과제3) 아래 코드는 자막 파일을 읽어 순번과 타이밍 정보를 제외하고 자막 내용만 별도의 텍스트 파일로 저장하는 코드의 일부이다.

첨부파일로 제공된 subtitles.zip 파일을 현재 작업디렉토리에 압출을 풀어 디렉토리 내의 모든 자막파일을 순회하여 작업을 처리하세요.

1. 대상 디렉토리를 기준으로 모든 하위 디렉토리를 함수 중에 os.walk() 함수가 있다. walk() 함수를 사용하는 사례와 동작방식에 대해서 기술하세요.
2. 자막파일의 순번라인과 타이밍정보를 포함한 라인을 식별하기 위해 정규식을 사용할 수 있다. 순번라인을 식별하는 정규식과 타이밍정보 라인을 식별하는 정규식을 각각 re.compile() 의 인자로 들어갈 수 있게 로우(raw) 스트링으로 작성하세요.
3. 아래 코드를 참조하여 자막파일에서 자막내용만을 추출하여 다시 파일로 쓰는 코드를 완성하세요. 단, 자막파일과 같은 경로에 자막내용만 추출된 파일을 자막파일명.txt 로 저장하세요.

import glob

import os

def get\_subtitle(file\_name):

sub\_title\_content = []

tmp\_line = ''

file = open(file\_name)파일

for line in file:

//TODO 여기를 작성하여 코드를 완성하세요

return sub\_title\_content

def make\_file\_and\_save(content, file\_name, ext='txt'):

file = open(file\_name + '.' + ext, 'a')

for line in content:

file.write('%s\n' % line)

file.close()

def clear\_dir():

file\_list = glob.glob('\*')

for item in file\_list:

if item.endswith('.txt'):

os.remove(item)

def main():

os.chdir('./Intro to Data Analysis Subtitles')

dir\_list = glob.glob('\*')

os.chdir('./' + dir\_list[0])

print('current dir is : ' + os.getcwd())

#

clear\_dir()

subtitle\_list = glob.glob('\*.srt')

for file\_name in subtitle\_list:

print(file\_name, "is working..")

subtitle\_content = get\_subtitle(file\_name)

make\_file\_and\_save(subtitle\_content, file\_name, 'txt')

print('job completed..')

if \_\_name\_\_ == '\_\_main\_\_':

main()