

Universidade Tecnológica Federal do Paraná
Departamento Acadêmico de Informática
Curso de Bacharelado em Sistemas de Informação

Relatório

Análise Exploratória

Bruno Wandrey
Daniel Vieira
Eduardo Darrazão

Curitiba
Abril, 2022

Universidade Tecnológica Federal do Paraná
Departamento Acadêmico de Informática
Curso de Bacharelado em Sistemas de Informação

Relatório

Análise Exploratória

Relatório referente à Disciplina Introdução à Ciência de Dados do Curso de Bacharelado de Sistemas de Informação da Universidade Tecnológica Federal do Paraná (UTFPR).

Curitiba
Abril, 2022

Sumário

1	Introdução	3
1.1	Tema e Equipe	3
1.2	Obtenção e processamento de dados	3
1.3	Procedimentos de limpeza e processamento	4
1.4	Ponderação do número de mortes	5
2	Cobertura e distribuição dos dados	7
2.1	Distribuição da População por faixa etária	7
2.2	Número de mortes por população	8
2.3	Políticas Públicas	12
2.3.1	Variação dos indicadores semanalmente nos anos de 2020 e 2021	13
3	Perguntas de pesquisa e explorações iniciais	16
3.1	Relação de mortes com as políticas públicas	17
3.1.1	Correlações das Mortes com os Indicadores	19
3.2	Metodologia de Análise	21
3.3	C1 School closing	21
3.4	C2 Workplace Closing	21
3.5	C3 Cancel public events	23
3.6	C4 Restrictions on gatherings	23
3.7	C5 Close public transport	23
3.8	C6 Stay at home requirements	24
3.9	C7 Restrictions on internal movement	24
3.10	C8 International travel controls	25
3.11	H1 Public information Campaigns	25
4	Discussão sobre os resultados	26
	Referências	30

1 Introdução

1.1 Tema e Equipe

A equipe *Dubruda* possuí três integrantes, todos cursando Bacharelado em Sistemas de Informação na UTFPR-Curitiba. Nossa trabalho tem como objetivo medir o impacto de políticas públicas na mortalidade por *COVID-19* nas capitais brasileiras e na maior cidade de cada estado além da capital.

1.2 Obtenção e processamento de dados

Os dados principais obtidos são do repositório *Brazil-covid-policy*^{*}. Este dataset é baseado no dataset *Oxford Covid-19 Government Response Tracker*[†]. O projeto examina as políticas de resposta à COVID-19 adotadas pelos governos federal, estaduais e municipais em mais de 50 cidades brasileiras, com base no sistema de codificação do Oxford COVID-19 Government Response Tracker, adaptado ao contexto brasileiro.

O dataset é um conjunto de indicadores sobre políticas públicas implantadas durante a pandemia de COVID-19, nos anos de 2020, 2021 e 2022. No dataset *Brazil-covid-policy*, os dados são indexados por dia, cidade, estado e jurisdição. Para políticas a níveis estaduais e nacionais, são utilizadas as jurisdições *STATE_* e *NAT_*, entretanto, como analisaremos os impactos a nível municipal, utilizaremos a jurisdição *CITY_TOTAL*, que expõe o maior nível de política implantada naquele município. Por exemplo, se o município tiver uma política de fechamento de escolas parcial, mas o governo estadual implantar o fechamento total das escolas, o nível de política exposto pelo indicador será o correspondente ao governo estadual.

Usamos principalmente os seguintes indicadores: *C1 School closing*, *C2 Workplace closing*, *C3 Cancel public events*, *C4 Restrictions on gatherings*, *C5 Close public transport*, *C6 Stay at home requirements*, *C7 Restrictions on internal movement*, *C8 International travel controls* e *H1 Public information campaigns*. Cada um desses indicadores funciona de forma similar, onde o valor zero corresponde a não implantação da política, e o maior valor corresponde a implantação total da política. Cada indicador é melhor descrito na seção 3, ou no codebook do repositório *Brazil-covid-policy*.

O dataset também contém dados referentes a indicadores de políticas econômicas, políticas do sistema de saúde e sobre vacinação. Contudo parte destes conjuntos de dados

^{*}disponível em <https://github.com/OxCGR/Brazil-covid-policy>

[†]Disponível em <https://github.com/OxCGR/covid-policy-tracker>

estão incompletos, ou não estão devidamente implementados. Vale ressaltar que no codebook do dataset brasileiro apenas trás informações dos indicadores citados inicialmente, e para os demais é necessário conferir no codebook da Oxford no qual foi baseado.

Usamos dados da mortalidade por Covid-19 oriundos da *National Center for Health Statistics*[†] para gerar uma distribuição estatística dos números brutos de mortes para projetar o perfil das mortes em cada faixa etária. Consideramos que a doença produziu um efeito na população brasileira similar ao efeito produzido na população estadunidense, em relação a taxa de mortalidade por idade. Para realizar a adaptação dos dados para o nosso caso, foi necessário criar uma relação dos dados que já possuímos com os dados demográficos do Brasil, obtivemos parte desses dados na página do IBGE e parte foi disponibilizada pelo orientador, sendo essa parte obtida através do portal DataSUS.

1.3 Procedimentos de limpeza e processamento

- *Brazil-covid-policy* dataset: Mantida somente as linhas onde a coluna 'Jurisdiction' era igual a 'CITY_TOTAL', que representa o maior nível de política implantada em cada cidade (vide segundo parágrafo do capítulo 2).
- *Brazil-covid-policy* dataset: Mantida somente as linhas onde a coluna 'CityCode' não era nula, removendo assim entradas referentes ao Governo Federal e aos estados, restando somente dados das cidades.
- *Brazil-covid-policy* dataset: Depois das devidas remoções de colunas que não serão utilizadas, restaram as seguintes que serão úteis: 'RegionName' (nome do estado), 'CityName' (nome da cidade), 'Date' (dia/mês/ano de cada dado), todos os indicadores (variam entre C1-8, E1-4, H1-8, V1-4), 'ConfirmedDeaths' (número total de mortes acumuladas) e 'StringencyIndex' (Índice de Rigor, representa a média simples dos 9 indicadores principais: C1-8 e H1).
- *Brazil-covid-policy* dataset: Há diversos campos nulos nos indicadores, mas como o fato de não existir dados para uma determinada cidade em uma determinada data para um conjunto de indicadores pode nos ser útil, todos as entradas nulas nas colunas de indicadores foram substituídas por -1. Este processo tem grande importância quando dividimos o dataset inicial (que possui frequência de dados diárias) em datasets de frequências semanais e anuais. Este tratamento de dividir por diferentes frequências captura a moda de cada indicador durante o respectivo período, onde o -1 captura quando não houve dados. Vale ressaltar que para cálculos como correlações durante um certo período os campos afetados são convertidos de volta para 'NaN' (nulos) para não causar erros nos resultados.

[†]Disponível em: <https://data.cdc.gov/d/9bhg-hcku>

- *Brazil-covid-policy* dataset: Para analisar as mortes por COVID-19 de forma temporal, precisamos construir uma nova coluna denominada 'NewDeaths', devido ao fato que as mortes (coluna 'ConfirmedDeaths') são apresentadas de forma cumulativa por cidade. Logo esta coluna que foi criada representa a diferença de 'ConfirmedDeaths' do (Dia,Cidade) que está sendo calculada, com a 'ConfirmedDeaths' desta cidade no dia anterior.

1.4 Ponderação do número de mortes

A ponderação (figura 1) do número de mortes foi feita considerando um ajuste por idade e tamanho da população, esse é um processo estatístico aplicado a taxas de mortalidade que permite comparar comunidades com diferentes estruturas etárias. Uma vez que quase todas as doenças possuem mortalidade diferentes em diferentes faixas etárias. Por exemplo, a maioria das doenças crônicas, incluindo a maioria dos cânceres, ocorre com mais frequência entre os idosos. A distribuição etária determina quais serão os problemas de saúde mais comuns em uma comunidade. Uma maneira de examinar os resultados em comunidades de diferentes tamanhos é calcular uma taxa de incidência ou mortalidade, que é o número de novos casos ou mortes dividido pelo tamanho da população. Em doenças crônicas e lesões, as taxas são geralmente expressas em termos de número de casos/óbitos por 100.000 pessoas por ano ([HEALTH, 1999](#)).

$$\left(\frac{\text{Number of deaths of the age group} \times 100.000}{\text{Population}} \right) \times \left(\frac{\text{Number of persons of the age group}}{\text{Population}} \right)$$

Figura 1 – Forma de ponderação usada no trabalho.

Usamos a distribuição de população por faixa etária no Brasil para estimar quantas pessoas morreram em cada uma das categorias, porque acreditamos que a distribuição de morte por faixa etária observada nos Estados Unidos ([STATISTICS, 2020](#)) poderia ser usada para projetar a distribuição de morte por faixa etária no Brasil, quando adaptássemos os tamanhos dos grupos em cada faixa etária para os valores brasileiros, ou seja, nós mantivemos a proporção de mortes por grupo e quando alteramos o tamanho do grupo para os valores brasileiros obtivemos a projeção para o Brasil.

	Grupo de risco								
	0 a 4	5 a 14	15 a 24	25 a 34	35 a 44	45 a 54	55 a 64	65 a 74	>=75
Média	6,86%	14,48%	16,60%	16,78%	15,91%	12,04%	8,96%	5,29%	3,08%
Média Acumulada	54,72%				27,94%	17,34%			
DP	1,18	2,25	1,90	1,21	1,01	1,28	1,77	1,45	1,08
Mortes	0,04%	0,03%	0,26%	1,11%	2,77%	6,68%	14,70%	23,00%	51,41%
Mortes Acumuladas	1,45%				9,45%	89,11%			
Mínimo	5,18%	10,88%	13,22%	14,48%	12,53%	8,61%	5,15%	2,56%	1,15%
25%	6,09%	12,83%	14,78%	16,10%	15,60%	11,65%	7,82%	4,43%	2,44%
Mediana	6,51%	14,30%	16,34%	16,70%	16,03%	12,46%	9,14%	5,31%	3,00%
75%	7,48%	15,44%	18,03%	17,23%	16,45%	12,87%	10,24%	6,06%	3,70%
Máximo	10,74%	20,86%	20,95%	21,31%	18,07%	13,59%	11,95%	8,68%	5,98%

Figura 2 – Tabela com resumo das distribuições de população por faixa etária do Brasil e a distribuição dos Estados Unidos de mortes por faixa etária.

