

Mineração de Dados para Auxiliar na Tomada de Decisões Estratégicas em Teleatendimentos do CISAM-UPE

Data Mining to Support Strategic Decision-Making in Telecare at CISAM/UPE

Daniel Germinio Aguiar¹
orcid.org/0009-0004-6584-0059

Thompson Henrique dos Santos Macedo Cysneiros³
orcid.org/0009-0006-8655-4398

Claudinalle Farias Queiroz de Souza²
orcid.org/0000-0003-1541-3089

Hemir da Cunha Santiago⁴
orcid.org/0000-0001-6307-067X

¹Faculdade Senac, Recife, Brasil. E-mail: daniel.aguiar@edu.pe.senac.br

²Faculdade de Enfermagem Nossa Senhora Das Graças, Universidade de Pernambuco, Recife, Brasil. E-mail: claudinalle.souza@upe.br

³Faculdade Nova Roma, Recife, Brasil. E-mail: thompson.cysneiros@novaroma.edu.br

⁴Escola Politécnica de Pernambuco, Universidade de Pernambuco, Recife, Brasil. E-mail: hemir.santiago@upe.br

DOI: [10.25286/repa.v11i1.3525](https://doi.org/10.25286/repa.v11i1.3525)

Esta obra apresenta Licença Creative Commons Atribuição-Não Comercial 4.0 Internacional.

Como citar este artigo pela NBR 6023/2018: Daniel Germinio Aguiar; Claudinalle Farias Queiroz de Souza; Thompson Henrique dos Santos Macedo Cysneiros; Hemir da Cunha Santiago. Mineração de Dados para Auxiliar na Tomada de Decisões Estratégicas em Teleatendimentos do CISAM-UPE. Revista de Engenharia e Pesquisa Aplicada, v.11, n. 1, p. 11-18, 2026.

RESUMO

Instituições de saúde pública enfrentam desafios contínuos na gestão da demanda e triagem. Em um sistema de saúde cada vez mais sobrecarregado e com o constante avanço das tecnologias de atendimento, a mineração de dados pode ser um forte aliado na saúde da mulher. Este estudo de caso aplicado no Centro Universitário Integrado de Saúde Amaury de Medeiros (CISAM-UPE), via Núcleo de Telessaúde (NUTES), utiliza a metodologia CRISP-DM para identificar padrões de atendimento. A análise abrangeu registros de janeiro de 2023 a abril de 2024 nas especialidades de Ginecologia Geral, Laqueadura e Histeroscopia, utilizando dashboards em Microsoft Excel e Power BI. Os resultados revelaram concentração de solicitações nas manhãs de segundas, terças e sextas-feiras, além da predominância de queixas de miomas uterinos e sangramento anormal. A aplicação do CRISP-DM permitiu a transformação de dados operacionais em informações açãoáveis, contribuindo para a otimização do agendamento médico.

PALAVRAS-CHAVE: Análise de dados; Plataforma de software; Instituição hospitalar; Teleatendimento.

ABSTRACT

Public health institutions face ongoing challenges in managing demand and triage. In an increasingly overburdened healthcare system and with the constant advancement of care technologies, data mining can be a powerful ally in women's health. This case study, applied at the Amaury de Medeiros Integrated University Health Center (CISAM-UPE) via the Telehealth Center (NUTES), uses the CRISP-DM methodology to identify patterns of care. The analysis covered records from January 2023 to April 2024 in the specialties of General Gynecology, Tubal Ligation, and Hysteroscopy, using dashboards in Microsoft Excel and Power BI. The results revealed a concentration of requests on Monday, Tuesday, and Friday mornings, as well as a predominance of complaints of uterine fibroids and abnormal bleeding. The application of CRISP-DM enabled the transformation of operational data into actionable information, contributing to the optimization of medical scheduling.

KEY-WORDS: Data analysis; Software platform; Healthcare institution; Tele Customer service.

