

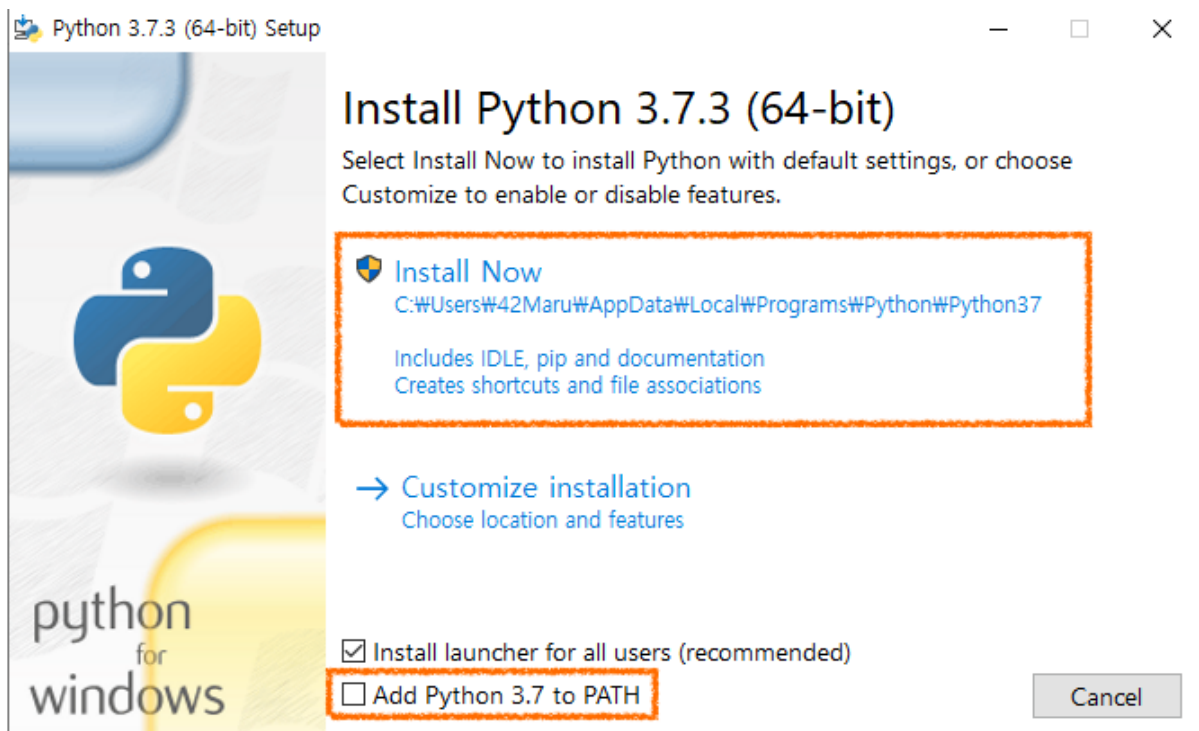


크롤러 설치 및 사용 메뉴얼

파이썬 셀레니움을 사용

python 설치 (설치파일로 설치)

<https://www.python.org/ftp/python/3.10.11/python-3.10.11-amd64.exe>



- 사진은 3.7이지만 3.10 버전으로 설치
- Add python 3.7 to PATH를 꼭 클릭

cmd 실행 후 아래 명령어로 셀레니움 설치

```
pip install selenium==4.10  
  
pip install webdriver_manager
```

크롤러 파이썬 파일 다운로드

https://drive.google.com/file/d/1RgGS4RysxBTsxSivxJLfEPoX9oraiO_v/view?usp=sharing

run.py를 비어있는 폴더를 만들어 집어넣음

그 뒤, cmd에 해당 폴더로 cd 명령어를 이용하여 들어가서 크롤러 실행

```
cd <폴더 디렉토리>  
  
python run.py
```

크롤러가 실행되면 검색어를 입력할 수 있고, 해당 검색어로 구글 이미지 검색을 함.

검색해야할 키워드는 4가지

- <꽃 이름> (영문명으로 해야함 ex: 빨간 장미 ⇒ red rose)
- <꽃 이름> flower (ex: red rose flower)
- bouquet of <꽃이름>(ex: bouquet of red rose)
- flowershop <꽃이름> (ex: flowershop red rose)

4가지 키워드로 검색해야 함으로, 하나의 꽃 종류마다 4번 run.py를 실행시켜야함

이때, 네트워크 오류로 인해 크롤링이 시작되지 않는 경우가 있으므로, 1분이 지나도 크롤링이 시작이 안되면 강제 종료 (cmd에서 컨트롤+c) 이후 재시도

4가지 키워드로 검색하면 약 1600개 ~ 2400개의 이미지가 imgs 폴더 내부에 저장됨.

imgs 폴더의 이름을 빨간 장미 변경전 같은 형태로 바꾸고, 다음 꽃 크롤링 시작

화면이 꺼지면 크롤링이 중단되기 때문에 아래를 신경써야함

- 노트북 덮개를 덮지 않는다
- 크롤링이 시작되면 윈도우 찾기에서 "화면이 꺼지는 시기 변경"을 검색하여, 모두 꺼지지 않음으로 변경,
- 윈도우 찾기에서 "화면 보호기" 검색 후 화면 보호기 비활성화

트러블 슈팅

셀레니움의 크롬 드라이버 사용

셀레니움 버전에 따라서, 크롬 드라이버를 따로 다운로드 받거나

이미지 다운로드 시간