

파이썬 설치, 주피터 노트북 사용법 및 경로 설정

이 문서에서는 파이썬 설치와 파이썬을 실행할 환경인 주피터 노트북의 사용법, 그리고 주피터 노트북에서 파일을 불러올 때 경로를 지정하는 방법에 대하여 알아볼 것 입니다.

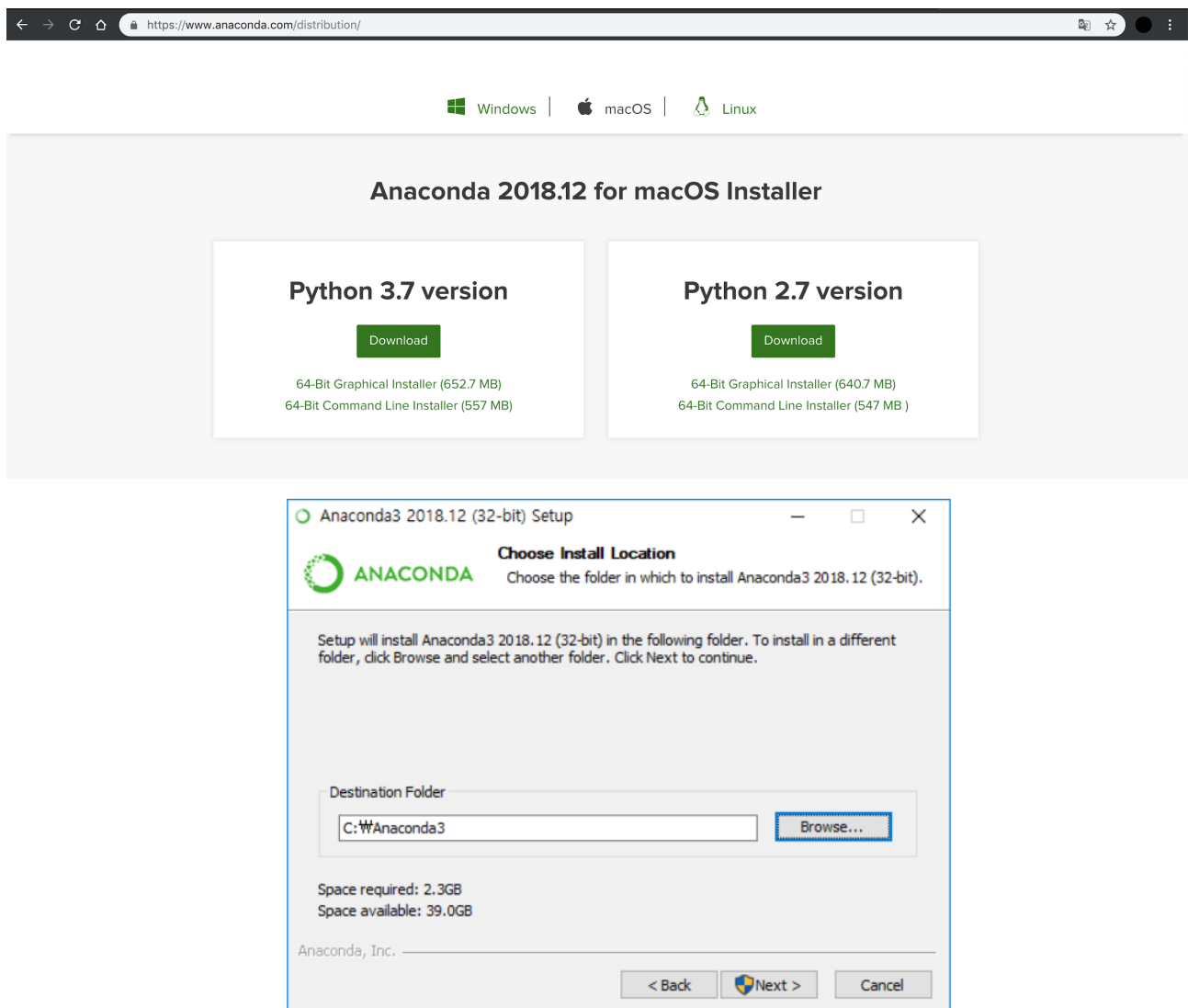
파이썬 설치

데이터 분석을 목적으로 파이썬을 설치하는 경우에는 아나콘다(Anaconda)를 통한 설치를 권장합니다. 아나콘다를 설치하면 파이썬이 자동으로 설치될 뿐만 아니라 데이터 분석에 관련된 패키지들을 매우 쉽게 설치하고 관리할 수 있기 때문입니다.

