

# 〔 COFFEE INDEX. 〕



6조 (육남매)

# 목차

001 프로젝트 개요







002 진행 과정

003 결과물 및 결론

004 부록



# 구성원 및 역할

| 구분 | 팀원명 | 역할  |  |
|----|-----|---|--|
| 팀원 | 김강섭 | 데이터 수집 및 분석<br>커피지수 도출<br>데이터 전처리 왕               |   |
| 팀원 | 김미리 | 회의록 정리<br>데이터 수집 및 분석<br>커피지수 도출<br><b>PPT</b> 제작 |   |
| 팀원 | 박민정 | 데이터 수집 및 분석<br>커피지수 도출<br>코드 취합 정리                |   |
| 팀장 | 서인혁 | 데이터 수집 및 분석<br>커피지수 도출<br>태블로 시각화                 |   |
| 팀원 | 신하진 | 데이터 수집 및 분석<br>커피지수 도출<br>깃허브 레파지토리               |   |
| 팀원 | 이호연 | 데이터 수집 및 분석<br>커피지수 도출<br><b>PPT</b> 발표           |  |

The background image shows a rustic interior space. A dark wooden table is in the foreground, surrounded by yellow chairs. The wall is made of exposed brick. On the left, there is a wooden shelf with various items and a hanging lamp. On the right, two long wooden skis are leaning against the wall. The text 'Part 1' and '프로젝트 개요' is centered on the brick wall, enclosed in a white bracket.

# Part 1

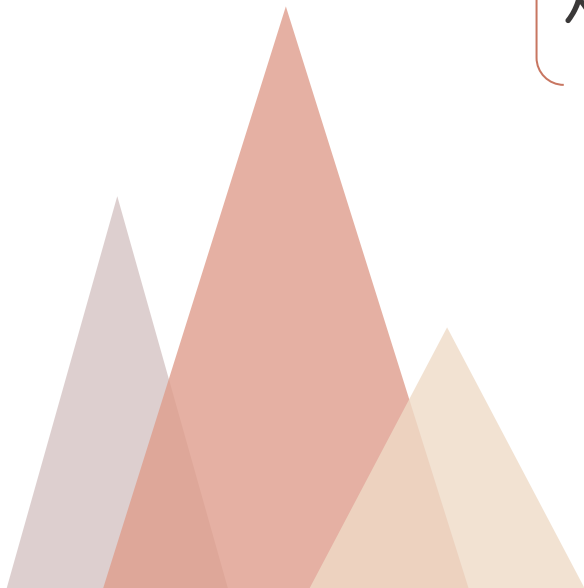
## 프로젝트 개요

# 프로젝트 주제

## 서울 지역의 인사이트 도출하는 커피지수 만들기

**선정 배경** : 커피전문 프렌차이즈의 지역별 매장 개수를 조사하면서,  
버거 지수처럼 흥미롭고 의미있는 지수를 만들어보자.

**목 적** : 커피 지수를 통해 지역구의 구매력을 평가할 수 있는  
경제지표로 활용될 수 있도록 인사이트를 도출하는 것





A top-down view of three hands holding coffee cups in a toast. The top hand holds a white cup with latte art. The bottom-left hand holds a dark cup with ice and coffee. The bottom-right hand holds a white cup with latte art. The background is a blurred wooden table with coffee-making tools.

Part 2

진행 과정

# 진행 과정

브랜드 선정

데이터 수집

지수 산정

인사이트 도출

# 1. 브랜드 선정

## [프랜차이즈 카페로 선택한 이유]

- 커피는 우리 일상생활에서 가장 쉽게 접할 수 있는 음료  
→ 실물경제지표로서 활용성이 높을 것이라 생각
- 프랜차이즈 매장은 동일 상품을 전국 매장 어디서나 볼 수 있으므로  
가격비교가 용이함

## [11개 브랜드 선정 기준]

- 최근(23년 1월) 브랜드 평판 순위 1~10위 브랜드 선정
- (12위) 엔제리너스는 송파구(잠실)에 많이 입점해있어, 지역 편차를 줄이기 위해 추가

<23년 1월 커피전문점 프랜차이즈 카페 순위>

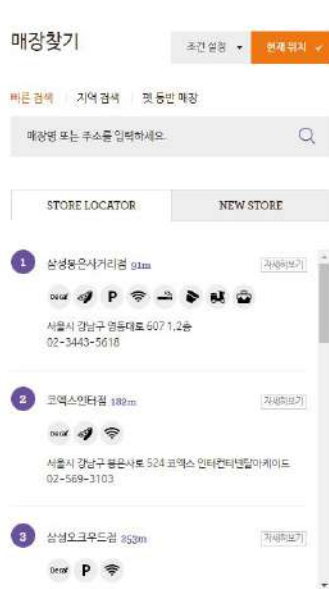
| 순위   | 프랜차이즈 카페 |
|------|----------|
| 1위   | 스타벅스     |
| 2위   | 메가커피     |
| 3위   | 이디야      |
| 4위   | 컴포즈커피    |
| 5위   | 빽다방      |
| 6위   | 커피빈      |
| 7위   | 파스쿠찌     |
| 8위   | 투썸플레이스   |
| 9위   | 폴바셋      |
| 10위  | 할리스      |
| (추가) | 엔제리너스    |



## 2. 데이터 수집

### [웹 크롤링]

Selenium, BeautifulSoup을 이용해 각 매장  
홈페이지를 크롤링하여 커피매장 데이터를 수집



### [데이터셋]

서울 열린데이터 광장, KOSIS 등 여러  
사이트에서 데이터셋을 다운받음

- ▶ 2022년도 서울시 생활인구, 유동인구
- ▶ 2022년도 서울시 공시지가
- ▶ 2022년도 서울시 대규모 점포
- ▶ 2022년도 서울시 재정자립도
- ▶ 2021년도 서울시 사업체 수, 종사자 수



서울 열린데이터 광장

공공데이터

통계

서울빅데이터

소식&참여

이용안내

### 데이터셋

Home > 공공데이터 > 공공데이터

찾고 싶은 데이터를 입력해 주세요.



상세 검색

통합 검색

☐ 결과 내 재검색



도시관리

활용사례(갤러리) 등록

URL 복사

목록 이동

### 서울시 개별공시지가 정보

개별공시지가는 매년 도시소제지 지가 조정률이 결정공시되는 1㎡당 토지(건물/㎡)로서 시가평형, 상업용형, 반자주형, 주상복합으로 인격임(시) 해당 지  
역의 일시 구상면적과 기준년월, 공시지가 평가일 제공되는 OpenAPI(2-F).  
※ OpenAPI 서비스는 2007년 우리 세종치과(가)에 이전(이전)되는 PaaS 서비스의 활용을 참조하시기 바랍니다.

## 2-1. 데이터 수집 결과

서울시 구별 커피 브랜드 갯수

| 시군구명1 | 브랜드명1 |     |      |       |     |     |       |        |      |     |     |
|-------|-------|-----|------|-------|-----|-----|-------|--------|------|-----|-----|
|       | 메가커피  | 백다방 | 스타벅스 | 엔제리너스 | 이디야 | 커피빈 | 컴포즈커피 | 투썸플레이스 | 파스쿠찌 | 폴바셋 | 할리스 |
| 강남구   | 21    | 17  | 88   | 1     | 44  | 50  | 19    | 42     | 12   | 12  | 16  |
| 강동구   | 22    | 10  | 17   | 1     | 26  | 4   | 12    | 12     | 1    | 2   | 3   |
| 강북구   | 20    | 9   | 6    | 2     | 20  | 1   | 2     | 7      |      |     | 2   |
| 강서구   | 35    | 17  | 25   | 3     | 42  | 4   | 27    | 19     | 5    | 4   | 2   |
| 관악구   | 18    | 11  | 12   | 4     | 28  | 2   | 13    | 8      | 3    | 1   | 6   |
| 광진구   | 14    | 10  | 19   | 2     | 24  | 4   | 16    | 12     | 1    |     | 5   |
| 구로구   | 28    | 13  | 14   | 1     | 29  | 2   | 19    | 11     | 1    | 4   | 5   |
| 금천구   | 13    | 8   | 13   | 2     | 23  | 2   | 16    | 11     |      |     | 2   |
| 노원구   | 28    | 14  | 14   | 2     | 29  | 3   | 8     | 11     | 2    | 1   | 3   |
| 도봉구   | 16    | 6   | 4    |       | 24  |     | 9     | 7      | 1    |     |     |
| 동대문구  | 18    | 7   | 10   | 1     | 25  | 3   | 10    | 9      |      | 1   |     |
| 동작구   | 14    | 6   | 11   | 3     | 22  |     | 9     | 5      |      |     | 6   |
| 마포구   | 27    | 11  | 35   | 3     | 34  | 7   | 15    | 25     | 1    | 3   | 10  |
| 서대문구  | 17    | 6   | 22   | 2     | 20  | 2   | 12    | 9      | 5    |     | 8   |
| 서초구   | 12    | 12  | 48   | 1     | 29  | 20  | 8     | 28     | 7    | 10  | 9   |
| 성동구   | 16    | 12  | 15   |       | 22  |     | 9     | 11     | 2    |     | 3   |
| 성북구   | 27    | 16  | 15   |       | 31  | 2   | 8     | 10     |      |     | 1   |
| 송파구   | 32    | 17  | 34   | 10    | 39  | 14  | 19    | 30     | 1    | 4   | 6   |
| 양천구   | 26    | 6   | 17   |       | 26  | 2   | 18    | 9      | 3    |     | 5   |
| 영등포구  | 26    | 10  | 41   | 4     | 41  | 11  | 14    | 21     | 3    | 6   | 8   |
| 용산구   | 6     | 8   | 25   | 1     | 14  | 4   | 1     | 9      | 4    | 2   | 4   |
| 은평구   | 24    | 10  | 13   | 5     | 26  | 1   | 15    | 6      |      | 1   | 3   |
| 종로구   | 16    | 8   | 40   | 5     | 31  | 15  | 8     | 12     | 3    | 3   | 11  |
| 중구    | 17    | 6   | 55   | 6     | 30  | 13  | 11    | 32     | 3    | 8   | 9   |
| 중랑구   | 19    | 10  | 8    | 1     | 29  | 1   | 10    | 10     |      |     | 1   |

### 3. 지수 산정

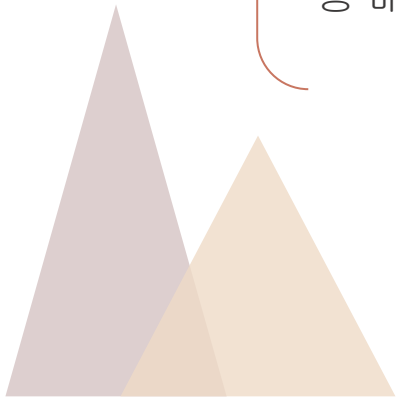
#### 팀원별 접근

| 팀원명 | 접근 방향                                     |
|-----|---|
| 김강섭 | 고가형 브랜드 / 저가형 브랜드 형태를 유지하며 각 브랜드에 가중치 곱하기 |
| 김미리 | 지역구별 브랜드 비율을 토대로 각 브랜드에 가중치 곱하기           |
| 박민정 | 각 지역별 평균 1ml 단가와 전국 평균 1ml단가 비교           |
| 서인혁 | 지역구별 총매출 대비                               |
| 신하진 | 지역의 낙후정도, 상업용 건물수 대비                      |
| 이호연 | 각 브랜드마다 연도별 증감 추세                         |

