

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

ARTHUR M. S DE OLIVEIRA

FELIPE ALVES BARBOZA

**BIG FIVE E COVID-19: UMA ANÁLISE DE DADOS DE PERFIS
PSICOMÉTRICOS DE PAÍSES E SUA RELAÇÃO COM A PANDEMIA**

CURITIBA

2022

**ARTHUR M. S DE OLIVEIRA
FELIPE ALVES BARBOZA**

**BIG FIVE E COVID-19: UMA ANÁLISE DE DADOS DE PERFIS
PSICOMÉTRICOS DE PAÍSES E SUA RELAÇÃO COM A PANDEMIA**

**BIG FIVE AND COVID-19: A DATA ANALYSIS OF PSYCHOMETRIC PROFILES
OF COUNTRIES AND THEIR RELATIONSHIP TO THE PANDEMIC**

Trabalho de Conclusão de Curso de Graduação apresentado como requisito para obtenção do título de Bacharel em Engenharia da Computação do Curso de Bacharelado em Engenharia da Computação da Universidade Tecnológica Federal do Paraná.

Orientador: Prof. Dr. Luiz Celso Gomes Jr.

Coorientador: Prof^a. Dra^a. Rita Berardi

CURITIBA

2022



[4.0 Internacional](https://creativecommons.org/licenses/by/4.0/)

Esta licença permite compartilhamento, remixe, adaptação e criação a partir do trabalho, mesmo para fins comerciais, desde que sejam atribuídos créditos ao(s) autor(es). Conteúdos elaborados por terceiros, citados e referenciados nesta obra não são cobertos pela licença.

**ARTHUR M. S DE OLIVEIRA
FELIPE ALVES BARBOZA**

**BIG FIVE E COVID-19: UMA ANÁLISE DE DADOS DE PERFIS
PSICOMÉTRICOS DE PAÍSES E SUA RELAÇÃO COM A PANDEMIA**

Trabalho de Conclusão de Curso de Graduação apresentado como requisito para obtenção do título de Bacharel em Engenharia da Computação do Curso de Bacharelado em Engenharia da Computação da Universidade Tecnológica Federal do Paraná.

Data de aprovação: 15/Dezembro/2022

Luiz Celso Gomes Júnior
Doutorado
Universidade Tecnológica Federal do Paraná

Rita Cristina Galarraga Berardi
Doutorado
Universidade Tecnológica Federal do Paraná

Leyza Baldo Dorini
Doutorado
Universidade Tecnológica Federal do Paraná

Luiz Guilherme Roth Cantarelli
Mestrado
Universidade de São Paulo

CURITIBA
2022

AGRADECIMENTOS

Somos extremamente gratos a todos que nos ajudaram na nossa formação acadêmica e agradecemos: ao Luiz Celso e à Rita Berardi, que foram a responsáveis por orientar nosso trabalho – obrigado por esclarecer inúmeras dúvidas e serem tão pacientes. Aos professores do curso de Engenharia da computação, que trilharam nossa formação e nos guiaram até esse momento. Aos amigos que fizemos no caminho. À UTFPR, que proporcionou nossa formação acadêmica. Às nossas famílias, que independente do momento são alicerces centrais. E em especial aos nossos pais, que se não fosse o apoio deles nada disso seria possível. Por fim, a todos que direta ou indiretamente fizeram parte de nossa formação, o nosso muito obrigado!

RESUMO

Traços de personalidade, no nível do indivíduo, são características subjetivas importantes, que em parte explicam as diferenças comportamentais observadas entre seres humanos. Em níveis agregados, evidências apontam que tais traços podem ser utilizados como uma forma de caracterizar diferentes culturas e que, inclusive, apresentam correlações com índices de liberdade política de sociedades. Em momentos críticos como o da atual pandemia de Covid-19, é importante compreender os fatores que podem influenciar o comportamento da população. No presente trabalho, foi investigada a relação entre os perfis psicométricos de diferentes países - analisados por meio de traços de personalidade agregados - e variáveis relacionadas às respostas à pandemia do novo coronavírus. Foram testadas três hipóteses por meio de diferentes abordagens estatísticas: a primeira destas consistiu em determinar se existe algum relacionamento significativo entre traços de personalidade de nações e taxas nacionais de vacinação por meio de modelos de regressão linear. A segunda hipótese procurou determinar a efetividade de políticas de isolamento social com base em perfis de personalidade. Técnicas de clusterização hierárquica foram aplicadas para a identificação de países com perfis psicométricos parecidos. Em seguida, por modelos de regressão espacial, a efetividade das políticas de restrição de mobilidade adotadas pelos países foi analisada por meio de regimes espaciais definidos pelos *clusters* encontrados. A terceira hipótese investigou, fazendo uso de modelos de regressão em dados de painel, a interação entre a resposta às políticas de restrição de mobilidade e os traços de personalidade nacionais.

Palavras-chave: big five; covid-19; análise; regressão.

ABSTRACT

Personality traits, at the level of the individual, are important subjective characteristics, which in part explain the observed behavioral differences between human beings. At aggregated levels, evidence points out that such traits can be used as a way of characterizing different cultures and that they even present correlations with indices of political freedom of societies. At critical moments such as the current Covid-19 pandemic, it is important to understand the factors that can influence population behavior. In the present work, the relationship between the psychometric profiles of different countries - analyzed using aggregated personality traits - and variables related to responses to the pandemic of the new coronavirus was investigated. Three hypotheses were tested using different statistical approaches: the first of these was to determine whether there is any significant relationship between personality traits of nations and national vaccination rates using linear regression models. The second hypothesis sought to determine the effectiveness of social isolation policies based on personality profiles. Hierarchical clustering techniques were applied to identify countries with similar psychometric profiles. Then, using spatial regression models, the effectiveness of mobility restriction policies adopted by countries was analyzed using spatial regimes defined by the *clusters* found. The third hypothesis investigated, using regression models on panel data, the interaction between the response to mobility restriction policies and national personality traits.

Keywords: big five; covid-19; analysis; regression.

LISTA DE FIGURAS

Figura 1 – Mapa de calor para vacinação	28
Figura 2 – Dendrograma dos traços de personalidade por país	29
Figura 3 – Dendrograma dividido entre os dois <i>clusters</i>	31
Figura 4 – <i>Violin plot</i> dos traços de cada <i>cluster</i>	32
Figura 5 – Gráfico da Interação entre Abertura à Experiência e C1	36
Figura 6 – Gráfico da Interação entre Extroversão e C7	36
Figura 7 – Mapa de calor para Agradabilidade	43
Figura 8 – Mapa de calor para Conscienciosidade	43
Figura 9 – Mapa de calor para Extroversão	44
Figura 10 – Mapa de calor para Neuroticismo	44
Figura 11 – Mapa de calor para Abertura à Experiência	45
Figura 12 – Gráfico da Interação entre Neuroticismo e C2	47
Figura 13 – Gráfico da Interação entre Neuroticismo e C7	47

LISTA DE TABELAS

Tabela 1 – Países analisados e suas informações de personalidade e vacinação	26
Tabela 2 – Modelos de regressão linear	30
Tabela 3 – Resultado da regressão espacial com regimes espaciais	33
Tabela 4 – Resultado do Teste de Chow apresentado a diferença global entre os regimes	33
Tabela 5 – Resultado dos Testes de Chow para os coeficientes da regressão	33
Tabela 6 – Regressão em dados de painel com efeitos de tempo	35

LISTA DE ABREVIATURAS E SIGLAS

Siglas

A	Agradabilidade
C	Conscienciosidade
E	Extroversão
IDH	Índice de Desenvolvimento Humano
N	Neuroticismo
O	Abertura à Experiência
OLS	<i>Ordinary Least Squares</i>
OxCGRT	Projeto <i>Oxford Covid-19 Government Response Tracker</i>
PFVPH	<i>People Fully Vaccinated per hundred</i>
PIB	Produto Interno Bruto
PVPH	<i>People Vaccinated per hundred</i>

SUMÁRIO

1	INTRODUÇÃO	11
2	REFERENCIAL TEÓRICO	13
2.1	Big Five	13
2.2	Conceitos estatísticos	13
2.2.1	Regressão linear	14
2.2.2	Regressões espaciais	16
2.2.3	Regressão linear em dados de painel	16
2.3	Clusterização hierárquica	17
3	TRABALHOS RELACIONADOS	19
4	METODOLOGIA	21
4.1	Variáveis utilizadas nos modelos	21
4.1.1	Fatores de personalidade agregados por país	21
4.1.2	Índices de vacinação dos países	21
4.1.3	Informações gerais de países – IDH, PIB e Coeficiente de Gini	22
4.1.4	Mobilidade	22
4.1.5	Indicadores de medidas governamentais	22
4.2	Análise Exploratória	23
4.3	Primeira Hipótese: Influência de fatores psicométricos na vacinação	23
4.4	Segunda Hipótese: Efetividade de medidas de contenção em função de <i>clusters</i> de personalidade	24
4.5	Terceira Hipótese: Interações entre traços de personalidade nacionais e a intensidade de políticas de contenção	24
5	RESULTADOS	26
5.1	Análise Exploratória	26
5.1.1	Visualização dos traços de personalidade nacionais	27
5.1.2	Visualização dos dados de vacinação por país	27
5.1.3	Clusterização hierárquica dos traços de personalidade agregados	28
5.1.4	Correlação entre vacinação e traços de personalidade por país	29
5.2	Modelagem e implementação	30
5.2.1	Primeira Hipótese: Influência de fatores psicométricos na vacinação	30

5.2.2	Segunda Hipótese: Efetividade de medidas de contenção em função de <i>clusters</i> de personalidade	31
5.2.3	Terceira Hipótese: Interações entre traços de personalidade nacionais e a intensidade de políticas de contenção	34
6	CONCLUSÃO	37
6.1	Resultados obtidos	37
6.2	Limitações do presente estudo	38
6.3	Trabalhos futuros	39
	REFERÊNCIAS	40
	APÊNDICE A MAPAS DE CALOR MUNDIAL DE TODOS OS TRAÇOS DE PERSONALIDADE	43
	APÊNDICE B GRÁFICOS DE INTERAÇÃO ENTRE INDICADORES DE POLÍTICAS DE CONTENÇÃO E TRAÇOS DE PERSONALIDADE	47

1 INTRODUÇÃO

A compreensão do comportamento de pessoas em contextos de crises sanitárias como o da pandemia da COVID-19 é um assunto de grande interesse para órgãos de saúde e entidades governamentais, pois possibilita que políticas públicas sejam tomadas de maneira mais eficiente. O estudo de Dabla-Norris, Khan e Lima (2021), por exemplo, aponta que indivíduos que apresentam maior confiança nas medidas preventivas implementadas por seus governos tendem a relatar maior adesão a comportamentos protetivos, como o uso de máscaras e isolamento social; além disto, diferenças significativas entre países são observadas nos índices de adesão a medidas protetivas. Torna-se importante, portanto, entender quais são os fatores responsáveis por influenciar nos comportamentos observados, tanto no nível individual quanto no nível de nações.

Por se tratar de uma crise global, a caracterização dos comportamentos que influenciam no seu aprofundamento ou contenção apresenta-se como uma tarefa difícil, uma vez que - dentre outros fatores complexos - estes são distribuídos geograficamente de forma heterogênea. E as diversas características socioeconômicas que compõem países, independentemente de atitudes individuais, respondem direta e substancialmente pela efetividade com a qual a crise será administrada.

Diferentemente de outras análises que investigaram a relação entre comportamentos e indicadores da pandemia, no presente estudo optou-se pela caracterização de países por meio de seus traços de personalidades agregados. Técnicas comumente empregadas consistem na elaboração de modelos de regressão que consideram apenas respostas a níveis de indivíduos e as correlacionam com alguma variável de interesse, sem contabilizar efeitos oriundos de relações geográficas e temporais. A taxonomia de personalidade dos cinco grande fatores (Abertura à Experiência, Conscienciosidade, Extroversão, Agradabilidade e Neuroticismo), no entanto, possibilita que um país seja caracterizado por sua "personalidade média". Neste sentido, partimos da premissa de que as heranças culturais e genéticas existentes entre diferentes países - traduzidas em parte pelos traços nacionais de personalidade - exercem influência significativa no comportamento dos cidadãos em uma situação desafiadora como a pandemia da COVID-19.

