

데이터 분석

k-Means 군집 기법을 활용한 최적입지 선정

(분석 대상: 미술관)

2024

문화 디지털혁신 및
문화데이터 활용 공모전

2024.05.08 ~ 2024.07.15

- 김민정 외 4인



TABLE OF CONTENTS

1.

개요

주제

- 분석 주제: 어린이 미술관 설립의 최적입지 선정
- 분석 Tool: Python, Tableau

분석 배경 및 필요성

2.

분석 내용 및 결과

EDA

1. 경기북부 입지 선정 이유
2. 파생 변수 생성
3. 상관관계 분석

Modeling

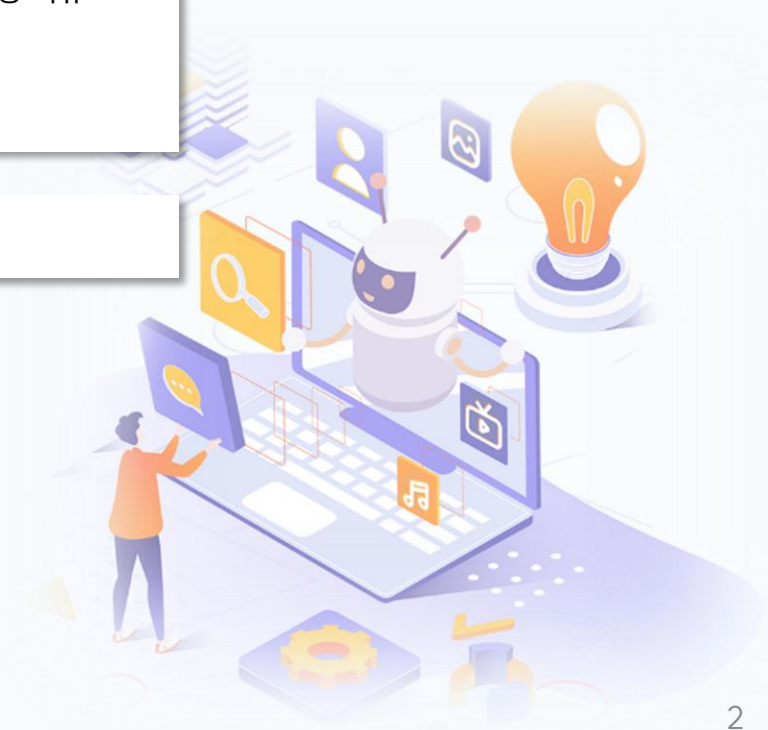
k-means 군집 분석

3.

시사점 및 기대효과

4.

문화 데이터 활용 및 출처



1. 개요

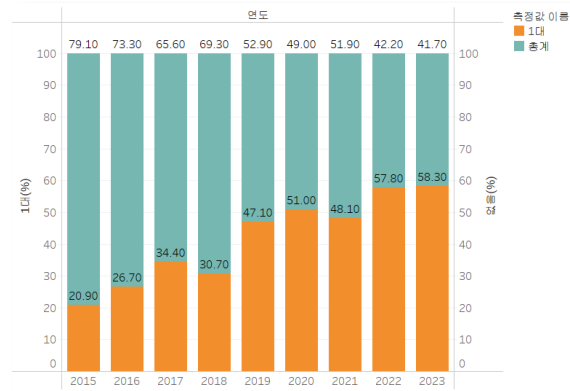
주제

- 분석 주제: **어린이 미술관 설립의 최적입지 선정**
- 분석 Tool: Python, Tableau

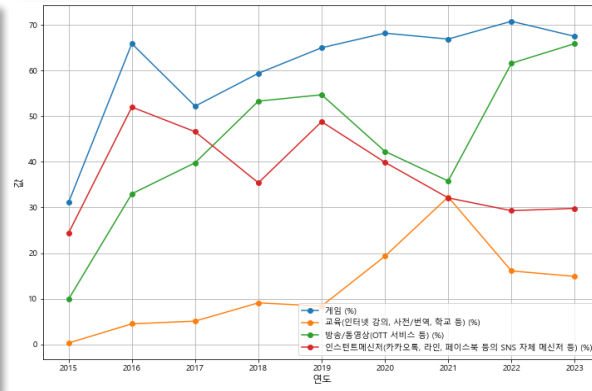
분석 배경 및 필요성

- 9세 미만 아동의 스마트폰 보유 비율의 증가
- 10대 미만 아동이 스마트폰으로 게임, 영상 매체 위주로 접하는 것으로 분석된다.