## 3-1. 지수 산정시 쟁점


### 총합계 영향 상쇄시키기

총 매장 개수가 많고, 적음에 영향을 받지 않는 지수를 만드는 것



## 3-2. 2가지 지수 접근 방법

### 1. 빅맥지수 관점으로 접근


$$\text{빅맥지수} = \frac{\text{각 나라의 빅맥가격}}{\text{미국의 빅맥가격}}$$

$$\text{전체 평균 1ml당 단가} = \frac{\text{브랜드별 1ml당 단가} \times \text{해당 브랜드 전국 매장수}}{\text{전국 전체 프랜차이즈 매장수}}$$

$$\text{지역구별 평균 1ml당 단가} = \frac{\text{브랜드별 1ml당 단가} \times \text{해당 브랜드 매장수}}{\text{지역구별 전체 프랜차이즈 매장수}}$$

$$\text{커피지수}_1 = \frac{\text{지역구별 평균 1ml당 단가}}{\text{전체 평균 1ml당 단가}}$$

| 순위 | 브랜드 명               | 1ml 당 원<br>(소수점 3 번째 자리에서 버림) |
|----|---------------------|-------------------------------|
| 1  | 커피빈                 | 14.08원                        |
| 2  | 폴바셋                 | 13.05원                        |
| 3  | 할리스                 | 12.71원                        |
| 4  | 스타벅스, 투썸플레이스, 엔제리너스 | 12.67원                        |
| 5  | 파스쿠찌                | 11.68원                        |
| 6  | 이디야                 | 8.33원                         |
| 7  | 빽다방                 | 3.75원                         |
| 8  | 메가커피, 컴포즈커피         | 2.64원                         |

ex) 어느 지역의 커피지수<sub>1</sub> > 1

- 전국 프랜차이즈 카페의 평균 단가보다 그 지역이 더 높은 단가를 가진다
- 해당 지역구의 구매력은 대한민국 전국 평균에 비해 고평가되어 있다

## 3-2. 2가지 지수 접근 방법

### 2. 자치구별 브랜드 비율로 접근

도시 발전 또는 경제  
수준이 높은  
지역구일수록 고가형  
브랜드가 차지하는  
비율이 높지 않을까?

$$\text{커피지수2} = \sum \frac{\text{해당 브랜드 매장수}}{\text{지역구별 전체 프랜차이즈 매장수}} \times \text{해당 브랜드 가중치}$$

#### 가중치

- ml당 원을 기준으로 중위값이 1이 되도록 조정
- 고가형일수록 높은 가중치를, 저가형일수록 낮은 가중치를 부여해서 차이가 드러나도록 접근

| 순위 | 브랜드 명               | 1ml 당 원<br>(소수점 3 번째 자리에서 버림) | 가중치<br>(ml당 원의 중위값을 1로 두기) |
|----|---------------------|-------------------------------|----------------------------|
| 1  | 커피빈                 | 14.08원                        | 1.11                       |
| 2  | 폴바셋                 | 13.05원                        | 1.03                       |
| 3  | 할리스                 | 12.71원                        | 1.00                       |
| 4  | 스타벅스, 투썸플레이스, 엔제리너스 | 12.67원                        | 1.00                       |
| 5  | 파스쿠찌                | 11.68원                        | 0.92                       |
| 6  | 이디야                 | 8.33원                         | 0.66                       |
| 7  | 빽다방                 | 3.75원                         | 0.30                       |
| 8  | 메가커피, 컴포즈커피         | 2.64원                         | 0.21                       |

## 4. 인사이트 도출

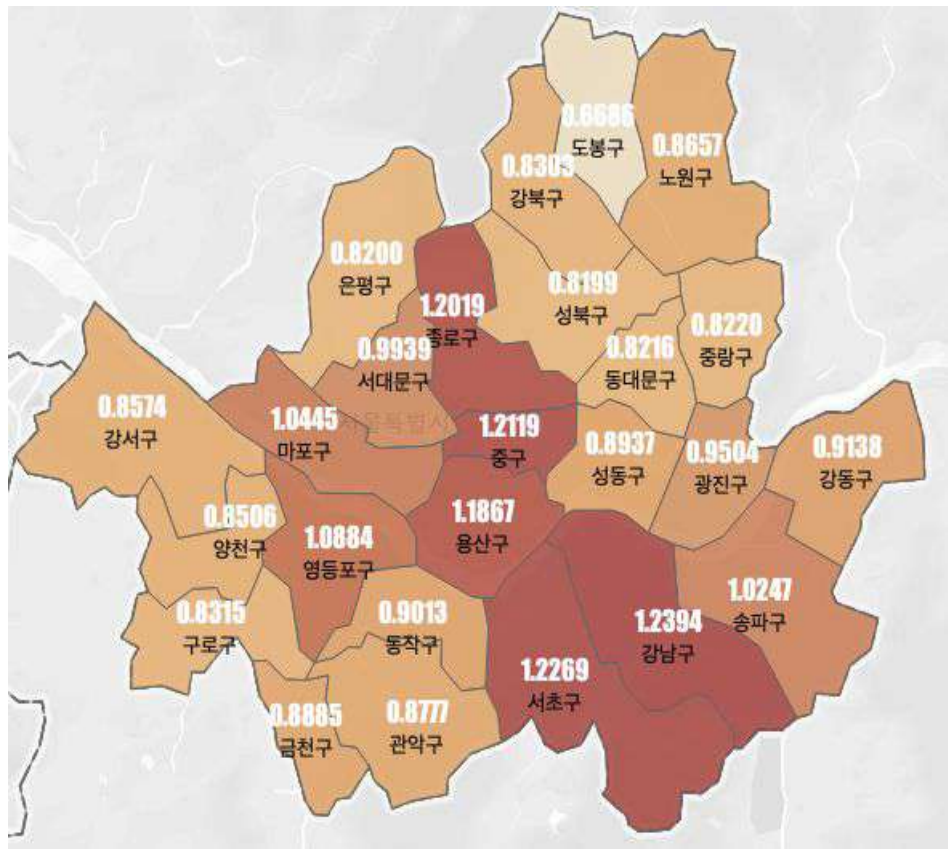
커피 지수가 높을수록 [     ] 다.





