



▶ 학습내용

01. 파일이 왜 필요한지 이해합니다.
02. 파일을 사용하기 위한 준비 방법을 학습합니다.
03. 파일의 내용을 읽고, 쓰는 방법을 학습합니다.
04. CSV 파일을 사용하는 방법을 학습합니다.

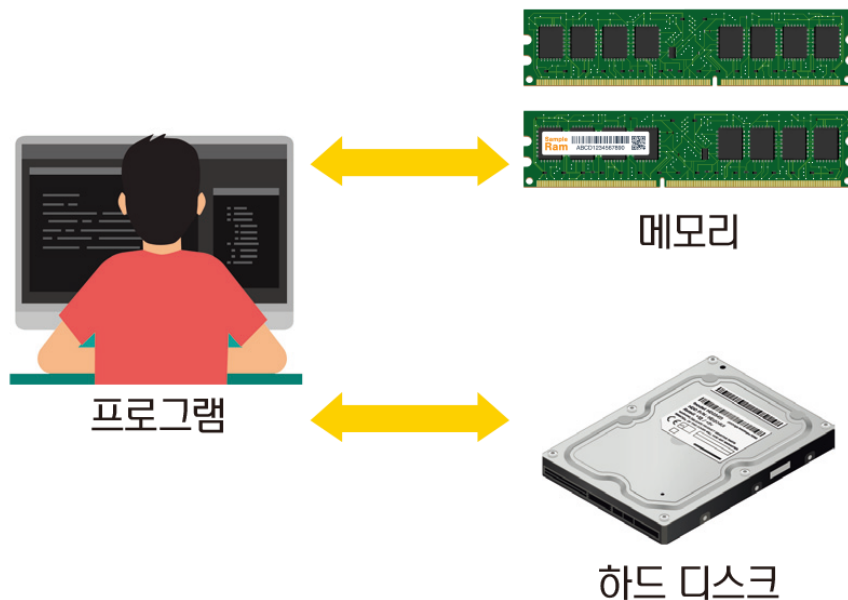
▶ LAB

01. 파일 복사하기
02. 연설문 데이터 분석
03. 평균 강수량 통계
04. 행맨

01. 파일은 왜 필요할까?



- 데이터를 오래 보관하고자 한다면 우리는 하드디스크와 같은 보조기억장치에 파일 형태로 저장해야 함.



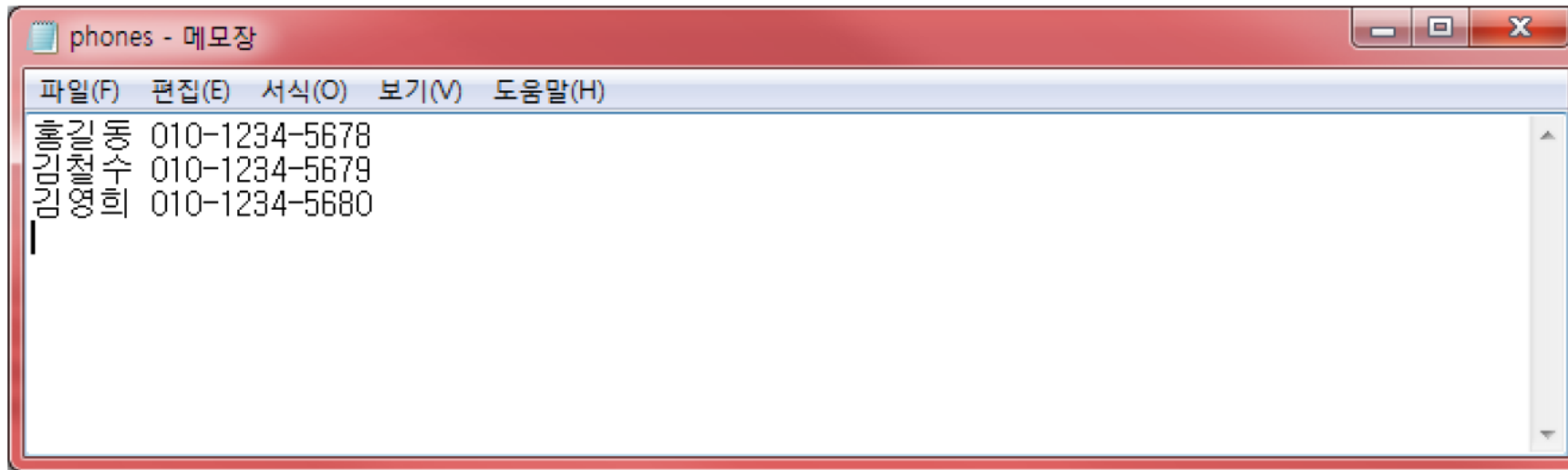
변수는 모두 메모리에 만들어지고 이것들은 전원이 꺼지면 모두 사라집니다.

