

「2023년 제11회 문화데이터 활용 경진대회」

분석보고서

분석 프로그램	<input checked="" type="checkbox"/> Python	<input type="checkbox"/> R	<input type="checkbox"/> Tableau	<input type="checkbox"/> 기타
---------	--	----------------------------	----------------------------------	-----------------------------

1) 분석 주제

※ 자유주제(문화데이터 분석을 통한 정책 현안 해결 또는 분석 인사이트 도출)

저희의 분석 주제는 "서울시 자치구별 장애인 삶의 질 향상을 위한 접근성과 시설 개선"입니다. 장애인의 삶의 질 향상을 위해서는 생활체육 활동 등을 통해 건강을 유지하고 문화 생활과 커뮤니티 활동을 통해 정신적 스트레스를 해소할 수 있어야 합니다. 또한, 이런 건강 생활 여건과 문화 커뮤니티 활동에 참여하기 위한 접근성 여건이 마련되어야 합니다. 저희는 장애인 삶의 질 현황을 나타내고, 정책 혹은 장애인 활동에 활용할 수 있게 하기 위해 아래 네 가지 지표를 산출하는 데이터 분석을 수행하였습니다.

1. 장애인 건강 생활지수: 이 지수는 장애인들의 건강 생활에 대한 접근성을 나타냅니다. 이를 측정하기 위해 생활체육 강좌의 수, 스포츠 동호회의 수, 그리고 복지시설의 수를 고려하였습니다.
2. 장애인 문화 커뮤니티 지수: 이 지수는 장애인들의 문화 활동 및 커뮤니티 참여에 대한 접근성을 나타냅니다. 이를 평가하기 위해 내부시설, 외부시설, 그리고 장애인 협회의 존재 여부를 고려하였습니다.
3. 신체적 장애인 이동성 지수 : 이 지수는 장애인들이 지하철을 이용할 때의 편의성을 나타냅니다. 이를 판단하기 위해 지하철 엘리베이터의 수, 전동보장구 급속 충전소의 유무, 그리고 지하철 휠체어 리프트의 수를 고려하였습니다.
4. 장애인 편의시설 설치 완료도: 이 지수는 장애인들이 일상 생활에서 이용할 수 있는 편의시설의 설치 정도를 나타냅니다. 이를 평가하기 위해 매개시설, 내부시설, 위생시설, 그리고 안내시설의 설치 완료도를 고려하였습니다.

위의 네 가지 지표를 활용하여 서울시의 자치구별로 각각의 지수를 계산하고, 이를 지도 상에 시각화하였습니다. 이를 통해 각 자치구의 장애인 삶의 질 향상을 위한 상황을 파악하고 인사이트를 도출할 수 있었습니다.

이와 같은 분석 방법을 통해 저희는 장애인들의 건강 생활, 문화 커뮤니티 접근성, 지하철 시설, 그리고 편의시설의 개선을 위한 자치구별 차이를 파악할 수 있었습니다. 이를 통해 서울시는 장애인 삶의 질 향상을 위한 정책과 개선 방향을 더욱 효과적으로 수립할 수 있을 것으로 기대합니다.

2) 분석의 배경 및 목적

※ 분석 배경(관련 현황, 문제점) 및 목적, 필요성 작성

장애인들의 집 밖 활동의 불편성에 대한 조사 결과에 따르면, 전체 응답자 중 약 46.6%가 집 밖 활동이 불편하다고 응답했습니다. 이는 전체의 절반에 가까운 비율로, 장애인들이 집 밖 활동에 어려움을 겪고 있다는 점을 보여줍니다. 그리고 집 밖 활동이 불편한 이유로는 장애 관련 편의시설 부족이 49.7%로 가장 높았습니다. 이를 통해 불편한 환경과 제한된 접근성으로 인해 집 밖 활동에 대한 불편함을 느끼는 장애인들이 많다는 것을 알 수 있습니다.

또한, 장애인들의 생활체육 참여율 역시 부족한 수준임을 문화체육관광부와 대한장애인체육회의 조사 결과에서 확인할 수 있었습니다. 조사 결과에 따르면, 주 2회 이상 집 밖에서 운동하는 장애인은 전체 응답자 중 26.6%에 불과하며, 생활체육 참여율은 60.8%로 나타났습니다. 이는 장애인 중 3명 중 2명 이상이 바깥에서의 운동을 하지 않는다는 것을 의미합니다.

이러한 배경 속에서 저희의 분석 목적은 서울시 자치구 별로 장애인들의 삶의 질 향상을 위한 접근성과 시설 개선에 대한 분석을 수행하는 것입니다. 이를 통해 서울시는 장애인들의 다양한 요구와 어려움을 파악하고, 보다 정확하고 효과적인 정책을 개발할 수 있습니다. 또한, 문제점이나 개선이 필요한 분야를 식별하여 그에 맞는 정책 방향성을 제시함으로써, 장애인들의 삶의 질 향상을 도모할 수 있습니다.

장애인들의 삶의 질 향상은 사회적 포용과 인권 보호의 중요한 요소입니다. 장애인들은 다양한 장애로 인해 일상생활에서 어려움을 겪을 수 있으며, 이로 인해 사회적 차별이나 접근성 문제 등이 발생할 수 있습니다. 이에 서울시는 장애인들의 권익을 보호하고 생활 환경을 개선하기 위해 다양한 정책을 추진하고 있습니다. 그러나 각 자치구별로 장애인 삶의 질에 대한 차이가 있을 수 있으며, 이를 파악하여 보완하는 것이 필요합니다. 따라서, 저희는 이러한 배경과 목적을 기반으로 분석을 수행하여 보다 포용적이고 향상된 장애인 삶의 질을 위한 정책 및 개선 방향을 제시하고자 합니다.

참고문헌 :

“장애 체육시설 부족의 아쉬움”, 한국기자협회(http://m.journalist.or.kr/m_article.html?no=53001)

“장애인의 사회활동 및 문화·여가활동 실태와 정책과제”, 보건복지포럼 2018년 9월 통권 제263호, 이민경

3) 활용 데이터 선정

※ 분석에 활용한 문화데이터(제공기관명, 데이터명 포함) 및 공공·민간 데이터 목록, 선정이

유

본 연구에서 활용한 데이터는 크게 두 가지 성격으로 구분됩니다. 먼저, 장애인을 위한 문화 커뮤니티 및 체육 관련 데이터입니다. 이에 해당하는 데이터는 다음과 같습니다. 괄호는 제공기관 명입니다.

