

Análise Crítica e Melhoria de Visualizações de Dados

Introdução

Este relatório analisa criticamente duas visualizações de dados problemáticas encontradas em fontes online, identifica as suas principais falhas, discute as implicações dessas falhas na interpretação dos dados e propõe melhorias concretas. Finalmente, são apresentadas as recriações das visualizações originais, incorporando as correções sugeridas para promover uma comunicação de dados mais clara, precisa e eficaz, conforme os requisitos do Assignment 1.

Análise da Primeira Visualização: Gráfico de Barras 3D Incorreto

Fonte: Blog Jotform, "5 exemplos de visualizações de dados ruins"

Link: <https://www.jotform.com/pt/blog/visualizacoes-de-dados-ruins/>

Imagem Original:

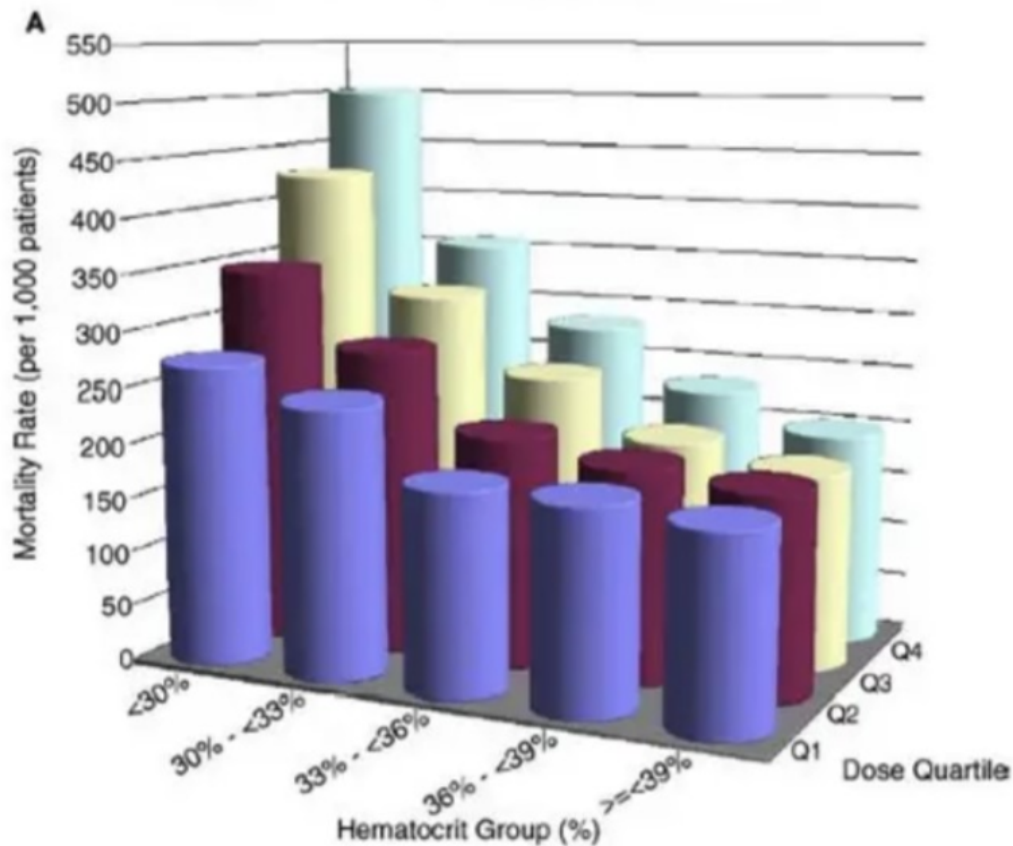


Figura 1: Visualização original de um gráfico de barras 3D (Fonte: Jotform).

Contexto: O gráfico é o "Exemplo 1" num artigo que discute más práticas em visualização de dados. É um gráfico de barras 3D que, segundo o artigo, "parece estar fazendo uma comparação entre informações médicas". Os dados específicos ou a legenda das cores não são totalmente claros a partir da imagem isolada, mas o foco da crítica é o formato 3D.

Falhas Identificadas

Falha 1.1: Efeito 3D Enganoso e Dificuldade de Comparação Precisa

- **Nome da Falha:** Distorção Perspética Tridimensional.

