



서울시 내의 고가 및 저가형 카페의 입점 분석을 통한 지역 불균형 탐구

5조 문과의 향기



1. 프로젝트 개요

- 프로젝트 개요
- 강남3구 동북4구 선정 이유

2. 프로세싱

- 데이터 수집 및 전처리 과정
- 데이터 모델

3. 데이터 분석

- 자치구 간 카페 점포 수 비교(1)
- 자치구 간 카페 점포 수 비교(2)
- Feature 정의
- Feature 간 상관계수
- 자치구 별 카페 feature 상관계수(1)
- 자치구 별 카페 feature 상관계수(2)

4. 기대효과 및 한계점, 개선 사항

5. 개발후기 및 느낀점

6. 출처

7. 부록

8. 사용도구

9. 질의응답



1. 프로젝트 개요

[주제]

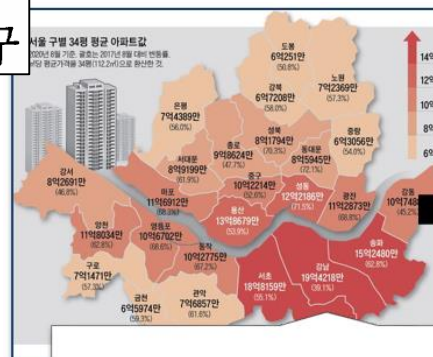
서울시 내의 고가 및 저가형 카페의 입점 분석

[배경]

- '서울은 돈 많으면 스타벅스, 적으면 백다방. 부자와 가난한 자 한 공간 있을 가능성 낮아...' (유현준 교수, 공원과 스타벅스의 차이 중 발췌)
- 고가형과 저가형 카페를 통해 서울시 자치구 간 부의 불균형을 볼 수 있을까?

서울시 자치구간 지역불균형 특성분석 연구

최근 서울시는 성장과 발전을 거듭하면서 자치구 간 물리적, 질적인 도시 요소에 불균형이 심화되면서 자치구 간 불균형 발전이 사회적 이슈로 부각되고 있다. 특히 자치구 간의 재정 격차, 고용 격차, 생활편의시설 및 삶의 질 격차 등은 서울시의 균형발전을 저해하는 요인으로 작용하고 있다. 이러한 자치구 간 불균형 현상을 해소하고자 서울시는 균형발전방안 및 정책들을 제시하고 있다. 새로운



서울시 집값



서울시 카페

1. 프로젝트 개요: 강남3구 동북4구 선정 이유



[강남 3구]

강남, 서초, 송파

강남과 강북지역의 공동주택단지 생활
환경 비교 평가 연구

조미정** ; 임영진*** ; 이명훈****

스타벅스 효과가 인근
표준지공시지가에 미치는 영향에
대한 연구

-서울 동북2권을 대상으로-

본 연구는 부동산 시장에서 발생하고 있는 스타벅스 입점에 따른 파급효과에 관심을 두었으며 이를 학문적인 근거 아래 학술적으로 증명하고자 한다. 연구 대상 지역은 서울시 생활권 계획상 베드타운으로 제시된 동북 2권(성북구, 노원구, 도봉구, 강북구)을 범위로 설정하였으며, 그 지역에 위치한 스타벅스 카페와 이디야 카페를 중심으로 반경 300m 이내 토지의 표준지공시지가를 중심으로 연구하였다. 연구 목적은 첫 번째,

[동북 4구]

강북, 노원, 도봉, 성북

1. 배경 및 목적

- 「도시재생 활성화 및 지원에 관한 특별법」(이하 “도시재생특별법”이라 한다.) 제13조(도시재생 전략계획의 수립)에 따라 창동·상계지역이 활성화지역으로 지정(15.12.10)됨

