|  |
| --- |
| 「2023년 제11회 문화데이터 활용 경진대회」참가신청서 |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| \* 해당란에 **☑** 표시 | | | | | | | |
| **공모 분야** | □ 제품·서비스 개발 | | □ 아이디어 기획 | | | | **☑** 데이터 분석 |
| **참가 구분** | □ 개인 | | | | **☑** 팀(기업) | | |
| **창업 구분** | □ 창업 (사업자등록번호 : , 법인등록일 : 년 월 일) | | | | | | |
| □ 창업예정 | | | | **☑** 해당없음 | | |
| **팀 명** | 앞3정1 | | | | | | |
| **제품·서비스 명**  **(분석 주제명)** | 서울특별시 자치구별 성별, 연령별 관심 여가 활동에 따른 문화체육시설 입지 제안 | | | | | | |
| **제품·서비스 개요**  **(분석 개요)** | 서울특별시의 자치구별 성별, 연령에 따른 관심 여가 활동을 알아보고, 해당 여가 활동과 관련있는 문화체육시설의 개수를 통해 문화체육시설의 분포가 적절한지 파악한 뒤 필요한 추가 시설의 입지를 제안한다. | | | | | | |
| **활용데이터 분야**  **※복수체크가능** | □ 문화예술  □ 문화유산  □ 문화산업  □ 관광  **☑** 체육 | | | | **☑** 문화홍보  □ 정책지원  □ 도서  □ 미디어·콘텐츠  **☑** 기타 ( 인구 ) | | |
| **활용데이터**  **정보**  **※복수기재가능** | **출처** | | | **제공기관명** | | **데이터명** | |
| 서울 열린 데이터 광장 | | | 서울특별시 | | 서울시 자치구별 연령별 인구(추계인구) 통계 | |
| 환경 빅데이터 플랫폼 | | | 지디에스컨설팅그룹 | | 가구 특성정보(+소득정보) | |
| 문화 빅데이터 플랫폼 | | | 컨슈머인사이트 | | 문화여가활동 특징 | |
| 문화 빅데이터 플랫폼 | | | 한국문화정보원 | | 문화역세권 | |
| 문화 빅데이터 플랫폼 | | | 국민체육진흥공단 | | 전국공공체육시설 데이터 | |
| <https://tinyurl.com/2qadx23k> | | | 민간데이터 | | 대한민국 최신 행정구역(SHP) | |
| 문화 빅데이터 플랫폼 | | | 컨슈머인사이트 | | 여가시간 활용 목적 | |
| 문화 빅데이터 플랫폼 | | | 한국문화정보원 | | 문화역세권 시설 성별 관심도 | |
| 문화 빅데이터 플랫폼 | | | 한국문화정보원 | | 문화예술체육소비 시군구별 인구 표준화 시설 점수 | |
| **참가자**  **정보** | **성 명** | **소 속** | | **연락처** | | **이메일** | |
| 박준희 | 고려대학교 통계학과 | | 010-5909-3711 | | jhpark0256@korea.ac.kr | |
| 오승주 | 고려대학교 통계학과 | | 010-6329-9589 | | seungjuniper@korea.ac.kr | |
| 이예지 | 고려대학교 통계학과 | | 010-4726-2895 | | ye\_ji@korea.ac.kr | |
| 전예린 | 고려대학교 통계학과 | | 010-9552-7203 | | jyr7203@korea.ac.kr | |
|  |  | |  | |  | |
| **이전 수혜 이력 및 입상 실적** | **년도** | **내용** | | | | | |
|  |  | | | | | |
|  |  | | | | | |

본인(팀)은 ‘2023년 제11회 문화데이터 활용 경진대회’ 참가와 관련하여 제출한 사항에 허위가 없으며, 유의사항을 숙지하고 진행에 필요한 사항에 성실히 응할 것을 동의합니다.

