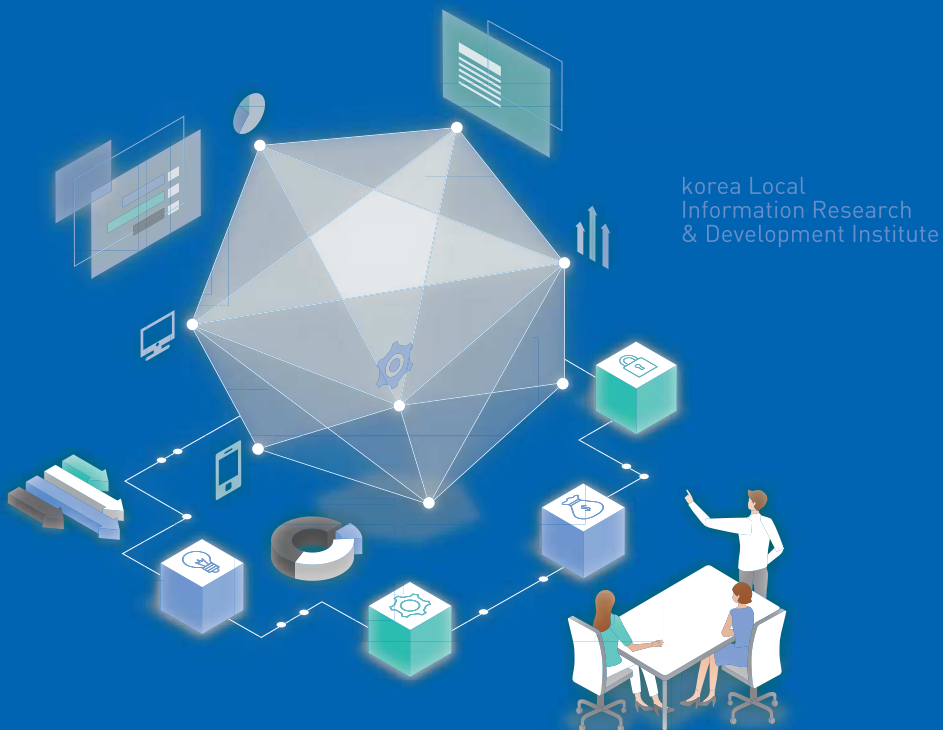


지방자치단체 빅데이터 분석 사례집



연구진

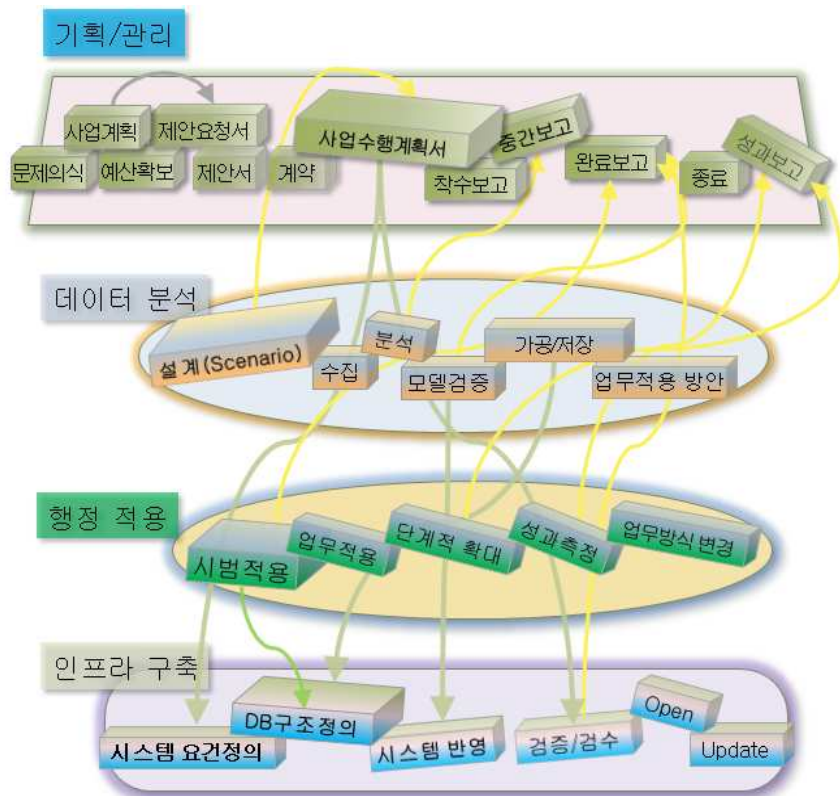
연구총괄	임완식	디지털기술부장
연구책임	김기홍	디지털기술부 선임
	강민선	디지털기술부 선임
공동집필	강두호(부산광역시),	김동한(국토연구원),
	장동익(경상남도),	노희섭(제주특별자치도),
	조기행(남양주시),	권목진(경기 광주시)

- 본 사례집의 내용 중 빅데이터 분석 추진 공무원의 의견 등은 개발원의 공식적인 입장이 아님을 밝힙니다.

사례집 활용 방안

- 본 사례집은 지방자치단체 빅데이터 분석 사업 활성화를 위해 우수 사례를 전반적인 프로세스와 주요 착안 사항을 중심으로 정리하여 타 기관에서 사업 추진 시 활용 가능하도록 구성하였습니다.

- (담당부서) 한국지역정보개발원 디지털기술부
- (협업기관) 부산광역시청, 경상남도청, 제주특별자치도청, 경기도 남양주시청, 광주시청
- (전체프로세스)



- 상기 프로세스 중 사례별로 반영이 가능한 부분만 포함하였고, 민감 정보는 삭제하였습니다.

2017년 12월 한국지역정보개발원 디지털기술부



안녕하십니까. 한국지역정보개발원 원장 손연기입니다.

우리원은 1997년 12월5일 지역정보화지원재단 설립에서 출발, 2003년 2월부터 5년간 자치정보화조합을 거쳐, 2008년 3월 이후 전자정부법에 기초한 법인으로 전환되어 올해로 창립 20주년을 맞이하였습니다.

우리원은 그 동안 지방자치단체와 행정안전부, 관련 유관기관과 함께 정보화를 통한 지방자치단체 행정효율 향상과 주민을 대상으로 하는 서비스 추진을 위해 끊임없이 노력해 왔습니다. 이러한 노력으로 ICT기반의 지방자치단체 행정서비스 개선, 지역주민 삶의 질 향상, 정보격차 해소 기여 등 명실상부한 국내 유일의 지역정보화 전문기관으로 우뚝 섰습니다.

특히, 지방자치분권 강화를 강조한 2017년 올해는 우리원에 큰 의미가 있는 한해였습니다. 지방분권을 실현하기 위한 기반으로 자치분권 로드맵이 마련되었다는 것은 중앙정부로부터 지방자치권의 실질적인 보장의 첫걸음이 시작되었음을 의미합니다.

이에 우리원에서는 다양한 정보화 관련 정책 및 ICBAM(사물인터넷, 클라우드 컴퓨팅, 빅데이터, 인공지능, 모바일)으로 대변되는 최신 ICT기술의 활용방안 등을 연구하여 최종 고객인 지방자치단체를 위한 자치정보화와 지식정보화 추진에 기여하고, 행정안전부와 함께 4차 산업혁명 시대의 지역정보화 혁신을 위해 헌신과 노력을 지속해 나가겠습니다.

아무쪼록 본 사례집이 지방자치단체 빅데이터 분석 업무의 커다란 발전을 위한 작은 밑거름이 되길 기원하며, 마지막으로 사례집이 발간되기까지 많은 도움을 주신 산·학·연 전문가 및 관계 공무원 여러분의 노고에 진심으로 감사드립니다.

2017년 12월

한국지역정보개발원 원장 **손 연 기**



지방자치단체 빅데이터 분석 사례집

Contents

특 집

제주특별자치도
공공빅데이터
활성화 정책

1. 빅데이터의 활용 절차와 지속가능성을 위한 요소들	3
2. 제주특별자치도의 빅데이터 활용 전략	5
3. 적극적으로 가치있는 데이터를 생산하는 지자체	6
4. 지역현안과 관련있는 민간의 데이터를 적극적으로 발굴 추진	16
5. 민간/공공데이터의 융합을 통한 지역 데이터 유통 체계 구축	21
6. 데이터의 활용을 적극적으로 유도하기 위한 Data Literacy 교육 추진	26
7. 향후과제	29

CHAPTER

I

부산광역시
도시양극화
분석 사례

가. 추진배경	34
나. 프로젝트 개요	34
다. 적용 프로세스	39
라. 기획·관리 단계	40
마. 데이터 분석 단계	53
바. 행정적용	64
사. 인프라 구축 단계	66

CHAPTER

II

경상남도 응급환자
골든타임 확보를 위한
빅데이터 분석 사례

가. 추진배경	78
나. 프로젝트 개요	79
다. 적용 프로세스	82
라. 기획·관리 단계	84
마. 데이터 분석 단계	89
바. 행정적용	99
사. 인프라 구축 단계	102



지방자치단체 빅데이터 분석 사례집

Contents

CHAPTER

III

남양주 보건 의료지표의
행정활용 사례

가. 추진배경.....	108
나. 프로젝트 개요.....	108
다. 적용 프로세스.....	113
라. 기획·관리 단계.....	114
마. 데이터 분석 단계.....	122
바. 행정적용.....	135
사. 인프라 구축 단계.....	137

CHAPTER

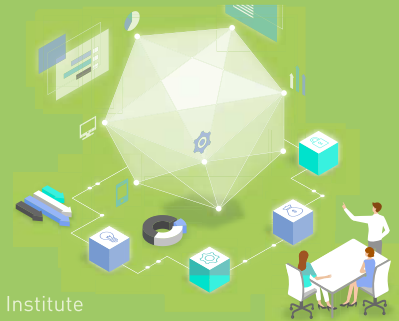
IV

광주시 당직민원
빅데이터 분석 사례

가. 추진배경.....	144
나. 프로젝트 개요.....	144
다. 적용 프로세스.....	147
라. 기획·관리 단계.....	148
마. 데이터 분석 단계.....	155
바. 행정적용.....	165
사. 인프라 구축 단계.....	172

특집

제주특별자치도 공공빅데이터 활성화 정책



korea Local Information Research & Development Institute

1. 빅데이터의 활용 절차와 지속가능성을 위한 요소들	3
2. 제주특별자치도의 빅데이터 활용 전략	5
3. 적극적으로 가치있는 데이터를 생산하는 지자체	6
4. 지역현안과 관련있는 민간의 데이터를 적극적으로 발굴 추진	16
5. 민간/공공데이터의 융합을 통한 지역 데이터 유통 체계 구축	21
6. 데이터의 활용을 적극적으로 유도하기 위한 Data Literacy 교육 추진	26
7. 향후과제	29

특집. 제주특별자치도 공공빅데이터 활성화 정책

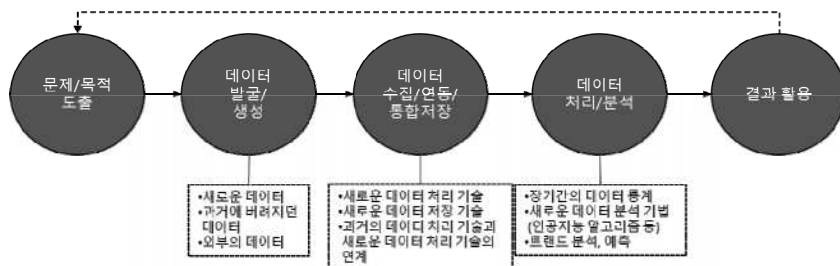
1. 빅데이터의 활용 절차와 지속가능성을 위한 요소들

빅데이터는 4차 산업혁명 시대의 원유(原油)로 일컬어진다. 4차 산업혁명을 주도하는 모든 기술의 기반에 빅데이터가 있기 때문이다. 공공영역에서 이루어지는 빅데이터에 대한 접근은 주로 '과학적 행정'을 지원하기 위한 형태이다. 또한, 지방자치단체의 경우 지역 내의 다양한 산업적 요구사항과 산업의 성장을 위해 빅데이터를 활용하기도 한다.

빅데이터를 지방자치단체 레벨에서 활용하려는 시도는 지난 수년간 국내에서 지속적으로 이루어졌다. 데이터 기반의 과학적 정책 수립과 행정이 해결해야하는 문제들에 대한 솔루션으로써 빅데이터는 큰 관심을 받았던 것이 사실이다.

그러나 지방자치단체가 추진한 대부분의 빅데이터 사업들은 단순한 빅데이터 플랫폼의 도입 또는 단발성의 분석과제 시도로 끝나거나, 과거의 단순 통계 또는 데이터 마이닝 형식의 프로젝트로 진행되고, 빅데이터에 대한 전문 인력의 조직 내 확보가 지지부진하다보니, 공공 영역에서의 빅데이터 실효성과 지속가능성에 대해서는 명쾌한 방향성을 제시하지 못한 것도 사실이다.

단발성의 빅데이터 프로젝트가 반복되는 이유는 지속가능한 빅데이터를 위한 요소들이 확보되지 않기 때문이다. 일반적인 빅데이터 활용 프로세스는 문제 발굴, 데이터 확보, 데이터 수집, 데이터 분석, 결과 활용의 절차로 이루어져있다.



이러한 절차의 본질은 데이터 자체의 가치에 집중하는 것에 있다. 빅데이터의 본질은 새로운 데이터에 대한 관점과 새로운 데이터 처리 기술을 이용하여 원하는 목적을 달성하는 일련의 데이터 활용 방식이라 할 수 있다. 절차 상 발생하는 다음과 같은 방식들은 모두 빅데이터 활용을 의미한다고 봐도 무방할 것이다.

- 과거에 사용하지 않던 데이터를 활용
- 목적에 맞는 새로운 데이터를 발굴/생성하여 활용
- 서로 다른 데이터를 연동하여 분석/활용
- 과거보다 장기간의 데이터를 활용
- 과거보다 저렴하지만 파워풀한 새로운 데이터 처리 기술을 활용

이러한 절차는 단일한 문제 해결의 과정에 빅데이터를 적용한 것으로 볼 수 있다. 전사적 관점이나 기관을 통틀어 빅데이터 활용을 극대화하기 위해서는 보다 상위의 요소들이 필요하다. 결과적으로 지속가능한 빅데이터를 위해서는 크게 4가지의 요소가 균형을 맞춰 갖춰져야 한다.

(1) 데이터 기반에 대한 기술(Data Technologies)

- 데이터를 수집/저장/분석하기 위한 기반에 대한 이해
- 데이터 솔루션/ 처리 아키텍처에 대한 이해

(2) 데이터 분석에 대한 기술(Data Analytics)

- 데이터를 활용해 분석할 수 있는 기법에 대한 이해
- 데이터와 분석기법에 대한 스킬

(3) 도메인(영역)에 대한 전문지식(Domain Knowledge)

- 데이터 분석을 통해 적용할 분야에 대한 전문지식
- 영역별 전문가 또는 자문단

(4) 종합적인 데이터 전략(Data Strategy)

- 데이터 기반, 분석, 적용분야에 대한 종합적인 전략
- 데이터 기술, 분석, 영역에 대한 하나의 통일된 적용 체계

2. 제주특별자치도의 빅데이터 활용 전략

제주특별자치도는 최근 3~4년간 급격한 변화를 겪고 있다. 인구는 매년 1만 명 이상씩 증가하여 2017년, 인구수는 67만 명에 달한다. 매년 급증한 관광객 수는 연간 1,500만 명에 달한다. 인구와 관광객의 증가에 따라 차량 등록 대수도 폭증하였으며, 교통체증, 부동산 가격 폭등에 따른 정주 환경의 악화, 환경 오염 등 다양한 사회적 문제에 직면해 있다. 제주특별자치도가 경험하지 못했던 문제들에 직면하면서, 시민들의 삶의 질은 급속히 악화되고 있다. 또한, 관광 영역을 중심으로 데이터와 모바일을 융합한 새로운 형태의 O2O기반 스타트업 기업들이 등장하면서 제주의 현황과 데이터들에 대한 요구사항들이 폭증하고 있다.

이러한 사회적 문제와 산업계의 성장지원은 단발적인 시도로 그쳐서는 충분한 성과를 거두는 것이 불가능하다. 때문에 단계적인 접근과 지속가능한 형태의 기반 구조들을 필요로 하며, 연속된 시도를 이끌어 내기 위해서는 중장기적인 로드맵과 정책들을 필요로 한다.

제주의 인구변화 추이



지속가능한 빅데이터 활용 체계 구축을 위해 제주특별자치도는 크게 4가지의 정책을 설계하여 추진하고 있다.

(1) 적극적으로 가치 있는 데이터를 생산하는 지자체 구축

- 공공 Wi-Fi 기반의 관광객 이동 경로 데이터 생산
- 공공 Wi-Fi 기반의 관광객 체류 시간 데이터 생산
- 표준 데이터 수집/저장/분석 체계 및 플랫폼 구축
- Silo 체계 파괴를 위한 행정 데이터의 Data Warehouse 구축

(2) 지역현안과 관련 있는 민간의 데이터를 적극적으로 발굴 추진

- 신용카드 데이터 분석을 통한 소비 패턴 분석
- 검색엔진, 내비게이션 로그 분석을 통한 개인 이동 요구사항 분석
- 지역 기업들의 데이터 발굴을 통한 공공데이터 전환
- 기업 MOU를 통한 데이터 확보 추진

(3) 민간의 데이터와 공공 데이터의 적극적인 융합을 통한 데이터 유통 체계 구축

- 지역 데이터 기반의 데이터 허브 구축
- 관광 특화형 데이터, 인프라 활용 플랫폼 구축

(4) 시민들의 데이터 활용을 적극적으로 유도하기 위한 Data Literacy 교육 추진

- 코딩 교육을 통한 사회 현안 해결 방식 교육
- 초/중/고 학생들과 사회 문제 해결을 위한 해커톤 추진

3. 적극적으로 가치 있는 데이터를 생산하는 지자체

제주특별자치도는 지자체 레벨에서 적극적으로 가치 있는 데이터를 생산하고, 이를 공공 데이터로 개방함으로써 정책 설계에 대한 데이터 생산 및 산업계의 비즈니스 지원을 위한 정책을 추진하고 있다.

이는 빅데이터 분석/활용에서 가장 큰 걸림돌이 되는 '쓸 만한 데이터가 없다'라는 문제를 해결하기 위한 것으로, 민간에서 생산하기에 과도한 비용이 들어가거나 행정 내/외부적으로 데이터 확보 자체가 불가능하지만 민관 요구사항이 높은 데이터를 지자체에서 직접 생산하여 공급함으로써 가치 있고 유용한 데이터의 부족 문제를 해결

하는 데 목적이 있다.

2015년에 이루어진 제주도청 및 민간기업과 데이터 요구사항에 대한 설문조사를 진행하였는데, 가장 많은 요구사항이 발생한 데이터 영역은 관광객의 이동경로 데이터였다. 민간이 보유하고 있는 이동경로 데이터의 확보를 위한 조사 결과, 정확히 요구사항과 맞아 떨어지는 데이터가 없었기 때문에 제주도청은 직접 이 데이터를 생산하기로 하였다. 데이터의 생산과 활용을 위해서 데이터 생산 구조와, 수집 및 저장 구조, 유통 구조에 대한 종합적인 설계가 진행되었고, 3개년에 걸친 구축 작업을 추진하기로 결정하였다.

(1) 1차년도 : 공공 Wi-Fi access log 수집/분석 모델에 대한 초기 모델 및 실험

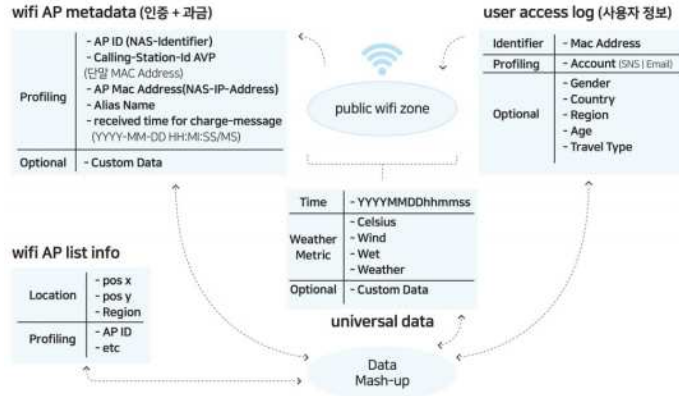
(2) 2차년도 : 공공 Wi-Fi 설치 착수 및 데이터 분석 모델 자동화

(3) 3차년도 : 공공 Wi-Fi 설치 확대 및 데이터 분석 모델 고도화(실시간 처리)

공공 Wi-Fi를 통해 접속한 관광객의 위치 정보를 확보하여, 관광객이 어떤 순서로 관광지를 방문하는지, 어느 관광지에 얼마만큼 머무르는지에 대한 데이터 생산을 위해 공공 Wi-Fi 설치 계획이 수립되었고 Wi-Fi access log를 이용한 분석 모델들이 설계되었다. Wi-Fi AP의 위치정보가 정리되었고, Wi-Fi AP에 접속하는 사용자로부터 기본적인 프로파일 정보(성별, 나이, 국가, 여행목적)를 수집하고 이를 기 보유하고 있는 기상청 데이터 등과 매쉬업(mash-up)하는 모델들이 설계되었다.

Wi-Fi access log를 이용한 이동경로 및 체류시간 추출

WiFi로 부터 생산되는 데이터



공공 Wi-Fi는 주요 관광지와 유동인구가 많은 지점, 대중교통(버스)등에 설치되었으며 순차적으로 확대해 나가고 있다.

제주지역 데이터 수집을 위한 공공 Wi-Fi 설치 확대

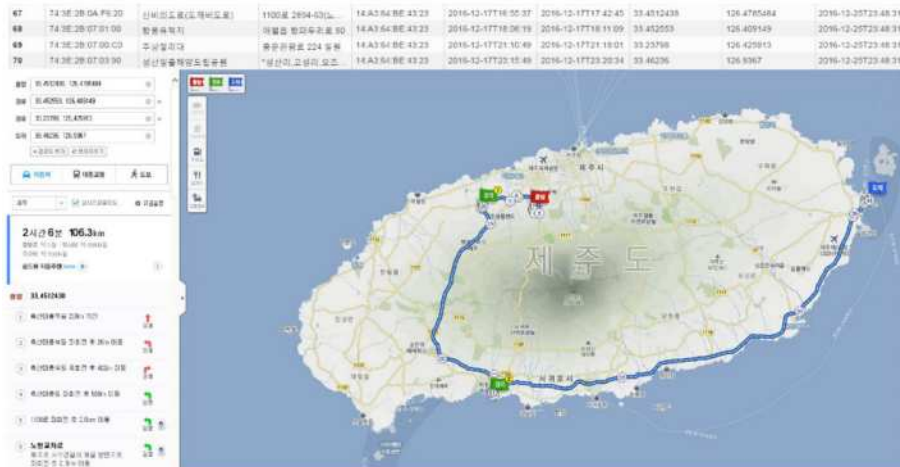


사용자의 Wi-Fi 접속정보는 자동으로 분석되어 이동경로 및 체류시간에 대한 분석을 Daily Batch로 분석하며, 향후 실시간 분석 모델로 개선될 예정이다.

아래는 초기 실험 단계에서 실험자가 이동하며 접속한 Wi-Fi access log를 지도위에 표출하여 실제 이동 경로와 수집된 데이터의 분석 결과를 비교하기 위한 화면이다. 접속을 시도한 Wi-Fi AP의 위치정보를 기반으로 카카오 지도에 위치정보 기반의 경로를 표현한 것이다.

Wi-Fi access log에 의한 이동경로 추출 실험 모델

저장/분석 플랫폼 수집 Data를 이용한 이동 경로에 대한 데이터 모델링 수행 결과



실험 결과 실제 이동 경로와 분석 결과가 거의 일치하였으며, 누적 수집된 데이터를 이용해 자동화된 분석모델을 구축하여 의미 있는 수준으로 중복되는 이동 경로를 추출한 결과로, 중국인과 내국인 관광객을 대상으로 구분하여 분석하였다. 이러한 데이터 분석결과는 관광 관련 부서와 유관부서에 제공되어 현황 분석 및 정책 수립의 기반 자료로 활용된다.

다음은 Wi-Fi access log를 기반으로 분석한 결과의 예시들이다. 중국인 관광객과 내국인 관광객들의 행동 패턴이 매우 상이함을 알 수 있다. 이러한 데이터를 요구하는 관광업계가 많기 때문에 해당 데이터는 공공 데이터로 개발되어 민간 비즈니스를 지원하게 될 것이다.

내국인 관광객과 중국인 숙박지역 분석 : 심야시간 Wi-Fi access log 분석

데이터 분석 결과



Wi-Fi access log 분석 기반 중국인 관광객의 주요 이동경로 추출

데이터 분석 결과



Wi-Fi access log 기반 내국인 관광객 이동경로 추출

데이터 분석 결과



이러한 데이터의 수집과 분석, 활용을 위해서 표준화된 빅데이터 플랫폼의 구축이 필요했다. 단순히 Wi-Fi access log 처리 뿐 아니라 제주의 다양한 사회적 문제 해결을 위한 데이터들의 수집, 분석, 활용 및 수집 데이터를 공공 데이터로의 공개 추진하는 부분과 데이터에 대한 이해도가 부족한 일반 부서에서도 현황을 파악할 수 있도록 시각화 구조에 대한 요구사항도 함께 해결해야 했다.

직접적으로 생산되는 데이터와 유관 기관/ 기업이 보유하고 있는 데이터들의 효율적인 수집을 위해 데이터 수집에 대한 표준화된 솔루션을 도입하였고, 빅데이터의 저장/분석을 위한 플랫폼이 구축되었으며, 분석결과의 효율적인 활용을 위해 DW(Data Warehouse) 구조와 시각화 플랫폼이 도입되었다.

효율적으로 지역기업이 유지보수 및 기반 기술에 대한 기술이전의 용이성을 위해 대부분의 솔루션은 오픈소스를 이용해 구축되었다.

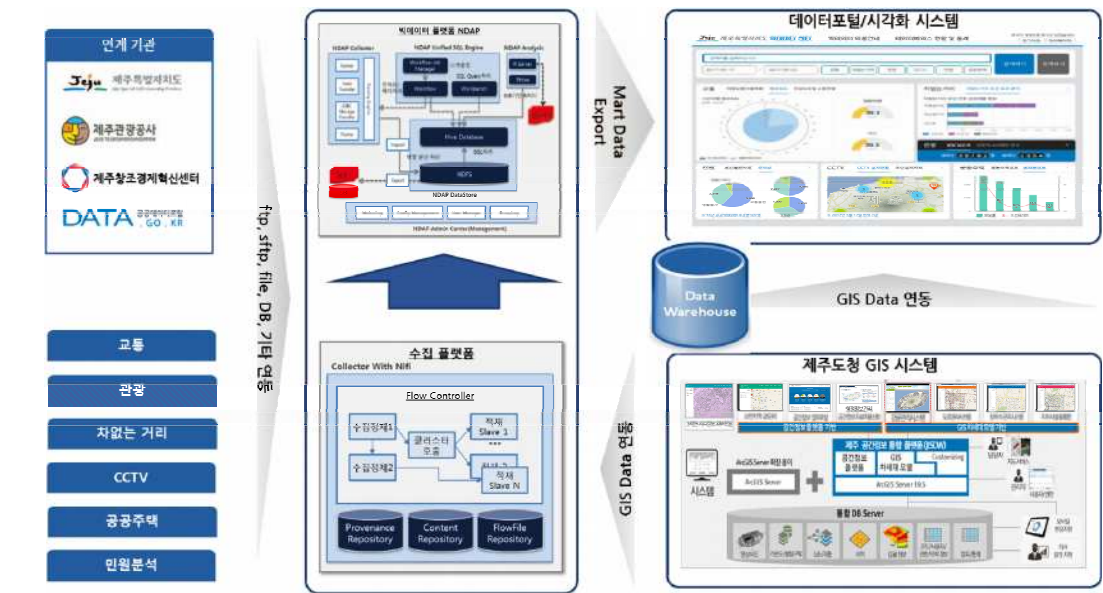
수집 플랫폼은 Apache Nifi를 활용하여, 프로토콜 및 전송방식에 대한 빠른 대응

및 기본적인 데이터 수집처리(ETL)을 GUI기반으로 처리할 수 있도록 하였으며, 저장/분석 구조는 Hadoop 2.x 기반의 분산 파일 시스템과 Hive, Spark, R을 이용해 분석 목적에 맞는 인터페이스를 활용할 수 있도록 구성하였다. 빅데이터의 분석결과를 빠르게 조회하고 활용할 수 있도록 Oracle과 같은 RDBMS로 분석결과가 자동 적재되도록 하였으며, RDBMS의 데이터와 공간정보 시스템과의 연계를 활용해 각 업무 분야에 맞는 시각화 모델들을 구성하였다.

이는 3개년에 걸친 로드맵으로 추진되었다.

- (1) 1차년도(2016): 기본 플랫폼 구축 - 오픈소스 기반의 수집/저장/분석 플랫폼 구축
- (2) 2차년도(2017): 플랫폼 고도화 - 대상 데이터 선별 및 데이터 포털/시각화 시스템 구축
- (3) 3차년도(2018): 플랫폼 활성화 - 공간정보 시스템 연계 및 업무부서의 플랫폼 활용 활성화

제주특별자치도의 표준 빅데이터 플랫폼 구조

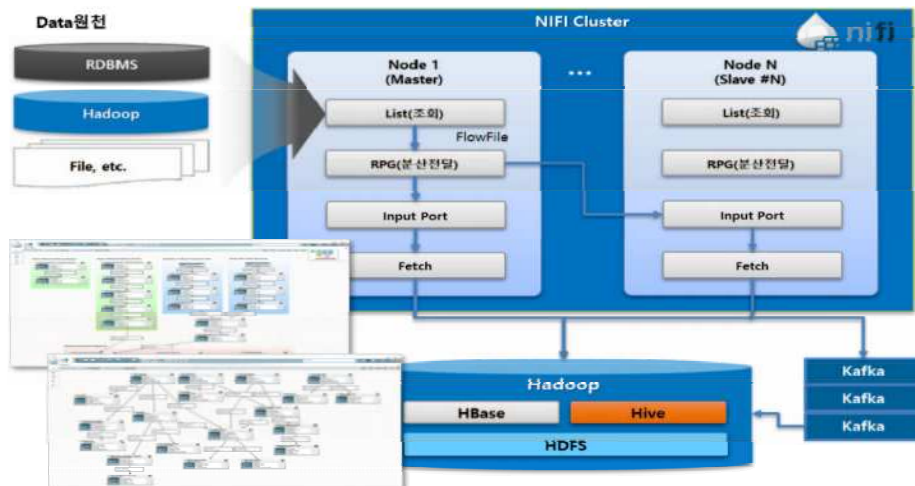


수집 플랫폼에 활용된 Apache Nifi는 오픈소스 기반으로 데이터 수집 출처와 프로토콜, 데이터 포맷에 대한 관리와 데이터 수집 시 발생하는 변환처리(ETL)에 대한 복잡한 job을 GUI기반으로 처리할 수 있도록 해준다. 이를 통해 교통, 관광, CCTV등 다

양한 지자체의 데이터와 유관기관의 데이터들을 빠르고 쉽게 수집관리 할 수 있게 되었다. 또한 향후 발생할 수 있는 새로운 데이터의 수집 조건에 대한 신속한 대응이 가능하게 되었다.

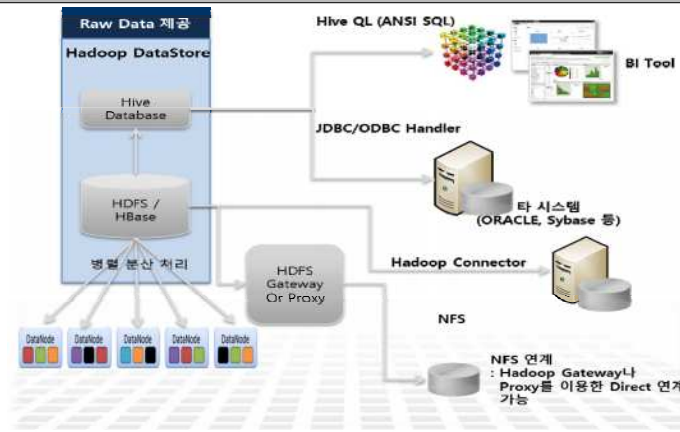
수집된 데이터는 메시지 큐를 통해 순차적으로 Hadoop 기반의 저장 플랫폼에 분산 저장되며, Hive 및 Spark, R등을 통해 구성된 분석 모델에 따라 분석을 수행하게 된다. 2017년 9월에 시행된 제주특별자치도의 대중교통 체계 개편에 따른 효과 분석 및 문제 분석에 버스카드 데이터, ITS(교통정보시스템)의 데이터들이 해당 플랫폼을 활용하여 수집/분석 처리되었다.

수집 플랫폼의 구조



Hadoop 분산 파일 시스템(HDFS: Hadoop Distributed File System)구조에서 분석 처리된 데이터들을 상시적으로 빠르게 활용하는 데에는 file I/O 퍼포먼스 등 여러 가지 오버헤드가 존재한다. 때문에 분석 결과를 I/O 퍼포먼스가 확보되어 있는 RDBMS로 옮기고 이를 다양한 업무 시스템과 연동하는 구조가 필요하다. 대량의 빅 데이터는 분산 파일 시스템 구조를 활용하지만, 상대적으로 볼륨이 적은 분석결과 데이터를 일반적인 데이터베이스로 옮겨 효율성을 극대화하는 조치이며 일반적으로 일컬어지는 Data Mart에 대한 구축 작업으로 볼 수 있다.

HDFS의 데이터 이동을 통한 데이터 활용 효율성 극대화

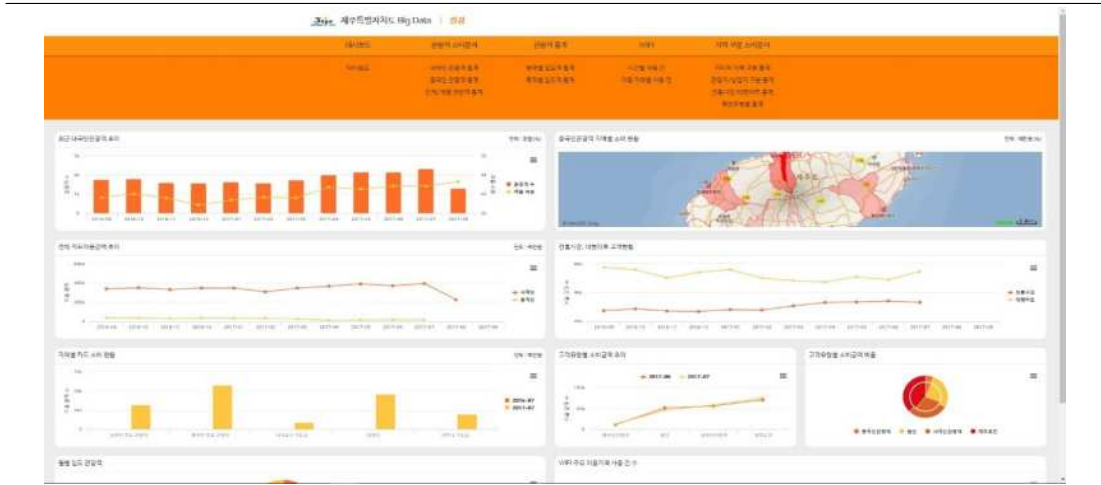


HDFS에 적재된 분석결과를 JDBC connector등을 통해 RDBMS로 옮겨 데이터 마트가 구축되면 이를 활용해 시각화에 대한 처리를 하고, 이를 각 업무부서나 일반 시민들에게 공개하게 된다. 다음은 수집된 각 영역의 데이터 분석결과를 데이터 마트로 변환하여 시각화 솔루션과 연계해 표출한 화면으로, 이 화면들은 담당 업무 부서에서 사용하게 된다.

아래는 분석결과 Data Mart를 시각화 솔루션과 연계하여 표출한 화면의 예시들이다.

각 영역의 데이터 분석 결과에 대한 시각화 연동화면 예시







가치 있는 데이터의 부족 현상 해소와 지자체 레벨의 데이터 생산 정책을 지속가능하게 추진하기 위해서는 생산된 데이터의 수집, 적재, 분석, 활용에 대한 전체적인 데이터 라이프 사이클에 대한 종합적인 고민이 필요하다.

제주특별자치도는 이를 위해, 데이터 생산을 위한 기반 인프라 구축부터 시작하여 지역 빅데이터의 표준 플랫폼 구조를 구축하여 '적극적으로 가치 있는 데이터를 생산하는 지자체'를 지향하고 있다.

4. 지역현안과 관련있는 민간의 데이터를 적극적으로 발굴 추진

공공이 가지고 있는 데이터의 경우, 물론 가치 있는 데이터도 다수 있지만 실제 활용하기에는 적합하지 않은 형태로 관리되고 있거나 품질의 문제, 개인정보 포함에 따른 제공/활용 상 어려움 등 다양한 문제들을 갖고 있기도 하다. 내부의 데이터를 활용할 수 없다면 외부(민간)의 데이터를 적극적으로 발굴하여 활용해야 한다. (서울특별시도 진행했던 빅데이터 기반의 심야좌석버스 노선 설계가 적합한 예일 것이다)

제주도 역시 공공기관의 데이터에 대한 인식 부족, 활용 의지의 부족 등으로 인해 고품질의 데이터가 매우 부족한 상황이었다. 내부 데이터에 대한 품질 개선 작업을 진행하는 동시에 사회 문제 해결 및 과학적 정책 설계를 위한 데이터들의 확보가 시급히 진행되어야 했다.

신용카드 기반 기간별, 지역별, 세그먼트별 패턴 분석



내국인, 중국인 관광객 세그먼트별 소비 패턴 분석

관광 유형					인당 일종별 소비금액 비중 Top 3 (H4, '16.3 - 8월)			카드 이용지 수 기준 소비 지역 Top 3 (Y2, '15.9~'16.8월)		
국적	상품 유형	제류 기간	방문 빈도	방문 지역	1위	2위	3위	1위	2위	3위
내국인	단체	단기	-	-	숙박	구매	식음료	제주시 도심	서귀포시 도심	성산읍
		장기	-	-	숙박	식음료	구매	제주시 도심	서귀포시 도심	성산읍
	개별	단기	-	-	교통	숙박	식음료	제주시 도심	서귀포시 도심	애월읍
		장기	-	-	식음료	숙박	교통	제주시 도심	서귀포시 도심	성산읍
중국인	단체	단기	최초방문	단독	숙박	구매	식음료	제주시 도심	서귀포시 도심	성산읍
				교차	숙박	구매	문화/레저	제주시 도심	서귀포시 도심	성산읍
			재방문	단독	숙박	구매	식음료	제주시 도심	서귀포시 도심	성산읍
				교차	숙박	구매	식음료	제주시 도심	서귀포시 도심	성산읍
		장기	최초방문	단독	숙박	구매	식음료	제주시 도심	서귀포시 도심	성산읍
				교차	숙박	구매	식음료	제주시 도심	서귀포시 도심	성산읍
			재방문	단독	숙박	구매	식음료	제주시 도심	서귀포시 도심	성산읍
				교차	숙박	구매	식음료	제주시 도심	서귀포시 도심	성산읍
	개별	단기	최초방문	단독	숙박	숙박	구매	제주시 도심	서귀포시 도심	애월읍
				교차	숙박	구매	숙박	제주시 도심	애월읍	서귀포시 도심
			재방문	단독	숙박	숙박	구매	제주시 도심	서귀포시 도심	애월읍
				교차	숙박	숙박	구매	제주시 도심	서귀포시 도심	애월읍
		장기	최초방문	단독	숙박	숙박	구매	제주시 도심	서귀포시 도심	안덕면
				교차	숙박	숙박	구매	제주시 도심	서귀포시 도심	안덕면
			재방문	단독	숙박	숙박	구매	제주시 도심	서귀포시 도심	안덕면
				교차	숙박	숙박	구매	제주시 도심	서귀포시 도심	안덕면

내국인, 중국인 관광객 세그먼트별 소비금액 변화 분석

관광 유형					일종별 소비금액 비중 변동 (H2, '15.3~8월 대비 H4, '16.3~8월, 기타 일종 제외)					
국적	상품 유형	제류 기간	방문 빈도	방문 지역	교통	문화/레저	숙박	숙박	식음료	유증
내국인	단체	단기	-	-	0.0%	-0.5%	-1.0%	-0.5%	0.0%	-0.1%
		장기	-	-	-0.6%	-0.5%	-0.9%	0.1%	0.0%	-0.2%
	개별	단기	-	-	-0.8%	-0.4%	-0.2%	0.1%	0.7%	-0.1%
		장기	-	-	-1.2%	-0.7%	-0.1%	0.1%	0.7%	0.0%
중국인	단체	단기	최초방문	단독	0.0%	-0.1%	-4.0%	8.8%	0.0%	0.0%
				교차	0.0%	-0.1%	-11.2%	18.1%	0.0%	0.0%
			재방문	단독	0.0%	-0.2%	-1.9%	2.8%	0.0%	-0.3%
				교차	0.0%	-0.2%	-7.7%	12.4%	0.0%	-0.1%
		장기	최초방문	단독	-0.5%	0.0%	-1.5%	5.2%	0.0%	-0.2%
				교차	0.0%	-0.2%	-8.1%	14.8%	0.0%	0.0%
			재방문	단독	-0.1%	-0.2%	1.0%	5.2%	0.0%	-0.5%
				교차	0.0%	-0.1%	-9.3%	14.0%	0.0%	-0.1%
	개별	단기	최초방문	단독	0.0%	0.1%	-0.2%	0.0%	-4.9%	0.6%
				교차	0.0%	-1.8%	-9.6%	11.0%	2.0%	0.0%
			재방문	단독	-0.1%	1.6%	3.6%	11.1%	-21.6%	0.9%
				교차	0.0%	0.4%	1.3%	20.1%	-22.7%	1.2%
		장기	최초방문	단독	0.1%	0.6%	2.8%	3.1%	-9.3%	0.6%
				교차	0.0%	0.5%	-4.1%	5.7%	-0.8%	1.3%
			재방문	단독	-0.2%	2.7%	1.0%	9.0%	-16.4%	0.4%
				교차	0.0%	-0.1%	-2.0%	7.2%	-3.8%	-0.3%

내국인 연령대별 소비금액 변화 분석

구분	14.9~15.8			최근 1년(15.9~16.8)			전년동기대비 증감		
	남성	여성	합계	남성	여성	합계	남성	여성	합계
연령대									
20대 미만	0.0%	0.0%	0.1%	0.1%	0.1%	0.2%	0.0%	0.0%	0.1%
20대	4.6%	6.8%	11.4%	5.7%	8.2%	13.8%	1.1%	1.4%	2.5%
30대	15.1%	13.4%	28.5%	15.2%	12.9%	28.1%	0.1%	-0.5%	-0.4%
40대	17.2%	9.9%	27.1%	17.1%	9.9%	27.0%	-0.1%	0.0%	-0.1%
50대	14.1%	7.3%	21.4%	13.5%	7.1%	20.6%	-0.7%	-0.2%	-0.9%
60대 이상	7.7%	3.9%	11.6%	6.8%	3.5%	10.4%	-0.9%	-0.3%	-1.2%
합계	58.7%	41.3%	100.0%	58.3%	41.7%	100.0%	-0.4%	0.4%	0.0%

주1) 국내 관광객의 제주지역 소비 기준

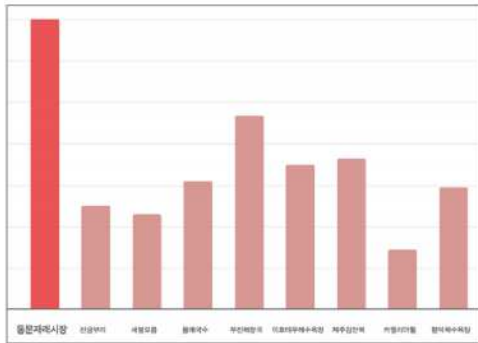
주2) 전년동기대비 증감 = ('15.9~'16.8) 비중 - ('14.9~'15.8)비중

이러한 데이터 분석 결과는 소비금액의 지속적인 증가가 예상되는 FIT(자유여행객) 대상의 정책 수립과 내국인 20대 관광객을 대상으로 한 마케팅 전략 등에 활용되었다.

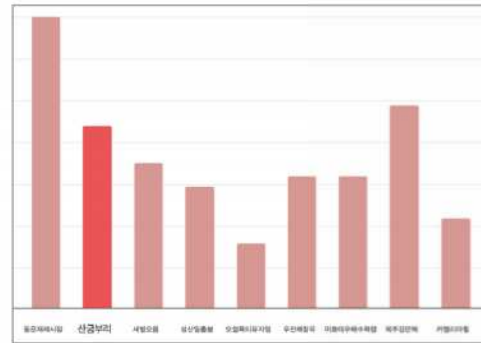
카드 데이터 이외에도 지역 현안에 대한 데이터들의 확보와 관련해서 지역기업들과의 협의가 지속적으로 진행되고 있다. 교통 문제 해결을 위해 카카오 등이 보유하고 있는 지역 내 모빌리티 데이터, 대기업들이 제공하고 있는 내비게이션 서비스와 렌터카에 장착된 내비게이션의 출발지·도착지 데이터 등에 대한 제공협력이 추진되고 있으며, 지도 서비스 등의 검색 로그 등 사용자들이 방문 또는 이동 실질적/잠재적 욕구를 포함하고 있는 데이터들의 확보들이 추진될 예정이다.

