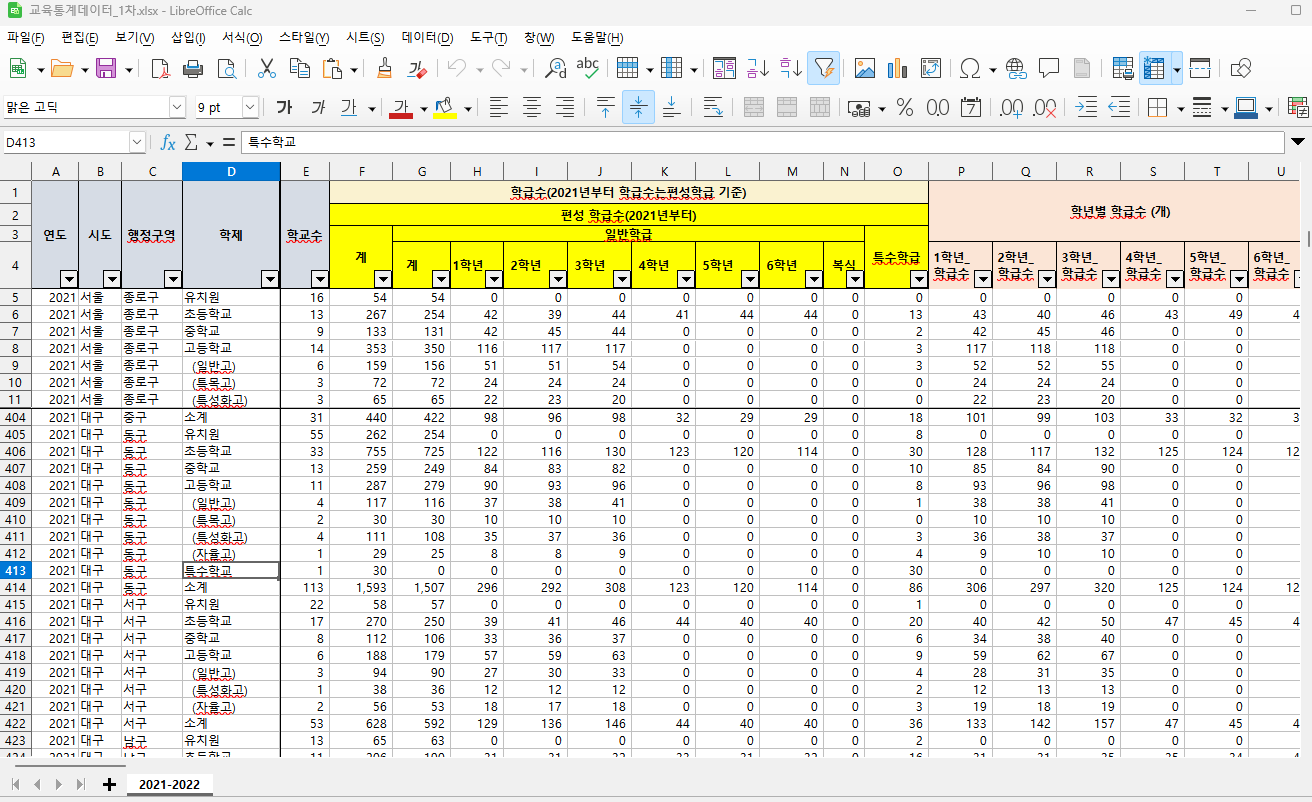
**전문상담 교사와 학업 중단자의 연관성**

빅데이터 4기

이예원 (개인보고서)

