

공공데이터 분석 기반 여성안전 시설지수 시각화 : 서울시 중심으로

김형균

국민대학교 소프트웨어학부 교수

Visualization of women's safety facility index based on public data analysis: Focusing on Seoul

Hyeong-Gyun Kim

Professor, Dept of Software, Kookmin University

요 약 본 논문에서는 서울시와 관련한 공공데이터를 이용해 여성 안전시설지수를 작성하고 시각화하였다. CPTED, 지역여성 안전지수 관련 8가지 데이터와 여성 5대 범죄피해자 데이터를 수집·분석하여 여성 안전시설지수를 작성하였다. 여성 안전시설지수 요소와 여성 범죄 피해자 수 간의 상관분석 결과 신뢰도 95% 수준에서 유의하다고 결과가 나온 'CCTV', '가로등', '여성안심지킴이집' 3가지 데이터를 주요 요소로 선정하였다. 여성 안전시설 지수 산출을 위한 주요 요소간의 상관계수를 이용해 가중치를 부여해서 구별 여성 안전시설 지수를 산출하고 파이썬의 Follium 라이브러리를 이용해 시각화하였다.

주제어 : 공공데이터, CPTED, 여성안전, 여성범죄, 시각화

Abstract In this paper, an index of women's safety facilities was created and visualized using public data related to Seoul. CPTED, the women's safety facilities index was created by collecting and analyzing eight data related to the local women's safety index and five major crime victims of women. As a result of the correlation analysis between the factors of the female safety facility index and the number of female crime victims, three data were selected as the main factors, "CCTV," "street lamps," and "female security guardians", which were found to be meaningful at the 95% level of reliability. The distinction women's safety facility index was calculated by weighting the correlation coefficient between the main factors for calculating the women's safety facility index, and visualized using Python's Follium library.

Key Words : Public data, CPTED, women's safety, women's crimes, visualization

1. 서론

'2019년 서울시 성(性)인지 통계'에 따르면 서울 거주 여성 중 50.3%는 우리 사회가 불안하다고 느낀다고 했다 [1]. 이러한 여성들의 불안감은 서울지방 경찰청이 집계한 '2018년 강력범죄 유형별 여성피해자 비율'을 보면 살

인·강도·방화·성폭력 등 강력범죄 피해자의 89.3%가 여성인 것으로 확인할 수 있었다[2, 3].

CPTED(Crime Prevention Through Environmental Design)는 '환경설계를 통한 범죄예방'의 약자로 범죄피해를 당할 잠재적 피해자를 보호하는데 목적이 있다[4]. 또한, 범죄의 구성요소인 피해자, 범죄자, 장소들 간의 상관

*Corresponding Author : Hyeong-Gyun Kim(multikim@kookmin.ac.kr)

Received January 8, 2021

Accepted April 20, 2021

Revised February 1, 2021

Published April 28, 2021

성을 분석하여 적절한 건축설계나 도시계획 등 주변 환경의 설계를 통해 잠재적인 범죄자의 범행기회를 제거하고자 하는데 목적이 있다[5, 6].

본 논문에서는 서울시와 관련한 공공데이터를 이용해 CPTED, 지역여성 안전지수 관련 8가지 데이터와 구별 여성 5대 범죄피해자 데이터를 수집·분석하여 여성 안전 시설지수를 작성하고 시각화하였다.

2. CPTED와 여성안전지수

본 논문에서는 서울시 여성 안전시설지수 산출을 위하여 Table 1과 같이 CPTED, 지역여성 안전지수와 관련된 8가지 데이터를 선별하고 경찰청에서 제공하는 서울시 구별 여성 5대 범죄 피해자 데이터를 수집하고자 한다. 각 데이터의 출처는 Table 2와 같다.

