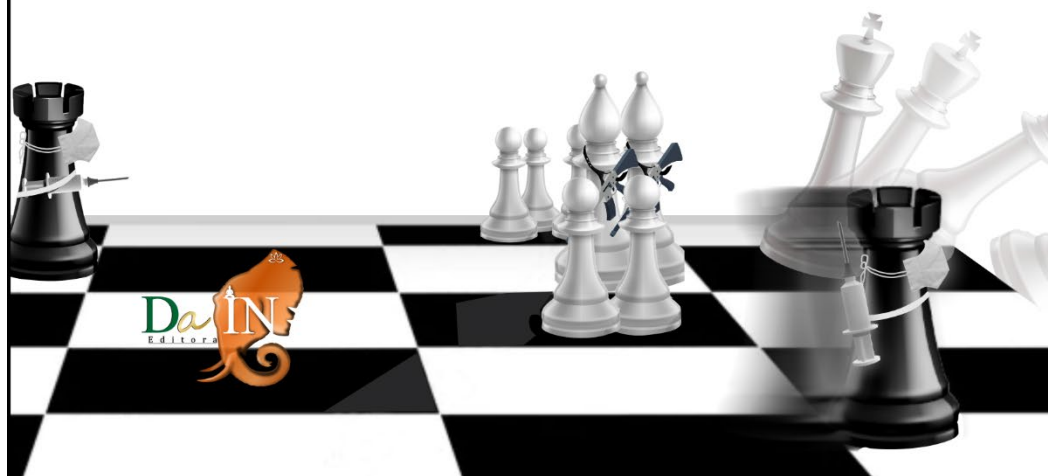
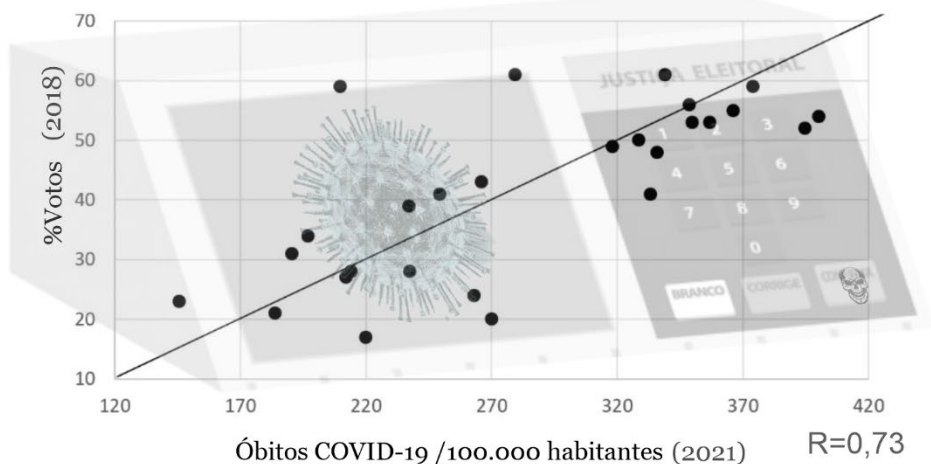


Óbitos & Votos:

Um estudo didático com os dados da COVID-19 no Brasil.

Paulo Wenderson Teixeira Moraes

Prof. Dr. Titular UNEB



Copyright © 2022 by Paulo Wenderson Teixeira Moraes

Nenhuma parte deste livro pode ser reproduzida, sejam qu
forem os meios empregados, a não ser com a permissão escrita do autor
e da editora, conforme a Lei nº 9610, de 19 de fevereiro de 1998.

Capa: Yago dos Santos Cruz

FICHA CATALOGRÁFICA

Biblioteca Professor Edivaldo Machado Boaventura – UNEB – Campus I

M823p Moraes, Paulo Wenderson Teixeira, 1978 –

Óbitos e votos: um estudo didático com os dados da COVID-19
no Brasil / Paulo Wenderson Teixeira: DaIN. – Salvador, 2022.
151 p. : il.

Inclui referências e anexos.
ISBN: 978-65-00-38334-8.

1. Estatística. 2. Covid-19 – Estatísticas - Brasil. 3. Doenças
transmissíveis – Estatística. 4. Saúde pública – Estatísticas –
Brasil. 5. Saúde pública – Aspectos sociais - Brasil I. Título.

CDD: 362.10981

Célia Maria da Costa – Bibliotecária – CRB5/918



<https://www.dain.art.br/obitosevotos>

Para Abílio, Jorge, Simon, Eduardos, Flávios
e Carlos, Gisela, Isabela...

Além da providência divina, este livro não seria possível sem o auxílio de uma série de pessoas: sou grato a todos que de alguma forma colaboraram para a conclusão dessa obra. Aos amigos que suportaram tanta discussão sobre dados estatísticos, à família que suportou o escritor em processo de criação, aos colegas de trabalho que compartilharam o tempo para avaliar os argumentos e análises. A Henrique Morais que gentilmente me presenteou com o prefácio. Ao professor Virgílio Bastos que me possibilitou a iniciação no mundo da ciência e me ensinou as primeiras análises em um banco de dados. A todos os meus professores, especialmente, a minha professora Olga Claro, exemplo de criatividade e experimentação no ensino da matemática. Finalmente, grato a minha mãe Adenilde, professora de Matemática que me introduziu na arte de brincar de resolver problemas!

When I Heard the Learn'd Astronomer

By Walt Whitman, 1892

When I heard the learn'd astronomer,
When the proofs, the figures, were ranged in columns
before me,
When I was shown the charts and diagrams, to add, divide,
and measure them,
When I sitting heard the astronomer where he lectured
with much applause in the lecture-room,
How soon unaccountable I became tired and sick,
Till rising and gliding out I wander'd off by myself,
In the mystical moist night-air, and from time to time,
Look'd up in perfect silence at the stars.

(WHITMAN, 1993)

Quando Eu Ouvi o Astrônomo Entendido

Por Walt Whitman

Quando ouvi o astrônomo entendido,
Quando as evidências, os algoritmos, foram ordenados em
tabelas diante de mim,
Quando me foram apresentados os mapas e diagramas a somar,
dividir e mensurar,
Quando, sentado, ouvia o astrônomo muito aclamado, no
auditório da conferência,
Senti-me logo inexplicavelmente cansado e nauseado,
Até que me levantei e saí, vagando sem rumo
No ar úmido e místico da noite, e repetidamente
Olhei, em silêncio perfeito, para as estrelas.

Sumário

Lista de Tabelas.....	10
Lista de Gráficos.....	11
Lista de Figuras	12
Prefácio	13
Apresentação.....	17
Crônica I: Mediana e Média	23
Crônica 2: Desvio Padrão	40
Crônica 3: A Curva Normal.	58
Crônica 4: O Número “z”	74
Crônica 5: Correlação	88
Crônica das Mortes Anunciadas.....	131
Referências	146
Anexo: Tabela Z da Distribuição Normal	150

Lista de Tabelas

Tabela 1: Votos em Bolsonaro e óbitos por COVID-19, por Estado ..	27
Tabela 2: Porcentagem de votos em Jair Bolsonaro em ordem decrescente e mortalidade por COVID-19 por Estado.	30
Tabela 3: Percentual de votos em Jair (2018) e número de óbitos COVID-19 por Estado (16/04/2021): ordem decrescente por votos X por óbitos.	37
Tabela 4: Número de óbitos em ordem decrescente por Estado e o desvio em relação à média.....	47
Tabela 5: Porcentagem de votos em Jair Bolsonaro (2018) em ordem decrescente e desvio em relação à média.	49
Tabela 6: Cálculo do desvio padrão do número de mortes por COVID-19 em 16 de abril de 2021	51
Tabela 7: Cálculo do desvio padrão da porcentagem de votos.	55
Tabela 8: Casos de COVID, em ordem crescente e óbitos por COVID, desvios em relação à média e o desvio padronizado, por Estado, 16/04/2021.	90
Tabela 9: Multiplicação dos desvios padronizados dos casos pelos óbitos por Estado.....	99
Tabela 10: Óbitos por COVID, em ordem crescente, e votos, com os desvios em relação à média e o desvio padronizado por Estado brasileiro, 16/04/21.....	102
Tabela 11: Desvio padronizado da média de óbitos (2021) e de votos (2018), e produto dos desvios padronizados por Estado.....	112
Tabela 12: Número de óbitos por COVID por 100 mil habitantes (16/04/2021), percentual de votos em Bolsonaro (2018), desvios em relação à média e o desvio padronizado, por Estado brasileiro.....	116
Tabela 13: Desvio padronizado da média de óbitos por 100 mil (2021) e de votos (2018), e produto dos desvios padronizados por Estado.....	121
Tabela 14: Número de óbitos por COVID por 100 mil habitantes (03/12/2021), percentual de votos em Bolsonaro (2018), desvios em relação à média e o desvio padronizado, por Estado brasileiro.....	124
Tabela 15: Desvio padronizado da média de óbitos por 100 mil (03/12/21) e de votos (2018), e produto dos desvios por Estado.....	126

Lista de Gráficos

Gráfico 1: Cinderela escolherá entre os sete irmãos, um dos seis Anões ou o Gigante?	35
Gráfico 2: O desvio em relação à média dos seis anões e um gigante.....	41
Gráfico 3: Nota 10 em português para 36 alunos.	43
Gráfico 4: Notas variadas em matemática para 36 alunos.....	43
Gráfico 5: histograma estimado da altura do homem brasileiro	60
Gráfico 6: Percentual de homens com 1,70 m ou mais	61
Gráfico 7: Porcentagem de homens abaixo de 1,65m ou acima de 1,83m.....	62
Gráfico 8: Principais áreas sobre a curva normal.....	63
Gráfico 9: Distribuição dos Dados de Estatura dentro da Curva de Normalidade Esperada.....	72
Gráfico 10: Porcentagem de homens de acordo com o intervalo de 1, 2 e 3 desvios padrões em relação à média.	73
Gráfico 11: Distribuição Leptocúrtica no país dos Abacates	75
Gráfico 12: Distribuição Platicúrtica no país dos Abacaxis.....	75
Gráfico 13: Distribuição de frequências mostrando o número de suicídios por faixa etária.	78
Gráfico 14: Casos de COVID-19 por Estado, Brasil, 2021.	92
Gráfico 15: Óbitos por COVID-19 por Estado, Brasil, 2021.	93
Gráfico 16: Desvio Padronizado dos casos de COVID por Estado, Brasil, 2021.	96
Gráfico 17: Desvio Padronizado dos óbitos por COVID, por Estado, Brasil, 2021.	97
Gráfico 18: Desvio da média de óbitos por COVID-19 por Estado, Brasil, 16/04/2021.....	104
Gráfico 19: Desvio da média do percentual de votos em Bolsonaro, por Estado, 2018.....	105
Gráfico 20: Desvio Padronizado da média de óbitos por COVID-19 por Estado, Brasil, 16/04/2021.....	106
Gráfico 21: Desvio padronizado do percentual médio de votos em Bolsonaro, por Estado, 2018.....	107
Gráfico 22: Desvio da média do percentual de votos em Bolsonaro, por Estado, em ordem decrescente 2018.....	118
Gráfico 23: Desvio da média dos óbitos por COVID por 100 mil habitantes, por Estado, 16/04/2021.....	119
Gráfico 24: Desvio da média de óbitos por 100 mil Habitantes, por Estado, 02/12/21.	125
Gráfico 25: Correlação de 0,95 entre casos e óbitos por COVID, por Estado, em 16/04/2021.....	128
Gráfico 26: Correlação de 0,73 entre votos em 2018 e óbitos por COVID por 100 mil habitantes, por Estado, em 03/12/2021.	129

Lista de Figuras

Figura 1: Exercício de preenchimento dos valores de menor porcentagem de votos e de menor número de óbitos	26
Figura 2: diferença em relação à média do percentual de votos de Santa Catarina (SC) e Rio Grande do Norte (RN).....	50
Figura 3: A Régua do desvio padrão dos óbitos por COVID-19 em 16/04/2021.	53
Figura 4: A régua do desvio padrão das porcentagens de votos por Estado nas eleições de 2018.	56
Figura 5: O elefante engolido pela jiboia.....	65
Figura 6: Interpretação forçada do elegante.....	66
Figura 7: Pôr do sol na Gamboa, Salvador-BA.....	67
Figura 8: Pôr do sol na Gamboa com curva normal, Salvador-BA	67
Figura 9: A curva normal em forma de sino se manifestando na Pedra Azul, em Domingos Martins-ES.....	68
Figura 10: A trilha e a tendência a trafegar pelo centro	69
Figura 11: Estrada e o desgaste do tráfego	70
Figura 12: Ilustração do intervalo de variação do “R” de Pearson... ..	100
Figura 13: Exemplo de correlações: em cima, uma correlação quase perfeita, representada por dois cavalos puxando uma carruagem; embaixo, exemplo da ausência de correlação, representada pela ideia de cavalos soltos cavalgando aleatoriamente.	111
Figura 14: Exemplo de variável confundidora.....	111
Figura 15: Correlação moderada de 0,62 entre votos (2018) e óbitos por 100 mil (2021)	120

Prefácio

Caro Leitor,

O livro que tenho a honra de te apresentar tem um conteúdo forte, polêmico e cheio de fundamentos científicos. Nele, Paulo Moraes conta a infeliz história de como ocorreu a gestão da Pandemia de Covid-19 no Brasil, apresentando dados estatísticos que mostram uma relação significativa entre o aumento do número de mortos e a escolha política, sugerindo e sublinhando que a gestão executiva nacional contribuiu para expor seus próprios eleitores ao vírus, dados corroborados por pesquisas nos Estados Unidos e em outros países que tiveram o dissabor de serem liderados por homens públicos que optaram por negar a existência da pandemia e se opuseram às recomendações científicas.

Apesar de tratar de um assunto denso, Paulo apresenta argumentos diretos, simples e inteligentes. Infelizmente, demonstrando que os brasileiros, durante a pandemia, levaram uma vida esmagada por aflições, infortúnios e conflitos evitáveis se houvesse bom senso. Logo no primeiro capítulo são apresentadas as estatísticas descritivas da pandemia, onde os brasileiros, por ficarem à mercê do contexto de desinformação, vão continuamente recebendo notícias trágicas e veem seus números de mortos aumentarem assustadoramente, chegando ao ponto de colapsar os sistemas de saúde.

A infelicidade das atitudes governamentais segue os seus passos, como se elas fossem gigantescos ímãs que atráíssem desgraças. Neste pequeno volume, são apresentadas informações que ilustram que os brasileiros tiveram que lidar com um vírus letal e ao mesmo tempo com um repulsivo “vilão” dominado por

suas crenças, cobiça e desinformação, deixando a população confusa e desorientada.

Viveu-se uma “Tempestade Perfeita”, como conhecemos muito bem em fenômenos e sistemas complexos: é quando um conjunto de fatores convergem em sincronia para promover um determinado efeito avassalador. Aqui o elevado grau de transmissibilidade do vírus, associado ao conhecido contexto econômico e social do brasileiro, aos aspectos culturais (um povo que gosta de grandes eventos que favorecem o contato entre pessoas) e, sobretudo, com uma gestão executiva nacional desastrosa na figura do presidente da república criaram o cenário destruidor da pandemia que colocou o Brasil em uma situação de “caos” na saúde coletiva, nos colocando como um dos piores países em números de casos e de mortos.

Coube a Paulo o triste dever de pôr no papel essas histórias lamentáveis. Mas teve a sensibilidade de lembrar que por trás dos números e das estatísticas existem pessoas. Familiares, nossos conhecidos no bairro e habitantes de cidades que têm amigos e parentes, e que por conta de todo o processo foram privados de suas vidas. Um desastre calamitoso, capturado pela percepção científica de Paulo. Por isso, esse livro faz justiça para um enorme número de pessoas que tiveram suas vidas e as vidas de seus parentes ceifadas durante a pandemia.

Não, esse não é um livro político, é um livro técnico e científico que trata de um fenômeno social. Lamentavelmente, ele apresenta um plano infeliz de gestão que foi servido aos brasileiros como um verdadeiro mingau frio e envenenado, servido como café da manhã à gente humilde do nosso país, que evita se envolver na política. Portanto, é um livro que deve ser estudado, relido e entendido na sua profundidade, mas é também uma obra apimentada com literatura brasileira. Não há nada que o impeça de largar o livro imediatamente e sair para outra leitura sobre coisas alegres, se é isso que você prefere. Entretanto, a leitura desse livro se apresenta como um dever moral de cada

brasileiro que presenciou os fatos ocorridos nos últimos anos no Brasil. É um livro para ler, reler e procurar entender. E a partir desse entendimento, tomar atitudes para evitar que essa história, vexatória e assustadora, se repita.

É na condição de Cientista Chefe da HSI - *Human and Social Innovation* e tendo um trabalho voltado para a ciência de dados que recomento a leitura atenta e o estudo da obra de Paulo, consciente da posição de quem pode julgar uma obra que está prefaciando e avalizá-la para o público a quem se destina. Paulo não iniciou a escrita desse livro durante a pandemia. Tendo iniciado sua vida no mesmo grupo de iniciação científica que o meu, ele seguiu com rigor e competência sua carreira científica, fez sua graduação, mestrado e doutorado em sequência, ensinou nas melhores universidades da Bahia e atualmente é professor titular da UNEB – Universidade do Estado da Bahia, onde também coordena o curso de psicologia. São mais de 27 anos de dedicação e preparo para que essa obra pudesse ser escrita.

Esse livro nos traz a consciência da nossa humilde condição humana, pela nossa incapacidade de controlar ou dominar os fenômenos sociais e políticos. Mesmo em questões já bem descritas e respondidas pela ciência. O universo do conhecimento se constrói pela ação transdisciplinar, através das sucessivas gerações. Necessita de um campo de prática e de modelos para que sejam colocados em teste. Com o retumbante fracasso que tivemos enquanto sociedade, na gestão dessa crise, fica claro que, apesar de conhecermos muito sobre esse fenômeno, ainda somos incapazes de exercer ações articuladas no contexto social, político e ideológico.

O que Paulo Moares se propôs foi descrever os fenômenos vinculados à COVID-19, foi dar um passeio pelos dados. Ofereceu generosamente as percepções de uma mente com formação científica sólida. Suas impressões, sensações e emoções acumuladas no mergulho que ele fez no tema e descreveu nestas páginas são uma dádiva didática para a

disseminação da ciência estatística e psicológica, incentivando o pensamento crítico fundamentado. Particularmente, eu deverei voltar para muitas abluções, estudos e imersões, como quem volta a uma fonte para desfrutar do cenário e beber água limpa. O termo revisitar aplica-se plenamente. Portanto, espero que você também aprecie essa leitura tão bem elaborada e apoiada na capacidade, competência, conhecimento, formação e renome de Paulo Moraes.

Prof. Henrique Moraes

Chief Scientist of HSI - Human and Social Innovation – Washington - DC.
Cientista de Dados e Psicometrista. PhD, Mestre e Esp. em Psicologia. Consultor da ONU e Diretor de Políticas Públicas e Sociais da HSI.

Apresentação

Ao longo da construção deste livro, um certo incômodo surgiu no ventre dos pensamentos daquele que está “segurando a pena” e derramando frases no papel, ou seja, teclando palavras no editor de texto: quando a luz da razão acende e a visão inexorável da consciência captura a fotografia do drama, torna-se angustiante tratar de mortes através de números sem rostos. O sentimento de impotência é a consequência inevitável da participação numa guerra épica de proporção mundial que por vezes paralisou a pena do artista e insinuou uma desistência prematura. Como avançar sem ferir os irmãos que se encontram no exército inimigo? Num sentido espiritual, todos nesse planeta compartilham alguma humanidade que corre nas veias junto com o vermelho necessário para a vida que caracteriza esse corpo material comum da espécie humana, quente e pulsante. A irmandade inscrita no sangue da raça humana é universal e faz sonhar com uma sociedade solidária e favorável ao desenvolvimento pleno de quem quer que queira crescer, indo de encontro àquela paz que vem junto com a justiça, sem os grilhões da escravidão ou das injustiças historicamente construídas.

Como fugir de uma guerra civil fratricida? Depois de assistir a qualquer filme sobre algum conflito armado, pergunta-se por quem os sinos dobram. Mas a guerra contra um vírus é de uma natureza tão exótica e inusitada que nenhum romancista ou cineasta poderiam antecipar os detalhes desse enredo melancólico. A pandemia de COVID-19 aconteceu concomitantemente a uma polarização da política e isso

diminuiu a capacidade de amortecer o impacto da doença e desacelerou as medidas preventivas necessárias para combater a disseminação do vírus, cuja transmissão avança de pessoa para pessoa.

Meditando sobre as motivações para entrar nesse campo de batalha, é possível encontrar consolo nos conselhos de Krishna a Arjuna, tal qual se lê no Bhagavad Gita. Arjuna se negava a entrar na luta, pois não queria ferir seus parentes que estavam no exército adversário. Na condição de guerreiro virtuoso, nenhuma flecha sairia do seu arco para ferir desnecessariamente qualquer semelhante. Ele agia sempre guiado pela força das ações iluminadas pelo bem e pela sabedoria advinda do conhecimento do carma: toda ação produz uma reação.

Mesmo observando o coração ingênuo e puro do seu pupilo, Krishna começa a lhe ensinar que a ação guiada pelo bem também é terrena e carrega em si os interesses próprios: receber o bem de volta. Por isso, não deixa de ser uma ação interessada visando o benefício próprio dos resultados. Fazer o bem é uma ação mais inteligente do que se deixar levar pela paixão ou pela ignorância. Mesmo sendo a ação humana mais inteligente, há alguns inconvenientes na postura insistentemente bondosa. Muitas vezes, a pessoa se apega ao bem e deixa de fazer o que é justo, tornando-se relapso com os deveres. Uma das principais motivações do “bonzinho” é não querer entrar em conflito com algum irmão, de “sangue” ou de “humanidade”, que poderia sofrer diante de sua própria atitude ou comportamento adverso. Entrar em conflito é inevitável na dinâmica natural da vida em sociedade, mas a forma de resolver as diferenças e desavenças é uma escolha que envolve alternativas: atacar violentamente o oponente? Encarar o problema de frente com sabedoria? Ou negar o conflito para manter uma reputação frágil, de pessoa pacata e bondosa, empurrando a sujeita para baixo do tapete?

Existe uma ação que não pertence ao plano terreno e não está ligada a qualquer resultado imediato no plano material. É a ação desinteressada de natureza divina e de caráter libertador, pois faz cumprir um dever sagrado. Após explicar isso, a cartada decisiva para convencer o inigualável arqueiro foi lembrá-lo de que é impossível atingir a alma de alguém com uma flecha, pois ela afeta apenas o corpo corruptível e temporário. Não se faz mal à alma de ninguém. Só a própria pessoa se afeta com a transitoriedade do que acontece ao seu redor. Se negar a entrar numa guerra para produzir a justiça motivado pela necessidade infantil de preservar o próprio ego, que apenas calcula os benefícios imediatos de se resguardar numa bondade excessiva, poderá produzir efeitos mais adversos do que enfrentar os fantasmas da culpa ou responsabilidade, que uma ação assertiva poderia acarretar. A Negação da luta seria vergonhoso para os próprios adversários que esperam do guerreiro a coragem de combater pelo que é verdadeiro, mesmo tendo seus mestres e parentes do outro lado das trincheiras.

O autor desse estudo didático não teve nenhum parente próximo abatido letalmente por causa da pandemia. Até o momento em que essas palavras foram digitadas, se registraram no país mais de seiscentos e oitenta mil mortes. Ele é filho de uma classe média que tem mais condições de se cuidar e de se isolar nos modernos apartamentos-prisões. Isso, no entanto, não minimizou o sentimento de dor relacionado com o sofrimento de muitos amigos e parentes que tiveram que se recolher em luto, devido ao falecimento de entes queridos e próximos a eles. E é pelo sentimento que a motivação se agiganta para utilizar a arma-flecha-literatura como estratégia de defesa diante de um inimigo insidioso. Impossível saber, definitivamente, se o conhecimento e a inteligência libertam, mas no mínimo proporcionam uma defesa contra o engano e o engodo.

Se cada palavra representasse uma flecha que mata, seria difícil chegar até o fim dessa prosa didática. Teria que haver

muita inspiração divina, em cada letra de cada palavra, para que mesmo ferindo amigos e parentes queridos, algum benefício seja alcançado e alguma consciência seja lançada para compreender esse fenômeno que assola os brasileiros. Isso é alcançado através de ideias fundamentadas e do estudo aprofundado em pesquisas amplamente revisadas por todos os cidadãos que queiram verificar com os seus próprios olhos. Mas o objetivo vai muito além de ferir: é preciso dotar a coletividade com as mesmas habilidades do arqueiro, para que o povo não seja ludibriado e possa ter consciência ampliada do que acontece ao redor e assim ter poder de reação e defesa.

Os cristãos sabem que a arma de Jesus foi a palavra, que foi mais forte e durável do que qualquer objeto perfurante ou cortante. É de se lamentar que uma comunidade religiosa queira pegar em armas e deixar as palavras de lado, deixar o conhecimento da ciência desamparado e regredir para uma fase infantil, onde as pessoas estão pressas em falsas concepções como a de que a Terra é plana. Se for para pegar em armas, que sejam as flexas-palavras da sabedoria, que iluminam as decisões e nos livram do infortúnio.

Sabendo que não posso afetar a alma de ninguém, pois o espírito é eterno e pertence a outro plano, reanimo a pena-flecha para que ela cumpra o seu papel educativo: informar, ensinar e convidar para a arte da reflexão e sabedoria. Na certeza de que todas as almas que desencarnaram permanecem na eternidade projetando a sua luz própria, foi tomada a liberdade de tornar o texto mais leve e irônico, como um bom livro didático que procura tirar lições das adversidades que a vida impõe. Dessa forma o estatístico pede licença aos parentes daquelas pessoas que continuam desencarnado, dia após dia por conta de um vírus, para tratar da morte de forma abstrata, como números convincentes que se tornam flechas num arco e apontam para um objetivo comum: vencer um adversário difícil e malicioso, comumente chamado pelo nome “Ignorância”.

O conhecimento pode libertar, mas também pode ser arma fatal usada para ludibriar, confundir e torturar. É possível ensinar e aprender tudo o que o ser humano construiu ao longo da história. Entretanto, não se tem garantia do que será feito com todo esse conhecimento. A Ética é uma tentativa de desenvolver valores honestos e que promovam a verdade no benefício da coletividade. A dificuldade é que não há regras precisas, como uma receita de bolo, e as pessoas possuem muitas perspectivas sobre a verdade dos fatos. Um limite para as atitudes de alguém é verificar o dano causado a outro indivíduo e à coletividade. Espero que a consciência desse limite favoreça as melhores decisões sobre a utilização do conhecimento proposto nos capítulos seguintes.

Este livro começa explicando a natureza da média e da mediana na primeira crônica. Esse primeiro passo será fundamental para ensinar o conceito mais complicado de desvio padrão, que será apresentado na crônica seguinte. A crônica 3 tratará de um número abstrato chamado “z” que auxilia no cálculo de probabilidades estatísticas de maneira a ser possível comparar dados que pertencem a escalas desiguais. Esses assuntos introdutórios constituem a base do conceito de correlação estatística, que é a tentativa de avaliar se dois fenômenos que variam no tempo fazem isso de forma conjunta e paralela ou de forma independente e casual. Se tivermos a certeza de que um fenômeno aconteceu anteriormente ao outro, é possível indicar uma relação de causalidade, em que X pode causar Y.

Nos fenômenos sociais, a ‘certeza absoluta’ em relação à determinação de um fenômeno isolado é uma ficção, pois normalmente se trata de causas associadas a múltiplos fatores de difícil mensuração e compreensão. Entretanto, é possível utilizar técnicas para estimar o quanto um fenômeno X influencia um resultado Y, não em termos de tudo ou nada, mas buscando uma proporção ou probabilidade de influência. Esse é um modo de

apresentar dados que facilita na tomada de decisão e previsão de acontecimentos futuros. Na gestão de grandes países, a observação das tendências permite a ação antecipada para evitar catástrofes. Às vezes, a ciência erra, quando o fenômeno é muito complexo e pouco conhecido, mas, muitas vezes, ela acerta e faz a humanidade prosperar e aumentar a expectativa de vida. Vamos juntos colocar mais um tijolo nesse precioso edifício da ciência e colaborar com a construção desse baluarte da vida na terra.

A construção dos argumentos e da análise conta com o auxílio e a participação do leitor que está apontando as lacunas e fomentando as discussões em tempo real e a tempo de trazer o barco a prumo, corrigindo a rota da viagem para um porto seguro. Cada capítulo foi postado a cada mês com um espaço para crítica do leitor. Lá é possível deixar comentários, propor correções e criticar as ideias apresentadas. É um convite para embarcar nessa jornada da construção de um livro. Espero que quando as velas forem recolhidas numa baía segura e o livro for para a impressão, nesta mesma página, eu tenha que digitar, letra após letra, inumeráveis leitores que se deram ao trabalho de contribuir com a obra. Será que já tem isso: escrever junto com o leitor? Bem, ninguém inventa nada, vamos então reescrever do nosso jeito!

13/08/2022

Paulo Wenderson Teixeira Moraes

Professor Titular UNEB

pwmoraes@yahoo.com



<https://www.dain.art.br/obitosevotos>

Crônica I: Mediana e Média

Deus não joga dados no universo. Essa é uma célebre frase de Einstein que foi cunhada para expressar a improbabilidade das energias e forças presente no cosmo terem sido engendradas como um mero capricho do acaso. Para apreender detalhadamente a profundidade dessa lição é necessário visitar o país das Maravilhas da Matemática Complexa. Ele é inacessível para a maior parte da população e requer a iniciação na linguagem utilizada nessa terra exótica, que é abstrata e gera significados ocultos em números impossíveis para um mero mortal, igual ao escritor que vos conduz nesse estudo didático. Felizmente, o objetivo dessa crônica é mais humilde: compreender os números mais simples que são utilizados na Estatística básica e introduzir o conceito de mediana e de média através de dois fenômenos populacionais registrados na realidade brasileira: os votos por Estado no primeiro turno das eleições de 2018 e o número de mortes por COVID-19 em 16 de abril de 2021.

A aposta é uma das formas das mais antigas de se colocar em risco toda uma fortuna. Mas se o Todo Poderoso não joga dados, certamente alguém joga. Na Bíblia há indícios de que as forças tenebrosas tentaram apostar com Ele e colocaram em xeque a obediência do servo Jô. Já no Mahabharata, Yudhistira (filho de Darma, deus da Justiça) perdeu tudo num tabuleiro de dados. Tudo mesmo. Seu reino, sua riqueza e sua mulher, da qual ele só possuía um quinto, uma vez que os outros quatro irmãos possuíam as outras partes da maravilhosa e bela Draupadí. No livro *Psiquê e as Cartas Mágicas* (MORAES, 2020) é possível saber um pouco mais sobre esse prodígio de uma flor com 5 maridos, deixando o próprio Jorge Amado perplexo com esse

ancestral caso de um relacionamento pós-moderno de múltiplas faces que surgiu na face da terra, talvez antes do dilúvio universal. Sua inquietação, vale ressaltar, não é pela grandeza numérica do matrimônio ou pela carência de licença poética para criar uma imagem abstrata de uma mulher dividida em cinco, mas pela disposição de um marido em apostar a própria mulher numa mesa de azar.

Então, quem joga dados são os próprios homens. Por não possuir todo o conhecimento das engrenagens da sorte, o indivíduo comum aposta na esperança de que a providência divina bata repentinamente na porta. Depois de lançados os dados aparecem os números que refletem um acontecimento que se manifestou. Num tabuleiro justo, não se pode voltar atrás e o que foi lançado está dado. É o momento em que as probabilidades se manifestam com suas consequências que lhes são peculiares, podendo inclusive acarretar o óbito do próprio jogador, se o jogo for uma roleta russa.

Morrer ou viver pode ser resolvido em um simples jogo de dados. Cada atitude que a pessoa toma é um lance decisivo na vida. Na democracia, o lance de dados decisivo é o voto para eleger um representante. Cada um faz suas apostas esperando colher um determinado resultado e faz suas estimativas e prognósticos. Nem sempre somos bons jogadores e por descuido ou falta de conhecimento, a população acaba escolhendo o número mais improvável para realizar o sonho do prêmio da loteria de um país bem governado.

A ilusão com a qual o sujeito se agasalha fornece os argumentos para manter firme a teimosia num determinado número, linha de ação ou jogo. O sociólogo Howard Becker (1960) fez a teoria dos *side-bets* com esse fenômeno. A pessoa, quando entra num curso de ação, investe recursos dos quais deseja algum retorno. Com o tempo, a tendência é que ela aumente suas apostas no rumo que tomou, tornando ainda mais difícil abandonar a decisão inicial, pois a sensação de prejuízo ou

frustração, diante de uma possível perda de tudo o que já foi investido, conduz a uma estratégia de recuperação que compromete, ainda mais, os recursos disponíveis. Quando a decisão se mostra recompensadora, o indivíduo segue tranquilo. Quando não, a teoria explica a bancarrota, na qual muitos indivíduos acabam entrando, pelo simples fato de não conseguirem mudar o rumo de suas ações, tendo em vista que já investiram muito tempo e dinheiro na esperança de alguma compensação. É o autoengano que produz esse comportamento irracional. O antídoto para isso é encarar a insuportável sensação de perda. Entretanto, o ser humano parece ter sido programado para detestar qualquer tipo de derrota. É muito ruim perder e, ao tentar evitar um pequeno fracasso, na vã ilusão de recuperar o dinheiro perdido, o sujeito pode aumentar o prejuízo. Muitos jogadores ficam presos nessa armadilha, tentando repetir a sensação da grande vitória conquistada no passado e aumentam as apostas num rumo fadado à frustração e ao fracasso.

