

Introdução ao Python

Pandas

Henrique Y. Shishido

Departamento de Computação
Universidade Tecnológica Federal do Paraná

Introdução

- O Pandas é uma biblioteca Python para trabalhar com conjunto de dados (datasets)
 - Análise
 - Limpeza
 - Exploração
 - Manipulação
- Por que usar Pandas?
 - Permite realizar análise de grande volume de dados e fazer conclusões baseadas em teorias estatísticas

O que o Pandas pode fazer?

- Obter respostas sobre um conjunto de dados como:
 - Há correlação entre os dados de duas ou mais colunas?
 - Qual é o valor médio de um conjunto de dados?
 - Maior valor?
 - Menor valor?

Instalação

- A instalação da biblioteca Pandas pode ser realizada por meio do comando:

```
$pip3 install pandas
```

Importação do pandas

- Uma vez instalado, a importação do Pandas é realizada por meio do comando `import`:

```
import pandas
```

Primeiro exemplo

Listing 1: Código 1

```
#Importação do pandas
import pandas

#Declaração de um dataset
dados = {
    'cliente': ["Jonas", "Kim", "Lucas"],
    'idade': [15, 39, 68]
}

#Conversão dos dados para Dataframe
df = pandas.DataFrame(dados)

print(df)
```

Saída do Código 1)

	cliente	idade
0	Jonas	15
1	Kim	39
2	Lucas	68