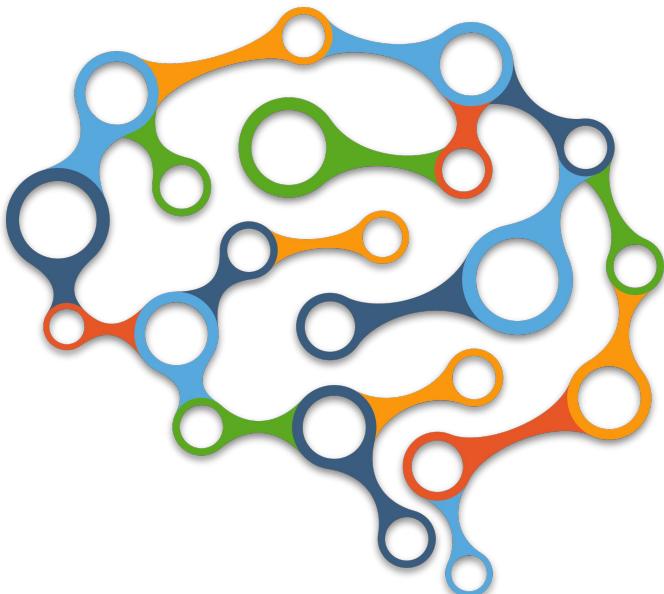


# Dos Dados à Visualização

Minha experiência na construção do painel da Covid-19

Adolfo Guimarães

@profadolfoguimaraes



# Quem sou eu?



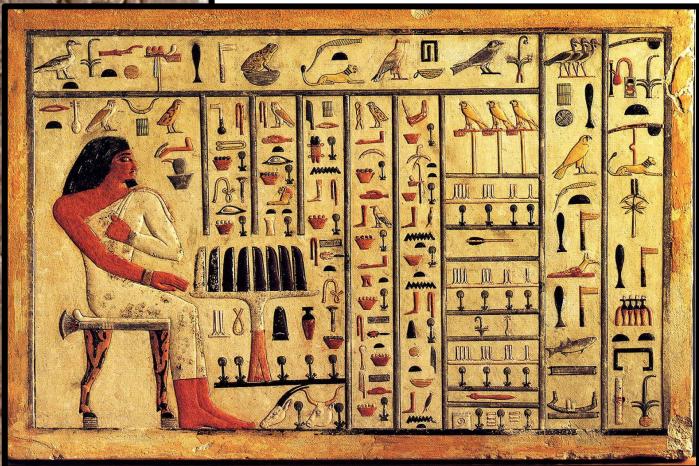
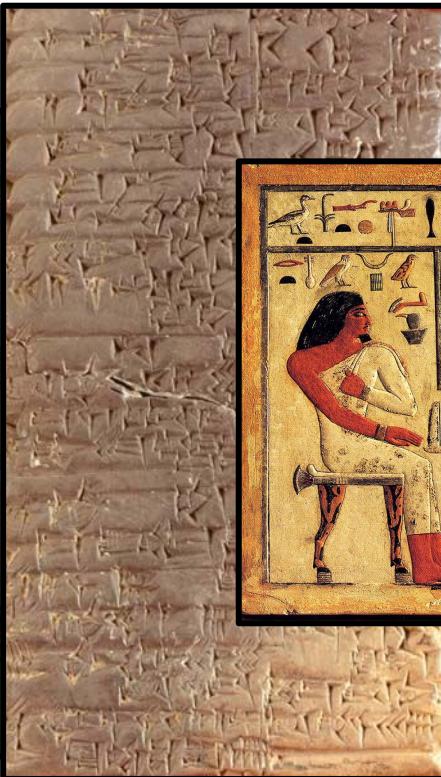
- Brasileiro / Nordestino / Sergipano.
- Ciência da Computação.
- Python.
- Inteligência Artificial.
- Acredito que a tecnologia pode e deve ser utilizada em prol da sociedade.

<https://github.com/adolfoguimaraes>

<https://twitter.com/adolfoguimaraes>

<https://instagram.com/profadolfoguimaraes>

# Não é de hoje que a humanidade produz dados ...



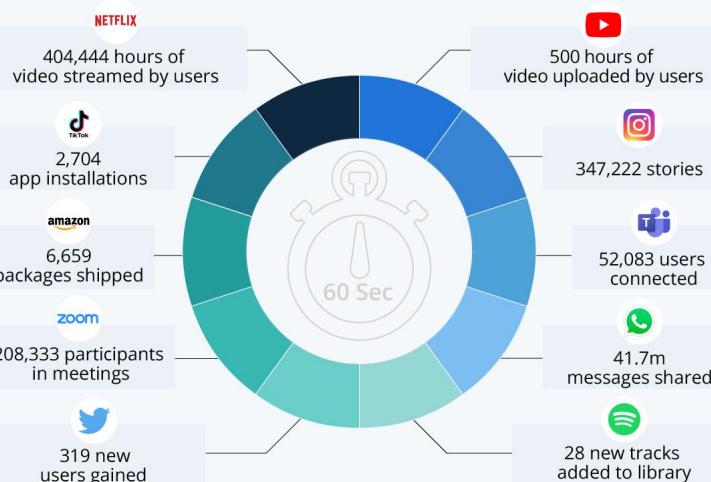
... nos mais variados formatos ...



# ... e continuamos produzindo MUITO dado.

## A Minute on the Internet in 2020

Estimated amount of data created  
on the internet in one minute



Source: Visual Capitalist



statista

The image displays three examples of open data portals:

- DATA.GOV**: Shows a search bar and a map filter for datasets ordered by popularity. One result is highlighted: "Department for the Aging (DFTA) Geriatric Mental Providers".
- dados.gov.br**: Shows a search bar and a sidebar with organization links like Banco Central do Br... (3305), Instituto Nacional ... (1386), and Agência Nacional de... (837). It also shows a count of 10,709 datasets found.
- Brasil.io**: Shows a search bar and a list of datasets, including COVID-19 and Estado da balneabilidade da costa baiana.
- EU Open Data Portal**: Shows a search bar and a large count of 15,735 datasets available. It includes tabs for Home, Data, Applications, Linked data, Visualisations, Developers' corner, and About.

