

중소규모 공연예술단체 활성화를 위한 키오스크 입지 선정

2021 데이터 청년 캠퍼스 동국대 B반 5조

01 배경 및 주제 소개

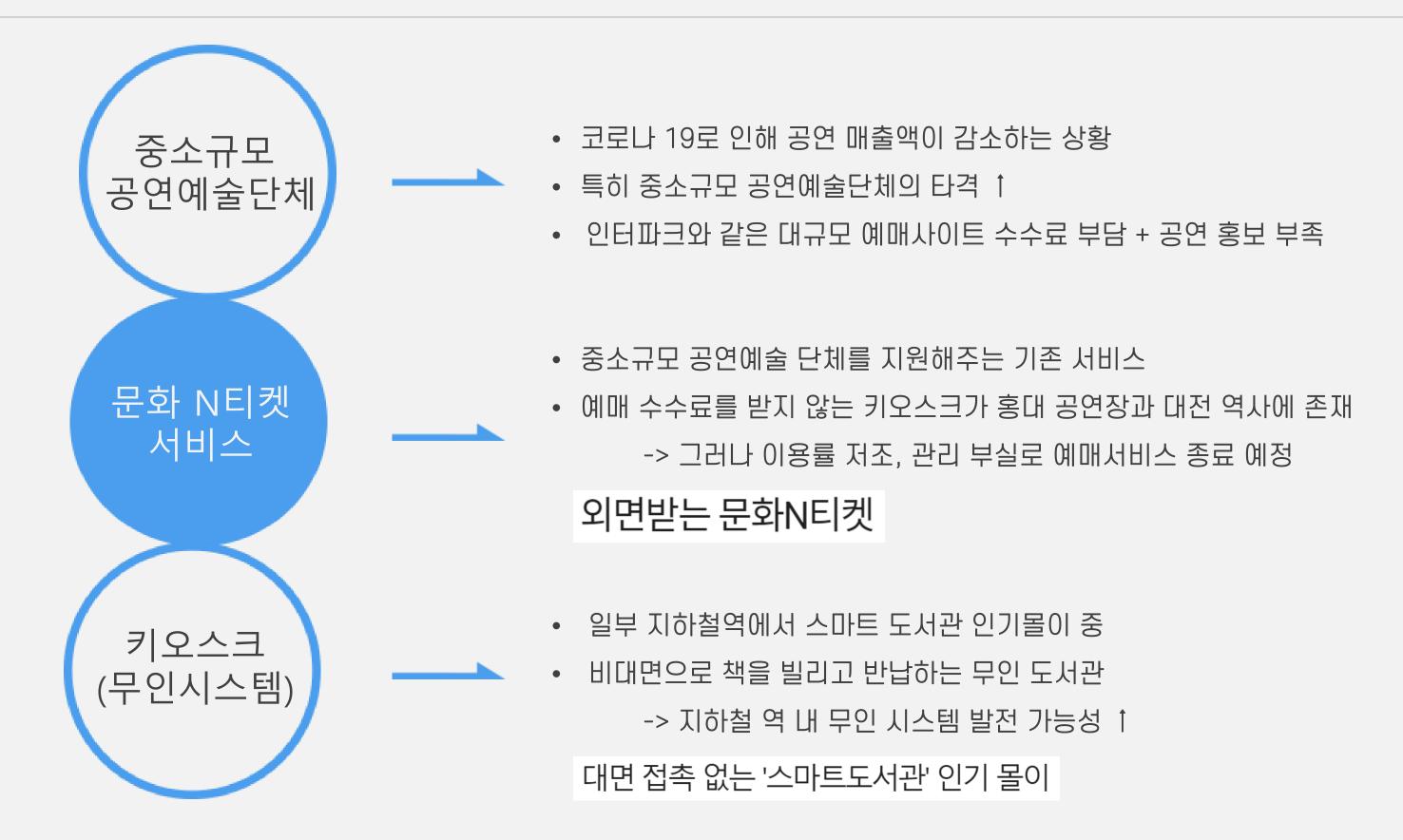
02 데이터 전처리 및 탐색

03 모델링 및 분석

04 분석 결과 정리

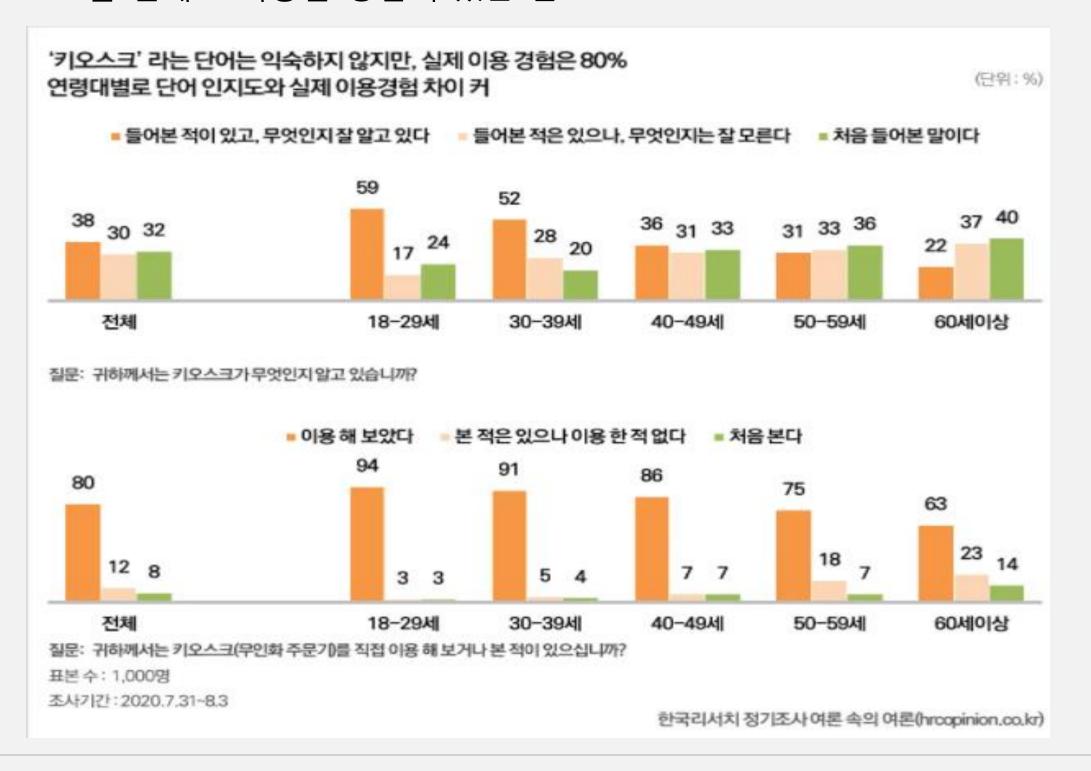
05 기대효과 및 발전방향

배경 및 주제소개



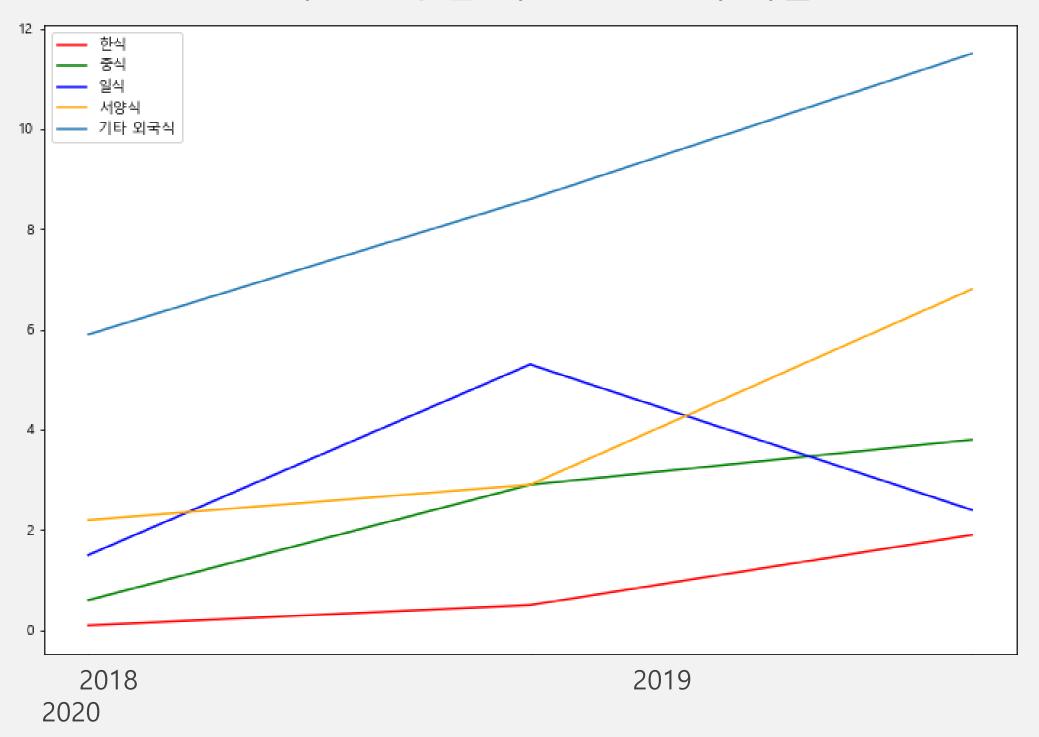
- 1) 외면 받는 문화 N티켓: http://www.ggilbo.com/news/articleView.html?idxno=831617
- 2) 대면 접촉 없는 '스마트도서관' 인기 몰이: https://www.koit.co.kr/news/articleView.html?idxno=87676

키오스크 이용인식 조사결과, "키오스크"라는 단어 낯설지만 실제 이용률은 80% 이상 10명 중 9명 이상이 키오스크를 실제로 이용한 경험이 있는 셈



^{1) [}기획] 키오스크(KIOSK)이용, 얼마나 편하신가요? – 키오스크 이용인식 조사: https://hrcopinion.co.kr/archives/16282