하드 디스크에 파일 형태로 저장하면 전원이 꺼지더라도 데이터가 보존됩니다.

01. 파일은 왜 필요할까?



실습용 파일 준비하기



02. 파일 열기와 닫기



파일을 연다.

파일에서 데이터를
읽거나 쓴다.

파일을 닫는다.

02. 파일 열기와 닫기



파일을 열고 닫는 구조는 다음과 같습니다.

파일 사용하기

```
infile = open("input.txt", "r")  
...  
infile.close()
```

파일 객체

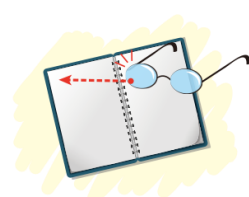
파일을 여는 모드(mode)

파일의 이름(name)

02. 파일 열기와 닫기



파일 모드	모드 이름	설명
"r"	읽기 모드 (read mode)	파일의 처음부터 읽습니다.
"w"	쓰기 모드 (write mode)	파일의 처음부터 씁니다. 파일이 없으면 생성됩니다. 만약 파일이 존재하면 기존의 내용은 지워집니다.
"a"	추가 모드 (append mode)	파일의 끝에 이어서 내용을 씁니다. 파일이 없으면 생성됩니다.
"r+"	읽기와 쓰기 모드	파일에 읽고 쓸 수 있는 모드입니다. 모드를 변경하려면 seek()가 호출되어야 합니다.



"r"

파일의 처음부터 읽는다.



"w"

파일의 처음부터 쓴다.
만약 파일이 존재하면
기존의내용이 지워진다



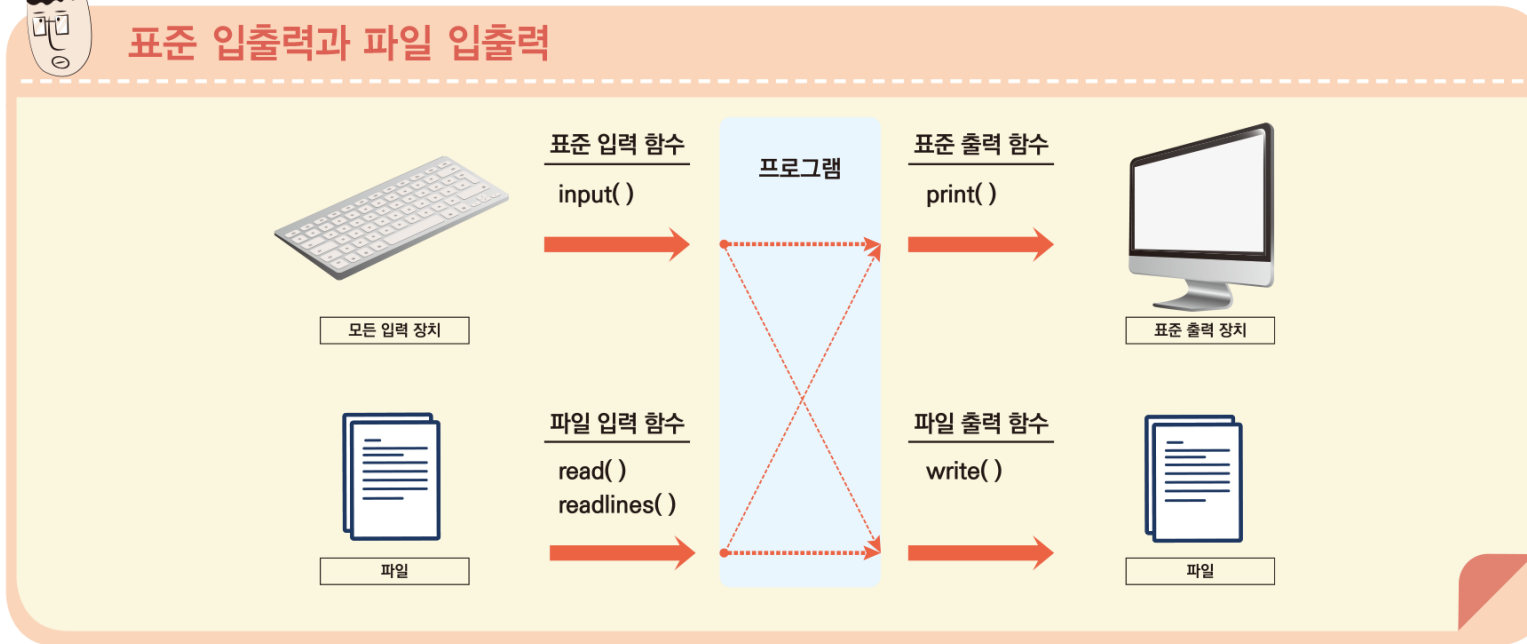
"a"