[연도별 스마트폰 보급율 (10대미만)]



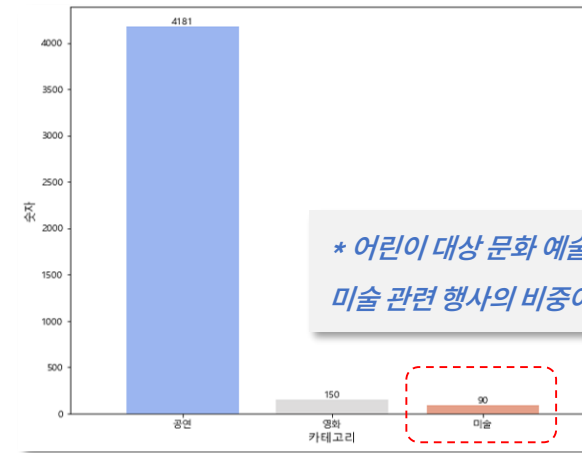
[연도별 스마트폰 어플 사용 현황 (10대미만)]



'미술'이 아동에게 미치는 긍정적 영향

- 영상 매체에 대조되게 '미술'은 아동의 발달에 상당한 긍정적 영향을 미침
1. **언어적 발달** - 다른 작품에 대한 이야기를 듣고 생각을 설명하거나, 내 작품과 그 과정에 대해 이야기하는 과정에서 언어 능력 향상
 2. **정서적 발달** - 자신의 생각과 감정을 자연스럽게 표현함으로써 사회적 상호작용 기회
 3. **인지적 발달** - 미술은 아이들의 창의력과 상상력을 자극하는 데 중요한 도구로서 다양성 존중과 개방적 태도로 다름을 수용하고 통합하는 능력 향상

[어린이 대상 문화예술 프로그램 수]

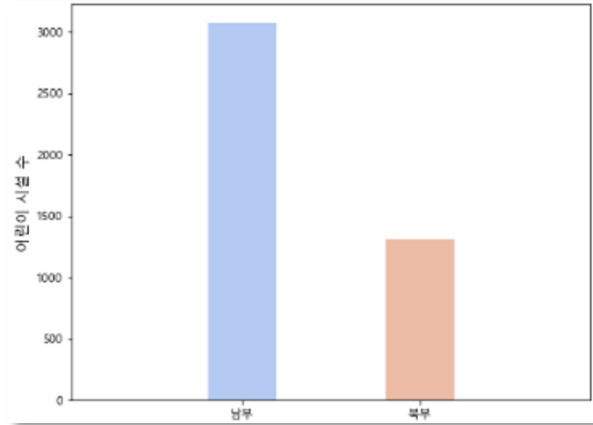


* 어린이 대상 문화 예술 프로그램의 수를 집계했을 때, 공연에 비해 미술 관련 행사의 비중이 매우 낮음

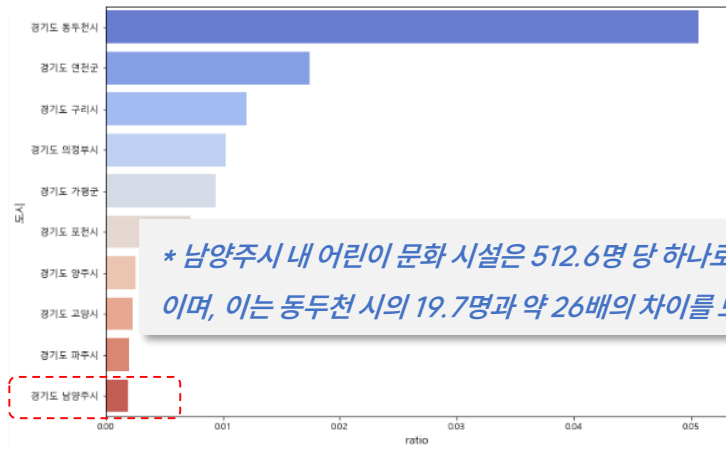
2. 분석 내용 및 결과 – EDA

'경기 북부' 입지 선정 이유

[경기 남부 / 북부 어린이 문화 시설 개수]



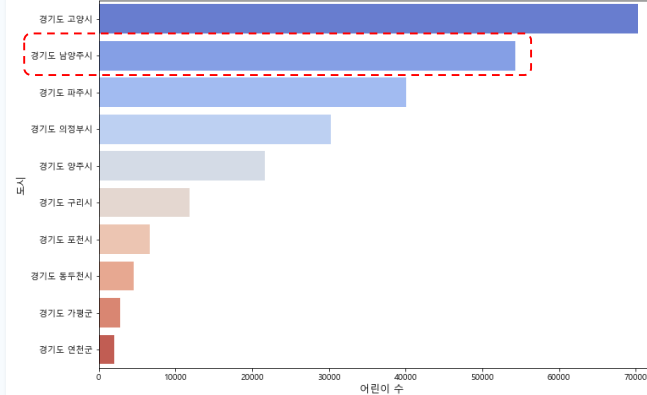
[경기 북부 시 별 어린이 문화 시설 비율]



