



Data Analytics Based Python

SECT5. 데이터구조 (Data Structure)

BigData Science Practical Training Course
Kim Jin Soo



- ◆ 리스트형, List Type
- ◆ 리스트형의 다양한 쓰임새
- ◆ 튜플형, Tuple Type
- ◆ 세트형, Set Type
- ◆ 사전형, Dictionary Type



❖ 데이터 구조, Data Structure

- 데이터를 활용 방식에 따라 조금 더 효율적으로 이용할 수 있도록 컴퓨터에 저장하는 방법들
- 여러 데이터를 하나의 집합 형태로 정의한 데이터형
- 여러 데이터 구조들의 구현을 가능하게 하고 있다.



❖ 리스트형 선언 및 색인

- 여러 값을 동일한 변수에 순차적으로 저장할 수 있는 용도로 사용
- 다른 프로그래밍 언어에서는 '배열(Array)'이라고 부름
- 리스트(List)형은 리스트 안의 값들을 쉼표(,)로 구분하고, 대괄호([])로 감싼다.



❖ 리스트 항목 추가, `append()` 함수

- 맨 뒤에 element 추가

❖ 리스트 항목 제거, `remove()` 함수

- element 값으로 제거

❖ 리스트 항목 삽입, `insert()` 함수

- List의 index 위치에 element 추가

❖ 리스트 항목 추출, `pop()` 함수

- List에서 index 번호로 element 추출 후 삭제



❖ 리스트 기능 중에 리스트를 자르는 방법

- 리스트형 변수명 우측의 대괄호 안에 숫자와 콜론기호(:)를 이용하여 자르면 된다.
- 변수명 [start : end]
 - 리스트형 변수의 start 색인부터 end-1 색인까지 자르세요.

리스트를 제어하는 방법



- ❖ 리스트형 복사하기
- ❖ 리스트형 데이터 합치기
- ❖ 리스트형 데이터 확장하기
- ❖ 리스트 삭제하기

리스트형의 다양한 쓰임새



- ❖ 리스트의 항목에는 어떠한 데이터형이라도 들어갈 수 있다.
- ❖ 동일 리스트라 하더라도 서로 다른 자료형을 넣을 수도 있다.
- ❖ 리스트 자체를 항목으로 갖는 '중첩 리스트'가 가능하다.
- ❖ 문자열 자르기



❖ 튜플형이란?

- 변경할 수 없는(Immutable) 리스트형
- 서로 다른 종류(Heterogeneous)의 데이터형으로 이루어진 항목들을 변수에 바로 풀어 쓰는 언패킹(Unpacking) 혹은 색인을 매기는 용도로 사용
- 각각의 항목을 쉼표(,)로 구분하지만,
- 대괄호 기호([])가 아닌 소괄호 기호(())로 감싸거나 아예 감싸지 않음



❖ 세트형이란?

- 색인에 의한 순서가 없으며 중복이 허용되지 않는 데이터들의 집합

❖ 세트형의 특징

- 데이터 순서는 중요하지 않다.
- 중복된 데이터를 허용하지 않는다.

❖ 세트형 생성은 반드시 `set()` 함수로만 가능하다.



❖ 사전형이란?

- 전체 항목이 정렬되지 않은 키(Key)와 값(Value)의 쌍으로 구성된 집합

❖ 사전형의 특징

- 숫자형이나 문자열형과 같이 변경할 수 없는(Immutable) 데이터형으로 개발자가 정의한 키에 의해 색인이 매겨진다.
- 다른 프로그램 언어에서는 일반적으로 '맵(Map)'이라고 부른다.

❖ 사전형 표기법

- 한 쌍의 키와 값을 '키:값' 형태로 중간에 콜론기호(:)를 사용하여 표기
- 각 쌍은 쉼표(,)로 구분한 뒤 전체 집합은 중괄호 기호({ })로 감싸는 형태



❖ 여러 데이터 구조를 제공할 수 있는 데이터형

- 리스트형 : 변경가능, 중복허용
- 튜플형 : 변경불가, 중복허용
- 세트형 : 변경가능, 중복불가
- 사전형 : 변경가능, 중복불가

❖ 각 자료형에 대한 선언

- Pocket = [3, 8, 4, 7]
- Movie = ('신과함께', '너의결혼식', '공작', '목격자')
- Color = { 'red', 'red', 'blue', 'green', 'blue' }
- Balls = { 'red':4, 'blue':3, 'green':5 }

❖ 외우지 말고 이해를 하려고 노력해라.