버스카드데이터와 카카오지도 검색 데이터 분석 결과 : 제주공항에서 출발 후 도착지

대중교통 이용 Top 9 방문지



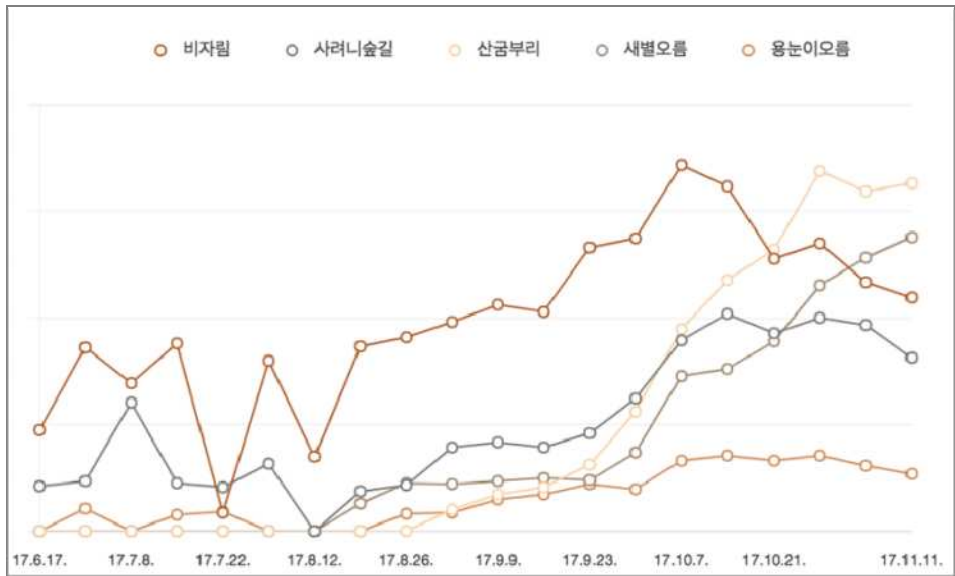
자동차 이용 Top 9 방문지



주요 POI별 카카오 검색 데이터 분석 결과 : 특정 지점의 연령/성별 시간별 인기도



주요 POI들의 카카오 검색 데이터 추이 분석 결과 : 기간별 오름 방문 수요



행정이 보유하고 있는 데이터나, 공공 데이터의 경우 대부분 실제로 발생한 이벤트에 대한 결과적 데이터인 경우가 많다. 이러한 데이터들이 사용자들의 실질적/잠재적 욕구를 담고 있는 데이터와 결합된다면 더욱 효율적인 정책 수립이 가능해 질 것이며, 민간 비즈니스에서 더욱 가치있게 활용되는 공공 데이터가 될 가능성이 크다.

제주특별자치도는 이러한 “민간과의 데이터 협력”을 통해 민간 비즈니스와 과학적 정책 수립 기반을 확보하기 위해 가치 있는 데이터의 확보에 집중하고 있다. 제주 지역 내에서는 이러한 노력으로 인해 다양한 기업 또한 지역사회에 대한 기여 측면에서 데이터 레벨의 협력을 긍정적으로 생각하는 추세가 증가하고 있다.

5 민간/공공데이터의 융합을 통한 지역 데이터 유통 체계 구축

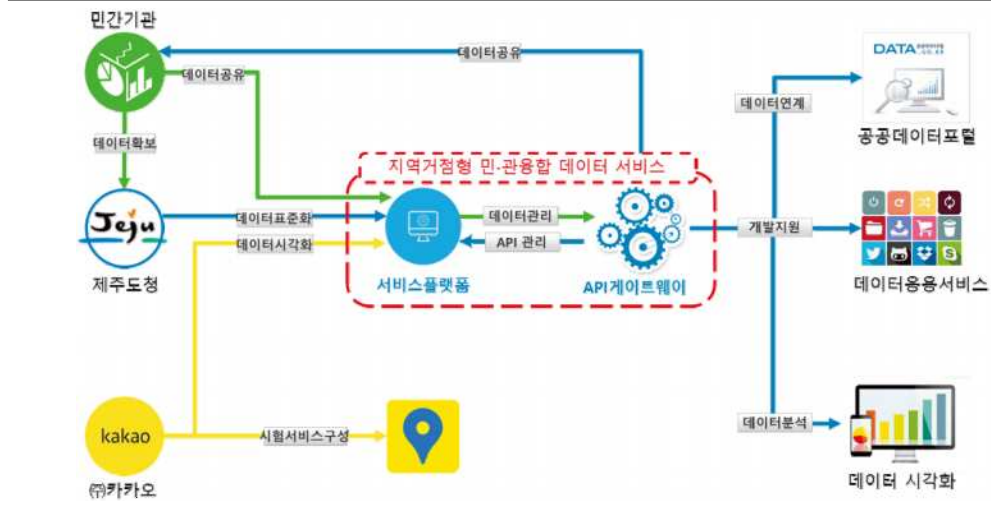
데이터의 확보, 수집, 분석, 활용 측면과 더불어 중요한 것은 데이터들에 대한 유통체계를 활성화함으로써 행정과 민간의 데이터 활용성을 극대화 하는 것이다. 이를 위해서는 Open API 등 데이터에 대한 접근 용이성을 확보하고, 데이터의 업로드 및 다운로드를 통한 손쉬운 데이터 교환 체계의 구축, 데이터에 대한 관리 구조 등이 필요하다.

지자체 입장에서 데이터에 대한 고민을 할 때는 지역 현안과 지역과 관련된 데이터(local intensive data)에 집중할 수밖에 없고, 지속가능한 데이터의 활용을 위해서는 이러한 데이터들의 효율적인 유통 구조에 대해 고민해야 한다. 또한, 공공과 민간이 각각 보유하고 있는 데이터는 각자의 영역에서 활용되고 있으며, 이는 개별적 영역에서의 목적에 따른 가치를 가질 뿐이다. 이와 같은 한계를 해결하기 위해서는 민간과 공공의 데이터가 각자의 영역을 벗어나 서로 융합되고 새로운 목적에 부합하는 가치를 가질 수 있도록 새로운 데이터로 구성될 필요가 있다.

제주특별자치도는 제주테크노파크, 카카오와 협력하여 이와 같은 민간과 공공의 데이터가 각각의 영역에만 존재하는 문제를 해결하기 위해, 2017년 과학기술정통부가 진행한 빅데이터 플래그십 시범사업으로 '지역 거점형 민·관 융합 데이터 서비스 표준모델 구축'을 제안하여 추진하였다.

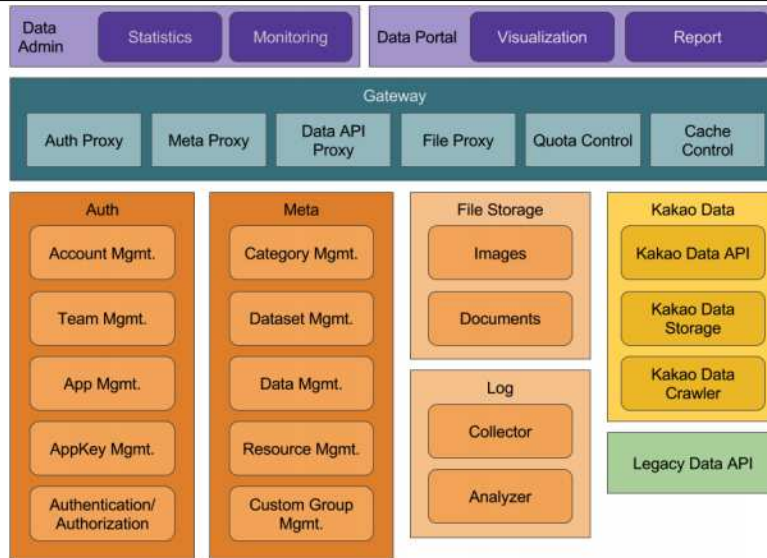
제주특별자치도가 직접 생산하고 있는 Wi-Fi 기반의 데이터들과 각종 행정 데이터, 공공기관이 보유하고 있는 다양한 데이터와 민간이 보유하고 있는 데이터들을 한 데 모아 유통할 수 있는 일종의 데이터 허브를 구축하고, 중장기적으로 이 데이터 허브를 지역 중심의 데이터 유통·활용의 거점으로 활용하는 것이 목적이었다.

지역거점형 민·관 융합 데이터 서비스 모델의 구조



자발적인 데이터 유통 구조의 확립을 위해 데이터의 주제별 그룹, 카테고리, 데이터 생산/관리 그룹의 관리 등 기본적인 데이터 관리 체계와 openAPI를 이용한 데이터 액세스 구조를 구축하였다.

데이터 허브의 기능 구조



데이터 레벨의 유통 뿐 아니라, 데이터 분석 결과를 자유롭게 공유할 수 있는 구조를 만들어 원시 데이터(raw data)뿐 아니라 민간이나 공공이 데이터 허브 내의 데이터를 이용해 분석한 결과 또한 유통할 수 있는 구조 또한 만들었다. 제주 데이터 허브는 민간과 공공은 데이터 허브를 통해 지역 내의 데이터에 접근, 활용하고 분석, 활용 결과를 다시 공개하게 함으로써 데이터의 생산 ⇒ 분석 ⇒ 활용 ⇒ 활용결과 재사용의 데이터 생태계 구축을 지향한다.

제주 데이터 허브 내의 데이터 분석결과 예시

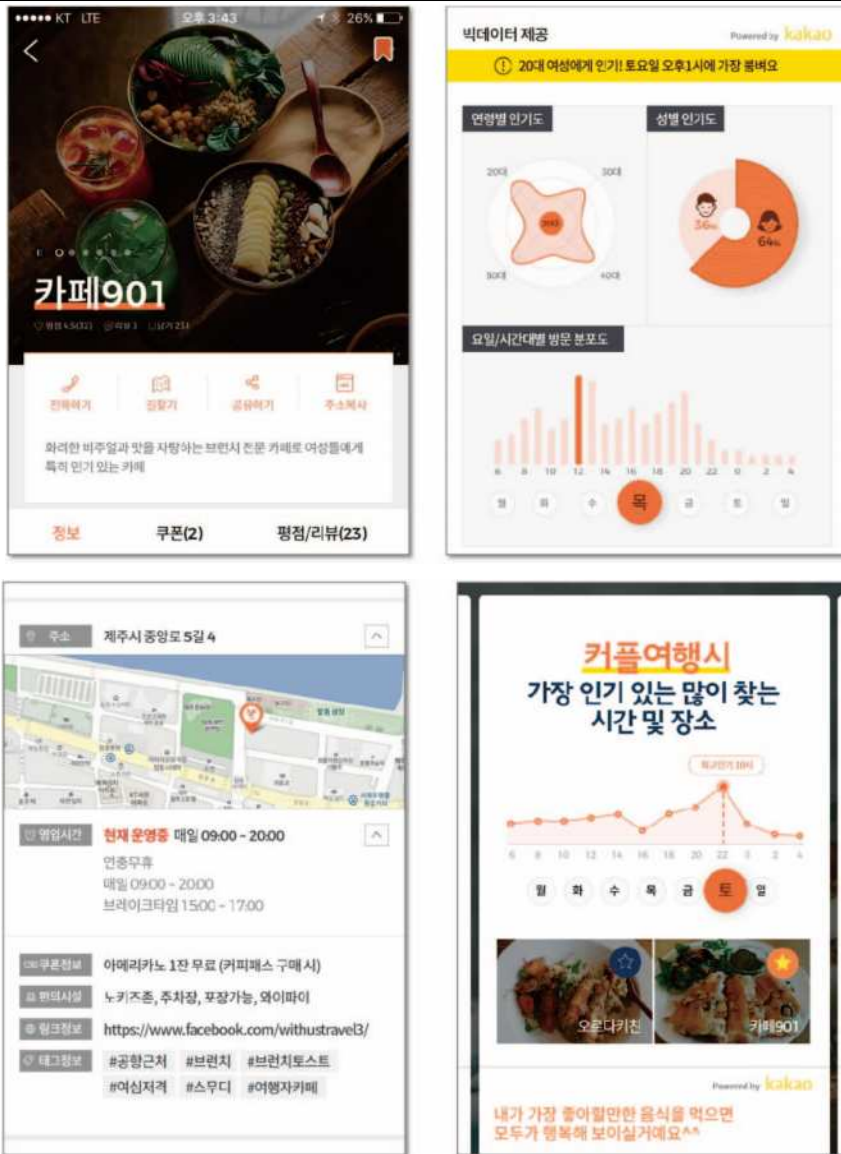


지역 거점의 데이터 허브는 지역 내의 데이터를 한 곳으로 통합함으로써 민관의 데이터 접근성을 향상시키고, 데이터 활용 사례들을 공유함으로써 지역 내의 데이터 사용성을 향상시키는 효과가 있다. 다만, 주 사용자가 될 지역 민간 기업들의 경우, 공공 데이터에 대한 요구사항이 강하나 구체적이지 못한 경우들이 많고, 요구사항과 부합되는 데이터가 제공되더라도 비즈니스에 제대로 활용 못하는 경우들이 다수 발생한다. 따라서 데이터 유통 및 활용을 강화하려면 데이터 허브의 구축 뿐 아니라 데이터 구축/활용에 대한 지역 민간기업 대상의 컨설팅이 동시에 이루어지는 것이 효과적이다.

제주특별자치도에서는 제주도청과 카카오가 민간 기업의 데이터 요구사항의 구체화와 데이터 활용에 대한 컨설팅을 공동으로 진행하여 지역 민간기업의 데이터 활용 사례(use case)들을 만드는 것을 빅데이터 플래그십 사업 진행과 동시에 추진하였다.

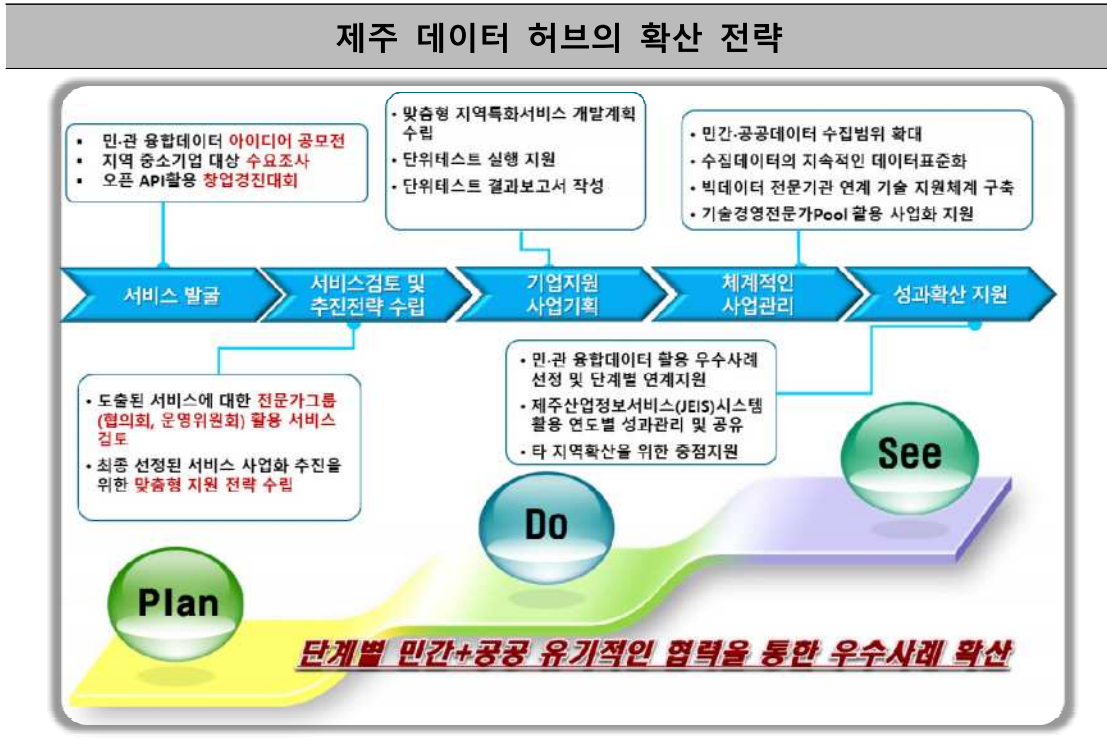
제주 데이터 허브를 통해 제공된 공개 데이터의 활용 방식을 지역 민간 기업에 컨설팅 함으로써, 주요 관광지 및 맛집에 대한 관광객의 선호도 분석 결과를 제공하고 이를 통한 관광지/맛집 추천 서비스에 제주 데이터 허브의 데이터가 활용되었다.

제주 데이터 허브의 데이터를 활용한 모바일 관광 정보 추천 서비스



제주 데이터 허브의 현재 모습은 지역 내의 데이터들을 통합 수집하고, 요구사항에 맞는 데이터들을 빠르게 검색/ 액세스 할 수 있는 구조이다. 향후에는 각자 존재하는 민간 데이터와 공공 데이터의 융합을 통해 새로운 가치의 데이터를 생산한 계획을 가지고 있다. 우선적으로 2018년에는 공공이 가지고 있는 대중교통과 교통량 데이터

와 함께 민간이 가지고 있는 내비게이션 데이터들을 융합하여 제주 지역 내의 모빌리티 데이터(mobility data)를 구축하는 것을 목표로 삼고 있다.



지자체, 지역 내 기관, 지역 내 민간기업의 데이터 협업을 통한 데이터 유통 구조 구축 사례로 제주도는 2017년 열린 혁신 우수기관으로 선정되었다.

6 데이터 활용을 적극적으로 유도하기 위한 Data Literacy 교육 추진

최근 정부는 시민 주도로 디지털 기술을 활용한 지역현안 발굴·해결을 위한 디지털 사회혁신을 추진하고 있다. 후쿠시마 원전 사고 이후 주민들이 스스로 방사능 측정 장치를 개발 보급한다거나, 개천의 수위와 안전상태의 실시간 확인 앱을 주민들이 직접 개발하여 보급하는 등, 시민 주도로 코드와 데이터를 활용한 문제 해결 시도들은 전 세계적으로 확산되는 추세이다.

시빅 해킹(Civic Hacking)으로 표현될 수 있는 이러한 디지털 기반의 사회혁신 운동은 각 지역의 커뮤니티를 중심으로 진행되고 있으며, 문제에 대한 발굴 및 데이터 기반의 분석, 활용을 위한 프로그램의 개발 등이 시민 커뮤니티 중심으로 이루어진다는 것이 특징이다.

제주도에서는 시민들의 코드와 데이터에 대한 기본 소양을 향상시키고 미래 세대들과 함께 사회적 문제를 발굴하고 해결하는 시도를 추진하고 있다. 전국 최초로 지자체 레벨에서 추진된 코딩교육인 <제주로 On 코딩>은 2016년 시범 사업을 시작으로 2017년에는 미래 세대들의 지역 현안에 대한 관심을 유도하고 코드와 데이터를 이용한 문제 해결에 대한 과정들을 교육하였다.

제주로 On 코딩 교육 체계 설계



<제주로 On 코딩>은 크게 4개의 영역으로 추진된다.

- (1) 찾아가는 교육: 초중고교를 방문한 기초 코딩 교육
 - (2) 지역 거점 교육: 제주 자기주도학습센터 등 교육시설을 활용한 중급 이상의 코딩 교육
 - (3) 학부모와 함께하는 워크숍: 코딩을 단순한 대학진학 수단인 아닌 미래시대의 새로운 교양임을 인식시키기 위한 교육
 - (4) 제주로 On 코딩 해커톤: 초중고교 학생들과 문제해결 및 협력에 대한 가치 교육
- 효율적인 Code/Data Literacy 교육을 위해서 지역 교육청 등 교육기관과 지역 기

업, 공공기관들이 모여 코딩교육 거버넌스 협의체를 구성하여 강사 양성, 코딩 교육 프로그램 기획, 교재 기획 등을 함께 추진하며 지역 내에서 이루어지는 코딩 교육 관련 이벤트들을 하나의 방향성으로 통합하는 체계를 함께 진행하였다. 단일 기관이나 단일 기업에서 지역사회 전반에 걸쳐 코딩 교육을 진행하는 것은 한계가 있는데다가 각기 진행되는 코딩 교육을 수준별, 단계별로 묶어 하나의 체계를 구성하기 위해서는 이러한 협의체 구성이 필수적이다.

제주로 On 코딩 거버너스 협의체



2017년에 진행된 제주로 On 코딩 해커톤에서는 고등학교 IT동아리를 중심으로 3개월간 사회적 문제 발굴 및 코드/데이터 기반의 솔루션을 개발해 제안하는 행사를 진행하였다. 유명 관광지 중심으로 이루어지고 있는 관광산업의 편중현상에 대한 해결을 위한 알려지지 않은 관광지에 대한 소개, 불법주차 문제 해결을 위한 주차장 안내 서비스, 이주민과 토착민들 사이에서 발생하는 정보 접근 차이에 대한 해결을 위한 정보 공유 서비스 등 다양한 사회적 문제에 대한 자발적인 분석 및 해결 방안들이 제시되어 큰 반향을 이끌어냈다.

제주로 On 코딩 해커톤에 출품된 프로젝트 설명 자료들의 일부

2조. 오늘 저녁은 코딩이야

제주 문화 여행 Jeju culture tour

이름: 제주문화여행 (Jeju Culture Tour) | 소개글: 제주도를 여행하는 사람들의 마음을 사로잡는 제주 문화 여행입니다. 제주도의 아름다운 자연 경관과 독특한 문화를 소개합니다. 제주도의 대표 음식과 전통 문화를 소개합니다. 제주도의 대표 관광지 소개합니다.

구현방법

- 1. HTML/CSS를 사용하여 웹페이지를 제작합니다 (index.html)
- 2. Bootstrap을 사용하여 레이아웃을 구성합니다 (index.html)
- 3. JavaScript를 사용하여 동적 효과를 추가합니다 (index.html)
- 4. jQuery를 사용하여 DOM 조작을 쉽게 합니다 (index.html)
- 5. Google Maps API를 사용하여 지도를 표시합니다 (index.html)

사용 프로그램

- Atom
- Visual Studio
- Sublime Text
- Bootstrap
- jQuery
- jQuery UI
- jQuery Validation Plugin
- jQuery Mobile
- jQuery EasyUI
- jQuery UI ThemeRoller
- jQuery UI Widget Factory
- jQuery UI Sortable
- jQuery UI Draggable
- jQuery UI Droppable
- jQuery UI Resizable
- jQuery UI Selectable
- jQuery UI Sortable
- jQuery UI Draggable
- jQuery UI Droppable
- jQuery UI Resizable
- jQuery UI Selectable

디자인

- 1. Bootstrap을 사용하여 레이아웃을 구성합니다 (index.html)
- 2. jQuery를 사용하여 DOM 조작을 쉽게 합니다 (index.html)
- 3. jQuery UI를 사용하여 UI 효과를 추가합니다 (index.html)
- 4. jQuery Validation Plugin을 사용하여 폼 유효성 검사를 합니다 (index.html)
- 5. jQuery Mobile을 사용하여 모바일 친화적인 디자인을 만듭니다 (index.html)

작업 방법

- 1. 팀원들끼리 아이디어를 내고 아이디어를 바탕으로 웹페이지를 제작합니다
- 2. 팀원들끼리 코딩을 하고 코딩한 내용을 팀원들끼리 검토합니다
- 3. 팀원들끼리 코딩을 하고 코딩한 내용을 팀원들끼리 검토합니다
- 4. 팀원들끼리 코딩을 하고 코딩한 내용을 팀원들끼리 검토합니다
- 5. 팀원들끼리 코딩을 하고 코딩한 내용을 팀원들끼리 검토합니다

4조. 코프이 데하나왔다

제주 주차장 Jeju parking service

이름: 제주주차장 (Jeju Parking Service) | 소개글: 제주도를 여행하는 사람들의 마음을 사로잡는 제주 주차장 서비스입니다. 제주도의 아름다운 자연 경관과 독특한 문화를 소개합니다. 제주도의 대표 음식과 전통 문화를 소개합니다. 제주도의 대표 관광지 소개합니다.

구현방법

- 1. Bootstrap을 사용하여 레이아웃을 구성합니다 (index.html)
- 2. jQuery를 사용하여 DOM 조작을 쉽게 합니다 (index.html)
- 3. jQuery UI를 사용하여 UI 효과를 추가합니다 (index.html)
- 4. jQuery Validation Plugin을 사용하여 폼 유효성 검사를 합니다 (index.html)
- 5. jQuery Mobile을 사용하여 모바일 친화적인 디자인을 만듭니다 (index.html)

사용 프로그램

- Atom
- Visual Studio
- Sublime Text
- Bootstrap
- jQuery
- jQuery UI
- jQuery Validation Plugin
- jQuery Mobile
- jQuery EasyUI
- jQuery UI ThemeRoller
- jQuery UI Widget Factory
- jQuery UI Sortable
- jQuery UI Draggable
- jQuery UI Droppable
- jQuery UI Resizable
- jQuery UI Selectable
- jQuery UI Sortable
- jQuery UI Draggable
- jQuery UI Droppable
- jQuery UI Resizable
- jQuery UI Selectable

디자인

- 1. Bootstrap을 사용하여 레이아웃을 구성합니다 (index.html)
- 2. jQuery를 사용하여 DOM 조작을 쉽게 합니다 (index.html)
- 3. jQuery UI를 사용하여 UI 효과를 추가합니다 (index.html)
- 4. jQuery Validation Plugin을 사용하여 폼 유효성 검사를 합니다 (index.html)
- 5. jQuery Mobile을 사용하여 모바일 친화적인 디자인을 만듭니다 (index.html)

작업 방법

- 1. 팀원들끼리 아이디어를 내고 아이디어를 바탕으로 웹페이지를 제작합니다
- 2. 팀원들끼리 코딩을 하고 코딩한 내용을 팀원들끼리 검토합니다
- 3. 팀원들끼리 코딩을 하고 코딩한 내용을 팀원들끼리 검토합니다
- 4. 팀원들끼리 코딩을 하고 코딩한 내용을 팀원들끼리 검토합니다
- 5. 팀원들끼리 코딩을 하고 코딩한 내용을 팀원들끼리 검토합니다

8조. 오토 코딩팀

Link+ Jeju 제주를 연결하다

이름: 링크플러스 (Link+ Jeju) | 소개글: 제주도를 여행하는 사람들의 마음을 사로잡는 링크플러스 서비스입니다. 제주도의 아름다운 자연 경관과 독특한 문화를 소개합니다. 제주도의 대표 음식과 전통 문화를 소개합니다. 제주도의 대표 관광지 소개합니다.

구현방법

- 1. Bootstrap을 사용하여 레이아웃을 구성합니다 (index.html)
- 2. jQuery를 사용하여 DOM 조작을 쉽게 합니다 (index.html)
- 3. jQuery UI를 사용하여 UI 효과를 추가합니다 (index.html)
- 4. jQuery Validation Plugin을 사용하여 폼 유효성 검사를 합니다 (index.html)
- 5. jQuery Mobile을 사용하여 모바일 친화적인 디자인을 만듭니다 (index.html)

사용 프로그램

- Atom
- Visual Studio
- Sublime Text
- Bootstrap
- jQuery
- jQuery UI
- jQuery Validation Plugin
- jQuery Mobile
- jQuery EasyUI
- jQuery UI ThemeRoller
- jQuery UI Widget Factory
- jQuery UI Sortable
- jQuery UI Draggable
- jQuery UI Droppable
- jQuery UI Resizable
- jQuery UI Selectable
- jQuery UI Sortable
- jQuery UI Draggable
- jQuery UI Droppable
- jQuery UI Resizable
- jQuery UI Selectable

디자인

- 1. Bootstrap을 사용하여 레이아웃을 구성합니다 (index.html)
- 2. jQuery를 사용하여 DOM 조작을 쉽게 합니다 (index.html)
- 3. jQuery UI를 사용하여 UI 효과를 추가합니다 (index.html)
- 4. jQuery Validation Plugin을 사용하여 폼 유효성 검사를 합니다 (index.html)
- 5. jQuery Mobile을 사용하여 모바일 친화적인 디자인을 만듭니다 (index.html)

작업 방법

- 1. 팀원들끼리 아이디어를 내고 아이디어를 바탕으로 웹페이지를 제작합니다
- 2. 팀원들끼리 코딩을 하고 코딩한 내용을 팀원들끼리 검토합니다
- 3. 팀원들끼리 코딩을 하고 코딩한 내용을 팀원들끼리 검토합니다
- 4. 팀원들끼리 코딩을 하고 코딩한 내용을 팀원들끼리 검토합니다
- 5. 팀원들끼리 코딩을 하고 코딩한 내용을 팀원들끼리 검토합니다

지역민들의 지역 현안에 대한 관심도를 이끌어 내는 것과 더불어, 실제 코드와 데이터를 통해 문제 해결에 접근하도록 Code/ Data literacy에 대한 교육은 향후 성인까지 확장하여, 다양한 시민들의 제안들을 통해 지역 현안을 자발적으로 해소하는 분위기를 만드는 데에 제주도의 코딩 교육의 주요 목표가 설정되어 있다.

공공 데이터와 민간데이터, 이를 유통/활용하기 위한 플랫폼 구조와 지역현안에 관심 있는 미래세대, 시민들의 코드/데이터 활용 능력을 향상시킴으로써 새로운 형태의 사회 혁신에 접근할 수 있을 것으로 기대하고 있다.

7. 향후과제

2015년부터 추진되었던 지속가능한 빅데이터 활용 체계를 만들기 위한 제주도의 정책들은 1) 지자체 레벨에서 가치 있는 데이터를 직접 생산, 2) 민간의 데이터를 발굴하고 공공 데이터와 융합하여 효율적인 데이터 유통체계 구축, 3) 시민들의 데이터 리터러시 확산이라는 3개의 방향성으로 추진되고 있다.

2016년에는 Wi-Fi 기반 이동경로 분석 및 데이터 생산에 대한 기초 실험과 표준 빅데이터 플랫폼의 기초 모델들이 구축되었고, 2017년까지 기초적인 플랫폼들과 체계들이 구축되었고, 주요 데이터들에 대한 식별 및 확보 작업들이 순차적으로 진행되어왔다. 또한, 데이터 유통 체계에 대한 실험 및 컨설팅을 통한 데이터 활용 사례 도출에 대한 실험들을 진행하였다.

향후과제로, 본격적인 데이터 활용/유통 체계를 구축하여 본격적인 지역 데이터 생태계를 구축하는 것과 업무부서 간 데이터 활용 프로세스 도입을 통한 과학적 행정 처리 모델에 대한 구축을 추진할 예정이다.

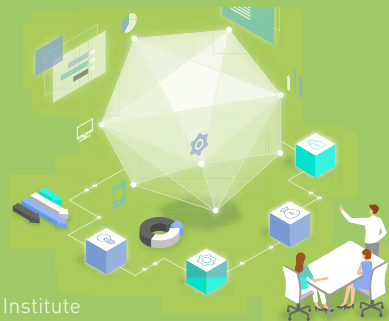
단발성 분석 과제나 1회성의 현황 분석을 위한 빅데이터 사업도 물론 가치가 있지만, 보다 광범위한 영역에서 지속가능한 빅데이터 활용 체계를 만드는 시도는 앞으로 제주특별자치도를 비롯하여 데이터에 대한 가치와 활용도를 끌어올리고자 하는 타 지자체에서도 추진되어야 할 것이다.

CHAPTER

I

지방자치단체 빅데이터 분석 사례집

부산광역시 도시양극화 분석 사례



korea Local Information Research & Development Institute

가. 추진배경	34
나. 프로젝트 개요	34
다. 적용 프로세스	39
라. 기획·관리 단계	40
마. 데이터 분석 단계	53
바. 행정적용	64
사. 인프라 구축 단계	66

I. 부산광역시 도시양극화 분석 사례

요 약

부산광역시에서는 ‘데이터 기반’의 사회 현안 해결과 이슈 모니터링을 위한 미래창조과학부·한국정보화진흥원 주관 ‘2017년도 빅데이터 플래그십(기획·검증) 시범사업’에 국책연구기관인 국토연구원과 신용·부채 데이터를 보유하고 있는 개인신용평가 전문회사인 코리아크레딧뷰로(KCB)와 공동 응모하여 ‘도시 양극화 모니터링 및 시뮬레이션’ 과제가 선정되어 추진하였다.

도시양극화는 우리나라의 경제 성장이 둔화하면서 사회·경제적 계층 이동의 기회도 감소함에 따라 도시 내에서 공간적으로 발생, 진행, 고착되고 있다. (도시공간의 부유화(젠트리피케이션)와 분리화(세그리게이션) 등 공간적인 차원에서 표출) 최근 도시양극화 문제에 대한 사회적 관심이 증대하고 있으나, 구체적인 데이터와 과학적인 방법에 기반하여 도시양극화의 실태를 보다 체계적으로 분석, 모니터링, 예측할 수 있는 과학적 정책지원 수단은 미흡한 상황이다. 이에 본 프로젝트에서는, 도시양극화 해소와 완화를 위해, 도시양극화를 진단, 모니터링, 예측할 수 있는 구체적인 데이터와 과학적 분석모형 등을 마련하고, 이를 국토·도시정책에 체계적으로 반영해 나갈 수 있는 제도를 개선을 위해 양극화 문제를 금융(신용/부채)·공간(지리정보 GIS) 빅데이터로 분석, 시뮬레이션하고 시각화 하고자하였다.

주요 분석 내용은 도시 양극화 분석을 위한 금융 빅데이터와 공간 빅데이터의 융합, 도시 양극화 분석 및 시뮬레이션 모형 개발, 도시 양극화 분석·시뮬레이션 및 결과 서비스를 위한 시범시스템 구축, 부산시에 대한 도시 양극화 분석 및 시뮬레이션 수행, 부산시 개방형 빅데이터 플랫폼을 연계·활용한 서비스 제공이며, 분석 결과는 부산시가 운영하는 개방형 빅데이터 플랫폼을 활용하여 확인 가능하다.

이번 프로젝트는 도시 양극화 대응을 위해 부산시가 전국에서 빅데이터 기반 과학적 분석방법으로 처음 시도한 사례로 의의가 있으며, 이를 통해 도시 계획 및 도시재생 의사결정, 지역 사회·경제적 변동 추이 파악 등에 적극 활용할 수 있었다.

가 추진배경

- 경제성장이 둔화되고 고착되면서 사회경제적 계층 간 이동의 기회도 함께 감소함에 따라 양극화 문제가 새로운 사회경제적 문제로 대두 및 심화되고 있음
 - 기획재정부는 2018년도 예산투자 4대 핵심 분야의 하나로 ‘양극화 대응’을 선정
- 사회 경제적 양극화는 국토·도시 등 공간적 차원에서 부유화(Gentrification)와 분리화(Segregation) 등과 같은 현상으로 표출
 - 기존에 비교적 낮은 가격의 주택지에 도시재개발 사업 등으로 인하여 주택가격이 상승 → 상승된 가격을 감당할 수 없는 저소득 계층의 이주와 동시에 그 자리를 고소득 계층들이 채우게 되면서 부유화(젠트리피케이션) 발생 → 기존 저소득 계층들은 가격 감당이 가능한 지역으로 이주하게 되면서 분리화(세그리게이션) 발생
- 도시양극화 문제에 대한 정확한 진단과 효과적인 미래예측을 위하여 체계적이고 구체적인 데이터 기반의 과학적 접근 필요
- 빅데이터 등의 새로운 데이터와 과학적 방법에 기반하여 도시양극화를 분석, 모니터링, 예측할 필요가 있으며, 이를 국토·도시정책의 수립과 집행에 활용할 필요가 있음

나 프로젝트 개요

- 개요
 - 미래창조과학부의 “2017년 빅데이터 플래그십” 공모 사업으로 부산광역시와 국토연구원 이 2017년 6월부터 6개월간 금융 빅데이터와 공간 빅데이터를 활용한 도시양극화 분석 및 시뮬레이션 프로젝트를 추진
- 사업의 주요 내용은 금융 빅데이터와 공간 빅데이터를 활용하여 부산광역시 도시양극화 실증분석을 수행하는 것임

- 결과를 토대로 기존 사회경제통계의 한계를 극복하고 실효성 있는 정책 수립 추진을 목적으로 실시
 - 서민주거안정, 서민주거환경개선, 다복동 사업도시재생뉴딜사업

■ 분석의 필요성

- 최근 도시양극화 문제에 대한 사회적 관심이 증대하고 있으나 주로 정성적인 측면에서 논의되고 있으며, 기존의 사회경제통계 자료 등을 대체하거나 보완하기 위한 데이터와 분석방법의 개발에 대한 노력은 아직 미흡
- 부산광역시는 공간적 분리정도가 심화되고 있는 추세로 주민생활안정 및 도시양극화 완화를 위한 과학적 접근 방식 필요
- 분석결과는 과학적 의사결정 자료로 활용할 수 있으며 향후에도 타 도시 및 중앙부처의 도시양극화 대응 관련 정책에 대한 시사점을 제시해 줄 수 있을 것으로 판단

■ 주요 분석 내용

- 도시양극화 실증 분석을 위해 부산광역시 약 270만 명에 대한 과거 5년 간 금융 데이터를 주소 지 단위로 구축 하고, 정부부처에서 생산·관리하고 있는 정부 공공데이터 등을 취득 후 가공
- 소득, 지가, 교통 접근성, 공공서비스 접근성, 유통서비스 접근성, 자연환경 등을 도시 양극화를 발생시키는 주요한 지표로 고려하여 도시양극화 발생 확률 분석

도시양극화의 개념과 정책개입 필요성

◆ 젠트리피케이션(Gentrification)이란?

- 상류층(신사)을 뜻하는 gentry 에서 파생된 말로 "낙후된 지역을 고급화하다"라는 뜻을 가지고 있지만, 현실에선 주택가격이 저렴한 낙후된 지역에 중산층 이상의 사람들이 몰리고 지역이 발전하면서 기존에 거주 중이던 원주민을 밀어내는 현상을 말한다. 영국의 학자 Ruths Glass가 최초로 사용함

◆ 세그리게이션(Segregation)이란?

- 서로 다른 계층 간 주거지 등이 분리되는 현상으로 서구 선진국 등에서는 주로 인종별 분리 현상이 나타나고 있으나 최근에는 소득계층에 따른 주거 분리가 나타나고 있으며 우리나라의 경우 인종 간 주거 분리 현상은 없으나 소득에 따른 분리 현상은 나타나고 있는 것으로 파악됨

◆ 경제양극화 완화의 필요성

- 경제양극화는 소득격차가 심화되는 것을 의미하는데 상위소득 계층보다는 하위소득 계층에 대한 부정적 파급효과가 크게 나타나게 되며 소득격차 완화를 통해 서민생활을 안정시키고 사회적 통합을 강화할 필요가 있음

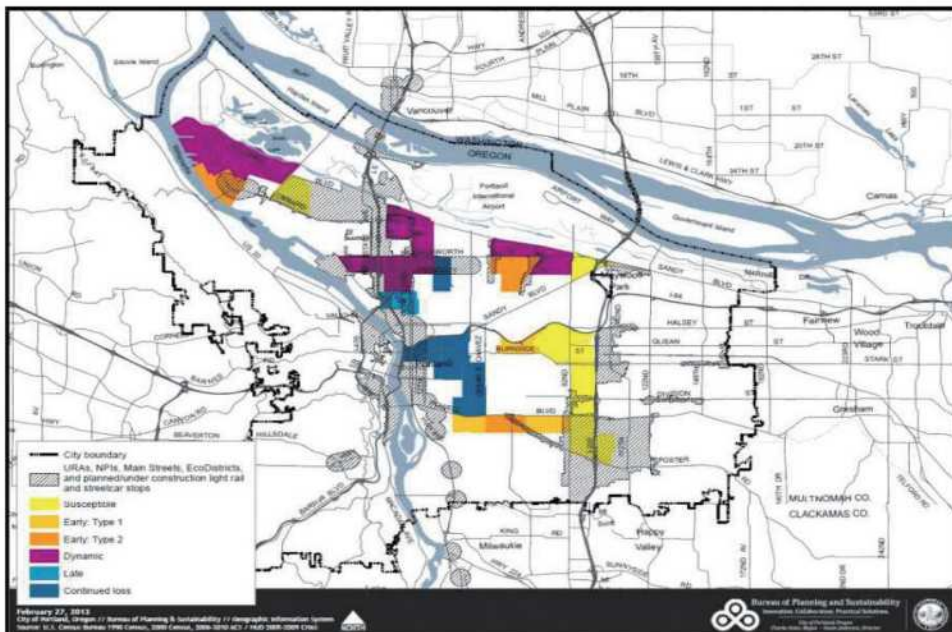
◆ 도시양극화 완화의 필요성

- 한편 경제양극화는 소득에 따른 주거분리를 통해 도시양극화 형태로 나타나게 되며, 상위소득 계층과 하위소득 계층의 주거지의 공간격차를 완화하여 서민주거안정, 서민주거환경개선을 도모할 필요가 있음

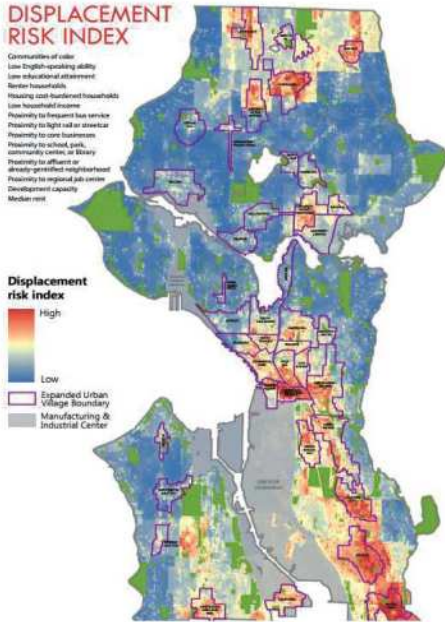
외국의 도시양극화 분석사례

◆ 미국 포틀랜드 사례

- 포틀랜드(Portland) 도시계획 및 지속가능발전부(Bureau of Planning and Sustainability: BPS)는 젠트리피케이션 취약성 분석을 수행하여 정책에 반영
- 인구통계 데이터(인종 및 인구 변화 등), 소득 및 교육수준, 주거 점유기간 및 가격변동 등의 공공데이터 활용
- 정부의 지원이 필요한 소생활권(neighborhoods) 정책분석을 위한 도구로 활용
- 분석결과를 통해 정책입안자는 정부의 개입이 필요한 지역을 보다 쉽게 선정할 수 있으며 결과를 토대로 주택수리 및 담보대출, 저렴한 주택공급, 소득에 알맞은 주거지를 선정할 수 있도록 5개년 계획을 수립



(계속) 외국의 도시양극화 분석사례



◆ 미국 시애틀 사례

- 시애틀 도시계획 및 도시개발부(Department of Planning and Development: DPD)는 2015 Growth and Equity 일환으로 도시 전역에 대한 비자발적 이주(displacement)의 위험도를 분석
- 임대료 인상 등에 대한 주민 취약성 편의시설에 대한 차별 토지이용규제에 근거한 개발역량을 이탈의 요인으로 판단 공공데이터 등을 활용하여 각 도시의 위험지수를 분석
- 취약성 지표(교육수준 인구통계 및 주택비용 부담 정도), 편의시설 지표(대중교통 주력사업 인프라), 개발역량 지표(토지이용규제 규칙 및 평균 임대료) 등을 도출
- 분석결과를 도시계획을 위한 기초자료로 활용, 성장과 형평이 균형을 이루는 포용적 도시성장을 위한 정책수립에 활용

◆ 외국 사례의 시사점

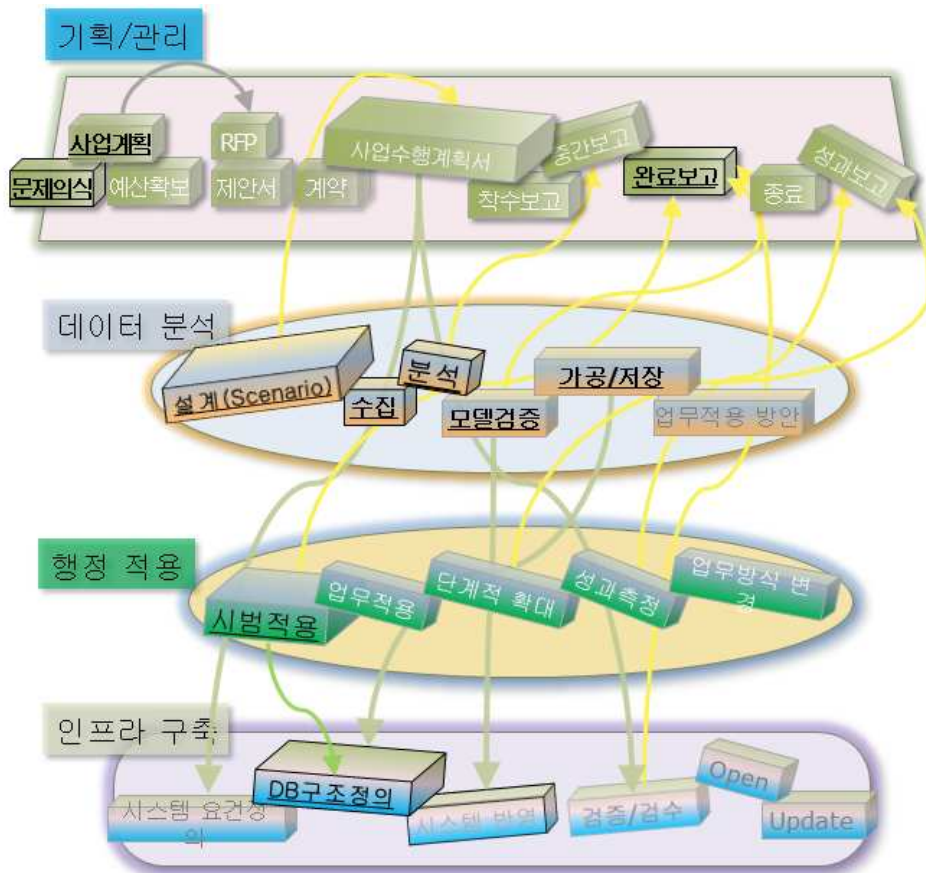
- 외국의 사례를 살펴보면, 도시양극화를 분석함에 있어 사회적 계층의 공간적 분포를 파악하기 위해 다양한 자료들을 활용하고 있는데, 주로 시공간적인 구체성에는 한계가 있는 인구통계 조사 자료를 활용
- 그러나 사회경제환경의 변화 속도가 상대적으로 빠른 우리나라의 경우, 실태를 보다 정확하게 반영하고 생활 체감적인 분석결과를 도출하기 위해서는 기존의 사회경제 통계자료의 한계를 극복하는 새로운 빅데이터를 활용할 필요가 있음

다 적용 프로세스

■ 부산광역시 도시양극화 분석 전체 프로세스

- 기획관리(문제의식 → 사업계획 → 완료보고)
- 데이터분석(분석 → 모델설계 → 데이터수집 → 분석·검증 → 데이터 가공·저장)
- 행정적용(시범적용)
- 시스템구축(DB구조정의)

