Application of Formal Concept Analysis and Data Mining to Characterize Infant Mortality in Two Regions of the State of Minas Gerais

Deivid Santos, Cristiane Nobre, Luis Zarate and Mark Song
Instituto de Ciências Exatas e Informática, Pontifícia Universidade Católica de Minas Gerais, Brazil

Keywords: Formal Concept Analysis, Infant Mortality, SIM, SINASC.

Abstract: Infant mortality is characterized by the death of children under one year, a problem that affects a large part of the world population. This article applies the Formal Concept Analysis (FCA), a mathematical technique used in data analysis to characterize infant mortality in two regions of Minas Gerais state - Brazil: Belo Horizonte and Vale do Jequitinhonha. The Metropolitan Region of Belo Horizonte has the best human development rate, and Vale do Jequitinhonha has the worst social equality. The relationships between attributes and victims are identified through association rules and implications.

1 INTRODUCTION

Infant mortality is a term used to describe the death of children during the first year of life. It is a circumstance that occurs all over the world, more regularly in poorer nations where one of the factors is low human development index and the lack of basic sanitation, leading to an absence of healthy livelihoods and the spreading of different diseases. The Infant Mortality Rate is characterized as the number of deaths of children under one year per thousand live births in a specific population and year. As such, it is an indicator that assesses the danger of a child conceived alive, dying before completing the first year of life.

Infant mortality can be understood as the sum of deaths occurring in the early neonatal period (0 to 6 days of life), late neonatal (7 to 27 days), and postneonatal period (28 days or more, up to a year) (UNICEF, 2020).

The study of infant mortality can reveal insights into which perspectives should be worked on in a population, with the aim that the extent of deaths can be reduced, which is a definite factor for a country's advancement. From a logical and social perspective, infant mortality can be used as an approach to assess networks and welfare strategies adopted in a determined region, as indicated by (Black et al., 2017). Besides that, the newborn's death studies can reveal perspectives around points of view, for example, the population's diseases and the connection between social disparity and infant mortality as indicated by (Hernandez et al., 2011). High estimates often reflect doubtful degrees of well-being, bad health care, and financial changes.

Several factors can be associated with infant mortality, as social and biological, presenting some differences between regions with huge social imbalance. According to data from (IBGE, 2021), the Brazilian Institute of Geography and Statistics, in Vale do Jequitinhonha the infant death numbers are increasing more than in other regions.

The federal government created the SIM (Mortality Information System), in 1975, in order to get data about the children's demise (Brazil, 2021a), and the SINASC (System of Information about Born Alive), in 1990, in order to get data about the children's alive (Brazil, 2021b). Both epidemiological reasonableness structures acquired a reputation for presenting the welfare, observing registers, and evaluating governmental programs.

In this article, we characterized infant mortality through SIM and SINASC repository data, applying Formal Concept Analysis (FCA) in two regions of Minas Gerais state: The Metropolitan Region of Belo Horizonte (BH) with more than 6 million people, being the third largest in population brazil and the sixtieth most populous metropolitan area in the world, and the Vale do Jequitinhonha (VJ), bathed by the 31 thousand miles of the Jequitinhonha River, is home to more than 950,000 Brazilians. The Vale Jequitin-

Santos, D., Nobre, C., Zarate, L. and Song, M.
Application of Formal Concept Analysis and Data Mining to Characterize Infant Mortality in Two Regions of the State of Minas Gerais.
DOI: 10.5220/0011039900003179
In *Proceedings of the 24th International Conference on Enterprise Information Systems (ICEIS 2022) - Volume 1*, pages 155-162
ISBN: 978-989-758-569-2; ISSN: 2184-4992

honha is divided into three micro-regions: Low, Middle, and High Jequitinhonha. The Lower Jequitinhonha comprises the micro-region of Almenara, the closest to the State of Bahia, while the Middle encompasses the regions of Pedra Azul and Araçuaí. Finally, closer to the Metropolitan Region of Belo Horizonte, there is the Alto Jequitinhonha. Minas Gerais is one of the 27 Brazilian states and has 580 cities. The Metropolitan Region of Belo Horizonte (BH) owns 35 of them and Vale Jequitinhonha (VJ), 51 (IBGE, 2021).

