

에타 강의 평점과 수강신청 경쟁률 사이의 관계

<숨겨진 ‘꿀강’ 찾기>

과 목 명	컴퓨팅 기초	팀 명	자율 4팀 너구리
지도교수	변해선 교수님	학 과 / 이 름	지구환경과학부 류영민, 황세웅



연구배경 및 필요성

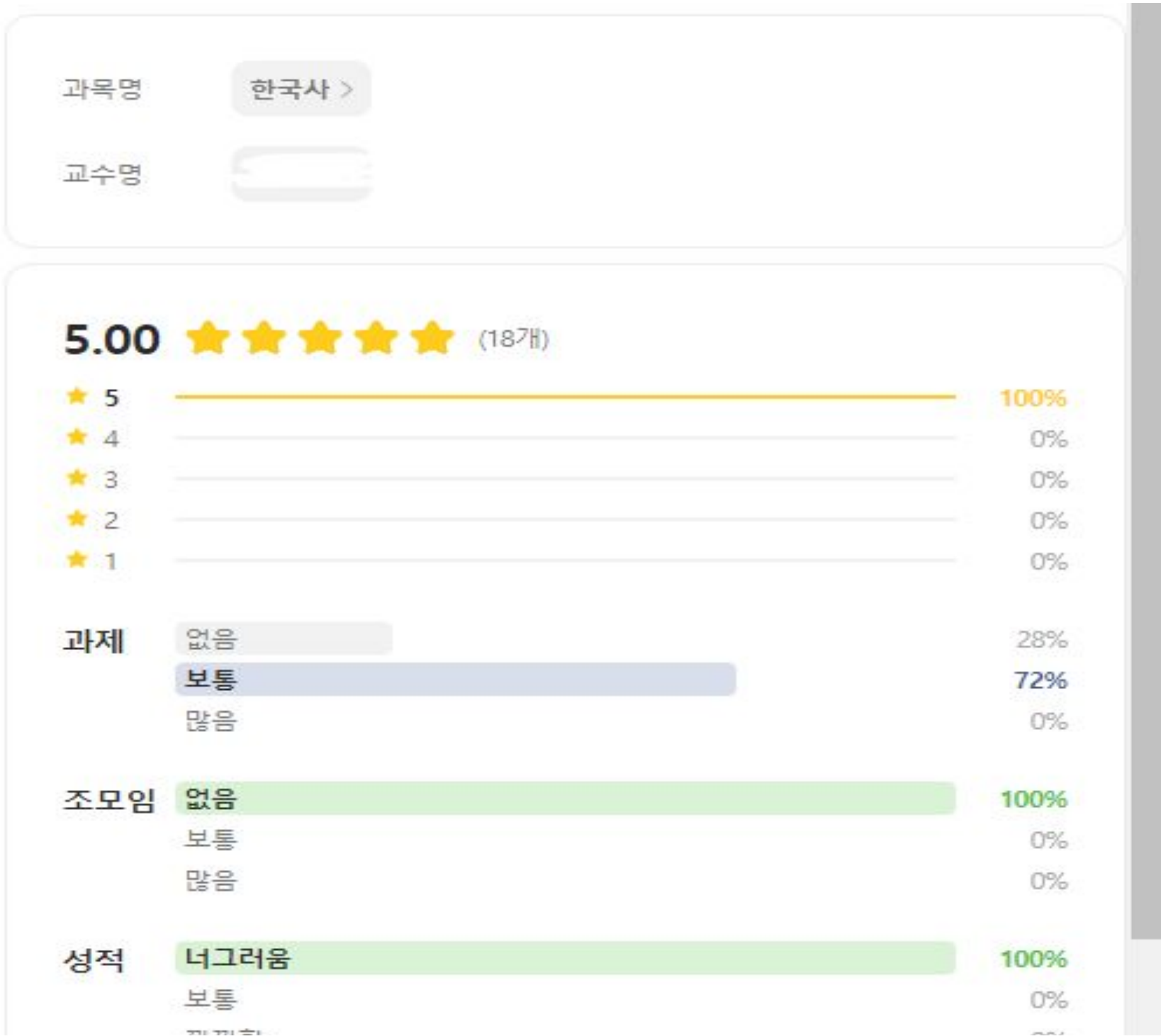
연구 목적

- 이번 프로젝트에서는 강의평이 좋은데 비해 그 경쟁률이 낮은 소위 ‘숨겨진 꿀강’을 찾는 것을 목적으로 한다.

