

# 여성 화장실 대기 줄, 데이터로 풀어본 불편의 비밀

## 1. 문제 인식

일상생활 속에서 우리는 종종 공중화장실 앞에서 긴 줄을 기다리는 경험을 합니다. 특히 여성 화장실을 이용할 때마다 남성 화장실에 비해 유독 긴 대기 줄을 마주하며 불편함을 느꼈습니다. 이러한 현상을 목격하며, “여성 화장실 이용 시간이 유독 긴 이유가 무엇일까?” 라는 단순한 호기심에서 본 분석을 시작하게 되었습니다.

이러한 궁금증은 단순히 개인적인 의문점을 넘어, 서울시 공중화장실의 전반적인 수급 현황을 파악하고, 여성 화장실 이용 시간이 길어지는 근본적인 요인들을 분석해보고자 하는 목표로 이어졌습니다. 하지만 화장실 내부에서의 실제 이용 행태를 직접 조사하거나 설문하는 것은 현실적인 제약으로 인해 불가능했습니다. 이에 따라 현실적으로 가능한 범위 내에서 문제에 접근할 대안적인 분석 방법이 필요했습니다.

## 2. 목표

본 프로젝트의 목표는 직접적인 이용 행태 분석의 한계를 극복하고, 실제 활용 가능한 공공 데이터를 활용하여 수급 현황을 간접적으로 진단하는 것입니다. 이를 위해 공공데이터포털에서 제공하는 서울시 공공 화장실 정보와 지역별 인구 현황 데이터를 활용합니다. 구체적으로는 화장실 칸수 대비 인구수를 비교함으로써 1인당 이용 가능한 칸수를 산출하고, 특히 이용이 집중되는 출/퇴근 시간에는 가중치를 부여하는 방식으로 현황을 분석하고자 합니다. 궁극적으로는 이러한 분석을 통해 시민들이 보다 쾌적하게 화장실을 이용할 수 있도록 실질적인 개선 방안을 모색하는 것을 목표로 합니다.

### 3. 데이터 소스 및 전처리 요약

#### [주요 활용 데이터셋]

- 서울시 공공화장실 현황 데이터: 공공데이터포털에서 다운로드 받은 서울시 내 공공화장실 상세 현황 데이터셋을 활용했습니다. 이 데이터는 '소재지도로명주소', '남성용-대변기수', '남성용-소변기수', '여성용-대변기수' 등 성별 및 유형(장애인용, 어린이용)별 칸수 정보와 위치 정보를 포함하고 있습니다.
- 서울시 인구 현황 데이터: 서울시 자치구별 및 행정동별 인구 현황 데이터를 성별로 구분하여 확보함으로써, 화장실 이용 수요를 추정하는 기초 자료로 활용했습니다.

#### [데이터 전처리 및 가공 과정]

##### 1. 화장실 칸수 데이터 정제 및 집계

- 원본 데이터에서 남성용 및 여성용 화장실 칸수(대변기, 소변기, 장애인용, 어린이용 포함) 관련 컬럼들을 대상으로 결측값은 0으로 처리하고, 간혹 존재하는 음수 값 또한 0으로 정제하여 데이터의 무결성과 일관성을 확보했습니다.
- 정제된 칸수 데이터를 바탕으로 '남성\_총칸수'(대변기+소변기), '여성\_총칸수'(대변기), '전체\_총칸수'를 새롭게 계산하여 핵심 분석 지표로 활용했습니다.

##### 2. 주소 기반 행정동 매핑 및 데이터 통합 (카카오 API 활용)

- 공공화장실 현황 데이터는 주소 정보만 포함하고 있어, 인구 데이터의 행정동 단위와 직접적으로 통합하기 어려운 문제가 있었습니다. 이를 해결하기 위해 카카오 REST API를 활용한 2단계 주소-행정동 매핑 프로세스를 구축했습니다.
- 대량의 데이터를 효율적이고 안정적으로 처리하기 위해 배치 처리 방식을 적용했으며, API 호출 한도 초과(429 Error)에 대비한 재시도 로직과 진행 상황 저장을 구현하였습니다.

