

Introdução ao Pandas

Prof. Matheus



Pandas

- Biblioteca Python para análise de dados
- Construída em cima da biblioteca numpy

Usabilidade

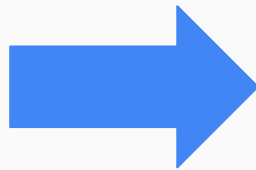
A biblioteca pandas é vastamente utilizada para:

- Leitura e escrita de dados tabulares em diversos formatos (csv, excel, etc.)
- Limpeza de dados (valores nulos, duplicados, etc.)
- Análise de dados (criação de gráficos, medidas estatísticas, etc.)
- Manipulação e modelagem de dados (agrupamentos, pivotagem, etc.)
- Integração de dados de diversas fontes (merge, join, concat, etc.)

Series - Conjunto de dados de 1 dimensão

Ingredientes:

- maisena
- farinha
- açúcar
- ovos
- manteiga
- castanhas



Ingredientes

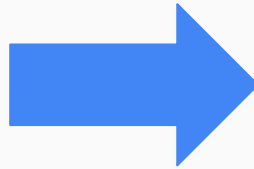
maisena
farinha
açúcar
ovos
manteiga
castanhas

Dados

Series - Conjunto de dados de 1 dimensão

Ingredientes:

- maisena
- farinha
- açúcar
- ovos
- manteiga
- castanhas



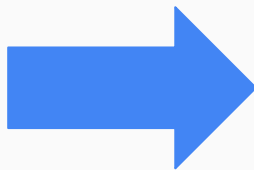
Índice	Ingredientes
0	maisena
1	farinha
2	açúcar
3	ovos
4	manteiga
5	castanhas

} Dados

Series - Conjunto de dados de 1 dimensão

Ingredientes:

- maisena
- farinha
- açúcar
- ovos
- manteiga
- castanhas



Índice Ingredientes

0	maisena
1	farinha
2	açúcar
3	ovos
4	manteiga
5	castanhas

Dados

Tipo: Texto

Series - Conjunto de dados de 1 dimensão

Quantidade:

- maisena: 200g
- farinha: 250g
- açúcar: 100g
- ovos: 2 unidades
- manteiga: 150g
- castanhas: 180g



Pos	Índice	qtd
0	maisena	200
1	farinha	250
2	açúcar	100
3	ovos	2
4	manteiga	150
5	castanhas	180

Dados

Tipo: número

DataFrame - Conjunto de dados de 2 dimensões

Índice	qtd
maisena	200
farinha	250
açúcar	100
ovos	2
manteiga	150
castanhas	180



Índice	Ingredientes
0	maisena
1	farinha
2	açúcar
3	ovos
4	manteiga
5	castanhas



Pos	Índice	Ingredientes	qtd
0	0	maisena	200
1	1	farinha	250
2	2	açúcar	100
3	3	ovos	2
4	4	manteiga	150
5	5	castanhas	180

Prática