음식점 업종별 키오스크 설치 비율



> 음식점 업종별 키오스크 사용여부

대부분의 음식점 업종에서 키오스크 설치 비율이 점점 증가하는 추세 앞으로 키오스크 사용이 더 늘어날 것으로 예상

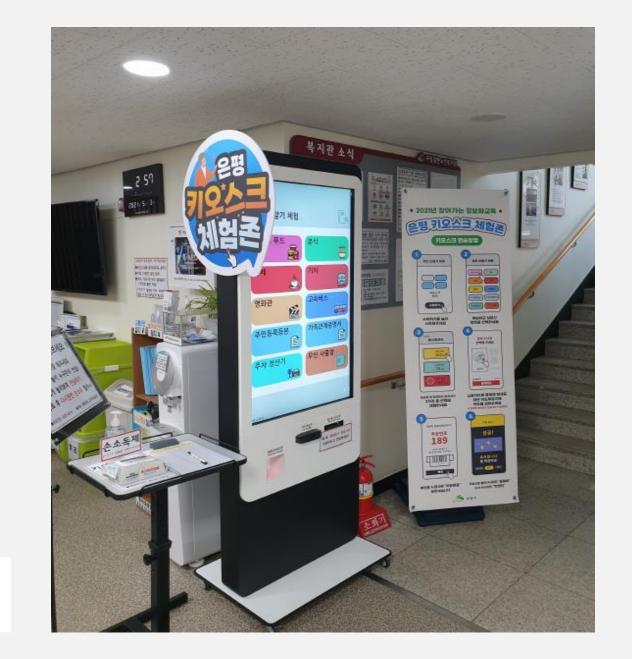


대구광역시, 관광지 여행정보 호텔 내 키오스크'로 한눈에

키오스크 산업 활성화



어르신들을 위한 교육용 '키오스크 체험존' 운영



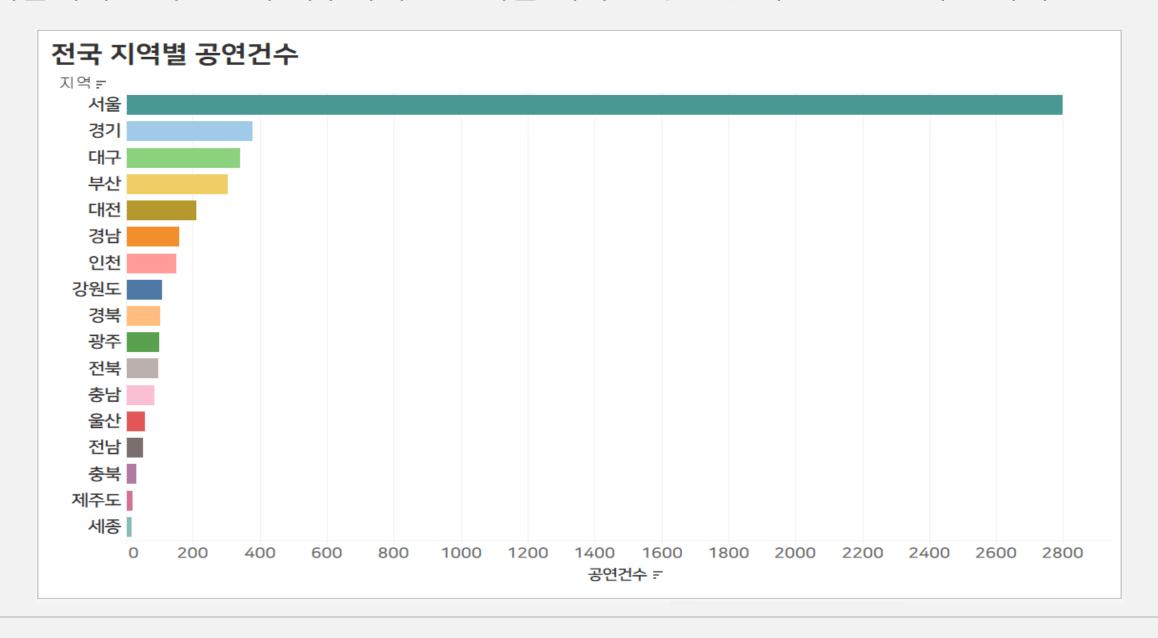
¹⁾ 대구광역시, 관광지 여행정보 호텔 내 키오스크로 한눈에: http://www.gukjenews.com/news/articleView.html?idxno=2264109

²⁾ 비대면 대세 달아오르는 키오스크: https://www.etnews.com/20210816000076

³⁾ 어르신들을 위한 교육용 '키오스크 체험존'운영: https://m.ecomedia.co.kr/news/newsview.php?ncode=1065575239257306

주제 소개

- "서울 지하철역 내에 중소규모 공연예술단체를 지원하는 키오스크를 설치 "
- ✓ 중소규모 공연예술 단체에게 수수료 부담 없는 무인 발권시스템 제공 및 공연 홍보
- ✓ 대부분 서울에서 문화공연이 이루어지므로 서울 지하철역 한정 키오스크 설치 입지 후보로 선정



데이터전처리및탐색

역사명

가락시장

가산디지털단지

강남

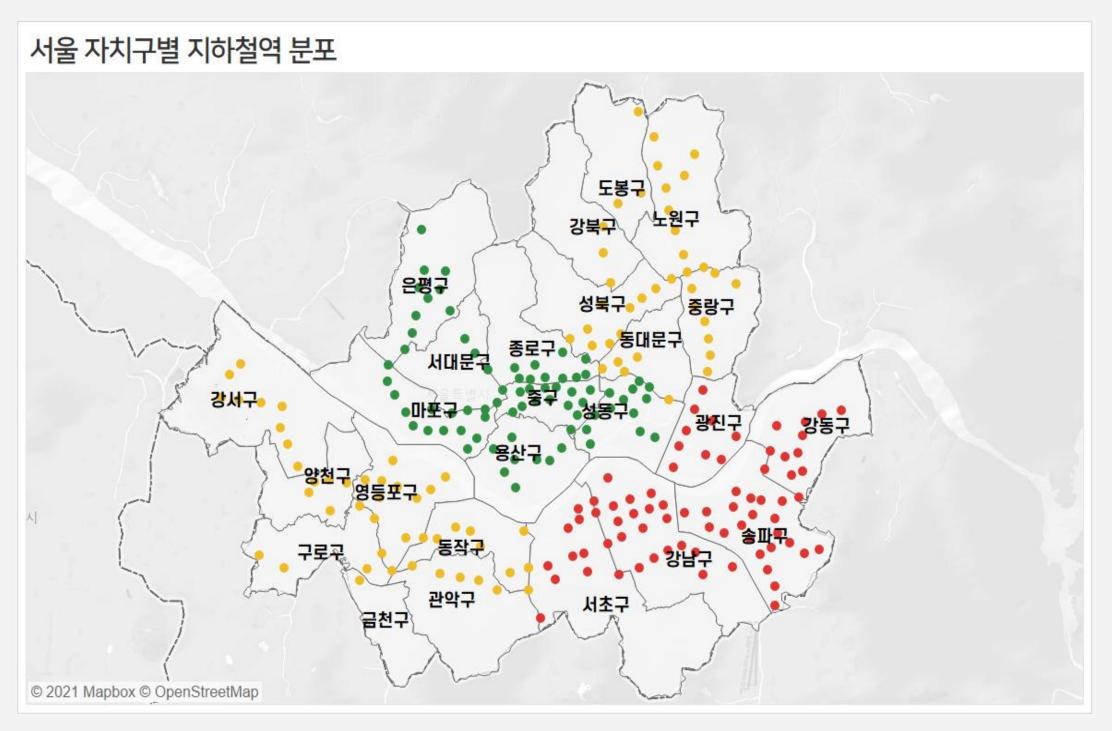
총 273개 역



화랑대

회현

효창공원앞



01 지하철 하차 인원

🖜 유동인구 홍보 지표 💂



하차인원 사용이유

- 승차인원은 키오스크 지나칠 확률 ↑
- 환승이 아닌 하차인원은 실제 귀가경로에 포함
- 저녁시간대는 주로 여가시간대로 활용

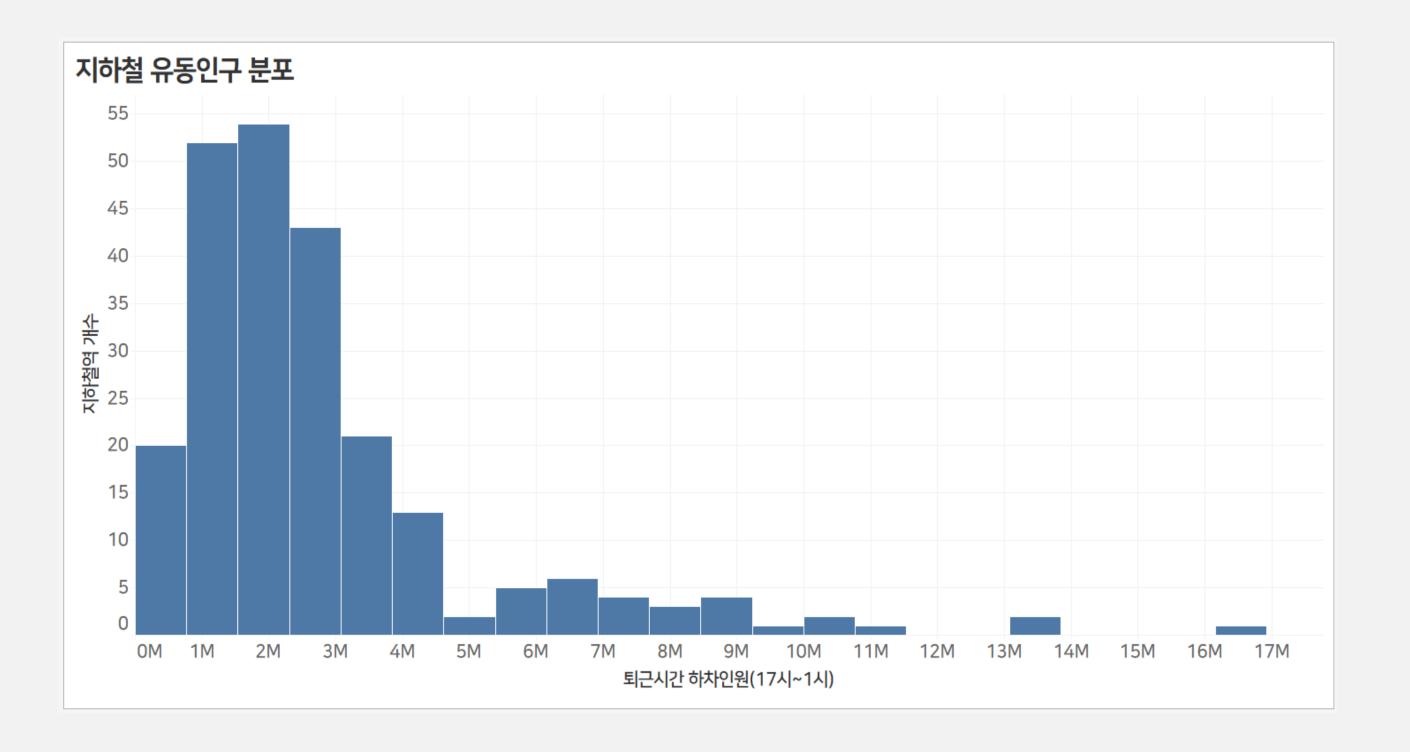
서울 지하철역 시간대별 승하차 인원 정보

서울 지하철역 퇴근 시간 하차인원(17시~1시)

	역사명	퇴근시간 하차인원
0	가락시장	2601550
1	가산디지털단지	1986397
2	강남	8523024
3	강남구청	1074613
4	강동	3522083
•••	***	***
229	홍제	3715896
230	화곡	6908265
231	화랑대	2264901
232	회현	811553
233	효창공원앞	1082067

