

**UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ  
DEPARTAMENTO ACADÊMICO DE ENGENHARIA DE PRODUÇÃO  
DEPARTAMENTO ACADÊMICO DE ENGENHARIA DE ELÉTRICA  
ENGENHARIA DE PRODUÇÃO  
ENGENHARIA ELÉTRICA**

**CAIO FERREIRA  
VINICIOS ROBERTO SOARES**

**ANÁLISE DE IMPACTO DOS FATORES ASSOCIADOS AO COMBATE  
E CONTENÇÃO DO SARS-COV-2 UTILIZANDO REDES BAYESIANAS**

**TRABALHO DE CONCLUSÃO DE CURSO**

**PONTA GROSSA  
2021**

**CAIO FERREIRA  
VINICIOS ROBERTO SOARES**

## **ANÁLISE DE IMPACTO DOS FATORES ASSOCIADOS AO COMBATE E CONTENÇÃO DO SARS-COV-2 UTILIZANDO REDES BAYESIANAS**

Trabalho de Conclusão de Curso apresentado como requisito para obtenção do título de Bacharel em Engenharia de Produção, Engenharia Elétrica, do Departamento Acadêmico de Engenharia de Produção, Departamento Acadêmico de Engenharia Elétrica, da Universidade Tecnológica Federal do Paraná (UTFPR).

Orientadora: Prof.<sup>a</sup> Dr.<sup>a</sup> Marcella S. Ribeiro Martins

Coorientador: Prof. Dr. Everton Melo



[4.0 Internacional](https://creativecommons.org/licenses/by-nc-nd/4.0/)

Esta licença permite download e compartilhamento do trabalho desde que sejam atribuídos créditos ao(s) autor(es), sem a possibilidade de alterá-lo ou utilizá-lo para fins comerciais. Conteúdos elaborados por terceiros, citados e referenciados nesta obra não são cobertos pela licença.

**PONTA GROSSA**

**2021**

**CAIO FERREIRA**  
**VINICIOS ROBERTO SOARES**

**ANÁLISE DE IMPACTO DOS FATORES ASSOCIADOS AO COMBATE E  
CONTENÇÃO DO SARS-COV-2 UTILIZANDO REDES BAYESIANAS**

Trabalho de Conclusão de Curso de Graduação apresentado como requisito para obtenção do título de Bacharel em Engenharia de Produção, Engenharia Elétrica, nos cursos de Engenharia de Produção, Engenharia Elétrica da Universidade Tecnológica Federal do Paraná (UTFPR).

Data de aprovação: 08/dezembro/2021

---

Hugo Valadares Siqueira  
Doutor em Engenharia Elétrica  
Universidade Tecnológica Federal do Paraná

---

Marcella Scoczynski Ribeiro Martins  
Doutora em Engenharia Elétrica e Informática Industrial  
Universidade Tecnológica Federal do Paraná

---

Virginia Helena Varotto Baroncini  
Doutora em Engenharia Elétrica e Informática Industrial  
Universidade Tecnológica Federal do Paraná

**PONTA GROSSA**  
**2021**

Dedicamos este trabalho às mulheres das  
nossas vidas, que sem o suporte e o amor  
delas, nada disso teria acontecido. Mara,  
Francisca, Emília e Geralda, nossa eterna  
dívida e gratidão.

## **AGRADECIMENTOS**

Este trabalho não poderia ser terminado sem a ajuda da nossa orientadora, Marcella, não só pela capacidade de nos ensinar e nos colocar no caminho certo, mas como a sua capacidade absoluta de amar e se doar a tantos, inclusive a nós. E por isso, pelo resto de nossas vidas, todos nós seremos gratos e devedores.

A nossas famílias, pela paciência, pelos anos de "mas falta muito ainda?", pela dor de não estarmos juntos em todos os aniversários, natais, festas. À minha mãe, Mara por todo o suporte em todos estes anos, mesmo com nossas dificuldades; avó Emília por todas as ligações de saudade; e irmão Douglas por ter me ajudado a dar o primeiro passo para me matricular na UTFPR e ter me incentivado a seguir meu caminho.

À Viviane por ter me acolhido e tratado como um filho mais velho e ter me ajudado tanto nos momentos mais difíceis.

Sem a ajuda dos nossos amigos, principalmente do Jhonatan por todo o suporte ao longo de todos estes meses. Ao Kiska, por toda a ajuda na vida nos últimos anos, que não foram fáceis (e pela geladeira cheia). Do começo ao fim estamos juntos, das necessidades mais brandas às mais difíceis e nada mudará isso.

Ao Roy e à Riza, porque sei que tudo que faço é por vocês.

Enfim, a todos os que de alguma forma contribuíram para a realização deste trabalho.

## RESUMO

FERREIRA, Caio; SOARES, Vinicios Roberto. **Análise de impacto dos fatores associados ao combate e contenção do Sars-CoV-2 utilizando Redes Bayesianas.** 2021. 61 f. Trabalho de Conclusão de Curso (Bacharelado em Engenharia de Produção, Engenharia Elétrica) – Universidade Tecnológica Federal do Paraná. Ponta Grossa, 2021.

A pandemia do Covid-19 é causada pela síndrome respiratória aguda grave coronavirus 2, ou SARS-CoV-2. Alertada pelo governo chinês em meados de 2020, o vírus que foi primeiramente identificado em Wuhan, na China, causou pânico global, assim como uma corrida contra o tempo para encontrar métodos plausíveis de se frear a proliferação da doença. A partir de março de 2020, a OMS, Organização Mundial da Saúde, declarou oficialmente que o vírus foi elevado ao nível de pandemia. O presente trabalho tem como objetivo principal enumerar em dados estes métodos e, via uma rede bayesiana, que utiliza da relação entre eventos anteriores a um evento posterior já conhecido - que no caso é a mortalidade que a doença causa como evento-fim - para identificar quais métodos utilizados possuem relação direta com a eficácia na minimização de mortes. O conjunto de dados que será utilizado é relativo aos países Alemanha, Brasil, Chile, Estados Unidos e Finlândia. Como resultados encontrou-se que na falta de vacina específica, a política mais eficaz é baseada no *lockdown*. Mesmo com a vacina disponível abrir mão de outras políticas de saúde pública não é o ação ótima para evitar um novo surto.

**Palavras-chave:** Covid-19. Inteligência Computacional. Rede Bayesiana. *Soft Computing*. Teorema de Bayes.

## ABSTRACT

FERREIRA, Caio; SOARES, Vinicios Roberto. **Impact Analysis of Factors associated with combat and containment of Sars-CoV-2 using Bayesian Networks.** 2021. 61 p. Bachelor Thesis (Bachelor's Degree in Industrial Engineering, Electrical engineering) – Universidade Tecnológica Federal do Paraná. Ponta Grossa, 2021.

The Covid-19 pandemic is caused by the severe acute respiratory syndrome coronavirus 2, or SARS-CoV-2. Alerted by the Chinese government in a press release of 2020, the virus which was first identified in Wuhan, China, has caused global panic as well as a race against time to find plausible methods to curb the spread of the disease. On March, 2020, based on these data, WHO, World Health Organization, officially stated that the virus was elevated to pandemic level. The main objective of this present work is to enumerate these methods in data and, via a Bayesian network, which uses the relationship between events prior to an already known subsequent event - which in this case will be the mortality that the disease causes as an end-event - to identify which methods have a direct relationship with the effectiveness in minimizing deaths. The utilized data will be used in relation to the countries of Brazil, Chile, Germany, Finland and the United States of America. As results, it was found that in the absence of a specific vaccine, the most effective policy is based on *lockdown*. Even with the vaccine available, giving up other public health policies is not the optimal action to prevent a new outbreak.

**Keywords:** Covid-19. *Computational Intelligence. Bayesian Network. Soft Computing. Bayes' Theorem.*

## LISTA DE ALGORITMOS

Algoritmo 1 – Algoritmo para Construção de Redes Bayesianas . . . . .	27
---	----

## LISTA DE ILUSTRAÇÕES

Figura 1 – Estrutura de um Vírus. . . . .	13
Figura 2 – O Sars-Cov-2 . . . . .	18
Figura 3 – Linha do tempo do começo da pandemia do COVID-19 . . . . .	18
Figura 4 – Total de Casos em milhões do COVID-19 no Brasil . . . . .	19
Figura 5 – Evolução da vacinação no Brasil até dezembro de 2021 (em milhões)	21
Figura 6 – Técnicas da Inteligência Computacional . . . . .	22
Figura 7 – O conceito englobado de Inteligência Artificial e Inteligência Compu- tacional . . . . .	24
Figura 8 – O conceito básico de um <i>Supervised Learning</i> . . . . .	25
Figura 9 – Exemplo de Fluxo de uma rede bayesiana . . . . .	26
Figura 10 – Delineamento da Pesquisa . . . . .	33
Figura 11 – Etapas . . . . .	35
Figura 12 – Grafo Acíclico direcionado . . . . .	35
Figura 13 – Probabilidade relacionada aos fatores e eventos exemplificado . . .	36
Figura 14 – Elementos <i>boxplot</i> . . . . .	39
Figura 15 – Boxplot dados Brasil . . . . .	39
Figura 16 – Boxplot dados . . . . .	40
Figura 17 – Grafo 2020 . . . . .	44
Figura 18 – Grafo 2021 . . . . .	46
Figura 19 – Histograma dados Brasil . . . . .	56
Figura 20 – Histograma dados . . . . .	57

## LISTA DE TABELAS

Tabela 1 – Delineamento da Pesquisa . . . . .	33
Tabela 2 – Exemplificação dos dados . . . . .	37
Tabela 3 – Descrição dos estados das variáveis . . . . .	37
Tabela 4 – Variáveis Auxiliares . . . . .	42
Tabela 5 – Categorias de diferença de morte agrupadas . . . . .	46
Tabela 6 – Probabilidade <i>contact_tracing</i> . . . . .	58
Tabela 7 – CPT <i>income_support contact_tracing</i> . . . . .	58
Tabela 8 – CPT <i>stay-at-home income_support</i> . . . . .	58
Tabela 9 – CPT <i>new_deaths stay-at-home</i> . . . . .	58
Tabela 10 – CPT <i>testing_policy income_support</i> . . . . .	58
Tabela 11 – CPT <i>face_coverings income_support</i> . . . . .	59
Tabela 12 – Probabilidade <i>vaccination_policy</i> . . . . .	59
Tabela 13 – CPT <i>people_vaccinated vaccination_policy</i> . . . . .	59
Tabela 14 – CPT <i>deaths-million_difference people_vaccinated</i> . . . . .	60
Tabela 15 – CPT <i>face_coverings deaths-million_difference</i> . . . . .	61
Tabela 16 – CPT <i>stay-at-home deaths-million_difference</i> . . . . .	61
Tabela 17 – CPT <i>testing_policy deaths-million_difference</i> . . . . .	61
Tabela 18 – CPT <i>contact_tracing deaths-million_difference</i> . . . . .	61
Tabela 19 – CPT <i>income_support deaths-million_difference</i> . . . . .	61

## LISTA DE ABREVIATURAS, SIGLAS E ACRÔNIMOS

### SIGLAS

AIDS	Síndrome da Imunodeficiência Adquirida, do inglês <i>Acquired Immunodeficiency Syndrome</i>
CL	Aprendizagem de conceito, do inglês <i>Concept learning</i>
CPT	Tabela de probabilidade conjunta, do inglês, <i>Conditional Probability Table</i>
DNA	Ácido desoxirribonucléico, do inglês <i>Deoxyribonucleic Acid</i>
GAD	Gráfico acíclico direcionado
GPL	Licença Pública Geral, do inglês <i>General Public License</i>
HIV	Vírus da imunodeficiência humana, do inglês <i>Human Immunodeficiency Virus</i>
IA	Inteligência artificial
IC	Inteligência computacional
LSTM	do inglês ( <i>Long Short-Term Memory</i> )
PGM	Modelo Probabilístico Gráfico, do inglês <i>Probabilistic Graphic Model</i>
OMS	Organização Mundial da Saúde, do inglês <i>World Health Organization</i> WHO
RNA	Ácido ribonucléico, do inglês <i>Ribonucleic Acid</i>
SARS	Síndrome respiratória aguda grave, do inglês <i>Severe Acute Respiratory Syndrome</i>
SL	Aprendizado Supervisionado, do inglês <i>Supervised Learning</i>
UTFPR	Universidade Tecnológica Federal do Paraná

### ACRÔNIMOS

COVID-19	doença causada pela síndrome respiratória aguda grave coronavírus 2
CoVs	coronavírus
OxCGRT	Rastreador Governamental de Resposta ao Coronavirus de Oxford, do inglês <i>Oxford Coronavirus Government Response Tracker</i>
Sars-CoV-2	síndrome respiratória aguda grave coronavirus 2

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>13</b>
1.1	CONTEXTUALIZAÇÃO DO PROBLEMA	13
1.2	PROBLEMA DE PESQUISA	15
1.3	OBJETIVOS	15
1.3.1	OBJETIVO GERAL	15
1.3.2	OBJETIVOS ESPECÍFICOS	15
1.4	JUSTIFICATIVA	16
1.5	DELIMITAÇÃO DO TEMA	16
<b>2</b>	<b>REFERENCIAL TEÓRICO</b>	<b>17</b>
2.1	O COVID-19	17
2.1.1	A PANDEMIA DO COVID-19	17
2.1.1.1	Pandemia no Brasil	19
2.1.2	Sintomas e Tratamento	20
2.2	INTELIGÊNCIA COMPUTACIONAL	21
2.2.1	Aprendizado de conceito e aprendizado supervisionado	23
2.3	REDES BAYESIANAS	25
2.3.1	Aprendizagem	28
2.3.1.1	Aprendizagem da estrutura	28
2.3.1.1.1	<i>Aprendizagem da estrutura usando algoritmo guloso</i>	28
2.3.1.1.2	<i>Aprendizagem da estrutura usando algoritmo Chow-Liu tree</i>	29
2.3.2	Inferência Probabilística	29
2.3.2.1	Distribuição Multinomial e Dirichlet	30
<b>3</b>	<b>METODOLOGIA</b>	<b>32</b>
3.1	CLASSIFICAÇÃO DA PESQUISA	32
3.2	MÉTODO DE EXECUÇÃO	33
3.2.1	Revisão bibliográfica	33
3.2.2	Pesquisa Experimental	34
3.2.3	Estudo de Caso/Análise	36
<b>4</b>	<b>RESULTADOS E DISCUSSÃO</b>	<b>41</b>
4.1	ANÁLISE DOS DADOS: 2020	44
4.2	ANÁLISE DOS DADOS: 2021	45
<b>5</b>	<b>CONCLUSÕES E PERSPECTIVAS</b>	<b>48</b>
	<b>REFERÊNCIAS</b>	<b>50</b>
	<b>APÊNDICES</b>	<b>55</b>
	<b>APÊNDICE A – HISTOGRAMAS</b>	<b>56</b>

	<b>APÊNDICE B – CONDITIONAL PROBABILITY TABLE (CPT)   TA-</b>	
	<b>BELAS DE PROBABILIDADE CONJUNTA . . . .</b>	<b>58</b>
B.1	2020 . . . . .	58
B.2	2021 . . . . .	59

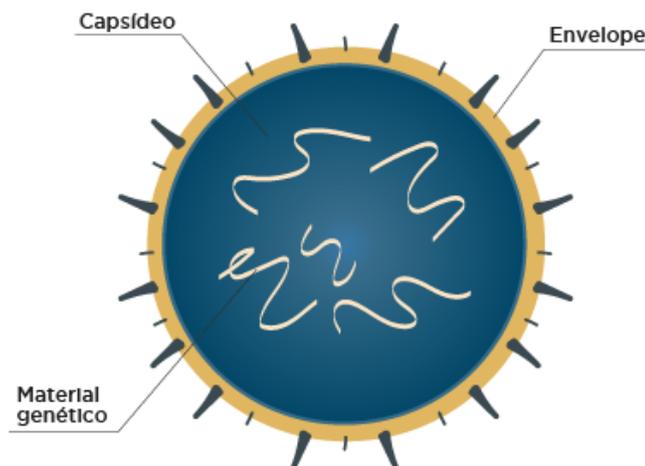
# 1 INTRODUÇÃO

## 1.1 CONTEXTUALIZAÇÃO DO PROBLEMA

A estrutura de um vírus, mostrada na Figura 1, apresenta em sua formação básica uma cápsula de proteínas e a apresentação de ácido nucleico, ou capsídeo, que pode ser uma base RNA, DNA ou de ambos (KONININ *et al.*, 2006). Os vírus, seres microscópicos, possuem segundo KONININ *et al.* (2006) a reputação de serem a causa de qualquer tipo de contágio, isto que a partir de eventos reais onde predominante houve grande número de morte e foram causadas por um, sem dúvida reforçaram essa reputação. O papel principal do vírus é “entregar seu genoma de DNA ou RNA na célula hospedeira para que o genoma possa ser expresso, ou seja, traduzido, pela célula” (KONININ *et al.*, 2006).