Table 1. CPTED and women's safety index analysis data

Area classification	Data name
CPTED	City market statistics in Seoul
	Statistics of road facilities in Seoul
	Statistics on housing status by year of construction in Seoul
Local Women's Safety Index	CCTV installation status by year in Seoul
	Statistics of National Basic Livelihood Security Recipients in Seoul
	Life Safety Statistics in Seoul
	Women's Relief Scout Usage Status in Seoul
	Women's Safety Guard House Information in Seoul
The five major crimes	Age and gender status of the five major crime victims by government office in Seoul

Table 2. Source of analysis data

Data name	Source of data
City market statistics in Seoul	http://data.seoul.go.kr/dataList/datasetView.do?infld=569&srvType=S&serviceKind=2
Statistics of road facilities in Seoul	https://data.seoul.go.kr/dataList/datasetView.do?infld=261&srvType=S&serviceKind=2
Statistics on housing status by year of construction in Seoul	https://data.seoul.go.kr/dataList/datasetView.do?infld=231&srvType=S&serviceKind=2
CCTV installation status by year in Seoul	https://data.seoul.go.kr/dataList/datasetView.do?infld=OA-2734&srvType=S&serviceKind=1¤tPageNo=1
Statistics of National Basic Livelihood Security Recipients in Seoul	http://data.seoul.go.kr/dataList/datasetView.do?infld=1&srvType=S&serviceKind=2

Life Safety Statistics in Seoul	https://data.seoul.go.kr/dataList/datasetView.do?infld=10406&srvType=S&serviceKind=2
Women's Relief Scout Usage Status in Seoul	https://data.seoul.go.kr/dataList/datasetView.do?infld=OA-14665&srvType=S&serviceKind=1¤tPageNo=1
Women's Safety Guard House Information in Seoul	https://data.seoul.go.kr/dataList/datasetView.do?infld=OA-15178&srvType=S&serviceKind=1¤tPageNo=1
Age and gender status of the five major crime victims by government office in Seoul	http://www.police.go.kr/portal/main/contents.do?menuNo=200529

3. 여성안전 시설지수 분석 및 시각화

3.1 데이터 분석 방법

앞서 제시한 CPTED, 지역여성 안전지수와 관련된 8가지 데이터와 구별 여성 5대 범죄 피해자 데이터를 수집한 후 다음과 같은 과정을 거쳐 데이터를 분석하게 된다.

첫째, 수집된 데이터를 각각 구별로 이용할 데이터 열만 남긴 후 오름차순으로 정렬한다.

둘째, 전처리 과정을 거친 데이터들을 통합하여 하나의 엑셀 파일로 추출한다.

셋째, 위의 결과로 통합된 데이터 엑셀 파일을 SPSS를 통해 '여성 범죄 피해자' 데이터와 나머지 8가지 데이터 간의 유의미성을 분석한다.

넷째, 유의미한 데이터를 가지고 신뢰도 분석을 한다. 표준화된 크론바흐 알파를 기준으로 가중치를 부여하여 '여성 안전시설지수'를 산출한다.

3.2 Folium을 이용한 지도 시각화

Folium은 'Open Street Map'과 같은 지도데이터에 'Leaflet.js'를 이용하여 위치정보를 시각화하기 위한 라이브러리이다[7]. 기본적으로 'GeoJSON' 형식 또는 'topoJSON' 형식으로 데이터를 지정하면, 오버레이를 통해 마커의 형태로 위치 정보를 지도상에 표현할 수 있다[8].

JSON은 JavaScript Object Notation의 약자로 JavaScript 문법에 영향을 받아 개발된 Lightweight한 데이터 표현 방식이다[9]. JSON은 데이터를 교환하는 한 포맷으로서 그 단순함과 유연함 때문에 널리 사용되고 있는데, 특히 웹 브라우저와 웹서버 사이에 데이터를 교환하는데 많이 사용된다[10].

Python은 기본적으로 JSON 표준 라이브러리(json)를 제공한다[11]. "import json" 을 사용하여 JSON 라

이브러리를 사용할 수 있고, Python 타입의 Object를 JSON 문자열로 변경할 수 있으며 JSON 문자열을 다시 Python 타입으로 변환할 수 있다[12].