Formal Concepts Analysis (FCA) is a mathematical technique that had its beginnings in 1982 with the work of Wille (Bernhard and Rudolf, 2012), who considered each element of a reticulated to be a formal concept, and the reticulated represented a hierarchy among the concepts. There is interest in applying FCA in several areas such as health, software engineering, data mining, among others. We use FCA in this work to extract association and implication rules.

2 BACKGROUND

2.1 Formal Concept Analysis

The FCA is being used to recognize patterns with the help of association rules and implications. The Formal Concept Analysis is composed of a set of elements such as formal context, formal concepts, and rules. The formal context is a triple $K(G,M,I)$, where G and M are, respectively, a set of objects and a set of attributes. I is an incidence relationship between G and M, where $I \subseteq G \times M$. Every incidence element is denoted as gIm or $(g,m) \in I$.

A formal concept from a formal context $K(G,M,I)$ is defined by a pair (A,B) where A is the extention, and B the intention - $A \subseteq G$, $B \subseteq M$. The pair (A,B) which defines the concept, need to follow the conditions where ($A = B$') e ($B = A$'). This relation is defined by the derivation operator as follows:

$$A' = \{g \in G \mid gIm \; \forall \; m \in B\} \quad (1)$$

$$B' = \{m \in M \mid gIm \; \forall \; g \in G\} \quad (2)$$

The extension A contains each object from G that has all the attributes from B, and the intention B owns all the attributes from M that belongs to all objects in A.

Association rules are dependencies between elements of a formal context. A rule $B \rightarrow C$ is valid only if every object that owns B attributes also contains C attributes. Formally, $B \rightarrow C$ only if $B, C \subseteq M$ and B' $\subseteq C$'. Given a rule r, and parameters s, c:

$$\text{s} = \text{supp}(r) = \frac{|A' \cap B'|}{|G'|}$$

is called the support of the rule r, and

$$\text{c} = \text{conf}(r) = \frac{|A' \cap B'|}{|A'|}$$

is its confidence. When $\text{conf}(r) = 100\%$, the rule is denoted as an implication.

Developed por Rudolf Wille in the 80's, the FCA is a applicated mathematics camp based in the concept and the conceptual hierarchy mathematization, (Bernhard and Rudolf, 2012). The Formal Concept Analysis considers the concepts as ways of intersubjectively comprehension in situations of oriented action for the porpoise. The concepts formalization must be clean and simple, but as well wide, so as the main aspects of a concept can have its references explicit in the formal model.

2.1.1 Lattice Miner

Lattice Miner developed by the LARIM research laboratory at Université du Québec en Outaouais under the supervision of Professor Rokia Missaoui. It allows the generation of clusters (called formal concepts) and association rules given a binary relation between a collection of objects and a set of attributes. The focus of Lattice Miner is on pattern (knowledge) discovery visualization, exploration and approximation through a lattice representation of either a flat or a nested structure (Missaoui and Emamirad, 2017). Version used in this work is (2.0).

3 RELATED WORKS

For countries in development, the child mortality until the first year still is a serious public health problem. Although the child mortality in Brazil has reduces, it continues to be high in different regions and cities.

In (Silva et al., 2017), they presents an FCA-based approach to identify the behavior of professionals registered on linkedIn through the database. Appropriate implications were applied to identify the skills that are needed to achieve a particular job position. Application of pre-processing techniques, transforming data into a formal concept and finally extracting appropriate implications using the PropIm algorithm analyzing the results through graphical representations.

In (Barbosa et al., 2014) it was proposed characterizing the 2008 child mortality in Vale Jequitinhonha region analyzing the demographic and socioeconomic data, the mother's historic, numbers of prenatal visits, data form a research done with the mothers. The Epi Info for error identification and posteriorly the Statistical Package for Social Science (SPSS)