# COLETA



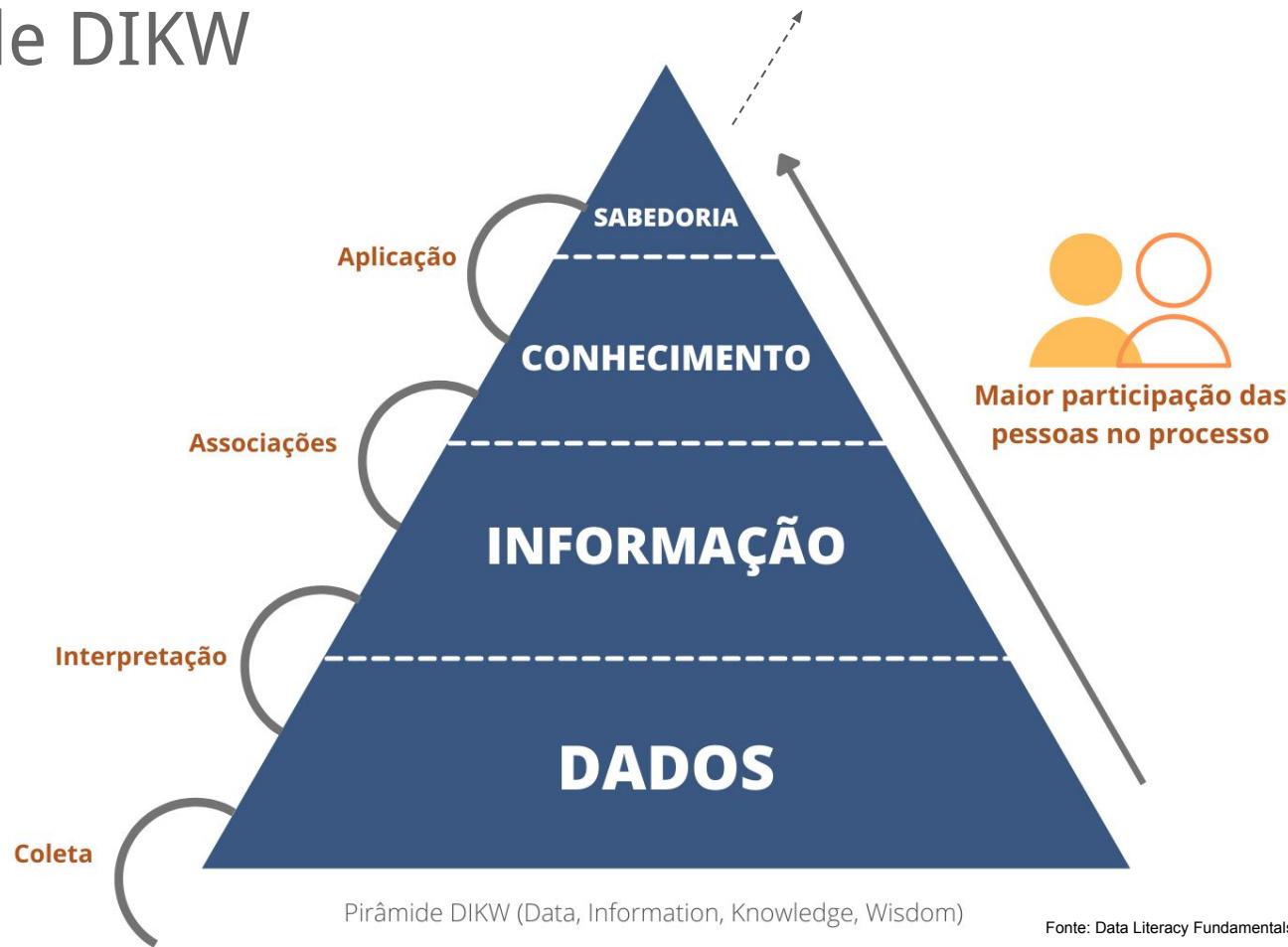
# ANÁLISE



# VISUALIZAÇÃO

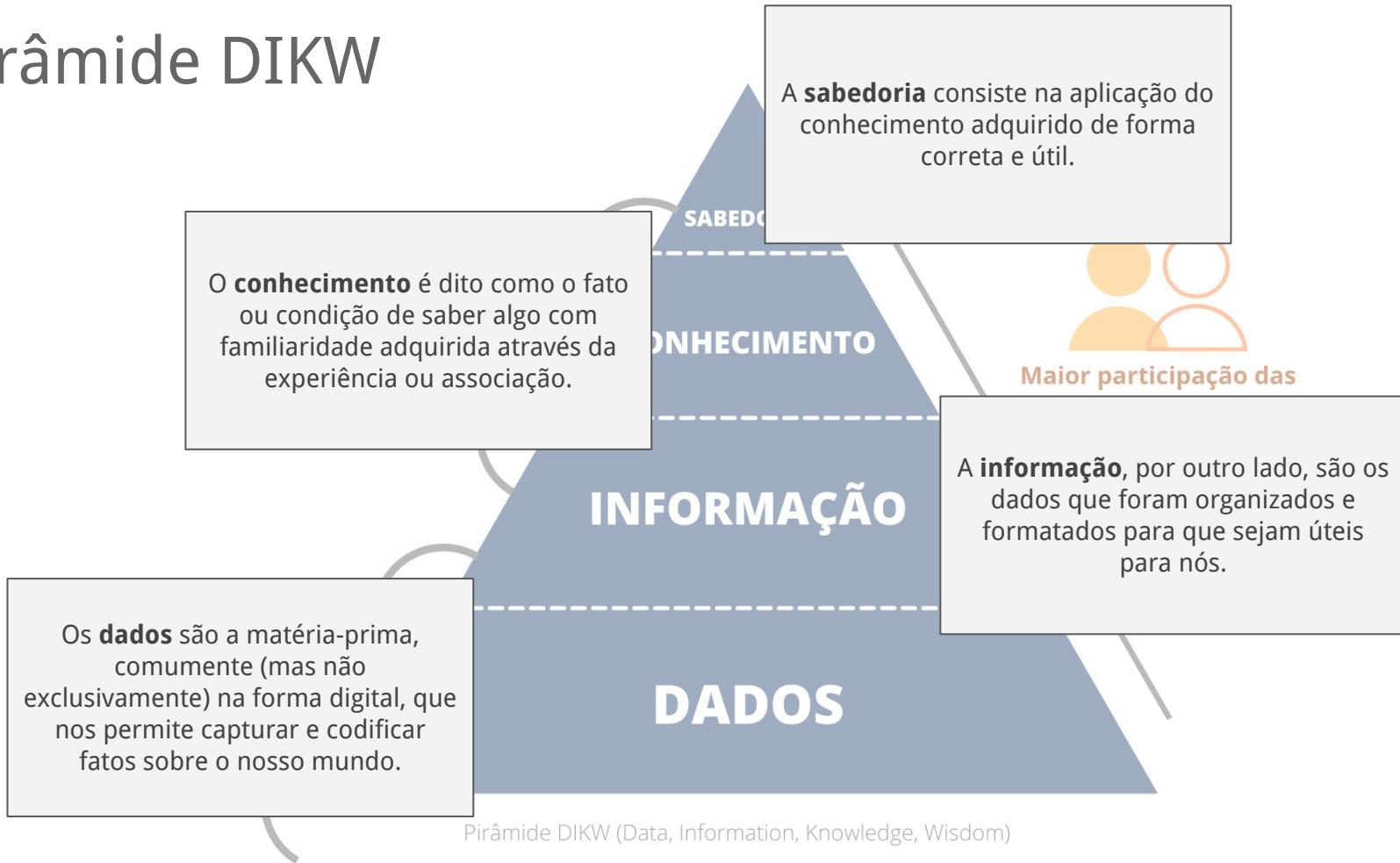


# Pirâmide DIKW

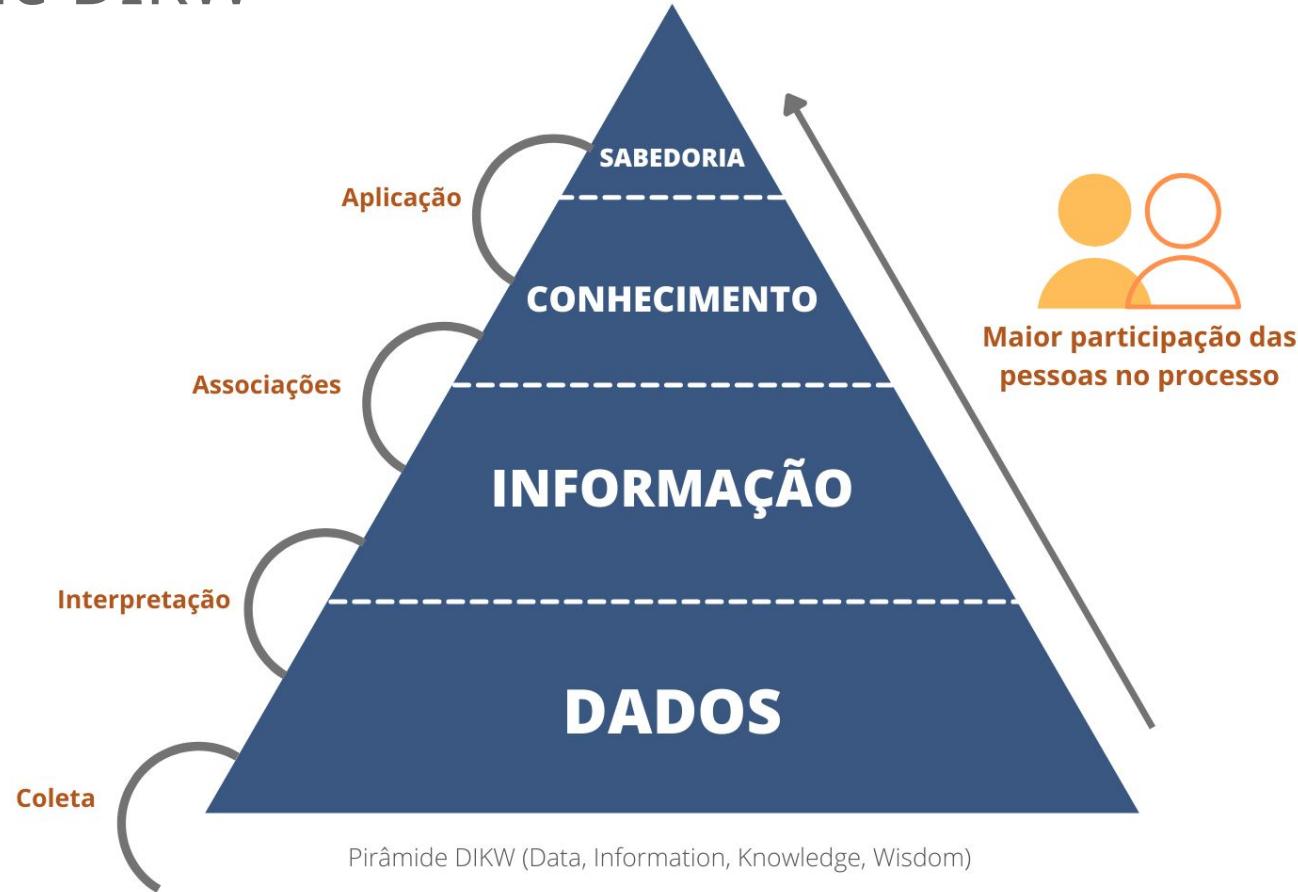


Fonte: Data Literacy Fundamentals: Understanding the Power & Value of Data  
Ben Jones  
<https://www.amazon.com.br/Data-Literacy-Fundamentals-Understanding-English-ebook>

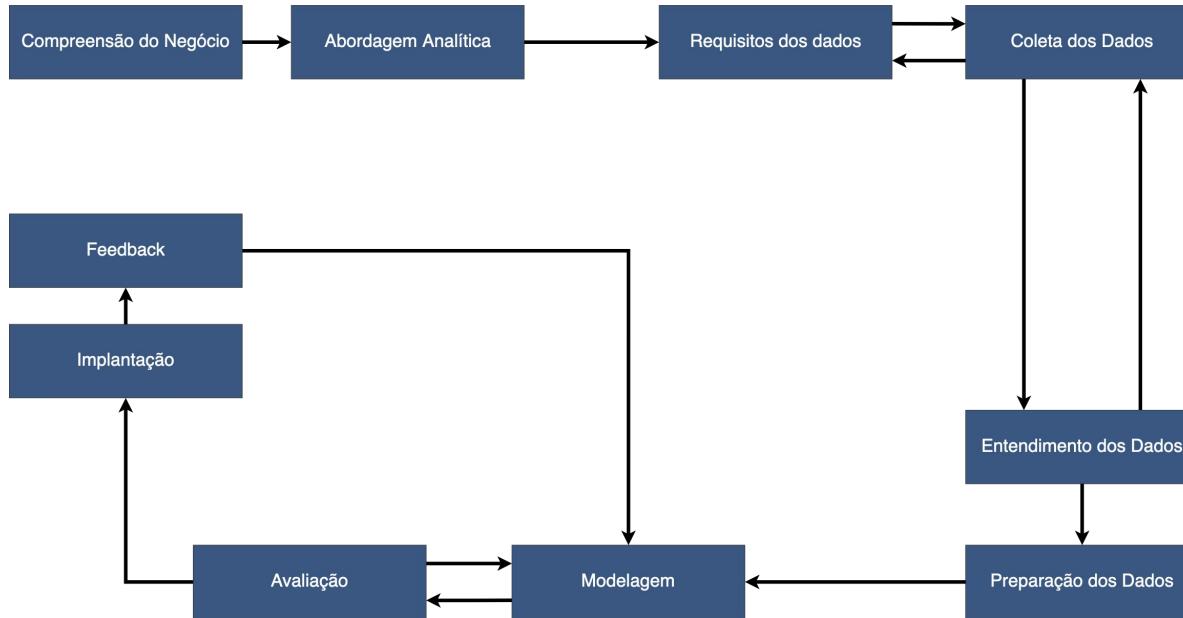
# Pirâmide DIKW



# Pirâmide DIKW



# Metodologia da Ciência dos Dados

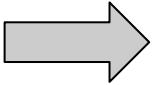


Metodologia proposta por John Rollins (Data Scientist, IBM Analytics, IBM)

<https://www.ibm.com/bigdatahub.com/blog/why-we-need-methodology-data-science>  
<https://www.slideshare.net/JohnBRollinsPhD/foundational-methodology-for-data-science>

# Metodologia da Ciência dos Dados

A metodologia proposta por John Rollins busca responder algumas questões chaves do processo de análise de dados.



