

UNIVERSIDADE EVANGÉLICA DE GOIÁS - UNIEVANGÉLICA
ENGENHARIA DE SOFTWARE

GULHERME SANTOS MENEZES
ITALO LUAN VALDIVINO VIEIRA

Desenvolvimento de agentes inteligentes para classificação de dados de Covid-19 para o Estado de Goiás.

Anápolis - GO

Dezembro, 2021

UNIVERSIDADE EVANGÉLICA DE GOIÁS - UNIEVANGÉLICA
ENGENHARIA DE SOFTWARE

GUILHERME SANTOS MENEZES
ITALO LUAN VALDIVINO VIEIRA

Desenvolvimento de agentes inteligentes para classificação de dados de
Covid-19 para o estado de Goiás.

Trabalho apresentado ao Curso de Engenharia de Software
da Universidade Evangélica de Goiás – UniEVANGÉLICA,
da cidade de Anápolis-GO como requisito parcial para
obtenção do Grau de Bacharel em Engenharia de Software.

Orientador (a): Prof. M.e William P. Santos Júnior.

Anápolis - Go
Dezembro, 2021

UNIVERSIDADE EVANGÉLICA DE GOIÁS - UNIEVANGÉLICA
ENGENHARIA DE SOFTWARE

GUILHERME SANTOS MENEZES
ITALO LUAN VALDIVINO VIEIRA

Desenvolvimento de agentes inteligentes para classificação de dados de
Covid-19 para o estado de Goiás.

Monografia apresentada para Trabalho de Conclusão de Curso de Engenharia de Software da Universidade Evangélica de Goiás - UniEVANGÉLICA, da cidade de Anápolis-GO como requisito parcial para obtenção do grau de Engenheiro(a) de Software.

Aprovado por:

Prof. M.e William P. Santos Júnior
(**ORIENTADOR**)

Henrique Valle de Lima
(**AVALIADOR**)

Anápolis, 10/12/2021

FICHA CATALOGRÁFICA

MENEZES, Guilherme Santos. **Desenvolvimento de agentes inteligentes para classificação de dados de Covid-19 para o estado de Goiás.** Anápolis 2021 (Universidade Evangélica de Goiás – UniEVANGÉLICA, Engenheiro(a) de software, 2021).

VIEIRA, Italo Luan Valdivino. **Desenvolvimento de agentes inteligentes para classificação de dados de Covid-19 para o estado de Goiás.** Anápolis 2021 (Universidade Evangélica de Goiás – UniEVANGÉLICA, Engenheiro(a) de software, 2021).

1. Covid-19. 2. Inteligência Artificial. 3. Pandas. 4. Python.

REFERÊNCIA BIBLIOGRÁFICA

MENEZES, Guilherme Santos, VIEIRA, Italo Luan Valdivino. **Desenvolvimento de agentes inteligentes para classificação de dados de Covid-19 para o estado de Goiás.** Anápolis, 2021. 37 p. Monografia - Curso de Engenharia de Software Universidade Evangélica de Goiás - UniEVANGÉLICA.

CESSÃO DE DIREITOS

NOMES DOS AUTORES: Guilherme Menezes, Italo Vieira

TÍTULO DO TRABALHO: Desenvolvimento de agentes inteligentes para classificação de dados de Covid-19 para o estado de Goiás

GRAU/ANO: Graduação / 2021

É concedida à Universidade Evangélica de Goiás - UniEVANGÉLICA, permissão para reproduzir cópias deste trabalho, emprestar ou vender tais cópias para propósitos acadêmicos e científicos. O autor reserva outros direitos de publicação e nenhuma parte deste trabalho pode ser reproduzida sem a autorização por escrito do autor.

Guilherme Santos Menezes, Italo Luan Valdivino Vieira
Anápolis, 10/12/2021

RESUMO

Os dados sobre o covid-19 são disponibilizados com intuito de passar as informações para todos, proporcionando uma massa de informações constante durante todos os dias, com isso surge a necessidade de filtrar esses dados de maneira fácil e rápida, tornando o objetivo desse estudo, a apresentação dos resultados de casos confirmados e de mortalidade do covid-19. Foram coletados dados através do site brasil.io de 247 cidades do estado de Goiás, utilizando a linguagem Python e as bibliotecas Pandas, Matplotlib e Folium. Obtivemos uma comparação dos casos confirmados entre cada região do estado, dentre elas foram destacadas 8 cidades que apresentaram o maior nível de contaminação e mortalidade durante o ano de 2021.

Palavras-chave: Covid-19. Inteligência Artificial. Pandas. Python.

ABSTRACT

Data on covid-19 are made available in order to pass as information to everyone, providing a constant mass of information every day, with this the need to filter this data easily and quickly, making the objective of this study, the presentation of the results of confirmed cases and mortality of covid-19. Data were collected through the brasil.io website from 247 cities in the state of Goiás, using a Python language and the Pandas, Matplotlib and Folium libraries. We obtained a comparison of confirmed cases between each region of the state, among them, 8 cities were highlighted that had the highest level of contamination and the highest level of mortality during the year 2021.

Keywords: Covid-19. Artificial intelligence. Pandas. Python.

LISTA DE ILUSTRAÇÕES

Figura 1 - Criação de DataFrame	19
Figura 2 - Criação de Series	20
Figura 3 - Pie Chart (Gráfico de pizza)	21
Figura 4 - Stamen Terrain	22
Figura 5 - Stamen Toner	22
Figura 6 - Base de dados completa	25
Figura 7 - Importação das bibliotecas	26
Figura 8 - Importando a base de dados em uma variável	26
Figura 9 - Filtragem da base de dados, exibindo apenas as colunas definidas	26
Figura 10 - Base de dados filtrada	27
Figura 11 - Top 8 cidades com maior índice de casos confirmados	28
Figura 12 - Top 8 cidades com maior índice de óbitos	29
Figura 13 - Código da plotagem do gráfico Pie Chart, Top 8 Confirmações	30
Figura 14 - Código da plotagem do gráfico Pie Chart, Top 8 Óbitos	30
Figura 15 - Código da plotagem do gráfico, passando as coordenadas	30
Figura 16 - Código feito para visualização dos popup	31
Figura 17 - Código para a marcação das regiões no mapa	32
Figura 18 - Gráfico Top 8 casos confirmados	33
Figura 19 - Gráfico Top 8 óbitos confirmados	34
Figura 20 - Mapa de Goiás com marcações das 8 cidades com maiores índices de casos confirmados	35
Figura 21 - Mapa de Goiás com marcações das 8 cidades com maiores índices de óbitos confirmados	35

LISTA DE ABREVIATURAS E SIGLAS

Siglas	Descrição
IA	Inteligência Artificial
Machine Learning	Aprendizado de Máquina
IBGE	Instituto Brasileiro de Geografia e Estatística