1 INTRODUÇÃO

O Centro Universitário Integrado de Saúde Amaury de Medeiros (CISAM/UPE) tem sido descrito como referência no atendimento à saúde da mulher no estado de Pernambuco, sendo líder em atendimento à distância com o Núcleo de Telessaúde (NUTES) incluído em sua estrutura institucional [1]. As iniciativas de telessaúde desenvolvidas pelo NUTES contribuem para a ampliação do alcance das ações do CISAM, ao viabilizar o acesso remoto de profissionais de saúde a consultas, orientações técnicas e processos de capacitação em diferentes regiões do estado.

A incorporação da ciência de dados nas operações de núcleos de telessaúde representa uma oportunidade para aprimorar a compreensão dos fluxos de atendimento e apoiar a gestão dos serviços [2]. O uso de análises descritivas e exploratórias possibilita identificar padrões de demanda, apoiar o planejamento da alocação de recursos e evidenciar gargalos nos processos assistenciais [3]. Essas análises são particularmente relevantes diante dos recorrentes tempos de espera, frequentemente associados à fragmentação e à baixa integração dos sistemas de informação em saúde [4]. Dessa maneira, a análise sistemática de dados pode contribuir para a geração de informações açãoáveis, auxiliando a tomada de decisão gerencial e favorecendo uma utilização mais eficiente dos recursos disponíveis.

O impacto da telessaúde também contribui para o fortalecimento da formação dos profissionais de saúde [5], especialmente em áreas mais remotas. Por meio de programas de educação continuada e treinamentos realizados pelo NUTES-CISAM, os profissionais podem atualizar seus conhecimentos e habilidades, melhorando a qualidade do atendimento prestado. Isso é particularmente relevante em Pernambuco, onde a distribuição geográfica e as desigualdades socioeconômicas podem dificultar o acesso a treinamentos presenciais. A telessaúde, portanto, não apenas facilita o acesso imediato a cuidados especializados, mas também é essencial para garantir que os profissionais de saúde estejam adequadamente preparados para atender às necessidades da população brasileira, principalmente após a necessidade de diversificação das estratégias de acesso aos serviços e o crescimento da telessaúde durante a pandemia de COVID-19 [6, 7, 8].

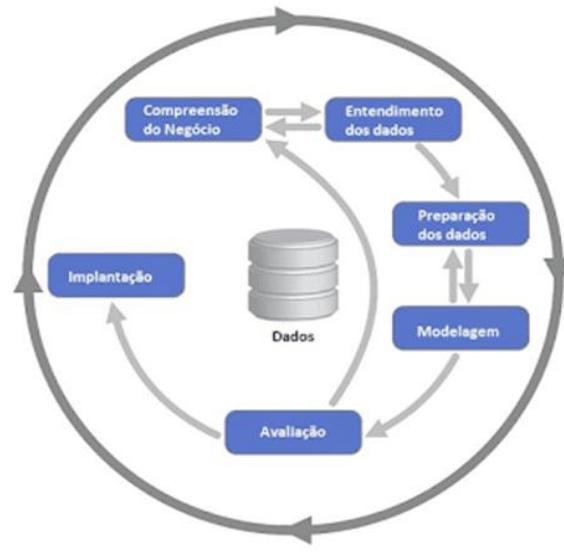
Este estudo de caso teve como objetivo examinar como a aplicação de técnicas de ciência de dados

pode melhorar a eficácia do atendimento, focando diretamente nos desafios relacionados ao tempo de espera e à eficiência dos serviços de saúde, avaliando o impacto da telessaúde na redução das desigualdades no acesso aos serviços e na otimização do atendimento.