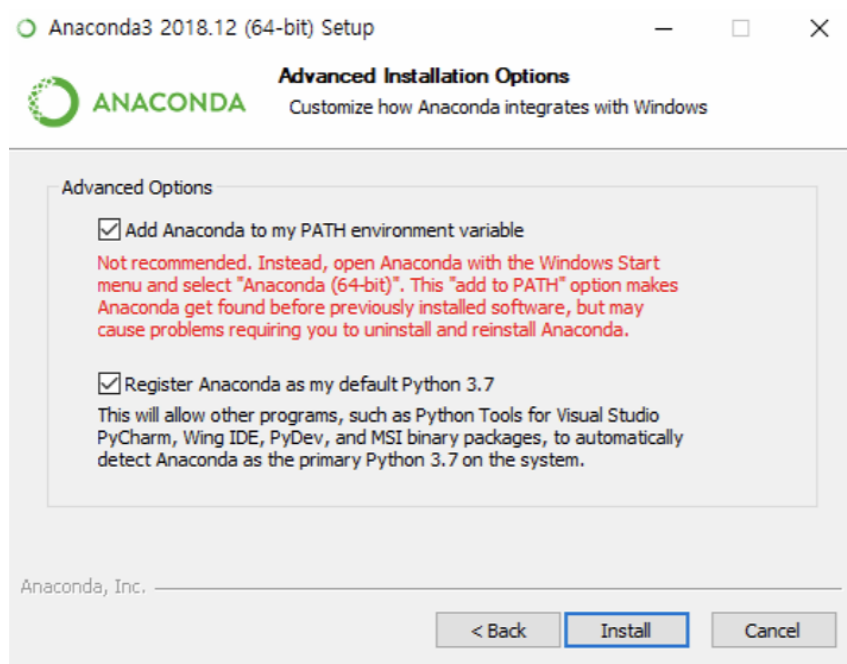
아나콘다를 통한 설치

구글에 anaconda를 검색 또는 <https://www.anaconda.com/distribution/> 주소를 들어가면 다음과 같은 화면을 볼 수 있습니다.

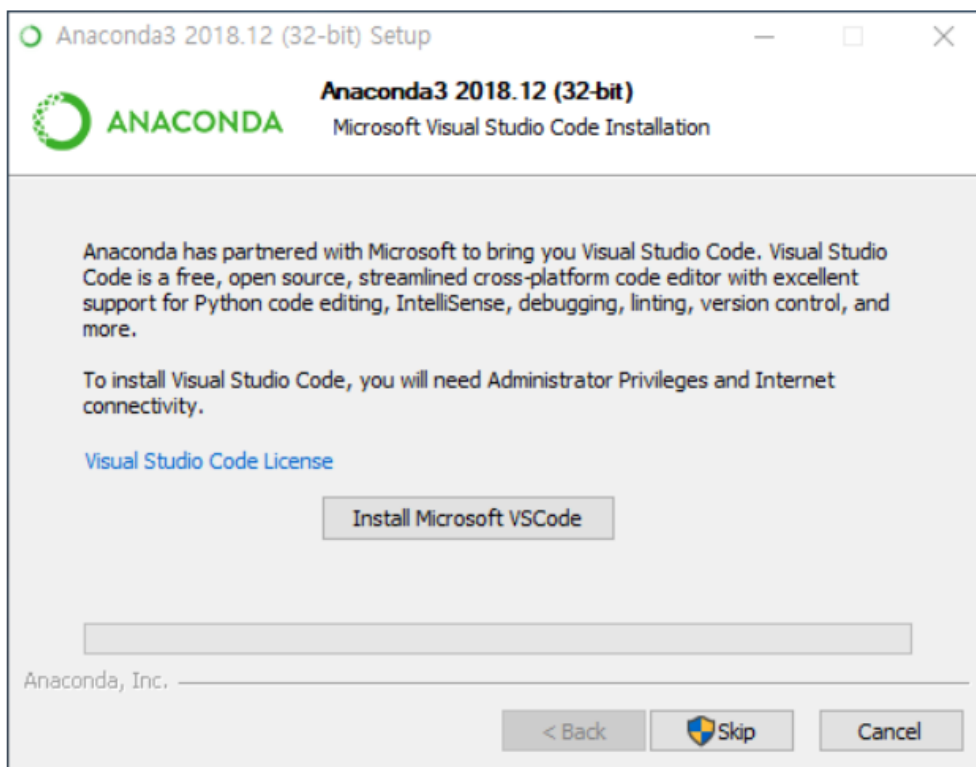
자신이 사용하는 운영체제에 맞게 Python 3.7 버전을 다운로드하시면 됩니다.



기본으로 설정되는 경로에 설치하는 것이 향후 에러를 방지하는데 좋습니다. 주의할 점은 설치 경로에 한글이 있는 경우 에러가 발생하기 때문에 반드시 설치 경로에 한글이 들어가지 않도록 해주세요.



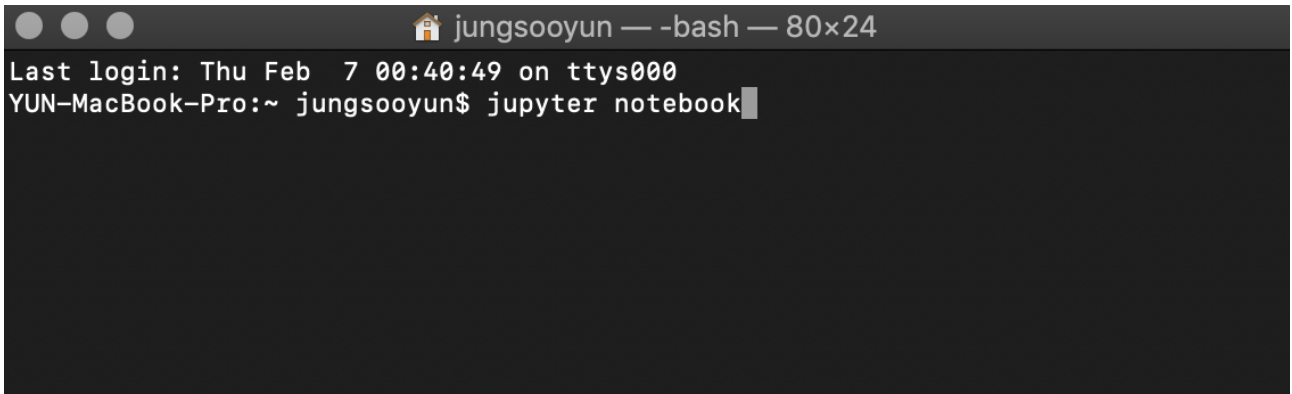
윈도우의 경우 위의 체크박스를 선택해주세요. 선택하지 않고 설치했더라도 큰 문제가 되지 않으니 계속 설치를 진행해주시면 됩니다.



