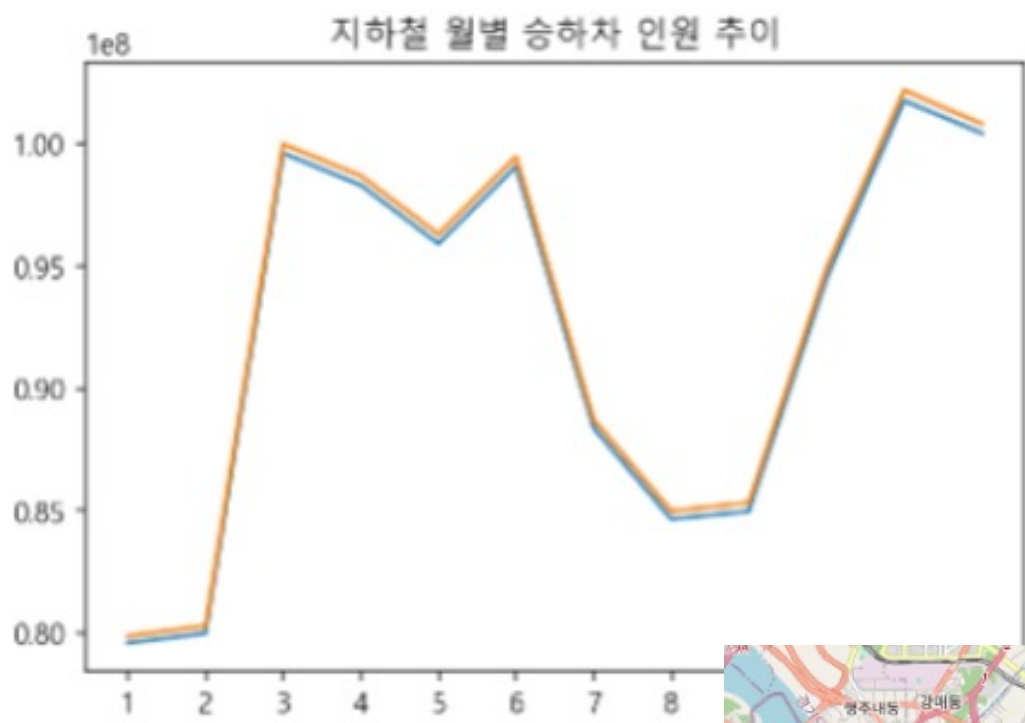
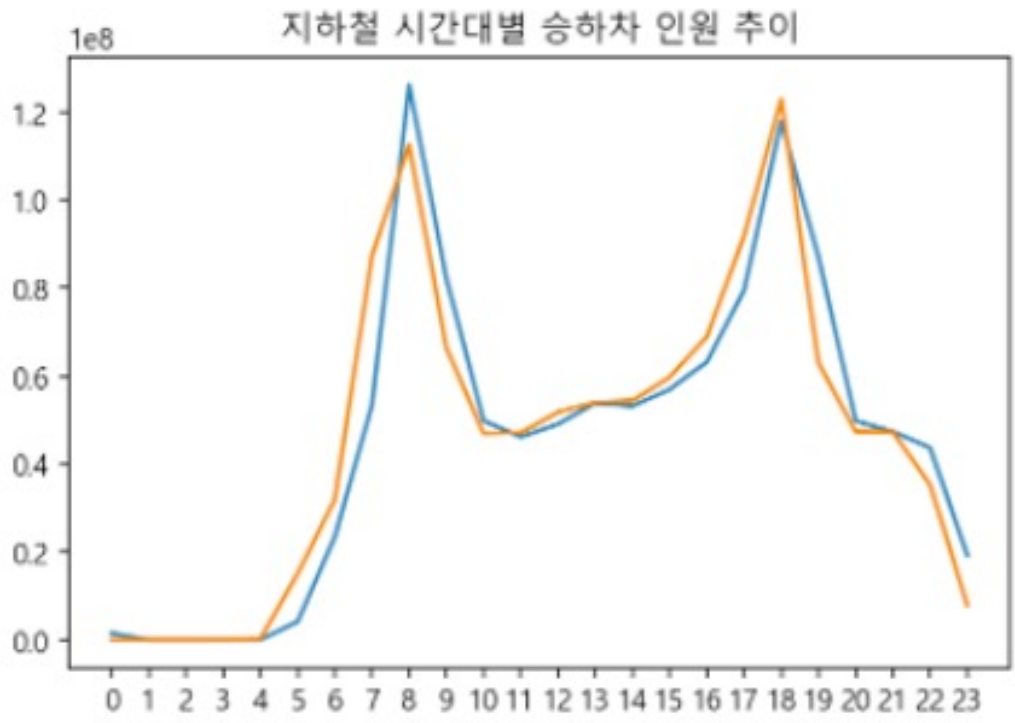
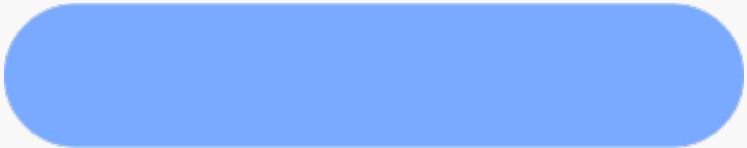

