

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 ...



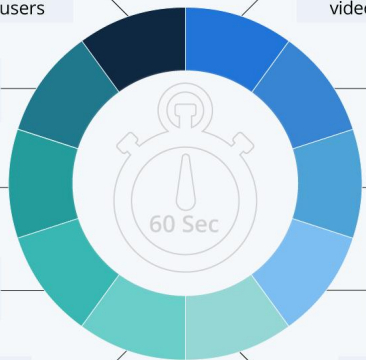
... nos mais variados formatos ...



... 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

DATA.GOV DATA TOPICS - RESOURCES STRATEGY DEVELOPERS CONTACT

DATA CATALOG

Search datasets...

Datasets ordered by Popular

Filter by location

193,205 datasets found

Department for the Aging (DFTA) Geriatric Mental Providers

City of New York - Listing of DFTA Geriatric Mental Health Centers of Operation.

Low-altitude aerial imagery (UAS) flights over Black E...

dados.gov.br

PORTAL BRASILEIRO DE DADOS ABERTOS

Dados | Organizações | Aplicativos | Inventários | Concursos | INDA | Perguntas frequentes

Conjuntos de dados

Organizações

Banco Central do Br... (3305)

Instituto Nacional ... (1386)

Agência Nacional de... (637)

10.709 conjuntos de dados encontrado(s)

Brasil.io

Datasets

COVID-19

Estado da balneabilidade da costa baiana

EU Open Data Portal

Access to European Union open data

EUROPA > EU Open Data Portal

Home Data Applications Linked data Visualisations Developers' corner About

Search datasets...

Show results with:

all of these words | any of these words | the exact phrase

Total datasets available: 15735

COLETA



ANÁLISE

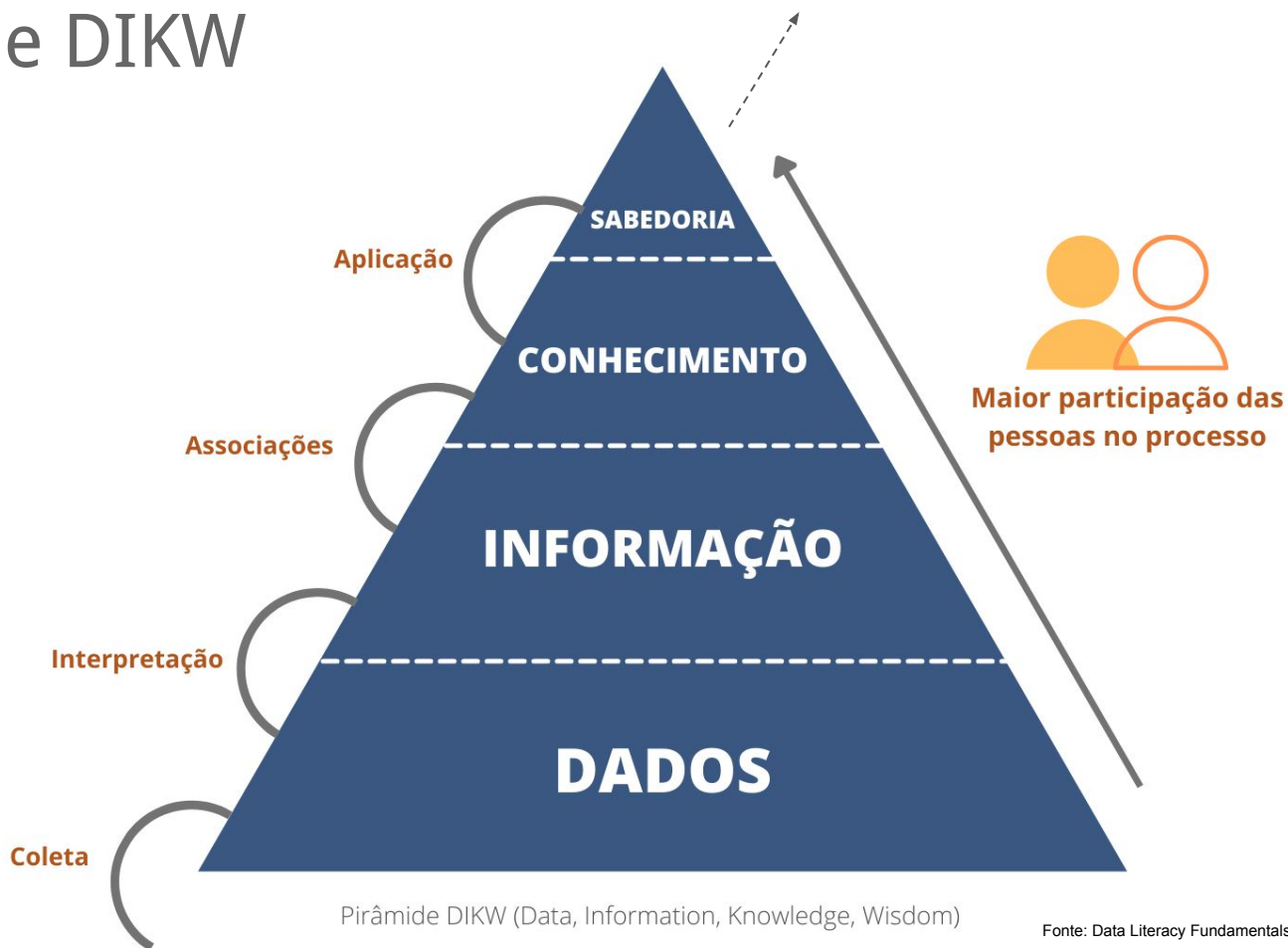


VISUALIZAÇÃO



Pirâmide DIKW

wisdom: the ability to use your knowledge and experience to make good decisions and judgments.

