pressure on healthcare resources, potentially leading to higher infant mortality rates. Conversely, improvements in healthcare access or socioeconomic conditions may mitigate these risks.

Understanding the relationship between fertility and infant mortality rates, along with the underlying factors shaping these trends, is important for identifying vulnerable populations and informing targeted public health interventions.

Question of interest

How have changes in fertility rates across different demographic groups in the U.S. influenced infant mortality rates from 2016 to 2023, and what socioeconomic or healthcare-related factors contribute to these trends?

The question is formulated to explore three key dimensions:

1. How have fertility rates changed across demographic groups from 2016 to 2023?
2. How have infant mortality rates changed over the same period, and how do these changes correlate with fertility rate trends?
3. What socioeconomic and healthcare-related factors explain the observed trends and correlations?

A satisfactory answer would

- describe how fertility rates have changed across demographic groups over time,
- highlight changes in infant mortality rates over the same period and their correlation with fertility rate trends
 - for example, do demographic groups with higher fertility rates also have higher infant mortality rates?
- identify and analyze socioeconomic and healthcare-related factors that explain the observed trends.

Data exploration and analysis

Let's begin the first step to answering our question by analyzing the general trends of both the fertility rate and infant mortality rate. The analysis is divided into several steps, including data preparation, exploratory data analysis (EDA), and visualization of trends across different demographic groups and regions. The formulas to finding the infant mortality rate and the fertility rate are respectively given below:

$$\text{Infant mortality rate} = \frac{\text{Number of infant deaths in a population}}{\text{Number of live births in the same population}} \times 1000$$

$$\text{Fertility rate} = \frac{\text{Number of live births to a population}}{\text{Population of women in the same population}} \times 1000,$$

Now that we have the formulas to calculate the infant mortality and fertility rates for a population, let's start doing some exploratory data analysis! We want to start by looking at the general trend lines for fertility rates and infant mortality rates from 2016-2023.

Handling missingness

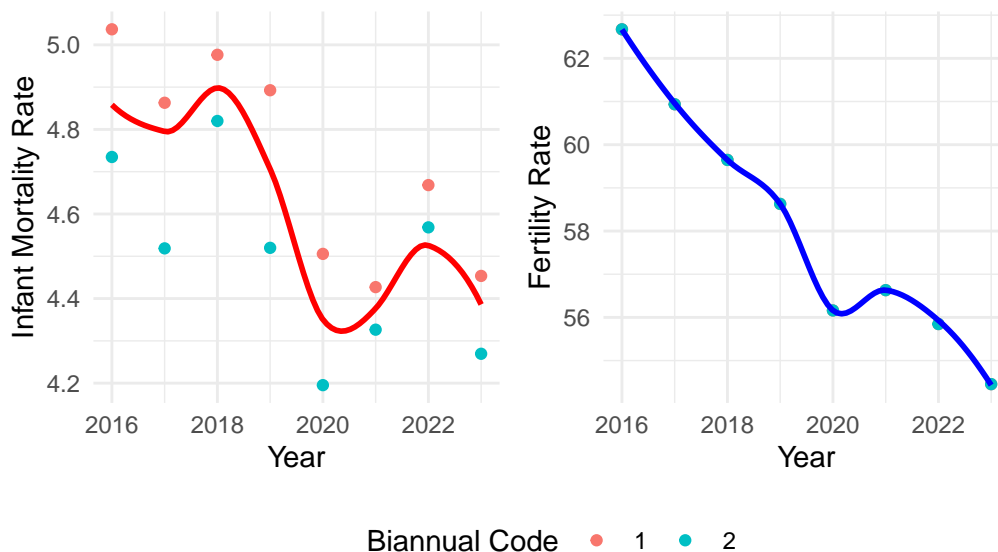
Missing values in both datasets were replaced with 0 using the `replace_na()` function. This was done under the assumption that missing values represent suppressed counts (due to privacy concerns) and can be treated as zero for aggregation purposes.

Exploratory Data Analysis

Overall trends

Average fertility and infant mortality rates were calculated for each year and biannual period, and trends were visualized using line charts with LOESS smoothing.

Infant Mortality vs Fertility Rates in the U.S. (2016–2023)



From the graphs above, we can see that the infant mortality rate appears to be on a general decline in the US from 2016-2023, showing a bit of fluctuations. The fertility rate, also showing a general decline, shows a clearer downward trend. This indicates a direct correlation between the fertility rate and infant mortality rate in the US—as fertility declines, infant mortality also trends downward overall.