부산광역시 적용 프로세스



라 기획 · 관리 단계

1. 기획 · 관리 단계 요약

단계	주요 내용
문제의식	<ul style="list-style-type: none"> ◆ 부산광역시의 경우 노후불량주거지역과 신규개발지역의 주거환경 격차와 공간적 분리 정도가 심하고, 공간적 위치에 따라 특정 지역에서의 상·하위 소득 격차의 비율이 높게 나타나는 등 도시양극화가 심한 것으로 판단 ◆ 도시양극화를 발생 시킬 수 있는 주요한 지표를 고려하여 발생확률을 사전에 예측하고 도시양극화에 대한 다양한 측면과 영향요인 등을 파악할 필요가 있음
사업계획	<ul style="list-style-type: none"> ◆ 데이터 연계확보계획 ◆ 데이터 확보 및 운영관리 계획 ◆ 데이터 분석 방법 및 분석인력 운용계획 ◆ 과제수행을 위한 연계협력방안 계획 ◆ 시스템구축계획
국비사업 공모	<ul style="list-style-type: none"> ◆ '2017년 빅데이터 플래그십 시범사업'은 미래창조과학부가 제4차 산업혁명에 따른 지능정보사회의 핵심 자원인 빅데이터의 활용 확산을 위해 정부부처, 지방자치단체, 공공기관 및 기업·협회 등 민간기관을 대상으로 공모 ◆ 부산광역시·국토연구원(국책연구기관)·KCB(금융 빅데이터 보유)가 공동 응모하여 최종 지원 대상으로 선정된 후 사업 수행
완료발표	<ul style="list-style-type: none"> ◆ 착수발표, 최종발표가 진행되었으며 진행과정에서 관련 담당부서(도시계획과, 도시재생과 등) 담당자와의 면담을 통해 의견수렴을 받음

2. 문제의식

- 경제양극화는 단순히 사회계층의 극단화 문제가 아니라, 국민의 일상생활이 영위되는 도시공간의 부유화와 분리화 등 공간적인 차원에서 표출
- 부산광역시의 경우 도시 내에서 소득에 따른 공간적 분리 정도가 심하고, 일부 특정 지역에서의 소득격차가 높게 나타남
 - 도시양극화를 발생시킬 수 있는 주요한 지표를 고려하여 발생확률을 사전에 예측하고 양극화 발생 요인 등을 파악할 필요가 있음

과제	<ul style="list-style-type: none"> ◆ 문제점 확인 ◆ 문제해결을 위해 필요한 활동(대안)을 나열 ◆ 필요한 활동(대안)을 통해 얻을 수 있는 효과를 확인
추진내용	<ul style="list-style-type: none"> ◆ 문제점 확인 <ul style="list-style-type: none"> • 개인과 지역의 소득에 관한 자료는 지방자치단체 정책수립 시 근본적인 판단자료로 활용 가능함 • 현재 제공되는 사회경제통계정보를 통해 기초자치단체장은 지역 내 상세 소득정보를 알 수 없음(데이터 가용성의 문제) • 기초자치단체 내 세부적인 공간단위 중심의 소득 데이터 분석이 필요 ◆ 문제해결을 위해 필요한 활동 나열 <ul style="list-style-type: none"> • 현재 보유한 소득 관련 데이터가 무엇인지 확인 : 시군구별 가계통계 등 • 정책수립 시 필요한 분석/지수 정보가 무엇인지 확인 : 소득정보 등 ◆ 금융 데이터 분석을 통해 얻을 수 있는 효과 확인 <ul style="list-style-type: none"> • 도시 관련 정책 수립 시 대상이 되는 공간의 소득분포 특성 파악 가능
Tip	<ul style="list-style-type: none"> ◆ 주민의 소득과 관련한 문제가 지역에 미치는 영향은 매우 큼 ◆ 현재까지 소득통계는 통계청에서 전국의 약 4,000 가구를 표본으로 조사하는 가계동향조사에 의해 생산되고 있으며, 시군구, 읍면동 같은 소지역을 중심으로 통계가 조사되어 개선되었으나, 실제 자치단체에서 활용하는데 한계는 있음

3. 사업계획

- 사업에 대한 요구사항을 정의하여 계획을 수립하고 확정하는 단계
- 사업범위를 결정하고 소요자원 및 예산 등을 문서화하여 사업내용을 구체화함
 - 사업계획 수립 시 현황 및 문제점, 추진목표, 전략, 추진일정, 사업내용, 소요예산, 산출물 및 사업의 주체와 분석결과물의 활용 주체에 대한 정의가 필요

과제	<ul style="list-style-type: none"> ◆ 구체적이고 명확한 사업계획 수립 필요 ◆ 사업의 필요성과 목적, 기대효과 등을 기반으로 사업계획을 수립
추진내용	<ul style="list-style-type: none"> ◆ 구체적이고 명확한 사업계획 수립 <ul style="list-style-type: none"> • 사업의 필요성 확인 : 분석내용 및 결과물의 활용 측면에서 사업추진 필요성 제시 ◆ 데이터 연계확보계획 <ul style="list-style-type: none"> • 이 과제는 금융 빅데이터(민간 데이터)와 공간 빅데이터(공공 데이터)의 융합을 통해 수행 • 과제 종료 후 분석내용의 검증 및 시연과 금융 빅데이터의 구조 등을 파악하기 위한 일반 공개 목적으로는 KCB를 통해 제한된 범위와 형태로 공유 • 도시양극과 분석 및 시뮬레이션의 안정적 검증과 이후의 활용 등을 고려 일정기간 제공은 보장 • 다만 재식별 가능성 및 민원 등을 고려, 향후 일반 개방되는 데이터는 대상 지역의 특정시점, 블록별(또는 법정동) 단위의 집계정보 형태의 제한된 방식으로 제공 가능 • 과제 수행에 필요한 각종 공간 데이터는 정부의 공공 데이터 개방 정책에 따라 무상으로 공개되고 있는 자료를 활용하며, 국가공간정보포털 또는 소관부처의 홈페이지 등을 통해 구독 가능 ◆ 데이터 확보계획 <ul style="list-style-type: none"> • 금융(개인 신용) 빅데이터 수집 및 이용 방안

- 안정적이고 정확한 데이터의 확보 가능성 및 본 사업을 위한 맞춤형 정보 추출 등을 위해 개인신용정보를 보유한 코리아크레딧뷰로(KCB)를 사업 참여기관으로 함과 동시에 데이터 추출 및 자료 해석 역량이 있는 코리아크레딧뷰로의 담당 분야 직원도 연구 인력에 포함함으로써 보다 탄력적이면서 내실 있는 데이터 확보가 가능
- 공간 데이터와 금융 데이터와의 융합방안
 - 공간단위로 데이터 융합이 가능할 수 있도록 물리적 주소 및 공간좌표(위경도 등)를 기반으로 개별 데이터를 집계하거나 결합함으로써 이종(異種)간 데이터 결합을 지원
- 정기적인 도시 양극화 현상 모니터링을 위해 사업종료 후 단·중기 기간까지(예시 : 2019년 12월말) 대상 지역 데이터를 업데이트 하여 파일 형태로 제공 예정
- ◆ 데이터 운영관리계획
 - 개인신용정보의 비식별화 처리
 - 기본적으로 본 사업에 제공되는 데이터는 이미 금융위원회와 사전협의된 신용정보를 바탕으로 한 지역단위 집계(통계) 데이터이므로 별도의 비식별화 처리는 불필요
 - 다만 개인신용정보의 유출 및 유출된 데이터에 대한 재식별화 가능성을 고려 정부의 개인정보 비식별화 처리가이드(2016년)에 따라 개인정보 재식별 우려 등을 고려하여 “준식별자”를 기반으로 K익명화 처리 후 데이터 제공
- ◆ 데이터 분석 방법 및 분석인력 운용계획
 - 도시 양극화 진단 및 모니터링을 위해 탐색적 공간분석(Exploratory Spatial Data Analysis) 수행
 - 분류분석, 군집분석, 패턴분석, 집중도분석 등
 - 도시 양극화 시뮬레이션을 위해 사회경제 계층별 입지선택 영향요인 분석
 - 로지스틱 회귀분석을 통한 영향요인(determinant) 분석과 요인별 계수 값(coefficient) 추정
 - 행위자 기반모형 방법론을 활용하여 도시양극화 시뮬레이션 수행
 - 현 추세 연장 시나리오와 정책대응 시나리오 하에서 시뮬레이션 수행 및 결과 비교
 - Leaflet, Polymaps 등 오픈소스 웹맵핑 소프트웨어를 활용하여 도시

	<p>양극화 분석 결과를 시각적으로 표현 및 제공</p> <ul style="list-style-type: none"> - 사용자환경(GUI)을 디자인하여 상호작용(Interactive) 가능한 탐색 도구 제공 - 국토교통부에서 개발하여 오픈데이터로 제공하고 있는 브이월드(vWorld) 지도를 호출하고 배경지도로 사용하여 분석 결과의 현실적 이해도와 설명력 제고 <p>• 분석인력 운용계획</p> <ul style="list-style-type: none"> - 부산광역시는 본 연구의 실증 및 확산을 위해 자체 빅데이터 전문관 및 도시계획재생 업무담당자를 중심으로 과제의 타당성과 효율성을 높이하고자 함 - 국토연구원은 국토도시 분야의 정책분석과 관련한 다양한 연구를 수행하여 왔으며, 해당분야에 대한 학문적 전문성과 실무적 경험이 풍부한 다수의 연구 인력이 재직 <p>◆ 과제수행을 위한 연계협력방안 계획</p> <ul style="list-style-type: none"> • 국토연구원(KRIHS)의 내부 연구심의회와 외부 자문 회의를 통해 실무 협의체를 구성운영하여 국토연구원 내외부의 정책 및 데이터 분석 전문가들과 협력을 • 통계화된 민간 개인신용정보의 개인정보 여부 관련 법제도 검토 계획 <p>◆ 시스템구축계획</p> <ul style="list-style-type: none"> • 분석 모형, 시뮬레이션 모형, 시각화 모형은 오픈소스 소프트웨어를 활용하여 개발함으로써 상용 소프트웨어 구입에 따른 비용을 절감 • 분석 및 시뮬레이션 모형 결과의 서비스는 부산광역시의 개방형 빅데이터 플랫폼과 연계하여 제공 • 부산광역시의 빅데이터 플랫폼을 연계·활용함으로써 서비스 구현과 제공을 위해 필요한 하드웨어 및 소프트웨어 비용을 절감하고, 공공부문의 유관사업간 연계·협력을 도모하여 궁극적으로는 국가차원에서 공공투자 효율성을 제고
<p>Tip</p>	<ul style="list-style-type: none"> • 데이터의 가공처리에 시간과 비용이 소요되는 만큼, 이후의 데이터 제공 문제는 추가 협의 필요 • 개인신용정보의 경우 신용정보법에 따라 엄격하게 관리되고 그 사용처가 제한되고 있으나, 공공기관 대상, 연구 목적 사업인 경우 신용정보법 제 28조, 10항 7조에 따라 신용정보를 보유한 참여기관(신용조사회사인 코리아크레딧뷰로)는 국토연구원 및 사업주체인 미래창조과학부, 전문가인 정보화진흥원에 신용정보제공은 가능

4. 공모 준비

- '2017년 빅데이터 플래그십 시범사업'은 미래창조과학부가 제4차 산업혁명에 따른 지능정보사회의 핵심 자원인 빅데이터의 활용확산을 위해 정부부처, 지방자치단체, 공공기관 및 기업협회 등 민간기관을 대상으로 자유 공모를 추진
- 부산광역시·국토연구원(국책연구기관)·KCB(금융 빅데이터 보유)가 공동 응모하여 2차례의 심사과정을 거친 후 최종 지원 대상으로 선정된 후 사업 수행

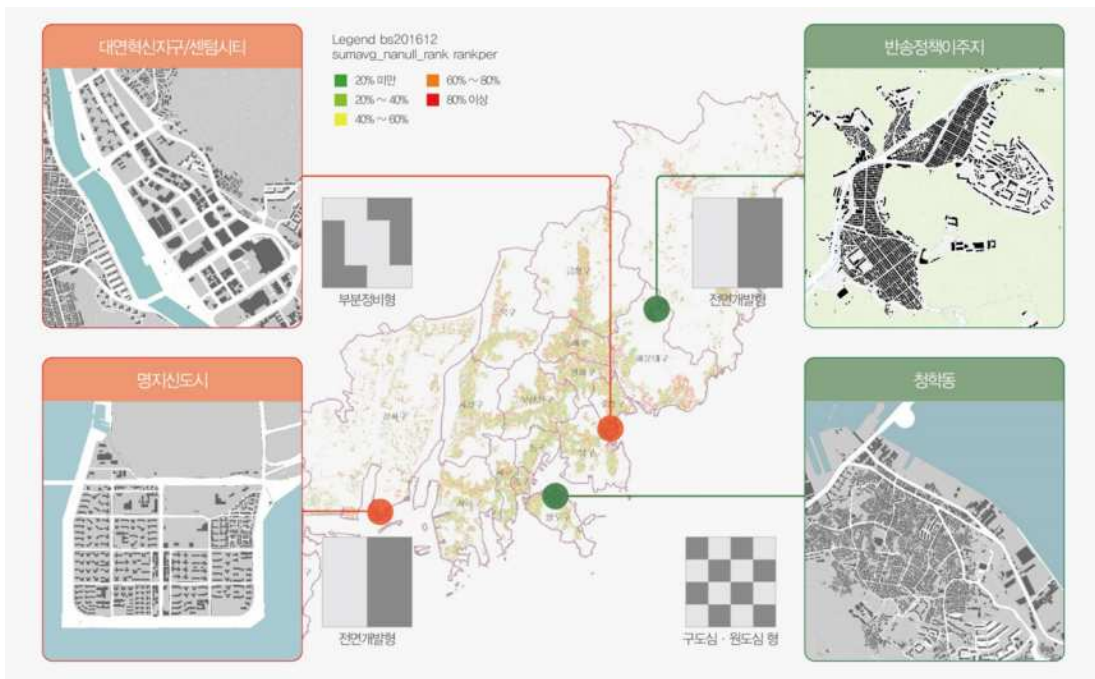
5. 완료보고

과제	◆ 분석 수행 결과물을 활용 주체에 맞게 업무 중심으로의 적용을 위한 활용방안 제시
추진 내용	<ul style="list-style-type: none"> ◆ 부산광역시를 대상으로 한 실증 적용 <ul style="list-style-type: none"> • 도시계획위원회 등 주기적으로 개최되는 도시 관리 및 도시개발사업 관련 의사결정 프로세스에 관련정보를 제공하고 도시 양극화 완화 및 사회 통합 제고를 위한 도시정책 의사결정에 활용하며 국토연구원에서 기 발간 중인 국토 그래픽 뉴스, 국토 정책브리프 등의 홍보수단으로 활용 ◆ 도시재생뉴딜사업 추진 시 동지내몰림 대응을 위한 가이드라인 마련 <ul style="list-style-type: none"> • 문재인 정부에서 중점적으로 추진하고자 하는 '도시재생뉴딜사업'에서는 각종 도시재생사업 추진 시 젠트리피케이션 대응을 주요한 추진전략의 하나로 제시하고 대응책을 마련하도록 하고 있음 ◆ 도시재생뉴딜사업 중 주거지 재생사업에 해당되는 우리동네살리기 사업과 주거지지원형 사업의 추진 시 이 연구의 결과를 활용하여 젠트리피케이션에 대응할 수 있도록 분석자료 및 분석지표 제공
Tip	◆ 현장적용에 대한 방안은 관련 협업부서와 충분한 사전협의를 거친 후 추진

- 부산광역시 도시계획과 및 도시재생과 등 관련 담당자 면담 및 국토연구원 내부의 정책 전문가와 다양한 연구기관 및 전문가 네트워크를 활용하여 협력체계 마련
 - 부산광역시청 도시계획과, 도시정비과, 노인복지과, 건축주택과 등
- 해당 결과에 대하여 착수발표, 결과발표가 진행
- 추가적으로 과학기술정보통신부 주최, 한국정보화진흥원 등이 주관하는 '4차 산업혁명 견인을 위한 2017 빅데이터 진흥주간 행사'에서 해당 내용을 논의
- 최종적으로는 부산광역시의 빅데이터 플랫폼을 연계·활용함으로써 서비스 구현과 제공을 위해 필요한 하드웨어 및 소프트웨어 비용을 절감하고, 공공부문의 유관사업간 연계·협력을 도모하여 궁극적으로 국가차원에서 공공투자 효율성 제고

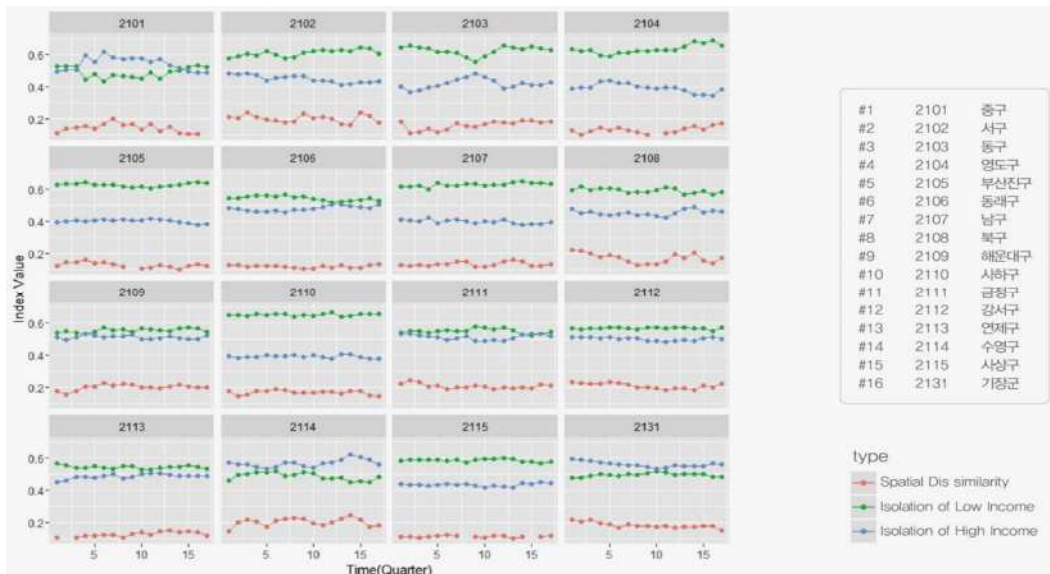
6. 결과물 샘플

- ◆ 부산광역시 개발 방식 및 장소에 따른 공간적 분리 형태
 - 소득구간별 공간분석 결과 상위 20% 소득계층은 주로 초고층 주상복합건축물과 대규모 아파트 단지를 중심으로 밀집
 - 해운대구 우동(마린시티), 남구 대연혁신지구(대연동), 강서구 명지국제신도시(명지동), 기장군 정관신도시(정관읍) 등 2000년대 이후에 신규 개발된 고층(복합) 아파트 단지를 중심으로 상위 소득계층이 군집하는 경향이 있음
 - 하위 20% 소득계층은 노후불량주택지역, 정책이주단지 등을 중심으로 밀집
 - 영도구, 동구 등 아파트의 비중이 비교적 낮고, 소규모 가구와 다세대주택의 비중이 높으며 차량 진입이 어려운 협소한 도로 등 생활 인프라 개선이 필요한 지역에 하위 소득계층이 광범위하게 거주



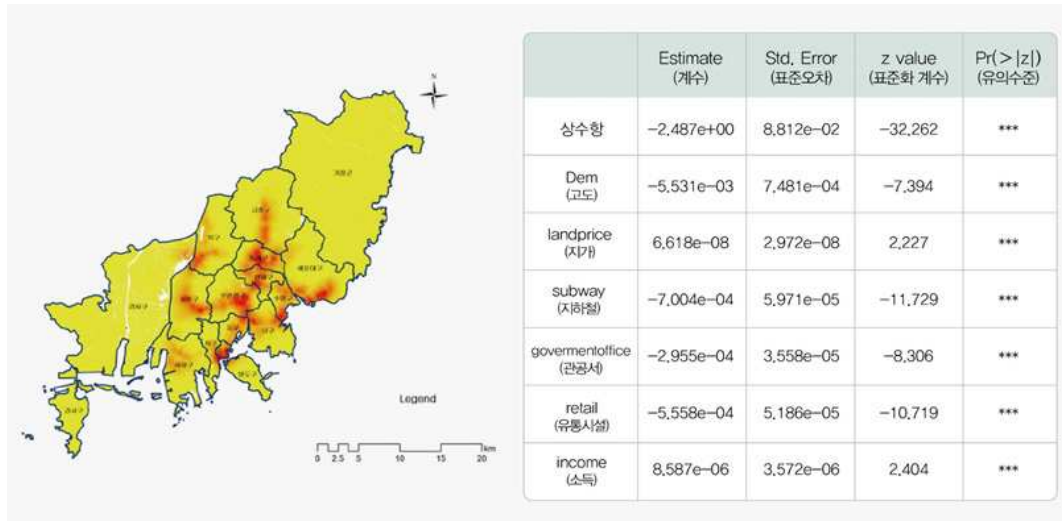
<주안점 2> 부산광역시 행정구역별 도시양극화 진단지수 개발

- 도시양극화의 정도를 계량적으로 진단하고 모니터링 할 수 있도록 공간적 비유사성 지수 (spatial dissimilarity index)와 공간적 고립지수(spatial isolation index)를 활용하여 도시 양극화 진단지수를 개발하고 행정구역별로 분석 수행
- 비유사성 지수는 공간적 현상의 균일성(evenness) 또는 군집성(clustering) 패턴을 측정할 수 있게 하며, 공간적 고립지수는 공간 내 특정 집단의 노출성(exposure) 또는 고립성 (isolation) 패턴을 측정할 수 있게 하는데, 각각 01 사이의 값을 가짐
- 비유사성 지수의 수치가 높을수록 상위소득과 하위소득의 공간적 분리 정도가 높다고 할 수 있으며, 공간적 고립 지수의 수치가 높을수록 해당 소득 계층의 고립도가 높다고 할 수 있음
- 예를 들어 가장 최근 시점을 기준으로 볼 때, 상위소득과 하위소득의 공간적 분리 정도는 해운대구에서 높게 나타남
- 소득계층별 고립정도는 상위 소득 계층의 경우 기장군과 수영구가 가장 높고, 하위 소득 계층의 경우 사하구와 중구의 고립도가 가장 높음



<주안점 3> 도시양극화 영향요인분석

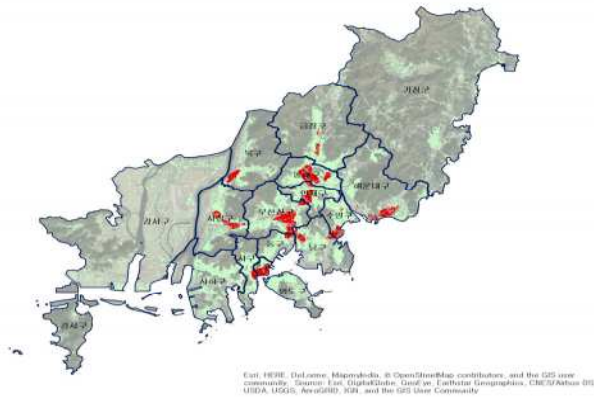
- 도시양극화에 대한 영향요인을 분석하고 이를 시뮬레이션 모형에 반영하기 위하여 로지스틱 회귀모형을 구성하고 분석을 수행
- 종속변수로는 과거 시행된 도시재개발 사업에 의해 소득이 증가한 지역, 즉 주거 젠트리피케이션이 발생한 지역을 고려
- 독립변수로는 소득, 지가, 교통 접근성, 공공서비스 접근성, 유통서비스 접근성, 자연환경 등을 고려
- 도출된 분석결과를 통해 도시양극화가 발생할 확률을 산정하고 공간 정보화함으로써 젠트리피케이션이 발생할 가능성이 높은 구체적인 위치를 도출



<주안점 4> 도시양극화 시뮬레이션

◆ 시나리오별 시뮬레이션 결과 1: 현 추세 연장 시나리오

- 이 시나리오는 현재와 같은 패턴으로 도시재개발사업이 진행되고 젠트리피케이션이 발생할 경우 미래에 나타나게 될 공간구조를 제시
- 즉, 노후불량주거지역 중 재개발 매력도가 높은 지역을 중심으로 대규모 고층아파트의 건설이 지속될 경우 상위 소득 계층의 입지선택이 활성화될 지역을 미리 파악



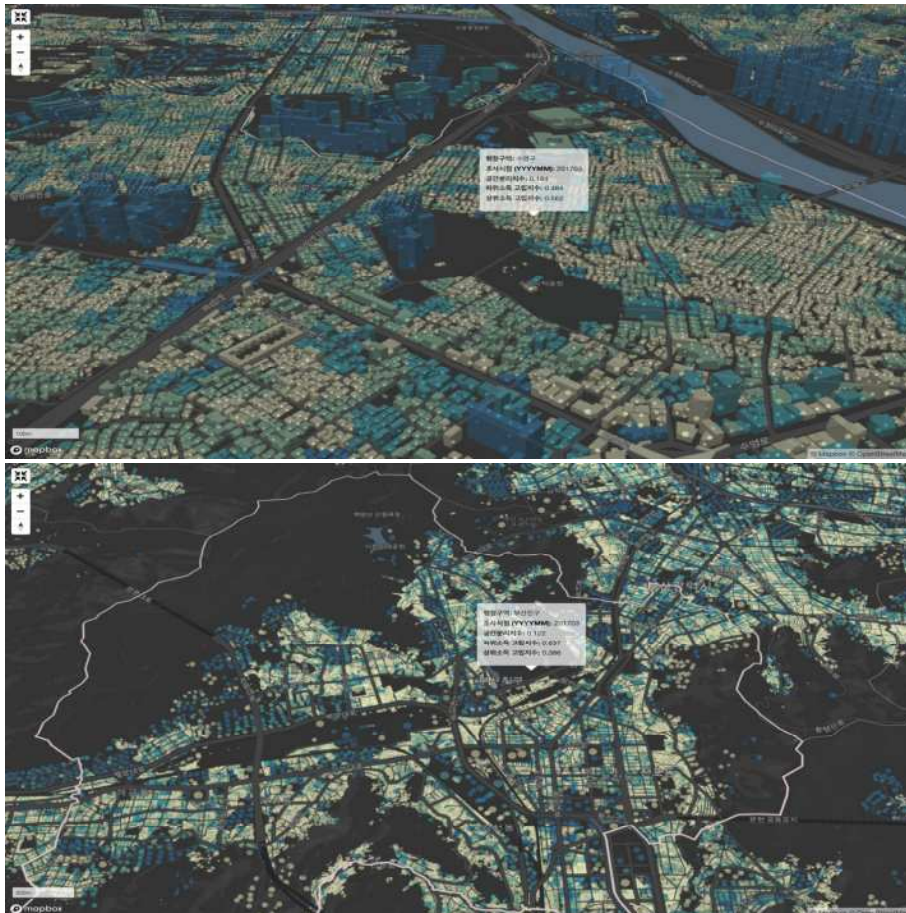
◆ 시나리오별 시뮬레이션 결과 2: 정책 시나리오

- 이 시나리오는 주거 젠트리피케이션을 방지하고 도시양극화를 완화하기 위한 정책 대안을 상정할 경우 미래에 나타나게 될 공간구조를 제시
- 노후불량주거지역에 대규모 주상복합아파트의 건설을 억제하고 소규모주택정비를 활성화할 경우 하위 소득 계층의 입지선택이 활성화될 지역을 미리 파악



<주안점 5> 분석 및 시뮬레이션 결과의 시각화

- 시각화는 인간 의사소통의 기본적인 수단의 하나이자, 분석 및 시뮬레이션 결과의 전달을 위해 매우 중요한 요소
- 분석된 결과가 주는 의미를 간결하게 전달하기 위해 색(Color)을 이용하여 분석에서 보여주고자 하는 Key Message를 강조
- 맵박스를 활용하여 분석 및 시뮬레이션 결과를 시각화하고자 하였음
- 맵박스는 사용자가 별도의 프로그래밍 과정 없이 기 제공되는 템플릿을 활용하여 공간정보를 시각화할 수 있는 차원부터, 자바스크립트 라이브러리 등을 활용하여 사용자가 필요에 따라 공간정보를 독창적으로 시각화 할 수 있는 수단까지 제공



- 다양한 형태의 빅데이터를 유자·관리하고 실시간으로 다양한 분석을 수행하여 그 결과를 시각화할 수 있는 빅데이터 플랫폼 서비스 제공
- 부산광역시는 빅데이터의 활용을 통해 내부행정을 효율화하고 대민서비스를 제고하기 위하여 개방형 빅데이터 플랫폼을 구축하고 관련된 정보를 제공하고자 함
- 부산광역시에서 구축하고 있는 빅데이터를 플랫폼 구축사업과 연계하여 도시양극화 분석 및 시뮬레이션 결과의 제공을 추진
- 이를 통해, 본 사업성과의 활용성을 제고하고 나아가 공공부문 사업간 연계효과 도모, 공공부문 예산절감 등을 도모하고자 함



마 데이터 분석 단계

1. 데이터 분석 단계 요약

단계	주요 내용
분석설계	<ul style="list-style-type: none"> ◆ 데이터를 분석하기 위한 계획을 세우는 것으로 분석모델을 설정하는 것 ◆ 분석목적과 활용목적을 명확히 하고 그에 적합한 방법론을 정하는 것
데이터 수집 및 가공	<ul style="list-style-type: none"> ◆ 분석에 필요한 데이터(raw data) 보유기관 확인 ◆ 적절한 협업을 통하여 필요 데이터 수집 ◆ 분석모델에 용이한 형태로 데이터 가공
데이터 분석	<ul style="list-style-type: none"> ◆ 1단계: 분류분석, 분포분석 ◆ 2단계: 도시양극화 진단지수 개발 ◆ 3단계: 도시양극화 영향요인 분석 ◆ 4단계: 도시양극화 시뮬레이션
데이터 시각화	<ul style="list-style-type: none"> ◆ 분석결과를 웹기반GIS 환경에서 정책결정자 및 일반시민에게 제공

2. 분석설계

- 분석설계란 결과를 도출하기 위해 필요한 전략을 세우는 일로, 분석을 계획하는 것
 - 분석설계에서는 분석을 통해 알아내고자 하는 대상(분석의 목적, 내용, 결과가 가지는 기대효과)을 명확하게 규정해야 함
 - 분석목적에 적합한 분석방법을 결정해야 함
 - 분석설계는 관련 주제에 대한 데이터 분석모델을 정하는 것임

<p>과제</p>	<ul style="list-style-type: none"> ◆ 분석결과의 활용 목적과 기대효과를 구체적으로 정의 ◆ 분석목적과 기대효과에 적합한 분석 방법론 선정 ◆ 분석 시나리오 작성 및 이에 근거하여 필요한 DATA의 종류, 유형 등 결정
<p>추진내용</p>	<ul style="list-style-type: none"> ◆ 분석결과의 활용목적과 기대효과를 구체적으로 정의 <ul style="list-style-type: none"> • 활용목적 : 부산광역시 도시양극화 분석의 목적은 소득격차에 따른 경제양극화가 공간적인 차원에서 도시양극화로 투영되는 정확한 현상을 파악하고, 서민주거안정과 서민주거환경개선을 위한 지방자치단체 정책 결정의 기반이 되는 구체적인 정보를 제공하기 위함 • 기대효과 : 도시양극화 분석을 통해, 정책결정의 기반(도시개발, 도시재생)을 마련하여 지역주민 맞춤형 행정서비스와 정책을 구현 ◆ 적합한 분석방법론 선정 <ul style="list-style-type: none"> • 금융 빅데이터를 분석하기 위한 프로젝트로, 소득 관련 원 자료를 취합하고 집계함 • 도시양극화에 영향을 미치는 요인을 파악하거나, 도시양극화를 예측하고자 할 때에는 로지스틱 회귀분석 방법을 사용할 수 있음 ◆ 분석 시나리오 및 분석에 필요한 데이터 항목, 유형 등을 결정 <ul style="list-style-type: none"> • 분석 시나리오 설정 : 연구내용을 명확하게 하고 분석에 불필요한 것들을 배제하는 등 원활한 분석활동을 위해 시나리오를 설정 ◆ 부산광역시 도시양극화 분석 시나리오 <ul style="list-style-type: none"> • (1단계) 市가 가장 관심을 가지는 주제와 관련한 핵심 수치(key Numbers)를 제시함 : 부산광역시 공간에 대한 소득분포 수준을 제시 • (2단계) 제시한 핵심 수치(Key Numbers)와 관련한 주제에 대한 기존 혹은 평소의 인식을 가설로 두고 분석하여 검증 : 상위소득 계층과 하위소득 계층의 소득격차에 따라 공간분리가 존재한다는 가설 하에 관련 데이터를 분석 • (3단계) 분석을 통해 보여주고자 하는 주요 메시지(Key Message)를 제시 : 도시양극화가 심한 지역을 확인하고, 해당 지역의 소득분포

	<p>패턴을 심층 분석</p> <ul style="list-style-type: none"> • (4단계) 市 내 특정지역을 예로 제시하고 관련 소득 데이터를 통해 종합적 분석한 후 분석결과를 일반화 : 도시양극화의 수준이 가장 높은 지역을 예시로 제시하고, 외부연계정보를 이용한 다차원적 분석을 통해 부산광역시 소득분포 수준을 범주화 <p>◆ 도시양극화 분석을 위해 필요한 데이터 분석 단위 설정</p> <ul style="list-style-type: none"> • 부산광역시는 개인정보비식별화 방침에 따라 데이터 분석 단위를 격자셀(grid cell)로 설정함 <p>◆ <참고> 격자셀(grid cell)이란?</p> <ul style="list-style-type: none"> • 분석의 정교성을 위해서는 주소지 단위인 포인트 데이터 형태로 사용하는 것이 이상적이나, 이 경우 개별 행위 주체에 관한 정보가 있어 개인정보유출에 대한 위험성이 있음 • 따라서 분석의 구체성을 확보하되, 개인정보 유출의 위험성을 방지하는 한편, 분석모델링의 효율성과 효과성을 고려하여 50m 셀 단위로 금융정보를 가공함 • 행안부에서 제공하는 국가격자망 데이터를 피쉬넷(fishnet) 기능으로 분할하여 50미터 격자망을 생성한 후 금융정보를 집계 처리: 개별 주체에 관한 정보는 없음
Tip	<p>◆ 분석설계는 분석을 위한 과정을 계획하는 것으로 "청사진"과 같음</p> <p>◆ 명확하고 구체적인 분석설계를 통해 분석 원래의 목적대로 진행해야 함</p>

3. 데이터 수집 및 가공

- 분석에 필요한 원자료(raw data)를 보유한 기관 등을 확인하고 적절한 협업을 통하여 데이터를 수집

과제	<ul style="list-style-type: none"> ◆ 분석에 필요한 데이터 리스트 및 보유기관 확인 및 수집 ◆ 수집 후 분석이 용이한 형태로의 데이터 가공
추진내용	<ul style="list-style-type: none"> ◆ 데이터의 수집 <ul style="list-style-type: none"> • 금융 빅데이터는 연구의 컨소시엄 기관으로 참여한 코리아크레딧뷰로(이하 KCB)를 통하여 구득 • 금융 빅데이터는 개별 금융기관이 개인에 대한 신용도를 평가하기 위한 목적으로 구축되는 것이나, 개인소득 등에 관한 구체적인 정보를 가지고 있어 다양한 사회경제분석 목적으로 활용가치가 높음 • 공간 빅데이터는 대부분 정부부처에서 생산·관리하고 있으며, 정부의 공공데이터 개방 정책에 따라 무상으로 공개되고 있음 • 온라인 다운로드, 오프라인 취득 등의 방법을 통해 필요한 데이터를 구득한 후 가공 ◆ 금융 빅데이터의 가공 <ul style="list-style-type: none"> • 사례 연구 대상지역인 부산광역시의 경우, 약 270만 명에 대한 과거 5년(2013~2017)간 데이터를 수집·가공·분석 • 금융 빅데이터는 원칙적으로는 개인에 대한 다양한 금융정보가 주소지 단위로 구축되어 있는 데이터임 • 따라서 금융 빅데이터는 다른 빅데이터와 마찬가지로 다양한 형태 또는 공간 단위로 가공하여 사용할 수 있음 • 대개의 경우 빅데이터를 활용할 때 분석의 용이성 등을 고려하여, 시·군·구, 읍면동 등 집계적인 공간단위를 기반으로 활용하는 경우가 많음 • 그러나 그렇게 할 경우, 빅데이터의 장점이라고 할 수 있는 미시적 분석이 어렵게 되므로, 이 연구에서는 개인정보를 보호하되 분석모델링의 효율성

	<p>과 효과성을 고려하여 50m 셀 단위로 금융정보를 가공함</p> <p>◆ 공간빅데이터의 가공</p> <ul style="list-style-type: none"> 공간 빅데이터는 자료별로 고유한 공간단위를 가지고 있어, 해당 공간 단위별로 자료를 구득한 뒤 필요에 따라 다양한 형태로 가공하여 사용 우리나라에서 생산되는 공간정보의 경우 서로 다른 좌표계를 사용하고 있는 경우가 있어(예: TM 좌표계 등), 공간정보의 분석을 위해서는 좌표계를 통일할 필요가 있으며 이 연구에서는 세계측지계(IRTf 2000)로 좌표변환 및 좌표계 통일
Tip	<p>◆ 분석의 정교성을 위해서는 주소지 단위인 포인트 데이터 형태로 사용하는 것이 이상적이나, 이 경우 개별 행위 주체에 관한 정보가 있어 개인정보유출에 대한 위험성이 있음</p> <p>◆ 공시지가 데이터는 주소지 단위의 정보를 필지단위에 지오코딩하여 장소성을 부여한 후 사용</p> <p>◆ 반면 지적 데이터는 기 구축된 자료를 구득한 후 추가적인 가공처리 없이 사용</p>

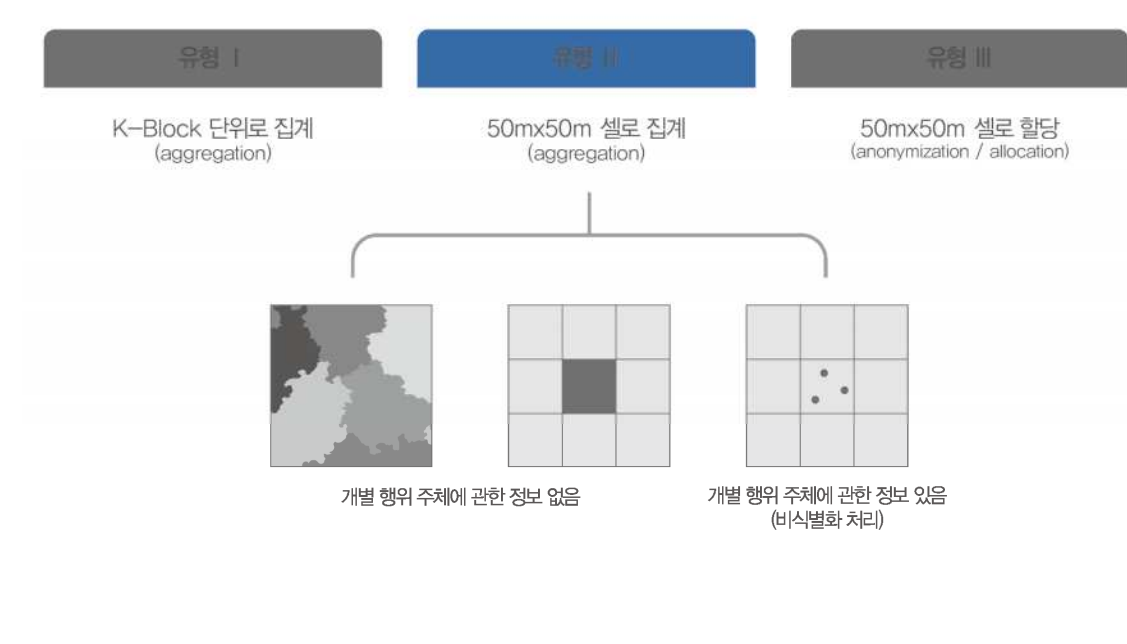
<참고> 금융·공간 빅데이터 구축 및 분석단위

- ◆ 금융 빅데이터는 원칙적으로는 개인에 대한 다양한 금융정보가 주소지 단위로 구축되어 있는 데이터임
 - 따라서 금융 빅데이터는 다른 빅데이터와 마찬가지로 다양한 형태 또는 공간 단위로 가공하여 사용할 수 있음
 - 대개의 경우 빅데이터를 활용할 때 분석의 용이성 등을 고려하여, 시군구, 읍면동 등 집계적인 공간단위를 기반으로 활용하는 경우가 많으나,
 - 그렇게 할 경우 빅데이터의 장점이라고 할 수 있는 미시적 분석이 어렵게 되므로, 이 연구에서는 개인정보를 보호하되 미시적인 분석이 가능하도록 다양한 공간단위에 대한 검토를 수행 (아래 표 참고)
 - 검토 결과 유형 II의 분석 활용성이 가장 좋은 것으로 판단하여 데이터를 구축함
- ◆ 금융 빅데이터 가공유형별 특징

유형 구분		유형별 특징
유형 I	마이크로 블록단위	<ul style="list-style-type: none"> • 마이크로 블록형태의 공간단위는 국내 대부분의 카드회사 등에서 활용하는 자료구축 및 분석단위로 상관분석을 위한 목적으로 구축된 후, 금융회사 등에서 보편적으로 활용이 되고 있음 • 그러나 공간구획의 기준 등이 임의적(adhoc)이며, 위치와 장소에 따라 공간단위의 크기가 상이한데 특히 인구밀도가 적은 도시외곽지역이나 비도시 지역에서는 분석 활용성이 떨어짐
유형 II	50m*50m 셀로 집계	<ul style="list-style-type: none"> • 분석의 정교성을 위해서는 주소지 단위인 포인트 데이터 형태로 사용하는 것이 이상적이나, 이 경우 개별 행위 주체에 관한 정보가 있어 개인정보유출에 대한 위험성이 있음 • 따라서 분석의 구체성을 확보하되, 개인정보 유출의 위험성을 방지하는 한편, 분석모델링의 효율성과 효과성을 고려하여 50m 셀 단위로 금융정보를 가공함 • 행안부에서 제공하는 국가격자망 데이터를 피쉬넷(fishnet) 기능으로 분할하여 50미터 격자망을 생성한 후 금융정보를 집계 처리: 개별 주체에 관한 정보는 없음
유형	50m*50m 셀로 할당	<ul style="list-style-type: none"> • 가공 유형 II는 개별주체에 관한 정보는 포함하지 않고 있어, 개별주체에 관한 정보를 포함하고 있는 가공 유형도 활용성을

Ⅲ		<p>검토함</p> <ul style="list-style-type: none"> 50미터 셀 단위로 개별주체에 관한 구체적인 기록을 할당하며 검토한 결과, 개인정보 유출의 방지를 위해 랜덤 노이즈를 발생시켜 사용하여야 하는 관계로, 공간분석 결과의 왜곡 가능성이 있어, 보다 심도 있는 검토 후에 활용이 필요할 것으로 판단되어 이 유형에 대한 활용은 보류
---	--	--

◆ 금융빅데이터 활용단위에 대한 검토



4. 데이터 분석

- 분류분석, 분포분석
- 도시양극화 진단지수 개발
- 도시양극화 영향요인 분석
- 도시양극화 시뮬레이션

과제	<ul style="list-style-type: none"> ◆ 1단계: 분류분석, 분포분석 ◆ 2단계: 도시양극화 진단지수 개발 ◆ 3단계: 도시양극화 영향요인 분석 ◆ 4단계: 도시양극화 시뮬레이션
추진내용	<ul style="list-style-type: none"> • 양극화는 소득격차로부터 비롯되는 사회경제적 현상이나, 소득격차에 따른 주거양극화(income segregation) 형태로 나타나게 되는데, 이때 각종 도시개발 사업에 의한 젠트리피케이션 현상 등에 의해 종합적으로는 도시양극화 현상으로 나타나게 되며, 도시 내 공간적 위치에 따라 다양한 형태로 발생 • 금융 빅데이터는 금융기관에서 활용하는 개인소득 등에 관한 사회경제 통계데이터이나 그리드 단위로 데이터를 구축하여 익명화, 비식별화 처리한 후, 이를 공간정보와 융합하고 분석을 수행하여 도시양극화의 다양한 측면과 영향요인 등을 파악 • 소득, 지가, 교통 접근성, 공공서비스 접근성, 유통서비스 접근성, 자연환경 등을 도시양극화를 발생시키는 주요한 지표로 고려하여 도시양극화 발생확률을 분석
Tip	<ul style="list-style-type: none"> ◆ 현장조사, 담당자면담 등을 통해 분석결과를 검증하면 결과의 타당성과 신뢰성을 제고할 수 있음

<참고> 왜 시뮬레이션을 하는가?

