

07. 파일 다루기

이번 장에서는 파이썬을 이용하여 파일로부터 데이터를 읽거나 반대로 프로그램의 결과값을 파일로 저장하는 방법을 배워보겠습니다. 참고로 알고리즘 트레이딩 분야에서는 통계 데이터를 틈틈이 파일로 저장하여 여러분만의 데이터베이스를 구축하는 것이 매우 중요합니다. 이번 장을 통해서 파이썬을 통한 파일 입/출력에 익숙해지기 바랍니다.

1) 파일 읽기

컴퓨터에서 파일이란 무엇인가요? 여러분이 프로그래밍에 대해서는 아직은 잘 모르지만, 한글 파일(hwp) 또는 마이크로소프트의 워드 파일(doc)이란 말은 종종 들어 봤을 겁니다. 우리가 자주 사용하는 문서 파일 이외에도 영화 파일(*.avi)이나 음악 파일(.mp3)들도 파일이란 용어를 사용하고 있습니다.

위키피디아를 참고해보면 파일은 컴퓨터 등의 기기에서 의미 있는 정보를 담은 논리적인 단위라고 합니다. 앞서 설명한 것처럼 컴퓨터에서 파일은 그 종류가 매우 다양합니다. 그중 가장 쉽게 만들 수 있는 파일 형태는 아마 텍스트 파일일 것입니다. 참고로 텍스트 파일은 윈도우의 메모장등을 이용하여 간단히 텍스트를 입력함으로써 쉽게 만들 수 있습니다.

파이썬을 이용한 파일 읽기에 앞서 윈도우 메모장을 이용하여 주식 매수 리스트를 텍스트 파일로 만들어 보도록 하겠습니다. 윈도우 10에서는 메모장 앱을 실행하려면 그림 7.1 과 같이 검색 기능을 활용하면 됩니다.