- 장애인(베리어프리) 실내,실외 문화생활정보 (한국문화정보원)
- 장애인 및 문화예술 단체 정보 (한국문화예술위원회)
- 장애인 생활체육 동호인 클럽 조회 (국민체육진흥공단)
- 장애인 생활체육교실 데이터 (국민체육진흥공단)
- 서울시 사회복지시설 목록 (서울시 복지정책실)

장애인이 지역 사회 내에서 원활하게 건강 생활을 이어가고, 문화를 향유할 수 있는 조건이 갖추어져 있는지 파악하기 위해 위와 같은 데이터를 활용하였습니다. 이에 더하여, 본 데이터 분석에서는 장애인의 접근성 및 시설 이용 편의성과 관련된 데이터를 선정하였습니다. 실질적으로 장애인이 생활 체육과 문화 생활을 누리기 위해서는 접근성이 충분히 보장되어야 하며 시설이 장애인을 고려한 시설이 확충되어 있어야 합니다. 따라서 실질적인 문화 및 체육 접근성을 종합적으로 고려하기 위해 아래와 같은 장애인 인프라 및 시설 데이터를 선정했습니다.

- 서울교통공사 엘리베이터 설치 정보 (서울교통공단)
- 서울교통공사 자치구별 지하철역 정보 (서울교통공단)
- 서울교통공사 휠체어리프트 설치현황 (서울교통공단)
- 서울시 전동보장구 급속충전소 (서울시 복지정책실)
- 서울시 장애인 편의시설 정보 (서울시 복지정책실)

추가로, 장애인 1인당, 즉 실질적으로 장애인 개개인이 권리를 제공받을 수 있을 만큼 여건이 갖추어져 있는지 파악하기 위해 자치구별 장애인 현황 데이터를 선정했습니다.

- 서울시 장애유형별 등급별 장애인 현황 (서울시)

4) 분석 내용 및 결과

1. 데이터 전처리 과정과 변수 설정

1-1. 신체적 장애인 대중교통 접근성 관련 변수 추출

서울시 자치구별 지하철역 정보는 자치구와 지하철의 역 리스트로 이루어져 있습니다. 본 데이터 분석에서는 지하철역과 자치구를 매핑하여 자치구별 접근성 지표를 산출하기 위한 변수를 만드는 과정에서 자치구별 지하철역 정보를 활용합니다. 따라서 <서울시 자치구별 지하철역 정보>를 ‘자치구’ 및 ‘역’ Column 형태의 데이터프레임으로 정리한 후 <서울교통공사 휠체어리프트 설치현황> 및 <서울교통공사 엘리베이터 설치 정보> 데이터를 병합하였습니다.

병합한 데이터프레임으로 자치구별 휠체어리프트 설치 및 엘리베이터 설치 현황 변수를 도출하기 위해, Python의 group_by() 함수로 데이터를 자치구별로 묶고, count() 함수를 사용하였습니다. <서울시 전동보장구 급속충전소>도 마찬가지로의 방식으로 group_by() 함수로 데이터를 자치구별로 묶고, count() 함수를 사용하여 자치구별 설치 현황 변수를 산출하였습니다.

1-2. 장애인 건강 생활 및 문화 커뮤니티 관련 변수 추출

본 데이터 분석에서 사용하는 장애인 건강 생활 및 문화 생활 관련 데이터를 모두 '자치구' 열을 포함하고 있습니다. 따라서 <장애인 생활체육교실 데이터> <장애인 실내 문화생활정보> <장애인 실외 문화생활정보> <장애인 생활체육 동호인 클럽> <장애인 사회복지시설> 데이터는 1-1과 같은 방식을 사용하여, `group_by()` 함수로 자치구별로 묶고 `count()`로 자치구별 현황을 산출하였습니다.

하지만 <장애인 및 문화예술 단체 정보>는 문화예술단체와 장애인 관련 단체가 함께 있어서 나머지 데이터와 달리 row를 선별하는 작업을 거쳤습니다. `str.contains()` 함수를 활용해 단체 이름에 '장애인'이 포함된 단체의 row를 추출했고, 이후 `group_by()` 함수로 자치구별로 묶고 `count()`로 자치구별 현황을 산출하는 과정을 거쳤습니다.

최종적으로 1-1과 1-2에서 처리한 데이터와 자치구별로 나누어진 <서울시 장애유형별 등급별 장애인 현황> <장애인 편의시설 정보>를 병합하여 데이터 분석에 활용할 Date Set을 완성했습니다.

```
merge_2_3 = pd.merge(sport_commun , merge_2_2, on = '자치구', how = 'outer')
✓ 0.0s

merge_2_3.fillna(0, inplace=True)
✓ 0.0s
```

1-3. 변수 생성 및 표준화를 통한 스케일 편향 조절

1-2까지 과정을 통해 만든 Data Set은 절대 수치를 반영하고 있습니다. 하지만 실질적인 장애인의 건강 생활, 문화 생활, 접근성을 파악하기 위해서는 장애인의 인구 대비 얼마나 잘 이뤄지고 있는지가 필요합니다. 따라서 본 분석에서는 우선 데이터 전처리 과정에서 Data Set의 변수를 일정한 비율로 바꾸는 작업을 진행했습니다.

먼저, <장애인 생활체육교실 데이터> <장애인 실내 문화생활정보> <장애인 실외 문화생활정보> <장애인 생활체육 동호인 클럽> <장애인 사회복지시설> <장애인 단체> 등을 “자치구별 장애인 인구수”로 나누어 장애인 1인당 생활 체육 및 문화 커뮤니티 인프라를 잘 누리고 있는 파악할 수 있는 변수를 만들었습니다. 다음으로, 지하철 <휠체어리프트 설치현황> <엘리베이터 설치 현황> <전동보장구 급속충전소>는 신체적 장애인의 접근성과 관련이 있으므로 <서울시 장애 유형별 장애인 현황>의 자치구별 ‘지체장애’ 인구수로 나누어 변수를 도출했습니다. <장애인 편의시설 정보>에는 설치의무대상 시설의 설치완료 비율이 포함되어 있기 때문에 해당 변수를 그대로 활용했습니다.