Exemplos de doenças humanas comuns causadas por vírus incluem o resfriado comum, gripe, varicela e herpes. Muitas doenças graves, como a raiva, o vírus Ebola, a AIDS (HIV), a gripe aviária e a SARS, são causadas por vírus. A capacidade relativa dos vírus de causar doenças é descrita em termos de virulência (KONININ *et al.*, 2006).

**Figura 1 – Estrutura de um Vírus.**



**Fonte: (PRAVALER, 2020)**

Uma via de transmissão comum é por meio de organismos transmissores de doenças, conhecidos como organismos vetores, por exemplo, de uma planta para insetos que se alimentam da seiva da planta, como pulgões; e os vírus em animais podem

ser transportados por insetos sugadores de sangue, como os mosquitos (KOONIN *et al.*, 2006). Os vírus da gripe se espalham pelo ar por meio da tosse e dos espirros (KOONIN *et al.*, 2006). O HIV, por exemplo, é um dos vários agentes patogênicos transmitidos por contato sexual e pela exposição a sangue infectado. A variedade de células hospedeiras que podem ser infectadas é chamada de "gama de hospedeiros". Isso pode ser restrito, o que significa que este é capaz de infectar poucas espécies, ou amplo, o que significa que é capaz de infectar muitas (GALLO, 1987).

O DNA ou RNA encontrado no núcleo do vírus pode ser de fita simples ou dupla. Constitui o genoma ou a soma total das informações genéticas de um vírus. Os genomas virais são geralmente pequenos em tamanho, codificando apenas para proteínas essenciais, como proteínas do capsídeo, enzimas e proteínas necessárias para a replicação dentro de uma célula hospedeira (KOONIN *et al.*, 2006).

Seres humanos tem pouca ou nenhuma imunidade a novos vírus (SAMPSON, 2020). Normalmente um novo microrganismo viral não se espalha entre espécies, mas pode acontecer de uma mutação ocorrer e uma praga que antes só atingia certos animais comece a infectar pessoas. Pode acontecer também contágio causado pelo surgimento de uma nova cepa ou subtipo de um patógeno que já tinha circulação conhecida.

Autoridades consideram uma doença ser uma epidemia quando o número de pessoas infectadas é maior do que o previsto para uma região específica. Caso esta infecção se torne comum em vários países ao mesmo tempo, ela pode se tornar uma pandemia.

A pandemia de gripe de 1918, que durou até 1919, foi uma pandemia de influenza de categoria 5 causada por um vírus influenza A grave e mortal (MORENS *et al.*, 2009). As vítimas geralmente eram jovens adultos saudáveis, em contraste com a maioria dos surtos de gripe, que afetam predominantemente pacientes jovens, idosos ou debilitados de outra forma (JOHNSON; MUELLER, 2002). Estimativas mais antigas dizem que matou de 40 a 50 milhões de pessoas, enquanto pesquisas mais recentes sugerem que pode ter matado até 100 milhões de pessoas, ou 5% da população mundial em 1918 (JOHNSON; MUELLER, 2002).

A vacinação é a forma mais barata e eficaz de prevenir infecções por vírus. As vacinas foram usadas para prevenir infecções virais muito antes da descoberta dos vírus

reais (MOORE *et al.*, 2006). Seu uso resultou em um declínio dramático na morbidade e mortalidade associada a infecções virais, como poliomielite, sarampo, caxumba e rubéola (MOORE *et al.*, 2006). As infecções por varíola foram erradicadas. As vacinas vivas contêm formas enfraquecidas do vírus, que não causam a doença, mas, no entanto, conferem imunidade. Esses vírus são chamados de atenuados (MOORE *et al.*, 2006).

A inteligência computacional vem sendo amplamente utilizada para o prognóstico de doenças no mundo (BICHINDARITZ *et al.*, 2010). Sua aplicação é vista especialmente em temas onde há ainda partes obscuras, como doenças coronárias e de fígado (JABBAR *et al.*, 2015).

Os fatores que podem afetar a saúde são diversos e heurísticos, ou seja, são caóticos, randômicos e não possuem um padrão definido. Assim, utilizar a inteligência computacional para analisar esses dados tem sido de extrema importância para a medicina moderna (JABBAR *et al.*, 2015).

## 1.2 PROBLEMA DE PESQUISA

A partir disto, temos como problema de pesquisa: **Não havendo uma relação clara entre as políticas de contenção e combate ao COVID 19 e a taxa de mortalidade causada pela doença, este trabalho tem por fim estimar e analisar essas correlações.**

## 1.3 OBJETIVOS

### 1.3.1 OBJETIVO GERAL

Este trabalho tem por objetivo analisar o impacto dos fatores associados ao combate e contenção do Sars-CoV-2 utilizando redes Bayesianas.

### 1.3.2 OBJETIVOS ESPECÍFICOS

- Extrair dados sobre a política de contenção e combate ao Covid-19 de diferentes países;

- Relacionar estes dados com o impacto causado (positiva ou negativamente);
- Modelar as relações do conjunto de dados através de redes-bayesianas;
- Implementar técnicas e algoritmos de estimação de estrutura das redes a fim de otimizar a relação causa/efeito entre os fatores e as políticas aplicadas.
- Implementar técnicas e algoritmos de estimação de parâmetros (probabilidades) da rede para realização de inferências e análise de resultado

#### 1.4 JUSTIFICATIVA

Pelo momento vivido pela população mundial, se tornou vital conceber quais fatores realmente afetaram a taxa de mortalidade de quaisquer regiões analisadas. Uma melhor análise dos fatores que realmente influenciaram no combate a pandemia do Covid-19 pode nos ensinar e demonstrar, para uma futura pandemia ainda não descoberta, o que é possível se fazer para limitarmos o avanço do vírus no planeta.

A partir da catástrofe econômica e social pela que passa o Brasil e o mundo, é de suma importância utilizarmos a mais diversas ferramentas para termos uma avaliação destes fatores se correlacionarem ou não, a partir de uma base de dados distinta, específica e limitada.

#### 1.5 DELIMITAÇÃO DO TEMA

O tema será delimitado apenas para os seguintes países: Alemanha, Brasil, Chile, Estados Unidos da América e Finlândia, para melhor qualidade dos dados obtidos e de sua veracidade. Há também maior controle sobre a porcentagem de mortes, vacinados e, por exemplo, os dias de *lockdown* efetivo de cada um desses países.

## 2 REFERENCIAL TEÓRICO

Esta seção apresenta uma síntese sobre o Covid-19, nome dado ao surto pandêmico do novo patógeno SARS-CoV-2, o que se entende ser a origem deste e conceitos básicos epidemiológicos. Desta forma, também será abordado sobre o que se entende de inteligência computacional, seus conceitos e histórico; assim como a aplicação de redes Bayesianas, qual seu valor para o presente trabalho, e seu funcionamento teórico.

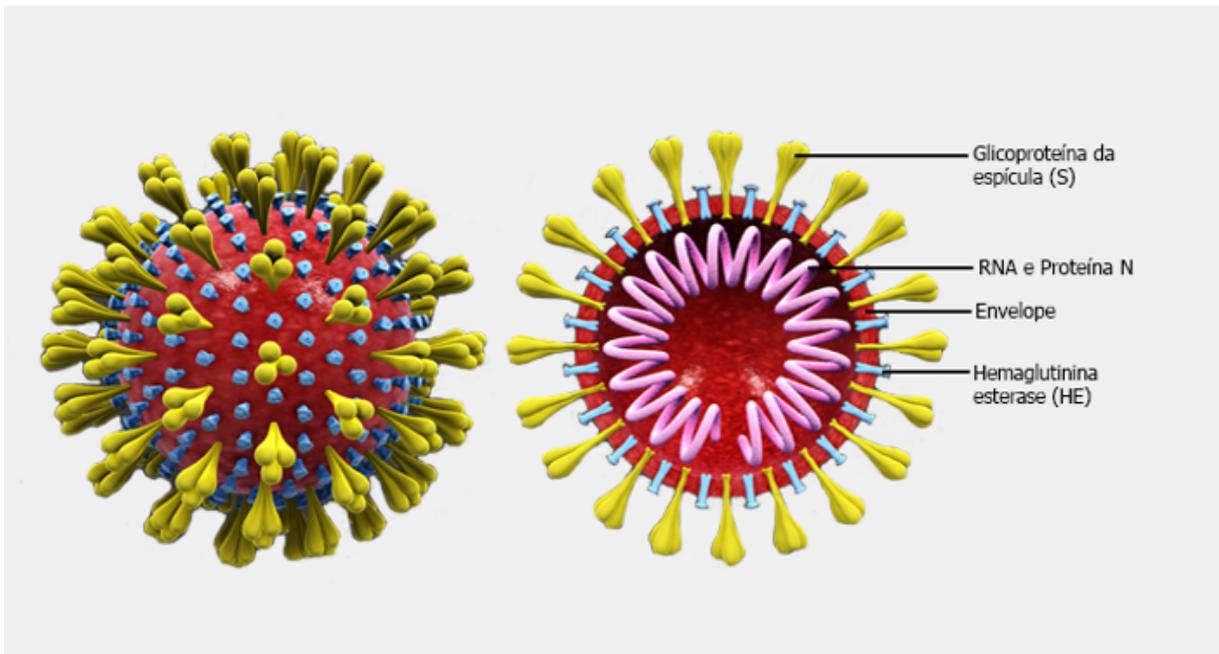
### 2.1 O COVID-19

Os coronavírus (CoVs), é uma grande família de vírus do tipo RNA de fita simples, que podem infectar animais e também humanos, causando doenças respiratórias, gastrointestinais, hepáticas e neurológicas (WU *et al.*, 2020). Como os maiores vírus de RNA conhecidos, os CoVs são divididos em quatro gêneros: alfa-coronavírus, beta-coronavírus, gama-coronavírus e delta-coronavírus (YANG; LEIBOWITZ, 2015). Novos coronavírus parecem surgir periodicamente em humanos, principalmente devido à alta prevalência e ampla distribuição, à grande diversidade genética e recombinação frequente de seus genomas e ao aumento das atividades humano-animais (ZHU *et al.*, 2020).

#### 2.1.1 A PANDEMIA DO COVID-19

A pandemia do COVID-19 é causada pela síndrome respiratória aguda grave coronavírus 2 (SARS-CoV-2) (WU *et al.*, 2020). O vírus, representado na Figura 2, foi identificado pela primeira vez em Wuhan, China, em dezembro de 2019 (ZHU *et al.*, 2020). Em janeiro de 2020, a China implementou um bloqueio – também conhecido como *lockdown* – em Wuhan, mas essa medida não conseguiu conter o vírus, o qual se espalhou para outras partes da China continental e, eventualmente, ao redor do mundo (KUO, 2020). A Organização Mundial da Saúde (OMS) declarou o vírus uma emergência de saúde pública de preocupação internacional em 30 de janeiro de 2020 e, posteriormente, uma pandemia em 11 de março de 2020 (MOREIRA; PINHEIRO, 2021).

**Figura 2 – O Sars-Cov-2**



Fonte: Adaptado de <https://www.scienficanimations.com/wiki-images/>

Desde 2021, variantes do vírus resultaram em novas ondas em vários países, com a variante Delta sendo a mais virulenta. Em 19 de julho de 2021, mais de 190 milhões de casos foram confirmados, com mais de 4,09 milhões de mortes confirmadas atribuídas ao COVID-19, tornando-se uma das pandemias mais mortais da história (KATELLA, 2021). De acordo com o *Worldometers* esses números no início de dezembro de 2021 já estavam em 267.642.979 casos confirmados e 5.291.050 mortes confirmadas. Assim, a linha do tempo desde a identificação do vírus em Wuhan até o momento em que a OMS declarou estado de pandemia se apresenta na Figura 3:

**Figura 3 – Linha do tempo do começo da pandemia do COVID-19**

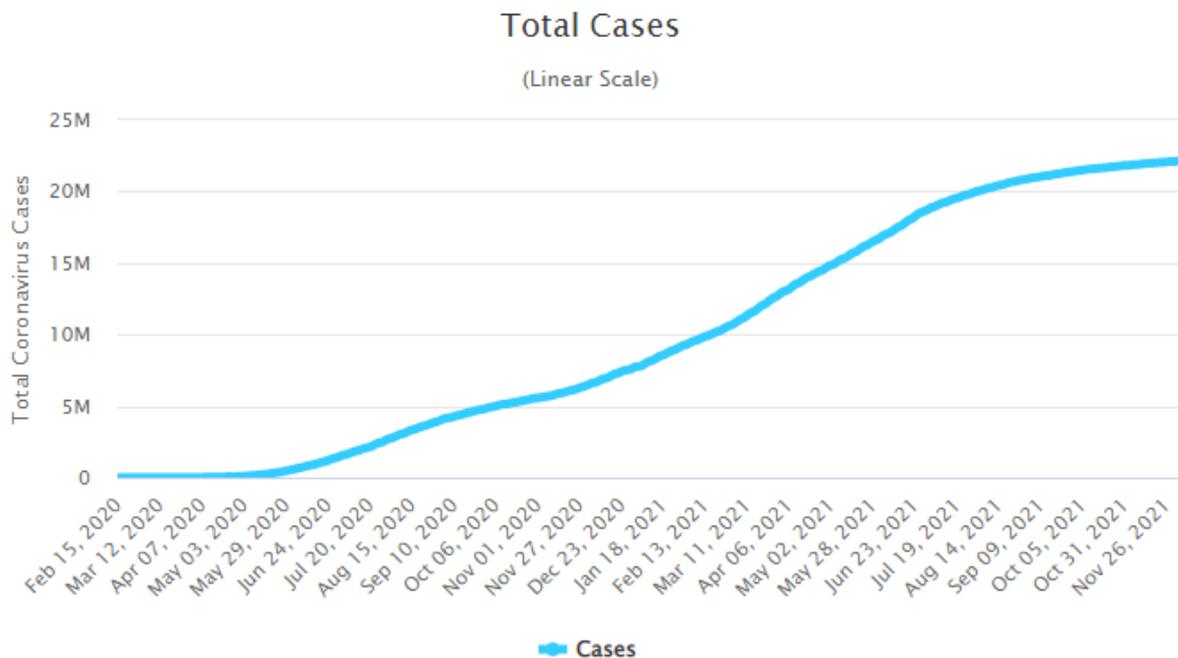


Fonte: Autoria própria, (2021).

### 2.1.1.1 Pandemia no Brasil

A pandemia do COVID-19 teve efeitos calamitosos em muitos países, e o Brasil foi um dos mais afetados. Mais de 22.105.872 de casos totais foram notificados, em uma evolução como apresentado na Figura 4 enquanto 615.020 pessoas faleceram devido ao COVID-19 no país, até dezembro de 2021, ocupando a terceira posição em casos totais e segunda posição em mortes (WORLDMETERS, 2021). O país enfrentou grandes entraves políticos, principalmente na esfera federal, para começar as primeiras medidas preventivas indicadas pela OMS (UOL, 2020). A primeira capital do país a aderir um lockdown parcial foi São Luís, no Maranhão, no dia 05 de maio de 2020 (BARIFOUSE, 2021), mas o primeiro caso de infecção do novo coronavírus foi indicado ainda no dia 26 de fevereiro de 2020, quase 3 meses antes da primeira medida mais rigorosa indicada pela OMS para combate à pandemia em uma capital do território brasileiro (G1, 2020).

**Figura 4 – Total de Casos em milhões do COVID-19 no Brasil**



Fonte: (WORLDMETERS, 2021).

### 2.1.2 Sintomas e Tratamento

A gravidade dos sintomas de COVID-19 é altamente variável, podendo se apresentar de forma imperceptível a fatal. A doença grave é mais provável em pacientes idosos, bem como naqueles que apresentam certas condições médicas subjacentes (comorbidades) (WHO, 2021). A COVID-19 é transmitido quando as pessoas respiram ar contaminado por gotículas e pequenas partículas transportadas pelo ar. O risco de inalá-los é maior quando as pessoas estão próximas umas das outras, mas eles podem ser inalados por distâncias mais longas, principalmente em ambientes fechados. A transmissão também pode ocorrer se respingado ou pulverizado com fluidos contaminados, nos olhos, nariz ou boca e, raramente, através de superfícies contaminadas. As pessoas permanecem contagiosas por até 20 dias e podem espalhar o vírus mesmo se não desenvolverem nenhum sintoma (WHO, 2021).