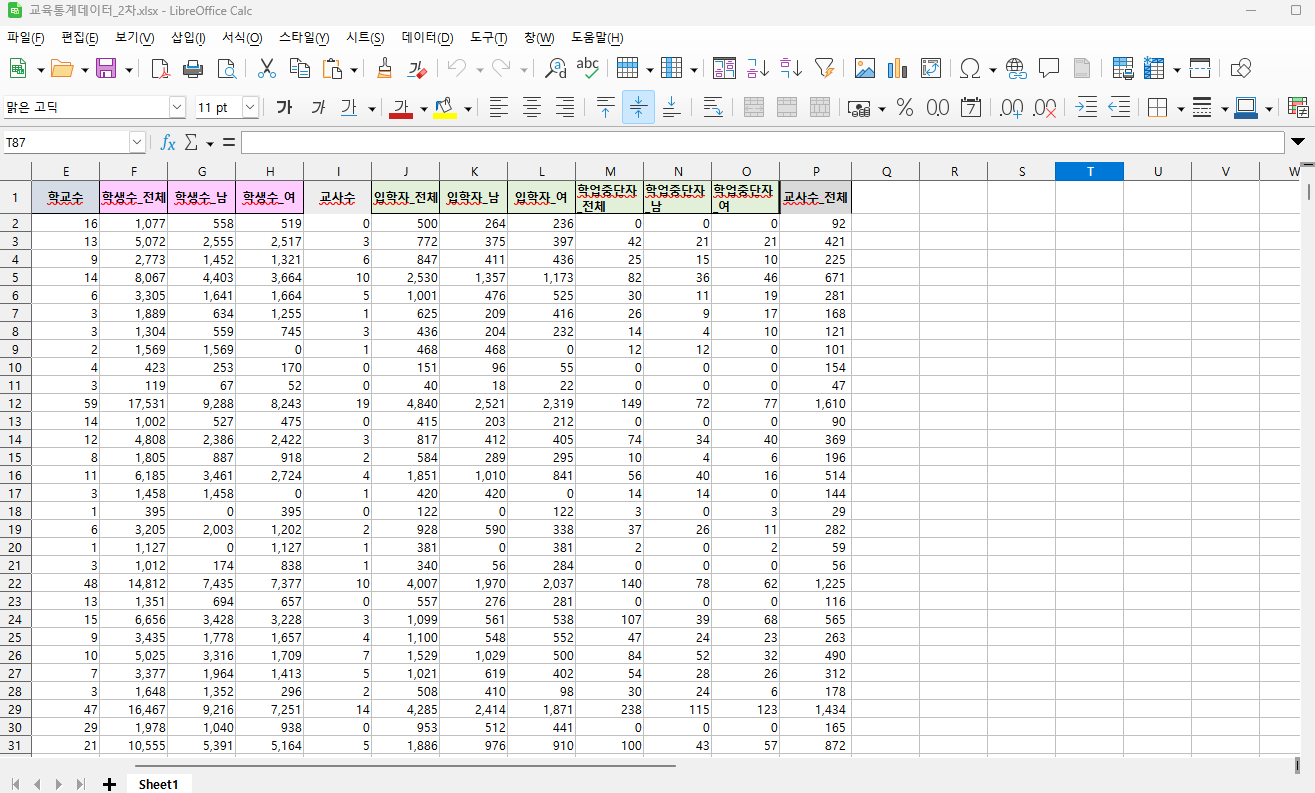
※ 팀 분석보고서는 팀원 ’김성민’님께서 올릴 예정입니다. 분석 코드는 하단에 github 주소로 올려놓았습니다. Input data는 같이 첨부하겠습니다.

* 서론:
  + 질문: 학교에 있는 전문상담 교사의 유무가 학업중단자의 수와 연관이 있을까?
  + 가설: 전문상담 교사의 수가 학교에 많을수록 학업중단자가 적을 것이다.
* 역할:
  + 데이터 전처리
  + 시각화 보조
* 방법:
  + input data = ‘시도별 행정구역별 교육통계 현황\_방통제외(1999-2022).xslx’
  + 전처리
  + 1차 전처리
    1. 원본 데이터에서 ‘2021-2022’ 이름의 2번째 시트만 추출
    2. ‘교육통계데이터\_1차.xlsx’로 저장



