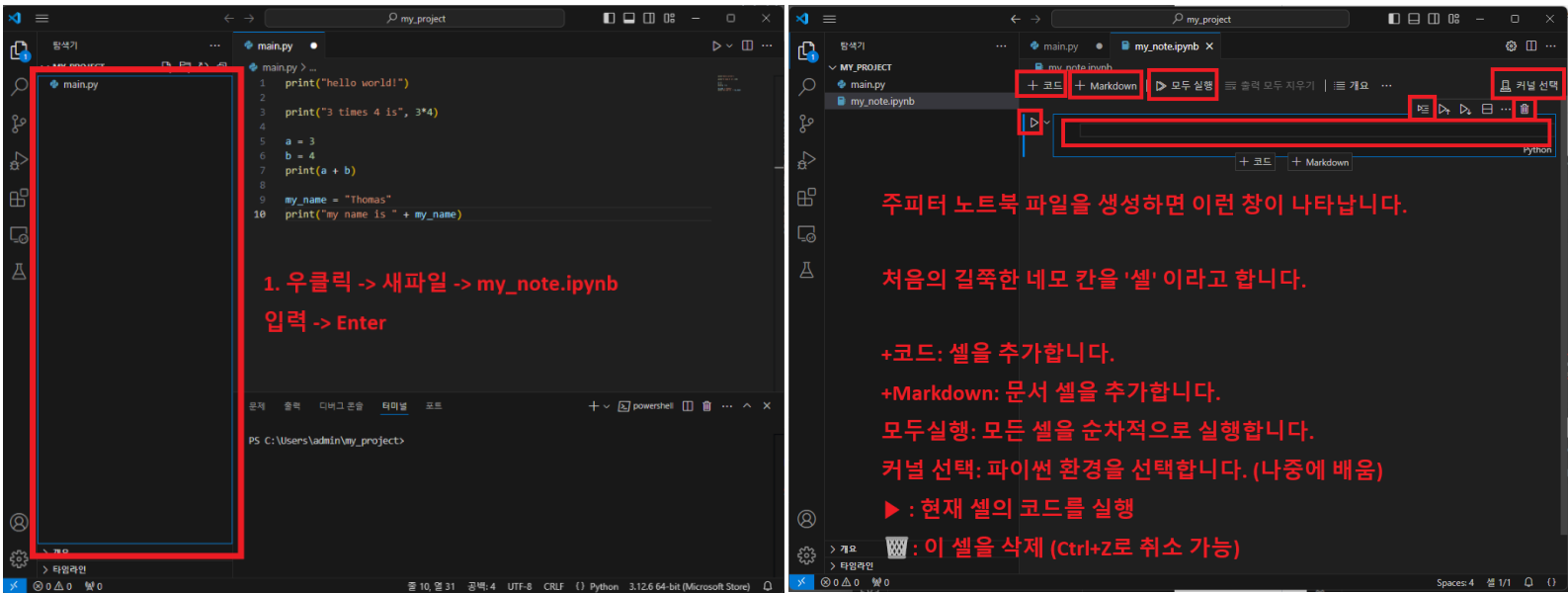


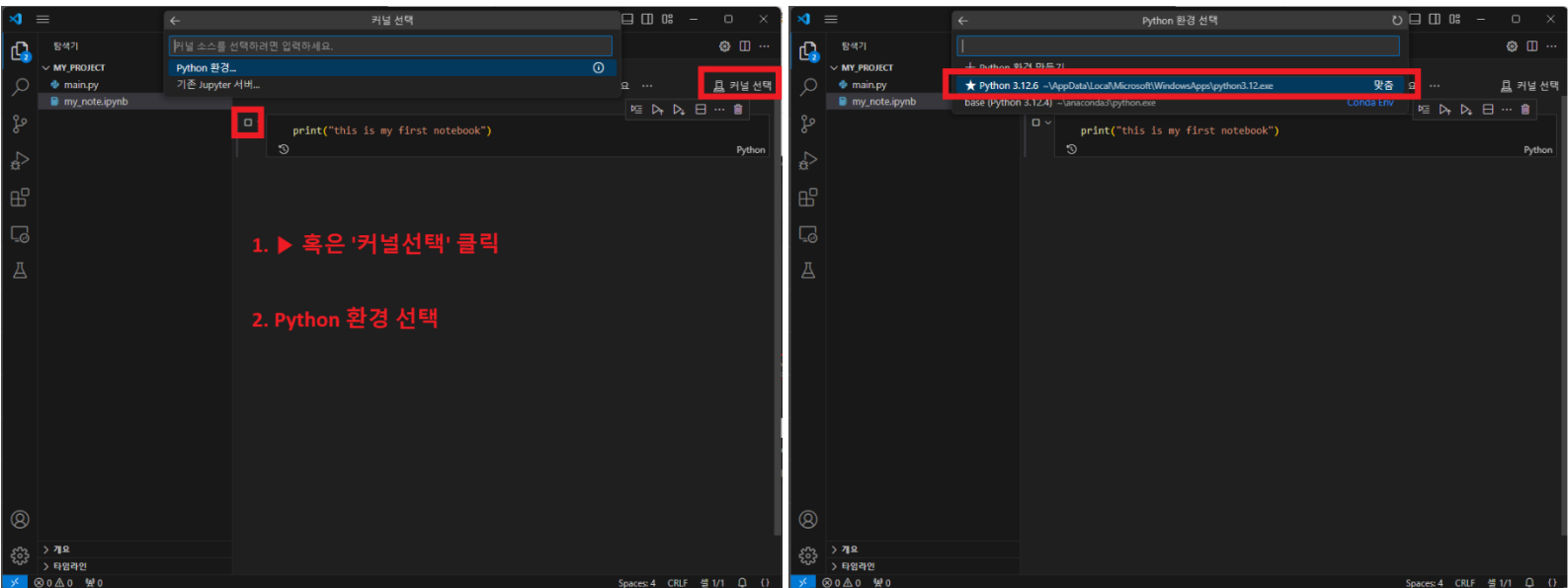
## 진행목표

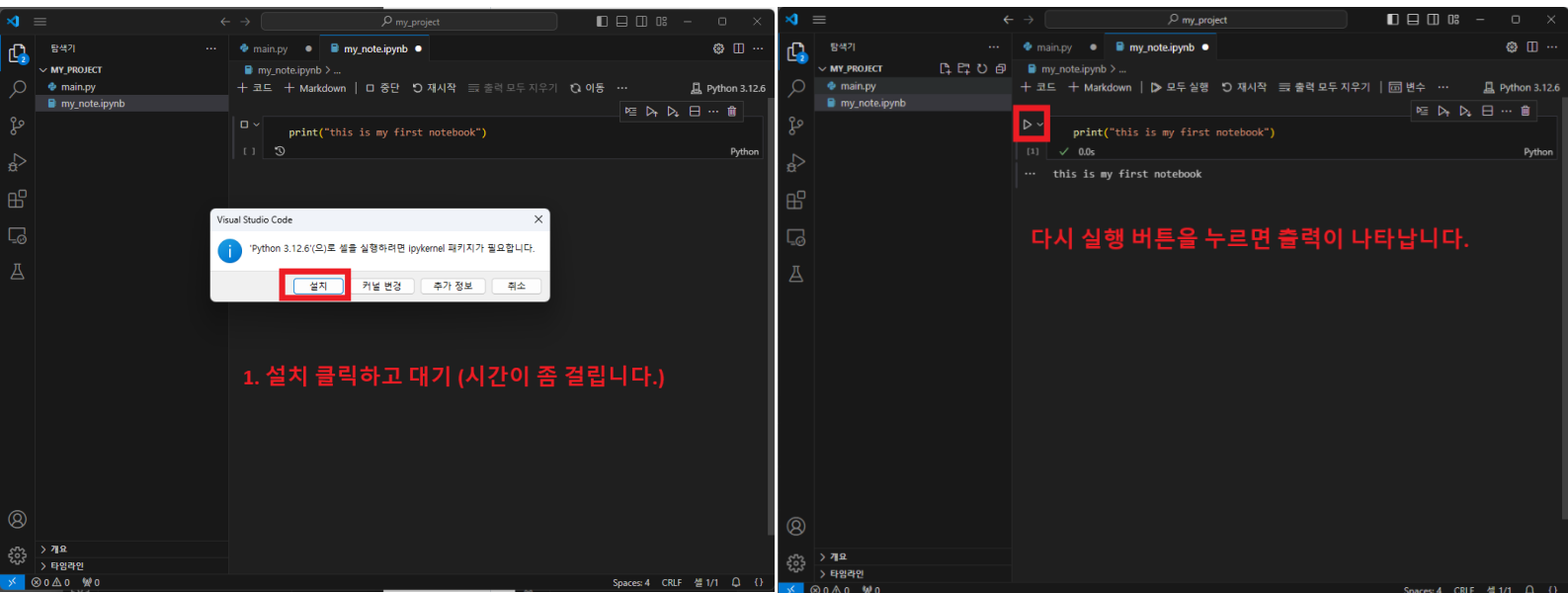