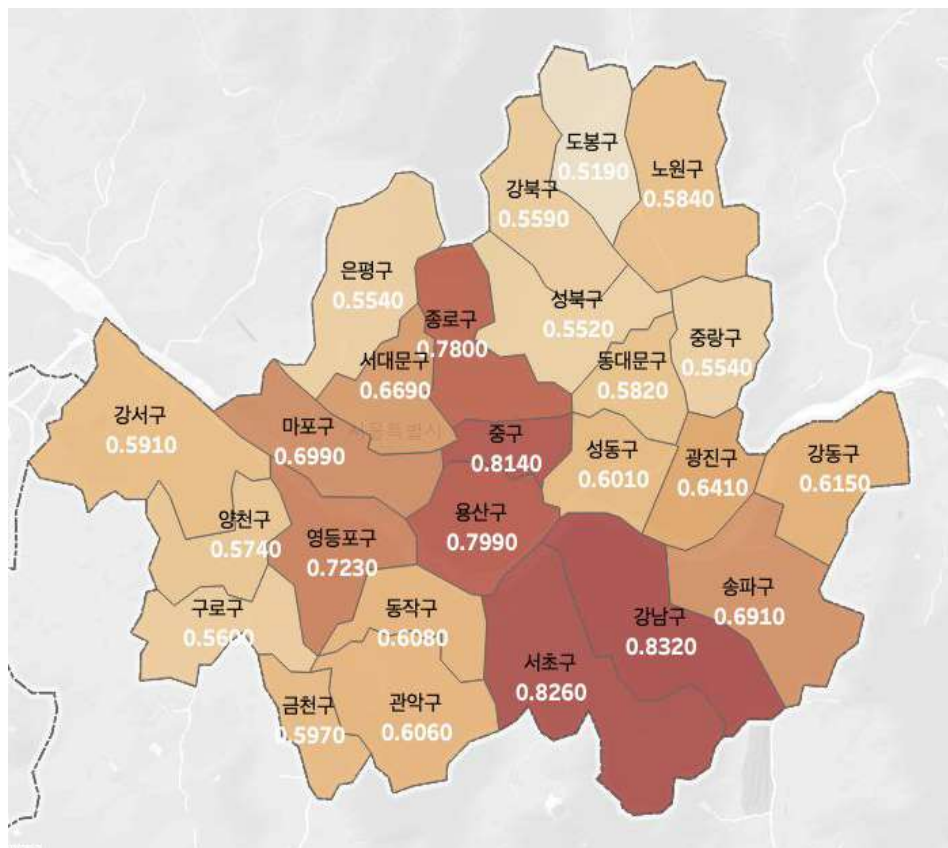
## 4-1. 커피지수 1 - 빅맥지수 관점

| 브랜드명 | 메가커피 | 맥다방 | 스타벅스 | 엔제리너스 | 이디야 | 커피빈 | 컴포즈커피 | 투썸플레이스 | 피스쿠피 | 콜바셋 | 할리스 | 합계   | 커피지수    |
|------|------|-----|------|-------|-----|-----|-------|--------|------|-----|-----|------|---------|
| 시군구명 |      |     |      |       |     |     |       |        |      |     |     |      |         |
| 강남구  | 21   | 17  | 88   | 1     | 44  | 50  | 19    | 42     | 12   | 12  | 16  | 322  | 1.23938 |
| 서초구  | 12   | 12  | 48   | 1     | 29  | 20  | 8     | 28     | 7    | 10  | 9   | 184  | 1.22693 |
| 중구   | 17   | 6   | 55   | 6     | 30  | 13  | 11    | 32     | 3    | 8   | 9   | 190  | 1.21187 |
| 용산구  | 6    | 8   | 25   | 1     | 14  | 4   | 1     | 9      | 4    | 2   | 4   | 78   | 1.18668 |
| 종로구  | 16   | 8   | 40   | 5     | 31  | 15  | 8     | 12     | 3    | 3   | 11  | 152  | 1.15762 |
| 영등포구 | 26   | 10  | 41   | 4     | 41  | 11  | 14    | 21     | 3    | 6   | 8   | 185  | 1.07286 |
| 마포구  | 27   | 11  | 35   | 3     | 34  | 7   | 15    | 25     | 1    | 3   | 10  | 171  | 1.03686 |
| 송파구  | 32   | 17  | 34   | 10    | 39  | 14  | 19    | 30     | 1    | 4   | 6   | 206  | 1.02469 |
| 합계   | 512  | 260 | 601  | 60    | 706 | 167 | 308   | 366    | 58   | 62  | 128 | 3230 | 1.00001 |
| 서대문구 | 17   | 6   | 22   | 2     | 20  | 2   | 12    | 9      | 5    | 0   | 8   | 103  | 0.99389 |
| 경진구  | 14   | 10  | 19   | 2     | 24  | 4   | 16    | 12     | 1    | 0   | 5   | 107  | 0.95044 |
| 강동구  | 22   | 10  | 17   | 1     | 26  | 4   | 12    | 12     | 1    | 2   | 3   | 110  | 0.91376 |
| 동작구  | 14   | 6   | 11   | 3     | 22  | 0   | 9     | 5      | 0    | 0   | 6   | 76   | 0.90128 |
| 관악구  | 18   | 11  | 12   | 4     | 28  | 2   | 13    | 8      | 3    | 1   | 6   | 106  | 0.90032 |
| 성동구  | 16   | 12  | 15   | 0     | 22  | 0   | 9     | 11     | 2    | 0   | 3   | 90   | 0.89374 |
| 금천구  | 13   | 8   | 13   | 2     | 23  | 2   | 16    | 11     | 0    | 0   | 2   | 90   | 0.88853 |
| 강서구  | 35   | 17  | 25   | 3     | 42  | 4   | 27    | 19     | 5    | 4   | 2   | 183  | 0.87580 |
| 노원구  | 28   | 14  | 14   | 2     | 29  | 3   | 8     | 11     | 2    | 1   | 3   | 115  | 0.86568 |
| 동대문구 | 18   | 7   | 10   | 1     | 25  | 3   | 10    | 9      | 0    | 1   | 0   | 84   | 0.86250 |
| 양천구  | 26   | 6   | 17   | 0     | 26  | 2   | 18    | 9      | 3    | 0   | 5   | 112  | 0.85059 |
| 구로구  | 28   | 13  | 14   | 1     | 29  | 2   | 19    | 11     | 1    | 4   | 5   | 127  | 0.83146 |
| 강북구  | 20   | 9   | 6    | 2     | 20  | 1   | 2     | 7      | 0    | 0   | 2   | 69   | 0.83027 |
| 종랑구  | 19   | 10  | 8    | 1     | 29  | 1   | 10    | 10     | 0    | 0   | 1   | 89   | 0.82201 |
| 은평구  | 24   | 10  | 13   | 5     | 26  | 1   | 15    | 6      | 0    | 1   | 3   | 104  | 0.82001 |
| 성북구  | 27   | 16  | 15   | 0     | 31  | 2   | 8     | 10     | 0    | 0   | 1   | 110  | 0.81994 |
| 도봉구  | 16   | 6   | 4    | 0     | 24  | 0   | 9     | 7      | 1    | 0   | 0   | 67   | 0.77011 |




## 4-1. 커피지수 2 - 자치구별 브랜드 비율 관점

|    | 시군구  | 메가커피  | 백다방   | 스타벅스  | 엔제리너스 | 이디야   | 커피빈   | 컴포즈커피 | 투썸플레이스 | 파스쿠찌  | 몰바셋   | 할리스   | 합계    |
|----|------|-------|-------|-------|-------|-------|-------|-------|--------|-------|-------|-------|-------|
| 0  | 강남구  | 0.014 | 0.016 | 0.273 | 0.003 | 0.090 | 0.172 | 0.012 | 0.130  | 0.034 | 0.038 | 0.050 | 0.832 |
| 1  | 서초구  | 0.014 | 0.020 | 0.261 | 0.005 | 0.104 | 0.121 | 0.009 | 0.152  | 0.035 | 0.056 | 0.049 | 0.826 |
| 2  | 중구   | 0.019 | 0.010 | 0.289 | 0.032 | 0.104 | 0.075 | 0.012 | 0.168  | 0.015 | 0.043 | 0.047 | 0.814 |
| 3  | 용산구  | 0.016 | 0.031 | 0.321 | 0.013 | 0.118 | 0.057 | 0.003 | 0.115  | 0.047 | 0.027 | 0.051 | 0.799 |
| 4  | 종로구  | 0.022 | 0.016 | 0.263 | 0.033 | 0.135 | 0.110 | 0.011 | 0.079  | 0.018 | 0.021 | 0.072 | 0.780 |
| 5  | 영등포구 | 0.030 | 0.016 | 0.222 | 0.022 | 0.147 | 0.065 | 0.016 | 0.114  | 0.015 | 0.033 | 0.043 | 0.723 |
| 6  | 마포구  | 0.033 | 0.019 | 0.205 | 0.018 | 0.131 | 0.046 | 0.018 | 0.146  | 0.006 | 0.019 | 0.058 | 0.699 |
| 7  | 송파구  | 0.033 | 0.025 | 0.165 | 0.049 | 0.125 | 0.075 | 0.019 | 0.146  | 0.005 | 0.020 | 0.029 | 0.691 |
| 8  | 서대문구 | 0.035 | 0.017 | 0.214 | 0.019 | 0.128 | 0.021 | 0.025 | 0.087  | 0.045 | 0.000 | 0.078 | 0.669 |
| 9  | 평균   | 0.036 | 0.026 | 0.170 | 0.018 | 0.156 | 0.042 | 0.021 | 0.107  | 0.015 | 0.015 | 0.038 | 0.644 |
| 10 | 광진구  | 0.028 | 0.028 | 0.178 | 0.019 | 0.148 | 0.041 | 0.032 | 0.112  | 0.008 | 0.000 | 0.047 | 0.641 |
| 11 | 강동구  | 0.042 | 0.027 | 0.155 | 0.009 | 0.156 | 0.040 | 0.023 | 0.109  | 0.008 | 0.019 | 0.027 | 0.615 |
| 12 | 동작구  | 0.039 | 0.024 | 0.145 | 0.039 | 0.191 | 0.000 | 0.025 | 0.066  | 0.000 | 0.000 | 0.079 | 0.608 |
| 13 | 관악구  | 0.036 | 0.031 | 0.113 | 0.038 | 0.174 | 0.021 | 0.026 | 0.075  | 0.026 | 0.009 | 0.057 | 0.606 |
| 14 | 성동구  | 0.037 | 0.040 | 0.167 | 0.000 | 0.161 | 0.000 | 0.021 | 0.122  | 0.020 | 0.000 | 0.033 | 0.601 |
| 15 | 금천구  | 0.030 | 0.027 | 0.144 | 0.022 | 0.169 | 0.024 | 0.037 | 0.122  | 0.000 | 0.000 | 0.022 | 0.597 |
| 16 | 강서구  | 0.040 | 0.028 | 0.137 | 0.016 | 0.152 | 0.024 | 0.031 | 0.104  | 0.025 | 0.023 | 0.011 | 0.591 |
| 17 | 노원구  | 0.051 | 0.037 | 0.122 | 0.017 | 0.166 | 0.029 | 0.015 | 0.096  | 0.016 | 0.009 | 0.026 | 0.584 |
| 18 | 동대문구 | 0.045 | 0.025 | 0.119 | 0.012 | 0.197 | 0.040 | 0.025 | 0.107  | 0.000 | 0.012 | 0.000 | 0.582 |
| 19 | 양천구  | 0.049 | 0.016 | 0.152 | 0.000 | 0.153 | 0.020 | 0.034 | 0.080  | 0.025 | 0.000 | 0.045 | 0.574 |
| 20 | 구로구  | 0.046 | 0.031 | 0.110 | 0.008 | 0.150 | 0.018 | 0.032 | 0.087  | 0.007 | 0.032 | 0.039 | 0.560 |
| 21 | 강북구  | 0.061 | 0.039 | 0.087 | 0.029 | 0.191 | 0.016 | 0.006 | 0.101  | 0.000 | 0.000 | 0.029 | 0.559 |
| 22 | 은평구  | 0.049 | 0.029 | 0.125 | 0.048 | 0.165 | 0.011 | 0.030 | 0.058  | 0.000 | 0.010 | 0.029 | 0.554 |
| 23 | 중랑구  | 0.045 | 0.034 | 0.090 | 0.011 | 0.215 | 0.012 | 0.024 | 0.112  | 0.000 | 0.000 | 0.011 | 0.554 |
| 24 | 성북구  | 0.051 | 0.044 | 0.136 | 0.000 | 0.186 | 0.020 | 0.015 | 0.091  | 0.000 | 0.000 | 0.009 | 0.552 |
| 25 | 도봉구  | 0.050 | 0.027 | 0.060 | 0.000 | 0.236 | 0.000 | 0.028 | 0.104  | 0.014 | 0.000 | 0.000 | 0.519 |







# Part 3

## 결과물 및 결론

# 1. 서울시 데이터셋 시각화

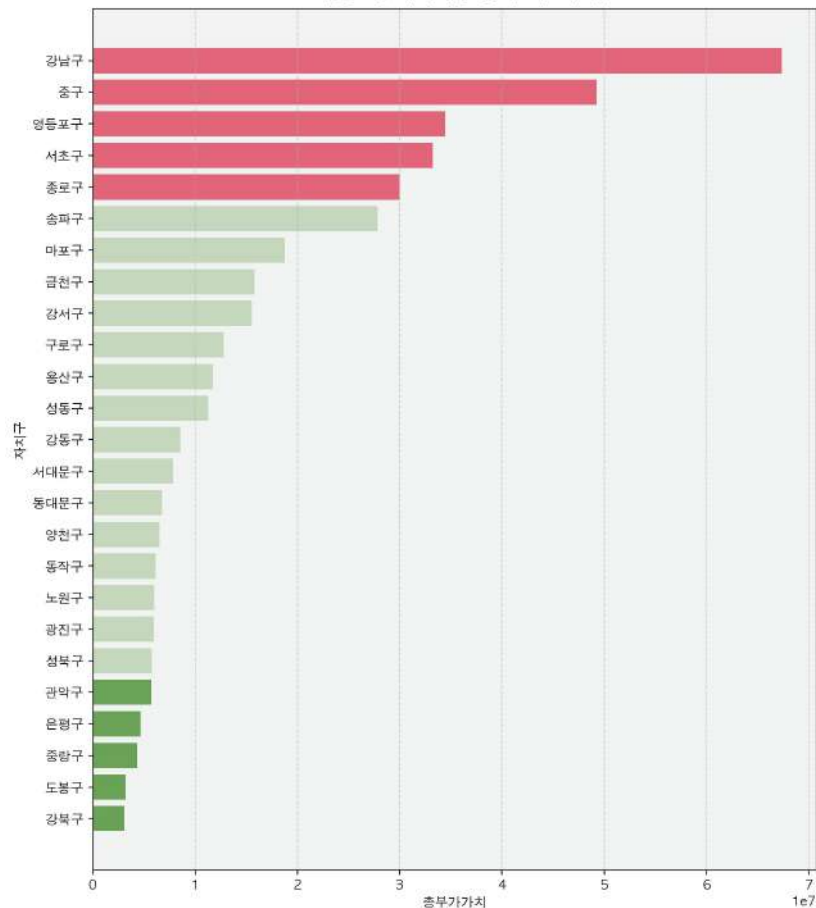
## 경제력 및 총부가가치 관련

1. 경제 : 총부가가치, 재정자립도
2. 인구 : 생활 인구, 이동 인구
3. 회사 : 사업체 수, 종업원 수, 대규모상점
4. 토지 : 공시지가

