

2022 환경데이터 분석·활용 공모전

# 머신러닝을 활용한 커피박 수거함 입지 최적화 분석

---

재사용 가능 자원 커피박 수거 활성화 방안

JazzIsHorse

# CONTENTS

---

## 01

### 기획 배경

1. 커피박 개요
2. 커피박 처리 현황
3. 커피박 활용 사례
4. 프로젝트 목표

## 02

### 데이터 분석

1. 데이터 수집 & 전처리
2. 데이터 분석 종합 및 시각화
3. 입지선정을 위한 군집분석

## 03

### 결론

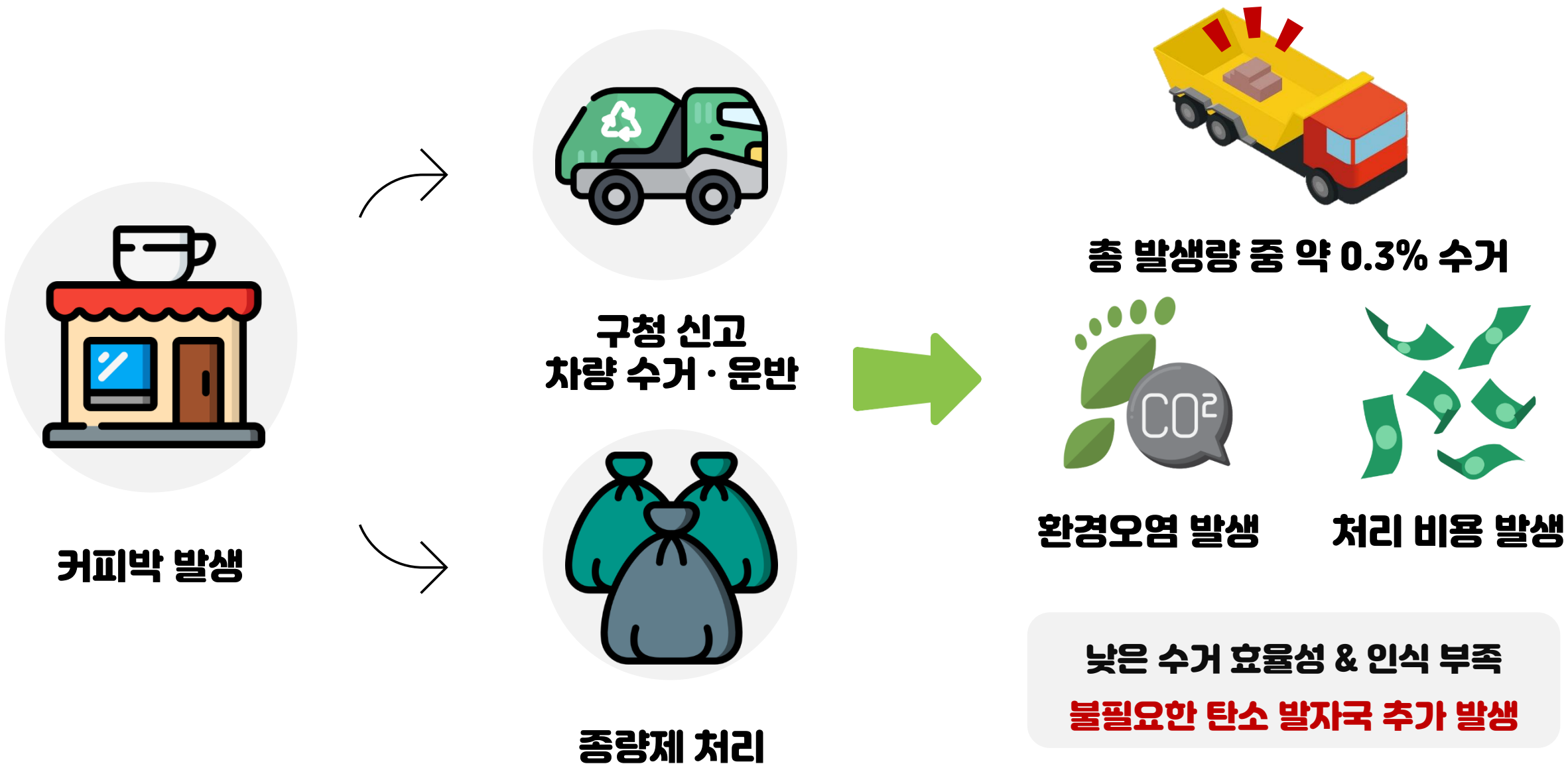
1. 군집분석 결과
2. 기대 효과(환경적 관점)
3. 기대 효과(경제적 관점)



**커피박이란?** 커피를 만들고 남은 '**커피 찌꺼기**'

아메리카노 한 잔에 약 15g의 커피원두 사용, 이 중 14.97g, 99.8%의 원두는 버려짐

연간 '**약 15만 톤**'의 커피박 발생

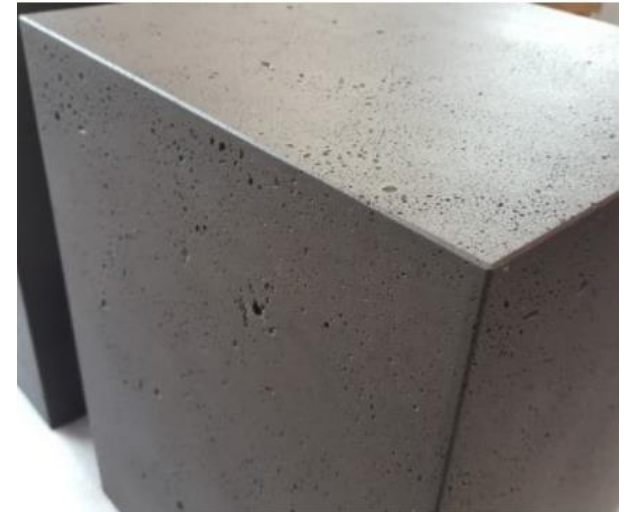




화장품



연료

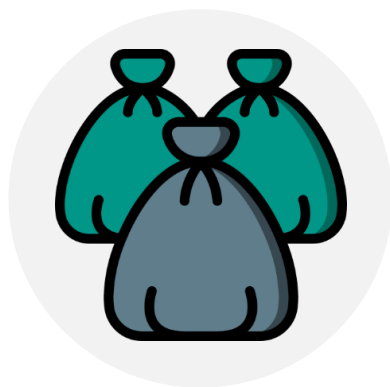


가구

커피박은 처리 과정에서 탄소 배출의 큰 문제로 작용  
그러나 커피박을 적절히 활용한다면 **많은 부가가치** 창출 가능!



