

Integração e Análise de Dados sobre Queimadas e Internações por Doenças Respiratórias no Brasil (2015–2025)

André F. da Silva¹, Héldon J. O. Albuquerque¹

¹Universidade Federal Rural de Pernambuco (UFRPE)
Caixa Postal 063 – 56.900-000 – Serra Talhada – PE – Brasil

{andre.fsilva2, heldon.jose}@ufrpe.br

Resumo. A pandemia de COVID-19 provocou uma redução atípica nas internações por doenças respiratórias registradas no Sistema de Informações Hospitalares TabNet, comprometendo a consistência das séries históricas e dificultando a análise integrada entre fatores ambientais e agravos respiratórios. Este trabalho analisa a relação entre queimadas e internações por doenças respiratórias no Brasil no período de 2015 a 2025, considerando o impacto da pandemia sobre os registros hospitalares. A abordagem metodológica baseia-se na integração e validação dos registros de internações hospitalares do SIH/TabNet, dos registros de vigilância epidemiológica de SRAG do SIVEP-Gripe/OpenDataSUS e dos dados de focos de queimadas disponibilizados pelo Instituto Nacional de Pesquisas Espaciais por meio do sistema BDQueimadas. Para viabilizar esse tratamento, o estudo emprega um processo de tratamento e consolidação de dados baseado em Python para padronização das bases, com suporte do Pentaho Data Integration no gerenciamento do fluxo e na agregação dos dados por estado e ano. A análise será apoiada pela construção de um Data Mart e pela aplicação de recursos de análise OLAP por meio do Power BI, com distinção explícita entre indicadores de internação hospitalar e indicadores de vigilância epidemiológica. Espera-se que essa abordagem permita uma avaliação mais transparente da relação entre queimadas e eventos respiratórios, contribuindo para o fortalecimento da vigilância em saúde e ambiental e oferecendo subsídios ao planejamento de ações voltadas à mitigação dos impactos das queimadas na saúde coletiva.

Palavras-chave: Queimadas; Doenças Respiratórias; Integração de Dados; Data Mart; Análise de Séries Temporais.

Abstract. The COVID-19 pandemic caused an atypical reduction in hospital admissions for respiratory diseases recorded in the Hospital Information System TabNet, compromising the consistency of historical time series and hindering integrated analyses between environmental factors and respiratory conditions. This study analyzes the relationship between wildfires and hospitalizations for respiratory diseases in Brazil from 2015 to 2025, considering the impact of the pandemic on hospital records. The methodological approach is based on the integration and validation of hospital admission records from SIH/TabNet, Severe Acute Respiratory Syndrome surveillance records from SIVEP-Gripe/OpenDataSUS, and wildfire hotspot data provided by the National Institute for Space Research through the BDQueimadas system. To support this

process, the study employs data preparation and consolidation procedures developed in Python for dataset standardization, using Pentaho Data Integration for managing the data flow and aggregating information by state and year. The analysis will be based on the construction of a Data Mart and the application of OLAP analytical resources using Power BI, with explicit methodological distinction between hospital admission indicators and epidemiological surveillance indicators. This approach is expected to enable a more transparent assessment of the relationship between wildfires and respiratory health events, contributing to environmental and health surveillance and supporting actions aimed at mitigating the health impacts of wildfires on the Brazilian population.

Keywords: Wildfires; Respiratory Diseases; Data Integration; Data Mart; Time Series Analysis.

1. Introdução

As queimadas representam um dos principais fatores de degradação ambiental em diversas regiões do mundo, influenciando processos ecológicos, climáticos e socioeconômicos. No Brasil, esses eventos se intensificam sobretudo durante o período de seca e estão associados tanto a fenômenos naturais quanto, principalmente, à ação humana, incluindo desmatamento, expansão agropecuária e manejo inadequado do solo. Segundo análises reunidas no *Curso Análise de Situação de Saúde Ambiental: ASISA-Queimadas* [Gracie et al. 2024], a recorrência e a distribuição espacial das queimadas no território brasileiro refletem padrões históricos de uso e ocupação do solo e constituem um relevante desafio para a gestão ambiental e para a vigilância em saúde.

A combustão da biomassa libera uma ampla gama de poluentes atmosféricos, tais como material particulado fino ($MP_{2,5}$), material particulado grosso (MP_{10}), monóxido de carbono (CO), dióxido de nitrogênio (NO_2) e dióxido de enxofre (SO_2). A exposição prolongada ou intensa a essas substâncias está associada a impactos negativos na qualidade do ar, redução da visibilidade, aumento do risco de incêndios subsequentes e alterações climáticas locais. Estudos como o de [Godoy 2023] evidenciam que concentrações elevadas de poluentes atmosféricos estão associadas a maior risco de agravamento de condições respiratórias, especialmente em centros urbanos e regiões próximas a áreas críticas de queima.

Do ponto de vista epidemiológico, a literatura tem demonstrado que a combinação entre queimadas e poluição atmosférica intensifica a carga de doenças respiratórias no Brasil. [Schroeder 2022] identificaram que a alta exposição ao fogo no estado do Pará esteve associada ao aumento de internações por doenças respiratórias e à maior mortalidade durante o período da pandemia de COVID-19. Resultados semelhantes foram observados por [Albuquerque et al. 2023], que apontaram distorções significativas nas séries históricas de internações por doenças respiratórias não relacionadas à COVID-19, especialmente a partir de 2020. Além disso, [da Silva 2023] destacam como inconsistências nos sistemas de informação podem comprometer a interpretação epidemiológica, evidenciando a necessidade de análises integradas e rigorosamente padronizadas.

Diante desse cenário, torna-se essencial desenvolver abordagens que integrem as diferentes bases de dados, identifiquem padrões temporais e espaciais e apoiem a tomada de decisão com ferramentas analíticas. Considerando as lacunas metodológicas identi-

cadados nos estudos anteriores — especialmente no que se refere à ausência de ambientes analíticos unificados e à falta de padronização entre múltiplas bases públicas — este estudo propõe uma integração sistemática entre dados ambientais e dados de saúde, criando as condições necessárias para análises mais consistentes acerca dos efeitos das queimadas nos agravos respiratórios das populações expostas.

Este trabalho tem como objetivo geral analisar a relação entre queimadas e internações por doenças respiratórias no Brasil, utilizando dados abertos de órgãos públicos de saúde e meio ambiente, abrangendo o período de 2015 a 2025. A análise será realizada por meio do desenvolvimento de uma ferramenta OLAP, permitindo a exploração interativa dos dados e a identificação de padrões e correlações relevantes. Para alcançar esse objetivo geral, o estudo define os seguintes objetivos específicos: 1. integrar, padronizar e consolidar os dados provenientes das bases TabNet, OpenDataSUS e BDQueimadas, garantindo sua consistência e possibilitando a construção de um Data Mart confiável para o período de análise; 2. avaliar a série histórica de internações por doenças respiratórias entre 2015 e 2025, utilizando os registros de SRAG do SIVEP-Gripe como fonte complementar de vigilância epidemiológica para contextualizar distorções decorrentes da pandemia; 3. investigar a associação entre a ocorrência de focos de queimadas e o aumento das internações por doenças respiratórias não relacionadas à COVID-19, aplicando métodos estatísticos de correlação às séries temporais integradas. Por fim, o estudo busca responder às seguintes questões de pesquisa:

- **QP1:** De que forma a aplicação de técnicas de Engenharia de Dados pode assegurar a integridade e a padronização das bases TabNet, OpenDataSUS e BDQueimadas, resultando em um Data Mart confiável para o período de análise?
- **QP2:** Qual é a magnitude da redução nas internações por doenças respiratórias observada a partir de 2020 na base TabNet, e de que modo a incorporação dos registros não associados à COVID-19 da base OpenDataSUS ajusta a curva da série histórica?
- **QP3:** Existe correlação estatisticamente significativa entre a ocorrência de focos de queimadas e o aumento das internações por doenças respiratórias não relacionadas à COVID-19 no Brasil, após a integração e o ajuste da série temporal?

