Exploring Newborn Trends in Brazil (2008-2021)

Ludmilla M O Sousa

05 December, 2023

Contents

Introduction

In Brazil, there are Health Information Systems, consisting of various databases available on the internet that can be used to support health management (NETO; CHIORO, 2021). These volumes of available data serve as powerful sources to assist managers in making informed decisions. Despite challenges related to data dispersion across different platforms, the integration of a data-driven culture can revolutionize health management, enabling the identification of trends and patterns, population health monitoring, outbreak and disease identification, more effective resource allocation, personalized treatments, and more (ANDRADE et al., 2012; KAUR; WASAN, 2006).

However, these valuable data sources are underutilized by managers due to the lack of an established information culture and a shortage of professionals skilled in handling data (QUITES, 2016). The lack of clarity and relevant insights can lead to misguided or inefficient choices, directly impacting the performance and effectiveness of actions taken. In an era of technological advancements, it is crucial to seek more efficient ways to incorporate these databases into decision-making, as this can significantly enhance the quality of health services offered.

In this context, data science plays a fundamental role in advancing health management. The various stages of the analytical process, from data collection to the implementation of solutions, play a crucial role in advancing healthcare services (SUBRAHMANYA et al., 2022). Thus, data coherence supports health managers in making informed decisions, promoting the development of healthcare services and, consequently, the Unified Health System (SUS).

Concerning pregnancy and childbirth, it is important to highlight that these are natural events in human reproduction, and adequate assistance during these moments is vital for the well-being of the woman, the baby, and everyone involved. Therefore, childbirth is the most performed procedure in Brazil (DATASUS, 2015). This reality makes the healthcare service related to childbirth a central area in health planning, deserving special attention and well-elaborated strategies to ensure the quality and accessibility of this fundamental service.

In the field of maternal and child health, national public policies and actions aim to ensure adequate assistance during childbirth and delivery (BRAZIL, 2000, 2004, 2007, 2010, 2011a). These initiatives include negotiation forums among managers, seeking the integration of healthcare services in a network (BRAZIL, 2011b).

However, Brazil still faces serious issues with access to childbirth, leading some women to visit more than one maternity hospital to find assistance. The increased distance to access childbirth services can result in a higher risk of infant mortality (ALMEIDA; SZWARCWALD, 2012). Another concerning aspect regarding childbirth in Brazil is the high rate of cesarean sections (55%), well above the World Health Organization's (WHO) recommended rate of 15%. It is known that the unnecessary performance of cesarean sections can have harmful effects on both the mother and the baby.

Given this scenario, data science can be used as a tool to monitor the childbirth situation in Brazil, identify risks, and promote improvements to enhance safety during childbirth. As I am part of the state health management in Bahia/Brazil, this research will be focused on childbirths conducted in Bahia. I hope that through this work, we can gather information to contribute to the advocacy for safe childbirth in Bahia.

Method

To conduct this research, data from the Public Health System's Hospital Information System (SIH) in Brazil were downloaded. These data are publicly available on the Datasus website (https://datasus.saude.gov.br/transferencia-de-arquivos/). The data is initially provided in DBC format, and it is first tabulated using a program called Tabwin, which is also in the public domain. Subsequently, the tabulated data is downloaded for importation into RStudio. Due to the enormous size of the database, it was necessary to reduce it for sharing on GitHub. In this research, it was placed within the repository https://github.com/millamonfort/BirthBA.git. The datasets available on GitHub, within the 'dataBithsBA' folder, are separated by year as each file was too large. They are provided in CSV format. After downloading the datasets, they are merged into a single dataset to facilitate analysis.

After the initial treatment, the dataset has the following characteristics:

```
## Rows: 1,761,691
## Columns: 10
## $ MUNIC_RES     <int> 291390, 291390, 291390, 291390, 291390, 291520, 291390, ~
## $ Municipio_res <chr> "Ipiau", "Ipiau", "Ipiau", "Ipiau", "Ipiau", "Itagiba", ~
## $ MUNIC_MOV     <int> 291390, 291390, 291390, 291390, 291390, 291390, 291390, ~
## $ Municipio_mov <chr> "Ipiau", "Ipiau", "Ipiau", "Ipiau", "Ipiau", "Ipiau", "I~
## $ PROC_REA      <int> 411010034, 411010034, 411010034, 411010034, 411010034, 4~
## $ DT_INTER      <int> 20110109, 20110110, 20110111, 20110112, 20110111, 201101~
## $ DT_SAIDA      <int> 20110111, 20110113, 20110113, 20110114, 20110113, 201101~
## $ IDADE         <int> 32, 28, 26, 31, 21, 23, 23, 26, 28, 20, 18, 23, 23, 32, ~
## $ DIAS_PERM     <int> 2, 3, 2, 2, 2, 4, 1, 1, 1, 1, 2, 2, 1, 1, 2, 2, 1, 1, 1,~
## $ CNES          <int> 2603055, 2603055, 2603055, 2603055, 2603055, 2603055, 26~
```

The dataset consists of 1761691 rows and 10 columns, where each row represents a record of a woman's childbirth.

