

CAROLINE LOPES LELES - SP3094928

ERICK CARVALHO BRITO – SP3097773

ELIZEU SANTIAGO BRAGA DOS REIS – SP3096491

MATHEUS ALMEIDA SOARES - SP3096556

SISTEMA DE INFORMAÇÃO SOBRE NASCIDOS VIVOS

TRABALHO DE ESPA5

SÃO PAULO

2024

SUMÁRIO

1. INTRODUÇÃO	3
2. ESTATÍSTICA DESCRITIVA	4
3. PROBABILIDADE	8
4. INFERÊNCIA	10
5. CONCLUSÃO	13

1. INTRODUÇÃO

O estudo da saúde de natalidade é essencial para entender e melhorar as condições de nascimento e em geral, os cuidados com recém-nascidos e mães. Com o objetivo de analisar dados vitais e identificar padrões relevantes, este relatório utiliza o conjunto de dados **Sistema de Informação sobre Nascidos Vivos – Sinasc - Nascidos Vivos – 2022**. Esta base de dados compila informações detalhadas sobre os nascimentos ocorridos no Brasil durante o ano de 2022, fornecendo uma rica fonte de dados para análises estatísticas.

Este trabalho aborda três principais áreas de análise:

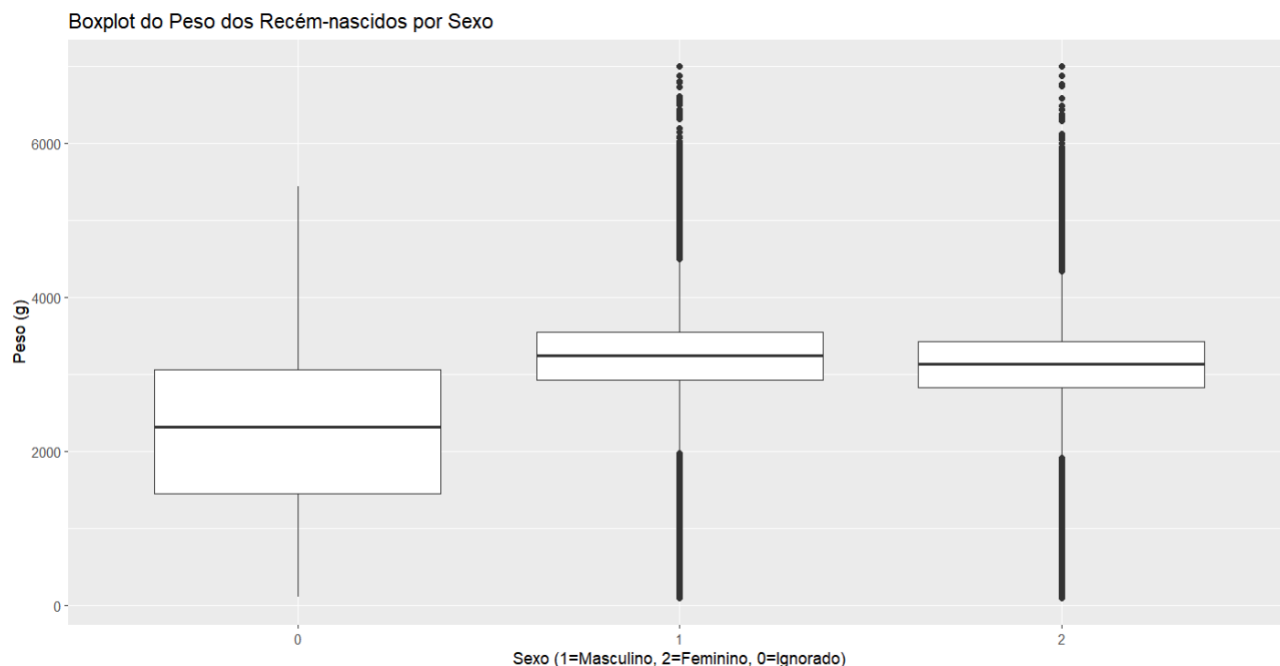
- **Estatística Descritiva:** A estatística descritiva foi utilizada para explorar distribuições de variáveis chave, como o peso dos recém-nascidos por sexo e a idade das mães por estado civil. Essas análises fornecem uma visão inicial das características básicas dos dados, ajudando a identificar tendências e padrões relevantes.
- **Probabilidade:** A probabilidade foi utilizada para avaliar a probabilidade de eventos importantes, como o baixo peso ao nascer e a ocorrência de anomalias congênitas, considerando diferentes fatores como o tipo de parto e a idade da mãe. A análise probabilística oferece uma compreensão mais profunda dos riscos associados a diferentes condições.
- **Inferência:** A interferência foi utilizada para testar hipóteses sobre a diferença no número de consultas pré-natais entre grupos etários das mães e a variação no peso dos recém-nascidos conforme a idade gestacional. Usando essa técnica, é possível identificar se as observações encontradas são estatisticamente significativas.