파일의 끝에 쓴다.
파일이 없으면 생성된다.

02. 파일 열기와 닫기



표준 입출력과 파일 입출력



03. 파일에서 전체 데이터 읽기



한 번에 파일의 모든 줄을 읽는 함수는 파일 객체의 read() 입니다.

코드

```
infile = open("d:\phones.txt", "r")  
lines = infile.read()  
print(lines)  
infile.close()
```

실행 결과

```
홍길동 010-1234-5678  
김철수 010-1234-5679  
김영희 010-1234-5680
```


03. 파일에서 전체 데이터 읽기



한 번에 파일의 모든 줄을 읽는 함수는 파일 객체의 `readlines()`도 있습니다.

코드

```
infile = open("d:\wwphones.txt", "r")  
lines = infile.readlines()  
print(lines)  
infile.close()
```

실행 결과

```
['홍길동 010-1234-5678\n', '김철수 010-1234-5679\n', '김영희 010-1234-5680\n']
```

04. 파일에서 한 줄씩 읽어오기



파일에서 한 번에 하나의 줄만 읽어오는 함수는 파일 객체의 `readline()`이 있습니다.

코드

```
infile = open("d:WWphones.txt", "r")  
  
line1 = infile.readline()  
print(line1)  
  
line2 = infile.readline()  
print(line2)  
  
infile.close()
```

실행 결과

```
홍길동 010-1234-5678  
  
김철수 010-1234-5679
```

04. 파일에서 한 줄씩 읽어오기



하지만!! 우리는 파일에 몇 줄의 내용이 들어있는지 미리 알 수가 없습니다.

코드 1

```
infile = open("d:WWphones.txt", "r")
line = infile.readline().rstrip()
while line != "":
    print(line)
    line = infile.readline().rstrip()

infile.close()
```

코드 2

```
infile = open("d:WWphones.txt", "r")

for line in infile:
    line = line.rstrip()
    print(line)

infile.close()
```

05. 파일에 데이터 쓰기



파일에 데이터를 쓰려면 `write()` 함수를 이용합니다.

코드

```
outfile = open("d:WWphones1.txt", "w")

outfile.write("전우치 010-1234-5678\n")
outfile.write("홍길동 010-1234-5679\n")
outfile.write("김서방 010-1234-5680\n")

outfile.close()
```

실행 결과 : 파일 phones1.txt

```
전우치 010-1234-5678
홍길동 010-1234-5679
김서방 010-1234-5680
```

05. 파일에 데이터 쓰기



파일의 기존 내용에 새로운 데이터를 이어서 쓸 수 있습니다.

코드

```
outfile = open("d:\WWphones1.txt", "a")

outfile.write("강감찬 010-1234-5681\n")
outfile.write("김유신 010-1234-5682\n")
outfile.write("정약용 010-1234-5683\n")

outfile.close()
```

실행 결과 : 파일 phones1.txt

```
전우치 010-1234-5678
홍길동 010-1234-5679
김서방 010-1234-5680
강감찬 010-1234-5681
김유신 010-1234-5682
정약용 010-1234-5683
```