To facilitate the identification of the type of delivery performed, a variable identifying the childbirth procedure by name was added to the database. Additionally, to enable temporal comparisons of childbirth volume,

variables for the month and year of admission (admMonth) and the year of admission (admYear) to the hospital were included.

At the conclusion of these procedures, which involved modifying the dataset, below is an overview of the updated dataset. It now comprises 15 columns.

```
## Rows: 1,761,691
## Columns: 15
## $ MUNIC_RES     <int> 291390, 291390, 291390, 291390, 291390, 291520, 291390, ~
## $ Municipio_res <chr> "Ipiau", "Ipiau", "Ipiau", "Ipiau", "Ipiau", "Itagiba", ~
## $ MUNIC_MOV     <int> 291390, 291390, 291390, 291390, 291390, 291390, 291390, ~
## $ Municipio_mov <chr> "Ipiau", "Ipiau", "Ipiau", "Ipiau", "Ipiau", "Ipiau", "I~
## $ PROC_REA      <chr> "411010034", "411010034", "411010034", "411010034", "411~
## $ DT_INTER      <int> 20110109, 20110110, 20110111, 20110112, 20110111, 201101~
## $ DT_SAIDA      <int> 20110111, 20110113, 20110113, 20110114, 20110113, 201101~
## $ IDADE         <int> 32, 28, 26, 31, 21, 23, 23, 26, 28, 20, 18, 23, 23, 32, ~
## $ DIAS_PERM     <int> 2, 3, 2, 2, 2, 4, 1, 1, 1, 1, 2, 2, 1, 1, 2, 2, 1, 1, 1,~
## $ CNES          <int> 2603055, 2603055, 2603055, 2603055, 2603055, 2603055, 26~
## $ name_PROC     <chr> "Cesarean section", "Cesarean section", "Cesarean sectio~
## $ DT_ADMISSION  <date> 2011-01-09, 2011-01-10, 2011-01-11, 2011-01-12, 2011-01~
## $ DT_EXIT       <date> 2011-01-11, 2011-01-13, 2011-01-13, 2011-01-14, 2011-01~
## $ admYear       <dbl> 2011, 2011, 2011, 2011, 2011, 2011, 2011, 2011, 2011, 20~
## $ admMonth      <date> 2011-01-01, 2011-01-01, 2011-01-01, 2011-01-01, 2011-01~
```

In this research, two prediction algorithms were applied to visualize the quantity of births in a few years. The first one was the "forecast," and the second one was the "random forest."

In the birth research, we employed the ARIMA (AutoRegressive Integrated Moving Average) method to forecast the total number of births in Bahia for the next 10 years, as well as predict trends for Vaginal Delivery and Cesarean Section for the next 20 years. Here's a simplified explanation:

- Data Preparation:We gathered data on the number of births for each month.
- Time Series Object: We organized this data into a time series, treating each month as a separate point in time.
- ARIMA Model: We applied an ARIMA model, which is a type of statistical model for time series data. The "auto" part means the model automatically selects the best configuration based on the data.
- Forecasting:The ARIMA model was then used to predict the number of births for the next 10 years.
- Visualization: Finally, we created a plot to visualize these predictions.

This approach helps us understand and anticipate trends in the number of births, aiding in future planning and decision-making.

We explored alternative predictive algorithms, utilizing Random Forest and Linear Regression models to forecast the number of births in Bahia over the next 20 years. Let's delve into the essential steps of this analysis.

We began by organizing the data, summarizing the monthly birth counts and extracting the corresponding years. This helped create a clear timeline of births over the years.

To build our predictive models, we divided the dataset into training and testing sets, with 80% of the data designated for training. In preparation for modeling, we added a quadratic term to capture potential non-linear trends in the birth data.

The Random Forest model was then employed to predict future births. This model is known for its ability to handle complex relationships and non-linear patterns. Simultaneously, we developed a Linear Regression model, which assumes a linear relationship between the variables.

Predictions were made for the next 20 years using both models, allowing us to compare their performance. The visualization of predictions and historical data provides insights into how well each model captures the expected trends in birth rates.

By employing these different modeling techniques, we aimed to gain a comprehensive understanding of the future birth scenario in Bahia, considering both linear and non-linear factors that might influence the birth rates. This dual-model approach enhances the reliability of our predictions and provides a more nuanced perspective on future birth trends.