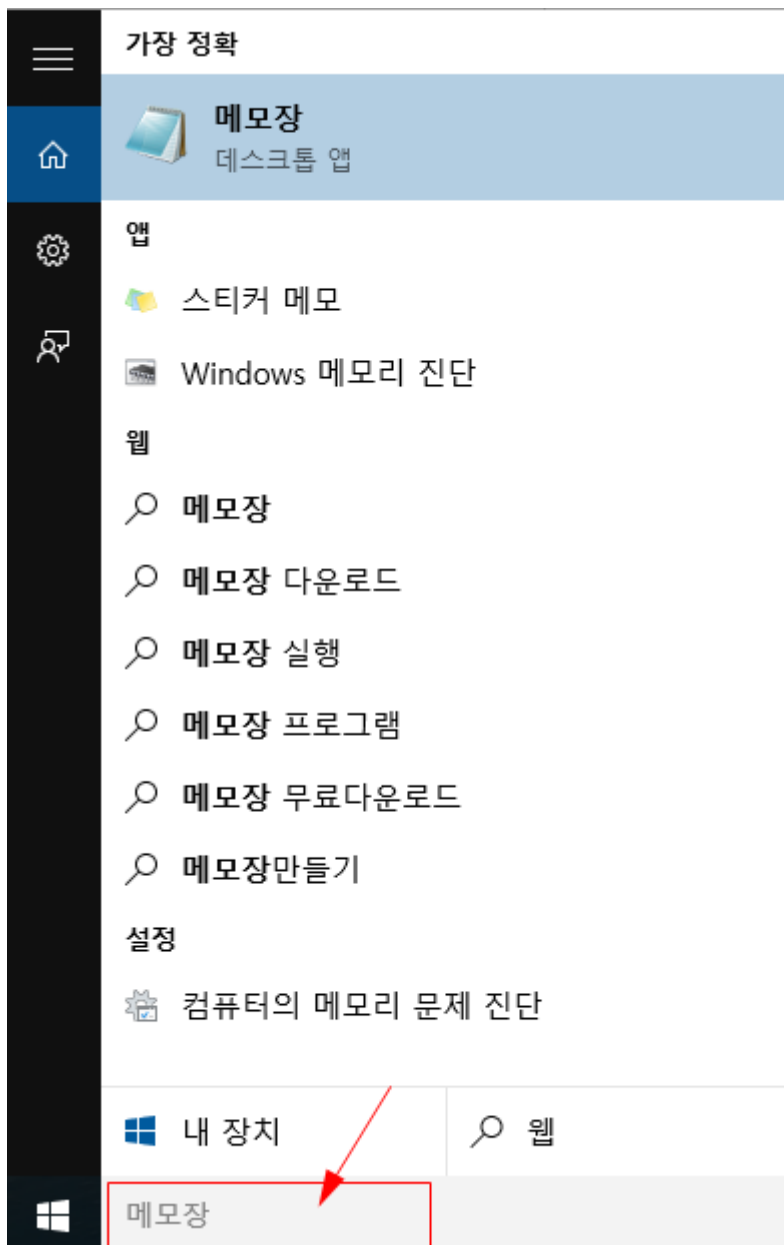


그림 7.1 윈도우 10 을 이용한 메모장 앱 실행

메모장 프로그램을 실행했다면 그림 7.2 와 같이 NAVER, 현대차, 기아차를 입력해보기 바랍니다. 이때 중요한 점은 구분자를 사용하여 종목을 서로 구분해줘야 한다는 점입니다. 참고로 저자는 그림 7.2 와 같이 종목명을 입력한 후 엔터키를 눌러 종목별로 구분될 수 있도록 했습니다.

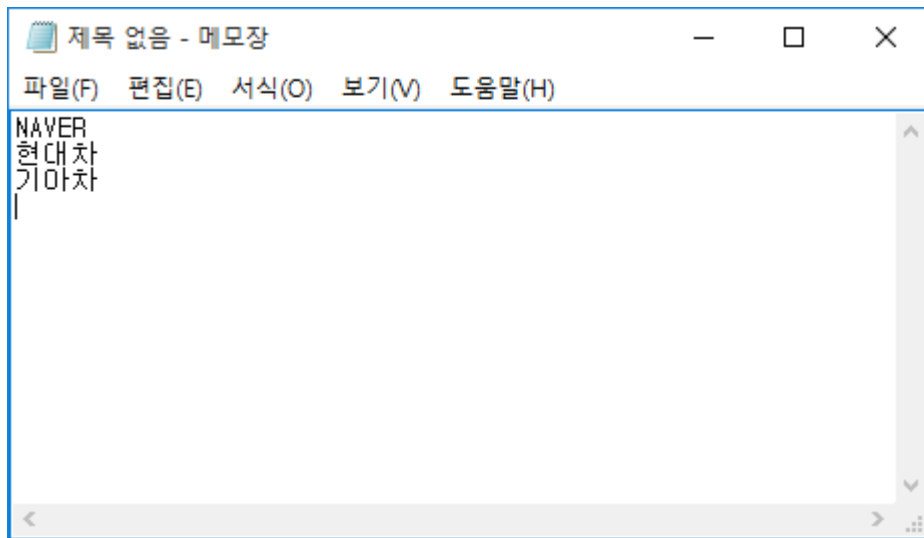


그림 7.2 매수 리스트

그림 7.2 와 같이 메모장에 매수 리스트를 작성한 뒤에 메모장의 파일 메뉴를 누른 후 다른 이름으로 저장을 클릭합니다. 그림 7.3 과 같은 다른 이름으로 저장화면에서 먼저 저장 경로로 바탕 화면을 선택한 후 파일 이름을 buy_list 로 입력하고 저장 버튼을 클릭합니다.