카페별 운반수거



생활폐기물로 처리



커피박 수거함 설치

본 프로젝트에서는 머신러닝을 활용한 최적의 커피박 수거함 입지 분석·적용  
커피박 처리 과정에서 발생하는 **환경오염, 경제적 비용**을 감소시키고자 함

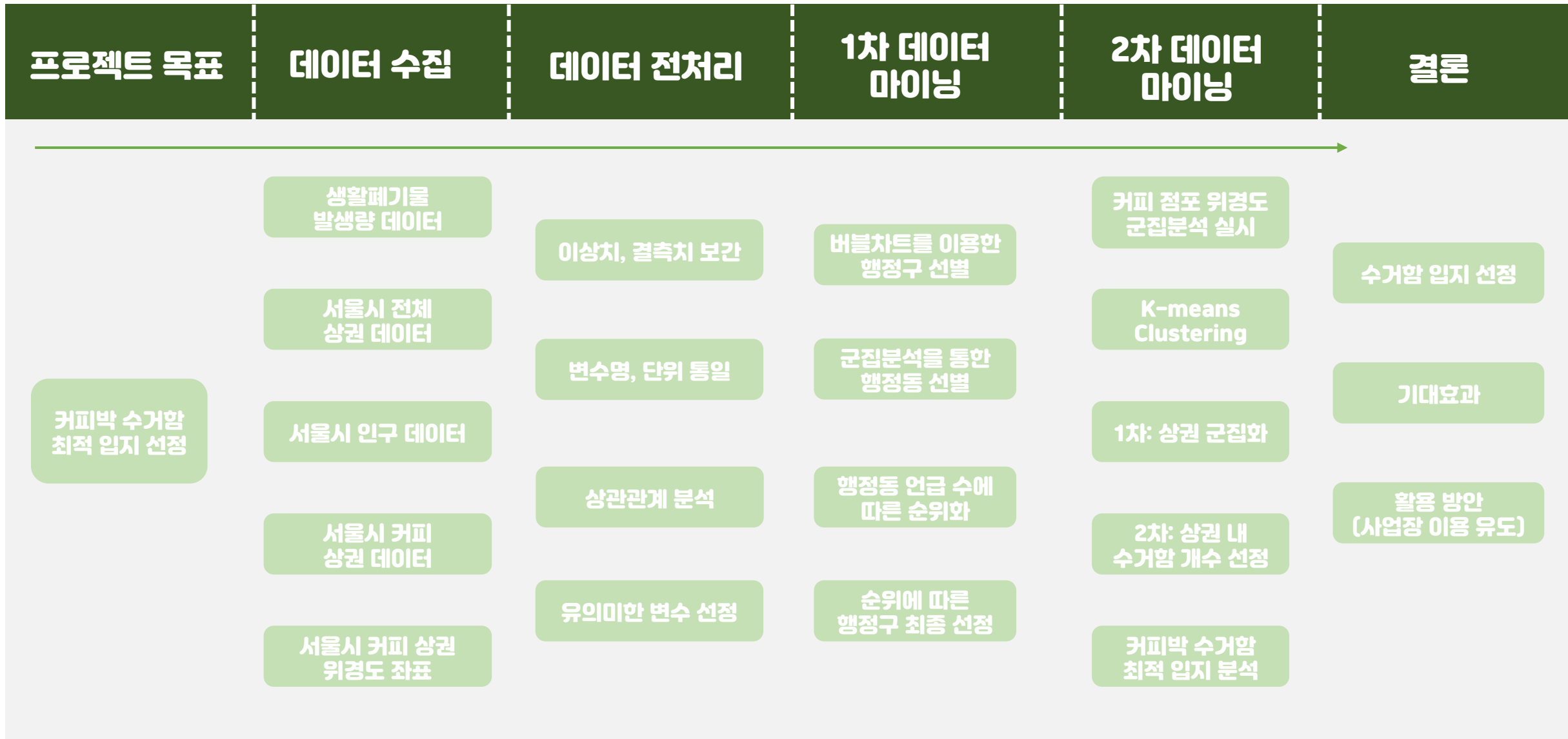




그림 1. 데이터 수집 경로

37	금천구	사업장배출 시설계폐기물	2.1	0.8	8.8	7.3	38.095238	82.954545
38	영등포구	생활폐기물	261.9	111.0	449.4	160.8	42.382589	35.781041
39	영등포구	사업장배출 시설계폐기물	38.7	38.6	56.5	56.0	99.741602	99.115044
40	동작구	생활폐기물	222.4	131.9	324.5	151.7	59.307554	46.748844
41	동작구	사업장배출 시설계폐기물	-	-	2.6	2.5	NaN	96.153846
42	관악구	생활폐기물	343.4	221.9	401.6	188.7	64.618521	46.987052
43	관악구	사업장배출 시설계폐기물	0.4	0.4	1.0	0.8	100.000000	80.000000
44	서초구	생활폐기물	319.1	114.0	504.8	142.6	35.725478	28.248811
45	서초구	사업장배출 시설계폐기물	-	-	7.3	4.8	NaN	65.753425
46	강남구	생활폐기물	468.0	225.0	694.9	213.1	48.076923	30.666283
47	강남구	사업장배출 시설계폐기물	586.2	234.3	408.5	240.6	39.969294	58.898409

