

# Onde o governo e as empresas conseguem dados?

**Dados Abertos** são informações públicas que qualquer pessoa pode acessar e usar gratuitamente para análise, pesquisa ou desenvolvimento de aplicativos. Eles são a base para muitas decisões importantes tomadas por governos e empresas.



## Transparência

Permite que cidadãos acompanhem e fiscalizem as ações do governo, promovendo a prestação de contas.



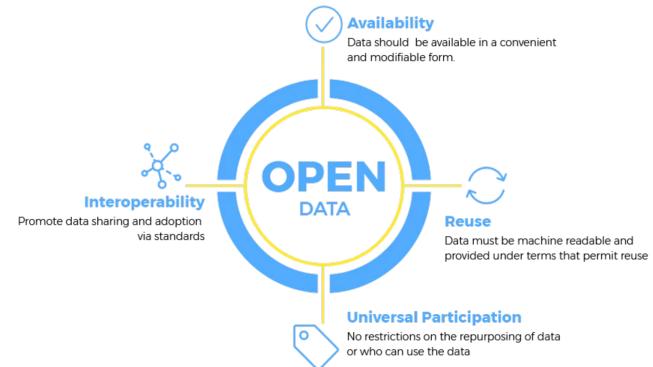
## Inovação

Possibilita que empresas e desenvolvedores criem novos produtos e serviços baseados em informações públicas.



## Cidadania

Incentiva a participação social e o engajamento cívico através do acesso à informação.



# Tipos de Dados: Estruturados vs Não Estruturados



## Dados Estruturados

São dados que seguem um **formato fixo** como tabelas e planilhas. São organizados em linhas e colunas, facilitando a análise e o processamento.

### Exemplos:

- Planilhas Excel com informações de vendas
- Bancos de dados SQL com registros de clientes
- Arquivos CSV com dados demográficos
- Tabelas com resultados de pesquisas

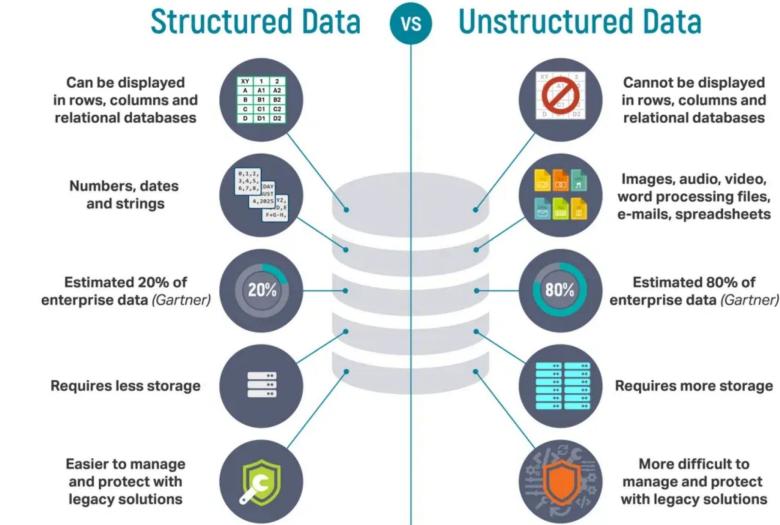


## Dados Não Estruturados

São dados **sem formato predefinido**, mais difíceis de processar e analisar automaticamente. Representam a maior parte dos dados disponíveis atualmente.

### Exemplos:

- Textos de redes sociais e e-mails
- Arquivos de áudio e vídeo
- Imagens e fotografias
- Documentos PDF sem tabelas



# Fontes de Dados Públicos: Nacionais



## Portal Brasileiro de Dados Abertos [dados.gov.br](http://dados.gov.br)

Portal oficial do governo federal que centraliza o acesso a [dados públicos governamentais](#).

### Exemplos de dados disponíveis:

- Orçamento federal e gastos públicos
- Indicadores de saúde e educação
- Dados ambientais e climáticos



## IBGE

[ibge.gov.br](http://ibge.gov.br)

O [Instituto Brasileiro de Geografia e Estatística](#) é a principal fonte de dados demográficos e econômicos do Brasil.