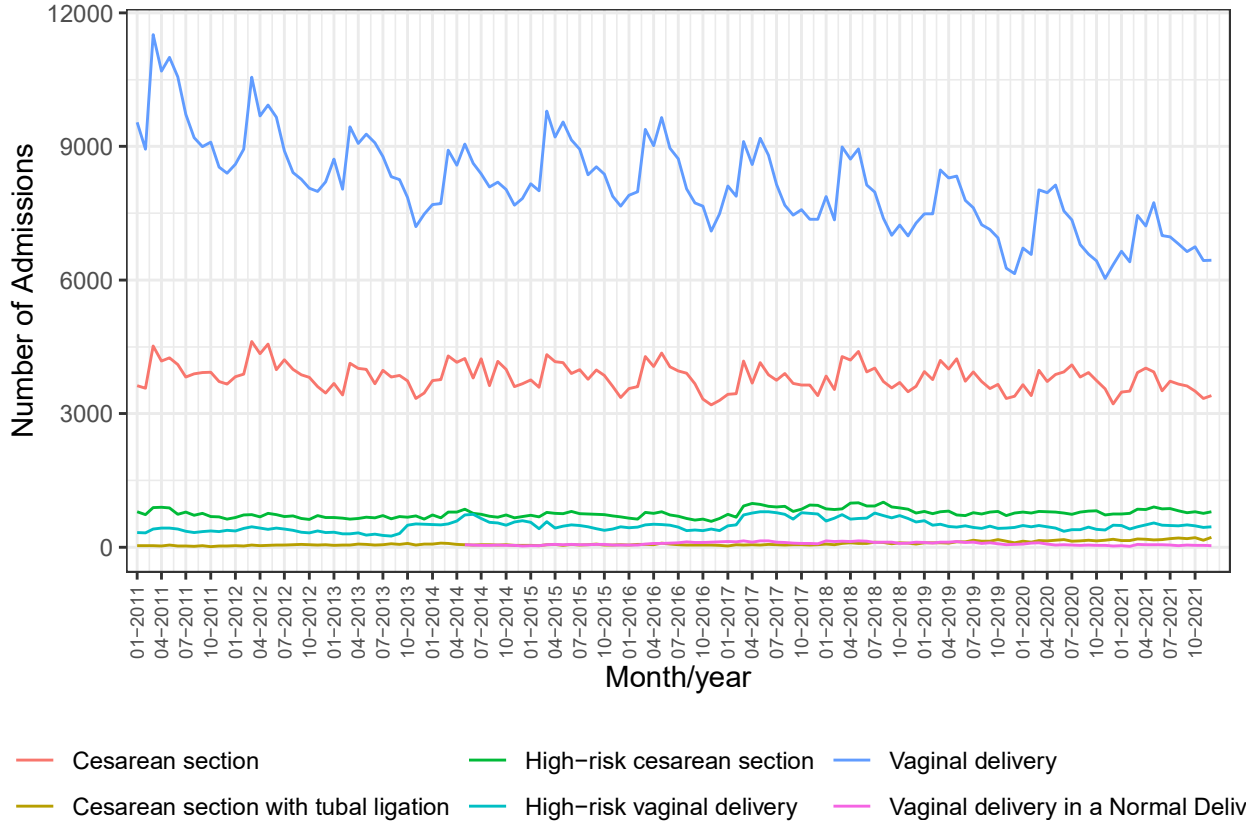
Result

Exploratory Analysis

For a better understanding of the data comprising the dataset we will be working with, we will begin with a brief data exploration. In this stage, we aim to gain an overview of the data, identify trends, and enhance our familiarity with its characteristics.

Following this, we will employ two predictive algorithms on the dataset.

Graph 1 - Delivery Admissions Over Time in Bahia, Brazil (2008-2021)



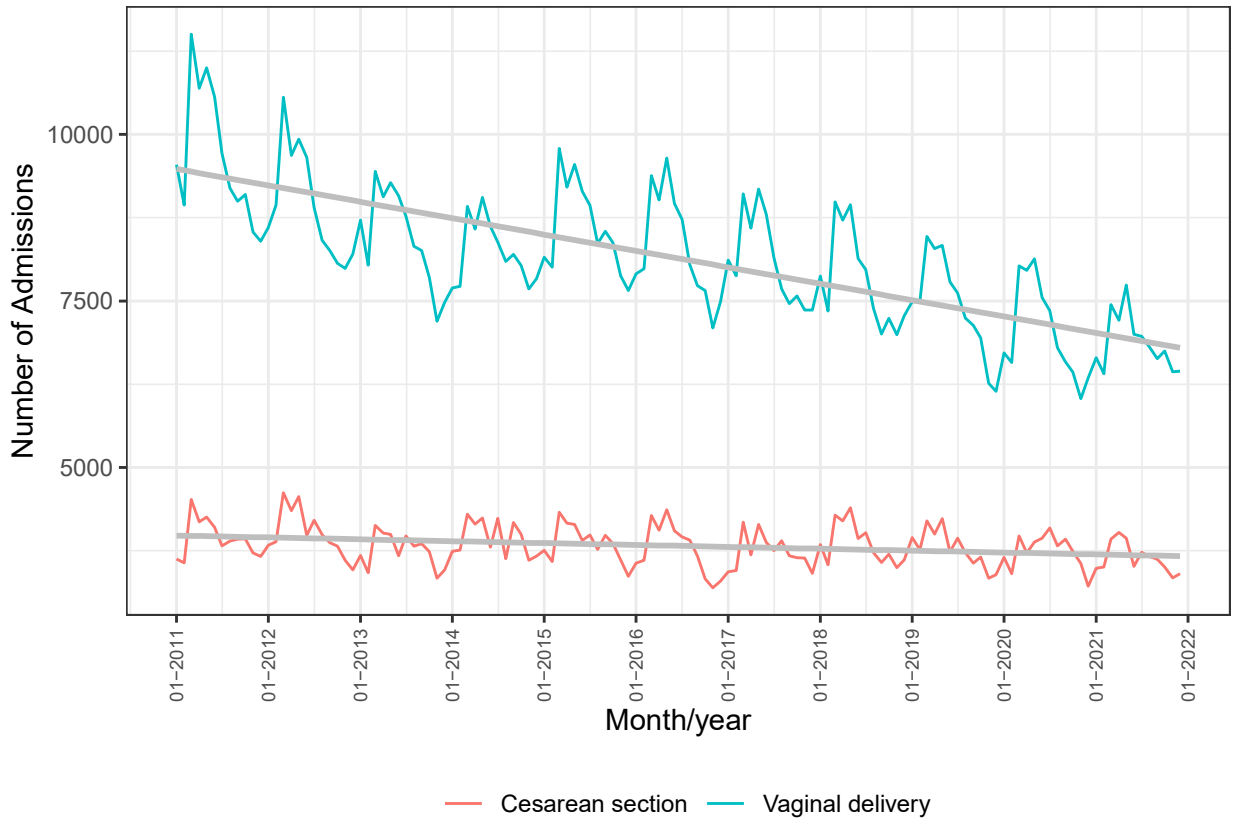
The analysis of childbirth types over the historical series in Bahia reveals that vaginal delivery is the most frequently performed, followed by cesarean section. This trend can be observed in Graph 1, which presents data from 1761691 childbirth records from 2011 to 2021. During this period, there was a significant decrease in the total number of records, dropping from 177386 in 2011 to 144359 in 2021, representing a decrease