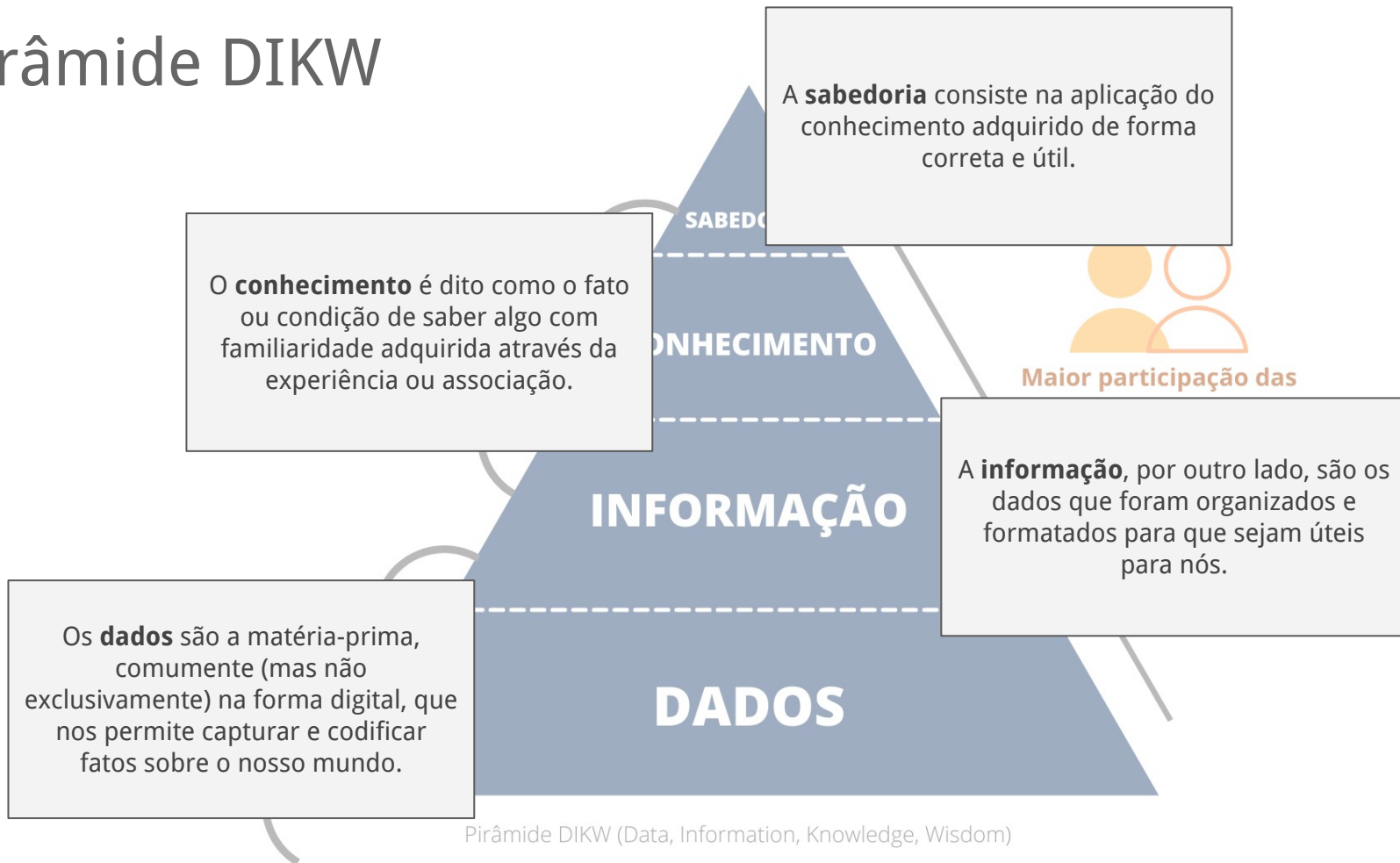


Pirâmide DIKW (Data, Information, Knowledge, Wisdom)

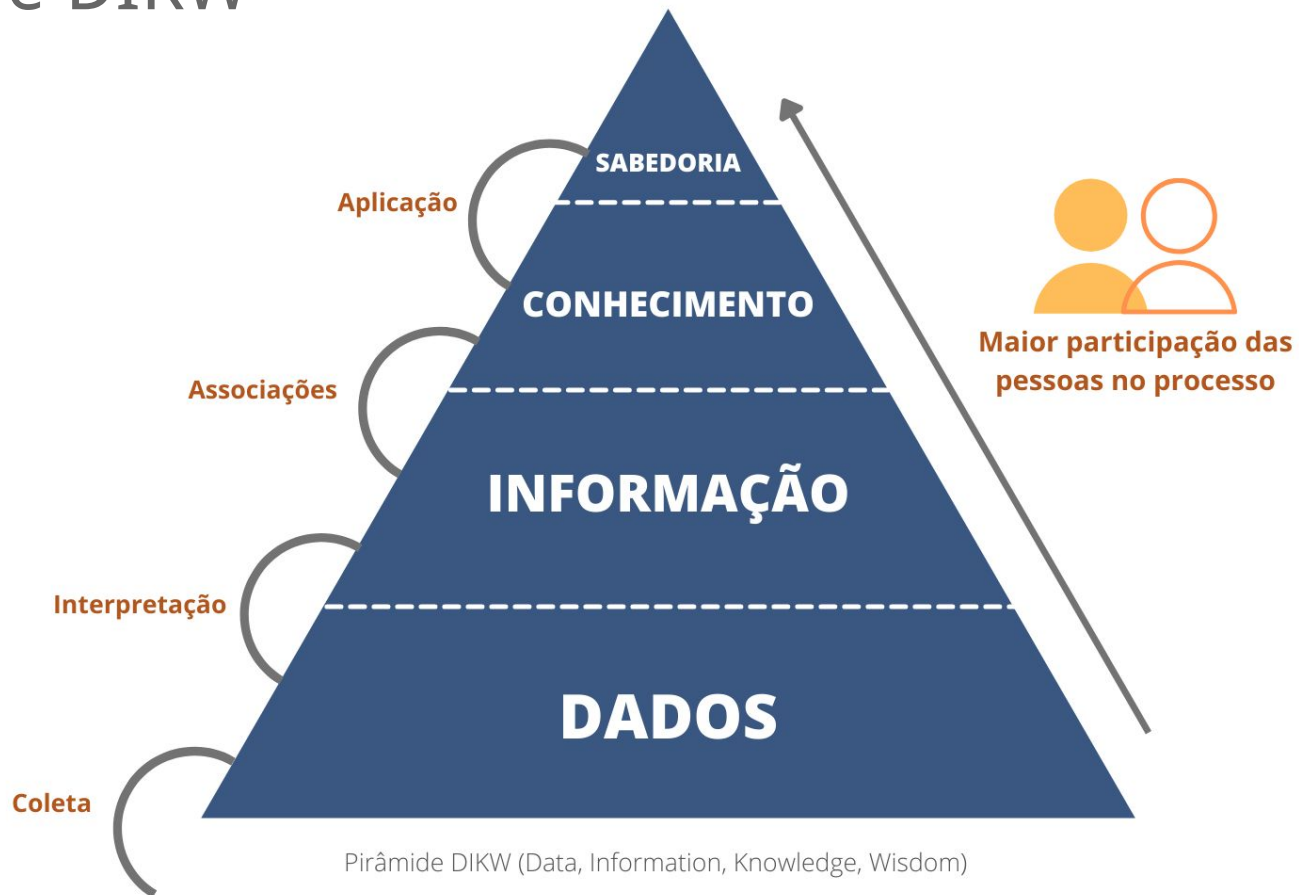
Fonte: Data Literacy Fundamentals: Understanding the Power & Value of Data
Ben Jones

<https://www.amazon.com.br/Data-Literacy-Fundamentals-Understanding-English-ebook>

