

UNIBRA – CENTRO UNIVERSITARIO BRASILEIRO
CURSO DE GRADUAÇÃO TECNOLÓGICO EM REDES DE
COMPUTADORES

JOÃO VITOR GUEDES MOREIRA

**A CIÊNCIA DE DADOS NO INÍCIO AO COMBATE DO
COVID-19**

RECIFE/2022

JOÃO VITOR GUEDES MOREIRA

A CIÊNCIA DE DADOS NO INÍCIO AO COMBATE DO COVID-19

Trabalho de Conclusão de Curso apresentado ao Centro Universitário Brasileiro – UNIBRA, como requisito parcial para obtenção do título de tecnólogo em Redes de Computadores.

Professora Orientadora: Msc. Ameliara Freire dos Santos de Miranda.

RECIFE/2022

Ficha catalográfica elaborada pela
bibliotecária: Dayane Apolinário, CRB4- 2338/ O.

M835c Moreira, João Vitor Guedes.
A ciência de dados no início ao combate do covid-19 / João Vitor Guedes
Moreira. - Recife: O Autor, 2022.
10 p.

Orientador(a): MSc. Ameliara Freire dos Santos de Miranda.

Trabalho de Conclusão de Curso (Graduação) - Centro Universitário
Brasileiro – UNIBRA. Tecnólogo em Redes de Computadores, 2022.

Inclui Referências.

1. Covid-19. 2. Ciência de dados. 3. Dados. I. Centro Universitário
Brasileiro. - UNIBRA. II. Título.

CDU: 004

Dedico este trabalho aos meus pais que sempre estiveram ao meu lado, me incentivando e orientando para que eu alcance os meus objetivos.

AGRADECIMENTOS

É com extrema felicidade que concluo esta etapa da minha vida, durante a qual me esforcei muito e que, hoje, vejo os frutos de todo meu esforço se concretizar neste trabalho. Deixo nítido aqui o quão difícil foi esta jornada na reta final do curso, quantos contratempos e obstáculos que se impuseram em meus caminhos, mas que hoje posso olhar para eles e ser grato. Muitos passaram, apenas passaram, outros nos acolheram e insistiram junto, na esperança de que, apesar dos pesares, daria certo. E deu! Àqueles que infelizmente não forem citados, desde já peço perdão. Entretanto, esta vitória é de todos nós que acreditamos.

Agradecer, primeiramente, a Deus e a professora e orientadora, Ameliara Freire Santos de Miranda, sem a qual este trabalho jamais haveria saído. Por todo o seu comprometimento em sonhar junto e por sua paciência de transformar uma simples pedra brilhante em diamante. Obrigado por me manter tranquilo quando nem eu mesmo confiava em minha capacidade e que, graças ao seu auxílio, concluo, de forma honrosa e primorosa, este trabalho.

Agradecer, também, aos professores, que tanto contribuíram para a minha formação, a todos, muito obrigado.

Aos meus familiares, por compreenderem os meus momentos e serem ponto de apoio durante todos esses dois anos e meio de curso. Sem eles, esse sonho não seria possível, obrigado por sonharem comigo desde sempre. Em especial, à minha mãe Fernanda Maria Guedes e Silva, e meu pai Odir Dias Moreira, por todo amor e carinho, essa vitória é nossa! Sem vocês não eu teria chegado tão longe. Em especial a minha namorada Ieda Correia Zarzar, e meus amigos Daniel Montalvão Rodrigues e Elbert Jorge por estarem sempre presentes em todos os momentos difíceis, sendo vocês, as peças-chave para manter os orientandos seguros e confiantes.

Aos meus amigos Allan Cleysson e Bruno Felipe que estiveram comigo durante esses dois anos e meio de curso, sempre juntos, apesar dos altos e baixos. Com eles, fui capaz de me reinventar e ser minha melhor versão.

Aos meus demais amigos Ruan Galvão e Caio Jucá que estiveram comigo durante os momentos difíceis da minha vida, por entenderem cada um deles.

Por fim, agradecer a mim, pela força de vontade e esforço que tive ao longo

do curso, para concluir está importante etapa da minha vida.

