

2020 서울시 빅데이터 캠퍼스 공모전

우리동네키움센터
최적입지 선정을 통한
돌봄 공백 해소



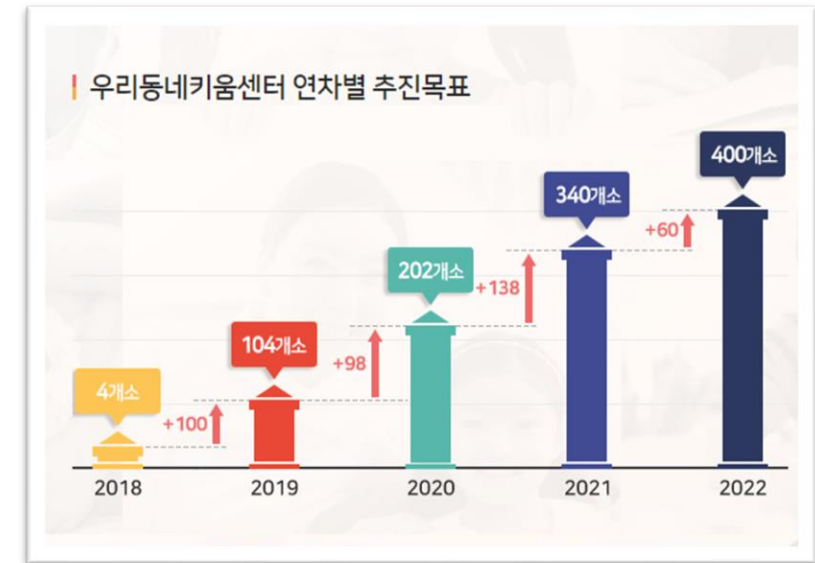
Table of contents

- 1 분석 배경**
- 2 수요산출 방법**
- 3 최적입지 추천**
- 4 결과**

2019년 기준 한국 초등학생 공적
돌봄 서비스 비율은 13.9%에 불과¹⁾



서울시는 이 위기를 해결하고자
우리동네키움센터 추가 설립 결정



그렇다면, 우리동네키움센터는 서울시 어느 장소에 추가해야 될까?

1. ‘우리동네키움센터’ 운영방향

‘우리동네키움센터’는 (이하 키움센터) 초등학생 연령의 아동을 대상으로
지역 내 방과 후 돌봄의 틈새공백을 메우는 데 그 목적이 있다. ²⁾

2. ‘우리동네키움센터’ 정체성

키움센터의 정체성은

- 1) 자격제한 및 검증 없이 모든 아동에게 선착순 개방하고,
- 2) 아동 및 학부모가 선택적으로 이용할 수 유연한 운영체계를 유지하는 것으로,
돌봄이 필요하나 대안이 없는 아동을 탄력적으로 수용하여 선행연구가 지적하는 사각지대를 메우는 목적에
적합한 모델이라 볼 수 있다. ³⁾

➡ 우리동네키움센터의 설립 목적은 돌봄의 사각지대에 놓여있는 초등학생을 돌보기 위함

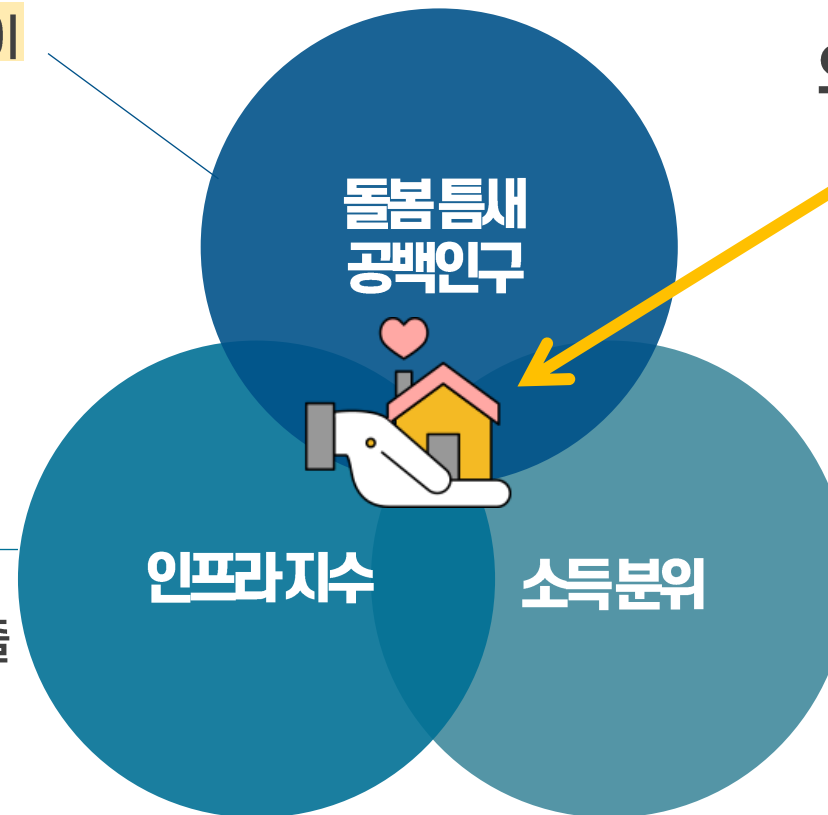
행정동 별 '수요지수' 산출 과정

우리동네키움센터 설립 취지에 맞게 **돌봄 인프라가 부족한 지역 중 돌봄의 사각지대에 놓인** 초등학생들이 많고 **대체시설(사적돌봄시설)에 보낼 여력이 적은 곳에** 우선 입지하도록 지수 산출

행정동 별 17-19시 사이
돌봄 사각지대에 놓인
초등학생 수 계산

우리동네키움센터
최적 입지

TF-IDF 알고리즘을 통해
행정동 별 돌봄 인프라 지수 산출



KNN 알고리즘을 통해
행정동 별 소득추정

TF-IDF 기반 돌봄 인프라 필요 지수

TF-IDF란? 여러 문서의 문서군이 있을 때 특정 단어의 특정 문서내에서 중요도를 나타내는 통계적인 수치. 문서를 행정동으로 치환하여 사용함.