폰트, 타이포그래피이(가) 표시된 사진

자동 생성된 설명2023년 7월 13일

신청인(대표자) 오승주

|  |
| --- |
| 「2023년 제11회 문화데이터 활용 경진대회」분석보고서 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| |  |  |  |  |  | | --- | --- | --- | --- | --- | | **분석 프로그램** | □ Python | **☑** R | □ Tableau | □ 기타 | |
| **1) 분석 주제** |
| 서울특별시의 자치구별 성별, 연령에 따른 관심 여가 활동을 알아보고, 해당 여가 활동과 관련있는 문화체육시설의 개수를 통해 문화체육시설의 분포가 적절한지 파악한 뒤 필요한 추가 시설의 입지를 제안한다. |
| **2) 분석의 배경 및 목적** |
| <전국 시도별, 서울시 자치구별 표준화 시설 점수>    다음은 ‘문화예술체육소비 시군구별 인구 표준화 시설 점수’ 데이터를 활용해 전국 시도별, 서울시 자치구별 표준화 시설 점수의 분포를 나타낸 막대그래프이다. 지역별 및 분류체계별 시설물 수 비교가 가능하도록 지역별 인구수 및 시설물 수 분포를 고려한 표준화 점수로 변환 제공한 인구(천명당)대비 문화여가시설 표준화 점수의 분포를 통해 전국 시도별로 문화예술체육 시설의 분포에 차이가 있음을 알 수 있으며, 표준화 시설 점수가 가장 높은 전라남도는 가장 낮은 세종특별자치시에 비해 약 60배 높은 것을 고려해보면 전국 시도별 차이가 극심한 것을 알 수 있다. 또한 서울시 내 자치구별로도 문화예술체육 시설의 분포가 차이가 있음을 알 수 있다.  이에 따라 서울특별시의 자치구별 관심 여가 활동을 알아보고 해당 여가 활동과 관련있는 문화체육시설의 위치를 파악해 봄으로써 현재 문화체육시설이 알맞게 분포하고 있는지 살펴보고, 문화체육시설이 적절하게 분포하지 않는 경우 이를 해결할 수 있는 방안을 도모하고자 한다. |
| **3) 활용 데이터 선정** |
| * [서울 열린 데이터 광장] 서울시 자치구별 연령별 인구(추계인구) 통계: 자치구별 인구 분포를 연령과 성별을 구분하여 표시 * [환경 빅데이터 플랫폼] 가구 특성정보(+소득정보): 서울시 자치구별 소득 분포 파악 * [문화 빅데이터 플랫폼] 문화여가활동 특징: 시도별 사람들의 연령, 성별에 따른 여가 활동의 수요를 파악 * [문화 빅데이터 플랫폼] 문화역세권, [문화 빅데이터 플랫폼] 전국공공체육시설 데이터, [민간데이터] 대한민국 최신 행정구역(SHP):서울시 자치구별 지도 및 문화시설 위치 파악 그림, 자치구별 인당 시설 수 계산 * [문화 빅데이터 플랫폼] 여가시간 활용 목적, [문화 빅데이터 플랫폼] 문화역세권 시설 성별 관심도: 성별, 연령, 소득별 여가시간 활용 목적 파악 * [문화 빅데이터 플랫폼] 문화예술체육소비 시군구별 인구 표준화 시설 점수: 지역별 및 분류체계별 시설물 수 비교가 가능하도록 지역별 인구수 및 시설물 수 분포를 고려한 표준화 점수로 변환 제공한 인구(천명당)대비 문화여가시설 표준화 점수 분포 파악 (전국, 서울시) |
| **4) 분석 내용 및 결과** |