Neste trabalho, três hipóteses foram testadas por meio de modelos com diferentes abordagens estatísticas. A primeira hipótese (Seção 4.3) consistiu em determinar se existe algum relacionamento significativo entre traços de personalidade de nações e taxas nacionais de vacinação. Para tanto, modelos de regressão linear multivariáveis foram implementados tendo como variáveis independentes as médias nacionais de cada traço da taxonomia Big-Five, e as taxas nacionais de vacinação como variável dependente. Variáveis de controle consideradas relevantes para a análise - como o coeficiente de Gini (BANK, 2020), Produto Interno Bruto (PIB) (BANK, 2021) e Índice de Desenvolvimento Humano (IDH) (UNDP, 2019) dos países sob análise - foram incluídas nos modelos.

Na segunda hipótese (Seção 4.4) foi investigada a efetividade de medidas governamentais de contenção da pandemia em função de *clusters* de personalidade. Inicialmente, técnicas de clusterização hierárquica foram aplicadas aos dados de personalidade a fim de identificar países com perfis de personalidade similares. Em seguida elaborou-se um modelo de regressão espacial fazendo uso do conceito de regimes espaciais definidos pelos *clusters* obtidos. Este modelo teve como variáveis dependentes os dados de variação de mobilidade do *Google Mobility Reports* (GOOGLE, 2022b), e índices de restrição de mobilidades do projeto (HALE *et al.*, 2021) como variáveis independentes.

Por fim, a terceira hipótese (Seção 4.5) teve como objetivo analisar a influência da interação entre traços de personalidade e intensidade das políticas de contenção nas variações de mobilidade. A hipótese foi testada por meio de modelos de regressão linear em dados de painel tendo como variáveis dependentes os dados de variação de mobilidade e como variáveis independentes os índices de restrição de mobilidade e o produto destes pelos traços agregados de personalidade.

O restante do texto está organizado da seguinte forma: no Capítulo 2 são apresentados a fundamentação teórica e conceitos básicos utilizados ao longo deste trabalho. No Capítulo 3, discute-se sobre estudos relacionados ao tema em questão e as diferentes abordagens adotadas por seus autores. A metodologia do estudo, descrição da implementação dos modelos e suas variáveis, e detalhes gerais das hipóteses testadas são descritas no Capítulo 4. O Capítulo 5 apresenta os resultados dos modelos implementados bem como a visualização dos dados sob análise. Finalmente, no Capítulo 6, discutem-se os resultados obtidos e elencam-se as limitações do presente estudo.

2 REFERENCIAL TEÓRICO

2.1 Big Five

A taxonomia do Big Five surge dentro da psicologia e decompõe a personalidade de um indivíduo em cinco grandes fatores, os quais podem ser derivados a partir de um questionário simples. Trata-se de um modelo amplamente estabelecido na área da psicometria, que se origina a partir da análise léxica de termos, em linguagem natural, utilizados pelas pessoas para descreverem umas às outras (JOHN; SRIVASTAVA *et al.*, 1999). Os traços podem ser brevemente descritos como (JOHN; SRIVASTAVA *et al.*, 1999)(VEDEL; POROPAT, 2017):

- Extroversão: traço associado a entusiasmo, senso de aventura, assertividade, sociabilidade e expressão de emoções positivas.
- Agradabilidade: traço associado a comportamentos pró-sociais, atitudes cooperativas, ternura, gentileza, confiança em outras pessoas e modéstia.
- Abertura à Experiência: traço associado à originalidade de pensamento, criatividade, interesse artístico, procura por novas experiências - principalmente orientadas a estímulos cognitivos - e curiosidade.
- Neuroticismo: traço associado à tendência de experimentar emoções negativas, ansiedade, nervosismo, tristeza, irritabilidade e medo.
- Conscienciosidade: traço relacionado à auto-disciplina, controle de impulsos, planejamento e estabelecimento de objetivos, adesão à regras, organização e priorização de tarefas.

2.2 Conceitos estatísticos

Na ciência de dados, a escolha dos modelos a serem desenvolvidos para análise de um problema deve ser feita com base na sua adequabilidade e interpretabilidade. Muitas vezes, ao tratar de problemas não-familiares, técnicas de análise de dados exploratórias são empregadas para informar a tomada de decisões a respeito dos modelos. Em outras circunstâncias, no caso de problemas de natureza bem estabelecida, técnicas específicas e tradicionalmente aplicadas na área de estudo do problema são utilizadas. Nesta seção são apresentados conceitos fundamentais a respeito das técnicas estatísticas utilizadas ao longo do desenvolvimento deste trabalho.

2.2.1 Regressão linear

Modelos de regressão linear são amplamente utilizados na ciência de dados e em áreas de pesquisa que envolvem grandezas quantitativas como a medicina e a psicologia. Tais modelos servem como base para a implementação de outros de maior complexidade - muitas vezes, técnicas não-lineares podem ser desenvolvidas por meio da generalização de técnicas lineares (HASTIE; TIBSHIRANI; FRIEDMAN, 2009). Sua principal propriedade, a linearidade, permite que a relação analisada seja implementada e interpretada de forma simples, o que em boa parte justifica a escolha por tais modelos. E uma vez que todos os modelos de regressão desenvolvidos neste trabalho são fundamentados por conceitos de regressão linear, torna-se importante evidenciar sua fundamentação matemática básica e as métricas estatísticas utilizadas para analisar sua adequabilidade. A estrutura geral de uma regressão linear multivariável, com $i = 1, \dots, n$ observações e k regressores pode ser descrita da seguinte maneira:

$$Y_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_k x_{ik} + \epsilon_i \quad (1)$$

E em forma matricial:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \quad (2)$$

Onde Y representa a variável dependente (regressante); β_j , $j = 0, 1, \dots, k$, os coeficientes da regressão x_i , as variáveis independentes (regressores), e o termo ϵ_i (épsilon), o erro entre os valores real e estimado. Diversas técnicas podem ser empregadas para se estimar os coeficientes da regressão. Neste trabalho, foram utilizadas implementações que fazem uso do método dos mínimos quadrados (*Ordinary Least Squares* (OLS)) (MONTGOMERY; RUNGER, 2018). Este método popularmente empregado consiste em minimizar a soma dos quadrados dos termos de erro:

$$L(\boldsymbol{\beta}) = \sum_{i=1}^n \epsilon_i^2 = \sum_{j=1}^n (Y_i - \beta_0 - \sum_{i=1}^k \beta_i x_{ij})^2 \quad (3)$$

Sendo que a equação 3 pode ser representada na forma matricial:

$$L(\boldsymbol{\beta}) = (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})^T (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}) \quad (4)$$

Uma vez que L é uma função de $\boldsymbol{\beta}$ - a matriz de coeficientes a serem estimados -, a minimização da soma dos quadrados dos termos de erro é realizada tomando-se a derivada de L em relação a $\boldsymbol{\beta}$:

$$\frac{\partial L(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = -2[\mathbf{X}^T (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})]^T \quad (5)$$

Os coeficientes estimados $\hat{\beta}$ são então obtidos igualando-se a equação 5 a zero:

$$\mathbf{X}^T(\mathbf{Y} - \mathbf{X}\beta) = 0 \quad (6)$$

$$\hat{\beta} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y} \quad (7)$$

Obtidos os coeficientes estimados, elabora-se o modelo ajustado:

$$\hat{\mathbf{Y}} = \mathbf{X}\hat{\beta} \quad (8)$$

Testes de significância da regressão são então conduzidos para determinar se há de fato um relacionamento linear entre regressor e regressantes. Para regressões lineares multivariáveis a testagem de hipóteses é geralmente feita sobre cada coeficiente β_i , assumindo como hipóteses nula e alternativa, respectivamente: $H_0 : \beta_i = 0$, $H_1 : \beta_i \neq 0$. Nesse caso, a estatística de testes utilizada para testar a hipótese corresponde à (MOLNAR, 2020):

$$t_0 = \frac{\hat{\beta}_i}{SE(\hat{\beta}_i)} \quad (9)$$

sendo $SE(\hat{\beta}_i)$ o erro padrão do coeficiente estimado $\hat{\beta}_i$. A relevância estatística de coeficientes pode ser determinada através do conceito de *P-values*, que são calculados por meio da soma das probabilidades nas duas caudas da distribuição t com $n - 1$ graus de liberdade (MONTGOMERY; RUNGER, 2018): $p = 2P(T_{n-1} > |t_o|)$. O conceito de *P-value*, amplamente empregado na estatística - e particularmente importante para a interpretação dos resultados obtidos neste trabalho - pode ser definido como o menor nível de significância que levaria à rejeição da hipótese nula (MONTGOMERY; RUNGER, 2018), de modo que, quanto menor for seu valor, maior é a credibilidade da hipótese alternativa.

Outras estatísticas comumente empregadas para se examinar a adequabilidade do modelo consistem no coeficiente de determinação R^2 , e no R^2 ajustado (R^2_{adj}), os quais são obtidos respectivamente por meio das relações:

$$R^2 = 1 - \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \quad (10)$$

$$R^2_{adj} = 1 - (1 - R^2) \frac{n - 1}{n - (k + 1)} \quad (11)$$

sendo \hat{Y}_i , os valores de Y obtidos por meio do modelo ajustado, \bar{Y} , a média das observações de Y e k , o número de regressores do modelo. O R^2 pode ser percebido como uma medida do quanto o modelo ajustado se adequa aos dados analisados, de modo que, $R^2 = 1$ representaria um modelo perfeitamente ajustado e $R^2 = 0$, um modelo que não explica nenhuma variação de

Y . Nota-se que à medida que variáveis são introduzidas ao modelo o seu R^2 tende a aumentar. Em virtude disto, o R^2 ajustado atua como uma estatística utilizada para controlar este efeito, penalizando o aumento provocado pela adição excessiva de variáveis.

2.2.2 Regressões espaciais

Modelos de regressão espacial consistem em técnicas estatísticas que introduzem estruturas espaciais e geográficas às análises de regressão (REY; ARRIBAS-BEL; WOLF, 2022). Tais modelos são especialmente úteis quando o objeto de estudo possui uma estrutura espacial explícita - como países, bairros e vizinhanças -, pois permitem que efeitos subjacentes, oriundos de relações geográficas, sejam devidamente incorporados na análise. No presente trabalho, implementou-se um modelo que faz uso do conceito de regimes espaciais, o qual permite que os regressores sejam relacionados a determinados grupos (regimes) a fim de que padrões heterogêneos, distribuídos geograficamente, sejam capturados pelo modelo. Uma regressão de regimes espaciais com n observações e p regimes pode ser expressa pela seguinte relação linear:

$$Y_{ij} = \alpha_j + x_{ij}\beta_j + \epsilon_{ij} \quad (12)$$

de modo que $i = 1, \dots, n$ e $j = 1, \dots, p$. Sendo que cada regime possui a sua própria constante α_j e coeficientes de regressão β_j . Um modelo simplificado com dois regimes pode ser representado em forma matricial como (ANSELIN; REY, 2014):

$$\begin{bmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{X}_1 & 0 \\ 0 & \mathbf{X}_2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \end{bmatrix} \quad (13)$$

sendo \mathbf{Y}_1 e \mathbf{Y}_2 matrizes de observações da variável dependente, de dimensões $n_1 \times 1$ e $n_2 \times 1$ respectivamente. \mathbf{X}_1 e \mathbf{X}_2 , de dimensões $n_1 \times k$ e $n_2 \times k$, são matrizes de observações de k variáveis independentes, β_1 e β_2 são as matrizes dos coeficientes da regressão, de dimensões $k \times 1$. E as matrizes ϵ_1 e ϵ_2 , de dimensões $n_1 \times 1$ e $n_2 \times 1$ representam os termos de erro da regressão.

