

UNIVERSIDADE FEDERAL DE SÃO JOÃO DEL-REI

Valney de Paiva Faria e Ygor Magalhães

**Análise da propagação da COVID-19 em
cidades do estado de Minas Gerais durante o
ano de 2020**

São João del-Rei

2020

Sumário

1	INTRODUÇÃO	2
1.1	Objetivos	3
2	METODOLOGIA	4
2.1	Os Dados	4
2.2	Modelagem da Rede	5
2.3	Experimentos	7
3	RESULTADOS	10
4	CONCLUSÃO	16
	REFERÊNCIAS	17

1 Introdução

O ano de 2020 foi acometido por uma das pandemias mais sérias de que se tem registro. Segundo o Centro de Ciência e Engenharia de Sistemas (Center for Systems Science and Engineering – CSSE) da Universidade Johns Hopkins ¹, até a data de escrita deste trabalho, haviam sido registrados mais de 65 milhões de casos de COVID-19 ao redor do mundo. E desse montante, mais de 1,5 milhões de pessoas tiveram suas vidas interrompidas pela doença. O Brasil, com seus 211 milhões de habitantes ² (6^a posição no globo), figura no ranking da COVID-19 com os números amargos de 6 milhões de casos (3^a posição no globo) e 175 mil mortes (2^a posição no globo). A doença, cuja origem circula nas redes como uma possível consequência do tráfico de animais silvestres e do desequilíbrio ambiental provocado pelo ser humano, teve os primeiros casos anunciados na região de Wuhan, uma pequena cidade localizada na região central da China e capital da província de Hubei. Transmitida a partir do ar, por meio de gotículas de saliva, e também do contato físico entre pessoas, o vírus se espalhou e fez vítimas por vários meses até que se detectasse o início da calamidade. Os sintomas, muito parecidos com os da gripe ([BRASIL, 2020](#)), ofuscaram a doença e permitiram o seu espalhamento por todos os países do mundo.

Na data de 11 de março de 2020, a Organização Mundial de Saúde (OMS) decreava situação de calamidade global ([VEJA, 2020](#)), identificando uma série de mudanças que deveriam ser aplicadas a todos os setores da sociedade a fim de minimizar os riscos de contágio da doença e diminuir os efeitos de curto e longo prazo que poderiam ser deflagrados caso não fosse possível conter o avanço do vírus. Devido aos riscos sociais e econômicos que podem surgir como efeito de uma pandemia, inúmeros estudos foram elaborados a fim de se calcular as taxas de propagação do vírus, conhecer seu DNA para elaborar uma vacina e entender os efeitos da crise nos países emergentes e quais os malefícios que uma adversidade dessa magnitude poderiam causar no mundo.

Seguindo a tendência, e buscando entender as consequências da pandemia no Brasil, o trabalho de ([ROMERO; SILVA, 2020](#)) realizou um estudo para identificar relações entre a desigualdade socioeconômica no estado do Ceará e os números de casos da doença provocada pelo Coronavírus. Para tal, identificaram como o Índice de Desenvolvimento Humano Municipal (IDH-M) como variável de representação para o nível socioeconômico dessas cidades e a taxa de casos de COVID-19 no estado, durante os três meses iniciais da pandemia no ano de 2020. Como resultado, os pesquisadores observaram que é possível relacionar o status socioeconômico com a ocorrência de casos de COVID-19 no estado.

¹ <https://coronavirus.jhu.edu/map.html>

² <https://www.ibge.gov.br/cidades-e-estados>

Porém, não foi possível mensurar que o IDH-M tenha sido o único fator para os números detectados.

1.1 Objetivos

Com base na premissa avaliada no estudo de (ROMERO; SILVA, 2020), pretende-se, por meio da aplicação de conceitos de Redes Complexas, identificar relações entre um conjunto escolhido dentre as 853 cidades do Estado de Minas Gerais, utilizando, para isso, o IDH-M, a taxa populacional e o números de casos de COVID-19 como parâmetros.