of 18.62%. Specifically regarding vaginal delivery, the number of records decreased from 116182 in 2011 to 82497 in 2021, a decrease of 29%. On the other hand, the number of cesarean sections increased from 47206 in 2011 to 43651 in 2021, showing an decrease of 7.5%. These results indicate a declining trend over time for vaginal delivery, which is more pronounced than the trend for cesarean sections, and the gap between these two delivery methods is narrowing.

An interesting insight to draw from this graph is the apparent presence of certain months with periodic patterns, indicating possible annual repetitions, particularly related to normal deliveries. These analyses will be further explored in the upcoming graphs.

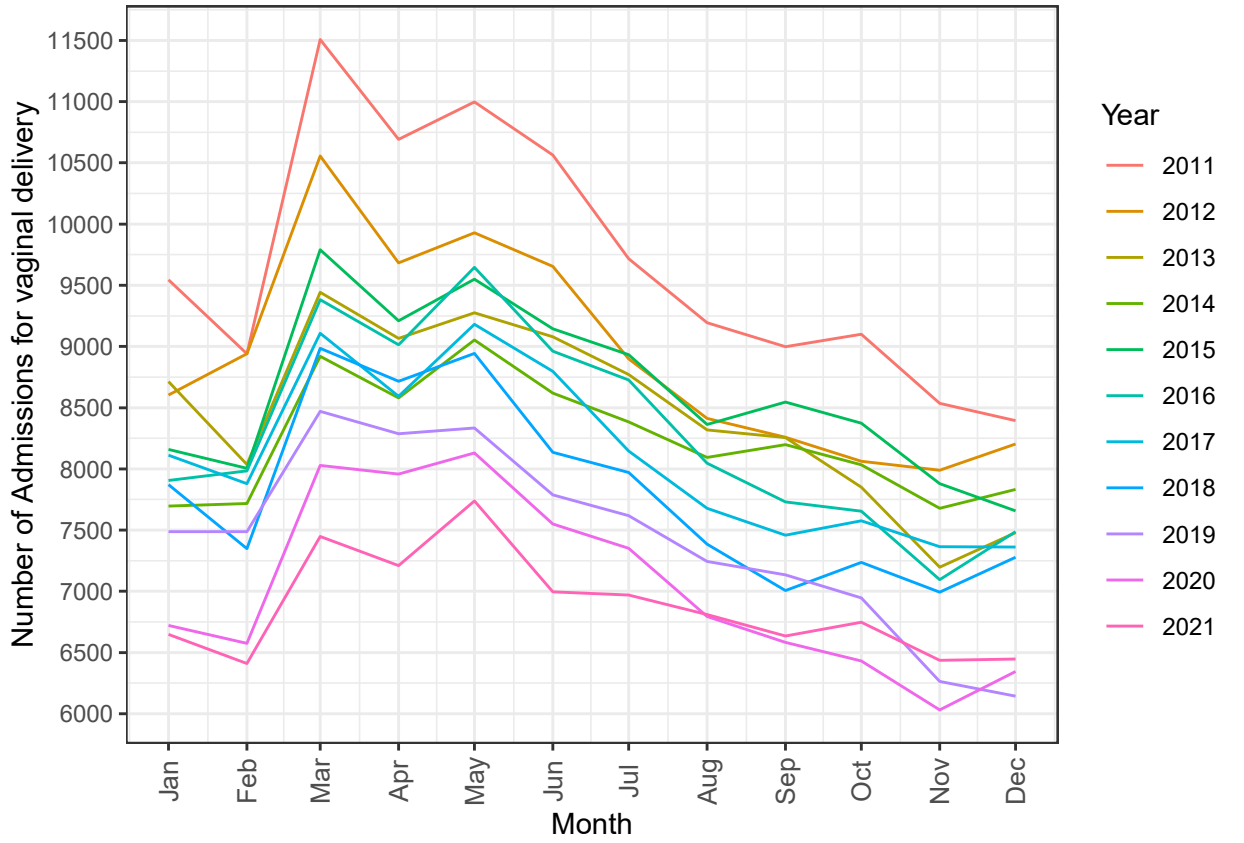
In Brazil, there is a decline in the fertility rate, and according to the data from the Hospital Information System (SIH) for deliveries, this trend is also observed in Bahia.