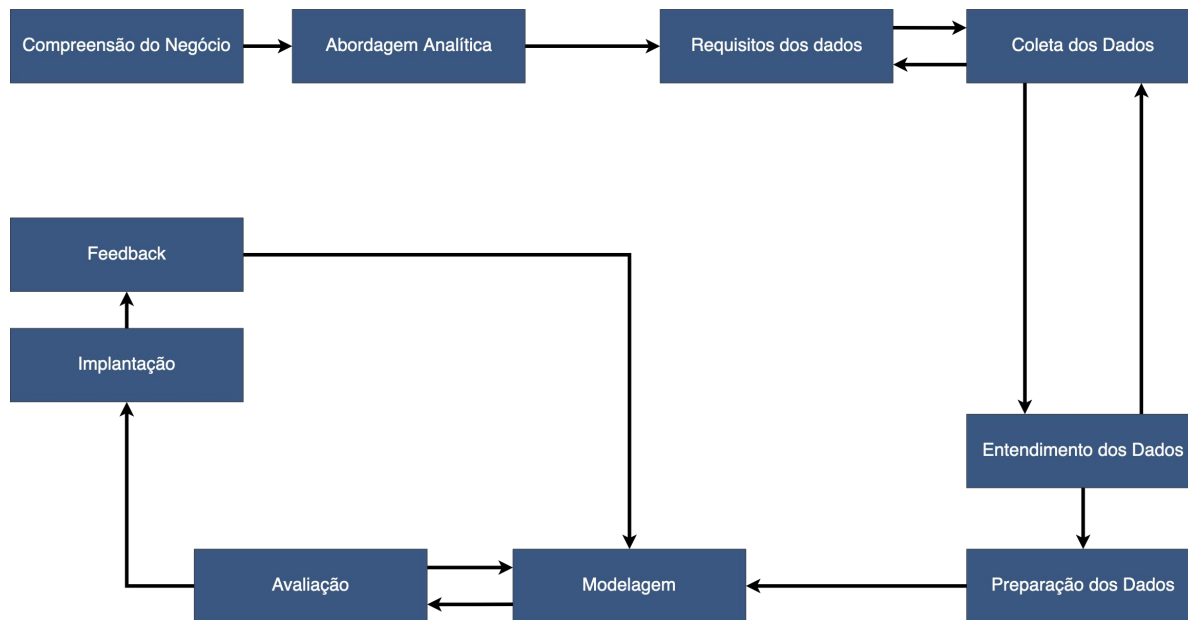
Pirâmide DIKW



Pirâmide DIKW



Metodologia da Ciência dos Dados



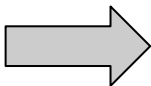
Metodologia proposta por John Rollins (Data Scientist, IBM Analytics, IBM)

<https://www.ibmbigdatahub.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?

Compreensão do Negócio

Abordagem Analítica

> É preciso ter **clareza** sobre o **problema** desde do início;
> Pensar na melhor abordagem analítica: **descritiva, diagnóstica, preditiva, prescritiva.**

