

## CORONAVÍRUS E *DATA SCIENCE*: UMA BREVE ANÁLISE DO PANORAMA MUNDIAL E UM COMPARATIVO ENTRE OS PAÍSES NÓRDICOS

Iria Guazzi Linden – UNIESP – (iria.guazzi@hotmail.com), Alana Marques de Morais – UNIESP – (alanamm.prof@gmail.com), Aline Marques de Morais – UNIESP – (alinemm.prof@gmail.com)

### RESUMO

Este trabalho tem o intuito de analisar os dados relativos ao COVID-19 no mundo e traçar um comparativo entre os países nórdicos. Os países nórdicos se apresentaram como um recorte relevante por incluir países com contextos semelhantes, mas com abordagens distintas no combate ao COVID-19, com destaque para Suécia. Para tanto, foi utilizada uma base de dados pública, que é atualizada diariamente com as informações oficiais publicadas pelos países. Para a análise exploratória foram utilizadas técnicas de *Data Science* atreladas à linguagem Python. O estudo identificou a relevância das metodologias de registro de casos e mortes pelo coronavírus para que se possa comparar os dados dos países de forma eficaz, sendo indispensável que existam dados fidedignos e tempestivos. Tendo em vista as semelhanças entre os países nórdicos, a diferença apresentada nos números da Suécia em frente aos demais países chamou a atenção e reforçou a ponderação acerca das políticas de restrição e seus efeitos no controle da pandemia.

Palavras-chave: Ciência dos Dados. Coronavírus. COVID-19. Suécia. Países Nórdicos.

### ABSTRACT

This work aims to analyze the data related to COVID-19 in the world and draw up a comparison among the Nordic countries. The Nordic countries were relevant because they included similar countries and similar contexts, but with different approaches in combating against the COVID-19, with emphasis on Sweden. To this end, a public database was used, which is updated daily with the official information published by countries. For the exploratory analysis, Data Science methods linked to the Python language were used. The study identified the relevance in the methodologies for registering cases and deaths from coronavirus so the comparison can be made effectively, and that reliable and timely data is essential. Because of the similarities among the Nordic countries, the differences presented in the numbers of Sweden compared to the other countries drew attention and reinforced the consideration about the restriction policies and their effects in the control of the pandemic.

Keywords: *Data Science*. Coronavirus. COVID-19. Sweden. Nordics Countries.

## 1 INTRODUÇÃO

A COVID-19 é a maior ameaça à saúde pública mundial desde a pandemia de Gripe espanhola, provocada pelo vírus H1N1, em 1918 (FERGUSON *et al.*, 2020). A sigla COVID significa *Corona Virus Disease* (Doença do Coronavírus) e o número 19 se refere ao ano 2019, quando os primeiros casos foram divulgados publicamente.

Essa doença foi identificada pela primeira vez em dezembro de 2019, na cidade de Wuhan – China, como um surto de pneumonia (ZHU, 2020). Em menos de 15 dias, já havia

notificação de casos do novo coronavírus no Japão, Coréia do Sul, Taiwan e Tailândia, mostrando a eficiência da contaminação para outros países. Em 11 de março de 2020, a OMS (Organização Mundial de Saúde) declarou a doença como uma pandemia global e, em 23 de março de 2020, o vírus já havia afetado 172 dos 195 países (SHAW, 2020).

À medida que a pandemia do COVID-19 avança, é possível esperar que os dados tenham um papel cada vez mais relevante nas operações públicas e privadas. Os dados podem subsidiar as tomadas de decisão e deliberações dos governantes ou seus representantes.

As estratégias de combate ao coronavírus são definidas por cada país individualmente, quer seja por dados históricos de situações que aconteceram no passado, ou por adotar medidas que funcionaram em outros países ou, ainda, por tentar uma abordagem definida pelos seus especialistas. Um país que chama a atenção por contrariar o direcionamento mundial de isolamento social é a Suécia. Além disso, seus números são muito distintos dos seus vizinhos Noruega e Dinamarca, que tiveram as primeiras mortes registradas no mesmo período que a Suécia (entre 11 e 14 de março). O Brasil segue em suas diretrizes de combate ao COVID-19 alguns posicionamentos semelhantes ao da Suécia, por isso se faz necessário entender o impacto de tais ações nos contextos locais.

As metodologias de registros adotadas pelos países afetam diretamente os resultados dos estudos. Muitas vezes é necessário realizar uma análise comparativa para entender as projeções do COVID-19 no cenário mundial. Contudo, a falta de um padrão de registros dificulta as comparações entre os países, bem como a identificação de correlação entre as medidas tomadas por cada nação e a diminuição (ou aumento) no número de casos e mortes.

Diante deste contexto, o presente artigo apresenta uma análise dos dados abertos, disponibilizados pelos países, acerca do comportamento do COVID-19 nas nações, bem como uma comparação das informações relativas apenas aos países nórdicos, em virtude de suas semelhanças geográficas, culturais, sociais e econômicas.

As próximas seções apresentam a fundamentação teórica na qual este estudo se baseia e que são fundamentais ao entendimento do artigo; a descrição da metodologia adotada pelo estudo; o desenvolvimento das análises e seus resultados; as conclusões obtidas e considerações finais do artigo; e ao final, uma lista das referências bibliográficas citadas.

## 2 FUNDAMENTAÇÃO TEÓRICA

A presente seção tem como intuito discutir e fundamentar conceitos importantes ao entendimento do presente estudo. Inicialmente, há o detalhamento sobre a terminologia *Data Science* e sua aplicação em processos decisórios, destacando os casos que obtiveram sucesso no controle do número de casos por meio do uso de tecnologias. Outro ponto abordado, é a metodologia de registros dos números pelos países, que não obedece a um padrão.