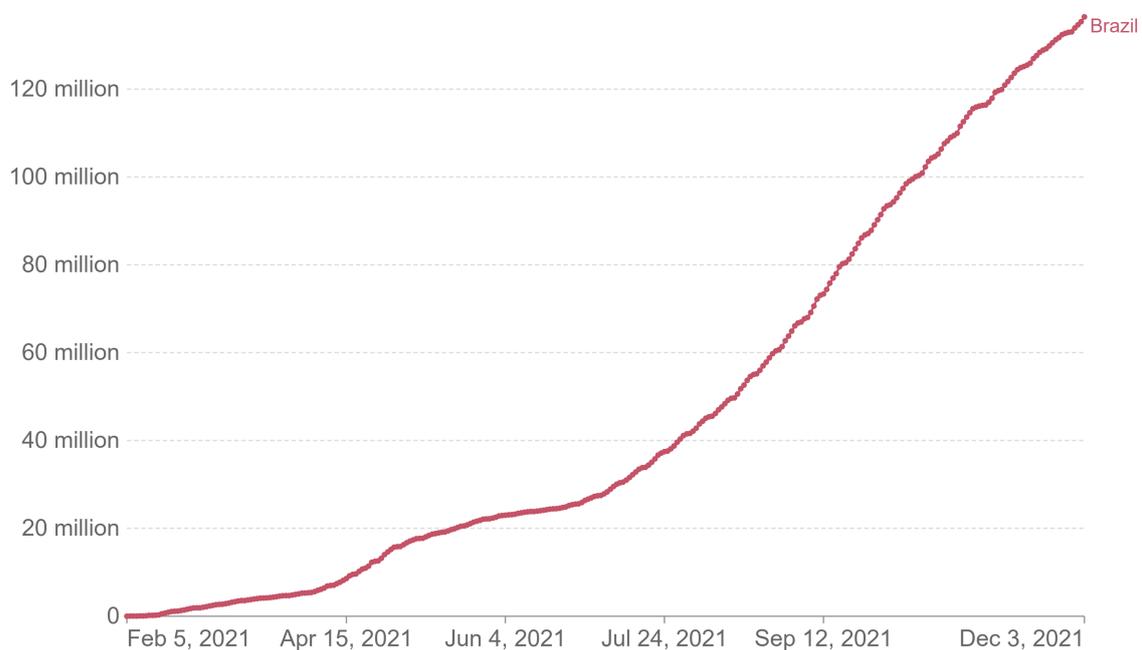
As medidas preventivas recomendadas incluem distanciamento social, uso de máscaras faciais em público, ventilação e filtragem de ar, lavagem das mãos, cobertura da boca ao espirrar ou tossir, desinfecção de superfícies e monitoramento e auto isolamento para pessoas expostas ou sintomáticas (WHO, 2019). Autoridades em todo o mundo responderam implementando restrições a viagens, bloqueios e quarentenas, controles de perigos no local de trabalho e fechamento de empresas. Inúmeras jurisdições também trabalharam para aumentar a capacidade de teste e rastrear os contatos dos infectados (WHO, 2021).

Várias vacinas foram desenvolvidas e amplamente distribuídas na maioria dos países desenvolvidos desde dezembro de 2020 (CREECH *et al.*, 2021). A primeira vacina a ser aprovada para uso emergencial pela OMS foi a vacina da Pfizer/BioNTech, em 03 de dezembro de 2020 (WHO, 2019). Desta forma, o primeiro país a aprovar o uso da vacina Pfizer/BioNTech foi o Reino Unido, dia 02 de dezembro de 2020 (UK, 2020).

Apesar da hesitação na vacinação da população mundial, que é considerada ainda um dos obstáculos para o fim gradual da pandemia (DROR *et al.*, 2020), até dezembro de 2021, 54,3 % da população mundial chegou a ser vacinada com ao menos uma dose, e apenas 42,7 % foi totalmente vacinada, com duas doses e terminando o ciclo imunológico (WORLDMETERS, 2021).

No Brasil, apesar de todos os entraves colocados pelo governo federal, a cultura de vacinação surpreendeu e a aceitação da vacina foi considerada altíssima. Foram administradas 318.561.549 doses, até 03 de dezembro de 2021. Desta forma, aproximadamente 136.438.552 de pessoas receberam primeira e segunda dose, completando o ciclo de imunização, uma porcentagem de 64,2 % da população. A evolução desse número pode ser vista na Figura 5. Em congruência, 17.665.501 de cidadãos já receberam a dose de reforço da vacina do covid-19. (OURWORLDINDATA, 2021).

**Figura 5 – Evolução da vacinação no Brasil até dezembro de 2021 (em milhões)**



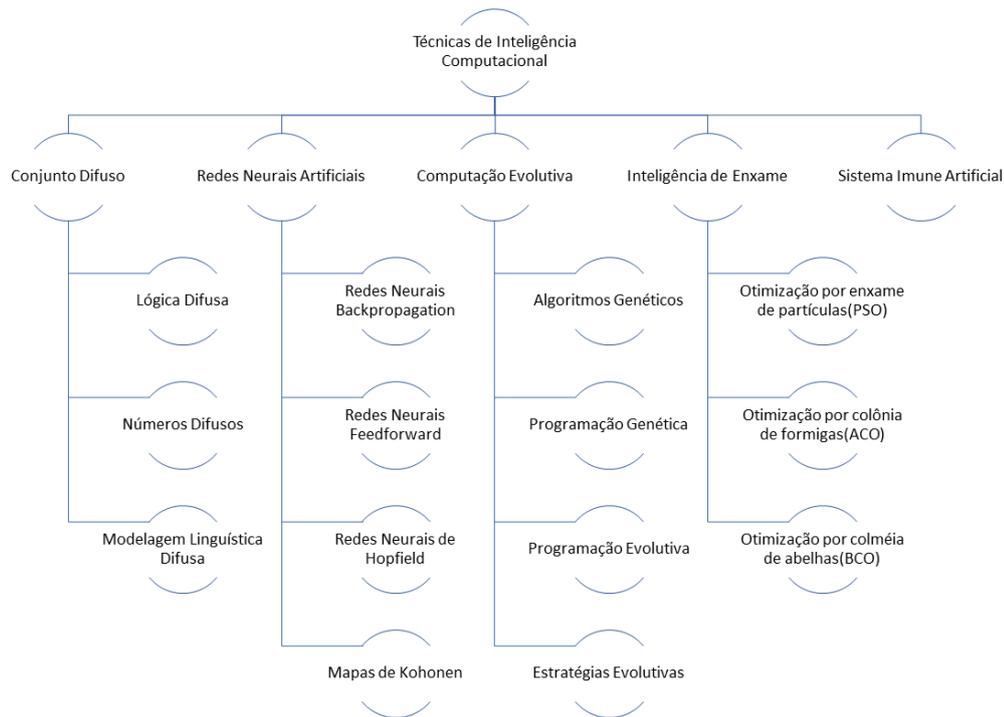
**Fonte: Adaptado de Our World in Data (2021).**

## 2.2 INTELIGÊNCIA COMPUTACIONAL

Segundo Lam *et al.* (2012), Inteligência computacional (IC) é a teoria, design, aplicação e desenvolvimento de paradigmas computacionais motivados biologicamente e linguisticamente. Tradicionalmente, os quatro principais pilares da IC são Redes Neurais, Sistemas *Fuzzy*(Difusos), Computação Evolutiva e Inteligência de Enxame.

No entanto, com o tempo, muitos paradigmas de computação inspirados na natureza, que incluem inspirações em comportamentos de espécies, sistemas de indivíduos entre outros, evoluíram, como pode ser visto na Figura 6 em que é mostrada a relação dos pilares junto às técnicas relativas que foram se juntando a estes.

**Figura 6 – Técnicas da Inteligência Computacional**



**Fonte: Adaptado de SADEGI et Al., (2017).**

Assim, a IC se tornou um campo em evolução e, atualmente, além dos três constituintes principais, abrange paradigmas de computação como inteligência ambiental, vida artificial, aprendizagem cultural, redes endócrinas artificiais, raciocínio social e redes artificiais de hormônios (LAM *et al.*, 2012). A inteligência computacional desempenha um papel importante no desenvolvimento de sistemas inteligentes de sucesso, incluindo simulações de estados, jogos e sistemas de desenvolvimento cognitivo. Nos últimos anos, com a facilidade na exploração de novos dados de pesquisa, houve uma explosão de pesquisas sobre *Deep Learning*, em particular redes neurais convolucionais (LAM *et al.*, 2012).

Mas foi Bezdek (1994), que introduziu a primeira definição clara de Inteligência Computacional, onde um sistema é chamado de inteligência computacional se este lida com dados de baixo nível, como dados numéricos, tem um componente de reconhecimento de padrão e não usa conhecimento no sentido de inteligência artificial (IA) e, adicionalmente, quando começa a exibir de forma adaptativa computacional, tolerância a falhas, velocidade de recuperação semelhante à humana e taxas de erro que se aproximam do desempenho humano.

Um entendimento importante a se destacar é a diferença entre inteligência

artificial e inteligência computacional. Segundo, Bezdek (1994), embora a Inteligência Artificial e a Inteligência Computacional busquem um objetivo semelhante de longo prazo, que é alcançar a inteligência geral, ou seja, a inteligência de uma máquina capaz de realizar qualquer tarefa intelectual que um ser humano pode; há uma diferença clara entre elas. Assim, a Inteligência Computacional é um subconjunto da Inteligência Artificial (BEZDEK, 1994). Existem dois tipos de inteligência de máquina: a artificial baseada em técnicas de computação pesada e a computacional baseada em métodos de computação flexível, que permitem a adaptação a muitas situações (ZADEH, 1994).

Computação pesada, segundo Zadeh, requer um modelo analítico com declaração precisa, muitas vezes requerendo muito tempo de computação. As técnicas de computação pesada funcionam segundo a lógica binária com base em apenas dois valores (os booleanos verdadeiro ou falso, 0 ou 1) nos quais os computadores modernos se baseiam. Um problema com essa lógica é que nossa linguagem natural nem sempre pode ser traduzida facilmente em termos absolutos de 0 e 1.

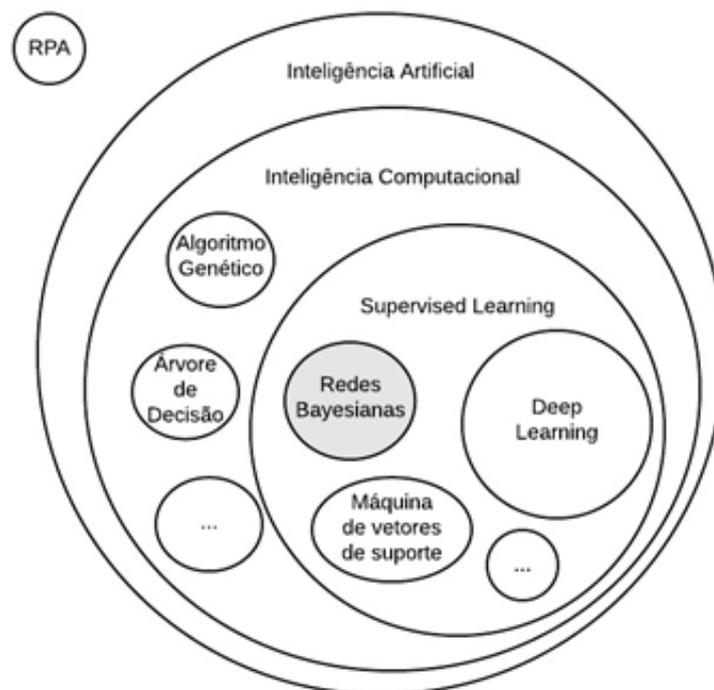
Já a computação flexível, que foi um termo proposto por Lotfi Aliasker Zadeh, como diz o próprio autor, é tolerante à imprecisão, incerteza, verdade parcial e aproximação. Na verdade, o modelo para computação flexível é a mente humana. Aqui incluem-se os conceitos fundamentais de redes neurais, lógica difusa e algoritmo genético que, por sua vez, oferecem a superioridade das capacidades humanas de resolução de problemas (Zadeh, 1994).

### 2.2.1 Aprendizado de conceito e aprendizado supervisionado

*Concept learning* (CL), ou numa tradução literal, aprendizado de conceito, é um conceito da psicologia definido como a busca e listagem de atributos que podem ser usados para distinguir objetos exemplares de não-exemplares em inúmeras categorias (BRUNE; AUSTIN, 1986). Simplificando, os conceitos são as categorias mentais que nos ajudam a classificar objetos, eventos ou ideias, construindo na compreensão de que cada objeto, evento ou ideia tem um conjunto de características relevantes comuns. Assim, a aprendizagem de conceitos é uma estratégia que necessita que o objeto compare e contraste grupos ou categorias que contêm características relevantes para o conceito com grupos ou categorias que não contêm características relevantes para o conceito.

Assim como os algoritmos genéticos se inspiram na natureza (LINDEN, 2006), o *supervised learning* (SL), ou em tradução literal, aprendizado supervisionado, se inspira na aprendizagem de conceito, que de maneira resumida aborda que um conceito já assimilado pode ser transferido, criando um paralelo matemático com o *machine learning*, subdividindo-se (MOHRI *et al.*, 2018). *Machine learning*, por sua vez, pode ser conceituado como um treinamento, utilizando-se de inteligência artificial, que ao fornecer informações a um conjunto de algoritmos, o permite aprender e melhorar com a experiência, sem ser explicitamente programado. Ou seja, o algoritmo utiliza-se de IA para aprender por si (MOHRI *et al.*, 2018). Uma apresentação conceitual de como estas técnicas se englobam pode ser visto na Figura 7.

**Figura 7 – O conceito englobado de Inteligência Artificial e Inteligência Computacional**

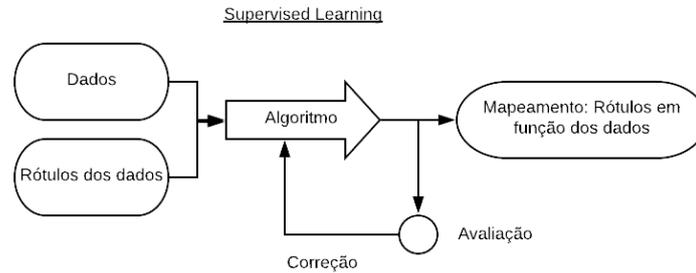


**Fonte: Autoria própria, (2021).**

Desta forma, o aprendizado supervisionado (SL) é a tarefa de aprendizado de máquina para aprender uma função onde se mapeia uma entrada para uma saída com base em pares de entrada-saída (*input-output*) de exemplos (MOHRI *et al.*, 2018) como visto na Figura 8. A analogia mais simples de se fazer é comparar essa forma de aprendizado com a maneira com que se ensina uma criança pequena. Com a criança é necessário repetir e mostrar o padrão até que ela assimile o conceito, em SL essa

assimilação é feita através de modelo matemáticos.

**Figura 8 – O conceito básico de um *Supervised Learning***



**Fonte: Autoria própria, (2021).**

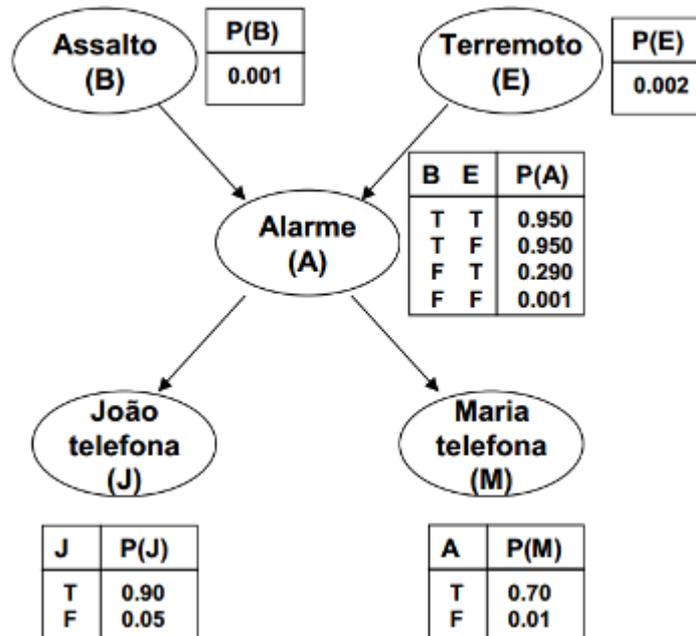
Ele infere uma função a partir de dados de treinamento rotulados que consistem em um conjunto de exemplos de treinamento (MOHRI *et al.*, 2018). Na aprendizagem supervisionada, cada exemplo é um par que consiste em um objeto de entrada (*input*) (normalmente um vetor) e um valor de saída desejado (*output*) (também chamado de sinal de supervisão). Um algoritmo de aprendizado supervisionado analisa os dados de treinamento e produz uma função inferida, que pode ser usada para mapear novos exemplos (CARUANA; NICULESCU-MIZIL, 2006). Um cenário ideal permitirá que o algoritmo determine corretamente os rótulos de classe para instâncias não vistas. Isso requer que o algoritmo de aprendizado faça uma generalização a partir dos dados de treinamento para situações invisíveis de uma forma "razoável". Essa qualidade estatística de um algoritmo é medida por meio do chamado erro de generalização (MOHRI *et al.*, 2018).