동북2생활권의 간선버스 노선 수는 서북생활권에 비해 절반 수준

지하철을 대신하는 간선버스 노선수 역시 최하위권

2호선이 지나지 않는 동북4구, 전철역 수도 최하위권

서울 각 권역을 통과하는 지하철노선도: 2호선이 지나지 않는 동북4구

■ 베드타운이었던 신시가지는 점차 노후화

- 동북부 320만의 입지적 중심이나 베드타운(주거중심)으로 지역이미지가 쇠락

2018년 12월 4일 [화]
13:00 ~ 19:00

I·SEOUL·U
내와 내리서울

플랫폼창동61
레드박스



주최·주관 | 서울시 동북4구 도시재생협력지원센터

협력 | 동북4구 행정협의회(성북, 강북, 도봉, 노원), 서울시 동북권 NPO센터

2. 프로세싱 : 데이터 수집 및 전처리 과정



데이터 수집	데이터 추출	데이터 전처리 방법
아파트 매매가격	아파트 면적 아파트 거래금액 면적 당 거래금액(파생변수)	결측치 제거 이상치 제거 엔코딩(D4:0,G3:1) 구별 평균치
상권정보	카페 별 점포 수 (저가-백다방,메가커피,컴포스 고가-스타벅스,커피빈,폴바셋)	평균치 계산 엔코딩(저가:0,고가:1)
유동인구 수	동 별 총 유동인구 수	이상치 제거 네이버API(주소변환)
지하철역 정보	카페와 지하철역 간의 거리(파생변수)	경도,위도 변환
버스정류장 정보	카페반경 160m 내 개수(파생변수)	경도,위도 변환

(2018~2020년, 서울시 기준)

2. 프로세싱 : 데이터 모델



Features (input)

Target Variable

D4, G3

아파트 면적 당 거래금액(파생변수)
카페 별 점포 수
동 별 총 유동인구 수
카페와 지하철역 간의 거리(파생변수)
카페 반경 160m 내 버스정류장 개수(파생변수)

