

Fundamentos de Visualização de Dados

Introdução Storytelling

Você provavelmente já ouviu falar a palavra storytelling, o que talvez você não saiba é que não somos naturalmente bons em storytelling. Esta palavra tem por definição a habilidade de contar histórias com uma narrativa envolvente para transmitir uma mensagem e criar conexões. Trata-se de aproveitar as histórias para envolver seu público ou esclarecer qualquer coisa (Knafllic, 2016; Souza, 2022).

Storytelling na Análise de Dados

No contexto de dados não seria diferente, utilizamos algumas ferramentas para a construção da história através de gráficos e tabelas. Você pode inserir alguns dados no Excel e criar um gráfico. Para muitos, a visualização de dados acaba aí. Isso pode estragar completamente uma história, muito interessante ou, pior, torná-la difícil ou impossível de entender (Knafllic, 2016). Ou seja, o trabalho de um Analista de dados é a coleta dos dados brutos, uma avaliação dos dados e por fim elaborar conclusões relevantes (*insights*) que podem ser utilizadas de maneiras significativas (Souza, 2020). Aqui separamos algumas dicas e boas práticas para facilitar na hora de montar um *dashboard*.

Ter estratégia para apresentar os dados é imprescindível, mas antes de começarmos a falar sobre o desenvolvimento do *dashboard* é necessário elucidar alguns conceitos, como análise exploratória e análise explanatória.

A visualização de dados começa muito antes do desenvolvimento de um *dashboard*. É importante uma análise dos dados brutos que chegam até você no primeiro momento. A análise **exploratória** é realizada pelo profissional responsável pela análise. É realizado uma “garimpagem” nos dados gerados para descobrir o que realmente é interessante e aproveitável ali (Gerola, 2020).”Quando fazemos uma análise exploratória, é como procurar pérolas em ostras. Talvez precisemos de 100 ostras para encontrarmos duas

pérolas”(Knafllic, 2016). No desenvolvimento do *dashboard* você não falará sobre as 100 ostras e sim das duas pérolas. Contar sobre as pérolas é **explicar** a sua análise, comunicar apenas o que é realmente importante (Gerola, 2020).

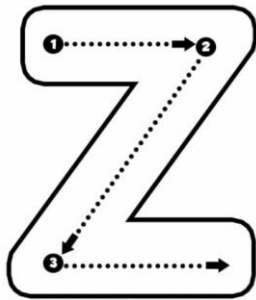
Conheça seu público e identifique seu objetivo

Você precisa saber o motivo pelo qual seu público precisa de um *dashboard*. Pois para atrair a atenção você precisa ter um propósito bem definido.

Para ajudar a conhecer o seu cliente e entender o seu objetivo, uma boa prática é antes de montar a visualização final, criar um **esboço** do resultado esperado em conjunto ao cliente. Este esboço não precisa ser uma representação fiel com a preocupação de estética, cores, itens, etc. É apenas para definir a posição dos gráficos e dos dados (Karpinski, 2022). Você pode fazer o esboço a mão ou então utilizar algum software para isso, abaixo temos alguns exemplos de ferramentas que podem auxiliar nesse processo :

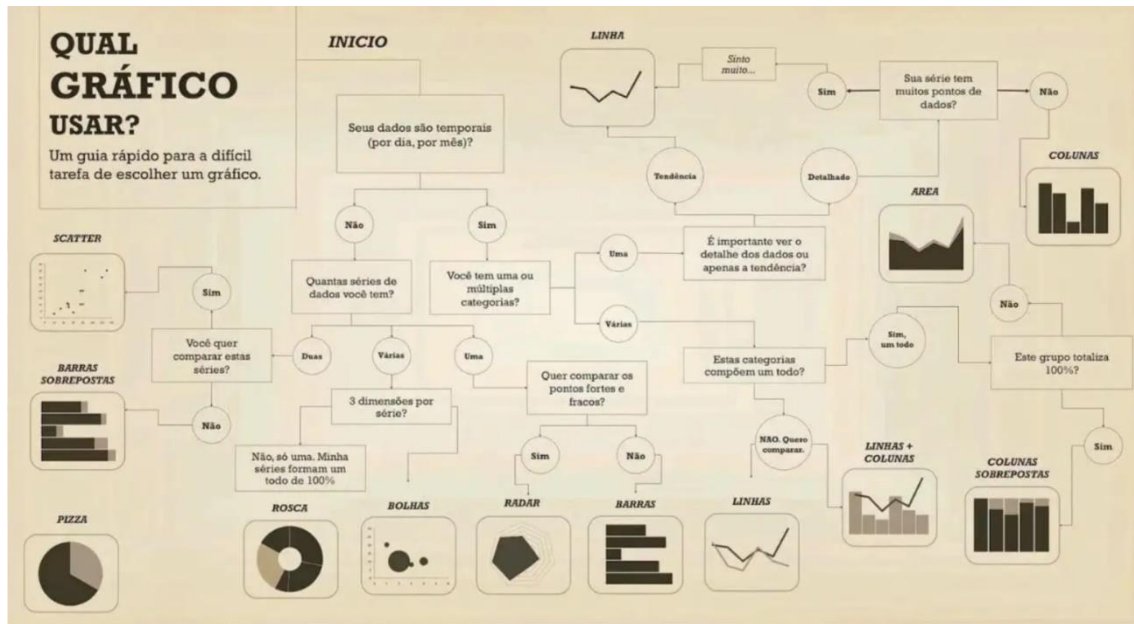
- [Figma](#)
- [Draw io](#)
- [Excalidraw](#)
- [MockFlow](#)
- [Wireframe](#)
- [Bizagi](#)

Para direcionar o seu desenho você pode fazer algumas perguntas, como por exemplo quais são os KPIs mais importantes. Os KPIs (Key Performance Indicator, ou em português Indicador – Chave de Desempenho) são indicadores dos valores quantitativos fundamentais para o acompanhamento e o melhor gerenciamento do nível de desempenho de uma empresa/cliente. Ou então qual a informação que o cliente destaca como fundamental. Uma dica de visualização é sempre colocar os dados mais importantes no topo da página, o usuário tende a deslizar o olhar em zigzag, seguindo um caminho em “z”, da esquerda para a direita (Karpinski, 2022).