### 2.3 REDES BAYESIANAS

Redes Bayesianas tem esse nome devido ao Teorema de Bayes, conceito probabilístico que descreve a probabilidade de um evento ocorrer, com base no conhecimento prévio das condições que podem estar relacionadas a este evento (JOYCE, 2003), como visto na Equação 1, criando assim um fluxo de realização de eventos, em que o mais importante é o conhecimento sobre as possíveis relações condicionais prévias ao evento-hipótese, como visto no exemplo da Figura 9.

$$P(A | B) = \frac{P(B | A) * P(A)}{P(B^*)} \quad (1)$$

Figura 9 – Exemplo de Fluxo de uma rede bayesiana



Fonte: (JASON, 2012)

onde  $P(A | B)$  é a probabilidade *a posteriori*, ou probabilidade condicionada, que se refere a probabilidade de  $A$  ocorrer se  $B$  tiver ocorrido.  $P(B | A)$  é encontrada através de um estimador de máxima verossimilhança, fixando o valor de  $B$  e encontrando a probabilidade condicionada de  $A$  ocorrer, caso  $B$  esteja no valor fixado.  $P(A)$  é a probabilidade de  $A$  ocorrer e  $P(B)$  é a probabilidade de  $B$  ocorrer.

Pearl (1988) afirmou que as redes bayesianas são modelos gráficos que contêm informações sobre relações de probabilidade causal entre variáveis e são frequentemente utilizadas para auxiliar na tomada de decisões. As relações de probabilidade causal em uma rede Bayesiana podem ser sugeridas por especialistas ou atualizadas usando o teorema de Bayes e novos dados sendo coletados. A estrutura de dependência inter variável é representada por nós (que representam as variáveis) e arcos direcionados (que representam as relações condicionais) na forma de um gráfico acíclico direcionado (GAD) (PEARL, 1988).

Em (PEARL, 1988) o autor apresenta um algoritmo mostrado no Algoritmo 1, baseando-se na estrutura e evidências, no qual dado um conjunto de variáveis discretas ordenadas, uma estrutura de rede bayesiana única é criada, adicionando às variáveis

a rede em sua ordem e acrescentando arcos para a formação da estrutura. Desta forma, as variáveis são conectadas impreterivelmente às variáveis anteriores da rede, garantindo que seja sempre acíclico.

#### Algoritmo 1 – Algoritmo para Construção de Redes Bayesianas

- 1: Escolha um conjunto de variáveis  $X_i$  que em suposição descreva o problema
- 2: Escolha uma ordem para as variáveis
- 3: **para** todas as variáveis em ordem **faça**
- 4: Escolha a variável  $X$  e adicione-a na rede
- 5: Determine os pais da variável  $X$  dentre os nós que já estão na rede, que satisfaça  $P(X_i|X_j, pais(X_i)) = P(X_i|pais(X_i))$
- 6: Construa a tabela de probabilidade condicional (TPC) para  $X$
- 7: **finaliza para**

Portanto, as redes bayesianas são majoritariamente utilizadas a partir de um evento-hipótese que já ocorreu e prever a probabilidade de que qualquer uma das inúmeras causas conhecidas possíveis tenha sido o fator contribuinte (BEN-GAL, 2008).

Seja  $Y = (Y_1, \dots, Y_M)$  um vetor de variáveis aleatórias, e seja  $y_m$  um valor de  $Y_m$ , o  $m$ -ésimo componente de  $Y$ . A representação de um modelo bayesiano é dada por dois componentes (LARRAÑAGA *et al.*, 2012): uma estrutura e um conjunto de parâmetros locais.

O conjunto de parâmetros locais  $\Theta$  contém, para cada variável, a distribuição de probabilidade condicional de seus valores dadas as diferentes configurações de valor para seus pais, de acordo com a estrutura  $B$ . A estrutura  $B$  para  $Y$  é um grafo acíclico direcionado que descreve um conjunto de dependências condicionais de todas as variáveis em  $Y$ .  $Pa_m^B$  representa o conjunto de pais (variáveis de onde uma seta de influência sai de  $B$ ) da variável  $Y_m$  no *Probabilistic Graphic Model* (PGM) ou na tradução Modelo Probabilístico Gráfico, cuja estrutura é dada por  $B$  (BENGOETXEA, 2002). Esta estrutura assume que  $Y_m$  é independente de seus não descendentes dado  $Pa_m^B$ ,  $m = 2, \dots, M$ , onde  $Y_1$  é o nó raiz. Portanto, uma rede bayesiana codifica uma fatoração para a probabilidade de massa como segue na Equação 2:

$$p(y) = p(y_1, y_2, \dots, y_n) = \prod_{m=1}^M p(y_m | pa_m^B) \quad (2)$$

### 2.3.1 Aprendizagem

De acordo com (ARA-SOUZA, 2010) o aprendizado estatístico está principalmente incluído ao processo de aprendizagem quando existem incerteza e variabilidade. Portanto, através de um conjunto de dados, é possível a utilização de um processo de estimação e validação do sistema em estudo, aplicando-se técnicas estatísticas necessárias ao domínio do problema. Diz (ARA-SOUZA, 2010), que devido à dificuldade da construção de uma Rede Bayesiana unicamente consultando um especialista, existe o interesse de se estimar todos os elementos da rede, estes sendo sua estrutura, e as probabilidades condicionais de cada tabela de probabilidade condicional (CPT), também chamadas de parâmetros ou elementos numéricos.

No caso de redes bayesianas, tanto a estrutura quanto os parâmetros numéricos podem ser aprendidos, caso se tenha uma base amostral suficiente significativa. Num primeiro passo se encontra uma estrutura  $S$ , que represente o problema e a seguir se aprendem os parâmetros numéricos  $\Theta$ .

#### 2.3.1.1 Aprendizagem da estrutura

Em seu texto, Ara-Souza (2010) diz que:

"A construção de uma Rede Bayesiana não é trivial, além de existirem vários métodos para a estimação de estruturas de rede através do conjunto de dados, os métodos podem ser influenciados por fatores como a ordem e escolha das variáveis que compõem o problema."

Neste trabalho escolheu-se trabalhar com dois algoritmos, *greedy search* (algoritmo guloso) e *Chow-Liu tree* (árvore de Chow-Liu), descritos nas seções 2.3.1.1.1 e 2.3.1.1.2. Os motivos destas escolhas estão explícitos na seções 4.1 e 4.2.

##### 2.3.1.1.1 Aprendizagem da estrutura usando algoritmo guloso

De maneira geral um *greedy search* ou algoritmo guloso é uma heurística para resolução de problema que tem como resposta o ótimo local para aquele problema. Neste experimento a função a ser otimizada se encontra na Equação 3.

$$P(D | M) + \frac{\log_2(|D|)}{2} \cdot |M| \quad (3)$$

Onde  $P(D|M)$  é a probabilidade dos dados considerado o modelo encontrado,  $|M|$  é o número de parâmetros no modelo,  $\log_2(|D|)/2$  é a penalidade e  $|D|$  é a soma dos pesos do exemplo.

### 2.3.1.1.2 Aprendizagem da estrutura usando algoritmo Chow-Liu tree

O método *Chow-Liu tree*, (CHOW; LIU, 1968) descreve uma distribuição de probabilidade conjunta  $P(X_1, X_2, \dots, X_n)$  como um produto de distribuições condicionais e marginais de segunda ordem.

O algoritmo mostra como selecionar termos de segunda ordem por aproximação do produto de modo que, entre todas essas aproximações de segunda ordem (árvores de dependência de primeira ordem), a aproximação construída  $P'$  tenha a divergência de Kullback – Leibler (KULLBACK; LEIBLER, 1951) mínima para a distribuição real  $P$ , e é, portanto, a aproximação mais próxima no sentido clássico da teoria da informação. A divergência de Kullback-Leibler entre uma aproximação de produto de segunda ordem e a distribuição real é mostrado mostrada na Equação 4.

$$D(P \parallel P') = - \sum I(X_i; X_{j(i)}) + \sum H(X_i) - H(X_1, X_2, \dots, X_n) \quad (4)$$

onde  $I(X_i; X_{j(i)})$  é a informação mútua entre a variável  $X_i$  e a mãe  $X_{j(i)}$  e  $H(X_1, X_2, \dots, X_n)$  entropia conjunta do conjunto de variáveis  $X_1, X_2, \dots, X_n$ .

Uma vez que os termos  $\sum H(X_i)$  e  $H(X_1, X_2, \dots, X_n)$  são independentes de qualquer ordem de relação na árvore, apenas a soma das informações pareais mútuas,  $\sum I(X_i; X_{j(i)})$  determina a qualidade da aproximação.

### 2.3.2 Inferência Probabilística

De acordo com (RUSSEL; NORVIG, 2004), a Inferência Probabilística é a computação a partir de evidências observadas de probabilidades posteriores para proposições de consulta.

- Inferência a partir da Distribuição de Probabilidade Conjunta Total: para qualquer proporsição  $a$ ,  $P(a)$  é a soma dos eventos atômicos  $w$  onde  $a$  ocorre:  $P(a) = \sum_{w:w|a} P(w)$ ;
- Inferência a partir do Teorema de Bayes: é a base de todos os sistemas de Inteligência Artificial. para inferência probabilística. Utiliza-se da Equação 1;
- Inferência a partir de Redes Bayesianas: é comumente utilizada para referenciar a atualização de probabilidades por toda a estrutura da rede dado um conjunto de evidências. Ou seja, segundo (ARA-SOUZA, 2010), trata-se de um mecanismo para cálculo da distribuição *posteriori* de probabilidade para um conjunto de variáveis, dado um conjunto de evidências, ou seja, variáveis aleatórias com valores instanciados.

### 2.3.2.1 Distribuição Multinomial e Dirichlet

Estas duas distribuições, aqui introduzidas, são amplamente utilizadas no contexto de Redes Bayesianas quando métodos de estimação bayesiana e por máxima verossimilhança são requeridos.

Considera-se uma variável aleatória  $X$  discreta que represente um experimento com  $K$  possíveis resultados, sendo que cada tipo de resultado possui uma probabilidade específica  $P(X = x_k) = \theta_k$  e  $\sum_{k=1}^K \theta_k = 1$ . Além disso, o experimento é repetido de forma independente  $N$  vezes, de forma que a variável  $N_k$  seja o número de vezes que o resultado  $x_k$  está presente na amostra com  $k = 1, \dots, K$ . Tem-se que a variável  $X$  segue distribuição Multinomial, sendo sua função densidade de probabilidade expressa pela equação 5 a seguir.

$$P(X|N, \theta_k) = \frac{N!}{x_1! x_2! \dots x_K!} \theta_1^{x_1} \theta_2^{x_2} \dots \theta_K^{x_K} \quad (5)$$

sendo  $\sum_{k=1}^K x_k = N$

Considerando o termo  $\frac{N!}{x_1! x_2! \dots x_K!}$  como normalizador tem-se a Equação a seguir.

$$P(X|N, \theta_k) \propto \theta_1^{x_1} \theta_2^{x_2} \dots \theta_K^{x_K} \quad (6)$$

Além disso, temos que para um vetor  $\theta_k = (\theta_1, \theta_2, \dots, \theta_K)$  de valores desconhecidos com  $\sum_{k=1}^K \theta_k = 1$ , pode-se assumir que  $\theta_k$  segue distribuição Dirichlet com

parâmetros  $\alpha = (\alpha_1, \dots, \alpha_K)$  com  $\alpha_k > 1$ , média  $E(\theta_k) = \alpha_k / \alpha_0$ , e função densidade de probabilidade expressa pela Equação 7, sendo  $\alpha_0 = \sum_{k=1}^K \alpha_k$ .

$$P(\theta_k|\alpha) = \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1)\dots\Gamma(\alpha_K)} \theta_1^{\alpha_1-1} \theta_2^{\alpha_2-1} \dots \theta_K^{\alpha_K-1} \quad (7)$$

Da mesma forma pode-se considerar o termo  $\frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1)\dots\Gamma(\alpha_K)}$  como normalizador. Assim tem-se a Equação 8.

$$P(\theta_k|\alpha) \propto \theta_1^{\alpha_1-1} \theta_2^{\alpha_2-1} \dots \theta_K^{\alpha_K-1} \quad (8)$$

Assumindo como *priori*  $P(\theta_k|\alpha)$  e como verossimilhança  $P(X|N, \theta_k)$ , tem-se que a *posteriori*  $P(\theta_k|X)$  é dada pela equação 9 a qual tem distribuição Dirichlet com parâmetros  $\alpha = (\alpha_1 + x_1, \dots, \alpha_K + x_K)$ , ou seja  $E(\theta_k) = \frac{\alpha_k + x_k}{\alpha_0 + N}$ , sendo  $\alpha_0 = \sum_{k=1}^K \alpha_k$ .

$$P(p|X) \propto \theta_1^{\alpha_1+x_1-1} \theta_2^{\alpha_2+x_2-1} \dots \theta_K^{\alpha_K+x_K-1} \quad (9)$$

Nota-se que neste caso a *posteriori* possui o mesmo tipo de distribuição que a *priori*, assim pode-se afirmar que a família Dirichlet é conjugada para amostras com distribuição Multinomial.

Como exemplo, pode-se considerar que os seguintes eventos podem ocorrer com um pacote enviado pelo correio: chegar em perfeito estado, chegar danificado ou perder-se pelo caminho. As probabilidades desses eventos são, respectivamente 0,7, 0,2 e 0,1. Foram enviados recentemente 10 pacotes pelo correio. Qual a probabilidade de 6 chegarem corretamente ao destino, 2 serem perdidos e os outros 2 avariados?

Defina as seguintes variáveis aleatórias:

$X_1$ : número de pacotes que chegaram corretamente e sem danos;  $X_2$ : número de pacotes que chegaram avariados;  $X_3$ : número de pacotes que se perderam pelo caminho.

Então  $X_1 + X_2 + X_3 = N = 10$  e  $\theta_1 + \theta_2 + \theta_3 = 0,7 + 0,2 + 0,1 = 1$ . Logo,

$$P(X_1 = 6, X_2 = 2, X_3 = 2) = \frac{10!}{6!2!2!} (0,7)^6 (0,2)^2 (0,1)^2 = 0,059$$

Neste trabalho encontramos os parâmetros por estimação Bayesiana, utilizando a distribuição de Dirichlet.

### 3 METODOLOGIA

Esta seção possui o objetivo de detalhar como o trabalho será executado, qual é a classificação da pesquisa, como será feita a modelagem do problema e como será elaborada a análise dos resultados.

#### 3.1 CLASSIFICAÇÃO DA PESQUISA

O experimento é considerado o melhor jeito de fazer pesquisa científica, afinal, o autor consegue avaliar com precisão o objeto de estudo e considerar as variáveis que influenciam nos resultados. Assim, esta pesquisa pode ser definida como uma pesquisa quantitativa.

Explica Fonseca (2002):

"Diferentemente da pesquisa qualitativa, os resultados da pesquisa quantitativa podem ser quantificados. Como as amostras geralmente são grandes e consideradas representativas da população, os resultados são tomados como se constituíssem um retrato real de toda a população alvo da pesquisa. A pesquisa quantitativa se centra na objetividade. Influenciada pelo positivismo, considera que a realidade só pode ser compreendida com base na análise de dados brutos, recolhidos com o auxílio de instrumentos padronizados e neutros. A pesquisa quantitativa recorre à linguagem matemática para descrever as causas de um fenômeno, as relações entre variáveis, etc. A utilização conjunta da pesquisa qualitativa e quantitativa permite recolher mais informações do que se poderia conseguir isoladamente".

A classificação do objetivo do trabalho pode ser feita por até três tipos: descritiva, explicativa ou exploratória. Desta forma, o presente trabalho tem seu objetivo definido como exploratório pois tem como objetivo demonstrar maior familiaridade de um certo problema para poder definir hipóteses de resolução (GIL *et al.*, 2002). E, por este motivo, será aplicado como um estudo de caso relativo a uma população definida.

Assim, sua natureza é considerada de pesquisa aplicada, pois possui o objetivo de gerar conhecimento para problemas específicos (SILVEIRA; CÓRDOVA, 2009).

De maneira geral, podemos sumarizar o presente trabalho via a Tabela 1:

**Tabela 1 – Delineamento da Pesquisa**

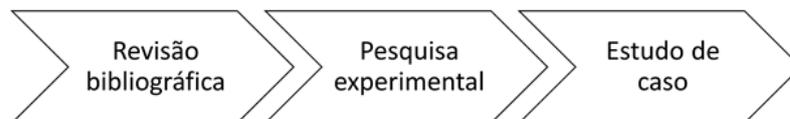
Atributo	Classificação
Abordagem	Quantitativa
Natureza	Aplicada
Objetivos	Exploratória
Procedimentos	Pesquisa bibliográfica e Estudo de caso/Análise de dados

**Fonte: Autoria própria, (2021).**

### 3.2 MÉTODO DE EXECUÇÃO

Com o objetivo de alcançar os objetivos estabelecidos no trabalho, um resumo de como a pesquisa ocorrerá pode ser representado pela Figura 10:

**Figura 10 – Delineamento da Pesquisa**



**Fonte: Autoria própria, (2021).**

Nas seções subsequentes será realizada uma explanação maior de cada uma destas etapas.