Feedback

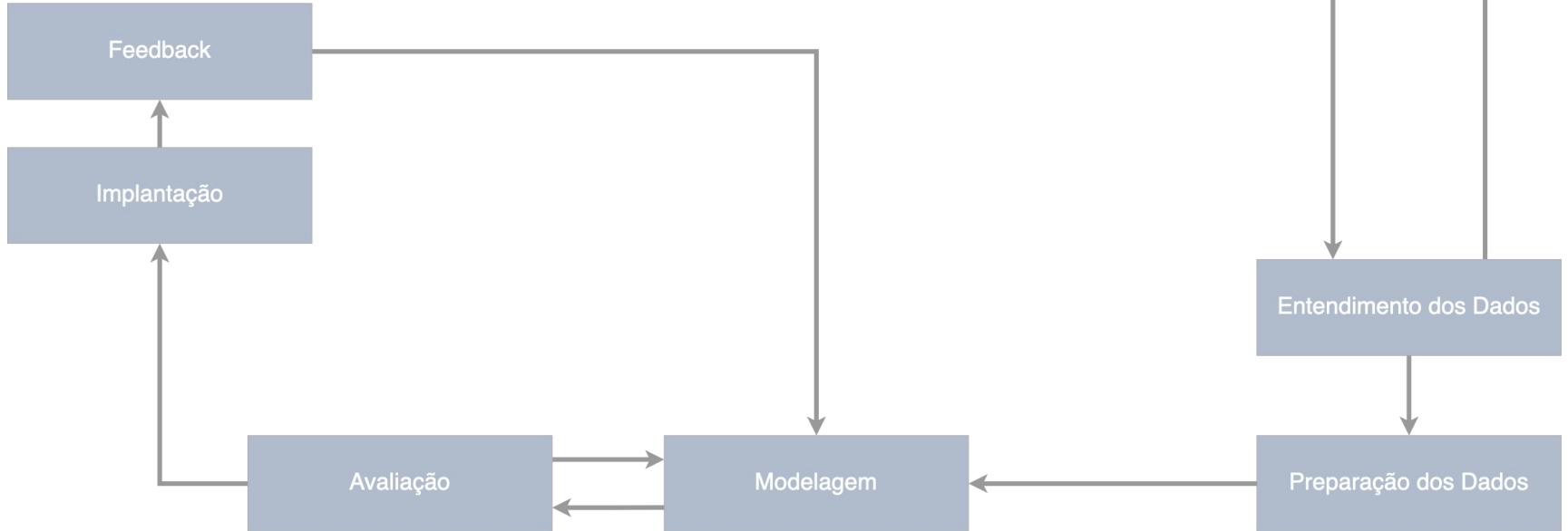
Implantação

Avaliação

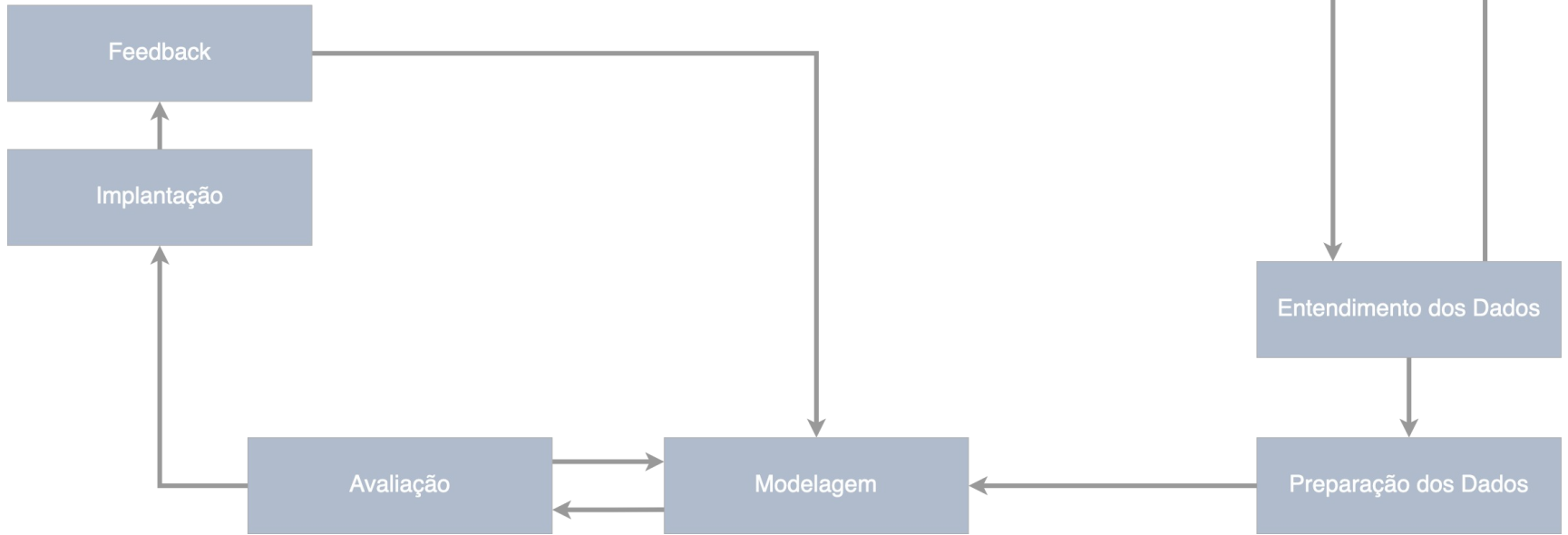
Modelagem

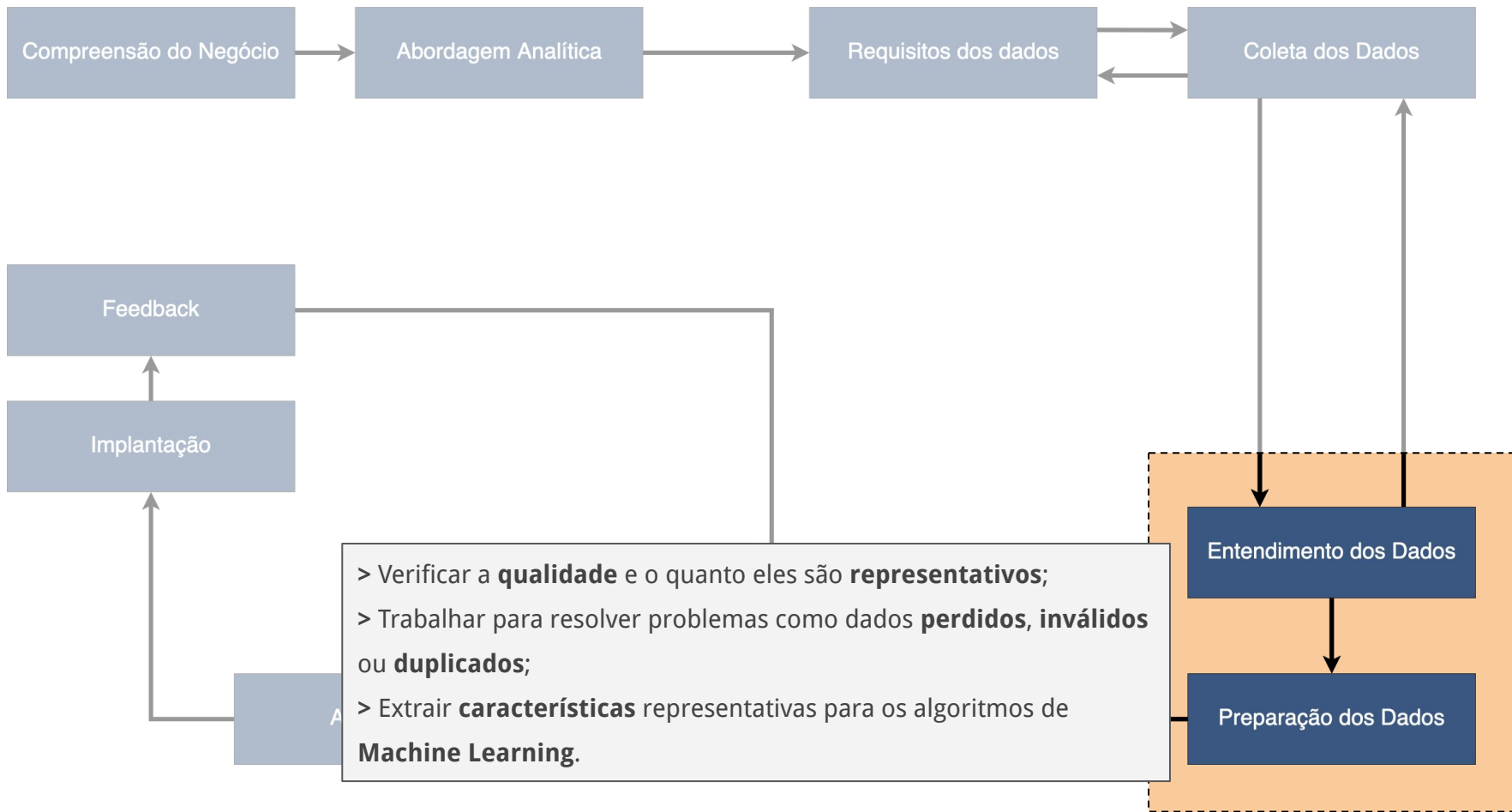
Entendimento dos Dados

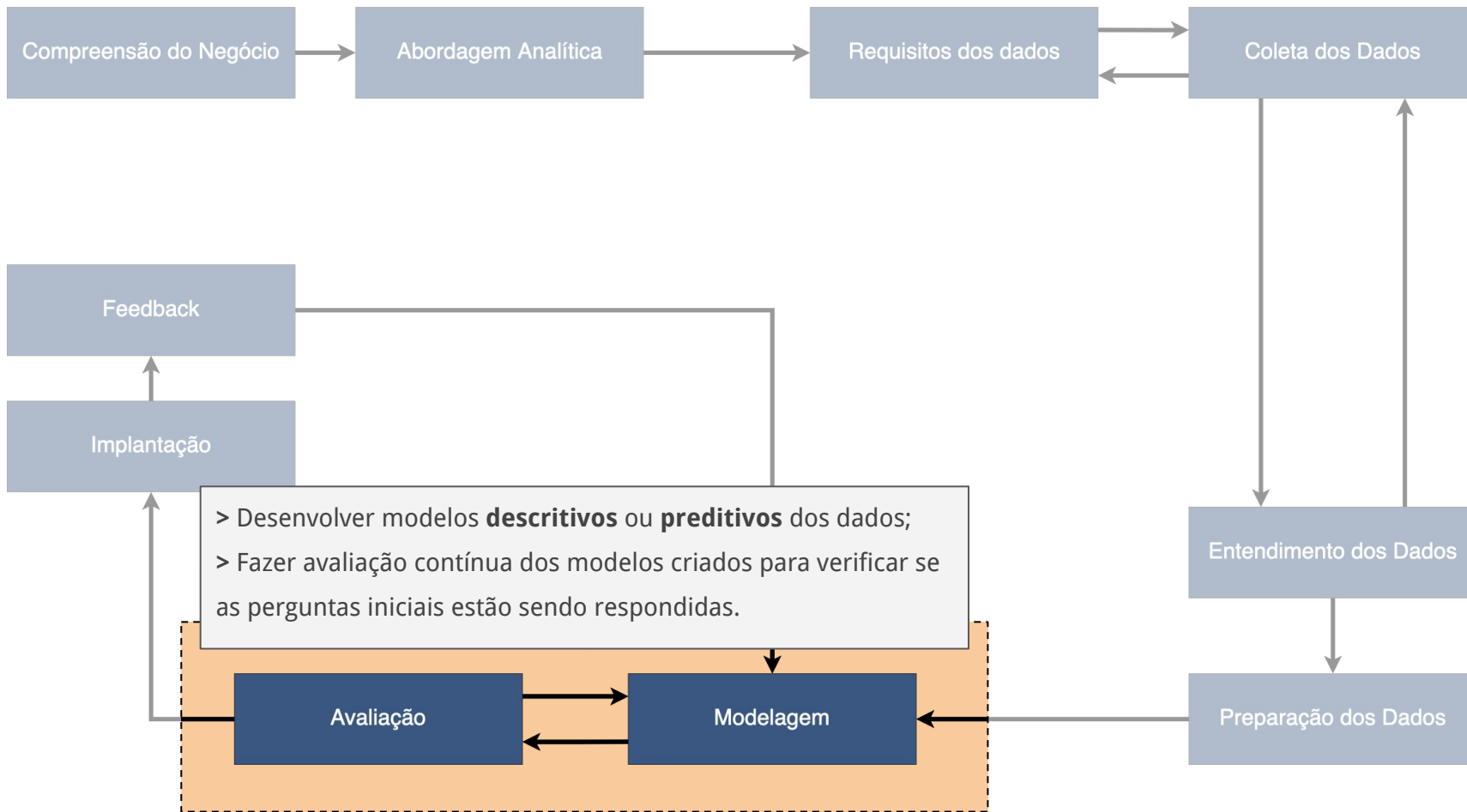
Preparação dos Dados

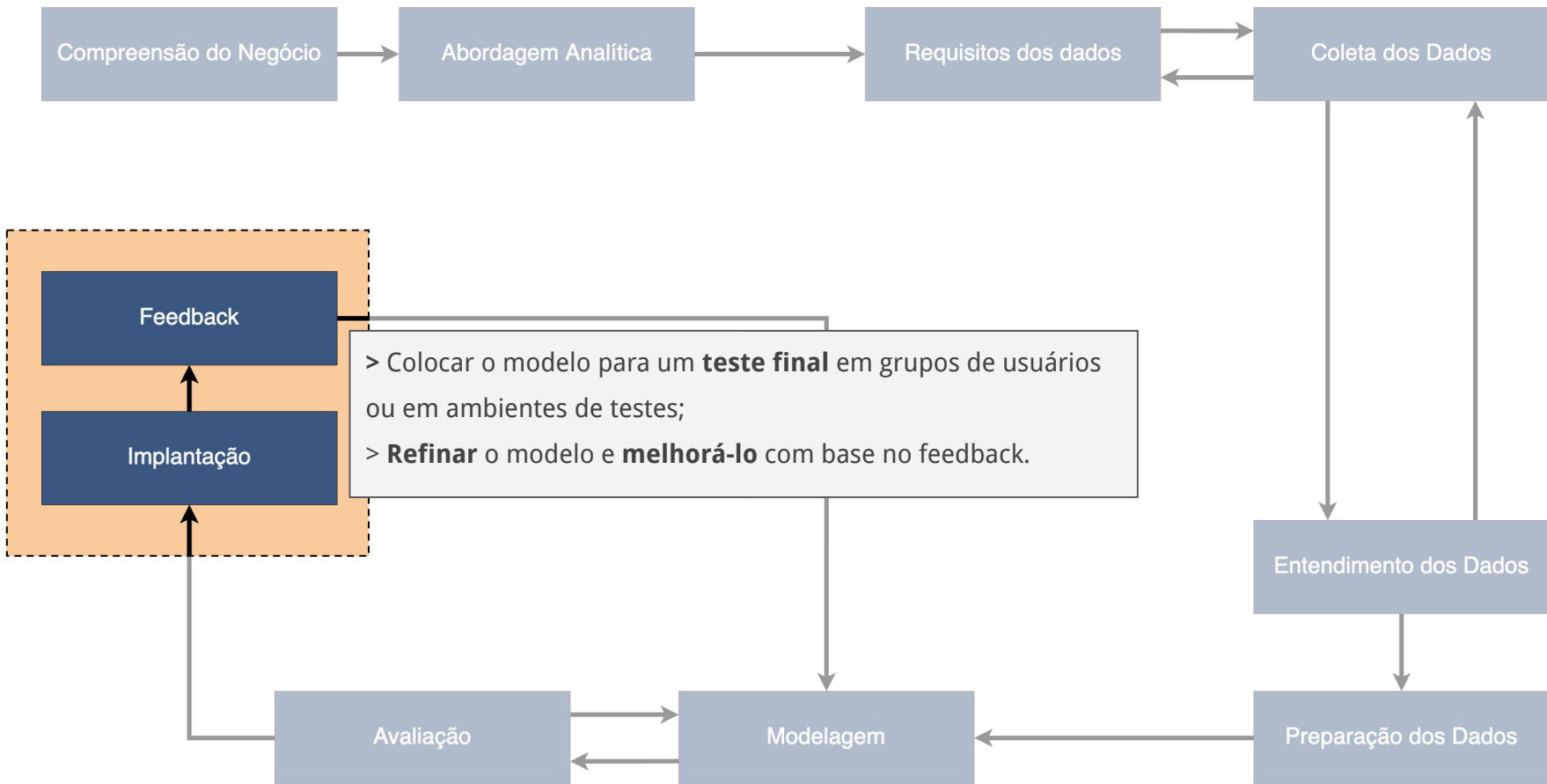


- > É preciso saber o que **coletar** e o que **não coletar**;
- > Avaliar a **coleta** e **redefinir requisitos**, se necessário, para uma nova coleta;









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?

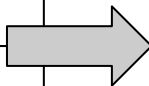
3. Encontrando respostas

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

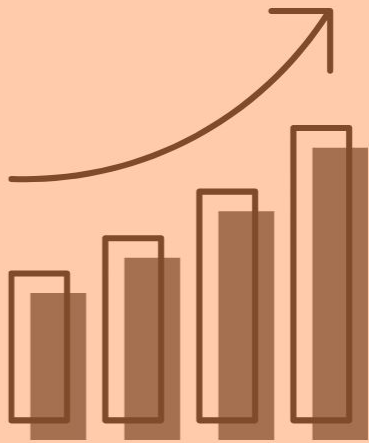
Entenda os dados

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?

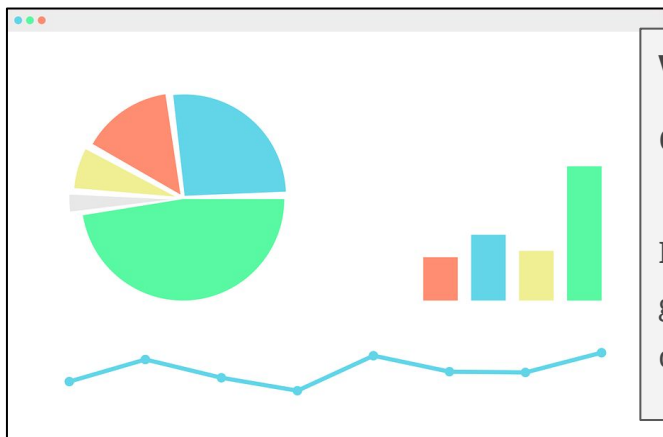


Dados coletados e analisados. Onde entra a visualização nisso tudo?

A visualização é inerente a diversas etapas do processo de análise de dados. Desde do 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 ...

Carte Figurative des pertes successives en hommes de l'Armée Française dans la campagne de Russie 1812-1813.

Dessiné par M. Minard, Inspecteur Général des Ponts et Chaussées en retraite Paris, le 20 Novembre 1869.

Les nombres d'hommes présents sont représentés par les hauteurs des zones colorées à raison d'un millimètre pour dix mille hommes; ils sont de plus écrits en travers des zones. Le rouge désigne les hommes qui entrent en Russie, le noir ceux qui en sortent. — Les renseignements qui ont servi à dresser la carte ont été puisés dans les ouvrages de M. M. Cbiers, de Légar, de Fesendac, de Chambray et le journal inédit de Jacob, pharmacien de l'Armée depuis le 28 Octobre. Pour mieux faire juger à l'œil la diminution de l'armée, j'ai supposé que les corps du Prince Ségne et du Maréchal Davoust qui avoient été détachés sur Minsk et Mohilew et ont rejoint vers Orscha et Witebsk, avoient toujours marché avec l'armée.