Este artigo será estruturado da seguinte forma: na Seção 2, serão apresentados os trabalhos relacionados e suas contribuições para este estudo; na Seção 3, será descrito o método proposto; na Seção 4, serão apresentados os experimentos a serem conduzidos e a análise dos resultados esperados; e, por fim, na Seção 5, serão apresentadas as conclusões e as perspectivas de trabalhos futuros.

2. Referencial Teórico

O estudo da relação entre queimadas e agravos respiratórios no Brasil envolve múltiplas dimensões: ambiental, epidemiológica, informacional e tecnológica. Para fornecer suporte conceitual ao desenvolvimento deste trabalho, esta seção apresenta um panorama sobre as queimadas no Brasil e seus impactos, os efeitos da poluição resultante na saúde respiratória, os principais sistemas de informação utilizados no monitoramento de saúde e meio ambiente e, por fim, fundamentos de engenharia de dados aplicados ao tratamento e integração de bases públicas.

2.1. Queimadas no Brasil e seus Impactos Ambientais

As queimadas no território brasileiro constituem um fenômeno complexo, associado tanto a fatores antrópicos quanto a condições climáticas que favorecem a combustão da biomassa. Segundo [Gracie et al. 2024], as queimadas possuem padrões espaciais e temporais diretamente relacionados ao uso e ocupação do solo, com maior concentração em biomas como Amazônia, Cerrado e Pantanal. Além disso, o curso ASISA-Queimadas destaca que eventos de fogo, especialmente em períodos de estiagem, desencadeiam processos de degradação ambiental que incluem perda de biodiversidade, alteração dos ciclos hidrológicos e emissões de gases poluentes.

O estudo de [Sena 2024] reforça que as queimadas impactam de forma significativa a integridade ecológica e as condições atmosféricas, contribuindo para o aumento das concentrações de partículas e gases tóxicos. Essas emissões possuem efeitos imediatos e cumulativos sobre a qualidade do ar e, consequentemente, sobre a saúde humana, especialmente em áreas urbanas ou próximas a regiões que registram alta incidência de focos de calor.

2.2. Infecções Respiratórias

As infecções respiratórias constituem um conjunto de doenças que acometem o trato respiratório superior ou inferior e figuram entre as principais causas de morbidade no sistema de saúde brasileiro. Conforme destacado por [Filho et al. 2017], essas condições podem ser desencadeadas por vírus, bactérias ou outros agentes infecciosos, apresentando amplo espectro clínico que varia desde quadros leves, como resfriados e rinofaringites, até formas mais graves, incluindo bronquiolite, pneumonia e exacerbações de doenças crônicas, como a DPOC.

A alta prevalência dessas infecções está diretamente relacionada à facilidade de transmissão dos agentes etiológicos e à exposição ambiental, sendo especialmente crítica em grupos vulneráveis — crianças, idosos, gestantes e pessoas com comorbidades. Além disso, sua evolução clínica é fortemente influenciada pela qualidade do ar, o que reforça a relevância desse grupo de doenças no contexto de análises que envolvem queimadas e poluentes atmosféricos.

Compreender esses aspectos clínicos é essencial para interpretar as oscilações observadas nas internações hospitalares ao longo da série histórica. Como a ocorrência e a gravidade das infecções respiratórias respondem rapidamente a mudanças na qualidade do ar, períodos marcados por maior concentração de poluentes tendem a refletir-se em aumentos nos registros de hospitalização. Assim, essa perspectiva clínica explica porque determinados períodos apresentam variações nos registros de internação, sobretudo quando coincidem com picos de poluição.

2.3. Efeitos das Queimadas na Saúde Respiratória

A relação entre exposição à fumaça das queimadas e o aumento de agravos respiratórios está amplamente documentada. [Caumo et al. 2022] apresentaram uma síntese das evidências das últimas duas décadas, demonstrando que a inalação de material particulado fino ($MP_{2,5}$), gases tóxicos e compostos orgânicos voláteis está associada a piora de quadros de asma, bronquite, doença pulmonar obstrutiva crônica (DPOC) e infecções

respiratórias agudas. Os autores reforçam que populações mais expostas, como crianças, idosos e indivíduos com comorbidades respiratórias, apresentam maior vulnerabilidade.

Da mesma forma, [Schroeder 2022] demonstrou que a alta incidência de queimadas no Pará está estatisticamente relacionada ao aumento das internações por doenças respiratórias, além de potencializar a gravidade de casos de COVID-19 em regiões críticas. [Godoy 2023] corroborou esses achados ao analisar padrões temporais de poluentes atmosféricos no estado de São Paulo e associá-los diretamente ao aumento das internações hospitalares nos dias subsequentes à exposição. Em conjunto, esses estudos evidenciam que a fumaça das queimadas é um fator determinante na dinâmica dos agravos respiratórios no Brasil.

2.4. Sistemas de Informação em Saúde e Meio Ambiente

A análise integrada entre queimadas e saúde respiratória depende de bases de dados confiáveis e atualizadas. [de Souza et al. 2020] destacam que o sistema de informação é o eixo estruturante da vigilância em saúde, sendo fundamental para monitorar padrões epidemiológicos, subsidiar políticas públicas e compreender emergências sanitárias. Os autores analisam a consolidação dos sistemas durante a pandemia de COVID-19, enfatizando desafios como duplicidades, inconsistências e defasamentos temporais nos registros.

No contexto ambiental, o BDQueimadas/INPE é a principal plataforma nacional para monitoramento de focos de calor detectados por satélites. Ele fornece dados diários e georreferenciados, permitindo análises espaciais e temporais sobre a dinâmica das queimadas. Quando comparados a sistemas de saúde como SIH/TabNet e SIVEP-Gripe/OpenDataSUS, surgem desafios de integração decorrentes de granularidade distinta (diária, municipal e estadual), nomenclaturas divergentes e diferentes modelos de codificação.

Essa heterogeneidade reforça a necessidade de processos sistemáticos de integração e padronização, essenciais para permitir análises combinadas e consistentes sobre indicadores ambientais e de saúde.

2.5. Engenharia de Dados Aplicada à Saúde e Meio Ambiente

A Engenharia de Dados desempenha papel crucial na integração, padronização e consolidação de grandes bases de dados provenientes de fontes distintas. [Sousa 2020], ao desenvolver um Data Warehouse e processos ETL para dados da saúde pública, destaca desafios como granularidade inconsistente, presença de dados ausentes, diferenças de nomenclatura e incompatibilidades estruturais. Esses problemas também estão presentes no trabalho de [Oliveira et al. 2025], que propõe um framework integrado para análise de dados climáticos e epidemiológicos, demonstrando que a combinação de fontes diversas permite análises mais robustas e preditivas sobre a dinâmica de agravos sensíveis ao clima.

Ambos os estudos evidenciam que processos de ETL (Extração, Transformação e Carga) são essenciais para viabilizar análises em ambientes analíticos, como Data Warehouses e Data Marts, ao garantir que os dados sejam armazenados de forma consistente, padronizada e orientada a análises temporais.

No contexto do presente trabalho, a Engenharia de Dados será aplicada para in-

tegrar as bases TabNet, OpenDataSUS e BDQueimadas, corrigir distorções introduzidas pela pandemia, padronizar unidades federativas, consolidar registros anuais e preparar as bases para análises exploratórias e analíticas com suporte a OLAP. Assim, a adoção de um processo estruturado de integração irá contribuir para ampliar a confiabilidade do conjunto de dados e permitir investigações mais precisas sobre a relação entre queimadas e agravos respiratórios.

3. Trabalhos Relacionados

Estudos recentes têm buscado compreender a relação entre queimadas e o aumento de doenças respiratórias no Brasil, explorando diferentes recortes espaciais, fontes de dados e enfoques metodológicos. Essas pesquisas refletem o crescente interesse científico em mensurar os impactos ambientais sobre a saúde humana, especialmente em regiões onde os episódios de queimadas são recorrentes e coincidem com picos de internações hospitalares por doenças respiratórias. Em comum, esses trabalhos contribuem para o entendimento dos padrões espaciais e temporais desses eventos, oferecendo subsídios relevantes para análises mais integradas entre os campos da saúde e do meio ambiente.

