

## **Covid-19 nas Olimpíadas de 2021: como a postura adotada pelos países contra a pandemia impactou nos seus resultados**

### **Covid-19 at the 2021 Olympic Games: how the stance adopted by countries against the pandemic impacted their results**

Giovani Falcometa da Cunha<sup>1\*</sup>, Emerson Augusto Priamo Moraes<sup>1</sup>, Sandro Roberto Fernandes<sup>1</sup>

---

#### **RESUMO**

Os Jogos Olímpicos são o maior evento esportivo do mundo, contando com a participação de mais de 200 países. Devido a essas proporções, é responsável pela movimentação de muito dinheiro, fato que o torna ainda mais relevante. Sabendo que a pandemia de Covid-19 impactou diversas áreas, inclusive a esportiva, este trabalho objetiva analisar se a postura adotada pelos países no combate à pandemia pode ter impactado seus resultados nos jogos de Tóquio 2021. Para as análises, diferentes indicadores foram agrupados para criar um índice geral referente à severidade com que os países atuaram contra a pandemia. Através de softwares de análise estatística e utilizando-se de testes de hipóteses, os valores foram comparados entre os diferentes grupos de países, separados previamente através de técnicas de clusterização. Após as análises, não foi possível estabelecer nenhuma relação entre esse grau de severidade adotado no combate à pandemia e os resultados olímpicos.

**Palavras-chave:** Covid-19; Pandemia; Olimpíadas; Tóquio 2021.

---

#### **ABSTRACT**

The Olympic Games are the biggest sporting event in the world, with more than 200 countries participating. Due to these proportions, it is responsible for moving a lot of money, a fact that makes it even more relevant. Knowing that the Covid-19 pandemic has impacted several areas, including sports, this work aims to analyze whether the posture adopted by countries in the fight against the pandemic may have impacted their results in the Tokyo 2021 games. For the analyses, different indicators were grouped together to create an overall index referring to the severity with which countries acted against the pandemic. Through statistical analysis software and using hypothesis tests, the values were compared between the different groups of countries, previously separated through clustering techniques. After the analyses, it was not possible to establish any relationship between this degree of severity adopted in the fight against the pandemic and the Olympic results.

**Keywords:** Covid-19; Pandemic; Olympics; Tokyo 2021.

---

---

<sup>1</sup> Instituto Federal de Educação, Ciência e Tecnologia do Sudeste de Minas Gerais – Campus Juiz de Fora

\*E-mail: giovanifcunha@gmail.com

## INTRODUÇÃO

Os Jogos Olímpicos são uma tradição com mais de 2000 anos. Originários da Grécia Antiga e modernizados no final do século XIX, hoje são considerados o maior evento esportivo do mundo e unem mais de 200 países em competições de diversas modalidades esportivas. Objetivando primordialmente a união e a paz entre as nações, outro fator que torna esse evento tão importante e relevante é o montante de dinheiro movimentado. Estima-se que os investimentos totais para Tóquio sediar os jogos, por exemplo, tenha sido de, aproximadamente, US\$ 15 bilhões (LAPORTA, 2021). Também vale destacar o valor gasto por patrocinadores, ou seja, empresas e instituições que esperam obter algum retorno com o evento: em 2021, esse montante ultrapassou os US\$ 3 bilhões (ANDRADE, 2021).

Com tanto dinheiro envolvido, os países estão sempre investindo e incentivando seus atletas, visando sempre melhorar seus resultados e, conseqüentemente, obter uma visibilidade positiva que pode ser convertida em retorno financeiro. Apesar disso, imprevistos podem ocorrer e influenciar todo o planejamento, como é o caso da pandemia global de Covid-19.

Originando-se na China e expandindo-se para o restante do mundo em poucos meses, a pandemia trouxe diversos prejuízos para o mundo globalizado nas mais diversas áreas. As estimativas do Fundo Monetário Internacional (FMI) são de que, até 2024, os custos relacionados à pandemia ultrapassem os US\$ 12.5 trilhões (SHALAL, 2022). E esse impacto chegou também às Olimpíadas: o adiamento dos jogos de 2020 para 2021 teve um custo estimado em quase US\$ 3 bilhões (LIES, 2021). Esse valor inclui, por exemplo, os gastos com a devolução de ingressos, agravada pela proibição da entrada de estrangeiros no país para assistirem ao evento. A conta ficou mais cara também para os patrocinadores, que precisaram desembolsar mais US\$ 200 milhões para estender contratos (ibid).

Visando combater o vírus e amenizar seus impactos, diversas medidas foram tomadas pelos países, desde campanhas de informação a políticas de distanciamento social (RITCHIE et al, 2020). Considerando que a pandemia impactou a preparação de atletas olímpicos (URBAŃSKI; SZELIGA; TASIEMSKI, 2021) e que estudos indicam que algumas dessas medidas podem impactar no desempenho dos atletas (JAGIM et al, 2020), o presente trabalho objetiva analisar a relação entre o rigor adotado na execução das medidas pelos países e os resultados obtidos nos jogos olímpicos de 2021.

## MATERIAIS E MÉTODOS

O processo metodológico iniciou-se com uma pesquisa bibliográfica para levantamento de dados referentes às medidas de combate à pandemia adotadas pelos países e pelos dados referentes aos seus desempenhos nos jogos olímpicos. Uma premissa utilizada foi a de que todos os dados deveriam ser oriundos de repositórios abertos, ou seja, que estivessem disponíveis para uso sem necessidade de autorização prévia e sem custos envolvidos. Com os dados coletados, seguiu-se para a definição das ferramentas que seriam utilizadas, preferencialmente de uso livre ou com versões gratuitas. Com esses pontos definidos, a metodologia utilizada para analisar e validar os resultados foi a de teste de hipóteses. Hipóteses a respeito da relação entre as medidas de combate à pandemia e os resultados olímpicos foram levantadas e testadas utilizando-se de testes e comparações estatísticas.

## SELEÇÃO E PREPARAÇÃO DOS DADOS

Os dados utilizados no trabalho foram separados em dois grupos: aqueles referentes aos resultados olímpicos, utilizados para agrupar os países de acordo com as medalhas obtidas; e aqueles relacionados às medidas de combate à pandemia adotadas por esses países.

Para o primeiro grupo, o site do Comitê Olímpico Internacional (COI) foi utilizado para obter os resultados dos jogos de 2021. Os resultados referentes aos anos anteriores, apesar de também disponíveis através do COI, foram retirados da plataforma Kaggle, que fornecia os dados já agrupados e mais bem organizados.

Os dados do segundo grupo foram majoritariamente retirados do projeto *Oxford Coronavirus Government Response Tracker* (OxCGRT). Esse projeto consiste no agrupamento de indicadores sobre medidas de combate à Covid-19 retirados de fontes oficiais dos países e organizados por centenas de pesquisadores ao redor do mundo. O OxCGRT foi importante por disponibilizar os dados referentes a cada dia, ou seja, com ele foi possível ter uma visão exata do período escolhido para estudo.