#### 3.2.1 Revisão bibliográfica

A pesquisa bibliográfica, como o próprio nome já diz, é realizada a partir de fontes bibliográficas, como livros e artigos científicos. O autor leva em consideração obras de outros autores para conduzir o estudo exploratório. E, para extrair informações, aplica a técnica de análise de conteúdo

A revisão bibliográfica foi realizada em duas fases, cada uma com um objetivo distinto. Primeiramente foi realizada uma revisão bibliográfica de caráter exploratório a respeito da propagação e concepção de um vírus, assim como de aplicações gerais de uma rede bayesiana, do que é inteligência computacional e do uso da mesma na área da saúde. Esta pesquisa foi realizada em livros, artigos científicos, teses e sites.

A seguir, foi realizada outra busca nas bases de dados Google Acadêmico e *ScienceDirect* com as seguintes strings de pesquisa: “Covid-19”, “Teorema de Bayes”, “Rede Bayesiana\*”, “Inteligência Computacional” e “*Soft Computing*”. Como filtro, foi selecionado que os resultados aparecessem por ordem de relevância e o número de citações em outros trabalhos.

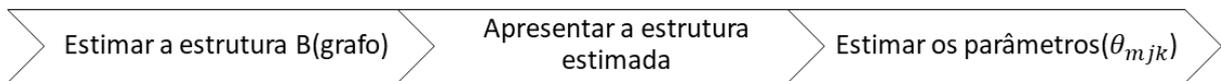
Desta pesquisa em base de dados, foram selecionados 27 resultados que auxiliaram na produção do trabalho. Esta seleção se deu primeiramente pela leitura do título dos resultados apresentados nas bases de dados, e posteriormente pela leitura do resumo dos artigos, para averiguar se o conteúdo dos materiais tinha relevância para o presente trabalho.

### 3.2.2 Pesquisa Experimental

A pesquisa que fornecemos sobre as respostas das políticas é proveniente do Oxford Coronavirus Government Response Tracker (OxCGRT), ou, em tradução literal, o Rastreador Governamental de Resposta ao Coronavírus de Oxford. Os dados atualizados são apresentados graficamente no site Our World in Data<sup>1</sup>. OxCGRT é um projeto de agrupamento contínuo de dados ativos. O rastreador apresenta dados coletados de fontes públicas por uma equipe de mais de cem alunos da Universidade de Oxford e funcionários de todas as partes do mundo. A OxCGRT coleta informações publicamente disponíveis sobre 17 indicadores de respostas do governo, abrangendo políticas de contenção e fechamento (como fechamento de escolas e restrições de movimento); políticas econômicas; e políticas do sistema de saúde (como regimes de teste). Mais detalhes sobre como essas métricas são medidas e coletadas estão disponíveis no documento de trabalho do projeto. (HALE *et al.*, 2021)

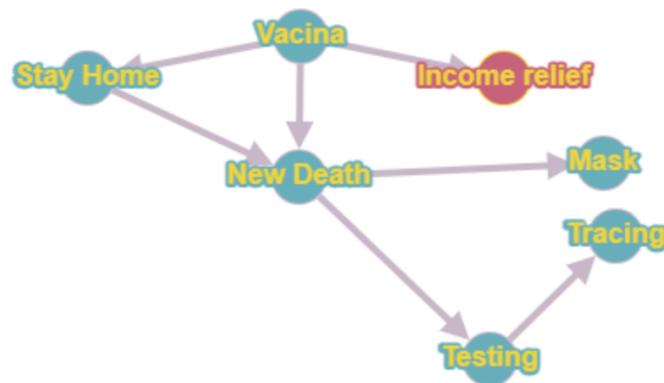
Após a realização desta coleta para base de dados, será aplicado o método de redes bayesianas para identificar possíveis correlações entre eventos que ocorreram chegando numa hipótese já pré-determinada. Este aprendizado de rede utiliza-se de uma hipótese já de conhecimento prévio do problema a qual combinam-se os fatores selecionados afim de determinar uma possível e provável hipótese final. Para isto, podemos referenciar as etapas do presente trabalho, seguindo as diretrizes apresentadas na seção 2.3.1 através de um fluxograma como descrito na Figura 11.

<sup>1</sup> <https://ourworldindata.org/policy-responses-covid>

**Figura 11 – Etapas**

Fonte: Autoria própria, (2021).

A rede Bayesiana será apresentada via um grafo acíclico direcionado com arestas (relações) entre cada um dos eventos nos levando até o evento-fim que procuramos determinar, ou seja, o ótimo de número de mortes (por minimização). Desta forma, no final dessa estrutura com todas as arestas, poderemos tomar uma decisão. A representação de um grafo acíclico direcionado é visto na Figura 12 para melhor descrição:

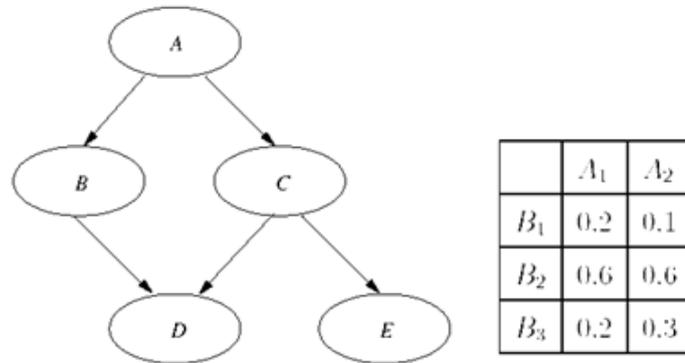
**Figura 12 – Grafo Acíclico direcionado**

Fonte: Autoria própria, (2021).

A Rede Bayesiana apresentará os valores de cada ponto relacionado em taxas de probabilidade( $\theta$ ) em relação aos eventos anteriores( $j$ ) ou posteriores (grafo acíclico direcionado), onde cada variável poderá assumir diferentes valores de estado( $k$ ), dependendo da sua natureza. Por exemplo, caso seja binária, a variável pode assumir dois valores, como (1 ou 2), ou (0 ou 1). Os arcos representam a dependência condicional entre estas e os nós são as próprias variáveis.

Portanto, a força da dependência entre os eventos será representada pelas probabilidades representadas por uma porcentagem de 0.0 a 1.0, como por exemplo na Figura 16, onde são apresentados as probabilidades condicionais relacionadas aos nós A (variável binária com possíveis valores de estado 1 ou 2) e B (variável com 3 possíveis estados 1, 2 ou 3).

**Figura 13 – Probabilidade relacionada aos fatores e eventos exemplificado**



Fonte: (VALTORTA *et al.*, 2008).

Nesse exemplo temos 2 variáveis ( $A$  e  $B$ ), logo  $m$  pode assumir 2 valores (vamos definir o valor 1 para  $A$  e 2 para  $B$ ), como o grafo estabelece que  $B$  é filho de  $A$ , vamos analisar as probabilidades de  $B$  ocorrer ( $\theta_2$ ). Para inferir qual a probabilidade de  $B$  assumir o estado  $k$  sabendo que  $A$  está no estado  $j$ , basta observar-se a célula  $(k,j)$ .

Neste exemplo para saber qual a probabilidade de  $B$  assumir o estado 3, se  $A$  estiver no estado 2, basta conferir a célula da terceira linha, segunda coluna, no caso 0,3 ou 30%.

Assim, pode-se com certeza probabilística entender quais fatores possuem relação com o nosso evento final discutido (Número de mortes) e quais fatores não possuem correlação probabilística. Desta forma, haverá um plano de ação que pode ser tomado em doenças do mesmo tipo.

### 3.2.3 Estudo de Caso/Análise

O objetivo do presente trabalho é realizar a análise de dados das políticas adotadas para contenção e combate ao Covid-19. Assim, realizada a coleta de dados, a fim de relacionar essas medidas com a minimização das mortes, a Tabela 2 e a Tabela 3 evidenciam em exemplo quais dados foram computados, a sua descrição, e por conseguinte, o nome das variáveis utilizadas na programação. As variáveis, portanto, foram as seguintes: suporte financeiro (*Income support*), rastreamento de contato (*Contact tracing*), política de testagem (*Testing policy*), política de confinamento, ou do chamado *lockdown* (*Stay at home requirements*), uso de proteção facial (*Face*

coverings) e política de vacinação (*Vaccination policy*).

**Tabela 2 – Exemplificação dos dados**

Políticas	Descrição
Income Support	Suporte <b>financeiro e/ou de alívio de contas</b> utilizado por governos ou entidades nos países afetados.
Contact Tracing	Processo de identificação de <b>indivíduos que estiveram em contato com pessoas infectadas</b> com o vírus em um estipulado período de tempo.
Testing Policy	Política de <b>disponibilidade de testagem</b> do COVID-19 à população.
Stay-at-home Requirements	Políticas relativas ao chamado <b>lockdown</b> . (Confinamento não eletivo motivado por questões de segurança pública.
Face Coverings	Utilização de <b>máscaras e proteção facial</b> com o objetivo de mitigar a propagação do vírus por vias aéreas.
Vaccination Policy	Política de <b>distribuição e disponibilidade de vacinas</b> para o combate do COVID-19.

Fonte: Autoria própria, (2021).

**Tabela 3 – Descrição dos estados das variáveis**

Variáveis	Estados
Income Support	1 - <b>Sem</b> suporte de renda
	2 - Suporte cobre <b>menos de 50% do salário</b> perdido
	3 - Suporte cobre <b>mais de 50% do salário</b> perdido
Contact Tracing	1 - <b>Sem</b> rastreamento
	2 - Rastreamento Limitado ( <b>alguns</b> mas não todos os casos são rastreados)
	3 - <b>Todos</b> os casos são rastreados
Testing policy	1 - <b>Sem</b> política de testagem
	2 - Testagem somente de <b>sintomáticos de grupos específicos</b> (trabalhadores essenciais, retornando de viagem, contato com caso confirmado, internados)
	3 - Qualquer <b>sintomático</b>
	4 - Testagem disponível ao público, <b>mesmo para assintomáticos</b>
Stay-at-Home Restrictions	1 - <b>Sem</b> medidas de restrição
	2 - <b>Recomendado</b> ficar em casa
	3 - <b>Necessário</b> ficar em casa exceto <b>saídas essenciais</b>
	4 - <b>Necessário</b> ficar em casa com <b>poucas exceções</b>
Face Coverings	1 - <b>Sem</b> política de uso de máscaras
	2 - Uso <b>recomendado</b>
	3 - Uso <b>obrigatório</b> em <b>alguns</b> espaços públicos
	4 - Uso <b>obrigatório</b> em <b>todos</b> os espaços públicos
	5 - Uso <b>obrigatório todo o tempo</b> fora da residência
Vaccination Policy	1 - <b>Nenhum</b>
	2 - Disponível para <b>um dos grupos</b> (trabalhadores essenciais, idosos e grupos de risco)
	3 - Disponível para <b>dois dos grupos</b> (trabalhadores essenciais, idosos e grupos de risco)
	4 - Disponível para os <b>três grupos</b> (trabalhadores essenciais, idosos e grupos de risco)
	5 - Todos os <b>três grupos e parcelas da população</b>
	6 - Acesso <b>universal</b>

Fonte: Autoria própria, (2021).

Para se obter significância dos dados, as categorias utilizadas são idênticas às usadas pela OxCGRT, porém os dados diários foram concatenados em médias mensais, arredondadas para cima, com adição do valor um porque o *software* utilizado para processamento dos dados, *MATLAB®*, plataforma privada de programação multi-paradigma e ambiente de computação numérica desenvolvido pela, não aceita valores de entrada zero.

O *software* proprietário possui diversas parcerias com instituições de ensino e pesquisa, fornecendo licenças com desconto ou até mesmo versões estudantis e acadêmicas que possuem poucas diferenças quando comparadas à versão comercial.

Este trabalho foi desenvolvido utilizando a versão acadêmica do programa com licença disponibilizada pela Universidade Tecnológica Federal do Paraná(UTFPR) em parceria com a *MathWorks®*.

A capacidade de processamento e visualização do programa podem ser complementadas utilizando as *toolboxes* disponibilizadas pela empresa desenvolvedora através de um processo simples de instalação. Entusiastas e pesquisadores também podem desenvolver suas ferramentas (não somente *scripts*, mas interfaces e *toolboxes*) e compartilhar com os usuários da plataforma aumentando ainda mais seu alcance, variedade e especificidade.

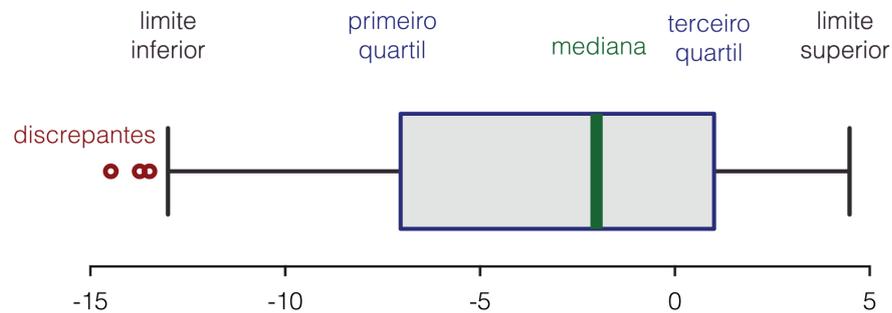
Inicialmente foram analisados os dados referentes às políticas adotadas no Brasil e as variâncias das amostras estão graficamente representadas de duas formas em histogramas e *boxplots*.

Um histograma é uma representação gráfica que organiza um grupo de pontos de dados em intervalos especificados, criado por Karl Pearson. O objetivo de um histograma é resumir graficamente a distribuição de um conjunto de dados univariados. Um histograma permite observar a distribuição da frequência de um conjunto de dados, demonstrando uma imagem relativizada de um padrão de distribuição, mapeado em categorias específicas (PEARSON, 1895).

Um *boxplot*, ou diagrama de caixas é uma forma padronizada de demonstrar a distribuição de certos dados com base em um padrão de cinco classes (*limite inferior*, *primeiro quartil*( $Q1$ ), *mediana*, *terceiro quartil*( $Q3$ ) e *limite superior*), representados na Figura 14. Pode ainda ajudar a encontrar valores discrepantes, além de mostrar simetria de dados, agrupação e distorções. Assim, é possível encontrar modos de

entender a variabilidade entre os dados utilizados (FRIGGE *et al.*, 1989).

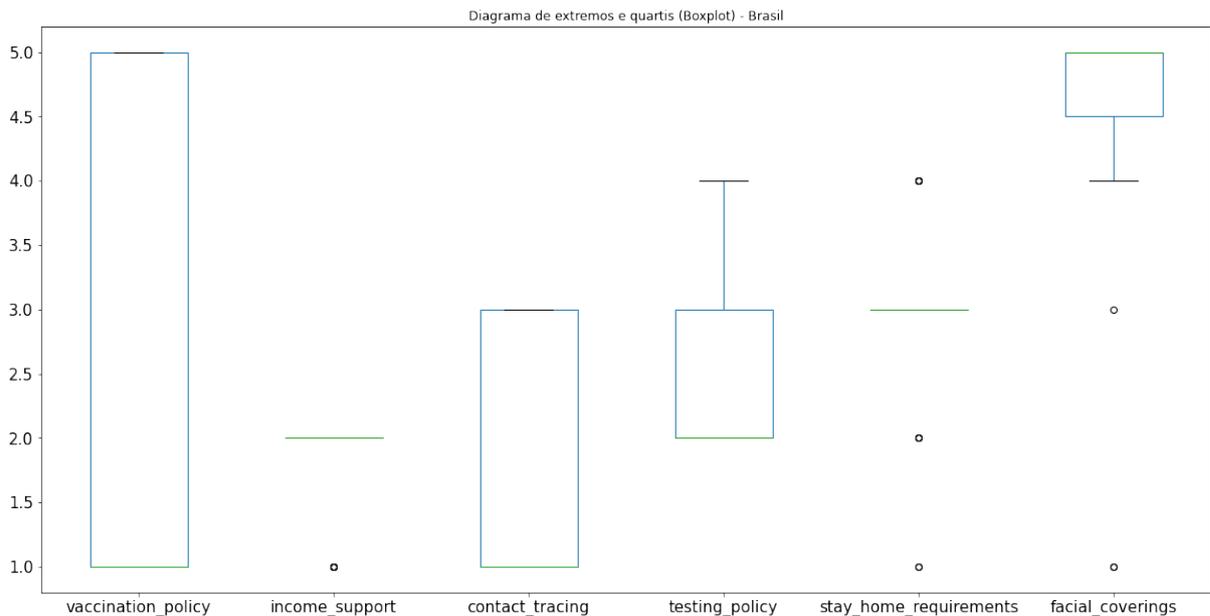
**Figura 14 – Elementos *boxplot***



Fonte: Autoria própria, (2021).

Os dados do Brasil podem ser observados na Figura 15, e no APÊNDICE A.

**Figura 15 – Boxplot dados Brasil**



Fonte: Autoria própria, (2021).