공적 돌봄시설이란? 우리동네키움센터 / 초등돌봄교실 / 지역아동센터 / 방과후아카데미

인프라 지수

$$\text{인프라지수} \propto \frac{\text{돌봄 시설 중요도} \times \text{시설 정원 비율}}{\text{사교육 기관 중요도}}$$

Idea1. 행정동 내의 초등학생 인구의 비율이 높을수록 공적 돌봄 시설이 더 필요하다. (돌봄시설 중요도)

Idea2. 기존 돌봄 시설의 정원이 많을수록 인프라의 수요가 꾸준하다고 가정하여 비례관계에 놓았다.

Idea3. 학원과 같은 사적 인프라가 돌봄 시설의 역할인 공적 돌봄을 대체한다. (역의 상관관계)

TF-IDF 기반 돌봄 인프라 필요 지수의 식

돌봄 시설 중요도와 사교육 기관 중요도는 TF-IDF 알고리즘을 활용하여 측정

돌봄 인프라 필요 지수

$$Infra = \frac{tfidf_{center} \times P_d \times C_d}{tfidf_{academy}}$$

$tfidf_{center}$: 행정동 별 돌봄센터(*center*) 중요도 (TF-IDF)

P_d : 행정동(*d*)별 5~14세 인구(*P*)비율

C_d : 서울시 모든 센터의 정원과 해당 행정동(*d*)에 있는 모든 돌봄시설 정원의 비율

$tfidf_{academy}$: 행정동 내 학원(*academy*)의 중요도 (TF-IDF)

➡ $Infra$ 가 높을수록 해당 행정동에 돌봄의 수요가 높은 것을 의미함

TF-IDF 기반 돌봄 인프라 필요 지수의 식

$$\text{센터 TF-IDF(중요도)} = \frac{\{\max(\text{특정 행정동의 시설 수}) - \text{해당 행정동의 시설 수}\} \times \log(\text{전체 행정동의 수})}{\log(\text{시설이 있는 행정동의 수})}$$

빈도수가 0인 경우를 위해 +1 과 log scaling 으로 Smoothing 작업

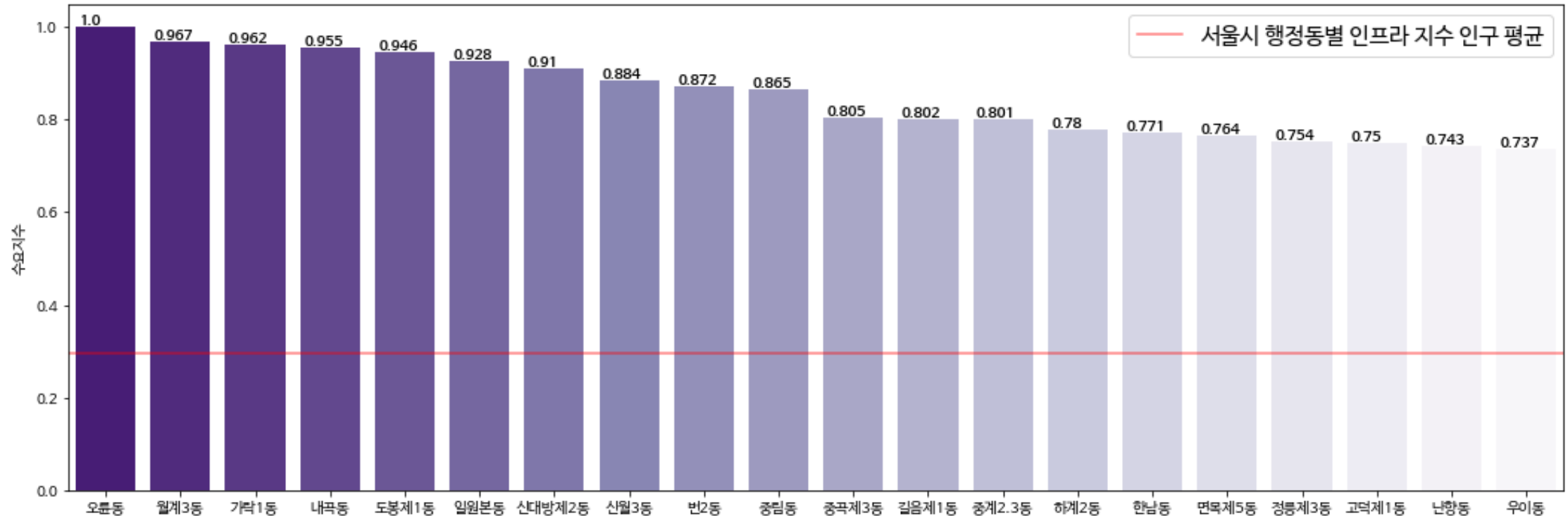
$$\text{시설 교육 기관 TF-IDF(중요도)} = \frac{(\text{특정 행정동의 학원 수} + 1) \times \text{전체 행정동의 수}}{1 + \text{학원이 있는 행정동의 수}}$$

서울시 행정동 별 돌봄 인프라 필요 지수 산출 표 Top5

순위	시군구명	행정동	인구비율 (5-14세)	센터 정원 비율	센터 TF-IDF	시설 교육 기관 TF-IDF	인프라지수
1	송파구	오륜동	0.0034	0.0014	1.1216	0.0213	2.58e-04
2	노원구	월계3동	0.0032	0.0048	0.6943	0.0425	2.49e-04
3	송파구	가락1동	0.0039	0.0016	1.0148	0.0260	2.48e-04
4	서초구	내곡동	0.0027	0.0029	0.9614	0.0307	2.46e-04
5	도봉구	도봉1동	0.0014	0.0040	0.8012	0.0189	2.44e-04

돌봄 인프라 필요 지수 행정동 별 시각화

서울시 행정동 별 돌봄 인프라 필요 지수 Top20



- TF-IDF 기반으로 계산한 인프라 필요 지수를 행정동 별로 나열
- 돌봄 인프라 필요지수가 가장 높은 동은 송파구 오륜동이다.

