

Homework 6: Pandas

A. Serie

1. Viết chương trình thực hiện các yêu cầu sau đây:

a) Tạo ra một Serie có dạng sau:

```
iPhone      10
Samsung Note 12
Samsung S    30
Nokia       100
dtype: int64
```

b) In ra giá trị của nhãn 'iPhone', 'Samsung S'

c) Tìm giá trị lớn nhất và nhỏ nhất trong Serie. In ra tên nhãn của dòng tương ứng.

d) In ra 2 dòng đầu tiên/2 dòng cuối trong Serie.

2. Viết chương trình thực hiện các yêu cầu sau đây:

a) Tạo ra một Serie có dạng sau:

```
Ke toan      200
Moi Truong   160
CNTT         800
Toan         100
dtype: int64
```

b) Sắp xếp Serie theo thứ tự tăng dần/ giảm dần.

c) Tạo ra một Serie mới bao gồm các thành phần có giá trị lớn hơn 150.

d) Tạo ra một Serie mới trong đó các giá trị của mỗi thành phần tăng thêm 200.

B. DataFrame

1. Cho một bảng dữ liệu có thông tin như sau:

	Animal	Weight	Average Age
Data 1	Tiger	258.5	16
Data 2	Lion	320.2	15
Data 3	Cat	3.3	12
Data 4	Whale	41000	33

Trong đó, các thông tin (**Animal** (tên động vật), **Weight** (trọng lượng trung bình tính theo đơn vị kg, **Average Age** (tuổi thọ trung bình) tính theo năm) là nhãn của mỗi cột. Cột **Animal** biểu diễn ở dạng chuỗi, cột **Weight** biểu diễn ở dạng số thực, cột **Average Age** biểu diễn dạng số nguyên.

Các thông tin (**Data 1, Data 2, Data 3, Data 4**) là nhãn của mỗi dòng.

- Viết chương trình tạo ra DataFrame lưu trữ nội dung bảng trên.
- In ra tên động vật có trọng lượng trung bình (weight) là lớn nhất/ nhỏ nhất.
- In ra tên động vật có tuổi thọ trung bình (Average Age) là lớn nhất/ nhỏ nhất.
- Từ DataFrame của câu (a) tạo ra DataFrame chỉ chứa 2 dòng (Data 1 và Data 4) và 2 cột là (Animal và Weight)
- In ra giá trị lớn nhất/ nhỏ nhất của cột dữ liệu **Weight**
- Tính giá trị trung bình (mean) của cột dữ liệu **Average Age**
- In 2 dòng đầu tiên, 2 dòng cuối ra màn hình.
- In ra màn hình các động vật có tuổi thọ trung bình (**Average Age**) lớn hơn 15.
- In ra màn hình các động vật có trọng lượng trung bình (**Weight**) nhỏ hơn 300.

2. Cho một bảng dữ liệu có thông tin như sau:

	Name	Origin	Price
Data 1	Grapefruit	Long An	5.5
Data 2	Coconut	Ben Tre	10.2
Data 3	Watermelon	Long An	6.7
Data 4	Banana	Ben Tre	2.5
Data 5	Dragon fruit	Long An	7.5

Trong đó, các thông tin (**Name** (tên loại trái cây), **Origin** (nơi xuất xứ), **Price** (giá tiền) tính theo đơn vị \$/kg. Cột **Name** biểu diễn ở dạng chuỗi, cột **Origin** biểu diễn ở dạng chuỗi, cột **Price** biểu diễn dạng số thực.

Các thông tin (**Data 1, Data 2, Data 3, Data 4, Data 5**) là nhãn của mỗi dòng.

- Viết chương trình tạo ra DataFrame lưu trữ nội dung bảng trên.
- In ra tên loại trái cây có giá tiền (Price) là lớn nhất/ nhỏ nhất.
- Tìm các loại trái cây có xuất xứ tại **Ben Tre**
- Từ DataFrame của câu (a) tạo ra DataFrame chỉ chứa 2 dòng (Data 2 và Data 5) và 2 cột là (Name và Price)
- In ra giá trị lớn nhất/ nhỏ nhất/trung bình của cột dữ liệu **Price**
- Tìm loại trái cây có xuất xứ tại **Ben Tre** và có giá tiền (Price) là cao nhất
- Tìm loại trái cây có xuất xứ tại **Long An** và có giá tiền (Price) là thấp nhất
- Tính tổng giá tiền (Price) của các loại trái cây có xuất xứ tại **Long An**
- Tính trung bình giá tiền (Price) của các loại trái cây có xuất xứ tại **Ben Tre**

