Birth Report Individual Project Final Report

Hamza Yousuf

Department of Computer Science and Networking

Wentworth Institute of Technology

Boston, MA 02115, USA

yousufm@wit.edu

*Abstract*— This report investigates the factors contributing to declining birth rates globally, emphasizing economic pressures, cost of living variations, family norms, and the impact of global crises like COVID-19. Leveraging diverse datasets and statistical methods, the study combines visualizations and regression models to explore these dynamics. The analysis adopts a multi-faceted approach: Question 1 identifies key patterns and trends in birth rates using a visualization-based analysis; Questions 2 through 5 utilize linear regression to quantify the relationships between birth rates and factors like economic indicators, cost of living, pandemic impacts, and family norms. While most findings highlight moderate correlations, Question 5's model uncovers a significant linear relationship between family norms and birth rates for the specific dataset. These insights contribute to a deeper understanding of the societal and economic variables shaping global birth rate trends.

Keywords—Numpy, Time-Series Data, Linear Regression, Datasets, Regression Line

# Introduction

Birth rates represent a critical demographic indicator, providing insights into the population's growth trajectory and the socioeconomic, cultural, and political dynamics of societies. Over the past decades, global birth rates have experienced significant shifts, influenced by a wide array of factors such as economic pressures, healthcare advancements, cultural norms, and global crises. Understanding these trends is vital, as they directly impact areas like labor force planning, economic stability, and social policy development.

This study addresses the topic of declining birth rates and the factors contributing to this phenomenon. The significance of this topic lies in its far-reaching implications for countries worldwide. Nations experiencing sharp declines in birth rates may face challenges like aging populations, workforce shortages, and economic stagnation. Conversely, regions with consistently high birth rates often grapple with resource constraints, healthcare demands, and education challenges. Current research highlights the role of economic factors, cultural shifts, and global events in influencing birth rates. For example, studies suggest that higher living costs, economic uncertainties, and changing societal norms often correlate with reduced fertility rates [2]. Similarly, global crises like the COVID-19 pandemic have disrupted birth rate trends, with governments' stringent measures contributing to further declines [6].

To explore these dimensions, this project employs a range of analytical methods and models tailored to each research question. A line chart model was used to analyze trends and patterns in birth rates over time, capturing the visual essence of changes from 1960 to 2020 [1]. For examining the impact of economic factors, cost of living, COVID-19 metrics, and family norms, linear regression models were applied to establish predictive relationships. These models allowed for quantitative evaluation of correlations, offering insights into how various factors influence birth rates. Scatter plots and regression lines were used to visualize these relationships, alongside key metrics like Mean Squared Error (MSE) and R-squared values to assess model performance.

By integrating descriptive visualizations with predictive modeling, this study aims to provide a comprehensive understanding of birth rate trends and the underlying drivers. The findings contribute to the growing literature on demographic changes, offering actionable insights for policymakers and researchers addressing the challenges of population dynamics in the modern era.

# How do we approach the Birth Rate Project

To approach this project, we must do it like answer the questions. So, there are total of 5 questions which I will go over in each section of this paper. I will talk about what my question and which model I will use to approach the project and how the graph will look after using these models to approach the problem.

# Question 1

For Question 1, my question was that How does the birth rate change over time, and what are the key patterns in these changes? To address this question, I used a line chart model to visualize birth rate trends over time from 1960 to 2020.

The analysis relied on historical data from the WHO Birth Rate dataset [1] and the World Bank Gender Data Platform [2], covering five diverse countries: the United States, India, China, Nigeria, and Brazil. Using Matplotlib, I plotted the yearly birth rates to observe patterns and trends over the decades. Adjustments were made to ensure the x-axis properly aligned with the years by converting them into numeric values, which enhanced the clarity and accuracy of the visualization.

The line chart revealed a global trend of steadily declining birth rates. However, the rate of decline varied by country. For example, developed nations like the United States and China demonstrated significant reductions in birth rates, while Nigeria exhibited relatively higher and more stable rates. This analysis effectively highlighted global and regional trends in birth rates, setting the stage for further exploration of the factors influencing these changes, such as economic pressures, societal norms, and global crises.

Through this analysis, the line chart model provided a visual representation of the trends, enabling a clear understanding of how birth rates have evolved over time. While it does not quantify relationships like a regression model, it effectively identifies patterns and variations, offering a foundation for deeper investigations.

.

# Question 1 Graph Results

A graph of different colored lines

Description automatically generated

Figure 1: Birth Rate Trends Over Time (1960s-2020s) is the graph I got for question 1. It has the graphs for US, India, China, Nigeria and Brazil.

