

<문제 해결 방안>

1. **Select_subject** 모듈 :

파일에 존재하는 과목을 읽어와서 저장해야 함. 중복값을 제거하는 것과 과목 분류와 세부 과목명을 짝짓는 것이 관건. → np의 **unique**함수를 사용해서 중복값을 제거하고 과목 분류와 세부 과목명을 한 배열에 저장하기로 결정함.

1. **csv** 파일을 읽어서 거기 있는 과목들을 추출함. 일단 과목 분류와 세부 과목이 있는 열만 모든 행에서 읽어온 후 거기서 중복된 값을 제거하여 [과목 분류, 세부 과목 명] 형태의 원소가 저장되어 있는 2차원 배열을 **numpy**모듈의 **ndarray** 자료형으로 생성해서 반환하는 함수가 있음.

2. 또한 사용자가 출력된 과목들을 보고 원하는 과목을 입력하도록 하는 함수가 있음.

3. 이 때 사용자의 입력값이 올바른지 확인하고 올바르다면 **True**, 아니면 **False**를 반환하는 함수가 있음.

1번, 3번 함수는 2번 함수 안에서 호출되고 2번 함수는 사용자의 입력 과목 [과목 분류, 세부 과목 명] 형태의 1차원 배열로 반환함.

2. **Drawinggraph** 모듈 :

numpy 형태로 데이터가 가공된다는 점을 고려하여 배열을 가장 잘 활용할 수 있는 시각화 모듈이 무엇이 있을까 생각해본 결과 **matplotlib** 모듈을 활용하여 그래프를 그리기로 결정하였음. 메인 함수에서 구성된 가공된 데이터를 활용하여 가로축을 표준점수, 세로축을 인원수로 설정하였고, 이에 따라 그래프를 꺾은선 그래프로 각 표준점수마다의 인원수의 경향성을 파악할 수 있도록 구성함. 성별에 따라 성적을 구분함으로써 그래프를 분석함에 있어 용이하도록 함.

3. **Extract** 모듈:

데이터를 가져와서 정리하고 저장하는 역할을 하는 모듈. 데이터를 가져와서 성별을 구분지어서 정리하고, 저장된 **narray**에서 유형, 과목, 성별을 입력해서 유저가 입력한 데이터를 가져오는 두 단계로 나뉜다. 첫 번째 단계에서는 **pandas** 라이브러리를 이용하여 **csv** 파일을 읽고 데이터로 저장한다. 그다음 저장한 데이터를 **numpy** 형태로 변형시키고, 남자 데이터는 1, 여자 데이터는 0으로 구분지어 데이터를 [영역, 유형, 표준편차, 명, 성별] 형태로 저장시킨다. 두 번째 단계에서는 **numpy array** 특성과 **boolean** 특성을 이용해 유저가 입력한 과목명, 과목유형, 성별에 따라 데이터를 필터링 해서 **return**한다.