2 Metodologia

Esta seção descreve os passos para a realização deste trabalho, incluindo a obtenção dos dados utilizados, a manipulação sofrida por eles e as técnicas de redes escolhidas para torná-los úteis à obtenção de informações acerca dos casos de COVID-19 nas cidades de Minas Gerais.

2.1 Os Dados

Para a realização das análises deste estudo, foram obtidas duas bases de dados a partir do site *Kaggle*¹. A base *Coronavirus - Brazil*² possui informações sobre os números de casos e mortes identificados em cada uma das cidades do Brasil no período de 27 de março de 2020 a 04 de novembro de 2020. Além disso, há uma relação de cada um dos municípios com os dados de suas respectivas latitudes e longitudes, código de cidade e estado, região, código de saúde da região, dentre outros. A base *Brazilian Cities*³ é uma coleção com 79 atributos sobre cada uma das cidades brasileiras com base em dados do IBGE de 2010. Consta com informações como população residente, população urbana, população rural, IDH-M, produção de alimentos, número de Lojas McDonald's, dentre outros. Apesar de inúmeros desses atributos apresentarem certa relevância no que tange análises socioeconômicas, iremos considerar apenas o IDH-M assim como feito no estudo referência para este trabalho.

O site *Kaggle* disponibiliza suas bases de dados gratuitamente, bastando apenas se cadastrar por meio de um *e-mail* válido. Após o cadastro, basta pesquisar pelo conteúdo de seu interesse e realizar o *download*. Os arquivos disponibilizados para este caso encontram-se em formato *.csv* e podem ser manipulados por quaisquer editores de texto disponíveis gratuitamente na *internet*. No caso deste trabalho, não houve a necessidade de armazenamento dos dados em um SGBD, já que as ferramentas escolhidas cumpriram suas funções em tempo viável ao desenvolvimento da aplicação.

Para realizar a extração dos dados, fez-se uso da linguagem *Python* e da biblioteca *Pandas*, uma ferramenta gratuita para manipulação e análise de grandes volumes de informações que permite criar tabelas numéricas manipuláveis e operar diretamente sobre elas. Por meio do uso de *Python* e *Pandas*, foi realizada a junção entre os arquivos utilizando como parâmetro o nome das cidades, comum entre as tabelas.

¹ www.kaggle.com

² <https://www.kaggle.com/unanimad/corona-virus-brazil>

³ <https://www.kaggle.com/crisparada/brazilian-cities>

2.2 Modelagem da Rede

Para a modelagem da rede como um grafo, primeiramente construiu-se, utilizando a Linguagem *C*, um *script* cuja função era selecionar os dados a serem utilizados e gerar a descrição GML (Graph Modeling Language) dos mesmos. A seleção das cidades se fez necessária para facilitar a visualização e obter informações passíveis de entendimento. Para garantir uma base suficientemente relevante, definiu-se como conjunto de testes uma lista contendo as 58 cidades de Minas Gerais com mais de 50 mil habitantes residentes, segundo o Censo IBGE 2010. Segue abaixo a lista contendo os municípios considerados relacionando-os seu IDH-M e população.

Tabela 1 – Lista de municípios de Minas Gerais por população ([IBGE, 2010](#))

Posição	Município	IDH-M	População 2010
1	Belo Horizonte	0.81	2263631
2	Contagem	0.76	595593
3	Uberlândia	0.789	580561
4	Juiz de Fora	0.778	499216
5	Betim	0.75	372678
6	Montes Claros	0.77	338940
7	Uberaba	0.772	285743
8	Ribeirão das Neves	0.684	272404
9	Governador Valadares	0.727	248674
10	Ipatinga	0.77	235407
11	Sete Lagoas	0.76	208184
12	Santa Luzia	0.715	191977
13	Divinópolis	0.764	190832
14	Ibirité	0.715	8253
15	Poços de Caldas	0.779	148198
16	Patos de Minas	0.765	124101
17	Sabará	0.731	119262
18	Pouso Alegre	0.774	118760
19	Varginha	0.778	118734
20	Barbacena	0.769	114563
21	Teófilo Otoni	0.701	112361
22	Conselheiro Lafaiete	0.76	110658
23	Araguari	0.773	101933
24	Itabira	0.76	101931
25	Passos	0.76	100405
26	Coronel Fabriciano	0.755	99429
27	Ubá	0.72	96248
28	Muriaé	0.73	92629
29	Ituiutaba	0.74	92119
30	Araxá	0.772	88054