<’교육통계데이터\_1차.xlsx’>

* + 2차 전처리
    1. ‘교육통계데이터\_1차.xlsx’에서 [‘연도’, ‘시도’, ‘행정구역’, ‘학제’, ‘교사수 - 전문상담교사(정규)’, ‘학교수’, ‘학생수계 - 전체 ’, ’ ‘학생수계 - 남자’, ‘학생수계 - 여자’, ‘입학자 - 입학자\_계’, ‘입학자 - 입학자\_남’, ‘입학자 - 입학자\_여’, ‘학업중단자 - 학업중단자\_계’, ‘학업중단자 - 학업중단자\_남’, ‘학업중단자 - 학업중단자\_여’, ‘교사수\_전체체’]의 총 16개의 칼럼 정보를 새로운 시트에 불러오기
    2. ‘교육통계데이터\_2차.xlsx’로 저장



<‘교육통계데이터\_2차.xlsx’>

* + 3차 전처리
    1. pandas 패키지 불러오기
    2. ‘교육통계데이터\_2차.xlsx’ 데이터 불러오기
    3. ‘연도’가 2022년인 데이터를 df2로 마스킹
    4. 모든 칼럼에 대해서 unique 함수를 이용하여 데이터 형태를 확인하기
    5. ‘학제’ 칼럼에서 데이터명 통일화 하기
    6. 통일화된 데이터를 다시 ‘학제’ 칼럼에 저장하기
       - replace 함수를 이용하여 ‘()’ 형태 삭제하기
       - ‘학제’ 칼럼 전체에 적용할 수 있도록 school\_convert(x) 함수 제작하기
    7. ‘학제’ 칼럼에서 '고등학교' ,'일반고', '특목고', '특성화고' 만 마스킹한 데이터를 d3에 저장하기
    8. ‘연도’ 칼럼 삭제하기
    9. index를 지우고 to\_excel을 통해 "교육통계데이터\_2022\_최종\_교사수전체포함함.xlsx" 으로 데이터 내보내기

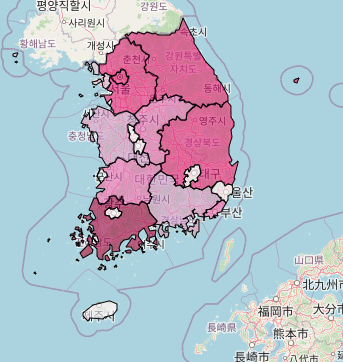


<최종 데이터셋>

|  |  |  |
| --- | --- | --- |
| 칼럼명 | 자료형 | 용어정의 |
| 시도 | object | 전체 17개의 시도 |
| 행정구역 | object | 각 시도의 207개의 행정구역 |
| 학제 | object | 고등학교 (일반고/ 특목고/ 특성화고/통합) 구분 |
| 학교수 | int | 전체 학교수 |
| 학생수\_전체 | int | 전체 학생수 |
| 학생수\_남 | int | 전체 남학생 수 |
| 학생수\_여 | int | 전체 여학생 수 |
| 교사수 | int | 전체 전문상담교사수  (전문상담교사; 전문상담 자격증을 가진 교원으로 현재 근무하는 학교에 전문상담교사로 발령받은 교원) |
| 교사수\_전체 | int | 전체 정규교원수  (교장, 교감, 수석교사, 보직교사, 일반교사, 특수교사, 전문상담교사, 사서교사, 실기교사, 보건교사, 영양교사) |
| 입학자\_전체 | int | 전체 입학생수 |
| 입학자\_남 | int | 전체 남자 입학생 수 |
| 입학자\_여 | int | 전체 여자 입학생 수 |
| 학업중단자\_전체 | int | 전체 학업중단자 수 |
| 학업중단자\_남 | int | 전체 남자 학업중단자 수 |
| 학업중단자\_여 | int | 전체 여자 학업중단자 수 |

