2. 데이터 타입의 종류

식별자 및 데이터 타입

식별자

- 식별자(identifier)는 언어에서 변수명, 함수명 등을 지정하는데 사용
- Python 에서는 식별자로 **한글, 영문자(letter), 숫자(digit), 밑줄(_)** 을 사용가능
- 식별자의 **시작은 숫자외 다른 문자**로 시작하며, 영문자는 **대소문자를 구분**함

데이터 타입 종류

• 파이썬은 많은 내장 데이터 타입(built-in data type)들을 가지고 있음

자료형	타입	설명	예
	int	정수형 데이터(integer)	<u>1,</u> 10, -50, -101
숫자	float	실수형 데이터(floating-point number))	1.25, -23.1546
	complex	복수소형 데이터(complex number)	3+3j
불린	bool	참/거짓(<u>boolean</u>)	True
문자열	str	문자열(string)	'Big_data'
리스트	list	순서가 있는 배열, 수정/추가/삭제가 가능한 구조	[1,2,3,'a','b']
튜플	tuple	순서가 있는 배열, 수정/추가/삭제가 불가능한 구조	(1,2,3,'a','b')
딕셔너리	dict	{key, value}키와 값의 쌍으로 데이터를 저장	{'Math':99,'English':88}
집합	set	키로 <u>구성되어 있는</u> 구조, 순서가 <u>없고 중복되지</u> 않는 값들의 집합	{'a','b','c'}

□ 숫자형 타입 : int, float, complex 확인

• type()함수: 변수의 타입 확인 가능

```
In [1]: a = 1
In [2]: b = 1.2345
In [3]: c = 3+4j
In [4]: type(a)
Out[4]: int
In [5]: type(b)
Out[5]: float
In [6]: type(c)
Out[6]: complex
In [7]: d = a + b + c
In [8]: d
Out[8]: (5.2345+4j)
```

□ 문자열(str) 타입

- 큰따음표(" "), 작은따음표(' '), 큰따음표세개(""" """), 작은따음표세개("' ''')로 문자 열 표현 가능
- 따음표 안에 숫자를 써도 문자로 인식
- len(): 문자열 길이

In [37]:	sa = 'Hello Bigdata'
In [38]:	print(sa)
	Hello Bigdata
In [39]:	sb = "Hello Bigdata"
In [40]:	print(sb)
	Hello Bigdata
In [41]:	sc = """Hello Bigdata"""
In [42]:	print(sc)
	Hello Bigdata
In [44]:	sd = '''Hello Bigdata'''
In [45]:	print(sd)
	Hello Bigdata

- 문자열 인덱싱(Indexing)
 - Indexing : 특정한 위치의 데이터만 골라내는 것 ([] 대괄호 사용)
 - 각각의 문자열은 위치에 따른 고유 번호를 가짐
 - a의 2를 인덱싱 할 때 세 번째 문자를 나타냄why) 파이썬에서는 0,1,2,3...으로 0부터 숫자를 셈
 - -1은 뒤에서 부터 숫자를 셈
 - 'this is a string' 이라는 **문자열 자료형의 인덱스 번호**

문	자	t	h	i	3		i	3		a		3	t	r	i	n	g
번	앞	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
호	뒤	-16	-15	-14	-13	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1

- 인덱싱 방법 : **변수이름[번호**]
 - 변수 a 에 문자 'this is a string' 입력
 - a[2] 는 3번째 있는 'i' 출력

In [59]:	a = 'this is a string'
In [60]:	print(a)
	this is a string
In [61]:	print(a[2])
	İ
In [62]:	print(a [0])
	t
In [63]:	print(a[-1])
	g
In [64]:	print(a [5])
	İ
In [65]:	print(a[4])

- 문자열 슬라이싱(Slicing)
 - Slicing : 자료 구조 내에서 여러 개의 항목에 접근
 - Slicing 방법 : **변수이름[시작번호:끝번호**]
 - 슬라이싱에서는'끝번호' 값은 제외하고 출력
 - 앞자리가 시작번호 뒷자리가 끝번호인데 **끝번호는 포함되지 않음**