Após a coleta preliminar de dados, foram feitos ajustes para melhor organizar e trabalhar com esses dados. Todas as bases utilizadas encontravam-se no formato .csv, padrão que facilita o manuseio de dados entre diferentes aplicações, mas que dificulta sua visualização. Assim, os dados foram convertidos para .xlsx e separados em colunas rotuladas, facilitando o entendimento de cada medida e suas gradações. Esses dados

também foram agrupados em um único arquivo, a partir do qual criou-se uma tabela principal para exibir todas as informações. Ainda foi necessário padronizar os nomes dos países que, vindos de diferentes fontes, apareciam com formatos e idiomas diferentes e poderiam gerar incongruências nas análises.

## **SOFTWARES**

O Orange, consiste em um conjunto de ferramentas para visualização e análise de dados. Desenvolvido na Universidade de Liubliana, sua principal vantagem é o fato de funcionar de forma mais visual, fornecendo blocos arrastáveis para que o usuário execute as ações, dispensando o uso de códigos e agilizando as análises. Também possui alguns complementos que podem ser instalados de acordo com a necessidade.

Power BI Desktop, versão gratuita do *software* da Microsoft, consiste em uma ferramenta robusta para análise e visualização de dados. O número de funcionalidades e a possibilidade de integração com diferentes fontes de dados tornam-se o ponto forte da ferramenta. Se necessário, também possibilita a criação de dashboards interativos, que facilitam a exibição e o entendimento das informações que precisam ser apresentadas.

O Excel, também da Microsoft, é um dos editores de planilhas mais completos do mercado, podendo ser acesso de forma gratuita na versão online. Versátil e utilizável em diversas aplicações, seu diferencial é a facilidade com que permite formatar dados, levando vantagem em relação a softwares mais robustos por não exigir tanto poder de processamento para essas tarefas.

Jamovi é uma ferramenta de código aberto para análises estatísticas. Criado a partir da colaboração de dezenas de desenvolvedores ao redor do mundo e com a premissa de ser um software “neutro”, sua principal característica é permitir que o usuário tenha acesso a diferentes métodos estatísticos lado a lado, podendo comparar os resultados de diferentes abordagens. Também conta com um conjunto de módulos extras para análises mais específicas.

## **ANÁLISE ESTATÍSTICA**

### **TESTE DE HIPÓTESES**

Em uma definição formal, podemos dizer que hipótese estatística é uma afirmação sobre os parâmetros de uma ou mais populações (MONTGOMERY; RUNGER, 2009). O teste de hipóteses nos permite identificar se dados amostrais de determinada população

evidenciam ou não determinada hipótese acerca da população total (BUSSAD; MORETTIN, 2010).

Iniciamos o teste definindo a hipótese a ser testada, que recebe o nome de hipótese nula ( $H_0$ ), e sua contraparte, chamada de hipótese alternativa ( $H_1$ ), que obrigatoriamente é verdadeira caso  $H_0$  seja falsa. A caracterização estatística do teste depende do nível de conhecimento do problema analisado (ibid), mas sua forma mais comum se dá pelo que chamamos de teste bilateral, representado por:

$$\begin{cases} H_0: \theta = x \\ H_1: \theta \neq x \end{cases}$$

Também existem os chamados testes unilaterais, que utilizam outras operações de comparação, tais como:

$$\begin{cases} H_0: \theta \leq x \\ H_1: \theta > x \end{cases} \quad \text{ou} \quad \begin{cases} H_0: \theta \geq x \\ H_1: \theta < x \end{cases}$$

Como em qualquer análise estatística, o resultado está sujeito a erros. No teste de hipóteses temos o erro do tipo I (probabilidade  $\alpha$ ), no qual rejeitamos a hipótese  $H_0$  mesmo ela sendo verdadeira; e o erro do tipo II (probabilidade  $\beta$ ), quando não rejeitamos a hipótese  $H_0$  e ela é, de fato, falsa.

$$\alpha = P(\text{rejeitar } H_0 \mid H_0 \text{ é verdadeira})$$

$$\beta = P(\text{não rejeitar } H_0 \mid H_0 \text{ é falsa})$$

Chamamos a probabilidade  $\alpha$  de nível de significância e seu valor é arbitrariamente definido, sendo usualmente fixado em 1%, 5% ou 0,1% (ibid). Esse valor auxilia na definição da chamada região crítica, que delimita os valores extremos que a amostragem pode assumir para aceitarmos ou não  $H_0$ . Rejeitamos  $H_0$  se o valor observado estiver dentro da região crítica; caso contrário, aceitamos  $H_0$ . Neste trabalho, fixou-se esse valor em 5%.

## COEFICIENTE DE VARIAÇÃO (CV)

O coeficiente de variação consiste em uma medida de dispersão utilizada para verificar a variabilidade de um conjunto de dados (BUSSAD; MORETTIN, 2010). Seu valor é calculado a partir da razão entre o desvio padrão ( $s$ ) e a média ( $\bar{x}$ ) da amostra e expresso em porcentagem:

$$CV = \frac{s}{\bar{x}} * 100$$

As classificações dos coeficientes utilizadas neste trabalho foram baseadas na metodologia proposta por (TUKEY, 1977). São elas:

**Tabela 1** - Classificações do coeficiente de variação (CV)

<b>Classificação</b>	<b>Intervalo</b>
Baixo	$CV \ll 26\%$
Médio	$26\% < CV \ll 33\%$
Alto	$33\% < CV \ll 37\%$
Muito Alto	$CV > 37\%$

Fonte: Elaboração própria, 2022.

### **LIMITE INFERIOR (LI) E LIMITE SUPERIOR (LS)**

Os conceitos de limite inferior e limite superior utilizados neste trabalho foram adaptados para servirem de parâmetro na comparação do número de medalhas. Aqui, eles indicam o mínimo e o máximo de medalhas que um país deve ter obtido em 2021 para que os resultados sejam considerados estáveis. Da mesma forma, caso esse valor seja inferior ou superior aos limites, o resultado será considerado, respectivamente, pior ou melhor. Foram calculados a partir da fórmula:

$$LI = \bar{x} - t * \frac{s}{\sqrt{n}} \qquad LS = \bar{x} + t * \frac{s}{\sqrt{n}}$$

Em que:

- $\bar{x}$ : média amostral;
- t: valor tabelado de acordo com significância;
- s: desvio amostral;
- n: tamanho amostral

### **TESTE T**

O teste t consiste em um método estatístico utilizado para a comparação de médias (GOMES, 1990). A partir do cálculo da chamada estatística t, tem como objetivo inferir se duas médias podem ser consideradas estatisticamente iguais. Podemos utilizar o conceito já definido de nível de significância, também chamado de nível de confiança, para avaliar o resultado de t: caso o valor encontrado seja menor que o nível de

significância definido, assumimos que as médias são iguais; caso contrário, assumimos que as médias são diferentes.