| [전국 시도별, 서울시 자치구별 표준화 시설 점수]  <데이터 전처리>  문화예술체육소비 시군구별 인구 표준화 시설 점수 데이터셋에서 공연시설, 전시시설, 지역문화시설, 생활스포츠시설, 전문스포츠시설, 공원에 해당하는 지역 인구 표준화 시설 점수를 추출해 전국 시군구별 점수의 총합을 구해 새로운 데이터셋을 만들었다. 이를 이용해 같은 시도별 데이터를 추출하고 그룹화해 시도별 표준화 시설 점수를 구하고, 서울시 데이터만 추출해 서울시 자치구별 표준화 시설 점수를 구해 시각화에 사용했다.  [성별, 연령대, 소득수준별 관심 여가 활동 분석]  <데이터 전처리>  ‘문화여가활동 특징’ 데이터셋에서 여가 활동의 수요를 파악하기 위해 필요한 변수인 성별, 연령대, 가구 소득 정도, 관심 여가 활동만 추출하고 연령대 변수와 가구 소득 정도 변수의 경우, 분석의 편의를 위하여 데이터 값을 간단한 숫자로 변경하였다: 연령대 - 2(20대), 3(30대), 4(40대), 5(50대), 6(60대), 가구 소득 - 1(300만원 미만), 2(300만원 이상 500만원 미만), 3(500만원 이상 700만원 미만), 4(700만원 이상). 관심 여가 활동의 목록에서 특징이 유사한 것끼리 분류하여 movie, walking, etc, arts&museum, sports\_viewing, in\_sports, workout, out-sports, music\_show로 간소화하였다. 각 목록이 기존 데이터셋에서 해당했던 목록은 아래와 같다.   * movie-: 영화관 관람, 연극-뮤지컬 관람 * walking: 걷기-속보-조깅 직접 하기 * etc: 문학 행사 관람, 전통예술 직접 활동 및 수강하기, 기타 문화예술 직접 활동 및 수강하기, 사진 직접 활동 및 수강하기, 기타 문화예술관람 * arts&museum: 미술 직접 활동 및 수강하기, 미술관-전시관 관람, 박물관 관람 * sports\_viewing: 해외 프로스포츠 영상 시청, 국내 프로스포츠 영상 시청, e스포츠 영상 시청, 골프 영상 시청 기타 스포츠 영상-현장 관람, 국내 프로스포츠 현장 관람, 격투 스포츠 현장 관람, 격투 스포츠 영상 시청 * in\_sports: 요가-필라테스 직접 하기, 수영 직접 하기, 볼링, 당구 포켓볼 직접 하기, 댄스스포츠 직접 하기, 줄넘기-맨손체조-스트레칭 직접 하기, 무도-격투기 직접 하기 * workout: 헬스-보디빌딩 직접 하기 * out-sports: 라켓스포츠 직접 하기, 등산 직접 하기, 겨울레저스포츠 직접 하기, 골프 직접 하기, 자전거-싸이클-산악자전거 직접 타기, 구기스포츠 직접 하기, 기타 스포츠-운동 직접하기 * music\_show: 음악 직접 활동 및 수강하기, 음악 공연 관람   <분석 기법>  성별, 연령대, 소득 수준 변수 모두 각각 두 개 이상의 범주를 가지는 범주형 변수이므로, 범주형 변수를 다루는 통계 모형인 VGLM(Vector Generalized Linear Model, 일반화 선형 모형)을 이용해 분석을 진행해야 한다. 일반화 선형 모형의 경우 설명변수와 반응변수 사이의 관계를 설명하는 데에 사용되며, 이 상황의 경우 여가 활동이라는 반응변수로 두고, 성별, 연령대, 소득수준을 설명변수로 가지는 세 개의 모델이 필요하다. 일반화 선형 모형 분석을 진행할 때에는 설명변수와 반응변수 각각에서 기준 수준(reference level)을 설정해 주어야 하며, 모형의 해석은 기준 수준과 다른 수준들 사이의 차이를 설명하는 것에 초점을 맞추어 진행된다. 예를 들어 성별 변수의 기준 수준이 ‘F(여성)’이고 반응변수의 기준 수준이 ‘out\_sports’인 상황에서의 arts\_museum의 Estimate 값이 -2.8176라면, “여성이 out\_sports에 비해 arts\_museum을 선호할 odds가 남성의 odds의 e^-2.8176배이다”라는 해석이 가능하다. 즉, Estimate 값의 부호가 음수라면 out\_sports에 비해 arts\_museum을 선호할 odds는 남성보다 여성이 더 높다는 의미가 된다.  <분석 결과>  - 성별에 따른 관심 여가 활동    성별에 따른 관심 여가 활동을 파악하기 위해 SEX 변수와 ACTIONS\_NEW 변수를 활용하였다.    위 결과는 reference level을 F(여성), out\_sports로 설정하였을 때의 일반화 선형 모형 분석 결과이다. 표의 Estimate 값의 부호와 p-value로 판단한 변수의 유효성을 함께 고려하여 남성과 여성이 out\_sports에 비해 상대적으로 더 선호하는 여가 활동을 구분하고자 하였다. 이후, reference level을 out\_sports가 아닌 다른 항목으로 변경하는 작업을 반복하며 여가 활동을 구분지었다.  추가적으로, 더 정확한 분류를 위해 ‘문화역세권 시설 성별 관심도’ 자료를 활용하여 각 성별이 각 여가 활동을 선호하는 정도를 파악해보고, 이를 반영하였다.  분석 결과, 각 성별에 따른 관심 여가 활동을 다음과 같이 정리할 수 있었다.  [남] sports\_viewing, out\_sports, workout  [여] walking, in\_sports, arts\_museum, music\_show  - 나이에 따른 관심 여가 활동  연령대별 관심 여가 활동을 알아보기 위해 AGE 변수와 ACTIONS\_NEW 변수를 활용하였다. 다만, AGE 변수가 항목이 5개인 범주형 변수라는 점에서 5개의 항목을 모두 구분하여 분석을 진행할 경우에는 결과 해석이 필요 이상으로 복잡해질 것이라는 판단 하에 5개 범주를 연령대별 사회적 활동을 고려해 3개 범주로 간소화하여 AGE\_NEW 변수를 생성하였다. AGE\_NEW 변수의 경우 1,2,3의 값을 가지며, 각각 대개 사회초년생인 2-30대, 어느 정도 사회에 자리를 잡은 4-50대, 그리고 은퇴를 앞두거나 은퇴를 한 상태인 60대의 연령대를 나타낸다.    위 결과는 reference level을 1(2-30대), sports\_viewing으로 설정하였을 때의 일반화 선형 모형 분석 결과이다. 표의 Estimate 값의 부호와 p-value로 판단한 변수의 유효성을 함께 고려하여 연령대별로 sports\_video에 비해 상대적으로 더 선호하는 여가 활동을 구분하고자 하였으나, 총 범주의 개수와 여가 활동의 종류의 수가 많다는 점을 고려하여 변수의 유효성보다 estimate 값의 부호를 우선적으로 고려하였다. 이후, reference level을 sports\_video가 아닌 다른 항목으로 변경하는 작업을 반복하며 여가 활동을 구분지었다.  분석 결과, 각 연령대에 따른 관심 여가 활동을 다음과 같이 정리할 수 있었다.  [2-30대] movie, sports\_viewing, workout  [4-50대] out\_sports, walking  [60대] walking, in\_sports  - 소득 수준에 따른 관심 여가 활동  소득 수준에 따른 관심 여가 활동을 알아보기 위해 INCOME 변수와 ACTIONS\_NEW 변수를 활용하였다. 우선 소득 수준에 대해 ‘무응답’으로 답한 데이터는는 제거하고 분석을 진행하였다. 위 단계에서와 마찬가지로, 4개의 항목을 가진 INCOME 변수의 구분이 어려울 것이라고 판단하여 300만원 이하/300~500만원/500~700만원/700만원 이상의 4개 범주를 500만원을 기준으로 간소화하여 2개의 범주를 가지는 INCOME\_NEW 변수를 생성하였다. INCOME\_NEW 변수의 경우 1,2의 값을 가지며, 각각 가구소득 정도가 500만원 미만인 데이터, 500만원 이상인 데이터를 나타낸다.    위 결과는 reference level을 1(가구소득 500만원 미만), music\_show로 설정하였을 때의 일반화 선형 모형 분석 결과이다. 