In	[13]:	1 b = '20200722Wed'
In	[14]:	1 print(b)
"	[14]	
		20200722Wed
In	[15]:	1 print(b[0:4]) #print([0:8])
		2020
In	[16]:	1 print(b[4:6])
		07
In	[17]:	1 print(b[:8]) #print([0:8])
		20200722
In	[18]:	1 print(b[8:]) #8번째 부터 끝까지
		₩ed
In	[19]:	1 print(b[:]) #전제
		20200722Wed
In	[20]:	1 print(b[:-5]) #-8世째까지
		202007

• 문자열 수정

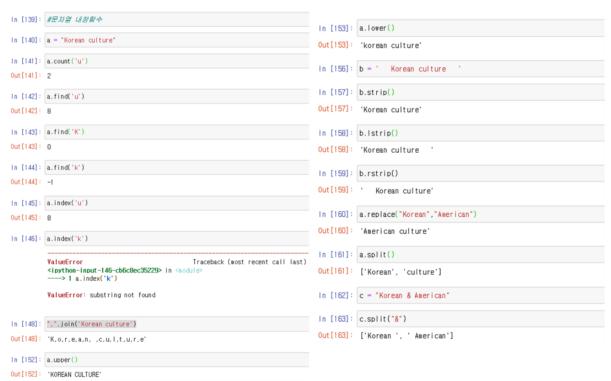
∘ replace()함수 이용하여 해당 문자 변경 가능

- 문자열 format
 - 。 %d: 숫자(정수)
 - %s: 문자 (뒤에 반드시 %"문자"와 같이 ""넣어줘야 함)
 - 어떤 형식이든지 다 받아줌
 - %d, %f 대신 %s 사용 가능
 - 。 %f: 소수 자리수 출력,
 - %.**숫자f**: 표시할 소수점 아래 자릿수
 - 。 여러 개의 문자열을 입력할 때 %(3, "all")와 같이 로 구분하여 추가
 - %10s: %와 s 사이에 숫자를 넣는 경우는 그 숫자 자릿수 공간에서 오른쪽 정렬을 해라는 뜻, 숫자앞에 마이너스(-)를 붙이면 왼쪽 정렬을 하라는 뜻임
 - 소수점 자릿수 표현
 - %0.5f -> 소수점 뒤에 표현될 숫자 개수를 지정(소수점 이하 5자리)
 - %20.5f -> 전체 20자리의 문자열 공간에서 오른쪽 정렬로 소수점 이하 5 자리





- 문자열에 자주 사용하는 함수들
 - 문자열.count('문자') : 한문자에 대한 개수를 샘
 - 문자열.find('문자') : 찾고자 하는 문자의 위치 파악(같은 문자가 있을 땐 가장 앞에 있는 문자 위치, 대소문자 구분함, -1이 결과값으로 나올 땐 해당 문자가 없음)
 - 문자열.index('문자'): find 함수와 같이 찾고자 하는 문자의 위치를 알려주지 만, 문자가 없을 땐 에러가 발생
 - ∘ ",".join('문자'): 문자 각각 사이를 ,로 구분지음
 - 。 문자열.upper(): 소문자를 모두 대문자로 표현
 - 。 문자열.lower(): 대문자를 모두 소문자로 표현
 - 문자열.strip(): 앞뒤에 양쪽의 공백을 모두 없애기
 - 。 문자열.lstrip(): 왼쪽의 공백만 없애기
 - 。 문자열.rstrip(): 오른쪽의 공백만 없애기
 - 문자열.replace("old문자","new문자"): old문자를 new문자로 변경
 - 문자열.split(): 문자열을 나누어 주는 함수. space, tab, enter등을 기준으로 문 자열을 나누어줌. 특정 값을 입력하면 해당 값 기준으로 나누어줌