A fórmula padrão para o cálculo de  $t$  é dada por:

$$t = \frac{\bar{x} - u_0}{\frac{s}{\sqrt{n}}}$$

Em que:

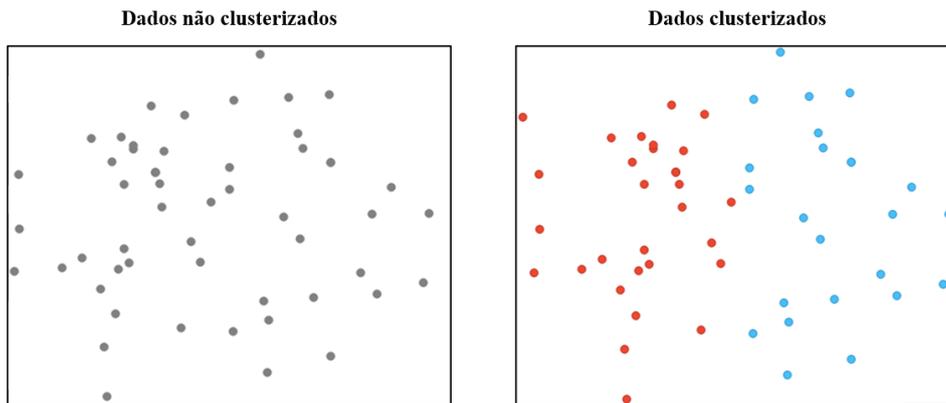
- $\bar{x}$ : média amostral;
- $u_0$ : média ou valor utilizado para comparação;
- $s$ : desvio amostral;
- $n$ : tamanho amostral

## CLUSTERIZAÇÃO (*CLUSTERING*)

Quando se trabalha com dados, uma das atividades mais importantes é a sua classificação ou agrupamento de acordo com características em comum (XU; WUNSCH, 2008) para simplificar e facilitar seu entendimento (EVERITT; LANDAU; LEESE, 2011). Uma das técnicas existentes para a realização dessa classificação é a clusterização (*clustering*), também chamada de análise de grupos, no qual os dados não estão previamente rotulados (JAIN; DUBES, 1988).

O processo de clusterização consiste em agrupar esses dados não rotulados em grupos denominados *clusters*, termo para o qual não existe um consenso universal a respeito da definição (EVERITT; LANDAU; LEESE, 2011). Para este trabalho será utilizada a definição de que um *cluster* é um conjunto de dados, parecidos entre si, que se difere de dados de outros *clusters* (ibid). Com isso, ao término da clusterização, podemos dizer que nossos dados estão rotulados (clusterizados).

**Figura 1** – À esquerda, temos um conjunto de dados não clusterizados; à direita, temos esses mesmos dados agrupados em dois *clusters* (azul e vermelho)



Fonte: Elaboração própria, 2022.

É possível encontrar uma grande variedade de algoritmos de clusterização na literatura, dentre eles o k-means, escolhido para ser utilizado neste trabalho.

### ALGORITMO K-MEANS

O algoritmo k-means é um dos mais conhecidos e utilizados na literatura, uma vez que possui muitas variações e é adaptável para diferentes contextos. Proposto por MacQueen (MACQUEEN, 1967), objetiva agrupar os dados em torno de k grupos de acordo com a distância média entre esses dados e o centro (*mean*) desses grupos predefinidos.

Assumindo que  $P = [p_1, \dots, p_n]$  seja o conjunto de dados a serem clusterizados e  $C = [c_1, \dots, c_k]$  o conjunto de centróides, o algoritmo funciona da seguinte forma:

1. O número inicial de *clusters* k é definido e validado (a validação é realizada, por exemplo, através do método da silhueta);
2. O centróide dos k são definidos de forma aleatória ou seguindo algum critério pré-definido;
3. A distância euclidiana entre cada dado e o centróide dos *clusters* é calculada. Para  $c = (c_1, c_2)$  e  $p = (p_1, p_2)$ , temos:

$$\text{dist}(c,p) = \sqrt{(c_1 - p_1)^2 + (c_2 - p_2)^2}$$

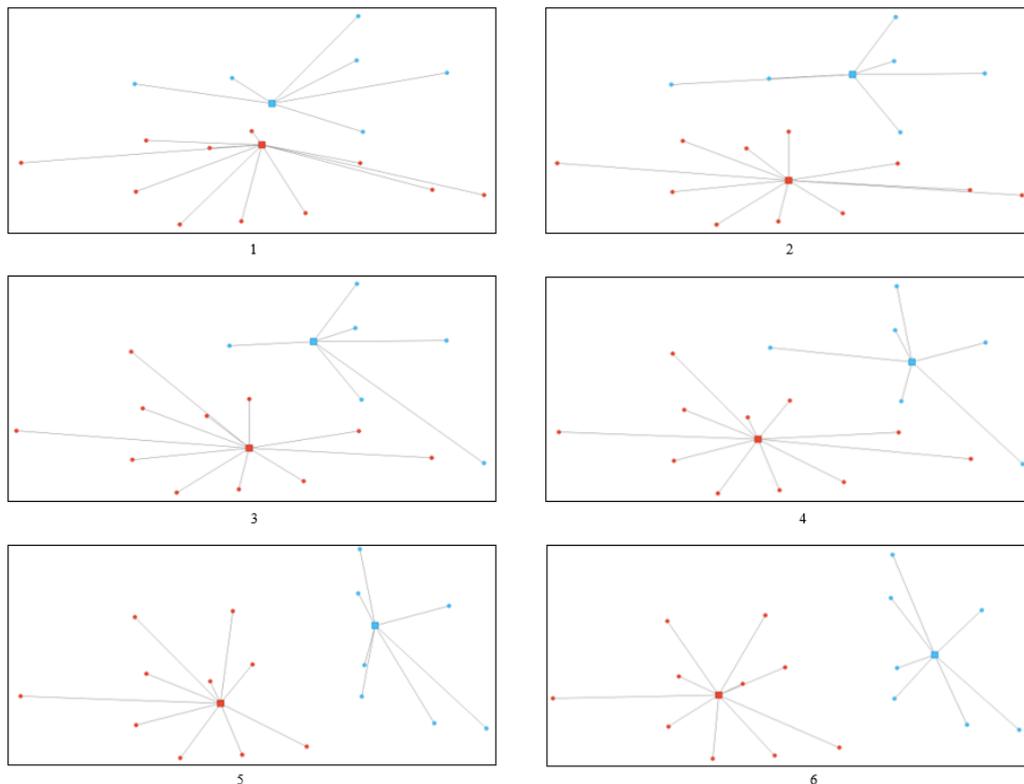
4. O dado é inserido ao *cluster* cuja distância tenha sido a menor;

5. Recalcula-se o centróide dos *clusters*, sendo S o conjunto de pontos pertencentes a um *cluster*:

$$c_i = \frac{1}{|S_i|} \sum_{x_i \in S_i} x_i$$

6. Os passos 2, 3 e 4 são iterados até que não haja mais alteração nos centróides.

**Figura 2** - Exemplificação gráfica do algoritmo k-means. Nela é possível perceber que, a cada iteração e até não ser mais possível, os pontos são realocados para o melhor *cluster*



Fonte: Elaboração própria, 2022.