Escolha um visual adequado

Após uma análise detalhada dos dados a escolhas dos KPIs e de quais informações você espera trazer é o momento de escolher os gráficos que irão compor o *dashboard*. A escolha dos gráficos parece uma tarefa simples, mas não é. Como foi citado no início deste documento não somos bons naturalmente com storytelling, ninguém quer escolher um gráfico ruim, mas acontece. Quando escolhemos um gráfico estamos facilitando o acesso à informação, a escolha de um gráfico equivocado pode gerar uma interpretação equivocada, abaixo existem dois guias que podem te ajudar a escolher qual gráfico utilizar em determinadas situações, lembrando que nem sempre todas as ferramentas de visualização são contempladas com todas as opções de gráfico abaixo (Souza, 2020):



[Link](#) para [acesso](#)

THE VISUALS REFERENCE

— FOR MICROSOFT POWER BI —

SEP. 2018
<http://sqlbi.com/visual-reference>

COMPARISON

Display measures compared by their magnitude

CHANGE OVER TIME

Display the changing trend of measures

RANKING

Display measures by their rank order

SPATIAL

Display measures over spatial maps

FLOW

Display a flow or dynamic relations

PART-TO-WHOLE

Display the parts of a measure

DISTRIBUTION

Display the distribution of a measure

CORRELATION

Display relations between measures

SINGLE

Display single values

FILTER

Control report filters

NARRATIVE

Tell a story with data

MISCELLANEOUS

Recommended
There is a better alternative
Don't use in the category
Built-in visual
Certified visual
Not required

2018 © SQLBI AND OKVIZ ARE TRADEMARKS OF SQLBI CORP. / POWER BI IS A TRADEMARK OF MICROSOFT CORP.

[Link](#) para [acesso](#)



[Link para acesso](#)

Você pode encontrar mais informações sobre gráficos nestes sites (Souza, 2022):

<https://datavizcatalogue.com/index.html>

<https://medium.com/geekculture/create-beautiful-graphs-with-python-4235f50b2adb>

<https://www.python-graph-gallery.com/scatter-plot/>

<https://bbc.github.io/rcookbook/>

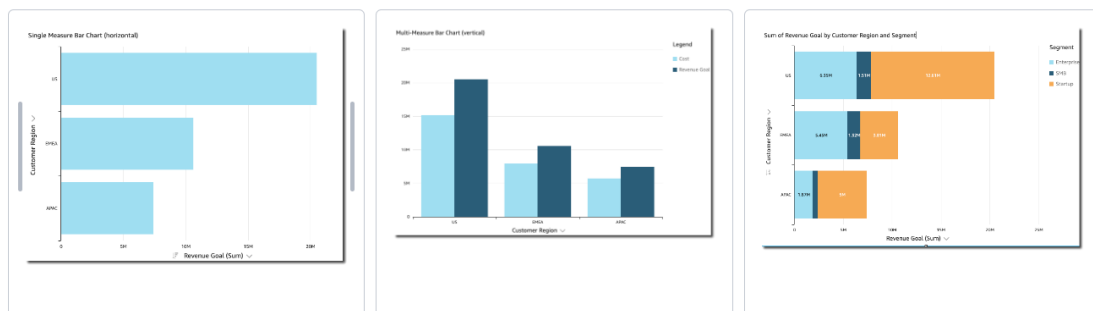
Escolha de gráficos é importante, porém outro ponto que deve ser levado em consideração é o **contexto**. Os números não mentem, mas eles não se sustentam sozinhos. E lembre-se, você é responsável por passar informações de maneira respeitosa.

Gráficos do QuickSight

O QuickSight conta com diversos tipos de gráficos para adicionar a uma visualização. Temos desde os componentes mais triviais, como barras, linhas e pizza até tabelas dinâmicas e gráficos de dispersão. Na sequência adicionaremos os tipos de gráfico mais comuns acompanhados do link das respectivas documentações.

Gráficos de Barra

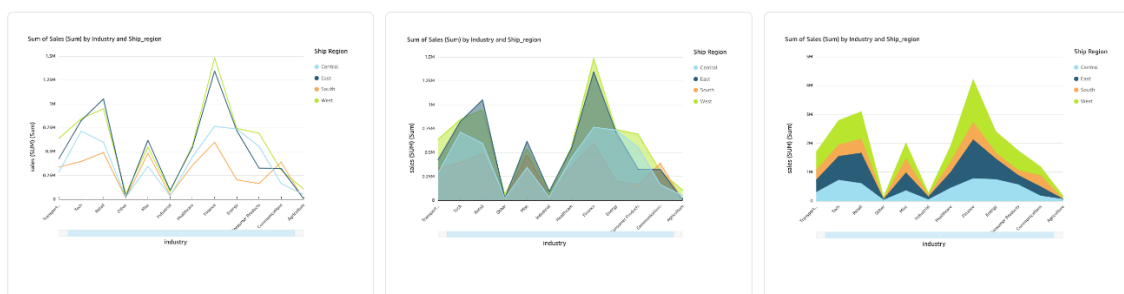
O gráfico de barras no geral é o mais versátil, além da facilidade em interpretá-lo. Podem ser construídos utilizando barras horizontais ou verticais com possibilidade de uma ou mais categorias.



[Documentação](#)

Gráficos de Linha

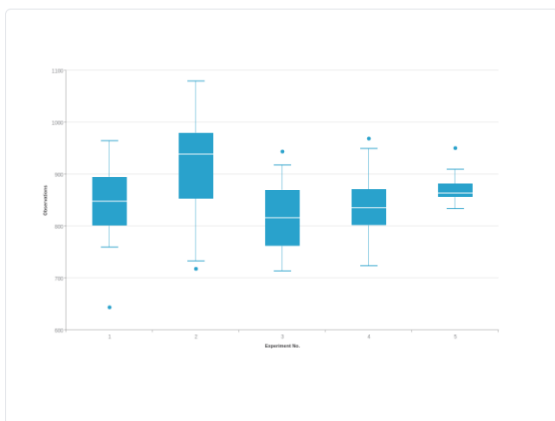
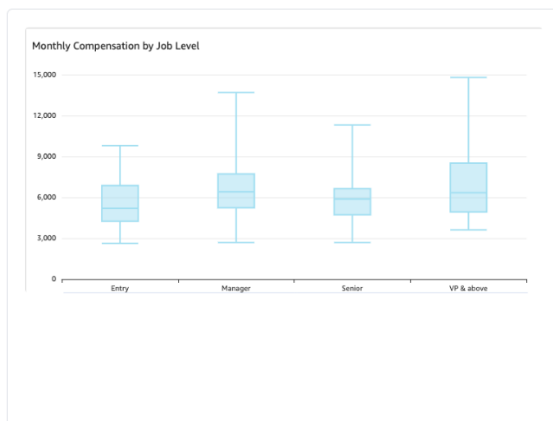
São comumente utilizados para apresentar as mudanças de uma métrica com o passar do tempo.



