



# Acesso à Saúde Mental no Estado de São Paulo: Uma Análise Exploratória Baseada em Dados do DATASUS

Larissa Yumi Tinem, Leo Tsuchida Hoçoya

Faculdade de Computação e Informática (FCI)

Universidade Presbiteriana Mackenzie – São Paulo, SP – Brasil

Sistemas de Informação – Faculdade de Computação e Informática (FCI) –

Universidade Presbiteriana Mackenzie – São Paulo, SP – Brasil

{10409094, 10332699}@mackenzista.com.br

**Resumo.** Este trabalho apresenta uma análise exploratória do acesso à saúde mental no estado de São Paulo, utilizando dados do DATASUS, complementados por informações do IBGE. O estudo enfoca a cobertura de Centros de Atenção Psicossocial (CAPS), Atenção Básica de Saúde e indicadores socioeconômicos regionais. A partir de técnicas de classificação descritiva com Inteligência Artificial, como clustering, as regiões do estado serão categorizadas quanto ao déficit de atendimento, permitindo identificar padrões, desigualdades e lacunas. Os resultados fornecerão subsídios para o planejamento de políticas públicas mais eficazes e a priorização de recursos na atenção psicossocial.

**Palavras-chave:** Saúde Mental, Acesso à Saúde, CAPS, Classificação Descritiva, Inteligência Artificial, Clustering

**Abstract.** This study presents an exploratory analysis of access to mental health services in São Paulo state, using data from DATASUS, complemented by information from the IBGE. The study focuses on the coverage of Psychosocial Care Centers (CAPS), Primary Health Care, and regional socioeconomic indicators. Using descriptive classification techniques with Artificial Intelligence, such as clustering, regions of the state will be categorized according to their service coverage deficits, allowing the identification of patterns, inequalities, and gaps. The results will provide insights to support more effective public policy planning and resource allocation in psychosocial care.

**Keywords:** Mental Health, Access to Health, CAPS, Descriptive Classification, Artificial Intelligence, Clustering.

## 1. Introdução

Conforme o artigo 196 da Constituição Federal de 1988, ‘a saúde é direito de todos e dever do Estado’ (Brasil, 1988). Entretanto, a efetivação desse direito enfrenta alguns desafios, como barreiras geográficas, má distribuição das unidades de saúde e

precariedade dos serviços oferecidos, principalmente em regiões mais afastadas dos grandes centros urbanos como regiões rurais ou periféricas, dificultando o acesso dessa população à cuidados básicos e emergenciais (Rocha et al., 2021).

Já segundo Dimenstein et al. (2021), a desigualdade ao acesso à saúde mental persiste mesmo após a implementação da Reforma Psiquiátrica e Rede de Atenção Psicossocial (RAPS), limitando o acesso equitativo aos serviços especializados e concentrando os recursos em regiões urbanas mais centrais. Essa desigualdade regional evidencia a necessidade de análise detalhada dos padrões de cobertura e da identificação de lacunas no acesso a serviços de saúde mental em todo o país.

A Política Nacional de Saúde Mental, apoiada na lei nº 10.216/2001, busca consolidar um modelo de atenção à saúde mental aberto e de base comunitária, promovendo o convívio com a família e a comunidade em vez do isolamento institucional, como predominava nos antigos manicômios.

Os Centros de Atenção Psicossocial (CAPS) têm como objetivo acolher pessoas com transtornos mentais, bem como indivíduos com transtornos causados pelo uso de álcool e outras drogas, promovendo sua integração social e familiar, apoioando o desenvolvimento da autonomia e oferecendo atendimento médico e psicológico. A principal característica desses serviços é inseri-los em um ambiente social e cultural concreto, chamado de “território”, que corresponde ao espaço da cidade onde ocorre o cotidiano dos usuários e de suas famílias. Dessa forma, os CAPS representam a estratégia central da reforma psiquiátrica no Brasil.

Estudos recentes indicam que a cobertura dos CAPS e a integração com a Atenção Básica são fatores determinantes para reduzir internações psiquiátricas e ampliar o acesso a tratamento adequados. Miliauskas et al. (2019), ao analisarem as regiões metropolitanas do Rio de Janeiro e de São Paulo, observaram que a ampliação desses serviços está associada à diminuição de hospitalizações, evidenciando a importância de políticas públicas que fortaleçam a rede de atenção psicossocial.

Diante desse cenário, o presente trabalho tem como objetivo realizar uma análise territorial do acesso à saúde mental no estado de São Paulo, utilizando dados do DATASUS combinados com informações populacionais e estruturais do IBGE. A investigação busca identificar padrões regionais de internações psiquiátricas e disponibilidade de serviços comunitários, examinando possíveis desigualdades na distribuição e intensidade da oferta.