# 총부가가치

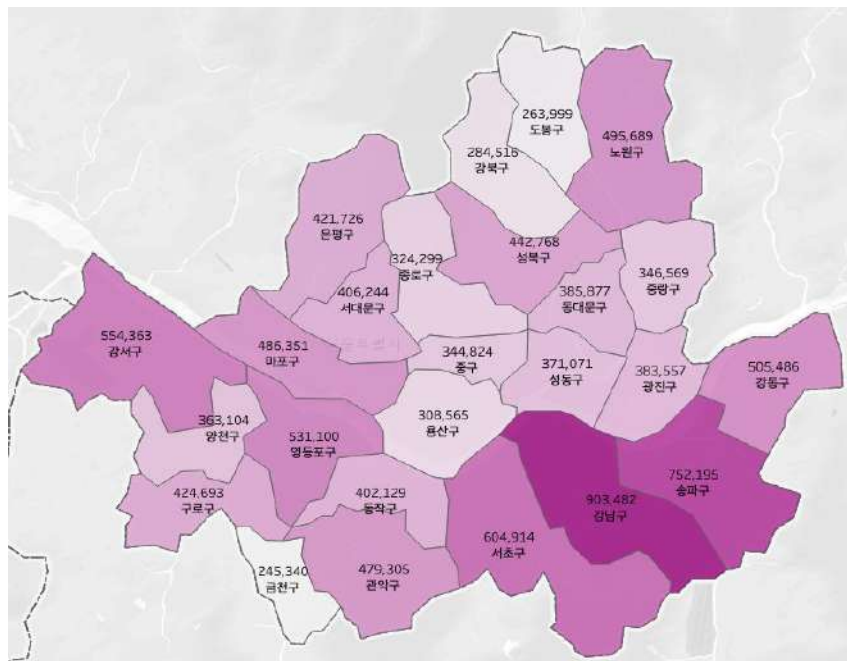
- 생산과정에서 부가된 가치생산과정을 거치면서 새롭게 더해진 가치인 순부가가치에 감가상각을 합한것
- 기업의 부가가치를 전국가적으로 집계한 것은 생산국민소득과 연결되므로 자치구별 총부가가치를 보면 각 지역의 구매력을 알아볼 수 있을 것이라 판단

서울 자치구별 총부가가치



# 생활 인구

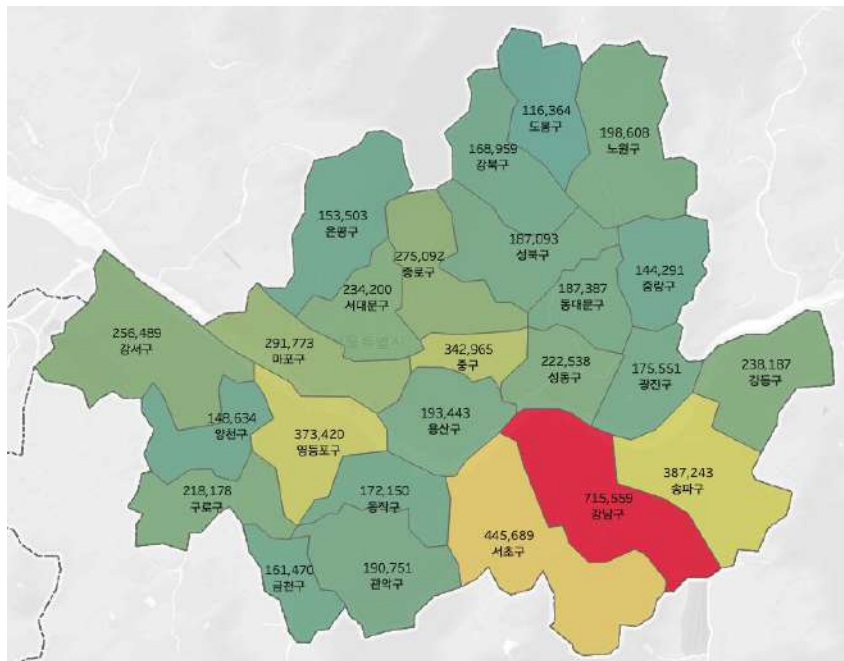
- 통신데이터로 특정 시점에 개인이 위치한 지역을 집계한 ‘현주인구’를 의미