돌봄 틈새공백이란?

집에 도착한 초등학생이 부모의 보살핌을 받기까지 기존의 공적 돌봄 프로그램으로 메꾸기 힘든 시간적 틈새

초등학생 하교 시간 ⁴⁾

평균 13시 - 15시

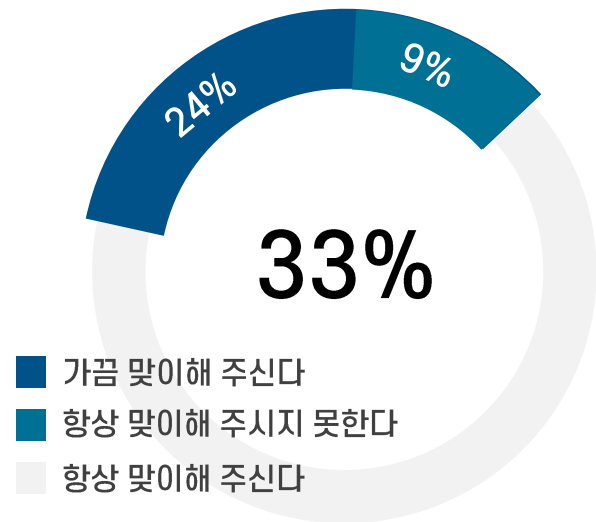
공적 돌봄 종료시간 ⁵⁾

평균 17시~18시

부모님의 퇴근시간 ⁶⁾

평균 19시

귀가 시 어른이 맞이해주지 못하는 초등학생 ⁷⁾ 수업을 마치고 가장 많이 있는 장소 (시간별) ⁸⁾



단위(%)	5시 이전	5~6시	6~7시
집	31.5	46.8	67.9
교과목 관련 학원	30.2	23.5	13.6
예체능 관련 학원	13.1	10.4	5.7
태권도 도장	5.7	6.0	4.9
동네 독서실 공부방 등	1.9	1.5	0.7
기타

집에서 19시까지 부모님을
기다려야하는
초등학생 인구 비율

$$67.9\% \times 33\% = 22.4\%$$

돌봄 틈새공백 인구 계산 수식

맞벌이 부모 슬하 초등학생 중, 7시 이전에 집에 혼자 있는 인구

→ 돌봄의 사각지대에 있는 초등학생 인구

$$NCP = K_h \times \left(\frac{P_h \times M_g \times D}{TP_h} \right) \times AR$$

NCP : 방임된 아이 인구(Neglected Child Population)

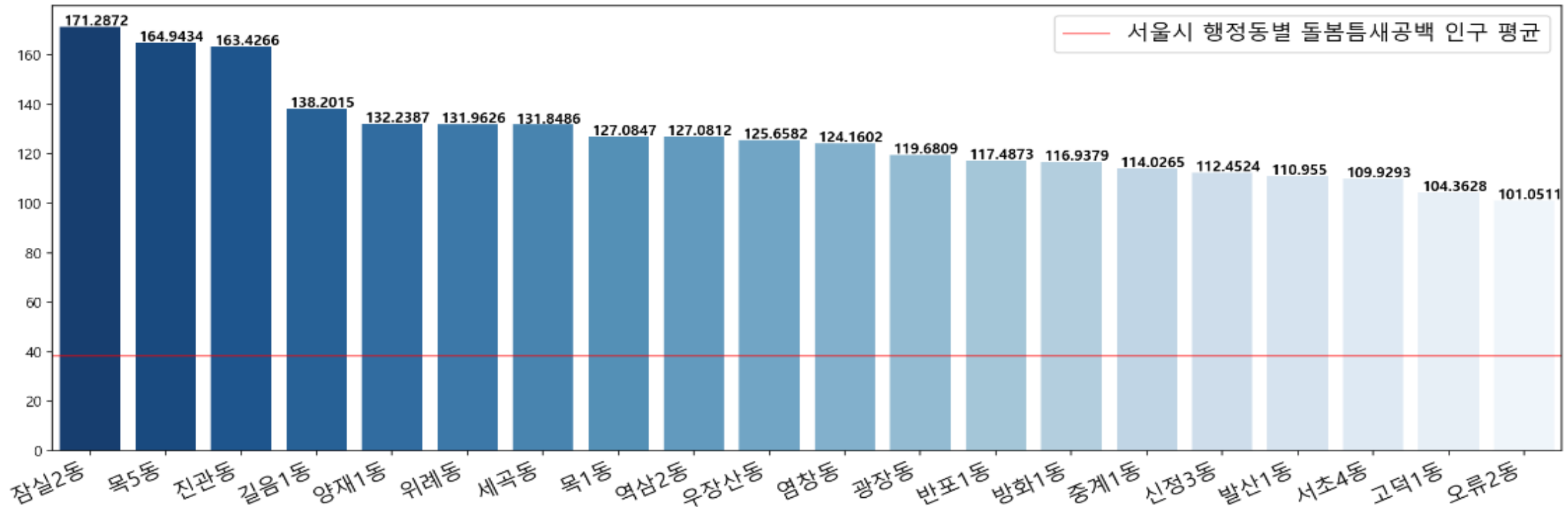
K_h : 행정동 별 5~14세 인구

AR : 19시 이전에 집에 도착하는 초등학생 비율 (67.9%) x
귀가 시, 어른이 맞이해주지 못하는 초등학생 비율 (33%)

