

SACI (Suicide Analytics for Counter Incidence): de Microdados à Metadados, Machine Learning e Análise de Correlação para a Prevenção do Suicídio em Escala Global

Instituição:

Colégio Visconde de Porto de Seguro – Unidade Panamby

R. Itapaiúna, 1355 - Panamby, São Paulo - SP, CEP: 05707-001

Estudante Autor do Projeto:

Andres Philipp Aeschlimann

Arthur Sandre de Almeida

Orientador:

Prof. Dr. Francisco Tupy

Coorientador: João Pedro Sassi Sandre

Data de início da pesquisa: 02/ Maio/ 2024

Data de fim da pesquisa: 10/ Outubro/ 2024

Sumário

Lista de Figuras.....	p.4
Lista de Tabelas.....	p.6
Resumo.....	p.7
Abstract.....	p.8
Zusammenfassung.....	p.9
1. Introdução.....	p.10
2. Justificativa.....	p.12
3. Hipótese e Problemática.....	p.16
3.1 Hipótese.....	p.16
3.2 Problemática.....	p.16
4. Objetivo.....	p.17
5. Objetos de estudo e revisão bibliográfica.....	p.18
5.1 A compreensão do que é Suicídio.....	p.18
5.2. Entendimento do que é pesquisa em Inteligência Artificial (IA).....	p.19
5.3. As variáveis, sua compreensão e importância para uma análise legítima e o porquê da sua utilização.....	p.20
5.4. As características da Pesquisa em estatística.....	p.21
5.5. Entendimento do que é pesquisa em ciência da informação.....	p.21
6. Métodos e Técnicas.....	p.24
6.1. Método.....	p.24
6.2. Considerações sobre a pesquisa.....	p.26
6.2.1. Considerações da Ética em pesquisa analisando suicídio	p.26
6.2.2. Delimitação da área de estudo (possibilidades e limites em relação à escala de abordagem).....	p.26
6.2.2.1. Análise do contexto local a partir da tabela.....	p.27
6.2.2.2. Análise do contexto regional a partir da tabela	p.28
6.2.2.3. Análise do contexto nacional a partir da tabela.....	p.29
6.2.2.4. Análise do contexto global a partir da tabela.....	p.30
6.3. Processos relacionados a base de dados.....	p.31
6.3.1. Entendimento básico do que é uma base de dados de qualidade.....	p.31
6.3.2. Definindo a amostra de pesquisa e a limitação em uma perspectiva global	p.31
6.3.3. Escolha das bases de dados.....	p.32
6.3.4. Softwares utilizados.....	p.33

6.3.5. Limpeza dos dados.....	p.34
6.4. Técnicas de Pesquisa.....	p.34
6.4.1. Machine learning.....	p.34
6.4.2. Análise estatística.....	p.35
6.5. Procedimentos da pesquisa.....	p.36
6.6. Confeção dos gráficos e imagens	p.37
6.7. Limitações em relação ao modelo.....	p.38
7. Desenvolvimento da pesquisa.....	p.39
7.1 Coleta de dados.....	p.39
7.2 Tratamento e junção das bases de dados.....	p.39
7.3. Descrição dos processos estatísticos.....	p.40
7.4 Análise exploratória dos dados e criação de <i>insights</i>	p.55
7.4.1. Distribuição taxa de suicídio.....	p.57
7.4.2. Análise do Índice de Desenvolvimento Humano e PIB per capita em relação à taxa de suicídio.....	p.57
7.4.3. Análise do Índice de Desemprego per capita em relação à taxa de suicídio.....	p.61
7.4.4. Análise do Índice de Liberdade Humana em relação à taxa de suicídio.....	p.63
7.4.5. Análise da Taxa de Muçulmanos em relação à taxa de suicídio.....	p.65
7.4.6. Análise da Taxa de Ateus em relação à taxa de suicídio.....	p.67
7.4.7. Análise do Índice de Temperatura em relação à taxa de suicídio.....	p.69
7.4.8. Análise cruzada da temperatura média em relação aos outros índices.....	p.71
7.4.9. Análise cruzada da taxa de ateus em relação aos outros índices.....	p.74
7.4.10. Mapa de Análise de Risco.....	p.76
7.5. Discussão.....	p.78
8. Considerações Finais.....	p.84
8.1. Futuro da pesquisa (curto e médio prazo).....	p.86
9. Bibliografia.....	p.87
10. Anexo.....	p.98

Lista de Figuras

Figura 1- Gráfico do escopo sistemático da sinergia entre os ODS.....	p.14
---	------

Figura 2- Fluxograma do desenvolvimento econômico a partir da melhora social.....	p.15
Figura 3- Mapa de calor da correlação entre variáveis não religiosas.....	p.44
Figura 4- Mapa de calor da correlação entre variáveis religiosas.....	p.45
Figura 5- Mapa de calor da correlação entre as variáveis.....	p.47
Figura 6- Distribuição taxa de suicídio.....	p.56
Figura 7 - Distribuição de Desenvolvimento Humano (IDH).....	p.57
Figura 8 - Distribuição PIB per capita.....	p.57
Figura 9- Distribuição IDH em relação ao suicídio.....	p.58
Figura 10 - Distribuição PBI per capita em relação à Taxa de Suicídio.....	p.60
Figura 11- Distribuição Desemprego.....	p.61
Figura 12- Distribuição de Desemprego em relação à taxa de suicídio.....	p.62
Figura 13- Distribuição índice de Liberdade Humana (HFI).....	p.63
Figura 14- Distribuição Índice de Liberdade Humana em relação à taxa de Suicídio.....	p.64
Figura 15- Distribuição Taxa de Muçulmanos.....	p.65
Figura 16- Distribuição Taxa de Muçulmanos em relação à taxa de suicídio.....	p.66
Figura 17- Distribuição Taxa de Ateus.....	p.67
Figura 18- Distribuição Taxa de Ateus em relação à taxa de suicídio.....	p.68
Figura 19- Temperatura média.....	p.69
Figura 20- Distribuição Temperatura média em relação à taxa de suicídio.....	p.70
Figura 21- Distribuição Temperatura média em relação à PIB per capita.....	p.71
Figura 22- Distribuição Temperatura média em relação ao Índice de Desenvolvimento Humano.....	p.72
Figura 23- Distribuição Temperatura média em relação ao Índice de Liberdade Humana (HFI).....	p.73
Figura 24: Distribuição Taxa de Ateus em relação à Índice de Liberdade Humana.....	p.74
Figura 25: Distribuição Taxa de Ateus em relação ao PIB per capita.....	p.75
Figura 26: Distribuição Taxa de Ateus em relação ao Índice de Liberdade Humana.....	p.76
Figura 27: Mapa de risco.....	p.77
Figura 28: Relações temperatura com suicídio.....	p.81

Lista de Tabelas

Tabela 1 - Processos da Ciência da Informação.....	p.23
Tabela 2- Tabela de Aplicação do Método Hipotético-Dedutivo ao tema.....	p.25
Tabela 3 - Análise em escala local.....	p.27
Tabela 4- Análise em escala regional.....	p.28
Tabela 5- Análise em escala nacional.....	p.29
Tabela 6- Análise em escala global.....	p.30
Tabela 7 - Tipo de distribuição com W e <i>p-value</i>	p.40
Tabela 8 - Resultados teste Kruskal-Wallis.....	p.42
Tabela 9- Resultados método AIC linear, quadrático e cúbico.....	p.50
Tabela 10- Resultados AIC Normal e Ponderado.....	p.50
Tabela 11 - Resultados AIC, R^2 e <i>P-value</i>	p.51
Tabela 12 - Resultados método AIC linear, quadrático e cúbico para o cruzamento de temperatura média em relação aos três índices.....	p.52
Tabela 13 - Resultados método AIC linear, quadrático e cúbico para o cruzamento de taxa de ateus em relação aos três índices.....	p.52
Tabela 14 - Resultados AIC Normal e Ponderado para o cruzamento de temperatura média em relação aos três índices.....	p.52
Tabela 15 - Resultados AIC Normal e Ponderado para o cruzamento de taxa de ateus em relação aos três índices.....	p.52
Tabela 16 - Resultados R^2 e <i>P-value</i> cruzamento temperatura média	p.53
Tabela 17 - Resultados R^2 e <i>P-value</i> cruzamento taxa de ateus.....	p.53
Tabela 18 - Resultado inicial regressão de Ridge	p.54
Tabela 19 - c de Ridge.....	p.54

Resumo

O projeto SACI (*Suicide Analytics for Counter Incidence*) surgiu da motivação de realizar contribuições por meio da Inteligência Artificial para prevenir o suicídio. Para isso, partimos de um pressuposto que verificasse um possível vínculo entre variáveis demográficas pela sinergia entre Ciência da Informação e *Machine Learning*, permitindo a criação de sistemas de visualização capazes de gerar *insights* para promover a compreensão do fenômeno e a tomada de decisões. Realizamos a análise de microdados de 149 países, por meio de um método estatístico que utilizou: Regressões (linear, polinomial e de Ridge) e Testes de Correlação e Distribuição de Dados (Spearman, Kruskal-Wallis e Shapiro-Wilk) para analisar os valores e entender a conexão entre uma série de variáveis e o suicídio. Para 7 variáveis identificadas (IDH, HFI, PIB per capita, desemprego, taxa de ateus, taxa de muçulmanos e temperatura média), conforme o que a Organização Mundial enfoca em relação à abordagem transnacional do problema, obtivemos valores significativos: p valores $< 0,05$ em todos os testes; coeficientes na análise de Spearman próximos ou maiores que 0,3; e um R^2 igual a 0,41. Esses resultados corroboraram nossa hipótese sobre a conexão entre as variáveis e o suicídio. O projeto foi capaz de transitar entre análises globais e aplicações locais, respeitando parâmetros éticos e alcançando as seguintes contribuições: visualização das análises (mapas de calor, mapas-múndi, *scatter plots* e mapa de análise de risco) para facilitar a interpretação dos resultados e promover iniciativas da prevenção do suicídio; além da disponibilização dos dados e algoritmos produzidos, assim como do protocolo desenvolvido. Dessa maneira, o presente estudo aprofundou as aplicações relacionadas à análise de dados para a palição do suicídio, demonstrando a factibilidade do seu uso tanto para o aprofundamento do conhecimento do fenômeno quanto para o desenvolvimento de estratégias preventivas.

Palavras-chave: Prevenção ao Suicídio; Data Analytics; Machine Learning.

Abstract

The SACI (Suicide Analytics for Counter Incidence) project arose from the motivation to make contributions through Artificial Intelligence to prevent suicide. To do this, we started from the premise of verifying a possible link between demographic variables through the focus of Information Science and Machine Learning, enabling the creation of visualization systems capable of generating insights to promote understanding of the phenomenon and decision-making. We analyzed microdata from 149 countries in 2019, using a statistical method that used: Regressions (linear, polynomial and Ridge) and Correlation and Data Distribution Tests (Spearman, Kruskal-Wallis and Shapiro-Wilk) to analyze the values and understand the connection between a series of variables and suicide. For 7 variables identified (HDI, HFI, GDP per capita, unemployment, rate of atheists, rate of Muslims and average temperature), in line with what the World Organization focuses on in relation to the transnational approach to the problem, we obtained significant values: $p < 0.05$ in all tests; coefficients in Spearman's analysis close to or greater than 0.3; and an R^2 equal to 0.41. These results corroborated our hypothesis about the connection between the variables and suicide. The project was able to move between global analyses, respecting ethical parameters and achieving the following contributions: visualization of the analyses (heat maps, world maps, scatter plots and risk analysis maps) to facilitate interpretation of the results and promote suicide prevention initiatives; and making the data and algorithms produced available, as well as the protocol developed. In this way, this study has deepened the synergy of applications related to data analysis for suicide prevention, demonstrating the feasibility of using it both to deepen knowledge of the phenomenon and to develop preventive strategies.

Keywords: Suicide Prevention; Data Analytics; Machine Learning.

Zusammenfassung

Das Projekt SACI (Suicide Analytics for Counter Incidence) entstand aus der Motivation heraus, durch künstliche Intelligenz einen Beitrag zur Suizidprävention zu leisten. Zu diesem Zweck sind wir von der Prämisse ausgegangen, eine mögliche Verbindung zwischen demografischen Variablen durch den Fokus der Informationswissenschaft und des maschinellen Lernens zu verifizieren, was die Schaffung von Visualisierungssystemen ermöglicht, die in der Lage sind, Erkenntnisse zu generieren, um das Verständnis des Phänomens und die Entscheidungsfindung zu fördern. Wir haben Mikrodaten aus 149 Ländern im Jahr 2019 mit Hilfe einer statistischen Methode analysiert: Regressionen (linear, polynomial und Ridge) sowie Korrelations- und Datenverteilungstests (Spearman, Kruskal-Wallis und Shapiro-Wilk), um die Werte zu analysieren und den Zusammenhang zwischen einer Reihe von Variablen und Selbstmord zu verstehen. Für sieben identifizierte Variablen (HDI, HFI, Pro-Kopf-BIP, Arbeitslosigkeit, Anteil der Atheisten, Anteil der Muslime und Durchschnittstemperatur) erhielten wir in Übereinstimmung mit dem Ansatz der Weltorganisation für das grenzüberschreitende Problem signifikante Werte: $p < 0,05$ in allen Tests; Koeffizienten in der Spearman-Analyse nahe bei oder größer als 0,3; und ein R^2 von 0,41. Diese Ergebnisse bestätigten unsere Hypothese über den Zusammenhang zwischen den Variablen und dem Suizid. Das Projekt konnte unter Einhaltung der ethischen Parameter zwischen globalen Analysen wechseln und folgende Beiträge leisten: Visualisierung der Analysen (Heatmaps, Weltkarten, Streudiagramme und Karten zur Risikoanalyse), um die Interpretation der Ergebnisse zu erleichtern und Initiativen zur Suizidprävention zu fördern, und Bereitstellung der erstellten Daten und Algorithmen sowie des entwickelten Protokolls. Auf diese Weise hat diese Studie die Synergie von Anwendungen im Zusammenhang mit der Datenanalyse für die Suizidprävention vertieft und die Möglichkeit aufgezeigt, sie sowohl zur Vertiefung der Kenntnisse über das Phänomen als auch zur Entwicklung von Präventionsstrategien zu nutzen.

Schlüsselwörter: Suizidprävention; Datenanalyse; Maschinelles Lernen.

1. Introdução

Ao considerarmos as problemáticas da sociedade atual, uma questão que surge é a relacionada ao suicídio e às variáveis que se associam a esse fenômeno. É evidente, ao observar a nossa realidade, que o suicídio, exacerbado pela forma como a sociedade se relaciona e por outros fatores diversos, impacta um número significativo de pessoas, sendo a principal causa de morte entre jovens de 15 e 29 anos (OMS, 2021). Adicionalmente, a cada 40 segundos, uma pessoa se mata, ou seja, se o leitor demorar em média 40 minutos em ler nosso relatório, aproximadamente 60 pessoas haverão se suicidado no mundo nesse período (OMS, 2019). Considerando a gravidade do problema, nos indagando sobre sua origem e causa e, respaldados pela literatura consultada, afirmamos que a temática do suicídio é um desafio de escala global que afeta sociedades e indivíduos e que requer investimentos de ordens diversas para paliá-lo (OMS, 2019).

Consequentemente, buscamos criar sistemas de visualização que reúnam informações que se associam ao suicídio, tais como variáveis e causas. Para isto propomos desenvolver uma pesquisa que utiliza a programação e a estatística para criar um sistema capaz de analisar e produzir dados.

Desta maneira, o projeto, que denominamos SACI (Suicide Analytics for Counter Incidence) propõe a análise e produção de dados sobre o suicídio, para contra incidência, mediante a reunião de microdados (dados detalhados sobre indivíduos e suas características) e metadados (informações que descrevem e contextualizam esses dados) associados a essa problemática, focando na facilidade de sua compreensão a partir de uma visualização clara e didática.

Devido sua natureza complexa, o suicídio possui diversas definições, resultantes principalmente pela forma com que cada estudioso entende a sua motivação (JOINER, 2005). Entretanto, analisando a literatura profissional e buscando uma denominação que expresse objetividade e possibilite a análise do fenômeno e de suas motivações no contexto global, nos deparamos com a definição de Cassorla (2019) que define o problema como um ato voluntário realizado pelo próprio indivíduo de forma proposital, e a partir de um meio que ele acredita ser letal, motivado por algum fato ou acontecimento que o leva a não querer preservar a vida. Além do entendimento do termo diretamente, existem teorias que tentam explicar profundamente como surgem as motivações que levam uma pessoa a se suicidar. A partir de uma análise sociológica, Durkheim (2000) afirma, que o suicídio tem motivações individuais e psicológicas; e, mediante, o entendimento das diferentes características que englobam o fenômeno é possível compreender as variáveis que a ele se relacionam.

Assim percebemos que, a fim de prevenir o suicídio, precisamos analisá-lo para entendê-lo de forma sistemática e objetiva. Segundo Millner, Robinaugh e Nock (2020) a realização de investigações descritivas e rigorosas das causas e comportamentos do suicídio, ou seja, das variáveis envolvidas com esta situação, é fundamental para a sua compreensão e prevenção.

Certos projetos já apontam a incidência de variáveis que se referem ao indivíduo – como idade, gênero, traços de personalidade e transtornos mentais – e à sociedade e economia – como grupos de pertencimento social, étnico e religioso, estresse laboral, no primeiro caso; PIB, inflação e taxa de desemprego no segundo (MOTILLON-TOUDIC et al., 2022) – para o suicídio. Mas, apesar de pesquisas recentes apontarem a utilidade de compreender esses fatores para a prevenção da problemática, faltam ainda estudos mais precisos e extensos sobre esses dados que permitam obter *insights* verdadeiramente úteis (TURECKI et al., 2019).

Neste trabalho, considerando a forma com que o suicídio é estudado por especialistas, como Durkheim (2000) e Joiner (2005), das conclusões que obtivemos das várias literaturas citadas no decorrer do projeto e da análise das bases de dados escolhidas, consideramos que é fundamental entender a correlação entre as variáveis: religião, temperatura, *human freedom index* (HFI), Produto Interno Bruto (PIB) per capita, inflação, desemprego e índice de Desenvolvimento Humano (IDH) com o suicídio (por 100 mil habitantes). Para alcançar esse objetivo, utilizaremos da ciência da informação, mais especificamente a ciência de dados, para analisar e classificar os microdados e metadados obtidos, tirando proveito das técnicas de aprendizado de máquina e ferramentas estatísticas: *Machine learning* supervisionado (Regressão polinomial, linear e de Ridge), teste de correlação Spearman, teste de Shapiro-Wilk e o teste de Kruskal-Wallis (OLIVEIRA, 2013; FREEDMAN, PISANI, PURVES, 2007; ZAR, 2005). Essas técnicas e ferramentas serão utilizadas para a análise prévia das variáveis e ampliação dos *insights* obtidos. Essas análises permitirão a identificação de padrões vinculados ao suicídio e fatores de risco, possibilitando a criação de formas de visualização mais completas e abrangentes (MEJIA, PIVATTO, 2023), como gráficos e mapas, possibilitando assim o compartilhamento desses dados de forma intuitiva e precisa (STACK, 2021).

2. Justificativa

A importância desta investigação é demonstrada a partir de observações do nosso entorno (escola, amigos, família) e por notícias do cotidiano. Assim mesmo, pela literatura profissional analisada, como a teoria expressada no clássico livro de Émile Durkheim, *O suicídio*. Estudo sociológico (2000), publicado por primeira vez em 1897, e por estudos recentes divulgados pela PubMed e pela Organização Mundial da Saúde (OMS) sobre o suicídio e suas tendências, que demonstram que aproximadamente 700 mil pessoas morrem anualmente por essa causa no mundo, sendo, como já citado, a quarta maior causa de óbito entre jovens de 15 a 29 anos. Esses dados destacam a urgência de estudos e medidas para combater o problema, sendo fundamental o reconhecimento causal do suicídio para preveni-lo. Porém, a dificuldade em realizar estudos que analisem o fenômeno de forma direta torna o descobrimento de causas um desafio, isto ao considerar a falta de proximidade entre quem estuda e quem é estudado (FRANKLIN, HUANG, BASTIDAS, 2019). Dessa forma, nos questionamos como encontrar dados ou tendências correlacionadas ao suicídio em escala mundial para criar intervenções eficazes para contra restar a problemática (MOTILLON-TOUDIC et al., 2022).

Em nosso estudo, constatamos que a análise de informações como a idade, o gênero, os transtornos mentais, o isolamento social e outras características individuais são variáveis importantes no estudo do fenômeno, já que auxiliam no entendimento do sujeito e as suas motivações. Entretanto, o estudo de valores relacionados apenas ao indivíduo, como as mencionadas anteriormente, além de já terem sido profundamente consideradas, é insuficiente para satisfazer análises completas que contribuam com a identificação de padrões relacionados ao suicídio.

Como mencionado por Durkheim (2000), existem quatro tipos de suicídio, que descreveremos posteriormente, que estão inteiramente relacionados a sociedade e a forma com que os indivíduos se relacionam. Portanto, é necessário entender o desenvolvimento da população como grupo social e as suas características fundamentais e evolutivas para a análise de dados relacionados ao fenômeno que estudamos (LEWIS, 2005). Nesse sentido, consideramos a importância de incorporar dados como religião e *human freedom index* (HFI) em nossa análise, assim como a variável ambiental temperatura média que incide no comportamento dos grupos humanos.

Por outro lado, os governos no geral, como reguladores da ordem social, por exemplo, têm papéis significativos nas taxas suicidas, como indicado no artigo “*Government spending, recession, and suicide: evidence from Japan*” (MATSUBAYASHI, SEKIJIMA, UEDA, 2020). Como mostrado nesse estudo, no Japão, o aumento de 1 % do PIB esteve relacionado com a diminuição de 0,2% dos casos de suicídio no país, portanto a ação do governo em relação à economia ou às políticas públicas, que influenciam indicadores como Produto Bruto Interno (PIB) per capita, inflação, desemprego e índice de Desenvolvimento Humano (IDH), pode exacerbar ou mitigar o impacto das recessões que

incidem diretamente no suicídio, percepção que também é defendida por Chen, Choi e Sawada (2008).

Apesar da existência de pesquisas que tentam identificar características e fatores de risco relacionados ao fenômeno, como os de Mitchel et al. (2021) e de Sinyor, Tse e Pirkis (2017) faltam projetos que consigam reunir as variáveis sociais, culturais e econômicas para uma interpretação que permita combater o suicídio efetivamente (TURECKI et al., 2019) e é nesse sentido que o nosso projeto pretende colaborar.

Adicionalmente, considerando o intuito de compartilhar os resultados disponíveis e direcionar esforços na prevenção do suicídio, nos focamos na criação de *insights* visuais, mais especificamente *scatter plots*, mapas-múndi e mapa de risco para facilitar a compreensão da problemática e tornar nossa pesquisa mais concreta.

Considerando a importância da análise de dados na prevenção do suicídio (BERNERT et al., 2020) e tendo em mente a criação de ferramentas que reúnam essas informações de modo prático, fazendo que delas possam ser extraídas percepções úteis, como as citadas anteriormente, e cumprindo os objetivos de um projeto de ciência da informação, de analisar, classificar e compartilhar os *insights* obtidos, consideramos utilizar os seguintes processos fundamentais: *Machine learning* supervisionado (regressões lineal, polinomial e de Bridge) e testes de correlação e distribuição de dados (teste de correlação de Sperman, o teste de Shapiro-Wilk e o teste de Kruskal-Wallis). A capacidade da inteligência artificial de analisar e classificar bases de dados volumosas rapidamente, identificar tendências entre os diversos dados e fazer previsões precisas, permite a aplicação dessa tecnologia para uma melhora social, em áreas como educação, saúde, agricultura e gestão de recursos (LEE, K, 2018), pelo fato de possibilitar a criação de padrões e de facilitar tarefas. O método estatístico também ajudará nessa análise, exemplificando a relação entre as variáveis e auxiliando na produção e compartilhamento dos dados (PEARCE, 1992). Assim, a qualidade e rapidez no processamento de informações, nos possibilitará a confecção de um projeto mais detalhado e prático, garantindo as etapas fundamentais de um projeto embasado na metodologia da ciência da informação já citadas.

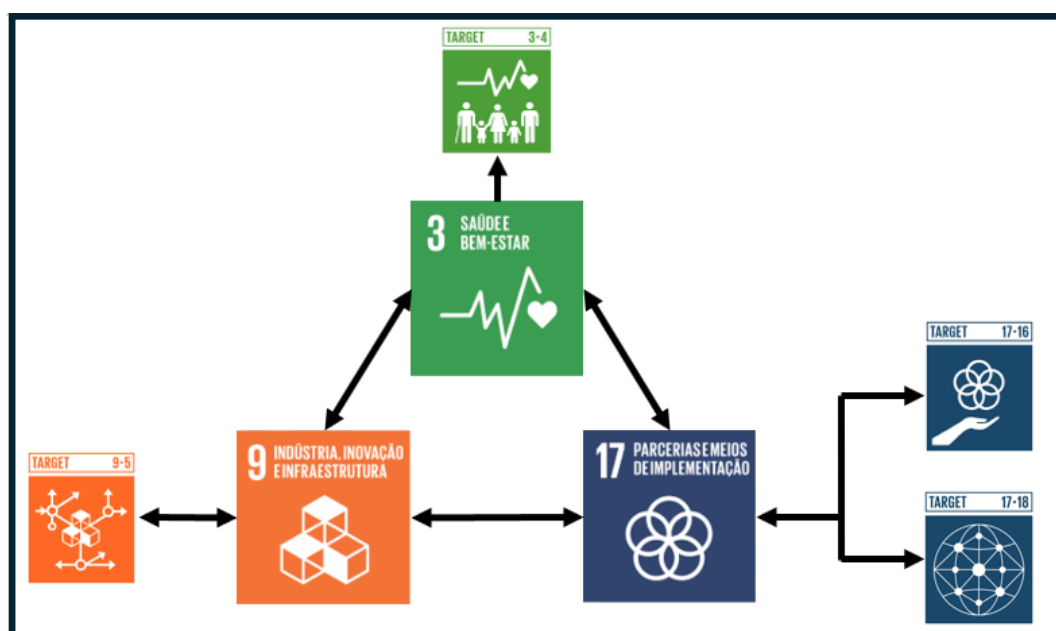
Reconhecemos que para resolver de forma factível o problema do suicídio, a consideração de metas como a dos Objetivos de Desenvolvimento Sustentável¹ (ODS) é fundamental, já que elas baseiam e direcionam ações mitigatórias em áreas como sustentabilidade, saúde e sociedade. Além disso, o cumprimento dos objetivos da terceira ODS para a Área de Saúde e Bem-estar², a partir do

¹ Os objetivos de Desenvolvimento Sustentável (ODS) formam parte de uma agenda adotada em 2015 durante a Cúpula das Nações para o Desenvolvimento Sustentável. Essa agenda deve ser cumprida até 2030.

² A meta indica “assegurar uma vida saudável e promover o bem-estar para todos, em todas as idades”.

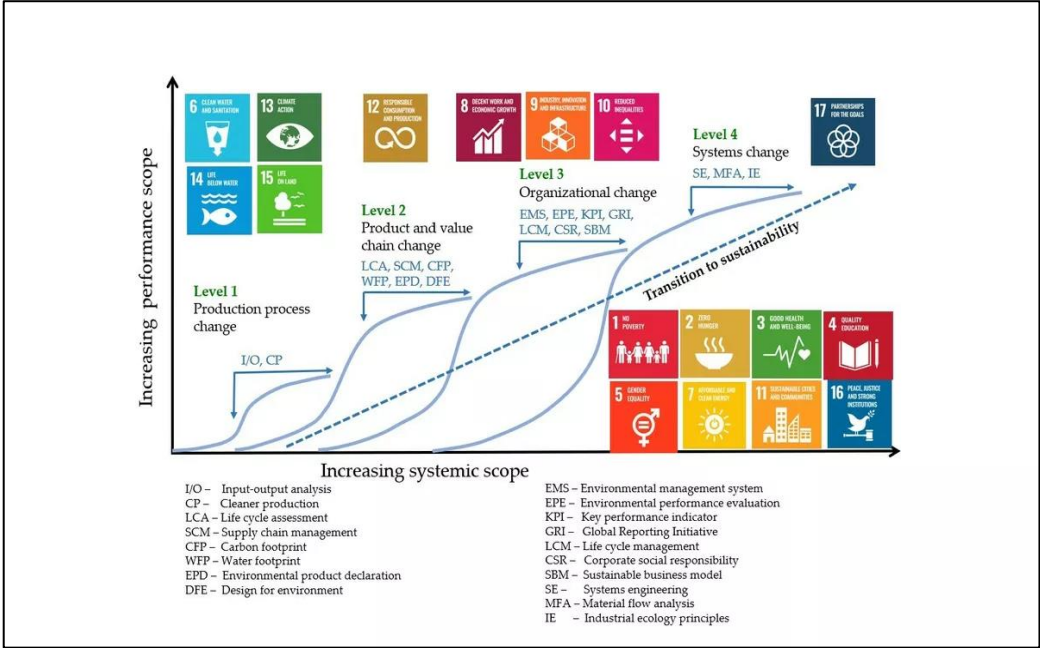
nosso projeto, teria impactos positivos em outras metas, já que a saúde, área diretamente vinculada à problemática que estudamos, conforme é destacado pelo esquema, (Figura 1) é a base para resoluções econômicas em áreas como indústria, inovação, infraestrutura, parcerias e meios de implementação aplicadas à sociedade e sua vida digna (ITAMARATY, 2024). Como mostrado pelo gráfico abaixo, o desenvolvimento econômico vá da mão da melhora social e permite a extensão das ações e iniciativas realizadas em busca da sustentabilidade e qualidade de vida. Alcançar as ODS em todos os tipos de performance, possibilita mudanças sociais e econômicas, evidenciando a importância de transformação em qualquer setor, a fim da melhora social. (Figura 2).