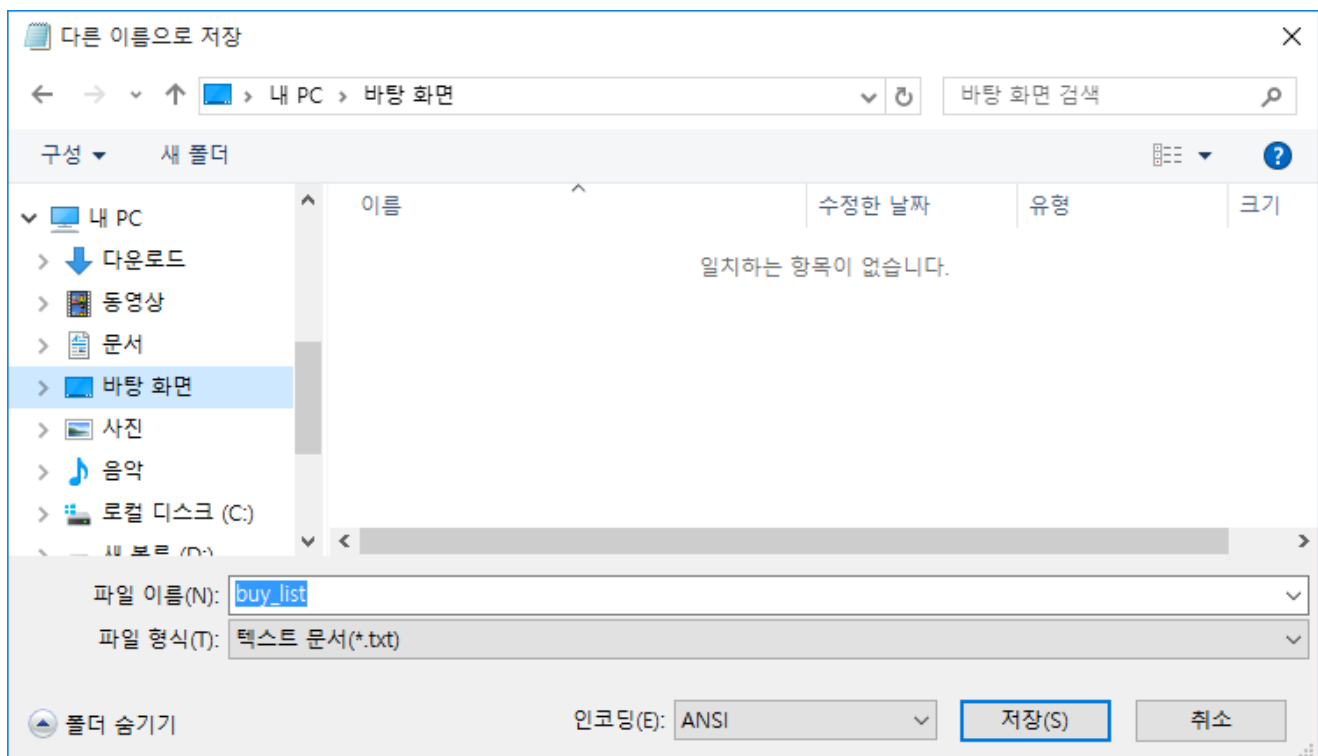


그림 7.3 텍스트 파일을 다른 이름으로 저장하기

윈도우 바탕화면을 보면 그림 7.4 와 같이 buy_list 라는 이름으로 텍스트 파일이 하나 생성된 것을 확인할 수 있습니다.



그림 7.4 바탕화면에 생성된 텍스트 파일

택배 기사가 여러분의 집으로 택배를 배달하는 경우를 상상해 봅시다. 택배 기사가 여러분에게 택배를 제대로 배달하려면 여러분의 집 주소를 정확히 알아야 합니다. 마찬가지로 파이썬으로 컴퓨터내의 파일을 읽기 위해서는 파일의 경로를 정확히 알아야 합니다.

앞서 저장한 buy_list.txt 파일 경로를 확인하기 위해서 해당 파일에 마우스 오른쪽 버튼을 클릭한 후 그림 7.5 와 같이 속성 메뉴를 선택합니다.

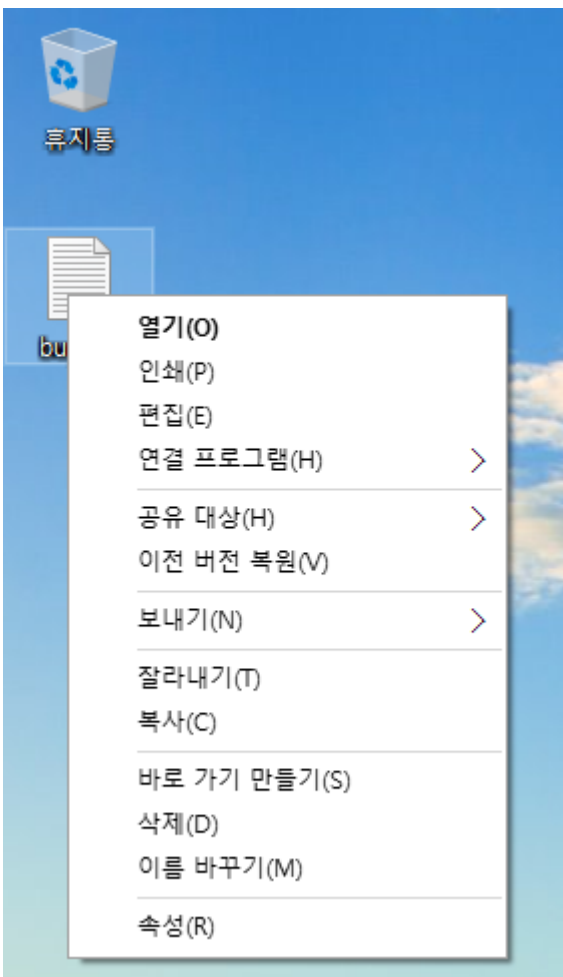


그림 7.5 파일 속성 메뉴

그림 7.6 과 같이 텍스트 파일에 대한 속성 창에서 위치 항목에 현재 파일 위치가 나타납니다. 그림 7.6 을 참조하면 현재 파일은 C:\Users\Jason\Desktop 이라는 디렉터리에 있으며 파일의 이름은 buy_list 이고 확장자는 .txt 입니다.

