**5장.**

**p.153 [CODE3] 수정**

※ 파란색으로 표시한 부분을 추가하였습니다.

|  |  |
| --- | --- |
| 01  02  03  04  05  06  07  08  09  10  11  12  13  14  15  16  17  18  19  20  21  22  23  24  25  26  27  28 | **def getTourismStatsService(**nat\_cd, ed\_cd, nStartYear, nEndYear**):**  jsonResult = []  result = []  **dataEND = "{0}{1:0>2}".format(str(nEndYear), str(12)) #데이터 끝 초기화**  **isDataEnd = 0 #데이터 끝 확인용 flag 초기화**  **for** year **in** **range(**nStartYear, nEndYear+1**):**  **for** month **in range(**1, 13**):**  **if(isDataEnd == 1): break #데이터 끝 flag 설정되어있으면 작업 중지.**  yyyymm = "{0}{1:0>2}".format(str(year), str(month))  jsonData = getTourismStatsItem(yyyymm, nat\_cd, ed\_cd) #[CODE 2]  **if (**jsonData['response']['header']['resultMsg'] == 'OK'**):**  #입력된범위까지 수집하지 않았지만, 더이상 제공되는데이터가 없는 마지막 항목인 경우---  **if** jsonData['response']['body']['items'] == ''**:**  **isDataEnd = 1** #데이터 끝 flag 설정  dataEND = "{0}{1:0>2}".format(str(year), str(month-1))  print("데이터 없음.... \n 제공되는 통계 데이터는 %s년 %s월까지입니다."  %(str(year), str(month-1)))  break  #jsonData를 출력하여 확인......................................................  print (json.dumps(jsonData, indent=4,  sort\_keys=True, ensure\_ascii=False))  natName = jsonData['response']['body']['items']['item']['natKorNm']  natName = natName.replace(' ', '')  num = jsonData['response']['body']['items']['item']['num']  ed = jsonData['response']['body']['items']['item']['ed']  print('[ %s\_%s : %s ]' %(natName, yyyymm, num))  print('----------------------------------------------------------------------')  jsonResult.append({'nat\_name': natName, 'nat\_cd': nat\_cd,  'yyyymm': yyyymm, 'visit\_cnt': num})  result.append([natName, nat\_cd, yyyymm, num])  **return (**jsonResult, result, natName, ed, dataEND**)** |

04행 수집할 데이터의 끝 날짜(**dataEND**)를 nEndYear의 12월로 설정한다.

05행 수집한 데이터의 끝인지 확인하기 위한 플래그(**isDataEnd**)를 0으로 설정한다.

08행 데이터 끝 플래그(**isDataEnd** )를 확인하여 플래그가 설정되어 있으면 작업을 중단한다.

09행 수집할 연도와 월을 여섯 자리로 맞추어 yyyymm에 저장한다.

10행 getTourismStatsItem()을 호출해 받은 월 데이터를 jsonData에 저장한다.

11행 응답 데이터가 정상인지 확인한다.

13~17행 ['items'] 항목에 값이 없으면 출입국관광통계 데이터가 아직 들어가지 않은 월이므로, **isDataEnd를 1로 설정한다. 그리고** 마지막 데이터가 있는 날짜(month-1)를 dataEND에 저장하고 데이터 수집 작업을 중단한다.

19행 수집한 월 데이터인 jsonData 내용을 확인할 수 있게 파이썬 셸 창에 출력한다.

20~21행 수집한 국가 이름인 ['natKorNm'] 항목의 값에서 띄어쓰기를 제거하고, natName에 저장한다.

22행 수집한 월의 데이터 수인 ['num'] 항목의 값을 num에 저장한다.

23행 수집한 출입국 구분 데이터인 ['ed'] 항목의 값을 ed에 저장한다.

26행 수집한 국가 이름(natName), 국가 코드(nat\_cd), 날짜(yyyymm), 데이터 수(num)를 딕셔너리 자료형으로 구성하여 jsonResult 리스트에 원소로 추가한다.

27행 수집한 국가 이름(natName), 국가 코드(nat\_cd), 날짜(yyyymm), 데이터 수(num)를 result 리스트에 원소로 추가한다.

28행 수집하여 정리한 데이터를 반환한다.

**p.155~156 [프로그램 5-2] openapi\_tour.py**

**…**

**#[CODE 3]**

**def getTourismStatsService(**nat\_cd, ed\_cd, nStartYear, nEndYear**):**

jsonResult = []

result = []

**dataEND = "{0}{1:0>2}".format(str(nEndYear), str(12)) #데이터 끝 초기화**

**isDataEnd = 0 #데이터 끝 확인용 flag 초기화**

**for** year **in** **range(**nStartYear, nEndYear+1**):**

**for** month **in range(**1, 13**):**

**if(isDataEnd == 1): break #데이터 끝 flag 설정되어있으면 작업 중지.**

yyyymm = "{0}{1:0>2}".format(str(year), str(month))

jsonData = getTourismStatsItem(yyyymm, nat\_cd, ed\_cd) #[CODE 2]

**if (**jsonData['response']['header']['resultMsg'] == 'OK'**):**

#입력된범위까지 수집하지 않았지만, 더이상 제공되는데이터가 없는 마지막 항목인 경우---

**if** jsonData['response']['body']['items'] == ''**:**

**isDataEnd = 1** #데이터 끝 flag 설정

dataEND = "{0}{1:0>2}".format(str(year), str(month-1))

print("데이터 없음.... \n 제공되는 통계 데이터는 %s년 %s월까지입니다."

%(str(year), str(month-1)))

break

#jsonData를 출력하여 확인......................................................

print (json.dumps(jsonData, indent=4,

sort\_keys=True, ensure\_ascii=False))

natName = jsonData['response']['body']['items']['item']['natKorNm']

natName = natName.replace(' ', '')

num = jsonData['response']['body']['items']['item']['num']

ed = jsonData['response']['body']['items']['item']['ed']

print('[ %s\_%s : %s ]' %(natName, yyyymm, num))

print('----------------------------------------------------------------------')

jsonResult.append({'nat\_name': natName, 'nat\_cd': nat\_cd,

'yyyymm': yyyymm, 'visit\_cnt': num})

result.append([natName, nat\_cd, yyyymm, num])

**return (**jsonResult, result, natName, ed, dataEND**)**

#[CODE 0]

…

**8장. 수정.**

**p.244 : 맨 아래**



**Tip 추가** 내부적으로 연결하여 사용하는 nltk의 리소스를 최초 한번은 다운로드 받아야 한다. 아나콘다에는 nltk가 기본으로 설치되어있으므로 설치(pip install) 할 필요 없이, import nltk 를 한 후에 nltk.download()를 하여 download 대화상자가 뜨면 모두 선택하고 [download] 버튼을 누르면 된다.