주피터 노트북(.ipynb 확장자) 사용법 익히기

## Step1. 주피터 노트북 만들기



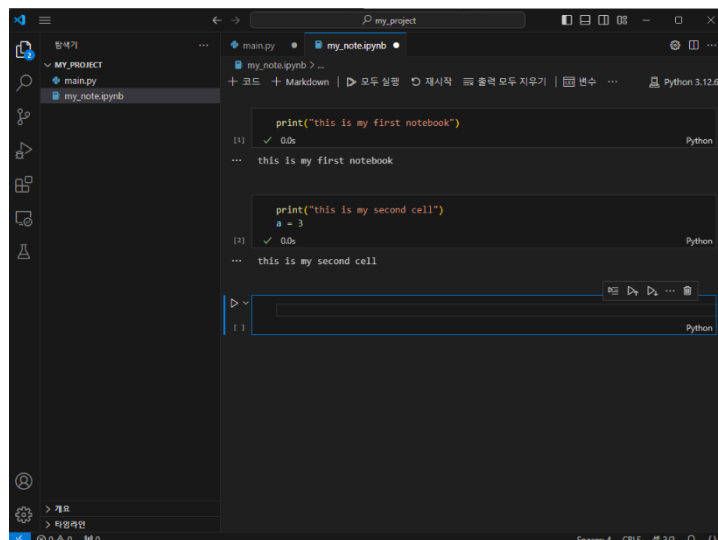
## Step2. 커널 선택



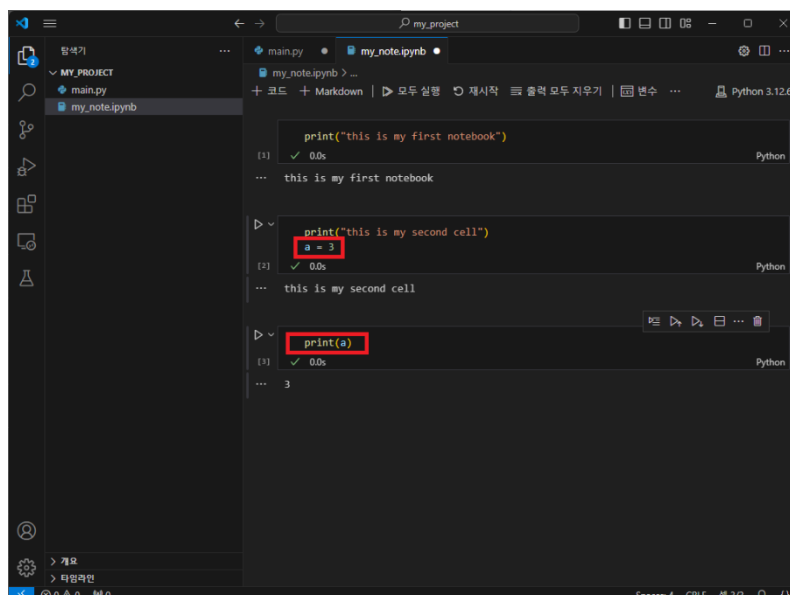