SUMÁRIO

1. INTRODUÇÃO	14
2. FUNDAMENTAÇÃO TEÓRICA	16
2.1. Inteligência Artificial.....	16
2.2. Machine Learning.....	17
2.3. Python.....	17
2.4 Fundamentos de Python para análise dos dados	18
2.5 Pandas.....	18
2.6 Dataframes	19
2.7 Séries	19
2.8 Matplotlib	20
2.9 Folium.....	21
2.10 Mineração de dados.....	23
3. METODOLOGIA	24
4. DESENVOLVIMENTO	25
4.1 Base de Dados.....	25
4.2 Plotagem do Mapa	30
5. RESULTADOS ALCANÇADOS.....	33
6. CONCLUSÃO E CONSIDERAÇÕES FINAIS	36
7. REFERÊNCIAS BIBLIOGRÁFICAS	37

1. INTRODUÇÃO

O mundo atualmente vem enfrentando uma grande pandemia resultante de um vírus que teve sua origem na cidade de Wuhan, na China, infectando inicialmente 153.648 habitantes, sendo 5.746 que resultaram em mortes e 146 países/territórios afetados até o dia 15 de março de 2020. (FREITAS et al., 2020) O vírus possui 80% de sua estrutura genética e idêntica ao SARS-CoV, que foi responsável pela pandemia de SARS em 2002. A sua forma de contaminação é por gotículas proveniente de espirros e tosses, porém já foi encontrado até mesmo em urina e fezes de pacientes contaminados. (GRONER et al., 2021) Os sintomas apresentados podem ser facilmente confundidos com outras doenças, devida as suas semelhanças, como a febre (87,9%), tosse seca (67,7%), fadiga (38,1%), produção de escarro (33,4%), dor de garganta (13,9%), calafrios (11,4), náuseas ou vômitos (5%), congestão nasal (4,8%), diarreia (3,7%) e congestão conjuntival (0,8%). (LIMA, 2020)

Os primeiros casos confirmados no Brasil foram no mês de fevereiro de 2021, neste mesmo mês o país declarou emergência de Saúde Pública de Importância Nacional (ESPIN), logo após foram iniciadas diversas ações com intuito de paralisar o avanço do vírus. A corrida pelas informações dos casos e óbitos confirmados são realizados pelo Ministério da Saúde brasileiro e entregues pelas Secretarias Estaduais de Saúde, ajudando a ter o conhecimento da doença e implementando ações para desacelerar os números de casos. (CAVALCANTE et al., 2020)

Assim, como em diversas áreas, faz-se necessário o aperfeiçoamento de tecnologias capazes de colaborar na luta contra o novo Covid-19, como o uso da inteligência artificial para analisar os dados de contágio da referida doença. De que forma essas análises podem contribuir para as classificações dos dados de indivíduos contaminados?

A pandemia estabeleceu uma nova realidade nos modelos de serviços para atuar frente a essa realidade de distanciamento social, tornando assim necessário a evolução tecnológica e um possível marco na revolução tecnológica do setor. E consiste a relevância em trabalhar esta temática, pois há a necessidade de disponibilizar dados concretos para toda a população e ferramentas para auxiliar o combate do Covid-19. (CELUPPI et al., 2021)

Devido à Inteligência Artificial, os estudos e tratamentos do Covid-19 evolui grandiosamente, minimizando as perdas e o cansaço físico dos profissionais da linha de

frente, sendo um aliado para os médicos e das políticas de saúde que lutam diariamente contra o vírus. (CELUPPI et al., 2021)

Um dos principais objetivos no auxílio do combate do vírus é a transmissão das informações de maneira clara e rápida. (NEVES, 2020)

Nos últimos anos a quantidade de dados aumentou muito além das expectativas, a quantidade alta de informações que são transmitidas, compartilhadas e encaminhadas todos os dias, criou a necessidade de filtragens maiores, para facilitar o entendimento das informações. (PORTO, 2014)

Assim o trabalho tem como objetivo geral a apresentação dos dados estatísticos já existentes sobre o Covid-19 disponibilizados publicamente no site Brasil.IO. Para alcançar o objetivo que foi proposto a este trabalho, precisamos concluir os seguintes objetivos específicos: (a) estudar e aplicar a linguagem de programação Python para realizar a análise dos dados, (b) analisar as bibliotecas Pandas e Matplotlib e aplicar para realizar a coleta dos dados e a visualização dos dados com maior representatividade, (c) realizar a plotagem do mapa utilizando a biblioteca Folium, (d) analisar a importância e a veracidade dos dados entregues pelo site Brasil.IO, (e) ressaltar a importância dos recursos tecnológicos, como a Inteligência Artificial na área da saúde.

Assim o trabalho se justifica pela relevância da ciência de dados no combate do Covid-19, contribuindo com informações sobre o vírus para servir de base e inspiração para pesquisas e experimentos futuros.

2. FUNDAMENTAÇÃO TEÓRICA

2.1. Inteligência Artificial

Existem tarefas que os seres humanos realizam que requer uma certa inteligência do indivíduo e algumas dessas tarefas podem ser feitas também através de operações computacionais e isso é chamado de Inteligência Artificial. Sistemas podem ser treinados de diversas formas e algumas delas o tornaram inteligente como por exemplo a capacidade de adquirir e armazenar conhecimentos, resolver problemas, expressar informações, planejar eventos, comunicar-se e aprender na língua falada. Os primeiros programas de xadrez foram programados para tomar decisões como qual peça mover e pra onde e isso envolve o aprendizado de máquina, já na inteligência artificial essas tomadas de decisões do sistema baseiam-se apenas em um conjunto de regras de codificação. (NETWORKS, 2018)

O desempenho da Inteligência Artificial tem como apoio a junção racional de maneiras gerais e automatizadas para agregar na hora de se pensar na resolução lógica de problemas. A IA se difere como disciplina científica pela sua generalidade, computabilidade, e combinabilidade lógica que está sempre presente nos seus métodos. Existe outras disciplinas científicas que também utilizam de técnicas inteligentes, porém de formas mais específicas, técnicas que funcionam, mas que as vezes não possuem um esclarecimento do raciocínio ou que não são planejadas num todo automatizado e por isso a IA se destaca e está sempre evoluindo de forma rápida. (SILVA, 2019)

Atualmente existem certos problemas que necessitam de mais de 24 horas para serem solucionados e o ser humano não foi feito para ser produtivo por todo esse tempo, é necessário também realizar as necessidades básicas como se alimentar, ir ao banheiro e principalmente ter horas de descanso. Porém por mais que essas pausas sejam necessárias elas podem acabar levando a atrasos e isso pode ser prejudicial de certa forma na entrega de alguma tarefa, no entanto é notório que o trabalho feito por uma Inteligência Artificial tem essa grande vantagem de não precisar ser interrompido em momento algum. (SILVA, 2019)

