

제10회 빅데이터기반
성남시 청소년 정책제안대회

 아이디어
맛.ZIP

제안대회 빅데이터 공개 자료집

1. 데이터의 역사

1 빅데이터의 역사

최초의 데이터 활용

우리가 찾아낸 가장 오래된 인류 문명 최초의 기록은 그림으로, 기원전 2만 2천년 경, 알타미라 동굴의 벽화로 추정됩니다.

현존하는 가장 오래된 수치의 기록은 기원전 1만 8천년 경의 엄대(Tally sticks, 나무 막대기나 뿔, 뼈 등의 재료에 홈을 파서 수치를 기록)입니다. 이것을 통해 고대 사람들은 식량 재고를 기록하는 등 수치를 기반으로 판단을 할 수 있었습니다.

이는 인류가 최초로 문자를 기록한 기원전 3천년 경 이집트의 파피루스 두루마리와 메소포타미아의 점토판보다 1만 5천년 앞서 있는 시점으로, 인류는 문자보다 수치를 먼저 기록해왔던 것이죠.

이후 기원전 2,400년 바빌로니아에서 인류 최초의 계산도구인 주판이 발명되었고, 서기 100년에 그리스에서 최초의 기계식 청동 계산기인 안티키테라가 발명되면서, 인류는 문명의 많은 부분에서 수치를 통해 사고하고 판단하게 됩니다. 이것이 데이터 활용의 시작입니다.



고대 인류가 수치를 기록하였던 엄대

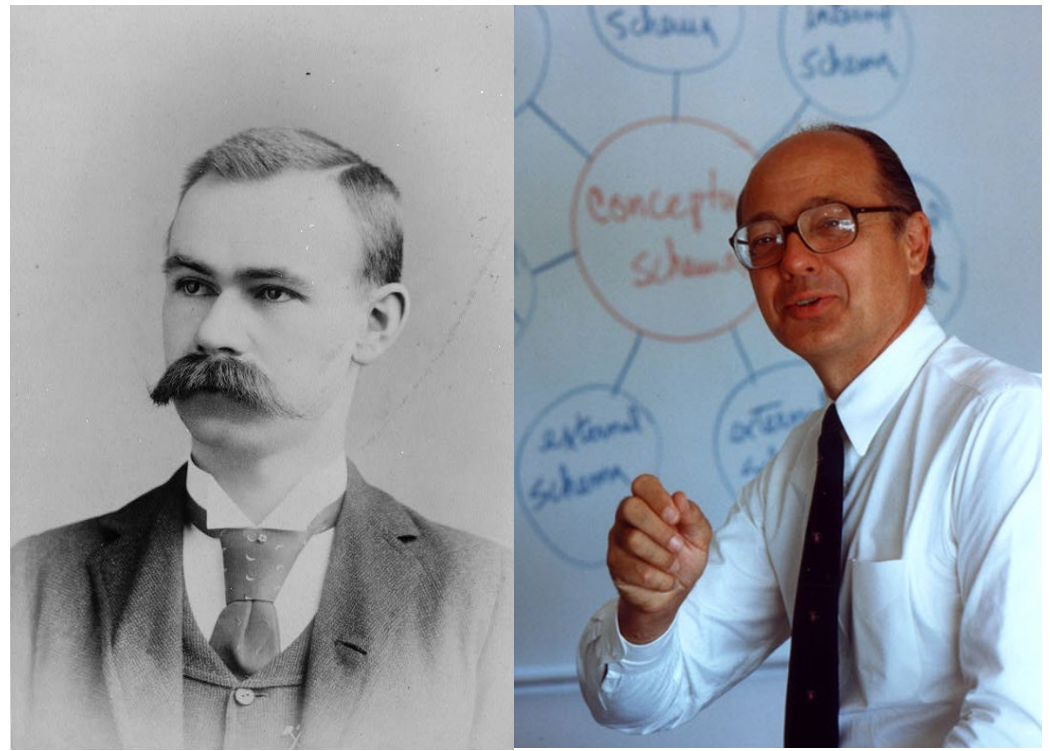
* 참고자료 - 산업통상자원부 ktech 인더스토리

1 빅데이터의 역사

컴퓨터, 데이터 시대 개막

독일계 미국인이자 통계학자인 허먼 홀러리스(Herman Hollerith)는 수백만 조각의 자료로부터 통계를 빠르게 도표화하기 위하여 천공 카드를 기반으로 한 공학용 도표 작성기를 개발하였습니다. 그가 개발한 작성기는 10차 조사까지 수년이 걸렸던 미국 통계청 인구조사의 데이터 처리 기간을 11차 조사인 1890년에 집계 분석 기간을 수개월 단위로 줄였습니다. 그는 전산제표기록회사(Computing-Tabulating-Recording Company)를 설립했고, 이 기업이 바로 오늘날 IBM입니다.

20세기에 들어 컴퓨터가 개발되면서 데이터를 저장할 수 있게 되면서 데이터의 시대는 한 단계 발전하였습니다. 1960년대 초반 제너럴 일렉트릭사의 찰스 바크만(Charles Bachman)은 자료의 조작과 저장, 관리를 프로그램과는 별개로 분리하는 현대적인 데이터베이스 개념을 처음으로 고안해냈고, 최초의 상용화 데이터베이스 관리 시스템(DBMS)인 IDS를 개발하여 제너럴 일렉트릭 사의 생산라인을 관리하는 MIACS 시스템에 활용하였습니다.



허먼 홀러리스(왼쪽)와 찰스 바크만(오른쪽)

* 참고자료 - 위키백과

1 빅데이터의 역사

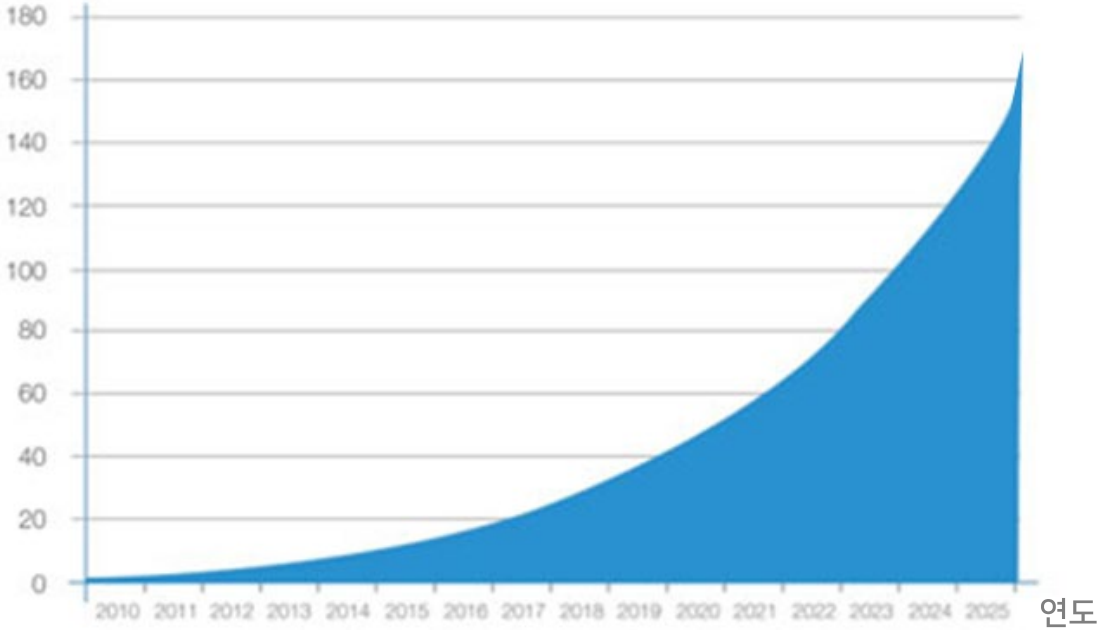
인터넷, 데이터 빅뱅

1969년, 미국 국방성은 정보의 안전한 관리를 위해 4개 대학에 서버를 두고, 아파넷(ARPAnet)이라는 네트워크를 만들었고, 1983년 민간에 사용하도록 개방하여 인터넷이 본격적으로 시작되었습니다.

한국에서는 1982년 5월 서울대학교와 한국전자통신연구원(ETRI)이 1200bps 전화선을 통해 SDN을 연결함으로써, 미국에 이어 세계에서 두 번째로 네트워크 망을 자체 개발하였습니다. 이러한 성과를 시작으로 2020년 현재, 우리는 전 세계 최고 수준의 인터넷 인프라 환경 속에서 생활하고 있습니다.

1983년 최초의 인터넷이 시작된 지 25년이 지난 2018년 전 세계에서 연간 생성되는 데이터 양은 33 제타 바이트(33조 기가 바이트) 수준에 도달하였으며, 2025년에 이르면 175제타 바이트까지 늘어날 것으로 예측되고 있습니다. 이제 우리 생활에서 빅데이터는 자연스러운 일상이 되었습니다.

제타 바이트



글로벌 연간 데이터 생성 추이(2010~2025, IDC)

* 참고자료 - 위키백과, IDC

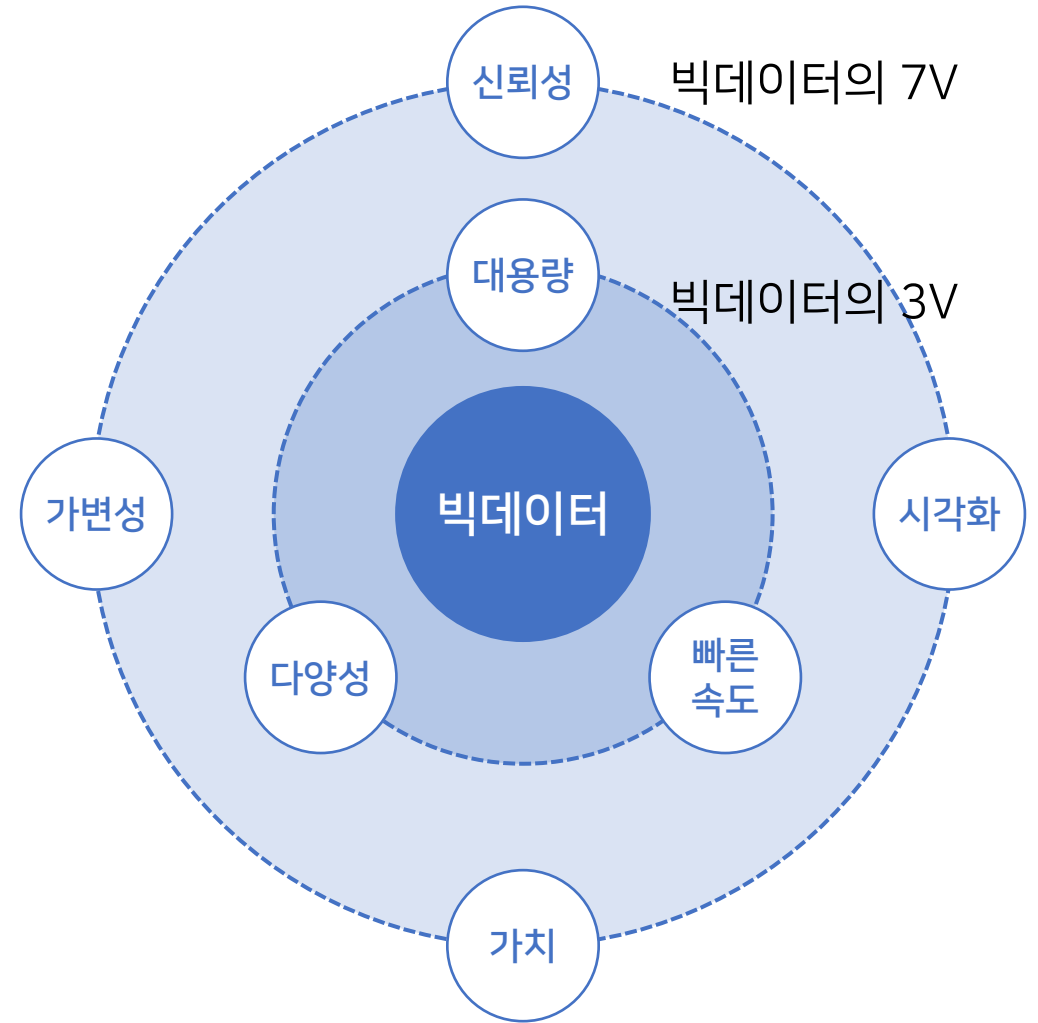
2. 빅데이터의 정의

2 빅데이터의 정의

빅데이터의 정의

빅데이터는 데이터 수집·저장·관리·분석 관점에서 기존 데이터베이스 역량을 넘어서는, 정형(형태가 일정한)·비정형(형태가 일정하지 않은) 대용량의 데이터를 말합니다. 빅데이터의 정의는 시대가 변화함에 따라 확장되고 있습니다.

빅데이터는 판단에 필요한 정보와 지식을 추출하는 관리·분석 기술을 필요로 하며, 대용량(Volume), 다양성(Variety), 빠른 속도(Velocity)의 3V의 특성을 갖습니다. 여기에 신뢰성(Veracity), 가변성(Variability), 시각화(Value) 가치를 더해 7V로 확장됩니다. 경우에 따라서는 정확성(Validity), 휘발성(Volatility)을 포함하기도 합니다. 대규모의 데이터를 처리하기 위해서는 컴퓨터의 처리 성능과 데이터를 전문적으로 다룰 수 있는 인력이 필요합니다. 데이터를 처리하는 방법, 기술에 대한 과학을 데이터 사이언스라고 부르며, 빅데이터 시대와 함께 각광받고 있습니다.



2 빅데이터의 정의

빅데이터의 한계

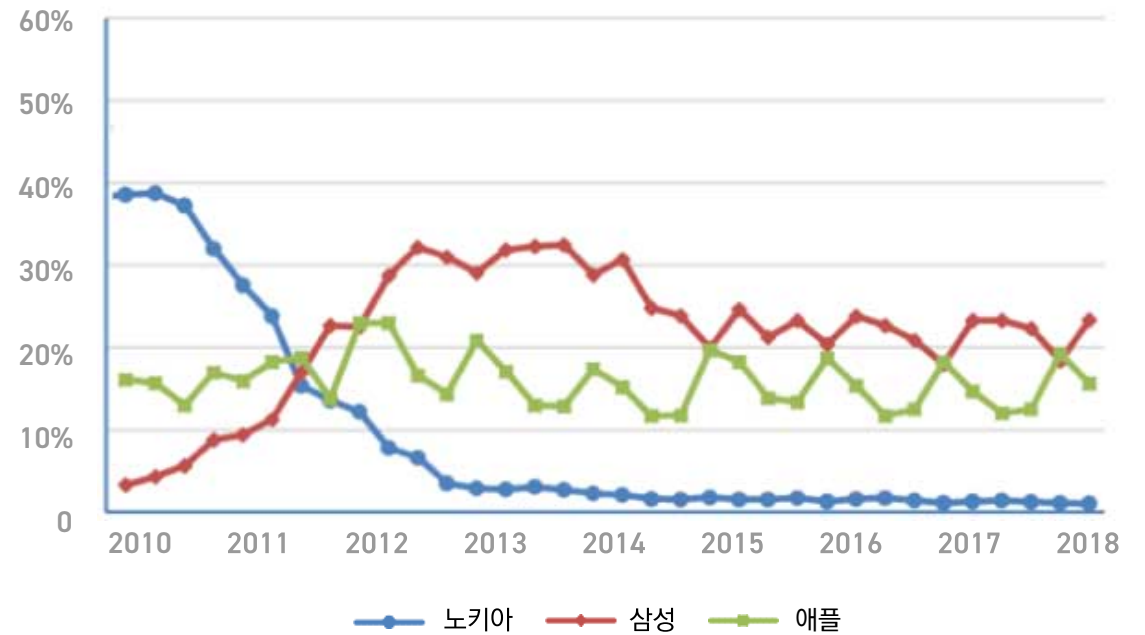
데이터가 무조건 많다고 가치를 갖는 것은 아니라는 관점에서 미국의 기술 인류학자 트리시아 왕(Tricia Wang)이 빅데이터와 함께 '두꺼운 데이터 접근'(Thick Data)이라는 개념을 제안하였습니다.

트리시아 왕은 대표적인 사례로 노키아의 사례를 제시합니다. 2007년 9월 새로운 개념의 스마트폰을 애플에서 발표한 시점에서 글로벌 시장 점유율 1위 휴대전화 기업이었던 노키아는 자사가 확보한 방대한 고객 데이터를 수집하였습니다.

그러나 이 데이터의 모집단은 노키아의 피쳐폰을 사용하던 고객이었고, 스마트폰에 열광하는 새로운 이용자를 반영하지 않았습니다. 트리시아 왕은 개발도상국에 직접 방문하여 실제 사람들이 스마트폰에 얼마나 열광하는지 소규모의 데이터만으로 통찰을 얻었고, 노키아에 전달하였지만, 노키아는 빅데이터만 믿었고, 변화의 흐름을 놓치게 되면서 애플과 삼성에게 세계 시장 점유율을 추월 당하고 시장에서 철수하게 됩니다. 중요한 것은 빅데이터 자체보다 이를 통해 얻게 되는 통찰입니다.

* 참고자료 - TED, Statista.com

노키아, 삼성, 애플 3사의 연도별 휴대전화 세계 시장 점유율 변화



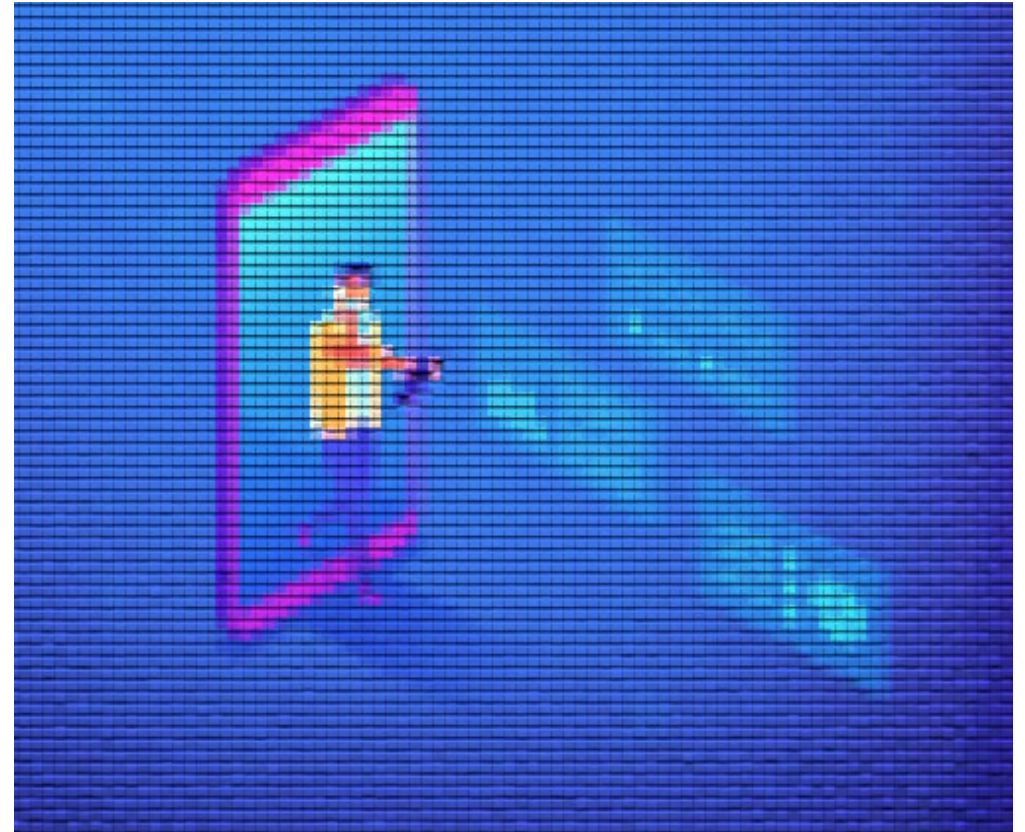
2 빅데이터의 정의

빅데이터의 올바른 활용

빅데이터의 정의와 빅데이터의 한계에 대해 살펴보았습니다. 빅데이터를 활용하는 핵심 목적은 탐색적이고 귀납적인 방법을 통해, 현재의 상황에 대한 새로운 발견을 하고, 이를 통해 우리가 앞으로 나아갈 방향에 대해 올바른 판단을 하는 것입니다. 정책제안처럼, 인간 행동에 대한 판단은 대체로 설문이나 인터뷰와 같은 방법들을 통해 이루어져 왔습니다.