<소주제> 숨겨진 ‘꿀강’의 기준

- 2023년 2학기(정규학기)에서 3학점 교양 중
- 에브리타임의 별점이 4.5 이상이면서
- 수강신청 사이트의 장바구니 경쟁률이 1.5 이하인 강의

ex) 한국사(고*진) – 강의평 5.00, 경쟁률 0.48



[학사] [교양] 한국사

고대진 | 역사학부(한국사학전공) | 043.007(005)

수강신청인원/정원(재학생) 44/60 (60) | 총수강인원 44 | 학점 3 | 월(17:00~18:15) 수(17:00~18:15)

취소여석



프로젝트 진행과정

- 웹 크롤링

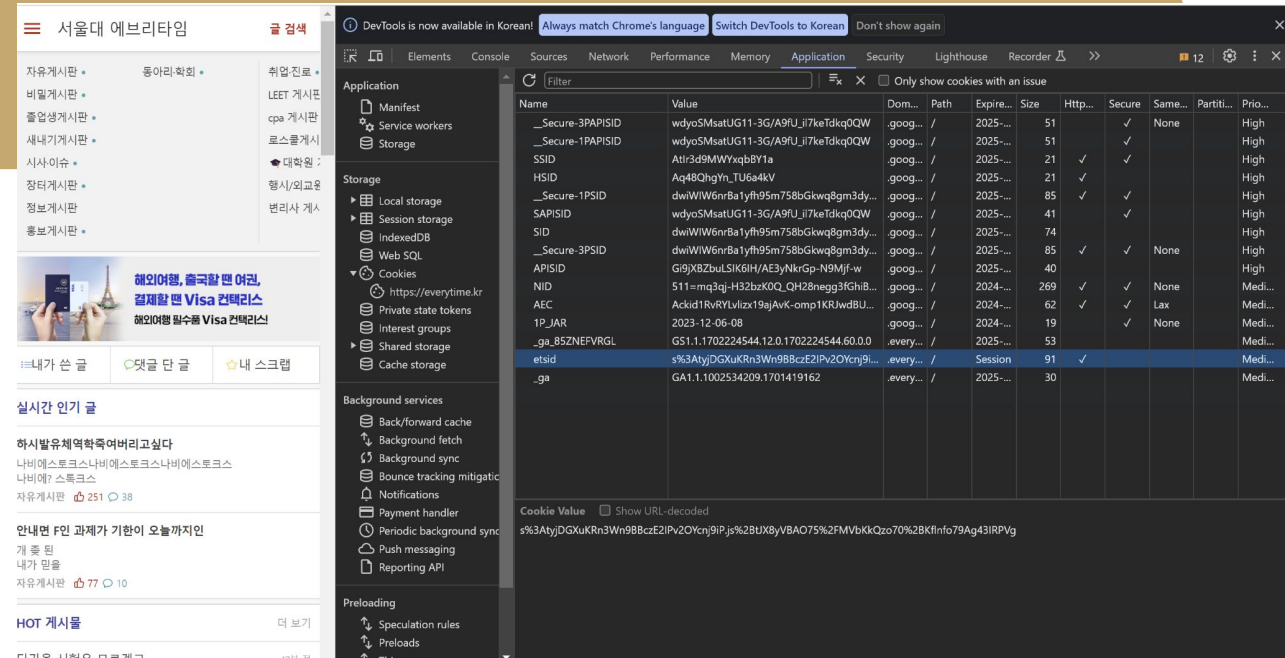
처음 selenium으로 크롤링 진행

-> 보안체계의 문제

-> 무한 스크롤링 과정에서의 실패

-> request & Beautifulsoup 이용

-> 데이터 프레임으로 저장 -> 확인을 위해 엑셀에 저장



과목명	과목코드	교수님	수강인원	장바구니	학점	과목명	과목코드	교수님	별점
환경과 건강047.025(00조경덕		52/100 (1034		3		언어와 인강0545.001€전영철		0	
현대국가와044.026(00최태현		50/70 (70) 73		3		고급한국어032.004 00홍은영		5	
경영학개론044.013(00서진영		8/80 (80) 19		3		중급한국어032.003 00이윤복		0	
고급영어: ₩0441.0013Nick Clem		13/20 (20) 6		3		중급한국어032.002 00전진호		0	
고급영어: ₩0441.0012박여선		18/20 (20) 24		3		고급한국어032.004 00이강혁		0	
고급영어: ₩0441.0014Nicholas V		17/20 (20) 18		3		한국인의 ₩041.004 00이은숙		5	
고급영어: ₩0441.0014윤수진		17/20 (20) 24		3		한글맞춤법041.018 00문숙영		4.39	
영어로 읽는041.028(00나은하		20/40 (40) 26		3		한국어 어학041.009 00최중원		0	