"If you can dream it, you can do it."

"Se você pode sonhar, você pode fazer"

(Walt Disney)

SUMÁRIO

<u>RESUMO</u>	8
<u>1 INTRODUÇÃO</u>	9
1.1 CIÊNCIA DE DADOS X ÁREAS RELACIONADAS.....	11
<u>2 A CIÊNCIA DE DADOS NA PANDEMIA</u>11
2.1 A CIÊNCIA DE DADOS NA PRÁTICA.....	14
<u>3 A IMPORTÂNCIA DA CIÊNCIA DE DADOS NA PANDEMIA</u>	14
<u>4 A IMPORTÂNCIA DOS RESULTADOS</u>	15
<u>5 CONSIDERAÇÕES FINAIS</u>	16
<u>6 REFERÊNCIAS.....</u>	17

A CIÊNCIA DE DADOS NO INÍCIO AO COMBATE DO COVID-19

João Vitor Guedes Moreira

Prof^a. Msc. Ameliara Freire Santos de Miranda

Resumo: Ao encontrar respostas ao Covid-19 para autoridades científicas e de saúde, a importância e o valor dos dados estão aumentando. Como é difícil diagnosticar infecções na população em geral, governos ou empresas privadas desenvolveram iniciativas apoiadas pela Ciência de Dados como forma de rastrear sintomas, pontos de contato e o deslocamento da doença para apoiar estratégias de vigilância e avaliação na vigilância do novo coronavírus. Embora essas medidas sejam importantes e necessárias, a quantidade e o tipo de dados pessoais coletados, processados, compartilhados e usados em nome da saúde pública, bem como o acompanhamento ou uso subsequente desses dados, têm desencadeado aspectos éticos, legais e técnicos questões.

Palavras-chave: Covid-19. Ciência de Dados. Dados.

1. INTRODUÇÃO

Em dezembro de 2019, a Organização Mundial de Saúde (OMS) foi alertada sobre uma possível pandemia que se dava início na China. Esse vírus tratava-se de uma nova cepa não identificada em seres humanos.

Com o nome científico de SARS-CoV2, o Covid-19 foi declarado como surto epidêmico no final de janeiro de 2020 e constituído como caso de Emergência de Saúde Pública de Importância Internacional (ESPII) – o mais alto nível de alerta considerado pela OMS.

De acordo com o Regulamento Sanitário Internacional (RSI), a ESPII é considerada “um evento extraordinário passível de construir risco a saúde pública para outros países devido a sua disseminação internacional de doenças; e que potencialmente precisa de uma resposta internacional coordenada e imediata”.

Na história mundial, essa é a sexta vez que ocorre uma ESPII, sendo as demais:

- Pandemia de H1N1 (2009);
- Disseminação Internacional de Poliovírus (2014);
- Surto de Ebola na África Ocidental (2014);
- Vírus da Zika e aumento no número de casos de microcefalia e outras malformações congênitas (2016);
- Surto de Ebola no Congo (2018).

Em março de 2020, a Covid-19 foi considerada uma pandemia pela OMS, termo utilizado para se referir à distribuição geográfica da doença e não à sua gravidade. Sendo assim, o vírus foi reconhecido como um surto que atingia vários países e regiões do mundo.

No Brasil, o primeiro caso confirmado ocorreu em fevereiro de 2020 no estado de São Paulo. Desde então, já foram registrados mais de 28 milhões e casos no país. Ultrapassando regiões como Estados Unidos e Reino Unido, o Brasil atingiu uma média maior que 7 mil mortes em apenas uma semana, sendo o país líder na média diária de óbitos.

Dentro do contexto da pandemia, diversos conflitos de gestão sanitária surgiram no país, em razão das linhas de medidas de segurança que seriam abordadas pelos níveis regionais e nacional. Em termos de gestão de saúde, estados e municípios passaram a definir e implementar suas próprias medidas de vigilância e controle, mesmo que essas fossem em contraposição as medidas recomendadas pelo governo federal.

Em relação aos estados, o Nordeste era a região que concentrava maior número de casos do país, 35,2% de casos do Covid-19, sendo o maior percentual entre as regiões.