### Exemplos de dados disponíveis:

- Censo Demográfico (população, educação)
- Pesquisa Nacional por Amostra de Domicílios
- Índices de inflação (IPCA, INPC)



# Fontes de Dados Públicos: Internacionais

**ONU**



**Data**

[data.un.org](http://data.un.org)

Portal de dados das **Nações Unidas** que reúne estatísticas globais sobre desenvolvimento.

## Exemplos de dados disponíveis:

- Indicadores de desenvolvimento sustentável
- Estatísticas demográficas mundiais
- Dados sobre comércio internacional



**Kaggle**  
**Datasets**

[kaggle.com/datasets](http://kaggle.com/datasets)

Plataforma que hospeda **milhares de conjuntos de dados** compartilhados pela comunidade.

## Exemplos de dados disponíveis:

- Dados de competições de ciência de dados
- Conjuntos de dados para aprendizado de máquina
- Datasets para visualização e análise



# Atividade A: Navegação Guiada nos Portais

Nesta atividade, vamos explorar os portais de dados abertos **individualmente ou em duplas**. O objetivo é navegar pelos portais, identificar conjuntos de dados interessantes e anotar informações relevantes.

## 1 Acessar o Portal

Abra o navegador e accese um dos portais: [dados.gov.br](https://dados.gov.br) ou [ibge.gov.br](https://ibge.gov.br). Observe a página inicial e os menus disponíveis.

## 2 Navegar pelas Categorias

Explore as diferentes seções e categorias de dados disponíveis (ex: Saúde, Educação, Economia). Clique em algumas categorias para ver os conjuntos de dados relacionados.

## 3 Identificar Conjuntos de Dados

Encontre pelo menos três conjuntos de dados que você considere interessantes. Não é necessário baixá-los ainda, apenas explore as descrições e metadados.

## 4 Anotar Informações

Para cada conjunto de dados identificado, anote o nome e o tema principal. Observe também o formato do arquivo (CSV, XLS, etc.) e a data da última atualização.

A screenshot of the gov.br open data portal homepage. The URL in the address bar is 'dados.gov.br/dados/conjuntos-dados'. The page features the 'gov.br' logo and navigation links for 'Órgãos do Governo', 'Acesso à Informação', 'Legislação', and 'Acessibilidade'. A blue header bar contains the text 'Dados Abertos' and 'gov.br - Governo Federal'. On the left, there's a sidebar with categories like 'Comunicações', 'Comércio e Serviços', 'Cultura', 'Defesa Nacional', 'Economia e Finanças', 'Educação' (which is highlighted in blue), 'Energia', and 'Selecionar um tema'. The main content area shows a large yellow banner with the word 'abertos'. Below the banner are several search and filter fields: 'Selecionar um formato', 'Selecionar a periodicidade de atualização', 'Selecionar uma palavra-chave', 'Selecionar se tem reuso', 'Selecionar uma licença', 'Objetivo de Desenvolvimento Sustentável', 'Possui dados de raça/etnia?', and 'Possui dados de gênero?'. At the bottom, there's a section titled 'Filtros de organização'.

15151 Conjuntos de dados abertos

# Atividade B: Exploração dos Dados

Nesta atividade, vamos explorar um conjunto de dados de forma sistemática para entender suas características principais. Esta etapa é crucial para a alfabetização em dados .

## 1 Escolha do Dataset

Cada grupo deve escolher um dataset para trabalhar. Pode ser um dos conjuntos de dados que você encontrou na atividade anterior.

## 2 Download e Abertura

Baixe o arquivo (geralmente em formato .csv) e abra-o no Google Sheets ou Excel. Lembre-se que CSV é um formato de arquivo que separa os dados por vírgulas.

## 3 Análise Inicial (em grupo)

- Identificar Variáveis:** Liste as colunas do dataset e explique o que cada uma representa.
- Tipos de Dados:** Identifique se os dados são numéricos, textuais ou de data.
- Valores Faltantes:** Localize dados ausentes ou vazios e estime a quantidade.

## 4 Criação do Quadro

O grupo deve organizar essas observações em um quadro simples, para apresentar as características principais do dataset e facilitar a análise posterior.