$\left(\frac{P_h \times M_g \times D}{TP_h} \right)$: 행정동 별 3040대 기혼 맞벌이 인구비율
(P_h : 행정동 별 30~40대 인구, M_g : 구 별 30~40대 기혼 비율, D : 30~40대 맞벌이 가구 비율 (49.7%), TP_h : 행정동 별 전체 인구)

서울시 행정동 별 돌봄 틈새 공백 인구 예측

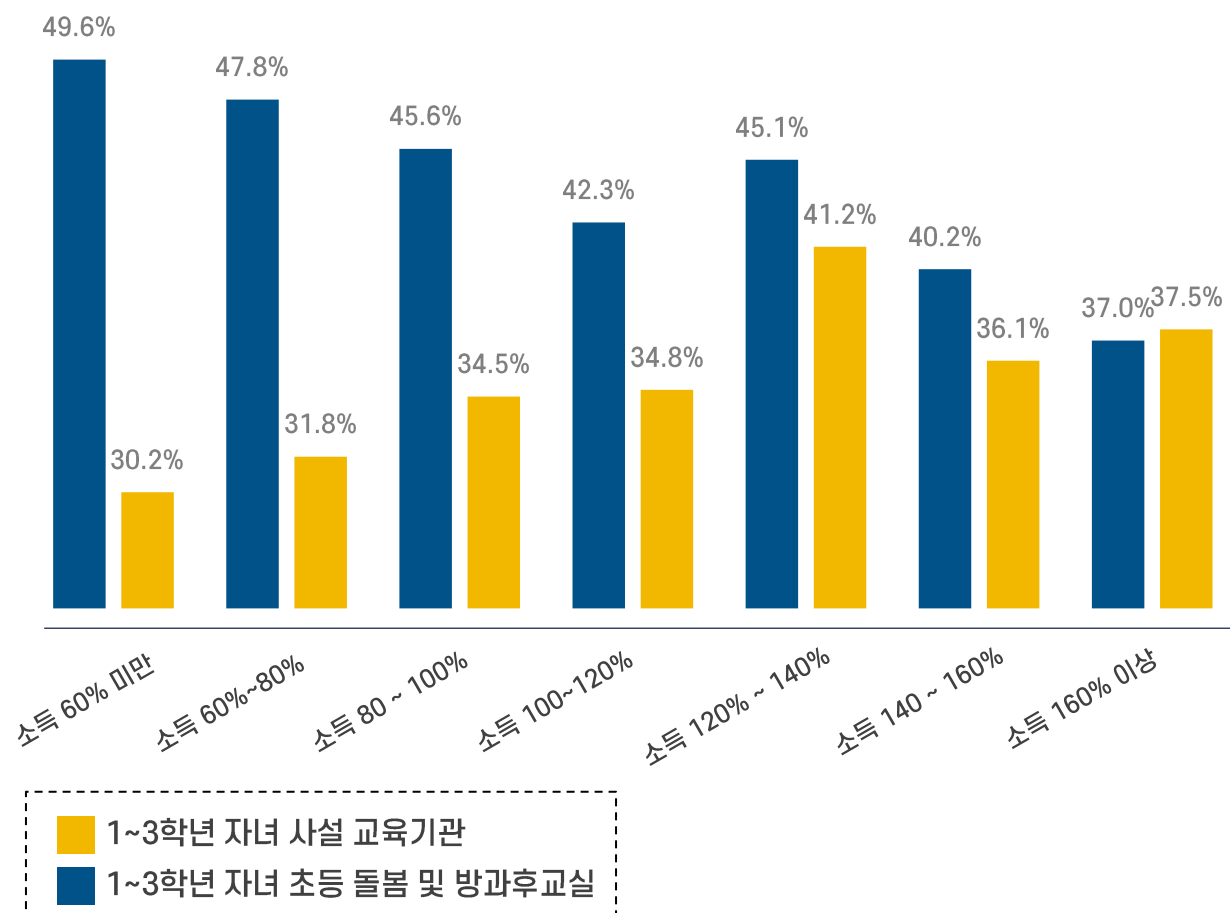
서울시 행정동 별 돌봄 틈새공백 Top20



- 수식에 따른 돌봄 틈새공백 인구를 행정동 별로 나열
- 돌봄 틈새공백 인구가 가장 높은 동은 송파구 잠실2동이다.