Para alcançar esses objetivos, será adotada uma abordagem que integra análise exploratória de dados (EDA), com foco na construção de indicadores padronizados, como taxa de internações e taxa de CAPS por 100 mil habitantes, e técnicas estatísticas que permitem compreender relações estruturais no território. Serão avaliadas correlações entre indicadores, mapeados outliers regionais e analisada a concentração de serviços em diferentes Diretorias Regionais de Saúde (DRS). Esses resultados fundamentam a etapa subsequente do estudo, que envolve a preparação dos dados para aplicação de métodos de aprendizado de máquina, particularmente a clusterização, permitindo identificar agrupamentos territoriais com características semelhantes de demanda e oferta em saúde mental.

Com essa abordagem, o trabalho busca revelar desigualdades estruturais, evidenciar regiões prioritárias para políticas públicas e contribuir para uma compreensão mais clara da organização dos serviços de saúde mental no estado de São Paulo.

## 2. Aspectos Éticos e Responsabilidade

O uso de Inteligência Artificial (IA) em análises de saúde pública envolve questões éticas importantes, especialmente quando se trabalha com dados sensíveis sobre indivíduos e comunidades. Neste estudo, os dados utilizados são provenientes de fontes oficiais, garantindo anonimização e proteção à privacidade, evitando a identificação de indivíduos. Ressalta-se que o respeito à confidencialidade e à segurança das informações é um requisito ético fundamental, alinhado às normas legais e regulatórias vigentes.

Além disso, o uso de algoritmos de IA, como *clustering* e árvores de decisão, deve ser feito com cuidado, garantindo que os resultados sejam claros, confiáveis e não prejudiquem grupos específicos. Os pesquisadores têm a responsabilidade de usar os insights obtidos para orientar políticas públicas e decisões estratégicas de forma ética, justa e socialmente responsável, sem reforçar desigualdades ou estigmatizar pessoas vulneráveis.

Além disso, é importante comunicar os resultados de maneira clara e acessível, evitando simplificações que possam gerar mal-entendidos sobre o acesso à saúde mental nas diferentes regiões. A ética no desenvolvimento da solução envolve não apenas proteger os dados, mas também usar as informações de forma consciente para promover equidade, inclusão e melhorar a atenção psicossocial no país.

## 3. Dataset, Análise Exploratória e Preparação dos Dados

### 3.1. Base de Dados

A base de dados utilizada neste estudo foi composta por duas fontes oficiais e complementares.

A primeira refere-se ao Sistema de Informações Hospitalares do SUS (SIH-SUS), disponibilizado pelo DATASUS, que registra as internações hospitalares financiadas pelo SUS em todo o país. Para este trabalho, os dados foram obtidos por meio da biblioteca *Pysus*, em ambiente *Google Colab*, possibilitando a automação do processo de download, leitura e conversão dos arquivos para o formato .parquet, o que facilita a manipulação em análises computacionais.

A segunda fonte corresponde ao Censo Demográfico Brasileiro, originalmente fornecido pelo Instituto Brasileiro de Geografia e Estatística (IBGE). Os dados utilizados foram obtidos a partir do dataset “*Population of Brazilian Municipalities*” disponibilizado por Daniel Komesu na plataforma *Kaggle*. Em relação aos dados sobre os Centros de Atenção Psicossocial (CAPS), foi utilizado o arquivo

estabelecimentos.csv, obtido a partir do Cadastro Nacional de Estabelecimentos de Saúde (CNES), mantido pelo Ministério da Saúde.

Como dado complementar, foram utilizados os códigos dos municípios por região, disponibilizados pelo Departamento Regional de Saúde do Estado de São Paulo na plataforma Dados Abertos SP. O arquivo utilizado foi codigos\_municipios\_regioes.csv

A base de dados completa foi armazenada no *Google Drive*. O link para acesso à pasta dos dados no Google Drive está presente no Github.

### **3.2. Análise Exploratória e Preparação dos Dados**

Nesta primeira etapa, utilizou-se como escopo uma amostra do conjunto de dados do SIH-SUS, especificamente para o estado de São Paulo no ano de 2022. O foco inicial foi uma análise superficial do conjunto de dados para entender sua estrutura.

Inicialmente, uma análise superficial foi realizada para compreender a estrutura e o conteúdo do *dataset*. Para isso, as funções *df.head()* e *df.tail()* foram utilizadas para visualizar, respectivamente, as 5 primeiras e últimas linhas do *dataset*. A disposição das colunas foi verificada com a listagem completa dos seus nomes. Além disso, foi verificado também o tipo de dado do conjunto de dados.

O processo de preparação avançou com a filtragem dos dados dos anos 2022, 2023 e 2024, garantindo que o escopo de análise se restringisse ao tema. Foram selecionadas apenas as internações relacionadas a transtornos mentais e comportamentais, conforme a classificação do CID-10 (códigos que variam de F00 a F99). Em seguida, filtrou-se também os registros referentes ao Estado de São Paulo, utilizando o código 35. Estes dados foram concatenados em um único arquivo .xlsx que serviu de base para as análises subsequentes.