Figura 1: Gráfico do escopo sistemático da sinergia entre os ODS



Fonte: ITAMARATY (2024). Organização: Autores.

Figura 2: Fluxograma do desenvolvimento econômico a partir da melhora social



Fonte: FET, KNUDSON, KEITSCH (2021). Organização: Autores.

3. Hipótese e Problemática

3.1 Hipótese

Apoiados nas observações que motivaram nosso trabalho, a hipótese deste estudo se expressa em duas perguntas:

1) Podemos utilizar *Machine learning* supervisionado (Regressão polinomial, linear e de Ridge), o teste de correlação de Spearman, o teste de Shapiro-Wilk e o teste de Kruskal- Wallis (PAIXÃO et al., 2022; BRUCE, BRUCE, GEDECK, 2020) para identificar a relação individual e coletiva entre as variáveis socioeconômicas e ambiental selecionadas e os índices do suicídio nos dados analisados?

2) Caso a identificação dessa relação for possível, será viável, a partir dessas informações, extrair, tratar e compartilhar dados importantes para a prevenção do suicídio em escala global?

3.2 Problemática

A problemática deste estudo, baseada nos conceitos de relevância e viabilidade discutidos por Gil (2002), reside na complexidade de compreender e analisar dados para a prevenção do suicídio, principalmente por limitações relacionadas à disposição dos dados e à literatura profissional que permeia o tema, tornando claro a necessidade de novas medidas para preveni-lo a partir da tecnologia (KREUZE. et al., 2017). Definimos assim a problemática do estudo: Como extrair dados úteis provenientes de fontes de informação a partir do *machine learning* e de métodos estatísticos para a prevenção do suicídio globalmente.

Identificamos o problema de acordo com os conceitos de viabilidade, relevância, novidade e oportunidade, seguindo abordagem hipotético-dedutiva baseada em Lakatos e Marconi (1992). Este problema específico é determinado por métodos analíticos e teóricos, levando em consideração a possibilidade de pesquisa e sua relevância para a existência humana. Além disso, a problemática do suicídio, em escala global, é destacada pela OMS (2019), ressaltando a importância de abordar esse tema.

4. Objetivo

Este estudo tem como meta principal analisar dados demográficos, econômicos e ambientais coletados de bancos de dados globais para identificar índices de correlação significantes em comparação com às taxas de suicídios pelo mundo. Mediante à esta análise, caso corroborada, desenvolveremos formas de visualizações com natureza gráfica (gráficos e mapas) que permitam facilitar o entendimento do suicídio. Desta maneira, queremos possibilitar a exposição da análise dos dados de forma clara e acessível que permitam servir de base para decisões estratégicas para a prevenção do suicídio perante as iniciativas de prevenção global.

Em linhas gerais trata-se de:

- Gerar informações mediante análises estatísticas via *Machine Learning* dos bancos de dados analisados.
- Estruturar as correlações encontradas e compiladas em visualizações gráficas (gráficos e mapas) que permitam um entendimento da problemática estudada de forma clara e acessível para visualização, interpretação e tomada de decisão relacionada à prevenção de suicídio em escala global.

5. Objetos de estudo e revisão bibliográfica

5.1. A compreensão do que é Suicídio

Devido a sua natureza complexa, e as dificuldades em estudá-lo, o suicídio não possui uma definição universalmente aceita, pois, como já mencionado, cada teoria que o aborda defende motivações diferentes para a sua realização (JOINER, 2005). Considerando o objetivo de nossa pesquisa, de analisar questões culturais e socioeconômicas relacionadas ao fenômeno, selecionamos pensamentos que buscam compreender as tendências suicidas socialmente e de forma objetiva. Além de identificar estudos que indiquem possíveis fatores de risco do suicídio.

Como já mencionado, o suicídio pode ser definido fundamentalmente como um ato deliberado executado pelo próprio indivíduo de forma consciente e intencional, cuja finalidade é a própria morte (CASSORLA, 2019).

Conforme Durkheim (2000), considerando o aspecto social do fenômeno, suas deliberações e origem, o suicídio pode ser categorizado em 4 modalidades, além de em outras cinco subcategorias. As categorias são:

- Suicídio Egoísta: ele acontece quando um indivíduo se sente desconectado ou isolado da sociedade. A falta de integração social e de apoio podem levar o ser humano a ter sentimentos de desesperança, eventualmente levando-o a escolher a morte.

- Suicídio Altruísta: ocorre em uma sociedade com alta taxa de integração, levando o sujeito achar que o certo é se suicidar para ajudar o seu grupo ou sociedade, exemplos incluem suicídios por motivos de honra como os dos Kamikazes.

- Suicídio Anômico: é relacionado a mudanças drásticas sociais ou econômicas, onde normas ou valores sociais são abalados, levando o sujeito a uma falta de propósito vital, esse tipo de suicídio geralmente acontece durante grandes crises econômicas ou troca de governos.

- Suicídio Fatalista: se trata de quando um indivíduo sofre opressão extrema onde acaba perdendo as esperanças de sair de tal opressão e acaba escolhendo tirar sua vida, geralmente ocorre em contextos de governos totalitários ou com prisioneiros de guerra.

As subcategorias:

- Suicídio Assistido: ele acontece quando um indivíduo ajuda uma pessoa a cometer suicídio, fornecendo-lhe os meios ou informações necessárias para que ela mesma possa realizar o ato de se suicidar. Mesmo sendo um tema controverso em diversos países, como a Suíça, é legalmente permitido.

- Suicídio Medicamentoso Assistido: ocorre quando um médico usa uma substância letal para causar a morte de uma pessoa, geralmente aludindo motivos de compaixão como evitar a dor crônica, nesse caso a pessoa que deseja se matar tem que estar de acordo com o procedimento.

- Suicídio Impulsivo: caracteriza-se por ser uma reação do sujeito sem ter um planejamento prévio para se suicidar. Pode ser causado por uma emoção que ele não consegue controlar, levando-o a se suicidar logo depois de sofrer com os sentimentos, geralmente acontece em traições ou alguma descoberta chocante para a vítima.

- Suicídio Crônico: ocorre de forma lenta e continua, onde o sujeito tem hábitos autodestrutivos e não consegue ou não quer mudá-los, geralmente ocorre por uso de substâncias químicas prejudiciais.

- Suicídio por Protesto: ocorre quando um indivíduo se suicida como um ato de crítica contra uma injustiça ou para chamar a atenção para uma causa específica, como forma de protesto.

Durkheim (2000) também explica o que motiva os indivíduos a terem tendências suicidas, assim como Joiner (2005) afirma que o suicídio se concretiza quando três fatores se combinam:

- Sentimento de ser um fardo;
- Sentimento de isolamento e estar desconexo;
- Capacidade de realizar o suicídio.

Além do entendimento de origem e determinações das tendências suicidas, a percepção dos fatores relacionados ao fenômeno, como os socioeconômicos é importante. O estudo de variáveis socioeconômicas, entre outras, é fundamental para o seu entendimento e para a realização de projetos, à medida que direcionam os esforços (MCLOUGHLIN, GOULD, MALONE, 2015).

A análise de diferentes estudos, que buscam definições e motivações diferentes para o fenômeno nos possibilita compreendê-lo na sua complexidade, à medida que descobrimos suas diversas nuances, promovendo uma visão mais abrangente em relação ao problema, e impactando assim, positivamente o nosso processo de cruzamento de dados e a busca por correlações.

5.2. Entendimento do que é pesquisa em Inteligência Artificial (IA)

A pesquisa em inteligência artificial, envolve a criação de *softwares*, que se fossem feitos por seres humanos necessitariam de colaboração entre muitos indivíduos e de bastante tempo para serem realizados. A IA, e mais especificamente o Aprendizado de máquina ou *Machine learning*, tornam a análise e interpretação de dados muito mais eficaz e rápida, auxiliando na confecção de projetos que

lidam com grandes entradas de informação (RUSSELL, NORVIG, 2010). Dentre os métodos de IA mais utilizados, estão o aprendizado supervisionado e não supervisionado, redes neurais artificiais, métodos de lógica e algoritmos genéticos (GOODFELLOW, BENGIO; COURVILLE, 2016). Na nossa pesquisa, devido a presença de dados de entrada e saída já rotulados, optamos pela utilização de um aprendizado de máquina supervisionado, mais especificamente a regressão de Ridge, que possibilita análises de correlação e previsões precisas (BISHOP, 2006). A regressão é um modelo estatístico que modela a relação entre uma ou várias variáveis dependentes e uma variável independente.

A consideração de questões como memória, viés de servidor, limpeza de dados e a análise correta de dados são fatores fundamentais na criação de modelos de inteligência artificial que necessitam acompanhar o desenvolvimento do *software* a fim de assegurar uma análise correta. (DALLY et al., 2003).

5.3. As variáveis, sua compreensão e importância para uma análise legítima e o porquê da sua utilização

Variáveis são características ou propriedades que podem ser medidas ou observadas e que variam entre objetos (pessoas, animais), esses dados são fundamentais em pesquisas, pois possibilitam descrever e entender fenômenos, a partir da estatística, que está inteiramente baseada em valores (FIELD, 2017). Dentro do ramo da estatística, as variáveis podem se relacionar de diferentes maneiras (CRESWELL, 2015):

- **Influência direta:** Quando dois valores estão diretamente relacionados, tendo tendências opostas ou congruentes.
- **Mediação:** Quando uma variável explica a relação entre outras duas variáveis, ou seja, são mediadoras entre valores dependentes e independentes.
- **Moderação:** Quando a variável impacta a relação entre outras duas variáveis.

No entendimento do suicídio utilizaremos dados que mostram as variáveis em escala global com o objetivo de poder compreender melhor e mais facilmente o fenômeno. Esses dados constam nas bases de dados que estudamos. As variáveis globais são de aspecto mais amplo, englobam comunidades, organizações e países, entre elas podemos considerar tendências econômicas, fluxos migratórios e mudanças demográficas, entre outras informações.

Embora a compreensão e utilização de variáveis individuais seja fundamental para compreender a natureza complexa do suicídio, já existem estudos que consideram profundamente

essas variáveis individuais, sendo que a forma de metrificação desses dados é diretamente relacionada ao suicídio. Como consequência disso, as variáveis individuais não serão foco da nossa análise, sim são consideradas as variáveis demográficas, sociais, econômicas e natural já citadas.

5.4. As características da Pesquisa em estatística

A pesquisa em estatística é um processo sistemático que envolve quatro fatores principais: coleta, análise, interpretação e apresentação de dados quantitativos ou qualitativos. O objetivo principal desse processo é descobrir padrões, tendências e correlações entre variáveis, permitindo a realização de pesquisas em diversas áreas, como saúde, economia, ciências sociais, entre outras (LANE et al., 2003). Cada uma dessas etapas desempenha um papel crucial na obtenção de resultados válidos e significativos.

5.5. Entendimento do que é pesquisa em ciência da informação

Com o objetivo de transformar metadados em variáveis úteis, buscamos compreender o campo da ciência da informação, para definir assim os processos que envolveram esta pesquisa.

Os processos da ciência da informação são: Coleta, armazenamento e administração, reconstrução e análise, interpretação e difusão dos dados. O objetivo é descobrir como as informações são geradas, organizadas e usadas. Isso permitirá a realização correta do projeto, envolvendo todos os processos fundamentais (OLIVEIRA, 2013). As definições dos processos são:

- Coleta de dados: para garantir precisão e relevância, processos de coleta de dados como entrevistas e análise documental devem ser realizados com cuidado. Para chegar a conclusões confiáveis, a qualidade dos dados é essencial. A escolha da técnica de coleta depende do objetivo da pesquisa e da natureza dos dados necessários (CRESWELL, 2015).

A qualidade dos dados coletados é a base para validar as conclusões da pesquisa. Por isso, é necessário garantir que os dados sejam representativos da população estudada e que sejam coletados de maneira imparcial, minimizando vieses ou tendências determinadas por forças externas. A seleção de uma amostra adequada, o treinamento dos coletores de dados e a utilização de instrumentos de coleta validados são passos essenciais nesse processo (HAIR et al., 2010);

- Administração e Armazenamento: após a coleta, as informações devem ser organizadas e armazenadas para facilitar a sua recuperação. Isso inclui a catalogação e classificação de informações em sistemas eficientes;

- Reconstrução e análise: a recuperação localiza os dados e ao transformá-los em *insights* valiosos, a análise permite a identificação de padrões e correlações.

- Interpretação e Difusão: a interpretação entende as informações analisadas. A disseminação promove a comunicação e a aplicação do conhecimento obtido.

Tabela 1. Processos da Ciência da Informação

Processo	Descritivo
Coleta de dados:	Para garantir precisão e relevância, processos de coleta de dados como entrevistas e análise documental devem ser realizados com cuidado. Para chegar a conclusões confiáveis, a qualidade dos dados é essencial
Administração e Armazenamento	Após a coleta, as informações devem ser organizadas e armazenadas para facilitar a recuperação. Isso inclui a catalogação e classificação de informações em sistemas eficientes.
Reconstrução e análise	A recuperação localiza os dados. Ao transformar os dados em <i>insights</i> valiosos, a análise permite a identificação de padrões e correlações.
Interpretação e Difusão	A interpretação entende as informações analisadas. A disseminação promove a comunicação e a aplicação do conhecimento obtido.

Fonte: CRESWELL (2015). Organização: Autores.

Os processos antes descritos refletem a natureza interdisciplinar da ciência da informação, que combina áreas como biblioteconomia e comunicação para resolver os problemas de informação contemporâneos. (CRESWELL, 2015).

6. Métodos e Técnicas

O foco principal de nosso estudo foi investigar a existência de dados sociais relacionados ao suicídio, a partir do cruzamento das variáveis independentes com esse fenômeno, assim como o do cruzamento entre elas, para ajudar na criação de parâmetros mais objetivos e corretos para análise de tendências suicidas.

Com esse intuito o método, os processos, as ferramentas e técnicas utilizadas durante a execução foram cuidadosamente escolhidos e planejados. Esse rigor foi essencial para avaliar e identificar as abordagens mais eficazes, garantindo a integridade, a validade e a relevância dos resultados obtidos. Assim mesmo, consideramos fundamental realizar algumas considerações em relação a critérios éticos aplicados, delimitações de escala e limitações do nosso trabalho.

6.1. Método

O nosso estudo visa compreender de forma sistemática a relação entre diversas variáveis independentes (religião, temperatura, *human freedom index* (HFI), produto bruto interno (PIB) per capita, inflação, desemprego e índice de desenvolvimento humano (IDH)) com o suicídio, utilizando técnicas de inteligência artificial e estatística, mais especificamente a regressão polinomial, linear (PAULA, 2004) e de Ridge (KUTNER et al., 2005), o coeficiente de Spearman, o teste de Kruskal-Wallis e de Shapiro-Wilk (Bruce; Bruce, 2020). Dessa maneira, buscamos compreender, relacionar e identificar variáveis atreladas as tendências suicidas para assim formular sistemas de visualizações, mais especificamente *scatter plots*, mapas de calor, mapas-múndi e um mapa de análise de risco, capazes de direcionar ações mitigatórias.

Em nossa pesquisa utilizamos uma abordagem hipotético-dedutiva, fundamentada na premissa metodológica do filósofo Popper (2002). Essa abordagem orientou nossa investigação, na medida em que nos desafiava a identificar lacunas entre o problema inicialmente formulado e a solução desejada. De início, formulamos hipóteses e premissas que conduziram nossa pesquisa. Essas hipóteses serviram como partida para a coleta de dados e análise. Em função dessa razão optamos por esta forma de abordagem, pela clareza do problema observado e a necessidade da solução encontrada.

Tabela 2: Tabela de Aplicação do Método Hipotético-Dedutivo ao tema.

Método Hipotético-Dedutivo	Aplicação na Pesquisa
1. Observação	Identificamos a problemática do suicídio e as suas possíveis causas fundamentais
2. Formulação de Hipótese	Levantamos a seguinte hipótese: Será possível criar um modelo de regressão polinomial junto ao coeficiente de Spearman capaz de identificar as principais variáveis relacionadas ao suicídio a fim de mitigar as tendências suicidas
3. Previsões Testáveis	Baseado na literatura e nos bancos de dados, criar um algoritmo que possa reconhecer padrões entre o suicídio e as variáveis analisadas.
4. Realização de Testes Empíricos	Coletar dados por meio de testes em ambiente computacional.
5. Refutabilidade	Testes de refutabilidade mediante ao falseamento de dados gerados pelo modelo, como comparações com outras pesquisas e métodos estatísticos para análise de precisão.
6. Análise Crítica	Análise de resultados dos testes para determinar se as relações e previsões são suficientes ou insuficientes.
7. Iteração e Desenvolvimento da Teoria	Com base nos resultados, revisar e desenvolver a teoria, repetindo o processo conforme necessário. Desta maneira se busca aprimorar o modelo e a aplicação dos dados.

Fonte: Autores.

Estruturando a pesquisa, ainda em Lakatos e Marconi (1992), temos itens que observamos e levamos em conta para a nossa abordagem de trabalho:

- Técnica: observação de fenômenos de forma direta e intensiva: pois, devido a fase inicial da pesquisa, focamos em trabalhar com o cruzamento de dados
- Observação: Análise dos dados obtidos para sua melhor seleção e avaliação dos dados
- Documentação Direta: os dados que foram produzidos estão associados ao ambiente computacional e as bases de dados utilizadas. Logo, neste primeiro momento os testes foram realizados em um ambiente totalmente controlado.

6.2. Considerações sobre a pesquisa

Considerando a sensibilidade do tema abordado, e planejamento necessários a fim de garantir a qualidade do projeto, este segmento do método prevê a análise e consideração de fenômenos e problemas pertinentes à pesquisa, tais como as considerações éticas e a delimitação da escala de abordagem.

6.2.1 Considerações da Ética em pesquisa analisando suicídio

A realização de estudos relacionados ao suicídio exige rigores éticos e morais devido a sua natureza sensível (WORLD HEALTH ORGANIZATION, 2014). A complexidade de conduzir estudos diretamente com indivíduos pelas limitações, nos motivou a encontrar soluções criativas para combater o suicídio sem enfrentar problemas de diretrizes. Assim, decidimos utilizar dados globais com variáveis sociais e econômicas relacionados a população. Apesar de uma maior facilidade ao tratar dados, é necessário considerar pontos para conduzir uma pesquisa humana e com moralidade (WORLD HEALTH ORGANIZATION, op. cit.). Tendo em consideração os problemas relacionados a este problema tão complexo, é necessário seguir e considerar alguns fatores para a realização de um projeto digno:

- Minimização de danos: As análises e afirmações geradas devem ser apresentadas de forma a não serem mal interpretadas e ou incentivar o ato deliberado do suicídio (AMERICAN PSYCHOLOGICAL ASSOCIATION, 2017).
- Revisão ética: Projetos que tratam de temas sensíveis e potencialmente perigosos, devem ser revisados por profissionais para assegurar que siga os diretrizes (CONSELHO NACIONAL DE SAÚDE, 2016).

6.2.2. Delimitação da área de estudo (possibilidades e limites em relação à escala de abordagem)

Para abordar de forma mais eficaz a complexidade do suicídio e as suas diversas causas, é necessário explorar diferentes escalas de investigação que nos permitam entender mais amplamente o fenômeno e como preveni-lo, possibilitando perspectivas mais complexas que auxiliam para o desenvolvimento de estratégias de prevenção mais robustas e eficazes (BEAUTRAIS, 2000). Assim, após a análise de estudos em diferentes escalas de atuação selecionamos as principais escalas de estudo:

Tabela 3: Análise em escala local

Escala	Prós	Contras
--------	------	---------

Local	- Informações detalhadas sobre cada variável estudada permitem uma análise aprofundada e específica, ajudando a identificar com precisão os grupos de risco e a personalizar as intervenções	- Dificuldade na generalização dos dados para além do contexto local pode limitar a aplicabilidade de descobertas e melhores práticas em outras áreas geográficas e culturais
-	- Possibilidade de realizar intervenções diretas e personalizadas com base em dados locais atualizados	- Quantidade limitada de dados disponíveis pode resultar em análises menos abrangentes e na não identificação de tendências emergentes
-	- Entendimento detalhada dos fatores de risco relacionadas ao indivíduo e a comunidade (Beautrais, 2000)	- Os possíveis recursos limitados dificultam o cruzamento e entendimento dos dados (Gunnell & Frankel, 1994)
-	- Aumento o engajamento da comunidade na prevenção do suicídio	- Possibilidade de resistências ou estigma em pequenas comunidades

Fonte: RODRIK, et al. (2011). BEAUTRAIS (2000). GUNNELL, FRANKEL (1994). Organização: Autores.

6.2.2.1. Análise do contexto local a partir da tabela:

A análise em nível local apresenta vantagens em relação ao detalhamento da análise, incluindo a obtenção de informações específicas sobre cada variável investigada, o que permite uma análise completa e direcionada. Isso aprimora a identificação precisa das populações em risco e a personalização das intervenções, aumentando assim a importância das medidas preventivas. Além disso, permite a implementação de intervenções diretas informadas por dados locais atuais, promovendo uma compreensão profunda dos fatores de risco pertinentes tanto aos indivíduos quanto à comunidade (BEAUTRAIS, 2000). Outro aspecto favorável é o maior envolvimento da comunidade nas iniciativas de prevenção do suicídio. No entanto, a análise localizada também encontra desvantagens, como o desafio de generalizar as descobertas além do local específico, o que restringe a aplicabilidade dos resultados em contextos alternativos. A escassez de dados pode levar a análises menos completas, complicando a tarefa de reconhecer padrões emergentes. Além disso, recursos limitados podem impedir a referência cruzada e a interpretação dos dados (GUNNELLL, FRANKEL, 1994). Além disso, em comunidades menores, pode existir resistência ou estigma associado às intervenções propostas.

Tabela 4: Análise em escala regional

Escala	Prós	Contra
Regional	- Identificação de tendências e padrões em áreas específicas	- Diversidade de dados entre regiões resultaria em uma dificuldade no estabelecimento de políticas públicas uniformes
-	- Permite a comparação entre diferentes regiões para melhores o entendimento causal	- Desigualdade de recursos entre diferentes regiões
-	- Maior volume de dados possíveis para serem analisados em relação ao nível local, aumentando a precisão das análises	- Necessidade de coordenação entre múltiplos governos locais ou municipais
-	- Desenvolvimento de políticas públicas mais abrangentes e precisas	Evitar vieses e generalizações erradas.

Fonte: GUNNELL (2005). HEGERL et al. (2006). HEGERL et al. (2008). Organização: Autores.

6.2.2.2. Análise do contexto regional a partir da tabela:

Analisar em escala regional, como mostrado na Tabela 2, amplia a compreensão dos padrões e tendências do suicídio, sendo essencial para a criação de políticas públicas mais específicas e abrangentes, adaptáveis a diferentes contextos. Estudos realizados por Hawton, Zahl e Weatherall (2003) apontam que essa perspectiva ajuda a identificar variações significativas nas taxas de suicídio entre diversas áreas, destacando fatores determinantes dessas discrepâncias. Dessa maneira, torna-se viável implementar intervenções mais eficazes, considerando as particularidades socioeconômicas, culturais e ambientais de cada localidade. Contudo, um desafio relevante consiste na necessidade de coordenação entre entidades governamentais locais e estaduais para assegurar uma resposta coesa e alinhada às demandas identificadas.

Tabela 5: Análise em escala nacional

Escala	Prós	Contra
Nacional	- Capacidade de implementar políticas de grande escala e abrangentes	- Pode ignorar diferenças importantes entre regiões e comunidades dentro do país

-	- Disponibilidade de recursos nacionais para iniciativas de prevenção	- Dificuldade em adaptar políticas nacionais para contextos locais específicos
-	- Dados estatísticos e epidemiológicos mais amplos	- Burocracia pode atrasar a implementação de políticas
-	- Possibilidade da realização de campanhas de largo alcance	- Pode haver diferença significativas entre os dados coletados em diferentes regiões

Fonte: MANN et al. (2005). WYMAN et al. (2010). HEGERL et al. (2008). Organização: Autores.

6.2.2.3. Análise do contexto nacional a partir da tabela:

No nível nacional, a análise do suicídio possibilita reunir dados de diferentes partes do país, a fim de criar políticas públicas abrangentes. Essa abordagem de estudo permite uma compreensão detalhada dos fatores de risco e das necessidades de prevenção em todo o país, possibilitando uma melhor distribuição dos recursos e esforços na criação de políticas públicas e de ações por parte de profissionais da saúde. Entretanto, um problema significativo está na necessidade em adaptar políticas nacionais para contextos mais específicos para garantir a eficácia e relevância das intervenções em todas as regiões do país (Tabela 4).

Tabela 6: Análise em escala global

Escala	Prós	Contra
Global	- Identificação de tendências e padrões globais	- Diferenças culturais e socioeconômicas tornam difícil a criação e alteração de políticas gerais.
-	- Compartilhamento de melhores práticas e recursos entre países	- Dados podem ser inconsistentes ou incompletos em diferentes países
-	- Colaboração internacional em pesquisa e desenvolvimento de políticas	- Dificuldade em coordenar e dividir ações entre os diferentes países e organizações internacionais
-	- Capacidade de abordar questões transnacionais, como crises econômicas e pandemias	- Pode haver desafios logísticos e políticos na implementação de estratégias globais

Fonte: ALICANDRO et al. (2019). ARENSMAN et al. (2020). BEAUTRAIS (2000). GUNNELL, FRANKEL (1994). Organização: Autores.

6.2.2.4. Análise do contexto global a partir da tabela:

A análise em global permite uma visão ampla sobre o suicídio, classificando-o como um fenômeno que ultrapassa fronteiras. Isso facilita a troca de melhores práticas, recursos e estratégias entre países, permitindo assim a criação de melhores políticas ao redor do mundo. Por meio da colaboração internacional em pesquisas e no desenvolvimento de políticas, torna-se viável adaptar e implementar estratégias de prevenção do suicídio embasadas em evidências reconhecidas globalmente. Entretanto, a diversidade cultural, socioeconômica e política entre as nações representa um desafio de grande magnitude, devido à falta de consideração nas peculiaridades de cada país, requerendo uma abordagem flexível e adaptável para assegurar que as estratégias globais sejam eficazes e ajustadas às demandas locais (Tabela 5).

Após as considerações das qualidades e empecilhos de cada forma de visualização, e levando principalmente em consideração a falta de variáveis relacionadas ao suicídio em contextos específicos, optamos por analisar o suicídio em escala global, possibilitando abranger dados de maior qualidade e propor soluções que abrangessem uma maior parcela da população.

6.3. Processos relacionados a base de dados

6.3.1. Entendimento básico do que é uma base de dados de qualidade:

Durante a seleção dos conjuntos de dados, buscamos compreender o que define uma base de dados de qualidade a partir da literatura profissional. Segundo Hastie, Tibshirani e Friedman (2009), dados de qualidade são cruciais para modelos estatísticos precisos, evitando conclusões incorretas. Dentre os fatores que os definem como de qualidade estão:

- Qualidade de dados: A utilização de dados de qualidade é fundamental para o desenvolvimento de modelos precisos. Dados ruidosos, incompletos ou com viés prejudicam a qualidade do programa.
- Dimensionamento dos dados: A quantidade de dados possui um impacto significativo na eficiência do modelo, podendo, entretanto, aumentar a necessidade computacional.
- Variedade de variáveis: Para projetos que buscam entender correlações, a existência de diversos tipos de variáveis que possam ser analisadas, impacta também a qualidade de um modelo.

6.3.2. Definindo a amostra de pesquisa e a limitação em uma perspectiva global:

Uma das questões fundamentais durante o processo de escolha de uma base de dados é definir a amostra que será analisada na pesquisa, sendo necessário levar em consideração diversos aspectos do programa e os objetivos. Apesar de uma maior quantidade de dados geralmente indicar maior precisão e robustez do modelo, há um ponto em que a adição de variáveis ou colunas se torna insignificante e pode até prejudicar o desempenho do modelo devido à complexidade e ao ruído adicional (FIGUEROA et al., 2012). Adicionalmente, a fim de realizar estimativas aproximadas, a utilização de métodos que prevejam a taxa de sucesso da amostra também são bons indicadores para a seleção da amostra (RAJPUT, WANG, CHEN, 2023).