이렇게 정리된 변수는 다음과 같습니다.

- 장애인 1인당 생활체육교실
 - 장애인 1인당 실내 문화공간
-

- 장애인 1인당 실외 문화공간
- 장애인 1인당 사회복지시설수
- 장애인 1인당 장애인 단체 수
- 장애인 1인당 생활체육 동호인 클럽 수
- 지체 장애인 1인당 지하철 휠체어리프트 수
- 지체 장애인 1인당 지하철 엘리베이터 설치대수
- 지체 장애인 1인당 전동보장구 급속충전소 수
- 장애인 편의시설 중 매개시설(장애인 통행 가능 접근로, 주차구역 등) 설치완료 비율
- 장애인 편의시설 중 위생시설(화장실, 욕실 등) 설치완료 비율
- 장애인 편의시설 중 안내시설(점자블록, 유도 및 안내설비 등) 설치완료 비율
- 장애인 편의시설 중 내부시설(복도, 승강기 등) 설치완료 비율

위의 변수는 각각 다른 특성에 의해 서로 다른 스케일을 가집니다. 따라서 데이터 분석을 할 때 변수의 중요성과 상관없이 특정 변수의 중요성이 스케일에 의해 좌우될 수 있습니다. 따라서 본 분석에서는 정규화 방식을 사용해 변수를 표준화하여 스케일에 따라 변수의 중요성이 희석되지 않도록 하였습니다. 아래는 표준화에 사용한 코드 및 결과입니다.

```
1 from sklearn.preprocessing import StandardScaler
2
3 scaler = StandardScaler()
4
5 scaler.fit(data_for_cultural_index1)
6
7 culture_data_standardized = scaler.transform(data_for_cultural_index1)
8
9 culture_data_standardized = pd.DataFrame(culture_data_standardized, columns=data_for_cultural_index1.columns)
```

✓ 0.0s

```
culture_data_standardized.index = data_seoul_gun['자치구']
```

✓ 0.0s

```
culture_data_standardized
```

✓ 0.0s

	스포츠 동호회 수	장애인 커뮤니티	내부시설수	외부시설수	2022년 생활체육 총강좌수	장애인 의료 사회복지시설 수
자치구						
강남구	-0.681171	-0.492425	1.092378	-0.532016	-0.696199	2.233552
강동구	0.039245	-0.554136	1.449023	-0.603789	0.258910	0.653392
강북구	0.691931	-0.539145	-0.146777	0.150767	-0.778516	-0.777986
강서구	1.306800	0.105933	0.000162	0.611432	-0.828378	-0.372596

```

scaler = StandardScaler()
scaler.fit(data_app1)

✓ 0.0s
StandardScaler()

app_data_standardized = scaler.transform(data_app1)
✓ 0.0s

app_data_standardized = pd.DataFrame(app_data_standardized, columns=data_app1.columns)
✓ 0.0s

app_data_standardized
✓ 0.0s

```

	매개시설_설치완료	내부시설_설치완료	위생시설_설치완료	안내시설_설치완료	기타시설_설치완료	지하철_엘리베이터 수	전동보장구_급속충전소	지하철_휠체어리프트 수
0	0.219463	-0.122472	0.146232	-0.090425	0.600597	0.117299	0.204934	-0.325188
1	0.704984	0.761202	0.587607	0.374476	-3.905686	-0.204121	-0.485030	-0.537476
2	0.704984	-2.851758	-1.190062	0.374476	0.342425	-0.759844	-0.504859	-0.870893
3	0.704984	0.594513	0.817301	0.374476	0.000000	-0.627050	-0.336887	-0.708531
4	0.111630	-1.392003	-1.080681	0.106999	0.000000	-0.737645	0.127487	-0.870893
5	0.704984	0.761202	0.249951	0.374476	0.600597	-0.123196	1.757973	-0.040793
6	0.185292	0.484356	-2.573184	0.374476	0.000000	-0.559700	-1.034555	0.498294

2. PCA를 활용한 장애인 시설 및 접근성 지수 산출

본 분석에서는 개요에서 제시한 네 가지 지표를 산출하기 위해 PCA(Principal Component Analysis)를 활용했습니다. PCA는 다차원 데이터의 차원을 축소하거나 변수의 공통된 요인을 추출하기 위해 사용하는 통계 도구입니다. X라는 데이터를 가지고 있다고 할 때, 해당 데이터를 Factor Loading 'A'와 Factor 'F' 두 가지 행렬로 Decomposition하여 최종적으로 본래 데이터인 X보다 낮은 차원의 Factor를 산출합니다. A와 F를 Decomposition하는 과정에서 본래 데이터인 X의 분산, 즉 정보를 최대한 보존하는 요인을 찾아내는 것이 PCA의 골자입니다.

장애인의 건강 생활 수준 / 문화 커뮤니티 접근 가능성 / 대중교통 접근성 / 편의시설 향유 정도를 나타낼 수 있는 변수는 많습니다. 위에서 데이터 전처리 과정을 거친 변수들 모두 지표로 활용할 수 있습니다. 하지만 종합적으로 현재 상황을 판단하고, 문제를 발견해 개선하기 위해서는 한 가지 측면이 아닌 종합적인 측면을 반영한 지표가 필요합니다. PCA는 다차원 데이터의 차원을 축소해 여러 변수를 종합적으로 고려할 수 있는 통계 도구입니다. 따라서 본 분석에서는 PCA를 활용해 지표를 산출하였습니다. PCA를 거치면, 여러 개의 주성분이 나옵니다. 이에 따라 여러 요인을 도출할 수 있습니다. 하지만 본 데이터 분석에서는 지표 도출이라는 분석의 목적에 따라 가장 큰 공통 분산, 즉 Eigen Value가 가장 큰 주성분만을 활용하였습니다.

각 지표에 포함될 변수는 전처리된 변수들의 특성에 따라 분류하였습니다. 먼저, 장애인 건강생활지수는 장애인의 생활체육 및 건강 유지 여건을 판단하는 지표이므로 “장애인 1인당 생활체육 교실”, “장애인 1인당 생활체육 동호인 클럽 수”, “장애인 1인당 사회복지시설수”를 포함하였습니다. 다음으로, 장애인 문화 커뮤니티 지수는 장애인의 문화 활동 및 커뮤니티 활동 접근성을 파악해야 하므로, “장애인 1인당 실내 문화공간”, “장애인 1인당 실외 문화공간”, “장애인 1인당 장애인 단체 수”, 세 가지 변수를 활용해 만들었습니다.

