대한상공회의소 광주인력개발원

# 광주 광산구(21개 행정구역) 데이터 분석 및 편의점 입지 선정

TEAM 2조 미정 (팀명)

이범규, 박의용, 주민석, 장은희





# 목차

- 01. 프로젝트 개요
- 02. 프로젝트 팀 구성 및 역할
- 03. 프로젝트 수행 절차 및 방법
- 04. 프로젝트 수행 결과
- 05. 자체 평가 의견

#### 프로젝트 개요

- ▶ 프로젝트 주제 및 선정 배경
  - 데이터 분석을 통한 프랜차이즈 입지 선정이 공통 주제였습니다.
  - 데이터가 광범위해지면 분석에 차질이 생길 것을 우려하여, 광주광역시 광산구로 **지역 한정**을 하였습니다.
  - 프랜차이즈 입지 선정 부분에서는 누구나 접하기 쉽고, 친근한 편의점으로 주제를 선정하였습니다.
- 프로젝트 개요
  - 편의점 입지 선정에 필요한 데이터 (접근성, 편의시설, 유동인구) 중점으로 데이터 수집 (목표).
  - 신뢰성 확보 (컨셉) 및 훈련내용과의 관련성 (이전 프로젝트에서 공공데이터 포털 중점 데이터 활용 능력 연습).
  - 이후 부족한 자료 크롤링 기술을 이용하여 데이터 (보완).
  - 데이터 분석한 내용을 시각화 해서 보여주고, 광산구 행정동 최적입지조건을 사용자에게 보여준다 (<mark>구현내용</mark>).
- ▶ 활용 장비 및 재료 (개발 환경)
  - Window OS, C++, Qt Creator, QGIS, MySQL.
- 프로젝트 구조
  - 서버 데이터베이스구축 및 클라이언트 데이터 분석에 필요한 데이터 요청 이후 시각화.
- 기대 효과
  - 사용자가 필요한 항목에 데이터 분석을 직접하는 느낌을 주고, 광산구 행정동 입지조건을 분석할 수 있다.

## 프로젝트 팀 구성 및 역할

훈련생	역할	담당 업무
박의용	팀장	▶ 서비스 시스템 <b>설계</b> (주제 선정, 데이터 수집 및 정제, 서비스 구조) ▶ 서버 시스템 보조 및 클라이언트 시스템 보조 (최종 테스트)
이범규	팀원	► 서버 시스템 구축 ( 클라이언트 요청 데이터 송신 ) ► 데이터 수집
주민석	팀원	▶ 데이터 베이스 총괄 구축 및 정제 ▶ 클라이언트 시스템 구축 ( 데이터 <b>분석</b> )
장은희	팀원	▶ 데이터 베이스 구축, 정제 보조 ▶ 클라이언트시스템 구축 (데이터 <b>시각화-차트</b> )

## 프로젝트 수행 절차 및 방법

구분	기간	<u>활</u> 동	비고
사전 기획	▶ 2/23(목) ~ 3/1(수) (7일)	▶ 프로젝트 기획 및 주제 선정 ▶ 기획안 작성	▶ 아이디어 선정
데이터 수집	▶ 2/27(월) ~ 3/1(수) (3일)	▶ 필요 데이터 및 수집 절차 정의 (159건) ▶ 외부 데이터 수집 (64건)	▶ 공공데이터 포털
데이터 전처리	▶ 3/2(목)~ 3/7(금) (6일)	▶ 데이터 정제 및 정규화	► PYTHON, QGIS, MySQL
모델링	▶3/3(금)~ 3/10(금) (8일)	▶ 모형 구현	► C++, Qt Creator
서비스 구축	▶ 3/11(토)~3/12(일) (2일)	▶ 데스크톱 시스템 설계 및 테스트	▶ 최적화, 오류 수정
총 개발기간	▶ 3/2(목)~ 3/12(일) (11일)	-	▶ 사전 기획 기간 미포함

