



# Grupo de Pesquisa em Educação, Trabalho e Formação Humana (GEFOR)

## OFICINA DE DADOS EDUCACIONAIS

Módulo 1 - Aula 1

# Fundamentos de dados

Ministrante: Leidiane Santos

Setembro/2025

# Conteúdo

---

- O que são dados
- Tipos de dados (estruturados e não estruturados)
- Tipos de variáveis e tipos de dados
- Como organizar uma análise de dados
- Perguntar aos dados

Atividade prática

# O que são dados?

---

- Dados são fatos ou estatísticas brutos que ainda não foram organizados, processados ou analisados de forma significativa. Eles podem ser números, palavras, imagens, sons ou qualquer outra coisa que possa ser capturada e representada de alguma forma.
- Por exemplo, o número "20", a palavra "aluno" ou uma imagem de um pôr do sol são todos dados.

# Como os dados se tornam informações?

---

- A transformação dos dados em informações ocorre quando eles são organizados, estruturados e interpretados de maneira significativa. Isso geralmente envolve algum tipo de processamento ou análise.
- Por exemplo, se eu tiver uma lista de números, posso calcular a média desses números. Esse resultado é uma informação derivada dos dados originais.

# Por que os dados são importantes na análise?

---

- Os dados são essenciais para a análise porque fornecem a matéria-prima sobre a qual podemos basear nossas descobertas e tomadas de decisão. Sem dados, estamos apenas especulando. Com dados, podemos identificar padrões, tendências e correlações que nos ajudam a entender melhor o mundo ao nosso redor e a fazer escolhas informadas.

# Dados estruturados

---

- Os dados estruturados são organizados em um formato predefinido e uniforme, geralmente em tabelas com linhas e colunas. Cada elemento de dados é categorizado e possui um tipo específico, facilitando a consulta e análise.
- Um exemplo comum de dados estruturados são os bancos de dados relacionais, onde as informações são armazenadas em tabelas com campos definidos. Linguagens de consulta como SQL são frequentemente usadas para acessar e manipular dados estruturados.

# Dados estruturados

A	B	C	D	
nome_candidato	total_valor_rece	total_doadores	total_votos	
CAMILA CRISTINA LANES DA SILVA	152,432	10	1881	
CAROLINE ARNS DE SANTA CRUZ ARRUDA	1,632,700	22	22360	
CHRISTIANE DE SOUZA YARED	3,041,200	2	32677	
DIOGO TADAO HARA FURTADO	1,000	1		
ELOY FASSI CASAGRANDE JUNIOR	60,969	9	1021	
FERNANDO DESTITO FRANCISCHINI	5,635,884	39	52340	
JOAO JOSE DE ARRUDA JUNIOR	1,558,750	3	21833	
JORGE GOMES DE OLIVEIRA BRAND	1,313,882	22	110977	
JOSÉ MARIA BONI	6,738	3	3509	
JOÃO GUILHERME OLIVEIRA DE MORAES	1,587,735	50	40505	
LETÍCIA LANZ DE SOUZA	64,652	10	3564	
MARISA LOBO FRANCO FERREIRA ALVES	348,350	23	18547	
PAULO RICARDO OPUSZKA	923,973	3	20537	
RAFAEL VALDOMIRO GRECA DE MACEDO	3,791,400	35	499821	
RENATO MOCELLIN	50,566	6	6305	
SAMARA GARRATINI	12,380	5	632	

# Dados não estruturados

---

- Os dados não estruturados não têm um formato pré-determinado e não se encaixam facilmente em tabelas ou esquemas tradicionais. Eles podem ser de natureza textual, como documentos de texto, e-mails, posts em mídias sociais, ou podem incluir dados multimídia, como imagens, vídeos e áudio.
- Devido à sua falta de estruturação, a análise de dados não estruturados geralmente requer técnicas mais avançadas, como processamento de linguagem natural (PLN) ou reconhecimento de padrões. Exemplos incluem posts em redes sociais, e-mails, transcrições de áudio ou vídeo.



# Dados semi-estruturados

---

- Os dados semi-estruturados estão em algum lugar entre dados estruturados e não estruturados. Eles contêm alguma estrutura, mas não são tão rígidos quanto os dados estruturados. Geralmente, os dados semi-estruturados são organizados em um formato que permite alguma flexibilidade na forma como são armazenados e consultados.
- Um exemplo comum de dados semi-estruturados são os documentos XML (Extensible Markup Language), que têm uma estrutura hierárquica, mas podem conter diferentes tipos de dados em diferentes partes do documento. Outro exemplo são os documentos JSON (JavaScript Object Notation), frequentemente utilizados em aplicações web para troca de dados entre sistemas.

## Estruturado

1001  
1010

1001  
0101

1100  
0110

0011  
1100

0110  
1001

0011  
1010

0011  
0011

0101  
1100

1001  
1001

## Não Estruturado



## Semi Estruturado



# Tipos de variáveis

---

- Uma variável é um atributo mensurável passível de mudança e...  
variação
- Podemos olhar para as variáveis por mais de um ponto de vista,  
como por exemplo:
  - Pela sua forma de operacionalização (variáveis qualitativas  
nominais e ordinais, quantitativa discreta e contínua)
  - De acordo com a escala de mensuração (nominal, ordinal,  
intervalar e de razão)
  - A partir da sua relação com outras variáveis (variável  
dependente e independente, explicativa e resposta etc.)