Um peso grande faz tensão no portador desse conhecimento que será apresentado a seguir de forma didática. Na mesma medida em que ele não gosta de perder, também não gosta de ser o veículo que comunica a derrota alheia. E nesse caso, estão todos no mesmo barco sofrendo com o naufrago acarretado por uma decisão coletiva equivocada. Ou teria sido acertada? Como mensurar o acerto de um voto? Enfim, ele se solidariza com aqueles que se encontram do outro lado da mesa e ainda não perceberam que são os bobos da vez. Tem um ditado no pôquer que diz assim: se você não sabe quem é o otário do jogo, se levante e vá embora, pois provavelmente é quem não sabe. O problema da situação atual é que todos perderam e ninguém saiu vencedor: os cidadãos foram feitos de bobos.

Não há mais cartas na manga e os dados da eleição de 2018 aqui no Brasil já foram lançados. Está disponível para todos. A pergunta que se faz é qual foi a consequência desse resultado que se apresentou? Impactou em óbitos ou em vidas?

É possível traçar alguma relação entre votos e óbitos? Como se diz no português vulgar, os dados falam por si mesmos. Na tabela 1, a seguir, tente observar alguma tendência entre os Estados que mais deram votos ao presidente Bolsonaro e a quantidade de mortes. O Estado que mais votou nele foi Santa Catarina (SC), com 61 por cento dos votos, lá embaixo da tabela, um pouco antes de São Paulo (SP). É possível perceber que os Estados se encontram ordenados por ordem alfabética das siglas e não dos nomes completos (SC antes de SP). Olhando para a coluna dos óbitos, se observa que o Estado com mais óbitos foi São Paulo (SP), com 87.326 registros de morte, enquanto Santa Catarina (SC) teve apenas 12.480 óbitos registrados, ambos na data de 16 de abril de 2021. Voltando para a coluna dos votos em 2018, SP obteve 48 por cento dos votos no Presidente. Nesse primeiro exercício de observação da tabela 1, o observador focou nos Estados que tiveram o maior valor em cada coluna e se constatou que Estados diferentes ficaram em primeiro lugar em cada uma delas. Mas não parece fácil fazer esse exercício sem as orientações prévias, pois não há uma ordenação adequada. A seguir, tente preencher o restante dos valores da parte de baixo da figura 1 com os Estados correspondentes aos menores valores:

ESTADO	Maior percentagem de votos em Bolsonaro em 2018	↑
SC	61	
	Maior número de mortes por COVID-19 em 16/04/2021	↑
SP	87.326	
	Menor percentagem de votos em Bolsonaro em 2018	↓
?	?	
	Menor número de mortes por COVID-19 em 16/04/2021	↓
?	?	

Figura 1: Exercício de preenchimento dos valores de menor percentagem de votos e de menor número de óbitos

Tabela 1: Votos em Bolsonaro e óbitos por COVID-19, por Estado

ESTADO	% votos Bolsonaro*	ÓBITOS**
AC	59	1.395
AL	31	3.921
AM	41	12.352
AP	39	1.456
BA	21	17.134
CE	20	16.075
DF	55	7.125
ES	50	8.581
GO	53	13.481
MA	23	6.783
MG	43	29.538
MS	56	5.099
MT	52	8.842
PA	34	11.635
PB	28	6.388
PE	27	13.102
PI	17	4.693
PR	53	20.182
RJ	54	40.716
RN	28	5.045
RO	59	4.737
RR	61	1.435
RS	49	22.977
SC	61	12.480
SE	24	3.903
SP	48	87.326
TO	41	2.348

* Fonte: TSE, Percentual de votos dados a Jair Bolsonaro no 1º turno das eleições presidenciais de 2018, considerando a soma total dos votos.

** Fonte: Óbitos registrados até 16/04/2021, às 18:10h no site <https://covid.saude.gov.br/>

Se você não perdeu muito tempo com a tabela 1, você é esperto e percebeu que os dados precisam ser trabalhados para poderem falar alguma coisa. É necessário ordenar os Estados por outra ordem que não seja a alfabética. Quem comenta sobre os dados é alguém que interpreta. A matemática só é perfeita nos modelos da abstração, mas quando ela é aplicada, está sujeita às imperfeições da matéria e aos erros de cálculo da humanidade. Se os dados não falam, alguém poderia argumentar que são os fatos que falam por si mesmos. É preciso lembrar que a palavra “fato” vem do latim e significa “feito”, fabricado. Possivelmente, o fato é algo produzido pelo homem e apresentado como verdade, trazendo alguma dúvida sobre a autonomia dos fatos em criar suas próprias realidades.

Procurando se situar um passo afrente da guerra de narrativas, a ciência utiliza da probabilidade na busca de se aproximar de algum fundamento para validar seus argumentos. Por outro lado, o povo quer uma verdade absoluta, nem que seja uma facilmente comprada no supermercado mais próximo, depois de uma propaganda apelativa para vender alguma verdade fabricada num alambique da esquina. A forma elegante do cientista se opor a isso sem dizer diretamente que alguém não tem razão em seus “fatos” e argumentos é declarar que ele mesmo pode se enganar com os seus cálculos e teorias. Isso cria uma empatia e uma humildade de reconhecer que não se tem a verdade absoluta numa máquina de calcular. Com esse cuidado da prudência, é possível fazer prognósticos. Mesmo assim, algumas vezes, a Estatística se surpreende com o surgimento do improvável e do imponderável. Para evitar um fracasso vergonhoso e a exposição vexatória de um revés qualquer, ela reconhece que pode errar, ela avisa antes e inverte a forma mais intuitiva de enunciar uma hipótese. Ao invés de dizer que “o número de mortes por COVID-19 em 2021 tem tudo a ver com a porcentagem de votos em 2018”, diz de forma mais comedida: “É provável que as diferenças entre os números de óbitos por

COVID-19 por Estados e as diferentes porcentagens da votação de 2018 em Bolsonaro não sejam mero fruto do acaso e existe uma associação entre esses dois fenômenos separados por 3 anos de diferença”. Essa é a hipótese experimental que anima a parte do eleitorado que votou contra Bolsonaro, os “comunistas”.

Tendo apresentado a hipótese de trabalho, outro procedimento comum é enunciar uma hipótese alternativa que é chamada de “nula”, pois normalmente é uma negação da hipótese experimental. Não tem nada a ver com voto nulo. É apenas um procedimento de estabelecer duas hipóteses que se contraponem. Pode ser enunciada da seguinte maneira: “É improvável existir alguma relação entre óbitos por COVID-19 em 2021 e a porcentagem de votos em Bolsonaro em 2018”. Essa é a hipótese que alegra o lado da torcida bolsonarista, constituída pelos denominados de “bolsominions”. O fato de ser “nula” não tem nenhum sentido pejorativo, é apenas um procedimento padrão estabelecido em muitas ciências que também avançam quando conseguem negar determinadas teorias e explicam com dados mais precisos o que antes se pensava como fato consolidado. Uma hipótese fraca facilmente é derrubada com o “tempo”, ou com a confrontação de outros experimentos que refutam, com dados comprobatórios, e impõem um “xeque mate” no tabuleiro da celeuma científica. Mas esse “fim da partida” pode ser o começo de outra e vem a necessidade de repensar continuamente o jogo no tabuleiro do xadrez acadêmico, que é a universidade em busca de produção de conhecimento científico.

Como expressa o mestre Cunha (2000):

Este desfiar de conjecturas tem o valor único de indicar quantos fatores remotos podem incidir numa questão que duplamente nos interessa, pelo seu traço superior na ciência, e pelo seu duplo significado mais íntimo no envolver o destino de extenso trato do nosso país.

A tabela 2, a seguir, é um primeiro esforço de ordenar os dados para desfiar as primeiras conjecturas e facilitar a observação de relações entre os números.

Tabela 2: Porcentagem de votos em Jair Bolsonaro em ordem decrescente e mortalidade por COVID-19 por Estado.

ESTADO	% votos Bolsonaro*	ÓBITOS**
RR	61	1435
SC	61	12480
AC	59	1395
RO	59	4737
MS	56	5099
DF	55	7125
RJ	54	40716
GO	53	13481
PR	53	20182
MT	52	8842
ES	50	8581
RS	49	22977
SP	48	87326
MG	43	29538
AM	41	12352
TO	41	2348
AP	39	1456
PA	34	11635
AL	31	3921
PB	28	6388
RN	28	5045
PE	27	13102
SE	24	3903
MA	23	6783
BA	21	17134
CE	20	16075
PI	17	4693

* Fonte: TSE, Percentual de votos dados a Jair Bolsonaro no 1º turno das eleições presidenciais de 2018, considerando a soma total dos votos.

** Fonte: Óbitos registrados 16/04/2021, <https://covid.saude.gov.br/>

Numa tentativa de fazer os dados falarem, o estatístico começa a organizar os números de uma forma mais didática para o entendimento do público. Um primeiro procedimento, ao abrir a arena da batalha entre as hipóteses, é ordenar os Estados pela porcentagem de votos obtidos pelo presidente Jair Messias Bolsonaro, do maior para o menor, e analisar a coluna dos óbitos para tentar perceber algum padrão associado.

Encabeçando a lista, estão empatados os Estados de Roraima (RR) e Santa Catarina (SC). De início, é possível perceber que a ordem alfabética foi dilacerada verticalmente explodindo num padrão aparentemente aleatório. O olhar agora é capturado pelo número 61 na primeira linha que representa o percentual de 61% de votos obtidos dos cidadãos roraimenses e dos Catarinenses, também chamados de “barrigas-verdes” (Não pesquisei o porquê, alguém poderia apresentar a explicação no espaço do leitor). São regiões caracterizadas por uma convicção fortíssima no líder que possui Messias no seu sobrenome.

Deslisando o olhar para a coluna adjacente à direita, um contraste evidente desfere o primeiro golpe na hipótese experimental: Roraima registou mil quatrocentos e trinta e cinco (1435) mortes por COVID-19 no dia 16 de abril de 2021, enquanto foram contabilizadas doze mil e quatrocentos e oitenta (12480) no Estado de Santa Catarina. Em termos de porcentagem, dividindo o primeiro número pelo segundo e multiplicando por 100, encontra-se o valor de apenas 11,5% de mortes no primeiro em relação ao segundo. O primeiro ponto marcado para os bolsonaristas vem desse comportamento aleatório dos números que parecem indicar nenhuma relação entre o precioso voto no “mito”, como é chamado o “Messias”, e as mortes por COVID-19 contabilizadas até 16 de abril de 2021.

O mesmo prodígio se repete no final da tabela. Os Estados povoados com menos bolsominions, que poderiam ser chamados de Estados de “esquerda” ou “comunistas” por um “bolsonarista raiz”, possuem 17 e 20, ou seja, em outras palavras,

o Piauí (PI) cedeu apenas 17% dos votos dos cidadãos para Bolsonaro e o Ceará (CE) apenas 20%. Olhando para os respectivos números de mortes, neste foram registrados um pouco mais de dezesseis mil óbitos (16075) enquanto naquele apenas quatro mil e seiscentos e noventa e três (4693).

O querido Estado de Minas (MG), quietinho, encontra-se no meio exato da tabela. Ele representa a **mediana**, ou seja, aquele que está no meio. São 27 Estados ao todo, sendo que 13 ficaram acima e os outros 13 abaixo, exatamente o mesmo número. Em outros termos, 13 mais 13 é igual a 26, somando com mais um que representa aquele que é mediano, inteiramos 27. Quando o grupo é par, a mediana fica prejudicada, pois dois indivíduos partilham a posição central. Não é o nosso caso e podemos seguir com o baile, guiados pelo exemplo do bom mineiro que fala pouco, come calado e canta bonito, não entra em discussões improdutivas e apenas aponta o caminho do meio. Lá foram registrados quarenta e sete por cento de votos no presidente (47%) e mais de vinte e nove mil óbitos registrados (29538).

Apenas olhar a tabela e tentar visualizar algum padrão é um processo cansativo e um tanto quanto improdutivo. Para estabelecer alguma ferramenta de comparação, o homem criou essa figura abstrata do ser médio que só habita nos interstícios da imaginação fantástica da razão pura e simples. Apesar do exemplar médio calculado para muitos fenômenos não ser facilmente encontrado nas ruas, circulando com seus pés firmes no solo da realidade, a média é um instrumento matemático que permite comparar os indivíduos de um determinado grupo entre si mesmos e assim classificá-los, comparando-os se estão acima ou abaixo do ideal aritmético ou, o que vem ao caso, avaliando se estão mais à esquerda ou à direita em relação a um ponto central.

Como, então, obtém-se esse ser fantástico da média que bem poderia ser incluído no rol dos seres imaginários da enciclopédia de Borges(2000)? Apenas somar e dividir! Vamos

então contar o número de Estados: acima já foi dito que são vinte e sete (27). É recomendável sempre contar novamente para checar se a razão não falhou e contou a mais ou a menos. Mas adiante será visto que isso aconteceu com o Estatístico quando fazia as contas e somente na terceira recontagem ele percebeu o erro. É preciso considerar que a mente humana falha, mas que é capaz de correção através da disciplina e do reconhecimento da possibilidade do equívoco.

O segundo passo demanda um esforço maior. Porém as máquinas podem fazer o trabalho pesado para a gente. Basta digitar os números numa planilha e programar a operação de somar cada número que foi digitado numa célula da tabela. No lápis e no papel, como o escritor aprendeu quando frequentou a escola, é preciso somar todos os números da coluna e gerar um somatório total que será dividido por 27. O número resultante é a média. Na tabela 3, a seguir, encontram-se o somatório total e a média. É praticamente a mesma tabela que a anterior, apenas com o acréscimo dessas informações. Observa-se que o somatório das percentagens de votos chega até 1127 (que não tem muito sentido por si só, tendo em vista que o número 100 representa 100 por cento dos votos e acima disso é apenas um artifício para chegar à média) e que os óbitos totais vão até trezentos e sessenta e oito mil e setecentos e quarenta e nove (368749). Ao dividir por 27, obtém-se o percentual médio aproximado de votos de 42% e o número médio de mortos de treze mil e seiscentos e cinquenta e sete (13657). Nenhum Estado listado possui esses valores, apenas o Estado abstrato que habita no país das maravilhas matemáticas.

Como se vê, o cálculo é relativamente simples. O difícil é encontrar o exemplar médio como uma realidade concreta. Se uma gigante de dois metros se casasse com um pigmeu de um metro, eles se encontrariam em algum momento na média deles que é um metro e meio. Ali, naquele ponto improvável, talvez o amor deles se realizasse plenamente. Por esse novo critério,

Minas Gerais perderia um pouco da sua temperança caracterizada pelo seu lugar mediano, ou seja, deixaria de ser aquele que se localiza exatamente no meio da tabela. Esse é o conceito de mediada: o indivíduo que está localizado no meio, um conceito intuitivo que rivaliza com a média. A média de 0,42 deslocou levemente esse Estado para a posição de “acima da média” e não mais de “exatamente” no meio. Ainda assim, permanece na mente a imagem de um mineiro que “come quieto” e que sem fazer barulho, permanece no seu lugar de equilíbrio sem que nenhum outro indivíduo reivindique para si o lugar de ponto central. Acredito que nem o Amazonas e nem o Tocantins querem subir nessa tabela de morte, então deixa quieto. Média e mediana se aproximam nesse conjunto de dados fazendo com que uma ou outra sejam relativamente intercambiáveis. Isso indica que os dados não estão distorcidos, pois quanto mais próximo a média for da mediana mais nítida é a percepção de uma distribuição equilibrada dos dados de um grupo.

A mediana é interessante quando há alguma distorção na média. Por exemplo, numa determinada família de sete filhos surgiu milagrosamente um irmão mutante com gigantismo, formando um conjunto improvável com seis indivíduos de um metro e quarenta centímetros (140cm) de altura e apenas um com 3 metros e 50 centímetros (275cm). Nesse conjunto surreal, a mediana representa melhor o grupo do que a média. O quarto indivíduo teria a altura de 1,40 metros, portanto, a mediana seria mais representativa do todo. A média aritmética seria de 1,70 metro, que é obtida com a soma de todas as alturas dos indivíduos, dividida pelo número total de elementos do conjunto, no caso sete. Na verdade, apenas um indivíduo possui altura maior do que um metro e setenta (170cm). Além de estar mais perto das nuvens, esse ser imaginário de altura singular, que sobrepujou o recorde de Robert Wadlow, cuja estatura alcançou os estonteantes 2,72 metros, puxa a tendência do grupo para o alto através de um artifício matemático. Se alguma Branca de

Neve estivesse procurando um par perfeito para sua altura de um metro e setenta (1,70m), fantasiando a medida exata do amor para lhe completar na solidão das noites românticas, poderia facilmente pensar que encontraria seu pretendente nesse grupo, quando na verdade só haveria um gigante no meio de seus seis irmãos anões (um anão, tecnicamente, tem abaixo de 1,40m). Eis um perigo de utilizar apenas a média para tomar decisões fundamentais na vida. Talvez por isso, ela rejeitou os sete e esperou pelo fantasmagórico e retardado príncipe encantado. No gráfico 1, a seguir, é possível explorar a distância que os indivíduos têm em relação à média.

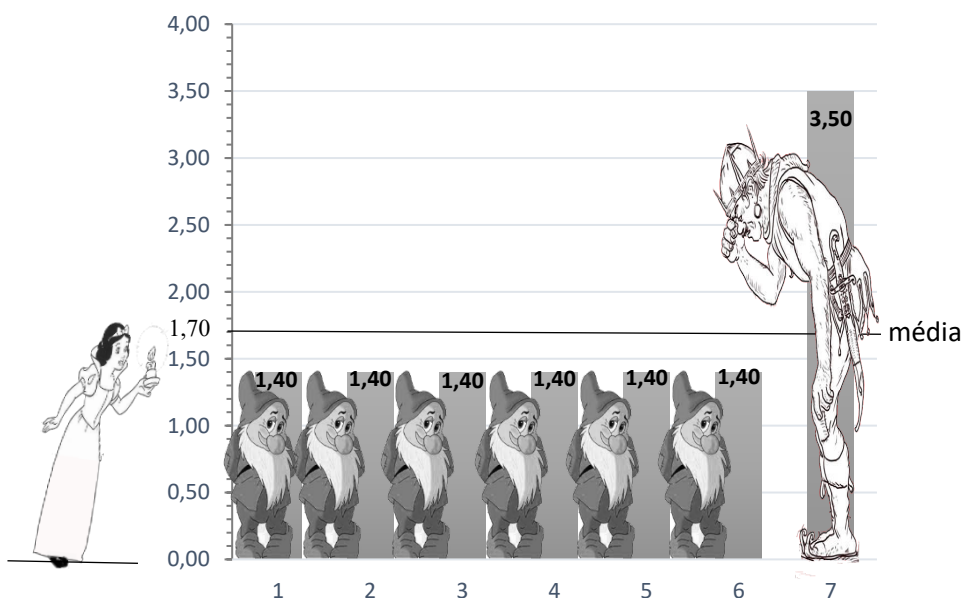


Gráfico 1: Cinderela escolherá entre os sete irmãos, um dos seis Anões ou o Gigante?

Uma forma de melhorar a visualização dos dados e da média é criar uma tabela de acordo com a ordem do número de mortes, gerando uma nova organização que acrescenta novas informações para a dinâmica dos dados. Na tabela 3, a seguir,

encontram-se do lado esquerdo, os Estados perfilados em ordem decrescente do percentual de votos no Presidente em 2018, enquanto do lado direito, vê-se os Estados organizados por ordem decrescente de mortes registradas por COVID-19 em 16/04/2121. Encabeçando a lista da esquerda, estão os Estados que mais ofereceram votos ao presidente: Roraima (RR) e Santa Catarina (SC) com 61%. Já no final da tabela, pelo lado esquerdo, vê-se o Piauí (PI) como o Estado que menos destinou votos, 17%. Passando para o lado direito, os dois Estados que se encontram no final são Roraima (RR) e Acre (AC): eles caíram do topo para a base ao trocar o ordenamento dos votos pelo dos óbitos. Ainda no ordenamento decrescente do número de óbitos colocado na parte direita da tabela 3, o Estado de São Paulo alcança a primeira posição, seguido do Rio de Janeiro e, pasmem, Minas Gerais. Seria precipitado retirar tudo o que foi dito anteriormente? Alguma coisa está errada com esses dados? Ou será mais um gol marcado para o time bolsonarista, fortalecendo a hipótese nula? Olhado para um lado e para o outro, parece não haver nenhuma relação entre os números absolutos de mortes e a votação. Os dois lados da tabela apresentam ordens muito próprias e distintas. Nenhum Estado permaneceu no mesmo lugar. Apenas se algum sergipano desconfiado olhar atentamente para a parte inferior da tabela, será constatado que seu povo é muito bem-educado e sabe permanecer no mesmo lugar quando está sendo observado, tentando dar alguma ordem para essa bagunça de dados desordenados aleatoriamente. É o único Estado que permanece na mesma colocação nas duas tabelas: o quinto com menor número de mortes e o quinto com menor porcentagem de votos no presidente em 2018. Talvez no futuro, esse dado possa ser relevante para o estudo do comportamento desse querido Estado que possui um povo acolhedor e elegante.

Tabela 3: Percentual de votos em Jair (2018) e número de óbitos COVID-19 por Estado (16/04/2021): ordem decrescente por votos X por óbitos.

ESTADO	% votos*	ÓBITOS**	Estado	Óbitos	% votos
RR	61	1435	SP	87326	48
SC	61	12480	RJ	40716	54
AC	59	1395	MG	29538	43
RO	59	4737	RS	22977	49
MS	56	5099	PR	20182	53
DF	55	7125	BA	17134	21
RJ	54	40716	CE	16075	20
GO	53	13481	GO	13481	53
PR	53	20182	PE	13102	27
MT	52	8842	SC	12480	61
ES	50	8581	AM	12352	41
RS	49	22977	PA	11635	34
SP	48	87326	MT	8842	52
MG	43	29538	ES	8581	50
AM	41	12352	DF	7125	55
TO	41	2348	MA	6783	23
AP	39	1456	PB	6388	28
PA	34	11635	MS	5099	56
AL	31	3921	RN	5045	28
PB	28	6388	RO	4737	59
RN	28	5045	PI	4693	17
PE	27	13102	AL	3921	31
SE	24	3903	SE	3903	24
MA	23	6783	TO	2348	41
BA	21	17134	AP	1456	39
CE	20	16075	RR	1435	61
PI	17	4693	AC	1395	59
Total	11,27	368749	Total	368749	11,27
Média	42	13657	Média	13657	42

Fontes: * TSE e **<https://covid.saude.gov.br/>

Os goianos (GO) também estão reivindicando um elogio pela sua postura uniforme: esse estado também permaneceu no mesmo lugar. Então, corrigindo: apenas dois Estados obedeceram à mesma colocação quando foram ordenados de acordo com duas diferentes categorias. Eles estão no planalto central que poderia estar influenciando nesse comportamento retilíneo. A vivência no Serrado, que tanto nos honra com sua fauna e flora ímpar no planeta, estaria relacionada com esse fenômeno? Aqui mais uma vez o caminho se abre para aventuras especulativas ainda sem fundamento em tradições de pesquisas consolidadas.

Algum bolsonarista poderia notar que a Bahia (BA) deu um salto do fim da fila na esquerda para o quinto lugar na direita. Isso se deve ao fato de o governador ser um “comunista”, como eles gostam de vociferar? Será mais um ponto marcado para a hipótese nula? Aqui talvez o raciocínio poderia até ser o inverso: poucos votos, mais óbitos. Mas essa percepção se esbarra no fato de que todos os outros colocados acima tem expressiva votação no presidente, fazendo da hipótese nula apenas a constatação de que o fenômeno do voto não está relacionado de forma clara e linear com o registro das mortes.

Parece que o rumo das análises ganhou uma direção imprevisível, como se o estatístico, comandante do barco, estivesse à deriva em alto mar. O que poderia nos socorrer? Alguma ferramenta de medição que nos dê alguma orientação? Algum astrolábio ou bússola que nos socorra? Sim! Temos ciência de alguns procedimentos que trazem luz e direção para essa navegação errante. Antes do GPS, o navegador procurava medir a distância das estrelas para o horizonte e, dessa maneira, conseguir indicadores para traçar a sua rota em mapas bastante rudimentares, se comparados com os computadores de bordo atuais. Imagine cruzar o Atlântico olhando as estrelas? Pode não ser tão preciso, mas certamente é bastante romântico! Talvez aí esteja o segredo do marinheiro, despertador do amor de muitas

mulheres: desbravar o mundo com um “GPS das cavernas” equipado em suas próprias cabeças, o fez desenvolver um mapa metal do coração feminino que o orienta na navegação dos mares perigosos da paixão.

O próximo instrumento a ser apresentado para amadurecer a análise é o desvio padrão, que será abordado na próxima crônica e dará um rumo mais adequado para essa viagem matemática no planeta da Estatística. Assim será possível começar a corrigir o rumo a tempo de alcançar o destino, são e salvos.

Crônica 2: Desvio Padrão

A média aritmética é um modelo tomado como referência para comparações em diversas áreas da sociedade. Longe de estar próxima de algum ideal, ela é uma referência conveniente em determinadas situações, mas impertinente em outras. No exemplo dos anões, no qual a Branca de Neve pretendeu encontrar sua “cara-metade” de um metro e setenta de altura num grupo imaginário de irmãos, seria possível traçar uma reta e calcular o quanto cada irmão desvia em relação à média. Então o modelo tem um erro em relação ao que é encontrado na realidade desse conto de fadas distópico. A maioria dos modelos encontra esse problema de ajuste. A ciência busca ajustar os erros para diminuir a discrepância entre a teoria e a experiência. Quanto menor o erro encontrado, maior é a capacidade de um modelo fazer alguma predição útil para a sociedade. Quando o erro é muito grande, é necessário reformular as teorias e aperfeiçoar as técnicas de contagem ou de observação para desenvolver modelos mais eficazes. No gráfico 2, a seguir, é possível ver que cada um dos anões teve um desvio de 0,30 metros, ou 30 centímetros, abaixo da média. Já o irmão mais novo, porém bem-dotado, teve um desvio para cima de um metro e oitenta centímetros (1,80m). Para calcular o desvio, é preciso diminuir da média o valor da altura de cada indivíduo. Por isso, o desvio de um deles é negativo, uma vez que é necessário tirar 1,80m da sua altura de 3,50m para alcançar a média de 1,70m que fica bem abaixo de sua estatura.

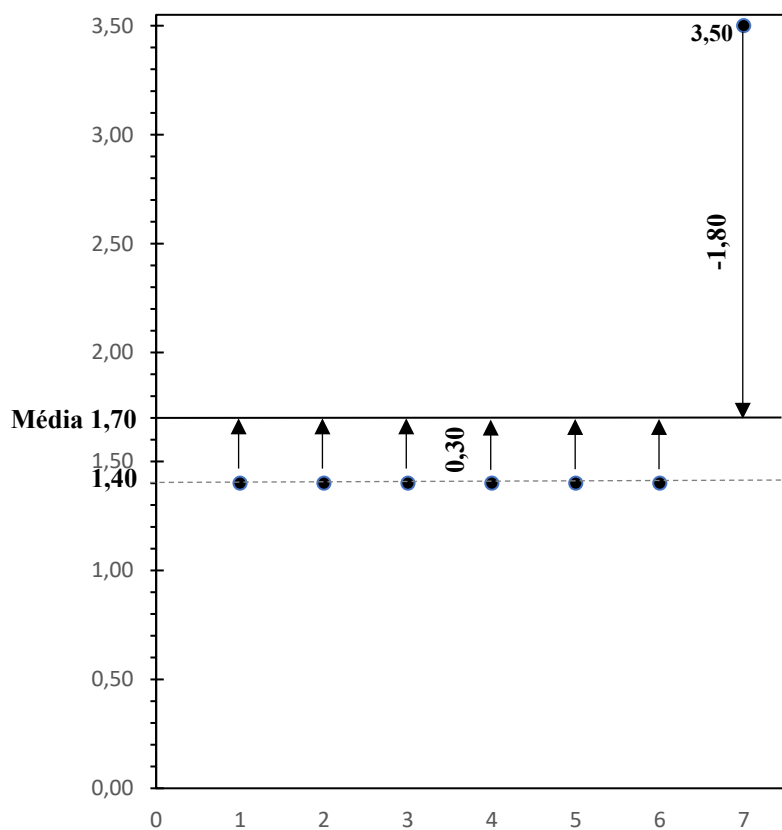


Gráfico 2: O desvio em relação à média dos seis anões e um gigante.

Já no campo escolar, se uma criança obteve uma nota 10 em português, mas todos os seus 35 colegas também tiraram dez, ela tende a sonegar essa informação dos pais, pois isso parece indicar que sua nota está mais relacionada com a bondade do professor, que acariciou a todos com a nota máxima do que com o seu empenho propriamente dito. Entretanto, vamos combinar que detalhes aritméticos são coisas chatas e os responsáveis não precisam ficar sabendo disso, tornando a vida mais simples e facilitando a negociação familiar de algum favorecimento no final de semana, como um presente ou um sorvete, afinal, foi uma nota 10! Dez é dez, isso é um fato! Entretanto, para ter sucesso nessa manobra de ludibriar os pais incautos, o aluno deve se certificar de que seus pais não leram esse livro ou de que não detêm um conhecimento mínimo de matemática. Quando distribuído em larga escala, facilmente se percebe que o “dez” pode não significar um resultado tão expressivo de um esforço hercúleo de concentração e estudo, criando uma banalização do símbolo máximo do desempenho escolar e o fato passa a ser falso (*fake*).

Numa situação adversa, porém, a média pode socorrer o aluno que não foi agraciado pela “sorte” de “receber” uma boa nota. Se ele tira uma nota cinco em matemática, pode se apoiar no fato de que a média da turma também foi essa, portanto, ele ainda ficou dentro de um intervalo razoável quando comparado ao grupo total de 36 alunos. O desempenho medíocre, portanto, pode ter sido decorrente da severidade do professor que fez uma avaliação mais difícil do que o nível de preparação da turma. Aqui a média pode arrefecer o humor dos pais e amenizar um possível castigo por conta de baixo desempenho escolar. Nos gráficos 3 e 4 a seguir, é possível observar dois comportamentos bem diferentes em português e matemática. Na primeira matéria, todos obtiveram nota dez. Já na segunda, apenas uma pessoa teve a nota máxima. A frequência com a qual as notas foram tiradas é muito diferente e gerou imagens distintas no gráfico de barras de frequência, também chamado de histograma.

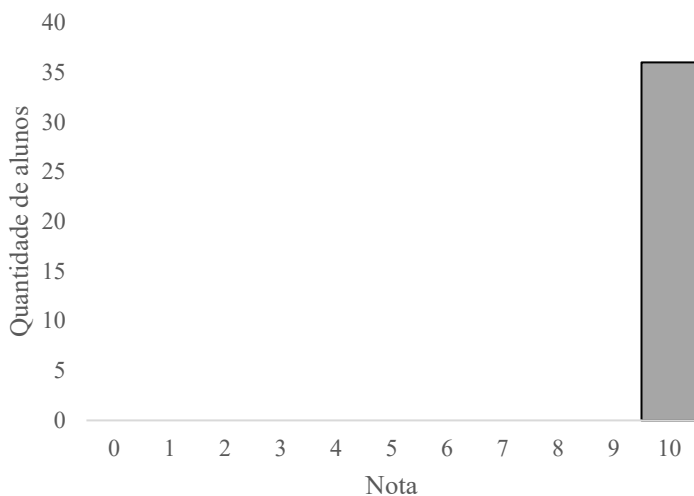


Gráfico 3: Nota 10 em português para 36 alunos.

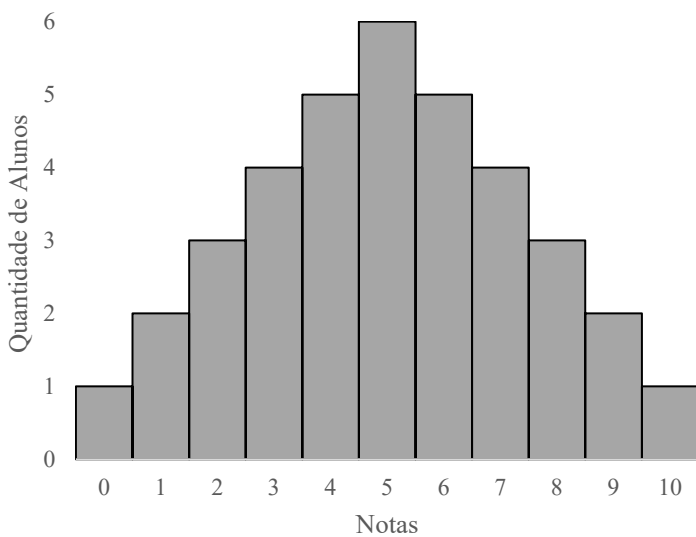


Gráfico 4: Notas variadas em matemática para 36 alunos.