예를 들어, 10대가 주로 사용하는 앱과 사용시간을 통해 올바른 모바일 사용에 대한 교육정책을 제안한다고 해봅시다. 조사와 인터뷰를 통해서 얻을 수 있는 데이터 양에는 한계가 있지만, 통신사의 앱 사용 데이터와 앱스토어 데이터를 수집할 수 있다면 시간을 훨씬 단축하면서도 정확한 데이터를 얻을 수 있겠죠. 반면, 조사와 인터뷰를 통해서 요즘 갑자기 유행하기 시작한 앱을 사용하는 이유와 같이 빅데이터로 찾을 수 없는 숨어 있는 시사점을 발견할 수 있을 것입니다.

결론적으로 우리는 기존의 설문, 인터뷰 등에서 얻게 되는 통계자료와 빅데이터를 함께 살펴볼 때, 더 나은 통찰에 도달할 수 있을 것입니다.



3. 빅데이터의 활용

3 빅데이터의 활용

빅데이터 활용 사례 1 - 자라

스페인 의류기업 인디텍스가 보유한 전세계 1위 SPA 패션 브랜드 자라(ZARA)는 그 어떤 패션 기업 브랜드보다 데이터를 잘 활용합니다. 자라가 판매하는 전 제품에는 RFID(Radio-Frequency Identification) 태그가 적용되어 있습니다. RFID 태그는 무선인식 시스템으로, 제품의 소재, 유형, 재고 등의 세부정보들이 저장되어 있고, 실시간으로 고객이 옷을 입어보거나 하는 다양한 데이터를 수집합니다.매장 판매기, 온라인 웹사이트 판매, 고객조사, PDA 기기 등에서 얻어지는 데이터를 수집하여 재고를 관리하고, 고객들이 선호하는 제품 정보를 수집합니다. 일반적인 패션 브랜드가 다음 시즌 예측을 위해 데이터를 사용하는 반면, 자라는 시즌 초기 물량을 적게 생산하면서, 고객 반응 데이터를 빠르게 분석하여, 해당 시즌이 진행되는 동안 생산, 재고 관리, 할인 등을 신속하게 진행합니다. 빅데이터 활용은 비용을 줄이고 이익률을 높여, 2017년 자라의 영업이익이익률은 17.2%로 같은 SPA 패션 브랜드인 H&M 12.4%, 유니클로 7.1%보다 월등히 높게 나타났습니다.



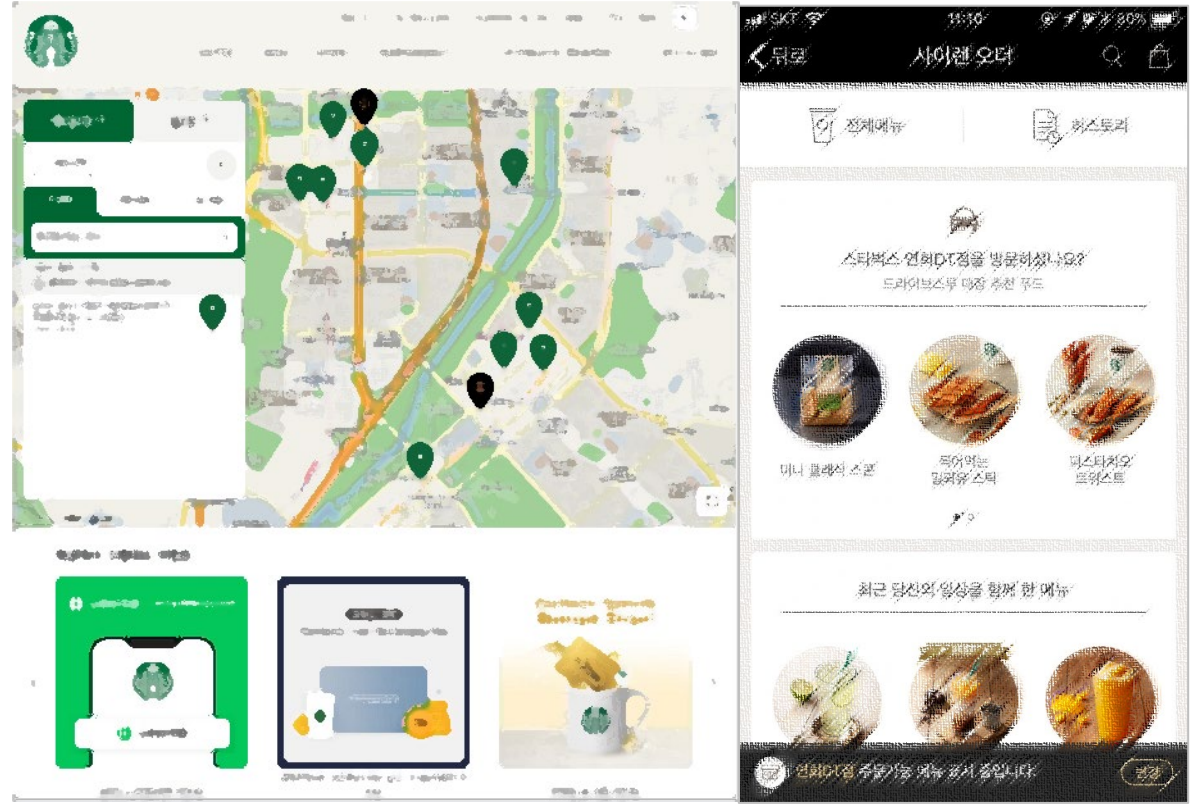
* 참고자료 - 인디텍스 홈페이지, 기업 소개자료

3 빅데이터의 활용

빅데이터 활용 사례 2 - 스타벅스

우리가 잘 아는 커피전문 글로벌 프랜차이즈 스타벅스는 IT기업보다도 빅데이터를 잘 활용하는 기업입니다. 스타벅스의 빅데이터 활용은 크게 2가지 측면에서 이루어집니다.

첫 번째는 매장의 위치를 선정하고, 매장을 늘리거나 줄이는 기준으로 빅데이터를 활용합니다. 성남 분당구 서현역 반경 300m에는 스타벅스 매장이 다섯 개나 됩니다. 스타벅스는 새 매장을 낼 때 유동인구와 같은 데이터가 뒷받침된다면 매장 간 거리가 가깝더라도 매장을 오픈합니다. 두 번째로는 고객의 취향을 분석하여 맞춤형 상품을 추천하는 과정에서 빅데이터를 활용합니다. 스타벅스는 모바일을 통해 주문할 수 있는 앱을 개발하였고, 앱을 통해 고객이 구매한 커피와 음식 데이터를 분석하여 고객이 원하는 메뉴를 추천하고 있습니다. 앱을 통한 주문 중에 37%는 빅데이터 기반 추천 메뉴에서 발생할 만큼 데이터 기반의 추천은 성과를 만들어 냅니다.



* 참고자료 - 스타벅스 홈페이지, 기업 소개자료

2 빅데이터의 활용

빅데이터 활용 사례 3 - 넷플릭스

넷플릭스는 인터넷을 통해 언제 어디서나 미디어 콘텐츠를 시청할 수 있는 전 세계 1위 OTT 서비스(Over The Top Service, 인터넷을 통해 방송 프로그램·영화·교육 등 각종 미디어 콘텐츠를 제공하는 서비스)로, 데이터 분석이 비즈니스 성공을 창출한 대표적인 사례입니다. 넷플릭스에서 빅데이터는 서비스 전 과정에서 매우 중요하게 다뤄지고 있습니다. 추천 알고리즘은 사용자 개인 정보가 아닌 영상 사용 패턴에 맞게 취향을 분석하여 영상을 추천합니다. 추천 알고리즘에 사용되는 데이터는 다양합니다. 시청자가 언제 영상을 멈추는지, 뒤로 돌려보는지, 멈춘 영상을 얼마 후에 또다시 재생하는지, 엔딩 크레딧이 올라간 이후의 행동은 무엇인지 등 넷플릭스는 사용자의 영상 습관을 데이터화 하여 패턴을 분석합니다. 시청자 행동의 75%가 넷플릭스의 추천에 기반할 정도로, 빅데이터를 활용하여 사용자 행동을 유도하고 있습니다. 넷플릭스는 오늘도 전세계 1억8800만 명, 국내 600만 명 사용자의 데이터를 수집·분석하고 있습니다.



* 참고자료 - 넷플릭스 홈페이지, 기업 소개자료, 보도자료

2 빅데이터의 활용

빅데이터 활용 사례 4 - 미국 보스턴 시

빅데이터 활용은 공익적 목적에서도 놀라운 결과를 만들어 냅니다. 2013년 4월 15일 미국 매사추세츠주 보스턴에서 개최된 마라톤대회의 결승선 직전에서 두개의 폭탄이 터져 관중들과 참가자 및 일반 시민들을 다치게 한 사건이 발생했습니다. 수만명이 모인 대회에, 범죄가 일어난 현장도 12개 블록에 걸쳐 있어 증거수집이 쉽지않았습니다. 그럼에도 불과 사흘 만에 테러용의자를 색출했습니다. 비결은 데이터에 있었습니다. FBI는 테러직후 주변 600여대의 CCTV 데이터를 수집하고, 시민들이 촬영한 사진과 영상 데이터를 비롯하여, 트위터와 페이스북 대화 등 클라우드소싱을 통해 수천여건의 데이터를 확보했습니다. 수집한 동영상을 분석하여 용의추정 인물을 유형화했고, 영상 판독과 부상자의 증언을 통해 확인하여 범인을 찾아냈습니다. 연관성 없어 보이는 정보라도 자세히 엮어보면 범죄 패턴 및 트렌드를 읽을 수 있는데, 그 분석기법이 바로 빅데이터입니다.



* 참고자료 - 국내외 보스턴 폭탄테러 관련 보도자료

2 빅데이터의 활용

빅데이터 활용 사례 5 - 강남경찰서

데이터는 이미 일어난 사건을 분석하는 '과거형' 수사에서, 예방 차원의 '미래형' 수사로 범죄 수사 방식을 바꾸고 있습니다. 서울 강남경찰서는 지난해 데이터를 활용하여 범죄 장소를 예측한 뒤 순찰을 강화해 절도, 폭력 등 범죄를 줄이는 데 성공했습니다.

서울 강남경찰서는 19년 4~6월 3개월간 공공 데이터를 기반으로, 강남 지역 절도·폭력범죄에 영향을 미칠 수 있는 인구밀도, 편의점 수 등 8개 변수 데이터를 바탕으로 관할 35개 블록의 범죄 위험도를 분석했습니다. 이 중 고위험지역으로 분석된 7개 블록에서 주간·야간으로 집중 순찰을 벌인 결과, 범죄가 확연히 줄어드는 효과가 발생했습니다.

'예측적 경찰 활동' 수행으로, 관내 절도 범죄는 1년 전 같은 기간보다 평균 31%, 폭력 범죄는 평균 14% 감소하였습니다.

데이터는 자원의 효율적 배분을 위한 의사결정을 지원하여, 우리가 보다 안전한 일상생활을 만드는 것에도 많은 역할을 하고 있습니다.



* 참고자료 - 강남경찰서 보도자료

4. 성남시 데이터 활용 사례

4 성남시 데이터 활용 사례

성남시 자율주행 빅데이터

성남시는 2019년 11월, 자율주행차 상용화를 위한 융합 빅데이터 구축·생산을 진행하고 있습니다. 전국 지방자치단체 중 자율주행 자동차를 제작·운영하는 것은 성남시가 최초입니다. 자율주행 자동차 상용화가 되려면, 운전자가 개입하지 않아도 자동차 멈춤, 가속, 추월 등의 명령을 수행해야 합니다. 이를 위해서는 자율주행차의 인공지능이 자율주행차 운행 데이터를 학습해야 합니다.

성남시는 데이터 확보를 위해 판교 일대와 시청 주변에서 자율주행차를 시범 운행하면서 실증테스트 정보, 차량 제어 데이터, 자동차와 실시간 정보를 주고받는 5G망 정보, 인공지능 학습용 자율주행 영상 정보 등을 '교통·자율주행 빅데이터 센터'로 보냅니다. 시청 8층에 있는 빅데이터 센터는 22㎡ 규모에 데이터 분석용 장비, GPU, 클라우드 관리용 서버, 클라우드 시스템, 데이터 통합관리 시스템 등을 갖추고 있습니다.



4 성남시 데이터 활용 사례

성남시청소년재단 이슈페이퍼

성남시청소년재단은 '청소년이 24시간 행복한 성남시'를 만들기 위해, 2019년부터 데이터 기반의 이슈 페이퍼를 분기마다 발간하고 있습니다. '온종일 돌봄 정책에 따른 청소년 방과후 아카데미의 새로운 방향 모색', '청소년 미래역량 개발을 위한 청소년 참여기구의 활성화 방안', '특성화고 청소년 성장을 위한 지역기반 통합적 지원 모델 사례', '성남시 미래역량 개발을 위한 청소년 참여기구의 활성화 방안', '코로나-19 확산에 따른 청소년활동의 대안', '청년기본법의 시행에 따른 성남시청소년재단 청년정책방향', '온(溫)&온(동네) 성남형 학습지원 시스템 구축'을 주제로, 지금까지 총 7개의 이슈 페이퍼가 발간되었습니다. 성남시청소년재단 이슈페이퍼는 구체적이고 현실적인 청소년 프로그램 운영과 청소년 지원을 목표로 설문조사, 통계자료, 데이터를 기반으로 작성되며, 필요한 청소년 프로그램과 지원을 위한 근거로 활용됩니다.

ISSN 2671-8294

2020-1호

발행인 신미서 | 발행일 2020년 3월 10일
발행처 성남시청소년재단

코로나바이러스감염증-19 등 법적 감염병 확산에 따른 청소년 활동의 대안과 전략

□ 법적 감염병의 예방적, 주격적 발생 및 지원사의 감염 확산 예방 대책과 청소년 활동

- 지난 2, 23, 코로나19의 급격한 확산 및 지역사회 전파에 따라 감염병 위기경보가 최고 단계인 '심각' 단계로 격상됨
- 2009년 신종플루 사태 이후 11년 만에 '중앙재난안전대책본부'를 설치하고 발부처 대응을 하고 있으며 성남시 역시 방역과 확산방지를 위해 총력 대응을 하고 있음.
- 2002년 사스 바이러스, 2009년 신종플루, 2015년 메르스 바이러스 등 법적 감염병은 반복적, 주기적으로 발생하고 있음.
- 최근 및 청소년시설 휴관 등 감염병 대응책은 별기피하게 16만 명의 성남시 청소년의 삶에도 심대한 영향을 미치고 있으며, 이에 대해 성남시 청소년 기관과 지원 인력이 요구됨.

2020 **2015** **2009**

코로나바이러스감염증-19 **중동호흡기증후군** **신종 인플루엔자**

- 전역에 모든 유·초·중·고등학교 3주간 개학 연기
- 성남시청소년재단 내 다중이용시설 휴 68일로 휴가휴 휴관 연장
- 성남시 120여개 학교
- 성남시청소년재단 모든 시설 7일간 휴관
- 성남시 유·초·중·고등학교 휴교
- 성남시청소년재단 내 모든 집단모임과 발원행사 시행 및 집단모임 휴교

□ 청소년 사각지대의 발생과 대상별 맞춤형 서비스 제공 필요

- 대규모-상기간의 휴교 및 휴관 상황에서, 아동-청소년의 생활 실태와 서비스 사각지대를 신속하게 파악하고, 청소년들의 수요에 대응하는 지원방안이 요구됨.
- 대안전략이 제정될 수 밖에 없는 상황이나, 청소년, 특히 취약계층의 청소년들을 지원할 수 있는 방안을 모색하고 실현 가능한 방안을 제시할 필요가 있음.
- 본 호에서는 성남시청소년재단 소속 참여기구 및 청소년방과후아카데미 참여 청소년을 대상으로 실시한 '코로나바이러스감염증-19(COVID-19) 확산에 따른 청소년 생활 실태조사' 결과를 중심으로 청소년 활동 사각지대에 따른 대상별 맞춤형 서비스 제공 등 개선방안을 제안함.

성남시 청소년 인구조형 현황

2019. 11월 기준, 만9세~24세 인구 현황

학교급	학교 수	2019년 3월 기준	
		인원	학수(학급)
계	288	4,191	226(1)
유치원	131	533	17
초등학교	72	1,775	72(1)
중학교	45	779	37
고등학교	36	1,076	30
특수학교	2	18	-
특수학교	2	-	477
특수학교	2	-	423

성남시청소년재단
SEONGNAM CITY YOUTH FOUNDATION

공백 기간 중 주요 활동유형 유령(n=410)

공백 기간 중 주요 활동유형 '가정'

'가정' 내 주요 활동 대상 유형(n=348)

코로나19 확산에 따른 공백 기간 중 청소년의 주요활동

• 청소년들의 공백 기간 중 주요 활동유형은 '학습(32.4%)', 'TV 및 유튜브 시청(26.6%)', '취미활동(15.8%)' 순으로 나타났다.

• 특히 초등학생(31.8%) 및 중학생(31.2%)의 경우 TV 및 유튜브 시청을 많이 하고 있음.

TV 및 유튜브 시청 증가, '아르바이트' 감소

'부모'의 부재에 따른 가정 돌봄 공백 발생

• 가정, 학교, 친척집, 돌봄시설, 기타

• '가정' 내 주요 활동유형 '가정'

• '부모'의 부재에 따른 가정 돌봄 공백 발생

성남시 청소년 인구조형 현황

2019. 11월 기준, 만9세~24세 인구 현황

학교급	학교 수	2019년 3월 기준	
		인원	학수(학급)
계	288	4,191	226(1)
유치원	131	533	17
초등학교	72	1,775	72(1)
중학교	45	779	37
고등학교	36	1,076	30
특수학교	2	18	-
특수학교	2	-	477
특수학교	2	-	423

성남시청소년재단
SEONGNAM CITY YOUTH FOUNDATION

지원 요청 유형(n=442)

학교급 지원 요청 유형(n=400)

• 청소년들의 주요 활동유형은 '학습(32.4%)'이 가장 많으나, '휴게 및 여가'의 비율이 가장 높음.

• 학교 급별로는 중·고·고·특수학교, 초등·유·초·중·고·특수학교 순으로 높음.

• '가정' 내 주요 활동 대상 유형은 '가족', '친척', '친구' 순으로 높음.

• 지원 요청은 '활동거의(52.3%)'가 가장 많았으며, '건강관리(20.1%)', '학습지원(13.1%)', '사생활(6.3%)', '문화예술(4.8%)' 순으로 나타났다.

• 청소년들의 수요를 반영하여 사각지대, 공극의 개편을 방지 하는 비대면 활동 프로그램의 개발 및 지원이 필요함.

성남시청소년재단
SEONGNAM CITY YOUTH FOUNDATION

4 성남시 데이터 활용 사례

성남 교육격차 해소를 위한 빅데이터 분석

성남시청소년재단과 성남형교육지원단은 지역 내 교육격차를 해소하고, 지역 특색을 살릴 수 있는 학교교육 지원사업을 발굴하기 위해 2019년 '성남교육격차 해소를 위한 빅데이터 분석' 연구를 진행하였습니다. '성남교육격차 해소를 위한 빅데이터 분석'은 크게 세 가지 주제를 통해 교육격차를 분석하고 맞춤형 지역 특화 정책을 제시하고 있습니다.

첫째로, 인구 감소와 유출로 인한 학생 수의 감소 양상을 분석하여, 학생 수가 많이 감소하는 지역과 학교에 대한 정책사업을 제시하였습니다.

둘째로, 만 10세~14세 연령대 성남시 학생들의 유동인구를 기반으로 학생이 유해환경에 노출되는 성남시 내 지역별 격차를 분석하여 학생이 유해환경에 노출되는 주요지역에 대한 정책사업을 제시하였습니다.

셋째로, 돌봄을 필요로 하는 만 0세~9세 연령대, 방과 후 활동이 필요한 만 10세~14세 연령대의 유동인구 데이터와 지역 돌봄 및 방과후 교육 시설 인프라를 연계 분석하여 지원이 필요한 주요지역을 발견했습니다.



