

# Análise da Tendência de Suicídio no Brasil: o Efeito da Pandemia de COVID-19

Luciana Prachedes<sup>1</sup>, Heder Bernardino<sup>1</sup>, Leonardo Martins<sup>2</sup>,  
Felipe Souza<sup>1</sup>, Rynane Wenecha<sup>3</sup>, Magda Dimenstein<sup>3</sup>

<sup>1</sup>Universidade Federal de Juiz de Fora (UFJF)  
Juiz de Fora – MG – Brasil

<sup>2</sup> Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio)  
Rio de Janeiro – RJ – Brasil

<sup>3</sup> Universidade Federal do Delta do Parnaíba (UFDPAr)  
Parnaíba – PI – Brasil

<sup>4</sup> Universidade Federal do Rio Grande do Norte (UFRN)  
Natal – RN – Brasil

luciananssantana@gmail.com, heder.bernardino@ufjf.br,  
leonardomartins@puc-rio.br, feliperafaeldesouza@gmail.com,  
ryannewenecha@hotmail.com, magda.dimenstein@ufrn.br

**Abstract.** *Suicide is one of the leading causes of death worldwide. Linked to the mental health conditions of a population, stressful global events can affect trends. This study analyzed suicide trends in Brazil between 2003 and 2022, focusing on the COVID-19 pandemic period. We used data from the Mortality Information System (SIM) of DATASUS (218,707 suicide occurrences). Suicide mortality rates were predicted for the pandemic years and beyond (2020-2022) using time series models (Prophet, SARIMA, and LSTM) trained with data up to 2019. Here, we compare the actual results with those forecasted to evaluate a possible impact of the pandemic. Our main findings show that the suicide rate in Brazil nearly doubled over two decades (from 4.45 to 8.11 per 100,000 inhabitants), with a pronounced increase during the COVID-19 pandemic. The time series models tested here performed well prior to 2020 but underestimated suicide peaks during the pandemic, underestimating the impact of COVID-19 on forecasting complex phenomena such as suicide. Our analysis highlighted the importance of public policies focused on mental health and reinforces the need for preventive actions, especially in the context of crises such as COVID-19.*

**Resumo.** *O suicídio é uma das principais causas de morte no mundo todo. Atrelado às condições de saúde mental de uma população, eventos globais estressantes podem afetar as tendências. Este trabalho analisou as tendências de suicídio no Brasil entre 2003 e 2022, com foco no período da pandemia de COVID-19. Foram usados dados do Sistema de Informação sobre Mortalidade (SIM) do DATASUS (218.707 ocorrências de suicídio). Foram previstas as taxas de mortalidade por suicídio nos anos da pandemia e seguintes (2020-2022) com modelos de séries temporais (Prophet, SARIMA e LSTM) treinados com dados até 2019. O intuito é comparar os resultados reais e os obtidos de forma a avaliar um possível impacto da pandemia. Os principais resultados indicaram que a taxa de suicídio praticamente dobrou ao longo de 20 anos (4,45 para*

*8,11 por 100 mil habitantes), com elevação mais expressiva durante a pandemia de COVID-19. Os modelos de séries temporais testados obtiveram bom desempenho em períodos anteriores a 2020, mas subestimaram os picos de suicídio durante a pandemia, reforçando o papel de eventos disruptivos na previsão de fenômenos complexos como o suicídio. A análise destacou a importância de políticas públicas voltadas para saúde mental e reforça a necessidade de ações preventivas, especialmente no contexto de crises como a COVID-19.*

## **1. Introdução**

O suicídio é reconhecido globalmente como um grave problema de saúde pública que afeta indivíduos, famílias e comunidades inteiras. De acordo com a Organização Mundial da Saúde (OMS), aproximadamente 700 mil pessoas morreram por suicídio em 2019, representando uma das principais causas de morte em todo o mundo<sup>1</sup>. No Brasil, as taxas de suicídio têm aumentado na última década, evidenciando a necessidade de uma compreensão mais aprofundada desse fenômeno no contexto nacional [Soares et al. 2023].

Diversos fatores estão associados ao suicídio, incluindo aspectos biológicos, psicológicos, sociais e culturais [Lopes et al. 2023]. Fatores demográficos como sexo, idade, raça/cor e nível de escolaridade desempenham papéis importantes na determinação dos grupos mais vulneráveis [Braun et al. 2023]. Além disso, eventos de grande impacto social e econômico, como a pandemia de COVID-19, podem exacerbar condições de risco e influenciar as tendências de suicídio em todo o país [Silva Junior et al. 2023]. Regiões com menores índices de desenvolvimento humano tendem a apresentar taxas mais elevadas, indicando a influência dos determinantes sociais na saúde [de Lyra Reis and Teixeira 2023]. Entretanto, ainda há lacunas na literatura sobre como eventos globais disruptivos, como pandemias, afetam as tendências de suicídio em países com diferenças socioeconômicas marcantes, como o Brasil.

