[골드문 8월 팀 프로젝트 회의록]

광양시 재활용품 자동수거기 최적입지 선정 과제

0809)

현재 진행 상황: 연구 목표 재설정 및 분석 방법론 결정

데이터를 보고 분석 방법을 떠올리는 것이 어려움 => 역으로 먼저 어떤 입지를 우선순위로 둘지 선정하고, 생각해 두었던 대로 결과를 만들어 내려면 어떻게 분석해야 할지 정하기.

1. 어떤 입지 선정해야 하는가?
2. “공동주택 단지 내(분석조건에 명시)” => “아파트” 위주로 선정해야 함.
3. 인구수가 많은 곳
4. 연령대 => 생산가능인구가 많은 곳.(쓰레기 배출량과 상관관계 확인)
5. 주거 형태 => 세대당 인구가 적은 곳.(가구가 많아질수록 쓰레기 배출량은 많아질 것.)
6. 쓰레기 배출량이 많은 곳.
7. 향후 추가 필요하면 추가.
8. 분석 방법은 어떤 방법을 취할 수 있는가?
9. 지표 개발은 어떻게 해야 할 것인가?
10. 세대당 인구수
11. 인구수
12. 세대 수
13. 쓰레기 배출량
14. 생산인구수 (연령대)
15. 성별 인구수 및 비율

https://jumin.mois.go.kr/#

문제 발생) 100x100 격자로 되어 있는 데이터가 인구수밖에 없는데다, 동별 쓰레기 발생량 데이터도 없는 상황. 이런 상황에서 배출량 예측을 어떻게 해서 각 격자에 특징을 매칭할지 모르겠음.

8월 10일 회의록

일단 쓰레기 발생량이 확실히 나온 것은 금호동, 광양시 전체

=>그럼 1차 회귀식을 광양시 전체로 보는 것이 맞는가?

금호동) 인구, 세대수, 성별 인구수, 쓰레기 발생량

주요 지표를 최대한 “인구사회학적” 특성으로 한정시키는 게 좋을 듯.

1차 회귀식 만들기 => 어차피 연구 목적은 “주거단지에서 나오는” 재활용품 수거 효율 높이는 것이 목적. 그렇다면 기타 거주 형태나 상가 건물, 거주지에서 나오는 것이 아닌 오피스/왔다갔다 지나가다 버리는 것/일회성 방문시 버리는 쓰레기 등등을 포함하지 않는 것으로 하자.

그렇다면 “아파트 재활용품 발생량”을 “해당 행정동 전체에서 나오는 재활용품 발생량”으로 볼 수 있음.

* 동별 인구사회학적 특성을 반영하여 회귀식 쓸 수 있음
* 이제 회귀식에 들어가는 독립변수 특징들을 100x100 격자에 매핑할 수만 있다면 격자별 쓰레기 발생량 예측해서 채울 수 있음.
* 마트나 편의점은 포함시키지 않기로.(왜냐면 소비활동이 일어나는 지역과 실제로 제품을 소비하는 지역(쓰레기 발생이 되는 지역)이 다르니까.)

100X100 격자에 어떻게 매핑하나요 그거 안 되면 어떡하나요 => 그럼 오늘까지 방법론 찾아보고 안 되면 다른 접근법 찾아봅시다

필요한 공부)

보간(크리깅) 방법 – 구글링 키워드: 격자별 인구추정 등..

추가 데이터 탐색)

전국 건물정보(?)

광양시 행정동 연령대별인구

<질문드렸던거 답변>

안녕하세요,

광양시 분석 요건에 보면

"공동주택 단지 내에 설치가 되어야 한다"라고 되어 있는데,

그러면 공원이나 기타 공공시설 등지는 설치 위치로 선정할 수 없나요?

감사합니다.

대상지는 공동주택 단지내에서 선정하여야 합니다.  
  
공원 등 기타 공공시설이 단지 내라면 가능하지만  
  
단지 외라하면 조건에 부합하지 않습니다.

2번데이터(2.광양시\_재활용품분리수거\_장려금지급내역.csv) 활용

1. 각 쓰레기를 배출량(kg)으로 환산 > Python에서 구현예정

* 장려금 지급내역의 단위 : 유리병(kg당 70원), 페트병(kg당 40원), 플라스틱류(kg당 30원), 폐형광등(kg당 150원), 필름류포장재비닐(kg당 200원), 폐건전지류(kg당 600원), 종이팩(kg당 200원), 소형가전(kg당 300원), 휴대폰(kg당 1000원)
* 금호동 재활용품 배출 통계 단위 : 캔(kg), 플라스틱(kg), 병(kg), 폐지(톤), 고철(kg)

1. 아파트명 획일된 기준으로 통일 > csv파일 수정 完

도움) 100X100 격자 보간방법 환희가 도와줌. 밥먹고 나서 다같이 보기.

EDA 초석

1. 분리수거장현황 시각화
2. 장려금 지급내역(아파트별 / 동별 평균), 금호동 재활용품통계 시각화
3. 인구데이터 시각화(연령대별, 성별별인구, 동별인구, 등등)
4. 폐기물 발생 및 처리 현황
5. 건물 용도별 시각화 > 아파트(주거단지)만 따로 강조

**분석 방법 최종 결정(1차)**

“아파트” 단위로 분석.(이유: 설치 위치는 반드시 “공동주택 단지 내” 이므로, 아파트가 아니면 설치할 근거가 있는 대상이 별로 없음.)

100X100 격자보다 매핑이 효율적일것으로 예상.

아파트별로

1. 세대당 인구수
2. 인구수
3. 세대 수
4. 쓰레기 배출량
5. 생산인구수 (연령대)
6. 성별 인구수 및 비율

데이터를 매핑.

세대 수는 이미 있고, 인구 수는 외부데이터(15.광양시\_건물정보.geojson) “건물연면적”의 정보를 활용하여, “1인당 주거면적(행정동의 “주거용도” 건물 총 면적/행정동별 총 인구)”을 구하여, 아파트 연면적을 통해 총 아파트 주거 인구 보간. 동일한 방식을 사용하여 나머지 피쳐를 구한 다음, 최종적으로 쓰레기 배출량을 예측함.

⇒ 2번 데이터(2.광양시\_재활용품분리수거\_장려금지급내역.csv) 수집이 균일하지 않게 되어 있으므로, 각 재활용품별 장려금을 배출량(kg)으로 환산한 뒤, “1가구당 쓰레기 배출량이 너무 적은 곳”, “표제부에는 있지만 배출량은 없는 곳”의 쓰레기 배출량을 예측하여 보간합니다.

마지막으로 완성된 데이터를 바탕으로 군집화(k-means 클러스터링) 후, 각 군집별 특성을 보고, 어떤 군집이 최적 군집인지 선택하여 재활용품 자동수거기 50개소의 위치를 선정합니다.

((선정된 위치들에 재활용품 수거기를 설치할 때, 반경 300m내에 3개 이상의 수거기가 겹치지 않도록 조정합니다.)) => 이건 우리 능력으로 가능하면 시도.