2.2.3 Regressão linear em dados de painel

Dados de painel são compostos por observações longitudinais de entidades ao longo de uma série temporal. Assim, diferentemente de uma regressão linear padrão - que lida apenas com o estado de entidades em um instante específico -, regressões em dados de painel permitem que a dimensão temporal seja incorporada na análise. A forma geral de uma regressão em

dados de painel, multivariável pode ser expressada como:

$$Y_{it} = \beta_0 + \sum_{j=1}^k \beta_j x_{j,it} + \epsilon_{it} \quad (14)$$

onde $i = 1, \dots, n$, sendo n o total de observações, e $t = 1, \dots, T$, os instantes de tempo em que as observações foram realizadas. Estas regressões podem ser modeladas de modo a observar efeitos fixos ao longo do tempo e/ou entidades. Isso pode ser realizado por meio da adição de variáveis *dummy* (binárias) para cada entidade e período de tempo analisados (HANCK *et al.*, 2021):

$$Y_{it} = \beta_0 + \sum_{j=1}^k \beta_j x_{j,it} + \sum_{j=2}^n \gamma_j G_{j,i} + \sum_{j=2}^T \delta_j D_{j,t} + \epsilon_{it} \quad (15)$$

sendo G_j e D_j os regressores *dummy* que representam, respectivamente, heterogeneidades entre entidades de $j = 1, \dots, n$ constantes ao longo do tempo, e efeitos fixos entre entidades mas variáveis ao longo dos intervalos $j = 1, \dots, T$. Os coeficientes podem ser estimados utilizando as técnicas de regressão linear e mínimos quadrados descritas na seção 2.2.1.

2.3 Clusterização hierárquica

Análises de *clusterização* podem ser definidas como métodos de aprendizado de máquina não-supervisionados que têm por objetivo dividir os dados sob análise em grupos interpretáveis e significativos (TAN; STEINBACH; KUMAR, 2013). Idealmente, membros de um mesmo *cluster* devem ser similares uns aos outros e substancialmente diferentes dos membros de outros grupos. Em sua essência, métodos de clusterização hierárquica aglomerativos organizam os dados de modo que grupos possam conter subgrupos. Partindo-se de um *cluster* inicial consistindo em um ponto isolado, os métodos aglomerativos iterativamente unem os *clusters* de maior proximidade uns dos outros, resultando, ao fim do processo, em apenas um *cluster* que contém todos os demais. Diversas métricas de distância podem ser utilizadas para determinar quais são os *clusters* de maior proximidade a serem aglomerados. No presente trabalho, para realizar a clusterização dos países em função de seus traços de personalidade agregados, optou-se pela utilização do método minimização de variância de Ward - uma vez que este apresentou os melhores resultados em termos de interpretabilidade. A métrica de distância deste método pode ser calculada da seguinte forma (VIRTANEN *et al.*, 2020):

$$d(u,v) = \sqrt{\frac{|v| + |s|}{T} d(v,s)^2 + \frac{|v| + |t|}{T} d(v,t)^2 - \frac{|v|}{T} d(s,t)^2} \quad (16)$$

Sendo $d(u,v)$ a distância entre u , um *cluster* recém formado e composto por *clusters* s e t , v , o *cluster* a ser aglomerado, e $T = |v| + |s| + |t|$, a soma das cardinalidades dos *clusters*

mencionados. As estruturas resultantes da aplicação destes métodos podem ser representada por árvores cujos nós e nós-filhos correspondem a *clusters* e *subclusters*, respectivamente. Uma forma conveniente de visualizar tais estruturas se dá por meio de *dendrogramas*, que consistem em diagramas em forma de árvores que apresentam: as relações hierárquicas entre *clusters*, a ordem em que foram aglomerados e as distâncias entre cada um deles.

3 TRABALHOS RELACIONADOS

No estudo de Götz et al. (GÖTZ *et al.*, 2021), foi analisada a interação entre os traços de personalidade de indivíduos, em micronível, e políticas públicas adotadas por governos, em macronível; a variável dependente utilizada, “Shelter-in-place”, consistiu em uma medida auto-relatada que indica o quanto os participantes da pesquisa observaram o isolamento social na última semana e varia numa escala de 0 (não se aplica) a 100 (se aplica muito). Os resultados obtidos neste estudo indicaram que todos os traços de personalidade, com exceção de Extroversão, correlacionam-se positivamente com o indicador de isolamento social. Também foi evidenciado que mesmo em países cujas medidas de combate à pandemia foram mais rígidas as diferenças individuais de personalidade foram capazes de prever comportamentos. Os dados utilizados nesta pesquisa foram obtidos do projeto “Measuring Worldwide COVID-19 Attitudes and Beliefs”¹, uma pesquisa global que avalia comportamentos e percepções dos participantes em relação à pandemia.

Noutro estudo que aborda tema semelhante, Dabla-Norris et al. (DABLA-NORRIS; KHAN; LIMA, 2021) analisaram índices de confiança e aceitação de políticas impostas por governos e sua relação com a adoção de medidas protetivas por parte de indivíduos. Os autores examinaram os dados do projeto “COVID-19 behavior tracker”², restringindo sua abordagem a 15 países. As análises realizadas revelaram que pessoas que apresentam maior confiança no governo tendem a relatar maior adesão ao uso de máscaras e isolamento social, quando os governos incentivam tais medidas. Também foram investigadas diferenças regionais nos índices de distanciamento social e uso de máscaras. Para tanto, os 15 países foram divididos nas regiões: Ásia, Américas, Europa do Norte e Europa do Sul. Diferenças regionais significativas foram observadas, sendo que, de acordo com os autores, dois terços destas são explicadas por fatores individuais e políticas governamentais. Contudo, uma análise de traços de personalidade não foi desenvolvida neste estudo. Diferentemente do dataset do estudo Götz et al. (GÖTZ *et al.*, 2021), a massa de dados de Dabla-Norris et al. (DABLA-NORRIS; KHAN; LIMA, 2021) contém, adicionalmente, perguntas relacionadas à aceitação de vacinas e apresenta maior ênfase em questões de confiança governamental. Entretanto, a disponibilidade de dados deste último é insatisfatória: diversos países de grande interesse para a análise proposta - como Brasil, Índia e China - não tiveram participação nos questionários relativos à vacinação.

Dentre artigos que realizaram análises de traços de personalidade agregados por países e sua relação com alguma variável de interesse, podemos citar o trabalho de Barceló, J. (BARCELO, 2017). Em seu estudo, o autor investiga a relação entre traços de personalidade nacionais e tipos de regimes políticos de diversos países. Medidas de personalidade do modelo Big Five de 47 países foram adquiridas do projeto de McCrae e Terracciano, “Personality profiles of cultures” (MCCRAE; TERRACCIANO, 2005), e correlacionadas com indicadores de liberdade

¹ Materiais do projeto podem ser encontrados em <https://osf.io/tuenj/>

² Dataset do projeto disponível em <https://github.com/YouGov-Data/covid-19-tracker/>

democrática. Foi observado, neste estudo, que os índices nacionais de “Abertura à Experiência” apresentam correlação estatisticamente significativa com indicadores de liberdade democrática, mesmo após a introdução de variáveis de controle.

4 METODOLOGIA

Neste trabalho são apresentadas uma análise exploratória do conjunto de dados, inicialmente, e três hipóteses principais. A primeira hipótese consiste em determinar se traços de personalidade agregados por países têm influência estatisticamente significativa nas taxas nacionais de vacinação. A segunda hipótese assume que países podem ser *clusterizados* de acordo com seus traços de personalidade nacionais e testa se a efetividade de políticas de contenção da pandemia varia em função dos perfis psicométricos de cada *cluster*. Por fim, a terceira hipótese estabelece que existem efeitos de interação entre traços de personalidade nacionais e a intensidade das políticas de contenção adotadas.

4.1 Variáveis utilizadas nos modelos

4.1.1 Fatores de personalidade agregados por país

Os valores dos traços de personalidade foram obtidos do projeto “Measuring Worldwide COVID-19 Attitudes and Beliefs” (FETZER *et al.*, 2020). Neste projeto, os traços do *Big-Five*: Abertura à Experiência (O), Conscienciosidade (C), Extroversão (E), Neuroticismo (N) e Agradabilidade (A) foram obtidos por meio da aplicação *online* do questionário *Ten-Item Personality Inventory* (GOSLING; RENTFROW; JR, 2003), cujos itens são pontuados através de uma escala Likert de sete pontos. A quantidade de respostas obtidas de cada país varia de forma expressiva – de 238 respostas na Hungria a 11568 no Brasil –, e ao todo somam em 100989.

4.1.2 Índices de vacinação dos países

A partir do repositório da Oxford “COVID-19 Dataset by Our World in Data” (MATHIEU E., 2021) – que dispõe em ordem cronológica os índices de vacinação de diferentes países – foram obtidos os dados mais recentes das taxas de vacinação dos países analisados. Esta base de dados define como “vacinados” os indivíduos que receberam pelo menos uma dose vacinal, e como “completamente vacinados” aqueles que receberam o número de doses prescrito pelo protocolo inicial de vacinação necessário para que estejam devidamente protegidos contra a doença. Uma vez que os dados de vacinação de cada país foram coletados em períodos distintos, estabeleceu-se que a data mínima para a análise partisse do dia 01/03/2022. Por apresentarem índices de vacinação muito inferiores (número de pessoas vacinadas menor que 35%) aos demais continentes, decidiu-se por retirar os dados de países Africanos das análises.

4.1.3 Informações gerais de países – IDH, PIB e Coeficiente de Gini

Valores de IDH, PIB e coeficiente de Gini foram obtidos respectivamente de *United Nations Development Programme* (UNDP, 2019) e *World Bank* (BANK, 2021), (BANK, 2020) ¹.

4.1.4 Mobilidade

O *Google COVID-19 Community Mobility Reports* apresenta tendências de variações de mobilidade para diferentes categorias de localidades (GOOGLE, 2022b). O *dataset* contém dados de geoposicionamento agregados, obtidos de dispositivos *Android* de indivíduos que habilitam a configuração *Location History*. As variações de mobilidade são calculadas em relação a um patamar base arbitrado como o valor mediano do período de cinco semanas entre os dias 3 de Janeiro a 6 de Fevereiro de 2020. As categorias de localidades são: *Grocery & pharmacy* (mercados e farmácias), *Parks* (parques), *Transit stations* (estações de trânsito), *Retail & recreation* (varejo e lazer), *Residential* (residencial) e *Workplaces* (locais de trabalho). No presente trabalho foram utilizados dados de Fevereiro de 2020 à Agosto de 2022 e, considerada como mobilidade do país a média das categorias de mobilidade desconsiderando residencial.