그림 2. 데이터 전처리 과정

## 데이터 수집

**행정구별:** 커피 전체 점포 수, 5년 커피집 생존율 ...

면적당 매출액 등 **13개 데이터셋 수집**

**행정동별:** 커피 점포 수, 평균영업기간(년),

5년 커피집 생존율, 추정 임대료 **4개 데이터셋 수집**

## 데이터 전처리

**데이터 확인:** 결측치 보간 · 제거

**데이터 통합:** 변수명, 단위 통일



배출량(톤/일)	1.0	0.7	0.7	-0.4	0.1	0.5	-0.2	0.4	0.1	-0.1	0.7	0.6	0.7
운영점포수 (개)	0.7	1.0	1.0	-0.3	0.2	0.7	-0.2	0.7	-0.3	-0.1	0.9	0.9	1.0
종사자수 (명)	0.7	1.0	1.0	-0.2	0.2	0.5	-0.2	0.6	-0.4	-0.0	1.0	1.0	0.9
평균영업기간 (년)	-0.4	-0.3	-0.2	1.0	0.3	-0.6	0.3	-0.6	0.2	0.5	-0.3	-0.2	-0.3
면적당매출액 (백만원/3.3m <sup>2</sup> )	0.1	0.2	0.2	0.3	1.0	-0.1	0.5	-0.2	0.2	0.5	0.1	0.1	0.1
경제활동 인구수	0.5	0.7	0.5	-0.6	-0.1	1.0	-0.3	0.6	-0.0	-0.4	0.5	0.4	0.8
3년 커피집 생존율	-0.2	-0.2	-0.2	0.3	0.5	-0.3	1.0	-0.4	0.4	0.8	-0.3	-0.2	-0.2
면적비(%)	0.4	0.7	0.6	-0.6	-0.2	0.6	-0.4	1.0	-0.5	-0.4	0.6	0.6	0.7
생활폐기물 재활용율	0.1	-0.3	-0.4	0.2	0.2	-0.0	0.4	-0.5	1.0	0.2	-0.4	-0.3	-0.1
5년 커피집 생존율	-0.1	-0.1	-0.0	0.5	0.5	-0.4	0.8	-0.4	0.2	1.0	-0.1	-0.0	-0.1
프랜차이즈 커피 점포 수	0.7	0.9	1.0	-0.3	0.1	0.5	-0.3	0.6	-0.4	-0.1	1.0	0.9	0.8
소상공인 커피 점포 수	0.6	0.9	1.0	-0.2	0.1	0.4	-0.2	0.6	-0.3	-0.0	0.9	1.0	0.8
생활폐기물 발생량	0.7	1.0	0.9	-0.3	0.1	0.8	-0.2	0.7	-0.1	-0.1	0.8	0.8	1.0