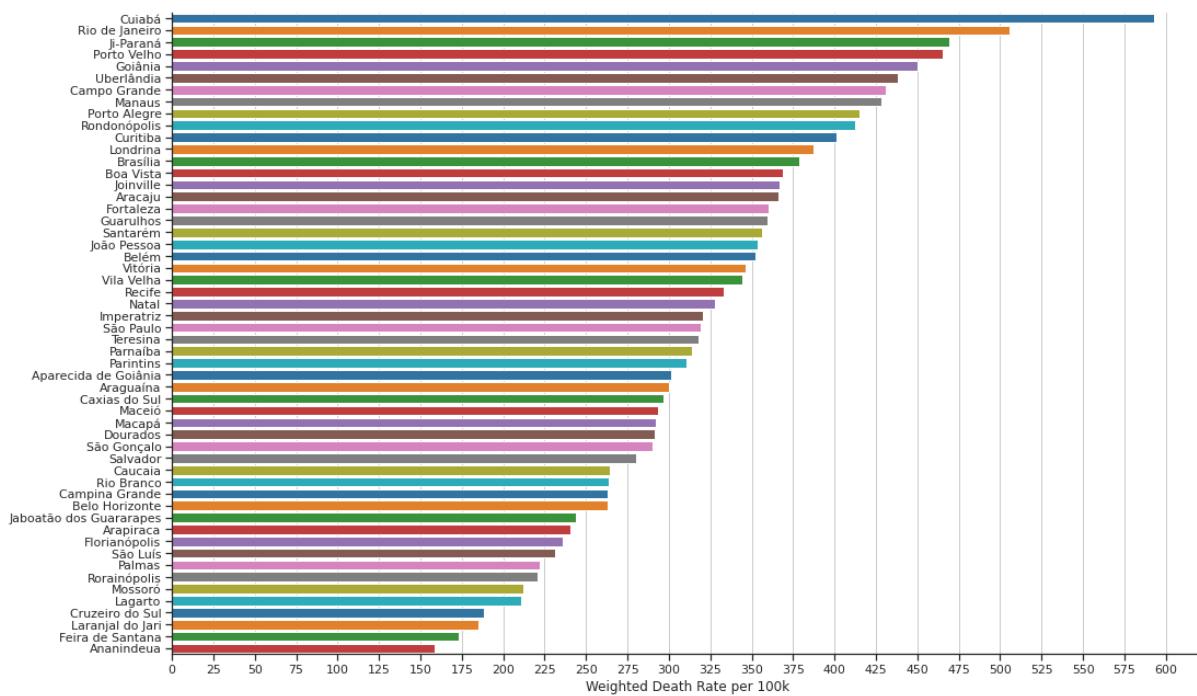


Figura 3 – Cidades avaliadas ordenadas por número ponderado de mortes.

2 Cobertura e distribuição dos dados

2.1 Distribuição da População por faixa etária

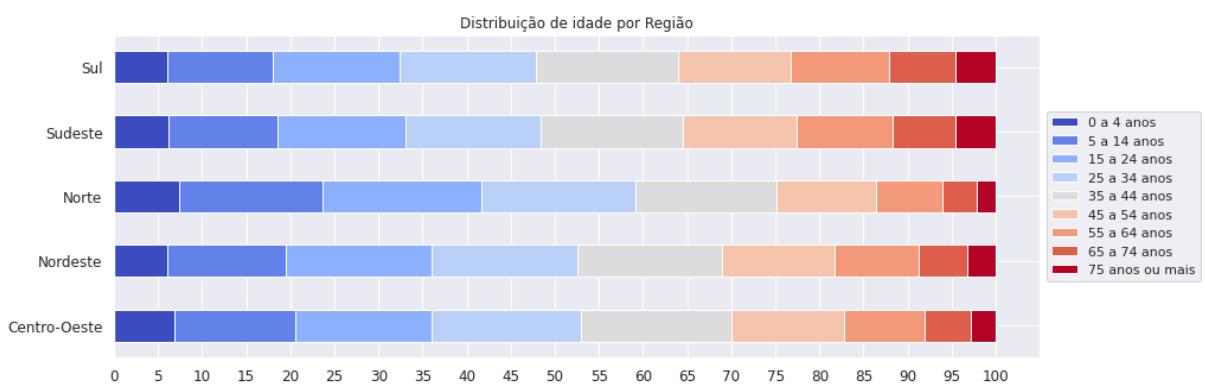


Figura 4 – Distribuição percentual das faixas etárias brasileiras selecionadas como categorias por região.

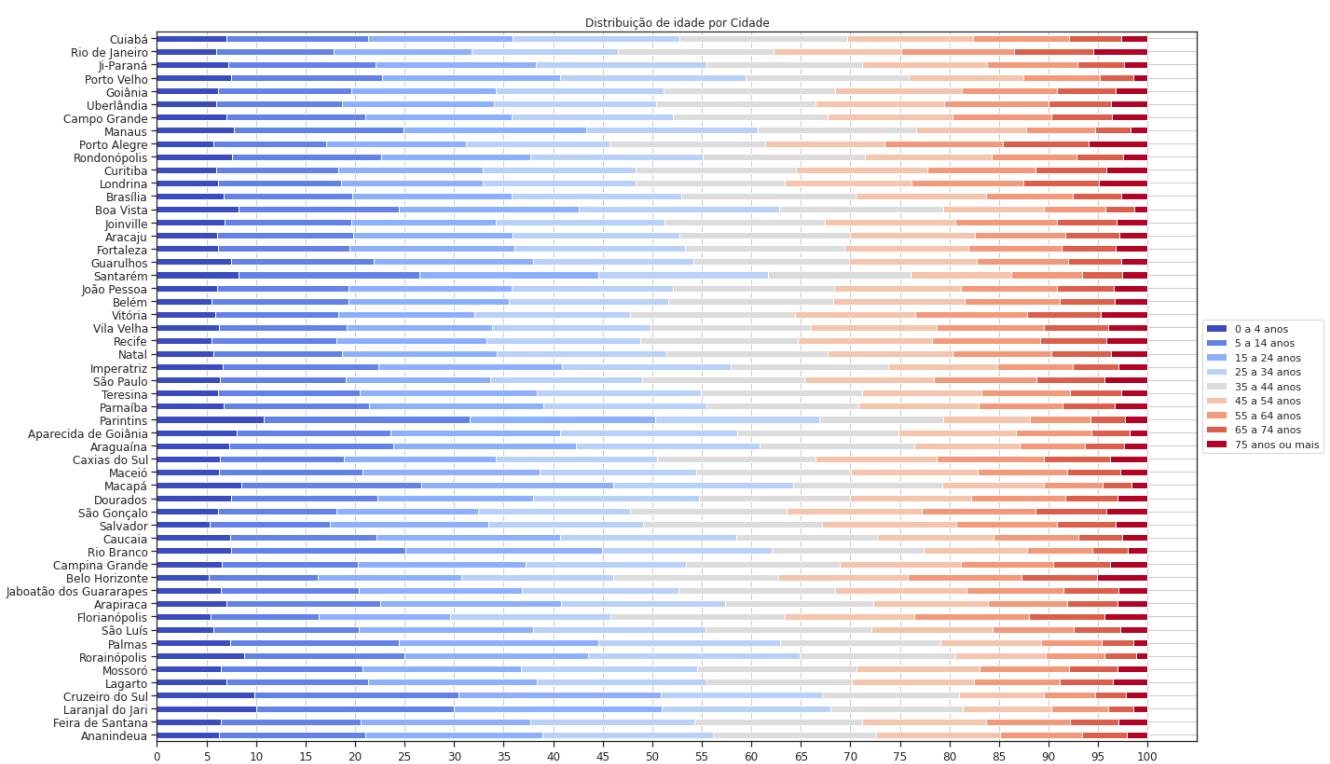


Figura 5 – Distribuição percentual das faixas etárias brasileiras selecionadas como categorias por cidade. Ordenadas por número de mortes.

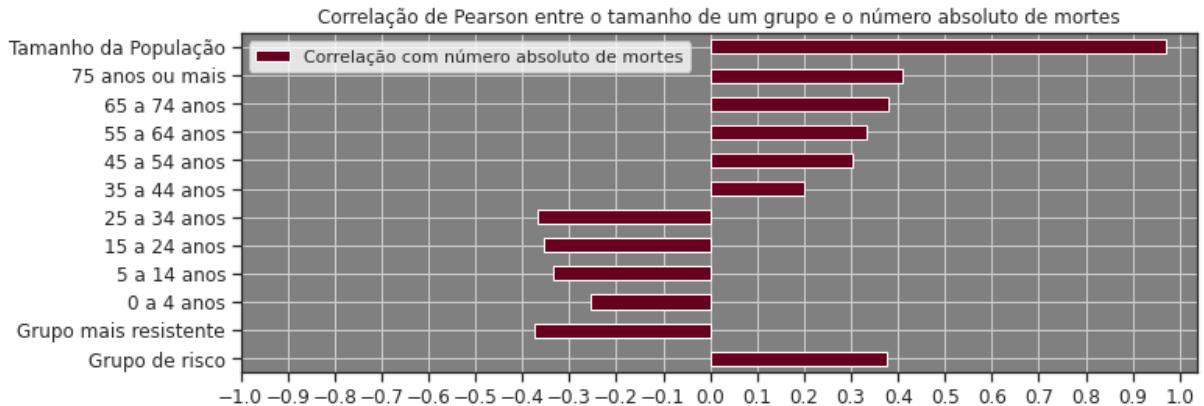


Figura 6 – Correlação entre número de mortes e grupos.



Figura 7 – Correlação entre número ponderado de mortes e grupos.

2.2 Número de mortes por população

Nas figuras 6 e 7 pode-se observar a importância de ponderação para comparar padrões entre resultados, o número de habitantes de uma cidade é um fator que determina muito fortemente o número de mortes, entretanto quando comparamos queremos estudar outros aspectos além do tamanho da população, dessa forma, a ponderação diminui o grau de influência do tamanho absoluto da população sobre o número de mortes. É possível perceber analisando o caso da região nordeste que apresentou um número de mortes, em geral, abaixo da média mesmo tendo a segunda maior população dentre as regiões avaliadas (figura 8).

O grupo de pessoas entre 15 e 24 anos é o mais decisivo individualmente para diminuir o número de mortes, entretanto a principal conclusão que podemos observar das figuras 6 e 7 é que a mortalidade por Covid-19 é altamente relacionada a idade da população, sendo muito mais inofensiva em grupos mais jovens e muito mais agressiva em grupos mais idosos.

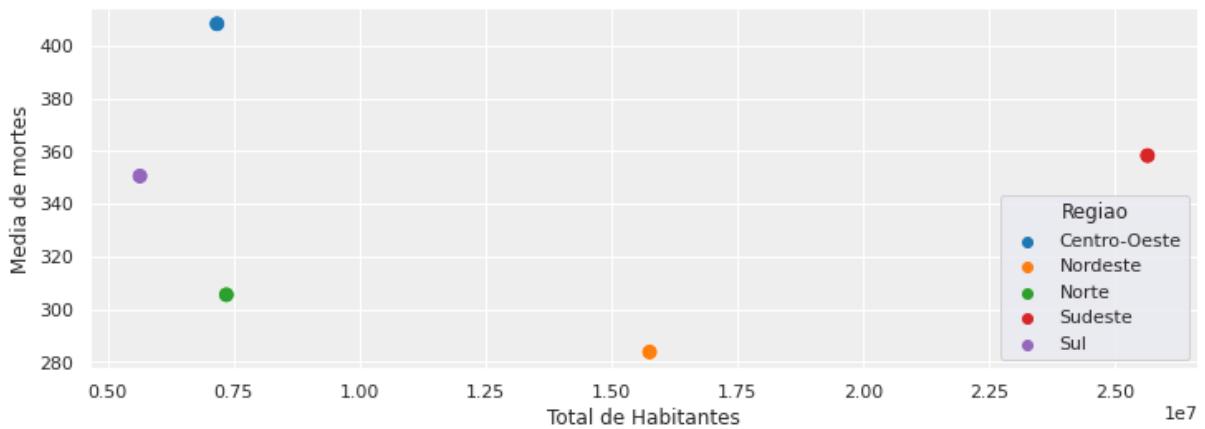


Figura 8 – Relação entre número de habitantes das regiões avaliadas e o número de mortes.