Em seguida, foram integrados dados complementares do Censo Demográfico, obtidos a partir do *dataset* “*Population of Brazilian Municipalities*” no *Kaggle*, considerando apenas os municípios do estado de São Paulo e o ano de 2022. Os dados sobre os Centros de Atenção Psicossocial (CAPS) foram extraídos do arquivo estabelecimentos.csv do CNES, filtrando registros cujo campo nome\_fantasia continha “CAPS” e pertencentes a municípios do estado de São Paulo. Complementarmente, utilizou-se o arquivo codigos\_municipios\_regioes.csv do Departamento Regional de Saúde do Estado de São Paulo, disponível na plataforma Dados Abertos SP, para analisar a distribuição dos municípios por região de saúde.

Foi realizada uma análise exploratória preliminar em todos os *datasets*, incluindo verificação de valores ausentes, duplicados e consistência dos dados. No caso do *dataset* de CAPS, foram exploradas distribuições numéricas e categóricas, como a quantidade total de CAPS por tipo, distribuição por município e categorias associadas. Para o *dataset* de internações, foram analisadas variáveis como idade, sexo, raça/cor e os dez principais diagnósticos de transtornos mentais e comportamentais. Esses procedimentos tiveram como objetivo garantir a qualidade, integridade e consistência dos dados antes de avançar para a modelagem e análise estatística, enquanto os

resultados obtidos a partir dessas análises serão apresentados na seção de Conclusões e Resultados.

#### **4. Referencial Teórico**

Este trabalho integra conceitos de direito à saúde pública, análise de dados e Inteligência Artificial (IA) para propor uma solução tecnológica ao problema da desigualdade no acesso aos serviços de saúde mental no Brasil.

##### **4.1. Direito e Acesso à Saúde Mental**

Segundo o artigo 196 da Constituição Federal de 1988, a saúde é um direito de todos e dever do Estado, devendo ser garantida por meio de políticas sociais e econômicas (BRASIL, 1988). No caso da Saúde Mental, a Lei nº 10.216/2001 estabeleceu a Política Nacional de Saúde Mental, que promove a atenção comunitária, o convívio familiar e a integração social.

Apesar dessas políticas, desigualdades regionais ainda persistem, afetando o acesso aos serviços especializados de saúde mental, como os Centros de Atenção Psicossocial (CAPS), leitos psiquiátricos, entre outros. Estudos indicam que a cobertura desses serviços é determinante para reduzir internações psiquiátricas e ampliar o acesso ao tratamento adequado (Miliauskas et al., 2019).

##### **4.2. Impacto do CAPS e Integração com a Atenção Básica**

Diversos estudos têm destacado a relevância dos CAPS como fator determinante para a redução de internações psiquiátricas e ampliação do acesso a cuidados em saúde mental. Miliauskas et al. (2019), ao investigarem as regiões metropolitanas de São Paulo e Rio de Janeiro, observaram que a expansão desses serviços contribuiu para a diminuição das hospitalizações, evidenciando a importância de políticas públicas que fortaleçam a rede de atenção psicossocial.

Além disso, a integração dos CAPS com a Atenção Básica é considerada fundamental para consolidar uma rede de cuidado mais resolutiva e equitativa, especialmente em regiões carentes de infraestrutura. Assim, compreender a distribuição desses serviços e sua cobertura populacional é um passo essencial para identificar desigualdades e subsidiar ações estratégicas em saúde mental.

##### **4.3. Análise Exploratória de Dados e Inteligência Artificial**

A análise exploratória de dados (AED) é um processo fundamental em qualquer estudo estatístico ou computacional, pois permite compreender a qualidade e as características dos dados antes da aplicação de técnicas mais complexas. Segundo Wickham e Grolemund (2017), a AED é um processo criativo, guiado por perguntas, que busca revelar padrões, verificar inconsistências e gerar hipóteses a partir da observação. A cada nova questão formulada, novos aspectos do conjunto de dados são revelados, aumentando a chance de descobertas relevantes.

Indicadores como número de CAPS por 100 mil habitantes, taxa de internações psiquiátricas e dados socioeconômicos podem ser combinados para identificar regiões com cobertura insuficiente e subsidiar políticas públicas mais efetivas (Dimenstein et al., 2021).

Além de EAD, técnicas de Inteligência Artificial (IA) podem ser aplicadas para analisar dados complexos de saúde pública, identificando padrões mais sutis e relações não lineares entre as variáveis. Por exemplo, clusterização pode ser usada para agrupar municípios com características semelhantes de cobertura e vulnerabilidade enquanto Random Forest pode ser usada para entender quais fatores socioeconômicos ou demográficos mais influenciam a desigualdade de acesso.