2 METODOLOGIA

A metodologia CRISP-DM (*Cross Industry Standard Process for Data Mining*) promove boas práticas e fornece às organizações a estrutura necessária para obter resultados mais rápidos e eficazes por meio da mineração de dados [9]. Desenvolvida em 1996, foi originalmente voltada para organizações industriais e comerciais que utilizavam a mineração de dados em suas operações de negócios. A metodologia CRISP-DM é composta por 6 fases: Compreensão do Negócio, Entendimento dos Dados, Preparação dos Dados, Modelagem, Avaliação e Entrega. Essas fases auxiliam as organizações a compreenderem o processo de mineração de dados, fornecendo um guia para o planejamento e a execução de um projeto. A Figura 1 ilustra as fases da metodologia. As setas 'internas' destacam as dependências mais importantes e frequentes entre as fases [9].

Figura 1 – As seis fases da metodologia CRISP-DM.



Fonte: [9].

2.1 Compreensão do Negócio

A fase de Compreensão do Negócio iniciou-se com o entendimento do contexto clínico e administrativo

do CISAM, com ênfase nos desafios relacionados ao atendimento especializado à saúde da mulher, especialmente nas áreas de maior complexidade, como ginecologia e obstetrícia. Para alinhar as necessidades organizacionais do CISAM, foram utilizadas técnicas de coleta de requisitos, como entrevistas e reuniões, com o objetivo de compreender como a gestão de agendamentos era realizada, em particular por meio de formulários eletrônicos de estrutura simplificada. Além de atender às demandas internas do CISAM, este trabalho teve como objetivo integrar os dados do NUTES, responsável pela gestão dos agendamentos de consultas, exames e procedimentos cirúrgicos das pacientes atendidas no CISAM. Esses agendamentos envolvem diferentes prioridades e tipos de atendimento, o que aumenta a complexidade da análise e reforça a necessidade de apoio analítico para gestão das filas de espera e à tomada de decisões. Por meio da aplicação de técnicas de mineração de dados, buscou-se obter informações relevantes anteriormente não explicitadas nos registros operacionais, como padrões sazonais de solicitações, áreas com maior concentração de demanda e procedimentos mais frequentemente solicitados. O objetivo final foi definido como a criação e utilização de dashboards integrando os dados obtidos, com a finalidade de apoiar a visualização e a interpretação das informações, proporcionando subsídios para a análise dos padrões de atendimento, dos perfis das pacientes e de possíveis tendências de demanda.

2.2 Entendimento dos Dados

A fase de Entendimento dos Dados iniciou-se com a coleta dos registros de teleagendamentos fornecidos pelo NUTES, abrangendo os procedimentos de Histeroscopia Cirúrgica, Cirurgia Geral em Ginecologia e Laqueadura Tubária. Os dados brutos, disponibilizados em formato de planilhas eletrônicas, continham variáveis demográficas, clínicas e administrativas. Nesta etapa, foram aplicadas técnicas de análise exploratória inicial para verificar a estrutura dos atributos e a distribuição preliminar dos registros. O foco metodológico consistiu no mapeamento dos tipos de dados e na identificação de atributos de qualidade, verificando a existência de valores nulos, erros de formatação e inconsistências de preenchimento que pudessem impactar a modelagem subsequente.

2.3 Preparação dos Dados

A fase de Preparação dos Dados consistiu na aplicação de técnicas de tratamento para converter a base bruta em um formato analítico, seguindo as diretrizes do CRISP-DM. O processo envolveu quatro subetapas:

Seleção: Foram aplicados filtros para manter apenas as variáveis com preenchimento suficiente e relevância clínica, descartando colunas com excesso de dados faltantes (acima de um limiar de corte predefinido).

Limpeza: Utilizaram-se scripts de tratamento no Power BI e Excel para remoção de caracteres espúrios, padronização de textos e tratamento de valores nulos.

Construção: Novas variáveis foram derivadas, como o cálculo da idade a partir da data de nascimento e a segmentação de horários em turnos (manhã/tarde/noite) a partir do carimbo de tempo.

Integração: As bases de dados distintas das três especialidades foram unificadas em um único *dataset* consolidado para permitir a análise transversal.

