

# Python cho Khoa học dữ liệu

## Bài 11: Kiểu dữ liệu tập hợp - Python Set

Hà Minh Tuấn  
hmtuan@hcmus.edu.vn

Khoa Toán - Tin học  
Trường Đại học Khoa học Tự nhiên, ĐHQG-HCM

Ngày 3 tháng 9 năm 2025

# Nội dung bài học

1. Giới thiệu về Sets

2. Các methods quan trọng

# Giới thiệu về Sets

- Set trong Python là tập hợp **không có thứ tự** và chỉ chứa các **phần tử duy nhất**.
- Có thể thay đổi (mutable), cho phép thêm hoặc xóa phần tử.
- Dùng để: kiểm tra thành viên, loại bỏ phần tử trùng lặp, và thực hiện các phép toán tập hợp.
- Được khai báo bằng dấu ngoặc nhọn {} hoặc hàm set().

# Tạo Set trong Python

```
# Tao set bang {}
my_set = {1, 2, 3, 4, 5}
print(my_set)

# Tao set bang ham set()
my_set2 = set([1, 2, 3, 4, 5])
print(my_set2)
```

```
{1, 2, 3, 4, 5}
{1, 2, 3, 4, 5}
```

## Set loại bỏ phần tử trùng lặp

```
my_set = {1, 2, 2, 3, 3, 4, 5, 5}  
print(my_set)
```

```
{1, 2, 3, 4, 5}
```

## Set chứa nhiều kiểu dữ liệu

```
mixed_set = {1, "hello", (1, 2, 3)}  
print(mixed_set)
```

```
{1, 'hello', (1, 2, 3)}
```

## Method add()

```
my_set = {1, 2, 3}  
my_set.add(4)  
print(my_set)
```

{1, 2, 3, 4}

## Method remove() và discard()

```
my_set = {1, 2, 3, 4}

# remove: lỗi nếu phần tử không tồn tại
my_set.remove(3)
print(my_set)

# discard: không lỗi nếu phần tử không tồn tại
my_set.discard(5)
print(my_set)
```

```
{1, 2, 4}
{1, 2, 4}
```

## Method pop() và clear()

```
my_set = {1, 2, 3}
removed = my_set.pop()
print("Phần tử bị lấy ra:", removed)
print("Còn lại:", my_set)

my_set.clear()
print(my_set)
```

Phần tử bị lấy ra: 1 (hoặc ngẫu nhiên)

Còn lại: {2, 3}

set()

## Method copy()

```
my_set = {1, 2, 3}  
new_set = my_set.copy()  
print(new_set)
```

{1, 2, 3}

## Kiểm tra phần tử trong Set

```
my_set = {1, 2, 3, 4}
if 2 in my_set:
    print("2 có trong set")
else:
    print("2 không có trong set")
```

2 có trong set

# Các phép toán tập hợp cơ bản

```
A = {1, 2, 3}  
B = {3, 4, 5}  
  
print(A | B)    # union  
print(A & B)    # intersection  
print(A - B)    # difference  
print(A ^ B)    # symmetric difference
```

```
{1, 2, 3, 4, 5}  
{3}  
{1, 2}  
{1, 2, 4, 5}
```

## Set Comprehension

```
squares = {x**2 for x in range(1, 6)}  
print(squares)  
  
even_set = {x for x in range(1, 11) if x % 2 == 0}  
print(even_set)
```

{1, 4, 9, 16, 25}  
{2, 4, 6, 8, 10}

# Frozen Set

```
my_frozen = frozenset([1, 2, 3])
print(my_frozen)
my_frozen.add(4)    # Lỗi!
```

```
frozenset({1, 2, 3})
```

```
AttributeError: 'frozenset' object has no attribute 'add'
```