마찬가지로 표의 Estimate 값의 부호와 p-value로 판단한 변수의 유효성을 함께 고려하여 연령대별로 music\_show에 비해 상대적으로 더 선호하는 여가 활동을 구분하고자 하였다. 이후, reference level을 music\_show가 아닌 다른 항목으로 변경하는 작업을 반복하며 여가 활동을 구분지었다.  분석 결과, 각 소득수준에 따른 관심 여가 활동을 다음과 같이 정리할 수 있었다.  [500 미만] movie, music\_show  [500 이상] in\_sports, out\_sports, walking, arts\_museum, sports\_viewing  [서울시 자치구별 지역별 인구 특성 분석]  -서울시 자치구별 성별 인구 분포 시각화  <데이터 전처리>  ‘대한민국 최신 행정구역’ 데이터를 활용하여 서울시 시군구 지도를 시각화한 뒤, ‘서울시 자치구별 연령별 인구(추계인구) 통계’ 데이터를 활용하여 성별 인구 분포를 나타내었다.  ‘서울시 자치구별 연령별 인구(추계인구) 통계’를 활용하여 각 자치구별로 남성, 여성의 수를 구한 뒤, 남성의 수/여성의 수로 MALEFEMALE\_RATIO라는 새로운 변수를 만들어주었다. 이때, 20세 이상의 인구만 고려하였다.  <분석 결과>    MALEFEMALE\_RATIO가 높을수록 지도에서는 더 진한 색으로 표시되었으며, 금천구가 여성에 비해 남성의 수가 가장 많은 것으로 나타났고, 서대문구와 마포구 등은 남성에 비해 여성의 수가 많은 것으로 나타났다.  -서울시 자치구별 연령대 분포  <데이터 전처리>  ‘서울시 자치구별 연령별 인구(추계인구) 통계’ 데이터를 활용해 자치구, 성별, 20~24세, 25~29세,30~34세, 35~39세, 40~44세, 45~49세, 50~54세, 55~59세, 60~64세, 65~69세, 70~74세, 75~79세 열을 활용해 해당 연령대의 인구수를 합해 새로운 변수인 20s\_30s, 40s\_50s, 60s\_70s 를 만들어 활용했다.  <분석 결과>    색이 진할수록 자치구에 해당 연령대의 수가 많음을 나타낸다.  -서울시 자치구별 평균 연소득 분포  <데이터 전처리>  ‘대한민국 최신 행정구역’ 데이터를 활용하여 서울시 시군구 지도를 시각화한 뒤, ‘가구 특성정보(+소득정보)’ 데이터를 활용하여 자치구별 평균 연소득을 시각화하였다. 행정동 이름에서 시/도와 시/군/구 이름을 얻어낸 뒤, 서울특별시에 해당하는 행의 평균 연소득만 추려냈다. 이후 시/군/구 이름으로 그룹 지어서 각 자치구별 평균 연소득을 계산했다. 단위는 10,000원이다.  <분석 결과>    강남구와 서초구의 평균 연소득은 1억을 넘어가는 것으로 나타났으며, 강남, 서초의 경우 4-50대가 많고 6-70대는 적다는 점을 고려할 때, 강남구와 서초구의 평균 연소득이 높게 나타나는 것은 타당하다.  2-30대는 사회초년생으로서 높은 연소득을 얻기가 쉽지 않고, 4-50대가 활발한 경제활동으로 평균 연소득이 높은 반면, 6-70대의 경우 소득이 적거나 없는 경우가 많은 것을 생각하면, 연령대와 소득 수준 사이에 관계가 있을 것이라 예상해볼 수 있다.  이에 따라 연령대와 소득 수준 사이에 관계가 있는지 파악하고자 카이제곱 검정을 통한 독립성 검정을 시행하였다. 카이제곱 검정은 두 개 이상의 범주형 변수가 서로 독립인지 확인하고자 할 때 적용할 수 있는 통계적 기법이다.  H0: 연령대와 소득 수준 사이에 관계가 없다 vs. H1: 연령대와 소득 수준 사이에 관계가 있다.    검정 결과, p-value < 0.05이기 때문에 귀무가설을 기각할 수 있으므로 연령대와 소득 수준 사이에 관계가 있다고 결론지을 수 있다. 