결과 제시 ① 데이터 수집

- 데이터 수집 방법
  - 공공 데이터 포털 (전체 159건 중 64건 수집)
  - 광주 빅데이터 플랫폼(참조)
  - 광주광산구홈페이지 데이터 시각화 자료 (참조)



박의용K2nd22 오늘 오전 11:39

네넵

지금

1페이지부터

16페이지까지있거든요

1페이지부터 5페이지 - 박의용, 6페이지-10페이지 - 주민석, 11페이지-16페이지 장은희 / 각자 맡아서 CSV하나씩 다열어보시고 필요한 데이터 라고 판단이되는것들 선 별해서 뽑아주세요 범규형은 나머지 카드사랑 유동인구관련 데이터 좀 해주시면 될 것 같습니다



박의용K2nd22 오늘 오전 11:47

지금 크게 데이터를 확보할 수 있는곳은 3곳입니다.

1. 공공데이터포털

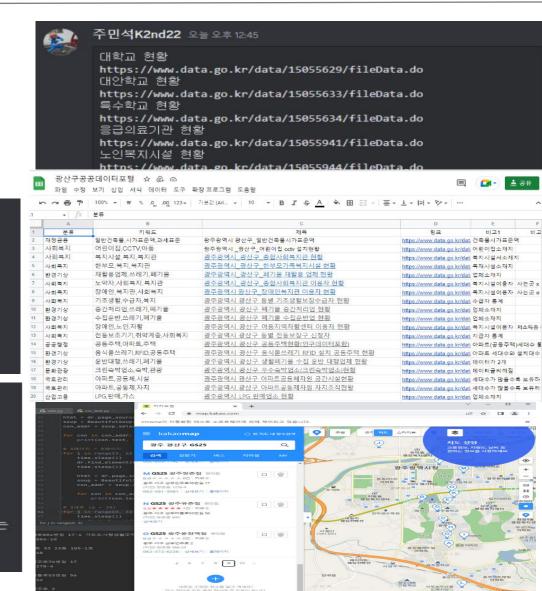
2.

광주빅데이터플랫폼

3.광주광산구 데이터시각화자료

(광산구 홈페이지)

위에서말한것처럼 저희는 공공데이터 기반으로 데이터를 찾을거구요, 거기서 찾는 노력을 다했는데 도무지 안나오는 데이터는 크롤링 기반으로 데이터 분석하는 방 향으로 생각하겠습니다



결과 제시 ② 데이터 수집 결과 및 분석 프로세스 회의

데이터 수집 결과 및 분석 프로세스 방안

## 분석 데이터 목록 (광주광역시 광산구)

<전체 공공데이터 159건의 CSV자료 중 총 70여개 데이터 수집>

분석 데이터 종류	제공기관
1. 인구현황 ( 동별, 성별, 세대수, 1인가구, 외국인 )	공공데이터 포털
2. 공동주택 및 오피스텔 현황	공공데이터 포털
3. 음식점 및 유흥주점·단란주점 현황	공공데이터 포털
4. 일반·상업시설, 복지·의료·치안·편의 시설 현황	공공데이터 포털
5. 교육시설 현황	공공데이터 포털
6. 토지, 건물기반 현황	공공데이터 포털
7. 유동인구 현황	공공데이터 포털

## 데이터 분석 프로세스

	1단계	2단계	3단계	4단계
데이터 수집	사회적 데이터 (인구밀집도, 편의시설)	경제적 데이터 ( 대중교통, 공시지 가, 유동인구 )	공간적 데이터 (도로,건물,토지)	상관관계 분석
데이터 전처리	데이터 셋 확인 데이터 추출/ 결합		데이터 통합	서버(C++) 데이터베이스 구축
데이터 분석	상관관계 분석	LQ(Location Quoient) 입지계수 도출	LQ(Location Quoient) 입지계수 분석	QT(C++) 데이터 시각화를 통한 분석

결과 제시 ③ 데이터 전처리(1)