## MÉTODO DA SILHUETA

Proposto por Rousseeuw (ROUSSEEUW, 1987), o método da silhueta (*silhouette method*) auxilia na determinação do número ideal de *clusters* para dada amostra de dados. São necessárias apenas duas premissas para esse método: que os dados estejam previamente clusterizados (independentemente do método) e que as distâncias entre os diferentes dados da amostra sejam conhecidas (ibid). Sua execução se dá por:

1. Assumindo que A seja um *cluster*, que  $i$  pertença a A e que  $q_A$  seja a quantidade de elementos de A, calculamos a média das distâncias entre  $i$  e os outros elementos pertencentes a A:

$$a(i) = \frac{1}{q_A - 1} \sum_{j \in A, i \neq j} d(i, j)$$

2. Assumindo que B seja um outro *cluster*, que  $i$  pertença a A, que  $j$  pertença a B e que  $q_A$  e  $q_B$  sejam as quantidades de elementos de cada *cluster*, calculamos a média das distâncias entre  $i$  e os elementos pertencentes a B:

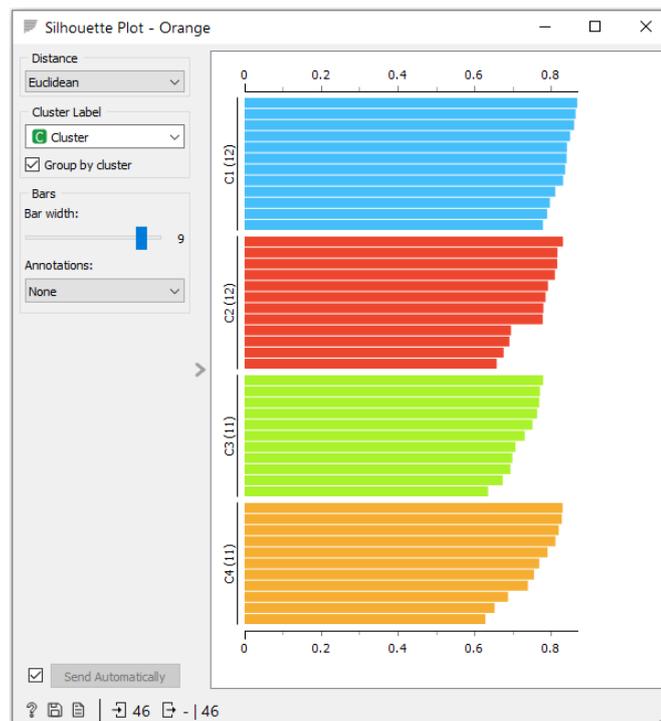
$$b(i) = \frac{1}{q_B} \sum_{j \in B} d(i, j)$$

3. Por fim, definimos o coeficiente de silhueta para  $i$  por:

$$c(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}$$

A média dos coeficientes de silhueta dos pontos de determinado *cluster* indicam o quão bem agrupados estão os dados. Assim, o número ideal de *cluster* é aquele cuja média dos coeficientes é mais próxima de 1.

**Figura 3** - Exemplo do resultado da aplicação do método da silhueta. Cada linha representa o coeficiente de determinado ponto.



Fonte: Elaboração própria, 2022.

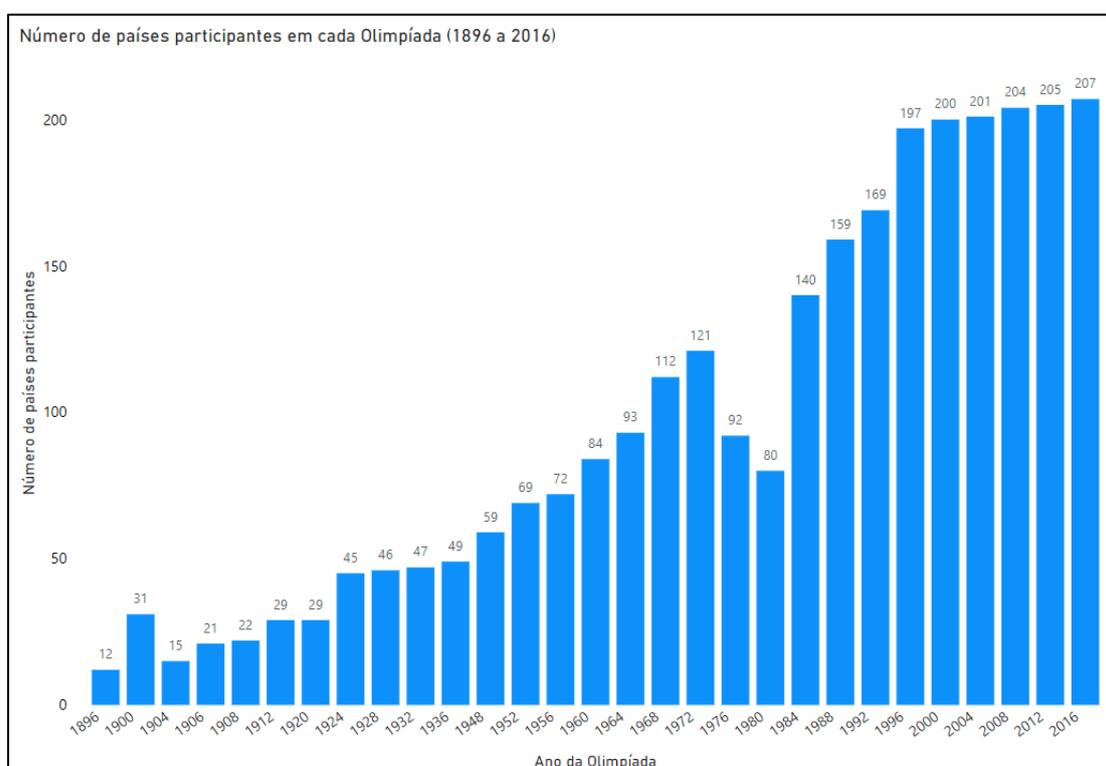
## DESENVOLVIMENTO

### AGRUPAMENTO E ANÁLISE DOS PAÍSES

Para ter-se uma visão do desempenho olímpico médio dos países, pensou-se em classificá-los com base no histórico de todos os jogos olímpicos modernos (de 1896 a 2016) e, a partir de então, usar o grupo de países com os melhores números de medalhas totais e medalhas de ouro. Porém, utilizar os dados dos mais de 100 anos de Olimpíadas apresentou as seguintes desvantagens:

- cada país iniciou sua participação em anos diferentes;
- a participação dos países nos jogos não era constante, variando o número de países em cada Olimpíada (Figura 4);
- países foram criados e extintos nesse período, impactando na distribuição de medalhas;
- eventos mundiais, tais como as duas grandes guerras, interferiram no andamento dos jogos.