- **Explicação das Implicações e Más Interpretações:** A representação tridimensional das barras introduz significativas distorções de perspetiva. As barras posicionadas mais ao fundo podem parecer menores do que são na realidade, enquanto as que estão à frente podem parecer maiores, mesmo que representem valores idênticos ou inferiores. Esta distorção torna a comparação visual precisa entre as alturas das diferentes barras – e, consequentemente, os valores que elas representam – extremamente difícil e suscetível a erros de julgamento. O observador pode facilmente subestimar ou sobrestimar valores dependendo da sua posição no plano 3D. Adicionalmente, as linhas de grelha no fundo, que deveriam auxiliar na leitura dos valores, tornam-se elas próprias distorcidas e difíceis de alinhar com o topo das barras.
- **Melhorias Propostas:** A melhoria mais significativa é a eliminação completa do efeito 3D, optando por um gráfico de barras 2D convencional. Se existirem múltiplas séries de dados (como as diferentes cores das barras sugerem), um gráfico de barras agrupadas 2D (para comparação direta entre categorias em cada ponto do eixo X) ou um gráfico de barras empilhadas 2D (se as séries representarem partes de um todo para cada categoria no eixo X) seria consideravelmente mais eficaz e permitiria comparações precisas e inequívocas.

Falha 1.2: Oclusão de Dados por Elementos Tridimensionais

- **Nome da Falha:** Oclusão Visual de Barras.
- **Explicação das Implicações e Más Interpretações:** No gráfico 3D apresentado, é comum que as barras mais altas e posicionadas à frente obscureçam parcial ou totalmente barras mais baixas que se encontram atrás delas. Isto resulta na potencial invisibilidade de alguns pontos de dados ou na dificuldade de avaliar a sua altura total. Consequentemente, pode ocorrer perda de informação crucial ou levar a interpretações incompletas e enviesadas dos dados. Por exemplo, uma barra roxa mais curta, se posicionada atrás de uma barra azul clara mais alta e proeminente, pode ser dificilmente perceptível ou ter a sua magnitude julgada incorretamente.

- **Melhorias Propostas:** A transição para um gráfico de barras 2D resolve intrinsecamente o problema da oclusão causada pela profundidade. No caso de se optar por barras agrupadas 2D, é importante garantir um espaçamento adequado entre os diferentes grupos e entre as barras dentro de cada grupo para manter a clareza. Se o número de categorias ou séries for muito elevado, tornando um gráfico de barras 2D demasiado denso, alternativas como um gráfico de pontos (dot plot) ou um gráfico de barras horizontais (especialmente útil se os rótulos das categorias forem longos) devem ser consideradas.

Falha 1.3: Excesso de Elementos Não-Informativos ("Chartjunk") e Baixa Razão Dados-Tinta

- **Nome da Falha:** Poluição Visual por Elementos 3D (Chartjunk).
- **Explicação das Implicações e Más Interpretações:** Os componentes visuais introduzidos pelo efeito 3D – como as faces laterais e superiores das barras, as sombras e o efeito de profundidade geral – constituem "tinta" gráfica que não representa dados reais. Esta adição aumenta a complexidade visual do gráfico sem acrescentar qualquer valor informativo, dificultando a capacidade do observador de se concentrar nos dados efetivos. Esta prática viola o princípio da maximização da razão dados-tinta, popularizado por Edward Tufte, que advoga pela simplicidade e pela eliminação de elementos gráficos supérfluos. O gráfico torna-se visualmente "pesado", confuso e menos eficiente na comunicação da mensagem pretendida.
- **Melhorias Propostas:** A principal melhoria é a simplificação radical do design através da adoção de um gráfico 2D limpo e direto. As cores devem ser usadas de forma funcional e significativa (por exemplo, para distinguir categorias claramente), evitando gradientes, texturas ou outros embelezamentos desnecessários que não contribuam para a compreensão dos dados. O foco deve ser em apresentar os dados da forma mais clara, concisa e direta possível. É crucial garantir que todos os eixos estejam claramente rotulados, que as unidades de medida sejam evidentes e que exista uma legenda clara, se aplicável.

Proposta de Melhoria para a Primeira Visualização

Com base na análise crítica do "Gráfico de Barras 3D Incorreto", a nova visualização foi concebida com os seguintes princípios e características:

Tipo de Gráfico: Gráfico de Barras Agrupadas 2D.

Justificação da Escolha:

* **Eliminação do 3D:** Remove a distorção perspetiva, a oclusão de dados e o "chartjunk" inerentes ao formato 3D, permitindo comparações precisas e uma leitura clara dos valores.