### 2.1 DATA SCIENCE EM PROCESSOS DECISÓRIOS

O volume de dados gerado nos últimos anos apresentou um crescimento exponencial e a melhoria na capacidade de processamento dos hardwares deram protagonismo à área de *Data Science* como uma ferramenta indispensável no suporte à tomada de decisão. Segundo Provost (2016), “*Data Science* é um conjunto de princípios fundamentais que norteiam a extração de conhecimento a partir de dados”.

Diante do atual cenário de pandemia, o setor de tecnologia oferece soluções no combate ao surto de coronavírus, principalmente nos segmentos de ciência de dados e inteligência artificial.

De acordo com Park (2020), a Coreia do Sul tem utilizado dados extraídos de câmeras de vigilância, obtido a localização de cidadãos por meio de celulares e monitorado as transações efetuadas através de cartões, a fim de mapear as possíveis conexões sociais entre pessoas saudáveis e os casos suspeitos. Além disso, eles também implementaram testes de diagnóstico rápidos e abrangentes (cerca de 250.000 testes). Em um segundo momento, caso esses pacientes suspeitos não pudessem ser encontrados, a capacidade de teste em massa não significaria muito. É exatamente neste ponto que a Coreia do Sul se destacou, pois o país se utilizou da análise e do cruzamento dos dados para descobrir onde estavam os pacientes conhecidos e testar qualquer pessoa que possa ter entrado em contato com eles. Desta maneira, todas as informações geradas a partir desses mapeamentos permitem que o movimento de um novo paciente possa ser comparado com o de pacientes anteriores.

A China também fez uso de diversas tecnologias no combate ao COVID-19, principalmente no que diz respeito ao controle do fluxo dos infectados, como por exemplo: rastreamento de contágio, localização por meio de celulares, identificação facial com controle de temperatura, além da criação de um “*health code*” que identifica as pessoas afetadas por meio de um aplicativo. Com a utilização dessa ferramenta, “o mapeamento *on-line* das pessoas afetadas pode ser feito e as pessoas podem evitar os agrupamentos onde as pessoas afetadas estão concentradas” (SHAW, 2020).

Cabe destacar que todos esses métodos geram informações precisas, transparentes e tempestivas acerca da epidemia e seu desenvolvimento e, em uma pandemia como a COVID-19, a tecnologia, a inteligência artificial e a *Data Science* são essenciais para ajudar as sociedades a lidar efetivamente com o surto, permitindo que os gestores tenham informações em tempo real e possibilitem a análise das melhores políticas a serem adotadas na redução dos impactos causados pelo vírus. Os casos de sucesso da China e da Coreia do Sul no controle da pandemia atestam que a tecnologia é uma forte aliada no entendimento e previsão dos efeitos das ações adotadas no combate à evolução da doença.

## 2.2 METODOLOGIAS DE REGISTRO E O IMPACTO NA ANÁLISE DOS DADOS

O acesso a dados precisos e completos tem um impacto direto nos resultados de quaisquer trabalhos relacionados a prevenção do COVID-19. Os números aumentam diariamente, o que faz com que o trabalho de controle e atualização dos dados sejam fundamentais para que as análises reflitam a real situação dos países, pois conclusões equivocadas podem acarretar em decisões ineficazes contra a expansão do vírus e os efeitos das más decisões só podem ser percebidas em um intervalo de duas semanas.

Provost (2016) enfatiza que “(...) precisamos ter em mente que os dados aos quais aplicamos as técnicas de *Data Science* são o produto de algum processo que envolveu decisões humanas”. Assim, um dos pontos que devem ser levados em consideração é a metodologia utilizada pelos países para os registros dos fatos relacionados ao coronavírus, principalmente no que se refere ao número de mortes.

Para a maioria dos países, o número de mortes por COVID-19 é registrado apenas nos casos de pacientes que se encontravam no hospital e que testaram positivo para o coronavírus. A Suécia, por exemplo, contabiliza as mortes pelo vírus apenas das pessoas que testaram positivo antes de falecer. Mesma metodologia é adotada pelos demais países Nórdicos. Já na Bélgica, o registro considera, inclusive, a morte de pessoas não hospitalizadas com suspeita de

infecção pelo vírus. Em se tratando de países mais pobres, essa disparidade deverá ser ainda maior, levando em consideração a baixa capacidade para efetuar testes na população (THE ECONOMIST, 2020).

Conforme análise feita pelo The Economist (2020) “As diferenças no registro das mortes por COVID-19 dificultam a comparação das estatísticas entre os países e podem significar que o número da Bélgica é mais realista ou até menor do que os dados sugeridos agora”. Neste sentido, o estudo mencionado adota uma abordagem que, para muitos, pode ser mais eficiente na mensuração dos efeitos causados pela pandemia, pois leva em conta o “excesso de mortes” que ultrapassam a média histórica para o mesmo local e época do ano. Na Itália, por exemplo, o número mortes que superam a média, no período de 25/02/2020 a 30/03/2020 foi de 25.807 e, conforme números oficiais, o número de mortes causadas por COVID-19 foi de 12.418. Ou seja, apenas 48% do total de “excesso de mortes” foi registrado como tendo sido causado pelo novo coronavírus.

Desta forma, para fins de pesquisas e modelos, é muito importante que se entenda a composição dos dados e a forma que são registrados na base que será analisada, tendo em vista que os resultados obtidos serão utilizados como base para a construção de soluções frente à pandemia.