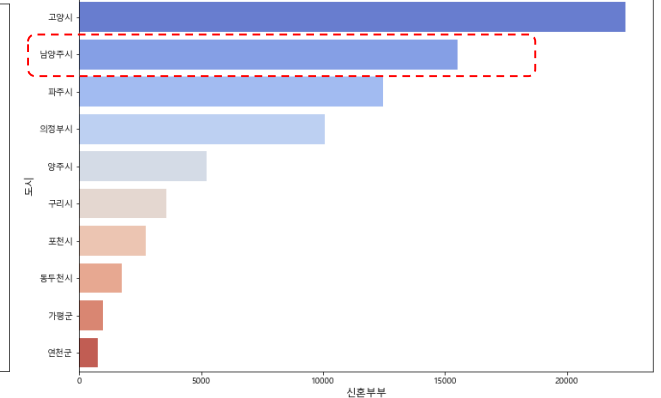
* x축: 비율 = (시 별 어린이 문화 시설 수) / (시 별 어린이 인구 수)

경기 북부 중 '남양주시' 입지 선정 이유

[경기 북부 9세 미만 어린이 수]



[경기 북부 신혼부부 수]



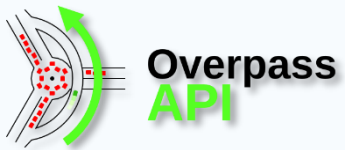
- 어린이 미술관이 위치할 지역으로 선정하는 데 있어 주요 타겟층인 9세 미만 어린이가 많이 거주하는 지역이 적합
- 추후 아이를 키울 확률이 높은 신혼 부부가 많이 거주하는 지역은 어린이 미술관의 적절한 입지를 고려할 수 있는 요소 중 하나
- 남양주시에 거주하는 9세 미만 어린이는 54,337명, 신혼 부부는 15,517쌍으로 경기 북부 중 높은 순위에 위치해 있음

2. 분석 내용 및 결과 – EDA

파생변수 생성

미술관의 성공 여부와 어린이 관련 지수 간 상관 관계 탐색을 위해 파생 변수를 생성했음

파생변수 이름	파생 변수 수식
CAI (어린이 문화 시설 접근성 지수)	(시 별 어린이 전용 문화시설 수 + 어린이를 위한 공연 정보 수 + 초등학교 수 + 유치원 수) / (시 별 지역 면적)
CFI (어린이 친화지수)	(9세 미만 어린이 인구 수 + 시 별 신혼 부부 수) / (해당 지역 인구 수)
CEI (어린이 교육 인프라 지수)	(초등학교 수 + 유치원 수) / (해당 지역 어린이 인구 수)
CFI (가족편의 지수)	(시 별 신혼 부부 수 + 주차장 수 + 유치원 현황 수) / (해당 지역 인구 수)
MSI (미술관 성공 지수)	연 미술관 관람객 수 / (해당 연도 미술관 소재지 총 인구 수)



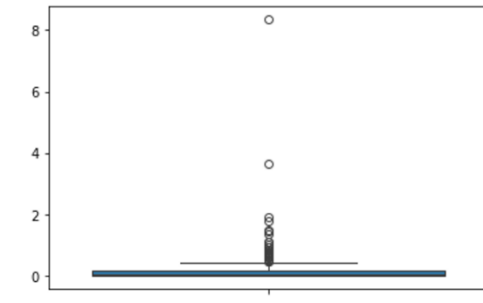
미술관의 성공 여부와 주변 편의 시설의 분포 간 상관 관계 탐색을 위해 위치 정보를 13개의 태그로 분류하는 Open Pass API를 이용해 각 미술관의 주변 2 km 이내 시설 정보를 수집함

13 Tags

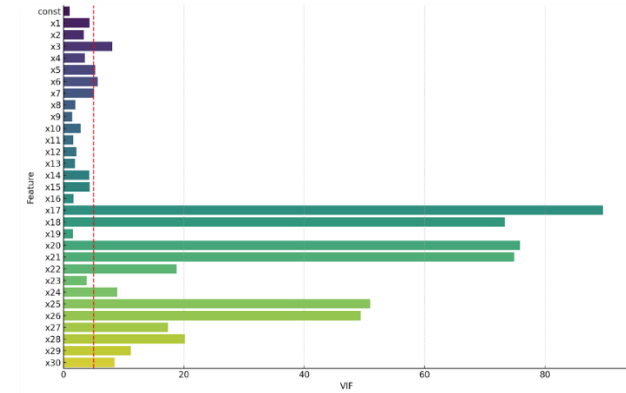
✓ '공공 편의 시설'	'도로 및 교통 요소'
'지명 및 장소'	'관광 명소'
'경계 및 행정 구역'	'수로'
'기타 건물'	'토지'
'상업 시설'	'여가 및 오락 시설'
'철로'	'자연 지형'
'인공 구조물'	

회귀분석 시도

미술관 성공 지수(MSI)를 종속 변수로, 어린이 관련 파생 변수와 주변 시설 변수를 독립 변수로 하는 회귀 분석을 적합하기 위해 회귀 모형의 기본 가정이 만족되는지 확인함



미술관 성공 지수 값의 분포를 살펴 보면, 거의 모든 값들이 상한인 약 0.44보다 작으며, 다른 미술관들은 이상치로 취급되는 데이터의 불균형 때문에 정규성 가정 불만족