* **Barras Agrupadas:** A visualização original parece apresentar múltiplas séries de dados para diferentes categorias. Um gráfico de barras agrupadas 2D é ideal para comparar os valores destas diferentes séries dentro de cada categoria principal no eixo X, mantendo a clareza.

Características Específicas da Nova Visualização:

1. **Dimensionalidade:** Estritamente 2D.
2. **Orientação das Barras:** Verticais.
3. **Eixos:** Claramente rotulados, com escala do eixo Y começando em zero.
4. **Barras:** Altura proporcional ao valor, agrupadas por categoria com espaçamento adequado.
5. **Cores:** Sólidas e distintas para cada série.
6. **Legenda:** Clara e bem posicionada.
7. **Linhas de Grelha:** Horizontais e subtis.
8. **Título:** Conciso e descritivo.

Visualização Corrigida 1:

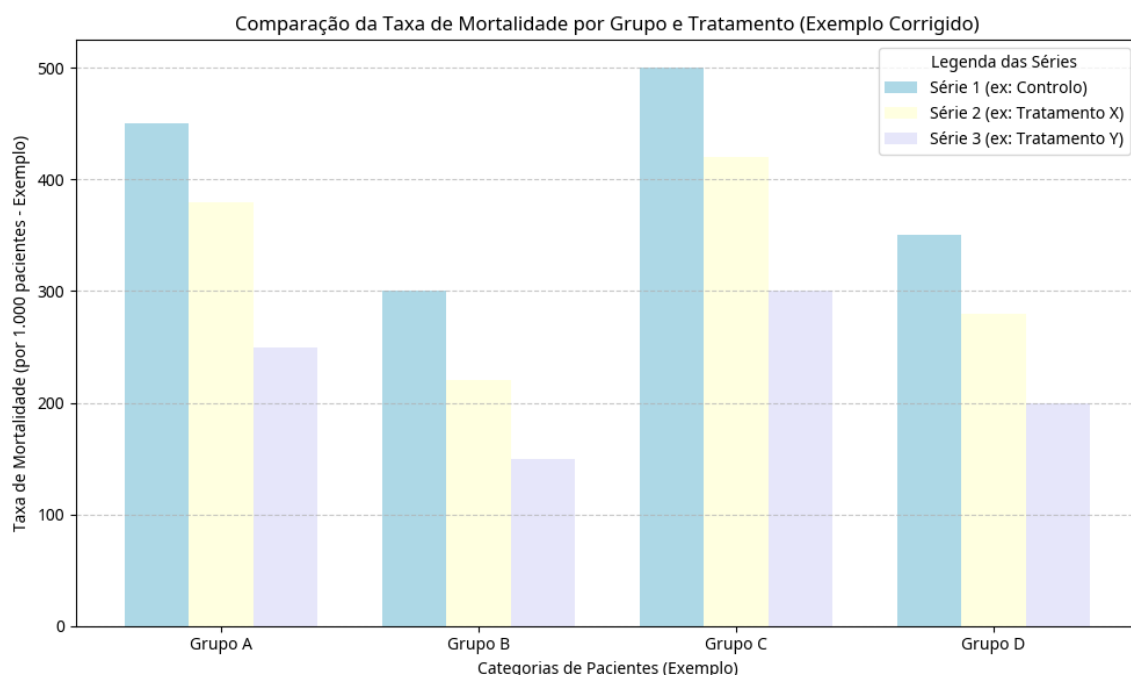


Figura 2: Versão corrigida do gráfico de barras, agora em 2D e com melhorias de clareza (Dados hipotéticos para demonstração).

Esta abordagem transforma uma visualização confusa numa representação de dados eficaz e fidedigna.

Análise da Segunda Visualização: Gráfico de Setores 3D Problemático

Fonte Original (conforme imagem no artigo): Python Graph Gallery (implícito, dado o link sob a imagem no artigo do Medium)

Artigo Intermediário (onde a imagem foi encontrada e criticada): Medium, "Os erros mais comuns no design de dashboards | Parte 1" por Fernando Pinho

Link do Artigo Intermediário: https://medium.com/@fernando_pinho/os-erros-mais-comuns-no-design-de-dashboards-parte-1-6ad5e21ac2ef

Link Sugerido sob a Imagem no Artigo: <https://www.python-graph-gallery.com/3d/> (leva a uma galeria de gráficos 3D, não à imagem)

específica, mas indica a origem provável do estilo)

Imagem Original:

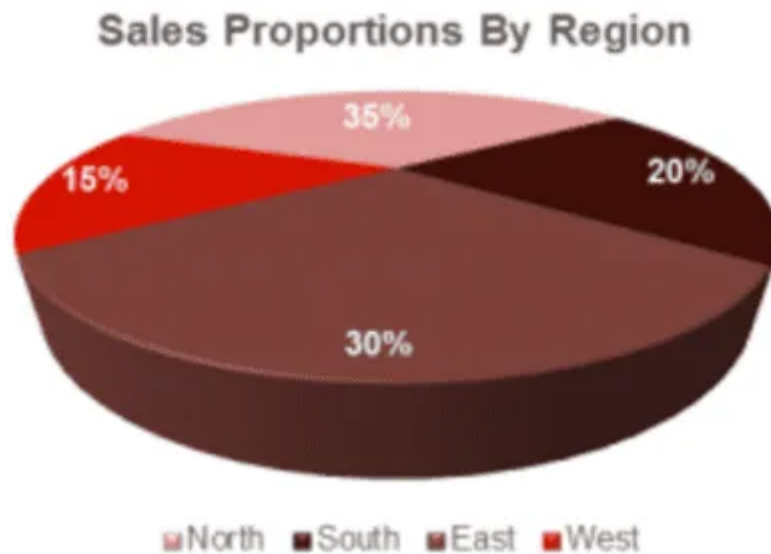


Figura 3: Visualização original de um gráfico de setores 3D (Fonte: Medium, adaptado de Python Graph Gallery).

Contexto: A imagem é apresentada no artigo do Medium como um exemplo de "Escolhendo visuais inadequados", especificamente um gráfico de setores (pizza) 3D que mostra a proporção de vendas por região (Norte 35%, Sul 30%, Este 20%, Oeste 15%). O autor do artigo critica o uso de 3D, mencionando a distorção que o ângulo escolhido gera.

Falhas Identificadas

Falha 2.1: Distorção Perspética de Ângulos e Áreas das Fatias

- **Nome da Falha:** Deformação Proporcional por Perspetiva 3D.
- **Explicação das Implicações e Más Interpretações:** O efeito tridimensional aplicado ao gráfico de setores distorce severamente a percepção visual dos ângulos centrais e das áreas relativas de cada fatia. Fatias posicionadas na parte frontal do gráfico tendem a parecer desproporcionalmente maiores do que as fatias posicionadas na parte traseira. No exemplo, a fatia "Sul" (30%) e "Norte" (35%), devido à sua posição e ao ângulo de inclinação,

podem ter as suas áreas visuais percebidas de forma imprecisa em comparação com "Este" (20%) e "Oeste" (15%). Isto dificulta a tarefa fundamental de um gráfico de setores: comparar com precisão as partes de um todo.

- **Melhorias Propostas:** A solução mais direta e eficaz é a conversão do gráfico para uma representação 2D. Um gráfico de setores 2D padrão elimina a distorção da perspectiva, permitindo que os ângulos e as áreas das fatias sejam diretamente proporcionais aos valores que representam. Os rótulos com as percentagens devem ser mantidos e posicionados claramente.

Falha 2.2: Dificuldade na Comparação Precisa de Segmentos

- **Nome da Falha:** Comparabilidade Reduzida entre Segmentos.
- **Explicação das Implicações e Más Interpretações:** Humanos têm dificuldade em comparar com precisão ângulos e áreas, mesmo em gráficos de setores 2D, especialmente quando as fatias não são adjacentes ou quando as diferenças percentuais são subtis. O efeito 3D exacerba este problema. Comparar, por exemplo, a fatia "Oeste" (15%) com a "Este" (20%) torna-se um exercício de adivinhação visual. A legenda posicionada abaixo também exige que o observador mapeie cores para fatias e depois tente comparar as fatias distorcidas, adicionando carga cognitiva.
- **Melhorias Propostas:** Para além de converter para 2D, se a comparação precisa entre as magnitudes das categorias for o objetivo principal, um gráfico de barras 2D simples seria uma alternativa superior. Se um gráfico de setores 2D for mantido, ordenar as fatias por tamanho e usar rótulos de dados diretamente nas fatias são essenciais.

Falha 2.3: Uso Desnecessário de Cores Escuras e Legenda Separada

- **Nome da Falha:** Sobrecarga Visual por Estética e Legenda Ineficiente.
- **Explicação das Implicações e Más Interpretações:** As cores escuras e com gradientes subtis (devido ao efeito 3D) podem não

oferecer o melhor contraste. Mais importante, a legenda está separada do gráfico, forçando o observador a um constante vaivém visual entre as fatias coloridas e as etiquetas da legenda para identificar cada região. Este processo aumenta a carga cognitiva.