No histograma, o eixo “Y”, que nesses dois gráficos foi denominado de “quantidade de alunos”, contabiliza a quantidade de ocorrências dos valores de uma determinada variável. No exemplo, a variável é a nota obtida e tem a amplitude de 0 a 10. No gráfico 3, como todos os alunos tiraram a mesma nota, o eixo “Y” vai até o número 40, para assim acolher a altura da barra que vai até 36 e corresponde ao número total de alunos da turma que tiraram nota dez. Já no gráfico 4, o eixo “Y” vai apenas até o “6”, pois esse número é suficiente para abarcar a altura da barra maior que fica na coluna 5 do eixo “X”, que representa a nota tirada. Portanto, seis alunos tiraram nota 5, cinco alunos tiraram nota 6, quatro nota sete, e assim sucessivamente.

Mas existem provas, ou variáveis, que não são registradas dessa forma discreta: ou uma nota ou outra, ou acertou ou errou uma questão. Pode ser uma avaliação com perguntas abertas em que o aluno pode ficar com notas entre os números inteiros. Nesses casos em que a variável é contínua e pode assumir mais valores, a barra observada no gráfico se torna um intervalo. Poderíamos, então, transformar a primeira barra em alunos que tiraram de “0” até “1”. Na segunda barra, os que tiraram de “1,1” até “2”. Repetindo esse procedimento até englobar todos os valores possíveis, se chegaria ao último agrupamento que seria dos alunos que obtiveram nota entre 9,1 e 10. Dessa outra forma, quem tirou uma nota “8,5”, por exemplo, seria contabilizado dentro da barra que representa a classe de valores compreendidos entre 8,1 e 9.

Olhando o gráfico 4 de matemática, um pai poderia ficar mais tranquilo ao ver que “15” colegas de seu filho estão abaixo dele e “6” estão iguaizinhos na mesma barra, ou seja, tiraram nota cinco. Com esses dados, ainda é possível pensar de forma probabilística: qual é a probabilidade de sortear uma criança desse grupo que tirou nota igual ou inferior a cinco? A resposta é dada ao somar os “15” que tiraram nota inferior com os “6” que

tiraram nota igual a cinco e dividir o somatório pela quantidade total de alunos, “36”.

$$p = \frac{21}{36} = 0,5833$$

O valor é de aproximadamente 0,58. A chance de escolher alguém com nota cinco ou inferior é de cinquenta e oito por cento (58%), aproximadamente. Traduzindo esse resultado, é possível dizer que 58% dos alunos ficaram numa condição igual ou inferior ao filho querido. Em termos contrários, calcula-se, também facilmente, a probabilidade de sortear algum aluno com nota superior a cinco, contabilizado a altura de cada barra seguinte a essa nota: $5+4+3+2+1 = 15$ e dividindo pelo número total de participantes do grupo: $15/36 = 0,4166$.

$$p = \frac{5 + 4 + 3 + 2 + 1}{36} = \frac{15}{36} = 0,4166$$

Assim, o filho pode dizer para o pai: “não fique triste papai, apenas quarenta e dois por cento (42%) dos alunos, aproximadamente, tiraram nota acima de cinco”. Portanto, histogramas podem também informar sobre a probabilidade da ocorrência de determinados valores de uma variável.

Da mesma maneira que o filho manipula os pais, muitas vezes a média é omitida ou explicitamente declarada a depender da conveniência. No caso da contabilização de mortes, os Estados sempre buscam estar abaixo da média, pois isso é um indicativo de que há um cuidado com a saúde para evitar mortes desnecessária. Isso não significa que a média é um número ideal. Se todo o trabalho de prevenção, medidas sanitárias e vacinação fosse feito com perfeição e antecipadamente, o ideal seria zero número de mortes. Mas as autoridades demoraram um bom

tempo para tomar providências e o vírus se espalhou de forma descontrolada, produzindo um estado de epidemia planetária, que mereceu o nome de pandemia. Tendo se espalhado e causado mortes, a tarefa que resta é verificar se a transmissão deu de forma uniforme e igual para todos os lugares ou se houve diferenças na velocidade e intensidade da disseminação.

A média se torna uma ferramenta para examinar a regularidade dos casos diagnosticados por região. Estar acima ou abaixo pode ser uma informação crítica que leva a consequências semelhantes à recompensa ou a punição dos pais, guardadas as devidas proporções. Governantes que obtiveram “nota baixa” na gestão da epidemia podem ser punidos nas próximas eleições e convidados a deixar o poder. Um indicativo de nota baixa é o número de mortes que pode ter saído do controle. Como o filho esperto, alguns governantes tentam minimizar o efeito de uma “nota baixa”, justificando os dados de contágio e morte como sendo reflexos de um fenômeno mais amplo e incontrollável, como se tirar nota boa fosse uma questão apenas de sorte e não de esforço e treinamento.

Para observar os “alunos-governantes” de forma mais acurada e comparar as suas performances na gestão dessa crise de saúde pública, o desvio é uma ferramenta interessante de organização das distâncias em relação à média. Na tabela 4, a seguir, é possível observar o número de mortes por Estado e seus respectivos desvios em relação à média. Em São Paulo, o número de mortes atribuídas ao coronavírus em 16 abril de 2020 foi de 87.326. Diminuindo a média de 13.657,37, chega-se ao número de 73.669, que significa o quanto esse Estado ficou acima da média. Quando o Estado tem menos mortes que a média, esse número fica negativo. Roraima, por exemplo, teve 4.737 mortes por COVID-19 nessa mesma data. Diminuindo da média, chega-se ao valor de -8.920 na coluna da direita, denominada de “Desvio”. Esse valor negativo significa que o Estado está 8.920 mortes abaixo da média.

Tabela 4: Número de óbitos em ordem decrescente por Estado e o desvio em relação à média.

	Óbitos	Desvio
SP	87.326	73.669
RJ	40.716	27.059
MG	29.538	15.881
RS	22.977	9.320
PR	20.182	6.525
BA	17.134	3.477
CE	16.075	2.418
Média	13.657,37	
GO	13.481	-176
PE	13.102	-555
SC	12.480	-1.177
AM	12.352	-1.305
PA	11.635	-2.022
MT	8.842	-4.815
ES	8.581	-5.076
DF	7.125	-6.532
MA	6.783	-6.874
PB	6.388	-7.269
MS	5.099	-8.558
RN	5.045	-8.612
RO	4.737	-8.920
PI	4.693	-8.964
AL	3.921	-9.736
SE	3.903	-9.754
TO	2.348	-11.309
AP	1.456	-12.201
RR	1.435	-12.222
AC	1.395	-12.262
Total	368.749,00	0,00

Olhando para o somatório total dos desvios se encontra o resultado zero. É um resultado esperado, pois se a média é um instrumento de definição de uma tendência central, os casos que ficaram acima da média tendem a anular os que ficaram abaixo dela. Por causa disso, é impossível calcular a média dos desvios, uma vez que zero dividido por qualquer número, dificilmente, sairá do zero. É preciso então fazer algum artifício no intuito de conseguir uma medida auxiliar para complementar o desvio bruto da média. Antes de adentrar no campo do cálculo do desvio padrão, é importante notar que os desvios são uma medida relacionada à média específica com a qual se está trabalhando. Mudando a medida, aparentemente altera a visualização da escala com a qual se está trabalhando. Por exemplo, na tabela 5 a seguir, é possível repetir o mesmo procedimento realizado com os óbitos utilizando os dados do percentual de votos no presidente em 2018. Se a média anterior refletia a contagem média de óbitos, agora a nova média expressa o percentual médio de votos entre os Estados. Apesar da diferença da medida, o procedimento é similar e produz efeitos comparáveis: criar uma referência para caracterizar o centro do grupo em relação a uma característica de interesse de observação. Enfim, primeiro observamos a média dos óbitos e depois a média do percentual de votos.

Os leitores mais atentos perceberam que anteriormente eu aproximei o percentual médio de votos para 42%, reduzindo assim os dígitos da fração para tornar o número mais elegante. Mas, para fazer o cálculo do desvio com precisão, a elegância pode gerar um resíduo de 0,07 quando realizado o somatório total dos desvios. Isso poderia atormentar aqueles mais atenciosos e dignos de São Tomé: além de precisar “ver para crer”, precisam calcular para ter certeza! Portanto, para se alcançar o “zerinho” redondo e perfeito demanda-se algum esforço: é necessário utilizar um número mais preciso com duas casas após a vírgula: 41,74.

Tabela 5: Porcentagem de votos em Jair Bolsonaro (2018) em ordem decrescente e desvio em relação à média.

Estado	% votos Jair	Desvio
SC	61	19,26
RR	61	19,26
RO	59	17,26
AC	59	17,26
MS	56	14,26
DF	55	13,26
RJ	54	12,26
PR	53	11,26
GO	53	11,26
MT	52	10,26
ES	50	08,26
RS	49	07,26
SP	48	06,26
MG	43	01,26
Média	41,74	
TO	41	-0,74
AM	41	-0,74
AP	39	-2,74
PA	34	-7,74
AL	31	-10,74
RN	28	-13,74
PB	28	-13,74
PE	27	-14,74
SE	24	-17,74
MA	23	-18,74
BA	21	-20,74
CE	20	-21,74
PI	17	-24,74
Total	1127	0,0

É possível observar que o maior desvio foi em Santa Catarina (SC), quase vinte pontos percentuais (19,26) acima da média geral de votos no então candidato. Já o Rio Grande do Norte (RN) teve um desvio negativo de -13,74. Aproximando ou arredondando a cifra, é possível dizer que esse Estado teve quase quatorze pontos percentuais abaixo da média geral de votos doados ao Presidente.

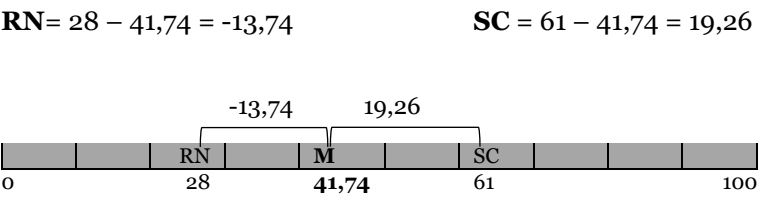


Figura 2: diferença em relação à média do percentual de votos de Santa Catarina (SC) e Rio Grande do Norte (RN)

Dessa forma, o desvio por si só é uma medida muito específica do material com o qual se está trabalhando. A fim de ter algum parâmetro de comparação, os cientistas propuseram um procedimento que transforma os desvios numa medida mais geral, possibilitando uma visão mais acurada dos dados. Eles deram o nome de desvio padrão ao procedimento de elevar ao quadrado os desvios que vimos anteriormente, somar todos eles, tirar a raiz quadrada e dividir pelo número de casos ou indivíduos. Na tabela 7, a seguir, é possível visualizar como o processo é realizado. A artimanha de elevar ao quadrado faz com que os números negativos se tornem positivos e o resultado da soma deixe de ser zero. Isso faz com que se encontre uma medida de quanto os indivíduos de uma amostra se afastam da média, em média. Fazendo uma analogia com os conceitos de segunda ordem de Bateson(1999), pensar sobre o pensar ou aprender a aprender, aqui se trata da média dos desvios da média.

Tabela 6: Cálculo do desvio padrão do número de mortes por COVID-19 em 16 de abril de 2021

Estado	Óbitos	Desvio	Desvio ²
SP	87.326	73.669	5.427.066.991,51
RJ	40.716	27.059	732.169.437,43
MG	29.538	15.881	252.194.397,43
RS	22.977	9.320	86.855.496,43
PR	20.182	6.525	42.570.791,80
BA	17.134	3.477	12.086.953,58
CE	16.075	2.418	5.844.933,03
Média	13.657	0	290.068.970,53
GO	13.481	-176	31.106,51
PE	13.102	-555	308.436,25
SC	12.480	-1.177	1.386.200,99
AM	12.352	-1.305	1.703.991,80
PA	11.635	-2.022	4.089.981,91
MT	8.842	-4.815	23.187.791,80
ES	8.581	-5.076	25.769.536,14
DF	7.125	-6.532	42.671.862,66
MA	6.783	-6.874	47.256.967,99
PB	6.388	-7.269	52.843.745,58
MS	5.099	-8.558	73.245.703,40
RN	5.045	-8.612	74.172.923,40
RO	4.737	-8.920	79.573.007,54
PI	4.693	-8.964	80.359.936,14
AL	3.921	-9.736	94.796.907,99
SE	3.903	-9.754	95.147.741,32
TO	2.348	-11.309	127.901.858,17
AP	1.456	-12.201	148.873.438,91
RR	1.435	-12.222	149.386.337,47
AC	1.395	-12.262	150.365.727,10
Total	368.749,00	0,00	7.831.862.204,30
Desvio Padrão			17.031,41

Os números começam a ficar estratosféricos, pois quando se eleva o desvio de São Paulo ao quadrado, por exemplo, se chega a uma cifra superior a 5 bilhões. Será que existe alguém que ainda faz esses cálculos no papel? Eu usei o Excel e confiei na máquina programada pela Microsoft para ganhar mais tempo na escrita. Para os amantes de Sudoku, palavras cruzadas e exercícios de matemática aí está uma oportunidade de colocar o lápis para trabalhar: é só elevar todos os desvios da média ao quadrado que, em cerca de uma hora de zelosa labuta, chega-se aos valores da coluna da direita (desvio²). Mas o trabalho não para aí: tem que somar todos esses resultados e ultrapassar os sete bilhões, que podem ser conferidos na penúltima linha da tabela. Aí é só dividir esse número por 27, que é o número de elementos do grupo, no caso os Estados. Finalmente, chega-se à cifra “lotérica” da média dos desvios elevados ao quadrado, cerca de 290 milhões. Pensando num prêmio de loteria, essa média se torna até um número apreciável e todo o cálculo pode ganhar um sentido grandioso nos mais belos sonhos. Para que o sonho perdure, é preciso convencer a razão de que uma chance em mais de 53 milhões de possibilidades ainda é uma possibilidade, bastando acertar aqueles 6 números tão desejados. Se isso for muito dinheiro para apenas um indivíduo, podemos então tirar uma raiz quadrada dessa média e, como num sonho, chegamos ao desvio padrão: raiz quadrada de 290.068.970,53 é igual à 17.031,41. Veja esquematicamente a seguir:

$$\text{desvio padrão} = \sqrt{\frac{7.831.862.204,30}{27}} = \sqrt{290.068.970,53} = \mathbf{17.031,41}$$

Se houver algo errado nesse cálculo, por favor, enderece as reclamações para a Microsoft que é a proprietária do Excel e a responsável por programar as fórmulas que fazem o show instantâneo dos cálculos automáticos de números tão grandes. Todo esse processo é muito discreto e acontece em milésimos de segundos, na caixa preta dos processadores dos computadores.

Aqueles que insistem em duvidar dos procedimentos matemáticos e científicos, podem nesse momento se entregar ao deleite de fazer os cálculos por si mesmo, voltando à Idade Média. Olha só, novamente a média, não tem como fugir dela...

Uma forma de usar o desvio padrão é construir uma escala na qual ele se torna uma espécie de medida. Colocando numa régua, o zero é o ponto da média, onde não há desvio. Aí é possível começar a contar em termos de desvios. Na figura 3, a seguir, é possível perceber onde cada Estado se posiciona nessa escala que pode ser chamada de régua do desvio padrão.

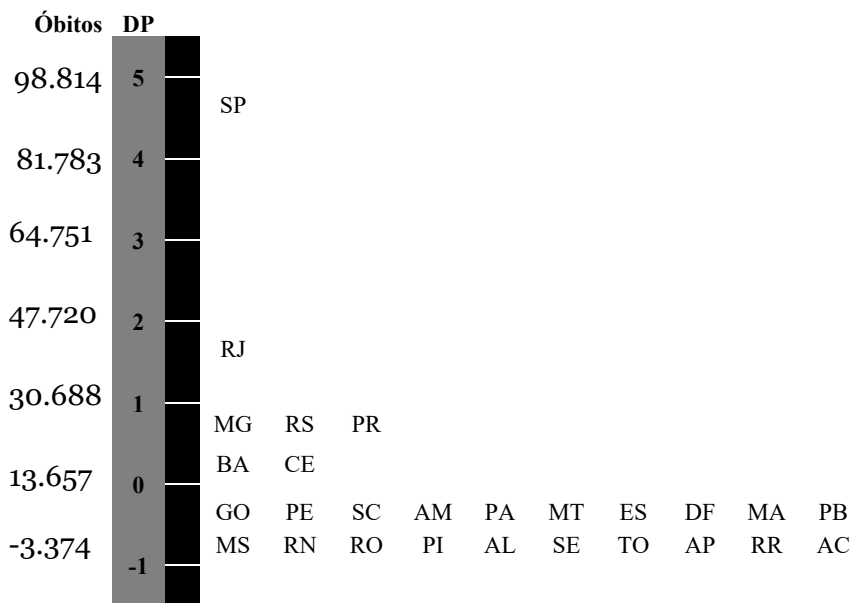


Figura 3: A Régua do desvio padrão dos óbitos por COVID-19 em 16/04/2021.

A régua mostra algumas anomalias nesse grupo de Estados que são pistas importantes para compreender esse conjunto de dados. Primeiramente, observa-se que ao descer menos um desvio padrão abaixo da média de 13.657 mortes, chega-se a um número negativo de -3.374. A operação é a

seguinte: Média menos o desvio padrão, ou seja, $13657 - 17.031,41 = -3.374$. Isso é um absurdo, pois ainda não foi relatado nenhum caso de pessoa que tenha ressuscitado depois do atestado de óbito por COVID-19 dado por algum médico. Assim, não tem nenhum Estado que tenha relatado esse prodígio de mortes negativas. O que se pode afirmar, entretanto, é a existência de 20 Estados que ficaram com o número de óbitos entre a média e um desvio padrão abaixo da média. Passando da média, ficaram 5 Estados entre ela e um desvio padrão acima (Média mais um desvio padrão, ou seja, $13657 + 17.031,41 = 30.688$). O Estado do Rio de Janeiro ficou quase exatamente há dois desvios padrões acima da média. E por fim, São Paulo ultrapassou os quatro desvios padrões acima da média, ficando a meio caminho entre o quarto e o quinto andar ou, em palavras mais formais, entre quatro e cinco desvios padrões acima da média. Essa distribuição de dados está alarmantemente anormal e algum fenômeno está explicando essa distorção. Há uma concentração no primeiro degrau da escala e um caso extremo no alto, parecido com o caso imaginário do gigante e os sete anões. Ao ler a crônica 1, muitos leitores já haviam alertado para isso e mataram a charada. Mas será em outra crônica que o mistério será revelado completamente, pois é necessário avançar didaticamente em alguns conceitos antes de caminhar nessa direção. Nem todos tem sua velocidade, caro e atencioso leitor, então vamos juntos (*Come together*).

O mesmo cálculo para o desvio pode ser feito para a porcentagem também. Assim poderemos comparar duas escalas diferentes com a mesma régua do desvio padrão e analisar o seu comportamento em uma e em outra situação. Na tabela 7, a seguir, é possível observar os mesmos fenômenos da última tabela se aplicando aos dados da porcentagem dos votos por Estado: o somatório dos desvios é igual a zero e elevando-se ao quadrado, os desvios negativos passam a ser positivos o que possibilita ter um somatório diferente de zero.

Tabela 7: Cálculo do desvio padrão da percentagem de votos.

Estado	% votos Jair	Desvio	Desvio ²
SC	61	19	361,00
RR	61	19	361,00
RO	59	17	289,00
AC	59	17	289,00
MS	56	14	196,00
DF	55	13	169,00
RJ	54	12	144,00
PR	53	11	121,00
GO	53	11	121,00
MT	52	10	100,00
ES	50	08	64,00
RS	49	07	49,00
SP	48	06	36,00
MG	43	1,26	1,59
Média	41,74		
AM	41	-0,74	0,55
TO	41	-0,74	0,55
AP	39	-2,74	7,51
PA	34	-7,74	59,91
AL	31	-10,74	115,35
PB	28	-13,74	188,79
RN	28	-13,74	188,79
PE	27	-14,74	217,27
SE	24	-17,74	314,71
MA	23	-18,74	351,19
BA	21	-20,74	430,15
CE	20	-21,74	472,63
PI	17	-24,74	612,07
Total	1127	0,00	5261,03
Desvio Padrão			13,70

Os dados da percentagem de votos se apresentaram de forma mais harmônica em torno da média, quando comparados com o caso dos óbitos. Aqui não houve casos nem tão abaixo da média, nem tão acima. A figura 4, a seguir, faz o mesmo exercício da figura 3, através dos dados da percentagem de votação no Presidente.

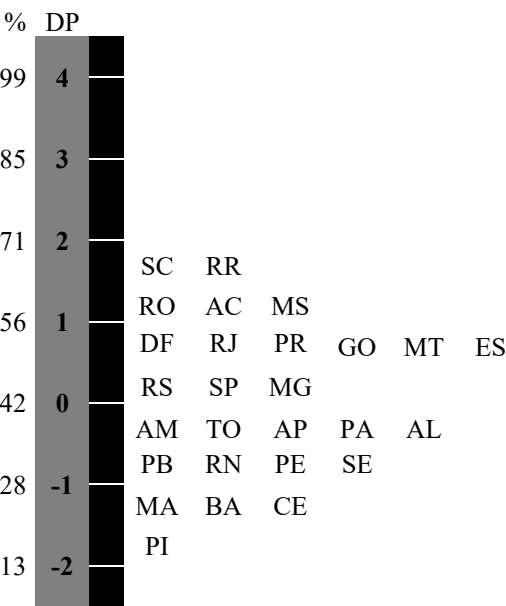


Figura 4: A régua do desvio padrão das percentagens de votos por Estado nas eleições de 2018.

Diferente da régua aplicada aos óbitos, nos votos não houve percentagem negativa para representar a subtração do desvio padrão da média. Será que só os “comunas” gostariam que aqui se encontrassem casos de Estados posicionados em um patamar abaixo de zero e, portanto, negativo? Assim seria possível reverter esses votos do passado e mudar toda história

num passe de mágica matemática, num processo que seria chamado de ressurreição do voto, renascendo para um outro candidato. Mas isso é mero delírio à esquerda.

Na próxima crônica, aprenderemos a visualizar a curva normal que foi ilustrada com essa régua do desvio padrão. Compreender essa curva é um passo preliminar para transformar um número de um conjunto de dados em um valor correspondente em termos de desvios padrões e comparar a um número especial chamado “z”. Cada Estado terá um número correspondente para ser posicionado de maneira mais precisa nessa “régua do desvio”. Até breve.

Crônica 3: A Curva Normal.

Os dados obtidos para uma análise podem ser agrupados em termos da frequência com que surgem dentro de classes de resultados, como foi visto no caso das notas escolares, na crônica 2. Para observar esse agrupamento, surgiram os gráficos de barras que fornecem uma boa visualização da distribuição dos dados. O gráfico 5, a seguir, foi construído com dados fictícios da altura dos homens brasileiros, baseado em dados originais do IBGE, segundo os quais a altura média do homem brasileiro é de um metro e setenta e três centímetros (1,73 metros ou 173 centímetros). De um total estimado de 100 milhões de homens, observa-se que a sexta maior barra se encontra na altura de 170 centímetros(1,70m), correspondendo à 6,8 milhões de indivíduos nessa condição. Para saber a probabilidade da Branca de Neve sortear um desses príncipes encantados de 1,70m, nessa população, basta dividir a quantidade de elementos desse subgrupo pelo número total de indivíduos da população de homens: $6,8 / 100 = 0,068$, ou seja, quase 7 por cento de chance de sortear um deles ao acaso. Se ela for menos exigente e disser que ele tem que ter apenas uma altura igual ou superior a essa, a operação é mais demorada, porém é tão simples quanto. Soma todos os valores das barras superiores a 169 centímetros, que corresponde à quantidade de indivíduos classificados com altura de 170 centímetros ou mais. A conta é essa: $6,8 + 7,3 + 7,5 + 7,6 + 7,5 + 7,3 + 6,8 + 6 + 5 + 4 + 3 + 2 + 1,5 + 1 + 0,6 + 0,4 + 0,3 + 0,2 + 0,1 + 0,08 + 0,06 + 0,04 = 0,7508$. Colocando em palavras, a Branca de Neve tem, aproximadamente, 75% de chance de sortear ao acaso um homem de estatura igual ou superior à 1,70 metros nessa população de brasileiros. Para o lado contrário, se ela desejasse alguém mais baixo do que ela, então teria que fazer

o somatório das colunas que estão à esquerda dessa altura. Pode fazer o exercício para conferir, mas tenho a impressão de que o resultado será de 25%, pois já achamos a percentagem anterior e sabemos que o total é 100%. Então, $100 - 75 = 25$. Essas percentagens podem ser representadas em termos da área correspondente sob a curva que é formada traçando uma linha que passa por cima de cada coluna que representa a frequência.

Como pode ser visto no gráfico 6, a área das barras corresponde a essas percentagens, demonstrando que é possível determinar as regiões do gráfico que correspondem a percentagens específicas. Observa-se também uma simetria ao dividir o gráfico no meio e uma curva pode ser traçada facilmente. Esses gráficos foram feitos de maneira a mostrar um desenho bem próximo de uma curva normal perfeita.

Já no gráfico 7, observa-se duas áreas representadas, uma à esquerda e outra à direita. Elas são iguaizinhas e mostram que é mais difícil encontrar uma pessoa com mais de 182 centímetros e abaixo de 165 centímetros. Apenas 4,3% dos indivíduos se encontram acima da primeira e outros 4,3% se encontram abaixo dessa segunda altura. Não faz parte do escopo desse livro demonstrar os cálculos que permitem chegar à conclusão de que a área sob a curva corresponde à probabilidade de ocorrência de um intervalo de frequência, pois envolve algumas operações mais complexas. Para o nosso objetivo, basta considerar o cálculo da percentagem realizado anteriormente que envolve a soma das frequências do intervalo desejado e divisão pelo número total de ocorrências.

No gráfico 8, encontram-se representadas duas áreas que são muito utilizadas para observar fenômenos estatísticos: uma delas tem a probabilidade de 68,26% e a outra, de 95,44%, correspondentes a intervalos específicos estabelecidos através do desvio padrão. O símbolo “ σ ” significa “desvio padrão”. No caso do primeiro intervalo, tem-se um desvio (σ) para direita e outro para esquerda ($-\sigma$). O segundo intervalo tem “ 2σ ” para cada lado.