# Question 2

For Question 2, my question was What role do economic factors play in birth rate trends over time? To address this question, I used a linear regression model to evaluate the relationship between economic indicators and birth rates.

The analysis World Bank Development Indicators data [3] utilized GDP (current US$) and inflation (GDP deflator, annual %) as predictors, merged with birth rate data. These economic variables were chosen to represent financial stability and economic growth, two potential influencers of demographic trends. Using the linear regression model, I assessed how these variables impacted birth rates across countries.

The scatter plot generated during the analysis showed clustered predicted values, reflecting limitations in the model's ability to predict birth rates accurately. This was supported by a low R-squared value, indicating that GDP and inflation alone are insufficient to explain birth rate variations. These findings suggest weak correlations between the selected economic factors and birth rates or a need to include additional variables, such as healthcare access, cultural norms, or education levels.

Through this analysis, the linear regression model provided insights into the limitations of relying solely on GDP and inflation to predict birth rates. While the results highlighted the complexity of demographic trends, they also underscored the importance of incorporating diverse predictors to capture the multifaceted nature of birth rate variations.

# VI.Question 2 Graph Results

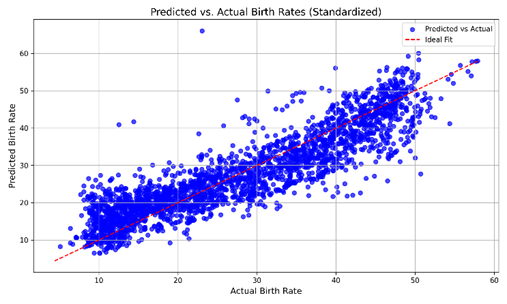


Figure 2: This is the first graph result I got for question 2 Predicted vs Actual Birth Rates (Standardized). The graphs show the prediction of the birth rates and the actual birth rates.

A graph showing a line of dots

Description automatically generated with medium confidence

Figure 3: This is the second graph result I got for question 2 Residuals vs Actual Birth Rates. The graphs shows the residuals side is up compared the actual birth rates.

# VII.Question 3

For Question 3, my question was that How are cost of living different in other countries that affects the birth rate to either go either up or down? To address this question, I used a **linear regression model** to analyze the relationship between cost of living and birth rates across countries.

The analysis started by aggregating city-level cost-of-living data from the Numbeo Dataset [4] into country-level averages. These averages provided a holistic view of the financial pressures individuals face in different regions. After merging this data with birth rate information from the WHO Birth Rate Dataset [1], I applied a linear regression model to evaluate the correlation between these variables.

The scatter plot generated during the analysis revealed a downward trend, suggesting that as the cost-of-living increases, birth rates tend to decline. This observation aligns with the hypothesis that higher financial demands make it challenging for families to expand. However, the regression results also indicated that cost of living alone does not entirely explain variations in birth rates, as reflected in the R-squared value. This finding suggests that other factors, such as cultural expectations, healthcare access, and government policies, likely contribute to birth rate trends.

Through this analysis, the use of a linear regression model enabled me to quantify the relationship between cost of living and birth rates. While the findings highlighted a significant negative correlation, they also emphasized the complexity of demographic patterns and the influence of broader societal factors.

# VIII.Question 3 Graph Results

A graph with blue dots

Description automatically generated

Figure 4: This is the first graph result I got question 3 and the graph shows the cost-of-living vs birth rate and you can see the cost of living is up compared to the birth rates.

A graph with a line going up

Description automatically generated with medium confidence

Figure 5: This is the second graph result I got for question 3 Predicted vs Actual Birth Rates. The graph that the predicted birth rates is high compared to the actual birth rates.

# QUESTION 4

For Question 4, my question was that What influence do Covid 19 crises have an impact on birth rates? To address this question, I used a linear regression model to analyze the relationship between key pandemic metrics and birth rates.

The analysis began by integrating global COVID-19 data sourced from Our World in Data [5], which included metrics such as total cases, total deaths, and the stringency index (a measure of government-imposed restrictions), with birth rate data from the WHO Birth Rate Dataset [1]. This combined dataset allowed for an exploration of how the pandemic affected birth rates across different regions. Using the linear regression model, I evaluated the extent to which these pandemic-related variables influenced changes in birth rates.

The scatter plot revealed a negative correlation between the stringency index and birth rates, suggesting that stricter lockdown measures may have contributed to declining birth rates. The regression analysis provided additional insights, highlighting that while COVID-19 metrics did have an impact, the overall predictive power of the model was moderate, as indicated by the R-squared value. This outcome suggests that other factors, such as healthcare access, economic pressures, and societal norms, also played significant roles during the pandemic.