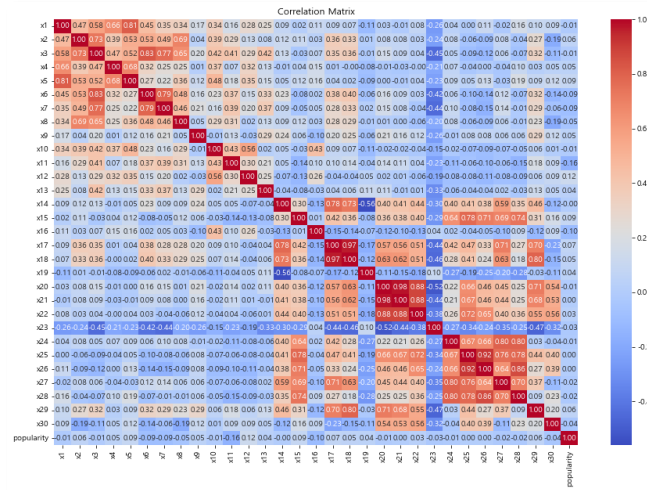


독립 변수들 간 분산 팽창 지수를 계산해보면 대부분의 변수들 간 다중공선성이 나타나는 것을 확인 (x축: VIF)

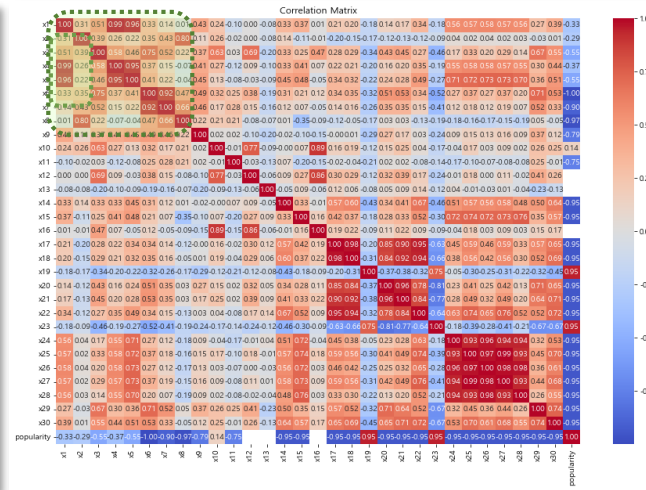
2. 분석 내용 및 결과 – EDA

상관관계 분석

높은 다중공선성, 정규성을 따르지 않아 회귀 분석을 사용하기에 부적합한 사례이므로 미술관 성공 지수(MSI)와 다른 변수 간 상관 분석을 진행



[Correlation of popularity < 0.44]



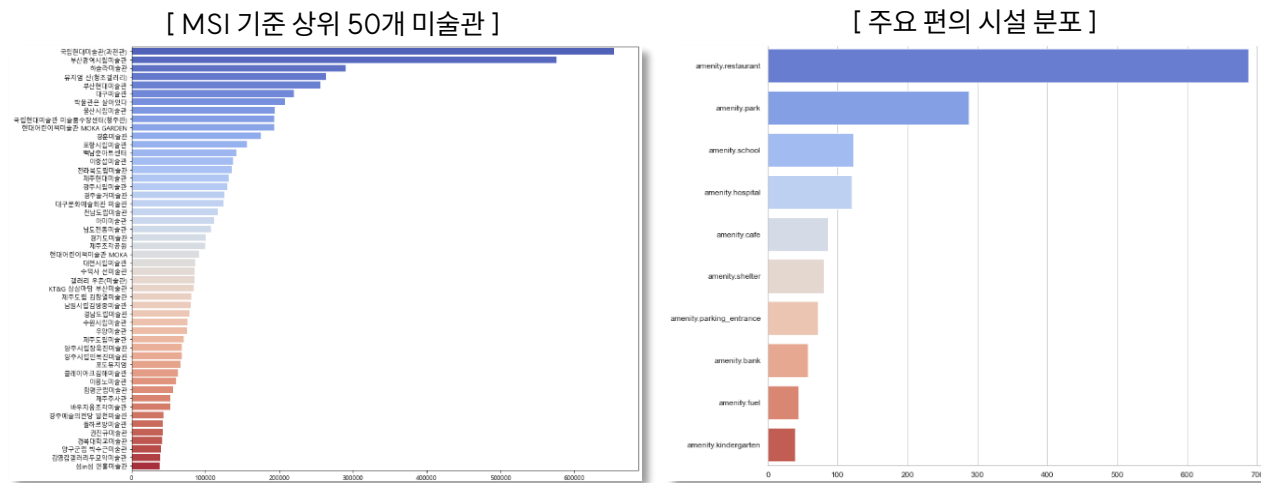
[Correlation of popularity >= 0.44]

- (좌) MSI 지수가 0.44 미만인 일반적인 미술관에서는 미술관 성공 지수와 유의미한 양의 상관을 보이는 변수가 존재하지 않았음.
- (우) 반면, MSI 지수가 0.44 이상에 속하는 성공하는 미술관에서는 CAI, CEI 변수가 미술관 성공 지수와 각각 0.98, 0.97의 상관 계수 값에서 강한 양의 상관을 확인 가능 (초록색 박스)
- 또한 공공 편의 시설과 상업 시설, 관광 명소 변수의 값들이 서로 강한 양의 상관 관계를 보이는 것에서 성공하는 미술관은 외부의 인적 유입을 가져오는 효과를 기대할 수 있음.