2.4 Análise e Modelagem dos Dados

A fase de Análise e Modelagem envolveu a aplicação de técnicas e ferramentas de ciência de dados para identificar padrões nos agendamentos do CISAM. O objetivo foi desenvolver modelos preditivos e de segmentação que otimizam o fluxo de atendimento e aprimoram a gestão dos agendamentos. Conforme o relatório, as ferramentas *Excel* e *Power BI* foram escolhidas devido à sua robustez na manipulação de dados e à capacidade de gerar visualizações interativas e detalhadas, essenciais para a análise dos padrões identificados. O *Power BI* permitiu a criação de dashboards intuitivos, que facilitaram a segmentação dos dados em diferentes visões: "Geral", "por Faixa Etária", "por Localidade" e "por Periodicidade". Essa abordagem permitiu uma análise abrangente dos agendamentos e suas características.

A análise abrangeu diversas dimensões, permitindo identificar perfis de pacientes e

tendências sazonais nos agendamentos. O uso de gráficos e visualizações, como gráficos de Pizza, Rosquinha e Linha Temporal, facilitou a visualização do fluxo de atendimentos e ajudou a identificar períodos com maior demanda.

2.5 Avaliação dos Resultados

A fase de Avaliação dos Resultados consistiu na validação técnica e de negócio dos modelos gerados. O processo foi estruturado em três níveis de verificação para garantir que os padrões identificados correspondem à realidade operacional do CISAM:

Validação Técnica: Foram aplicadas métricas comparativas entre a base bruta e a base tratada para mensurar o ganho de completude e consistência. Verificou-se a redução percentual de campos nulos e a correção de tipagens de dados (datas e horários) para assegurar a confiabilidade estatística das visualizações.

Validação de Negócio: Os dashboards preliminares foram submetidos à revisão dos gestores do NUTES e do CISAM. Esta etapa utilizou filtros dinâmicos de segmentação (faixa etária, localidade e periodicidade) para testar hipóteses e descobrir novos padrões, confirmando se as ferramentas atendiam aos requisitos de suporte à decisão definidos na fase de Compreensão do Negócio.

Verificação de Usabilidade: Testou-se a interatividade das visualizações, com *slicers*, para garantir que os usuários finais pudessem extrair insights autônomos sobre a distribuição de demanda e alocação de recursos sem necessidade de intervenção técnica constante.

2.6 Implantação

A implantação dos *dashboards* no CISAM foi planejada em duas etapas. A primeira consiste em implementar uma ferramenta de gestão que apoie a tomada de decisões, melhore a administração de recursos e otimize o atendimento prestado pelo serviço. Para isso, serão utilizados o método de tratamento de dados desenvolvido na Fase 3 e o método de modelagem de dados desenvolvido na Fase 4, ambos agora adaptados para as planilhas do *Google*, facilitando a atualização do banco de dados. Em relação às métricas, a estrutura de modelagem será mantida nos *dashboards*

interativos, organizados em visões específicas previamente definidas.

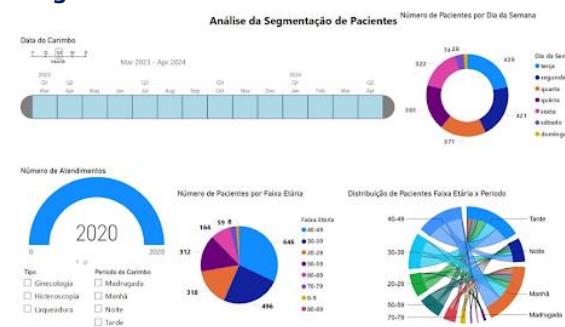
A segunda etapa tem como objetivo promover a transparência, oferecendo ao público um *dashboard* acessível, alimentado semestralmente com dados consolidados e estáticos. Esse *dashboard* apresentará métricas e indicadores sobre o desempenho do atendimento no CISAM, como a média mensal de atendimentos, um mapa de calor mostrando as regiões mais atendidas, um gráfico de barras com as cidades que mais recebem atendimentos e a média mensal de serviços mais solicitados. A intenção é fornecer uma visão clara e segura das operações do CISAM, destacando sua atuação abrangente em diversas cidades, regiões e especialidades, sem comprometer a privacidade dos dados.