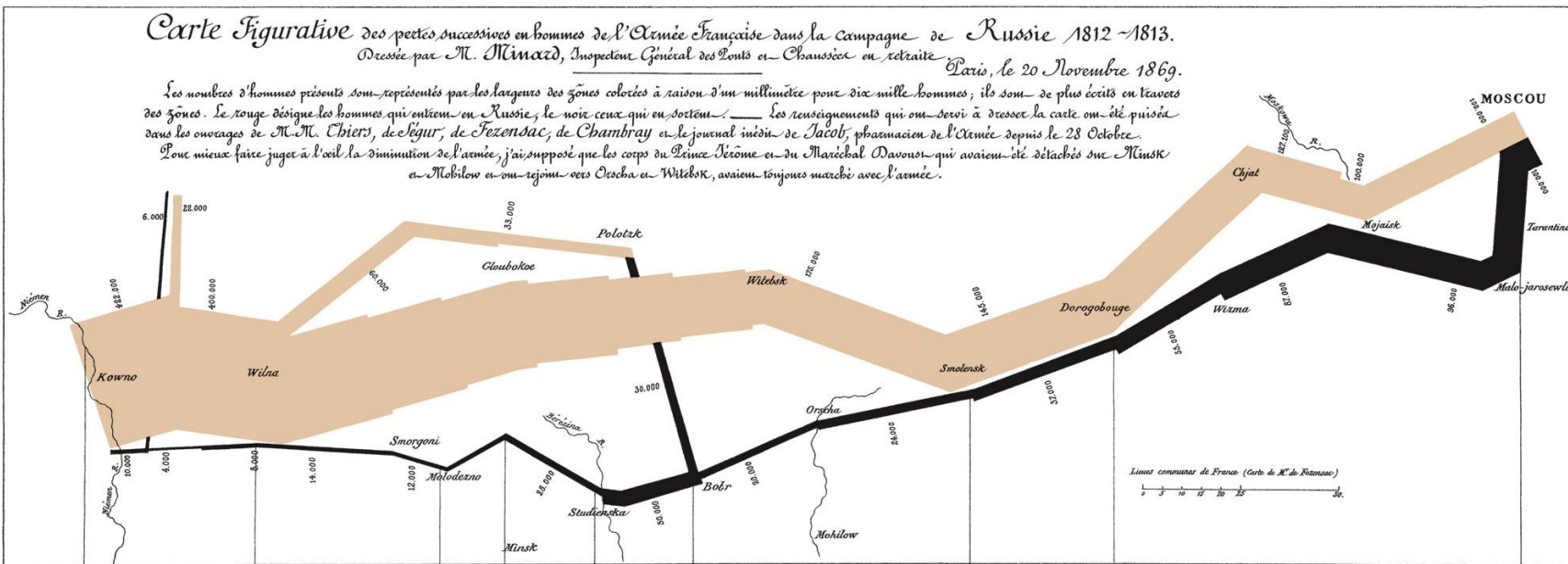
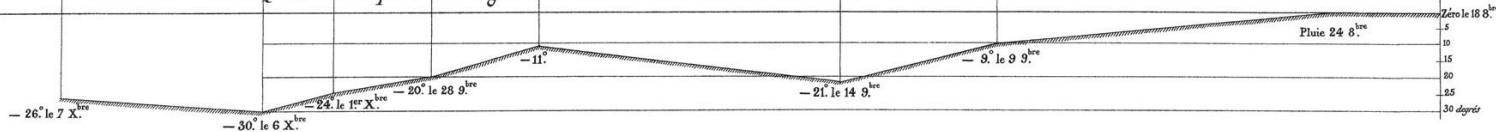


TABLEAU GRAPHIQUE de la température en degrés du thermomètre de Réaumur au dessous de zéro.

Les Cosaques passent au galop le Niémen gelé.

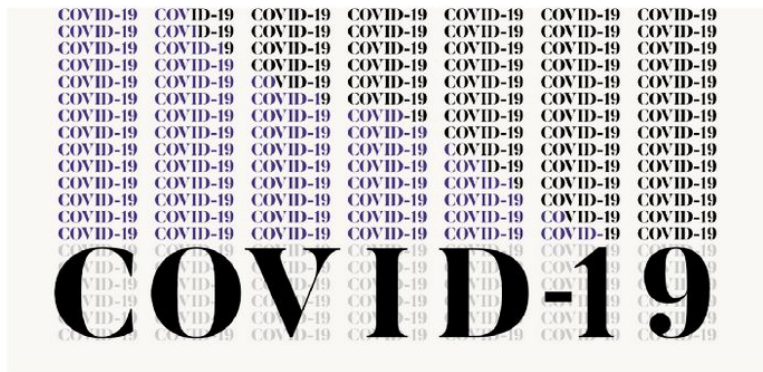


Mas chamou muita atenção pelo atual momento ...



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>

Desafios iniciais para construção do painel



🔍 Onde encontrar a base de dados da covid? ✕ 🗂

Pesquisa Google

Estou com sorte

Disponibilizado pelo Google em: [English](#)

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

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

Fonte dos dados de vacinação: <https://coronavirusbra1.github.io>

Criado por @adolfoguimaraes

10.331.767

casos confirmados

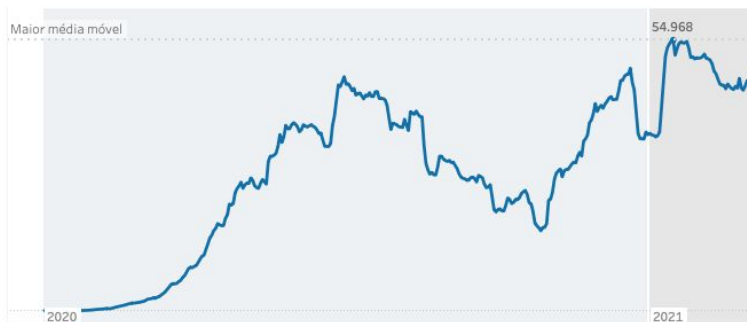
250.190

óbitos confirmados

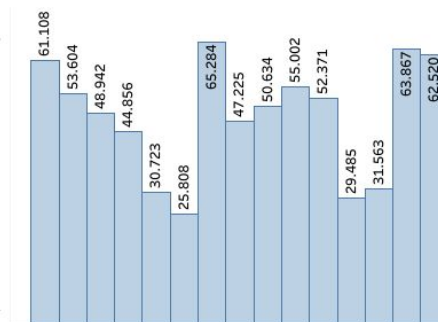
6.202.055

vacinados (1ª dose)

Casos



Casos por dia (últimos 15 dias)



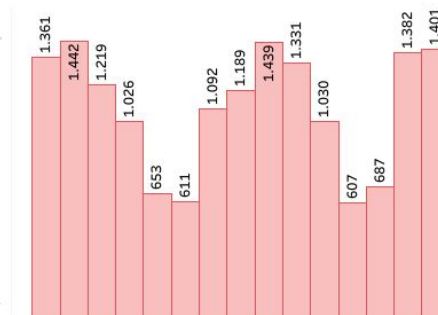
Selecione um estado para filtrar

- 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)