소득 별 희망하는 자녀 돌봄 시설

기혼여성의 1~3학년 초등학교 자녀 돌봄 희망장소 소득 별 응답률

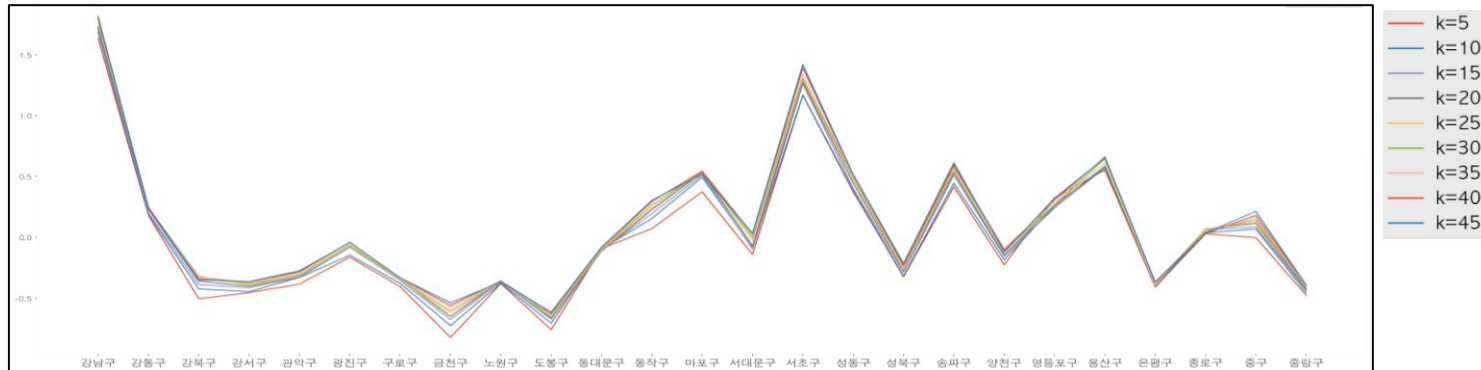


- 1 소득에 따라 공적 돌봄 시설 수요와 사설 교육기관 수요가 반비례
- 2 소득이 낮을수록 공적 돌봄 시설에 대한 수요가 높은 경향
- 3 인프라 수요 지수에 쓰인 사설교육기관 중요도가 소득과도 연관이 있음을 파악하고 지역 별 소득을 함께 고려

KNN이란? 관측 값마다의 유클리드 거리 또는 맨하튼 거리를 기반으로 한 기계학습 알고리즘.
유클리드 거리 기반으로 클러스터링 내 k개의 이웃한 훈련 데이터(위도, 경도)로 입력을 구성하며 소득 추정.

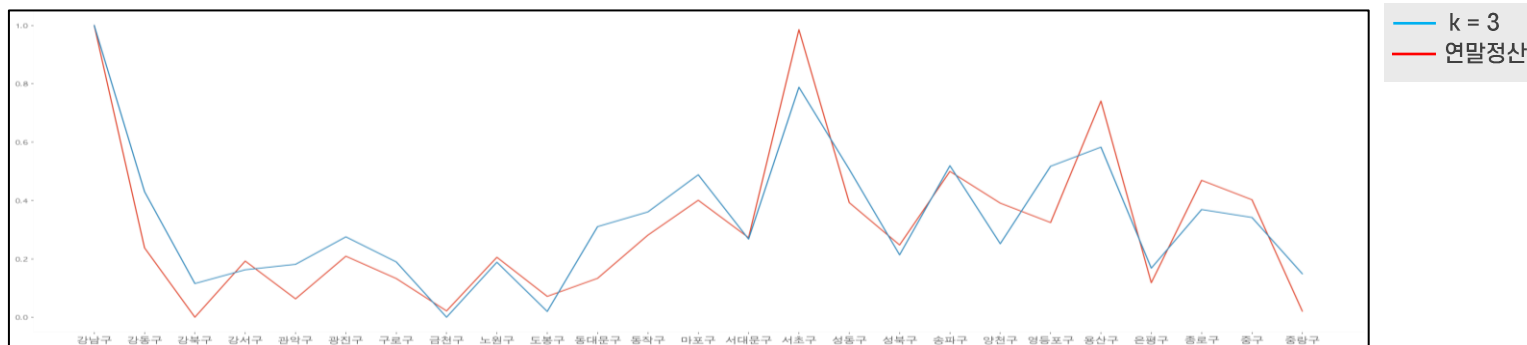
활용 데이터 행정동 별 소득 수준 데이터를 구하기 어려워
2017년 10월부터 2020년 9월까지의 아파트 매매 실거래가 데이터(약 24만개)를 활용
2019년 연말정산 근로소득 구 별 평균 데이터를 이용하여 검증

K값(이웃 데이터 개수) 별 자치구별 추정 소득 (x축: 자치구, y축: knn 추정소득(k=5~45))



K값에 따른 자치구별 소득 경향은
대체적으로 비슷하지만,
행정구 별로 k값에 따라 추정 소득 차이 존재

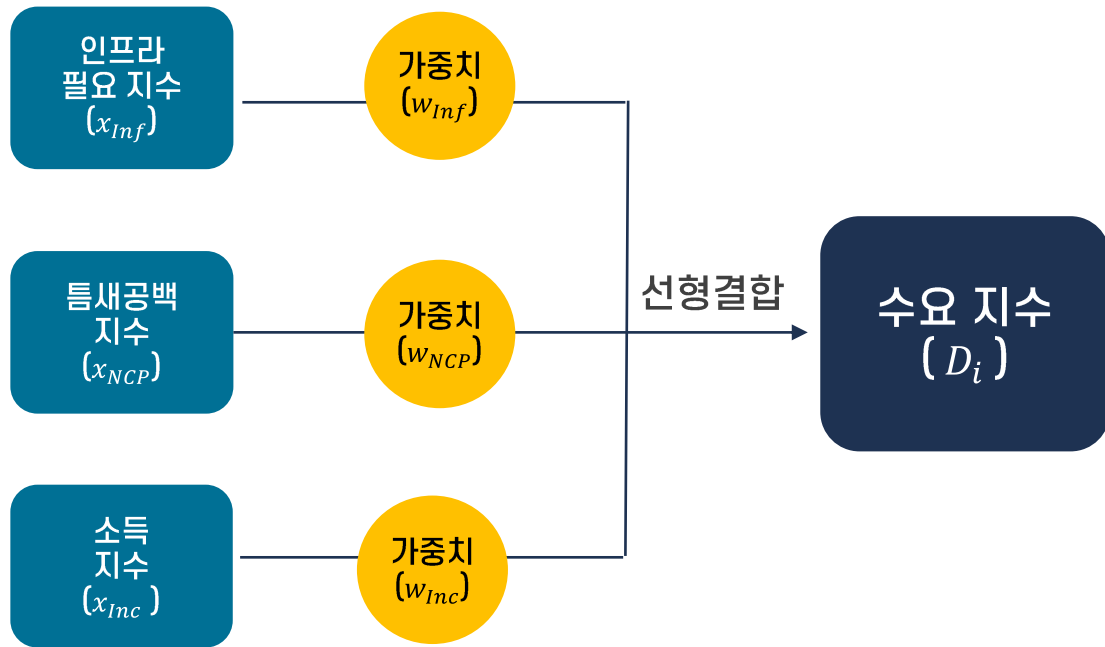
연말정산 근로소득 vs KNN 추정 소득 (x축: 자치구, y축: knn(추정소득), 연말정산(평균소득))



