

Uma visualização da evolução da pandemia e das vacinações contra COVID-19 nas maiores economias do mundo

Hanna Rodrigues Ferreira

Julho 2021

1 Introdução

A pandemia de COVID-19, ou pandemia de coronavírus, é uma pandemia em curso de COVID-19, uma doença respiratória causada pelo coronavírus da SARS-CoV-2. Em dezembro de 2019, o vírus tem o primeiro caso conhecido da doença em Wuhan, na China. Em 11 de março de 2020, a OMS classificou-a como pandemia. Em 15 de julho de 2021, 188 302 197 casos foram confirmados, e com 4 065 528 mortes atribuídas à doença, tornando-se uma das pandemias mais mortais da história.[11] Atualmente mais de 12% da população global já foi totalmente vacinada.

Analizamos dados de casos, mortes e sobre a vacinação do coronavírus de 229 países retirados do site Our World in Data em 15/06/2021.[4] Para construir visualizações interpretáveis e informativas, através do tratamento e seleção dos dados utilizando os pacotes dplyr e Hmisc para tratamento dos dados e ggplot2 para visualizações no R.[8][1][9] Com a finalidade de explorar aspectos da atual situação da pandemia e das campanhas vacinais contra a COVID-19 nos países com as 15 maiores economias do mundo. Tais como uma tendência nos dados da pandemia, e outros fenômenos interessantes que serão inferidos com a visualização.

2 Objetivos

Explorar a atual situação da pandemia e das campanhas vacinais contra a COVID-19 dos países com as 15 maiores economias do mundo através de visualizações. Deste modo, o trabalho terá continuidade, fazendo-se uso de outros recursos e ferramentas computacionais.

3 Metodologia

Analizamos dados de casos, mortes e sobre a vacinação de 229 países retirados do site Our World in Data em 15/06/2021.[4] O dataset consiste-se em 60 colunas e 94 743 linhas, onde cada linha representa um dia desde o início da contagem. Os dados foram tratados com o pacote dplyr e o pacote Hmisc do R, obtendo-se como colunas os valores acumulados de casos, mortes e pessoas totalmente vacinadas, além de outras medidas auxiliares como tamanho populacional e continente, em que cada linha representa cada um dos 229 países.[8][1] Veremos a seguir como estas variáveis auxiliares foram utilizadas para aumentar a interpretabilidade da visualização destes dados.

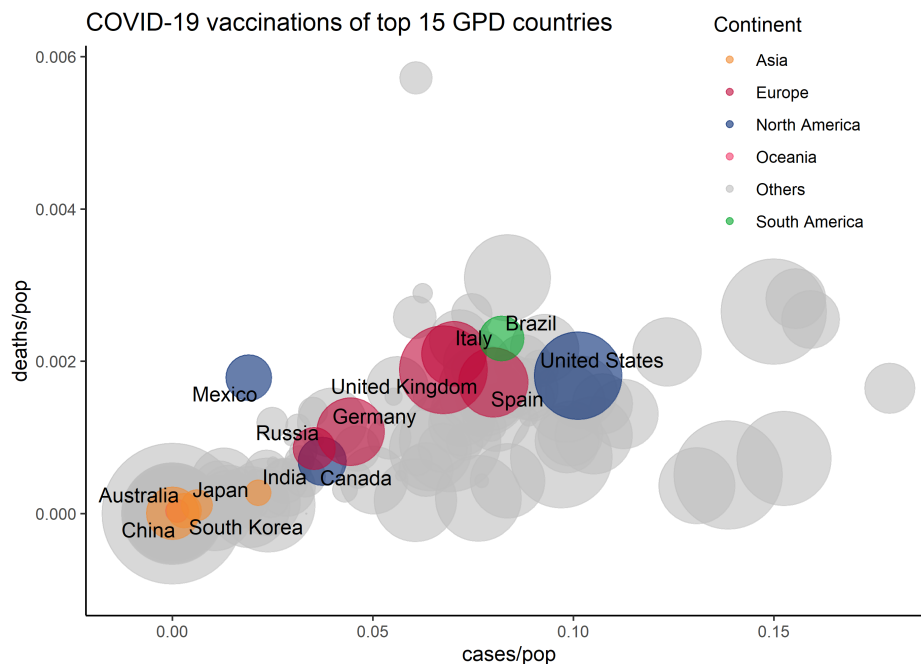
Como exemplo motivador, o número de mortes pelo coronavírus no Brasil atualmente é mais de 500 mil, e em outro país como Espanha é pouco mais de 81 mil mortes, não podemos simplesmente comparar estes países somente através dessa métrica.[4] Porque é importante considerar que o Brasil é aproximadamente 4,5 vezes mais populoso que a Espanha, este argumento vale de modo análogo também para o número de casos ou de pessoas vacinadas.[10] Portanto a escolha de fazer as variáveis da visualização como as porcentagens populacionais das métricas e não com seus valores absolutos torna-se a melhor solução para comparar países com dimensões populacionais tão distintas.