## 1. Do problema à abordagem

Qual o problema que você está tentando resolver?

Como você pode usar os dados para responder a questão levantada?

## 2. Trabalhando com os dados

Qual dado você precisa para responder sua questão?

De onde vem os dados e como coletá-los?

Os dados que você coletou são representativos para o problema?

O quanto de trabalho é necessário para manipular e trabalhar com os dados?

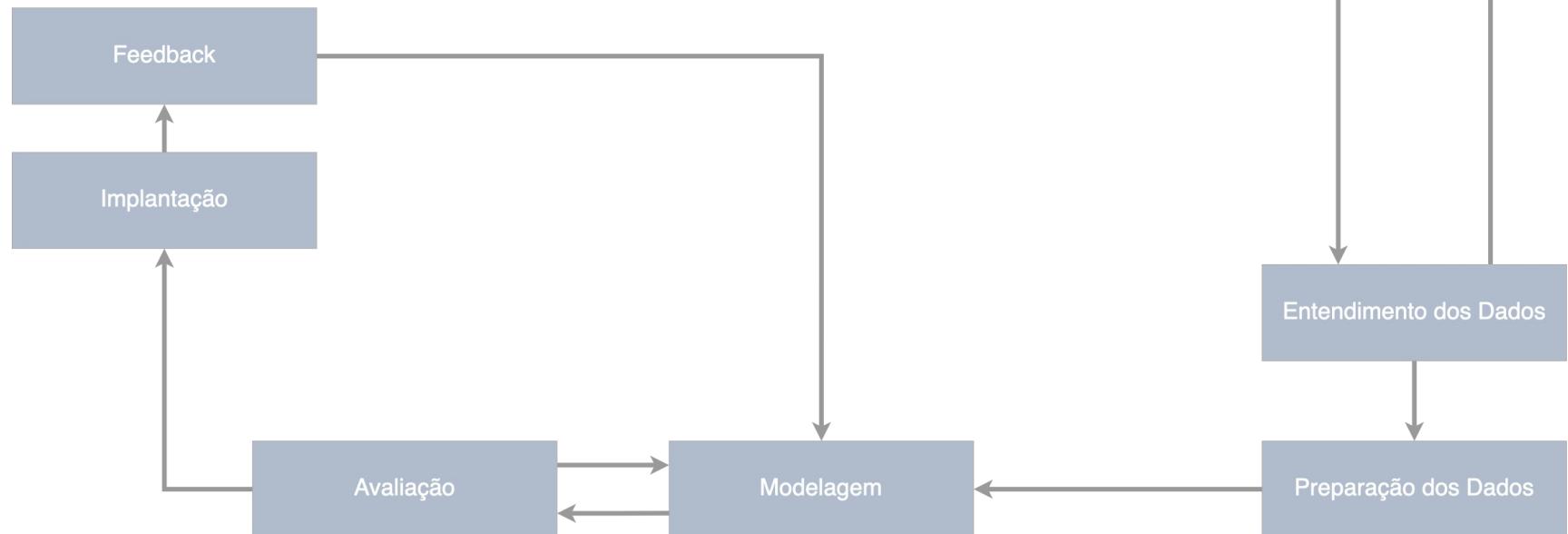
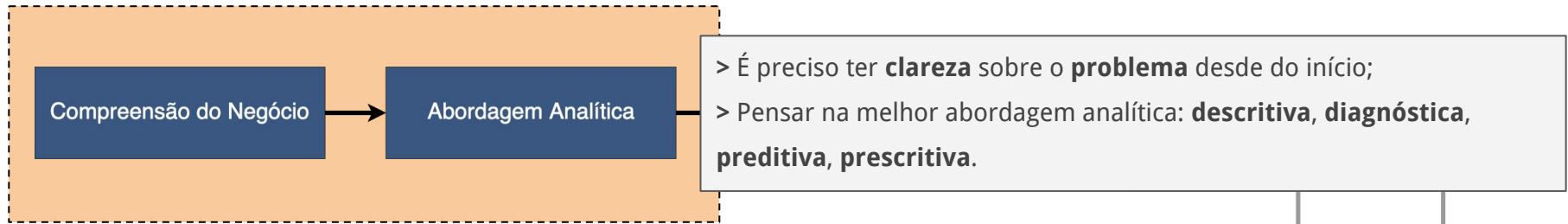
## 3. Encontrando respostas

Qual a melhor forma de visualizar os dados para encontrar respostas?

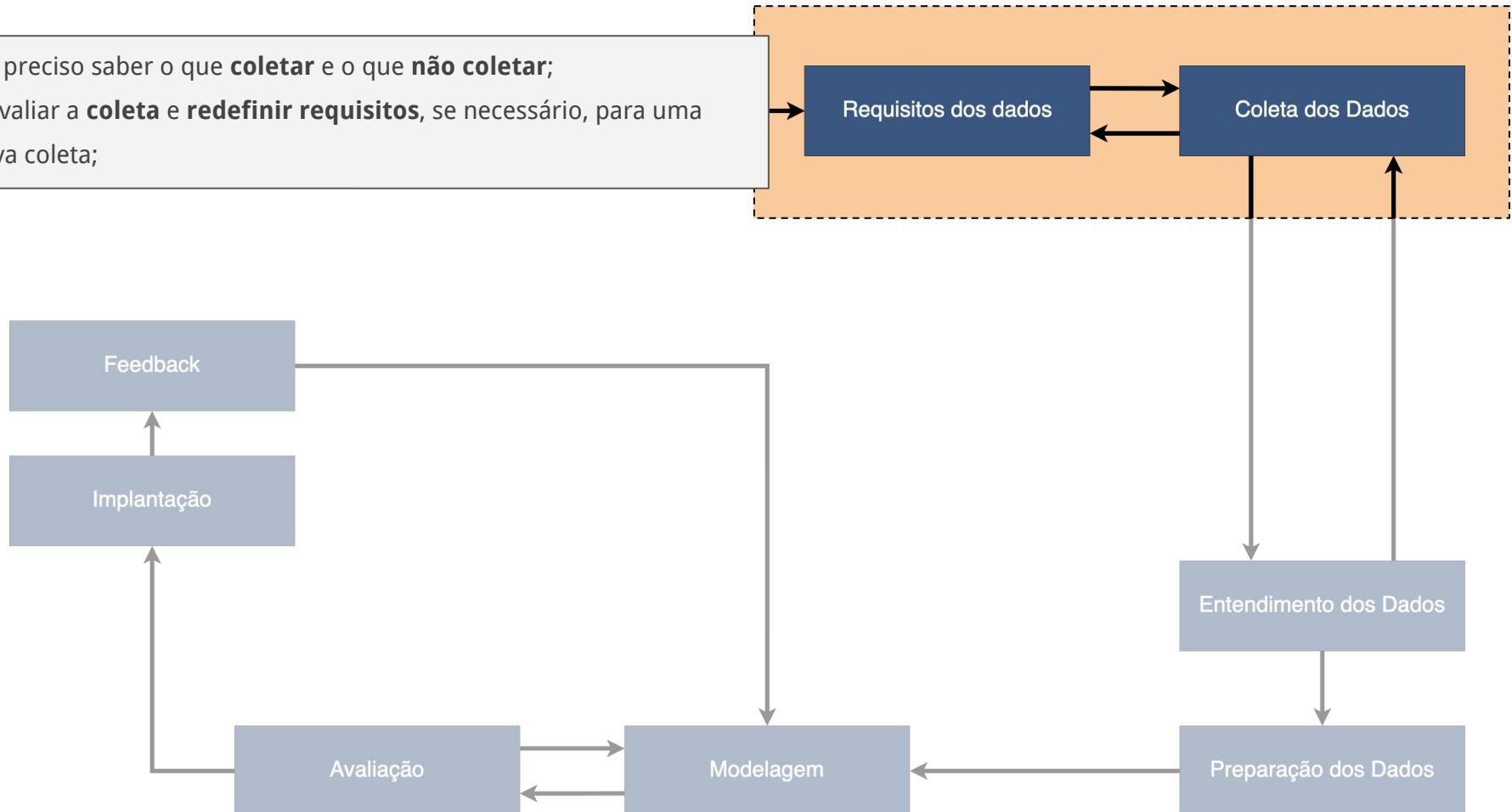
O modelo criado responde o que se quer ou precisa ser ajustado?