[Documentação](#)

Gráfico BoxPlot

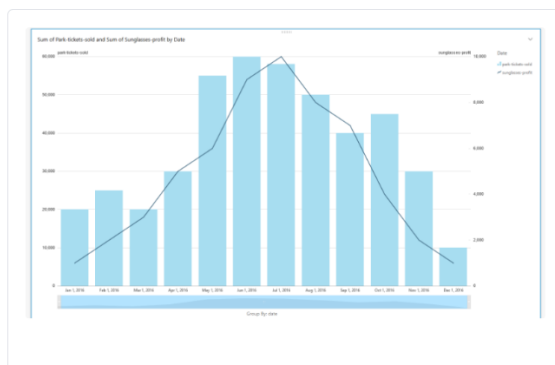
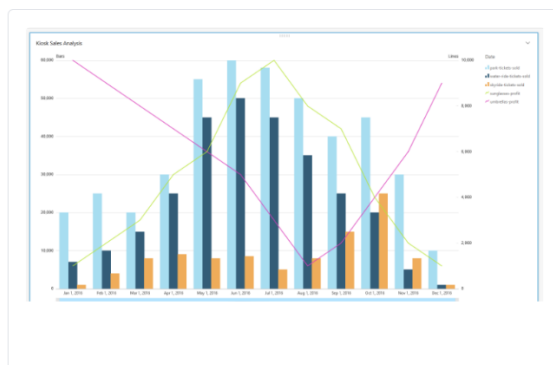
É um tipo de gráfico muito utilizado em estatística descritiva, pois identifica de forma visual 5 importantes valores: **máximo**, **mínimo**, **mediana**, **quartil inferior** e **quartil superior**. Você pode encontrar uma explicação sobre as características deste gráfico [aqui](#).



Documentação

Gráficos de linhas e colunas (combo)

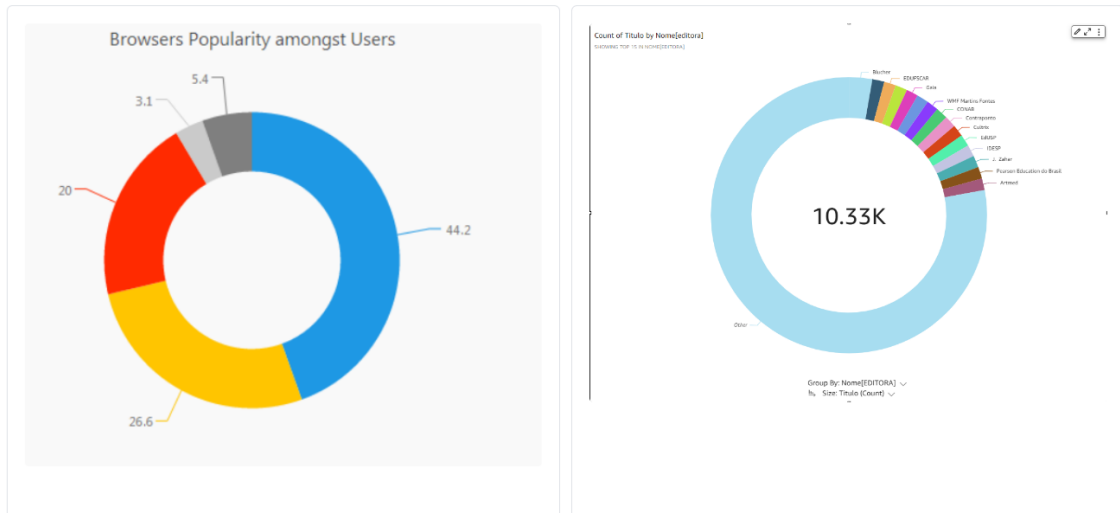
Permitem exibir a comparação entre diferentes categorias associada com mudanças ao longo do tempo.



Documentação

Gráficos Donut (Rosquinha)

São utilizados para comparar o quanto cada valor de uma dimensão representa sobre o total.



[Link para acesso](#)

Gráfico Heat maps

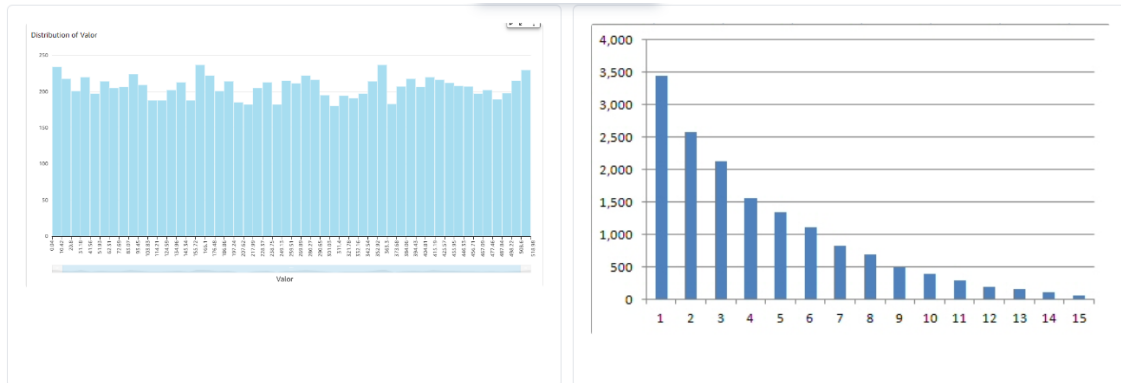
São gráficos que apresentam uma medida em relação a sua intersecção com duas dimensões. Aplica-se uma variação de coloração para indicar se o valor aproxima-se do máximo (mais escuro) ou do mínimo (mais claro).



[Documentação](#)

Histograma

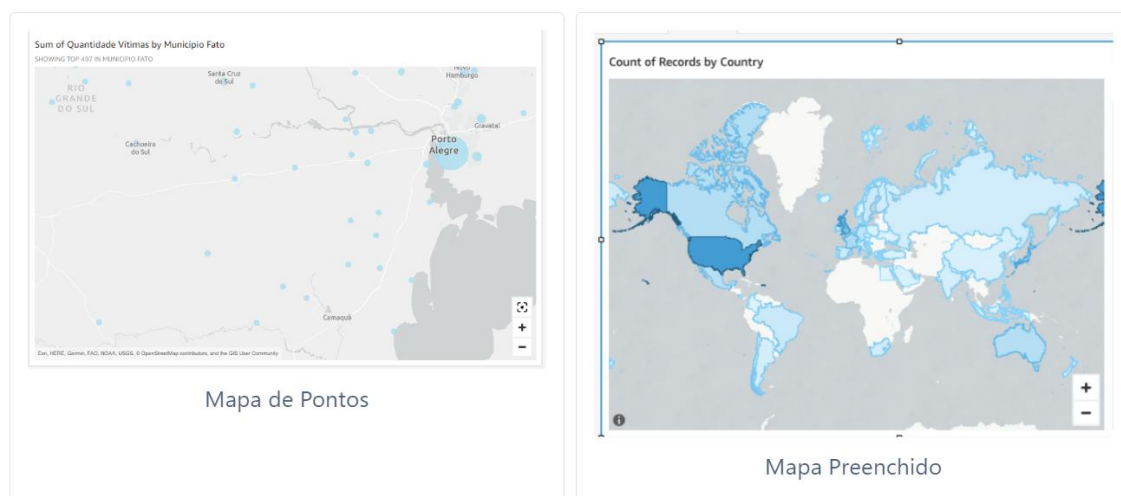
Histogramas permitem representar graficamente a distribuição de frequência de uma métrica.