'json', 'folium'을 импорт(import)하고 JSON 파일의 경로를 'geo_path'에 저장한 후 json.load함수를 이용해 geo_str 객체를 생성한다[13]. JSON 파일은 통계청 데이터를 기반으로 하여 작성한 JSON 형식 행정구역 데이터를 사용하였다[14].

```
> import json
> import folium
> geo_path = 'data/skorea_municipalities_geo_simpl
e.json'
> geo_str = json.load(open(geo_path, encoding='utf-8'))
```

3.3 서울시 5대 범죄 여성 피해 현황 시각화

서울지역 여성 안전시설지수 분석에 앞서 구별 5대 범죄 여성 피해 현황을 분석해 보았다. 정보공개포털을 활용하여 서울시 관서별 5대 범죄 피해자 성별 데이터를 요청해 제공받았다. 제공받은 데이터에서 관서별 남녀 피해자 데이터만 엑셀 파일로 옮긴 후, 관서별 데이터를 구별 데이터로 통합해 seoul_crime_num 데이터 프레임에 저장한 자료는 Table 3과 같다.

folium.Map.Map() 메소드에 중심 좌표값을 서울특별시청의 위도, 경도로 지정하고 zoom_start 속성은 11로, tiles 속성은 'Stamen Toner' 로 지정해 맵을 생성하였다[15].

```
> map = folium.Map(location=[37.566345, 126.977893], zoom_start=11, tiles='Stamen Toner')
geo_data 속성에 JSON파일을 로드한 geo_str를 지정하고, data 속성에 구별로 5대 범죄 개수를 카운트해 놓은 seoul_crime_num 데이터프레임을 지정한다.
fill_color 속성은 'YlGnBu'으로 지정하고 맵의 choropleth함수를 설정한다[16].
> map.choropleth(geo_data=geo_str, data=seoul_crime_num,
columns=['gu','count'],fill_color='YlGnBu',key_on= 'feature.id')
> map
```

이와 같은 명령에 의해 생성된 서울시의 구별 5대 범죄 여성 피해현황을 Fig. 1과 같이 시각화하였다.

Table 3. Current status of victims of female crime in Seoul

Gu name	Victim of female crime
Gangnam-gu	3019
Gangdong-gu	1499
Gangbuk-gu	1328
Gangseo-gu	2052
Gwanak-gu	2036
Gwangjin-gu	1625
Guro-gu	1704
Geumcheon-gu	1211
Nowon-gu	1582
Dobong-gu	719
Dongdaemun-gu	1412
Dongjak-gu	1299
Mapo-gu	2018
Seodaemun-gu	1145
Seocho-gu	1767
Seongdong-gu	871
Seongbuk-gu	1192
Songpa-gu	2027
Yangcheon-gu	1371
Yeongdeungpo-gu	2114
Yongsan-gu	1317
Eunpyeong-gu	1504
Jongno-gu	1278
Jung-gu	1514
Jungnang-gu	1628

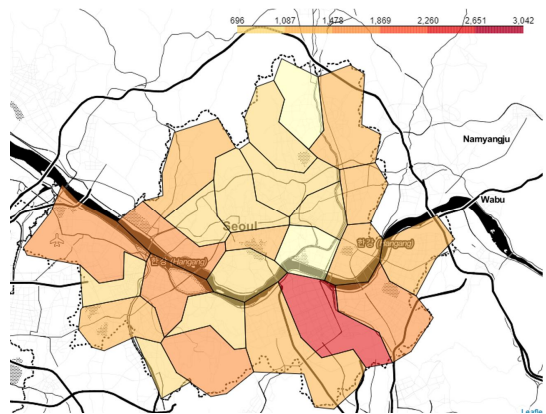


Fig. 1. current status of female victims in the 5 major crimes in Seoul

3.4 데이터 요소 선정

여성 안전시설지수 산출을 위한 데이터 요소 선정은 다음과 같이 진행하였다.