- ◆ 공간분석(spatial analysis)은 현재 상태에 대한 정보와 지식을 제공하나, 발생 가능한 미래에 대한 지식은 제공하지 않음
- ◆ 도시정책 및 계획은 현재의 상태를 기반으로 미래의 상태에 대한 의사결정을 수행하는 분야이며, 따라서 미래의 공간 상태에 대한 지식과 정보가 필요
- ◆ 도시모형(urban model)은 이러한 배경 하에 1960년대 라우리(Lowry) 모형을 필두로 발생한 뒤, 도시공간변화에 대한 원인과 결과를 예측하기 위한 목적으로 광범위하게 개발되고 활용되어 왔음
 - ※ 공간상호작용모형(spatial interaction model), 계량경제모형(econometrics model), 셀룰라 오토마타모형(cellular automata model), 행위자기반모형(agent based model) 등
- ◆ 과거에는 통계구역 단위의 자료와 모형이 주로 활용되었으나, 최근에는 보다 미시적인 셀 단위의 자료와 모형이 주로 활용됨
- ◆ 이 연구에서는 행위자 기반 모형을 활용하여 도시양극화 현상을 시뮬레이션 하고자 함
- ◆ 행위자 기반 모형은 복잡계 시스템(complexity systems)을 연구하기 위한 모델링 방법론으로 태동하였으나, 개별 행위자의 구체적인 상호작용을 기반으로 대상 시스템의 변화를 파악할 수 있는 장점이 있어, 자연과학 및 사회과학 전반으로 활용이 확산되고 있음
- ◆ 행위자 기반 모형은 행위규칙을 통해 행위자 간 또는 행위자와 환경 간의 상호작용을 규정하는 것으로 이를 통해 대상 시스템의 변화를 표현하며, 행위자(agent), 환경(environment), 행위규칙(decision making behaviour) 등의 구성요소를 가짐
- ◆ 행위자 기반 모형 자체가 각 구성요소들에 대한 구체적인 알고리즘을 제공하는 것은 아니며, 개별 모형에 따라 독자적이고 독창적인 알고리즘을 적용할 수 있음
- ◆ 도시모형 분야에서의 행위자 기반 모형은 다른 도시모형 방법론과 달리 도시의 공간변화를 유발하는 개별 행위주체(가구주, 개발자 등)의 입지선택행위를 기반으로 도시공간변화를 예측할 수 있는 장점이 있음

5. 데이터 시각화

과제	<p>◆ 분석결과를 웹기반GIS 환경에서 정책결정자 및 일반시민에게 제공</p>
추진내용	<ul style="list-style-type: none"> • 도시양극화 분석결과를 도시정책 및 계획을 담당하는 정책결정자, 실무담당자, 일반시민 등이 쉽게 접근하고 활용할 수 있도록 하기 위해서 웹기반 GIS 시각화 구현 필요 • 웹기반 GIS 시각화는 1990년대 후반 이후 관련 기술과 솔루션이 급격히 발전하였으며, 현재는 오픈소스 소프트웨어 및 상용 소프트웨어를 활용한 웹기반 GIS 시각화가 가능 • 특히 Leaflet, Polymaps, MapBox(이하 맵박스) 등은 간결하고 가벼운 라이브러리 등을 통해 복잡하고 고용량의 공간정보를 손쉽게 시각화하는 수단으로 활용이 급증하고 있음 • 이 연구에서는 맵박스를 활용하여 분석 및 시뮬레이션 결과를 시각화하고자 하며, 이를 부산광역시 빅데이터 홈페이지와 연계하여 서비스를 제공 • 맵박스는 사용자가 별도의 프로그래밍 과정 없이 기 제공되는 템플릿을 활용하여 공간정보를 시각화할 수 있는 차원부터, 자바스크립트 라이브러리 등을 활용하여 사용자가 필요에 따라 공간정보를 독창적으로 시각화 할 수 있는 수단까지 제공 • 이 연구에서는 양자를 절충하여 도시양극화 분석 및 시뮬레이션 결과를 시각화
Tip	<p>◆ 공간 빅데이터의 시각화를 위해 오픈소스 GIS 소프트웨어 활용 가능</p>

<참고> 왜 시각화가 중요한가?

- ◆ 시각화는 인간 의사소통의 기본적인 수단의 하나이자, 분석 및 시뮬레이션 결과의 전달을 위해 매우 중요한 요소임
- ◆ 정보를 일방향으로 전달하는 수단(means)일 뿐만이 아니라 양방향 의사소통의 중요한 매개체(medium)이기도 함
- ◆ Eppler and Burkard(2006)는 시각화의 유형을 과학적(scientific) 시각화, 정보(information) 시각화, 지식(knowledge) 시각화로 구분하였으며, 과학적 시각화와 정보 시각화는 객관적인 사실관계의 전달에 중점을 두는 반면, 지식 시각화는 가치관, 통찰력, 경험 등의 전달과 공유에 의미가 있다고 파악하였음
- ◆ 마찬가지로 빅데이터의 시각화는 분석결과를 활용하는데 매우 중요한 역할을 수행
- ◆ 분석에 대한 모든 결과를 살펴보는 것은 제약이 따르기에 효과적이고 효율적인 시각화 필요: 과학적 시각화
- ◆ 분석결과에 대한 특징(패턴 등)을 쉽게 파악하고 결과를 문서 또는 리포트 등으로 활용: 정보 시각화
- ◆ 분석 결과로 도출된 인사이트를 통해 사회현안 문제 논의: 지식 시각화
- ◆ 특히, 이 연구에서 분석·활용하는 금융·공간 빅데이터의 시각화를 위해서는 미시적인 단위의 공간정보를 효율적으로 시각화 할 수 있는 적절한 수단 필요

바

행정적용

1. 행정적용 단계 요약

단계	주요 내용
시범적용	<ul style="list-style-type: none"> ◆ 데이터를 분석을 통해 나온 결과를 통해 행정 개선, 시민 서비스 향상의 목적으로 실제 업무에 적용하기 전에 사전 효과성 검증 차원에서 진행하고, 분석 데이터 적용 이전과 이후에 대한 업무의 효율성 등을 측정하기 위한 지표를 만들고 진행해야 함(업무 처리 시간의 축소, 업무 투입 인력의 감축, 비용 감소, 처리 프로세스 단축 등)
업무 적용	<ul style="list-style-type: none"> ◆ 시범 분석을 통해 효과성이 검증된 분석 기법을 실제 업무에 적용하는 단계
단계적 확대	<ul style="list-style-type: none"> ◆ 확대 방안(1) : 적용 업무의 확대 ◆ 확대 방안(2) : 적용 조직(부서)의 확대 ◆ 확대 방안(3) : 분석 대상과 시기의 확대
성과측정	<ul style="list-style-type: none"> ◆ 분석사업 실시 이전과 이후의 성과를 측정하기 위한 방안을 마련하고, 시범적용 시 도출했던 지표를 보다 구체화하여 정기적으로 측정할 수 있도록 함
업무방식 변경	<ul style="list-style-type: none"> ◆ 분석사업 적용을 통해 업무방식을 개선하여 일상화 되도록 견인

2. 부산광역시 시범적용 사례

활용방안 및 업무 담당자 사업 이해도 제고

◆ 부산광역시에 시범적용을 위해 사전에 관련부서 업무담당자 특강 및 개별 면담을 통해 빅데이터와 사업에 대한 이해도, 분석 결과 활용방안을 제고하였다.

- ‘빅데이터 이해 및 도시계획 분야 활용방안’ 특강 실시
 - 일시·장소 : 2017.10.13.(금) 14:00~16:00, 시청 회의실(12층)
 - 참 석 자 : 도시계획실 업무담당자 45명
 - 강 사 : 국토연구원 김동한(박사)



- ‘분석 결과 활용방안 및 사업 활성화’ 업무담당자 개별 면담
 - 일시·장소 : 2017.10.20.(금) 10:00~16:00, 시청 회의실(3층)
 - 참 석 자 : 도시재생과 등 6개부서 9명, 국토연구원
 - 주요내용
 - ✓ 도시계획 및 도시재생
 - ✓ 노인 복지
 - ✓ 청년 주택 등

사 인프라 구축 단계

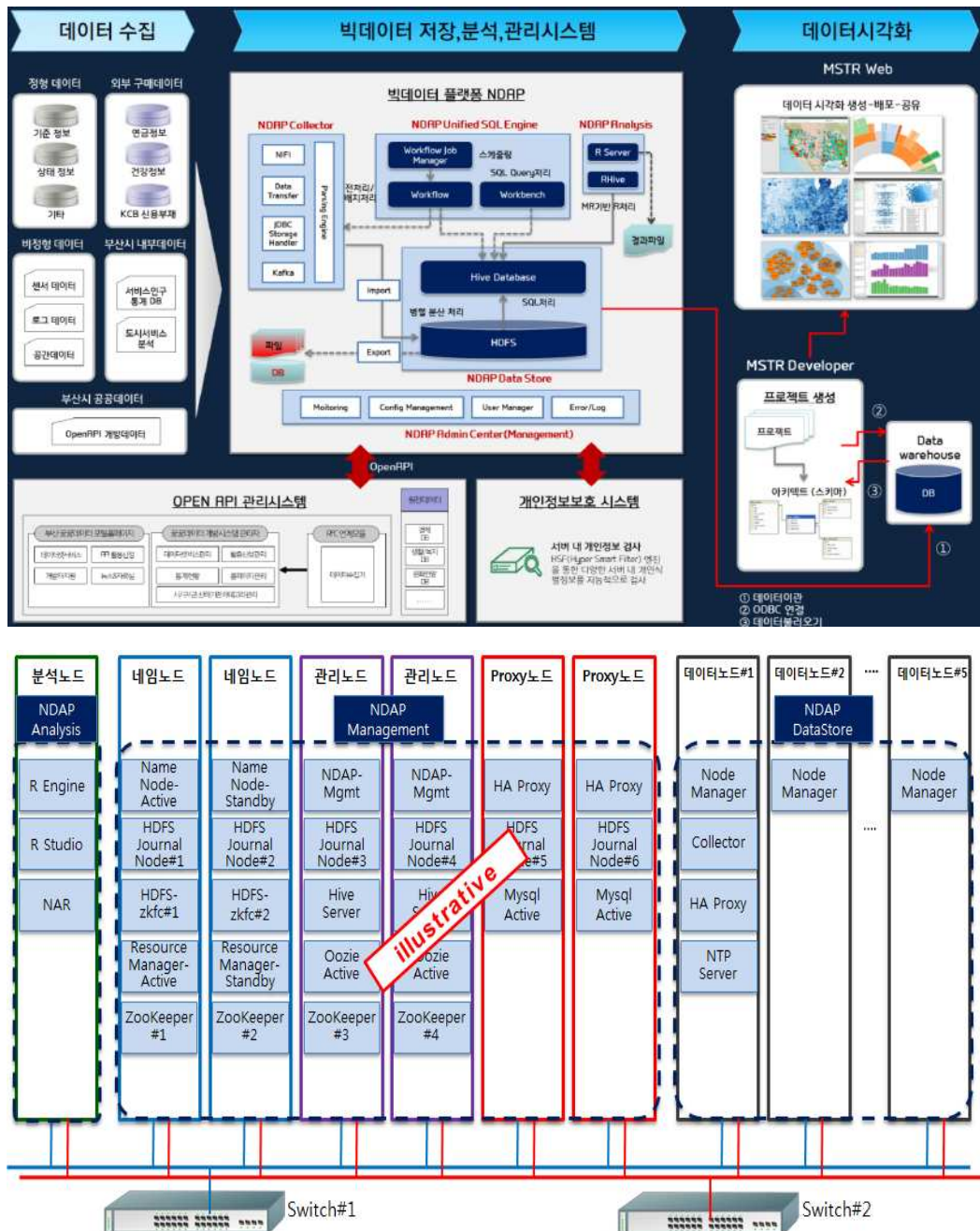
1. 인프라 구축 단계 요약

단계	주요 내용
시스템 요건 정의	<ul style="list-style-type: none"> ◆ 인프라 구축을 위한 정의 단계, 수집되는 데이터의 양, 분석 및 처리해야 하는 양, 저장해야 하는 양 등을 고려하여 서버 및 스토리지 사양을 결정 ◆ 분석사업의 특징을 고려하여 도입해야 할 소프트웨어를 선정 (메인 분석도구, GIS도구, BI솔루션, 실시간분석 도구, 리포팅 툴 등) ◆ 직접 서버를 구매할지, 클라우드 형태로 구현할지 장·단점을 파악할 필요가 있음
DB 구조 정의	<ul style="list-style-type: none"> ◆ 수집되는 데이터에 대하여 테이블 수 등 실제 분석 시행을 위한 DB 구조를 정의
시스템 반영	<ul style="list-style-type: none"> ◆ 실제 인프라를 구성하여 분석 알고리즘과 데이터를 입력하여 적용하는 단계
검증/검수	<ul style="list-style-type: none"> ◆ 데이터 분석이 원하는 결과값이 도출 될 때까지 알고리즘 등을 보정하는 과정, 데이터가 제대로 연계되어 입력되는지 확인 ◆ 데이터 정합성을 확신할 경우 검수
오픈	<ul style="list-style-type: none"> ◆ 분석 시스템 오픈
업데이트	<ul style="list-style-type: none"> ◆ 유지관리 사업 진행 후 필요 시 서버 용량, 분석 도구 구매 등

2. 개방형 빅데이터 플랫폼 인프라 구축

<p>시스템 구축 필요성</p>	<ul style="list-style-type: none"> ◆ 부산광역시는 과학적·지능적인 분석·예측 정보제공을 통한 선제적 시정 현안 대응과 정책 개발 등 신뢰성 있는 데이터 중심 스마트 시정 구현 ◆ 4차 산업혁명의 데이터 경제 중요성 및 빅데이터 시대를 맞아 부산시정·행정자료의 공유를 통한 수요자(시민·기업) 중심의 공공 빅데이터 서비스가 필요성이 대두 ◆ 사회현안 해결의 실증적 정책지원과 과학지능적 시정 구현을 위한 공공 빅데이터 활용과 통합 데이터기반 서비스 및 부서별 빅데이터 분석과제 추진으로 인한 중복투자(HW, SW)를 방지 ◆ 누구나 손쉽게 신규 과제 발굴·분석 및 결과 공유가 가능한 개방형 통합 빅데이터 플랫폼을 구축 ◆ 2017년도에 구축된 이 시스템을 통해 '금융 빅데이터와 공간 빅데이터를 활용한 도시양극화 분석 및 시뮬레이션' 분석사업을 연계 구축
<p>추진 목표</p>	<ul style="list-style-type: none"> ◆ 빅데이터의 효과적 활용을 위한 활용체계 구축과 분석결과(시각화) 공유 ◆ 효율적 정책발굴과 시민 공유·공감을 위한 개방형 플랫폼 구축 ◆ 예산 중복투자 방지를 위한 공동 활용 가능한 빅데이터 플랫폼 개발 ◆ 공공 빅데이터의 통합 활용과 개방 서비스를 위한 인프라 구축
<p>추진 전략</p>	<ul style="list-style-type: none"> ◆ 다양한 데이터 수집을 위한 빅데이터 저장·분석 플랫폼 구축 ◆ 부산사·외부 데이터를 저장·분석할 수 있는 빅데이터 플랫폼을 구축하고 Scaleout 구조로 확장성과 비용 효율성 및 가용성을 겸비한 Hadoop 기반의 저장·처리 시스템 기반 마련 ◆ 연계 분석을 위한 다양한 데이터의 수집 및 적재 <ul style="list-style-type: none"> • 부산사·외부데이터(국민연금정보·의료정보·신용/부채정보·부산공공데이터포털 주요 오픈API·부산도시서비스포털 등)를 저장하고 연계·분석 할 수 있는 데이터 셋 확보 기반 마련 ◆ 반응형 공공 빅데이터 포털시스템 구축(목표 시스템) <ul style="list-style-type: none"> • 부산시 빅데이터 포털시스템을 구축하고, 공공데이터 및 빅데이터를 하나의 시스템에서 사용자가 쉽게 접근하여 이용할 수 있는 기반 마련 • 누구나 언제 어디서나 이용 가능한 공공 개방형 빅데이터 플랫폼 구축

<참고> 개방형 빅데이터 구성도 및 흐름도



시스템
구축
개요

◆ HW-SW 물품 구입

- ① 구입물품 : 서버 15대, 개인정보 필터링 1조
- ② 사업기간 : 2017.4. ~ 5.
- ③ 사업금액 : 118백만원(전액 시비)
- ④ 계약방법 : 조달청 제3자 단가계약

◆ SW 개발

- ① 사업명 : 부산광역시 개방형 빅데이터 플랫폼 구축
- ② 사업기간 : 2017.5. ~ 10.(150일)
- ③ 사업금액 : 436백만원(전액 시비)
- ④ 계약방법 : 경쟁입찰(조달청 협상에 의한 계약)
- ⑤ 주요내용

부산시 빅데이터 반응형홈페이지 개발 (<http://bigdata.busan.go.kr>)

◆ 외부 빅데이터(신용·부채) 구입 및 분석결과 시각화 106종

- 외부자료 : 국민연금(연령별 취업 업종, 월평균 급여 등) 13종, 건강보험심사평가원(과목별 진료비 현황, 진료비 총액) 6종, 국민건강보험공단(건강보험 급여현황, 비만지수 등) 23종

※ 외부자료는 기관별 업무 협조로 데이터 수집(비용 없음)

- 구입자료 : KCB 신용·부채데이터(2014년 ~ 2018년, 지역별(구군별·읍면동별) 평균연소득, 가구소득수준 등) 28종
- 내부자료 : 부산시 공공데이터 오픈API형태(실시간 도로 정체구간, 대기질 정보, CCTV 정보 등) 17종, 도시서비스 분석(SKT이통사 유동인구의 지역에 따른 성별, 연령별, 시간대별 등, 현대카드 신용결제) 12종, 보건환경연구원(악취정보, 수질환경, 토양환경 등) 7종

◆ 기존 '부산공공데이터포털(<http://data.busan.go.kr>), 부산도시서비스포털(<http://sgis.busan.go.kr>)와 '개방형 빅데이터 플랫폼' 자료 이관 연계 및 시각화

- 부산공공데이터포털 : 오픈API 형태의 빅데이터 분석 가능 항목
- 부산도시서비스포털 : SKT통신사 유동인구, 현대카드 매출정보

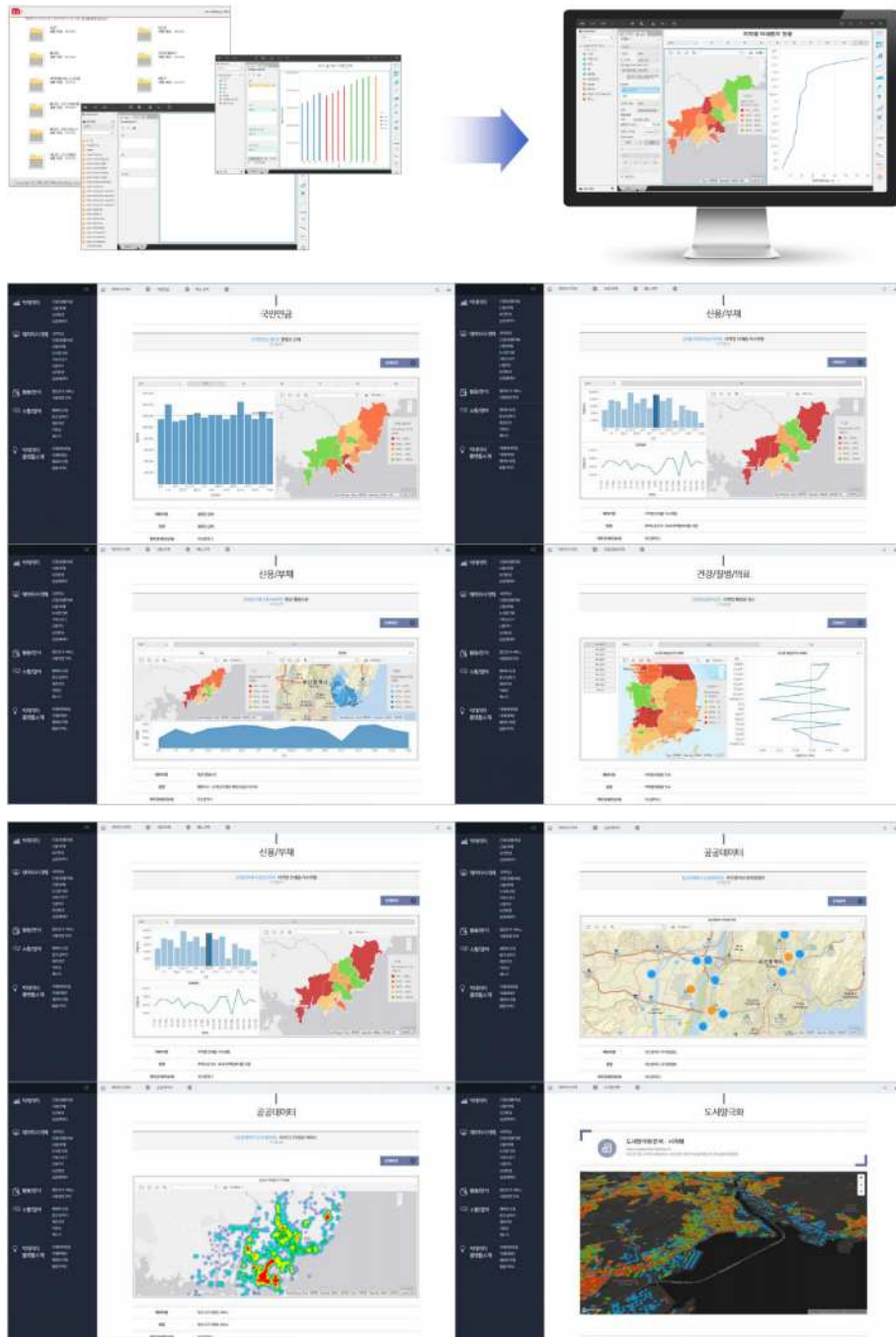
◆ 데이터 증가 및 기능 추가 등 확장이 용이한 기존(위 가)항목 인프라 증설 및 보강

	<ul style="list-style-type: none"> • SW : 빅데이터 플랫폼 솔루션(NDAP) 13조, 전자정부표준프레임워크(RFC, 빅데이터 홈페이지 관리) 1조, 공공데이터 API관리(RFC, 공공데이터 개방·연계) 1조, 개인정보보호(웹필터) 1조, BI(MicroStrategy, 데이터 시각화) 1조 • HW : 서버 16대(관리 노드 2대, 네임 노드 2대, 데이터 노드 6대, 분석 노드 1대, 프락시 노드 2대, BI 노드 1대, WAS 1대, 개인정보관리 1대), L3스위치 2대, KVM 1개, 시스템 랙 1개 <p>※ 홈페이지 데이터는 부산시 공공데이터시스템의 DB 공동 활용</p> <ul style="list-style-type: none"> ◆ 사업 효율성·전문성 강화를 위한 자문회의(교수·연구원·전문가 5명) 구성 운영 4회 ◆ 빅데이터 시범과제 실무T/F팀(시청 일자리창출과, 건강증진과, 경제기획과, 4개구 등 13명) 운영 4회 ◆ 빅데이터 플랫폼 활용 촉진 등 이용활성화를 위한 저명인사 교육 특강 2회, 설명회 개최 1회, 활용 교육 5일
기타 시스템 연관사업 추진	<ul style="list-style-type: none"> ◆ 도시 양극화 분석사업을 포함하여 해운대 IOT 실증단지 4대분야 30개 서비스 구축인 '글로벌 스마트시티 실증단지 조성' 사업과 2017년 사업을 추진 중인 'GIS기반 부산시 정책지도(다복동, 소상공인, 클린에너지 분야 등) 구축' ◆ '관광객(7개 언어 : 국문, 영문, 중문, 일문, 러시아어, 베트남어, 태국어) 이동경로(인원수, 체류시간 등) 빅데이터 분석' 사업이 개방형 빅데이터 플랫폼과 연계하여 2018년에 완료 될 예정
시스템 분석, 활용 예시	<ul style="list-style-type: none"> ◆ 관광지 무료 와이파이 설치 <ul style="list-style-type: none"> • 부산시 공공데이터인 무료 와이파이 설치현황에 따르면, 해운대 해수욕장은 12개 이상 설치 • (반면) 광안리 해수욕장과 태종대에는 1개소씩 설치되어 있어, 추가 설치가 필요 ◆ 비만과 소득수준 관계 <ul style="list-style-type: none"> • 건강보험공단과 KCB 신용·부채의 부산시 비만과 소득수준 관계를 빅데이터로 분석해 보면, 대체적으로 비만율과 소득수준은 반비례하지만 지역에 따라 연령별 비만율 보정이 필요하고, 기장군의 경우 비만율과 소득수준이 같이 높고 젊은층의 비만율이 타 구에 비해 높음

<참고> 부산시 개방형 빅데이터 홈페이지 스펙

물품명	수량	규격(서버 1대별)	물품명	수량	규격(서버 1대별)
관리 노드	2	<ul style="list-style-type: none"> · CPU : 6core 2.4Ghz×1 · 메모리 : 64GB · 디스크 : 0.6TB · NC : 10GB 2포트×1 · OS : CentOS 6.8×1 	BI 노드	1	<ul style="list-style-type: none"> · CPU : 6core 2.4Ghz×1 · 메모리 : 64GB · 디스크 : 1.8TB · NC : 10GB 2포트×1 · OS : CentOS 6.8×1
네임 노드	2	<ul style="list-style-type: none"> · CPU : 6core 2.4Ghz×1 · 메모리 : 128GB · 디스크 : 0.6TB · NC : 10GB 2포트×1 · OS : CentOS 6.8×1 	WAS	1	<ul style="list-style-type: none"> · CPU : 6core 2.4Ghz×1 · 메모리 : 16GB · 디스크 : 0.6TB · NC : 10GB 2포트×1 · OS : CentOS 6.8×1
데이터 노드	6	<ul style="list-style-type: none"> · CPU : 6core 2.4Ghz×1 · 메모리 : 64GB · 디스크 : 24.6TB · NC : 10GB 2포트×1 · OS : CentOS 6.8×1 	개인 정보 관리	1	<ul style="list-style-type: none"> · CPU : 6core 2.4Ghz×1 · 메모리 : 64GB · 디스크 : 1.2TB · NC : 10GB 2포트×1 · OS : CentOS 6.8×1
분석 노드	1	<ul style="list-style-type: none"> · CPU : 6core 2.4Ghz×1 · 메모리 : 512GB · 디스크 : 0.6TB · FIO :1600GB×2 · NC : 10GB 2포트×1 · OS : CentOS 6.8×1 	L3	2	<ul style="list-style-type: none"> · NC : 10GB 48port
프락시 노드	2	<ul style="list-style-type: none"> · CPU : 6core 2.4Ghz×1 · 메모리 : 64GB · 디스크 : 1.8TB · NC : 10GB 2포트×1 · OS : CentOS 6.8×1 	KVM	1	<ul style="list-style-type: none"> · 16port

<참고> 부산시 빅데이터 홈페이지 분석결과 예시



<참고> 기타 시스템 연관사업 추진



빅데이터 이해와 활용을 위한 전문가 특강
일시 : 2017. 8. 18.(금) 14:00
장소 : 대저리대청
주최 : 부산광역시

금융·공간 빅데이터를 활용한 도시양극화 분석

- 도시양극화 분석 및 시뮬레이션 결과에 대한 시각화
- 분석, 시뮬레이션 결과를 활용하기 위한 부산시 빅데이터 플랫폼 연계

※ NIA 2017년 빅데이터 플래그십 사업, 국토연구원 수행

글로벌 스마트시티 실증단지 조성

- 해운대 IoT 실증단지 4대분야 30개 서비스 구축
- 스마트시티플랫폼 일부 서비스를 부산시 빅데이터 플랫폼과 연계

※ 과학기술정보통신부, 부산시+SKT텔레콤 컨소시엄 수행

GIS기반 부산시 정책지도 구축

- 다북동, 소상공인 지원사업, 클린에너지 정책분야 빅데이터 분석 및 GIS기반 정책지도 제작

관광객 이동경로 빅데이터 분석

- 부산을 방문하는 관광객의 주요 이동경로, 인원수, 체류시간 분석

부산시 빅데이터 플랫폼

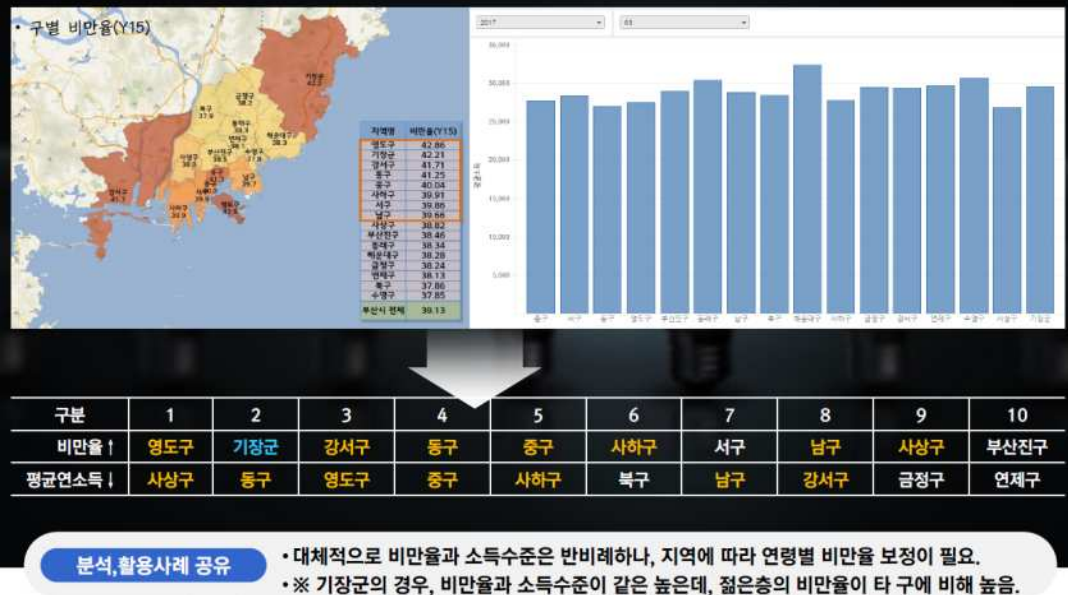
고 < 직업, 소득, 부채수준, 소비수준 > 저

<참고> 활용예시

관광지를 대상으로 무료와이파이 설치시, 최적지는?



비만과 소득수준 관계



CHAPTER

II

지방자치단체 빅데이터 분석 사례집

경상남도 응급환자 골든타임 확보를 위한 빅데이터 분석 사례



korea Local Information Research & Development Institute

가. 추진배경	78
나. 프로젝트 개요	79
다. 적용 프로세스	82
라. 기획·관리 단계	84
마. 데이터 분석 단계	89
바. 행정적용	99
사. 인프라 구축 단계	102

II. 경상남도 응급환자 골든타임 확보를 위한 빅데이터 분석 사례

요 약

경상남도는 신속하고 최적화된 구급서비스 체계 구축 방안을 도출한 빅데이터 분석 결과를 구조구급 정책에 활용한다.

응급환자 골든타임 확보를 위한 빅데이터 분석에서 2010년부터 2016년 7월까지의 구조구급시스템 데이터 약 263만 건을 비롯해 총 400만 건 이상의 데이터 분석을 통해 응급출동 현황, 응급환자 발생 유형, 3대 응급환자(심정지, 뇌혈관질환, 중증외상) 골든타임 확보율, 도내 시·군별 3대 응급환자 고위험군, 3대 응급환자 골든타임 확보율 제고를 위한 구급차 재배치 및 신규 배치 시 우선순위 등에 대한 결과를 도출했다.

응급출동 현황은 소방서, 안전센터 및 시·군별로 구분하여 분석했다. 응급환자 발생 유형은 구급활동일지를 기준으로 질병과 질병외로 구분하여 연도별·월별·시군별로 주로 발생하는 응급환자 유형을 분석했으며, 연도별·지역별 3대 응급환자 분포 및 도내 3대 응급환자 고위험군을 분석했다. 또한 연도별·지역별 응급출동 소요시간 및 3대 응급환자별 골든타임 확보 비율과 지역별 골든타임 확보 애로요인 등을 분석하였다.

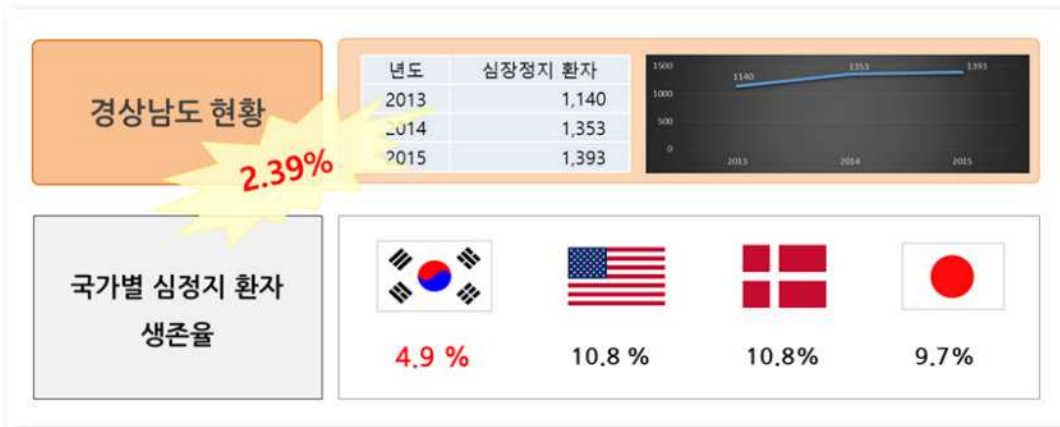
빅데이터 분석 결과는 도내 3대 응급환자 발생 고위험군 및 지역, 오인신고 많은 지역에 예방 홍보강화 및 효과적인 대응 방안(예: 도민 대상 응급상황 대비 수시 교육 실시 등) 마련 등에 활용하여 소중한 생명을 살릴 수 있는 기회를 놓치지 않도록 할 예정이다.

가

추진배경

- 2014년 “투명한 정부”, “유능한 정부”, “서비스 정부”를 지향하는 정부 3.0 추진 계획의 주요 추진 전략의 하나로 “빅데이터를 활용한 과학적 행정 구현” 이 선정된 이후 다양한 시범 사업과 이를 통한 확산 사업이 전국적으로 확산 적용 되어 경남에서도 이에 대한 적극적인 수용을 위해 관련 기관 협의 지속
- 경상남도에서는 도민의 안전과 지역 경제 활성화를 위한 정책 발굴의 과학적 근거 자료를 제시하여 향후 지속적인 과학 행정의 기반을 구축하기 위하여 만관이 보유한 다양한 데이터를 활용하여 대 도민 서비스 창출을 목적으로 2016년 빅데이터 시범 사업을 통해 과학 행정을 위한 첫 발걸음을 내 딛고자 사업 기획
- 심장정지 환자에게 골든타임(4분 이내)은 생사를 결정짓는 중요한 시안으로 빠른 응급 처치가 필수적으로 최적화된 구급서비스체계가 필요하나 경남은 연간 약 1,300명 이상 심장정지 환자 발생대비, 생존율은 2.39% 정도에 불과한 상태 (경남소방본부 통계)
 - ※ '13년도 1,140명 → '14년도 1,353명 → '15년도 1,393명
 - ※ 심장정지 환자 생존율 : 미국 10.8%, 덴마크 10.8%, 일본 9.7%, 한국 4.9%
- 심장정지 환자 등 응급환자의 생사를 결정짓는 골든타임 확보를 위해 필요 자원 (구급차, 응급 처치 물품 등) 및 응급 구조사의 신속한 출동을 위한 최적화된 구급서비스 체계 구축 방안 도출이 시급한 상황
- 과거 응급 환자 발생 빈도와 연령, 병력 등의 정보를 통합 분석하여 위험 지역 및 대상군을 분류하고 출동 시간 및 병원 이송 시간 감소를 위한 최적 구급기지 및 자원 배치 방안 제시가 필요

경상남도 및 국가별 심장정지 환자 생존율 [출처: 경상남도 소방본부]

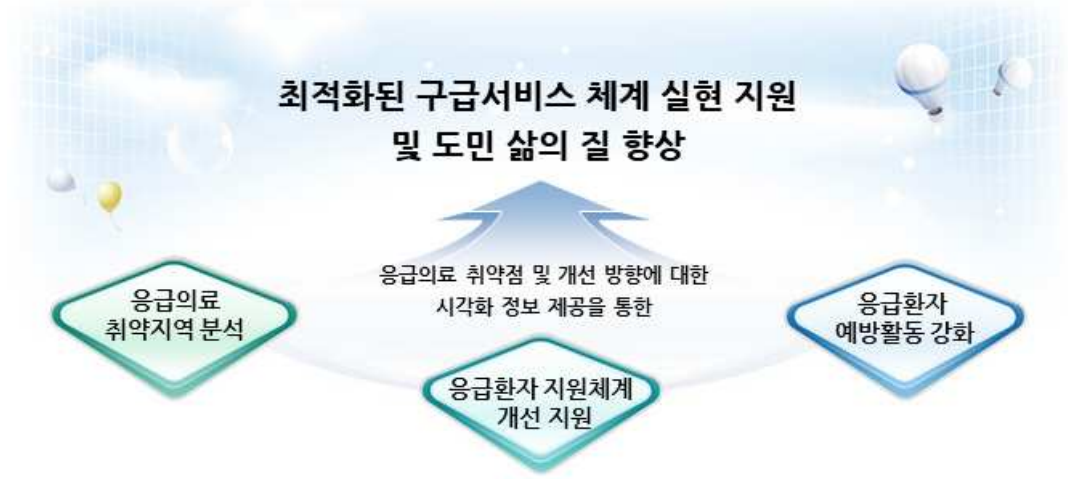


4

프로젝트 개요

- 본 프로젝트는 과거 응급 환자 발생 빈도와 연령, 병력 등 의 다양한 정보를 빅데이터 분석 기술을 활용하여 통합 분석, 위험 지역 및 대상군을 분류하고 출동 시간 및 병원 이송 시간 감소를 위한 최적 구급기지 및 자원, 심장정지 환자 등 응급환자의 생사를 결정짓는 골든타임 확보를 위해 필요 자원(구급차, 응급 처치 물품 등) 및 응급 구조사의 신속한 출동을 위한 최적화된 구급서비스 체계 구축 방안을 도출하고자 하는 것으로 아래의 분석 결과를 사업 목표로 설정
- 응급환자 구급 활동 빈도와 위치정보, 응급실 도착시간과 이동거리 등을 분석하여 최적화된 응급구조 모델 구현
- 정확한 구급수요 조사를 통한 수요가 많은 지역에 구급차 배치
- 농촌지역 등 의료취약지역 구급서비스 구축 방안 마련
- 구급활동에 필요한 각종 정보를 분석하여 최적화된 응급구조 모델을 구현함으로써 고품질 구급의료 서비스 제공

경상남도 응급환자 골든타임 확보를 위한 빅데이터 분석 목표



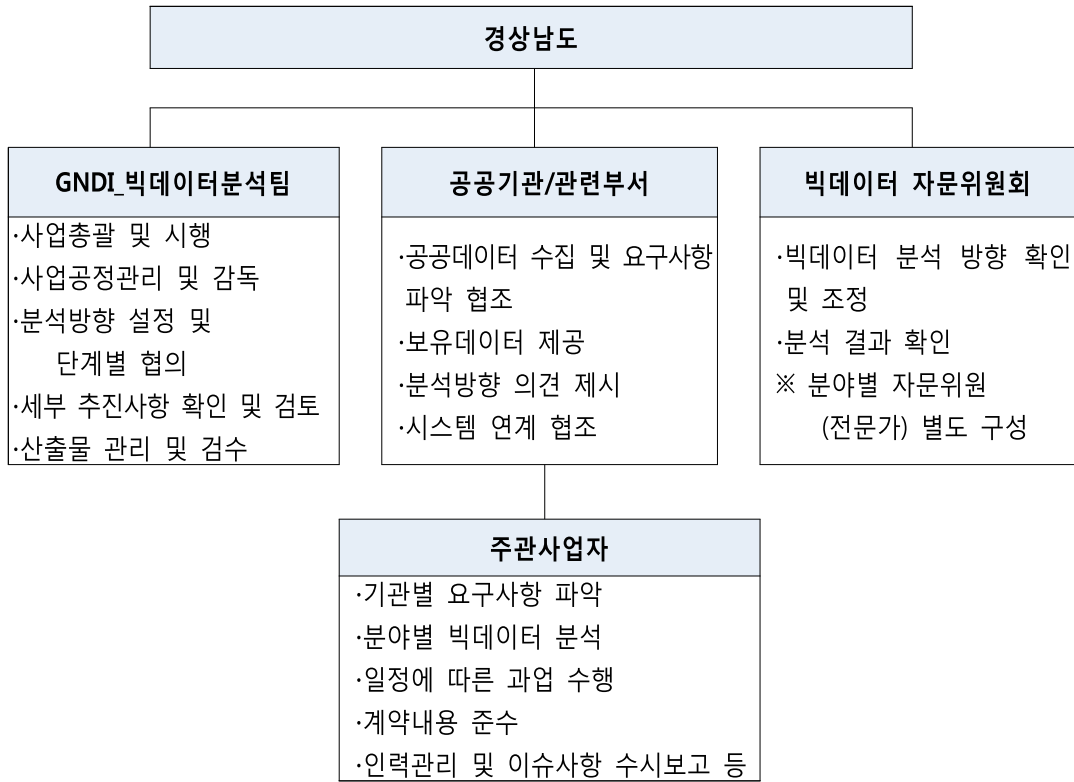
■ 기술적 측면

- 정형/비정형 빅데이터의 분석 처리 속도 향상을 위한 시스템 구축
- 글로벌 표준 기반의 오픈소스 활용
- 프로토타입 개발 후 반복 테스트를 통한 안정적 서비스 환경 구축
- 향후 관리의 용이성 및 연계의 지속성을 위한 연계기술의 표준화

■ 관리적 측면

- 경상남도 소방본부 정책 방향에 부합하는 실증 결과 도출
- 진행상황에 대한 점검 및 결과에 대한 전문가 의견 수렴을 거쳐 방향을 설정하고, 최적의 방안이 마련되도록 추진상황을 점검
- 효과적 데이터 수집 및 방향 설정을 위해 사업 관련 기관과의 협업체계를 구성하여 추진

응급환자 골든타임 확보를 위한 빅데이터 분석 사업 추진 체계

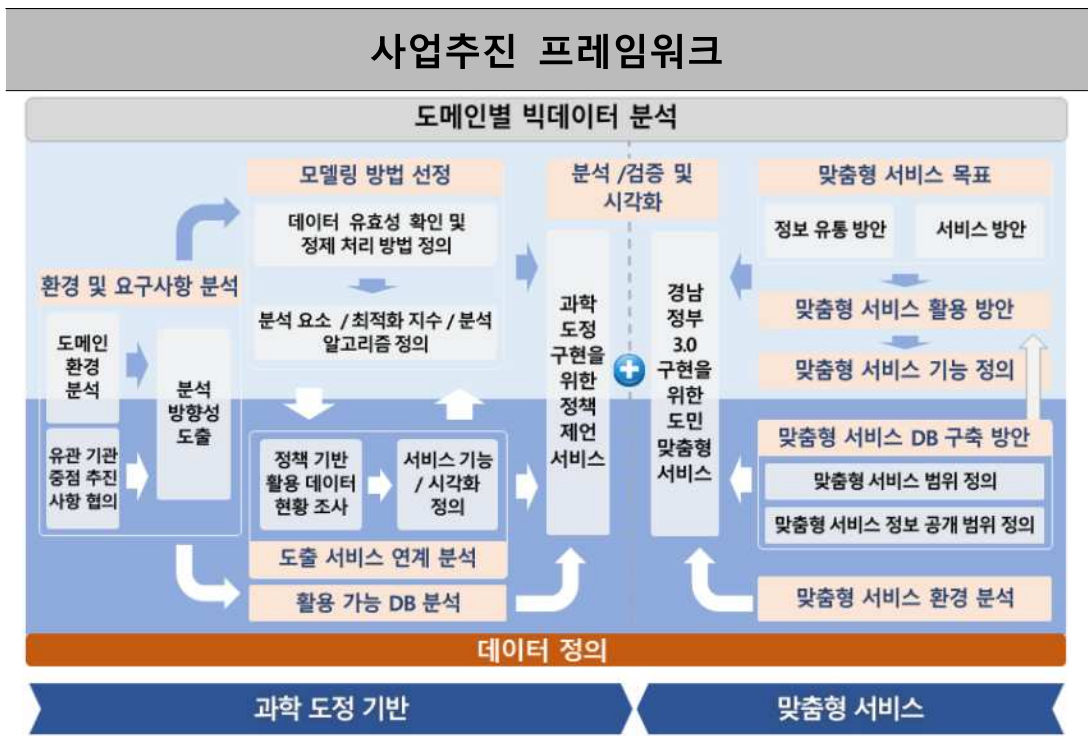


■ 경상남도 응급환자 골든타임 확보를 위한 빅데이터 분석 사업 추진 일정

구 분	추진일정						
	M	M + 1	M + 2	M + 3	M + 4	M + 5	
○ 현황 및 요구사항 분석	■				■		
○ 데이터 수집 및 연계		■	■	■			
○ 데이터 모델링			■	■	■		
○ 분석 알고리즘 개발 및 적용				■	■	■	
○ 결과 도출(시각화)				■	■	■	
○ 사용자테스트 및 안정화						■	■
○ 착수 및 중간 보고회	▲				▲		
○ 사용자 설명회							▲
○ 사업종료 및 결과보고회							▲

다 적용 프로세스

- 성공적 사업 수행을 위해, 업무 담당자 및 의사결정자의 적극적 사업 참여, 전문가 참여를 통한 충분한 사업내용 검토, 원활한 협업수행을 위한 환경 제공, 유관 기관 및 업체들과의 긴밀한 협조체계 구축 등이 전제되어야 하고 이를 효과적으로 수행하기 위하여 사업 추진 프레임워크를 정의하고 다음의 기준에 따라 수행
 - 프로젝트 수행 절차 표준화 및 합리화
 - 시스템의 효율적인 적용
 - 시스템 품질향상을 위하여 단계별 적용프로세스를 정의하여 각 단계별 산출물 점검을 통한 사업 목표 도달 가능성을 점검한 후 다음 단계로 이행