5. 유동인구 데이터와 활용 사례

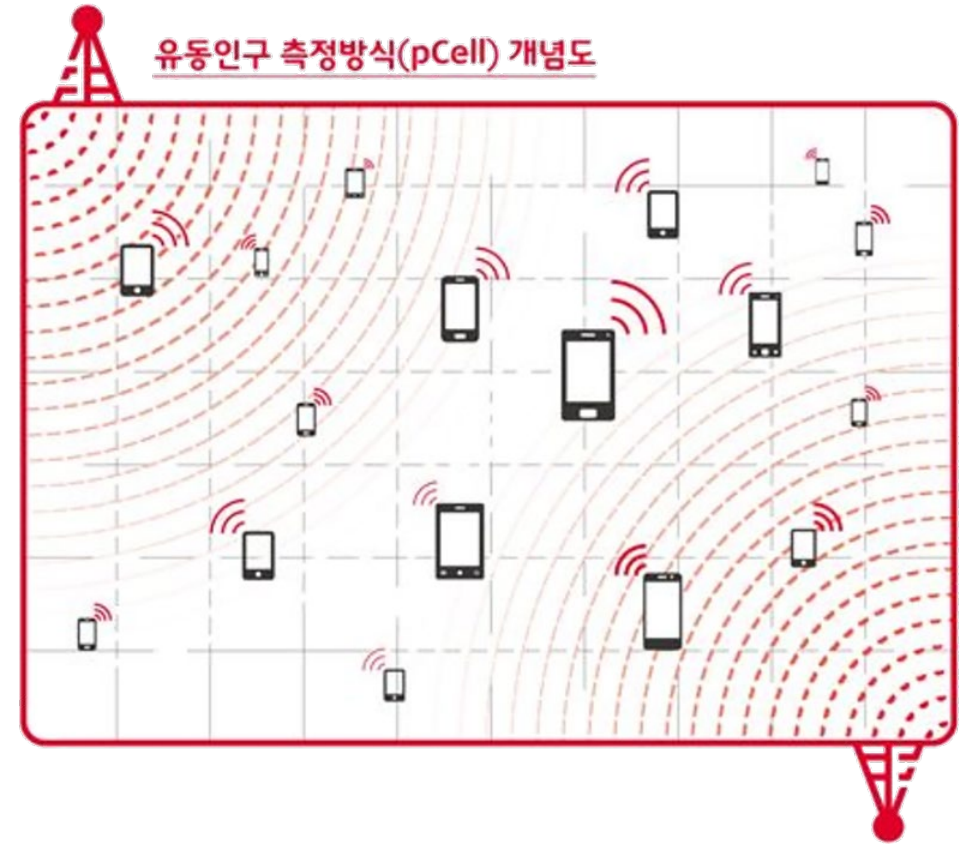
5 정책 아이디어 발견을 위한 유동인구 데이터

유동인구 데이터 소개

이번 대회에 제공될, 유동인구에 대해 소개해 드리겠습니다. 모바일이 생활의 일부가 된 지금, 휴대전화는 우리의 필수품이 되었습니다. 우리의 휴대전화와 통신사 기지국은 우리가 휴대전화를 사용하고 있는 시간에도, 사용하지 않는 시간에도 지속적으로 신호를 주고받습니다. 이 과정에서 통신사 기지국은 전파의 위치를 시간대별, 장소별로 측정하여 가로세로 50m 격자 단위로 인구를 집계합니다. 이 때 50m 격자를 셀(Cell)이라고 표현합니다.

지역 주변 이동통신 기지국 셀을 활용해, 지역을 가로세로 50m 면적을 기준으로 격자형으로 나누고, 각각의 공간에 미치는 이동통신 기지국의 신호 세기를 기준으로 측정합니다.

유동인구가 많을 수록 해당 지역에 특정 시간대에 많은 사람이 있다는 의미가 됩니다. 많은 사람이 다니는 것이 직접적인 의미를 가지는 상권, 여행 등의 분야에서는 이미 유동인구를 활용한 다양한 분석이 진행되고 있습니다.



* 참고자료 - SK텔레콤 유동인구 소개자료

5 정책 아이디어 발견을 위한 유동인구 데이터

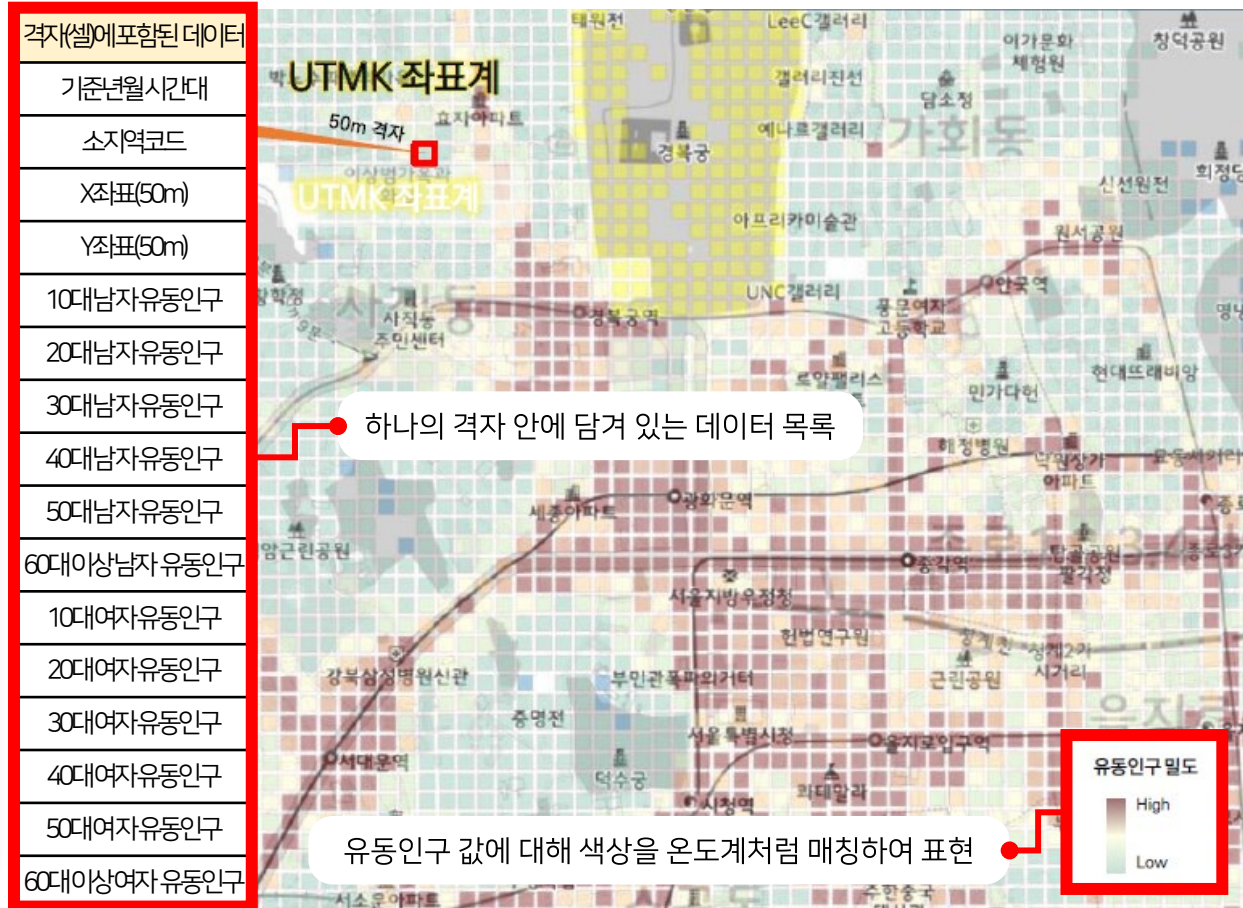
유동인구 데이터를 분석하고 시각화하는 과정

우측의 그림은 유동인구 데이터와 시각화 자료의 예시입니다. 가로세로 50m 작은 격자에는 기준년월 시간대, 소지역코드, X좌표, Y좌표, 10세 단위 연령별, 성별 유동인구의 수치가 담겨 있습니다.

원래 데이터는 일반적인 엑셀 데이터처럼 각 격자의 ID가 있고, ID마다 위에 언급한 데이터가 표현되어 있는 표 형태로 되어 있습니다.

연령별, 성별 유동인구의 수치에 맞게 색상 값을 지정하여 유동인구의 밀도를 시각적으로 표현할 수 있습니다. 이러한 데이터 시각화 기법을 열지도 시각화(Heatmap)라고 합니다. 그림 하단에 있는 범례에는 유동인구밀도에 대한 색상 지정을 표현합니다.

각 격자에 포함된 X좌표와 Y좌표를 통해, 우리는 이 색상 값을 지도의 좌표와 일치시켜 지도 상에 표현할 수 있습니다. 우측 그림을 살펴보면, 지도가 가로세로 50m 단위의 격자로 구분되어 있고, 격자마다 색상이 표현되어 있으며, 격자의 색상의 의미가 우측 하단의 범례에 지정되어 있습니다.



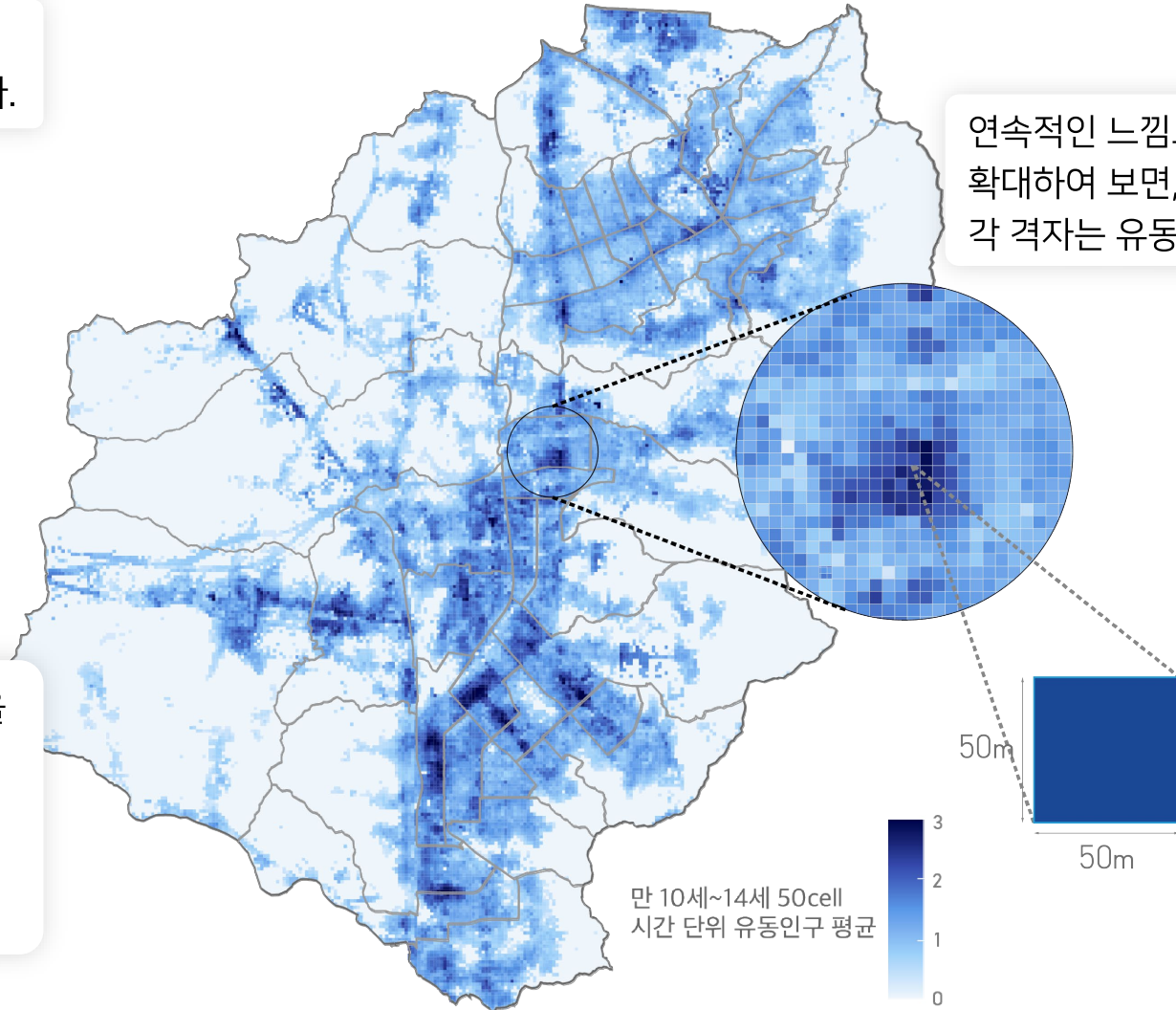
* 참고자료 - SK텔레콤 유동인구 소개자료

5 정책 아이디어를 위한 유동인구 데이터

이제 성남시 지도에 유동인구를 표현해 보겠습니다.

해당 지도는 특정 시간대, 만 10세~14세 유동인구를 성남시 지도에 시각화한 것입니다.

각 격자는 가로세로 50m이며, 유동인구 값을 갖고 있으므로, 특정한 행정동 내의 유동인구 전체 평균을 알 수도 있고, 반경 300m 내의 유동인구 평균값이 높은 지점이 어디인지도 알 수 있습니다.



연속적인 느낌으로 표현된 전체 지도를 확대하여 보면, 각각의 격자가 있습니다. 각 격자는 유동인구 값을 갖고 있죠.

5 정책 아이디어를 위한 유동인구 데이터

유동인구를 통한 분석이 어떻게 이루어지는지, 앞서 소개한 성남 교육격차 해소를 위한 빅데이터 분석 사례를 살펴보겠습니다.

학교 영역을 제외한 학기 중 일과 시간대 주요 지역

(학기 중 평일, 09:00~14:00)

학교 영역을 제외하고, 하루 일과 시간대 일 평균 만 10세~14세 연령 인구 활동이 가장 높은 수치를 기록한 5개의 단일 격자와 반경 300m 활동량



1. 한국잡월드 진입로 인근

반경 300m 내 일과 시간대 일 평균 1,063.5명, 1시간당 212.7명 활동 (측정 격자 104개, 단일 격자 최대 67.6명)

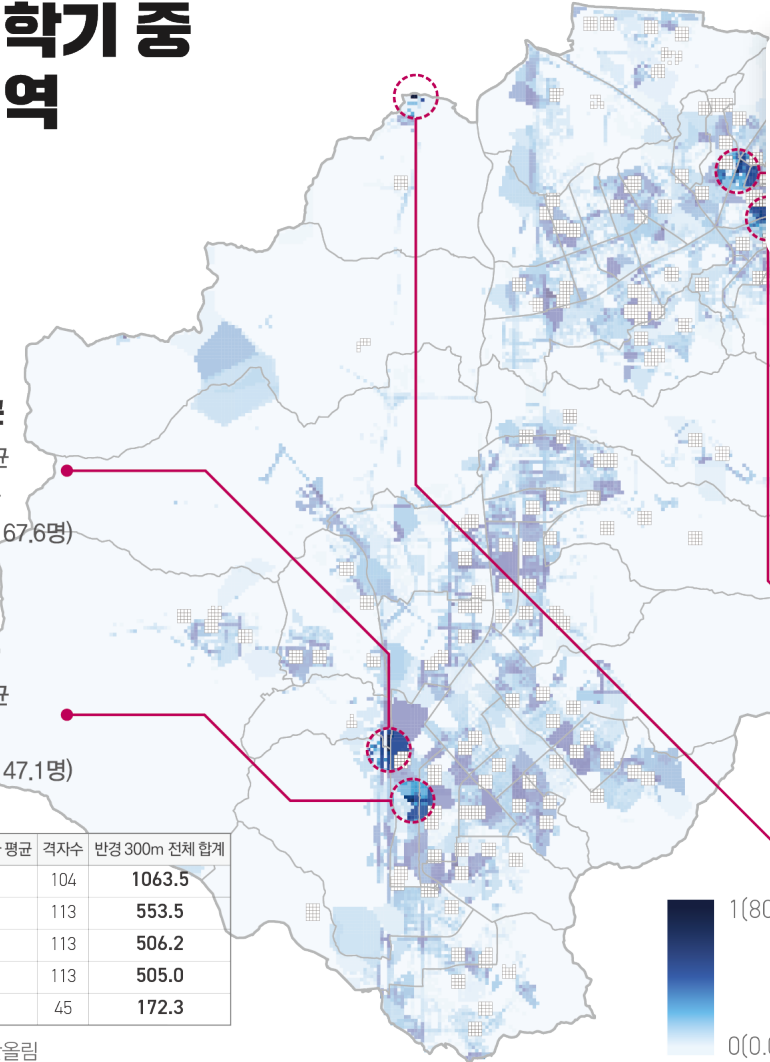


4. 정자동 카페 거리 인근

반경 300m 내 일과 시간대 일 평균 505.0명, 1시간당 101.0명 활동 (측정 격자 113개, 단일 격자 최대 47.1명)

구분	x좌표	y좌표	단일 격자 최댓값	반경 300m 격자 평균	격자수	반경 300m 전체 합계
1. 한국잡월드 진입로 인근	320752	530743	67.6	10.2	104	1063.5
2. 남한산성역 인근	325902	539293	55.6	4.9	113	553.5
3. 신구대 사거리 인근	326352	538593	46.8	4.5	113	506.2
4. 정자동 카페 거리 인근	321102	529993	47.1	4.5	113	505.0
5. 신촌동 세곡천 인근	321152	540393	80.3	3.8	45	172.3

※ 반경 300m 전체값 기준 순(단위: 명), 소수점 둘째 자리에서 반올림
 ※ 단일 격자의 최댓값은 주요 상위 지역의 중심 좌표를 도출하기 위해 산출
 ※ 활동하는 인구의 수를 의미하므로 추가 정보 제공을 위해 반경 300m 전체값 산출



유동인구 격자에는 값이 있으므로 반경과 시간대, 연령을 지정하여 반경 내 유동인구가 가장 높은 지역을 찾을 수 있습니다.

2. 남한산성역 인근

반경 300m 내 일과 시간대 일 평균 553.5명, 1시간당 110.7명 활동 (측정 격자 113개, 단일 격자 최대 55.6명)



3. 신구대 사거리 인근

반경 300m 내 일과 시간대 일 평균 506.2명, 1시간당 101.2명 활동 (측정 격자 113개, 단일 격자 최대 46.8명)



5. 신촌동 세곡천 인근

반경 300m 내 일과 시간대 일 평균 172.3명, 1시간당 34.5명 활동 (측정 격자 45개, 단일 격자 최대 80.3명)



5 정책 아이디어를 위한 유동인구 데이터

지역환경도 마찬가지로의 방법으로 함께 분석하여 시각화 한다면 어떨까요?

성남시 역내 청소년 유해 환경 밀도

성남시 학생 유해 환경 시각화 (50m×50m 격자별 정규화 지수 시각화)

유해시설도 주소를 변환하는 과정을 통해 좌표를 알 수 있으므로, 같은 격자에 유해 환경의 밀도를 시각화할 수 있습니다.



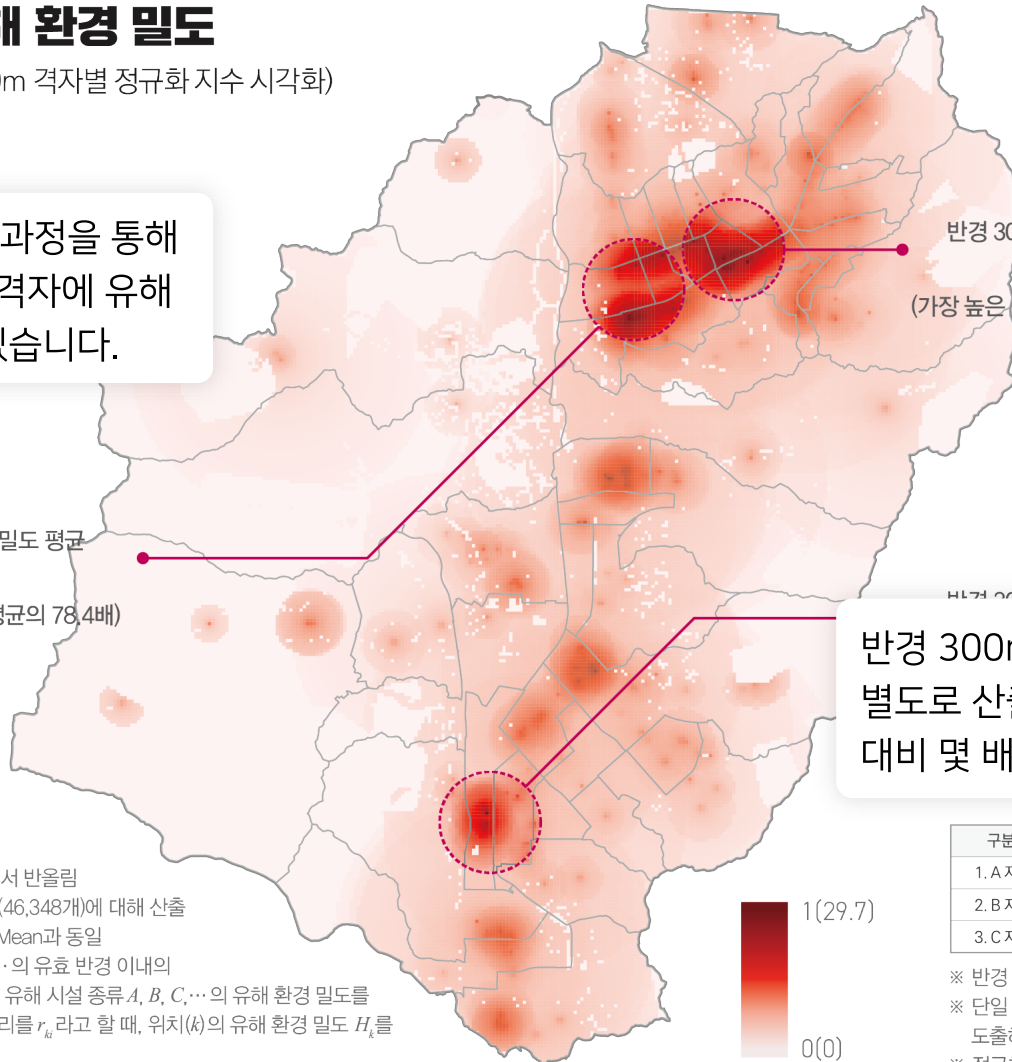
2. B 지역

반경 300m 내 유해 환경 밀도 평균
성남시 평균의 29.6배
(가장 높은 격자 성남시 평균의 78.4배)

Minimum	Mean	Maximum
0	0.31	29.7

- ※ Minimum, Mean, Maximum 값은 소수점 둘째 자리에서 반올림
- ※ 한번이라도 만 10세~14세 활동이 측정된 모든 격자(46,348개)에 대해 산출
- ※ 모든 격자에 대해 연산을 수행하여 유효값의 평균이 Mean과 동일
- ※ 성남시 내 임의의 위치(k)의 유해 시설 종류 A, B, C...의 유효 반경 이내의 유해 시설 종류 A, B, C...가 n_a, n_b, n_c, \dots 개 있는 경우, 유해 시설 종류 A, B, C...의 유해 환경 밀도를 각각 a, b, c, ...라고 하고, 위치(k)와 유해 시설과의 거리를 r_{ki} 라고 할 때, 위치(k)의 유해 환경 밀도 H_k 를

$$H_k = \sum_{i=1}^{n_a} \frac{a}{r_{ki}} + \sum_{i=1}^{n_b} \frac{b}{r_{ki}} + \sum_{i=1}^{n_c} \frac{c}{r_{ki}} + \dots \text{로 정의}$$



1. A 지역

반경 300m 내 유해 환경 밀도 평균
성남시 평균의 31.4배
(가장 높은 격자 성남시 평균의 94.5배)



3. C 지역

반경 300m 내에서 수치가 높은 지역을 별도로 산출할 수 있으며, 성남시 평균값 대비 몇 배인지도 알 수 있습니다.