## 행정구 최종 변수 선정

생활폐기물 발생량

면적당 매출액

경제활동 인구수

5년 커피집 생존율



버블차트(Bubble Chart) 활용한 분석 결과 시각화  
수집된 '행정구별 데이터'를 바탕으로 상위 분포 행정구 선정

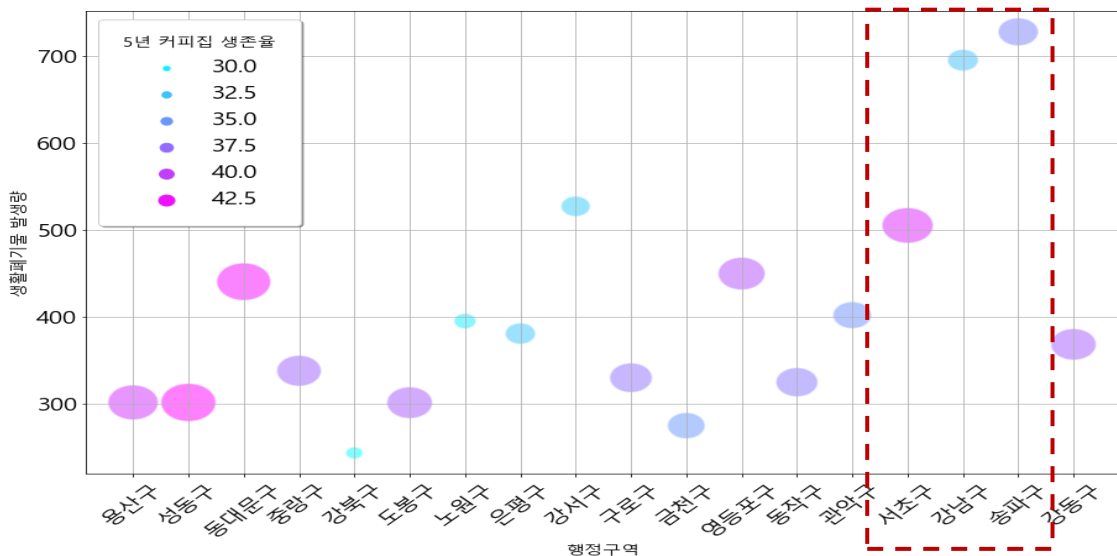


그림 3. 생활폐기물 발생량 - 5년 커피집 생존율

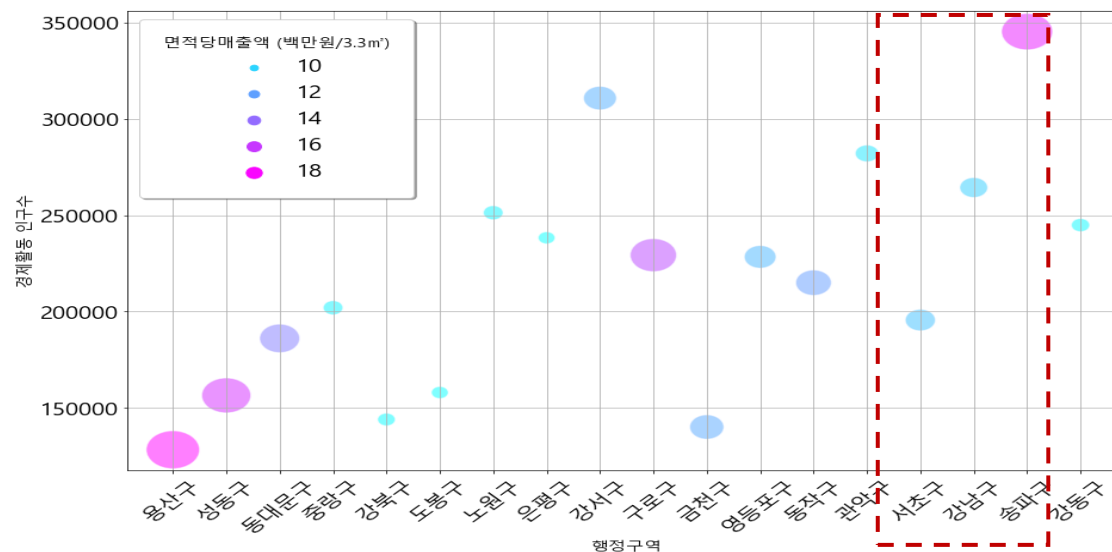


그림 4. 경제활동 인구수 - 면적당 매출액

선정 변수 4개 중 2개씩 교차 분석을 실시  
상위 분포 행정구 1차 선정



K-평균 알고리즘(K-means Clustering)을 활용  
수집된 ‘행정동별 데이터’를 바탕으로 상위 군집 선정

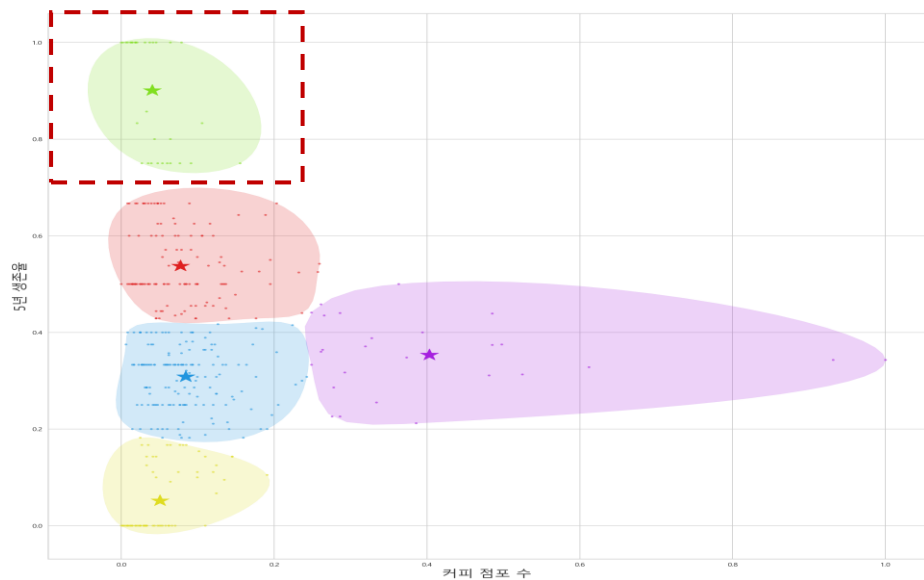


그림 4. 커피 점포 수 – 5년 커피점포 생존율

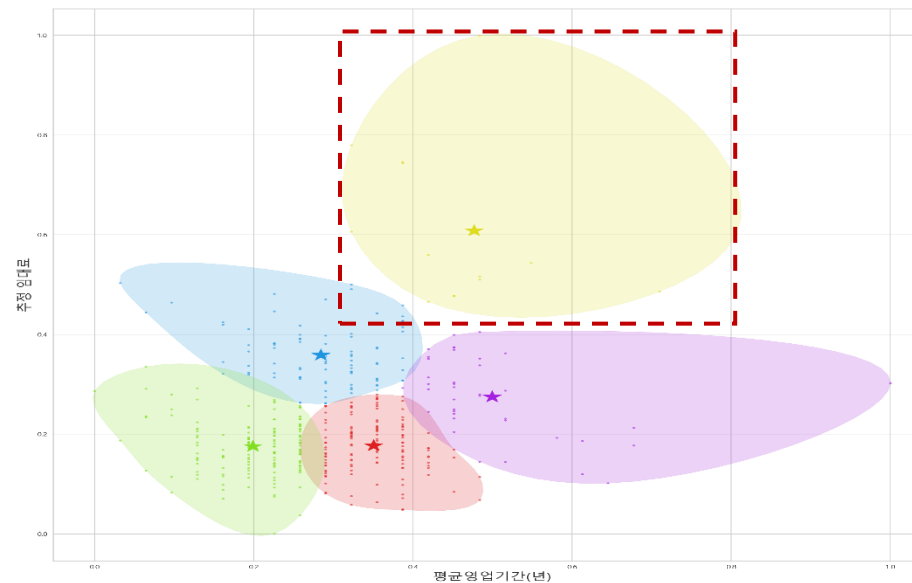


그림 5. 평균 영업기간 – 추정 임대료

4개의 선정 변수 중 2개씩 교차 분석 실시  
5개 군집화를 기준으로 선정 군집에 포함된 행정동 카운트

	행정구역	연급횟수
0	서울특별시 서초구 서초4동	4
1	서울특별시 송파구 방이2동	4
2	서울특별시 관악구 보라매동	4
...	...	...
52	서울특별시 용산구 이촌1동	1
53	서울특별시 강남구 일원본동	1
54	서울특별시 종로구 무악동	1

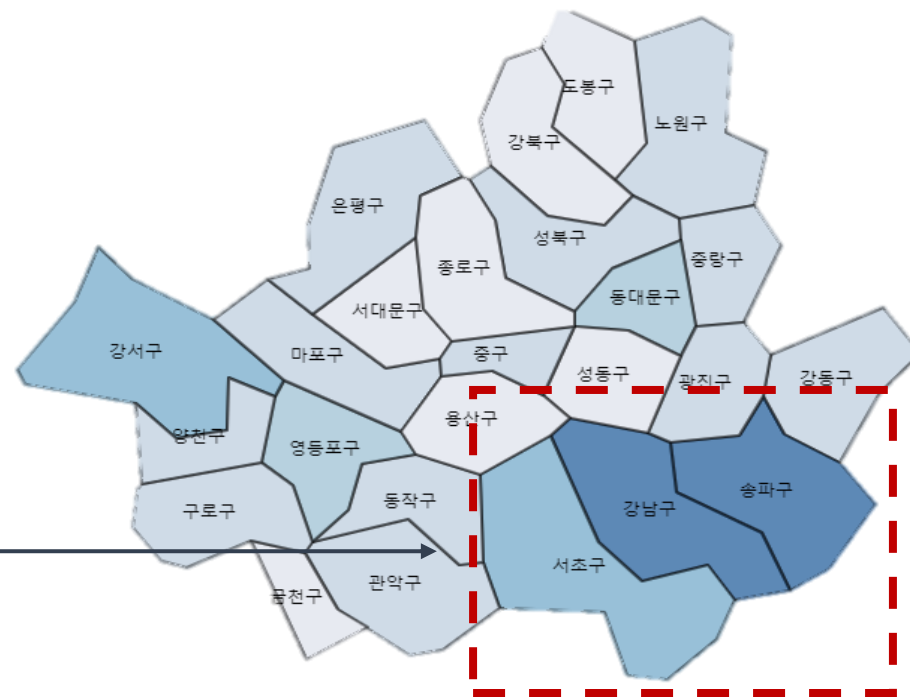
### 그림 6. 행정구 언급횟수 카운트

행정구역	면적비(%)
1 종로구	4.0
2 중구	1.6
3 용산구	3.6
...	...
23 강남구	6.5
24 송파구	5.6
25 강동구	4.1

## 그림 7. 행정구별 면적비

행정구역	연급횟수	면적비(%)	면적당 카운트
0 송파구	10	5.6	1.785714
1 강남구	11	6.5	1.692308
2 서초구	4	7.8	0.512821