		Street lamp	CCTV	Urbanization area ratio	Severity of garbage disposal	Reconstruction (before 1999)	Basic Livelihood Security Recipient	Women's Relief Scout Usage Status	Women's Safe Keeper House	Victim of female crime
Street lamp	Pearson Care	1	.323	-.067	-.094	.347	-.139	-.016	.752	.664
	Significance probability		.115	.749	.653	.090	.508	.938	.000	.000
	N	25	25	25	25	25	25	25	25	25
CCTV	Pearson Care	.323	1	.111	.065	.120	-.097	-.029	.538	.429
	Significance probability	.115		.597	.758	.568	.645	.891	.006	.032
	N	25	25	25	25	25	25	25	25	25
Urbanization area ratio	Pearson Care	-.067	.111	1	.159	-.459	-.453	-.053	-.111	-.181
	Significance probability	.749	.597		.448	.021	.023	.800	.599	.387
	N	25	25	25	25	25	25	25	25	25
Severity of garbage disposal	Pearson Care	-.094	.065	.159	1	-.238	-.164	-.450	-.229	-.088
	Significance probability	.653	.758	.448		.253	.434	.024	.270	.675
	N	25	25	25	25	25	25	25	25	25
Reconstruction (before 1999)	Pearson Care	.347	.120	-.459	-.238	1	.700	-.117	.305	.360
	Significance probability	.090	.568	.021	.253		.000	.579	.138	.077
	N	25	25	25	25	25	25	25	25	25
Basic Livelihood Security Recipient	Pearson Care	-.139	-.097	-.453	-.164	.700	1	-.124	.002	.151
	Significance probability	.508	.645	.023	.434	.000		.555	.994	.471
	N	25	25	25	25	25	25	25	25	25
Women's Relief Scout Usage Status	Pearson Care	-.016	-.029	-.053	-.450	-.117	-.124	1	-.074	-.204
	Significance probability	.938	.891	.800	.024	.579	.555		.726	.328
	N	25	25	25	25	25	25	25	25	25
Women's Safe Keeper House	Pearson Care	.752	.538	-.111	-.229	.305	.002	-.074	1	.772
	Significance probability	.000	.006	.599	.270	.138	.994	.726		.000
	N	25	25	25	25	25	25	25	25	25
Victim of female crime	Pearson Care	.664	.429	-.181	-.088	.360	.151	-.204	.772	1
	Significance probability	.000	.032	.387	.675	.077	.471	.328	.000	
	N	25	25	25	25	25	25	25	25	25

Fig. 2. Results of correlation analysis between women's safety facility index and women's crimes

Table 4. Basic Score for Women's Safety Facility Index

Gu name	Street lamp	CCTV	Women's SafeKeeper House
Gangnam-gu	23617->25	6799->25	259->25
Gangdong-gu	13131->16	2133->8	83->21
Gangbuk-gu	7396->1	1352->2	66->16
Gangseo-gu	11134->12	1718->4	64->14
Gwanak-gu	8145->4	4606->24	74->19
Gwangjin-gu	8074->3	1755->6	57->7
Guro-gu	12335->15	3924->23	62->13
Geumcheon-gu	8934->5	2181->9	27->1
Nowon-gu	15489->22	2252->11	62->13
Dobong-gu	11437->2	2819->1	57->6
Dongdaemun-gu	13472->18	2222->10	89->22
Dongjak-gu	9124->7	2561->13	50->6
Mapo-gu	13214->17	2491->12	83->21
Seodaemun-gu	9731->10	3865->22	37->3
Seocho-gu	19467->23	2669->15	108->23
Seongdong-gu	9422->8	3650->21	41->4
Seongbuk-gu	8941->6	3174->17	74->19
Songpa-gu	20943->24	1719->5	112->24
Yangcheon-gu	11977->13	3570->19	35->2
Yeongdeungpo-gu	15324->21	2628->14	69->17
Yongsan-gu	12132->14	3400->18	60->10
Eunpyeong-gu	10906->11	3580->20	58->9
Jongno-gu	14953->20	2751->16	62->13
Jung-gu	14379->19	1801->7	66->16
Jungnang-gu	9481->9	1505->3	58->9

첫째, 회귀분석을 통해, 8가지 데이터 요소들과 '여성

범죄 피해자 수' 데이터 간의 유의미함을 확인

둘째, 상관분석을 통해, 8가지 데이터 요소들과 '여성 범죄 피해자 수' 데이터 간의 연관성을 확인

셋째, 신뢰도 95% 수준에서 유의한 결과를 가지고 여성 안전시설지수 산출

Fig. 2는 여성 안전시설지수 요소와 여성 범죄 피해자 수 간의 상관분석 결과를 보여주고 있다. 신뢰도 95% 수준에서 유의미 하다고 결과가 나온 'CCTV', '가로등', '여성안심지킴이집' 3가지 데이터를 가지고 여성 안전시설 지수를 산출하였다.