2019년 연말정산 과세대상근로소득
행정구 별 평균을 검증 값으로 활용.
이와 가장 유사한 추이를 보이는 k값
을 선정해 행정동 별 소득 추정

수요지수

선형결합을 통한 수요지수 산출



수요지수

$$D_i = w_{Inf}x_{Inf} + w_{NCP}x_{NCP} + w_{Inc}x_{Inc}$$

변수 (x) : 인프라 필요지수·틈새공백지수·소득지수 각각을
백분위 기준으로 1~10분위 범주화

가중치(w) : $w_{Inf}=2, w_{NCP}=2, w_{Inc}=1$

소득 가중치는 소득에 따른 수요의 왜곡을 막을 뿐, 수요
에 직접적인 영향을 미치지 않으므로 타 지수들의 가중치
보다 작은 1로 설정

수요지수(D_i) : 선형 결합 결과로 각 지역 별(i) 수요지수 산출

➡ 산출된 지역 별 수요 지수를 활용하여 돌봄 시설 최적 입지 선정

행정동 별 우리동네키움센터 필요개수

수요지수에 기반한 행정동 별 우리동네키움센터 필요개수 추정

2022년까지 건립 예정인 400개를 행정동 별로 몇 개 소를 건설해야 하는지
선형결합 결과의 중요도 비중으로 분배하여 계산

추정 우리동네키움센터
필요 개수

=

400

X

행정동 별 수요지수

∑

행정동 별 수요지수

(우리동네키움센터 건설 예정 개수)

행정동 별 우리동네키움센터 공급 미달량

구	행정동	인프라지수	돌봄 공백	틈새 인구	소득분위	수요지수	추정 센터 필요 개수	현재 센터 개수	공급 미달량
강서구	화곡1동	9	9	10	46	1.578	0	1.578	
광진구	광장동	10	10	5	45	1.544	0	1.544	
성북구	종암동	8	10	9	45	1.544	0	1.544	
...	
노원구	상계5동	5	6	9	31	1.063	2	-0.937	
노원구	상계6.7동	7	9	5	37	1.269	3	-1.731	
성동구	마장동	9	5	5	33	1.132	3	-1.868	