### 그림 8. 면적 대비 언급 횟수 순위화



### 그림 9. 서울시 행정구

**커피박 수거함 필요 지역 선정**  
**서초구, 강남구, 송파구**

▶ 선정구별 커피 점포 위·경도를 활용하여 적정 군집 수 산출  
K-Elbow를 통해 적정 군집 수(K)개 선정

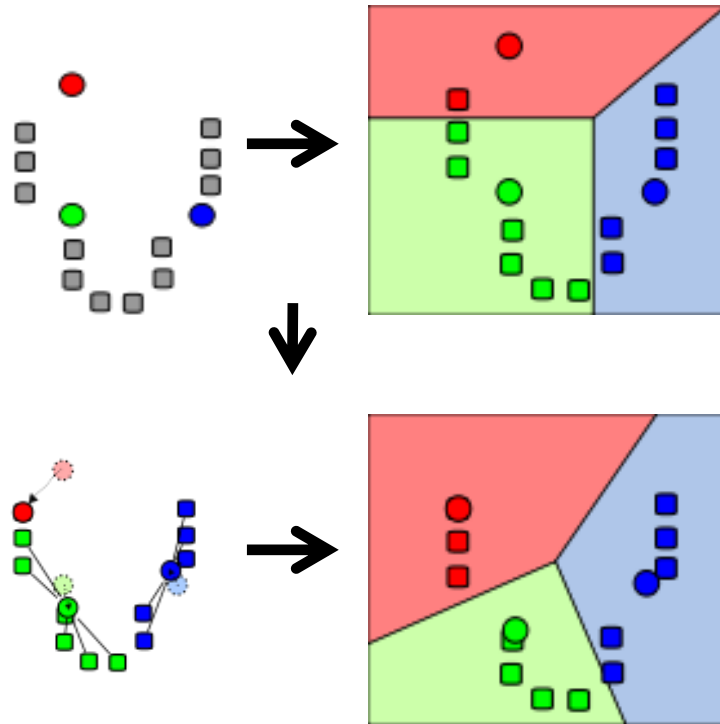


그림 10. 군집분석 원리

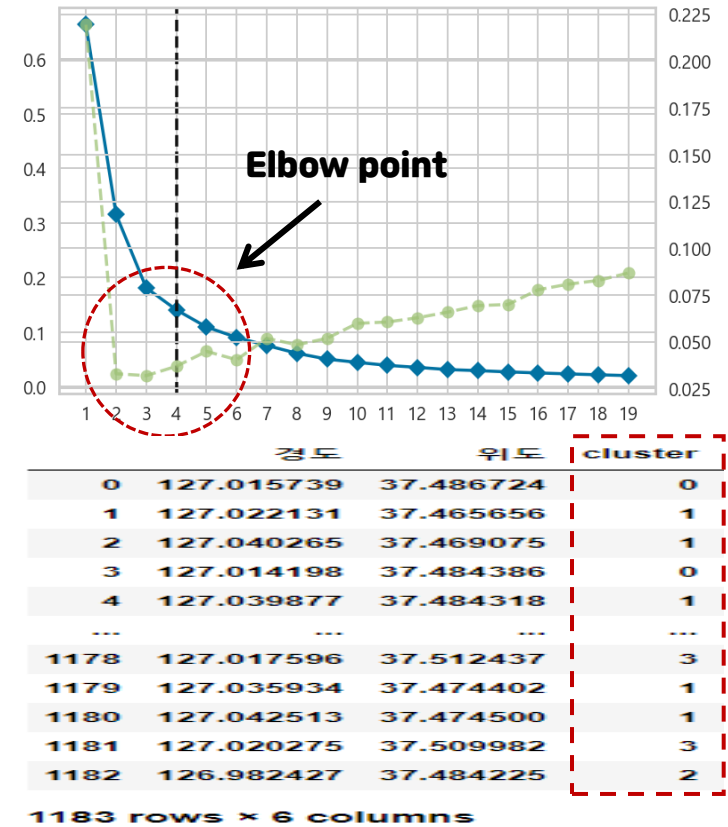


그림 11. 군집화 과정

커피 점포 간 **최단거리** 분석을 통한 군집화