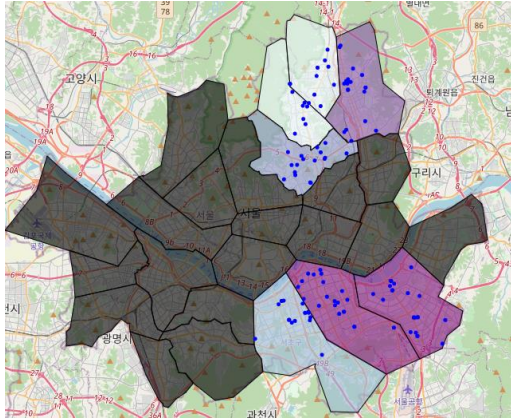


저가형 카페
고가형 카페

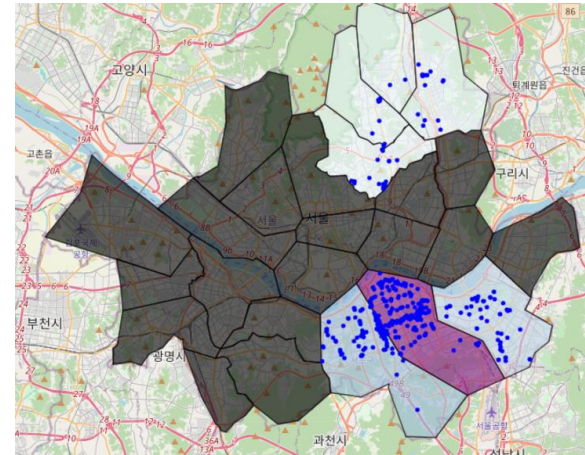
D4
G3

- 저가: 백다방, 메가커피, 컴포스
- 고가: 스타벅스, 커피빈, 폴바셋
- D4: 동북4구
- G3: 강남3구

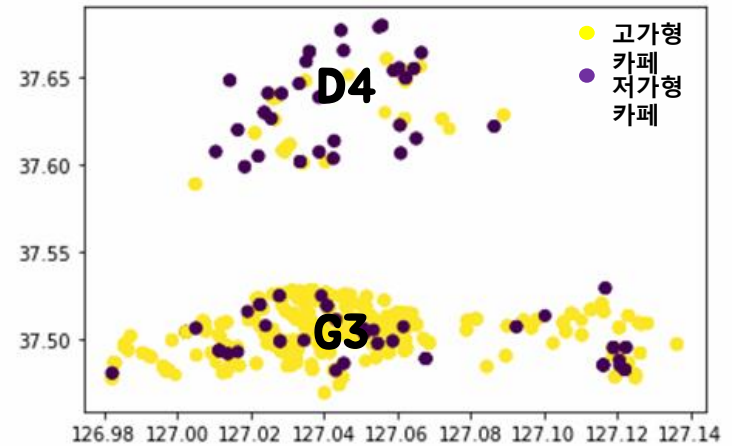
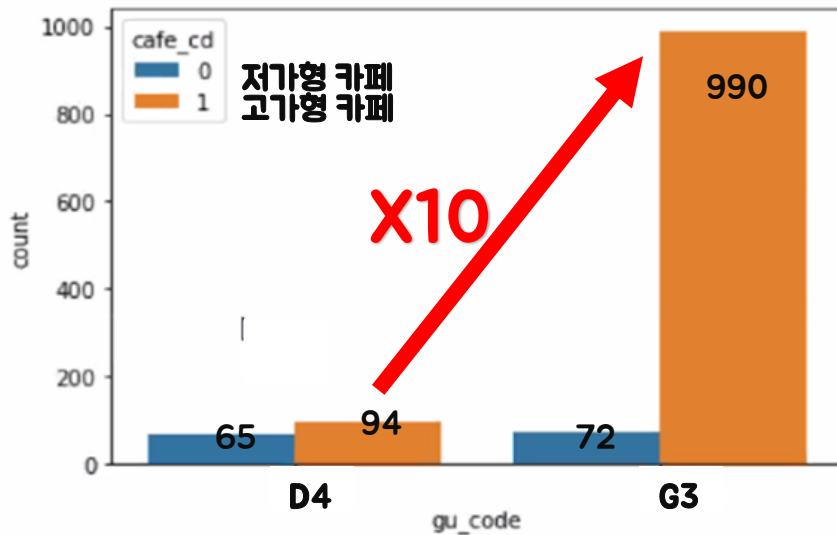
3. 데이터 분석 : 자치구 간 카페 점포 수 비교(1)



- 저가형 카페 분포



- 고가형 카페 분포



3. 데이터 분석 : 자치구 간 카페 점포 수 비교(2)