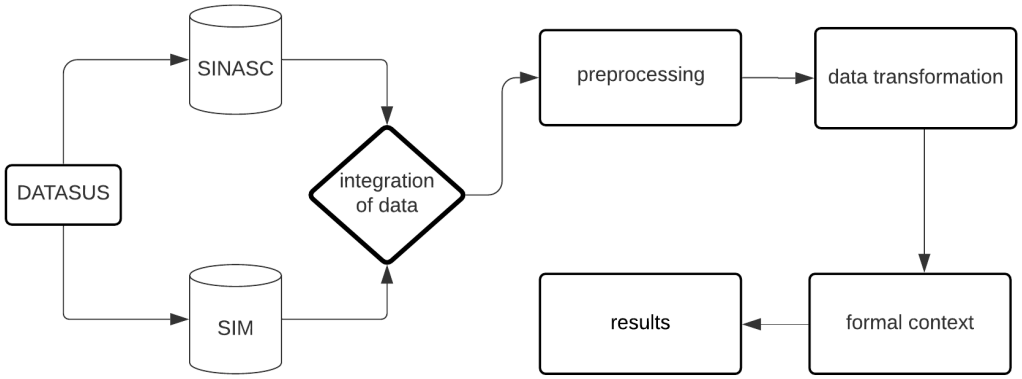

Figure 1: Methodology.

program for making statistics analysis. The study, despite of the small data show, points that prenatal period causes are determiners for child mortality, although the social problems and difficulties considering health access has a huge interference as well. In other (Barbosa et al., 2019) work it was proposed analyzing the child mortality focusing the evitability and social vulnerability between 2009 and 2014, as well in Vale Jequitinhonha region. The deaths were classified and two groups: the evitable and inevitable ones, following the CID-10 international disease classification. For the vulnerability analysis it was used the IVS, the social vulnerability index. With the data statistics analysis it was created a table that shows the importance of the causes about biological factor from the mother, and that evitable deaths are attributed as an access problem and the lack of assistance by the region health services.

In (Bonatti et al., 2020) it was proposed utilizing polynomial regression models for the temporal series study that analyzes the child mortality tendency in Mato Grosso state, Brazil, between 2007 to 2016. The result showed that regions had a child death reduction, but maintained the same in a few other regions.

The Formal Concept Analysis use was utilized (Ferreira et al., 2021) to extract a knowledge from a base of 120 women who did breast cancer chemotherapy, and they were separated in 2 groups to compare antimedical or non-olanzapine treatments. The Lattice Minner 2.0 was used for the rules extraction. The results pointed capacity to help doctors and researchers for proposals about the medicine administration according to the patient response. In (Ananias et al., 2021) it was proposed the classic FCA to the Triadic Concept Analysis, proposing an approach to find triadic concepts, allowing exploration and visualization, giving a tool for the data comprehension. Utilized in (Noronha et al., 2020), the triadic analysis allows the temporal evolution analysis from the clinical conditions in order to identify profiles applying the triadic rules that allows extracting human aging information.

Although many papers present solutions regarding this problem, most of them lack generalism. Some related articles propose different methods. Others address more specific issues by selecting a few attributes. In some papers, it is impossible to identify which procedures were adopted based on the information provided. In this paper, we present and apply a mathematical method of a binarized base for using the FCA, which generates association and implication rules, allowing us to determine infant mortality on aspects of mortality and survival, drawing a broad picture for each region.

4 METHODOLOGY

Although there are a lot of researches about the formal concept analysis applied in the health area, just a few explore the infant mortality rate characterization in the Metropolitan Region of Belo Horizonte and the Vale Jequitinhonha region, settled in Minas Gerais, Brazil. This article investigates infant mortality (IMR) using: data collection, analysis, exploration, selection and attribution transformation, and context and rules extraction.

The first step was to collect data from the Unique Health System Informatics Department (DATASUS), unifying the databank. The strategic organ inside the Brazilian government Health Ministry is responsible for collecting, processing, and sharing information related to health. Posteriorly, an analysis and attributes exploration was made to understand and identify those non-relevant to the theme to reduce the dimensionality to work on the FCA approach.

In this article, we used the database from DATASUS, considering the instances from 2006 to 2019 (for each region, DATASUS updates the information).

The Unique Health System (SUS) has a more extensive and variated database given by DATASUS. To obtain the data, we used the Mortality Information System (SIM) (Brazil, 2021a), and the System of Information about Born Alive (SINASC) to gather information about infant mortality (Brazil, 2021b). Both datasets acquired a reputation for spreading fundamental welfare information, observing registers, and evaluating governmental programs. They were unified so that a child classification, that survived or not their first life year, could be done.

