

## 파일 입출력

## 파일 읽고 쓰기



출처: R, Python 분석과 프로그래밍의 친구 <http://rfriend.tistory.com>

## 문자열로 형태로 파일 읽기

- `f.read()`는 파일을 전부 읽은 문자열을 리턴

```
f = open('README.txt', 'r')
data = f.read()
print(data)
f.close()
```

- 파일에서 한 라인씩 가져오기

```
f = open('README.txt', 'r')
for line in f:
    print(line)
f.close()
```

## 리스트 형태로 파일 읽기

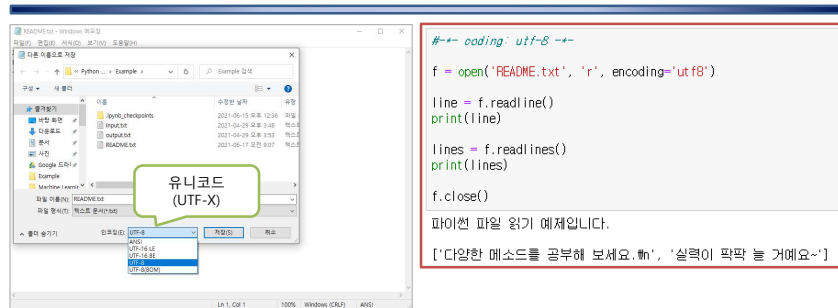
- `f.readline()`
  - 읽기 모드로 열린 파일의 한 라인을 읽어서 문자열로 리턴
  - 반복문을 이용하여 리스트 생성
- `f.readlines()`
  - 읽기 모드로 열린 파일의 모든 라인을 한꺼번에 읽어서 리스트로 리턴

원도우즈 인코딩 방식 (ANSI, CP949)

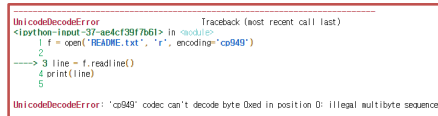
```
f = open('README.txt', 'r', encoding='cp949')
line = f.readline()
print(line)
lines = f.readlines()
print(lines)
f.close()

파이션 파일 읽기 예제입니다.
['다양한 메소드를 공부해 보세요.\n', '실력이 팍팍 늘 거예요~']
```

## 다른 인코딩 방식의 파일 읽기



인코딩/디코딩 방법이 다른 경우 예외 발생



## CSV 파일 읽기

### • CSV 파일

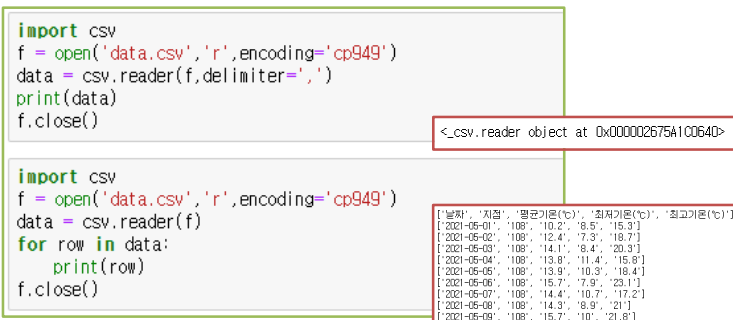
- 콤마로 분리된 값(Comma Separated Values)으로 구성된 파일
- 엑셀 데이터는 CSV 파일로 저장 가능

날짜	지점	평균기온(°C)	최저기온(°C)	최고기온(°C)
2021-05-01	108	10.2	8.5	15.3
2021-05-02	108	12.4	7.3	18.7
2021-05-03	108	14.1	8.4	20.3
2021-05-04	108	13.8	11.4	15.8
2021-05-05	108	13.9	10.3	18.4
2021-05-06	108	15.7	7.9	23.1
2021-05-07	108	14.4	10.7	17.2
2021-05-08	108	14.2	8.9	21
2021-05-09	108	15.7	10	21.8
2021-05-10	108	13.2	11.7	14.6
2021-05-11	108	18.5	9.4	25.7
2021-05-12	108	22	16.9	27.7
2021-05-13	108	22.9	16.6	29.6
2021-05-14	108	24.1	18	30.8
2021-05-15	108	22.1	20.8	26.2
2021-05-16	108	18.4	14.8	20.8
2021-05-17	108	15.3	14.2	16.7
2021-05-18	108	17.9	13.6	24.5