O estudo de [da Silva 2023] investigou as notificações e subnotificações de casos de COVID-19 no estado do Rio de Janeiro, evidenciando como erros de registro e inconsistências nos sistemas de informação afetam a compreensão dos dados e, consequentemente, o planejamento de ações em saúde pública. Já [Albuquerque et al. 2023] avaliaram a influência da pandemia de COVID-19 nas taxas de internações hospitalares, óbitos intra-hospitalares e letalidade por doenças respiratórias não associadas à COVID-19 no Brasil, entre 2015 e 2020. Os autores observaram uma redução significativa — cerca de 42% — nas internações durante o ano de 2020, indicando um possível viés nas séries históricas em decorrência da pandemia.

[Schroeder 2022] analisou a relação entre incêndios florestais e agravos respiratórios no estado do Pará durante o período de 2015 a 2020, utilizando dados do BDQueimadas, do TabNet — referentes às internações por doenças respiratórias classificadas segundo o CID-10 — e do OpenDataSUS, que disponibiliza registros sobre COVID-19. O estudo empregou técnicas de agrupamento (*k-means*) e modelagem temporal com o modelo ARIMAX para avaliar a associação entre queimadas e indicadores de saúde. Os resultados demonstraram que a ocorrência de incêndios esteve associada ao aumento das internações por doenças respiratórias e à maior mortalidade por COVID-19 em regiões com alta exposição ao fogo.

[Godoy 2023] investigou padrões espaciais e temporais de poluentes atmosféricos no estado de São Paulo, aplicando técnicas de agrupamento e regras de associação temporais para identificar correlações entre episódios críticos de poluição e variáveis meteorológicas. Os resultados revelaram que concentrações elevadas de material particulado fino ($MP_{2,5}$), material particulado grosso (MP_{10}) e ozônio troposférico (O_3) estiveram fortemente associadas ao aumento das internações por doenças respiratórias, especialmente nos dias subsequentes à exposição.

Os quatro trabalhos analisados convergem ao demonstrar que: 1) queimadas e poluição atmosférica elevam o risco de agravos respiratórios; 2) há limitações importantes nos sistemas de informação; 3) métodos estatísticos e modelos temporais são essenciais

para investigar relações entre poluição e saúde.

No entanto, nenhum desses estudos integra simultaneamente os dados do SIH/TabNet, SIVEP-Gripe e BDQueimadas em um *Data Mart* unificado ou emprega um processo completo de Engenharia de Dados com um pipeline de ETL totalmente estruturado, como proposto neste trabalho.

3.1. Estudo Comparativo

Para sintetizar as características metodológicas, objetivos e limitações dos trabalhos identificados, foi realizada uma análise comparativa entre os quatro estudos mais relevantes: [da Silva 2023], [Albuquerque et al. 2023], [Schroeder 2022] e [Godoy 2023]. A Figura 1 apresenta a visão consolidada dessa comparação.

	Objetivo	Motivação	Referencial Teórico	Trabalhos Relacionados	Método	Principais Resultados	Trabalhos Futuros
SCHROEDER (2022)	Analisar a associação Queimadas e DR/COVID-19.	Usar <i>Clustering</i> para segmentar cidades por risco e otimizar gestão.	Focos de Incêndios, Poluentes, Mineração de Dados	Agrupamento Geográfico (K-means). Modelagem Estatística PM2.5/COVID-19.	Agrupamento (K-means) e Modelagem de Séries Temporais (ARIMAX).	Incêndios levam ao aumento de internações DR e maior mortalidade por COVID-19.	Inclusão de variáveis socioeconômicas e ampliação da capacidade preditiva do modelo.
Da Silva (2023)	Investigar erros de notificação de casos de COVID-19 no RJ.	Avaliar a fidedignidade dos dados e o impacto nas políticas de saúde.	Erros Crônicos de Notificação e Fragilidade do SUS.	Estudos sobre subnotificação/atraso de casos no Brasil.	Triangulação de Dados Secundários (DATASUS) e análise de consistência.	Erros inviabilizaram estratégias de saúde pública.	Aprimoramento do controle de fidedignidade e da integridade dos registros oficiais.
Albuquerque (2023)	Avaliar a influência da pandemia nas internações e letalidade por DR não-COVID.	Investigar mudanças nos padrões de internação devido à sobrecarga do SUS.	Sobrecarga Sistêmica e Oferta/Demanda de Serviços de Saúde.	Queda de internações (Câncer, DCV); Interrupção generalizada no acesso ao atendimento.	Estudo Observacional com análise de Séries Temporais (2015-2020).	Redução de 42% nas Internações; Aumento de 60% na Letalidade IH.	Investigação de erros de codificação diagnóstica e subnotificação de internações.

Figura 1. Comparação entre os trabalhos analisados, destacando objetivos, métodos, limitações e contribuições.

A análise comparativa evidencia que cada estudo enfoca dimensões distintas do problema. [da Silva 2023] enfatiza problemas de inconsistência, subnotificação e falhas estruturais nos sistemas de informação em saúde, apontando limitações que prejudicam a confiabilidade das séries temporais, especialmente em períodos críticos. [Albuquerque et al. 2023] analisam os impactos da pandemia nos registros hospitalares e demonstram que as internações por doenças respiratórias foram distorcidas a partir de 2020, reforçando a necessidade de filtragem e correção contextualizada dos dados. [Schroeder 2022] investiga diretamente a associação entre queimadas e agravos respiratórios no Pará, utilizando métodos de modelagem temporal que demonstram relação estatisticamente significativa entre os focos de calor e o aumento subsequente de internações. [Godoy 2023], por sua vez, analisa séries temporais de poluentes atmosféricos no estado de São Paulo e demonstra que a elevação de material particulado e gases tóxicos está diretamente relacionada a internações por doenças respiratórias nos dias seguintes à exposição.

Apesar das contribuições relevantes, nenhum desses trabalhos realiza a padronização e integração completa das bases TabNet, OpenDataSUS e BDQueimadas; tampouco implementa um ambiente analítico unificado baseado em um *Data Mart* com suporte a operações OLAP. Assim, o presente estudo avança em relação à literatura ao

propor uma abordagem integrada que combina múltiplas bases públicas, aplica princípios de Engenharia de Dados e estabelece um processo sistemático de ETL para garantir consistência, comparabilidade e robustez analítica.

Nesse contexto, o presente trabalho complementa os estudos anteriores ao propor uma abordagem de integração e análise de dados que aprimora os processos de extração, transformação e carga (ETL), favorecendo a padronização e a consistência das informações provenientes de diferentes fontes públicas. Essa integração visa ampliar o potencial analítico e contribuir para uma compreensão mais precisa da relação entre queimadas e internações por doenças respiratórias no Brasil.

4. Materiais e Métodos

4.1. Dataset

A abordagem metodológica a ser adotada neste estudo consiste na integração de bases ambientais e de saúde pública, aliada à aplicação de técnicas de Engenharia de Dados para a padronização, consolidação e análise das informações. Serão apresentados os conjuntos de dados a serem utilizados, os indicadores e atributos definidos para a análise, a estrutura do modelo dimensional proposto e o processo de extração, transformação e carga (ETL) previsto para viabilizar a construção de um ambiente analítico consistente.

A análise dos dados será estruturada a partir de um modelo dimensional, inspirado nos princípios descritos por [Kimball and Ross 1996]. O período selecionado para a investigação compreenderá os anos de 2015 a 2025, permitindo observar tendências históricas e tratar possíveis distorções nos dados ocorridas durante o período da pandemia de COVID-19.

As informações utilizadas serão obtidas de três fontes públicas: o Sistema de Informações Hospitalares (SIH/TabNet), que disponibiliza registros de internações por doenças respiratórias conforme a Classificação Internacional de Doenças – 10ª Revisão (CID-10); o Sistema de Informação da Vigilância Epidemiológica da Gripe (SIVEP-Gripe), utilizado para identificar casos de Síndrome Respiratória Aguda Grave (SRAG) não associados à COVID-19; e o Instituto Nacional de Pesquisas Espaciais (INPE), responsável pelos dados de focos de queimadas detectados por satélite em todo o território brasileiro.