- 프레임워크의 단계를 세분화한 후 단계별 수행 태스크와 세부 태스크를 정의하여 체계적인 사업 수행을 담보하고 품질 관리의 용이성을 도모

● 사업추진 프레임워크 세부 단계 및 수행 태스크

단계명	수행 Task	세부수행 Task
환경 및 현황 분석	현황 및 요구사항 분석	◆ 현황 및 문제점 분석
	환경 분석 및 방향성 도출	◆ 필요데이터 정의 및 현황 분석
		◆ 필요데이터 수집 방안 협의
서비스 기획 및 설계	서비스 기획	◆ 목표시스템 도출
		◆ 서비스 상세 기능 정의
	DB 설계 및 데이터 확보	◆ 서비스 UI/UX 기획
		◆ DB 설계
		◆ 과거데이터 확보
	데이터 분석	◆ 실시간데이터 수집
		◆ 응급환자 골든타임관련 데이터 유형화 분석
서비스 개발	알고리즘 설계	◆ 응급환자 골든타임 알고리즘 설계
	시스템 개발	◆ 데이터 통합 수집 시스템 구축
		◆ 수집 데이터 연계 및 융합 DB 구축
		◆ 분석 알고리즘 개발
		◆ 서버 모듈 개발(ETL)
		◆ front 개발(CSS) 및 서버 연동(JS)
	시스템 시각화 구현	◆ 응급환자 골든타임 확보 페이지 개발
		◆ 응급환자 골든타임 확보 시각화 개발
	디자인 개발	◆ 디자인 컨셉 도출
		◆ 디자인 개발
	검증 및 테스트	◆ 알고리즘 적용 결과 테스트
		◆ front 기능 테스트
		◆ 통합 테스트
		◆ 시스템 안정화
교육 및 기술이전	서비스 활용 교육	◆ 사용자 매뉴얼 제작
		◆ 사용자 설명회
	데이터 및 서버 이전	◆ 모듈 및 소스 이관
		◆ DB 이관

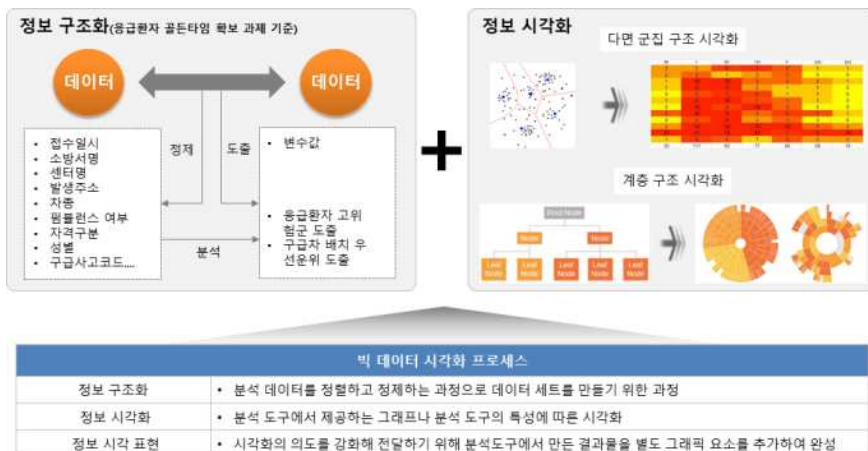
라 기획 · 관리 단계

- 응급환자 골든타임 확보를 위한 빅데이터 분석 사업의 기획 의도와 사업자 제안 사항을 기반으로 목표시스템 개요를 정의하여 사업 수행의 기준으로 설정한 후 사업 수행 프레임워크에 따라 주요 단계별 수행 절차를 정의

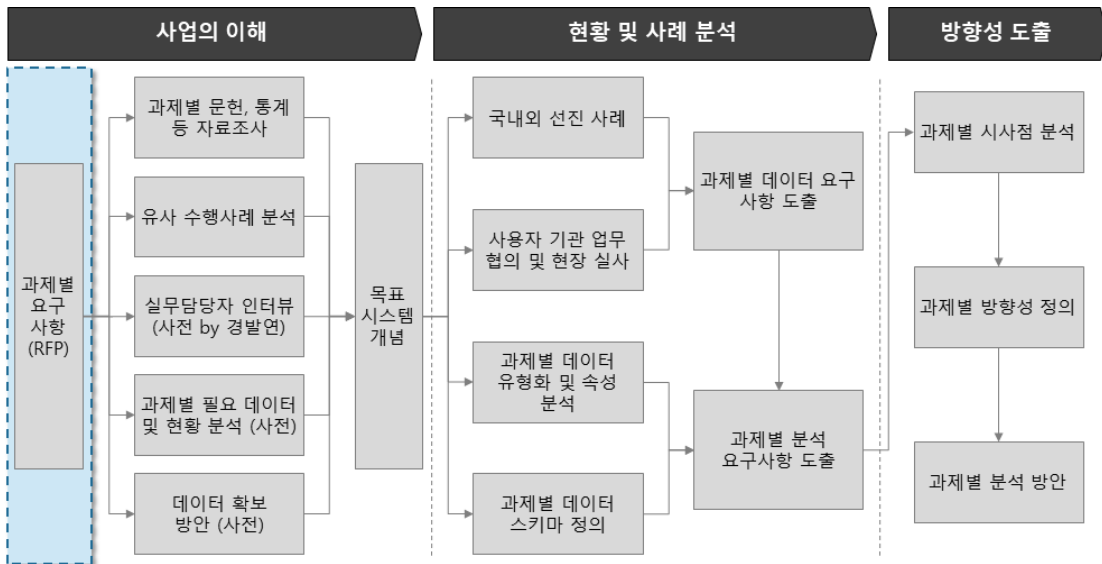
● 목표시스템 핵심 요소 정의

시스템 핵심 요소 (데이터, 인프라, 서비스)에 대한 정의 및 검토		
데이터	<ul style="list-style-type: none"> 기초 지역 정보: 지역 인구수, 성비, 연령층, 밀도, 지역 소득 분포, 거주 형태 등 도메인 정보: 지역 도메인 현황, 통계 정보, 일반 도메인 현황, 통계 정보 지표 정의 정보: 응급출동정보, 3대응급질환정보, 지역건강정보, 국가응급진료정보, 기상 정보 	응급환자 분석 플랫폼
인프라	<ul style="list-style-type: none"> 매쉬업 플랫폼: 정의된 외부 데이터의 수집 및 빅데이터 분석을 위한 저장관리 체계 분석플랫폼: 서비스 제공을 위하여 매쉬업 플랫폼의 데이터를 다양한 빅데이터 분석 방법을 사용하여 분석요소에 대한 영향력/연관/통계분석 등을 수행하는 시스템 체계 개방형: 분석플랫폼에 의해 도출된 결과를 사용자가 이해하기 쉽고 활용 가능한 형태로 시각화 하고 데이터나 분석 결과 또는 시각화 결과를 외부 시스템과 연동 할 수 있도록 지원하는 시스템 체계 	
서비스	<p>빅데이터 분석저장 시스템에 저장된 다차원 데이터를 독자적인 데이터 분석 모델을 통해 인사이트 제공</p> <ul style="list-style-type: none"> 과학 도정 구현을 위한 정책 제안 서비스 <ul style="list-style-type: none"> 빅데이터 분석을 통한 도정 현안 사항 해결의 신뢰성 확보 빅데이터 활용 활성화를 위한 지자체 분석 우수모델 제시 및 확산 발생할 수 있는 갈등 사항에 대한 선제적 대응으로 민원 발생 건수 및 민원 지속 감소 대 도민 대상 온라인 서비스 제공 <ul style="list-style-type: none"> 민관간 다양한 데이터 공유·활용을 통한 업무 생산성 향상 및 부가가치 창출 기회 도민 관련 정책 결정 과정 공개로 지방정부 신뢰도 향상 경상남도 공통 활용 기반을 통해 기초자치단체간 정보비대칭 문제를 해소 하고, 개방과 협력을 통한 지역별 대민 서비스의 상향 평준화 유도 	

● 목표시스템 서비스 요소 정의



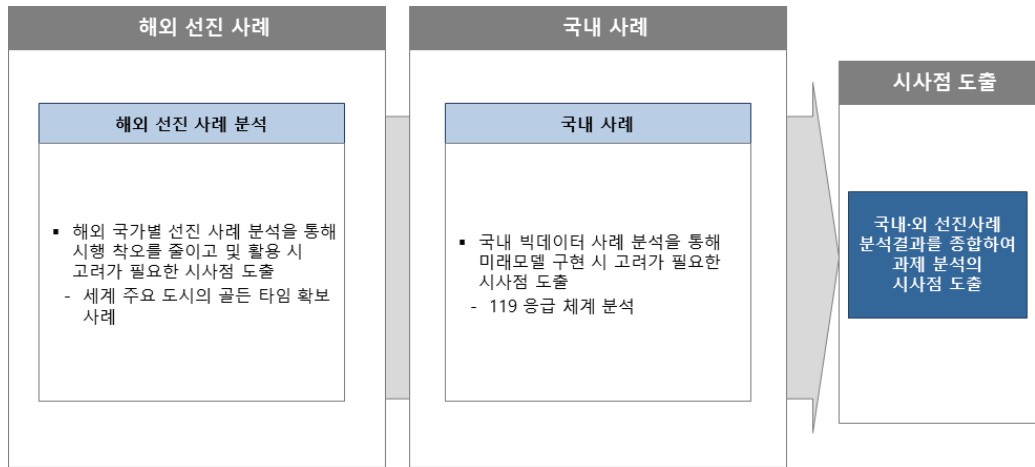
- 응급환자 골든타임 확보를 위한 빅데이터 분석 사업의 데이터 정의, 분석 방법 및 모델 정의, 결과 검증 및 대·도민 서비스 등의 방향성 도출을 위하여 방향성 도출을 위한 프레임워크를 경상남도과 사업자간의 협의를 통해 아래와 같이 정의하고 시행
- 방향성 도출을 위한 프레임워크에는 서비스 활용 부서의 현안 파악과 국내외 유사 사례를 분석하여 활용 가능한 데이터를 중심으로 도출 가능한 결과와 그 활용에 중점을 두고 시행
 - 목표시스템 방향성 도출을 위한 프레임워크



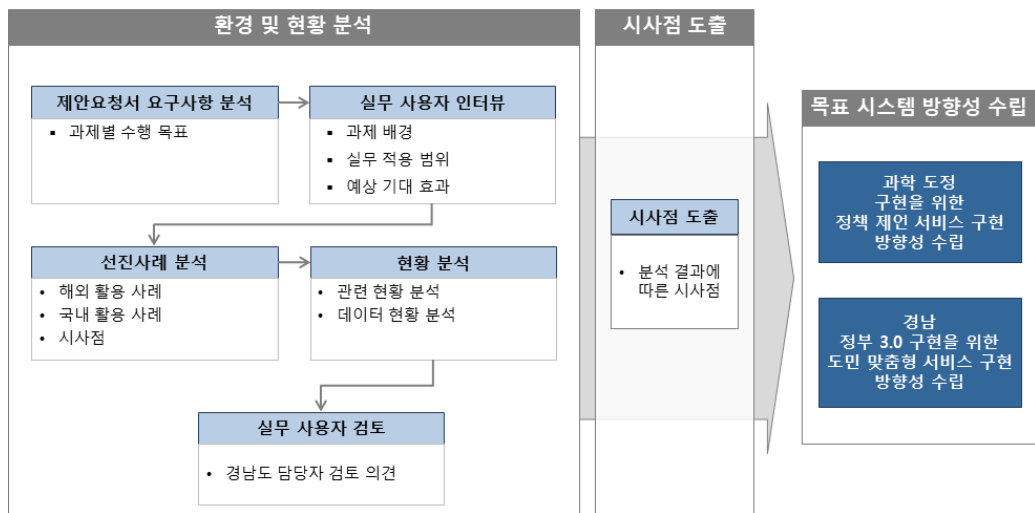
- 목표시스템 방향성 도출을 위한 프레임워크중 시사점 도출을 위한 “국내외 유사 사례 분석” 과 구체적인 방향성 도출을 위한 과정은 별도의 프레임워크를 통해 정량화 할 수 있는 결과를 도출하도록 하였음
- 방향성 도출을 위한 프레임워크는 사업추진 배경과 목적에 부합하도록 구성
 - 행복한 경남‘ 경남 비전 달성 지원
 - 데이터 기반의 과학적 분석 방법으로 현실성 있는 결과 도출

- 실무담당자가 사업 추진 및 정책 방향 설정에 활용할 수 있는 서비스 제공
- 도민이 편의성 등을 체감할 수 있는 서비스 제공
- 도민 체감형 공공 서비스의 질 개선

사례 분석 프레임워크



분석 시스템 구축 방향성 도출 프레임워크



- 정의한 프레임워크에 따라 실무 담당자 및 관련 기관 관계자의 면담을 통해 빅데이터 시스템의 분석 과제와 모델 구성에 필요한 요구사항들을 정의하였음

- 업무 관계자 면담 내용

인터뷰 내용 및 요구 사항		주요 Needs
인터뷰 및 실무 협의	경남 소방본부 창원 소방본부 경남응급의료지원센터	<ul style="list-style-type: none"> ○ 응급 환자의 생존율을 향상 시킬 수 있는 골든타임 확보를 할 수 있도록 대책을 마련할 수 있는 근거 필요 ○ 새로운 구급대 신설 등은 비용은 물론 시간이 필요하니 중, 장기 대책임 ○ 현 상황에서 구급차의 동적 배치, 장비 추가 배정 등 단기적으로 해결할 수 있는 방안 필요 ○ 예방 활동을 위한 홍보 자료 (지역별, 시간대별, 일별 위험 지역 등) 필요
요청 사항	<ul style="list-style-type: none"> • ‘심정지 환자의 생존율’ 확보가 구조구급 분야에서 전국적으로 이슈가 되고 있으며, 골든타임 확보를 위한 단순 물리적 환경 확장 이외의 방안이 시급한 실정으로 빅데이터 분석을 통해 방안이 나왔으면 함 • 센터 추가 조성 등 비용이 소요되는 해결책 보다는 기존의 자원을 최적화 하여 해결 가능한 방안을 먼저 고려 	
활용 방안	<ul style="list-style-type: none"> • 골든타임 확보를 위한 정책 자료 <ul style="list-style-type: none"> ✓ 경남 18개 시, 군별 골든타임 확보 현황 및 저해 요인 ✓ 서비스 수준 표준화를 위한 주요 장비 수급 계획 • 응급환자 예방을 위한 홍보 자료 <ul style="list-style-type: none"> ✓ 응급 질환별 주요 발생 지역 ✓ 응급 환자 발생 주요 패턴 	

- 또한 현재 사용 중인 시스템 현황과 제공 가능한 정보 형태 및 면담 결과에 따라 분석에 필요한 데이터를 정리하고 수집 가능성 및 수집 방법 그리고 세부 절차에 대한 업무협조 방안과 일정 계획 수립

● 필요데이터 정의

데이터 내용	데이터 출처
구급 일지(일자, 출동시간, 이동시간, 발생지역, 처치내용, 사용 장비 등)	경남소방본부
안전센터 안전 장비 현황 (구급시설 당, 장비별 개수)	경남소방본부 창원소방본부
응급구급시설(구급 시설) 정보 (위치, 규모, 커버 지역, 장비 내역)	경남소방본부 창원소방본부
응급의료시설(병원) 정보 (위치, 병원 규모, 전문의 수, 병동 수 등)	경남소방본부
기상 데이터(지역별 기온, 날씨)	기상청
응급의료 관련 데이터 (발생 일시, 환자 연령, 환자 주소지, 발생 질환)	경남소방본부
응급환자 세부 상황데이터	경남소방본부
도민 건강 데이터	보건복지부
경남 3대 응급환자* 예후정보	국립중앙의료원
도민 인구 데이터	통계지리정보서비스

※ 3대 응급 질환 : 심정지, 중증외상, 뇌심혈관질환

■ 국내외 주요 사례에 대한 조사를 통해 사업 목적 및 사용 데이터 그리고 분석 방법의 유사성을 비교 분석하여 분석 모델 수립을 위한 기초 자료로 활용

해외 사례

제목	Time-dependent ambulance allocation considering data-driven empirically required coverage
기관	• Management, Operations Research - Ruhr-University Bochum, Germany
목적	• 정확하며 동적인 최적화된 EMS(Emergency Medical Services) 모델 생성 지원
대상	구급차 재배치가 필요한 지역
데이터 및 방법	<ul style="list-style-type: none"> • 데이터(독일 : 보훔) - 위치, 면적, 인구수, 출동 횟수, 응급자 수, 출동 내역 • 최적화 모델 생성 - 과거 사고 발생 데이터를 이용한 모든 지역을 출동할 수 있는 응급시설 위치와 동적인 응급자 위치 도출
효과	<ul style="list-style-type: none"> • EMS 도착시간 감소 • 동일한 자원의 효율적 사용
시사점	<ul style="list-style-type: none"> • 현 구급차의 동적인 배치를 통한 골든 타임 확보 • 구급차 장비 표준화로 동일한 서비스 제공 • 동일 조건하의 불확실성 제거 증명

1) 출처: Time-dependent ambulance allocation considering data-driven empirically required coverage

국내 사례

제목	구급차 배치 방안을 고려한 시나리오 기반 구급차 입지 최적화 알고리즘 연구(논문)
기관	한국과학기술원 산업 및 시스템 공학과
목적	응급의료체계 수요의 커버리지를 최대화하는 구급차 대기 위치 선정
대상	구급차 재배치가 필요한 지역
데이터 및 방법	<ul style="list-style-type: none"> • 데이터 - 2010년 1, 2월의 대전 지역 구급일지 자료 • 분석방법 - 특정 응급환자 발생 순서에 대하여 최단 거리 구급차 배치 방안을 기반으로 응급환자에게 구급차 할당 - 시간간격 변동성을 고려한 구급차 입지 최적화 모델
효과	<ul style="list-style-type: none"> • 기존의 구급차 입지 선정 모델과 비교하여 골든 타임 이내 도착 비율을 향상
시사점	<ul style="list-style-type: none"> • 응급 시나리오 기반의 구급차 최적 배치 위치 결정 이론 • 응급 수요를 최대한 수용하는 방안 도출

2) 출처: 대한산악공학회/공통학술대회 논문집, 2015.4 13-18 (6 pages)
http://www.dbpia.co.kr/journal/ArticleDetail/NODE06277951

기존 자원의 활용성 최대

응급 서비스 수준 표준화

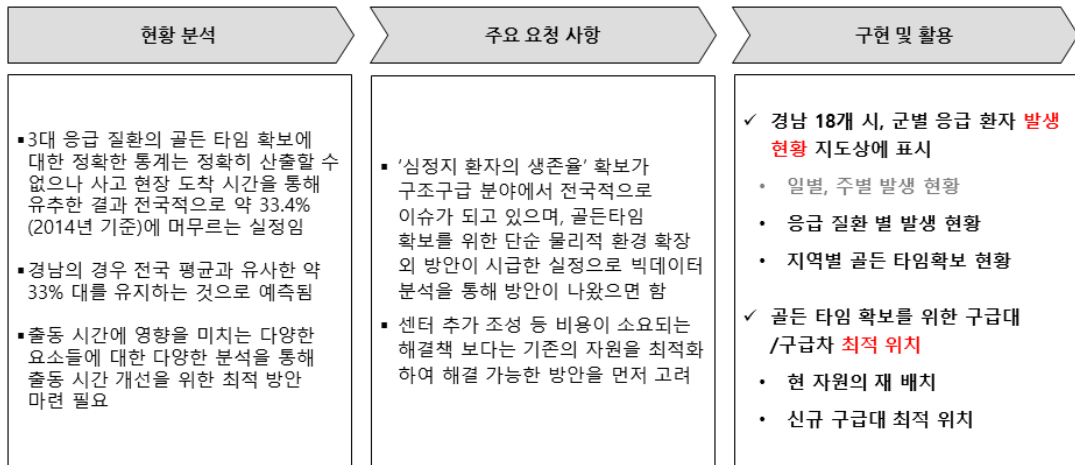
불확실성 도출 및 요인 추출

- “목표시스템 방향성 도출을 위한 프레임워크”를 통해 “응급환자 골든타임 확보를 위한 빅데이터 분석”의 구체적인 방향성을 확보하고 이를 바탕으로 수집한 데이터의 정렬과 정제, 분석 모델 정의 및 적용 알고리즘 결정 그리고 검증 방법에 대한 구체적인 방향성을 확보하였음

- 응급환자 골든타임 확보를 위한 빅데이터 분석 구축 방향성

구축 목적

경남 응급환자의 생존을 향상을 위하여 119 구조구급대의 골든타임 확보에 저해가 되는 요인을 분석하여 제시하고 이를 통해 가능한 다양한 해결책을 찾을 수 있는 과학적인 근거를 제시 한다. 또한, 지역별, 일별, 시간별 응급환자 발생 위험 지역을 분석하여 예방 활동에 활용하도록 함



마 데이터 분석 단계

- “응급환자 골든타임 확보를 위한 빅데이터 분석”을 위해 수집된 데이터는 주요 데이터 약 392만 건과 참고 데이터 약 300MB 로 시스템에서 추출한 CSV 파일과 API 를 이용한 XML 형태로 수집하여 Hadoop에 저장
- 분석 대상 데이터의 범위와 기준은 아래와 같음
 - 데이터 범위 : 2010.1월~2016.6월
 - 단, 3대 응급환자 예후데이터 2013, 2014, 2015 3년간 데이터 활용

- 데이터 기준 : 경상남도 구조구급과 응급출동내역 데이터 기준 타 데이터 매칭
- 현황 데이터 기준 : 2016. 7월 기준
- 수집데이터 현황

데이터 명	파일수	건수	데이터 출처
구조구급시스템데이터	8	약 263만 건	경남소방본부
세부 상황표 데이터	4	약 10만 건	경남소방본부 창원소방본부
구급인력데이터	5	약 1천3백건	경남소방본부 창원소방본부
인구통계데이터	2	약 7천 건	경남소방본부
기상데이터	1	약 116만 건	기상청
지역건강데이터	20	333 MB	보건복지부
경남 3대 응급환자 예후정보	4	약 2만2천 건	국립중앙의료원

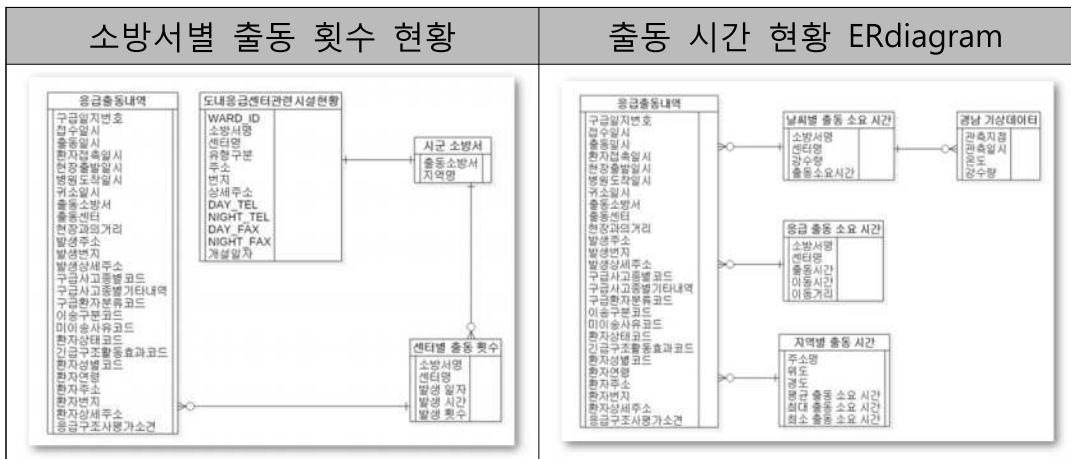
- 수집된 데이터는 단계별 정제 과정을 거쳐 분석에 사용할 수 있는 형태의 DB로 변환되어 저장되어 분석 모델의 요구 사항에 따라 분석 data set으로 구성
- 주요 데이터의 정제 과정은 기본적인 입력 오류 제거와 중복 요소 제거 과정을 거친 후, 결손 항목의 보정과 결합, 분리 등의 분석 오류 제거를 위한 정제 처리
- 결손 항목의 경우 항목이 포함된 파일의 다른 항목(레코드)을 검색하여 관련 자료 또는 유사 자료 (예, 구급사고종별코드, 구급사고종별기타내역, 응급구조사평가 조건 등)를 활용하여 보정, 관련 항목이 없거나 판단이 어려울 경우 제거하였음

원시 데이터명	속성 명	처리 내용
경남소방 본부 응급출동 내역	출동 지역	- 상세 주소지 없이 "경상남도"만 작성이 되어 있는 레코드 제거 - 경상남도 이외 지역 출동 데이터 존재, 해당 레코드 삭제
	출동 소방서명	소방서 명이 없는 경우는 출동 주소지를 이용하여 소방서명 생성 ex)경상남도 창원시의 경우 창원 소방서로 입력 - "김해소방서"의 경우 2015년 7월경 "김해서부소방서"와 "김해동부소방서"로 분리되어 2010년부터 2015년 7월까지 "김해소방서"로 작성된 데이터를 "김해서부소방서"와 "김해동부소방서"로 분류 - 분류를 위해 "김해서부소방서" 소속 센터와 "김해동부소방서" 소속 센터를 이용하여 김해소방서 매핑테이블을 아래와 같이 생성

		<table><tr><th>소방서명</th><th>센터명</th></tr><tr><td>김해동부소방서</td><td>구조대</td></tr><tr><td>김해동부소방서</td><td>내외119안전센터</td></tr><tr><td>김해동부소방서</td><td>대동119지역대</td></tr><tr><td>김해동부소방서</td><td>동상119안전센터</td></tr><tr><td>김해동부소방서</td><td>북부119안전센터</td></tr><tr><td>김해동부소방서</td><td>삼정119안전센터</td></tr><tr><td>김해동부소방서</td><td>상동119안전센터</td></tr><tr><td>김해동부소방서</td><td>생림119안전센터</td></tr><tr><td>김해서부소방서</td><td>장유119안전센터</td></tr><tr><td>김해서부소방서</td><td>주촌119안전센터</td></tr><tr><td>김해서부소방서</td><td>진례119안전센터</td></tr><tr><td>김해서부소방서</td><td>진영119안전센터</td></tr><tr><td>김해서부소방서</td><td>한림119안전센터</td></tr></table>	소방서명	센터명	김해동부소방서	구조대	김해동부소방서	내외119안전센터	김해동부소방서	대동119지역대	김해동부소방서	동상119안전센터	김해동부소방서	북부119안전센터	김해동부소방서	삼정119안전센터	김해동부소방서	상동119안전센터	김해동부소방서	생림119안전센터	김해서부소방서	장유119안전센터	김해서부소방서	주촌119안전센터	김해서부소방서	진례119안전센터	김해서부소방서	진영119안전센터	김해서부소방서	한림119안전센터
소방서명	센터명																													
김해동부소방서	구조대																													
김해동부소방서	내외119안전센터																													
김해동부소방서	대동119지역대																													
김해동부소방서	동상119안전센터																													
김해동부소방서	북부119안전센터																													
김해동부소방서	삼정119안전센터																													
김해동부소방서	상동119안전센터																													
김해동부소방서	생림119안전센터																													
김해서부소방서	장유119안전센터																													
김해서부소방서	주촌119안전센터																													
김해서부소방서	진례119안전센터																													
김해서부소방서	진영119안전센터																													
김해서부소방서	한림119안전센터																													
	출동 소방서명	<ul style="list-style-type: none">- 김해 동부 소방서와 김해 서부 소방서의 경우 안전센터 및 지역대 이름에 구분 값("신")이 존재함- 기존 김해소방서 매핑 테이블과 유일성을 위해 구분 값 제거																												
	접수일시	<ul style="list-style-type: none">- 2010년도 1월부터 2016년 6월 30일까지의 범위 이외의 값 제거- 시각 데이터는 YYYYMMDDHHMMSS형태로 존재(14자리), 시각 데이터형태에 만족하지 않는 레코드 제거																												
	접수 시간 출동 시간 환자 접촉 시간	<ul style="list-style-type: none">- 접수시각, 출동시각, 환자접촉시각의 경우 14자리의 문자열로 기록- 접수시각 중 14자리가 아닌 데이터 제거(1, 구급일지번호만 존재)- 출동시각 중 14자리가 아닌 데이터 제거(107, 12자리)- 접수시각 중 14자리가 아닌 데이터 제거 및 미기록 값 제거 (미기록 : 244969, 14자리 이하 2개) 917400 레코드 수- 정제된 데이터를 YYYY/MM/DD HH:MM:SS로 변환 수행																												
도내출동 차량현황	장비 종류	<ul style="list-style-type: none">- 펌프차의 경우는 구급 장비로 활용되고 있지 않으나, 향후 펌블런스로 사용가능성이 있기 때문에 사용- 구급차의 경우 특수, 일반 등의 속성과 크기에 상관없이 모두 인명구조로 사용가능하기 때문에 구급차라는 하나의 카테고리로 재정렬																												
심정지 환자세부 상황표	데이터 병합	<ul style="list-style-type: none">- 심정지 원시데이터의 경우 2012년을 기준으로 2개의 데이터로 나누어져 있으며, 두 개의 데이터는 서로 다른 속성 값을 가짐- 분석에 사용할 데이터를 각각 추출하고 통합하여 하나의 심정지 환자 통합 테이블 생성																												

중심 증외상 노혈관 질환 심정지	데이터 병합	- 출동 내역과 3대 질환 발생 현황의 데이터 분석을 위해 보고서 번호로 조인 연산을 통해 출동 내역과 동일 기간의 발생 환자 추출
구급사 종별 분류	데이터 보정	- 동물/곤충의 경우 종별코드에서 동물/곤충, 동물/곤충 기타 내역 분류 - 기타 종별 내역과 응급사 구조 소견에서 곤충 및 동물 분석

- 데이터 정제 후 현황 정보 제공을 위하여 수집 데이터를 현황 보고서 시각화에 맞도록 DB화하여 저장
- 현황 정보는 경상남도에서 발생하는 응급 출동 전체에 대한 요소별 현황을 파악하기 위한 것으로 현업에서 각종 보고서 작성 및 대 도민 서비스의 기초 자료로 활용
- 현황 보고서를 위하여 5개의 카테고리로 구분하여 DB set를 구성하고 각 DB의 연관관계를 ERdiagram 으로 구성
 - 소방서별 출동 횟수 현황
 - 도내 출동 차량 현황
 - 도내 출동 인력 현황
 - 3대 응급 질환 출동 현황
 - 출동 시간 현황



- “응급환자 위험군 분석”을 위하여 2010년부터 2016년까지의 3대 응급환자 출동

정보를 통계 분석하여 분석 연령대를 상세 정보데이터 기준, 전체 출동 대비 5% 이상의 비율을 차지하는 50대에서 80대까지의 연령대에 대하여 분석

● 연령대별 3대 응급 출동 비율

	남성	여성	합계
10대 이하	0%	0%	0%
10대	1%	1%	3%
20대	2%	2%	4%
30대	3%	3%	6%
40대	7%	4%	11%
50대	12%	6%	18%
60대	10%	6%	16%
70대	12%	10%	22%
80대	7%	9%	17%
90대	1%	2%	2%
100대	0%	0%	0%
기타	0%	0%	0%
총합	56%	44%	100%

- 각 응급 질환별 결정 계수를 사용하여 지역별로 예측한 연령별 발생 건수와 2015년 실제 자료와의 비교 검증을 통해 회귀분석을 통한 연령별 위험질환의 예측 확률을 검증
- 회귀분석을 통하여 연령대가 3대 응급질환에 미치는 영향을 검증
- 3대 응급질환별 환자 발생 자료와 연령간의 관계를 표현할 수 있는 회귀 모형식 간의 차이(잔차제곱)가 최소화 되는 결정 계수 (회귀 계수)를 도출

속성 값	회귀계수
70대 여성	-0.04138540
80대 여성	-0.01311675
40대 남성	-0.01531420
50대 남성	0.02623844
60대 남성	-0.03004949
70대 남성	0.11198972

- 각 응급 질환별 결정 계수를 사용하여 지역별로 예측한 연령별 발생 건수와 2015년 실제 자료와의 비교 검증을 통해 회귀분석을 통한 연령별 위험질환의 예측 확률을 검증



- 3대 응급 질환의 예측률이 각각 최소 60% 이상으로 분석 결과를 지역별, 연령별, 응급 질환별 위험군 알림 자료로 활용
- 응급자원(응급 상황 발생 시 출동 가능한 차량과 인원)은 3대질환자의 생존율과 직접적으로 관련이 있는 중요한 요소의 하나임
- 응급자원 배치는 3대 응급질환 환자가 발생하였을 경우 골든타임 이내 출동, 현장 조치 후 의료기관으로 이송을 위해 적정자원을 적소에 배치하는 것으로 응급자원의 효율성에 영향을 미침
- 경남지역의 3대 응급환자 생존율을 향상시키기 위해서는 기존 응급 자원의 배치를 변경하여 효율을 높이는 방법과 신규 자원을 추가하여 가용한 자원을 증가시키는 방법을 고려할 수 있음
- 기존 응급자원 재배치를 통하여 효율을 높이는 경우 한정된 자원으로 최적의 효

율을 얻을 수 있는 위치를 분석을 통하여 도출하여야 하며, 신규 자원 추가시의 경우에도 무한정 자원을 늘릴 수 없어 증가분에 따라 최고의 효율을 낼 수 있는 장소를 분석하여 찾아내야 함

- 신규 자원의 증가는 타당성 조사, 예산 심의 및 편성, 시행에 이르기까지의 일련의 과정에 소요되는 시간이 필요하여 중장기적인 시각에서 준비하여야 함
- 기존 자원의 재배치의 경우 신규 자원 증가에 비해 시간과 비용 소요가 적어 단기적인 시각에서 효율성을 검증할 수 있음.
- 경남 “응급환자 골든타임 확보를 위한 빅데이터 분석” 사업은 단기는 물론 중장기적인 관점에서의 해결책 제시를 위하여 기존 응급자원의 재배치와 신규 응급자원의 추가 시 최적 위치 모두에 대해 분석을 실시하였음
- 기존 응급자원 재배치의 경우와 신규 자원 추가 시 최적 위치 도출의 두 경우 모두 후보지를 제한하여 실제 배치 가능성과 계산 과정의 최적화를 도모하였음
후보지로는 공공 기관을 대상으로 기존 배치 지역과 중복되지 않는 곳을 선정
- 분석을 위한 데이터는 현황 정보와 다른 정제 과정과 보정 작업을 하여 분석에 적합한 데이터 세트를 구성

데이터 세트	처리 내용
사고 발생 데이터	<ul style="list-style-type: none"> - 이송 데이터와 차량 장비 데이터 중 차량은 존재하지 않지만 응급출동은 있는 내역을 제거 (차량이 존재하지 않는데 출동한 내역으로 기준이 되는 차량 숫자에 문제가 발생. 따라서 해당 데이터는 공정성을 위해 제거) - 구급대에 차량이 1대 존재하는데 복귀전에 또 다른 출동기록이 있는 경우의 데이터 제거

중복 출동 데이터	<ul style="list-style-type: none"> - 보고서 번호는 다르나, 접수시간과 출동소방서, 연령, 성별이 같은 경우 동일사건의 중복으로 간주하고 제거 (키 : 접수시간, 소방서, 연령, 성별) - 보고서 번호와 연령은 다르나, 접수시간, 출동시간, 환자접촉시간, 병원도착시간 등이 동일한 경우 환자는 여러 명이 존재하나 하나의 응급차로 이송한 것으로 간주하고 하나의 데이터로만 사용 (키 : 접수시간, 출동시간, 환자접촉시간, 병원출발시간, 병원도착시간, 귀소시간, 출동 소방서)
주소 좌표 변환	<ul style="list-style-type: none"> - 출동 위치를 지도상에 표시하기 위하여 주소좌표 변환 API 사용 (Geo coding API, 도로명주소, 지번 주소 모두 지원) - 출동 데이터에 번지까지의 데이터가 모두 나오지 않고 읍/면/동까지만 기재되어 있는 경우, 좌표 변환 시 군청이나 면사무소와 같은 지역 대표 위치로 반환 됨. 이러한 경우 분석 시뮬레이션을 돌릴 경우 후보지에 유리한 결과로 도출되기 때문에 작성된 주소지가 모두 만족하지 못하는 경우 해당 데이터는 제거
병원 데이터	<ul style="list-style-type: none"> - 병원의 주소지는 이송 결과 데이터의 병원 명을 이용하여 도출하나 사람이 입력하면서 문제가 발생하며, 동일한 병원에 대하여 다른 표현으로 인해 병합이 어려움 (예, 대구 W병원의 경우, 더블유병원, W 병원, 대구 W병원, w 병원 등) - 이 경우 특수문자 제거, 공백 제거 후 부분 문자열을 비교하여 병원명 도출, 2개 이상의 병원이 존재할 경우 최근거리 병원으로 가정 - 경남 지역 소재 병원이 아닌 경우 예외적인 경우로 보고 데이터 제거
이동 시간 데이터	<ul style="list-style-type: none"> - 사건 접수 후 병원까지 이송하는데 소요 시간 <ol style="list-style-type: none"> 1. 접수에서 출동 시작까지의 출동 준비 시간 2. 출동 시작부터 환자 접촉까지의 출동 시간 3. 환자 접촉부터 병원으로 출발 전까지의 환자 조치 시간 4. 병원으로 출발부터 병원 도착까지 병원 이동시간 5. 병원에서 출동지로 복귀 시간 - "1", "3" 소요 시간의 경우 출발지가 바뀌어도 변하지 않는 시간으로 분석 시 재계산 대상에서는 제외되나 전체 골든타임 계산에는 포함 - 각 지점간의 거리가 500m 이하인 경우 10km/h, 그 이상의 경우 60km/h 로 이동 속도 가정

■ 기존자원 재배치 분석 결과 최적 위치(샘플)

시/군/구	현 출동 지역	분석 결과 최적 재배치 지역
거제시	하청119지역대	장평동사무소
	옥포119안전센터	옥포1동사무소
	거제119안전센터	고현동사무소
	연초119안전센터	수양동사무소
	동부119안전센터	옥포2동사무소
	신현119안전센터	장승포119안전센터
	장승포119안전센터	능포동사무소
	연초119안전센터	아주동사무소
거창군	위천119안전센터	대평119안전센터
	가조119안전센터	남하면사무소
	신원119지역대	북상면사무소
	대평119안전센터	거창읍사무소
	대평119안전센터	웅양면사무소
고성군	거류119안전센터	현장대응단
	고성119안전센터	영현면사무소
	영오119지역대	동해119지역대
	회화119안전센터	고성119안전센터
	고성119안전센터	회화119안전센터

■ 기존자원 재배치 적용 시 골든타임 개선 효과 (시뮬레이션 결과)

○ 심정지 5분

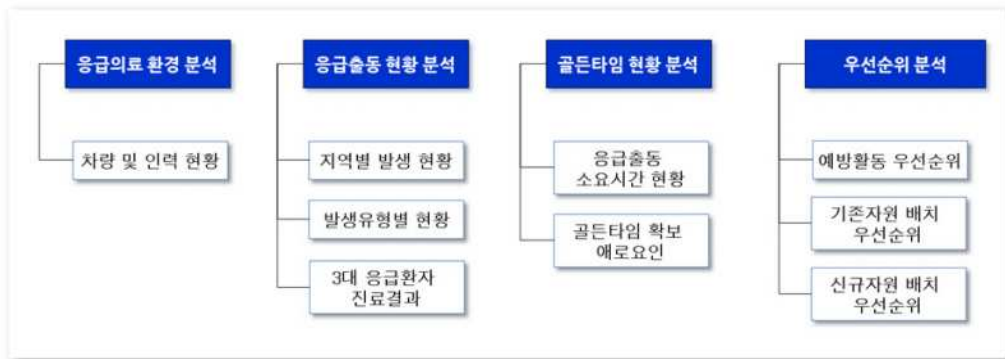
시/군/구	재배치 후 확보율	기존 골든타임 확보율	개선 효과
남해군	100.0%	14.9%	571.8%
창원시 마산합포구/회원구	100.0%	31.0%	222.1%
창원시 진해구	100.0%	42.5%	135.1%
거제시	88.4%	25.4%	247.6%
창원시 의창구/성산구	57.5%	41.1%	39.8%
김해동부	46.2%	28.9%	59.8%
진주시	44.6%	29.5%	51.3%
하동군	39.1%	19.2%	103.3%
사천시	38.8%	32.2%	20.5%
양산시	37.0%	28.7%	29.1%
통영시	31.0%	26.7%	16.3%
김해서부	28.5%	21.6%	31.8%
거창군	28.4%	24.2%	17.0%
밀양시	24.2%	23.0%	4.8%
창녕군	23.8%	23.2%	2.2%
산청군	23.2%	15.7%	47.8%
의령군	21.9%	21.3%	3.0%
합천군	18.3%	15.3%	19.3%
함안군	17.9%	16.0%	11.7%
함양군	15.1%	14.8%	2.1%
고성군	14.8%	14.5%	2.4%



바 행정적용

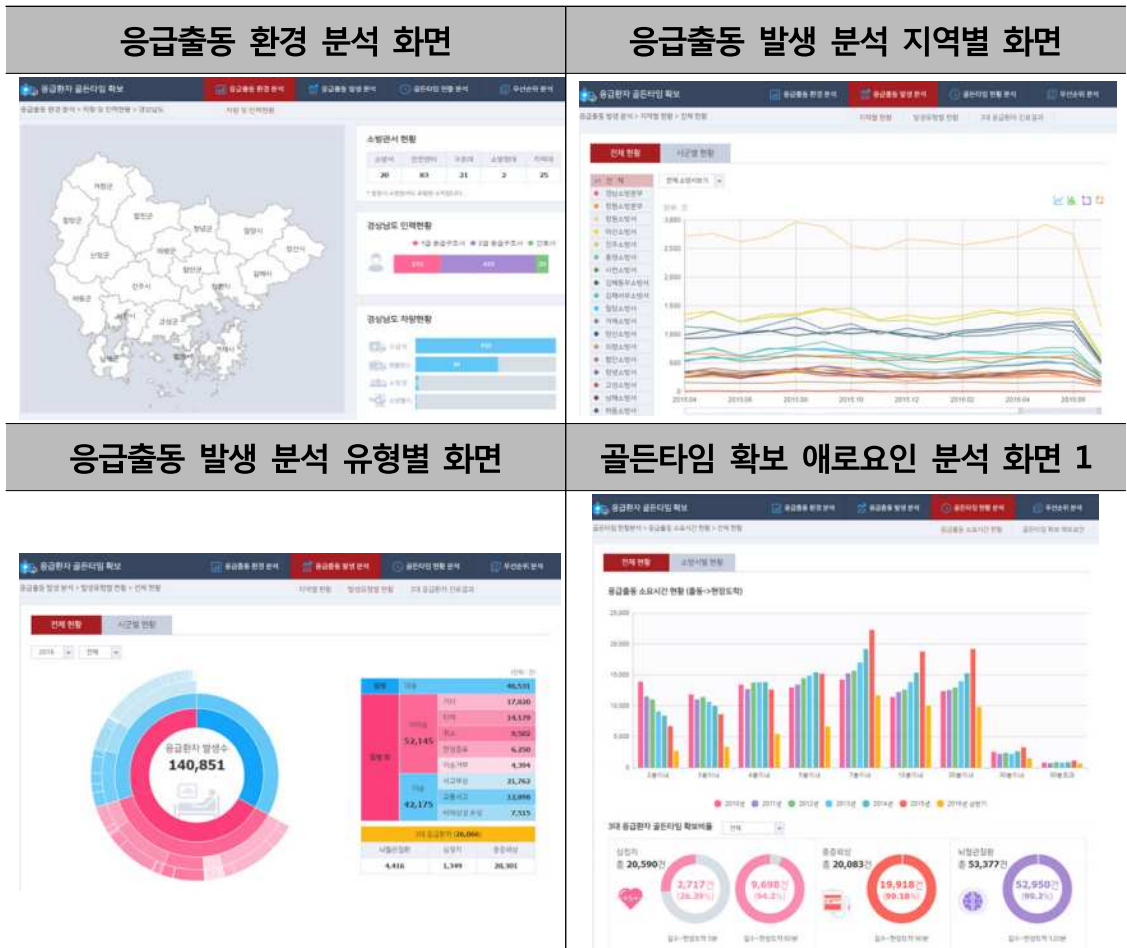
- 빅데이터 분석 결과는 행정 활용을 위한 담당 공무원 활용과 대 도민 홍보용으로 분리하여 각각의 홈페이지를 구성하여 서비스 제공
- 제공할 분석 내용은 소방행정에 대한 현황 정보와 출동 기록 현황의 시계열 분석, 3대 응급 환자 골든타임 현황 및 애로 요인 분석 그리고 예방 활동 및 응급 자원 배치 분석에 대한 내용 등 4가지의 카테고리로 분류하여 제공

골든타임 확보를 위한 빅데이터 분석 결과 서비스 항목



- 빅데이터 분석 결과를 이용한 근거 기반의 과학행정을 위한 행정 서비스는 로그인 페이지를 제공하여 분석 결과에 대한 상세한 내용과 사건 사고 현황 정보 등에 대한 정확한 정보를 제공
- 또한, 골든타임 확보 현황과 애로 요인에 대한 정보를 제공하여 개선점을 도출할 수 있는 시사점 제공
- 응급 자원 재배치 시 효과 추천 위치를 검색할 수 있으며 신규 자원 추가 시 규모별 최적 위치와 효과를 확인할 수 있도록 구성

- 대 도민 서비스의 경우 분석 내용 중 지역별, 질환별 중 발생 연령대 정보를 제공하여 사전에 예방 정보 제공
- 행정서비스를 위한 서비스 구성은 아래와 같이 직관적이고 보기 수월하도록 인포그래픽 형태로 제공하며 표시 정보는 엑셀 파일로 다운이 가능하여 각종 보고서 및 정책 수립 시 활용 가능하도록 제공



골든타임 확보 애로요인 분석 화면 2



발생위험군 분석화면



기존자원 배치 우선지역 분석 화면



신규자원 배치 우선지역 분석 화면

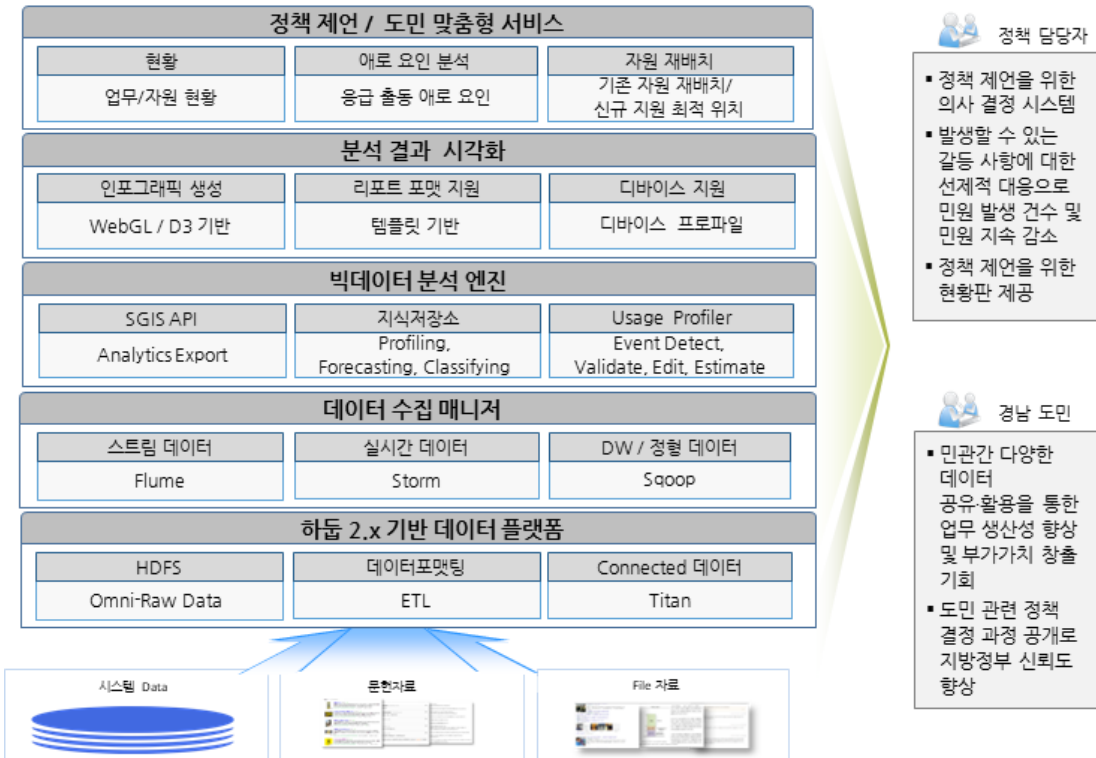


- 경남소방본부는 분석 결과 활용을 위해 6개월에 걸친 사전 조사와 준비 작업을 거쳐 2017년 5월 1일부터 도내 23개 지역에 'M불런스' 이동구급대 운영을 실시
- 'M불런스'는 이동(Mobile)의 'M'과 구급차(ambulance)의 합성어로 빅데이터 분석 결과를 반영한 골든타임확보 취약지역의 요일, 시간대별로 재배치 우선 지역에 구급차를 이동 배치하여 선제적 대응을 통해 골든타임 확보율을 향상시키기 위한 맞춤형 구급서비스

사

인프라 구축 단계

- 경남의 응급환자 골든타임 확보를 위한 빅데이터 분석은 글로벌 표준의 오픈소스 소프트웨어 (Hadoop) 기반의 Scale-out 방식의 빅데이터 플랫폼 구축



- 향후 알고리즘과 프로그래밍 언어가 개선되어지면 최소한의 수정만으로 적용이 가능하도록 주요 분석 단계를 모듈식으로 설계
 - 개방형 구조로 Tier별 확장/장애/분산 구조 채택
 - 시스템 성능을 위한 병렬/부하분산 구조
 - System S/W 와 Business Application을 Common Service Framework과 분리하여 독립성을 최대한 확보
 - 시스템 S/W(Middleware/OS/DBMS)의 변화에 따른 Application의 변경을 최소화

- 검증된 모듈을 활용함으로써 시스템의 성능, 안정성 확보
- DB I/O를 통해 개발자의 숙련도에 관계없이 DB Access 성능 보장
- 분석을 위해 사용된 데이터는 모두 공공데이터를 기반으로 하고 있으며 분석 모델링을 위해 사용된 알고리즘과 프로그램 언어 등 100% 공개 소프트웨어를 사용
- 데이터 정제, 분석 세트 구성 및 분석 알고리즘에 대한 정의서 매뉴얼화하여 시범 사업의 결과를 다른 지역으로의 확산 및 개선이 용이하도록 제공

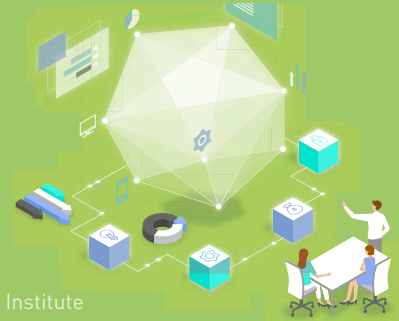


CHAPTER

III

지방자치단체 빅데이터 분석 사례집

남양주 보건 의료지표의 행정활용 사례



korea Local Information Research & Development Institute

가. 추진배경	108
나. 프로젝트 개요	108
다. 적용 프로세스	113
라. 기획·관리 단계	114
마. 데이터 분석 단계	122
바. 행정적용	135
사. 인프라 구축 단계	137

III. 남양주 보건 의료지표의 행정활용 사례

요 약

남양주시에서는 보건의료 행정 서비스가 효율적으로 기획되지 못하였으며 그 원인은 지역보건의료 계획의 지표가 계층 및 지역별로 세분화하여 제공되지 않는 것에 기인한다는 문제의식을 가지고 있었다. 질병정보와 건강검진 정보를 관리하고 있는 국민건강보험공단이 제공하는 지역별 보건의료 지표 역시 시군 구별로 편성되어 있다는 점도 감안해야 할 사항이었다.