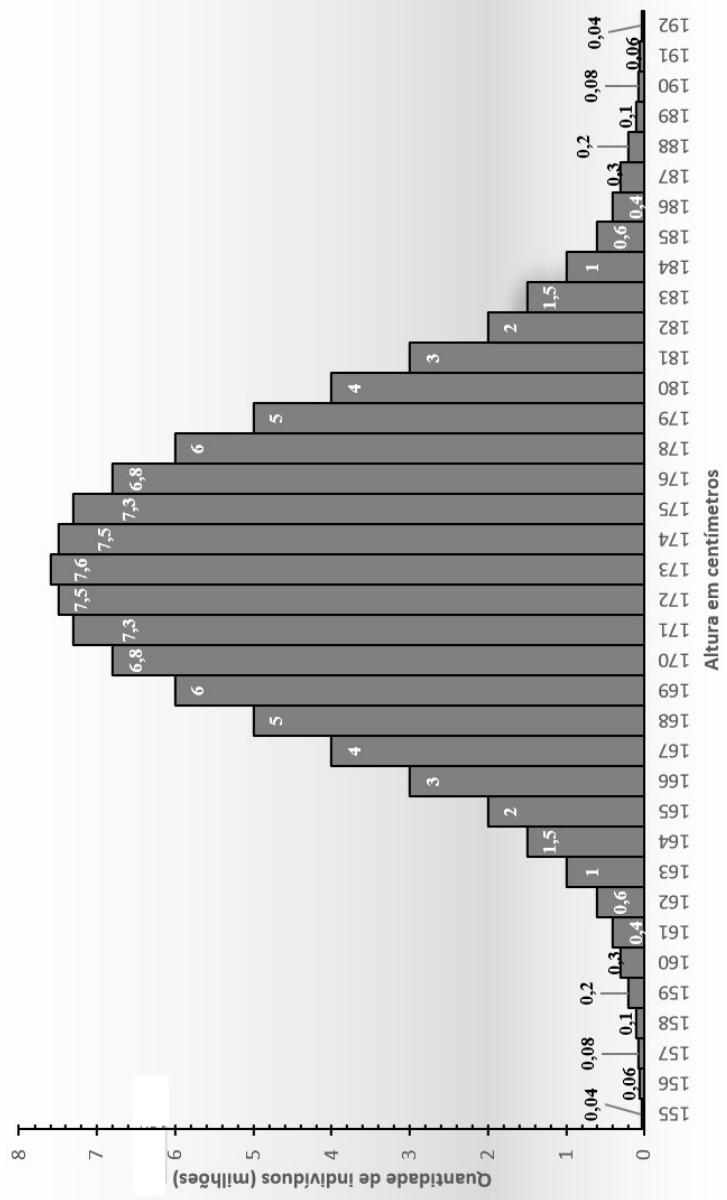


Gráfico 5: histograma estimado da altura do homem brasileiro

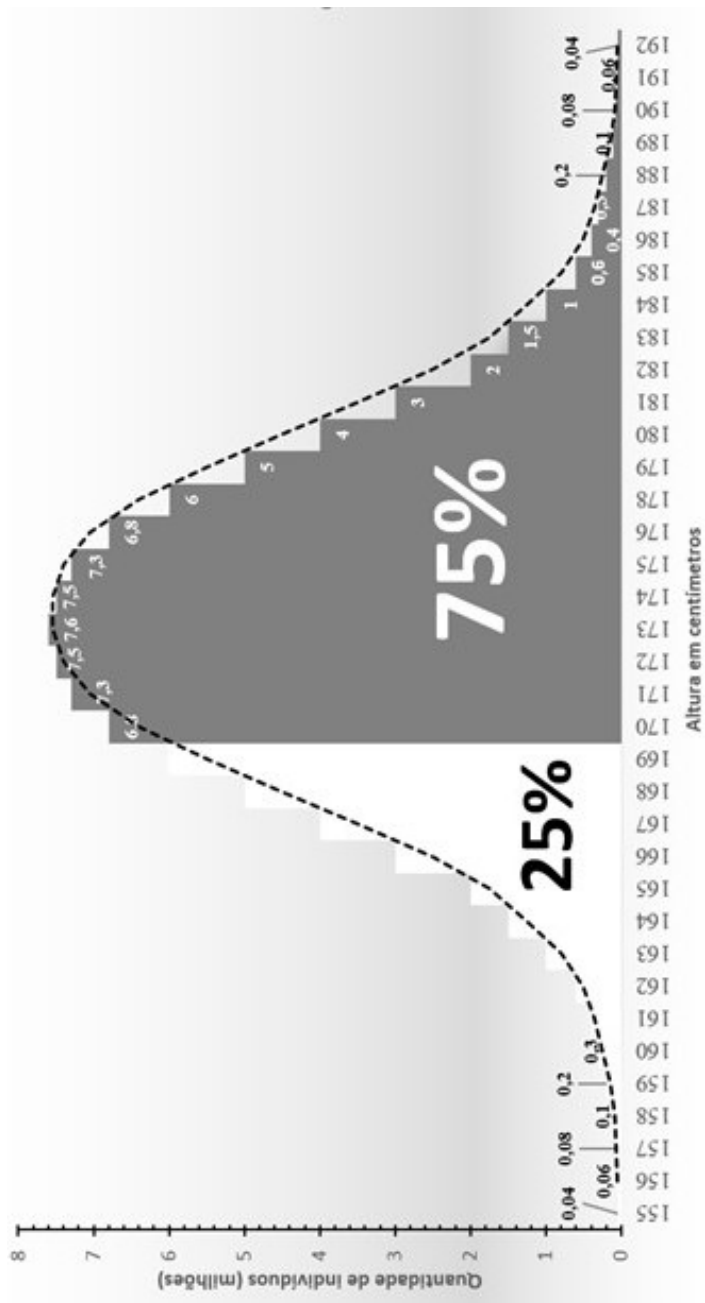


Gráfico 6: Percentual de homens com 1,70 m ou mais

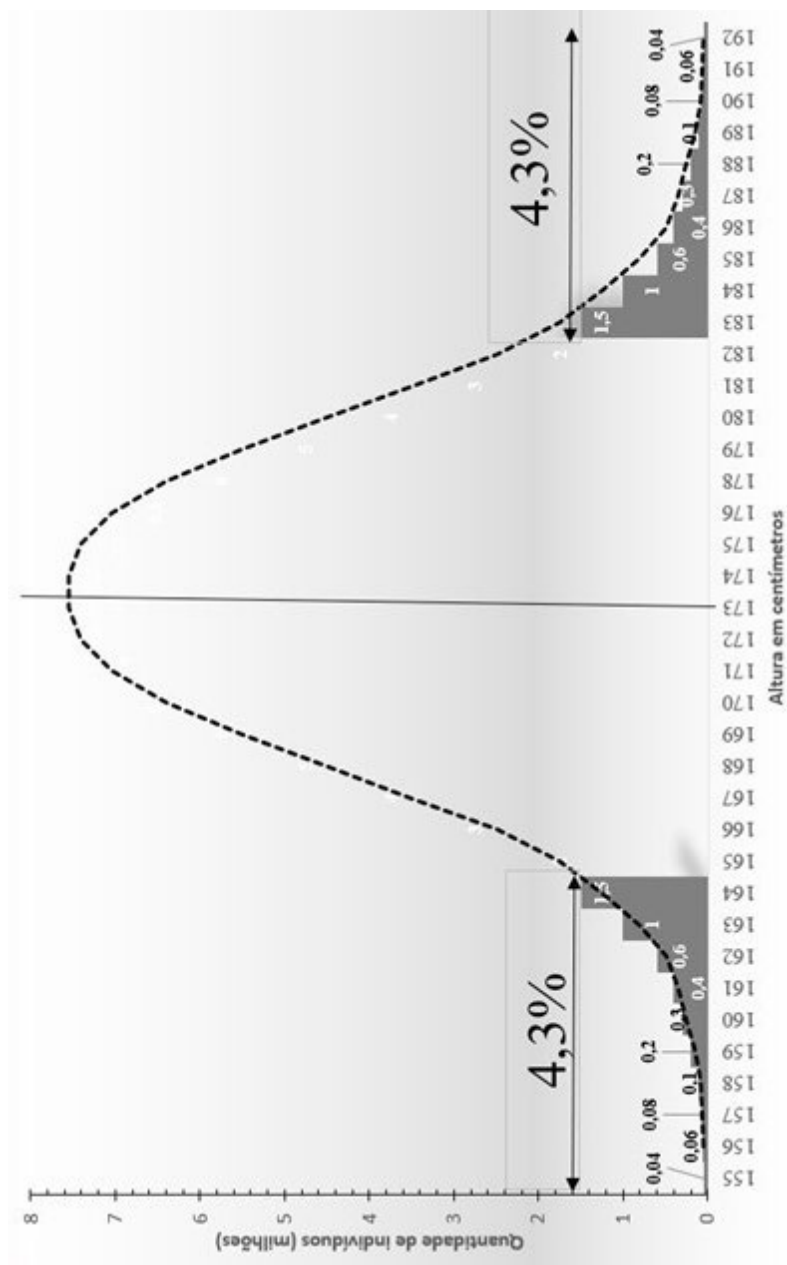


Gráfico 7: Percentagem de homens abaixo de 1,65m ou acima de 1,83m

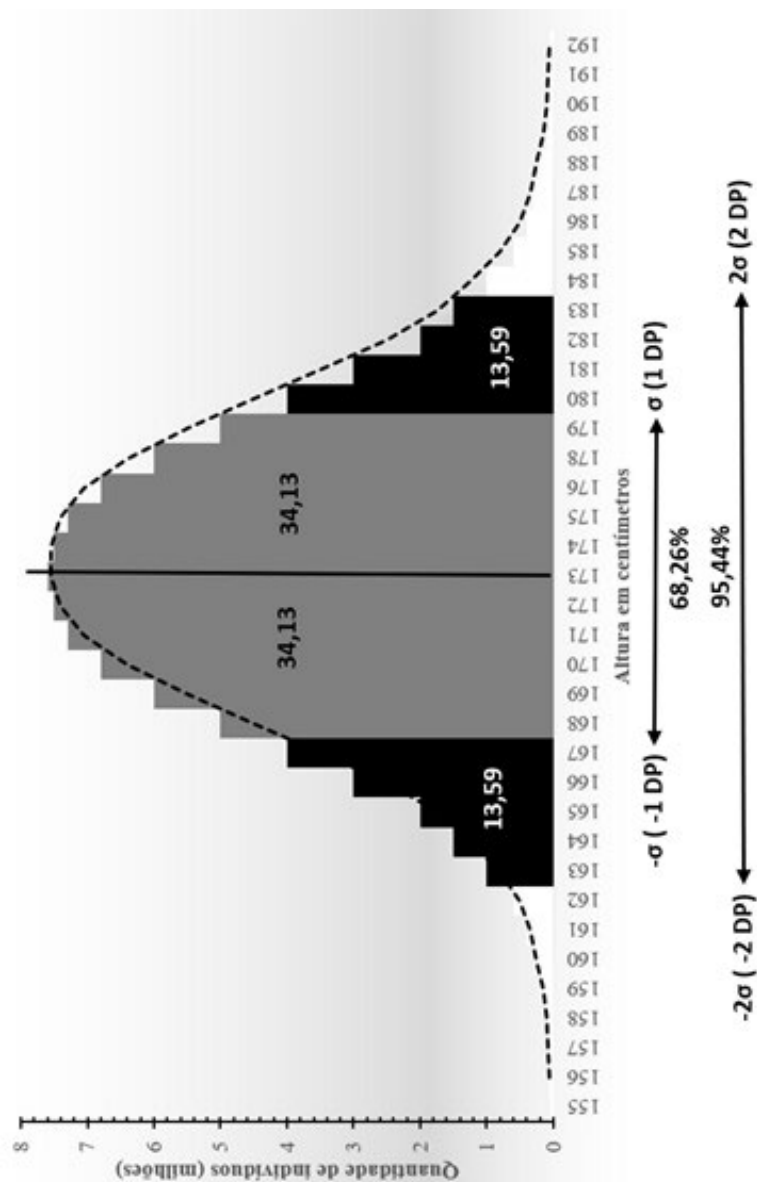


Gráfico 8: Principais áreas sobre a curva normal

Se a Branca de Neve quisesse reunir seus anões teria alguma dificuldade. Eles se encontram, tecnicamente, abaixo de 145 centímetros e nem aparecem nos gráficos. Para fins práticos, podemos considerar “anões” aqueles que possuem altura menor que a última quantidade presente no gráfico, 155 centímetros, lembrando que não se calcula um ponto específico de probabilidade de ocorrência, mas um intervalo para se chegar a uma porcentagem. Então, $0,4/100 = 0,004$. Isso significa que a chance de se sortear um “anão” entre os 100 milhões de homens é de 0,4 por cento, ou seja, para cada 1.000 homens, há 4 “anões”. Fazendo a conta: $1.000/4 = 250$, chega-se a 1 “anão” para cada 250 homens. Para encontrar a probabilidade de sortear os 7 “anões”, basta multiplicar esse valor 7 vezes: $250 \times 7 = 1750$, ou seja, para cada 1750 homens, existem 7 deles abaixo de 156 centímetros de altura. Quanto mais baixa ou mais alta é uma pessoa, mais difícil é encontrá-la, pois ela se perde entre os milhões que estão mais próximos da média.

Fenômenos que se expressam em grandes números revelam a tendência central em torno de uma média. Quanto mais distante da média, mas raro vai se tornando o acontecimento observado. Estudando a distribuição das frequências de muitos fenômenos, uma regularidade se revelou impressionante: quando somamos e diminuimos um desvio padrão da média, se estabelece um intervalo dentro do qual se encontram cerca de 68% das observações de fenômenos que obedecem à distribuição normal. Se somarmos e diminuirmos dois desvios da média, outro intervalo se manifesta com cerca de 95% dos casos de uma determinada variável. Por fim, no terceiro desvio padrão para mais e para menos, o intervalo contém praticamente a totalidade dos dados: 99,9%. Seria necessária uma matemática mais avançada, que não domino, para compreender por que o universo se manifesta dessa forma. Mas podemos nos deliciar em encontrar a curva normal, descrita no gráfico 8, nos mais inusitados fenômenos.

A curva se parece com um sino e esse tornou-se o seu nome popular. “Muita gente” vê sinos e curvas normais para onde quer que olhe, até mesmo num pôr do sol. Quando “essa gente” é estatístico então, resolve-se facilmente o enigma proposto pelo pequeno príncipe de Antoine de Saint-Exupéry (1994). Observe a figura 5 a seguir. O que você vê?

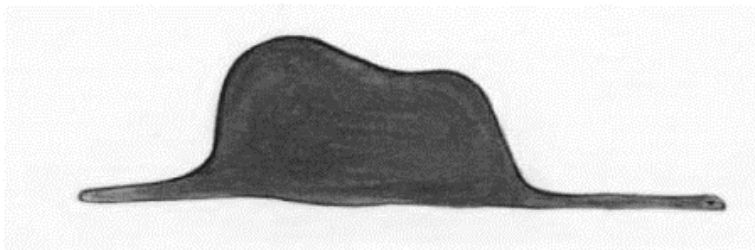


Figura 5: O elefante engolido pela jiboia

Fonte: (DE SAINT-EXUPÉRY, 1994)

O pequeno príncipe apresentou a interpretação alternativa dessa curva normal como sendo um elefante, que teria sido engolido por uma suposta jiboia. Uma imaginação muito criativa e lúdica, mas na verdade se tratava da ingestão muito rara de uma curva normal, que foi engolida acidentalmente pela jiboia apreciadora da estatística. Um pouco deformada pelas contorções estomacais, do difícil processo digestivo da matemática, é claro! Mas, ainda sim, uma curva quase simétrica em forma de um sino. Relatos mitológicos afirmam que um sino caiu de uma torre bastante alta, numa igreja abandonada na floresta amazônica, sendo depois consumido pela diligente cobra com tendências estatísticas. Para tentar confundir o leitor, na figura 6, o elefante se apresenta de maneira convincente, mas depois de treinamento estatístico, é mais difícil embarcar em *fake news*! Antigamente, muitos confundiam a figura com um chapéu, entretanto, depois de meditar no fenômeno da curva normal, não restam mais dúvidas.

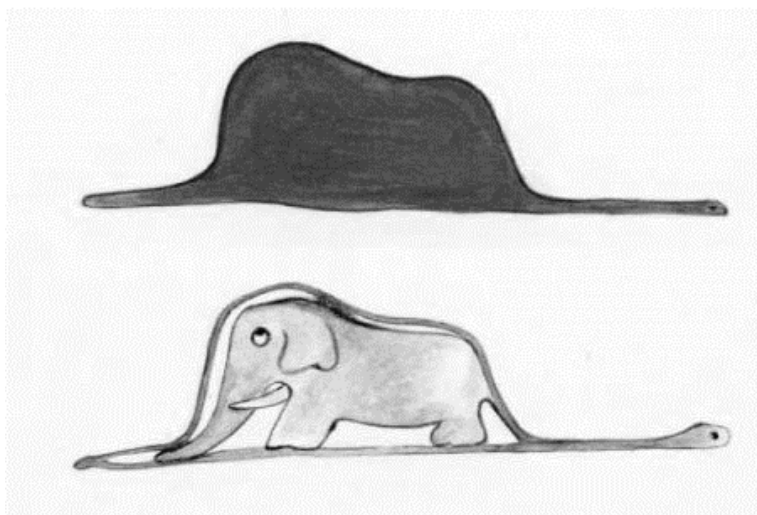


Figura 6: Interpretação forçada do elegante

Fonte: (DE SAINT-EXUPÉRY, 1994)

Essa figura 6 foi uma interpretação bastante forçada do pequeno príncipe, pois existem muitas alternativas para compreender esse fenômeno natural da ingestão espontânea da estatística.

Vamos mudar o foco e olhar para paisagens, ampliando a perspectiva visual. Na figura 7, a seguir, um belíssimo pôr do sol em preto e branco revela a concentração dos raios no centro, que representa a média. Na medida em que dirigimos o olhar para as extremidades, a luz diminui gradativamente. Uma autêntica curva em forma de sino.



Figura 7: Pôr do sol na Gamboa, Salvador-BA

Fonte: fotografia do autor

Para os incautos, fui obrigado a traçar a curva normal no pôr do sol na figura 8, a seguir, profanando uma ideia abstrata com um traço grosseiro da materialidade.



Figura 8: Pôr do sol na Gamboa com curva normal, Salvador-BA

Fonte: fotografia do autor

Muitas vezes, a didática do professor revela uma realidade empobrecida pelas linhas grosseiras do conhecimento que vara a paisagem para demonstrar teoremas e convencer a audiência. Talvez sejam esses cortes na beleza natural que tanto faz Walt Witman ficar enjoado e o compele a sair do auditório para evitar os gráficos. Procurando vencer essa aversão do poeta pela linha concreta, o professor solicita alguma liberdade poética para expressar suas ideias de maneira espontânea e assim conquistar o coração do aluno para as coisas estatísticas.

Continuando nosso passeio pela natureza, a formação de uma pedra também pode acompanhar a forma de um sino. A Pedra Azul localizada em Domingos Martins, no Espírito Santo, é um exemplo clássico de montanha de pedra que escolheu para si a forma de uma curva normal. Na figura 9, pode conferir com seus próprios olhos:



Figura 9: A curva normal em forma de sino se manifestando na Pedra Azul, em Domingos Martins-ES

Fonte: fotografia do autor

Voltando do nosso voo xamânico pelos delírios da mente fértil do escritor, é possível visualizar o fenômeno da tendência para a média quando observamos o desgaste que acontece pela utilização de um piso ou terreno. Por exemplo, o desgaste que forma uma trilha no meio de uma mata. Observa-se uma tendência de as pessoas trafegarem, a pé ou de bicicleta, pelo meio de uma via de terra, gerando o caminho bem delimitado no meio, que corresponde à média. Poucas pessoas caminham pelas beiradas e o terreno ali não se desgasta tanto e a vegetação se recompõe mais facilmente. Portanto, na figura 10, mais uma vez é possível visualizar esse fenômeno da tendência ao centro.



Figura 10: A trilha e a tendência a trafegar pelo centro

Fonte: Arquivos do autor

Em estradas, os carros em média trafegam dentro de uma tendência. Ali, onde os pneus mais passam, cria-se uma deformação. Mas eles não passam exatamente no mesmo lugar, mas dentro de uma faixa que varia num limite de segurança, para

não irem nem demasiadamente para esquerda e provocar um acidente, ou demasiadamente para a direita e sair pelo acostamento da estrada. Avançando gradativamente para fora da deformação, o asfalto se desgasta menos. Na figura 11, observamos uma via de asfalto na qual a parte mais clara está mais desgastada, pois os quatro pneus passam predominantemente dentro de um espaço que mantém o automóvel na mão adequada para evitar o choque com os carros que trafegam em sentido contrário. A faixa de asfalto que está mais escura é por onde raramente o pneu de um carro passa. Indo para as extremidades, o tráfego é menos provável, o que leva a menos desgastes da pista.



Figura 11: Estrada e o desgaste do tráfego

Fonte: Arquivos do autor

Cada passagem de um pneu é uma excelente analogia para a ideia de um experimento no qual observamos algum fenômeno. Geralmente não é possível observar todos os pneus que vão passar pela pista, assim como é impossível observar, de maneira onipresente, do início ao fim, qualquer que seja a variável que um cientista esteja estudando. Uma forma mais prática de observar é retirar uma amostra reduzida de uma

população para fins de análise e produção de informações úteis. No exemplo da altura do brasileiro, é um tanto quanto inviável medir a altura de todos os indivíduos da população. Mas uma amostra mais razoável de cinco mil pessoas é uma tarefa executável com algum empenho. Os estatísticos descobriram uma informação bastante útil: na medida em que mais amostras são extraídas de uma população, a média do conjunto delas tende a expressar a média real. Para não repetir infinitamente o processo de amostragem, eles sistematizaram o Teorema Central do Limite. Assim, a média das médias das amostras tende a corresponder à média da população. Como ninguém em sã consciência brinca de ficar fazendo amostras, uma dedução lógica do Teorema é que quanto maior for a amostra, maior é a probabilidade de encontrar uma média mais próxima da média da população. Foi assim que Silva e Monteiro (2009) encontraram, numa amostra de 5305 homens entre 15 e 87 anos de idade, nos vinte principais aeroportos do país, a média de estatura de 173,1 centímetros de altura (ou 1,73m). Um dos segredos para cravar essa média igual à da população de homens, em torno de 100 milhões de indivíduos, é construir uma boa amostra. Para isso, é necessário evitar escolher as pessoas por alguma preferência e utilizar algum critério ao acaso. Por exemplo: medir as pessoas durante quatro horas de embarques, em aeroportos nas capitais dos Estados brasileiros, até atingir um número razoável de medições.

Por que alguém se daria ao trabalho de medir tais pessoas, hein Silva? Esse estudo buscou conhecer as medidas dos brasileiros que viajam de avião e assim determinar um tamanho adequado de poltrona para atender ao maior número possível de passageiros de uma maneira confortável. Para aqueles que são muito altos, poderia ter uma ou duas poltronas especiais, pois, dificilmente, embarcam mais de dois grandalhões de 2 metros de altura por voo, a não ser que seja um time de basquete indo para algum torneio em outro Estado. Nos anexos desse trabalho,

encontra-se o esboço de uma curva normal da amostra. O gráfico 9, a seguir, é uma reprodução de um gráfico retirado de lá (SILVA; MONTEIRO, 2009). Percebam que ele usou intervalos de 10 centímetros que levou a uma grande concentração dentro das classes de altura (160-170; 170-180; 180-190). Se houvesse um pouco mais de esforço, construindo classes com intervalos menores de altura, o resultado seria mais elegante e as barras acompanhariam graciosamente a curva guia.

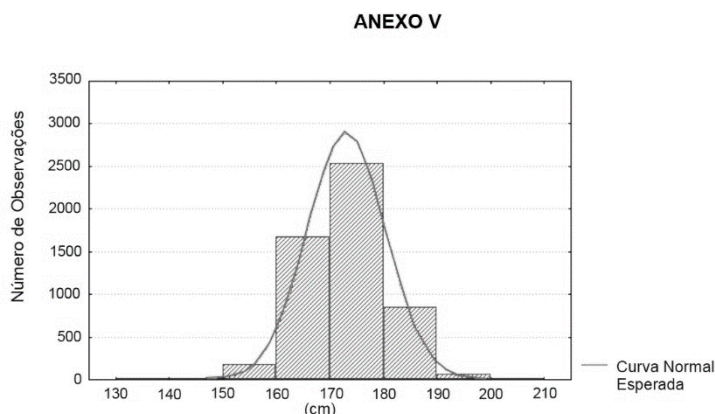


Gráfico 9: Distribuição dos Dados de Estatura dentro da Curva de Normalidade Esperada

Fonte: (SILVA; MONTEIRO, 2009)

Caminhando para as extremidades, as estaturas vão se tornando raras. Pode-se dizer que a quase totalidade se encontra entre 160 e 190 centímetros. O menor sujeito encontrado tinha 144 centímetros e o mais alto 200 centímetros. O desvio padrão foi de 7,3 centímetros. Aplicando aquela ideia de somar e diminuir o desvio padrão da média, encontramos os seguintes intervalos no gráfico 10:

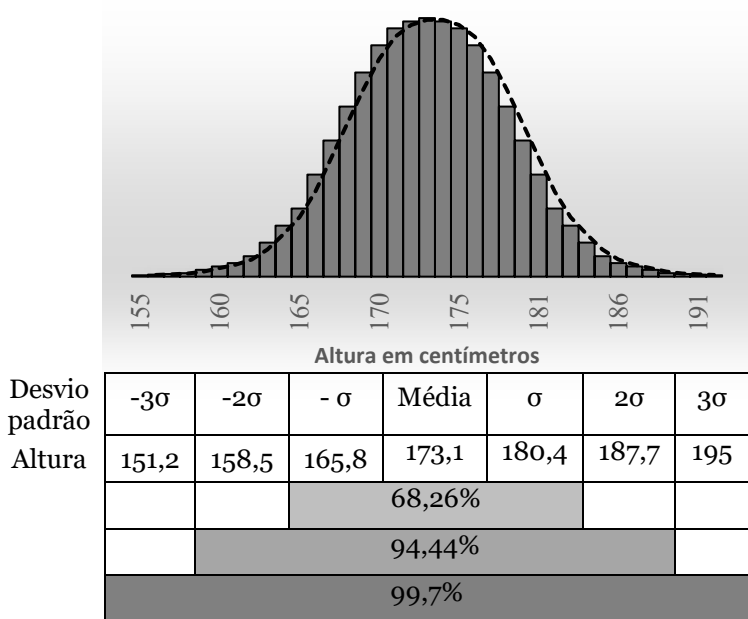


Gráfico 10: Porcentagem de homens de acordo com o intervalo de 1, 2 e 3 desvios padrões em relação à média.

Observando o gráfico 10, podemos ter uma ideia aproximada de que a maior parte dos homens terá entre 165 e 181 centímetros de altura. Assim, é possível projetar assentos mais adequados à população brasileira.

Vamos concluir essa crônica por aqui, pois fiquei muito contente com a ausência de temas que nos trazem sentimentos de tristeza e luto. Por um instante contemplamos o pôr do sol e revisitamos o pequeno príncipe. Um refresco para se preparar para o que há de vir na crônica 4 que trata do número “z”.

Crônica 4: O Número “z”

É possível usar o desvio padrão para observar se os dados estão bem agrupados em torno da média ou se estão mais dispersos, estabelecendo assim um critério de comparação de diferentes populações através de um gráfico. Quando o desvio padrão é grande, o gráfico resultante das frequências observadas tende a ficar bastante achatado e assim compromete “a forma de sino”. Se o desvio padrão for pequeno, o “sino” se afina em torno da média. A curtose é justamente essa avaliação do grau de achatamento da curva: se a curva resultante de uma distribuição de dados possui as caldas grossas, ela é denominada de distribuição platicúrtica, indicando uma dispersão maior dos dados em relação à média. Do contrário, quando a variação dos dados em relação à média é muito pequena, o gráfico resultante da distribuição da frequência dos dados agrupados em classes ou intervalos tende a se afinar ou tornar-se pontiagudo. A essa deformação se dá o nome de distribuições leptocúrticas. As duas formas principais de curtose estão representadas nos gráficos 11 e 12, onde se alucina a distribuição de 100 observações de um fenômeno qualquer, em cada um dos dois diferentes países de maravilhas frugais, o país dos Abacates e o outro dos Abacaxis. Os dados variam de 0 a 100 em alguma medida indefinida ou delirante. As observações foram agrupadas de 10 em 10 “libras de imaginação poética” e foram contadas as ocorrências fictícias dentro de cada classe de observação fantástica do nada. A aberração resultante não deixa nada a desejar aos famosos gráficos observados por Walt no século XIX que tanto lhe causaram náuseas numéricas e o fez contemplar o céu estrelado como remédio para a aversão matemática do poeta.

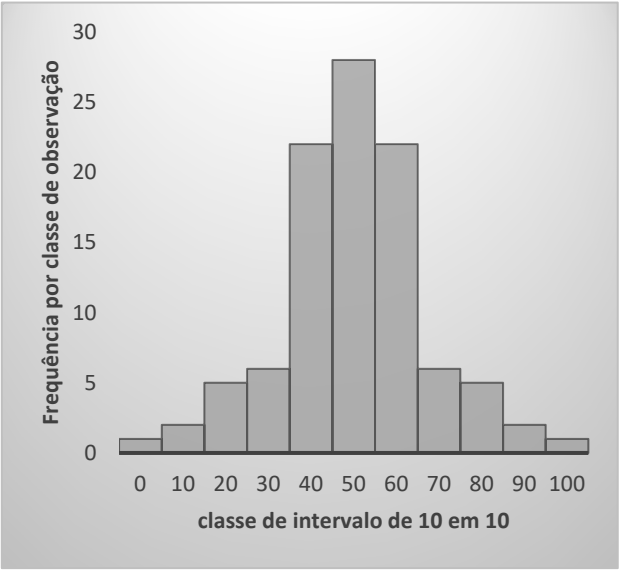


Gráfico 11: Distribuição Leptocúrtica no país dos Abacates
Fonte: Altos montes da imaginação poética escondidos no Excel

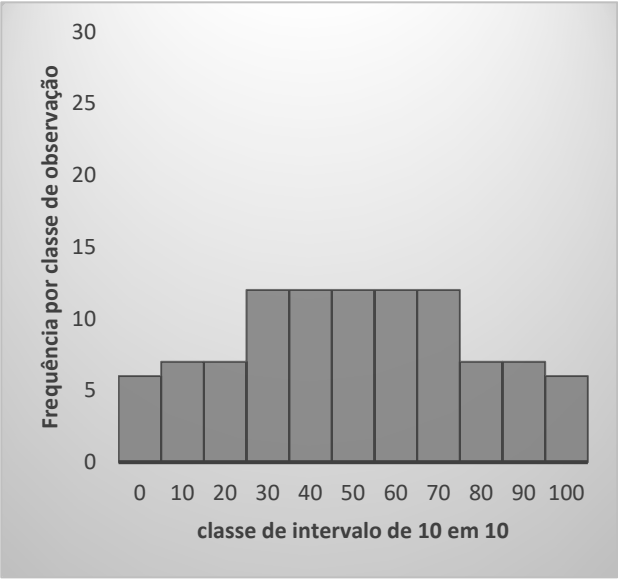


Gráfico 12: Distribuição Platocúrtica no país dos Abacaxis
Fonte: Platô fantástico delirado no Excel

Além de esticar para cima ou espremer violentamente para o chão, o sino que habita a curva normal pode ser deslocado em seu eixo para a direita ou esquerda. Esse tipo de deformação de uma distribuição de dados em relação a uma distribuição normal, denomina-se assimetria. A curva normal obedece cegamente à rígida regra da normalidade obsessiva da estatística das coisas. Entretanto, muitas vezes na sociedade humana, os fenômenos se apresentam como verdadeiros danados na terra e se recusam a vestir a camisa de força em forma de sino.

É por conta dessa revolta dos dados que as pesquisas dão à luz a rebentos tortos e disformes, de difícil compreensão. Mesmo assim, cientistas persistentes, amorosamente e politicamente engajados, acolhem generosamente aquilo que no contexto popular seria percebido como um simples aborto. Sob a bandeira da inclusão incondicional, adotam suas crias prematuras, encarando a sua frente de difícil decifração através dos vitrais de acrílico de alguma incubadora experimental. Em algum momento de desilusão, eles passam a mão por cima de seus cabelos sedosos e os chamam de seres singulares que demandam pesquisas posteriores para serem incluídos no rol dos seres razoáveis. Na medida em que os colegas replicam os estudos, na esperança de consertar o erro de nascimento através da produção de clones melhorados, o cidadão comum estica a cabeça sob o muro da maternidade acadêmica para vislumbrar algum motivo ou explicação justa para trazer ao mundo formas distorcidas e imperfeitas de contagens que desobedecem à lei da curva normal. Enfim, o sino perfeito raramente é encontrado na maior parte das pesquisas, principalmente com a escassez de recursos para estudar grandes levantamentos de dados, sendo a situação mais grave nas ciências humanas e sociais. Portanto, se o filho da pesquisa é feio, em parte isso se deve à política governamental que abandona a mãe ciência ao azar da desnutrição.

Para explicar uma assimetria positiva da curva normal, uma opção é recorrer a imagens convincentes, principalmente quando retiradas de uma obra consagrada mundialmente, com um estilo didático singular e que se tornou um livro de cabeceira de cientistas psicólogos que gostam de números (FIELD, 2009). Nele é possível encontrar muitos exemplos de dados e pesquisas na área da psicologia, ensinando a tratar os números com o carinho que eles merecem através de um programa de computador chamado SPSS (*Statistical Package for the Social Sciences*). Assim, o livro didático “Descobrimos a Estatística com o uso do SPSS” vem se tornando popular na disseminação da estatística em psicologia.

Para ilustrar distorções na curva normal, o autor Andy Field (2009) propõe o seguinte gráfico sobre os suicídios em Beachy Head, que é um rochedo grande e ventoso na costa de Sussex, Inglaterra, próximo à residência dele. No gráfico 13, se observa uma inclinação da curva para a esquerda de quem vê, distorcendo uma possível idealização de um sino perfeitamente simétrico. Quando o cume do sino fica à esquerda da média, denomina-se de assimetria positiva. Ao contrário, o fato do gráfico ser deslocado para direita da média indica uma assimetria negativa. Então, o gráfico a seguir possui uma assimetria positiva.

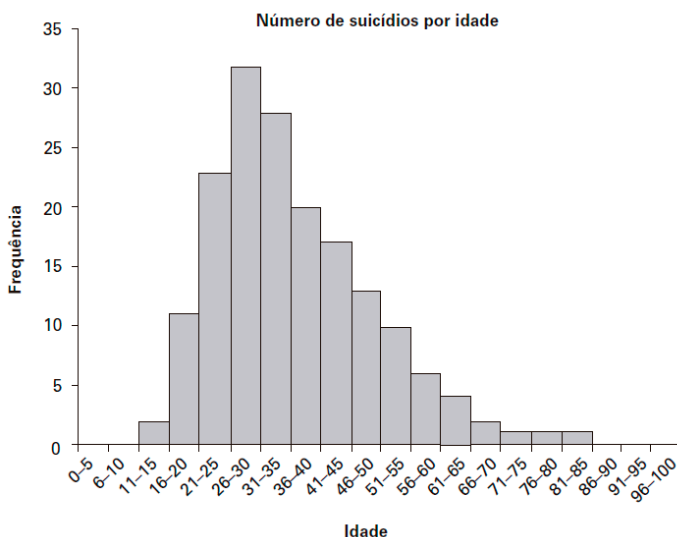


Gráfico 13: Distribuição de frequências mostrando o número de suicídios por faixa etária.

Fonte: Field (2009, p. 40)

De acordo com Field(2009), foram 171 suicídios ao todo. O pico de suicídios se dá entre os 25 e 30 anos nesse rochedo de Sussex. Na medida em que caminhamos para a direita, no aumento da idade, a quantidade de casos diminui, talvez devido à tendência das pessoas mais velhas caminharem com mais prudência, longe da beira do penhasco. Ele não resiste e desabafa pelo fato de se encontrar nessa faixa etária:

...o que é bastante preocupante para mim (embora isso signifique que se eu conseguir passar por este ano a vontade irá diminuir no ano que vem!). O gráfico também nos diz que, por exemplo, poucas pessoas acima de 70 anos cometeram suicídio em Beachy Head. (p. 40)

Será mesmo que ele pode ficar esperançoso ao passar pelos 30 anos? Com os dados apresentados ele tem motivos para se alegrar? Em seguida ele faz a seguinte pergunta:

Afirmei anteriormente que podíamos pensar nas distribuições de frequência em termos de distribuição de probabilidade. Para explicar isso, imagine que alguém pergunte a você o seguinte: **“qual é a probabilidade de uma pessoa de 70 anos cometer suicídio em Beachy Head?”**. Qual seria a sua resposta? As chances são, se você olhar a distribuição de frequências, de sua resposta ser “pouco provável”, pois há apenas duas pessoas do grupo de 171 suicidas com idade em torno de 70 anos. E se alguém perguntar “qual é a probabilidade de uma pessoa de 30 anos cometer suicídio?”. De novo, olhando o gráfico, você poderá dizer “é bastante provável”, pois 32 dos 171 suicidas eram pessoas com idade em torno de 30 anos (isso é mais do que uma em cada cinco pessoas que cometeram suicídio). (p. 40, grifo nosso)

E para você, de acordo com o gráfico “qual é a probabilidade de uma pessoa de 70 anos ou mais cometer suicídio em Beachy Head?” Vamos lá, hora de exercitar o raciocínio probabilístico.

A ideia da probabilidade é pensar se um evento não irá ocorrer ou se há certeza de que vai acontecer. No primeiro caso, costuma-se usar o “0” (zero) para representar a ausência de alguma chance de algo acontecer. No segundo, usa-se o “1” (um) para indicar a certeza total de que algo irá acontecer. Quando alguém que jogou na loteria diz que há 0% de chance de tirar na loteria no dia seguinte, a pessoa está ignorando o fato de que há uma chance em mais de 50 milhões de possibilidades, portanto, não é exatamente nenhuma chance ou 0%, mas algo próximo disso 0,000000019%. Por outro lado, dizer que é certo que o sol nascerá amanhã, é o mesmo que afirmar com 100% de certeza, ou seja, é certo que irá acontecer. Isso ignora a presença de buracos negros no universo que podem milagrosamente engolir o sistema solar num piscar de olhos ou do dia para a noite.

Apesar de boa parte da população viver em busca de certezas, que nem sempre são tão certas assim, no dia a dia, pessoas mais cautelosas procuram estimar probabilidades mais razoáveis para guiar suas decisões entre 0 e 1, ou seja, entre 0% e 100% de chances de algum fenômeno ocorrer. Dificilmente, na ciência, se estabelece alguma “certeza absoluta” baseada em probabilidades, por conta dessa cautela.