Graph 2 - Linear Regression of Vaginal Delivery and Cesarean Section in Bahia (2011-2021)



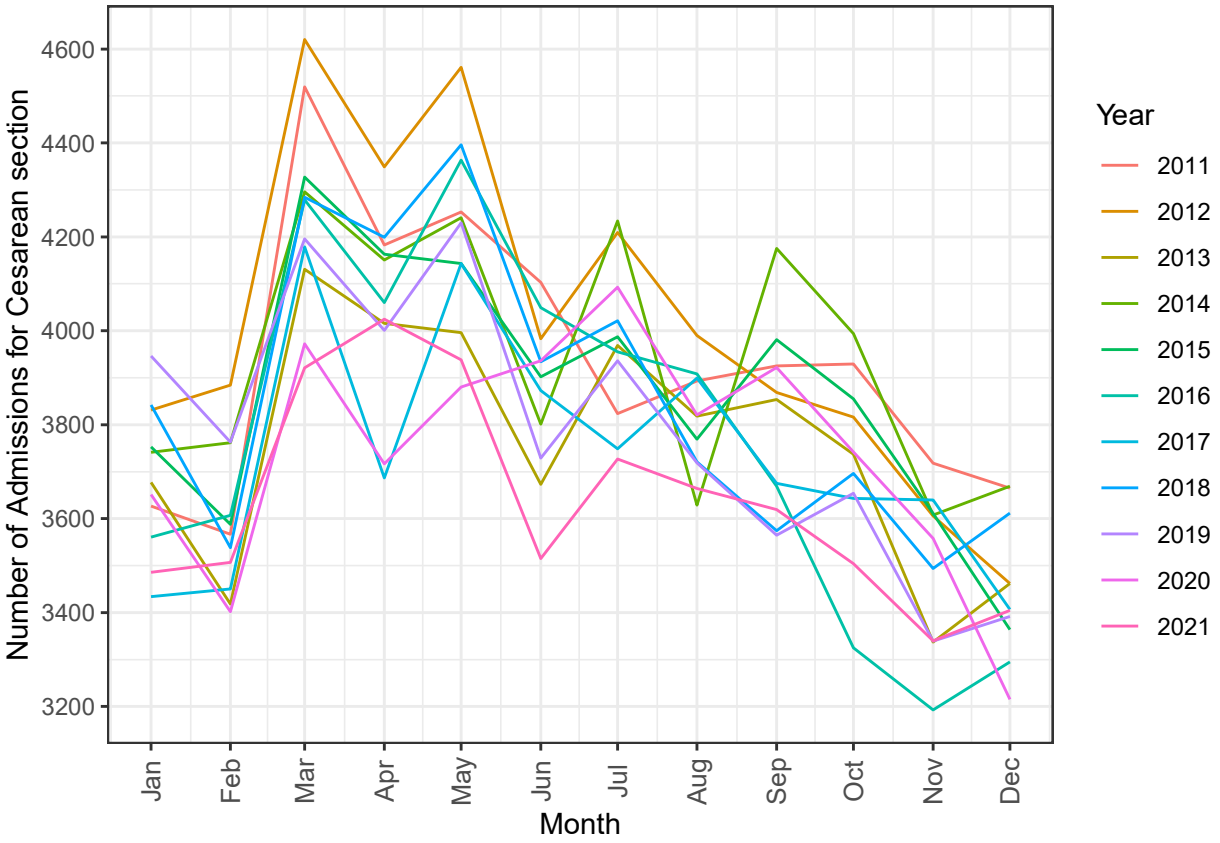
In Graph 2, the decreasing trend of vaginal deliveries becomes more evident, showing a more pronounced decline compared to cesarean sections.

Graph 3 - Monthly Count of Normal Delivery Hospitalizations. Bahia, from 2008 to 2021



For a more in-depth analysis of monthly patterns, we present the graphs in Figure 3, illustrating the variation in normal deliveries within the SIH over the months. We can observe a significant increase in both deliveries and births in the months of March and May, across all analyzed years, followed by a decline after May. Additionally, these graphs more clearly highlight the decrease in the number of normal deliveries over the years.