##### 3. 칸수 대비 인구수 산출 및 수요 가중치 부여

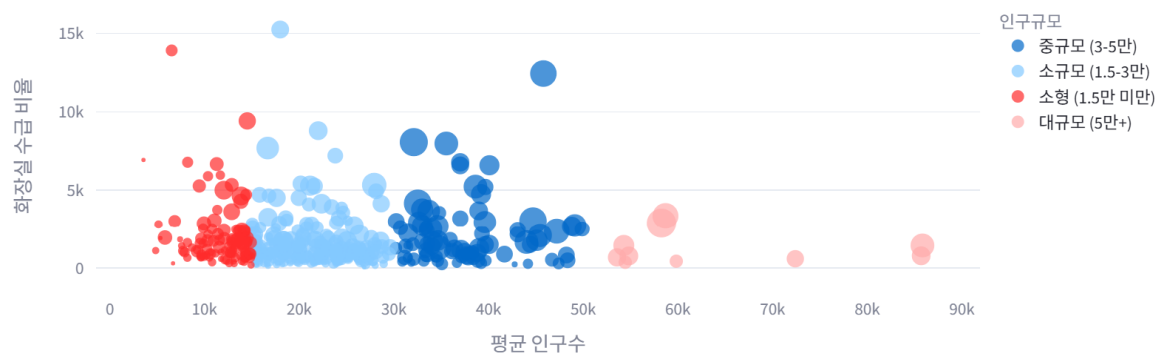
- 통합된 데이터를 바탕으로 각 행정동별 남성 인구수 대비 남성 화장실 칸수, 여성 인구수 대비 여성 화장실 칸수를 비교하여 '1인당 이용 가능한 화장실 칸수'를 산출했습니다. 이는 성별에 따른 화장실 시설 접근성을 정량적으로 파악하는 핵심 지표입니다.

- 더 나아가, 실제 화장실 이용량이 특정 시간대(특히 출/퇴근 시간)에 집중되는 경향을 반영하기 위해 해당 시간대의 이용 수요에 대한 가중치를 부여하여 분석의 현실성을 높였습니다. 이는 정적인 인구수 비교만으로는 파악하기 어려운 시간대별 수급 불균형을 분석하는 데 중요한 역할을 합니다.

#### 4. 핵심 분석 결과 요약

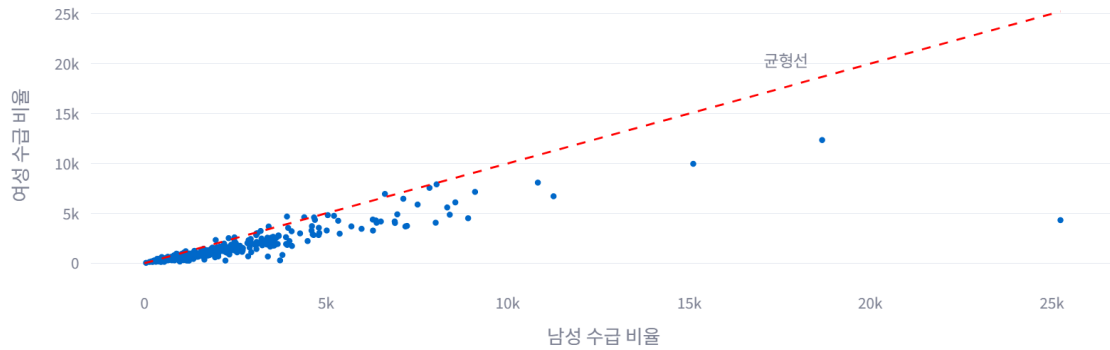
##### [화장실 사막 지역 발굴]

인구 규모 vs 화장실 수급 비율



"이 그래프는 지역의 인구 규모(X축)와 화장실 수급 비율(Y축) 간의 관계를 시각화 합니다. 특히, 인구 규모가 큰(오른쪽 부분) 지역임에도 불구하고 화장실 수급 비율이 현저히 낮은 점들이 다수 발견됩니다. 이는 인구 밀집 지역일수록 화장실 시설의 양적인 공급이 부족하여 '화장실 사막' 현상이 발생하고 있음을 보여줍니다."