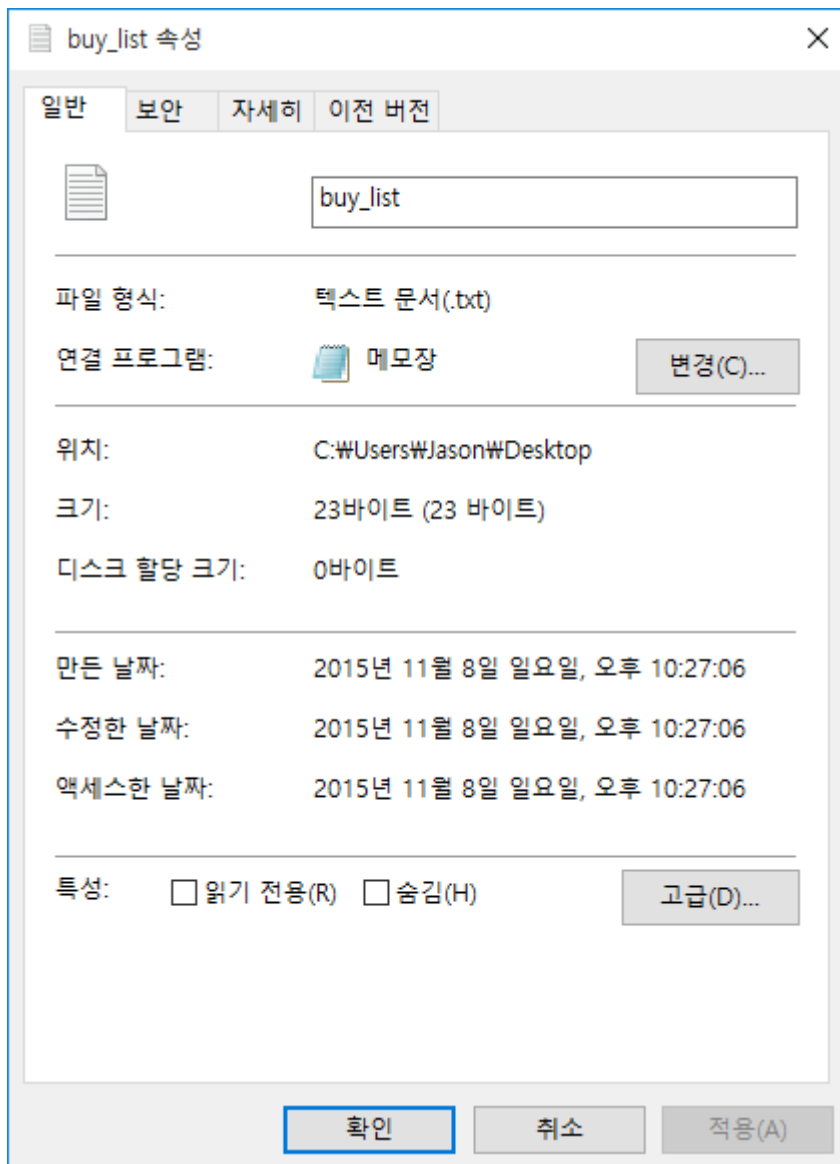


그림 7.6 파일의 위치, 이름 및 확장자 확인

여러분이 컴퓨터에 저장된 워드 파일을 읽을 때를 생각해봅시다. 보통 워드 프로그램을 실행한 후 열기 메뉴를 이용하여 워드 파일을 엽니다. 파일이 열린 후에는 해당 문서를 읽을 수 있습니다. 문서를 다 읽은 경우에는 역시 닫기 메뉴를 이용하여 해당 문서를 닫습니다. 파이썬에서 파일을 읽고 쓰는 것도 이와 같습니다. 파일을 읽거나 쓰려면 파일을 먼저 열어줘야 합니다.