Assim, a integração da EDA com técnicas de IA permite uma análise mais completa e estratégica, oferecendo subsídios para planejamento e tomada de decisão em políticas públicas de saúde mental, especialmente em contextos de desigualdade regional.

## 5. Metodologia

Este estudo exploratório visa analisar o acesso à saúde mental no estado de São Paulo a partir de dados estruturados provenientes de fontes oficiais governamentais. Foram utilizados dados do SIH-SUS, com informações sobre internações em saúde mental por município, do IBGE, com dados populacionais e demográficos, do CNES, com registros de Centros de Atenção Psicossocial (CAPS), e do Departamento Regional de Saúde de São Paulo, com a distribuição de municípios por região de saúde. Todos os dados foram integrados em um conjunto único, com tratamento de inconsistências, valores ausentes e anonimização quando necessário.

A análise exploratória envolveu a verificação da estrutura e consistência dos *datasets*, incluindo inspeção das primeiras e últimas linhas, checagem de tipos de variáveis, identificação de valores ausentes e duplicados, e revisão da disposição das colunas. No *dataset* de CAPS, foram filtrados os registros cujo nome continha “CAPS” e pertencentes a municípios do estado de São Paulo, seguida da análise da distribuição por tipo de CAPS, município e variáveis categóricas. No SIH-SUS, foram selecionadas internações relacionadas a transtornos mentais e comportamentais (CID-10 F00 a F99), restringindo-se ao estado de São Paulo, e realizadas análises de distribuição das variáveis idade, sexo, raça/cor e principais diagnósticos. O *dataset* do IBGE foi filtrado para incluir apenas municípios do estado no censo de 2022, enquanto o arquivo de distribuição regional do DRS-SP permitiu analisar a organização geográfica dos municípios.

Para identificar padrões e desigualdades na oferta de serviços, foi aplicada uma abordagem de clusterização das DRS utilizando o algoritmo *K-Means*, considerando como métricas CAPS por 100 mil habitantes e taxa de internações por 100 mil habitantes. Antes da clusterização, as variáveis foram padronizadas e inspecionadas quanto à presença de outliers, utilizando Z-score, e o número ideal de clusters foi definido com base no método do cotovelo e no *silhouette score*. Cada cluster foi caracterizado por estatísticas descritivas das variáveis CAPS, internações e população, bem como pelo perfil sociodemográfico, incluindo idade média, proporção de homens e mulheres e composição racial.

Foram também realizadas análises de correlação entre CAPS, características sociodemográficas e taxa de internações. Para compreender os fatores que influenciam as internações, foi aplicada uma regressão linear múltipla, permitindo quantificar o impacto de variáveis como CAPS, idade média, proporção de homens e composição racial. Complementarmente, um *Random Forest* foi utilizado para determinar a importância relativa das variáveis na previsão da taxa de internações. Além disso, testes estatísticos

(ANOVA) foram conduzidos para verificar diferenças significativas entre clusters, garantindo que os agrupamentos refletissem diferenças reais entre regiões.

Todas as operações foram realizadas no ambiente *Google Colab* com *Python*, utilizando bibliotecas como *Pysus* para extração de dados, *Pandas* e *NumPy* para manipulação, *Matplotlib* e *Seaborn* para visualização, e *Scikit-learn* para implementação de algoritmos de clusterização, regressão e classificação.

Essa metodologia integrada de análise exploratória, clusterização, regressão e testes estatísticos permite caracterizar padrões regionais de cobertura e demanda em saúde mental e fornecer subsídios interpretáveis para gestores públicos, auxiliando na priorização de investimentos e no planejamento estratégico da Rede de Atenção Psicossocial (RAPS), com o objetivo de reduzir desigualdades no acesso aos serviços especializados.

## 6. Resultados

A análise inicial concentrou-se na construção de indicadores padronizados por Departamento Regional de Saúde (DRS) para compreender a distribuição territorial tanto da oferta de serviços quanto da demanda por internações psiquiátricas no estado de São Paulo. Foram calculadas duas métricas principais: a taxa de internações psiquiátricas por 100 mil habitantes e a taxa de Centros de Atenção Psicossocial (CAPS) por 100 mil habitantes. Esses indicadores revelaram um cenário fortemente heterogêneo, evidenciando disparidades significativas entre as regiões.

Observou-se que as DRS de São João da Boa Vista e Marília apresentaram as maiores taxas de internações de todo o estado, configurando-se como polos regionais que absorvem fluxos provenientes de municípios menores e com menor capacidade de atendimento. Em contraste, a Grande São Paulo exibiu taxas relativamente baixas, o que pode estar associado à maior disponibilidade de serviços ambulatoriais, maior diversificação da rede e mecanismos mais robustos de atenção territorial. Na outra ponta, a DRS de Registro apresentou a menor taxa do estado, o que pode refletir baixa demanda registrada, limitações estruturais que restringem a procura por internações ou deslocamento sistemático de pacientes para outras regiões mais equipadas. O gráfico 1 abaixo evidencia essas análises.