구분	단일 격자 최댓값	반경 300m 격자 평균	평균 대비 배율
1. A 지역	29.7	9.8	31.4
2. B 지역	24.6	9.3	29.6
3. C 지역	25.0	3.4	10.8

- ※ 반경 300m 평균 대비 배율 순, 소수점 둘째 자리에서 반올림
- ※ 단일 격자의 최댓값은 주요 상위 지역의 중심 좌표를 도출하기 위해 산출
- ※ 정규화된 지수값이므로 반경 300m 합계값은 특별한 의미가 없으므로, 산출하지 않고 평균 대비 배율을 산출

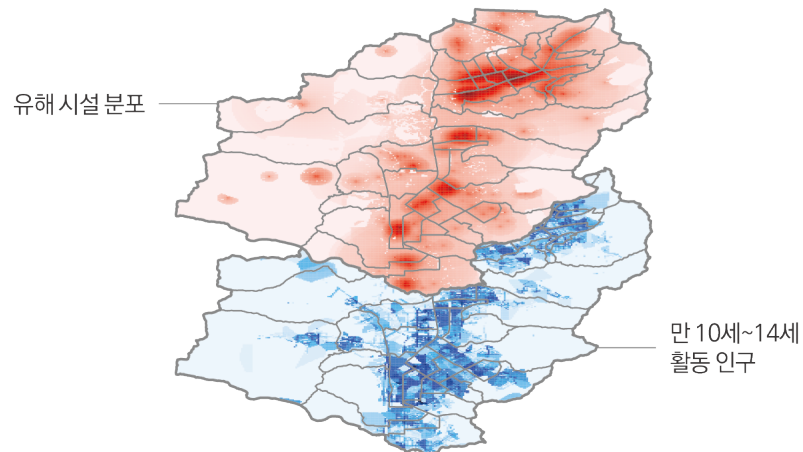
5 정책 아이디어를 위한 유동인구 데이터

유동인구와 유해환경을 결합하면 새로운 가치를 얻을 수 있습니다.

만 10세~14세 유해 환경 노출도

유해 환경과 활동인구를 결합해 유해 환경 노출 정도 분석

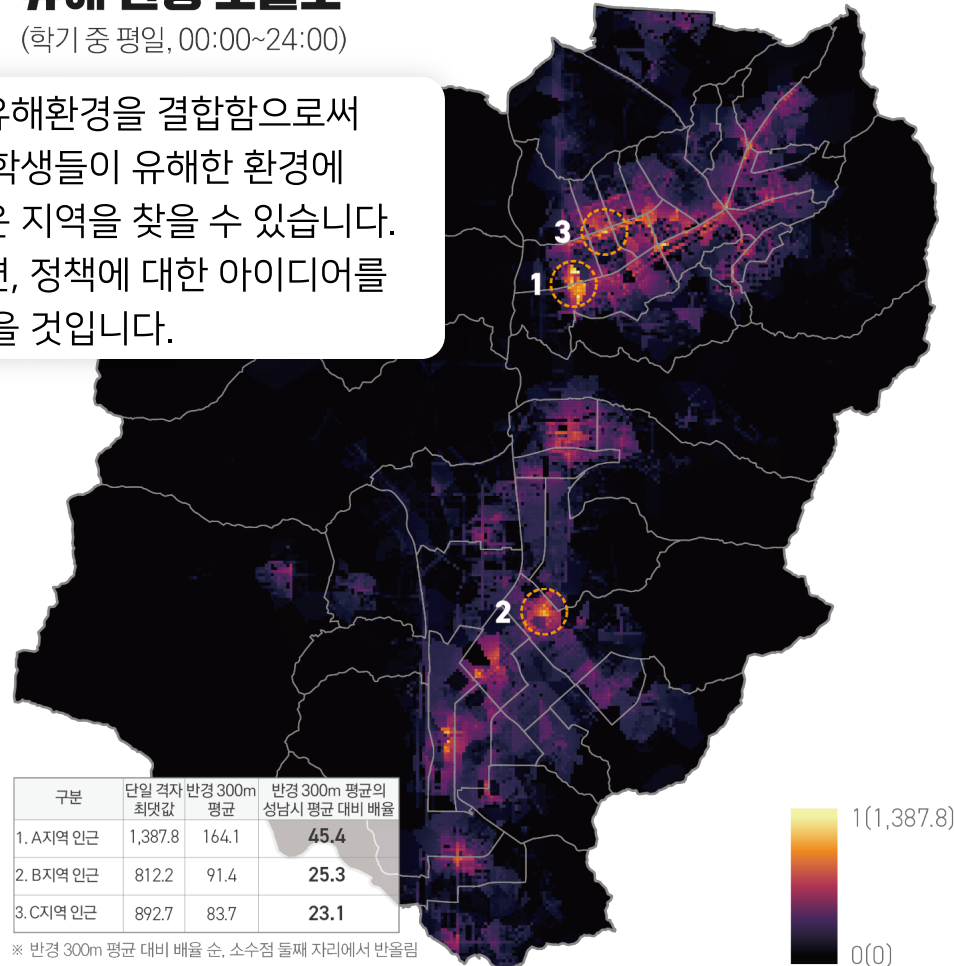
- 유해 환경 시각화와 활동 인구 시각화를 결합하여 성남시 임의의 격자에 학생이 유해 환경 근처에 노출되어 있는 정도의 차이를 분석
- 유해 환경 노출도는 단일 격자에서 만 10세~14세 학령 인구 활동이 많을수록, 유해 환경 밀도(H_i)가 높을수록 높다고 가정하고, 두 값의 곱연산으로 정의
- 방과 후 시간대, 저녁 활동 시간대, 심야 수면 시간대의 학생 활동 인구 (만 10세~14세)와 결합하여 시간대별 유해 환경 노출도 산출, 소수점 둘째 자리 반올림
- 단일 격자 노출도가 높은 지점(Spot)을 시간대별로 최상위 5개를 추출하여 격자 반경 300m 노출도 평균이 가장 높은 3개 지역 산출(단일 격자의 순위와 격자 반경 300m 평균값의 순위는 다를 수 있음)



만 10세~14세 전체 시간대 유해 환경 노출도

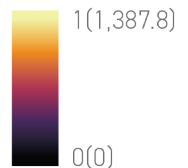
(학기 중 평일, 00:00~24:00)

유동인구와 유해환경을 결합함으로써 해당 연령대 학생들이 유해한 환경에 노출되기 쉬운 지역을 찾을 수 있습니다. 지역을 찾으면, 정책에 대한 아이디어를 생각할 수 있을 것입니다.



구분	단일 격자 반경 300m 최댓값	반경 300m 평균	반경 300m 평균의 성남시 평균 대비 배율
1. A지역 인근	1,387.8	164.1	45.4
2. B지역 인근	812.2	91.4	25.3
3. C지역 인근	892.7	83.7	23.1

※ 반경 300m 평균 대비 배율 순, 소수점 둘째 자리에서 반올림
 ※ 단일 격자의 최댓값은 주요 상위 지역의 중심 좌표를 도출하기 위해 산출
 ※ 정규화된 지수값이므로 반경 300m 합계 은 특별한 의미가 없으므로, 산출하지 않고 평균 대비 배율을 산출



6. 정책제안대회를 위한 유동인구 분석 데이터

6 정책제안대회를 위한 유동인구 분석 데이터

빅데이터 기반 정책대회 데이터 구성

데이터 구성은 다음과 같습니다.

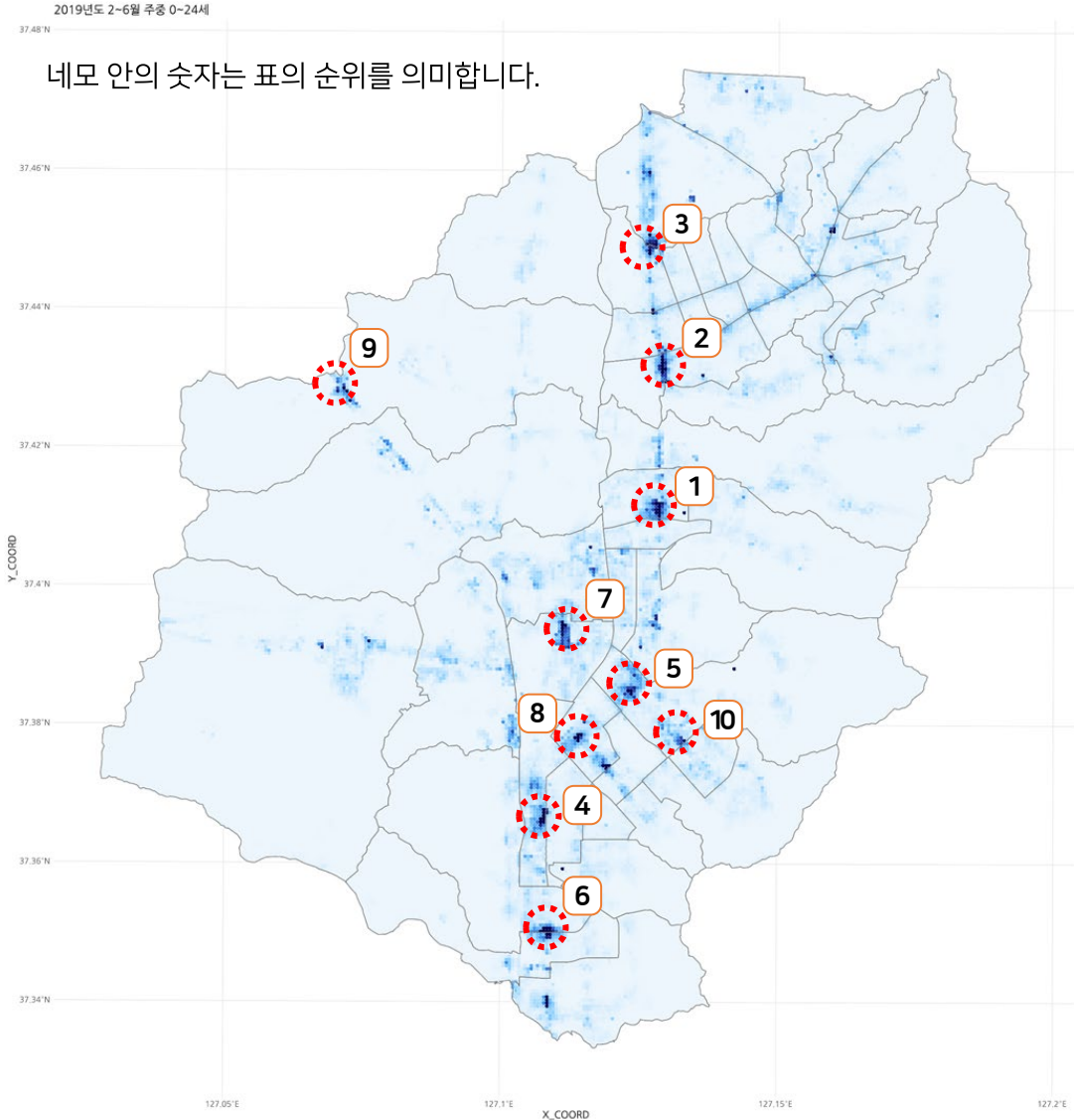
- 장소: 성남시 전체 지역, 가로세로 50m 단위 격자
- 기간 및 시간대:
 1. 코로나19 상황인 2월~6월, 2019년과 2020년, 24시간 평균
 2. 코로나19가 심각했던 3월~4월, 2019년과 2020년, 사람들이 주로 활동하는 09시~18시 평균
- ※ 각 기간에서 주중과 주말 구분
- 연령대 구분: 만 0세~24세(아동, 청소년층), 만 25세~49세(청장년층), 만 50세 이상(중노년층)으로 구분
- 데이터 시각화 방법: 유동인구 값(밀도) 색상 지정 열지도 시각화
- 제공 데이터: ①과 ②의 주중과 주말에 대해, 2019년 데이터, 2020년 데이터, 2019년과 2020년의 각 격자별 유동인구 변화 데이터, 각 지도마다 높은 지점의 위치와 값, 성남시 평균 값과의 배율 수치 제공

빅데이터 기반 정책대회 데이터 목록 코로나19 상황인 2월~6월, 2019년과 2020년, 24시간 평균

- 1-1-1.~1-1-3. 주중 만 0세~24세 2019년 유동인구, 2020년 유동인구, 2019년과 2020년 유동인구 변화량
- 1-2-1.~1-2-3. 주말 만 0세~24세 2019년 유동인구, 2020년 유동인구, 2019년과 2020년 유동인구 변화량
- 2-1-1.~2-1-3. 주중 만 25세~49세 2019년 유동인구, 2020년 유동인구, 2019년과 2020년 유동인구 변화량
- 2-2-1.~2-2-3. 주말 만 25세~49세 2019년 유동인구, 2020년 유동인구, 2019년과 2020년 유동인구 변화량
- 3-1-1.~3-1-3. 주중 만 50세 이상 2019년 유동인구, 2020년 유동인구, 2019년과 2020년 유동인구 변화량
- 3-2-1.~3-2-3. 주말 만 50세 이상 2019년 유동인구, 2020년 유동인구, 2019년과 2020년 유동인구 변화량

6 정책제안대회를 위한 유동인구 분석 데이터 2월~6월, 24시간

1-1-1. 만 0세~24세 코로나 발생 전인 2019년 2월~6월 주중 유동인구, 지도 상에 표시된 원은 반경 300m 유동인구 평균값 상위 10개 지역



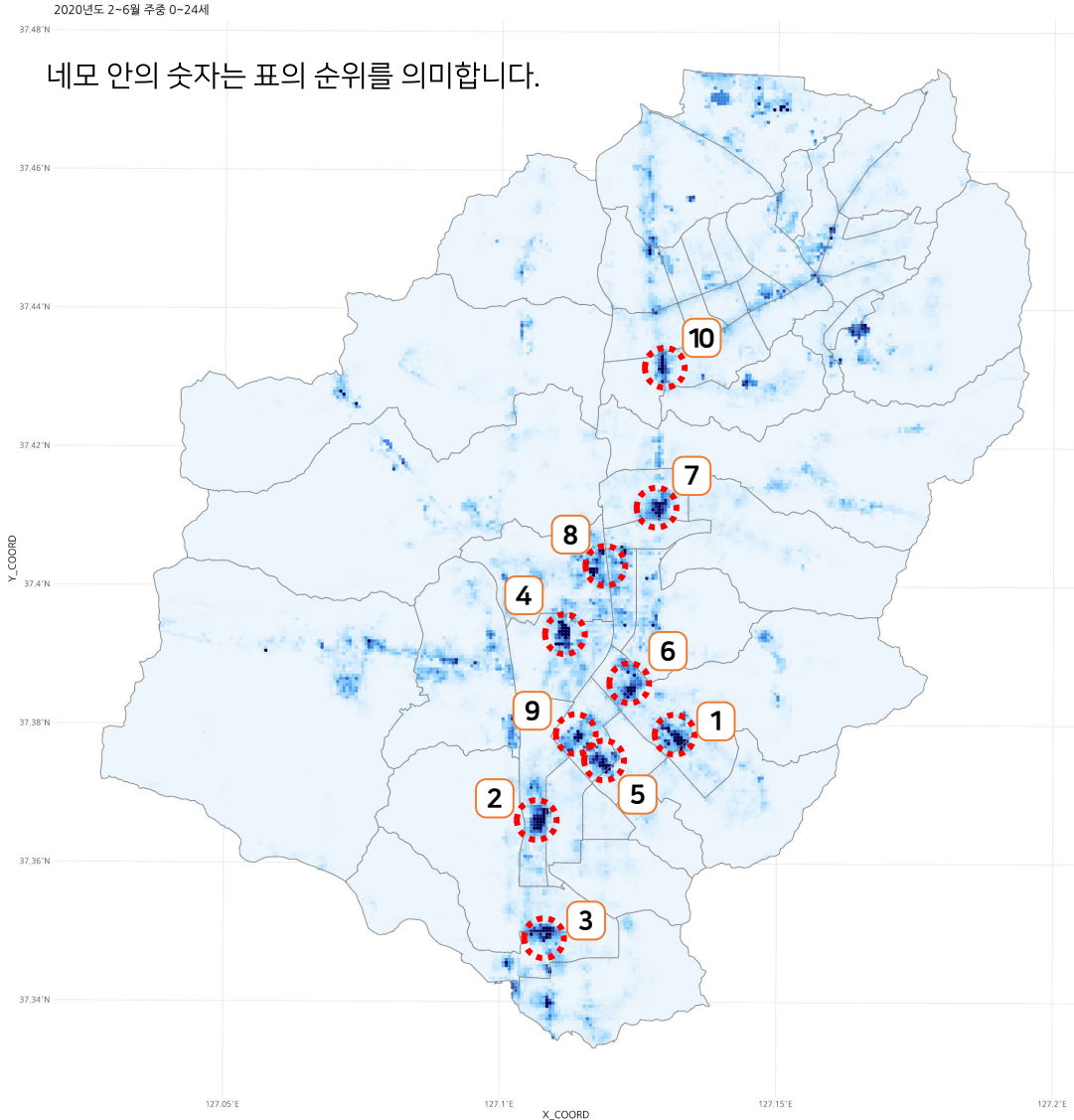
해당 조건의 성남시 유동인구 평균값: 71.5명

순위	반경 300m 평균 유동인구	성남시 평균대비	지역 위치
1	691.4	9.7배	분당선 야탑역 인근 지역
2	686.1	9.6배	분당선 모란역 인근 지역
3	635.0	8.9배	수인분당선 가천대역 인근 지역
4	620.1	8.7배	분당선 정자역 인근 지역
5	615.5	8.6배	분당선 서현역 인근 지역
6	605.4	8.5배	분당선 미금역 인근 지역
7	532.4	7.4배	신분당선 판교역 인근 지역
8	434.5	6.1배	분당선 수내역 인근 지역
9	376.9	5.3배	경부고속도로 성남 고등동 진입로 인근 지역
10	345.1	4.8배	분당어린이공원 인근 지역

◦ 유의사항: 평균값을 산출할 때, 유동인구가 측정된 셀만 전체 셀 개수로 집계함. 31

6 정책제안대회를 위한 유동인구 분석 데이터 2월~6월, 24시간

1-1-2. 만 0세~24세 코로나 발생 후인 2020년 2월~6월 주중 유동인구, 지도 상에 표시된 원은 반경 300m 유동인구 평균값 상위 10개 지역