06. 파일에 단어 단위로 읽어오기



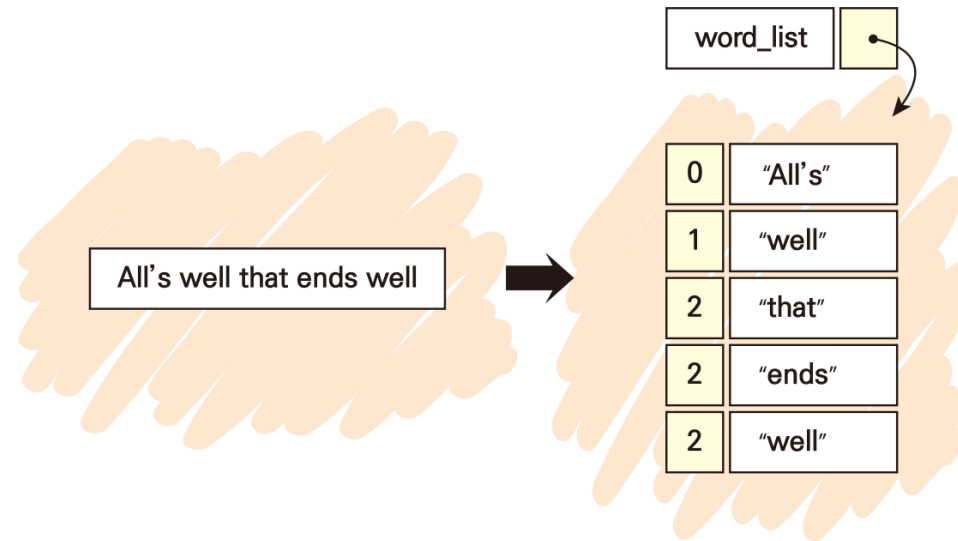
split() 함수를 이용하여 문자열을 분리해 보겠습니다.

코드

```
>>> a = "All's well that ends well"  
>>> a.split()
```

실행 결과

```
['All's', 'well', 'that', 'ends', 'well']
```



07. CSV 파일을 사용해 보기

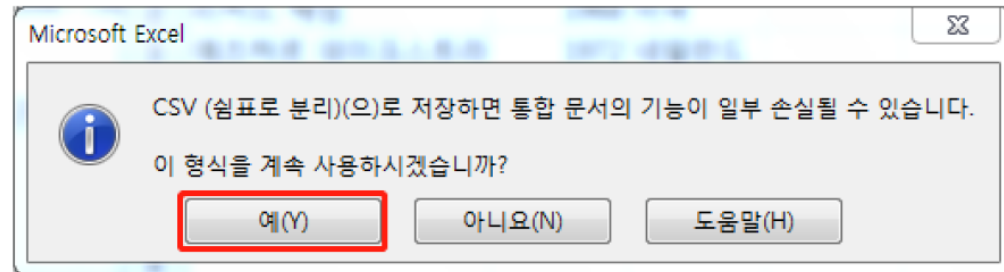


- CSV 파일을 처리하기 위해서는 CSV 모듈을 import 해야함



CSV 파일을 엑셀을 이용하여 만들어 보겠습니다. 아래와 같이 엑셀에서 자료를 작성한 후에 파일을 “input.csv”로 저장합니다.

	A	B	C	D
1	팀버너스리	2016	영국	
2	리처드 해밍	1968	미국	
3	에츠허르 데이크스트라	1972	네덜란드	
4	더글러스 엥겔바트	1997	미국	
5				



07. CSV 파일을 사용해 보기



split() 함수를 이용하여 문자열을 분리해 보겠습니다.

코드

```
import csv
f = open('d:\WWinput.csv','r')
data = csv.reader(f)

for line in data:
    print(line)

f.close()
```

실행 결과

```
['팀버너스리', '2016', '영국']
['리처드 해밍', '1968', '미국']
['에츠허르 데이크스트라', '1972', '네덜란드']
['더글러스 엥겔바트', '1997', '미국']
```


Lab. 파일 복사하기



실행 결과

입력 파일 이름: d:\phones.txt

출력 파일 이름: d:\temp.txt



Lab. 연설문 데이터 분석



- 연설문의 총 단어 수와 단어들의 빈도를 분석하는 프로그램을 작성하여 데이터 분석을 해보세요.

input - 메모장

파일(F) 편집(E) 서식(O) 보기(V) 도움말(H)