**Figura 4** – Número de participantes em cada Olimpíada (1896 a 2016)



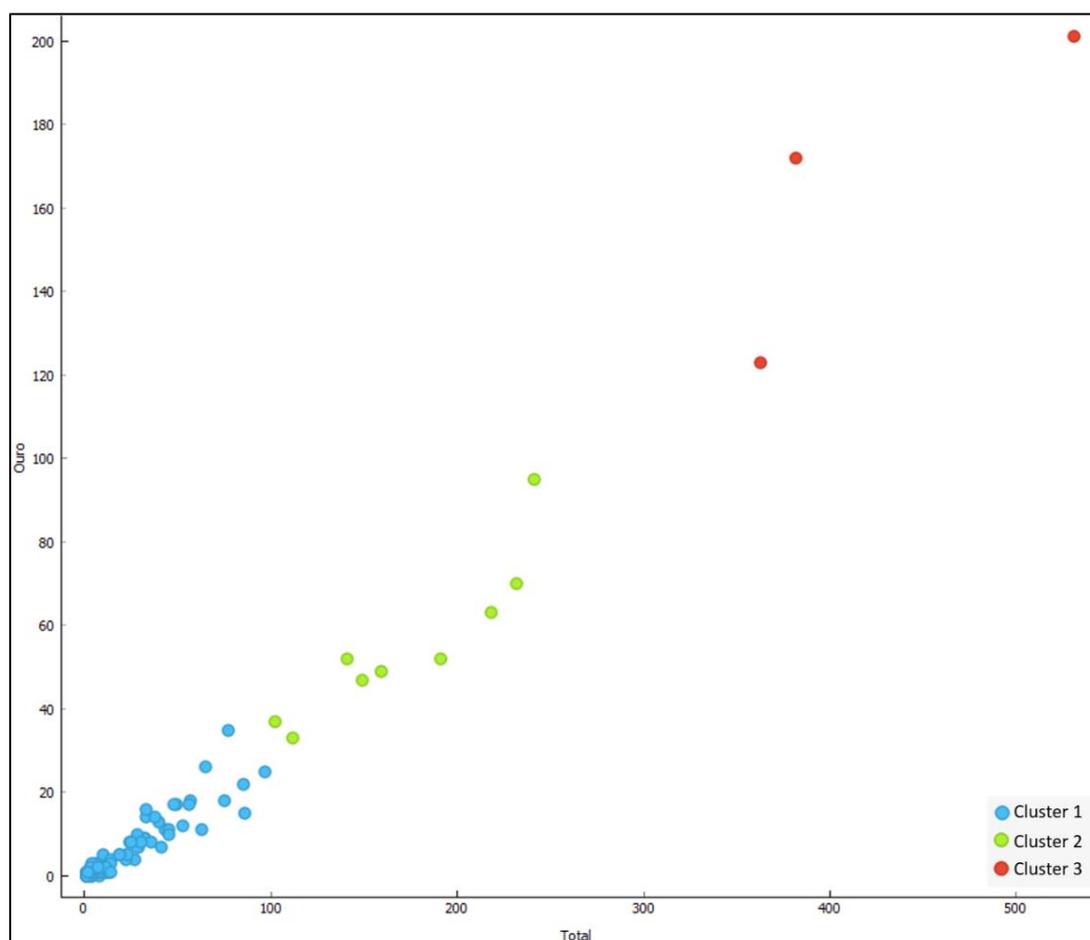
Fonte: Elaboração própria, 2022.

Diante desses fatos, optou-se por fazer a seleção levando em consideração apenas os resultados das últimas cinco Olimpíadas (2000, 2004, 2008, 2012, 2016). Com isso, resolveu-se não apenas os problemas listados, mas também foi possível obter resultados mais condizentes com o cenário olímpico atual. Por essa análise, conseguimos um total de 119 países com pelo menos uma medalha.

Também foi preciso validar se os países possuíam registros referentes às medidas de combate à pandemia. Países sem essas informações não poderiam ser selecionados, independentemente dos resultados olímpicos. Assim, depois de aplicado este filtro, ficamos com um total de 108 países para serem analisados.

Através da ferramenta *Orange*, que utiliza o algoritmo de clusterização k-means, esses países foram agrupados em 3 *clusters* de acordo com seu somatório total de medalhas e seu somatório de medalhas de ouro:

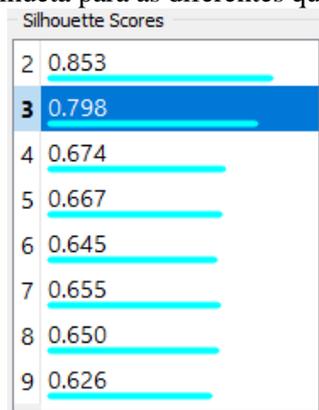
**Figura 2** – Agrupamento dos países nos 4 *clusters*



Fonte: Elaboração própria, 2022.

Vale destacar que o número ideal de *clusters*, de acordo com o Método da Silhueta, seria 2, uma vez que apresenta a pontuação mais próxima de 1. Porém, visto que um número maior de *clusters* seria mais interessante para a análise do trabalho, optou-se pelo agrupamento em 3 *clusters*, cujo valor de silhueta é o mais próximo daquele encontrado para 2 *clusters*.

**Figura 3** - Coeficiente de silhueta para as diferentes quantidades de *clusters* testadas



Fonte: Elaboração própria, 2022.

É possível perceber que existe uma discrepância muito grande nos resultados obtidos pelos países nas Olimpíadas. Enquanto o *cluster* 1, que abriga os países com resultados mais inexpressivos, agrupa um total de 96 países, o *cluster* 3, que representa os países de alto desempenho, tem apenas 3. A distribuição completa dos países nos *clusters* ficou da seguinte forma:

- *Cluster* 1: Afeganistão, África do Sul, Arábia Saudita, Argélia, Argentina, Áustria, Azerbaijão, Bahamas, Bahrein, Barbados, Bélgica, Bielorrússia, Botsuana, Brasil, Bulgária, Burundi, Camarões, Canadá, Catar, Cazaquistão, Chile, Chipre, Colômbia, Costa Rica, Croácia, Dinamarca, Egito, Emirados Árabes Unidos, Equador, Eritreia, Eslováquia, Eslovênia, Espanha, Estônia, Etiópia, Fiji, Filipinas, Finlândia, Gabão, Geórgia, Grécia, Guatemala, Hong Kong, Hungria, Índia, Indonésia, Irã, Irlanda, Islândia, Israel, Jamaica, Jordânia, Kosovo, Kuwait, Letônia, Lituânia, Malásia, Marrocos, Maurício, México, Moçambique, Moldávia, Mongólia, Níger, Nigéria, Noruega, Nova Zelândia, Panamá, Paraguai, Polônia, Porto Rico, Portugal, Quênia, Quirguistão, República Checa, República Dominicana, Romênia, Singapura, Síria, Sri Lanka, Sudão,

- Suíça, Tailândia, Tadjiquistão, Togo, Trindade e Tobago, Tunísia, Turquia, Uganda, Uruguai, Uzbequistão, Venezuela, Vietnã, Suécia, Ucrânia e Zimbábwe;
- *Cluster 2*: Alemanha, Austrália, Coreia do Sul, Cuba, França, Grã-Bretanha, Itália Japão e Países Baixos;
  - *Cluster 3*: China, Estados Unidos e Rússia.