No caso da pergunta anterior, há uma tentação de pegar o número de pessoas acima de 70 anos que cometeram suicídio e dividir pelo total de suicídios. Então se descobre que 2 dividido por 171 dá uma probabilidade bem pequena de suicídio nessa idade, no rochedo. Para ser mais específico, a chance é de 0,0117, ou seja, menos de 2%. Mas aqui está o problema: a pergunta não pode ser respondida com os dados fornecidos. Essa resposta caberia para outra pergunta: “Qual a probabilidade de os suicidas no rochedo especificado ter idade maior do que 70 anos?”. Para responder à pergunta inicial, seria necessário saber qual é a população de pessoas maiores de 70 anos que frequentam a região de Sussex e então fazer o cálculo. Se por exemplo, existirem apenas 2 pessoas acima de 70 anos na população exposta ao rochedo, isso indica uma realidade trágica na qual todos acima dessa idade cometeram suicídio nesse lugar. Essa informação levaria à uma política pública que impedisse idosos de acessarem o local devido ao risco observado no passado que poderia indicar uma tendência de repetição no futuro. Por outro lado, se fossem mais de 10 mil pessoas acima dos 70 anos na população relacionada com o problema, a informação torna-se completamente diferente e não justificaria uma atenção especial para evitar idosos na região ou impedi-los de apreciar a paisagem, pois o risco observado no passado teria sido muito reduzido para essa faixa etária. Por isso, na quarta edição do livro, o autor corrigiu a pergunta: “Quão provável é uma pessoa que cometeu suicídio no rochedo da praia ter mais de 70 anos?” (FIELD, 2013). Isso me alegra enormemente, pois este livro, que

te conduz até a antessala do paraíso artificial das estatísticas, também deve conter uma série de erros que poderão ser corrigidos nas edições futuras.

A probabilidade funciona como uma estimativa do que poderá acontecer no futuro e é uma ferramenta útil para planejamento, porém é limitada na produção da certeza de algum acontecimento. No exemplo dado, o padrão de suicídio do passado não é garantia de repetição no futuro. A estimativa é uma esperança de que o passado seja repetido no presente e avance pelo futuro mantendo uma tendência de longo prazo. Mas a esperança não é uma definição absoluta do que virá a ocorrer. E se algum “milagre” se manifestar impactando nos dados futuros ou simplesmente as causas do fenômeno se modificarem com o tempo? Tais “acidentes” ou coisas não previstas surpreendem, de vez em quando, e ignoram as tendências de longo prazo ou as probabilidades mais bem calculadas.

Muita informação estatística veiculada na mídia leva a população a tirar conclusões precipitadas quando calcula probabilidades de maneira descuidada ou distorcida, levando a própria razão a cometer o suicídio antes mesmo que o recipiente dela possa fazer as sinapses corretas. Ainda assim, números bem arrumados indicam fenômenos prováveis e potencializam a intuição sobre os fatos do cotidiano.

As explicações fazem mais sentido quando escoltadas por algarismos bem-educados e apresentáveis. No caso dos dados de Sussex, apesar de terem sido adaptados pelo autor, eles foram inspirados em estatísticas gerais de suicídio da região e indicam um problema observado através dos séculos: jovens tendem a passar por uma fase de risco onde “acidentes mortais” acontecem com mais frequência do que em outras idades. O romance de Goethe ficou célebre por retratar essa fase da vida através dos sofrimentos do jovem Werther (GOETHE, 2010), que depois de publicado ainda gerou uma “epidemia” de suicídios.

Olhando novamente para o gráfico 13, se vê que a maior barra está fixada no intervalo que vai de 26 a 30 anos. Tem 32 pessoas com idade dentro desse intervalo. Ao dividir 32 por 171 se encontra aproximadamente 0,19. Esse número significa que quase vinte por cento ($0,19 \times 100 = 19$, aproximadamente 20%) dos casos de suicídio no rochedo em Sussex é cometido por pessoas de 26 a 30 anos, ou seja, praticamente 1 em cada 5 pessoas está nessa faixa etária (20 em 100 é o mesmo que 1 em 5, pois $100/20 = 5$). As barras, então, podem ser convertidas em probabilidades de ocorrência e fornecer informações que confirmam a intuição da observação direta do gráfico.

Em cada gráfico de cada distribuição seria necessário fazer um cálculo para ter essa informação de probabilidade, pois as medidas variam entre populações distintas e entre instrumentos com escalas diferentes. Para facilitar a leitura dos dados de maneira a amenizar os efeitos dessas diferenças, os estatísticos construíram uma tabela de probabilidades baseada numa distribuição perfeita, portanto, abstrata. Essa tabela (em anexo) auxilia a encontrar uma probabilidade traduzida em termos de desvios padrões. Observe, na parte de cima dela, que a média do gráfico de ilustração é “o” (zero). Há um intervalo qualquer representado entre “o” e “z”, cuja área está com a cor cinza. É possível calcular a probabilidade de algum número cair nesse intervalo, numa contagem que se organiza de maneira semelhante a uma curva normal. Por exemplo, quando “z” é igual a “1”, é possível localizar o valor da probabilidade correspondente. No destaque feito na tabela, encontra-se o número 0,3413 correspondente ao $z = 1$. Isso representa 34,13% da área total da curva, ou seja, se a área total do sino representa 100% ou a totalidade das probabilidades, 34,13% desse sino, ou dessa área, equivalem a esse intervalo.

Compare a tabela em anexo com o gráfico 10 anterior, colocando no lugar do “o” a média de altura. A ideia é a mesma, sendo que é necessário comparar a distribuição ideal do desvio

padrão, onde se encontra o número “z”, com as frequências de altura. Na crônica 3, foi demonstrado, em linhas gerais, esse raciocínio com as alturas de cerca de 5000 homens brasileiros, cuja média era de 173,1 centímetros. A ideia é reduzir a média de qualquer distribuição que se assemelhe à uma curva normal a “o” (zero) e então poder calcular probabilidades específicas com a seguinte fórmula:

$$z = \frac{x - \bar{x}}{\sigma}$$

Onde “x” é o valor da variável de algum caso escolhido, “ $\bar{x}$ ” é a média da variável e “ σ ” é o desvio padrão.

Resgatando os dados da altura média, é possível reduzir a média de 173,1 centímetros para o centro correspondente ao “o” e depois encontrar a probabilidade de encontrar um homem nessa amostra com altura entre um desvio acima e um abaixo da média. Para isso, encontramos os valores correspondentes: com desvio padrão de 7,3 metros, diminuimos e somamos esse valor da média e encontramos 165,8 e 180,4 metros, respectivamente. Substituindo esses valores pelo “x” na fórmula:

$$z_1 = \frac{165,8 - 173,1}{7,3} = -1$$

$$z_2 = \frac{180,4 - 173,1}{7,3} = 1$$

Já descobrimos que a probabilidade do intervalo que vai de “o” até “1” é de 34,13%. Como a curva normal é simétrica, o intervalo de “-1” até “o” terá a mesma porcentagem. Se nossa área vai de “-1” até “1”, basta somar as duas metades e encontrar 68,26%. Esse é o valor que se encontra no gráfico 10

representando a probabilidade de encontrar algum caso no intervalo entre um desvio padrão para mais e para menos da média.

Outra forma de usar a tabela é tentar encontrar a probabilidade de algum evento raro acontecer. No exemplo do suicídio aqui apresentado, é possível calcular a probabilidade de alguém acima de uma determinada idade estar entre os casos de morte. Escolhendo a idade de 70 anos e supondo que a idade média de pessoas que cometeram o suicídio é de 36 com desvio padrão de 13 anos. Para encontrar o número “z” correspondente basta substituir esses números na fórmula: $z = (70-36)/13 = 2,62$. Procurando na tabela a probabilidade correspondente a esse valor de “z” = 2,62, encontra-se 0,4956. Observe que na primeira coluna está apenas o número “z” com uma casa decimal (2,6). A outra parte está na terceira coluna (0,02). Fazendo o cruzamento da linha com a coluna é que se chega à probabilidade de 49,56%. Mas esse valor corresponde ao intervalo que vai de “0” até “2,62”. O que interessa é o intervalo que vai de “2,62” até o fim do gráfico, que representa as pessoas com mais de 70 anos. Para chegar nessa probabilidade, basta lembrar que a metade do sino compreende 50 por cento. Se tirar dos 50% a probabilidade encontrada, o resultado indicará a probabilidade da área restante que representa a resposta correta: $50 - 49,56 = 0,44$. Enfim, menos de 1% das mortes foram de pessoas com mais de 70 anos, ou mais precisamente, 0,44%.

Esses dados foram utilizados para uma compreensão didática da mecânica das probabilidades. Mas observar fenômenos e calcular suas probabilidades envolve uma complexidade maior. Não é tão simples contar os suicídios na vida real. Quantos dos suicídios poderiam ter sido computados como acidentes depois de uma análise mais profunda? A pessoa efetivamente escorregou ou se jogou? Por outra via, quantos “acidentes” poderiam ser classificados como suicídios? E se algum outro fator estaria explicando esses acidentes/ suicídios?

Por exemplo, se uma pessoa estiver sob domínio de alguma droga, é possível atribuir à substância a responsabilidade pelo ocorrido? Será que a pessoa contabilizou o risco na hora de ingerir um alucinógeno sem supervisão adequada? Muitas explicações alternativas poderiam surgir para explicar o fenômeno das mortes no rochedo em Sussex. Imaginem que um pesquisador descobrisse que um cogumelo local produz uma alucinação muito forte em quem inala suas partículas que são lançadas ao ar no final da primavera. Isso poderia explicar o fato de algumas pessoas se atirarem no precipício pensando estar pedalando na bicicleta de Albert Hoffmann. Voltando à idade média, poderia ser criada uma explicação de que o próprio diabo habita a região, hipnotiza os visitantes e rouba-lhes a alma. A mais delirante explicação é pensar em contatos alienígenas que estariam produzindo esse fenômeno do suicídio aparente. Mas antes de concluir essa enumeração de causas fabulosas que tornam mais complexa a inocente ação de observar os fenômenos, não é possível deixar de fora o caso das duas irmãs de Psiquê que se atiraram inadvertidamente do precipício sem o auxílio de Zéfiro, primo de Eros (MORAES, 2020).

Muitas vezes, uma discussão começa sem a análise criteriosa dos fundamentos básicos da observação da pesquisa. Assim foi difundido um clássico da sociologia: “o suicídio” de Durkheim (1982). Desde a edição original em 1897, há uma tendência de pensar que entre os protestantes é maior o número de suicídios, tendo em vista a demonstração realizada com números convincentes de registros analisados pelo sociólogo. Porém, muitas críticas foram feitas e dúvidas foram lançadas sobre a forma como se registraram as notificações da morte suspeita de uma pessoa. (MERLLIÉ, 1995)

E se a tradição protestante tiver uma tendência a responsabilizar mais o indivíduo por suas ações do que outra tradição mais latina, como a católica? Para efeito de comparação, pense num simples incêndio na própria residência: entre

brasileiros pode ser percebido como acontecimento “acidental” e entre alemães como “intencional”, produzindo prisão na Alemanha e lamentação pela infelicidade se tivesse ocorrido no Brasil. Foi o caso do jogador de futebol Breno*. A maneira como cada cultura condiciona seus indivíduos a perceber a responsabilidade por um ato, principalmente quando envolve risco de vida, interfere na forma como se registra o acontecimento. Mas nem por isso é preciso jogar fora os números e sim continuar a busca por informações para encontrar explicações mais plausíveis da manifestação de determinadas contagens e entender o que elas estão indicando.

O número “z”, entretanto, não tem nada a ver com isso. Ele estabelece probabilidades abstratas perfeitas, auxilia no cálculo de intervalos de confiança e apoia a criação de níveis de significância de algum resultado estatístico encontrado. Mas o significado de tais resultados é uma tarefa de interpretação do pesquisador e do público que consome as estatísticas. Não é um trabalho simples e definitivo, mas sujeito a revisões e críticas. Enfim, o “z” é uma boa letra e pode ser alcançada com algum esforço, mesmo estando no final da fila, lá no terminal alfabético.



*Breno Vinícius Rodrigues Borges, Wikipédia

Crônica 5: Correlação

O número “z” é uma ferramenta essencial para perceber a correlação entre duas variáveis, ou seja, observar se dois fenômenos “correm” juntos ou separados, de maneira análoga a dois cavalos que podem correr unidos puxando a mesma carruagem ou em bando, correndo livremente. Na primeira analogia, quando a carruagem anda devagar, os dois cavalos diminuem a velocidade conjuntamente e, quando o cocheiro a acelera, os cavalos apressam a cavalgada na mesma intensidade. Já em bando, os cavalos podem as vezes cavalgar de maneira sincronizada no mesmo ritmo e outras vezes não, movimentarem-se ora de forma mais aleatória, ora de forma coordenada, em uma direção ou em várias, muitas vezes em sentido contrário.

Para compreender se existe uma relação entre duas variáveis, ou seja, se elas variam juntas como dois cavalos amarrados numa mesma carruagem ou se variam aleatoriamente independente uma da outra, como se estivessem em um bando de cavalos imprevisíveis, o nosso estudo sobre COVID-19 fornece alguns dados didáticas para explicar em números essa metáfora equina.

Primeiramente, colocam-se lado a lado dois cavalos de tamanhos diferentes, mas que parecem cavalgar na mesma direção e com a mesma intensidade na data de 16/04/2021:

- 1) Número de casos de COVID-19 por Estado
- 2) Número de óbitos por COVID-19 por Estado

Intuitivamente, aposta-se que esses dois cavalos, apelidados de “caso & óbito”, estão puxando a mesma carruagem

e indo para o mesmo destino. A suposição básica é a de que quanto maior o número de casos registrados da doença, maior é o número de óbitos ocasionados por ela. Quando “caso” sobe a montanha, “óbito” corre atrás e vice-versa: quando “óbito” desce, “caso” vem junto na mesma batida. Isso indica que quanto maior for o contágio pela doença, maior tende a ser o impacto do número de mortes que acomete a população nos diferentes Estados. Levando em consideração as mesmas características de mortalidade e taxa de transmissão, essa relação é esperada. O primeiro passo é colocar lado a lado os dados dessas duas variáveis e compará-las visualmente para averiguar se essa tendência se revela nos dados oriundos dos registros oficiais.

Observando as colunas “Casos” e “Óbitos” na tabela 8, é possível notar que os casos estão organizados em ordem crescente e que os números de óbitos por Estado quase ficaram na mesma posição, se não fossem alguns Estados que contrariaram a ordem, como o Estado de Sergipe que fez cair todas as especulações sergipanas feitas na tabela 3, pelo escritor imprudente e afoito. Na coluna de óbitos, se observa outros Estados que também saíram um pouco da ordem, no quesito de óbitos, mas foram apenas pequenas inversões de lugares na ordem crescente. Já o Rio de Janeiro foi o Estado que mais chamou a atenção, pois deveria estar antes apenas de São Paulo.

Então, nessa primeira investida do olhar atencioso, constata-se que os cavalos “Óbito & Caso” estão, de alguma forma, ligados à mesma carruagem “COVID-19”, apesar de não se tratar de uma ligação absolutamente perfeita. Em alguns momentos, alguns Estados saíram do ritmo e atrapalharam o compasso coordenado dos demais. Entretanto, isso não fez a carruagem cair num precipício ou capotar no caos indefinido. De maneira geral, o estatístico continua apostando que os números de casos e os de óbitos estão variando conjuntamente.

Tabela 8: Casos de COVID, em ordem crescente e óbitos por COVID, desvios em relação à média e o desvio padronizado, por Estado, 16/04/2021.

ES	Casos	Desvio da M. Casos	Em DP	Óbitos	Desvio da M. Óbitos	Em DP
AC	75.047	-437.266	-0,83	1.395	-12.262	-0,72
RR	93.259	-419.054	-0,79	1.435	-12.222	-0,72
AP	102.836	-409.477	-0,78	1.456	-12.201	-0,72
TO	151.909	-360.404	-0,68	2.348	-11.309	-0,66
AL	164.404	-347.909	-0,66	3.921	-9.736	-0,57
SE	188.659	-323.654	-0,61	3.903	-9.754	-0,57
RO	203.300	-309.013	-0,59	4.737	-8.920	-0,52
RN	211.050	-301.263	-0,57	5.045	-8.612	-0,51
PI	224.790	-287.523	-0,55	4.693	-8.964	-0,53
MS	235.010	-277.303	-0,53	5.099	-8.558	-0,5
MA	254.189	-258.124	-0,49	6.783	-6.874	-0,4
PB	277.793	-234.520	-0,44	6.388	-7.269	-0,43
MT	338.716	-173.597	-0,33	8.842	-4.815	-0,28
AM	362.401	-149.912	-0,28	12.352	-1.305	-0,08
DF	364.452	-147.861	-0,28	7.125	-6.532	-0,38
PE	378.852	-133.461	-0,25	13.102	-555	-0,03
ES	413.834	-98.479	-0,19	8.581	-5.076	-0,3
PA	448.256	-64.057	-0,12	11.635	-2.022	-0,12
GO	521.333	9.020	0,02	13.481	-176	-0,01
CE	609.849	97.536	0,18	16.075	2.418	0,14
RJ	697.763	185.450	0,35	40.716	27.059	1,59
SC	852.350	340.037	0,64	12.480	-1.177	-0,07
BA	855.871	343.558	0,65	17.134	3.477	0,2
PR	902.626	390.313	0,74	20.182	6.525	0,38
RS	915.558	403.245	0,76	22.977	9.320	0,55
MG	1.266.271	753.958	1,43	29.538	15.881	0,93
SP	2.722.077	2.209.764	4,19	87.326	73.669	4,33
M	512.313,15	0	0	13.657,37	0	0
DP(σ)	527.320,09			17.031,41		

Fonte: pesquisa realizada em 16/04/2021 em covid.saude.gov.br.

Uma segunda prova pode ser obtida a partir da elaboração de dois gráficos com os dados da tabela 8. No gráfico 14, observa-se os casos de COVID-19 por Estado brasileiro em ordem crescente e os desvios em relação à média geral de 512.313,15 casos. Esta última está representada pela linha contínua paralela ao eixo “X” do gráfico. As linhas tracejadas que partem de cada ponto em direção à média representam os desvios de cada Estado. Já na outra página, observando o gráfico 15, é possível se enganar pensando que se trata do mesmo gráfico anterior e da mesma curva crescente, pois, no geral, ambos são bastante parecidos e quase gêmeos univitelinos. Apenas alguns Estados sofreram uma espécie de mutação no “DNA” que impediu a perfeita correspondência em todas as partes dos “corpos gráficos”. Mas tal diferença não significou nenhuma aberração grosseira ou desvio gritante. O Estado do Amazonas, por exemplo, ficou levemente abaixo da média no primeiro gráfico e praticamente na média do segundo. Já o Rio de Janeiro deu um salto maior acima da média, no gráfico 15, enquanto Santa Catarina, que estava um pouco acima da média no gráfico 14, no seguinte deslizou para a média. No geral, é difícil destacar qualquer divergência drástica e seria aceitável afirmar que 95% do “DNA” deles é idêntico e deriva do mesmo “pai”.

Uma das diferenças está na escala de milhões no gráfico dos casos e em milhares, no dos óbitos. Com grandezas tão distintas, seria justo comparar uma diferença de mortes em relação à média de óbitos com a diferença de casos em relação à média de casos? É justo fazer essa comparação no Acre (AC), por exemplo, onde a diferença de mortes em relação à média de óbitos é negativa, igual a -12.262, enquanto a diferença de casos no mesmo Estado, em relação à média dos casos, é de -437.266? E no Rio de Janeiro, onde foram contabilizados 697.763 casos, ficando acima da média em 185.450, enquanto foram registrados 40.716 óbitos, ficando 27.059 acima da média de mortes? Será que é possível comparar grandezas tão distintas?

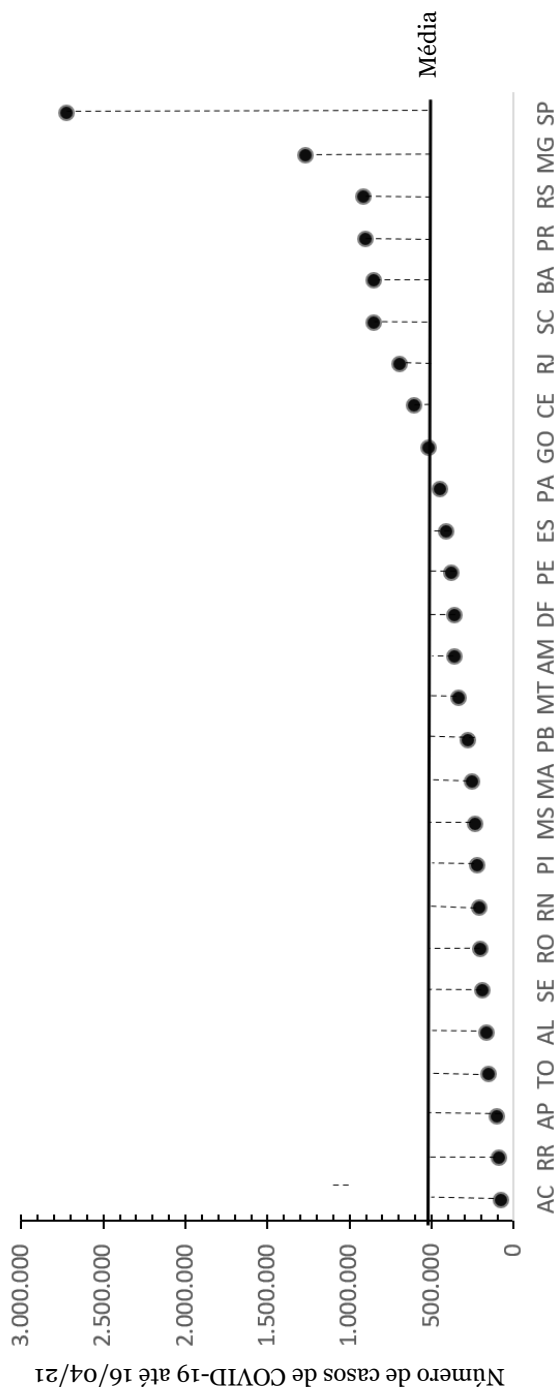


Gráfico 14: Casos de COVID-19 por Estado, Brasil, 2021.

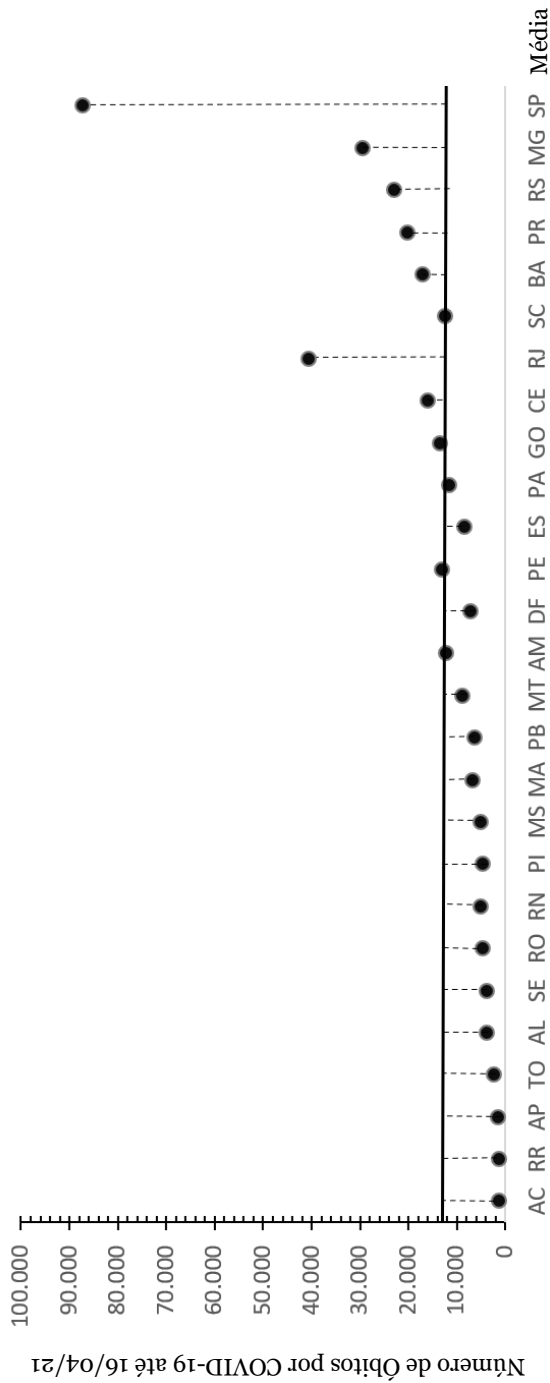


Gráfico 15: Óbitos por COVID-19 por Estado, Brasil, 2021.

No caso do Estado mais populoso que puxa a curva para cima, em São Paulo (SP) foram registrados 2.209.764 de casos acima da média geral, sendo que foram registrados 73.669 óbitos acima da média, na mesma data de 16/04/2021. Os primeiros estão na casa dos milhões e os segundos no campo dos milhares. Mesmo assim, a impressão de sincronia se mantém, as duas variáveis, “casos” e “óbitos”, parecem correr na mesma direção e com a mesma intensidade. Entretanto, as diferenças ou desvios em relação à média são relativos a cada escala considerada e dificultam a comparação justa entre os dois conjuntos de dados.

Para resolver esse problema, é possível reduzir essas diferenças a uma medida comum, tornando a comparação mais adequada e proporcional. O artifício para conseguir essa proeza é calcular qual a posição em termos de desvio padrão que cada Estado ocupa em cada uma das variáveis. Para isso, basta aplicar a fórmula trabalhada na crônica 4, para cada um dos Estados, substituindo as incógnitas pelos valores encontrados na tabela 8. Por exemplo: “z” é igual à diferença entre o valor da variável “x” e a média “ $\bar{x}$ ”, dividida pelo desvio padrão “ σ ”, como a seguir:

$$z = \frac{x - \bar{x}}{\sigma}$$

Então, para o Acre, o número “z” correspondente ao seu número de casos é calculado da seguinte forma:

$$z = \frac{75.047 - 512.313,15}{527.320,09} = \frac{-437.266}{527.320,09} = -0,83$$

Já o número “z” correspondente ao número de óbitos nesse mesmo Estado é calculado de maneira similar:

$$z = \frac{1.395 - 13.657,37}{17.031,41} = \frac{-12.262}{17.031,41} = -0,72$$

Observe que esses números colocados na fórmula acima são encontrados na primeira linha da tabela 8. O valor de -0,83 significa que esse Estado tem um número de casos abaixo da média, porém afastado dela por menos de um desvio padrão, ou seja, -0,83 de desvio padrão. O número de óbitos ficou igualmente abaixo da média e, também, com menos de um desvio padrão da média de mortes, mas especificamente, há -0,72 de desvio padrão da média. O mesmo procedimento deve ser feito para cada linha que representa cada Estado. Para ilustrar um caso com variação positiva em relação à média, a seguir apresentamos o cálculo para o Estado de São Paulo, a última linha listada. O número “z” para o Estado de São Paulo correspondente ao seu número de casos é calculado assim:

$$z = \frac{2.722.077 - 512.313,15}{527.320,09} = \frac{2.209.764}{527.320,09} = 4,19$$

Já o número “z” correspondente ao número de óbitos desse Estado é calculado de maneira similar:

$$z = \frac{87.326 - 13.657,37}{17.031,41} = \frac{73.669}{17.031,41} = 4,33$$

Nesse Estado, o desvio foi positivo e representa que o número de casos ficou muito acima da média, pois foram mais de 4 desvios padrões acima. Ele ficou exatamente a 4,19 desvios padrões acima da média de casos dos Estados nacionais. O número de óbitos também seguiu a mesma tendência: 4,33 desvios padrões acima da média. Agora a medida da distância em relação à média ficou compatível entre as duas variáveis, casos e óbitos, sendo possível fazer uma comparação mais justa. Os dados da tabela 8 foram usados para construir os gráficos 16 e 17, a seguir. Ambos estão na escala de desvio padrão.

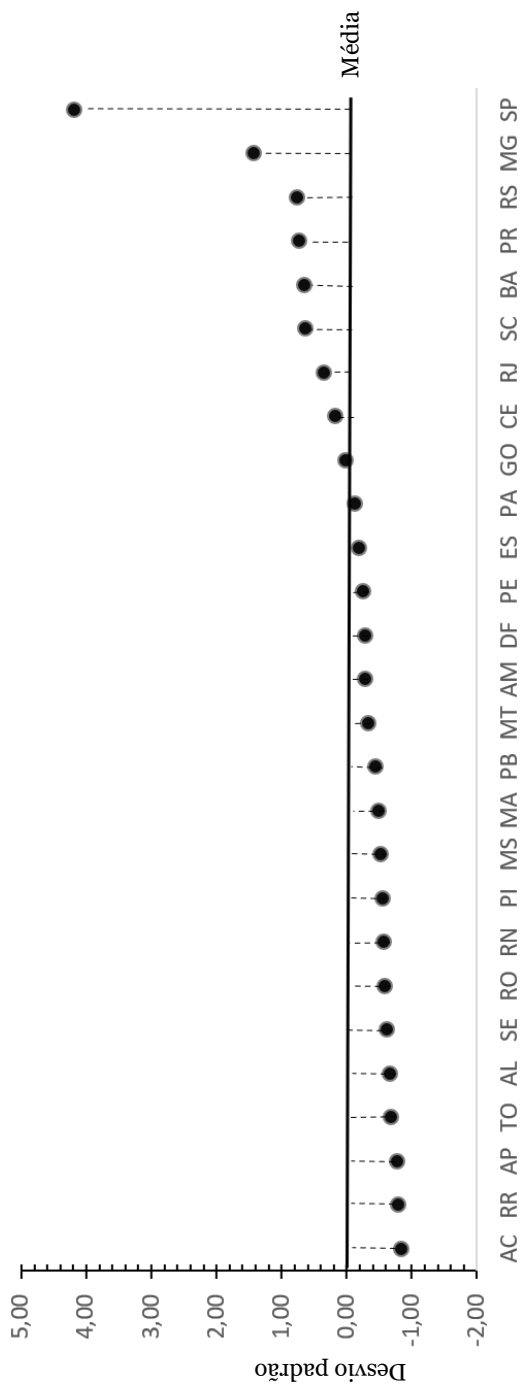


Gráfico 16: Desvio Padronizado dos casos de COVID por Estado, Brasil, 2021.

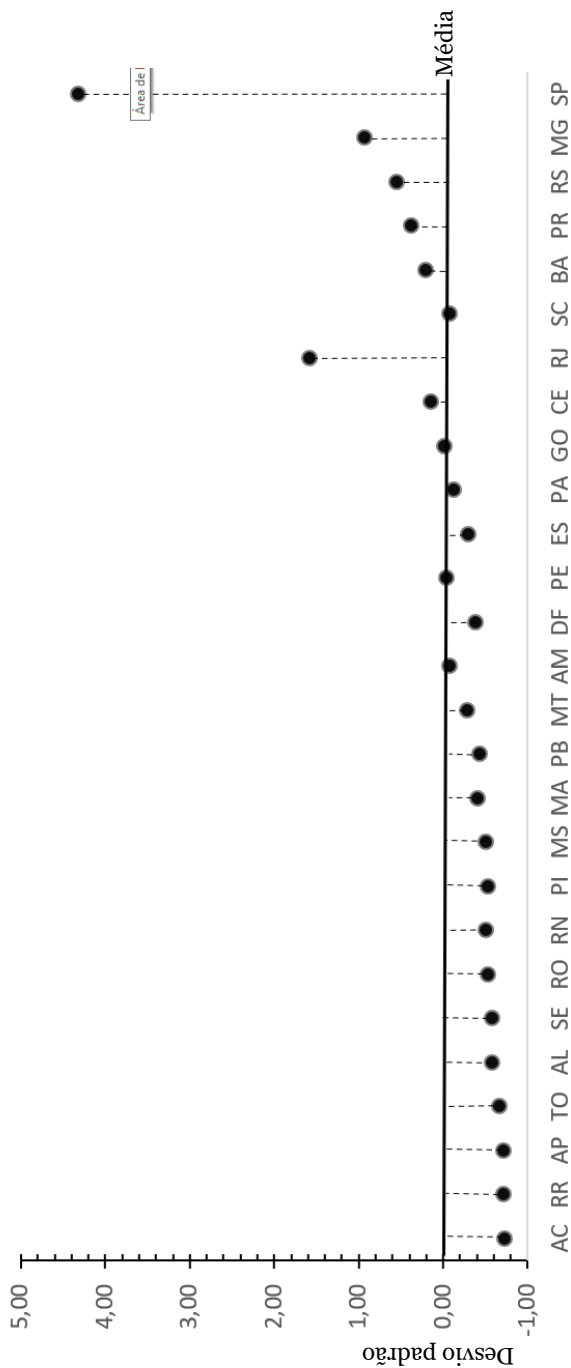


Gráfico 17: Desvio Padronizado dos óbitos por COVID, por Estado, Brasil, 2021.

Observe que cada ponto dos gráficos 16 e 17 corresponde a um valor retirado da coluna “Em DP” da tabela 8 que significa a distância da média em termos de desvio padrão. Ambos gráficos estão na escala de desvio padrão e confirmam a percepção até aqui constatada de que, no geral, são dois fenômenos que caminham juntos na mesma direção, iguais a cavalos que além de atados na mesma carruagem parecem até serem irmãos quase univitelinos, duas criaturas notáveis e comuns no reino da imaginação fantástica de Luis (BORGES, 2000), e aqui denominados de “gráficos quase univitelinos”, dignos de compor a enciclopédia de Borges.