신체적 장애인 이동성 지표는 “지체 장애인 1인당 지하철 휠체어리프트 수”, “지체 장애인 1인당 지하철 엘리베이터 설치대수”, “지체 장애인 1인당 전동보장구 급속충전소 수”, 그리고 장애인 편

의시설 설치완료 지수는 "매개시설(장애인 통행 가능 접근로, 주차구역 등) 설치완료 비율", "위생 시설(화장실, 욕실 등) 설치완료 비율", "안내시설(점자블록, 유도 및 안내설비 등) 설치완료 비율", "내부시설(복도, 승강기 등) 설치완료 비율"을 통해 만들었습니다.

아래 코드는 차례로 PCA를 통해 장애인 건강생활지수, 장애인 문화 커뮤니티 지수, 신체적 장애인 이동성 지수, 장애인 편의시설 설치완료 지수를 만드는 과정 및 결과입니다.

<장애인 건강생활지수>

```
col1 = ['스포츠_동호회_수', '2022년_생활체육_총강좌수', '장애인_의료_사회복지시설_수']
col2 = ['장애인_커뮤니티', '내부시설수', '외부시설수']

import pandas as pd
from sklearn.decomposition import PCA

pca = PCA(n_components=3)
principal_components = pca.fit_transform(culture_data_standardized[col1])

loadings1_1 = pd.DataFrame(pca.components_.T, columns=['PC1', 'PC2', 'PC3'], index=col1)

loadings1_1

relevant_components = loadings1_1 # Select the components that align with the concept of life quality
# Calculate index scores
index_scores_sport = pd.DataFrame(principal_components.dot(relevant_components.T))

index_scores_commun['자치구'] = data_seoul_gun['자치구']

Culture_commun_Index = index_scores_commun.rename(columns = {0 : "Culture_commun_Index"})[['자치구', 'Culture_commun_Index']]

Culture_commun_Index
```

	PC1	PC2	PC3
스포츠_동호회_수	0.711703	0.052134	0.700543
2022년_생활체육_총강좌수	0.685513	-0.269397	-0.676385
장애인_의료_사회복지시설_수	0.153461	0.961617	-0.227469

자치구	Culture_commun_Index	
0	강남구	-0.492425
1	강동구	-0.554136
2	강북구	-0.539145
3	강서구	0.105933
4	관악구	-0.310574
5	광진구	-0.856037
6	구로구	0.353351
7	금천구	-0.856037
8	노원구	-0.653675
9	동작구	-0.099290
10	마포구	-0.429282
11	서초구	1.765337
12	성동구	-0.371731
13	성북구	0.091974

<장애인 문화 커뮤니티 지수>

```
pca = PCA(n_components=3)
principal_components = pca.fit_transform(culture_data_standardized[col2])

loadings1_2 = pd.DataFrame(pca.components_.T, columns=['PC1', 'PC2', 'PC3'], index=col2)
```

```
loadings1_2
```

	PC1	PC2	PC3
장애인_커뮤니티	0.607055	0.424358	0.671866
내부시설수	0.487267	-0.866657	0.107128
외부시설수	0.627739	0.262345	-0.732884

```
relevant_components = loadings1_2 # Select the components that align with the concept of life quality

# Calculate index scores
index_scores_commun = pd.DataFrame(principal_components.dot(relevant_components.T))
```

```
index_scores_commun['자치구'] = data_seoul_gun['자치구']
```

```
index_scores_commun['자치구'] = data_seoul_gun['자치구']
```

```
culture_commun_index = index_scores_commun.rename(columns = {0 : "Culture_commun_index"})[['자치구', 'culture_commun_index']]
```

```
culture_commun_index
```

	자치구	Culture_commun_index
0	강남구	-0.492425
1	강동구	-0.554136
2	강북구	-0.539145
3	강서구	0.105933
4	관악구	-0.310574
5	광진구	-0.856037
6	구로구	0.353351
7	금천구	-0.856037
8	노원구	-0.653675
9	동작구	-0.099290
10	마포구	-0.429282
11	서초구	1.765337
12	성동구	-0.371731
13	성북구	0.091974

<신체적 장애인 이동성 지수>

```
pca = PCA(n_components=3)
principal_components = pca.fit_transform(app_data_standardized[app2_list])

loadings3 = pd.DataFrame(pca.components_.T, columns=['PC1', 'PC2', 'PC3'], index=app2_list)
```

```
loadings3
```

	PC1	PC2	PC3
지하철_엘리베이터_수	0.682107	-0.140040	-0.717718
전동보장구_급속충전소	0.319106	0.940111	0.119840
지하철_휠체어리프트_수	0.657952	-0.310772	0.685944

```
relevant_components = loadings3[['PC1', 'PC2', 'PC3']] # Select the components that align with the concept of life quality

# Calculate index scores
index_scores = pd.DataFrame(principal_components.dot(relevant_components.T))
```

```
index_scores['자치구'] = data_seoul_gun['자치구']
```


Physical_Accessibility_Index		
	Physical_Accessibility_Index	자치구
0	0.117299	강남구
1	-0.204121	강동구
2	-0.759844	강북구
3	-0.627050	강서구
4	-0.737645	관악구
5	-0.123196	광진구
6	-0.559700	구로구
7	-0.790187	금천구
8	-0.376219	노원구
9	-0.135822	동작구
10	0.531676	마포구
11	0.319765	서초구
12	0.085398	성동구
13	-0.515010	성북구
14	0.508822	송파구
15	-0.674384	양천구
16	-0.019506	영등포구

<장애인 편의시설 설치완료 지수>

```
import pandas as pd
from sklearn.decomposition import PCA

pca = PCA(n_components=4)
principal_components = pca.fit_transform(app_data_standardized[app1_list])

loadings2 = pd.DataFrame(pca.components_.T, columns=['PC1', 'PC2', 'PC3', 'PC4'], index=app1_list)
```