## CSV 파일 읽기

### • csv 라이브러리 이용

- 파일 열고 닫는 것은 일반 파일과 동일
- 파일 읽을 때 csv.reader() 이용



## 파일 쓰기

### • f.write(data)

- 쓰기 모드로 열린 파일에 데이터를 씀

```
f = open('new.txt', 'w')

for i in range(1,11):
    data = "{0:d} line.\n".format(i)
    f.write(data)

f.close()
```

new.txt

```
1 line.
2 line.
3 line.
4 line.
5 line.
6 line.
7 line.
8 line.
9 line.
10 line.
```

## 추가 모드로 파일 쓰기

- 'w' 옵션으로 file을 열면 기존의 내용은 사라짐
- 기존의 내용에 새로운 내용을 추가하고자 할 땐 'a' 옵션을 사용

```
f = open('new.txt', 'a')

for i in range(11,21):
    data = "{0:d} line.\n".format(i)
    f.write(data)

f.close()
```

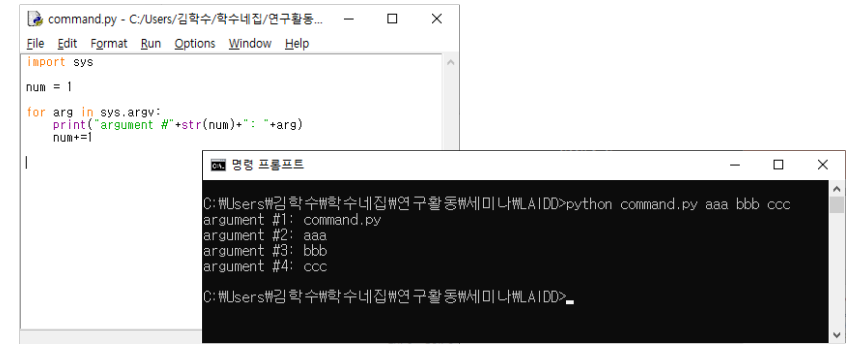
new.txt

```
1 line.
2 line.
3 line.
...
10 line.
11 line.
12 line.
...
19 line.
20 line.
```

## 명령행 인자 (command line argument)

### • 명령행 인자: sys 모듈 이용

- Dos나 Unix shell에서 Python명령어를 실행하면서 입력인수를 넣어줄 수 있음



## 실습

- 명령행 인자(프로그래밍 제외)를 입력 받아 모두 대문자로 변환하는 프로그램을 작성하시오.

?

## 실습

- "README.txt" 파일에 있는 각 단어의 수를 계산하여 출력하는 프로그램을 작성하시오.

README.txt - Windows 메모장

파일(F) 편집(E) 서식(O) 보기(V) 도움말(H)

파이션 파일 읽기 예제입니다.  
다양한 예소드를 공부해 보세요.  
실력이 팍팍 늘 거예요~  
단어 수를 카운트하는 파이션 파일 읽기 실습 예제입니다.

```
파이션 2
파일 2
읽기 2
예제입니다. 2
다양한 1
예소드를 1
공부해 1
보세요. 1
실력이 1
팍팍 1
늘 1
거예요~ 1
단어 1
수들 1
카운트하는 1
실습 1
```

?

## 실습

---

- Unix의 cp (copy) 명령어를 python으로 구현하시오.
  - 예: python cp.py file1 file2
  - file1의 내용을 file2에 복사 (file2는 새로 생성)



Edited by Harksoo Kim

## 질의응답

---

# Q&A

Homepage: <http://nlp.konkuk.ac.kr>  
E-mail: [nlpdrkim@konkuk.ac.kr](mailto:nlpdrkim@konkuk.ac.kr)



Edited by Harksoo Kim