공급 미달량
상위 3개 동

공급 미달량
하위 3개 동

행정동 별 우리동네키움센터 필요개수와 현황비교

공급 미달량을 기준으로 5단계 범주화

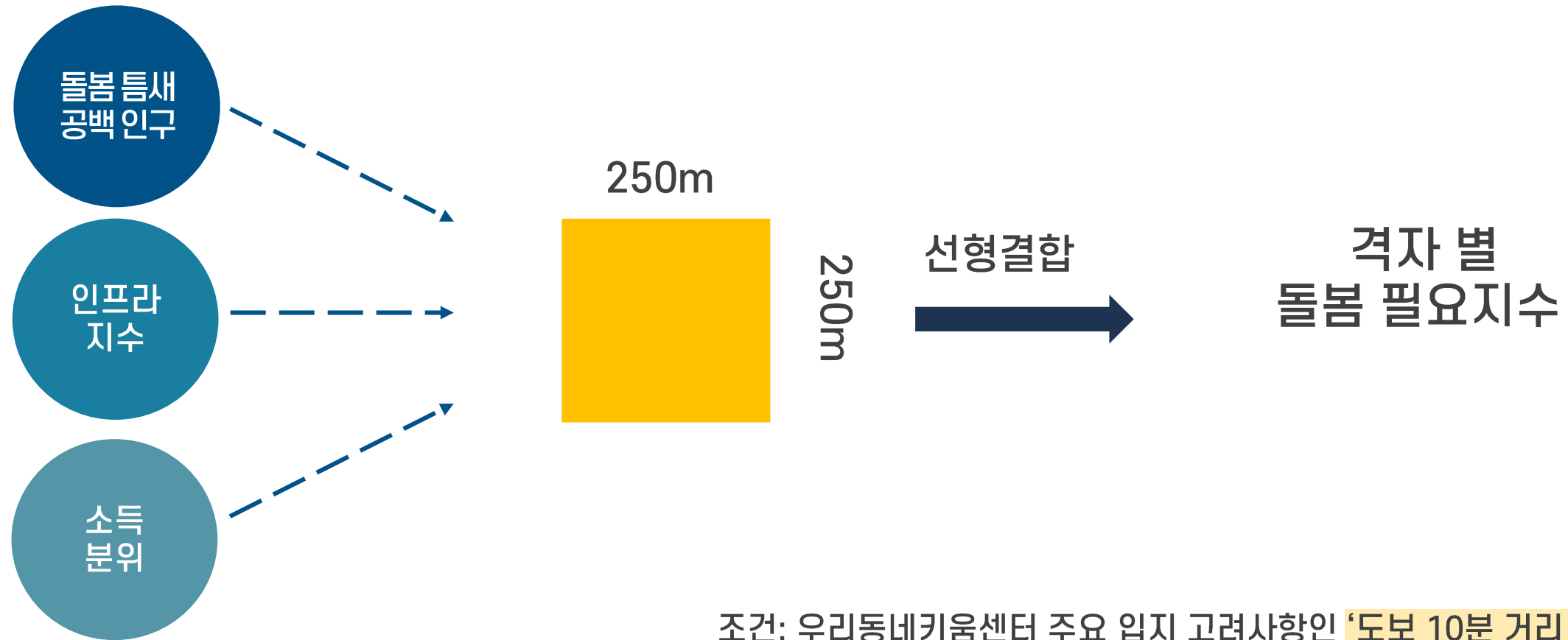
공급 미달량 그룹화 기준 각 점수의 평균

순위	행정동	공급 미달량 그룹화 기준 각 점수의 평균						
		공급 미달량 우선순위	인프라지수	돌봄 틈새 인구	소득분위	수요 지수	추정 키움센터 필요개수	현재 키움 센터 개수
1순위	화곡1동, 종암동, 광장동 외 73개							
2순위	면목2동, 돈암2동, 마천1동 외 88개							
3순위	방배2동, 서초2동, 양재2동 외 81개							
4순위	역삼1동, 대치4동, 반포4동 외 84개							
5순위	명동, 여의동, 청담동 외 83개							
		1	8.474	8.092	6.079	39.211	1.345	0.000
		2	6.198	6.022	6.121	30.560	1.048	0.000
		3	4.095	4.774	5.452	23.190	0.796	0.000
		4	4.057	4.195	5.552	22.057	0.757	0.218
		5	4.965	4.686	4.326	23.628	0.811	0.837

공급 미달량을 기준으로 행정동들을 5단계 범주화 한 결과,
1순위와 5순위 공급 미달량이 **평균 1.372개** 정도 큰 차이를 보임

행정동 별 우리동네키움센터 필요개수

각 행정동 내 250m x 250m 격자 기준으로 돌봄 필요지수를 계산

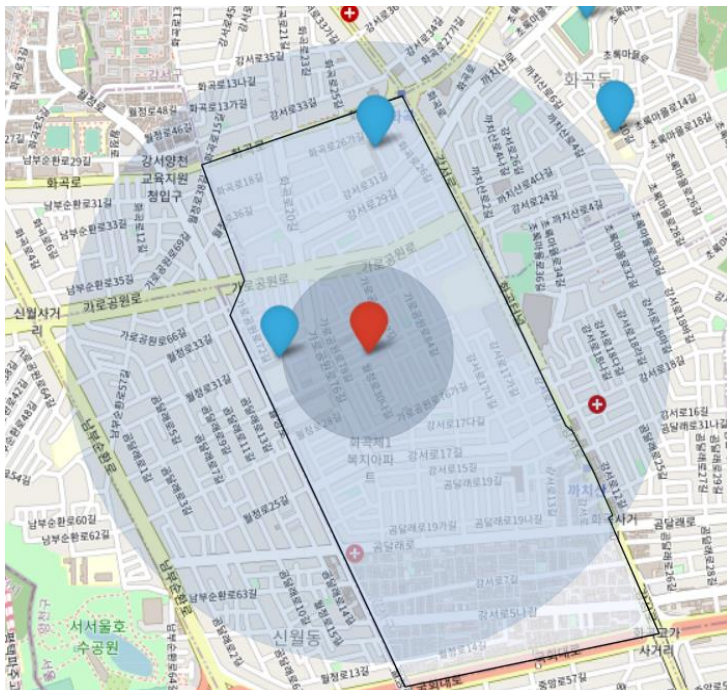


조건: 우리동네키움센터 주요 입지 고려사항인 '도보 10분 거리 내 입지'는 채정병 외(2008)⁹⁾를 참고하여 초등학교로부터 반경 881m 내 입지하도록 설정