역사와 철학032.009(00김현경		5/30 (30) 7		3		21세기 한₩041.011 00이혜정		2.67	
동서문명의L0547.0004최소영		23/30 (30) 36		3		한국의 한자041.017 00홍진욱		5	
초급스페인032.048(00박은영		16/20 (20) 33		3		한국현대시041.010 00김유중		4.67	
초급스페인032.048(01박도란		17/20 (20) 45		3		한글맞춤법041.018 00배윤정		4.86	
초급스페인032.048(00강성식		15/20 (20) 29		3		드라마의 0042.002 00박상은		5	
중급스페인032.050(00조영실		16/20 (20) 23		3		한국의 신화042.004 00김준희		0	
스페인어권042.014(00박도란		30/40 (40) 46		3		창작의 세계042.003 00염승숙		0	
라틴아메리041.037(00김용훈		30/55 (55) 24		3		초급한국어L0441.001€김영규		0	
시사스페인032.054(00김선영		14/15 (15) 15		3		한국문학과L0545.000€최희진		4.43	
초급스페인032.048(00김희진		19/20 (20) 22		3		한국문학과L0545.000€정병설		2.6	
초급스페인032.048(00정수현		20/20 (20) 53		3		한국 문학L0545.0007윤병용		0	
스페인어권042.014(00장재준		20/38 (38) 16		3		미디어 리트0545.0014이후인		5	
이태리어 2032.058(00이기철		16/20 (20) 33		3		한국문학과L0545.0015김대중		0	
이중언어사L0545.0013양수경		45/60 (60) 49		3		초급한문 1032.005 00주기평		4.59	
히브리어 1L0441.0002권성달		13/20 (20) 21		3		초급한문 1032.005 00신지영		4.88	
몽골어 1 032.061(00이선아		10/20 (20) 9		3		초급한문 1032.005 00서용준		4.22	