따라서 이번 분석에서는 소득 수준을 제외하고 연령대만을 고려하기로 한다.  [문화체육시설 분포 및 분석]  -서울시 문화체육시설 위치 시각화  <데이터 전처리>  ‘대한민국 최신 행정구역’ 데이터를 활용하여 서울시 시군구 지도를 시각화한 뒤, ‘문화역세권’ 데이터와 ‘전국공공체육시설’ 데이터를 활용하여 서울시 시군구 지도 위에 문화체육시설의 위치를 표현하였다.  ‘문화역세권’ 데이터는 위도와 경도가 바뀌어 있는 시설들이 존재하여 위도와 경도를 수정해주었고, ‘대한민국 최신 행정구역’ 데이터는 좌표 변환을 통해 WGS84 좌표계로 위도와 경도를 나타내주었다.  ‘문화역세권’ 데이터와 ‘전국공공체육시설’ 데이터에서는 ‘서울’ 혹은 ‘서울특별시’에 위치해있는 시설들의 종류와 위도, 경도를 사용하여 시설의 위치와 종류를 표시해주었다.  <분석 결과>    상당히 많은 수의 체육시설이 서울시에 위치해있다는 것을 확인할 수 있었다. 관악구, 용산구 등 일부 구들은 체육시설이 적게 위치해 있으나, 서초구, 도봉구 등 일부 구에는 체육시설이 상대적으로 많이 분포해있는 것을 확인할 수 있었다.  박물관은 종로구와 중구에 대다수가 위치해 있었으며, 양천구와 중랑구에는 위치해있지 않은 것을 확인할 수 있었다.  전반적으로 공공도서관과 박물관을 제외한 나머지 문화시설들은 적은 것으로 보인다. 체육시설 역시 간이운동장이 다수를 차지하는 것을 확인할 수 있다.  [자치구별 인당 문화시설수]  -성별, 연령대 적용  성별의 경우, 각 자치구별 여성과 남성의 인구 수를 구한 뒤 남성 대비 여성의 비율을 구하였다. 분석 결과, 금천구를 제외한 나머지 24개의 구 모두 남성보다 여성의 수가 더 많았으나, 모든 구에서의 성별 차이가 그렇게 크지 않음을 발견할 수 있었다. 이에 따라 성별 차이가 큰 구들만을 고려하는 것이 올바른 분석의 방향이라고 생각해 남성 대비 여성의 비율이 전체 구에서 구한 남성 대비 여성의 비율의 평균보다 작은 자치구들에서만 분석을 진행하기로 한다.  연령대의 경우, 2-30대, 4-50대, 그리고 6-70대의 인구수를 각각 구한 뒤, 해당하는 구의 전체 인구 대비 연령대별 비율을 구해 연령대별 비율이 가장 높은 연령대를 선정했다.  연령대 변수의 경우에도 각 연령대의 인구 수 차이가 크지 않음을 발견하고, 각 자치구에서 연령대별 비율이 최대인 연령대와 두 번째로 큰 연령대 사이의 차이를 구한 뒤, 25개 값 결과의 평균을 구해 비율값의 차가 평균값보다 큰 지역들에서만 분석을 진행하기로 하였다.  소득의 경우, boxplot을 그렸을 때 강남구와 서초구가 이상치로 판별되었으나, 앞서 설명한 것처럼 강남, 서초의 경우 4-50대가 많고 6-70대는 적다는 점을 고려할 때 자치구별 소득 분포와 자치구별 연령대 분포 사이에 높은 연관성이 있다고 판단하여 이후 분석에서 소득은 고려하지 않기로 결정하였다.  위 과정의 결과로 성별 변수의 경우 종로구 외 15개의 자치구가 선택되었으며, 연령대 변수의 경우 중구, 광진구, 동대문구 외 11개의 자치구가 선택되었다. 마지막으로 두 조합의 교집합을 구하여 분석 대상이 될 자치구 8곳(중구, 광진구, 강북구, 도봉구, 노원구, 마포구, 서초구, 강남구)을 선정할 수 있었다.  다음으로는 각 자치구의 인구 특성을 고려하여 선호하는 여가 활동 범주를 알아보았다. 예를 들어, 중구의 경우 남성에 비해 여성의 수가 많았고, 2-30대의 비중이 가장 높았기 때문에 walking, in\_sports, arts\_museum, music\_show 와 movie, sports\_viewing, workout에 대한 수요가 높을 것이라고 판단하였다.  