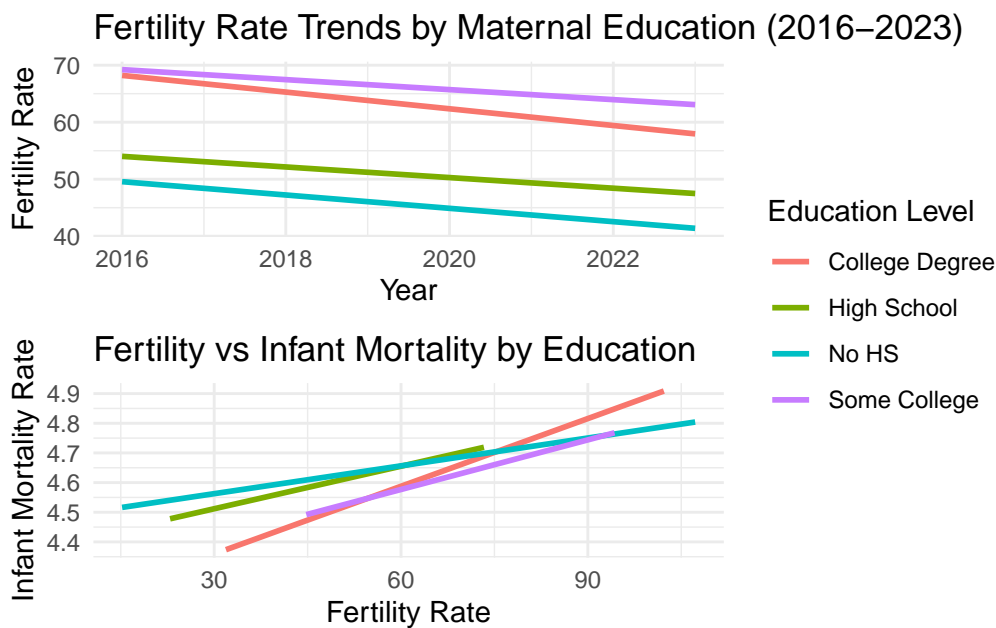
Historically, higher fertility rates have often been associated with higher infant mortality rates, particularly in developing regions where access to healthcare is inconsistent. In the United States, where healthcare systems are more stable, a declining fertility rate may indicate fewer high-risk births, improved maternal health, and enhanced prenatal care—potentially contributing to lower infant mortality rates. However, the observed fluctuations in infant mortality

suggest that additional factors, such as disparities in healthcare access, economic conditions, and racial/ethnic inequities, also significantly influence these trends.

Demographic and socioeconomic factors

The analysis focused on three key factors: education level, medicaid coverage, and race/ethnicity. For each factor, the data was reshaped using `pivot_longer()` to create long-format data, and fertility rates were calculated by dividing the number of births by the corresponding population and multiplying by 1,000. The next following graphs will utilize linear smoothing rather than LOESS smoothing to cut out the noise and clearly show the general trends and relationships between our factors.