A motivação para este estudo nasce na necessidade de explorar o potencial do aprendizado de máquina na compreensão de questões delicadas como o suicídio. Essas ferramentas têm sido amplamente usadas para examinar grandes volumes de dados e identificar padrões, identificando relações entre variáveis que não seriam facilmente detectadas por análises tradicionais [Nurmi and Lohan 2021]. A abordagem adotada aqui é relevante no cenário brasileiro, onde há escassez de análises aprofundadas sobre os efeitos da pandemia de COVID-19 na saúde mental e nas taxas de suicídio. Com a pandemia elevando os riscos à saúde mental [Bosi and Alves 2023], é preciso investigar seu efeito sobre as ocorrências de suicídio. Desta forma, a proposta aqui é avaliar modelos preditivos para a quantidade de suicídios ao longo do tempo e comparar essas previsões com o ocorrido durante a pandemia. Como resultado, foi possível observar que todos os modelos subestimaram as ocorrências, levando a um possível entendimento de que o aumento da quantidade de suicídios foi maior do que foi possível prever.

## **2. Trabalhos Relacionados**

Soares et al. [Soares et al. 2023] teve como objetivo descrever a tendência das taxas de suicídio no Brasil de 2011 a 2020 e investigar a associação dessas taxas com variáveis relacionadas à pandemia de COVID-19. O estudo utilizou um método ecológico de série

---

<sup>1</sup> <https://www.who.int/publications/i/item/9789240026643>, acessado em março de 2025.

temporal para analisar a mortalidade por suicídio, com foco em 2020, o primeiro ano da pandemia. A taxa de mortalidade por suicídio por 100 mil habitantes foi o desfecho principal. O estudo empregou a regressão generalizada de Prais-Winsten para calcular a tendência das taxas de suicídio ao longo do período e utilizou o modelo ARIMA para prever as taxas de suicídio em 2020. Além disso, investigou a associação entre as taxas de suicídio e variáveis como mortalidade por COVID-19, desemprego, recebimento de auxílio emergencial e taxas de distanciamento social. Os resultados mostraram uma taxa de suicídio de 6,68 por 100 mil habitantes em 2020. As taxas de suicídio aumentaram em ambos os sexos e em todas as regiões do país, especialmente no Sul. No entanto, o estudo não encontrou evidências de um aumento considerável de suicídios em 2020 em relação ao número esperado durante o primeiro ano da pandemia. O estudo concluiu que, embora as taxas de suicídio tenham aumentado de 2011 a 2020, especialmente entre homens e nas regiões Sul, não houve um aumento inesperado no primeiro ano da pandemia, com a tendência se mantendo com o crescimento dentro do esperado em 2020.

No contexto da Coreia do Sul, um trabalho investigou o efeito da pandemia de COVID-19 nas tendências de suicídio no país por meio de uma análise de séries temporais [Ryu et al. 2022]. Usaram o Prophet para gerar modelos de previsão com base no número mensal de mortes por suicídio na Coreia do Sul entre 1997 e 2018. Validaram os modelos comparando com os números de suicídios de 2019 e previram os números de suicídios para 2020. A comparação entre os números esperados e os observados durante a pandemia indicou que o total de suicídios durante a pandemia não se desviou significativamente das projeções feitas com base no período pré-pandemia. No entanto, o número de suicídios entre mulheres e pessoas com menos de 34 anos ultrapassou significativamente o nível esperado. O estudo concluiu que a pandemia de COVID-19 não aumentou significativamente a taxa geral de suicídios, mas que pode ter contribuído para o aumento de suicídios entre mulheres e jovens.

Tanaka et al. [Tanaka and Okamoto 2021] analisou o impacto da pandemia de COVID-19 na mortalidade por suicídio no Japão via registros mensais de suicídios. Usando *difference-in-difference estimation*, os pesquisadores descobriram que as taxas mensais de suicídio caíram 14% durante os primeiros cinco meses da pandemia (fevereiro a junho de 2020). Essa queda pode ter ocorrido devido a fatores complexos, como subsídios do governo, redução das horas de trabalho e fechamento das escolas. No entanto, as taxas mensais de suicídio aumentaram 16% durante a segunda onda da pandemia (julho a outubro de 2020), com um aumento maior entre mulheres (37%) e crianças e adolescentes (49%). O estudo propõe que, embora os impactos negativos da pandemia de COVID-19 possam persistir a longo prazo, modificadores como os subsídios governamentais podem não ser sustentáveis. Assim, sugere que a prevenção eficaz do suicídio, especialmente entre populações vulneráveis, é importante para a saúde pública.

### **3. Métodos**

#### **3.1. Modelagem de Séries Temporais**

A série temporal é definida como uma sequência de observações de uma ou mais variáveis registradas em intervalos de tempo sucessivos, frequentemente espaçados de forma uniforme [Araghinejad 2014]. A análise de séries temporais se concentra na modelagem e previsão desses dados. A natureza sequencial das séries temporais permite a captura de padrões temporais, tendências, ciclos e outras estruturas.

Uma característica das séries temporais é a autocorrelação, na qual as observações em diferentes momentos podem estar correlacionadas [Lomnicki and Zaremba 1957]. Essa dependência temporal requer técnicas especializadas e que considerem a estrutura sequencial dos dados. Modelos estatísticos clássicos, como os autorregressivos de médias móveis (ARMA), têm sido usados para capturar padrões lineares em séries estacionárias [Zhang 2018]. Para séries não estacionárias, onde as propriedades estatísticas mudam ao longo do tempo, modelos como o autorregressivo integrado de médias móveis (ARIMA) e suas extensões sazonais (SARIMA), são aplicados [Sirisha et al. 2022].