The database of infant mortality of the region of Belo Horizonte taken from (DATASUS) and the Vale Jequitinhonha region went through pre-processing where irrelevant attributes, outliers were filtered and in the final stage of database preparation was balanced according to (Table 1), since an unbalanced class leads to the classifier being biased towards the majority class (Prati et al., 2003).

Table 1: Balancing uneven datasets using undersampling.

class	BH Region	VJ Region
alive	1069565 → 1843	107858 → 87
dead	1843 → 1843	87 → 87

The attributes are numerical and categorical, containing absent data and others with inconsistency thatwere analyzed. The World Health Organization (WHO) warns of the social and biological factors associated with infant mortality. They also use some of these attributes in the works of (Soares et al., 2021) and (Ananias et al., 2021). Based on this information, selected the following attributes were: *baby weight, APGAR1, APGAR5, gestation, birth, type of pregnancy, congenital anomaly, marital status, mother's age, the number of sons and daughters alive, the number of sons and daughter dead, mother's schooling, race,and the number of antenatal visits* (Sridevi and Nirmala, 2016).

The attribute "weight" is a continuous numerical that indicates the baby's weight, in grams, when born. For the data discretization, we used the DATASUS tabulation (Brazil, 2017), which consists in weighing the baby at birth, in grams: superior or equal to 4000, between 3000 and 3999, 2500 and 2999, 1500 and 2499, 1000 and 1499, 500 and 999, and less them 500. It is possible to see the data discretization in ascending order (Figure 2). It shows that the percentage of newborns in both the Belo Horizonte Metropolitan Area and the Jequitinhonha Valley area is different from 3000 to 3999 grams. We can also see differences in the weights above 4000 grams and below 500 grams scales between 500 to 999 grams and 2500 to 2999 grams.

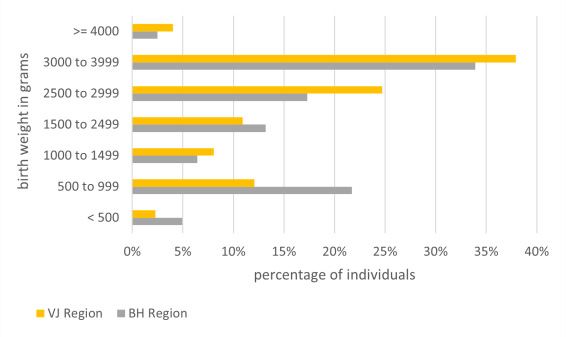

Figure 2: Weight of babies in the two regions evaluated (BH region and VJ region).

The APGAR 1 (Figures 3) and APGAR 5 (Figures 4) are two measures widely used in the newborn context evaluation. They are indicators used to evaluate five objective signals of the newborn: (skin color, reflective irritability, heart rate, breathing, and muscular tonus), each item is assigned a score of 0 to 3, the score of each item is added and a maximum score of 10 points is obtained (Sykes et al., 1982).

The attribute "APGAR" scale is directly related to mortality in the first 28 days of the baby's life. A score between 8 and 10 indicates a normal condition of the baby, a score between 4 and 7 the scale is considered moderate that needs special care, a score of 0 to 3 is severe and immediate resuscitation procedures are required. The discretization was done based on the apgar scale (Sykes et al., 1982). (Figure 3) shows that between regions there are important differences for both apgar1 considered good 8 to 10 and bad apgar1 0 to 3. In (Figure 4) apgar5 on the scale of 0 to 3 shows equivalence in the data between the two regions. Scales 4 to 7 and 8 to 10 show minor differences.

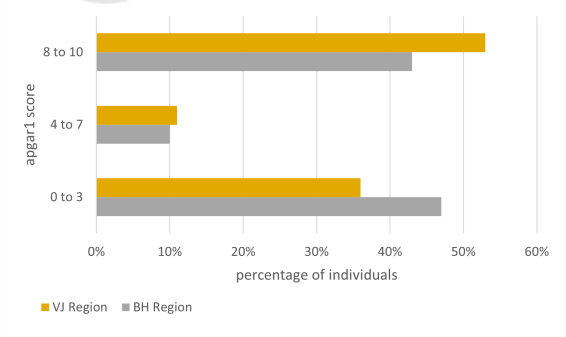

Figure 3: APGAR 1 value in the two regions evaluated (BH region and VJ region).