Como não existe um consenso a respeito da melhor maneira de classificar o desempenho de um país em Olimpíadas, ou seja, se um país deve ser ranqueado de acordo com seu número de medalhas de ouro ou de acordo com seu total de medalhas (GREENE, 2018), optou-se, neste trabalho, por utilizar o número total de medalhas. Uma vez que muitos países não possuem sequer uma medalha de ouro, entendeu-se que, dessa forma, seria possível incluir mais países na análise.

Comparando os resultados de 2021 com a média dos jogos anteriores, é possível notar que determinados países mantiveram um resultado estável, enquanto outros tiveram uma melhora ou piora. Para evitar que uma variação pequena no número de medalhas impactasse nessa classificação, foi utilizado um intervalo criado a partir do limite inferior (LI) e do limite superior (LS) para avaliar esses resultados (R):

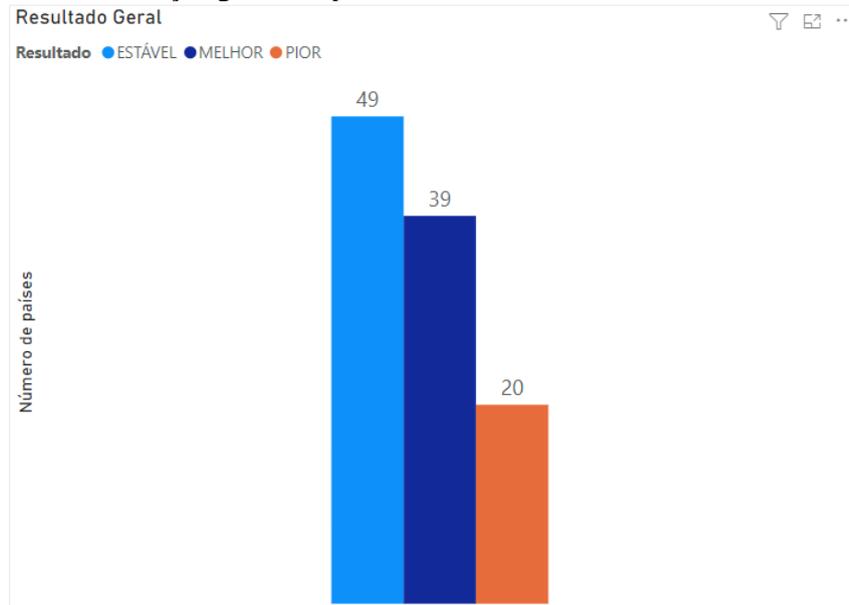
**Tabela 2** – Classificação adotada para os resultados de 2021

<b>Resultado 2021</b>	<b>Classificação</b>
R < LI	Pior
LI < R < LS	Estável
LS < R	Melhor

Fonte: Elaboração própria, 2022.

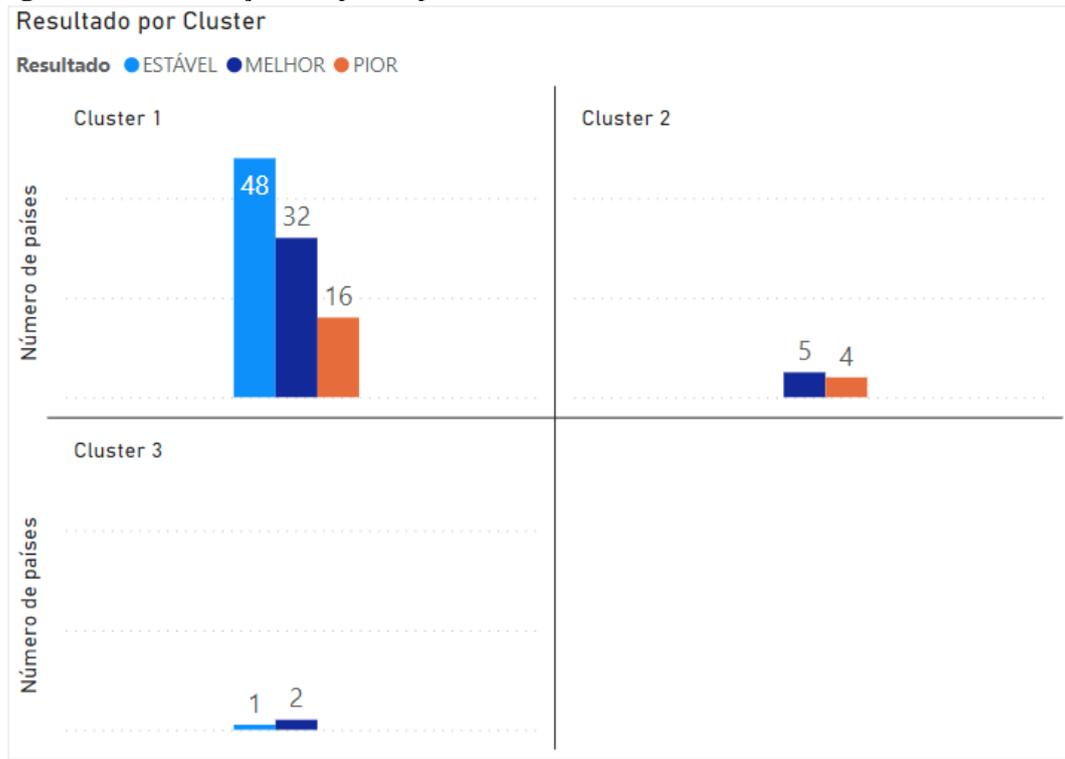
A distribuição dos países levando em consideração essa classificação ficou da seguinte forma:

**Figura 4** - Classificação geral dos países de acordo com os resultados obtidos em 2021



Fonte: Autoria própria, 2022.

**Figura 5** - Classificação dos países por *cluster* de acordo com os resultados obtidos em 2021



Fonte: Autoria própria, 2022.

Com esses gráficos é possível perceber que uma quantidade significativa de países, cerca de 55%, teve um desempenho fora do padrão dos anos anteriores, seja para melhor ou para pior. Também fica perceptível que a maioria dos países que mantiveram os resultados estáveis, guardadas as devidas proporções, pertencem ao *cluster* 1, ou seja, aqueles de menor desempenho olímpico.