2. ESTATÍSTICA DESCRITIVA

O primeiro ponto analisado foi a distribuição de peso dos recém-nascidos por sexo. Essa análise é útil para entender diferenças de crescimento entre esses recém-nascidos, além de ser possível planejar cuidados neonatais personalizados e até antecipar potenciais complicações de saúde.

Calculamos a média, mediana e desvio padrão do peso dos recém-nascidos por sexo utilizando **summary()** e visualizamos a distribuição através de um **boxplot** com **ggplot2**:

```
> summary(DNOPEN22$PESO)
  Min.   1st   Qu.   Median   Mean   3rd   Qu.   Max.
  100    2875   3195   3153    3500   7000
> summary(DNOPEN22[DNOPEN22$SEXO == 1, ]$PESO)# Masculino
  Min.   1st   Qu.   Median   Mean   3rd   Qu.   Max.
  100    2925   3250   3205    3555   7000
> summary(DNOPEN22[DNOPEN22$SEXO == 2, ]$PESO)# Feminino
  Min.   1st   Qu.   Median   Mean   3rd   Qu.   Max.
  100    2830   3140   3100    3435   7000
> summary(DNOPEN22[DNOPEN22$SEXO == 0, ]$PESO)# Ignorado
  Min.   1st   Qu.   Median   Mean   3rd   Qu.   Max.
  120  1456  2320  2266  3061  5435
```



Através dos resultados, é possível identificar que recém-nascidos do sexo masculino apresentaram pesos variando entre 100 e 7000 gramas, com uma média de 3205 gramas e mediana de 3250 gramas. Além disso, o primeiro quartil é 2925 gramas e o terceiro quartil é 3555 gramas.

Já recém-nascidos do sexo feminino, também tiveram pesos entre 100 e 7000 gramas. Já a média é de 3100 gramas, a mediana de 3140 gramas e os quartis são de 2830 e 3435 gramas. Em geral, é notório que elas possuem pesos ligeiramente mais baixos em relação aos meninos.

Para os recém-nascidos com sexo não identificado, os pesos variaram de 120 a 5435 gramas, com uma média de 2266 gramas e mediana de 2320 gramas. Os quartis são 1456 e 3061 gramas, mostrando uma maior dispersão e menores valores centrais de peso em comparação com os outros sexos.

O segundo questionamento levantado foi qual é a distribuição de idade das mães por estado civil. Esse questionamento é relevante para entender como diferentes contextos sociais e relacionais influenciam a idade em que as mulheres se tornam mães, permitindo uma análise mais profunda das condições socioeconômicas.

Nessa análise, calculamos a média, mediana e desvio padrão da idade das mães por estado civil, utilizando `summary()` e visualizamos a distribuição através de um **boxplot** com **ggplot2**:

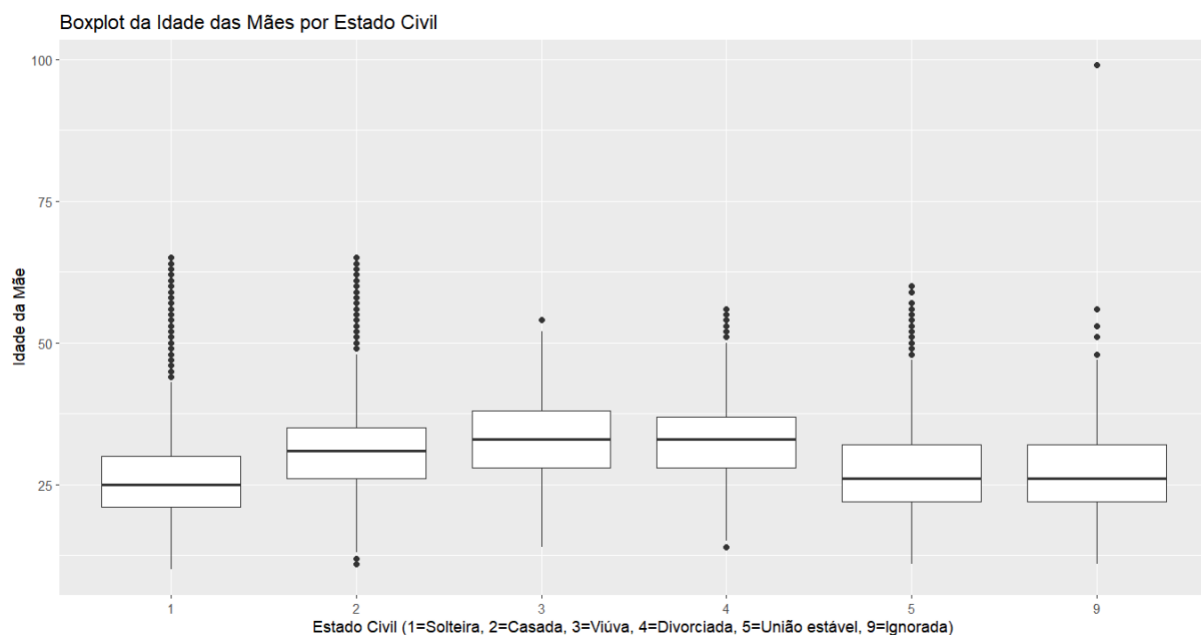
```
> summary(DNOPEN22$IDADEMAE)
```

Min.	1st	Qu.	Median	Mean	3rd	Qu.	Max.
------	-----	-----	--------	------	-----	-----	------

```

10.00      22.00      27.00      27.56      33.00      99.00
> summary(DNOPEN22[DNOPEN22$ESTCIVMAE == 1, ]$IDADEMAE) # Solteira
  Min.   1st   Qu.   Median     Mean   3rd   Qu.     Max.
10.00    21.00   25.00   25.69   30.00   65.00
> summary(DNOPEN22[DNOPEN22$ESTCIVMAE == 2, ]$IDADEMAE) # Casada
  Min.   1st   Qu.   Median     Mean   3rd   Qu.     Max.
11.00    26.00   31.00   30.67   35.00   65.00
> summary(DNOPEN22[DNOPEN22$ESTCIVMAE == 3, ]$IDADEMAE) # Viúva
  Min.   1st   Qu.   Median     Mean   3rd   Qu.     Max.
14.00    28.00   33.00   32.49   38.00   54.00
> summary(DNOPEN22[DNOPEN22$ESTCIVMAE == 4, ]$IDADEMAE) # Divorciada
  Min.   1st   Qu.   Median     Mean   3rd   Qu.     Max.
14.0    28.0   33.0   32.6    37.0   56.0
> summary(DNOPEN22[DNOPEN22$ESTCIVMAE == 5, ]$IDADEMAE) # União estável
  Min.   1st   Qu.   Median     Mean   3rd   Qu.     Max.
11.00    22.00   26.00   26.73   32.00   60.00
> summary(DNOPEN22[DNOPEN22$ESTCIVMAE == 9, ]$IDADEMAE) # Ignorada
  Min.   1st   Qu.   Median     Mean   3rd   Qu.     Max.
11.0   22.0   26.0   26.8   32.0   99.0

```



Através dos resultados, identifica-se que, em geral, a média de idade das mães é 27,56 anos, com uma mediana de 27 anos. A maioria das mães está no intervalo interquartil de 22 a 33 anos.

Tratando-se de estado civil, mães solteiras possuem uma média de 25,69 anos e mediana de 25 anos. O primeiro quartil é 21 anos, e o terceiro quartil é 30 anos, indicando que a maioria das mães solteiras são mais jovens. Já mães casadas possuem uma média de 30,67 anos e mediana de 31 anos. O intervalo interquartil é de 26 a 35 anos. Em comparação com as solteiras, elas tendem ter mais idade.

As viúvas e divorciadas possuem uma média de 32,49 e 32,6 e a mesma mediana, de 33 anos. O intervalo interquartil é de 28 a 37 anos. Em geral, elas possuem mais idade que as solteiras e casadas.

Identifica-se também mães com união estável, que possuem uma média de 26,73 anos e mediana de 26 anos. O intervalo interquartil é de 22 a 32 anos, colocando-as em uma faixa etária intermediária, mais próximas das mães solteiras em termos de idade.

Por fim, tem-se o estado civil ignorado com uma média de 26,8 anos e mediana de 26 anos. O intervalo interquartil é de 22 a 32 anos. A ampla faixa etária reflete a diversidade deste grupo, onde o estado civil não foi especificado.