NIA

DATA MENTORING

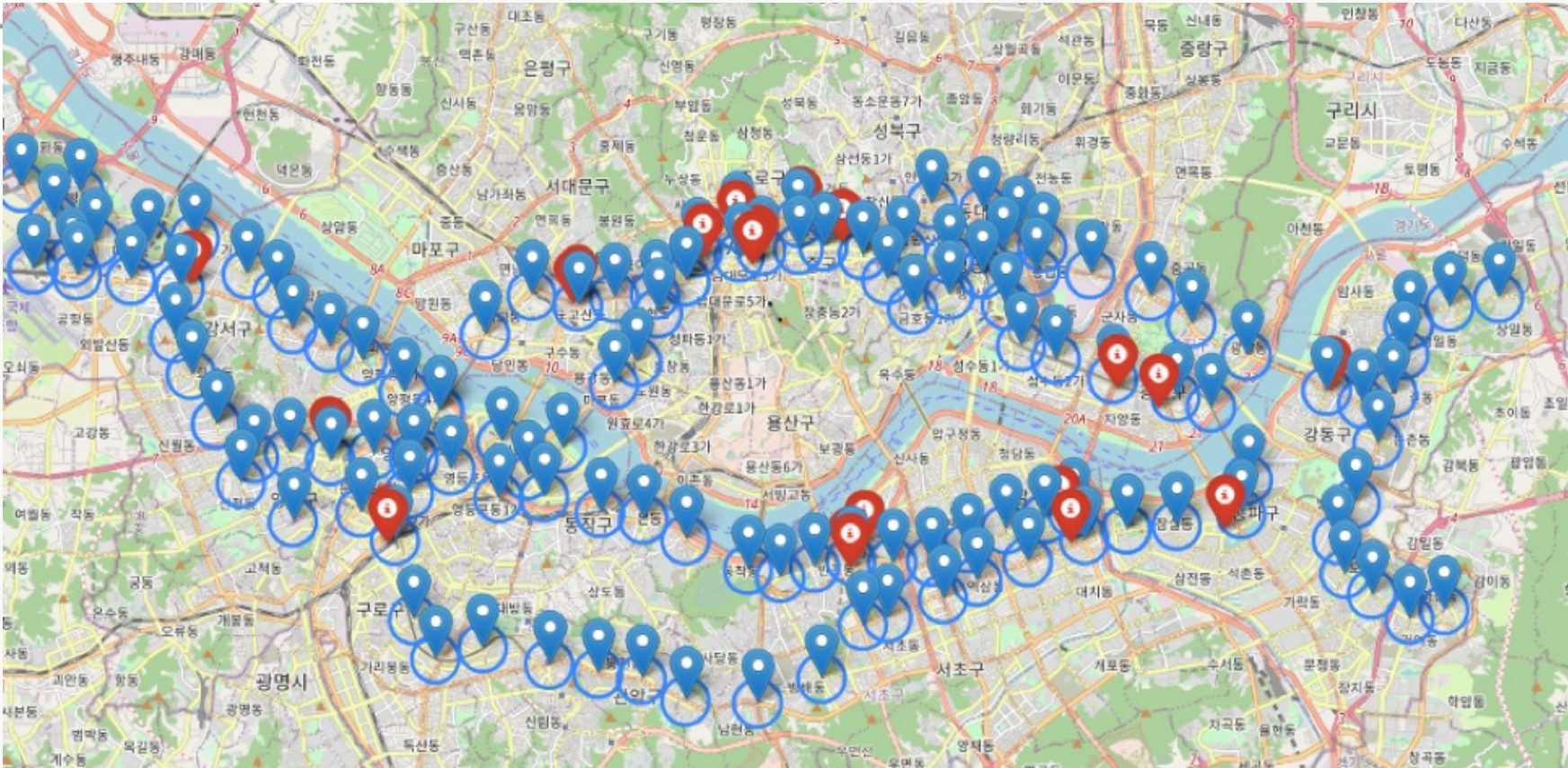
NATIONAL INFORMATION
SOCIETY AGENCY
MENTORING





지하철 역 500m 내 백화점 시각화

지하철 시간대별 승하차 인원 추이



데이터 분석 방향성

분석시 중요하게 생각하는 것

- ✓ 구글링
- ✓ 분석해야할 데이터의 사전조사 및
기본지식
- ✓ 데이터 선정시 전체 대변할 수 있는
정당한 이유 & 합리적 이유



근거 없는 분석

- ✓ 데이터 분석 시 시행착오
근거없는 주관적인 생각만을 가지고 데이터 및
변수 선택 후 프로젝트 진행
-> 안좋은 결과 얻어냄
- ✓ 노하우
데이터 및 변수 선정에는 근거가 있어야 한다는
노하우 생긴 상태



데이터 전처리 시 의문 데이터

- ✓ 의문이 드는 데이터
 - 그냥 흘리는 것이 아닌 이상치 확인
 - 전처리가 의미있게 되어지는 것인지 수정
 - 위의 과정을 통해 분석을 진행 -> 좋은 결과

데이터 활용 방향

서울시 지하철 혼잡도 분석 및 새로운 역 개통시 승객 수요 예측

- ✓ 환승인원 정보 데이터
혼잡도 예측에 있어서 가장 중요한 지표
- ✓ 지하철 이용통계 데이터 및 다양한 공간데이터
 - 우수 중소기업
 - 수도권 지하철
 - 서울시 중,고,대학교