The attribute "gestational" (Tabela 2) consists of the time, in weeks, that the pregnancy lasted, was discretized used the DATASUS tabulation into six cate-

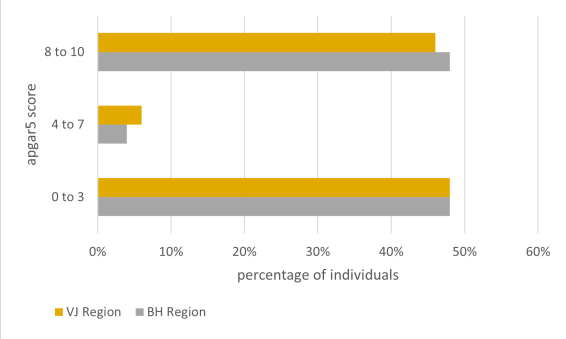

Figure 4: APGAR 5 value in the two regions evaluated (BH region and VJ region).

Table 2: Gestation period of mother in the two regions evaluated (BH region and VJ region).

Attribute	BH region	VJ region
42 weeks and more	1%	4%
37 to 41 weeks	54%	66%
32 to 36 weeks	13%	9%
28 to 31 weeks	9%	5%
22 to 27 weeks	22%	14%
less than 22 weeks	2%	2%

gories: less than 22 weeks, 22 to 27 weeks; 28 to 31 weeks; 32 to 36 weeks; 37 to 41 weeks and 42 weeks and then grouped: preterm or premature pregnancies (less than 37 weeks), the gestation of expected duration normal the gestational term is the period between 37 to 42 weeks, pregnancies exceeding 42 weeks are called poster which happens less frequently. According to (WHO, 2021), attention to care is needed in preterm and post-term babies.

Table 3: Type of Childbirth in the two regions evaluated (BH region and VJ region).

Attribute	BH region	VJ region
normal	51%	69%
cesarean	49%	31%

The attribute “Child-birth” (Table 3) can be classified as vaginal (or usual) or cesarean. While in the Metropolitan Region Belo Horizonte the proportion between vaginal and cesarean birth is balanced at almost 50%, in Vale do Jequitinhonha, the difference is that the vaginal birth is 2 times more common than the cesarean one.

Table 4: Type of pregnancy in the two regions evaluated (BH region and VJ region).

Attribute	BH region	VJ region
only	93%	91%
twins	6%	9%
triple or more	1%	0%

The attributes “Birth type” indicates one of thefollowing categories: only, twins, triple, and more babies (Table 4).

Table 5: Congenital anomaly of babies in the two regions evaluated (BH region and VJ region).

Attribute	BH region	VJ region
no anomaly	89%	91%
some anomaly	11%	9%

The attribute “Congenital anomaly” indicates that the child had some anomaly. Congenital refers to the existence at or before birth. It could occur during pregnancy and could be detected before or after birth (Table 5). Approximately 50% of all cannot be linked to a specific disease, there are some known genetic, environmental or risk factors.The great majority is no anomaly birth.

Table 6: Marital status in the two regions evaluated (BH region and VJ region).

Attribute	BH region	VJ region
single	52%	48%
married	43%	31%
consensual unions	3%	21%
divorced	2%	0%

The attribute “Marital status” identifies the mother’s social status divided into four categories: single, married, divorced and consensual unions. As it does not contain data in the widow category, this attribute was excluded from the database (Table 6).

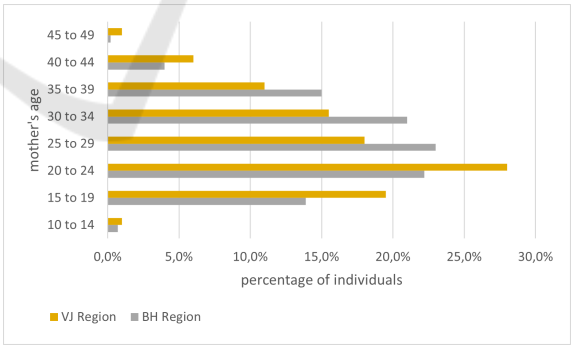

Figure 5: Age of mothers in the two regions evaluated (BH region and VJ region).