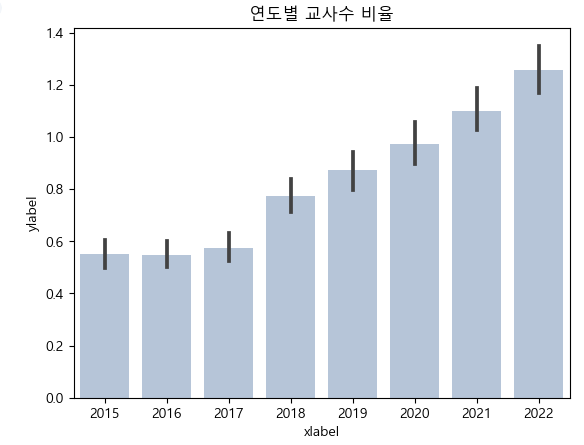
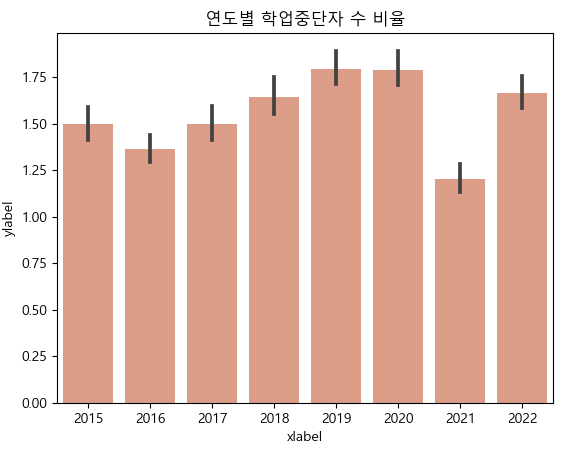
<최종 칼럼 정보보>

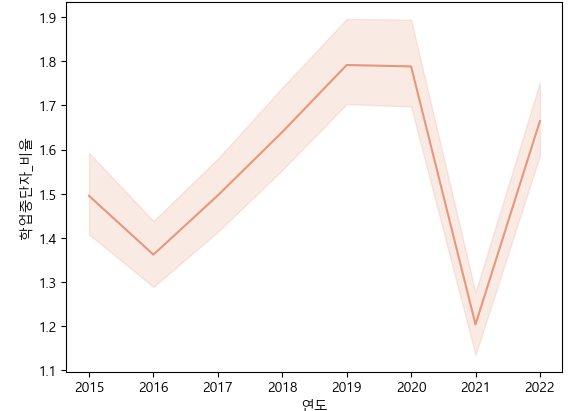
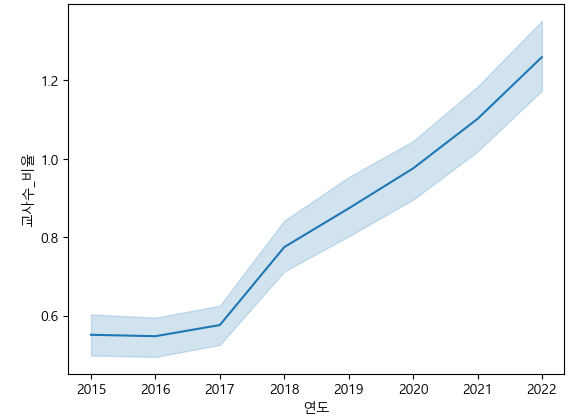
* + 시각화-보조
  + 시도별 상담교사수 비율 그래프
    1. 데이터 불러오기
    2. ‘학제’ 칼럼에서 ‘고등학교’만 마스킹
    3. 서울 -> 서울특별시, 부산->부산광역시 와 같이 ‘시도’ 형태를 바꿔서 ‘시도\_new’ 칼럼에 저장
    4. ‘교사수’를 ‘교사수\_전체’로 나눈 뒤 100을 곱해서 ‘교사수\_비율’ 칼럼에 저장
    5. ‘시도\_new’ 칼럼을 기준으로 ‘교사수\_비율’의 합계를 구하여 df\_teacher 변수에 저장
    6. folium 패키지와 json 패키지를 import 하기기
    7. 시도별 json 데이터를 이용하여 geo data를 불러오기
    8. Choropleth 맵을 이용하여 시도별 상담교사수 비율을 시각화 하기