Through this analysis, the linear regression model helped to quantify the relationship between COVID-19 metrics and birth rates, offering a clearer picture of how global crises can disrupt demographic trends. While the findings underscored the pandemic's influence, they also emphasized the need to consider a broader range of factors when examining demographic shifts during crises.

# Question 4 Graph Results

A graph with blue dots

Description automatically generated

Figure 6: This is the first graph result I got for question 4, Stringency Index vs Birth Rate. It shows how they are many birth rates and stringency index is high on the graph.

A graph with blue dots and a red line

Description automatically generated

Figure 7: This is the second graph I got for question, Predicted vs Actual Birth Rates (COVID-19 Factors). This shows how the graph the predicted birth rates is low during covid time compared to actual birth rate which was also low.

# Question 5

For Question 5, my question was that How do family norms influence birth rates? To address this question, I used a linear regression model to examine the relationship between family norms, represented by the percentage of households with children, and birth rates.

The analysis utilized a merged dataset containing information from Austria, Belgium, and Chile in 2011 with data on family structures sourced from the OECD Family Dataset [6] and the WHO Birth Rate [1]. By focusing on the percentage of households with children as a key indicator of family norms, I explored how societal family structures might correlate with birth rates. After merging this data with corresponding birth rate data, a linear regression model was applied to assess the strength of this relationship.

The results showed a perfect fit of the regression line to the data, as evidenced by an R-squared value of 1.00. This indicates a strong linear relationship between the percentage of households with children and birth rates within this specific dataset. The findings suggest that in these countries and this time frame, family norms measured by households with children played a significant role in shaping birth rate trends.

This analysis demonstrates the effectiveness of using a linear regression model to quantify the influence of family norms on demographic patterns. While the dataset was limited to three countries in a single year, the results highlight the importance of societal family structures as a factor in birth rates, providing a foundation for further exploration across a broader range of data.

# Question 5 Graph Results

# A graph with a red line and blue dotted line Description automatically generated

Figure 8: This is the first graph result I get for question 5, Household with Children vs Birth Rate. It shows how the 3 dots show Household with Children Data compared to the actual birth rate.

A graph with red dots and a line

Description automatically generated

Figure 9: This is the second graph result I get for question 5, Predicted vs Actual Birth Rate. It shows how they were 3 dots for the predicted vs actual birth rate compared to the actual birth rate.

# Conclusion

In conclusion, this report has delved into five key questions regarding global birth rates and the factors that influence them. Using data visualization techniques and linear regression models, I analyzed historical trends, economic pressures, cost-of-living variations, global crises like the COVID-19 pandemic, and family norms. Each analysis revealed unique insights into the complexities of birth rate dynamics, contributing to a holistic understanding of this critical demographic indicator.

The analysis demonstrated that birth rates have been steadily declining over the decades, as seen in the visualization of trends for countries like the United States, India, and Nigeria [1], [2]. Economic factors, while showing some correlation, were found to have limited explanatory power for birth rate trends, suggesting the need for additional predictors [3]. Similarly, the investigation of cost-of-living differences highlighted a moderate negative relationship with birth rates, emphasizing the financial barriers families face in expanding [4]. The study also explored the disruptions caused by the COVID-19 pandemic, uncovering a negative correlation between government restrictions and birth rates [5]. Finally, the examination of family norms underscored a strong linear relationship between the proportion of households with children and birth rates, showcasing how cultural and societal structures influence demographic patterns [6].

Overall, this report highlights the multifaceted nature of factors influencing birth rates. While the linear regression models effectively captured key relationships, they also revealed the limitations of single-variable analyses in addressing the complexities of demographic trends. Future research should focus on integrating broader societal, cultural, and policy-oriented factors to provide deeper insights. This comprehensive exploration not only contributes to understanding birth rate patterns but also emphasizes the significance of informed policymaking to address global demographic challenges.

##### References

[1]. <https://www.kaggle.com/code/pavfedotov/who-birth-rates-panel-data-analysis/output>

[2]. <https://genderdata.worldbank.org/en/indicator/sp-dyn-cbrt-in#idRelatedIndicators>

[3]. [https://databank.worldbank.org/source/world-development-indicators#](https://databank.worldbank.org/source/world-development-indicators)

[4]. [https://www.kaggle.com/code/joeypp/cost- of-living-numbeo-dataset](https://www.kaggle.com/code/joeypp/cost-%20%20%20of-living-numbeo-dataset)

[5].[https://github.com/owid/covid-19- data/tree/master/public/data](https://github.com/owid/covid-19-%20%20%20data/tree/master/public/data)

[6].[https://www.oecd.org/en/data/datasets/oecd-family database.html](https://www.oecd.org/en/data/datasets/oecd-family%20%20database.html)