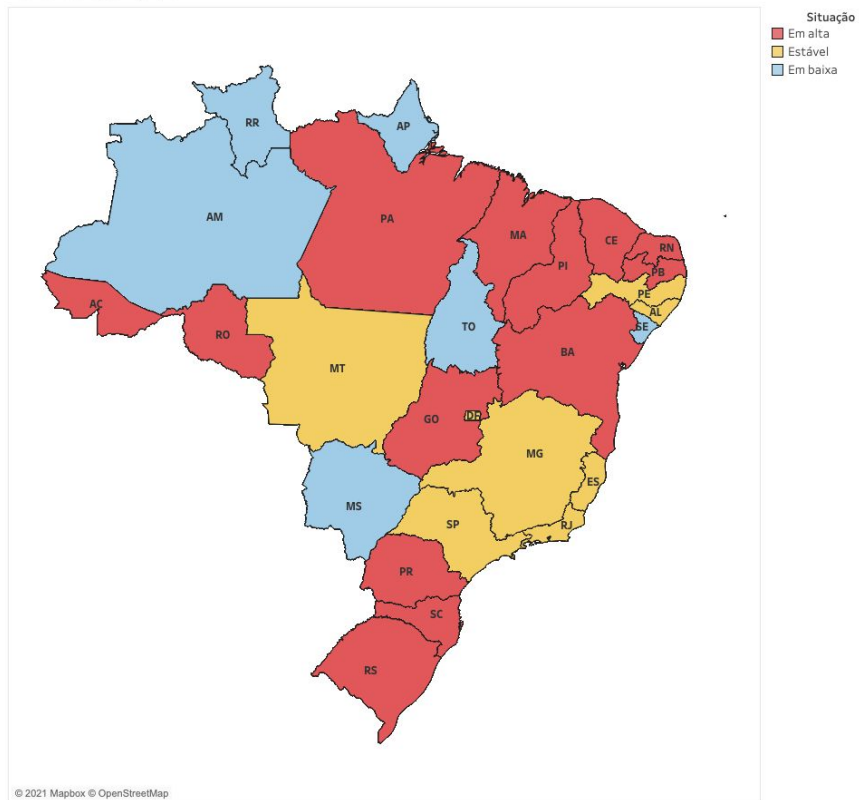


Mapa da situação de cada estado

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

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

Data da Atualização: 24/02/2021

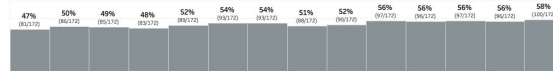


Ocupação das UTIs em Sergipe

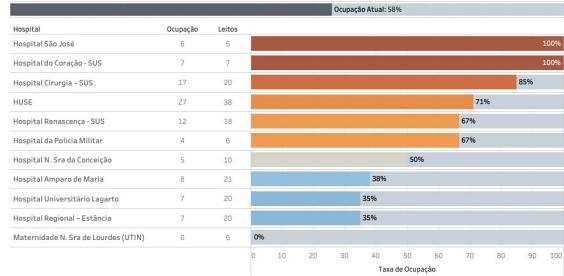
Fonte dos dados: Secretaria de Estado da Saúde
Data da Atualização: 23/02/2021

UTI Pública

Histórico dos últimos 14 dias:

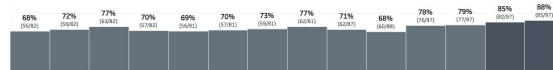


Ocupação Atual:

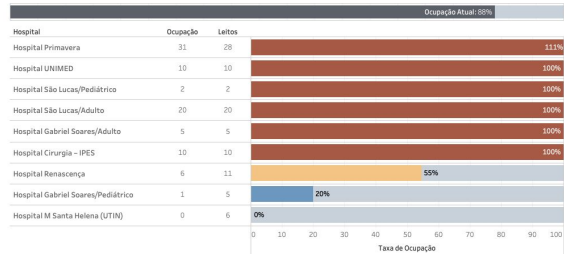


UTI Privada

Histórico dos últimos 14 dias:



Ocupação Atual:



Criado por @adoifegumarães

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

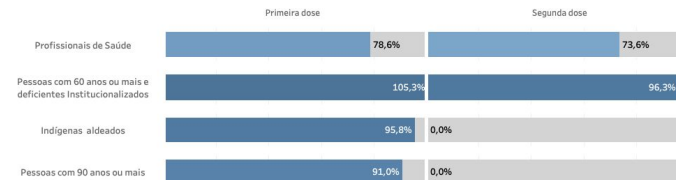
1,91%

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.



Vacinação por grupos

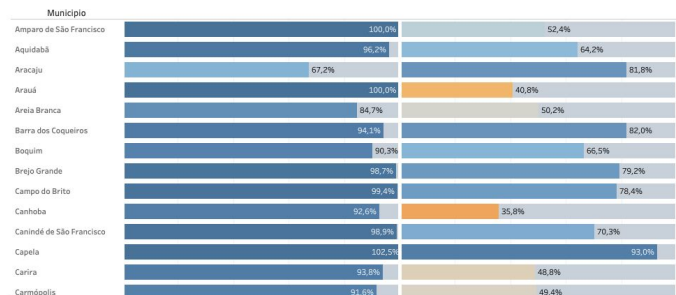
% de aplicação das doses enviadas para cada grupo. Passe o mouse nas barras para ver mais informações.



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 @adolfofoguimaraes

Painel dos Microdados da Vacinação COVID-19

Fonte de dados: arquivo disponibilizado pelo <https://brasil.io> (Atualização: 23/02/2021)

Painel criado por @adolfoguimaraes (<https://www.twitter.com/adolfoguimaraes>)

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

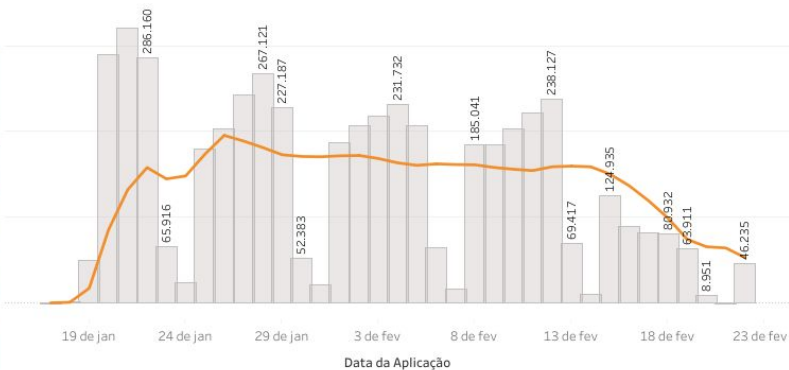
(% da população vacinada)

clique no estado para filtrar os dados

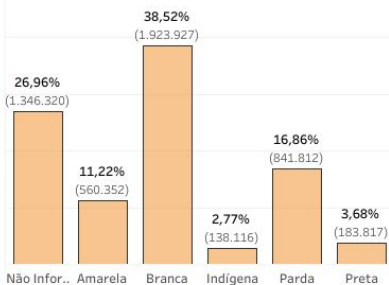
Acre	2,01%
Alagoas	1,90%
Amapá	1,94%
Amazonas	4,66%
Bahia	1,72%
Ceará	1,28%
Distrito Federal	2,88%
Espírito Santo	2,54%
Goias	2,22%
Maranhão	1,83%
Mato Grosso	2,36%
Mato Grosso do Sul	2,38%
Minas Gerais	1,88%
Pará	1,45%
Paraíba	2,59%
Paraná	1,44%
Pernambuco	2,08%
Piauí	1,79%
Rio de Janeiro	1,79%
Rio Grande do Norte	2,39%
Rio Grande do Sul	2,87%
Rondônia	2,12%
Roraima	2,57%
Santa Catarina	1,56%
São Paulo	3,61%
Sergipe	1,58%
Tocantins	2,56%

Aplicação da primeira dose

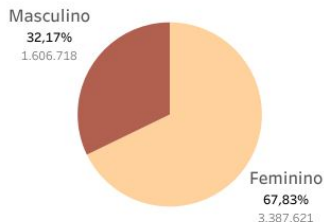
(por dia)



Etnia



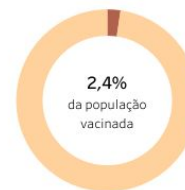
Sexo Biológico



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

COVID-19

Boletins informativos e casos do coronavírus por município por dia

Fonte original: Secretarias de Saúde estaduais

Liberado por: Álvaro Justin e decimas dos colaboradores

Código-fonte: <https://github.com/brasilio/covid19-br>

Licença: Creative Commons Attribution-ShareAlike 4.0 International (CC BY-SA 4.0)

Links relacionados: Boletins PB, Boletins SP, Boletins RJ, Boletins MG, Boletins RS, Boletins MT, Boletins MS, Boletins BA, Boletins MA, Boletins PA, Boletins PE, Informações sobre a coleta de dados (manual), 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

Tabuleiros: boletim, caso, caso_fut, obito, autorizo

Informações úteis

Esta 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 o repêto o dado do dia anterior). Caso queira ver a tabela original (sem repêto 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.