Com isso, evidencia-se que as mães solteiras e em união estável tendem a ter filhos em idades mais jovens, refletindo em possíveis barreiras ou diferentes pressões sociais em relação ao planejamento familiar. Em contraponto, as mães casadas têm filhos em idades mais avançadas. Isso pode ser atribuído a uma maior estabilidade econômica e a um planejamento familiar mais estruturado.

As mães viúvas e divorciadas mostram idades ainda mais altas, com médias superiores a 32 anos. Estes perfis refletem em uma combinação de experiências de vida que incluem a viuvez ou o divórcio, influenciando a decisão de ter filhos em uma fase mais madura da vida.

3. PROBABILIDADE

O primeiro ponto analisado nesse tópico foi a probabilidade de um recém-nascido ter baixo peso ao nascer ($< 2500\text{g}$) considerando o tipo de parto (normal ou cesárea).

Calculamos a probabilidade de um recém-nascido ter baixo peso ao nascer, separando pelo tipo de parto e utilizando a função **prop.table()** em uma tabela de frequências.

```
>                                baixo_peso_normal
[1]                                0.08671756
>                                baixo_peso_cesarea
[1] 0.09978912
```

Analisando os resultados, identifica-se que a probabilidade de um recém-nascido ter baixo peso ao nascer é de aproximadamente **8,67%** (0,0867). Já em cesárea, essa probabilidade é ligeiramente superior, em torno de **9,98%**. Como, em geral, partos cesáreos são frequentemente escolhidos em situações de risco, acaba fazendo sentido sua porcentagem ser maior que a de parto normal.

Essa análise destaca como é crucial considerar cuidadosamente o tipo de parto escolhido, não só para garantir a saúde imediata da mãe, mas também para entender como isso pode afetar o peso e a saúde do recém-nascido. Isso ajuda os profissionais de saúde e os futuros pais a planejarem os nascimentos com mais informações e segurança.

O segundo ponto analisado foi a probabilidade de um recém-nascido nascer com anomalia considerando a idade da mãe.

Calculamos a probabilidade de um recém-nascido nascer com anomalia, separando por faixas etárias da mãe e utilizando **prop.table()** em uma tabela de frequências.

	1 Sim	2 Não	9 Ignorado
0-19	0.008988888	0.982699401	0.008311710
20-29	0.008502770	0.984378971	0.007118258
30-39	0.009630548	0.983625159	0.006744293
40-49	0.017299501	0.975939109	0.006761391
50-59	0.010666667	0.981333333	0.008000000

Através dos resultados, identifica-se que mães mais jovens, entre 0 e 19 anos, apresentam uma probabilidade de 0,90%, enquanto aquelas na faixa de 20 a 29 anos têm uma probabilidade ligeiramente menor de 0,85%. A partir dos 30 anos, a probabilidade aumenta gradualmente, atingindo 0,96% para mães de 30 a 39 anos. No entanto, há um aumento mais acentuado entre as mães de 40 a 49 anos, com uma probabilidade de 1,73%, e uma leve queda para 1,07% entre as mães de 50 a 59 anos.