앞서 파일의 위치를 확인했으니 이번에는 파이썬 IDLE 를 실행하여 워드에서 문서를 여는 것처럼 파이썬으로 텍스트 파일을 열어보겠습니다. 파이썬에서 파일을 열기 위해서는 open 이라는 내장 함수를 사용하면 됩니다.

다음은 바탕화면에 저장한 buy_list.txt 파일을 파이썬으로 여는 코드입니다. open 함수는 여러 개의 인자를 받을 수 있는데 첫 번째 인자로 파일의 경로를 입력받습니다. 파일의 경로를 입력할 때 주의할 점은 디렉터리의 구분자로 '₩' 하나를 사용하는 것이 아니라 '₩₩' 또는 '/'를 사용해야 한다는 것입니다.

open 함수의 두 번째 인자에는 파일을 열기위한 모드를 입력 받습니다. 예를 들어 파일을 읽기 모드로 열려면 'r'을 넣어주면 되고 텍스트 파일인 경우에는 't'를 넣어주면 됩니다. open 내장 함수는 파일에 대한 정보를 담고 있는 객체를 반환 해주는데 이를 바인딩하기 위해 f 라는 이름의 변수를 사용했습니다.

```
>>> f = open('c:\\Users\\Jason\\Desktop\\buy_list.txt', 'rt')
```

파이썬에서 open 함수의 반환값을 변수로 바인딩한 후에는 해당 변수를 통해 파일을 조작할 수 있습니다. 예를 들어 열린 파일에서 데이터를 읽기 위해서는 readlines() 메서드를 사용하면 됩니다.

```
>>> lines = f.readlines()
```

```
>>> lines
```

```
['NAVER\\n', '현대차\\n', '기아차\\n']
```

```
>>>
```

readlines() 메서드는 파일을 읽어서 각 라인을 리스트에 넣은 후 리스트를 리턴해줍니다. 위 코드를 참조하면 lines 라는 변수가 바인딩하는 리스트에는 그림 7.2 에서 입력한 문자열이 저장되어 있음을 확인할 수 있습니다. 앞서 배운 for 문을 활용하여 파일에서 읽어온 매수 종목을 화면에 출력해봅시다. 다음과 같이 for 문을 사용하면 lines 이라는 변수가 가리키고 있는 리스트 값을 읽어서 하나씩 화면에 출력할 수 있습니다. 참고로 앞서 배운 것처럼 for 문에서 사용한 line 이라는 변수는 lines 리스트에 있는 원소를 하나씩 바인딩하는 역할을 하는 것으로 이름이 꼭 line 일 필요는 없습니다.

```
>>> for line in lines:
```

```
    print(line)
```

```
NAVER
```

```
현대차
```

```
기아차
```

```
>>>
```

실행 결과를 살펴보면 종목명이 출력된 후 빈 줄이 하나씩 존재하는 것을 볼 수 있습니다. print()라는 출력 함수가 문자열을 출력할 때 자동으로 줄 바꿈을 하는데 기존에 lines 라는 리스트에 있는 문자열에도 줄 바꿈을 의미하는 값인 '\n'가 들어 있기 때문입니다. 이럴 때는 print() 함수에 추가 인자를 넣어서 자동으로 줄 바꿈 되지 않도록 하면 됩니다.

```
>>> for line in lines:
```

```
    print(line, end="")
```

NAVER

현대차

기아차

```
>>>
```

또는 print() 함수는 그대로 두고 lines 리스트에 있는 원소에서 '\n' 기호를 제거함으로써 중복으로 줄 바꿈 되지 않도록 코드를 작성할 수도 있습니다.

```
>>> for line in lines:
```

```
    nline = line.split('\n')[0]
```

```
    print(nline)
```

NAVER

현대차

기아차

```
>>>
```

어떤가요? 파이썬으로 파일을 읽는 것이 생각보다 어렵지 않지요?

2) 파일 쓰기

7.1 절에서는 하드 디스크에 이미 존재하는 텍스트 파일을 파이썬에서 열고 텍스트 파일 내의 데이터를 읽어 보았습니다. 이번에는 파이썬을 이용하여 하드디스크에 존재하지 않는 새로운 텍스트 파일을 생성하고 해당 파일에 데이터를 쓰는 작업을 해보겠습니다.

파이썬에서 파일 쓰기는 파일 읽기와 매우 비슷합니다. 워드에서 새로운 글을 작성하는 경우를 생각해보면 다음과 같습니다.

1) 빈 문서 파일을 만든다.