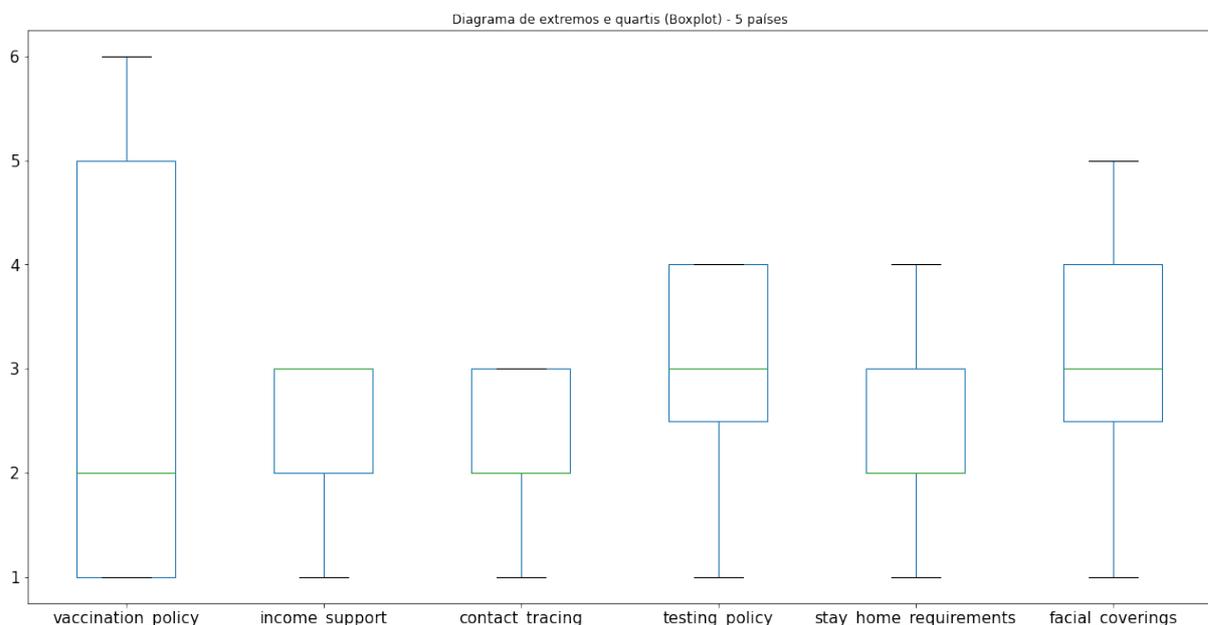
Para aumentar a diversidade de amostras foram coletados dados de mais cinco países, escolhidos a partir de alguns critérios. Porém, majoritariamente, foram escolhidos porque o Brasil não possuía entradas ou possuía poucas nas categorias que estes países se destacavam. A partir de uma análise entre as entradas para cada um dos níveis, temos que as melhores possibilidades para se compôr a amostra são:

- Finlândia: Possui entradas significativas no nível mais alto de *income\_support* (quarto maior), o nível 2. Assim como variedade de entrada em todos os níveis de *facial\_coverings*;

- Chile: Possui entradas significativas no nível 3 de *stay\_at\_home*, o segundo mais restritivo;
- Estados Unidos da América: O país possui o segundo maior nível de *testing\_policy*, nível 3, assim como é o pior no nível mais baixo de *contact\_tracing*, nível 1;
- Alemanha: Possui nível de *vaccination\_policy* (nível 1) ainda em dezembro de 2020. Foi um dos países mais adiantados na imunização via vacinação do mundo.

Os dados dos 5 países estão na Figura 16 e no APÊNDICE A.

**Figura 16 – Boxplot dados**



**Fonte: Autoria própria, (2021).**

Para a etapa de análise, a partir do *output* da rede, foi-se observado se há correlação de dados, se há coesão entre os indicadores e as relações posicionadas, assim como se estes fazem sentido com a realidade relatada por cada um dos países utilizado na base de dados.

A partir disto, poderão ser utilizados esses resultados para aplicação prévia de outras doenças epidemiológicas ou mesmo de menor escala (endemia), como um novo surto da gripe H1N1 ou outra doença que seja relacionada às síndromes respiratórias e que possua a mesma relação de infecção com o Covid-19, até mesmo outras variantes do coronavírus.

## 4 RESULTADOS E DISCUSSÃO

Nesta seção apresentar-se-á os resultados alcançados a partir da utilização dos algoritmos de estimação, assim como, a partir da estrutura do grafo direcionado acíclico, demonstrar as probabilidades dos eventos a partir de inferências probabilísticas escolhidas.

Os dados de *vaccination\_policy* apresentam alta discrepância entre os anos de 2020 e 2021. São 56 entradas da categoria 1 (sem política de vacinação) e 4 entradas da categoria 1 em 2020 e apenas 1 entrada da categoria 1 (sem política de vacinação) em 2021. Com base nessa informação decidiu-se separar os dados dos dois anos e realizar duas análises distintas, assim relativizando o foco da vacinação para o segundo ano.

Uma observação importante a ser feita é que apenas a política de disponibilidade de vacina não necessariamente implica em pessoas sendo de fato vacinadas. Numa relativização onde esta disponibilidade caminha diretamente em conjunto a porcentagem de pessoas vacinadas, não seria necessário uma variável auxiliar, porém isto não acontece. Algumas regiões enfrentam problemas logísticos, seja de entrega quanto de distribuição; a disponibilidade também não implica que a mentalidade pessoal esteja inferida, ou seja, a escolha do indivíduo em ser imunizado ou não (negacionismo).

Para que uma análise mais condizente fosse feita, foram então adicionadas à observação variáveis auxiliares, Tabela 4, referentes ao número de mortes pela doença confirmadas (*new\_deaths*), número de pessoas vacinadas (*people\_fully\_vaccinated*) e à diferença entre o número de mortos com o mês posterior (*deaths\_millions\_diff*).

Na realização do experimento, optou-se por não fazer nenhuma pressuposição de correlação a não ser que o evento final de 2020 seja o número de mortes e o evento inicial de 2021 seja *vaccination\_policy*. Utilizando essa abordagem o primeiro passo do experimento é encontrar os grafos que dão estrutura às redes bayesianas.

Para gerar a estrutura do grafo foi utilizada uma biblioteca da linguagem de programação *Python*, *Pomegranate*<sup>1</sup>, um pacote que implementa modelos probabilísticos rápidos e flexíveis que variam de distribuições de probabilidade individuais a modelos de composição, como redes Bayesianas e modelos ocultos de Markov.

<sup>1</sup> <https://pomegranate.readthedocs.io/en/latest/>

Tabela 4 – Variáveis Auxiliares

Variáveis auxiliares	Estados
New deaths (casos de morte confirmados por milhão de habitantes)	1 - De 0 a 0,1
	2 - Entre 0,1 e 0,2
	3 - Entre 0,2 e 0,5
	4 - Entre 0,5 e 1
	5 - Entre 1 e 2
	6 - Entre 2 e 5
	7 - Entre 5 e 10
	8 - Entre 10 e 20
	9 - De 20 a maior que 50
People fully vaccinated per hundred (Pessoas totalmente vacinadas por mil habitantes)	1 - De 0 a 10
	2 - Entre 10 e 100
	3 - Entre 100 e 250
	4 - Entre 250 e 500
	5 - Entre 500 e 1.000
	6 - Entre 1.000 e 1.500
	7 - Entre 1.500 e 2.000
	8 - Entre 2.000 e 2.500
Deaths millions diff (Diferença no número de mortes por milhão de habitantes do mês seguinte em relação ao mês atual)	1 - Entre -130 e -75
	2 - Entre -75 e -50
	3 - Entre -50 e -40
	4 - Entre -40 e -30
	5 - Entre -30 e -20
	6 - Entre -20 e -17
	7 - Entre -17 e -15
	8 - Entre -15 e -10
	9 - Entre -10 e -2
	10 - Entre -2 e 0
	11 - Entre 0 e 0,5
	12 - Entre 0,5 e 1
	13 - Entre 1 e 2
	14 - Entre 2 e 8
	15 - Entre 8 e 10
	16 - Entre 10 e 15
	17 - Entre 15 e 30
	18 - Entre 30 e 50
	19 - Entre 50 e 75
	20 - Entre 75 e 200

Fonte: Autoria própria, (2021).

Com a estrutura da rede definida foi utilizada uma *toolbox*<sup>2</sup> de *MATLAB*® para gerar a distribuição de Dirichlet e as CPT's do modelo.

A *toolbox* foi escrita e mantida por Kevin Murphy de 1997 a 2014 e encontra-se atualmente disponível para download em seu repositório na plataforma *GitHub*.

O código é vastamente documentado, orientado a objeto e de acesso gratuito, fazendo com que seja uma excelente ferramenta para o ensino pesquisa e prototipagem. Iniciantes no assunto e pesquisadores mais experientes que buscam resultado rápido podem se beneficiar muito com a solução.

<sup>2</sup> <https://github.com/bayesnet/bnt>

Uma lista com as principais capacidades do pacote desenvolvido por Murphy pode ser encontrada em <sup>1</sup>.

Segundo o autor<sup>1</sup> os principais motivos que o levaram a disponibilizar de forma gratuita a ferramenta foram:

- Que fosse criada uma comunidade como a comunidade *Linux*, onde mais pessoas contribuíssem com seus próprios códigos e o pacote crescesse;
- Desenvolver tendo em mente que mais pessoas irão ler o código exercita a habilidade de documentar o que está sendo feito além de ter mais pessoas para ajudar a corrigir eventuais *bugs*;
- Facilitar com que pesquisas sejam replicadas;
- Prover à comunidade científica ferramentas básicas de inferência e aprendizagem para que a pesquisa seja focada em questões e processos mais complexos.

Ele menciona ainda os motivos de ter escolhido uma plataforma proprietária<sup>2</sup> ao invés de uma de código aberto. Destacando que na época do início do desenvolvimento, a principal alternativa gratuita, o *Octave*<sup>3</sup>, não oferecia suporte a *arrays*, *cell arrays*, objetos, etc que fossem multidimensionais. Mas como em janeiro de 2010 essa funcionalidade já estava implementada, a *toolbox* provavelmente seria em grande parte compatível com o *software* que utiliza licença GNU General Public License (GPL), tornando a ferramenta 100% gratuita.

Os dados compilados neste trabalho e também os scripts utilizados para gerar os resultados estão disponíveis no repositório dos autores<sup>4</sup>, livre para acesso para que sejam replicados os experimentos e ideia é que progressivamente todo processo seja transferido para a linguagem *python* para que seja executada de forma *online*.

O intuito é continuar a proposta do desenvolvedor da *toolbox* e tornar o acesso a recursos básicos de modelagem e prototipagem de redes bayesianas cada vez mais amplo a fim de estimular a formulação de hipóteses mais complexas.

<sup>1</sup> <http://bayesnet.github.io/bnt/docs/majorFeatures.html>

<sup>1</sup> <https://github.com/bayesnet/bnt/wiki/WhyFree>

<sup>2</sup> <https://github.com/bayesnet/bnt/wiki/WhyMatlab>

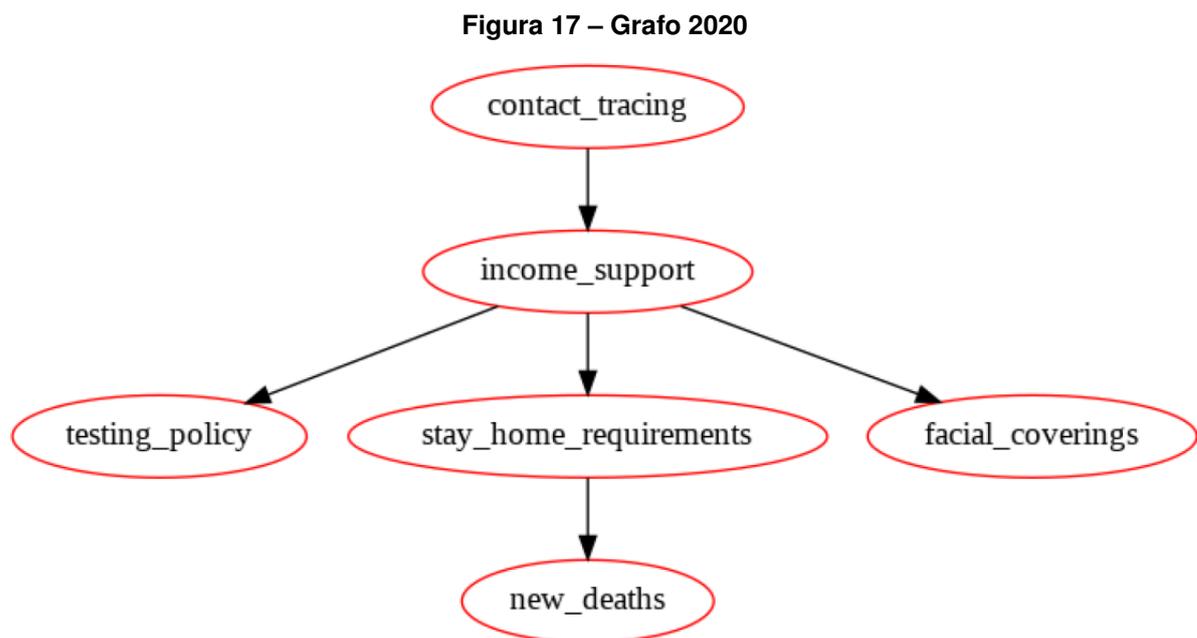
<sup>3</sup> <https://www.gnu.org/software/octave/index>

<sup>4</sup> <https://github.com/vrsoares/bbn>

#### 4.1 ANÁLISE DOS DADOS: 2020

O algoritmo utilizado para encontrar a estrutura que melhor representasse os dados de 2020 foi o *greedy search*, essa escolha se deu motivada pelo fato do algoritmo conseguir encontrar uma estrutura que ligasse todos os nós e tivesse como saída o nó que representa o total do número.

A estrutura encontrada pelo algoritmo pode ser vista na Figura 17:



**Fonte: Autoria própria, (2021).**

As CPT's do modelo estão nas tabelas de 6 a 11 no APÊNDICE B.1.

Agora, utilizaremos inferências probabilísticas para exemplificarmos e colocarmos possíveis casos. Para 2020, possuímos as seguintes situações:

- Caso *Stay-at-home* esteja no nível 3 (Obrigatório, menos para coisas essenciais), como estará a probabilidade para cada nível de morte (*new\_deaths*):

Como observado na tabela 7, temos que as probabilidades para cada nível de morte são:

	1	2	3	4	5	6	7
Nível 3	5,74%	5,74%	15,1%	13,02%	17,7%	24,98%	1,05%

Assim, podemos observar que, para o nível 3 da variável *Stay-at-Home*, as maiores probabilidades da variável *newdeaths* se encontram entre o nível 3 e 6,

mostrando que existe mitigação no número de mortes, porém dentro das poucas restrições impostas nesse nível ainda não é suficiente para uma diminuição mais drástica.

- Seja a variável de suporte financeiro (income support) nível 3 (o mais alto), como estará a probabilidade para cada nível de *stay-at-home*.

Como observado na tabela 6, temos que:

	1	2	3	4
Nível 3	45,34%	12,85%	13,13%	28,69%

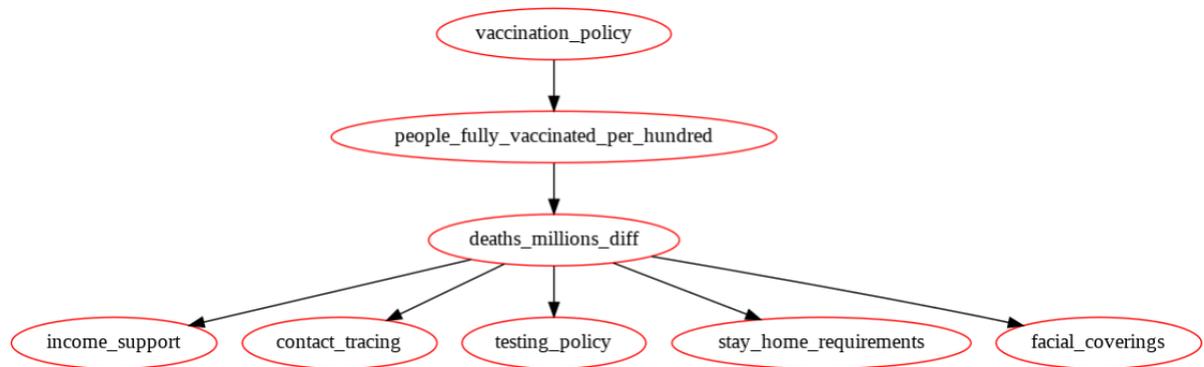
Assim, podemos observar que, para o nível 3 da variável *income\_support*, a maior probabilidade da variável *Stay-at-Home* está no nível mais baixo, sem restrição podendo indicar que a principal preocupação com a medida é econômica e não necessariamente humanitária. Alguns países demonstraram facilidade em oferecer este auxílio, porém não necessariamente existiram restrições mais proibitivas de *lockdown*.

