

ENCE
Bacharelado em Estatística
2024.1

Usos e abusos da Estatística

Trabalho realizado em grupo em atendimento à disciplina Introdução à Estatística - EST011

Grupo:

Anna Maria Scofano

Bernardo Monteiro Rocha

Viviane Bastos de Oliveira Moura

Fernando Chahine Barros

Rodrigo Ferreira Freire de Souza

Luiz Rodrigues Nunes Junior

Abril / 2024

Introdução

O presente trabalho é apresentado como requisito da matéria Introdução à Estatística.

Anna Maria Scofano (202410091-03);

Bernardo Monteiro Rocha (202410221-26)

Viviane Bastos de Oliveira Moura (202411001-00)

Fernando Chahine Barros (202410631-59)

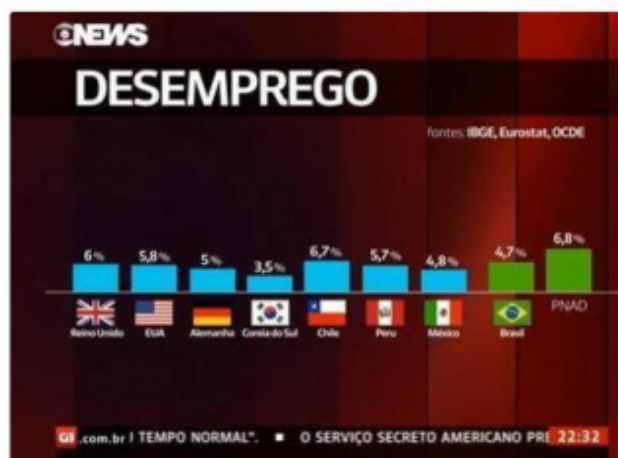
Rodrigo Ferreira Freire de Souza (202410031-72)

Luiz Rodrigues Nunes Junior (202411091-87)

1. Análise crítica

a) Os dois gráficos a seguir foram utilizados na reportagem da Globo News. Você acha que a informação está mostrada de forma adequada? O que mais chama sua atenção nos gráficos?

