

다양한 FILE I/O

PYTHON 파일 입출력

◆ CSV 파일

- >> 데이터 값을 쉼표(,)로 구분하는 파일
- >> Comma Separated Values 약자
- >> 쉼표(,)로 열 구분, 줄바꿈으로 행(row) 구분
- >> TSV(Tab), SSV(Space) 데이터 파일도 존재

PYTHON 파일 입출력

◆ CSV 파일

형태 → 데이터,데이터,데이터

1,13.2,9.1,blue

1,98,2.8,gray

2,10.5,81.3,red

1,8.8,5.21,yellow

1;13.2;9.1;blue

1;98;2.8;gray

2;10.5;81.3;red

1;8.8;5.21;yellow

PYTHON 파일 입출력

◆ CSV 파일

import csv

- <https://docs.python.org/ko/3.7/library/csv.html?highlight=csv>
- csv 파일 처리 표준 라이브러리

csv.reader(file) ➔ 읽은 csv 데이터 문자열 리스트 반환

csv.writer(file) ➔ csv 파일 데이터 쓰기 위한 객체 반환

writer_obj.writerow(row) ➔ 한 줄 쓰기

writer_obj.writerows([row, row, ..]) ➔ 여러 줄 쓰기

PYTHON 파일 입출력

◆ EXCEL파일

>> MS사의 Excel

>> xlsx 확장자지원

>> 설치

```
pip install openpyxl== 3.0.1
```

<https://pypi.org/project/openpyxl/3.0.1/>

PYTHON 파일 입출력

◆ JSON 파일

- >> JavaScript Object Notation 약자
- >> 자바스크립트에서 사용하는 객체 표기 방법
- >> 다양한 프로그래밍 언어에서 데이터 교환에 사용
- >> 인코딩/디코딩 표준으로도 사용
- >> <https://docs.python.org/ko/3.7/library/json.html?highlight=json#module-json>

PYTHON 파일 입출력

◆ JSON 파일

형태 : [{ 키:값, 키:값, 키:{ 키:값, 키:값 } }]

```
[ { 'id':1,  
  'name':'jane',  
  'data' : { 'major':'science',  
             'grade': 1 }  
}, { 'id':2,  
  'name':'Tom',  
  'data' : { 'major':'math',  
             'grade': 2 }  
} ]
```

PYTHON 파일 입출력

◆ JSON 파일

```
import json
```

```
with open('../Data/test.json', 'w') as f:  
    # json파일 쓰기  
    json.dump(json_data , f)
```

```
with open('../Data/test.json', 'r') as f:  
    # json파일 읽기  
    json_data = json.load(f)
```


바이너리 FILE I/O

PYTHON 파일 입출력

◆ 바이너리 파일

- 이진 파일 또는 바이너리 파일(binary file)
- 컴퓨터 저장과 처리 목적 위해 이진 형식으로 인코딩된 데이터 파일
- 문서 편집기로 열었을 경우 알아볼수 없는 문자들

[바이너리 데이터 파일]