No contexto da pesquisa focada na análise de dados globais, enfrentamos várias limitações devido à ausência de dados de certos países com capacidades limitadas ou que optam por não fornecer métricas, geralmente substituídas por estimativas. Essas limitações representam desafios para a produção de estatísticas mundiais, devido à falta de dados abrangentes e confiáveis. Assim, é necessário avaliar os dados disponíveis a fim de definir se uma pesquisa é viável (MAHLER, SERAJUDDIN, MAEDA, 2022). Adicionalmente, levando em consideração a falta de informação sobre o suicídio, como expressado pela própria OMS (2024), dados muito fora do padrão, e sem explicação aparente, foram vetados da pesquisa a fim de garantir uma maior qualidade dos resultados. A fim de garantir a qualidade e precisão do modelo, buscamos trabalhar com a maior quantidade possível de dados sem comprometer sua qualidade, segundo as definições de Figueroa e considerando limitações relacionadas a poder computacional.

6.3.3. Escolha das bases de dados:

Durante o processo da escolha das bases de dados, foi necessário a seleção de dados seguros e confiáveis, buscando abordar a maior quantidade de valores sem perder a qualidade. Nesse contexto, optamos por utilizar conjuntos de dados provenientes da Organização Mundial da Saúde (OMS, 2024), do Banco Mundial (BM, 2024), do Instituto Cato (CATO, 2024), da Administração Nacional Oceânica e Atmosférica (NOAA, 2024) e das Nações Unidas (NU, 2024) adaptados para o Kaggle (KAGGLE, 2024). O anterior, por serem consideradas como algumas das principais comunidades relacionadas a Inteligência Artificial.

As bases de dados escolhidas foram:

1) *O Human Development World Index* é um conjunto de dados global que oferece uma visão sobre o IDH em diversos países entre 1990 e 2019. Os dados foram obtidos do Relatório de Desenvolvimento Humano (RDH) e criado pelo analista de dados sênior Sourav Banerjee, detentor de cursos de aprendizado de máquina.

2) *World GDP (GDP, GDP per capita, and annual growths)*. Essa base de dados trata sobre PIB, PIB per capita e suas variações em diversos países entre 1960 e 2020, dados obtidos do Banco Mundial. Foi desenvolvida pelo profissional Ozgur Cem Tas, formado pela Universidade Técnica de Istambul e aluno de diversos cursos sobre análise de dados e aprendizado de máquina.

3) O conjunto *Global Inflation Dataset* possui dados sobre a inflação em 196 países entre os anos 1980 e 2024 obtidos pelo Banco Mundial. Formulada pelo analista Sazidul Islam, formado pela Universidade de Dhaka e com experiência como trabalhador da Unicef.

4) A base de dados *Unemployment dataset* analisa a taxa de desemprego em mais de 200 países entre 1991 e 2019. Foi criada pela Anjali Pant, analista de dados com bacharelado em tecnologia pela Universidade de Anusandhan.

5) *World Religion Projections 2010 to 2050* é um *data frame* que contém informações sobre a porcentagem de pessoas que fazem parte de uma religião, contendo oito “ramos” de religião: Budistas, Cristãos, Religiões tradicionais, Hindus, Judeus, Muçulmanos, outras religiões e ateus (apresentados como Não Afiliados na base de dados). Ademais dos dados de 2010 e 2020, a base de dados possui também estimativas do futuro. A base de dados foi desenvolvida por Utkarsh Singh, analista de dados formado pela Universidade Bharati Vidyapeeth.

6) O conjunto *The Weather of 187 Countries in 2020*, possui dados da temperatura média por dia de 187 países no ano de 2020. Os dados, foram obtidos da NOAA (Administração Nacional Oceânica e Atmosférica) e redigida por Amir Sedaghati, analista de dados e especialista em IA, formado pela universidade Azad.

7) Devido às limitações em relação aos dados do suicídio, optamos por obter os dados referentes ao suicídio diretamente do site da OMS. Assim, escolhemos a base de dados *Crude suicide rates (per 100 000 population) (SDG 3.4.2)*, que contém dados da taxa de suicídio por 100 mil habitantes dívida por sexo e abrangendo informações entre os anos 2019 e 2000.

8) A base de dados *The Human Freedom Index* aborda o Índice de Liberdade Humana e outras variáveis relacionadas no ano de 2019 em 162 países. Ela foi desenvolvida por Guillermina Schneider, analista de dados formada na Universidade Nacional de Rosario e com mestrado em ciência da informação pela Universidade George Washington. Os dados, foram juntados da Instituição Cato.

Devido às limitações respectivas a disposição de dados, poder computacional e definição do universo amostral, foram utilizados dados próximos à 2019. Adicionalmente, o presente estudo foi capaz de analisar 149 países, abrangendo um total de 91% da população mundial. A não consideração dos demais países correu devido à falta de dados, ou à ausência de credibilidade das informações disponíveis.

6.3.4. Softwares utilizados

Para o cruzamento de dados, utilizaremos o Python (PYTHON SOFTWARE FOUNDATION, 2020) junto ao compilador Colab (2024). Para as análises estatísticas aprofundadas será utilizado o software R (2024) e para a criação dos gráficos referentes à estatística será utilizado o software Tableau, com o intuito de gerar insights mais visuais e complexos. (TABLEAU, 2024)

6.3.5. Limpeza dos dados

Durante o desenvolvimento da pesquisa, o processo de limpeza e normalização de dados é fundamental para o bom andamento do projeto. O procedimento se inicia no ambiente Python com a importação de bibliotecas como Pandas e Numpy (MCKINNEY, 2023), seguida pelo carregamento dos dados de um arquivo salvo no formato .CSV (trata-se de um arquivo de texto, que contém valores organizado por linhas e dados separados por vírgulas). Após uma análise inicial, são identificados e tratados os valores faltantes, que podem ser excluídos ou preenchidos (LITTLE, RUBIN, 2019). Posteriormente, são verificadas possíveis duplicações e valores atípicos, utilizando técnicas como o Z-score (HAIR et al., 2010) para identificação e correção. É fundamental assegurar que cada coluna possua o tipo de dado adequado, fazendo, por exemplo, a conversão de *strings* em valores numéricos (WICKHAM, GROLEMUND, 2017). Por fim, os dados tratados são exportados para um novo

arquivo .CSV, preparados para análises mais aprofundadas, o que aprimora a qualidade dos resultados.

6.4. Técnicas de Pesquisa

Considerando que a pesquisa se divide em dois processos, abordaremos as técnicas utilizadas na obtenção das correlações e resultados estatísticos e as visualizações formuladas posteriormente a estatística.

6.4.1. Machine learning

- Regressão Linear: é um método estatístico utilizado para encontrar a relação entre uma variável dependente e uma ou mais variáveis independentes em uma linha reta. Usamos este método para entender como múltiplas variáveis se relacionam com o suicídio, fornecendo *insights* mais completos e relações mais significantes

- Regressão Polinomial: é um método estatístico utilizado para modelar a relação entre uma variável dependente e uma ou mais variáveis independentes, permitindo que essa relação seja representada por uma equação polinomial. Usamos este método, para obter mais *insights* e entende qual é a capacidade dos nossos dados para prever o suicídio.

- A regressão Ridge é um método estatístico avançado usado para modelar a relação entre uma variável dependente e múltiplas variáveis independentes, especialmente onde há multicolinearidade entre essas variáveis. Ao incorporar um termo de penalidade na equação de regressão, a regressão de crista permite que os analistas obtenham estimativas de coeficientes mais estáveis e confiáveis, mesmo que os dados sejam altamente correlacionados.

A análise dos dados obtidos no R permitirá uma interpretação detalhada dos resultados, ajudando a revelar padrões e relações significativas entre as variáveis. Com base nas explicações, será possível desenvolver estratégias e intervenções mais eficazes para prevenir o suicídio, contribuindo para um melhor entendimento e resposta a esse problema.

6.4.2. Análise estatística

Após a análise dos dados que serão utilizados na pesquisa foi definida a metodologia que seria implementada na estatística e conduziria a pesquisa. Dentre os escolhidos estão:

- Shapiro-Wilk: é um método estatístico utilizado para verificar se uma amostra de dados segue uma distribuição normal. Ele calcula uma estatística W, que compara os valores ordenados da amostra

com os valores esperados de uma distribuição normal (KLUWER, 2024). Usamos este método para verificar a distribuição dos dados, determinando se eles podem ser analisados por métodos paramétricos ou não paramétricos.

- Kruskal-Wallis: é um teste estatístico não paramétrico utilizado para comparar três ou mais grupos independentes para determinar se eles vêm da mesma população ou se pelo menos um grupo é significativamente diferente dos outros. Utilizaremos esse método, para comparar a distribuição de uma variável em quatro grupos, para ver o quão diferente valor da variável nesses diferentes grupos.

-Correlação de Spearman: é um método estatístico utilizado para verificar a existência de uma relação entre duas variáveis, independentemente de elas seguirem uma distribuição normal. Usamos este método para fazer uma relação entre as variáveis socioeconômicas e o suicídio.

6.5. Procedimentos da pesquisa

Após a implementação do teste de Shapiro-Wilk, utilizado unicamente para definir o tipo de distribuição dos dados, aplicamos respectivamente da seguinte maneira os métodos estatísticos para a extração de resultados:

Teste de Kruskal-Wallis:

O teste de Kruskal-Wallis será empregado para dar uma ideia inicial da correlação entre as variáveis e o suicídio, verificando como um índice independente afeta a distribuição do fenômeno em diferentes grupos. Ele nos ajudará na criação de hipótese e no entendimento inicial das correlações, direcionando os esforços para os demais métodos estatísticos.

Coeficiente de Correlação de Spearman:

O coeficiente de Correlação de Spearman será empregado para medir relação entre duas variáveis, independentemente de seguirem uma distribuição normal. No contexto deste projeto, ele ajudará a identificar quais variáveis estão mais fortemente correlacionadas com as taxas de suicídio. Essa análise preliminar é crucial para a seleção de variáveis a serem incluídas em modelos mais complexos, como a regressão polinomial. (OSTERTAGOVA, 2012).

Regressão polinomial e linear:

As regressões, polinomial e linear, serão utilizadas para modelar a relação entre o suicídio e múltiplas variáveis independentes, permitindo a captura de tendências lineares e não lineares nos

dados. Essa técnica é particularmente útil quando a relação entre as variáveis não é simplesmente linear, proporcionando uma melhor adequação aos dados e, consequentemente, previsões mais precisas. (BISHOP, 2006)

Regressão de Ridge:

A regressão de Ridge será usada para modelar a relação entre suicídio e múltiplas variáveis independentes para permitir uma análise mais precisa na possível presença de multicolinearidade. Esta técnica é particularmente eficaz quando os preditores são altamente correlacionados, o que pode dificultar a interpretação dos coeficientes nos modelos de regressão tradicionais. Ao aplicar a regressão descrita, podemos incluir um termo de penalidade que ajuda a estabilizar as estimativas dos coeficientes, resultando em um modelo mais confiável e geral.

Em relação as etapas que envolveram a formulação do resultado, estão respectivamente:

Primeira etapa: Primeiro coletamos microdados e metadados relacionados ao suicídio, abrangendo 149 países, divididos em oito bases de dados. A limpeza de dados foi realizada para garantir a qualidade e consistência das informações. Utilizamos técnicas estatísticas descritivas para compreender a distribuição inicial de variáveis como PIB per capita, taxa de desemprego e HFI. Esta etapa é fundamental para se familiarizar com os dados e identificar possíveis anomalias ou valores discrepantes.

Segunda etapa: Após a preparação dos dados, aplicou-se o teste de Shapiro-Wilk para avaliar a normalidade das distribuições das variáveis. Este teste é fundamental para determinar quais variáveis podem ser analisadas usando métodos paramétricos. Em seguida, utilizamos o teste de Kruskal-Wallis para comparar grupos e identificar diferenças significativas nas taxas de suicídio relacionadas a variáveis categóricas. Esta abordagem garante que apenas os dados mais relevantes sejam considerados para análise posterior, garantindo assim que os resultados sejam robustos e representativos.

Terceira etapa: Nesta fase, aplicamos as regressões linear, polinomial e de Ridge, para explorar melhor a relação entre variáveis relevantes e taxas de suicídio. A análise envolve a criação de modelos que sintetizem as relações observadas, identificando assim padrões mais profundos das interações entre variáveis.

6.6. Confeção dos gráficos e imagens

Após a realização da pesquisa, elaboramos gráficos e mapas para facilitar a compreensão dos resultados obtidos, processo descrito por Tufte (2020) como fundamental para revelar *insights* que não são visíveis a partir apenas de dados não processados.

Assim, considerando as ideias e métodos explicados por Tufte e Few (2012), decidimos realizar mapas-múndi, *scatter plots* e um mapa de risco para facilitar a visualização dos resultados obtidos. O modelo de risco foi feito com o intuito de aplicar os resultados dos testes estatísticos e aprendizado de máquina em uma solução gráfica, que permitisse fazer estimativas do suicídio e entender possíveis limitações referentes aos dados da pesquisa. As visualizações, foram formuladas no Tableau a partir dos dados obtidos e ordenadas em *data frames*.

6.7. Limitações em relação ao modelo

Embora a pesquisa busque uma compreensão abrangente do suicídio globalmente a partir de variáveis independentes que se referem a população e economia, é necessário estabelecer limitações.

Variáveis não incluídas:

O estudo visa compreender principalmente variáveis socioeconômicas, entretanto outros dados importantes como qualidade de saúde e influências culturais não serão abordadas diretamente.

Tempo analisado e países incluídos:

Na pesquisa, foram analisados 149 países (91 % da população mundial), os demais países, não foram analisados devido à falta de dados ou credibilidade das informações já existentes. Em relação ao tempo estudado, buscando abranger a maior quantidade de pessoas possíveis, e analisar dados de qualidade, optamos por analisar dados anteriores à pandemia de Covid, mais especificamente de 2019, levando em consideração a falta de dados pós-pandemia. No caso da variável ambiental Temperatura média, pela indisponibilidade de dados, foram utilizados os de 2020, por serem mais próximos ao ano de 2019.

7. Desenvolvimento da pesquisa

Antes de debruçarmos na interpretação dos resultados de nossa pesquisa, mostramos o processo que seguimos para a obtenção dos dados e sua análise.

7.1. Coleta de dados

Durante o processo de obtenção dos dados, devido às limitações relacionadas às informações existentes sobre o suicídio em nível global, foi necessária uma busca extensa dentro da plataforma Kaggle a fim de garantir dados confiáveis e que abrangessem uma quantidade considerável de informações em essa escala. Para o Banco Mundial (2022), os dados devem abranger ao menos metade da população mundial, universo que excedemos em nossa pesquisa já que abrangemos 91 %.

7.2. Tratamento e junção das bases de dados

Com as bases de dados já definidas, as exportamos para o Colab a fim de tratar, juntar e integrar os *data frames* para a facilitação de suas análises posteriores. Assim conseguimos reunir informação pertinente a 149 países. Utilizamos um conjunto de dados abrangente, que inclui informações sobre o Índice de Desenvolvimento Humano (IDH), o Índice de Liberdade Humana (HFI) o Produto Interno Bruto (PIB) per capita, a inflação, a taxa de desemprego, a temperatura média anual por país, a porcentagem das religiões presentes por país e as taxas de suicídio por 100 mil habitantes. Outras variáveis como crescimento do PIB, e outras que estão presentes nessas bases de dados, foram descartadas devido a que fogem do escopo de nossa pesquisa ou não aportam elementos importantes para nosso estudo. As bases de dados utilizadas estão presentes no anexo a partir de um *google drive*.

O processo de limpeza dos dados foi caracterizado pela remoção de dados faltantes, como caracteres incorretos e informações desnecessárias. Adicionalmente, as bases de dados tiveram seu design alterado, removendo as colunas por ano, e estabelecendo uma coluna única para cada variável. A fim de facilitar a junção dos dados, as colunas compartilhadas foram renomeadas para um nome comum (País e Ano) e organizadas de tal forma a tornar possível a integração de cada variável socioeconômica e a taxa de suicídio por 100.000 habitantes.

Para garantir a consistência dos dados, as colunas de cada base foram renomeadas e padronizadas para os seguintes termos: *Pais*, *Ano*, *IDH*, *PIB_per_capita*, *Taxa_de_inflacao*, *Desemprego* e *Taxa_de_suicidio_por_100k*. Este processo de padronização visou facilitar a integração das informações e assegurar a precisão da análise. Por fim criando uma base de dados única, denominada “df_final”.

7.3. Descrição dos processos estatísticos

Após o tratamento dos dados, já descrito, a base de dados final “df_final” foi exportada para o software R, onde os processos estatísticos e a obtenção dos resultados foram realizadas. As conclusões e explicações referentes as informações foram baseadas conforme as citações já expostas na metodologia. Entre os métodos aplicados e seus respectivos resultados estão:

- Teste de Shapiro-Wilk

Primariamente as análises de correlação entre as variáveis socioeconômicas e o suicídio, este teste foi conduzido para testar a distribuição dos dados com o intuito de definir os métodos que seriam utilizados posteriormente na pesquisa.

Tabela 7: Tipo de distribuição com W e *p-value*

Variável	W	<i>p-value</i>	Tipo de distribuição
PIB per capita	0.7278721	$p < 0,001$	Não normal
Desemprego	0.8373332	$p < 0,001$	Não normal
Taxa de inflação	0.1905451	$p < 0,001$	Não normal
IDH	0.9444966	$p < 0,001$	Não normal
HFI	0.9722308	$p < 0,001$	Não normal
Temperatura média	0.9190146	$p < 0,001$	Não normal
Budistas	0.2799062	$p < 0,001$	Não normal
Cristãos	0.8401889	$p < 0,001$	Não normal
Religiões Folclóricas	0.3101249	$p < 0,001$	Não normal
Hindus	0.2687689	$p < 0,001$	Não normal
Judeus	0.0562147	$p < 0,001$	Não normal
Muçulmanos	0.6866329	$p < 0,001$	Não normal
Ateus (Não Afilhados)	0.6980486	$p < 0,001$	Não normal
Outras Religiões	0.1405999	$p < 0,001$	Não normal

Nota: $p < 0,001$ (extremamente significativo); $p < 0,01$ (altamente significativo); $p < 0,05$ (significativo)

Fonte: Autores.

O *p-value*, valor fornecido, indica a probabilidade de se observar uma diferença tão grande ou maior do que a que foi observada sob a hipótese nula. No contexto desse teste, a hipótese nula afirma que a população possui distribuição normal, enquanto a hipótese alternativa sugere que a população não segue uma distribuição normal. Assim, um valor de *p* menor que 0,05 indica que os dados não são normais e a que a tese primária do teste está errada; O valor w também indica a normalidade dos dados a partir de um valor entre 0 e 1, em que quanto maior, mais normal os dados são.

Portanto, a partir da análise foi possível concluir que todas as variáveis têm um valor de p menor que 0,05 indicando assim que são não normalmente distribuídas. A partir deste texto concluímos que os demais processos estatísticos deveriam ser não paramétricos.

- Teste de Kruskal-Wallis

O teste de Kruskal-Wallis foi realizado com o intuito de fornecer uma dimensão inicial da relação entre as variáveis socioeconômicas e o suicídio, a partir das comparações de distribuições do método. No contexto da nossa pesquisa, os países foram divididos em quatro grupos, a partir de quartis definidos pelas taxas de suicídio. Após essa divisão, foi comparado, a partir do teste, a diferença da distribuição das variáveis nos grupos, fornecendo o *p-value*, que mostra a probabilidade de a hipótese nula ser verdadeira, indicando, portanto, se a variável relacionada condiciona a distribuição da taxa de suicídio. Os valores observados a partir de nossos dados foram:

Tabela 8: Resultados teste Kruskal-Wallis

Variável	<i>P-value</i>	Interpretação
PIB per capita	$p < 0,001$	Diferença extremamente significativa
Desemprego	$p < 0,001$	Diferença extremamente significativa
Taxa de inflação	0.142	Sem diferença significativa
IDH	$p < 0,001$	Diferença extremamente significativa
HFI	$p < 0,001$	Diferença extremamente significativa
Temperatura média	$p < 0,001$	Diferença extremamente significativa
Budistas	0.9048	Sem diferença significativa
Cristãos	0.4599	Sem diferença significativa
Religiões Tradicionais	0.1898	Sem diferença significativa
Hindus	0.2997	Sem diferença significativa
Judeus	0.5695	Sem diferença significativa
Muçulmanos	0.006623	Diferença significativa
Ateus	$p < 0,001$	Diferença extremamente significativa
Outras Religiões	0.3535	Sem diferença significativa

Nota: $p < 0,001$ (extremamente significativo); $p < 0,01$ (altamente significativo); $p < 0,05$ (significativo)

Fonte: Autores

No contexto deste processo estatístico, a hipótese nula indica que a mediana de todos os grupos é igual, enquanto a hipótese alternativa indica que ao menos uma mediana é diferente.

Interpretação dos valores de *p-value*:

Após a análise dos resultados, levando em consideração apenas as variáveis com p-valor menor que 0,05, definimos estas variáveis como relevantes na distribuição do suicídio:

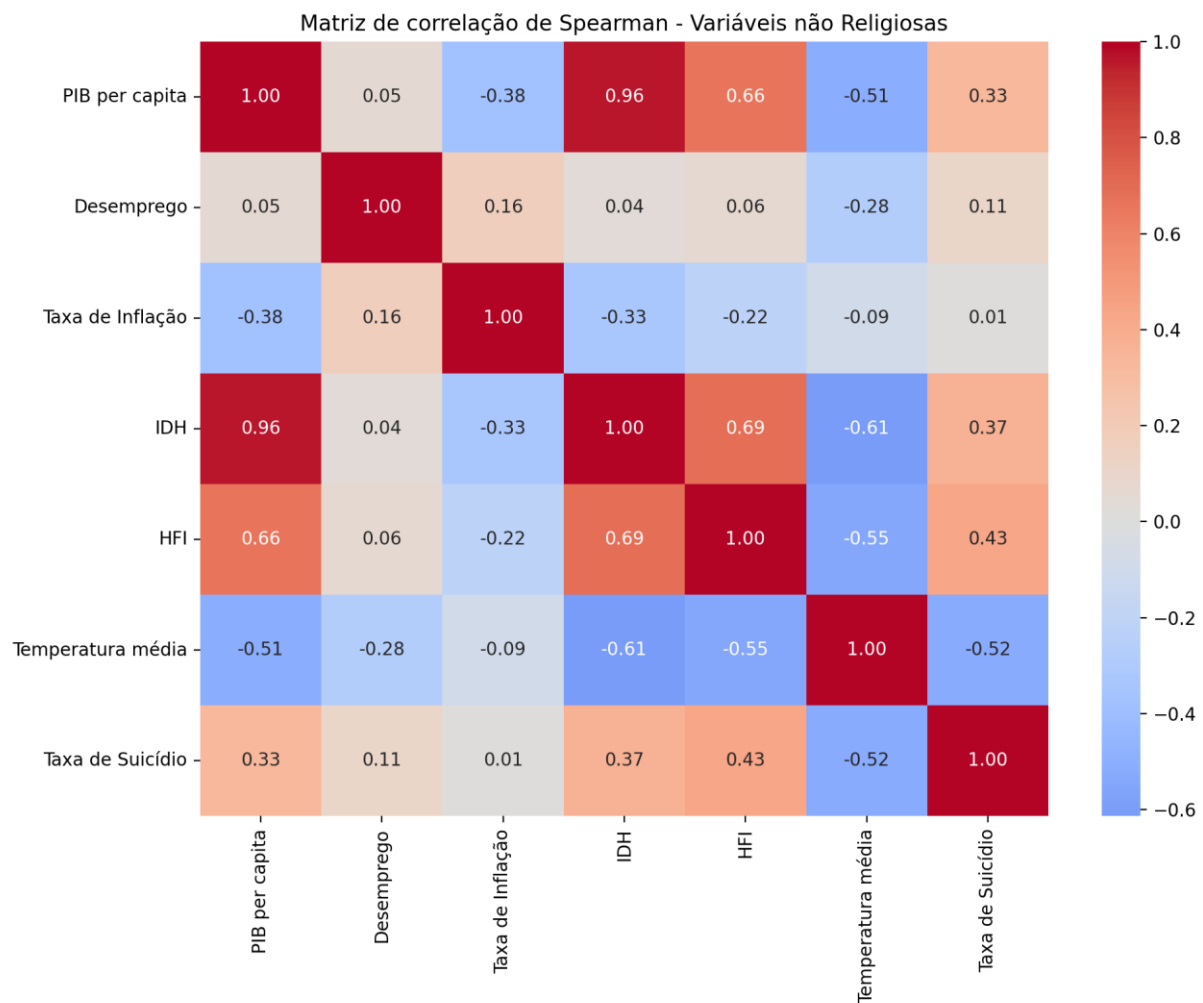
- PIB per capita (diferença altamente significativa)
- Desemprego (diferença altamente significativa)
- IDH (diferença altamente significativa)
- HFI (diferença altamente significativa)
- Temperatura média (diferença altamente significativa)
- Taxa de muçulmanos
- Taxa de ateus (diferença altamente significativa)

Com o teste pudemos concluir que as variáveis acima, mais especialmente a temperatura média e a taxa de ateus estão relacionadas com a distribuição do suicídio. Essas classificações facilitarão as análises posteriores reduzindo o número de análises feitas.

- Correlação de Spearman

A correlação de Spearman foi feita para entender de maneira concreta a relação entre as variáveis socioeconômicas e o suicídio, indicando sua força e direção a partir do coeficiente obtido. Adicionalmente, cruzamos as variáveis estudadas entre si para obter *insights* capazes de explicar melhor os valores e relações obtidas primeiramente. A fim de facilitar a observação dos dados, dividimos as variáveis em dois grupos (variáveis religiosas e não religiosas) e criamos mapas de calor a partir das suas respectivas matrizes. O coeficiente do método varia entre -1 e 1 e quanto mais próximo aos limites, maior é a correlação. Um valor negativo, indica que a tendência da variável é oposta ao suicídio, e quando positivo, indica que ambos os valores variam na mesma direção. Ao contrário do teste de Kruskal-Wallis, esta estatística fornece um coeficiente que quantifica diretamente a correlação entre duas variáveis, facilitando assim, sua compreensão e análise. Assim, realizamos os gráficos de matriz para interpretar os resultados:

Figura 3: Mapa de calor da correlação entre variáveis não religiosas



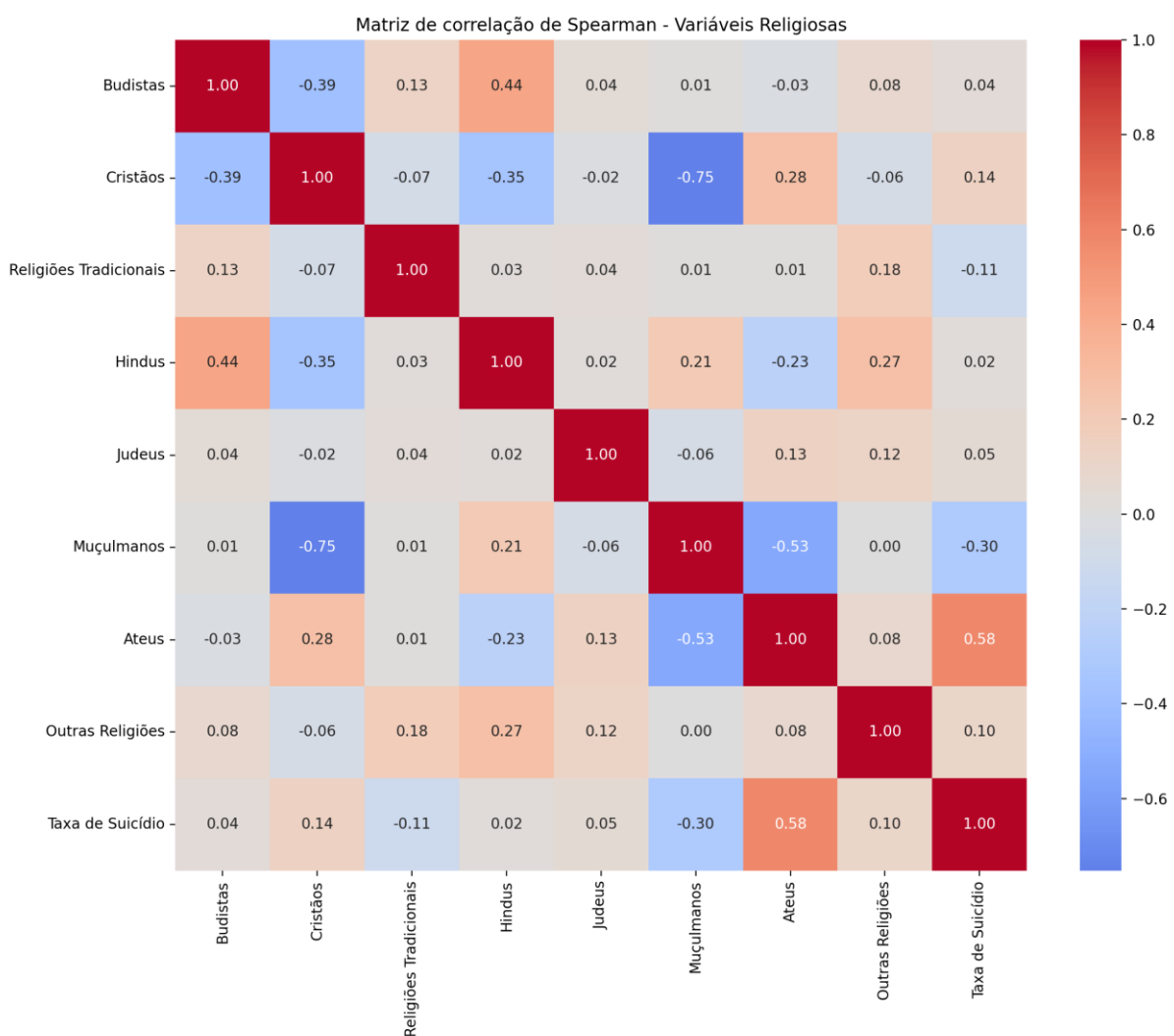
Fonte: Autores

Na análise de correlação das variáveis não religiosas, e comparando com o teste de Kruskal-Wallis, percebemos que as mesmas variáveis, PIB per capita, IDH, HFI, Temperatura média e Desemprego possuem uma correlação significativa com a taxa de suicídio. A taxa de inflação não possui correlação significativa, razão pela qual não consideraremos essa variável em nossa análise. Os coeficientes das variáveis significativas são, respectivamente:

- Desemprego: 0,11;
- PIB per capita: 0,32;
- IDH: 0,36;
- HFI, 0,41;
- Temperatura média: -0,52.