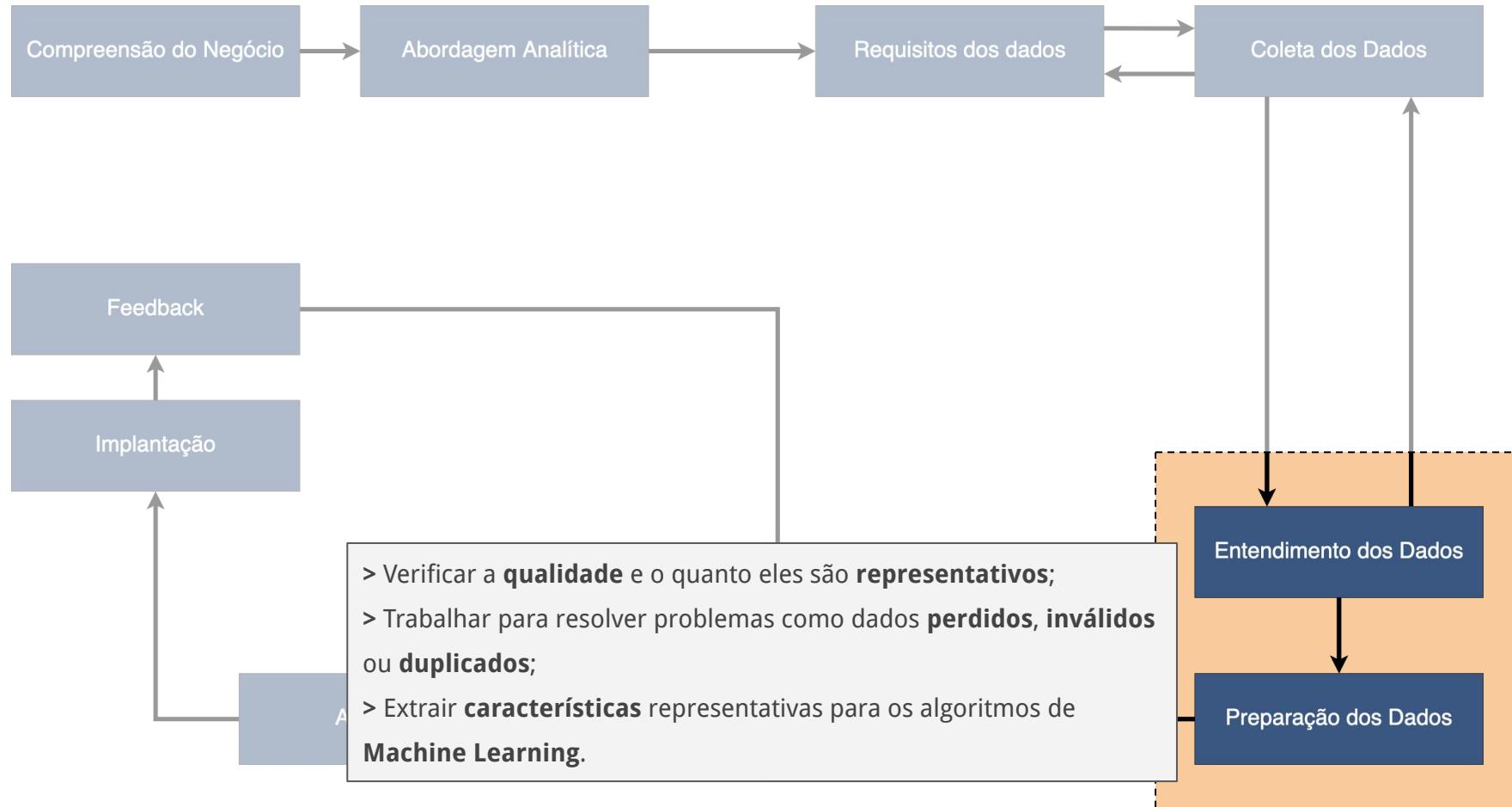
O modelo pode ser usado na prática?

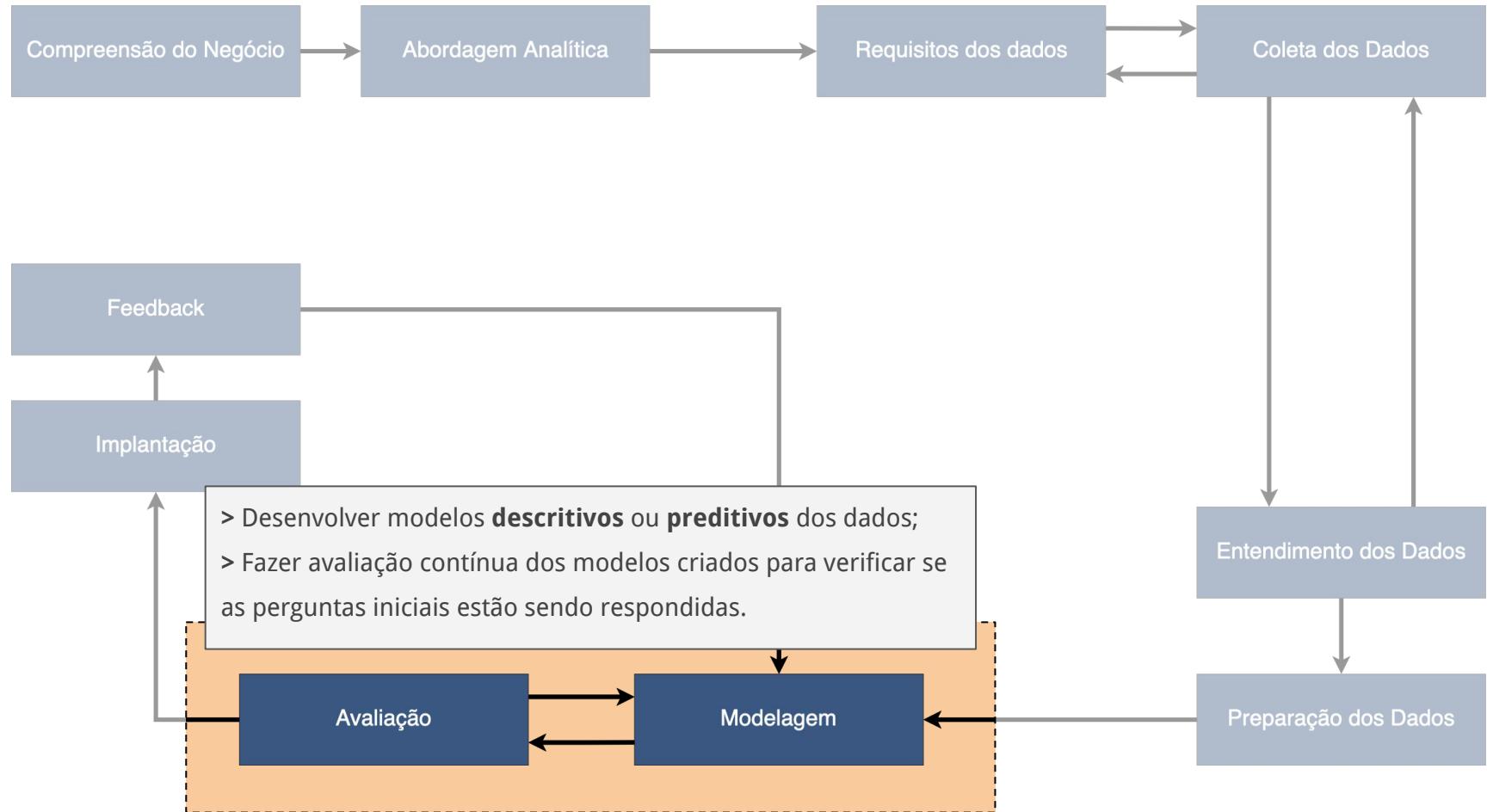
É possível ter feedback das perguntas respondidas?