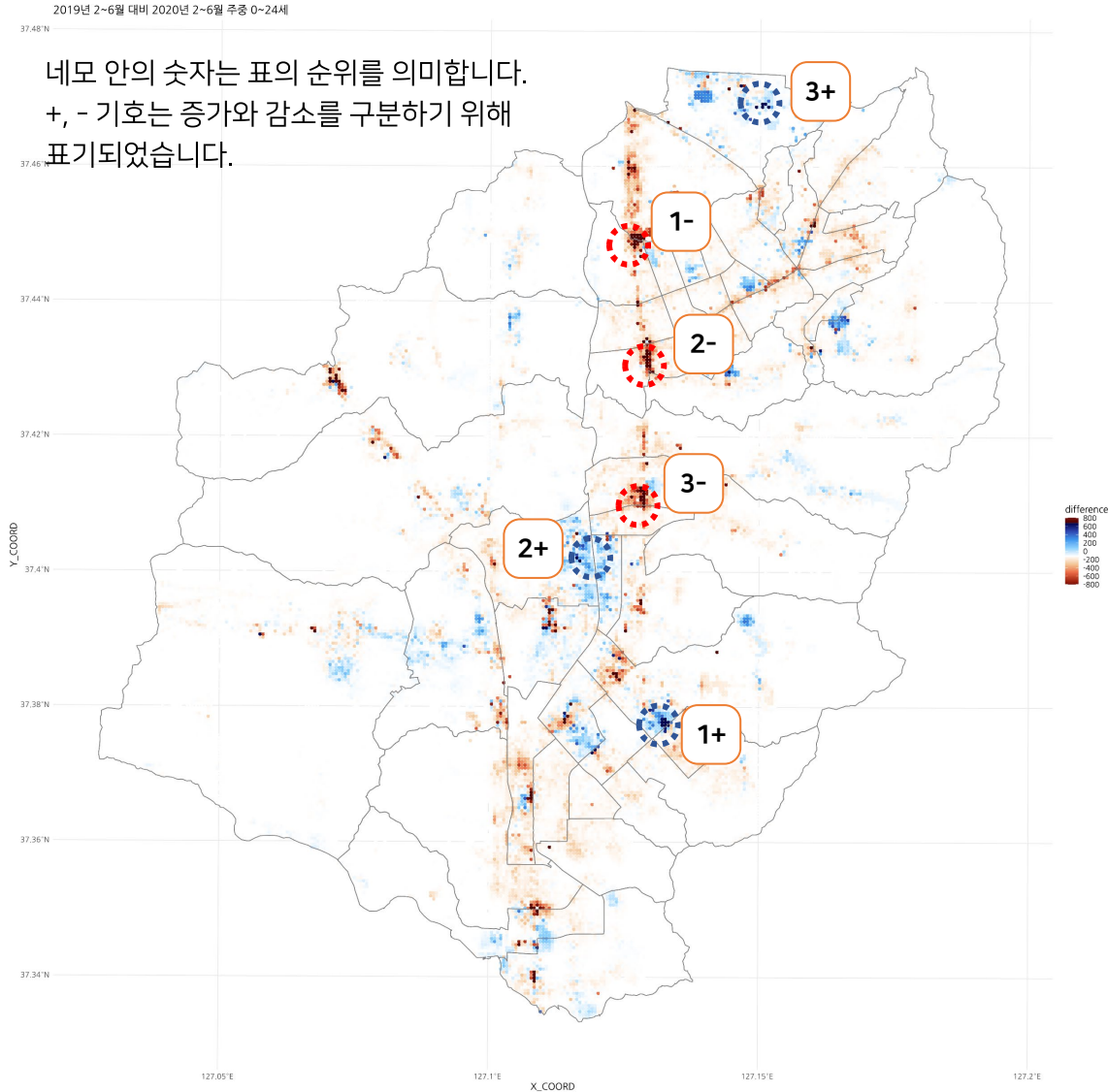
해당 조건의 성남시 유동인구 평균값: 57.5명

순위	반경 300m 평균 유동인구	성남시 평균대비	지역 위치
1	523.7	9.1배	분당어린이공원 인근 지역
2	487.4	8.5배	분당선 정자역 인근 지역
3	480.2	8.3배	분당선 미금역 인근 지역
4	464.4	8.1배	신분당선 판교역 인근 지역
5	438.2	7.6배	분당중앙공원 인근 지역
6	434.5	7.6배	분당선 서현역 인근 지역
7	405.0	7.0배	분당선 야탑역 인근 지역
8	388.4	6.8배	아름마을, 봇들마을 경계 인근 지역 분당·수서간 도시고속화도로 인근 지역
9	328.2	5.7배	분당선 수내역 인근 지역
10	322.7	5.6배	분당선 모란역 인근 지역

◦ 유의사항: 평균값을 산출할 때, 유동인구가 측정된 셀만 전체 셀 개수로 집계함.

6 정책제안대회를 위한 유동인구 분석 데이터 2월~6월, 24시간

1-1-3. 만 0세~24세 2019년과 2020년 2월~6월 주중 유동인구 변화, 지도 상에 표시된 원은 반경 300m 유동인구가 증가하거나 감소한 상위 3개 지역



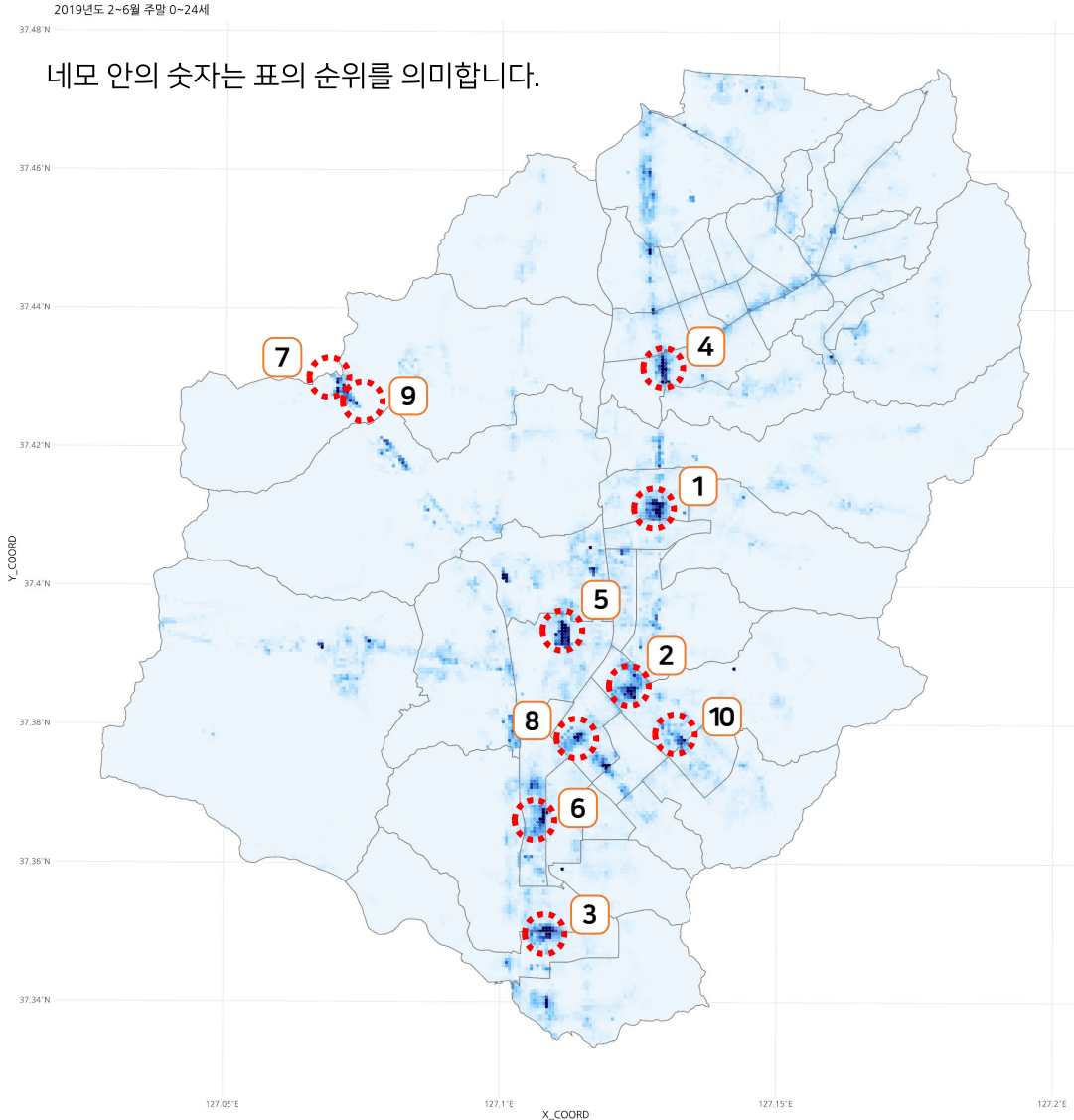
해당 조건의 성남시 유동인구 변화의 평균값: -13.3명

구분	순위	반경 300m 평균 유동인구 변화	지역 위치
증가	1+	180.2	분당어린이공원 인근 지역
	2+	179.2	아름마을, 봇들마을 경계 인근 지역 분당·수서간 도시고속화도로 인근 지역
	3+	149.7	위례파출소 옆 공터 인근 지역
감소	1-	-302.4	수인분당선 가천대역 인근 지역
	2-	-364.1	분당선 모란역 인근 지역
	3-	-389.1	분당선 야탑역 인근 지역

◦ 유의사항: 2019년과 2020년 각각의 지도에서는 평균값을 산출할 때, 유동인구가 측정된 셀만 전체 셀 개수로 집계하였으며, 2019년과 2020년의 변화를 측정할 때는 2019년에 집계된 셀과 2020년에 집계된 셀을 모두 합하여 전체 셀 개수로 집계하여, 변화의 평균값을 산출하였으므로 각 연도별로 산출한 평균값의 차이와 값이 다를 수 있음.

6 정책제안대회를 위한 유동인구 분석 데이터 2월~6월, 24시간

1-2-1. 만 0세~24세 코로나 발생 전인 2019년 2월~6월 주말 유동인구, 지도 상에 표시된 원은 반경 300m 유동인구 평균값 상위 10개 지역



해당 조건의 성남시 유동인구 평균값: 31.5명

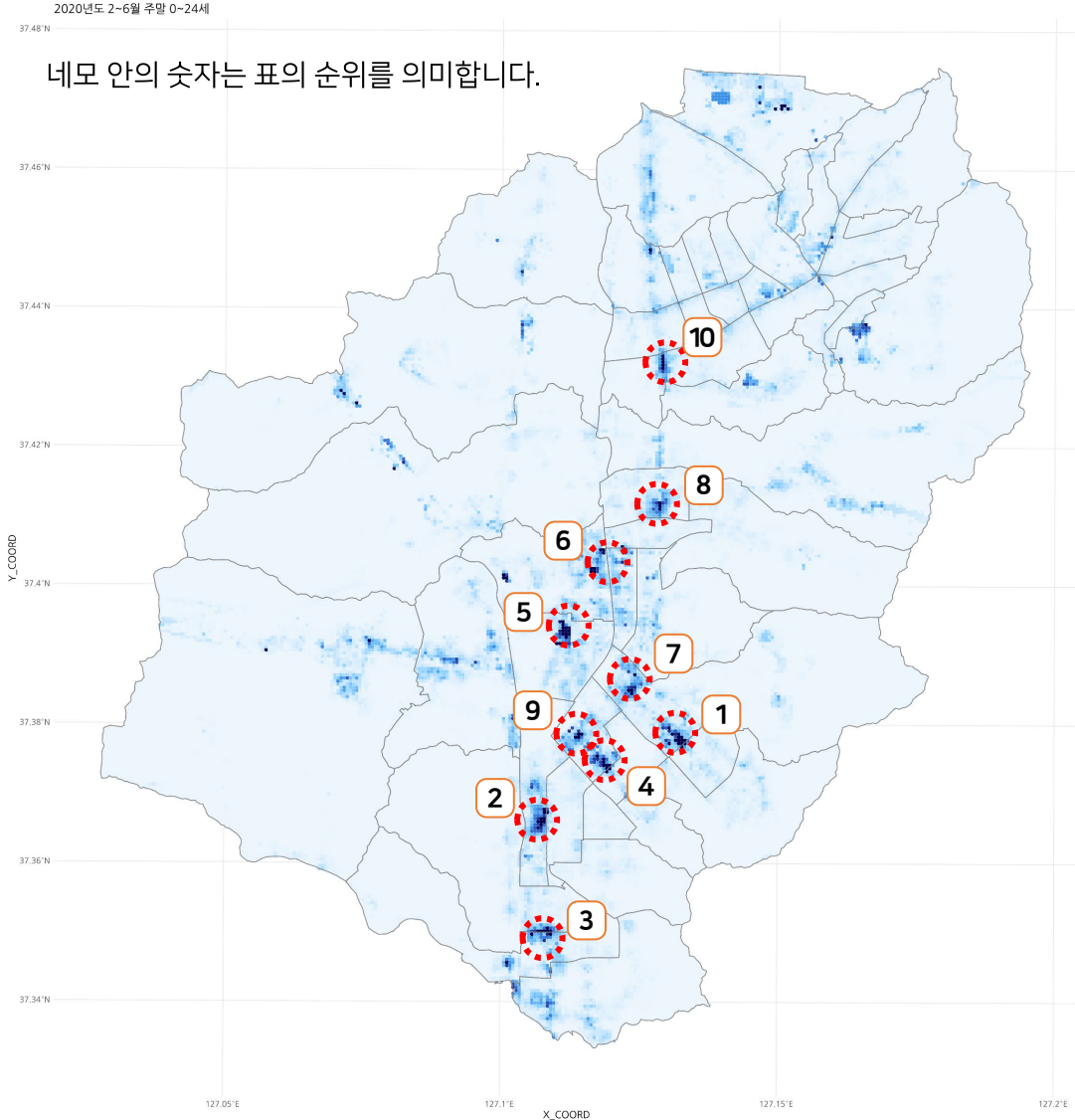
순위	반경 300m 평균 유동인구	성남시 평균대비	지역 위치
1	302.7	9.6배	분당선 야탑역 인근 지역
2	282.0	9.0배	분당선 서현역 인근 지역
3	269.5	8.6배	분당선 미금역 인근 지역
4	262.1	8.3배	분당선 모란역 인근 지역
5	253.9	8.1배	신분당선 판교역 인근 지역
6	250.0	7.9배	분당선 정자역 인근 지역
7	217.7	6.9배	경부고속도로 성남 고등동 진입로 인근 지역
8	169.4	5.4배	분당선 수내역 인근 지역
9	151.0	4.8배	경부고속도로 성남 고등동 진입로 인근 지역
10	149.2	4.7배	분당어린이공원 인근 지역

◦ 유의사항: 평균값을 산출할 때, 유동인구가 측정된 셀만 전체 셀 개수로 집계함.

6 정책제안대회를 위한 유동인구 분석 데이터

2월~6월, 24시간

1-2-2. 만 0세~24세 코로나 발생 후인 2020년 2월~6월 주말 유동인구, 지도 상에 표시된 원은 반경 300m 유동인구 평균값 상위 10개 지역



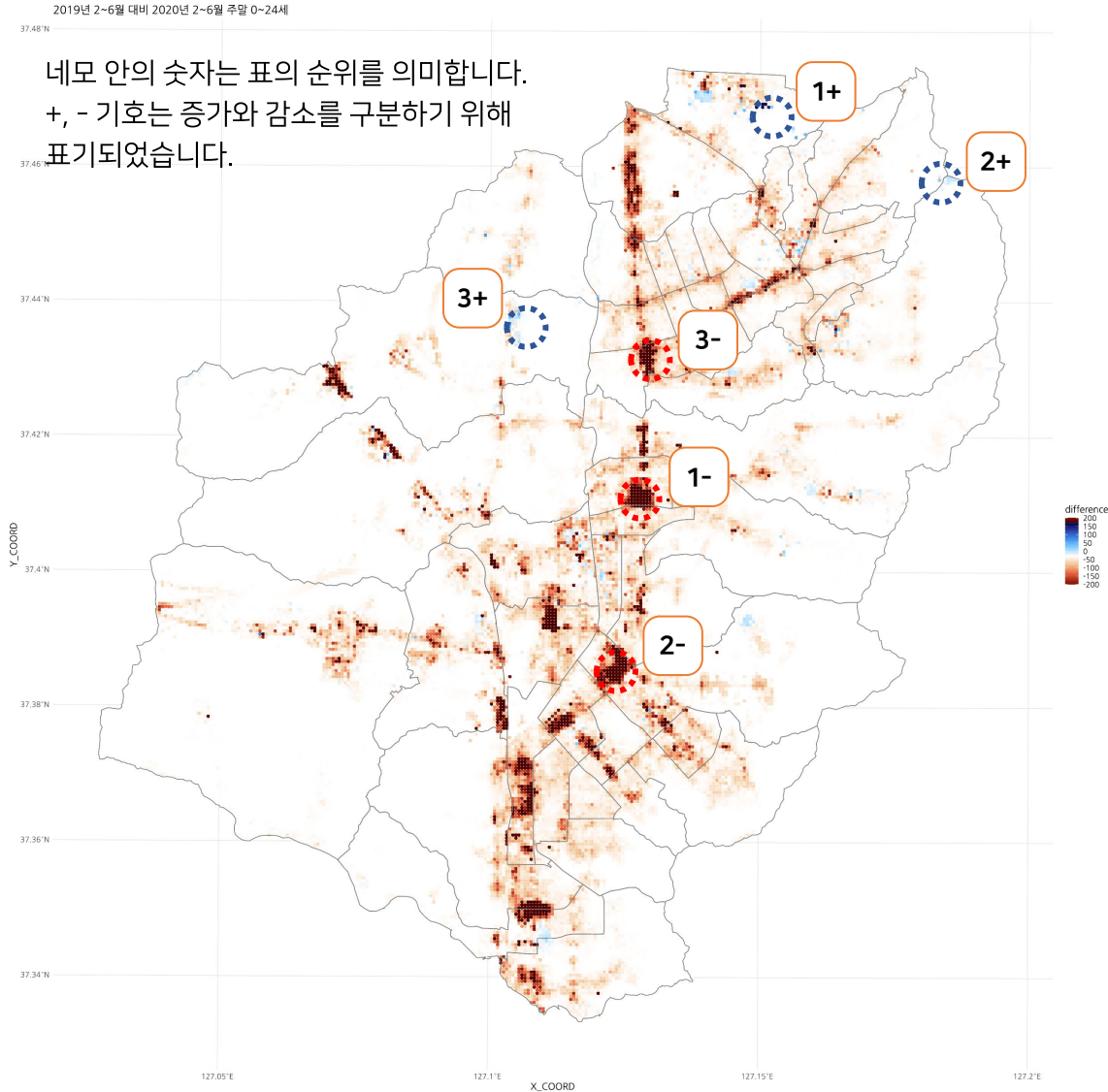
해당 조건의 성남시 유동인구 평균값: 9.7명

순위	반경 300m 평균 유동인구	성남시 평균대비	지역 위치
1	83.0	8.6배	분당어린이공원 인근 지역
2	70.4	7.3배	분당선 정자역 인근 지역
3	70.2	7.2배	분당선 미금역 인근 지역
4	68.6	7.1배	분당중앙공원 인근 지역
5	67.6	7.0배	신분당선 판교역 인근 지역
6	63.3	6.5배	아름마을, 봇들마을 경계 인근 지역 분당·수서간 도시고속화도로 인근 지역
7	61.8	6.4배	분당선 서현역 인근 지역
8	52.1	5.4배	분당선 야탑역 인근 지역
9	49.4	5.1배	분당선 수내역 인근 지역
10	47.7	4.9배	분당선 모란역 인근 지역

◦ 유의사항: 평균값을 산출할 때, 유동인구가 측정된 셀만 전체 셀 개수로 집계함.