Tabela 2 – Lista de municípios de Minas Gerais por população (IBGE, 2010)

Posição	Município	IDH-M	População 2010
31	Lavras	0.782	87422
32	Vespasiano	0.688	84080
33	Itajubá	0.787	80074
34	Itaúna	0.758	79872
35	São João del Rei	0.758	79654
36	Pará de Minas	0.725	78927
37	Timóteo	0.77	78539
38	Nova Lima	0.813	76049
39	Paracatu	0.74	73521
40	João Monlevade	0.758	73080
41	Patrocínio	0.73	71918
42	Nova Serrana	0.715	69346
43	Alfenas	0.76	68964
44	Viçosa	0.775	67078
45	Caratinga	0.706	67042
46	Curvelo	0.713	66960
47	Cataguases	0.75	66352
48	Três Corações	0.74	65609
49	Manhuaçu	0.69	63209
50	Ouro Preto	0.741	61519
51	Janaúba	0.696	60294
52	São Sebastião do Paraíso	0.722	59711
53	Unaí	0.736	59265
54	Formiga	0.755	58961
55	Pedro Leopoldo	0.757	55232
56	Esmeraldas	0.67	55040
57	Pirapora	0.731	51887
58	Ponte Nova	0.72	51046

Após a seleção do conjunto de cidades, calculou-se as porcentagens de casos de COVID-19 em relação ao seu número de habitantes. O valor correspondente à porcentagem foi multiplicado por 1000 para facilitar a visualização após a geração das redes. Foram armazenados na estrutura GML esse e outros parâmetros, como apresentados a seguir:

- *id*: Um identificador único para cada nó;
- *cidade*: O nome do município em questão;
- *casos*: O número de casos registrado para um município em um determinado mês;
- *idh*: O IDH-M do respectivo município;
- *populacao*: A população total do município com base no Censo do IBGE de 2010.

Após a geração da estrutura GML, utilizou-se da coleção de bibliotecas *iGraph* para realizar a leitura das informações e montar a rede. A montagem seguiu a seguinte configuração:

- Vértices: Representam as cidades;
- Tamanho do Vértice: Representa a porcentagem de casos de COVID-19 em relação à população total da cidade;
- Arestas: Para as ligações entre vértices, optou-se por definir uma faixa considerando o IDH-M de cada uma dos municípios. Se uma cidade possuir uma variação de IDH-M de 0,05 para mais ou para menos em relação às outras $n - 1$ cidades, haverá uma aresta entre elas. Supondo, por exemplo, o IDH-M de uma cidade A como sendo 0,78, todas as cidades com IDH-M entre 0,73 e 0,83 terão uma aresta para A .

2.3 Experimentos

Após a etapa de seleção das cidades e dos cálculos realizados, este estudo considerou ao todo os seguintes valores para cada item:

Tabela 3 – Características do grafo e seus respectivos valores

Característica	Valor
Nós	58
Arestas	1198
Densidade	0,72
Grau Médio	41,3
Coeficiente de Clustering	0,87

Para a realização da análise topológica, optou-se pela métrica de Centralidade, a qual produziu os resultados denotados nas Tabelas 2.3 e 2.3, para cada uma das 58 cidades. A cidade de Ituiutaba possui 1.0 como valor de centralidade, indicando que seu respectivo vértice é o mais importante em relação às ligações analisadas para atingir outros vértices.