Com o avanço das técnicas de aprendizado de máquina, modelos mais complexos foram introduzidos para lidar com não linearidades e interações sofisticadas em séries temporais. Redes neurais recorrentes (RNNs), particularmente aquelas com unidades de memória de curto e longo prazo (LSTM) e redes neurais recorrentes bidirecionais, têm demonstrado eficácia na captura de dependências de longo alcance e padrões complexos em dados temporais [Dubey et al. 2021]. Além disso, modelos baseados em atenção e transformadores têm sido explorados para melhorar a capacidade de modelar relacionamentos temporais de maneira mais flexível [Mohammadi Farsani and Pazouki 2020].

### 3.2. Prophet

O Prophet é um modelo de séries temporais desenvolvido pela Meta com o objetivo de ser robusto e de fácil uso [Rafferty 2021]. Projetado para lidar com dados que apresentam tendências não lineares, sazonalidades múltiplas e efeitos de feriados, o Prophet baseia-se em um modelo aditivo onde diferentes componentes são combinados para capturar padrões complexos nos dados temporais [Sivaramakrishnan et al. 2022]. Uma característica distintiva do Prophet é sua abordagem orientada ao usuário, permitindo que especialistas no domínio influenciem o processo de modelagem. Os usuários podem incorporar conhecimento específico, como datas de feriados relevantes, ajustes manuais nas sazonalidades ou alterações nos parâmetros de crescimento. Essa flexibilidade facilita a adaptação do modelo a diferentes contextos e melhora a precisão das previsões em cenários onde informações adicionais estão disponíveis [Taylor and Letham 2018].

### 3.3. SARIMA

O modelo *Seasonal Autoregressive Integrated Moving Average* (SARIMA) é uma extensão do modelo ARIMA que incorpora componentes sazonais para modelar e prever séries temporais que exibem padrões sazonais regulares [Dubey et al. 2021]. O SARIMA é muito usado em estatística e econometria para capturar tanto a dinâmica de curto prazo quanto as flutuações sazonais de longo prazo em dados temporais [Kaczmarczyk 2017].

A formulação do modelo SARIMA é denotada como  $SARIMA(p, d, q) \times (P, D, Q)$ , onde  $p$  representa a ordem autorregressiva não sazonal,  $d$  é o grau de diferenciação não sazonal e  $q$  denota a ordem de médias móveis não sazonais. Os termos  $P$ ,  $D$  e  $Q$  correspondem às ordens autorregressiva, de diferenciação e de médias móveis sazonais, respectivamente, enquanto  $s$  indica o período sazonal (por exemplo,  $s = 12$  para dados mensais com sazonalidade anual). Os componentes não sazonais (AR, I, MA) capturam a dependência linear de curto prazo entre observações em diferentes momentos, enquanto os componentes sazonais (SAR, SI, SMA) modelam padrões que se repetem em intervalos fixos ao longo do tempo. A diferenciação ( $d$  e  $D$ ) é aplicada para tornar a série estacionária, removendo tendências e componentes sazonais de nível, o que é importante para a aplicação adequada de modelos autorregressivos [Feng et al. 2022].

### 3.4. LSTM

As redes neurais artificiais são modelos inspirados no cérebro humano, projetadas para reconhecer padrões complexos e realizar tarefas como classificação e regressão [Walczak 2019]. São formadas por camadas de neurônios artificiais interconectados, onde cada neurônio processa suas entradas, aplica uma função de ativação e transmite o resultado para os neurônios da próxima camada. O aprendizado em redes neurais é alcançado através do ajuste iterativo dos pesos sinápticos via algoritmos de otimização.

As redes neurais recorrentes (RNNs) são redes neurais que podem lidar com dados sequenciais ou temporais, onde a ordem das informações é importante [Caterini et al. 2018]. Diferentemente das redes *feedforward* tradicionais, as RNNs possuem conexões cíclicas que permitem que informações de estados anteriores influenciem o processamento atual. Essa característica confere às RNNs uma forma de memória interna, possibilitando a modelagem de dependências temporais e padrões sequenciais em tarefas como processamento de linguagem natural, reconhecimento de fala e análise de séries temporais [Jelodar et al. 2020].

No entanto, as RNNs tradicionais enfrentam limitações ao lidar com dependências de longo prazo devido ao problema do gradiente desaparecente ou explosivo durante o treinamento [Yu et al. 2019]. Para superá-las, foram desenvolvidas arquiteturas recorrentes como a *Long Short-Term Memory* (LSTM). As LSTMs foram introduzidas em 1997 como uma solução para o problema do gradiente desaparecendo, permitindo que a rede aprenda dependências de longo prazo de maneira eficaz [Hochreiter 1997]. A arquitetura LSTM introduz uma estrutura de memória especializada, denominada célula LSTM, que utiliza mecanismos de portas para controlar o fluxo de informações. O mecanismo de portas das LSTMs permite que a rede retenha ou ignore informações conforme necessário, adaptando-se dinamicamente às características dos dados sequenciais [Landi et al. 2021].

## 4. Experimentos Computacionais

O processo metodológico adotado é descrito nesta seção. A base de dados e a modelagem de séries temporais adotada são discutidas. O código fonte em *Python* do tratamento dos dados, das análises e dos modelos está publicamente disponível<sup>2</sup>.