01 지하철 하차 인원

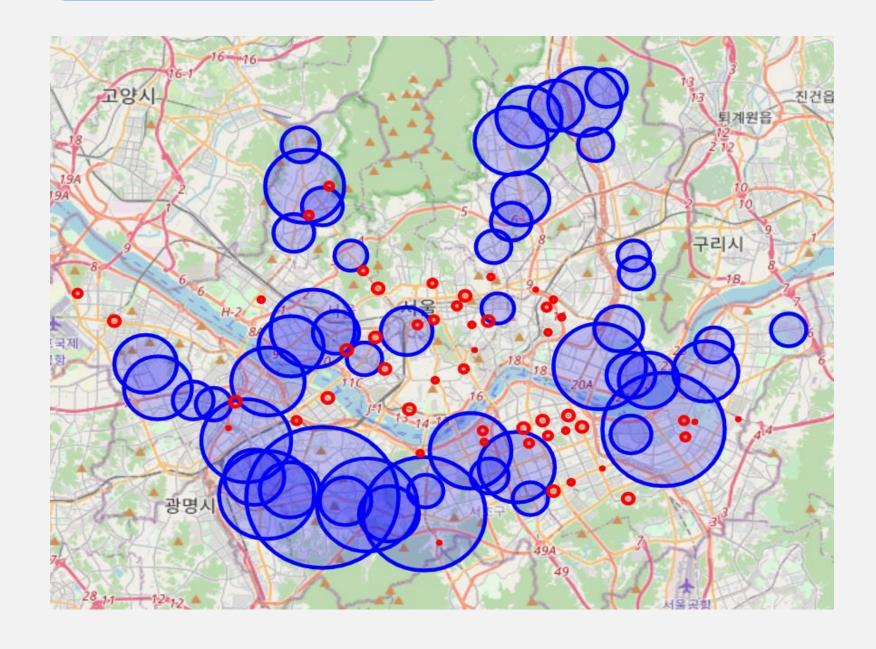
─ 지하철 유동인구 구간별 지하철역 개수



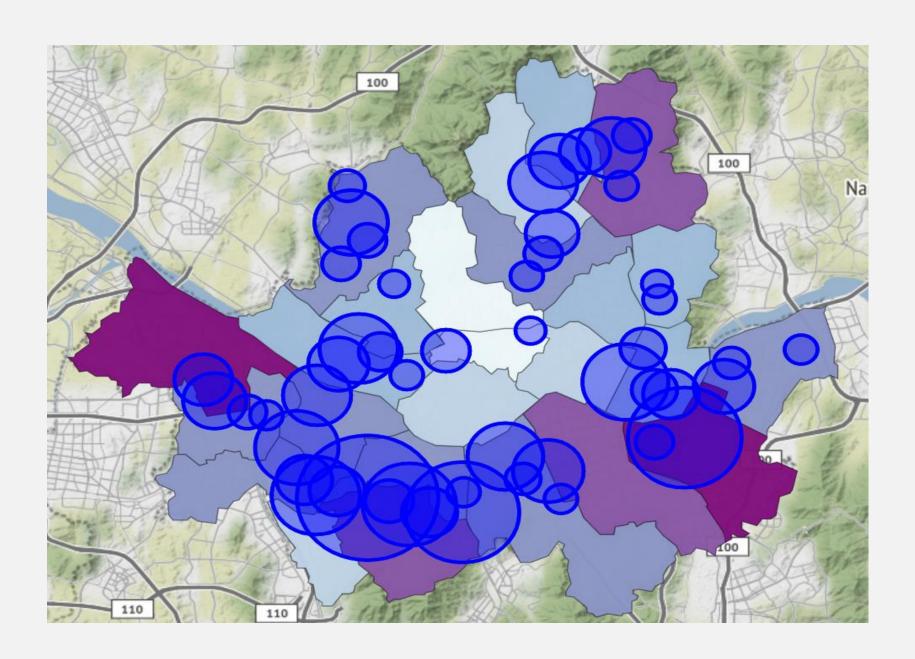
	퇴근시간 하차인원
Count	234.00
Mean	2855032.06
Std	2486569.98
Min	0.00
25%	1284361.25
50%	2090592.50
75%	3296192.50
Max	16652081.00

01 지하철 하차 인원

→ 지하철 하차인원과 실제 거주인원 비교



하차인원 상위 50개 역 + 하위 50개 역



하차인원 상위 50개 역 + 자치구별 인구수 : 강서구, 영등포구, 금천구, 관악구, 구로구

02 공연장수 대비

→ 문화 활성화 지표

총 예매건수 / (총 공연건수 / 총 공연장수)** 3km 이내

Ш

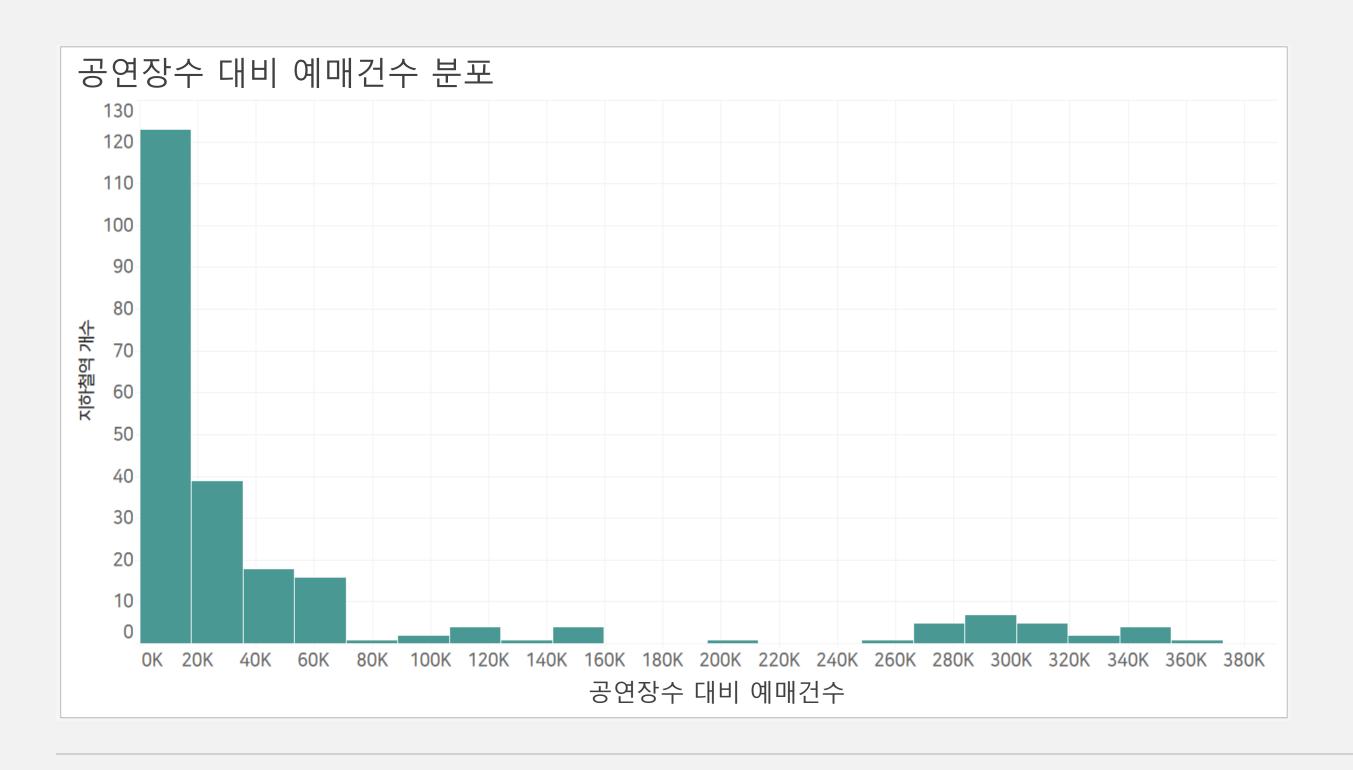
공연장수 대비 예매건수 평균



	역사명	공연장수 대비 예매건수
0	가락시장	6544.167
1	가산디지털단지	2501.423
2	강남	32536.280
3	강남구청	49059.080
4	강동	2966.129
•••	•••	•••
229	홍제	579.562
230	화곡	0.000
231	화랑대	4693.000
232	회현	114347.600
233	효창공원앞	25577.780

02 공연장수 대비

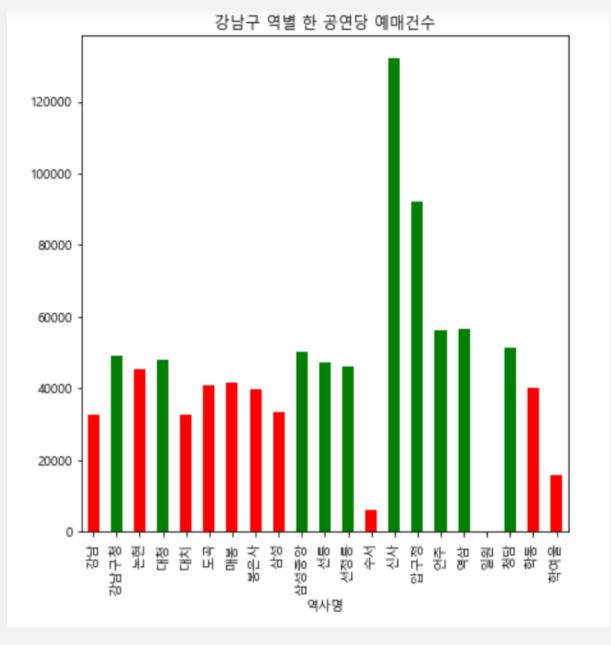
─ 공연장수 대비 예매건수 구간별 지하철역 개수

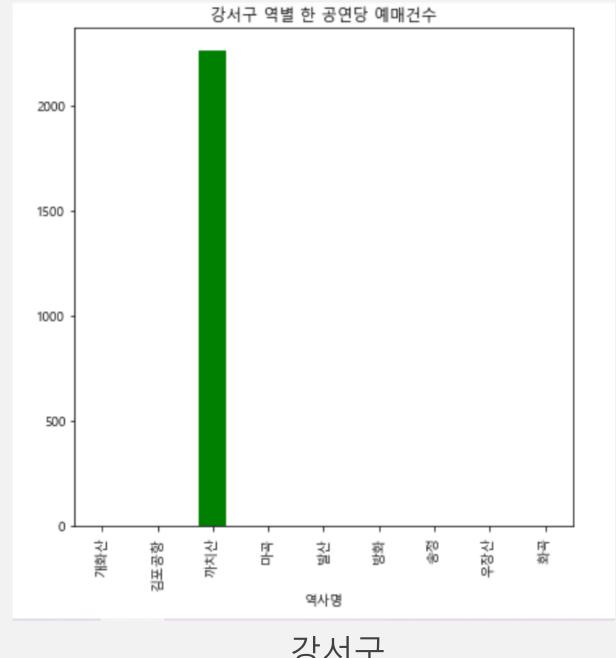


	공연장수 대비 예매건수
Count	234.000000
Mean	54114.946032
Std	92803.960890
Min	0.000000
25%	2451.713000
50%	14962.430000
75%	47092.552500
Max	365279.200000