4.2. Indicadores

A definição dos indicadores de análise — número anual de internações hospitalares, número anual de casos de SRAG não associados à COVID-19 e número anual de focos de queimadas — se apoiará em estudos prévios, como [da Silva 2023], [Albuquerque et al. 2023], [Godoy 2023] e [Schroeder 2022]. Essas metodologias orientam a seleção das variáveis e métricas adotadas. Os principais indicadores, suas descrições e respectivas fontes de dados estão apresentados na Figura 2.

Indicador	Descrição	Fonte de Dados
Número de internações por doenças respiratórias	Total de internações hospitalares registradas por doenças respiratórias (CID-10: J00–J99). Os dados estão disponíveis na granularidade mensal e por estado, sendo posteriormente agregados por ano para análise.	DATASUS/SIH (TabNet)
Casos de Síndrome Respiratória Aguda Grave (SRAG) não associados à COVID-19	Quantidade de registros de SRAG classificados como influenza, outro vírus respiratório, outro agente etiológico ou não especificado (valores 1–4 da coluna “classificação final dos casos”), excluindo os casos confirmados de COVID-19 (valor 5). Os dados estão na granularidade diária e por cidade, sendo agregados por estado e ano.	SIVEP-Gripe (OpenDataSUS)
Número de focos de queimadas	Total de focos de queimadas detectados por satélite. Os dados estão na granularidade diária e por cidade, sendo agregados por estado e ano para compatibilizar a granularidade com as demais fontes.	INPE (BDQueimadas)

Figura 2. Indicadores de análise, com suas descrições e fontes de dados.

4.3. Atributos

Os casos de SRAG não associados à COVID-19 serão identificados a partir da coluna “classificação final dos casos” da base SIVEP-Gripe, conforme o dicionário oficial do sistema (2023). Os códigos 1 (SRAG por influenza), 2 (SRAG por outro vírus respiratório), 3 (SRAG por outro agente etiológico) e 4 (SRAG não especificado) serão considerados na análise, enquanto as linhas com código 5, referentes a casos confirmados de COVID-19, serão excluídas.

Como as bases INPE e SIVEP-Gripe possuem granularidade diária e municipal, seus registros serão posteriormente agregados por estado e por ano, de modo a garantir compatibilidade com o SIH/TabNet, que já se encontra estruturado em nível estadual e mensal. Os dados do SIVEP-Gripe serão analisados como fonte complementar de vigilância epidemiológica, sem substituição direta dos registros hospitalares do SIH/TabNet, uma vez que os sistemas adotam finalidades e campos temporais distintos.

4.4. Modelo Dimensional

O modelo dimensional seguirá um esquema estrela, composto principalmente pela tabela `fato_fato_queimadas_saude`, responsável por consolidar as métricas ambientais, hospitalares e epidemiológicas utilizadas na análise, incluindo totais de queimadas, internações por diferentes grupos de doenças respiratórias, registros de SRAG/COVID-19, óbitos, precipitação, dias sem chuva, taxa de fogo e taxa de mortalidade. Complementarmente, a tabela `fato_datasus_cid` permitirá o detalhamento das internações hospitalares por CID. Essas tabelas `fato` se relacionam com as dimensões `dim_tempo`, `dim_localidade`, `dim_bioma` e `dim_cid`, possibilitando consultas analíticas por período, Unidade Federativa, região, bioma e classificação da doença. Essa estrutura visa facilitar a exploração interativa dos dados por ferramentas OLAP e está apresentada na Figura 3.

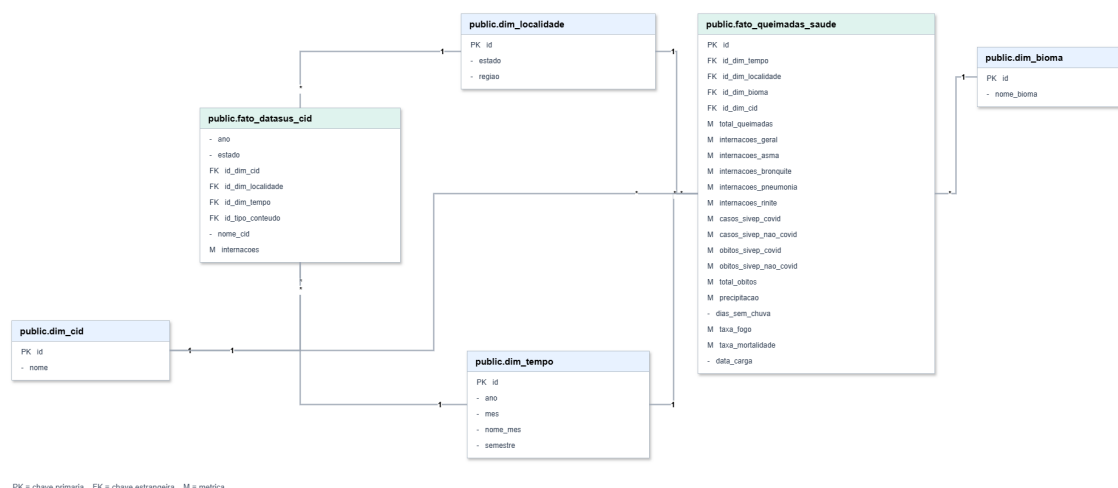


Figura 3. Modelo dimensional do Data Mart.

4.5. Extração, Transformação e Carga (ETL)

O processo de integração dos dados ocorrerá em duas etapas principais. No primeiro estágio, os arquivos serão carregados por meio do Django Admin, onde scripts em Python realizarão a leitura, verificação e correção de valores ausentes (NaN), além da padronização dos nomes dos estados. Os registros do SIVEP-Gripe serão tratados de acordo com a classificação final dos casos e analisados separadamente dos indicadores de internação hospitalar, especialmente nos recortes relacionados à COVID-19 e às demais etiologias de SRAG. Os dados tratados serão temporariamente armazenados em dicionários Python e posteriormente salvos no banco de dados PostgreSQL, que também receberá a estrutura do modelo dimensional.

No segundo estágio, o Pentaho Data Integration (PDI) realizará agregações temporais e espaciais no PostgreSQL, uniformizando todas as bases na granularidade anual e estadual e consolidando as três fontes em um ambiente integrado de análise. Esses procedimentos garantem a consistência entre as três fontes de dados e possibilitam a construção de um Data Mart confiável. Por fim, o Power BI será conectado ao PDI para criação de visualizações e painéis interativos, permitindo análises comparativas entre queimadas e internações hospitalares por doenças respiratórias ao longo de todo o período investigado. A arquitetura completa do processo está ilustrada na Figura 4.

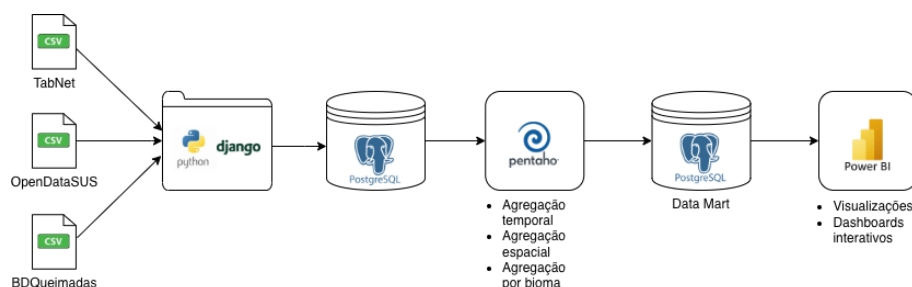


Figura 4. Fluxo de ingestão, padronização, agregação e análise envolvendo arquivos CSV, PostgreSQL, Pentaho Data Integration e Power BI.

4.6. Validação dos Resultados

A validação dos resultados foi estruturada como uma etapa de conferência entre os indicadores consolidados no Power BI e fontes oficiais utilizadas como referência para dados ambientais, hospitalares e epidemiológicos. Essa verificação teve como objetivo avaliar a aderência dos totais agregados por ano e Unidade Federativa, bem como identificar divergências decorrentes de diferenças metodológicas entre os sistemas consultados. Para isso, foram considerados o Banco de Dados de SRAG do OpenDataSUS/SIVEP-Gripe, o Sistema de Informações Hospitalares do SUS (SIH/SUS) via TABNET/DATASUS, o portal DATASUS, o portal TerraBrasilis/Queimadas e o boletim técnico Infoqueima do Programa Queimadas/INPE [Brasil. Ministério da Saúde 2026, Brasil. Ministério da Saúde. DATASUS 2026b, Brasil. Ministério da Saúde. DATASUS 2026a, Instituto Nacional de Pesquisas Espaciais 2026, Instituto Nacional de Pesquisas Espaciais 2022].