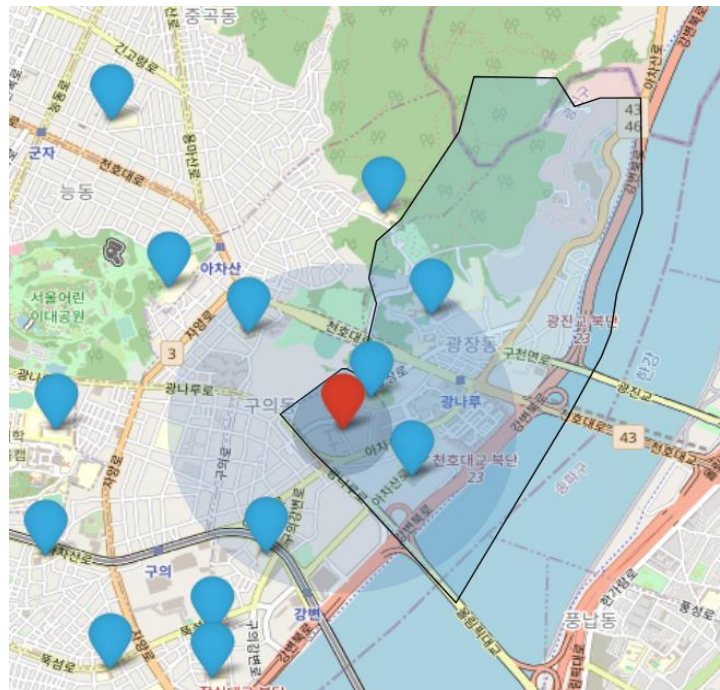
격자 별 필요지수

공급 미달량 상위 3개 동 최적입지 선정 결과

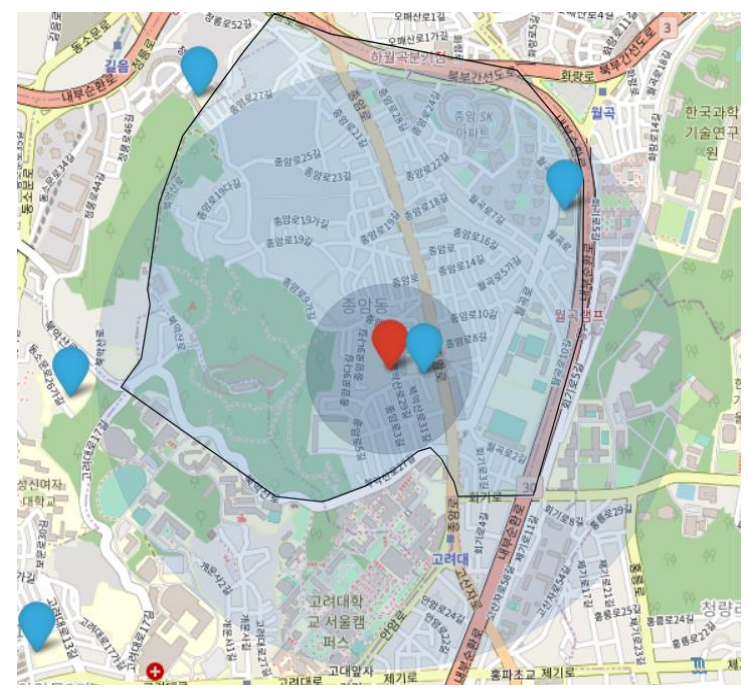
우선입지 후보 (1): 강서구 화곡1동



우선입지 후보 (2): 광진구 광장동



우선입지 후보 (3): 성북구 종암동



초등학교



추천 최적입지 기준 좌표



입지 추천(반경 250m)



초등학생 도보 10분 거리(반경 881m)

기대효과

1. 돌봄 틈새 공백을 고려한 최적 입지 분석으로 초등 돌봄 사각지대 해소
2. 적절한 입지 선정을 통해 경력 단절 문제 등 학부모들이 겪을 수 있는 사회적 문제 해결
3. 우리동네키움센터 최적 입지 분석을 통해 서울시의 합리적인 예산 운영 가능

발전방향

1. 데이터 확보에 한계(ex.행정동 별 소득)가 있어 지수를 간접 데이터로 구했으나 서울시 내 메타 데이터를 이용한다면 보다 정확하고 구체적인 입지 추천이 가능할 것으로 예상
2. 향후 일반형 · 융합형 · 거점형과 같은 우리동네키움센터의 종류를 구분하여 분석한다면 유형 별 입지 추천도 가능할 것으로 예상

행정동 별 우리동네키움센터 필요개수 추정

데이터	제공기관	기준년도
초등돌봄교실	교육부	2020
서울시학교별 정원(유초중고)	서울시교육청	2020
행정동경계	통계청	2020
서울시행정동/법정동코드	행정안전부	2020
전국 초등학교위치표준데이터	공공데이터 포털	2020. 10
행정동별 주거인구	빅데이터 캠퍼스	2020.05
서울시우리동네키움센터(Selenium 크롤링)	우리동네키움센터	2020. 10. 05
서울시아파트매매실거래가	국토교통부	2017. 10 - 2020. 09
서울특별시학원정보	서울특별시 교육청	2020. 04. 05
서울주소지연말정산신고	국세청 교육통계	2019
기혼여성의초등학생자녀돌봄희망 장소	국가통계포털	2018
서울시혼인상태별인구(15세 이상)통계	서울 열린데이터 광장	2015
서울시예술학원	우리마을 가게상권분석	2020. 06. 30
서울시세대원수별세대수(동별)통계	서울 열린데이터 광장	2017. 12
귀가시부모님등 어른의맞이횟수	국가통계포털	2018
2018한국의워킹맘보고서	KB금융지수 경영연구소	2018
맞벌이가구비율	e나라지표	2017
은평구격자별인구	국토정보플랫폼	2020. 04
수업을마치고가장많이있는장소(시간별)	국가통계포털	2018
서울시방과후돌봄지역사회협력방안	서울연구원	2018

참고 자료

1) 방현덕, "박원순, 초등자녀 맡아주는 '키움센터' 2022년까지 400개 신설", 연합뉴스, 2019.03.06, <https://www.yna.co.kr/view/AKR20190306072051004>

2) & 3) 전미양, & 이혁수. (2018). 온종일 돌봄을 위한 '키움센터' 운영 매뉴얼 (안) 개발 연구. *서울시 여성가족재단 연구사업보고서*, 1-101.

4) 양난주. "초등교육과 온종일 돌봄". 한겨레. 2020.06.22. <http://www.hani.co.kr/arti/opinion/column/950459.html>

5) 이해숙, & 이영주. (2018). 서울시 방과 후 돌봄 지역사회 협력방안. *서울연구원 정책과제연구보고서*, 1-135.

6) KB 금융경영연구소. (2018). 2018 한국의 워킹 맘 보고서.

7) 8) KOSIS 국가통계포털. (2018). 2018 청소년 방과 후 활동 수요 및 현황 조사.

9) 채정병 외. (2008). 연령에 따른 보행의 시간적, 공간적 요소에 관한 연구.

진찬우, & 이건학. (2015). 행복주택 최적 입지 선정에 관한 연구: 다목적 공간 최적화 접근*. *한국도시지리학회지*, 제18권 2호(81-95).

분석 언어 및 도구