시는 문제해결을 위하여 남양주시 보건소 담당 팀장들과 협의하여 지자체 요구 사항을 정리하고, 국민건강보험공단과 협의 후 지역별 보건의료 지표의 개선 프로젝트에 참여하여 지자체 의견을 개진하였다. 지역보건의료계획에 소지역별 지표를 제시하고 각 프로그램별 대상과 계층을 명시할 수 있는 세부 방안 또한 마련하였다.

분석단계에서는, 국민건강보험 데이터를 분석하여 읍면동별 담당자가 맞춤형 행정 서비스 계획을 세울 수 있도록 분석모델을 설정하고 국민건강보험의 건강 지표 분석방법과 분석결과에 대한 보건소의 행정활용목적을 명확히 하고 그에 적합한 방법론을 정하였다. 국민건강보험의 지역의료이용지표를 통해 건강 지표 및 만성질환 지표에 대한 데이터를 확보하고 필요에 따라 맞춤형 DB를 요청하여 수집하고 분석할 수 있도록 가공하였다. 수집하여 가공한 데이터를 적절한 도구 BI(Business Intelligence SW)를 이용하여 분석하고 분석결과를 검증하였다. 분석결과를 바탕으로, 빅데이터 분석이 지속적으로 이루어질 수 있도록 데이터를 정비하여 빅데이터 시스템에 저장하고 현업의 활용도 향상을 위해 대시보드도 구성하였다.

이러한 데이터 분석 활동을 기반으로 흡연율 감소를 위한 다양한 캠페인 및 예방 교육 활동을 위해 읍면동별, 성별, 연령별 상세 관리 대상을 정의하여 시범 적용할 수 있었다.

가 추진배경

- 지자체 보건행정 서비스 개선의 최우선 선결과제인 시민의 질병과 건강에 관한 소지역 단위별 지표와 Profile정보를 확보
 - 시군구 단위에서 읍면동 단위로 : 인구규모를 고려하지 않고 시군구 전체에 Parameter 하나를 부여하는 기존 방식에서 한 단계 더 내려간 읍/면/동 단위의 지표와 Profile을 확보
- 보건소의 각종 프로그램별 정확한 Targeting(대상정의, 대상규모)을 통해 기획의 효율을 높이고 성과의 측정을 가능하게 함
 - 세분화된 차원(Targeting) : 읍면동별/연령대별(5세단위)/성별/보험료기준별/직역별/연도별

나 프로젝트 개요

- 개 요
 - 제6차 지역보건의료 계획(2014~2018)을 근간으로 남양주시는 매년 연 단위 지역보건의료 계획을 수립하여 행정서비스를 기획, 시행하고 있음
 - 만성질환 등에 대한 지표와 건강증진 프로그램을 위한 지표 등을 주민센터 단위의 서비스에 맞도록 상세화 할 필요가 있음
 - 국민건강보험공단의 지역별 보건의료 지표서비스를 활용하여 Micro Targeting이 가능하도록 개선함
- 소지역 단위 보건의료 지표의 필요성
 - 지역보건의료 계획에는 시군구 단위로 단일한 보건과 의료 관련 지표(Parameter)가 제공되어 읍면동 단위 혹은 주민센터 단위의 맞춤형 행정서비스 기획에 장애가 되어 왔음
 - 국민건강보험 공단이 제공하고 있는 지역별 보건의료 지표를 지자체에서 사용하기 편리하도록 개선하여 활용사례를 확대할 필요가 있음

■ 주요 내용: 소지역 단위 보건의료 지표

- 관리대상 질환: 고혈압, 당뇨 등 주요 만성질환 및 암 등 15종
- 지표 종류: 건강검진 수검률, 건강위험요인, 만성질환 의료이용 및 관리 등 56종

<참고> 남양주시 제6기 지역보건의료계획



■ 보건의료계획 수립 배경

- 지역의 실정에 맞는 지역보건의료계획을 수립하여 지역주민의 보건의료서비스 질을 향상시키고 궁극적으로 주민 건강 향상에 기여
- 남양주시 보건의료계획의 슬로건은 “시민이 행복한 건강도시 남양주”이며 이를 위한 3가지 축이 제시되고 있는데, 그 각각은 다음과 같음
 - 보건의료 서비스 강화
 - 보건의료 전달체계 구축 및 활성화
 - 보건기관 자원 확충

- 한편, 남양주시는 위에 열거한 3가지 축을 실현하기 위하여 15개의 구체적 프로그램들을 추진해 나가겠다는 전략체계 아래, 매년 세부적인 예산과 일정을 세워 행정서비스를 제공해 나가고 있음

관점	현황
보건의료수요측면	◆ 일반현황 (인구분포 및 증감)
	◆ 지역의 건강수준(남양주시 건강행태)
보건의료공급측면	◆ 지역사회 주민 및 전문가 보건사업 수요도 파악
	◆ 지역보건체계 <ul style="list-style-type: none"> • 보건소 조직/인력/예산 • 보건소 주민 만족도 • 보건의료기관 의료인력수준 • 지역사회 협력수준

- 지역사회 현황 분석

- 3대 추진 분야

추진 분야	추진 과제
지역주민 건강향상을 위한 지역보건 서비스 강화	◆ 건강생활실천을 위한 NCD 예방
	◆ 스트레스, 우울 및 자살예방을 위한 정신건강관리
	◆ 지역별, 계층별, 연령별 건강격차 감소를 통한 건강 형평성 향상
지역보건 전달체계 구축 및 활성화	◆ 건강한 삶을 위한 지역사회 보건의료자원의 역량강화 및 연계 활성화
지역보건기관 자원 확충	◆ 보건의료 서비스의 효과적 제공을 위한 인프라 확보

- 분야별 중장기 계획

추진 과제	추진 사업
건강 생활 실천을 위한 NCD 예방	◆ 건강생활교실, 구강보건사업, 아토피천식 예방관리, 심뇌혈관질환 예방관리, 암관리사업, 건강검진사업, 진료사업
스트레스, 우울 및 자살예방을 위한 정신건강관리	◆ 정신보건사업, 치매관리사업
지역별, 계층별, 연령별 건강격차감소를 통한 건강 형평성 향상	◆ 여성 어린이 건강증진사업, 방문건강관리사업, 감염병 예방관리사업, 결핵관리사업, 예방관리사업(예방접종), 에이즈/성매개감염병 예방관리사업
건강한 삶을 위한 지역사회 보건의료자원의 역량강화 및 연계 활성화	◆ 지역보건 전달체계 개선, 지역사회 자원 협력 및 역량
보건의료 서비스의 효과적 제공을 위한 인프라 확보	◆ 권역별 건강생활지원센터 확충, 보건소 장비 보강 계획, 보건소 조직 및 체계 정비, 인력의 양적 확충 및 질적 강화, 예산 확충 및 보강

● 지표개선 방안

- 남양주시 전체 단위로 나와 있는 흡연율을 16개 읍면동 및 남녀와 연령대 단위로 세분화된 지표로 개선하였다.
- 사업대상의 중점 관리계층을 명시하기 위하여 세분화된 지표 생성 (Targeting)

현행	개선
개괄적인 수치로 현황 파악대상 : 전체 시민	◆ 읍면동별, 연도별, 연령별, 특성별 상세 현황 ◆ 지표 간 연계분석(다차원) ◆ 대상 구체화

● 개선방안 총괄

부문	현행	개선
현황분석	시단위의 단일지표만 제시	읍면동/연령대/성별/소득 별로 구체적인 대상 인구 수와 지표를 제시
분야별 성과지표 및 목표	개괄적인 목표값만 부여 세부단위의 성과지표 없음	
중장기 추진과제	시단위의 목표치만 제시	
세부사업 구성	관련 프로그램의 시행횟수 위주 제시	

● 개선방안사업별

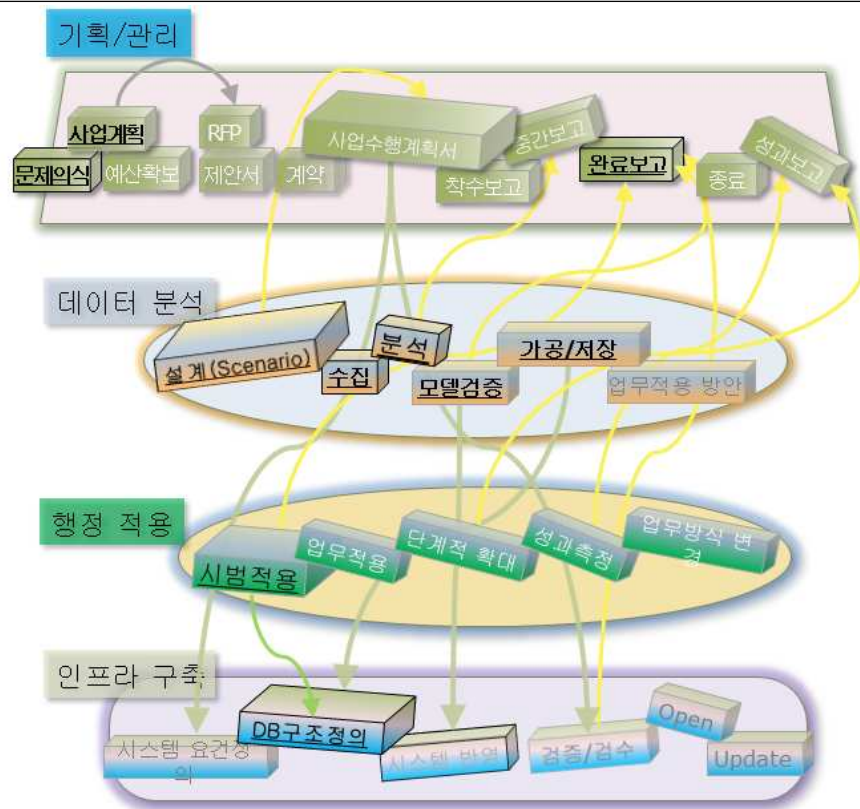
사업부문	현행	개선
금연지원서비스	금연 클리닉 등록자수 (시단위)	세부 흡연율을 제시
심뇌혈관질환 예방관리	고혈압/당뇨병 등록률 (시단위, 누적)	고혈압 의료 이용률을 제 시하여 등록률의 목표치 를 설정하게 함
암관리사업	국가암조기검진 수검률 (시단위)	세부 수검률을 제시

다 적용 프로세스

■ 보건 의료 지표 분석 전체 프로세스

- 기획관리(문제의식 → 사업계획 → 완료보고)
- 데이터분석(분석모델설계 → 데이터수집 → 분석·검증 → 데이터가공·저장)
- 행정적용(시범적용)
- 시스템구축(DB구조정의/대시보드구성)

남양주시 적용 프로세스



라 기획 · 관리 단계

1. 기획 · 관리 단계 요약

단계	주요 내용
문제의식	<ul style="list-style-type: none"> ◆ 남양주시 보건의료 행정 서비스가 효율적으로 계획되지 못한 원인은 지역보건의료 계획의 지표가 계층/지역별로 제공되지 않는 것임 ◆ 질병정보와 건강검진 정보를 관리하고 있는 국민건강보험공단이 제공하는 지역별 보건의료 지표 역시 시군구별로 편성되어 있음
사업계획	<ul style="list-style-type: none"> ◆ 남양주시 보건소 담당 팀장들과 협의하여 지자체 요구사항을 정리하고 국민건강보험공단과 협의하여 지역별 보건의료 지표의 개선 프로젝트에 참여하여 지자체 의견을 개진함 ◆ 지역보건의료계획에 소지역별 지표를 제시하고 각 프로그램별 대상과 계층을 명시할 수 있는 세부 방안을 마련함
완료보고	<ul style="list-style-type: none"> ◆ 분석결과물의 활용을 통해 개선되어 지는 점을 강조하고 다른 행정서비스 영역의 공감대 형성을 유도함 ◆ 단체장(남양주시/국민건강보험공단)의 관심을 힘입어 다음 단계의 보건 의료행정 서비스를 도모함

2. 문제의식

■ 지방자치단체의 맞춤형 보건의료 정책에 필요한 핵심 정보는 '지표'

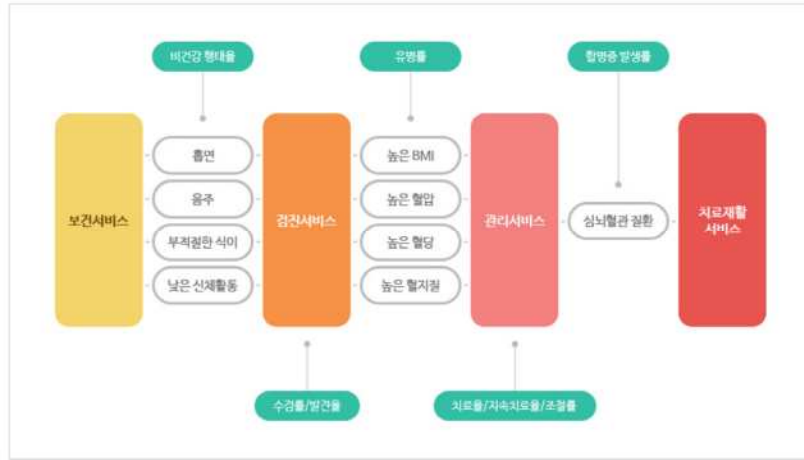
- 지자체가 담당하고 있는 지역을 주요 차원별로 세분화하여 각 계층별로 기존의 프로그램을 맞춤화하여 실행하기 위해서 '지표'의 세분화가 필수

■ 질병 및 건강 관련 정보가 모아지고 있는 기관인 국민건강보험공단은 건강보험자료 공유 서비스(<https://nhiss.nhis.or.kr>)를 통해 시군구 단위의 단일 지표를 제공해 오고 있음

- 지자체의 의견을 반영하여, 최근(2017년 11월), 읍면동별 연령대별(5세 단위)로 세분화한 정보를 제공하기 시작함

과제	<ul style="list-style-type: none"> ◆ 문제점 확인 ◆ 문제해결을 위해 필요한 활동(대안)을 나열 ◆ 필요한 활동(대안)을 통해 얻을 수 있는 효과를 확인
추진내용	<ul style="list-style-type: none"> ◆ 문제점 확인 <ul style="list-style-type: none"> • 지역보건의료계획에 시군구 단위의 지표만 제공되어 읍면동별 담당자의 맞춤 행정서비스 기획이 불가능함 • 현재(2016년 12월) 제공되는 국민건강보험공단 건강보험 자료 공유서비스를 통해서 상세 지표를 제공 받을 수 없음 • 기초자치단체의 보건 의료 행정 서비스 프로그램의 효율을 높이고 계층별 형편에 맞게 개선하려면 세분화된 지표가 필요 ◆ 문제해결을 위해 필요한 활동 나열 <ul style="list-style-type: none"> • 현재 적용되고 있는 “지역보건 의료 계획(제6기)”의 내용 검토 • 맞춤형 정책수립을 위해 필요한 ‘차원의 세분화’에 대한 정의 : 읍면동별 5세 단위 연령대별, 추세비교, 타 시군 및 광역지표와 비교 ◆ 맞춤형 보건 의료 지표를 통해 얻을 수 있는 효과 <ul style="list-style-type: none"> • 보건 의료 행정 서비스 프로그램별 기획 가능: 대상 인구 추정 및 세부 프로파일(Demographics)에 맞도록 개선 • 모니터링 지표의 세분화 : 읍면동별, 연령대별, 연도별, 타시군/광역 등의 지표대비 비교 • 보건 의료 행정 서비스 프로그램의 개선: 세부 계층별 프로그램 시행→ 세부 계층별 성과 측정→ 프로그램에 반영
Tip	<ul style="list-style-type: none"> ◆ 지자체의 보건 의료 정책의 중요성에 비해 관련한 지역별 계층별 지표의 제공이 매우 부실한 실정임 ◆ 지자체가 보유/관리하고 있는 다른 행정정보들(예: 주민등록 인구, 전출입, 복지, 영유아, 교통, 의료시설, 체육문화시설 이용)과 융합하여 분석하면 보건 의료 행정 서비스의 획기적인 발전을 기대할 수 있음

<참고1> 국민건강보험공단의 의료이용지표



구분		지표 종류
건강검진 수검률 지표		1차일반·생애검진수검률, 2차일반·생애검진수검률, 암검진통합수검률, 위암검진수검률, 대장암검진수검률, 유방암검진수검률, 자궁경부암검진수검률, 간암검진수검률
건강위험요인 지표		흡연율, 음주율, 신체활동률, 비만율, 고도비만율, 초고도비만율, 복부비만율, 대사증후군 발견율
만성질환 의료이용 지표	고혈압	고혈압 의료이용률, 고혈압 신규의료이용률
	당뇨병	당뇨병 의료이용률, 당뇨병 신규의료이용률
	이상지질혈증	이상지질혈증 의료이용률, 이상지질혈증 신규의료이용률
	고혈압·당뇨병	고·당 의료이용률, 고·당 신규의료이용률
	고혈압·당뇨병·이상지질혈증	고·당·지 의료이용률, 고·당·지 신규의료이용률
	허혈성심질환	허혈성심질환 전체의료이용률, 허혈성심질환 입원의료이용률
	뇌혈관질환	뇌혈관질환 전체의료이용률, 뇌혈관질환 입원의료이용률
	만성신부전	만성신부전 전체의료이용률, 만성신부전 입원의료이용률
만성질환 관리지표	COPD	COPD 전체의료이용률, COPD 입원의료이용률
	간경화증	간경화증 전체의료이용률, 간경화증 입원의료이용률
만성질환 관리지표	골다공증	골다공증 전체의료이용률, 골다공증 입원의료이용률
	질병치료 및 합병증 지표	1년후 고혈압 투약순응률, 1년후 고혈압 정기방문율, 1년후 고혈압 조절률, 2년후 고혈압 조절률, 1년후 당뇨병 투약순응률, 1년후 당뇨병 정기방문율, 1년후 당뇨병 조절률, 2년후 당뇨병 조절률, 1년후 이상지질혈증 투약순응률, 1년후 이상지질혈증 정기방문율, 1년후 이상지질혈증 조절률, 2년후 이상지질혈증 조절률, 1년후 고당 정기방문율, 1년후 고당지 정기방문율, 1년내 심뇌신 신규의료이용률, 2년내 심뇌신 신규의료이용률, 3년내 심뇌신 신규의료이용률 등
항생제관리지표		항생제사용량·항생제처방률

3. 사업계획

- 사업에 대한 요구사항을 정의하여 계획을 수립하고 확정하는 단계
- 사업범위를 결정하고 소요자원 및 예산 등을 문서화하여 사업내용을 구체화함
 - 사업계획 수립 시 현황 및 문제점, 추진목표, 전략, 추진일정, 사업내용, 소요예산, 산출물 및 사업의 주체와 분석결과물의 활용 주체에 대한 정의가 필요
- 남양주시 지역별 보건의료 지표분석은 시범사업과 본 사업으로 나눠 실시함

과제	<ul style="list-style-type: none"> ◆ 구체적이고 명확한 사업계획 수립 필요 ◆ 분석내용 및 결과 활용 측면의 필요성을 강조하여 의사결정자에게 보건의료지표 분석사업 추진의 필요성을 인지시키는 것이 중요함
추진내용	<ul style="list-style-type: none"> ◆ 구체적이고 명확한 사업 계획 수립 <ul style="list-style-type: none"> • 사업의 필요성 확인 : 분석내용 및 결과물의 활용 측면에서 사업 추진 필요성 제시 <ul style="list-style-type: none"> ※ 의사결정권자의 이해는 예산승인을 받기 위해 중요함 ※ 남양주시의 경우 맞춤형 보건의료 행정 서비스에 대한 단체장의 관심과 의지가 매우 높았음 ※ 남양주시 보건소의 소장과 담당 팀장들이 적극적으로 필요성을 개진했을 뿐 아니라 요건정의에 직접 참여 ◆ 추진주체의 구분 명확화 <ul style="list-style-type: none"> • 분석사업의 추진주체와 분석 결과물을 활용하는 주체를 명확히 구분해야 함 <ul style="list-style-type: none"> ※ 지역보건의료계획의 주관부서: 보건소 ※ 맞춤형 지표를 만들고 시스템에 반영하는 부서: 빅데이터 팀 ※ 지자체의 의견을 수렴하여 기존의 지표를 개선해 준 기관: 국민건강보험공단(빅데이터 운영실) ※ 지역보건의료 행정서비스에 적용: 보건소 ◆ 분석목적 및 방법의 제시 <ul style="list-style-type: none"> • 분석목적과 방법을 결정할 때는 분석결과를 활용하는 현업부서와 사전에 협의하여 진행하는 것이 중요함

4. 완료보고

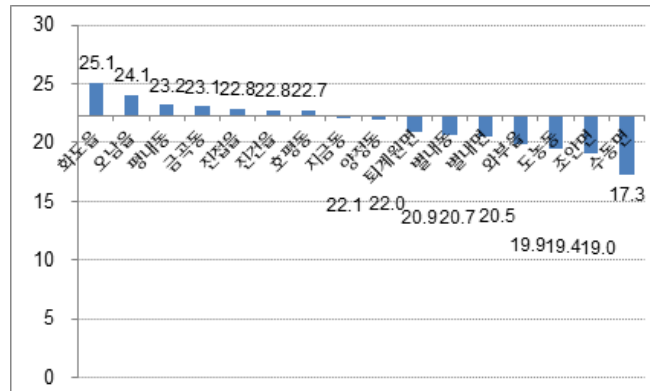
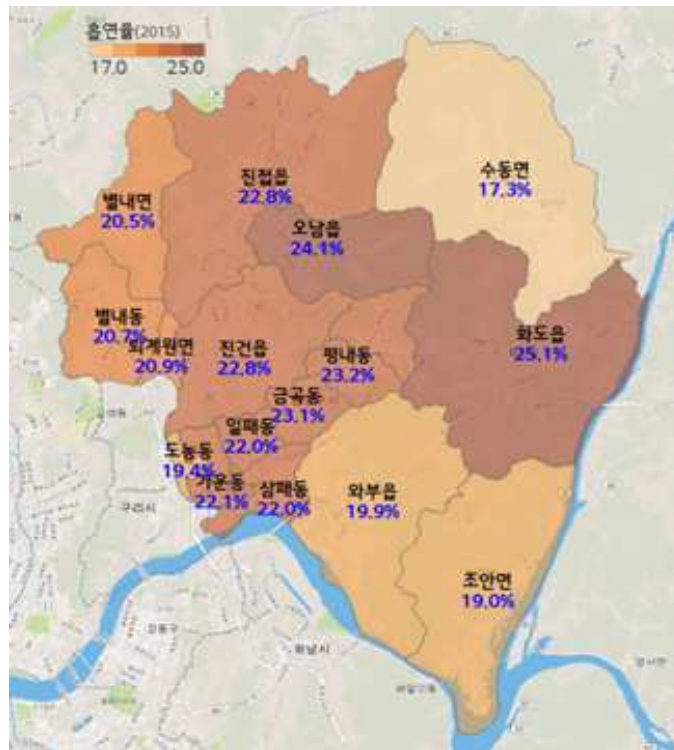
- 현업 부서장들과의 분석 결과물 활용에 대한 공감대 형성을 위한 의사소통의 방법으로 완료보고회를 실시
- 참석자들의 이해도를 높이고 공감대 형성을 위하여 분석 결과를 시나리오 형태로 재구성하여 제시
- 소지역별 보건의료 지표 분석의 활용 방안 및 개선방안을 논의

과제	<ul style="list-style-type: none"> ◆ 분석 결과물을 활용 주체의 눈높이에 맞게 업무적용 중심의 시나리오 형태로 재구성 ◆ 빅데이터 분석 결과에 대한 활용방안 및 효과적인 데이터 분석을 위한 빅데이터 시스템 개선방안 논의
추진내용	<ul style="list-style-type: none"> ◆ 발표 시나리오의 구성 <ul style="list-style-type: none"> • 분석결과에 대한 업무적용 시나리오를 만들어 분석결과를 재구성 • '주요질병의 읍면동별 현황'을 실제 데이터를 통해 보여주고, 연령대별/연도별 질병율에 대한 데이터 분석과 특정지역(별내동)을 예시로 하여 주민등록 정보를 이용한 서비스 대상 시민의 숫자 추정 분석결과를 제시 ◆ 활용방안 및 개선방안 제시 <ul style="list-style-type: none"> • 소지역단위 보건의료 지표분석 결과에 대한 보건소의 활용예시를 제시 • 금번 분석에서 부족했던 점을 도출하고 개선점 논의 • 이종데이터와의 융이한 연결 필요, 담당자 문서에 필요한 수치(표) 생성, 분석에 맞게 설계된 데이터 구조로의 변경 등 관련 빅데이터 시스템에 대한 개선사항을 제시
Tip	<ul style="list-style-type: none"> ◆ 산출물 및 현장적용방안은 관련 현업부서와 사전 협의 및 확인을 거친 후 추진

5. 결과물 샘플

주안점 1 : 분석의 시각화

- 분석된 결과가 주는 의미를 보다 간결하게 전달하기 위해 색(Color)을 이용하여 분석에서 보여주고자 하는 Key Message를 강조
- 읍면동별 흡연율 데이터를 그래프와 지도에 표현(단위:%)



주안점 2 : 연계분석을 통한 활용도 제고

- 다른 외부 혹은 이중 정보를 연계한 분석을 통해 단순 자료의 집계가 아닌 다차원 연계 분석으로 발전시킴으로써 정책정보로서의 가치와 활용도가 훨씬 높아짐
- 주민등록인구와의 연계분석을 통해 데이터 분석의 활용도와 가치를 향상시키고자 함
→ 우선관리 대상 인구를 추정
- 이외에도 주민등록 인구, 전출입, 복지, 영유아 보육, 교통, 의료시설, 체육문화시설 이용 등과의 연계를 통해서 활용도를 높이고자 함

지역 (단위:명,)	남성						여성					
	20 대	30 대	40 대	50 대	60 대	70 세~	20 대	30 대	40 대	50 대	60 대	70 세~
전체	40,002	47,414	62,750	52,392	29,743	21,212	36,914	49,921	61,186	49,490	31,966	29,694
Target 합	3,966	3,920	4,015	4,007	348	552	438	700	687	343	79	191
비율	9.9%	8.3%	6.4%	7.6%	1.2%	2.6%	1.2%	1.4%	1.1%	0.7%	0.2%	0.6%
화도읍	2,703	3,985	4,993	3,097	1,108	407	337	700	687	343	196	75
오남읍	1,559	1,874	2,883	1,496	358	279	105	198	268	208	38	36
평내동	1,133	1,361	1,874	922	271	148	224	198	237	112	24	50
금곡동	911	836	486	815	348	136	113	126	85	101	79	27
진접읍	2,472	3,392	4,324	2,341	672	450	357	402	428	246	49	115
진건읍	748	832	1,132	1,017	371	160	215	85	94	116	29	57
호평동	1,155	1,859	2,276	881	306	91	179	227	199	99	18	19
지금동	555	684	841	501	186	65	136	63	38	51	11	18
양정동			235	233	72				11	17	3	
퇴계원면	1,125	1,071	1,164	840	427	161	163	106	143	83	71	17
별내동	1,684	3,071	1,969	1,217	500	310	119	355	188	99	66	
별내면	622	593	779	509	289	82	72	126	52	43	22	8
와부읍	1,461	1,824	2,703	1,611	476	273	219	223	202	104	47	28
도농동	1,274	1,210	1,312	911	301	139	68	122	114	35	20	4
조안면			266	104	111					9		
수동면	189	161	243	248	108	109			16	34	25	40

마 데이터 분석 단계

1. 데이터 분석 단계 요약

단계	주요 내용
분석 설계	<ul style="list-style-type: none"> ◆ 국민건강보험 데이터를 분석하여 읍면동별 담당자가 맞춤형 행정 서비스 계획을 세울 수 있도록 분석모델을 설정하는 것 ◆ 국민건강보험의 건강지표 분석방법과 분석결과에 대한 보건소의 행정활용목적에 명확히 하고 그에 적합한 방법론을 정하는 것이 중요
데이터 수집	<ul style="list-style-type: none"> ◆ 국민건강보험의 지역의료이용지표를 통해 건강 지표 및 만성질환 지표에 대한 데이터를 확보하고 필요에 따라 맞춤형 DB를 요청하여 수집하고 분석할 수 있도록 가공
데이터 분석, 검증	<ul style="list-style-type: none"> ◆ 수집하여 가공한 데이터를 적절한 도구 (Business Intelligence SW)를 이용하여 분석하고 분석결과를 검증
데이터 가공 저장	<ul style="list-style-type: none"> ◆ 분석결과를 바탕으로, 빅데이터 분석이 지속적으로 이루어질 수 있도록 데이터를 정비하여 빅데이터 시스템에 저장 ◆ 현업의 활용도 향상을 위해 대시보드 구성

2. 분석 설계

- 분석설계란 결과를 도출하기 위해 필요한 전략을 세우는 일로, 분석을 계획하는 것
 - 분석설계에서는 분석을 통해 알아내고자 하는 대상(분석의 목적, 내용, 결과가 가지는 기대효과)을 명확하게 규정해야 함
 - 분석 목적에 적합한 분석모델을 결정해야 함
 - 분석설계는 관련 주제에 대한 데이터 분석모델을 정하는 것임
- 통계적 분석 방법을 쓸 것인가를 고려

- 분석 단위를 결정(성별, 연령, 지역, 보험료기준, 직역구분)
- 분석을 위한 적절한 통계 SW의 사용 여부를 함께 고려함

과제	<ul style="list-style-type: none"> ◆ 분석결과와 활용 목적과 기대효과를 구체적으로 정의 ◆ 분석목적과 기대효과에 적합한 분석 방법론 선정 ◆ 분석 시나리오 작성 및 이에 근거하여 필요한 DATA의 종류, 유형 등 결정
추진내용	<ul style="list-style-type: none"> ◆ 분석 결과의 활용목적과 기대효과를 구체적으로 정의 <ul style="list-style-type: none"> • 활용목적 <ul style="list-style-type: none"> 데이터 분석기반의 지역보건 의료계획 수립 읍면동별 계층별 맞춤형 보건의로 프로그램 시행 건강지표, 만성질환지표 분석의 틀 마련 • 기대효과 <ul style="list-style-type: none"> 읍면동별 맞춤 보건의로 행정서비스 지원 보건 의료 행정 서비스의 질 향상 빅데이터 분석 및 현업 활용을 통한 업무 효율성 향상 및 데이터 기반의 업무 방식 혁신 • 적합한 분석방법 선정 <ul style="list-style-type: none"> 지표별 성별, 연령, 지역, 보험료기준, 직역구분별 빈도를 집계 하고 항목별 교차분석표를 작성하여 유의미한 특성을 파악 ◆ 분석 시나리오 및 분석에 필요한 데이터 항목, 유형 등을 결정 <ul style="list-style-type: none"> • 분석 시나리오 설정 : 연구내용을 명확하게 하고 분석에 불필요한 것들을 배제하는 등 원활한 분석활동을 위해 시나리오를 설정 ◆ 건강 지표 및 만성질환 지표 분석 시나리오 <ul style="list-style-type: none"> • (1단계) 주제와 관련한 핵심 수치(key Numbers)를 제시함 : 읍면동별 건강지표인 흡연율, 비만율, 대사증후군발견율, 암수검률, 권장신체활동 수행율과 만성질환 지표인 고혈압 의료 이용률, 당뇨병 의료 이용률에 대한 지표값 제시 • (2단계) 제시한 핵심 수치(Key Numbers)와 관련한 주제에 대한 기존 혹은 평소의 인식을 가설로 두고 분석하여 검증 : 읍면동별, 성별, 연령별, 보험료기준별, 직역구분별 기준에 따라 건강지표 및 만성질환 지표의 발생률 차이가 있을 것이라는 가설하에 관련 데이터 분석 • (3단계) 분석을 통해 보여주고자 하는 주요 메시지(Key Message)를 제시

	<p>: 건강지표 및 만성질환 지표 가운데 읍면동별, 성별, 연령별 가장 높은 발생률을 나타내는 지역과 대상을 확인하고 우선관리 대상 선정 방법을 통해 대상별 보건의료행정 서비스 프로그램 대상을 선정함</p> <ul style="list-style-type: none"> • (4단계) 특정 보건의료지표를 예로 제시하고 지표를 분석하여 우선관리 대상 선정 방법 하는 등 분석방법을 일반화: 특정 보건의료지표에 대해 기준년도 인구에 우선관리 대상 선정등에 사용된 방법을 전체 읍면동에 대해서 일반화하여 적용 <p>◆ 분석값 정의</p> <ul style="list-style-type: none"> • 데이터 분석에 필요한 항목과 수치를 정함 <p>측정 항목</p> <ul style="list-style-type: none"> • 건강지표 발생률(흡연율, 비만율, 대사증후군 발건율, 암수검률, 권장신체활동 수행률, • 만성질환 지표 발생률(고혈압 의료이용률, 당뇨병 의료 이용률) <p>분석 범주</p> <ul style="list-style-type: none"> • 읍면동, 성별, 연령대, 보험료기준, 직역기준 <p>우선관리 대상 선정을 위한 읍면동별 인구수</p>
Tip	<ul style="list-style-type: none"> ◆ 분석설계는 분석을 위한 과정을 계획하는 것으로 "청사진"과 같음 ◆ 명확하고 구체적인 분석설계를 통해 분석 원래의 목적대로 진행해야 함

3. 데이터 수집 및 분석

- 보건의료지표 분석에서는 별도의 표본조사가 필요 없이, 원 자료를 수집하여 분석함
- 분석에 필요한 원 자료(RAW DATA)를 보유한 기관 등을 확인하여 적절한 방법으로 수집
 - 보건의료지표(건강지표 및 만성질환 지표)를 분석하기 위해서 국민건강보험의 보건의료지표 데이터를 수집

과제	<ul style="list-style-type: none"> ◆ 데이터 보유 주체(기관) 파악 및 수집 ◆ 수집된 다양한 데이터를 단일한 형태로 가공
추진내용	<ul style="list-style-type: none"> ◆ 데이터 보유 주체(기관) 확인 및 수집 <ul style="list-style-type: none"> • 지역보건의료지표 : 국민건강보험 • 읍면동별 인구수 : 남양주시 • 지역보건의료지표는 최근 5년, 읍면동별 인구수는 최근 3년치 데이터를 수집하여 분석 ◆ 수집한 정보의 가공 <ul style="list-style-type: none"> • 보유기관별로 제공된 다양한 형태의 데이터를 분석할 수 있도록 같은 단위의 척도(Scale)로 맞추어 구성하여 하나의 DB에 저장하고(오라클 DB), Tableau (BI SW)로 분석할 수 있도록 가공 • 국민건강보험에서 확보한 보건의료지표 데이터를 수집하여 분석에 필요한 형태로 가공하여 남양주시 빅데이터 분석 시스템에 저장함
Tip	<ul style="list-style-type: none"> ◆ 정부 혹은 공공기관이 보유한 정보를 요청할 경우, 사용 목적, 데이터 관리 방안, 내용, 형태 등을 구체적으로 작성하여 요청해야 함

4. 데이터 분석 · 검증

- 수집하여 가공한 정보(DATA)를 적절한 BI(Business Intelligence) 소프트웨어를 이용하여 분석함
- 수집된 정보(DATA)의 품질을 확인하고, 예상되는 분석 산출물에 대한 정의 필요
- 분석결과가 도출될 때마다, 현업담당자와의 의사소통을 통해 검증 필요

과제	<ul style="list-style-type: none"> ◆ 분석 산출물에 대한 정의 ◆ 수집 정보(DATA)에 대한 품질 검증 실시 ◆ 분석 설계 시 작성했던 분석 시나리오에 따라 도출하고자 한 주제별 분석 실시 ◆ 분석 결과 확인을 통한 분석방법 검증
추진내용	<ul style="list-style-type: none"> ◆ 분석 산출물에 대한 정의 <ul style="list-style-type: none"> • 산출물 정의 <p>보건의료지표별 특성을 분석하여 분석표를 산출 분석결과를 바탕으로 읍면동별 우선관리 대상을 선정</p> • 산출물의 시각화 여부 <p>그래프, 표 등을 사용하여 전달력이 강화되고 현업의 업무 활용도가 높은 대시보드 화면 구성 현업과 인터뷰를 통해 수시로 현업의 요구사항 수렴</p> ◆ 수집 정보(DATA)에 대한 품질 검증 실시 <ul style="list-style-type: none"> • 국민건강보험공단에서 수집된 데이터는 읍면동별, 성별, 연령, 보험료기준, 직장·지역구분으로 집계된 집계성 데이터이므로 별도의 품질 검증이 필요하지 않음 • 읍면동별 인구정보의 경우 정상범위에 포함되지 않는 오류정보(예 : 200세 생존 인구 정보 등) 제외, 다른 출처의 동일정보와 비교하여 품질 검증(예: 통계청 총인구수와 비교 등) ◆ 분석시나리오에 따른 주제별 분석 실시

	<ul style="list-style-type: none"> • 보건의료 지표 분석에 대한 시나리오 기반으로 분석 • 16개 읍면동별 인구특성 및 의료보험가입 특성에 따라 보건의료 지표 발생률을 분석하여 어느 지역에서 가장 발생률이 높고 어떤 대상에서 발생률이 높은지 파악 함 • 우선관리 대상 선정 방법에 따라 보건의료지표별 우선 관리대상을 선정하고 일반화 함 <p>◆ 분석 결과 확인을 통한 분석방법 검증</p> <ul style="list-style-type: none"> • 각 주제별 분석의 결과가 도출될 때마다 현업 담당자와 내용을 검토하여 확인 과정을 거치고, 다른 출처의 동일 자료와의 비교 등을 통해 분석방법 검증 • 분석 방법 및 산출물 검증을 통해 오류 발생 시 내용을 수정하여 재 분석하는 과정을 반복, 유의미한 결과가 도출되도록 함
Tip	<p>◆ 수집 정보(DATA) 품질 검증 시 해당 업무 담당자와 협의하면 오류 파악이 수월함</p> <p>◆ 주제별 데이터 분석 시 결과가 나올 때마다 해당 업무 담당자와 함께 결과를 확인 및 검증하는 것이 필요</p> <ul style="list-style-type: none"> • 유의할 것은 관련 업무의 문제점이나 기존에 발표된 데이터 오류를 지적하고자 하는 게 아님을 분명히 하고 빅데이터 분석의 목적을 명확히 제시해야 함

■ 데이터 분석·검증 단계

1단계 : 주제와 관련한 핵심 수치(key Numbers)를 확인

- 보건의료지표별 발생률을 먼저 확인 함
 - 읍면동별 보건의료지표별 발생률

흡연율, 비만율, 대사증후군 발견율, 암수검률, 권장신체활동수행율
고혈압 의료이용률, 고혈압 신규의료이용률, 1년 후 고혈압 투약 순응율
당뇨병 의료이용률, 당뇨병 신규의료이용률, 1년 후 당뇨병 투약 순응율
 - 읍면동별 성별 연령별 인구분포
- Key Message와 일관된 Insights를 도출
 - 읍면동별, 성별, 연령별, 보험료기준별, 직장·지역구분별 보건의료지표별 발생률 차이 있음

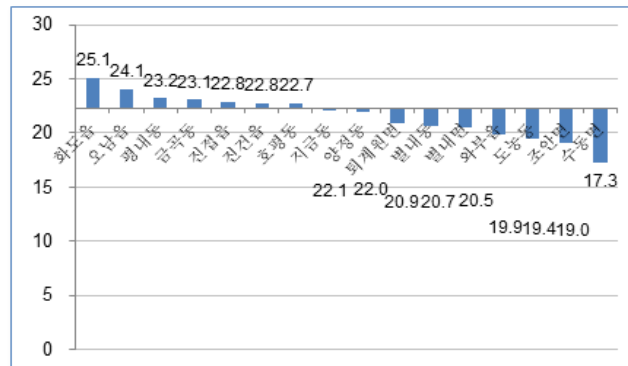
지표		발생률(2016년)
건강지표	흡연율	22.47%
	비만율	43.68%
	대사증후군발견율	23.21%
	암수검률	40.23%
	권장신체 활동 수행율	20.91%
만성질환지표	고혈압의료이용률	18.36%
	당뇨병의료이용률	0.94%

2단계 : 제시한 핵심 수치와 관련된 주제를 분석하고 검증

◆ 보건의료지표에 대한 분석 실시

흡연율에 대한 읍면동별 분포

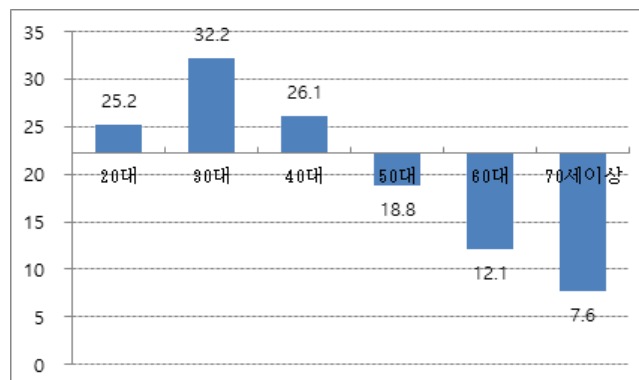
(단위 : %)