Four score and seven years ago, our fathers brought forth upon this continent a new nation, conceived in liberty and dedicated to the proposition that all men are created equal.
Now we are engaged in a great civil war, testing whether that nation or any nation so conceived and so dedicated can long endure. We are met on a great battlefield of that war.
We have come to dedicate a portion of that field as a final resting place for those who here gave their lives that this nation might



Lab. 연설문 데이터 분석



input - 메모장

파일(F) 편집(E) 서식(O) 보기(V) 도움말(H)

Four score and seven years ago, our fathers brought forth upon this continent a new nation, conceived in liberty and dedicated to the proposition that all men are created equal.
Now we are engaged in a great civil war, testing whether that nation or any nation so conceived and so dedicated can long endure. We are met on a great battlefield of that war.
We have come to dedicate a portion of that field as a final resting place for those who here gave their lives that this nation might

output - Window...

파일(F) 편집(E) 서식(O) 보기(V) 도움말(H)

a 7
above 1
add 1
advanced 1
ago 1
all 1
altogether 1

Lab. 연설문 데이터 분석



생각 1 : 프로그램의 순서를 생각해 봅니다.

게티즈버그 연설 원본 파일 준비



입력 및 출력 파일 열기



파일에서 데이터를 추출하여 단어별로 분리하고 정리(영어 소문자 처리, 콤마 · 쉼표 삭제 처리)

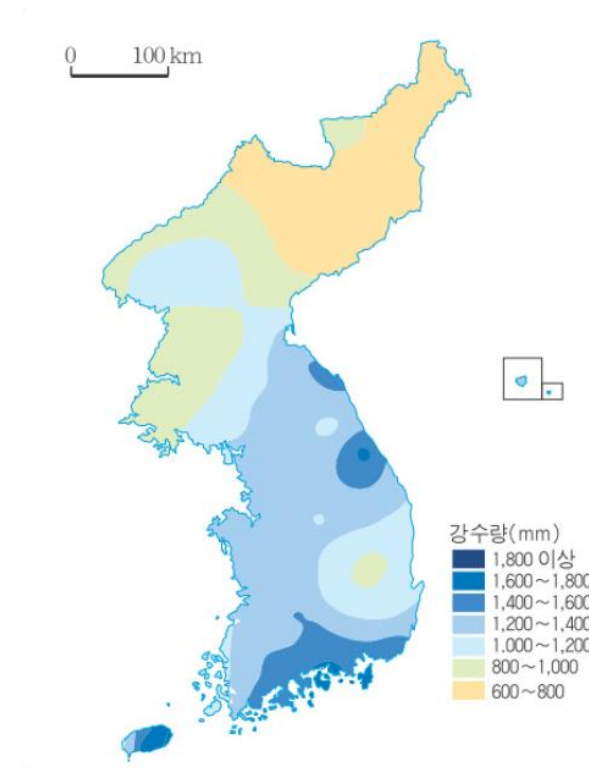
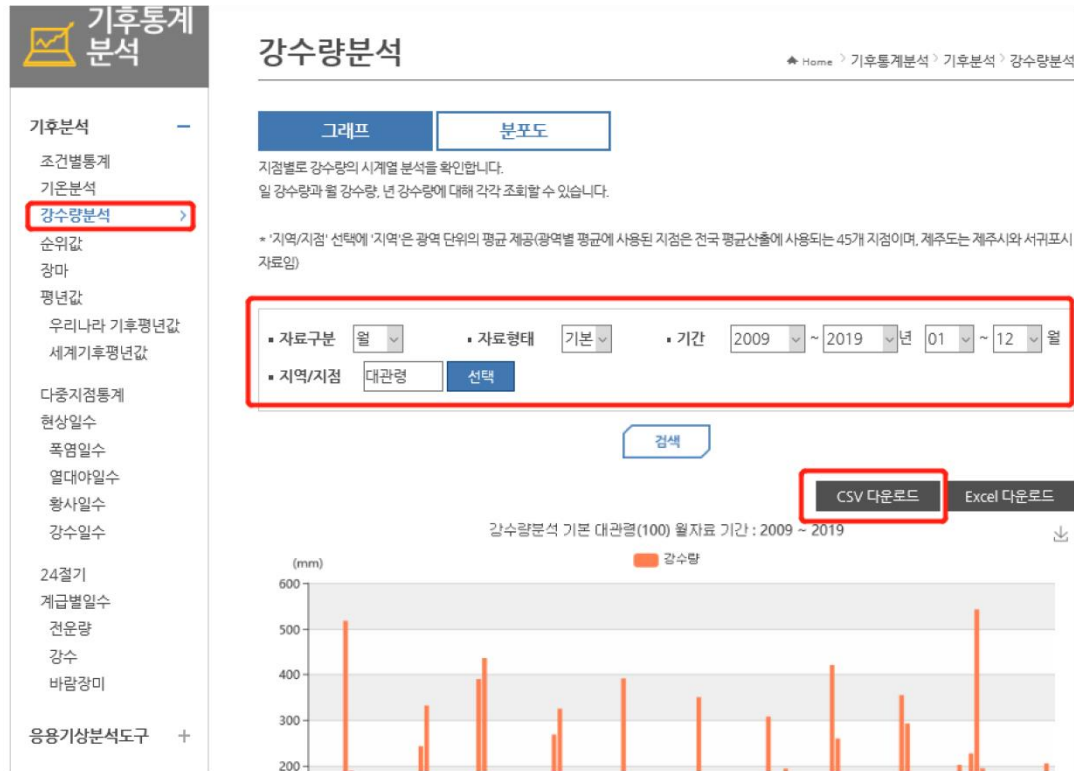