✓ 0.0s

```
loadings2
```

✓ 0.0s

	PC1	PC2	PC3	PC4
매개시설_설치완료	-0.513371	-0.436309	-0.738970	0.002725
내부시설_설치완료	-0.530352	0.464594	0.096723	0.702512
위생시설_설치완료	-0.547623	0.422702	0.128243	-0.710623
안내시설_설치완료	-0.394064	-0.644290	0.654310	0.038510

```

relevant_components = loadings2[['PC1','PC2','PC3','PC4']] # Select the components that align wi

# Calculate index scores
index_scores = pd.DataFrame(principal_components.dot(relevant_components.T))

✓ 0.0s

index_scores['자치구'] = data_seoul_gun['자치구']

✓ 0.0s

Facility_Index = index_scores.rename(columns={0 : "Facility_Index"})[['Facility_Index','자치구']]

✓ 0.0s

Facility_Index
✓ 0.0s

```

	Facility_Index	자치구
0	0.219463	강남구
1	0.704984	강동구
2	0.704984	강북구
3	0.704984	강서구
4	0.111630	관악구
5	0.704984	광진구
6	0.185292	구로구
7	0.704984	금천구
8	0.704984	노원구

3. 시각화 자료를 통한 결과 분석

3-1. 주요 변수 분포 시각화



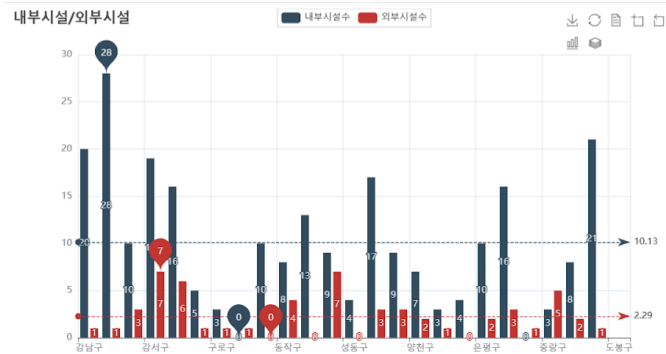
스포츠 동호회/장애인 커뮤니티 수 line chart

```

1 import pandas as pd
2 from pyecharts import options as opts
3 from pyecharts.charts import Line
4 import os
5
6 # CSV 파일 경로
7 csv_file_path = "/content/drive/MyDrive/문화재단/장애인_문화스포츠_인포란_모음 (1).csv"
8
9 # 원하는 HTML 파일의 경로와 파일명 설정
10 html_file_path = os.path.join("/content/drive/MyDrive/문화재단", "sports_community.html")
11
12 # CSV 파일 로드
13 df = pd.read_csv(csv_file_path)
14
15 # 내부시설/외부시설 수 데이터 추출
16 data1 = df[["스포츠_동호회수"]].tolist()
17 data2 = df[["장애인_커뮤니티"]].tolist()
18
19 # 행 이름 추출
20 row_names = df[["자치구"]].tolist()
21
22 # Line 그래프 생성
23 line = Line()
24
25 .add_xaxis(row_names)
26 .add_yaxis("스포츠_동호회수", data1, mark_line_opts=MarkLineOpts(data_opts=MarkLineOpts(type="average")), markpoint_opts=MarkPointOpts(data_opts=MarkPoint(type="max"), opts=MarkPoint(type="min"))))
27 .add_yaxis("장애인_커뮤니티", data2, mark_line_opts=MarkLineOpts(data_opts=MarkLineOpts(type="average")), markpoint_opts=MarkPointOpts(data_opts=MarkPoint(type="max"), opts=MarkPoint(type="min"))))
28 .set_global_opts(title_opts=TitleOpts(title="스포츠_동호회수/장애인_커뮤니티"), toolbar_opts=ToolbarOpts())
29
30 line.render(html_file_path)
31
32 # "/content/drive/MyDrive/문화재단/sports_community.html"

```

장애인 스포츠 동호회 수/ 장애인 커뮤니티 구별 개수에 대해 line 그래프로 나타내고 각각의 평균, 최대값, 최소값에 대해 나타내었습니다. 이를 통해 각 자치구별로 장애인 스포츠 동호회와 장애인 커뮤니티를 4.67개와 2.04개를 평균적으로 가지고 있다는 사실을 알 수 있습니다.



내부시설/외부시설 bar chart

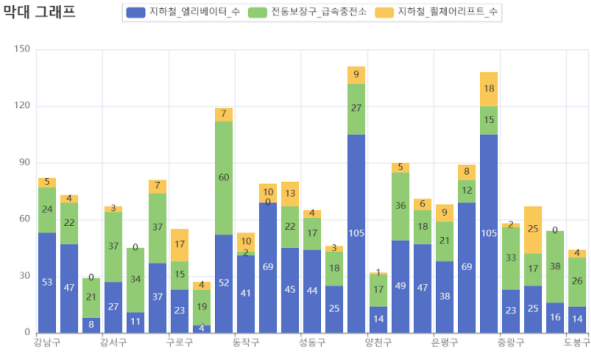
```

1 import pandas as pd
2 from pyecharts import options as opts
3 from pyecharts.charts import Bar
4 import os
5 import pyecharts
6
7 # CSV 파일 경로
8 csv_file_path = "/content/drive/MyDrive/문화재단/장애인_문화스포츠_인포란_모음 (1).csv"
9
10 # 원하는 HTML 파일의 경로와 파일명 설정
11 html_file_path = os.path.join("/content/drive/MyDrive/문화재단", "indoor_outdoor.html")
12
13 # CSV 파일 로드
14 df = pd.read_csv(csv_file_path)
15
16 # 내부시설/외부시설 수 데이터 추출
17 data1 = df[["내부시설수"]].tolist()
18 data2 = df[["외부시설수"]].tolist()
19
20 # 행 이름 추출
21 row_names = df[["자치구"]].tolist()
22
23 # Bar 그래프 생성
24 bar = Bar()
25
26 .add_xaxis(row_names)
27 .add_yaxis("내부시설수", data1, linestyle_opts=LineStyleOpts(color="#3399CC"), mark_line_opts=MarkLineOpts(data_opts=MarkLineOpts(type="average")), markpoint_opts=MarkPointOpts(data_opts=MarkPoint(type="max"), opts=MarkPoint(type="min")))
28 .add_yaxis("외부시설수", data2, linestyle_opts=LineStyleOpts(color="#CC3333"), mark_line_opts=MarkLineOpts(data_opts=MarkLineOpts(type="average")), markpoint_opts=MarkPointOpts(data_opts=MarkPoint(type="max"), opts=MarkPoint(type="min")))
29 .set_global_opts(title_opts=TitleOpts(title="내부시설/외부시설"), toolbar_opts=ToolbarOpts())
30
31 bar.render(html_file_path)

```

장애인 내부시설/외부시설의 구별 개수를 막대 그래프로 나타내고 평균, 최대값, 최소값을 표시 하였습니다. 이 그래프를 통해 이를 통해 각 자치구별로 장애인 내부시설/외부시설을 10.13개와 2.29개를 평균적으로 가지고 있다는 사실을 알 수 있습니다. 또한 강동구는 28개의 내부시설을 보유하고 있고, 강서구는 7개의 외부시설을 보유하고 있어, 각 자치구별로 가장 많이 보유하고 있다는 사실을 알 수 있습니다.