3 RESULTADOS

A análise dos dados referentes às requisições de atendimento nas especialidades de Ginecologia Geral, Laqueadura Tubária e Histeroscopia (janeiro de 2023 a abril de 2024) foi estruturada em quatro eixos analíticos: Visão Geral, Faixa Etária, Localidade e Periodicidade. As ferramentas Microsoft Excel e Power BI foram utilizadas para o processamento de grandes volumes de dados, permitindo a identificação de padrões de sazonalidade e perfil epidemiológico.

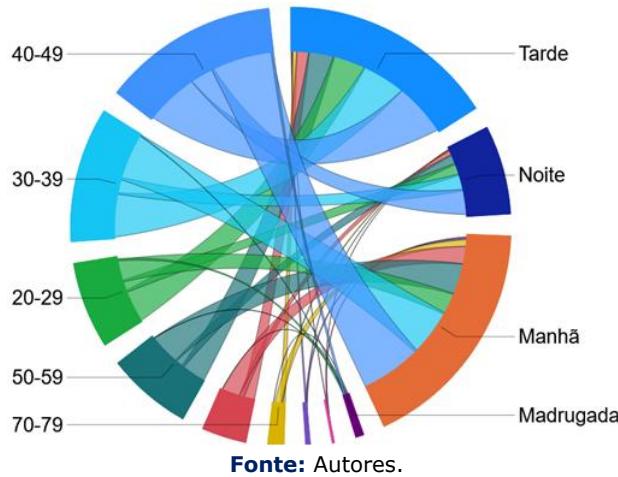
Esses benefícios ajudam a transmitir a visão do analista de forma clara e eficaz, permitindo uma melhor tomada de decisões. A Figura 2 mostra um exemplo de uso de *slicers* usados dentro do *dashboard* de análise, com eles é possível o próprio usuário final personalizar os gráficos de um modo interativo e dinâmico.

Figura 2 – Dashboard de Análise de Faixa Etária.



pacientes com idade média de 41 anos nas solicitações de Ginecologia Geral.

Figura 3 – Gráfico de Corda distribuição de pacientes x período x faixa etária.



A análise geoespacial (Figuras 4 e 5) indicou que a demanda não é uniforme, concentrando-se significativamente na Região Metropolitana do Recife, com destaque para bairros como Água Fria. Essa visualização granular permite à gestão do CISAM direcionar campanhas de saúde e recursos preventivos para as zonas de maior pressão assistencial, saindo de uma abordagem generalista para uma gestão baseada em evidências locais.

Figura 4 – Dashboard de Análise de Localidade.

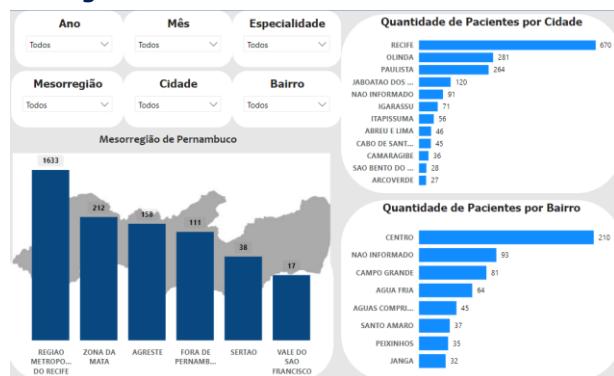
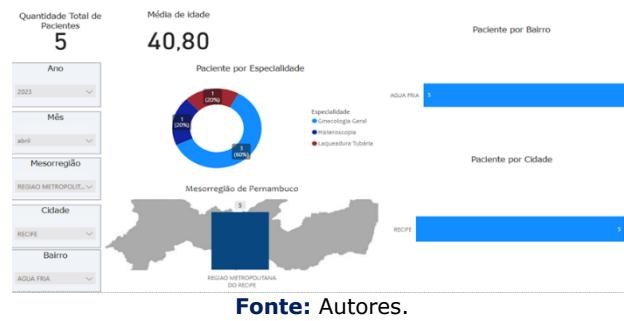
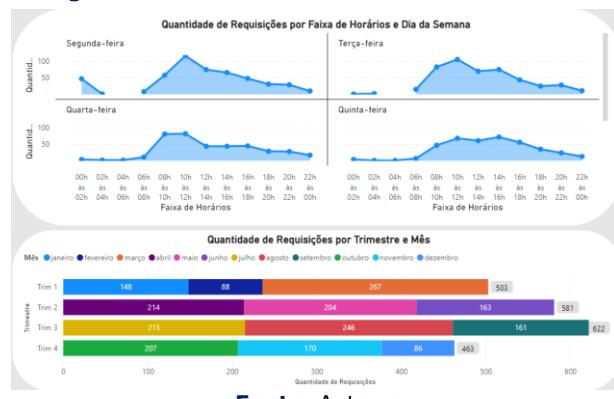


