

Kungfu Pandas

Lê Huỳnh Đức

2021-05-26

Contents

Lời nói đầu

Giới thiệu cuốn sách

Cài đặt Jupyter Lab

Cài đặt Pandas

Chapter 1

Cấu trúc và kiểu dữ liệu

1.1 Series

Trong Pandas, **Series** là mảng 1 chiều bao gồm một danh sách giá trị, và một mảng chứa index của các giá trị. Trong dữ liệu dạng bảng, mỗi Series được xem như là một cột của bảng đó. Cách đơn giản để tạo 1 series như sau

```
s = pd.Series(data, index=None, name=None)
```

Trong đó **data** có thể có dạng:

- `numpy.ndarray`, `List`
- Python `dict`
- `Scalar`

index có thể truyền hoặc không, tùy vào dạng của **data** mà **index** sẽ được định nghĩa khác nhau. **name** là tên của **Series**, giá trị này cũng không nhất thiết phải truyền vào.

Các cách khởi tạo

Khởi tạo Series bằng array Khi không truyền giá trị **index**, **Series** sẽ mặc định index của nó là 1 mảng số nguyên từ 0 đến `len(data) - 1`

```
In [1]: pd.Series(data=[0, 1, 2], index=["a", "b", "c"], name="meow")
Out[1]:
a      0
b      1
c      2
Name: meow, dtype: int64
```