A 포스터 피드백

B 지금 데이터 병합 문제

1. 상권데이터들
   1. 관악구 신사동, 강남구 신사동 데이터를 새로 만들어야함.
   2. 상권 최종데이터 (중복되는 구, 동 데이터 지우기, 결측치 많아서 (제외 또는 평균대체), 지하철역, 정류장의 경우 0개인지, null인지 파악 ( 1 데이터> 전처리 데이터> 최종병합데이터)
2. 음주 사고 데이터
   1. 구별로 데이터 병합 필요.
   2. 사고 건수 데이터(Y)만 상권데이터(Xs)와 병합하기
   3. 음주운전 신상정보 데이터를 요일과 시간대로 추출 가능한지 확인

C 이후 분석과정

1. 상관분석 : 사고 건수에 따른 상권 요인들 확인
2. 회귀분석 : 각 상권데이터만으로 사고율 예측 가능한지
3. 군집분석 : (사고가 높은 것끼리 각 지역별 특성을 분류 or 전체 범주에서 사고가 자주 발생하는 지역과 아닌지역 분류)
4. 시뮬레이션: ( 정책 시행 전후의 사고 발생률을 예측하는 몬테카를로 시뮬레이션.. 사용할지 고민중)

분석 흐름과 유사한 샘플

(원인 분석후 -> 특징 추출 -> 맞춤형 제안)

<https://youtu.be/KufWqQJUVRE?si=ROsx7MxNAtazQ6gf>