The attribute “Mother’s age” (Figure 5) is a continuous numeral that indicates how old the baby’s mother is, the discretization based on DATASUS tabulation in eight categories: 10 to 14 years, 15 to 19 years, 20 to 24 years, 25 to 29 years, 30 to 34 years, 35 to 39 years, 40 to 44 years, and 45 to 49 years. According to (WHO, 2021), adolescent mothers under 20 years and mothers older than 35 years should

receive special attention during pregnancy, for several implications.

Table 7: Number Living Children (BH region and VJ region).

Attribute	BH region	VJ region
0	51,9%	40%
1 a 2	40,2%	43%
3 a 4	6,5%	8%
5 a 6	1%	7%
7 a 8	0,4%	2%

The attribute “Number of living children” (Table 7) was discretized according to the DATASUS tabulation into five categories: 0 living children, between 1 and 2 children, 3 to 4 children, 5 to 6 children, and 7 to 8 living children. The number of children is linked to Family Planning being a set of actions that help men and women to plan the arrival of children, and also to prevent unplanned pregnancy, thinking about the conditions for their education and proper care.

Table 8: Number Dead Children (BH region and VJ region).

Attribute	BH region	VJ region
0	86%	85%
1 a 2	13%	13%
3 a 4	1%	2%

The attribute “number of dead children” (Table 8) were discretized, but due to lack of data, other data ranges were excluded, leaving three categories: 0 dead children, 1 to 2 dead children, and 3 to 4 dead children.

Table 9: Mother’s Schooling (BH region and VJ region).

Attribute	BH region	VJ region
1 to 3 years old	2%	13%
4 to 7 years old	18%	33%
8 to 11 years old	61%	48%
12 years and more	19%	6%

The attribute “Mother’s schooling” (Table 9), shows some outliers, knowing that those instances were removed. The attribute ”Mother’s schooling” indicates the mother’s level of education in years. The discretization was based on the DATASUS tabulation (Brazil, 2017). The categories are 1 to 3 years, 4 to 7 years, 8 to 11 years, and 12 years and more. According (WHO, 2021), the categories below 8 years of study are considered low education.

The attribute ”Race” (Table 10) has four categories: brown, white, black, and yellow. Therefore, it is an essential indicator for analyzing the complex social scenario according to race/color. The percentage of brown people in the Vale Jequitinhonha region is higher if we consider the Belo Horizonte region, which has a higher rate of white people.

Table 10: Race categorization in the two regions evaluated (BH region and VJ region).

Attribute	BH region	VJ region
brown	67%	83%
white	28%	12%
black	4%	4%
yellow	1%	1%

All attributes were analyzed and modified according to individual data analysis, discretized attributes with missing or irrelevant data were excluded, generating the categorized base.

Table 11: Part of the formal context.

child-birth-normal	child-birth-cesarean	anomaly-not	anomaly-yes	weight-3000-3999	weight-2500-2999	weight-1500-2499	weight-1000-1499	weight-4000-more	weight-500-less-than
x		x			x				
	x	x		x					
	x	x			x				
x		x		x					
x		x					x		

From the categorization, it was possible to create a formal context by transforming the base with binarized variables (Table 11).

We applied the Lattice Miner software, an FCA-oriented tool, to generate the association and implication rules based on the formal context created. We separated them into two scenarios, being living and dead. We defined a minimum support of 30% and a confidence level of 30%, to generate more rules. We selected only some of the implication rules that had the confidence of 100%, higher support rate being the consequence alive or dead. In (Table 12) and (Table 13) we selected for the two regions three main rules for the baby survival scenario and three main rules for the baby mortality scenario.

5 RESULTS AND DISCUSSION

When using the FCA, support and confidence was defined with values of at least 30% for the scenario in the metropolitan region of Belo Horizonte, the first analysis was considered the dead class, generating,

this way, sixty-three (63) rules, later the Alive class generated one hundred-sixty-nine (169) rules. For the Jequitinhonha Valley region, the first scenario generated one-hundred-fifty-eight (158) rules for the dead class, and for the alive class two hundred and eight-hundred-sixty-six (866) rules.

It was possible to collect the results in *xml* format and analyze the rules according to the form A $\rightarrow$ B, A being the premise, and B the consequence, generating the support (sup) and confidence (conf) data. (Table 12) generates implication rules pointing attributes.