Education level

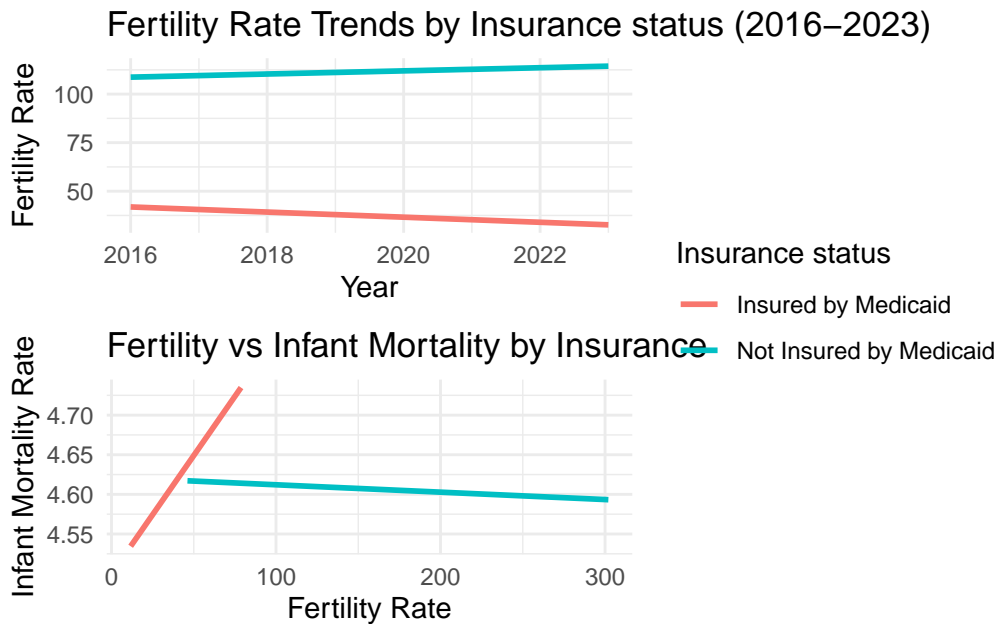


From the graph above of fertility rate stratified by maternal education level, we can observe that all education levels show a gradual decline in fertility rates over the observed period (2016-2023). Women with some college education have the highest fertility rate, followed by those with a college degree, then a high school diploma, and no high education with the lowest fertility rate. Surprisingly, higher education levels generally correlate with higher fertility rates, though all groups have shown a downward trend.

The graph following shows the linear relationship between fertility and infant mortality. There's a positive correlation between fertility rates and infant mortality rates across education levels. “Some College” and “College Degree” have the highest fertility rates and corresponding higher infant mortality rates. “No HS” has the lowest fertility rates but also

shows the lowest infant mortality rates among the groups. While fertility rates increase with education, infant mortality rates follow the same upward trend, indicating a possible confounding factor or socioeconomic variable influencing both.

Medicaid coverage

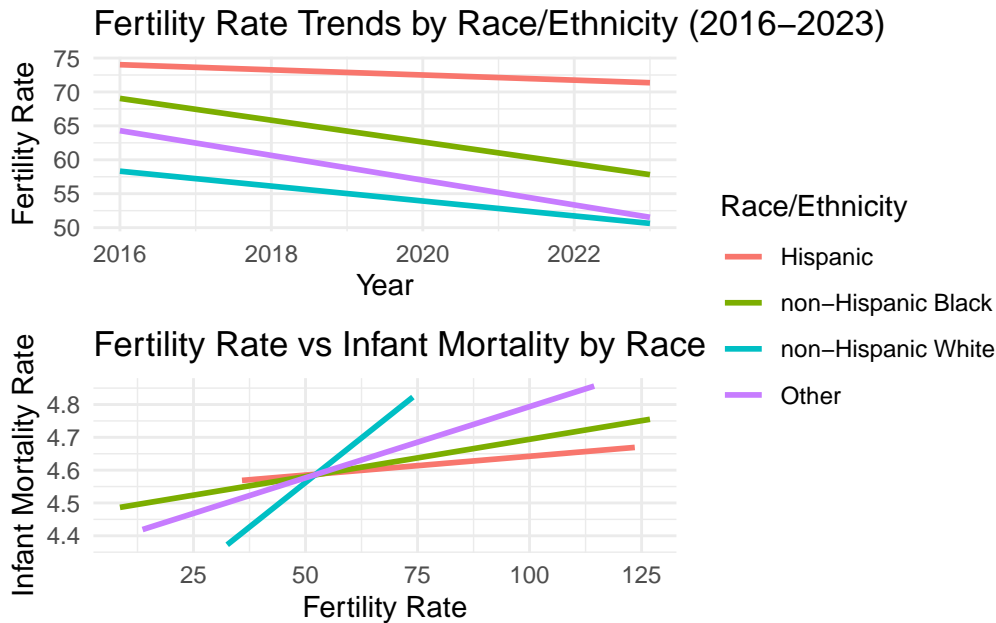


The fertility rates for the population of women whose births are insured by Medicaid shows a declining trend over time, while fertility rates for the population of women whose births are *not* insured by Medicaid are much higher and shows a slight inclining trend over time.

Possible explanations for these trends could be due to certain socioeconomic or demographic factors. For example, Medicaid coverage is primarily for low-income individuals, who may face greater economic instability, higher living costs, and career-related pressures that discourage having children. Women not covered by Medicaid are more likely to have a private health insurance or higher incomes, which might give them greater financial security to have more children. Another factor could be that women covered by Medicaid are often younger, lower-income individuals and may delay having children due to financial insecurity.

