대전시 공영자전거 대여소 인프라 분석 및 최적 모델 제안

김건민, 김태우, 이동원, 최현준

대전대신고등학교 FINDER(KE)

ABSTRACT: 본 연구는 대전광역시에서 운영 중인 공공 자전거 대여 서비스인 타슈의 자전거 대여소 위치 최적화를 위한 모델을 구현하는 것을 목표한다. 타슈 서비스는 시민의 건강 증진과 친환경 교통수단 활성화를 목표로 시작되었지만, 자전거 대수와 수요의 불균형, 대여소 과밀화 등 다양한 문제를 겪고 있다. 본 연구에서는 데이터 전처리와 k-means clustering을 통해 타슈 대여소의 위치를 분석하고 최적화하였다. 타슈 대여소별 분포율을 계산하고 각 지역의 주요 시설 인프라를 고려해 최적의 위치를 도출하였다. 본 연구는 데이터 기반 분석을 통해 대전시의 공공자전거 정책 개선 방안을 제시한다.

서론 ( Introduction )

대전시는 2008년부터 시민 건강 증진과 친환경 교통수단 활성화를 목적으로 전국 최초의 무인자전거 대여 서비스인 '타슈'를 운영해 왔다. 하지만 서비스 운영 중 자전거 대수와 수요 간의 불균형, 특정 지역의 대여소 밀집, 이용 패턴을 고려하지 않은 운영 정책 등 여러 문제점이 드러났다. 이러한 문제를 해결하기 위해 타슈의 입지 인프라를 분석하고, 다양한 데이터를 활용하여 최적 모델을 구현할 필요성이 제기되었다. 이에 따라 본 탐구에서는 탐구 제안 주제인 전기차 충전소 최적화 모델 구현 대신, 학생들이 더 공감할 수 있는 ‘타슈 대여소 입지 최적화 모델 구현’을 주제로 선정하였다. 이는 타슈가 본 탐구를 진행하는 우리 학생들에게 자주 사용되는 교통수단으로, 학생들의 생활과 밀접한 관련이 있기 때문이다. 따라서 이를 통해 우리가 실제 공감하는 문제에 대한 해결방안을 모색하고자 한다.

재료 및 방법 ( Materials and Methods )

1. 데이터 수집

데이터 대전광역시\_타슈정보(2023,대전광역시), 대전교통공사\_공영자전거타슈대여소현황(2023,대전교통공사)를 공공데이터포털(data.go.kr)에서 데이터 활용 신청을 통해 수집한다. 위 데이터들을 통해 대전시에 존재하는 약 300개의 타슈 대여소의 주소, 이름, 위치코드 등을 확인할 수 있다. 추가로 folium 라이브러리에서 제공하는 분포도 시각화에 활용되는 행정동 경계를 표시하기 위하여 주소기반산업지원서비스에서 구역의 도형 자료를 신청하여 수집하고, geodata.geojson이라는 이름으로 저장하였다.

2. 데이터 전처리

UTF-8 인코딩 형식으로 저장된 CSV 데이터를 pandas 라이브러리를 이용하여 불러와 데이터프레임으로 변환하였다. 본 데이터프레임은 '위치' 열을 포함하고 있으며, 해당 열만을 추출하여 새로운 데이터프레임 tasu에 저장하였다. 추가로, 데이터의 순서를 재정렬하기 위해 인덱스를 초기화했다. 주소 데이터를 활용하여 각 위치의 지리적 좌표(위도와 경도)를 획득하기 위해 geokakao 라이브러리를 사용하여 tasu 데이터프레임에 포함된 주소 정보를 바탕으로 위도(decimalLatitude)와 경도(decimalLongitude) 정보를 추가하였다. 변환 작업이 완료된 후, 좌표 정보를 포함한 새로운 CSV 파일 tasu\_position.csv로 저장하였다. 생성된 tasu\_position.csv 파일에서 좌표 정보(decimalLatitude, decimalLongitude)에 결측치가 존재할 수 있으므로, 이를 제거하여 데이터의 정확성을 확보하였다. 결측치가 제거된 데이터프레임은 인덱스를 다시 초기화하여 데이터의 순서를 재정렬하였다.



< 데이터 전처리 Python 코드>

3. 타슈 대여소 분포도 시각화

Python의 pandas 라이브러리를 이용하여 CSV 파일을 데이터프레임으로 불러왔으며, GeoJSON 파일은 json 라이브러리를 이용하여 읽어들였다. Folium 라이브러리를 이용하여 대전광역시를 중심으로 하는 지도를 생성하였다. 지도의 초기 중심 좌표는 대전광역시의 중심부로 설정하였으며, 초기 확대 수준은 12로 설정하였다. 대전광역시의 행정 구역 경계를 시각화하기 위해 GeoJSON 데이터를 필터링하였다. GeoJSON 데이터에서 'adm\_nm' 속성에 '대전광역시'가 포함된 행정 구역만을 추출하여 새로운 GeoJSON 객체를 생성하였다. 이 객체를 Folium의 GeoJson을 이용하여 지도에 추가하였으며, 경계선의 스타일은 빨간색(color: 'red'), 두께는 3(weight: 3), 채우기 불투명도는 0(fillOpacity: 0)으로 설정하였다. 정제된 타슈 대여소 데이터를 이용하여 각 대여소의 위치를 지도에 마커로 표시하였다. 데이터프레임에 각 대여소의 위도(decimalLatitude)와 경도(decimalLongitude)를 추출하고, Folium의 Marker 클래스를 이용하여 지도에 추가하였다. 마커의 색상은 파란색(color: 'blue'), 모양은 자전거 모양(icon: 'bicycle'), 팝업에는 대여소의 위치 이름을 표시하였다.



< 분포도 시각화 Python 코드>

4. 타슈 대여소 분포율 정의 및 계산

타슈 대여소의 분포율은 특정 범위 내에서 타슈 대여소의 수를 인구 수로 나눈 값을 나타낸 것으로 정의할 수 있으며, 인구 1만명 당 분포율을 나타내려면 10000을 곱해준 것으로 표현할 수 있다. 이를 수식으로 나타내면 다음과 같다.