누적 막대 그래프



지하철 엘리베이터, 휠체어 리프트, 급속충전소 누적 막대차트

```

import pandas as pd
from pyecharts import options as opts
from pyecharts.charts import Bar

# CSV 파일 경로
csv_file_path = "/content/drive/MyDrive/문화데이터/강변인, 문화스포츠, 의프라, 모음 (1).csv"

# CSV 파일 로드
df = pd.read_csv(csv_file_path)
# 원하는 HTML, 파일의 경로를 파일명 설정
html_file_path = os.path.join("/content/drive/MyDrive/문화데이터", "stacked_bar.html")

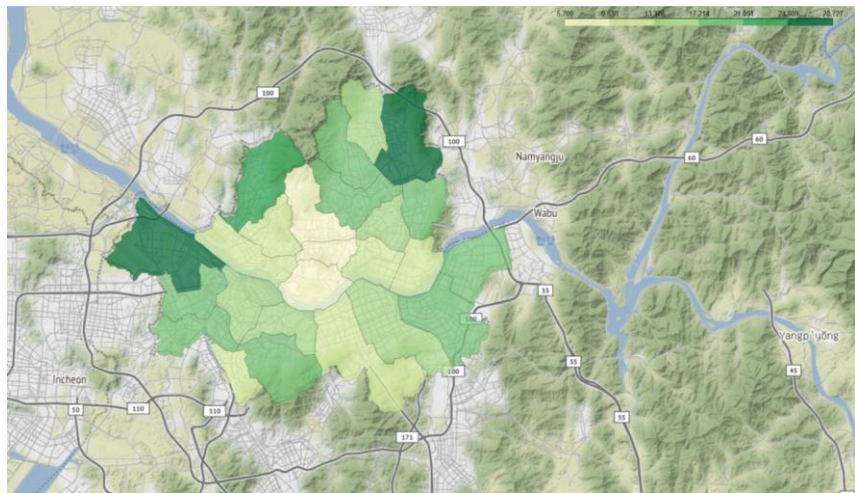
# 행 이름 설정
row_names = df["자치구"].tolist()
# x, y, z 축 (데이터) 설정
v_data1 = df["지하철_엘리베이터_수"].tolist()
v_data2 = df["전동보장구_급속충전소"].tolist()
v_data3 = df["지하철_휠체어리프트_수"].tolist()

# 막대 그래프 생성
bar = (
    Bar()
    .add_xaxis(row_names)
    .add_yaxis("지하철_엘리베이터_수", v_data1, stack="stack1")
    .add_yaxis("전동보장구_급속충전소", v_data2, stack="stack2")
    .add_yaxis("지하철_휠체어리프트_수", v_data3, stack="stack3")
    .set_series_opts(label_opts=opts.LabelOpts(is_show=False)) # 막대 위에 값 표시
    .set_global_opts(title_opts=opts.TitleOpts(title="누적 막대 그래프"))
)

# 그래프 렌더링
bar.render(html_file_path)

```

지하철 엘리베이터 수 , 전동보장구 급속 충전소, 지하철 휠체어 리프트 수에 대해 각 구마다 누적 막대그래프로 나타내어 자치구별 지하철역 접근성을 시각화였습니다.