### 4.1. Fonte dos Dados

Os dados utilizados neste estudo foram extraídos da Plataforma de Dados de Pesquisa em Saúde (PCDAS)<sup>3</sup>, que disponibiliza acesso a bases de dados de saúde pública no Brasil. Foram usados dados do Sistema de Informação sobre Mortalidade (SIM) do Departamento de Informática do SUS (DATASUS), que reúne dados sobre óbitos no país.

Foi realizada uma consulta SQL para selecionar os registros de óbitos por suicídio entre 2003 e 2022, com base nos códigos de causa básica relacionados ao suicídio. Seguindo o padrão de boletins epidemiológicos do Ministério da Saúde, consideraram-se suicídios os óbitos em indivíduos com 5 anos de idade ou mais, com causa básica registrada com códigos X60-X84 e Y87.0 da Classificação Internacional de Doenças, décima revisão (CID-10) [de Vigilância em Saude 2024]. Essa consulta resultou em 218.707 registros, incluindo informações sobre data do óbito, causa básica, sexo, raça/cor, escolaridade, local de ocorrência, município de residência, estado da ocorrência e idade do

<sup>2</sup><https://github.com/lunascas/sbcas25>

<sup>3</sup><https://pcdas.icict.fiocruz.br>

óbito. As variáveis escolhidas foram baseadas nas variáveis que aparecem em boletins epidemiológicos do governo e possuíam mais de 60% de preenchimento.

## **4.2. Modelagem de Séries Temporais**

Para analisar as tendências temporais dos suicídios e avaliar se houve impacto da pandemia de COVID-19, foram utilizados três modelos de séries temporais. Primeiro foi feita uma agregação mensal onde os dados foram agregados por mês, resultando em uma série temporal mensal de quantidade de suicídios de janeiro de 2003 a dezembro de 2022. Não haviam valores faltantes. Utilizar dados mensais na modelagem de séries temporais para análise de suicídios é interessante porque permite identificar padrões sazonais e tendências de longo prazo com mais clareza. Essa granularidade ajuda a capturar variações mensais significativas, como as influenciadas por eventos sociais ou econômicos. Além disso, os dados mensais oferecem um equilíbrio entre a simplicidade dos dados anuais e o ruído excessivo dos dados diários, sendo ideais para detectar impactos mensuráveis, como o efeito da pandemia de COVID-19, sobre as taxas de suicídio.

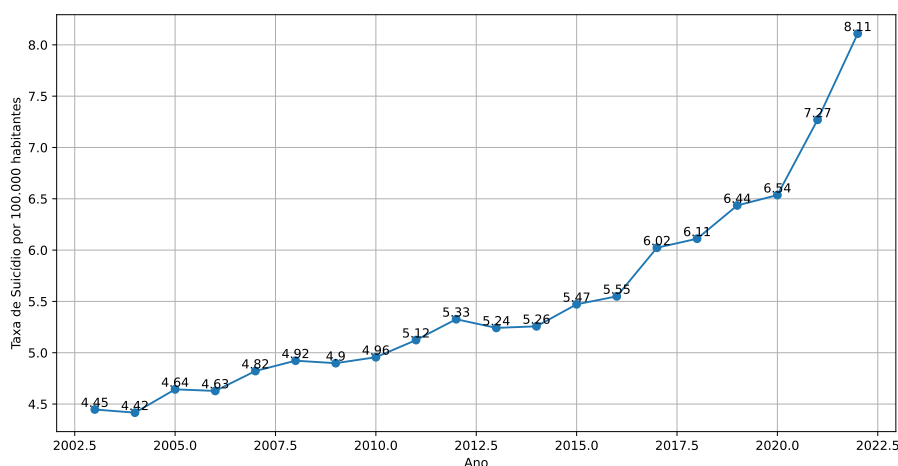
Os três modelos utilizados (Prophet, SARIMA, LSTM) seguiram o mesmo padrão de treino com dados até 2018 e janela de informação de 12 meses antes. Uma janela de 12 meses para a modelagem de séries temporais é justificada pela captura de padrões sazonais anuais, comuns em diversos tipos de dados como saúde e economia [Box et al. 2015]. Essa janela equilibra a memória de curto e longo prazo, permitindo que o modelo identifique tanto tendências recentes quanto padrões sazonais recorrentes, sem sobrecarregar o processo com dados antigos ou irrelevantes.

Os modelos foram executados para gerarem previsões mensais de 2019 a 2022. Os dados de 2019 foram usados como dados de validação e os dados de 2020 a 2022 como dados de teste. Os modelos foram avaliados segundo as métricas MAE, MSE, RMSE, e MAPE. Além disso, os resultados passaram por testes de significância estatística em duas etapas. Primeiro, os resíduos (diferença entre os valores observados e previstos) são calculados. Um teste de Shapiro-Wilk é aplicado para verificar se os resíduos seguem uma distribuição normal. Em seguida, um teste-t é realizado para verificar se a média dos resíduos é diferente de zero, o que pode indicar que as previsões são diferentes dos valores reais. Se o p-valor for menor que 0,05, rejeita-se a hipótese nula, significando que há uma diferença significativa.

### **4.2.1. Características Específicas dos Modelos**

Os modelos foram trabalhados com um intervalo de confiança de 95% para garantir uma margem de segurança razoável para lidar com a variabilidade inerente aos dados de séries temporais, sem ser excessivamente conservador. No modelo Prophet, foi considerada a análise dos componentes da previsão, como a tendência e a sazonalidade, oferecendo uma análise dos fatores que influenciam as previsões do modelo.