### 3 METODOLOGIA

A presente seção tem o intuito de apresentar a metodologia percorrida na execução do estudo de caso exploratório executado no presente trabalho. Os dados analisados estão relacionados ao COVID-19 nos países e foram obtidos no portal *Our World in Data* (<https://ourworldindata.org/coronavirus>), que atualiza diariamente suas bases, de acordo com as novas informações que são fornecidas pelas nações. O acesso ao sítio ocorreu em 09 de junho de 2020.

Os trabalhos relacionados ao *Data Science* seguem etapas bem definidas, conforme pode ser observado na figura 1. Assim, o estudo se inicia com a coleta de dados no portal mencionado (Etapa 1), este é uma fonte de informação relevante e confiável. A partir da base selecionada, haverá o pré-processamento e limpeza dessa amostra, para que sejam selecionadas apenas as variáveis relevantes à pesquisa (Etapa 2). “Durante a análise e a modelagem dos dados, um período significativo de tempo é gasto em sua preparação: carga, limpeza, transformação e reorganização.” (MCKINNEY, 2018, p.248).

Em seguida, inicia-se a Etapa 3, ou seja, o processo de análise exploratória, no qual há a inspeção dos dados e suas propriedades, visando identificar padrões, possíveis anomalias, bem como verificar suposições. Essa análise busca construir uma narrativa a partir das informações obtidas e o cientista de dados se utiliza da parte visual como meio facilitador na compreensão da história contada (VANDERPLAS, 2016). Para esta etapa, utilizou-se a linguagem Python que, segundo McKinney (2018), nos últimos dez anos passou a ser uma das linguagens mais importantes em *Data Science*, aprendizado de máquina (*machine learning*) e desenvolvimento de softwares em geral, no ambiente acadêmico e no mercado.

Figura 1 - Etapas da *Data Science*.



Fonte: Autor (2020)

A análise aprofundada (Etapa 4) contou com o estudo das variáveis: testes, casos, mortes, população, PIB *per capita*, índice de rigidez, entre outras. Conforme destacado anteriormente, as regras de contabilização podem mudar de país para país, fato que dificulta as análises comparativas. Assim, o estudo levou em consideração a evolução histórica no panorama mundial e, em seguida, no que se refere aos países Nórdicos (Suécia, Noruega, Dinamarca, Finlândia e Islândia). A partir da fonte de dados, também foi possível efetuar o cálculo do coeficiente de letalidade por país. Os resultados obtidos por meio da análise retratada anteriormente serão apresentados no tópico a seguir.

Por fim, esta é uma pesquisa descritiva que, segundo Walliman (2015), almeja examinar situações de modo a estabelecer um padrão que pode ocorrer em outros sob as mesmas circunstâncias.

### 3.1 FERRAMENTAS

O *Data Science* envolve princípios, processos e técnicas para compreender fenômenos por meio da análise (automatizada) de dados (PROVOST, 2016). Na presente pesquisa, foram executados experimentos desenvolvidos por meio da linguagem de programação Python, que é uma linguagem interpretada, de alto nível, permite análise, processamento e visualização de dados. Através da linguagem, foram utilizadas bibliotecas - coleção de funções e classes - que facilitam os processos de análise de dados como pandas e matplotlib.

O Pandas oferece estruturas de dados de alto nível e funções, projetada para fazer com que trabalhar os estruturados ou tabulares seja rápido, fácil e expressivo (MCKINNEY, 2018). Já a matplotlib é a biblioteca que nos permite fazer plotagens e gerar as visualizações dos dados.

O interpretador da linguagem Python utilizado foi o *Jupyter Notebook*, que é uma plataforma de código aberto que permite a criação de documentos executam código em Python ou outras linguagens, permitindo a fácil leitura, execução e replicação deles. Destaca-se o ambiente interativo e sua facilidade para a integração com bibliotecas como Matplotlib e Pandas (MCKINNEY, 2018).

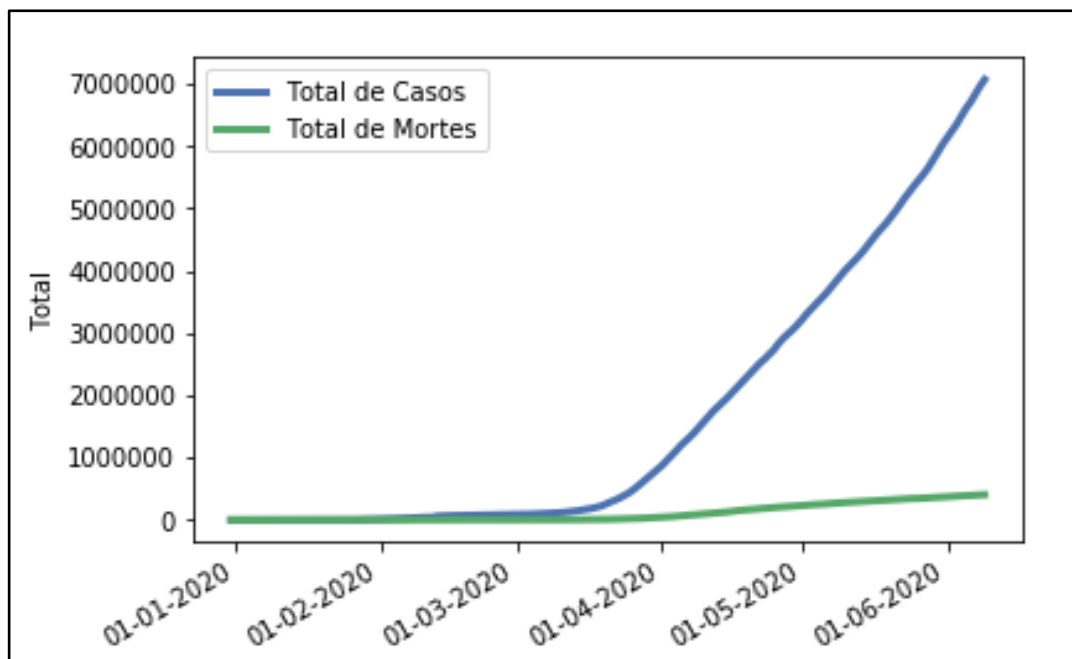
## 4 DESENVOLVIMENTO

Nesta seção, será apresentada a análise exploratória dos dados, levando-se em consideração o contexto dos países na pandemia. Primeiro, serão apreciados os dados relativos aos números globais e, em seguida, uma avaliação dos países Nórdicos foco da análise.