02 공연장수 대비

→ 자치구별 공연장수 대비 예매건수 비교





25개 자치구별 그래프 시각화 강남구 최하위역 > 강서구 최상위역

> 자치구별 기준 X 전체 역 중에서 상위 역 채택

강남구

강서구

03 신용판매 이용금액

🖜 소비규모 지표 🔙



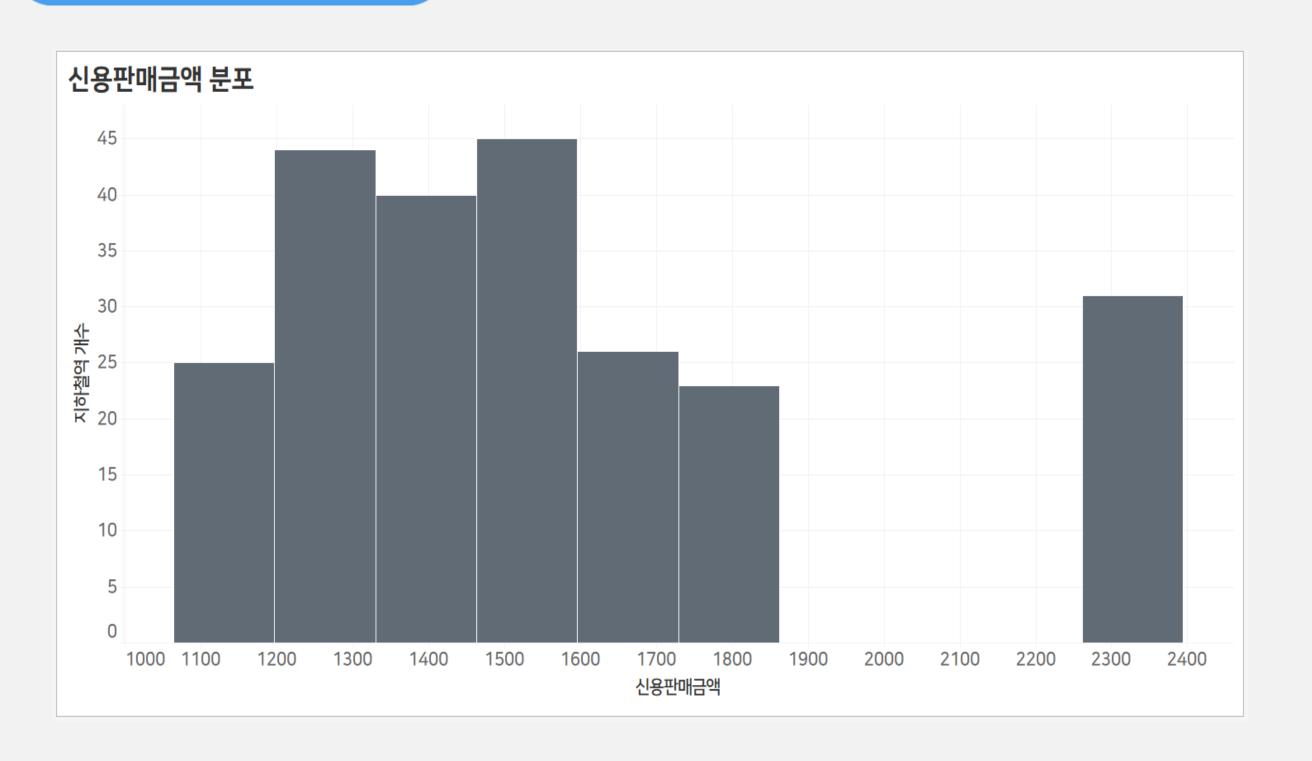
전국 시군구 단위 신용판매금액

서울 지하철역이 속한 자치구의 신용판매금액

	역사명	자치구	신용판매 이용금액
0	가락시장	송파구	1801
1	가산디지털단지	금천구	1092
2	강남	강남구	2326
3	강남구청	강남구	2326
4	강동	강동구	1518
229	홍제	서대문구	1273
230	화곡	강서구	1467
231	화랑대	노원구	1456
232	회현	중구	1196
233	효창공원앞	용산구	1620

03 신용판매 이용금액

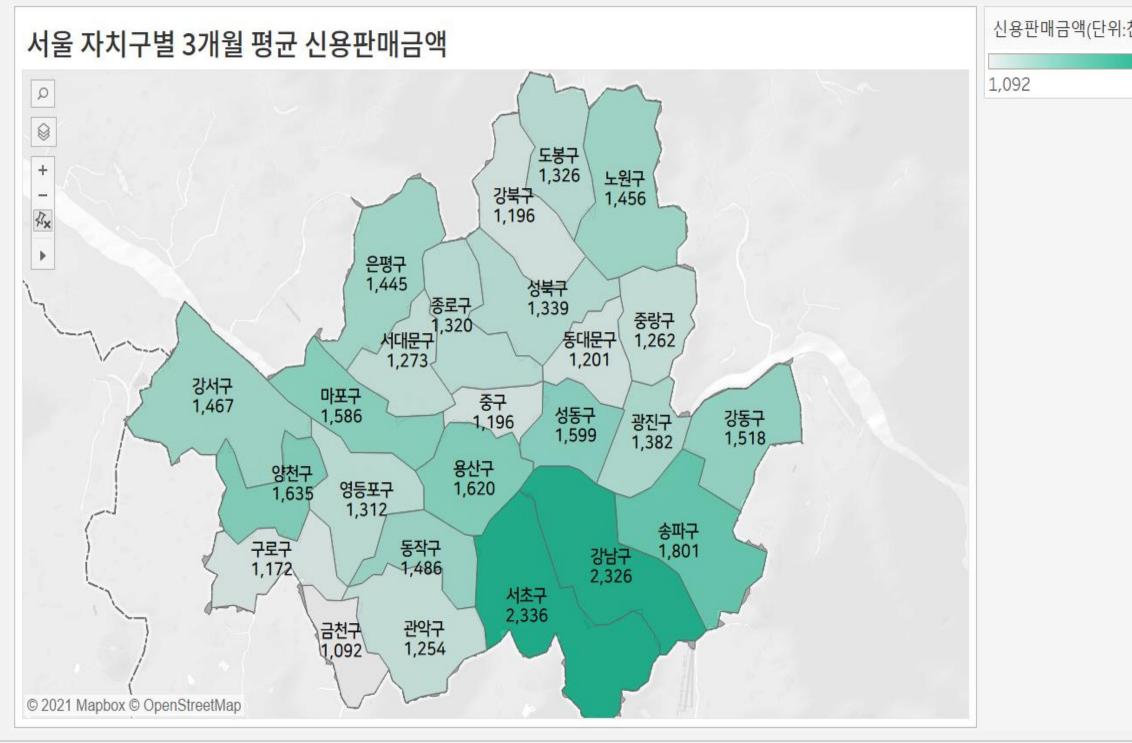
→ 신용판매금액 구간별 지하철역 개수



	신용판매금액
Count	234.000000
Mean	1566.431624
Std	345.654591
Min	1092.000000
25%	1320.000000
50%	1467.000000
75%	1631.250000
Max	2336.000000

03 신용판매 이용금액

서울 자치구별 신용판매금액





- > 신용판매금액 높은 자치구 강남구, 서초구
- > 신용판매금액 낮은 자치구 금천구, 구로구

04 문화예술관람

→ 문화예술 선호도 지표



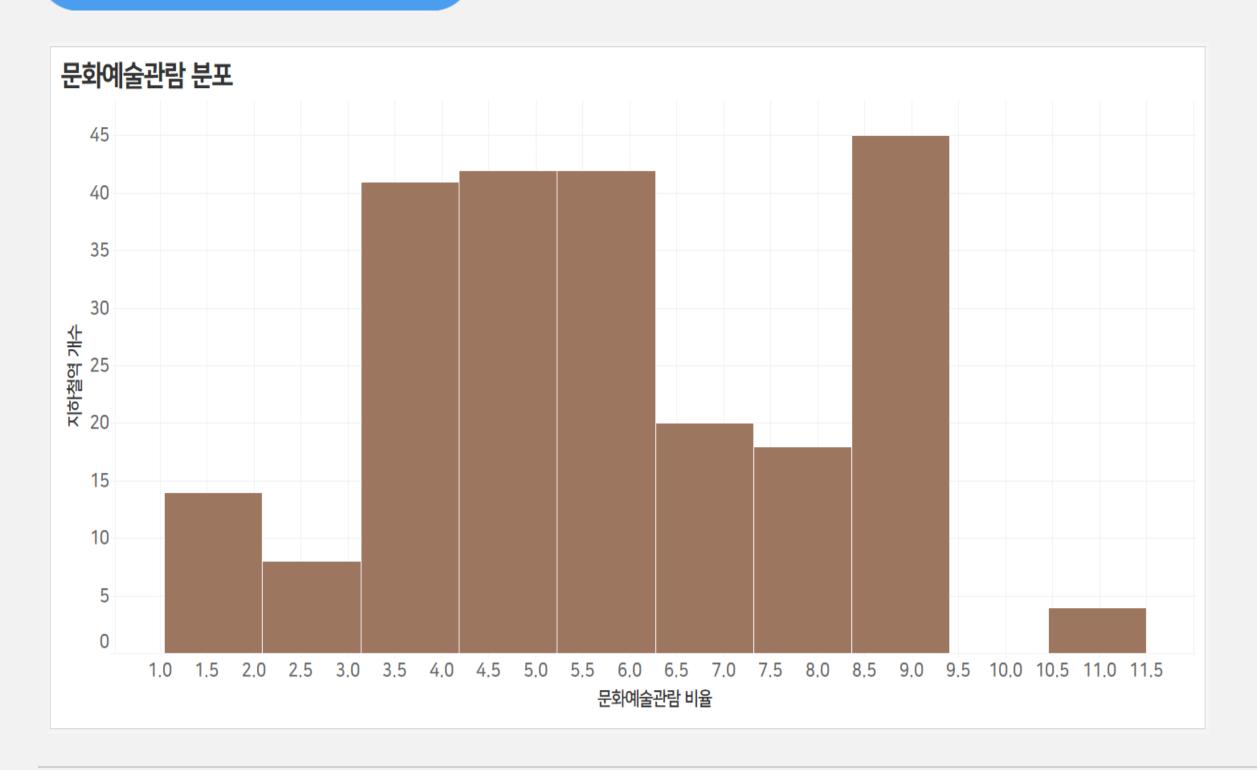
서울시 주말·휴일 여가활동 통계

서울 지하철역이 위치한 자치구의 문화예술관람 비율

	역사명	자치구	문화예술관람
0	가락시장	송파구	4.8
1	가산디지털단지	금천구	7.0
2	강남	강남구	9.0
3	강남구청	강남구	9.0
4	강동	강동구	4.1
•••			•••
229	홍제	서대문구	10.8
230	화곡	강서구	5.3
231	화랑대	노원구	5.3
232	회현	중구	1.6
233	효창공원앞	용산구	2.2