▶ 데이터 전처리(1) PYTHON Pandas(판다스)를 이용하여 불필요한 값 정제

## 데이터 전처리 정제 방안

(불필요한 값 제거)



사회적 데이터 (인구밀집도, 편의시설)

: 위치선정에 필요한 칼럼 제외한 다른 값 제거



경제적 데이터 ( 대중교통, 유동인구 )

: 상업 지역 제외한 변수 제거



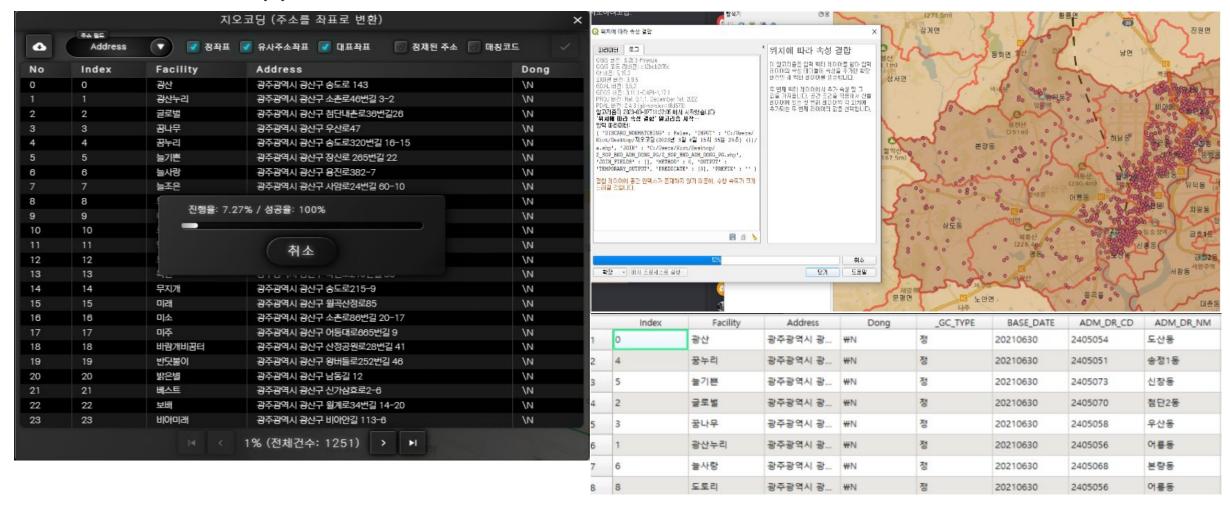
공간적 데이터 ( 도로 ,건물,토지 )

: 제외한 변수들을 채택



결과 제시 ③ 데이터 전처리(2)

▶ 데이터 전처리(2) 지오코딩 및 QGIS 를 이용하여 주소에 따른 행정동 칼럼 추가



결과 제시 ④ 데이터 분석 방법 및 최종 모형안

LQ (Location Quotient : ) 입지계수 분석

# 데이터 분석 방법 ( 광산구 전체 ) Location Quotient (입지계수) 분석

● 용어풀이 · 73 ●

#### 입지계수

입지계수(location quotient: LQ)는 어떤 지역의 산업에 대해 전국의 동일산업에 대한 상대적인 중요도를 측정하는 방법으로서 그 산업의 상대적인 특화 정도를 나타낸 지수다. 이 계수를 이용하면 한 지역의 특화 산업을 쉽게 분석할 수 있다. 따라서 입지계수는 경제기반모형(economic base model)과 투입산출모형 (input-output model) 등의 지역경제 분석모형에서 기반(특화)산업과 비기반산업을 나누는 분류기준으로 널리 사용되고 있다.