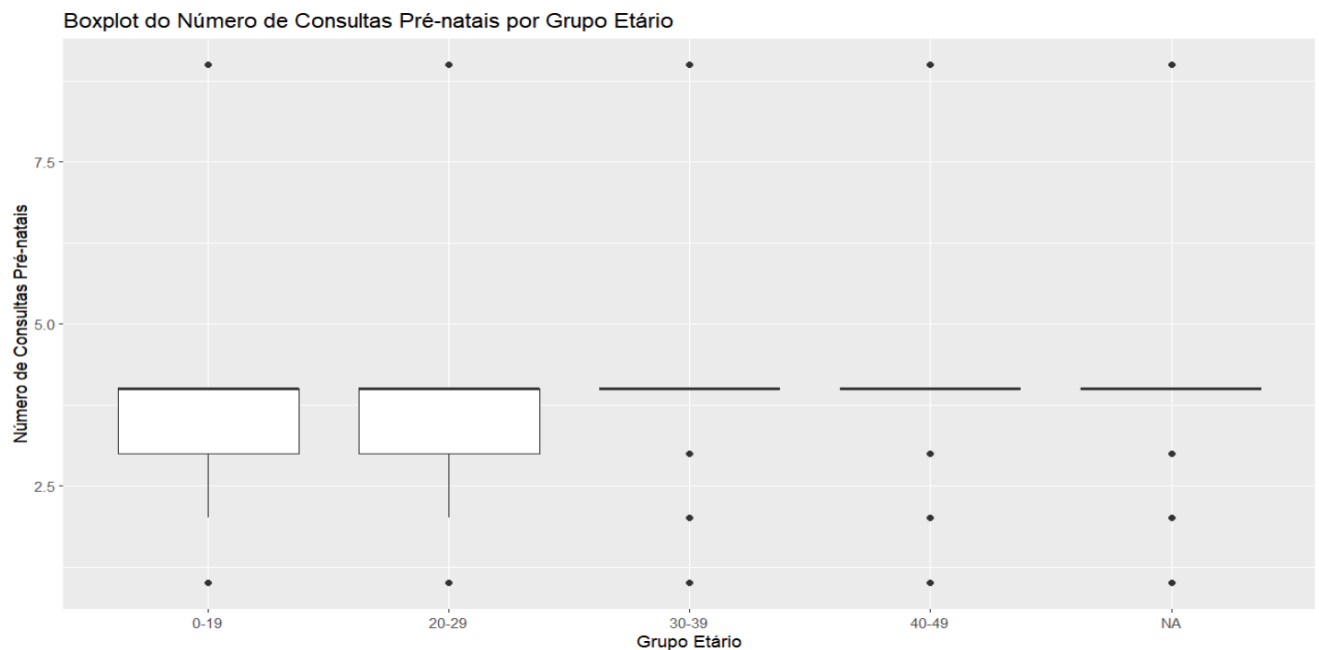
Mães mais jovens, abaixo de 20 anos, têm a menor probabilidade, enquanto mães mais velhas, especialmente aquelas com 40 anos ou mais, mostram uma probabilidade maior. Este padrão pode refletir fatores como maior incidência de complicações genéticas associadas ao envelhecimento materno, além de possíveis condições médicas crônicas que aumentam com a idade.

Por isso, é de suma importância considerar a idade materna como um fator de risco significativo na avaliação de anomalias ao nascimento. Esse conhecimento é fundamental para mitigar esses riscos e proporcionar suporte adequado durante o pré-natal e o parto, promovendo melhores resultados para as mães e seus filhos.

4. INFERÊNCIA

O primeiro ponto de análise foi verificar se existe diferença significativa entre o número médio de consultas pré-natais realizadas por mães de diferentes grupos etários.

Realizamos um teste para verificar essa diferença utilizando **aov()** e visualizamos os resultados com um **boxplot**.



Aplicando um teste de análise de variância (ANOVA) temos o resultado abaixo:

Df	Sum	Sq	Mean	Sq	F	value	Pr(>F)
GRUPO_ETARIO		3	13106	4369	8686	<2e-16	***
Residuals		2523413	1269208				1

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

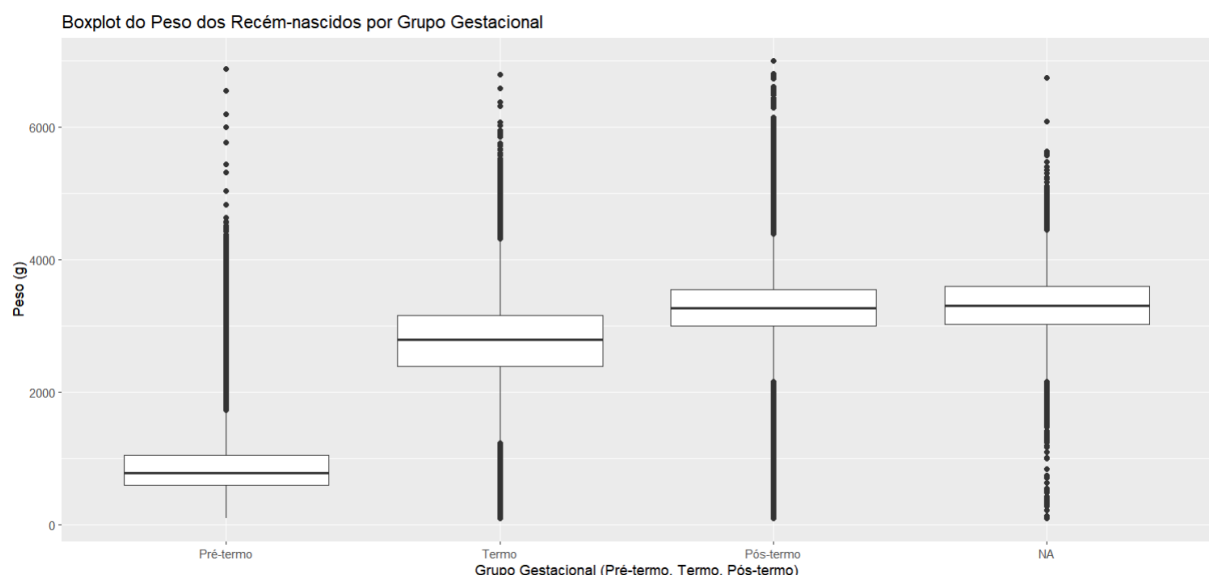
401 observations deleted due to missingness