4.1.5 Indicadores de medidas governamentais

O projeto *Projeto Oxford Covid-19 Government Response Tracker (OxCGRT)* (*Oxford Covid-19 Government Response Tracker*) Hale *et al.* (2021) reúne diversas informações a respeito de medidas governamentais tomadas para conter o avanço da pandemia. Com início em janeiro de 2020, dados de 180 países foram coletados e indicadores foram computados com base na rigidez das políticas adotadas por governos. Os indicadores são classificados em cinco categorias: *C – containment and closure policies* (políticas de contenção e fechamento), *E – economic policies* (políticas econômicas), *H – health system policies* (políticas de sistema de saúde), *V – vaccination policies* (políticas de vacinação) e *M – miscellaneous policies* (políticas diversas). Neste trabalho, foram utilizados os indicadores da categoria “C”, que possuem maior relevância para a análise de dados de mobilidade, sendo eles:

- C1: Fechamento de escolas
- C2: Fechamento de locais de trabalho
- C3: Cancelamento de eventos públicos

¹ O Coeficiente de Gini mede a desigualdade econômica em um país. Especificamente, é a medida de quanto a distribuição de renda (ou, em alguns casos, despesas de consumo) se desvia de uma distribuição perfeitamente igual entre indivíduos ou famílias dentro de uma economia. Um coeficiente de Gini de 0 representa igualdade perfeita, enquanto um índice de 100 implica desigualdade perfeita

- C4: Restrição de encontros
- C5: Fechamento de transporte público
- C6: Confinamento domiciliar
- C7: Restrição do movimento interno

Para cada um dos países, são informadas as datas na qual a medida foi implementada e sua intensidade, sendo representada por um intervalo discreto de 0 a um número inteiro maior que 1. Sendo que 0 implica em nenhuma restrição e o limite superior a maior restrição encontrada. Para C1, por exemplo, a codificação é representada por:

- 0: Nenhuma medida
- 1: Recomendação de fechar as escolas ou manter aberta com restrições
- 2: Requerir fechamento completo de pelo menos um nível (Ex.: Fechamento de escolas públicas ou ensino fundamental/médio)
- 3: Requerir fechamento de todos os níveis

4.2 Análise Exploratória

Antes de realizar a modelagem das hipóteses, teve-se em mente produzir uma visualização dos dados obtidos. Primeiramente, investigando a distribuição dos traços de personalidade pelos países, e aplicando algoritmos de clusterização no conjunto de dados dos traços para identificar grupos de países com perfis de personalidade similares. Também foi realizada a visualização das taxas de vacinação ao redor do globo e identificadas, por meio de uma matriz de correlações, as associações das taxas de vacinação com os traços de personalidade agregados.

4.3 Primeira Hipótese: Influência de fatores psicométricos na vacinação

Assume-se como hipótese principal que os traços de personalidade agregados por país correlacionam-se com taxas nacionais de vacinação. Para testá-la, uma regressão linear multivariável, OLS (*ordinary least squares*), foi realizada, tendo como parâmetros:

- Variáveis independentes: Fatores de personalidade agregados por país (O,C,E,A,N, de acordo com o Big-Five)
- Variáveis dependentes: Pessoas vacinadas a cada cem habitantes (PVPH)
- Variáveis de controle: IDH, renda e coeficiente de Gini.

O modelo utilizado para testar esta hipótese pode ser expressado como:

$$Y = \beta_0 + \beta_1 O + \beta_2 C + \beta_3 E + \beta_4 A + \beta_5 N + \beta_6 HDI + \beta_7 GDP + \beta_8 Gini \quad (17)$$

4.4 Segunda Hipótese: Efetividade de medidas de contenção em função de *clusters* de personalidade

A segunda hipótese consiste em identificar perfis de personalidade de países e determinar a sua relação com a efetividade das políticas de contenção. A hipótese foi testada por meio de técnicas de clusterização hierárquica aplicadas ao conjunto de dados de personalidade agregados por país. Modelos de regressões espaciais foram então implementados, fazendo-se uso de regimes espaciais definidos pelos diferentes *clusters* de personalidade obtidos (REY; ARRIBAS-BEL; WOLF, 2022). As variáveis dependentes deste modelo consistiram em na média dos dados de variação de mobilidade (adquiridos do *Google Mobility Reports* (GOOGLE, 2022a)) de todas as categorias de localidade, com exceção da residencial, e as variáveis independentes foram os índices de restrições de mobilidade obtidos do projeto OxCGRT (HALE *et al.*, 2021).

Expressa-se o modelo de regressão linear utilizado nos regimes espaciais da segunda hipótese da seguinte maneira:

$$Y_{ij} = \alpha_j + \sum_{k=1}^7 \beta_{k,j} C_{k,ij} \quad (18)$$

onde C_k ($k = 1, \dots, 7$), representam os indicadores da categoria “C”, $i = 1, \dots, n_j$ sendo n_j o total de observações no *cluster* j e $j = 1, 2$, os *clusters* de personalidade obtidos.

4.5 Terceira Hipótese: Interações entre traços de personalidade nacionais e a intensidade de políticas de contenção

A terceira hipótese também assume que traços de personalidade têm efeito sobre as medidas de contenção adotadas pelos países e, além de analisar os dados das séries temporais de variação de mobilidade e indicadores de políticas, inclui variáveis de interação entre cada traço e política analisados. A hipótese foi testada por meio de modelos de regressão em dados de painel, tendo como regressores: os indicadores da categoria “C” – utilizados também na segunda hipótese –, e o produto dos traços de personalidade agregados pelos mesmos indicadores. Assim como na segunda hipótese, o regressando consistiu na média dos dados de variação de mobilidade de todas as categorias de localidade, excetuando-se a residencial. Ao todo, cinco modelos foram implementados – um para cada traço –, e suas fórmulas podem ser descritas por:

$$Y_O = \beta_0 + \beta_O O + \sum_{i=1}^7 \beta_i C_i + \sum_{i=1}^7 \beta_{i+7} C_i \times O \quad (19)$$

$$Y_C = \beta_0 + \beta_C C + \sum_{i=1}^7 \beta_i C_i + \sum_{i=1}^7 \beta_{i+7} C_i \times C \quad (20)$$

$$Y_E = \beta_0 + \beta_E E + \sum_{i=1}^7 \beta_i C_i + \sum_{i=1}^7 \beta_{i+7} C_i \times E \quad (21)$$

$$Y_A = \beta_0 + \beta_A A + \sum_{i=1}^7 \beta_i C_i + \sum_{i=1}^7 \beta_{i+7} C_i \times A \quad (22)$$

$$Y_N = \beta_0 + \beta_N N + \sum_{i=1}^7 \beta_i C_i + \sum_{i=1}^7 \beta_{i+7} C_i \times N \quad (23)$$

As fórmulas acima (Fórmulas 19-23) variam também em relação aos períodos de tempo e entidades analisados, mas por brevidade tais detalhes foram omitidos.

5 RESULTADOS

5.1 Análise Exploratória

Inicialmente, conduziu-se uma análise exploratória sobre os conjuntos de dados para a construção e futura análise dos indicadores. Foram investigadas as associações entre os traços de personalidade agregados por países e seus respectivos índices de vacinação: pessoas vacinadas a cada cem habitantes (PVPH) e pessoas completamente vacinadas a cada cem habitantes (PFVPH).

Na Tabela 1 são apresentados dados de traços agregados de personalidade, PVPH e PFVPH dos países do conjunto analisado.

Tabela 1 – Países analisados e suas informações de personalidade e vacinação

País	O	C	E	N	A	n	PVPH	PFVPH	Data
Argentina	5.04	5.00	4.29	3.54	4.93	905.00	89.66	80.86	2022-03-25
Austrália	5.17	5.33	3.94	3.28	4.96	931.00	86.02	81.71	2022-03-24
Áustria	5.26	5.29	4.40	2.98	4.89	1088.00	75.33	72.86	2022-03-25
Bélgica	5.28	5.12	4.28	3.19	4.55	571.00	79.34	78.43	2022-03-23
Brasil	5.20	4.89	3.99	4.05	4.80	11568.00	84.45	74.62	2022-03-25
Bulgária	5.13	5.17	4.17	3.45	4.79	332.00	15.79	15.76	2022-04-03
Canada	5.24	5.45	3.95	3.26	5.05	2829.00	85.83	81.75	2022-03-24
Chile	5.10	5.40	4.45	3.37	5.20	545.00	92.96	90.29	2022-03-22
China	5.08	4.77	3.92	3.21	5.19	395.00	88.26	85.89	2022-03-21
Colômbia	5.15	5.26	4.40	3.24	5.31	1851.00	81.47	67.29	2022-03-25
Republica Tcheca	5.16	5.12	3.98	3.28	4.67	263.00	64.81	63.94	2022-03-25
Dinamarca	5.34	5.35	4.60	2.69	5.25	507.00	83.28	81.48	2022-03-24
Republica Dominicana	5.27	5.53	4.51	2.99	5.22	554.00	65.55	54.03	2022-03-25
Equador	5.12	5.27	4.37	3.22	5.05	303.00	82.91	76.57	2022-03-23
Finlândia	5.17	5.20	4.55	2.72	5.02	792.00	81.18	77.55	2022-03-23
França	5.18	5.09	4.17	3.48	4.55	2741.00	80.07	77.76	2022-03-24
Alemanha	5.19	5.27	4.26	3.01	4.98	10094.00	75.86	75.22	2022-03-25
Grécia	5.14	5.34	3.87	3.64	5.03	349.00	76.11	73.16	2022-03-25
Hungria	5.42	5.43	4.20	3.63	4.80	238.00	66.44	64.20	2022-03-24
Índia	5.16	5.22	4.00	3.22	5.19	988.00	70.36	59.19	2022-03-25
Indonésia	5.21	4.70	3.57	3.79	4.85	1607.00	70.69	56.57	2022-03-24
Irlanda	5.06	5.23	3.97	3.15	4.99	711.00	81.61	80.33	2022-03-24
Israel	5.11	5.50	3.71	3.16	4.78	407.00	72.14	65.95	2022-03-25
Itália	4.71	5.17	3.89	3.32	4.98	1856.00	84.01	79.17	2022-03-25
Japão	4.84	4.89	3.86	3.11	4.41	567.00	81.12	79.77	2022-03-24
Malásia	4.95	5.09	3.71	3.27	4.84	531.00	83.79	78.70	2022-03-25
México	5.21	5.18	4.54	3.27	5.12	3325.00	65.68	61.13	2022-03-25
Países Baixos	5.31	5.34	4.32	2.93	4.32	1418.00	78.35	72.01	2022-03-19
Nova Zelândia	5.14	5.35	4.05	3.15	5.02	359.00	83.56	78.72	2022-03-25
Noruega	5.18	5.28	4.40	3.09	4.95	303.00	79.25	73.67	2022-03-24
Peru	5.28	5.20	4.42	3.07	5.13	2036.00	84.81	75.90	2022-03-22
Filipinas	5.14	5.11	3.84	3.53	4.69	748.00	50.53	33.64	2022-03-17

(continua)

Tabela 1 – Países analisados e suas informações de personalidade e vacinação**(continuação)**

País	O	C	E	N	A	n	PVPH	PFVPH	Data
Polônia	4.60	5.23	4.09	3.24	4.84	482.00	59.74	59.09	2022-03-25
Portugal	5.27	5.30	4.22	3.29	5.28	556.00	95.04	92.60	2022-03-10
Catar	5.07	5.56	3.92	3.37	4.83	1273.00	88.40	88.40	2022-03-17
Romania	5.37	5.61	4.28	3.16	4.85	805.00	28.62	27.91	2022-04-03
Rússia	4.85	4.81	3.58	3.77	4.13	3404.00	54.57	49.83	2022-03-25
Singapura	4.78	5.26	3.76	3.29	4.93	417.00	91.74	90.92	2022-03-21
Eslováquia	5.19	5.35	4.27	3.59	4.50	610.00	51.74	50.69	2022-03-20
Coreia do Sul	4.84	5.01	3.62	3.28	4.61	292.00	87.60	86.68	2022-03-25
Espanha	5.25	5.19	4.46	3.07	5.26	2323.00	88.00	85.82	2022-03-23
Suécia	5.54	5.03	5.04	2.88	5.38	5861.00	76.94	74.94	2022-03-25
Suíça	5.15	5.48	4.52	3.22	4.74	4200.00	69.78	68.78	2022-03-24
Tailândia	5.26	5.14	3.75	3.39	4.57	309.00	78.92	71.64	2022-03-25
Turquia	5.37	5.14	4.54	3.58	3.89	2860.00	67.93	62.27	2022-03-25
Ucrânia	4.89	5.09	3.98	3.80	4.49	1454.00	36.19	35.02	2022-02-23
Reino Unido	5.06	5.26	3.95	3.25	4.92	11263.00	77.35	72.36	2022-03-24
Estados Unidos	5.30	5.36	3.90	3.26	5.07	11417.00	76.85	65.45	2022-03-24
Uruguai	4.93	5.29	4.22	3.10	5.02	242.00	85.26	81.51	2022-03-25
Venezuela	5.29	5.50	4.53	2.76	5.53	661.00	77.19	49.77	2022-03-11
Vietnã	4.83	5.30	3.49	3.72	5.63	848.00	81.44	79.20	2022-03-22

Fonte: Autoria própria (2022).