 $LQ = \frac{j \times 9 i \cdot \text{Vel } \text{SANA} / j \times 9 \text{SANA}}{2 \times 3 \cdot \text{SANA} / 2 \times 3 \text{SANA}}$ 

다만 입지계수는 산업제품의 수요패턴이 전국적으로 동일하다고 가정하여 상대적으로 지수를 계산하 기 때문에 결과를 해석하는 데 신중을 기해야 한다. 예를 들어 한 지역의 전체 산업규모가 크지 않은 경우 이면 특정 산업이 조금만 집중하여도 해당 산업이 특화산업으로 분석될 수 있다는 문제점이 있다. 따라서 보완설명이 필요하도록 전체 산업규모를 나타내는 지표와 같이 고려되어야 한다.

https://library.krihs.re.kr > ART000030445616 PDF

■ 입지계수 중심지이론

#### <표 6-1> 지역별 입지상(LQ) 분석결과

도소매 숙박음식 운수창고 금융 부동산 교육 보건 개인서비스

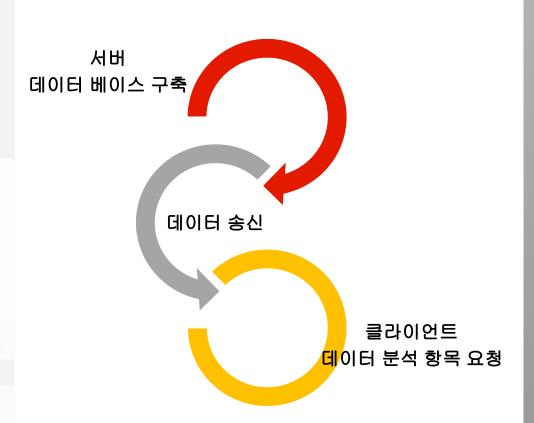
서울	1.17	0.83	1.15	1.24	1.62	0.77	1.02	0.94
부산	1.09	1.09	1.48	1.04	0.84	0.97	1.11	0.96
대구	1.08	0.99	1.00	0.93	0.84	0.96	1.18	1.02
인천	0.86	0.92	0.94	0.79	0.78	0.96	1.01	1.04
광주	1.11	0.97	1.09	1.24	1.00	1.20	1.23	1.14
대전	1.10	1.10	1.05	1.11	1.60	1.24	1.15	1.10

제6장 출점적합성 평가 ■ 191

## 출처

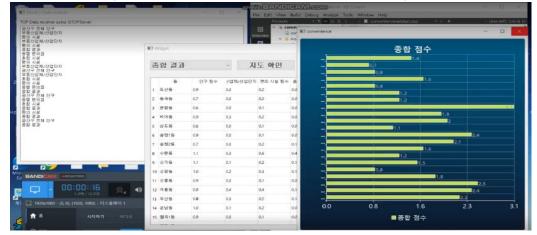
http://naurea.co.kr > file 20160610233818 Pr

- ■총서4. 프랜차이즈 입지 및 상권분석 연·
- 연구부문, 집필자, 주요 경력, 성 명, 제1권, 프랜차이즈 경영원론... 점포분석, 제5장 성 제1절 상권분석 프로세스, 1, 상권 및 업지조사 프로세스.

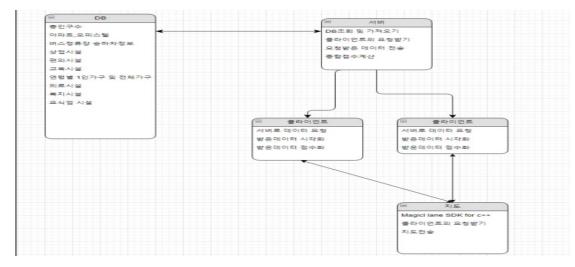


결과 제시 ⑤ 팀원별 수행 과정

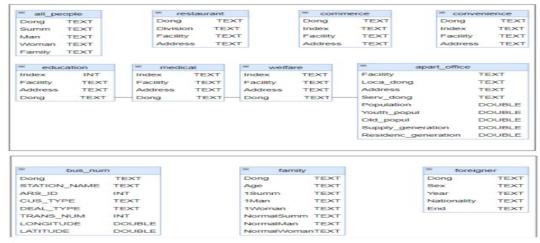
박의용 (전체 서비스 설계 및 테스트)