A variável **GRUPO_ETARIO** (que representa diferentes faixas etárias das mães) apresenta uma diferença significativa no número médio de consultas pré-natais realizadas. Isso é indicado pelo valor extremamente baixo de p-value ($<2e-16$), o que significa que há uma evidência estatística forte de que pelo menos um dos grupos etários tem um número médio de consultas pré-natais significativamente diferente dos outros.

Com base nos resultados, podemos afirmar com um alto nível de confiança estatística que há diferenças significativas no número médio de consultas pré-natais realizadas por mães de diferentes grupos etários. Essa descoberta é clinicamente relevante porque o número de consultas pré-natais pode impactar diretamente a saúde materna e fetal, o acompanhamento da gravidez e a detecção precoce de complicações.

O segundo ponto de análise foi verificar se existe diferença significativa no peso dos recém-nascidos entre diferentes faixas de idade gestacional?

Realizamos um teste para verificar diferenças significativas no peso dos recém-nascidos entre diferentes faixas de idade gestacional, utilizando `aov()` e visualizamos os resultados com um boxplot.



Aplicando um teste de análise de variância (ANOVA) temos o resultado abaixo:

	Df	Sum	Sq	Mean	Sq	F	value	Pr(>F)
GRUPO_GESTACIONAL			2	1.865e+11	9.323e+10	389865	<2e-16	***
Residuals				2506314	5.993e+11		2.391e+05	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

17501 observations deleted due to missingness

Após ajustar o modelo ANOVA, olhamos para o p-valor associado ao fator GESTACAO. Um p-valor significativo (< 0.05) indica que há evidência estatística de que pelo menos um dos grupos de idade gestacional tem um peso médio significativamente diferente dos outros.

A variável GRUPO_GESTACIONAL (que representa diferentes faixas de idade gestacional) apresenta uma diferença significativa no peso dos recém-nascidos. Isso é indicado pelo valor extremamente baixo de p-value ($< 2e-16$), o que significa que há uma evidência estatística forte de que pelo menos um dos grupos de idade gestacional tem um peso médio significativamente diferente dos outros.

F-value: O valor de F é muito alto (389865), o que também sugere uma forte evidência estatística contra a hipótese nula de que não há diferença entre os grupos.

Portanto, com base nesses resultados, podemos concluir que há uma diferença significativa no peso dos recém-nascidos entre diferentes faixas de idade gestacional. Isso significa que a idade gestacional tem um impacto importante no peso dos bebês ao nascer, e as diferenças observadas não são simplesmente devido ao acaso.

5. CONCLUSÃO

Baseado nas análises realizadas, é possível concluir que diversos fatores relacionados aos recém-nascidos e suas mães apresentam variações significativas que merecem atenção em políticas de saúde pública e práticas clínicas.

A distribuição de peso dos recém-nascidos por sexo revela diferenças sutis, porém relevantes, que podem indicar necessidades específicas de cuidados dependendo do gênero do recém-nascido. A idade materna se mostrou crucial ao analisar a probabilidade de anomalias ao nascimento, demonstrando um aumento significativo de risco em mães com mais idade. Além disso, a análise da quantidade de consultas pré-natais por faixa etária evidencia que a idade das mães também impacta diretamente nos cuidados pré-natais recebidos, ressaltando a importância de estratégias personalizadas para melhorar os resultados de saúde materno-infantil.

Essas descobertas sublinham a necessidade contínua de monitoramento e adaptação das práticas de saúde para garantir o bem-estar tanto das mães quanto dos recém-nascidos em diferentes contextos demográficos e sociais. Em geral, são dados que devem ser analisados para que se tenha uma melhoria na saúde pública do Brasil.