For individuals insured by Medicaid, as fertility rates increase, infant mortality also increases sharply. On the other hand, there is a slight negative correlation between fertility rates and infant mortality for individuals not insured by Medicaid, meaning higher fertility is associated with a small decline in infant mortality. This suggests a possible disparity in healthcare quality, access, or other socioeconomic factors affecting infant mortality differently based on insurance status. As an example, women with private insurance (likely in the non-Medicaid group) might receive better prenatal care and maternity support. However, mothers covered by Medicaid may face higher maternal health risks, more stress, and limited healthcare provider choices

Race/ethnicity



The racial/ethnic group of women with the highest fertility rate over the given time period was Hispanic women with a slight decline in the trend. Non-Hispanic Black, non-Hispanic White, and “Other” racial groups all show a steeper decline in fertility rates, with non-Hispanic White women having the lowest rates.

Higher fertility rates correlate with higher infant mortality rates across all racial/ethnic groups. Non-Hispanic White women show the sharpest increase in infant mortality as fertility rises, suggesting that as fertility increases, infant health outcomes may worsen disproportionately for this group. Hispanic women have a consistently high infant mortality rate, even at lower fertility levels, which aligns with well-documented disparities in maternal and infant health outcomes for this group.

Conclusion

The analysis reveals important patterns in fertility and infant mortality rates across different demographic groups in the U.S. from 2016 to 2023. Across all education levels, fertility rates have been gradually declining. Interestingly, women with some college education have the highest fertility rates, followed by those with a college degree, then a high school diploma, and finally those with no high school education. This contradicts the common expectation that higher education leads to lower fertility. However, what remains consistent is the positive correlation between fertility rates and infant mortality—when fertility rises, so does infant

mortality, regardless of education level. This suggests that other factors, such as healthcare access and socioeconomic conditions, may be at play.

Insurance status also plays a major role. Women whose births are covered by Medicaid, a program primarily for low-income individuals, show declining fertility rates over time. However, when fertility does rise in this group, infant mortality rises sharply. On the other hand, women not covered by Medicaid have higher fertility rates overall, and their infant mortality rates remain relatively stable or even decline slightly as fertility increases. This points to disparities in healthcare access, maternal health risks, and overall support systems. Women with private insurance may have better prenatal care, lower stress levels, and greater access to medical resources, whereas those relying on Medicaid might face greater challenges in accessing high-quality maternity care.

Racial and ethnic differences further highlight disparities in maternal and infant health. Hispanic women consistently have the highest fertility rates, but they also experience persistently high infant mortality rates, even at lower fertility levels. Non-Hispanic White women, despite having the lowest fertility rates, show the steepest rise in infant mortality when fertility increases. This suggests that certain groups face unique maternal health challenges, whether due to socioeconomic barriers, healthcare inequalities, or other structural factors.

Ultimately, these findings underscore the deep connections between fertility, healthcare access, and infant health outcomes. While overall fertility rates are declining, infant mortality remains a pressing issue, particularly for economically disadvantaged and marginalized groups. Addressing these disparities requires improving healthcare access, ensuring high-quality maternal care for all women, and tackling the socioeconomic challenges that influence reproductive health. Moving forward, policies and programs that provide better prenatal and postnatal care, reduce healthcare inequities, and support maternal well-being will be crucial in improving infant survival rates and ensuring healthier outcomes for future generations.

The role of Medicaid coverage in shaping both fertility and infant mortality feels like one of the most important factors in this entire analysis. The differences in infant health outcomes between mothers on Medicaid and those who aren't covered by it raise real concerns about healthcare access and quality. It makes me wonder—what challenges are these mothers facing? Are they struggling to find doctors who accept Medicaid? Are they getting the prenatal and postnatal care they need? Are financial stress and limited healthcare options affecting their babies' health?

I want to explore these questions further because they go beyond just numbers on a graph—they reflect the real experiences of mothers and families. If Medicaid-covered births come with a higher risk of infant mortality, then understanding why is crucial. Maybe it's a matter of policy, healthcare provider shortages, or gaps in maternal care. Whatever the reasons, getting to the root of this issue could help shape better healthcare solutions and improve outcomes for mothers and their babies, especially for those who rely on Medicaid the most.