De acordo com Marques a IA está bastante envolvida nos meios corporativos já que nesses lugares as tarefas costumam ser rotineiras e repetitivas, podendo ocasionar cansaço

e desânimo nos funcionários envolvidos, e com uma máquina este processo pode se tornar menos cansativo. (MARQUES, 2017)

Então podemos dizer que a Inteligência Artificial substitui o ser humano no desenvolvimento de algumas tarefas e isso pode influenciar na perda de empregos de algumas pessoas, conforme é dito no livro Inteligência Artificial de Russel e Norvig (2013p.1188):

[...] Alguém poderia argumentar que milhares de trabalhadores foram demitidos por esses programas de IA, mas, de fato, se não houvesse os programas de IA esses trabalhos não existiriam porque o trabalho humano adicionaria um custo inaceitável às transações. Até agora, a automação por meio da tecnologia de IA criou mais empregos do que eliminou, e criou empregos mais interessantes e com remuneração mais elevada.

2.2. Machine Learning

O Aprendizado de Máquina e Reconhecimento de Padrões é um dos campos mais populares da IA. O foco nesta área é desenvolver softwares que podem aprender com os dados que lhe são fornecidos. Com isso é possível fazer inferências e previsões em novos dados baseados nos modelos de aprendizagem. Então seus métodos têm como foco principal os dados, se o conjunto de dados for muito pequeno, pode limitar o modelo de aprendizagem. (NETWORKS, 2018)

Dentro as diversas subáreas do Aprendizado de Máquina está o método de Agrupamento, este que separa os dados em grupos homogêneos, geralmente aplicado em um grande 14 agrupamento de informações, de modo a agrupar esses dados baseado nas suas características importantes do problema. Para que um agrupamento possa ser considerado eficaz, é importante saber que os elementos de um grupo devem conter características parecidas, e cada diferente grupo deve possuir dados bem diferentes dos outros. (NETWORKS, 2018)

2.3. Python

A primeira versão do Python foi lançada na década de 1980 pelo matemático Van Rossum, e concluindo-se em 1991, sendo a versão “Python” 1.0, com intuito otimizar a leitura de códigos e gerar produtividade para quem está desenvolvendo. (BOITA et al., 2020)

A linguagem possui diversas aplicações desde desenvolvimento web, interface

gráfica do utilizador, games 3D entre outras, entregando uma ampla variedade de possibilidade de trabalho. (BOITA et al., 2020)

É importante ressaltar que Python é uma linguagem de fácil entendimento, tendo acesso a toda documentação aberta, por ter uma sintaxe simples e de fácil compreensão ela vem ganhando notoriedade nos últimos anos. (PYTHON SOFTWARE FOUNDATION, 2021)

A instalação da plataforma Windows inclui toda a biblioteca padrão e geralmente fornece componentes complementares. Para sistemas operacionais parecidos a Unix, é fornecido pacotes, sendo necessário a utilização de ferramentas de empacotamento que são fornecidas no próprio sistema operacional. (PYTHON SOFTWARE FOUNDATION, 2021).

No início um grande problema que o Python sofria era a falta de bibliotecas que auxiliassem na análise de dados e no processo de visualização dos dados, conforme o passar dos tempos foi se desenvolvendo uma comunidade grande e ativa de processamento científico e análise de dados, gerando assim popularidade para a linguagem. (MCKINNEY, 2015)

2.4 Fundamentos de Python para análise dos dados

Um ponto positivo do Python, é a sua capacidade de integração com outras linguagens, e a quantidade ampla de bibliotecas para auxiliar em seus tratamentos. As bibliotecas que serão apresentadas a seguir oferecem funcionalidades otimizadas e completas para trabalhar com ciência de dados. (BOSCHETTI, 2015)

2.5 Pandas

A biblioteca Pandas teve seu surgimento em 2010, e desde lá ele vem auxiliando e dando a possibilidade do Python de trabalhar com uma quantidade de formato de dados gigantescas, como arquivos csv, tsv e banco de dados SQL. Com os dados em mãos você tem a capacidade de manipular, alinhar e realizar comparações facilmente. (CHEN, 2018)

Os dados quando carregados, são exibidos em linhas e colunas, que são conhecidos como DataFrame, semelhante à uma tabela do Excel. Existem também as Séries, que são um matrizes unidimensional rotuladas capazes de armazenar dados de qualquer tipo. (MCKINNEY, 2015)

2.6 Dataframes

Dataframes são estruturas de dados bidimensional de forma tabular em linhas e colunas e mutável em tamanho. Para realizar a criação de um Dataframe é necessário passar os dados para seu construtor. (MCKINNEY, 2015)

Figura 1 - Criação de DataFrame

```
In [19]: import pandas as pd

In [20]: # Criação da lista dos alunos
listaAlunos = {'Name':['Italo', 'Guilherme', 'Pedro', 'João', 'Eliseu'],
               'Age':['22','21','25','28', '30'],
               'Cidade':['Goianésia', 'Goianésia', 'Jaraguá', 'Cocalzinho de Goiás', 'Ouro Verde']}
```

```
In [21]: # Transformando a lista em um DataFrame
dfListaAlunos = pd.DataFrame(listaAlunos)
```

```
In [22]: # Visualizando o dataframe criado
dfListaAlunos
```

Out[22]:

	Name	Age	Cidade
0	Italo	22	Goianésia
1	Guilherme	21	Goianésia
2	Pedro	25	Jaraguá
3	João	28	Cocalzinho de Goiás
4	Eliseu	30	Ouro Verde

Fonte: Próprio autor, 2021

2.7 Séries

São matrizes unidimensionais capazes de armazenar dados de qualquer valor, é bem parecida com listas em Python. Podemos interpretar como um array, porém possui índices que são criados pelo pandas. (MCKINNEY, 2015)

Figura 2 - Criação de Series

```
import pandas as pd

# Criando uma Series com pandas
SeriesList = pd.Series([2,4,6,8,15,10,9])
print(SeriesList)
```