#### 4.2 ANÁLISE DOS DADOS: 2021

Para uma Rede Bayesiana ser adequada, ela deve ser perfeita, ou seja, todos os arcos devem expressar corretamente as dependências entre as variáveis. Desta forma, é fácil notar que para a construção de uma Rede Bayesiana devemos escolher uma ordem correta para as variáveis, pois diferentes ordens podem gerar Redes Bayesianas diferentes. Desta forma, em (RUSSEL; NORVIG, 2004) os autores sugerem que primeiramente consideremos as variáveis possíveis a serem raízes e suas variáveis independentes, a seguir as demais variáveis.

Portanto, diante desta restrição, precisamos encontrar a estrutura que melhor representasse os dados de 2021, então foi escolhido como método *Chow-Liu tree* em tradução livre árvore de Chow-Liu. Um segundo método foi escolhido para o ano de 2021 pois, já com a vacinação presente, era necessário montar uma estrutura que um nó, agora nó raiz, fosse exatamente a variável de *vaccination\_policy*. Assim, para um nó raiz de estrutura, o método mais utilizado e mais rápido é o método de Chow-Liu (SUZUKI, 2012).

A estrutura encontrada pelo algoritmo pode ser vista na Figura 18.

**Figura 18 – Grafo 2021**

Fonte: Autoria própria, (2021).

As CPT's do modelo estão nas tabelas de 12 a 19 no APÊNDICE B.2.

Agora, utilizaremos inferências probabilísticas para exemplificarmos e colocarmos possíveis casos. Para 2021 porém, temos a introdução de uma variável auxiliar, diferença no número de mortes (*death\_millions\_diff*):

Para melhorar a visualização dos dados as categorias de diferença no número de mortes (*death\_millions\_diff*) foram agrupadas de acordo com a Tabela 5.

**Tabela 5 – Categorias de diferença de morte agrupadas**

<i>deaths_millions_diff</i>	1 - Diminuição Drástica (-130 a -20)
	2 - Diminuição (-20 a -2)
	3 - Estável (-2 a 2)
	4 - Aumento (2 a 30)
	5 - Aumento Drástico (30 a 200)

Fonte: Autoria própria, (2021).

Nesse cenário possuímos as seguintes situações:

- Caso a variável política de vacinação (*vaccination\_policy*) esteja nível 6 (Universal), como estará a probabilidade de cada nível da variável pessoas completamente imunizadas(*people\_fully\_vaccinated*).

Como observado na tabela 11:

	1	2	3	4	5	6	7	8
Nível 6	8,58%	33,30%	5,73%	3,82%	0,02%	11,43%	22,84%	14,28%

Onde o nível com maior probabilidade a partir de *vaccination\_policy* nível 6 é o nível 2, de 33,30% seguido do nível 7 com 22,84%, contrariando o que se espera numa análise crua que a probabilidade fosse crescente no nível 1 ao 8. O pico no nível 2 pode indicar que em alguns países, exista uma resistência à vacinação.

- Caso pessoas completamente imunizadas (*people\_fully\_vaccinated*) esteja nível alto (nível 6, 7, 8) estará a diferença no número de mortes (*death\_millions\_difference*):

	1	2	3	4	5	6	7	8	9	10
<i>Níveis 6,7,8</i>	1,67%	3,67%	3,33%	6,33%	6,67%	5,00%	6,00%	5,33%	5,67%	4,33%
	11	12	13	14	15	16	17	18	19	20
	3,33%	1,67%	4,00%	8,00%	4,33%	4,33%	6,67%	6,00%	8,33%	5,67%

Agrupando em 5 categorias:

	1	2	3	4	5
Nível 6,7,8	21,67%	22,00%	13,33%	23,33%	20,00%

Os altos números nas extremidades (aumento e queda) com a estabilidade (3) mostrando a menor probabilidade pode indicar que uma grande parcela da população vacinada pode diminuir muito o número de mortos mas que ainda existem picos no número de mortos, o que pode ser motivado por novas variantes do vírus.

## 5 CONCLUSÕES E PERSPECTIVAS

Neste trabalho foi realizada uma pesquisa experimental a fim de se entender quais são as possíveis influências a uma hipótese de mortalidade dentro de uma rede bayesiana, com o objetivo de concluirmos quais fatores de contenção ao vírus do COVID-19 realizaram tal influência em minimizar a variável da mortalidade neste conjunto de dados.

As metodologias utilizadas foram baseadas nas estruturas de uma rede bayesiana, utilizando métodos *greedy* e *Chow-Liu*. Os resultados nos demonstraram incertezas sobre os cenários, principalmente pela relativização que a estrutura bayesiana demonstra.

A partir de inferências escolhidas pelos autores, conseguimos identificar brevemente que, em 2020 por exemplo, existe uma grande correlação entre suporte financeiro (*income\_support*) e políticas de *lockdown* (*stay-at-home*), que por sua vez influência diretamente na nossa variável responsável pela mortalidade (*new\_deaths*). Portanto, contrariando, por exemplo, as recomendações do governo federal brasileiro, a contenção via *lockdown* (*stay\_at\_home*) é a mais indicada para reduzir de forma substancial as mortes pelo vírus em 2020.

Já em 2021, a introdução de uma variável de vacinação (*vaccination\_policy*) aponta uma melhora no diferencial de mortes de um mês a outro, a variável auxiliar *death\_diff*, juntamente ao nível crescente da disponibilidade da vacinação no mundo. Porém, a relação entre a parcela da população vacinada e número de mortes mostram que mesmo com a população vacinada podem haver picos de morte.

Há um forte indicativo que as políticas de mitigação ainda são necessárias diante do cenário demonstrado. Por exemplo, o surgimento de novas variantes pode não ocorrer em uma população com alta taxa de imunização, ao invés disso, podem surgir em países com baixa porcentagem vacinal da população e serem transportadas para esses países com altas porcentagens de vacinação, diminuindo assim a efetividade das vacinas. Desta forma, mostra-se ser indicado ainda para o ano de 2021 medidas de contenção já estipuladas, como o uso de máscara de proteção facial, redução de eventos com população elevada (e.g. Carnaval, feriados comemorativos de fim de ano) como forma preventiva a fim de evitar que uma nova variante se dissemine e que o

resultado dos trabalhos feitos pela sociedade civil de contenção, principalmente no que diz respeito a eficácia da imunização populacional, seja atenuado por um novo pico ou onda de contaminação.

Ficam como recomendações para trabalhos futuros o fato que o OxCGRT mantém sempre os dados atualizados incentivando fortemente que novas compilações com dados sempre recentes e análises mais minuciosas, por exemplo, concatenando dados por semana sejam realizadas. Outra abordagem seria regionalizar os dados ou até mesmo usar os dados diário. Também é possível utilizar de outros modelos matemáticos e de inteligência computacional para gerar resultados comparativos com esse trabalho ou mesmo complementares.

Um exemplo de aplicação seria utilizar uma rede LSTM (*Long Short-Term Memory*) na análise desse problema, já que essas redes são especializadas em dados sequenciais, onde o próximo dado tem íntima relação com o dado atual.

## REFERÊNCIAS

ARA-SOUZA, Anderson Luiz. **Redes Bayesianas: Uma Introdução aplicada a Credit Scoring**. 2010. Relatório Técnico — Universidade Federal de São Carlos, São Carlos, 2010.

BARIFOUSE, Rafael. **Coronavírus: primeira capital do Brasil em lockdown tem ruas lotadas e trânsito intenso**. 2021. Disponível em: <https://www.bbc.com/portuguese/brasil-52497230>. Acesso em: 18 de ago.2021.

BEN-GAL, Irad. Bayesian networks. **Encyclopedia of statistics in quality and reliability**, Wiley Online Library, v. 1, 2008.

BENGOETXEA, Endika. **Inexact graph matching using estimation of distribution algorithms**. 2002. Tese (Doutorado) — Citeseer, 2002.

BEZDEK, James. *What is Computational Intelligence?* **Computational Intelligence: Imitating Life**, J.M Zurada, R.J Marks II and C.J Robinson (Eds) *IEEE Press*, p. 1—12, 1994.

BICHINDARITZ, Isabelle; VAIDYA, Sachin; JAIN, Ashlesha (Eds.). *Computational intelligence in healthcare 4: Advanced methodologies*. In: \_\_\_\_\_. [S.l.]: Springer, 2010.

BRUNE, Jerome Seymour; AUSTIN, George Allen. *A study of thinking*. **Transaction publishers**, 1986.

CARUANA, Rich; NICULESCU-MIZIL, Alexandru. *An Empirical Comparison of Supervised Learning Algorithms*. **Proceedings of the 23rd international conference on Machine learning**, p. 161—168, Jun 2006. ISSN 10.1145/1143844.1143865.

CHOW, C.; LIU, C. Approximating discrete probability distributions with dependence trees. **IEEE Transactions on Information Theory**, v. 14, n. 3, p. 462—467, 1968.

CREECH, C Buddy; WALKER, Shannon C; SAMUELS, Robert J. *SARS-CoV-2 vaccines*. **Jama**, American Medical Association, v. 325, n. 13, p. 1318—1320, 2021.

DROR, Amiel A.; EISENBACH, Netanel; TAIBER, Shahar; MOROZOV, Nicole G.; MIZRACHI, Matti; ZIGRON, Asaf; SROUJI, Samer; SELA, Eyal. *Vaccine hesitancy: the next challenge in the fight against COVID-19*. **European Journal of Epidemiology**, p. 1—5, July 2020. ISSN 10.1007/s10654-020-00671-y.

FONSECA, João José Saraiva da. **Apostila de metodologia da pesquisa científica**. [S.l.]: João José Saraiva da Fonseca, 2002.

FRIGGE, Michael; HOAGLIN, David C; IGLEWICZ, Boris. Some implementations of the boxplot. **The American Statistician**, Taylor & Francis, v. 43, n. 1, p. 50–54, 1989.

G1. **Primeiro caso de COVID-19 no Brasil ocorreu em SP e completa seis meses nesta quarta**. **Globo**. 2020. Disponível em: <https://g1.globo.com/sp/sao-paulo/noticia/2020/08/26/primeiro-caso-confirmado-de-covid-19-no-brasil-ocorreu-em-sp-e-completa-seis-meses-nesta-quarta-feira.html>. Acesso em: 18 de jul.2021.

GALLO, Robert C. *The AIDS Virus*. **Scientific American**, v. 256, n. 1, p. 46–57, 1987.

GIL, Antonio Carlos *et al*. **Como elaborar projetos de pesquisa**. [S.l.]: Atlas São Paulo, 2002. v. 4.

HALE, Thomas; ANGRIST, Noam; GOLDSZMIDT, Rafael; KIRA, Beatriz; PETHERICK, Anna; PHILLIPS, Toby; WEBSTER, Samuel; CAMERON-BLAKE, Emily; HALLAS, Laura; MAJUMDAR, Saptarshi *et al*. A global panel database of pandemic policies (oxford covid-19 government response tracker). **Nature Human Behaviour**, Nature Publishing Group, v. 5, n. 4, p. 529–538, 2021.

JABBAR, M. A.; DEEKSHATULU, B. L.; CHANDRA, Priti. *2015 IEEE International Conference on Engineering and Technology (ICETECH)*. **Computational intelligence technique for early diagnosis of heart disease**, p. 1–6, 2015.

JASON, Janderson. **Recapitulando | PETNews - Redes Bayesianas**. 2012. <http://www.dsc.ufcg.edu.br/~pet/jornal/setembro2012/materias/recapitulando.html>. (Acesso em: 15 dez.2021).

JOHNSON, Niall; MUELLER, Juergen. *Updating the Accounts: Global Mortality of the 1918-1920 "Spanish" Influenza Pandemic*. **Bulletin of the history of medicine**, v. 76, p. 105–15, 02 2002. ISSN 10.1353/bhm.2002.0022.

JOYCE, James. *Bayes' theorem*. **Stanford Encyclopedia of Philosophy**, 2003.

KATELLA, Kathy. **5 Things To Know About the Delta Variant**. 2021. Disponível em: <https://www.yalemedicine.org/news/5-things-to-know-delta-variant-covid>. Acesso em: 18 de ago.2021.

KOONIN, E.V.; SENKEVICH, T.G.; DOLJA, V.V. *The ancient Virus World and evolution of cells*. **Biol Direct**, n. 1, p. 1–27, Set 2006. ISSN 10.1186/1745-6150-1-29.

KULLBACK, S.; LEIBLER, R. A. On information and sufficiency. **The Annals of Mathematical Statistics**, Institute of Mathematical Statistics, v. 22, n. 1, p. 79–86, 1951. ISSN 00034851. Disponível em: <http://www.jstor.org/stable/2236703>.

KUO, Lily. **Coronavirus: panic and anger in Wuhan as China orders city into lockdown | China | The Guardian**. 2020. Disponível em: <https://www.theguardian.com/world/2020/jan/23/coronavirus-panic-and-anger-in-wuhan-as-china-orders-city-into-lockdown>. Acesso em: 19 de jul. 2021.

LAM, BH. K.; LING, Steve S. H.; NGUYEN, Hung T. *Computational Intelligence and Its Applications: Evolutionary Computation, Fuzzy Logic, Neural Network and Support Vector Machine Techniques*. In: \_\_\_\_\_. [S.l.]: IMPERIAL COLLEGE PRESS, 2012.

LARRAÑAGA, Pedro; KARSHENAS, Hossein; BIELZA, Concha; SANTANA, Roberto. A review on probabilistic graphical models in evolutionary computation. **Journal of Heuristics**, Springer, v. 18, n. 5, p. 795–819, 2012.

LINDEN, Ricardo. **Algoritmos Genéticos**. [S.l.: s.n.], 2006. ISBN 8574523739.

MOHRI, Mehryar; ROSTAMIZADEH, Afshin; TALWALKAR, Ameet. **Foundations of Machine Learning**. [S.l.: s.n.], 2018.

MOORE, Zack S; SEWARD, Jane F.; LANE, J. Michael. *Smallpox*. **Lancet**, v. 367, n. 7, p. 425—435, Feb 2006. ISSN 10.1016/S0140-6736(06)68143-9.

MOREIRA, Ardilhes; PINHEIRO, Lara. **OMS declara pandemia de coronavírus | Coronavírus | G1**. 2021. Disponível em: <https://g1.globo.com/bemestar/coronavirus/noticia/2020/03/11/oms-declara-pandemia-de-coronavirus.ghtml>. Acesso em: 17 de jul.2021.

MORENS, David M.; FOLKER, Gregory K.; FAUCI, Anthony S. *What Is a Pandemic?* **The Journal of Infectious Diseases**, v. 200, n. 7, p. 1018—1021, Oct 2009. ISSN 10.1086/644537.

OURWORLDINDATA. **Share of people vaccinated against COVID-19**. 2021. <https://ourworldindata.org/covid-vaccinations?country=BRA>. (Acesso em: 04 dez.2021).

PEARL, Judea. Chapter 2 - bayesian inference. In: \_\_\_\_\_. **Probabilistic Reasoning in Intelligent Systems**. San Francisco (CA): Morgan Kaufmann, 1988. p. 29–75. ISBN 978-0-08-051489-5. Disponível em: <https://www.sciencedirect.com/science/article/pii/B978008051489500084>.

PEARSON, Karl. **Contributions to the Mathematical Theory of Evolution. II. Skew Variation in Homogeneous Material.** Zenodo, 1895. Disponível em: <https://doi.org/10.1098/rsta.1895.0010>.

PRAVALER. **Vírus, bactérias, fungos e protozoários - Entenda as diferenças.** 2020. Disponível em: <https://www.pravaler.com.br/virus-bacterias-fungos-e-protozoarios-entenda-as-diferencas/>. Acesso em: 01 de ago. 2021.

RUSSEL, S. J.; NORVIG, P. **Inteligência Artificial.** São Paulo: Campus, 2004.

SAMPSON, D.O. Stacy. **Pandemics: Past, present, and future.** 2020. <https://www.medicalnewstoday.com/articles/148945>. (Acesso em: 15 dez.2021).

SILVEIRA, Denise Tolfo; CÓRDOVA, Fernanda Peixoto. A pesquisa científica. **Métodos de pesquisa. Porto Alegre: Editora da UFRGS, 2009. p. 33-44,** 2009.

SUZUKI, Joe. The bayesian chow-liu algorithm. *In: The Sixth European Workshop on Probabilistic Graphical Models. [S.l.: s.n.], 2012. p. 315–322.*

UK. **UK medicines regulator gives approval for first UK COVID-19 vaccine.** 2020. Disponível em: <https://www.gov.uk/government/news/uk-medicines-regulator-gives-approval-for-first-uk-covid-19-vaccine>. Acesso em: 14 de jul.2021.

UOL. **Coronavírus: Bolsonaro diz que lockdown 'não dá certo' e volta a criticar governadores.** 2020. Disponível em: <https://noticias.uol.com.br/saude/ultimas-noticias/redacao/2020/05/14/bolsonaro-diz-que-lockdown-nao-da-certo-e-volta-a-criticar-governadores.htm>. Acessado em 19 de jul. 2021.