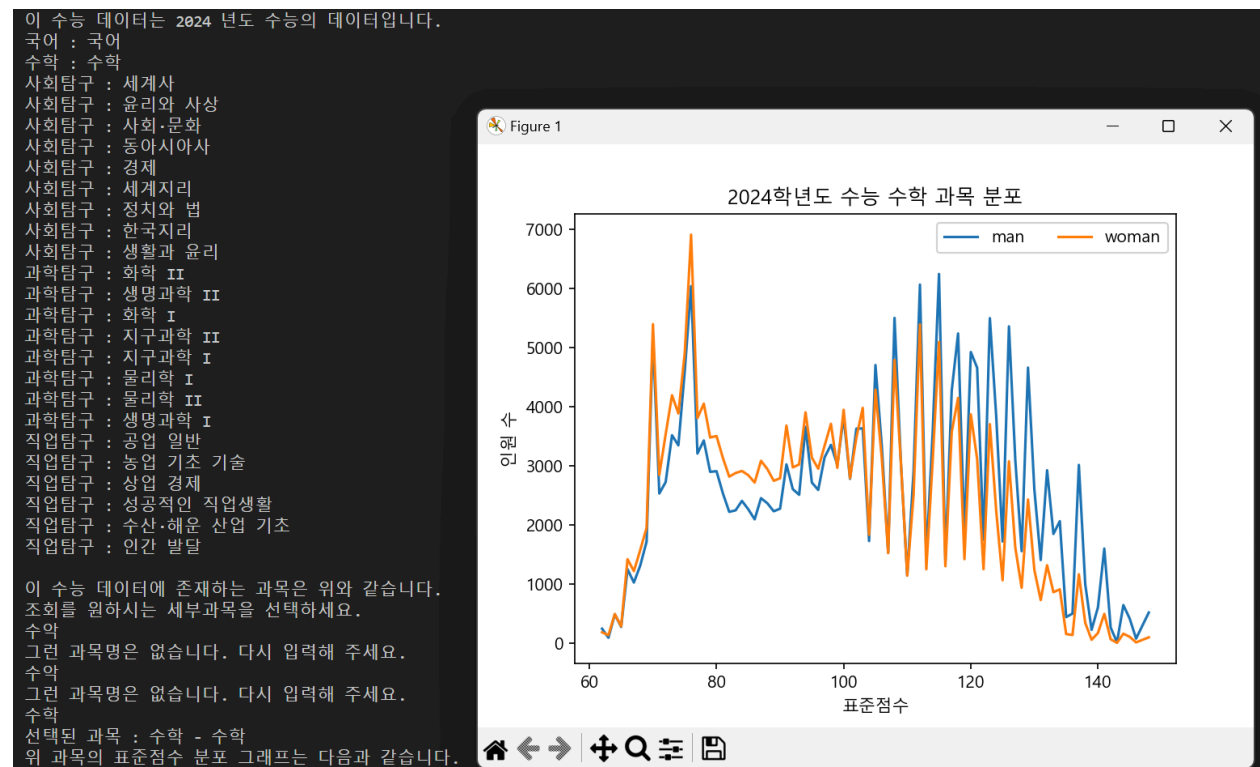
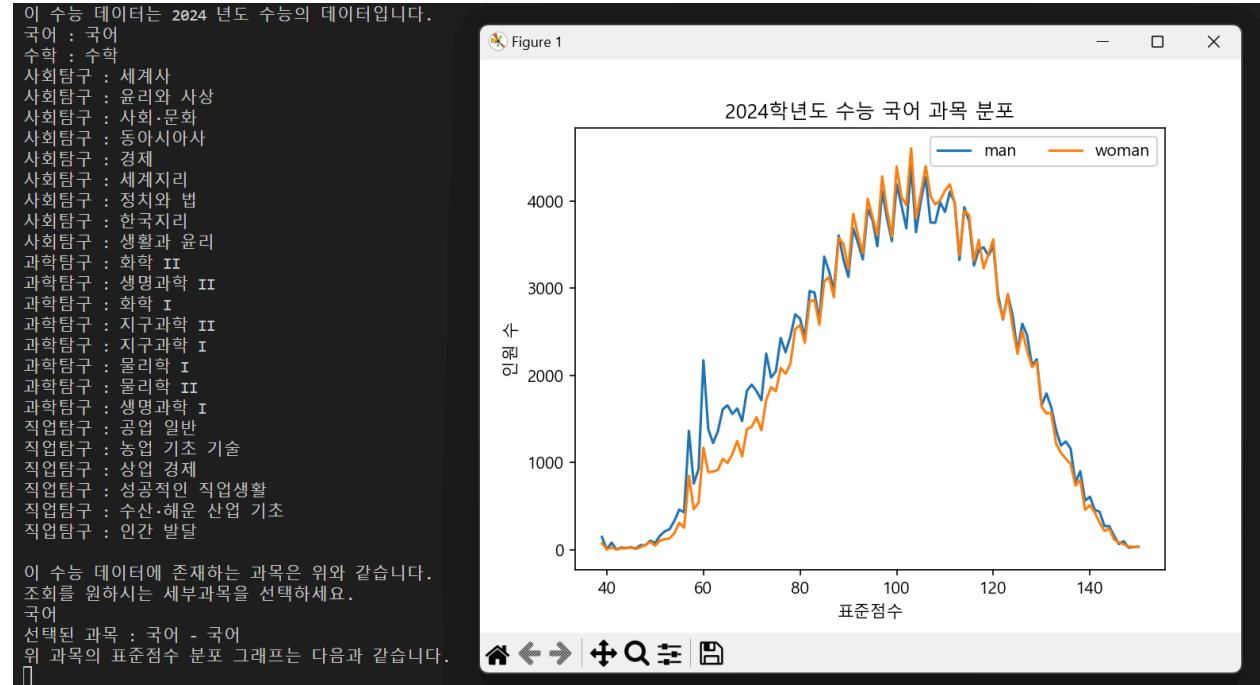
4. 메인 :

csv 파일을 **open**한 후 예외처리를 통해 정상적으로 파일을 열지 못한 경우 프로그램이 멈추는 대신 에러메시지를 내보내도록 함.

파일이 열린 경우 **select_subject** 모듈의 과목 출력하는 함수를 호출하여 존재하는 과목을 출력하고 사용자의 입력을 받음. 또 **extract** 모듈의 데이터 읽어오는 함수를 호출하여 **csv**파일의 데이터를 **numpy** 모듈의 자료형으로 변환하여 저장함.