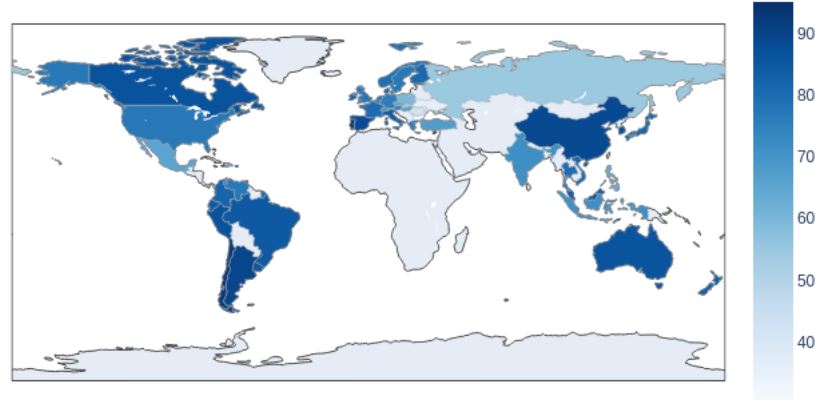
5.1.1 Visualização dos traços de personalidade nacionais

A partir do *dataset* foi possível observar a distribuição dos traços de personalidade através do globo. No apêndice A, os traços nacionais são dispostos em mapas de calor. É possível perceber que as escalas de cada traço variam e dependendo da figura observada, as discrepâncias entre eles ficam maiores, como é o caso da Figura 10, onde o Brasil é o número um em Neuroticismo no mundo com uma média de 4,05 e o segundo lugar é a Ucrânia com 3,8.

5.1.2 Visualização dos dados de vacinação por país

A Figura 1 demonstra como as taxas de vacinação distribuem-se pelo globo, isto é, pessoas que receberam pelo menos uma dose da vacina. Tais taxas variam desde 51.74%, na Eslováquia, até 95.04% em Portugal.

Figura 1 – Mapa de calor para vacinação

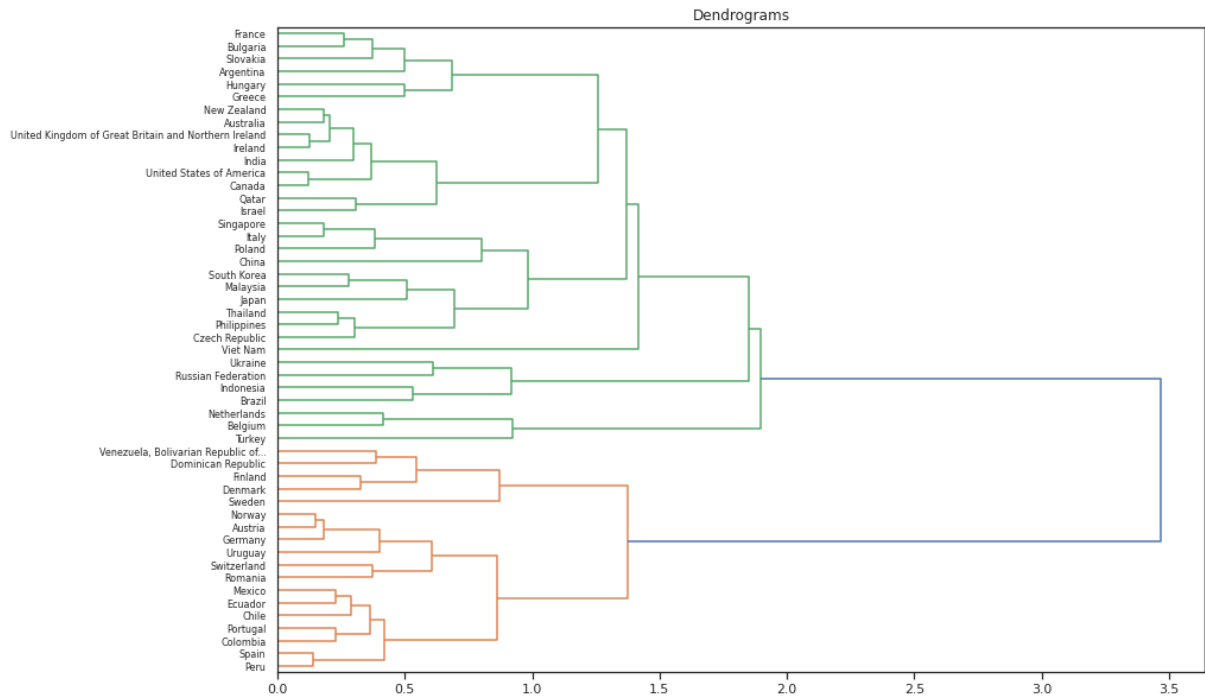


Fonte: Adaptado de (MATHIEU E., 2021).

5.1.3 Clusterização hierárquica dos traços de personalidade agregados

O dendrograma da Figura 2 apresenta os resultados da *clusterização hierárquica* realizada sobre os traços de personalidade agregados. A clusterização foi implementada por meio do módulo de clusterização hierárquica da biblioteca de Python *scipy* (VIRTANEN *et al.*, 2020), utilizando o método de Ward como de cálculo de distâncias entre *clusters*. Nota-se que países étnica e historicamente similares são agrupados em *clusters* de maior proximidade; vide Nova Zelândia, Austrália, Reino Unido, Irlanda, Índia, Estados Unidos da América e Canadá - todos marcados por uma mesma herança britânica. Similarmente, pode-se observar que países escandinavos, hispânicos, germânicos e asiáticos aglomeram-se em grupos próximos uns dos outros.

Figura 2 – Dendrograma dos traços de personalidade por país



Fonte: Autoria própria (2022).

5.1.4 Correlação entre vacinação e traços de personalidade por país

No Quadro 1, estão dispostas as correlações entre os traços de personalidade e os índices de vacinação. Observa-se pela última coluna os valores pertinentes ao estudo: dentre os traços, Agradabilidade (A) apresenta a maior correlação, de 0,38, com os indicadores de vacinação. O traço Neuroticismo (N) também aparentou uma correlação negativa com os índices de vacinação, porém com uma intensidade menor de -0,267.

Quadro 1 – Correlações entre variáveis

	O	C	E	A	N	PVPH
O	1	0,257	0,586	0,116	-0,256	-0,029
C	0,257	1	0,296	0,299	-0,376	-0,027
E	0,586	0,296	1	0,235	-0,539	-0,013
A	0,116	0,299	0,235	1	-0,372	0,38
N	-0,256	-0,376	-0,539	-0,372	1	-0,267
PVPH	-0,029	-0,027	-0,013	0,38	-0,267	1

Fonte: Autoria própria (2022).

5.2 Modelagem e implementação

5.2.1 Primeira Hipótese: Influência de fatores psicométricos na vacinação

A regressão OLS descrita na seção de Metodologia foi implementada por meio da biblioteca de Python *Statsmodels* (SEABOLD; PERKTOLD, 2010). Quatro modelos foram implementados visando analisar a contribuição individual de cada variável de controle no poder explicativo da regressão. Os valores de índice de Gini, obtidos do Banco Mundial (BANK, 2020), foram selecionados de modo que o ano mais recente de sua produção fosse posterior ao ano 2000. Três países (Cingapura, Qatar e Nova Zelândia) tiveram de ser excluídos dos modelos 3 e 4 por não possuírem dados do índice de Gini disponíveis na fonte em que os valores dos outros países foram coletados. Na Tabela 2, estão dispostos os resultados dos modelos:

Tabela 2 – Modelos de regressão linear

	(1)	(2)	(3)	(4)
A	21.032*** (6.231)	18.732*** (5.960)	13.352** (6.294)	18.758*** (5.986)
C	-17.714* (9.869)	-18.255* (10.002)	-19.747* (10.594)	-24.071** (9.678)
E	-9.634 (7.531)	-7.416 (7.664)	-8.953 (7.867)	-10.764 (7.308)
N	-7.487 (8.662)	-6.602 (9.015)	-18.492** (8.506)	-10.222 (8.426)
O	4.514 (11.586)	1.743 (11.633)	6.488 (12.176)	8.823 (11.295)
const	58.803 (90.552)	117.312 (82.603)	164.614* (84.392)	34.554 (89.808)
Gini			0.349 (0.296)	0.841*** (0.306)
GDP		0.000** (0.000)		-0.000 (0.000)
HDI	55.771** (24.907)			90.339** (42.309)
Observations	51	51	48	48
R²	0.312	0.303	0.286	0.444
Adjusted R²	0.218	0.208	0.182	0.330
Residual Std. Error	12.219 (df=44)	12.301 (df=44)	12.546 (df=41)	11.355 (df=39)
F Statistic	3.325*** (df=6; 44)	3.183** (df=6; 44)	2.741* (df=6; 41)	3.892*** (df=8; 39)
Note:	*p<0.1; **p<0.05; ***p<0.01			

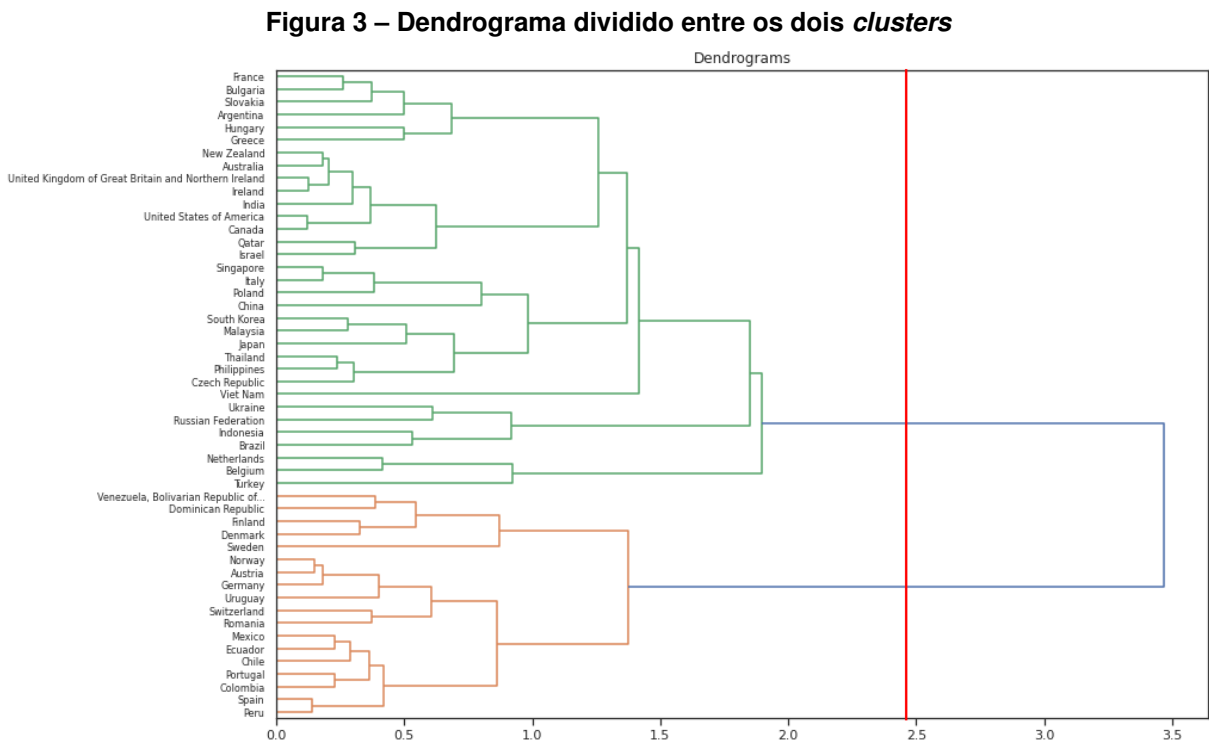
Fonte: Autoria própria (2022).