Tabela 4 – Valores de Centralidade por Cidade

Cidade	Centralidade
Alfenas	0.961
Araguari	0.875
Araxá	0.875
Barbacena	0.929
Belo Horizonte	0.339
Betim	0.982
Caratinga	0.485
Cataguases	0.982
Conselheiro Lafaiete	0.961
Contagem	0.961
Coronel Fabriciano	0.972
Curvelo	0.697
Divinópolis	0.961
Esmeraldas	0.107
Formiga	0.972
Governador Valadares	0.916
Ibirité	0.398
Ipatinga	0.929
Itabira	0.961
Itajubá	0.727
Itaúna	0.961
Ituiutaba	1.0
Janaúba	0.398
João Monlevade	0.961
Juiz de Fora	0.834
Lavras	0.748
Manhuaçu	0.310
Montes Claros	0.929
Muriaé	0.971
Nova Lima	0.339
Nova Serrana	0.718
Ouro Preto	0.999
Paracatu	0.999
Pará de Minas	0.916
Passos	0.961
Patos de Minas	0.929
Patrocínio	0.971
Pedro Leopoldo	0.961
Pirapora	0.971
Poços de Caldas	0.834
Ponte Nova	0.819
Pouso Alegre	0.875
Ribeirão das Neves	0.288
Sabará	0.971

Tabela 5 – Valores de Centralidade por Cidade

Cidade	Centralidade
Santa Luzia	0.718
São João del Rei	0.961
São Sebastião do Paraíso	0.819
Sete Lagoas	0.961
Teófilo Otoni	0.442
Timóteo	0.929
Três Corações	0.999
Ubá	0.819
Uberaba	0.875
Uberlândia	0.727
Unaí	0.981
Varginha	0.834
Vespasiano	0.310
Viçosa	0.875

3 Resultados

Nesta seção, serão analisados os resultados obtidos a partir das análises feitas sobre o conjunto de dados selecionado levando em conta a base coletada por meio do *site Kaggle* e Censos do IBGE. As tabelas constituem *rankings* contendo os 20 municípios com maiores porcentagens de casos em relação à população e os períodos de análise correspondem aos picos de casos de COVID-19 registrados no Brasil e cuja informação está disponibilizada nos *sites Wikipédia*¹ e *Google*².



Figura 1 – Picos de casos de COVID-19 no Brasil

Pode-se observar que, para esse caso de estudo, não há como detectar com facilidade uma relação direta entre a proporção de casos de COVID-19 e o IDH-M das cidades consideradas. A cidade de Uberlândia, por exemplo, possui um IDH-M alto em comparação com os outros municípios analisados. No entanto, desde o início da pandemia, já figurava no *ranking* das 10 cidades com maiores números de casos em Minas Gerais. Em julho, a cidade alcançou a posição 2 do *ranking*, e ao final do período analisado, em Novembro, possuía um número de casos proporcional a mais de 6% de sua população. Em contrapartida, cidades como Ribeirão das Neves (IDH-M: 0,684) e Vespasiano (IDH-M: 0,688), de IDH-M mais baixos, não tiveram aumentos tão agressivos.

Caso se considerasse o número de habitante como um fator determinante para o aumento de casos da doença, seria necessário avaliar outros fatores, visto que a cidade de Belo Horizonte, com seus mais de 2 milhões de habitantes, figura no *ranking* das 20 cidades mais afetadas somente até o pico registrado de agosto. A partir de então, a proporção de casos em relação ao número de habitantes se torna menor do que em outras cidades.

¹ <https://pt.wikipedia.org/>

² <http://www.google.com/covid19>

Tabela 6 – Ranking dos 20 municípios com maior porcentagem de casos em relação a população e seu respectivo IDH na data 31 de Março de 2020

Posição	Nome	porcentagem	IDH
1	Nova Lima	0.026	0.813
2	Belo Horizonte	0.007	0.81
3	Divinópolis	0.005	0.764
4	Juiz de Fora	0.005	0.778
5	Sabará	0.003	0.731
6	São João del Rei	0.003	0.758
7	Manhuaçu	0.002	0.69
8	Uberlândia	0.002	0.789
9	Unaí	0.002	0.736
10	Alfenas	0.001	0.76
11	Betim	0.001	0.75
12	Contagem	0.001	0.76
13	Coronel Fabriciano	0.001	0.755
14	Governador Valadares	0.001	0.727
15	Muriaé	0.001	0.73
16	Patrocínio	0.001	0.73
17	Poços de Caldas	0.001	0.779
18	Santa Luzia	0.001	0.715
19	Sete Lagoas	0.001	0.76
20	Timóteo	0.001	0.77