[Documentação](#)

Gráficos de Mapa

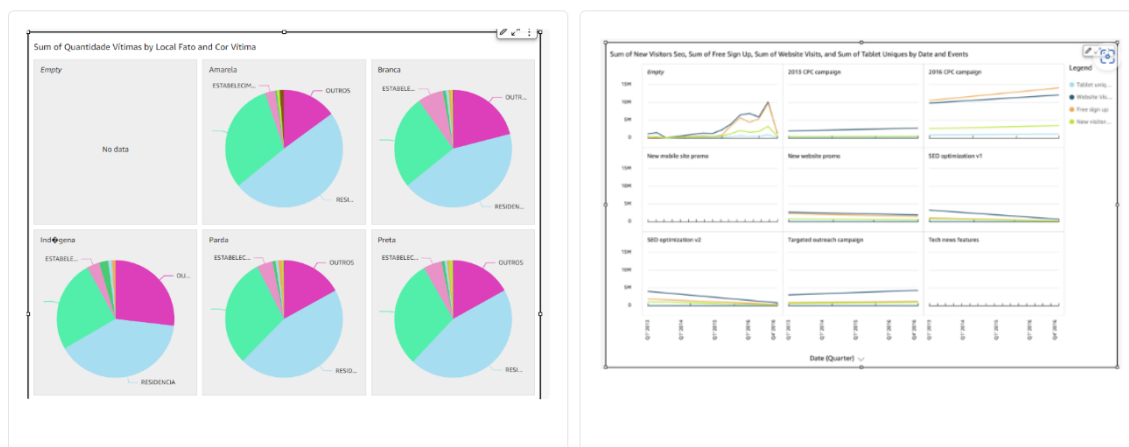
QuickSight permite criar dois tipos de gráfico no formato de mapas: *de pontos* e *preenchido*. Mapa de pontos apresentam valores para cada localização, diferenciando valores pelo tamanho do ponto. Já um mapa preenchido apresenta a diferença de valores de cada região alterando as tonalidades de cores. Em ambos é pre-requisito para uso a existência de informações de localização no *dataset*.



[Documentação](#)

Múltiplas pequenas visualizações

Gráficos de linha, barras e pizza podem ser customizados para apresentar várias pequenas visualizações, criadas a partir de um campo de categorização (***Small multiples***). É um recurso especialmente útil para apresentar uma série de visualizações que evidenciam o impacto do campo de categorização nos dados.



Documentação

Tabelas do QuickSight

Tabelas

Utilizada para apresentar dados em formato tabular. Permite algumas customizações, como alterar cor de fundo de linha, coluna ou texto e adicionar ícone quando os dados se enquadrarem em determinadas condições de formatação condicional.

Conditional formatting

Column: Select a column

Conditions: Collapse all

- Sexo Vítima: Background - Solid Greater than 2
- Cor Vítima: Background - Solid Greater than 2
- Tipo Fato: Background - Solid Greater than 2

Actions: Add text color, Add icon, Delete

Field wells

Group by: Município Fato, Sexo Vítima, Cor Vítima

Value: Tipo Fato (Count)

Sheet 2

Count of Tipo Fato by Município Fato, Sexo Vítima, and Cor Vítima

Município Fato	Sexo Vítima	Cor Vítima	Tipo Fato
PONTE PRETA			3
PONTE PRETA	Feminino	Branca	1
PONTE PRETA	Masculino	Branca	3

Context menu: Conditional formatting, Show totals, Export to CSV, Export to Excel, Duplicate visual, Actions, Delete

Tabelas Dinâmicas

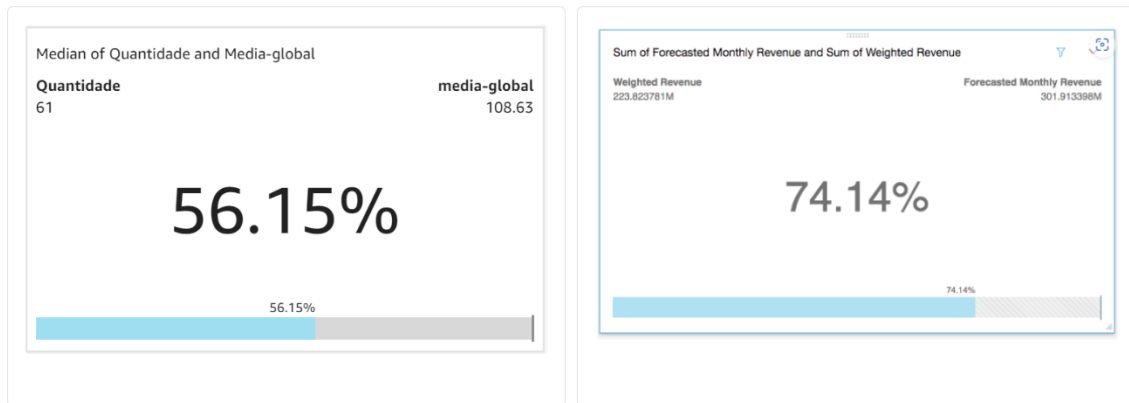
Tabelas dinâmicas são úteis para apresentar métricas com base na intersecção de dimensões. Tem características muito similares às tabelas dinâmicas de softwares de planilha eletrônica.