이후, 정의한 분포율의 정의를 바탕으로 대전광역시의 구별 타슈 대여소 분포율을 계산하는 python 코드를 작성하였다. 대전광역시 2023년 구별 인구 수 데이터를 활용하여 적용하였다. 인구수 데이터에 표시된 명칭 ‘도시 + 구(예 : 대전광역시 서구)’를 split 문법을 활용하여 나누고, expand를 True로 설정하여 다수의 칼럼으로 분할하였다. 이와 유사한 과정으로 인구수 데이터의 쉼표를 제거하고 정수형으로 변환하고, 타슈 대여소 또한 구별로 분류하여 value\_counts 문법을 활용해 대여소의 수를 데이터화하였다. 이후, 처리한 두 데이터 프레임을 병합하고, 대전광역시 전체 데이터의 분포율을 추가하여 테이블로 시각화하였다.

5. 모델 선정

(1) K-means

K-means는 주어진 데이터를 k개의 군집으로 분류하는 비지도 학습 알고리즘으로, 각 데이터 포인트를 최근접 중심에 할당하는 식으로 동작한다. 각 군집의 데이터는 서로 유사한 특성을 지니며, 각 군집 내의 데이터 간 거리를 최소화하고 군집 간 거리를 최대화하는 것을 핵심 목표로 진행한다. 이를 수식으로 표현하면 다음과 같다.

(n개의 데이터 오브젝트 (x1,x2, … xn) 집합에서 거리 차의 제곱을 최대로 하는 k개의 집합 (s1,s2, … , sk)으로 분할한다, )

(arg란 argument의 약자이며, 어떤 함수가 최솟값 또는 최댓값을 가질 때, 그 값을 만드는 변수를 구하는 것을 나타낸다.)

각 데이터로부터 각 클러스터들의 유클리드 거리를 계산한다.

K-means는 군집의 개수(K)를 사전에 설정해야 하며, 이는 분석가의 판단이나 특정 기법을 통해 결정된다. 본 탐구에서는 추가로 대여소를 설치할 위치를 사전에 지정할 수 있다는 점과 각 대여소가 위치한 지역의 특성과 인프라가 다르다는 점을 고려하여, 2차원 위도와 경도 데이터를 기반으로 군집화가 가능하다고 판단했다. 이러한 이유로, K-means 군집화 모델이 적합하다고 보아 이를 채택하였다.

(2) DBSCAN

DBSCAN(Density-Based Spatial Clustering of Applications with Noise)은 주어진 데이터를 밀도 기반으로 클러스터링하는 비지도 학습 알고리즘이다. 이 알고리즘은 밀도가 높은 데이터 포인트 집합을 클러스터로 정의하며, 클러스터 외부에 위치한 데이터 포인트는 노이즈 또는 이상치로 간주한다. DBSCAN은 각 포인트에서 설정된 반경(ε) 내에 있는 이웃 포인트의 수(minPts)를 기준으로 밀도 연결성을 판단하여 클러스터를 형성한다. 이를 수식으로 표현하면, n개의 데이터 오브젝트 (x₁, x₂, ... , xₙ) 집합에서 반경 ε 내에 포함된 포인트의 개수에 따라 밀집된 지역을 클러스터로 정의하며, ε는 반경, minPts는 클러스터를 형성하기 위한 최소 포인트 수를 나타낸다. ε 와 minPts는 DBSCAN알고리즘에서 클러스터링의 품질과 결과에 직접적인 영향을 미치기 때문에 이 값을 선정하는 것이 매우 중요하다. 그러나 ε 와 minPts는 따로 구하는 식이 없기 때문에 Heuristic 방법을 이용하여 구한다. DBSCAN의 원 논문(Martin Ester, Hans-Peter Kriegel, Jorg Sander, Xiaowei Xu, 1996, "A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise", KDD-96)에 따르면 2차원 데이터에서 minPts를 4로 설정하는 것을 권장하고 있다.

Eps는 sorted k-dist graph를 그린 후 Elbow method 를 사용해서 꺽이는 부분의 k-dist 를 Eps 로 결정하는 방법으로 구한다. sorted k-dist graph는 k\_NN (k-Nearest Neighbor)의 거리 k-dist를 구하고, k-dist 를 내림차순으로 정렬하여, X축은 정렬된 점들 별로 Y축에는 k-dist 를 그려서 그릴 수 있다.

DBSCAN은 클러스터 개수를 사전에 설정할 필요가 없으며, 이는 자전거 대여소가 위치한 지역의 밀도와 사용 패턴이 균일하지 않을 경우 큰 이점으로 작용한다. 또한 특정 지역에서 자전거 대여소 이용이 적은 경우 이를 노이즈로 처리할 수 있으며, 밀도가 높은 지역에서는 자연스럽게 클러스터를 형성할 수 있다. 대전시 자전거 인프라는 다양한 지역에 걸쳐 있으며, 각 지역의 인프라와 교통 상황이 상이하기 때문에 클러스터의 모양이 비정형적일 가능성이 크다.

이러한 이유로, DBSCAN 알고리즘은 노이즈 처리와 다양한 밀도 수준을 고려할 수 있다는 점에서 본 탐구의 대전시 공용 자전거 인프라 분석에 적합하다고 판단되었다. 특히, 2차원 위도와 경도 데이터를 기반으로 밀도 기반 클러스터링을 통해 자전거 대여소의 효율적인 설치와 운영을 위한 적절한 분석 방법으로 DBSCAN을 채택하였다.

6. 대여소 입지 최적 모델 구현 – Kmeans

공영 자전거를 이용하는 주된 요인 중 하나인 주요 시설의 위치 좌표 또한 최적화 모델에 활용할 수 있는 중요한 데이터로 판단하여 주요 시설을 임의로 특정하고, 이에 대한 데이터들을 수집하였다.

- 교육 시설(초등학교, 중학교, 고등학교, 학원가) : 전국초중등학교위치표준데이터(2024,교육부), 동부교육지원청 학원 및 교습소 현황(2023,교육부), 서부교육지원청 학원 및 교습소 현황(2023,교육부)

- 상업 지역(쇼핑몰, 대형마트, 주요 상점가) : 대전광역시\_대규모점포 현황(2023, 대전광역시)

- 공공 시설(도서관, 체육시설, 보건시설) : 대전광역시시설관리공단\_시설물 현황정보(2024, 대전광역시시설관리공단)

- 교통 시설(지하철역, 버스 정류장) : 대전광역시\_버스정류장 현황(2022,대전광역시), 대전시 도시철도역 관리 데이터

K-means 클러스터 분석은 데이터를 군집으로 분류하는 비지도 학습 기법의 하나로, 각 데이터 포인트를 k개의 군집 중 하나에 속하게 하는 알고리즘이다. K-means 클러스터 분석의 단계는 다음과 같다.