### 4.1 COVID-19 NO MUNDO

A pandemia do coronavírus está afetando 213 países e territórios ao redor do mundo. Segundo informações da OMS, o mundo teve, no dia 07/06/2020, o número mais alto de novas infecções diárias por COVID-19 desde o início da pandemia, representando 136.409 (cento e trinta e seis mil, quatrocentos e nove) novos diagnósticos.

Figura 2 - Total de casos e mortes no mundo



Fonte: Autor (2020)

Até 09 de junho de 2020 o total de casos confirmados alcançou o número de 7.069.278 de registros da doença. Já a quantidade de mortos, até a referida data, atingiu a marca de 405.587 de mortes. De acordo com a figura 2, os casos ultrapassaram 2 milhões de casos em abril, e em aproximadamente 2 meses depois, chegaram a 7 milhões casos de COVID-19.

Outro ponto de destaque é que, inicialmente, os casos confirmados coincidem com as mortes. Ou seja, é um forte indicador que os testes eram realizados apenas nos casos relacionados à óbito no mundo.

### 4.2 TESTES E COEFICIENTE DE LETALIDADE

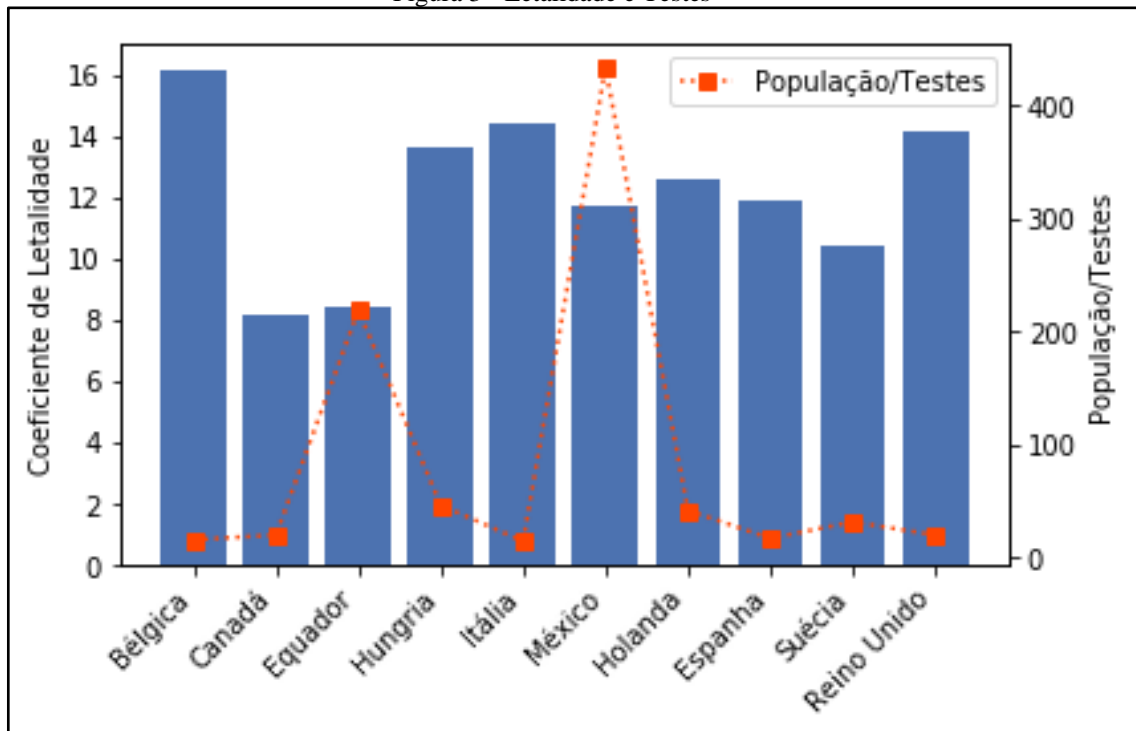
Nenhum país sabe ao certo o total de pessoas infectadas com COVID-19. As diferenças nos números de testes realizados significam que a quantidade de casos pode estar subestimada

em alguns países. Os números apresentados na variável “testes” se referem aos infectados que tiveram resultados positivos em seus exames de laboratório. Assim, estes números estão diretamente ligados à quantidade de testes que o país realiza, ou seja: sem testes, sem dados.

Os testes são fundamentais para que as medidas de controle contra a expansão do coronavírus sejam tomadas. Eles permitem identificar a evolução da doença naqueles que foram infectados, bem como os tratamentos que estão sendo utilizados e seus resultados. Além disso, é possível concluir se as medidas adotadas estão surtindo efeito e, com isso, se há uma retração na quantidade de contaminações, já que este é o principal objetivo para o controle da doença.

O coeficiente de letalidade é representado pelo total de mortes em comparação com o total de casos confirmados. A figura 3 apresenta os 10 países “mais letais”. Além disso, foi incluída a informação da relação entre a população e o total de testes realizados. A Bélgica, por exemplo, testou 1 a cada 15 habitantes; em contrapartida, o México testou 1 em cada 432.