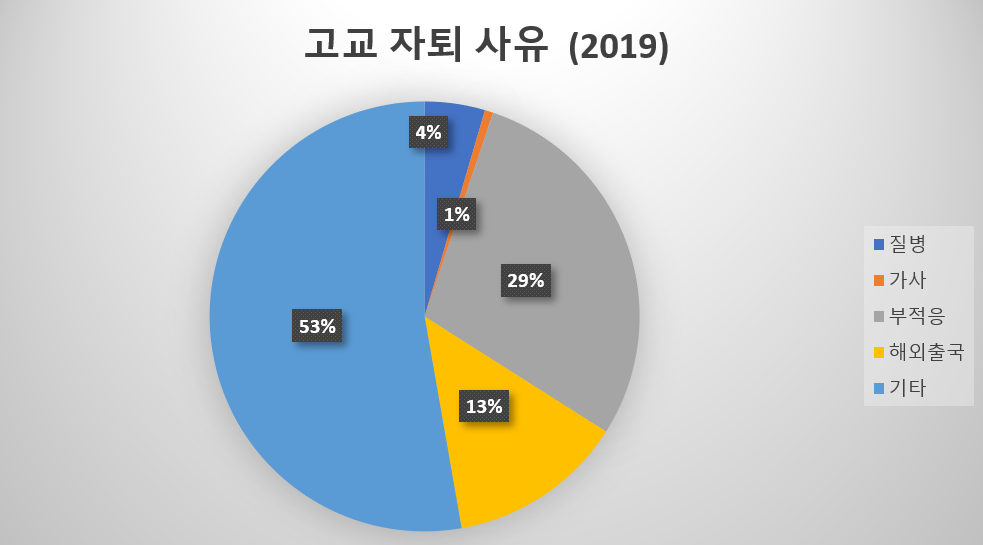
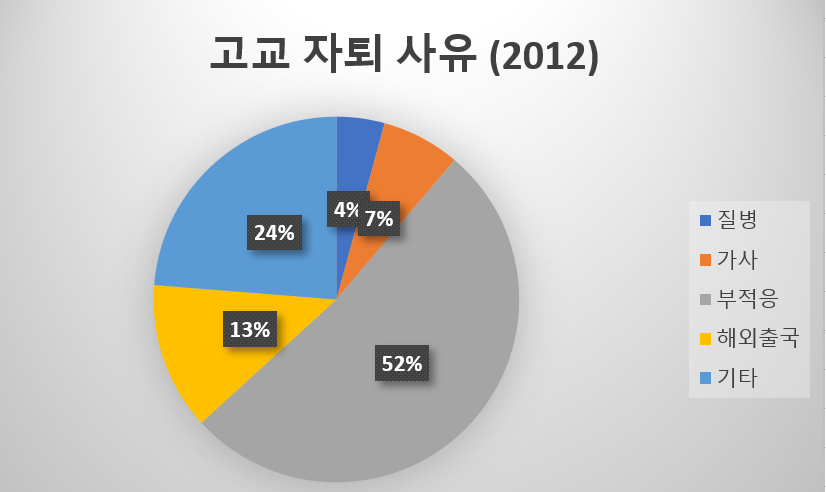
<시도별 상담교사수 비율 지도 시각화화>

* + 연도별 학업중단자 비율 및 상담교사 비율 그래프
    1. 데이터 불러오기
    2. 2021-2022년 데이터와 2015-2020년 데이터 합치기
    3. 고등학교만 마스킹하기
    4. 학업중단자 비율 칼럼 생성하기
    5. 교사수 비율 칼럼 생성하기
    6. 연도별 학업중단자 비율, 교사수 비율을 막대그래프와 선그래프로 그려보기기