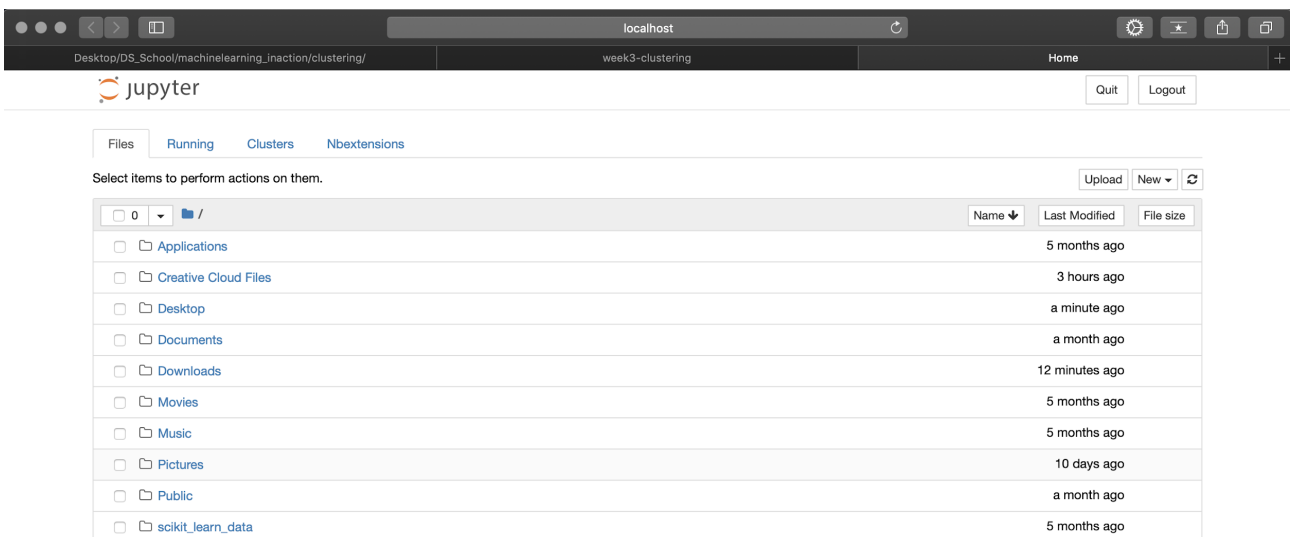
이 부분은 Skip을 클릭해주시면 됩니다.

맥OS에서 주피터 노트북 실행하기

맥 OS의 경우 Command + Space 버튼을 누르고 터미널을 검색한 뒤, 이를 실행 하시면 됩니다. 그 다음, jupyter notebook을 입력하고 실행하시면 주피터 노트북을 실행할 수 있습니다.

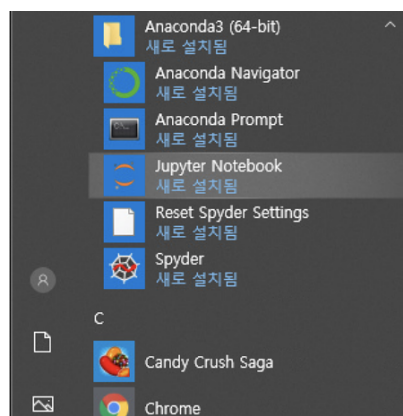


위 명령어를 Enter 를 눌러 실행하면 아래와 같은 창이 뜨게 됩니다.



윈도우(Windows)에서 주피터 노트북 실행하기

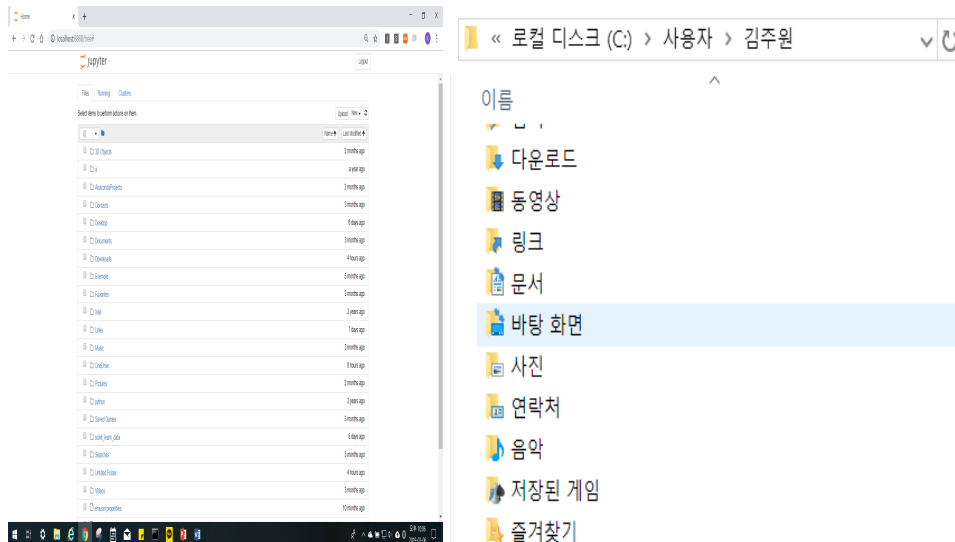
Window 버튼을 누른 뒤, anaconda3 폴더를 찾은 다음 jupyter notebook을 클릭합니다. 또는 윈도우 버튼을 누른 뒤 jupyter notebook을 검색해도 됩니다.