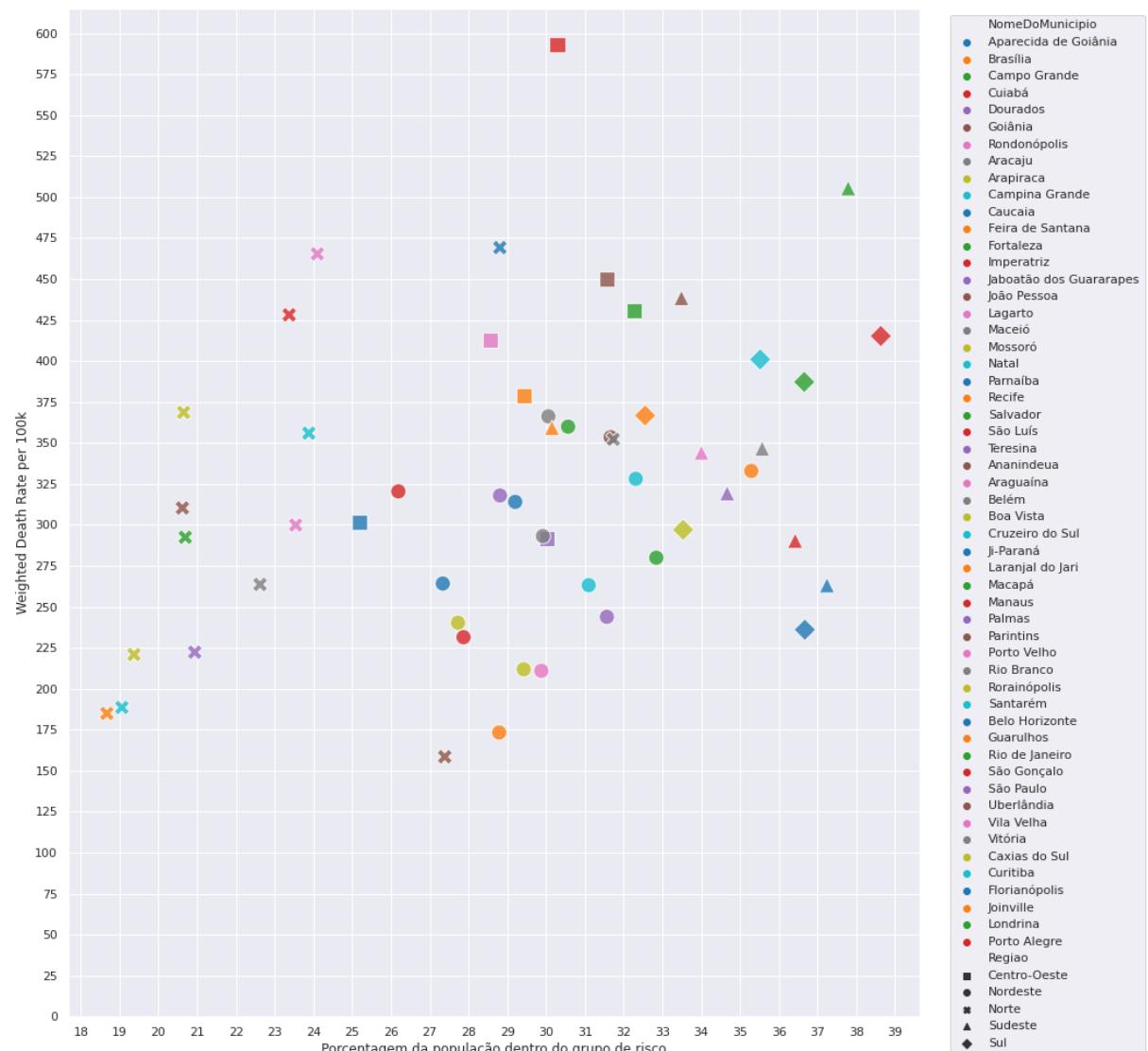


Figura 9 – Visualização da relação entre o tamanho do grupo de risco e o número de mortes de cada cidade.

Mortes por Covid-19 pelo Brasil

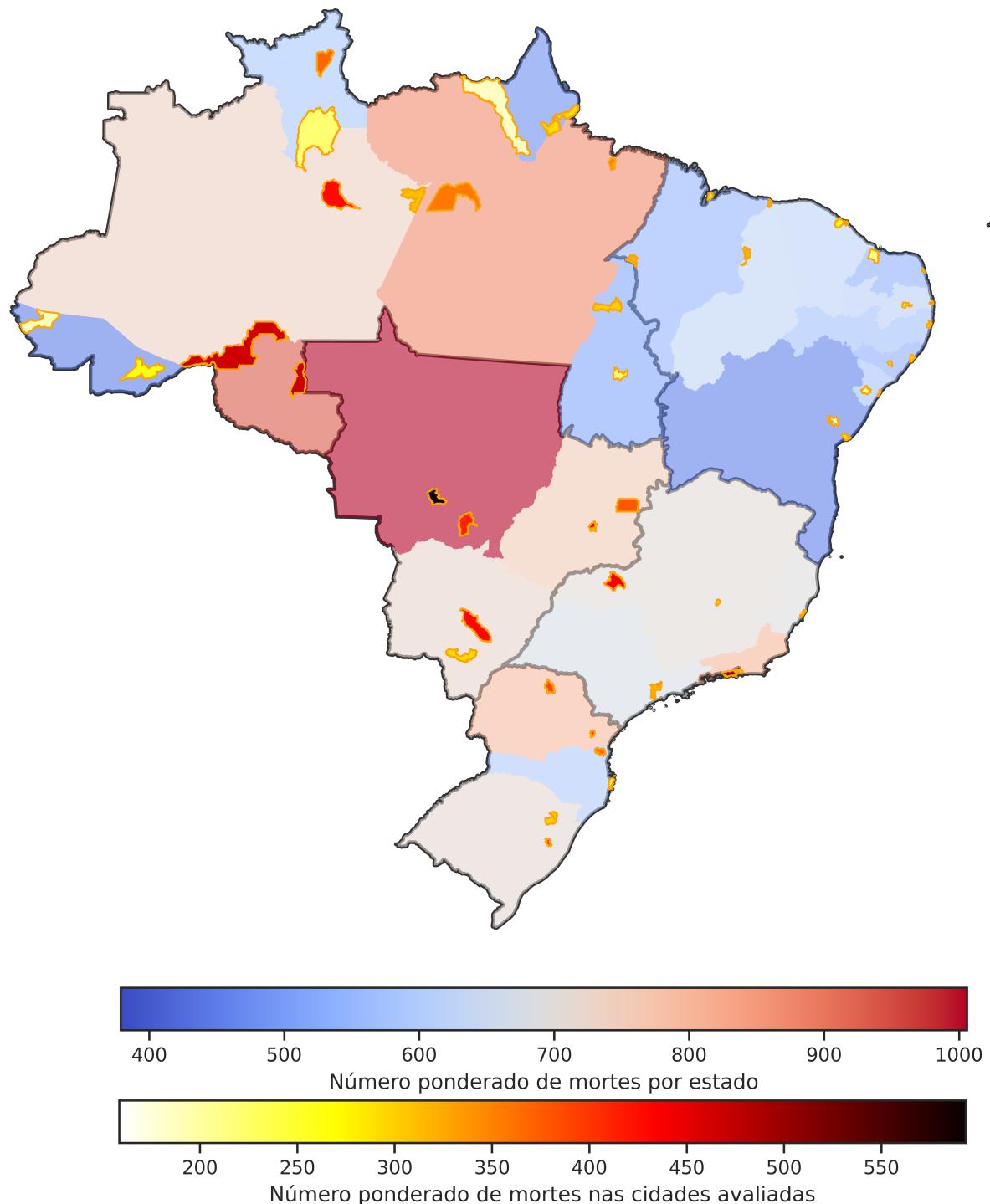


Figura 10 – Mortes no Brasil por estado e por cidade avaliada. O contorno em laranja das cidades avaliadas no *dataset* representa a média da escala de baixo, permitindo uma comparação rápida. Para os estados, é projetada a média das cidades avaliadas respectivamente. A linha em cinza translúcido divide as regiões do Brasil.

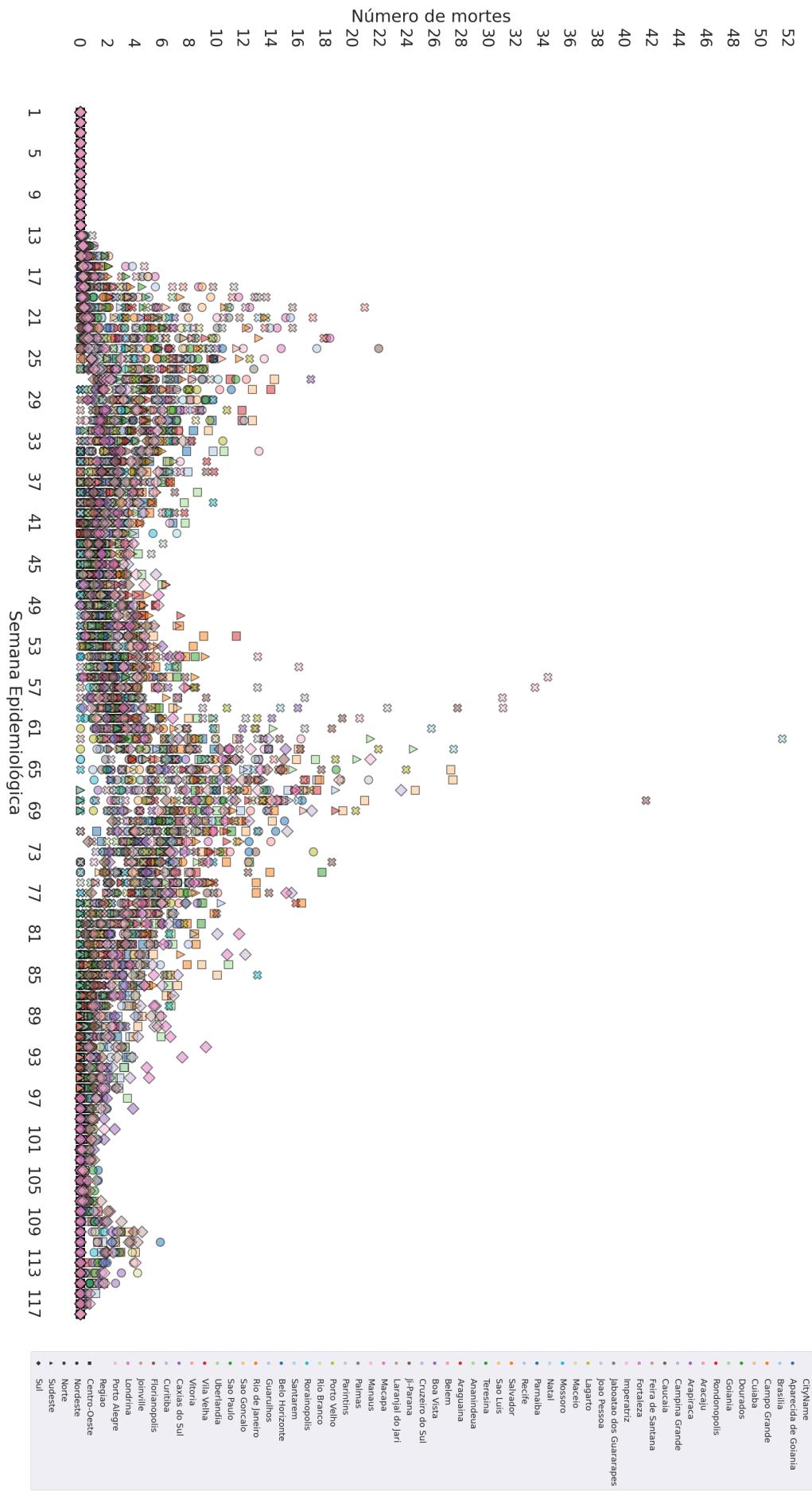


Figura 11 – Caption

2.3 Políticas Públicas

Inicialmente precisávamos saber quais são as cidades presentes no dataset, que foram encontradas visualizando os campos únicos na coluna 'CityName'. Encontramos 54 cidades, que representam a capital de cada estado e a maior cidade que não seja a capital.