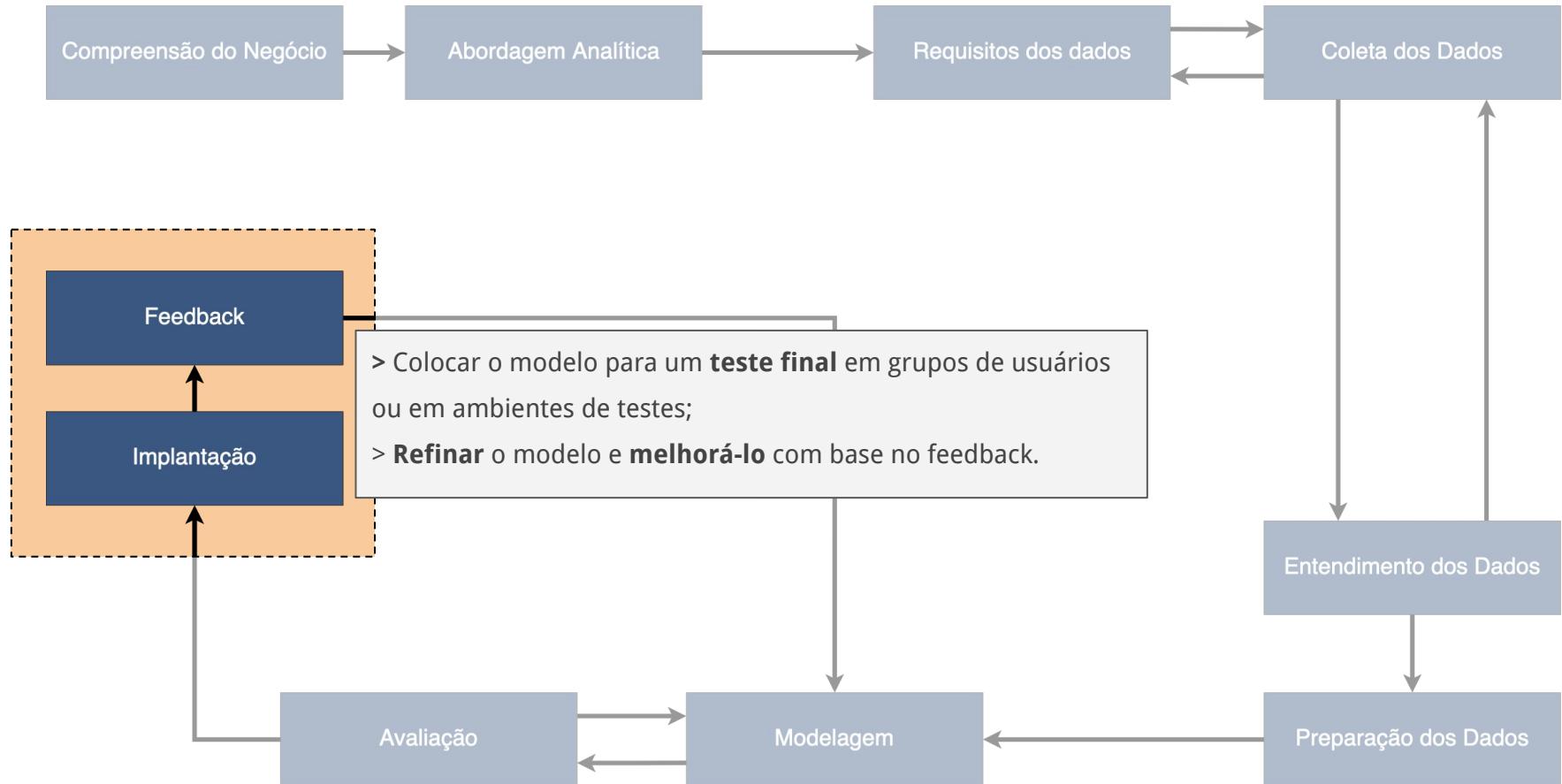


- > É preciso saber o que **coletar** e o que **não coletar**;
- > Avaliar a **coleta** e **redefinir requisitos**, se necessário, para uma nova coleta;









# Metodologia da Ciência dos Dados

## 1. Do problema à abordagem

Qual o problema que você está tentando resolver?

Como você pode usar os dados para responder a questão levantada?

## 2. Trabalhando com os dados

Qual dado você precisa para responder sua questão?

De onde vem os dados e como coletá-los?

Os dados que você coletou são representativos para o problema?

O quanto de trabalho é necessário para manipular e trabalhar com os dados?

## 3. Encontrando respostas

Qual a melhor forma de visualizar os dados para encontrar respostas?

O modelo criado responde o que se quer ou precisa ser ajustado?

O modelo pode ser usado na prática?

É possível ter feedback das perguntas respondidas?

# Metodologia da Ciência dos Dados

## 1. Do problema à abordagem

Qual o problema que você está tentando resolver?

Como você pode usar os dados para responder a questão levantada?

## 2. Trabalhando com os dados

Qual dado você precisa para responder sua questão?

De onde vem os dados e como coletá-los?

Os dados que você coletou são representativos para o problema?

O quanto de trabalho é necessário para manipular e trabalhar com os dados?

## 3. Encontrando respostas

Qual a melhor forma de visualizar os dados para encontrar respostas?

O modelo criado responde o que se quer ou precisa ser ajustado?

O modelo pode ser usado na prática?

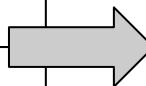
É possível ter feedback das perguntas respondidas?

Pense como um cientista de dados

Formule problemas concretos

Colete e analise os dados

Entenda a questão central a ser respondida



# Metodologia da Ciência dos Dados

## 1. Do problema à abordagem

Qual o problema que você está tentando resolver?

Como você pode usar os dados para responder a questão levantada?

## 2. Trabalhando com os dados

Qual dado você precisa para responder sua questão?

De onde vem os dados e como coletá-los?

Os dados que você coletou são representativos para o problema?

O quanto de trabalho é necessário para manipular e trabalhar com os dados?

## 3. Encontrando respostas

Qual a melhor forma de visualizar os dados para encontrar respostas?

O modelo criado responde o que se quer ou precisa ser ajustado?

O modelo pode ser usado na prática?

É possível ter feedback das perguntas respondidas?

Pense como um cientista de dados

Formule problemas concretos

Colete e analise os dados

Entenda a questão central a ser respondida

Determine os requisitos dos dados

Colete o dado apropriado

Entenda os dados

Prepare-os para os modelos de Machine Learning

# Metodologia da Ciência dos Dados

## 1. Do problema à abordagem

Qual o problema que você está tentando resolver?

Como você pode usar os dados para responder a questão levantada?

## 2. Trabalhando com os dados

Qual dado você precisa para responder sua questão?

De onde vem os dados e como coletá-los?

Os dados que você coletou são representativos para o problema?

O quanto de trabalho é necessário para manipular e trabalhar com os dados?

## 3. Encontrando respostas

Qual a melhor forma de visualizar os dados para encontrar respostas?

O modelo criado responde o que se quer ou precisa ser ajustado?

O modelo pode ser usado na prática?

É possível ter feedback das perguntas respondidas?

Pense como um cientista de dados

Formule problemas concretos

Colete e analise os dados

Entenda a questão central a ser respondida

Determine os requisitos dos dados

Colete o dado apropriado

Entenda os dados

Prepare-os para os modelos de Machine Learning

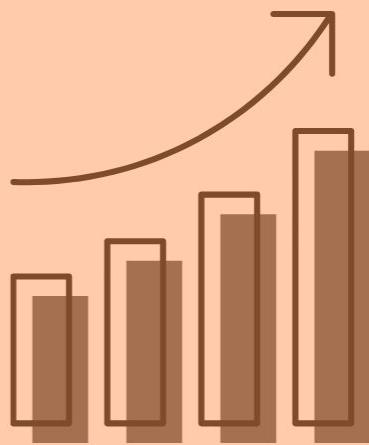
Avalie e implemente o modelo

Receba e analise os feedbacks

Use os feedbacks para melhorar o modelo



Dados coletados e  
analisados. Onde entra a  
visualização nisso tudo?

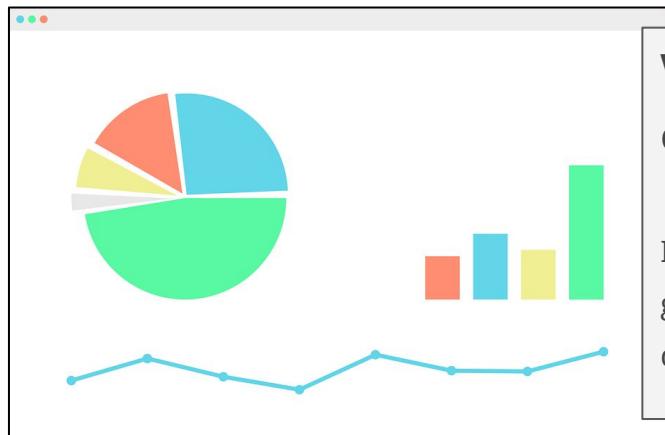


# Dados coletados e analisados. Onde entra a visualização nisso tudo?

A visualização é inerente a diversas etapas do processo de análise de dados. Desde o seu uso para entender os dados a visualização das análises e apresentação dos resultados.

# O que é visualização?

Consiste no estudo e na criação de representações visuais para os dados com o intuito de ajudar as pessoas a realização tarefas como mais eficácia.



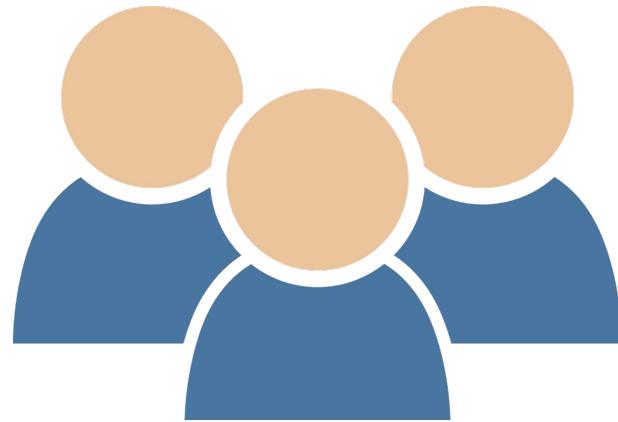
**Visual Literacy:** capacidade de expor e interpretar dados de forma visual.

Isso vai além de construir gráficos. É preciso saber escolher o melhor gráfico para um determinado tipo de dados e entender a mensagem que os dados querem transmitir.

