

빅데이터 사이언스 실무자 양성과정

< 파일럿 프로젝트 착수보고서 >

주 제	지역 인프라 분석으로 인프라 발전 방안 모색		
팀 명	취미요? 파이썬이요.	일 자	2019년 9월 10일
팀 장	손하늘<lab10922@gmail.com>	팀 원	양한솔, 류경아

1. 과제 개요

수도과 지방의 인프라 차이로 발생하는 인구 밀집 현상이 심화되고 있습니다. 인구 밀집 현상을 완화시키려면 지방의 인프라가 발전해야 하는데 현재로서는 미흡한 수준. 이번에 저희 팀에서 인구에 비해 인프라가 턱없이 부족한 지역을 선별, 어떤 방향으로 인프라를 개선시킬지 분석하려 합니다.

2. 데이터 수집 전략

- 인프라는 사회인프라(병원, 교육, 여가, 문화 등)으로 제한
- 영역은 광역시로 제한
- 광역시 별 인구 데이터 수집
- 인프라의 지표 중 하나로 유명 프랜차이즈 분점을 채택. 스타벅스, 버거킹, 시네마(CGV, 롯데시네마), 서브웨이 지역 별 지점 차이 웹크롤링
- 공공데이터에서 병원, 도서관, 학교, 버스 정류장(교통편) 등 데이터 수집
- 토지 매매가 데이터 수집

3. 데이터 분석 전략

- 인프라가 가장 좋은 서울과 다른 광역시 비교
- 수집한 인구 데이터와 인프라 데이터를 시각화하여 인구에 비해 인프라가 가장 부족한 지역 채택
- 토지 매매가를 분석하여 채택된 지역의 최적 인프라 조성 구역 선택하기 (시간이 남다면 인프라에 따른 아파트 매매 거래가 상관관계 분석)

4. 결과 보고 예측 및 실행 계획

- 9/9 ~ 9/10 공공데이터 수집
- 9/11 ~ 9/14 데이터 전처리 / 시각화
- 9월 둘째주 주말 오프라인 모임 (시각화 및 피드백)
- 9/16~ 2차 시각화 진행 및 정리

- **결과 도출 전략 및 가설**

저희의 가설은

1. 인프라가 좋은 곳은 인구가 많고 인프라가 적은 곳은 인구가 적어 지역 경제 발전이 늦다.
 2. 즉 인프라를 높여야 사람이 몰려 유동인구가 상승하고 지역 경제 발전이 이뤄질 수 있다.
- 만약 가설이 성립되면 그 가설에 따라 경제 발전을 수반하는 유동인구 상승을 위해 가장 부족한 지역부터 발전시키는 것을 목표로 발전 전략을 세우겠습니다. 그리고 추가적으로 그 지역에 서울에 비해 비율적으로 가장 부족한 인프라를 발전 목표 1순위로 두겠습니다.

- **파일럿 프로젝트를 통해 얻고 싶은 것**

수업한 내용을 바탕으로 새로운 데이터 탐색에 적용해보고 싶습니다.

스스로에게 부족한 점을 보완한 결과물을 완성해보고 싶습니다.

팀 프로젝트인 만큼 협력해서 더 시너지를 내고 싶습니다.

감사합니다

- **결과보고서 제출 과제**

- 1) 발표자료 : 프리스타일, 파워포인트 12page 내외
- 2) 소스코드 : 데이터수집, 데이터분석은 모듈별 별도 파일 작업후 (폴드를 하나의 파일로 압축)
- 3) 기타자료 : 그외 프로젝트 산출물