As variáveis econômicas significativas estão relacionadas de maneira positiva com o suicídio, diferente da temperatura média, que segue uma distribuição contrária ao fenômeno. A partir do cruzamento das variáveis socioeconômicas entre si concluímos que o IDH e o PIB per capita, e mais especificamente suas distribuições são extremamente semelhantes, já que possuem um coeficiente de 0,96. Devido a essa semelhança a análise das duas variáveis será feita de forma conjunta a fim de facilitar suas interpretações e futuramente direcionar de maneira mais eficaz esforços relacionados a elas. Em relação ao desemprego, apesar de não ter obtido um coeficiente relevante em comparação as demais variáveis, nos demais testes ele obteve resultados significativos que podem ser considerados úteis na prevenção do suicídio.

Figura 4: Mapa de calor da correlação entre variáveis religiosas



Fonte: Autores

Na análise das variáveis religiosas, percebemos que as variáveis taxam de muçulmanos e taxa de ateus, as mesmas consideradas como relevantes no teste anterior, são significativamente

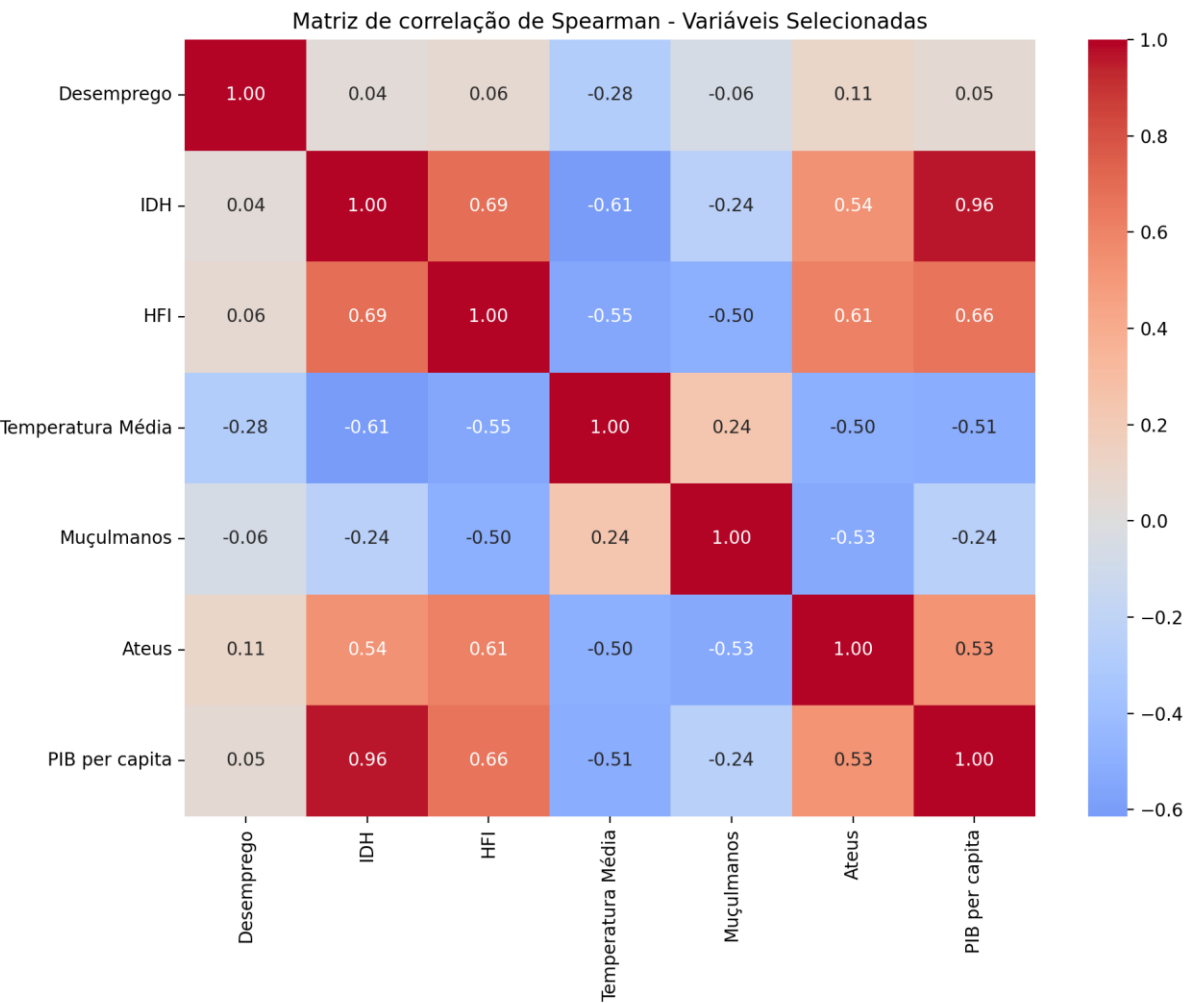
relacionadas ao problema. As demais variáveis não possuem valores significativos para a análise do suicídio. Os coeficientes relevantes obtidos são:

- Taxa de muçulmanos: -0,29;
- Taxa de não afilhados: 0,57;

Apesar de algumas variáveis apresentarem relações entre si, a literatura revisada não forneceu evidências que estabeleçam uma conexão clara entre diferentes religiões e como essas relações poderiam contribuir para conclusões úteis na prevenção do suicídio, portanto, não serão abordadas.

Além das análises nas duas divisões (religiosas e não religiosas), decidimos juntar os dados relevantes dos dois grupos a fim de perceber outras correlações significantes e gerar mais *insights*.

Figura 5: Mapa de calor da correlação entre as variáveis



Fonte: Autores

Após a análise cruzada, percebemos que a temperatura média está relacionada de maneira negativa e significativa com o PIB per capita, IDH e HFI, indicando que países com altos valores nesses índices tem de maneira geral uma menor temperatura. Além do clima, a taxa de ateus também está relacionada com essas três variáveis, mas de maneira positiva, indicando também no geral, que países com os índices mencionados maiores, possuem taxa de ateus maiores. Assim respaldados pela literatura que será discutida posteriormente, percebemos que as duas variáveis explicativas (temperatura e taxa de ateus) podem estar relacionadas com o índice de suicídio mais elevado nas regiões mais desenvolvidas. As correlações obtidas foram:

Temperatura média em relação aos três índices:

- IDH: -0,60;
- PIB per capita -0,51;
- HFI: -0,55;

Taxa de ateus em relação aos três índices:

- IDH: 0,53;
- PIB per capita 0,53;
- HFI: 0,60;

Como descrito anteriormente, esses cruzamentos foram abordados nas demais etapas da pesquisa e foram relacionadas com a literatura existente.

Assim, após as análises a partir do teste de Spearman, chegamos à conclusão de que as seguintes variáveis estão relacionadas ao suicídio:

- Desemprego: 0,13;
- PIB per capita: 0,32;
- IDH: 0,36;
- HFI, 0,41;
- Temperatura média: -0,52;
- Taxa de muçulmanos: -0,29;
- Taxa de não afilhados: 0,57;

Identificamos também, a correlação cruzada entre as variáveis socioeconômicas, sendo elas:

Temperatura média em relação aos três índices:

- IDH: -0,60;
- PIB per capita -0,51;
- HFI: -0,55;

Taxa de ateus em relação aos três índices:

- IDH: 0,53;
- PIB per capita 0,53;
- HFI: 0,60;

- Regressão polinomial, linear e de Ridge:

As regressões foram utilizadas a fim de obter *insights* mais profundos e avaliar a qualidade dos nossos dados em prever o suicídio, deixando mais claro, portanto, a significância das variáveis do estudo. As regressões linear e polinomial foram usadas para a análise individual entre as variáveis e o suicídio e para o cruzamento entre elas, enquanto a regressão de Ridge permitiu entender de maneira simultânea a relação entre as variáveis e o suicídio, mostrando a relevância entre as informações estudadas na pesquisa e o fenômeno. Os valores utilizados nestes testes foram filtrados a partir da correlação de Spearman e o teste de Kruskal-Wallis, a fim de promover a eficiência e qualidade do projeto. Os cruzamentos obtidos no teste anterior também serão aplicados no modelo de regressão para a criação de *insights* mais profundos.

- Resultados das regressões individuais entre as variáveis socioeconômicas e o suicídio:

Para definir qual modelo de regressão seria utilizado nas variáveis, foi utilizado o método AIC (Critério de Informação de Akaike) usado para avaliar a eficácia e otimização de um modelo a partir do valor resultante, que quanto menor for, indica uma maior qualidade. No contexto da nossa pesquisa, geramos três tipos de regressão para cada variável (linear, polinomial quadrática e polinomial cúbica) e verificamos o valor de AIC, definindo qual modelo possuía o menor valor. Apesar da possibilidade de considerar o erro quadrático e *p-value* para a seleção do modelo, essas técnicas podem não identificar problemas, como o *overfitting*, o que prejudica gravemente a regressão.

Assim, fizemos os testes e definimos os modelos:

Tabela 9: Resultados método AIC linear, quadrático e cúbico

Variável	AIC Linear	AIC Quadrático	AIC Cúbico
PIB per capita	1060.524	1061.882	1063.568
Desemprego	1055.177	1056.656	1056.233
IDH	1058.779	1059.268	1061.020
HFI	1053.542	1055.538	1057.433
Temperatura média	1038.280	1040.280	1042.214
Muçulmanos	1046.721	1048.622	1050.598
Ateus (Não Afilhados)	1045.101	1042.442	1043.958

Fonte: Autores

Após as análises dos resultados, concluímos que para a maioria das variáveis, uma regressão linear é mais eficaz. Entretanto, para a taxa de ateus, uma regressão polinomial quadrática será utilizada.

Adicionalmente, durante o processo de análise dos dados para as criações dos modelos, percebemos que na distribuição das variáveis com o suicídio, alguns valores do fenômeno se afastavam da média, prejudicando a precisão das regressões. Portanto, decidimos verificar como o AIC se alterava com a remoção desses *outliers*, a fim de melhorar a eficácia do modelo, verificada novamente pelo método.

Os resultados foram:

Tabela 10: Resultados AIC Normal e Ponderado

Variável	AIC Normal	AIC Ponderado
PIB per capita	1060.524	942.8111
Desemprego	1055.177	949.5742
IDH	1058.779	933.6456
HFI	1053.542	931.8068
Temperatura média	1038.280	901.327
Muçulmanos	1046.721	950.598
Ateus (Não Afilhados)	1045.101	1042.442

Fonte: Autores

Após essas análises, concluímos que os modelos ponderados são mais eficazes, e, portanto, serem utilizados para a análise dos resultados.

Assim, fizemos as regressões e obtemos os resultados referente a correlação entre as variáveis e o suicídio e a precisão dos modelos:

Tabela 11: Resultados R^2 e P -value variáveis individuais

Variável	R ² ajustado	<i>p-value</i>
PIB per capita	0.05797	$p < 0,001$
Desemprego	0.01392	0.0816
IDH	0.1145	$p < 0,001$
HFI	0.1255	$p < 0,001$
Temperatura média	0.2882	$p < 0,001$
Taxa de muçulmanos	0.1402	$p < 0,001$
Taxa de ateus	0.3483	$p < 0,001$

Nota: $p < 0,001$ (extremamente significativo); $p < 0,01$ (altamente significativo); $p < 0,05$ (significativo)

Fonte: Autores

Para as regressões no geral, o valor de p indica a probabilidade de a hipótese nula ser verdadeira, que afirma não existir significância entre a variável independente e a dependente, enquanto a hipótese alternativa sugere que existe uma correlação significativa entre as variáveis.

Praticamente todas as variáveis estão relacionadas significativamente com o suicídio, com exceção do desemprego de forma parcial, que não obteve um *p-value* igual a 0,05, apesar de ainda ser considerado significativo por ter obtido um valor de p menor que 0,1. Apesar dos testes das regressões terem dados resultados semelhantes aos métodos anteriores, foi possível notar que algumas variáveis menos correlacionadas do que outras são mais eficazes na predição do suicídio, como a taxa de muçulmanos em comparação ao HFI, IDH e PIB per capita que possui um r quadrático melhor. Mostrando assim que sua distribuição é mais previsível, e, portanto, no contexto da nossa pesquisa, ela pode ser mais útil na prevenção do suicídio.

Desta maneira, a análise apenas da correlação não é a maneira mais eficaz de avaliar a qualidade de uma variável na análise do suicídio, já que a capacidade em prever o fenômeno também é fundamental para amenizá-lo. O r quadrático será analisado posteriormente, a partir da regressão de Ridge a fim de avaliar a capacidade das variáveis em prever o suicídio.

- Resultados da regressão individual dos cruzamentos entre as variáveis socioeconômicas:

Para estes cruzamentos, realizamos os mesmos métodos para avaliar a qualidade dos modelos, comparando os valores de AIC. Entretanto, como utilizamos duas variáveis dependentes (temperatura média e taxa de muçulmanos) dividimos os resultados em duas tabelas. Assim, obtemos os seguintes resultados:

Tabela 12: Resultados método AIC linear, quadrático e cúbico para o cruzamento de temperatura média em relação aos três índices

Variável	AIC Linear	AIC Quadrático	AIC Cúbico
IDH	951.4527	947.0954	946.9620
PIB per capita	985.9218	982.9679	983.1069
HFI	968.3200	947.6375	945.5827

Fonte: Autores

Tabela 13: Resultados método AIC linear, quadrático e cúbico para o cruzamento de taxa de ateus em relação aos três índices

Variável	AIC Linear	AIC Quadrático	AIC Cúbico
IDH	1123.688	1114.284	1114.033
PIB per capita	1126.655	1120.061	1118.757
HFI	1117.905	1097.177	1093.047

Fonte: Autores

Após a análise dos resultados, chegamos à conclusão que para o PIB per capita e HFI, um modelo cúbico é mais eficaz dentre os três métodos. Entretanto, para o IDH, o modelo quadrático é melhor.

De maneira semelhante a análise das variáveis com o suicídio, optamos também por verificar a qualidade dos modelos com e sem outliers. Assim, fizemos essa comparação também:

Tabela 14: Resultados AIC Normal e Ponderado para o cruzamento de temperatura média em relação aos três índices:

Variável	AIC Normal	AIC Ponderado
IDH	1114.284	1029.473
PIB per capita	1120.061	1011.860
HFI	1093.047	1007.998

Fonte: Autores

Tabela 15: Resultados AIC Normal e Ponderado para o cruzamento de taxa de ateus em relação aos três índices:

Variável	AIC Normal	AIC Ponderado
IDH	1114.033	1027.035
PIB per capita	1120.061	1011.860
HFI	1093.047	1007.998

Fonte: Autores

De maneira semelhante as análises realizadas anteriormente, a remoção dos outliers foi positiva na melhoria da eficiência. Assim, consideraremos os resultados dos modelos ponderados.

Após as regressões, obtemos os seguintes resultados:

Tabela 16: Resultados R^2 e P -value cruzamento temperatura média

Variável	P -Value	R^2 ajustado
IDH	2e-16	0.4088

PIB per capita	0.0279	0.2397
HFI	2e-16	0.4143

Nota: $p < 0,001$ (extremamente significativo); $p < 0,01$ (altamente significativo); $p < 0,05$ (significativo)

Fonte: Autores

Tabela 17: Resultados R^2 e P -value cruzamento taxa de ateus

Variável	P -Value	R^2 Ajustado
IDH	2e-16	0.2654
PIB per capita	0.00377	0.2296
HFI	2e-16	0.3632

Nota: $p < 0,001$ (extremamente significativo); $p < 0,01$ (altamente significativo); $p < 0,05$ (significativo)

Fonte: Autores

A partir da tabela, foi possível concluir que os três índices estão correlacionados com as duas variáveis analisadas de maneira semelhante, obtendo p -values menores que 0,05. Entretanto, as variáveis preveem e explicam a temperatura média melhor em relação a taxas de ateus. De qualquer forma, ambos resultados são relevantes e devem ser considerados na prevenção do suicídio.

- Resultados da regressão de Ridge entre as variáveis socioeconômicas e o suicídio:

Após as regressões individuais, decidimos fazer a regressão de Ridge a fim de entender como a variação de mais de uma variável influem no suicídio. Em uma primeira instância, buscamos utilizar no modelo todas as variáveis e posteriormente filtramos os dados utilizados a fim de entender qual modelo é mais eficaz no contexto dos nossos dados.

Assim, o modelo que utilizou todas as variáveis obteve o seguinte resultado:

Tabela 18: Resultado inicial regressão de Ridge

Variável	Coeficiente	<i>P-value</i>
PIB per capita	-2.102e-05	0.4779
Desemprego	0.1690	0.0261
Taxa de Inflação	0.001656	0.9268
IDH	-2.798	0.5218
HFI	-0.4155	0.4614
Temperatura Média	-0.3819	1.09e-06
Budistas	-0.5804	0.1504
Cristãos	-0.5791	0.1492
Religiões Folclóricas	-0.5770	0.1792
Hindus	-0.5314	0.1990
Judeus	-0.6048	0.1388
Muçulmanos	-0.6121	0.1261
Ateus (Não Afilhados)	-0.4769	0.2454
Outras Religiões	0.1226	0.8690

Nota: $p < 0,001$ (extremamente significativo); $p < 0,01$ (altamente significativo); $p < 0,05$ (significativo)

Fonte: Autores

R^2 Ajustado: 0.01108669

No modelo geral, foi atribuído à maioria das variáveis um peso semelhante, em que a variável com peso mais importante foi a taxa de judeus devido ao seu maior coeficiente. As distribuições parecidas dos pesos não foram positivas para o modelo, já que o R quadrático ajustado foi muito baixo. De qualquer forma, a temperatura média e desemprego obtiveram resultados significantes, o que poderia indicar uma possível explicação da variação a partir dessas duas variáveis.

Devido ao resultado insatisfatório, decidimos remover as variáveis consideradas insignificantes nos testes a fim de verificar como o modelo mudaria. Após alguns testes, chegamos a um melhor modelo que consegue explicar o suicídio da maneira mais eficaz com os dados disponíveis na pesquisa. Os resultados foram:

Tabela 19: Resultado otimizado da regressão de Ridge

Variável	Coeficiente	<i>P-value</i>
Desemprego	0.16218	0.02481

IDH	-1.55972	0.66121
HFI	-0.64641	0.20925
Temperatura Média	-0.31762	7.90e-06
Muçulmanos	-0.03950	0.00334
Não Afilhados	0.11916	0.00216

Nota: $p < 0,001$ (extremamente significativo); $p < 0,01$ (altamente significativo); $p < 0,05$ (significativo)

Fonte: Autores

R^2 Ajustado: 0.4170135

Na regressão ajustada, foi obtido valores mais significantes, principalmente em relação ao R ajustado. O PIB per capita, apesar de relevante foi removido do modelo, já que devido a sua semelhança com o IDH, ele não era utilizado devido aos ajustes realizados pela multicolinearidade. O IDH e HFI, não obtiveram *p-values* muito relevantes, entretanto foram úteis na análise do suicídio.

Adicionalmente, o desemprego, apesar de não ter obtido resultados significantes como outras variáveis, foi útil em explicar a variação do suicídio.

Portanto, a utilização de variáveis sem correlações aparentes pode ser prejudicial para a análise do suicídio, e não devem ser consideradas nas criações de medidas. As variáveis presentes no modelo são as mais úteis na explicação do suicídio, e devem ser analisadas de maneira especial.

7.4. Análise exploratória dos dados e criação de *insights*

Com os resultados expostos previamente, foi possível definir as variáveis relevantes para o estudo do suicídio, para a partir da análise das diversas formas em que elas se relacionam, gerar visualizações gráficas que facilitem a sua compreensão e orientem conclusões para direcionar esforços.

Após a consideração dos métodos estatísticos e seus respectivos resultados, identificamos as seguintes variáveis como significativas na análise do problema estudado:

- Índice de Desenvolvimento Humano (IDH) e PIB per capita;
- Desemprego;
- Índice de Liberdade Humana (HFI);
- Temperatura média;
- Taxa de muçulmanos;
- Taxa de ateus.

Ademais das relações entre as variáveis e o suicídio, analisaremos mais profundamente os cruzamentos obtidos entre os valores estudados. Assim, encontramos também os cruzamentos que serão analisados mais profundamente:

- Temperatura média em relação ao IDH, PIB per capita e Índice de liberdade Humana (HFI)
- Taxa de ateus em relação ao IDH, PIB per capita e Índice de liberdade Humana

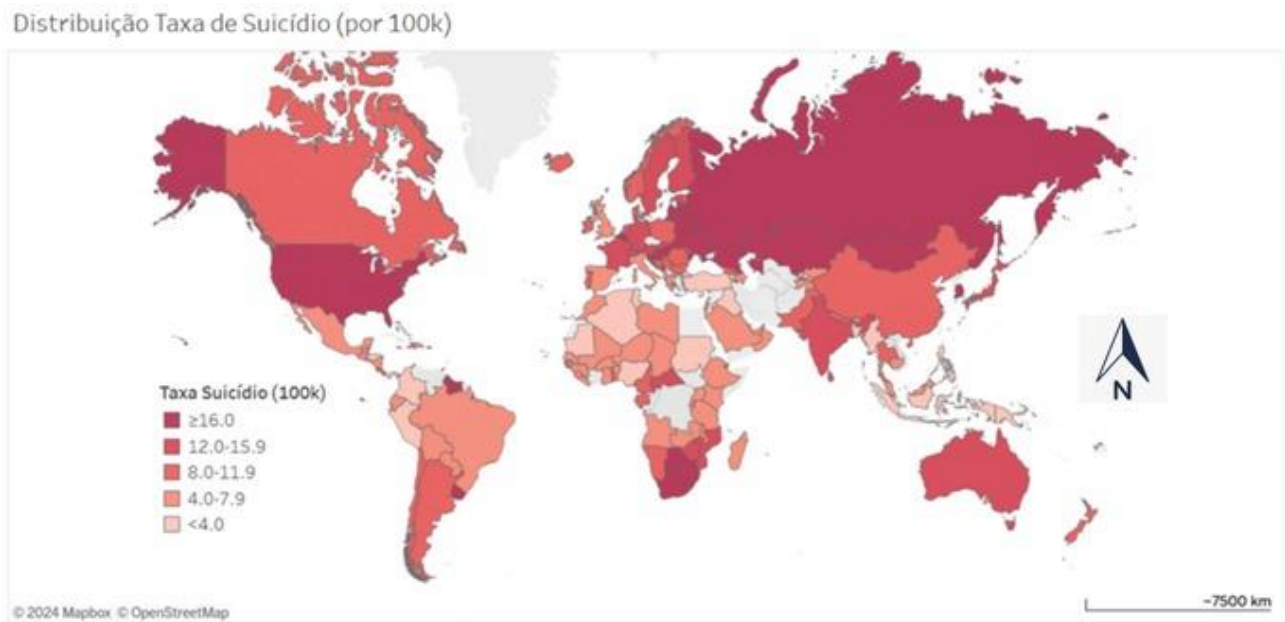
Uma vez escolhidas, realizamos uma análise individual para cada variável com o suicídio ou cruzamento, levando em consideração as conclusões obtidas a partir dos resultados. Devido aos processos estatísticos já terem sido explicados de maneira detalhada, as explicações referentes aos gráficos darão apenas uma ideia geral do padrão reconhecido e sua respectiva importância.

Adicionalmente, como ponto final da nossa pesquisa, decidimos empreender e desenvolver um mapa de análise de risco para o suicídio, considerando todos os valores estudados e seus resultados, a fim de gerar um sistema de visualização central que direcione os esforços para as criações de medidas, almejando entender as limitações de nossa análise, e perceber, a eficácia dos dados analisados em diferentes regiões.