〈카페 브랜드〉

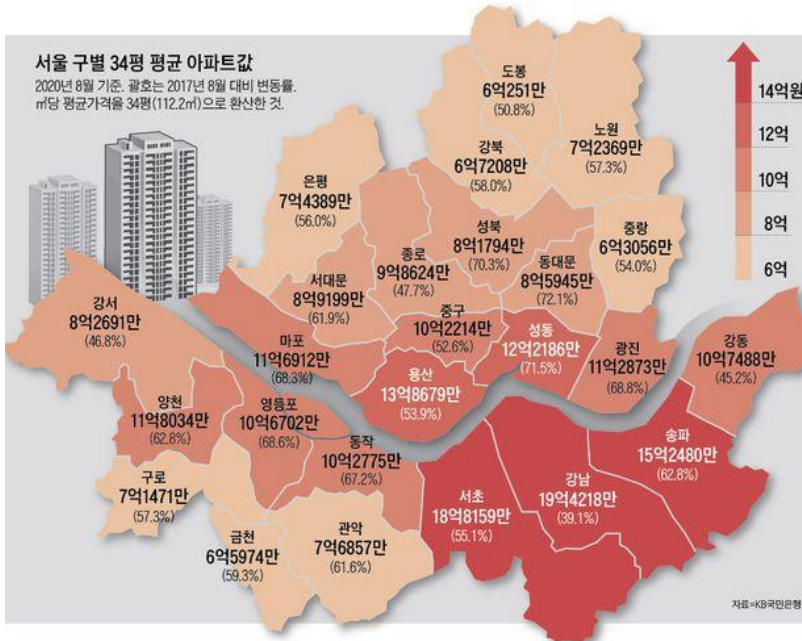


〈카페 위치〉

3. 데이터 분석 : Feature 간 상관관계



	lng	lat	year	bus_cnt	HubDist	cafe_cd	pop_avg	byear	average_sizetrade	gu_code
cafe_cd	-0.053112	-0.349579	-0.136847	0.023901	-0.022581	1.000000	-0.057371	0.009806	0.275141	0.363623
pop_avg	-0.009402	0.077520	0.748386	-0.016198	0.007692	-0.057371	1.000000	-0.017264	0.204599	-0.044465
byear	-0.225542	0.152287	-0.033662	0.201596	-0.076566	0.009806	-0.017264	1.000000	-0.403069	-0.143238
average_sizetrade	-0.068085	-0.699179	0.189133	-0.189776	0.021190	0.275141	0.204599	-0.403069	1.000000	0.739100
gu_code	0.028104	-0.949999	-0.066751	-0.192514	-0.011414	0.363623	-0.044465	-0.143238	0.739100	1.000000
cafe_cnt	0.016655	-0.006119	-0.041658	0.012414	0.015206	0.010178	-0.043408	0.006545	0.014686	0.011078



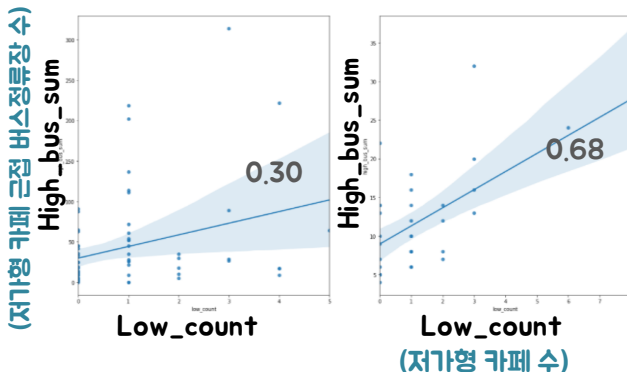
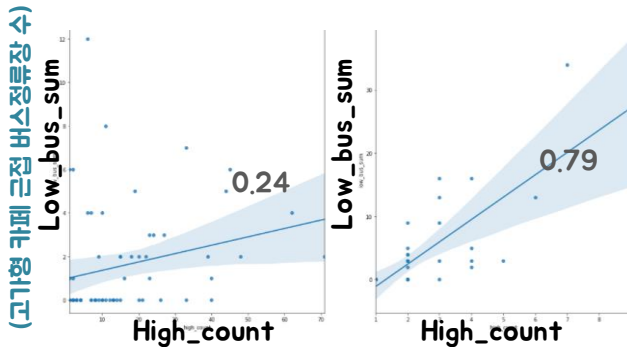
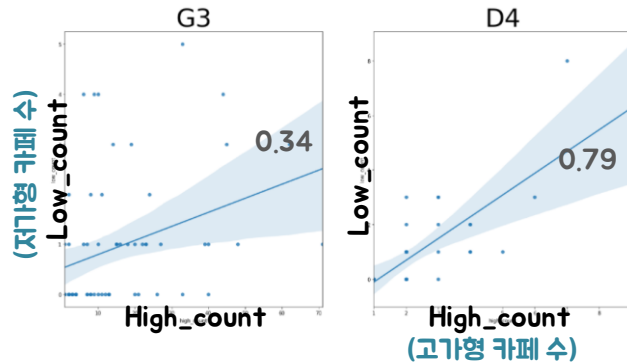
- 위도 ↓ -> 강남 3구 쪽 -> 면적 당 거래금액 ↑
- 자치구와 면적 당 거래금액의 양의 상관관계
- 위도 ↓ -> 강남 3구 쪽 -> 고가형 카페 ↑
- 자치구와 카페 전략의 양의 상관관계

- Encoding -

자치구 : D4 -0, G3-1

카페 : 저가 -0, 고가-1

3. 데이터 분석 : 자치구 별 카페 feature 상관관계수(1)



<G3>

	average_sizetrade	high_count	low_count
average_sizetrade	1.000000	0.152302	0.087590
high_count	0.152302	1.000000	0.338915
low_count	0.087590	0.338915	1.000000
high_bus_sum	0.060388	0.900350	0.307803
low_bus_sum	0.347714	0.243107	0.713672
pop_avg	0.204880	0.186569	0.231860