ㄱ. 초기화: 군집의 개수 k를 설정하고, 초기 군집 중심(centroid)을 무작위로 선택한다.

ㄴ. 할당 단계: 각 데이터 포인트를 가장 가까운 군집 중심에 할당한다. 이는 유클리드 거리(Euclidean distance)를 사용하여 계산한다.

ㄷ. 업데이트 단계: 각 군집에 할당된 데이터 포인트의 평균을 계산하여 새로운 군집 중심으로 설정한다.

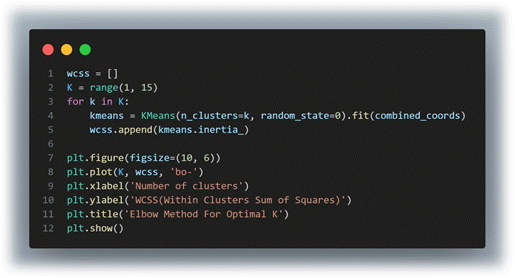
ㄹ. 반복: 군집 중심이 변화하지 않을 때까지(또는 변화가 매우 미미할 때까지) 2번과 3 번 단계를 반복한다.

K-means 군집 알고리즘에서 K 값을 선정하는 보편적인 방법은 WCSS(Within Clusters Sum of Squares)를 활용해 WCSS 값과 K 값의 변화에 따른 그래프를 그려 팔처럼 굽어지는 특정 지점을 선정하는 Elbow Method이다.

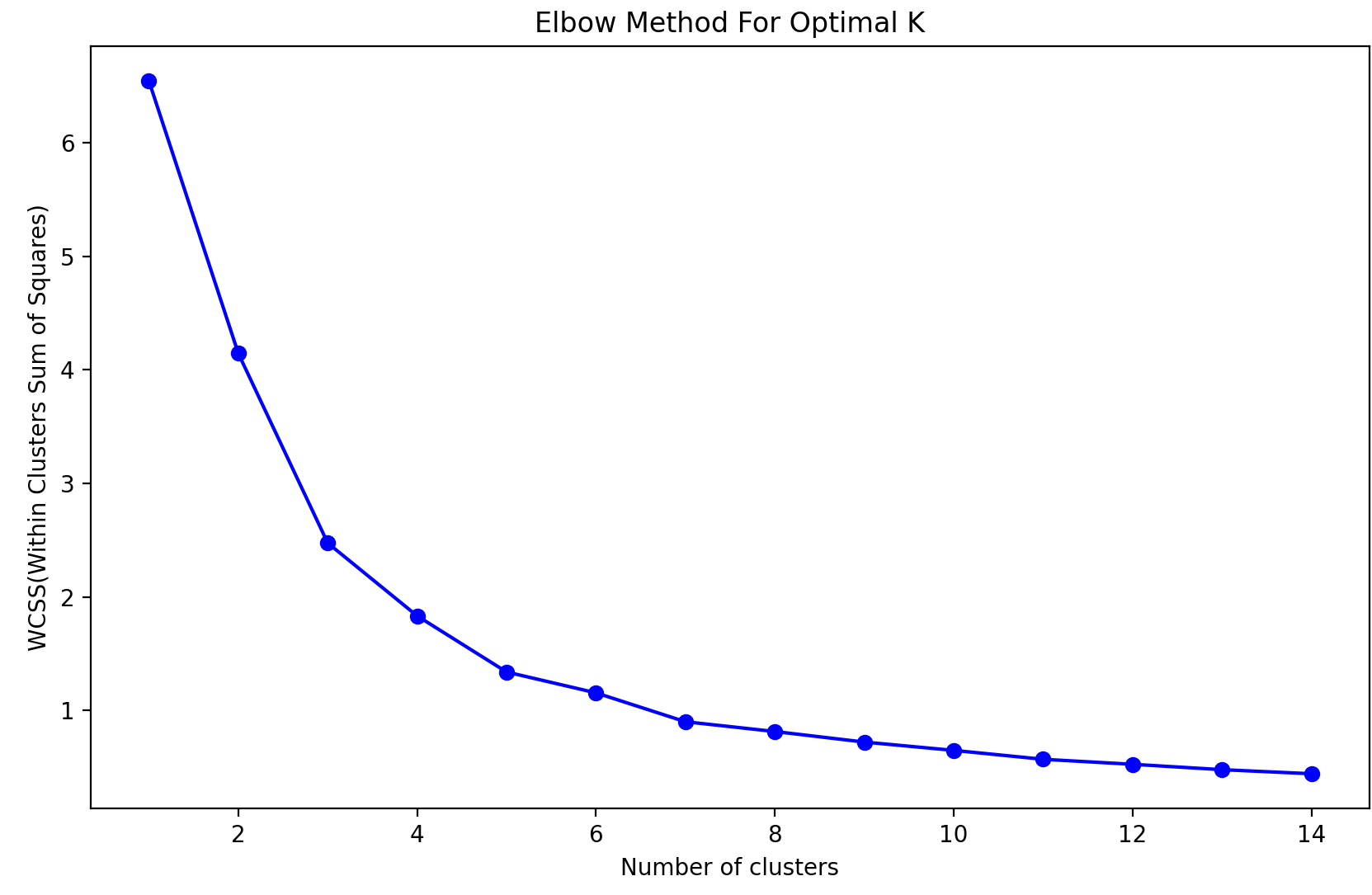
C = 클러스터의 중심 값

d = 클러스터 내에 있는 데이터

본 탐구에서는 sklearn 라이브러리의 cluster을 이용하여 kmeans의 적정 k 값을 찾은 과정을 진행하였다.