jupyter notebook을 클릭하여 실행하면 검은 커널 창이 뜨고 약간의 시간이 지나면 인터넷 브라우저에 주피터 노트북 창이 켜집니다. **검은 커널 창**을 끄시면 주피터 노트북이 더 이상 작동하지 않기 때문에 주피터 노트북을 사용하시는 동안에는 끄지 말아주세요!

주피터 노트북에서 파이썬 실행 파일 생성 및 실행하기

주피터 노트북은 파일 탐색기의 다른 버전이라고 생각하시면 편합니다. 주피터 노트북에서 폴더나 파일을 삭제, 이동, 또는 편집 등을 하는 것은 기존에 컴퓨터에서 하던 것처럼 동일하게 적용됩니다. 주피터 노트북을 처음 실행했을 때 나오는 폴더는 맥의 경우 Users/(사용자명) 폴더, 윈도우의 경우 C:/Users/(사용자명) 폴더입니다.



위의 화면은 윈도우의 사용자 폴더입니다.

여기서 주피터 노트북 우측 상단의 New 버튼을 클릭해 python 3 실행 파일을 만들 수 있습니다. 그러면 사용자 폴더 내에 Untitled.ipynb 란 파일이 생성되면서 파일이 실행됩니다. 주피터 노트북 파일의 확장자는 ipynb입니다. 원하시는 폴더에 파일을 만들고 싶으시면 해당 폴더로 들어가신 다음 New 버튼을 클릭해 Python 3 파일을 클릭하면 됩니다. 파일을 이동하려면 탐색기 내에서 평소와 같이 이동하면 됩니다.

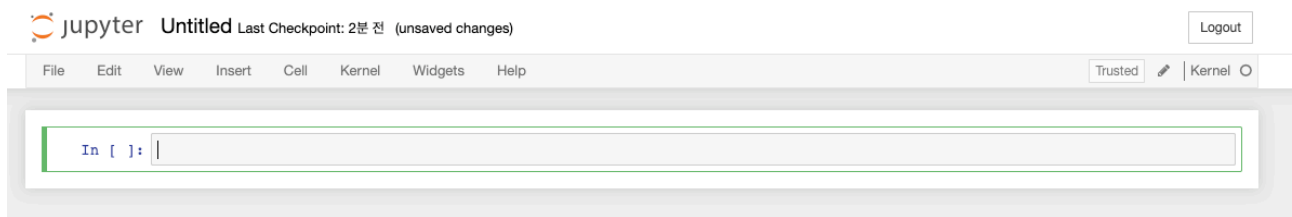


주피터 노트북 기초

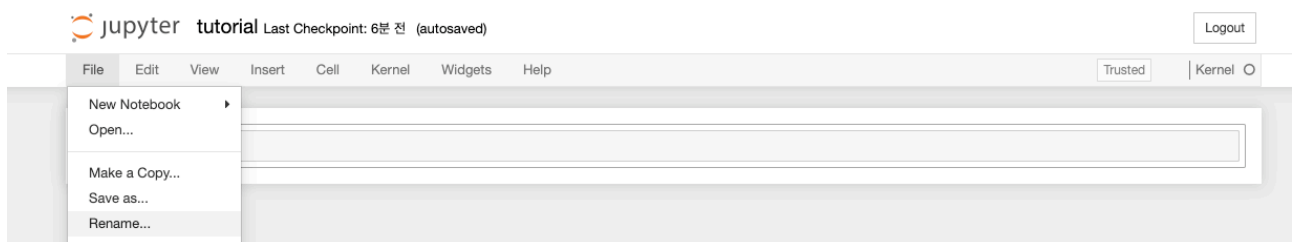
바탕 화면(Desktop)에 DS_School이란 폴더를 만들고 이 폴더 안에 Python 3파일인 Untitled.ipynb 을 만들어 놓았습니다. 같은 폴더 내에 tutorial.csv 파일도 만들어 놓았습니다.

/ Desktop / DS_School			Name	Last Modified	File size
..				몇 초 전	
tutorial.ipynb				13분 전	555 B
tutorial_data.csv				몇 초 전	42 B

Untitled.ipynb 를 클릭하면 파일이 실행됩니다. 여기서 주의하셔야 할 것은 주피터 노트북 파일(.ipynb)은 주피터 노트북에서만 열리기 때문에 탐색기에서 더블 클릭 등을 하셔도 실행되지 않습니다.



우선 주피터 파일의 이름을 tutorial로 바꾸도록 하겠습니다. 왼쪽 상단의 File을 클릭하고 Rename..을 클릭하시거나 탐색기에서 평소와 같이 직접 파일 이름을 바꿔주셔도 됩니다.