0	2
1	4
2	6
3	8
4	15
5	10
6	9

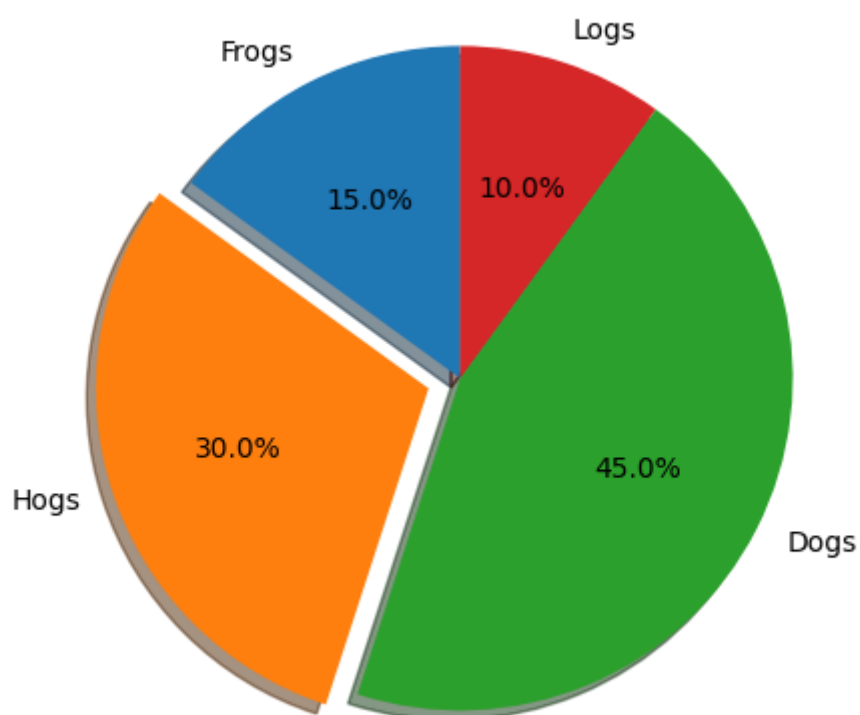
```
dtype: int64
```

Fonte: Próprio autor, 2021

2.8 Matplotlib

Matplotlib é uma biblioteca Python bastante conhecida e utilizada para realizar plotagens 2D e 3D, geração de resultados gráficos, relatórios quantitativos de deslocamentos e esforços apropriados para apresentação. (MCKINNEY, 2015)

Permite a criação de gráficos de diversos formatos e também realizar anotações personalizadas da maneira que preferir. A biblioteca te dá a possibilidade de utilizar diversos tipos de gráficos, sendo eles: gráficos de barras, gráficos de plot, Scatterplot e o que foi utilizado no trabalho, pie chart, mais conhecido como gráfico de pizza. (TOSI, 2009)

Figura 3 - Pie Chart (Gráfico de pizza)

Fonte: matplotlib.org (s.d)

2.9 Folium

A biblioteca Folium realizou uma junção das forças do ecossistema Python com a capacidade de visualizações de mapas que a biblioteca Leaflet.js fornece, proporcionando facilidade na hora de realizar as manipulações dos dados e visualização em mapa. (AMORIM, 2021)

Folium possibilita a plotagem de diversos estilos de mapas, sendo eles: Stamen Terrain, Stamen Toner. Cada um trazendo um layout diferente ao mapa.

Figura 4 - Stamen Terrain



Fonte: python-visualization.github.io/ (s.d)

Figura 5 - Stamen Toner



Fonte: python-visualization.github.io/ (s.d)

2.10 Mineração de dados

Com a popularização da Internet, temos todos os dias um download muito intenso de informações, embora não tenhamos conscientemente essa percepção. Por exemplo, as redes sociais precisam apenas passar o dedo no feed de notícias para apresentar uma grande quantidade de dados para nós. O conteúdo nem sempre é relevante para o conteúdo que nos interessa, por isso podemos optar por filtrar todo esse conteúdo e nos concentrar apenas nas páginas e nas pessoas que atendem aos nossos gostos. (GALVÃO, 2009)

A mineração de dados está filtrando grandes quantidades de informações e passando-as por uma peneira tecnológica que revela padrões consistentes e informações úteis para necessidades específicas de negócios. Este é o primeiro passo para verificar uma estratégia verdadeiramente útil, obviamente, esta tarefa é impossível de ser realizada apenas com mão de obra. Isso porque a Internet superou as expectativas das pessoas e é impossível dominá-la sem a ajuda de outras ferramentas igualmente poderosas. (CAMILO, 2009)

Portanto, podemos dizer que a mineração de dados é um dos elementos da transformação digital e é muito importante para empresas de qualquer setor. Utilizando algoritmos de aprendizagem e a uma velocidade bastante rápida, a mineração de dados é um sistema baseado em aprendizagem de máquina, nesta base que tenta comprovar a coleta de informações, tendências de consumo e interações geradas por clientes potenciais. (GALVÃO, 2009)

3. METODOLOGIA

O estado de Goiás assim como outros estados do Brasil foi bastante afetado pelo o vírus, sendo o nono estado com maior número de casos de óbitos. (O POPULAR, 2021) Possui também a quarta posição em maior taxa de ocupação de leitos de UTIs do Brasil. (OLIVEIRA, 2021)

Pelo fato dos dados do COVID-19 se fundamentar em números, sua exibição geralmente é feita através de gráficos que apresentam resultados como o número de casos confirmados, número de casos em idosos, o número de óbitos confirmados e os que estão sob acompanhamento médico, gráficos esses que podem ser classificados de diferentes maneiras possibilitando assim meios de realizar comparações entre dois deles para poder apresentar possíveis resultados. Será dividido nas seguintes fases: coleta dos dados, classificação dos dados e entrega dos resultados.

Foi utilizada a base de dados disponibilizada pelo Brasil.IO, esse que disponibiliza informações sobre diversos assuntos que acontecem no Brasil, entre eles nós selecionamos os dados referente ao Covid-19. Esses dados foram coletados pela biblioteca Pandas, foram realizadas comparações e classificações destes dados com a biblioteca chamada Matplotlib e por último plotamos o mapa do estado de goiás com a biblioteca Folium, para apresentar as áreas que houve maior nível de contágio e maior índice de mortalidade.

4. DESENVOLVIMENTO

4.1 Base de Dados

O conjunto de dados obtidos através da base de dados do Brasil.IO, é composto por 11 colunas, sendo elas: Data, Estado, Cidade, Tipo de local, Confirmações, Mortes, Última Atualização, População estimada, Código IBGE, confirmados/100k hab. e Mortes/confirmados. A primeira coluna registra a data das atualizações, que é apresentada pelo ano, mês e dia que a informação foi coletada, a segunda coluna exibe o nome dos estados do país, a terceira as cidades existentes, a quarta difere o tipo do local, se é cidade, estado ou país, a quinta trás a somatória do número de casos confirmados, a sexta trás o número total de mortes comprovadas, a sétima que diz se aquela atualização foi a mais recente publicada, a oitava que mostra a quantidade estimada de habitantes daquela região, a nona que é a identificação dos municípios definido pelo IBGE, a décima que traz o número de casos confirmados por 100 mil habitantes, e a última mostra o número de mortes por cada caso confirmado.