O modelo SARIMA teve os dados primeiramente submetidos a um teste de dickey-fuller aumentado, para verificar se a série é estacionária. Depois, foi ajustado com dois conjuntos de parâmetros principais:  $order=(p, d, q)$  que são os parâmetros ARIMA não sazonais, onde  $p$  é a ordem do componente autoregressivo e tem valor 1,  $d$  é o grau de diferenciação e tem valor 1, e  $q$  é a ordem do componente de média móvel com valor 1; e  $seasonal\_order=(P, D, Q, m)$  que refere-se aos parâmetros sazonais do SARIMA, com  $P$  sendo o componente autoregressivo sazonal com valor 1,  $D$  a diferenciação sazonal com



**Figura 1. Taxa de Suicídio no Brasil (2003-2022).**

valor 1,  $Q$  o componente de média móvel sazonal também com valor 1, e  $m$  com valor 12 indicando a sazonalidade anual (dados mensais). As opções `enforce_stationarity` e `enforce_invertibility` foram adicionadas como *false* e permitem que o modelo aceite estados estacionários ou invertíveis, mas não forçam essas condições. Após ajustado o teste de dickey-fuller aumentado, verificou-se que a série é estacionária.

O modelo LSTM teve os dados primeiramente normalizados linearmente entre 0 e 1. Em seguida, foi criada uma função para gerar sequências de tempo, onde os últimos 12 meses de dados são usados para prever o mês seguinte. O modelo foi definido com 50 unidades ocultas e uma camada densa, e foi treinado usando o otimizador *Adam* e a função de perda *mean squared error*. O treinamento do modelo foi realizado com 100 épocas, incluindo a técnica de *early stopping* para interromper o treinamento caso a perda no conjunto de validação não melhorasse por 5 épocas consecutivas, evitando *overfitting*.

## 5. Resultados e Discussões

A Figura 1 apresenta a taxa de suicídio no Brasil entre os anos de 2003 e 2022, revelando uma tendência de crescimento, com algumas fases de estabilidade e outras de aumento mais acentuado. Entre 2003 e 2010, a taxa de suicídio permaneceu relativamente estável, variando entre 4.4 e 5.0 por 100.000 habitantes, com um crescimento lento e modesto. Esse período pode indicar um contexto em que os fatores de risco para suicídio, como condições socioeconômicas e saúde mental, estavam relativamente equilibrados, sem grandes mudanças que afetassem drasticamente os números.

A partir de 2011, observa-se um crescimento gradual e constante na taxa de suicídio, que atinge 5.55 em 2016. Esse aumento pode ser explicado por diversas variáveis, como mudanças econômicas, aumento de distúrbios mentais e sociais, ou melhorias no registro e notificação dos casos de suicídio. No entanto, foi a partir de 2017 que a taxa de crescimento se intensificou, com a taxa de suicídio subindo para 6.11 em 2019, antes da pandemia de COVID-19. Após 2020, nota-se uma aceleração ainda mais acentuada, com a taxa subindo de 6.54 para 8.11 por 100.000 habitantes até 2022. Esse aumento pode estar associado aos impactos negativos da pandemia, com particular impacto na saúde mental da população. Ao longo dos 20 anos analisados, a taxa de suicídio no Brasil praticamente dobrou, passando de 4.45 para 8.11 por 100.000 habitantes.

**Tabela 1. Valores obtidos para as métricas de erro e estatísticas (2019-2022). As colunas “test-t” e “Shapiro” contêm os p-valores destes testes.**

Ano	Modelo	MAE	MSE	RMSE	MAPE (%)	teste-t	Shapiro
2019	Prophet	44.64	2571.27	50.71	3.96	0.32	0.34
	SARIMA	<b>39.76</b>	<b>2173.45</b>	<b>46.62</b>	<b>3.49</b>	0.02	0.78
	LSTM	45.65	3157.20	56.19	3.98	0.07	0.93
2020	Prophet	46.89	3806.51	61.70	4.07	0.28	0.847
	SARIMA	<b>45.12</b>	<b>3545.48</b>	<b>59.54</b>	<b>3.81</b>	0.03	0.468
	LSTM	79.81	9193.78	95.88	6.68	0.09	0.480
2021	Prophet	<b>87.53</b>	<b>11479.54</b>	<b>107.14</b>	<b>6.52</b>	0.04	0.008
	SARIMA	132.81	21574.82	146.88	10.04	0.01	0.007
	LSTM	187.30	42691.53	206.62	14.11	0.08	0.474
2022	Prophet	<b>121.39</b>	<b>18909.19</b>	<b>137.51</b>	<b>8.66</b>	0.03	0.396
	SARIMA	182.08	36286.01	190.49	13.12	0.01	0.404
	LSTM	259.90	73097.75	270.37	18.69	0.10	0.734