Seguem as visualizações, que serão discutidas posteriormente:

7.4.1. Distribuição taxa de suicídio

Figura 6. Distribuição taxa de suicídio



Fonte: Autores

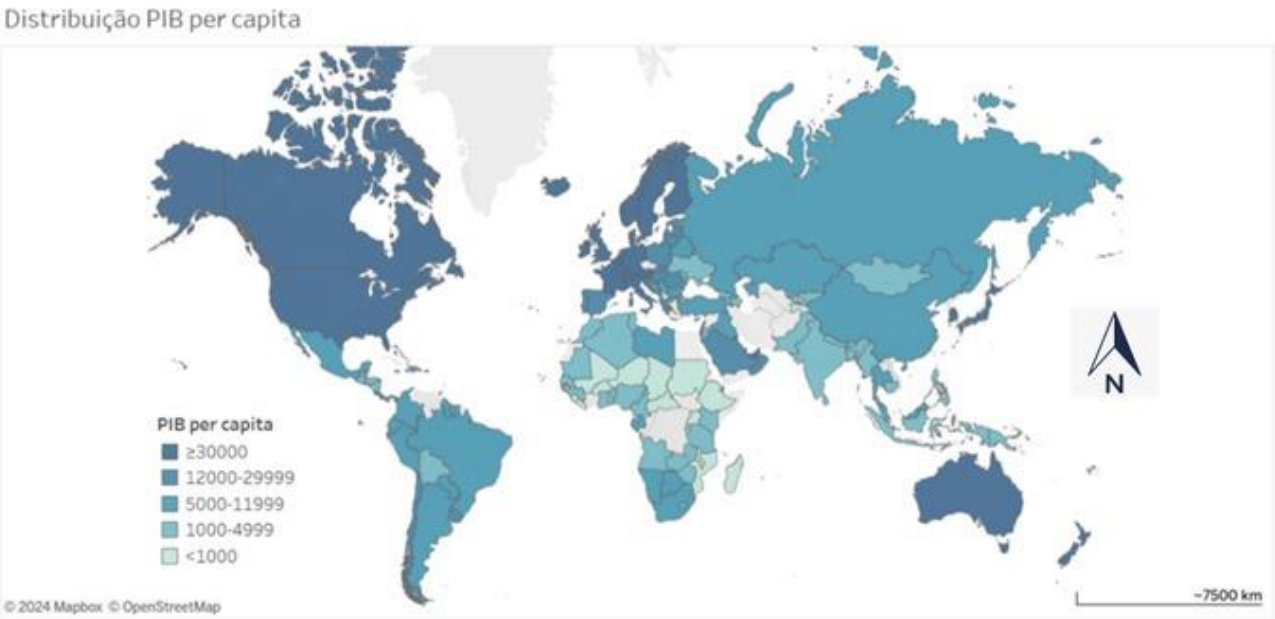
7.4.2. Análise do Índice de Desenvolvimento Humano e PIB per capita em relação à taxa de suicídio:

Figura 7: Distribuição de Índice de Desenvolvimento Humano (IDH)



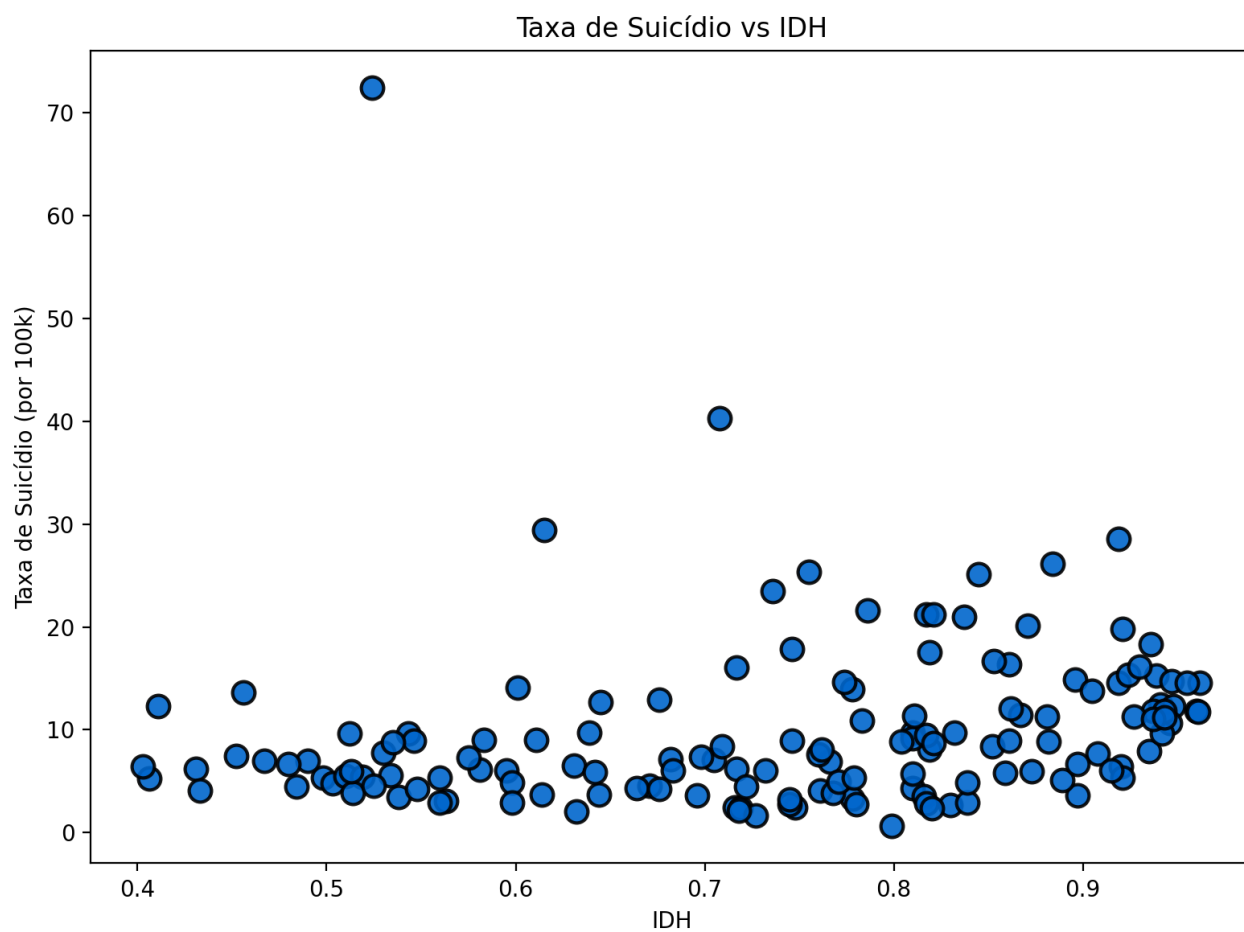
Fonte: Autores

Figura 8: Distribuição PIB per capita



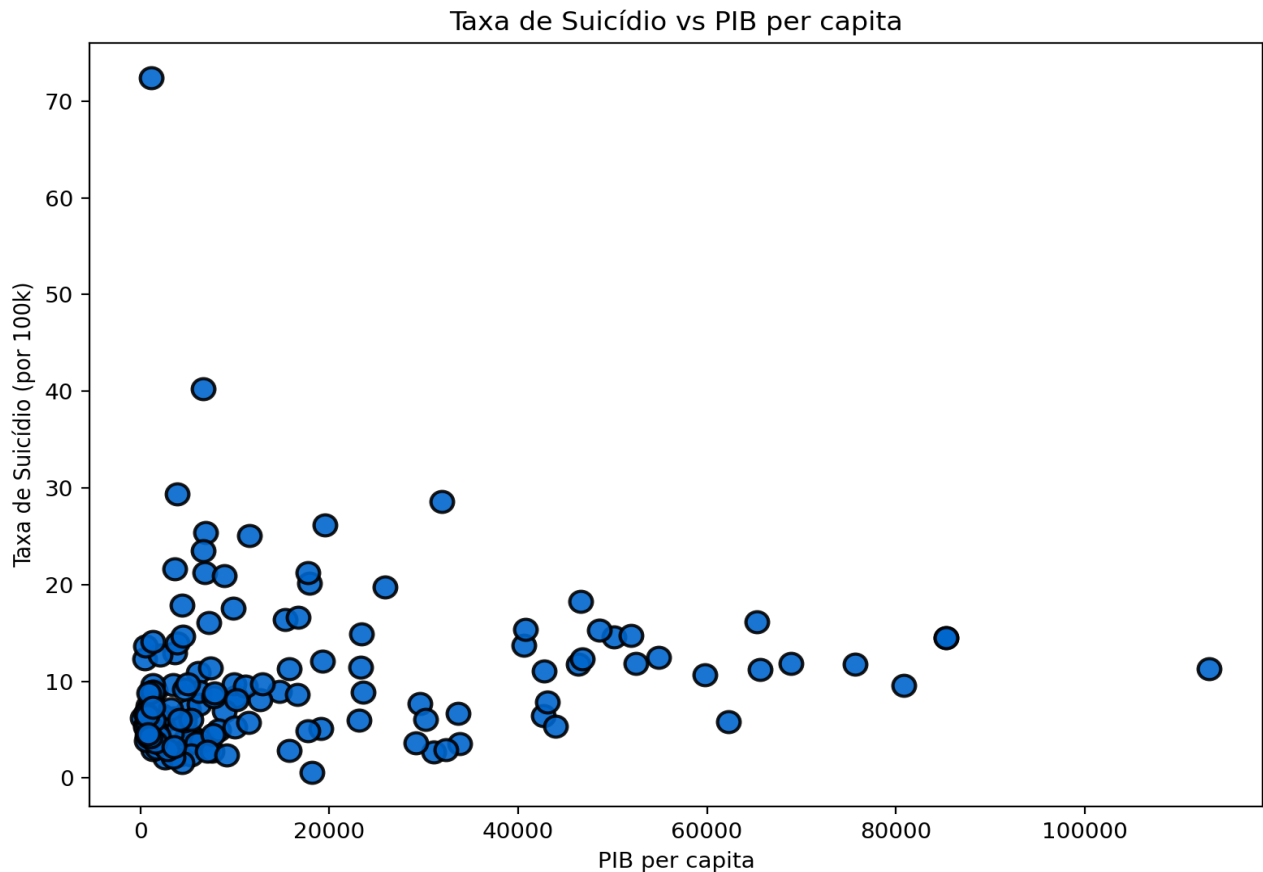
Fonte: Autores

Figura 9: Distribuição IDH em relação ao suicídio



Fonte: Autores

Figura 10: Distribuição PBI per capita em relação à Taxa de Suicídio

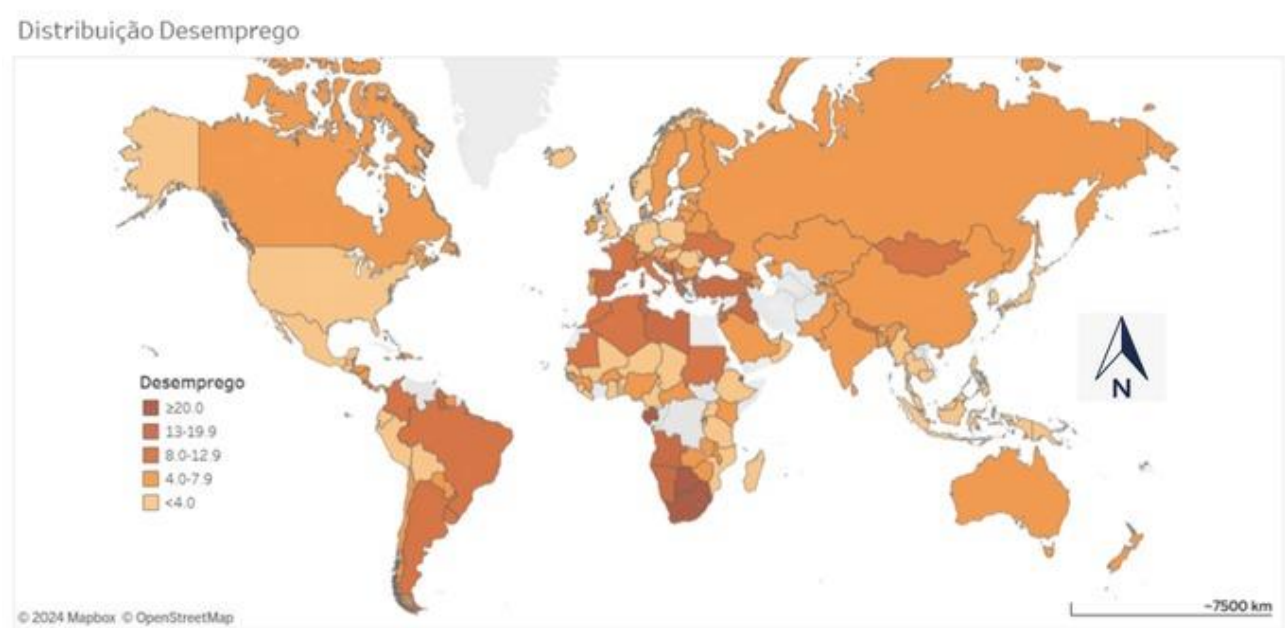


Fonte: Autores

Após a análise das distribuições das variáveis em relação ao suicídio, foi possível perceber que tanto o IDH e o PIB per capita estão parcialmente relacionados de maneira positiva com o suicídio e que as duas variáveis possuem distribuições semelhantes com o suicídio. O IDH e PIB per capita são respectivamente a quarta e quinta variáveis mais relevantes com o suicídio. Apesar da correlação, os valores do r ajustado nas regressões foi menor em relação a outras variáveis, indicando que existe uma alta variabilidade. De qualquer forma, é necessário levar esses índices em consideração, entendendo que países com esses perfis podem estar mais sujeitos a tentativas de suicídio.

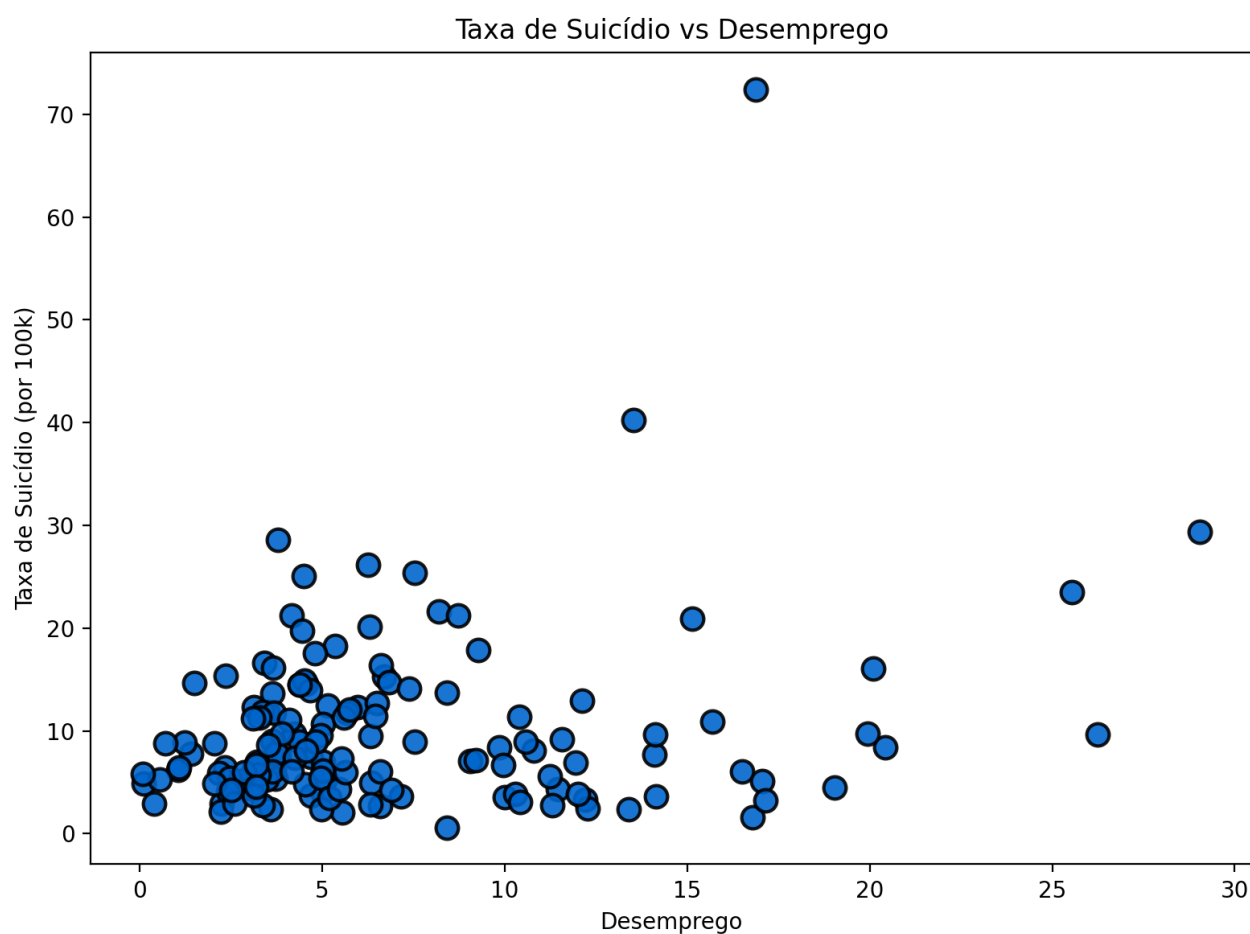
7.4.3. Análise do Índice de Desemprego per capita em relação à taxa de suicídio:

Figura 11: Distribuição Desemprego



Fonte: Autores.

Figura 12: Distribuição Desemprego em relação à taxa de Suicídio



Fonte: Autores.

Apesar do desemprego não ter obtido valores muito significantes como as demais variáveis, e ter sido considerada a variável menos relevante, a sua distribuição em relação ao suicídio ainda é presente e pode facilitar a prevenção do suicídio. O desemprego, possui uma correlação positiva em relação ao suicídio, indicando, portanto, que um aumento em sua taxa, está associado com o aumento do suicídio. Adicionalmente, o desemprego foi capaz de explicar de maneira mais eficiente a variabilidade do fenômeno em relação a outras variáveis, tornando assim sua consideração mais relevante.

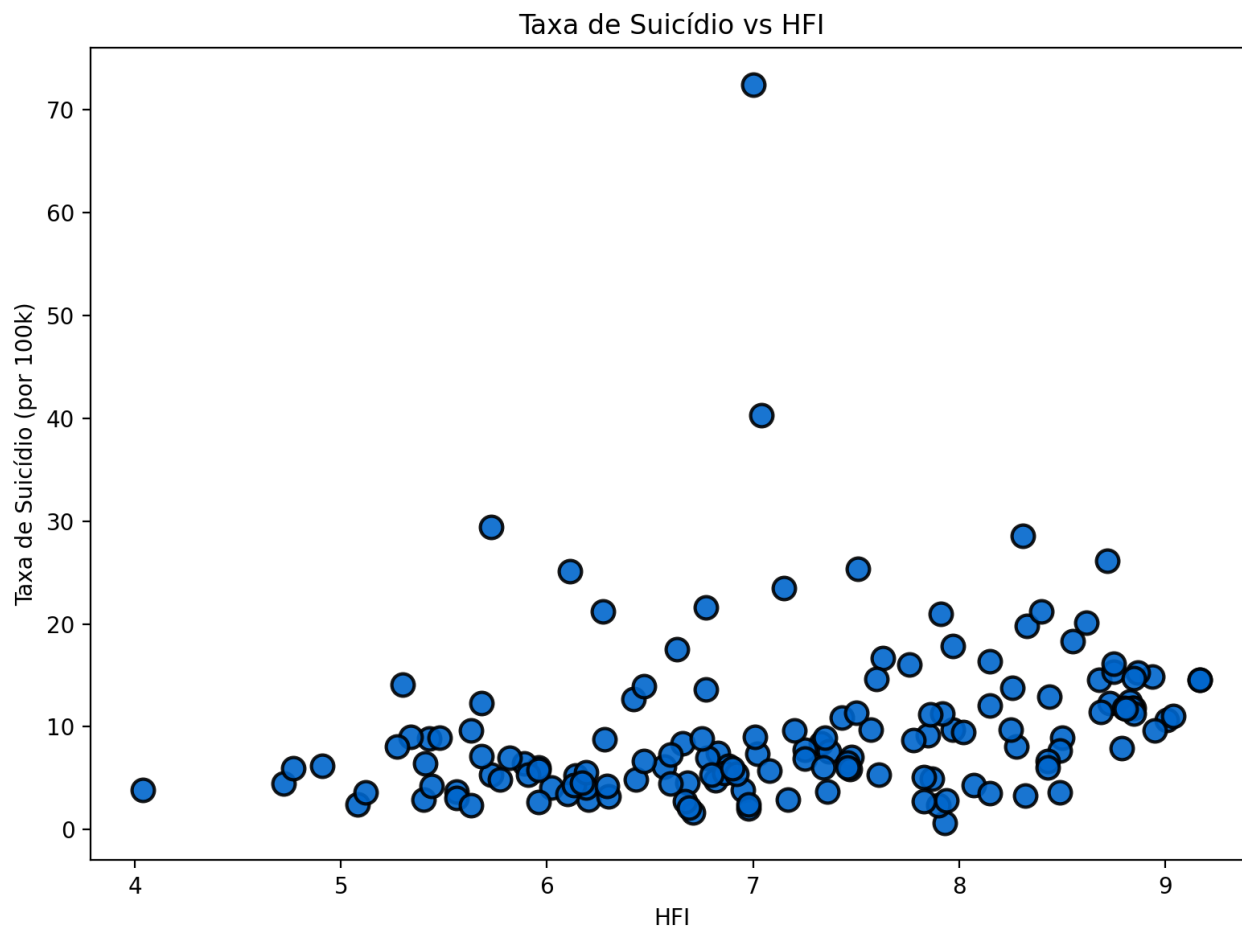
7.4.4. Análise do Índice de Liberdade Humana em relação à taxa de suicídio:

Figura 13. Distribuição Índice de Liberdade Humana (HFI)



Fonte: Autores.

Figura 14: Distribuição Índice de Liberdade Humana em relação à taxa de Suicídio



Fonte: Autores.

O HFI, de maneira semelhante ao PIB per capita e o IDH, está relacionado positivamente com o suicídio e se distribui de maneira próxima ao suicídio, o que indica que a consideração do HFI também é necessária a fim de garantir a diminuição das taxas de suicídio. É importante ressaltar, que apesar da similaridade com os índices citados, as distribuições são diferentes, e precisam ser feitas de maneira diferente. Adicionalmente, o HFI é mais relevante em relação as outras variáveis socioeconômicas, além de explicar a variabilidade do fenômeno melhor, portanto, necessita ter um enfoque maior.

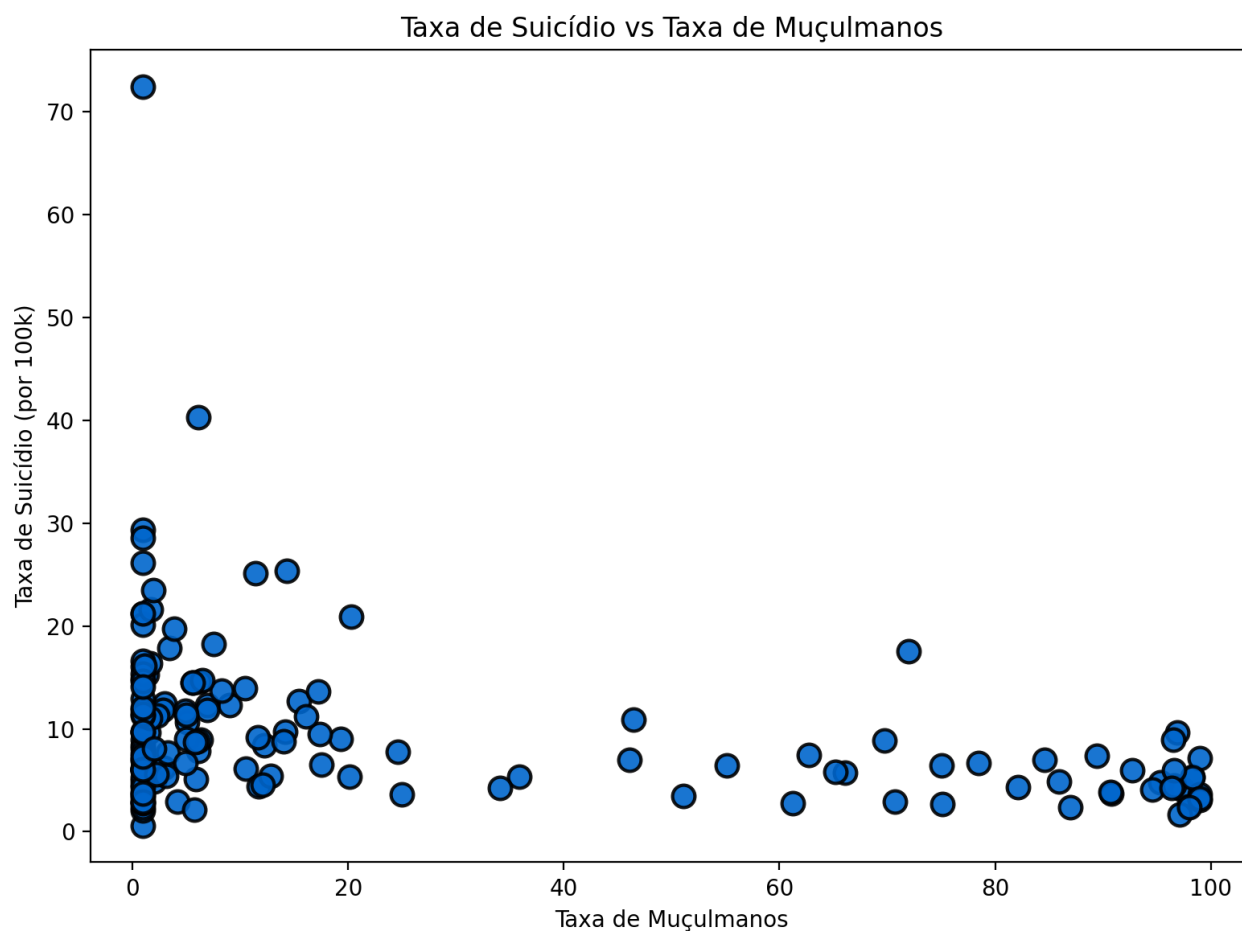
7.4.5. Análise da Taxa de Muçulmanos em relação à taxa de suicídio:

Figura 15: Distribuição Taxa de Muçulmanos



Fonte: Autores.

Figura 16: Distribuição Taxa de Muçulmanos em relação à taxa de suicídio

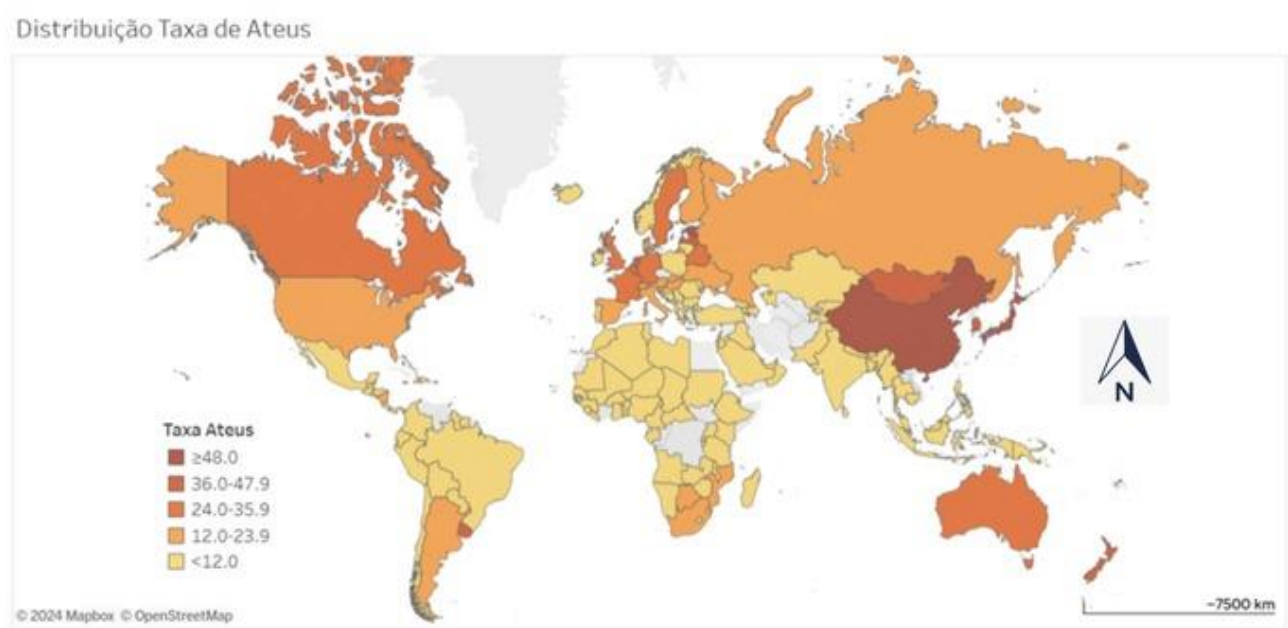


Fonte: Autores.

A partir da análise da taxa de muçulmanos, foi possível perceber a sua correlação negativa em relação ao suicídio, indicando assim, que países com uma taxa maior de pessoas afilhadas à essa religião possuem de maneira geral uma menor taxa de suicídio. Adicionalmente, a taxa de muçulmanos foi capaz de explicar a variabilidade do suicídio melhor em relação a outras variáveis consideradas mais relacionadas. Portanto, sua análise é relevante.

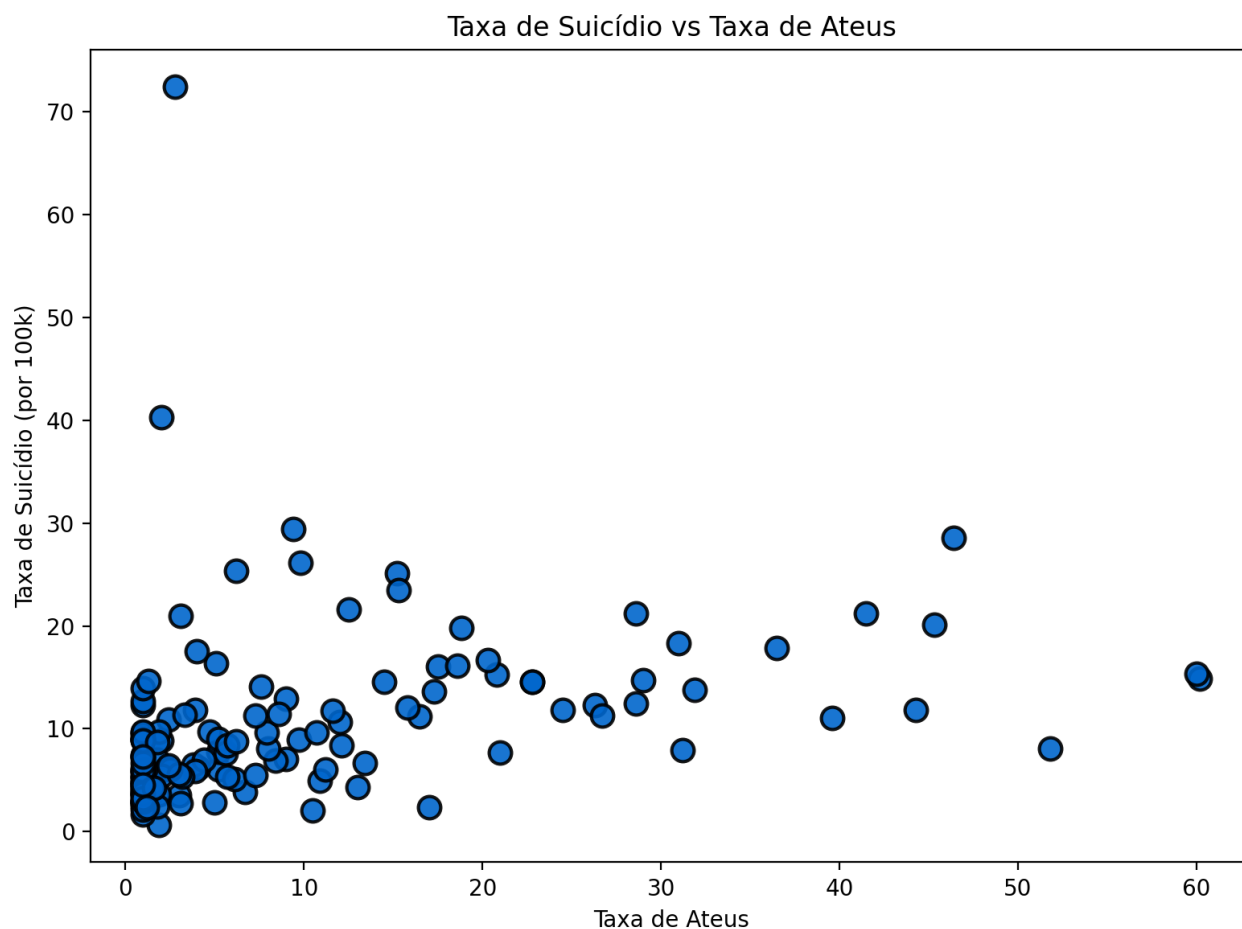
7.4.6. Análise da Taxa de Ateus em relação à taxa de suicídio:

Figura 17: Distribuição Taxa de Ateus



Fonte: Autores

Figura 18: Distribuição Taxa de ateus em relação à taxa de suicídio

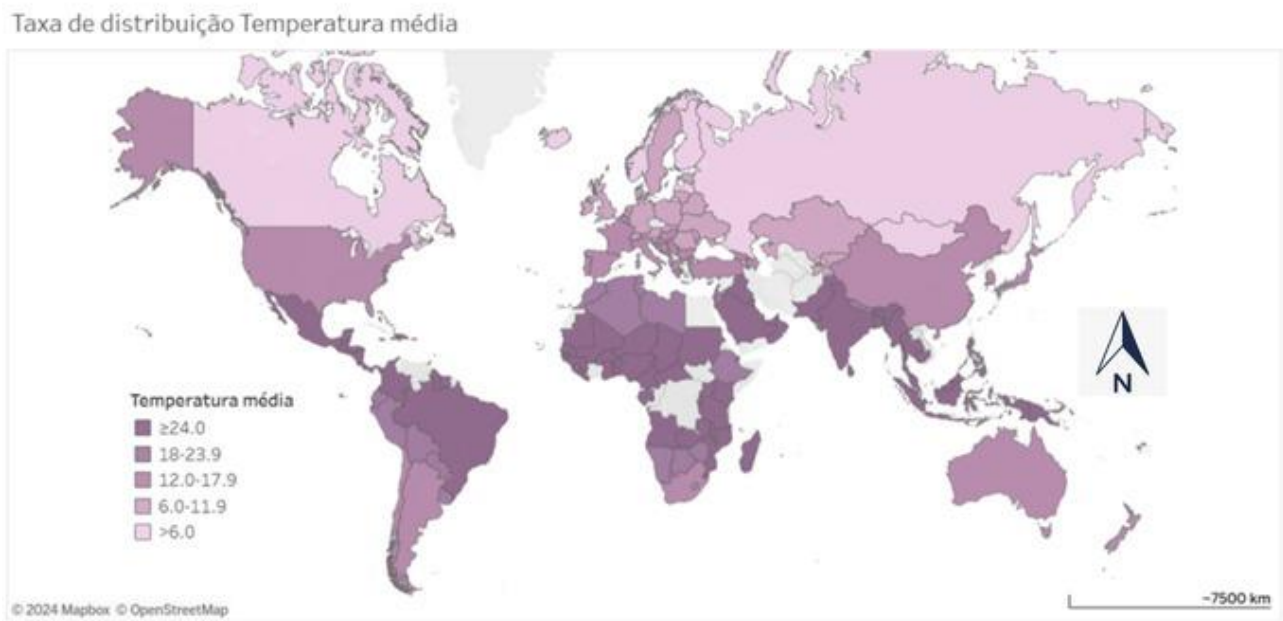


Fonte: Autores.