## MEDIDAS DE COMBATE À COVID-19

Em relação à Covid-19, foram utilizados 16 indicadores, com gradações diferentes, para análise. O período utilizado compreende os 569 dias entre 1º de janeiro de 2020 e 21 de julho de 2021 (data anterior ao início das Olimpíadas). Em alguns casos, os países não disponibilizaram dados referentes a determinados dias, então foi considerado, para esses casos, que as medidas não foram tomadas.

**Tabela 3** – Conjunto de indicadores utilizados no cálculo do índice geral

<b>Indicador</b>	<b>Número de gradações</b>
Alívio de dívidas	3
Apoio financeiro governamental	3
Campanhas públicas de informação	3
Controle de viagens internacionais	5
Controle de viagens internas	3
Fechamento de escolas	4
Fechamento de locais de trabalho	4
Rastreamento de pessoas contaminadas	3
Restrições no transporte público	3
Restrições para eventos públicos	3
Restrições para reuniões públicas	5
Restrições para sair de casa	4
Política de testagem	4
Uso obrigatório de máscara	5
Política de vacinação	6
Disponibilidade de vacina	6

Fonte: Elaboração própria, 2022.

As gradações representam os  $n$  níveis que determinado indicador pode ter recebido e, independentemente da quantidade de níveis, os mais elevados representam que as medidas foram mais rígidas. Por exemplo, quando falamos das campanhas públicas de informação, os países poderiam receber, a cada dia, uma das três diferentes classificações, sendo elas: inexistência de campanhas; campanhas isoladas; e campanhas coordenadas.

Para cada indicador foi desenvolvido um índice para representar por quanto tempo e em quais níveis cada país aplicou determinada medida. Esse índice levou em consideração a quantidade de dias que cada gradação do indicador foi utilizada, dando-se um peso maior para níveis mais altos (de 1 a  $n$ ). O índice foi calculado através de uma média ponderada e, a fim de que todos estivessem no mesmo intervalo (0 a 100) e facilitasse a visualização, dividido por 10.

Os índices individuais de cada medida foram somados para obter um índice geral para cada país. É válido destacar que esse índice visa mostrar o quão rígidas foram as medidas de cada país, ou seja, não necessariamente indica se essas medidas foram efetivas. Assim, não é possível afirmar que um país com índice maior tenha sido mais efetivo no combate à pandemia do que um com índice menor.

## ANÁLISE ESTATÍSTICA

Tendo definido e classificado quais países seriam analisados e calculado um índice geral em relação à severidade das medidas de combate à Covid-19, foi possível iniciar as análises a que este trabalho se propunha. Diferentes hipóteses foram criadas e testadas para verificar a possível relação entre os resultados olímpicos e o rigor das medidas tomadas por cada país.

O primeiro teste foi realizado para verificar o percentual de variação entre os índices dos países de forma geral, não considerando os *clusters*. Dessa forma temos:

$$\begin{cases} H_0: \text{o coeficiente de variação dos índices é maior que 26\%} \\ H_1: \text{o coeficiente de variação dos índices não é maior que 26\%} \end{cases}$$

O coeficiente de variação encontrado foi de aproximadamente 15% (baixo), indicando que a  $H_0$  não é verdadeira. Esse resultado evidencia que não existe uma variação tão grande entre os índices ao se fazer uma comparação geral.

O próximo teste foi semelhante, porém realizando a verificação da variação dentro de cada *cluster* de forma isolada:

$$\begin{cases} H_0: \text{o coeficiente de variação dos índices dentro de cada } cluster \text{ é maior que } 26\% \\ H_1: \text{o coeficiente de variação dos índices dentro de cada } cluster \text{ não é maior que } 26\% \end{cases}$$

Analisando cada *cluster* separadamente, o resultado se mantém, indicando que não existe uma variação grande dentro dos *clusters* ( $H_0$  falsa). Os valores são diferentes entre si, mas continuam dentro da variação considerada baixa. É importante considerar que o número de países em cada *cluster* varia bastante, mas realizando o mesmo procedimento para todos, o resultado foi o exposto na tabela a seguir:

**Tabela 4** – Variações encontradas para cada *cluster* individualmente

<i>Cluster</i>	Variação
1	14,97%
2	11,20%
3	13,41%

Fonte: Elaboração própria, 2022.

Também foi realizado um teste para verificar se a média dos índices de cada *cluster* é diferente da média geral. O intuito era verificar se a baixa variação se manteria utilizando uma análise diferente:

$$\begin{cases} H_0: \text{a média dos índices de cada } cluster \text{ é diferente da média geral} \\ H_1: \text{a média dos índices de cada } cluster \text{ não é diferente da média geral} \end{cases}$$

**Figura 6** – Média dos índices e resultado do teste t por *cluster*

Descriptives					
	N	Mean	SD	Minimum	Maximum
Cluster 1	96	30.0	4.49	11.9	38.9
Cluster 2	9	30.8	3.45	26.8	36.6
Cluster 3	3	34.0	4.56	29.6	38.7

One Sample T-Test				
		Statistic	df	p
Cluster 1	Student's t	-0.388	95.00	0.699
Cluster 2	Student's t	0.534	8.00	0.608
Cluster 3	Student's t	1.466	2.00	0.280

Fonte: Elaboração própria, 2022.

Uma vez que o valor p, visível na tabela acima, é maior que 0,05, podemos concluir que não existe diferença significativa entre a média dos índices de cada *cluster* e a média geral ( $H_0$  falsa). Porém, é importante destacar que o *cluster* 1, que representa os países de baixo desempenho, possui amostras com índices muito menores que a média. Esses valores são baixos porque representam países que não possuíam informações referentes a todos os indicadores agrupados e, por consequência, receberam um valor padrão mínimo. Uma vez que a quantidade desses países sem indicadores é pequena (2 no total), a remoção deles não impactou significativamente na média.

Também foi realizado um teste levando em consideração a classificação que indicou se os resultados de 2021, em comparação com os anos anteriores, foram estáveis, melhores ou piores:

$$\begin{cases} H_0: \text{a média dos índices dos grupos é diferente da média geral} \\ H_1: \text{a média dos índices dos grupos não é diferente da média geral} \end{cases}$$

**Figura 10** – Média dos índices e resultado do teste t de acordo com os resultados de 2021

Descriptives				
	N	Mean	Minimum	Maximum
ESTÁVEL	49	30.3	11.9	38.9
MELHOR	25	32.6	29.1	38.7
PIOR	20	29.2	18.0	37.0

One Sample T-Test				
		Statistic	df	p
ESTÁVEL	Student's t	0.319	48.0	0.751
MELHOR	Student's t	5.235	24.0	< .001
PIOR	Student's t	-0.764	19.0	0.454

Fonte: Elaboração própria, 2022.

Neste teste, diferentemente do anterior, é possível perceber que houve uma diferença entre as médias ( $H_1$  verdadeira), mas só para o grupo de países que obtiveram um resultado melhor nos jogos de 2021.