Figura 6 - Base de dados completa

Data	Estado	Cidade	Tipo de Local	Confirmações	Mortes	Última Atualização?	População estimada	Código IBGE	Confirmados/100k hab	Mortes/Confirmados
2021-08-20	GO	Abadia de Goiás	city	1727	37	False	8958.0	5200050.0	19278.85689	0.0214
2021-08-20	GO	Abadiânia	city	495	41	False	20461.0	5200100.0	2419.23660	0.0828
2021-08-20	GO	Acreúna	city	2486	58	False	22546.0	5200134.0	11026.34614	0.0233
2021-08-20	GO	Adelândia	city	227	4	False	2516.0	5200159.0	9022.25755	0.0176
2021-08-20	GO	Água Fria de Goiás	city	321	10	False	5793.0	5200175.0	5541.17038	0.0312
...
2021-08-20	GO	Varjão	city	14	14	False	3838.0	5221908.0	14877.54039	0.0245

Fonte: Próprio autor, 2021

Com a base de dados em mãos foi necessário realizar a importação das bibliotecas para a visualização dos dados.

Figura 7 - Importação das bibliotecas

```
# Importando as bibliotecas
import pandas as pd
import matplotlib.pyplot as plt
import folium
```

Fonte: Próprio Autor, 2021.

Após realizar a importação das bibliotecas necessárias, foi necessário armazenar a base de dados em uma variável para que fosse possível ter acesso aos dados e fazer as manipulações nos mesmos.

Figura 8 - Importando a base de dados em uma variável

```
# Armazenando o arquivo csv do dia 20/09 em uma variável
dataState_day20_08 = pd.read_csv('dados/covid 20-08.csv')
```

Fonte: Próprio Autor, 2021.

Com a variável criada foi possível dar início ao tratamento dos dados coletados, o primeiro passo é diminuir a quantidade de coluna, filtrando apenas pelos dados definidos sendo eles: data, cidade, confirmações e mortes.

Figura 9 – Filtragem da base de dados, exibindo apenas as colunas definidas

```
# Casos de óbito e casos confirmados no estado - dia 20 de setembro
dataDeathsConfirmation_20_08 = dataState_day20_08[["Data", "Cidade", "Confirmações", "Mortes"]]
# Apresentando os casos de óbitos e casos confirmados
dataDeathsConfirmation 20 08
```

Fonte: Próprio Autor, 2021

Figura 10 - Base de dados filtrada

Data	Cidade	Confirmações	Mortes
2021-08-20	Abadia de Goiás	1727	37
2021-08-20	Abadiânia	495	41
2021-08-20	Acreúna	2486	58
2021-08-20	Adelândia	227	4
2021-08-20	Água Fria de Goiás	321	10
...
2021-08-20	Varjão	14	14

Fonte: Próprio Autor, 2021

Com os dados filtrados foi possível dar início ao tratamento dos dados, onde foi realizada as comparações e apresentada as oito cidades com maiores níveis de casos confirmados e casos de óbitos, e colocamos eles em ordem decrescente.

Figura 11 - Top 8 cidades com maior índice de casos confirmados

Data	Cidade	Confirmações	Mortes
2021-08-20	Goiânia	186015	6148
2021-08-20	Aparecida de Goiânia	78285	1630
2021-08-20	Anápolis	46085	1664
2021-08-20	Rio Verde	29514	761
2021-08-20	Luziânia	28814	440
2021-08-20	Senador Canedo	21422	362
2021-08-20	Trindade	13336	456
2021-08-20	Valparaíso de Goiás	13315	366

Fonte: Próprio Autor, 2021

Figura 12 - Top 8 cidades com maior índice de óbitos

Data	Cidade	Confirmações	Mortes
2021-08-20	Goiânia	186015	6148
2021-08-20	Anápolis	46085	1664
2021-08-20	Aparecida de Goiânia	78285	1630
2021-08-20	Rio Verde	29514	761
2021-08-20	Trindade	13336	456
2021-08-20	Itumbiara	12739	444
2021-08-20	Luziânia	28814	440
2021-08-20	Jataí	13307	389

Fonte: Próprio Autor, 2021

Logo após isso, plotamos os dados manipulados em gráfico, que apresentaram as porcentagens de forma precisa e com qualidade de impressão as oito cidades do estado de Goiás que contém os maiores níveis de casos confirmados e óbitos. Utilizamos o parâmetro `explode` do `pyplot.pie` para realizar a separação da cidade com maior índice.

Figura 13 - Código da plotagem do gráfico Pie Chart, Top 8 Confirmações

```
# Plotagem dos graficos das 8 cidades com maiores niveis de confirmações
exp = [0.2,0,0,0,0,0,0,0]
plt.pie(top8Confirmacoes["Confirmações"], labels=top8Confirmacoes["Cidade"], autopct="%1.1f%%", explode= exp,
radius=1.5, shadow=True)
plt.show()
# Grafico de casos confirmados
```

Figura 14 - Código da plotagem do gráfico Pie Chart, Top 8 Óbitos

```
# Plotagem dos graficos das 8 cidades com maiores niveis de óbitos
exp = [0.2,0,0,0,0,0,0,0]
plt.pie(top8Mortes["Mortes"], labels=top8Mortes["Cidade"], radius=1.5, explode = exp,
shadow=True, autopct="%1.1f%%")
plt.show()
```

Fonte: Próprio Autor, 2021

4.2 Plotagem do Mapa

Para realizar a visualização do mapa, foi utilizado a biblioteca Folium, com ela tivemos a possibilidade de apresentar um mapa passando determinada coordenada, como foi realizado na figura 12.

Figura 15 - Código da plotagem do gráfico, passando as coordenadas

```
mapa = folium.Map(
    location=[-16.3780762, -48.786904],
    zoom_start=9
)
mapa.add_child(folium.LatLngPopup())
```

Fonte: Próprio Autor, 2021

Para realizar a marcação no mapa, utilizamos a propriedade *popup*, onde você replica para todas as cidades, alterando apenas as coordenadas e os dados da cidade, como é demonstrado na Figura 13.