Em Pernambuco, as repostas à pandemia vieram rapidamente após confirmado os primeiros casos no estado. Em março, o governo do estado publica o seu primeiro Decreto Lei proibindo eventos com mais de 50 pessoas, seguido por vários outros decretos publicados nos meses subsequentes com medidas de distanciamento social, isolamento e restrições de contato.

E foi a partir dessa perspectiva de crise sanitária que o mundo passou a falar sobre Ciência de Dados e a sua importância. Mas afinal, do que se trata?

A ciência de dados combina vários campos, incluindo estatísticas, métodos científicos, inteligência artificial (IA) e análise de dados para extrair valor dos dados. Aqueles que praticam ciência de dados são chamados de cientistas de dados e combinam uma variedade de habilidades para analisar dados coletados da Web, de smartphones, clientes, sensores e outras fontes para obter insights acionáveis.

A ciência de dados abrange a preparação de dados para análise, incluindo limpeza, agregação e manipulação dos dados para realizar análises avançadas de dados. As aplicações analíticas e os cientistas de dados podem revisar os resultados para descobrir padrões e permitir que os líderes de negócios obtenham insights informados.

Por esse motivo, a Ciência de Dados tem sido tão explorada atualmente. Graças a sua rica abrangência de coleta de dados e armazenamento através da tecnologia, diversos setores passaram a usá-la como mecanismo benfeitor a sociedade.

1.1. CIÊNCIA DE DADOS X ÁREAS RELACIONADAS

Para que este artigo fique ainda mais completo e entendível sobre a Ciência de Dados e como ela é utilizada, será elencado abaixo a diferença entre o tema e outros dois termos que também estão relacionados à área, como a Inteligência Artificial (IA), *Machine Learning* e *Deep Learning* que são frequentemente confundidos.

Aqui está uma análise simples sobre a função de cada um:

- Inteligência Artificial (IA): trata-se de um comportamento computadorizado semelhante ao comportamento humano.
- Ciência de Dados: subconjunto da IA e se refere mais às áreas sobrepostas de estatísticas, métodos científicos e análise de dados, todas as quais são usadas para extrair significados e *insights* de dados.
- *Machine Learning*: outro subconjunto da IA e consiste nas técnicas que permitem aos computadores descobrir coisas a partir dos dados e fornecer aplicações de IA.
- *Deep Learning*: subconjunto da *machine learning* que permite aos computadores resolverem problemas mais complexos.

2. A CIÊNCIA DE DADOS NA PANDEMIA

O trabalho feito pela Ciência de Dados durante o período da pandemia está na investigação sobre informações pertinentes a Sars-COV2 que possam ser assimiladas a questões já abordadas em sociedade e na correlação de dados, como o número de casos e óbitos, há elementos implícitos durante este período. Desta forma, foi possível notar que, independentemente das suas circunstâncias, fatores socioeconômicos agravavam ainda mais para a disseminação da doença mundo à fora.

De acordo com Patel (2020), pessoas com status socioeconômico menos favorecidas, eram mais facilmente acometidas pelo vírus. Visto que muitas delas não poderiam deixar os seus empregos ou não tinham a oportunidade de migrar para o

home office. Além disso, suas condições de moradia também facilitavam a propagação e aquisição da doença.

Desta forma, por meio da Ciência de Dados, modelos de predição de dados sobre a Covid-19 foram desenvolvidos como forma de demonstrar o impacto da quarentena, e outras medidas, na redução do número de casos e, conseqüentemente, de óbitos.

Também através da Ciência de Dados foi possível apresentar a criação de repositórios de pesquisas e data sets que eram atualizados diariamente para a criação, desenvolvimento e implementação de estratégias ao combate da Sars-COV2. Repositório este que serviu de plataforma de conhecimento para toda a população civil sobre o novo coronavírus.

Portanto, levando em consideração todas as informações citadas acima, a união do amplo conhecimento sobre a Covid-19, juntamente com as estatísticas disponibilizadas pelo IBGE, é possível validar a seguinte informação: “o novo coronavírus não discrimina cor, gênero, classe social ou qualquer status estabelecido socialmente”.