## 선정된 군집 수(K)에 따른 1차 군집화 결과

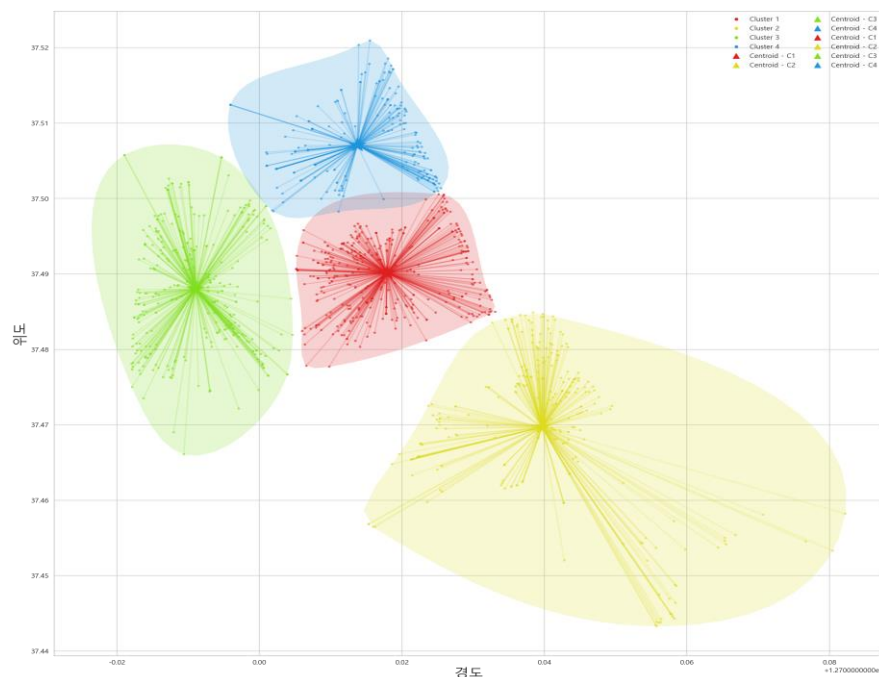


그림 12. K개수 만큼 군집화

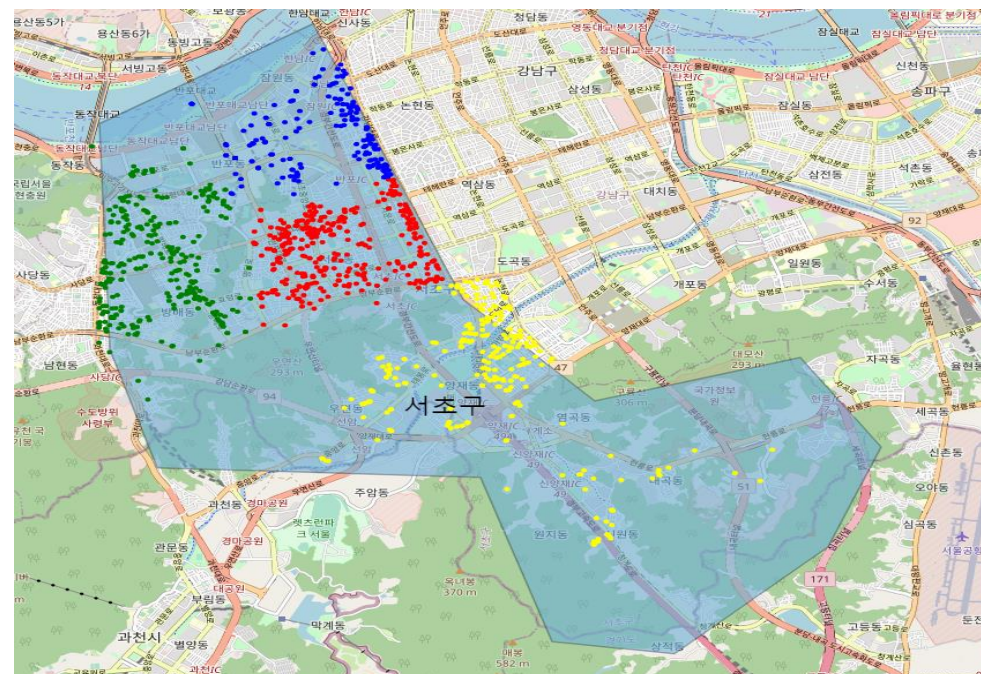


그림 13. 1차 군집 지도 시각화

✓ 행정구 내 **K개의 커피 상권**으로 세분화



1차 선정된 커피 상권들을 대상으로 2차 군집화 실시

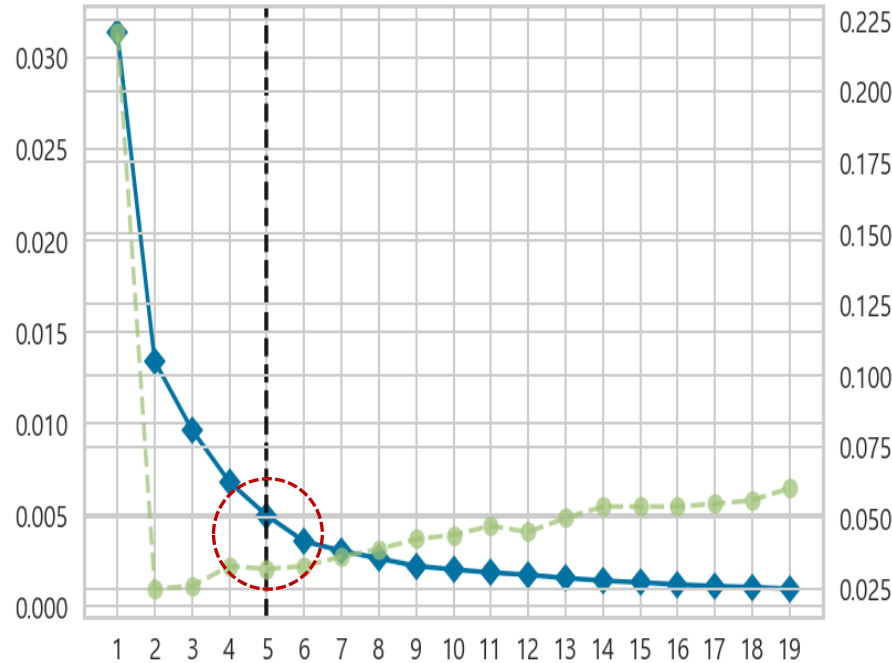


그림 14. K-Elbow 군집개수 선정