생활인구

# 유동 인구

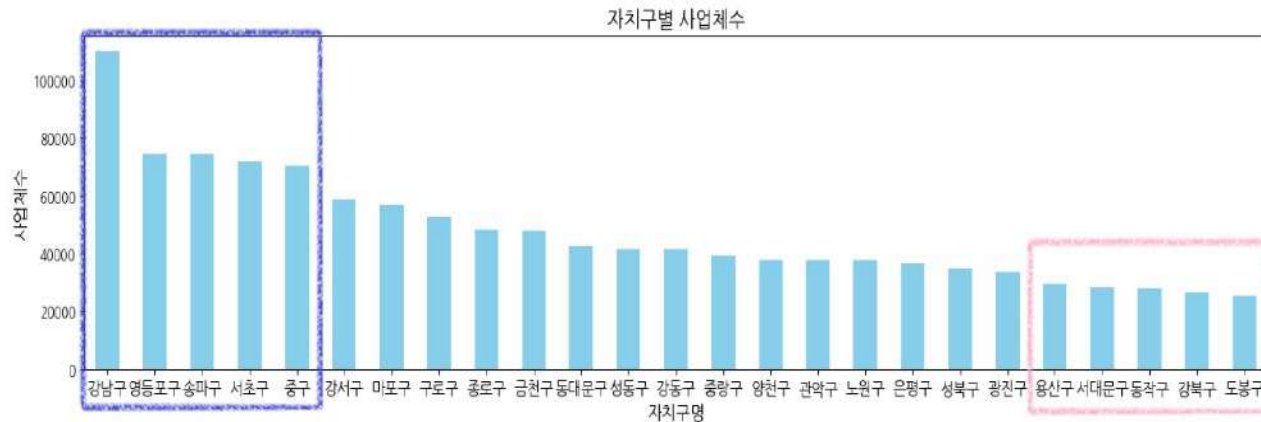
- 자치구별 일일 최대 이동인구수를 기준으로 유동인구를 정의



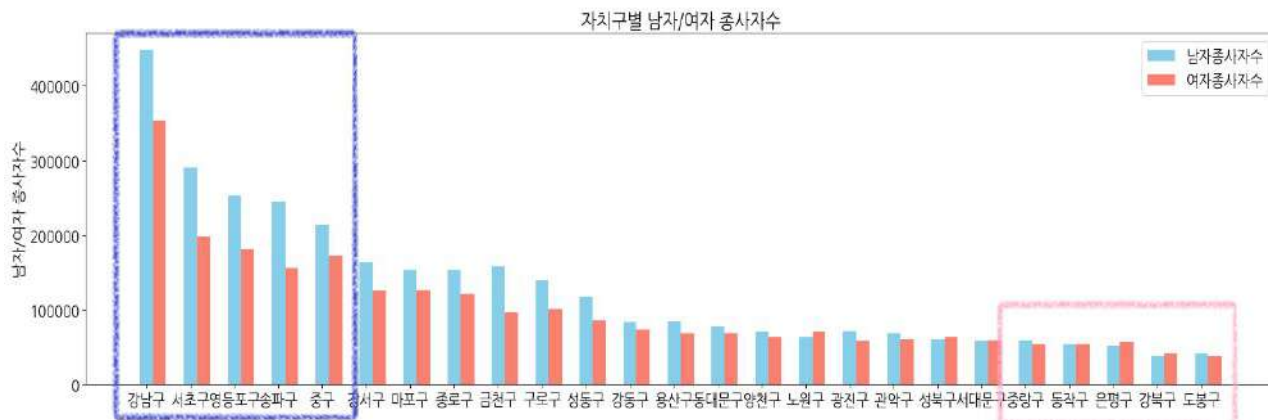
유동인구



# 사업체 수

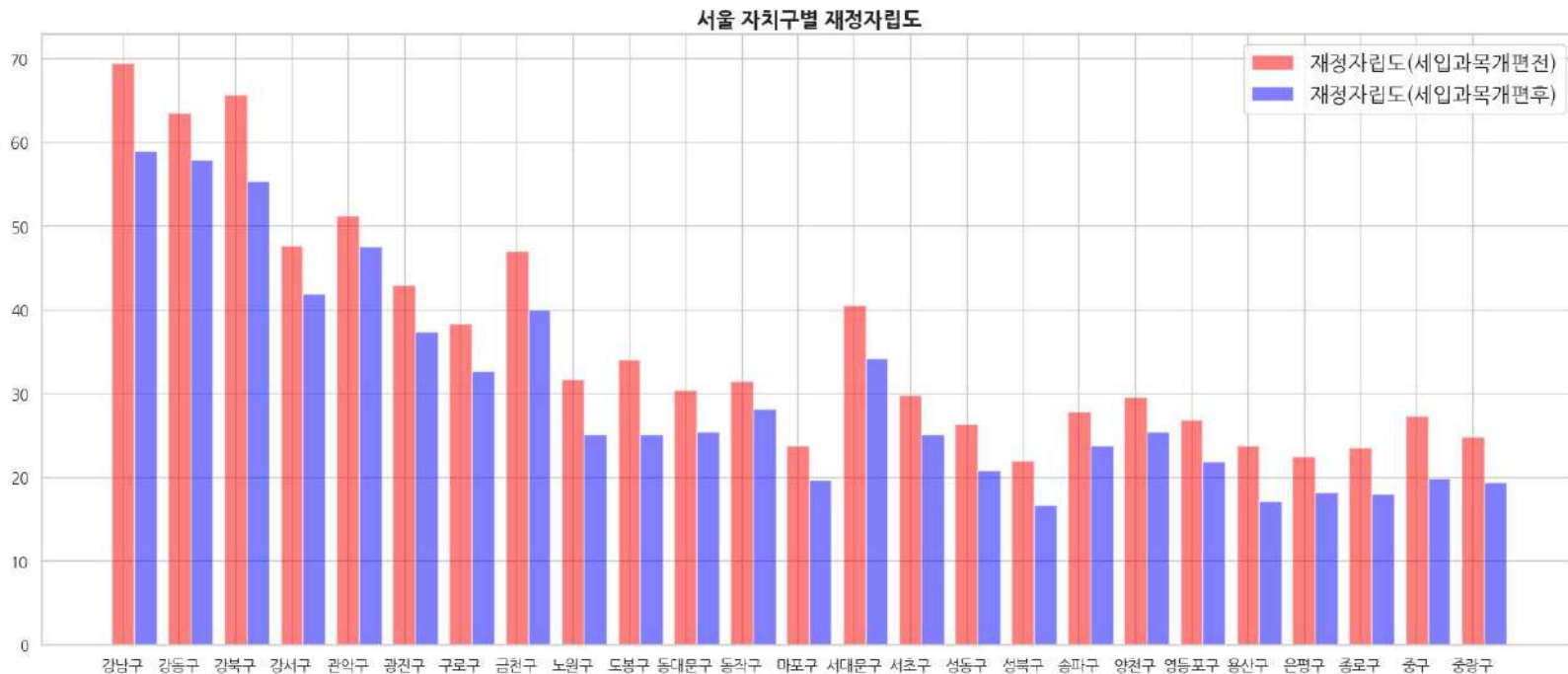


# 종사자 수



# 재정자립도

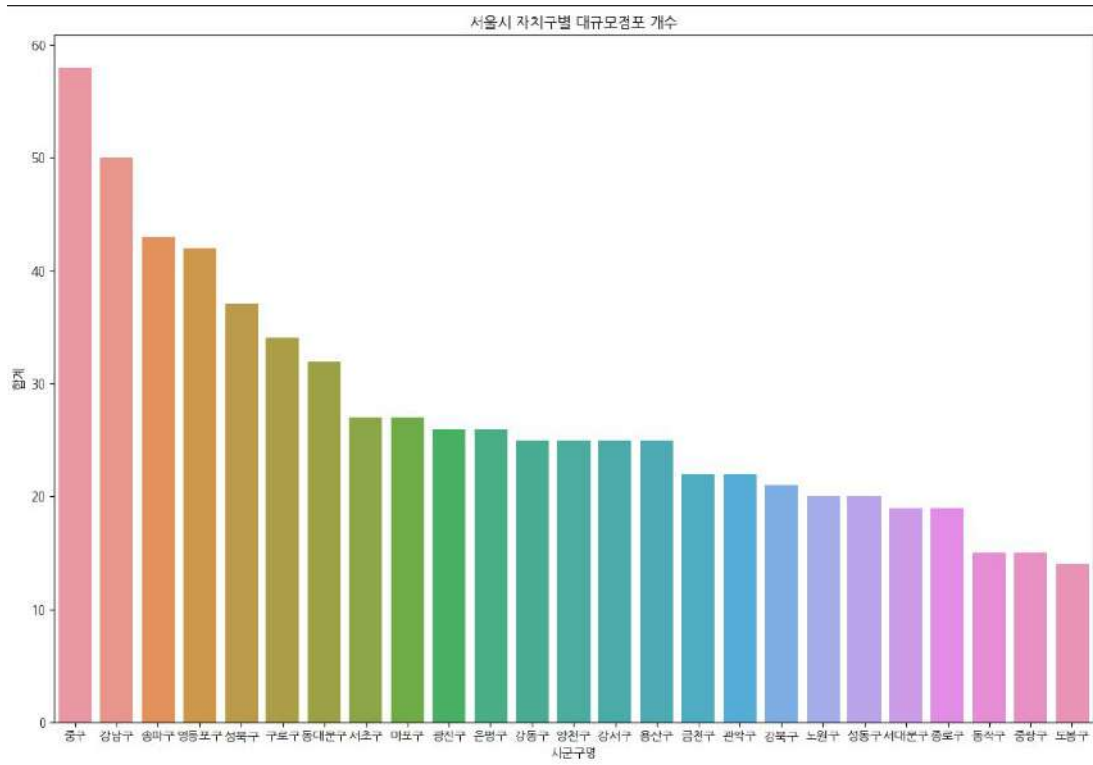
- 재정수입의 자체 충당 능력을 나타내는 세입분석지표
- 재정자립도가 높을수록 재정운영의 자립능력이 우수함을 의미



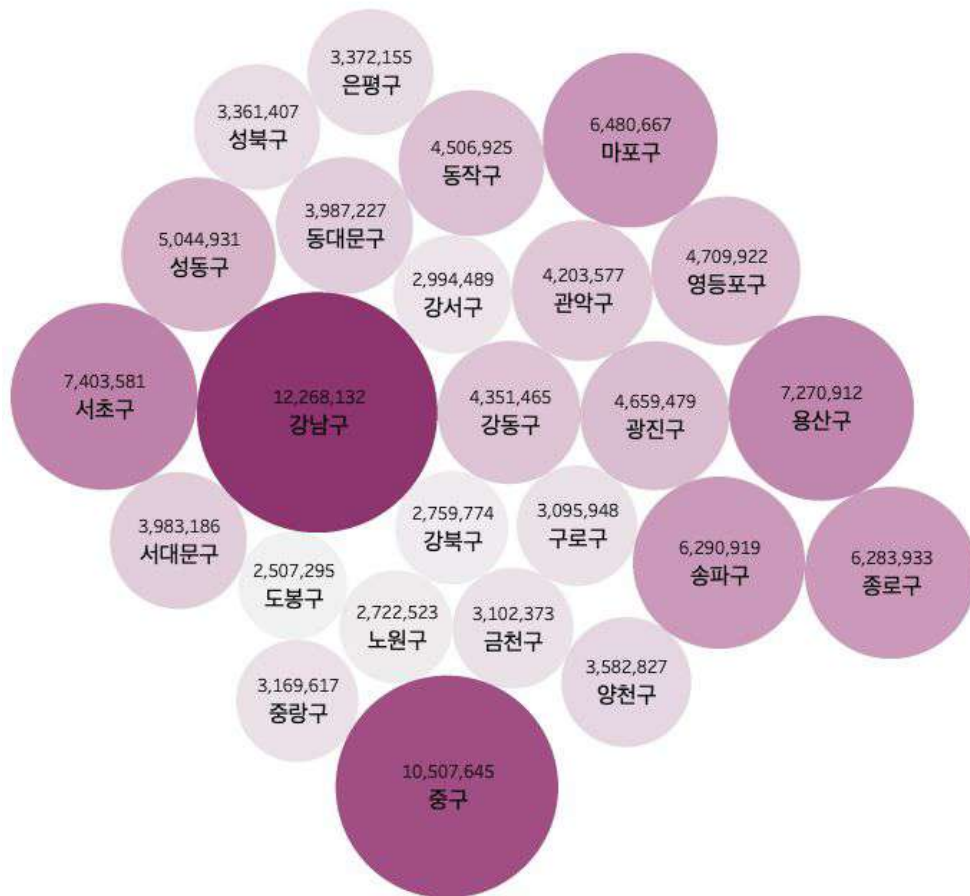
# 대규모 점포

- 항목은 백화점, 쇼핑센터, 복합쇼핑몰, 대형마트, 시장, 전문점, 그 밖의 대규모점포 와 구분없음으로 구성됨

- 대규모 점포의 개수가 많을수록 상권이 발달한 지역이므로 자치구의 구매력을 반영할 수 있음



# 공시지가



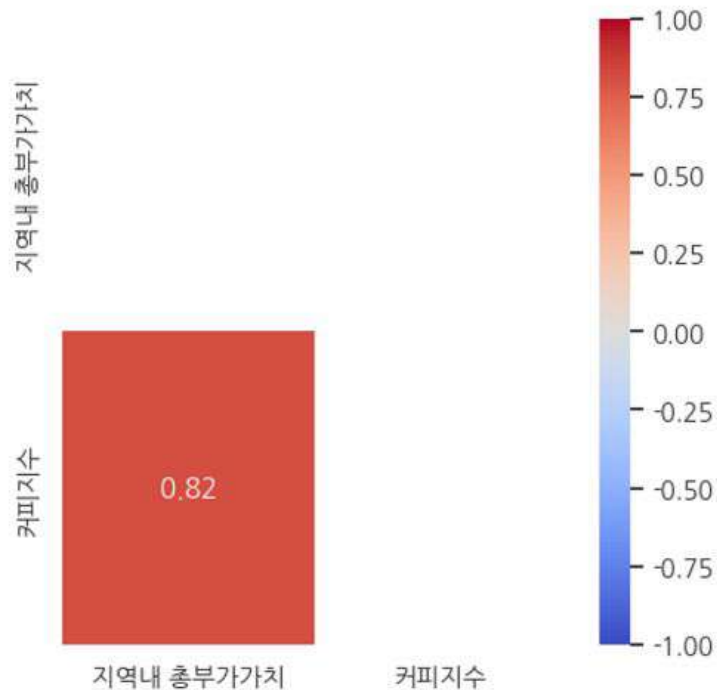
## 2. 커피지수와 상관관계

(+) / (-) 상관관계 분석

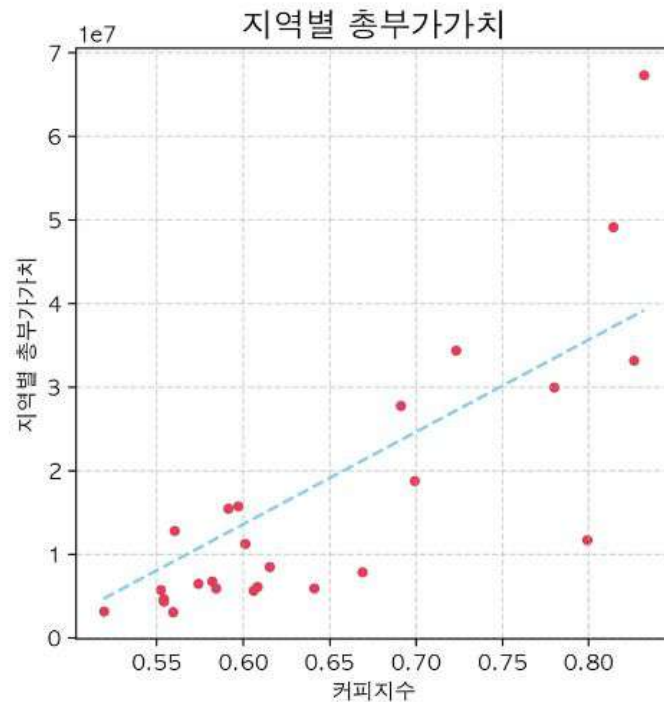


# 총부가가치

[커피지수 1]

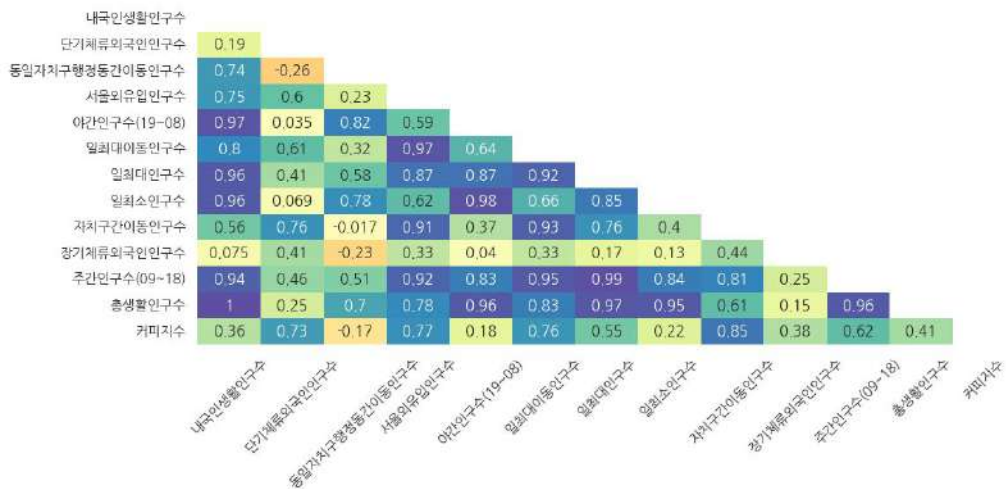


[커피지수 2]