Um estatístico elaborou uma medida para sintetizar essas observações em apenas um número. O nome dele é Pearson e dá nome à medida de correlação que utiliza desses desvios padronizados para calcular um índice que avalia o quanto duas variáveis estão ligadas ou relacionadas. Essa medida varia em um intervalo de menos um (“-1”) até um (“1”). Quanto mais próximo dos extremos, maior é a indicação de que existe uma relação entre as variáveis. Se próximo de “-1”, há uma indicação de relação negativa, ou seja, se uma variável aumenta a outra tende a diminuir na mesma proporção. Do contrário, quanto mais perto de “1”, a relação é positiva, ou seja, quanto maior o valor de uma variável, maior tende a ser o valor da outra, por caso observado. Quanto mais próximo de zero (“0”), maior é o indício de que não existe relação alguma entre elas.

Pelas observações anteriores, é de se esperar que o número calculado resulte num valor mais próximo de “1”. Para calcular o que se convencionou chamar de “R” de Pearson, basta multiplicar o desvio padronizado de uma variável por Estado pelo desvio padronizado da outra variável do mesmo Estado. Na tabela 9, encontramos esse procedimento já realizado. Os valores foram aproximados para duas casas decimais, auxiliando na visualização. Na coluna “DP¹ X DP²” encontra-se o resultado de cada multiplicação por Estado. (Ex. AC = -0,83 x -0,72 = 0,60)

Tabela 9: Multiplicação dos desvios padronizados dos casos pelos óbitos por Estado.

ES	DP ¹ casos	DP ² óbitos	DP ¹ X DP ²
AC	-0,83	-0,72	0,60
RR	-0,79	-0,72	0,57
AP	-0,78	-0,72	0,56
TO	-0,68	-0,66	0,45
AL	-0,66	-0,57	0,38
SE	-0,61	-0,57	0,35
RO	-0,59	-0,52	0,31
RN	-0,57	-0,51	0,29
PI	-0,55	-0,53	0,29
MS	-0,53	-0,5	0,27
MA	-0,49	-0,4	0,20
PB	-0,44	-0,43	0,19
MT	-0,33	-0,28	0,09
AM	-0,28	-0,08	0,02
DF	-0,28	-0,38	0,11
PE	-0,25	-0,03	0,01
ES	-0,19	-0,3	0,06
PA	-0,12	-0,12	0,01
GO	0,02	-0,01	0,00
CE	0,18	0,14	0,03
RJ	0,35	1,59	0,56
SC	0,64	-0,07	-0,04
BA	0,65	0,2	0,13
PR	0,74	0,38	0,28
RS	0,76	0,55	0,42
MG	1,43	0,93	1,33
SP	4,19	4,33	18,14
Total	-	-	25,58
R de Pearson			0,95

Na penúltima linha e na última coluna, encontra-se o somatório dessas multiplicações que é igual a 25,58. Ao dividir esse valor pelo número de Estados, 27, encontra-se o famigerado “R” de Pearson, que para essas variáveis foi de aproximadamente 0,95, muito próximo de uma correlação perfeita, que seria igual a 1.

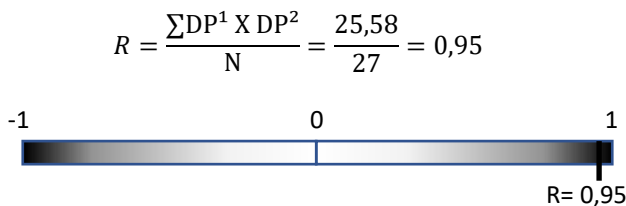


Figura 12: Ilustração do intervalo de variação do “R” de Pearson

O primeiro pensamento que ocorre à mente é que o número de casos provoca o número de mortes, como uma conclusão natural desse procedimento. Entretanto, a confirmação é apenas de uma correlação entre as variáveis sem definir qual delas é a causa da outra. Outro fenômeno poderia estar explicando tanto os casos quanto as mortes, ou seja, uma terceira variável confundidora pode explicar tanto a variação de uma quanto a da outra, sendo a causa da covariação entre ambas. Assim, para estabelecer uma relação de causalidade em que uma variável causa a outra é necessário um árduo trabalho de observação experimental em que o cientista faz registros ao longo do tempo e depois de muitas testagens estabelece uma comprovação fundamentada de uma relação de causalidade. Muitos fenômenos não podem ser objeto de um experimento, pois se trata de vidas humanas. Nessa situação, a correlação pode ser um instrumento adequado e pouco invasivo.

Num experimento clássico, há a manipulação de variáveis para constatar a causa de um fenômeno. Sem

manipulação, devido a implicações éticas, resta ao pesquisador fazer inferências a partir do que já está dado, comparando com diferentes manifestações ao longo do tempo e buscando teorias para auxiliar nas explicações. Para começar a estudar um fenômeno complexo, a correlação é apenas um passo inicial para coletar pistas em um emaranhado de investigações científicas. No geral, é um procedimento que não implica em manipulações perigosas no fenômeno objeto de uma pesquisa. Encontrar uma correlação significativa entre duas variáveis é um procedimento desejável na ciência, pois indica que determinada teoria faz sentido ou que o resultado da pesquisa está sendo corroborado com dados de observação. Entretanto, é comum ao longo das análises não encontrar correlações significativas. Isso não é necessariamente ruim, pois pode ser uma prova para refutar alguma teoria estabelecida e fazer o conhecimento científico avançar.

Então, como seriam duas variáveis que não estão correlacionadas? Como seriam dois cavalos assíncronos ou descoordenados? Como seria observar duas variáveis cujos gráficos apresentam imagens muito distintas? Para apresentar esses dois cavalos danados com comportamentos bem distintos, existe duas variáveis que parecem não estar vinculadas pelo mesmo laço que as prenderiam na mesma carruagem: o número de óbitos por COVID-19 por Estado, em 16/04/2021 e o percentual de votos em Bolsonaro por Estado nas eleições de 2018. Tais números se encontram a seguir, na tabela 10. Lá é possível observar que a coluna “óbitos” se encontra em ordem crescente, do Estado que registrou menos mortes, para aquele no qual mais mortes aconteceram devido a essa infecção virulenta. Quando os olhos saltam para a coluna “votos”, torna-se muito difícil obter qualquer lógica na distribuição do percentual de votos em Jair Bolsonaro em 2018 em função dos óbitos absolutos por Estado.

Tabela 10: Óbitos por COVID, em ordem crescente, e votos, com os desvios em relação à média e o desvio padronizado por Estado brasileiro, 16/04/21.

ES	Óbitos	Desvio da M. Óbitos	Em DP	Votos (%)	Desvio da M. de %	Em DP
AC	1.395	-12.262	-0,72	59	17,26	1,23
RR	1.435	-12.222	-0,72	61	19,26	1,37
AP	1.456	-12.201	-0,72	39	-2,74	-0,19
TO	2.348	-11.309	-0,66	41	-0,74	-0,05
SE	3.903	-9.754	-0,57	24	-17,74	-1,26
AL	3.921	-9.736	-0,57	31	-10,74	-0,76
PI	4.693	-8.964	-0,53	17	-24,74	-1,76
RO	4.737	-8.920	-0,52	59	17,26	1,23
RN	5.045	-8.612	-0,51	28	-13,74	-0,98
MS	5.099	-8.558	-0,5	56	14,26	1,01
PB	6.388	-7.269	-0,43	28	-13,74	-0,98
MA	6.783	-6.874	-0,4	23	-18,74	-1,33
DF	7.125	-6.532	-0,38	55	13,26	0,94
ES	8.581	-5.076	-0,3	50	8,26	0,59
MT	8.842	-4.815	-0,28	52	10,26	0,73
PA	11.635	-2.022	-0,12	34	-7,74	-0,55
AM	12.352	-1.305	-0,08	41	-0,74	-0,05
SC	12.480	-1.177	-0,07	61	19,26	1,37
PE	13.102	-555	-0,03	27	-14,74	-1,05
GO	13.481	-176	-0,01	53	11,26	0,8
CE	16.075	2.418	0,14	20	-21,74	-1,55
BA	17.134	3.477	0,2	21	-20,74	-1,47
PR	20.182	6.525	0,38	53	11,26	0,8
RS	22.977	9.320	0,55	49	7,26	0,52
MG	29.538	15.881	0,93	43	1,26	0,09
RJ	40.716	27.059	1,59	54	12,26	0,87
SP	87.326	73.669	4,33	48	6,26	0,44
M	13.657,37	0	0	41,7407	0	0
DP	17.031,41			14,07		

Nos quatro primeiros Estados, há um número compatível de mortes, entre 1300 e 2400. Entretanto, no Acre, Jair obteve 59 por cento dos votos e em Roraima 61 por cento, enquanto no Amapá e Tocantins, os votos foram de 39 e 41 por cento, respectivamente. O Rio de Janeiro teve metade das mortes do que São Paulo, no entanto, obteve 54 por cento de votos em Bolsonaro contra 48 por cento no outro Estado.

Pernambuco e Goiás tiveram um pouco mais de 13000 mortes, mas o primeiro teve 27 por cento de votos no presidente enquanto o segundo 54. Uma coisa parece nada ter a ver com a outra, como cavalos independentes correndo à vontade num pasto aberto e sem limites. Quando o número de óbitos absoluto de um Estado é maior ou menor isso não indica nenhuma tendência correspondente no percentual de votos nas eleições de 2018. Dessa maneira, as duas variáveis parecem não estar associadas ou correlacionadas e provavelmente o “R” de Pearson deve estar próximo de “0” (zero).

Para ser justo com a primeira análise, é necessário transformar essa tabela 10 nos mesmos gráficos avaliados anteriormente. No gráfico 18, é observado a curva crescente dos óbitos por COVID-19 por Estado, registrados até a data de 16/04/21. A ordem dos Estados agora está de acordo com os óbitos, o que fez alterar a sequência deles em relação aos gráficos anteriores. Por exemplo, o Distrito Federal que estava depois do Estado do Amazonas, agora se encontra antes dele e logo após o Estado do Maranhão. A linha tracejada representa o desvio de cada Estado em relação à média de óbitos. A média é a linha contínua paralela ao eixo do “X”, que por sua vez, é a linha cinza que cruza o eixo “Y” no número zero “0”.

Quando o olho muda o foco para o gráfico 19, não há nenhum padrão claro que possa ser estabelecido. É como se houvesse uma desorganização, de um gráfico para outro, e os números tivessem sido postos aleatoriamente no segundo.

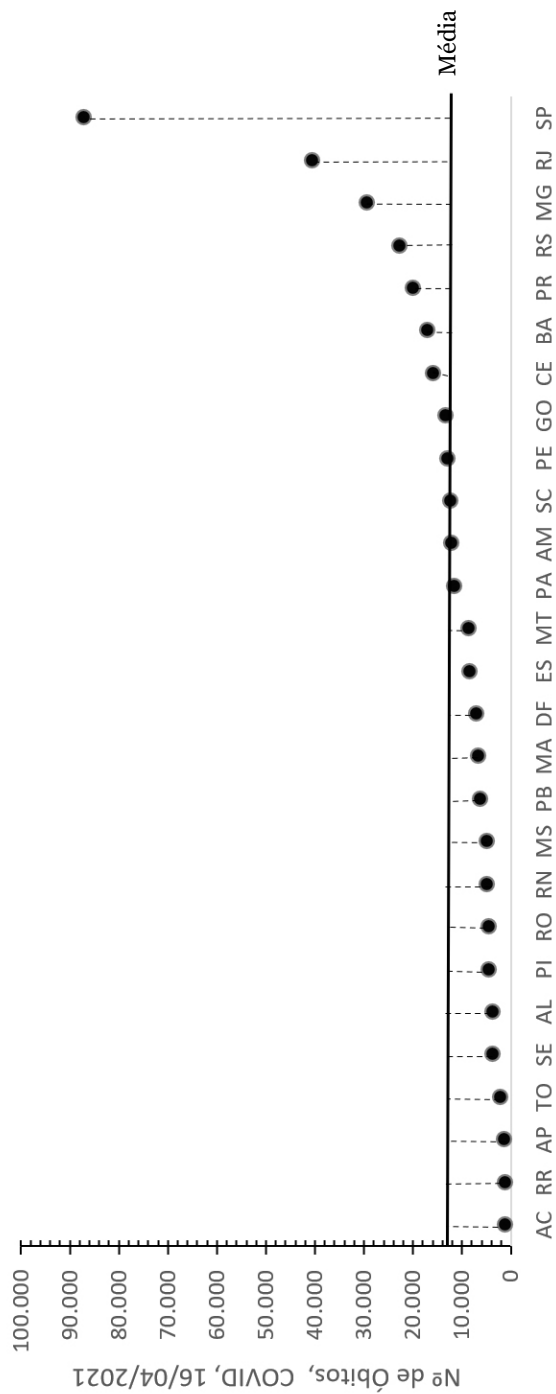


Gráfico 18: Desvio da média de óbitos por COVID-19 por Estado, Brasil, 16/04/2021.

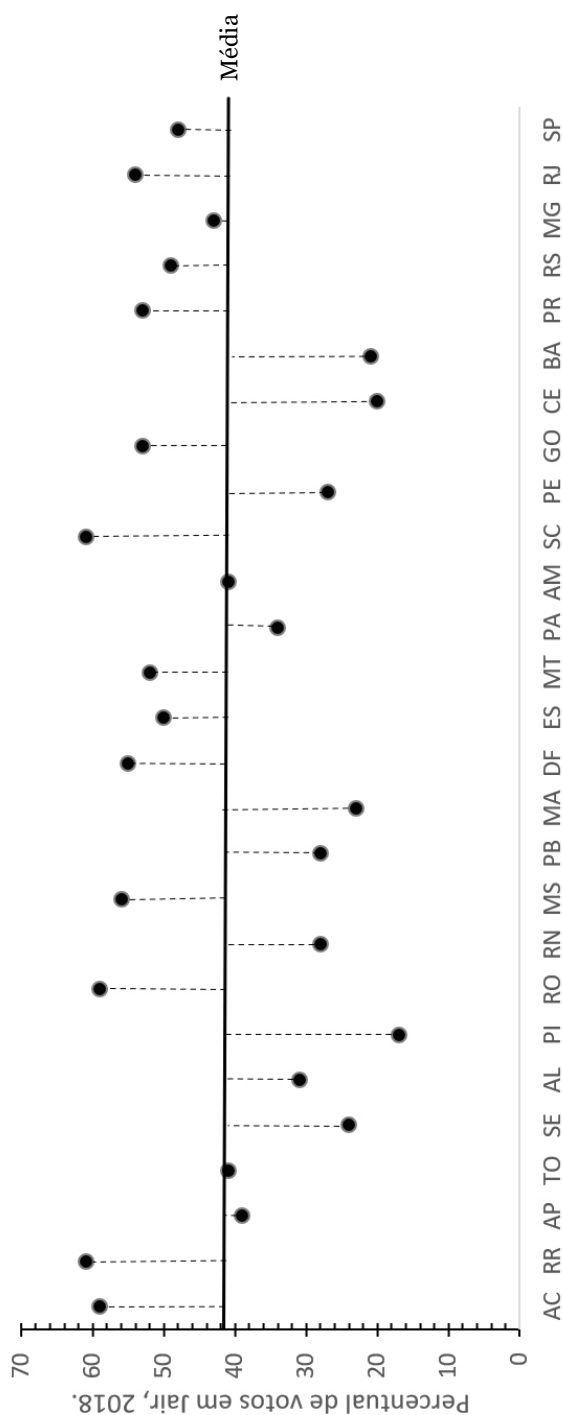


Gráfico 19: Desvio da média do percentual de votos em Bolsonaro, por Estado, 2018.

Fonte: TRE

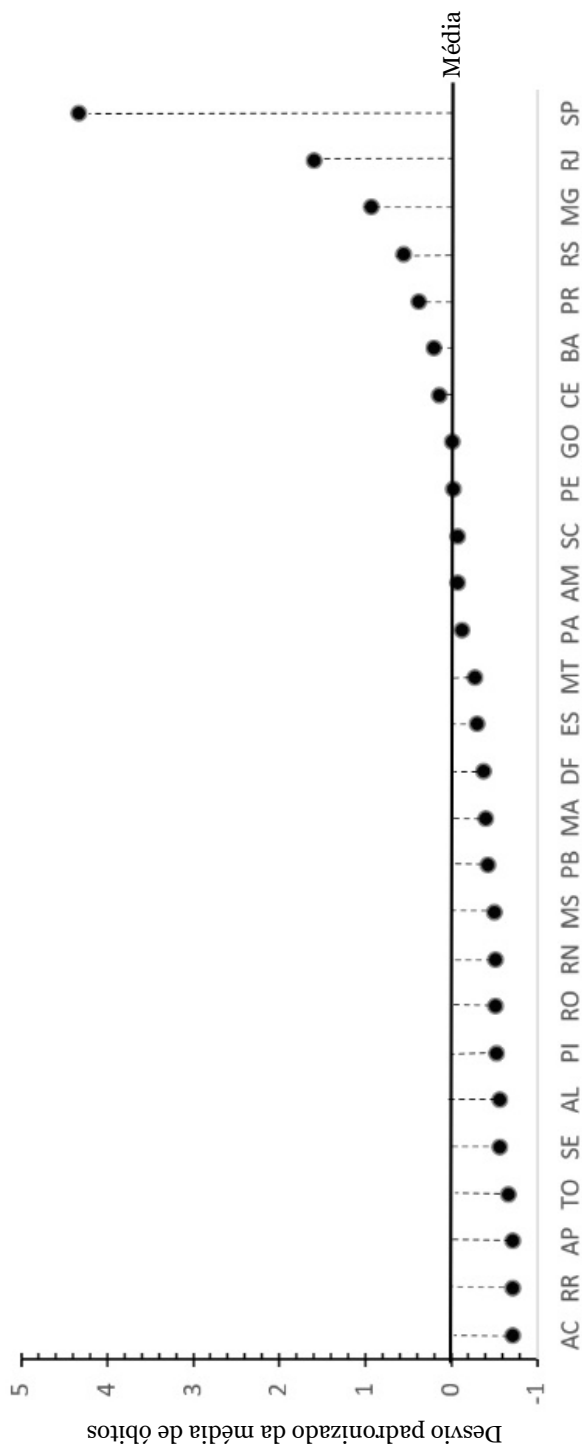


Gráfico 20: Desvio Padronizado da média de óbitos por COVID-19 por Estado, Brasil, 16/04/2021.

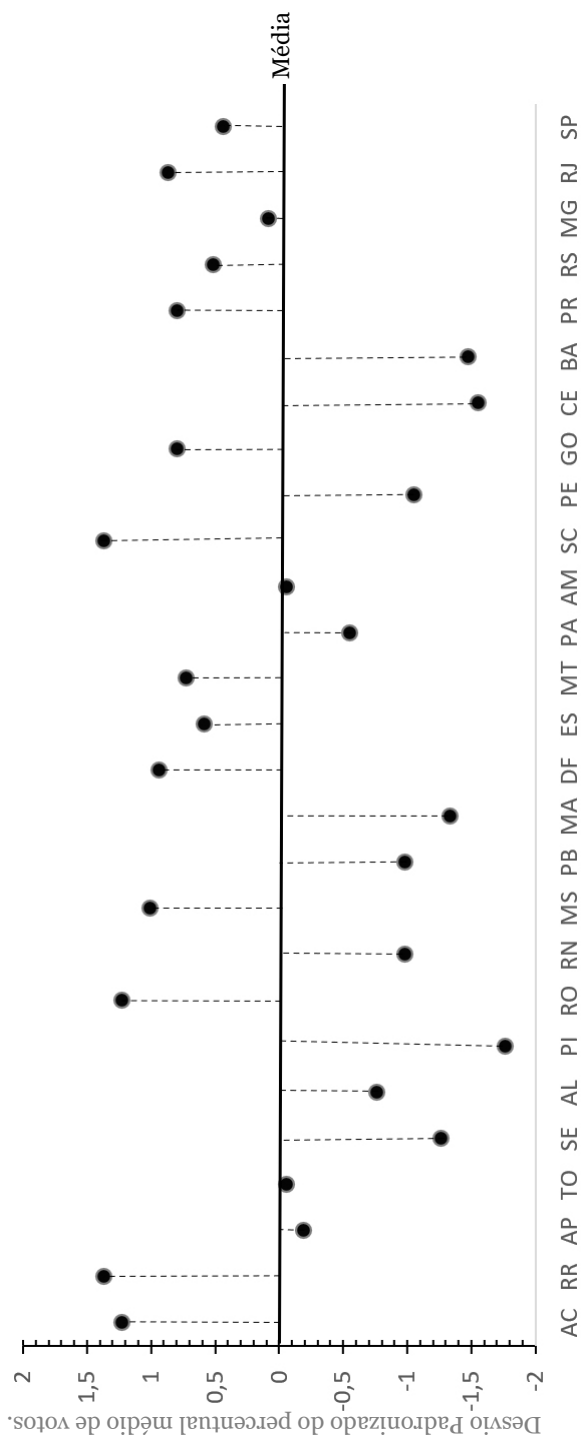


Gráfico 21: Desvio padronizado do percentual médio de votos em Bolsonaro, por Estado, 2018.

Observe que a ordem dos Estados permanece a mesma, do Acre até São Paulo, mas enquanto num gráfico encontra-se uma nítida ordem crescente, no outro os pontos explodem de maneira abrupta em todas as direções. Essa disparidade indica que não há correlação entre essas duas variáveis. Analisando alguns casos, vê-se que o Acre (AC) ficou abaixo da média no primeiro gráfico e bem acima no segundo, similar ao que aconteceu em Roraima (RR) e diferente do Estado seguinte, Amapá (AP) que ficou abaixo da média nos dois gráficos, sendo que no segundo, bem próximo da média. O Estado do Amazonas (AM) ficou na média nos dois gráficos, mas Santa Catarina (SC), ficou na média apenas no primeiro e bem acima da média no segundo. No caso desses gráficos, a impressão de cavalos correndo aleatoriamente é mais apropriada para descrever o movimento dos dados: quando um cavalo vai para direita, o outro fica parado, outro anda para cima e outro corre para baixo, outras vezes, correm juntos, variando aleatoriamente o padrão de movimento sem uma coordenação do conjunto.

Mas alguém poderia argumentar que é muito complicado comparar óbitos com votos, pois estão em medidas completamente diferentes e são fenômenos que não podem ser postos numa mesma balança. Enquanto os óbitos são o resultado da contagem ou de registros obtidos através do atestado de óbito, assinado por um médico que determina a causa da morte, os votos são as manifestações da população em apoio a um determinado candidato, ou seja, uma atitude favorável indicando quem deve ser eleito para um cargo num sistema democrático. A soma das atitudes registradas nos votos determina se um candidato tem a maioria dos votos e está apto a assumir um cargo eletivo. Como alho pode ter alguma relação com bugalhos?

A partir da tabela 10, obtém-se os desvios padronizados dessas variáveis, sendo possível construir os gráficos 20 e 21. No geral, permanecem as mesmas percepções anteriores: não há relação entre elas.

Apesar de serem fenômenos aparentemente distantes, ambos são contagens e são expressos em números que podem ser transformados em desvios padronizados. Nesse sentido, o mesmo procedimento realizado para compatibilizar a grandeza das variáveis “casos” e “óbitos”, será utilizado para tornar “óbitos” e “votos” comparáveis através de uma escala que relativize os seus valores numa mesma proporção.

$$\text{Se: } z = \frac{x - \bar{x}}{\sigma},$$

então, o número “z” para o Estado do Acre, correspondente ao seu percentual de votos é o seguinte:

$$z = \frac{59 - 41,74}{14,07} = \frac{17,26}{14,07} = 1,23$$

O número “z” correspondente ao número de óbitos nesse mesmo Estado é calculado de maneira similar:

$$z = \frac{1.395 - 13.657,37}{17.031,41} = \frac{-12.262}{17.031,41} = -0,72$$

Esse valor positivo de 1,23 referente ao desvio padronizado do percentual de votos já se encontra em contradição com o valor negativo de -0,72, que é o desvio padronizado dos óbitos no Acre. Isso significa que enquanto os óbitos ficaram abaixo da média, mais especificamente um pouco menos de um desvio padrão abaixo da média, os votos nesse mesmo Estado ficaram acima da média, um pouco mais de um desvio padrão acima da média. Já no Estado de São Paulo, olhando para os seus resultados na tabela 11, observa-se que ambas as variáveis ficaram acima da média, sendo que os óbitos ficaram bem acima da média, acima de 4 desvios padrões da média, enquanto os votos, que também ultrapassaram a média

de votos por Estado, conquanto com muito menos intensidade, ficaram cerca de meio desvio padrão acima da média (0,44). Enquanto a multiplicação dos escores do Acre resultará em um número negativo, em São Paulo, o resultado será positivo. Quando o processo de somar todas as multiplicações for executado, os valores negativos tendem a anular os positivos. Assim, o somatório geral vai se aproximar de zero.

O último passo da análise é calcular o “R” de Pearson. Na tabela 11, encontram-se os valores necessários para essa operação. Na coluna “D.P.¹ X D.P.²” estão os valores obtidos pela multiplicação das duas colunas anteriores, linha por linha, Estado por Estado. Embaixo dessa coluna, está o somatório de todas essas multiplicações, que deu 3,40. Ao dividir esse somatório pelo número de Estados, encontra-se o “R” de Pearson, que foi de 0,13. A seguir, a fórmula do cálculo, onde DP¹ representa o desvio padronizado da média da variável 1 (Óbitos) e DP² o desvio padronizado da média da variável 2 (Votos).

$$R = \frac{\sum(DP^1 \cdot DP^2)}{N} = \frac{3,4}{27} = 0,13$$

Esse valor de 0,13 é bastante próximo de zero “0”, indicando que não há uma relação entre óbitos absolutos por Estado e votos em Bolsonaro em 2018. Então, a seguir, na figura 13, um diagrama didático para desenhar um exemplo de correlação quase perfeita e outro exemplo da inexistência de correlação. Na parte de cima, a correlação de 0,95 entre o número de casos e de Óbitos por COVID-19 foi caracterizada por uma carruagem puxada por dois cavalos que serviu como a metáfora para caracterizar essa forte relação. Já na parte de baixo, a ausência de correlação entre votos em 2018 e óbitos absolutos por Estado foi representada por cavalos soltos que descrevem, numa imagem, o “R” de Pearson de apenas 0,13, próximo a zero.

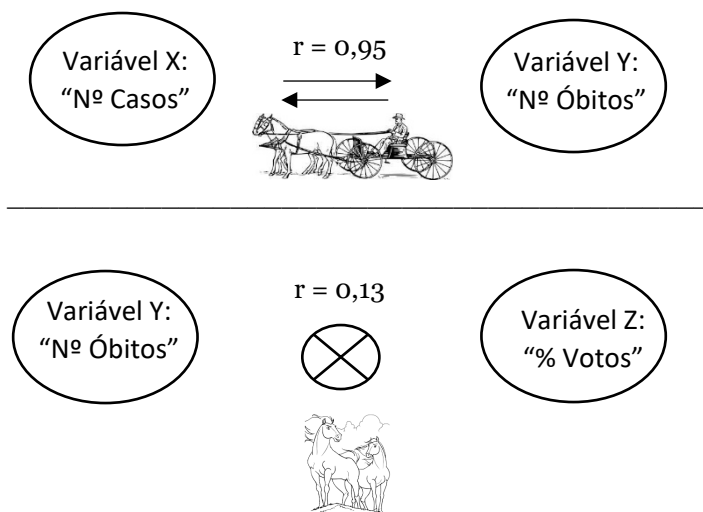


Figura 13: Exemplo de correlações: em cima, uma correlação quase perfeita, representada por dois cavalos puxando uma carruagem; embaixo, exemplo da ausência de correlação, representada pela ideia de cavalos soltos cavalgando aleatoriamente.

Como foi dito anteriormente, a correlação não determina uma relação de causalidade. Outra variável pode estar influenciando o comportamento das variáveis estudadas "óbitos" e "casos". Por exemplo, o tamanho da população pode estar influenciando tanto o número de casos, quanto o número de óbitos, como ilustrado na figura 14 a seguir:

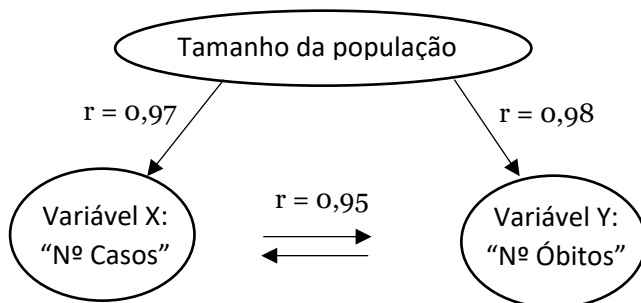


Figura 14: Exemplo de variável confundidora.

Tabela 11: Desvio padronizado da média de óbitos (2021) e de votos (2018), e produto dos desvios padronizados por Estado.

ES	D.P. ¹ da M. de Óbitos	D. P ² . da M. de Votos	D.P. ¹ X D.P ² .
AC	-0,72	1,23	-0,89
RR	-0,72	1,37	-0,99
AP	-0,72	-0,19	0,14
TO	-0,66	-0,05	0,03
SE	-0,57	-1,26	0,72
AL	-0,57	-0,76	0,43
PI	-0,53	-1,76	0,93
RO	-0,52	1,23	-0,64
RN	-0,51	-0,98	0,50
MS	-0,5	1,01	-0,51
PB	-0,43	-0,98	0,42
MA	-0,4	-1,33	0,53
DF	-0,38	0,94	-0,36
ES	-0,3	0,59	-0,18
MT	-0,28	0,73	-0,20
PA	-0,12	-0,55	0,07
AM	-0,08	-0,05	0,00
SC	-0,07	1,37	-0,10
PE	-0,03	-1,05	0,03
GO	-0,01	0,80	-0,01
CE	0,14	-1,55	-0,22
BA	0,2	-1,47	-0,29
PR	0,38	0,80	0,30
RS	0,55	0,52	0,29
MG	0,93	0,09	0,08
RJ	1,59	0,87	1,38
SP	4,33	0,44	1,91
Total	-	-	3,40
R de Pearson			0,13

Observe que as correlações entre o tamanho da população de cada Estado com o número de casos e o número de óbitos é ainda maior do que a correlação entre essas duas últimas variáveis entre si: 0,98 e 0,97 contra 0,95, respectivamente. Essa informação é relevante para avaliar comparações feitas entre os Estados. Não é justo, por exemplo, comparar um Estado semelhante a São Paulo, com uma população grande, com um Estado pequeno, como o Acre. Seria preciso retirar o peso dessa variável populacional para analisar a possibilidade de qualquer outra relação. Ela pode estar funcionando como aquela famigerada “variável confundidora” que impede o estatístico de visualizar uma relação legítima entre duas variáveis. Na epidemiologia, os pesquisadores convencionaram calcular as mortes dividindo-as por grupos de 100 mil habitantes ou mil habitantes, criando assim padrões de comparação que possibilitam relativizar o tamanho da população. Assim é possível comparar países ou Estados sem comprometer a análise e concentrando a atenção na variável de interesse para o estudo.

Se um bolsonarista leu até aqui, é recomendável que pare imediatamente e dissemine a meia verdade encontrada nas páginas anteriores. Será um excelente material para *fake News* que por instantes fortuitos, em redes sociais volúveis, acariciaria o ego dos simpatizantes com narrativas enviesadas.

Lamento se ainda algum teimoso “bolsonarista-raiz” insiste em ler o livro até o fim. Ou ele terá que se converter para o “comunismo” ou viverá em dissonância cognitiva, que é uma espécie de incoerência entre duas proposições, percepções, ações ou pensamentos incompatíveis: “como o meu voto pode estar correlacionado com óbitos? Como uma atitude minha de votar no Messias pode gerar morte?”. As revelações das próximas páginas são desesperadoramente óbvias e assustadoramente simples para a compreensão de qualquer bom aluno do ensino básico da educação brasileira e pode levar um simpatizante do presidente a rever sua atitude e considerar a gravidade dos números.

Para eliminar a incoerência nas próprias ideias, pensamentos e ações, o indivíduo pode reavaliar suas posições iniciais e como numa metamorfose ambulante, corrigir o rumo para um caminho mais coerente e racional que resolva as contradições que são insuportáveis. Infelizmente, a solução para a dissonância cognitiva, muitas vezes, é resolvida também com a negação de uma das partes incoerentes ou contraditórias. A incoerência entre as próprias ações, pensamentos e ideias pode ser resolvida contorcendo os fatos ou reinterpretando as ações. “Serás tu a negar a matemática por três vezes antes do alvorecer do próximo capítulo, Pedro? Será que, realmente, os números não mentem? Contra fatos não existem argumentos?” Em tempos de negação da ciência, a astúcia da mente dá nó em “pingo d’água” e os fatos são costurados com a inventividade e a transcendência dos argumentos. Não se trata de menosprezar a inteligência de alguns em compreender os cálculos estatísticos. Talvez seja inteligência até demais, pois para ser capaz de enganar a si próprio, o sujeito necessita de argúcia. Afinal, acreditar numa fábula envolve uma capacidade de imaginar e criar todo um cenário provável e convincente que é projetado na realidade e orienta a pessoa na direção desejada.