A	B	C	D	E	F	G	H	I	J	K
1	Invoice ID	Branch	City	Customer	Gender	Product	Unit price	Quantity	Tax %	Total
2	750-67-84:A	Yangon	Member	Female	Health	ani	74.69	7	26.1415	548.9715
3	226-31-30:C	Naypyitaw	Normal	Female	Electronic		15.28	5	3.82	80.22
4	631-41-31:A	Yangon	Normal	Male	Home	and	46.33	7	16.2155	340.5255
5	123-19-11:A	Yangon	Member	Male	Health	ani	58.22	8	23.288	489.048
6	373-73-79:A	Yangon	Normal	Male	Sports	anc	86.31	7	30.2085	634.3785
7	699-14-30:C	Naypyitaw	Normal	Male	Electronic		85.39	7	29.8865	627.6165
8	355-53-59:A	Yangon	Member	Female	Electronic		68.84	6	20.652	433.692
9	315-22-56:C	Naypyitaw	Normal	Female	Home	and	73.56	10	36.78	772.38
10	665-32-91:A	Yangon	Member	Female	Health	ani	36.26	2	3.626	76.146
11	692-92-55:B	Mandalay	Member	Female	Food	and	54.84	3	8.226	172.746
12	351-62-08:B	Mandalay	Member	Female	Fashion	anc	14.48	4	2.896	60.816
13	529-56-39:B	Mandalay	Member	Male	Electronic		25.51	4	5.102	107.142
14	365-64-05:A	Yangon	Normal	Female	Electronic		46.95	5	11.7375	246.4875
15	252-56-26:A	Yangon	Normal	Male	Food	and	43.19	10	21.595	453.495
16	829-34-39:A	Yangon	Normal	Female	Health	ani	71.38	10	35.69	749.49
17	299-46-181:B	Mandalay	Member	Female	Sports	anc	93.72	6	28.116	590.436
18	656-95-93:A	Yangon	Member	Female	Health	ani	68.93	7	24.1255	506.6355
19	765-26-69:A	Yangon	Normal	Male	Sports	anc	72.61	6	21.783	457.443
20	329-62-151:A	Yangon	Normal	Male	Food	and	54.67	3	8.2005	172.2105
21	319-50-33:B	Mandalay	Normal	Female	Home	and	40.3	2	4.03	84.63
22	300-71-46:C	Naypyitaw	Member	Male	Electronic		86.04	5	21.51	451.71
23	371-85-57:B	Mandalay	Normal	Male	Health	ani	87.98	3	13.197	277.137

# Atividade C: Elaboração de Perguntas Investigáveis

Com base na análise inicial do dataset, cada grupo deve propor **perguntas investigáveis** que possam ser respondidas usando os dados disponíveis.



## O que é uma Pergunta Investigável?

É uma pergunta específica que pode ser respondida através da análise dos dados disponíveis, estabelecendo relações entre variáveis ou identificando padrões.

### Exemplo:

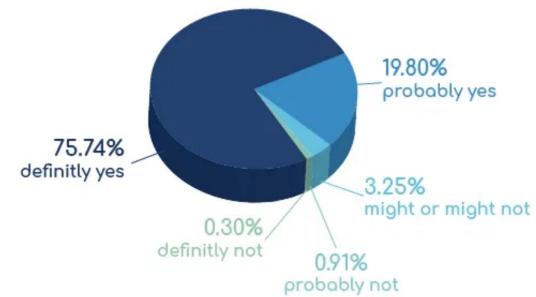
"Se o dataset for sobre saúde: Existe uma relação entre o número de hospitais em uma cidade e a taxa de mortalidade infantil?"



## Dicas para Boas Perguntas

- ✓ Seja específico e claro na formulação
- ✓ Verifique se os dados disponíveis permitem responder à pergunta
- ✓ Busque relações entre diferentes variáveis do dataset
- ✓ Considere perguntas que possam gerar insights relevantes

Do you believe that critical thinking skills are necessary in today's world?



# Atividade D: Análise Crítica e Debate em Grupo

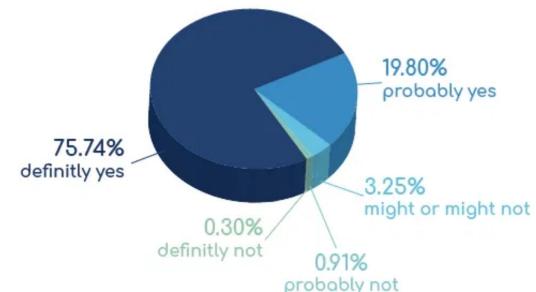
Nesta atividade, vamos estimular o **pensamento crítico** com um debate em grupo sobre as possíveis limitações dos dados encontrados.