Figura 5 – Distribuição de pacientes por cidade, bairro e mesorregião no ano de 2023.



A análise temporal, Figura 6, evidenciou gargalos operacionais críticos. Identificou-se uma concentração de requisições no período matutino, com picos acentuados às segundas, terças e sextas-feiras. Ao saber que a demanda real apresenta sazonalidade semanal a alocação de plantões médicos fica mais clara. A identificação e segmentação de variações sazonais e geográficas é fundamental para a alocação de recursos e a gestão eficiente dos serviços de saúde [10].

Figura 6 – Dashboard de Análise de Periodicidade.



A mineração de texto aplicada aos campos de “motivo da consulta” resultou na geração de uma nuvem de palavras (ver Figura 7), evidenciando a frequência relativa das principais queixas clínicas. Termos como “sangramento”, “mioma” e “cirurgia” apresentaram alta recorrência, indicando a predominância de demandas associadas a condições que usualmente requerem maior atenção clínica na triagem. Em contraste com análises baseadas exclusivamente no volume de atendimentos, essa abordagem exploratória baseada em texto permitiu incorporar elementos qualitativos dos relatos dos pacientes, contribuindo para uma compreensão mais ampla da natureza da demanda reprimida. Esses achados reforçam a relevância de estratégias de triagem que considerem, além da ordem de chegada, a presença de terminologia clínica sensível nos registros iniciais.

Figura 7 – Gráfico de Word Cloud.



Fonte: Autores.

4 CONCLUSÕES

A aplicação da metodologia CRISP-DM possibilitou a integração e a organização dos dados provenientes dos prontuários de Histeroscopia Cirúrgica, Cirurgia Ginecológica e Laqueadura Tubária em um modelo único, viabilizando a análise sistemática da demanda assistencial. As análises realizadas evidenciaram padrões consistentes de periodicidade para as especialidades estudadas, com maior concentração de atendimentos no período da manhã e em dias úteis, especialmente às segundas, terças e sextas-feiras, além de variações mensais que podem subsidiar o planejamento operacional ao longo do ano. Os resultados obtidos indicam que a exploração estruturada dos dados permite ampliar a compreensão sobre o comportamento da demanda e apoiar o planejamento da alocação de recursos humanos e assistenciais. A utilização de ferramentas como Microsoft Excel e Power BI mostrou-se adequada para a condução de análises descritivas e para a construção de visualizações interativas, facilitando a interpretação dos padrões identificados e o suporte à tomada de decisão gerencial. Durante o processo analítico, foram identificadas limitações relacionadas à qualidade dos dados, como valores ausentes e inconsistências de formatação, ressaltando a importância do aprimoramento contínuo dos processos de registro e da incorporação de bases históricas mais extensas e de múltiplas fontes. A segmentação das análises por motivos de consulta, faixa etária, localidade e periodicidade contribuiu para uma visão mais abrangente da demanda, destacando a relevância de abordagens analíticas integradas para a gestão de serviços de saúde. Por fim, o estudo demonstra que a aplicação sistemática do CRISP-DM, pode

beneficiar sistemas de atendimento hospitalar por meio da tomada de decisões baseada em dados.