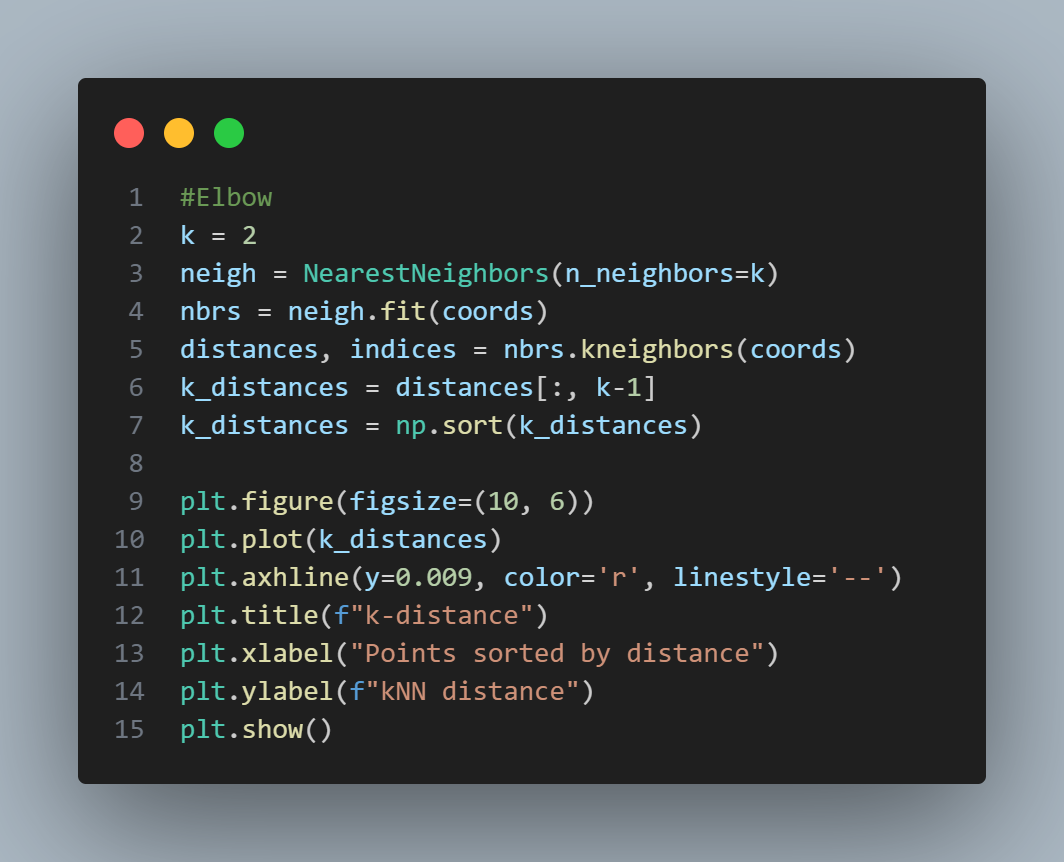
< 최적 K값 도출 Python 코드>



< K값에 따른 WCSS 변화 그래프>

위 그래프 출력 결과에 따라 k=5, 7로 설정하여 각각 k-means 군집 분석을 진행하였다.

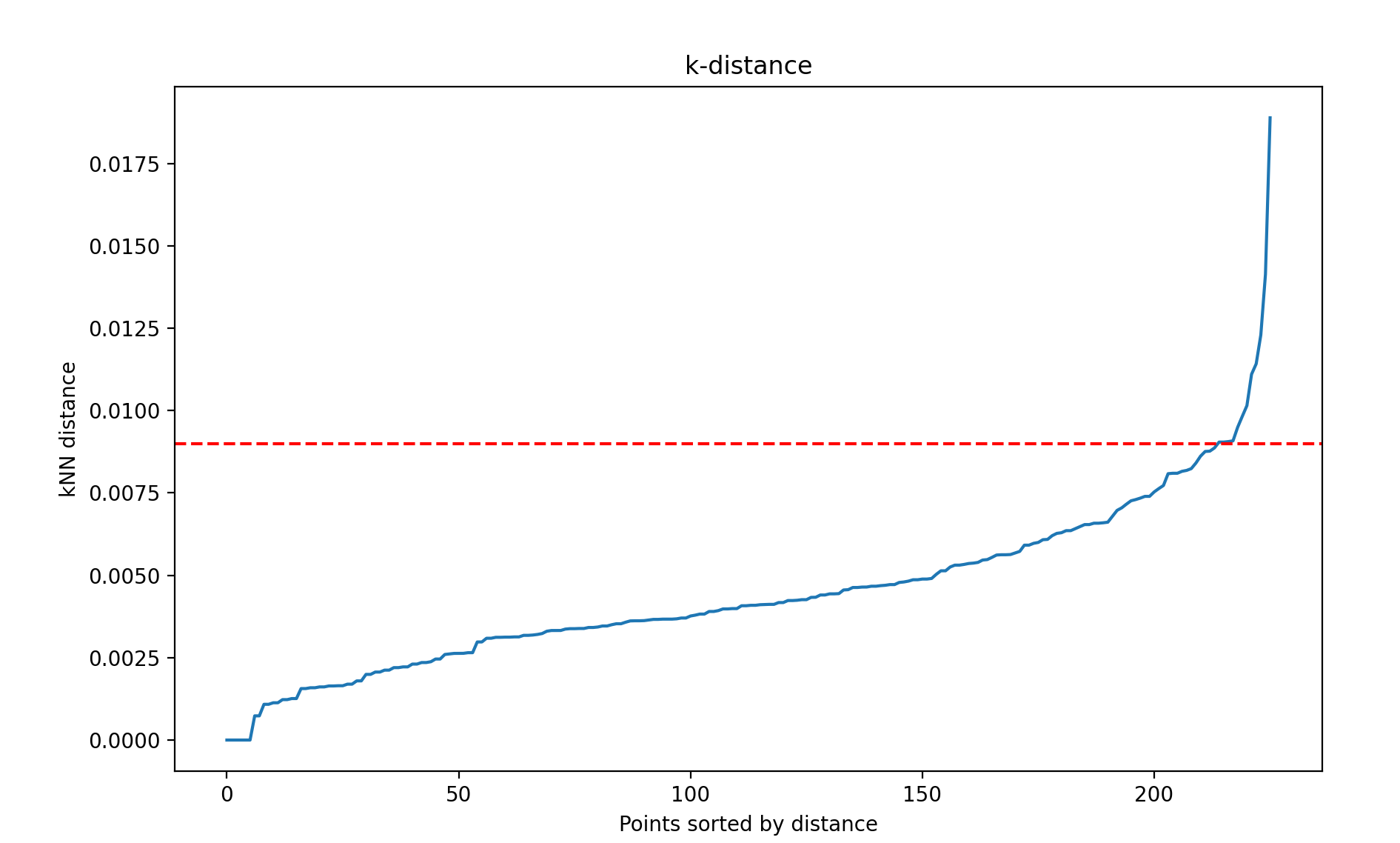
7. 대여소 입지 최적 모델 구현 – DBSCAN



< KNN 거리 Elbow method Python 코드>

Python의 sklearn.neighbors와 numpy를 활용해 knn 알고리즘을 구현하였다.

coord로 설정된 좌표 데이터들을 활용하여 knn을 적용시키고, 적용한 값에 대하여 오름차순 정렬하여 pandas 라이브러리를 활용하여 그래프로 시각화하였다.