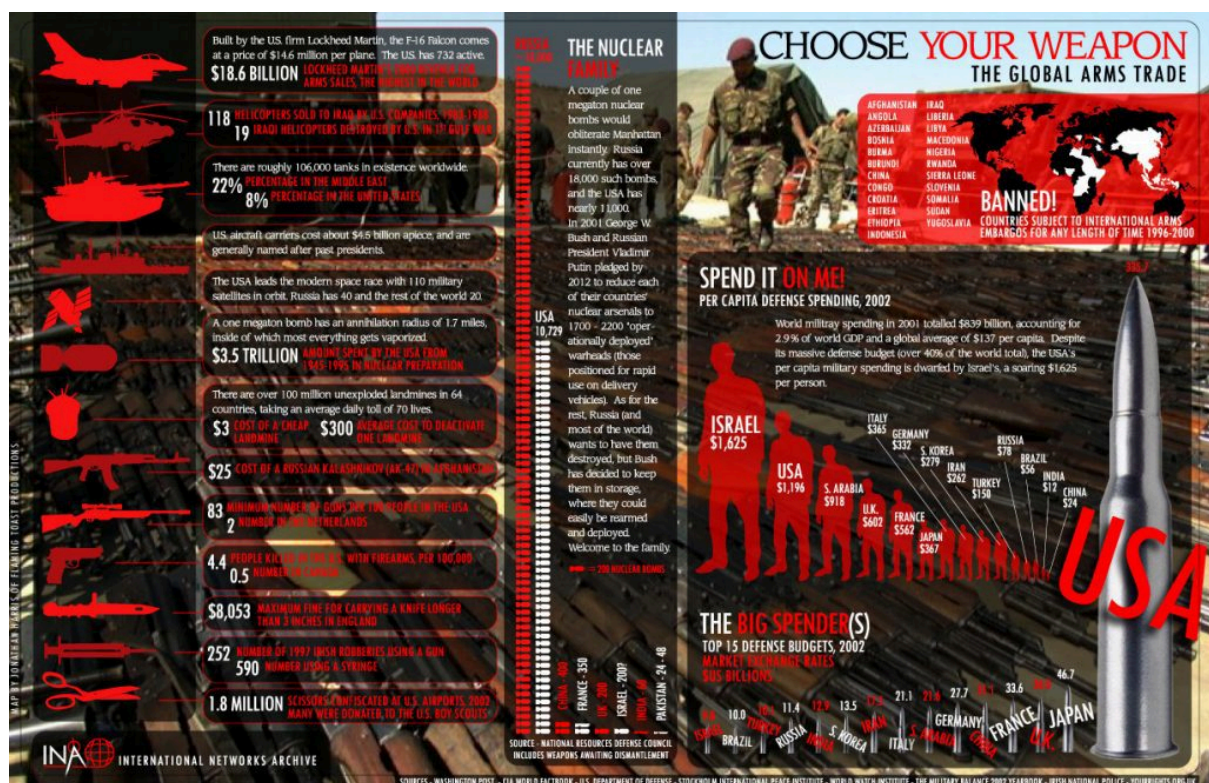
Faça uma análise crítica



Primeiro gráfico: É um gráfico apresentado em um programa jornalístico da empresa Globo mostrando a taxa de desemprego em alguns países. Nota-se que a barra correspondente ao Brasil é maior na sua altura em relação a países com um desemprego maior, como por exemplo, o Chile, mesmo com o número sendo exibido acima de cada barra. A escolha pela produção de colocar a barra referente à taxa de desemprego maior do que apareceria caso tivessem respeitado a escala das barras, pode indicar interesse em manipular a atenção do espectador levando-o a acreditar que a situação do desemprego no país está pior do que nos demais apresentados no gráfico. Outro ponto que chama a atenção é a taxa de desemprego apontada pela PNAD (Pesquisa Nacional de Amostras de Domicílio) que numericamente é próxima à taxa chilena e, ainda assim, apresenta visivelmente a intenção de mostrar uma taxa maior. Este fato corrobora que não houve exatamente um erro na elaboração do gráfico, mas sim, uma tentativa de desvio de atenção do espectador.

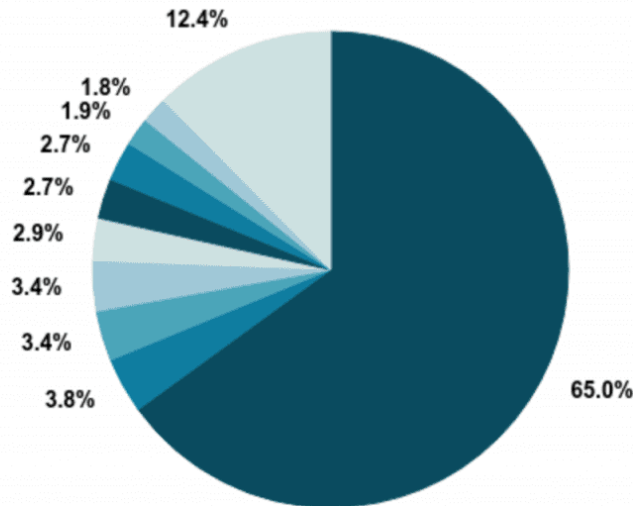
Segundo gráfico: É outro gráfico apresentado em um programa jornalístico da empresa Globo, no qual a inflação do país no ano de 2013 é apresentada com um tamanho da barra maior que em outros anos em que a inflação estava maior. A escala foi ignorada pela produção dando uma falsa ideia que a situação em 2013 estava pior do que em anos anteriores.

1b. Encontre pelo menos 02 gráficos extraídos de Jornais, Revistas, Internet, etc, que você acredita contenham uma quantidade demasiada de adornos desnecessários que possam mascarar a mensagem proporcionada pelos dados. Você deverá indicar a Fonte de Dados.



Foreign investment in China (2018)

■ Hong Kong ■ Singapore ■ British Virgin Islands ■ South Korea
■ Cayman Islands ■ Japan ■ Germany ■ USA
■ UK ■ Others



Source: Invest in China

BBC

Primeiro gráfico: Há uma quantidade exagerada de texto, as partes mais importantes poderiam estar destacadas em um lado e a parte mais detalhada em outro como um texto complementar. Não há muito contraste da fonte vermelha com o fundo do infográfico tornando a leitura difícil. O uso de figuras humanas para identificar o gasto per capita por cada país na defesa não foi eficiente em demonstrar a diferença entre esses gastos, pois cada figura tem uma largura diferente.

Fonte: <https://www.princeton.edu/~ina/infographics/>

Segundo gráfico: Um gráfico de setores com mais de 5 setores se torna difícil para diferenciar cada um deles, a escolha de cores com pouco contraste não ajuda a diferenciar cada parte fazendo com que a pessoa tenha que retornar para a legenda para entender qual porcentagem pertence a determinado país, o que torna a leitura dos dados lenta e difícil, indo de encontro ao objetivo de um gráfico. Um gráfico de Pareto seria mais objetivo e fácil de entender.

Fonte:

<https://www.reddit.com/media?url=https%3A%2F%2Fi.redd.it%2Fat03gpevcd45l.png>

1c. Apresente também 02 gráficos que apresentem problemas: explicita os problemas identificados.