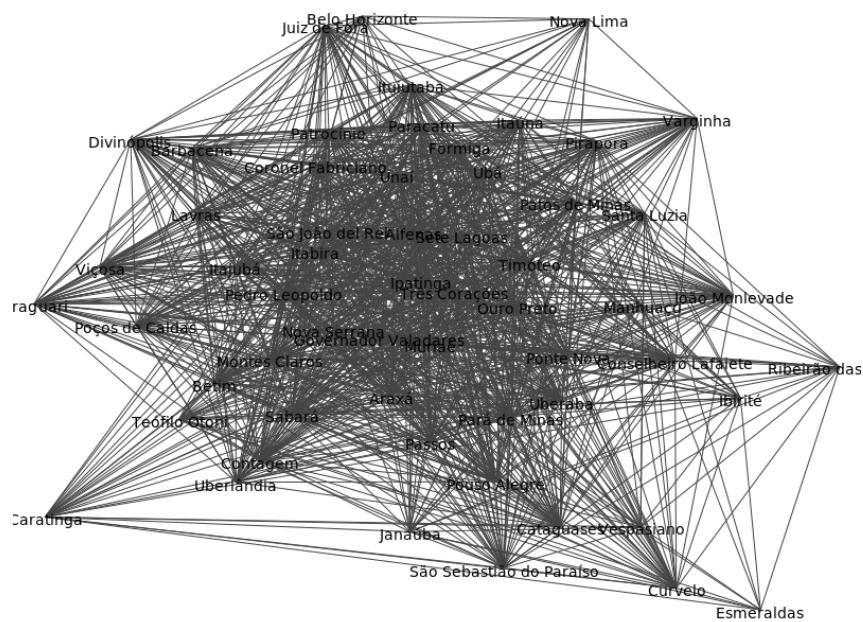


Figura 2 – 31 de Março de 2020

Tabela 7 – Ranking dos 20 municípios com maior porcentagem de casos em relação a população e seu respectivo IDH na data 31 de Julho de 2020

Posição	Nome	Porcentagem	IDH
1	Ipatinga	2.343	0.77
2	Uberlândia	2.145	0.789
3	Muriaé	1.942	0.73
4	Unaí	1.905	0.736
5	Coronel Fabriciano	1.886	0.755
6	Nova Lima	1.736	0.813
7	Governador Valadares	1.559	0.727
8	Araguari	1.423	0.773
9	Itabira	1.268	0.76
10	Timóteo	1.247	0.77
11	Teófilo Otoni	1.193	0.701
12	Caratinga	1.14	0.706
13	Patos de Minas	1.039	0.765
14	Manhuaçu	0.935	0.69
15	Belo Horizonte	0.896	0.81
16	Janaúba	0.861	0.696
17	Ribeirão das Neves	0.826	0.684
18	Paracatu	0.8	0.74
19	Vespasiano	0.783	0.688
20	Pouso Alegre	0.781	0.774