< KNN Elbow method 실행 결과>



< DBSCAN 적용 시각화 코드>

DBSCAN 알고리즘은 주어진 데이터를 밀도 기반으로 클러스터링하는 비지도 학습 기법으로, 각 데이터 포인트를 밀도에 따라 클러스터로 그룹화하고 밀도가 낮은 포인트는 노이즈로 처리한다. 특정 반경(ε) 내에 포함된 데이터 포인트의 개수(minPts)를 기준으로 밀집된 지역을 클러스터로 정의하기 때문에 대전시 타슈 대여소가 위치한 지역의 밀도와 교통 패턴이 균일하지 않을 경우 적절히 활용될 수 있다.

DBSCAN 알고리즘의 주요 장점은 클러스터의 개수를 사전에 설정할 필요가 없으며, 비정형적 형태의 클러스터를 탐지할 수 있다는 점이다. 이를 통해 자전거 대여소의 위치 분석 시, 인프라와 교통 상황이 다르게 분포된 지역의 특징을 반영할 수 있다고 판단하여 각 대여소의 위치를 포함한 다양한 시설의 2차원 위도와 경도 데이터를 기반으로 클러스터링을 수행하였다.

DBSCAN 알고리즘을 적용하기 전, k-distance 플롯을 그려 Elbow Method를 통해 적절한 ε 값을 도출하였다. 플롯은 k-NN 거리 기반으로 생성되었으며, k-distance의 꺾이는 지점을 ε 값으로 설정하였다. 본 연구에서는 minPts 값을 4로 설정하였으며, 이는 DBSCAN의 원 논문(Martin Ester 외, 1996)에서 2차원 데이터에 적합하다고 권장된 값이다. 최종적으로 ε 값을 0.009로 설정하여 대여소와 주요 시설의 좌표 데이터를 클러스터링하였다. 밀도가 높은 지역은 하나의 클러스터로 정의되며, 이 외의 포인트는 노이즈로 간주되었다. 이러한 클러스터 중심 좌표는 자전거 대여소의 추가 설치를 위한 참고 지점으로 활용될 수 있다.

결과 ( Results )

대전광역시 전체의 타슈 대여소 분포율은 10,000명당 약 1.81개의 대여소로 나타났으며, 이를 기준으로 각 구의 분포율을 비교 분석하였다. 대덕구와 유성구는 각각 2.224785와 2.093335의 분포율을 보이며, 10,000명당 대여소 수가 상대적으로 높았다. 이는 이 두 구에 타슈 대여소가 비교적 잘 분포되어 있으며, 인구 대비 대여소 수가 많음을 시사한다. 따라서, 이들 구에서는 대여소의 추가 설치보다는 현재 대여소의 활용도와 효율성을 점검하는 것이 중요할 것으로 판단된다. 반면, 서구는 1.560986의 낮은 분포율을 보였다. 서구는 인구가 가장 많은 구(467,653명)임에도 불구하고 대여소 분포율이 낮아, 타슈 대여소가 부족할 가능성이 크다. 이에 따라, 서구에는 대여소를 추가 설치하여 접근성을 높이는 것이 필요하다. 특히 서구의 주요 생활권 및 교통 중심지에 대여소를 추가하는 것이 효과적일 것이다. 동구와 중구는 각각 1.646083과 1.735771의 분포율을 보이며, 대전광역시 전체 평균에 가깝거나 약간 낮은 수준을 나타냈다. 이들 구에서는 소규모의 추가 대여소 설치가 필요할 수 있으며, 인구가 밀집된 지역이나 교통 결절점에 대여소를 추가 설치함으로써 효율성을 높일 수 있을 것으로 보인다.

표 1. 구별 타슈 대여소 분포율 계산

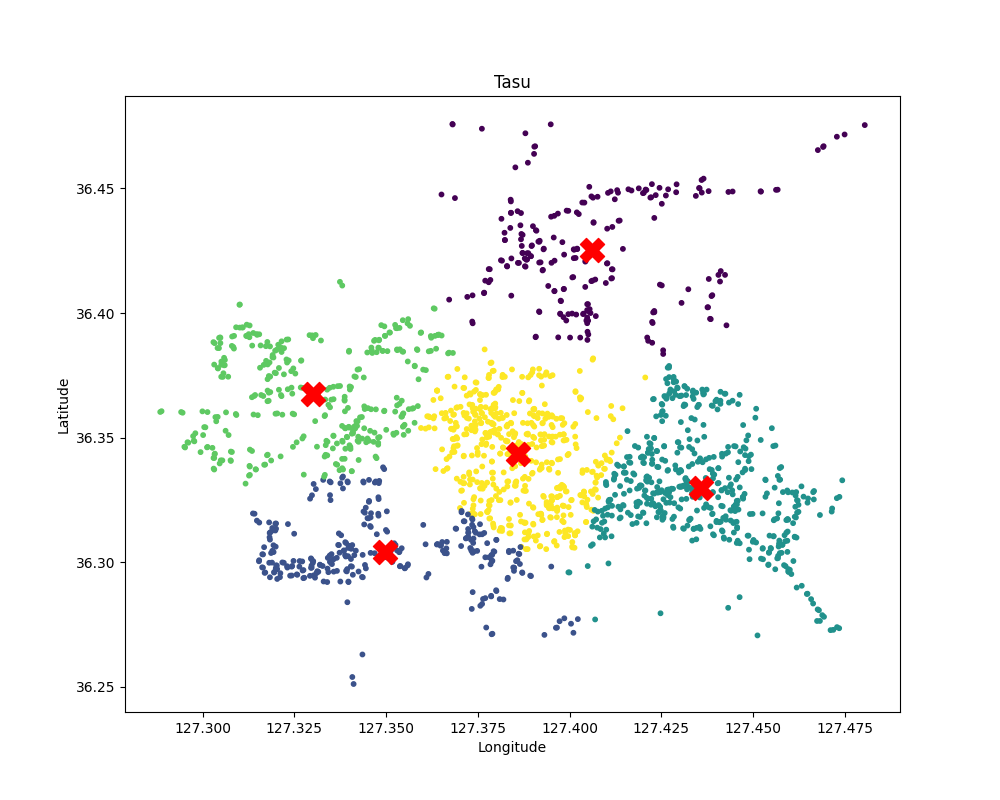
|  |  |  |  |
| --- | --- | --- | --- |
| District | Total\_  Population | Tasu\_Count | Distribution\_Rate |
| 대전광역시  동구  중구  서구  유성구  대덕구 | 1444898  218701  224684  467653  363057  170803 | 262  36  39  73  76  38 | 1.813277  1.646083  1.735771  1.560986  2.093335  2.224785 |