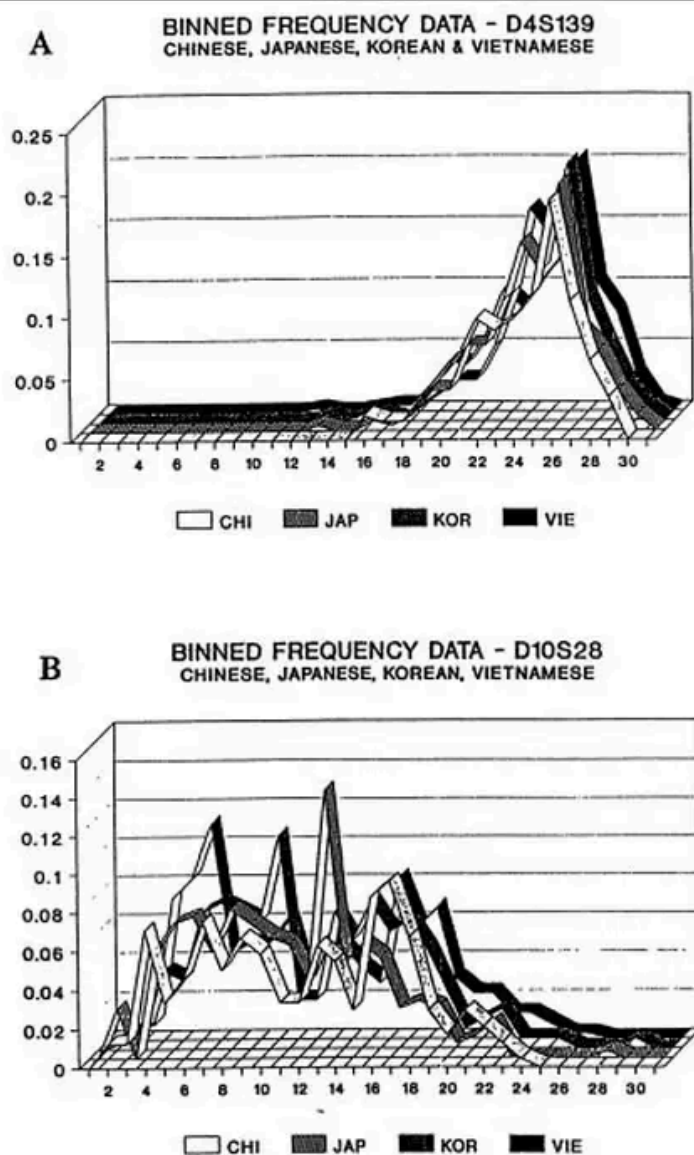


FIG. 4. Fixed bin distribution (histogram) for two loci and four Asian subpopulations (used with permission from John Hartmann): the boundaries of the 30 bins (vertical axis) are determined by the FBI; these bins are not of equal length. Sample sizes (numbers of individuals) for Chinese, Japanese, Korean and Vietnamese are 103, 125, 93 and 215 for D4S139 and 120, 137, 100 and 193 for D10S28. The horizontal axis is the bin number; bins are not of equal length.

Primeiro gráfico: Estes gráficos foram extraídos de um artigo publicado em 1994 pela Revista *Statistical Science*, no qual discute-se o uso de material genético (DNA) como prova em casos criminais como estupro e assassinato desde meados da década de 1980. Os gráficos apresentam a distribuição de probabilidade dos alelos (formas alternativas de um determinado gene que ocupam a mesma posição cromossomos homólogos) nas posições D4S139 e D10S28, representadas para as quatro subpopulações de asiáticos. De uma forma geral, gráficos tridimensionais tornam sua interpretação mais trabalhosa para o cérebro, é uma complexidade desnecessária, fica difícil relacionar as quantidades aos países correspondentes. Neste exemplo, um gráfico de linhas bidimensionais seria uma

opção mais adequada para que sejam percebidas as diferenças entre as populações e as quantidades exatas. Outra observação refere-se ao uso de padrões gráficos na legenda para diferenciação das subpopulações que não são visíveis na escala da publicação. A diferenciação por cores seria mais adequada, porém, há de se considerar os padrões editoriais das revistas científicas pois nem todas admitem outros padrões de cores além do preto e branco, pelo menos àquela época (1994).

Fonte: Roeder, Kathryn. "DNA Fingerprinting: A Review of the Controversy." Statistical Science, vol. 9, no. 2, 1994, pp. 222–47. JSTOR, <http://www.jstor.org/stable/2246327>. Accessed 29 Mar. 2024.