6 정책제안대회를 위한 유동인구 분석 데이터 2월~6월, 24시간

1-2-3. 만 0세~24세 2019년과 2020년 2월~6월 주말 유동인구 변화, 지도 상에 표시된 원은 반경 300m 유동인구가 증가하거나 감소한 상위 3개 지역



해당 조건의 성남시 유동인구 변화의 평균값: -20.8명

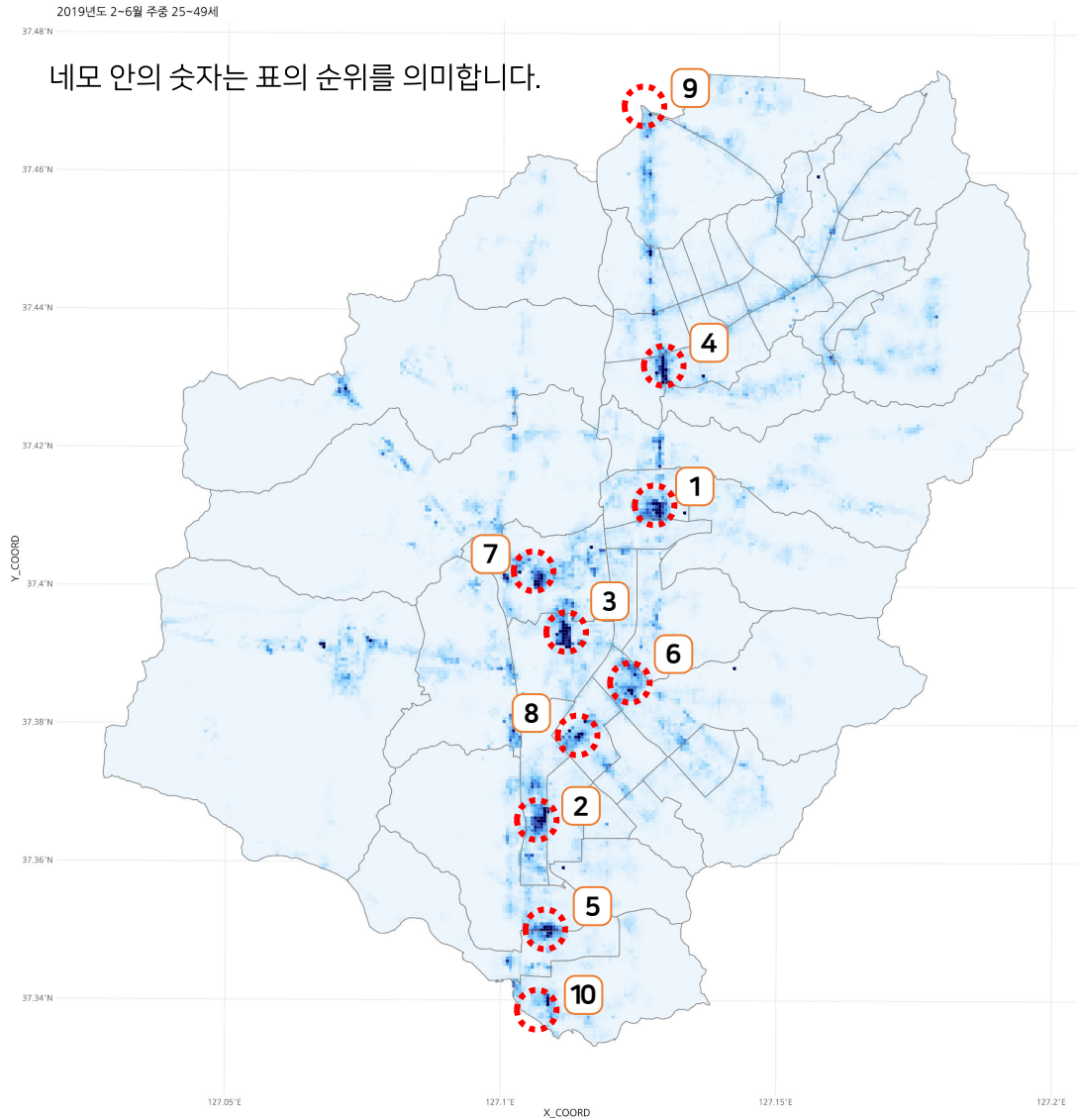
구분	순위	반경 300m 평균 유동인구 변화	지역 위치
증가	1+	17.8	위례파출소 옆 공터 인근 지역
	2+	10.0	신남성 산책로 인근 지역
	3+	7.8	성남 공군체력단련장 인근 지역
감소	1-	-251.4	분당선 야탑역 인근 지역
	2-	-221.8	분당선 서현역 인근 지역
	3-	-214.5	분당선 모란역 인근 지역

◦ 유의사항: 2019년과 2020년 각각의 지도에서는 평균값을 산출할 때, 유동인구가 측정된 셀만 전체 셀 개수로 집계하였으며, 2019년과 2020년의 변화를 측정할 때는 2019년에 집계된 셀과 2020년에 집계된 셀을 모두 합하여 전체 셀 개수로 집계하여, 변화의 평균값을 산출하였으므로 각 연도별로 산출한 평균값의 차이와 값이 다를 수 있음.

6 정책제안대회를 위한 유동인구 분석 데이터

2월~6월, 24시간

2-1-1. 만 25세~49세 코로나 발생 전인 2019년 2월~6월 주중 유동인구, 지도 상에 표시된 원은 반경 300m 유동인구 평균값 상위 10개 지역



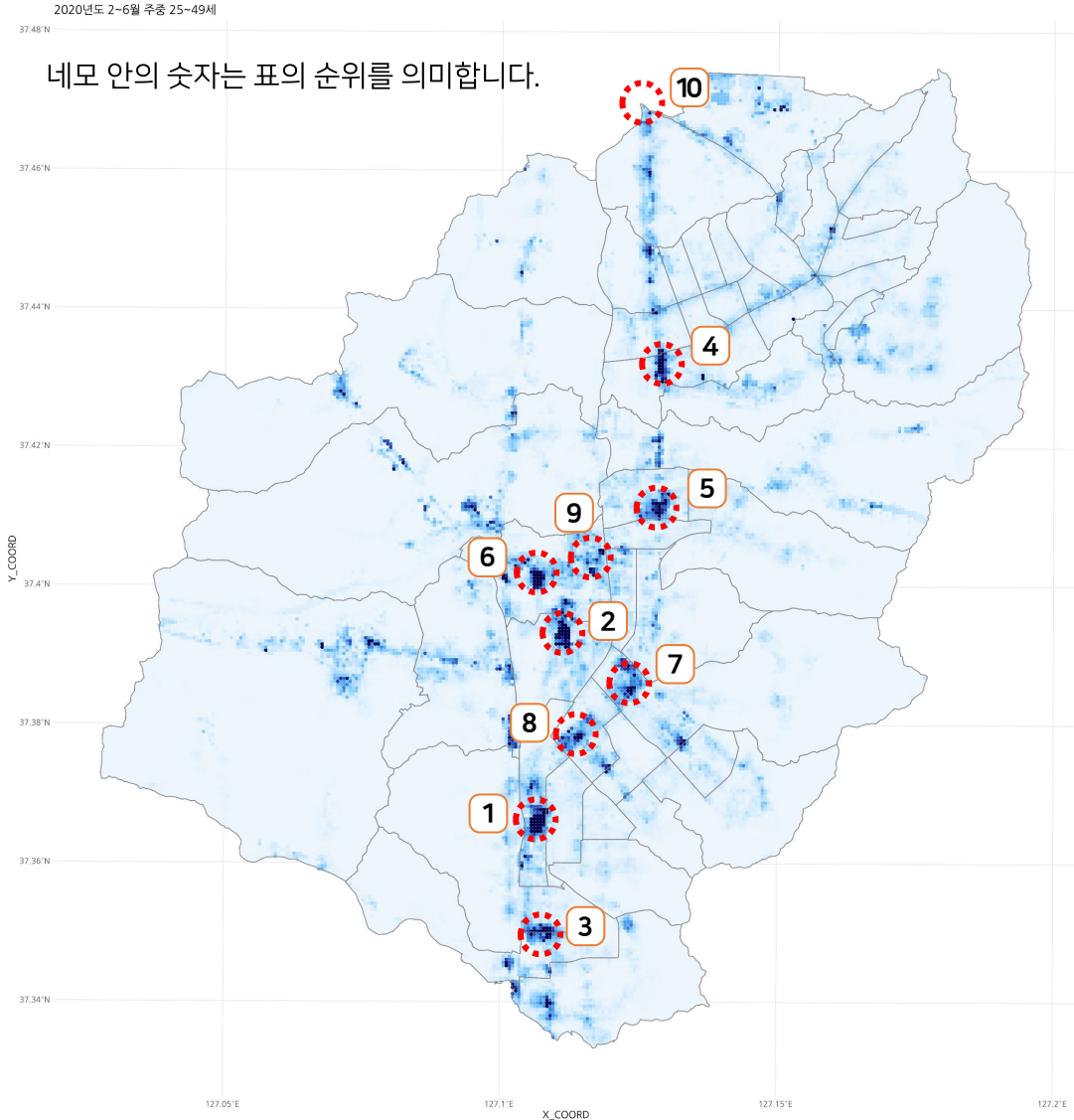
해당 조건의 성남시 유동인구 평균값: 384.6명

순위	반경 300m 평균 유동인구	성남시 평균대비	지역 위치
1	3423.6	8.9배	분당선 야탑역 인근 지역
2	3410.4	8.9배	분당선 정자역 인근 지역
3	3402.9	8.8배	신분당선 판교역 인근 지역
4	3307.2	8.6배	분당선 모란역 인근 지역
5	2802.6	7.3배	분당선 미금역 인근 지역
6	2629.9	6.8배	분당선 서현역 인근 지역
7	2515.0	6.5배	판교 테크노밸리 테크노중앙사거리 인근 지역
8	2109.9	5.5배	분당선 수내역 인근 지역
9	1729.3	4.5배	분당선 복정역 인근 지역 분당·수서간 도시고속화도로 인근 지역
10	1688.5	4.4배	분당선 오리역 인근 지역

◦ 유의사항: 평균값을 산출할 때, 유동인구가 측정된 셀만 전체 셀 개수로 집계함. 37

6 정책제안대회를 위한 유동인구 분석 데이터 2월~6월, 24시간

2-1-2. 만 25세~49세 코로나 발생 후인 2020년 2월~6월 주중 유동인구, 지도 상에 표시된 원은 반경 300m 유동인구 평균값 상위 10개 지역



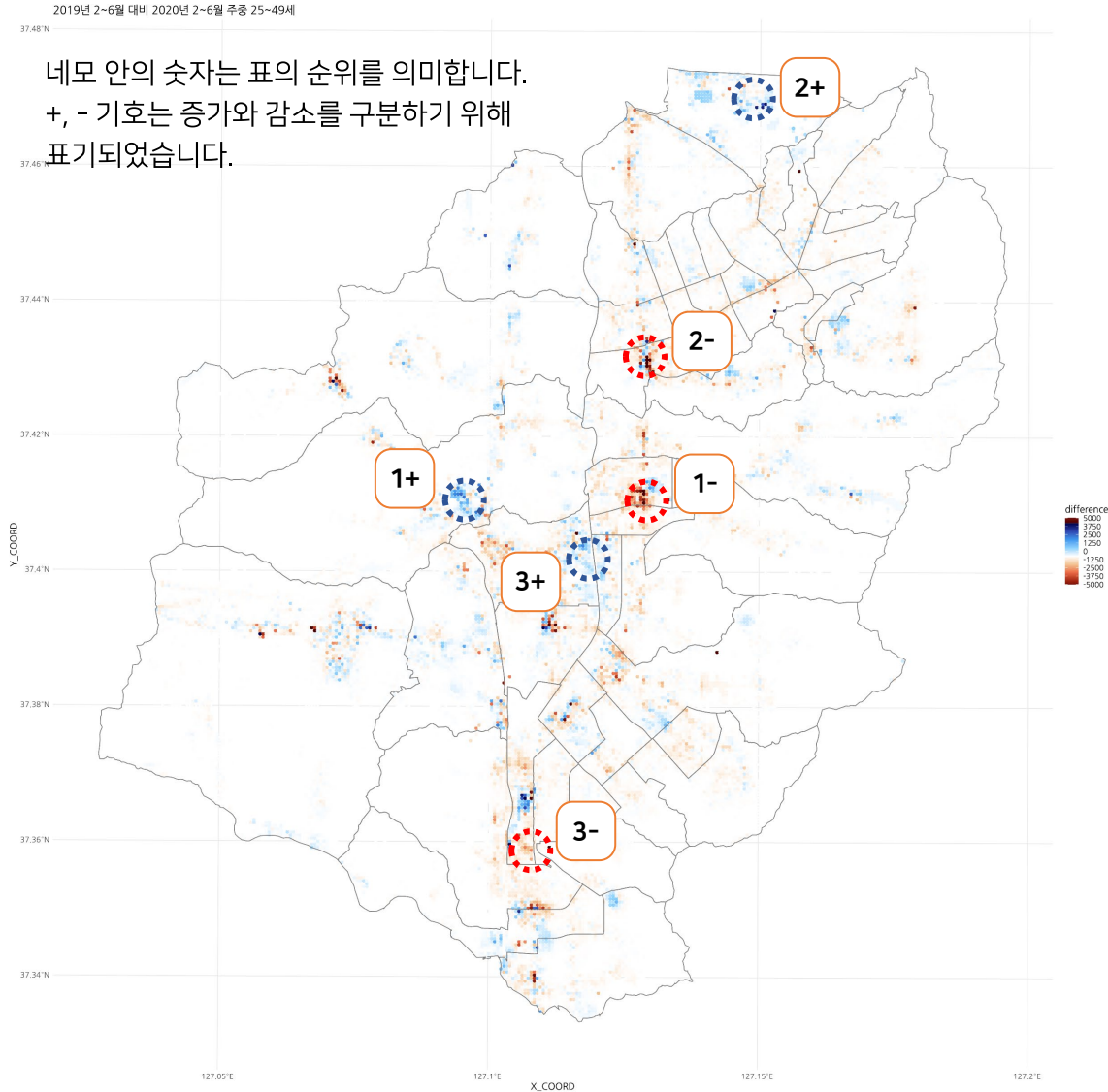
해당 조건의 성남시 유동인구 평균값: 352.4명

순위	반경 300m 평균 유동인구	성남시 평균대비	지역 위치
1	3543.2	10.1배	분당선 정자역 인근 지역
2	3146.6	8.9배	신분당선 판교역 인근 지역
3	2533.7	7.2배	분당선 미금역 인근 지역
4	2417.0	6.9배	분당선 모란역 인근 지역
5	2394.5	6.8배	분당선 야탑역 인근 지역
6	2232.2	6.3배	판교테크노밸리 테크노중앙사거리 인근 지역
7	2231.6	6.3배	분당선 서현역 인근 지역
8	1867.0	5.3배	분당선 수내역 인근 지역
9	1488.5	4.2배	판교테크노밸리 붓들사거리 인근 지역
10	1408.2	4.0배	분당선 복정역 인근 지역 분당·수서간 도시고속화도로 인근 지역

◦ 유의사항: 평균값을 산출할 때, 유동인구가 측정된 셀만 전체 셀 개수로 집계함.

6 정책제안대회를 위한 유동인구 분석 데이터 2월~6월, 24시간

2-1-3. 만 25세~49세 2019년과 2020년 2월~6월 주중 유동인구 변화, 지도 상에 표시된 원은 반경 300m 유동인구가 증가하거나 감소한 상위 3개 지역



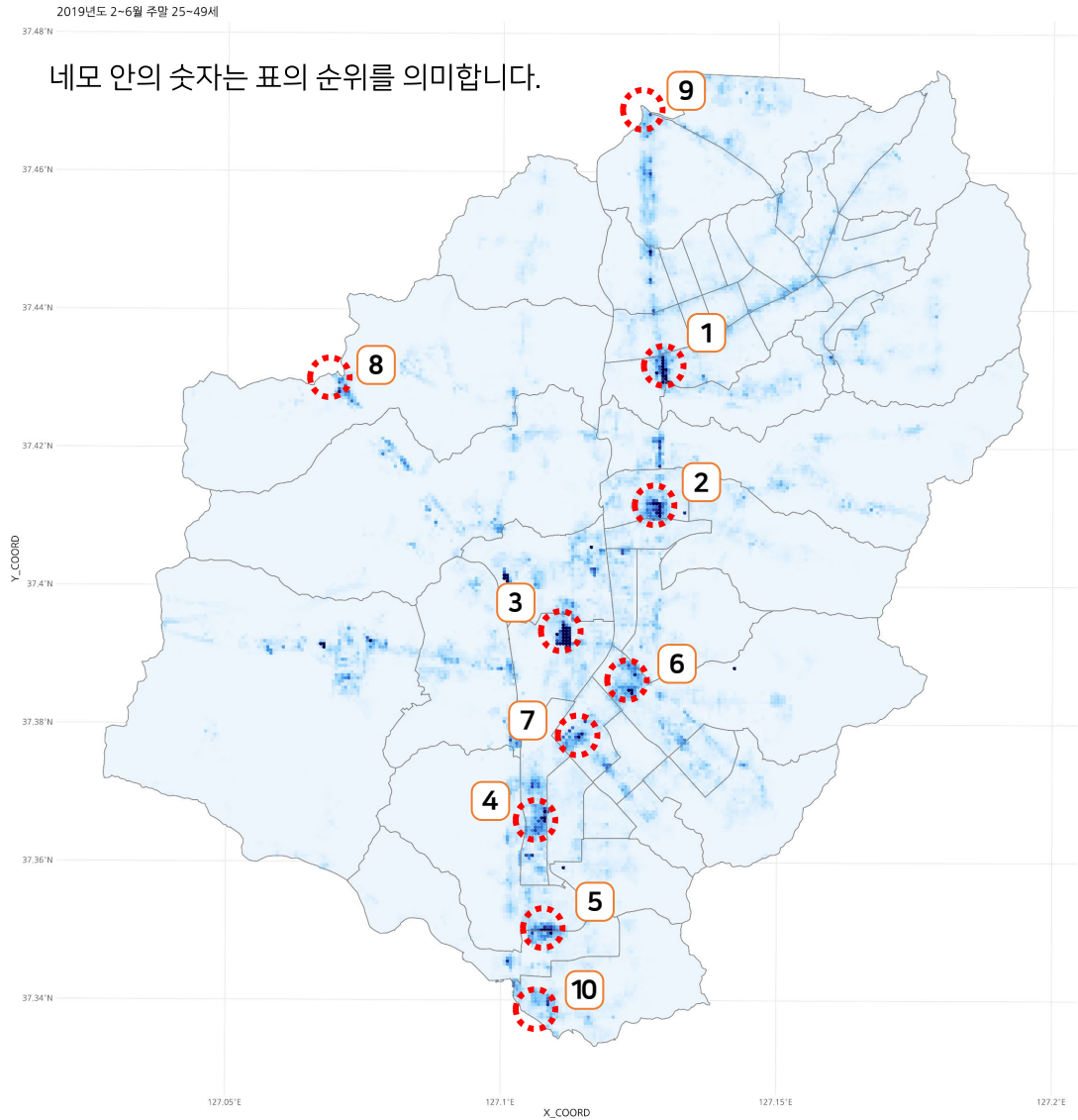
해당 조건의 성남시 유동인구 변화의 평균값: -32.8명

구분	순위	반경 300m 평균 유동인구 변화	지역 위치
증가	1+	585.2	판교제2테크노밸리 인근 지역 경부고속도로 인근 지역
	2+	457.6	위례파출소 옆 공터 인근 지역
	3+	380.4	아름마을, 봇들마을 경계 인근 지역 분당·수서간 도시고속화도로 인근 지역
감소	1-	-1191.4	분당선 야탑역 인근 지역
	2-	-890.2	분당선 모란역 인근 지역
	3-	-515.8	불정교 사거리 인근 지역

◦ 유의사항: 2019년과 2020년 각각의 지도에서는 평균값을 산출할 때, 유동인구가 측정된 셀만 전체 셀 개수로 집계하였으며, 2019년과 2020년의 변화를 측정할 때는 2019년에 집계된 셀과 2020년에 집계된 셀을 모두 합하여 전체 셀 개수로 집계하여, 변화의 평균값을 산출하였으므로 각 연도별로 산출한 평균값의 차이와 값이 다름.