<D4>

	average_sizetrade	high_count	low_count
average_sizetrade	1.000000	-0.128806	0.082813
high_count	-0.128806	1.000000	0.790772
low_count	0.082813	0.790772	1.000000
high_bus_sum	-0.301967	0.790257	0.627791
low_bus_sum	0.104859	0.785778	0.953995
pop_avg	0.141883	0.428084	0.514784

High_count - 고가형 카페 수

Low_count - 저가형 카페 수

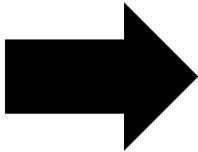
Low_bus_sum - 저가형 카페 근접 160m 내 버스정류장

High_bus_sum - 고가형 카페 근접 160m 내 버스정류장

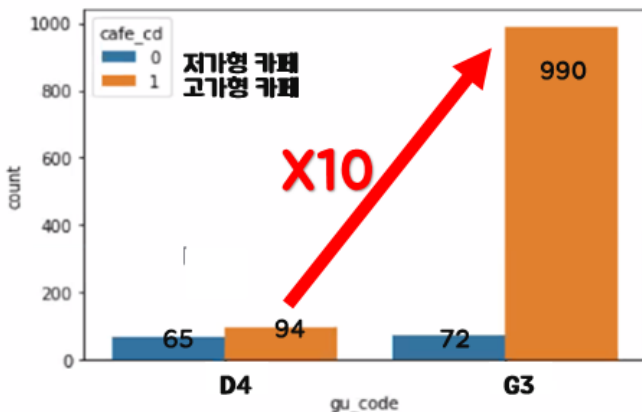
4. 기대 효과



- 서울 자치구 간 **고가형 카페의 점포수 차이**
- 서울 자치구 간 **아파트 매매 가격 차이**
- 서울 자치구 간 **대중교통 접근성 차이**



- 카페의 가격 정책 별 입점 지역 제안 가능
- 정부에 지역 간 균형 잡힌 발전을 위한 개발 정책의 필요성을 제안
-> 고가형 카페를 통해 지역간 생활 편의시설(삶의 질) 불균형 정도를 확인 가능



1) 서울숲의 아파트가격에 대한 영향

서울숲광장에서 각 아파트단지 중심점까지 거리로 측정되는 서울숲 변수는 3가지 모형 모두에서 통계적으로 유의한 영향을 주며, 아파트가격에 영향을 미치는 가장 강력한 변수의 하나인 것으로 나타났다(상대적인 영향력을

최근 서울시는 성장과 발전을 거듭하면서 자치구 간 물리적, 질적인 도시 요소에 불균형이 심화되면서 자치구 간 불균형 발전이 사회적 이슈로 부각되고 있다. 특히 자치구 간의 재정 격차, 고용 격차, 생활편의시설 및 삶의 질 격차 등은 서울시의 균형발전을 저해하는 요인으로 작용하고 있다. 이러한 자치구 간 불균형 현상을 해소하고자 서울시는 균형발전방안 및 정책들을 제시하고 있다. 새로운



4. 한계점 및 향후 개선 사항

[반영하지 못한 요인들]

- 카페의 입점 위치를 결정하는 요인
 - + 요인 : 대학교의 수, 상가 개수 등
 - - 요인 : 연령, 카페 밀집도 등
- 지역 불균형을 야기할 수 있는 다른 요소
 - 주거여건 격차, 재정 격차 등
- 유동인구 데이터에서 인구통계학적 특성(예:나이, 성별 등)

[연구의 확장성 필요]

- 서울의 특정 구(G3와 D4) 만의 비교를 통해 지역 간의 불균형을 분석
 - 분석범위를 확장하여 현 모델이 다른 지역에서도 적용 가능한지 확인이 필요

[지역특성 고려 필요]

- 유동인구에서 송파구의 이질적 특성을 고려하지 못함
 - 강남, 서초: 오피스 밀집
 - 송파: 놀 거리 밀집 (놀이동산과 야구장, 올림픽 공원 등)