- 남양주시 전체 평균 흡연율 보다 높은 지역은 화도읍, 오남읍, 평내동이며 전체 평균보다 흡연율이 낮은 지역은 별내동, 별내면, 외부읍, 도농읍, 조안면, 수동면임

흡연율에 대한 연령별 분포

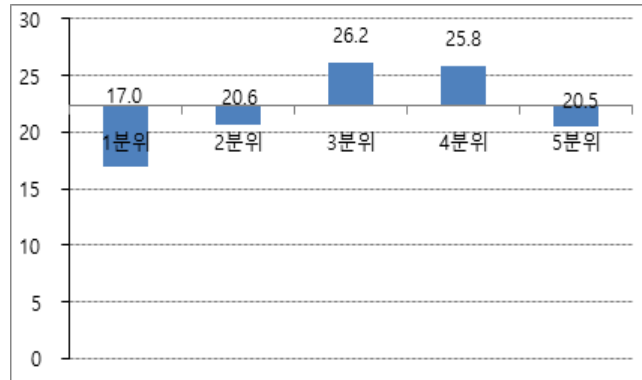
(단위 : %)



- 연령별로 살펴보면 30대의 흡연율이 가장 높고 50대부터는 흡연율이 낮음

흡연율에 대한 보험료 기준별 분포

(단위:%)



- 3분위, 4분위의 흡연율이 높음

◆ 흡연율에 대한 지역별 성별 연령별 교차 분석 실시

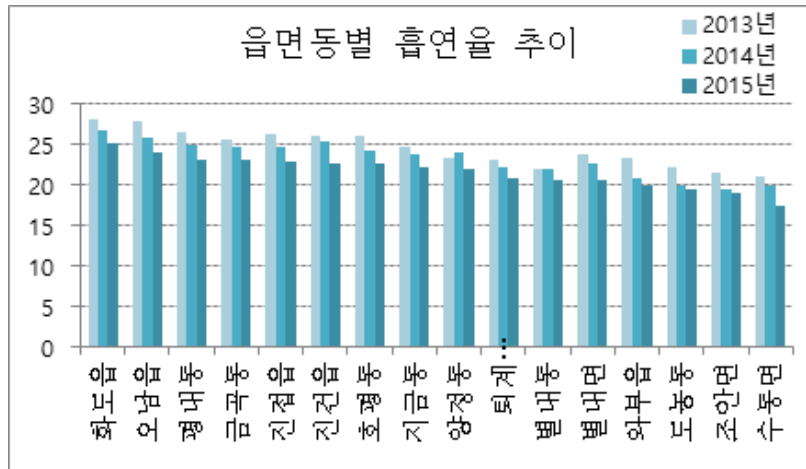
지역 (단위:%)	남성						여성					
	20대	30대	40대	50대	60대	70세~	20대	30대	40대	50대	60대	70세~
전체	46.6	50.6	45.5	33.7	21.8	14.2	6.9	6.0	4.8	3.6	2.4	1.7
화도읍	47.2	53.0	49.7	38.9	26.1	15.1	7.0	8.3	7.4	5.2	3.9	1.3
오남읍	52.7	55.3	49.9	36.5	19.7	17.5	5.5	6.2	5.4	5.1	1.6	1.6
평내동	52.1	49.2	44.1	33.9	22.0	15.5	9.6	5.9	5.1	4.0	2.1	2.7
금곡동	52.0	54.5	48.5	37.5	28.8	16.0	7.4	7.9	6.0	4.5	4.7	1.3
진접읍	48.0	51.5	46.1	34.0	19.5	16.1	7.2	5.6	5.1	3.8	2.1	2.7
진건읍	42.1	53.1	50.7	36.6	26.1	16.5	10.8	6.4	4.7	4.5	2.1	3.1
호평동	49.6	50.3	43.0	31.5	21.8	9.8	7.9	5.3	4.2	3.8	1.5	1.3
지금동	49.6	50.1	44.5	36.0	24.1	14.2	11.2	4.7	3.1	3.6	1.9	2.4
양정동			56.8	35.8	20.6					4.4	1.7	
퇴계원면	47.5	49.7	46.6	31.0	25.7	16.4	7.2	4.5	5.1	2.8	3.9	1.2
별내동	42.7	49.6	39.3	27.8	19.2	15.3	4.1	4.8	2.9	2.3	2.2	0.0
별내면	48.7	49.4	43.6	31.8	26.2	10.1	5.5	9.3	3.2	2.6	1.6	0.8
와부읍	38.3	46.3	41.8	31.6	18.2	11.3	5.4	4.7	3.3	2.4	1.5	1.0
도농동	47.0	47.1	42.4	30.8	17.5	12.3	3.7	4.8	3.4	1.5	1.5	0.9
조안면			50.9	25.2	19.8					1.9		
수동면	35.4	41.3	36.9	29.7	16.9	11.8			3.1	4.0	3.7	4.1

3단계 : 분석을 통해 제시하고자 하는 메시지를 제시

◆ 증감 추이 분석을 통한 메시지 파악

흡연율 최근 3년간 추이

(단위 : %)



- 대부분 3년간 흡연율이 감소하고 있으나 특정한 읍면동에서는 흡연율 감소율이 낮으므로 집중적인 관리 필요

◆ 읍면동 가운데 흡연율이 증가한 지역과 전년도 대비 흡연율 감소가 낮은 지역 파악

흡연율 증감

(단위 : 명,%)

지역	인구수 (2016년)	흡연율		감소율 (14년 vs 15년)
		2014년	2015년	
전체	662,154	23.89	22.25	1.64
화도읍	104,816	26.76	25.07	1.69
오남읍	55,206	25.91	24.07	1.84
평내동	41,788	24.84	23.20	1.64
금곡동	21,912	24.79	23.06	1.73
진접읍	96,312	24.77	22.82	1.95
진건읍	27,428	25.31	22.77	2.54
호평동	50,256	24.20	22.73	1.47
지금동	19,288	23.90	22.12	1.78
양정동	5,297	24.13	21.99	2.14

퇴계원면	32,192	22.14	20.89	1.25
별내동	64,526	22.03	20.69	1.34
별내면	21,165	22.74	20.50	2.24
와부읍	69,264	20.93	19.87	1.06
도농동	39,181	19.87	19.43	0.44
조안면	4,413	19.43	19.04	0.39
수동면	9,110	19.83	17.29	2.54

- 읍면동별 흡연율과 전년대비 흡연율 감소정도 그리고 인구수를 고려하여 흡연율에 대한 관리가 우선적으로 실시되어야 할 지역 선정

- ◆ 선정된 우선관리 대상 지역에 대해 성별, 연령별 교차분석에 의한 중점관리 대상의 규모 파악

우선관리대상 선정

(단위 : 명)

지역	남성						여성					
	20대	30대	40대	50대	60대	70세~	20대	30대	40대	50대	60대	70세~
전체	40,002	47,414	62,750	52,392	29,743	21,212	36,914	49,921	61,186	49,490	31,966	29,694
Target 計	3,966	3,920	4,015	4,007	348	552	438	700	687	343	79	191
%	9.9%	8.3%	6.4%	7.6%	1.2%	2.6%	1.2%	1.4%	1.1%	0.7%	0.2%	0.6%
화도읍	2,703	3,985	4,993	3,097	1,108	407	337	700	687	343	196	75
오남읍	1,559	1,874	2,883	1,496	358	279	105	198	268	208	38	36
평내동	1,133	1,361	1,874	922	271	148	224	198	237	112	24	50
금곡동	911	836	486	815	348	136	113	126	85	101	79	27
진접읍	2,472	3,392	4,324	2,341	672	450	357	402	428	246	49	115
진건읍	748	832	1,132	1,017	371	160	215	85	94	116	29	57
호평동	1,155	1,859	2,276	881	306	91	179	227	199	99	18	19
지금동	555	684	841	501	186	65	136	63	38	51	11	18
양정동			235	233	72				11	17	3	
퇴계원면	1,125	1,071	1,164	840	427	161	163	106	143	83	71	17
별내동	1,684	3,071	1,969	1,217	500	310	119	355	188	99	66	
별내면	622	593	779	509	289	82	72	126	52	43	22	8
와부읍	1,461	1,824	2,703	1,611	476	273	219	223	202	104	47	28
도농동	1,274	1,210	1,312	911	301	139	68	122	114	35	20	4
조안면			266	104	111					9		
수동면	189	161	243	248	108	109			16	34	25	40

**4단계 : 특정 보건의료지표를 예로 제시하고 지표를 분석하여
우선관리대상 선정 방법 하는 등 분석방법을 일반화(표준화)**

- 특정 보건의료지표에 대해 기준년도 인구에 우선관리 대상 선정등에 사용된 방법을 전체 읍면동에 대해서 일반화하여 적용
- 선정된 지역의 성별, 연령별 교차 분석한 지표를 통해 각 읍면동별 중점관리 계층 선정

(단위 : %)

지역	남성						여성					
	20대	30대	40대	50대	60대	70세	20대	30대	40대	50대	60대	70세
전체	46.6	50.6	45.5	33.7	21.8	14.2	6.9	6.0	4.8	3.6	2.4	1.7
화도읍	47.2	53.0	49.7	38.9	26.1	15.1	7.0	8.3	7.4	5.2	3.9	1.3
오남읍	52.7	55.3	49.9	36.5	19.7	17.5	5.5	6.2	5.4	5.1	1.6	1.6
평내동	52.1	49.2	44.1	33.9	22.0	15.5	9.6	5.9	5.1	4.0	2.1	2.7
금곡동	52.0	54.5	48.5	37.5	28.8	16.0	7.4	7.9	6.0	4.5	4.7	1.3
진접읍	48.0	51.5	46.1	34.0	19.5	16.1	7.2	5.6	5.1	3.8	2.1	2.7
진건읍	42.1	53.1	50.7	36.6	26.1	16.5	10.8	6.4	4.7	4.5	2.1	3.1
호평동	49.6	50.3	43.0	31.5	21.8	9.8	7.9	5.3	4.2	3.8	1.5	1.3
지금동	49.6	50.1	44.5	36.0	24.1	14.2	11.2	4.7	3.1	3.6	1.9	2.4
양정동			56.8	35.8	20.6					4.4	1.7	
퇴계원면	47.5	49.7	46.6	31.0	25.7	16.4	7.2	4.5	5.1	2.8	3.9	1.2
별내동	42.7	49.6	39.3	27.8	19.2	15.3	4.1	4.8	2.9	2.3	2.2	0.0
별내면	48.7	49.4	43.6	31.8	26.2	10.1	5.5	9.3	3.2	2.6	1.6	0.8
와부읍	38.3	46.3	41.8	31.6	18.2	11.3	5.4	4.7	3.3	2.4	1.5	1.0
도농동	47.0	47.1	42.4	30.8	17.5	12.3	3.7	4.8	3.4	1.5	1.5	0.9
조안면			50.9	25.2	19.8					1.9		
수동면	35.4	41.3	36.9	29.7	16.9	11.8			3.1	4.0	3.7	4.1

5. 데이터 가공·저장

- 이 단계에서의 데이터 가공 및 저장은 해당 주제와 관련된 빅데이터 분석이 지속적으로 이루어질 수 있도록 DB를 정비하는 것을 말함
- 검증된 분석결과를 바탕으로 최초 수집된 데이터에서 불필요한 데이터를 제외하고, 필요한 분석 항목 및 결과를 도출할 수 있도록 수정된 분석 모델에 따라 데이터를 가공 하여 저장하는 것을 의미함

과제	<ul style="list-style-type: none"> ◆ 빅데이터 분석 모델 확정 ◆ 수정된 분석 모델에 따른 데이터 가공 및 저장
추진내용	<ul style="list-style-type: none"> ◆ 빅데이터 분석 모델 확정 <ul style="list-style-type: none"> • 보건의료지표관련 데이터 확정 <ul style="list-style-type: none"> - 보건의료지표 발생률(성별, 연령별, 보험료기준별, 직역구분별) • 남양주시 인구분포 <ul style="list-style-type: none"> - 읍면동별 인구수(전체, 성별, 연령별) ◆ 빅데이터 분석 결과 불필요한 데이터 삭제 <ul style="list-style-type: none"> • 시스템 유지를 위하여나 단순 로그기록 등의 데이터 삭제 • 블록단위 가공 후 원천 데이터 모두 삭제 ◆ 수정된 분석 설계에 따라 데이터 가공 및 저장
Tip	<ul style="list-style-type: none"> ◆ 장비 도입에 대해서는 신중하게 고려 <ul style="list-style-type: none"> • 데이터를 저장하기 위한 인프라는 지속적인 데이터 발생 및 저장 후 분석을 추진하기 위해서 고려해야 함 • 일반적으로 일회성으로 정리되는 분석사업에 대해서는 별도의 인프라 구축까지는 고려하지 않아도 무방

바 행정적용 단계

1. 행정적용 단계 요약

단계	주요 내용
시범적용	<ul style="list-style-type: none"> ◆ 국민건강보험 데이터 분석을 통해 나온 결과를 통해 행정 활용, 보건서비스의 질 향상의 목적으로 실제 업무에 적용 하기 전에 사전 효과 검증 차원에서 진행 ◆ 분석 결과 적용 이전과 이후에 대한 업무의 효율성 등을 측정하기 위한 성과지표를 만들고 진행해야 함(읍면동별 지표별 구체적인 관리대상 정의로 캠페인 프로그램 효과 증진 등)
업무 적용	<ul style="list-style-type: none"> ◆ 시범 분석을 통해 효과성이 검증된 분석 기법을 실제 업무에 적용하는 단계
단계적 확대	<ul style="list-style-type: none"> ◆ 확대 방안(1) : 적용 업무의 확대 ◆ 확대 방안(2) : 적용 조직(부서)의 확대 ◆ 확대 방안(3) : 분석 대상과 시기의 확대
성과측정	<ul style="list-style-type: none"> ◆ 분석사업 실시 이전과 이후의 성과를 측정하기 위한 방안 마련 ◆ 시범적용 시 도출했던 지표를 보다 구체화하여 정기적으로 측정할 수 있도록 함
업무방식 변경	<ul style="list-style-type: none"> ◆ 분석결과를 대시보드형태로 적용하여 업무방식을 개선하고 일상화 되도록 건인

2. 남양주 시범적용 사례

- 국민건강보험 데이터 분석을 기반으로 흡연을 감소를 위한 다양한 캠페인 및 예방 교육 활동을 위해 읍면동별, 성별, 연령별 상세 관리 대상 정의

과제	<ul style="list-style-type: none"> ◆ 국민건강보험 데이터를 활용한 흡연율 상세 관리 대상 파악
추진내용	<ul style="list-style-type: none"> ◆ 활용 데이터(국민건강보험 지역의료지표) <ul style="list-style-type: none"> • 읍면동별 흡연율 • 성별, 연령별 흡연율 • 보험료 기준별, 직역구분별 흡연율 • 읍면동별 인구수(3년 치) ◆ 분석 방법 <ul style="list-style-type: none"> • 기초통계분석을 통해 읍면동별, 성별, 연령별, 보험료기준별, 직역구분별 흡연율을 파악 • 교차분석을 통해 각각의 조건에 따른 흡연율 파악 성별x연령별 흡연율, 성별x 보험료기준별 흡연율 읍면동별 x 성별 x 연령별 흡연율 • 분석범주별 기초통계분석 결과와 읍면동별 최근 3개년간의 추이 및 전년대비 증감정도를 파악하여 우선관리대상 선정 기준 정의 • 분석결과를 바탕으로 2017년 흡연율을 추정하고 2017년말 인구수에 적용 • 우선관리 대상 선정 기준을 적용하여 읍면동별 해당 의료지표의 우선 관리대상 규모 추정 성별, 연령별, 지역별 건강관리지표 및 만성질환지표별 우선 관리 대상규모 추정 ◆ 분석 결과 <ul style="list-style-type: none"> • 읍면동별로 보면 OO동, OO동에서 흡연율이 높았고 성별로는 남성, 연령별로는 30대가 높았음 • 흡연율 증감 추이를 보면 OO동은 지속적으로 흡연율이 증가하고 OO동은 정체이며 OO동은 감소를 나타냄 • 흡연율 우선관리가 요구되는 계층은 OO동 50대 남성, 30~50대 여성, OO동 40~70세 남성, OO동 60대남여임
Tip	<ul style="list-style-type: none"> • 현업 담당자들의 의견을 수렴하고, 필요성 등을 좀 더 확보하여 정책적으로 추진 필요

사 인프라 구축 단계

1. 인프라 구축 단계 요약

단계	주요 내용
시스템 요건 정의	<ul style="list-style-type: none"> ◆ 인프라 구축을 위한 정의 단계 ◆ 수집되는 데이터의 양, 분석 및 처리해야하는 데이터의 양, 저장해야 하는 데이터의 양 등을 고려하여 서버 및 스토리지 사양을 결정 ◆ 분석사업의 특징을 고려하여 도입해야 할 소프트웨어를 선정(메인 분석도구, BI솔루션, 실시간분석 도구, 리포팅 툴 등) ◆ 직접 서버를 구매할지, 클라우드 형태로 구현할지 장단점을 파악할 필요가 있음 ◆ 남양주시의 경우 기존에 구축된 빅데이터 분석시스템이 있어서 그 시스템을 사용함 ◆ BI솔루션을 이용하여 시각화를 위한 대시보드 화면을 구성함 대시보드 화면은 현업의 업무 활용도를 높일 수 있는 내용으로 구성함
DB 구조 정의	<ul style="list-style-type: none"> ◆ 수집되는 데이터에 대하여 테이블 수 등 실제 분석 시행을 위한 DB 구조를 정의
시스템 반영	<ul style="list-style-type: none"> ◆ 실제 인프라를 구성하여 분석 알고리즘과 데이터를 입력하여 적용하는 단계
검증/검수	<ul style="list-style-type: none"> ◆ 데이터 분석이 원하는 결과값이 도출 될 때까지 알고리즘 등을 보정하는 과정 ◆ 데이터가 제대로 연계되어 입력되는지 확인 ◆ 데이터 적합성을 확신할 경우 검수
오픈	<ul style="list-style-type: none"> ◆ 분석 시스템 오픈
업그레이드	<ul style="list-style-type: none"> ◆ 유지관리 사업 진행 후 필요 시 서버 용량, 분석 도구 구매 등

2. DB 구조 정의

- 보건의료지표의 분석을 위해서 국민건강보험공단의 의료이용지표 데이터를 수집

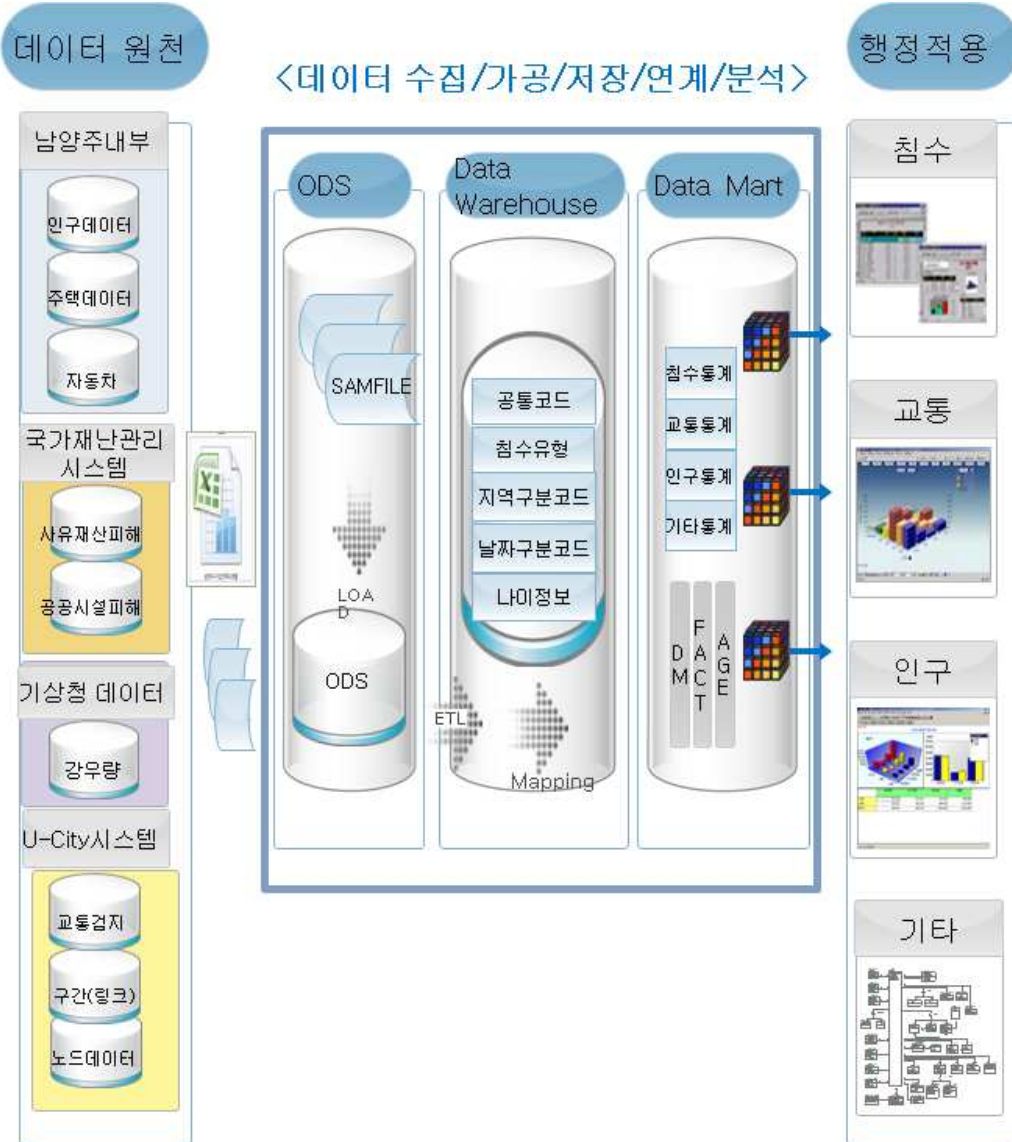
하여 분석에 활용하였으며 별도 서버 구매 등이 없었음

분석을 위한 데이터의 지속적인 저장을 위해 데이터의 속성을 정의하고 저장 기준, 주기 등을 정의하였음

- 빅데이터 분석을 완료한 후 관련 주제에 대한 지속적인 빅데이터 분석을 위하여 데이터를 효과적으로 이용할 수 있도록 DB(Database)의 데이터 속성을 정의함
- 빅데이터 분석 DB에 저장되는 데이터 항목과 속성 정의는 데이터 분석 단계에서와 같이 설계하고 정의하는 것이 효율적임

과제	<ul style="list-style-type: none"> ◆ 데이터 속성 정의 ◆ 지속적인 저장 주기와 데이터 마감 기준 정의 ◆ 측정 기준 정의
추진내용	<ul style="list-style-type: none"> ◆ 보건의료지표 분석을 위한 데이터 속성 정의 <ul style="list-style-type: none"> • 데이터 가공 및 저장단계에서 정해진 데이터 항목뿐만 아니라, 분석을 위한 기본 데이터 항목을 정의함 • 기본항목 : 연령, 성, 지역, 보험료구분, 직역구분 <ul style="list-style-type: none"> - 연령은 1세단위, 5세 단위로 저장 함 - 성별 : 남/여 - 지역 : 16개 읍면동 - 보험료 구분 : 의료급여, 1분위~5분위 - 직역구분 : 의료급여, 직장의료보험, 지역의료보험 ◆ 데이터 저장 기준 및 주기를 정의 <ul style="list-style-type: none"> • 저장기준 <ul style="list-style-type: none"> - 국민건강보험 데이터 : 보건의료 지표 업데이트 기준을 따름 (1년) - 인구현황 데이터 (1개월) ◆ 측정기준 정의 <ul style="list-style-type: none"> • 데이터를 효과적으로 처리하여 의미 있는 분석결과를 도출하기 위해, 데이터를 측정, 수집할 때의 기준을 정의 • 데이터의 기간, 종류, 크기 등을 미리 고려하여 DB 구조를 정의함으로써, 업무 파트에서 활용 시 효과적으로 데이터를 처리할 수 있도록 함
Tip	<ul style="list-style-type: none"> ◆ DB구조정의를 빅데이터 분석 과정에서 정리하는 것이 보다 수월함

<참고> DB구조 정의에 따른 남양주시 빅데이터 분석 시스템 구조

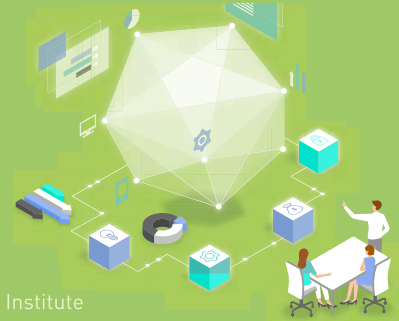


CHAPTER

IV

지방자치단체 빅데이터 분석 사례집

광주시 당직민원 빅데이터 분석 사례



korea Local Information Research & Development Institute

가. 추진배경	144
나. 프로젝트 개요	144
다. 적용 프로세스	147
라. 기획·관리 단계	148
마. 데이터 분석 단계	155
바. 행정적용	165
사. 인프라 구축 단계	172

IV. 광주시 당직민원 빅데이터 분석 사례

요약

경기도 광주시는 행정안전부 국가정보자원관리원과의 협업으로 야간과 휴일 등 취약시간에 발생하는 민원을 빅데이터로 분석하였다. 그 분석 결과를 토대로 당직운영 체계를 합리적으로 개선하였고, 그 결과 신속한 민원처리가 가능해져 민원인의 만족도가 크게 향상될 전망이다.

빅데이터 분석을 통해서 야간 또는 휴일 동안 접수된 민원에 대해 민원 담당 부서(실국, 읍면동과 단위) 별로 하루 동안 평균적으로 처리하는 민원의 유형과 건수를 쉽게 확인할 수 있도록 구현하였다. 또한, 기존에는 담당자가 수작업으로 내용을 확인하고 분류하던 것을 빅데이터를 통해 대중교통, 가로등, 소음 등 24개 유형으로 자동 분류하였다.

분석 결과 민원이 집중되는 시간과 유형을 알 수 있었다. 평일 야간의 경우에는 18시부터 21시까지 민원이 집중되었고, 가로등 고장이나 설치 및 버스 증차, 운행시간 조정 등 교통민원이 가장 많았다. 또한, 휴일에는 오전 7시부터 9시 사이에 집중되었고 공사장의 소음이나 먼지, 통행불편 그리고 차량의 소음 등 대중교통에 대한 민원이 많았다. 한편, 요일에 따른 차이는 거의 없었지만 토요일에 집중되는 것으로 나타났다.

경기도 광주시는 이러한 분석결과를 바탕으로 기존 민원처리에 소요되는 장시간의 대기비용을 대폭 줄이고 민원인별 맞춤형 서비스를 제공하게 되어 민원 만족도가 크게 향상될 전망이다. 향후에도 주기적으로 빅데이터 분석을 실시하여 당직근무 체계를 개선하는 등 민원인의 편의를 높여나갈 계획이다.

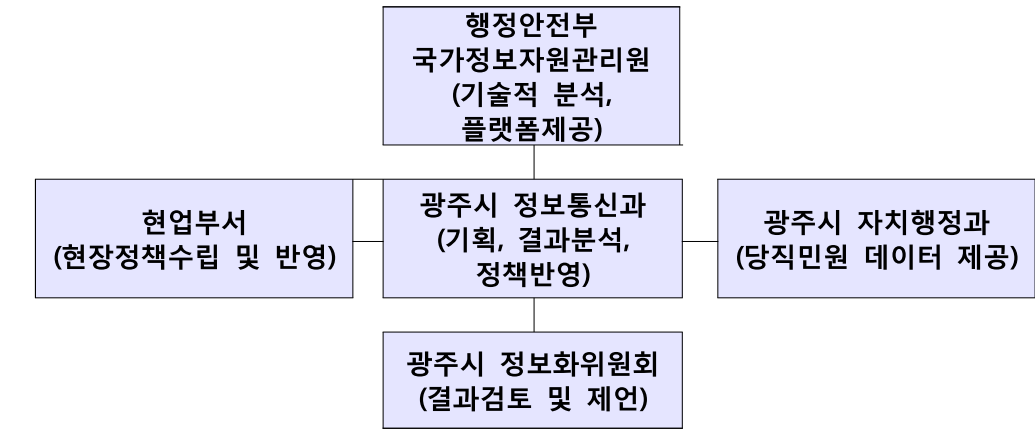
가 추진배경

- 사회 환경변화 및 시민수요 다변화에 따라 행정의 복잡화·다양화 시대에 직면
- 광주시 현안과 갈등 이슈의 합리적 해결방안의 일환으로, 행정안전부 국가정보자원관리원과 협업하여 「광주시 당직민원」에 대한 빅데이터 분석 수행을 통해 과학적인 정책수립 및 선제적 대응체계를 지원함으로써 「시민 중심의 책임 행정 구현」 '일 잘 하는 유능한 정부'를 실현 하고자 함

나 프로젝트 개요

- 개요
 - 분석과제명 : 당직민원 분석을 통한 취약시간대 효과적인 민원처리 방안 연구
 - 행정안전부 국가정보자원관리원과 경기도 광주시가 협업하여 야간과 휴일 등 취약시간에 발생 하는 민원을 빅데이터로 분석
 - 분석 결과를 토대로 당직운영 체계를 합리적으로 개선함에 따라 신속한 민원처리가 가능해져 민원인의 만족도 향상
- 분석의 필요성
 - 개인의 다양한 기대와 수요에 유연하게 대응할 수 있는 행정서비스에 대한 요구 증대
 - 최근, 일과시간 이후 민원발생 접수 및 민원 해결 요구 증가
 - 민원해결의 한계로 시민과의 갈등 초래 및 당직 근무자의 업무부담 가중
 - 취약시간대 효과적인 민원해결 방안 강구 필요
- 추진체계
 - 광주시 정보통신과와 행정안전부 국가정보자원관리원과의 협업으로 기술적 분석 및 플랫폼은 행정안전부 국가정보자원관리원에서 담당하고 기획, 결과분석, 보고서 작성 및 정책반영은 광주시 정보통신과에서 담당하여 효과적인 분석사업의 진행 구현

빅데이터 분석 추진체계



■ 추진기간 : 2016. 7. ~ 2017. 6.

■ 소요예산 : 해당 없음(행정안전부 국가정보자원관리원과 협업 추진)

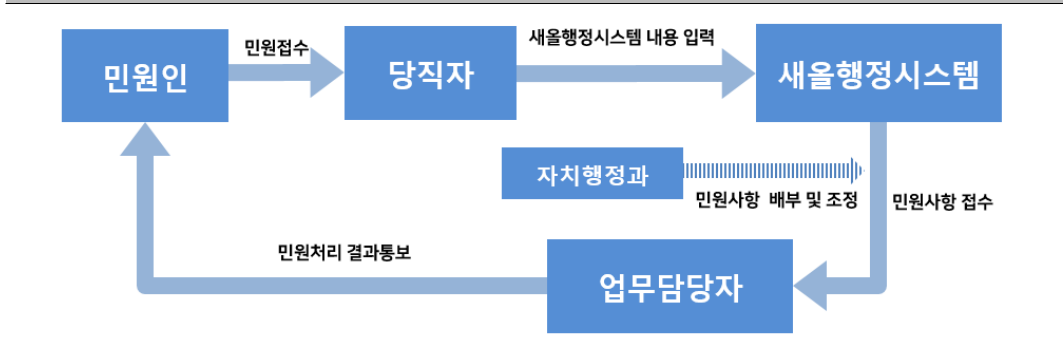
■ 분석데이터 : 당직민원(본청분)

● 데이터 추출 : 서울행정시스템 > 복무관리 > 당직관리 > 당직민원

< 당직민원 >

- 시간범위 : (숙직)18:00~익 일 09:00, (일직) 공휴일 09:00~18:00
- 내용범위 : 당직근무 중 발생한 **미해결 민원**으로 담당부서 이관 민원

당직민원 처리 흐름도

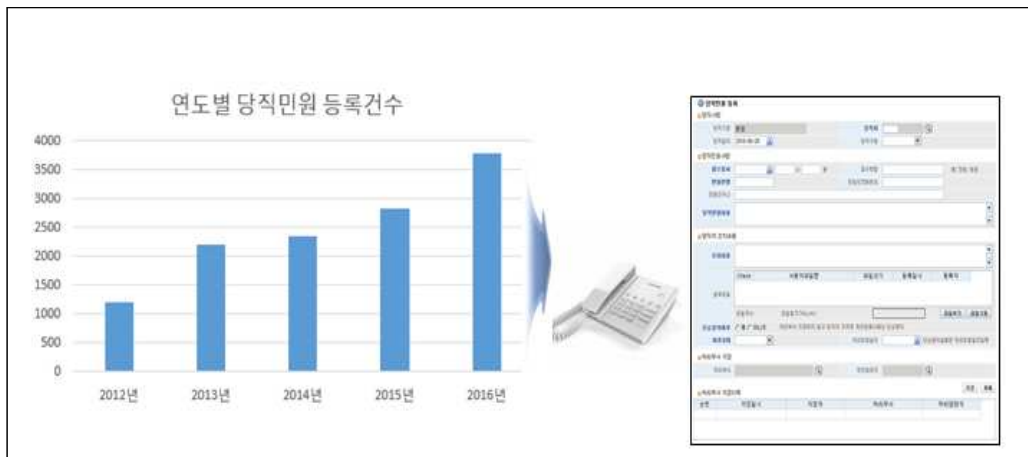


■ 주요내용

- 분석 데이터 양 : 2012년 ~ 2016년 (최근 5년간)

분석 데이터 양

연도	건수(건)	기간
2016	3,798	2016.1.1.~12.31.
2015	2,833	2015.1.1.~12.31.
2014	2,354	2014.1.1.~12.31.
2013	2,201	2013.1.1.~12.31.
2012	1,208	2012.3. ~ 12.31.
총합계	12,395	



■ 분석 데이터셋(예시)

당직 구분	접수일	접수 시간 (시)	접수 시간 (분)	접수 방법	민원인 주소	민원내용	처리완료 일자	처리 부서
숙직	20160628	22	30	전화	쌍령동	소음 때문에 시끄러워요	20160629	환 경 보호과

■ 분석내용

- <키워드 분석> 민원내용별로 추출된 키워드 빈도수 분석, 행정 및 정책과 관련성 있는

상위 키워드 추출 및 세부 연관 분석

- <시간별 분석> 시간별, 요일별, 월별, 연도별 민원 빈도 및 민원 키워드 추출
- <지역별 분석> 지역별(읍·면·동 단위) 민원 빈도수 및 Top키워드 추출, 민원발생 연관 분석
- <유형별 분석> 안전, 대중교통, 주차차, 도로, 대기, 소음, 쓰레기, 복지, 행정일반 등으로 분류 추출(담당부서별)

다 적용프로세스

■ 경기도 광주시 당직민원 분석 전체 프로세스

- 기획관리(문제의식 → 아이디어 제안 → 사업추진 → 최종보고 → 종료)
- 데이터분석(데이터수집 → 데이터가공 → 검증 → 저장)
- 행정적용(결과보고 → 관련부서회의 → 현장정책반영 → 성과보고)

광주시 적용 프로세스

기획/관리	문제의식	아이디어 제안	사업추진	최종보고	종료
데이터 분석	데이터수집	데이터가공	검증	저장	
행정적용	결과보고	관련부서회의	현장정책반영	성과보고	

라 기획 · 관리 단계

1. 기획 · 관리 단계 요약

단계	주요내용
문제의식	<ul style="list-style-type: none"> ◆ 개인의 다양한 기대와 수요에 유연하게 대응할 수 있는 행정서비스에 대한 요구 증대 <ul style="list-style-type: none"> • 최근, 일과시간 이후 민원발생 및 민원 해결 요구 증가 • 기존당직 운영체계로는 민원발생에 대한 효과적인 대응 부족 • 일과 시간 이후에도 즉시 민원처리를 원하는 24시간 서비스 체계에 대한 수요 존재 ◆ 민원의 다양성과 수량의 점진적 증가에도 현재의 당직 체계는 유연한 변화 없이 그대로 유지되어왔음 ◆ 당직자가 느끼는 취약시간 대 당직민원의 양과 강도의 체감 심각성이 점진적으로 증가하고 있으나, 구체적인 사례나 통계 등 근거가 미흡한 상황 ◆ 당직근무 중 민원인의 긴급 처리요청 사항의 처리지연 발생으로 시민이 느끼는 안정상 불안감 심화 ◆ 즉시 처리되지 못한 민원은 2차 민원을 발생시켜 시민의 불편을 가중시킴
사업추진	<ul style="list-style-type: none"> ◆ 2016.1.12. : 2016년 공공 빅데이터 분석사업 과제 제출 (광주시 → 행정안전부 국가정보자원관리원) <ul style="list-style-type: none"> • 제출과제명 : 광주시「당직민원 분석을 통한 취약시간대 효과적인 민원 처리 방안연구」 ◆ 빅데이터 분석 지원 사업 조사에 당직민원 사항에 대한 문제점과 분석의 필요성 어필 ◆ 2016.7.14. : 2016년 공공 빅데이터 분석 지원기관으로 선정 및 분석추진(행정안전부 국가정보자원관리원→광주시) ◆ 행정안전부 분석지원 사업으로 지정됨에 따라 행정안전부 국가정보자원관리원과의 협업을 통한 분석 추진

완료보고	<ul style="list-style-type: none"> ◆ 광주시 당직민원 빅데이터 분석결과 보고회 개최 <ul style="list-style-type: none"> • CEO 대면 보고를 통한 문제점 부각 및 정책의 실질적 반영 유도 ◆ 효과적인 민원해결을 위한 민원처리 담당자 회의 등 개최 <ul style="list-style-type: none"> • 당직 민원 빅데이터 분석결과 및 정책제언 등 설명 • 취약시간대 민원 대응 문제점 진단 및 해결방안 모색을 위한 토의
------	---

2. 문제의식

- 지속적인 인구증가 및 지역개발은 민원발생률과 비례하는 것으로 나타났으며, 사회 환경의 변화로 민원유형에도 큰 변화를 가져 오고 있는 것으로 나타남
 - 행정수요 급증에 따른 지속적인 민원 증가
 - 유기동물, 사체 등 새로운 민원의 대두
 - 24시간 즉시 민원처리 요구사항의 증대
 - 시의 급속한 팽창으로 대응방식에 대한 근본적인 고려가 필요
- 현장 당직근무 중 체감적으로 느끼는 당직민원의 횟수와 강도 등의 점진적 증대에 대한 환경과 내용 변화에 대한 의식
- 당직근무 중 현장 민원에 대한 근무자의 즉시 대응 부재 발생 시 이는 바로 민원인의 곤란과 불편으로 직결됨

■ 기존 당직근무 체계

● 일직/숙직

구분	총 근무 인원	근무대상	배정인원	근무시간	비고
일반당직 (숙직)	4명	5급	1명(재택)	평일/휴일 18:00 ~ 09:00	
		6급	1명(남)		
		7 ~ 9급	2명(남)		
일반당직 (일직)	5명	5급	1명	휴일 09:00 ~ 18:00	
		6급	1명(여)		
		7 ~ 9급	2명(여)		
		운전직	1명		
365 민원실운영	1명	7 ~ 9급	1명	휴일 09:00 ~ 18:00	주민등록, 인간 지적미원발급 처리
휴일 환경기동반	1명	6 ~ 9급	1명	휴일 09:00 ~ 18:00	소음 등 환경 민원 처리

※ 상황유지 근무자(자치행정과, 재난당직자) 제외

● 민원처리 상황반 운영현황

민원구분	처리부서	근무시간	근무인원	비고
광고물단속	주택 정책과	휴일(토) 09시 ~ 18시 휴일(일) 09시 ~ 16시	광고협회	
		봄,가을,겨울 07시 ~ 11시 여름 : 06시 ~ 10시	해병전우회	
주정차단속	교 통 정책과	토/일요일 10:00 ~ 14:00 평일 07:00 ~ 15:00 09:00 ~ 17:00 14:00 ~ 22:00	시간선택제	
노점상	도 로 관리과	주말, 공휴일 09:00 ~ 18:00	고엽제	

과제	<ul style="list-style-type: none"> ◆ 문제점 확인 ◆ 문제해결을 위한 필요한 활동을 나열
추진내용	<ul style="list-style-type: none"> ◆ 문제점 확인 <ul style="list-style-type: none"> • 당직근무 중 일어나는 민원사례에 대한 구체적 내용을 수집 확인 • 당직민원 근무 사항에 대한 체감적 어려움과 민원처리 현장근무의 어려움 • 당직근무의 어려움은 시민의 불편으로 직결 ◆ 당직근무 중 문제점 나열 <ul style="list-style-type: none"> • 취약시간대 즉시처리 민원 발생 시 대응 부재현상 발생 • 일과시간 이후 민원 발생 시 해당 민원의 대응 부재에 따른 2차 민원 발생 • 당직자의 즉시처리 불가능 민원(주정차단속, 도로 보수 등)에 대한 시민 불편 가중 • 일과시간 이후 집중 발생 민원처리를 위한 소요인력과 예산수립을 위한 근거 데이터필요 • 필요인력과 예산 확보를 통해서 취약시간대 발생민원 해결을 위한 인력 구성 및 대민서비스 필요
Tip	<ul style="list-style-type: none"> ◆ 현장에서 당직 근무를 통한 대민서비스로 일어나는 현장중심의 마인드로 문제점과 필요사항 파악

3. 사업추진

- 최초 빅데이터 분석 대한 경험이 전무한 상황에서 문제점을 알고 있으면서도 빅데이터 분석 사업에 대한 접근이 어려웠던 상황
 - 행정안전부 빅데이터 분석사업 수요조사를 통한 최초 분석사업 시도
- 분석사업의 과제선정에 따른 최초 시작의 어려움 극복
 - 행정안전부 빅데이터 분석사업의 하나로 선정됨으로써 빅데이터 분석추진 경험 부족에 대한 부담을 경감시키고 분석추진 가능
- 행정안전부 국가정보자원관리원의 빅데이터 전문가들과의 협업으로 빅데이터 분석

추진의 노하우 습득

- 행정안전부 국가정보자원관리원과의 사업관련 회의 추진으로 빅데이터 분석에 대한 개념파악 가능
- 사업 추진에 따른 전문가의 분석 가이드를 통해 빅데이터 분석에 대한 추진 동력 확보

과제	<ul style="list-style-type: none"> ◆ 빅데이터에 대한 최초 개념과 경험부족으로 시작의 어려움 해결 필요 ◆ 사업에 대한 동기 부여와 추진 동력 확보를 위한 의사결정라인의 관심 유도 방안 마련 필요 ◆ 협업 기관 및 부서 간 효율적 협조 도출 필요
추진내용	<ul style="list-style-type: none"> ◆ 2016.1.12. : 2016년 공공 빅데이터 분석사업 과제 제출 (광주시 → 행정안전부 국가정보자원관리원) <ul style="list-style-type: none"> • 제출과제명 : 광주시「당직민원 분석을 통한 취약시간대 효과적인 민원 처리 방안연구」 ◆ 2016.6.30. : 분석지원 가능 과제 선정 사전 통보에 따른 데이터의 관리 주체인 자치행정과와 분석사업을 위한 데이터 제공 가능여부 협의 ◆ 2016.7. 6. : 행정안전부 국가정보자원관리원 방문 협의 <ul style="list-style-type: none"> • 사업목적, 활용계획 등 설명, 데이터 제공 방법 및 진행방안 협의 ◆ 2016.7.14. : 「2016년도 공공 빅데이터 추가 분석지원 기관 선정」통보 (행정안전부→광주시) ◆ 2016. 7.15. : 당직민원 데이터 추출 및 민감 정보 제거 작업 및 자료제공 <ul style="list-style-type: none"> • 서울행정시스템을 통한 당직민원 데이터 추출 ◆ 2016.12.31. : 당직민원 데이터 기술적 분석 및 시각화 완료 <ul style="list-style-type: none"> • 공통기반 빅데이터 분석시스템인 '혜안'을 통한 분석 및 실시간 시각화를 통한 데이터 분석결과 내용 도출 ◆ 2017.5.2. : 당직민원 빅데이터 분석 결과의 정책 활용을 위한 담당자 회의 개최 <ul style="list-style-type: none"> • 결과로 도출된 데이터를 분석하여 현재 광주시의 당직민원 현황을 분석 • 분석에 따른 정책반영 및 효과적인 제언을 위해 상위 민원사항 담당직원들과 대책회의 및 정책반영을 위한 토론회 개최 ◆ 2017.6.7. : 당직민원 빅데이터 분석 결과보고 및 정책 활용 계획 수립 <ul style="list-style-type: none"> • 최종 결과보고서를 통한 다양한 정책제언

	<ul style="list-style-type: none"> ◆ 2017.6.19. : 당직민원 빅데이터 분석 결과보고회의 개최 <ul style="list-style-type: none"> • CEO 대면보고를 통하여 정책결정의 효과성 증대 • 각 부서 담당팀장, 담당자등과 함께 자유토론을 통하여 효율적인 정책 반영 도모
Tip	<ul style="list-style-type: none"> ◆ 각 부처 및 전문기관들의 분석사업 공모와 수요조사에 적극적 참여를 통한 빅데이터 추진 초기단계 어려움 극복 가능 ◆ 기술적 분석에 대한 행정안전부 국가정보자원관리원의 적극적 협조를 통해 빅데이터 사업의 접근성 해결 가능

4. 완료보고

■ 정책결정자와 관련 현업부서의 담당자들과의 분석 결과물 공유와 분석결과의 정책 활용을 위해서 완료보고회를 실시

- 최고결정권자에 대한 보고를 통해 전체적인 관심도 고취
- 보고회 전 사전 담당자 회의를 통한 실질적 보고회의가 되도록 노력

과제	<ul style="list-style-type: none"> ◆ 사업 완료에 따른 산출물의 지속적 효과성 제고 ◆ 결과보고서의 현업부서 활용을 위한 방법론적 고민 필요 ◆ 정책 반영을 위한 정책반영 계획의 별도 수립 필요
추진내용	<ul style="list-style-type: none"> ◆ 정책 활용을 위한 관련부서 회의 지속적 개최 추진 <ul style="list-style-type: none"> • 추진주제 : 당직민원 빅데이터 분석결과의 정책 반영 방안 • 추진기간 : 2017. 1. ~ 5. • 추진부서 : 정보통신과, 자치행정과외 관련부서 • 추진사항 <ul style="list-style-type: none"> - 2017. 3. 31. : 당직 민원 분석결과에 따른 1차 총괄부서(자치행정과) 업무 회의 개최 - 분석결과 공유 및 추진방안 모색 - 2016년도 전체 추가 민원분석, 주요 민원 부서 담당자 회의 개최 제안 - 2017. 4. 1. ~ 30. : 2016년도 전체 당직 민원 추가 분석추진(행정안전부)