6 정책제안대회를 위한 유동인구 분석 데이터 2월~6월, 24시간

2-2-1. 만 25세~49세 코로나 발생 전인 2019년 2월~6월 주말 유동인구, 지도 상에 표시된 원은 반경 300m 유동인구 평균값 상위 10개 지역



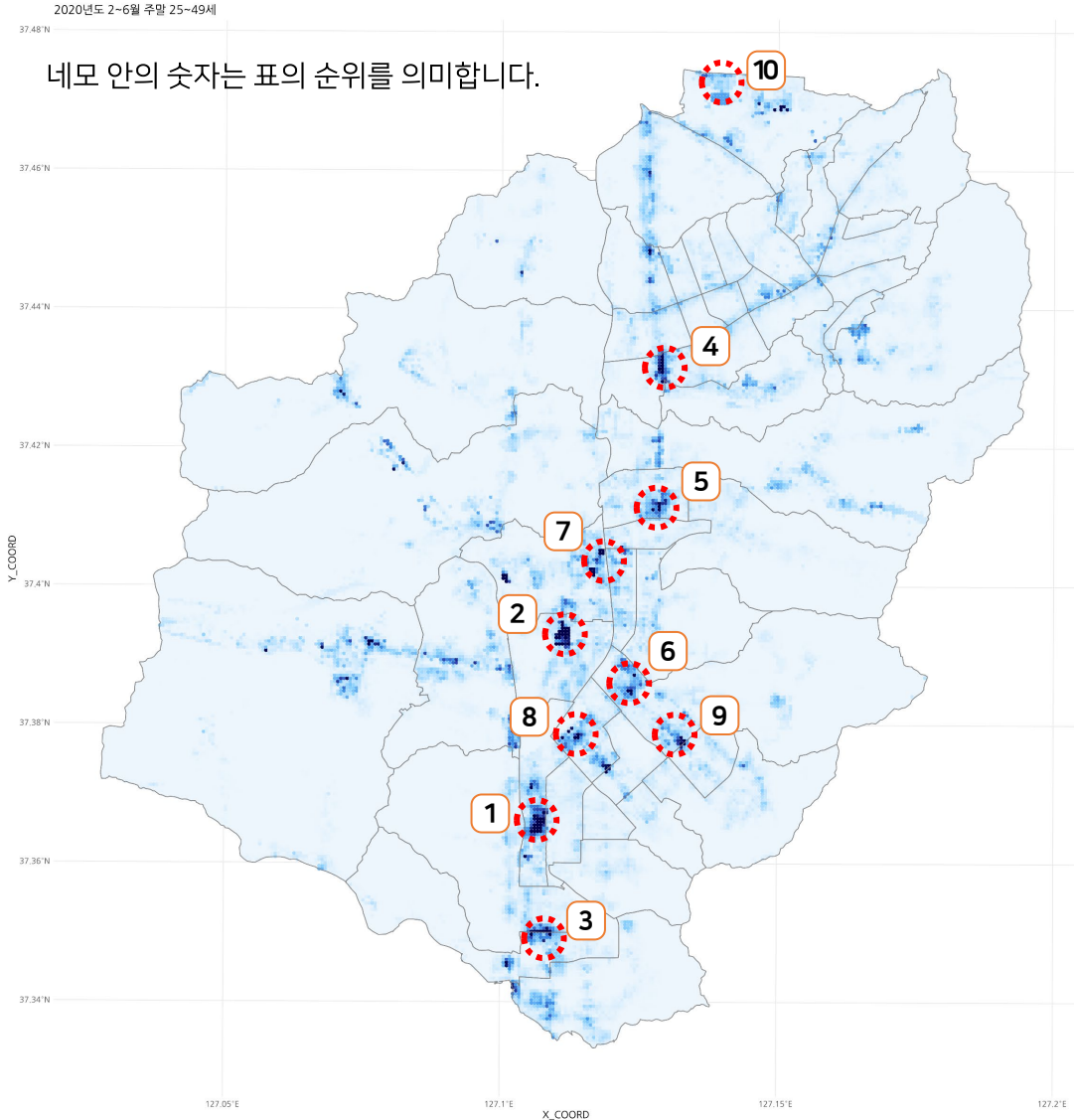
해당 조건의 성남시 유동인구 평균값: 133.5명

순위	반경 300m 평균 유동인구	성남시 평균대비	지역 위치
1	1157.6	8.7배	분당선 모란역 인근 지역
2	1096.9	8.2배	분당선 야탑역 인근 지역
3	1093.2	8.2배	신분당선 판교역 인근 지역
4	927.6	6.9배	분당선 정자역 인근 지역
5	908.0	6.8배	분당선 미금역 인근 지역
6	892.6	6.7배	분당선 서현역 인근 지역
7	639.5	4.8배	분당선 수내역 인근 지역
8	612.4	4.6배	경부고속도로 성남 고등동 진입로 인근 지역
9	610.9	4.6배	분당선 복정역 인근 지역 분당·수서간 도시고속화도로 인근 지역
10	558.2	4.2배	분당선 오리역 인근 지역

◦ 유의사항: 평균값을 산출할 때, 유동인구가 측정된 셀만 전체 셀 개수로 집계함. 40

6 정책제안대회를 위한 유동인구 분석 데이터 2월~6월, 24시간

2-2-2. 만 25세~49세 코로나 발생 후인 2020년 2월~6월 주말 유동인구, 지도 상에 표시된 원은 반경 300m 유동인구 평균값 상위 10개 지역



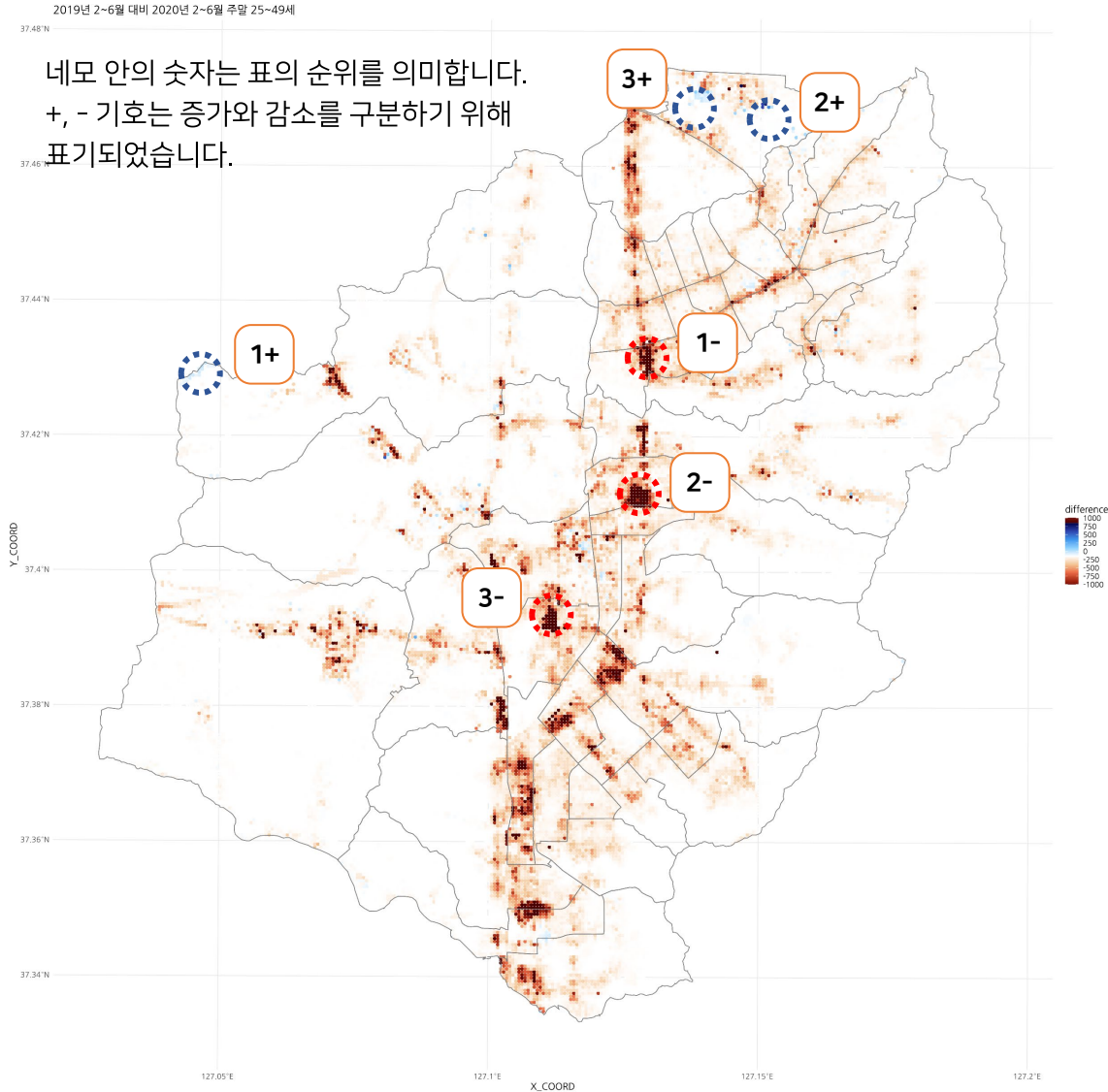
해당 조건의 성남시 유동인구 평균값: 45.0명

순위	반경 300m 평균 유동인구	성남시 평균대비	지역 위치
1	368.7	8.2배	분당선 정자역 인근 지역
2	279.4	6.2배	신분당선 판교역 인근 지역
3	315.4	7.0배	분당선 미금역 인근 지역
4	202.8	4.5배	분당선 모란역 인근 지역
5	219.3	4.9배	분당선 야탑역 인근 지역
6	232.8	5.2배	분당선 서현역 인근 지역
7	234.1	5.2배	아름마을, 봇들마을 경계 인근 지역 분당·수서간 도시고속화도로 인근 지역
8	275.4	6.1배	분당선 수내역 인근 지역
9	183.3	4.1배	분당어린이공원 인근 지역
10	171.9	3.8배	위례중앙광장 인근 지역

◦ 유의사항: 평균값을 산출할 때, 유동인구가 측정된 셀만 전체 셀 개수로 집계함.

6 정책제안대회를 위한 유동인구 분석 데이터 2월~6월, 24시간

2-2-3. 만 25세~49세 2019년과 2020년 2월~6월 주말 유동인구 변화, 지도 상에 표시된 원은 반경 300m 유동인구가 증가하거나 감소한 상위 3개 지역



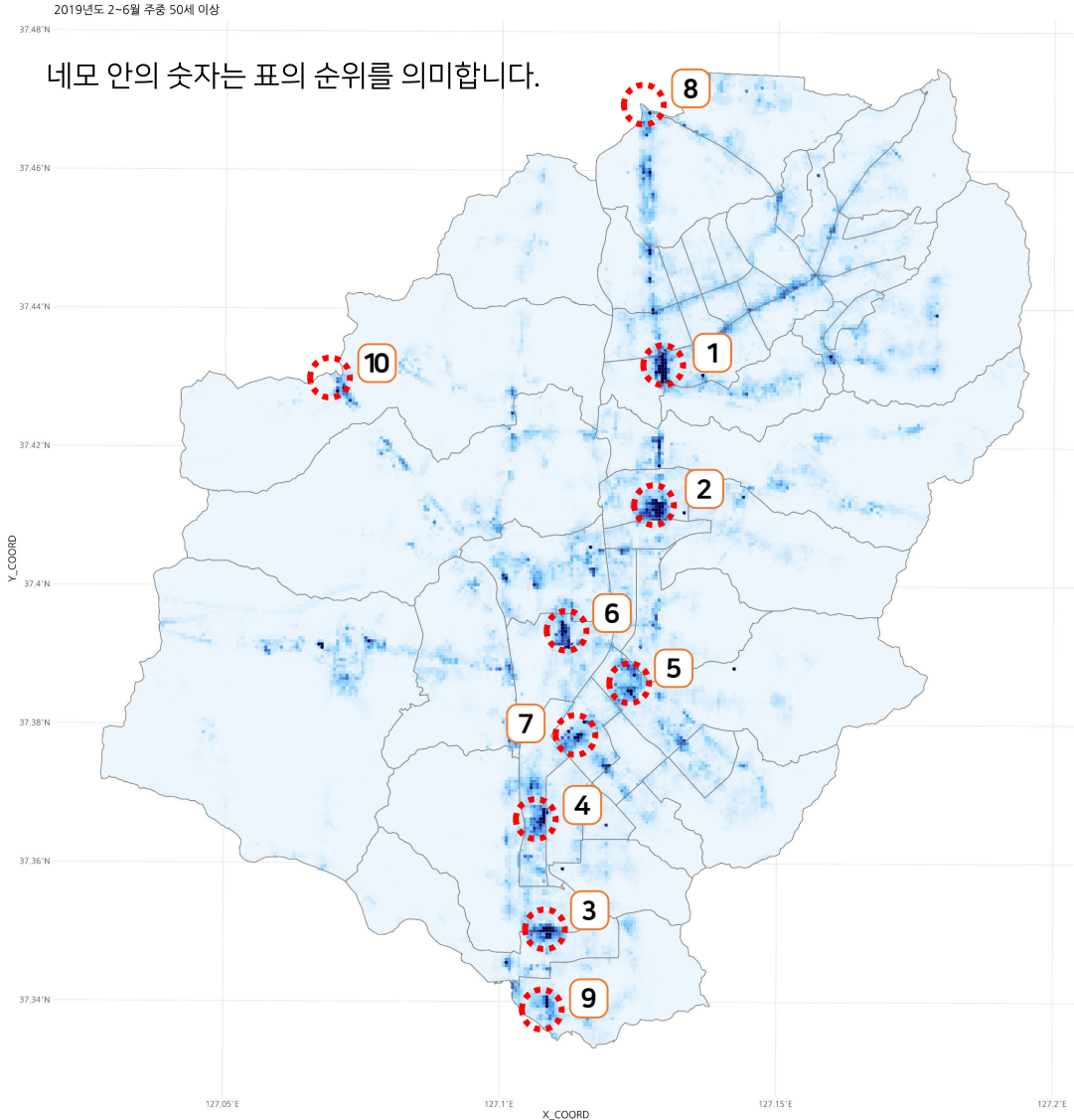
해당 조건의 성남시 유동인구 변화의 평균값: -84.8명

구분	순위	반경 300m 평균 유동인구 변화	지역 위치
증가	1+	52.8	청계산 매봉 산책로 인근 지역
	2+	52.0	위례파출소 옆 공터 인근 지역
	3+	16.3	위례역사수변공원 인근 지역
감소	1-	-882.2	분당선 모란역 인근 지역
	2-	-864.9	분당선 야탑역 인근 지역
	3-	-779.0	신분당선 판교역 인근 지역

◦ 유의사항: 2019년과 2020년 각각의 지도에서는 평균값을 산출할 때, 유동인구가 측정된 셀만 전체 셀 개수로 집계하였으며, 2019년과 2020년의 변화를 측정할 때는 2019년에 집계된 셀과 2020년에 집계된 셀을 모두 합하여 전체 셀 개수로 집계하여, 변화의 평균값을 산출하였으므로 각 연도별로 산출한 평균값의 차이와 값이 다름.

6 정책제안대회를 위한 유동인구 분석 데이터 2월~6월, 24시간

3-1-1. 만 50세 이상 코로나 발생 전인 2019년 2월~6월 주중 유동인구, 지도 상에 표시된 원은 반경 300m 유동인구 평균값 상위 10개 지역



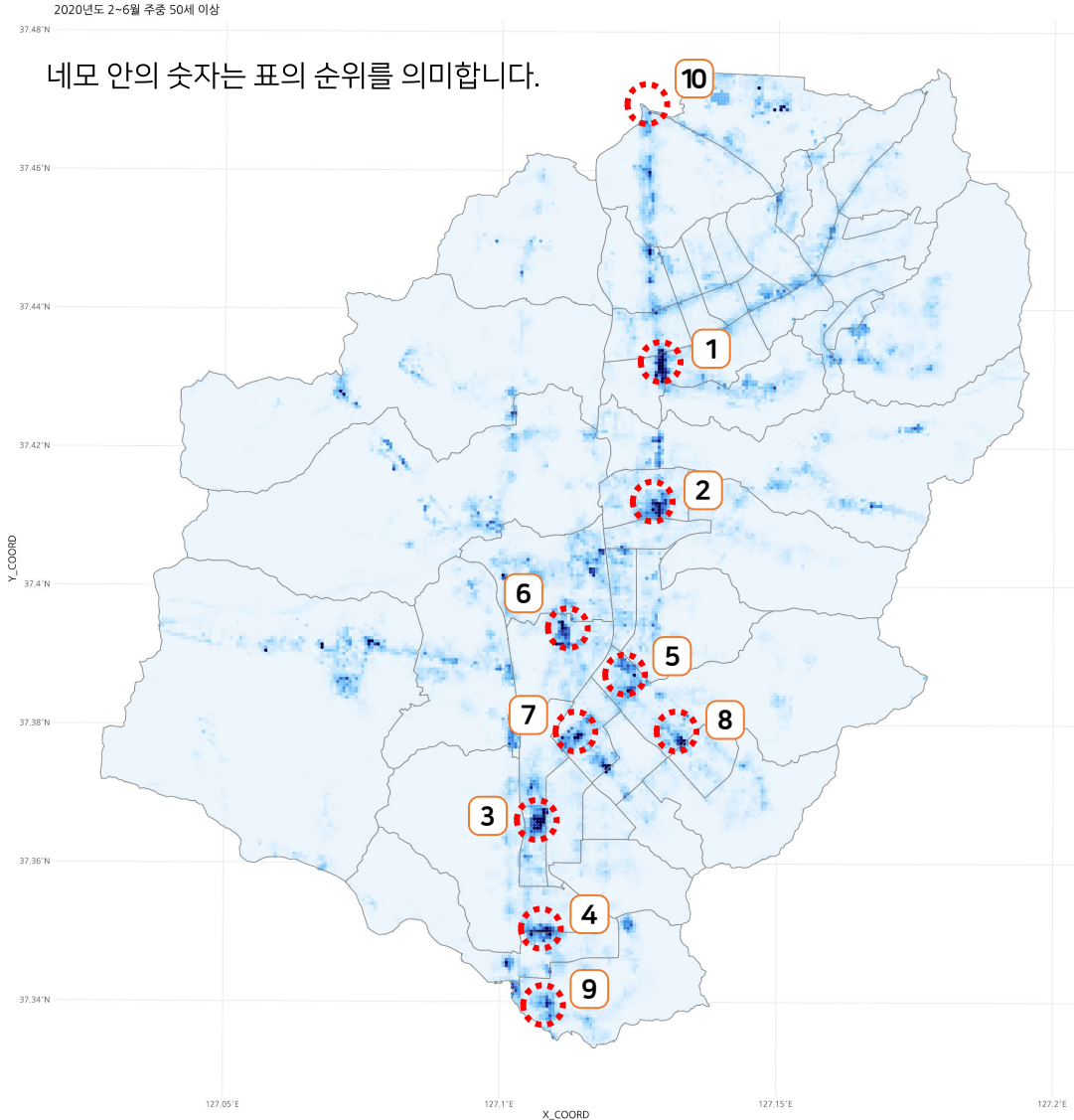
해당 조건의 성남시 유동인구 평균값: 291.6명

순위	반경 300m 평균 유동인구	성남시 평균대비	지역 위치
1	3007.5	10.3배	분당선 모란역 인근
2	2783.5	9.5배	분당선 야탑역 인근
3	2070.8	7.1배	분당선 미금역 인근
4	1929.7	6.6배	분당선 정자역 인근
5	1825.9	6.3배	분당선 서현역 인근
6	1625.8	5.6배	신분당선 판교역 인근
7	1536.5	5.3배	분당선 수내역 인근
8	1311.1	4.5배	분당선 복정역 인근 지역 분당·수서간 도시고속화도로 인근 지역
9	1277.8	4.4배	분당선 오리역 인근
10	1119.9	3.8배	경부고속도로 성남 고등동 진입로 인근 지역

◦ 유의사항: 평균값을 산출할 때, 유동인구가 측정된 셀만 전체 셀 개수로 집계함.