Figura 3 - Letalidade e Testes



Fonte: Autor (2020)

Excetuando-se México e Equador, os demais países estão entre os mais ricos do mundo. Em um primeiro momento, é possível questionar se o Coronavírus é uma doença que atinge apenas os países mais ricos, ou se apenas estes conseguem identificar o vírus em virtude da quantidade de testes que são capazes de efetuar.

Dos países que menos testam, proporcionalmente em relação ao total de sua população (quadro 1), é possível notar que todos se destacam por possuírem baixos valores de PIB *per capita* (Produto Interno Bruto pela quantidade de habitantes), o que os classifica como países pouco desenvolvidos. A Nigéria é o pior país, dos que disponibilizam dados, neste sentido: apenas 1 a cada 2.579 pessoas são testadas.

Quadro 1 - Países que menos testam por população

País	Total Mortes	Total Casos	Total Testes	PIB per capita	População	População / Testes
------	--------------	-------------	--------------	----------------	-----------	--------------------

<b>Nigéria</b>	361	12.801	79.948	5.338,45	206.139.587	2.579,42
<b>Myanmar</b>	6	244	39.211	5.591,60	54.409.794	1.387,61
<b>Indonésia</b>	1.883	32.033	281.653	11.188,74	273.523.621	971,14
<b>Etiópia</b>	27	2.156	152.334	1.729,92	114.963.583	754,68
<b>Zimbábue</b>	4	287	21.354	1.899,77	14.862.927	696,02
<b>Quênia</b>	85	2.862	97.340	2.993,03	53.771.300	552,40
<b>Brasil</b>	36.455	691.758	485.000	14.103,45	212.559.409	438,26
<b>México</b>	14.053	120.102	297.963	17.336,47	128.932.753	432,71
<b>Uganda</b>	0	646	112.678	1.697,71	45.741.000	405,94
<b>Bangladesh</b>	930	68.504	410.931	3.523,98	16.468.938	400,77

Fonte: Autor (2020)

Com as análises dos dados apresentados, as hipóteses mais prováveis que se podem suscitar são: a) o coronavírus demorou mais para chegar a esses países porque são menos conectados, levando-se em consideração tráfego de pessoas (aeroportos, portos, vias terrestres); ou b) esses países ainda não foram capazes de investir o suficiente em testes para saber a real situação de contaminação.

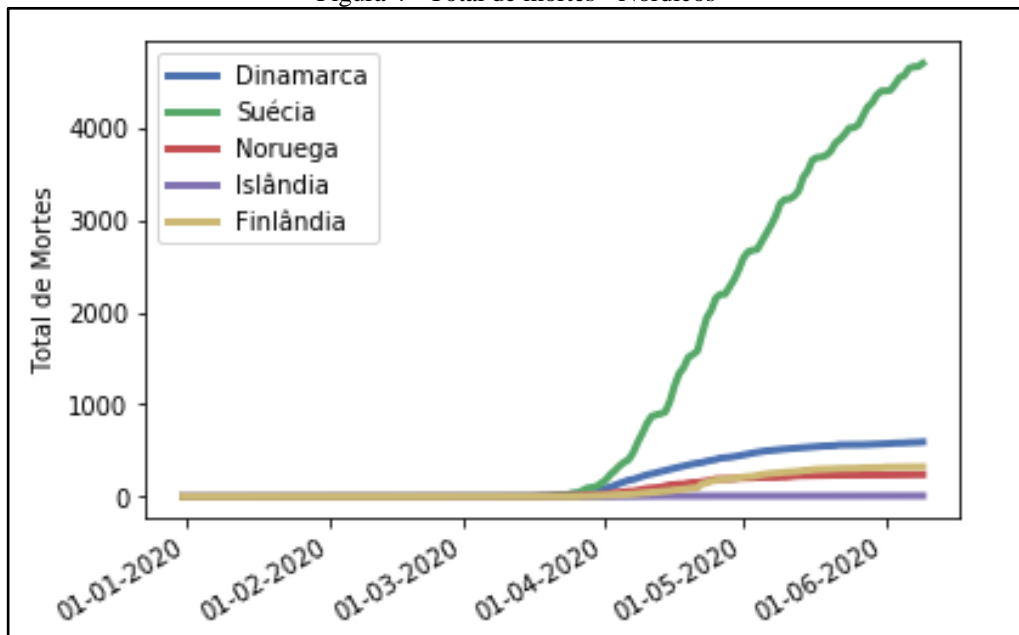
### 4.3 OS NÓRDICOS E SEU *OUTLIER* SUÉCIA

O termo *outlier* pode ser definido como todo aquele objeto que se distancia de forma considerável do restante dos dados (HAN; PEI; KAMBER, 2011). Um dos pontos que chama a atenção durante a análise exploratória dos dados é a enorme diferença dos números da Suécia em relação aos outros países Nórdicos (Noruega, Dinamarca, Finlândia e Islândia). Por serem geográfica, cultural, social e economicamente parecidos, seria esperado que houvesse uma harmonia nos dados deste grupo de países.

Um dos primeiros dados que fazem a Suécia se destacar é no que diz respeito ao total de mortes, sendo evidente que os números, além de ser muito maior quando comparado aos demais, tais índices não param de crescer (figura 4). Até o dia 09 de junho de 2020, a Suécia já tinha acumulado um total de 4.694 mortes, enquanto os demais países Nórdicos juntos totalizavam 1.165. Ainda, de acordo com a figura 3, verifica-se que a Suécia se encontra entre os 10 países mais letais (total de mortes sobre o total de casos).