	<ul style="list-style-type: none"> - 2017. 5. 2. : 주요 민원(상위 5위) 담당부서 담당자 회의 개최 - 분석결과 설명, 부서의견 및 개선방안 등 토의 - 주요 민원에 대한 세부적인 민원처리 실태 파악 제안 - 2017. 5. 2. ~ 5. 30. : 현장부서 상황 파악 - 관련 부서별 민원 발생 주요 요인, 민원처리 실태 및 문제점 등 의견수렴, 타 시·군 대응 사례 조사 실시 <p>◆ 광주시 당직민원 빅데이터 분석결과 보고회 개최</p> <ul style="list-style-type: none"> • 정책결정자의 보고회 참석을 통해 보고회 및 결과공유의 중요성 부각 • 당직 민원 빅데이터 분석결과 및 정책제언 등 설명 • 취약시간대 민원 대응 문제점 진단 및 해결방안 모색을 위한 토의
Tip	<p>◆ 분석결과 보고서 공유 및 해당 부서에 정책 반영 유도</p> <ul style="list-style-type: none"> • 관련 부서 대상 결과보고 회의 개최 추진(자치행정과, 관련 부서 등) <p>◆ 빅데이터 분석결과와 활용가치 제고를 위해 지속적 관리</p> <ul style="list-style-type: none"> • 반기별, 매년 주기적 민원 추이 분석 및 제공 • 추가적 상세 분석자료 업데이트를 통해 분석결과와 지속적 효용성 보장

5. 결과물 샘플

■ 행정안전부 빅데이터 공통기반 플랫폼(헤안)의 전문분석 사례에서 볼 수 있음

- 시각화 대시보드는 처리부서(비율), 처리부서(건수), 민원분류(비율), 민원분류(건수), 민원 상세조회와 분류별 시각화를 통하여 당직민원의 발생 추이와 현황을 파악하도록 하였음
- 처리부서(비율) : 처리부서별 민원처리 비율과 순위를 시각화
- 처리부서(건수) : 처리부서별 민원처리 건수와 순위를 시각화
- 민원분류(비율) : 민원유형별 비율과 순위를 시각화
- 민원분류(건수) : 민원유형별 건수와 순위를 시각화
- 민원상세조회 : 당직민원 현황의 종합적 분석내용의 시각화



단계	주요내용																		
분석 설계	<ul style="list-style-type: none"> ◆ 분석을 위한 기본적 당직민원 카테고리 정의 ◆ 당직민원 분석을 위한 필요데이터 정의 ◆ 당직민원 분석을 위한 분석방법 정의 <ul style="list-style-type: none"> • (키워드 분석, 시간별 분석, 지역별 분석, 유형별 분석) 																		
데이터 수집	<ul style="list-style-type: none"> ◆ 분석을 위한 데이터 정의 <table border="1"> <thead> <tr> <th>당직 구분</th><th>접수일</th><th>접수 시간 (시)</th><th>접수 시간 (분)</th><th>접수 방법</th><th>민원인 주소</th><th>민원내용</th><th>처리완료 일자</th><th>처리 부서</th></tr> </thead> <tbody> <tr> <td>숙직</td><td>20160628</td><td>22</td><td>30</td><td>전화</td><td>쌍령동</td><td>소음 때문에 시끄러워요</td><td>20160629</td><td>한 경 보호과</td></tr> </tbody> </table>	당직 구분	접수일	접수 시간 (시)	접수 시간 (분)	접수 방법	민원인 주소	민원내용	처리완료 일자	처리 부서	숙직	20160628	22	30	전화	쌍령동	소음 때문에 시끄러워요	20160629	한 경 보호과
당직 구분	접수일	접수 시간 (시)	접수 시간 (분)	접수 방법	민원인 주소	민원내용	처리완료 일자	처리 부서											
숙직	20160628	22	30	전화	쌍령동	소음 때문에 시끄러워요	20160629	한 경 보호과											

	<ul style="list-style-type: none"> ◆ 데이터 추출의 범위 지정 : 2012. 3. ~ 2016. 12. 31.(최근 5년간) ◆ 데이터 추출을 위한 사전협의 : 당직민원 업무의 주관 부서인 자치행정과 사전협의 ◆ 데이터 추출 <ul style="list-style-type: none"> • 서울행정시스템의 데이터 사용을 위한 사용승인 • 데이터 추출을 위한 스크립트 요청 • 추출스크립트 실행 요청 • 추출된 데이터의 정제(개인정보 제거 등) • 전처리과정 수행(부서매핑, 주소매핑, 불용어 정의 등) ◆ 데이터의 카테고리 정의 <ul style="list-style-type: none"> • 7개의 대분류와 24개의 소분류를 기준으로 키워드를 정의
데이터 분석	<ul style="list-style-type: none"> ◆ 분석에 따른 시각화 데이터의 해석과 보고서 작성 <ul style="list-style-type: none"> • 연도별 추이, 유형별 분석, 키워드 분석 등 시각화 결과에 대한 분석 • 현재 당직민원 현황과 분석결과의 융합 분석을 통한 분석보고서의 현장성 확보 ◆ 분석에 따른 보고서로만 그치지 않도록 부서의견의 청취(관련부서 회의 개최) 및 현장의 의견을 반영한 분석 및 보고서 작성

■ 데이터 추출 절차

데이터 추출

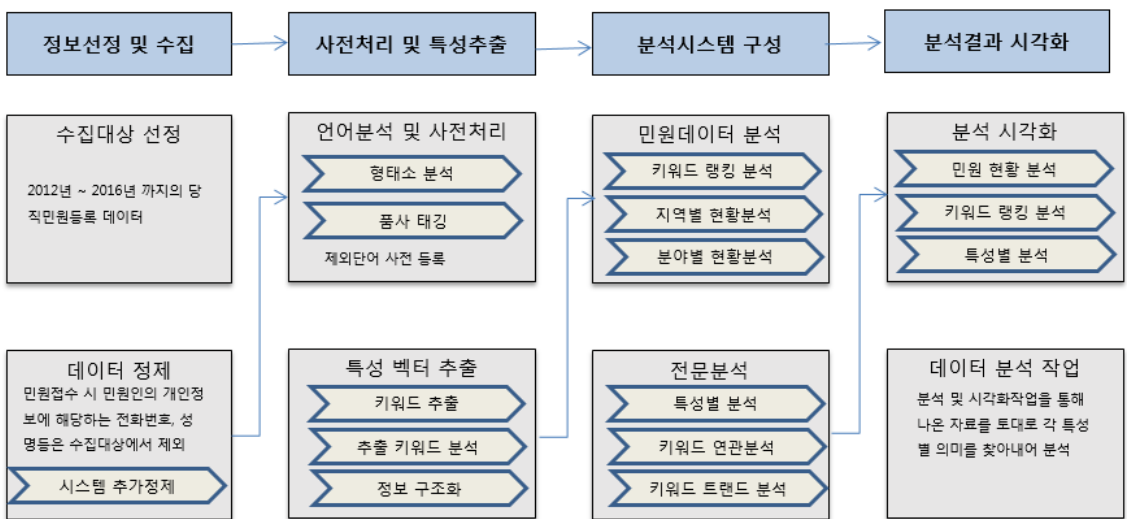


< 당직민원 >

- 시간범위 : (숙직)18:00~익 일 09:00, (일직) 공휴일 09:00~18:00
- 내용범위 : 당직근무 중 발생한 **미해결 민원**으로 담당부서 이관민원

- 서울행정시스템 내의 당직민원 접수 내역의 빅데이터 분석 사용을 위한 승인 협조(자치행정과)
- 서울행정시스템에서 데이터 추출을 위한 스크립트 요청(서울행정사업단)
- 당직민원 데이터 추출(정보통신과)
- 추출된 데이터의 민감 정보 제거 등 정제작업(분석사업 담당자)
- 정제된 데이터의 제공 및 분석 작업(행정안전부 국가정보자원관리원)

■ 분석 진행절차



- 서울행정시스템에서 분석대상 엑셀 파일을 추출
- 데이터 안에 포함된 개인정보를 삭제하고, 카테고리를 매핑하고 행정구역, 담당부서 등 추가 정보 추출 등 데이터 정제
- 형태소 분석을 위해 민원 내용에서 형태소를 추출하고 품사를 태깅하는 등 언어처리를 진행 (비속어 등 분석에 부적합한 단어 미리 사전에 등록하여 추출되지 않도록 처리)
- 분석결과를 시각화하여 보여줄 수 있는 데이터 분석 시스템 구성
- 기간별(연별/월별/시간대별 등), 지역별, 분야별 등 분석결과 추출
- 해당 결과에서 민원의 성격을 대표할 수 있는 의미를 찾아내어 분석 결과작성

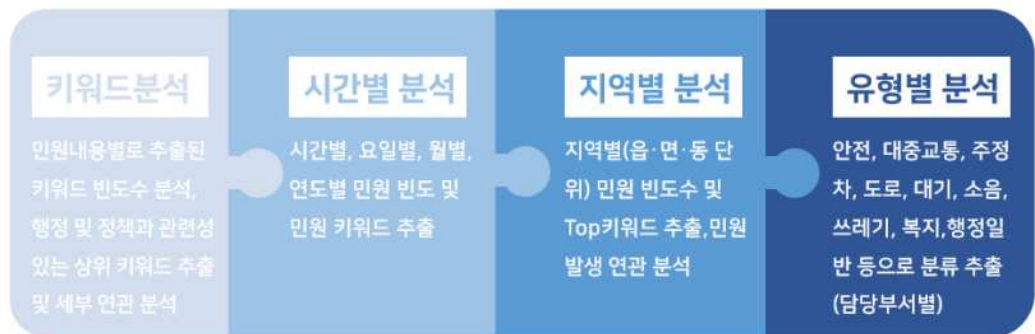
■ 카테고리 정의

- 원활한 데이터 분석을 위해 자체적으로 카테고리 체계를 만들었으며, 7개의 대분류와 24개의 소분류를 기준으로 키워드를 정의



- 분석 중 추출된 87개 부서명을 현재 부서명으로 현행화 매핑
 - 과거(건설도시국 건설과, 경제산업국 교통행정과, 문화공보담당관 등) 부서를 현재부서(안전건설국 건설과, 경제산업국 교통정책과, 공보담당관 등)으로 현행화하여 매핑자료 제공

■ 분석내용



■ 시각화된 데이터

● 시각화된 데이터는 총 다섯 개의 대시보드로 구성됨

- 대시보드 1 - 처리보드(비율)

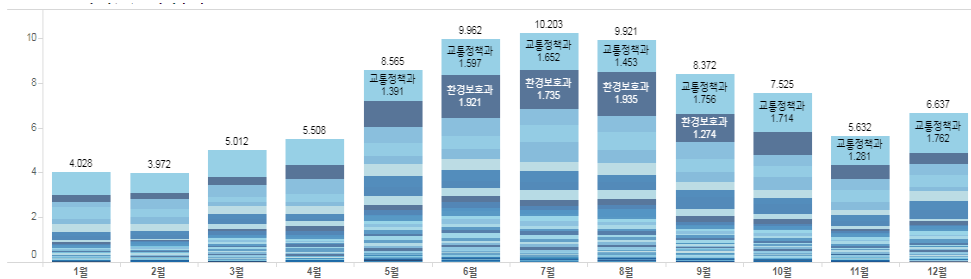
☞ 대시보드 1의 주제 : 민원분류, 처리부서, 처리건수, 처리일수와의 상관관계를 비율로 시각화하여 보여줌

✓ 민원 대분류(7종)과 민원 소분류(24종), 연도별로 동적으로 변화시키면서 변화상태를 볼 수 있음

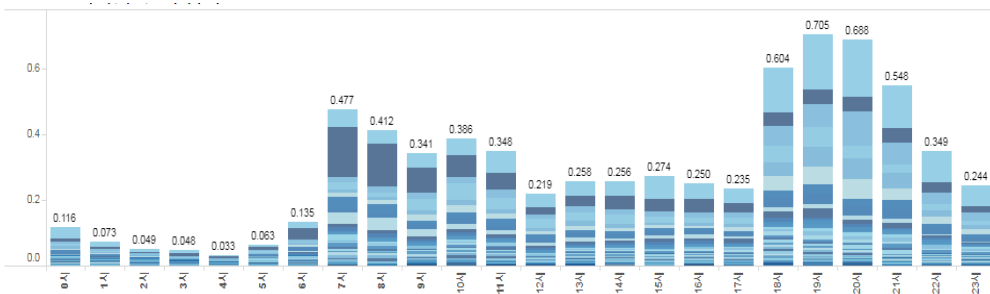
☞ 처리부서별 현황(처리국) : 처리국 별로 상위 10개를 보여줌

☞ 처리부서별 현황(처리과) : 처리과 별로 상위 20개를 보여줌

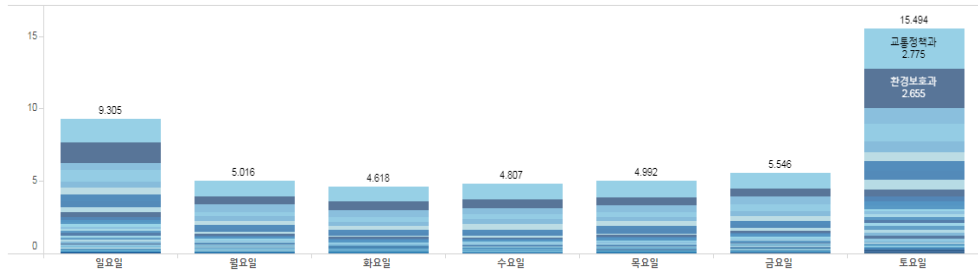
☞ 민원발생 추이를 월별로 분류하여 보여주고 있음



☞ 민원발생 추이를 시간별로 분류하여 보여주고 있음



☞ 민원발생 추이를 요일별로 분류하여 보여주고 있음



- 대시보드 2 - 처리부서(건수)

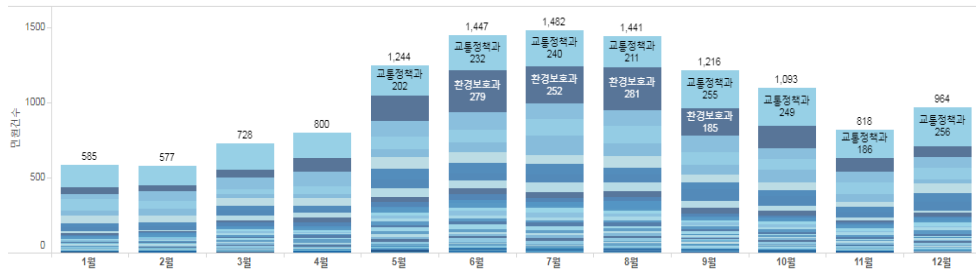
☞ 대시보드 2의 주제 : 민원분류, 처리부서, 처리건수, 처리일수와의 상관관계를 건수로 시각화하여 보여줌

- ✓ 민원 대분류(7종)과 민원 소분류(24종), 연도별로 동적으로 변화시키면서 변화 상태를 볼 수 있음

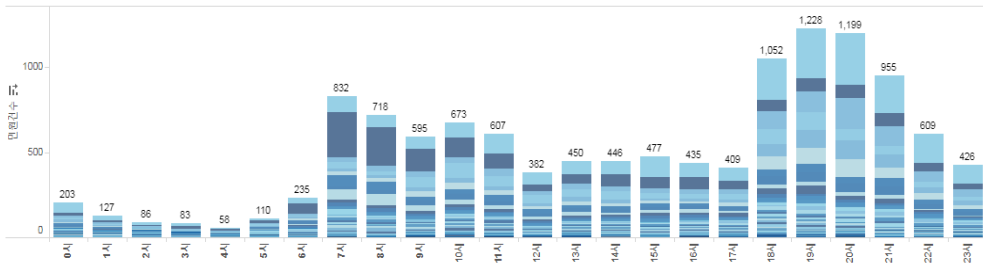
☞ 처리부서별 현황(처리국) : 처리국 별로 상위 10개를 보여줌

☞ 처리부서별 현황(처리과) : 처리과 별로 상위 20개를 보여줌

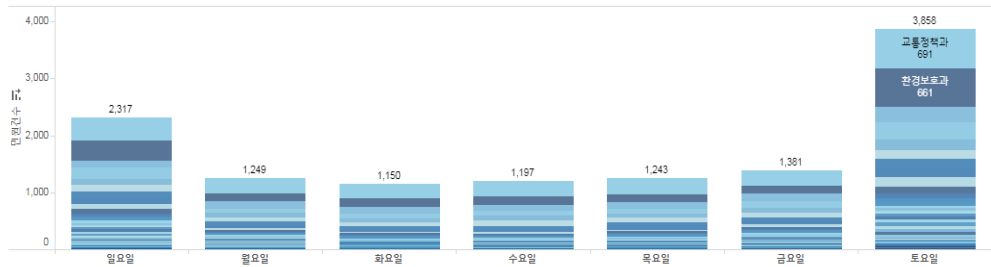
☞ 민원발생 추이를 월별로 분류하여 보여주고 있음



☞ 민원발생 추이를 시간별로 분류하여 보여주고 있음



☞ 민원발생 추이를 요일별로 분류하여 보여주고 있음



- 대시보드 3 - 민원분류(비율)

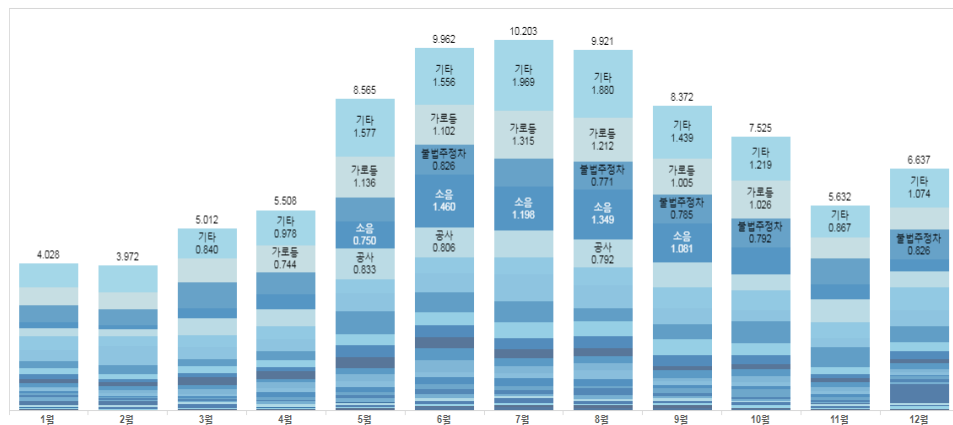
☞ 대시보드 3의 주제 : 민원분류, 처리부서, 처리건수, 처리일수와의 상관관계를 민원분류를 메인으로 하여 민원분류별 비율로 시각화하여 보여줌

- ✓ 상위부서(14종)과 하위분류(54종), 연도별로 동적으로 변화시키면서 변화상태를 볼 수 있음

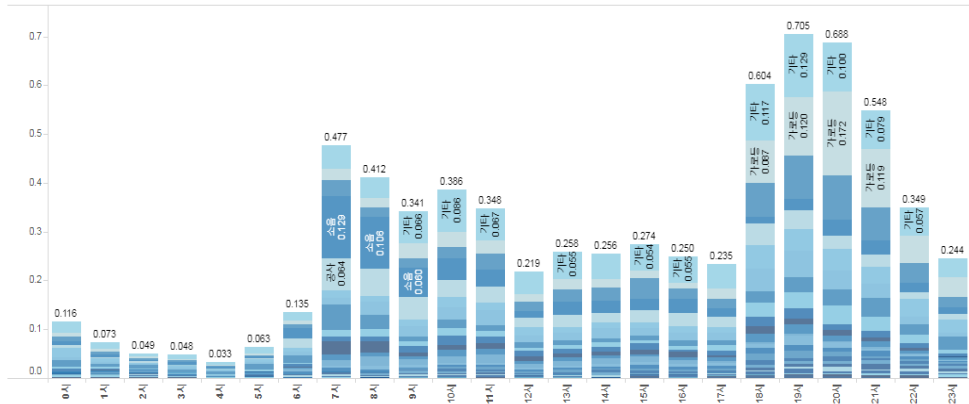
☞ 민원대분류 현황을 부서별로 비율을 시각화하여 보여줌

☞ 민원소분류 현황을 부서별로 비율을 시각화하여 보여줌

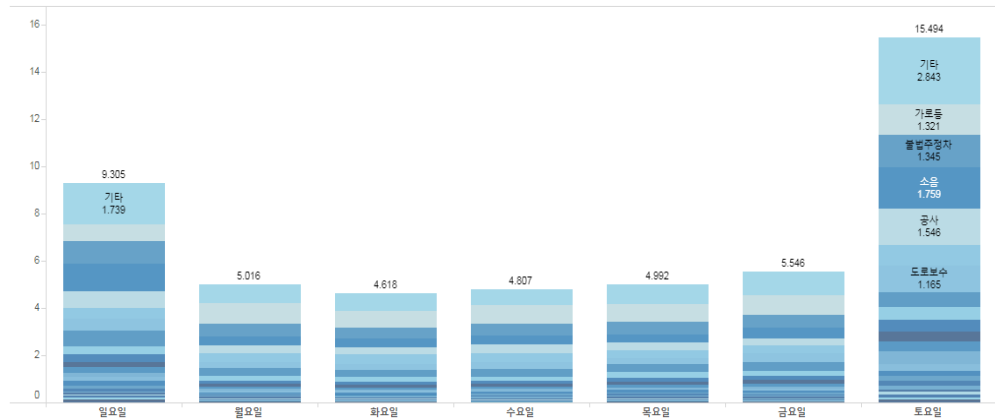
☞ 민원발생 추이를 월별로 민원분류에 따라 시각화하여 보여주고 있음



☞ 민원발생 추이를 시간별로 민원분류에 따라 시각화하여 보여주고 있음



☞ 민원발생 추이를 요일별로 민원분류에 따라 시각화하여 보여주고 있음



- 대시보드 4 - 민원분류(건수)

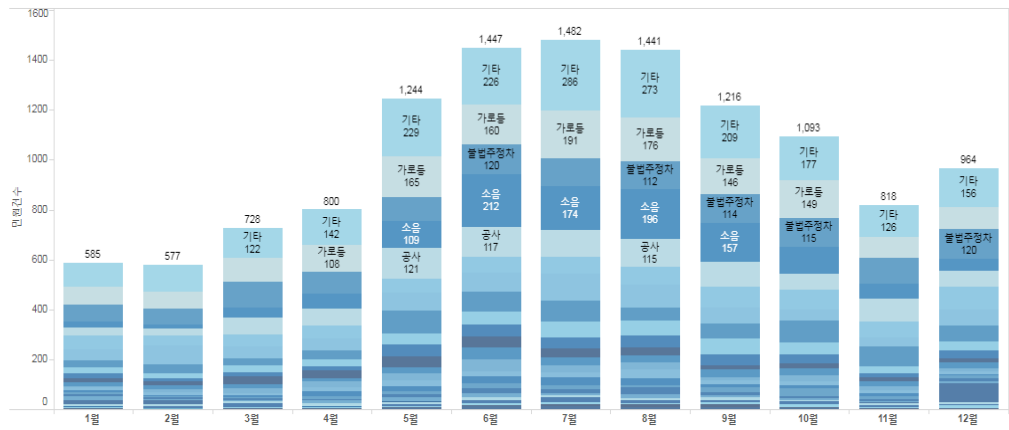
☞ 대시보드 4의 주제 : 민원분류, 처리부서, 처리건수, 처리일수와의 상관관계를 민원분류를 메인으로 하여 민원분류별 건수로 시각화하여 보여줌

✓ 상위부서(14종)와 하위분류(54종), 연도별로 동적으로 변화시키면서 변화 상태를 볼 수 있음

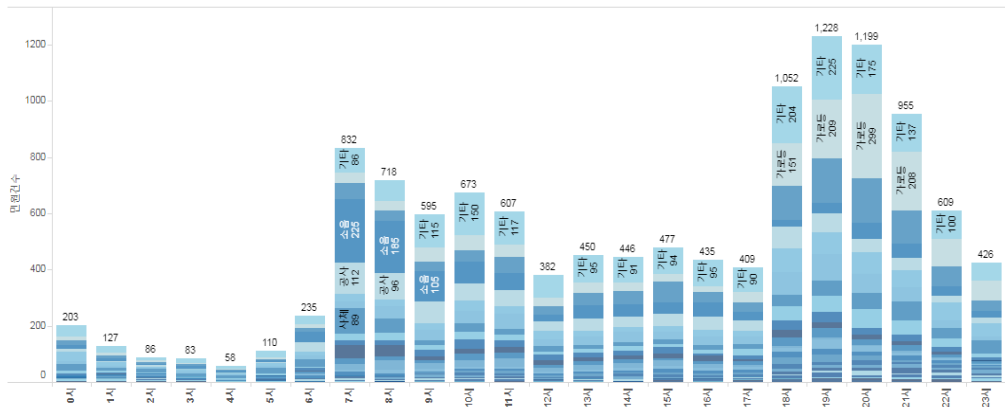
☞ 민원대분류 현황을 부서별로 건수를 시각화하여 보여줌

☞ 민원소분류 현황을 부서별로 건수를 시각화하여 보여줌

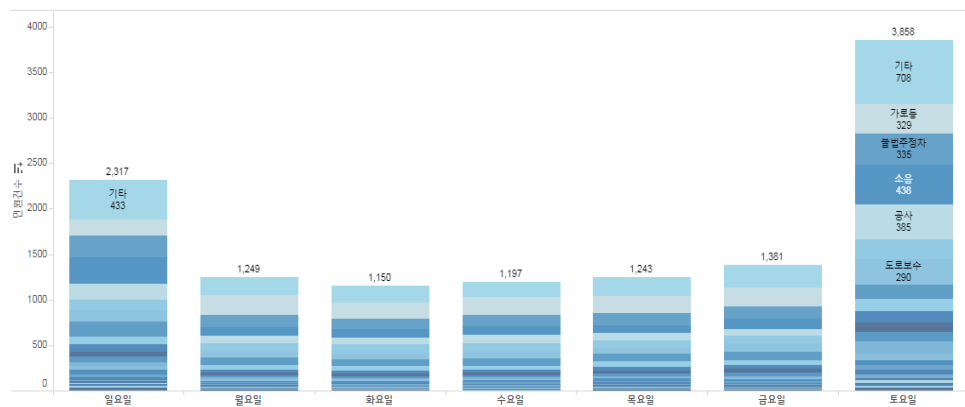
☞ 민원발생 추이를 월별로 민원분류에 따라 시각화하여 보여주고 있음



민원발생 추이를 시간별로 민원분류에 따라 시각화하여 보여주고 있음

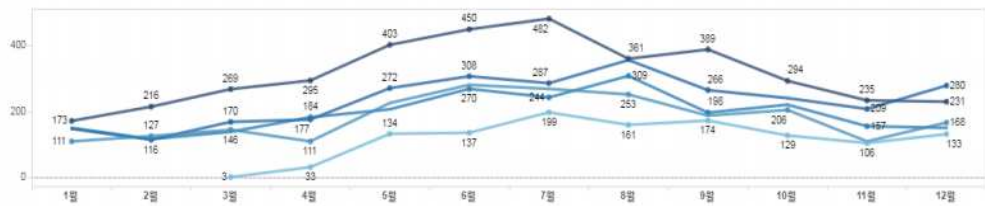


민원발생 추이를 요일별로 민원분류에 따라 시각화하여 보여주고 있음

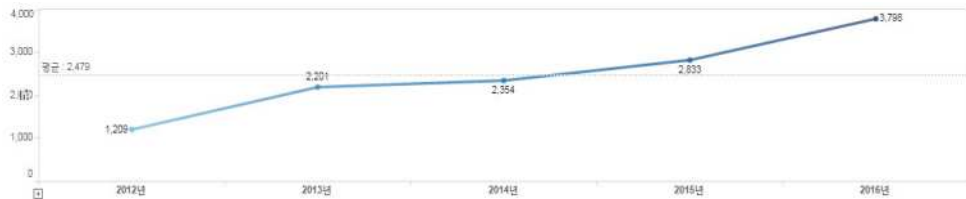


- 대시보드 5 - 민원상세조회

☞ 민원건수의 연도별, 월별로 추세를 분석하여 보여줌



☞ 연도별 민원의 총합의 변화를 볼 수 있음



☞ 부서별 처리유형과 상위처리 부서를 한눈에 볼 수 있음

☞ 대분류와 소분류로 나누어 어떤 분류의 민원이 상위에 있는지 한눈에 볼 수 있음

- 최초 비율에 따른 분류를 대시보드로 구성하였으나 추후 사업진행 과정에서 건수에 따른 대시보드 분류의 필요성이 대두됨에 따라 추후 구성됨
- 가장 많이 사용되는 대시보드는 민원상세조회로 가장 큰 그림을 직관적으로 파악이 가능하여 연도별 추이와 유형별 추이 등을 년, 월, 시간, 요일, 계절, 지역, 민원처리 분류에 따라 다양한 옵션으로 추이를 파악하고 분석할 수 있음

바

행정적용

1. 행정적용 단계 요약

단계	주요내용
업무적용	<ul style="list-style-type: none"> ◆ 「공공 빅데이터 분석 사업(당직민원분석)」관련 관련부서 회의 ◆ 광주시 당직민원 빅데이터 분석결과 보고회 및 효과적인 민원해결을 위한 회의 개최 <ul style="list-style-type: none"> • 실질적 회의를 위한 정책결정자와 업무담당자 간의 격의없는 토론 • 정책반영 가능 제안사항 등 결과 도출 ◆ 업무반영을 위한 관련부서의 활동 <ul style="list-style-type: none"> • 가로등 : 가로등 민원처리 단축 계획 수립 반영 • 불법주정차 <ul style="list-style-type: none"> - 기초질서(주정차 분야) 확립 연계 종합대책 추진 - 매주 목요일 특별단속 활동 중 • 당직개선 <ul style="list-style-type: none"> - 당직 통화 중 자동녹취 기능 시스템 개선 - 당직자 안전을 위한 야간 경광등, 조끼 등 안전장비 확보 - 당직자 업무 매뉴얼 개선
성과측정	<ul style="list-style-type: none"> ◆ 주기적 당직민원 분석을 통한 정책반영의 효과 측정 <ul style="list-style-type: none"> • 반영된 정책들의 추진사항 점검을 통한 지속적인 효과를 유도 • 매년 주기적 분석보고서 제출 및 성과보고회 개최
업무방식 변경	<ul style="list-style-type: none"> ◆ 분석결과를 토대로 현업부서의 업무방식을 개선 ◆ 분석결과가 효과적으로 정책에 반영되고 시민의 생활환경 개선에 실제적으로 반영될 수 있도록 추진 ◆ 1회성 분석과 이벤트가 아닌 실질적이고 지속적인 업무로 정착되기 위한 주기적 보고서 작성 ◆ 당직민원 분석뿐만 아니라 기타 빅데이터 분석사항(주택인허가 분석, 인구분석, 새울민원 분석 등)들과의 복합, 연계 분석을 통한 거시적 관점에서의 문제점과 원인 분석

7. 정책개선

■ 정책개선 성과

● 가로등 민원 해소 대책

- 시민의 안전을 위해 신속 조치를 위한 보수 장비 등 확충
- 효율적인 관리를 위해 장기적으로 정보통신기술을 활용한 가로등 관리시스템의 단계적 도입 확대
 - ☞ 가로등 양방향시스템 구축을 위한 사업 추진(553백만원 확보)

● 불법주정차 민원 해소 대책

- 인력 및 장비 확충을 통한 신속한 단속 시스템 구축
 - ☞ 지도단속 인력 증원(2명) 및 장비 확충(단속차량 1대)
 - ☞ 휴일 지도단속 운영시간 확대 (10시~14시 => 09시~17시)
 - ☞ 민원발생률이 높은 지역 선정, 단속요원 상주 근무실시로 즉시 민원 해결(4개지역)
 - ☞ 주정차 금지구역 내 방범용 CCTV 활용 단속시스템 구축 (7개소 / 40백만원)
- 지속적인 민원발생 지역을 중심으로 주차시설 확충, 거주자 우선주차제 등 추진
 - ☞ 노상주차장 추가 확충 (주요 주택가, 상가지역 등 3개소)
- 주변 주차장, 불법주정차 단속지역 정보 제공 및 지속적인 계도 등을 통한 시민 의식 개선
- 야생동물 출몰(로드 킬) 지역 GIS데이터 구축
 - 야생동물 출몰(로드 킬) 빈도가 높은 지역 안내표지판 설치(14개소)
 - 공간정보 활용지도 구축 및 생활지리 웹 포털 연계 정보 제공
- 당직 운영 개선 검토
 - 당직자 현장 출동 시 안전장치 확보 필요
 - ☞ 도로 위 사체, 적치물 등 제거 시 안전사고 발생 우려
 - ☞ 안전장비(야광 조끼, 안전봉, 장화 등) 제공
- 새롭게 대두되고 있는 민원사항 및 조치방법의 신속 공유와 주기적인 매뉴얼 갱신으로 원활한 민원 안내 지원

- 사체발생, 멧돼지 출현(농작물 피해 방지 조치), 대규모 건축공사, 시민불편이 예상되는 각 부서, 읍·면·동 추진 중인 공사 등

☞ 신규 공무원 대상 당직근무요령 교육 실시로 실무 적응 지원

☞ 컬러링 시스템 개선으로 당직민원 자동 녹취 시스템 구축(10백만원)

☞ 당직 매뉴얼 개편 및 새울 시스템 내 코너 개설 운영

● 주기적인 당직민원 분석의 활성화

- 각종 민원에 대한 주기적 분석으로 “시민의 관심사”, “경향” 등을 수시로 파악, 정책 수립에 반영
- 민원분석 시 직원 애로사항, 문제점 등 병행 분석으로, 내·외부 시스템 동시 개선

● 빅데이터 업무 활성화 기반 마련

- 팀 명칭 변경 (종전) 정보기획팀 ⇒ (변경) 데이터기획팀
- 팀 인력 충원 (1명)

■ 정책개선 노력

● 분석사업의 효과 극대화 고민

- 빅데이터 분석 사업의 결과가 정책에 효과적으로 반영하기 위한 방법 고민
- 원타임 프로젝트, 전시행정, 예산낭비 등 부정적 이미지 탈피 고민
- 정보기술의 이해 및 과학적 자료로의 신뢰성 제고

● 실질적 결과보고서 산출을 위한 노력

- 효과적인 정책반영이 이루어지기 위해서 최종 결과보고서 보고 전 관련부서 회의 10회 이상 개최
- 실제 업무를 수행하는 실무 담당자의 면담을 통한 현장 내용 파악 및 공감대 형성
- 조직과 인력 등 시정을 관할하는 담당자와의 면담을 통한 현장부서 지원 협의

● 실무부서 담당자 회의 등의 도출된 결과의 취합

● 실질적인 결과보고서 산출을 위해 회의 결과 등을 반영하여 10회 이상 결과보고서 초안의 수정

● 수요자 중심의 분석 추진

- 현업부서와의 지속적인 소통과 공유

- ☞ 이해와 의견수렴 과정을 통한 부정적 시각 해소
- ☞ 분석목표의 현실적 접근을 통한 신뢰성 제고
- 이슈화를 통한 정책반영 실행 유도
 - 실무자와 최고 결정권자와 토론회 마련
 - ☞ 부시장과 실무담당자와 격의 없는 토론 진행
 - ☞ 민원처리 실태 진단 및 문제점 부각
- 대내외 결과 공표
 - 언론보도
 - “혜안” 활용 분석결과보고서 공개
 - ☞ 혜안(www.insight.go.kr) > 전문분석 > 분석사례 > [경기도 광주시] 당직민원 분석을 통한 취약 시간대 효과적인 민원처리 방안 연구
 - ☞ 메뉴에서 결과분석 및 시각화 내용 확인 가능
 - 새울시스템 전 직원 열람 및 다운로드 상시 가능
- 지속적인 모니터링
 - 정책 반영 모니터링 및 보고체계 구축으로 실효성 제고
 - 인사이동 시 담당자 교육 실시로 정책 연계 유도
- 관련부서 공감대 및 협의를 이끌어 내기 위한 정무적 노력
 - 관련부서 공감대 형성을 위한 실무자 면담
 - 관련부서 공감대 형성을 위한 팀장급 회의
 - 자유로운 분위기 속에서 격의 없는 토론 진행
- 결정권자에게 사실적 보고 및 자유로운 분위기의 보고회 진행
 - 사실적 보고를 통한 최고결정권자와의 공감대 형성
 - 자유로운 보고회 진행을 통한 진솔한 토론의 장 마련

■ 정책개선 절차

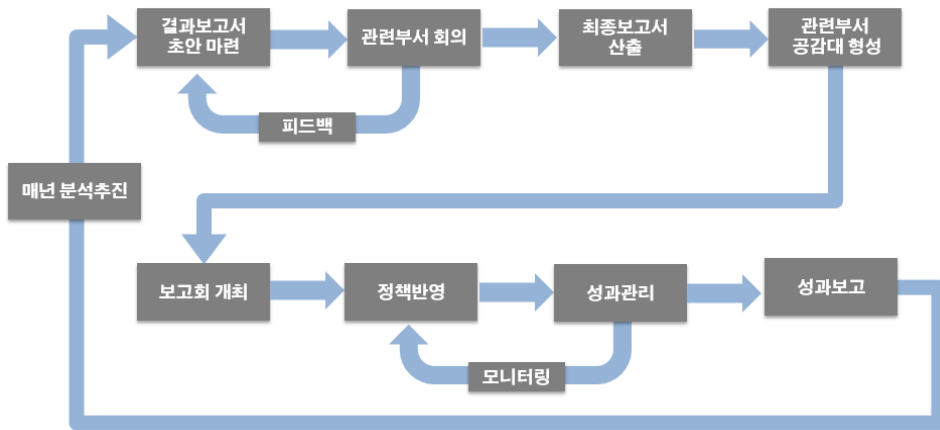
- 결과보고서 초안 마련 → 관련부서 회의 → 실질적인 최종보고서 산출 시까지 피드백 → 회의 반복 → 관련부서 공감대 형성 → 최고결정권자 보고회 진행 → 사실적 보고

및 자유로운 보고회 토론 → 보고회 결과의 효과성을 위해 모니터링 → 정책반영사항
모니터링 및 독려 지속 → 성과관리 → 성과보고 → 매년 분석 추진

■ 향후추진 사항

- 성과도출이 1회성에 그치지 않도록 매년 당직민원 분석
- 매년 분석사항을 과년도와 비교 분석 추진
- 정책반영 사항들의 효과성 검증
- 새로운 유형들의 당직민원에 대한 파악

정책반영 프로세스(요약)



<참고> 분석결과(요약)

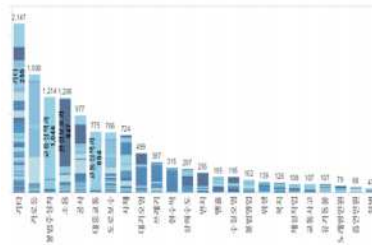
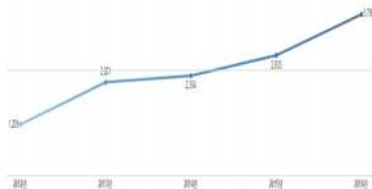
- ◆ 전체 민원발생건수 : 12,395건, 매년 지속적 증가
 - ※ 2016년 민원발생 최고(3,798건), 2013년(2,833건) 대비 72.6%(1,597건)
- ◆ 민원발생 상위(5위) : 가로등(1,500건), 불법주정차(1,214건), 소음(1,206건), 공사(977건), 대중교통(775건) 순
 - ※ 2013년 대비 가장 크게 증가한 민원 : 사체 385%(70건→306건), 불법주정차 124.7%(174건→391건), 공사 116.3%(153건→346건), 소음 101.9%(212건 → 428건)
- ◆ 시간대별
 - 하루 중 오후 7시~8시 사이 가로등, 불법주정차, 대중교통 민원 집중 발생, 오전 7시 소음 민원 최다
 - 일주일 중 토요일 민원 최고, 평일 가로등, 주말 소음민원 최고
 - 1년 중 6월, 계절별 여름철 소음 민원 발생 최고
- ◆ 민원분야별 : 환경 분야(소음, 공사, 사체 등) 가장 높고, 교통 분야(불법주정차, 대중교통, 도로보수 등), 안전 분야(가로등, 유기동물 등) 순
 - 환경 분야 민원 발생이 가장 높고, 교통정책과 민원처리량이 최다임. 평균 민원처리 일수는 4.087일
- ◆ 민원처리 부서별(상위 3위) : 교통정책과, 환경보호과, 도로관리과 순
- ◆ 민원처리 일수 : 전체 평균 처리일수 4.087일
 - 2015년(2,933건/4.553일), 2016년(3,798건/3.284일)로 전년 대비 민원은 29.5%(865건) 증가했으나, 평균 민원처리일수는 1.269일 단축
 - 상위 5위 민원 중 불법주정차, 소음, 공사, 대중교통 민원 처리일수는 전년 대비 단축
 - 가로등, 불법주정차, 소음, 공사 민원 발생은 지속적으로 증가한 반면, 대중교통 민원은 2015년 대비 감소한 것으로 나타남
 - 평균 처리일수 추이를 보면, 소음 민원의 경우 2015년 대비 민원발생은 크게 증가(161건)하였으나, 처리기간은 3.531일 단축된 것으로 나타남
 - 불법주정차, 공사 민원의 경우도 민원발생은 증가하였으나, 처리기간은 각각 1.608일과 0.993일 단축된 것으로 나타남
- ◆ 민원 발생지역
 - 읍면지역 : 가로등, 소음, 공사 민원이 가장 많음

- 동지역 : 가로등, 소음, 불법주정차 민원 최고
- ◆ 시사점
 - 상위 5종 민원에 대한 적극적인 해결 방안 모색 필요
 - 가로등, 불법주정차, 소음, 공사, 대중교통
 - 현업부서 중심의 근본적인 개선 대책 마련 시급
 - 도로관리과, 교통정책과, 환경보호과
 - 취약 시간대 민원발생 지속 증가
 - 생활여건 변화에 따른 새로운 행정수요 대응 방안 필요

01 연도별 발생추이

2012.3.~2016까지 총 12,395건
(일 평균 7.1건)의 민원 발생

매년 지속적인 증가 추세
2016년 (3,798건) 가장 최고
2013년 대비 72.6%(1,597건) 상승



가로등, 불법 주정차, 소음, 공사, 대중
교통 민원 순으로 높게 나타남

가로등, 소음, 불법 주정차 매년 높게 발생

2012~2015년까지 **가로등 민원** 최고
2016년 **소음민원** 가장 높게 나타남

2013년 대비 가장 크게 증가한 민원
동물사체 385%, 불법 주정차 124.7%,
공사 116.3%, 소음 101.9% 순

02 시간대별 발생추이

하루 중 오후 7시 부터 8시 사이
저녁 시간대 집중 발생

가로등, 불법주정차, 대중교통 민원 최다

오전 시간대 7시부터 8시, 소음, 공사, 동물사체
민원 높게 발생



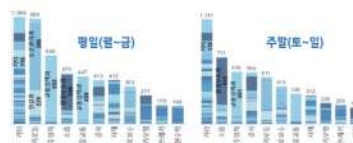
시간대별 주요 민원 발생 현황



일주일 중 토요일 민원 발생 최고
주말 민원 49.8% 차지

평일 가로등, 불법주정차, 소음 민원

주말 소음, 불법주정차, 공사 민원이 가장 높음



사 인프라 구축 단계

1. 인프라 구축 단계 요약

단계	주요내용
시스템 사용	◆ 빅데이터 공통기반 해안 활용

■ 공통기반 빅데이터 플랫폼(해안)을 활용

- 본 사례는 광주시와 행정안전부 국가정보자원관리원이 공동으로 분석 작업을 하였으며 분석에는 기 구축 활용 중인 빅데이터 공통기반 플랫폼인 ‘해안’시스템을 활용하였음

■ 공통기반 빅데이터 플랫폼(해안)의 전문분석 사례에 등록

- 해안(www.insight.go.kr) > 전문분석 > 분석사례 > [경기도 광주시] 당직민원 분석을 통한 취약 시간대 효과적인 민원처리 방안 연구
- 결과보고서 공개 설정으로 필요 직원들의 참고가능

해안 빅데이터공통기반 서비스소개 데이터검색 일반분석 분석모델 전문분석 지자체이슈 커뮤니티 비식별조치

분석사례 플랫폼신청

분석사례

과제명: 검색

요청기관: 기관검색

과제유형: [전체] 분석기법: [전체]

• 총 170 건

등록일	과제유형	과제명	요청기관	추천수/조회수
2017-08-25	분석과제	GIS를통한 일리리아 우선방역 후보지 빅데이터 분석	정보통신과	12/134
2017-08-22	분석과제	진매구 원도실 유동인구 분석	장원시	15/154
2017-07-27	분석과제	김포시 인구증가 소분분석	정보통신과	11/96
2017-03-06	분석과제	차량주행용 빅데이터 분석을 통한 이용정보 제공	행정지침과	105/1604
2016-10-18	분석과제	소셜 분석을 통한 "부정청탁 및 금품 등 수수의 금지에 관한 법률" 집행사항 파악	빅데이터분석과	20/1432
2016-10-10	분석과제	[경기도 광주시] 당직민원 분석을 통한 취약 시간대 효과적인 민원처리 방안 연구	광주시	78/1024

지방자치단체 빅데이터 분석 사례집

2018년 1월 인쇄

2018년 1월 발행

발행인 손연기

발행처 한국지역정보개발원

03923 서울특별시 마포구 성암로 301

전화: (02)2031-9100 팩스: (02)2031-9360

발간등록번호 B551982-2017-000008-01