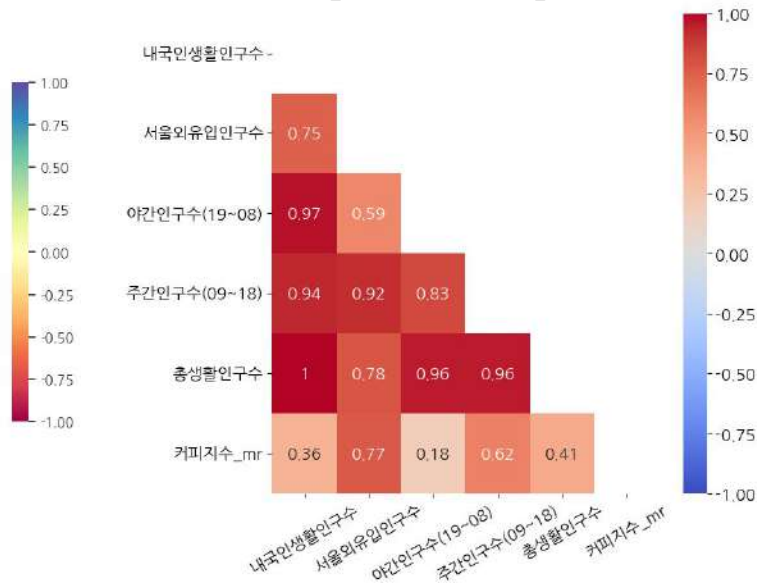


# 인구

[커피지수 1]



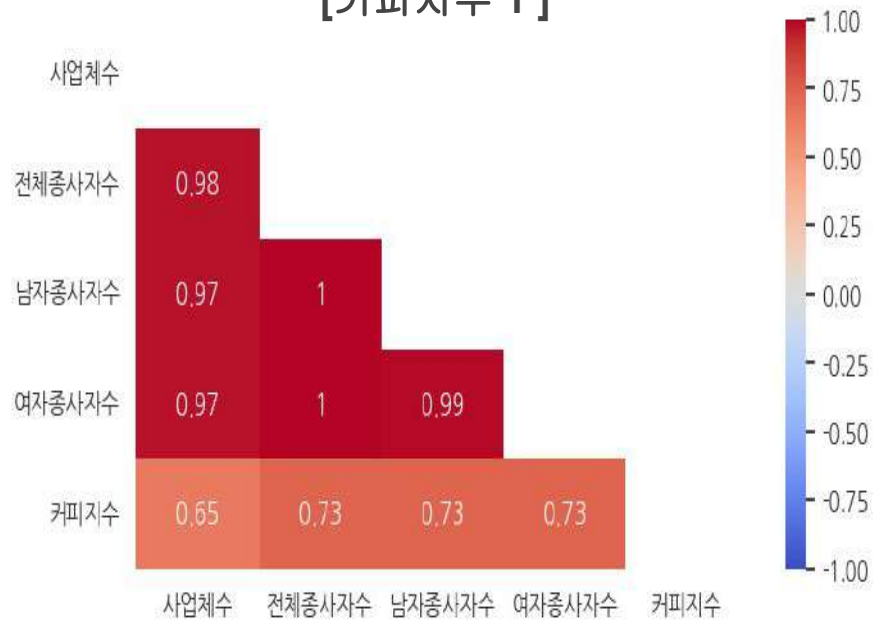
[커피지수 2]



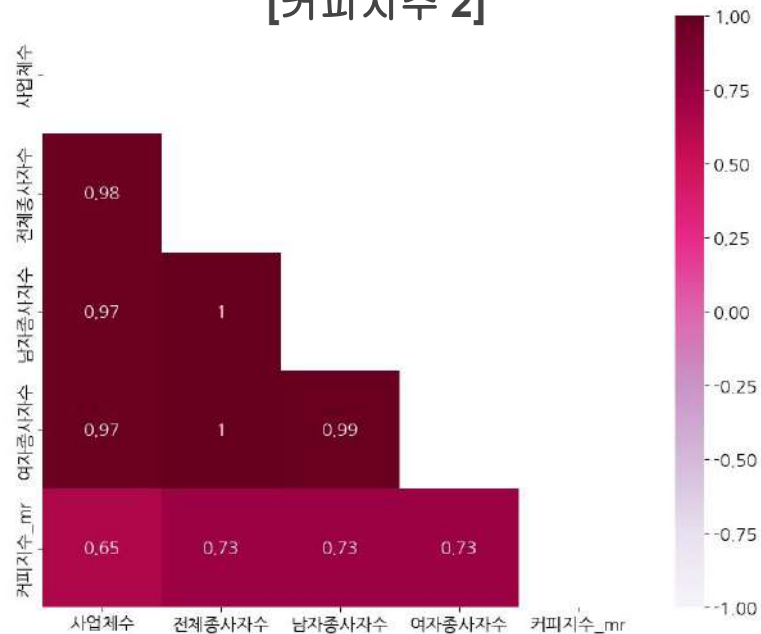


# 사업체 수, 종사자 수

[커피지수 1]

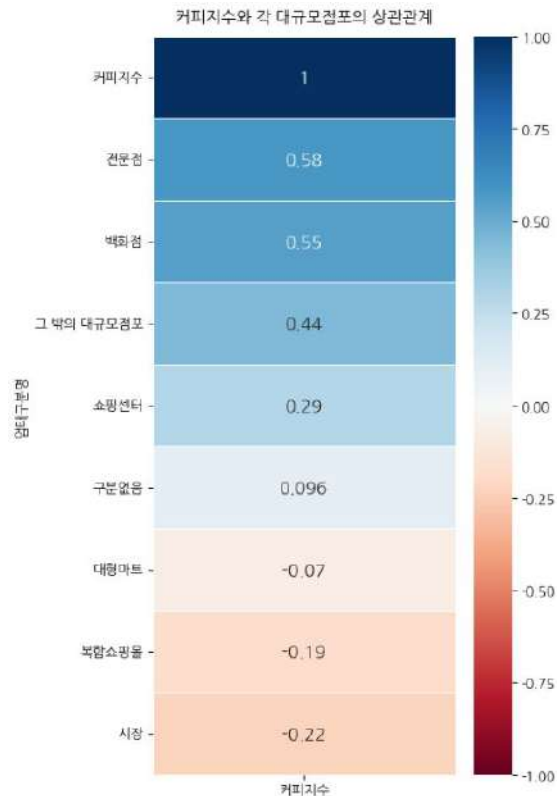


[커피지수 2]

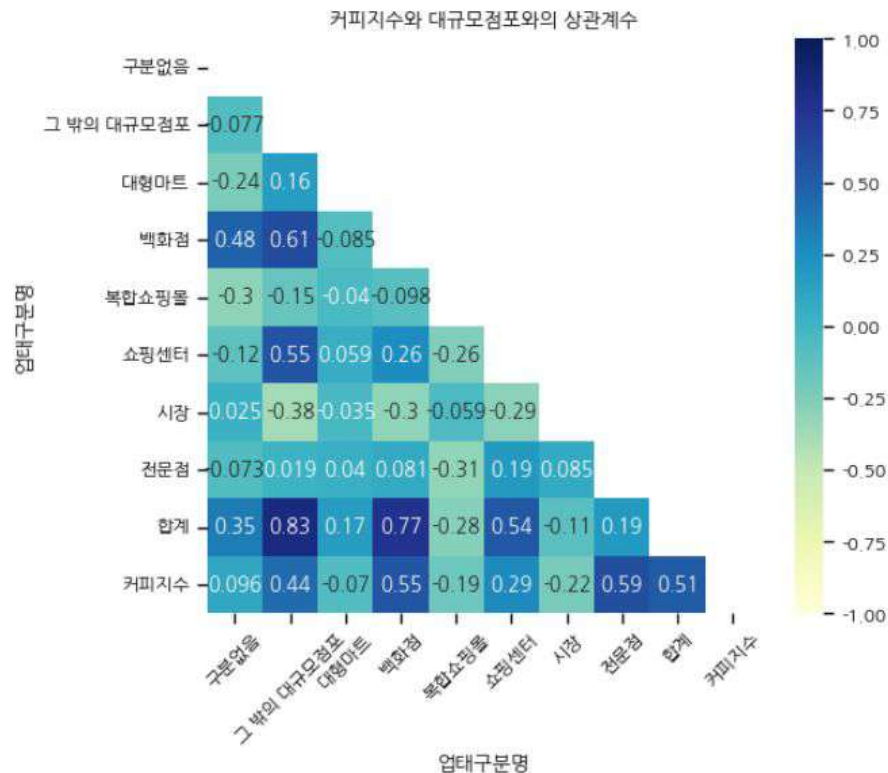


# 대규모점포

[커피지수 1]

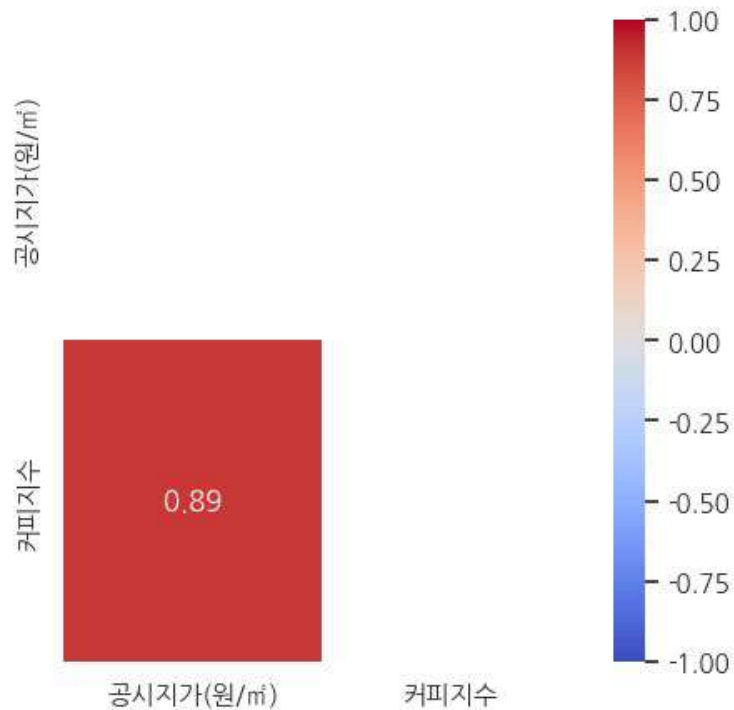


[커피지수 2]

