

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE MINAS GERAIS

Instituto de Ciências Exatas e Informática – ICEI

Breno Rezende Silva de Abreu

Caio César de Oliveira

Caio Martins Montandon

Gabriel Henrique Costa Soares

Rafael Hashimoto Sanches Barbosa

**DESIGUALDADES REGIONAIS NO TRATAMENTO HOSPITALAR DE
CRIANÇAS NO SUS:**

uma análise das internações, custos e mortalidade infantil entre 0 e 9 anos

Belo Horizonte

2025

Breno Rezende Silva de Abreu
Caio César de Oliveira
Caio Martins Montandon
Gabriel Henrique Costa Soares
Rafael Hashimoto Sanches Barbosa

**DESIGUALDADES REGIONAIS NO TRATAMENTO HOSPITALAR DE
CRIANÇAS NO SUS:
uma análise das internações, custos e mortalidade infantil entre 0 e 9 anos**

Artigo Científico apresentado aos alunos da
Graduação Presencial Assíncrona em Engenharia da
Computação da Pontifícia Universidade Católica de
Minas Gerais.

Orientadores: Prof. Sandro Jerônimo de Almeida
Prof. Kenia Aparecida Caires Cardoso

Área de concentração: Análise exploratória de dados

Belo Horizonte
2025

SUMÁRIO

INTRODUÇÃO	4
Definição do Problema	4
Objetivo	4
Hipóteses a serem investigadas.....	4
EXTRAÇÃO DE DADOS.....	6
DICIONÁRIO DE DADOS	8
UNIFICAÇÃO DAS BASES	9
ANÁLISE EXPLORATÓRIA DE DADOS	10
Relações dos dados com as hipóteses	10
Conclusão estatística.....	11
CONCLUSÃO.....	12

INTRODUÇÃO

As **desigualdades regionais na saúde pública brasileira** constituem um desafio **histórico e persistente**. Em um país de dimensões continentais como o Brasil, a **oferta** e a **qualidade** dos serviços de saúde variam consideravelmente entre as diferentes regiões, refletindo **disparidades socioeconômicas, de infraestrutura e de acesso aos serviços básicos**. Essas desigualdades tornam-se especialmente preocupantes quando afetam populações **vulneráveis**, como as **crianças**. A **morbidade hospitalar infantil**, nesse contexto, é um importante indicador da eficácia dos serviços prestados durante a internação e pode revelar **falhas estruturais** nos sistemas regionais de saúde.

Definição do Problema

A presente investigação busca responder à seguinte pergunta orientada a dados: **como a eficácia da assistência hospitalar prestada durante a internação pública de crianças de 0 a 9 anos varia entre as regiões do Brasil, considerando indicadores de morbidade hospitalar como número de internações, óbitos, tempo de permanência e valores gastos por internação?** Essa questão surge da necessidade de compreender como o contexto regional influencia os resultados hospitalares infantis e quais padrões podem ser identificados por meio da **análise dos dados públicos disponíveis**.

Objetivo

Este estudo tem como principal objetivo **analisar a desigualdade regional nos indicadores de saúde hospitalar para o público infantil (0 a 9 anos)**, com base nos dados da **Tabela de Morbidade Hospitalar do SUS (DataSUS)** e nos dados disponibilizados pelo **Instituto Brasileiro de Geografia e Estatística (IBGE)**, referentes ao ano de **2024**. A meta é identificar disparidades na qualidade e na eficácia do atendimento hospitalar entre as diferentes regiões do país, contribuindo para o entendimento das lacunas existentes e fornecendo subsídios para a formulação de políticas públicas mais equitativas.

Hipóteses a serem investigadas

A partir do problema delineado, foram elaboradas as seguintes hipóteses a serem investigadas:

- Regiões com **menor investimento** médio por internação apresentam **maior taxa de mortalidade** hospitalar infantil.
- O **tempo médio de permanência hospitalar** pode ser **maior em regiões com menor estrutura de saúde**, refletindo ineficiência no tratamento.

Essas hipóteses **orientam** a análise dos dados e buscam explorar possíveis relações entre os indicadores hospitalares e as condições estruturais e socioeconômicas regionais. A **validação ou refutação** dessas hipóteses poderá **revelar padrões relevantes** para a compreensão das **desigualdades na assistência à saúde infantil no Brasil**.

EXTRAÇÃO DE DADOS

Escolhemos relacionar a **faixa etária de 0 a 9 anos** com os diversos atributos sobre **saúde e morbidade hospitalar** do SUS, filtrando, junto disso, a **desigualdade regional** a fim de descobrir a relação entre essa e a eficácia das internações.

Sendo assim, utilizamos os dados presentes no site [DataSUS](#), onde tivemos que gerar tabelas avulsas, relacionando faixa etária e distribuição regional com os outros atributos, individualmente. **Os dados coletados foram correspondentes ao ano de 2024.**

Para atender a faixa etária de 0 a 9 anos tivemos que modificar o modelo de faixas utilizado pela base fornecida pelo DataSUS e, para isso, utilizamos o **ChatGPT**, o instruindo a realizar a soma das linhas a fim de formar um novo modelo de faixa etária. Atente-se a valores que tratam de **médias ponderadas**.