- ✓ 유동인구 데이터, 주거시설 데이터
지하철 이용비율 70-80% 출퇴근

집객 시설

영화관 | 백화점
주거시설
학교 | 회사
버스 정류장
학생 종사자 수

쇼핑, 문화생활 등과 같은 해당역의 통행 목적 특성 반영
아침, 저녁에 유동인구가 가장 많음 (출퇴근)
회사, 학교가 많은 곳은 유입(하차)가 많음
역 주변 정류장 수 = 지하철 - 버스 간 환승의 용이성 나타냄 (역의 접근성 좋음)
최대 승차 인원 기록하는 시간대 목적 -> 통근, 통학

지원금 활용 방안

지하철 관련 데이터분석 프로젝트 진행 시
가장 핵심데이터 : 환승인원 데이터 | 다양한 공간데이터
유료 데이터 -> 데멘토에서 지원받는 지원금 적절히 활용 예정

기획

기획 예정

✓ 데이터셋 가공, 전처리, EDA

» 환승인원, 지하철 수요, 서울시 공간데이터, 유동인구데이터
각종 데이터셋 가공 및 전처리

» EDA 등 분석 초기단계는 어느정도 틀이 짜여진 상태

멘토링 활동 목표

✓ 정확도가 나름 높은 예측 모델 만들기 원함

✓ 예측모델이 완성된다면, 실제 9호선에 확장되는 개통역의 수요승객수 예측

» 예측 모델을 만들기 까지의 과정
-> 부족한 부분을 멘토님에게 피드백 받으며 보완

» 더 나아가 서울시 지하철 시간대별 혼잡도를 예측
-> 유의미한 인사이트를 도출