Nos dados ambientais, a comparação indicou compatibilidade entre os valores apresentados no painel e os boletins oficiais do INPE. Como exemplo, o boletim Infoqueima de junho de 2022 registrou 2.210 focos de fogo ativo no estado de Mato Grosso, mesmo total observado no painel para a mesma Unidade Federativa e período. Esse resultado reforça a consistência da base de queimadas utilizada no Data Mart em relação aos produtos oficiais do Programa Queimadas/INPE.

Nos dados de saúde, a conferência exigiu maior cautela, pois SIH/SUS e SIVEP-Gripe possuem finalidades distintas. O SIH/SUS registra internações hospitalares no âmbito do SUS, enquanto o SIVEP-Gripe é um sistema de vigilância epidemiológica voltado ao registro de casos e óbitos por Síndrome Respiratória Aguda Grave (SRAG), incluindo COVID-19, influenza e outros vírus respiratórios. Assim, embora ambos estejam relacionados a agravos respiratórios, seus registros não representam exatamente o mesmo evento e não devem ser comparados como medidas equivalentes.

A comparação com o TABNET/DATASUS apresentou diferença entre o total oficial consultado e o valor consolidado no painel. Em 2024, o total nacional consultado no TABNET para internações por doenças do aparelho respiratório foi de 1.363.903 registros, enquanto o painel apresentou 1.341.677 internações, diferença absoluta de 22.226 registros, correspondente a aproximadamente 1,63% do total oficial. Essa diferença pode estar associada a filtros aplicados no painel, atualizações posteriores da base, critérios de agrupamento por CID/lista de morbidade, exclusões por inconsistência ou diferenças no recorte temporal utilizado na extração.

Por outro lado, a comparação dos dados do SIVEP-Gripe com os dados unificados disponibilizados pelo Ministério da Saúde revelou diferenças mais expressivas em determinados recortes. No estado da Bahia, em 2020, o painel apresentou 23.831 casos, enquanto os dados unificados consultados indicaram 15.361 casos para o mesmo estado e ano. Essa divergência não invalida a análise, mas evidencia que os critérios utilizados no painel e nos produtos oficiais não são necessariamente equivalentes. Entre os fatores explicativos estão o uso da data de internação no painel, enquanto boletins e painéis epidemiológicos frequentemente organizam os registros por semana epidemiológica de início dos sintomas, notificação ou encerramento; diferenças entre casos notificados, confirmados, encerrados e classificados; revisões contínuas das

notificações; e filtros específicos definidos nos scripts oficiais do Ministério da Saúde [Brasil. Ministério da Saúde 2026, Brasil. Ministério da Saúde 2022].

Indicador validado	Fonte/evidência	Valor	Interpretação
Internações respiratórias em 2024	TABNET/DATASUS (evidência no Anexo A)	1.363.903	Total oficial consultado para morbidade hospitalar por doenças do aparelho respiratório.
Internações respiratórias em 2024	Painel Power BI (evidência no Anexo A)	1.341.677	Valor consolidado no painel, com diferença de 22.226 registros em relação ao TABNET.
Focos de queimadas em Mato Grosso, junho/2022	Boletim Infoqueima/INPE (evidência no Anexo A)	2.210	Total publicado no boletim técnico mensal do Programa Queimadas/INPE.
Focos de queimadas em Mato Grosso, junho/2022	Painel Power BI (evidência no Anexo A)	2.210	Valor reproduzido integralmente no painel, indicando aderência à fonte oficial.

Figura 5. Comparação entre valores oficiais e valores consolidados no painel Power BI.

Dessa forma, os dados do SIVEP-Gripe foram tratados como indicadores de vigilância epidemiológica de SRAG, e não como medida direta de internações hospitalares. Para a variável de internações, a fonte mais adequada permanece sendo o SIH/SUS, por meio do TABNET/DATASUS, uma vez que esse sistema registra a produção hospitalar do SUS. O SIVEP-Gripe, por sua vez, contribui para contextualizar o comportamento dos casos graves e da COVID-19 no período pandêmico, mas sua comparação direta com internações deve ser apresentada como limitação metodológica.

A utilização do Power BI, portanto, foi orientada por uma leitura crítica das bases públicas. O painel não foi empregado apenas como ferramenta de visualização, mas como ambiente de integração, filtragem e comparação exploratória entre bases oficiais. Essa etapa permitiu distinguir indicadores com maior aderência aos boletins oficiais, como os focos de queimadas do INPE, daqueles que exigem ressalvas metodológicas, como os registros do SIVEP-Gripe. Essa distinção aumenta a transparência da análise e reduz o risco de interpretações inadequadas sobre a relação entre queimadas, hospitalizações e registros epidemiológicos de SRAG/COVID-19.

5. Resultados

A análise exploratória foi conduzida a partir dos painéis desenvolvidos no Power BI, alimentados pelo Data Mart descrito na Seção 4. As visualizações foram utilizadas para verificar a coerência temporal e espacial dos indicadores, identificar padrões preliminares e orientar a validação com relatórios e boletins oficiais. Os resultados permitem avaliar a consistência das bases integradas e os limites de comparação entre indicadores ambientais, hospitalares e epidemiológicos.

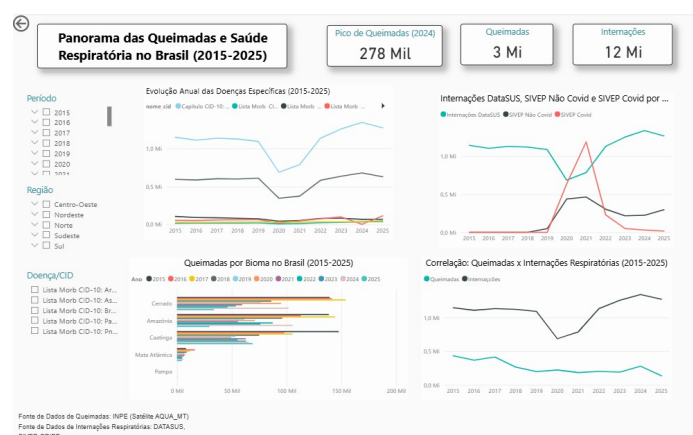


Figura 6. Visão geral do painel de queimadas e saúde respiratória no Brasil (2015–2025).

5.1. Integração das Bases e Avaliação da Consistência (QP1)

A etapa referente à QP1 evidenciou que a integração entre BDQueimadas, SIH/SUS e SIVEP-Gripe exige padronização cuidadosa de granularidade, recorte temporal e unidade territorial. No caso das queimadas, os dados agregados por bioma e Unidade Federativa apresentaram coerência com os boletins do INPE, reforçando a consistência do processo de carga e agregação. A Figura 7 apresenta a distribuição territorial dos focos de queimadas por bioma e estado, permitindo observar a concentração dos registros em Unidades Federativas associadas à Amazônia e ao Cerrado.

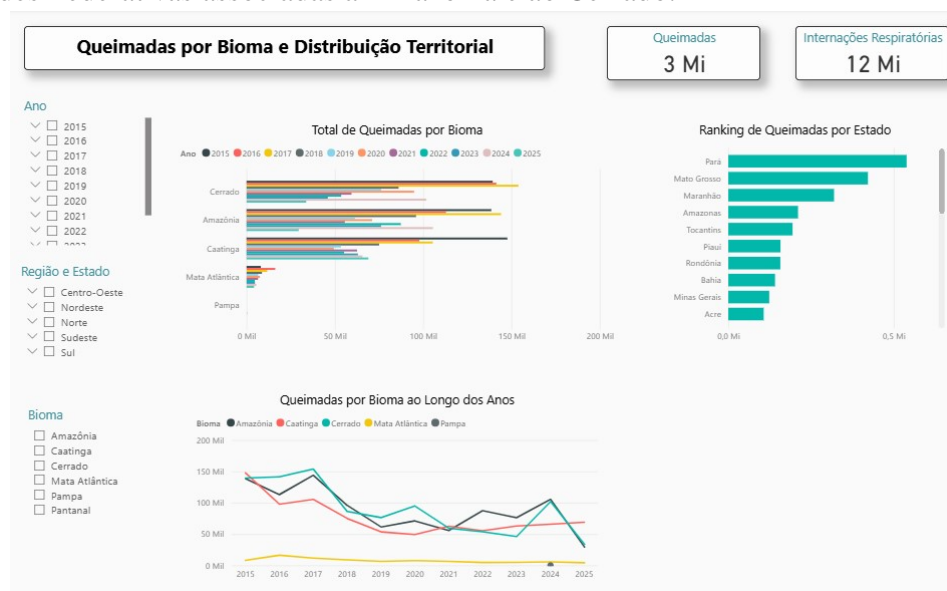


Figura 7. Distribuição das queimadas por bioma e Unidade Federativa no painel Power BI.