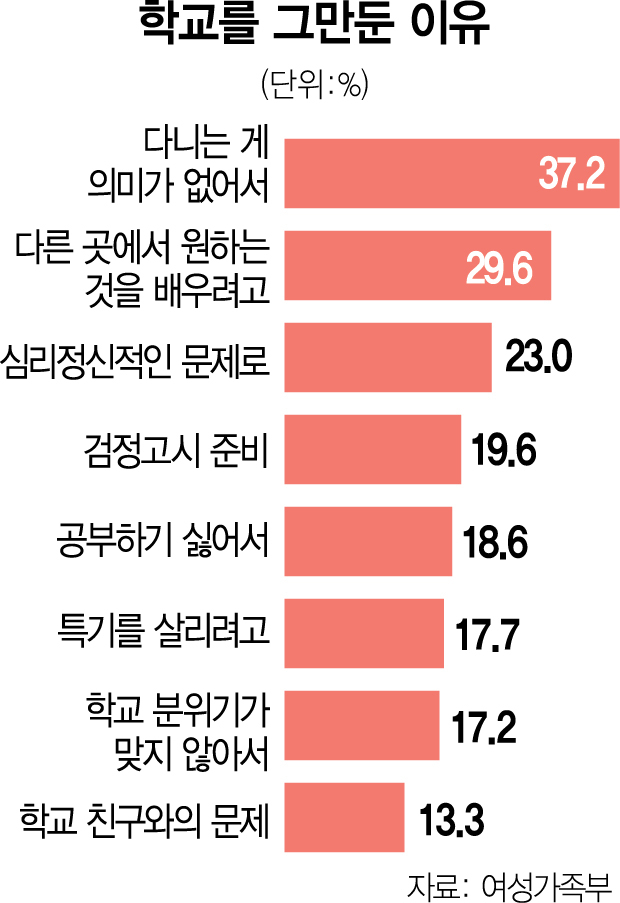
 

* 결론:
  + 피어슨 상관관계 분석 결과 p>0.05 으로 상담교사와 학업중단자 사이에 상관관계가 없음으로 나온다.
  + 이에 두 변수가 정말 서로 무의미한 변수인지 살펴보기로 한다.
    1. 기사 살펴보기: 기사를 살펴본 결과 2009년부터 Wee프로젝트(학생 위기상담·지원 사업)를 실시한 것으로 보인다. 이에 따라 상담교사는 학업중단자의 수와 상관없이 매년 늘어난 것으로 보인다.
    2. 연도별 그래프 살펴보기: 2015~2022까지의 연도별 그래프를 살펴보니 실제로 학업중단자의 수와 상관없이 교사수는 계속해서 늘어나고 있는 것으로 보인다. 특히, 2021년에 코로나로 인해 학교에서 원격수업을 진행함으로써 학업중단자의 수가 줄어든 것으로 보이는데 그와 상관없이 상담교사 수는 위 정책으로 계속해서 늘어나는 것을 볼 수 있다.
  + 위 근거들로 보아 학업중단자와 상담교사수의 통계적인 상관관계는 없는 것으로 보인다. 이에 우리의 제안은 학업중단자를 줄이기 위해 무작위적으로 상담교사를 늘리지말고 학업중단자에게 보다 실질적인 도움이 될 수 있는 방안을 생각해보기로 하고 이를 제안한다.
* 제안:
  + 2012년과 2019년 고교 자퇴 사유 비교