구별 장애인 수 지도에 시각화

```
[ ] pip install folium

Requirement already satisfied: folium in /usr/local/lib/python3.10/dist-packages (0.14.0)
Requirement already satisfied: branca>=0.5.0 in /usr/local/lib/python3.10/dist-packages (from folium) (0.5.0)
Requirement already satisfied: Jinja2>=2.9 in /usr/local/lib/python3.10/dist-packages (from folium) (3.1.2)
Requirement already satisfied: numpy in /usr/local/lib/python3.10/dist-packages (from folium) (1.22.4)
Requirement already satisfied: requests in /usr/local/lib/python3.10/dist-packages (from folium) (2.27.1)
Requirement already satisfied: MarkupSafe>=2.0 in /usr/local/lib/python3.10/dist-packages (from Jinja2>=2.9->folium) (2.1.3)
Requirement already satisfied: urllib3<1.27, >=1.21.1 in /usr/local/lib/python3.10/dist-packages (from requests->folium) (1.26.16)
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.10/dist-packages (from requests->folium) (2023.5.7)
Requirement already satisfied: charset-normalizer<=2.0.0 in /usr/local/lib/python3.10/dist-packages (from requests->folium) (2.0.12)
Requirement already satisfied: idna<4, >=2.5 in /usr/local/lib/python3.10/dist-packages (from requests->folium) (3.4)

[ ] import folium
import pandas as pd
import json
import re

geojson= https://raw.githubusercontent.com/southkorea/seoul-maps/master/estat/2019/son/seoul-municipalities_geo-sample.json'

csv_file_path = '/content/drive/MyDrive/문화데이터/장애인_문화스포츠츠_인프라_모음 (1).csv'
df = pd.read_csv(csv_file_path)

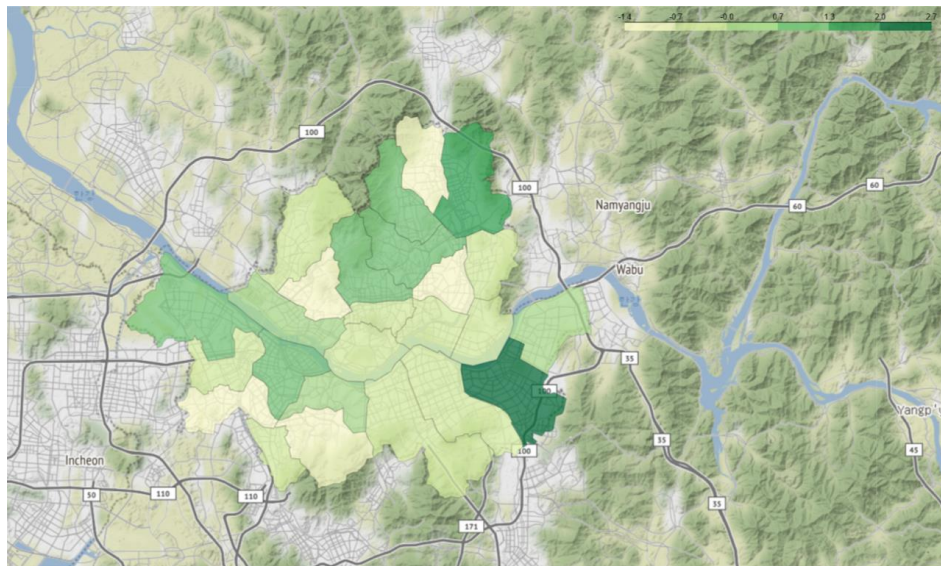
df2=df[['자치구', '장애인_합계']]
df2.columns=['name', 'values']

m=folium.Map(
    location=[37.559819, 126.963896],
    tiles='Stamen Terrain'
)

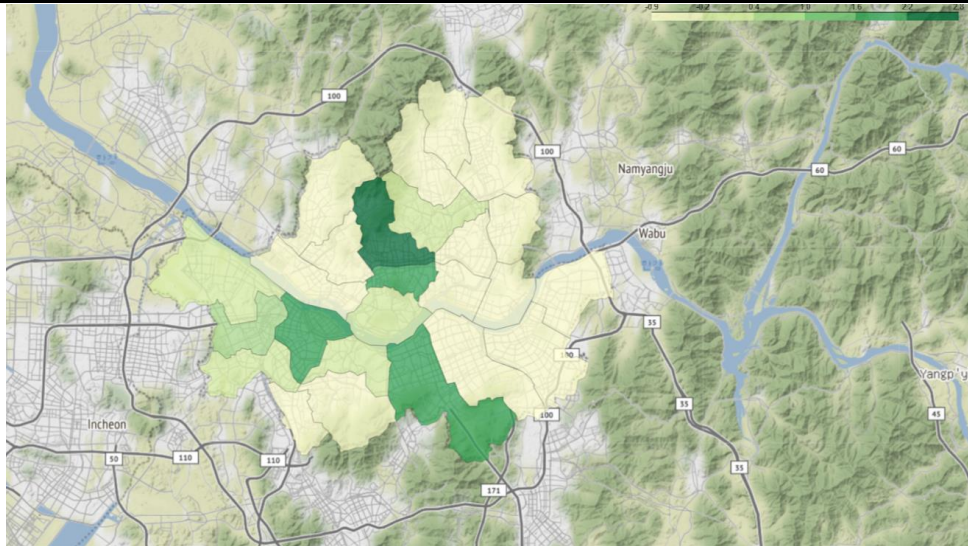
folium.Choropleth(
    geo_data=geojson,
    name='choropleth',
    data=df2,
    columns=['name', 'values'],
    key_on='properties.name',
    fill_color='YlGn',
    fill_opacity=0.7,
    line_opacity=0.2
).add_to(m)
m.save('/content/drive/MyDrive/문화데이터/Map.html')
```

장애인 인구 수를 서울특별시 지도에 나타내어 자치구별 장애인 인구를 나타내었습니다. 이 지도를 통해 강서구와 노원구의 장애인 인구 수가 높다는 것을 알 수 있습니다.