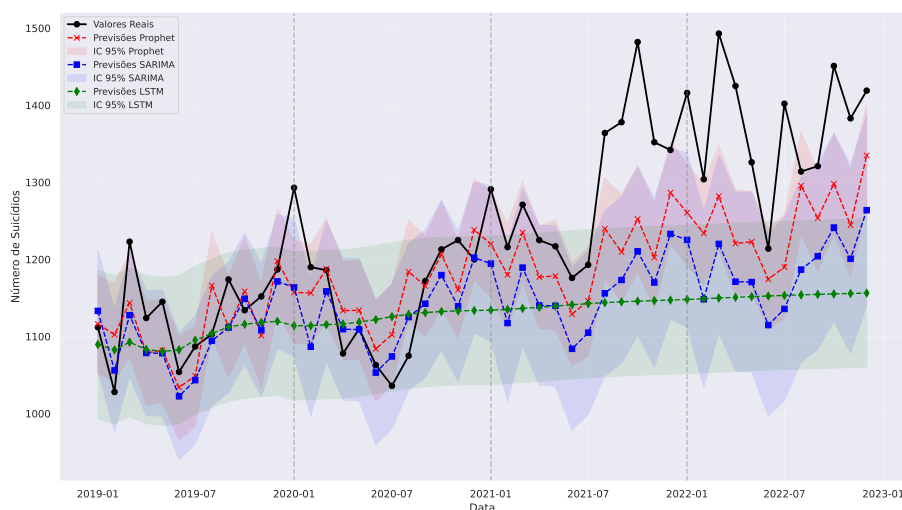
A Tabela 1 mostra os resultados do Prophet, SARIMA e LSTM, e os testes estatísticos. Observando o ano de 2019, SARIMA obteve os menores valores de MAE (39.76), MSE (2173.45) e RMSE (46.62), sugerindo que ele gerou resultados de melhor qualidade. Isso significa que, em média, as previsões do SARIMA são mais próximas dos valores reais do que as previsões do Prophet e do LSTM.

Ainda para 2019, o LSTM apresentou o maior MSE (3157.20) e RMSE (56.19), indicando maior variação nos erros de previsão. No entanto, sua MAE (45.65) e MAPE (3.98%) são semelhantes às do Prophet, o que sugere que o LSTM cometeu mais erros grandes. O Prophet, embora não tenha o melhor desempenho, apresentou valores similares aos do SARIMA, indicando que é uma boa alternativa, com MAPE de 3.96%.

O SARIMA apresentou um p-valor de 0.02 no teste-t, sugerindo que o modelo fez previsões significativamente diferentes dos valores observados, mas em um sentido que pode ser considerado adequado dada a baixa magnitude dos erros. O Prophet e o LSTM têm p-valores de 0.32 e 0.07, indicando que suas previsões não são significativamente diferentes dos valores observados. O teste de Shapiro-Wilk mostra que, para todos os modelos, os resíduos seguem uma distribuição normal, dado que todos os p-valores são superiores a 0.05 (0.34 para Prophet, 0.78 para SARIMA e 0.93 para LSTM).

SARIMA se destaca como o melhor modelo em termos de precisão (menores valores de MAE, MSE e RMSE), embora seus resíduos sejam estatisticamente diferentes de zero. Prophet e LSTM, embora tenham desempenho ligeiramente inferior, apresentam resíduos que não são significativamente diferentes de zero. Em termos de normalidade dos resíduos, todos os modelos têm bom desempenho, o que sugere que os erros são distribuídos de maneira apropriada, reforçando a qualidade das previsões. Portanto, SARIMA se mostra superior em precisão, enquanto Prophet e LSTM têm uma estabilidade maior em termos de resíduos não significativos.

Em contrapartida, com base na análise das métricas de erro entre 2020 e 2022, pode-se observar uma tendência de deterioração no desempenho preditivo de todos os modelos ao longo do tempo. Enquanto o SARIMA apresentava os melhores resultados em 2019, houve uma mudança significativa nos anos seguintes, com o Prophet emergindo



**Figura 2. Valores previstos pelos modelos e valores reais para o mesmo período.**

como o modelo mais resiliente em 2021 e 2022, apresentando erros menores (MAPE de 6.52% em 2021 e 8.66% em 2022). O LSTM, demonstrou o pior desempenho em todos os anos analisados, com erros que chegaram a quase 19% em 2022.

Os valores reais de suicídios, as previsões feitas pelos modelos SARIMA, Prophet e LSTM, e os respectivos intervalos de confiança são apresentados na Figura 2 para o intervalo de 2019 a 2022. O gráfico fornece uma visão ampla da quantidade de suicídios no início e durante a pandemia de COVID-19, que começou em 2020. A partir do início de 2020, observa-se um aumento nos valores reais de suicídios, com picos acentuados.

Todos os modelos subestimaram a incerteza real, pois os valores reais ficaram fora dos intervalos de confiança das previsões (95%). A baixa cobertura sugere que fatores externos causaram uma mudança estrutural que nenhum dos modelos conseguiu capturar adequadamente, algo esperado considerando que foram treinados com dados pré-pandemia (2003-2018), mas que não foi visto com os dados de 2019, onde obtiveram bom resultado. Embora o SARIMA se aproxime dos valores reais em alguns momentos, ele ainda subestima os picos que ocorrem durante 2020 e 2021. Prophet e LSTM também mostram dificuldades em prever os picos, e o LSTM, em particular, apresenta previsões mais suavizadas, falhando em capturar a volatilidade dos dados. Durante 2020 e 2021, os dados reais exibem mais variações bruscas em comparação com os períodos anteriores, tornando mais difícil prever as quantidades de suicídios durante a pandemia.