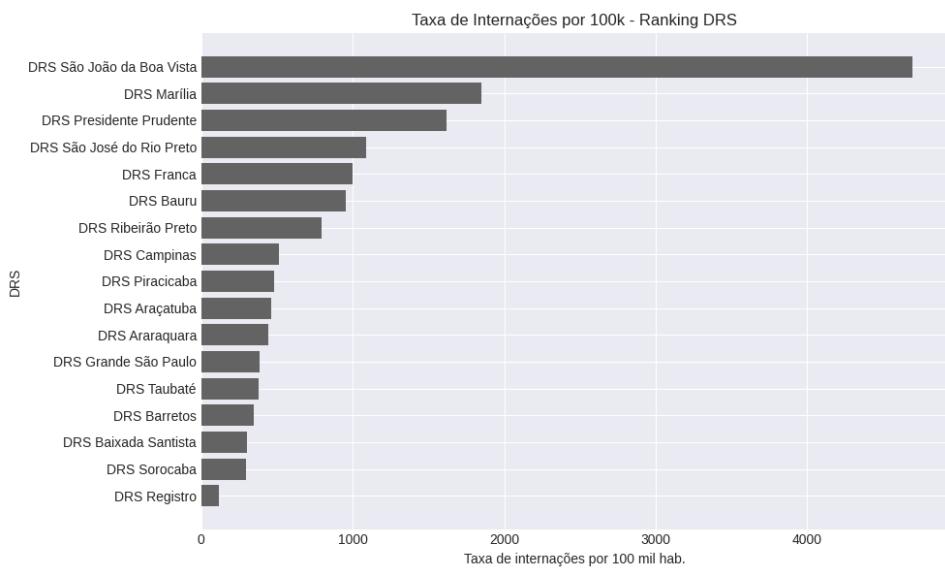


Gráfico 1 - Taxa de Internações por 100 mil habitantes

A distribuição dos CAPS também é desigual, mesmo acompanhando um pouco onde acontecem mais internações. As regiões com maior número de CAPS por habitante geralmente são lugares mais populosos ou que servem como referência para cidades vizinhas. No entanto, ter mais CAPS não significa, automaticamente, que eles estão reduzindo mais as internações, muitas vezes é apenas uma resposta ao volume maior de pessoas que precisam de atendimento nesses locais.

Com o objetivo de investigar a relação entre a oferta de CAPS e as internações psiquiátricas, foi conduzida uma análise de correlação entre as taxas padronizadas. O coeficiente encontrado ( $r = 0,629$ ) indica uma correlação positiva moderada a forte, sugerindo que a presença de CAPS não está associada à redução das internações, mas sim à coexistência em regiões de maior demanda. Esse resultado reforça a interpretação de que a rede funciona mais de forma reativa do que preventiva: os serviços são alocados onde a demanda já é elevada, e não necessariamente onde sua presença poderia atuar de forma preventiva na redução das internações. Assim, em vez de indicar substitutividade entre os níveis de atenção, os dados apontam que oferta e necessidade clínica se acumulam nos mesmos lugares.

A análise estatística também identificou a DRS de São João da Boa Vista como um outlier significativo, tanto em relação à taxa de CAPS por 100 mil habitantes quanto à taxa de internações (como mostrado no gráfico 1), apresentando valores que ultrapassam dois desvios-padrão da média estadual. Esse comportamento extremo sugere que tal região desempenha um papel singular na rede estadual, com possível impacto direto na estruturação dos agrupamentos durante as próximas etapas de modelagem não supervisionada.

Por fim, a integração dos resultados territoriais com a análise exploratória prévia reforça a coerência geral do cenário. Perfis etários concentrados na população economicamente ativa, maior prevalência entre homens, desigualdades raciais, predominância de grandes centros urbanos como referência regional e distribuição desigual dos CAPS convergem para a compreensão de um sistema pressionado, concentrado e estruturalmente assimétrico. Esses achados fornecem a base necessária

para a etapa subsequente de modelagem por *clustering*, que buscou identificar padrões latentes e agrupar as regiões segundo características multivariadas de demanda e oferta.

A aplicação do algoritmo de *clustering* permitiu identificar três grupos distintos de DRS no estado de São Paulo, cada um refletindo diferentes combinações entre oferta de serviços, intensidade da demanda e características populacionais. A segmentação expôs padrões claros de concentração regional, assim como disparidades estruturais relevantes para o planejamento de políticas públicas em saúde mental.