GRÁFICO 1¹ – Número de casos novos de COVID-19 notificados no Brasil, de 26 de fevereiro a 6 de abril de 2020.

¹ Fonte: Ministério da Saúde (BR). Secretaria de Vigilância em Saúde. Bol Epidemiol. 2020

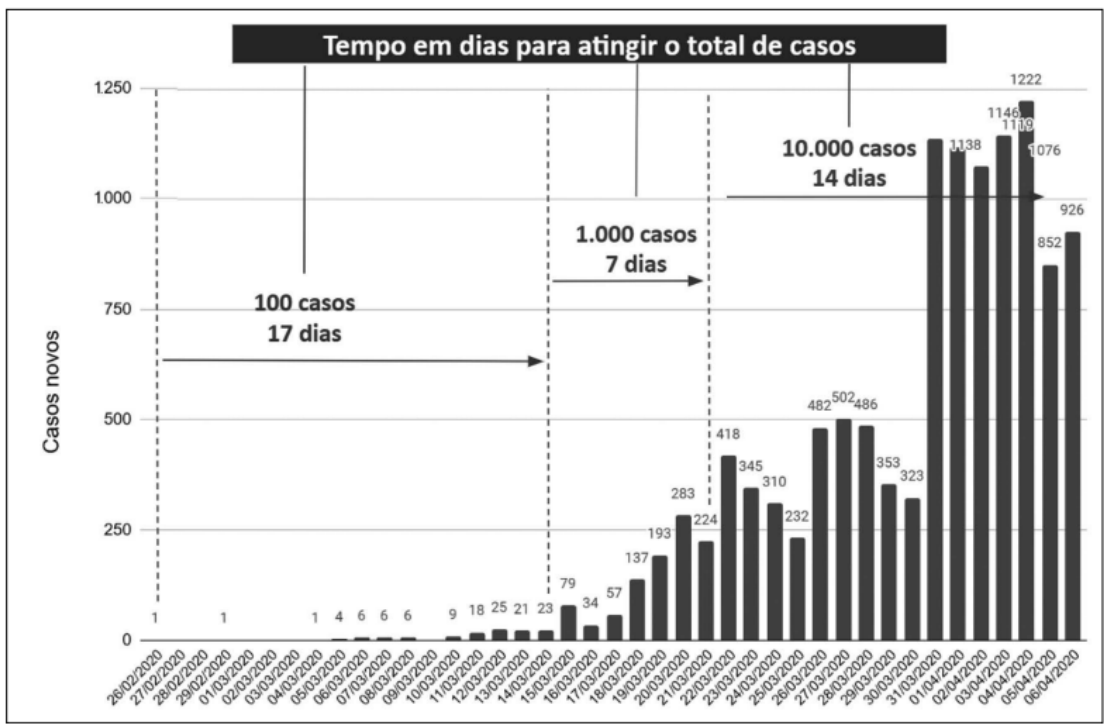
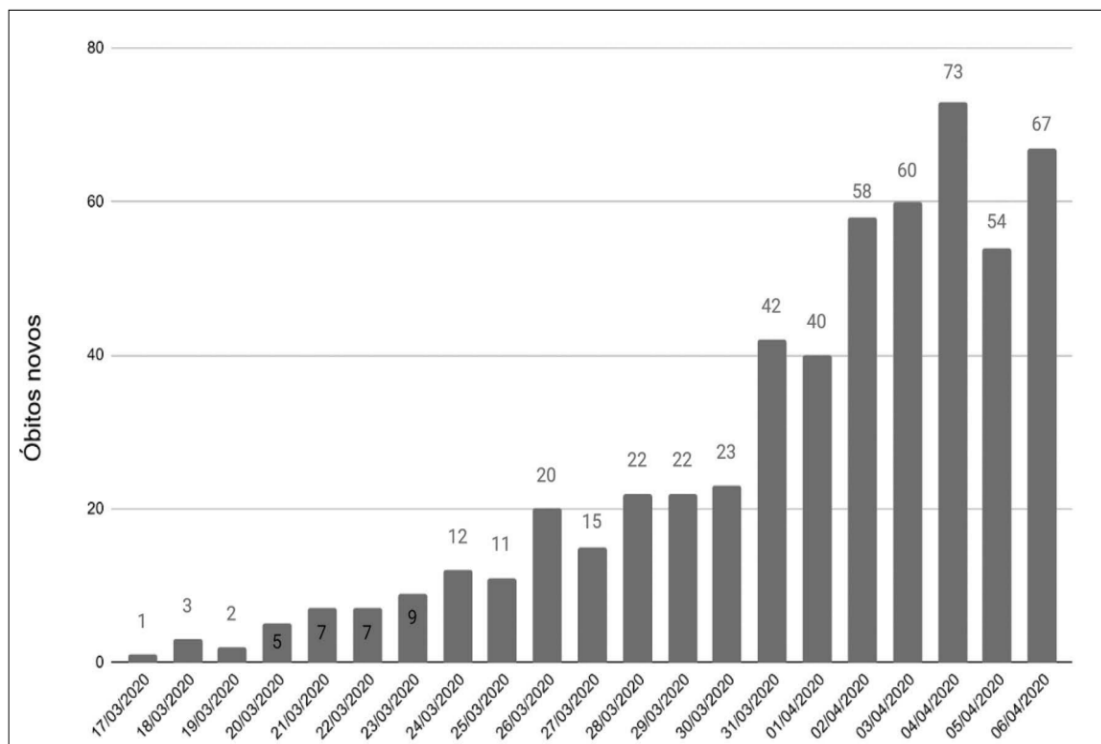