**Quadro 2.2 – Resumo dos tipos de variáveis**

	Tipo	Características	Exemplos	Funções Formais
<b>Qualitativa</b>	<b>Nominal ou de atributos</b>	Apenas para identificar pessoas, objetos ou categorias.	Cor de cabelo, estado civil, nome, marca de carro.	Equivalência, igualdade.  "="
	<b>Ordinal ou de ordem</b>	Respostas podem ser ordenadas em uma dimensão própria.	Escolaridade. Desempenho de governo. Ordem de preferência, de chegada, status social, escala de Likert. (tem moda e mediana, mas não a média)	Além de igualdade, identifica relação de superioridade ou inferioridade. Diferenças de grandezas.  ">" ; "<"
<b>Quantitativa</b> Discreta: números inteiros; Contínua: números quebrados	<b>Intervalar</b>	Além de poder ser ordenada em uma dimensão específica há intervalos com tamanho conhecido que podem ser comparados.	Escala de Likert, estimativa de distâncias. Escala de QI.	Além das anteriores permite operações aritméticas nas diferenças entre os números que representam os eventos. Determina igualdade de intervalos.
	<b>Proporcional ou de Razão</b>	Além das características da escala anterior, existe ainda um ponto zero absoluto.	Salário, altura, tempo gasto em uma tarefa.	Além das anteriores permite operações aritméticas nos próprios códigos/números. Determina igualdade de relações.

Tipo	Características	Escala	Operações
Qualitativa nominal	Identifica, nomeia e classifica pessoas, objetos ou categorias	Somente nominal	Classificação; operações de igualdade e diferença
Qualitativa ordinal	Além do anterior, estabelece alguma forma de ordenamento em uma dimensão própria	<i>Ordinal</i> : Existe um ordenamento entre as categorias, apesar do tamanho ou distância entre elas não ser conhecido	Classificação e ordenamento; operações de igualdade, diferença e diferença entre grandezas
Quantitativa discreta	Permite a contagem e contabilização de pessoas, objetos ou categorias; os valores são inteiros e não negativos	<i>Intervalar</i> : pode ser ordenada em uma dimensão específica, com intervalos de tamanho conhecido que podem ser comparados.	Permite operações aritméticas nas diferenças entre os números que representam os eventos investigados
Quantitativa contínua	Podem assumir um número infinito de valores possíveis que podem ser associados a pontos em uma escala contínua	<i>Intervalar</i> : pode ser ordenada em uma dimensão específica, com intervalos de tamanho conhecido que podem ser comparados. <i>De razão</i> : pode ou não ter zero não arbitrário	Permite operações aritméticas nas diferenças entre os números que representam os eventos investigados

# Tipos de variáveis e tipos de dados

---

- Temperatura
- Pressão arterial
- Pobreza
- Crescimento econômico

# Tipos de dados

---

- Mas ainda precisamos prestar atenção em um aspecto adicional: a **forma como o computador armazena as informações**.
  - Muitas vezes precisamos “traduzir” os tipos de variáveis. Assim temos certeza que o nosso software ou linguagem de programação de escolha vai fazer o que precisamos que ele faça. Variáveis qualitativas (nominais ou ordinais) podem ser representadas por números. Contudo, esses números são apenas contêineres ou representações do que a variável de fato significa; não são números de verdade, daqueles que podemos fazer operações aritméticas. Vamos ver mais sobre isso.

# Tipos de variáveis e tipos de dados

---

- Contexto:

- **Programação e matemática**

Uma variável é um espaço de armazenamento simbolicamente representado por um nome, que está associado a um valor ou informação. Em programação e matemática, variáveis são usadas para armazenar e manipular dados. Elas servem como contêineres para valores que podem ser modificados durante a execução de um programa.



# Tipos de variáveis e tipos de dados

---

- Em linguagens de programação como Python e outras, por exemplo, você pode criar uma variável chamada "idade" e atribuir a ela um valor específico, como 25:

`idade = 25`

- Agora, a variável "idade" contém o valor 25 e pode ser usada em operações ou cálculos dentro do programa.

Tipo	Características
Inteiro (Integer)	Representa números inteiros, como -1, 0, 1, 2, etc. Geralmente é declarado como 'int' em muitas linguagens.
Ponto Flutuante (Float)	Representa números com parte decimal, como 3,14 ou -0,5. Em muitas linguagens, é declarado como 'float' ou 'double'.
Caractere (Character ou Char)	Representa um único caractere, como 'a', '1' ou '@'. Geralmente é declarado como 'char'.
String	Uma sequência de caracteres. Pode representar texto, como "Olá, mundo!". Em muitas linguagens, é declarado como 'string' ou 'str'.
Booleano (Boolean)	Pode ter apenas dois valores: 'true' ou 'false'. É comumente usado em expressões lógicas.
Date	Um tipo de dado especializado para representar datas. Em muitos bancos de dados e linguagens de programação, o tipo date facilita operações como a comparação de datas, cálculos de diferença entre datas, formatação e validação de entradas de data.
Outros	Lista (list ou array); dicionário (dictionary ou map); tupla (tuple); conjunto (set); objeto (object)

# O que é GitHub?

---

- É uma plataforma baseada na web que utiliza o Git, um sistema de controle de versão. É como um serviço de hospedagem para projetos de código.
- **Controle de Versão:** Guarda um histórico completo de todas as alterações feitas no código, permitindo que você volte para versões anteriores a qualquer momento.
- **Colaboração:** Facilita o trabalho em equipe, permitindo que várias pessoas trabalhem no mesmo projeto de forma organizada e sem conflitos.
- **Portfólio:** Funciona como um portfólio online onde desenvolvedores podem mostrar seus projetos.

# Conta e repositório do curso

---

<https://github.com/anevsantos/dados-educ-ufpa> – Material do curso e Scripts de consulta