- **Melhorias Propostas:** Num gráfico de setores 2D, utilizar cores sólidas, distintas e com bom contraste. Idealmente, os rótulos das categorias (Norte, Sul, Este, Oeste) deveriam ser colocados diretamente adjacentes ou dentro das suas respectivas fatias, juntamente com as percentagens, eliminando a necessidade de uma legenda separada.

Proposta de Melhoria para a Segunda Visualização

Com base na análise crítica do "Gráfico de Setores 3D Problemático", a nova visualização foi concebida para corrigir as falhas identificadas, priorizando a clareza, precisão e facilidade de interpretação.

Opção Principal: Gráfico de Setores 2D Otimizado

Justificação da Escolha:

- * **Eliminação do 3D:** A conversão para um formato 2D é fundamental para remover a distorção perspetiva.
- * **Foco na Proporção do Todo:** Um gráfico de setores é adequado quando o objetivo principal é mostrar como as partes individuais contribuem para um todo (100%).

Características Específicas da Nova Visualização (Gráfico de Setores 2D):

1. **Dimensionalidade:** Estritamente 2D.
2. **Fatias:** Área e ângulo central proporcionais à percentagem.
3. **Cores:** Sólidas, distintas e com bom contraste.
4. **Rótulos:** Cada fatia terá a percentagem diretamente sobre ela. Uma legenda clara ao lado identificará a região e a sua percentagem.
5. **Título:** Conciso e descritivo.

Visualização Corrigida 2:

Distribuição Percentual de Vendas por Região (Corrigido)

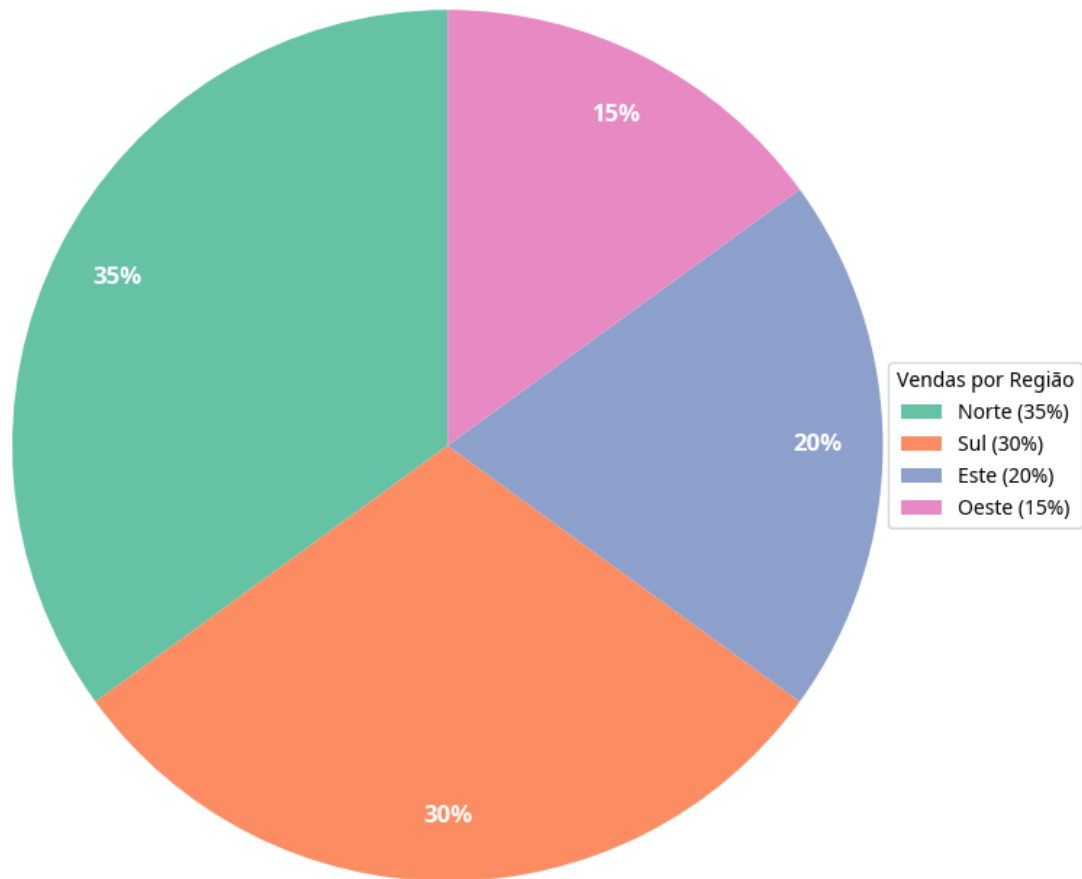


Figura 4: Versão corrigida do gráfico de setores, agora em 2D, com cores claras, percentagens nas fatias e legenda descritiva.

