

210911 멘토링

📅 Date	@2021년 9월 10일
📋 Purpose	Template

- 발표자료

https://s3-us-west-2.amazonaws.com/secure.notion-static.com/3b4a2b4c-cb9e-4aee-9831-9057d9a432f9/4조_할수있조_DS_210911.pptx

https://s3-us-west-2.amazonaws.com/secure.notion-static.com/70966f1f-5c0e-4901-af49-f50fb137d99b/4조_할수있조_DE_210911.pptx

▼ 대본

p.4

정부는 수소차 보급 확산을 위해 수소충전소를 2022년에는 310기 2030년에는 660기 까지 구축할 계획을 가지고 있음

그러나 현재 수소차 충전소 부족 문제로 인해 수소차 확산이 지연되고 있음

수소차 보급확산 목표를 위한 효율적인 수소충전소 입지 분석을 수행하고자 함

p.6

저희는 MCLP(최대커버링 입지모델) 적용하기 위해 먼저 변수별 가중치를 선형회귀로 산정하고자 하였습니다.

저희는 X피쳐로 인구현황, 자동차 등록대수로 하였고

Y 타겟 변수는 앞서 설명한 바와 같이 수소충전소가 서울 내에 4개까지만 설치되어 있어 기존 LPG 충전소를 기반으로 회귀 분석을 수행하고자 합니다.

선형회귀를 통해 데이터를 기반으로 지역특성을 충전소에 맞게 적절한 가중치 를 부여 하여 객관성을 부여하고자 하였습니다.

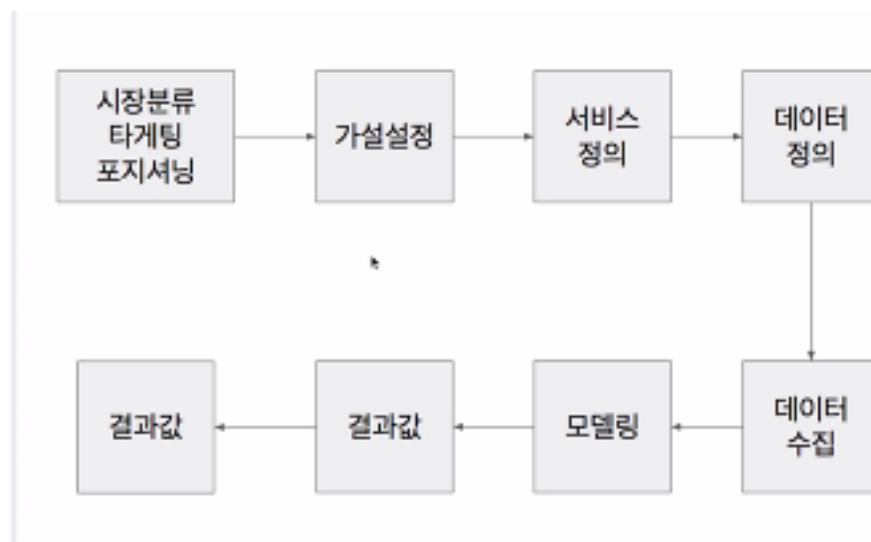
이를 위해 1kmx1km 격자로 나누어져 있는 인구현황 데이터에서 각 격자의 가운데 중심점을 찍고, 격자 중심점에 정규화된 인구 수 정보를 할당하였습니다.

이후 서울시 동별 자동차등록현황을 활용하여 각 지역에 해당하는 point를 모두 추출 후, 인구격자 중심점에 자동차등록대수 정보를 할당하였습니다.

이후 타겟 변수인 LPG 충전소를 할당하고 회귀분석을 수행할 예정입니다.

MCLP 를 적용하기 전

먼저 법적 규제에 맞추어 1종 보호시설인 학교, 유치원과 수소충전소 위치를 제거하고 MCLP(Maximal Covering Location Problems)는 최대커버링 입지모델을 적용할 예정입니다.



1. 데이터 사이언스와 데이터 엔지니어링분들이 같이 공유할 수 있는 공통 정리 부분을 만들어 주었으면 좋겠다.

(6조의 그래프 같은 느낌인거 같아요. >> 추측)

2. 데이터에 대한 구체적인 수치를 이야기를 해주는 것이 좋음.

데이터를 "어디서 몇 건"을 수집하였으며 "몇 개의"변수를 사용하였느냐.

(00데이터에서 00개의 자료를 가져왔고, 00개의 변수를 뽑아서 정리하였다.)

3. 그래프



4. 주제 선정 이유? 서비스 대상 정부기관/공공기관으로 선정 이유?

변수선정이유? 기존에 있는 서비스를 다시 구축하려고 하는 이유?

—> 이런 질문들이 많이 들어올 것임. 이에 대비하는 부분이 중요함.

주제, 가설, 데이터, 데이터 가져오는 이유, 모델 등등의 이유를 “팀원 모두”가 설명할 수 있어야 한다.

= 웹페이지 만드는 것 혹은 모델을 직접 돌리는 것도 중요하지만 “왜??”가 정말 중요하다. “왜??”에 팀원 모두가 합의하고 모두 명확히 설명할 수 있는 부분이 필요하다.

5. 뉴스나 주식을 넣으려고 하는 이유를 명확히 해야 한다.

6. 입지선정 옵션(LPG, 1종 보호시설 등)의 설정 이유를 다시 생각해야 된다.

연구논문이 아니라 서비스 제공임을 생각하고 실제 사용자가 필요한 옵션을 고려해야 된다.

위험지역이 과연 실제 사용자의 관심사일까? 1종보호시설을 피하는 것은 당연한 것

• Q&A

Q) 더 추가하면 좋을 거 같은 아이디어는 뭐가 있나요??

A) 서비스를 더 구축하기 보다는 지금 진행하고 있는 것들을 더 깊게 들어가는 것을 추천한다.

우선 기간이 짧기도 하고 면접 시에 왜 이런 서비스를 제공하였냐라는 질문을 꼭 받을 것이다.

김상인 멘토님(DS):서비스를 구축하기 보다

1. 모든 팀원들이 이 서비스를 이해하였는가?
2. 그리고 이런 서비스 혹은 모델을 왜 사용하였는가? 라는 논리적 이유를 명확히 할 필요가 있다.

조태원 멘토님(DE):새로운 것을 구축하지 마시고 1번(입지선정 옵션)에 더욱 집중하는 것을 추천한다.

뉴스,주식 보다는 입지선정 옵션에 더욱 집중하자.

추천하는 옵션은 서울시에서 강북/강남 을 나눠보는 건 어떨까?

강북에 30% 강남에 70% 을 넣어보기 이런 식이 어떨까? K=1~30까지 모두 보여주자.

상위 랭킹 3개구의 수소충전소 개수 및 비율 제시

(슬라이드 바(개수 좌우)를 조절하여 최적입지 제시 되도록 → 태블로로 구현 가능)