Resultado: A inteligência artificial apenas fez a soma e escreveu a tabela, porém, **não foi capaz** de gerar o arquivo .xlsx ou .csv, que seria necessário para os próximos processos.

Com isso, recorremos ao **Grok** e apresentamos a necessidade de transformar a tabela, até então em formato de texto, em uma com **formato de arquivo**. A ferramenta foi então orientada a utilizar **python** para isso.

Resultado: [Código funcional em python](#) & [faixa etaria X atributos.csv](#)

Após isso, repetimos o processo para formar a base responsável pela **distribuição regional**.

Resultado: [Código funcional em python](#) & [distribuição regional X atributos](#)

A **Taxa de Internamento** foi criada com o objetivo de estimar a **proporção da população** regional que passou por internações hospitalares em determinado período. Essa taxa é expressa em **percentual** e permite comparar a demanda por serviços hospitalares entre as diferentes regiões do Brasil, considerando o tamanho populacional de cada uma.

O cálculo foi feito a partir do **número total de internações por região, dividido pela população estimada da respectiva região em 2024**, conforme dados do IBGE. Os valores foram multiplicados por 100 para representar o resultado em percentual.

Assim, a taxa indica, por exemplo, que 8,02% da população da Região Sul foi internada, enquanto no Sudeste essa proporção foi de 6,18%, refletindo **possíveis diferenças no perfil demográfico, acesso ao sistema de saúde ou prevalência de doenças**.

Os dados populacionais por região foram obtidos a partir das **Estimativas da População 2024** divulgadas pelo **IBGE**, com data de referência em 1º de julho de 2024. Essas informações estão

disponíveis no portal oficial do IBGE:
https://ftp.ibge.gov.br/Estimativas_de_Populacao/Estimativas_2024.

Em seguida, **refinamos** a análise para o público-alvo (0 a 9 anos), utilizando a mesma fórmula, porém com base no **número de internações e na população estimada especificamente dessa faixa etária por região**.

Para isso, **cruzamos** dois conjuntos de dados:

- **Internações hospitalares por região e faixa etária:** extraídos da **Tabela de Morbidade Hospitalar do SUS (DataSUS)**.
- **População por faixa etária e região em 2024:** obtidos a partir das **Projeções da População do IBGE**, com data de referência em **1º de julho de 2024**.

As taxas obtidas representam, portanto, a **proporção de crianças** de 0 a 9 anos que foram **internadas em cada região do país**. Essa abordagem permite uma **análise mais fiel** da demanda por serviços hospitalares infantis, levando em consideração as **diferenças demográficas regionais**.

fontes: [Projeção percentual etária](#)
[ESTIMATIVAS DA POPULAÇÃO RESIDENTE NO BRASIL](#)

Resultado: [Tabela que ilustra claramente a taxa de internação e população por região](#)

DICIONÁRIO DE DADOS

A fim de **melhor organizar os dados** adquiridos e utilizados, viu-se necessária a criação de um **dicionário de dados**. Dessa forma, a definição de cada elemento em uso, seu tipo, formato em que é apresentado, atributos e usos podem ficar mais claros a qualquer interessado.

Esse foi feito com assistência do **ChatGPT**, que foi **supervisionado** durante todo o processo de forma a garantir que as **informações necessárias** e **formato de tabela** estivessem **de acordo com o esperado**.

Dessa forma, o **dicionário final** apresenta os **atributos** utilizados, que **tipo de dado** eles são e uma breve **descrição** deles.

Resultado: [Dicionário](#) classificando corretamente os tipos de dados.

UNIFICAÇÃO DAS BASES

A fim de obter uma visão mais clara sobre as desigualdades regionais no tratamento de crianças de **0 a 9 anos**, buscamos **unificar** ambas as bases de dados. Para isso, partimos do **percentual de indivíduos** nessa faixa etária **em relação ao total** e **aplicamos esses dados às regiões**, construindo assim uma tabela da faixa etária analisada com regiões X e seus atributos específicos. Em busca de **representar os dados regionais pela faixa etária alvo**, utilizamos o **ChatGPT** para calcular o valor dos **novos atributos** e gerar uma **nova tabela**.

Resultado: Tabela em texto.

Agora, utilizando da IA **Perplexity**, foi requisitado que gere um código em python que crie a tabela.

Resultado: Tabela no formato .csv.

Manualmente **ajustamos** o número de óbitos que foi **incoerente** e **implantamos a taxa de internação** nas demais bases.

Resultado: [Código funcional em python](#) & [distribuição regional X publico alvo](#)

ANÁLISE EXPLORATÓRIA DE DADOS

Relações dos dados com as hipóteses

Com o objetivo de **identificar relações** entre as **hipóteses** apresentadas e os **dados** adquiridos durante o período de pesquisa, fizemos uma **revisão** de nossas hipóteses com um **novo olhar** e com maior ciência da **situação da saúde pública infantil brasileira**. Dessa forma, essas são as **relações encontradas** pelo nosso grupo em contraste a nossas **hipóteses iniciais**:

Hipótese: Regiões com **menor investimento** médio por internação apresentam **maior taxa de mortalidade** hospitalar infantil.