O Cluster 0, composto por Franca, Marília, Presidente Prudente e Registro, apresentou um perfil intermediário em relação à oferta de CAPS e ao volume de internações. Apesar de reunir apenas quatro DRS, representa mais de 2,8 milhões de habitantes e acumula cerca de 39 mil internações. Sua taxa média de CAPS por 100 mil habitantes é relativamente elevada (2,30), mas acompanha uma taxa de internações ainda mais alta (1146,62), sugerindo que a presença dos serviços não tem gerado efeito protetor claro. O perfil demográfico desse cluster é relativamente jovem, com idade média de 38,7 anos e predominância masculina acentuada (60,7%). A composição racial indica maior proporção de pessoas pardas em comparação aos demais clusters, o que reflete características sociodemográficas típicas de regiões interioranas e de menor densidade populacional.

O Cluster 1, formado por 12 DRS, incluindo localidades expressivas como Grande São Paulo, Campinas, Sorocaba, São José do Rio Preto e Baixada Santista, concentra quase toda a população do estado (mais de 40 milhões de habitantes) e a maior parte da estrutura assistencial, com 517 CAPS. Apesar disso, apresenta a menor taxa média de CAPS por 100 mil habitantes, cerca de 1,50, sinalizando subdimensionamento relativo frente ao porte populacional. Ainda assim, suas taxas de internação (536,15 por 100k) são significativamente menores que as dos outros clusters, o que sugere maior diversificação dos serviços, maior capilaridade da rede territorial ou melhor capacidade de manejo ambulatorial. O perfil demográfico é mais envelhecido e mais equilibrado em termos de sexo, com idade média de 40,3 anos e 56,9% de usuários masculinos. A composição racial desse grupo também apresenta maior heterogeneidade urbana.

Já o Cluster 2 destaca-se como um caso extremo, uma verdadeira “ilha estatística”: composto exclusivamente pela DRS de São João da Boa Vista, ele concentra simultaneamente a maior oferta relativa de CAPS (2,86 por 100k) e, de forma muito mais expressiva, a maior taxa de internações do estado (4695,65 por 100k). Trata-se de um outlier estrutural, cuja intensidade de demanda supera em magnitude as demais regiões, inclusive quando ajustada pela população. Demograficamente, apresenta o perfil mais envelhecido (idade média de 43,7 anos), maior concentração de homens (70,8%) e maior proporção de pessoas brancas. Esses resultados apontam para uma dinâmica regional altamente singular, possivelmente ligada a características da rede hospitalar local, fluxos intermunicipais intensos ou concentração histórica de serviços psiquiátricos.

Indicador	Cluster 0	Cluster 1	Cluster 2
Nº de DRS	4	12	1
CAPS por 100k (média)	2.30	1.50	2.86
CAPS por 100k (mín–máx)	2.08–2.52	1.00–1.87	2.86–2.86
Internações por 100k (média)	1146.62	536.15	4695.65
Internações por 100k (mín–máx)	118–1850	296–1089	4695–4695
População total	2.8 M	40.8 M	0.8 M
Total CAPS	63	517	23
Total internações	39,671	188,308	37,773

Tabela 1 - Quadro Comparativo de Clusters

Para aprofundar a compreensão dos padrões identificados pelos clusters, foram avaliadas correlações bivariadas entre variáveis demográficas e a taxa de internações. Os resultados mostraram que a proporção de homens é o fator que mais se correlaciona com a taxa de internações ( $r = 0,671$ ), indicando que regiões com maior predominância masculina tendem a registrar maior demanda hospitalar. A taxa de CAPS também apresentou correlação positiva e elevada ( $r = 0,629$ ), reforçando a interpretação de que os serviços são alocados em regiões de demanda já consolidada, e não como instrumento de prevenção. Variáveis como idade média ( $r = 0,476$ ) e proporção de pessoas brancas ( $r = 0,314$ ) apresentam relações moderadas com o volume de internações, enquanto a proporção de pessoas pardas apresenta correlação negativa fraca.

A análise multivariada reforçou esse panorama. O modelo de regressão linear apresentou excelente poder explicativo ( $R^2 = 0,771$ ), indicando que idade média, proporção masculina e densidade de CAPS são fortes preditores das internações regionais. O coeficiente da variável *caps\_100k* (+1228,53) revela que, longe de reduzir as internações, a maior oferta acompanha, e possivelmente responde, a maior demanda pré-existente, como visto no gráfico 3 abaixo.