- 데이터 시각화 전 사전 작업

엑셀 저장 이전 학점==3인 과목만 추출

-> 본 수강신청 인원 추출을 위해 수강인원에서 '/'와 '(' 사이에 있는 값 추출해서 수강인원 열에 다시 저장

-> 장바구니 열을 수강인원 열로 나누어 '경쟁률' column을 만들

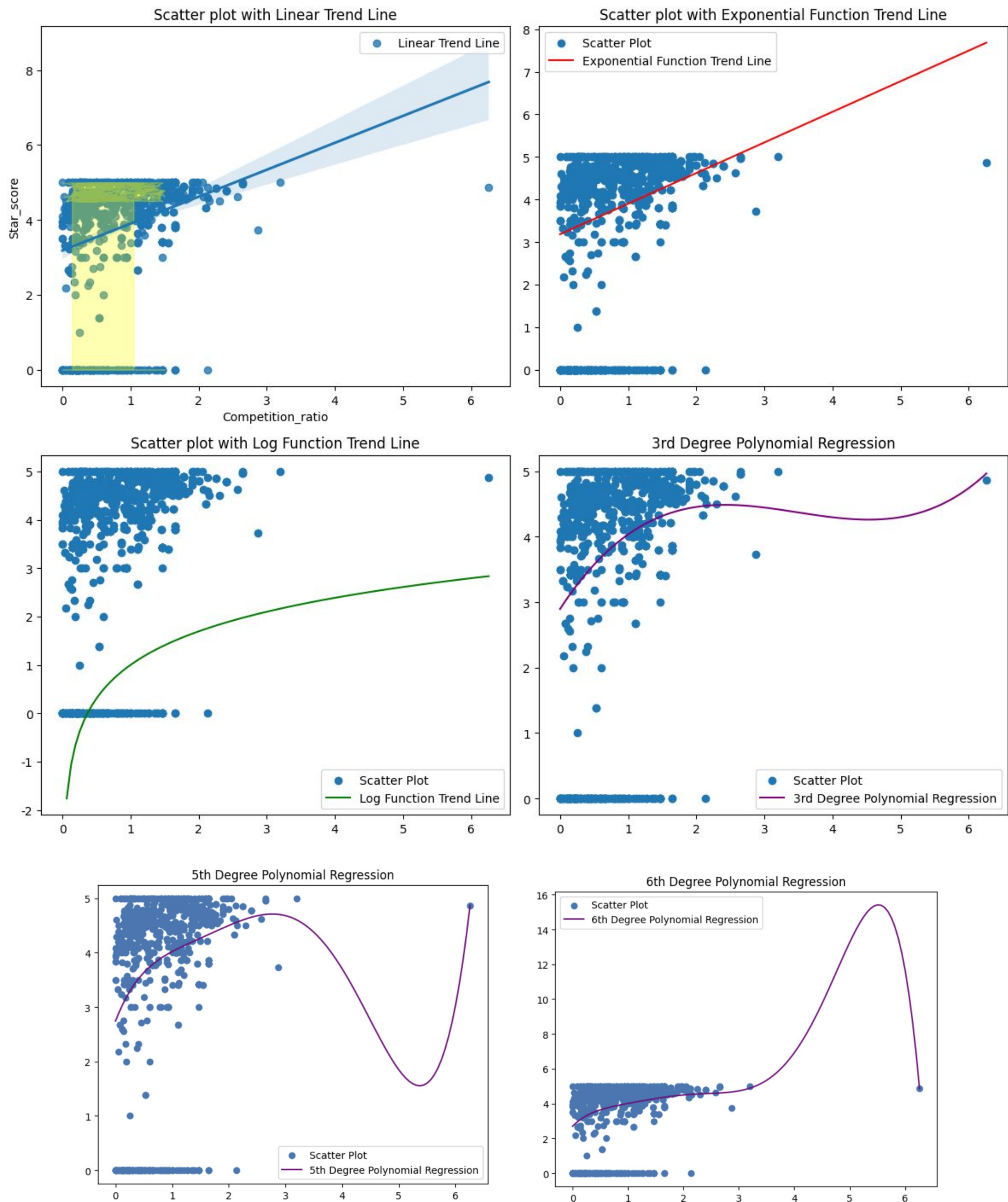
-> 두 데이터프레임 sugang_df와 everytime_df를 하나의 df로 합침.
merge함수와 how 매개변수 이용(how=inner)

- 데이터 출처 : 서울대학교 수강신청 사이트(2023-2) & everytime 서울대 강의평



분석 방법 및 프로젝트 성과

에타 강의 평점(y)과 수강신청 경쟁률(x) 사이의 상관관계(산점도)