5.2. Tendências Históricas de Queimadas e Internações (QP2)

As séries temporais indicam oscilações importantes nos focos de queimadas ao longo do período analisado, com maior participação dos biomas Amazônia, Cerrado e Caatinga. A Figura 8 sintetiza essa evolução anual e permite identificar anos de maior intensidade do fogo, que serão posteriormente comparados aos indicadores hospitalares e epidemiológicos.



Figura 8. Evolução anual dos focos de queimadas por bioma no painel Power BI.

A análise regional também demonstrou diferenças expressivas entre estados com maior número de focos de queimadas e estados com maior volume de internações respiratórias. Conforme apresentado na Figura 9, os maiores totais de queimadas concentram-se principalmente em estados da região Norte e do Centro-Oeste, enquanto as internações respiratórias apresentam maior peso em Unidades Federativas populosas, como São Paulo e Minas Gerais. Esse resultado reforça a necessidade de considerar a dimensão territorial e populacional na interpretação das associações.

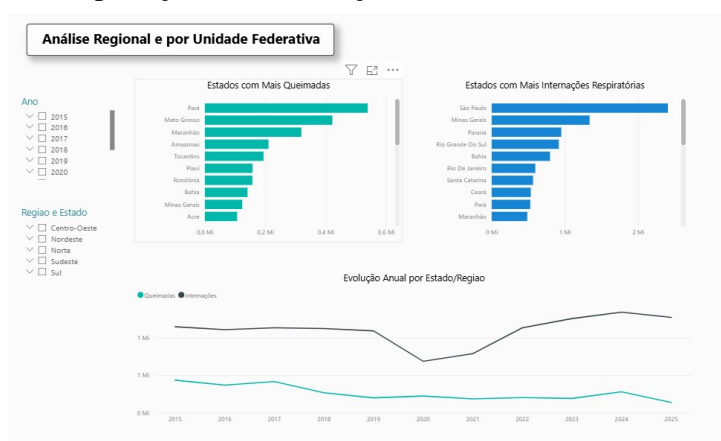


Figura 9. Análise regional e por Unidade Federativa dos indicadores de queimadas e internações respiratórias.

Os resultados também evidenciam o impacto da pandemia de COVID-19 sobre a série histórica de internações por doenças respiratórias registrada no SIH/SUS. Observa-

se uma redução expressiva dos registros a partir de 2020, comportamento compatível com os achados de [Albuquerque et al. 2023], que identificaram diminuição aproximada de 42% nas internações por doenças respiratórias não associadas à COVID-19 durante o primeiro ano da pandemia. No presente estudo, essa descontinuidade reforça a importância de interpretar os indicadores hospitalares em conjunto com os registros epidemiológicos do SIVEP-Gripe. Embora esses registros não sejam empregados como substitutos das internações do SIH/SUS, sua incorporação ao ambiente analítico permite contextualizar os efeitos da pandemia sobre a série histórica, contribuindo para uma interpretação mais consistente da evolução temporal dos indicadores respiratórios.

5.3. Comparação entre DATASUS e SIVEP-Gripe no Período Pandêmico

A comparação entre DATASUS e SIVEP-Gripe evidenciou que os dois sistemas não devem ser tratados como equivalentes. A Figura 10 mostra que os registros do SIVEP-Gripe apresentam comportamento fortemente influenciado pelo período pandêmico, especialmente em 2020 e 2021, enquanto a série de internações do SIH/SUS segue outra lógica de registro. Essa diferença reforça a decisão metodológica de utilizar o SIH/SUS como fonte principal para internações hospitalares e o SIVEP-Gripe como fonte complementar para contextualização epidemiológica da SRAG e da COVID-19.

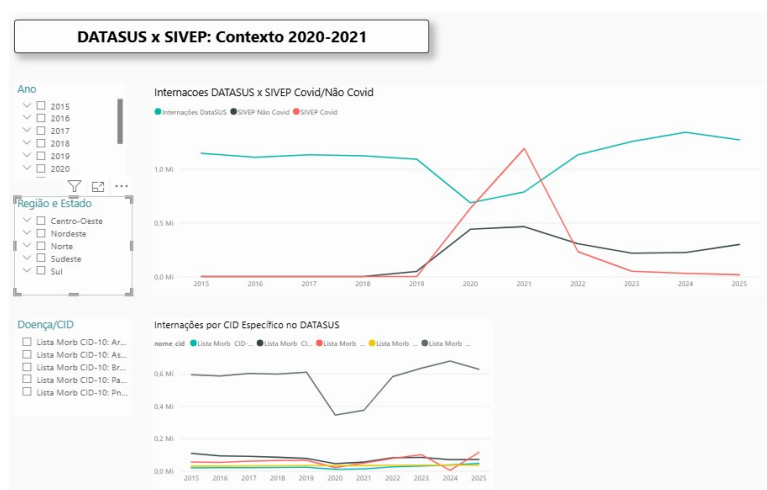


Figura 10. Comparação exploratória entre internações do DATASUS/SIH e registros de SRAG do SIVEP-Gripe.

5.4. Associação entre Focos de Queimadas e Internações (QP3)

A investigação da associação entre queimadas e internações foi realizada por meio dos coeficientes de correlação de Pearson e Spearman, aplicados aos dados mensais alinhados entre focos de queimadas e internações respiratórias. A análise considerou 3.440 observações, abrangendo o período de 2015 a 2025, sem aplicação de defasagem temporal entre as variáveis. Também foram considerados os p-valores associados aos testes, a fim de avaliar a significância estatística dos coeficientes obtidos. Os resultados estão sintetizados na Figura 11.

Método	Coefficiente	p-valor	Amostra	Período	Defasagem
Pearson	-0,033086	0,052	3.440	2015–2025	0 mês
Spearman	0,163475	<0,001	3.440	2015–2025	0 mês

Figura 11. Resultados dos testes de correlação entre focos de queimadas e internações respiratórias.

O coeficiente de Pearson apresentou valor próximo de zero (-0,033086; $p=0,052$), indicando ausência de correlação linear relevante ao nível usual de 5% de significância. Já o coeficiente de Spearman apresentou valor positivo, porém fraco (0,163475; $p<0,001$), sugerindo uma associação monotônica estatisticamente significativa, mas de baixa intensidade. Esses resultados indicam que a relação entre focos de queimadas e internações respiratórias não ocorre de forma direta e homogênea quando observada no conjunto nacional de dados mensais, podendo ser influenciada por fatores como população exposta, estrutura da rede hospitalar, distribuição regional dos focos de queimadas, sazonalidade climática, critérios de registro e vulnerabilidades locais.

Dessa forma, os testes de correlação reforçam a necessidade de interpretar os resultados em conjunto com as análises regionais e com a validação das fontes oficiais. Embora os coeficientes não indiquem associação forte no agregado nacional, eles não descartam a existência de relações específicas em determinados estados, biomas, períodos de seca ou grupos populacionais mais vulneráveis.

6. Discussão

Os resultados demonstram que a integração de bases ambientais, hospitalares e epidemiológicas amplia a capacidade de análise sobre a relação entre queimadas e saúde respiratória, mas exige cautela metodológica. A validação com fontes oficiais indicou aderência total dos dados de queimadas ao boletim Infoqueima/INPE no recorte de Mato Grosso em junho de 2022, enquanto os dados de internações apresentaram diferença entre o painel e o TABNET/DATASUS em 2024. Essa diferença reforça a necessidade de registrar explicitamente os filtros, períodos de extração, agrupamentos por CID e critérios de atualização adotados no Data Mart.