Graph 4 - Monthly Count of Cesarean Section Hospitalizations. Bahia, from 2008 to 2021

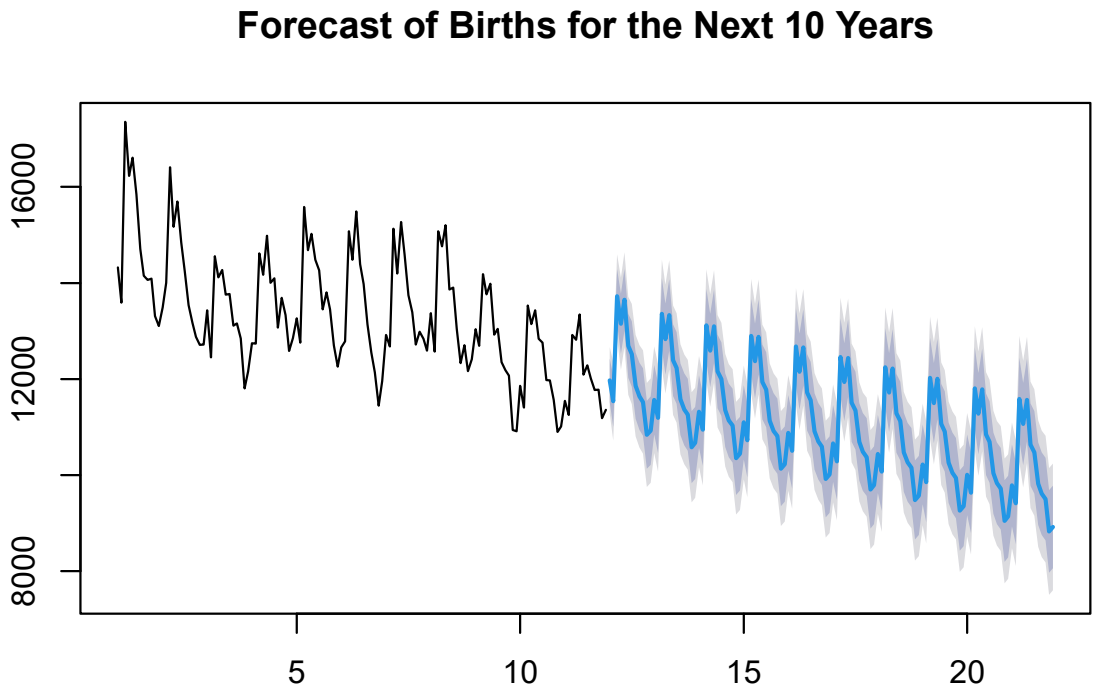


According to Graph 4, we can observe a certain monthly pattern in births, although not as evident as in vaginal deliveries. Another important insight from this graph is the decline over the years, which is much less pronounced than the decline in vaginal births.

Predictive models

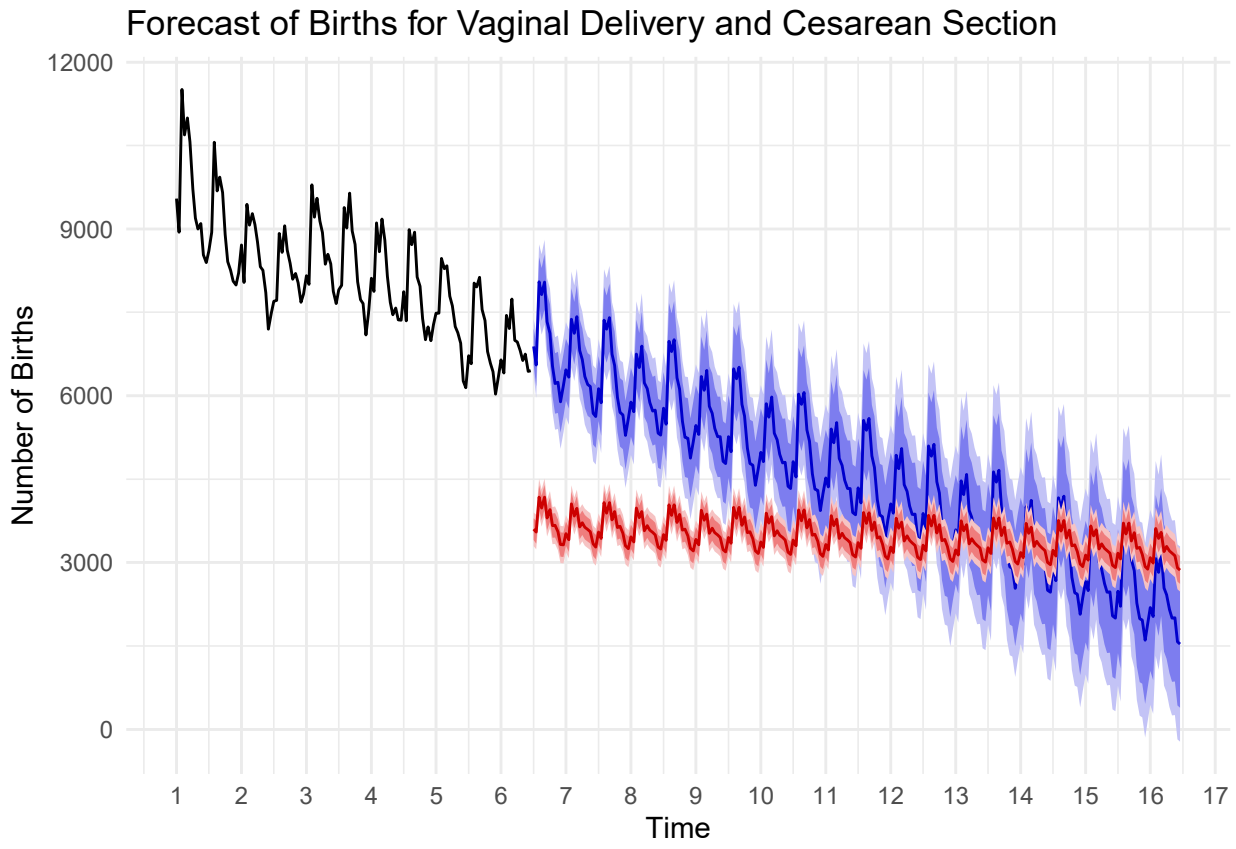
In this section, we will explore some graphs that depict forecasts related to childbirth in Bahia.