프로젝트 기대효과

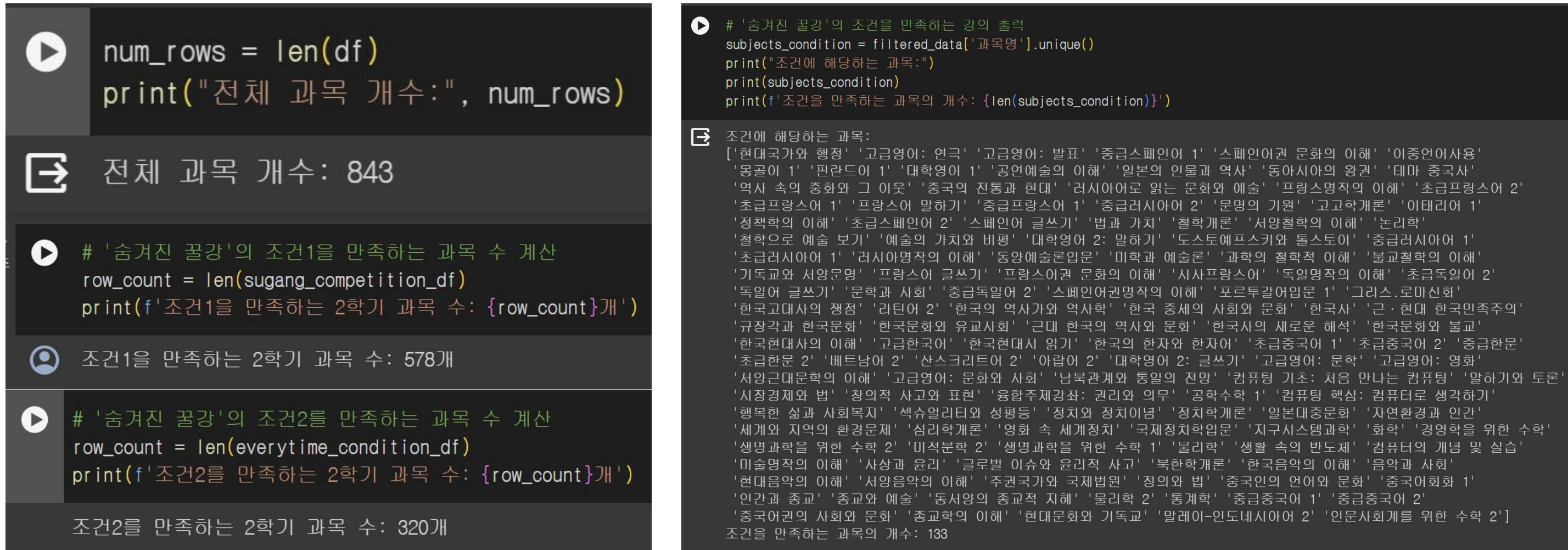
분석 결과, 에타 강의 평점과 수강신청 경쟁률의 관계는 선형, 로그, 지수, 다항식 중 5차 다항식으로 가장 잘 fitting된다.
또한, 숨겨진 ‘꿀강’의 조건을 만족하는 과목은 133개를 찾았다.

-기존 코드에서 조건 1만 만족했던 강의는 580개였다.

- 숨겨진 ‘꿀강’은 다음과 같다(대표강의 몇개).

현대국가와 행정, 지구시스템과학, 공연 예술의 이해, 스페인어권 문화의 이해...

-코드를 완성하고 찾아본 보완점은 다음과 같다.



1. 각 강의 정보에 들어가보면 과제, 조모임, 성적을 주는 정도가 나와있는데, 이 부분 또한 조사에 포함
→더 정확도 높은 결과
2. 강의평이 n(특정 숫자)개 이상인 강의들로 한정하여 웹크롤링을 진행
→더 신뢰도 높은 결과

기대 효과: 이 프로젝트를 통해 학생들이 숨겨진 꿀강을 알고, 이 강의들과 경쟁률 사이의 관계를 파악함으로써 내년의 수강신청 성공률을 높일 것으로 전망된다.