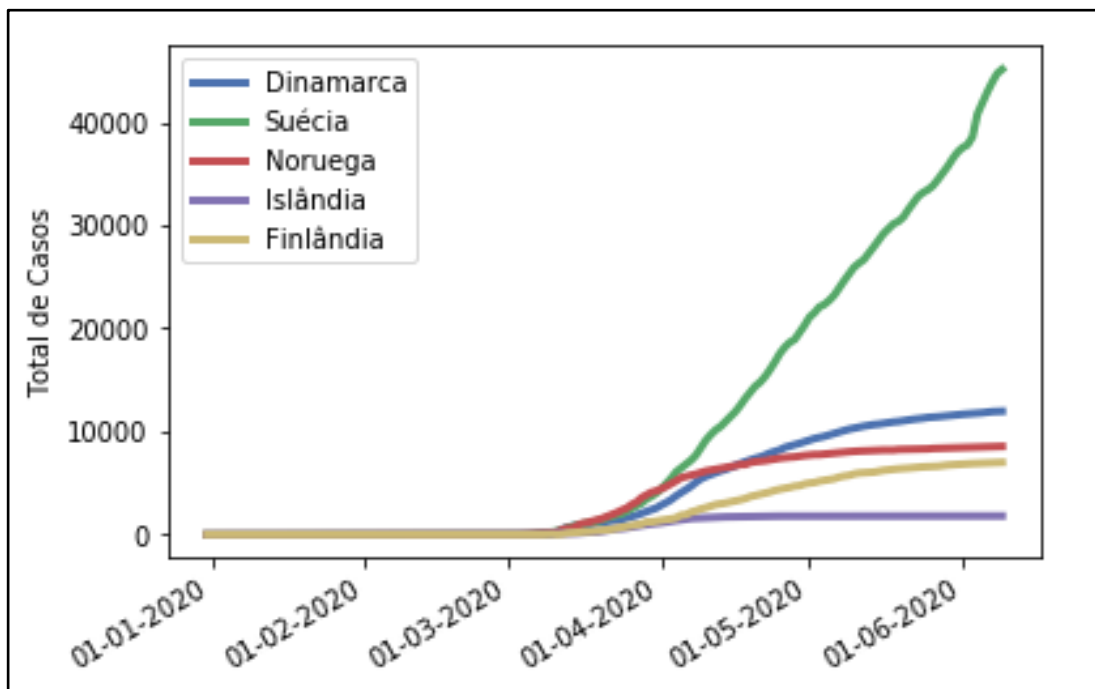
Figura 4 - Total de mortes - Nórdicos



Fonte: Autor (2020)

Conforme pode ser observado na figura 5, excetuando-se a Suécia, os demais países nórdicos conseguiram controlar a expansão da epidemia, considerando que o número total de casos se mostra constante ou com uma pequena variação, como a Dinamarca.

Figura 5 - Total de Casos - Nórdicos



Fonte: Autor (2020)

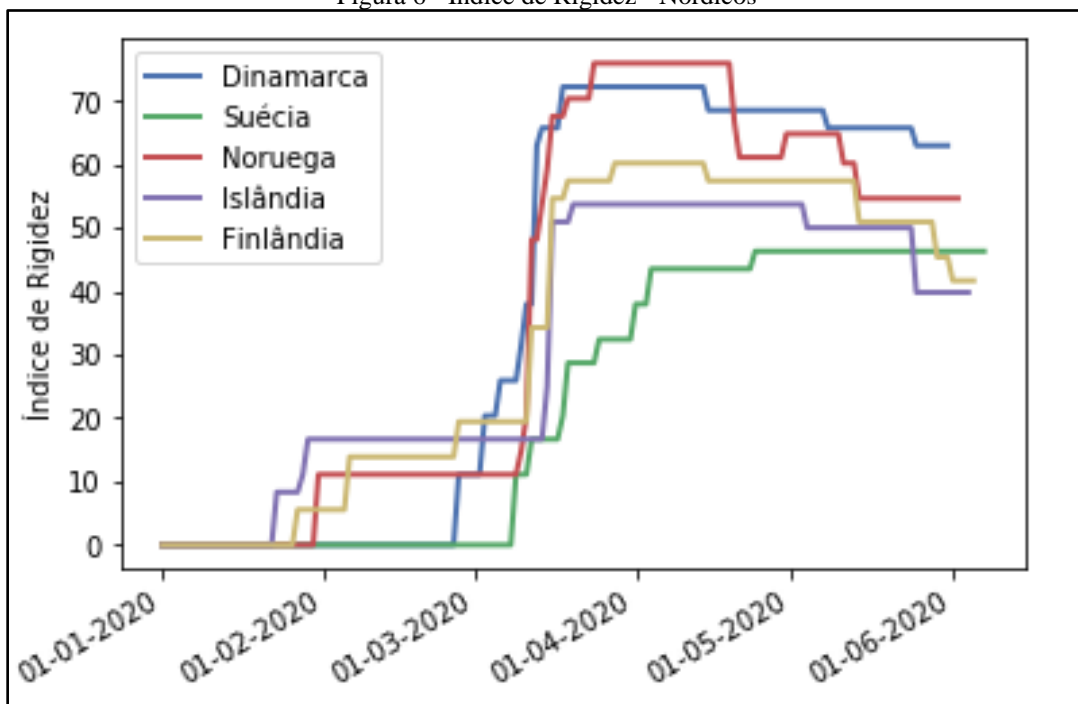
### 4.3.1 POLÍTICAS DE RESTRIÇÃO

Uma das variáveis constante na base de dados utilizada no presente artigo é o *Government Stringency Index* (Índice de Rigidez Governamental), que é uma medida criada pelo *Our World in Data*, composta de 9 métricas: fechamento de escolas, fechamento de locais de trabalho, cancelamento de eventos públicos, restrições a reuniões públicas, fechamentos de transportes públicos, requisitos de “ficar em casa”, campanhas de informação pública, restrições de movimentos internos e controle de viagens internacionais. O índice vai de 0 a 100, sendo importante destacar que ele por si só não significa necessariamente que os resultados alcançados após as medidas sejam melhores nos países cujo índice são mais altos.