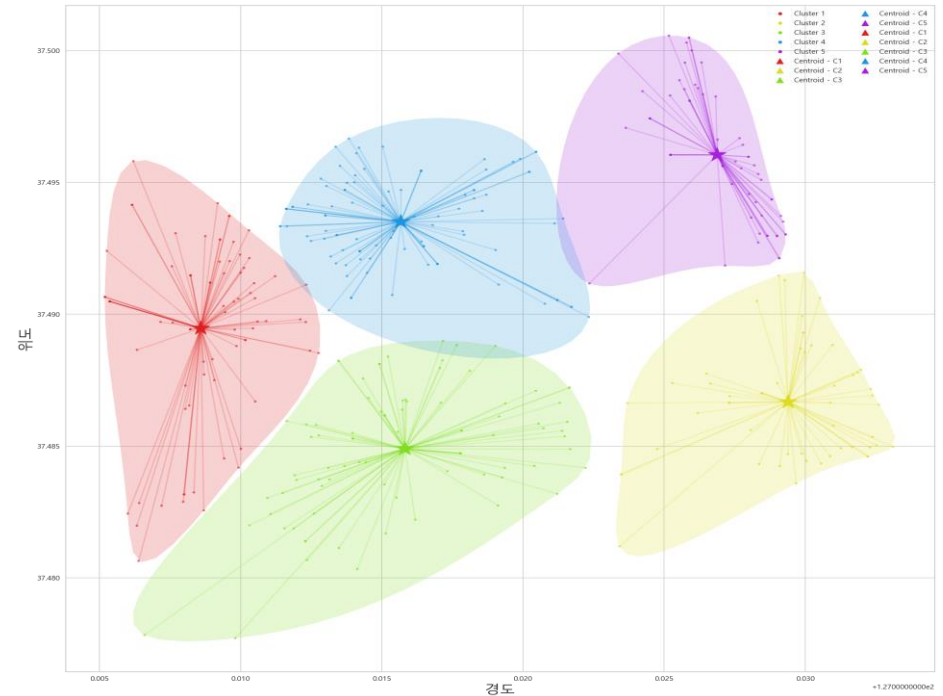


그림 15. K개수 만큼 군집화

✓ 선정된 커피 상권 내 **K개의 커피박 수거함 위치 & 개수 선정**





## 커피 상권 A 내 2차 군집화 결과

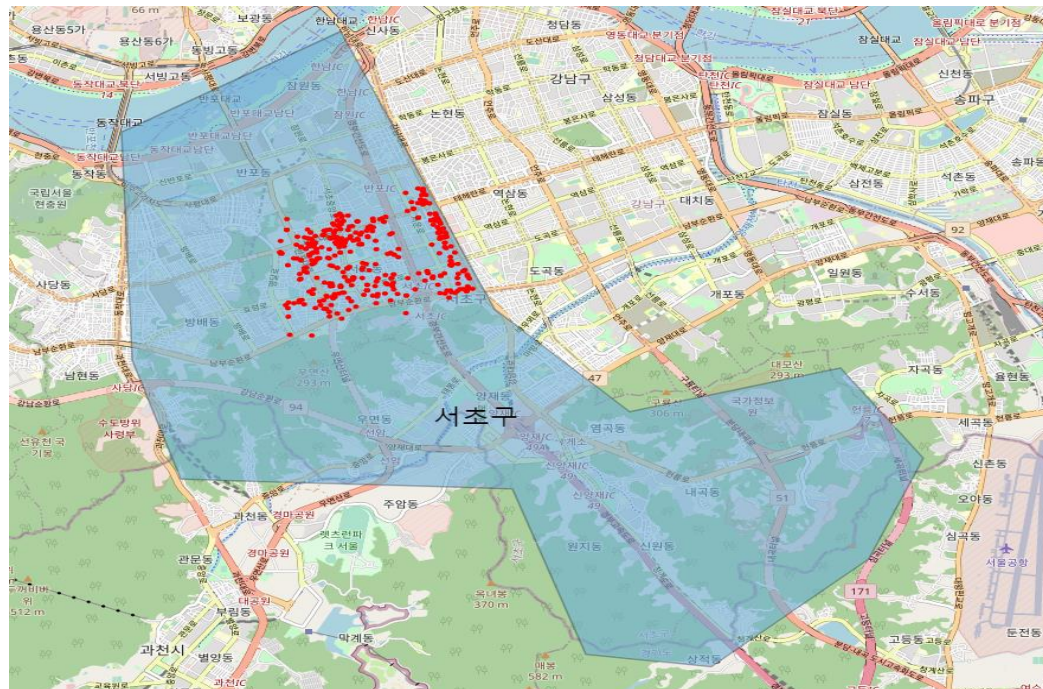


그림 16. 1차 군집 분석을 통한 상권 A

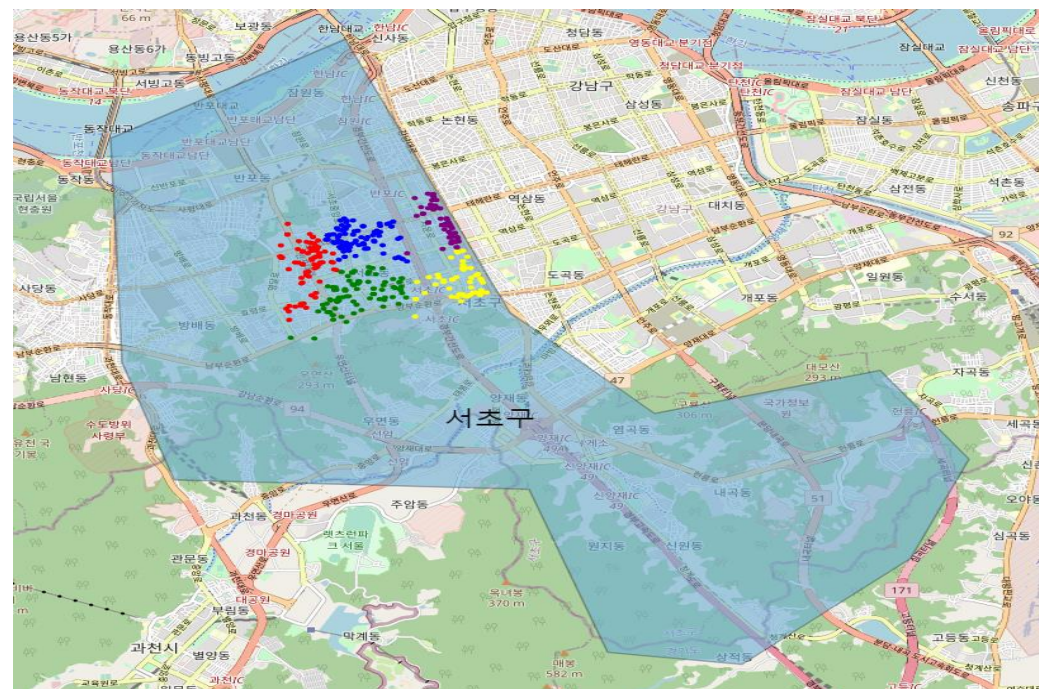
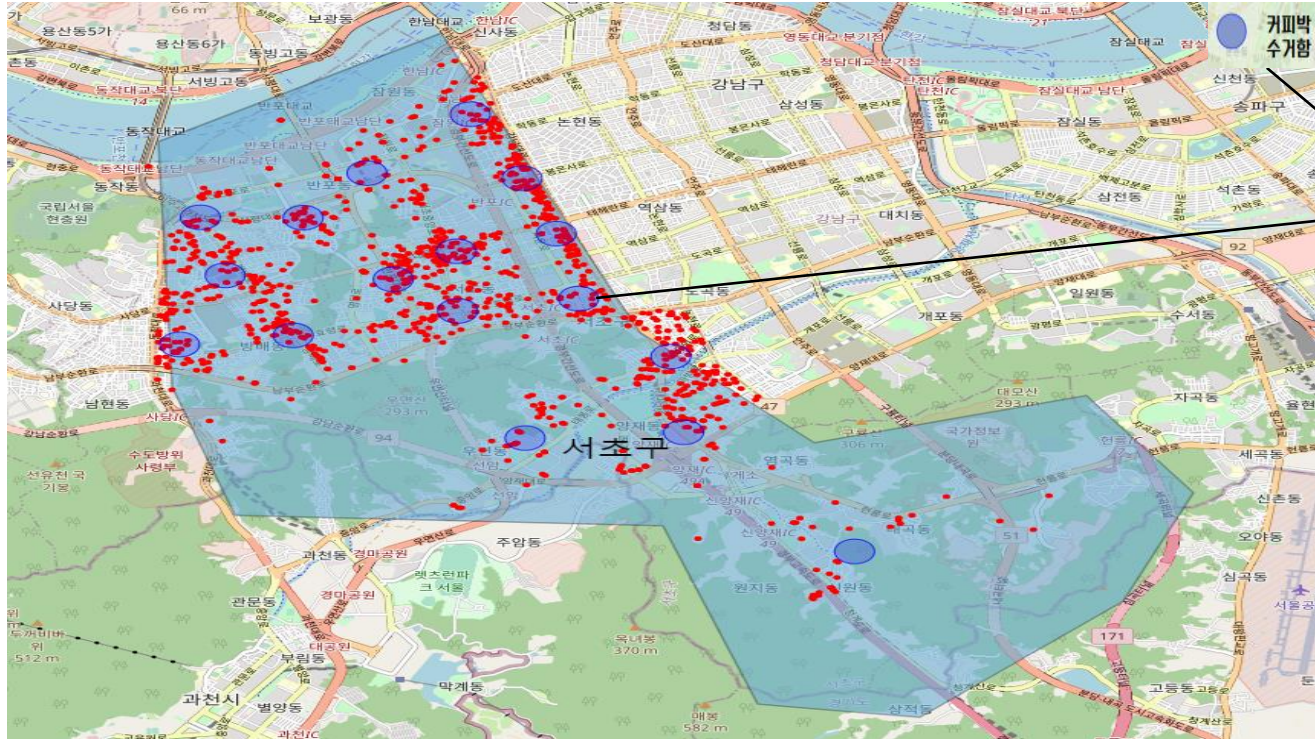