Figura 16 - Código feito para visualização dos *popup*

```
# Goiânia
folium.Marker(
    [-16.6946667,-49.2671275],
    popup='<i> 6148 casos de óbitos confirmados <i>',
    tooltip='Clique aqui para ver a quantidade de casos em Goiânia',
    icon=folium.Icon(color='red')
).add_to(mapa)
|
# Anápolis
folium.Marker(
    [-16.328446678676315, -48.95244663669904],
    popup='<i> 1664 casos de óbitos confirmados <i>',
    tooltip='Clique aqui para ver a quantidade de casos em Anápolis',
    icon=folium.Icon(color='red')
).add_to(mapa)

# Aparecida de Goiânia
folium.Marker(
    [-16.8650,-49.2462],
    popup='<i> 1630 casos de óbitos confirmados <i>',
    tooltip='Clique aqui para ver a quantidade de casos em Aparecida de Goiânia',
    icon=folium.Icon(color='red')
).add_to(mapa)
```

Fonte: Próprio Autor, 2021

Para finalizarmos utilizamos a propriedade `CircleMarker` que dá a possibilidade de colocar marcações de regiões com em um raio atribuído, sendo colorida de uma cor passada, como é demonstrado na Figura 14. Realizamos o mesmo trecho de código para cada uma das cidades.

Figura 17 - Código para a marcação das regiões no mapa

```
# Goiania
folium.CircleMarker(
    [-16.6946667, -49.2671275],
    radius=30,
    color='#DC6F57',
    fill_color='#DC6F57'
).add_to(mapa)

# Anápolis
folium.CircleMarker(
    [-16.328446678676315, -48.95244663669904],
    radius=30,
    color='#DC6F57',
    fill_color='#DC6F57'
).add_to(mapa)

# Aparecida de Goiania
folium.CircleMarker(
    [-16.8729, -49.2435],
    radius=30,
    color='#DC6F57',
    fill_color='#DC6F57'
).add_to(mapa)
```

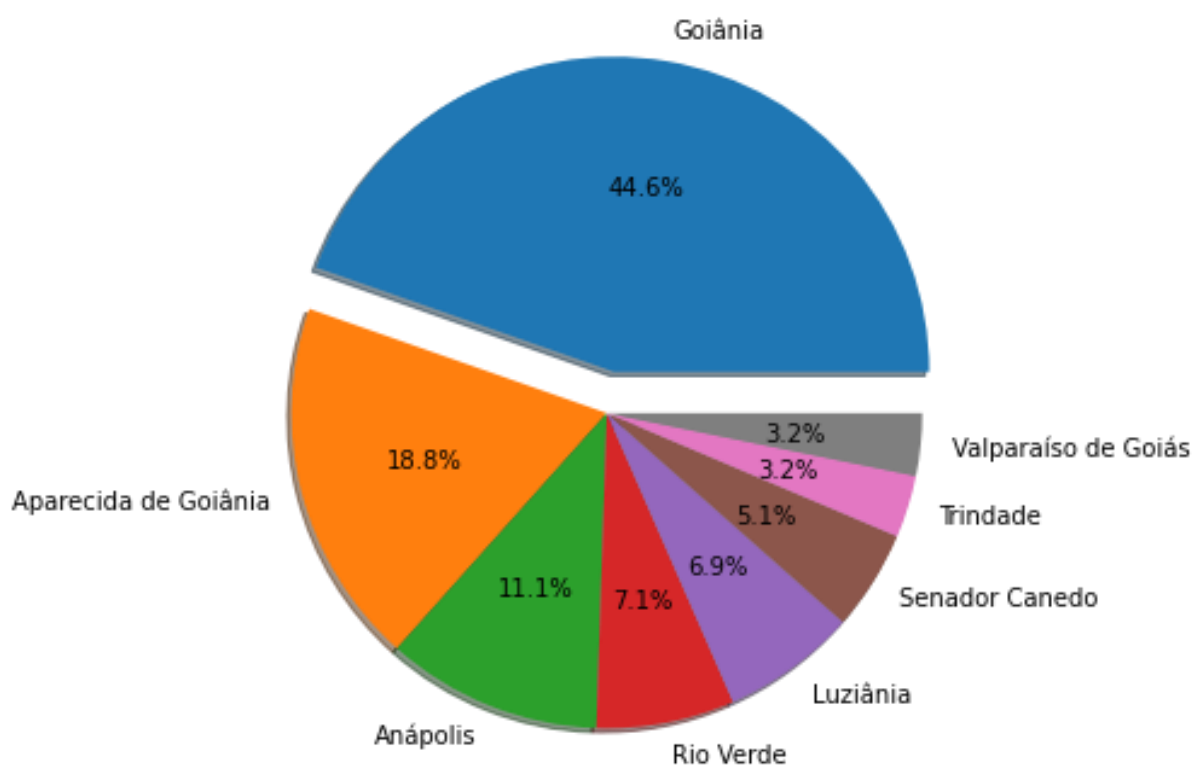
Fonte: Próprio Autor, 2021

5. RESULTADOS ALCANÇADOS

Foram coletadas as informações do dia 20 de agosto de 2021, onde apresentamos as informações de cada cidade do estado de Goiás, demonstrado na Figura 3, além disso realizamos o tratamento de algumas informações em específico como a data, cidade, casos confirmados e casos de óbitos, como foi demonstrado na Figura 9.

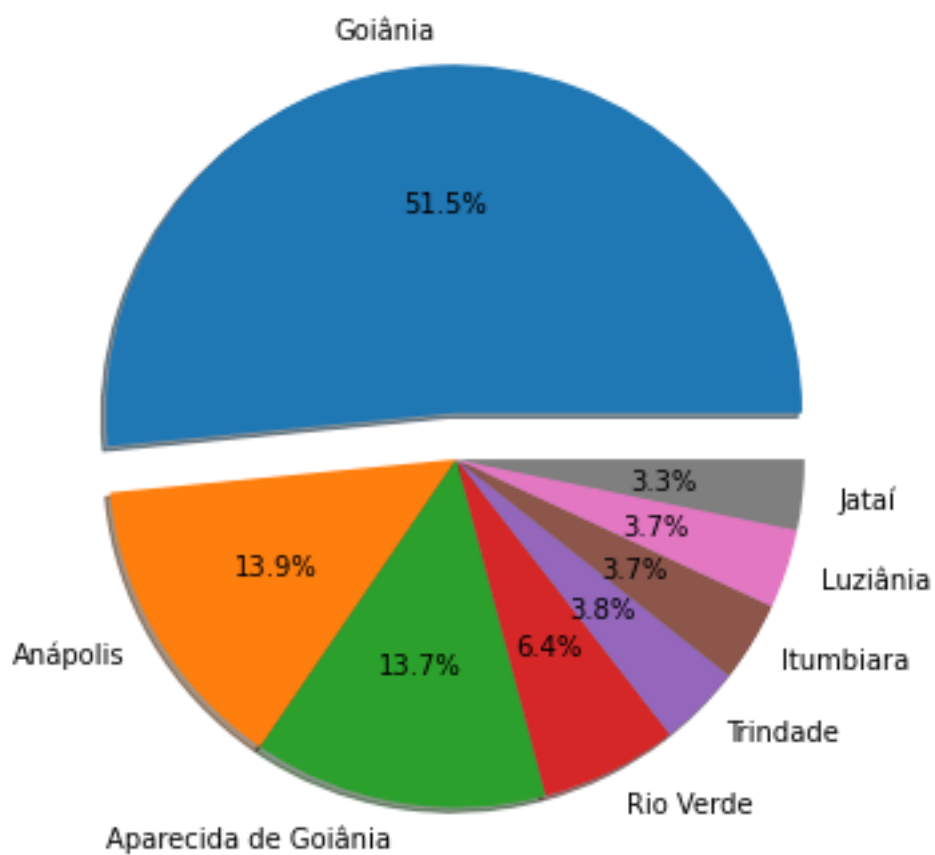
Com as informações selecionadas foram realizadas as classificações das 8 regiões que apresentavam os maiores índices de casos confirmados e casos de óbitos. Sendo aprestando em forma de gráfico essas regiões, o tipo de gráfico escolhido foi o pie Chart, conhecido como gráfico de pizza.