□ 불린(Boolean) 타입

- bool()함수: **참/거짓을 식별**
- bool() 함수 이용시 문자열, 리스트, 튜플, 딕셔너리가 비어져있으면 False가 되고 그렇지 않으면 True가 됨
- bool() 함수 이용시 숫자일 때 0이면 Flase이고 그렇지 않으면 True가 됨



□ 리스트(LIST) 타입

• 리스트는 대괄호[] 안에 각 요소(숫자, 문자, 숫자+문자, 리스트 안에 리스트 등) 값들은 쉼표(,)로 구분

```
In [164]: #리스트 생성
In [165]: a = [] #a=/ist()
In [166]: b = [1,2,3,4,5]
In [167]: c = ['red', 'blue', 'yellow']
In [168]: d = [1,2,'red','blue']
In [169]: e=[1,2,['red','blue']]
In [170]: print(a)
          []
In [171]: print(b)
          [1, 2, 3, 4, 5]
In [173]: print(c)
          ['red', 'blue', 'yellow']
In [174]: print(d)
          [1, 2, 'red', 'blue']
In [175]: print(e)
          [1, 2, ['red', 'blue']]
```

- 리스트 인덱싱
 - 리스트 인덱싱도 문자열 인덱싱과 마찬가지로 위치를 기반으로 인덱싱함(**0부 터 시작**)
 - 인덱싱 방법 : **변수이름[번호**]

```
In [177]: b[0] + b[2] #첫번째 요소와 세번째 요소 함
In [164]: #리스트 생성
                                                                          Out[177]: 4
In [165]: a = [] #a=/ist()
                                                                          In [178]: b[-1] #뒤에서 첫번째
                                                                          Out [178]: 5
In [166]: b = [1,2,3,4,5]
                                                                          In [179]: a = [1,2,3,4,5,['a','b','c','d'],7,8,9]
In [167]: c = ['red','blue','yellow']
In [168]: d = [1,2,'red','blue']
                                                                                   [1, 2, 3, 4, 5, ['a', 'b', 'c', 'd'], 7, 8, 9]
                                                                          In [181]: a[0]
In [169]: e=[1,2,['red','blue']]
                                                                          Out[181]: 1
In [170]: print(a)
                                                                          In [182]: a[-1]
          []
                                                                          Out[182]: 9
In [171]: print(b)
                                                                          In [185]: a[5] #List 안에 6번째
          [1, 2, 3, 4, 5]
                                                                          Out[185]: ['a', 'b', 'c', 'd']
                                                                          In [186]: a[5][2] #List 안에 6번째 인덱상 해서 다시 한번 세번째 요소 인덱상
In [173]: print(c)
          ['red', 'blue', 'yellow']
In [174]: print(d)
          [1, 2, 'red', 'blue']
In [175]: print(e)
          [1, 2, ['red', 'blue']]
In [176]: b[0] #b의 첫번째 값
Out [176]: 1
```

• 리스트 슬라이싱

- Slicing 방법 : **변수이름[시작번호:끝번호**]
 - 슬라이싱에서는'끝번호' 값은 제외하고 출력
 - 앞자리가 시작번호 뒷자리가 끝번호인데 **끝번호는 포함되지 않음**

```
In [187]: #슬라이성
                                                                       In [197]: a[0:6:1] #0:6의 안에서 한자리 수마다 출력해라.a[0:6] 동일한 결과
                                                                       Out[197]: [1, 2, 3, 4, 5, ['a', 'b', 'c', 'd']]
In [188]: b = [1,2,3,4,5]
In [190]: b[0:4] #첫번째 부터 네번째 까지 실행
                                                                       In [198]: a[0:6:2] #0:6의 안에서 두자리 수마다 출력해라
Out [190]: [1, 2, 3, 4]
                                                                       Out[198]: [1, 3, 5]
In [191]: c = "12345"
                                                                       In [199]: a[0:6:3] #0:6의 안에서 세자리 수마다 출력해라
                                                                       Out[199]: [1, 4]
In [192]: c[0:4]
Out [192]: '1234'
                                                                       In [200]: a[5]
In [193]: b[:3] #첫번째 부터 세번째 까지 실행
                                                                       Out[200]: ['a', 'b', 'c', 'd']
Out[193]: [1, 2, 3]
                                                                       In [201]: a[5][:2]
In [194]: b[3:] #네번째 부터 끝까지
                                                                       Out [201]: ['a', 'b']
Out[194]: [4, 5]
In [195]: a = [1,2,3,4,5,['a','b','c','d'],7,8,9]
In [196]: a[0:6]
Out[196]: [1, 2, 3, 4, 5, ['a', 'b', 'c', 'd']]
```