Sugere-se um reflexo das condições adversas e instáveis impostas pela pandemia, o que dificulta a previsão precisa por meio de modelos estatísticos tradicionais e até mesmo por redes neurais artificiais. Em 2022, os valores reais de suicídios parecem continuar elevados e, embora haja uma leve queda em alguns meses, o padrão de picos altos e frequentes continua. Isso sugere que os impactos da pandemia podem ter efeitos prolongados. Ainda, pode-se dizer que sejam efeitos residuais, como o aumento de problemas de saúde mental e o esgotamento das redes de suporte social e econômico.

Após 2020, o crescimento populacional do Brasil desacelerou. Entre 2010 e 2022, a população brasileira cresceu 6,5%, chegando a 203,1 milhões de habitantes em 2022,

com uma taxa de crescimento anual média de 0,52%—a menor taxa já registrada no país<sup>4</sup>. Assim, o aumento na quantidade de suicídios após a pandemia dificilmente pode ser explicado apenas pelo crescimento populacional, que foi modesto nesse período. Após 2021, o SARIMA e o Prophet conseguem capturar a tendência geral de alta nos números de suicídios, mas falham em prever com precisão os picos mais acentuados. Esses picos podem refletir crises momentâneas ou eventos específicos, difíceis de serem capturados por modelos baseados apenas em dados históricos. O LSTM mantém sua tendência de previsões suavizadas, continuando a subestimar as flutuações reais, sugerindo que ele não foi capaz de se ajustar à mudança no comportamento dos dados após a pandemia.

Nenhum dos modelos (Prophet, SARIMA, ou LSTM) foi capaz de prever com precisão os aumentos bruscos durante a pandemia, ainda que demonstraram ser bons preditores no ano de 2019, refletindo a complexidade e o impacto inesperado que esse evento global teve na saúde mental. Mesmo após a fase crítica da pandemia, os dados sugerem que os efeitos prolongados da crise continuam a influenciar os números de suicídios, resultando em uma tendência geral de alta. Isso ressalta a necessidade de mais estudos e modelos que levem em consideração os pontos mencionados.

## 6. Conclusões e Trabalhos Futuros

Este trabalho realizou uma análise das tendências de suicídio no Brasil entre os anos de 2003 e 2022, com foco especial no impacto da pandemia de COVID-19. Para as análises, foram usados dados do SIM/DATASUS. Observou-se um aumento gradual nas taxas de suicídio no período estudado. A pandemia parece ter exacerbado as tendências de suicídio, com diferença acentuada entre as previsões e os valores observados durante 2020 a 2022. Sugere-se que os efeitos adversos da pandemia na saúde mental, incluindo isolamento social, insegurança econômica e acesso limitado a serviços de saúde, podem ter contribuído para este aumento. Além disso, o aumento dificilmente pode ser explicado apenas pela expansão populacional, dado seu sutil crescimento neste período.

Há a necessidade de ampliar o acesso a serviços de saúde mental, especialmente em períodos de crise como a pandemia. Programas de apoio psicológico e intervenções precoces podem mitigar os riscos associados ao suicídio. Estratégias específicas devem ser desenvolvidas para os grupos mais vulneráveis. Campanhas de conscientização e programas comunitários podem ser eficazes nesses contextos. O suicídio é um fenômeno complexo que requer abordagens multidisciplinares e ações coordenadas entre governos, profissionais de saúde e sociedade civil.

A dependência de registros oficiais sujeita este trabalho a subnotificações. As conclusões refletem tendências gerais e podem não capturar as nuances regionais ou culturais presentes em todo o território brasileiro. Embora modelos avançados tenham sido usados, nenhum modelo pode capturar completamente a complexidade do comportamento humano, e os resultados devem ser interpretados dentro das limitações metodológicas.

Por fim, estes achados reforçam a importância de modelos preditivos em saúde pública, mas evidenciam que eventos globais disruptivos exigem estratégias adicionais de análise. A inclusão de variáveis contextuais (econômicas, sociais e de saúde mental) podem melhorar a robustez dos modelos. Propõe-se futuramente analisar a distribuição

---

<sup>4</sup><http://bit.ly/3FaQAPs>, acessado em setembro de 2024.

geográfica, faixas etárias e o papel de fatores exógenos, como índices de desemprego. Essa abordagem pode contribuir para intervenções públicas mais direcionadas e eficazes.

## Agradecimentos

Os autores agradecem o apoio financeiro do CAPES, CNPq, FAPEMIG e UFJF. Contamos também com o apoio da Chamada Pública Nº 21/2023 “Estudos Transdisciplinares em Saúde Coletiva” do (Decit/SECTICS/MS) e CNPq, Bolsa Jovem Cientista do Nosso Estado (FAPERJ), e Bolsa de Incentivo à Produtividade da PUC-Rio. Agradecemos também à Fiocruz pela concessão do acesso aos dados através da PCDaS.