A população é muito bem-educada para acreditar em fábulas e não para pensar por si própria. É educada para acreditar na ciência e não para ser cientista, acreditar nos números do economista e do médico e não para calcular por si própria e tirar as próprias conclusões. Esse é o paradigma da civilização cristã: as pessoas são estimuladas a acreditar em Deus. Um iniciado nos mistérios da espiritualidade estranha essa exótica pergunta: “Você acredita em Deus?”. Ele responderia: “Como é possível acreditar em algo que você não conhece?”. Em uma entrevista, o psicólogo Jung afirmou que não acreditava em Deus. Logo em seguida ele acrescentou: “Eu tenho a convicção”. Uma sociedade de oprimidos necessita da fé para continuar a viver. Mas a fé é apenas um degrau no caminho da consciência

plena e realizada. Muitos líderes manipulam a fé do povo em benefício próprio e conseguem o prodígio de distorcer até os números mais óbvios que surgem da observação dos assuntos públicos, registrados pelos órgãos oficiais. Vamos, então, desenhar para eles o retrato falado dos criminosos que estão relacionados com o número de óbitos: os votos.

Na tabela 12, encontra-se o número de óbitos por 100 mil habitantes, por Estado Brasileiro na data de 16/04/2021. As maiores taxas estão no alto da tabela, acima de 200. Já na parte de baixo da tabela, todos os Estados encontram-se com menos de 200 mortes por 100 mil cidadãos. Bahia, Ceará e Piauí ficaram no final da coluna e tiveram respectivamente 115, 176 e 143 mortes por mil habitantes. Deslocando o olhar para a coluna dos votos, observa-se que eles estão dispostos em ordem decrescente, do Estado com mais votos em Bolsonaro, Roraima 61%, até aquele com menor percentual de votos, Piauí 17%. De maneira grosseira, a “olho nu”, percebe-se alguma relação entre aqueles Estados com maior percentual de votos e a incidência de mortes por 100 mil habitantes. As outras colunas trazem os dados necessários para a construção dos gráficos e para o cálculo do “R” de Pearson: o desvio em relação à média de óbitos e o respectivo desvio padronizado, o desvio em relação à média do percentual de votos em Jair e o respectivo desvio padronizado, por Estado.

A maior taxa acontece no Estado do Amazonas, 298 mortes a cada 100 mil amazonenses. Ele se destacou devido à falta de oxigênio hospitalar para atender ao crescente número de pessoas internadas por conta da COVID-19. Ali houve maior descontrole da pandemia decorrente de uma política equivocada quanto à condução dos cuidados sanitários que evitariam a propagação do vírus e diminuiriam as chances de internamento hospitalar e consequente aumento do número de óbitos.

Tabela 12: Número de óbitos por COVID por 100 mil habitantes (16/04/2021), percentual de votos em Bolsonaro (2018), desvios em relação à média e o desvio padronizado, por Estado brasileiro.

ES	Óbitos / 100 mil	Desvio da M. Óbitos	Em DP	Votos (%)	Desvio da M. de %	Em DP
RR	237	56	1,16	61	19,26	1,37
SC	174	-6	-0,13	61	19,26	1,37
AC	158	-22	-0,46	59	17,26	1,23
RO	267	86	1,77	59	17,26	1,23
MS	184	3	0,06	56	14,26	1,01
DF	236	56	1,15	55	13,26	0,94
RJ	236	55	1,14	54	12,26	0,87
GO	192	12	0,24	53	11,26	0,80
PR	177	-4	-0,08	53	11,26	0,80
MT	254	73	1,51	52	10,26	0,73
ES	214	33	0,68	50	8,26	0,59
RS	202	22	0,44	49	7,26	0,52
SP	190	10	0,20	48	6,26	0,44
MG	140	-41	-0,84	43	1,26	0,09
AM	298	118	2,42	41	-0,74	-0,05
TO	149	-31	-0,64	41	-0,74	-0,05
AP	172	-8	-0,17	39	-2,74	-0,19
PA	135	-45	-0,93	34	-7,74	-0,55
AL	118	-63	-1,30	31	-10,74	-0,76
PB	159	-21	-0,44	28	-13,74	-0,98
RN	144	-37	-0,75	28	-13,74	-0,98
PE	137	-43	-0,89	27	-14,74	-1,05
SE	170	-11	-0,22	24	-17,74	-1,26
MA	96	-85	-1,74	23	-18,74	-1,33
BA	115	-65	-1,34	21	-20,74	-1,47
CE	176	-4	-0,09	20	-21,74	-1,55
PI	143	-37	-0,76	17	-24,74	-1,76
M	180,43	0	0	41,74	0	0
DP	48,55			14,07		

Um exercício simples, mas convincente, é cortar a tabela ao meio, bem no coração do Brasil, em Minas Gerais. Calculando a média do grupo superior, encontra-se aproximadamente 209 mortes por 100 mil habitantes causadas por COVID-19. Já na metade inferior, mesmo com a presença do Estado do Amazonas, a média cai para 155, aproximadamente.

Os Estados do Nordeste estão expressivamente concentrados na parte de baixo e os do sul e sudeste estão acima, deixando qualquer atlas geográfico de cabeça para baixo: como pode o sul ficar em cima e o nordeste em baixo? O que causou essa anomalia não foi nenhuma hecatombe, mas simplesmente o fato de serem os Estados Nordestinos aqueles que menos ofereceram seus votos ao suposto Messias e assim se posicionaram na parte de baixo da tabela.

No gráfico 22, a seguir, observa-se uma linha descendente que indica a ordem decrescente dos Estados alinhados segundo o percentual de votos obtidos por Jair M. Bolsonaro. As linhas tracejadas representam os desvios em relação à média. Roraima e Santa Catarina foram os Estados que ficaram mais acima da média, com 19 pontos percentuais acima da média de 42 por cento, aproximadamente. Já o Ceará e o Piauí encontram-se no extremo oposto, pois têm respectivamente -22 e -25 pontos percentuais abaixo dessa média.

Passando a atenção para o gráfico 23, o padrão não se revela tão nítido quanto as correlações encontradas anteriormente. Mas é precipitado afirmar categoricamente que não há relação nenhuma. Observe o Estado de Minas Gerais no centro do gráfico. Com exceção do Estado do Amazonas que demanda uma análise particular, todos os Estados à direita encontram-se abaixo da média de 180,43 óbitos por 100 mil habitantes, representada pela linha contínua paralela ao eixo “X” do gráfico.

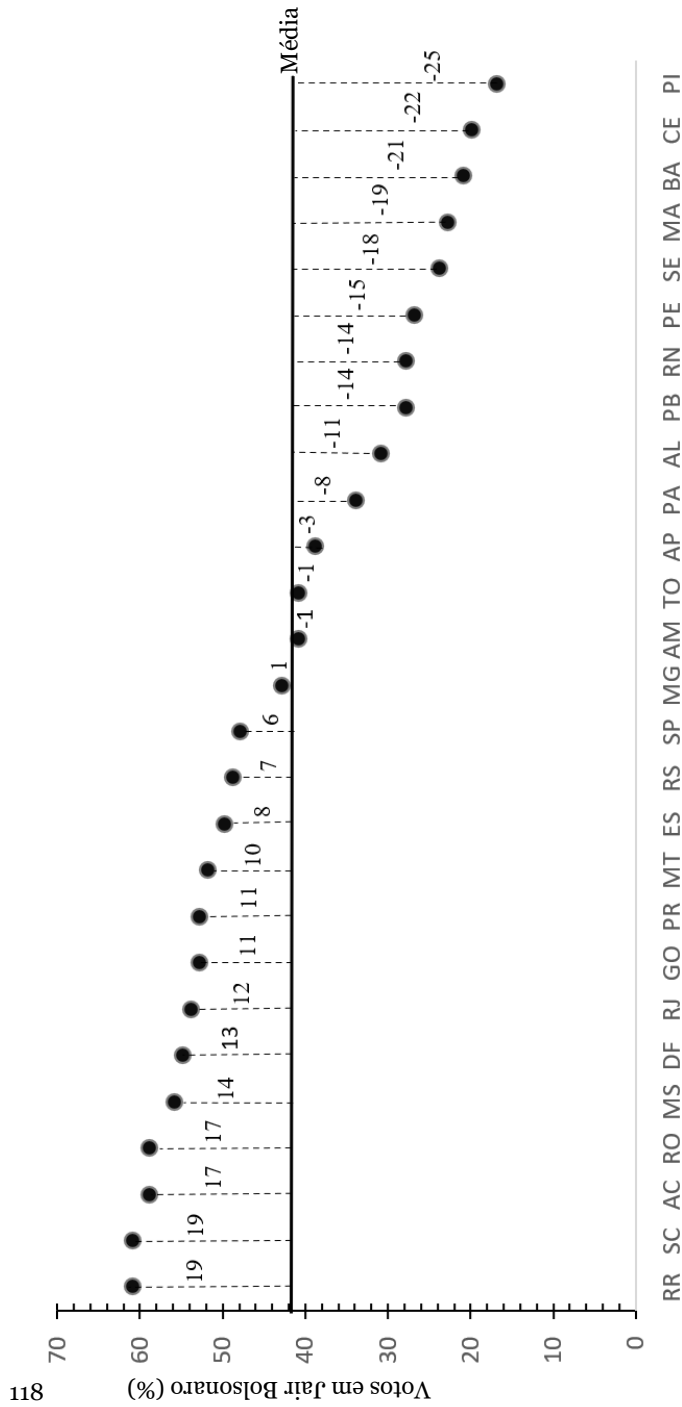


Gráfico 22: Desvio da média do percentual de votos em Bolsonaro, por Estado, em ordem decrescente 2018.

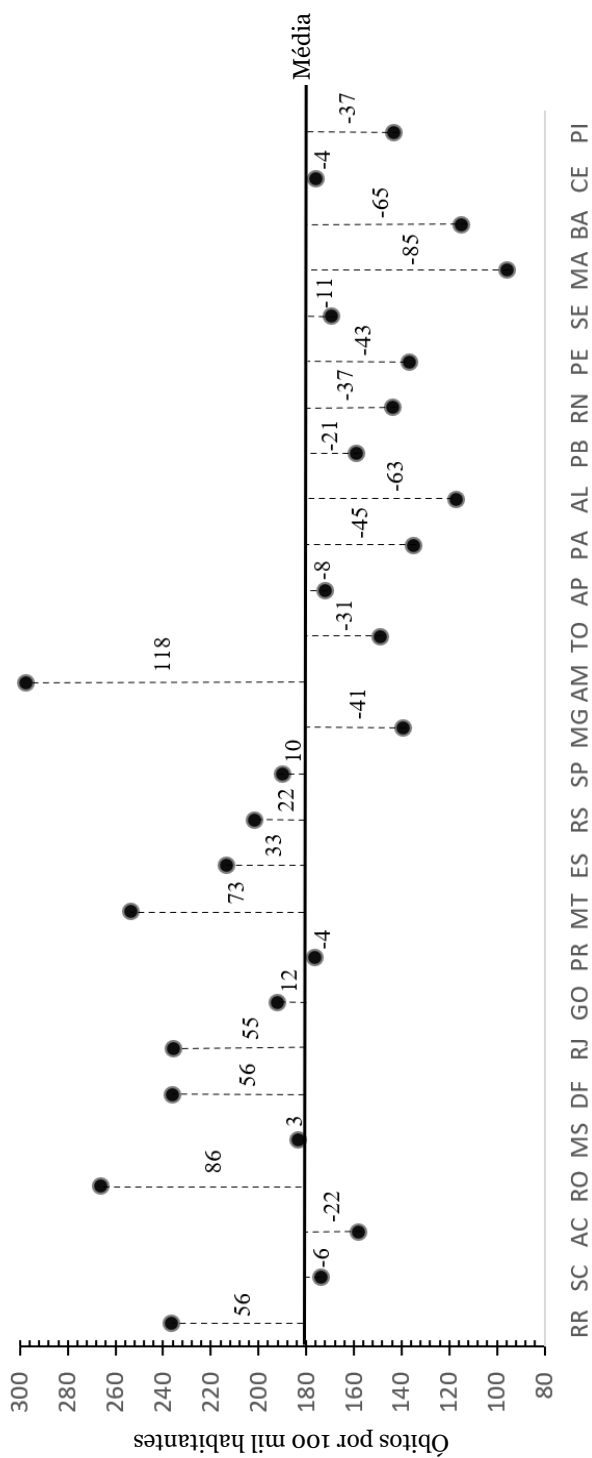


Gráfico 23: Desvio da média dos óbitos por COVID por 100 mil habitantes, por Estado, 16/04/2021.

Já do lado esquerdo de quem vê, apenas dois Estados encontram-se levemente abaixo da média, Santa Catarina e Paraná com, respectivamente, -6 e -4 mortes por 100 mil, ou seja 6 e 4 mortes abaixo da média e um bem mais sensivelmente abaixo da média, o Acre com -22. De fato, a maioria esmagadora do grupo de Estados da esquerda está acima da média no que tange ao número de mortes por 100 mil indivíduos. São também os Estados nos quais o atual presidente da república obteve a mais expressiva quantidade de votos.

Os gráficos demonstram que há alguma correlação entre essas duas variáveis, mas não se trata de uma correlação tão perfeita quanto aquelas que foram detectadas anteriormente. Se aquelas foram correlações bastante fortes, essa se revela uma correlação moderada.

Na tabela 13, a seguir, encontram-se os dados necessários para o cálculo do “R” de Pearson. O procedimento é o mesmo que o realizado anteriormente: multiplicar os desvios padronizados de “óbitos por 100 mil” com os de “votos em Jair” e realizar o somatório de todas essas multiplicações por Estado. No final da coluna “D.P.¹x D.P.²”, encontra-se esse somatório que foi igual a 14,64. Esse número dividido por 27, o número de Estados, é igual a 0,62, ou seja, “R” = 0,62. Tendo em vista que o valor de “R” varia de -1 até 1, esse valor representa uma correlação moderada entre as duas variáveis analisadas.

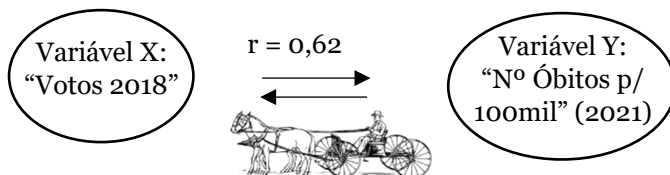


Figura 15: Correlação moderada de 0,62 entre votos (2018) e óbitos por 100 mil (2021)

Tabela 13: Desvio padronizado da média de óbitos por 100 mil (2021) e de votos (2018), e produto dos desvios padronizados por Estado.

ES	D.P. ¹ da M. de Óbitos p/ 100mil	D.P. ² da M. de Votos	D.P. ¹ X D.P. ²
AC	1,16	1,37	1,59
RR	-0,13	1,37	-0,18
AP	-0,46	1,23	-0,56
TO	1,77	1,23	2,17
AL	0,06	1,01	0,06
SE	1,15	0,94	1,08
RO	1,14	0,87	0,99
RN	0,24	0,80	0,19
PI	-0,08	0,80	-0,06
MS	1,51	0,73	1,10
MA	0,68	0,59	0,40
PB	0,44	0,52	0,23
MT	0,20	0,44	0,09
AM	-0,84	0,09	-0,08
DF	2,42	-0,05	-0,13
PE	-0,64	-0,05	0,03
ES	-0,17	-0,19	0,03
PA	-0,93	-0,55	0,51
GO	-1,30	-0,76	0,99
CE	-0,44	-0,98	0,43
RJ	-0,75	-0,98	0,73
SC	-0,89	-1,05	0,93
BA	-0,22	-1,26	0,28
PR	-1,74	-1,33	2,32
RS	-1,34	-1,47	1,98
MG	-0,09	-1,55	0,14
SP	-0,76	-1,76	1,34
Soma	0	0	16,64
R			0,62

A Estatística dificilmente estabelece relações de causalidade, tendo em vista a necessidade de experimentação para determinar causas cientificamente comprovadas. Entretanto, devido à questão temporal, intuitivamente, constata-se que “os votos” aconteceram primeiro, com quase três anos de antecedência em relação aos óbitos. Com esse fato em tela, a maneira correta de enunciar esse achado é dizer que uma boa parte da variação dos óbitos por 100 mil habitantes por Estado é explicada pela variação no percentual de votos em 2018. Diferente das correlações anteriores, nas quais quase 100 por cento da variação dos óbitos absolutos e dos casos de COVID-19 é explicada pelo tamanho da população, no caso dos votos, um pouco menos de 40 por cento da variação dos óbitos por 100 mil é explicada pelos votos de 2018.

Alguém poderia indagar: “E se esse resultado não passou de uma coincidência ou acidente ao acaso?” Ano passado, quando se começou a redação desse livro, a construção dos gráficos e os cálculos estatísticos, havia aproximadamente 368.000 mortes registradas por COVID-19. Foi na data de 16 de abril de 2021 que o interesse do pesquisador se fixou em estudar os dados oficiais e calcular a correlação entre votos e óbitos. “Talvez mais de sete meses depois, em 02/12/2021, os dados tenham se ajeitado e essa correlação tenha sumido!” insiste um incrédulo negacionista. A experiência indica o contrário, pois quando grandes tendências se revelam em grandes números, dificilmente elas somem simplesmente sem dar satisfação. Para ser didático, a seguir uma atualização dos dados trabalhados anteriormente.

Na tabela 14, se encontram os dados atualizados de mortes por 100 mil habitantes até 02/12/2021 às 17:47. É a mesma tabela anterior com a diferença da atualização dos

dados. No final da tabela se encontra a média dessa variável que aumentou para 282,2, um aumento de 100 mortes em relação à média anterior. O desvio padrão também aumentou para um pouco mais de 70. O Amazonas não se destaca mais como um caso extremo (compare o gráfico 23 com o 24). A maior taxa agora está com o Rio de Janeiro, com mais de 400 mortes por 100 mil habitantes. As maiores taxas continuam no alto da tabela, enquanto embaixo estão as menores, acompanhando a coluna do percentual de votos em Bolsonaro, que está em ordem decrescente do maior para o menor. Na coluna dos desvios é possível perceber que da metade da tabela para cima encontram-se desvios positivos em relação à média, com exceção de, apenas, dois Estados: Santa Catarina, levemente abaixo da média -2,86, e o Acre com -72,46 abaixo da média. A partir de Minas Gerais, todos os desvios são negativos, com exceção do Estado do Amazonas que teve um desvio positivo em relação à média, 50,94. Isso indica que a mesma tendência observada anteriormente com os dados de 16/04/2021 se repete com mais vigor em 03/12/2021.

Tabela 14: Número de óbitos por COVID por 100 mil habitantes (03/12/2021), percentual de votos em Bolsonaro (2018), desvios em relação à média e o desvio padronizado, por Estado brasileiro.

ES	Óbitos / 100 mil	Desvio da M. Óbitos	Em DP	Votos (%)	Desvio da M. de %	Em DP
RR	339,1	56,94	0,80	61	19,26	1,37
SC	279,3	-2,86	-0,04	61	19,26	1,37
AC	209,7	-72,46	-1,02	59	17,26	1,23
RO	374,0	91,84	1,29	59	17,26	1,23
MS	348,6	66,44	0,93	56	14,26	1,01
DF	366,1	83,94	1,18	55	13,26	0,94
RJ	400,2	118,04	1,65	54	12,26	0,87
GO	349,8	67,64	0,95	53	11,26	0,80
PR	356,9	74,74	1,05	53	11,26	0,80
MT	394,6	112,44	1,58	52	10,26	0,73
ES	328,4	46,24	0,65	50	8,26	0,59
RS	317,9	35,74	0,50	49	7,26	0,52
SP	335,8	53,64	0,75	48	6,26	0,44
MG	265,8	-16,36	-0,23	43	1,26	0,09
TO	249,2	-32,96	-0,46	41	-0,74	-0,05
AM	333,1	50,94	0,71	41	-0,74	-0,05
AP	237,0	-45,16	-0,63	39	-2,74	-0,19
PA	196,8	-85,36	-1,20	34	-7,74	-0,55
AL	190,5	-91,66	-1,28	31	-10,74	-0,76
PB	213,9	-68,26	-0,96	28	-13,74	-0,98
RN	237,3	-44,86	-0,63	28	-13,74	-0,98
PE	211,9	-70,26	-0,98	27	-14,74	-1,05
SE	263,0	-19,16	-0,27	24	-17,74	-1,26
MA	145,6	-136,56	-1,91	23	-18,74	-1,33
BA	183,7	-98,46	-1,38	21	-20,74	-1,47
CE	270,1	-12,06	-0,17	20	-21,74	-1,55
PI	219,9	-62,26	-0,87	17	-24,74	-1,76
M	282,2	0	0	41,74	0	0
DP	71,35			14,07		

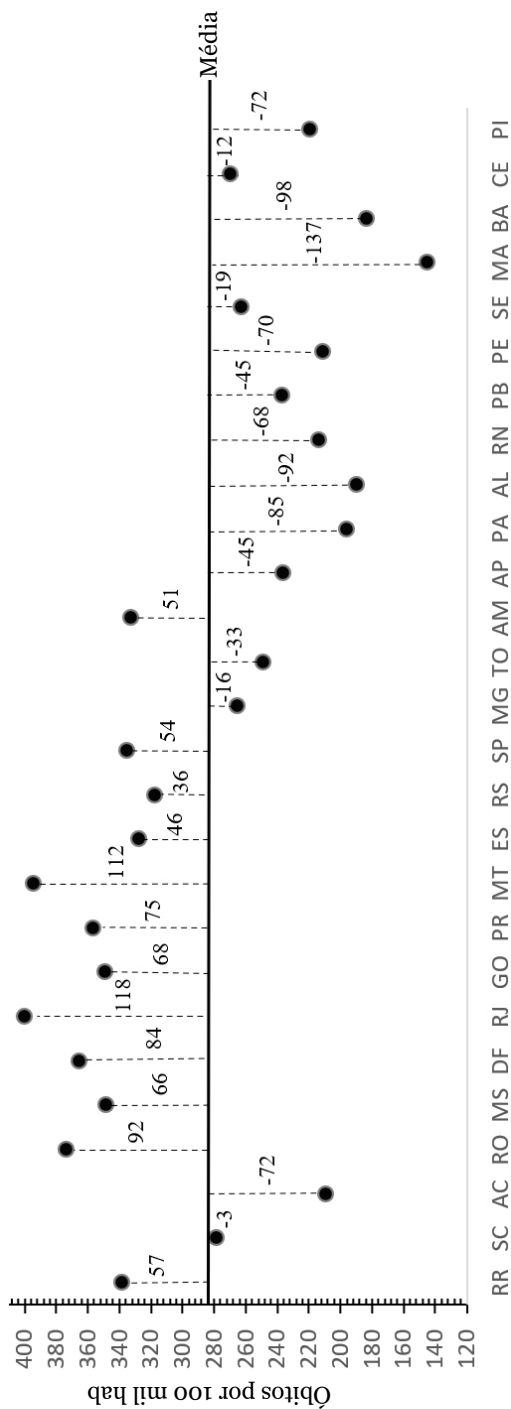


Gráfico 24: Desvio da média de óbitos por 100 mil Habitantes, por Estado, 02/12/21.

Tabela 15: Desvio padronizado da média de óbitos por 100 mil (03/12/21) e de votos (2018), e produto dos desvios por Estado.

ES	D.P. ¹ da M. de Óbitos p/ 100mil	D.P. ² da M. de Votos	D.P. ¹ X D.P. ²
RR	0,80	1,37	1,09
SC	-0,04	1,37	-0,05
AC	-1,02	1,23	-1,25
RO	1,29	1,23	1,58
MS	0,93	1,01	0,94
DF	1,18	0,94	1,11
RJ	1,65	0,87	1,44
GO	0,95	0,80	0,76
PR	1,05	0,80	0,84
MT	1,58	0,73	1,15
ES	0,65	0,59	0,38
RS	0,50	0,52	0,26
SP	0,75	0,44	0,33
MG	-0,23	0,09	-0,02
TO	-0,46	-0,05	0,02
AM	0,71	-0,05	-0,04
AP	-0,63	-0,19	0,12
PA	-1,20	-0,55	0,66
AL	-1,28	-0,76	0,98
PB	-0,96	-0,98	0,93
RN	-0,63	-0,98	0,61
PE	-0,98	-1,05	1,03
SE	-0,27	-1,26	0,34
MA	-1,91	-1,33	2,55
BA	-1,38	-1,47	2,03
CE	-0,17	-1,55	0,26
PI	-0,87	-1,76	1,53
Soma	0	0	19,61
R			0,73

Observa-se que a correlação aumentou nesse intervalo de sete meses de pandemia, quando as mortes chegaram há mais de 615.000 óbitos por COVID-19, consolidando a observação da tendência. Elevando-se ao quadrado esse valor de $R = 0,73$, encontra-se o coeficiente de determinação, uma estimativa em termos percentuais para avaliar o quanto uma variável pode prever a outra. Nessa edição desse livro, não foi possível explicar matematicamente esse procedimento, mas ele é bastante didático para compreender o tamanho do problema. Então:

$$R^2 = 0,73 \times 0,73 = 0,53$$

Isso significa que a variação dos votos em 2018 explica mais de 50% da variação dos óbitos por 100 mil habitantes em dezembro de 2021. Em outras palavras, isso significa uma responsabilidade muito grande da variável X em explicar a variável Y. É claro que não foi a materialidade de apertar um botão que provocou, proporcionalmente, mais mortes em uns Estados do que em outros. Mas esse gesto é uma sinalização da influência política que impacta nas ações daqueles que estão sujeitos à liderança da pessoa que foi eleita. Este fenômeno nos faz observar o impacto das ideias do líder nas grandes massas. No começo da pandemia, o movimento negacionista da ciência ganhou corpo e influenciou o comportamento daqueles que estão mais vulneráveis a tal discurso, que foi capitaneado pelo presidente Bolsonaro. Houve menosprezo do potencial da máscara em prevenir a transmissão, rebelou-se contra o distanciamento social e depois se criou uma aversão ao processo de vacinação. Esta correlação pode ser um indicativo de que essa influência foi mais efetiva nos lugares com maior influência do bolsonarismo, tornando a disseminação do vírus mais rápida e mais abrangente nesses contextos, impactando assim no número de mortes por 100 mil.

Uma forma tradicional de visualizar a correlação é o gráfico de dispersão, analisado a seguir no gráfico 25. Nele, é possível visualizar uma reta na diagonal que representa uma correlação perfeita. Os pontos são os resultados dos Estados em relação ao número de casos no eixo “Y” e ao número de óbitos no eixo “X”. Observa-se como os pontos tendem a se aglomerar em torno da reta representando uma correlação forte e intensa. Vê-se o Estado de São Paulo destacado no alto à direita do gráfico, tendo resultados altos nas duas variáveis, como esperado na tendência observada e confirmando a observação dos gráficos anteriores.

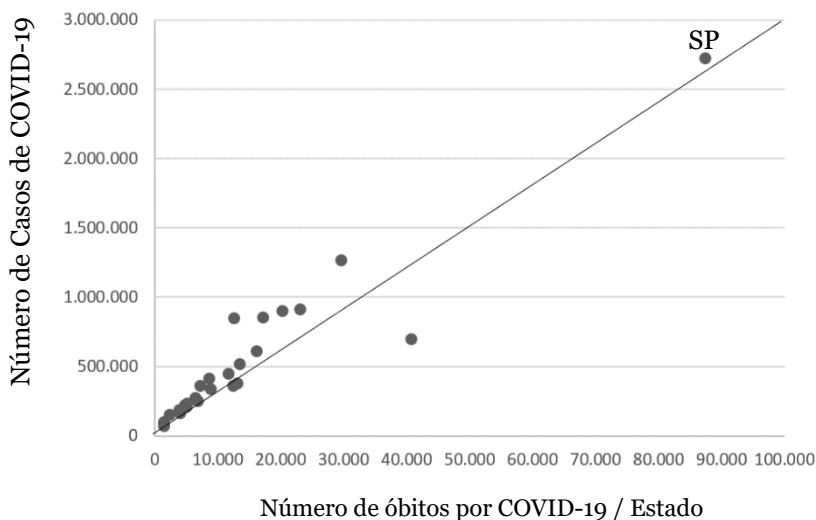


Gráfico 25: Correlação de 0,95 entre casos e óbitos por COVID, por Estado, em 16/04/2021.

Já no gráfico 26, a seguir, não se vê toda essa força observada no gráfico anterior, mas é possível perceber uma tendência moderada de aglomeração em torno da reta diagonal que representa uma correlação perfeita de 1. Alguns Estados estão mais próximos da reta, enquanto outros estão mais distantes.

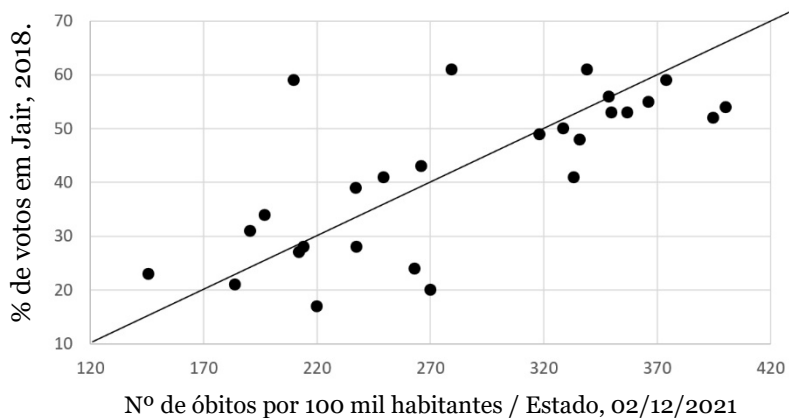


Gráfico 26: Correlação de 0,73 entre votos em 2018 e óbitos por COVID por 100 mil habitantes, por Estado, em 03/12/2021.

Esse retrato da correlação não é tão nítido quanto o da correlação anterior, mas ainda é uma representação clara de uma tendência que impactou o país e produziu um efeito significativo na distribuição das mortes. Seria muito difícil justificar esse resultado como fruto do acaso, pois padrões em grandes números normalmente se estabelecem de forma consistente e condicionado por fatores intimamente relacionados, inviabilizando a tese de aleatoriedade. Seria o mesmo que dizer que Deus jogou dados e a distribuição de mortes impactou mais nos territórios habitados por seres fanáticos por seu Messias. Seria uma provação para o povo de Deus? Uma limpeza dos mais fracos? Ou simplesmente irresponsabilidade da liderança que subestimou os dados científicos que previram a catástrofe? É possível questionar as interpretações e até mesmo vasculhar em busca de explicações mais plausíveis para esses números. Mas questionar a natureza dos dados é como enfiar a cabeça num buraco de avestruz para não ver a luz do sol. Teria havido erro nos dados do TRE? Mas aí Bolsonaro não deveria ter sido eleito, contrariando as pesquisas de opinião? Ou os dados oficiais de óbitos por 100 mil habitantes estão equivocados? Teria uma

conspiração de médicos, em todos os Estados, estabelecido um acordo para combinar a causa da morte como COVID-19 nos atestados de óbito com maior frequência em uns Estados do que em outros? Eles teriam calculado uma proporção adequada entre casos e óbitos para construir o fenômeno que estamos presenciando com o desenrolar da pandemia?

Talvez seja possível recorrer aos negacionistas norte-americanos, uma vez que um fenômeno parecido aconteceu no solo deles: a questão das medidas de prevenção à disseminação do vírus, como distanciamento social, se tornou uma questão partidária (ALLCOTT; BOXELL; CONWAY; GENTZKOW *et al.*, 2020; GAO ; RADFORD 2021). Os republicanos que apoiaram Trump tiveram uma tendência a negar o distanciamento social como medida de prevenção e nessas áreas houve maior taxa de morte e de infecção pelo vírus. (GOLLWITZER; MARTEL; BRADY; PÄRNAMETS *et al.*, 2020). Nessa pesquisa de Gollwitzer *et al.* (2020), foram estudados dados de georeferenciamento oriundos 15 milhões de *smartphones* por dia, aproximadamente. Nos condados onde prevaleceram os votos em Trump contra Hillary Clinton, em 2016, os celulares exibiram 14% menos distanciamento social entre março e maio de 2020. Essa diferença se revelou associada com maiores taxas de infecção e mortalidade por COVID-19.

Após perder as eleições, o ex-presidente Trump mudou alguma de suas opiniões e inclusive aderiu ao processo de vacinação. Isso é um sinal de que o negacionismo pode ceder diante das evidências dos dados científicos e do castigo das urnas nas eleições seguintes. Para infelicidade do Brasil, as eleições serão apenas em 2022 o que ainda dá algum poder aos negacionistas nacionais, atrasando a mudança de “*mindset*” (mentalidade). Enfim, o assunto começa a deslizar para o campo da psicologia e da política, quando os dados já não bastam por si mesmos para explicar a situação. Será necessário considerar esse fato no capítulo final, numa crônica das mortes anunciadas.

Crônica das Mortes Anunciadas

Não se pode basear nada na opinião: antes de tudo, é preciso destruí-la. Ela é o primeiro obstáculo a ser superado... O espírito científico proíbe que tenhamos uma opinião sobre questões que não compreendemos, sobre questões que não sabemos formular com clareza.