São elas: Ananindeua, Aparecida de Goiânia, Aracaju, Araguaína, Arapiraca, Belém, Belo Horizonte, Boa Vista, Brasília, Campina Grande, Campo Grande, Caucaia, Caxias do Sul, Cruzeiro do Sul, Cuiabá, Curitiba, Dourados, Feira de Santana, Florianópolis, Fortaleza, Goiânia, Guarulhos, Imperatriz, Jaboatão dos Guararapes, Ji-Paraná, João Pessoa, Joinville, Lagarto, Laranjal do Jari, Londrina, Macapá, Maceió, Manaus, Mossoró, Natal, Palmas, Parintins, Parnaíba, Porto Alegre, Porto Velho, Recife, Rio Branco, Rio de Janeiro, Rondonópolis, Rorainópolis, Salvador, Santarém, São Gonçalo, São Luís, São Paulo, Teresina, Uberlândia, Vila Velha e Vitoria.

Após os processos de limpeza e processamento, descobrimos que os principais indicadores que pretendemos analisar possuem uma grande quantidade de dados.

Tabela 1 – Campos e quantidade de dados não vazios, no dataset filtrado por jurisdição
CITY_TOTAL

Campo	# de dados	Campo	# de dados
CountryName	44334	C1 School closing	40872
CountryCode	44334	C2 Workplace closing	40870
RegionName	44334	C3 Cancel public events	40869
RegionCode	44334	C4 Restrictions on gatherings	40839
CityName	44334	C5 Close public transport	40844
CityCode	44334	C6 Stay at home requirements	40844
Jurisdiction	44334	C7 Restrictions on internal movement	40869
Date	44334	C8 International travel controls	40861
H1 Public info. campaigns	40873	H2 Testing policy	40950
H3 Contact tracing	40828	H4 Emergency invest. in healthcare	411
H5 Investment in vaccines	468	H6 Facial Coverings	40267
H7 Vaccination policy	40474	H8 Protection of elderly people	38783
E1 Income support	40460	E2 Debt/contract relief	40160
E3 Fiscal measures	417	E4 International support	0
V1,V2,V3,V4	0	ConfirmedDeaths	30928
StringencyIndex	40865		

Tabela 2 – Indicadores, suas variâncias e possíveis valores, no dataset filtrado por jurisdição *CITY_TOTAL*

Indicador	Variância	Faixa de valores (inteiros)
C1 School closing	0.98	0-3
C2 Workplace closing	0.74	0-3
C3 Cancel public events	0.48	0-2
C4 Restrictions on gatherings	2.46	0-4
C5 Close public transport	0.36	0-2
C6 Stay at home requirements	0.51	0-3
C7 Restrictions on internal movement	0.71	0-2
C8 International travel controls	1.63	0-4
H1 Public information campaigns	0.12	0-3

2.3.1 Variação dos indicadores semanalmente nos anos de 2020 e 2021

Para analisar como os indicadores variaram semanalmente neste período, capturamos a moda semanal do indicador 'StringencyIndex' que representa a média simples dos nove indicadores principais que estamos estudando, C1 a C8 e H1. Isso é descrito na equação abaixo, onde k é o número de indicadores componentes em um índice e I_j é a pontuação do subíndice para um indicador individual.

$$index = \frac{1}{k} \sum_{j=1}^k I_j$$

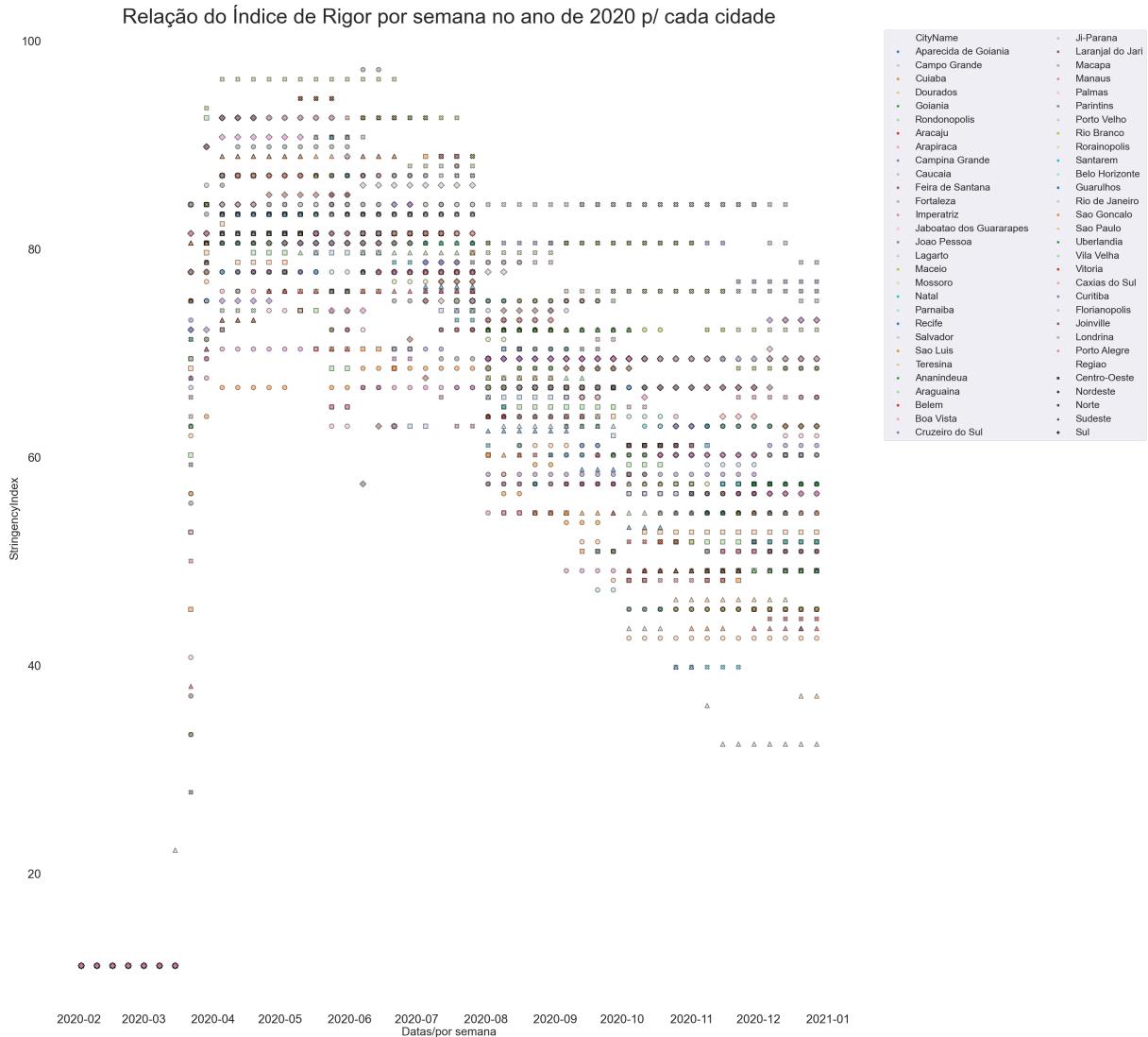


Figura 12 – Podemos perceber aqui como os indicadores variaram desde o início da pandemia até o final do ano de 2020. Por exemplo, percebemos que no início cidades do Norte do Brasil já começaram bem rigorosas, e a maioria se manteve assim. Já cidades do Nordeste foram bem descrepantes, algumas sendo bem rigorosas e outras pouco.

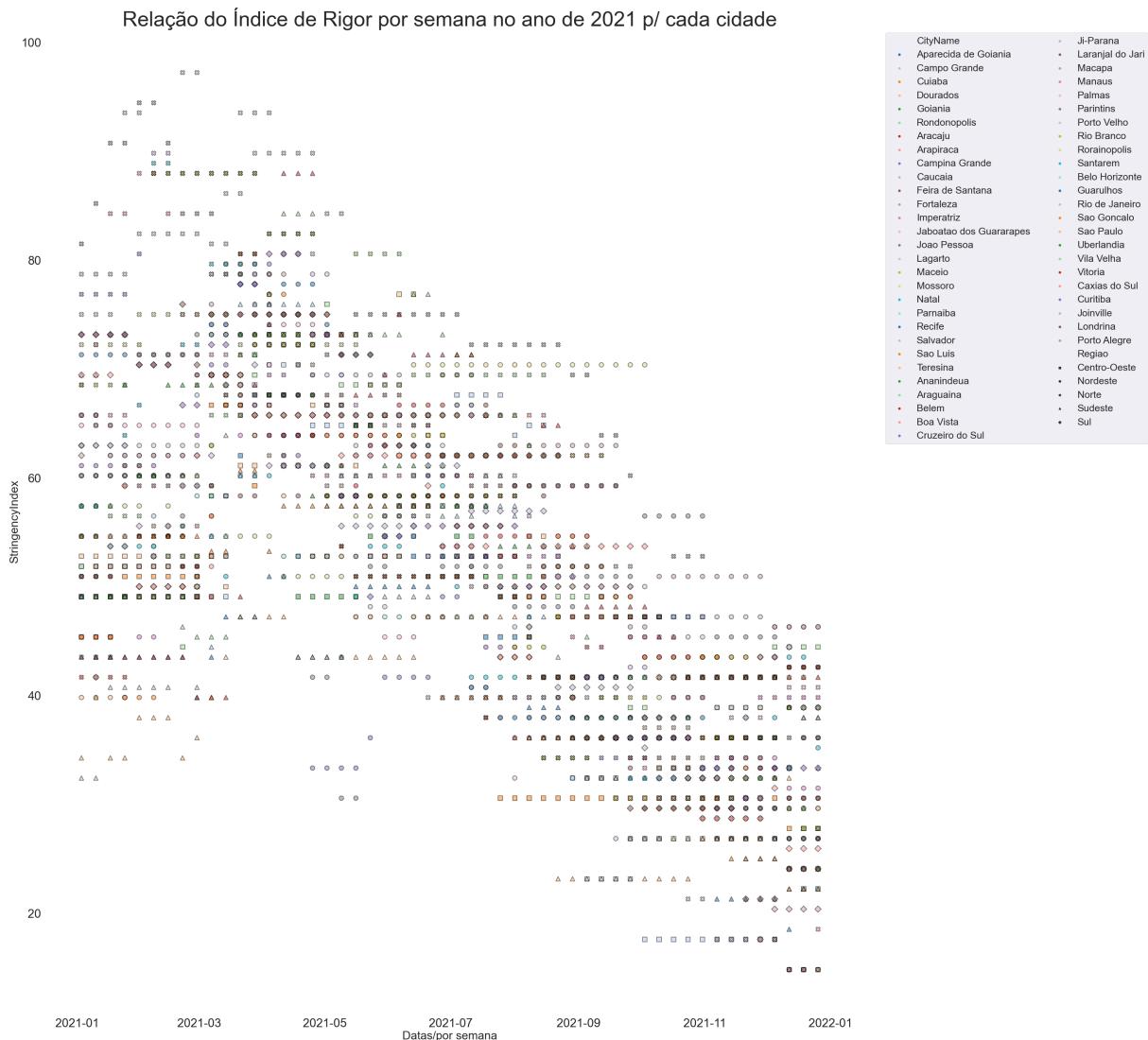


Figura 13 – No ano de 2021 houveram muito mais mortes devido a pandemia, e é nítido que o Norte se manteve rigoroso quanto as políticas. No geral o Sudeste foi condescendente, com as médias mais baixas de rigor. As cidades do Nordeste se mantiveram discrepantes, umas relativamente rigorosas e outras com políticas bem brandas.