Observa-se que em todos os modelos o fator Agradabilidade apresentou, dentre os traços, a relação mais significativa com a taxa de vacinação por cem mil habitantes. A adição de todas as variáveis de controle (Modelo 4) resultou em um modelo com valor de R^2 ajustado de 33.0%; sendo que Agradabilidade ($\beta_A = 18.758, SE = 5.986, p < 0.01$), co-

eficiente de Gini ($\beta_{Gini} = 0.841, SE = 0.306, p < 0.01$) e IDH ($\beta_{IDH} = 90.339, SE = 42.309, p < 0.05$) apresentaram relação positiva com a taxa de vacinação, e Conscienciosidade ($\beta_C = -24.071, SE = 9.678, p < 0.05$), negativa. A nível de significância de $p < 0.05$ o traço Neuroticismo apresentou relação negativa com a variável-alvo (Modelo 3) ao considerar como variável de controle apenas o coeficiente de Gini. Outra interação significativa e negativa ($p < 0.05$) foi encontrada para o traço Conscienciosidade (Modelo 4), após serem introduzidas todas as variáveis de controle. De maneira geral, os resultados apontam que, independentemente das variáveis socio-econômicas de um país, o traço Agradabilidade agregado está associado com sua taxa de vacinação.

5.2.2 Segunda Hipótese: Efetividade de medidas de contenção em função de *clusters* de personalidade

Dois *clusters* resultaram da divisão – representada na Figura 3 – estabelecida com base nos *subclusters* de maior distância entre si. Observa-se que o *cluster* com menor número de elementos é composto principalmente por países de origens norte-européia e hispânica.



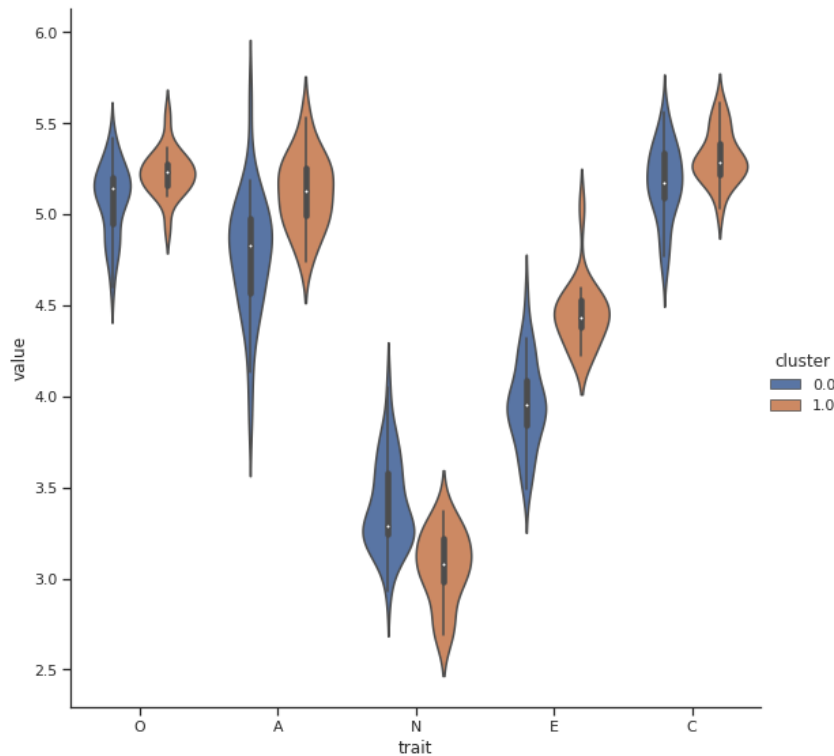
Fonte: Autoria própria (2022).

Para identificar sob quais traços essa divisão se baseou, foram utilizados *violin plots*, que caracterizam as medianas e densidades de probabilidade dos traços em cada *cluster*. Observa-se pela Figura 4 que o traço *Openness* (Abertura à experiência) apresenta medianas similares entre os dois *clusters* - 5.08 e 5.22 respectivamente no *cluster* 0 e 1 -, o traço *Conscientious-*

ness (Conscienciosidade) também possui medianas parecidas, sendo elas 5.17 e 5.31 respectivamente. Já os outros traços apresentam diferenças entre os *clusters* de no mínimo 0.33.

Nesse sentido, pode-se dizer que em relação ao *cluster* 0 os países do *cluster* 1 são caracterizados por maiores Agradabilidade e Extroversão, e menor Neuroticismo; e os países dos dois grupos são similares em relação aos traços Abertura à experiência e Conscienciosidade.

Figura 4 – Violin plot dos traços de cada *cluster*



Fonte: Autoria própria (2022).

Fazendo uso da biblioteca *pysal* foi implementado – por meio de *sprege.OLS_Regimes* – o modelo de regressão espacial com regimes espaciais, utilizando a média das mobilidades como variável dependente, os índices de restrições de mobilidade impostas pelos países como variáveis independentes e o número do *cluster* como associação de regime. Pela tabela 3 é possível perceber que os indicadores de maior relevância estatística para a variação de mobilidade são C1, C2, C3 e C6 e C7. Nota-se que, para o *cluster* 0, todos os indicadores apresentaram relação negativa com a variação de mobilidade e, no *cluster* 1, C4 e C5 apresentaram relação positiva.

Tabela 3 – Resultado da regressão espacial com regimes espaciais

	cluster_0			cluster_1		
	Coefficient	Std. Error	P-value	Coefficient	Std. Error	P-value
CONSTANT	16.568516	0.584205	9.963002e-167	17.690763	0.732723	2.630566e-123
C1M_School closing	-4.248997	0.380675	1.156346e-28	-6.779189	0.536368	3.487418e-36
C2M_Workplace closing	-3.993131	0.484151	1.944330e-16	-2.483120	0.720750	5.743459e-04
C3M_Cancel public events	-3.084542	0.639454	1.441892e-06	-4.784742	0.856095	2.378022e-08
C4M_Restrictions on gatherings	-0.183820	0.251251	4.644287e-01	1.751544	0.374949	3.052682e-06
C5M_Close public transport	-0.046221	0.474121	9.223415e-01	3.005024	0.771382	9.894560e-05
C6M_Stay at home requirements	-4.400504	0.455662	6.461948e-22	-6.013377	0.636070	4.482925e-21
C7M_Restrictions on internal movement	-1.055232	0.441453	1.686072e-02	-4.380181	0.587052	9.709576e-14

Fonte: Autoria própria (2022).

A fim de determinar se existem diferenças significativas entre os regimes (*clusters*) da regressão, Testes de *Chow* foram conduzidos no nível global, e entre os coeficientes de cada regime. A Tabela 4 representa a diferença global entre os regimes, demonstrando – à nível de significância $p = 3.606 \cdot 10^{-16}$ – que os dois modelos são estatisticamente diferentes entre si. A Tabela 5 apresenta as estatísticas de *Chow* para cada coeficiente da regressão, revelando que apenas os indicadores C2, C3 e C4 diferem de forma significativa ($p < 0.0055$) entre os modelos.

Tabela 4 – Resultado do Teste de Chow apresentado a diferença global entre os regimes

Statistic	P-value
90.543	3.606e-16

Fonte: Autoria própria (2022).

As análises revelaram que C2 e C3 foram os únicos indicadores que ao mesmo tempo apresentaram diferenças significativas entre regimes e relações estatisticamente relevantes na variação de mobilidade, de modo que o efeito de C3 foi mais intenso no *cluster* 1 e o de C2, no *cluster* 0.

Tabela 5 – Resultado dos Testes de Chow para os coeficientes da regressão

	Statistic	P-value
CONSTANT	7.891929	0.004966
C1M_School closing	1.592139	0.207021
C2M_Workplace closing	13.462710	0.000243
C3M_Cancel public events	7.709230	0.005494
C4M_Restrictions on gatherings	7.965285	0.004768
C5M_Close public transport	0.029764	0.863026
C6M_Stay at home requirements	3.056279	0.080426
C7M_Restrictions on internal movement	0.000004	0.998496

Fonte: Autoria própria (2022).

5.2.3 Terceira Hipótese: Interações entre traços de personalidade nacionais e a intensidade de políticas de contenção

Apenas com os resultados da segunda hipótese, não foi possível discriminar quais traços diminuem efetivamente a mobilidade e determinar quais políticas governamentais são adequadas para cada perfil de país.

Para determinar o impacto de cada traço na mobilidade, foi realizada uma regressão em painel com a biblioteca *linearmodels* utilizando o modelo *PanelOLS*, utilizando variáveis de interação entre cada política adotada e traços de personalidade agregados dos países, de forma a ressaltar o impacto de cada traço no resultado final. O período temporal foi dividido em semanas, resultando em 130 intervalos de tempo e as entidades são representadas por cada país. As entidades são representadas por cada país, e como variável dependente foi utilizada a variação de mobilidade. A regressão foi realizada de modo a controlar efeitos constantes entre entidades e variáveis ao longo do tempo.

A Tabela 6 apresenta os resultados dos modelos de regressão por painel com efeitos de tempo. O Modelo 0 é constituído apenas pelos indicadores C1-C7, e os Modelos 1-5 introduzem – além dos indicadores – os traços de personalidade e variáveis de interação. No Modelo 1, por exemplo, foram introduzidas as interações com o traço Abertura à Experiência (O), de modo que interação do indicador C1 com o este traço é representada por C1_O. É possível perceber pelo Modelo 0 que, em contraste com resultados da segunda hipótese, apenas C1, C2 e C6 apresentaram relevância estatística. De modo geral, a significância dos indicadores varia substancialmente de acordo com os modelos.

Os modelos que apresentaram interações significativas com os traços de personalidade foram aqueles influenciados por Abertura à Experiência (O), Extroversão (E) e Neuroticismo (N). No Modelo 1, a interação mais significativa de todas, C1_O, é relevante a nível de $p = 0.01$. À nível de significância $p = 0.05$, três variáveis de interação mostraram-se relevantes: C7_E, C2_N e C7_N – Modelos 3 e 4, respectivamente.

Para interpretar o relacionamento da variáveis de interação com a variação de mobilidade, foram traçados gráficos de interação (*interaction plots*)¹. A Figura 5 revela a interação do indicador C1 com o traço Abertura à Experiência (O). Cada ponto corresponde a um país, de modo que a coordenada x representa a média nacional do traço “O” e a coordenada y , a média de variação de mobilidade do país em função da intensidade da política C1. É possível perceber pelos gráficos que há grande variação na forma como as políticas de contenção afetaram a variação de mobilidade de diferentes países. Portanto, a natureza das interações aparenta ter elevada complexidade.