Gaston Bachelard (2003, p. 18).

A mente funciona através da construção de modelos, por meio dos quais é possível organizar a percepção da realidade, classificar coisas, categorizar acontecimentos, fazer previsões, tomar decisões e verificar a presença de algum erro. A palavra “cognição” define esse processo de avaliação no qual o sujeito analisa as pessoas e o ambiente para tomar decisões baseado em interpretações e conclusões próprias. A memória auxilia na comparação dos dados da realidade a fim de distinguir o falso do verdadeiro, verificando a autenticidade da informação para corrigir algum equívoco eventual nessa complexidade de operações.

Um dos tijolos básicos que sustentam a cognição é denominado de “esquemas”, que é um padrão elaborado na mente para atribuir características e classificar tudo ao redor da pessoa. O indivíduo sempre está processando dados da percepção através de esquemas, ou seja, computando alguma informação, classificando e localizando fenômenos em alguma categoria ou incluindo coisas dentro de algum esquema cognitivo. O processo de perceber e raciocinar pode, então, ser comparado ao funcionamento de um computador que conta números e transforma-os através de operações predeterminadas. Essa programação trabalha com códigos que convertem os *inputs* (entradas/insumos) em *outputs* (saídas/produtos). Em outras palavras (*in other words, please be true...*), dados brutos são tratados e tornam-se informações úteis. Isso pode ser uma boa

metáfora para os esquemas que são utilizados para fazer uma leitura da realidade através da assimilação do ambiente e, em seguida, execução das operações lógicas típicas da inteligência.

A programação da mente, porém, é mais dinâmica do que os computadores mais avançados, uma vez que os esquemas são desenvolvidos ao longo do tempo e em consonância com o desenvolvimento social mais amplo. As pessoas compartilham uma determinada realidade na medida em que os seus esquemas cognitivos modelam a percepção de maneira similar. A cultura é a expressão desses esquemas compartilhados que mudam com o tempo. Toda cultura tem uma dimensão mais estática e outra mais dinâmica. Quando a dimensão estática prevalece, aumentam as chances dos esquemas se tornarem estereótipos e do indivíduo agir automaticamente baseado em preconceitos. Na medida em que os preconceitos causam problemas sociais e emergem da inconsciência para as discussões e lutas políticas, mudanças podem levar à transformação da sociedade e, conseqüentemente, à revisão de antigos preconceitos consolidados em esquemas cognitivos antiquados e rígidos.

Apesar da mente gerar uma realidade que aparenta uma uniformidade e padronização, o mundo real é dinâmico e desafiador. A busca por modelos perfeitos para ver o que acontece fora do cérebro é uma tentação e uma compulsão ligada à necessidade de controlar o ambiente e manter uma sensação de segurança oriunda da previsibilidade. Entretanto, um modelo relativamente estável para lidar com o ambiente e a sociedade pode se tornar um obstáculo para se adaptar às mudanças inevitáveis. O fato é que os preconceitos são ferramentas para reconhecer o mundo, categorizar e classificar, mas também podem ser prisões que impedem a visão ampliada do real.

Os grandes problemas da filosofia dialogam com essa ideia de prisão inconsciente baseada em ilusões convincentes que impedem o próprio sujeito de perceber que está preso. Existe uma ancestral dúvida do ser humano em relação à sua

capacidade de ver o mundo com precisão e atestar a sua autenticidade ou realidade: nossos pensamentos refletem a realidade ou a inventam? Vemos o mundo ou uma cópia rabiscada dele? Até que ponto as imagens que carregamos no cérebro representam adequadamente a verdade dos objetos e pessoas que habitam fora dele? Existe realidade além do que é imaginado na mente do ser humano? Indagações como essas levaram ao jogo da caverna de Platão. Essa é uma antiga metáfora para o teste da realidade, que para muitos não tem uma solução final e condena o homem a nunca ter acesso à realidade como ela verdadeiramente é. Ou seja, a verdade seria sempre relativa à cela da prisão na qual o sujeito habita e de onde ele projeta o seu olhar através de um ângulo peculiar de invenção. No limitado pátio da realidade, onde os prisioneiros perambulam inocentes sem saber qual crime cometeram e nem quanto tempo vai durar a reclusão inconsciente e forçada, às vezes, eles compartilham o mesmo ângulo ou perspectiva da realidade. As pessoas geralmente não percebem que se encontram numa prisão psicológica e nem imaginam a gravidade do crime que cometeram para serem condenadas a essa reclusão. O fato de ter nascido é o crime capital que condena o ser humano ao aprisionamento dentro de seu próprio cérebro.

Na formação do organismo, a evolução dotou o indivíduo de mapas cognitivos para poder navegar de maneira eficiente na “prisão” da realidade. Os esquemas são imagens mentais que representam o que acontece fora do organismo e auxilia na sobrevivência do ser vivo, pois funciona como um mapa da realidade, fundamental para navegar pela “caverna-prisão”. Os mapas cognitivos são representações dinâmicas e resumidas da imensidão do universo que constrange a mente. Tendo em vista que o cérebro tem recursos de energia e de tempo limitados, para ele trabalhar com eficiência e viabilidade ele precisa resumir esse universo, destacando o que é mais relevante e descartando o que não é útil, em cada situação vivida. Esses critérios de utilidade e

relevância não são necessariamente universais e dependem da cultura, do indivíduo e da história: o que conta para uns, pode ser irrelevante para outros, o que aprisiona uns, liberta outros.

A introdução nesse assunto da cognição tem uma razão de ser: esse livro que começou como uma demonstração estatística, deslizou para os recantos da psicologia, pois a simples apresentação de dados numéricos se revelou um problema político que, para alguns, pode facilitar a compreensão da situação e, para outros, impossibilitar o diálogo. Quando os dados da COVID-19 trabalhados nesse livro foram apresentados a alguns simpatizantes de Bolsonaro, a reação de questionamento foi automática. Por outro lado, a aceitação foi espontânea e natural por aqueles que antipatizam com a figura do presidente. Os mapas cognitivos desses dois grupos são diferentes, conduzindo os sujeitos a percepções diametralmente opostas do mesmo dado estatístico, como se habitassem “cavernas de realidade” diferentes. E não adianta tentar convencer os prisioneiros acorrentados em cavernas do contrário de suas opiniões. Platão advertiu sobre o perigo de alguém afirmar que a realidade é mais ampla do que as paredes de pedra onde a luz é projetada para criar as imagens fictícias do mundo mental. Tal indivíduo “questionador” poderia ser enquadrado como louco e possivelmente ser condenado à “morte por incompreensão”.

Ao entrar na caverna, o professor se encontra com os seus maiores temores: existe didática suficiente para desenhar num quadro os argumentos necessários para a transformação de algumas ideias equivocadas de seus alunos? Quantos gráficos, quantas tabelas e quantos números são necessários para estabelecer uma luz sobre a realidade dos fatos? E quando os alunos não conhecem a matemática básica e duvidam dos números que o professor oferece para a análise? Com tantas dúvidas sobre a eficácia de suas próprias ferramentas, o professor começa a questionar a validade de suas ideias e a

indagar se não é ele mesmo que está preso na caverna-manicômio de um conto fantástico no estilo do alienista de Machado de Assis (ASSIS, 2014).

O fantasma da apofenia é uma preocupação constante para o professor-pesquisador, pois a mente é programada para encontrar padrões em dados que são aleatórios e mero efeito do acaso. O conceito de apofenia surgiu para descrever um sintoma de doença mental que leva o sujeito a tirar conclusões fantásticas e paranoicas a partir de dados que não confirmam adequadamente tais conclusões. A necessidade de dar sentido aos fenômenos que acoçam a mente pode precipitar o sujeito em teorias da conspiração ou ativar a imaginação fantástica que conecta “alhos com bugalhos”. Será que o professor foi sequestrado por uma rede conspiratória de “comunistas” que formata o pensamento da ciência moderna? No final das contas, o aluno começa a indagar se a terra é plana, pois a experiência dos próprios olhos não permite assegurar o modelo esférico e, talvez, o homem ainda não tenha ido à lua para ver como é lá de cima. Afinal de contas, será que tudo não passa de narrativas projetadas em cavernas platônicas?

Por esse ponto de vista, adulterar dados da COVID-19 parece ser o menor dos problemas diante desse retorno a ideias da idade média, a um mundo tradicional em que aparentemente se vivia melhor sem essa ciência enganadora que não oferece “certezas absolutas e finais”. Além disso, existe um obstáculo bíblico (Matheus 7:3) para reconhecer falhas nas próprias teorias adotadas: é mais fácil ver um cisco no olho do vizinho, do que um vergalhão no próprio olho. Quando alguém calcula qualquer coisa está sujeito à erro. Na Estatística é estabelecido a ideia de Erro do tipo I: dizer que existe uma correlação, por exemplo, quando na verdade os números apenas revelam uma manifestação do acaso. O que salva a ciência é a capacidade de submeter os achados das pesquisas ao escrutínio dos pares, que

refazem os cálculos e as experiências para, gradativamente, validar ou refutar os achados originais.

Esse processo de construção coletiva do conhecimento é complexo e avança de acordo com o esforço da comunidade científica. Porém, mesmo sem fundamentação científica, com o advento da sociedade midiática, desenvolveu-se um pessimismo à direita e à esquerda, a ponto de muitos pregarem esse retorno a um ideal perdido de um passado nebuloso. Pinker (2018) propõe que não há fundamentação em dados científicos para esse pessimismo e propõe lançar nova luz sobre o obscurantismo contemporâneo. O iluminismo surgiu como reação a essa tendência negacionista da ciência, na idade média. Agora, nos vemos novamente desafiados a fortalecer a razão e a ciência frente a uma geração digital que dissemina com rapidez os acontecimentos midiáticos baseados em cálculos pessoais, que levam em conta, predominantemente, o retorno de exposição nas mídias e ganhos financeiros. Já as informações fundamentadas na comunidade científica, apesar de serem mais lentas, envolvem procedimentos e cálculos que podem ser replicados em qualquer parte do planeta e assim avaliadas externamente com maior precisão e controle.

Intuitivamente, todo ser vivo faz “contas”, cálculos e “estimativas”. Um predador avalia o custo-benefício de perseguir uma determinada presa a partir de suas expectativas de sucesso. No reino animal isso é feito de maneira instintiva e automática através de sequências de estimulações bioquímicas que evoluíram para tornar o animal eficiente. Já os seres humanos desenvolveram, além disso, esse sofisticado equipamento de análise que é a mente, capaz de fazer contas complexas e realizar grandes empreendimentos com a coordenação de inúmeros indivíduos organizados em sociedade, numa comunidade aberta ao diálogo e à contestação racionalmente fundamentada.

O ato de calcular demanda todo um aparato de pressupostos que muitas vezes o próprio indivíduo não percebe.

São teorias e modelos ideais vividos como verdades absolutas que condicionam a forma de ver os fenômenos ao redor, mas que não habilitam a enxergar as falhas do próprio conhecimento. A tendência de encontrar um erro mínimo nos olhos de outra pessoa, um cisco, e não conseguir enxergar o elefante que sustenta a trave nos próprios olhos foi aqui denominado de “obstáculo bíblico”. No contexto religioso, é a humildade que faz a correção desse olhar arrogante. Em outras palavras, no contexto da filosofia da ciência, o “obstáculo epistemológico” (BACHELARD, 2003) é contornado com o apoio do olhar atencioso de um outro que possui uma perspectiva diferente e nos auxilia a visualizar as falhas no processo de construção do conhecimento. Ouvir o outro falar sobre a “trave” que está no próprio olho é um ato de humildade e de amadurecimento para superar o egocentrismo que o ser humano carrega dentro de si e assim relativizar os modelos ideais que constroem a inteligência coletiva.

Os modelos idealizados na mente de cada um auxiliam no processo cognitivo de comparar os dados que são lançados na consciência de uma pessoa. Um dos procedimentos utilizados pela mente para gerenciar a vida é a comparação com modelos ideais. Mesmo não existindo no plano material, os ideais auxiliam na medição e controle dos fenômenos. Para comparar qualquer coisa é preciso ter a competência de identificar os desvios em relação a algum modelo. Quando o desvio é grande, isso chama a atenção da pessoa, se ela foi treinada para prestar atenção aos sinais da diferença. Mais uma vez, é o contexto cultural que vai definir os critérios de inclusão ou exclusão de processos classificatórios ou eliminatórios. Por exemplo, até que ponto uma pessoa é incluída em uma categoria ou em outra? A partir de que ponto, uma pessoa se torna um desviante do modelo hegemônico em determinado período da existência de alguma sociedade? O obstáculo epistemológico se levanta quando tais modelos ideais são tomados como a própria

realidade e não um simples conceito para organizar a experiência e fazer as comparações a partir de convenções construídas coletivamente.

É de se imaginar que algum leitor se dirija para a porta do PROCOM para dar uma queixa, baseado no direito do consumidor, a respeito desse livro que prometeu explicações estatísticas e está entregando, numa venda casada, elocubrações psicológicas. Antes que ele adentre nos meandros da burocracia sem fim e esse processo se alongue indefinidamente nos corredores dos tribunais da justiça comum, eu tentarei justificar esse processo arrastado como forma de evitar uma condenação prévia: será ele inocente ou culpado? (KAFKA, 2003)

Essa guinada no tema do livro se deve a uma razão elementar: não adiantou utilizar da lógica estatística para alcançar a compreensão dos simpatizantes do atual governo para demonstrar as evidências de um crime que produz morte em série com maior intensidade em determinadas regiões do país. Esse é um fenômeno psicológico de cognição social e de liderança que será estudado por muito tempo, mesmo após a mitigação da pandemia.

A explicação mais sensata é que nas regiões de maior influência de um líder, as pessoas tendem a seguir os seus ideias e exemplos. Não importa se tais exemplos sejam bons ou ruins, sejam válidos ou ilegítimos, se são baseados na ciência ou em alguma seita. Alguns líderes conseguem captar tendências e expressá-las com maestria, sendo assim seguidos nas condições mais improváveis. O poder não está só no líder, mas também na cultura que o cerca. Um líder é um cara que resolve problemas em um determinado contexto cultural (SCHEIN, 2009) e as pessoas adotam sua maneira de solucionar “quebra-cabeças” como exemplo a ser copiado. Édipo, por exemplo, resolveu a charada da Esfinge e se tornou um curioso caso de Rei que aparentemente desconhecia sua origem. Durante um tempo, Tebas prosperou e se livrou de sua praga, mas não para sempre,

pois o fantasma do assassino incestuoso rondava a cidade. Quando a solução proposta não funciona mais, é possível haver uma ruptura e abertura para a mudança do líder. Mas o processo de mudança cultural é complexo e as vezes demorado.

Para Freud, Édipo era o sujeito que nada sabia. Mas para Foucault (2002), ele é o sujeito que tudo sabe e que tudo pode, o modelo do tirano tão combatido na Grécia clássica, com o advento da democracia que sobrepujou a Grécia arcaica. A partir do momento em que o povo, representado pelo camponês e pelo servo, apresenta o seu testemunho, o déspota que acusava a todos de seu próprio crime é condenado a cumprir a pena que ele mesmo tinha estipulado.

Não é fácil classificar Foucault à esquerda ou à direita: ele é escorregadio e esquivo. Ao descrever o lugar que o tirano ocupou nesse contexto, ele também faz uma crítica à idealização do filósofo enquanto salvador e detentor do saber apartado do poder, relegando o político ao papel daquele que apenas tem interesses no poder desprovido de qualquer saber, como se o primeiro tivesse uma verdade sem poder e o segundo, um poder sem verdade nenhuma. Ao contrário disso, não seria possível essa tal coisa de um saber desinteressado em batalhas linguísticas. Para aqueles que admiram Foucault à esquerda, é extremamente desconfortante os representantes da direita usarem estratégias discursivas ensinadas por ele, se estabelecendo assim no centro dessa “guerra de narrativas”. Enfim, a tirania não desaparece com a morte do tirano, mas apenas ocupa novas formas de controle social oriundas das práticas estabelecidas numa sociedade.

Muitos adeptos que votaram no atual governo ainda não se convenceram dos erros cometidos que impactaram no atual quadro de calamidade pública em diversas áreas da administração pública, uma vez que se trata de um “mito”. Já os que estão contrários, demonizam a figura desse líder, pintando o seu retrato como um maníaco psicopata que não possui empatia

com o sofrimento alheio, tendo em vista que se constitui como o clássico tirano. Essa explicação pela patologia é extremamente sedutora para todos os lados: para a esquerda tem-se uma explicação lógica e catártica para o que é visto e ouvido, enquanto que para a direita há uma sólida justificativa para se livrar de qualquer vestígio de culpa, se o mandato presidencial acabar tragicamente: a culpa era dele que estava lá no poder, ou melhor, nem o presidente é culpado, pois ele estava doente.

O humano é equipado com essa capacidade de organizar o mundo dentro de “caixinhas”, estabelecendo critérios do que é saudável e o que é patológico, normal e anormal. Tais critérios variam de acordo com as épocas e costumes. Muitas vezes, o normal se confunde com a média, criando uma aliança entre os estatísticos e os administradores, entre os quais foi forjado um pacto para buscar controlar aqueles que se encontram à margem e que relutam em entrar no rol dos normais. Os marginais dão um trabalho fenomenal para serem enquadrados a fim de não perturbarem a tranquilidade daqueles que gentilmente habitam na média, nas praias da normalidade e na monótona paisagem do deserto ideal. (Tendo em vista que não há ser orgânico vivente perfeitamente ajustado ao modelo ideal e que na perfeição não está incluído nenhum ser mortal, tal lugar deve ser, portanto, um deserto!)

Mesmo não tocando na perfeição, a humanidade desenvolveu técnicas para se aproximar dela. Avaliando a distância em relação ao modelo, é possível realizar prodígios, como por exemplo, medir o quanto um indivíduo está distante de um ponto central em relação a alguma qualidade qualquer. O quanto o leitor está distante da média? Essa é uma pergunta que envolve autoconhecimento e conhecimento dos outros com os quais se compara. O caminho do Delfos que leva a conhecer a si mesmo, entretanto, é tortuoso e cheio de armadilhas.

Cada um é carregado com uma boa dose de tendenciosidade, que pode ser definida como a propensão a

cometer equívocos em benefício próprio, utilizando os esquemas de maneira a favorecer alguma vantagem pessoal. É da natureza humana julgar a partir de seus próprios esquemas, preconceitos e estereótipos. Isso tem relação com a teoria da evolução das espécies, na qual o indivíduo busca vantagem para si na competição pela sobrevivência. Quando essa tendência se torna crônica, a competição pode dificultar a cooperação, prejudicando os alicerces da vida em sociedade. Essa ideia darwinista (Darwin é um dos principais autores da teoria da evolução) foi a base do darwinismo social que justificou a lei dos mais fortes e legitimou a desigualdade através do argumento de que os mais aptos são mais bem sucedidos na sociedade e merecem ser mais ricos. Ironicamente, o matemático, estatístico e fundador da eugenia Francis Galton que pregava essa ideia se revelou estéril e não passou adiante os seus valorizados genes que lhe conferiam QI muito superior à média:

É interessante o fato de Galton, que fundou a eugenia e acreditava que só as pessoas muito inteligentes deviam se reproduzir, não ter tido filhos. Ao que parece, o problema era genético; nenhum dos seus irmãos os teve. (SCHULTZ; SCHULTZ, 2004)

O normal e o patológico foi tema do professor de Foucault, Caguilhem(2002). As sociedades modernas se revelaram uma indústria de normalização dos indivíduos, atribuindo a pessoa certa para o lugar certo, de acordo com a habilidade e a inteligência de cada um, e obtendo o máximo de rendimento que cada indivíduo pode entregar. A filosofia levantou diversas críticas ao poder central que coage os indivíduos a viverem papéis predeterminados estrategicamente organizados, para drenar as energias coletivas, a fim de aumentar a produção de riquezas e a concentração do poder numa dada minoria. Isso é um traço de psicopata admirado em larga escala. Aplicando isso à questão da saúde mental, pode-se indagar: e se

muitas pessoas pensam de acordo com um psicopata, isso não se torna uma forma normal de ver a vida? A estatística pode ser uma tirana bem-posicionada nessa hora, legitimando o comportamento perverso através de uma média que normaliza a ação violenta que não se preocupa com os mais “fracos”.

Outra complicação prática em classificar o tirano-mito como psicopata é relativa ao impacto social disso: se ele é apoiado por parcela significativa da sociedade, esta parcela estaria toda doente? Que poder teria o professor para julgar toda essa gente como louca? Novamente o espírito de Simão Bacamarte espreita, atento, aguardando o momento de agir. Para não correr o risco de ficar preso no manicômio construído com as próprias mãos, a alternativa é responsabilizar os indivíduos pelas suas ações e fazer com que eles arquem com as consequências. Essa é, então, uma questão de justiça e não de medicina. Mas a justiça também é uma questão de poder. A sociedade brasileira, de certa maneira, está de mãos atadas se as autoridades não enquadrarem o delito cometido dentro das penalidades previstas.

Ao ler Foucault(2002) de forma mais cuidadosa, é difícil determinar a sua lateralidade, se ele tende mais ao conservadorismo ou ao progresso. Por ser escorregadio, o filósofo muitas vezes evita rótulos que comprometam a sua flexibilidade de pensamento e imobilize a sua inteligência. Ele ensina que na história das formas de resolver litígios, a sociedade em diferentes contextos lançou mão de muitos artifícios singulares para chegar a um veredito numa batalha judicial. Seria o caso de estabelecer um inquérito ou uma inquisição para caçar o bruxo? Ou estabelecer uma banca examinadora para avaliar a periculosidade do sujeito e sua sanidade mental? Na dúvida, o escritor começa a refletir sobre o poder que está buscando com esse dispositivo estatístico que condena a partir de análises de correlação: o que isso significa? Qual poder o pesquisador está buscando ao usar os números como arma de estabelecimento de uma verdade? A resposta vem com o romantismo que é próprio

dos amantes das letras: ele está buscando o poder que tem Zeca Chapéu Grande, aqui possuído por Santa Bárbara:

E não foi com espanto que vi naquela mesma noite, antes de todos os outros encantados chegarem e se abrigarem no seu corpo, santa Barbara girar, gritar e parar com sua espada apontada para o prefeito, a quem fez honras, como se cumprimentasse um monarca, mas também como se se dirigisse a um súdito, para lhe pedir, na frente da audiência, que cumprisse a promessa feita no passado – e que não recorde de sabermos – de construir uma escola para os trabalhadores. O prefeito olhou desconcertado, esboçando um sorriso sem graça, quando se viu diante do olhar das quarenta famílias que moravam em Água Negra. Quase compassivo, recordando das graças e temendo a má sorte que teria, dependendo do esforço empreendido para realizar a ordem dada pela encantada, aquiesceu. (VIEIRA JUNIOR, 2019)

Enquanto as professoras não detêm esse poder transcendente da espada de Santa Bárbara para submeterem e coagirem as autoridades a investirem em educação e assim habilitar a população a entender das letras e dos números, nos resta construir mais um documento que fique para a posteridade como registro de fenômenos danados que aconteceram na face da terra. A seguir, o veredito provisório desse processo inglório.

Tecnicamente, um trabalho estatístico se conclui ao corroborar ou rejeitar a hipótese nula de trabalho. A seguir, reproduzimos novamente a hipótese nula (H_0) e a hipótese experimental (H_1) objeto da análise:

H_0 : “É improvável existir alguma relação entre óbitos por COVID-19 em 2021 e a porcentagem de votos em Bolsonaro em 2018”

H_1 : “É provável que as diferenças entre os números de óbitos por COVID-19 por Estados e as diferentes porcentagens da votação de 2018 em Bolsonaro não sejam mero fruto do acaso e existe uma associação entre esses dois fenômenos separados por 3 anos de diferença”

Sumariamente, a Estatística rejeita a hipótese nula e corrobora a hipótese experimental, ou seja, foi encontrada uma associação moderada entre votos em Bolsonaro em 2018 e mortes por 100 mil habitantes em 2021, por Estado brasileiro. Essa associação foi mensurada através da comparação com um modelo baseado na simples média aritmética e que produz um indicador de correlação.

O modelo baseado na média é uma espécie de ideal entre outros. O ideal mais desejável seria o menor número possível de mortes ocasionadas pelo coronavírus. Na dificuldade de estimar qual seria esse número ideal em diferentes cenários possíveis, a média observada se torna um valor de referência para comparar as diferentes manifestações desse fenômeno da pandemia, tendo em vista a complexidade de encontrar um modelo perfeito de análise. Se todos adotassem medidas preventivas para não disseminar o vírus, o modelo ideal das mortes por COVID-19 no Brasil tenderia a zero. Esse é um ideal utópico demais devido à dificuldade de ser implantado. Não tendo sido tomadas tais medidas, resta adotar como parâmetro o que efetivamente ocorreu e estabelecer algumas constatações:

1. Quanto maior a porcentagem de votos em Bolsonaro, em 2018, maior tende a ser a taxa de mortes por COVID-19 por 100 mil habitantes por Estado em 2021;
2. Essas diferenças nas taxas de mortalidade podem ser explicadas pela não adoção de medidas preventivas pelos seguidores de Bolsonaro que copiaram os exemplos do líder e negaram a ciência;

Felizmente, não é papel do psicólogo, nem do professor e nem do cientista julgar e dar o veredicto final. Isso pode ser feito apenas pelo povo nas urnas da próxima eleição, na qual os votos podem se livrar dessa carga amaldiçoada dos óbitos. Parte do problema do convencimento da população para votar em outro candidato que não esteja associado com a morte é devido ao arrebatamento da contemplação das estrelas e rejeição dos gráficos do astrônomo erudito. Em quem será que o poeta Walt Witman votará? Eu compartilho com ele esse romantismo do céu estrelado, que é bem mais convidativo e reconfortante do que as palestras matemáticas ou os 26 gráficos desse livro. Mas preciso lhe dizer, Walt, que as estrelas são vistas com o coração, mas os buracos negros só são perceptíveis com a matemática.

O poeta nos faz viajar pela beleza das letras do cosmos e o cientista propõe a viagem através do buraco de minhoca. O poeta e o cientista viajam por rotas paralelas pelo universo e em algum ponto do infinito eles se tocam e ouvem juntos a harmonia celestial. Olhando só para as estrelas, o cidadão está sujeito a topar em algumas pedras no plano terreno ou pisar em merda. Porém, olhando só para o chão, o eleitor fica exausto e deprimido com a densidade “dos números da terra” sob seus pés. Se o poeta faz sonhar, o cientista evita as pedras no caminho, com seus cálculos, gráficos e figuras. Não desista da matemática do astrônomo entendido, Walt, haveremos de encontrar alguma associação entre as rotas paralelas do entendimento humano. Entre letras e números, o conhecimento há de florir e nos trazer o mel da sabedoria e da prosperidade.

Referências

ALLCOTT, H.; BOXELL, L.; CONWAY, J.; GENTZKOW, M. *et al.* Polarization and public health: Partisan differences in social distancing during the coronavirus pandemic. **Journal of Public Economics**, 191, 2020.

ASSIS, M. D. O alienista. São Paulo: Companhia das Letras 2014.

BACHELARD, G. **A formação do espírito científico - contribuição para uma psicanálise do conhecimento**. Rio de Janeiro: Contraponto, 2003.

BATESON, G. **Steps to an Ecology of mind**. Chicago: The University of Chicago Press, 1999.

BECKER, H. S. Notes on the concept of commitment. **The American journal of Sociology**, 66, n. 1, p. 32-40, 1960.

BORGES, J. L. **O livro dos seres imaginários**. São Paulo: Globo, 2000.

CANGUILHEM, G. **O normal e o patológico**. Rio de Janeiro: Forense Universitária, 2002.

CUNHA, E. D. **Os Sertões: campanha de Canudos**. 39 ed. Rio de Janeiro: Publifolha, 2000.

DE SAINT-EXUPÉRY, A. **O pequeno príncipe**. Rio de Janeiro: Agir, 1994.

DURKHEIM, E. **O suicídio: estudo sociológico**. São Paulo: Editorial Presença, 1982.

FIELD, A. **Descobrendo a estatística usando o SPSS**. 3 ed. Porto Alegre: Artmed, 2009.

FIELD, A. **Discovering statistics using SPSS**. 5 ed. London: Sage, 2013.

FOUCAULT, M. **A verdade e as formas jurídicas**. Rio de Janeiro: NAU, 2002. 85859336487.

GAO , J.; RADFORD , B. J. Death by political party: The relationship between COVID-19 deaths and political party affiliation in the United States. **World Med.&Health Policy**, 13, n. 2, p. 224 – 249, 2021.

GOETHE, J. W. **Os sofrimentos do jovem Werther**. Porto Alegre: L&PM, 2010.

GOLLWITZER, A.; MARTEL, C.; BRADY, W. J.; PÄRNAMETS, P. *et al*. Partisan differences in physical distancing are linked to health outcomes during the COVID-19 pandemic. **Nat Hum Behav**, 4, p. 1186–1197, 2020.

KAFKA, F. **O processo**. São Paulo: Folha de São Paulo., 2003.

MERLLIÉ, D. Suicídios: modos de registros. *In*: BESSON, J.-L. (Ed.). **A ilusão das estatísticas**. São Paulo: UNESP, 1995. p. 113-134.

MORAES, P. W. T. **Psiquê e as cartas mágicas**. Salvador: DaIN, 2020.

PINKER, S. **now: the case for reason, science, humanism, and progress**. New York: Penguin Books Limited/Viking, 2018.

SCHEIN, E. H. **Cultura organizacional e liderança**. São Paulo: Atlas, 2009. 978-85-224-5497-6.

SCHULTZ, D.; SCHULTZ, S. **História da psicologia moderna**. São Paulo: Cultrix, 2004.

SILVA, S. C. D.; MONTEIRO, W. D. **Levantamento do perfil antropométrico da população brasileira usuária do transporte aéreo nacional: projeto conhecido**. Agência Nacional de Aviação Civil, 2009.

VIEIRA JUNIOR, I. **Torto arado**. São Paulo: Todavia, 2019.

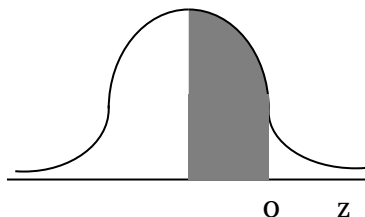
WHITMAN, W. *Leaves of grass*. New York: Modern Library 1993.

P. S.: O esforço solitário para escrever esse livro impossibilitou a ampliação do estudo para os municípios do país. Outro ponto a ser explorado é o fato de que o Brasil tem cerca de 3% da população mundial e em torno de 11% dos casos de morte por COVID-19. Só com esse dado é possível escrever outro livro para tentar compreender as razões que levam um país a ficar tão acima da média enquanto outros permanecem abaixo. Equipes de cientistas completarão esse trabalho de analisar os grandes números a partir de análises multivariadas e destruir os frágeis argumentos das mentiras que são proferidas de maneira irresponsável, sorrateira e criminosas.

P. W. T. M.

Anexo: Tabela Z da Distribuição Normal

Probabilidade da área entre “o” e “z”.



Z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2257	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2517	0,2549
0,7	0,2580	0,2611	0,2642	0,2673	0,2704	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2995	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1,2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4306	0,4319
1,5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
1,6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1,7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633

1,8	0,4641	0,4649	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4699	0,4706
1,9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4761	0,4767
2,0	0,4772	0,4778	0,4783	0,4788	0,4793	0,4798	0,4803	0,4808	0,4812	0,4817
2,1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
2,2	0,4861	0,4864	0,4868	0,4871	0,4875	0,4878	0,4881	0,4884	0,4887	0,4890
Z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
2,3	0,4893	0,4896	0,4898	0,4901	0,4904	0,4906	0,4909	0,4911	0,4913	0,4916
2,4	0,4918	0,4920	0,4922	0,4925	0,4927	0,4929	0,4931	0,4932	0,4934	0,4936
2,5	0,4938	0,4940	0,4941	0,4943	0,4945	0,4946	0,4948	0,4949	0,4951	0,4952
2,6	0,4953	0,4955	0,4956	0,4957	0,4959	0,4960	0,4961	0,4962	0,4963	0,4964
2,7	0,4965	0,4966	0,4967	0,4968	0,4969	0,4970	0,4971	0,4972	0,4973	0,4974
2,8	0,4974	0,4975	0,4976	0,4977	0,4977	0,4978	0,4979	0,4979	0,4980	0,4981
2,9	0,4981	0,4982	0,4982	0,4983	0,4984	0,4984	0,4985	0,4985	0,4986	0,4986
3,0	0,4987	0,4987	0,4987	0,4988	0,4988	0,4989	0,4989	0,4989	0,4990	0,4990
3,1	0,4990	0,4991	0,4991	0,4991	0,4992	0,4992	0,4992	0,4992	0,4993	0,4993
3,2	0,4993	0,4993	0,4994	0,4994	0,4994	0,4994	0,4994	0,4995	0,4995	0,4995
3,3	0,4995	0,4995	0,4995	0,4996	0,4996	0,4996	0,4996	0,4996	0,4996	0,4997
3,4	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4998
3,5	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998
3,6	0,4998	0,4998	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3,7	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3,8	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3,9	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000