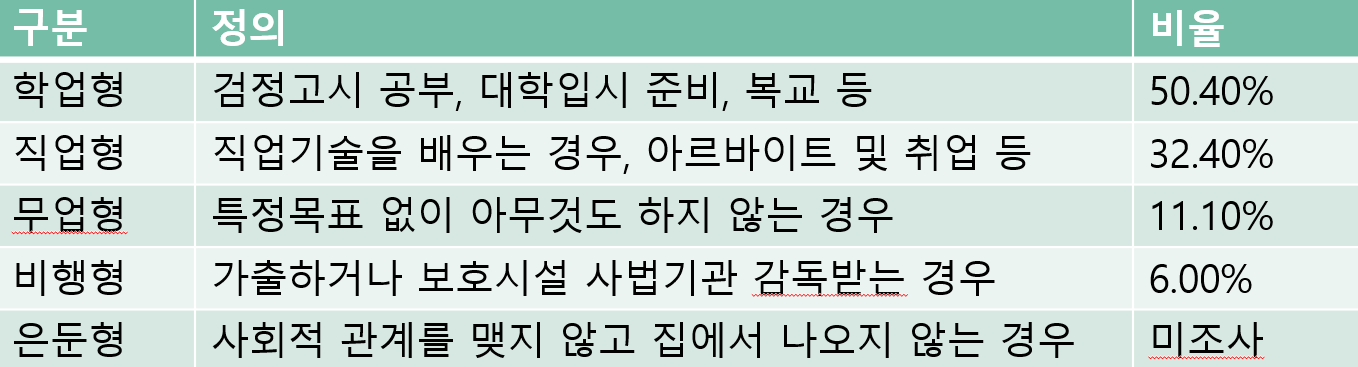


2019년 자료를 살펴보면 2012년 자료에서는 52%를 차지했던 부적응이 29%로 줄고 24%였던 기타는 53%로 늘었다. 과거와 달리 최근에는 부적응으로 인한 자퇴보다 기타사유가 많다.

* + 학업중단 상세 사유



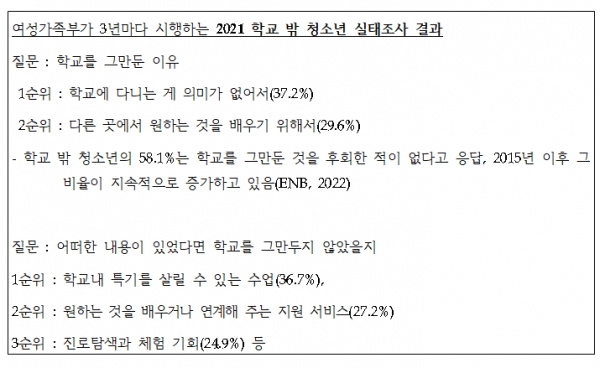
* + 자퇴 후 학생들의 현황



윤철경, 성윤숙, 유성렬, 김강호(2015)

위의 자료들을을 보면, 학교를 그만둔 이유가 ‘다니는 게 의미가 없어서’라는 이유가 대부분이며 자퇴 후 학생들의 현황을 살펴보면 전체의 반은 학업을 준비한다. 이는 학생들이 학교를 다니는 이유에 대해 확립이 되어있지 않는다고 볼 수 있다. 그저 다니라고 해서 다니는 학교에 불가할 뿐이다. 또한, 더 이상 자퇴는 부적응의 회피 수단이 아닌 또 다른 입시 전략 중 하나라고 볼 수 있다. 학교에서 다른 과목이나 수행평가 등으로 시간을 쓰는 대신 응시 과목에 집중하는 것이다.

* + 학교 밖 청소년 실태조사 결과(2021)



2021년에 시행한 학교 밖 청소년 실태조사 결과, 어떠한 내용이 있었다면 학교를 그만두지 않았을지에 대한 질문에 대부분의 학생이 자신의 특기와 진로탐색에 대한 수업 및 프로그램 지원을 원했다.

기존에는 학업 중단 예방 정책으로 학업중단 숙려제, 학교(Wee 클래스 전문상담 교사) 또는 Wee센터, 청소년상담복지센터 등의 기관들에서 상담 등의 심리 상담 위주의 해결책을 제시했다.

하지만, 현재는 이러한 정책들이 효과가 없는 것으로 나타났다. 따라서 직업프로그램, 고교학점제 등과 같이 학생들 개별 특성을 고려한 진로 관련 프로그램이 만들어져야 한다.

특히, 자신이 원하는 공부(과목)을 배울 수 있는 교육인 고교학점제도는 학업 중단 문제를 개선할 수 있다는 기대감을 가진다. 소수의 학생들의 의견과 다양한 학생들의 입장을 생각하여 학교에서 학생이 원하는 것을 배울 수 있도록 제도가 구체화 되어야 한다.

* 코드 참조
  + <https://github.com/ye-bong/project.git>
* 출처
  + <http://www.sobilife.com/news/articleView.html?idxno=35135>
  + <https://www.sedaily.com/NewsView/26606IHXZQ>