3 Perguntas de pesquisa e explorações iniciais

O objetivo da análise é encontrar possíveis indicadores de como as políticas públicas sobre COVID-19 afetam a quantidade de óbitos em razão do mesmo. Portanto, desenvolvemos as seguintes questões para análise do dataset

Tabela 3 – Indicadores de políticas públicas e Perguntas de pesquisa

Indicador de política pública	Pergunta de pesquisa
C1 School closing	Como o fechamento de escolas afeta a quantidade de óbitos causados pelo COVID-19?
C2 Workplace closing	Como o fechamento de ambiente de trabalhos afeta a quantidade de óbitos causados pelo COVID-19?
C3 Cancel public events	Como o cancelamento de eventos públicos afeta a quantidade de óbitos causados pelo COVID-19?
C4 Restrictions on gatherings	Como a imposição de restrições de aglomerações afeta a quantidade de óbitos causados pelo COVID-19?
C5 Close public transport	Como o fechamento de transporte público afeta a quantidade de óbitos causados pelo COVID-19?
C6 Stay at home requirements	Como a imposição de quarentena afeta a quantidade de óbitos causados pelo COVID-19?
C7 Restrictions on internal movement	Como as restrições em movimentos internos afeta a quantidade de óbitos causados pelo COVID-19?
C8 International travel controls	Como o controle de viagens internacionais afeta a quantidade de óbitos causados pelo COVID-19?
H1 Public information campaigns	Como campanhas de conscientização afetam a quantidade de óbitos causados pelo COVID-19?

3.1 Relação de mortes com as políticas públicas

Para melhor compreender como as mortes se relacionam com as políticas públicas, selecionamos as modas do Índice de Rigor (StringencyIndex) de cada cidade nos anos de 2020 e 2021, e para correlacioná-las calculamos o número de novas mortes ponderadas (por 100mil habitantes e faixa etária) nestes anos.

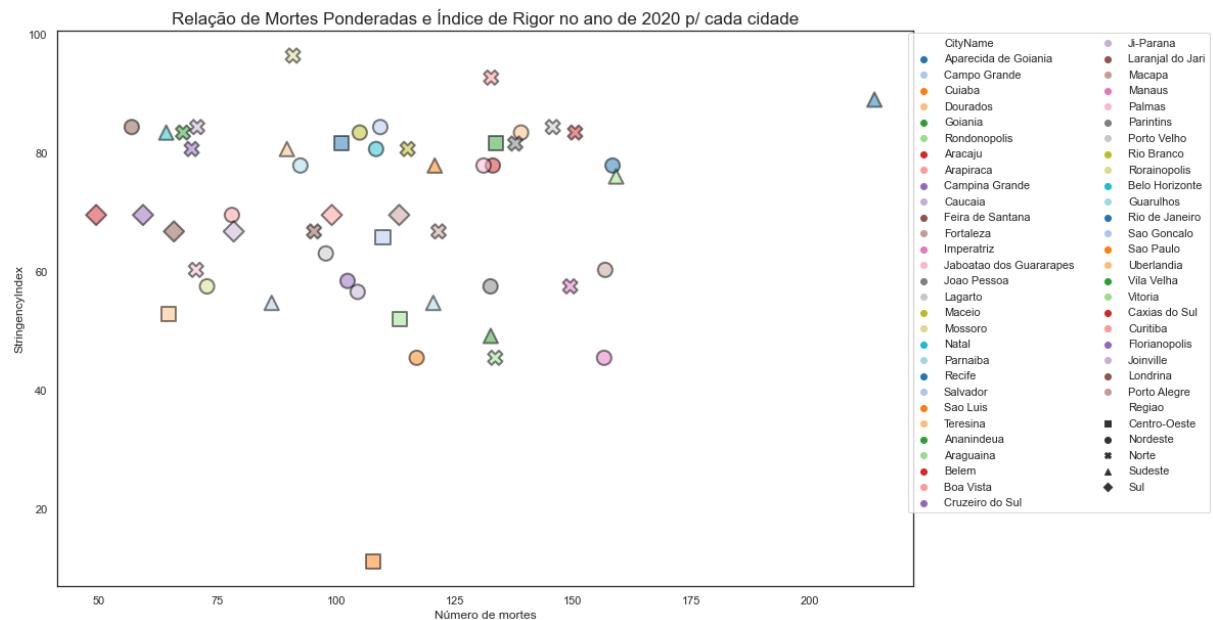


Figura 14 – Nota-se que neste primeiro ano de pandemia a cidade que foi menos rigorosa disparadamente foi Cuiabá, Centro-Oeste, contudo não teve muitas mortes. Rio de Janeiro, Sudeste, apesar de ter sido bem rigorosa morreu muitas pessoas. É possível perceber também que na parcela de cidades que morreram menos pessoas (0-80) há uma tendência a uma maior concentração na parte superior do eixo Y, ou seja, mais concentrado em cidades mais rigorosas.

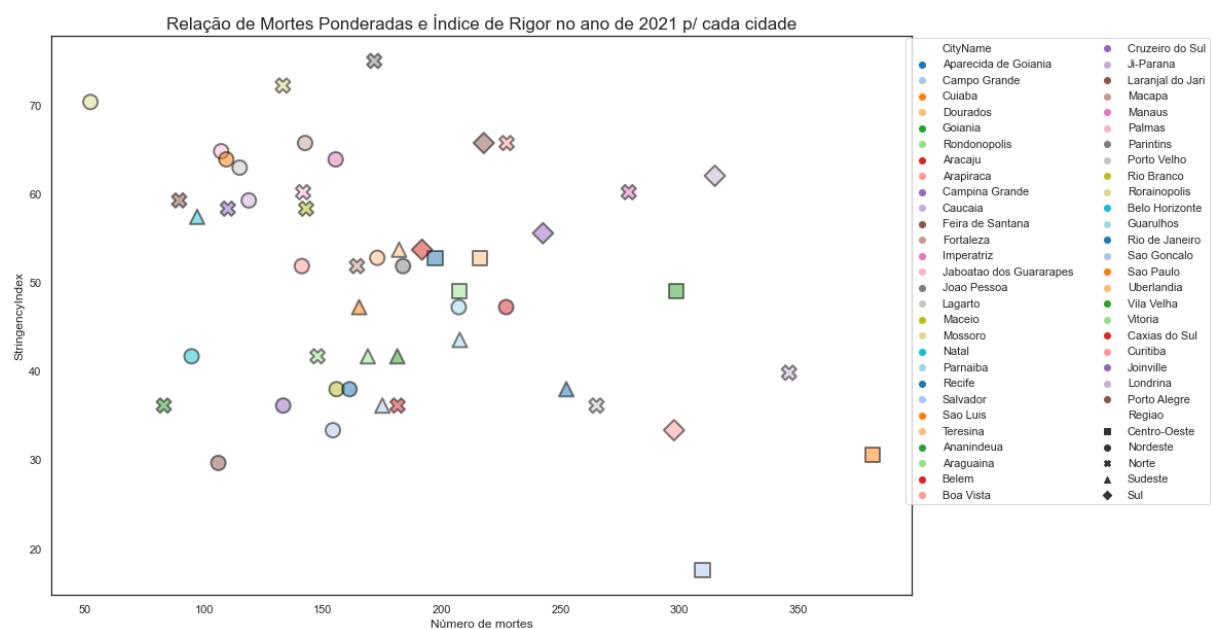


Figura 15 – No ano de 2021, onde houve mais óbitos por COVID-19, a possível correlação entre cidades mais rigorosas tiveram menos mortes fica mais evidente. No geral as cidades que mais morreram pessoas foram bem lenientes no rigor das políticas públicas, enquanto as mais rigorosas no geral se saíram bem melhor.

3.1.1 Correlações das Mortes com os Indicadores

Correlacionamos a moda de cada indicador e a quantidade de novas mortes ponderadas nos anos de 2020 e 2021. É interessante salientar que *possivelmente* as correlações dos indicadores com a quantidade de mortes são correlações espúrias, isto é, uma relação estatística existente entre duas variáveis, mas onde não existe nenhuma relação causa-efeito entre elas. Essa correlação espúria não necessariamente se refere a todos os indicadores, é algo que ainda será estudado.

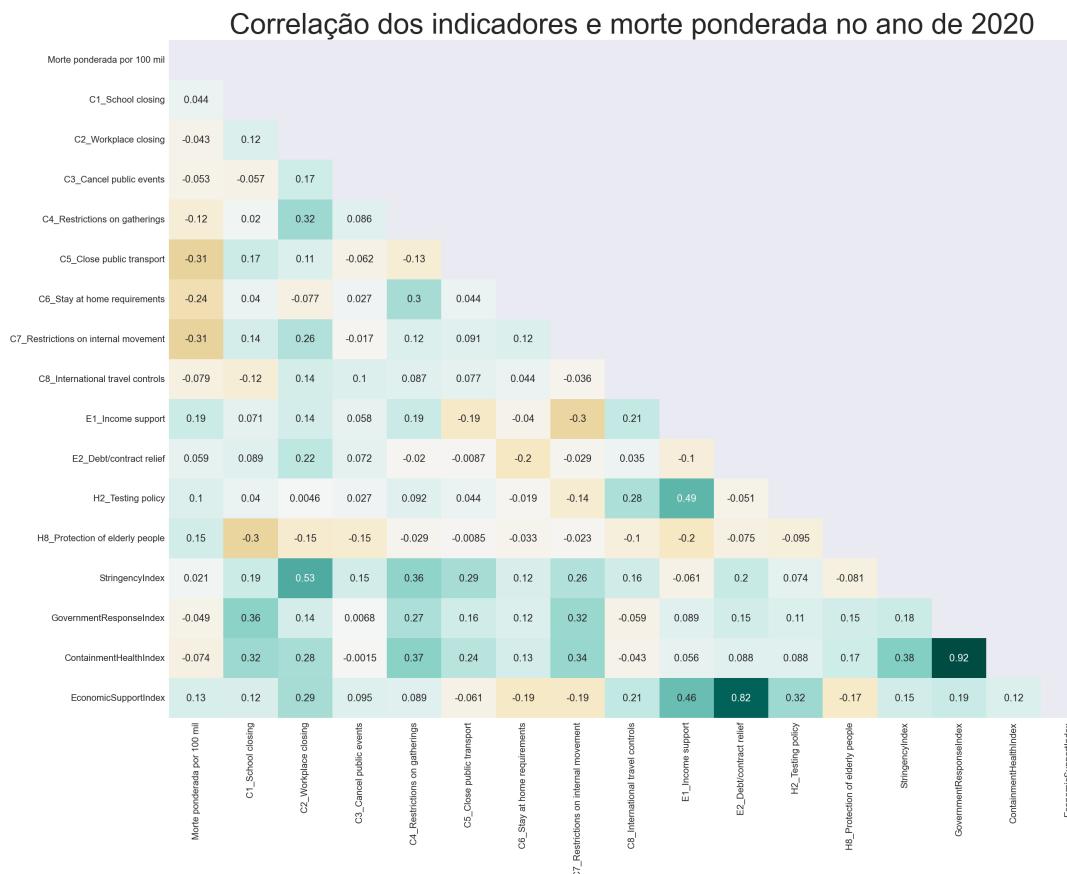


Figura 16 – No primeiro ano da pandemia *aparentemente* as políticas que tiveram um maior impacto na redução da mortalidade por COVID-19 foram: fechamento dos transportes públicos, restrições na movimentação entre cidades/estados e recomendações de não sair de casa.