그림 17. K개수 만큼 상권 A 2차 군집화

✓ 커피 상권 A 내 **K개의 커피박 수거함 위치 & 개수 선정**



## 군집 중심점(Centroids) 좌표 수집

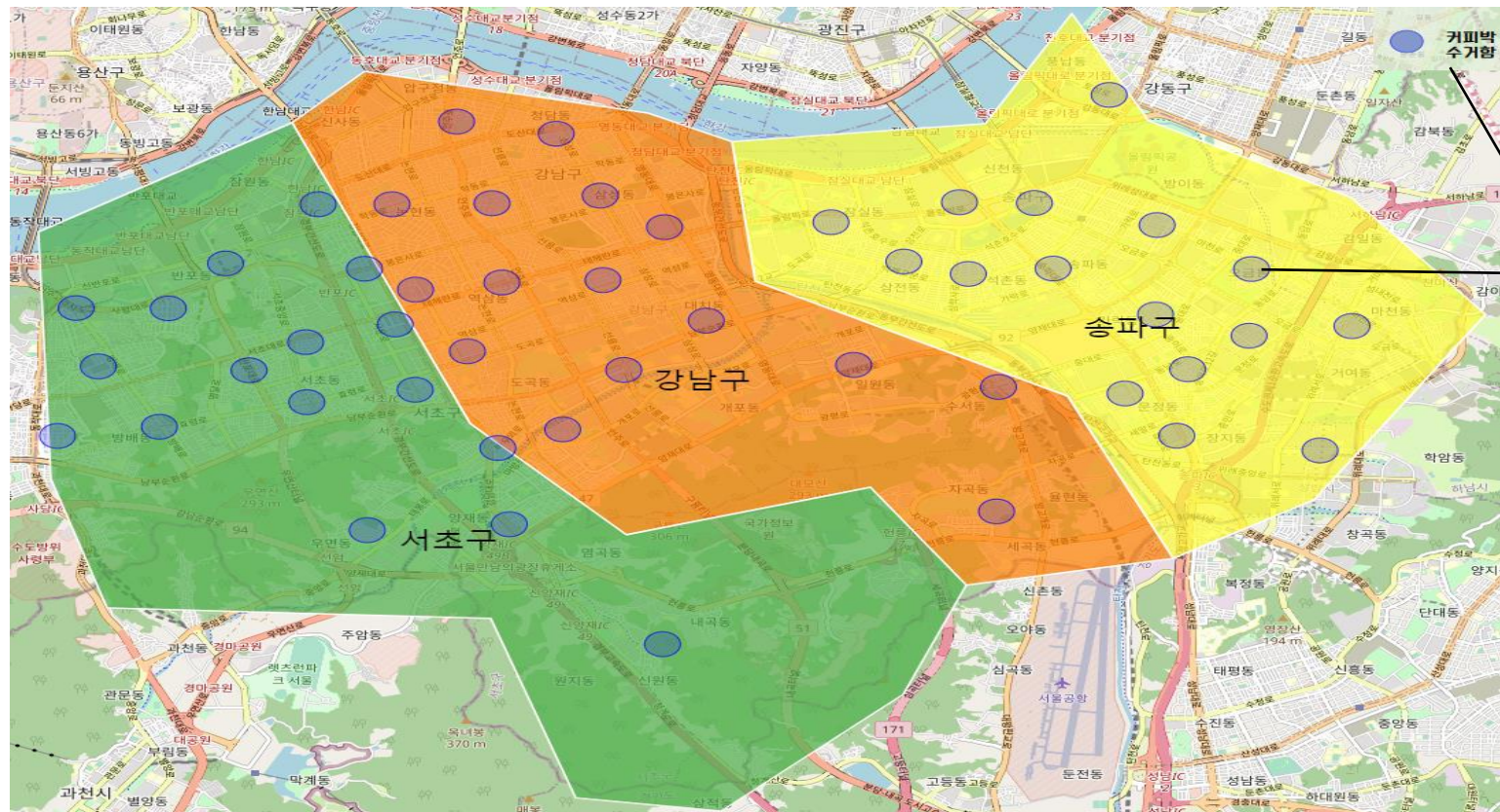


### 중심점(Centroids)

커피 점포간 '최단 거리'를 고려한  
커피박 수거함 설치 최적 입지

그림 18. 커피박 수거함 입지 분석 결과

✓ 행정구별 최종 커피박 수거함 설치 위치 & 개수 입지 분석



**강남 3구 대상 입지 분석 결과**  
**총 49개 커피박 수거함 설치**

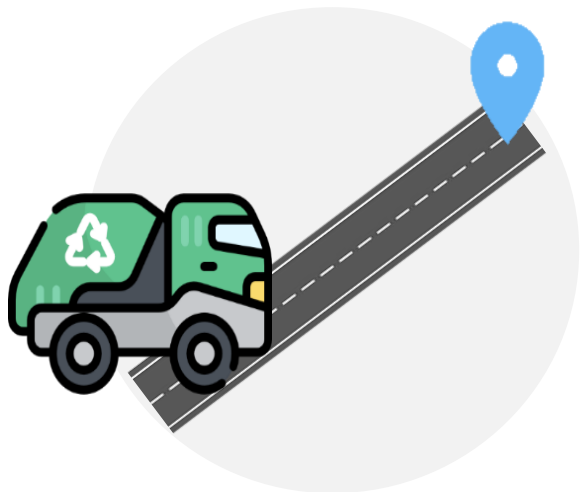
그림 19. 강남3구 커피박 수거함 최적 입지

- ✓ 수 백 곳의 매장 개별 방문에서 49곳으로 수거 과정의 탄소발자국 대폭 감소 기대
- ✓ 커피박 재활용을 용이하게 해 처리 비용 감소 및 탄소 발생 저감에 이바지 가능





환경오염 개선 효과



커피박 수거 방법 개선 효과

2020년 기준 커피박 발생량 **약 35만 3296톤**

매립 · 소각 처리하는데 약 39억 대 자동차 매년 CO<sub>2</sub> 발생



- ✓ 커피박 수거함을 통해 재활용을 용이하게 함
- ✓ 커피박 **처리 비용 감소 및 탄소 발생 저감**에 이바지 가능

구청 신고 후 차량 매장 개별 방문 > 수거 · 운반  
비효율 & 불필요한 동선 발생



- ✓ 커피 수거함은 '거점형 방식' 채택 - **불필요한 동선 감축**
- ✓ 이를 통해 **수거 · 운반 시 비용 및 탄소발자국 저감** 일조



경제적 문제 개선 효과

2020년 기준 커피박 처리 시 **약 350억원**이 소요



- ✓ 커피박 수거함 방식 - **종량제 봉투 사용량 감소**
- ✓ 사업주 쓰레기 처리 비용 감소 - **가게 운영비 절감**



커피박 수거함 리워드 혜택

국내 '수퍼빈' / 유럽 'Pfand' 수거 방식 운영



- ✓ '**리워드형 수거 방식**'을 차용해 커피박 수거함 사용 촉구
- ✓ 사업주는 가게 운영비 절감과 리워드 **이중 수혜 가능**
- ✓ 이를 통해 **커피박 수거·운반 및 처리 비용 저감** 일조