## Perguntas para Debate

- "Os dados são recentes? Qual a data da última atualização?"
- "Os dados cobrem todas as regiões do Brasil ou apenas algumas?"
- "A fonte é confiável? O que a fonte diz sobre a metodologia de coleta?"
- "Os dados podem ter algum tipo de viés? Quais grupos podem estar sub-representados?"

Do you believe that critical thinking skills are necessary in today's world?



## Objetivo da Atividade

Estimular o pensamento crítico sobre a qualidade e representatividade dos dados, desenvolvendo a capacidade de avaliar fontes de informação e reconhecer limitações antes de tirar conclusões.

# Atividade E: Estudo de Caso Guiado - Análise do IBGE (Parte 1)

Nesta atividade prática, vamos analisar dados reais do IBGE sobre a população brasileira. Siga o passo a passo enquanto demonstramos na tela.

## 1 Encontrar o Dataset Estruturado

Acesse o SIDRA - Sistema IBGE de Recuperação Automática, uma ferramenta poderosa para consulta de dados estatísticos:

<https://sidra.ibge.gov.br/Tabela/6579>

## 2 Entender o Conteúdo da Tabela

Esta tabela contém dados sobre a **População Residente**, um indicador fundamental para entender a demografia do país. Observe que os dados estão organizados por:

- **Brasil:** Os dados são para o país inteiro
- **Ano:** O ano em que a população foi medida
- **População Residente:** O valor da população em milhões

## 3 Explorar a Interface do SIDRA

Familiarize-se com a interface do SIDRA. Observe as opções de filtro, visualização e download disponíveis no topo da página. Esta plataforma permite personalizar a consulta de dados de diversas formas.

AlanTaranti/**sidrapy**

A library that provides a python interface for the  
IBGE SIDRA API.



3  
Contributors

71  
Used by

75  
Stars

5  
Forks



# Atividade E: Estudo de Caso Guiado - Análise do IBGE (Parte 2)

Continuando nosso estudo de caso, vamos agora [carregar e analisar](#)

[os dados](#) do IBGE sobre População Residente no Google Sheets ou

Excel.

## 2 Entender e Baixar os Dados

Na página da tabela do SIDRA, observe as colunas e o que representam:

Brasil: Os dados são para o país inteiro

Ano: O ano em que a população foi medida

População Residente: O valor da população em milhões

Para baixar, clique no ícone de "download" (seta para baixo) e selecione "CSV".



Invoice ID	Branch	City	Customer Type	Gender	Product Line	Unit Price	Quantity	Tax %	Total	Date
750-67-84	A	Yangon	Member	Female	Health care	74.69	7	26.1415	548.9715	1/05/201
226-31-30	C	Naypyitaw	Normal	Female	Electronic	15.28	5	3.82	80.22	3/08/201
631-41-31	A	Yangon	Normal	Male	Home and Garden	46.33	7	16.2155	340.5255	3/03/201
123-19-11	A	Yangon	Member	Male	Health care	58.22	8	23.288	489.048	1/27/2019
373-73-79	A	Yangon	Normal	Male	Sports and Outdoors	86.31	7	30.2085	634.3785	2/08/201
699-14-30	C	Naypyitaw	Normal	Male	Electronic	85.39	7	29.8865	627.6165	3/25/2019
355-53-59	A	Yangon	Member	Female	Electronic	68.84	6	20.652	433.692	2/25/2019
315-22-56	C	Naypyitaw	Normal	Female	Home and Garden	73.56	10	36.78	772.38	2/24/2019
665-32-91	A	Yangon	Member	Female	Health care	36.26	2	3.626	76.146	1/10/201
692-92-55	B	Mandalay	Member	Female	Food and Beverage	54.84	3	8.226	172.746	2/20/2019
351-62-08	B	Mandalay	Member	Female	Fashion accessories	14.48	4	2.896	60.816	2/06/201
529-56-39	B	Mandalay	Member	Male	Electronic	25.51	4	5.102	107.142	3/09/201
365-64-05	A	Yangon	Normal	Female	Electronic	46.95	5	11.7375	246.4875	2/12/201
252-56-26	A	Yangon	Normal	Male	Food and Beverage	43.19	10	21.595	453.495	2/07/201
829-34-39	A	Yangon	Normal	Female	Health care	71.38	10	35.69	749.49	3/29/2019
299-46-18	B	Mandalay	Member	Female	Sports and Outdoors	93.72	6	28.116	590.436	1/15/2019
656-95-93	A	Yangon	Member	Female	Health care	68.93	7	24.1255	506.6355	3/11/201
765-26-69	A	Yangon	Normal	Male	Sports and Outdoors	72.61	6	21.783	457.443	1/01/201
329-62-15	A	Yangon	Normal	Male	Food and Beverage	54.67	3	8.2005	172.2105	1/21/2019
319-50-33	B	Mandalay	Normal	Female	Home and Garden	40.3	2	4.03	84.63	3/11/201
300-71-46	C	Naypyitaw	Member	Male	Electronic	86.04	5	21.51	451.71	2/25/2019
371-85-57	B	Mandalay	Normal	Male	Health care	87.98	3	13.197	277.137	3/05/201