이 범주 분류 기준에 따르면, 중구, 광진구, 마포구는 선호하는 여가 활동 범주가 walking, in\_sports, arts\_museum, music\_show, movie, sports\_viewing, workout 가 되고, 나머지 5곳인 강북구, 도봉구, 노원구, 서초구, 강남구는 선호하는 여가 활동 범주가 walking, in\_sports, arts\_museum, music\_show, out\_sports가 된다.  이를 바탕으로 각 여가 활동 범주에 속하는 문화체육시설을 선정해, 각 자치구별로 인당 시설수를 계산했다. 여가 활동 범주에 해당하는 문화체육시설은 아래와 같으며, ‘문화역세권’, ‘전국공공체육시설’ 데이터를 활용했다.   * walking: 간이운동장, 육상경기장 * in\_sports: 수영장, 생활체육관, 투기체육관 * arts\_museum: 미술관, 박물관 * music\_show: 문예회관 * movie: 문예회관 * sports\_viewing: 국궁장, 축구장, 야구장, 승마장, 빙상장, 테니스장, 수영장, 하키장, 사이클경기장, 육상경기장, 양궁장 * workout: 기타체육시설 * out-sports: 생활체육관, 테니스장, 빙상장, 하키장, 골프연습장, 사이클경기장, 구기체육관, 축구장, 야구장, 국궁장, 승마장, 수영장, 육상경기장, 양궁장, 간이운동장   ‘문화역세권’ 데이터의 지방문화원/문화의집은 서울시 지도 위에 나타내본 결과, 서울시 바깥을 벗어난 것이 확인되었다. 따라서 데이터셋에 있는 위도/경도에 오류가 있다고 판단하여 제외하고 인당시설수를 계산하였다.  ‘전국공공체육시설’ 데이터의 기타시설은 다양한 종류의 시설이 함께 묶여 있어 특정 여가 활동 종류로 포함시키기에 무리가 있다고 판단하여 제외하고 인당 시설수를 계산하였다.  중구, 광진구, 마포구는 walking, in\_sports, arts\_museum, music\_show, movie, sports\_viewing, workout 에 해당하는 시설 수를 합하고, 각 구별 인구 수인 106832, 292249, 298895로 나누어주었다. 같은 방법으로 강북구, 도봉구, 노원구, 서초구, 강남구는 walking, in\_sports, arts\_museum, music\_show, out\_sports 에 해당하는 시설 수를 합하고, 각 구별 인구 수인 238105, 248388, 394826, 303397, 402011로 나누어주었다.        [제언]  (1) 문화체육시설 추가 설치  서울시 평균 인당 시설수보다 각 구별 관심 여가 활동의 평균 인당 시설수가 적을 때 특정 문화체육시설의 추가적인 설치가 필요하다고 판단하였다. 따라서, 다음과 같이 서울시 평균 인당 시설수를 구하였다.  -서울시 평균 인당 시설수      광진구의 경우, 남성에 비해 여성의 수가 많고 2-30대가 많아 walking, in\_sports, arts\_museum, music\_show, movie, sports\_viewing, workout 활동에 관심에 많다는 결론이 나왔다. 그러나 walking, arts\_museum, music\_show, movie, sports\_viewing 활동과 관련된 문화체육시설의 인당 시설수가 각각 평균 인당 시설수에 비해 0.00001036819, 0.0000199363, 0.0000000626, 0.0000000626, 0.0000110241 정도 낮은 것으로 나타났다. 따라서 해당 구에 추가적으로 간이운동장, 육상경기장, 미술관, 박물관, 문예회관, 축구장, 야구장, 테니스장, 수영장 등을 추가로 설치하여 인당 시설수를 높이는 방안을 생각할 수 있다.  (2) 특정 연령대나 성별의 사람들이 선호하는 여가 활동을 모두 즐길 수 있는 공간 제공  문화체육시설 추가 설치가 어려운 경우, 하나의 시설에서 다양한 여가 활동을 즐길 수 있도록 공간이나 프로그램을 제공하는 방안도 생각해볼 수 있다. 예를 들어, 마포구의 경우 music\_show, movie 활동에 관심도가 높지만, music\_show, movie 관련 시설의 인당 시설수는 서울시 평균 인당 시설수보다 낮은 것으로 나타났다. 이런 경우, music\_show와 관련된 문화시설에서 movie 관련 활동을 할 수 있는 다용도 공간이나 프로그램을 함께 제공해주거나, 반대로 movie와 관련된 문화시설에서 music\_show 관련 활동을 할 수 있는 다용도 공간이나 프로그램을 제공해준다면 마포구의 music\_show, movie와 관련된 시설의 인당 시설수를 높일 수 있다. |
| **5) 시사점 및 기대효과** |
| -문화체육시설의 적절한 분포 파악: 서울시 자치구별 관심 여가 활동과 문화체육시설의 위치를 비교함으로써, 어떤 자치구에서 어떤 종류의 문화체육시설이 부족하게 분포되어 있는지 확인할 수 있다. 이는 해당 자치구의 주민들이 원하는 여가 활동을 즐길 수 있는 시설이 부족하다는 의미로 볼 수 있으며, 이를 보완하기 위한 새로운 문화체육시설 입지의 필요성을 인지할 수 있다.  -정책 개발과 시설 개선 방안 마련: 문화체육시설의 분포 상황을 파악한 후, 시설이 부족한 자치구나 부족한 종류의 시설을 대상으로 정책 개발과 시설 개선 방안을 도모할 수 있다. 예를 들어, 부족한 자치구에 새로운 문화체육시설을 건립하거나 기존 시설을 개선함으로써 주민들의 다양한 여가 활동을 지원할 수 있다.  <활용 방안 및 적용 대상>  이러한 기대효과를 통해 자치구별 관심 여가 활동과 문화체육시설의 분포 상황을 파악하고, 문제점을 인식하며, 개선 방안을 모색함으로써 더 나은 문화체육시설 환경을 조성할 수 있을 것이다. 또한 성별, 연령대별 선호 여가생활의 경우, 전국 단위의 데이터를 활용하여 진행하였으므로, 다른 시/도에도 해당 결과를 적용할 수 있을 것으로 기대된다. |

|  |
| --- |
| 참가 서약서 |

|  |
| --- |
| 본인(팀)은 “2023년 제11회 문화데이터 활용 경진대회”에 출품하며 아래 사항을 숙지하고, 허위사실 기재 및 타인의 권리를 침해하는 등의 행위로 인하여 손해를 발생시키는 경우, 본인의 귀책으로 인하여 발생되는 손해에 관한 손해배상책임이 본인에게 있음을 확인합니다.  1. 이미 채택된 제안과 동일한 것, 표절 및 복제 등의 지적재산권 침해 작품, 타 공모전 입상작품 등은 심사에서 제외되며, 이에 따른 모든 책임은 참가자에게 있음  2. 제출한 작품이 제3자의 권리(소유권, 저작권, 이용권)를 침해하였거나 이와 관련한 분쟁이 발생한 사실이 없으며, 이로 인하여 발생하는 법적인 책임은 출품자에게 있음  3. 수상 이후 위반 사실이 밝혀질 경우 수상 취소 및 상금 환수(자진반납)에 이의를 제기하지 않음  본인은 유의사항을 충분히 숙지하였으며 대회진행에 필요한 주관기관의 요구사항에 성실히 응할 것에 동의합니다.  2023년 7월 13일  폰트, 타이포그래피, 그래픽, 서예이(가) 표시된 사진  자동 생성된 설명  폰트, 타이포그래피이(가) 표시된 사진  자동 생성된 설명 서약자 팀(기업)명 성명 박준희  폰트, 타이포그래피, 화이트, 서예이(가) 표시된 사진  자동 생성된 설명 성명 오승주  폰트, 타이포그래피, 로고, 서예이(가) 표시된 사진  자동 생성된 설명 성명 이예지  성명 전예린    한국문화정보원장 귀중 |