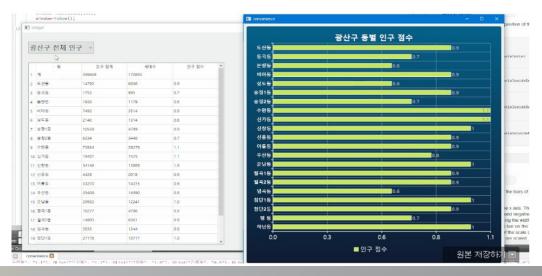
▶ 이범규 (서버 및 데이터 송신 )



주민석 (데이터 총괄 및 분석)

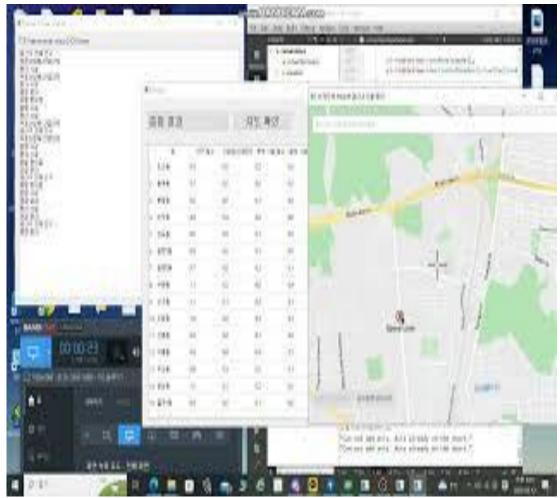


장은희 (데이터 시각화)



결과 제시 ⑤ 시연 동영상





### 자체 평가 의견 및 후기

- 처음 프로젝트기획 당시, 팀원 전체 프로그래밍 언어(C++)과 환경(Qt Creator)에 익숙치 않다고 판단하여, 입지분석 데이터 구축 후에 프로그램 초기 구현을 먼저하고, 세부적으로 주제(편의점)에 맞게 데이터 및 프로그램을 구현하려고 하였으나, 세부적 진행단계로 나아가지 못하였다.
  - \* 모델 평가 결과, 완성도가 50.00%로 완성도 향상을 위해 모델 추후 개선 필요

	훈련생	역할	후기
			▶전체적으로 팀원들 간 코드,데이터 모든 부분에서 유기적 협력을 하기 위해 짧은 시간
	박의용	팀장	고민을 많이하였다. 결과적으로는 완성도가 프로젝트 의도에 부합하지 못했다고 생각은
	518	<u> </u>	하지만, 프로젝트 진행 과정 전반에 팀원들 각자 맡은 파트에 본인 포함 할 수 있는 만큼
			최대한의 노력을 하였고, 성장하였다고 생각이 들어, 이 부분에서는 만족한다.
		팀원	▶C++, 또 pyqt가 아닌 qt6로 서버 클라이언트 구축을 하는것은 처음이라서 그것에 초점을
	이범규		두다보니 디테일한 다양한 데이터를 가공하지 못한것이 아쉽지만 전반적으로 계획한
			시간 및 내용으로 구현이 잘되서 다행이라고 생각합니다.
		팀원	▶익숙하지 않은 언어와 기능들을 이용해 프로젝트를 진행하는만큼 사전준비 과정에서
	주민석		많이 긴장했지만 팀원들이 각자의 역할을 잘 수행했고 개인적으로 새로운 기능을
			다양하게 사용해볼 수 있어서 굉장히 유익했던 프로젝트였다고 생각한다.
		팀원	▶빅데이터 분석 주제로 인한 고민을 거쳐 광산구 상권분석을 통한 상권 추천 주제를
	장은희		선택하여 데이터 분석을 배웠습니다. 과정지향적 시각화에 대한 중요성을 바탕으로 <b>Q</b> t
	· 이 는의	급권	프레임워크를 이용한 QChart 구현을 한 프로그램을 만들며 많은 경험과 지식을
_			얻었습니다.