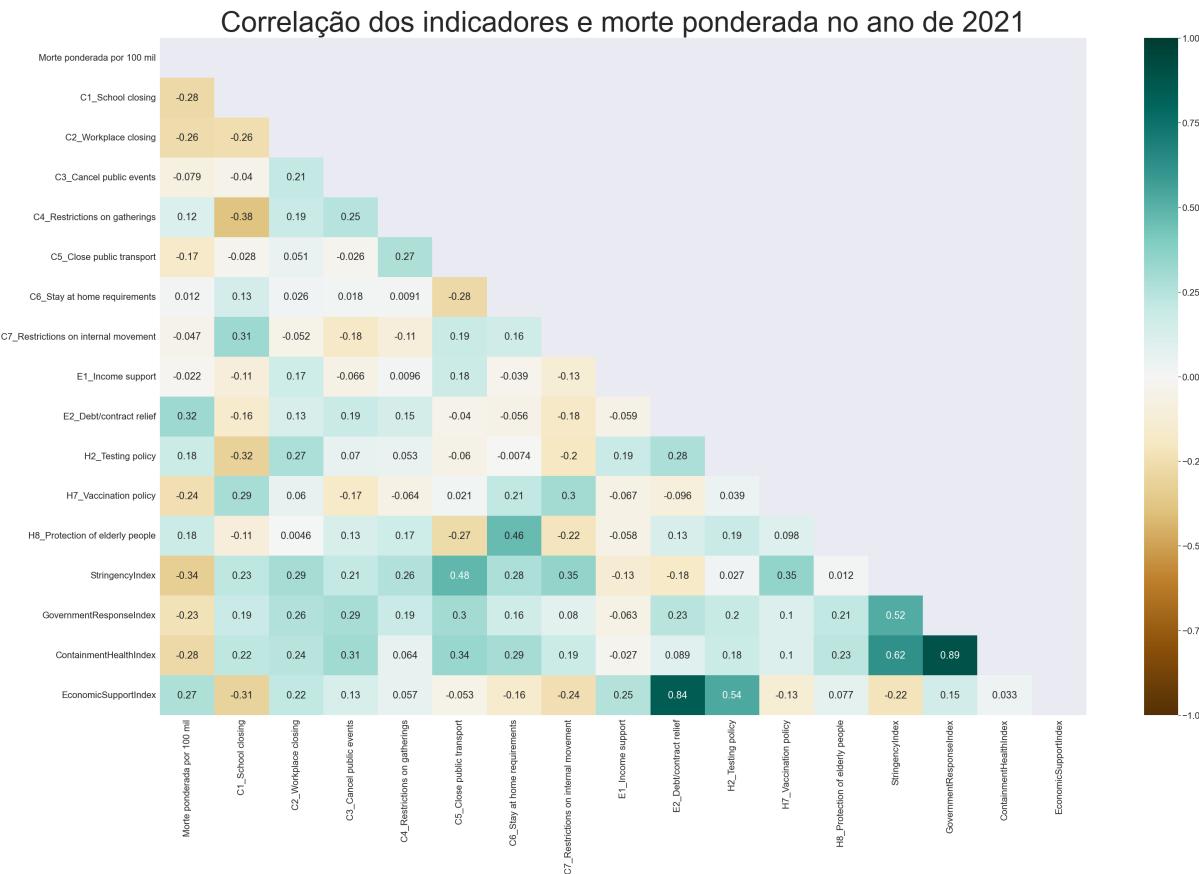


Figura 17 – No segundo ano as políticas que *aparentemente* foram eficazes no primeiro já não são mais, entretanto outras se mostraram relevantes: fechamento de escolas, fechamento de lugares de trabalho e políticas de vacinação. O índice de rigor (StringencyIndex) também se mostrou relevante, lembrando que o mesmo diz respeito aos 9 indicadores principais (C1-8,H1). Outros índices também se mostraram potencialmente relevantes, entretanto não estamos nos aprofundando nos indicadores que eles representam. Para uma compreensão do que cada índice representa é possível encontrar no projeto *Oxford COVID-19 Government Response Tracker (index methodology)*.

3.2 Metodologia de Análise

Utilizamos três conjuntos de valores para correlacionar com o valor escalar de cada indicador de política pública: 1. O número de novas mortes diárias por COVID-19 de cada município, por município, 2. O número de casos de COVID-19 totais, por município, e 3. O número de mortes totais por COVID-19 de cada município

Em cada uma dessas categorias, foi utilizado uma média móvel com intuito de apaziguar as discrepâncias em taxas de reportamento dos dados. Para cada um desses conjuntos, analisaremos o efeito das medidas em 5, 10, 15, 30, 35, 40, 45, 50, 55, 60, 90 e 120 dias após o dia em que o indicador exibe que um certo nível de política pública está em efeito.

3.3 C1 School closing

C1 School closing é o indicador que descreve o quanto restrito foi a abertura de escolas naquele período, com um valor entre 0,1,2,3, ditando o nível da restrição. Podemos, portanto, utilizar este valor diretamente como grandeza escalar, e obter a correlação deste com o número de mortes

C1_School Closing	+5D	+10D	+15D	+20D	+25D	+30D	+35D	+40D
newDeathsPer100k	0.2379	0.2409	0.2462	0.2520	0.2540	0.2512	0.2465	0.2421
totalCasesPer100k	-0.5724	-0.5511	-0.5212	-0.4881	-0.4532	-0.4182	-0.3845	-0.3497
deathsPer100k	-0.6335	-0.6083	-0.5733	-0.5346	-0.4936	-0.4524	-0.4128	-0.3724

C1_School Closing	+45D	+50D	+55D	+60D	+90D	+120D
newDeathsPer100k	0.2371	0.2310	0.2208	0.2062	0.0829	-0.0526
totalCasesPer100k	-0.3158	-0.2822	-0.2492	-0.2168	-0.0607	0.0124
deathsPer100k	-0.3331	-0.2949	-0.2577	-0.2216	-0.0563	0.0082

3.4 C2 Workplace Closing

C2 Workplace Closing é o indicador que descreve o quanto restrito foi a abertura de ambientes de trabalho naquele período, com um valor entre 0,1,2,3, ditando o nível da restrição. Podemos, portanto, utilizar este valor diretamente como grandeza escalar, e obter a correlação deste com o número de mortes

C2_Workplace Closing	+5D	+10D	+15D	+20D	+25D	+30D	+35D	+40D
newDeathsPer100k	0.2791	0.2928	0.3034	0.3095	0.3098	0.3051	0.2975	0.2892
totalCasesPer100k	-0.4206	-0.4027	-0.3794	-0.3537	-0.3269	-0.3007	-0.2744	-0.2474
deathsPer100k	-0.5017	-0.4799	-0.4515	-0.4201	-0.3873	-0.3558	-0.3241	-0.2917

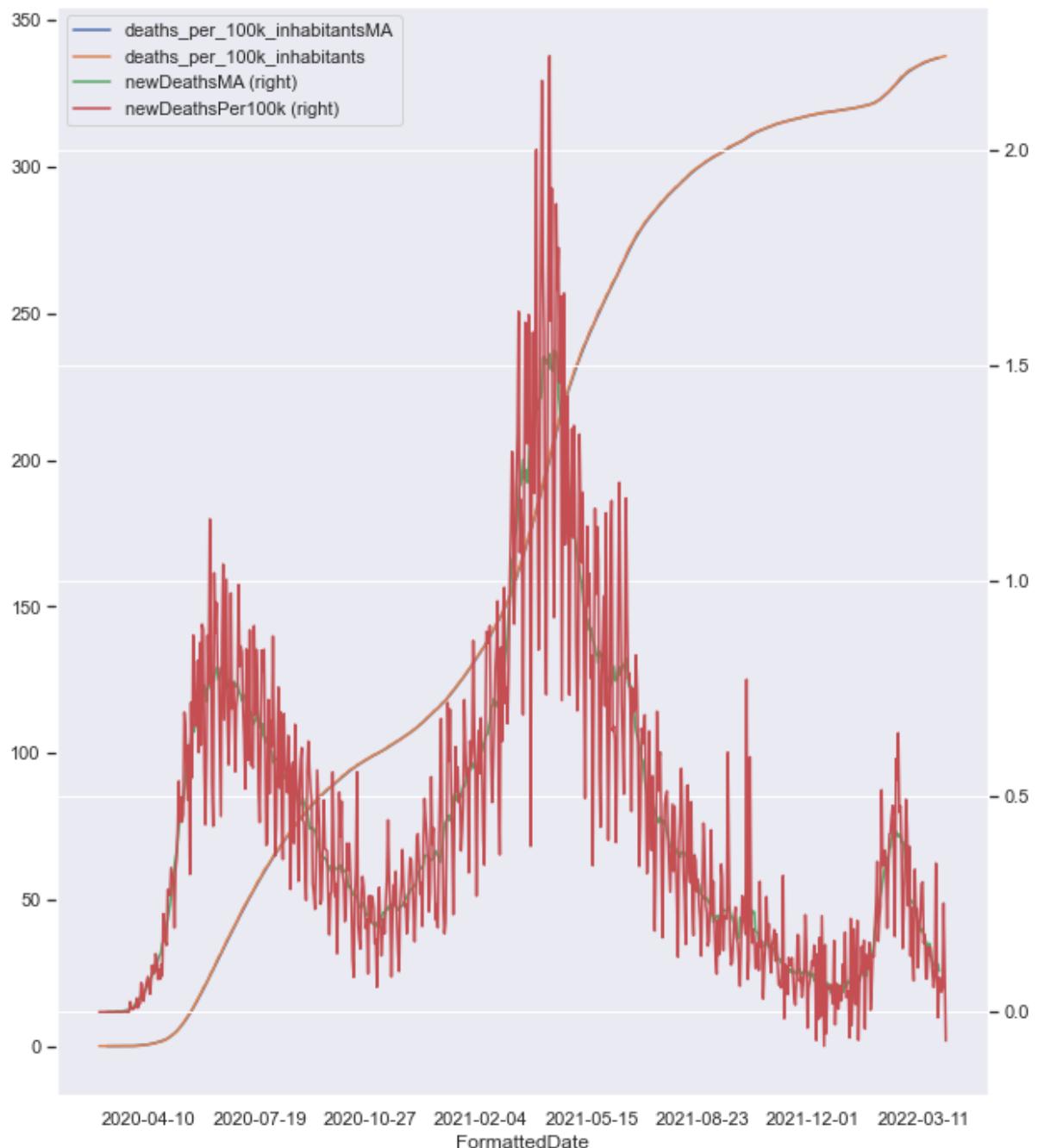


Figura 18 – Conjuntos de dados utilizados.