# O que é visualização?

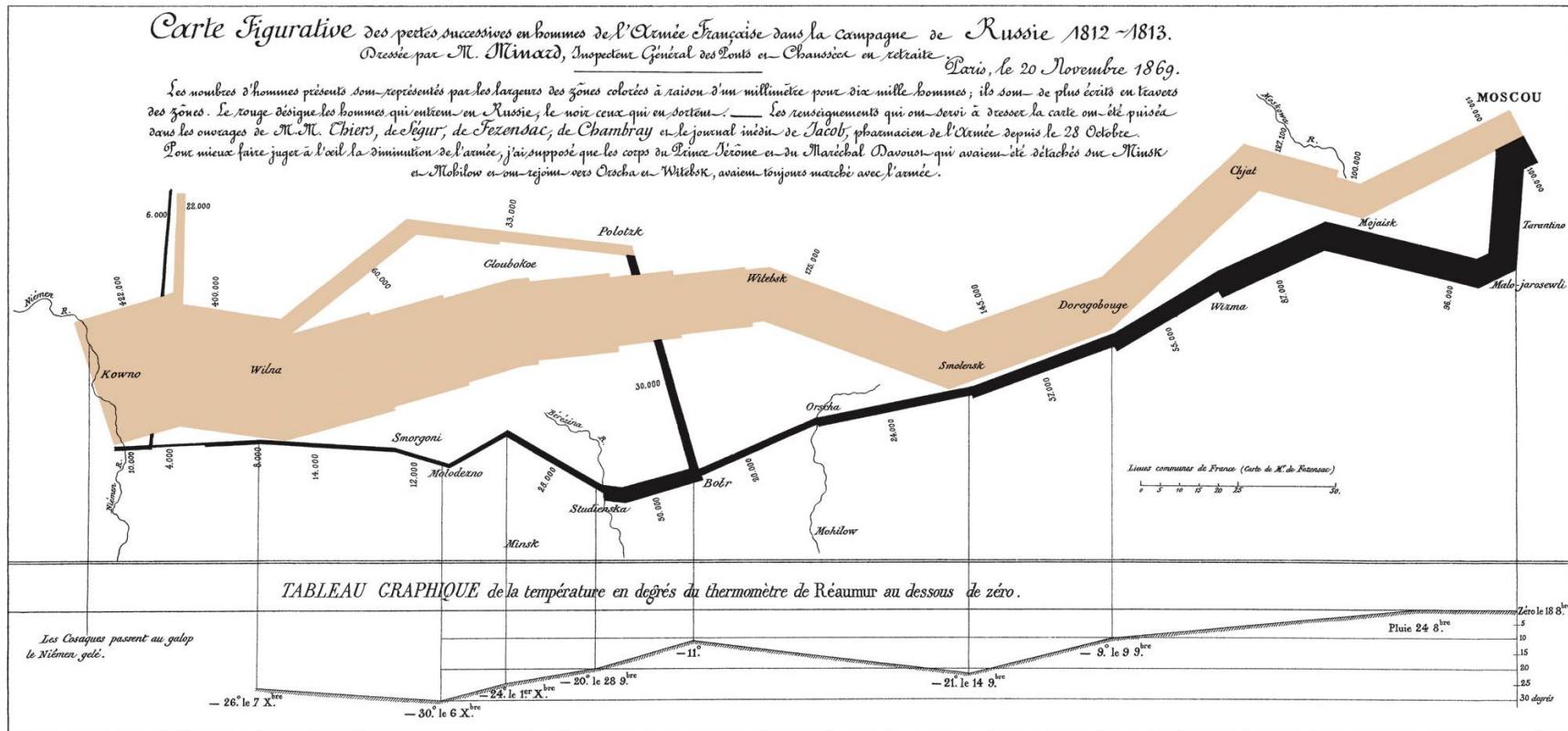
## Ainda sobre a definição de visualização ...

- > A visualização é adequada quando há a necessidade de aumentar as capacidades humanas ao invés de substituí-las por métodos computacionais de tomada de decisão;
- > As possibilidades de designs para construção de visualização (idiomas) são enormes e implica em questões de como criar e como interagir com representações visuais;
- > A maioria das possibilidades de design para construção das visualizações são ineficazes, portanto avaliar a eficácia de um design é muito importante, portanto não é uma tarefa simples;
- > Deve-se levar em consideração 3 limitações: os computadores, os humanos e os meios de visualização;
- > O uso de uma visualização deve ser analisada em termos de "por que o usuário precisa dela", quais dados vão ser exibidos e como será projetada.



**"A escolha do melhor design para as visualizações de dados, vai muito além de questões técnicas. O fator humano, o público e os objetivos são determinantes na seleção da visualização que será adotada".**

# Visualizar dados não é algo novo ...

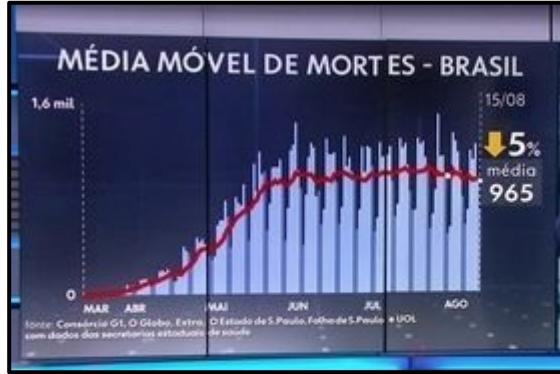


Avec la permission de l'Imp. S<sup>e</sup> Marius S<sup>e</sup> Gédé à Paris.

Minard's Map (1869)

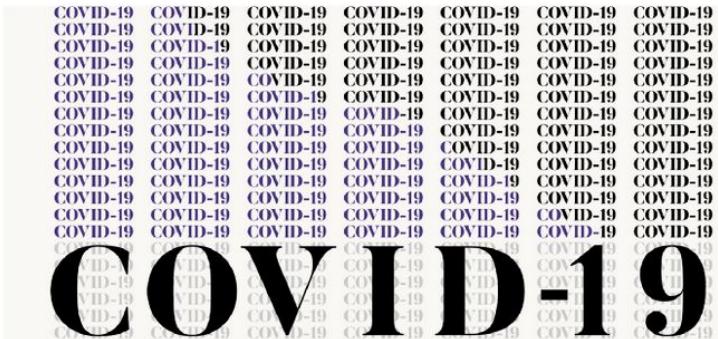
<https://bigthink.com/strange-maps/229-vital-statistics-of-a-deadly-campaign-the-minard-map>

# Mas chamou muita atenção pelo atual momento ...



Todos os dias vemos gráficos de barra, média móvel, choropleth maps e outras visualizações na tv aberta.

# Visualizando dados da COVID-19



## Ten Considerations Before You Create Another Chart About COVID-19

To sum it up — #vizresponsibly; which may mean not publishing your visualizations in the public domain at all

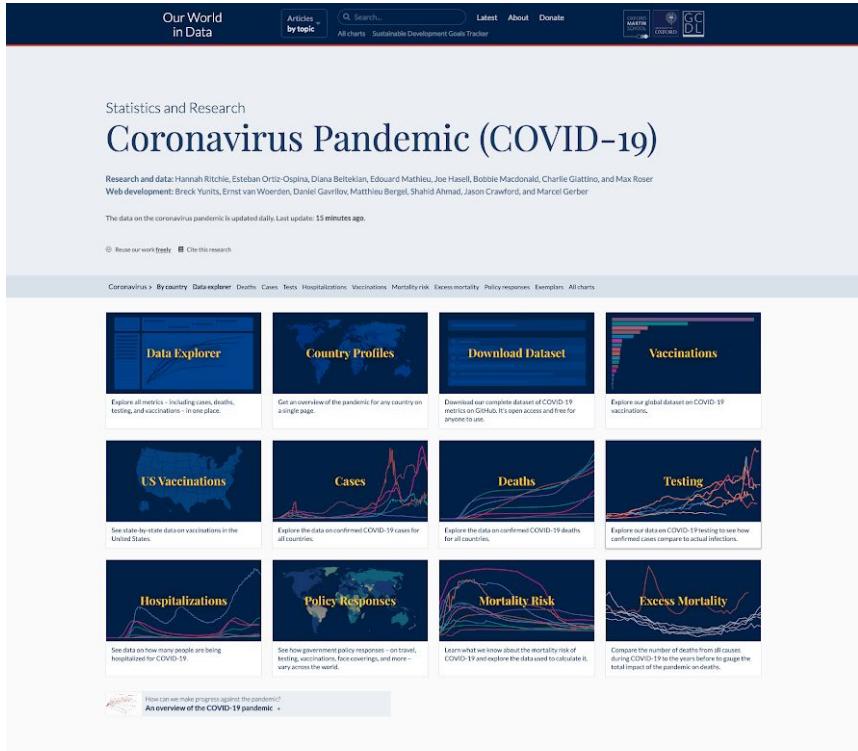
 Amanda Makulec [Follow](#)  
Mar 11, 2020 · 9 min read



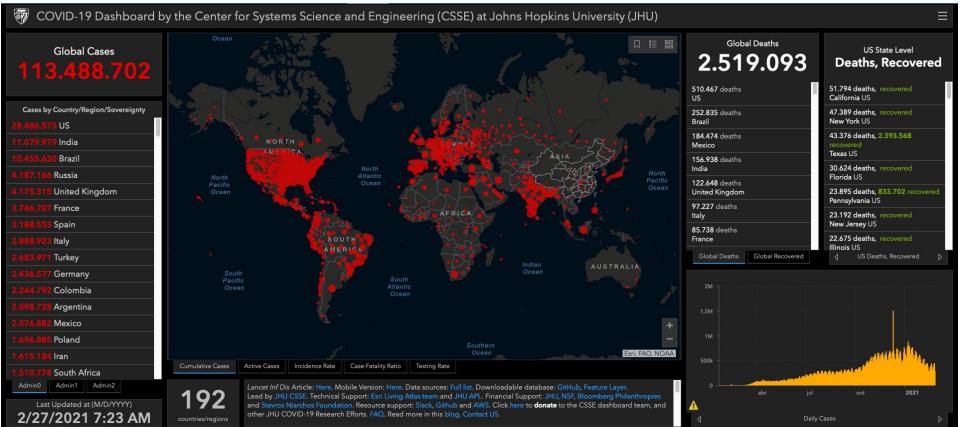
<https://escoladedados.org/2020/04/graficos-covid19/>

<https://medium.com/nightingale/ten-considerations-before-you-create-another-chart-about-covid-19-27d3bd691be8>

# Visualizando dados da COVID-19



<https://ourworldindata.org/>



<https://coronavirus.jhu.edu/>

# Desafios iniciais para construção do painel



## Painel COVID-19 (Atualizado em: 24/02/2021)

Fonte dos dados: <https://brasil.io>

Fonte dos dados de vacinação: <https://coronavirusbrasil.github.io>

Criado por @adolfoguimaraes

**10.331.767**

casos confirmados

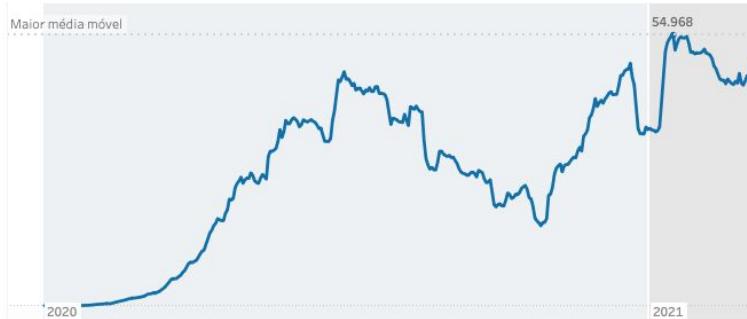
**250.190**

óbitos confirmados

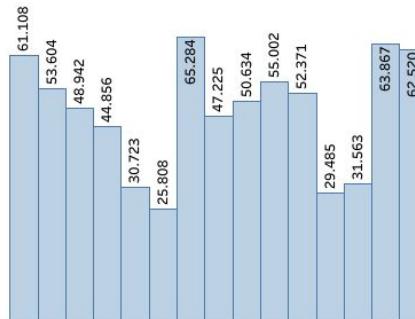
**6.202.055**

vacinados (1ª dose)