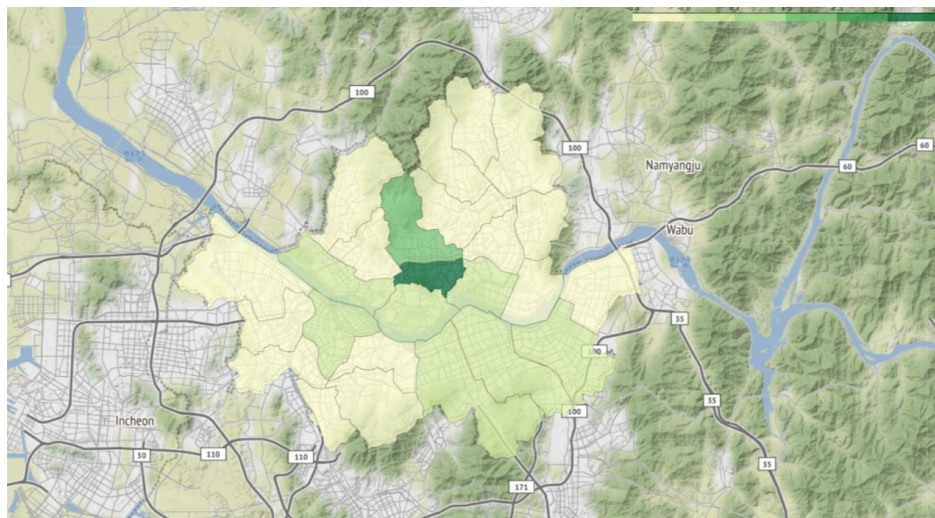
3-2. PCA를 통해 도출한 지표 시각화



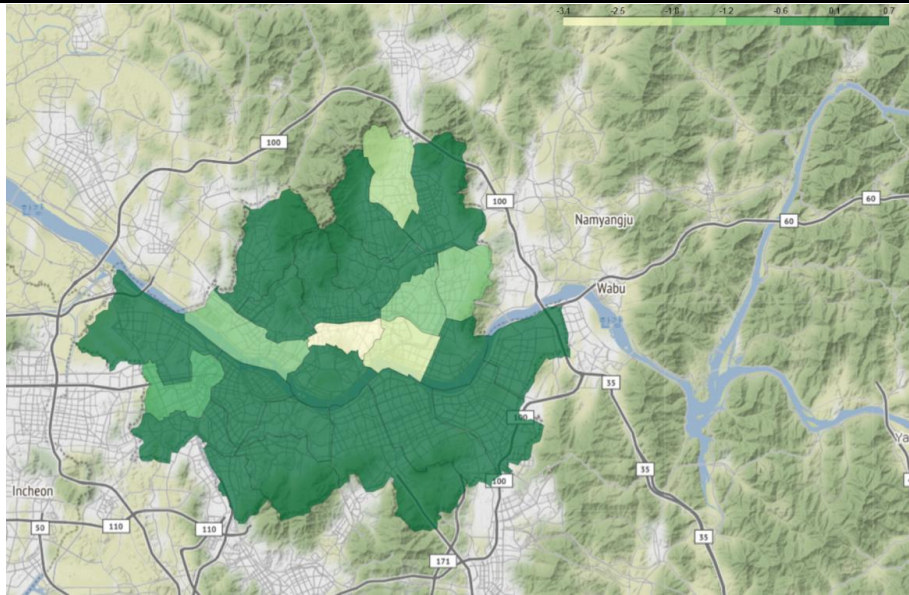
<장애인 건강 생활 지수>



<장애인 문화 커뮤니티 지수>



<신체적 장애인 이동성 지수>



<장애인 편의시설 설치 완료도>

장애인 건강 생활 지수는 문화 커뮤니티 지수에 비해 전반적으로 높은 분포를 보입니다. 지역별로 장애인들이 향유할 수 있는 문화 생활 및 커뮤니티 수준의 편차가 크다는 의미입니다. 따라서 전반적으로 장애인의 삶의 질을 향상시키기 위해서는 문화 커뮤니티 지수가 상대적으로 높게 나타난 지역을 제외한 나머지 지역들을 타겟으로 문화 커뮤니티 영역의 여건 보장이 가능하도록 정책 역량을 집중해야 합니다.

또한, 서울시 중구는 장애인 문화 커뮤니티 지수에서 다른 지역에 비해 상대적으로 높은 수준이지만, 편의시설 설치 완료 수준은 상대적으로 낮은 편입니다. 해당 지역에서 장애인들이 참여할 수 있는 문화생활과 커뮤니티 여건은 충분하지만, 편의시설 인프라가 덜 갖춰진 상황입니다. 따라서 실질적으로 문화 커뮤니티를 통한 장애인의 삶의 질 향상을 이끌어내기 위해서는 해당 지역의 편의시설 인프라 구축이 필요합니다. 중구와 유사하게 노원구는 건강 생활 지수는 높은 편이지만, 신체적 장애인 이동성에서는 취약합니다. 따라서 이동성 개선을 통해 실질적으로 건강 생활 여건을 영위할 수 있도록 지역 환경을 개선해야 합니다.

이와 같이 위의 네 가지 지표는 장애인 건강 개선 및 문화 생활 개선 정책을 설계하는 과정에서 정책의 타겟을 설정하고, 미흡한 면을 파악하는 데 유용하게 사용할 수 있습니다.

5) 시사점 및 기대효과

※ 분석 결과 적용 시 기대효과 및 시사점, 파급 효과

※ 분석 결과 활용 방안 및 적용 대상

본 분석은 장애인들의 집 밖 활동의 불편여부 조사, 생활체육 참여율 조사 결과에 따라 서울시의 장애인 인구수, 인프라 시설, 문화 및 사회 관련 시설을 분석하여 자치구별 장애인 친화도를 측정하였습니다. 분석 결과를 바탕으로 본 시사점은 다음과 같습니다.

첫째, 장애인에 대한 바깥 활동이 불편하다는 현실을 근거로, 필요한 인프라 개선 및 정책 지원

이 요구됩니다. 그렇게 함으로써, 장애인들이 안전하고 편리하게 이동하고, 다양한 문화 및 사회 활동에 참여하는 데 기여할 수 있게 됩니다. 이를 개선하기 위해 각 자치구별로 차별화된 시설 및 프로그램을 제공하는 정책을 개발하고 실행할 필요가 있습니다.

둘째, 장애인들이 충분한 생활체육 참여를 하지 못하는 문제를 해결하기 위해, 다양한 프로그램과 시설을 제공하고 홍보하는 데 중점을 둘 필요가 있습니다.

셋째, 서울시의 각 자치구별 장애인 인프라와 시설 제공에 따른 삶의 질 차이를 효과적으로 해소하기 위해 지속적인 분석과 평가가 필요함을 제시합니다. 구체적인 문제점과 개선 방안을 도출할 수 있는 데이터 분석과 평가를 통해 지속적인 정책 개선이 이루어져야 합니다.

다음은 분석을 통해 주목할 수 있는 기대효과입니다. 첫째, 장애인 인프라 및 시설 개선으로 인한 삶의 질 향상이 이루어질 것입니다. 서울시의 자치구별 장애인 친화도 분석 결과와 현실적인 피드백을 바탕으로, 각 자치구에서 편의시설과 접근성을 개선하여 장애인들의 삶의 질 향상을 도모할 수 있을 것입니다.

둘째, 조사 및 분석을 통해 파악한 문제점들에 대한 적절한 정책 개발로 서울시가 보다 실질적인 개선을 이룰 수 있습니다. 2023년 보건복지부에서 장애인건강과를 신설함에 힘입어 필요성과 효과성이 확보된 정책들을 도입할 수 있는 기회라고 볼 수 있습니다.

셋째, 충분한 생활체육 참여 기회 제공을 통해 장애인들의 건강과 정신적 복지가 향상됩니다. 건강한 신체 활동이 순환 효과를 가져, 정신적 측면에서 긍정적 변화를 가져올 수 있습니다. 이를 통해 사회의 건강한 발전을 도모하게 됩니다. 마지막으로 서울시 전체적으로 장애인 인식 개선이 이루어질 것으로 기대할 수 있습니다. 정부와 각종 단체, 기관의 노력은 시민들의 인식에 큰 변화를 가져올 수 있으며, 이를 통해 전반적인 사회적 분위기 변화가 이루어질 것입니다.

분석의 결과로 산출한 지표를 바탕으로 정책을 도입한다면 장애인들이 겪고 있는 불편함을 해소함과 동시에 커뮤니티를 활성화함으로써 장애인의 삶의 질 향상, 인권 보호와 사회적 편견 감소를 이끌어 낼 것이며, 우리 사회를 더욱 포용적이고 안목이 넓은 곳으로 만드는 데 크게 기여할 수 있을 것입니다.