• 리스트 연산

```
In [202]: #완산
                                                                                        In [215]: c = ['good'] #리스트로 변경
In [203]: a=[1,2,3,4]
                                                                                        In [216]: b+c
In [204]: b=[6,7,8,9]
                                                                                        Out[216]: [6, 7, 8, 9, 'good']
In [205]: a+b
                                                                                        In [217]: #수정
Out[205]: [1, 2, 3, 4, 6, 7, 8, 9]
                                                                                        In [218]: a[3]
In [206]: a+2
                                                                                        Out [218]: 4
Out[206]: [1, 2, 3, 4, 1, 2, 3, 4]
                                                                                        In [219]: a[3]=9
In [208]: Ien(a)
Out [208]: 4
                                                                                        In [220]: print(a)
In [209]: k=[12345678]
                                                                                                  [1, 2, 3, 9]
In [210]: Ien(k)
                                                                                        In [221]: n='1234'
Out [210]: 1
                                                                                        In [222]: n[3]
In [211]: n='12345678'
                                                                                        Out [222]: '4'
In [212]: len(m)
                                                                                        In [223]: n[3]='9' #문자열 수정 불가
Out [212]: 8
In [213]: c='good'#문자일
In [214]: b+c #리스트 + 문자열
          TypeError Traceback (most recent call last)
<ipython-input-214-5862b4a280fe> in <module>
----> 1 b+c #리스트 + 문자열
         TypeError: can only concatenate list (not "str") to list
```

```
| Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition | Solition
```

• 리스트 기초 연산 함수

o sum:합, max:최대값, min:최소값, del:삭제, str:문자형으로 변경



- append: 리스트 추가, sort: 올림차순으로 정렬(숫자, 문자 정렬가능)
- reverse: 현재 리스트의 순서를 반대로 출력
- o index: 해당 값의 위치를 출력, 없는 값 입력하면 에러발생
- insert: 특정 위치에 값을 추가



- remove: 특정 값을 삭제(list안에 동일한 값이 여러개 있을 때 맨 앞의 값만 삭제)
- pop: 빈칸 입력시 리스트의 마지막 값을 보여주고 삭제시킴. 입력시 삭제 요소의 위치 지정

```
In [279]: e=['3','4','3','4','5']
In [280]: e.remove('4') #'4'를 제거(여러개 있으면 맨 앞의값 제거)
In [281]: print(e)
        ['3', '3', '4', '5']
In [282]: e.remove('4') #'4'를 제거
In [283]: print(e)
        ['3', '3', '5']
In [284]: e.pop() #리스트에서 마지막 요소 보여주고 삭제
Out [284]: '5'
In [285]: print(e)
        ['3', '3']
In [286]: e.pop(1) #삭제시킬 요소 입력, 1번삭제(2번째요소)
Out [286]: '3'
In [287]: print(e)
        ['3']
In [288]: e.pop(0)
0@[2@ount: 동일한 값 개수 세기
In [2@xtendt(e)리스트의 요소 추가 (괄호 안에 리스트만 올 수 있음, append와의
    차이미구분)
In [290]: c = [6,6,6,6,6,7,8,9]
In [292]: c.count(6) #6의 갯수 세기
Out [292]: 5
In [293]: c.extend([4,5]) #괄호안에 리스트만 올수 있음.리스트의 요소만 추가됨
In [294]: print(c)
          [6, 6, 6, 6, 6, 7, 8, 9, 4, 5]
In [295]: c.append([4,5])
In [296]: print(c)
          [6, 6, 6, 6, 6, 7, 8, 9, 4, 5, [4, 5]]
```