### Casos



### Casos por dia (últimos 15 dias)



Selecionar um estado para filtrar

- Acre
- Alagoas
- Amapá
- Amazonas
- Bahia
- Ceará
- Distrito Federal
- Espírito Santo
- Goiás
- Maranhão
- Mato Grosso
- Mato Grosso do Sul
- Minas Gerais
- Pará
- Paraíba
- Paraná
- Pernambuco
- Piauí
- Rio de Janeiro
- Rio Grande do Norte
- Rio Grande do Sul
- Rondônia
- Roraima
- Santa Catarina
- São Paulo
- Sergipe
- Tocantins

### Óbitos



### Óbitos por dia (últimos 15 dias)

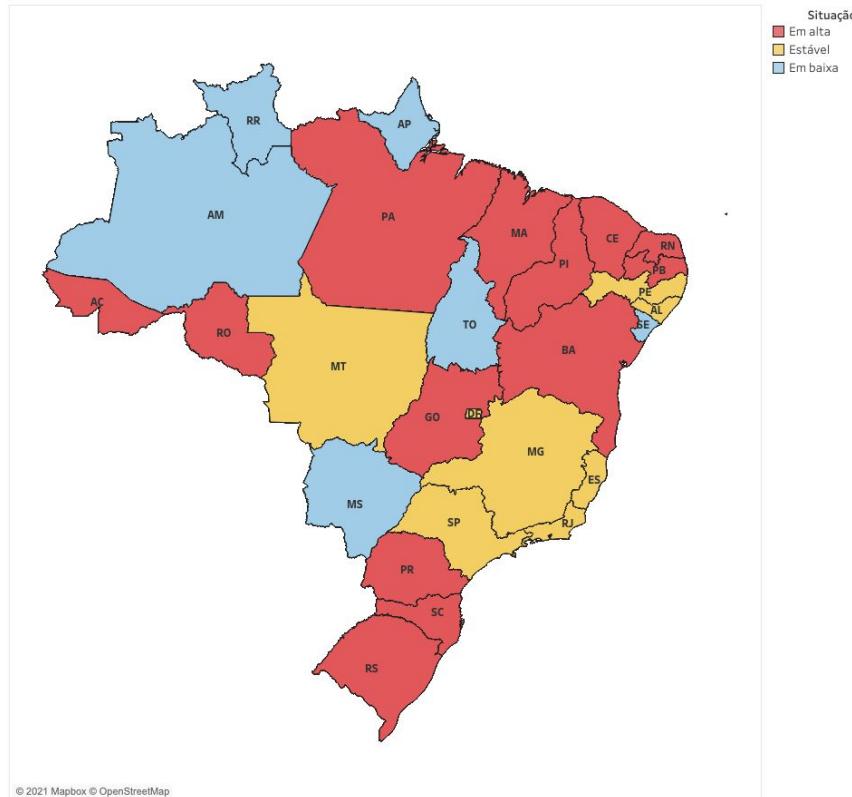


## Mapa da situação de cada estado

(em relação a média de óbitos comparada à 14 dias atrás)

Fonte dos dados: <https://brasil.io>

Data da Atualização: 24/02/2021



## Ocupação das UTIs em Sergipe

Fonte dos dados: Secretaria de Estado da Saúde

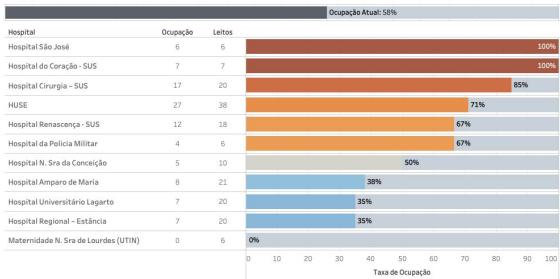
Data da Atualização: 21/02/2021

### UTI Pública

Histórico dos últimos 14 dias:



Ocupação Atual:

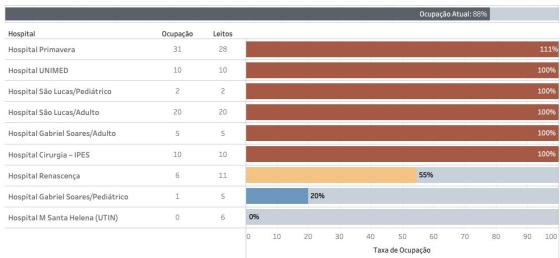


### UTI Privada

Histórico dos últimos 14 dias:



Ocupação Atual:



Criado por @adoteguimaraes

## Vacinação em Sergipe

Fonte dos dados: Secretaria de Estado da Saúde  
Data da Atualização: 23/02/2021

Já foram vacinadas

**44.264**

pessoas em Sergipe

% da população de SE vacinada

1,91%

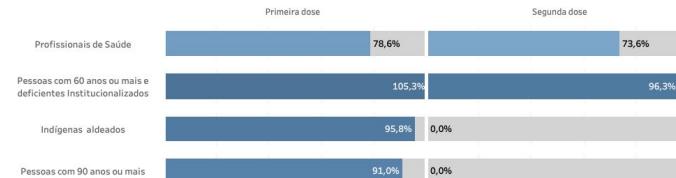
Em Sergipe, das 54.877 vacinas enviadas para primeira dose, já foram aplicadas 44.264. No caso da segunda dose, já foram aplicadas 20.979 das 28.259 enviadas.

80,7% da primeira dose aplicada

74,2% da segunda dose aplicada

## Vacinação por grupos

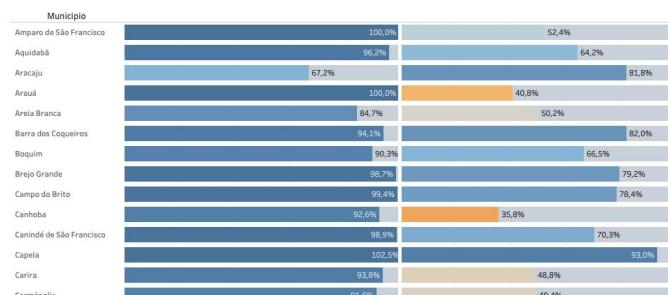
% de aplicação das doses enviadas para cada grupo. Passe o mouse nas barras par ver mais informações.



## Detalhe por município

% de aplicação das doses enviadas para os municípios. Passe o mouse nas barras para ver as informações por grupo.

Filtrar por município  
Tudo



Criado por @adolfoguimaraes

## Painel dos Microdados da Vacinação COVID-19

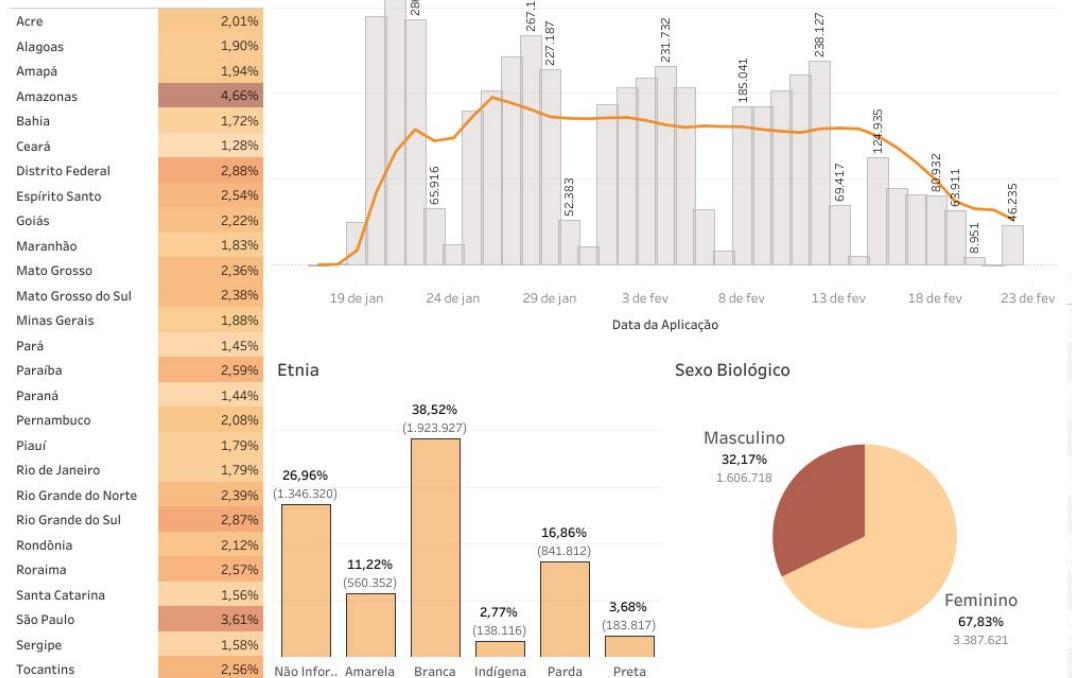
Fonte de dados: arquivo disponibilizado pelo <https://brasil.io> (Atualização: 23/02/2021)  
 Painel criado por @adolfoquimaraes (<https://www.twitter.com/adolfoquimaraes>)

Esse painel usa os microdados disponibilizados pelo SUS e pré-processados pelo site Brasil.io. Os dados ainda possuem uma diferença para os números totais apresentados por outros meios de divulgação.  
 Um painel completo com os dados de casos, óbitos e vacinação pode ser encontrado em  
<https://covid19.data2learning.com>