A agregação anual e estadual, embora necessária para compatibilizar bases com granularidades distintas, limita a interpretação dos resultados. Esse nível de agregação pode ocultar variações mensais, sazonalidades, episódios críticos de fumaça, diferenças entre municípios e deslocamentos temporais entre exposição ambiental e internação hospitalar. Além disso, Unidades Federativas mais populosas tendem a concentrar maior volume absoluto de internações, o que pode reduzir a comparabilidade direta com estados que possuem grande número de focos de queimadas, mas menor população ou menor disponibilidade de serviços hospitalares.

Os testes de correlação reforçam essa leitura. O coeficiente de Pearson não indicou associação linear relevante no conjunto mensal agregado, enquanto o coeficiente de Spearman apontou associação monotônica fraca, ainda que estatisticamente significativa. Portanto, os resultados não devem ser interpretados como ausência de efeito das queimadas sobre a saúde respiratória, mas como evidência de que a relação é complexa e depende de recortes espaciais, temporais e populacionais mais específicos. Estudos futuros podem aprofundar essa análise com defasagens temporais, normalização por população, taxas de

internação, análise municipal e estratificação por bioma ou período seco.

Os resultados obtidos apresentam convergência com estudos anteriores ao demonstrar a relevância da integração de bases públicas para ampliar a confiabilidade das análises em saúde e meio ambiente. Em consonância com [Albuquerque et al. 2023], verificou-se que a pandemia provocou descontinuidades importantes na série histórica de internações respiratórias, exigindo cautela na interpretação dos indicadores hospitalares. Da mesma forma, os resultados obtidos são consistentes com [Schroeder 2022], ao evidenciar que a relação entre queimadas e agravos respiratórios é influenciada por múltiplos fatores ambientais, demográficos e epidemiológicos. Nesse contexto, os testes de correlação realizados neste estudo indicaram uma associação monotônica positiva de baixa intensidade no conjunto de dados analisado, enquanto as visualizações desenvolvidas no painel permitiram complementar essa análise por meio da exploração dos indicadores sob diferentes recortes espaciais e temporais. Além disso, a abordagem baseada em Engenharia de Dados amplia as contribuições da literatura ao integrar, padronizar, validar e disponibilizar diferentes bases públicas em um único ambiente analítico, favorecendo análises mais transparentes, reprodutíveis e passíveis de expansão em estudos futuros.

7. Conclusão

O presente trabalho desenvolveu uma abordagem integrada para análise de dados sobre focos de queimadas e agravos respiratórios no Brasil, no período de 2015 a 2025, a partir da consolidação de bases públicas ambientais, hospitalares e epidemiológicas. A construção do ambiente analítico permitiu organizar dados provenientes do BDQueimadas/INPE, do SIH/SUS via TABNET/DATASUS e do SIVEP-Gripe/OpenDataSUS em uma estrutura orientada à análise, com apoio de processos de ETL, modelo dimensional e visualizações em Power BI.

Os resultados obtidos demonstram a importância da padronização e da validação metodológica em estudos que combinam bases de naturezas distintas. No caso dos focos de queimadas, a conferência com o boletim Infoqueima/INPE indicou compatibilidade entre os totais apresentados no painel e os valores publicados em boletins oficiais, reforçando a consistência da base ambiental utilizada. Em relação aos dados de internações respiratórias, a comparação com o TABNET/DATASUS apresentou diferenças residuais, consideradas compatíveis com possíveis variações de filtros, atualizações, agrupamentos por CID/lista de morbidade e recortes temporais.

A análise também evidenciou que os registros do SIVEP-Gripe não devem ser interpretados como substitutos diretos das internações hospitalares. As divergências observadas entre o painel e os dados unificados do Ministério da Saúde, especialmente em recortes por Unidade Federativa e ano, indicam que esse sistema deve ser utilizado como fonte de vigilância epidemiológica de SRAG, e não como medida equivalente à produção hospitalar registrada no SIH/SUS. Essa distinção é fundamental para evitar interpretações inadequadas sobre o período pandêmico e para preservar a coerência metodológica do estudo.

As visualizações desenvolvidas no Power BI permitiram identificar padrões relevantes na distribuição territorial e temporal dos focos de queimadas, com destaque para biomas e Unidades Federativas com maior concentração de registros. Também foi

possível observar que a relação entre queimadas e internações respiratórias não ocorre de forma simples ou homogênea entre os estados, uma vez que depende de fatores como população exposta, estrutura da rede hospitalar, critérios de registro e diferenças regionais de vulnerabilidade.

Em relação às questões de pesquisa propostas, os resultados obtidos permitiram responder aos objetivos estabelecidos. A QP1 foi atendida por meio da implementação de um processo estruturado de Engenharia de Dados, contemplando a integração, padronização, validação e consolidação das bases SIH/SUS, SIVEP-Gripe e BDQueimadas em um Data Mart orientado à análise. A QP2 evidenciou que a pandemia de COVID-19 provocou uma descontinuidade relevante na série histórica de internações por doenças respiratórias, tornando necessária a utilização dos registros do SIVEP-Gripe como fonte complementar para contextualizar esse comportamento, sem substituir os indicadores hospitalares do SIH/SUS. Por sua vez, a QP3 demonstrou que a associação entre focos de queimadas e internações respiratórias apresentou correlação monotônica positiva de baixa intensidade, estatisticamente significativa, indicando que essa relação é influenciada por múltiplos fatores e deve ser interpretada de forma integrada às características ambientais, demográficas e epidemiológicas do período analisado.

Conclui-se que a integração entre Engenharia de Dados, bases públicas oficiais e ferramentas de análise visual contribui para ampliar a transparência e a qualidade da investigação sobre os impactos das queimadas na saúde respiratória. Ao explicitar as limitações dos sistemas de informação e diferenciar indicadores hospitalares de indicadores de vigilância epidemiológica, o estudo oferece uma base mais consistente para análises futuras, tomada de decisão em saúde pública e fortalecimento das ações de vigilância ambiental e epidemiológica.

Referências

- Albuquerque, D. d. A. R. d., Melo, M. D. T. d., and Sousa, T. L. F. e. a. (2023). Internações hospitalares e taxas de mortalidade por doenças respiratórias não covid-19 no sistema público de saúde do brasil durante a pandemia de covid-19: um estudo observacional nacional. *Jornal Brasileiro de Pneumologia*.
- Brasil. Ministério da Saúde (2022). Boletins epidemiológicos covid-19: 2022. Ministério da Saúde. Acesso em: 2 jul. 2026.
- Brasil. Ministério da Saúde (2026). Banco de dados da síndrome respiratória aguda grave (srag) – 2019 a 2026. Portal de Dados Abertos do SUS. Acesso em: 2 jul. 2026.
- Brasil. Ministério da Saúde. DATASUS (2026a). Datasus: Departamento de informática do sus. Ministério da Saúde. Acesso em: 3 jul. 2026.
- Brasil. Ministério da Saúde. DATASUS (2026b). Sistema de informações hospitalares do sus (sih/sus): morbidade hospitalar do sus por local de internação. TAB-NET/DATASUS. Acesso em: 2 jul. 2026.
- Caumo, S., Gioda, A., da Silva Carreira, R., Jacobson, L. d. S. V., and Hacon, S. d. S. (2022). Impactos na saúde humana causados pela exposição a incêndios florestais: as evidências obtidas nas últimas duas décadas. *Revista Brasileira de Climatologia*, 30(18):182–218.
- da Silva, M. H. A. (2023). Os erros das notificações e as subnotificações de casos de covid-19: Revisão integrativa. *Revista Saber Digital*, 16(3):e20231605.
- de Souza, M. L., Ichihara, M. Y. T., and Sena, S. O. L. (2020). *Sistemas de informação para a COVID-19*, volume 2. EDUFBA, Salvador.
- Filho, E. B. d. S., Silva, A. L., Santos, A. O., Dall’acqua, D. S. V., and Souza, L. F. B. (2017). Infecções respiratórias de importância clínica: uma revisão sistemática: Respiratory infections of clinical importance: A systematic review. *REVISTA FIMCA*, 4(1):7–16.
- Godoy, R. L. d. (2023). *Relação entre poluentes do ar e doenças respiratórias: análise por agrupamento de dados espaciais e regras de associação temporais*. Tese (doutorado), Universidade Estadual de Campinas (UNICAMP), Limeira, SP.
- Gracie, R., Muzy, J., Xavier, D., Rulli, J., Moraes, T., Peixoto, J., Oliveira, B., Lopes, R., Rodrigues, D. d. O., and Rocha, V. (2024). *Curso Análise de Situação de Saúde Ambiental: ASISA-Queimadas*. Edições Livres, Rio de Janeiro.
- Instituto Nacional de Pesquisas Espaciais (2022). Infoqueima: Boletim técnico mensal de monitoramento de focos de fogo ativo, junho de 2022. Programa Queimadas do INPE. Acesso em: 2 jul. 2026.
- Instituto Nacional de Pesquisas Espaciais (2026). Portal do monitoramento de queimadas e incêndios florestais. TerraBrasilis – Programa Queimadas/INPE. Acesso em: 3 jul. 2026.
- Kimball, R. and Ross, M. (1996). *The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling*. John Wiley & Sons, New York, NY.