2. 분석 내용 및 결과 – EDA

상관관계 분석

Overpass API를 이용하여 방문객 수 상위 10개 미술관 주변 2km 반경의 어린이 관련 시설 개수를 조사



- 이들 주변에는 음식점, 공원, 학교, 병원, 카페, 쉼터, 주차장, 은행, 주유소, 유치원 등이 주로 위치하고 있음
- 앞의 상관 분석 결과를 토대로 **유치원, 초등학교, 공원, 도서관, 주차장, 키즈카페, 문화 시설**을 군집 분석의 변수로 선택함

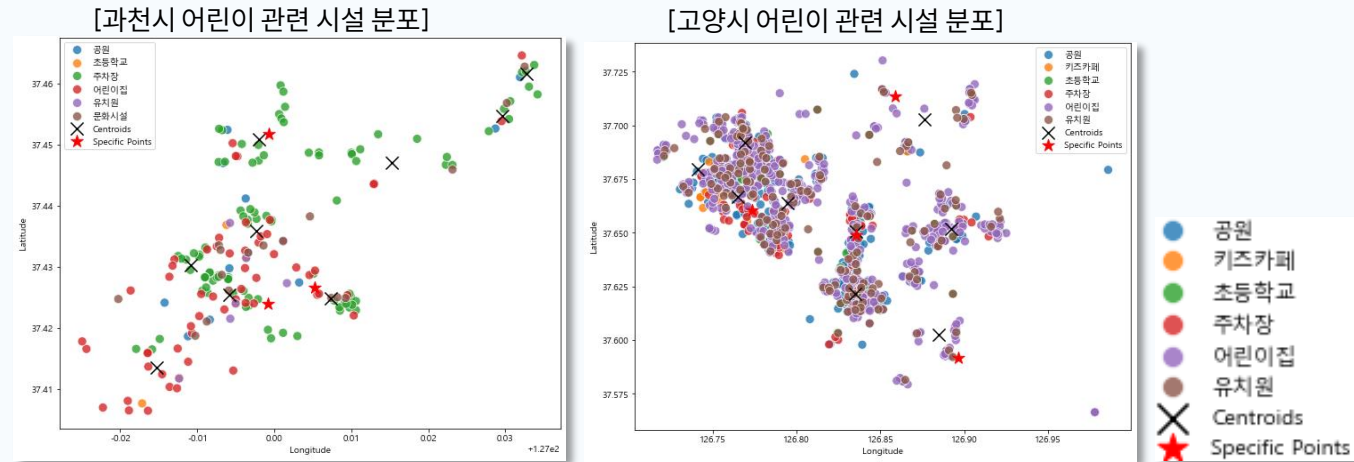
2. 분석 내용 및 결과 – Modeling

k-means 군집 분석을 통한 입지 선정 계획의 유의성 확인

k-means 란?

- 데이터를 K개의 군집(Cluster)으로 묶는(Clustering) 알고리즘
- 가깝게 위치하는 데이터를 비슷한 특성을 지닌 데이터로 여기고 같은 군집으로 군집화