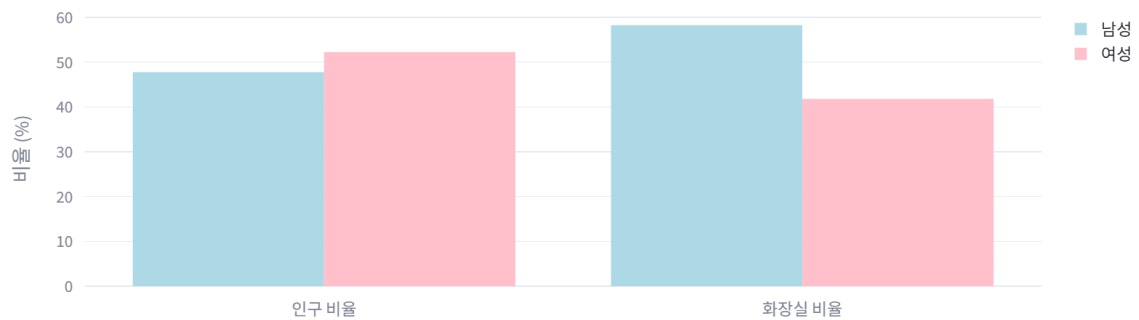
### 성별 화장실 수급 비교



"이 그래프는 각 지역별 남성 화장실 수급 비율(X축) 대비 여성 화장실 수급 비율(Y축)을 보여줍니다. 대다수의 지역이 남녀 수급 균형선(점선) 아래에 위치하고 있으며, 이는 서울시 전반적으로 여성 화장실 칸수가 남성 화장실에 비해 상대적으로 부족함을 명확히 시사합니다. 이러한 불균형은 여성 이용객의 대기 시간 증가에 직접적인 원인이 됩니다."

### [가설검정]

#### 인구 vs 화장실 성별 비율



- 귀무가설: 화장실의 성별 비율(남성 칸수 비율 vs 여성 칸수 비율)이 인구의 성별 비율과 동일하다.
- 대립가설 : 화장실의 성별 비율이 인구의 성별 비율과 동일하지 않다 (화장실 성별 분포에 구조적 불균형이 존재한다).

“p-value가 0.00e+00으로 나타나, 귀무가설을 강력하게 기각합니다. 이는 서울시 화장실의 성별 분포가 인구의 성별 분포와 통계적으로 유의미하게 다르다는 것을 의미합니다. '인구 vs 화장실 성별 비율' 막대 그래프 에서도 인구 성별 비율은 남녀가 거의 비슷하지만, 화장실 성별 비율은 남성 칸수의 비율이 여성 칸수의 비율보다 훨씬 높은 것으로 나타나, 여성 화장실의 구조적인 공급 부족 문제를 통계적으로 뒷받침합니다.”

여성 화장실 수급 현황



- 귀무가설: 여성 균형 지수(여성 수급 비율이 1 미만, 즉 부족한 경우)가 0.5일 확률이 무작위이다 (체계적 불균형이 없다).
- 대립가설 : 여성 균형 지수(여성 수급 비율이 1 미만)가 0.5보다 클 확률이 체계적이다 (여성에게 체계적 불균형이 존재한다).

“p-value가 0에 매우 가까운 7.22e-116으로 나타나, 귀무가설을 강력하게 기각합니다. 이는 여성 화장실이 부족한 지역의 비율이 우연이 아닌, 통계적으로 유의미하게 높다는 것을 의미합니다. 함께 제시된 파이 차트에서도 전체 지역의 95.6%가 여성에게 불이익이 있는 지역으로 나타나, 여성 화장실 수급 불균형이 서울시 전역에 걸쳐 체계적으로 존재함을 통계적으로 입증합니다.”

## 5.인사이트 및 정책 제안

서울시 공중화장실의 수급 현황과 시민들의 체감 불편에 대한 중요한 인사이트가 도출되

었으며, 이는 실질적인 정책 제안으로 이어집니다.

가장 먼저, 서울시 공중화장실의 성별 수급에 심각한 불균형이 존재하는 것으로 나타났습니다. 분석 결과에 따르면, 서울시 전역에서 여성 화장실의 칸수 대비 수급 비율이 남성에 비해 현저히 낮게 나타났습니다. 이러한 성별 불균형은 인구 밀집 지역일수록 더욱 두드러지는 경향을 보였으며, 이는 여성 이용객의 화장실 대기 시간 증가에 직접적인 영향을 미치는 주요 원인으로 분석됩니다.

두 번째 인사이트는 인구 밀집 지역 내 '화장실 사막'의 존재입니다. 인구 규모가 큰 일부 지역임에도 불구하고 화장실 시설의 양적인 공급이 매우 부족하여 '화장실 사막'으로 분류되는 지역들이 식별되었습니다. 이러한 지역들은 전반적인 시민들의 화장실 이용에 불편을 야기하며, 시급한 시설 확충이 필요함을 시사합니다.