Imagem 2¹ – Número de óbitos por COVID-19 no Brasil, de 17 de março a 6 de abril de 2020.



2.1. A CIÊNCIA DE DADOS NA PRÁTICA

Com o a explosão de dados sobre o avanço da pandemia e no processamento de informações a respeito da doença, foi possível utilizar a Ciência de Dados como base para auxiliar diversas tomadas de decisões. No Brasil, por exemplo, diversos órgãos publicaram informações e dados pertinentes aos dados da Covid-19, dentre eles o Ministério da Saúde, órgão de poder federal, e também possibilitou o estudo de casos responsáveis pelo impacto socioeconômico através do Instituto Brasileiro de Geografia e Estatística (IBGE), órgão também federal.

Como pode ser visto no site do [governo federal](#)², o portal do Sistema Único de Saúde (SUS) brasileiro foi tomado por uma gigantesca data base responsável por informar a sociedade acerca dos novos casos, suspeitas e óbitos.

Além disso, de acordo com o portal núcleo do conhecimento, em seu artigo “Ciência de Dados: Ferramentas para o combate à Covid-19”, as principais ferramentas para análise partiram da obtenção de dados sobre a doença, bibliotecas de pacotes a partir de softwares desenvolvidos como:

- Pandas;
- Numpy;
- Matplotlib;
- Folium;
- Prophet;
- Pwa;

E a partir deles colher dados matemáticos que, por se tratar de uma ciência exata, permite a interdisciplinaridade como método de solucionar problemas e encontrar soluções lógicas.

3. A IMPORTÂNCIA DA CIÊNCIA DE DADOS NA PANDEMIA

A fim de realizar a análise de dados sem precedentes, o mercado de Ciência de Dados durante a pandemia foi a maior fornecedora de tecnologia de predição de

² <https://www.gov.br/pt-br/apps/coronavirus-sus>

dados ao governo como forma de estimar a quantidade de leitos de hospitais e UTIs que seriam necessários.

Além disso, a Ciência de Dados foi responsável pela criação de um painel centralizador de informações bem estruturadas, organizados em planilha, e não estruturadas, informações na rede e na internet.

“A ciência de dados foi usada para identificar rapidamente as informações mais importantes e deixar as pessoas mais livres de trabalhos que podem ser automatizados” (Guedes, Rychard. Let’s Code. 2020).

