

Pensamento Computacional e Lógica de Programação

Prof. Dr. Álvaro Campos
Ferreira

Monitoria e teste semana que vem

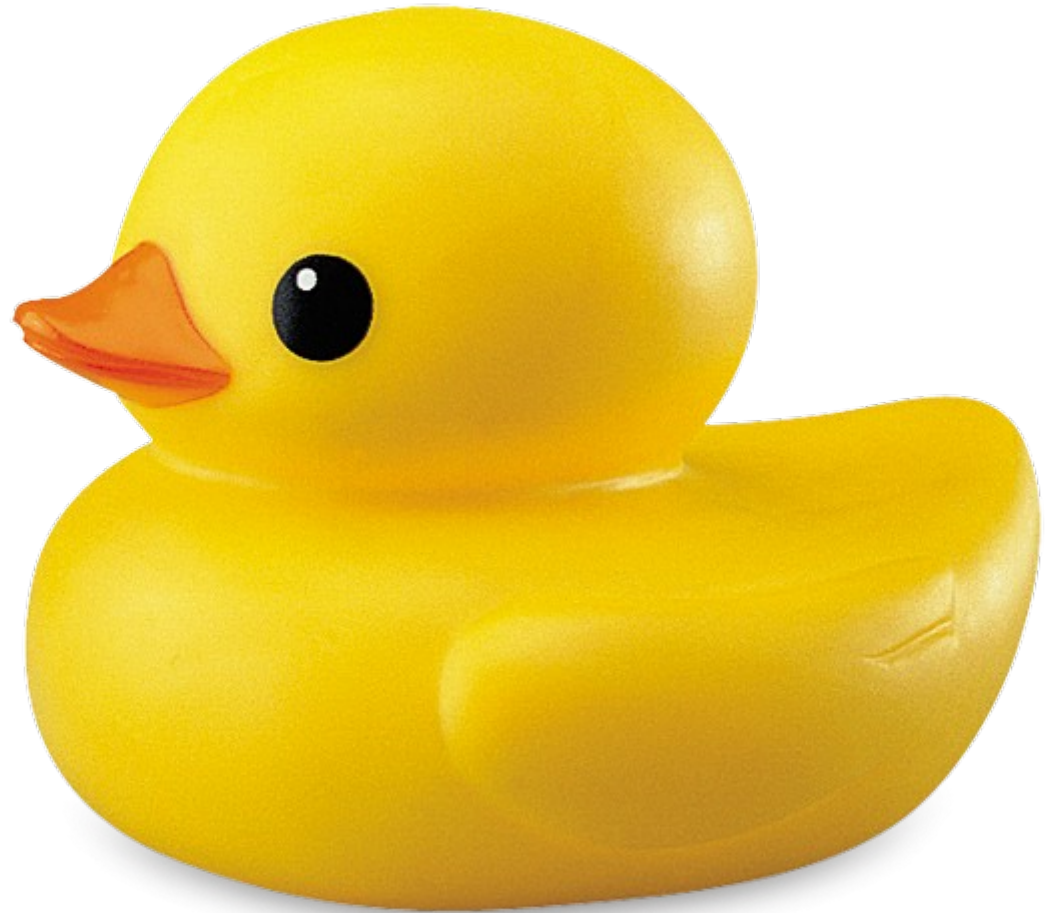
Dados

Estruturas de dados

Patinho de borracha

Patinho de borracha

Sempre que possível,
vamos explicar para
o patinho o que o
código faz ou
pretende fazer.



Pandas e DataFrames

Pandas e Dataframes

Vamos usar o Pandas para acessar os nossos dados sobre Pokemons.

```
import pandas as pd  
file = "pokemon_data.csv"  
df = pd.read_csv(file)  
print(df)
```


Análise exploratória de dados

Análise exploratória

Etapas de transformação do dado.



Análise exploratória

Etapas da estatística descritiva.



Indexação de DataFrames

O DataFrames é organizado em colunas que podem ser acessadas a partir de seu nome.

Para determinar os nomes das colunas, pode-se escrever na tela os primeiros valores com a função `columns`

```
df.columns
```

Indexação de DataFrames

Para selecionar uma coluna de um DataFrame, a sintaxe é a mesma de um dicionário, com o nome da coluna no lugar da chave:

```
df["Name"]
```

Para selecionar elementos, use a indexação de listas:

```
df["Name"][0]
```

Indexação de DataFrames

Slices ou fatias são seleções de mais de um elemento de um objeto ao mesmo tempo.

A faixa de índices é indicada entre o sinal “:”.

Ou seja, para selecionar os dez primeiros elementos, usa-se:

```
df["Name"][0:10]
```

Funções de DataFrames

DataFrames possuem diversas funções para facilitar a análise exploratória e de estatística descritiva. Algumas das funções que vamos utilizar são:

- `describe()`
- `plot()`

Funções de DataFrames

Funções de estatística descritiva:

- describe()
- count()
- sum()
- mean()
- median()
- mode()
- std()
- min()
- max()
- abs()

Funções de DataFrames

Algumas funções importantes para DataFrames são:

- groupby()
- sort_values()
- filter()
- value_counts()
- columns
- head()
- tail()

**Quais perguntas
responder?**

Quais perguntas responder?

A partir de agora, a análise vai acontecer através de uma série de perguntas cujas respostas podem ser obtidas através da análise e interpretação dos dados.

Definir bem as perguntas é útil para guiar a análise exploratória.

Quais perguntas responder?

Algumas perguntas que podemos responder de nosso conjunto de dados sobre Pokemons:

- Quantos pokemons estão no conjunto?
- Quantos de cada tipo?
- Quantos por geração?

Quais perguntas responder?

Outras perguntas:

- Qual é o Pokemon com o maior ataque?
- Qual é o com a maior defesa?
- Em média, qual é o tipo de Pokemon com maior ataque? E defesa?
- Quais são os Pokemons mais fortes de cada tipo e geração?

Onde encontrar dados?

- 1) Google Data Set Search
- 2) OpenDataSUS
- 3) Kaggle
- 4) Webscrapping



INSTITUTO BRASILEIRO DE ENSINO,
DESENVOLVIMENTO E PESQUISA