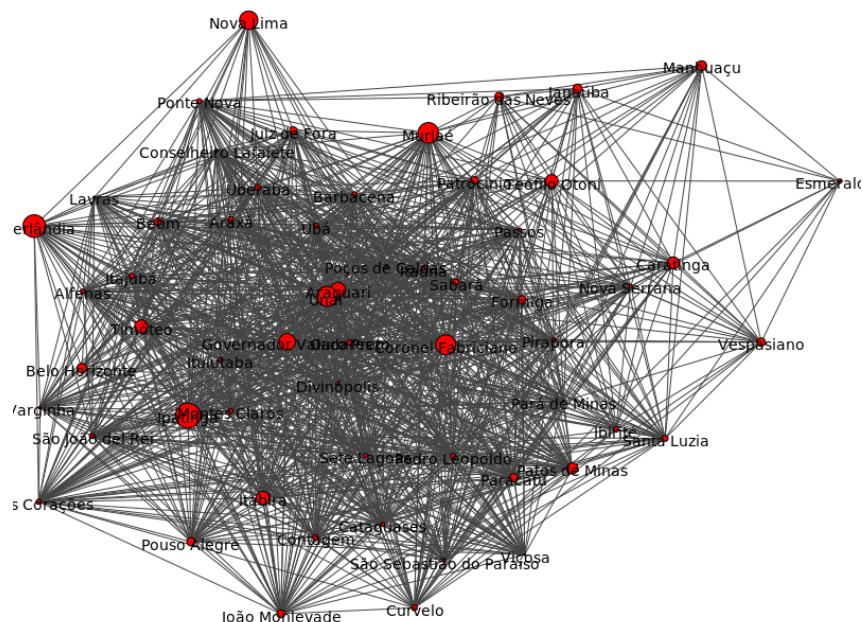


Figura 3 – 31 de Julho de 2020

Tabela 8 – Ranking dos 20 municípios com maior porcentagem de casos em relação a população e seu respectivo IDH na data 31 de Agosto de 2020

Posição	Nome	Porcentagem	IDH
1	Uberlândia	3.446	0.789
2	Nova Lima	3.207	0.813
3	Unaí	3.187	0.736
4	Ipatinga	3.179	0.77
5	Coronel Fabriciano	2.802	0.755
6	Governador Valadares	2.458	0.727
7	Muriaé	2.429	0.73
8	Araguari	2.251	0.773
9	Timóteo	1.928	0.77
10	Caratinga	1.918	0.706
11	Itabira	1.852	0.76
12	Teófilo Otoni	1.814	0.701
13	Paracatu	1.692	0.74
14	Manhuaçu	1.669	0.69
15	Ituiutaba	1.629	0.74
16	Montes Claros	1.586	0.77
17	Pousos Alegre	1.549	0.774
18	Patrocínio	1.535	0.73
19	Belo Horizonte	1.482	0.81
20	Patos de Minas	1.461	0.765

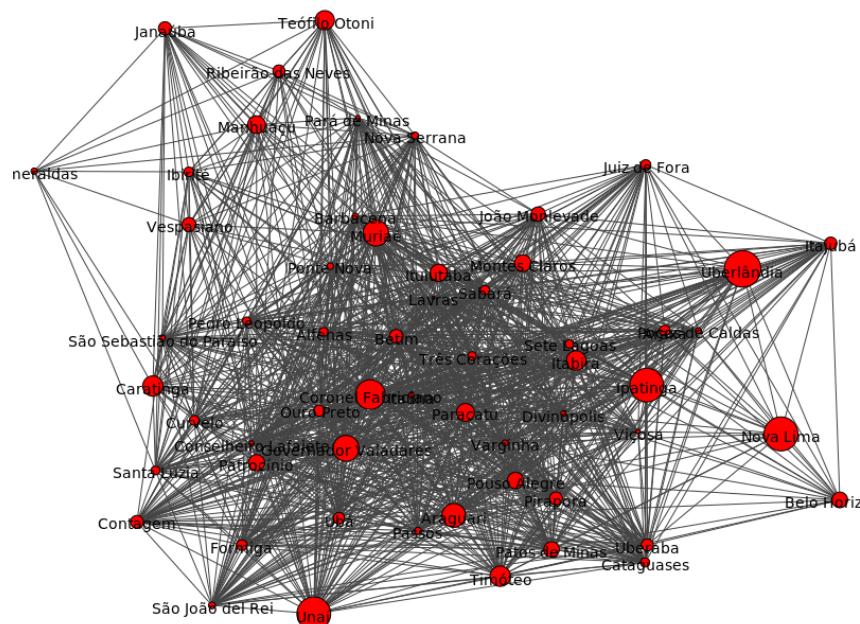


Figura 4 – 31 de Agosto de 2020

Tabela 9 – Ranking dos 20 municípios com maior porcentagem de casos em relação a população e seu respectivo IDH na data 31 de Outubro de 2020