A partir da análise da taxa de ateus, foi possível observar uma correlação positiva com o suicídio. Sugerindo que os países com maiores proporções de ateus têm taxas de suicídio mais elevadas em geral. Essa porcentagem, é a variável mais relevante do estudo, além de explicar de maneira considerável o suicídio. Assim, a avaliação da distribuição de ateus é importante para considerar o risco do suicídio e preveni-lo.

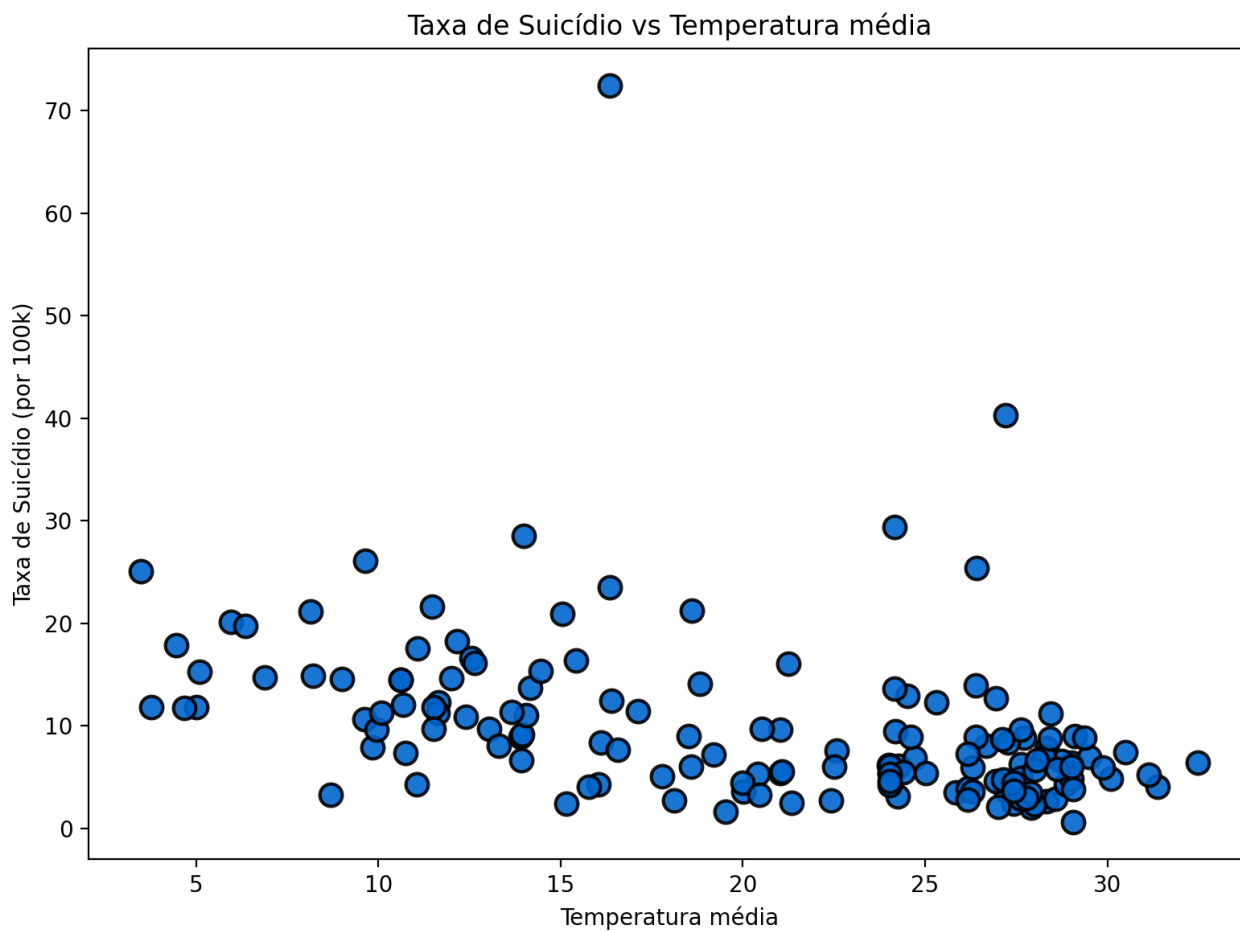
7.4.7. Análise do Índice de Temperatura em relação à taxa de suicídio:

Figura 19: Temperatura média



Fonte: Autores.

Figura 20: Distribuição Temperatura média em relação à taxa de suicídio

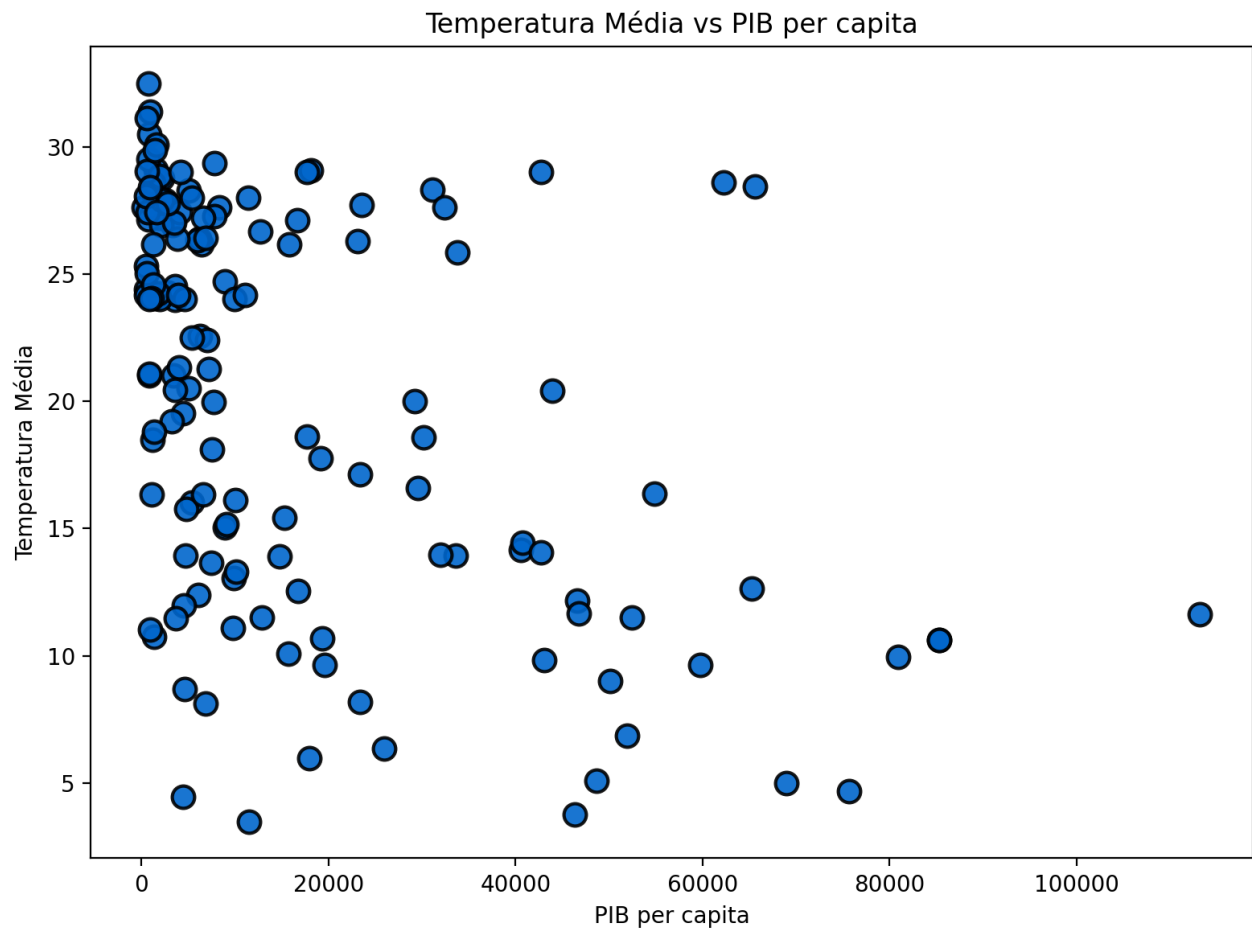


Fonte: Autores

A partir da análise, foi possível perceber a correlação negativa entre a temperatura média, indicando assim que a diminuição da temperatura está associada com o aumento do suicídio. A temperatura média, foi considerada a segunda variável mais relevante do estudo, além de ser útil em explicar a variância da problemática. Portanto, sua análise é fundamental.

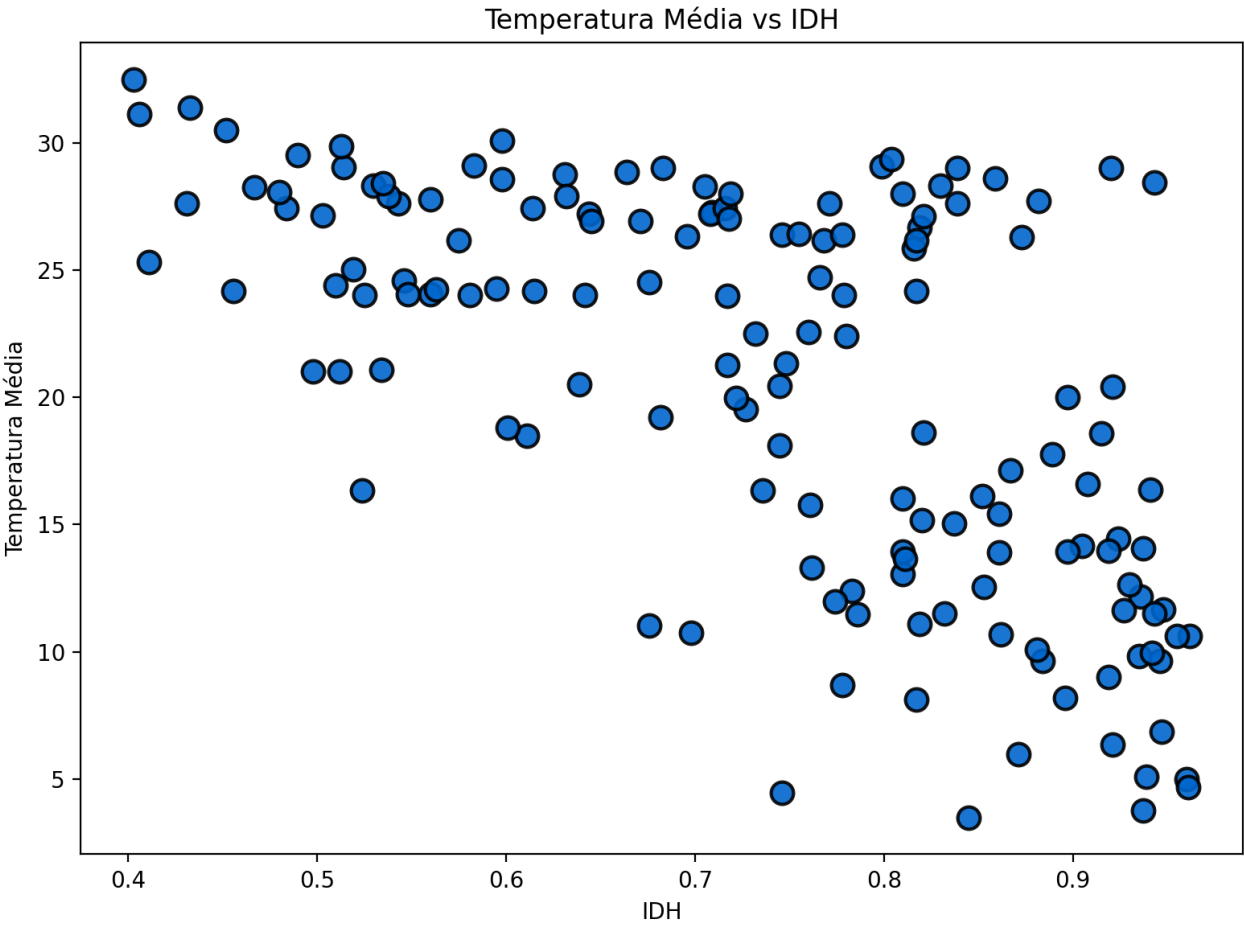
7.4.8. Análise cruzada da temperatura média em relação aos outros índices

Figura 21: Distribuição Temperatura média em relação à PIB per capita



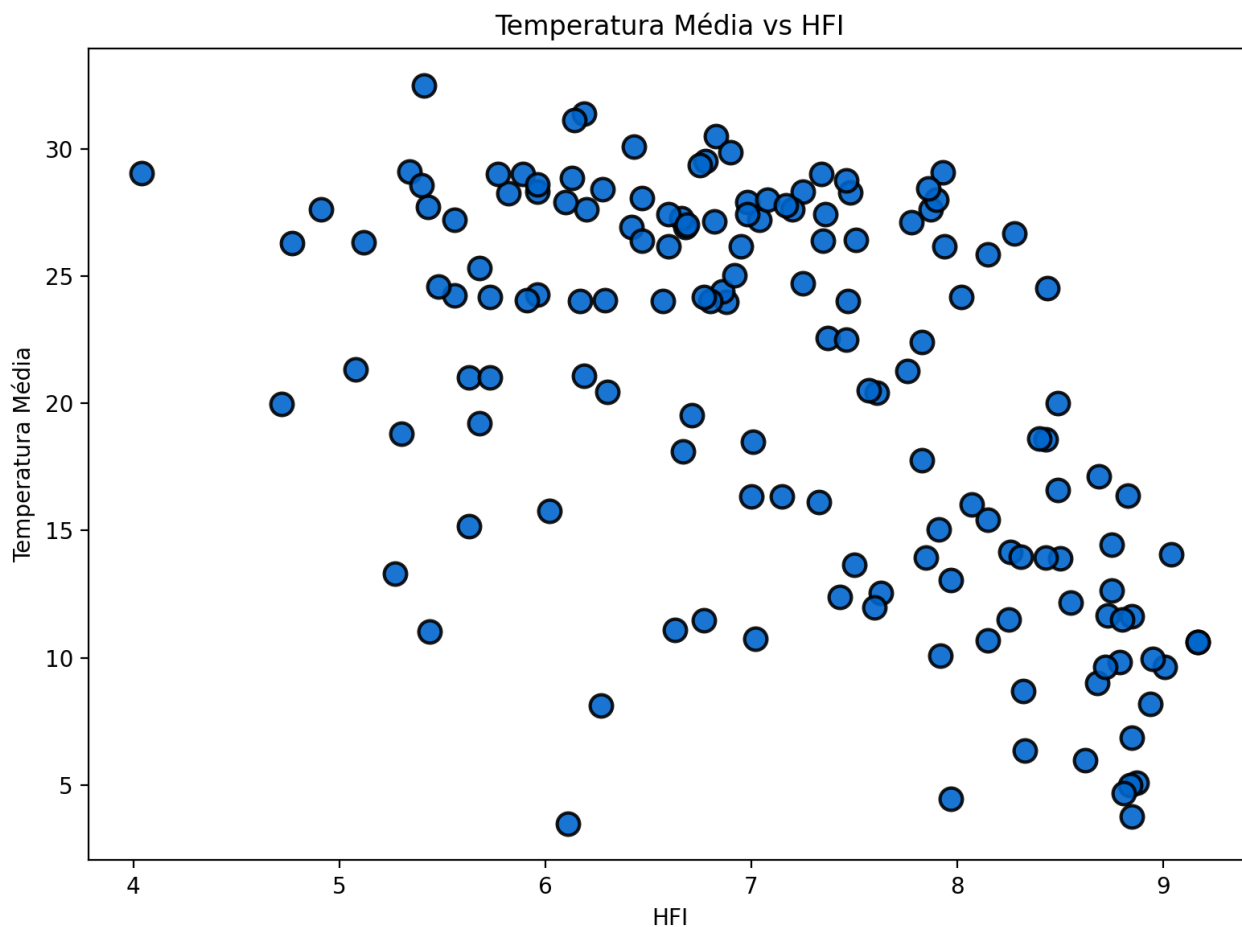
Fonte: Autores.

Figura 22: Distribuição Temperatura média em relação ao Índice de Desenvolvimento Humano



Fonte: Autores.

Figura 23: Distribuição Temperatura média em relação ao Índice de Liberdade Humana (HFI)

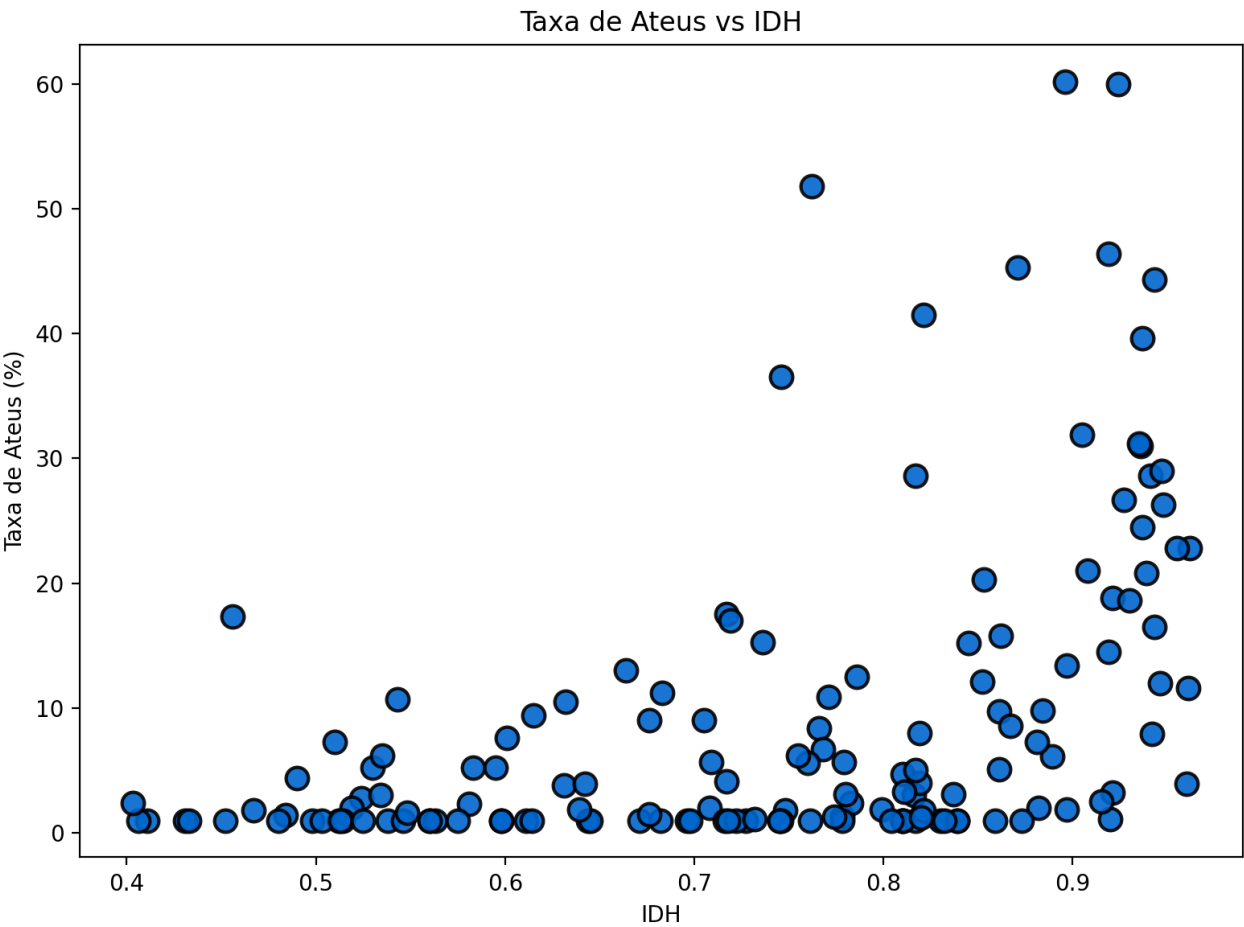


Fonte: Autores.

A partir desse cruzamento, fomos capazes de entender como os países com os índices altos (IDH, PIB per capita e HFI) possuem no geral uma temperatura média baixa. Indicando assim, que isso é um perfil de países com os índices altos, sugerindo assim, que a temperatura condiciona, ao menos que parcialmente, a correlação positiva entre as três variáveis e o suicídio, tendo em conta a relação temperatura baixa e a taxa do fenômeno.

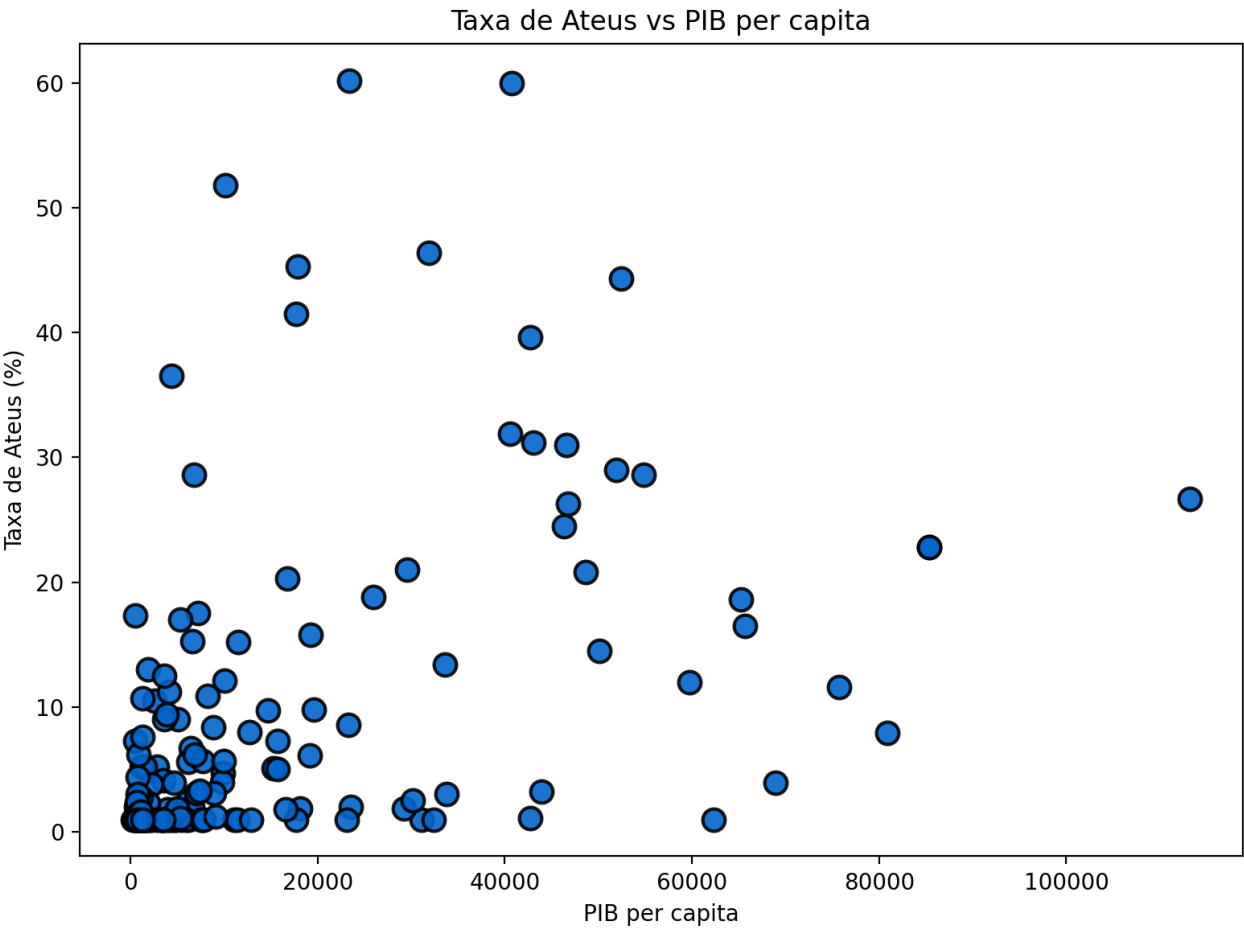
7.4.9. Análise cruzada da taxa de ateus em relação aos outros índices

Figura 24: Distribuição Taxa de Ateus em relação à Índice de Liberdade Humana



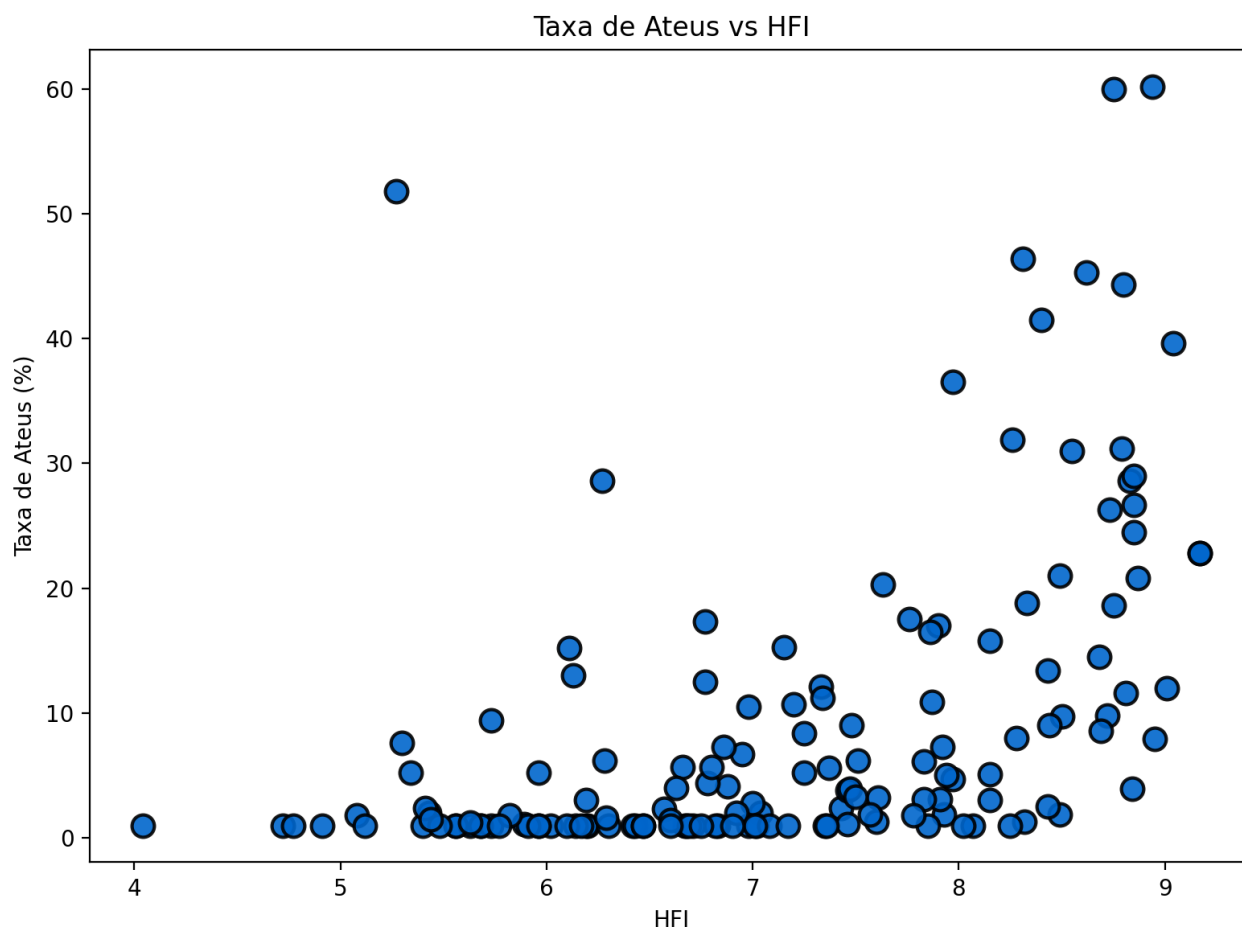
Fonte: Autores

Figura 25: Distribuição Taxa de Ateus em relação ao PIB per capita



Fonte: Autores

Figura 26: Distribuição Taxa de Ateus em relação ao Índice de Liberdade Humana



Fonte: Autores

De maneira semelhante ao cruzamento anterior, países com os três índices altos, possuem no geral uma taxa de ateus maior. Perfil esse que está associado com a taxa de suicídio,

7.4.10. Mapa de Análise de Risco

Após as análises individuais entre cada variável e o suicídio, decidimos empreender e criar um Mapa de Análise de Risco a fim de avaliar a proximidade dos nossos resultados em relação ao suicídio e concluir como nossos dados diferem da realidade e em que contexto isso acontece.