세 번째로, 시간대별 수요 변동에 따른 불편 심화 현상이 확인되었습니다. 출퇴근 시간과 같이 유동 인구가 집중되는 특정 시간대에 가중치를 적용한 결과, 여성 화장실의 체감 수급 부족이 더욱 심화되는 경향을 보였습니다. 이는 정적인 공급량 분석을 넘어 동적인 수요 변화에 대한 고려가 필요함을 시사합니다.

이러한 인사이트들을 바탕으로 다음과 같은 정책 도입이 필요할 것으로 생각됩니다.

첫째, 분석에서 명확히 드러난 여성 화장실의 상대적 부족을 해소하기 위해 여성 화장실 칸수 확충을 최우선적으로 고려해야 합니다. 신규 공중화장실 설치 시 여성 화장실 칸수 비율을 적극적으로 상향 조정하고, 기존 시설의 리모델링 시에도 여성 칸수 확충을 우선적으로 추진할 필요가 있습니다.

둘째, 인구 규모 대비 화장실 수급 비율이 현저히 낮은 '화장실 사막'으로 식별된 지역에 대한 공중화장실 신설 및 증설 투자를 우선적으로 시행하여 지역 간 수급 불균형을 해소해야 합니다. 마지막으로, 출퇴근 시간 등 특정 시간대에 화장실 이용 불편이 심화되는 지역의 경우, 스마트 화장실 시스템 도입(이용 가능 칸수 안내), 이동식 화장실 배치 고려, 인근 상업 시설과의 협력 방안 모색 등 피크 타임 수요를 분산하거나 임시적으로 공급을 늘릴 수 있는 방안을 적극적으로 검토해야 합니다.

## 6.기술 스택

-주요 프로그래밍 언어: Python

-데이터 수집 및 전처리:

Pandas: 데이터프레임 조작, 정제, 통합 등 핵심적인 데이터 핸들링에 활용

Requests: 카카오 REST API 연동을 통한 주소-행정동 매핑

Tqdm: API 호출 배치 처리 시 진행 상황 시각화

python-dotenv: 카카오 API 키 등 민감 정보 보안 관리

-데이터 시각화 및 분석:

Matplotlib / Seaborn: 분석 결과를 직관적으로 이해할 수 있는 다양한 그래프 생성

## 7.향후 개선 계획

'화장실 사막' 지역, 특히 인구 밀집 지역이면서 화장실 수급이 현저히 부족한 곳에 대한 공중화장실의 신설 및 증설을 최우선으로 추진해야 합니다.

단순히 칸수만 늘리는 것을 넘어, 화장실의 접근성, 안전성(비상벨, CCTV 설치 유무), 위생 상태, 편의 시설(기저귀 교환대 등) 등 질적인 측면도 함께 고려하여 시민 만족도를 높이는 방향으로 개선이 이루어져야 합니다.

## 8.분석의 한계점 개선:

### 1. 개인/상업 시설 화장실 활용 미반영

본 분석은 공공데이터포털에서 제공하는 '공중화장실' 데이터만을 활용했습니다. 그러나 서울시에는 백화점, 쇼핑몰, 음식점 등 개인 소유의 상업 시설 내 화장실이 매우 많고 시민들이 일상생활에서 이러한 시설을 빈번하게 이용합니다. 이러한 사설 화장실의 이용 현황이 분석에 반영되지 못했다는 점은 실제 화장실 수급 상황을 과소평가할 수 있는 중요한 한계점입니다.

### 2. 실시간/동적 이용 패턴의 정교화

현재는 출퇴근 시간 등 특정 시간대에 가중치를 부여하는 방식으로 수요를 추정했으나, 실제 화장실 이용 패턴은 시간대별 유동 인구, 대중교통 이용량, 지역 축제 등 훨씬 더

복합적이고 동적으로 변화합니다. 향후에는 통신사 유동 인구 데이터, 대중교통 승하차 데이터 등 보다 정교한 실시간 또는 동적 데이터를 연계하여 화장실 이용 수요를 예측하고, 이에 따른 최적의 화장실 배치 및 운영 전략을 수립할 수 있도록 분석 모델을 고도화해야 합니다.

### 3.사용자 경험 데이터 부족

화장실 내부에서의 실제 이용 행태(평균 이용 시간, 대기열 길이 등)나 사용자 설문 조사 데이터는 현실적인 제약으로 인해 본 분석에 직접 포함되지 못했습니다. 향후 가능하다면 소규모 표본 조사 등을 활용하여 실제 이용 행태 데이터를 보완한다면 분석의 깊이를 더욱 심화할 수 있을 것입니다.