

빅데이터 프로그래밍

5회차: 파일 및 csv 다루기 - 1

목차

5회차: 파일 및 csv 다루기 - 1

- 복습
- 파일 다루기
 - ▶ 텍스트 파일 읽기
 - ▶ 텍스트 파일 쓰기
 - ▶ CSV 파일 다루기

복습

지난 강의 때 뭐했지..?

01

■ 함수, 클래스

복습

02

지난 강의 때 뭐했지..?

■ 함수, 클래스

```
class Email():
    from_email = ''
    to_email = ''
    subject = ''
    contents = ''

    def send_mail(self):
        print('Send to ' + self.to_email)
```

복습

03

지난 강의 때 뭐했지..?

■ 함수, 클래스

```
from my_email import Email
from my_news import News
from my_excel import Excel

m_email = Email()
m_news = News()
m_excel = Excel()

news_list = m_news.find_news('python')

m_email.from_email = 'alghost.lee@gmail.com'
m_email.to_email = 'yskim@fastcampus.com'
m_email.subject = 'Dear. '

for news in news_list:
    m_email.contents = m_email.contents + news + '\n'

m_email.send_mail()

m_excel.excel_file = 'result.xlsx'
m_excel.save_to_excel(news_list)
```

파일 다루기

어디에 쓰일까?

04

■ 파일 입출력

- 텍스트 파일에 대한 읽기/쓰기를 다룸
- 어디에 쓰일까?
 - ▶ 텍스트파일의 데이터를 읽고 쓰기 위해
 - ▶ 자동화 스케줄링시 입력값/결과값으로 사용
 - 특정파일의 값을 입력값으로, 특정파일에 결과값을 씀
 - ▶ csv 파일의 데이터를 읽고 쓰기 위해

파일 다루기

텍스트 파일 읽기

05

■ 파일을 읽어보자

- 파일 입출력은 기본적으로 텍스트 파일을 의미
- 텍스트 편집기를 이용하여 아무 내용이나 작성한 후 코드에서 읽어보자
 - ▶ 아래 작성된 텍스트를 data.txt 라는 파일명으로 저장

```
1 안녕하세요 .  
2 파일 읽기 테스트를 위한 글입니다 .  
3 여러분들은 다른 텍스트를 작성해보세요 :)  
4 후후
```

파일 다루기

텍스트 파일 읽기

06

■ 파일을 읽어보자

- 파일 입출력은 기본적으로 텍스트 파일을 의미
- 텍스트 편집기를 이용하여 아무 내용이나 작성한 후 코드에서 읽어보자
- encoding: 파일을 읽을 때 사용할 문자 해석 방법
 - ▶ utf-8: 윈도우를 제외하고 많이 일반적으로 쓰이는 방법
 - ▶ cp949: 윈도우에서 사용하는 방법
 - ▶ 하지만, **vscode**에서는 'utf-8'을 사용

```
datafile = open('data.txt', 'r', encoding='utf-8')
data = datafile.read()
print(data)
```


파일 다루기

텍스트 파일 읽기

07

■ 파일을 읽어보자

- 파일 입출력은 기본적으로 텍스트 파일을 의미
- 텍스트 편집기를 이용하여 아무 내용이나 작성한 후 코드에서 읽어보자

```
alghost:4th Alghost$ python3 007.py
안녕하세요.
파일 읽기 테스트를 위한 글입니다.
여러분들은 다른 텍스트를 작성해보세요 :)
후후
```

파일 다루기

텍스트 파일 읽기

08

■ 파일을 한줄씩 읽어보자

- 파일 입출력은 기본적으로 텍스트 파일을 의미
- 텍스트 편집기를 이용하여 아무 내용이나 작성한 후 코드에서 읽어보자

```
datafile = open('data.txt', 'r', encoding='utf-8')  
  
line = 'init'  
while line:  
    line = datafile.readline()  
    print(line)
```

파일 다루기

텍스트 파일 읽기

09

■ 파일을 한줄씩 읽어보자

- 파일 입출력은 기본적으로 텍스트 파일을 의미
- 텍스트 편집기를 이용하여 아무 내용이나 작성한 후 코드에서 읽어보자

```
alghost:4th Alghost$ python3 009.py  
안녕하세요.
```

```
파일 읽기 테스트를 위한 글입니다.
```

```
여러분들은 다른 텍스트를 작성해보세요 :)
```

```
후후
```

파일 다루기

텍스트 파일 쓰기

10

■ 파일에 써보자

- 파일 입출력은 기본적으로 텍스트 파일을 의미
- 사용자 입력을 받아서 입력받은 내용을 파일에 써보자
 - ▶ 입력받은 내용을 textfile.txt에 저장

```
user_input = input('User input: ')\ndfile = open('textfile.txt', 'w', encoding='utf-8')\ndfile.write(user_input+'\n')
```

파일 다루기

텍스트 파일 쓰기

11

■ 파일에 써보자

- 파일 입출력은 기본적으로 텍스트 파일을 의미
- 사용자 입력을 받아서 입력받은 내용을 파일에 써보자

```
alghost:4th Alghost$ python3 011.py
User input: alghost
alghost:4th Alghost$ cat textfile.txt
alghost
```

파일 다루기

텍스트 파일 쓰기

12

■ 파일에 추가로 써보자

- 파일 입출력은 기본적으로 텍스트 파일을 의미
- 사용자 입력을 받아서 입력받은 내용을 파일에 추가로 써보자

```
user_input = input('User input: ')\ndfile = open('textfile.txt', 'a', encoding='utf-8')\ndfile.write(user_input+'\n')
```

파일 다루기

텍스트 파일 쓰기

13

■ 파일에 추가로 써보자

- 파일 입출력은 기본적으로 텍스트 파일을 의미
- 사용자 입력을 받아서 입력받은 내용을 파일에 추가로 써보자

```
alghost:4th Alghost$ python3 013.py
User input: 처음 입력입니다
alghost:4th Alghost$ python3 013.py
User input: 두번째
alghost:4th Alghost$ python3 013.py
User input: 세번째
alghost:4th Alghost$ cat textfile.txt
alghost
처음 입력입니다
두번째
세번째
```

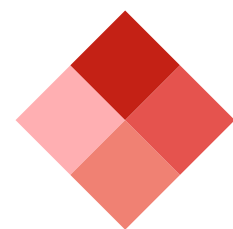
파일 다루기

csv 파일 다루기

14

■ CSV 파일이란

- Comma separated values
- 콤마로 컬럼이 구분되고 개행으로 데이터가 구분되는 형태
- 탭으로 컬럼을 구분하는 TSV 파일도 있음
- split 과 join을 통해 실습



Good Bye

See you next time