Rules number 6 and 5 indicate weight from 500 to 999 grams and a premature pregnancy around 22 to 27 weeks in the scenario for children who died. Rules number 1, 2, and 3 indicate attributes as pregnancy from 37 to 41 weeks, and antenatal visits are the most explicit. In the scenario for living children, the most reliable rule is 100%, and the most supportive is 86%.

Table 12: Main rules generated in Belo Horizonte Region.

N.	Rules	Sup	Conf
1	**IF** antenatal visits 7, gestation 37 to 41 weeks and weight 3000 to 3999 grams **THEN Living**	41%	100%
2	**IF** gestation 37 to 41 weeks **THEN Living**	86%	100%
3	**IF** antenatal visits 7 **THEN Living**	72%	100%
4	**IF** apgar1 score of 0 to 3 **THEN Child death**	42%	100%
5	**IF** gestation 22 to 27 weeks **THEN Child death**	43%	100%
6	**IF** gestation 22 to 27 weeks, weight 500 to 999 grams **THEN Child death**	33%	100%

In the region of Vale Jequitinhonha the rules in the scenario for living children (Table 13) they point different attributes as gestation time, only child, and mother schooling. The rules selected were the 100% confidence ones and the rule with higher support is 85%.

Rules number 1, 2, and 3 point to gestation between 37 and 41 weeks and mother's education above 8 years as crucial attributes for child survival. The feature of a mother's schooling does not appear for the Metropolitan Region of Belo Horizonte, showing it to be a decisive factor in the Vale Jequitinhonha Region.

For the rules in the scenario for living children they point different attributes as gestation time, only child, and mother schooling. The rules selected were the 100% confidence ones and the rule with higher support is 85%.

Rule number 4 points out that even with an Apgar score between 8 to 10, which is considered good, the baby dies in 48% of cases. This is expected with an Apgar score between 0 and 3, which, as rule 5 shows, happens in 40% of the cases. Finally, rule number 6 points out the factor of the mother's poor education as one of the factors for the baby's death.

Table 13: Main rules generated in Vale Jequitinhonha Region.

N.	Rules	Sup	Conf
1	**IF** gestation 37 to 41 weeks **THEN Living**	85%	100%
2	**IF** gestation 37 to 41 weeks and only child **THEN Living**	82%	100%
3	**IF** mother schooling 8 to 11 years old **THEN Living**	58%	100%
4	**IF** apgar1 score of 8 to 10 **THEN Child death**	48%	100%
5	**IF** apgar5 score of 0 to 3 **THEN Child death**	40%	100%
6	**IF** mother schooling 4 to 7 years old**THEN Child death**	48%	100%

In the Metropolitan Region of Belo Horizonte, attributes such as weight, apgar, and gestation are linked to the biological factors of the mother. Still, with proper nutritional monitoring and adequate prenatal care, the three attributes above can be reduced and minimized since they are often among preventable causes. On the other hand, in the Vale Jequitinhonha region, the situation is a social issue, presenting a low number of prenatal consultations due to poor health structure and inadequate education of some mothers, causing poor financial conditions of part of the population.

6 CONCLUSIONS

The association rules showed the characteristics of infant mortality taking into account variables such as weight and gestation time as determinants of child survival, mortality is higher taking into account

weight, premature pregnancy and APGAR with a value between 0 and 3 , within the rules, it is remarkable that for the metropolitan region of Belo Horizonte, even if the child has an unfavorable weight, the recovery is better than the region of Vale Jequitinhonha, which has higher mortality with higher weights.

For future work, the necessity to create new scenarios and use different tools is essential to identify and understand infant mortality in the regions of Minas Gerais and Brazil.

REFERENCES

Ananias, K. H., Missaoui, R., Ruas, P. H., Zarate, L. E., and Song, M. A. (2021). Triadic concept approximation. *Information Sciences*, 572:126–146.

Barbosa, T. A. G. d. S., Coelho, K. R., Andrade, G. N. d., Bittencourt, S. D. d. A., Leal, M. d. C., and Gazzinelli, A. (2014). Determinantes da mortalidade infantil em municípios do vale do jequitinhonha, minas gerais, brasil. *Revista Mineira de Enfermagem*, 18(4):907–922.