앞선 데이터 전처리 과정에서 진행했던 주소 – 좌표 변환 과정을 통해 수집한 CSV데이터들에 decimalLatitude, decimalLongitude 열을 각각 생성하고 지리정보 데이터 처리의 기하하적 연산과 시각화 등을 돕는 geopandas 라이브러리를 통해 geodataframe을 생성하여 병합하였다. 기존에 존재하던 타슈 대여소의 좌표 데이터들을 참고하여 병합한 데이터와 함께 k-means 군집 분석의 데이터포인트로 활용하였다.

1. K-means 군집 분석 결과



<데이터 병합 및 k-means 군집 분석 Python 코드>



< 주변 시설 인프라 활용 최적화, k=5 >

표 2. 주변 시설 인프라 활용 좌표 추출, (k=5)

|  |  |  |
| --- | --- | --- |
|  | Latitude | Longitude |
| 1  2  3  4  5 | 36.425472  36.304248  36.329578  36.367316  36.343447 | 127.406026  127.349625  127.435932  127.330216  127.385966 |

< 주변 시설 인프라 활용 최적화, k=7 >

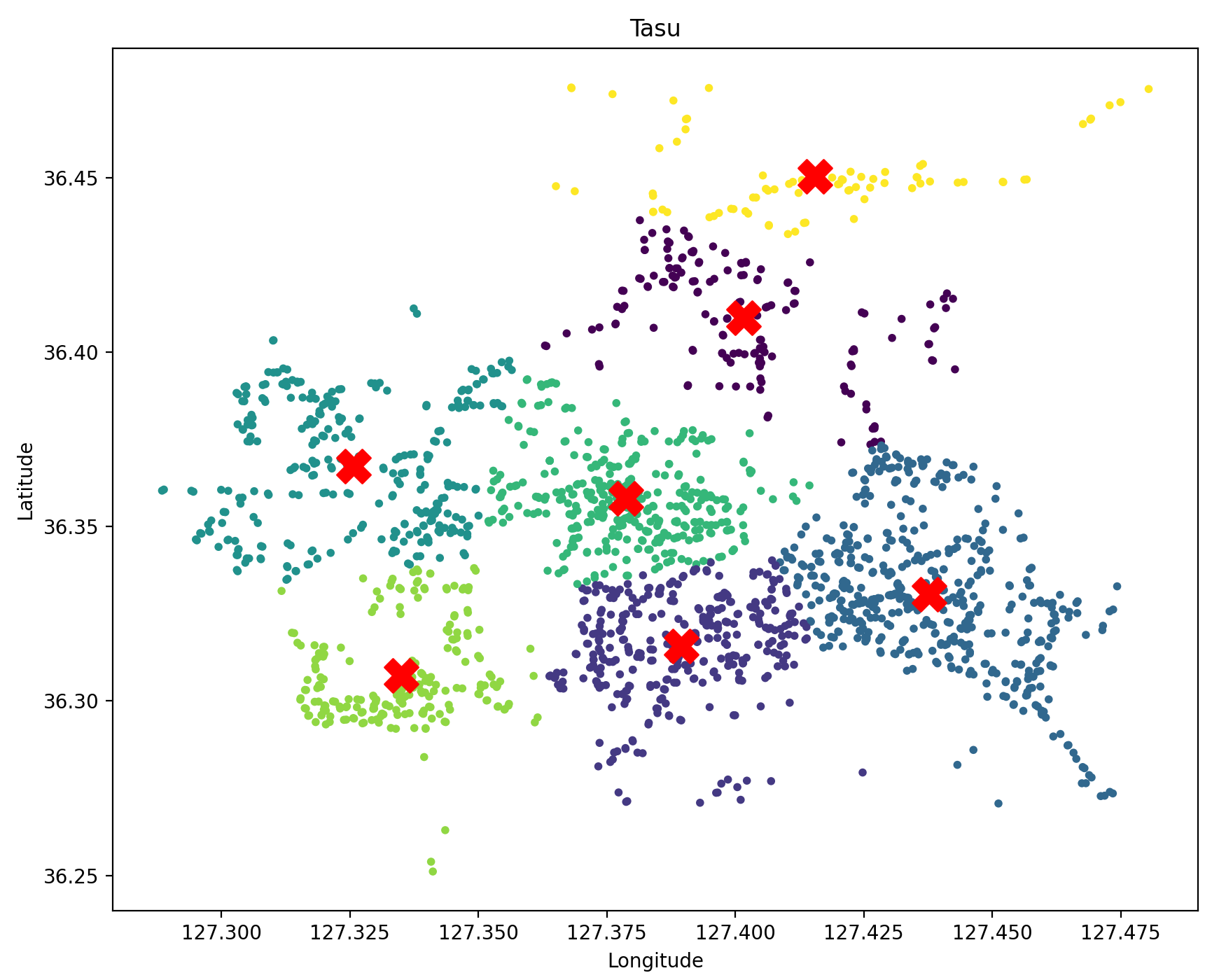


표 3. 주변 시설 인프라 활용 좌표 추출, (k=7)