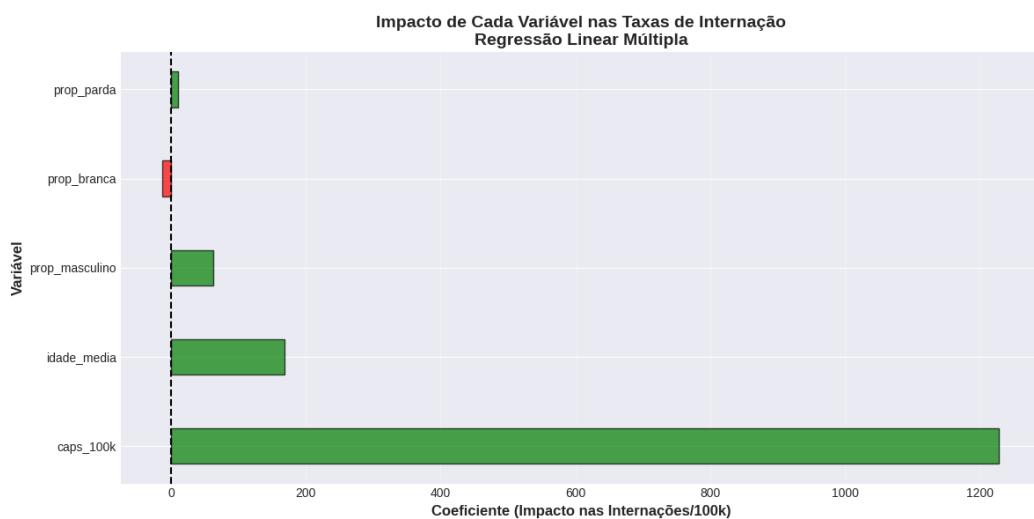


Gráfico 2 - Impacto das Variáveis com Regressão Linear Múltipla

Já o *random forest* elevou ainda mais o desempenho ( $R^2 = 0,892$ ), destacando como variáveis mais relevantes: idade média (30,8%), proporção masculina (29,6%) e taxa de CAPS (28,7%), como visto no gráfico 2. Esses achados convergem na interpretação de que fatores demográficos e estruturais desempenham papel central na definição dos padrões regionais de internação.

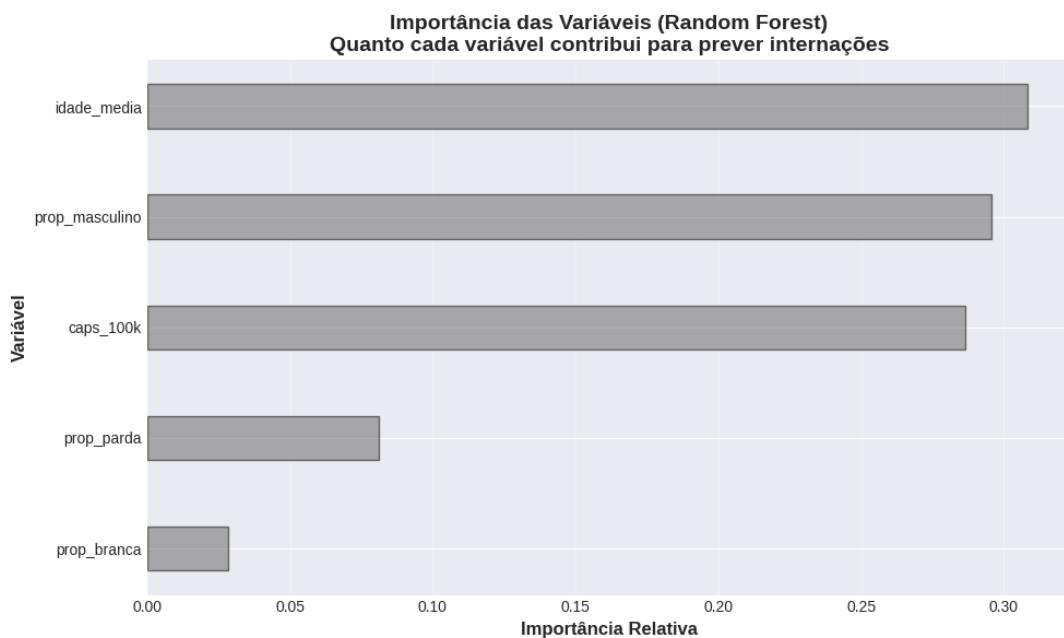


Gráfico 3 - Análise da Importância das Variáveis

Por fim, testes estatísticos comparando variáveis demográficas entre clusters indicaram ausência de diferenças significativas em idade e sexo, reforçando que as diferenças mais marcantes entre os grupos estão na dinâmica das internações e na oferta de serviços, não no perfil populacional. Em contrapartida, a taxa de CAPS entre os clusters apresentou diferença significativa, destacando seu papel estruturante na classificação obtida.