Neste sentido, a análise e construção de gráficos trabalhados durante a pandemia foram utilizados como parte do mecanismo de entendimento e desenvolvimento de habilidades de leitura e interpretação de dados. Vale ressaltar que, durante o período de pandemia, a Ciência de Dados se fortaleceu e se mostrou ferramenta de grande apoio às gestões de crise sanitária, trazendo para a sociedade o *dashboard* mais acessado na história da humanidade, o mapa de monitoramento em tempo real do número de casos e óbitos pelo coronavírus.

Em outras palavras, as ciências, em todas as suas áreas, foram responsáveis por direcionar e realizar pesquisas de levantamento de números e informações que salvaram vidas e que serviram na gestão de crise em diversas regiões do mundo.

De acordo com André Filipe Batista, professor e coordenador do Centro de Data Science do Instituto de Ensino e Pesquisa (Insper): “a crise sanitária foi a demonstração de um evento complexo, no qual não tínhamos precedentes para explicar o dinamismo do vírus e seu comportamento. Então, conseguimos fazer modelos que nos dessem com exatidão o número de casos que tínhamos. Mas foi por meio de dados de ocorrência da doença que a gente pôde entender a dimensão da pandemia”.

4. A IMPORTÂNCIA DOS RESULTADOS

Como visto anteriormente neste artigo, a captação e coleta de dados através de estudos baseados em informações pertinentes fomentam o que chamamos de *Big Data*.

Através desse amplo conceito, pode-se dizer que o conjunto de dados relevantes, a partir de determinado volume, passa a ser tão grande que não é possível ser gerenciado através de uma base de dados relacional. Desta forma, é necessário que especialistas analisem e discorram acerca dos dados e relativizem quais são realmente necessários para a base de estudos.

Durante a pandemia, alguns conceitos da Ciência de Dados foram utilizados como base de entendimento e captação de resultados, são eles:

- Volume: quantidade inimaginável de dados gerados por segundo e também dados já existentes que precisam ser analisados e cruzados;
- Velocidade: como esses dados passaram a ter valor dentro de um pequeno espaço de tempo através da percepção de fraudes e dados irreais;
- Variedade: busca por fontes variadas para obtenção de novos dados através do cruzamento deles;
- Valor: extração de valor, sendo este o mais importante, visto que um dado só irá as vias finais se for realmente pertinente.

Desta forma, através destes conceitos e de uma rede favorável, tivemos avanços e captações reais disponíveis para as mais diversas áreas da sociedade. É importante ressaltar como os resultados da Ciência de Dados é essencial para a sociedade, visto que através dela temos amplo conhecimento de forma democrática e entendível.

5. CONSIDERAÇÕES FINAL

Diante de todo o trabalho produzido e por meio de pesquisas, foi possível entender ainda mais sobre a relevância do estudo acerca da Ciência de Dados e como ela foi importante para todos nós durante a pandemia iniciada em 2020.

Mesmo com o vasto arsenal de conhecimento que possuímos hoje em dia, graças ao advento da internet e da democratização do saber através dela, podemos notar o quão esta ciência se enraizou e fez parte do cotidiano de muitos sem ao

menos sequer ser reconhecida. Em contrapartida, foi possível perceber que falar sobre Ciência de Dados já não é mais um mistério e que, por mais que não seja de amplo conhecimento de que ela faz parte daquele processo, muitos já reconhecem o seu valor dentro da esfera socioeconômica.

É por isso que, diante deste estudo, é perceptível quais novos direcionamentos a sociedade têm tomado ao longo dos tempos e que os seus níveis de intelectualidade estão cada vez mais dissolvidos não só nas grandes áreas, mas também nas áreas emergentes que buscam a sua estabilidade.

Desta forma, é notório que mesmo com a percepção ou não de que tais resultados são feitos através da Ciência de Dados, o seu estudo vai além do entendimento comum e que o assunto está ainda mais interligada as questões socioeconômicas e as diversas áreas do saber. No Brasil, esta Ciência ainda está em expansão e ainda há muito a acrescentar durante a trajetória político-social do país.

6. REFERÊNCIAS

AFINAL, o que é ciência de dados? 17 nov. 2021. Disponível em: <https://www.tecmundo.com.br/mercado/228839-ciencia-dados.htm>. Acesso em: 27 jun. 2022.