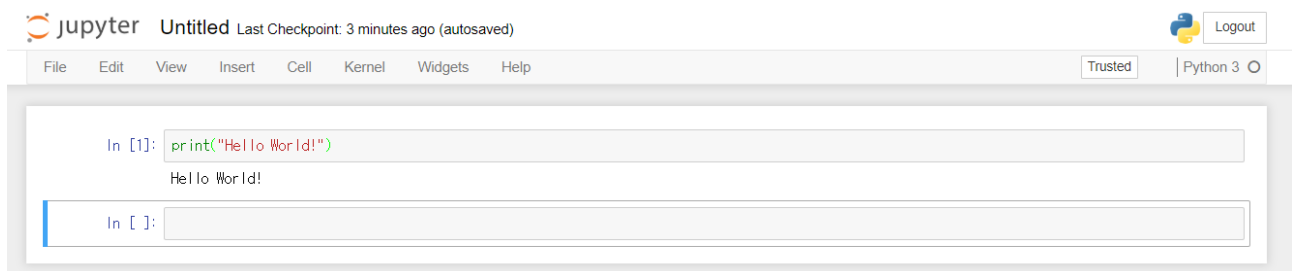
파일의 이름이 tutorial로 바뀐 것을 확인할 수 있습니다.

주피터 노트북 화면상의 박스 하나를 셀(cell)이라 부릅니다. 셀에서는 코드(code) 또는 마크다운(markdown) 등을 작성할 수 있으며 셀마다 실행(run)이 가능합니다.

셀 왼쪽 끝의 파란색 표시는 **선택 상태**를 나타내며 이는 Enter를 누르거나 마우스로 셀에서 코드를 작성하는 부분을 클릭하면 초록색으로 바뀝니다. 초록색 상태는 코드 **편집 상태**를 의미하며 코드를 작성할 수 있게 됩니다. 다시 파란색 표시의 선택 상태로 돌아오려면 셀 바깥 부분을 클릭하거나 키보드의 Esc를 누르면 됩니다.

천천히 여러 버튼들을 클릭하며 기능들을 탐색해보시는 것도 좋습니다 :)

이제 코드를 실행해 보겠습니다. 먼저 셀을 편집상태로 바꾼 뒤 print("Hello World!")를 입력해주세요. 그리고 나서 Shift + Enter를 누르면 해당 셀이 실행되며 다음 셀로 넘어가게 됩니다.



이번에는 code를 실행하는 것이 아닌 markdown을 작성하도록 하겠습니다. markdown은 쉽게 메모장이라고 생각하시면 됩니다. 우선 셀을 파란색(선택 상태)로 해주시고 키보드에서 m을 눌러주세요.

```
In [1]: print("Hello World!")  
Hello World!
```

위와 같이 셀 왼쪽의 In [] : 이 사라진 것을 확인할 수 있습니다. 이제 아무 글을 입력해주시고 Shift + Enter 를 눌러 실행해주세요. 저는 Nothing is impossible ! 을 입력하였습니다.

```
In [1]: print("Hello World!")  
Hello World!  
  
Nothing is impossible !
```

위와 같이 깔끔하게 글이 써진 것을 확인하실 수 있을 겁니다. 이번에는 글의 맨 앞에 #을 넣고 다시 실행시켜 주세요. #을 쓰신 다음 꼭 한칸을 띄우고 실행해야 합니다.

```
In [1]: print("Hello World!")  
Hello World!  
  
Nothing is impossible !
```

#은 위와 같이 글씨를 크고 굵게 만들어 줍니다. #의 개수가 많아질수록 글씨는 작아집니다.

여기 까지 jupyter notebook의 기본 기능들에 대하여 간단하게 알아 보았습니다. 마지막으로 많이 사용하는 단축 키들을 정리해 보겠습니다..

- Esc : 셀 편집 상태를 선택 상태로 전환
- Enter : 셀 선택 상태를 편집 상태로 전환
- 셀 선택 상태에서 a : 선택된 셀을 위에 새로운 셀 추가
- 셀 선택 상태에서 b : 선택된 셀을 아래에 새로운 셀 추가
- 셀 선택 상태에서 m : 선택된 셀을 markdown 상태로 전환
- 셀 선택 상태에서 y : 선택된 셀을 code 편집 상태로 전환
- 셀 선택 상태에서 x : 셀 잘라내기
- 셀 선택 상태에서 c : 셀 복사하기
- 셀 선택 상태에서 v : 셀 붙여넣기
- 셀 선택 상태에서 dd : 셀 제거하기
- Ctrl + Enter : 셀 실행
- Shift + Enter : 셀 실행 후 아래 셀 선택
- Alt + Enter : 셀 실행 후 아래에 새로운 셀 추가

csv 파일 불러오기 및 경로 설정하기

실습 중에는 csv 포맷으로 되어있는 데이터들을 자주 불러오게 될 예정입니다. 이 때, 파이썬의 패키지인 판다스(pandas)의 read_csv()라는 기능을 사용하여 불러올 수 있습니다. 위에서 만든 tutorial.ipynbb 파일과 같이 DS_School 폴더 내에 있는 tutorial.csv 파일을 불러오도록 하겠습니다.

tutorial.csv 파일은 다음과 같습니다.