A Suécia é o único país nórdico e um dos poucos países da Europa que não fechou suas escolas imediatamente ou introduziu medidas de distanciamento social rigorosas e compulsórias aos seus cidadãos. Diferentemente dos outros países, onde os líderes políticos sempre estão à frente das decisões (com apoio de suas equipes técnicas), na Suécia a figura principal é o epidemiologista que está à frente da Autoridade Nacional de Saúde Pública, que é a agência do governo responsável pelas orientações de medidas a nível nacional. As medidas de distanciamento social, além de não serem rígidas (figura 6), são apenas recomendações.

A figura 6 demonstra quão rígidas foram as medidas adotadas, de acordo com o período. Assim, fica evidente que a Suécia foi o último país nórdico a adotar qualquer medida restritiva e, além disso, mesmo com todas as mortes tão superiores aos demais países, continua sendo o país com as medidas mais brandas.

Figura 6 - Índice de Rigidez - Nórdicos



Fonte: Autor (2020)

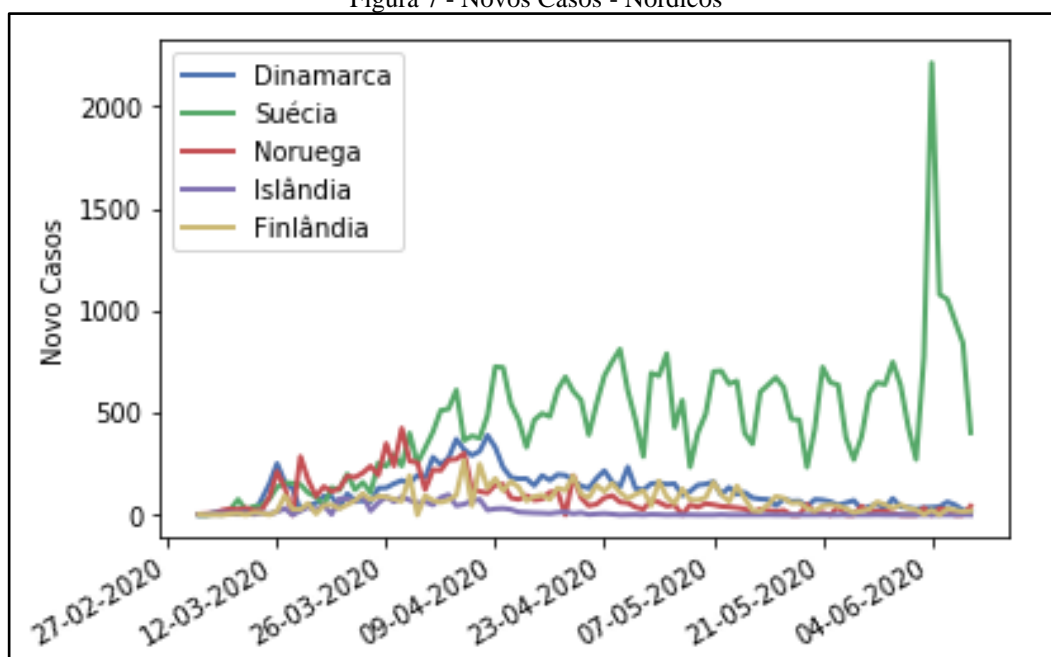
Ao analisar a figura 7, observa-se que a Noruega conseguiu reduzir a quantidade de novos casos já em meados de abril, mesmo tendo adotado medidas mais rígidas de isolamento no começo de março. A Suécia ainda não conseguiu obter sucesso na redução de novos casos.

A primeira medida sueca de isolamento aconteceu apenas em 17 de março, e foi uma recomendação para o encerramento das escolas secundárias e de nível superior. O governo tenta uma abordagem diferente, não adota medidas de confinamento, no intuito de criar imunidade

de grupo (ou efeito rebanho), cujo objetivo principal é que as pessoas sejam infectadas e se tornem imunes ao contágio.

Outra medida que chama a atenção é a não recomendação do uso de máscaras de proteção, com a alegação de que isso poderia representar uma falsa sensação de segurança. Não se pode afirmar se o objetivo de se criar esta imunidade em grupo está sendo alcançado, mas quando se compara a atual situação da Suécia frente aos outros nórdicos, fica visível a discrepância entre os números.

Figura 7 - Novos Casos - Nórdicos



Fonte: Autor (2020)

## 5. CONSIDERAÇÕES FINAIS

A tecnologia está se mostrando fundamental na luta contra a COVID-19 e futuras pandemias. Além de poder apoiar a construção de modelagem e prever o fluxo de uma pandemia; inteligência artificial, *Big Data*, aprendizado de máquina e outras tecnologias podem analisar dados de maneira rápida e eficaz para ajudar os humanos nas linhas de frente a descobrir as estratégias mais eficientes e responder a esta e às futuras pandemias.

Com a rápida expansão e democratização de dados e análises avançadas, o *Data Science* é capaz de trazer as ferramentas, técnicas e procedimentos que foram desenvolvidos na última década para suportar esse desafio sem precedentes.

A principal fonte de dados é proveniente de testes, ou seja, os países que testam pouco possuem menos dados e, conseqüentemente, suas análises poderão ser menos precisas. Levando-se em consideração o panorama mundial, verifica-se que a forma de registro impacta diretamente nas análises. Assim, quando há a comparação de dois ou mais países, deve-se observar se as metodologias adotadas são similares.

