

# Computer Science Advanced

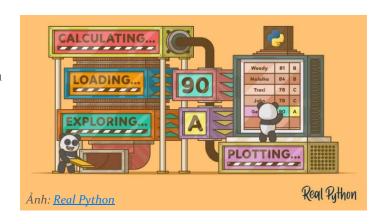
BÀI 11. XỬ LÝ DỮ LIỆU TRONG PANDAS

# 1. Thay Đổi Giá Trị trong DataFrame

Để thay đổi giá trị của một hoặc một số ô dữ liệu trong DataFrame, ta có thể sử dụng phép gán trên thuộc tính loc và iloc. Ví du:

#### <u>Dữ liệu ví dụ trong df</u>:

	Name	Score	Attempts	Qualify
0	Anastasia	12.5	1	yes
1	Dima	9.0	3	no
2	Katherine	16.5	2	yes
3	James	NaN	3	no
4	Emily	9.0	2	no



Ngoài ra, *DataFrame* còn có các thuộc tính **at** và **iat** có thời gian chạy nhanh hơn **loc** và **iloc** để *truy* xuất duy nhất một ô dữ liệu.

```
Cú pháp: df.at[<tên dòng>, <tên cột>]
    và df.iat[<thứ tự dòng>, <thứ tự cột>]

Ví dụ: df.at[1, 'Score']
    df.iat[1, 2] = 3
```

# THAY THẾ GIÁ TRỊ

DataFrame hỗ trợ thay thế một giá trị bất kỳ bằng một giá trị khác. **Cú pháp**:

```
df.replace(to replace, value, [inplace=...])
```

Tham số inplace là một giá trị boolean.

- False (mặc định): không thay đổi df mà trả về một DataFrame khác chứa thay đổi
- True: thay đổi df và trả về None

Name Score Attemps Qualify O Anastasia 12.5 1 True 1 Dima 9.0 3 False 2 Katherine 16.5 2 True 3 James NaN 3 False 4 Emily 9.0 2 False	>>	df :			
1 Dima 9.0 3 False 2 Katherine 16.5 2 True 3 James NaN 3 False		Name	Score	Attemps	Qualify
2 Katherine 16.5 2 True 3 James NaN 3 False	0	Anastasia	12.5	1	True
3 James NaN 3 False	1	Dima	9.0	3	False
	2	Katherine	16.5	2	True
4 Emily 9.0 2 False	3	James	NaN	3	False
	4	Emily	9.0	2	False

# 2. Thay Đổi Cột và Dòng



## Thêm cột

```
df[<tên cột>] = <Series hoặc list>
Cột được thêm phải số giá trị bằng số dòng trong df.
```

#### Ví du:

```
df['Rank'] = [2, 3, 1, 5, 3]
df['Rank'] = pd.Series([2, 3, 1, 5, 3])
```

# Xóa cột

```
df.drop(columns=..., [inplace=...])
columns: tên cột hoặc danh sách các cột.
```

#### Ví du:

### Thêm dòng

Series phải có index giống với df. Dict phải có các key giống tên các cột.

#### Ví du:

## Xóa dòng

```
df.drop(index=..., [inplace=...])
index: tên dòng hoặc danh sách các dòng.

Ví dụ:
df.drop(index=1, inplace=True)
df.drop(index=[1, 3], inplace=True)
```