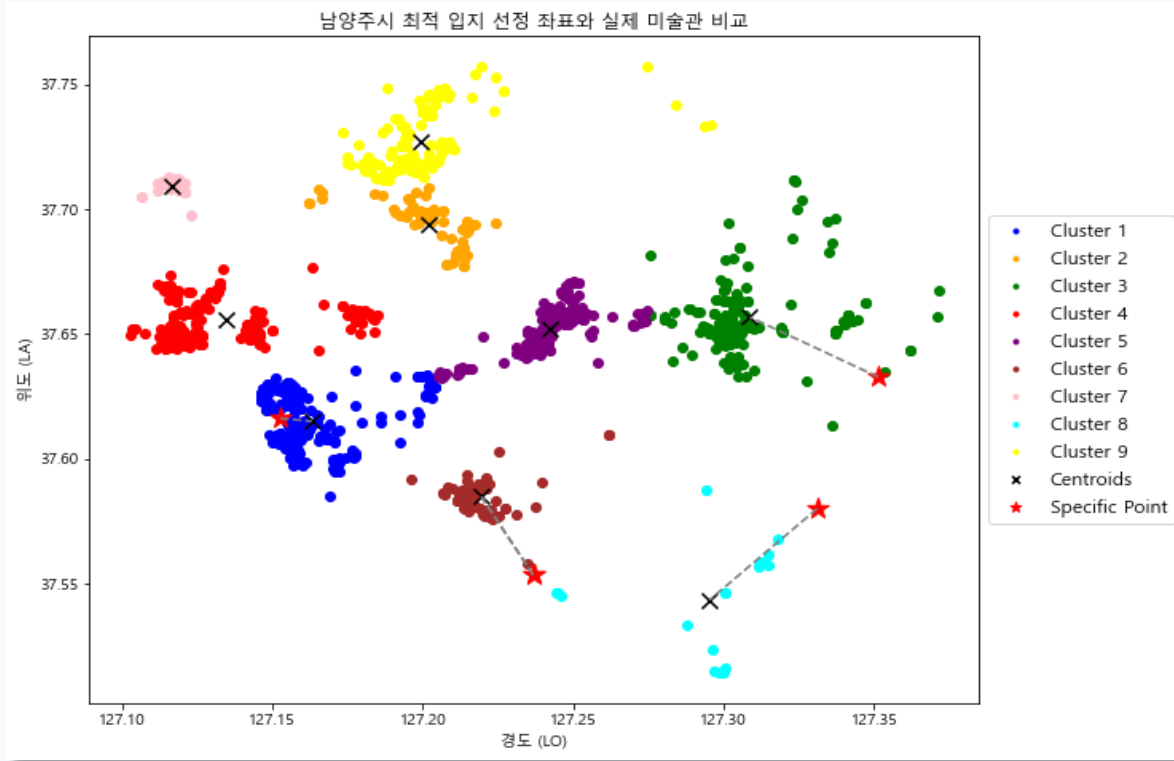
- 앞서 선택한 주요 시설 변수를 통한 군집 분석의 결과가 실제 미술관의 최적 입지를 추천하는지 그 유의성을 평가하기 위해 경기도 내 과천시와 고양시에서 먼저 k-means 군집 분석을 진행함
- 과천시 모델 Silhouette 계수는 약 0.46, 고양시 모델 Silhouette 계수는 0.44로, 군집 분석의 모델은 유의