☐ 튜플(TUPLE) 타입

- 튜플 타입은 리스트 타입과 비슷
- 차이점은 **리스트는 대괄호** []를 **사용하고 튜플은 소괄호()를 사용(**괄호생략가능) 하며, **리스트는 수정이 가능하지만 튜플은 수정이 불가능함**

- 튜플에서는 괄호안에 한 개의 요소를 가질 때는 (1,)와 같이 마지막에 콤마 필요
- 콤마(,)를 붙이지 않으면 튜플이 아닌 정수 타입이 생성
- 프로그램을 실행되는 동안 **그 값이 항상 변화지 않아야 한다**면, 튜플을 사용하고 값이 자주 바뀐다면 리스트를 사용

```
In [297]: #튜플의 생성
In [298]: a = tuple()
In [299]: a=()
In [300]: type(a)
Out[300]: tuple
In [303]: b = (1,) #하나만 입력시 만드시 콤마 추가
In [306]: type(b)
Out[306]: tuple
In [307]: c = (1,2,3)
In [308]: type(c)
Out[308]: tuple
In [309]: d = 1,2,3 #괄호 생략 가능
In [310]: type(d)
Out[310]: tuple
In [311]: e = 1,2,('ab','cd'),3,4
In [313]: type(e)
Out[313]: tuple
```

• 튜플은 수정 및 변경 불가

```
In [311]: e = 1,2,('ab','cd'),3,4
In [313]: type(e)
Out[313]: tuple
In [314]: e[0]
Out [314]: 1
In [315]: del e[0]
         TypeError
                                               Traceback (most recent call last)
         <ipython-input-315-88f9a3ba7b63> in <module>
            --> 1 del e[0]
         TypeError: 'tuple' object doesn't support item deletion
In [316]: e.remove[1]
         AttributeError
                                               Iraceback (most recent call last)
         <ipython-input-316-92eade322d4d> in <module>
           ---> 1 e.remove[1]
         AttributeError: 'tuple' object has no attribute 'remove'
In [318]: e[:] #요소 전체 출력
Out[318]: (1, 2, ('ab', 'cd'), 3, 4)
In [319]: c+e
Out[319]: (1, 2, 3, 1, 2, ('ab', 'cd'), 3, 4)
In [320]: c*2
Out[320]: (1, 2, 3, 1, 2, 3)
In [321]: len(c)
Out[321]: 3
In [322]: x,y,z=(1,2,3)
 In [324]: print(x,y,z)
             123
 In [325]: print(x)
             1
```

☐ 사전(DICT)타입

- key와 value를 한 쌍으로 가지고 있는 자료형
- {key1:value1,key2:value2,key3:value3...}

• 위치자리에 key값을 입력하면 value 출력

```
In [326]: #사전의 생성
In [328]: #{key1,:value1,key2:value2,...}
In [329]: a ={'gender':'Female', 'age':30, 'name':'Kristina'}
In [330]: type(a)
Out[330]: dict
In [332]: a['gender'] #위치자리에 key를 입력하면, key해당되는 value가 반환됨
Out[332]: 'Female'
In [333]: a['age']
Out [333]: 30
In [334]: a[0] #기존 방법처럼 위치를 지정하면 에러남
                                                 Traceback (most recent call last)
          <ipython-input-334-6a1284577a36> in <module>
          ----> 1 a[0]
          KeyError: 0
 In [1]: a['Krestina']
                                                 Traceback (most recent call last)
          <ipython-input-1-24c09df5fdd1> in <module>
          ----> 1 a['Krestina']
          NameError: name 'a' is not defined
```