사용자의 입력값이 저장된 변수를 **extract** 모듈에 있는 원하는 과목, 성별의 데이터만 읽어오는 함수의 인수로 넣어서 따로 만든 **numpy**의 **ndarray**자료형에 저장하고 이것을 **Drawinggraph** 모듈의 그래프 그리는 함수의 인수로 넣어서 그래프를 출력함.

<실행 결과>



이 수능 데이터는 2024 년도 수능의 데이터입니다.

국어 : 국어

수학 : 수학

사회탐구 : 세계사

사회탐구 : 윤리와 사상

사회탐구 : 사회·문화

사회탐구 : 동아시아사

사회탐구 : 경제

사회탐구 : 세계지리

사회탐구 : 정치와 법

사회탐구 : 한국지리

사회탐구 : 생활과 윤리

과학탐구 : 화학 II

과학탐구 : 생명과학 II

과학탐구 : 화학 I

과학탐구 : 지구과학 II

과학탐구 : 지구과학 I

과학탐구 : 물리학 I

과학탐구 : 물리학 II

과학탐구 : 생명과학 I

직업탐구 : 공업 일반

직업탐구 : 농업 기초 기술

직업탐구 : 상업 경제

직업탐구 : 성공적인 직업생활

직업탐구 : 수산·해운 산업 기초

직업탐구 : 인간 발달

이 수능 데이터에 존재하는 과목은 위와 같습니다.

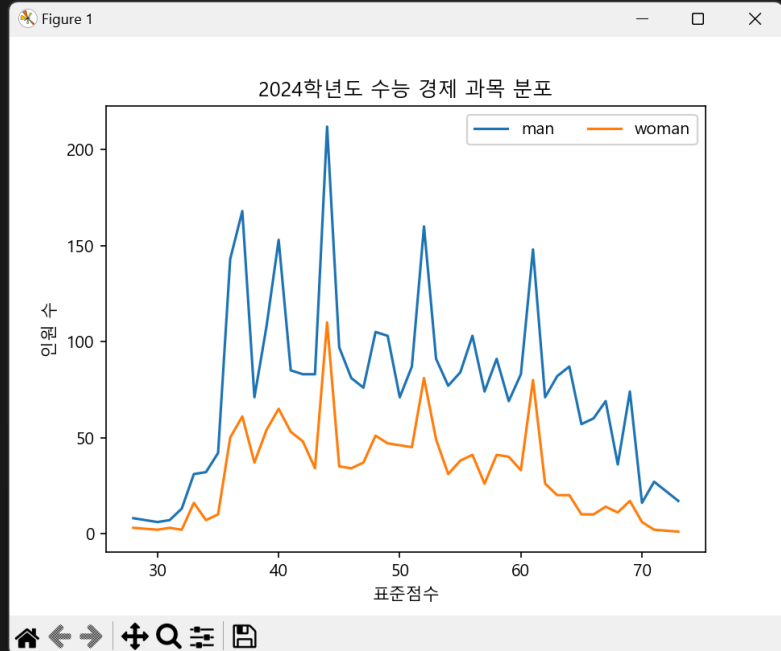
조회를 원하시는 세부과목을 선택하세요.

경제

선택된 과목 : 사회탐구 - 경제

위 과목의 표준점수 분포 그래프는 다음과 같습니다.

□



이 수능 데이터는 2024 년도 수능의 데이터입니다.

국어 : 국어

수학 : 수학

사회탐구 : 세계사

사회탐구 : 윤리와 사상

사회탐구 : 사회·문화

사회탐구 : 동아시아사

사회탐구 : 경제

사회탐구 : 세계지리

사회탐구 : 정치와 법

사회탐구 : 한국지리

사회탐구 : 생활과 윤리

과학탐구 : 화학 II

과학탐구 : 생명과학 II

과학탐구 : 화학 I

과학탐구 : 지구과학 II

과학탐구 : 지구과학 I

과학탐구 : 물리학 I

과학탐구 : 물리학 II

과학탐구 : 생명과학 I

직업탐구 : 공업 일반

직업탐구 : 농업 기초 기술

직업탐구 : 상업 경제

직업탐구 : 성공적인 직업생활

직업탐구 : 수산·해운 산업 기초

직업탐구 : 인간 발달

이 수능 데이터에 존재하는 과목은 위와 같습니다.

조회를 원하시는 세부과목을 선택하세요.

생명과학 II

선택된 과목 : 과학탐구 - 생명과학 II

위 과목의 표준점수 분포 그래프는 다음과 같습니다.

□