## 3 Carregar e Analisar no Google Sheets/Excel

Abra um novo arquivo no Google Sheets ou Excel e importe o arquivo CSV baixado.

**Pergunta:** "Qual foi a evolução da população brasileira entre 2010 e 2020?"

**Cálculo de Crescimento:** Use a fórmula exemplo:

$$= (U6-A6)/A6$$

## 4 Aplicar a Lógica ao Seu Próprio Dataset

Volte ao dataset que seu grupo escolheu no Bloco 2 e tente responder a uma das perguntas que vocês criaram, usando a mesma lógica do exemplo do IBGE.

Identifique as variáveis relevantes, aplique fórmulas simples e interprete os resultados.

# Atividade F: Apresentação e Troca de Experiências

Cada grupo apresentará suas descobertas para a turma. Esta é uma oportunidade para compartilhar insights e aprender com as experiências dos colegas.

## 1 Resumo do Dataset

Apresente qual foi o dataset escolhido e o que ele contém. Explique brevemente a origem dos dados e sua relevância.

## 2 Perguntas de Pesquisa

Compartilhe as perguntas investigáveis que o grupo criou e explique por que elas são importantes ou interessantes.

## 3 Limitações e Desafios

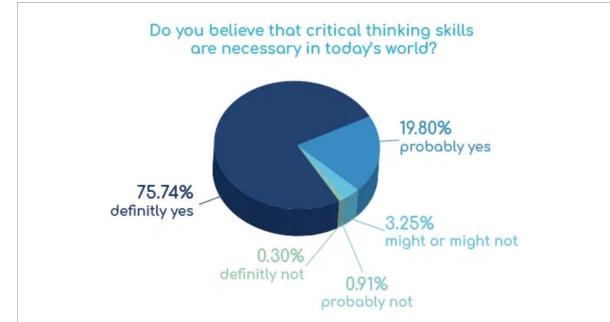
Discuta as limitações do dado (confiabilidade, cobertura geográfica, atualidade) e os desafios encontrados durante a análise.

## 4 Descoberta Principal

Se conseguiram, compartilhem uma descoberta inicial sobre o dataset que consideram relevante ou surpreendente.

## 5 Reflexão

Explique como esse dado pode ser usado para gerar impacto social ou entender um problema real da sociedade.



## Tempo de Apresentação

Cada grupo terá de 7 a 8 minutos para apresentar. Organize bem o tempo para cobrir todos os pontos importantes.

# Reflexão Coletiva e Encerramento



## Reflexão Coletiva

Vamos discutir a importância da **cidadania baseada em dados** e como o acesso à informação pode transformar nossa participação na sociedade.

### Perguntas para Discussão:

- ① Como o acesso a esses dados pode nos ajudar a fiscalizar o governo?
- ② Onde mais na nossa comunidade podemos aplicar essa lógica de análise para resolver problemas?



## Principais Aprendizados

- ✓ A importância dos dados abertos para transparência, inovação e cidadania
- ✓ Fontes confiáveis para encontrar dados públicos nacionais e internacionais
- ✓ Métodos para explorar, analisar e questionar criticamente os dados
- ✓ O poder de usar dados para tomar decisões informadas e exercer a cidadania