파일에서 데이터를 추출하여 단어별 빈도를 딕셔너리에 저장 & 총 단어 수 계산



딕셔너리 결과 파일로 출력 및 파일 닫기

Lab. 평균 강수량 통계



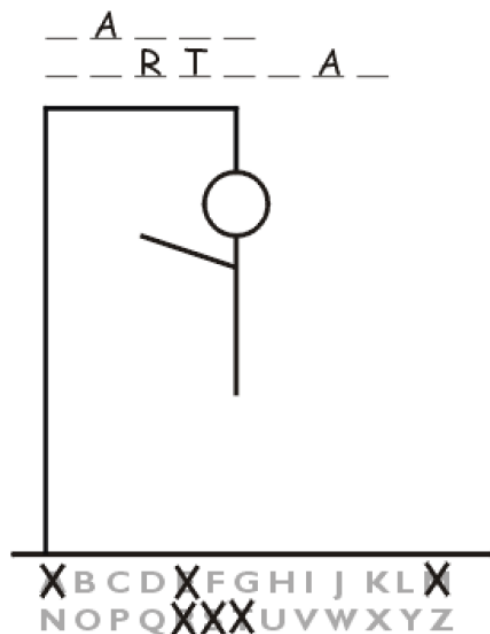
Lab. 평균 강수량 통계



실행 결과

강원도 2009년 01월부터 2019년 09월까지의 총 강수량: 13666.799999999996

강원도 2009년 01월부터 2019년 09월까지의 평균 강수량: 105.94418604651159



실행 결과 1

단어를 추측하세요: a
['_', 'a', '_', '_', '_', 'a', '_']
단어를 추측하세요: s
['_', 'a', '_', '_', '_', 'a', '_']
단어를 추측하세요: h
['_h', 'a', '_', '_', '_', 'a', '_']
단어를 추측하세요: n
['_h', 'a', 'n', '_', '_', 'a', 'n']
단어를 추측하세요: g
['_h', 'a', 'n', 'g', '_', 'a', 'n']
단어를 추측하세요: m
['_h', 'a', 'n', 'g', 'm', 'a', 'n']
성공입니다.

실행 결과 2

단어를 추측하세요: a
['_', '_', '_', '_', '_', '_']
단어를 추측하세요: b
['_', '_', '_', '_', '_', '_']
...(중략)...
단어를 추측하세요: h
['_', '_', '_', 'h', '_', '_']
단어를 추측하세요: i
['_', '_', '_', 'h', '_', '_']
단어를 추측하세요: j
['_', '_', '_', 'h', '_', '_']
실패하였습니다.



생각 1 : 프로그램의 순서를 생각해 봅니다.

문제 파일 준비하기



문제 파일 열어 무작위로 단어를 1개 선정하기



모든 단어를 '_'로 문제 출제 후 사용자로부터 한 글자 입력받기



입력받은 글자가 단어에 있으면 해당 위치에서 글자가 보이고
나머지는 '_'로 처리함/ 시도횟수를 하나 줄임



시도횟수가 0보다 작아지거나 글자를 맞추면 프로그램을 종료