Graph 5 - Predictive Model for Birth Trends in Bahia, Brazil for the Next 10 Years



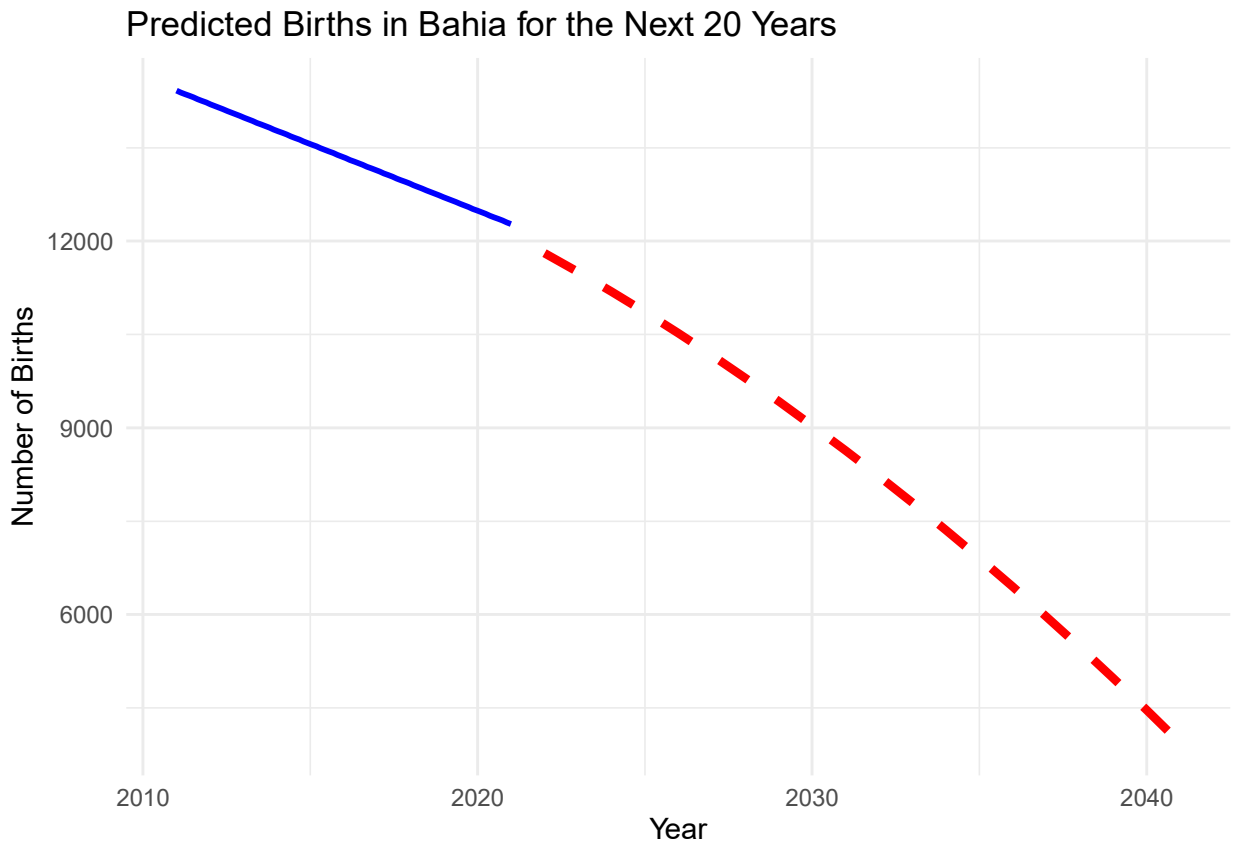
The graph 5 suggests a potential decline in the number of births in the future based on historical patterns. This information could be valuable for policymakers, healthcare providers, and researchers who are interested in understanding and planning for demographic changes.