여성 안전시설 지수 산출을 위한 기본점수를 부여하기 위해 상관분석 결과에서 유의미한 결과를 보인 'CCTV', '가로등', '여성안심지킴이집' 3가지 데이터의 설치 숫자를 기준으로 구별로 1위에서 25위에 25점에서 1점을 부여하였다.(설치 수가 많을수록 점수가 높음).

Table 4는 이렇게 산출된 여성 안전시설 지수 기본점수를 구별로 보여주고 있다.

3.5 여성 안전시설 지수 산출 및 시각화

Table 5는 여성 안전시설 지수 산출을 위한 주요 요소간의 상관행렬을 보여주고 있다. 이에 근거하여 각 항목간의 가중치를 다음과 같이 부여하여 구별 여성 안전 시설 지수를 산출하였다.

$$\begin{aligned}
 &[\text{OO구의 여성안전시설지수}] \\
 &= [\text{OO구의 여성안심지킴이집 점수}] * 1.772 \\
 &+ [\text{OO구의 가로등 점수}] * 1.664 \\
 &+ [\text{OO구의 CCTV 점수}] * 1.429
 \end{aligned}$$

Table 5. Correlation matrix between major elements

Correlation coefficient	Women's Safe Keeper House	Street lamp	CCTV	Victim of female crime
Women's safe KeeperHouse	1.000	0.752	0.538	0.772
Street lamp	0.752	1.000	0.323	0.664
CCTV	0.538	0.323	1.000	0.429
Victim of femalecrime	0.772	0.664	0.429	1.000

Table 6은 이렇게 산출된 구별 여성 안전시설 지수를 보여주고 있다.

Table 6. Index of Women's Safety Facilities in Seoul

Gu name	Women's Safety Facility Index Safe Keeper House
Gangnam-gu	122
Gangdong-gu	75
Gangbuk-gu	33
Gangseo-gu	20
Gwanak-gu	75
Gwangjin-gu	26
Guro-gu	81
Geumcheon-gu	23
Nowon-gu	75
Dobong-gu	15
Dongdaemun-gu	83
Dongjak-gu	41
Mapo-gu	83
Seodaemun-gu	53
Seocho-gu	100
Seongdong-gu	50
Seongbuk-gu	68
Songpa-gu	90
Yangcheon-gu	52
Yeongdeungpo-gu	85
Yongsan-gu	67
Eunpyeong-gu	63
Jongno-gu	79
Jung-gu	70
Jungnang-gu	35

Fig. 3은 서울시의 구별 여성 안전시설 지수를 시각화한 결과를 보여주고 있다.

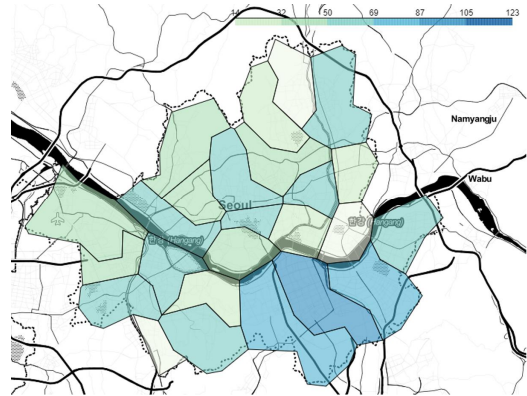


Fig. 3. Visualization of the index of women's safety facilities in Seoul

4. 결론

본 논문에서는 서울시와 관련한 공공데이터를 이용해 CPTED, 지역여성 안전지수 관련 8가지 데이터와 구별 여성 5대 범죄피해자 데이터를 수집·분석하여 여성 안전 시설지수를 작성하고 시각화하였다.

여성 안전시설지수 요소와 여성 범죄 피해자 수 간의 상관분석 결과 신뢰도 95% 수준에서 유의미하다고 결과가 나온 'CCTV', '가로등', '여성안심지킴이집' 3가지 데이터를 주요 요소로 선정하였다. 여성 안전시설 지수 산출을 위한 주요 요소간의 상관계수를 이용해 가중치를 부여해서 구별 여성 안전시설 지수를 산출하고 파이썬의 Follium 라이브러리를 이용해 시각화하였다.

