

## 5/1 회의 정리록

데이터 전처리  
필요없는 열 삭제  
: 역 ID, 호선

연도별을 무시하고, 월별부터 정렬을 하기  
-> 개강, 방학 등의 사유로 특정 월의 차이가 있을것으로 예상

승차원 총 104315 - 하차원은 108856  
탄 사람 보다 하차한 사람이 더 많다  
-> 버스 환승 후 지하철 탄 사람의 데이터는 누락되기 때문에 승차원의 합이 적은 것으로 예상

몰리는 역 위치의 이유 찾기  
-> 기업 위치, 역의 시설 / 위치 데이터 찾기  
-> 서울시 행정동 단위 거주인구 데이터를 가지고 지역별로 인구 확인해서 이 역에서 왜 사람이 많이 타는지