과천시와 고양시 내 미술관 좌표와 분석의 클러스터 중심점을 비교한 결과, 대부분의 미술관이 중심점의 1km 이내에 존재함을 확인함

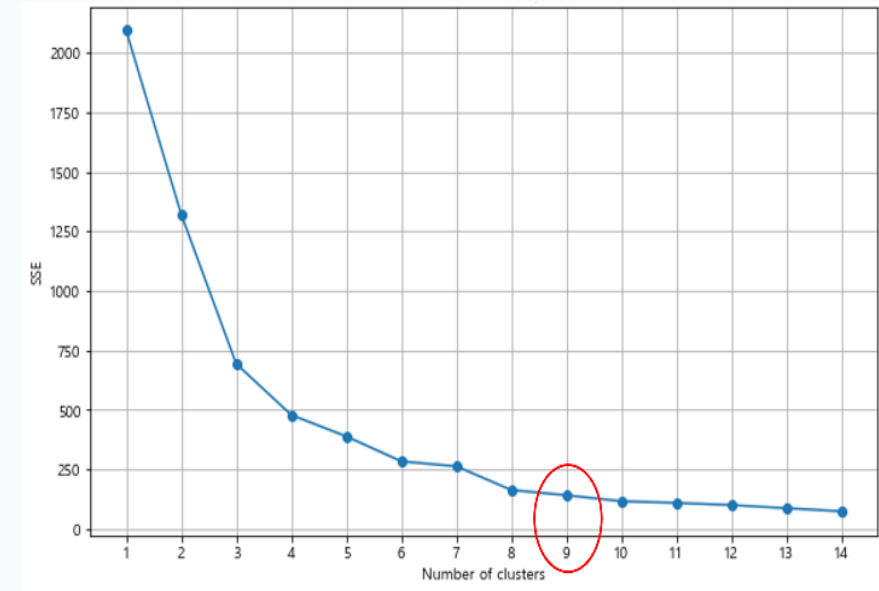
2. 분석 내용 및 결과 – Modeling

남양주시 k-means 입지 선정 군집 모델 구축



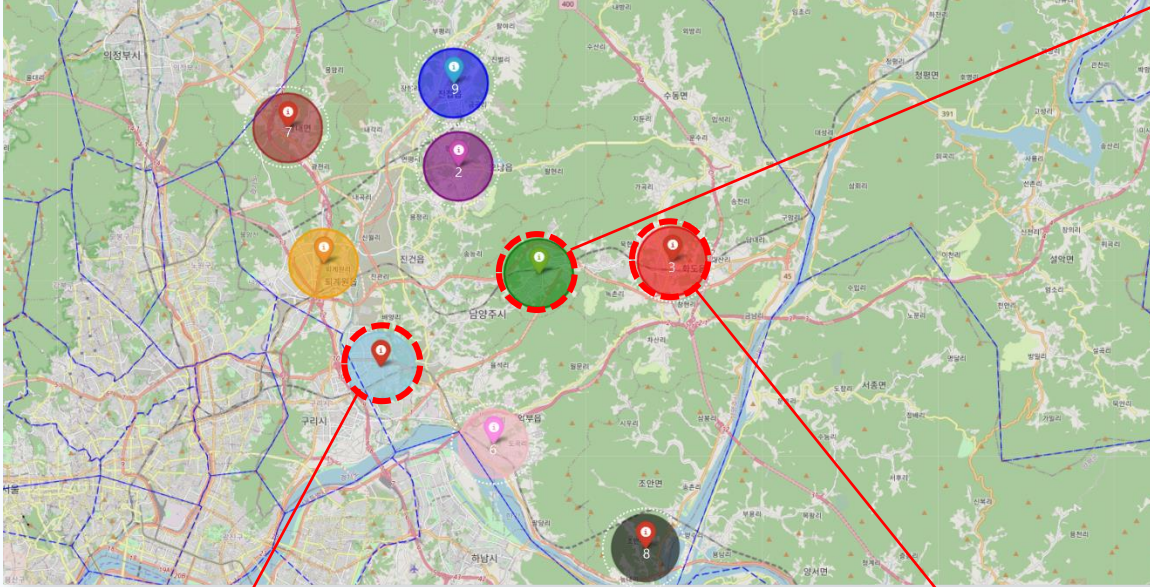
* 좌표 내용: 공원, 키즈카페, 초등학교, 주차장, 어린이집, 유치원, 문화시설

최적 군집 수(k) 결정을 위해 엘보우 기법(Elbow method) 사용



2. 분석 내용 및 결과 – Modeling

남양주시 K-means 입지 선정 군집 모델 구축 (검토 사항)



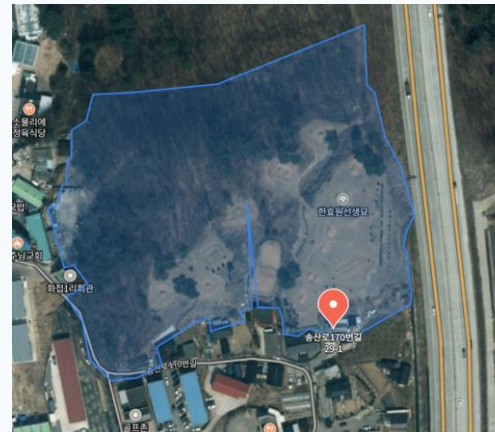
Cluster 1

중심 좌표와 1.3km 내에 '모카가든 미술관'이 존재하여
최종 입지 선정에서 제외



Cluster 5

민선8기 문화분야의 공약사업인 '종목별 생활체육시설 확충 및 운영지원 확대' 사업으로 현재 평내 체육 문화 시설 착공으로 입지 선정에서 제외

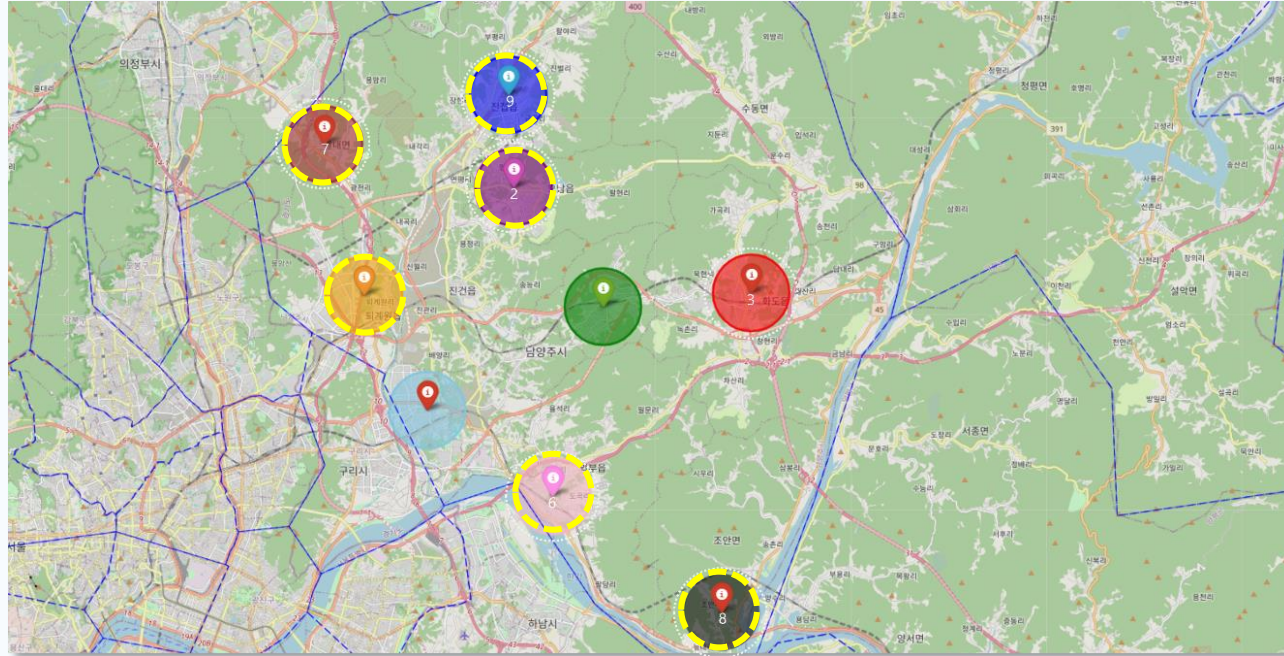


Cluster 4

영의정 한효원 선생의 묘가 있기 때문에 입지 선정에서 제외

2. 분석 내용 및 결과 – Modeling

결론



▶ 앞의 세 곳을 제외한 Cluster 2, Cluster 3, Cluster 6, Cluster 7, Cluster 8,
Cluster 9에 어린이 미술관 설립을 위한 입지로 제시해 볼 수 있음