- Oliveira, F., Santos, D., Bohm, B., and Santana, B. (2025). Framework integrado para análise de dados climáticos e epidemiológicos: Potencial para monitoramento. In *Anais do XVI Workshop de Computação Aplicada à Gestão do Meio Ambiente e Recursos Naturais (WCAMA)*, pages 340–343, Porto Alegre, RS, Brasil. SBC.
- Schroeder, L. (2022). Associação das queimadas com doenças respiratórias e complicações da covid-19 no estado do pará, brasil. Dissertação (mestrado), Universidade do Vale do Rio dos Sinos (UNISINOS), São Leopoldo, RS.
- Sena, P. S. d. (2024). Incêndios culturais: queimadas, biodiversidade e saúde ambiental na região metropolitana do vale do paraíba e litoral norte de são paulo. *Revista Ângulo*, 1(168).
- Sousa, M. V. d. C. (2020). Data warehouse e etl para dados da saúde pública: uma aplicação prática. Monografia (Bacharelado em Ciência da Computação) – Instituto de Matemática e Estatística, Universidade de São Paulo. Trabalho de Formatura Supervisionado – MAC499.

Anexo A — Evidências Visuais de Validação

Morbidade Hospitalar do SUS - por local de internação - Brasil	
Internações por Ano atendimento e Unidade da Federação	
Capítulo CID-10: X. Doenças do aparelho respiratório	
Período: 2024-2025	
Ano atendimento	Total
2024	1.363.903
Fonte: Ministério da Saúde - Sistema de Informações Hospitalares do SUS (SIH/SUS)	

Figura A.1. Consulta ao TABNET/DATASUS para internações por doenças do aparelho respiratório em 2024.

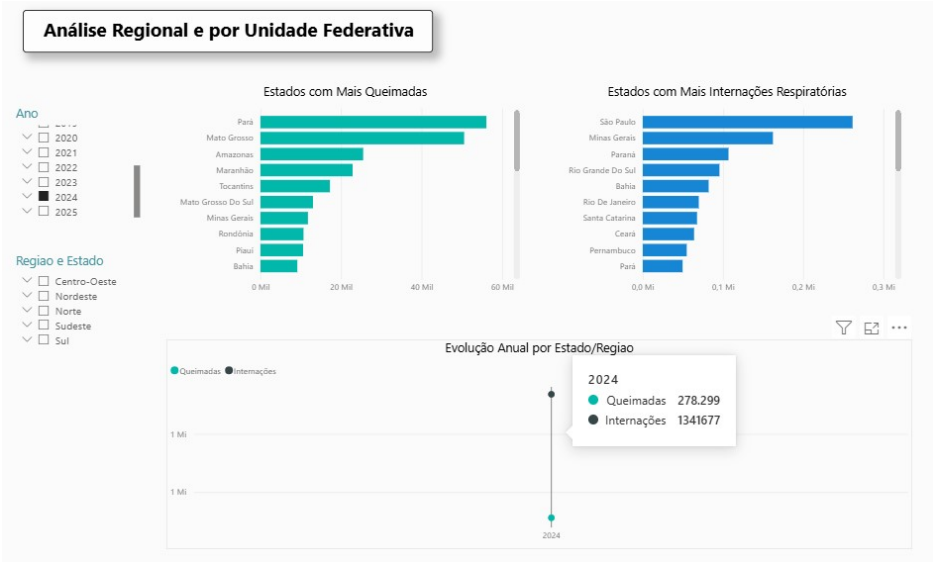


Figura A.2. Valor de internações respiratórias em 2024 consolidado no painel Power BI.

INFOQUEIMA - Boletim Técnico Mensal de Monitoramento de Focos de Fogo Ativo			Junho 2022
A Tabela 1.1 ilustra a comparação percentual da quantidade de focos de fogo ativo em relação ao período de 01/junho até 30/junho do ano anterior. É possível verificar que 19 estados apresentaram menor quantidade de focos comparado ao ano anterior, enquanto outros 7 estados apresentaram um aumento na quantidade de focos, no qual podem ser vistos na Tabela 1.1 abaixo.			
Tabela 1.1: Quantidade de focos de fogo ativo por estado em junho/2022 em comparação com o mesmo período de 2021, segundo o satélite de referência			
Estado	Focos em 2022	Focos em 2021	%
Mato Grosso	2210	2185	1
Tocantins	1370	1149	19
Maranhão	917	979	-6
Minas Gerais	665	523	27

Figura A.3. Registro de focos de queimadas no boletim Infoqueima/INPE de junho de 2022.

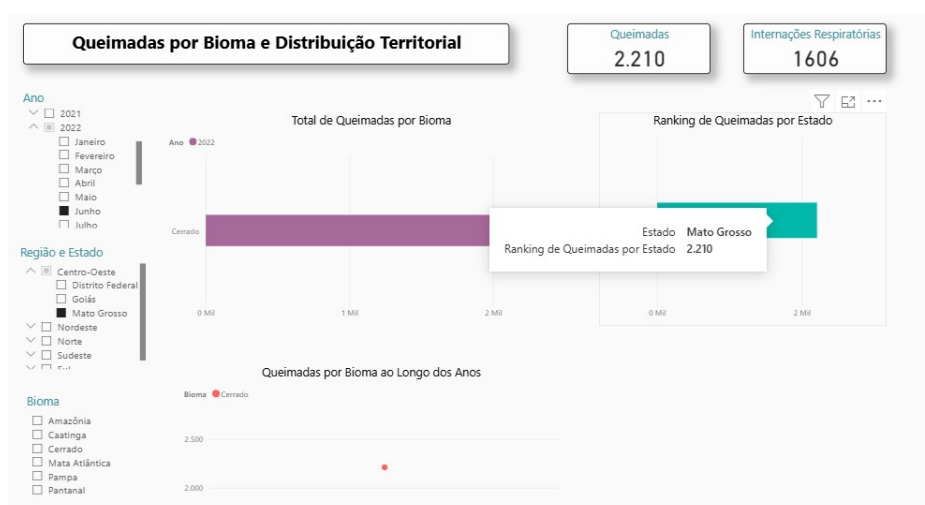


Figura A.4. Valor de focos de queimadas em Mato Grosso, junho de 2022, consolidado no painel Power BI.

```
(venv) PS C:\Users\andre\Documents\sistemas de informação\projetos\healthfire> venv\Scripts\py
Correlation results saved to: correlation_tests\results\correlation_results.csv
Aligned monthly data saved to: correlation_tests\results\aligned_monthly_data.csv
method coefficient sample_size start_year end_year lag_months
pearson -0.033086 3440 2015 2025 0
spearman 0.163475 3440 2015 2025 0
(venv) PS C:\Users\andre\Documents\sistemas de informação\projetos\healthfire>
```

Figura A.5. Resultado dos testes de correlação de Pearson e Spearman executados no ambiente de análise.