Customer Region	Date	Service Line > Consumption Channel					
		Billing			HR		
		Web	Mobile	API	Web	Mobile	API
APAC	2014	\$257K	\$435K	\$519K	\$762K	\$925K	\$999K
	2015	\$471K	\$609K	\$954K	\$1,419K	\$1,742K	\$1,888K
	2016	\$819K	\$1,391K	\$1,851K	\$2,499K	\$3,059K	\$3,399K
EMEA	2014	\$279K	\$609K	\$757K	\$1,111K	\$1,389K	\$1,679K
	2015	\$609K	\$1,009K	\$1,212K	\$1,789K	\$2,189K	\$2,589K
	2016	\$909K	\$1,609K	\$1,979K	\$2,907K	\$3,589K	\$3,879K
US	2014	\$709K	\$1,209K	\$1,417K	\$2,149K	\$2,847K	\$2,894K
	2015	\$1,253K	\$2,129K	\$2,509K	\$3,787K	\$4,879K	\$5,044K
	2016	\$2,181K	\$3,729K	\$4,369K	\$6,557K	\$8,111K	\$8,769K

[Documentação](#)

Cartões de KPI

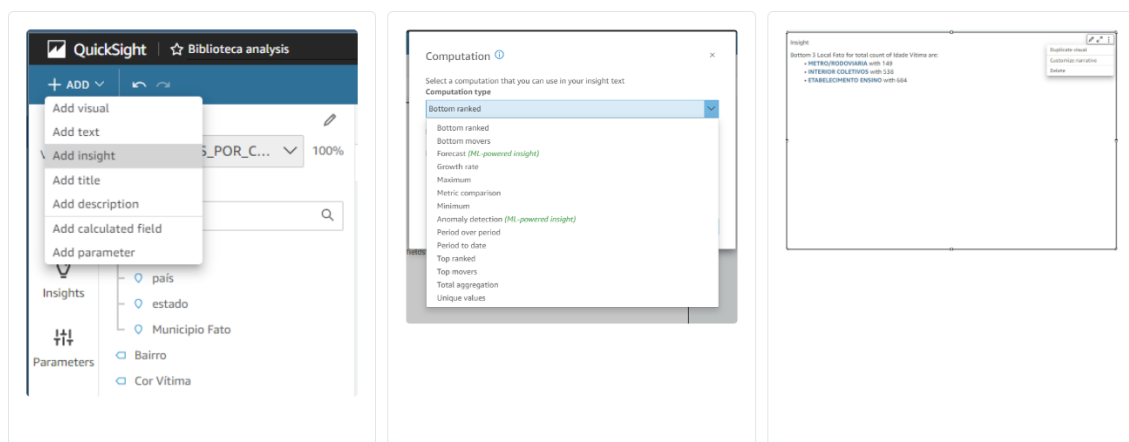
É um tipo de objeto visual que permite comparar uma métrica em relação ao seu valor de referência. São exemplos de aplicação a comparação entre resultado de vendas em relação ao projeto, volume total de compras de um item em relação ao mesmo período do ano anterior, entre outros.



[Documentação](#)

Insights

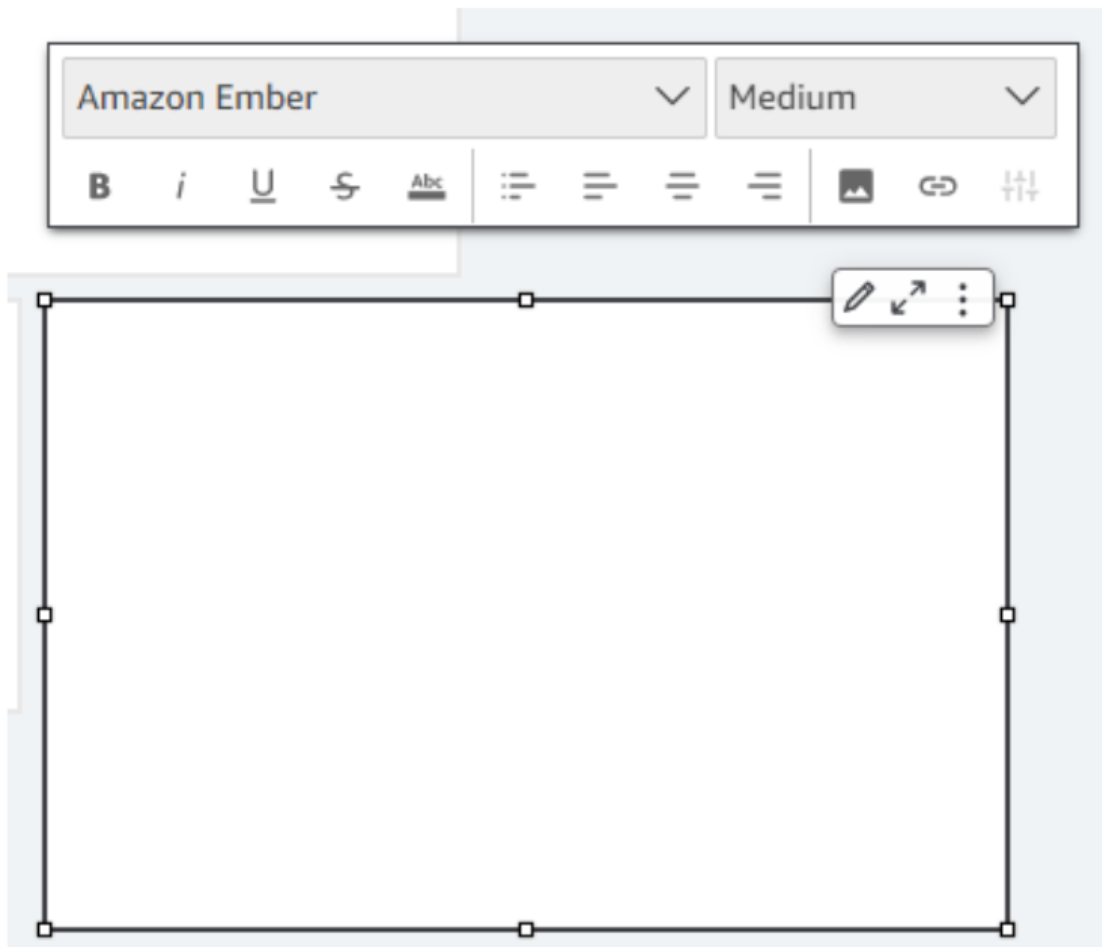
São computações analíticas sugeridas ou customizadas que podem ser adicionadas à análise. Estão disponíveis a partir do menu *Add Insight*, junto às demais opções de adicionar painel, texto, título, etc. Podemos selecionar uma computação pronta ou customizar nossa própria narrativa através da opção *Customize narrative* do menu suspenso.