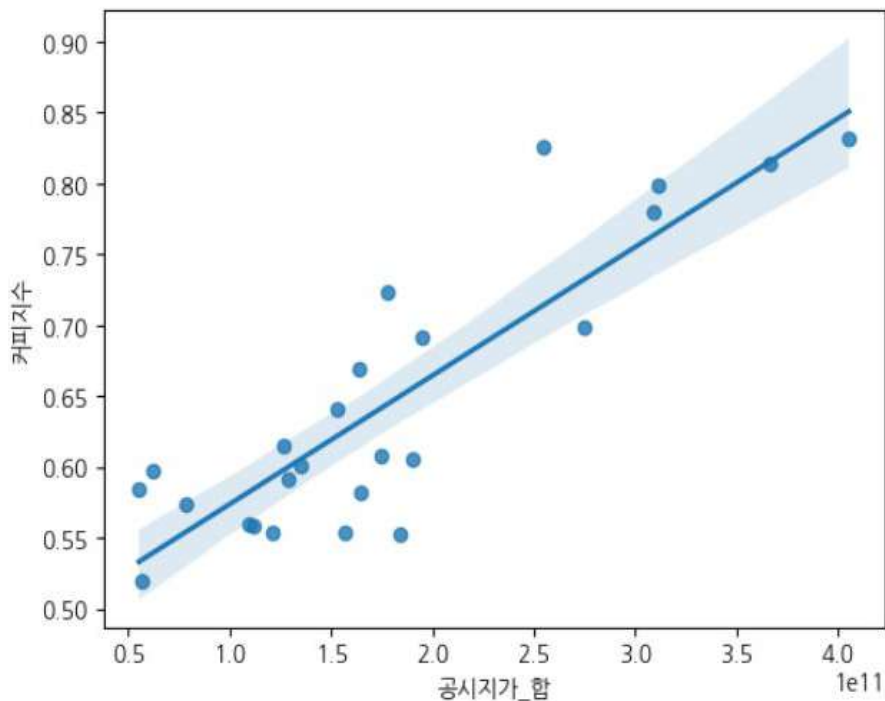


# 공시지가

[커피지수 1]



[커피지수 2]



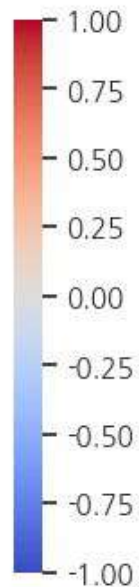
# 재정자립도

[커피지수 1]

재정자립도(세입과목개편전)

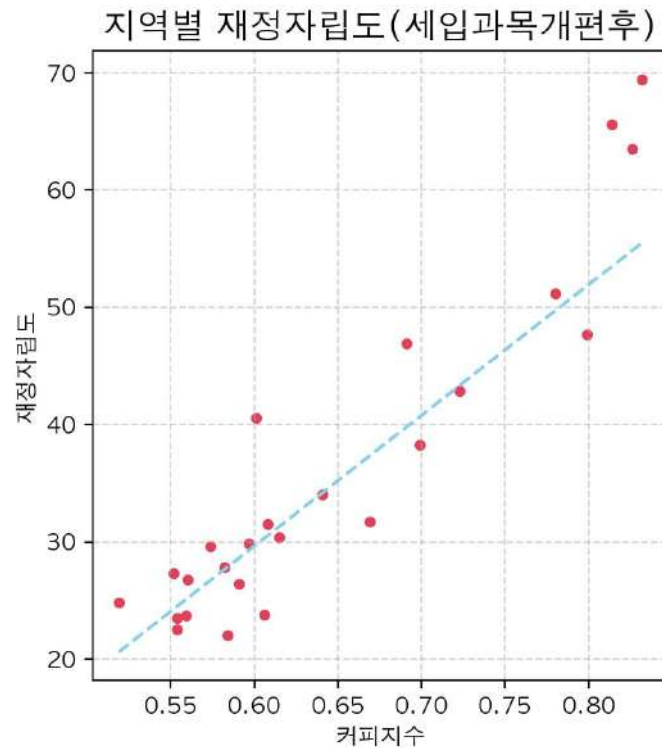
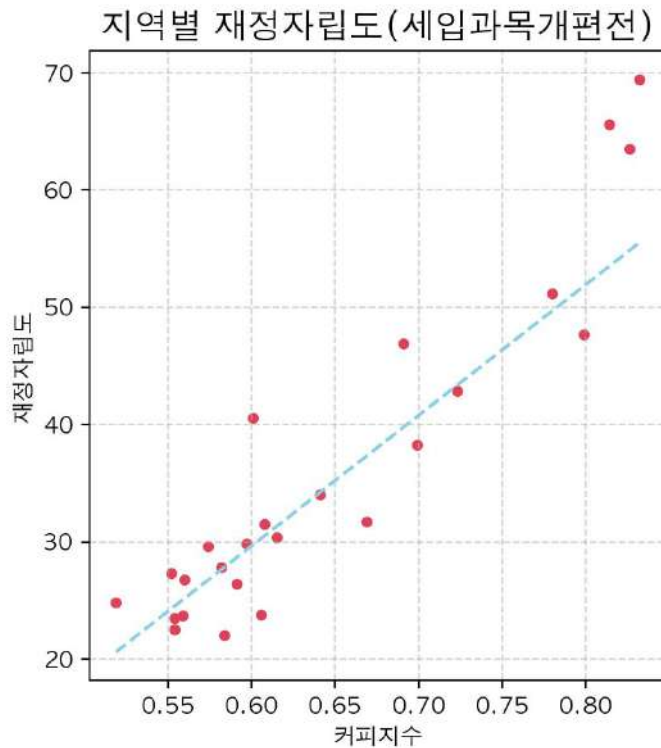
재정자립도(세입과목개편후)

커피지수



# 재정자립도

[커피지수 2]



### 3. 대표 커피지수 결정 기준

#### 통계적 검정

‘커피지수’와 ‘재정자립도’ 그리고 ‘총부가가치’와의 회귀 분석을 통해 설명력이 높은 것



# 재정자립도

## [커피지수 1]

$$\widehat{\text{재정자립도}} = -50.7996 + 84.5571 * \text{커피지수} + \varepsilon$$

|           | coef     | std err | t      | P> t  | [0.025  | 0.975]  |
|-----------|----------|---------|--------|-------|---------|---------|
| Intercept | -50.7996 | 6.176   | -8.225 | 0.000 | -63.577 | -38.023 |
| x         | 84.5571  | 6.377   | 13.259 | 0.000 | 71.364  | 97.750  |

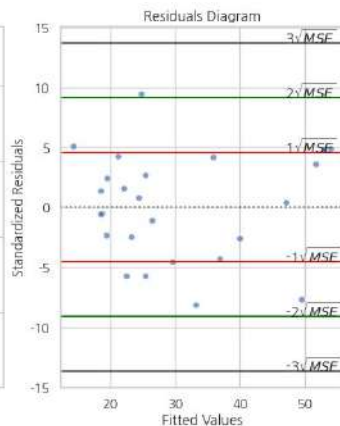
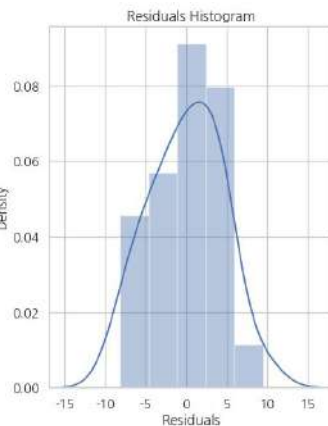
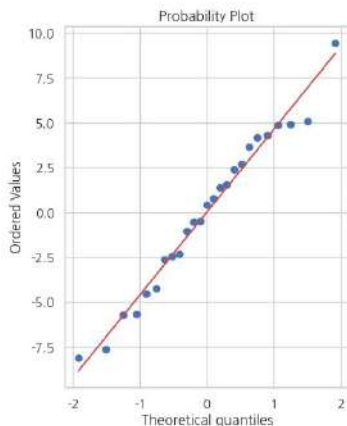
|               |       |     |       |
|---------------|-------|-----|-------|
| R-squared     | 0.884 | AIC | 148.7 |
| Durbin-Watson | 2.028 | BIC | 151.1 |

## [커피지수 2]

$$\widehat{\text{재정자립도}} = -50.6514 + 134.4070 * \text{커피지수} + \varepsilon$$

|           | coef     | std err | t      | P> t  | [0.025  | 0.975]  |
|-----------|----------|---------|--------|-------|---------|---------|
| Intercept | -50.6514 | 6.952   | -7.286 | 0.000 | -65.033 | -36.270 |
| 커피지수      | 134.4070 | 10.658  | 12.611 | 0.000 | 112.359 | 156.455 |

|               |       |     |       |
|---------------|-------|-----|-------|
| R-squared     | 0.874 | AIC | 154.4 |
| Durbin-Watson | 1.978 | BIC | 156.8 |





# 지역내 총 부가가치

## [커피지수 1]

지역내 총부가가치 =  $(-7e + 07) + (8.989e + 07) * \text{커피지수} + \varepsilon$

|           | coef      | std err  | t      | P> t  | [0.025    | 0.975]    |
|-----------|-----------|----------|--------|-------|-----------|-----------|
| Intercept | -7e+07    | 1.28e+07 | -5.479 | 0.000 | -9.64e+07 | -4.36e+07 |
| x         | 8.989e+07 | 1.32e+07 | 6.814  | 0.000 | 6.26e+07  | 1.17e+08  |

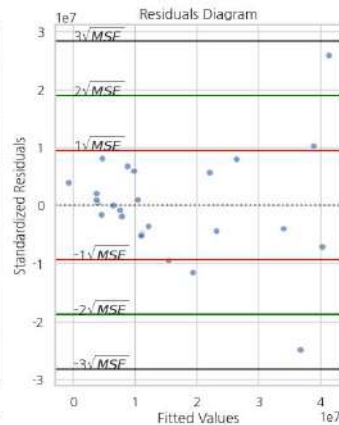
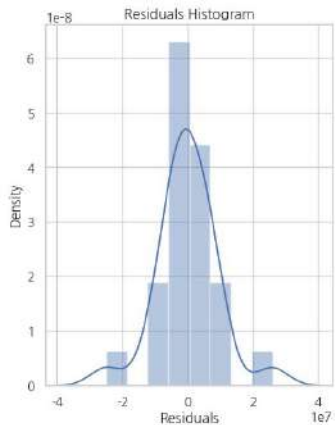
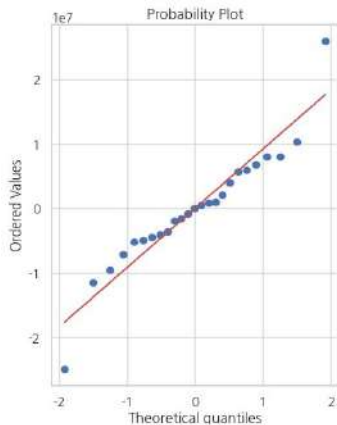
|               |       |     |       |
|---------------|-------|-----|-------|
| R-squared     | 0.669 | AIC | 875.8 |
| Durbin-Watson | 2.012 | BIC | 878.3 |