A análise do gráfico de dispersão “**Valor Médio por Internação x Taxa de Mortalidade**” revela que essa hipótese **não se sustenta** de forma consistente **como imaginávamos**. A região **Norte**, por exemplo, apresenta o **menor valor médio por internação** (~R\$1.190) e a **menor taxa de mortalidade** hospitalar infantil (~0,79%), informação que **contradiz** diretamente **nossa hipótese inicial**.

Se fosse verdadeiro o caso de a taxa de mortalidade **realmente** ser maior nessa região, poderíamos inferir então que o **investimento** nas internações tem **grande influência** nos serviços prestados à essas faixas etárias, **mas esse não foi o caso**.

Isso é **ainda mais evidente** após analisar que, em contrapartida, o **Sudeste**, com um dos **maiores custos por internação** (~R\$1.820), registra também a **maior taxa de mortalidade** (~1,42%).

Esses dados sugerem que o maior investimento médio por internação **não está, necessariamente, associado a melhores resultados clínicos**, como a redução da mortalidade. Isso pode indicar que **outras variáveis**, como complexidade dos casos, desigualdade no acesso e eficiência dos serviços, **tem influência maior** sobre esses desfechos.

Hipótese: O tempo médio de permanência hospitalar pode ser **maior em regiões com menor estrutura de saúde**, refletindo ineficiência no tratamento.

A análise do gráfico “**Tempo Médio de Permanência Hospitalar por Região**” **também não confirma plenamente** essa hipótese. O **tempo médio mais elevado** é observado nas regiões **Sudeste** (5,2 dias) e **Nordeste** (5,1 dias), que **não necessariamente** representam as regiões com **menor estrutura de saúde**, especialmente no caso do **Sudeste**, que concentra parte **significativa** dos recursos hospitalares do país.

Enquanto isso, a região **Norte**, que tradicionalmente enfrenta **maiores desafios em infraestrutura**, apresenta o **menor tempo médio** de permanência (4,7 dias). Esse dado pode refletir **limitações no suporte hospitalar** (ex.: *alta precoce por indisponibilidade de leitos*) ao **invés de eficiência**, e não necessariamente uma permanência **adequada** à complexidade dos casos.

Considerações complementares:

- O **Sudeste lidera** em número **absoluto de internações e óbitos hospitalares** infantis. A alta taxa de mortalidade e o maior tempo médio de permanência **podem indicar** que **a região concentra casos mais graves** ou sofre com **sobrecarga** do sistema.
- **Sul** apresenta **boa performance relativa**: possui **alta taxa de internação**, mas **mortalidade** infantil **inferior à do Sudeste**, apesar do **valor médio** por internação ser o **mais alto** (~R\$2.000), o que **pode indicar melhor eficiência** no uso dos recursos.
- **Centro-Oeste** e **Nordeste** apresentam resultados **intermediários**, com **mortalidade próxima da média** nacional e **valores médios por internação abaixo da média**.

Conclusão estatística

As **medidas estatísticas** mostram que os **indicadores variam**, mas **não confirmam** as **hipóteses** iniciais.

O menor valor médio por internação e menor taxa de mortalidade hospitalar no Norte podem indicar subnotificação ou menor acesso, não melhor qualidade.

O tempo de permanência é bastante uniforme, sugerindo protocolos semelhantes.

CONCLUSÃO

A análise dos dados revela que, embora existam diferenças regionais nos indicadores de saúde hospitalar infantil, as hipóteses iniciais não foram confirmadas pelos resultados, mas, na verdade, refutadas por eles, revelando que não há uma relação linear direta entre investimento médio por internação e taxa de mortalidade hospitalar infantil. O menor investimento médio por internação não se traduziu em maior mortalidade hospitalar, e o tempo de permanência hospitalar mostrou-se praticamente uniforme entre as regiões, contrariando a expectativa de ineficiência nas áreas com menor estrutura.

Esses achados sugerem que a eficácia da assistência hospitalar infantil no Brasil é influenciada por múltiplos fatores, indo além do simples volume de investimento ou da estrutura disponível. Aspectos como acesso ao serviço, perfil dos casos atendidos, qualidade dos dados e protocolos clínicos adotados e outros parecem desempenhar papel relevante nas diferenças observadas. Assim, os resultados sugerem que a efetividade da assistência hospitalar infantil no SUS varia significativamente entre as regiões, e depende de um conjunto mais complexo de fatores que os inicialmente imaginados.

Ainda assim, essa análise se mostra de grande importância, porque mesmo que os fatores analisados não tenham grande influência direta sobre a taxa de mortalidade, a disparidade observada entre eles nas diferentes regiões brasileiras ainda é uma realidade existente, comprovada pelos dados. Isso abre espaço para a implementação de diversas políticas públicas em cada região, com o objetivo de torná-las mais similares entre si e buscar melhorar ainda mais o serviço público de saúde ao público infantil no futuro.