	A	B	C
1	Name	Value	
2	DS_School	1	
3	Korea	2	
4	python	3	
5			

```
In [1]: print("Hello World!")
Hello World!

Nothing is imposible !

In [2]: import pandas as pd

In [3]: pd.read_csv("tutorial_data.csv")

Out[3]:
```

	Name	Value
0	DS_School	1
1	Korea	2
2	python	3

```
In [ ]:
```

위와 같이 tutorial.csv 파일이 불러와진 것을 확인할 수 있습니다.

하지만, tutorial.csv 파일과 주피터 노트북 파일(.ipynb)이 파일이 같은 폴더 내에 존재하지 않는다면 실행시 다음과 같은 에러가 발생합니다.

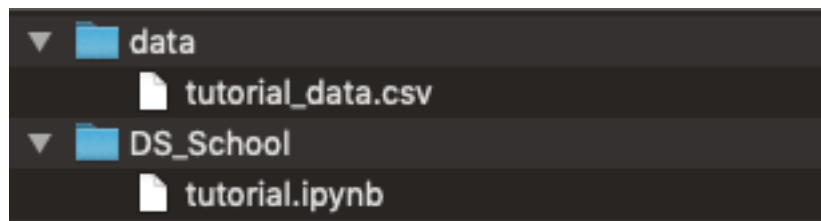
```
In [4]: pd.read_csv("tutorial.csv")

-----
FileNotFoundError                                Traceback (most recent call last)
<ipython-input-4-bd4ed20d0894> in <module>
----> 1 pd.read_csv("tutorial.csv")
```

이런 FileNotFoundError 라는 에러는 read_csv()가 파일을 찾지 못하여 발생합니다. 따라서 파일을 불러올 때는 불러올 파일이 어디에 위치해 있는지 read_csv()에 정확하게 입력해줘야 합니다. 이 때 파일의 위치를 **경로**라고 합니다. 파일의 경로를 표현하는 방식은 2가지로 **상대경로**와 **절대경로**가 있습니다. 하지만 기본적으로 실행하고 있는 주피터 노트북 파일과 불러올 csv 파일이 같은 폴더 내에 있다면 위의 예시와 같이 "파일명.csv" 만 입력해도 됩니다.

상대경로 입력법

상대경로는 실행하고 있는 주피터 노트북 파일을 기준으로 불러올 csv 파일의 위치를 표현하는 방식입니다. 즉 실행하고 있는 주피터 노트북 파일을 기준으로 불러올 csv 파일을 가는 과정(경로)을 나타냅니다.



위와 같이 바탕화면(Desktop)의 data 폴더 내에 존재하는 tutorial_data.csv 파일을 불러오도록 하겠습니다. 실행하고 있는 tutorial.ipynb 파일은 DS_School 폴더 안에 있습니다.

```
In [4]: pd.read_csv("../data/tutorial_data.csv")
```

Out[4]:

	Name	Value
0	DS_School	1
1	Korea	2
2	python	3

위와 같이 read_csv() 안에 경로를 입력하면 됩니다. 여기서 ../은 상위 폴더(여기서는 Desktop이 됩니다.)로 나가는 것을 의미합니다. 그다음 data/ 는 data 폴더로 들어가는 것을 의미합니다. 따라서 상대경로 "../data/tutorial_data.csv"가 의미하는 것은 "실행하고 있는 주피터 노트북 파일을 기준으로 상위 폴더로 간다음 거기에 존재하는 data 폴더에 tutorial_data.csv 파일이 존재한다." 입니다.

이와 같이 실행하고 있는 주피터 노트북 파일을 기준으로 불러올 csv 파일까지 접근 하는 과정을 나타내는 경로 표현 방식이 상대경로입니다.

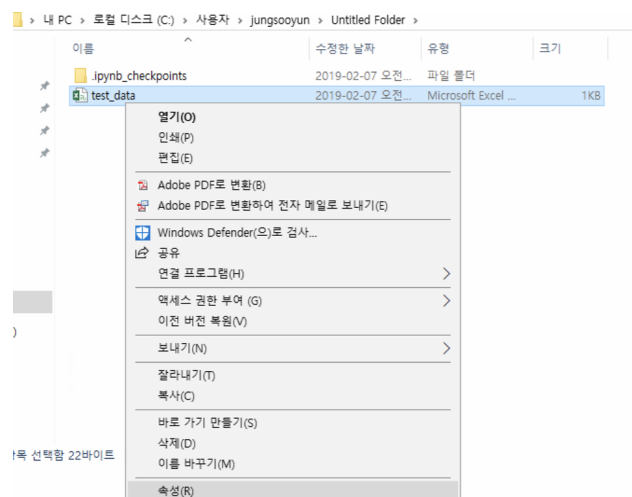
절대경로 입력법

주피터 노트북 파일과 불러올 csv 파일의 거리가 매우 먼 경우에는 상대 경로를 입력하기 어렵습니다. 이러한 경우에는 절대 경로를 입력하는 것이 편합니다.

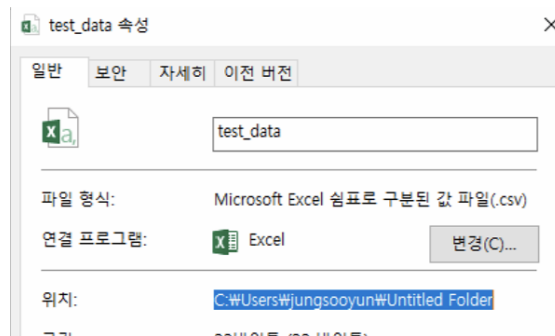
절대경로란 불러올 파일이 속해있는 드라이브(Drive)를 시작으로 파일까지의 접근 과정을 나타내는 경로입니다. 상대경로와 비슷하지만 시작하는 기준이 주피터 노트북 파일이 아닌 Drive 라는 차이가 있습니다.