Segundo gráfico: O gráfico acima foi divulgado pela campanha eleitoral do PSDB para governador do Estado de São Paulo em 2018. Apesar de ter-se baseado numa pesquisa encomendada pela Band ao IBOPE em maio de 2018, o gráfico produzido com os dados gerados apresenta uma série de distorções: a) a barra com os 22% de intenções de votos para o candidato do PSDB é maior que os 40% de votos nulos e brancos; b) a escala vertical não representa a porcentagem de intenções de votos indicadas pelas barras; e c) a barra dos 10% de eleitores que não sabiam em quem votar é menor que os 4% que pretendiam votar no candidato do PT. Após várias críticas, a campanha do referido candidato retirou o infográfico do ar indicando que a publicação foi feita após terem recebido o material de militantes. A matéria completa pode ser encontrada no link [Matéria UOL de Lucas Borges Teixeira \(30/05/2018\)](#).

2. Pesquisas Amostrais:

Visite a página do IBGE www.ibge.gov.br e explore a página buscando conhecer as pesquisas que o IBGE realiza e os diferentes Bancos de Dados disponíveis para uso. Você deverá fazer uma resenha do Censo Demográfico e PNADc . O foco é a metodologia da pesquisa, o plano (desenho) amostral e o elenco de características investigadas.

Resenha: Censo Demográfico e PNAD Contínua

Censo Demográfico:

O Censo Demográfico é uma das pesquisas mais importantes realizadas pelo Instituto Brasileiro de Geografia e Estatística (IBGE). Ele é conduzido a cada dez anos e tem como objetivo principal fornecer informações detalhadas sobre a população brasileira, suas características socioeconômicas e demográficas, bem como sobre as condições de vida e as estruturas domiciliares.

Metodologia da Pesquisa:

- O Censo Demográfico é um levantamento realizado em todo o território nacional, com a participação de todos os municípios brasileiros.
- Utiliza uma abordagem censitária, ou seja, tenta abranger toda a população, não apenas uma amostra dela.
- A coleta de dados é feita por meio de questionários aplicados diretamente aos moradores dos domicílios, abordando uma ampla gama de tópicos, como idade, sexo, educação, trabalho, renda, condições de habitação, entre outros.
- Os dados coletados são tabulados e analisados para produzir estatísticas confiáveis e abrangentes sobre a população brasileira.

Plano (Desenho) Amostral:

- Por se tratar de um censo, o Censo Demográfico não utiliza um plano amostral no sentido tradicional, já que busca abranger toda a população.
- No entanto, a logística de coleta e processamento dos dados segue um planejamento detalhado, que envolve a divisão do território em setores censitários e a alocação de equipes de recenseadores para cada área.

Elenco de Características Investigadas:

- O Censo Demográfico investiga uma ampla variedade de características da população brasileira, incluindo:
 - Demográficas: idade, sexo, estado civil, migração, entre outros.
 - Socioeconômicas: educação, ocupação, renda, condições de trabalho, acesso a serviços públicos, entre outros.
 - Habitacionais: tipo de domicílio, condições de moradia, posse de bens duráveis, entre outros.

PNAD Contínua:

A Pesquisa Nacional por Amostra de Domicílios Contínua (PNAD Contínua) é uma pesquisa amostral realizada pelo IBGE que fornece informações contínuas sobre a população brasileira, suas características socioeconômicas e demográficas. Ao contrário do Censo, a PNAD Contínua é realizada anualmente e utiliza uma amostra representativa da população.

Metodologia da Pesquisa:

- A PNAD Contínua utiliza uma abordagem amostral, selecionando uma amostra representativa de domicílios em todo o país.
- A coleta de dados é feita por meio de entrevistas domiciliares, onde são aplicados questionários aos moradores, abordando uma variedade de tópicos, semelhantes aos do Censo Demográfico.
- Os dados são tabulados e analisados para produzir estimativas confiáveis sobre a população brasileira e suas características.

Plano (Desenho) Amostral:

- O plano amostral da PNAD Contínua envolve a seleção de uma amostra probabilística de domicílios em todo o país, utilizando técnicas de amostragem estratificada e por conglomerados.
- A amostra é atualizada regularmente para garantir sua representatividade e precisão.

Elenco de Características Investigadas:

- Assim como o Censo Demográfico, a PNAD Contínua investiga uma ampla gama de características da população brasileira, incluindo:
- Demográficas: idade, sexo, estado civil, migração, entre outros.
- Socioeconômicas: educação, ocupação, renda, acesso a benefícios sociais, entre outros.
- Habitacionais: tipo de domicílio, condições de moradia, acesso a serviços básicos, entre outros.

Conclusão: Tanto o Censo Demográfico quanto a PNAD Contínua são pesquisas fundamentais para o conhecimento da população brasileira. Enquanto o Censo fornece uma visão detalhada da população a cada dez anos, a PNAD Contínua complementa essas informações, fornecendo dados anuais que permitem acompanhar tendências e mudanças na sociedade brasileira ao longo do tempo. Ambas as pesquisas utilizam metodologias rigorosas e planos amostrais bem elaborados para garantir a qualidade e a representatividade dos dados coletados.