CIÊNCIA de dados na educação: a importância durante a pandemia. 29 mar. 2020. Disponível em: <https://fundacaotelefonicao.org.br/noticias/ciencia-de-dados-importancia-educacao-pandemia/>. Acesso em: 11 maio 2022.

CORONAVÍRUS Brasil. Atualizado em 25 maio 2022. Disponível em: <https://covid.saude.gov.br/>. Acesso em: 25 maio 2022.

CIÊNCIA de dados na educação: a importância durante a pandemia. 29 mar. 2020. Disponível em: <https://fundacaotelefonicao.org.br/noticias/ciencia-de-dados-importancia-educacao-pandemia/>. Acesso em: 11 maio 2022.

CIÊNCIA de dados aplicada à pandemia do coronavírus no Brasil, uma análise socioeconômica. 21 jul 2021. Disponível em: <https://repositorio.unesp.br/handle/11449/213821>. Acesso em 26 jun. 2022.

CIÊNCIA de dados: ferramentas para o combate à covid-19. 22 jun. 2020. De Brito, Arlindo Matheus Santiago. Disponível em: <https://www.nucleodoconhecimento.com.br/tecnologia/ciencia-de-dados>. Acesso em: 26 jun. 2022.

DO NASCIMENTO, M. G., Iorio, G., Thome, T. G., Medeiros, A. A., Mendonça, F. M., Campos, F. A., David, J. M., Stroele, V., and Dantas, M. A. (2020). Covid-19: A digital transformation approach to a public primary healthcare environment. Technical report, EasyChair.

GOVERNO do Estado de Pernambuco. Decreto nº 48.809, de 14 de março de 2020. Regulamenta, no Estado de Pernambuco, medidas temporárias para enfrentamento da emergência de saúde pública de importância internacional decorrente do coronavírus, conforme previsto na Lei Federal nº 13.979, de 6 de fevereiro de 2020. Diário Oficial do Estado de Pernambuco 2020; 14 mar.

HASSANIN, A. (2020). Coronavirus origins: genome analysis suggests two viruses may have combined. In World Economic Forum

MAPA Completo das Funções em Ciência de Dados - Data Science Academy. 29 abr. 2022. Disponível em: https://blog.dsacademy.com.br/mapa-completo-das-funcoes-em-ciencia-de_dados/. Acesso em: 12 maio 2022.

MINISTÉRIO da Saúde. Portaria nº 188, de 3 de fevereiro de 2020. Declara Emergência em Saúde Pública de importância Nacional (ESPIN) em decorrência da Infecção Humana pelo novo Coronavírus (2019-nCoV). Diário Oficial da União 2020; 4 fev.

MINISTÉRIO da Saúde. Plano de Contingência Nacional para Infecção Humana pelo novo Coronavírus COVID-19. <https://portalarquivos2.saude.gov.br/images/pdf/2020/fevereiro/13/plano-contingencia-coronavirus-COVID19.pdf>. Acesso em: 11 maio 2022.

O que é ciência de dados? Disponível em: <https://www.oracle.com/br/data-science/what-is-data-science/>. Acesso em: 25 jun. 2022.

PIRES RR. Os efeitos sobre grupos sociais e territórios vulnerabilizados das medidas de enfrentamento à crise sanitária da COVID-19: propostas para o aperfeiçoamento da ação pública. (Nota técnica 33). <http://repositorio.ipea.gov.br/handle/11058/9839>. Acesso em: 12 maio 2022.

PATEL, J., Nielsen, F., Badiani, A., Assi, S., Unadkat, V., Patel, B., Ravindrane, R., and Wardle, H. (2020). Poverty, inequality and covid-19: the forgotten vulnerable. *Public health*, 183:110.

SOHRABI C, Alsafi Z, O'Neill N, Khan M, Kerwan A, Al-Jabir A, et al. World Health Organization declares global emergency: a review of the 2019 novel coronavirus (COVID-19). *Int J Surg* 2020; 76:71-6

SISTEMA de computação petaflopica do sinapad, 2020. <https://sdumont.incc.br/>. Acesso em: 12 maio 2022.