C2_Workplace Closing	+45D	+50D	+55D	+60D	+90D	+120D
newDeathsPer100k	0.2809	0.2723	0.2627	0.2501	0.1255	-0.0419
totalCasesPer100k	-0.2207	-0.1949	-0.1699	-0.1455	-0.0137	0.0452
deathsPer100k	-0.2601	-0.2294	-0.1999	-0.1713	-0.0318	0.0157

3.5 C3 Cancel public events

C1 Workplace Closing é o indicador que descreve como ocorreu a restrição à eventos públicos naquele período, com um valor entre 0,1,2, ditando o nível da restrição. Podemos, portanto, utilizar este valor diretamente como grandeza escalar, e obter a correlação deste com o número de mortes

C3_Cancel public events	+5D	+10D	+15D	+20D	+25D	+30D	+35D	+40D
newDeathsPer100k	0.2186	0.2226	0.2278	0.2333	0.2378	0.2397	0.2388	0.2374
totalCasesPer100k	-0.3524	-0.3381	-0.3133	-0.2824	-0.2497	-0.2171	-0.1862	-0.1564
deathsPer100k	-0.4048	-0.3876	-0.3582	-0.3215	-0.2826	-0.2442	-0.2071	-0.1713

C3_Cancel public events	+45D	+50D	+55D	+60D	+90D	+120D
newDeathsPer100k	0.2366	0.2340	0.2281	0.2181	0.0986	-0.0606
totalCasesPer100k	-0.1268	-0.0983	-0.0709	-0.0443	0.0957	0.1564
deathsPer100k	-0.1358	-0.1015	-0.0688	-0.0373	0.1136	0.1745

3.6 C4 Restrictions on gatherings

C4 Restrictions on gatherings é o indicador que descreve como foram restritas aglomerações naquele período, com um valor entre 0,1,2,3,4, ditando o nível da restrição. Podemos, portanto, utilizar este valor diretamente como grandeza escalar, e obter a correlação deste com o número de mortes

C4_Restrictions on gatherings	+5D	+10D	+15D	+20D	+25D	+30D	+35D	+40D
newDeathsPer100k	0.2082	0.2094	0.2129	0.2167	0.2199	0.2220	0.2218	0.2203
totalCasesPer100k	-0.3500	-0.3422	-0.3293	-0.3128	-0.2958	-0.2802	-0.2640	-0.2470
deathsPer100k	-0.4048	-0.3876	-0.3582	-0.3215	-0.2826	-0.2442	-0.2071	-0.1713

C4_Restrictions on gatherings	+45D	+50D	+55D	+60D	+90D	+120D
newDeathsPer100k	0.2192	0.2175	0.2134	0.2067	0.1210	-0.0100
totalCasesPer100k	-0.2290	-0.2108	-0.1928	-0.1738	-0.0528	0.0124
deathsPer100k	-0.1358	-0.1015	-0.0688	-0.0373	0.1136	0.1745

3.7 C5 Close public transport

C5 Close public transport é o indicador que descreve com qual intensidade o fechamento de transporte público foi naquele período, com um valor entre 0,1,2, ditando

o nível da restrição. Podemos, portanto, utilizar este valor diretamente como grandeza escalar, e obter a correlação deste com o número de mortes

C5_Close public transport	+5D	+10D	+15D	+20D	+25D	+30D	+35D	+40D
newDeathsPer100k	0.2378	0.2350	0.2327	0.2303	0.2239	0.2115	0.1990	0.1875
totalCasesPer100k	-0.3037	-0.2935	-0.2817	-0.2686	-0.2547	-0.2373	-0.2185	-0.1990
deathsPer100k	-0.3329	-0.3193	-0.3041	-0.2874	-0.2701	-0.2498	-0.2283	-0.2067

C5_Close public transport	+45D	+50D	+55D	+60D	+90D	+120D
newDeathsPer100k	0.1772	0.1662	0.1554	0.1445	0.0608	-0.0071
totalCasesPer100k	-0.1791	-0.1592	-0.1392	-0.1190	-0.0095	0.0556
deathsPer100k	-0.1850	-0.1635	-0.1424	-0.1216	-0.0152	0.0446

3.8 C6 Stay at home requirements

C6 Stay at home requirements é o indicador que descreve o quanto intenso as ordens de quarentena da população geral foi naquele período, com um valor entre 0,1,2,3, ditando o nível da restrição. Podemos, portanto, utilizar este valor diretamente como grandeza escalar, e obter a correlação deste com o número de mortes

C6_Stay at home requirements	+5D	+10D	+15D	+20D	+25D	+30D	+35D	+40D
newDeathsPer100k	0.3801	0.3806	0.3789	0.3752	0.3659	0.3509	0.3325	0.3123
totalCasesPer100k	-0.1773	-0.1648	-0.1503	-0.1344	-0.1174	-0.0959	-0.0735	-0.0502
deathsPer100k	-0.2356	-0.2171	-0.1961	-0.1734	-0.1500	-0.1232	-0.0956	-0.0679

C6_Stay at home requirements	+45D	+50D	+55D	+60D	+90D	+120D
newDeathsPer100k	0.2925	0.2743	0.2571	0.2376	0.1115	-0.0224
totalCasesPer100k	-0.0267	-0.0034	0.0193	0.0421	0.1641	0.2442
deathsPer100k	-0.0406	-0.0139	0.0119	0.0370	0.1611	0.2312

3.9 C7 Restrictions on internal movement

C7 Restrictions on internal movements é o indicador que descreve o quanto restritas foram as viagens nacionais naquele período, com um valor entre 0,1,2, ditando o nível da restrição. Podemos, portanto, utilizar este valor diretamente como grandeza escalar, e obter a correlação deste com o número de mortes

C7_Restrictions on internal movements	+5D	+10D	+15D	+20D	+25D	+30D	+35D	+40D
newDeathsPer100k	0.2168	0.2204	0.2230	0.2251	0.2256	0.2229	0.2194	0.2137
totalCasesPer100k	-0.3745	-0.3683	-0.3592	-0.3478	-0.3354	-0.3209	-0.3059	-0.2896
deathsPer100k	-0.4364	-0.4270	-0.4142	-0.3989	-0.3822	-0.3634	-0.3440	-0.3236

C7_Restrictions on internal movement	+45D	+50D	+55D	+60D	+90D	+120D
newDeathsPer100k	0.2047	0.1928	0.1789	0.1634	0.0347	-0.0873
totalCasesPer100k	-0.2729	-0.2562	-0.2390	-0.2215	-0.1287	-0.0740
deathsPer100k	-0.3030	-0.2828	-0.2627	-0.2428	-0.1465	-0.0986

3.10 C8 International travel controls

C8 International travel controls é o indicador que descreve o quanto restritas foram as viagens nacionais naquele período, com um valor entre 0,1,2, ditando o nível da restrição. Podemos, portanto, utilizar este valor diretamente como grandeza escalar, e obter a correlação deste com o número de mortes

C8_Restrictions on internal movements	+5D	+10D	+15D	+20D	+25D	+30D	+35D	+40D
newDeathsPer100k	0.1064	0.1185	0.1332	0.1497	0.1657	0.1792	0.1916	0.2009
totalCasesPer100k	-0.5862	-0.5828	-0.5738	-0.5602	-0.5443	-0.5257	-0.5064	-0.4867
deathsPer100k	-0.6227	-0.6180	-0.6070	-0.5909	-0.5721	-0.5499	-0.5270	-0.5036

C8_International travel controls	+45D	+50D	+55D	+60D	+90D	+120D
newDeathsPer100k	0.2067	0.2083	0.2057	0.1988	0.0882	-0.0527
totalCasesPer100k	-0.4656	-0.4439	-0.4220	-0.3998	-0.2847	-0.2363
deathsPer100k	-0.4790	-0.4540	-0.4292	-0.4046	-0.2857	-0.2448

3.11 H1 Public information Campaigns

H1 Public information Campaigns é o indicador que descreve a intensidade das campanhas de informação pública naquele período, com um valor entre 0,1,2, ditando o nível da restrição. Podemos, portanto, utilizar este valor diretamente como grandeza escalar, e obter a correlação deste com o número de mortes

C8_Restrictions on internal movements	+5D	+10D	+15D	+20D	+25D	+30D	+35D	+40D
newDeathsPer100k	0.0791	0.0701	0.0597	0.0468	0.0337	0.0310	0.0283	0.0267
totalCasesPer100k	-0.2748	-0.2527	-0.2286	-0.2020	-0.1767	-0.1671	-0.1587	-0.1514
deathsPer100k	-0.1664	-0.1448	-0.1213	-0.0953	-0.0713	-0.0661	-0.0624	-0.0589

C8_International travel controls	+45D	+50D	+55D	+60D	+90D	+120D
newDeathsPer100k	0.0259	0.0258	0.0279	0.0372	0.1278	0.0892
totalCasesPer100k	-0.1445	-0.1360	-0.1223	-0.0989	0.0907	0.2121
deathsPer100k	-0.0554	-0.0503	-0.0402	-0.0195	0.1713	0.2933

4 Discussão sobre os resultados

Percebe-se uma correlação negativa geral mediana (em torno de 0.3 à 0.6 em valor absoluto) de todos os indicadores em relação ao número de casos total por cem mil habitantes de cada município, e o total de casos por cem mil habitantes de cada município. Esta correlação diminui ao longo do tempo, indicando que, no geral, as políticas tendem à perder eficácia com o passar do tempo, e, como é negativa, indica que as políticas de fato reduzem os indicadores (número de casos total por cem mil habitantes de cada município, e o total de casos por cem mil habitantes de cada município) .