[Documentação](#)

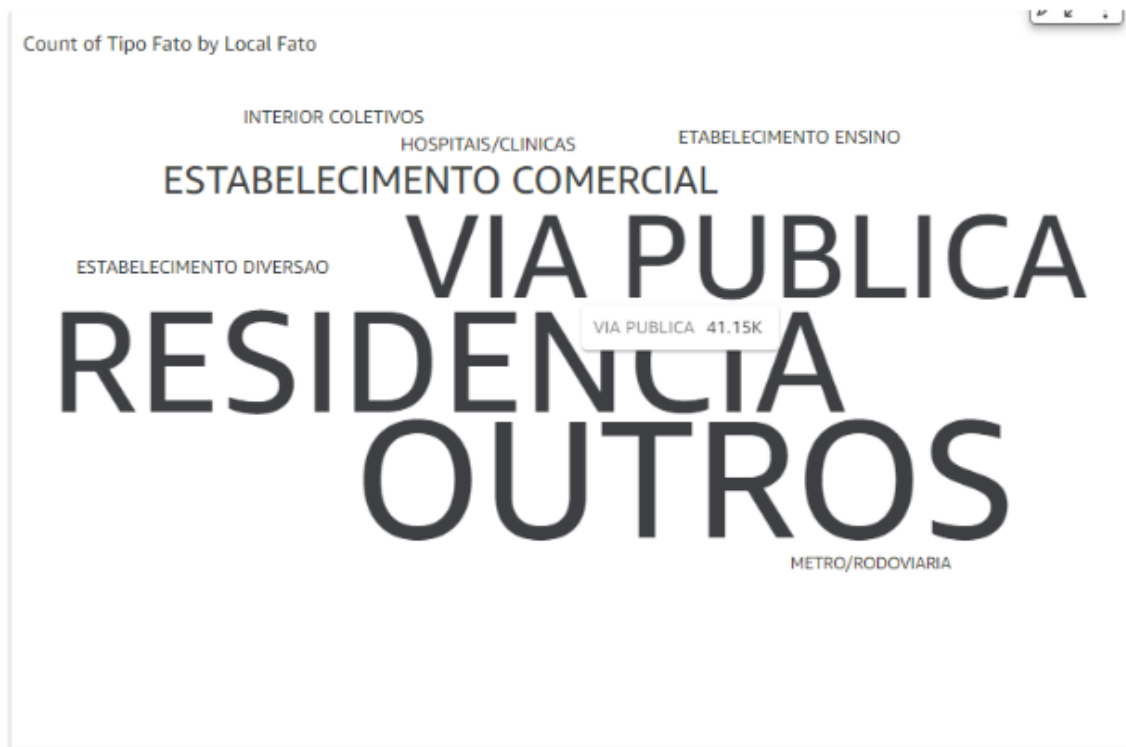
Caixas de Texto

São objetos utilizados para adicionar texto à visualização. Disponibilizam ferramentas de edição, com opção de adicionar imagens, hiperlinks e formatações típicas de editores de texto.



Nuvem de palavras

Apresentam palavras obtidos a partir de uma dimensão, expressando sua importância (quantidade de ocorrências) pelo tamanho.



Desenvolvimento do *Dashboard*

Cores

Em algumas situações as cores do tema/*dashboard* são pré-definidas de acordo com cada empresa/cliente. Quando não há essa possibilidade você será responsável por escolher a paleta de cores que será utilizada. É importante manter a consistência, você não precisa utilizar apenas uma cor para compor um *dashboard* inteiro, porém misturar muitas cores também pode ser um problema. Afinal “O layout e as palavras podem dizer uma coisa, mas as cores podem dizer outra”.

Aqui temos um exemplo onde se usa quatro cores distintas e ainda assim se mantem a consistência, onde as cores dos produtos A, B e C são mantidas independente do gráfico (O produto A sempre será da cor rosa, o B da cor verde e o C da cor laranja).



Fonte: hashtagtreinamentos

Nem sempre é fácil escolher as melhores cores para compor um *dashboard*, pra isso existem algumas ferramentas que podem facilitar (Karpinski, 2022):

- [Adobe Color](#)
- [My Color Space](#)
- [Color Hexa](#)
- [Color Designer](#)
- [Coolors](#)
- [Paleta de Cores](#)
-

Tipografia

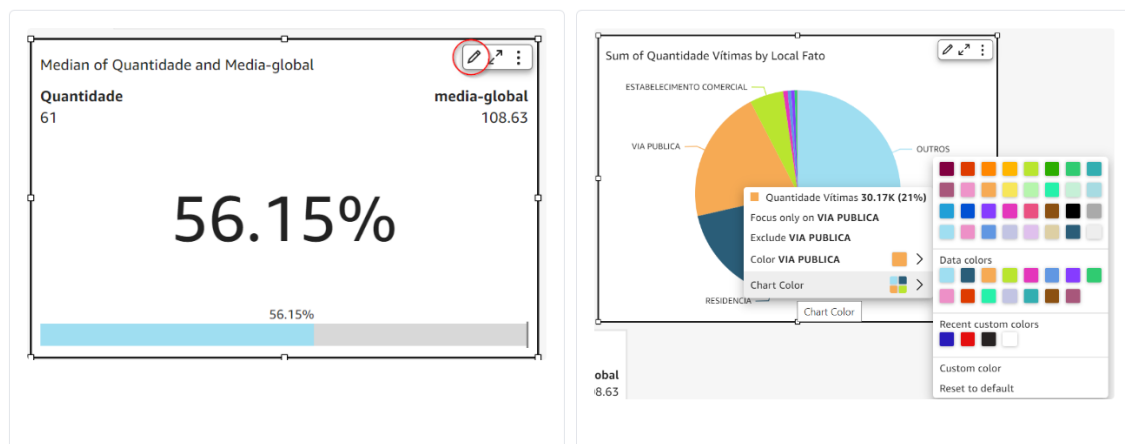
A tipografia abrange todo o estudo, criação e aplicação dos caracteres, estilos, formatos e arranjos visuais das palavras. Por serem a base na comunicação escrita eles precisam ser bem trabalhados para estarem em sintonia com a mensagem que você deseja passar (Rallo, 2018).

Como nas cores, a tipografia pode ser pré-definida, quando isso não acontece se mantém regra da consistência, o ideal é escolher uma tipografia e mantê-la

no *dashboard* inteiro. O tamanho da letra também é importante, escolher letras muito pequenas podem dificultar a leitura, assim como o contrario (letras muito grandes) pode chamar a atenção de maneira equivocada.

Formatar Visualizações no QuickSight

Toda visualização adicionada a uma análise apresenta um conjunto de opções de customização. Para acessá-las é preciso clicar no ícone em formato de lápis no canto superior direito do painel. As opções de formatação são exibidas à esquerda e dependem do tipo de objeto selecionado. Também podemos customizar cores em alguns tipos de gráfico clicando com o botão direito sobre uma área específica.