분류	파일 종류
이미지 파일	- jpg, png,
오디오 파일	- mp3, mp4, ...
실행 파일	- exe, bin,

PYTHON 파일 입출력

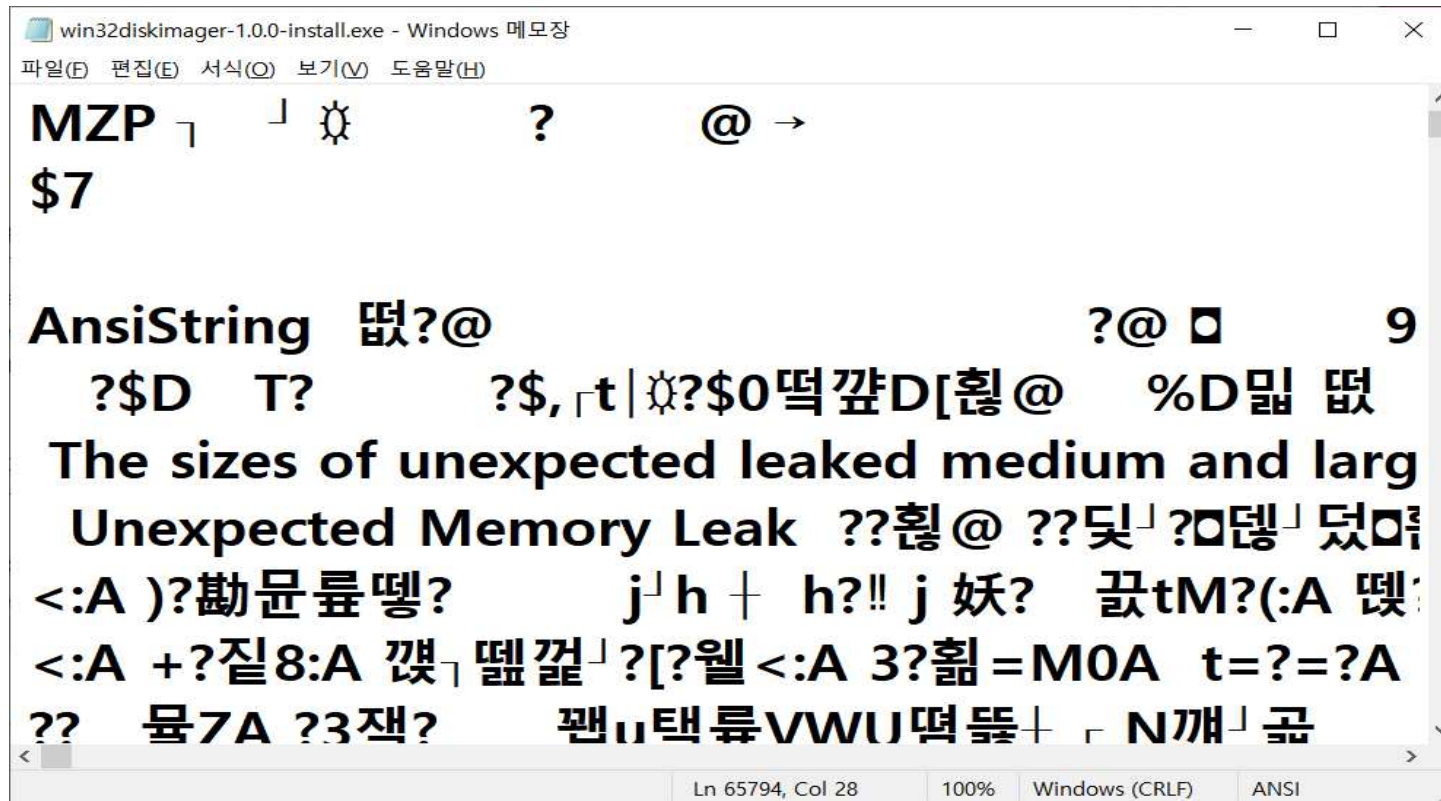
◆ 바이너리 데이터 타입

bytes 데이터 타입

- 1바이트 단위 값을 연속적으로 저장하는 시퀀스 자료형
- 1바이트 => 8비트
- 0~255(0x00~0xFF)까지 정수 사용

PYTHON 파일 입출력

◆ 바이너리 파일



win32diskimager-1.0.0-install.exe - Windows 메모장

파일(F) 편집(E) 서식(O) 보기(V) 도움말(H)

MZIP ǀ ǂ ǃ ? @ →
\$7

AnsiString 덧?@ ?@ 9
?\$D T? ?\$, ǂt|ǃ?\$0떡깟D[훤@ %D밍 덧
The sizes of unexpected leaked medium and larg
Unexpected Memory Leak ??훤@ ??딛 ǂ ǃ덧 ǂ덧
<:A)?勘문릍덧? j ǂh + h?!! j 妖? ǃtM?(:A 덧
<:A +?질8:A ǃ ǃ덧 ǃ ǃ[?웰<:A 3?훤=M0A t=?=?A
?? ㄴ7A ?3잭? ǃ ǂ덧 ǃVWU떡덧+ ǂ N개 ǃ ǃ

Ln 65794, Col 28 100% Windows (CRLF) ANSI

PYTHON 파일 입출력

◆ 바이너리 데이터 타입

bytes 객체 생성

bytes(숫자):	숫자만큼 0으로 채워진 바이트 객체 생성
bytes(반복가능한객체)	반복 가능한 객체로 바이트 객체 생성
bytes(b'바이트객체')	바이트 객체로 바이트 객체 생성

PYTHON 파일 입출력

◆ 바이너리 모듈

import struct

- C언어의 구조체를 구현한 모듈
- 파일이나 네트워크 연결에 사용하는 이진 데이터 다루는 모듈

Format	C Type	Python type	Standard size	Notes
x	pad byte	no value		
c	char	bytes of length 1	1	
b	signed char	integer	1	(1), (2)
B	unsigned char	integer	1	(2)
?	_Bool	bool	1	(1)
h	short	integer	2	(2)
H	unsigned short	integer	2	(2)
i	int	integer	4	(2)
I	unsigned int	integer	4	(2)
l	long	integer	4	(2)
L	unsigned long	integer	4	(2)
q	long long	integer	8	(2)