Posição	Nome	Porcentagem	IDH
1	Uberlândia	6.066	0.789
2	Nova Lima	5.35	0.813
3	Unaí	4.547	0.736
4	Araguari	4.334	0.773
5	Ituiutaba	4.198	0.74
6	Ipatinga	4.186	0.77
7	Coronel Fabriciano	3.905	0.755
8	Muriaé	3.761	0.73
9	Governador Valadares	3.475	0.727
10	Montes Claros	3.371	0.77
11	Paracatu	3.366	0.74
12	Timóteo	3.094	0.77
13	Manhuaçu	3.001	0.69
14	Itabira	2.935	0.76
15	Patrocínio	2.91	0.73
16	Alfenas	2.846	0.76
17	Caratinga	2.774	0.706
18	Uberaba	2.747	0.772
19	Pouso Alegre	2.74	0.774
20	Araxá	2.724	0.772

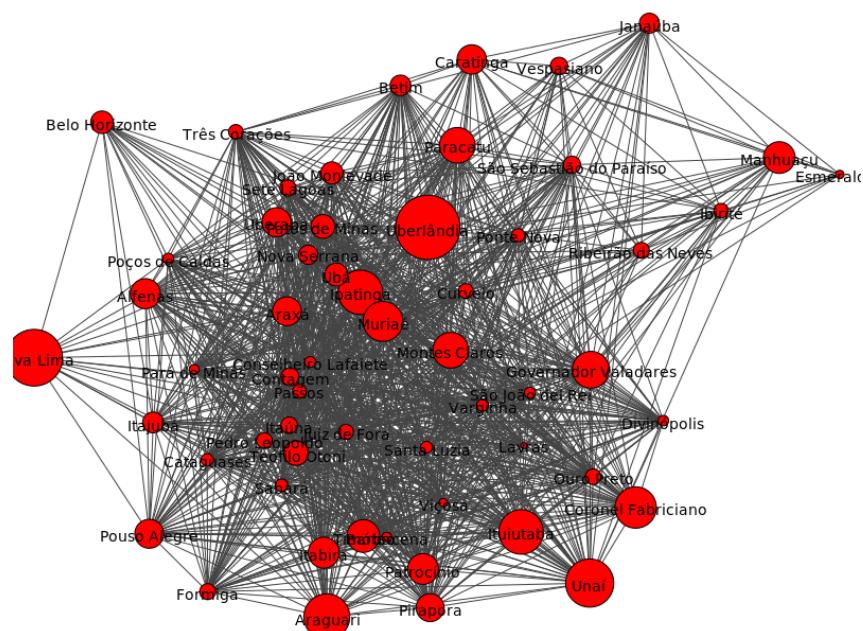


Figura 5 – 31 de Outubro de 2020

Tabela 10 – Ranking dos 20 municípios com maior porcentagem de casos em relação a população e seu respectivo IDH na data 4 de Novembro de 2020

Posição	Nome	Porcentagem	IDH
1	Uberlândia	6.138	0.789
2	Nova Lima	5.35	0.813
3	Unaí	4.568	0.736
4	Araguari	4.345	0.773
5	Ipatinga	4.252	0.77
6	Ituiutaba	4.212	0.74
7	Coronel Fabriciano	3.924	0.755
8	Muriaé	3.781	0.73
9	Governador Valadares	3.504	0.727
10	Montes Claros	3.395	0.77
11	Paracatu	3.38	0.74
12	Timóteo	3.167	0.77
13	Manhuaçu	3.028	0.69
14	Itabira	3.016	0.76
15	Patrocínio	2.939	0.73
16	Alfenas	2.909	0.76
17	Caratinga	2.794	0.706
18	Uberaba	2.793	0.772
19	Pouso Alegre	2.763	0.774
20	Araxá	2.757	0.772