절대 경로는 편하게 확인할 수 있는 방법이 있습니다.

(윈도우)



위와 같이 불러올 csv 파일을 오른쪽 클릭한 뒤 속성을 클릭하면 다음과 같이 파일의 절대 경로를 볼 수 있습니다.

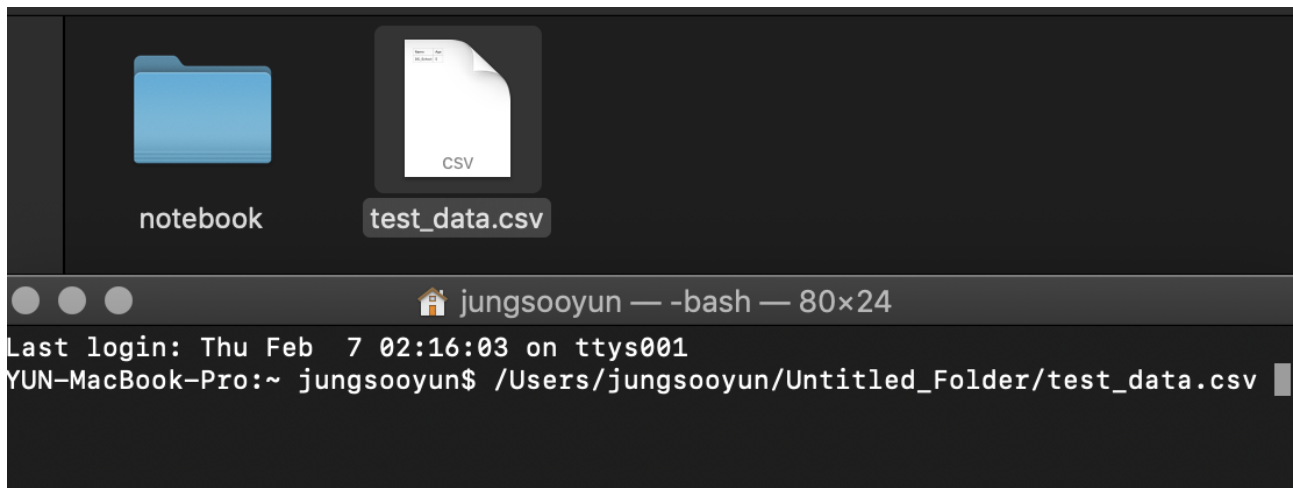


C 드라이브에서 시작하는 절대경로를 확인할 수 있습니다. 이 경로를 복사한 아래와 같이 원화 표시(₩)를 슬래시 (/)로 바꾸고 마지막에 "/파일이름.csv"을 붙여서 입력하시면 됩니다.

```
pd.read_csv("C:/Users/jungsooyun/Untitled Folder/test_data.csv")
```

(맥OS)

맥 OS의 경우는 파일을 찾은 뒤 이를 텍스트 에디터나 터미널에 드래그 해주면 다음과 같이 경로가 표시됩니다.



위 경로를 그대로 복사해서 입력해주시면 됩니다.

주피터 노트북 종료 및 재실행

마지막으로 주피터 노트북 종료에 대하여 알아보겠습니다. 주피터 노트북은 검은 커널창까지 모두 종료해야 완전히 종료됩니다. 주피터 노트북은 종료되는 순간 실행되었던 메모리는 모두 사라지지만 코드는 그대로 남습니다. 주피터 노트북을 종료후 다시 실행한 다음 위에서 만들었던 tutorial.ipynb 파일을 다시 열면 아래와 같습니다.

```
In [1]: print("Hello World!")
Hello World!

Nothing is imposible !

In [2]: import pandas as pd

In [4]: pd.read_csv("../data/tutorial_data.csv")
Out[4]:
```

	Name	Value
0	DS_School	1
1	Korea	2
2	python	3

```
In [ ]:
```

얼핏 보기에는 달라진게 없어보입니다. 하지만 `pd.read_csv("../data/tutorial_data.csv")`이 적혀 있는 셀만 실행하면 아래와 같은 에러가 뜹니다.

```
In [2]: import pandas as pd

In [1]: pd.read_csv("../data/tutorial_data.csv")

-----
NameError                                Traceback (most recent call last)
<ipython-input-1-4ff0f7414cc1> in <module>
----> 1 pd.read_csv("../data/tutorial_data.csv")

NameError: name 'pd' is not defined
```

우선 왼쪽에서 `In : [1]` 을 볼 수 있습니다. 이는 해당 셀이 첫번째로 실행되었다는 것을 뜻합니다. 그리고 맨 아래에 `NameError : name 'pd' is not defined` 에러를 확인할 수 있는데 이는 위의 셀의 `import pandas as pd` 가 실행되지 않았기 때문에 발생합니다. 종료전에 실행하였던 메모리가 모두 날라간 것을 확인할 수 있습니다. 따라서 주피터 노트북을 종료후 다시 실행하였다면 코드를 처음부터 다시 실행해야 합니다.

HTML 파일 실행하기

HTML은 Hyper Text Markup Language의 약자이며 WWW(World Wide Web)에서 하이퍼 텍스트 문서를 작성하는데 사용되는 기본언어입니다. HTML 파일(.html)은 HTML로 작성된 파일을 뜻합니다.