Existe uma correlação positiva baixa de todos os indicadores de baixa intensidade (0 à 0.3 em valor absoluto) com a quantidade de mortes novas por cem mil habitantes, a qual ainda não conseguimos explicar

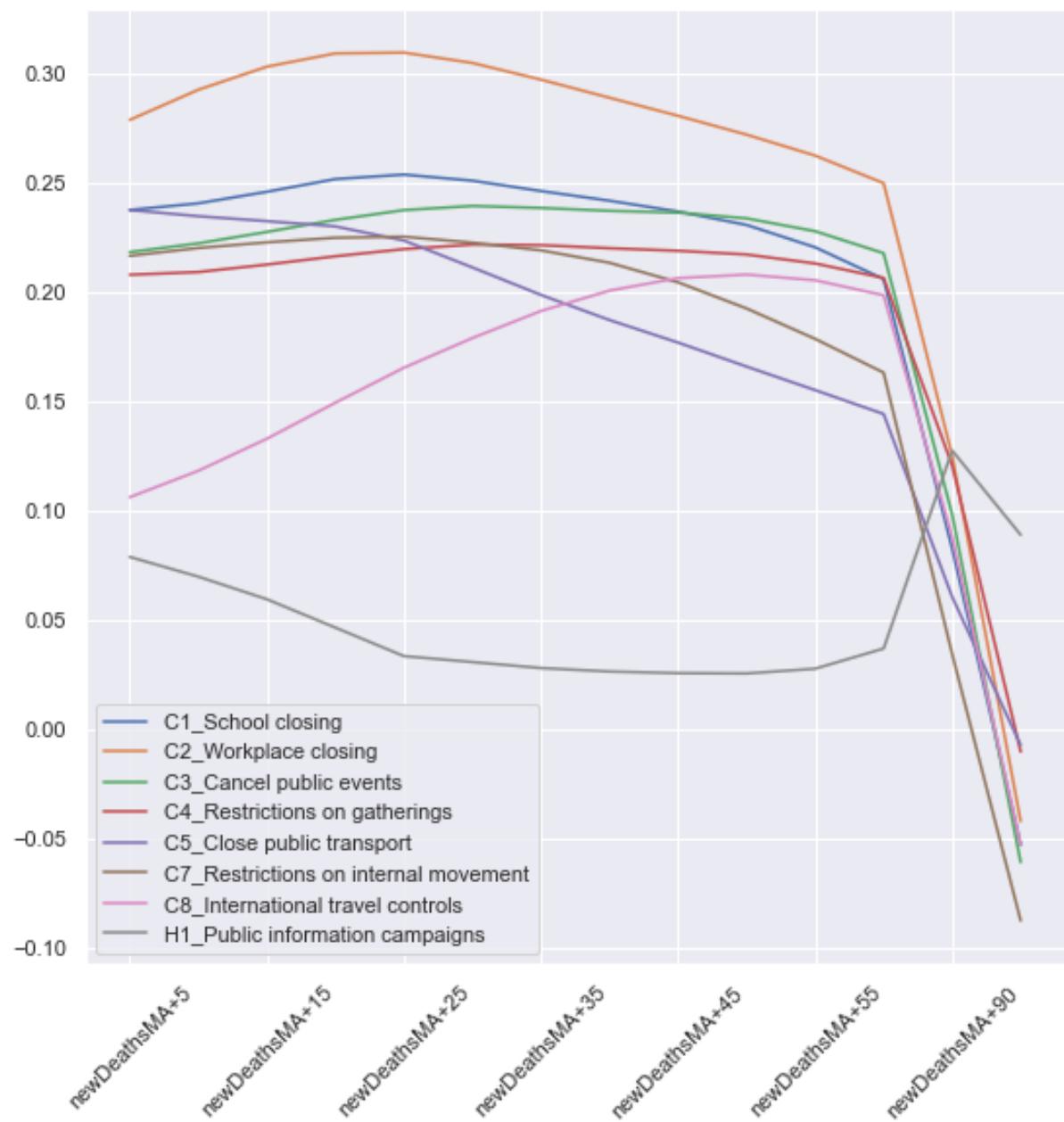


Figura 19 – Correlações entre medidas e a quantidade de novas mortes por 100 mil habitantes.

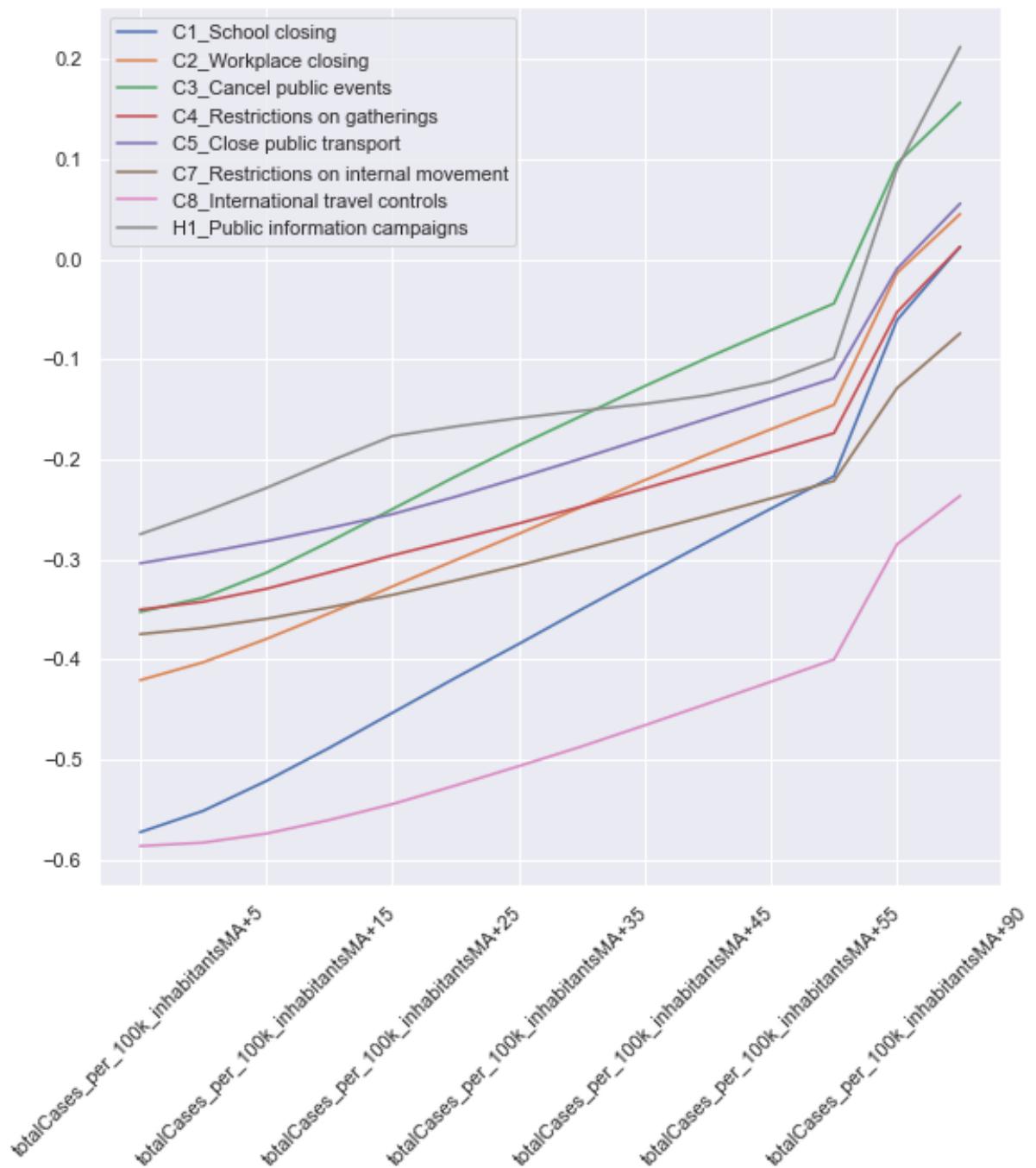


Figura 20 – Correlações entre medidas e a quantidade de casos totais por 100 mil habitantes.

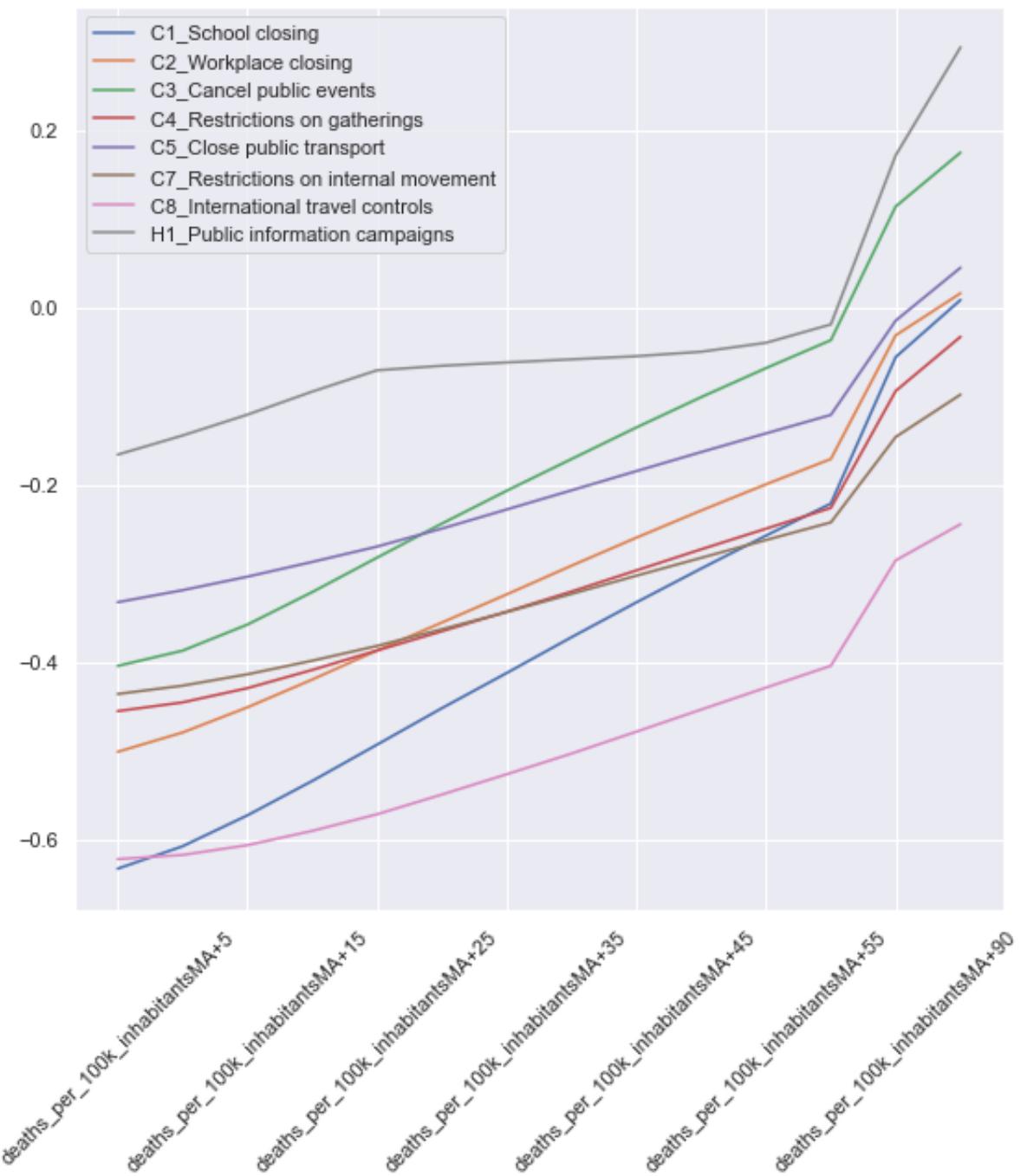


Figura 21 – Correlações entre medidas e a quantidade de mortes totais por 100 mil habitantes.

Referências

HEALTH, D. of. **Age-Adjusted Rates - Statistics Teaching Tools**. 1999. Disponível em: <<https://www.health.ny.gov/diseases/chronic/ageadj.htm>>. Citado na página 5.

STATISTICS, N. C. for H. **Provisional COVID-19 Deaths by Sex and Age**. 2020. Disponível em: <<https://data.cdc.gov/NCHS/Provisional-COVID-19-Deaths-by-Sex-and-Age/9bhg-hcku>>. Citado na página 5.