VALTORTA, Marco; BYRNES, John; HUHNS, Michael; ROHWER, Richard. *Combining Facts and Expert Opinion in Analytical Models via Logical and Probabilistic Reasoning Summary.* 12 2008.

WHO. **Question and Answers Hub.** 2019. Disponível em: <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/interactive-timeline>. Acesso em: 14 de jul.2021.

WHO. **Question and Answers Hub.** 2021. Disponível em: <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/question-and-answers-hub/q-a-detail/coronavirus-disease-covid-19>. Acesso em: 17 de jul.2021.

WORLDOMETERS. **Total Coronavirus cases in Brazil**. 2021. Disponível em: <https://www.worldometers.info/coronavirus/country/brazil/>. Acesso em: 18 em ago.2021.

WU, Di; WU, Tiantian; LIU, Qun; YANG, Zhicong. *The SARS-CoV-2 outbreak: what we know*. **International Journal of Infectious Diseases**, Elsevier, v. 94, p. 44–48, 2020.

YANG, Dong; LEIBOWITZ, Julian L. *The structure and functions of coronavirus genomic 3 and 5 ends*. **Virus research**, Elsevier, v. 206, p. 120–133, 2015.

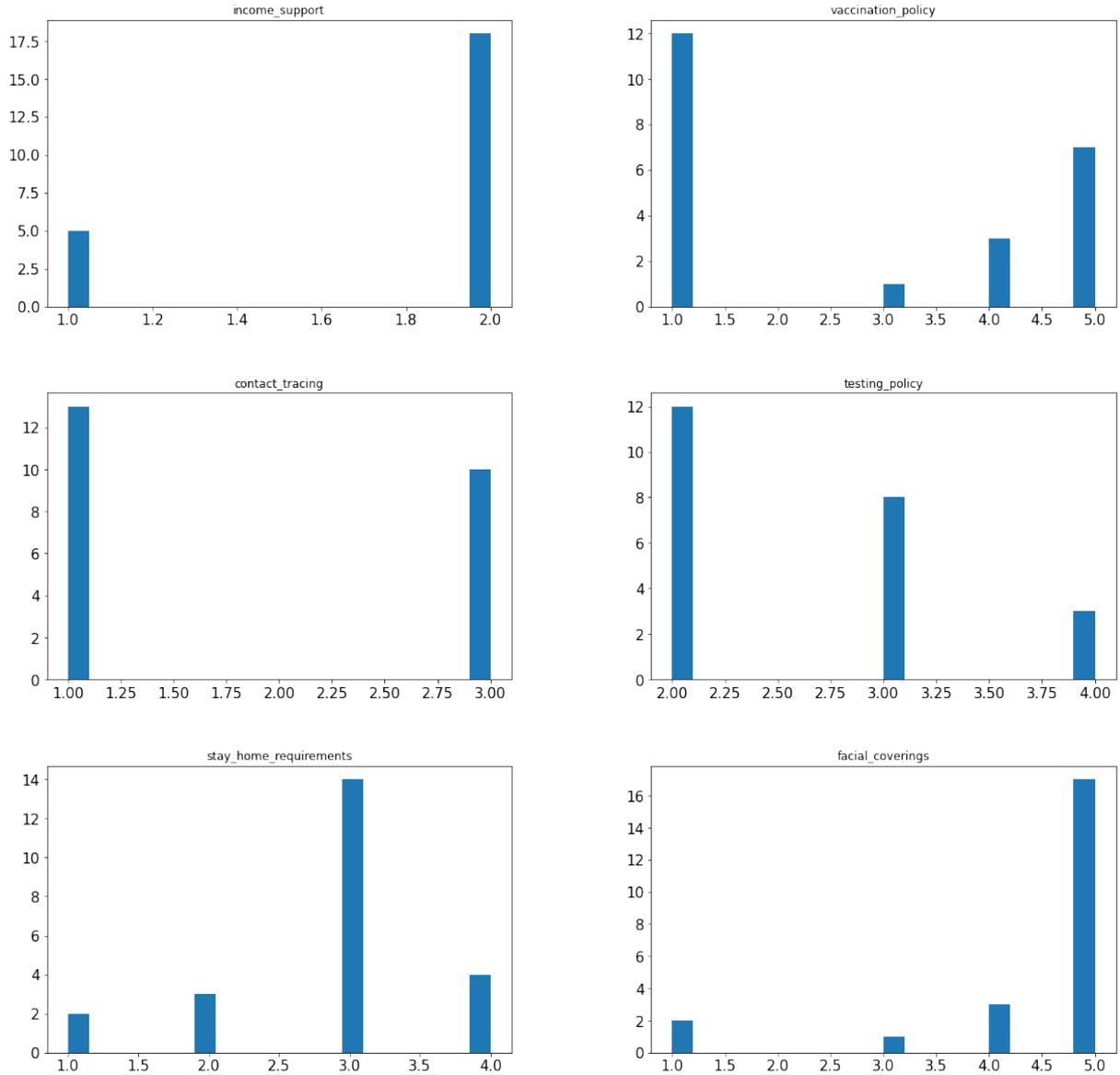
ZADEH, Lotfi A. *Soft Computing and Fuzzy Logic*. **IEEE Softw.**, IEEE Computer Society Press, Washington, DC, USA, v. 11, n. 6, p. 48–56, nov. 1994. ISSN 0740-7459. Disponível em: <https://doi.org/10.1109/52.329401>.

ZHU, Na; ZHANG, Dingyu; WANG, Wenling; LI, Xingwang; YANG, Bo; SONG, Jingdong; ZHAO, Xiang; HUANG, Baoying; SHI, Weifeng; LU, Roujian *et al*. *A novel coronavirus from patients with pneumonia in China, 2019*. **New England journal of medicine**, Mass Medical Soc, 2020.

## **APÊNDICES**

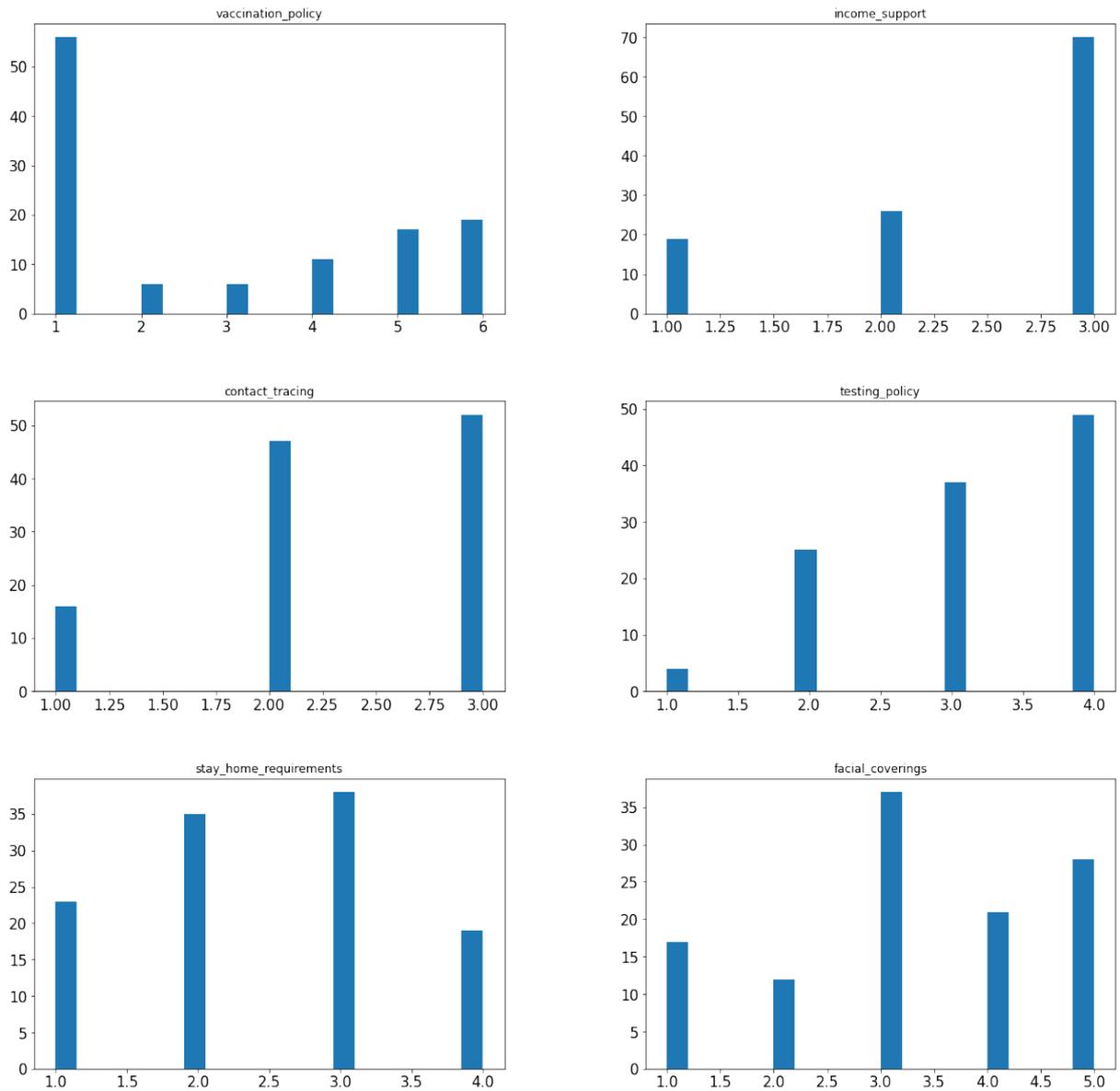
## APÊNDICE A – HISTOGRAMAS

**Figura 19 – Histograma dados Brasil**



Fonte: Autoria própria, (2021).

**Figura 20 – Histograma dados**



**Fonte: Autoria própria, (2021).**

## APÊNDICE B – CONDITIONAL PROBABILITY TABLE (CPT) | TABELAS DE PROBABILIDADE CONJUNTA

B.1 2020

**Tabela 6 – Probabilidade *contact\_tracing***

<i>contact_tracing</i>	P
1	0.2781
2	0.3689
3	0.3530

**Fonte: Aatoria própria.**

**Tabela 7 – CPT *income\_support|contact\_tracing***

<i>income_support contact_tracing</i>	1	2	3
1	0.1477	0.7154	0.1369
2	0.0301	0.6095	0.3604
3	0.1673	0.2805	0.5522

**Fonte: Aatoria própria.**

**Tabela 8 – CPT *stay-at-home|income\_support***

<i>stay-at-home income_support</i>	1	2	3	4
1	0.3780	0.0726	0.2253	0.3241
2	0.1607	0.0785	0.2275	0.5333
3	0.4534	0.1285	0.1313	0.2869

**Fonte: Aatoria própria.**

**Tabela 9 – CPT *new\_deaths|stay-at-home***

<i>new_deaths stay-at-home</i>	1	2	3	4	5	6	7
1	0.1438	0.1233	0.1643	0.0446	0.1027	0.0412	0.0548
2	0.0003	0.0107	0.1354	0.1146	0.1769	0.1457	0.0834
3	0.0574	0.0574	0.1510	0.1302	0.1770	0.2498	0.0105
4	0.1785	0.0191	0.1000	0.1262	0.0976	0.0453	0.2071

**Fonte: Aatoria própria.**

**Tabela 10 – CPT *testing\_policy|income\_support***

<i>testing_policy income_support</i>	1	2	3	4
1	0.2073	0.0996	0.5127	0.1804
2	0.0670	0.4282	0.1033	0.4014
3	0.2787	0.1749	0.4261	0.1203

**Fonte: Aatoria própria.**

Tabela 11 – CPT *face\_coverings|income\_support*

<i>face_coverings income_support</i>	1	2	3	4	5
1	0.3509	0.0096	0.2431	0.1982	0.1982
2	0.2562	0.1014	0.3231	0.1282	0.1912
3	0.2950	0.2431	0.1885	0.1094	0.1640

Fonte: Aatoria própria.

B.2 2021

Tabela 12 – Probabilidade *vaccination\_policy*

<i>vaccination_policy</i>	P
1	0.0951
2	0.0361
3	0.1470
4	0.4817
5	0.1350
6	0.1051

Fonte: Aatoria própria.

Tabela 13 – CPT *people\_vaccinated|vaccination\_policy*

<i>ppl_vaccinated vacc_policy</i>	1	2	3	4	5	6	7	8
1	0.0317	0.1473	0.2104	0.1789	0.0212	0.1053	0.2104	0.0948
2	0.3047	0.0006	0.2218	0.0006	0.0559	0.1388	0.2771	0.0006
3	0.0681	0.0885	0.1768	0.1768	0.1836	0.0341	0.1225	0.1496
4	0.1680	0.1224	0.1162	0.0830	0.1494	0.0394	0.0892	0.2323
5	0.0002	0.2665	0.1185	0.2147	0.3849	0.0002	0.0002	0.0150
6	0.0858	0.3330	0.0573	0.0382	0.0002	0.1143	0.2284	0.1428

Fonte: Aatoria própria.

**Tabela 14 – CPT deaths-million\_difference|people\_vaccinated**

<i>deaths people_vaccinated</i>	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	0.09	0.09	0.03	0.07	0.00	0.11	0.08	0.11	0.08	0.03	0.03	0.04	0.01	0.04	0.01	0.07	0.04	0.00	0.04	0.04
2	0.09	0.02	0.10	0.04	0.06	0.03	0.12	0.04	0.03	0.01	0.04	0.08	0.05	0.04	0.01	0.01	0.12	0.01	0.07	0.03
3	0.04	0.02	0.02	0.03	0.14	0.03	0.07	0.15	0.03	0.02	0.00	0.05	0.03	0.09	0.02	0.10	0.00	0.04	0.11	0.02
4	0.03	0.05	0.03	0.09	0.11	0.08	0.01	0.03	0.05	0.09	0.03	0.03	0.10	0.08	0.02	0.04	0.04	0.01	0.08	0.02
5	0.05	0.11	0.06	0.09	0.07	0.01	0.00	0.10	0.01	0.01	0.00	0.04	0.11	0.01	0.06	0.08	0.03	0.06	0.06	0.03
6	0.00	0.00	0.04	0.10	0.12	0.00	0.06	0.06	0.02	0.02	0.02	0.00	0.04	0.10	0.08	0.04	0.12	0.06	0.12	0.02
7	0.04	0.10	0.02	0.01	0.07	0.03	0.05	0.03	0.08	0.04	0.01	0.02	0.04	0.07	0.02	0.09	0.05	0.06	0.07	0.10
8	0.01	0.01	0.04	0.08	0.01	0.12	0.07	0.07	0.07	0.07	0.07	0.03	0.04	0.07	0.03	0.00	0.03	0.06	0.06	0.05

**Fonte: Autoria própria.**

Tabela 15 – CPT *face\_coverings\deaths-million\_difference*

<i>face_coverings\deaths-million_difference</i>	1	2	3	4	5
1	0.2330	0.1204	0.2871	0.1502	0.2093
2	0.2864	0.2314	0.1910	0.1315	0.1597
3	0.2770	0.2414	0.1076	0.1809	0.1930
4	0.2142	0.1906	0.1400	0.2720	0.1832
5	0.3749	0.1281	0.1226	0.2537	0.1207

Fonte: Autoria própria.

Tabela 16 – CPT *stay-at-home\deaths-million\_difference*

<i>stay-at-home\deaths-million_difference</i>	1	2	3	4
1	0.2969	0.3257	0.2710	0.1063
2	0.2300	0.1970	0.2914	0.2816
3	0.2544	0.2484	0.2720	0.2253
4	0.2293	0.2505	0.2118	0.3084
5	0.3820	0.2056	0.2377	0.1747

Fonte: Autoria própria.

Tabela 17 – CPT *testing\_policy\deaths-million\_difference*

<i>testing_policy\deaths-million_difference</i>	1	2	3	4
1	0.2027	0.2734	0.1693	0.3546
2	0.2331	0.1963	0.4486	0.1220
3	0.2266	0.1909	0.2505	0.3320
4	0.2830	0.2339	0.2030	0.2801
5	0.3203	0.0730	0.3221	0.2846

Fonte: Autoria própria.

Tabela 18 – CPT *contact\_tracing\deaths-million\_difference*

<i>contact_tracing\deaths-million_difference</i>	1	2	3
1	0.4405	0.2727	0.2868
2	0.2553	0.3751	0.3695
3	0.2513	0.3778	0.3709
4	0.4244	0.1556	0.4199
5	0.2331	0.3836	0.3833

Fonte: Autoria própria.

Tabela 19 – CPT *income\_support\deaths-million\_difference*

<i>income_support\deaths-million_difference</i>	1	2	3
1	0.3836	0.3821	0.2342
2	0.1157	0.2679	0.6164
3	0.3669	0.4754	0.1577
4	0.2939	0.3706	0.3355
5	0.2605	0.2394	0.5002

Fonte: Autoria própria.