## Referências

- Araghinejad, S. (2014). *Time Series Modeling*. Springer.
- Bosi, M. L. M. and Alves, E. D. (2023). Distanciamento social em contextos urbanos na pandemia de covid-19: desafios para o campo da saúde mental. *Physis: Revista de Saúde Coletiva*, 33:e33007.
- Box, G. E., Jenkins, G. M., Reinsel, G. C., and Ljung, G. M. (2015). *Time series analysis: forecasting and control*. John Wiley & Sons.
- Braun, B. F., Anjos, G. O., Fonseca, T. M. A., Trevisan, E. R., and de Souza Castro, S. (2023). Perfil epidemiológico dos casos de tentativa de suicídio: revisão integrativa. *SMAD, Rev. Eletrônica Saúde Mental Álcool e Drogas (em Português)*, 19(1):112–122.
- Caterini, A. L., Chang, D. E., Caterini, A. L., and Chang, D. E. (2018). Recurrent neural networks. *Deep neural networks in a mathematical framework*, pages 59–79.
- de Lyra Reis, J. G. and Teixeira, K. H. (2023). Os determinantes socioeconômicos do suicídio nos municípios do brasil. In *Anais do Encontro Nacional da Associação Brasileira de Estudos Regionais e Urbanos (ENABER)*.
- de Vigilância em Saude, S. (2024). Boletim epidemiológico, volume 55, nº 04. Acessado em setembro de 2024.
- Dubey, A. K., Kumar, A., García-Díaz, V., Sharma, A. K., and Kanhaiya, K. (2021). Study and analysis of sarima and lstm in forecasting time series data. *Sustainable Energy Technologies and Assessments*, 47:101474.
- Feng, T., Zheng, Z., Xu, J., Liu, M., Li, M., Jia, H., and Yu, X. (2022). The comparative analysis of sarima, facebook prophet, and lstm for road traffic injury prediction in northeast china. *Frontiers in public health*, 10:946563.
- Hochreiter, S. (1997). Long short-term memory. *Neural Computation MIT-Press*.
- Jelodar, H., Wang, Y., Orji, R., and Huang, S. (2020). Deep sentiment classification and topic discovery on novel coronavirus or covid-19 online discussions: Nlp using lstm recurrent neural network approach. *IEEE Journal of Biomedical and Health Informatics*, 24(10):2733–2742.
- Kaczmarczyk, P. (2017). Microeconometric analysis of telecommunication services market with the use of sarima models. *Dynamic Econometric Models*, 17:41–57.
- Landi, F., Baraldi, L., Cornia, M., and Cucchiara, R. (2021). Working memory connections for lstm. *Neural Networks*, 144:334–341.

- Lomnicki, Z. and Zaremba, S. (1957). On the estimation of autocorrelation in time series. *The Annals of Mathematical Statistics*, 28(1):140–158.
- Lopes, L. O. R., Jesus, R. S. M., Souza, R. S. B., and Teodoro, M. L. M. (2023). Fatores de risco e associados ao comportamento suicida no brasil: uma revisão sistemática. *Tempus Psicológico*, 6(2).
- Mohammadi Farsani, R. and Pazouki, E. (2020). A transformer self-attention model for time series forecasting. *Journal of Electrical and Computer Engineering Innovations (JECEI)*, 9(1):1–10.
- Nurmi, J. and Lohan, E. S. (2021). Systematic review on machine-learning algorithms used in wearable-based ehealth data analysis. *IEEE Access*, 9:112221–112235.
- Rafferty, G. (2021). *Forecasting Time Series Data with Facebook Prophet: Build, improve, and optimize time series forecasting models using the advanced forecasting tool*. Packt Publishing Ltd.
- Ryu, S., Nam, H. J., Jhon, M., Lee, J.-Y., Kim, J.-M., and Kim, S.-W. (2022). Trends in suicide deaths before and after the covid-19 outbreak in korea. *PLoS One*, 17(9):e0273637.
- Silva Junior, A. P. d., Silva Júnior, F. J. G. d., Sales, J. C., Monteiro, C. F. d. S., Miranda, P. I. G., et al. (2023). Estratégias para prevenção e posvenção do suicídio em tempos de pandemia de covid-19. *Interface-Comunicação, Saúde, Educação*, 27:e230181.
- Sirisha, U. M., Belavagi, M. C., and Attigeri, G. (2022). Profit prediction using arima, sarima and lstm models in time series forecasting: A comparison. *IEEE Access*, 10:124715–124727.
- Sivaramakrishnan, S., Fernandez, T. F., Babukarthik, R., and Premalatha, S. (2022). Forecasting time series data using arima and facebook prophet models. In *Big data management in Sensing*, pages 47–59. River Publishers.
- Soares, F. C., Stahnke, D. N., and Levandowski, M. L. (2023). Tendência de suicídio no brasil de 2011 a 2020: foco especial na pandemia de covid-19. *Revista Panamericana de Salud Pública*, 46:e212.
- Tanaka, T. and Okamoto, S. (2021). Increase in suicide following an initial decline during the covid-19 pandemic in japan. *Nature human behaviour*, 5(2):229–238.
- Taylor, S. J. and Letham, B. (2018). Forecasting at scale. *The American Statistician*, 72(1):37–45.
- Walczak, S. (2019). Artificial neural networks. In *Advanced methodologies and technologies in artificial intelligence, computer simulation, and human-computer interaction*, pages 40–53. IGI global.
- Yu, Y., Si, X., Hu, C., and Zhang, J. (2019). A review of recurrent neural networks: Lstm cells and network architectures. *Neural computation*, 31(7):1235–1270.
- Zhang, M. (2018). Time series: Autoregressive models ar, ma, arma, arima. *University of Pittsburgh*.