No conjunto, os resultados revelam um sistema profundamente desigual, em que regiões de alta demanda concentram tanto maior oferta de serviços quanto maiores taxas de internações, enquanto áreas densamente povoadas como a Grande São Paulo operam com suboferta relativa, porém com melhor desempenho em internações per capita. O achado mais crítico é a coexistência de CAPS mais numerosos em regiões que também apresentam maior carga de internações, indicando uma rede assistencial que reage à demanda, mas que pode não estar funcionando de forma plenamente resolutiva. Esses elementos fornecem base sólida para recomendações de redistribuição de serviços, revisão de estratégias regionais e aprofundamento investigativo nas DRS que aparecem como extremos estatísticos.

## 7. Conclusão

O estudo teve como objetivo compreender a distribuição territorial das internações psiquiátricas no estado de São Paulo e avaliar como a rede de atenção psicossocial, especialmente os CAPS, se relaciona com esse padrão. A proposta central era verificar se a oferta de serviços territoriais estaria associada à redução das internações ou se, na prática, ambos os fenômenos se concentravam nos mesmos espaços, indicando um sistema que opera mais pela resposta a demandas já existentes do que pela prevenção estruturada.

Os resultados mostraram que o cenário é marcado por fortes desigualdades regionais. Foram identificadas diferenças expressivas entre os DRS, tanto na taxa de

internações quanto na oferta de CAPS, com destaque para São João da Boa Vista e Marília como regiões de alta intensidade e para Registro como região de baixa demanda registrada. A correlação positiva entre CAPS por habitante e internações reforça que a presença de serviços não implica redução direta das internações, mas sim coexistência em regiões pressionadas por maiores necessidades assistenciais.

Com isso, os resultados esperados foram parcialmente alcançados. Esperava-se encontrar algum indício de relação inversa entre oferta de CAPS e internações, o que não se confirmou. No entanto, a investigação conseguiu cumprir o propósito de revelar o padrão territorial real, oferecendo justificativas consistentes: a organização da rede ainda está fortemente associada à demanda histórica, e não à prevenção ou ao equilíbrio regional. Além disso, a identificação de outliers como a DRS de São João da Boa Vista contribuiu para evidenciar que certos territórios exercem papel desproporcional na rede, algo que poderá orientar futuras políticas públicas.

Em síntese, o estudo cumpriu sua função ao oferecer uma visão clara, comparativa e crítica da estrutura atual da atenção em saúde mental no estado. Embora nem todos os efeitos desejados, como evidenciar substitutividade entre CAPS e internações, tenham emergido, os achados fornecem base sólida para aprofundar a análise por meio de técnicas de clusterização e para subsidiar decisões de planejamento mais alinhadas às reais necessidades de cada região.

## **8. Endereço do GitHub e Vídeo da Apresentação**

O repositório contendo o material analítico está disponível em: <https://github.com/tinemyumi/saude-mental-datasus> e o vídeo da apresentação está postado na plataforma de vídeos Youtube acessível por este link: <https://youtu.be/bnfQOuicw7Q>.

## **9. Referências bibliográficas**

BRASIL. *Constituição da República Federativa do Brasil de 1988*. Brasília, DF: Presidência da República, [2024]. Disponível em: <https://www.gov.br/conselhonacional-de-saude/pt-br/acesso-a-informacao/legislacao/outrasnormativas/constituicao-federal.pdf/view>.

BRASIL. *Lei nº 10.216, de 6 de abril de 2001*. Dispõe sobre a proteção e os direitos das pessoas portadoras de transtornos mentais e redireciona o modelo assistencial em saúde mental. Diário Oficial da União: seção 1, Brasília, DF, 6 abr. 2001. Disponível em: [http://www.planalto.gov.br/ccivil\\_03/leis/leis\\_2001/l10216.htm](http://www.planalto.gov.br/ccivil_03/leis/leis_2001/l10216.htm).

MILIAUSKAS, C. R.; FAUS, D. P.; JUNKES, L.; RODRIGUES, R. B.; JUNGER, W. *Associação entre internações psiquiátricas, cobertura de CAPS e atenção básica em regiões metropolitanas do RJ e SP, Brasil*. Ciência & Saúde Coletiva, v. 24, n. 5, p. 1935-1944, maio 2019.

Rocha HA, Reis IA, Santos MAC, Melo APS, Cherchiglia ML. *Internações psiquiátricas pelo Sistema Único de Saúde no Brasil ocorridas entre 2000 e 2014*. Rev Saude Publica. 2021;55:14. <https://doi.org/10.11606/s15188787.2021055002155>

DIMENSTEIN, M. et al. *Equidade e acesso aos cuidados em saúde mental em três estados nordestinos*. Ciência & Saúde Coletiva, v. 26, n. 5, p. 1727–1738, maio 2021.

WICKHAM, H.; GROLEMUND, G. *R for Data Science: import, tidy, transform, visualize, and model data*. 1. ed. Sebastopol: O'Reilly Media, 2017. Disponível em: <https://r4ds.had.co.nz/exploratory-data-analysis.html>.