Assim, seguindo a metodologia de Rocha e Nunes (2020), utilizamos os resultados obtidos a partir dos métodos estatísticos para a criação de um mapa capaz de revelar *insights* mais profundos em relação ao fenômeno. De maneira similar ao estudo, avaliamos a relação entre as variáveis socioeconômicas e o suicídio para a criação da visualização, seguindo este processo:

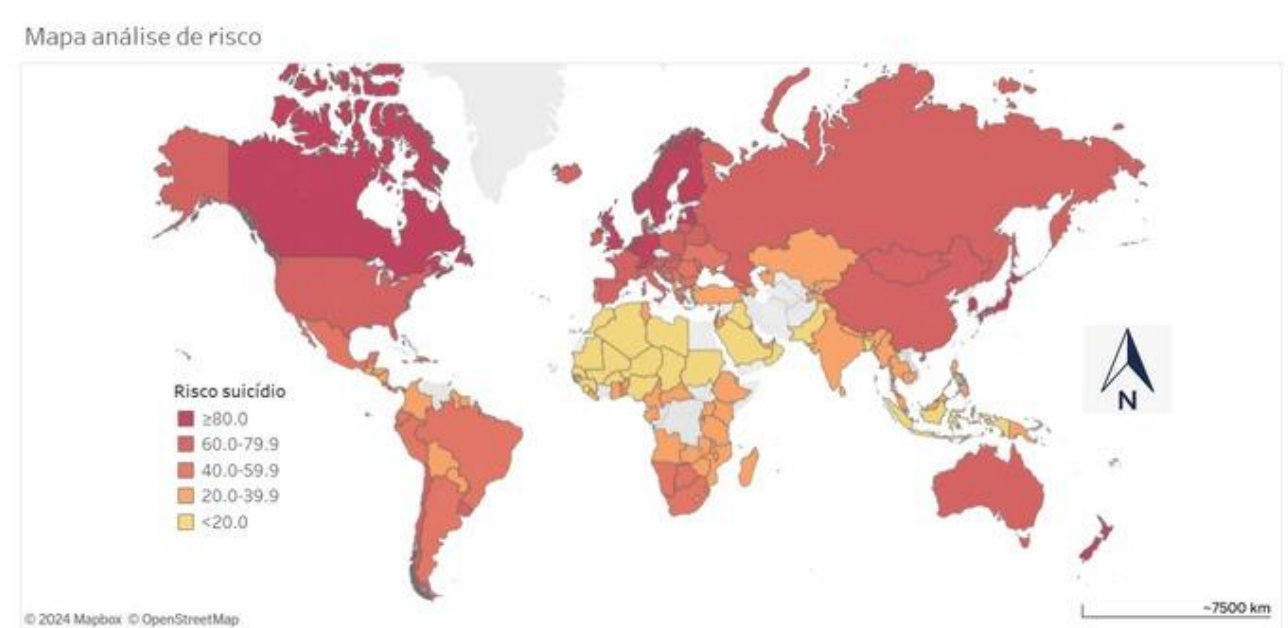
- Normalização dos dados, transformando os valores de cada variável em um número entre 0 e 100, para assim deixar o peso entre os dados igual.

- Criação de um coeficiente a partir do critério de Spearman e o *p-value* da regressão individual (coeficiente de Spearman dividido pelo *p-value*). Com o objetivo de atribuir peso avaliando risco e probabilidade.
- Multiplicação dos valores normalizados com o coeficiente e sua somatória a fim de obter a taxa de risco.
- Segunda normalização. Transformando os dados novamente em valores entre 0 e 100 para facilitar a compreensão dos resultados.
- Criação de 4 grupos definidos pela taxa de risco e confecção do mapa para a criação de insights.

O método, foi definido considerando o processo de confecção de análise de risco, que compreende a relação e probabilidade entre as variáveis e o suicídio.

Assim, confeccionamos o Mapa de Risco e o comparamos com o Mapa de Suicídio também dividido por grupo:

Figura 27: Mapa de risco



Fonte: Autores

Após a análise entre o mapa de risco, e o mapa com a taxa de suicídio concreta, foi possível analisar que, em certas regiões, a taxa de risco se aproxima do valor real, principalmente na Europa e América do Norte. Indicando, assim, a eficiência em parte do nosso projeto da análise do suicídio. Entretanto, em outras regiões como América do Sul e a África, o risco se afastou muito da realidade, mais especificamente em países como o Lesoto e Guiana, que apesar de serem os dois países com maior taxa de suicídio na pesquisa, foram classificados como de risco intermediário.

Portanto, a análise das variáveis e seus respectivos pesos são válidos e podem auxiliar na prevenção, entretanto, em casos mais específicos, principalmente nas regiões citadas, é necessário identificar questões mais específicas (sociais e culturais) que expliquem a taxa de suicídio exacerbada.

7.5. Discussão

Considerando diversas bibliografias que tratam sobre as variáveis socioeconômicas do suicídio, podemos realizar algumas comparações e reflexões sobre os resultados atingidos em nossa pesquisa, a seguir:

De forma contrária ao que se percebe no senso comum, a análise de dados nos leva à conclusão de que o maior desenvolvimento econômico dos países, que pode ser medido pelos seus PIB, não é um índice que indica necessariamente a diminuição da taxa de suicídios. Nesse sentido, os dados observados na Figura 10 nos levariam a questionar pesquisas que vinculam o aumento do suicídio com dificuldades de ordem econômica como a de Stack (2000) ou a de Innamorati (2010), Tamburello e Lester (2010) que associam as taxas de suicídio a índices macro e socioeconômicos, e cujos resultados indicam uma relação entre esses índices e o suicídios; bem como a de Chen, Choi e Sawada (2008), na qual afirmam que um nível de renda e o crescimento econômico elevado se associam à redução do suicídio.

Mas, então, como se justificam os nossos resultados? Associando-nos às percepções de Rodriguez et al. (2014), estudiosos que analisam pesquisas a favor e em oposição à associação positiva do suicídio aos índices econômicos elevados (melhor economia, mais suicídio), eles, à par que corroboram essa associação positiva, afirmam que ela pode ser consequência de dois fatores:

1. O debilitamento dos vínculos sociais dos indivíduos, que acompanham os processos de desenvolvimento social (estes autores fazem menção aos processos de modernização que Durkheim indicou como motivadores do suicídio). Nesse sentido, defendem que a necessidade de vínculos sociais é o resultado de um processo evolutivo prolongado, demonstrando que a sua falta provoca psicopatologias que podem provocar suicídio (2014).

2. Em segundo lugar mencionam o excesso de abundância existente nas sociedades mais desenvolvidas, situação que, se positiva em numerosos aspectos, por outro lado, seria contrária ao aprendizado evolutivo do ser humano, dado em contextos de escassez relativa e que fomentou as capacidades criativas para resolver situações adversas. Para os pesquisadores, o funcionamento biológico humano está configurado para resolver as problemáticas, e desde esse ponto de vista circuitos de recompensa e motivação produzem sensação de satisfação ao conseguir logros. Não

atingir conquistas, em relação a necessidades, sejam básicas ou abstratas, poderia fazer com que as pessoas não superem sensações de perda de sentido, o que poderia causar depressão e suicídio. Contudo, é importante destacar a advertência que fazem esses estudiosos, no sentido que é uma perspectiva evolutiva a que explica suas observações de uma forma complementar, sendo necessários outros estudos para provar essas hipóteses, a partir da melhora da qualidade dos registros em nível mundial, já que os dados utilizados possuem muitas vezes limitações estatísticas, assim como pelo fato da impossibilidade de individualizar os dados para cada caso.

No sentido de fazer uma discussão mais aprofundada, um outro gráfico obtido (Figura 18) mostra a correlação direta entre taxa de ateus e suicídio, indicando que a ausência de filiação a uma prática religiosa traz como consequência o aumento desse fenômeno. Autores como Peralta Gallego (2018), ao aplicar e analisar autópsias psicológicas, obteve resultados que confirmam a hipótese de que a religião tem um papel protetor em relação ao suicídio, indicando uma tendência a diminuir essa ação autodestrutiva à medida que aumenta a prática religiosa. Esse pesquisador, indica as manifestações de diferentes religiões em relação à prática suicida, sendo que os judeus e os cristãos a consideram proibida tanto pela Bíblia, quanto pelos mandamentos. Por outro lado, afirma a existência de uma tendência de mudança nessa ideia de condenar o suicídio por esses credos, isso porque foram sendo incorporadas perspectivas de perdão e compaixão divina e humana.

Já na religião muçulmana, aparece um mandado mais contundente em relação à sua proibição, indicando Peralta Gallego que a prática é considerada ilegal em vários países islâmicos. Todavia, esclarece que o suicídio também é sub-informado nessas nações, situações, ambas, que explicariam o resultado obtido na nossa análise, que indica uma baixa correlação entre religião muçulmana e casos de suicídio (Figura 16). Assim mesmo, Peralta Gallego indica um aspecto que transcende a religiosidade como um sistema do que é permitido ou proibido ou daquilo que é recompensado ou punido consequência de um comportamento incorreto: manifesta a importância da religião no sentido de pertencimento a uma comunidade o que reforça, refletimos, àquela ideia de Rodríguez et al. (2014) no sentido da importância dos vínculos sociais para evitar o suicídio.

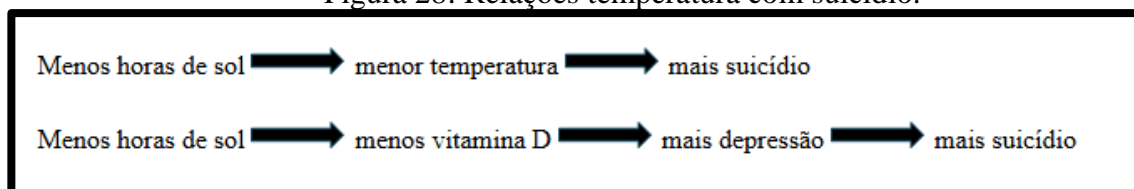
Acudindo ao trabalho de Lawrence, Oquendo e Stanley (2016), eles coincidem com Peralta Gallego sobre o fato de a filiação a uma religião proteger contra as tentativas de suicídio, embora, advertem, que a ideia suicida não é afetada por essa pertença. Assim ser religioso pode inibir a execução do suicídio, isto pelo acesso que a pessoa tem a uma comunidade de apoio, com crenças comuns, que promovem pensamentos esperançosos e formas de interpretar o sofrimento (nessas comunidades se procura dar um sentido ao sofrimento). Contudo, existem ressalvas. Os autores indicam que a relação entre religião e suicídio é complexa, porque, embora sejam percebidas as vantagens mencionadas, o fato de pertencer a filiações religiosas minoritárias pode provocar

sentimentos de isolamento e uma consequente depressão motivadora de iniciativa suicida. Também, em sociedades nas quais existe proibição ao exercício da religião as crenças são menos efetivas para evitar o suicídio, assim como se uma religião provoca culpa desmedida, ou se a pessoa se sente distante da divindade ou abandonada por sua comunidade, isso pode aumentar o risco. Resgatamos, do estudo de Lawrence, Oquendo e Stanley principalmente a ideia de que os resultados obtidos das pesquisas sobre a contribuição da religião podem ajudar a gerar ações educativas, assim como estimular para que os médicos ajudem os pacientes religiosos a acessar apoios que os confortem nas suas crenças. Também indicam a sua importância para fomentar nas autoridades das comunidades religiosas e nos seus fiéis a consciência sobre a utilidade da sua ação, com o objetivo de potencializá-la.

Uma outra variável que mereceu a nossa atenção é a relação entre temperatura e suicídio. A análise de dados (Figura 20) indica que existe uma relação direta: menor temperatura corresponde a uma maior taxa de suicídio, o que corroboraria a percepção popular sobre isso. Entretanto, ao aprofundar a questão em bibliografia acadêmica, não encontramos pesquisas científicas que confirmem essa relação, mas encontramos várias que indica o contrário: que o aumento de temperaturas pelo aquecimento global, e não a sua diminuição, pode ser associada ao incremento das taxas de mortalidade por suicídio (GRAXIOLA-ROBLES et al., 2013)

Analisando outras fontes de pesquisa, no intuito de explicar as nossas observações, encontramos um vínculo comprovado entre uma menor quantidade de horas de sol, com a baixa produção de vitamina D, e por isso com um aumento de depressão (MANZANOS et al. 2022), o que entendemos poderia provocar o aumento do suicídio. O anterior nos leva a realizar o seguinte raciocínio: a temperatura indicada nos dados analisados poderia ser resultante de uma maior ou menor incidência do sol com a consequente repercussão na produção de vitamina D. De ser assim se justificariam as seguintes relações:

Figura 28: Relações temperatura com suicídio.



Fonte: MANZANOS, et al. 2020. Organização: Autores.

Assim, observamos que a relação entre temperatura e suicídio não seria consequência direta da influência de uma sobre outro, mas do fato de que a menor temperatura e a menor produção de

vitamina D dependerem da quantidade de horas de sol, então consideramos que os resultados obtidos por nós deveriam ser corroborados ou refutados a partir de outras análises.

Por outro lado, a Figura 21 indica uma relação direta entre aumento do PIB e diminuição da temperatura para uma proporção importante de países. Pensamos que essa interação poderia ser indicadora de uma potencialização da taxa de suicídio, pois, embora percebamos que essa afirmação não se cumpre de forma completa, sim representa a realidade de uma boa parcela dos países. Entre as exceções se encontra o caso de Lesoto o país com maior índice de suicídio mundial que à sua vez possui um PIB baixo e uma temperatura alta. No caso de Lesoto, segundo manifesta o psiquiatra Taiwo Lateef, representante da ONG LifeLine na África, esse índice é consequência da estigmatização dessa prática por motivos sociais, religiosos e culturais. A prevenção do suicídio encontra barreiras importantes no continente africano, pois em 11 países essa prática é considerada ilegal, o que faz com que ao sofrimento da pessoa com ideação suicida se some a impossibilidade de solicitar ajuda, pelas consequências que isso teria para ela mesma e a sua família (LIFE LINE, 2024).

Um outro caso que se opõe ao vínculo entre PIB alto, baixa temperatura e alta taxa de suicídio é o dos países nórdicos. Como explica a matéria Minuto Saúde Mental do Jornal da USP (2021) as taxas nesses países, com exceção da Finlândia, estão por debaixo da média do continente europeu (MINUTO, 2021). Para o prof. João Paulo Machado de Souza, essa associação se deve a um preconceito consolidado a partir da década de 1960, quando o presidente norte-americano Eisenhower tentou-se vincular o *Welfare State*, política do Estado de Bem-Estar Social desenvolvido por Suécia, com uma alta taxa de suicídio, como forma de deslegitimar esse tipo de política, afirmação que nas décadas posteriores teve um extenso uso político. A discussão anterior indica como a problemática que nos preocupa é um fenômeno social multifatorial, como o próprio Souza menciona na matéria, complexa, dependente de múltiplas variáveis que deve ser cada vez mais estudados para desenvolver ações e estratégias para combatê-lo.

No Gráfico de Análise Risco (Figura 25), no qual consideramos todas as variáveis analisadas, que multiplicamos por um coeficiente de relação e comparamos com a taxa de suicídio verdadeira, comprovamos que para a Europa, a América do Norte e a Austrália, a taxa de Risco é semelhante à Taxa de Suicídio o que indica que nesses lugares existe coincidência com a realidade. Nada entanto, em outras regiões acontece o contrário, por exemplo em Lesoto, país com a maior taxa de suicídio, mas que apresenta em nosso gráfico uma taxa de Risco não tão alta. O anterior significa que para este país existem outras variáveis importantes, de ordem cultural por exemplo, que não foram consideradas em nossa pesquisa. No caso de Lesoto, Granjero, Moshoeshoe e Townsley (2024) indicam que trabalhadores de saúde desse país observam que as causas do suicídio são muitos complexas, entre elas mencionam as doenças mentais, as adições, mas também a sensação de perda,

a solidão, as dores crônicas e o abuso que a população suporta. A essas variáveis, que adquirem nuances particulares, se somam fatores únicos característicos do país, como a falta de perspectivas para os jovens, a migração a outros lugares onde existe trabalho e as tensões familiares de diversas origens, assim como o tabu que implica o tratamento ou o simples fato de admitir os problemas de saúde mental. O preconceito é tal, explicam os autores, citando a experiência do psicoterapeuta Liao Mochaba, que aqueles que sofrem esses problemas preferem escondê-los, nem sequer mencionando-os na terapia, pois não querem ser identificados “como loucos” (GRANJERO, MOSHOESHOE, TOWNSLEY, 2024), e até preferem consultar médicos tradicionais, antes que psiquiatras. O modelo de família na qual o homem é o encarregado de sustentá-la não colabora, pois impulsiona os indivíduos a emigrar para poder trabalhar, provocando sensação de abandono naqueles que ficam. Já para aqueles que não conseguem emigrar para trabalhar, a sensação de ser um fardo, sentimento que como já explicado pode estimular o suicídio, é muito presente.

Consequência de toda a discussão anterior, identificamos que a consideração das variáveis individuais nos estudos destinados a paliar a problemática do suicídio, não estudadas por nós como já explicamos, e das variáveis socioeconômicas e ambiental, foco do nosso trabalho, deve se somar a análise e interpretação das variáveis culturais. Essa incorporação, requer um trabalho particular para cada um dos países, no melhor dos casos para regiões com culturas parecidas, já que as culturas são diversas e provocam nuances que devem ser consideradas nas políticas de prevenção do suicídio. Considerar a diversidade cultural, embora faça mais complexa a análise, é fundamental, pois seu reconhecimento colabora no desenvolvimento de políticas públicas mais ajustadas à realidade dos indivíduos. Indicamos neste sentido a importância de estabelecer políticas de educação, conscientização e informação sobre a problemática, que devem não apenas permear a escala global, mais também as outras: é na escala mais próxima do indivíduo que as políticas das outras escalas reverberam com mais potência.

Considerando os temas discutidos e as comparações realizadas com outras pesquisas, fica claro a validade de nosso trabalho, levando em consideração as correlações obtidas que condizem com a realidade e o ambiente científico.

8. Considerações Finais

Levando em consideração a hipótese e objetivo do estudo, que consideram a possibilidade da utilização dos métodos estatísticos mencionados anteriormente, a fim de identificar correlações individuais e coletivas entre as variáveis socioeconômicas e o suicídio, foi possível cumprir as metas planejadas. Porém, discorreremos mais a frente, também, sobre as problemáticas enfrentadas no estudo tais como a falta de alguns dados e a ausência de literatura profissional em relação à escala global.

Sobre o objetivo e hipótese da pesquisa: fomos capazes de definir 6 variáveis (Desemprego, IDH, HIF, taxa de muçulmanos, taxa de ateus e temperatura média) relacionadas de maneira considerável com o suicídio, relação essa comprovada a partir do teste de Kruskal-Wallis, a correlação de Spearman e as regressões. Obtendo em algumas análises valores que indicavam a relevância das amostras, como os *p-values* menores que 0,05 conseguidos nas regressões, apesar de na correlação de Spearman nenhuma variável chegar a uma correlação forte (0.7 a 0.89) (Bruce; Bruce, 2020). A partir da análise coletiva conseguimos criar modelos capazes de prever de maneira precisa o suicídio, obtendo um *r* quadrático ajustado próximo a 0.42, indicando que o modelo consegue prever 42% a variabilidade do suicídio (LIRA, 2004), indicando assim a importância e qualidade dos dados na prevenção do fenômeno. Adicionalmente, obtemos dois cruzamentos relevantes (temperatura média e taxa de ateus em relação ao IDH, PIB per capita e HFI) que explicaram mais profundamente a distribuição do suicídio em relação às três variáveis; e tornaram o entendimento do suicídio por parte da pesquisa mais complexo. A partir desses resultados, fomos capazes de criar mapas de distribuições, *scatter plots* e um mapa de risco para facilitar a visualização das correlações. Os dois primeiros métodos (mapas de distribuição e *scatter plots*), possibilitaram visualizações mais claras de cada variável de maneira individual, enquanto o mapa de risco, mostrou como a análise feita pela pesquisa se difere da taxa de suicídio, evidenciando as limitações da pesquisa e a aplicação da análise das variáveis socioeconômicas por regiões.

Sobre o que fora descoberto em relação ao objeto do estudo, é necessário recapitular as seguintes questões já apresentadas na discussão: O suicídio, em relação às variáveis socioeconômicas no contexto global, está altamente relacionado a duas variáveis, temperatura média e a taxa de ateus, obtendo na maioria dos testes de *p-value* o dobro de correlação em relação às demais variáveis e na correlação de Spearman valores superiores 0.5 (Tabela 16 e Figura 5). Assim, levando em consideração a nossa pesquisa e a bibliografia relacionada já discutida, o uso dessas variáveis é fundamental para paliar o suicídio. Em relação aos tipos de variáveis estudadas, os valores econômicos relevantes são correlacionados de maneira positiva com o suicídio, principalmente aqueles que no senso comum já discutido são associados de maneira negativa com esse evento (IDH, PIB per capita e HFI). As variáveis religiosas, apesar de em sua maioria não terem associação

relevante com o suicídio, indicaram que a ausência de fé está associada com o aumento da taxa de suicídio, classificando a religião como fator protetor, porém de acordo com a bibliografia estudada, a ausência de fé não é um fator preponderante, mas sim a relação entre a religiosidade e fatores adjacentes como a sociabilidade que a religião permite, embasando aquilo que autores como Peralta Gallego (2018) e Lawrence, Oquendo e Stanley (2016) afirmam em suas pesquisas.

Além, disso a relação entre as variáveis e o suicídio nos cruzamentos entre os dados socioeconômicos, possibilitaram um entendimento mais profundo da distribuição do fenômeno, pois no contexto dos resultados obtidos, países com o IDH, PIB per capita e HIF maior possuem uma menor temperatura e maior taxa de ateus, duas características associadas com o aumento do suicídio, e que, respaldados pela pesquisa, indicam que essas distribuições fazem com que os três índices se correlacionem de maneira positiva com o suicídio.

Dessa maneira, tivemos nossa hipótese corroborada, já que conseguimos através de microdados alcançar metadados que exemplifique de modo claro e direto às correlações validadas por estudos pregressos. Assim, o mapa que se embasava na taxa de suicídio por 100 mil habitantes, apenas demonstra a complexidade e heterogeneidade dessa prática autodestrutiva entre países supostamente semelhantes com resultados dispares. Logo, nosso estudo demonstrou com nossas análises, outros panoramas dentro da escala global; e, também, permitindo fatores decisórios dentro desta abordagem para prevenção, com isso temos hoje um maior fator de compreensão inclusive para as outras abordagens de escalas de maior proximais.

Ainda que a pesquisa sobre a prevenção de suicídio seja algo que demande muito mais esforços e análises, o escopo do nosso projeto permitiu trazer contribuições inéditas, e com isso todo e qualquer esforço que resulte no salvamento de uma vida é mais do que válido e, acreditamos ainda que esse passo inicial permita resultados mais assertivos no futuro e muitas mais vidas salvas.

8.1 Futuro da pesquisa (curto e médio prazo)

Em função dos resultados e das expectativas que foram atingidas, observamos a possibilidade de evolução e inovação na análise e validação do projeto:

- Validação do projeto mediante questionário: Formularemos um questionário que será entregue à profissionais de diferentes áreas como Saúde e Políticas Públicas a fim de melhorar a aplicabilidade do projeto. Buscaremos entender que dados abordar, como analisá-los e como tratá-los com o objetivo de otimizar os resultados já encontrados e que serão futuramente formulados.

- Patente e publicações com *Creative Commons*, para garantir a distribuição gratuita e acessível da informação de nossa pesquisa.
- Disponibilização dos resultados: Compartilhamento dos métodos e técnicas em um repositório.
- Análise em escala nacional: Mudaremos a escala de abordagem de índice global para o índice nacional, buscando entender de maneira mais concreta o suicídio avaliando aspectos culturais próprios dos países. A princípio, consideraremos apenas o Brasil, mas também caso haja a possibilidade, averiguaremos outros países que tenham pesquisas que possamos usar como lastro de validação.

9. Bibliografia

ABDI, H. The Bonferonni and Šidák Corrections for Multiple Comparisons. In: Neil Salkind (Ed.) **Encyclopedia of Measurement and Statistics**. Thousand Oaks: Sage, 2007. Disponível em: <chrome-extension://efaidnbnmnibpcajpcglclefindmkaj/https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=8b4300f253644d49c778c037ee614b9cf42a908c>. Acesso em: 15 jul. 2024.

ALICANDRO, G.; MALVEZZI, M.; GALLUS, S.; LA VECCHIA, C.; NEGRI, E.; BERTUCCIO, P. Worldwide trends in suicide mortality from 1990 to 2015 with a focus on the global recession time frame. **International Journal of Public Health**, [S. l.], v. 64, n. 5, p. 785-795, Jun. 2019. DOI:10.1007/s00038-019-01219-y. Disponível em: <http://dx.doi.org/10.1007/s00038-019-01219-y>. Acesso em: 14 jun. 2024.

AMERICAN PSYCHOLOGICAL ASSOCIATION. **Ethical principles of psychologists and code of conduct**. 2017. Disponível em: <https://www.apa.org/ethics/code/ethics-code-2017.pdf>. Acesso em: 2 out. 2024.

ARENSMAN, E.; DE LEO, D.; SCOTT, V.; PIRKIS, J (E.ds.). **Suicide and Suicide Prevention From a** ALVES, J. E. **Os países mais religiosos são os mais pobres e os menos religiosos os mais ricos. Disponivel em:** <https://www.ihu.unisinos.br/categorias/188-noticias-2018/584060-os-paises-mais-religiosos-sao-os-mais-pobres-e-os-menos-religiosos-os-mais-ricos>. Acesso em 30 set. 2024.. Acesso em 30 set. 2024.

BANCO MUNDIAL. **Saúde Mental e Suicídio: Relatórios e Dados**. Disponível em: <https://www.worldbank.org/en/topic/health/brief/mental-health>. Acesso em: 18 jun. 2024.

BANERJEE, S. Human Development World Index [base de dados]. 2024. **Kaggle**. Disponível em: <https://www.kaggle.com/datasets/iamsouravbanerjee/human-development-index-dataset>. Acesso em: 27 ago. 2024.

BEAUTRAIS, A. Fatores de risco para suicídio e tentativa de suicídio entre jovens. **The Australian and New Zealand Journal of Psychiatry**, [Sidney], v. 34, n. 3, p. 420-436, Jun. 2000. DOI: 10.1080/j.1440-1614.2000.00691.x. Disponível em: <http://dx.doi.org/10.1080/j.1440-1614.2000.00691.x>. Acesso em: 15 jul. 2024.