3. 시사점 및 기대효과

시사점

- 분석에 따르면 최근 유아와 아동의 스마트폰 사용이 급증하고, 영상 매체에 대한 과의존 경향이 심해짐에 따라 언어 발달 및 사회적 상호 작용에 부정적인 영향이 있을 것이라 우려되고 있음

발달 시기 잦은 미디어 노출은 두뇌 성장을 방해하므로, 이와 대조되는 정적 활동과 교육을 위해 가정 단위에서 노력할 뿐만 아니라 시군구 역시 적극적인 투자를 해야 함

- 경기도 내 어린이를 대상으로 하는 문화 시설이 비교적 불균형한 분포를 보이고 있음

경기 북부는 남부에 비해 어린이 대상 문화 시설이 약 2.3배 적으며, 특히 남양주시는 어린이 수와 신혼 부부 수가 많은 것에 비해 어린이 대상 문화 시설은 내부 수요를 온전히 만족시키지 못하는 상황이라 확인되고 있음

- 성공하는 미술관은 CAI 지수, CEI 지수와 관련된 편의 시설인 공원, 주차장, 도서관, 유치원, 초등학교, 키즈카페 등이 많은 구역 근처에 분포하고 있음

이는 일반적으로 미술관이 많은 관람객을 유치하는 것은 많은 어린이가 쉽게 접근할 수 있도록 하는 것과 높은 상관 관계가 있음을 의미함.

기대효과

- 어린이 미술관 설립을 통해 어린이 대상 문화 시설의 지역적 불균형 해소
- 이는 문화적 참여의 기초를 다질 수 있게 하며, 결과적으로 지역 사회의 문화적 수준을 높이는 데에 기여
- 어린이 미술관은 가족 단위의 외부 활동을 장려
- 지역 경제가 활성화되는 데 긍정적인 영향을 기대할 수 있음 (일례로 홍대 혹은 한남동 등의 사례를 고려하면 지역 내 예술가들의 유입으로 인한 소문화의 형성과 지역 내 풍경의 변화는 다양한 일자리를 창출하고 상권을 형성함)



4. 문화데이터 활용도 및 출처

분류	제공 기관	데이터명	내용	URL
분석 배경 및 목적	한국문화정보원	아이랑 문화 생활 정보	어린이 미술관 시설이 다른 문화 시설에 비해 적음을 시각화	https://url.kr/vx5gkj
	한국문화정보원	어린이를 위한 공연 정보	어린이 대상 문화 프로그램은 대부분 공연이며 미술의 사례는 적다는 점을 시각화	https://url.kr/5otqln
	한국미디어 패널 조사	10대 미만 스마트폰 및 키즈폰 보유 대수	스마트폰 보급률 변화 시각화	https://url.kr/wnxp9v
	한국미디어 패널 조사	자주 사용하는 스마트기기 애플리케이션 1순위 (10대 미만)	어린이 스마트폰 사용 행태 조사	https://url.kr/wafbmb
지역 선정	통계청	시군구 별 신혼 부부 수	최적 입지 지역 선정 기준 최적 입지 관련 파생 변수 생성	https://shorturl.at/hJBXu
	주민 등록 인구 통계	행정동 별 연령별 인구 현황	최적 입지 관련 파생 변수 생성	https://shorturl.at/p8DLI
	레드타이	어린이 대상 문화 공간	최적 입지 관련 시설 변수 생성	https://shorturl.at/FliVG
	한국문화정보원	국내 문화체육관광 분야 국립 도립 군립 및 도시 내 공원 데이터	최적 입지 관련 파생 변수 생성	https://shorturl.at/HAAHE
	공공데이터포털	전국 초중등학교 위치 표준 데이터	최적 입지 관련 파생 변수 생성	https://shorturl.at/U9Efv
	공공데이터포털	전국 주차장 정보 표준 데이터	최적 입지 관련 파생 변수 생성	https://url.kr/lpzibq
	Overpass API	API 객체 데이터	최적 입지 선정의 변수 선택을 위한 주변 편의 시설 분포 탐색	Rest API
	한국문화정보원	국내 문화 체육 관광 분야 미술관 시설 및 현황 통계 데이터 (방문객 등 통계 포함)	상관 분석 및 군집 분석	https://url.kr/piuayo