## Step3. 셀 추가하기

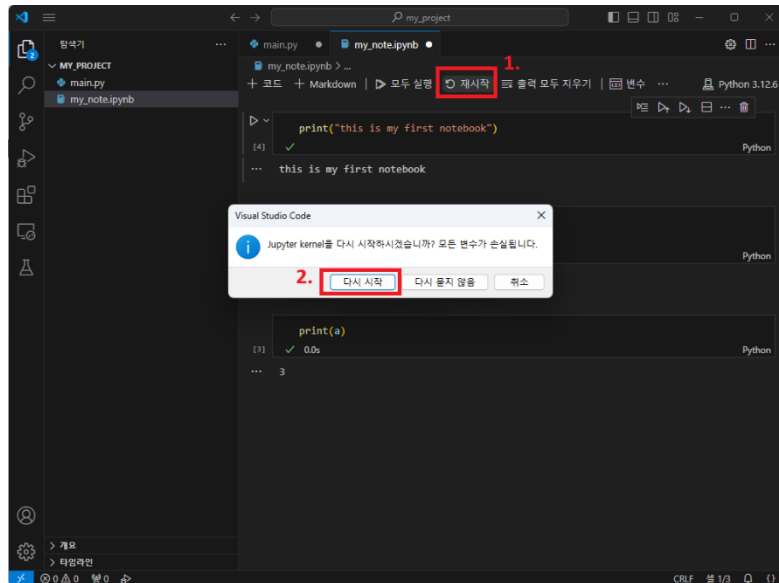
1. 셀 상, 하단에 마우스 커서를 가져대면 '+코드' 버튼이 나타납니다.  
셀을 추가하고 코드를 입력하고 실행해봅시다.