5. 개발 후기 및 느낀점



이름	개발후기 및 느낀점
강주영	실제 데이터를 통해 전 처리 및 시각화를 해보니, 코드 실력이 따라와주지 않아 너무 힘들었습니다. 다음 프로젝트에는 가능하면 비/전공자가 적절하게 섞여 프로젝트 할 수 있으면 좋겠습니다.
김민형	데이터로 하는 팀프로젝트는 처음이라서 많이 헤매기도 했고 논리적 설명이나 코딩부분에서 막혀서 힘들었던 경험도 종종 있었습니다. 그러나 팀원분들이 프로젝트에 진심을 다해서 임해주셨기에 좋은 결말을 맺은 거 같습니다.
정길종	과정 중 첫 과제를 좋은 분들과 함께 할 수 있어 좋았습니다. 과제를 통해 배웠던 것을 복습하고 오류를 수정하는 과정에서 많은 공부가 되었습니다.
한유정	프로젝트를 처음 해볼 뿐만 아니라 코딩이나 서칭에 미숙하니깐 어려움이 많았습니다. 그래도 같이 하는 조원분들께서 이해해주고 제 의견에도 많은 관심을 주셔서 좀 더 재미있게 프로젝트에 임할 수 있었던 것 같습니다. 그래도 다음 프로젝트 때에는 전공자와 비전공자가 함께 할 수 있으면 좋겠습니다.

6. 출처



- ◆ 저가(검포즈, 뽕다방, 메가커피), 고가형(스타벅스, 커피빈, 폴바셋) 카페 자료
- 카페 위치
<https://www.data.go.kr/data/15083033/fileData.do>
공공데이터 포털, 한국 소상공인시장진흥공단
- ◆ 아파트(매매)_실거래가 자료
- 면적당 거래 금액
<http://rtdown.molit.go.kr/>
국토교통부의 실거래가 공개시스템
- ◆ 서울시 우리마을가게 상권분석서비스(상권배후지-추정유동인구)
- 행정구 별 유동인구
<https://www.data.go.kr/data/15076417/fileData.do>
공공데이터 포털, 서울특별시
- ◆ 서울교통공사 노선별 지하철역 정보
- 지하철 역 정보
<http://data.seoul.go.kr/dataList/OA-15442/S/1/datasetView.do>
서울교통공사 스마트정보처에서 제공
- ◆ 서울특별시 버스정류소 위치정보
- 버스 정류장 정보
<http://data.seoul.go.kr/dataList/OA-15067/S/1/datasetView.do>
서울특별시 도시교통실 보행친화기획관 교통정보과에서 제공
- ◆ 논문
- 이재원, 이우종 (2007). 서울시 자치구간 지역불균형 특성분석 연구. 국토계획, 42(5), 81-94
- 조미정, 임영진, & 이명훈. (2016). 강남과 강북지역의 공동주택단지 생활환경 비교 평가 연구. Journal of Korea Planning Association-Vol, 51(2), 31-53.
- 김호진. (2020). The Impact of Starbucks on Adjacent Assessed Land Price-The Case of Four Northeast Districts of Seoul (Doctoral dissertation, 서울대학교 대학원).
- 신상영, 김민희, 목정훈(2006). 서울숲 조성이 주택가격에 미치는 영향 14p
- ◆ 보고서 및 도서
- 서울특별시, 창동*상계 도시재생활성화계획(도시경제기반형) 2017, 2p, 6p
- GIS United (2014), 공공정책을 위한 빅데이터 전략지도 성공하는 지자체를 위한 GIS 분석. 더숲
- ◆ 신문
- 사라지는 서민주택... 서울 25개구 모두 아파트 평균값 6억 넘었다
- [#서울커피맵] 커피공화국..카페 점 찍으면 서울지도가 뜬다

7. 부록



[Feature 정의]

Feature	정의
lng	경도
lat	위도
year	데이터 연도
bus_cnt	카페 반경 160m 내에 있는 버스 정류장의 수
Hub_Dist	카페로부터 지하철역까지의 거리
cafe_cd	카페 유형 encoding(저가형:0, 고가형:1)
pop_avg	평균 유동인구 수
byear	아파트 건축연도
average_sizetrade	면적 당 평균 거래금액
gu_code	서울 지역별 encoding(D4구:0, G3구:1)

8. 사용도구



matplotlib

QGIS



NumPy



seaborn

Pandas



Microsoft
Excel



Folium



python™



wordcloud



scikit

learn

9. 질의응답



Q & A



감사합니다.