### Estado (% da população vacinada)

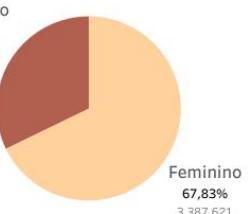
### Aplicação da primeira dose (por dia)

clique no estado para filtrar os dados



### Etnia

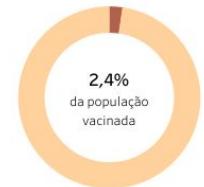
### Sexo Biológico



### Já foram vacinadas

**4.994.344**

pessoas com a primeira dose



### Grupos de Vacinados

Trabalhadores de Saúde	62,59%
Faixa Etária	27,97%
Povos Indígenas	3,81%
Pessoas de 60 anos ou mais institucionalizadas	2,63%
Não Informado	2,22%
Pessoas com Deficiência	0,29%
Comorbidades	0,23%
Povos e Comunidades Tradicionais	0,15%
Forças de Segurança e Salvamento	0,07%
Trabalhadores da Educação	0,04%

# Bases Utilizadas

**Brasil.io**

**COVID-19**

Boletins informativos e casos do coronavírus por município no dia

Fonte original: Secretaria de Saúde estaduais  
Liberado por: Álvaro Justen e dezenas de colaboradores  
Código fonte: <https://github.com/covid19br/covid19br>  
Licença: Creative Commons Attribution-ShareAlike 4.0 International (CC BY-SA 4.0)  
Links referenciados: Boletins PR, Boletins SP, Boletins RO, Boletins MG, Boletins RS, Boletins MT, Boletins MS, Boletins BA, Boletins PE, Informações sobre a coleta de dados (manuais), Boletins AC, Boletins AL, Boletins AM, Boletins AP, Boletins CE, Boletins ES, Boletins GO, Boletins MA, Boletins PA, Boletins PB, Boletins PI, Boletins RJ, Boletins RN, Boletins RR, Boletins SC (1), Boletins SC (2), Boletins DF (1), Boletins DF (2), Boletins SE, Boletins TO, Boletins RJ (2), Documentação da API, Perguntas e respostas sobre os dados, Portal da Transparência do Registro Civil

Tabelas: boletim, caso, caso\_full, óbito, cartorio.

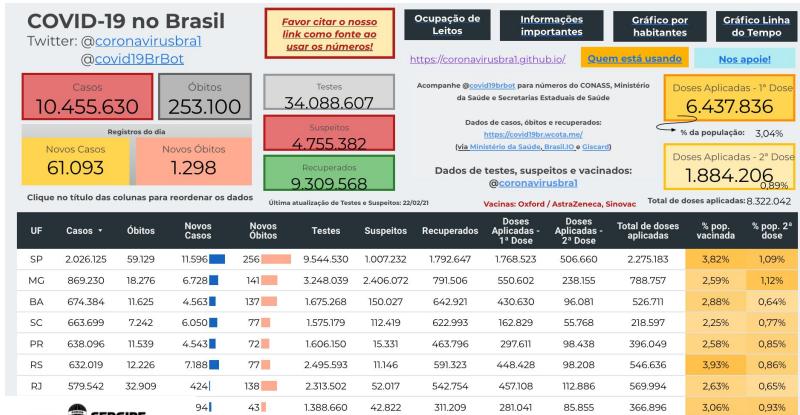
**Informações úteis**

Essa tabela possui os casos confirmados e óbitos obtidos dos boletins das Secretarias Estaduais de Saúde (SES). Os dados foram enriquecidos, de forma que a partir do momento em que um município confirma um caso, ele sempre aparecerá nessa tabela (mesmo que para uma determinada data a SES não tenha liberado o boletim - nesse caso é repeteido o dado do dia anterior). Caso queira ver a tabela original (sem repeteção e com datas faltantes), visite [caso](#).  
Além de acessar os dados por essa interface você pode também [baixar o dataset completo](#) ou [acessá-lo via API](#).

Acesso também ao [Palavrão COVID-19](#)

**ATENÇÃO:** Essa é uma versão beta e nos encorajamos dúvidas sobre os dados:  
Saiba mais sobre esse projeto no [projeto Brasil.io COVID-19](#).  
Leia a documentação desse dataset.  
Leia nossa FAQ sobre esse dataset.  
Caso tenha verificado dado incorreto, verifique se nas "Observações" do boletim para o dia em questão existem informações sobre.  
Caso você não esteja conseguindo abrir o CSV no Excel, [aprenda a fazê-lo](#).

Quer baixar os dados completos ou precisa acessá-los de maneira automatizada? Antes de mais nada, [leia a documentação](#)



<https://brasil.io/>



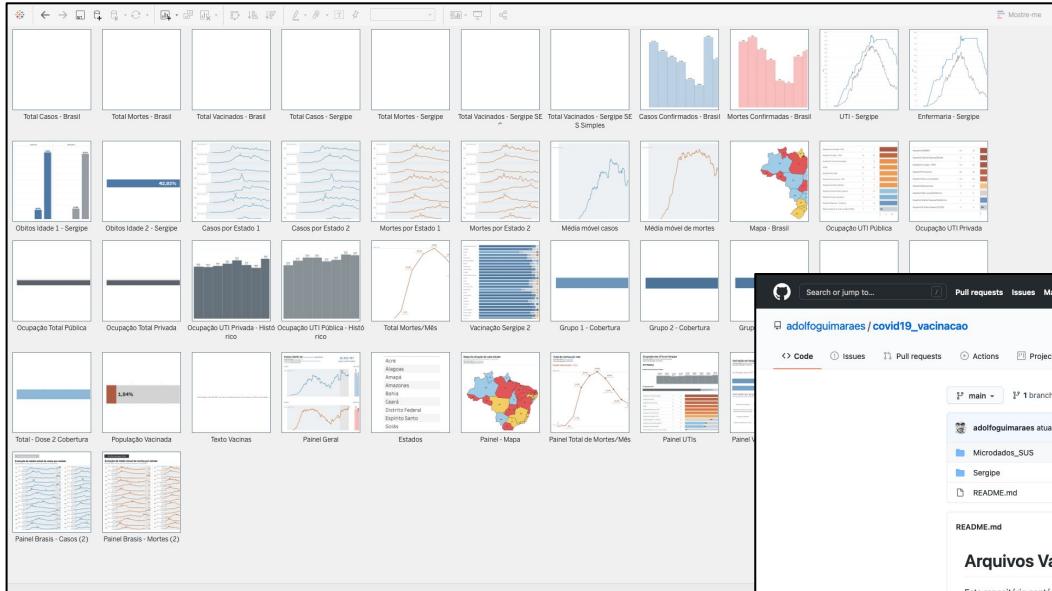
<https://coronavirusbra1.github.io/>

<https://todoscontraocorona.net.br/>

# Ferramentas Utilizadas



# Próximos passos



adolfoguimaraes / covid19\_vacinação

Code Issues Pull requests Actions Projects Wiki Security Insights Settings

main · 1 branch · 0 tags

adolfoguimaraes atualização sergipe 23/02

Microdados\_SUS Atualização no arquivo de microdados 6 days ago

Sergipe atualização sergipe 23/02 3 days ago

README.md Atualização do README 6 days ago

README.md

### Arquivos Vacinação

Este repositório contém os códigos que estou utilizando para processar arquivos relacionados à vacinação da COVID-19.

- Sergipe: Scripts para o processamento do arquivo de vacinação do estado de Sergipe.
- Microdados\_SUS: Scripts para o processamento do arquivo de microdados da vacinação disponibilizados pelo SUS.

-- Desenvolvido por @adolfoguimaraes ([Twitter](#) | [Instagram](#))

About

Este repositório contém os códigos que estou utilizando para processar arquivos relacionados à vacinação da COVID-19.

Issues

Pull requests

Actions

Projects

Wiki

Security

Insights

Settings

Code

Unwatch 1 Star 0 Fork 0

No releases published Create a new release

Releases

No packages published Publish your first package

Packages

Languages

Jupyter Notebook 100.0%

© 2021 GitHub, Inc. Terms Privacy Security Status Docs Contact GitHub Pricing API Training Blog About

# E para finalizar ...



... isso não é mágica ...

# E para finalizar ...



... isso não é  
ficção científica ...

# E para finalizar ...



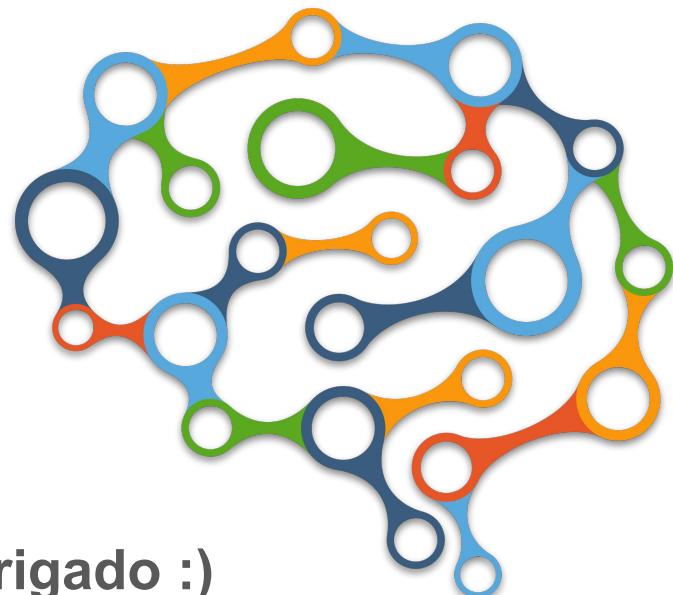
... isso é computação!

# Dos Dados à Visualização

Minha experiência na construção do painel da Covid-19

Adolfo Guimarães

@profadolfoguimaraes



Obrigado :)