6 정책제안대회를 위한 유동인구 분석 데이터 2월~6월, 24시간

3-1-2. 만 50세 이상 코로나 발생 후인 2020년 2월~6월 주중 유동인구, 지도 상에 표시된 원은 반경 300m 유동인구 평균값 상위 10개 지역



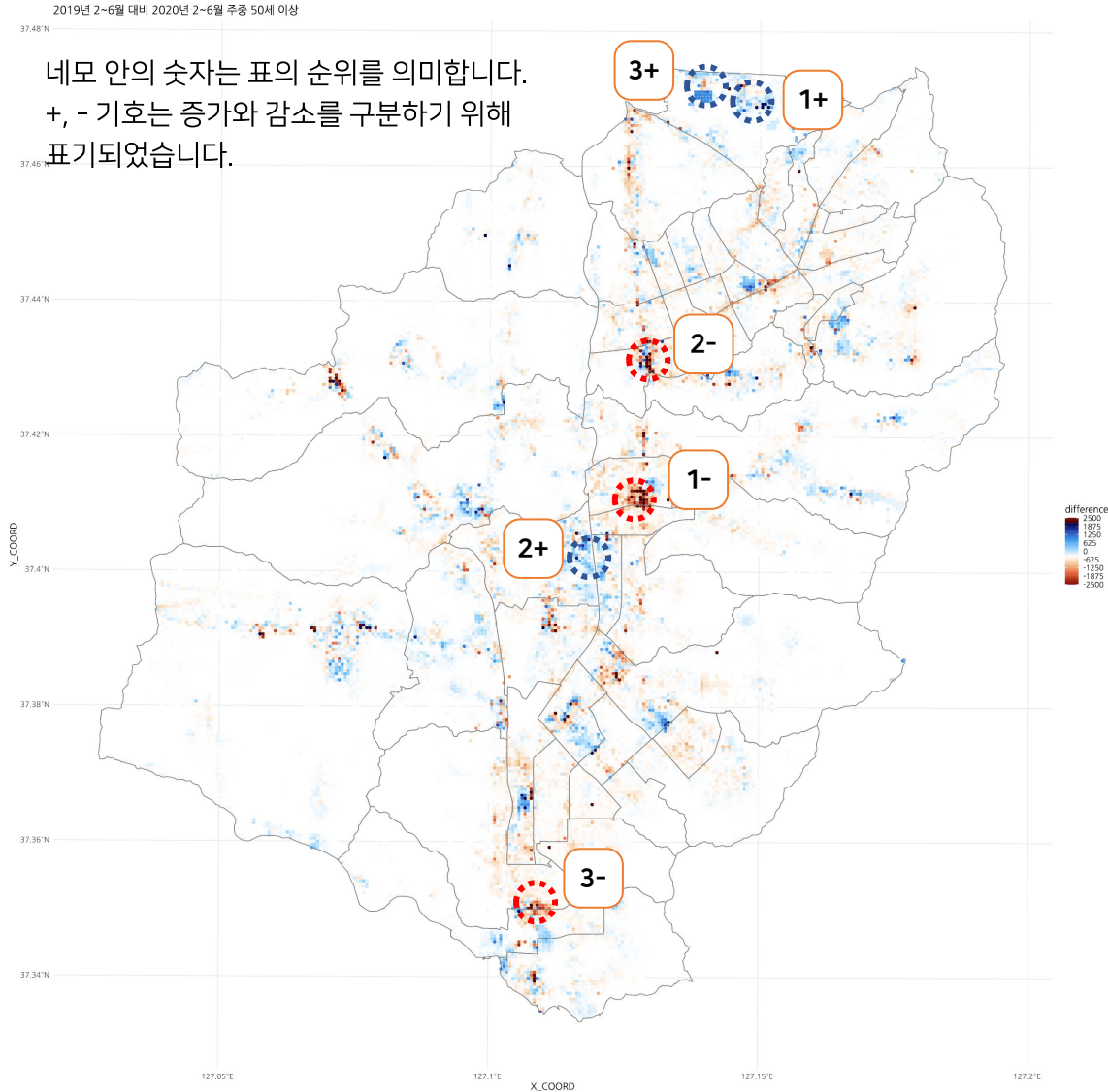
해당 조건의 성남시 유동인구 평균값: 283.0명

순위	반경 300m 평균 유동인구	성남시 평균대비	지역 위치
1	2305.0	8.1배	분당선 모란역 인근 지역
2	1951.1	6.9배	분당선 야탑역 인근 지역
3	1924.2	6.8배	분당선 정자역 인근 지역
4	1813.7	6.4배	분당선 미금역 인근 지역
5	1536.7	5.4배	분당선 서현역 인근 지역
6	1456.1	5.1배	신분당선 판교역 인근 지역
7	1367.4	4.8배	분당선 수내역 인근 지역
8	1178.2	4.2배	분당어린이공원 인근 지역
9	1163.3	4.1배	분당선 오리역 인근 지역
10	1151.0	4.1배	분당선 복정역 인근 지역 분당·수서간 도시고속화도로 인근 지역

◦ 유의사항: 평균값을 산출할 때, 유동인구가 측정된 셀만 전체 셀 개수로 집계함.

6 정책제안대회를 위한 유동인구 분석 데이터 2월~6월, 24시간

3-1-3. 만 50세 이상 2019년과 2020년 2월~6월 주중 유동인구 변화, 지도 상에 표시된 원은 반경 300m 유동인구가 증가하거나 감소한 상위 3개 지역



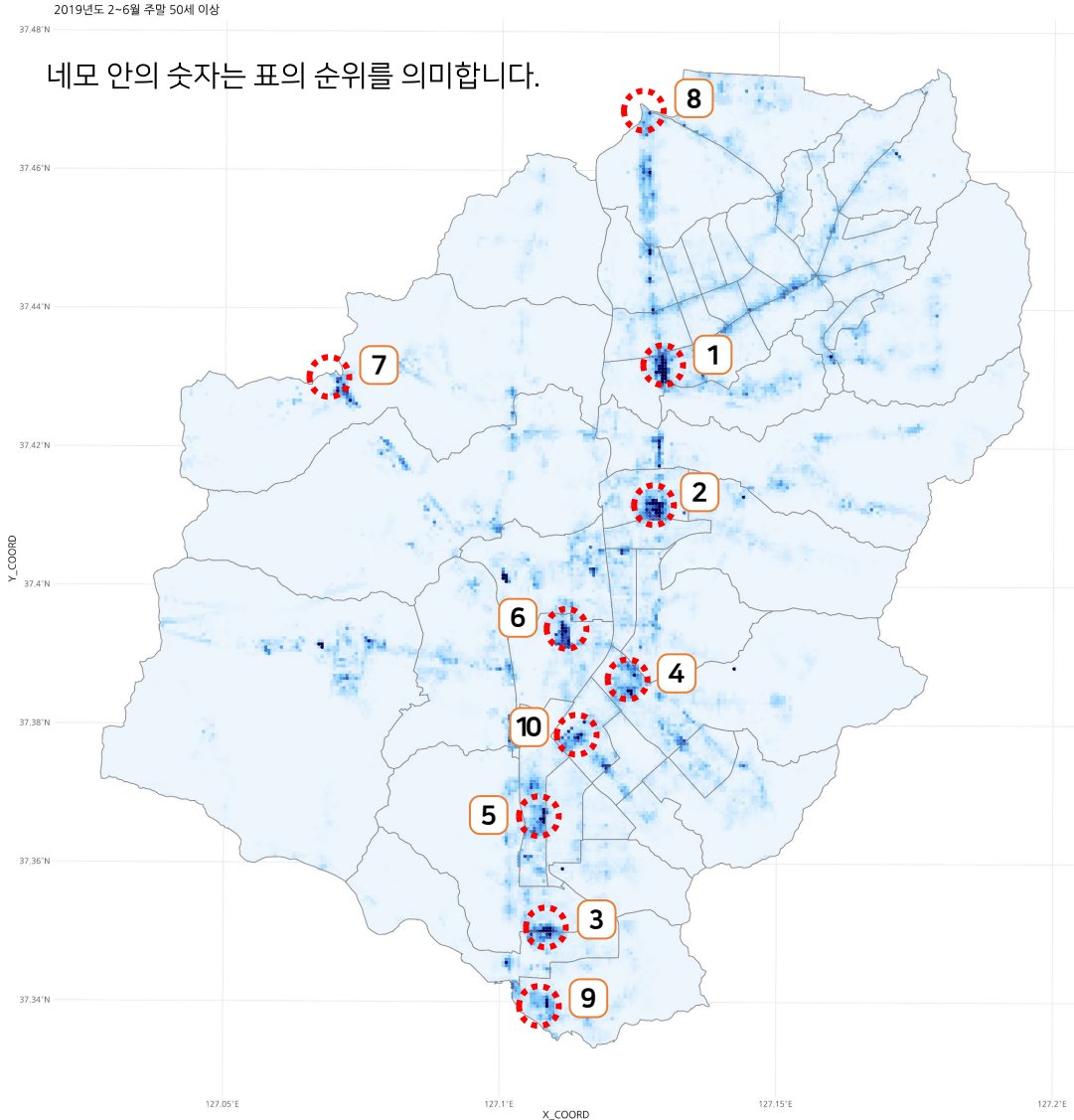
해당 조건의 성남시 유동인구 변화의 평균값: -10.7명

구분	순위	반경 300m 평균 유동인구 변화	지역 위치
증가	1+	467.8	위례파출소 옆 공터 인근 지역
	2+	366.7	아름마을, 봇들마을 경계 인근 지역 분당·수서간 도시고속화도로 인근 지역
	3+	329.5	위례중앙광장 인근 지역
감소	1-	-918.4	분당선 야탑역 인근 지역
	2-	-713.8	분당선 모란역 인근 지역
	3-	-394.1	분당선 미금역 인근 지역

◦ 유의사항: 2019년과 2020년 각각의 지도에서는 평균값을 산출할 때, 유동인구가 측정된 셀만 전체 셀 개수로 집계하였으며, 2019년과 2020년의 변화를 측정할 때는 2019년에 집계된 셀과 2020년에 집계된 셀을 모두 합하여 전체 셀 개수로 집계하여, 변화의 평균값을 산출하였으므로 각 연도별로 산출한 평균값의 차이와 값이 다름.

6 정책제안대회를 위한 유동인구 분석 데이터 2월~6월, 24시간

3-2-1. 만 50세 이상 코로나 발생 전인 2019년 2월~6월 주말 유동인구, 지도 상에 표시된 원은 반경 300m 유동인구 평균값 상위 10개 지역



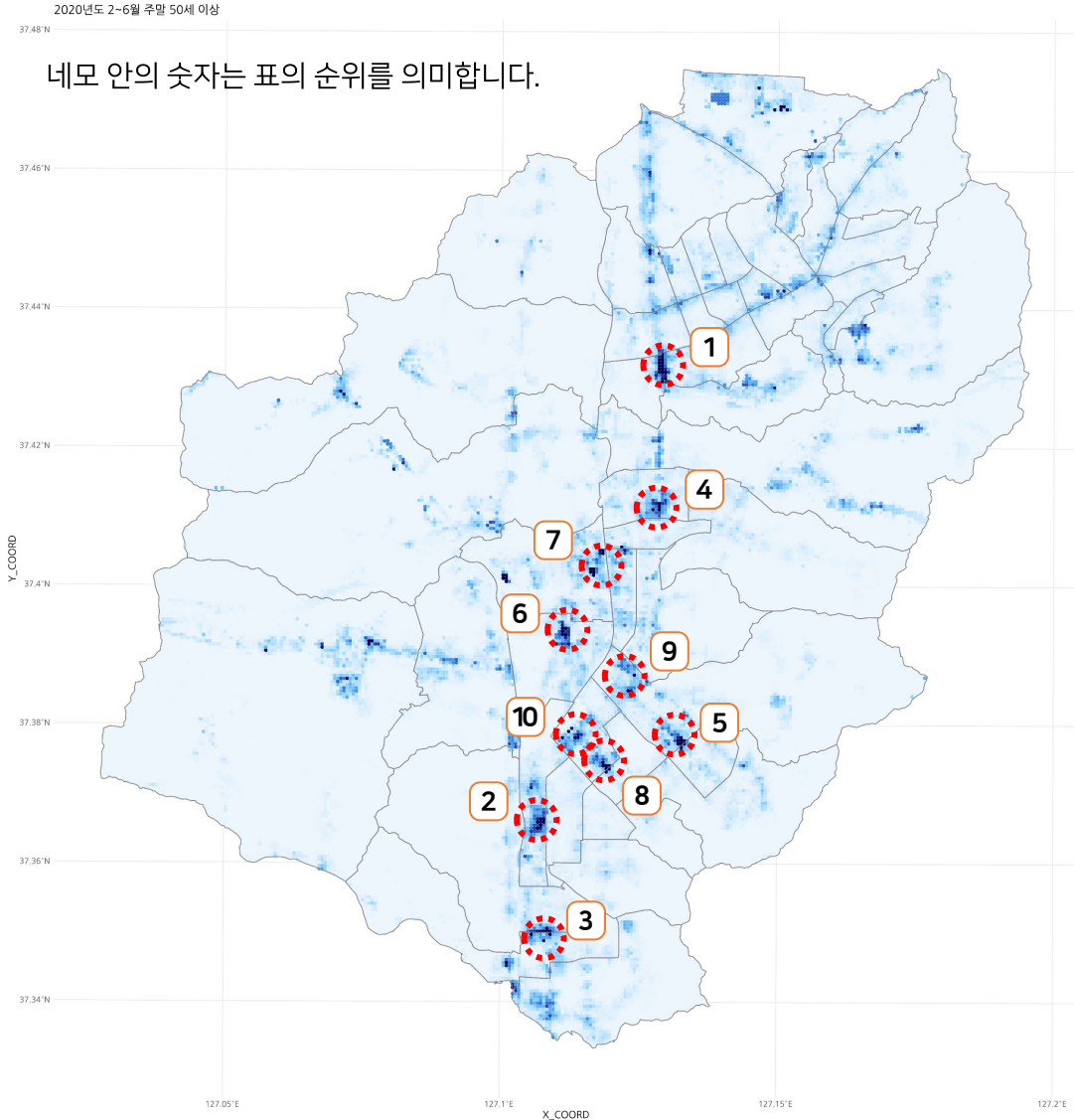
해당 조건의 성남시 유동인구 평균값: 108.5명

순위	반경 300m 평균 유동인구	성남시 평균대비	지역 위치
1	1144.1	10.5배	분당선 모란역 인근 지역
2	962.7	8.9배	분당선 야탑역 인근 지역
3	679.3	6.3배	분당선 미금역 인근 지역
4	630.5	5.8배	분당선 서현역 인근 지역
5	596.6	5.5배	분당선 정자역 인근 지역
6	585.2	5.4배	신분당선 판교역 인근 지역
7	557.6	5.1배	경부고속도로 성남 고등동 진입로 인근 지역
8	471.2	4.3배	분당선 복정역 인근 지역 분당·수서간 도시고속화도로 인근 지역
9	470.8	4.3배	분당선 오리역 인근 지역
10	465.6	4.3배	분당선 수내역 인근 지역

◦ 유의사항: 평균값을 산출할 때, 유동인구가 측정된 셀만 전체 셀 개수로 집계함.

6 정책제안대회를 위한 유동인구 분석 데이터 2월~6월, 24시간

3-2-2. 만 50세 이상 코로나 발생 후인 2020년 2월~6월 주말 유동인구, 지도 상에 표시된 원은 반경 300m 유동인구 평균값 상위 10개 지역



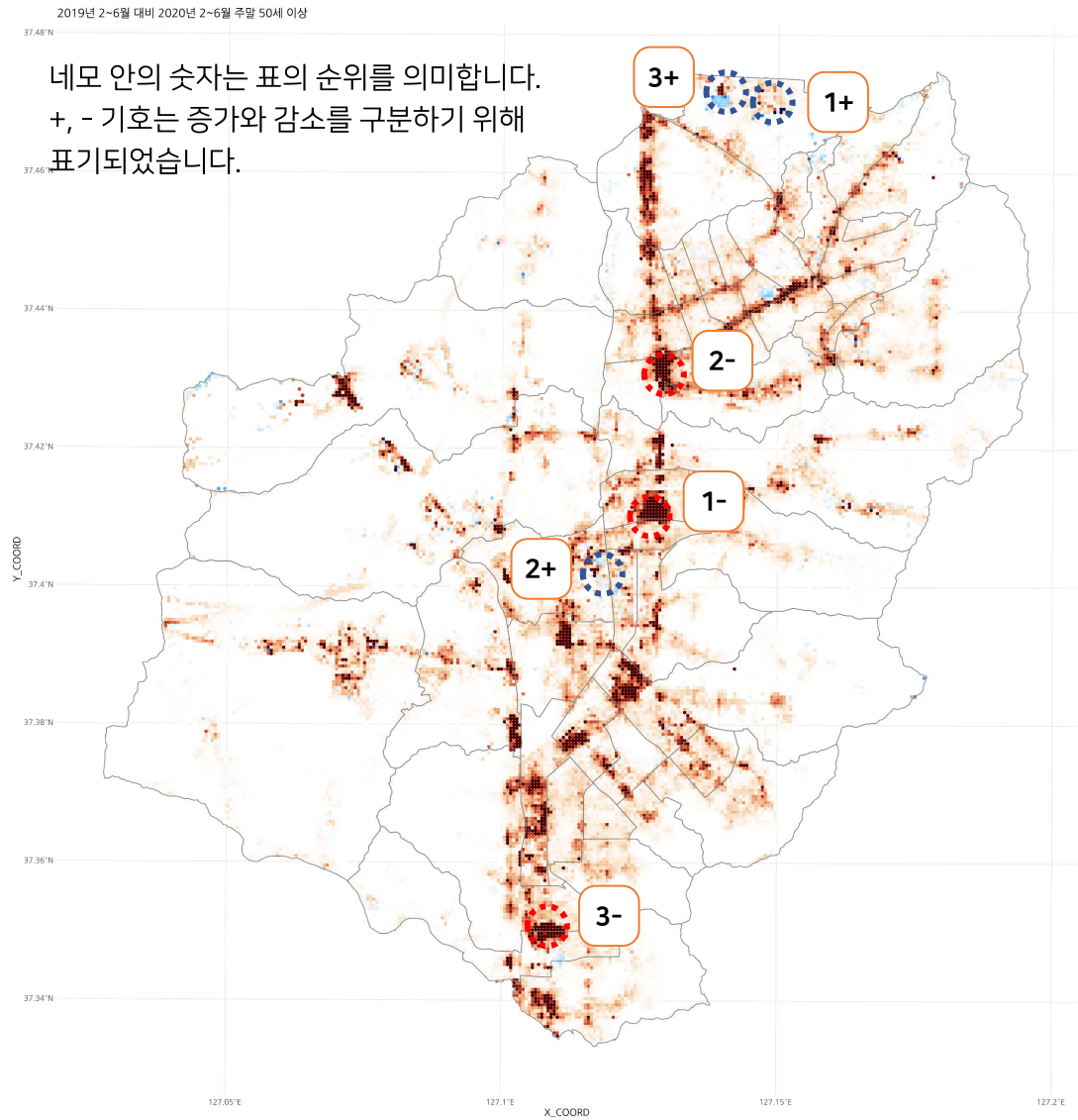
해당 조건의 성남시 유동인구 평균값: 38.6명

순위	반경 300m 평균 유동인구	성남시 평균대비	지역 위치
1	269.5	7.0배	분당선 모란역 인근 지역
2	208.4	5.4배	분당선 정자역 인근 지역
3	197.7	5.1배	분당선 미금역 인근 지역
4	194.8	5.0배	분당선 야탑역 인근 지역
5	179.1	4.6배	분당어린이공원 인근 지역
6	170.0	4.4배	신분당선 판교역 인근 지역
7	170.0	4.4배	아름마을, 봇들마을 경계 인근 지역 분당·수서간 도시고속화도로 인근 지역
8	161.1	4.2배	분당중앙공원 인근 지역
9	160.6	4.2배	분당선 서현역 인근 지역
10	153.8	4.0배	분당선 수내역 인근 지역

◦ 유의사항: 평균값을 산출할 때, 유동인구가 측정된 셀만 전체 셀 개수로 집계함.

6 정책제안대회를 위한 유동인구 분석 데이터 2월~6월, 24시간

3-2-3. 만 50세 이상 2019년과 2020년 2월~6월 주말 유동인구 변화, 지도 상에 표시된 원은 반경 300m 유동인구가 증가하거나 감소한 상위 3개 지역



해당 조건의 성남시 유동인구 변화의 평균값: -67.1명

구분	순위	반경 300m 평균 유동인구 변화	지역 위치
증가	1+	53.9	위례고운초등학교 인근 지역
	2+	30.4	아름마을, 봇들마을 경계 인근 지역 분당·수서간 도시고속화도로 인근 지역
	3+	25.0	위례중앙광장 인근 지역
감소	1-	-874.6	분당선 야탑역 인근 지역
	2-	-768.6	분당선 모란역 인근 지역
	3-	-486.7	분당선 미금역 인근 지역

◦ 유의사항: 2019년과 2020년 각각의 지도에서는 평균값을 산출할 때, 유동인구가 측정된 셀만 전체 셀 개수로 집계하였으며, 2019년과 2020년의 변화를 측정할 때는 2019년에 집계된 셀과 2020년에 집계된 셀을 모두 합하여 전체 셀 개수로 집계하여, 변화의 평균값을 산출하였으므로 각 연도별로 산출한 평균값의 차이와 값이 다름.

6 정책제안대회를 위한 유동인구 분석 데이터

빅데이터 기반 정책제안을 위한 팁

1. 대회를 주최하는 성남시청소년재단을 포함하여 성남시에는 다양한 기관이 있습니다. 성남시의 기관 홈페이지에는 개별 기관이 공공 정책 사업을 진행하는 과정에서 작성한 다양한 자료집과 통계, 데이터 등을 제공하고 있으므로 여러분이 생각하는 정책과 관련된 성남시의 기관의 홈페이지를 방문해 보는 것이 도움이 될 수 있습니다.
 2. 검색은 새로운 정보를 얻을 수 있는 최고의 방법입니다. 언론사에서 발표한 다양한 기사들 중에는 통계 자료와 데이터, 인터뷰를 담고 있는 기사가 있습니다. 또한 통계 자료와 데이터를 직접 검색 과정에서 얻게 될 수도 있습니다.
 3. 유동인구 데이터로 제시된 지역을 지도 서비스를 통해 검색해보면서 해당 지역의 특성을 살펴본다면, 독창적인 정책 아이디어를 위한 통찰을 얻을 수 있습니다. 가까운 곳이라면 직접 해당 지역을 방문해보는 것도 좋은 관찰이 될 수 있습니다.
- 여러분의 좋은 정책 아이디어가 성남시를 더 나은 도시로 만듭니다!



제10회 빅데이터기반
성남시 청소년 정책제안대회

 아이디어
맛.ZIP

여러분의 특별한 정책 아이디어를 기다립니다!