¹ Os gráficos de interação para os coeficiente C2_N e C7_N encontram-se no Apêndice B

Tabela 6 – Regressão em dados de painel com efeitos de tempo

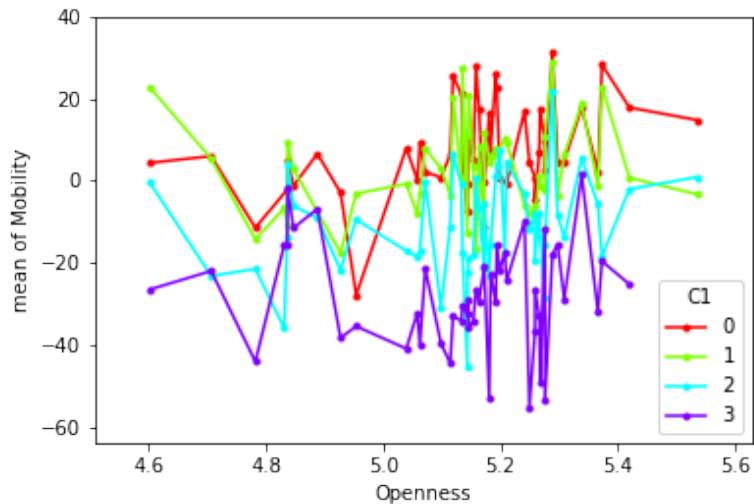
variable	0	1	2	3	4	5
const	10.841***	-18.951	41.912	-27.886	-22.491	-4.415
C1	-2.046**	44.101***	17.639	7.901	-3.173	17.302
C2	-3.558***	-19.333	-38.781*	-12.248	30.159**	-1.787
C3	-1.058	-74.393**	1.843	-1.129	-10.330	-15.124
C4	-0.262	13.084	5.682	-2.788	8.664	3.269
C5	-1.603	-45.713	-6.065	-21.769	3.039	-2.354
C6	-4.404***	-1.917	-26.730	7.272	-13.900	-22.595*
C7	-1.188	27.122	3.307	29.207**	-34.110***	12.999
O		5.810				
C1_O		-9.026***				
C2_O		3.143				
C3_O		14.267*				
C4_O		-2.629				
C5_O		8.577				
C6_O		-0.487				
C7_O		-5.472				
C			-5.852			
C1_C			-3.797			
C2_C			6.745			
C3_C			-0.610			
C4_C			-1.129			
C5_C			0.858			
C6_C			4.293			
C7_C			-0.873			
E				9.100		
C1_E				-2.413		
C2_E				2.241		
C3_E				-0.037		
C4_E				0.618		
C5_E				4.977		
C6_E				-2.876		
C7_E				-7.335**		
N					10.414	
C1_N					0.190	
C2_N					-10.238**	
C3_N					2.723	
C4_N					-2.742	
C5_N					-1.466	
C6_N					2.993	
C7_N					9.956**	
A						3.072
C1_A						-3.992
C2_A						-0.282
C3_A						2.888
C4_A						-0.737
C5_A						0.124
C6_A						3.743
C7_A						-2.901
R ²	0.217448	0.237404	0.22157	0.242562	0.243743	0.224725

*p<0.1; **p<0.05; ***p<0.01;

Fonte: Autoria própria (2022).

De maneira geral, os gráficos de interação (Figuras 5, 6, 12 e 13) revelam que quanto mais intensa foi a política de contenção mais negativa foi a variação de mobilidade, o que indica que as medidas atingiram o objetivo desejado. Há indícios de linearidade no relacionamento da variação de mobilidade com Abertura à Experiência para C1 em nível 3, especialmente no intervalo de 5.0 a 5.4.

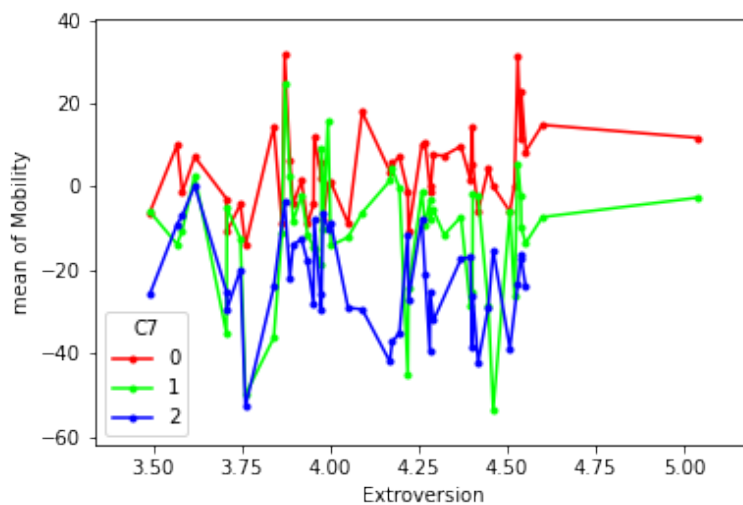
Figura 5 – Gráfico da Interação entre Abertura à Experiência e C1



Fonte: Autoria própria (2022).

A Figura 6 aponta também para um possível relacionamento linear negativo entre Extroversão e variação de mobilidade para C7 em nível 2, e relacionamento linear positivo para C7 em nível 0.

Figura 6 – Gráfico da Interação entre Extroversão e C7



Fonte: Autoria própria (2022).

6 CONCLUSÃO

6.1 Resultados obtidos

Os resultados dos modelos de regressão linear da primeira hipótese demonstram que certos traços de personalidade agregados por países apresentam, de fato, relações estatisticamente significativas com as taxas nacionais de vacinação. O traço Agradabilidade apresentou a relação mais significativa de todas; mais especificamente, relacionou-se de forma significativa e positiva independentemente da introdução de variáveis de controle. Este resultado concorda com o que a literatura de psicologia tradicionalmente relata sobre o traço Agradabilidade: pessoas que apresentam níveis mais altos deste traço são caracterizadas por comportamentos pró-sociais e maior obediência; o que potencialmente explicaria uma adesão maior aos protocolos de vacinação por parte destes indivíduos.

Ainda, nos resultados da primeira hipótese, outros relacionamentos menos significativos foram revelados para o traço Conscienciosidade – após a adição de todas as variáveis de controle – e para Neuroticismo no Modelo 3; ambos apresentaram coeficientes negativos. No entanto, os mecanismos que explicam estas relações não parecem claros. Esperaria-se no caso de Conscienciosidade, justamente por ser um traço associado à adesão a regras e padrões, que este apresentasse relação positiva com as taxas de vacinação. Similarmente, no caso de Neuroticismo, era esperado que este traço também apresentasse relação positiva com os índices de vacinação, por caracterizar indivíduos com maior aversão a riscos e propensão a experimentar medo e ansiedade – características de natureza protetiva em contextos de prevenção de transmissão patógenos. Futuras análises deverão, portanto, investigar os sub-aspectos destes traços por meio da aplicação de testes de personalidade mais extensos.

Em relação à segunda hipótese, a regressão espacial revelou que parte substancial dos indicadores de contenção e fechamento relaciona-se de forma significativa com a variação de mobilidade. No entanto, não foi possível determinar claramente como os *clusters* de personalidade dos países influenciam na efetividade das medidas tomadas. É possível ver, porém, que existem, entre os dois grupos, diferenças estatisticamente significativas nas respostas a certas medidas – notadamente, fechamento de locais de trabalho (C2), cancelamento de eventos públicos (C3) e restrições a aglomerações (C4).

Por fim, a terceira hipótese teve como objetivo investigar a interação entre os traços de personalidade nacionais e as medidas de contenção, e sua influência na variação de mobilidade. Os resultados das regressões em dados de painel revelaram coeficientes de interação significativos nos modelos influenciados pelos traços: Abertura à Experiência, Extroversão e Neuroticismo. A interpretação dos efeitos de interação, no entanto, se mostrou bastante complexa – especialmente em relação aos coeficientes C2_N e C7_N (Figuras 12 e 13). Os gráficos de interação dos coeficientes C1_O e C7_E (Figuras 5 e 6) apresentam indícios de relacionamentos

lineares – entre os traços e a variação de mobilidade – que variam conforme a intensidade das políticas de contenção.

De maneira geral, apesar dos resultados dos testes das hipóteses segunda e terceira não diferenciarem claramente quais medidas políticas são mais adequadas aos países com base em seus perfis de personalidade, eles apontam para efeitos estatisticamente relevantes dos traços de personalidade nacionais na efetividade das medidas de contenção tomadas pelos países. Além disso, as evidências apontaram que os perfis de personalidade encontrados são substancialmente diferentes a ponto de produzirem efeitos significativos na efetividade de certas medidas de contenção.

6.2 Limitações do presente estudo

O principal fator limitante do estudo foi a qualidade e disponibilidade dos dados utilizados. Nesse sentido, destacam-se:

- Países com índices de vacinação destoantes em relação aos demais – como no caso de países africanos.
- A documentação dos dados de mobilidade do *Google* carece de explicações sobre quais tipos de estabelecimentos se enquadram em cada categoria de mobilidade;
- A mobilidade referencial das 5 semanas, entre 3 de Janeiro a 6 de Fevereiro de 2020, pode não se enquadrar como um patamar ideal, pois não leva em conta diversos fatores como a sazonalidade de países, que é o caso do Brasil, tendo uma mobilidade altíssima nesse período principalmente por conta do turismo. Além disso alguns governos, como o caso da China, já estavam começando a aplicar medidas de contenção, ocasionando num patamar basal adulterado.
- Ausência de dados de mobilidade para países como a China.
- O *dataset* OxCGRT apresenta diversas limitações metodológicas, principalmente em relação a forma e a frequência com as quais os dados dos indicadores são atualizados.
- A presença de efeitos de colinearidade nos indicadores de contenção dificulta a interpretação dos resultados dos modelos.
- Os dados de personalidade foram mensurados a partir de um inventário reduzido com 10 itens, o que resulta em medidas de menor definição - em relação à inventários mais extensos -, e impede que traços sejam decompostos a nível de suas sub-facetas do *Big-Five*.
- Apenas 55 países apresentaram mais de 200 respondentes ao questionário de personalidade.

A somatória da falta de dados acumula um fator de erro significativo no estudo, impactando a fidelidade de seu resultado.

6.3 Trabalhos futuros

Almejando resultados de maior precisão, futuras análises deverão principalmente preocupar-se em coletar os dados de personalidade por meio de inventários mais extensos. Idealmente, os testes deveriam ser aplicados presencialmente, a fim de evitar vieses resultantes de auto-aplicação. Além disso, seria de interesse coletar informações a respeito das cidades e regiões de origem dos respondentes, permitindo que os traços médios sejam analisados com relação às subdivisões dos países. Isso possibilitaria que as análises de clusterização obtivessem maior granularidade, possivelmente resultando em grupos geográficos de maior significado. No caso do *dataset* de dados de mobilidade – que apresenta para alguns países informações sobre variações regionais de mobilidade – a maior granularidade dos dados permitiria que heterogeneidades subnacionais dos países fossem capturadas de forma apropriada.

Em relação a segunda e terceira hipóteses, técnicas de análises de componentes principais poderiam ser aplicadas para controlar os efeitos de colinearidade dos indicadores de contenção. Dessa forma, as interações entre as medidas de contenção e os perfis de personalidade dos países possivelmente se tornariam mais interpretáveis. A adição de variáveis de controle aos modelos de regressão espacial e em dados de painel também poderia ter sido realizada para que efeitos socioeconômicos fossem contabilizados.

Enfim, de posse dos resultados apresentados no presente estudo, pode-se concluir que, de fato, os traços de personalidade agregados dos países relacionam-se de forma significativa com as respostas às medidas de contenção da pandemia. A clara identificação dos fatores responsáveis por estas interações é portanto importante para que recomendações sobre políticas de contenção da pandemia possam ser elaboradas de acordo com ambiente psicométrico particular de cada país. Os governos poderiam então fazer uso de tais informações para orientar a aplicação de medidas de controle em eventuais contextos de crises sanitárias, potencialmente reduzindo seus impactos de forma mais eficiente.