## [커피지수 2]

지역내 총부가가치 =  $(-7.02e + 07) + (1.388e + 08) * \text{커피지수} + \varepsilon$

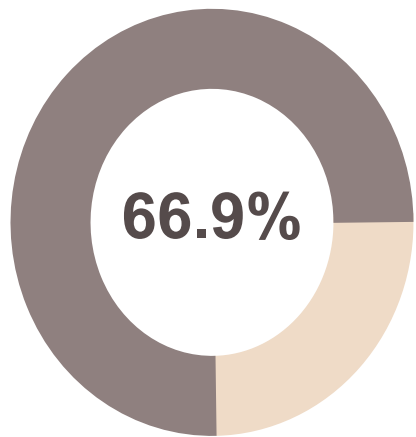
|           | coef      | std err  | t      | P> t  | [0.025    | 0.975]    |
|-----------|-----------|----------|--------|-------|-----------|-----------|
| Intercept | -7.02e+07 | 1.29e+07 | -5.435 | 0.000 | -9.69e+07 | -4.35e+07 |
| 커피지수      | 1.338e+08 | 1.98e+07 | 6.754  | 0.000 | 9.28e+07  | 1.75e+08  |

|               |       |     |       |
|---------------|-------|-----|-------|
| R-squared     | 0.665 | AIC | 876.1 |
| Durbin-Watson | 2.011 | BIC | 878.6 |

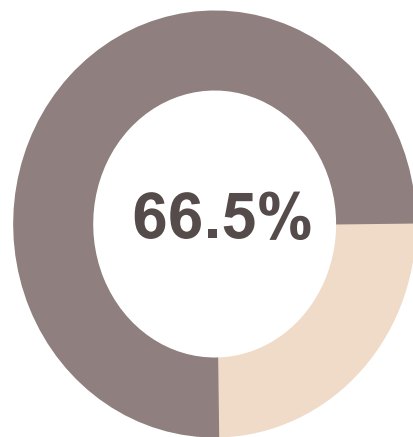


총부가가치 기준

## 대표 커피지수 선정



커피지수1



커피지수2

# 우리팀의 대표 커피지수



## 커피지수 1

$$\text{커피지수1} = \frac{\text{지역구별 평균 1ml당 단가}}{\text{전체 평균 1ml당 단가}}$$

# 커피지수 활용 방향

## 경제 지표로써 활용

- 커피지수를 통해 어느 지역구의 구매력이 대한민국 전국 평균에 비해 저평가되었는지, 고평가되었는지 판단할 때 유용하게 쓰일수 있다
- 커피만으로도 각 지역의 물가수준을 간접적으로 가늠해볼 수 있는 수단이 될 수 있다.

# 커피지수 한계점

사회현상에 대한 모든 변수를 고려하지 않아 “불완전성”을 가진다

- 단순히 가격과 매장수외에 지역마다 각각 다른 특징들이 있음
- 시간적, 능력적 한계로 다른 다양한 변수를 분석하지 못함
- 커피지수로 지역별 구매력에 대해서 간략히 파악을 하는 것은 가능하지만, 다른 지수나 자료가 보완될 필요가 있음

The background image shows a rustic interior space. A dark wooden table is in the foreground, surrounded by yellow chairs. The wall is made of exposed brick. On the left, there's a wooden shelf with a glass lantern and some boxes. On the right, two wooden skis are leaning against the wall. The lighting is warm and ambient, with a few pendant lights hanging from the ceiling.

# Part 4

## 부록

# 회고 및 개선점

| 구분 | 팀원명 | 회고  |
|----|-----|---|
| 팀원 | 김강섭 | 무엇을 할 것인지, 왜 하는 것인지, 누가 할 것인지 명확하게 결정한 후 진행하지 못 한 부분이 있었다.  |
| 팀원 | 김미리 | 데이터 전처리의 중요성을 체감한 시간이었다.<br>(번거롭게 해드려서 죄송했어요 흑흑...ㅠㅠ)       |
| 팀원 | 박민정 | 분석에 시작하기 전에 분석목표를 명확히 정하는 게 너무 중요한 거 같다.                    |
| 팀장 | 서인혁 | 데이터 전처리가 너무 중요했다.. ,팀원들간의 의사소통이 중요했다..                      |
| 팀원 | 신하진 | 코드 정리하면서 작성할 필요가 있다.<br>회의에서 정확한 목표를 설정하고 작업에 들어가는 것이 중요하다. |
| 팀원 | 이호연 | 문제 정의의 필요성을 절실히 느꼈다.  |





**부록** : 우리가 시도해본 것들...

저가형 대비 고가형

- 동식

$$\text{커피지수} = \sum \frac{\text{가충치 반영된 지역구별 국가형 틀렌드 매장}}{\text{가충치 반영된 지역구별 국가형 틀렌드 매장}}$$

가충치: 원/ml 최저가를 1로 했을 때

- 가증치, 일

| 커피빈    | 가격(원) | 기본사이즈(ml) | 결과물   |
|--------|-------|-----------|---|
| 콜바셋    | 5000  | 355       | 전국 시<br>high_pr<br>low_pr<br>df_sido<br>브랜드 |
| 할리스    | 4700  | 360       |   |
| 스타벅스   | 4500  | 354       |   |
| 투썸플레이스 | 4500  | 355       |   |
| 연재리너스  | 4500  | 355       |   |
| 파스쿠찌   | 4500  | 355       |   |
| 이디야    | 3200  | 385       |   |
| 백다방    | 1500  | 384       |   |
| 메가커피   | 1500  | 400       |   |
| 컴포즈커피  | 1500  | 567       |   |
|        |       | 567       |   |

## 결과물

## 전국 시도별 커피지수

```
high_price = (df_sido['커피믹스'] * 5.32 + df_sido['루블렛'] * 4.94 + df_sido['할라소'] * 4.81
              + df_sido['스타벅스'] + df_sido['푸링클레머슨'] + df_sido['연희로비스트']) * 4.79 + df_sido['파도']
low_price = df_sido['알로즈커피'] + df_sido['베터커피'] + df_sido['맥다양'] * 1.42 + df_sido['이디야'] * 3.1

df_sido['커피지수'] = high_price / low_price
```

| 시도명     | 프랜차이즈 | 캐주얼리 | 백화점 | 스타벅스 | 연세대학교 | 이디야 | 커피빈 | 립모즈커피 | 투썸플레이스 |
|---------|-------|------|-----|------|-------|-----|-----|-------|--------|
| 대구광역시   | 30    | 57   | 79  | 21   | 71    | 1   | 34  |       |        |
| 대전광역시   | 45    | 77   | 61  | 12   | 45    | 1   |     |       |        |
| 경상북도    | 39    | 37   | 54  | 19   | 70    |     |     |       |        |
| 제주특별자치도 | 33    | 32   | 26  | 0    |       |     |     |       |        |
| 서울특별시   | 512   | 260  | 604 |      |       |     |     |       |        |
| 충청남도    | 95    |      |     |      |       |     |     |       |        |

커피지수 구하기

- 피지수 구하기**
- 매장별 점포수의 증감 현황을 파악해본다(프랜차이즈, 독립점포)
  - 톱 10 매장 증감 현황을 보니, 커피빈, 언제리너스, 할리스총 3개사가 매출 비중이 높고, 감소 추세인 곳은 7곳이다.
  - 매장들의 증감 크기에 따라 각각의 상수를 곱하여 인자를 준비.
  - (증가 추세 매장의 합 / 감소 추세 매장의 합) 식으로? 또는 몇몇 별첨으로?
- | E         | C    | D  | E    | F    | G    | H    | I    | J    |
|-----------|------|----|------|------|------|------|------|------|
| 매장별 변화율   | \$45 | 32 | 1262 | 116  | 483  | -55  | 2875 | 3005 |
| 에디나 변화율   | 577  | 45 | 1378 | 125  | 373  | 68.9 | 0.07 |      |
| 스타벅스 변화율  |      |    | 1141 | 121  | 554  | -71  | 2651 |      |
| 엔제리너스 변화율 |      |    | 1262 | 116  | 483  | -55  | 2875 |      |
| 할리스 변화율   |      |    | 116  | 483  | -55  | 2875 | 3005 |      |
| 커피빈 변화율   |      |    | 483  | -55  | 2875 | 3005 |      |      |
| 에디나 변화율   |      |    | 68.9 | 0.07 |      |      |      |      |

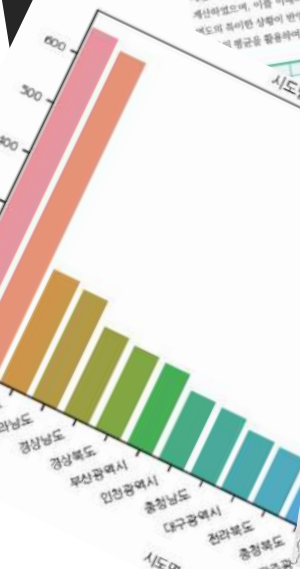
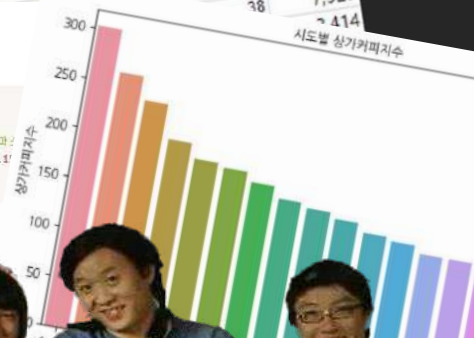
[illegible]


### 3 커피지수4

- ### 3 커피지수4
- 매장수를 면적으로 나누게 되면 산지나 공지가 있어서 오차가 발생한다는 의견
  - 부동산으로 접근했을 때 카페는 대부분 상가에 있기 때문에 상가 정보가 필요하다고 생각
  - 전국 상가정보(중대형, 소규모, 집합상가), 오피스상권 데이터들 가져옴
  - 시군구가 아닌 상권별로 나눠져있어서 도 단위의 전국데이터에만 사용가능
  - 상가의 동수, 평균층수, 평균면적이 있으나 평균층수와 연면적은 지역의 특성을 나타내지 못한다고 판단
  - 상가의 건물동수와 오피스상권의 건물동수의 총합을 커피지수1로 나눠보면 유의미한 결과를 도출하지 않을까?
- 상가가 아닌곳에 카페가 있는 경우를 포함하지 못한다.
- | 상가건물동수 | 오피스상권건물동수 | 상가건물동수+오피스상권건물동수 | 상가건물면적 | 오피스상권면적 | 상가건물면적+오피스상권면적 | 커피지수1 | 커피지수4 |
|--------|-----------|------------------|--------|---------|----------------|-------|-------|
| 404    | 8,352     | 8,756            | 7,927  |         |                |       |       |

상가의 건물동수와 오피스상권의 건물동수 비교를 보면, 오피스상권이 상가보다 4배 이상 증가한 반면, 상가의 건물동수는 오히려 10% 정도 감소했다. 이는 상가가 아닌 곳에 카페가 있는 경우를 포함하지 못한다.

| 스상가 | 합     |
|-----|-------|
| 404 | 8,352 |
| 38  | 7,927 |
|     | 2,411 |





**Thank You**  
감사합니다