Barbosa, T. A. G. d. S., Gazzinelli, A., and Andrade, G. N. d. (2019). Mortalidade infantil evitável e vulnerabilidade social no vale do jequitinhonha, minas gerais, brasil. *Revista Mineira de Enfermagem*, 23:1–8.

Bernhard, G. and Rudolf, W. (2012). *Formal concept analysis: mathematical foundations*. Springer Science & Business Media.

Black, R. E., Liu, L., Oza, S., Hogan, D., Perin, J., Rudan, I., ELawn, J., Cousens, S., and Mathers, C. (2017). Global, regional, and national causes of child mortalityin 2000–13, with projections to inform post-2015 priorities: an updated systematic analysis. Lancet.

Bonatti, A. F., Silva, A. M. C. d., and Muraro, A. P. (2020). Mortalidade infantil em mato grosso, brasil: tendência entre 2007 e 2016 e causas de morte. *Ciência & Saúde Coletiva*, 25:2821–2830.

Brazil (2017). Mortalidade geral – 1996 a 2015 notas técnicas. website: http://tabnet.datasus.gov.br/cgi/sim/Mortalidade_Geral_1996_2012.pdf.

Brazil (2021a). Sistema informação sobre mortalidade. website: http://tabnet.datasus.gov.br/cgi/tabcgi.exe?sim/cnv/obt10uf.def.

Brazil (2021b). Sistema informação sobre nascidos vivos. website: http://tabnet.datasus.gov.br/cgi/deftohtm.exe?sinasc/cnv/nvuf.def.

Ferreira, L., Nobre, C., Zárate, L., and Song, M. (2021). Study of the evolution of antiemetic treatment through the application of triadic formal concept analysis. In *Anais do IX Symposium on Knowledge Discovery, Mining and Learning*, pages 1–7. SBC.

Hernandez, A. R., Silva, C. H. d., Agranonik, M., Quadros, F. M. d., and Goldani, M. Z. (2011). Análise de tendências das taxas de mortalidade infantil e de seus fatores de risco na cidade de porto alegre, rio grande do sul, brasil, no período de 1996 a 2008. volume 27, pages 2188–2196. SciELO Brasil.

IBGE (2021). Instituto brasileiro de geografia e estatística. website: https://www.ibge.gov.br/cidades-e-estados/mg.html.

Missaoui, R. and Emamirad, K. (2017). Lattice miner-a formal concept analysis tool. In *14th International Conference on Formal Concept Analysis*, page 91.

Noronha, M. D., Rodrigues, M. W., Ribeiro, C. E., Nobre, C. N., Song, M. A., and Zárate, L. E. (2020). Characterization of long-lived and non-long lived profiles through biclustering. In *Proceedings of the 35th Annual ACM Symposium on Applied Computing*, pages 473–476.

Prati, R. C., Batista, G., and Monard, M. C. (2003). Uma experiência no balanceamento artificial de conjuntos de dados para aprendizado com classes desbalanceadas utilizando análise roc. Proc. of the Workshop on Advances & Trends in AI for Problem Solving.

Silva, P. R. C., Dias, S. M., Brandão, W. C., Song, M. A., and Zárate, L. E. (2017). Formal concept analysis applied to professional social networks analysis. In *Proceedings of the 19th International Conference on Enterprise Information Systems - Volume 1: ICEIS,*, pages 123–134. INSTICC, SciTePress.

Soares, W. L., Zárate, L. E., Song, M. A., and Nobre, C. N. (2021). Characterizing infant mortality using machine learningtechniques: a case study in two brazilian states -santa catarina and amapá. volume 7, pages 45269–45290.

Sridevi, S. and Nirmala, S. (2016). Anfis based decision support system for prenatal detection of truncus arteriosus congenital heart defect. *Applied Soft Computing*.

Sykes, G., Johnson, P., Ashworth, F., Molloy, P., Gu, W., Stirrat, G., and Turnbull, A. (1982). Do apgar scores indicate asphyxia? *The Lancet*, 319(8270):494–496.

UNICEF (2020). Neonatal mortality. website: https://data.unicef.org/topic/child-survival/neonatal-mortality/.

WHO (2021). World health organization. website: https://www.who.int/data/gho/data/themes/topics/topic-details/GHO/child-mortality-and-causes-of-death.