• 사전에서 추가 및 삭제하기

```
In [12]: b = {1:'good'} #key자리에 정수값 1을 넣고 dictionary 만든, 실행가능
                                                                       In [25]: d = {(1,3,5):'a'} #튜플(바괄수없음)은 key자리에 사용가능
In [13]: print(b)
                                                                       In [26]: a = {'gender':'Femal'}
                                                                        In [27]: a['age']=30 #age = 30 key value 추가
In [17]: c = {'a':[1,3,5]} #Key자리에 문자열, value자리에 리스트 [1,3,5]입력
                                                                       In [28]: print(a)
In [18]: print(c)
                                                                                 {'gender': 'Femal', 'age': 30}
        {'a': [1, 3, 5]}
                                                                        In [29]: a['name'] = 'Kristina' #key value 추가
In [21]: d = {'a':(1,3,5)} #튜플도 덕셔너리 지정 가능
                                                                        In [30]: print(a)
In [22]: print(d)
                                                                                 {'gender': 'Femal', 'age': 30, 'name': 'Kristina'}
        {'a': (1, 3, 5)}
In [23]: c = {[1,3,5:'a']} #리스트(수정가능)는 key자리에는 사용 불가
                                                                       In [32]: a['a'] = [1,3,5] #[1,3,5] 리스트 추가
          File "<ipython-input-23-fd9f75c50466>", line 1
c = {[1,3,5:'a']} #리스트(수정가능)는 key자리에는 사용 불가
                                                                       In [33]: print(a)
                                                                                 {'gender': 'Femal', 'age': 30, 'name': 'Kristina', 'a': [1, 3, 5]}
        SyntaxError: invalid syntax
                                                                       In [34]: del a['a'] #사전의 삭제, 해당key와 value 한쌍이 삭제
                                                                       In [35]: print(a)
                                                                                 {'gender': 'Femal', 'age': 30, 'name': 'Kristina'}
                                                                       In [36]: del a['name']
                                                                       In [37]: print(a)
                                                                                 {'gender': 'Femal', 'age': 30}
```

• 사전 함수

∘ keys: 전체 key출력

∘ values: key값 출력

∘ items: 전체 key와 value를 쌍으로 출력

。 clear: 딕셔너리안에 요소 삭제

o get: key값(value) 출력

```
In [51]: #사전 함수
                                                                              In [66]: a.clear() #덕셔너리안에 모든 요소들을 삭제
In [61]: a ={'gender':'Female', 'age':30, 'name':'Kristina'}
                                                                              In [67]: printer(a)
In [62]: a.keys() #key 출력,파이썬3,0이상부터 적용
                                                                                      NameError 1
<ipython-input-67-5255153360e2> in <module
                                                                                                                             Traceback (most recent call last
Out[62]: dict_keys(['gender', 'age', 'name'])
In [63]: list(a.keys())
                                                                                      NameError: name 'printer' is not defined
Out [63]: ['gender', 'age', 'name']
                                                                              In [68]: a ={'gender':'Female', 'age':30, 'name':'Kristina'}
In [64]: a.values() #전체 value값 출력
Out[64]: dict_values(['Female', 30, 'Kristina'])
                                                                              In [69]: a.get('gender') #a['gender']
In [65]: a.items()#key, value(key값)를 묶은 한쌍을 출력
Out[65]: dict_items([('gender', 'Female'), ('age', 30), ('name', 'Kristina')])
                                                                             In [70]: a.get('height','?') #key가 딕쳐너리 안에 아무값 없을때, ?로 표현
                                                                             In [71]: 'gender' in a # 값이 있으면 Ture
                                                                             Out [71]: True
                                                                             In [72]: 'height' in a # 200 22 Palse
                                                                             Out [72]: False
```