04 문화예술관람

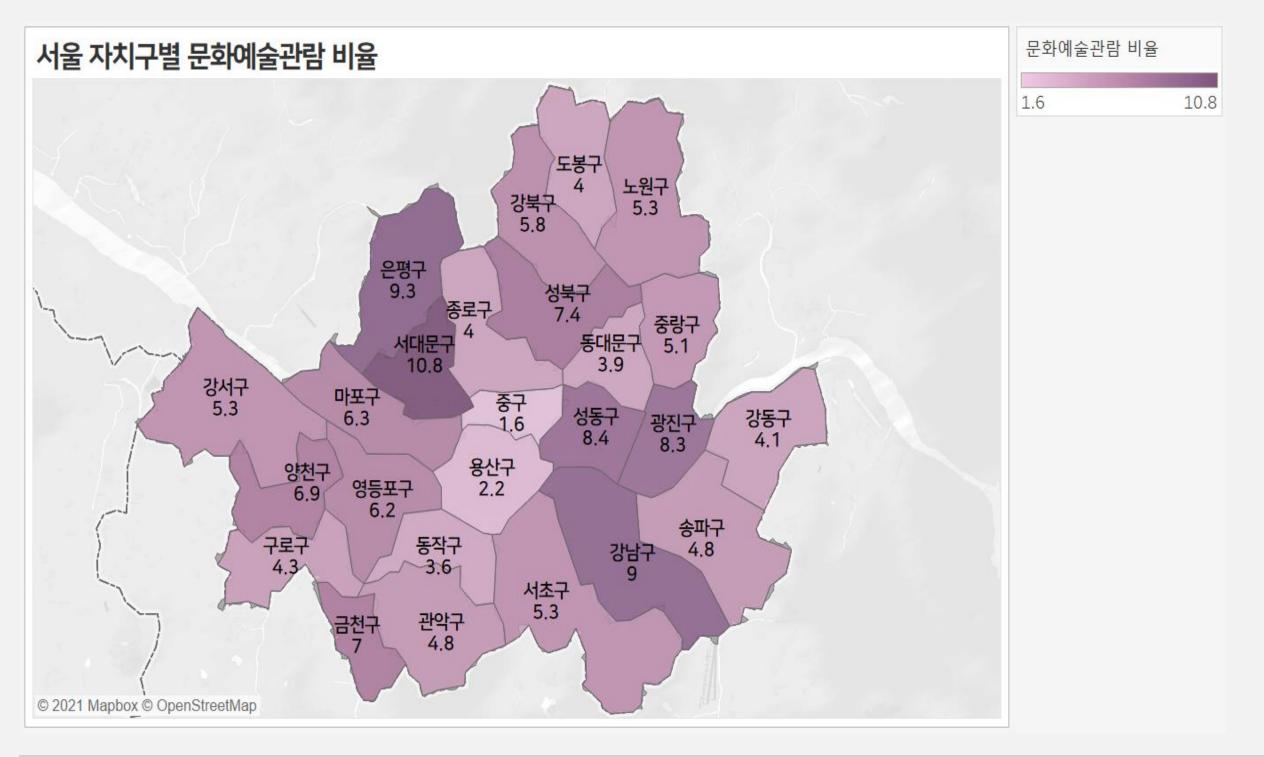
─ 문화예술관람 구간별 지하철역 개수



	문화예술관람
Count	234.000000
Mean	5.758974
Std	2.252558
Min	1.600000
25%	4.100000
50%	5.300000
75%	7.400000
Max	10.800000

04 문화예술관람

─ 서울 자치구별 문화예술관람 비율 시각화



> 문화예술관람 높은 자치구 서대문구, 은평구

> 문화예술관람 낮은 자치구 중구, 용산구

모델링및분석

─ get_score , get_grade 함수

역사명	자치구	퇴근시간 하차인원	신용판매금액	문화예술관람	공연장수 대비 예매건수
가락시장	송파구	2601550	1801	4.8	6544.167
가산디지털단지	금천구	1986397	1092	7	2501.423
강남	강남구	8523024	2326	9	32536.28
강남구청	강남구	1074613	2326	9	49059.08
강동	강동구	3522083	1518	4.1	2966.129
강동구청	강동구	2071184	1518	4.1	10639.86
강변	광진구	6585359	1382	8.3	22995.98
개롱	송파구	1384945	1801	4.8	7281.5
•••	•••	•••			•••
혜화	종로구	3371394	1320	4	299432.2
홍대입구	마포구	9185956	1586	6.3	18941.53
홍제	서대문구	3715896	1273	10.8	579.562
화곡	강서구	6908265	1467	5.3	0
화랑대	노원구	2264901	1456	5.3	4693
회현	중구	811553	1196	1.6	114347.6
효창공원앞	용산구	1082067	1620	2.2	25577.78

- 1. 변수별로 점수를 나눌 기준 값 계산
- 2. 변수별로 정해진 기준 값에 따라 점수 부여
- 3. 최종 점수가 높은 상위 지하철역을 키오스크 설치 후보로 선정

─ get_score , get_grade 함수

```
def get_score(level, data):
    score = []
    for j in range(len(data)):
        for i in range(len(level)):
            if data[j] <= level[i]:</pre>
                 score.append(i+1)
                break
            elif data[j] > max(level):
                 score.append(len(level)+1)
                break
            else:
                 continue
    return score
```

Def:

level 안에 있는 원소를 기준으로 1~len(level)+1까지 점수를 부여하는 함수

Params:

level = 튜플 or 리스트 타입의 숫자형 데이터 반드시 오름차순으로 정렬 예) [1,2,3,4,5] O, [5,4,3,2,1] X, [1,3,2,10,4] X data = 점수를 부여할 데이터

Return:

점수를 담고 있는 리스트 반환

get_score([10,50,100],[5,15,70,120])

[1, 2, 3, 4]

─ get_score , get_grade 함수

def get_grade(df, num_class, tick_point, col_map):

Def:

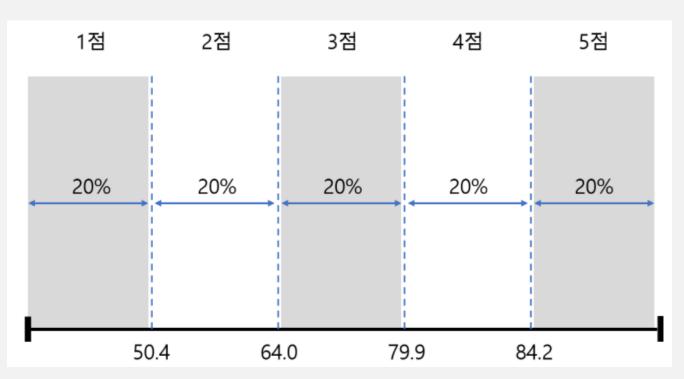
각 데이터에 대한 열의 변수 값이 주어지면 변수별로 점수를 계산하여 주어진 데이터 오른쪽에 추가

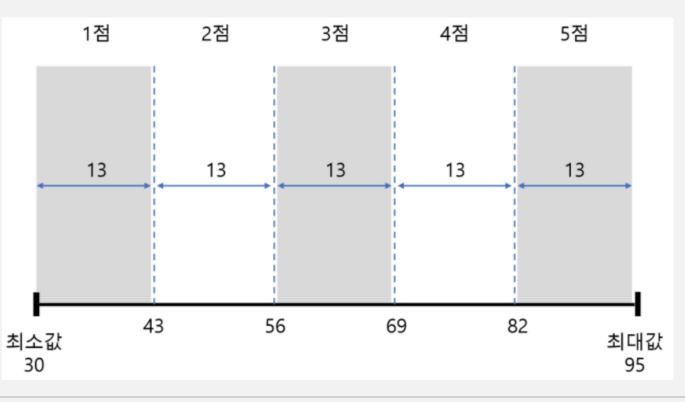
Params:

df = pandas.DataFrame
num_class = 점수 개수
tick_point = 변수별로 점수를 나눌 기준
(분위수, 최대최소, 리스트 직접 입력)
col_map = 각 점수에 해당하는 칼럼명

Return:

pandas.DataFrame





─ get_score , get_grade 함수

col_map={'score_1':'퇴근시간 하차인원','score_2':'공연장수 대비 예매건수', 'score_3':'신용판매금액', 'score_4':'문화예술관람'}

tick_point={'score_1':'quantile', 'score_2':'quantile', 'score_3':'quantile', 'score_4':'quantile'}

num_class= {'score_1':5, 'score_2':5, 'score_3':3, 'score_4':3}

변수별로 정해진 기준에 따라 점수 부여

		공연장수 대비 예매건수	신용판매금액	
0	문화예술관람	2	3	1
1	3	2	1	3
2	5	4	3	3
3	1	4	3	3
4	4	2	2	1
229	4	1	1	3
230	5	1	2	2
231	3	2	2	2
232	1	5	1	1
233	1	4	3	1