Figura 18 - Gráfico Top 8 casos confirmados



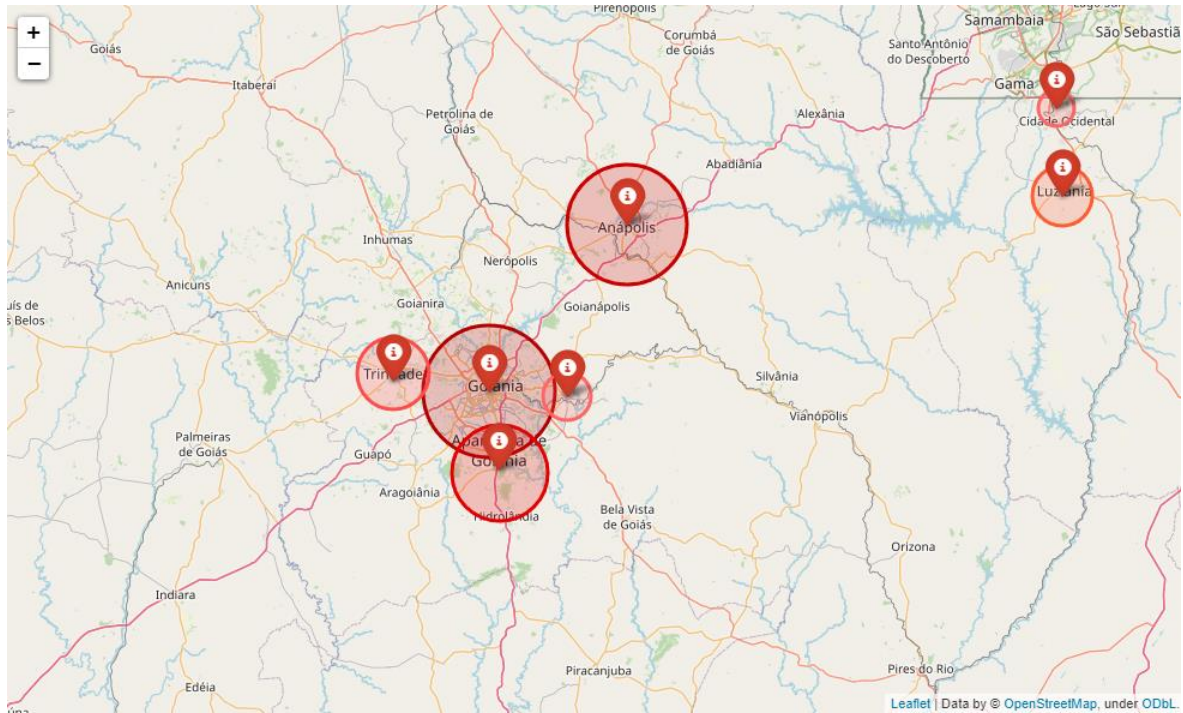
Fonte: Próprio Autor, 2021

Figura 19 - Gráfico Top 8 óbitos confirmados

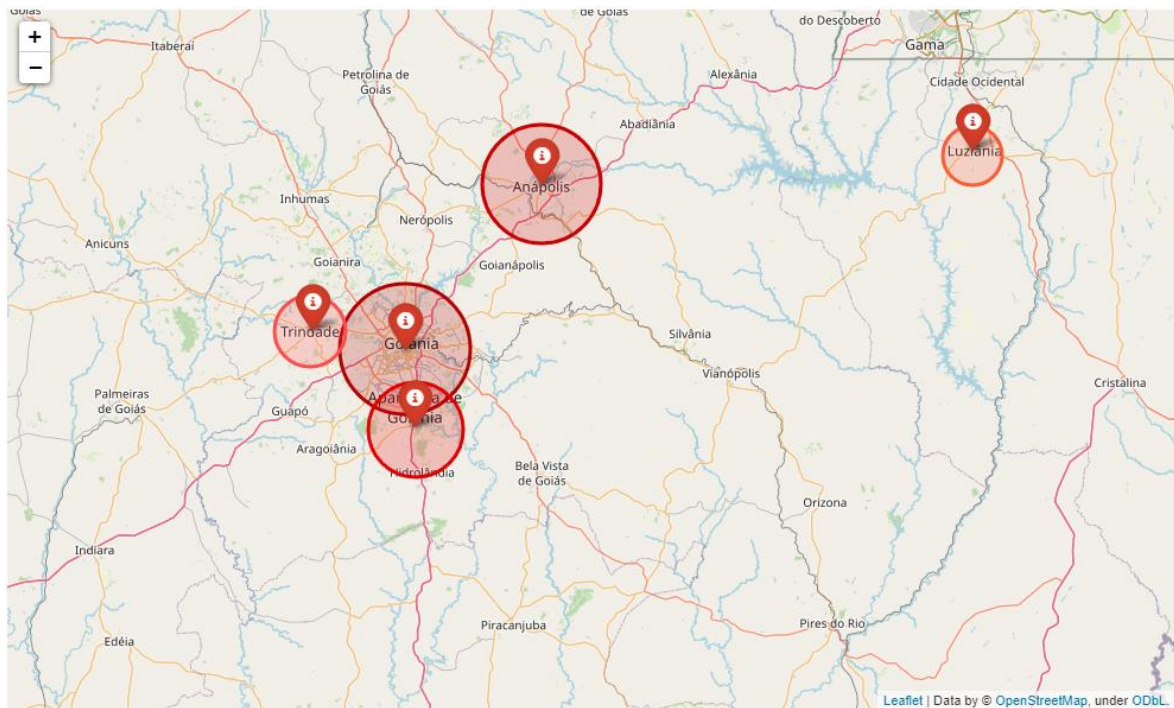


Fonte: Próprio Autor, 2021

Após os resultados em gráficos, realizamos a plotagem do mapa, onde apresentamos popups em cada uma da cidade, ao ser clicado aparecem as informações que foram tratadas nos gráficos na determinada cidade.