Acesse também o **Portal COVID-19**

ATENÇÃO antes de nos enviar dúvidas sobre os dados:

Saiba mais sobre esse dataset e o projeto Brasil.io COVID-19.

Leia a documentação dessa tabela.

Leia nossa FAQ sobre esse dataset.

Caso tenha verificado dados incorretos, 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/>

COVID-19 no Brasil

Twitter: [@coronavirusbr](https://twitter.com/coronavirusbr)
[@covid19BrBot](https://twitter.com/covid19BrBot)

Favor citar a nossa link como fonte ao usar os números!

Ocupação de Leitos **Informações importantes** **Gráfico por habitantes** **Gráfico Linha do Tempo**

<https://coronavirusbr1.github.io/> **Quem está usando** **Nos apoie!**

Acompanhe [@covid19BrBot](https://twitter.com/covid19BrBot) para números do CONASS, Ministério da Saúde e Secretarias Estaduais de Saúde

Dados de casos, óbitos e recuperados:
<https://covid19-br.netlify.com/>
 (via [Ministerio da Saúde, Brasil.io](https://twitter.com/MinisterioDaSaude) e [Giscard](https://twitter.com/Giscard))

Dados de testes, suspeitos e vacinados:
[@coronavirusbr](https://twitter.com/coronavirusbr)

Doses Aplicadas - 1ª Dose
6.437.836
% da população: 3,04%

Doses Aplicadas - 2ª Dose
1.884.206
0,893%

Dados de testes, suspeitos e recuperados:
 Casos: 10.455.630 | Óbitos: 253.100 | Testes: 34.088.607
 Suspensos: 4.755.382 | Recuperados: 9.309.568

Registros do dia
 Novos Casos: 61.093 | Novos Óbitos: 1.298

Clique no título das colunas para reordenar os dados

Última atualização de Testes e Suspeitos: 22/03/21

Vacinas: **Oxford / AstraZeneca, Sinovac** Total de doses aplicadas: 8.322.042

UF	Casos	Óbitos	Novos Casos	Novos Óbitos	Testes	Suspeitos	Recuperados	Doses Aplicadas - 1ª Dose	Doses Aplicadas - 2ª Dose	Total de doses aplicadas	% pop. vacinada	% pop. 2ª dose
SP	2.026.125	59.129	11.596	256	9.544.530	1.007.232	1.792.647	1.768.523	506.660	2.275.183	3,82%	1,09%
MG	869.230	18.276	6.728	141	3.248.039	2.406.072	791.506	550.602	238.155	788.757	2,59%	1,12%
BA	674.384	11.625	4.563	137	1.675.268	150.027	642.921	430.630	96.081	526.711	2,88%	0,64%
SC	663.699	7.242	6.050	77	1.575.179	112.419	622.993	162.829	55.768	218.597	2,25%	0,77%
PR	638.096	11.539	4.543	72	1.606.150	15.331	463.796	297.611	98.438	396.049	2,58%	0,85%
RS	632.019	12.226	7.188	77	2.495.593	11.146	591.323	448.428	98.208	546.636	3,93%	0,86%
RJ	579.542	32.909	424	138	2.313.502	52.017	542.754	457.108	112.886	569.994	2,63%	0,65%
			94	43	1.388.660	42.822	311.209	281.041	85.855	366.896	3,06%	0,93%



<https://coronavirusbr1.github.io/>

PORTARIA/PROTOCOLO
SANITÁRIO

TRANSPARÊNCIA

NOTÍCIAS

DECRETOS

BOLETINS

ÁREA TÉCNICA

REDE DE ATENÇÃO À SAÚDE

VACINA SERGIPE

BOLETIM VACINA COVID-19

Clique aqui

BOLETIM COVID-19

<p>CONFIRMADOS</p> <p>150.042</p>	<p>NEGATIVADOS</p> <p>164.760</p>
<p>CURADOS</p> <p>141.391</p>	<p>ÓBITOS</p> <p>2.947</p>

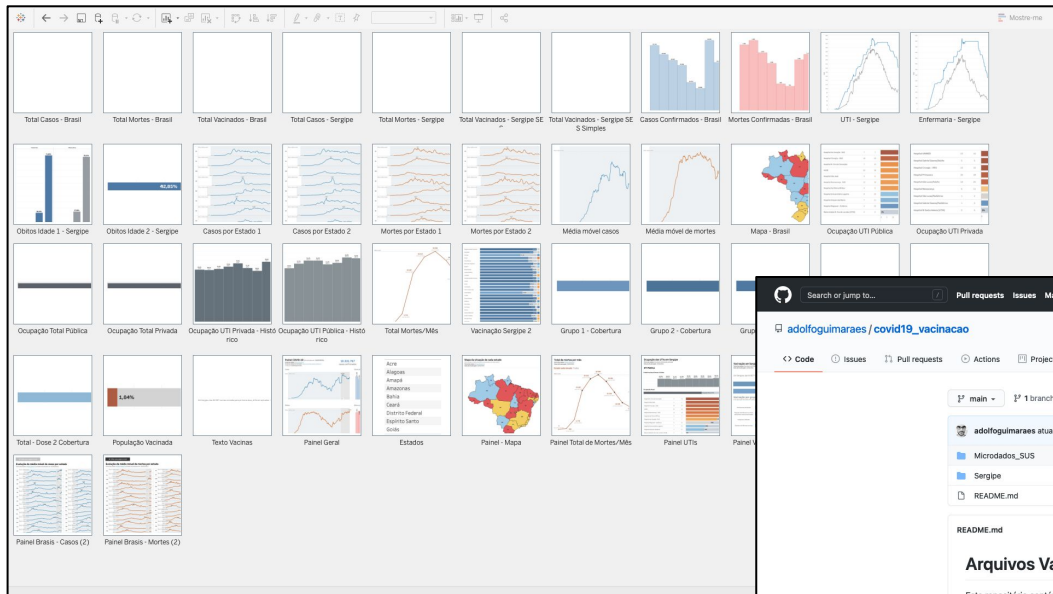
MESMO DEPOIS DE VACINADO MANTENHA OS

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

Ferramentas Utilizadas



Próximos passos



Search or jump to... Pull requests Issues Marketplace Explore

adolfofoguimaraes / covid19_vacinacao

Unwatch 1 Star 0 Fork 0

Code Issues Pull requests Actions Projects Wiki Security Insights Settings

main 1 branch 0 tags Go to file Add file Code

adolfofoguimaraes atualização sergipe 23/02 a592c1c 3 days ago 17 commits

- 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 @adolfofoguimaraes (Twitter | Instagram)

About

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

Releases

No releases published
Create a new release

Packages

No packages published
Publish your first package

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

@profadolfoaguimaraes



Obrigado :)