Graph 6 - Predictive Model for Vaginal Delivery and Cesarean Section Trends in Bahia, Brazil for the Next 20 Years



This graph above visualizes the forecast of births for Vaginal Delivery and Cesarean Section over the next 20 years, highlighting the distinct trends between the two delivery methods. The most notable feature is the divergence in trends between Vaginal Delivery and Cesarean Section. The line for Vaginal Delivery shows a sharp and consistent decrease, indicating a declining trend over the forecast period.In contrast, the line for Cesarean Section remains relatively stable and does not exhibit the same sharp decline. This suggests that while Vaginal Delivery is decreasing, Cesarean Section rates may be maintaining a more consistent level. The graph hints at a historical point where both Vaginal Delivery and Cesarean Section lines match each other.The graph raises questions about the societal factors influencing the choice of delivery method. Factors such as cultural shifts, medical advancements, and changes in maternal preferences may contribute to the observed patterns.

Graph 7 - Predictive Model with random forest and linear regression for Birth Trends in Bahia, Brazil for the Next 20 Years



The graph 7 illustrates a projection of births in Bahia for the next 20 years based on a Random Forest regression model with a quadratic term. While the model suggests a decreasing trend, it is crucial to note that the future birth rate is subject to various factors, and the observed decline may not necessarily continue linearly. Birth rates can be influenced by dynamic societal, economic, and healthcare factors, and the projection serves as a model-based estimation rather than a definitive prediction. Understanding the potential for stabilization or changes in influencing factors is important for a more comprehensive interpretation of the projected trend.

Conclusion

In this research, we present patterns of childbirth in Bahia and the decreasing trend over the years. It is essential to observe the childbirth model being adopted in Brazil; data shows that the more interventionist model is predominant, as evidenced by the sharp decline in vaginal births and a certain stability in cesarean section deliveries.

We also highlighted some future trends in these births. Both are declining, but one algorithm alone is not sufficient to predict the future. Nevertheless, it is crucial for shaping public policies.

References

ALMEIDA, W. S.; SZWARCWALD, C. L. Mortalidade infantil e acesso geográfico ao parto nos municípios brasileiros. Revista de Saúde Pública, v. 46, n. 1, p. 68–76, 2012.

ANDRADE, C. T. DE et al. A importância de uma base de dados na gestão de serviços de saúde. Einstein (São Paulo), v. 10, n. 3, p. 360–365, set. 2012.

BRASIL. Portaria no 569, de 01 de junho de 2000 - Institui o Programa de Humanização no Pré-natal e Nascimento, no âmbito do Sistema Único de Saúde. Brasília: Ministério da Saúde: [s.n.]. Disponível em: http://bvsms.saude.gov.br/bvs/saudelegis/gm/2000/prt0569_01_06_2000.html.

BRASIL. Pacto Nacional pela redução da Mortalidade Materna e Neonatal. Brasília: Ministério da Saúde: [s.n.]. Disponível em: http://www.saude.ba.gov.br/dab/Pacto_Aprovado_na_tripartite.pdf.

BRASIL. Lei no 11.634, de 27 de dezembro de 2007. Dispõe sobre o direito da gestante ao conhecimento e a vinculação à maternidade onde receberá assistência no âmbito do Sistema Único de Saúde. Brasília: Casa Civil: [s.n.]. Disponível em: http://www.planalto.gov.br/ccivil_03/_Ato2007-2010/2007/Lei/L11634.htm. Acesso em: 7 jul. 2021.

BRASIL. Pacto pela Redução da Mortalidade Infantil no Nordeste e Amazônia Legal: 2009-2010. Brasília: Ministério da Saúde, 2010.

BRASIL. Portaria no 1.459, de 24 de junho de 2011 - Institui, no âmbito do Sistema Único de Saúde - SUS - a Rede Cegonha. Brasília: Ministério da Saúde: [s.n.]. Disponível em: http://bvsms.saude.gov.br/bvs/saudelegis/gm/2011/prt1459_24_06_2011.html.

BRASIL. Resolução no1, de 29 de setembro de 2011 - Estabelece diretrizes gerais para a instituição de Regiões de Saúde no âmbito do Sistema Único de Saúde (SUS), nos termos do Decreto No 7.508, de 28 de junho de 2011. Brasília, 2011b. Disponível em: https://bvsms.saude.gov.br/bvs/saudelegis/cit/2011/res0001_29_09_2011.html. Acesso em: 6 jul. 2021

KAUR, H.; WASAN, S. K. Empirical Study on Applications of Data Mining Techniques in Healthcare. Journal of Computer Science, v. 2, n. 2, p. 194–200, 2006.

QUITES, H. F. DE O. Barreiras do uso da Informação em Saúde na tomada de decisão municipal: uma Revisão de Literatura. Revista Eletrônica Gestão & Saúde, v. 07, n. Supl. 7, p. 1011–22, 2016.

NETO, G. C. C.; CHIORO, A. Afinal, quantos Sistemas de Informação em Saúde de base nacional existem no Brasil? Cadernos de Saude Publica, v. 37, n. 7, p. e00182119, 2021.

SUBRAHMANYA, S. V. G. et al. The role of data science in healthcare advancements: applications, benefits, and future prospects. Irish Journal of Medical Science (1971 -), v. 191, n. 4, p. 1473–1483, 16 ago. 2022.