향후 이렇게 산출된 여성 안전시설 지수와 시각화 자료를 활용해 서울시의 여성안전 관련 정책 시행 시 우선 적용 대상지역을 선택하는데 도움이 될 것으로 생각한다.

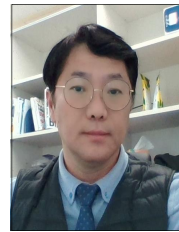
REFERENCES

- [1] W. S. Park. (2019). *Gender statistics in Seoul in 2019*. <https://opengov.seoul.go.kr>
- [2] S. Y. Oh. (2017). Study of Causes and Solutions of Violent Crimes against Second-class Citizens - Focus on Women's Violent Crimes -. *Korean Police Studies Review*, 16(3), 225-250.
- [3] J. H. Ki. (2015). A Study of Impact of Urban Population Characteristics on Violent Crimes. *Journal of the Korean Regional Development Association*, 27(1), 107-124.

- [4] J. H. Ryu. (2020). CPTED and Fear of Crime. *Journal of Korean Public Police and Security Studies*, 17(1), 15-32.
- [5] J. H. Jang. (2018). Study on Effectiveness of CPTED for preventing crimes : Based on panel data analysis of 25 districts in Seoul. *Korean Police Studies Review*, 17(4), 267-292.
- [6] M. H. Cho. (2013). "A plan to secure school safety by applying CPTED". Master's Thesis. Yongin University Graduate School.
- [7] M. Lee. (2018). *Map Visualization-Folium*. https://ericnjennifer.github.io/python_visualization/2018/01/21/PythonVisualization_Chapt6.html
- [8] M. G. Kim, M. H. Lee. (2016). A Visualization Method of Spatial Information based on Web Map Service. *Journal of Digital Convergence*, 14(2), 209-216.
- [9] Psychoria. (2019). *JSON data format handling in Python*. <https://psychoria.tistory.com/703>
- [10] S. W. Park. (2015). A Study on the Visualization of Images in the Mind through Texts and Maps - Based on My Works. *Journal of Korean Society of Media & Arts*, 13(1), 23-34.
- [11] J. A. Kim, M. G. Kim. (2019). Effect of data visualization education with using Python on computational thinking of six grade in elementary school. *Journal of The Korean Association of Information Education*, 23(3), 197-206.
- [12] D. H. Lee. (2015). An Alternative Approach for Implementing Interactive Media Contents - a Case Study of Teaching Computer Game Programming using Python. *Journal of Korean Society of Media & Arts*, 13(1), 145-156.
- [13] Y. S. Lee. (2018). Python-based Software Education Model for Non-Computer Majors. *Journal of the Korea Convergence Society*, 9(3), 73-78.
- [14] J. B. Lee. (2019). Analysis of Tensor Processing Unit and Simulation Using Python. *The Journal of the Institute of Internet, Broadcasting and Communication*, 19(3), 165-171.
- [15] K. M. Kim, H. S. Kim. (2014). A Case Study on Necessity of Computer Programming for Interdisciplinary Education. *Journal of Digital Convergence*, 12(11), 339-348.
- [16] W. S. Moon. (2018). Analysis of error data generated by prospective teachers in programming learning. *Journal of The Korean Association of Information Education*, 22(2), 205-212.

김형균(Hyeong-Gyun Kim)

[충남대]



- 2004년 2월 : 조선대학교 컴퓨터공학과(공학박사)
- 2004년 3월 ~ 2007년 8월 : 동강대학교 컴퓨터정보과 초빙교수
- 2012년 3월 ~ 2016년 8월 : 조선대학교 컴퓨터공학과 객원교수
- 2017년 3월 ~ 2019년 8월 : 서울여자대학교 SW교육혁신센터 교수
- 2019년 9월 ~ 현재 : 국민대학교 SW학부 교수
- 관심분야 : IOT 시스템 통합, 데이터분석
- E-Mail : multikim@kookmin.ac.kr