- get_score , get_grade 함수

인구점수 = 지하철 하차인원 점수 + 신용판매금액 점수

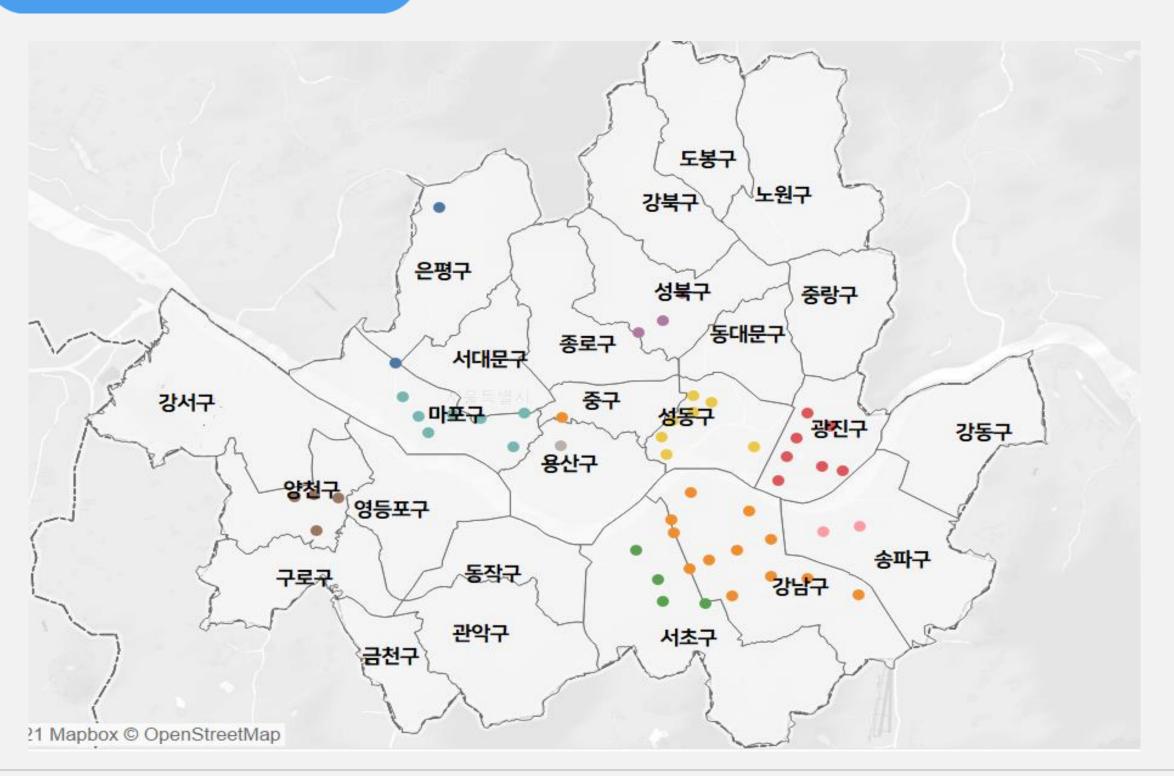
문화점수 = 공연장수 대비 예매건수 점수 + 문화예술관람 점수

최종점수 = (인구점수 + 문화점수)/2 ** 총 8점 만점

-> 최종점수가 6점 이상인 역을 키오스크 설치 상위 후보로 선정

	역사명	문화점수	인구점수	최종점수
0	강남	7	8	7.50
1	신사	8	7	7.50
2	목동	7	8	7.50
3	오목교	7	8	7.50
4	압구정	8	7	7.50
***	***	***	***	***
229	개화산	3	3	3.00
230	신내	4	2	3.00
231	마곡	3	3	3.00
232	여의나루	4	2	3.00
233	온수	2	4	3.00

→ 상위 후보 지하철역 시각화



> 최종점수가 6점 이상인 지하철역은 총 50개로 나타남



K-Means clustering

연령대별 장르 선호도 국악, 무용, 뮤지컬, 연극, 클래식/오페라



인터파크 인기공연 예매율 크롤링 (2019-2021)

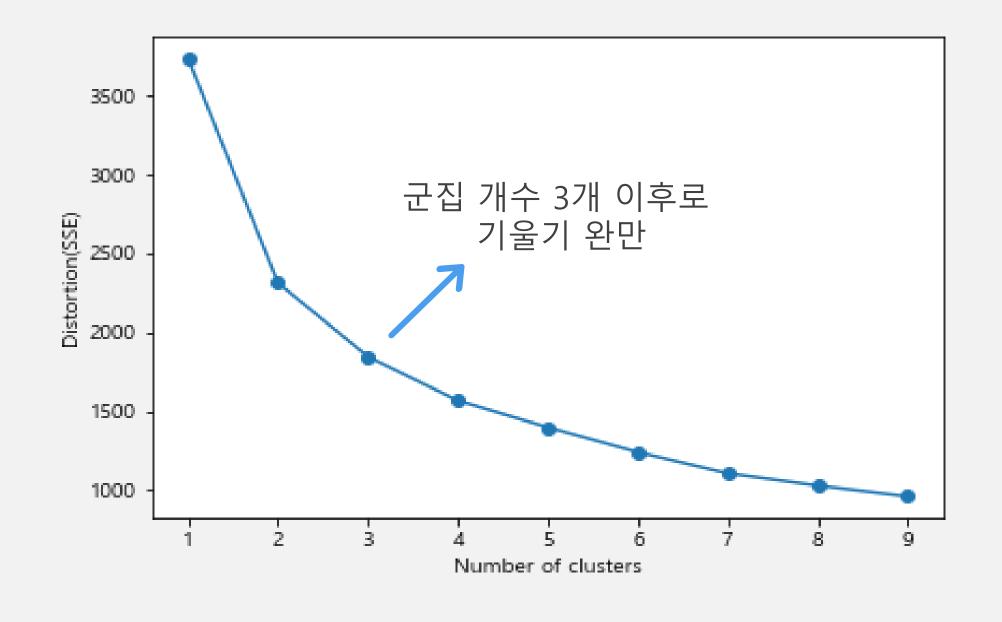
공연명	공연장소	10대	20대	30대	40대	50대	공연년도	공연유형
안성맞춤 남사당공연- 2019	안성맞춤 남사당공연장	2	5.9	35.6	47.5	8.9	2019	국악
음악극〈적로〉	서울돈화문국악 당	4.2	34.8	30.4	18.2	9.2	2019	국악
2019 토요명품공연	국립국악원 우면당	14.5	39.1	12.7	24.9	7.7	2019	국악
•••	•••	•••	•••	•••	•••	•••	•••	•••
베어 더 뮤지컬(bare the musical)	두산아트센터 연강홀	11.5	49.8	28.7	7.5	1.7	2020	뮤지컬
뮤지컬 〈스웨그에이지 : 외쳐, 조선!〉	홍익대 대학로 아트센터 대극장	3.5	38.5	39.8	14.1	3.1	2020	뮤지컬
뮤지컬〈웃는 남자〉	예술의전당 오페라극장	12.5	48.8	20.3	12.8	4.5	2020	뮤지컬
•••		•••			•••	•••	•••	•••
〈포르테 디 콰트로 언플러그드 콘서트: 더 클래식〉 - 서울	롯데 콘서트홀	0.4	14	28.1	37.6	16.1	2021	클래식/오페라

K-Means clustering

- K-Means clustering
 대표적인 비지도학습 군집화 알고리즘
- Elbow method
 다양한 군집 수에 대해 성능 평가 지표(SSE)를 도시하여 최적의 군집 수 선택하는 방법
- Silhouette method
 한 군집 안의 데이터들을 다른 군집과 비교해 유사성을 나타내는 지표

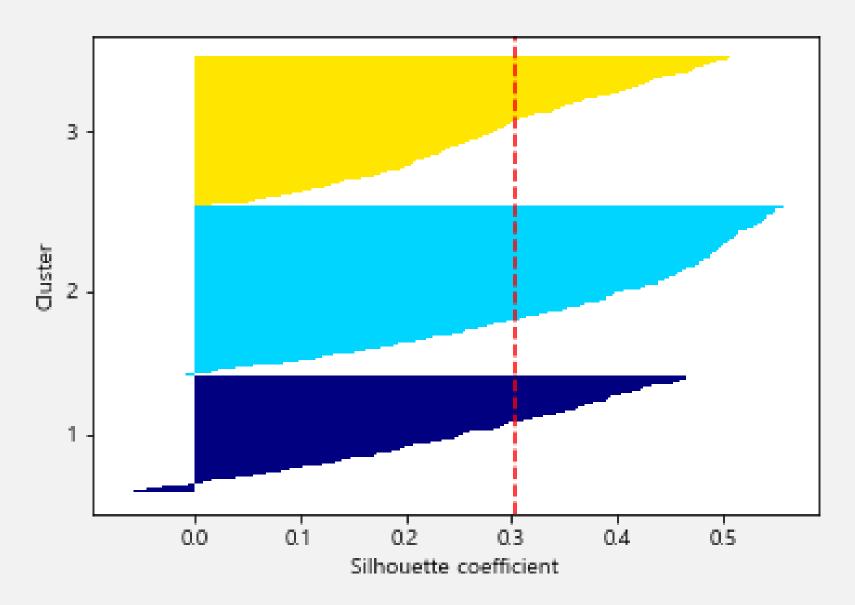
K-Means clustering

Elbow method 군집을 추가로 늘려가면서 군집내 변동성이 급감하는 군집 개수에서 최적 군집 수가 결정되는 경우가 일반적



K-Means clustering

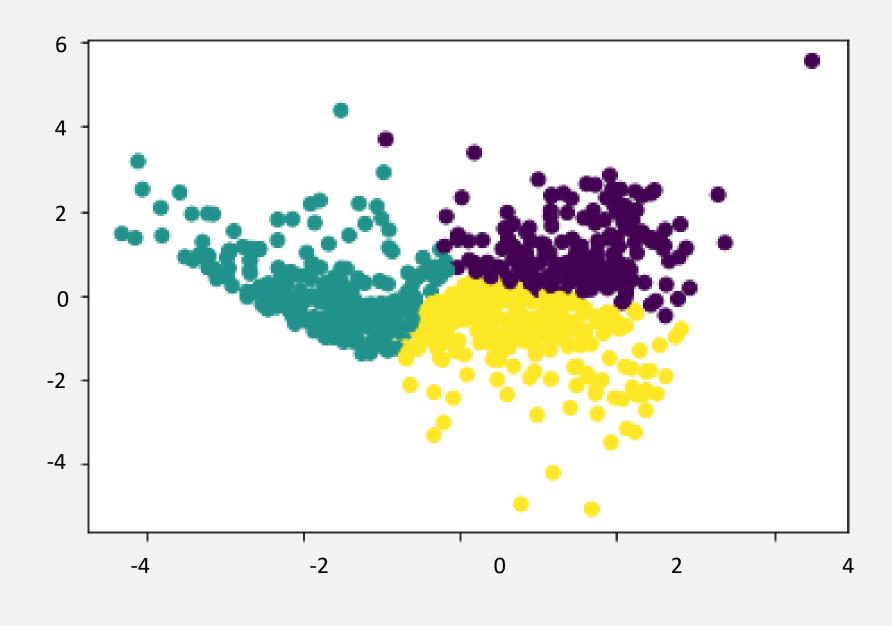
Silhouette method : 실루엣 계수는 -1 부터 1사이의 값을 가지고 1에 가까울수록 클러스터링 결과가 잘 부합함 같은 군집 내 거리가 짧을 수록, 다른 군집과의 거리는 멀수록 좋음