## CONSIDERAÇÕES FINAIS E TRABALHOS FUTUROS

O objetivo deste trabalho era analisar a relação entre os resultados olímpicos dos países nos jogos de 2021 e o rigor adotado nas medidas de combate e prevenção à pandemia de Covid-19. Sabendo da existência de estudos mostrando uma influência negativa de determinadas medidas no desempenho de atletas, entendeu-se que analisar como os países adotaram medidas de forma geral poderia indicar resultados semelhantes. O índice criado e utilizado nos testes foi idealizado para que todas as medidas fossem consideradas de forma equivalente. É importante destacar que esse índice representa o quão rígida foi a aplicação dessas medidas e não necessariamente se elas foram eficazes no combate à pandemia.

As variações encontradas, apesar de diferentes em cada grupo, foram baixas para todos eles, não permitindo assumir como verdadeira a relação entre as medidas e os resultados. Por mais que tenha sido observado que a média dos índices de países que obtiveram resultados melhores em 2021 foi diferente da média geral, também não é possível afirmar uma relação de causa e efeito. A diferença foi pequena e pode ter sido influenciada pelo quantidade de países em cada grupo (aqueles com resultados melhores

representam apenas 23% da amostra). Reforça essa ideia o fato de que, ao analisar os países individualmente, percebe-se índices semelhantes distribuídos entre todos os grupos, ou seja, os maiores e menores índices não são exclusivos de determinado grupo.

Apesar do foco deste trabalho não ter sido analisar a relação entre uma medida específica e os resultados olímpicos, algumas hipóteses foram levantadas a respeito do tema. Caso houvesse um agrupamento das medidas de acordo com características em comum e as análises fossem realizadas levando em consideração esses grupos menores, teríamos um resultado diferente? Levar em consideração uma única medida ou grupos menores, por exemplo, aquelas que impactam diretamente a capacidade de treino dos atletas, demandaria uma variedade diferente de dados, porém é uma ideia para trabalhos futuros.

Os resultados olímpicos de um país podem sofrer influência de diversos fatores. A análise desses outros fatores também não fazia parte do escopo deste trabalho, mas entende-se que, quanto mais dados puderem ser agregados em uma análise, mais precisos tendem a ser os resultados. Sendo a pandemia de Covid-19 um assunto recente e em aberto, espera-se que muitos mais dados sejam disponibilizados sobre o tema. Somar esses dados aos já disponíveis, bem como aqueles referentes a outros fatores que impactam nos resultados olímpicos e refazer as análises também é uma proposta de trabalho futuro.

## REFERÊNCIAS

ANDRADE, A. Olimpíadas de Tóquio arrecadam US\$ 3,1 bilhões em patrocínios. **Poder 360**, 2021. Disponível em: <https://www.poder360.com.br/olimpiada/olimpiadas-de-toquio-arrecadam-us-31-bilhoes-em-patrocínios/>. Acesso em: 17 de set. de 2021.

BUSSAB, W. O.; MORETTIN, P. A. **Estatística Básica**. 6ª ed. São Paulo: Saraiva, 2010.

EVERITT, B.; LANDAU, S.; LEESE, M. **Cluster Analysis**. 5ª ed. Chichester: Wiley, 2011.

GOMES, F. P. **Curso de Estatística Experimental**. 13ª ed. São Paulo: Nobel, 1990.

GREENE, N. 2018. The Great Olympics Debate: Should We Rank Countries by Gold Medals or Total Medal Count? **Slate**, 2018. Disponível em: <<https://slate.com/culture/2018/02/should-we-rank-olympic-performance-by-gold-medals-or-total-medal-count.html>>. Acesso em: 18 de jan. de 2022.

JAGIM, A. R.; LUEDKE, J.; FITZPATRICK, A.; WINKELMAN, G.; ERICKSON, J. L.; ASKOW, A. T; CAMIC, C. L. The Impact of COVID-19-Related Shutdown Measures on the Training Habits and Perceptions of Athletes in the United States: A Brief Research Report. **Frontiers in Sports and Active Living**, p. 208, 2020.

JAIN, A. K.; DUBES, R. C. **Algorithms for Clustering Data**. 1ª ed. Nova Jersey: Prentice-Hall, 1988.

LAPORTA, T. Olimpíadas Tóquio 2021: de patrocínio a medalhas, os bilhões gastos na pandemia. **Invest News**, 2021. Disponível em: <<https://investnews.com.br/negocios/olimpiadas-2021-toquio-bilhoes-gastos-pandemia/>>. Acesso em: 16 de set. de 2021.

LIES, E. Factbox: Money, money, money: the cost of Tokyo's pandemic-delayed Olympics. **Reuters**, 2021. Disponível em: <<https://www.reuters.com/business/media-telecom/money-money-money-cost-tokyos-pandemic-delayed-olympics-2021-06-10/>>. Acesso em: 21 de maio de 2022.

MACQUEEN, J. B. (1967). Some Methods for classification and Analysis of Multivariate Observations. **5th Berkeley Symp. Math. Statist. Probability**. 1967. p. 281-297.

MONTGOMERY, D. C.; RUNGER, G. C. (2009), **Estatística Aplicada e Probabilidade para Engenheiros**. 4ª ed. Rio de Janeiro: LTC, 2009.

RITCHIE, H.; MATHIEU, E.; RODÉS-GUIRAO, L.; APPEL, C.; GIATTINO, C.; ORTIZ-OSPINA, E.; HASSEL, J.; MCDONALD, B.; BELTEKIAN, D.; ROSER, M. Coronavirus Pandemic (COVID-19). **Our World in Data**, 2020. Disponível em: <<https://ourworldindata.org/coronavirus>>. Acesso em: 12 de setembro de 2021.

ROUSSEEUW, P.J. Silhouettes: A Graphical Aid to the Interpretation and Validation of Cluster Analysis. **Journal of computational and applied mathematics**, v. 20, p. 53-65, 1987.

SHALAL, A. IMF sees cost of COVID pandemic rising beyond \$12.5 trillion estimate. **Reuters**, 2022. Disponível em: <<https://www.reuters.com/business/imf-sees-cost-covid-pandemic-rising-beyond-125-trillion-estimate-2022-01-20/>>. Acesso em: 21 de maio de 2022.

Tukey, J. W. **Exploratory Data Analysis**. 1ª ed. Massachusetts: Addison-Wesley, 1977.

URBAŃSKI, P.; SZELIGA, Ł.; TASIEMSKI, T. Impact of COVID-19 pandemic on athletes with disabilities preparing for the Paralympic Games in Tokyo. **BMC research notes**, v. 14, n. 1, p. 1-5, 2021.

XU, R.; WUNSCH, D. **Clustering**. 10ª ed. Nova Jersey: John Wiley & Sons, 2008.

*Recebido em: 06/07/2022*

*Aprovado em: 12/08/2022*

*Publicado em: 18/08/2022*