BERNERT, R.; HILBERG, A.; MELIA R.; PAIK, J.; SHAH, N.; ABNOUSI, F.. Artificial Intelligence and Suicide Prevention: A Systematic Review of Machine Learning Investigations. **International Journal of Environmental Research Public Health**, [S. l.], v. 17, n. 16, p. 5929, Ago. 2020. DOI: 10.3390/ijerph17165929. Disponível em: <http://dx.doi.org/10.3390/ijerph17165929>. Acesso em: 18 jun. 2024.

BISHOP, Ch. **Pattern Recognition and Machine Learning**. Singapore: Springer, 2006.

- BRUCE, P.; BRUCE, A.; GEDECK, P. **Estatísticas práticas para cientistas de dados: mais de 50 conceitos essenciais usando R e Python**. [S.l.]: O'Reilly Media, 2020.
- CASSORLA, R. Em busca do objeto idealizado. **Revista Brasileira de Psicoanálise**, v. 53, n. 4, p. 49-65, 2019. Disponível em: <http://pepsic.bvsalud.org/pdf/rbp/v53n4/v53n4a04.pdf>. Acesso em: 4 jun. 2024.
- CHEN, J.; CHOI, Y.; SAWADA, Y. How is Suicide Different in Japan?. **Japan and the World Economy**, Tokio, v. 21, n. 2, p. 140-150, 2009. DOI:10.1016/J.JAPWOR.2008.06.001. Disponível em: <http://dx.doi.org/10.1016/J.JAPWOR.2008.06.001>. Acesso em: 15 mai. 2024.
- COENEN, A.; KIM, Y.; KRUK, M.; et al. Precisão diagnóstica de uma abordagem de aprendizado de máquina para reserva de fluxo fracionada baseada em angiografia tomográfica computadorizada coronária: Resultado do consórcio MACHINE. *Circulation: Cardiovascular Imaging*, v. 11, n. 6, 2018. Disponível em: <http://dx.doi.org/10.1161/circimaging.117.007217>.
- CORTOPASSI, L. E. Impacto da Liberdade Econômica e da Concentração de Renda na Qualidade de vida das Nações. 2000 – 2019. Disponível em: https://bdm.unb.br/bitstream/10483/36415/1/2022_LucasEickhoffCortopassi_tcc.pdf. Acesso em: 30 set. 2024.
- CRESWELL, J.. **A concise introduction to mixed methods research**. Thousands Oaks: Sage, 2015.
- DAFOE, W. Principles of exercise testing and interpretation. **The Canadian Journal of Cardiology**, v. 23, n. 4, p. 274. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2647882/>. Acesso em: 10 jul. 2024.
- DALLY, W.; LABONTE, F.; DAS, A.; HANRAHAN, P. Merrimac: supercomputing with streams. In: THE 2003 ACM/IEEE CONFERENCE ON SUPERCOMPUTING, 2003, [S. l.]. **Proceedings [...]**. p. 35. DOI: 10.1145/1048935.1050187. Disponível em: <http://dx.doi.org/10.1145/1048935.1050187>. Acesso em: 15 ago. 2024.
- DOMARADZKI, J. The Werther effect, the Papageno effect or no effect? A literature review. **International Journal of Environmental Research and Public Health**, v. 18, n. 5, p. 2396, 2021. Disponível em: <https://doi.org/10.3390/ijerph18052396>. Acesso em: 2 jul. 2024.
- DURKHEIM, E. **O suicídio**. Estudo de Sociologia. São Paulo: Martins Fontes, 2000. Disponível em: https://www.academia.edu/26063553/DURKHEIM_%C3%89mile_O_Suic%C3%ADdio. Acesso em: 23 jun. 2024.
- FET, A; KNUDSON, H.; KEITSCH, M. Sustainable Development Goals and the CapSEM Model. In: Fet, A.M. (org.). **Business Transitions: A Path to Sustainability**. Cham: Springer, 2023.

https://doi.org/10.1007/978-3-031-22245-0_3. Disponível em: https://link.springer.com/chapter/10.1007/978-3-031-22245-0_3#DOI . Acesso em: 11 jul. 2024.

FEW, S. *Show Me the Numbers: Designing Tables and Graphs to Enlighten*. Oakland: Analytics Press, 2012.

FIELD, A. **Discovering Statistics Using IBM SPSS Statistics**. London: Sage Publications, 2017. Disponível em: <https://www.scirp.org/reference/referencespapers?referenceid=3309174>. Acesso em: 11 jul. 2024.

FIGUEIREDO FILHO, D.; SILVA JUNIOR, J. Desvendando os mistérios do coeficiente de correlação de Pearson (r). **Revista Brasileira de Ensino de Física**, [S.l.], v. 18, n. 1, p. 69-79, 2009. Disponível em: <https://periodicos.ufpe.br/revistas/politica hoje/article/view/3852>. Acesso em: 3 jul. 2024.

FIGUEROA, R.; ZENG-TREITLER, Q.; KANDULA, S.; NGO, L. Predicting sample size required for classification performance. **BMC Medical Informatics and Decision Making**, v. 12, p. 8, 2012. DOI: 10.1186/1472-6947-12-8. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/22336388/>. Acesso em: 15 ago. 2024.

FREEDMAN, D., PISANI, R.; PURVES, Roger. **Statistics**. New York: Norton, 2007.

FRANKLIN, J.; HUANG, X.; BASTIDAS, D. Virtual reality suicide: Development of a translational approach for studying suicide causes. **Behavior Research and Therapy**, v. 117, p. 82-89, 2019. DOI:10.1016/j.brat.2018.12.013. Disponível em: <https://doi.org/10.1016/j.brat.2018.12.013>. Acesso em: 14 jul. 2024.

GAXIOLA-ROBLES, R.; CELIS, A.; LABRADA-MARTAGON, V.; DIAZ-CASTRO, S.; ZENTENO, T. Incremento de la temperatura ambiental y su posible asociación al suicidio en Baja California Sur (BCS) 1985-2008. **Salud Mental**, México, v. 36, n. 5, set. 2013. Disponível em: <https://www.redalyc.org/pdf/582/58228970011.pdf>. Acesso em: 2 out. 2024.

GIL, A. **Como elaborar projetos de pesquisa**. São Paulo: Atlas, 2002. Disponível em: <https://docente.ifrn.edu.br/mauriciofacanha/ensino-superior/redacao-cientifica/livros/gil-a.-c.-como-elaborar-projetos-de-pesquisa.-sao-paulo-atlas-2002./view>. Acesso em: 26 mai. 2024.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep learning**. Cambridge: The MIT Press, 2016.

GOOGLE. Google Colaboratory. Disponível em: <https://colab.research.google.com/>. Acesso em: 30 mai.2024.

GUNNELL, D.; FRANKEL, S. Prevention of suicide: aspirations and evidence. **British Medical Journal**, [S. l.], v. 308, n. 6938, p. 1227-1233, mai. 1994. DOI: 10.1136/bmj.308.6938.1227. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2540097/>. Acesso em: 4 out. 2024.

GUNNELL, David; LEWIS, Glyn. Studying suicide from the life course perspective: Implications for prevention. **The British Journal of Psychiatry**, [S. l.], v. 187, p. 206–208, set. 2005. DOI: [10.1192/bjp.187.3.206](https://doi.org/10.1192/bjp.187.3.206). Disponível em: <http://dx.doi.org/10.1192/bjp.187.3.206>.

Acesso em: 26 mai. 2024.

HAIR, J.; BLACK, W.; BABIN, B.; ANDERSON, R. **Multivariate Data Analysis: A Global Perspective**. Upper Saddle River: Pearson, 2010. Disponível em: <https://www.scirp.org/reference/referencespapers?referenceid=3309174>. Acesso em: 11 jul. 2024.

HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. **The Elements of Statistical Learning: Data Mining, Inference, and Prediction**. Springer Series in Statistics. New York: Springer, 2009. Disponível em: <https://www.sas.upenn.edu/~fdiebold/NoHesitations/BookAdvanced.pdf>. Acesso em: 15 ago. 2024.

HEGERL, U.; ALTHAUS, D.; SCHMIDTKE, A.; NIKLEWSKI, G. The alliance against depression: 2-year evaluation of a community-based intervention to reduce suicidality. **Psychology Medicine**, [S. l.], v. 36, n. 9, p. 1225-1233, mai. 2006. DOI: [10.1017/S003329170600780X](https://doi.org/10.1017/S003329170600780X). Disponível em: <https://dx.doi.org/10.1017/S003329170600780X>. Acesso em: 14 jun. 2024.

HEGERL, U.; WITTMANN, M.; ARENSMAN, E.; VAN AUDENHOVE, C.; BOULEAU, J. The ‘European Alliance Against Depression (EAAD)’: A multifaceted, community-based action programme against depression and suicidality. **The World Journal of Biological Psychiatry**, [S. l.], v. 9, n. 1, p. 51- 58, 2008. Disponível: <http://www.mentalhealthpromotion.net/resources/hegerl-et-al-2008.pdf>. Acesso em: 12 jun. 2024.

HEGERL, U.; WITTENBURG, L. Focus on Mental Health Care Reforms in Europe: The European Alliance Against Depression: A Multilevel Approach to the Prevention of Suicidal Behavior. **Psychiatric Services**, [S. l.], v. 60, n. 5, p. 596-599, mai. 2009. DOI: [10.1176/ps.2009.60.5.596](https://doi.org/10.1176/ps.2009.60.5.596). Disponível em: <https://doi.org/10.1176/ps.2009.60.5.596>. Acesso em: 15 ago. 2024.

HAWTON, K.; ZAHL, D.; WEATHERALL, R. Suicide following deliberate self-harm: long-term follow-up of patients who presented to a general hospital. **British Journal of Psychiatry**, Londres, v. 182, p. 537-542, jun. 2003. DOI: [10.1192/bjp.182.6.537](https://doi.org/10.1192/bjp.182.6.537). Disponível em: <https://doi.org/10.1192/bjp.182.6.537>. Acesso em: 14 jun. 2024.

HONCHAR, O. et al. A prognostic model and pre-discharge predictors of post-COVID-19 syndrome after hospitalization for SARS-CoV-2 infection. **Frontiers in Public Health**, v. 11, p. 1276211, 2023. Acesso em: 17 jul. 2024.

INNAMORATI, M; TAMBURELLO, A; LESTER, D. Inequalities in suicide rates in the European Union's elderly: trends and impact of macro-socioeconomic factors between 1980 and 2006. **Canadian Journal of Psychiatry**, [S. l.], v. 55, n. 4, 2010; p. 55:229–38. DOI:

10.1177/070674371005500406. Disponível em: <https://doi.org/10.1177/070674371005500406>. Acesso em: 24 jul. 2024.

ISLAM, S. Global Inflation Dataset. Annual inflation rate of 196 countries (1980- 2024) [base de dados]. 2024. **Kaggle**. Disponível em: <https://www.kaggle.com/datasets/sazidthe1/global-inflation-data>. Acesso em: 20 set. 2024.

ITAMARATY. Agenda 2030. Disponível em: http://www.itamaraty.gov.br/images/ed_desenvsust/Agenda2030-completo-site.pdf. Acesso em: 5 mai. 2024.

JAIN, A. **Algorithms for clustering data**. Englewood Cliffs, N. J.: Prentice Hall, 1988.

JAIN, A. Data clustering: 50 years beyond K-means. **Pattern Recognition Letters**, [S. l.], v. 31, n. 1, jun. 2010, p. 651-666.

JOINER, T. **Why people die by suicide**. Cambridge: Harvard University Press, 2005.

KREUZE, E.; JENKINS, C.; GREGOSKI, M.; YORK, J.; MUELLER, M.; LAMIS, D.; RUGGIERO, K. Technology-enhanced suicide prevention interventions: A systematic review. **Journal of Telemedicine and Telecare**, [S. l.], v. 23, n. 6, p. 605–617, 2017. DOI: 10.1177/1357633x16657928. Disponível em: <https://doi.org/10.1177/1357633x16657928>. Acesso em: 24 jul. 2024.

LAKATOS, E.; MARCONI, M. **Fundamentos de metodologia científica**. São Paulo: Atlas, 2003. Disponível em: https://docente.ifrn.edu.br/olivianeta/disciplinas/copy_of_historia-i/historia-ii/china-e-india/view. Acesso em: 25 mai. 2024.

LANE, D.; SCOTT, D.; HEBL, M.; GUERRA, R.; OSHERSON, D.; ZIMMER, H. **Introduction to statistics**. 2003. Disponível em: <https://citeseerx.ist.psu.edu/documentrepid=rep1&type=pdf&doi=a8182909e4e95c530f4849ba75f74c7c633651b0>. Acesso em: 15 ago. 2024.

LAWRENCE, R.; OQUENDO, M.; STANLEY, B. Religion and Suicide Risk: A Systematic Review. **Archives of Suicide Research**, [S. l.], v. 20, n. 1, p. 1-21, 2016. DOI: 10.1080/13811118.2015.1004494. Disponível em: <https://doi.org/10.1080/13811118.2015.1004494>. Acesso em: 25 mai. 2024.

LIFE LINE INTERNATIONAL. Why kind of man cries? The country with the highest suicide rate in the world. 2024. Disponível em: <https://lifeline-international.com/es/lesotho-looking-for-hope-behind-the-highest-suicide-rate-in-the-world/>. Acesso em: 1 out. 2024.

LIRA, S. Análise de correlação: abordagem teórica e de construção dos coeficientes com aplicações. 2004. Dissertação (Mestrado em Estatística) – Universidade Federal do Paraná, Curitiba, 2004.

LITTLE, Roderick; RUBIN, Donald. **Statistical Analysis with Missing Data**. Hoboken: Wiley, 2019, DOI:10.1002/9781119482260. Disponível: <https://onlinelibrary.wiley.com/doi/book/10.1002/9781119482260>. Acesso em: 14 jul 2024.

MAHLER, D.; SERAJUDDIN, U.; MAEDA, H. When is there enough data to create a global statistic? **World Bank Blogs**, [S. l.], 10 maio 2022. Disponível em: <https://blogs.worldbank.org/en/opendata/when-there-enough-data-create-global-statistic>. Acesso em: 15 jul. 2024.

MANN, J. Suicide prevention strategies: a systematic review. **JAMA**, Chicago, v. 294, n. 16, p. 2064-2074, 26 out. 2005. Disponível em: <https://jamanetwork.com/journals/jama/article-abstract/201761>. Acesso em: 10 out. 2024.

MANZANOS, I.; MARTINO, P.; AUDISIO, E.; BONET, J. Vitamin D: Between the Brightness of the sun and the Darkness of Depression. **Revista Colombiana de Psiquiatria**, Bogota, v. 51, n. 3, 2022. DOI: 10.1016/j.rcp.2020.08.002. <https://doi.org/10.1016/j.rcp.2020.08.002>. Acesso em: 25 mai. 2024.

MEJIA, J.; PIVATTO, V. Mortalidade por suicídio no Brasil: análise temporal (2010-2021) e comparação com os primeiros dois anos da pandemia de COVID-19. **Revista Ciências Saúde**, [S. l.], v. 13, n. 3, p. 31-39, set. 2023. Disponível em: <https://pesquisa.bvsalud.org/portal/resource/pt/biblio-1510847>. Acesso em: 12 jul. 2024.

MCKINNEY, W. **Python para análise de dados: tratamento de dados com Pandas, NumPy e IPython**. São Paulo: Novatec, 2023.

MATSUBAYASHI, T.; SEKIJIMA, K.; UEDA, M. Government spending, recession, and suicide: evidence from Japan. **BMC Public Health**, [S. l.], v. 20, n. 1, p. 243, 2020. DOI: 10.1186/s12889-020-8258-0. Disponível em: <https://doi.org/10.1186/s12889-020-8258-0>. Acesso em: 14 jun. 2024.

MCLOUGHLIN, A.; GOULD, M.; MALONE, K.. Global trends in teenage suicide: 2003-2014. **Review QJM**, [S. l.], v. 108, n. 10, p765-80, out. 2015. DOI: 10.1093/qjmed/hcv026. Epub 2015. Disponível em: <https://doi.org/10.1093/qjmed/hcv026>. Epub 2015. Acesso em: 15 ago. 2024.

MILLNER A., ROBINAUGH, D.; NOCK, M. Advancing the Understanding of Suicide: The Need for Formal Theory and Rigorous Descriptive Research. **Trends Cogn Sci**, [S. l.], 24, n. 9, p. 704-716, 2020. DOI: 10.1016/j.tics.2020.06.007. Disponível: <https://doi.org/10.1016/j.tics.2020.06.007>. Acesso em: 14 jun. 2024.

MINUTO SAUDE MENTAL. 9: É mito que as taxas de suicídio são maiores nos países nórdicos. In: **Jornal da USP**, São Paulo, 8 abr. 2021. Disponível em: <https://jornal.usp.br/podcast/minuto-saude->

mental-9-e-mito-que-as-taxas-de-suicidio-sao-maiores-nos-paises-nordicos/. Acesso em: 2 out. 2024.

MIRANDA, S.; CASTRO, R. Uso do Google Scholar na pesquisa acadêmica. **Revista Brasileira de Biblioteconomia e Documentação**, v. 18, n. 1, p. 15–26, 2022. Disponível em: <https://revistas.ufg.br/rbbd/article/view/59125>. Acesso em: 5 jul. 2024.

MITCHELL, S.; CERO, I.; LITTLEFIELD, A. K.; & BROWN, S. L. Using categorical data analyses in suicide research: Considering clinical utility and practicality. **Suicide & Life-Threatening Behavior**, 51(1), 76-87, 2021. DOI: <https://doi.org/10.1111/sltb.12670>.

MOTILLON-TOUDIC, C.; WALTER, M.; SÉGUIN, M.; CARRIER, J.; BERROUGUET; LEMEY, C. Social isolation and suicide risk: Literature review and perspectives. **European psychiatry: the journal of the Association of European Psychiatrists**, v. 11, n. 65. DOI: 10.1192/j.eurpsy.2022.2320. Disponível em: <https://doi.org/10.1192/j.eurpsy.2022.2320>. Acesso em: 25 mai. 2024.

MÜLLER, J.; JUNG, M.; HILLER, M.; KALISZ, G. Suicidal behavior and socio-economic factors in Europe: A systematic review. *Social Psychiatry and Psychiatric Epidemiology*, v. 57, n. 8, p. 1401–1420, 2022. Disponível em: <https://doi.org/10.1007/s00127-021-02163-y>. Acesso em: 6 jul. 2024.

OMS. Relatório Mundial sobre Prevenção do Suicídio. 2021. Disponível em: https://www.who.int/mental_health/suicide-prevention/en/ Acesso em: 2 jun. 2024.

OMS. Suicide worldwide in 2019 Global Health Estimates. Disponível em: <https://iris.who.int/bitstream/handle/10665/341728/9789240026643-eng.pdf?sequence=1>. 2019. Acesso em: 2 jun. 2024.

OS PAISES mais e menos religiosos do planeta. **Instituto Humanitas Unisinos - IHU**. Disponível em: <https://ihu.unisinos.br/categorias/169-noticias-2015/541803-os-paises-mais-e-menos-religiosos-do-planeta>. Acesso em: 1 out. 2024.

OSTERTAGOVA, E. Modelling using Polynomial Regression. **Procedia Engineering**, [S. l.], v. 48, 2012, p. 500-506. DOI: <https://doi.org/10.1016/j.proeng.2012.09.545>.

PAIXÃO, G. M. M.; SANTOS, B. C.; ARAUJO, R. M.; RIBEIRO, M. H.; MORAES, J. L.; RIBEIRO, A. L. Machine Learning na Medicina: Revisão e Aplicabilidade. **Arquivos Brasileiros de Cardiologia**, v. 118, n. 1, p. 95-102, jan. 2022. DOI: 10.36660/abc.20200596

PANT, A. Unemployment dataset [base de dados]. 2022. **Kaggle**. Disponível em: <https://www.kaggle.com/datasets/pantanjali/unemployment-dataset>. Acesso em: 23 set. 2024.

PARKER, R. **Survey research methods**. 4. ed. Los Angeles: Sage Publications, 2014. Disponível em: <https://www.sagepub.com/en-us/nam/survey-research-methods/book246659>. Acesso em: 20 jul. 2024.

PAULA, G. A.. **Modelos de regressão: com apoio computacional**. São Paulo: IME-USP, 2004. Disponível em: https://www.ime.unicamp.br/~cnaber/Livro_MLG.pdf. Acesso em: 14 jul. 2024.

PEARCE, S. Introduction to Fisher (1925) Statistical Methods for Research Workers. In: KOTZ, S., JOHNSON, N.L. (eds) **Breakthroughs in Statistics**, [S.l.], Springer Series in Statistics. New York: Springer, 1992. DOI: [10.1007/978-1-4612-4380-9_5](https://doi.org/10.1007/978-1-4612-4380-9_5). Disponível em: https://doi.org/10.1007/978-1-4612-4380-9_5. Acesso em: 25 mai. 2024.

PERALTA GALLEGO, F. Estudio de la relación entre suicidio y religión. **Interpsiquis, XIX Congreso Virtual Internacional de Psiquiatría**, 2018. Disponível em: [https://psiquiatria.com/congresos/pdf/1-1-2018-1-pon5\[1\].pdf](https://psiquiatria.com/congresos/pdf/1-1-2018-1-pon5[1].pdf). Acesso em: 15 ago. 2024.

POPPER, K. **The Logic of Scientific Discovery**. Londres- New York: Routledge, 2002. Disponível em: <https://philotextes.info/spip/IMG/pdf/popper-logic-scientific-discovery.pdf>. Acesso em: 6 jul. 2024.

PYTHON SOFTWARE FOUNDATION. PSF Annual Report. 2020. Disponível em: <https://www.python.org/psf/annual-report/2020/>. Acesso em: 14 set. 2024.

R CORE TEAM. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Viena, 2017. Disponível em: <https://www.R-project.org/>. Acesso em: 2 out. 2024

RAJPUT, D.; WANG, W.; CHEN, C. Evaluation of a decided sample size in machine learning applications. **Bioinformatics**, Londres, v. 24, n. 1, 2023. DOI: [10.1186/s12859-023-05156-9](https://doi.org/10.1186/s12859-023-05156-9). Disponível em: <https://doi.org/10.1186/s12859-023-05156-9>. Acesso em: 15 ago. 2024.

RODRÍGUEZ, J; SANTANDER, J.; AEDO, A.; ROBERT, S. Suicidio en el mundo: ¿puede un enfoque evolutivo explicar la relación entre tasas de suicidio y variables asociadas a calidad de vida?. **Revista Chilena de Neuro-psiquiatría**, Santiago, v. 52, n. 4, dec. 2014. DOI: [10.4067/S0717-92272014000400002](https://doi.org/10.4067/S0717-92272014000400002). Disponível em: <http://dx.doi.org/10.4067/S0717-92272014000400002> . Acesso em: 2 out. 2024.

RODRIGUES, F. **Metodologia da Pesquisa Científica**. São Paulo: Atlas, 2016. Disponível em: <https://docente.ifrn.edu.br/paulofreitas/ensino-superior/livros/metodologia-da-pesquisa-cientifica/view>. Acesso em: 9 jul. 2024.

RUSSELL, S.; NORVIG, P. Artificial intelligence: a modern approach. **Bioinformatics**, [S. l.], v. 14, n. 24, p. 48, feb. 2023. DOI: [10.1186/s12859-023-05156-9](https://doi.org/10.1186/s12859-023-05156-9). Disponível em: <https://doi.org/10.1186/s12859-023-05156-9>. Acesso em: 14 jun. 2023.

SALDANHA, F.; PINTO, R. **A análise de dados no estudo de caso: uma abordagem prática. Revista Brasileira de Metodologia**, v. 6, n. 2, p. 124–139, 2021. Disponível em: <https://www.scielo.br/j/ramsci/a/JK8tPL7WjtHcRjZy3P4rVxL/?lang=pt>. Acesso em: 15 jul. 2024.

SCHNEIDER, G. The Human Freedom Index. A global measurement of personal, civil and economic freedom [base de dados]. 2022. **Kaggle**. Disponível em: <https://www.kaggle.com/datasets/gsutters/the-human-freedom-index>. Acesso em: 20 set. 2024.

SEDAGHATI, A. The weather of 187 Countries [base de dados]. 2023. **Kaggle**. Disponível em: [https://www.kaggle.com/datasets/amirhoseinsedaghati/the-weather-of-187-countries-in-](https://www.kaggle.com/datasets/amirhoseinsedaghati/the-weather-of-187-countries-in-2020?select=the+weather+of+187+countries+in+2020.csv)

2020?select=the+weather+of+187+countries+in+2020.csv. Acesso em: 27 ago. 2024. STACK, S.

Suicide: A 15 year review of the sociological literature. Part 1: cultural and economic factors. **Suicide Life Threatment Behaviour**, [S. l.], v. 30, p. 145-6, 2000. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/10888055/>. Acesso em: 15 ago. 2024.

SILVA, J.; OLIVEIRA, M.; PEREIRA, C. Utilização dos testes estatísticos Kruskal-Wallis e Mann-Whitney para avaliação de sistemas de solos reforçados com geotêxteis. *Matéria*, Rio de Janeiro, v. 27, n. 2, 2022. DOI: 10.1590/1517-7076-RMAT-2021-45351. Disponível em: <https://doi.org/10.1590/1517-7076-RMAT-2021-45351>. Acesso em: 2 out. 2024.

SILVA, M. H. **Análise estatística para pesquisadores**. Porto Alegre: Bookman, 2015. Disponível em: <https://www.bookman.com.br/livro/analise-estatistica-para-pesquisadores>. Acesso em: 12 jul. 2024.

SINGH, U. World Religion Projections 2010 to 2050. Estimate Religious composition of 198 countries and territories for 2010 to 2050 [base de dados]. 2023. **Kaggle**. Disponível em: <https://www.kaggle.com/datasets/utkarshx27/world-religion-projections>. Acesso em: 27 ago. 2024.

SINYOR, M.; TSE, R.; & PIRKIS, J. Global trends in suicide epidemiology. **Current Opinion in Psychiatry**, [S. l.], v. 30, n. 1, p. 1-6, jan. 2017. DOI: 10.1097/YCO.0000000000000296. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/27845946/>. Acesso em: 6 jul. 2024.

STACK, S. Contributing factors to suicide: Political, social, cultural and economic. **Preventive Medicine**,

TAS, O. World GDP (GDP, GDP per capita, and annual growths) [base de dados]. 2021. **Kaggle**. <https://www.kaggle.com/datasets/zgrcemta/world-gdpgdp-gdp-per-capita-and-annual-growths>. Acesso em: 20 set. 2024.

TUFTE, E. *Seeing with Fresh Eyes: Meaning, Space, Data, Truth*. Cheshire: Graphics Press, 2020.

TURECKI, G.; BRENT, D.; GUNNELL, D.; O'CONNOR, P; OQUENDO, M.; PIRKIS, J.; STANLEY, B.. Suicide and suicide risk. **Nature Reviews**, [S. l.], v. 24; n. 5, out. 2019. DOI: 10.1038/s41572-019-0121-0. Disponível em: <http://dx.doi.org/10.1038/s41572-019-0121-0>. Acesso em: 25 mai. 2024.

WASSERMAN, D.; RIHMER, D.; RUIESCU, M.; SARCHIAPONE, M. The European Psychiatric Association (EPA) guidance on suicide treatment and prevention. **European Psychiatry**, v. 27, n. 2, p. 129- 141, feb. 2012. DOI: <http://dx.doi.org/10.1016/j.eurpsy.2011.06.003>. Epub 2011 Dec 1.

WICKHAM, H.; G., Garrett. **R for data science: import, tidy, transform, visualize, and model data**. Beijin: O'Reilly, 2017.

WORLD HEALTH ORGANIZATION. **Preventing suicide: a global imperative**. Disponível em: http://apps.who.int/iris/bitstream/10665/131056/1/9789241564779_eng.pdf. Acesso em: 6 jul. 2024.

WYMAN, P.; BROWN, H.; MURRAY, L.; SCHMEELK-CONE, M; WALSH, X; WANG, W. An outcome evaluation of the Sources of Strength suicide prevention program delivered by adolescent peer leaders in high schools. **American Journal of Public Health**, [S. l.], v. 100, n. 9, p. 1653-61, set. 2010. DOI: 10.2105/AJPH.2009.190025. Epub 2010 Jul 15.

WILSON, J. A. **Linear regression: Theory and applications**. 2. ed. New York: Wiley, 2017. Disponível em: <https://www.wiley.com/en-us/Linear+Regression%3A+Theory+and+Applications%2C+2nd+Edition-p-9781119309682>. Acesso em: 8 jul. 2024.

ZAR, J. Spearman Rank Correlation. **Encyclopedia of Biostatistics**, [S. l.], 15 July 2005. DOI: <https://doi.org/10.1002/0470011815.b2a15150>. Acesso em: 14 jun. 2024.

10. Anexo

Base de dados com os países e código em r:

<https://onedrive.live.com/?authkey=%21AOKeBr4V0Axta%2DI&id=DC1AA098DEA58C98%21317635&cid=DC1AA098DEA58C98>