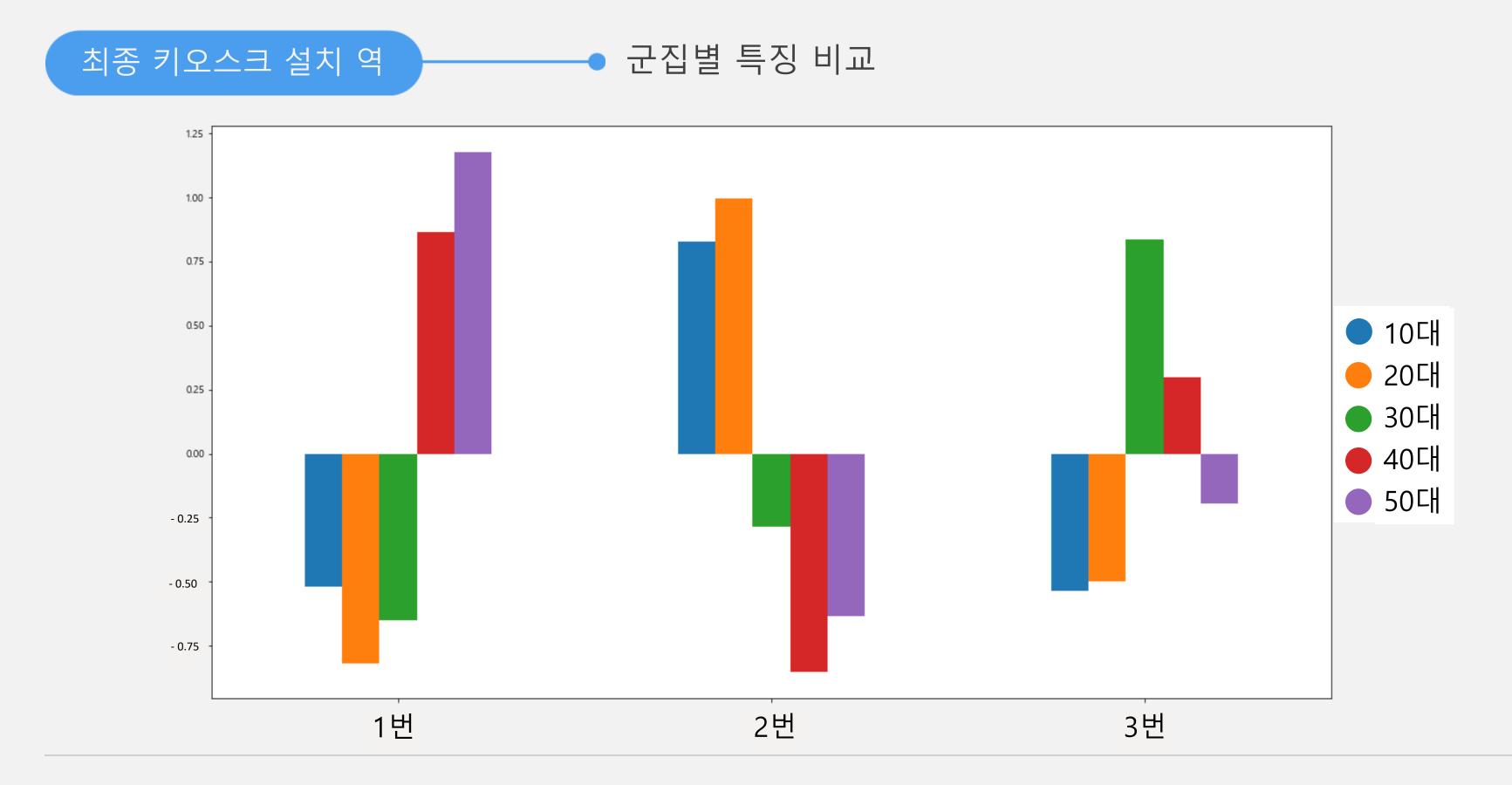
3개의 군집에서 모두 실루엣 계수가 0 이상이고 평균은 0.3으로 나타났으므로 군집화가 잘 이루어졌다고 평가할 수 있음

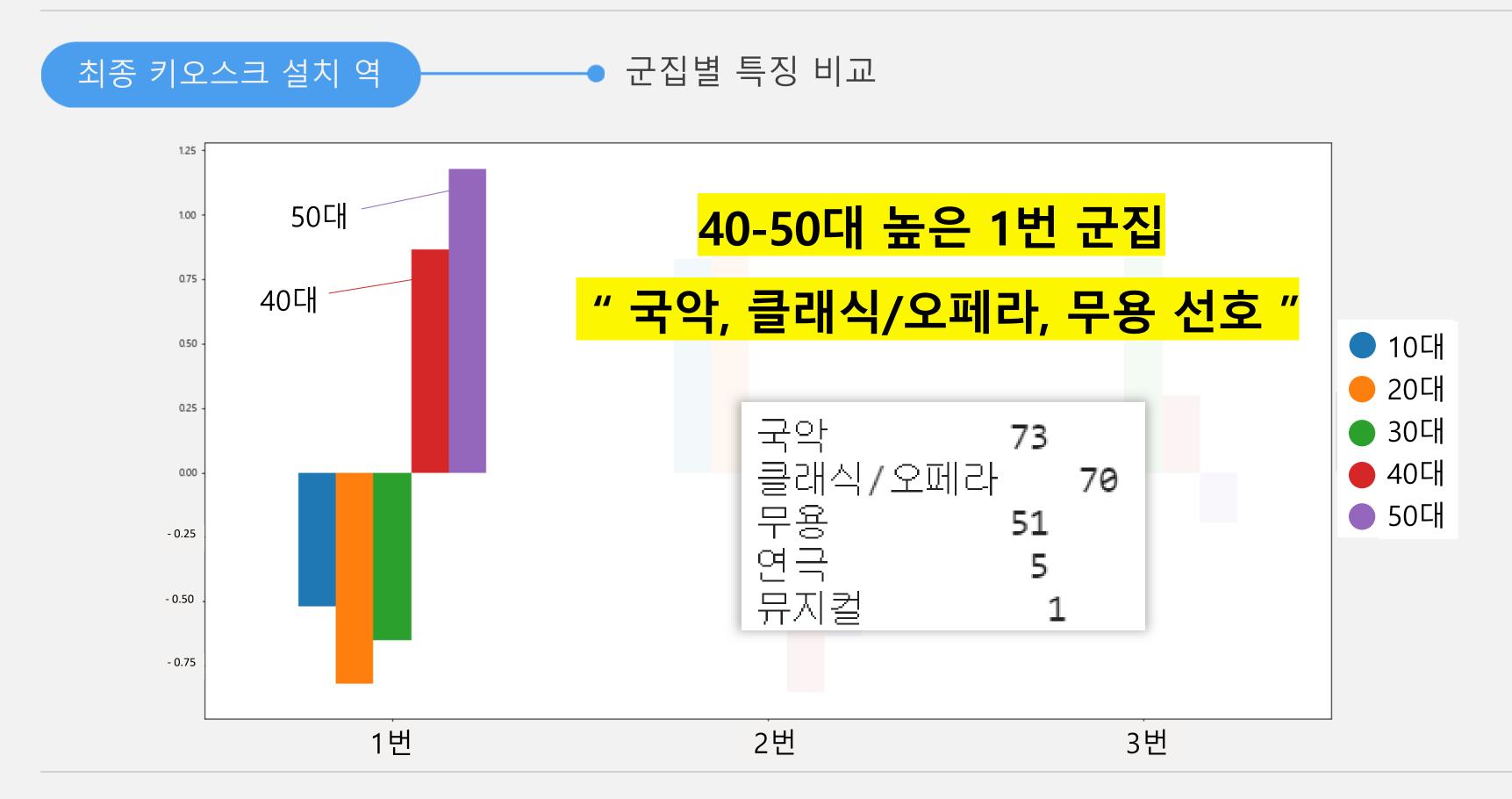
K-Means clustering

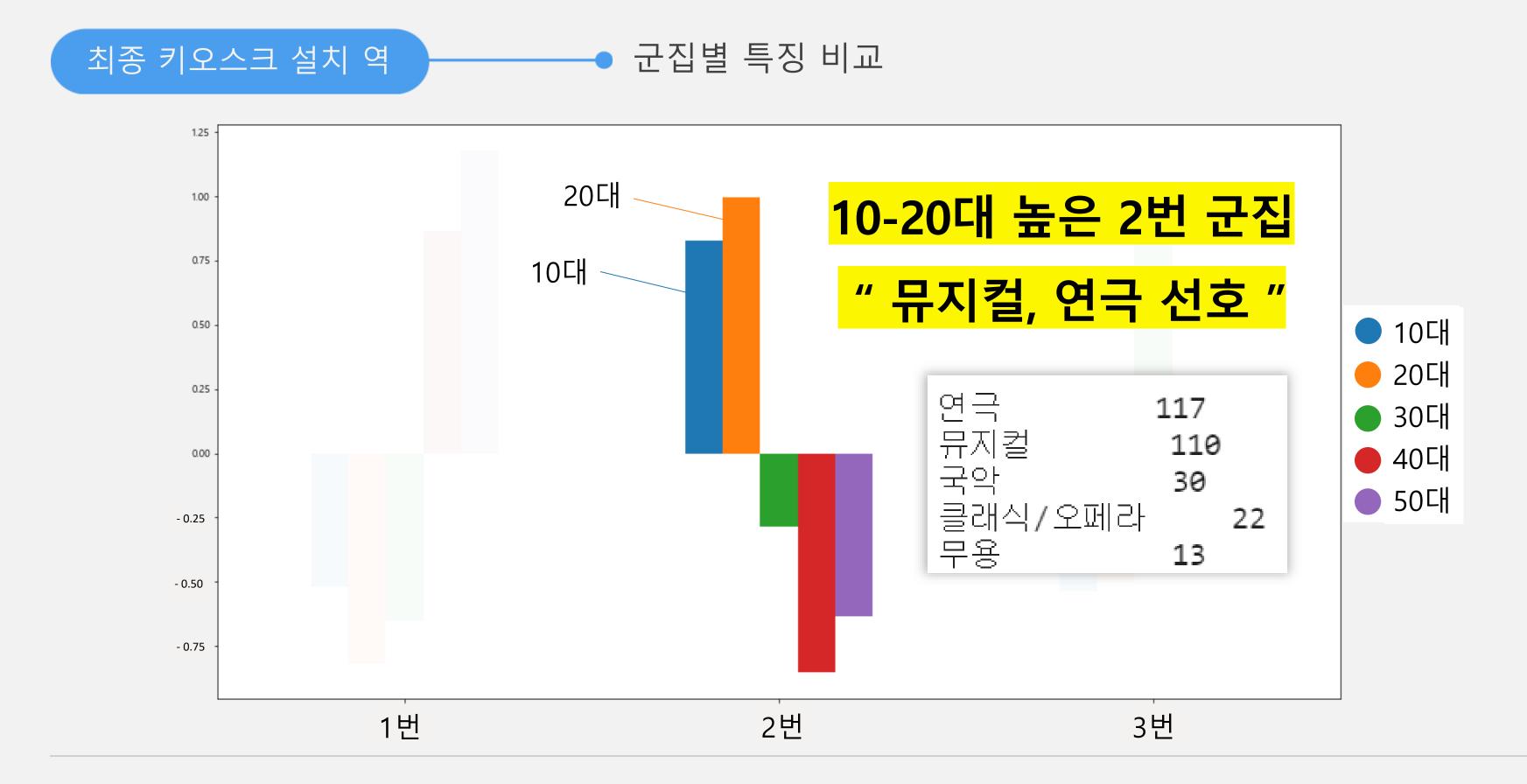
주성분 분석(PCA): 주어진 데이터의 정보를 최대한 유지하면서 차원을 축소하는 기법