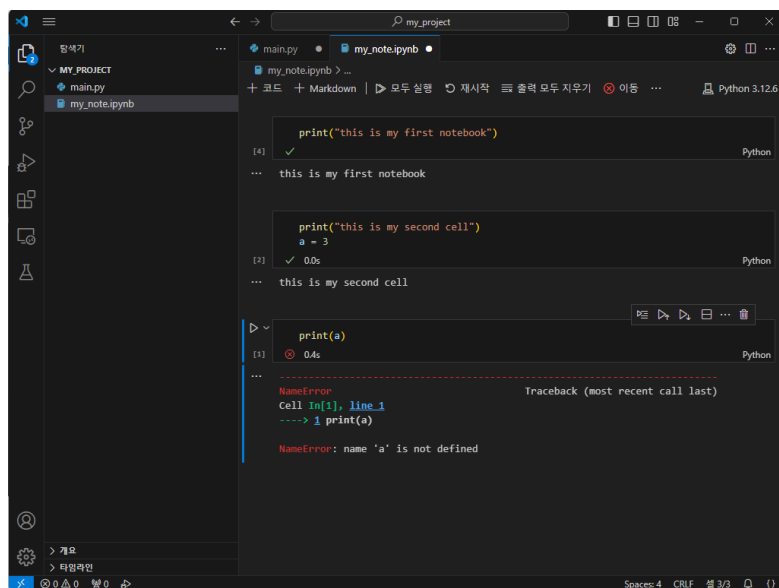
2. 주피터 노트북은 이전에 실행된 코드의 실행결과를 메모리에 저장해두고 있습니다.  
(2 번째 셀에서 선언된 변수 a를 다시 사용할 수 있음)



3. 재시작(Restart) 버튼을 누르면 파이썬 인터프리터를 초기화해 현재 저장된 상태를 초기화합니다.



4. 재시작 버튼을 누르고 가장 먼저 세번째 셀을 다시 실행시켜보면 **NameError** 가 발생합니다.  
(변수 a가 선언이 되지 않았으므로)  
다시 두번째 셀을 실행시킨 뒤, 세번째 셀을 실행시키면 에러가 사라지고 정상적으로 출력이 됩니다.

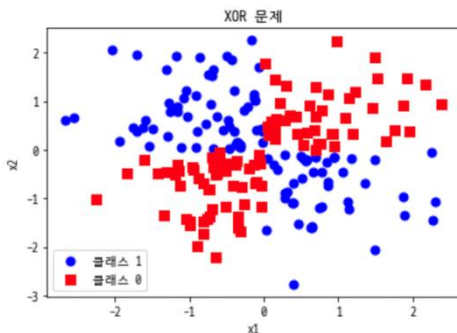


## 주피터 노트북이란?

- 데이터과학, 머신러닝, 통계학, 이미지 처리 등 다양한 분야에서 널리 사용되는 오픈소스 웹 어플리케이션
- 코드를 '셀' 단위로 작성해 각각 독립적으로 실행할 수 있다.
- 따라서 코드를 실행할 때마다 결과를 저장하고 재사용할 수 있어 편리
- 마크다운(Markdown) 을 사용해 설명 텍스트, 이미지, 수식을 추가할 수 있다.
- Matplotlib, seaborn 같은 라이브러리를 사용해 데이터를 시각화 할 수 있음

## 예시)

```
In [1]: 1 np.random.seed(0)
2 X_xor = np.random.randn(200, 2)
3 y_xor = np.logical_xor(X_xor[:, 0] > 0, X_xor[:, 1] > 0)
4 y_xor = np.where(y_xor, 1, 0)
5 plt.scatter(X_xor[y_xor == 1, 0], X_xor[y_xor == 1, 1],
6             c='b', marker='o', label='클래스 1', s=50)
7 plt.scatter(X_xor[y_xor == 0, 0], X_xor[y_xor == 0, 1],
8             c='r', marker='s', label='클래스 0', s=50)
9 plt.legend()
10 plt.xlabel("x1")
11 plt.ylabel("x2")
12 plt.title("XOR 문제")
13 plt.show()
```



### LinearRegression Example

To go over this model, we'll use the `medical_insurance.csv` dataset

- contains information about 1338 people (age, sex, bmi, whether they have children or not, etc.)
- we want to see if we can predict their insurance cost

```
In [165]: 1 import pandas as pd
2 import numpy as np
3 from sklearn.preprocessing import OneHotEncoder
4 from sklearn.preprocessing import StandardScaler
5 from sklearn.model_selection import train_test_split
6 from sklearn.model_selection import cross_val_score
7 from sklearn.linear_model import LinearRegression
8 from sklearn.metrics import mean_squared_error as mse
9 from pandas.plotting import scatter_matrix
10 import matplotlib.pyplot as plt
11 import seaborn as sns
12
13 %matplotlib inline
14 sns.set()
```

Let's read the data and do some exploratory data analysis

```
In [166]: 1 data = pd.read_csv('data/medical_insurance.csv')
2 data.head()
```

```
Out[166]:
```

	age	sex	bmi	children	smoker	region	charges
0	19	female	27.900	0	yes	southwest	16884.92400
1	18	male	33.770	1	no	southeast	1725.55230