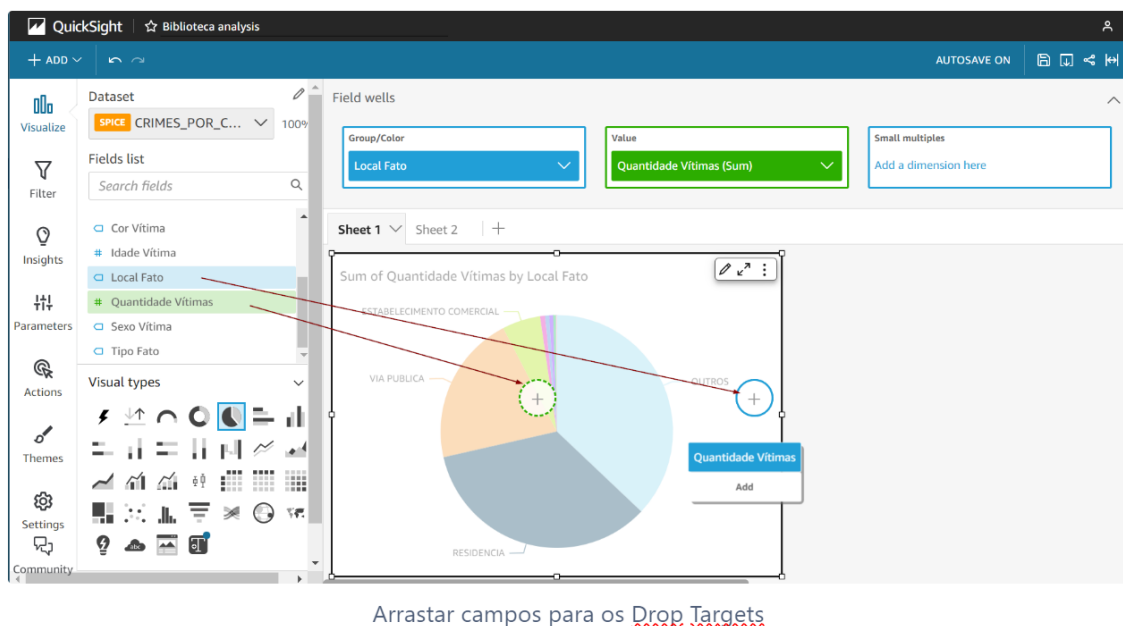
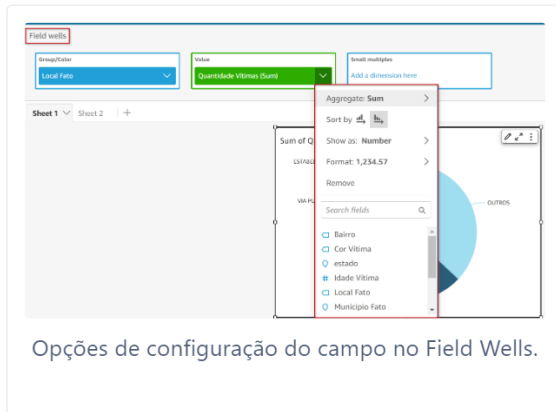
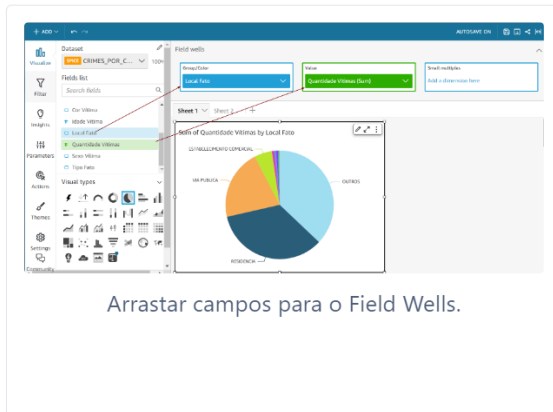


[Documentação](#)

Field wells

Boa parte dos elementos visuais adicionados a uma análise solicitam informar campos no *Field Wells*. É por meio destas configurações que criamos as séries dos gráficos, por exemplo. Logo, os tipos e quantidades de campos solicitados dependem do objeto visual adicionado ao painel. Para adicionar campos aos espaços de interesse, basta arrastá-los do *dataset*. Também é possível fazer

o *drag and drop* diretamente no painel, em espaços predefinidos chamados *drop targets*. Estando no *Field Wells*, algumas configurações podem ser realizadas no campo, como tipo de ordenação, formato de apresentação, função aplicada e tipo de dado.



Temas

Temas são coleções de configurações que se aplicam a várias visualizações e *dashboards*. Além dos temas nativos, podemos criar outros através do editor de temas. É importante considerar que toda a análise deve ter um tema vinculado. Utilizamos temas para uniformizar e adequar as visualizações a requisitos estéticos.



Depois de tudo pronto...

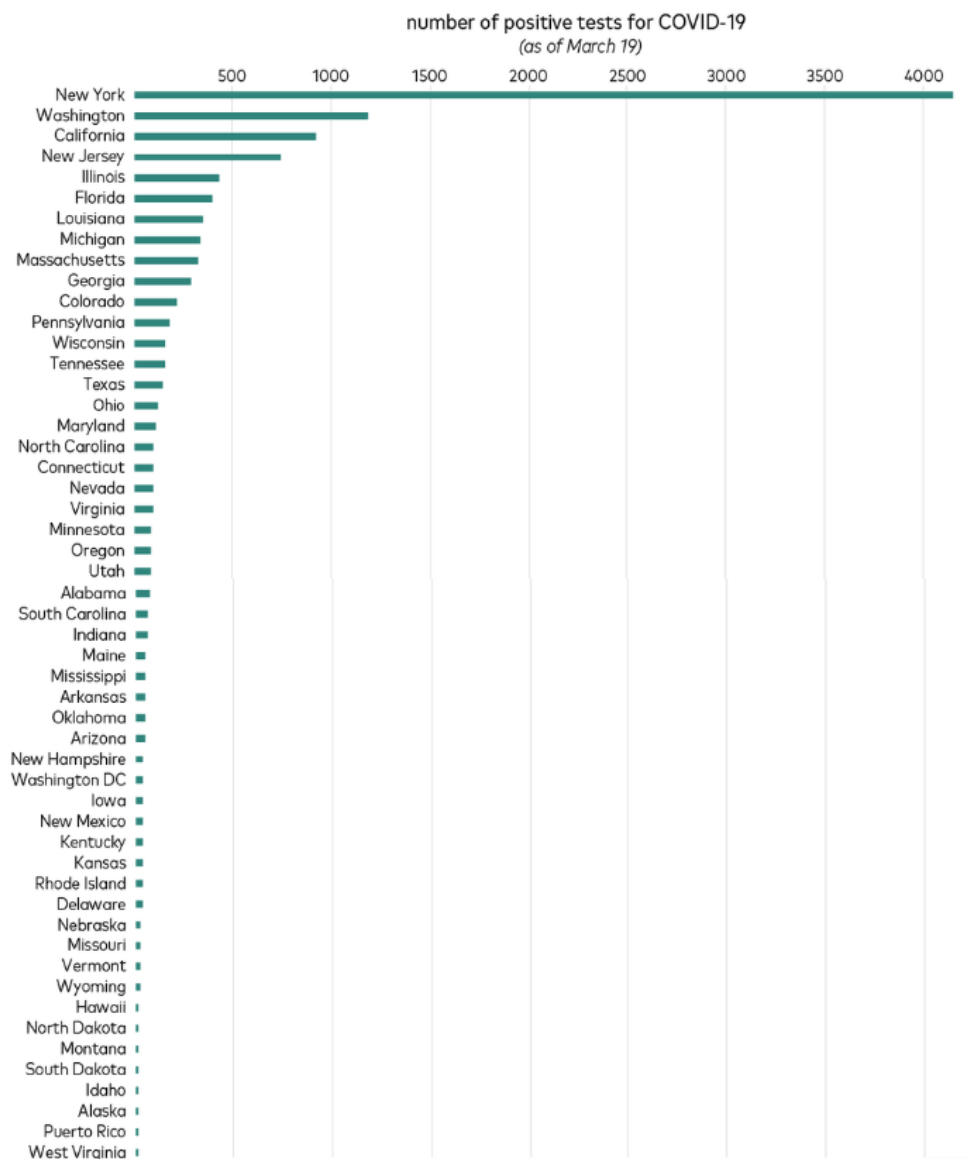
Segundo Karpinski (2022) você pode realizar um “Check” ao concluir seu *Dashboard* para ter certeza que passou por todos os pontos que são importantes:

- **Cor** — Contraste adequado entre background e elementos.
- **Espaço e alinhamento** — Espaços padronizados entre elementos e entre elementos e margens. Alinhamento dos visuais e títulos.
- **Gráficos** — Escolha adequadamente os gráficos e suas propriedades (cor, rótulo, eixo, legenda, etc).
- **Ortografia** — Revise todos os elementos de texto.
- **Usabilidade** — Simplifique! O *Dashboard* deve ser fácil de usar.
- **Acessibilidade** — Não se esqueça de verificar tamanho dos elementos de texto e combinação de cores caso tenha usuários daltônicos. É fundamental conhecer quem vai usar o Dash!
- **Consistência** — Defina um padrão de estilo e siga em todo o *Dashboard*. Isso inclui cor, fonte, nomes de métricas, efeitos de borda, sombra, margens, ícones, etc. (Karpinski, 2022).

O que não fazer!

Exemplos de más escolhas de infográficos:

1º Caso: A falta de contexto: The Atlantic, nos fornecem a quantidade de infectados de COVID-19 nos Estados dos Estados Unidos. Porém de uma maneira limitada, isso porque a pesquisa não leva em consideração por exemplo os infectados que não fazem o teste e consequentemente o infográfico nos leva a uma falsa impressão de que apenas alguns estados tiveram pouquíssimos casos de COVID-19.



Source: The COVID Tracking Project



Fonte: Venngage

2º Caso: A escolha do visual: O gráfico de pizza definitivamente não é o preferido entre os profissionais de visualização de dados, isso porque o gráfico de pizza tem segmentos com tamanhos parecidos o que torna difícil enxergar a diferença entre eles. Neste caso que também se trata da quantidade de casos da COVID-19 nos estados dos Estados Unidos, além da quantidade de categorias e consequentemente a quantidade de cores, os segmentos não tem o valor exato o que dificulta ainda mais de entender de fato quantos casos existem em cada estado.

Confirmed coronavirus cases in the UK

Total cases: 8,077

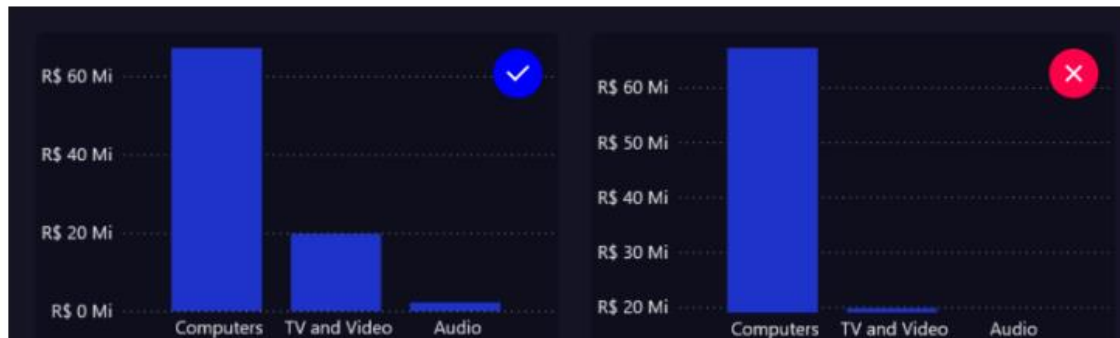
Northern Ireland Scotland Wales London Midlands South East North West
North East and Yorkshire East of England South West



Fonte:

Venngage

3º Caso: Se atente aos detalhes: Aqui temos dois gráficos de barras, que no geral são os preferidos pois são bons para demonstrar dados categóricos, a diferença entre eles é onde se inicia o eixo. O gráfico onde o eixo não inicia em 0 dá a impressão de que não há dados na categoria Áudio.



Fonte: PowerBI Experience

Existem inúmeros outros exemplos do mau uso de infográficos, por isso a atenção ao escolher as visualizações mais adequadas para os eu objetivo.

Cuidado!

Cuidado com a quantidade de gráficos, gráficos com muitas legendas, com muitos auxiliares de rótulos, com muitos elementos, tentem a deixar o *dashboard* carregado de informações o que pode causar confusão para quem irá trabalhar com ele. Claro que “cada caso é um caso”, e por isso a importância de alinhamento para entender qual é o objetivo do *dashboard*.

Referencias

GEROLA, Leticia. **Menos é mais: análise exploratória x análise explanatória.** 2022. Disponível em: <https://medium.com/joguei-os-dados/menos-é-mais-análise-exploratória-x-análise-explanatória-6d6c491e70cc> . Acesso em: 07 jan. 2023.

KARPINSKI, Leonardo. **Storytelling.** 2020. Disponível em: <https://powerbiexperience.com/pt/storytelling/> . Acesso em: 07 jan. 2023.

KNAFLIC, Cole Nussbaumer. **Storytelling com dados**. Rio de Janeiro: Alta Books, 2016. 237 p.

RALLO, Rafael. **Tipografia: como usar um dos pilares do Design Gráfico a seu favor**. 2018. Disponível em: <https://rockcontent.com/br/blog/tipografia/> . Acesso em: 07 jan. 2023.

SOUZA, Alex. **Storytelling e Dataviz em Análise de Dados**. 2020. Disponível em: <https://medium.com/blog-do-zouza/storytelling-em-análise-de-dados-f708cca115bb> . Acesso em: 07 jan. 2023.