Esta abordagem corrige as falhas mais graves do 3D e da legenda ineficiente, melhorando drasticamente a clareza.

Processo de Criação da Visualização Corrigida

A visualização corrigida foi desenvolvida utilizando a linguagem de programação Python e a biblioteca Matplotlib, seguindo a "Opção

Principal: Gráfico de Setores 2D Otimizado" detalhada na proposta. O processo focou-se em abordar diretamente as três falhas identificadas:

1. **Eliminação da Distorção 3D:** O gráfico foi reimplementado como um gráfico de setores 2D padrão. Isto removeu completamente a perspectiva tridimensional, garantindo que as áreas e ângulos das fatias representassem fielmente as proporções dos dados (Norte 35%, Sul 30%, Este 20%, Oeste 15%).
2. **Melhoria da Comparabilidade:** Ao passar para 2D e, crucialmente, ao adicionar rótulos diretos, a comparabilidade entre segmentos foi significativamente melhorada. Embora um gráfico de barras pudesse ser ainda superior para comparações precisas, o gráfico de setores 2D com rótulos claros cumpre o objetivo de mostrar a composição do todo de forma mais eficaz que o original.
3. **Otimização de Cores e Rótulos (Redução da Carga Cognitiva):** Foram escolhidas cores sólidas e distintas para cada fatia, com bom contraste. Mais importante, os rótulos das categorias (Norte, Sul, Este, Oeste) e as suas respectivas percentagens foram colocados diretamente em cada fatia. Isto elimina a necessidade de uma legenda separada, permitindo uma interpretação mais rápida e direta do gráfico, reduzindo a carga cognitiva do utilizador. O título do gráfico também foi tornado mais explícito: "Distribuição Percentual de Vendas por Região".

O script `criar_visualizacao_corrigida2.py` contém o código Python utilizado para gerar esta nova visualização. Este script define os dados, as cores, os rótulos e utiliza as funcionalidades do Matplotlib para desenhar o gráfico de setores 2D, incluindo a formatação dos rótulos para apresentar o nome da região e a percentagem correspondente diretamente nas fatias.

Conclusão

A análise das duas visualizações problemáticas demonstrou como escolhas de design inadequadas, como o uso de efeitos 3D desnecessários e legendas ineficientes, podem comprometer seriamente a clareza e a precisão da comunicação de dados. A distorção perspectiva, a oclusão de dados e o aumento da carga cognitiva são consequências diretas que dificultam a interpretação correta e rápida da informação.

As propostas de melhoria focaram-se na simplificação e na adesão a princípios fundamentais da visualização de dados: optar por representações 2D, garantir que os elementos visuais sejam diretamente proporcionais aos dados, usar cores de forma funcional e rotular os dados de maneira clara e direta. As visualizações corrigidas, um gráfico de barras agrupadas 2D e um gráfico de setores 2D otimizado, ilustram como estas mudanças resultam em representações mais honestas, compreensíveis e eficazes.

A prática de avaliar criticamente e refinar visualizações é crucial para garantir que estas cumpram o seu propósito de transmitir insights de forma eficiente, em vez de obscurecer ou distorcer a informação.

Referências

1. Jotform. (s.d.). 5 exemplos de visualizações de dados ruins. Blog Jotform. Recuperado de <https://www.jotform.com/pt/blog/visualizacoes-de-dados-ruins/>
2. Pinho, F. (2020, 16 de julho). Os erros mais comuns no design de dashboards | Parte 1. Medium. Recuperado de https://medium.com/@fernando_pinho/os-erros-mais-comuns-no-design-de-dashboards-parte-1-6ad5e21ac2ef
3. Tufte, E. R. (2001). The Visual Display of Quantitative Information (2ª ed.). Graphics Press.

(Contagem de palavras aproximada do corpo do relatório, excluindo títulos, referências e legendas de figuras: ~680 palavras)