Definido como uma extensão do gráfico de dispersão, o gráfico de bolhas é comumente usado para visualizar relacionamentos entre três variáveis numéricas. Cada bolha em um gráfico representa um único ponto de dados. Os valores de cada bolha são codificados por sua posição horizontal, sua posição vertical e o tamanho da bolha.[12] Utilizamos a ggplot2 em nosso caso de estudo, definindo o eixo horizontal como a parcela da população que contrai COVID-19, o eixo vertical como a parcela da população que morreu por conta dela e o tamanho das bolhas como a porcentagem da população totalmente vacinada, onde cada bolha representa um país.[9]

Ao tentarmos tirar conclusões relevantes acerca da visualização olhando apenas para 229 bolhas parece pouco informativo. Desta forma, decidimos colorir as bolhas de acordo com o continente de cada país e adicionar a cada bolha o rótulo do nome de seu respectivo país. Todavia, o que antes parecia ser um gráfico pouco informativo, agora com todas as bolhas coloridas e com rótulos se tornou um excesso de informações, diminuindo a interpretabilidade. A fim de selecionar certos países para analisar, filtrei nosso Gráfico de Bolhas os 15 países de maior PIB de 2021.[7] Ou seja, as informações adicionais só foram para estes países do nosso recorte.

Relevante citar que optamos por manter todas as demais nações, em cinza, para ter uma noção global do comportamento desses países do recorte. Além disso, as cores dos continentes foram inspiradas nas bandeiras dos países em questão. Por exemplo, o laranja utilizado no continente asiático é o mesmo da bandeira indiana e o verde referente à América do Sul é o mesmo da bandeira brasileira.

4 Resultados parciais



Primeiramente, tanto dentro do nosso recorte quanto de um modo geral as bolhas possuem uma tendência linear crescente da parcela de casos e mortes, com variações mais comuns da porcentagem de mortos para menos. Um outlier chamativo neste contexto é o México, pois sua parcela de mortes foge a linha de tendência para mais. Este fato mostra que talvez os hospitais mexicanos estejam sobrecarregados,[3] e conseqüentemente culminaria numa maior taxa de mortalidade pela COVID-19.

Perceba também que a bolha do Brasil possui um tamanho pequeno se comparada com outros países do nosso recorte com parcelas semelhantes, tais como Estados Unidos, Espanha, Itália e Reino Unido. Desta forma, mostrando uma possível campanha vacinal lenta se comparada aos países citados anteriormente. Tornando-se assim um fator que contribuiu para o surgimento de novas cepas, devido a grande circulação do vírus no país.[6]

Outra coisa interessante a se ressaltar é a concentração desses países asiáticos perto origem do gráfico, ou seja, com poucos casos e mortes pelo coronavírus em comparação com sua população. O que mostra que conseguiram lidar bem a pandemia em seus respectivos países, controlando a pandemia na primeira onda.[5] Uma possível contribuição para esse fenômeno é pelo fato destes países já terem passado por surtos epidêmicos anteriores ao COVID-19, como por exemplo a SARS, síndrome respiratória aguda grave, em 2002.[2]

References

- [1] CRAN. Package Hmisc. <https://cran.r-project.org/web/packages/Hmisc/index.html>.
- [2] Valor Econômico. As lições da Ásia para conter a epidemia. <https://valor.globo.com/mundo/noticia/2020/03/23/as-lico-es-da-asia-para-conter-a-epidemia.ghtml>.
- [3] Forbes. Why Are So Many People Dying Of Covid-19 In Mexico? <https://www.forbes.com/sites/nathanielparishflannery/2020/09/03/why-are-so-many-people-dying-of-covid-19-in-mexico/?sh=3366742a18e7>.
- [4] Our World in Data. Coronavirus Pandemic (COVID-19). <https://ourworldindata.org/coronavirus>.
- [5] El País. Por que a Ásia está melhor que a Europa na pandemia? O segredo está no civismo. <https://brasil.elpais.com/internacional/2020-10-30/por-que-a-asia-esta-melhor-que-a-europa-na-pandemia-o-segredo-esta-no-civismo.html>.
- [6] El País. Ritmo lento na vacinação contra a covid-19 no Brasil favorece novas cepas do vírus. <https://brasil.elpais.com/brasil/2021-02-03/ritmo-lento-na-vacinacao-contra-a-covid-19-no-brasil-favorece-novas-cepas-do-virus.html>.
- [7] World Population Review. GPD Ranked by Country 2021. <https://worldpopulationreview.com/countries/countries-by-gdp>.
- [8] tidyverse. A Grammar of Data Manipulation - dplyr. <https://dplyr.tidyverse.org/>.
- [9] tidyverse. Create Elegant Data Visualizations Using the Grammar of Graphics - ggplot2. <https://ggplot2.tidyverse.org/>.
- [10] Wikipédia. Lista de países por população. https://pt.wikipedia.org/wiki/Lista_de_pa%C3%ADses_por_popula%C3%A7%C3%A3o.
- [11] Wikipédia. Pandemia de COVID-19. https://pt.wikipedia.org/wiki/Pandemia_de_COVID-19.
- [12] Storytelling with data. what is a bubble chart? <https://www.storytellingwithdata.com/blog/2021/5/11/what-is-a-bubble-chart>.