6 AGRADECIMENTOS

Os autores agradecem à Fundação de Amparo à Ciência e Tecnologia do Estado de Pernambuco (FACEPE) pelo apoio financeiro concedido para a realização deste trabalho, por meio do processo ARC-0441-1.03/23, vinculado ao Edital 05/2023 - Programa de Extensão Tecnológica 2023, para desenvolvimento do projeto: "Ciência de Dados e Inteligência Artificial para Auxiliar na Tomada de Decisões em Agendamentos de Atendimentos no CISAM" no período de 01/02/2024 a 30/06/2024. Este financiamento foi essencial para o desenvolvimento e a concretização desta pesquisa.

Os autores também agradecem ao NUTES-CISAM pelo provimento dos dados de teleatendimento, assim como pelo apoio técnico ao longo da execução de experimentos e análises fundamentais para os resultados obtidos.

REFERÊNCIAS

- [1] UNIVERSIDADE DE PERNAMBUCO – UPE.** *Cisam/UPE, referência no atendimento à mulher em Pernambuco, recebe a visita da secretária estadual de Ciência, Tecnologia e Inovação.* UPE, 5 jun. [s.d.]. Disponível em: <https://www.upe.br/noticias/cisam-upe,-refer%C3%A3ncia-no-atendimento-%C3%A0-mulher-em-pernambuco,-recebe-a-visita-da-secret%C3%A1ria-estadual-de-ci%C3%A3ncia,-tecnologia-e-inova%C3%A7%C3%A3o.html>. Acesso em: 22 dez. 2025.

[2] GHEORGHE, Mihaela; PETRE, Ruxandra. Integrating data mining techniques into telemedicine systems. **Informatica Economica**, v. 18, n. 1, p. 120, 2014.

[3] FUGULIN, Fernanda Maria Togeiro; GAIDZINSKI, Raquel Rapone; KURCGANT, Paulina. Sistema de classificação de pacientes: identificação do perfil assistencial dos pacientes das unidades de internação do HU-USP. **Revista Latino-Americana de Enfermagem**, v. 13, p. 72-78, 2005.

[4] FOLLEN, Marilyn et al. Implementing health information technology to improve the process of health care delivery: a case study. **Disease Management**, v. 10, n. 4, p. 208-215, 2007.

- [5] NOVAES, Magdala de Araújo et al. Tele-educação para educação continuada das equipes de saúde da família em saúde mental: a experiência de Pernambuco, Brasil. **Interface-Comunicação, Saúde, Educação**, v. 16, p. 1095-1106, 2012.
- [6] SARTI, Thiago Dias; ALMEIDA, Ana Paula Santana Coelho. Incorporação de telessaúde na atenção primária à saúde no Brasil e fatores associados. *Cadernos de saude publica*, v. 38, p. PT252221, 2022.
- [7] OLIVEIRA, Daniela et al. Prediction of covid-19 diagnosis based on openehr artefacts. **Scientific Reports**, v. 12, n. 1, p. 12549, 2022.
- [8] FREIRE, Mariana Prado et al. Telemedicine in healthcare access during the covid-19 pandemic: a scoping review. **Revista de Saúde Pública**, v. 57, p. 4s, 2023.
- [9] SHEARER, Colin. The CRISP-DM model: the new blueprint for data mining. **Journal of data warehousing**, v. 5, n. 4, p. 13-22, 2000.
- [10] RODRIGUES, Nádia Cristina Pinheiro et al. Padrões espaciais e temporais da mortalidade infantil e seus componentes no Rio de Janeiro. **Saúde em Debate**, v. 47, p. 531-545, 2023.