2) 문서에 내용을 기록한다.

3) 문서 파일을 닫는다.

파이썬에서 파일 쓰기도 기본적으로 위의 3 단계를 따릅니다. 단계별로 코드를 작성하면서 파이썬으로 파일 쓰기를 익혀보겠습니다.

먼저 새로운 파일을 엽니다. 이때 중요한 점은 텍스트 파일이 생성될 위치 (경로)와 파일 이름을 지정해줘야 한다는 것입니다. 앞서 확인했던 윈도우의 바탕화면 경로에 `sell_list.txt` 라는 빈 문서를 만들어 보겠습니다.

```
>>> f = open('c:\\Users\\Jason\\Desktop\\sell_list.txt', 'wt')
```

위 코드를 보면 읽기 모드로 파일을 열었던 것과 유사합니다. 차이점은 파일 이름이 `sell_list.txt` 로 변경됐고, `open()` 함수의 두 번째 인자로 `'wt'`를 사용했습니다.



그림 7.7 생성된 텍스트 파일 확인

파이썬에서 파일을 쓰기 모드로 열면 그림 7.7 과 같이 하드디스크에 실제 파일이 생성되는 것을 확인할 수 있습니다. 물론 아직 파일 안에는 아무런 내용도 없는 상태입니다.

열려있는 파일에 실제로 데이터를 쓰기 위해서는 write 메서드를 사용해야 합니다. 다음과 같이 write 메서드를 이용하면 파일에 데이터를 쓸 수 있습니다.

```
>>> f.write('삼성전자\n')
5
>>> f.write('SK 하이닉스\n')
7
>>> f.close()
>>>
```

위 코드를 살펴보면 '삼성전자'와 'SK 하이닉스'라는 문자열을 파일에 쓸 때 종목 구분을 위해 줄 바꿈 기호 ('\n')를 사용했습니다. 파일에 데이터를 쓴 후에는 close() 함수를 사용해서 파일을 닫아 주면 됩니다.

파이썬을 이용한 파일 쓰기가 정상적으로 실행했는지 확인하기 위하여 바탕화면에 새로 생성된 sel_list.txt 파일을 열어보기 바랍니다. sel_list.txt 파일 내용에 '삼성전자'와 'SK 하이닉스'라는 종목명이 그림 7.8 처럼 나타난다면 정상적으로 실행된 것입니다.

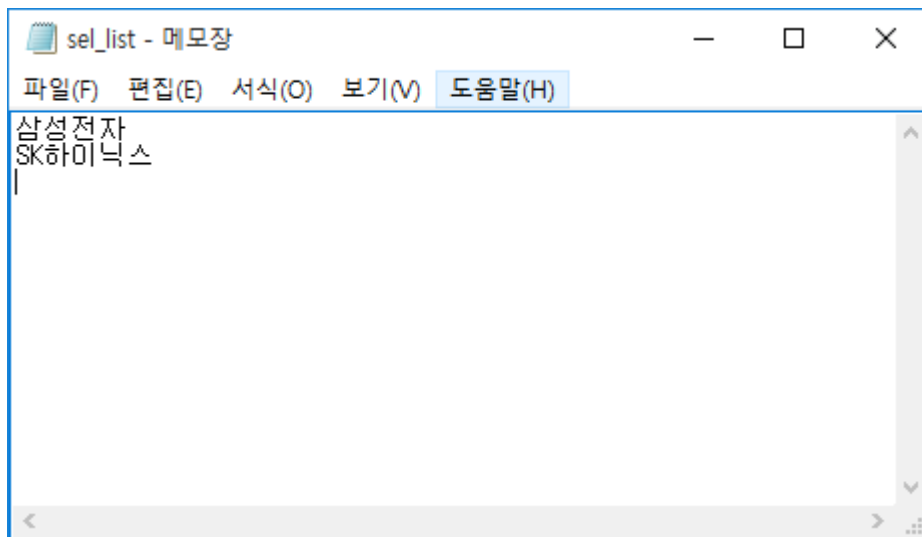


그림 7.8 self_list 파일 내용 확인

1) 연습문제

문제 7-1

1 부터 10 까지의 숫자를 각 라인 단위로 파일에 출력하는 프로그램을 작성하세요.

- 생성되는 파일의 이름은 number.txt 이다.

문제 7-2

사용자로부터 경로를 입력 받은 후 해당 경로에 있는 디렉터리와 파일 목록을 flist.txt 라는 파일로 출력하는 함수를 작성하세요.