REFERÊNCIAS

- ANSELIN, L.; REY, S. J. **Modern spatial econometrics in practice: A guide to geoda, Geodaspace and Pysal**. [S.l.]: Geoda Press LLC, 2014.
- BANK, W. **Gini index**. 2020. Disponível em: <https://data.worldbank.org/indicator/SI.POV.GINI>.
- BANK, W. **GDP per capita**. 2021. Disponível em: <https://data.worldbank.org/indicator/NY.GDP.PCAP.CD>.
- BARCELO, J. National personality traits and regime type: A cross-national study of 47 countries. **Journal of Cross-Cultural Psychology**, Sage Publications Sage CA: Los Angeles, CA, v. 48, n. 2, p. 195–216, 2017.
- DABLA-NORRIS, E.; KHAN, H.; LIMA, F. What determines social distancing? evidence from advanced and emerging market economies. **IMF Working Paper, forthcoming**, 2021.
- FETZER, T. *et al.* **Global behaviors and perceptions in the COVID-19 pandemic**. OSF, 2020. Disponível em: <https://osf.io/3sn2k/>.
- GOOGLE. **COVID-19 Community Mobility Report**. Google, 2022. Disponível em: <https://www.google.com/covid19/mobility/>.
- GOOGLE. **Mobility Report CSV Documentation**. Google, 2022. Disponível em: https://www.google.com/covid19/mobility/data_documentation.html?hl=en.
- GOSLING, S. D.; RENTFROW, P. J.; JR, W. B. S. A very brief measure of the big-five personality domains. **Journal of Research in personality**, Elsevier, v. 37, n. 6, p. 504–528, 2003.
- GÖTZ, F. M. *et al.* How personality and policy predict pandemic behavior: Understanding sheltering-in-place in 55 countries at the onset of covid-19. **American Psychologist**, American Psychological Association, v. 76, n. 1, p. 39, 2021.
- HALE, T. *et al.* A global panel database of pandemic policies (oxford covid-19 government response tracker). **Nature Human Behaviour**, v. 5, n. 4, p. 529–538, 2021.
- HANCK, C. *et al.* **Introduction to econometrics with R**. 2021. Disponível em: <https://www.econometrics-with-r.org/>.
- HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. **The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition**. Springer New York, 2009. (Springer Series in Statistics). ISBN 9780387848587. Disponível em: <https://books.google.com.br/books?id=tVljmNS3Ob8C>.
- JOHN, O. P.; SRIVASTAVA, S. *et al.* The big five trait taxonomy: History, measurement, and theoretical perspectives. **Handbook of personality: Theory and research**, v. 2, n. 1999, p. 102–138, 1999.
- MATHIEU E., R. H. O.-O. E. e. a. A global database of covid-19 vaccinations. **Nat Hum Behav**, Nature, v. 5, p. 947–953, 2021.
- MCCRAE, R. R.; TERRACCIANO, A. Personality profiles of cultures: aggregate personality traits. **Journal of personality and social psychology**, American Psychological Association, v. 89, n. 3, p. 407, 2005.

MOLNAR, C. **Interpretable Machine Learning**. Leanpub, 2020. ISBN 9780244768522.
Disponível em: <https://books.google.com.br/books?id=jBm3DwAAQBAJ>.

MONTGOMERY, D. C.; RUNGER, G. C. **Applied Statistics and probability for Engineers**.
7th. ed. [S.l.]: Wiley, 2018.

REY, S. J.; ARRIBAS-BEL, D.; WOLF, L. J. **Geographic Data Science with python**. 2022.
Disponível em: <https://geographicdata.science/book/intro.html#>.

SEABOLD, S.; PERKTOLD, J. statsmodels: Econometric and statistical modeling with python.
In: 9th Python in Science Conference. [S.l.: s.n.], 2010.

TAN, P.; STEINBACH, M.; KUMAR, V. **Introduction to Data Mining: Pearson New
International Edition PDF eBook**. Pearson Education, 2013. ISBN 9781292038551.
Disponível em: <https://books.google.com.br/books?id=jC6pBwAAQBAJ>.

UNDP. **Human Development Index (HDI)**. 2019. Disponível em: <https://hdr.undp.org/en/indicators/137506>.

VEDEL, A.; POROPAT, A. E. Personality and academic performance. **Encyclopedia of
personality and individual differences**, Springer Cham, Switzerland, p. 1–9, 2017.

VIRTANEN, P. *et al.* SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python.
Nature Methods, v. 17, p. 261–272, 2020.

**APÊNDICE A – Mapas de calor mundial de todos os traços de
personalidade**

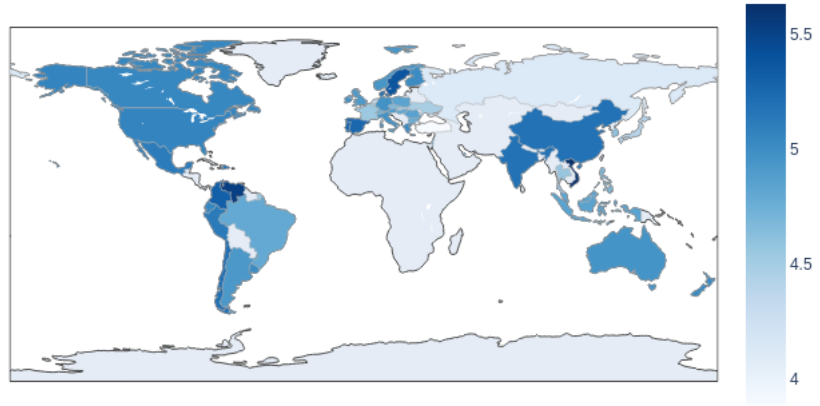


Figura 7 – Mapa de calor para Agradabilidade

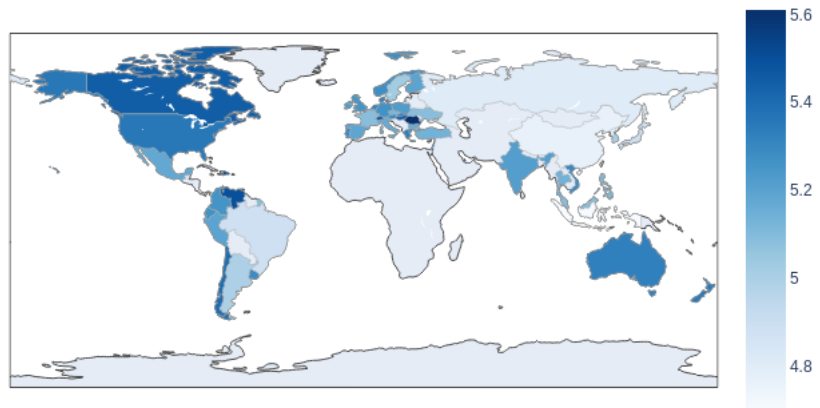


Figura 8 – Mapa de calor para Conscienciosidade

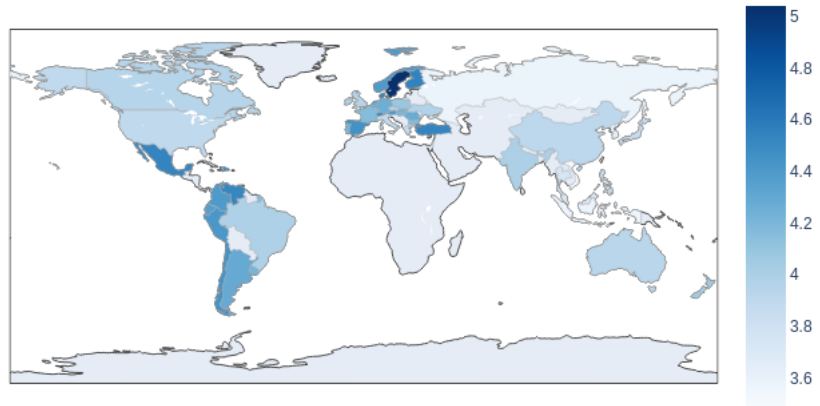


Figura 9 – Mapa de calor para Extroversão

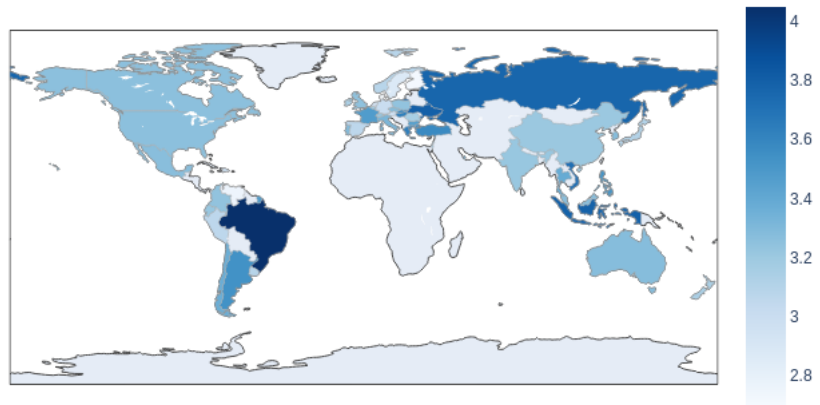


Figura 10 – Mapa de calor para Neuroticismo

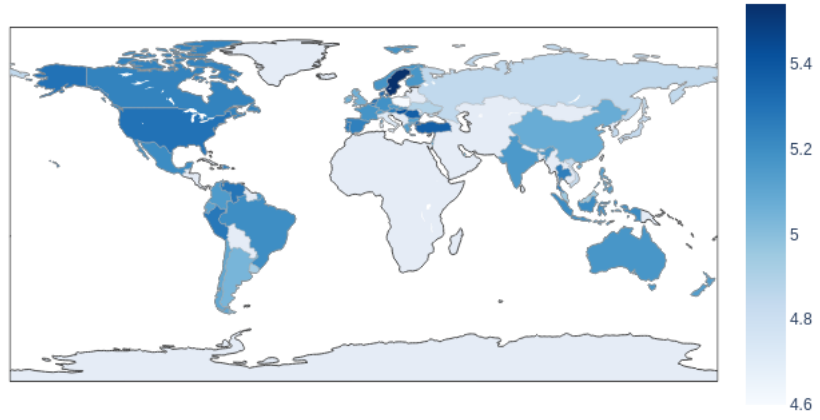
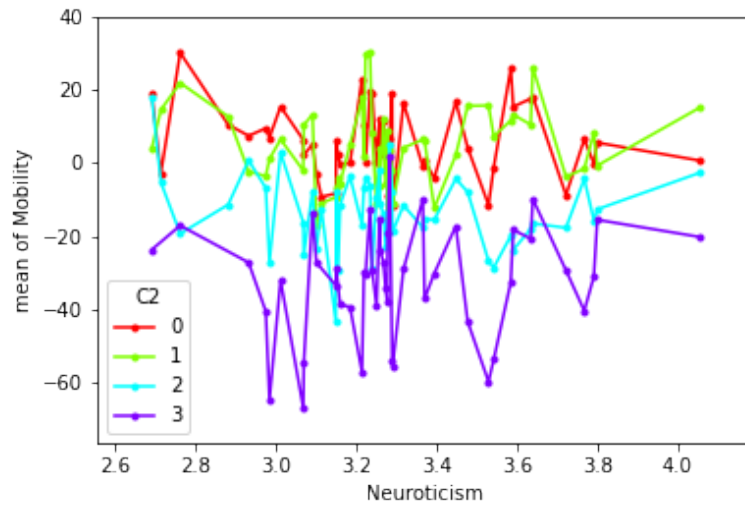


Figura 11 – Mapa de calor para Abertura à Experiência

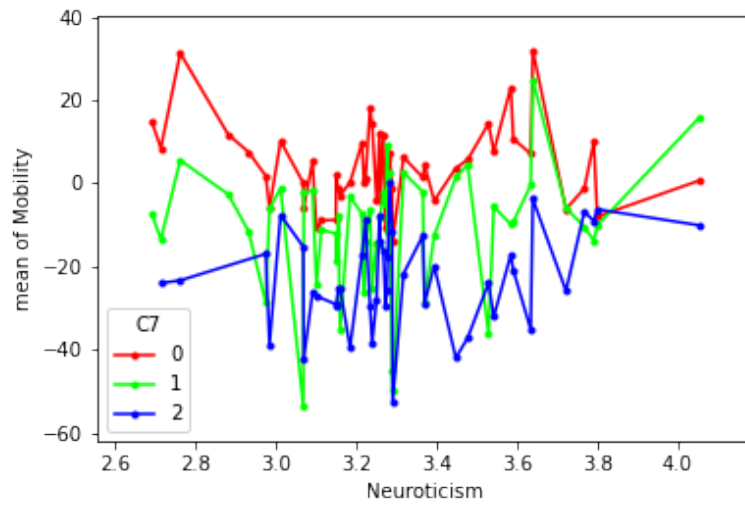
**APÊNDICE B – Gráficos de interação entre indicadores de políticas de
contenção e traços de personalidade**

Figura 12 – Gráfico da Interação entre Neuroticismo e C2



Fonte: Autoria própria (2022).

Figura 13 – Gráfico da Interação entre Neuroticismo e C7



Fonte: Autoria própria (2022).