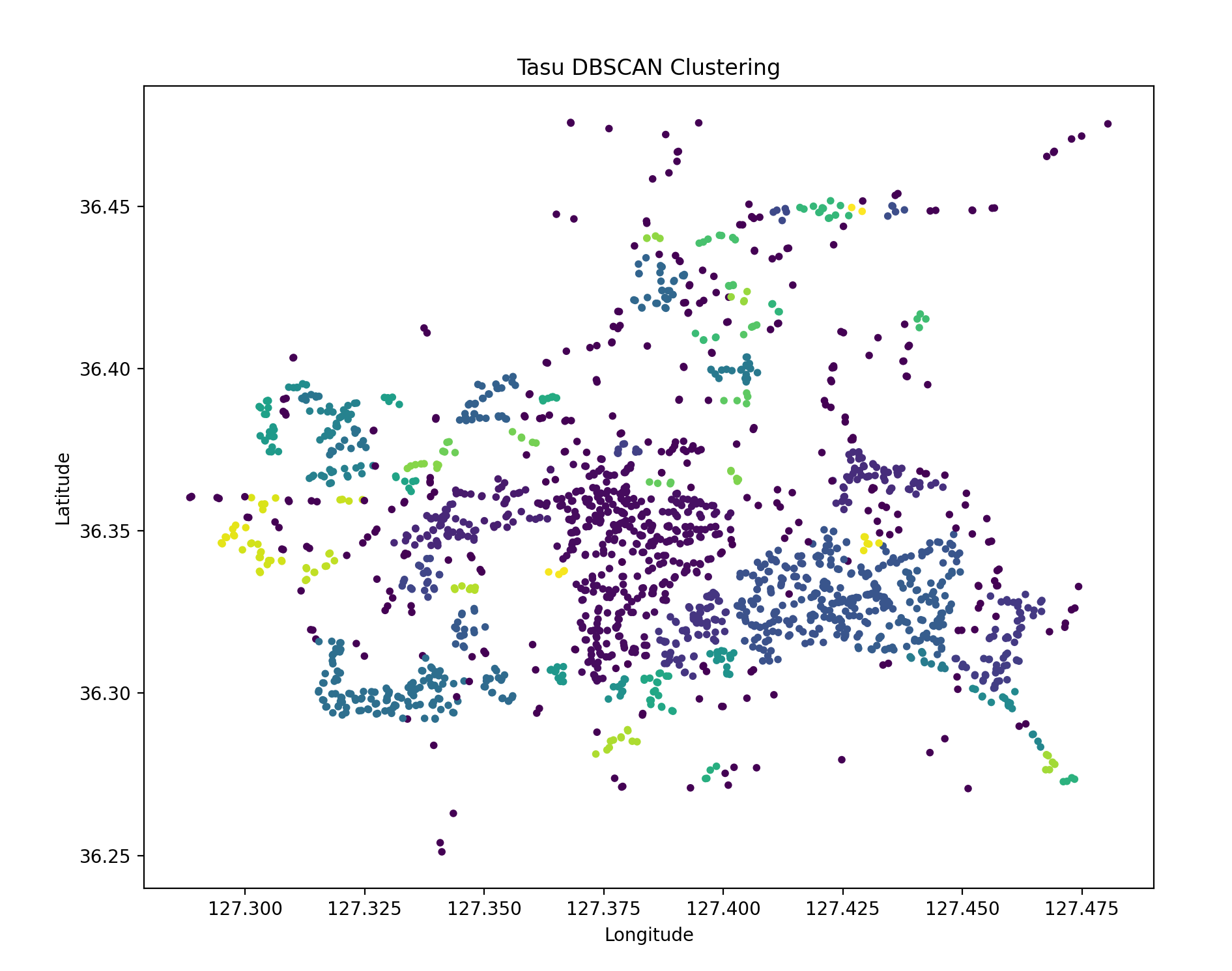
|  |  |  |
| --- | --- | --- |
|  | Latitude | Longitude |
| 1  2  3  4  5  6  7 | 36.409888  36.315632  36.330521  36.367322  36.358112  36.307359  36.450339 | 127.401599  127.389556  127.437701  127.325694  127.378693  127.334985  127.415533 |

K-means 군집 분석 결과, 총 7개의 군집 중심점이 도출되었으며 각 군집은 다양한 공공시설과 타슈 기존 대여소 위치의 데이터를 종합하여 특정 지리적 구역으로 분류되었다.

군집 1의 중심 좌표(위도 36.426602, 경도 127.406328)는 교통과 공공 서비스 시설이 밀집된 지역으로, 대중교통 접근성이 높고, 군집 2의 중심 좌표(위도 36.314381, 경도 127.387802)는 교육 기관과 상업 시설이 근접한 지역으로 해석되며, 군집 3의 중심(위도 36.367539, 경도 127.325781)은 공공장소와 학교가 집중된 지역으로, 주민들의 공공 서비스 이용이 활발한 중심지임을 확인할 수 있다. 군집 4의 중심(위도 36.358931, 경도 127.378454)은 다양한 시설과 타수 위치가 균형 있게 배치된 지역으로, 인프라의 안정적인 분포가 특징이다. 군집 5의 중심 좌표(위도 36.307618, 경도 127.334776)는 주거지와 공공 서비스의 근접성을 나타내며, 주민들의 이동 및 서비스 이용의 효율성을 높이는 지역으로 판단된다. 군집 6의 중심(위도 36.341616, 경도 127.427158)은 주요 상업 및 교통 지점과 가까운 곳으로 상업 활동이 활발하고 교통 편의성이 높은 곳으로 보인다. 마지막으로, 군집 7의 중심(위도 36.312399, 경도 127.452172)은 공공장소와 타슈 대여소가 고르게 분포된 지역으로, 공공 서비스 및 교육 시설이 넓게 퍼져있다는 특징이 있다.

이렇게 k-means를 활용한 군집 분석을 통해 도출된 중심 좌표들의 지리적, 상황적 요인을 분석하는 과정으로 주민들의 타슈 대여소 이용을 더욱 최적하게 하고, 타슈 대여소 분포에 대한 근거를 제공할 수 있을 것이다.

