

과제 **6 (ver1.1)**

문제풀이 보고서

2023100085 신희원

2023100086 전하은

2023100096 김익희

1. 파트 분배 과정

팀원이 세명이니 과제설명 pdf에 나온 그대로 데이터 추출 파트(A), 그래프 구현 파트 (B), 메인 모듈 작성 파트(C) 이렇게 세 파트로 설계함. 각각 김익희 (2023100096), 신희원 (2023100085), 전하은(2023100086) 이 전담하기로 결정하였음.

2. A_part의 반환 값 설정 (build_data.py)

1. 사전에 나는 대화를 기반으로 아래 사진과 같은 형식으로 데이터를 반환하기로 함.

```
"과목명1": {  
    "표준점수1": [남학생수1, 여학생수1],  
    "표준점수2": [남학생수2, 여학생수2],  
    ...  
},
```

2. csv 모듈 import

- csv 파일 속의 데이터를 효율적으로 다루기 위해 csv 모듈 불러오기

3. 변수 설정 및 파일 불러오기

- 추출한 데이터를 저장할 딕셔너리인 "data"를 생성
- open 메서드를 이용하여 csv 파일을 불러오기
- csv 모듈에 있는 reader를 사용하여 csv 파일의 각 행을 리스트로 변환하여 각의 줄별로 reader에 데이터가 저장됨

4. 사전에 계획한 형식대로 데이터 추출

- 반복문을 사용하여 모든 행과 열에 있는 데이터를 갖고 온다
- data[line[1]] : 이중 딕셔너리로 반환할 것이기 때문에 key값이 과목명인 딕셔너리를 생성
- 안쪽 딕셔너리는 점수가 key값이 되도록 하고, value값은 남학생 수(line[3]), 여학생 수(line[4])가 리스트의 형식으로 들어가도록 딕셔너리 업데이트
- 각 열에 저장된 데이터를 점수와 인원은 정수형이 되도록 딕셔너리에 저장

5. 사전에 약속된 형식으로 추출한 data 딕셔너리를 반환하고 함수(build_data) 종료

6. version 1.0과의 차이점

- 차이점이 있는 부분은 아래 코드와 같음

```
while True:  
    year_input = input("연도 선택(2020, 2021, 2022, 2023) : ")  
    if year_input in ['2020', '2021', '2022', '2023']:  
        break  
    else:  
        print("잘못된 입력입니다.")  
  
    input_file = year_input + '1231.csv'  
    f = open(input_file, 'r', encoding = 'euc-kr')
```

- 파일명이 정해져있던 version 1.0과 달리, version 1.1에서는 연도를 입력받아 해당 연도에 맞는 파일을 처리해야했음
- 따라서 반복문을 사용해 연도를 입력받았고, 이 과정에서 잘못된 입력이 들어올 경우 다시 입력하도록 처리했음.
- 또한 string 연산자를 이용해서 연도에 맞는 파일이 처리되도록 처리했음

3. B_part의 매개변수 설정 (Draw_graph.py)

그래프를 그리기에 앞서 그래프로 무엇을 시사하고 싶은지 확인을 해야 그래프의 종류를 설정 할 수 있으니 해당 사항을 파악해 보았다. 우리는 이 활동으로 표준점수에 따른 남녀 학생들의 분포 우세 정도를 파악하는 것이 목표라고 생각하여 아래 그래프같은 영역그래프로 그리기로 결정했다.

데이터를 받는 것은 앞서 1.0버전과 똑같이 설정하였고, 하나 달라진 것은 남녀의 분포 퍼센테이지를 확인하고 싶은 것이니 y 값을 기존 학생수에서 (알고 싶은 성별의 학생 수) / (전체 학생 수) * 100으로 백분율로 바꾸어 표기하기로 하였다. 여기서 남, 여 두 y값 모두 위의 식으로 구하면 전체 퍼센테이지가 99% 혹은 101%가 되어 가시적으로 보기 깔끔하지 않은 그래프가 나오게 되니 남자 y 값은 (남 학생 수) / (전체 학생 수) * 100 으로 여자 y값은 100 - (남자 y값)으로 설정하였다.

+) 영역그래프로 하면 기존에 설정한 red, blue의 색이 너무 채도가 높아 눈의 피로감을 높이고 가독성이 떨어져 색을 조금더 부드러운 색을 설정해 주었다.

+) 범례를 설정하여 조금 더 해당 색이 어떤 것을 가르키는 것인지 직관적으로 알 수 있게 설정했다.

+) 조금 더 수치를 정확히 가능할 수 있도록 격자 설정을 켜 두었다.

4. C_part의 메인모듈 설계 (main.py)

1. build_data.py와 Draw_graph.py 파일을 import해서 불러오기

2. 사용자에게 과목명 입력받기 (숫자로 입력받도록 변경)

- ver1.0에서는 사용자로부터 직접 “국어”와 같이 텍스트로 입력받도록 했지만, ver1.1 과제 설명 파일에 나와 있는 “실행 예시”와 동일해지도록 입력받도록 했지만, ver1.1 과제 설명 파일에 나와 있는 “실행 예시”와 동일해지도록 숫자로 입력받도록 함.
- build_data()를 통해 추출한 데이터에서 과목명은 걸 딕셔너리의 key값이므로, 그 key값들을 list로 만든 후, 앞에 번호를 붙여서 다시 리스트로 만든다. (for문 사용)
- input을 받기 위해 앞서 만든 리스트를 출력하면서 입력을 숫자로 받는다.

3. 나머지는 ver 1.0 과 동일하므로 설명 생략

- DrawG의 매개변수가 ver1.0 과 동일하기 때문에 따로 수정할 부분이 없었음.

5. 결과 화면

수학