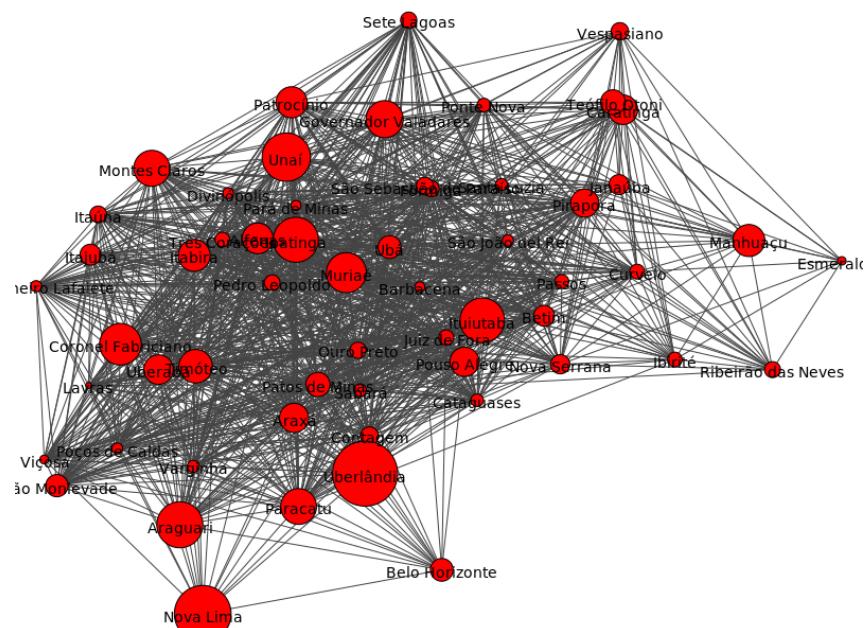


Figura 6 – 04 de Novembro de 2020

4 Conclusão

Este trabalho teve como premissa básica analisar o conteúdo de bases de dados do site *Kaggle* e comparar os resultados com o trabalho de (ROMERO; SILVA, 2020) para verificar se o Índice de Desenvolvimento Humano Municipal (IDH-M) é fator delimitante para a maior ocorrência de casos de COVID-19 em uma determinada região. No contexto deste estudo, optou-se por avaliar os casos de COVID-19 nas cidades do Estado de Minas Gerais devido ao tamanho do conjunto e relevância para os autores. Após a realização dos testes, verificou-se, assim como no trabalho de (ROMERO; SILVA, 2020), que existe uma relação direta entre o IDH-M e o número de casos de COVID-19 em uma determinada cidade, mas ele sozinho, não possui representatividade suficiente para justificar as ocorrências em cidades cujo IDH-M é maior que a média. Pode-se levar em consideração fatores como decisões governamentais, presença de centros comerciais, vias de alta rotatividade, dentre outros.

A partir deste trabalho, abre-se portas para discussões acerca de quais parâmetros são de fato os responsáveis pelos números de casos de COVID-19 nas cidades mencionadas. Essa abordagem traz uma oportunidade de criação de novos estudos mais amplos considerando o Brasil como um todo em suas análises. Difundir trabalhos desta natureza são de fundamental importância para que os dias da pandemia fiquem mais próximos do fim.

Referências

- BRASIL, M. d. S. *Sobre a doença*. 2020. Disponível em: <<https://coronavirus.saude.gov.br/sobre-a-doenca>>. Citado na página 2.
- IBGE. *Sinopse do Censo Demográfico - Minas Gerais*. 2010. Disponível em: <<https://censo2010.ibge.gov.br/sinopse/index.php?uf=31>>. Acesso em: 2020-12-04. Citado 2 vezes nas páginas 5 e 6.
- ROMERO, J. A. R.; SILVA, F. A. M. da. Relação entre as condições socioeconômicas e a incidência da pandemia da covid-19 nos municípios do ceará. *Boletim de Conjuntura (BOCA)*, v. 3, n. 7, p. 85–95, 2020. Citado 3 vezes nas páginas 2, 3 e 16.
- VEJA, R. *OMS decreta pandemia do novo coronavírus. Saiba o que isso significa*. 2020. Disponível em: <<https://saude.abril.com.br/medicina/oms-decreta-pandemia-do-novo-coronavirus-saiba-o-que-isso-significa>>. Citado na página 2.