(2) DBSCAB 군집 분석 결과



DBSCAN 군집 분석 결과에 따라 타슈 대여소의 위치와 공공시설 좌표의 분포를 보다 유연하게 분류할 수 있었으며, 데이터의 밀집도를 기준으로 서로 다른 군집을 식별할 수 있었다. 이는 K-means와 달리 비구형의 군집도 효과적으로 감지할 수 있어 보다 복잡한 지리적 분포를 분석하는 데 적합하다.

분석 결과를 통해 71개의 군집의 중심이 도출되었다. 각 군집은 공공시설과 타수 위치가 밀집된 구역을 중심으로 형성되었으며, 그 외의 지역은 밀도가 낮아 군집화되지 않고 '노이즈'로 판별되었다. 군집 별 주요 특징은 다음과 같이 정리해 볼 수 있다.

군집 1 (위도 36.375455, 경도 127.391485): 도시 중심부에 위치하며, 공공 서비스와 교통의 접근성이 뛰어난 지역으로 분석된다. 이 지역은 대중교통 이용이 활발하고 인구 밀집도가 높은 것으로 판단된다. 군집 5 (위도 36.361551, 경도 127.351947): 공공시설과 교육 기관이 인접해 있는 지역으로, 주민들의 생활 편의성과 학업 활동이 집중된 구역이다. 군집 9 (위도 36.366999, 경도 127.429655): 도시 외곽에 위치한 군집으로, 인프라가 비교적 넓게 분포하며 주거지와 상업 시설의 혼재가 나타난다. 이는 도시 개발 및 확장 가능성을 보여준다. 군집 11 (위도 36.320057, 경도 127.394016): 상대적으로 작은 규모의 군집으로, 주변에 공공 서비스와 상업 시설이 적절히 분포되어 있어, 지역 주민의 생활 활동에 기여할 수 있는 구역이다. 군집 16 (위도 36.448051, 경도 127.412033): 비교적 외곽에 위치하며, 넓은 지역에 걸쳐 공공 시설과 주거지가 분포되어 있는 것으로 분석된다. 이 지역은 향후 도시 확장 시 개발의 잠재력을 가진다. 군집 25 (위도 36.302114, 경도 127.352885): 공공 서비스와 교육 시설이 밀집된 구역으로, 주민과 학생들의 접근성이 뛰어난 지역이다. 군집 30 (위도 36.309939, 경도 127.442697): 도시 경계선에 위치하며, 공공 서비스의 접근성이 상대적으로 낮을 수 있지만, 도심과의 연결성을 유지하는 지역이다. 군집 45 (위도 36.275297, 경도 127.397173) 및 군집 46(위도 36.273249, 경도 127.472329): 이 군집들은 타수 대여소의 주요 밀집 지역은 아니지만, 공공 서비스와의 접근성 확보를 위해 중요할 수 있는 외곽 지역으로, 도시 외부 지역의 교통과 연계되는 중요한 지점을 나타낸다.

군집 72 (위도 36.449056, 경도 127.427948): 가장 북쪽에 위치한 군집으로, 교외 지역의 공공 서비스와 대중교통의 분포를 반영하며, 교외 거주자들의 서비스 접근성을 보여준다. 밀도 기반으로 군집을 분류하여 특정 구역 내의 공공 서비스와 타슈의 상호 연관성을 파악하는 데 정보를 제공할 수 있으며, 이를 통해 도심 내의 인프라 개선이나 공공 서비스의 효율적 배치를 위한 도시 계획에 유용한 자료로 활용될 수 있다. 이 분석은 또한 군집화되지 않은 외곽 지역에 대한 추가 탐색의 필요성을 제안하며, 도시 확장 시 고려할 수 있는 기초 데이터를 제공할 수 있다.

결론(Conclusion)

이러한 최적 모델의 결과는 타슈 대여소 신설을 결정하는 데에 근거를 제공할 수 있는 조언자의 기능을 할 수 있을 뿐만 아니라 타슈 이용에 영향을 미치는 실데이터를 기반으로 진행한 과정이라는 점에서 이러한 과정을 통하여, 보다 효율적인 타슈 정책 운용에 대한 가능성을 제시할 수 있을 것이다. 타슈 대여소의 최적화는 운영 비용 절감과 효율성 향상을 목표로 한다. 이러한 인구 분포와 대여소 분포를 고려한 최적화를 통해 과밀하거나 부족한 문제를 해결하고 균형 잡힌 운영 체계를 구축할 수 있을 것이다. 데이터 분석을 활용하면 대여소 배치와 자전거 배분의 효율성을 높여 자전거 재배치 작업의 빈도를 줄이고, 인력 및 자원의 관리 효율성을 도모할 수 있다. 이는 자동차 이용을 줄여 교통 혼잡을 완화하고 대기 오염을 감소시키며, 시민들의 건강 증진과 의료비 절감에도 기여한다. 궁극적으로 시민들의 삶의 질 향상과 지속 가능한 도시 교통 환경 조성에 도움이 될 것이다.

참고 문헌REFERENCES

[1] 문현수 외 1명, 「대전시 공공 자전거(타슈) 공개 데이터 시각화 및 분석」, 「정보과학회 컴퓨팅」의 논문지, 2016년, 253 – 267쪽

[2] 김경옥 and 이창환. (2021). 공유자전거 시스템의 이용 예측을 위한K-Means 기반의 군집 알고리즘. 정보처리학회 논문지, 10(5), 169-178.

[3] 이우진, 전준헌. (2021). 머신러닝을 활용한 공공자전거 대여소의 입지 예측에 대한 연구. 한국IT정책경영학회 논문지, 13(4), 2553-2559.

[4] [1]K. Cho, S.-S. Lee, and D. Nam, “Forecasting of Rental Demand for Public Bicycles Using a Deep Learning Model,” The Journal of The Korea Institute of Intelligent Transport Systems, vol. 19, no. 3. The Korea Institute of Intelligent Transport Systems, pp. 28–37, 30-Jun-2020.

[5] Classifying Location Points as Daily Activities using Simultaneously Optimized DBSCAN-TE Parameters. Gregory S. Macfarlane, Gillian Riches, Emily K. Youngs, Jared A. Nielsen

[6] Martin Ester, Hans-Peter Kriegel, Jorg Sander, Xiaowei Xu, 1996, "A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise", KDD-96