군집분석 결과를 차원 축소한 후 2차원 산점도를 그려서 임의의 기준에 따라 군집이 3개로 잘 나누어짐을 확인







1번

3번

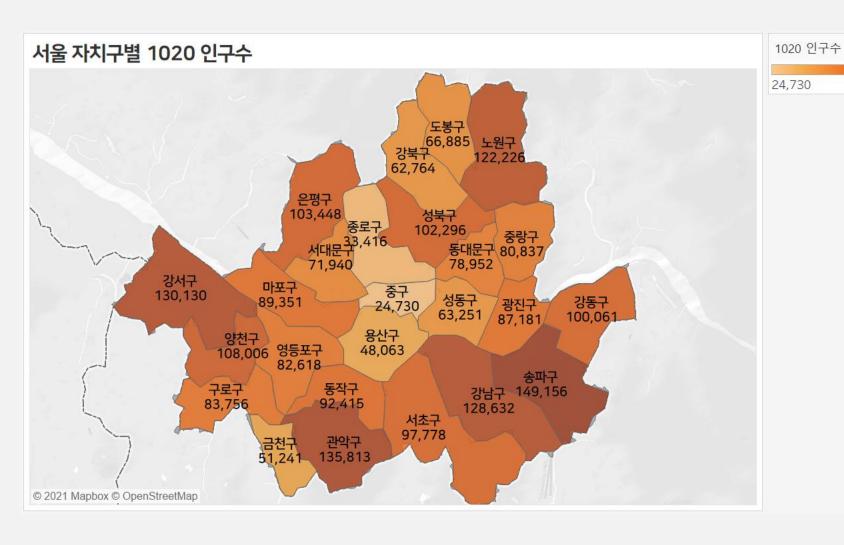
군집별 특징 비교 최종 키오스크 설치 역 1.25 30대 높은 3번 군집 30대 " 모든 공연장르 고르게 선호 " - 10대 _ 20대 86 30대 클래식/오페라 56 40대 0.00 47 50대 D0.25 뮤지컬 39 연극 28 D0.50 D0.75

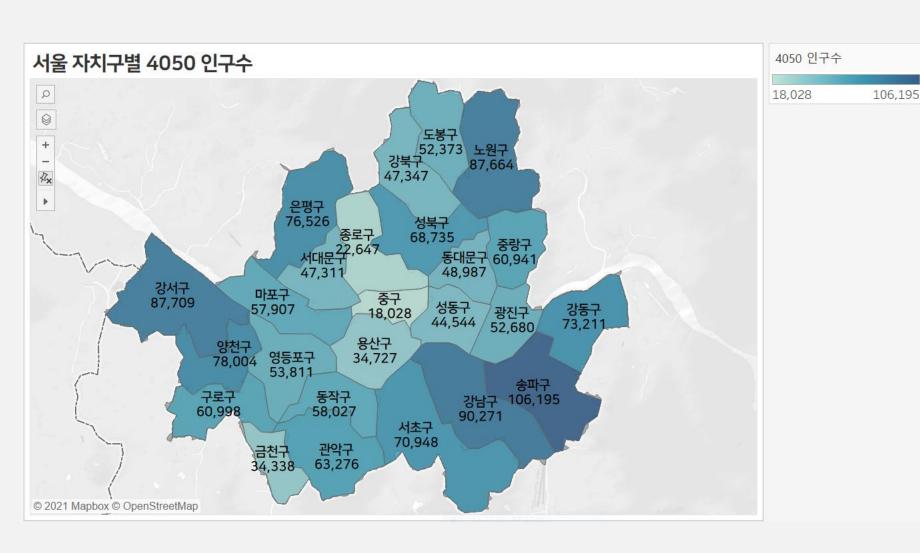
2번

─ 연령대별 인구 분포

149,156

군집분석 결과 특정 장르 선호가 나타난 10-20대와 40-50대 인구분포 시각화





- > 대부분 1020 인구수가 높은 지역과 4050 인구수가 높은 지역이 동일
- > 마포구는 1020 인구수가 평균 이상이지만 4050 인구수는 평균 미만

최종 키오스크 설치 역 제외

1차에서 선정된 50개의 역 중 1020, 4050 인구수 모두 평균 미만인 자치구에 위치한 지하철 역 제외

광진구 - 구의, 건대입구, 군자, 강변, 아차산, 어린이대공원, 뚝섬유원지

성동구 - 왕십리, 상왕십리, 금호, 옥수, 성수, 신금호, 행당

용산구 - 숙대입구

중구 - 서울역

> 총 16개 역 제외

→ 군집분석 결과 활용

1차에서 선정된 50개의 역 중 1020, 4050 인구수 모두 평균 이상인 자치구 (27개 역)



"모든 공연장르" 홍보효과 ↑ 최적 키오스크 설치 역으로 추천

강남구 - 강남, 신사, 압구정, 역삼, 삼성, 선릉, 대청, 논현, 대치, 매봉, 수서, 청담

서초구 - 고속터미널, 양재, 교대, 남부터미널

양천구 - 목동, 오목교, 양천구청, 신정

성북구 - 성신여대입구, 길음, 한성대입구

송파구 - 잠실, 잠실새내

은평구 - 구파발, 디지털미디어시티

→ 군집분석 결과 활용

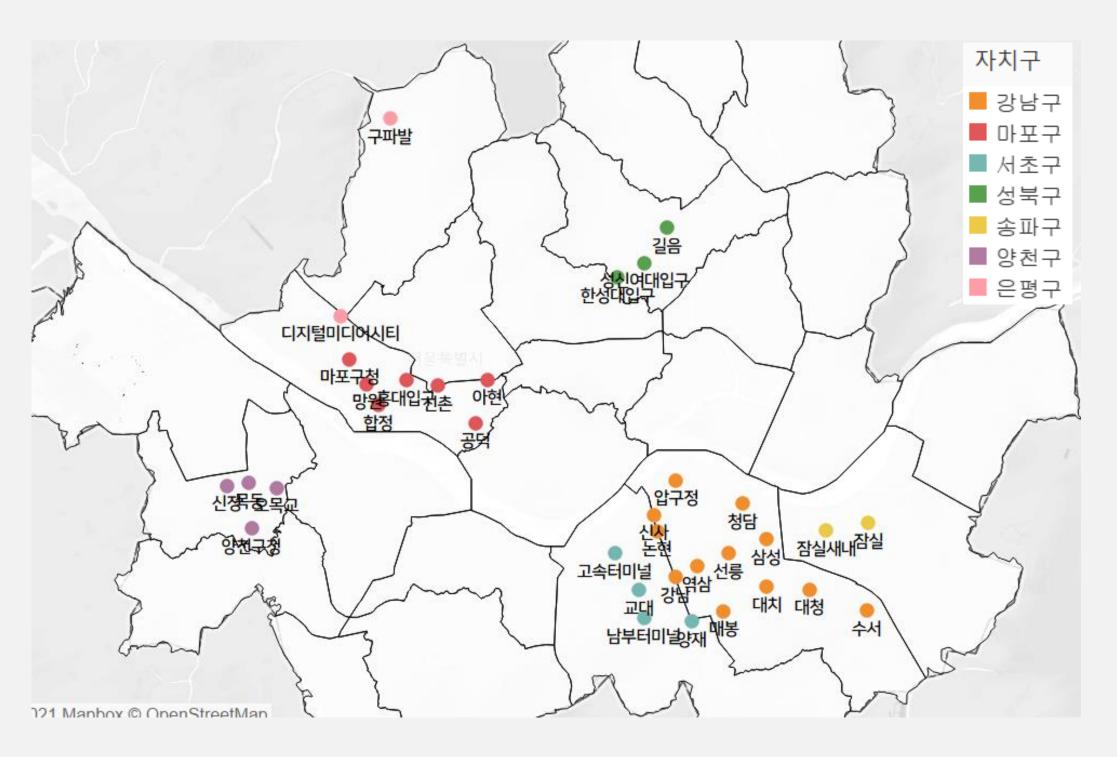
1차에서 선정된 50개의 역 중 1020 인구수만 평균 이상인 자치구 (7개 역)

마포구 - 신촌, 홍대입구, 합정, 아현, 망원, 마포구청, 공덕

키오스크 설치 시 1020 연령대에게 선호도가 높은 "뮤지컬, 연극" 공연 위주의 진출 추천

분석결과 정리

최종 키오스크 설치 역 추천 시각화



> 총 34개역을 키오스크 설치 최적 입지로

최적입지로 선정된 34개역 중에서 키오스크 설치를 우선적으로 실시

강남구, 송파구는 4050 인구수가 9만명 이상 1020 인구수가 12만명 이상이므로 키오스크 이용이 가장 활발하고 공연 홍보효과 높을 것으로 예상

기대효과 및 발전방향

기대효과

중소규모 공연예술의 활성화

: 국내 문화 공연의 발전에 기여

비대면 24시간 무인 발권시스템의 편리성

: 자신이 원하는 시간에 발권 가능

누구나 접근할 수 있다는 지하철역의 특성

: 기존 문화공연에 관심있는 고객 이외에 잠재적 고객 유입 예상

키오스크에 공연 홍보영상 송출

: 중소규모 단체의 홍보 부족 문제 보완



발전방향

키오스크 이용률을 높이기 위한 혜택 방안 제시

: 키오스크에서 발권 시 포토 티켓 제공, 할인 이벤트 등

키오스크 설치 후 예매현황 데이터 활용

: 설치 전후 효과 비교 분석

서울 뿐만 아니라 타 지역으로의 확대 방향 고안

: 서울과 달리 지하철 역이 적은 지역은 다른 설치 기준 설정



THANK YOU!

중소규모 공연예술단체 활성화를 위한 키오스크 입지 선정

주요 데이터

https://www.culture.go.kr/bigdata/user/data_market/detail.do?id=6bb822b0-4a54-11eb-af9a-4b03f0a582d6 https://www.culture.go.kr/bigdata/user/data_market/detail.do?id=b316ae32-244c-45d1-8a6a-4b283f6b7cb9 http://data.seoul.go.kr/dataList/OA-12914/S/1/datasetView.do https://data.seoul.go.kr/dataList/10260/S/2/datasetView.do https://www.bigdata-culture.kr/bigdata/user/data_market/detail.do?id=78c60e99-06c8-435d-9d2f-06469adbc6b0 https://data.seoul.go.kr/dataList/419/S/2/datasetView.do

참고 자료

https://zephyrus1111.tistory.com/13?category=820621 https://stricky.tistory.com/283 https://velog.io/@sset2323/07-02.-%EA%B5%B0%EC%A7%91-%ED%8F%89%EA%B0%80Cluster-Evaluation https://sean-parkk.github.io/study/Study_statistics_7/#k-%ED%8F%89%EA%B7%A0-%ED%81%B4%EB%9F%AC%EC%8A%A4%ED%84%B0%EB%A7%81 https://dailyheumsi.tistory.com/144

이미지

https://www.pexels.com/ko-kr/https://unsplash.com/