Figura 20 - Mapa de Goiás com marcações das 8 cidades com maiores índices de casos confirmados

Fonte: Próprio Autor, 2021

Figura 21 - Mapa de Goiás com marcações das 8 cidades com maiores índices de óbitos confirmados

Fonte: Próprio Autor, 2021

6. CONCLUSÃO E CONSIDERAÇÕES FINAIS

Através dos métodos aplicados e das análises feitas, foi possível coletar informações importantes que estavam espalhadas de forma implícita na internet, reunindo assim dados relevantes e apresentando-os de forma visual através de mapas e gráficos, tornando mais claro a leitura e o entendimento dos elementos apresentados, facilitando assim o entendimento dessas informações para toda a população interessada nesses dados atualizados e de fácil acesso.

A base de dados utilizada tem como objetivo ressaltar as informações mais graves sobre a situação do covid-19 nos estados do Brasil. Logo foi escolhido o estado de Goiás como referência para extração dessas informações e os dados definidos para tratamento foram as cidades do estado de Goiás que contém o maior nível de contaminação em uma escala geral e também o número total de mortes comprovadas.

Além de ajudar na análise dos dados do covid-19 e reunir essas informações para trazê-las de forma mais evidente para os que se interessam, a pesquisa teve como objetivo também agregar informações e resultados sobre o covid-19 no banco de pesquisas da internet, incentivando assim novas análises e buscas sobre esse tema que é tão importante e que tem um alto nível de risco para toda a sociedade.

7. REFERÊNCIAS BIBLIOGRÁFICAS

- AMORIM, Fabrício Rosa; SCHMIDT, Marcio Augusto Reolon. **WebMapa: Desenvolvimento de Mapas Web por um Complemento QGIS**. Revista Brasileira de Cartografia, v. 73, n. 3, p. 842-854, 2021.
- BOITA, J. et al. **Python 101**. Universidade Federal de Santa Maria - Campus Cachoeira do Sul: pesquisa, ensino e extensão. Ciência e Natura, v. 42, p. e25, 2020.
- BOSCHETTI, A. and Massaron, L. **Python data science essentials**. Packt Publishing Ltd. 2015.
- CAMILO, C. O., & Silva, J. C. D. Mineração de dados: Conceitos, tarefas, métodos e ferramentas. Universidade Federal de Goiás (UFG), 1(1), 1-29, 2009.
- CAVALCANTE, J. R. et al. **COVID-19 no Brasil: evolução da epidemia até a semana epidemiológica 20 de 2020**. Epidemiologia e serviços de saúde: revista do Sistema Unico de Saúde do Brasil, v. 29, n. 4, Brasília, 2020.
- CELUPPI, I. C. et al. **Uma análise sobre o desenvolvimento de tecnologias digitais em saúde para o enfrentamento da COVID-19 no Brasil e no mundo**. Cadernos de saúde pública, v. 37, n. 3, Rio de Janeiro, 2021.
- CHEN, Daniel Y. **Análise de dados com Python e Pandas**. Novatec Editora, 2018.
- FREITAS, A. R. R.; NAPIMOGA, M.; DONALISIO, M. R. **Análise da gravidade da pandemia de Covid-19**. Epidemiologia e serviços de saúde : revista do Sistema Unico de Saúde do Brasil, v. 29, n. 2, p. e2020119, 2020.
- FREITAS, R. A. B. DE et al. **Prospecção Científica sobre Epidemiologia e Prevenção da Covid-19 Aliada à Inteligência Artificial**. Cadernos de Prospecção, v. 13, n. 2 COVID-19, p. 543, 2020.
- GALVÃO, Noemi Dreyer; MARIN, Heimar de Fátima. Técnica de mineração de dados: uma revisão da literatura. Acta Paulista de Enfermagem, v. 22, p. 686-690, 2009.
- GRONER, M. F. et al. **Effects of Covid-19 on male reproductive system**. International braz j urol : official journal of the Brazilian Society of Urology, v. 47, n. 1, p. 185–190, São Paulo, 2021.
- JAKOVLJEVIC, M. et al. **Covid-19 pandemia and public and global mental health from the perspective of global health security**. Psychiatria Danubina, v. 32, n. 1, p. 6–14, Croácia, 2020.
- O Popular. **Goiás tem 5ª maior letalidade por Covid-19 no Brasil | 20 de Maio de 2021 às 19h55**. Disponível em: <https://opopular.com.br/noticias/cidades/goi%C3%A1s-tem-5%C2%AA-maior-letalidade-por-covid-19-no-brasil-1.2253917>. Acesso em: 21 out. 2021.
- PORTO, Fábio; ZIVIANI, Artur. **Ciência de dados**. III Seminário de Grandes Desafios da Computação no Brasil, Rio de Janeiro, RJ, 2014.
- LIMA, C. M. A. DE O. **Informações sobre o novo coronavírus (COVID-19)**. Colégio Brasileiro de radiologia e diagnóstico por imagem, v. 53, n. 2, p. V–VI, São Paulo, 2020.

MARQUES, José Roberto. **Inteligência artificial: vantagens e desvantagens quanto ao seu uso.2017**. Disponível em:
<<https://www.ibccoaching.com.br/portal/artigos/inteligencia-artificial-vantagens-desvantagens-quanto-seu-uso/>> Acesso em: 10 set. 2019

MCKINNEY, Wes, et al. **Pandas: a foundational Python library for data analysis and statistics**. Python for high performance and scientific computing, 2011, 14.9: 1-9.

NETWORKS, M. N. **Deep Learning Inferences With**. Learning, n. April, p. 1–15, Brasília, 2018.

NEVES, B. C. **Metodologias, ferramentas e aplicações da inteligência artificial nas diferentes linhas do combate a Covid-19**. Folha de Rostto, v. 6, n. 2, p. 44-57. 2020.

OLIVEIRA, **Goiás tem a maior média móvel de mortes por Covid-19 do Brasil, apontam dados do consórcio** | 17 de Março de 2021 às 14h31. Disponível em:
<https://g1.globo.com/go/goias/noticia/2021/03/17/goias-tem-a-maior-media-movel-de-mortes-por-covid-19-do-brasil-apontam-dados-do-consorcio.ghtml>. Acesso em: 20 out. 2021.

SILVA, J. A. S.;MAIRINK, C. H. P.**Inteligência artificial: aliada ou inimiga**. LIBERTAS: Rev. Ciênci. Soc. Apl., Belo Horizonte, v. 9, n. 2, p. 64-85, ago./dez. 2019.

TOSI, Sandro. **Matplotlib for Python developers**. Packt Publishing Ltd, 2009.