Com relação aos países nórdicos, por serem similares em vários aspectos como social, cultural, geográfico e econômicos os resultados das comparações são mais fáceis de serem entendidos pois se espera um padrão entre eles. A partir do momento que a Suécia apresenta

dados discrepantes dos demais países, foi possível identificar o que aconteceu neste país que não aconteceu nos outros Nórdicos. Assim, passou-se a entender que as políticas de restrição podem ter sido fator determinante no número de casos e mortes, tendo em vista que os suecos adotaram medidas mais brandas e mais tardias que os demais.

Sem dados, não é possível entender a pandemia. Somente com uma base dados é possível saber como a doença está evoluindo, qual o impacto da pandemia na vida das pessoas em todo o mundo e se as medidas adotadas pelos países são bem-sucedidas ou não.

A partir das análises efetuadas foi possível entender o quão importante são os dados para que as melhores decisões sejam tomadas. Além disso, tornou-se evidente a relevância dos testes na população, para que haja transparência e veracidade na situação de cada país em meio ao surto do COVID-19.

Pelo fato de estarmos vivendo a pandemia, as análises ainda podem sofrer alterações. Muitos dados ainda serão gerados e muitas pesquisas ainda devem ser realizadas para que se possa certificar a eficiência das medidas adotadas pelas nações, bem como quais as diretrizes devem ser adotadas caso outra pandemia surja no futuro.

Durante o desenvolvimento do trabalho notou-se que outros estudos podem surgir a partir das análises observadas neste. Como citado, alguns países fazem poucos testes, o que não permite saber a real situação deflagrada pelo coronavírus. Assim, poderia ser feito um estudo nesses países para identificar, a partir de aprendizagem de máquina, qual a quantidade de mortes esperada para o período e analisar o montante excedente. Outro estudo pertinente seria utilizar técnicas de aprendizagem de máquina para tentar prever o comportamento nos países nórdicos.

## REFERÊNCIAS

COHEN, Joshua. Underreporting Of COVID-19 Coronavirus Deaths In The U.S. And Europe (Update). **Forbes**, 14 abr. 2020. Healthcare. Disponível em: <https://www.forbes.com/sites/joshuacohen/2020/04/14/underreporting-of-covid-19-deaths-in-the-us-and-europe/#2604f6f082d7>. Acesso em: 25 mai. 2020.

DZUROSKA, Filip. Coronavirus: A Big Data lesson from South Korea. **Towards data science**, 28 abr. 2020. Disponível em: <https://towardsdatascience.com/coronavirus-a-big-data-lesson-from-south-korea-5bb703b8b0ae>. Acesso em: 21 mai. 2020.

GIELOW, Igor. Tecnologia usada no combate à pandemia de coronavírus ameaça privacidade. **Folha de S. Paulo**, 05 abr. 2020. Coronavírus. Disponível em: <https://www1.folha.uol.com.br/mundo/2020/04/tecnologia-usada-no-combate-a-pandemia-de-coronavirus-ameaca-privacidade.shtml>. Acesso em: 25 mai. 2020.

HAN, Jiawei; PEI, Jian; KAMBER, Micheline. **Data mining: concepts and techniques**. Elsevier, 2011.

MCKINNEY, Wes. **Python para análise de dados: Tratamento de dados com Pandas, NumPy e IPython**. Novatec Editora, 2018.

PROVOST, Foster. **Data Science para Negócios**. Rio de Janeiro: Alta Books, 2016. 408 p.

PARK, Sangchul; CHOI, Gina Jeehyun; KO, Haksoo. **Information technology–based tracing strategy in response to COVID-19 in South Korea—privacy controversies.** JAMA, 2020.

PUEYO, Tomas. Coronavirus: The Hammer and the Dance. **Medium**, 19 mar. 2020. Coronavirus. Disponível em: <https://medium.com/@tomaspueyo/coronavirus-the-hammer-and-the-dance-be9337092b56>. Acesso em: 22 mai. 2020.

ROSER, Max; RITCHIE, Hannah; ORTIZ-OSPINA, Esteban; HASELL, Joe. Coronavirus Pandemic (COVID-19). **Our World In Data**, 2020. Disponível em: <https://ourworldindata.org/coronavirus>. Acesso em: 09 jun. 2020.

SHAW, Rajib; KIM, Yong-kyun; HUA, Jinling. Governance, technology and citizen behavior in pandemic: Lessons from COVID-19 in East Asia. **Progress in disaster science**, p. 100090, 2020.

SONN, Jung Won. South Korea's success in controlling coronavirus is due to acceptance of surveillance. **MedicalXpress**, 2020. Disponível em: <https://medicalxpress.com/news/2020-03-south-korea-success-coronavirus-due.html>. Acesso em: 22 mai. 2020.

TRACKING covid-19 excess deaths across countries, **The Economist**, 15 abr. 2020. Disponível em: <https://www.economist.com/graphic-detail/2020/04/16/tracking-covid-19-excess-deaths-across-countries>. Acesso em: 25 mai. 2020.

VANDERPLAS, Jake. **Python data science handbook: Essential tools for working with data.** O'Reilly Media, Inc., 2016.

WALLIMAN, Nicolas. **Métodos de pesquisa.** São Paulo: Saraiva, 2015.

ZHU, Na; ZHANG, Dingyu; WANG, Wenling; LI, Xingwang; YANG, Bo; SONG, Jingdong, et al. A novel coronavirus from patients with pneumonia in China, 2019. **New England Journal of Medicine**. 20 fev. 2020. Disponível em: <http://doi.org/10.1056/NEJMoa2001017>. Acesso em: 22 mai. 2020.