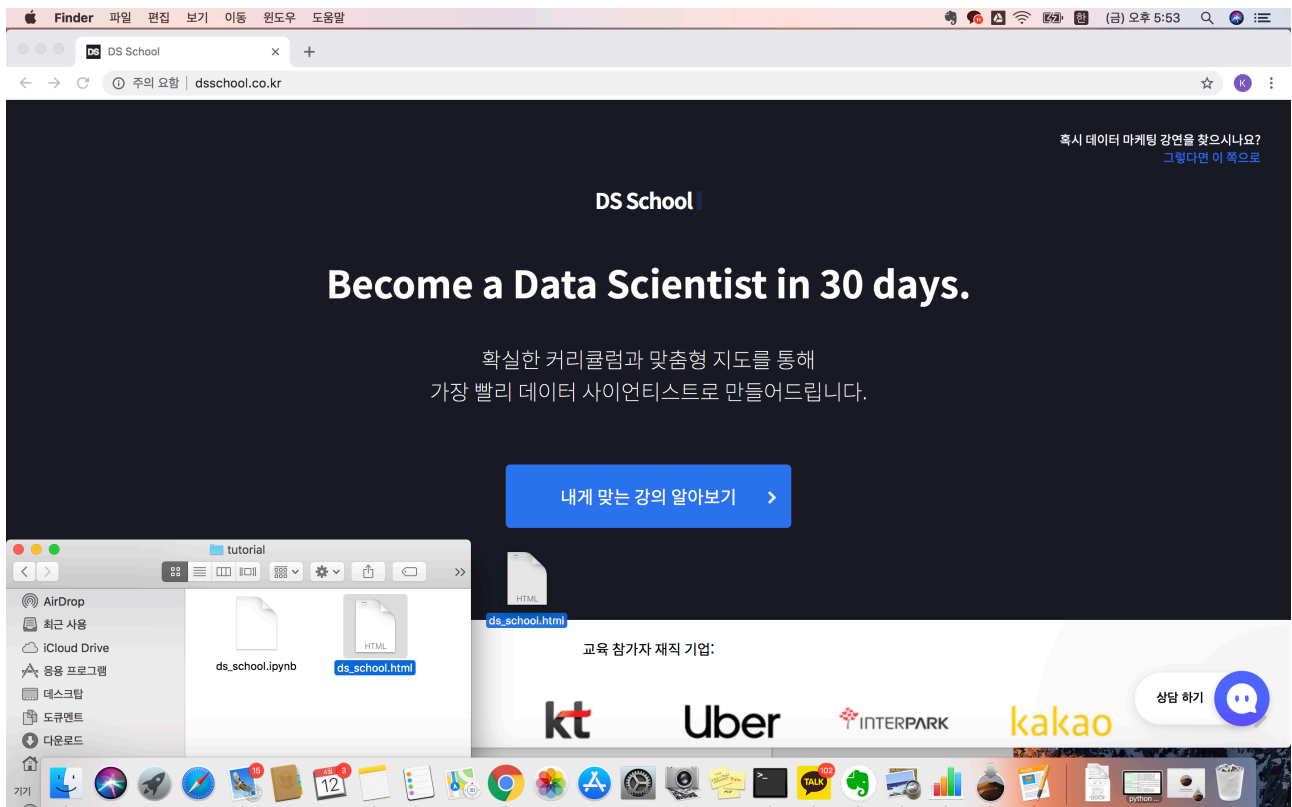
수업 후에 제공드리는 자료중에 HTML 파일이 있기 때문에 HTML 파일을 실행하는 방법을 알아보겠습니다.

우선 HTML 파일을 어떻게 열고 주피터 노트북 파일(.ipynb)와는 어떻게 다른지 알아보겠습니다.



ds_school.html

위의 ds_school.html 파일을 실행해 보겠습니다. 실행하는 방법은 크게 두가지가 있습니다.
첫번 째로 아래의 이미지처럼 브라우저(익스플로어, 크롬 등)를 실행시켜 주신 뒤 ds_school.html 파일을 브라우저 화면에 드래그 하면 됩니다. 브라우저의 종류는 상관 없습니다.



두번 째는 해당 파일을 더블 클릭한 후 연결할 브라우저(익스플로어, 크롬 등)를 선택하면 됩니다. 연결할 브라우저가 기본으로 설정되어 있으면 바로 브라우저를 통해 열립니다.

파일이 열리면 아래와 같습니다.



열린 파일을 보면 주피터 노트북 파일(.ipynb)과 거의 비슷해 보이지만 수정이 불가능합니다. 즉 주피터 노트북 파일의 읽기 전용 버전이라고 생각하시면 됩니다.

HTML 파일은 주피터 노트북을 통해서 열지 않아도 되기 때문에 모바일에서도 열 수 있습니다. 따라서 이동중에서 편리하게 핸드폰으로 열어보실 수 있는 것이 장점입니다. DS School에서도 수강생 여러분께서 편하게 수업 자료를 보시며 복습하실 수 있게 HTML 파일로 제공해 드리고 있습니다. 만약 다른 주피터 노트북 파일(.ipynb)를 HTML 파일로 변환하고 싶다면 아래의 그림과 같이 file을 클릭한 다음 Download as -> HTML을 클릭하면 됩니다.

